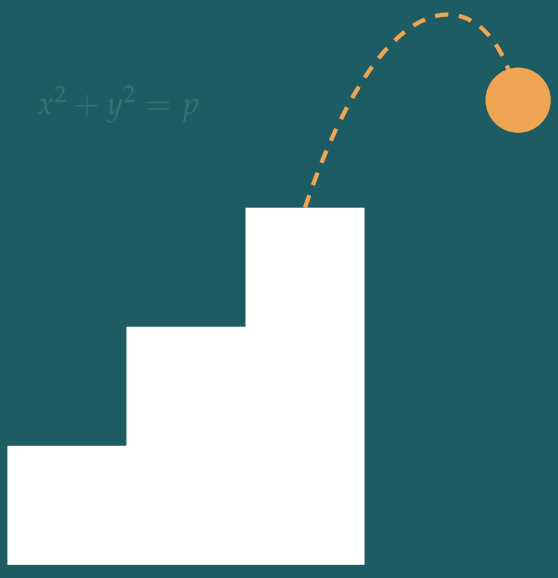
STEP 備試 · 大學數學

Beyond STEP

STEP I, II 及 III

1988 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

1988 年詳細解答

Keith McNulty

Beyond STEP：以歷屆 STEP 試題為門，走進大學數學

1988 年詳細解答

1988 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 曲線描繪 純數學 . . . 3

第 2 題 駐點 純數學 . . . 6

第 3 題 幾何極值的界限 純數學 . . . 10

第 4 題 直線交點的計數 純數學 . . . 13

第 5 題 級數與生成函數 純數學 . . . 17

第 6 題 複數與幾何 純數學 . . . 20

第 7 題 二次函數導數的界限 純數學 . . . 24

第 8 題 三維幾何 純數學 . . . 27

第 9 題 代換積分法 純數學 . . . 30

第 10 題 拋體運動 力學 . . . 33

第 11 題 質心 力學 . . . 36

第 12 題 變力作用下的運動 力學 . . . 39

第 13 題 彈性繩與碰撞 力學 . . . 43

第 14 題 刪失正態分佈的期望值 概率與統計 . . . 47

第 15 題 項分佈的假設檢驗 概率與統計 . . . 51

第 16 題 連續分佈與條件化 概率與統計 . . . 54

STEP II

第 1 題 部分分式與級數 純數學 . . . 58

第 2 題 對稱代數方程 純數學 . . . 61

第 3 題 根的對稱性 純數學 . . . 64

第 4 題 阿根圖上的軌跡 純數學 . . . 68

第 5 題 單位根與正切 純數學 . . . 73

第 6 題 雙曲函數不等式 純數學 . . . 77

第 7 題 代換積分法 純數學 81
第 8 題 矩陣與特徵向量 純數學 85
第 9 題 子群與併集 純數學 89
第 10 題 拋物面的向量幾何 純數學 92
第 11 題 繩纏繞柱子 力學 96
第 12 題 墜落的鏈條 力學 100
第 13 題 摩擦力與平衡 力學 103
第 14 題 轉動的剛體 力學 106
第 15 題 離散均勻分佈變量之和 概率與統計 110
第 16 題 聯態分佈 概率與統計 114

STEP III

第 1 題 曲線描繪與界限 純數學 117
第 2 題 線性遞推關係 純數學 121
第 3 題 複數軌跡 純數學 124
第 4 題 弧長的最優化 純數學 128
第 5 題 遞推關係與貸款 純數學 131
第 6 題 逐次微分 純數學 134
第 7 題 微分方程與正交性 純數學 137
第 8 題 拋物線的切線與法線 純數學 141
第 9 題 群元素的階 純數學 146
第 10 題 蹤曲線 純數學 149
第 11 題 摩擦及摩擦力的梯子平衡 力學 153
第 12 題 碰撞 力學 156
第 13 題 空氣阻力的拋體 力學 160
第 14 題 滑金屬線上的運動 力學 164
第 15 題 望值的遞推關係 概率與統計 168
第 16 題 幾何分佈 概率與統計 172

詳細解答

STEP I 1988，第 1 題

純數學 • 曲線描繪

> **STEP I 1988，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/88-s1-q1

關鍵思路

方程 $n^m = m^n$ 把兩個變數分別放在底數和指數的位置，糾纏在一起，處理起來並不方便。取對數(logarithm)後得到 $m\ln n = n\ln m$，再除以 mn 便可把兩者完全分開：$\frac{\ln n}{n} = \frac{\ln m}{m}$。因此我們要找兩個不同的整數，使 h 在這兩處取相同的值，而圖像的形狀正好告訴我們這樣的數對可以出現在哪裏。

解答

h 的圖像

用商法則（quotient rule）求導，

$$\mathrm{h}'(x) = \frac{x\cdot\frac{1}{x} - \ln x}{x^2} = \frac{1-\ln x}{x^2}.$$

由於 $x^2 > 0$，$\mathrm{h}'(x)$ 的正負號與 $1-\ln x$ 的正負號相同。所以 h 在 $(0,\mathrm{e})$ 上嚴格遞增（strictly increasing），在 (e,∞) 上嚴格遞減（strictly decreasing），並在 $x = \mathrm{e}$ 處取得極大值（maximum），其值為

$$\mathrm{h}(\mathrm{e}) = \frac{1}{\mathrm{e}}.$$

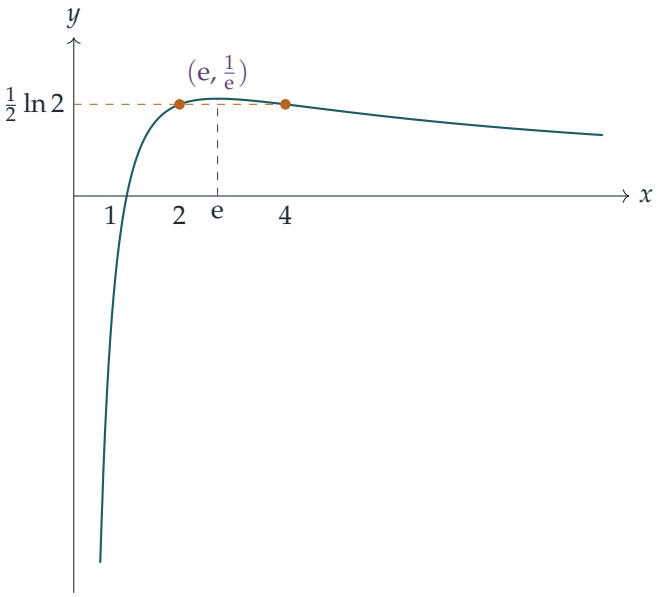
再求導一次，

$$\mathrm{h}''(x) = \frac{-\frac{1}{x}\cdot x^2 - (1-\ln x)\cdot 2x}{x^4} = \frac{2\ln x - 3}{x^3},$$

所以在 $x = \mathrm{e}^{3/2} \approx 4.48$ 處有一個拐點（point of inflection），該處 $\mathrm{h} = \frac{3}{2}\mathrm{e}^{-3/2} \approx 0.335$。曲線在這一點的左方是凹（concave）的，在右方是凸（convex）的。

接着看兩端的變化趨勢和函數的正負：

- 當 $x \to 0^+$ 時，$\ln x \to -\infty$ 而 $\frac{1}{x} \to +\infty$，所以 $\mathrm{h}(x) \to -\infty$：$y$-軸是一條漸近線（asymptote）。
- $\mathrm{h}(1) = 0$，而 $\mathrm{h}(x)$ 的正負號與 $\ln x$ 相同：在 $(0,1)$ 上為負，當 $x > 1$ 時為正。
- 當 $x \to \infty$ 時，$\frac{\ln x}{x} \to 0$（對數的增長比 x 的任何正數次冪都要慢），而且該處 $\mathrm{h}(x) > 0$，所以曲線從上方趨近 x-軸。



方程 $n^m = m^n$

由於 m 和 n 都是正數，等式兩邊均為正，可以取對數：

$$n^m = m^n \iff m\ln n = n\ln m \iff \frac{\ln n}{n} = \frac{\ln m}{m} \iff \mathrm{h}(n) = \mathrm{h}(m),$$

其中中間一步除以了 $mn > 0$。所以我們要找滿足 $\mathrm{h}(m) = \mathrm{h}(n)$ 的相異正整數。由對稱性，不妨設 $m < n$。

想一想 圖像上哪些位置可以有兩個不同的點處於同一高度？

由於 h 在 $(0, \mathrm{e}]$ 上嚴格遞增，在 $[\mathrm{e}, \infty)$ 上嚴格遞減，它在這兩個區間上各自都是一對一（one-to-one）的。所以，輸出相同的兩個相異輸入必定分處 e 的兩側：$m < \mathrm{e} < n$。由於 m 是正整數而 $\mathrm{e} \approx 2.718$，故 $m = 1$ 或 $m = 2$。

- 若 $m = 1$，則 $\mathrm{h}(n) = \mathrm{h}(1) = 0$，由此必有 $\ln n = 0$，即 $n = 1 = m$，不合要求。（直接來看：$n^1 = 1^n$ 給出 $n = 1$。）

- 若 $m = 2$，則需要一個整數 $n \geq 3$ 使得 $\mathrm{h}(n) = \frac{1}{2}\ln 2$。而 $n = 4$ 正好符合，因為

$$\mathrm{h}(4) = \frac{\ln 4}{4} = \frac{2\ln 2}{4} = \frac{\ln 2}{2} = \mathrm{h}(2),$$

  也就是 $4^2 = 16 = 2^4$。由於 h 在 $[\mathrm{e}, \infty)$ 上嚴格遞減，它在該處取 $\frac{1}{2}\ln 2$ 這個值至多一次，所以 $n = 4$ 是唯一的可能。

因此，滿足 $n^m = m^n$ 的相異正整數對只有

$$(m, n) = (2, 4) \quad \text{及} \quad (m, n) = (4, 2).$$

驗算

用電腦以精確的整數運算搜尋所有 $1 \leq m, n \leq 300$，只找到 $(2, 4)$ 和 $(4, 2)$。數值上，$\mathrm{h}(2) = \mathrm{h}(4) = 0.346574\ldots$，僅略低於極大值 $1/\mathrm{e} = 0.367879\ldots$，與圖像所示吻合。你也可以看出為甚麼 $\mathrm{h}(3)$ 得不出任何解：$\mathrm{h}(3) \approx 0.366$ 大於 $\mathrm{h}(2)$，它在 e 左方的對應點嚴格介乎 2 與 e 之間，所以不是整數。

常見錯誤

- 靠試數找到 $(2, 4)$ 便就此停筆。題目要求找出*所有*數對，所以用單調性（monotonicity）論證再沒有其他數對，才是答案的核心。
- 畫草圖時忘記漸近性態：曲線在 y-軸附近直墜至 $-\infty$，但當 $x \to \infty$ 時只是緩緩趨向 0，此後不再穿過橫軸。

STEP 之外

如果容許實數，則 $(1, \mathrm{e})$ 內的每個 a 都恰好有一個對應的 $b > \mathrm{e}$ 使得 $a^b = b^a$，而這些數對可以簡潔地參數化（parametrise）為 $a = (1 + \frac{1}{t})^t$，$b = (1 + \frac{1}{t})^{t+1}$，其中 $t > 0$；當 $t \to \infty$ 時兩者都趨向 e，也就是圖像頂端兩邊相會之處。我們用過的極限 $\frac{\ln x}{x} \to 0$ 是比較增長率的第一個例子，大學的第一門分析課程會把它嚴格化；參見 Abbott, *Understanding Analysis*（一本分析學入門教科書）。

詳細解答

STEP I 1988，第 2 題

純數學 • 駐點

> **STEP I 1988，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/88-s1-q2

關鍵思路

把 g 看成一個複合函數（composition）：$\mathrm{g} = h \circ \mathrm{f}$，其中 $h(t) = t + 1/t$。由鏈式法則（chain rule）可知，g' 只會在 $\mathrm{f}' = 0$ 或 $h'(\mathrm{f}) = 0$ 之處為零，而後者即 $\mathrm{f} = \pm 1$。把 $\cos 2x$ 用 $\sin x$ 表示，f 便成為 $\sin x$ 的二次式（quadratic），其大小和形狀由此一目了然。

解答

導數

用鏈式法則對 $\mathrm{g} = \mathrm{f} + \mathrm{f}^{-1}$ 求導，

$$\mathrm{g}'(x) = \mathrm{f}'(x) - \frac{\mathrm{f}'(x)}{\mathrm{f}(x)^2} = \mathrm{f}'(x)\left(1 - \frac{1}{\mathrm{f}(x)^2}\right).$$

再用積法則（product rule）求導一次，並利用 $\dfrac{\mathrm{d}}{\mathrm{d}x}\left(-\mathrm{f}^{-2}\right) = 2\mathrm{f}^{-3}\mathrm{f}'$，

$$\mathrm{g}''(x) = \mathrm{f}''(x)\left(1 - \frac{1}{\mathrm{f}(x)^2}\right) + \frac{2\,\mathrm{f}'(x)^2}{\mathrm{f}(x)^3}.$$

f 的大小

利用 $\cos 2x = 1 - 2\sin^2 x$，並設 $s = \sin x$，

$$\mathrm{f}(x) = 5 + 2s - 2s^2 = \tfrac{11}{2} - 2\left(s - \tfrac{1}{2}\right)^2.$$

當 s 取遍 $[-1,1]$ 時，此式在 $s=\frac{1}{2}$ 處最大（為 $\frac{11}{2}$），而在區間中離 $\frac{1}{2}$ 最遠的端點，即 $s=-1$ 處最小，其值為 1。因此

$$1 \le \mathrm{f}(x) \le \tfrac{11}{2} \quad \text{對所有} x,$$

而且只有當 $\sin x=-1$ 時才有 $\mathrm{f}(x)=1$。特別地，f 永不為零，所以 g 處處有定義。

駐點

$\mathrm{g}'(x)=0$ 當且僅當 $\mathrm{f}'(x)=0$ 或 $\mathrm{f}(x)^2=1$；滿足此條件的點就是駐點(stationary point)。

f′ 的零點。我們有

$$\mathrm{f}'(x) = -2\sin 2x + 2\cos x = 2\cos x\,(1-2\sin x),$$

它在 $\cos x=0$ 或 $\sin x=\frac{1}{2}$ 時為零；在 $[0,2\pi]$ 內即

$$x = \tfrac{\pi}{6},\ \tfrac{\pi}{2},\ \tfrac{5\pi}{6},\ \tfrac{3\pi}{2}.$$

$\mathrm{f}=\pm 1$ 的點。由於 $\mathrm{f}\ge 1$，這即是 $\mathrm{f}=1$，亦即 $\sin x=-1$，亦即 $x=\frac{3\pi}{2}$，而這一點已在上面的列表之中。

所以 g 在 $[0,2\pi]$ 內恰好有四個駐點。(兩個端點不在其中：$\mathrm{g}'(0)=\mathrm{g}'(2\pi)=2\cdot\frac{24}{25}\neq 0$。)

駐點的分類

我們有 $\mathrm{f}''(x)=-4\cos 2x-2\sin x$。在每一個駐點都有 $\mathrm{f}'=0$，所以 g'' 的第二項消失，得 $\mathrm{g}''=\mathrm{f}''\,(1-\mathrm{f}^{-2})$。

- $x=\frac{\pi}{6}$ 和 $x=\frac{5\pi}{6}$：此時 $\sin x=\frac{1}{2}$，$\cos 2x=\frac{1}{2}$，所以 $\mathrm{f}=\frac{11}{2}$ 且 $\mathrm{f}''=-2-1=-3$。於是 $\mathrm{g}''=-3\big(1-\frac{4}{121}\big)<0$：兩點都是**極大點**（maxima），並且

$$\mathrm{g}=\tfrac{11}{2}+\tfrac{2}{11}=\tfrac{125}{22}.$$

- $x=\frac{\pi}{2}$：此時 $\sin x=1$，$\cos 2x=-1$，所以 $\mathrm{f}=5$ 且 $\mathrm{f}''=4-2=2$。於是 $\mathrm{g}''=2\big(1-\frac{1}{25}\big)=\frac{48}{25}>0$：這是**極小點**（minimum），並且 $\mathrm{g}=5+\frac{1}{5}=\frac{26}{5}$。
- $x=\frac{3\pi}{2}$：此時 $\mathrm{f}=1$，所以 $1-\mathrm{f}^{-2}=0$，同時 $\mathrm{f}'=0$。因此 $\mathrm{g}''=0$。

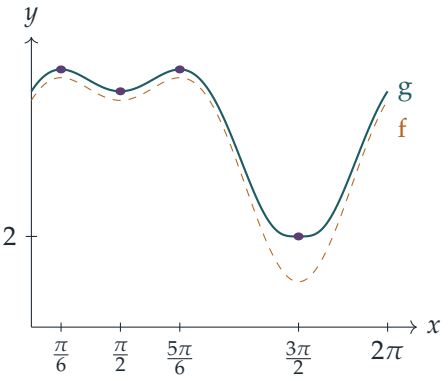
> **想一想** 二階導數判別法（second-derivative test）在 $x=\frac{3\pi}{2}$ 處失效。還有甚麼方法可以判斷？

回到 $h(t)=t+1/t$。對於 $t>0$，

$$t+\frac{1}{t}-2=\frac{(t-1)^2}{t}\ge 0,$$

等號只在 $t = 1$ 時成立。由於 $\mathrm{f} > 0$，由此可得對每一個 x 都有 $\mathrm{g}(x) \geq 2$，而等號恰好在 $\mathrm{f}(x) = 1$ 時成立。在 $x = \frac{3\pi}{2}$ 處有 $\mathrm{g} = 2$，所以這是**極小點**（事實上是 g 的全域最小值（global minimum））。

總結。極大點在 $x = \frac{\pi}{6}, \frac{5\pi}{6}$（值為 $\frac{125}{22}$）；極小點在 $x = \frac{\pi}{2}$（值為 $\frac{26}{5}$）和 $x = \frac{3\pi}{2}$（值為 2）。



驗算

- 由於 h 在 $t > 1$ 時遞增，凡是 $\mathrm{f} > 1$ 之處，g 都應該隨 f 一同升降；圖像證實 f（虛線）和 g 有相同的轉向點（turning point）。
- 數值上，$\mathrm{g}(\frac{\pi}{6}) \approx 5.6818 = \frac{125}{22}$ 而 $\mathrm{g}(\frac{\pi}{2}) = 5.2$；在 $[0, 2\pi]$ 上對 g 作細密掃描，只在 $\frac{\pi}{6}, \frac{\pi}{2}, \frac{5\pi}{6}, \frac{3\pi}{2}$ 找到轉向點，最小值為 2.000。
- g'' 的公式與在隨機選取的點上以有限差分（finite difference）計算的二階導數吻合。

常見錯誤

- 忘記因子 $1 - \mathrm{f}^{-2}$：在本題中它沒有帶來新的駐點，但這只是因為 $\mathrm{f} = 1$ 恰好發生在 $\mathrm{f}' = 0$ 之處。你仍然必須檢查這個因子。
- 由 $\mathrm{g}''(\frac{3\pi}{2}) = 0$ 就斷定該點是拐點（point of inflection）。二階導數為零，單憑這一點並不能說明甚麼。

STEP 之外

$x = \frac{3\pi}{2}$ 這個退化點在複合函數中十分典型：如果 f 取極小值之處恰好也是 h 取極小值之處，$h \circ \mathrm{f}$ 在該處就會「更平坦」，二階導數亦隨之為零。大學的數學分析（analysis）以泰勒定理（Taylor's theorem）取代二階導數判別法，按第一個非零的導數來為這類點分

類，或者（像我們所做的那樣）直接用不等式來判斷。Abbott 的 *Understanding Analysis* 嚴謹地建立中值定理（mean value theorem）和泰勒定理，是了解這些內容的好去處。

詳細解答

STEP I 1988，第 3 題

純數學 • 幾何極值的界限

> **STEP I 1988，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/88-s1-q3

關鍵思路

先問一個問題：正方形內的一點若離 O 很遠，它可以在哪裏？它必定被擠進對角附近的一小塊區域。如果 P 和 Q 都離 O 很遠，兩點便同在這一小塊區域內，因此彼此相距很近。$\sqrt{6}-\sqrt{2}$ 這個數的大小恰好令這個論證成立，而它來自一個對稱地放在正方形內的等邊三角形（equilateral triangle）。

解答

以厘米為單位，取 O 為原點，正方形為 $0 \le x \le 1$、$0 \le y \le 1$。記

$$s = \sqrt{6}-\sqrt{2}, \qquad s^2 = 6+2-2\sqrt{12} = 8-4\sqrt{3} \approx 1.072,$$

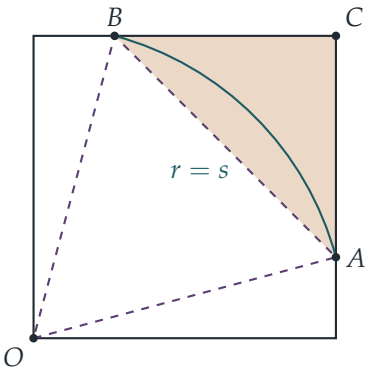
並設 $t = 2-\sqrt{3} \approx 0.268$。定義以下各點：

$$A = (1,t), \qquad B = (t,1), \qquad C = (1,1).$$

於是

$$\begin{aligned} OA^2 = OB^2 &= 1+(2-\sqrt{3})^2 = 8-4\sqrt{3} = s^2, \\ AB^2 &= 2(1-t)^2 = 2(\sqrt{3}-1)^2 = 8-4\sqrt{3} = s^2, \\ AC = BC &= 1-t = \sqrt{3}-1 \approx 0.732 < s. \end{aligned}$$

所以 OAB 是正方形內一個邊長為 s 的等邊三角形。



藍綠色的弧是以 O 為圓心、半徑為 s 的圓；正方形在這條弧以外的部分，全在塗上陰影的三角形 ABC 之內。

第 1 步：離 O 遠的點都在三角形 ABC 內

直線 AB 的方程是 $x+y=1+t$。考慮正方形在這條直線靠 O 一側的部分，即 $\{x+y\le 1+t\}$。這是一個五邊形，頂點為

$$O=(0,0),\quad (1,0),\quad A,\quad B,\quad (0,1).$$

函數 x^2+y^2 是凸函數（convex function），所以它在這個凸五邊形上的最大值必在某個頂點取得。它在各頂點的值依次是 0、1、s^2、s^2、1，其中最大的是 s^2。因此，對正方形內的每一點，

$$x+y\le 1+t \implies x^2+y^2\le s^2$$

都成立。反過來說：如果正方形內的一點 (x,y) 與 O 的距離大於 s，則 $x+y>1+t$；又因為 $x\le 1$ 及 $y\le 1$，這一點必在三角形 ABC 內。

> ? **想一想** 為甚麼三角形內兩點之間的最大距離，必定是它其中一條邊的長度？

固定 Q。距離 $|PQ|$ 是 P 的凸函數，所以在三角形上，當 P 是頂點時它取得最大值。現在固定這個頂點，讓 Q 變動：根據同樣的論證，Q 也可以取為頂點。因此，三角形內兩點之間的最大距離就是某兩個頂點之間的距離，也就是它最長的邊。

第 2 步：導出矛盾

假設 $OP>s$ 而且 $OQ>s$；否則結論已經成立。根據第 1 步，P 和 Q 都在三角形 ABC 內。這個三角形最長的邊是 $AB=s$（另外兩邊的長度是 $\sqrt{3}-1<s$）。所以根據上面方框內的結果，

$$PQ\le s.$$

由此可見，無論在哪一種情況，OP、PQ、QO 之中至少有一個不超過 $s = (\sqrt{6} - \sqrt{2})$ 厘米，這正是要證明的結論。

這個界已是最佳

取 $P = A$ 及 $Q = B$，便恰好有 $OP = PQ = QO = s$，所以任何更小的數都不行。由於 $\cos 15^\circ = \frac{\sqrt{6}+\sqrt{2}}{4}$，我們有 $s = \sec 15^\circ$：極端情況的三角形就是關於對角線對稱放置的等邊三角形，其中 OA 和 OB 各與正方形的一條邊成 15° 角。

驗算

- $s \approx 1.0353$，比正方形的邊長稍大，又遠小於對角線的長度 $\sqrt{2} \approx 1.414$，這是合理的。
- 數值上，在正方形內隨機抽取 10^6 對 P, Q，$\min(OP, PQ, QO)$ 從未超過 s（最大的約為 1.0235）。以爬山法（hill-climbing）搜尋這個最小值的最大可能值，結果收斂到 $P \approx (1, 0.26795)$、$Q \approx (0.26795, 1)$，其值為 1.035276，與 A、B 和 s 吻合。

常見錯誤

- 只論證了當 $OP, OQ > s$ 時 $\angle POQ < 60^\circ$。這是對的，但單憑這一點並不能令 PQ 成為最短的邊，所以若不再多下工夫，證明仍未完成。
- 只檢查對稱的情況。這只說明這個界不能再改進；它並沒有證明這個界對所有 P 和 Q 都成立。

STEP 之外

這是一道關於點的裝填的小型極值問題（extremal problem）：一個區域內的三點，可以同時彼此相距多遠？這類問題會引向裝填（packing）與覆蓋（covering）問題，也引向集合的*直徑*（diameter）這個概念：一個集合的直徑與它的凸包（convex hull）的直徑相同。正是這個事實令第 2 步變得輕而易舉。想做更多同類問題，可參看 Siklos 的 *Advanced Problems in Mathematics*。

詳細解答

STEP I 1988，第 4 題

純數學 • 直線交點的計數

> **STEP I 1988，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/88-s1-q4

關鍵思路

不要試圖一次過想像所有線段。不如改問：兩條線段（segment）何時會相交？答案只取決於它們的端點（endpoint）在兩條軸上的排列次序，而每一個交叉點因此都對應於在每條軸上各選兩點。只要能數出單獨一條線段上的交叉點，那些極端的線段以及最後的平均值，只需小心點算便可得出。

解答

把 y 軸上的點按與原點 O 的距離由小至大依次記為 $B_1, \ldots, B_m$，x 軸上的點亦同樣記為 $A_1, \ldots, A_n$。以 s_{ij} 表示線段 B_iA_j。

第 (i) 部分

m 個點 B_i 中的每一個都與 n 個點 A_j 中的每一個相連，所以共有

$$mn$$

條線段。

第 (ii) 部分

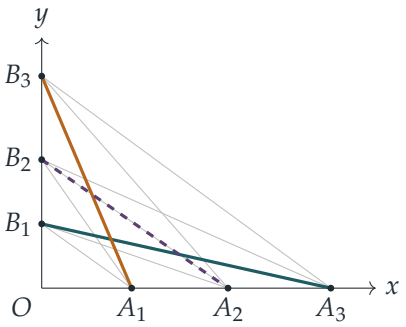
取兩條不同的線段 s_{ij} 和 s_{kl}。

- 若 $i = k$ 或 $j = l$，它們有一個公共端點。由同一點出發的兩條相異線段只會在該點相遇，而這類相遇點不計算在內。
- 若 $i < k$ 且 $j < l$，則 B_i 和 A_j 都與 O 位於直線 B_kA_l 的同一側（它們在三角形 OB_kA_l

的兩條邊 OB_k 和 OA_l 上，而不在 B_k 或 A_l 處）。若一條線段的兩個端點都嚴格位於某直線的同一側，整條線段便留在該側，所以 s_{ij} 永不與 s_{kl} 相遇：它「嵌套」在下方。

- 若 $i < k$ 且 $j > l$，則 B_i 與 O 位於直線 B_kA_l 的同一側，而 A_j 在另一側。所以 s_{ij} 穿過該直線。同理，A_l 位於直線 B_iA_j 靠 O 的一側，而 B_k 在另一側。每條線段都把另一條線段的兩個端點分隔開，所以這兩條線段恰好交於一點，而該點並非端點。

$i > k$ 的情況相同，只是角色互換。



在圖中（$m = n = 3$），藍綠色的 s_{13} 與琥珀色的 s_{31} 的端點次序相反，所以它們相交。虛線 s_{22} 亦與它們各自相交。

因此，每一對交於內點的線段，都來自選取兩點 B_i, B_k（$i < k$）和兩點 A_j, A_l（$j < l$），然後取 s_{il} 和 s_{kj}。反過來，每一個這樣的選擇都恰好給出一對相交的線段。由於沒有三條線段通過同一個內點，不同的線段對給出不同的點。因此交點（point of intersection）的數目為

$$\binom{m}{2}\binom{n}{2} = \frac{m(m-1)n(n-1)}{4}.$$

平均值

一條線段上的交叉點。根據上述法則，s_{ij} 與 s_{kl} 相交，當且僅當 $k < i,\ l > j$ 或 $k > i,\ l < j$。所以 s_{ij} 上的交點數目為

$$c_{ij} = (i-1)(n-j) + (m-i)(j-1).$$

當 $m = n$ 時的極端線段。設 $N = n-1$，$a = i-1$，$b = j-1$，則 $0 \le a, b \le N$，且

$$c_{ij} = a(N-b) + (N-a)b = N(a+b) - 2ab = \tfrac{1}{2}\big(N^2 - (N-2a)(N-2b)\big),$$

這可以展開驗證。由於 $|N-2a| \le N$ 及 $|N-2b| \le N$，乘積 $(N-2a)(N-2b)$ 介乎 $-N^2$ 與 N^2 之間。因此 $0 \le c_{ij} \le N^2$，並且：

- $c_{ij} = N^2$ 僅當 $(N-2a)(N-2b) = -N^2$，即 $\{a,b\} = \{0,N\}$。這就是 s_{1n} 和 s_{n1} 兩條線段。
- $c_{ij} = 0$ 僅當 $(N-2a)(N-2b) = N^2$，即 $a = b = 0$ 或 $a = b = N$。這就是 s_{11} 和 s_{nn} 兩條線段。

（這裏 $N \geq 2$，所以四條線段互不相同。）所以恰好有兩條線段上的點數最多，為 $(n-1)^2$，亦恰好有兩條線段上的點數最少，為 0。移除它們後，剩下 $n^2 - 4$ 條線段。

想一想 究竟是對甚麼取平均？

我們數的是移除那四條線段之後仍然存在的交點，即兩條剩餘線段之間的交叉點，每點只數一次。然後除以剩餘線段的數目。（如果你改為把每條剩餘線段上的交叉點加起來，每個點都會被數兩次，而且一些已失去的點仍會被計入。這樣會得出 $\frac{1}{2}(n-1)^2$，並非題目所給的答案。見「常見錯誤」。）

失去的點。線段 s_{11} 和 s_{nn} 上沒有交點，所以移除它們不會失去任何點。線段 s_{1n} 和 s_{n1} 上各有 $(n-1)^2$ 個點。它們互相相交（$1 < n$ 且 $n > 1$），所以恰好有一點同時位於兩者之上。根據容斥原理（inclusion–exclusion），失去的點數為 $2(n-1)^2 - 1$。

剩餘的點。由第 (ii) 部分取 $m = n$，剩餘的點數為

$$\begin{aligned}\frac{n^2(n-1)^2}{4} - 2(n-1)^2 + 1 &= \frac{(n-1)^2(n^2-8)+4}{4} \\ &= \frac{n^4 - 2n^3 - 7n^2 + 16n - 4}{4} \\ &= \frac{(n-2)(n^3 - 7n + 2)}{4}.\end{aligned}$$

至於因式分解（factorisation），注意 $n = 2$ 時該四次多項式為 $16 - 16 - 28 + 32 - 4 = 0$，所以 $n - 2$ 是一個因式。把 $(n-2)(n^3 - 7n + 2) = n^4 - 2n^3 - 7n^2 + 16n - 4$ 乘開即可確認。

平均值。除以剩餘的 $n^2 - 4 = (n-2)(n+2)$ 條線段，並約去 $n - 2 \neq 0$（因為 $n \geq 3$），得平均值為

$$\frac{(n-2)(n^3 - 7n + 2)}{4(n-2)(n+2)} = \frac{n^3 - 7n + 2}{4(n+2)},$$

正如所求。

驗算

當 $n = 3$ 時，公式給出 $\frac{8}{20} = \frac{2}{5}$。直接計算：移除 $s_{11}, s_{33}, s_{13}, s_{31}$ 後，剩下 $s_{12}, s_{21}, s_{22}, s_{23}, s_{32}$。其中只有 s_{12} 與 s_{21}，以及 s_{23} 與 s_{32}，端點次序相反，所以 5 條線段上共有 2 個點，得出

$\frac{2}{5}$。我們亦用電腦驗算：在兩條軸上隨機放置點，以精確算術找出每一個交叉點，確認了 mn、$\binom{m}{2}\binom{n}{2}$ 和 c_{ij}。此外亦確認了平均值公式對 3 至 29 的每一個 n 都成立。

常見錯誤

- 把每條剩餘線段上的交叉點加起來。每個點都位於兩條線段上，所以這樣會重複計算，而且仍然包括了與被移除線段共有的點。結果是 $\frac{1}{2}(n-1)^2$，而非題目所給的公式。
- 忘記 s_{1n} 和 s_{n1} 互相相交，因而減去 $2(n-1)^2$ 而非 $2(n-1)^2-1$。
- 把交叉點數成 $\binom{mn}{2}$ 對線段：大部分線段對不是共用一個端點，就是互相嵌套，根本不會相交。

STEP 之外

本題的核心是一個雙射（bijection）：交叉點與在每條軸上各選兩點的方法一一對應。這是組合數學（combinatorics）的基本工具。同一幅圖亦出現於完全二部圖（complete bipartite graph）$K_{m,n}$ 的「交叉數」（crossing number）之中。Zarankiewicz 的畫法（正是這一種，只是每條軸的兩側都有點）用了 $\lfloor \frac{m}{2} \rfloor \lfloor \frac{m-1}{2} \rfloor \lfloor \frac{n}{2} \rfloor \lfloor \frac{n-1}{2} \rfloor$ 個交叉點，而這是否總是最少的可能數目，至今仍是未解決的問題（open problem）。想多做以雙射計數的題目，可看 Siklos 的 *Advanced Problems in Mathematics*；至於圖論（graph theory），可看 Wilson 的 *Introduction to Graph Theory*。

詳細解答

STEP I 1988，第 5 題

純數學 • 級數與生成函數

STEP I 1988，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/88-s1-q5

關鍵思路

把每個因式展開成等比級數（geometric series）（即指數為 -1 的二項式定理（binomial theorem）），然後相乘：c_m 就是一個有限等比級數之和，有簡潔的閉合形式（closed form）。只要把 c_m^2 寫成 $(a^2)^m$、$(ab)^m$ 和 $(b^2)^m$ 這三個純冪的組合，最後的級數便會分拆成三個等比級數。餘下的工作，就是小心處理分式的代數運算，並留意每個級數在哪裏收斂（converge）。

解答

求 c_m

根據指數為 -1 的二項式定理，

$$(1-ax)^{-1} = \sum_{k=0}^{\infty} a^k x^k \quad (|ax|<1), \qquad (1-bx)^{-1} = \sum_{j=0}^{\infty} b^j x^j \quad (|bx|<1).$$

由於 $0<a<b$，條件 $b|x|<1$ 給出 $a|x|<b|x|<1$，所以兩個展開式都成立。兩個級數都絕對收斂（converge absolutely），因此我們可以把它們相乘，再按 x 的冪次合併各項。x^m 的係數來自滿足 $k+j=m$ 的各對指數：

$$c_m = \sum_{k=0}^{m} a^k b^{m-k} = b^m + ab^{m-1} + \cdots + a^{m-1}b + a^m.$$

這是一個等比級數，首項為 b^m，公比為 a/b，共有 $m+1$ 項。由於 $a \neq b$，它的和是

$$c_m = \frac{b^{m+1} - a^{m+1}}{b-a}.$$

（等價地，$(b-a)c_m = b^{m+1} - a^{m+1}$，因為把乘積展開後，各項會逐一相消（telescope）。）

c_m^2 的公式

兩邊平方，並利用 $(b-a)^2=(a-b)^2$，得

$$c_m^2=\frac{\left(b^{m+1}-a^{m+1}\right)^2}{(b-a)^2}=\frac{a^{2m+2}-2(ab)^{m+1}+b^{2m+2}}{(a-b)^2},$$

這正是所要求的結果。

級數 $\sum c_m^2x^m$

上一條公式的要點，在於每一項都是一個常數乘以一個 m 次冪：

$$c_m^2x^m=\frac{1}{(a-b)^2}\left(a^2\,(a^2x)^m-2ab\,(abx)^m+b^2\,(b^2x)^m\right).$$

所以 $\sum c_m^2x^m$ 是三個等比級數的組合，公比分別為 a^2x、abx 和 b^2x。

? **想一想** 對哪些 x，這三個等比級數全部收斂？

我們需要 $|a^2x|<1$、$|abx|<1$ 和 $|b^2x|<1$。由於 $0<a^2<ab<b^2$，最後一個條件最強，並且蘊涵另外兩個。所以三個級數全部收斂，當且僅當 $|x|<1/b^2$。這時把每個等比級數求和，得

$$\sum_{m=0}^{\infty}c_m^2x^m=\frac{1}{(a-b)^2}\left(\frac{a^2}{1-a^2x}-\frac{2ab}{1-abx}+\frac{b^2}{1-b^2x}\right).$$

把所有項通分，公分母為 $(1-abx)(1-a^2x)(1-b^2x)$。分子是

$$N=a^2(1-abx)(1-b^2x)-2ab(1-a^2x)(1-b^2x)+b^2(1-a^2x)(1-abx).$$

我們按 x 的冪次整理：

$$\begin{aligned}
\text{常數項：}\quad & a^2-2ab+b^2=(a-b)^2,\\
\text{係數，對應}x\text{：}\quad & -a^2(ab+b^2)+2ab(a^2+b^2)-b^2(a^2+ab)\\
& \qquad =a^3b-2a^2b^2+ab^3=ab(a-b)^2,\\
\text{係數，對應}x^2\text{：}\quad & a^2\cdot ab^3-2ab\cdot a^2b^2+b^2\cdot a^3b=0.
\end{aligned}$$

所以 $N=(a-b)^2(1+abx)$。因式 $(a-b)^2$ 可以約去，餘下

$$c_0^2+c_1^2x+c_2^2x^2+\cdots=\frac{1+abx}{(1-abx)(1-a^2x)(1-b^2x)}\qquad\text{適用於}-\frac{1}{b^2}<x<\frac{1}{b^2}.$$

這個區間不能再擴大。在 $x=\pm1/b^2$ 處，各項滿足

$$|c_m^2x^m|=\frac{\left(b^{m+1}-a^{m+1}\right)^2}{(b-a)^2\,b^{2m}}=\frac{b^2\left(1-(a/b)^{m+1}\right)^2}{(b-a)^2}\to\frac{b^2}{(b-a)^2}\neq 0,$$

所以級數在該處發散（diverge）；當 $|x|>1/b^2$ 時級數同樣發散，因為這時各項會無限增大。

驗算

取 $a = 1$，$b = 2$。這時 $c_m = 2^{m+1} - 1$，所以 $c_0, c_1, c_2 = 1, 3, 7$，而直接計算可得 $\frac{1}{(1-x)(1-2x)} = (1 + x + x^2 + \ldots)(1 + 2x + 4x^2 + \ldots) = 1 + 3x + 7x^2 + \cdots$。它們的平方是 $1, 9, 49$。另一方面，$\frac{1}{(1-x)(1-2x)(1-4x)} = 1 + 7x + 35x^2 + \cdots$，再乘以 $1 + 2x$ 便得到 $1 + 9x + 49x^2 + \cdots$，與預期相符。數值上，取 $a = 0.7$、$b = 1.3$ 和 $x = 0.3$，把 $c_m^2 x^m$ 加至很多項，得 $4.163878\ldots$，與閉合形式的值吻合。

常見錯誤

- 把區間寫成 $b|x| < 1$，即照抄第一部分的條件。新級數的公比是 a^2x、abx、b^2x，所以條件應為 $b^2|x| < 1$。
- 讓 c_m 停留在求和的形式。閉合形式 $(b^{m+1} - a^{m+1})/(b - a)$ 正是令題目餘下部分得以完成的關鍵。
- 在 N 中 x 的係數出現正負號錯誤。老老實實地把每個乘積逐項展開，比嘗試取巧更快。

STEP 之外

級數 $\sum c_m x^m$ 是數列 c_m 的生成函數（*generating function*），而這道題說明：如果一個數列有有理的生成函數，那麼由它各項的平方組成的數列也是如此。這是兩個有理冪級數的阿達馬乘積（*Hadamard product*）的一個特例，而這種乘積必定是有理的。這裏求得的區間是收斂半徑（*radius of convergence*）的一個例子：有理函數的冪級數（power series）一直收斂到最近的極點（pole）為止，在這裏就是 $x = 1/b^2$。關於生成函數，Wilf 的 *generatingfunctionology* 有十分精彩的闡述；至於冪級數及其收斂性，Abbott 的 *Understanding Analysis* 有嚴謹細緻的討論。

詳細解答

STEP I 1988，第 6 題

純數學 • 複數與幾何

> **STEP I 1988，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/88-s1-q6

關鍵思路

在阿根圖（Argand diagram）上，向量 $\overrightarrow{P_jP_k}$ 就是複數（complex number）$z_k - z_j$，所以向量相等變成差相等，而垂直（perpendicular）則變成「一個差除以另一個差所得的商是純虛數（purely imaginary）」。把題目給定的數值代入後，條件 (i) 會以 z 表示出 z_4 和 w，而條件 (ii) 則決定了 z 的虛部（imaginary part）。這時會發現六邊形關於其中心對稱，凸性（convexity）和最後的角度因而都容易處理。

解答

把題目的陳述翻譯成代數

由於 $\overrightarrow{P_jP_k}$ 由 $z_k - z_j$ 表示，陳述 (i) 即是說

$$z_2 - z_1 = z_4 - z_5 \qquad 及 \qquad z_3 - z_2 = z_5 - z_6.$$

至於 (ii)，兩個非零向量 u 和 v 互相垂直，當且僅當它們的輻角（argument）相差 $\pm\frac{\pi}{2}$，亦即 u/v 是純虛數。各點互不相同，所以 $z_4 - z_2 \neq 0$ 及 $z_6 - z_3 \neq 0$，而陳述 (ii) 即是說

$$\frac{z_4 - z_2}{z_6 - z_3} \text{ 是純虛數，} \qquad 等價於 \qquad \operatorname{Re}\left[(z_4 - z_2)\,\overline{(z_6 - z_3)}\right] = 0.$$

(i) 的推論

把 (i) 的兩條等式相加，z_2 和 z_5 兩項互相抵消：

$$z_3 - z_1 = z_4 - z_6, \qquad 所以 \qquad z_4 - z_3 = z_6 - z_1,$$

這正是 $\overrightarrow{P_3P_4} = \overrightarrow{P_1P_6}$，即所要證明的結果。

六邊形

現在設 $z_1 = 0$、$z_2 = 1$、$z_3 = z$、$z_5 = \mathrm{i}$ 及 $z_6 = w$。由 (i) 的第一條等式得

$$z_4 = z_5 + z_2 - z_1 = 1 + \mathrm{i},$$

而由第二條等式得 $z - 1 = \mathrm{i} - w$，所以

$$w = 1 + \mathrm{i} - z.$$

留意 $z_1 + z_4 = z_2 + z_5 = z_3 + z_6 = 1 + \mathrm{i}$：相對的頂點關於點 $\frac{1}{2}(1 + \mathrm{i})$ 對稱，所以這個六邊形有對稱中心（centre of symmetry）。

至於 (ii)，$z_4 - z_2 = \mathrm{i}$，而 $z_6 - z_3 = w - z = 1 + \mathrm{i} - 2z$。商 $\mathrm{i}/(1 + \mathrm{i} - 2z)$ 是純虛數，當且僅當 $1 + \mathrm{i} - 2z$ 是實數，亦即 $1 - 2\,\mathrm{Im}(z) = 0$。所以

$$\mathrm{Im}(z) = \tfrac{1}{2}, \qquad \text{於是} \qquad \mathrm{Im}(w) = 1 - \tfrac{1}{2} = \tfrac{1}{2}.$$

設 $z = a + \frac{1}{2}\mathrm{i}$，其中 a 為實數。那麼 $w = (1 - a) + \frac{1}{2}\mathrm{i}$，所以 $\mathrm{Re}(w) = 1 - \mathrm{Re}(z)$ 恆成立。六個頂點是

$$(0,0), \quad (1,0), \quad (a, \tfrac{1}{2}), \quad (1,1), \quad (0,1), \quad (1 - a, \tfrac{1}{2}) :$$

即單位正方形在兩條垂直邊的半高處各加上一個「尖角」。

想一想 怎樣才可以不靠圖形來檢驗凸性？

順着次序沿六邊形走一圈。六邊形是凸的，當且僅當我們在每個頂點都向同一方向轉彎。在某個頂點的轉向，可用進入該頂點的邊 u 與離開該頂點的邊 v 的叉積（cross product）$\mathrm{Im}(\bar{u}v) = u_x v_y - u_y v_x$ 的正負號來量度。各邊是

$$P_1P_2 = (1,0), \quad P_2P_3 = (a-1, \tfrac{1}{2}), \quad P_3P_4 = (1-a, \tfrac{1}{2}),$$
$$P_4P_5 = (-1,0), \quad P_5P_6 = (1-a, -\tfrac{1}{2}), \quad P_6P_1 = (a-1, -\tfrac{1}{2}),$$

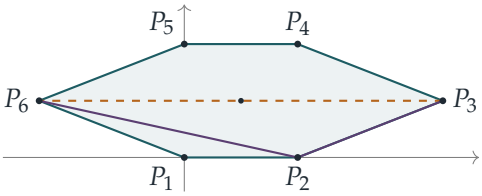
而在 $P_2, P_3, P_4, P_5, P_6, P_1$ 的叉積依次為

$$\tfrac{1}{2}, \qquad a - 1, \qquad \tfrac{1}{2}, \qquad \tfrac{1}{2}, \qquad a - 1, \qquad \tfrac{1}{2}.$$

無論 a 取甚麼值，其中四個都是正數，所以六邊形是凸的，當且僅當 $a - 1 > 0$。（若 $a = 1$，則 P_3 位於線段 P_2P_4 上，得不到真正的六邊形。）從幾何上看，P_3 必須位於正方形的邊 $x = 1$ 的右方，而根據對稱性，P_6 必須位於 $x = 0$ 的左方。所以條件是

$$\boxed{\mathrm{Re}(z) > 1 \quad \text{及} \quad \mathrm{Re}(w) = 1 - \mathrm{Re}(z) \ \ (\text{所以} \mathrm{Re}(w) < 0)}$$

（連同 $\mathrm{Im}(z) = \mathrm{Im}(w) = \frac{1}{2}$，這已由 (ii) 所決定）。



距離 P_3P_6

在 P_2 處，角的兩臂是

$$z_3 - z_2 = (a-1) + \tfrac{1}{2}\mathrm{i}, \qquad z_6 - z_2 = -a + \tfrac{1}{2}\mathrm{i}.$$

對於角 $\theta = \angle P_3P_2P_6$（介乎 0 與 π 之間），兩臂的點積（dot product）給出 $|P_2P_3|\,|P_2P_6|\cos\theta$，而叉積的大小給出 $|P_2P_3|\,|P_2P_6|\sin\theta$：

$$\text{點積} = -a(a-1) + \tfrac{1}{4} = \tfrac{1}{4}(1 + 4a - 4a^2), \qquad \text{叉積} = (a-1)\cdot\tfrac{1}{2} - \tfrac{1}{2}\cdot(-a) = \tfrac{1}{2}(2a-1).$$

當 $a > 1$ 時叉積為正數，所以它正好就是含 $\sin\theta$ 的那一項，於是

$$\tan\theta = \frac{\frac{1}{2}(2a-1)}{\frac{1}{4}(1+4a-4a^2)} = \frac{2(2a-1)}{1+4a-4a^2}.$$

令它等於 $-\frac{2}{3}$：

$$3(2a-1) = -(1+4a-4a^2) \quad\Longrightarrow\quad 4a^2 - 10a + 2 = 0 \quad\Longrightarrow\quad 2a^2 - 5a + 1 = 0,$$

所以 $a = \frac{1}{4}(5 \pm \sqrt{17})$。取負號得 $a \approx 0.22$，不符合凸性條件 $a > 1$；取正號得 $a \approx 2.28 > 1$。（這時分母 $1 + 4a - 4a^2$ 是負數，所以 θ 是鈍角，正切為負數時正應如此。）

最後，P_3 和 P_6 的虛部相同，所以

$$P_3P_6 = |z - w| = |2a - 1| = 2a - 1 = \frac{5+\sqrt{17}}{2} - 1 = \boxed{\frac{3+\sqrt{17}}{2}} \approx 3.56.$$

如果不要求六邊形是凸的，用同樣的方法（這時分子要改為 $|2a-1|$）會多得一個解 $a = \frac{1}{4}(-1 - \sqrt{17})$。這只是鏡像 $a \mapsto 1 - a$，它把 z 和 w 互換，所得的距離同樣是 $\frac{1}{2}(3 + \sqrt{17})$。

驗算

- 對稱性：$z_1 + z_4 = z_2 + z_5 = z_3 + z_6$，所以結果 $\overrightarrow{P_3P_4} = \overrightarrow{P_1P_6}$ 不過是中心對稱六邊形的第三對對邊。

- 取 $a = \frac{1}{4}(5+\sqrt{17}) \approx 2.2808$，直接計算 $\arg\left((z_6 - z_2)/(z_3 - z_2)\right)$，所得的角的正切是 -0.66667，而 $|z - w| \approx 3.5616 = \frac{1}{2}(3+\sqrt{17})$。
- 對由 -5 至 5 的 a 以數值方法檢驗凸性，確認六邊形是凸的當且僅當 $a > 1$。

常見錯誤

- 弄錯向量的方向：$\overrightarrow{P_5P_4}$ 是 $z_4 - z_5$，而不是 $z_5 - z_4$。
- 把垂直寫成「斜率之積為 -1」，結果遺漏了鉛垂的情況。這裏 $\overrightarrow{P_2P_4}$ 是鉛垂的，而用商來檢驗就能乾淨利落地處理這種情況。
- 保留了根 $a = \frac{1}{4}(5-\sqrt{17})$，但它並不給出凸六邊形（事實上這時 $\tan\theta > 0$）。

STEP 之外

把點看成複數，幾何就變成代數：平移是加法，旋轉和縮放是乘法。「u/v 是純虛數」這個檢驗，正是角度化為輻角的一個例子。以叉積的正負號檢驗凸性，是計算幾何（computational geometry）中標準凸包（convex hull）演算法的基礎。Needham 的 *Visual Complex Analysis*（《複分析：可視化方法》）把這種以幾何看待複數的觀點發揮得淋漓盡致。

詳細解答

STEP I 1988，第 7 題

純數學 • 二次函數導數的界限

> **STEP I 1988，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/88-s1-q7

本題在 *Beyond STEP* 第 3.6 節斜率的界中有完整的引導式講解：https://beyondstep.keithmcnulty.org/zh-hant/ch03/markov-bound.html

關鍵思路

二次函數（quadratic）由它在三個點的取值完全確定，所以 f' 可以由 $f(1)$、$f(-1)$ 和 $f(0)$ 重新構造出來。一旦把 f' 寫成這個形式，假設 $|f| \le 1$ 便控制了三個係數中的每一個，餘下的工作交給三角不等式（triangle inequality）就行了。至於最後一部分，我們考察 f 取哪些值時，證明中的每一個不等式都變成等式。

解答

$f'(x)$ 的恆等式

我們有 $f(1) = a+b+c$、$f(-1) = a-b+c$ 和 $f(0) = c$。把右邊按 x 的冪次整理：

$$\begin{aligned} &f(1)\left(x+\tfrac{1}{2}\right) + f(-1)\left(x-\tfrac{1}{2}\right) - 2f(0)x \\ &\quad = \left[(a+b+c)+(a-b+c)-2c\right]x + \tfrac{1}{2}\left[(a+b+c)-(a-b+c)\right] \\ &\quad = 2ax + b. \end{aligned}$$

由於 $f'(x) = 2ax+b$，這就是所求的恆等式（identity）。

上界 $|f'(x)| \le 4$

假設只要 $|x| \le 1$，就有 $|f(x)| \le 1$。特別地，$|f(1)|$、$|f(-1)|$ 和 $|f(0)|$ 全都不超過 1。任取滿足 $|x| \le 1$ 的 x。由恆等式和三角不等式，

$$\begin{aligned} |f'(x)| &\le |f(1)|\,\left|x+\tfrac{1}{2}\right| + |f(-1)|\,\left|x-\tfrac{1}{2}\right| + 2|f(0)|\,|x| \\ &\le \left|x+\tfrac{1}{2}\right| + \left|x-\tfrac{1}{2}\right| + 2|x|. \end{aligned}$$

> **想一想** 當 $|x| \le 1$ 時，$\left|x+\frac{1}{2}\right| + \left|x-\frac{1}{2}\right|$ 最大可以是多少？

這是 x 到 $\frac{1}{2}$ 和到 $-\frac{1}{2}$ 的距離之和。若 $|x| \le \frac{1}{2}$，則 x 位於這兩點之間，距離之和恰好是 1。若 $|x| > \frac{1}{2}$，兩點都在 x 的同一側，距離之和是 $2|x|$。無論哪種情況，當 $|x| \le 1$ 時它都不超過 2。另外 $2|x| \le 2$，所以

$$|f'(x)| \le 2 + 2 = 4 \qquad \text{對於} |x| \le 1,$$

這正是要證明的結果。

（另一個看法：f' 是線性的，所以在 $[-1,1]$ 上它的絕對值（modulus）在端點處最大，而在 $x = \pm 1$ 處的上界是 $|f'(\pm 1)| \le \frac{3}{2} + \frac{1}{2} + 2 = 4$。）

等式成立的一個例子

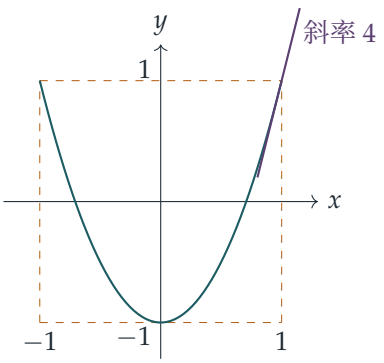
要在 $x = 1$ 處取得等式，上面的每一步在該處都必須恰好取等。在 $x = 1$ 處，係數 $x+\frac{1}{2}$、$x-\frac{1}{2}$ 和 $-2x$ 分別是 $\frac{3}{2}$、$\frac{1}{2}$ 和 -2。因此，要得到 $f'(1) = +4$，我們希望

$$f(1) = 1, \qquad f(-1) = 1, \qquad f(0) = -1.$$

由此得 $c = -1$、$a + b + c = 1$ 和 $a - b + c = 1$，所以 $b = 0$ 而 $a = 2$：

$$a = 2, \quad b = 0, \quad c = -1, \qquad f(x) = 2x^2 - 1.$$

現在檢查條件。對於 $|x| \le 1$，我們有 $0 \le x^2 \le 1$，所以 $-1 \le 2x^2 - 1 \le 1$，即 $|f(x)| \le 1$。而 $f'(x) = 4x$，所以 $f'(1) = 4$，且 $x = 1$ 在區間之內。這就是我們需要的例子。



曲線 $y = 2x^2 - 1$（藍綠色）始終留在正方形 $|x| \le 1$、$|y| \le 1$ 之內，由頂邊盪到底邊再盪回頂邊，因此它在兩端達到了可能的最大陡峭程度。

（$\pm(2x^2 - 1)$ 中任何一個都給出 $|f'(\pm 1)| = 4$。選取 $f(x) = 2x^2 - 1$ 會得到 $f'(1) = +4$，而 $-(2x^2 - 1)$ 在 $x = -1$ 處也是如此。）

驗算

- 試用 $f(x) = x^2 + x + 1$：$f(1) = 3$、$f(-1) = 1$、$f(0) = 1$。在 $x = 2$ 處，恆等式給出 $3 \cdot \frac{5}{2} + 1 \cdot \frac{3}{2} - 2 \cdot 2 = 5 = f'(2)$。
- 對 20,000 個二次函數作隨機搜尋，每個都經縮放使得 $\max_{|x|\le 1} |f(x)| = 1$，結果從未發現 $\max_{|x|\le 1} |f'(x)| = 2|a| + |b|$ 超過 4（最大的約為 3.91），而 $2x^2 - 1$ 則恰好達到 4。

常見錯誤

- 把 $\left|x + \frac{1}{2}\right|$ 和 $\left|x - \frac{1}{2}\right|$ 分開估計，各以 $\frac{3}{2}$ 為上界。這樣只能得到 $|f'(x)| \le 5$。這兩個距離必須合起來一併估計。
- 提出例子後，只在 0 和 ± 1 處檢查，而沒有在整個區間上驗證 $|f(x)| \le 1$。

STEP 之外

這是馬爾可夫不等式（Markov's inequality）在二次情況下的特例：若 n 次多項式（polynomial）p 在 $[-1, 1]$ 上滿足 $|p| \le 1$，則在該區間上 $|p'| \le n^2$。等式由切比雪夫多項式（Chebyshev polynomial）$T_n(\cos\theta) = \cos n\theta$ 取得，而 $2x^2 - 1 = T_2(x)$。利用多項式在選定點上的取值來把它寫出，就是拉格朗日插值法（Lagrange interpolation），它是數值分析（numerical analysis）和逼近理論（approximation theory）的基本工具。想做更多同類風格的題目，可參閱 Siklos 的 *Advanced Problems in Mathematics*（進階數學難題集）。

詳細解答

STEP I 1988，第 8 題

純數學 • 三維幾何

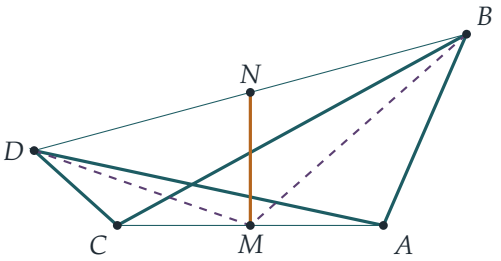
> **STEP I 1988，第 8 題**
>
> 繼續閱讀之前，請先看題目（原題為英文）。題目載於 STEP Support Programme 的試題資料庫：
>
> step.maths.org/questions/88-s1-q8

關鍵思路

每條對角線都把四邊形分成兩個三角形，而對邊相等令這兩個三角形全等（congruent）（SSS）。全等三角形的中線（median）相等，所以一條對角線的中點與另一條對角線的兩個端點等距。若某點與一條線段的兩個端點等距，把它與該線段的中點連起來，所得的直線便垂直於該線段：這其實就是等腰三角形（isosceles triangle）的性質。

解答

設 M 為 AC 的中點，N 為 BD 的中點。



上圖是一個空間四邊形（skew quadrilateral）的投影；虛線 MB 和 MD 是下文用到的中線，琥珀色的線段是 MN。

首先， $M \neq N$

只有當 M 和 N 是不同的點時，直線 MN 才有意義。假如 $M = N$，對角線 AC 和 BD 便會在一個公共點互相平分。過同一點的兩條直線位於同一平面上，所以 A、B、C、D

全都在該平面上，而 $ABCD$ 就會是一個（平面的）平行四邊形（parallelogram）。由於 $ABCD$ 是空間四邊形，故 $M \neq N$。「空間」這個條件正正用在這裏：平行四邊形的對邊相等，但兩個中點卻重合。

MN **垂直於** BD

考慮三角形 ABC 和 CDA，它們有公共的對角線 AC。我們有

$$AB = CD, \qquad BC = DA, \qquad CA = AC,$$

所以這兩個三角形全等（SSS），其中 $A \leftrightarrow C$、$B \leftrightarrow D$、$C \leftrightarrow A$。在這個對應之下，三角形 ABC 中由 B 到 AC 中點的中線，對應於三角形 CDA 中由 D 到 CA 中點的中線。兩條中線的另一端都是同一點 M，所以

$$MB = MD.$$

想一想 為甚麼全等三角形的中線必定相等？

中線完全由邊長決定。例如，根據餘弦定理（cosine rule）（或阿波羅尼奧斯定理（Apollonius's theorem）），三角形 ABC 中由 B 引出的中線 m 滿足 $4m^2 = 2AB^2 + 2BC^2 - AC^2$。邊長相同，中線也就相同，所以 $MB^2 = \frac{1}{4}(2AB^2 + 2BC^2 - AC^2) = \frac{1}{4}(2CD^2 + 2DA^2 - CA^2) = MD^2$。

現在 M 與 B 和 D 等距，而 N 是 BD 的中點。三角形 MNB 和 MND 有 $MB = MD$、$NB = ND$，以及公共邊 MN，所以它們全等（SSS），而 $\angle MNB$ 與 $\angle MND$ 兩角相等。又因為 B、N、D 在同一直線上，這兩個角互補（supplementary）。所以兩者都是直角：$MN \perp BD$。（這兩個三角形是真正的三角形，並非退化（degenerate）：假如 M 在直線 BD 上而又有 $MB = MD$，M 便會是中點 N，而這個情況我們已經排除了。）

MN **垂直於** AC

把兩條對角線的角色對調，同樣的論證依然成立。三角形 ABD 和 CDB 有公共的對角線 BD，而且

$$AB = CD, \qquad AD = CB, \qquad BD = DB,$$

所以它們全等，其中 $A \leftrightarrow C$、$B \leftrightarrow D$、$D \leftrightarrow B$。由 A 和由 C 引出的中線，另一端都是 BD 的中點 N，所以 $NA = NC$。由於 M 是 AC 的中點，而且 $N \neq M$，由上文的等腰三角形論證可得 $MN \perp AC$。

因此，連接兩條對角線中點的直線垂直於每一條對角線，這正是所要證明的。

或者用向量

取 M 為原點，於是以位置向量（position vector）表示時有 $C = -A$，而 $\overrightarrow{MN} = \frac{1}{2}(\mathbf{b} + \mathbf{d})$。把題目給出的兩個等式平方，

$$AB^2 = CD^2 \iff |\mathbf{b}|^2 - 2\mathbf{a}\cdot\mathbf{b} = |\mathbf{d}|^2 + 2\mathbf{a}\cdot\mathbf{d},$$
$$BC^2 = AD^2 \iff |\mathbf{b}|^2 + 2\mathbf{a}\cdot\mathbf{b} = |\mathbf{d}|^2 - 2\mathbf{a}\cdot\mathbf{d}.$$

兩式相減得 $4\mathbf{a}\cdot\mathbf{b} = -4\mathbf{a}\cdot\mathbf{d}$，即 $\mathbf{a}\cdot(\mathbf{b}+\mathbf{d}) = 0$。由於 $\overrightarrow{AC} = -2\mathbf{a}$，這就是說 $\overrightarrow{MN} \perp AC$。改為把原點放在 N，便可用同樣的方法得到 $MN \perp BD$。

驗算

試取 $A = (1,0,0)$、$B = (1,2,1)$、$C = (-1,0,0)$、$D = (-1,-2,1)$。這時 $AB = CD = \sqrt{5}$，$BC = AD = 3$，而且四點不共面（純量三重積（scalar triple product）$\overrightarrow{AB}\cdot(\overrightarrow{AC}\times\overrightarrow{AD})$ 等於 8）。這裏 $M = (0,0,0)$，$N = (0,0,1)$，所以 MN 沿着 z 軸，而 AC 沿着 x 軸，$\overrightarrow{BD} = (-2,-4,0)$ 則沒有 z 分量：兩者都垂直於 MN。我們還測試了 20 000 個對邊相等的隨機空間四邊形；MN 與任何一條對角線之間夾角的餘弦值，從未超過 4×10^{-12}。

常見錯誤

- 沒有解釋為甚麼 $M \neq N$。缺少了「空間四邊形」這個條件，結論便不成立，因為平行四邊形根本沒有直線 MN。
- 聲稱「根據對稱性」有 $MB = MD$，卻沒有指明是哪兩個全等三角形。
- 只證明了與其中一條對角線垂直。第二條需要用另一對全等三角形。

STEP 之外

在上面的例子中，繞 z 軸（即直線 MN）的半周旋轉（half-turn）把 A 與 C 互換，也把 B 與 D 互換。這在一般情況下也成立：每一個對邊相等的空間四邊形，繞 MN 旋轉 180° 後都會與自身重合。（這個旋轉固定 M，並把每一個垂直於 MN 的方向反轉，所以它把 A 送到 C；同樣地，它把 B 送到 D。）反過來說，有了這個半周旋轉，對邊相等便一目了然。把幾何事實看成某種對稱（symmetry）——即空間的一個等距變換（isometry）——的結果，是大學一年級幾何和羣論（group theory）的核心思想；Allenby 的 *Rings, Fields and Groups*（環、域與羣）闡述了羣論方面的內容，而 Siklos 的 *Advanced Problems in Mathematics*（高等數學難題）則有更多 STEP 風格的幾何題。

詳細解答

STEP I 1988，第 9 題

純數學 • 代換積分法

STEP I 1988，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/88-s1-q9

關鍵思路

只要選對着手的部分，兩個積分都可以用標準技巧解決。在 (i) 中，$\ln x$ 與 x 的冪相乘，適合使用分部積分法（integration by parts）：對對數求導，對數便會消失。在 (ii) 中，棘手的是平方根，所以我們把它設為新變量。$\cos x\,\mathrm{d}x$ 正好是代換（substitution）所需要的，而餘下的是一個有理函數（rational function），可以用部分分式（partial fractions）處理。

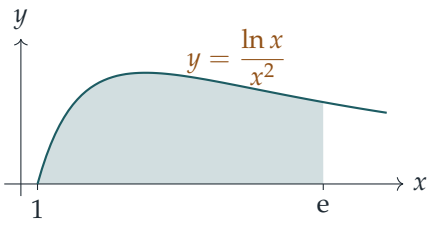
解答

第 (i) 部分

取 $u = \ln x$ 及 $\dfrac{\mathrm{d}v}{\mathrm{d}x} = \dfrac{1}{x^2}$ 作分部積分，於是 $\dfrac{\mathrm{d}u}{\mathrm{d}x} = \dfrac{1}{x}$ 及 $v = -\dfrac{1}{x}$：

$$\begin{aligned}\int_1^{\mathrm{e}} \frac{\ln x}{x^2}\,\mathrm{d}x &= \left[-\frac{\ln x}{x}\right]_1^{\mathrm{e}} + \int_1^{\mathrm{e}} \frac{1}{x^2}\,\mathrm{d}x \\ &= \left(-\frac{1}{\mathrm{e}} - 0\right) + \left[-\frac{1}{x}\right]_1^{\mathrm{e}} \\ &= -\frac{1}{\mathrm{e}} + 1 - \frac{1}{\mathrm{e}} = 1 - \frac{2}{\mathrm{e}}.\end{aligned}$$

這裏用到 $\ln \mathrm{e} = 1$ 及 $\ln 1 = 0$。答案是正數，這是必然的：被積函數（integrand）在 $(1, \mathrm{e}]$ 上為正。



陰影部分的面積為 $1-2/\mathrm{e}\approx 0.264$。

第 (ii) 部分

被積函數在 $\sin x\neq 0$ 且 $1+\sin x>0$ 時有定義。設

$$u=\sqrt{1+\sin x},\qquad \text{於是}u^2=1+\sin x,\quad 2u\,\mathrm{d}u=\cos x\,\mathrm{d}x,\quad \sin x=u^2-1.$$

這裏 $u>0$，而且因為 $\sin x\neq 0$，所以 $u\neq 1$。於是

$$\int\frac{\cos x}{\sin x\sqrt{1+\sin x}}\,\mathrm{d}x=\int\frac{2u}{(u^2-1)\,u}\,\mathrm{d}u=\int\frac{2}{u^2-1}\,\mathrm{d}u.$$

? **想一想** $\cos x$ 到哪裏去了？我們需要把它寫成 u 的式子嗎？

不需要。整個 $\cos x\,\mathrm{d}x$ 都由 $2u\,\mathrm{d}u$ 取代，這正是這個代換如此乾淨利落的原因。假如分子是別的東西，我們便要用到 $\cos x=\pm\sqrt{1-\sin^2 x}$，還得應付隨之而來的種種正負號問題。

由部分分式得 $\dfrac{2}{u^2-1}=\dfrac{1}{u-1}-\dfrac{1}{u+1}$，所以

$$\int\frac{2}{u^2-1}\,\mathrm{d}u=\ln|u-1|-\ln|u+1|+C=\ln\left|\frac{u-1}{u+1}\right|+C.$$

換回 x，

$$\int\frac{\cos x}{\sin x\sqrt{1+\sin x}}\,\mathrm{d}x=\ln\left|\frac{\sqrt{1+\sin x}-1}{\sqrt{1+\sin x}+1}\right|+C.$$

絕對值（modulus）符號不可省略：當 $-1<\sin x<0$ 時，我們有 $0<u<1$，而 $\dfrac{u-1}{u+1}$ 是負數。（嚴格來說，由於定義域被 $\sin x$ 的零點分割成多個區間，在被積函數有定義的每個區間上，常數可以各不相同。）

把分數的分子和分母同乘以 $u-1$，可得一個等價的形式。由於 $(u+1)(u-1)=u^2-1=\sin x$，答案亦可寫成

$$2\ln\left|\sqrt{1+\sin x}-1\right|-\ln|\sin x|+C.$$

驗算

- 用辛普森法則（Simpson's rule）得 $\int_1^{\mathrm{e}} x^{-2}\ln x\,\mathrm{d}x \approx 0.264241$，與 $1-2/\mathrm{e} = 0.264241\ldots$ 吻合。
- 對於 (ii)，把答案求導：由鏈式法則（chain rule），$\dfrac{\mathrm{d}}{\mathrm{d}x}\ln\left|\dfrac{u-1}{u+1}\right| = \dfrac{2}{u^2-1}\cdot\dfrac{\cos x}{2u}$，這正是被積函數。數值上，用辛普森法則算得被積函數的 $\int_{0.3}^{1.2}$ 為 0.926006，與原函數（antiderivative）在兩端的差一致。在 $\sin x < 0$ 的區間上，結果同樣吻合，例如在 $[-1.2,-0.2]$ 上，積分為 -2.361027。

常見錯誤

- 在 (i) 中，反過來選取 $u = 1/x^2$ 及 $\mathrm{d}v/\mathrm{d}x = \ln x$。這樣做原則上可行，但繁複得多；應該對對數求導。
- 在 (ii) 中，漏掉絕對值符號，以致答案在 $\sin x < 0$ 時沒有定義。
- 漏掉 $2u\,\mathrm{d}u$ 中的因子 2，令答案減半。

STEP 之外

第 (ii) 部分是一個一般原理的例子：若被積函數是 x 與 $\sqrt{ax+b}$ 的有理函數，經代換 $u = \sqrt{ax+b}$ 後便會變成 u 的有理函數，而有理函數總可以用部分分式來積分。這裏 $\sin x$ 扮演了 x 的角色。在大學一年級的分析（analysis）課程中，你會看到故事的另一面：那些為代換法和分部積分法提供依據的定理，以及積分常數（constant of integration）為何可以在定義域的各個分離區間上有所不同。Bartle 與 Sherbert 的 *Introduction to Real Analysis*（實分析入門）對這兩方面都有細緻的論述。

詳細解答

STEP I 1988，第 10 題

力學 • 拋體運動

> **STEP I 1988，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/88-s1-q10

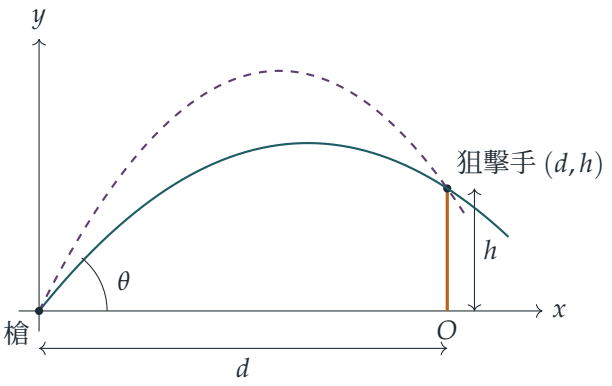
關鍵思路

把問題反過來想。設槍與樹腳的水平距離為 d，然後問：是否存在任何一個仰角，使速率為 v 的子彈剛好經過樹頂？以 $\tan\theta$ 表示的軌跡方程，是 $\tan\theta$ 的二次方程，所以答案為「是」的充要條件，就是該二次方程有實根。由判別式（discriminant）得出的條件，原來是一個只涉及 d 的不等式。

解答

模型

我們把子彈視為只受重力作用的質點（particle）（不計空氣阻力），由地面以速率 v、與水平成 θ 的仰角（angle of elevation）射出，其中 $0 < \theta \leq \frac{\pi}{2}$。由於槍的位置未知，設它與樹腳 O 的距離為 $d \geq 0$，並以槍為原點建立坐標軸，x 軸水平指向樹，y 軸鉛垂向上。於是樹頂就是點 (d, h)。



擊中狙擊手

經過時間 t 後，子彈位於 $x = vt\cos\theta$，$y = vt\sin\theta - \frac{1}{2}gt^2$。若 $\theta < \frac{\pi}{2}$，我們以 $t = x/(v\cos\theta)$ 消去時間，並利用 $\sec^2\theta = 1 + \tan^2\theta$，得到軌跡（trajectory）

$$y = x\tan\theta - \frac{gx^2}{2v^2}\left(1 + \tan^2\theta\right).$$

子彈擊中狙擊手，當且僅當 (d,h) 位於這條曲線上。先假設 $d > 0$。記 $T = \tan\theta$ 及 $k = \dfrac{gd^2}{2v^2} > 0$，條件就是

$$kT^2 - dT + (h+k) = 0.$$

每個實數 T 都恰好是唯一一個 $\theta \in (-\frac{\pi}{2}, \frac{\pi}{2})$ 的 $\tan\theta$。不過，根會不會是負數呢？兩根之積為 $(h+k)/k > 0$，兩根之和為 $d/k > 0$，所以只要有實根，兩根都必為正數，對應真正的仰角 $0 < \theta < \frac{\pi}{2}$。因此，存在可行的仰角，當且僅當這個二次方程有實根，亦即其判別式非負：

$$\begin{aligned} d^2 - 4k(h+k) \geq 0 &\iff d^2 \geq \frac{2gd^2}{v^2}\left(h + \frac{gd^2}{2v^2}\right) \\ &\iff v^4 \geq 2ghv^2 + g^2d^2 \qquad \text{（除以 } d^2 > 0\text{，再乘以 } v^4\text{）} \\ &\iff d^2 \leq \frac{v^2}{g^2}\left(v^2 - 2gh\right). \end{aligned}$$

? **想一想** 假如槍就在樹腳，即 $d = 0$，以致上面的軌跡方程派不上用場，那又如何？

這時子彈必定是鉛垂向上射出的（$\theta = \frac{\pi}{2}$），而它能到達高度 h，當且僅當 $v^2 \geq 2gh$；這可由 $v^2 = u^2 - 2gs$ 及末速率不小於零得出。這正是 $d = 0$ 時的同一條件，所以該不等式涵蓋了所有情況。

由於 $d \geq 0$，取非負平方根便得

$$d \leq \frac{v}{g}\sqrt{v^2 - 2gh}.$$

所以槍位於水平地面上的某一點，而該點與樹腳的距離至多為 $(v/g)\sqrt{v^2-2gh}$：它位於以樹腳為圓心、以此為半徑的圓之內（極端情況下則在圓上），正如所求。平方根是實數，因為我們知道這一槍確實發生了：取 $d = 0$ 時，不等式迫使 $v^2 \geq 2gh$，這不過是說子彈有足夠的能量升到高度 h。

這個條件也是充分的。從圓內任何一點出發，該二次方程都有兩個正實根，所以有兩個可能的仰角（一個較平、一個較陡的射擊，如圖所示）；而在圓上，兩者合而為一。

驗算

- 當 $h = 0$ 時，半徑為 v^2/g，正是大家熟悉的水平地面上的最大射程。當 $v^2 \to 2gh$ 時，半徑縮小至 0：槍只有在樹腳筆直向上射擊，才僅僅能夠擊中狙擊手。
- 單位：$(v/g)\sqrt{v^2}$ 的單位是 $(\mathrm{m\,s^{-1}})^2/(\mathrm{m\,s^{-2}}) = \mathrm{m}$。
- 數值上，取 $g = 9.8$、$v = 20$ 及 $h = 8$，半徑約為 31.8 m。從 $d = 31.8\,\mathrm{m}$ 出發，以唯一容許的仰角 $\tan\theta = v^2/(gd)$ 逐步推算運動，子彈到達樹的位置時高度為 8.000 m。再對所有仰角逐一掃描，可確認圓外百分之幾的點無法擊中目標，而圓內的點則可以。

常見錯誤

- 只用最大射程 v^2/g 或只用最大高度。狙擊手位於高度 h，所以兩個坐標必須同時考慮。
- 忘記檢查 $\tan\theta$ 的根是否對應水平以上的仰角，或忽略了 $d = 0$ 的鉛垂射擊。

STEP 之外

對於固定的 v，由槍可以射到的點填滿了拋物線 $y = \dfrac{v^2}{2g} - \dfrac{gx^2}{2v^2}$ 以下的區域；這條拋物線是整族軌跡的包絡線（envelope），又稱「安全拋物線」（parabola of safety）。我們的判別式條件，正正就是 $h \leq$ 這條包絡線在 $x = d$ 處的高度。曲線族的包絡線，在變分法（calculus of variations）和一階微分方程的理論中還會再次出現。Kleppner 與 Kolenkow 合著的 *An Introduction to Mechanics*（力學入門教材），以同樣嚴謹的方式處理拋體問題，是很好的進階讀物。

詳細解答

STEP I 1988，第 11 題

力學 • 質心

> **STEP I 1988，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/88-s1-q11

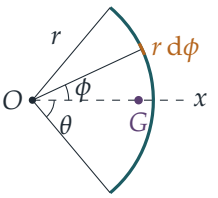
關鍵思路

對於圓弧，由對稱性可知質心（centre of mass）位於平分圓弧的半徑上，只需對角度作一次積分，便可求出質心與圓心的距離。至於框架，最值得注意的是它以半圓上的每一點着地都能平衡。地面的反作用力（reaction）總是沿半圓的半徑作用，而同時位於所有這些半徑上的點只有圓心。因此整個框架的質心必定是半圓的圓心，再對該點取力矩（moments），便得到一條關於 h 的方程。

解答

均勻圓弧

設圓弧的圓心為 O，每單位長度的密度為 ρ。從平分圓弧的半徑起量度角 ϕ，則圓弧為 $-\theta \leq \phi \leq \theta$。取 x 軸沿該平分線。由於圓弧關於此軸對稱，質心必在此軸上，故只需求 $\bar{x}$。



介乎 ϕ 與 $\phi + \mathrm{d}\phi$ 之間的線元長度為 $r\,\mathrm{d}\phi$，故質量為 $\rho r\,\mathrm{d}\phi$，而它的 x 坐標為 $r\cos\phi$。總

質量為 $\rho r \cdot 2\theta$。因此

$$\bar{x} = \frac{\displaystyle\int_{-\theta}^{\theta} r\cos\phi \; \rho r \, \mathrm{d}\phi}{2\rho r\theta} = \frac{r\big[\sin\phi\big]_{-\theta}^{\theta}}{2\theta} = \frac{r\sin\theta}{\theta}.$$

所以質心位於平分圓弧的半徑上，與圓心的距離為

$$\boxed{\bar{x} = \frac{r\sin\theta}{\theta}}$$

對於半圓弧，$\theta = \frac{\pi}{2}$，此距離為 $\dfrac{2r}{\pi}$，這正是下文要用到的情況。

框架：質心必須在哪裏？

我們把框架視為剛體（rigid body），由每單位長度密度同為 ρ 的均勻幼桿組成，並假設地面是水平的。設 O 為半圓的圓心，亦即長方形底邊的中點。

> **想一想** 為甚麼「以半圓上任何一點着地都處於平衡（equilibrium）」就能完全確定質心的位置？

假設框架以半圓上的點 P 着地而靜止。地面在 P 點與半圓相切，所以通過 P 的鉛垂線垂直於該切線：它就是半徑所在的直線 OP。作用於框架的力只有兩個，即作用於質心 G 的重量和作用於 P 的反作用力。要達至平衡，兩力必須沿同一直線作用，所以 G 位於通過 P 的鉛垂線上，亦即直線 OP 上。這對半圓上的每一點 P 都成立。兩個不同的點 P 給出兩條通過 O 的不同直線，而它們只相交於 O。因此 $G = O$。

反過來說，若 $G = O$，則對每一點 P，重量都沿 OP 作用，剛好穿過接觸點，所以框架無論以哪一點接觸地面都處於平衡。因此題目的條件恰好等同於「G 位於 O」。

對底邊取力矩

對底邊所在的直線取力矩，高度由 O 起向上量度（故半圓在該直線之下）。各部分如下：

部分	質量	其質心的高度	力矩
頂桿	$2r\rho$	$2h$	$4rh\rho$
兩條側桿	$2 \times 2h\rho$	h	$4h^2\rho$
底桿	$2r\rho$	0	0
半圓	$\pi r\rho$	$-\dfrac{2r}{\pi}$	$-2r^2\rho$

半圓的質心由圓弧公式取 $\theta = \frac{\pi}{2}$ 得出，而且它在 O 的下方。由於 G 的高度為 0，令總力矩等於零，再除以 2ρ：

$$2h^2 + 2rh - r^2 = 0.$$

這是一條關於 h 的二次方程，其根為

$$h = \frac{-2r \pm \sqrt{4r^2 + 8r^2}}{4} = \frac{(-1 \pm \sqrt{3})\, r}{2}.$$

由於 $h > 0$，我們取正根：

$$\boxed{h = \frac{(\sqrt{3} - 1)\, r}{2} \approx 0.366\, r}$$

留意底桿對它本身所在直線的力矩為零。無論框架是否包括底桿，答案都一樣，所以對圖形的理解稍有不同亦無妨。

驗算

- 當 $\theta \to 0$ 時，$\frac{r \sin\theta}{\theta} \to r$：極短的圓弧，其質心就在圓弧本身之上。對於整個圓，$\theta = \pi$，我們得到 0，即圓心，正如對稱性所要求。
- 取 $h = \frac{(\sqrt{3}-1)r}{2}$：$2h^2 = (2 - \sqrt{3})r^2$ 而 $2rh = (\sqrt{3} - 1)r^2$，兩者相加為 r^2，符合要求。
- 我們把每條桿分成 20 000 小段，以數值方法計算質心。取 $r = 1$ 及 $h \approx 0.36603$，質心與 O 的距離小於 10^{-9}。把 h 增加 10%，質心便上移約 0.03，框架就不能再平衡了。

常見錯誤

- 誤用 $\frac{4r}{3\pi}$，那是半圓形薄片（lamina）的結果；半圓形金屬線應使用 $\frac{2r}{\pi}$。
- 以角度而非長度作加權：線元的質量是 $\rho r\, \mathrm{d}\phi$，因子 r 之所以能約去，只因為它同樣出現在總質量之中。
- 沒有給出理由便斷言 G 位於 O。論證的關鍵在於反作用力總是沿半徑作用。

STEP 之外

圓弧的積分是把質心表示為沿曲線的線積分（line integral）$\int \mathbf{r}\, \rho\, \mathrm{d}s$ 的最簡單例子。大學一年級的課程會進一步利用二重積分和三重積分處理曲面和立體，而帕普斯（Pappus）定理則把這些質心與旋轉體的面積和體積聯繫起來。這個平衡論證也是搖擺穩定性背後的想法：物體擱在曲面上時，若其質心低於曲率中心（centre of curvature），它便是穩定的。Kleppner 與 Kolenkow 合著的 *An Introduction to Mechanics*（力學入門）正是以這種精神處理質心和剛體平衡。

詳細解答

STEP I 1988，第 12 題

力學 • 變力作用下的運動

> **STEP I 1988，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/88-s1-q12

關鍵思路

對每一種方法，求出溜冰者朝岸邊最遠能前進多少；只要 d 大於其中至少一種方法的這個距離，她就安全，反之則不然。轉彎是典型的圓周運動（circular motion）：一個大小不變、與速度垂直的力會令她沿固定半徑的圓運動，而這個半徑正是她仍會向前推進的距離。煞停則是變力（variable force）問題：力取決於速率，但我們要求的是距離，這正提示我們把加速度寫成 $v\,\frac{\mathrm{d}v}{\mathrm{d}x}$。

解答

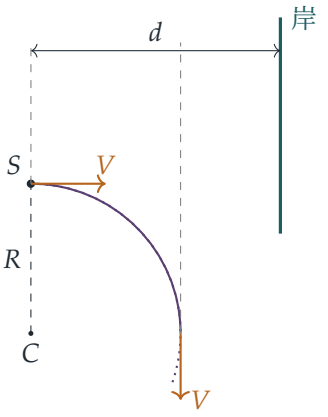
模型

我們把溜冰者視為質量為 M 的質點（particle），置於光滑（smooth）的水平冰面上。她在水平方向所受的唯一的力就是她自己施加的那個力，因此重量與法向反作用力（normal reaction）互相抵消，往後不必再考慮。我們由她的起點朝岸邊量度距離 x，所以岸邊位於 $x = d$。

方法一：沿圓弧轉彎

與速度垂直的力不做功，所以她的速率保持為 V。一個大小恆為 F 的力，若始終與速率恆為 V 的速度成直角，便提供了沿半徑為 R 的圓運動所需的向心力（centripetal force），其中

$$F = \frac{MV^2}{R}, \qquad \text{所以} \qquad R = \frac{MV^2}{F}.$$



她由 S 出發，正對着岸邊前進，所以圓心 C 位於通過 S 且與岸邊平行的直線上，與 S 相距 R。轉過角度 θ 之後，她朝岸邊前進的距離為 $x = R\sin\theta$，當 $\theta = \frac{\pi}{2}$ 時達到最大值 R；這時她正與岸邊平行地運動，其後便會彎離岸邊。所以她能避開岸邊的條件是

$$d > \frac{MV^2}{F}.$$

（若恰好 $d = MV^2/F$，她會剛好擦過岸邊，沿着岸邊運動。）

方法二：煞停

現在力沿着她的運動直線正向後作用，所以她沿直線朝岸邊運動，並逐漸減速。由於我們要求的是距離而不是時間，把加速度寫成 $v\frac{\mathrm{d}v}{\mathrm{d}x}$，由牛頓（Newton）第二定律得

$$Mv\frac{\mathrm{d}v}{\mathrm{d}x} = -\frac{F(V^2+v^2)}{2V^2}.$$

分離變數（separating the variables），並由 $x = 0$（該處 $v = V$）積分至她停下的位置 $x = X$：

$$\int_V^0 \frac{v}{V^2+v^2}\,\mathrm{d}v = -\frac{F}{2MV^2}\int_0^X \mathrm{d}x,$$
$$\left[\tfrac{1}{2}\ln(V^2+v^2)\right]_V^0 = -\frac{FX}{2MV^2},$$
$$\tfrac{1}{2}\ln\frac{V^2}{2V^2} = -\tfrac{1}{2}\ln 2 = -\frac{FX}{2MV^2},$$

所以她的停止距離（stopping distance）為

$$X = \frac{MV^2}{F}\ln 2.$$

想一想 她真的會停下來嗎？還是她的速率只會趨近於零，而她永遠在緩緩向前挪動？

煞停的力永遠不會低於 $\frac{1}{2}F$，所以她的速率會在有限時間內降至零。具體來說，把 $M\frac{\mathrm{d}v}{\mathrm{d}t} = -\frac{F}{2V^2}(V^2+v^2)$ 積分可得 $\frac{2MV}{F}\left(\arctan 1 - \arctan\frac{v}{V}\right) = t$，所以她在 $t = \pi MV/(2F)$ 時停下，其間走過有限的距離 X。（一旦靜止，她便不再施力。）因此當 $d > X$ 時，煞停可以救她；若 $d = X$，她到達岸邊時速率為零，只是碰到岸邊而沒有任何撞擊。

比較兩種方法

由於 $\ln 2 \approx 0.693 < 1$，停止距離比轉彎半徑短：煞停是較佳的方法，而且凡是轉彎行得通的情況，煞停也行得通。因此，她能夠避免撞上岸邊，當且僅當

$$\boxed{d > \frac{MV^2}{F}\ln 2}$$

（而 $d = \frac{MV^2}{F}\ln 2$ 是臨界情況，這時她到達岸邊時剛好靜止）。

質點假設

真實的溜冰者是有大小的：她的身軀和手臂會伸到質心（centre of mass）前方一段距離，所以實際上她所需的 d 要比上述數值大，多出的大約就是這段伸展的長度。更重要的是，這些力來自她的冰刀對冰面的作用。要轉彎，她必須向彎內傾側；要急煞，她必須把冰刀扭至橫向於運動方向（即冰球式急停，hockey stop），這會令她的身體轉向側面。這類轉動與姿勢，正是質點模型所忽略的。模型還暗中假設了冰面確實能夠提供這些力，儘管冰面被稱為光滑；即使沿冰刀方向沒有摩擦力，冰刀在橫向仍能抓緊冰面。

驗算

- 單位：MV^2/F 的單位是 $\mathrm{kg\,m^2\,s^{-2}/N} = \mathrm{m}$，是長度，正如所應得的。
- 上下界：煞停的力介乎 $\frac{1}{2}F$ 與 F 之間。恆力 F 會令她在距離 $MV^2/(2F)$ 內停下，而恆力 $\frac{1}{2}F$ 則為 MV^2/F；事實上確有 $\frac{1}{2} < \ln 2 < 1$。
- 數值驗證：隨機選取 M、V、F（例如 $R = MV^2/F = 7.0747$），以龍格－庫塔（Runge–Kutta）積分法逐步計算兩種運動，得出轉彎時最遠前進距離為 7.0747，煞停時停止距離為 $4.9038 = R\ln 2$，停止時間為 $\pi MV/(2F)$。

常見錯誤

- 把圓心放在運動直線上（或放在岸邊），因而得出前進距離為 $2R$ 或 0；其實圓心應在她的旁邊，前進距離才是 R。
- 處理煞停時使用 $\frac{\mathrm{d}v}{\mathrm{d}t}$，然後苦於求不出距離；用 $v\,\frac{\mathrm{d}v}{\mathrm{d}x}$ 只需積分一次便可得出。
- 給出兩個各自獨立的條件，卻沒有加以比較。題目問的是她能夠脫險的 d 值所成的集合，所以我們需要的是兩種方法中較佳的一種。

STEP 之外

寫出 $\ddot{x} = v\,\frac{\mathrm{d}v}{\mathrm{d}x}$ 是邁向能量方法的第一步：對於只取決於位置的力，它就是功能定理（work–energy theorem），而在大學力學中，它會發展成勢（potential）以及描述運動的相平面（phase plane）圖。溜冰者在轉向與煞停之間的抉擇，也是一個簡單的最優控制（optimal control）問題：怎樣運用有限的力來避開障礙物，才是最好的方法？Kleppner and Kolenkow, *An Introduction to Mechanics*（力學入門教科書）正好在高一個層次上討論取決於速度的力和圓周運動。

詳細解答

STEP I 1988，第 13 題

力學 • 彈性繩與碰撞

STEP I 1988，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/88-s1-q13

關鍵思路

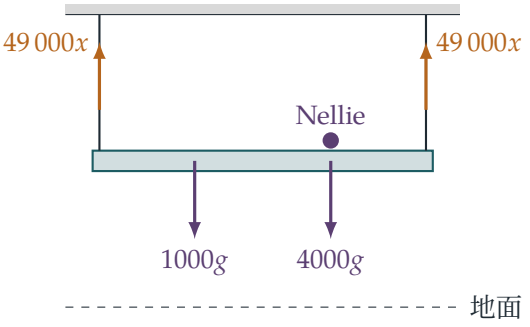
整個過程分為三個階段，每個階段各有自己的守恆定律（conservation law）。Nellie 落在木板上是一次突然發生的非彈性碰撞（inelastic collision），所以動量（momentum）守恆，但能量並不守恆。下降至地面的過程只涉及重力和彈性繩，所以能量守恆。Nellie 離開木板之後，能量再次守恆，而我們要問的是：儲存的彈性勢能是否足以把木板提升至超過繩子回復自然長度（natural length）的位置。

解答

我們把木板看成一個始終保持水平的剛體（rigid body），於是可以把它當作一個質量為 1000 kg、沿鉛垂方向運動的質點（particle）。Nellie 是一個質量為 4000 kg 的質點，而繩子是遵守虎克定律（Hooke's law）的輕彈性繩（light elastic string）。兩條繩子的伸長量（extension）x 始終相同，所以它們合共施加向上的力

$$2 \cdot \frac{490\,000}{10}\, x = 98\,000\, x \text{ N}.$$

因此這對繩子的作用就如一條勁度（stiffness）為 $K = 98\,000 \text{ N m}^{-1}$ 的彈簧，其彈性勢能（elastic energy）為 $\frac{1}{2}Kx^2 = 49\,000\, x^2$ J。



起始位置

Nellie 到來之前，木板懸掛着處於平衡（equilibrium）狀態，所以伸長量 e_0 滿足

$$98\,000\,e_0 = 1000 \times 9.8 \quad \Longrightarrow \quad e_0 = 0.1\text{ m}.$$

落在木板上

Nellie 落在木板上之後便留在上面，所以這是一次完全非彈性碰撞（perfectly inelastic collision）。撞擊歷時極短，而繩子的張力（tension）和各重量都是有限的力，所以它們在撞擊期間的衝量（impulse）可以忽略不計。因此動量守恆，而緊接撞擊之後的共同速率 v 滿足

$$4000 \times 5 = 5000\,v \quad \Longrightarrow \quad v = 4\text{ m s}^{-1}.$$

想一想 我們可以從 Nellie 到達的一刻起就直接使用能量守恆嗎？

不可以。一頭大象落在木板上並留在上面時，大量動能（kinetic energy）會轉化為熱和聲音：這裏 $\frac{1}{2}\cdot 4000\cdot 5^2 = 50\,000\,\text{J}$ 變成了 $\frac{1}{2}\cdot 5000\cdot 4^2 = 40\,000\,\text{J}$，即損失了 $10\,000\,\text{J}$。能量只是從緊接落下之後的一刻起才守恆。

下降過程

設 h 為木板最初離地面的高度。木板和 Nellie（總質量 5000 kg）以速率 $4\,\text{m s}^{-1}$ 出發，在低了距離 h 的地面處靜止下來，該處的伸長量為 $0.1 + h$。做功的力只有重力和繩子的拉力，所以

$$\underbrace{\tfrac{1}{2}\cdot 5000\cdot 4^2}_{\text{動能損失}} + \underbrace{5000\cdot 9.8\,h}_{\text{重力勢能損失}} = \underbrace{49\,000\left[(0.1+h)^2 - 0.1^2\right]}_{\text{彈性勢能增加}},$$

$$40\,000 + 49\,000\,h = 49\,000\,(0.2h + h^2).$$

兩邊除以 49 000 並整理，得

$$h^2 - 0.8\,h - \tfrac{40}{49} = 0, \qquad \text{所以} \qquad h = 0.4 + \sqrt{0.16 + \tfrac{40}{49}} = \frac{14 + 2\sqrt{299}}{35},$$

由於 $h > 0$，這裏取正根。因此

$$h \approx 1.39\ \text{m}.$$

在這個階段中繩子一直保持拉緊，因為木板下降時伸長量只會增加。

Nellie 離開之後

Nellie 輕輕地走下木板，所以只剩下木板（1000 kg）靜止在地面處，伸長量為

$$X = 0.1 + h \approx 1.488\ \text{m}.$$

如果木板上升得足夠高，使伸長量降至零，繩子便會鬆弛（slack），這表示木板要上升高度 X。繩子拉緊期間能量守恆，因此只要儲存的彈性勢能大於勢能的增加，木板便會到達該高度：

$$\tfrac{1}{2}KX^2 = 49\,000\,X^2 \approx 108\,500\ \text{J}, \qquad 1000 \cdot 9.8 \cdot X \approx 14\,600\ \text{J}.$$

前者遠大於後者，所以木板到達自然長度位置時仍在向上運動，動能約為 93 900 J（速率約為 $13.7\,\text{m}\,\text{s}^{-1}$）。在這一點之上，繩子要比自然長度更短才行，所以它們是鬆弛的。**會，繩子會變得鬆弛**。

> **想一想** 我們不可以直接用簡諧運動嗎？

可以，只要繩子仍然拉緊。木板獨自繞着其平衡伸長量 0.1 m 作簡諧運動（simple harmonic motion，簡稱 SHM），而它從平衡位置之下 $h \approx 1.39$ m 處出發，所以振幅（amplitude）是 1.39 m。振動的最高點會在伸長量為 $0.1 - 1.39 < 0$ 之處，而這對拉緊的繩子來說是不可能的。所以繩子必定在木板到達最高點之前鬆弛。（事實上，只要 $h > 0.1$，SHM 的論證便已給出答案。）

繩子一旦鬆弛，木板便只在重力作用下自由運動，再上升 $13.7^2/(2 \times 9.8) \approx 9.6$ m。這小於 10 m，所以木板（僅僅！）不會在回落之前碰到繩子的固定點。

驗算

- 把 $h = 1.3881$ 代入能量方程：兩邊都等於 108 017 J（準確至最接近的焦耳）。
- 由 $x = 0.1$、$\dot{x} = 4$ 出發，以數值方法逐步計算 $\ddot{x} = g - Kx/5000$，得出木板下降 1.3881 m 後停下。

- 逐步計算 Nellie 離開後木板的運動，顯示伸長量降至零時木板正以 $13.7\,\mathrm{m\,s^{-1}}$ 向上運動，而最高點在地面之上 $11.07\,\mathrm{m}$，與 $X + 9.58$ 吻合。
- 單位：$490\,000\,\mathrm{N} / 10\,\mathrm{m}$ 給出 $\mathrm{N\,m^{-1}}$，所以能量方程中每一項的單位都是焦耳。

常見錯誤

- 在落下的一刻使用能量守恆，這樣得出的共同速率是 $\sqrt{20} \approx 4.47\,\mathrm{m\,s^{-1}}$，而不是 $4\,\mathrm{m\,s^{-1}}$。
- 由自然長度起計算彈性勢能，卻忘記了初始伸長量 $0.1\,\mathrm{m}$：增加的能量是 $49\,000[(0.1+h)^2 - 0.1^2]$，而不是 $49\,000\,h^2$。
- 只用了一條繩子的勁度，而不是兩條。

STEP 之外

這對繩子是一條非線性彈簧（nonlinear spring）：拉長時它會拉，但壓縮時卻不能推。這類會在不同運動方程之間切換的系統，稱為分段線性（piecewise-linear）振子或「碰撞」振子（“impact” oscillator），即使是簡單的例子，其行為也可以出乎意料地複雜。Kleppner 與 Kolenkow 的 *An Introduction to Mechanics*（力學入門教科書）對能量、碰撞和振動有嚴謹的論述，而這些正是本題的全部基礎。

詳細解答

STEP I 1988，第 14 題

概率與統計 • 刪失正態分佈的期望值

> **STEP I 1988，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/88-s1-q14

關鍵思路

變量 X_M 其實就是 $\min(X, M)$：把正態變量（normal variable）中所有高於 M 的部分一律壓低到 M。它的期望值（expectation）可拆成兩部分：一個在 $x < M$ 上的積分，以及一個常數乘以尾部概率（tail probability）；而這個積分很容易計算，因為 $x\phi(x)$ 除了正負號之外，正是 $\phi(x)$ 的導數。遊客的故事只是同一個想法換了個包裝。一隊巴士所接載的乘客人數是 $\min(\text{需求}, \text{載客量})$，經正態近似（normal approximation）之後，這正好就是一個 X_M。

解答

X_M 的期望值

由於在 $\{X < M\}$ 上 $X_M = X$，而在 $\{X \geq M\}$ 上 $X_M = M$，

$$\mathrm{E}(X_M) = \int_{-\infty}^{M} x\,\phi(x)\,\mathrm{d}x \;+\; M\,\mathrm{P}(X \geq M) = \int_{-\infty}^{M} x\,\phi(x)\,\mathrm{d}x \;+\; M(1-\Phi(M)).$$

> **想一想** 有沒有一個函數，其導數是 $x\phi(x)$？

有。由 $\phi(x) = \frac{1}{\sqrt{2\pi}}\mathrm{e}^{-x^2/2}$，按鏈式法則（chain rule）可得 $\phi'(x) = -x\phi(x)$。因此

$$\int_{-\infty}^{M} x\,\phi(x)\,\mathrm{d}x = \Big[-\phi(x)\Big]_{-\infty}^{M} = -\phi(M) + 0,$$

因為當 $x \to -\infty$ 時 $\phi(x) \to 0$。故此

$$\mathrm{E}(X_M) = -\phi(M) + M(1-\Phi(M)),$$

這正是所求的結果。

作為合理性檢查，令 $M \to \infty$ 可得 $\mathrm{E}(X_M) \to 0 = \mathrm{E}(X)$，因為 $M(1-\Phi(M)) \to 0$。令 $M \to -\infty$ 則得 $\mathrm{E}(X_M) \approx M$，與預期相符。

巴士

設 N 為某一次到訪中被指往巴士站的遊客人數。我們假設 1024 名遊客各自獨立地被指路，往兩邊的概率各為 $\frac{1}{2}$。於是

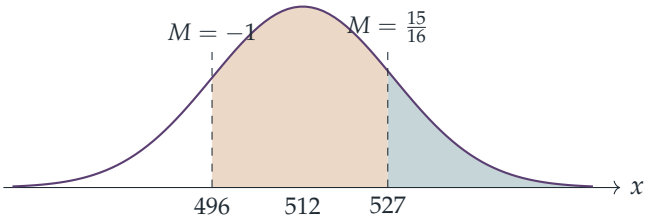
$$N \sim \mathrm{B}(1024, \tfrac{1}{2}), \qquad \mathrm{E}(N) = 512, \qquad \mathrm{Var}(N) = 1024 \cdot \tfrac{1}{2} \cdot \tfrac{1}{2} = 256 = 16^2.$$

由於 $n = 1024$ 這麼大，N 可以非常準確地用 $512 + 16X$ 來近似，其中 X 服從標準正態分佈（standard normal distribution）。

有 16 輛巴士時，載客量是 $16 \times 31 = 496$。有 17 輛時則是 527。載客量為 C 的車隊接載 $\min(N, C)$ 名乘客，所以它每次到訪的利潤（以 vloska 計）是

$$\begin{aligned}\min(N,C) &\approx \min(512 + 16X,\ C)\\ &= 512 + 16\min\Big(X, \frac{C-512}{16}\Big) = 512 + 16\,X_M, \qquad \text{其中} M = \frac{C-512}{16}.\end{aligned}$$

因此 $C = 496$ 對應 $M = -1$，而 $C = 527$ 對應 $M = \frac{15}{16}$。



當需求落在琥珀色的區域時，第 17 輛巴士接載的是超出 496 的那部分乘客。當需求落在藍綠色的尾部時，它載滿 31 人。

利用 $\phi(1) \approx 0.24197$ 和 $\Phi(1) \approx 0.84134$，並留意 $1 - \Phi(-1) = \Phi(1)$，

$$\mathrm{E}(X_{-1}) = -\phi(1) - \big(1 - \Phi(-1)\big) = -\phi(1) - \Phi(1) \approx -0.24197 - 0.84134 = -1.08332.$$

利用 $\phi(0.9375) \approx 0.25707$ 和 $1 - \Phi(0.9375) \approx 0.17425$，

$$\mathrm{E}(X_{15/16}) \approx -0.25707 + 0.9375 \times 0.17425 \approx -0.25707 + 0.16336 = -0.09371.$$

因此，多一輛巴士令每次到訪的期望利潤增加

$$16\big(\mathrm{E}(X_{15/16}) - \mathrm{E}(X_{-1})\big) \approx 16 \times 0.98961 \approx 15.83 \text{ vloska}.$$

每年有 50 次到訪，所以多一輛巴士的價值大約是

$$50 \times 15.83 \approx 792 \text{ vloska（每年）}.$$

這就是值得考慮的每年賄款上限。付出的金額一旦超過期望的額外利潤，合作社平均而言就會得不償失。

> **想一想** 既然 N 是離散的，我們不是應該使用連續性校正（continuity correction）嗎？

這裏並不需要。連續性校正的用處，在於近似 $\mathrm{P}(N \le c)$ 這類概率，因為這時被取平均的函數在整數處有跳躍。函數 $k \mapsto \min(k, C)$ 沒有跳躍，不作校正的正態近似對它已經極為準確。以二項分佈（binomial distribution）作精確計算，結果是 791.68，與上面的數字吻合至小數點後兩位。把載客量移動 $\frac{1}{2}$ 反而會令結果變差，得出 775。

驗算

- 以中點法則作數值積分（midpoint quadrature）計算 $\int_{-\infty}^{M} x\phi(x)\,\mathrm{d}x + M(1-\Phi(M))$，對多個 M 值所得的結果都與公式吻合至小數點後八位。以 4×10^5 個樣本對 $\min(X, M)$ 進行蒙地卡羅模擬（Monte Carlo simulation），結果亦吻合。
- 對 $\mathrm{B}(1024, \frac{1}{2})$ 作精確求和，得出每次到訪的額外期望利潤為 15.8337，即每年 791.68，與正態近似相符。
- 粗略的合理性檢查：只要 $N > 496$，多出的那輛巴士就有收入，這種情況出現的概率約為 0.84，而且它往往滿座。因此每次到訪的數字比 31 的一半略多，是合理的。

常見錯誤

- 忘記換算比例：$\min(N, C)$ 是 $512 + 16X_M$，所以 $\mathrm{E}(X_M)$ 之差必須先乘以標準差（standard deviation）16，再乘以 50。
- $M = -1$ 時的正負號錯誤：$M(1-\Phi(M))$ 一項是 $-(1-\Phi(-1)) = -\Phi(1)$，而不是 $-(1-\Phi(1))$。

STEP 之外

$\mathrm{E}\min(X, M)$ 這個量是一個刪失（censored）正態變量的期望值。與它密切相關的表達式，例如 $\mathrm{E}(X-M)^+ = \phi(M) - M(1-\Phi(M))$，是存貨理論（inventory theory，即「報販問題」（newsvendor problem））以及布萊克–舒爾斯（Black–Scholes）期權定價的骨

幹。恆等式 $\mathrm{E}\min(N, C+1) - \mathrm{E}\min(N, C) = \mathrm{P}(N > C)$ 提供了一個簡潔的方法，讓人看清每多一個單位載客量的價值。想進一步了解這些概念，可參閱 Grimmett 與 Welsh 的 *Probability: An Introduction*（概率論入門教材）。

詳細解答

STEP I 1988，第 15 題

概率與統計 • 二項分佈的假設檢驗

STEP I 1988，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/88-s1-q15

關鍵思路

把「超過一半的得分來自犯規」化為犯規次數與入球數之間的一個簡單不等式（inequality）。以入球數作為條件（conditioning）之後，餘下的工作只是數一數有多少個等可能的犯規次數符合要求。至於第二部分，要問的是：按照舊分佈，在 300 次試驗中預期會出現多少次九次犯規的情況？假如甚麼都沒有改變，觀察到的次數又有多出人意料？

解答

概率

就某場比賽中的某一支球隊而言，設 G 為該隊的入球數，F 為對方向該隊犯規的次數。那麼 G 與 F 獨立（independent），且

$$\mathrm{P}(G=0),\ \ldots,\ \mathrm{P}(G=4) \;=\; \tfrac{1}{9},\ \tfrac{2}{9},\ \tfrac{3}{9},\ \tfrac{2}{9},\ \tfrac{1}{9}, \qquad \mathrm{P}(F=f)=\tfrac{1}{10} \quad (f=0,1,\ldots,9).$$

該隊從入球得 $2G$ 分，從犯規得 F 分，合共 $2G+F$ 分。該隊超過一半的得分來自犯規，當且僅當

$$F > \tfrac{1}{2}(2G+F) \iff F > 2G.$$

（假如該隊一分也沒有得到，則 $F = 0 = 2G$，這種情況不應計算在內，而嚴格不等式正好把它排除了。）

給定 $G = g$，符合要求的 F 值為 $2g+1,\ldots,9$，即十個等可能的值當中的 $9-2g$ 個。因此，利用 F 與 G 的獨立性，

$$\mathrm{P}(F > 2G \mid G=g) = \frac{9-2g}{10}, \qquad g=0,1,2,3,4,$$

得出 $\frac{9}{10}, \frac{7}{10}, \frac{5}{10}, \frac{3}{10}, \frac{1}{10}$。根據全概率公式（law of total probability），

$$\mathrm{P}(F > 2G) = \frac{1\cdot 9 + 2\cdot 7 + 3\cdot 5 + 2\cdot 3 + 1\cdot 1}{9\times 10} = \frac{45}{90} = \frac{1}{2}.$$

答案如此簡潔，是有原因的：入球數的分佈關於 2 對稱，而 $(9-2g)/10$ 是 g 的線性函數，所以它的平均值等於它在 $g=2$ 的值，即 $\frac{1}{2}$。

犯規次數的分佈有沒有改變？

一支球隊的犯規次數，就是對手被犯規的次數，所以在舊規則下，每支球隊在一場比賽中犯規 9 次的概率為 $\frac{1}{10}$。設 X 為 300 個球隊場次當中，球隊犯規 9 次的場次數目。我們檢驗

$$H_0:\ p = \tfrac{1}{10}\ \text{（沒有改變）} \qquad \text{對} \qquad H_1:\ p < \tfrac{1}{10},$$

其中 p 是一支球隊在一場比賽中犯規 9 次的概率。備擇假設（alternative hypothesis）是單側（one-sided）的，因為修改規則的目的，是令犯規次數偏高的情況變得較不可能出現。

? **想一想** 這裏採用二項分佈（binomial）模型是否合理？

我們需要這 300 個球隊場次互相獨立，而且每個場次的概率 p 都相同。題目已說明同一場比賽之內的獨立性，而假設不同場次的比賽之間也互相獨立，亦屬合理。因此在 H_0 之下，$X \sim \mathrm{B}(300, \frac{1}{10})$，且

$$\mathrm{E}(X) = 30, \qquad \mathrm{Var}(X) = 300\times\tfrac{1}{10}\times\tfrac{9}{10} = 27, \qquad \text{標準差}\sqrt{27} \approx 5.20.$$

我們觀察到 $X = 12$，比期望值 30 少了 18，相當於約 3.5 個標準差（standard deviation）。採用正態近似（normal approximation）並加上連續性校正（continuity correction），

$$\mathrm{P}(X \le 12) \approx \Phi\left(\frac{12.5 - 30}{\sqrt{27}}\right) = \Phi(-3.37) \approx 0.0004.$$

精確的二項概率還要更小，約為 0.0001。無論採用哪一個數值，假如甚麼都沒有改變，這樣極端的結果在一千個球季之中也出現不到一次。它遠遠超出常用的 5% 甚至 1% 顯著性水平（significance level），所以我們拒絕 H_0。這*確實*是有力的證據，顯示犯規次數的分佈已經改變，而且是朝着所期望的方向改變：犯規九次的場次變得較為罕見。

有兩點保留值得說明。這個檢驗只顯示 P(9 次犯規) 下降了，並沒有說明分佈的其餘部分如何改變（儘管那是最自然的解釋）。此外，統計學也無法證明這個下降是由修改規則*造成*的，因為兩個球季之間可能還有其他事情改變了。

驗算

► 入球數的概率 $\frac{1}{9}(1+2+3+2+1)$ 之和為 1，理應如此。

- 以精確分數逐一列舉全部 5×10 個數對 (g, f)，得出 $\mathrm{P}(F > 2G) = \frac{1}{2}$；而模擬 200,000 場比賽的結果為 0.4997。
- 直接把二項概率加起來，對於 $X \sim \mathrm{B}(300, 0.1)$ 得出 $\mathrm{P}(X \leq 12) = 9.2 \times 10^{-5}$。正態近似（$3.8 \times 10^{-4}$）與平均值為 30 的泊松近似（Poisson approximation）（1.7×10^{-4}）都一致顯示這個概率極小。

常見錯誤

- 把 F 與 G 比較，而不是與 $2G$ 比較：一個入球值兩分。
- 計算 $F \geq 2G$ 而不是 $F > 2G$。「超過一半」是嚴格的不等關係，而一分未得的球隊，並沒有任何得分來自犯規。
- 以 150 場比賽作為樣本大小，而不是 300 個球隊場次；或者在計算之前忘記寫明零假設（null hypothesis）和二項分佈模型。

STEP 之外

第二部分是對二項比例所作的假設檢驗（hypothesis test），而我們求得的尾部概率就是 p 值（p-value）。到了大學，你會學到似然比（likelihood ratio）檢驗和卡方檢驗（chi-squared test），它們會把觀察所得的整個犯規次數分佈與舊分佈作比較，而不只是比較其中一個值。Blitzstein 與 Hwang 合著的 *Introduction to Probability*（概率論入門）把概率方面的內容（條件概率、二項分佈與正態近似）講解得十分清楚。

詳細解答

STEP I 1988，第 16 題

概率與統計 • 連續分佈與條件化

> **STEP I 1988，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/88-s1-q16

關鍵思路

第一部分是一道歸一化（normalisation）與積分的練習：題目只給出密度函數（density）到相差一個常數的程度，所以我們必須要求總概率為 1，從而求出這個常數；而 $t^2/(1+t^2)^2$ 的積分，只要把它拆成 $1/(1+t^2)$ 與 $t/(1+t^2)$ 的導數，便迎刃而解。其後的檢驗制度只是一棵簡短的分情況樹形圖。只有車齡 4 和 5 時的檢驗才有影響（車齡 3 之前沒有汽車會生鏽），所以我們按汽車在第四次檢驗時是否已經生鏽來分情況，而最後一部分就是在同一棵樹形圖上運用貝葉斯定理（Bayes' theorem）。

解答

開始生鏽時間的分佈

設汽車在車齡 $3+T$ 時開始生鏽，其中 $T \geq 0$ 的密度函數為 $f(t) = k\,t^2/(1+t^2)^2$，k 為某常數。關鍵的積分來自以下觀察：

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{t}{1+t^2}\right) = \frac{1-t^2}{(1+t^2)^2}, \qquad \text{故} \qquad \frac{t^2}{(1+t^2)^2} = \frac{1}{2}\left(\frac{1}{1+t^2} - \frac{1-t^2}{(1+t^2)^2}\right),$$

只要把括號內的式子通分，以 $(1+t^2)^2$ 為公分母，便可驗證：分子為 $(1+t^2)-(1-t^2) = 2t^2$。因此

$$\int_0^t \frac{s^2}{(1+s^2)^2}\,\mathrm{d}s = \frac{1}{2}\left(\arctan t - \frac{t}{1+t^2}\right).$$

當 $t \to \infty$ 時，$\arctan t \to \frac{\pi}{2}$ 而 $t/(1+t^2) \to 0$，所以在 $[0,\infty)$ 上的積分為 $\frac{\pi}{4}$。要使 f 成為密度函數，我們需要 $k \cdot \frac{\pi}{4} = 1$，故 $k = 4/\pi$。於是

$$F(t) = \mathrm{P}(\text{開始生鏽時車齡小於} 3+t) = \mathrm{P}(T < t) = \frac{2}{\pi}\left(\arctan t - \frac{t}{1+t^2}\right), \quad t \geq 0.$$

（留意 $F(t) \to 1$：每輛汽車終究都會生鏽。）

我們需要以下兩個數值：

$$F(1) = \frac{2}{\pi}\left(\frac{\pi}{4} - \frac{1}{2}\right) = \frac{1}{2} - \frac{1}{\pi} \approx 0.1817,$$
$$F(2) = \frac{2}{\pi}\left(\arctan 2 - \frac{2}{5}\right) \approx 0.4502.$$

第五周年檢驗不合格

車齡 3 之前沒有汽車生鏽，所以車齡 1、2 和 3 時的檢驗必定合格（恰好在第三次檢驗那一刻開始生鏽的概率為零）。汽車要在車齡 5 時不合格，必須在車齡 4 時合格，然後在車齡 5 時不合格，而不合格的前提是車齡 5 時已經生鏽。鏽蝕永遠不會消失，所以這件事有兩種互斥（disjoint）的發生方式。

- 車齡 4 時已經生鏽（概率為 $F(1)$）。它以概率 $\frac{1}{2}$ 通過第四次檢驗，然後在仍然生鏽的情況下，以概率 $\frac{1}{2}$ 在第五次檢驗不合格。貢獻：$\frac{1}{4}F(1)$。
- 在車齡 4 與 5 之間開始生鏽（概率為 $F(2) - F(1)$）。它必定通過第四次檢驗，並以概率 $\frac{1}{2}$ 在第五次檢驗不合格。貢獻：$\frac{1}{2}\big(F(2) - F(1)\big)$。

這裏我們按題目的原意假設：一輛生鏽汽車在各次檢驗的結果彼此獨立（independent），亦與開始生鏽的時間獨立。

> **想一想** 為甚麼我們不用乘上通過車齡 1、2 和 3 時各次檢驗的概率？

因為這些檢驗都在任何汽車有可能生鏽之前進行，所以每輛汽車都會通過；該因子為 1。

把兩種情況相加，

$$\begin{aligned}\mathrm{P}(\text{檢驗不合格時的車齡為5}) &= \tfrac{1}{4}F(1) + \tfrac{1}{2}\big(F(2) - F(1)\big) = \tfrac{1}{2}F(2) - \tfrac{1}{4}F(1)\\ &= \frac{1}{\pi}\left(\arctan 2 - \frac{2}{5}\right) - \frac{1}{8} + \frac{1}{4\pi}\\ &= \boxed{\frac{1}{\pi}\left(\arctan 2 - \frac{3}{20}\right) - \frac{1}{8} \approx 0.1797.}\end{aligned}$$

已知在第五次檢驗後被銷毀，求第四次檢驗時已生鏽的概率

一輛在第五次檢驗後隨即被銷毀的汽車，正正就是在該次檢驗不合格（而之前各次全部合格）的汽車，亦即剛才求得的事件。設 A 為汽車在第四次檢驗時已經生鏽的事件，D 為它在第五次檢驗不合格的事件。上面第一點所述的正是 $A \cap D$，所以

$$\mathrm{P}(A \cap D) = \tfrac{1}{4}F(1) = \frac{1}{8} - \frac{1}{4\pi} = \frac{\pi - 2}{8\pi}.$$

根據條件概率（conditional probability）的定義，

$$\mathrm{P}(A \mid D) = \frac{\mathrm{P}(A \cap D)}{\mathrm{P}(D)} = \frac{\dfrac{\pi - 2}{8\pi}}{\dfrac{1}{\pi}\left(\arctan 2 - \dfrac{3}{20}\right) - \dfrac{1}{8}}.$$

分子和分母同乘以 8π，

$$\boxed{\mathrm{P}(A \mid D) = \frac{\pi - 2}{8\arctan 2 - \frac{6}{5} - \pi} = \frac{5(\pi - 2)}{40\arctan 2 - 6 - 5\pi} \approx 0.253.}$$

所以在五年時被銷毀的汽車當中，只有約四分之一是在已經生鏽的情況下僥倖通過第四次檢驗的；大部分都是在第五年內才生鏽。

驗算

- $F(0) = 0$，而且當 $t \to \infty$ 時 $F(t) \to 1$，符合分佈函數（distribution function）應有的性質。對 $\frac{4}{\pi}\int_0^2 t^2/(1+t^2)^2\,\mathrm{d}t$ 使用辛普森法則（Simpson's rule）得 0.45018，與 $F(2)$ 吻合。
- 模擬 400 000 輛汽車（抽樣得出開始生鏽的時間，再進行每年的檢驗），得出五年時不合格的概率為 0.1789，條件概率為 0.255，而精確值分別為 0.1797 和 0.2528。
- 合理性檢查：$\mathrm{P}(D)$ 必須小於 $\frac{1}{2}F(2) \approx 0.225$（即車齡五年時已生鏽然後不合格的概率，當中不考慮較早被淘汰的情況），而事實的確如此。

常見錯誤

- 忘記歸一化：密度函數只是與 $t^2/(1+t^2)^2$ 成正比，漏掉因子 $4/\pi$ 會令其後每個答案都出錯。
- 把「車齡 5 時已生鏽」當作概率為 $F(2)$ 的單一情況，然後乘以 $\frac{1}{2}$。這樣做忽略了那些在車齡 4 時已生鏽、並在該次檢驗被淘汰的汽車。
- 在最後一部分，以「車齡 5 時已生鏽」而非「車齡 5 時不合格」作為條件。

STEP 之外

這裏的函數 F 是一個分佈函數，而比值 $f(t)/(1-F(t))$ 稱為風險率（hazard rate），即已知汽車至今仍未生鏽的條件下，瞬間生鏽的可能性；它是存活分析（survival analysis）與可靠性理論（reliability theory）的核心對象。最後一部分是濾波（filtering）的一個小

例子：從不完美的觀測（各次檢驗）推斷隱藏的狀態（是否生鏽）。Grimmett 與 Welsh 的 *Probability: An Introduction*（概率論入門）正是以這種精神講解連續分佈與條件概率。

詳細解答

STEP II 1988，第 1 題

純數學 • 部分分式與級數

> **STEP II 1988，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/88-s2-q1

關鍵思路

把 f 分解成部分分式（partial fractions），整個問題便歸結為把兩個簡單的分式 $\frac{1}{x-1}$ 和 $\frac{1}{x-2}$ 展開成等比級數（geometric series）。唯一真正要決定的，是每個分式應該朝哪個方向展開。等比級數 $\sum r^n$ 只在 $|r|<1$ 時收斂（converge），所以對每個分式，我們都把分母兩項之中絕對值較大的一項提出來。當 $|x|$ 先後越過 1 和 2 時，這個選擇便會改變，這就是為甚麼會有三個不同的展開式。

解答

部分分式

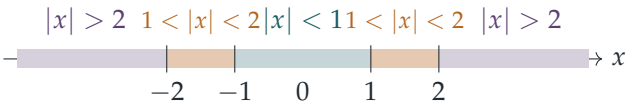
設 $\dfrac{1}{(x-1)(x-2)} = \dfrac{A}{x-1} + \dfrac{B}{x-2}$。由遮蓋法（cover-up rule）得 $A = \frac{1}{1-2} = -1$ 及 $B = \frac{1}{2-1} = 1$，所以

$$\mathrm{f}(x) = \frac{1}{x-2} - \frac{1}{x-1}.$$

我們需要的兩件基本工具都來自等比級數

$$\frac{1}{1-r} = \sum_{n=0}^{\infty} r^n \qquad (|r|<1),$$

其中 r 取 x、$\frac{x}{2}$、$\frac{1}{x}$ 或 $\frac{2}{x}$ 其中之一，視乎在所討論的範圍內哪一個的模（modulus）小於 1。



範圍 $|x|<1$

這裏 $|x|<1$，同時 $|\frac{x}{2}|<\frac{1}{2}<1$，所以兩個分式都可以展開成 x 的冪：

$$-\frac{1}{x-1}=\frac{1}{1-x}=\sum_{n=0}^{\infty}x^n,$$
$$\frac{1}{x-2}=-\frac{1}{2}\cdot\frac{1}{1-\frac{x}{2}}=-\frac{1}{2}\sum_{n=0}^{\infty}\left(\frac{x}{2}\right)^n.$$

兩式相加，得

$$\mathrm{f}(x)=\sum_{n=0}^{\infty}x^n-\frac{1}{2}\sum_{n=0}^{\infty}\left(\frac{x}{2}\right)^n\qquad(|x|<1),$$

正是所求。

範圍 $1<|x|<2$

第二個分式維持不變：我們仍然有 $|\frac{x}{2}|<1$，所以 $\frac{1}{x-2}=-\frac{1}{2}\sum_{n\geq0}(\frac{x}{2})^n$，與之前完全一樣。

> **想一想** 為甚麼第一個分式不能繼續使用 $\sum x^n$？

因為現在 $|x|>1$，各項 x^n 並不趨向零，所以 $\sum x^n$ 發散（diverge）。我們改為從分母提出較大的一項 x，餘下的公比是 $\frac{1}{x}$，而 $|\frac{1}{x}|<1$：

$$-\frac{1}{x-1}=-\frac{1}{x}\cdot\frac{1}{1-\frac{1}{x}}=-\frac{1}{x}\sum_{n=0}^{\infty}x^{-n}=-\sum_{n=0}^{\infty}x^{-(n+1)}=-\sum_{n=1}^{\infty}x^{-n},$$

其中最後一步只是把 $n+1$ 重新記作 n。把兩部分相加，得

$$\mathrm{f}(x)=-\sum_{n=1}^{\infty}x^{-n}-\frac{1}{2}\sum_{n=0}^{\infty}\left(\frac{x}{2}\right)^n\qquad(1<|x|<2),$$

正是所求。

範圍 $|x|>2$

第一個分式沿用上一個範圍的展開式，因為 $|\frac{1}{x}|<\frac{1}{2}<1$。但現在 $|\frac{x}{2}|>1$，所以第二個分式也要用同樣的方法倒轉過來展開：提出 x，餘下公比 $\frac{2}{x}$，而 $|\frac{2}{x}|<1$：

$$\frac{1}{x-2}=\frac{1}{x}\cdot\frac{1}{1-\frac{2}{x}}=\frac{1}{x}\sum_{n=0}^{\infty}\left(\frac{2}{x}\right)^n=\sum_{n=0}^{\infty}2^nx^{-(n+1)}=\sum_{n=1}^{\infty}2^{n-1}x^{-n}.$$

因此，當 $|x|>2$ 時，

$$\mathrm{f}(x)=\sum_{n=1}^{\infty}2^{n-1}x^{-n}-\sum_{n=1}^{\infty}x^{-n}=\sum_{n=1}^{\infty}(2^{n-1}-1)x^{-n}.$$

$n=1$ 的係數是 $2^0-1=0$，所以這個級數其實由 x^{-2} 開始：

$$\mathrm{f}(x)=\sum_{n=2}^{\infty}(2^{n-1}-1)x^{-n}=\frac{1}{x^2}+\frac{3}{x^3}+\frac{7}{x^4}+\frac{15}{x^5}+\cdots \qquad (|x|>2).$$

（兩個級數各自收斂，所以逐項合併是合理的。）無論寫成兩個級數的形式，還是合併後的形式，都是正確答案。

驗算

- 主導項。當 $|x|$ 很大時，$\mathrm{f}(x)=\frac{1}{x^2-3x+2}\approx\frac{1}{x^2}$，與最後一個展開式的首項吻合。$x^{-1}$ 的係數等於零並非巧合：兩個部分分式的 $\frac{1}{x}$ 項大小相等、符號相反。
- 代入一個值。當 $x=3$ 時，$\mathrm{f}(3)=\frac{1}{2\cdot1}=\frac{1}{2}$，而 $\sum_{n\ge2}(2^{n-1}-1)3^{-n}=\frac{1}{2}\sum_{n\ge2}(\frac{2}{3})^n-\sum_{n\ge2}(\frac{1}{3})^n=\frac{1}{2}\cdot\frac{4}{3}-\frac{1}{6}=\frac{1}{2}$。同樣地，$\mathrm{f}(\frac{1}{2})=\frac{4}{3}$ 和 $\mathrm{f}(\frac{3}{2})=-4$ 亦與頭兩個展開式吻合。
- 數值上，在各自範圍內隨機選取數百個點（正負皆有），三個級數（各取 4000 項求和）與 f 的值都相符，誤差約為 10^{-14}。

常見錯誤

- 在 $\frac{1}{x-2}=-\frac{1}{2}\cdot\frac{1}{1-x/2}$ 中弄錯正負號：提出來的因子是 -2，不是 2。
- 乘以 $\frac{1}{x}$ 之後，求和指標差了一：$\frac{1}{x}\sum_{n\ge0}x^{-n}$ 由 x^{-1} 開始，所以重新標記後的和由 $n=1$ 開始。
- 沒有檢查公比的模是否小於 1，就在每個範圍都使用同一個展開式。

STEP 之外

這三個展開式，是複分析（complex analysis）中洛朗級數（Laurent series）在實變數上的投影。對於複數 z，函數 $\frac{1}{(z-1)(z-2)}$ 在 $|z|<1$、$1<|z|<2$ 和 $|z|>2$ 這三個區域內各有不同的洛朗展開式：這三個區域是一個圓盤和兩個圓環（annulus），以通過極點（pole）的圓為邊界，而 $|x|=1$ 和 $|x|=2$ 在複平面上正好變成這些圓。這類展開式中 z^{-1} 的係數稱為留數（residue），原來它們決定了圍道積分（contour integral）的值。Needham 的 *Visual Complex Analysis*（《複分析：可視化方法》）以優美的幾何方式講解了冪級數（power series）和洛朗級數。

詳細解答

STEP II 1988，第 2 題

純數學 • 對稱代數方程

STEP II 1988，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/88-s2-q2

關鍵思路

先去分母，再把兩條方程相減。右邊的平方差（difference of two squares）會產生因式 $y-z$，正好與左邊出現的 $y-z$ 對應，而假設 $y \neq z$ 恰恰讓我們可以把它約去。此後，一切都歸結為一條關於 x、y 和 z 的對稱式（symmetric expression），餘下的工作就交給對稱性完成。

解答

證明 $x+y+z=a$

由於 $y \neq 0$ 及 $z \neq 0$，我們可以把兩邊乘開：

$$2ay - 3y^2 = (z-x)^2, \tag{1}$$
$$2az - 3z^2 = (x-y)^2. \tag{2}$$

由 (1) 減去 (2)，並在右邊運用平方差，

$$\begin{aligned} 2a(y-z) - 3(y^2-z^2) &= (z-x)^2 - (x-y)^2 \\ &= \big[(z-x)-(x-y)\big]\big[(z-x)+(x-y)\big] \\ &= (y+z-2x)(z-y). \end{aligned}$$

左邊是 $(y-z)\big[2a-3(y+z)\big]$，所以

$$(y-z)\big[2a-3(y+z)\big] = -(y-z)(y+z-2x).$$

想一想 究竟在哪一步需要用到 $y \neq z$？

就在這裏：因為 $y - z \neq 0$，我們可以把它除去，得到

$$2a - 3y - 3z = 2x - y - z \quad \Longrightarrow \quad 2a = 2x + 2y + 2z,$$

所以 $x + y + z = a$，正如所求。

證明第三條方程

把 $a = x + y + z$ 代回 (1)：

$$y(2x + 2y + 2z - 3y) = (z - x)^2 \implies 2xy - y^2 + 2yz = z^2 - 2zx + x^2,$$

整理後得

$$x^2 + y^2 + z^2 - 2xy - 2yz - 2zx = 0. \tag{3}$$

最引人注目的是，(3) 對 x、y 和 z 是對稱的。現在看看我們要證明甚麼。由於 $x \neq 0$，方程 $2a - 3x = \dfrac{(y-z)^2}{x}$ 等價於

$$x(2a - 3x) = (y - z)^2,$$

而代入 $a = x + y + z$ 後，這就是

$$x(2y + 2z - x) = y^2 - 2yz + z^2 \iff x^2 + y^2 + z^2 - 2xy - 2yz - 2zx = 0,$$

這正是 (3)。因此

$$2a - 3x = \frac{(y-z)^2}{x}.$$

(事實上，三條方程都只是 (3) 連同 $a = x + y + z$，各自從其中一個變量的角度寫出來而已。)

當 $y = z$ 時，最後一條方程可以成立嗎？

可以。若 $y = z$，題目給出的兩條方程變成同一條方程 $2a - 3y = \dfrac{(y-x)^2}{y}$，而最後一條方程則變成 $2a - 3x = 0$，即 $a = \frac{3}{2}x$。把 $a = \frac{3}{2}x$ 代入第一條方程，

$$3x - 3y = \frac{(x-y)^2}{y} \iff 3y(x - y) = (x - y)^2 \iff (x - y)(x - 4y) = 0.$$

所以，舉例說，$x = y = z = 1$ 及 $a = \frac{3}{2}$ 滿足全部三條方程：

$$2a - 3y = 0 = \frac{(z-x)^2}{y}, \qquad 2a - 3z = 0 = \frac{(x-y)^2}{z}, \qquad 2a - 3x = 0 = \frac{(y-z)^2}{x}.$$

另一個例子是 $x=4$、$y=z=1$、$a=6$：這時頭兩條方程都有 $2a-3y=9=\frac{(1-4)^2}{1}$，而 $2a-3x=0=\frac{0^2}{4}$。

所以，最後一條方程並非僅當 $y\neq z$ 時才成立：條件 $y\neq z$ 保證它成立，但當 $y=z$ 時它也可以成立。當 $y=z$ 時失去的是這個保證。例如 $x=1$、$y=z=2$、$a=\frac{13}{4}$ 滿足題目給出的兩條方程（兩邊各等於 $\frac{1}{2}$），但 $2a-3x=\frac{7}{2}\neq 0$。還要留意，第一個例子中 $x+y+z=3\neq a$，所以當 $y=z$ 時，連第一個結論也可能不成立。

驗算

- 取 $x=9$、$y=4$、$z=1$，它們滿足 (3)：$81+16+1-72-8-18=0$。這時 $a=14$，而 $2a-3y=16=\frac{(1-9)^2}{4}$、$2a-3z=25=\frac{(9-4)^2}{1}$ 及 $2a-3x=1=\frac{(4-1)^2}{9}$。三條方程全部成立。
- 一次數值檢驗在 10,000 組隨機有理數三元組上精確地驗證了恆等式 $(1)-(2)=(y-z)\big(2a-2(x+y+z)\big)$，並在 (3) 上約 1,900 個滿足 $y\neq z$ 的隨機點確認了全部三條方程，另加上文三個 $y=z$ 的例子。

常見錯誤

- 除以 $y-z$ 而不說明為何可以這樣做。題目之所以給出這個假設，全是為了這一步。
- 最後一部分只答「不可以」或「可以」而沒有舉例。只要明確給出一組 $y=z$ 的三元組，問題便解決了。
- 以為 $y=z$ 會令最後一條方程不成立。它只是令保證消失而已。

STEP 之外

方程 (3) 是一個二次型（quadratic form），它的零點集是 (x,y,z) 空間中的一個雙錐面（double cone）。若 $x,y,z>0$，設 $x=p^2$、$y=q^2$、$z=r^2$：則 (3) 可因式分解為

$$-(p+q+r)(-p+q+r)(p-q+r)(p+q-r)=0,$$

這是邊長為 p,q,r 的三角形面積平方的 16 倍（希羅公式，Heron's formula）。所以 (3) 的意思是 $\sqrt{x},\sqrt{y},\sqrt{z}$ 構成一個退化三角形（degenerate triangle），例如上面驗算中的 $\sqrt{9}=\sqrt{4}+\sqrt{1}$。二次型及其分類是大學一年級線性代數（linear algebra）的標準課題；參見 Strang，*Introduction to Linear Algebra*（線性代數入門教科書）。

詳細解答

STEP II 1988，第 3 題

純數學 • 根的對稱性

> **STEP II 1988，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/88-s2-q3

關鍵思路

不要先從係數入手，而應先考慮根的集合。題目的性質是說，根的集合在運算 $k \mapsto 1/k$ 之下保持不變，所以兩個根要麼互換，要麼各自不動。加入第二個運算 $k \mapsto 1-k$ 後，同樣的推理可以再用一次。至於三次方程，把這兩個運算反覆施行，由一個根最多可得出六個值，而三次方程只容得下三個根：因此各根必須落在特殊的點上，使六個值之中有些重合。

解答

倒數性質

設兩根為 α 和 β（可能相等，也可能是複數），則 $\alpha+\beta=-b$ 及 $\alpha\beta=c$。首先，0 不可能是根，因為 0^{-1} 不存在；所以 $c\neq 0$。

映射（map）$k \mapsto 1/k$ 必須把每個根送到一個根，所以 $1/\alpha \in \{\alpha, \beta\}$。

- 若 $1/\alpha=\beta$，則 $\alpha\beta=1$，即 $c=1$。反過來，若 $c=1$，則兩根皆非零，且 $\beta=1/\alpha$，所以每個根的倒數（reciprocal）就是另一個根，而這性質對任何實數 b 都成立。（這包括 $b=\mp 2$ 時的重根（repeated root）± 1，以及 $|b|<2$ 時位於單位圓（unit circle）上的複數根。）
- 若 $1/\alpha=\alpha$ 但 $1/\alpha\neq\beta$，則 $\alpha=\pm 1$ 且 $\alpha\beta\neq 1$。把這性質應用於 β：它的倒數不是 α（因為 $\alpha\beta\neq 1$），所以 $1/\beta=\beta$，於是 $\beta=\pm 1$ 亦然。由於 $\alpha\beta\neq 1$，兩根必定是 1 和 -1，得出 $x^2-1=0$：$b=0$，$c=-1$。這方程確實具有該性質。

所以該性質成立，當且僅當

$$c = 1 \ (b \text{ 為任何實數}), \qquad \text{或} \qquad b = 0,\ c = -1.$$

加入性質 $k \mapsto 1-k$

以 $1-k$ 取代 $1/k$，重複同樣的論證：要麼 $1-\alpha = \beta$，於是 $\alpha+\beta = 1$；要麼兩根都不動，即 $1-\alpha = \alpha$ 及 $1-\beta = \beta$，於是 $\alpha = \beta = \frac{1}{2}$，同樣有 $\alpha+\beta = 1$。無論哪種情況都有 $b = -1$；反過來，$b = -1$ 使兩根成為 α 和 $1-\alpha$，所以該性質成立。

$b = 0$、$c = -1$ 這情況不符合此條件，所以我們需要 $c = 1$ 及 $b = -1$：

$$b = -1, \qquad c = 1,$$

得出 $x^2 - x + 1 = 0$，其根為 $\frac{1}{2} \pm \frac{\sqrt{3}}{2}\mathrm{i}$。每個根的模（modulus）都是 1，實部都是 $\frac{1}{2}$，所以 $1/k$ 和 $1-k$ 都等於共軛複數（complex conjugate）$\bar{k}$，即另一個根。

三次方程

現設 $x^3 + px^2 + qx + r = 0$ 的係數為實數，並同時具有兩個性質。一如之前，0 不是根，1 也不是（否則 $1-1=0$ 便會是根）。

實係數三次方程至少有一個實根，稱之為 a。反覆施行這兩個運算，以下每一個都是根：

$$a, \quad \frac{1}{a}, \quad 1-a, \quad \frac{1}{1-a}, \quad 1-\frac{1}{a} = \frac{a-1}{a}, \quad \frac{1}{(a-1)/a} = \frac{a}{a-1}.$$

（第四個是第三個的倒數，第五個是 1 − 第二個，第六個是第五個的倒數。）這是六個實數，但三次方程最多只有三個相異的根，所以其中必有兩個重合。

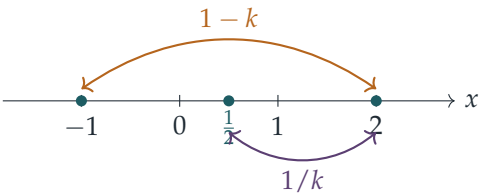
> **想一想** 「其中兩個相等」這句話，怎樣確定 a 的值？

把這六個映射稱為 $f_1, \ldots, f_6$，其中 f_1 為恆等映射（identity）。每一個映射的逆映射（inverse）都在這列表之中：$\frac{1}{k}$、$1-k$ 和 $\frac{k}{k-1}$ 各自是自己的逆映射，而 $\frac{1}{1-k}$ 和 $\frac{k-1}{k}$ 則互為逆映射。還可以驗證，把任意兩個映射複合（compose），所得的仍是列表中的映射。所以若 $f_i(a) = f_j(a)$ 而 $i \neq j$，施以 f_i^{-1} 便得 $a = g(a)$，其中 $g = f_i^{-1} \circ f_j$ 是六個映射之一，且不是恆等映射。所以 a 是五個非恆等映射其中之一的不動點（fixed point）：

$$\frac{1}{a} = a \iff a = \pm 1, \qquad \frac{1}{1-a} = a \iff a^2 - a + 1 = 0,$$
$$1-a = a \iff a = \tfrac{1}{2}, \qquad \frac{a-1}{a} = a \iff a^2 - a + 1 = 0,$$
$$\frac{a}{a-1} = a \iff a(a-2) = 0.$$

二次式 $a^2 - a + 1$ 沒有實根（它就是先前的那個二次式），而 $a \neq 0, 1$。因此 $a \in \{-1, \frac{1}{2}, 2\}$。

無論 a 是其中哪一個，上面六個值恰好取遍 -1、2 和 $\frac{1}{2}$：映射 $1-k$ 把 -1 和 2 互換，並保持 $\frac{1}{2}$ 不動；而 $1/k$ 則把 2 和 $\frac{1}{2}$ 互換，並保持 -1 不動。



所以 -1、2 和 $\frac{1}{2}$ 是該三次方程的三個相異根。首一（monic）三次多項式由它的三個根完全決定，所以它必定是

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

反過來，這個三次方程確實具有兩個性質，因為它的根集 $\{-1, \frac{1}{2}, 2\}$ 在 $k \mapsto 1/k$ 和 $k \mapsto 1-k$ 之下都被送回自身。所以

$$x^3 - \tfrac{3}{2}x^2 - \tfrac{3}{2}x + 1 = 0$$

是唯一符合條件的三次方程，正如題目所求。

驗算

代入 $x=2$：$8-6-3+1=0$；$x=-1$：$-1-\frac{3}{2}+\frac{3}{2}+1=0$。各根之積為 $-1\cdot 2\cdot\frac{1}{2} = -1 = -r$，各根之和為 $\frac{3}{2} = -p$。留意這三次方程的係數，除正負號外，順讀和倒讀都一樣（$1, -\frac{3}{2}, -\frac{3}{2}, 1$）：這種迴文（palindromic）形狀正是 $k \mapsto 1/k$ 所必然導致的。以電腦測試 20,000 組隨機的 (b, c)，結果印證了倒數條件；而隨機選取遠離 $0, \pm 1, \frac{1}{2}, 2$ 的實數 a，每次都得出六個相異的值。

常見錯誤

- 第一部分只答「$c=1$」：每個根都是自身倒數的那種情況，還會多出 $b=0$、$c=-1$。
- 忽略了 $\frac{1}{2} \pm \frac{\sqrt{3}}{2}\mathrm{i}$ 這一對根，它在兩個映射之下都是封閉的。正是由實根出發，才能把它排除：實係數三次方程必有一個實根，而該實根的六個值不可能包含這兩個數。
- 只證明了題目給出的三次方程符合條件；題目要求的是證明它是唯一的一個。

STEP 之外

六個映射 $k \mapsto k$, $1/k$, $1-k$, $\frac{1}{1-k}$, $\frac{k-1}{k}$, $\frac{k}{k-1}$ 在複合運算下構成一個羣（group），它是對稱羣（symmetric group）S_3 的一個複本；它置換 $0, 1, \infty$ 這三個點，而那六個值就是四個點的交比（cross-ratio）所有可能的取值。點算的那一步正是軌道－穩定子定理（orbit–stabiliser theorem）：只有在被某個非恆等元素固定的點上，軌道（orbit）的長度才會少於六。Allenby 的 *Rings, Fields and Groups*（環、域與羣）闡述羣在集合上的作用，而 Needham 的 *Visual Complex Analysis*（複分析的視覺化講解）則把這些映射展示為平面上的莫比烏斯變換（Möbius transformation）。

詳細解答

STEP II 1988，第 4 題

純數學 • 阿根圖上的軌跡

> **STEP II 1988，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/88-s2-q4

關鍵思路

把複數寫成 $a+\mathrm{i}b$，再把「實數」或「虛數」這個條件化為一條關於 a 和 b 的實方程。這樣便會在阿根圖（Argand diagram）上得到一條熟悉的曲線。接着，把平方看成平面上的一個映射（map），將 x 和 y（或 p 和 q）用 a 和 b 表示，然後消去參數，便可找出該曲線被映射到哪裏。

關於題目的印刷文字，要說明一點：原文寫的是 w^2-2x，但 x 在該處尚未引入，而最後部分（出現 t^2-2t）顯示了題目原意的條件。我們把它理解為w^2-2w 是實數。

解答

第 (i) 部分

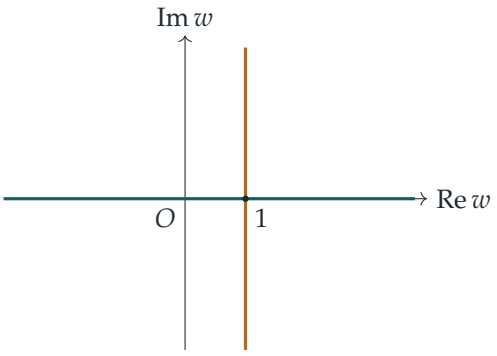
設 $w=a+\mathrm{i}b$，其中 a,b 為實數。則

$$w^2-2w=(a^2-b^2-2a)+\mathrm{i}\,(2ab-2b)=(a^2-b^2-2a)+2\mathrm{i}\,b(a-1).$$

這個數是實數，當且僅當其虛部（imaginary part）為零，即

$$b(a-1)=0,\qquad \text{即}\qquad b=0\ \text{或}\ a=1.$$

因此 w 的軌跡（locus）是實軸（real axis）連同垂直線 $\mathrm{Re}\,w=1$。兩條直線相交於 $w=1$。



第 (ii) 部分

我們把兩條直線分開來逐一平方。

實軸。若 $b = 0$，則 $w = a$ 是實數，而 $w^2 = a^2$。所以 $y = 0$ 及 $x = a^2 \geq 0$，而當 a 取遍所有實數時，a^2 會取遍 $[0, \infty)$ 內的每一個值。軌跡的這一部分是非負的 x 軸：即半直線（half-line）$y = 0$，$x \geq 0$。

直線 $a = 1$。若 $w = 1 + \mathrm{i}b$，則

$$w^2 = 1 - b^2 + 2\mathrm{i}b, \qquad \text{所以} \qquad x = 1 - b^2, \quad y = 2b.$$

利用 $b = \frac{1}{2}y$ 消去參數，得

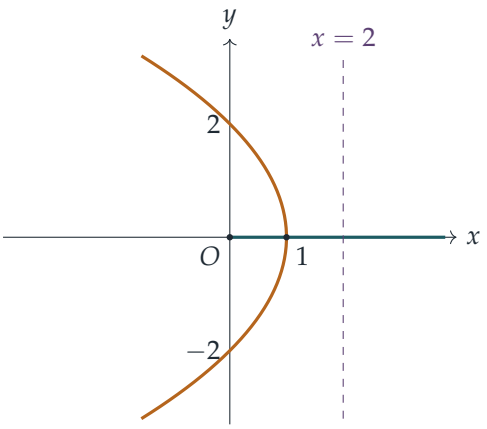
$$x = 1 - \tfrac{1}{4}y^2, \qquad \text{或} \qquad y^2 = 4(1 - x) = -4(x - 1).$$

只要取 $b = \frac{1}{2}y$，這條曲線上的每一點都能取到。它是一條拋物線（parabola），其軸沿 x 軸，頂點（vertex）在 $(1, 0)$，開口朝向 x 的負方向。與標準式 $Y^2 = 4cX$ 比較，這裏 $c = -1$，而 $X = x - 1$。所以焦點（focus）在 $X = -1$，即原點；準線（directrix）是 $X = 1$，即直線 $x = 2$。事實上 $x^2 + y^2 = x^2 + 4 - 4x = (2 - x)^2$，所以曲線上每一點與 O 的距離都等於它與直線 $x = 2$ 的距離。拋物線與 y 軸相交於 $(0, \pm 2)$。

? **想一想** 為甚麼這條半直線不會延伸到 x 的負值？

w 平面上的實軸被覆蓋了兩次（分別由 a 和 $-a$），但實數的平方永遠不會是負數。所以整條實軸對摺到非負的那一半上。同樣地，直線 $\mathrm{Re}\, w = 1$ 只覆蓋拋物線一次，因為 $b \mapsto 2b$ 是一對一（one-to-one）的。

所以 (x, y) 的軌跡是拋物線 $y^2 = 4(1 - x)$（焦點在原點，準線為 $x = 2$），連同半直線 $y = 0$，$x \geq 0$。這條半直線由焦點出發，穿過頂點 $(1, 0)$；該頂點是 $w = 1$ 的像，亦即第 (i) 部分兩條直線的交點的像。



最後部分

設 $t = a + \mathrm{i}b$。這裏的「虛數」是指純虛數（purely imaginary），所以我們需要 $t^2 - 2t$ 的實部（real part）為零：

$$a^2 - b^2 - 2a = 0, \qquad \text{即} \qquad (a-1)^2 - b^2 = 1.$$

所以 t 位於一條以 1 為中心的等軸雙曲線（rectangular hyperbola）上，其頂點在 $t = 0$ 和 $t = 2$。在這條曲線上，$b^2 = a^2 - 2a = a(a-2) \geq 0$，所以 $a \leq 0$ 或 $a \geq 2$。

現在由 $t^2 = p + \mathrm{i}q$ 得

$$p = a^2 - b^2 = 2a, \qquad q = 2ab,$$

其中我們用了雙曲線的方程 $a^2 - b^2 = 2a$ 來化簡 p。所以 $a = \frac{1}{2}p$，於是

$$q^2 = 4a^2b^2 = 4a^2(a^2 - 2a) = p^2\left(\tfrac{1}{4}p^2 - p\right), \qquad \text{所以} \qquad 4q^2 = p^3(p-4).$$

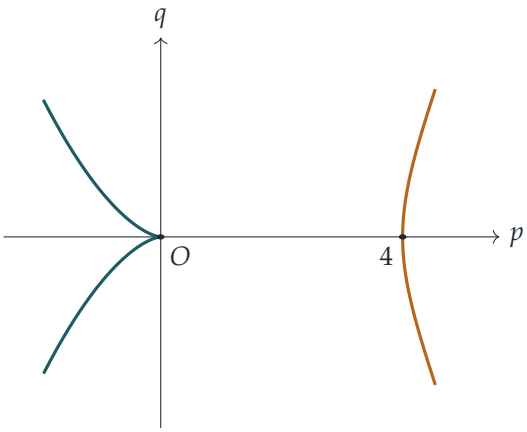
條件 $a \leq 0$ 或 $a \geq 2$ 變成 $p \leq 0$ 或 $p \geq 4$，這正好就是 $p^3(p-4) \geq 0$ 成立的範圍。反過來，在這條曲線上任取一點。若 $p \neq 0$，設 $a = \frac{1}{2}p$ 及 $b = q/p$。則 $b^2 = q^2/p^2 = \frac{1}{4}p(p-4) = a^2 - 2a$，所以 t 在雙曲線上，而且 $t^2 = (a^2 - b^2) + 2\mathrm{i}ab = p + \mathrm{i}q$。若 $p = 0$，則 $q = 0$，而這一點來自 $t = 0$。所以軌跡就是整條曲線

$$4q^2 = p^3(p-4).$$

它的形狀：

- 它對 p 軸對稱，因為方程中只出現 q^2。
- 在 $0 < p < 4$ 的範圍內沒有任何點。曲線分為兩段，分別來自雙曲線的兩支。
- 右邊一段由 $(4, 0)$ 開始，該點是 $t = 2$ 的像。在該點附近 $q^2 \approx 16(p-4)$，所以切線（tangent）是垂直的。

- 左邊一段通過原點，即 $t = 0$ 的像。在該點附近 $q^2 \approx -p^3$，所以 $|q| \approx |p|^{3/2}$，形成一個指向右方的尖點（cusp），兩臂都與 p 軸相切。
- 當 $|p|$ 很大時，$|q| \approx \frac{1}{2}p^2$，所以兩段都像拋物線般向外張開。



（圖中縱向比例經過壓縮。藍綠色的一段來自雙曲線上 $\operatorname{Re} t \le 0$ 的一支，琥珀色的一段來自 $\operatorname{Re} t \ge 2$ 的一支。）

驗算

- 第 (i) 部分：$w = 1 + 2\mathrm{i}$ 給出 $w^2 - 2w = (-3 + 4\mathrm{i}) - (2 + 4\mathrm{i}) = -5$，這是實數。第 (ii) 部分：它的平方 $-3 + 4\mathrm{i}$ 滿足 $16 = 4(1 - (-3))$。
- 最後部分：$t = -1 + \mathrm{i}\sqrt{3}$ 滿足 $1 - 3 + 2 = 0$。它的平方是 $-2 - 2\sqrt{3}\,\mathrm{i}$，而 $4q^2 = 48 = (-8)(-6) = p^3(p - 4)$。
- 我們以數千個隨機點作了數值檢驗，確認每個軌跡都映射到所述的曲線上，而且每條曲線上的每一點都能取到。

常見錯誤

- 在第 (i) 部分忘記了實軸，以致在第 (ii) 部分漏掉那條半直線。又或者畫了整條 x 軸，而不是只畫 $x \ge 0$ 的部分。
- 沒有利用 $a^2 - b^2 = 2a$ 來得出 $p = 2a$。不用這一步的話，消去參數的過程會繁複得多。
- 在最後的草圖中漏掉了 $0 < p < 4$ 的空隙，或原點處的尖點。

STEP 之外

平方映射 $z \mapsto z^2$ 是最簡單的保角映射（conformal map）。它把角度加倍，這就是為甚麼這裏的直線和雙曲線會變成拋物線和帶尖點的曲線。在 $z = 0$ 處導數為零，角度不再保持不變，尖點正是由此而來。在流體流動和靜電學中，有一個標準技巧，就是利用這類映射把難處理的區域變成容易處理的區域。Needham 的 *Visual Complex Analysis*（《視覺複分析》）以圖像把這些映射解釋得十分精彩。

詳細解答

STEP II 1988，第 5 題

純數學 • 單位根與正切

STEP II 1988，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/88-s2-q5

關鍵思路

設 $z = \cos\theta + \mathrm{i}\sin\theta$，並用二項式定理 (binomial theorem) 展開 z^7。其虛部 (imaginary part) 是 $\sin 7\theta$，把它除以 $\cos^7\theta$ 後，就變成 $t = \tan\theta$ 的多項式。七次單位根（roots of unity）令 $\sin 7\theta$ 等於零，於是我們毫不費力就得到一個 7 次多項式的七個根。之後，正切之積不過是各根之積，可以直接從常數項讀出。

解答

六次方程的根

由棣美弗定理(de Moivre's theorem)，$(\cos\theta + \mathrm{i}\sin\theta)^7 = \cos 7\theta + \mathrm{i}\sin 7\theta$。設 $c = \cos\theta$，$s = \sin\theta$。用二項式定理展開左方，並只保留含 $\mathrm{i}s$ 奇數次冪的項，得

$$\sin 7\theta = 7c^6 s - 35c^4 s^3 + 21c^2 s^5 - s^7.$$

只要 $\cos\theta \neq 0$，我們就可以除以 c^7，並設 $t = \tan\theta$：

$$\frac{\sin 7\theta}{\cos^7\theta} = 7t - 35t^3 + 21t^5 - t^7 = -t\,(t^6 - 21t^4 + 35t^2 - 7). \qquad (*)$$

$z^7 = 1$ 的根是 $z = \cos\theta + \mathrm{i}\sin\theta$，其中 $\theta = \frac{2k\pi}{7}$，而對這些角有 $\sin 7\theta = \sin 2k\pi = 0$，即 $z^7 = 1$ 的虛部。改用 $\theta = \frac{k\pi}{7}$，$k = 0, 1, \ldots, 6$，會較為簡潔，這同樣給出 $\sin 7\theta = \sin k\pi = 0$（而且得到的正切值集合相同，因為 tan 的週期是 π）。

對上述每一個 θ：

- $\cos\theta \neq 0$，因為 $\frac{k\pi}{7} = \frac{\pi}{2}$ 需要 $2k = 7$；所以 $(*)$ 成立；

► $\sin 7\theta = 0$，所以由 $(*)$ 可知，$t = \tan\frac{k\pi}{7}$ 這個數是 $p(t) = t^7 - 21t^5 + 35t^3 - 7t$ 的根。

$\tan\frac{k\pi}{7}$（$k = 0, \ldots, 6$）這七個值互不相同：各角 $0, \frac{\pi}{7}, \ldots, \frac{6\pi}{7}$ 是 $[0, \pi)$ 中互不相同的點，沒有一個等於 $\frac{\pi}{2}$，而 tan 在 $[0, \pi) \setminus \{\frac{\pi}{2}\}$ 上是一對一（one-to-one）的（在 $[0, \frac{\pi}{2})$ 上為正且遞增，在 $(\frac{\pi}{2}, \pi)$ 上為負，同樣遞增，兩段的取值範圍分別是 $[0, \infty)$ 和 $(-\infty, 0)$）。

? **想一想** 我們怎樣知道已經找到所有的根，而不只是其中一部分？

一個 7 次多項式最多有 7 個根。我們已找到 7 個互不相同的根，所以這就是全部的根。$k = 0$ 給出 $t = 0$，即來自因式 t 的那個根。因此 $t^6 - 21t^4 + 35t^2 - 7 = 0$ 的根恰好是

$$t = \tan\frac{k\pi}{7}, \qquad k = 1, 2, 3, 4, 5, 6,$$

也就是 $\pm\tan\frac{\pi}{7}$、$\pm\tan\frac{2\pi}{7}$、$\pm\tan\frac{3\pi}{7}$，因為 $\tan\frac{(7-k)\pi}{7} = -\tan\frac{k\pi}{7}$。

乘積 $\tan\frac{2\pi}{7}\tan\frac{4\pi}{7}\tan\frac{6\pi}{7}$

對於偶數次的首一多項式（monic polynomial），各根之積等於常數項。所以

$$\left(\tan\tfrac{\pi}{7}\right)\left(-\tan\tfrac{\pi}{7}\right)\left(\tan\tfrac{2\pi}{7}\right)\left(-\tan\tfrac{2\pi}{7}\right)\left(\tan\tfrac{3\pi}{7}\right)\left(-\tan\tfrac{3\pi}{7}\right) = -7,$$

亦即

$$\left(\tan\tfrac{\pi}{7}\tan\tfrac{2\pi}{7}\tan\tfrac{3\pi}{7}\right)^2 = 7.$$

三個角都在 $(0, \frac{\pi}{2})$ 內，所以每個正切值都是正數，而乘積就是正平方根：$\tan\frac{\pi}{7}\tan\frac{2\pi}{7}\tan\frac{3\pi}{7} = \sqrt{7}$。

現在利用 $\tan(\pi - x) = -\tan x$：$\tan\frac{4\pi}{7} = -\tan\frac{3\pi}{7}$ 及 $\tan\frac{6\pi}{7} = -\tan\frac{\pi}{7}$。兩個負號互相抵消，所以

$$\tan\frac{2\pi}{7}\tan\frac{4\pi}{7}\tan\frac{6\pi}{7} = \tan\frac{2\pi}{7}\tan\frac{3\pi}{7}\tan\frac{\pi}{7} = \sqrt{7},$$

正如所求。

$n = 9$ 和 $n = 11$ 的情況

同樣的論證適用於任何奇數 $n = 2m + 1$。取 $(\cos\theta + \mathrm{i}\sin\theta)^n$ 的虛部，再除以 $\cos^n\theta$，得

$$\frac{\sin n\theta}{\cos^n\theta} = \sum_{r=0}^{m}(-1)^r\binom{n}{2r+1}t^{2r+1} = t\,q_n(t),$$

其中 q_n 的次數是 $2m = n - 1$，首項係數（leading coefficient）是 $(-1)^m$，常數項是 $\binom{n}{1} = n$。具體寫出來，

$$q_9(t) = t^8 - 36t^6 + 126t^4 - 84t^2 + 9,$$
$$q_{11}(t) = -t^{10} + 55t^8 - 330t^6 + 462t^4 - 165t^2 + 11.$$

與之前完全一樣（n 是奇數，所以沒有任何 $\frac{k\pi}{n}$ 等於 $\frac{\pi}{2}$），q_n 的根就是 $n-1$ 個互不相同的數 $\tan\frac{k\pi}{n}$，$k=1,\dots,n-1$，它們可配對成 $\pm\tan\frac{k\pi}{n}$，其中 $k=1,\dots,m$。由於次數是偶數，q_n 各根之積是（常數項）/（首項係數）：

$$(-1)^m\prod_{k=1}^{m}\tan^2\frac{k\pi}{n}=\frac{n}{(-1)^m}\quad\Longrightarrow\quad\prod_{k=1}^{m}\tan\frac{k\pi}{n}=\sqrt{n},$$

這裏取正根，因為這些角全都在 $(0,\frac{\pi}{2})$ 內。當 $n=9$ 時，這就是 q_9 的常數項 9 除以首項係數 1；當 $n=11$ 時，則是 $11/(-1)=-11=(-1)^5\prod\tan^2$。無論哪種情況，$\tan\frac{k\pi}{n}$ 對 $k=1,\dots,m$ 的乘積，在 $n=9$ 時是 3，在 $n=11$ 時是 $\sqrt{11}$。

想一想 題目所求的乘積，其角是 $\frac{2\pi}{n},\frac{4\pi}{n},\dots$，而不是 $\frac{\pi}{n},\frac{2\pi}{n},\dots$。我們怎樣由一個轉到另一個？正負號又會怎樣變化？

每個偶數倍 $2j\le m$ 本來就在 $1,\dots,m$ 這列數之中。每個偶數 $2j>m$（即 $2j>\frac{n}{2}$）給出 $\tan\frac{2j\pi}{n}=-\tan\frac{(n-2j)\pi}{n}$，而 $n-2j$ 是奇數，且不超過 m。由 1 到 m 的每個數都恰好這樣出現一次，因為偶數來自第一類，奇數來自第二類。所以所求的乘積是 $\pm\sqrt{n}$，而介乎 $\frac{n}{2}$ 與 n 之間的每個偶數各帶來一個負號。

- $n=9$：偶數有 $2,4,6,8$；其中大於 4.5 的是 $6,8$，共兩個負號。所以
$$\tan\frac{2\pi}{9}\tan\frac{4\pi}{9}\tan\frac{6\pi}{9}\tan\frac{8\pi}{9}=\sqrt{9}=3.$$
- $n=11$：偶數有 $2,4,6,8,10$；其中大於 5.5 的是 $6,8,10$，共三個負號。所以
$$\tan\frac{2\pi}{11}\tan\frac{4\pi}{11}\tan\frac{6\pi}{11}\tan\frac{8\pi}{11}\tan\frac{10\pi}{11}=-\sqrt{11}.$$

（當 $n=7$ 時，大於 3.5 的偶數是 $4,6$，共兩個負號，所以該乘積是正數。）

驗算

- $n=9$ 有一個簡單的局部驗算：$\tan\frac{6\pi}{9}=\tan\frac{2\pi}{3}=-\sqrt{3}$，而數值上 $\tan\frac{2\pi}{9}\tan\frac{4\pi}{9}\tan\frac{8\pi}{9}\approx 0.8391\times 5.6713\times(-0.3640)\approx -1.732=-\sqrt{3}$，所以總乘積是 3。
- 用計算機計算，$\tan\frac{2\pi}{7}\tan\frac{4\pi}{7}\tan\frac{6\pi}{7}\approx 2.645751$，而 $\sqrt{7}\approx 2.645751$；$n=11$ 的乘積是 $\approx -3.316625=-\sqrt{11}$。
- 代入 $\tan\frac{\pi}{7}\approx 0.4816$，得 $t^6-21t^4+35t^2-7\approx 0$，與預期相符。

常見錯誤

- 除以 $\cos^7\theta$ 時沒有檢查它不等於零，或者沒有論證七個正切值互不相同。缺少這幾點，就未能證明已經找到所有的根。

- 把 $n = 11$ 的正負號弄錯：乘積的平方很容易求得，但正負號要靠數清楚有多少個負的正切值來決定。
- 忘記 $a_d t^d + \cdots + a_0$ 各根之積是 $(-1)^d a_0/a_d$；這在 $n = 11$ 時尤其重要，因為其首項係數是 -1。

STEP 之外

$\tan \frac{k\pi}{n}$ 這些數都是代數數（algebraic number），而多項式 q_n（當 n 是質數時）就是它們在有理數上的最小多項式（minimal polynomial）。像 $\prod \tan \frac{k\pi}{n} = \sqrt{n}$ 這類恆等式，其實是高斯和（Gauss sums）的影子；高斯和說明 $\sqrt{\pm p}$ 位於由 p 次單位根生成的域（field）之中，這是代數數論（algebraic number theory）的一塊基石。Davenport 的 *The Higher Arithmetic* 是進入這一系列想法的平易入門讀物，而 Needham 的 *Visual Complex Analysis* 則讓人從幾何角度感受單位根。

詳細解答

STEP II 1988，第 6 題

純數學 • 雙曲函數不等式

> **STEP II 1988，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/88-s2-q6

關鍵思路

三個函數在 $x = 0$ 時的值都是零，所以只需證明它們在 $x > 0$ 時遞增（increasing）便可。題目中各導數（derivative）是刻意安排的，使每一部分都為下一部分鋪路：(ii) 的導數包含 (i) 的函數，而 (iii) 的導數則是 (ii) 經重新縮放後的版本。至於 f，對數微分法（logarithmic differentiation）把 f′ 的正負號化為 (iii) 中函數的正負號，這顯然正是設置該部分的用意。

解答

以下我們一直使用這個原理：若 $F(0) = 0$ 且對所有 $x > 0$ 都有 $F'(x) > 0$，則對所有 $x > 0$ 都有 $F(x) > 0$。（根據中值定理（mean value theorem），存在某個 $\xi \in (0, x)$ 使得 $F(x) = F(x) - F(0) = xF'(\xi)$。）

第 (i) 部分

設 $g(x) = x - \tanh x$。則 $g(0) = 0$，且

$$g'(x) = 1 - \operatorname{sech}^2 x = \tanh^2 x,$$

這裏用了 $1 - \tanh^2 x = \operatorname{sech}^2 x$。當 $x > 0$ 時，我們有 $\tanh x > 0$，所以 $g'(x) > 0$。因此當 $x > 0$ 時，$x - \tanh x > 0$。

第 (ii) 部分

設 $h(x) = x\sinh x - 2\cosh x + 2$。則 $h(0) = 0 - 2 + 2 = 0$，且

$$h'(x) = \sinh x + x\cosh x - 2\sinh x = x\cosh x - \sinh x = \cosh x\,(x - \tanh x).$$

當 $x > 0$ 時，$\cosh x > 0$，而且由第 (i) 部分知 $x - \tanh x > 0$，所以 $h'(x) > 0$。因此當 $x > 0$ 時，$h(x) > 0$。

第 (iii) 部分

設 $k(x) = 2x\cosh 2x - 3\sinh 2x + 4x$。則 $k(0) = 0$，且

$$\begin{aligned} k'(x) &= 2\cosh 2x + 4x\sinh 2x - 6\cosh 2x + 4 \\ &= 4x\sinh 2x - 4\cosh 2x + 4. \end{aligned}$$

? **想一想** 這個組合我們在哪裏見過？

令 $u = 2x$，上式便是 $2u\sinh u - 4\cosh u + 4 = 2\,(u\sinh u - 2\cosh u + 2) = 2h(u)$。所以

$$k'(x) = 2h(2x) > 0 \quad \text{對於} x > 0,$$

這是由第 (ii) 部分得出的，因為 $2x > 0$。因此當 $x > 0$ 時，$k(x) > 0$。

f 沒有轉向點

當 $x > 0$ 時，$\mathrm{f}(x)$ 中每個因子都是正數，所以 $\mathrm{f}(x) > 0$，我們可以取對數：

$$\ln \mathrm{f}(x) = \ln x + \tfrac{1}{3}\ln\cosh x - \ln\sinh x.$$

求導，得

$$\frac{\mathrm{f}'(x)}{\mathrm{f}(x)} = \frac{1}{x} + \frac{\sinh x}{3\cosh x} - \frac{\cosh x}{\sinh x} = \frac{3\sinh x\cosh x + x\sinh^2 x - 3x\cosh^2 x}{3x\sinh x\cosh x}.$$

現在把分子用倍角表示，利用 $\sinh x\cosh x = \frac{1}{2}\sinh 2x$、$\sinh^2 x = \frac{1}{2}(\cosh 2x - 1)$ 和 $\cosh^2 x = \frac{1}{2}(\cosh 2x + 1)$：

$$\begin{aligned} 3\sinh x\cosh x + x\sinh^2 x - 3x\cosh^2 x &= \tfrac{3}{2}\sinh 2x + \tfrac{x}{2}(\cosh 2x - 1) - \tfrac{3x}{2}(\cosh 2x + 1) \\ &= \tfrac{3}{2}\sinh 2x - x\cosh 2x - 2x \\ &= -\tfrac{1}{2}\big(2x\cosh 2x - 3\sinh 2x + 4x\big). \end{aligned}$$

於是

$$\mathrm{f}'(x) = -\frac{\mathrm{f}(x)\,\big(2x\cosh 2x - 3\sinh 2x + 4x\big)}{6x\sinh x\cosh x}.$$

當 $x > 0$ 時，由第 (iii) 部分知括號內的式子為正，而 $\mathrm{f}(x)$、x、$\sinh x$ 和 $\cosh x$ 全部為正。所以對每個 $x > 0$ 都有 $\mathrm{f}'(x) < 0$：f' 永不為零，故 f 沒有轉向點（turning point），事實上 f 在 $x > 0$ 上嚴格遞減。

草圖

我們需要知道曲線在兩端的表現。

當 $x \to 0^+$ 時：$x/\sinh x \to 1$ 且 $(\cosh x)^{1/3} \to 1$，所以 $\mathrm{f}(x) \to 1$。更仔細地說，$(\cosh x)^{1/3} = 1 + \frac{1}{6}x^2 + O(x^4)$ 而 $x/\sinh x = 1 - \frac{1}{6}x^2 + O(x^4)$，所以 x^2 項互相抵消；每個級數再多取一項，便得

$$\mathrm{f}(x) = 1 - \tfrac{1}{45}x^4 + O(x^6).$$

所以曲線離開（不包括在內的）點 $(0,1)$ 時非常平坦，斜率為零。

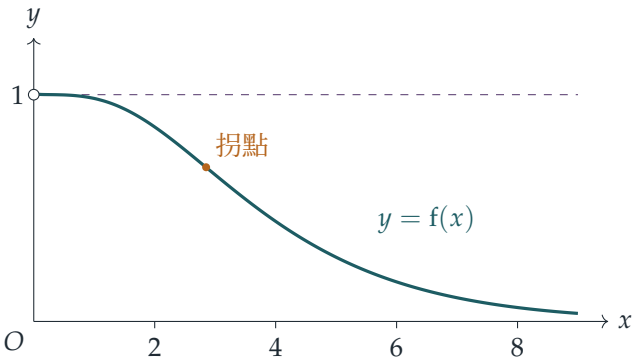
當 $x \to \infty$ 時：$\cosh x \sim \frac{1}{2}e^x$ 且 $\sinh x \sim \frac{1}{2}e^x$，所以

$$\mathrm{f}(x) \sim \frac{x\,(\frac{1}{2}e^x)^{1/3}}{\frac{1}{2}e^x} = 2^{2/3}\,x\,e^{-2x/3} \to 0,$$

因為指數函數的增長勝過 x 的冪。x-軸是一條漸近線（asymptote），而由於 $\mathrm{f} > 0$，曲線始終在它的上方。

> **想一想** 曲線開始時平坦，結束時也平坦，卻一直遞減。這必然導致甚麼？

在 0 附近，曲線向下彎（它是 $1 - x^4/45$）；而當 x 很大時，它是一個衰減的指數函數，向上彎。所以兩者之間必定至少有一個拐點（point of inflection）；數值計算顯示恰有一個，在 $x \approx 2.85$ 附近，該處 $\mathrm{f} \approx 0.68$。一些函數值：$\mathrm{f}(1) \approx 0.983$，$\mathrm{f}(2) \approx 0.858$，$\mathrm{f}(4) \approx 0.441$，$\mathrm{f}(6) \approx 0.174$。



驗算

- 在 $x = 1$ 處：$1 - \tanh 1 \approx 0.238$、$\sinh 1 - 2\cosh 1 + 2 \approx 0.089$ 和 $2\cosh 2 - 3\sinh 2 + 4 \approx 0.643$，全部為正。
- 恆等式 $\mathrm{f}'/\mathrm{f} = -k(x)/(6x\sinh x\cosh x)$ 可在任何一點檢驗：在 $x = 1$ 處兩邊都約為 -0.0592，而 f 的數值導數亦與此吻合。
- 數值上，當 x 很小時 $(1 - \mathrm{f}(x))/x^4 \approx 0.0222 \approx \frac{1}{45}$，而當 x 很大時 $\mathrm{f}(x)/(2^{2/3}xe^{-2x/3}) \to 1$。

常見錯誤

- 只證明導數為正，卻忘記檢查函數在 $x = 0$ 時為 0；兩者缺一不可。
- 試圖直接用商法則和積法則對 f 求導。這樣做是可行的，但對數導數簡潔得多，而且能讓人看清它與 (iii) 的聯繫。
- 把曲線畫成穿過 x-軸，或從原點出發。曲線從高度 1 開始（但取不到這個值），並且一直保持為正。

STEP 之外

由 $F(0) = 0$ 和 $F' > 0$ 證明 $F(x) > 0$ 的技巧，正是中值定理的實際應用；而把它連環使用（如本題般依次經過 k'、h'、g'），是分析學中證明函數之間不等式的標準方法。展開式 $\mathrm{f}(x) = 1 - x^4/45 + \cdots$ 讓我們初嘗運算冪級數（power series）的滋味，大學一年級的分析課程會為此奠定嚴格的基礎。要閱讀這兩方面的內容，Abbott 的 *Understanding Analysis* 是個好選擇。

詳細解答

STEP II 1988，第 7 題

純數學 • 代換積分法

STEP II 1988，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/88-s2-q7

關鍵思路

這裏每一個等式都來自一次代換（substitution），難處在於找出合適的代換。第一步需要一個把 $[1,2]$ 送到 $[0,1]$、並把 $2-2x+x^2$ 變成含有 $1+t^2$ 的式子的代換：莫比烏斯映射（Möbius map）$x=2/(1+t)$ 恰好做到這一點。一旦出現 $1+t^2$，取 $t=\tan\theta$ 便是順理成章的一步。以 θ 表示的被積函數（integrand）關於 $\theta=\frac{1}{8}\pi$ 對稱，而在半個區間上把 $t=\tan\theta$ 還原，便會帶出 $\tan\frac{1}{8}\pi=\sqrt{2}-1$。最後的推論是把同樣的代換反過來使用。

解答

全篇之中，k 是任意實常數。所有要取冪的底數在所涉及的範圍內都是正數，所以非整數次冪不會造成任何問題。

從 $[1,2]$ 到 $[0,1]$

設 $x=\dfrac{2}{1+t}$，於是當 $x=1$ 時 $t=1$，當 $x=2$ 時 $t=0$，並且 $\mathrm{d}x=-\dfrac{2}{(1+t)^2}\,\mathrm{d}t$。那麼

$$2-2x+x^2=\frac{2(1+t)^2-4(1+t)+4}{(1+t)^2}=\frac{2(1+t^2)}{(1+t)^2},\qquad x^{k+1}=\frac{2^{k+1}}{(1+t)^{k+1}}.$$

所以被積函數變成

$$\frac{2^k(1+t^2)^k}{(1+t)^{2k}}\cdot\frac{(1+t)^{k+1}}{2^{k+1}}\cdot\frac{2}{(1+t)^2}=\frac{(1+t^2)^k}{(1+t)^{k+1}},$$

而 $\mathrm{d}x$ 中的負號把積分上下限由 $\int_1^0$ 對調為 $\int_0^1$：

$$I = \int_0^1 \frac{(1+t^2)^k}{(1+t)^{k+1}}\,\mathrm{d}t.$$

三角形式

現在設 $t = \tan\theta$，其中 θ 由 0 變到 $\frac{1}{4}\pi$。那麼 $1+t^2 = \sec^2\theta$，$\mathrm{d}t = \sec^2\theta\,\mathrm{d}\theta$，並且

$$1 + t = \frac{\cos\theta + \sin\theta}{\cos\theta} = \frac{\sqrt{2}\cos\left(\frac{1}{4}\pi - \theta\right)}{\cos\theta},$$

這裏用到 $\cos\theta + \sin\theta = \sqrt{2}\left(\cos\frac{1}{4}\pi\cos\theta + \sin\frac{1}{4}\pi\sin\theta\right)$。被積函數變成

$$\begin{aligned}\sec^{2k}\theta \cdot \frac{\cos^{k+1}\theta}{\left[\sqrt{2}\cos\left(\frac{1}{4}\pi - \theta\right)\right]^{k+1}} \cdot \sec^2\theta &= \frac{\cos^{-(k+1)}\theta}{\left[\sqrt{2}\cos\left(\frac{1}{4}\pi - \theta\right)\right]^{k+1}} \\ &= \frac{1}{\left[\sqrt{2}\cos\theta\cos\left(\frac{1}{4}\pi - \theta\right)\right]^{k+1}},\end{aligned}$$

因為 $\cos\theta$ 的冪次合計為 $-2k + (k+1) - 2 = -(k+1)$。兩個餘弦在 $[0, \frac{1}{4}\pi]$ 上都是正數，所以這對任何實數 k 都成立。這就得出 I 的第二個形式。

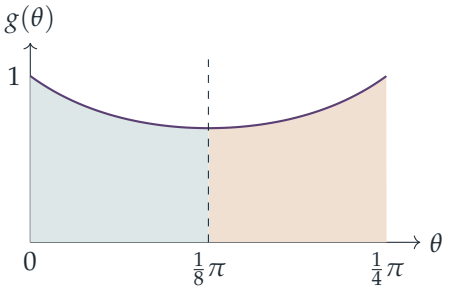
想一想 為甚麼在 $[0, \frac{1}{4}\pi]$ 上的積分應該是在 $[0, \frac{1}{8}\pi]$ 上的積分的兩倍？

把以 θ 表示的被積函數記作 $g(\theta)$。把 θ 換成 $\frac{1}{4}\pi - \theta$，只不過是把 $\cos\theta$ 和 $\cos(\frac{1}{4}\pi - \theta)$ 這兩個因子對調，所以 $g(\frac{1}{4}\pi - \theta) = g(\theta)$：$g$ 的圖像關於 $\theta = \frac{1}{8}\pi$ 對稱。（你也可以直接看出這一點：$\sqrt{2}\cos\theta\cos(\frac{1}{4}\pi - \theta) = \frac{1}{2}\left(1 + \sqrt{2}\cos(2\theta - \frac{1}{4}\pi)\right)$，它只取決於 θ 與 $\frac{1}{8}\pi$ 的距離。）嚴格來說，代入 $\phi = \frac{1}{4}\pi - \theta$，

$$\int_{\pi/8}^{\pi/4} g(\theta)\,\mathrm{d}\theta = \int_0^{\pi/8} g(\tfrac{1}{4}\pi - \phi)\,\mathrm{d}\phi = \int_0^{\pi/8} g(\phi)\,\mathrm{d}\phi,$$

所以 $[0, \frac{1}{4}\pi]$ 的兩半貢獻相等，而

$$I = 2\int_0^{\pi/8} \frac{\mathrm{d}\theta}{\left[\sqrt{2}\cos\theta\cos\left(\frac{1}{4}\pi - \theta\right)\right]^{k+1}}.$$



圖中所示是 $k=1$ 時的 g：兩個着色的半邊互為鏡像。

由此得出積到 $\sqrt{2}-1$ 的積分

現在在 $[0,\frac{1}{8}\pi]$ 上把代換 $t=\tan\theta$ 反過來做。上面計算的每一步都保持不變；只有上限移到了 $t=\tan\frac{1}{8}\pi$。由 $\tan\frac{1}{4}\pi=\dfrac{2\tau}{1-\tau^2}=1$，其中 $\tau=\tan\frac{1}{8}\pi$，可得 $\tau^2+2\tau-1=0$，所以 $\tau=-1\pm\sqrt{2}$，而正根就是 $\tan\frac{1}{8}\pi=\sqrt{2}-1$。因此

$$I=2\int_0^{\sqrt{2}-1}\frac{(1+t^2)^k}{(1+t)^{k+1}}\,\mathrm{d}t,$$

這正是所求。把這個結果與同一被積函數的 $I=\int_0^1$ 比較，我們發現在 $[0,\sqrt{2}-1]$ 上的積分和在 $[\sqrt{2}-1,1]$ 上的積分*兩者都*等於 $\frac{1}{2}I$。這就是我們接下來要用的事實。

推論

左方。寫成 $\dfrac{2-2x^2+x^4}{x^2}=\dfrac{2-2u+u^2}{u}$，其中 $u=x^2$。那麼 $\mathrm{d}u=2x\,\mathrm{d}x$，所以 $\dfrac{\mathrm{d}x}{x}=\dfrac{\mathrm{d}u}{2u}$，而 u 由 1 變到 2：

$$\int_1^{\sqrt{2}}\left(\frac{2-2x^2+x^4}{x^2}\right)^k\frac{1}{x}\,\mathrm{d}x=\frac{1}{2}\int_1^2\frac{(2-2u+u^2)^k}{u^k}\cdot\frac{1}{u}\,\mathrm{d}u=\frac{1}{2}I.$$

右方。被積函數是 $\dfrac{(2-2x+x^2)^k}{x^{k+1}}$，與 I 的定義中的相同，但積分區間是 $[1,\sqrt{2}]$。再次使用第一個代換 $x=2/(1+t)$。這時 $x=\sqrt{2}$ 對應 $1+t=\sqrt{2}$，即 $t=\sqrt{2}-1$，所以

$$\begin{aligned}\int_1^{\sqrt{2}}\left(\frac{2-2x+x^2}{x}\right)^k\frac{1}{x}\,\mathrm{d}x&=\int_{\sqrt{2}-1}^1\frac{(1+t^2)^k}{(1+t)^{k+1}}\,\mathrm{d}t\\&=I-\int_0^{\sqrt{2}-1}\frac{(1+t^2)^k}{(1+t)^{k+1}}\,\mathrm{d}t=I-\tfrac{1}{2}I=\tfrac{1}{2}I.\end{aligned}$$

兩方都等於 $\frac{1}{2}I$，所以兩者相等，這就是所要證明的結果。

驗算

- $k = 0$：$I = \int_1^2 \frac{1}{x}\,\mathrm{d}x = \ln 2$，而最後那個恆等式的兩方都是 $\int_1^{\sqrt{2}} \frac{1}{x}\,\mathrm{d}x = \frac{1}{2}\ln 2$，正如所料。
- $k = 1$：$I = \int_1^2 \left(\frac{2}{x^2} - \frac{2}{x} + 1\right)\mathrm{d}x = 2 - 2\ln 2 \approx 0.61371$。對 I 的全部五個形式，以及最後那個恆等式的兩方，使用辛普森法則（Simpson's rule）計算，結果吻合至小數點後十位（每一方都是 ≈ 0.30685）。對於非整數和負數的 k，例如 $k = 0.5$、$k = -0.7$ 和 $k = 4.3$，結果同樣吻合。

常見錯誤

- 把 $\cos\theta + \sin\theta$ 寫成單一個餘弦時漏掉因子 $\sqrt{2}$，結果多出一個 $2^{(k+1)/2}$。
- 試圖求出 I 的值。對一般的 k，它並沒有閉合形式（closed form），也不需要有：每一步都只是變數代換或對稱性。
- 在推論中，沒有留意到右方只是 I 的一部分，需要用上「由此得出」的那個結果才能完成。

STEP 之外

映射 $x \mapsto 2/(1+t)$ 是一個莫比烏斯變換（Möbius transformation），又稱分式線性變換（fractional linear transformation）。這類映射把圓和直線送到圓和直線，對於由二次式構成的被積函數，它們是最自然的變數代換。找出被積函數隱藏的對稱性，就像這裏的 $\theta \mapsto \frac{1}{4}\pi - \theta$，是計算沒有初等原函數（antiderivative）的定積分（definite integral）時的常用技巧。想進一步了解莫比烏斯變換，可閱讀 Needham 的 *Visual Complex Analysis*（以圖像方式講解複分析）。想看更多這種風格的積分，可參考 Siklos 的 *Advanced Problems in Mathematics*（進階數學難題集）。

詳細解答

STEP II 1988，第 8 題

純數學 • 矩陣與特徵向量

STEP II 1988，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/88-s2-q8

關鍵思路

模型中的一年是一個線性映射（linear map），因此相當於乘以一個 2×2 矩陣（matrix）M。比例永遠不變的種羣，就是 M 只會把它按比例放大或縮小的種羣，也就是一個*特徵向量*（*eigenvector*）。一旦求得兩個特徵向量，任何初始種羣都可以分拆成沿這兩個方向的兩部分，而每一部分每年只是乘以它自己的特徵值（eigenvalue）。於是，長遠的命運就取決於哪個特徵值較大，以及它前面的係數的正負。

解答

保持不變的比例

把種羣數量寫成列向量（column vector），模型中的一年就是

$$\begin{pmatrix} A' \\ B' \end{pmatrix} = M \begin{pmatrix} A \\ B \end{pmatrix}, \qquad M = \begin{pmatrix} \frac{1}{4} & \frac{3}{4} \\ -\frac{1}{2} & \frac{3}{2} \end{pmatrix}.$$

比例 $A : B$ 保持不變，當且僅當存在某個數 λ 使得 $(A', B') = \lambda(A, B)$，即 (A, B) 是 M 的特徵向量。特徵值滿足

$$\det(M - \lambda I) = \left(\tfrac{1}{4} - \lambda\right)\left(\tfrac{3}{2} - \lambda\right) + \tfrac{3}{8} = \lambda^2 - \tfrac{7}{4}\lambda + \tfrac{3}{4} = 0,$$

所以 $4\lambda^2 - 7\lambda + 3 = (\lambda - 1)(4\lambda - 3) = 0$，得出 $\lambda = 1$ 或 $\lambda = \frac{3}{4}$。

- $\lambda = 1$：由第一行得 $\frac{1}{4}A + \frac{3}{4}B = A$，所以 $A = B$。比例為 1，而這樣的種羣每年都完全不變。
- $\lambda = \frac{3}{4}$：由第一行得 $\frac{1}{4}A + \frac{3}{4}B = \frac{3}{4}A$，所以 $A = \frac{3}{2}B$。比例為 $\frac{3}{2}$；以第二行驗算：$\frac{3}{2}B - \frac{1}{2} \cdot \frac{3}{2}B = \frac{3}{4}B$。兩個物種每年都縮減至原來的 $\frac{3}{4}$。

不用矩陣也可以得到同樣的結果：若 $A = rB$，則新的比例為 $\dfrac{\frac{1}{4}r + \frac{3}{4}}{\frac{3}{2} - \frac{1}{2}r}$，令它等於 r 便得 $2r^2 - 5r + 3 = 0$，所以 $r = 1$ 或 $r = \frac{3}{2}$。無論用哪種方法，比例只有在等於 1 或 $\frac{3}{2}$ 時才能保持不變，正如題目所要求。我們取 $x = 1$ 和 $y = \frac{3}{2}$。

其後各年

設

$$\begin{pmatrix} a \\ b \end{pmatrix} = p \begin{pmatrix} 1 \\ 1 \end{pmatrix} + q \begin{pmatrix} \frac{3}{2} \\ 1 \end{pmatrix}.$$

比較各分量，得 $p + \frac{3}{2}q = a$ 和 $p + q = b$。兩式相減，$\frac{1}{2}q = a - b$，所以

$$q = 2(a - b), \qquad p = b - q = 3b - 2a.$$

由於 M 是線性的，一年之後第一部分乘以 1，第二部分乘以 $\frac{3}{4}$；n 年之後（只要模型仍然適用），我們有

$$\begin{pmatrix} a_n \\ b_n \end{pmatrix} = (3b - 2a) \begin{pmatrix} 1 \\ 1 \end{pmatrix} + 2(a - b)\left(\frac{3}{4}\right)^n \begin{pmatrix} \frac{3}{2} \\ 1 \end{pmatrix},$$

即

$$a_n = (3b - 2a) + 3(a - b)\left(\tfrac{3}{4}\right)^n,$$
$$b_n = (3b - 2a) + 2(a - b)\left(\tfrac{3}{4}\right)^n.$$

由於 $(\frac{3}{4})^n \to 0$，種羣數量趨向 $(3b - 2a,\ 3b - 2a)$，這是平衡點（equilibrium）所組成的直線 $a = b$ 上的一點。

分類之前，先確認一件可以放心的事：土豚永遠不會出問題。若 $a_n \geq 0$ 且 $b_n > 0$，則 $a_{n+1} = \frac{1}{4}a_n + \frac{3}{4}b_n > 0$，所以只要水牛仍然存活，土豚的數量也保持為正數。因此唯一需要留意的是 b_n 的正負，而我們只需考慮平面上 $a \geq 0$、$b > 0$ 的區域。

> **想一想** b_n 會不會有一段時間跌至零以下，然後回升並趨向一個正的極限（limit）？

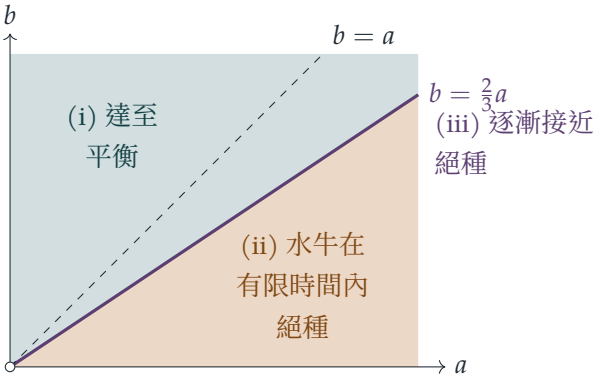
不會。在 b_n 的公式中，只有 $(\frac{3}{4})^n$ 與 n 有關，而它由 1 穩步遞減至 0。所以 b_n 由 $b_0 = b$ 單調（monotonic）地移向它的極限 $3b - 2a$，而每個 b_n 都介乎 b 與 $3b - 2a$ 之間。這樣分類就乾淨利落了。

三個區域

(i) 若 $3b - 2a > 0$，即 $b > \frac{2}{3}a$，則 b 和 $3b - 2a$ 都是正數，所以每個 b_n 都是正數，模型永遠適用。種羣趨向平衡點 $(3b - 2a,\ 3b - 2a)$，而這個平衡點並非零。這是直線 $b = \frac{2}{3}a$ 上方的區域（包括 b 軸，即 $a = 0$）。

(ii) 若 $3b-2a<0$，即 $b<\frac{2}{3}a$，則 $a>\frac{3}{2}b>b$，所以 $a-b>0$，而 b_n 嚴格遞減並趨向負數 $3b-2a$。因此對某個有限的 n 有 $b_n\le 0$：水牛在有限時間內絕種。這是直線 $b=\frac{2}{3}a$ 與 a 軸之間的區域。

(iii) 若 $3b-2a=0$，即 $b=\frac{2}{3}a$ 且 $a>0$，則 $p=0$，種羣純粹是 $(\frac{3}{2},1)$ 的倍數：$b_n=b(\frac{3}{4})^n$，它對每個 n 都是正數，但趨向 0。水牛（以及土豚）逐漸接近絕種，卻永遠不會真正絕種。這只是半直線（half-line）$b=\frac{2}{3}a$，$a>0$。



虛線 $b=a$ 由已經處於平衡狀態的種羣組成；區域 (i) 內的每個起點都會向它移動，並且沿着與直線 $b=\frac{2}{3}a$ 平行的方向滑動（$(\frac{3}{2},1)$ 方向的分量不斷縮小，而 $(1,1)$ 分量則保持不變）。

驗算

- 取 $n=0$，公式給出 $a_0=3b-2a+3a-3b=a$ 和 $b_0=3b-2a+2a-2b=b$，正如所應得的結果。
- 由 $(a,b)=(1,3)$ 出發，它位於區域 (i)。按模型計算一年，得 $(\frac{1}{4}+\frac{9}{4},\ \frac{9}{2}-\frac{1}{2})=(\frac{5}{2},4)$，而公式給出 $7-6(\frac{3}{4})=\frac{5}{2}$ 和 $7-4(\frac{3}{4})=4$。極限為 $(7,7)$。
- 我們用電腦以分數作精確運算，由 3000 個隨機起點迭代這個模型，確認了閉合形式（closed form）在 40 年內都成立，也確認了 $3b-2a$ 的正負在每一個情況下都能正確預測結果，包括 $b=\frac{2}{3}a$ 上的點，例如 $(3,2)$ 在五年後變成 $(\frac{729}{1024},\frac{243}{512})$，即起點的 $(\frac{3}{4})^5$ 倍。

常見錯誤

- 把比例上下顛倒：題目問的是土豚對水牛的比例，所以答案是 1 和 $\frac{3}{2}$，而不是 1 和 $\frac{2}{3}$。

- 忘記模型在 $b_n \leq 0$ 時便告終止，於是聲稱每個種羣都趨向 $(3b - 2a, 3b - 2a)$，連「極限」為負數的情況也包括在內。
- 把區域 (iii) 說成一塊面積：它只是半直線 $b = \frac{2}{3}a$，即 (i) 與 (ii) 之間的邊界。

STEP 之外

這是一個離散線性動力系統（discrete linear dynamical system），而這裏的方法（先對角化（diagonalise），再把每個特徵值取 n 次方）正是大學一年級線性代數計算 M^n 的做法。特徵值 1 給出一條由不動點（fixed point）組成的直線，而特徵值 $\frac{3}{4} < 1$ 則給出一個不斷衰減的方向；把時間改為連續，並以 $e^{\lambda t}$ 取代 λ^n，同樣的圖像就支配着線性微分方程組（systems of linear differential equations）。Strang 的 *Introduction to Linear Algebra* 有一章很好地講解差分方程（difference equation）和矩陣的冪，而 Boyce 與 DiPrima 的 *Elementary Differential Equations* 則處理連續的版本。

詳細解答

STEP II 1988，第 9 題

純數學 • 子群與併集

> **STEP II 1988，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/88-s2-q9

本題在 *Beyond STEP* 第 6.6 節子羣的併集中有完整的引導式講解：https://beyondstep.keithmcnulty.org/zh-hant/ch06/subgroup-unions.html

關鍵思路

用反證法（proof by contradiction）。假如兩個子羣（subgroup）都不是整個羣，就取一個只屬於 G_2 的元素，再取一個只屬於 G_1 的元素，然後問：它們的乘積落在哪裏？無論把乘積放進哪一個子羣，乘法的封閉性（closure）和逆元（inverse）都會導出矛盾。換成三個子羣，乘積就多了一個去處。最自然的嘗試，是找一個每個元素都是自身逆元的羣。

解答

兩個子羣

假設 G 的每個元素都屬於 G_1 或 G_2，並為了導出矛盾，假設 $G_1 \neq G$ 且 $G_2 \neq G$。

由於 $G_1 \neq G$ 而 $G_1 \subseteq G$，存在元素 $a \in G$ 使得 $a \notin G_1$。每個元素都屬於 G_1 或 G_2，所以 $a \in G_2$。

同理，由於 $G_2 \neq G$，存在元素 $b \in G$ 使得 $b \notin G_2$，於是 $b \in G_1$。

因此我們有

$$a \in G_2, \quad a \notin G_1, \qquad\qquad b \in G_1, \quad b \notin G_2.$$

現在考慮乘積 $ab \in G$。按假設，它屬於 G_1 或 G_2。

- 若 $ab \in G_1$：由於 $b \in G_1$ 而 G_1 是子羣，故 $b^{-1} \in G_1$，於是由封閉性得

$$a = (ab)\,b^{-1} \in G_1,$$

  這與 $a \notin G_1$ 矛盾。

- 若 $ab \in G_2$：由於 $a \in G_2$，故 $a^{-1} \in G_2$，於是

$$b = a^{-1}(ab) \in G_2,$$

  這與 $b \notin G_2$ 矛盾。

兩種情況都不可能發生，所以我們的假設不成立：$G_1 = G$ 或 $G_2 = G$。 ■

想一想 我們在哪裏用到了 G 是有限羣這個條件？

完全沒有用到。這個論證只用到子羣的性質（封閉性和逆元），對任何羣都成立，不論有限還是無限。換句話說：一個羣永遠不會是兩個真子羣（proper subgroup）的併集（union）。

三個子羣

取 H 為克萊因四元羣（Klein four-group）

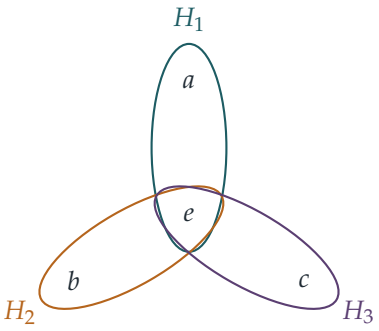
$$H = \{e,\ a,\ b,\ c\}, \qquad a^2 = b^2 = c^2 = e, \quad ab = ba = c,\ bc = cb = a,\ ca = ac = b.$$

具體來說，H 就是 $\{(0,0),(1,0),(0,1),(1,1)\}$ 在每個坐標各自模 2 相加之下所成的羣，也可以看成非正方形的長方形的對稱羣（symmetry group）（恆等變換、兩個反射和半周旋轉）。

每個非單位元的元素的階（order）都是 2，所以

$$H_1 = \{e,a\}, \qquad H_2 = \{e,b\}, \qquad H_3 = \{e,c\}$$

每一個都是子羣：它包含 e，在乘法下封閉（唯一非平凡的乘積是 $a \cdot a = e$，b 和 c 的情況相同），而且每個元素都是自身的逆元。它們各有 2 個元素，所以沒有一個等於 H。但它們合起來包含了 e、a、b 和 c，也就是 H 的全部元素。



所以三個真子羣是可以覆蓋一個羣的。另一個例子是四元數羣（quaternion group）$Q_8 = \{\pm 1, \pm i, \pm j, \pm k\}$，它是三個循環子羣（cyclic subgroup）$\{\pm 1, \pm i\}$、$\{\pm 1, \pm j\}$ 和 $\{\pm 1, \pm k\}$ 的併集，每個子羣的階都是 4。

留意第一部分的論證為何在這裏失效：當 $a \notin H_2$ 而 $b \notin H_1$ 時，乘積 $ab = c$ 可以溜進第三個子羣 H_3。

驗算

- 以窮舉法搜尋下列各羣的所有子羣：階為 1 至 12 的循環羣、S_3、S_4、階為 8 的二面體羣（dihedral group）、Q_8、$\mathbb{Z}_2^2$ 和 $\mathbb{Z}_2^3$（共 1278 個由真子羣組成的有序對），結果找不到任何一對的併集等於整個羣。
- 同一次搜尋亦確認 $\{e, a\} \cup \{e, b\} \cup \{e, ab\}$ 就是整個克萊因四元羣，而 Q_8 是 $\langle i \rangle$、$\langle j \rangle$、$\langle k \rangle$ 的併集。有趣的是，所測試的循環羣，以及 S_3 和 S_4，沒有一個是三個真子羣的併集。
- 對第一部分作一個簡單的合理性檢查：在循環羣中，生成元（generator）不屬於任何真子羣，所以循環羣無論用多少個真子羣都無法覆蓋。

常見錯誤

- 用計數來論證：兩個真子羣各自最多有 $\frac{1}{2}|G|$ 個元素，但它們都含有 e，所以併集最多有 $|G| - 1$ 個元素。這個計數證明是正確的（根據拉格朗日定理（Lagrange's theorem)），而且簡短，但它確實需要有限性和拉格朗日定理；兩者都必須寫明。
- 取了 $a \notin G_1$ 和 $b \notin G_2$，卻忘記使用 $a \in G_2$ 和 $b \in G_1$。缺少這兩個事實，就推不出矛盾。
- 給出的三個子羣並非全是真子羣，或者根本不是子羣（例如 $\{e, a, b\}$，由於 $ab = c$，它並不封閉）。

STEP 之外

哪些羣是三個真子羣的併集？Scorza 的一個定理（1926 年）給出了答案：恰好是那些有商羣（quotient）同構（isomorphic）於克萊因四元羣的羣，這也正是克萊因羣成為基本例子的原因。同樣的「不能是兩個的併集」這一事實也出現在線性代數（linear algebra）中：無限域（infinite field）上的向量空間（vector space）永遠不會是有限個真子空間（proper subspace）的併集。若想找一本羣論的入門書，可以試試 Allenby 的 *Rings, Fields and Groups*（《環、域與羣》）。

詳細解答

STEP II 1988，第 10 題

純數學 • 拋物面的向量幾何

> **STEP II 1988，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/88-s2-q10

關鍵思路

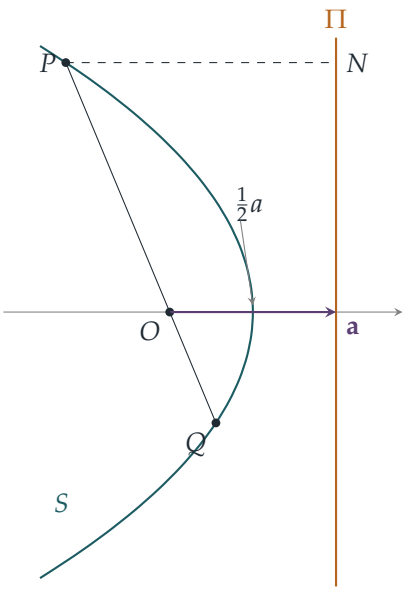
方程中的兩個部分各有清晰的幾何意義：r 是與 O 的距離，而 $(a^2 - \mathbf{a}\cdot\mathbf{r})/a$ 是與平面 $\mathbf{r}\cdot\mathbf{a} = a^2$ 的有向距離（signed distance）。因此 S 是一個旋轉拋物面（paraboloid of revolution），以 O 為焦點（focus），以該平面為準面（directrix）。至於其餘部分，可沿着通過 O 的直線來處理：把直線 OP 上的點寫成 $\lambda\mathbf{p}$，把通過 O 且同時垂直於 $\mathbf{a}$ 和 $\mathbf{p}$ 的直線上的點寫成 $\mu\,\mathbf{a}\times\mathbf{p}$，然後代入 S 的方程。λ 的正負號十分重要，因為 $|\lambda\mathbf{p}| = |\lambda|p$。

解答

軌跡

設 Π 為平面 $\mathbf{r}\cdot\mathbf{a} = a^2$。它的單位法向量（unit normal）是 $\hat{\mathbf{a}} = \mathbf{a}/a$，與 O 的距離為 a，並位於 $\mathbf{a}$ 所指的一側。由點 $\mathbf{r}$ 到 Π 的垂直距離是

$$d(\mathbf{r}) = |\,a - \hat{\mathbf{a}}\cdot\mathbf{r}\,| = \frac{|a^2 - \mathbf{a}\cdot\mathbf{r}|}{a}.$$



上圖顯示 S 被一個包含 $\mathbf{a}$ 的平面所截出的截面。點 P 與 O 的距離，等於它與其在 Π 上的垂足 N 的距離。曲面與軸相交於 $\frac{1}{2}\mathbf{a}$，即 O 與 Π 的正中間。

S 上的點是等距的。若 $\mathbf{a}\cdot\mathbf{r}+ar=a^2$，則 $a^2-\mathbf{a}\cdot\mathbf{r}=ar\geq 0$。因此

$$d(\mathbf{r})=\frac{a^2-\mathbf{a}\cdot\mathbf{r}}{a}=\frac{ar}{a}=r=|OR|.$$

等距的點都在 S 上。假設 $r=d(\mathbf{r})$，即 $ar=|a^2-\mathbf{a}\cdot\mathbf{r}|$。若 $a^2-\mathbf{a}\cdot\mathbf{r}\geq 0$，這正是 S 的方程。否則 $\mathbf{a}\cdot\mathbf{r}>a^2$，而該條件表示 $\mathbf{a}\cdot\mathbf{r}=ar+a^2>ar$。這與 $\mathbf{a}\cdot\mathbf{r}\leq ar$（柯西–施瓦茨不等式，Cauchy–Schwarz）矛盾。所以在 Π 另一側的點都不會等距，而 S 恰好就是與 O 和 Π 等距的所有點所組成的集合。這是一個旋轉拋物面，其旋轉軸是通過 O 且方向為 $\mathbf{a}$ 的直線。

以下兩個推論稍後會用到。對於 S 上的點 P，

$$\mathbf{a}\cdot\mathbf{p}=a^2-ap. \qquad (*)$$

此外 $a^2=\mathbf{a}\cdot\mathbf{p}+ap\leq 2ap$，所以 $p\geq\frac{1}{2}a$，從而 $2p-a\geq 0$。等號只在 $\mathbf{p}$ 是 $\mathbf{a}$ 的正數倍時成立，即 P 是頂點（vertex）$\frac{1}{2}\mathbf{a}$ 的時候。

第二個交點 Q

直線 OP 上的點是 $\lambda\mathbf{p}$，其中 $|\lambda\mathbf{p}|=|\lambda|p$。代入 S 的方程得

$$\lambda\,\mathbf{a}\cdot\mathbf{p}+a|\lambda|p=a^2.$$

想一想 為甚麼 Q 不可能與 P 位於 O 的同一側？

若 $\lambda > 0$，左方是 $\lambda(\mathbf{a}\cdot\mathbf{p} + ap) = \lambda a^2$，所以 $\lambda = 1$，這就是 P 本身。（從幾何上看，由焦點出發的射線與拋物面最多相交一次。）因此 Q 對應 $\lambda < 0$，這時 $|\lambda| = -\lambda$。利用 $(*)$，

$$\lambda(\mathbf{a}\cdot\mathbf{p} - ap) = a^2 \quad\Longrightarrow\quad \lambda(a^2 - 2ap) = a^2 \quad\Longrightarrow\quad \lambda = -\frac{a}{2p-a}.$$

只要 $2p > a$，即 P 不是頂點，這個值便如所要求般是負數。（在頂點處，直線 OP 就是軸，它與 S 只相交一次。）因此

$$\boxed{\mathbf{q} = -\frac{a}{2p-a}\,\mathbf{p}} \qquad \text{及} \qquad |OQ| = \frac{ap}{2p-a}.$$

點 T 和 T'

通過 O 且同時垂直於 $\mathbf{a}$ 和 $\mathbf{p}$ 的直線是 $\mathbf{t} = \mu\,\mathbf{a}\times\mathbf{p}$。（我們需要 $\mathbf{a}\times\mathbf{p} \neq \mathbf{0}$，所以 P 不在軸上。）由於 $\mathbf{a}\cdot\mathbf{t} = 0$，$S$ 的方程簡化為 $at = a^2$，即 $|\mathbf{t}| = a$。所以 T 和 T' 位於 S 與通過 O 且平行於 Π 的平面相交所得的圓上，並且

$$|\mu|\,|\mathbf{a}\times\mathbf{p}| = a.$$

想一想 在不知道 $\mathbf{a}$ 與 $\mathbf{p}$ 之間夾角的情況下，怎樣求出 $|\mathbf{a}\times\mathbf{p}|$？

利用恆等式（identity）$|\mathbf{a}\times\mathbf{p}|^2 = a^2p^2 - (\mathbf{a}\cdot\mathbf{p})^2$，它其實就是 $\sin^2\theta = 1 - \cos^2\theta$ 的另一種寫法。然後代入 $(*)$：

$$\begin{aligned}|\mathbf{a}\times\mathbf{p}|^2 &= a^2p^2 - (a^2 - ap)^2\\ &= a^2p^2 - a^4 + 2a^3p - a^2p^2\\ &= a^3(2p-a).\end{aligned}$$

因此

$$|\mu| = \frac{a}{a^{3/2}\sqrt{2p-a}} = \frac{1}{\sqrt{a(2p-a)}} = \frac{1}{\sqrt{2ap-a^2}},$$

而 T 和 T' 的位置向量（position vector）是 $\pm\dfrac{1}{\sqrt{2ap-a^2}}\,\mathbf{a}\times\mathbf{p}$，正如所求。

三角形 PQT 的面積

P、O、Q 三點共線（collinear），而 O 在 P 與 Q 之間。向量 $\mathbf{t}$ 平行於 $\mathbf{a}\times\mathbf{p}$，所以它垂直於 $\mathbf{p}$，從而垂直於直線 PQ。因此 OT 是三角形 PQT 中由 T 引出的高，其長度為 a。底邊是

$$|PQ| = p + \frac{ap}{2p-a} = \frac{p(2p-a)+ap}{2p-a} = \frac{2p^2}{2p-a}.$$

因此

$$\text{面積} = \tfrac{1}{2}\cdot|PQ|\cdot|OT| = \tfrac{1}{2}\cdot\frac{2p^2}{2p-a}\cdot a = \frac{ap^2}{2p-a},$$

正如所求。把 T 換成 T'，結果同樣成立。

驗算

- 取 $\mathbf{a} = (2,0,0)$ 和 $P = (-1.25,3,0)$，即圖中的那一點。這時 $\mathbf{a}\cdot\mathbf{p} + ap = -2.5 + 2(3.25) = 4 = a^2$，所以 P 在 S 上。公式給出 $\mathbf{q} = -\frac{2}{4.5}\mathbf{p} = (0.556,-1.333,0)$，而 $\mathbf{a}\cdot\mathbf{q} + aq = 1.111 + 2.889 = 4$，所以 Q 也在 S 上。
- 取 P 為正焦弦（latus rectum）的一個端點，於是 $\mathbf{p}\perp\mathbf{a}$ 且 $p = a$。這時 $\mathbf{q} = -\mathbf{p}$，由對稱性可知這是正確的。面積是 $\frac{1}{2}(2a)(a) = a^2$，而公式給出 $a\cdot a^2/a = a^2$。
- 兩個焦距滿足 $\dfrac{1}{OP} + \dfrac{1}{OQ} = \dfrac{1}{p} + \dfrac{2p-a}{ap} = \dfrac{2}{a}$，這正是拋物線經典的焦弦（focal chord）性質。
- 在數值上，我們隨機選取了 20 000 組 $\mathbf{a}$ 和 P，檢驗了 Q、T、T' 和面積。最大誤差約為 10^{-11}。

常見錯誤

- 對所有 λ 都寫成 $|\lambda\mathbf{p}| = \lambda p$。這樣唯一的解就是 $\lambda = 1$，便無法找到 Q。
- 在第一部分忘記證明逆命題（converse）。位於 Π 另一側的點滿足 $|a^2 - \mathbf{a}\cdot\mathbf{r}| = \mathbf{a}\cdot\mathbf{r} - a^2$，我們必須排除這些點。
- 用 $\overrightarrow{PQ}$ 與 $\overrightarrow{PT}$ 的向量積（cross product）來計算面積。這樣做是可行的，但比起留意到 $OT\perp PQ$ 要冗長得多。

STEP 之外

這就是圓錐曲線（conic）的焦點–準線定義，只是放在三維空間中，並且不依賴坐標。設 e 為離心率（eccentricity），則曲面 $e\,\mathbf{a}\cdot\mathbf{r} + ar = e\,a^2$（與 O 的距離等於與 Π 的距離的 e 倍）給出其他以 O 為一個焦點的旋轉二次曲面（quadric of revolution）。關係式 $1/OP + 1/OQ = 2/a$ 是極方程（polar equation）$\ell/r = 1 + e\cos\theta$ 的向量形式，而這個方程在軌道力學（orbital mechanics）中佔有核心地位。若想進一步學習向量代數，Strang 的 *Introduction to Linear Algebra*（線性代數入門）是很好的下一步；Siklos 的 *Advanced Problems in Mathematics*（高等數學難題）則收錄了更多這類風格的題目。

詳細解答

STEP II 1988，第 11 題

力學 • 繩纏繞柱子

> **STEP II 1988，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/88-s2-q11

關鍵思路

繩子未纏上圓柱的部分，總是沿切線（tangent）離開圓柱。因此質點的位置就是「接觸點加上一個長度等於未纏繞繩長的切向量」，而這很容易用角 θ 表示出來。求導之後，有兩項互相抵消，速度原來與繩子垂直。於是張力（tension）不做功，速率保持在初始值，而速率公式就成為一條關於 θ 的可分離變量微分方程（separable differential equation）。

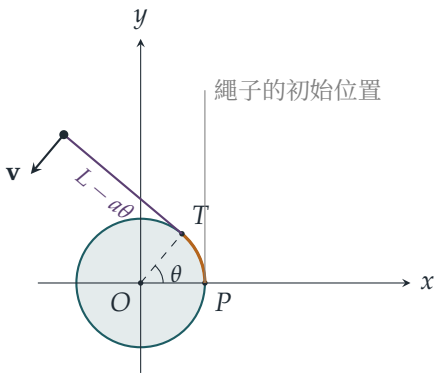
解答

建立模型

我們把質點看成光滑桌面上的一個質點（point mass），繫着一條始終拉緊的輕而不可伸長的繩子（light inextensible string）。重力與法向反作用力（normal reaction）都沿鉛垂方向，互相抵消，所以唯一的水平力就是張力。

取以 O 為原點的坐標軸，並把 P 放在 $(a, 0)$。繩子起初垂直於 OP，所以它在 P 處與圓柱相切，我們設質點從 (a, L) 出發。（若取鏡像的情況，只是把 y 坐標的正負號反轉。）當繩子沿逆時針方向纏上圓柱，纏過角 θ 之後，長度 $a\theta$ 的一段貼在由 P 起的圓弧上，而繩子離開圓柱的位置是

$$T = a(\cos\theta,\ \sin\theta).$$



繩子未纏上的部分長 $L-a\theta$，沿 T 處的切線伸展。指向繩子展開方向的單位切向量（unit tangent）是 $(-\sin\theta,\ \cos\theta)$：它垂直於 OT，而且當 $\theta=0$ 時等於 $(0,1)$，與起始位置吻合。

位置與速度

因此質點的位置向量（position vector）是

$$\mathbf{r}=a\begin{pmatrix}\cos\theta\\ \sin\theta\end{pmatrix}+(L-a\theta)\begin{pmatrix}-\sin\theta\\ \cos\theta\end{pmatrix}=\begin{pmatrix}a\cos\theta-(L-a\theta)\sin\theta\\ a\sin\theta+(L-a\theta)\cos\theta\end{pmatrix}.$$

利用鏈式法則（chain rule）和乘積法則（product rule）對 t 求導，

$$\begin{aligned}\dot{\mathbf{r}}&=a\dot\theta\begin{pmatrix}-\sin\theta\\ \cos\theta\end{pmatrix}-a\dot\theta\begin{pmatrix}-\sin\theta\\ \cos\theta\end{pmatrix}+(L-a\theta)\dot\theta\begin{pmatrix}-\cos\theta\\ -\sin\theta\end{pmatrix}\\&=-(L-a\theta)\dot\theta\begin{pmatrix}\cos\theta\\ \sin\theta\end{pmatrix}.\end{aligned}$$

首兩項互相抵消：接觸點沿圓柱移動所帶來的效果，恰好被未纏上的繩子縮短所抵銷。

由於 $(\cos\theta,\sin\theta)$ 是單位向量，$L-a\theta>0$，而且繩子纏上圓柱期間 $\dot\theta>0$，所以速率是

$$|\dot{\mathbf{r}}|=(L-a\theta)\dot\theta,$$

正是所求。

碰到圓柱所需的時間

想一想 繩子一直拉着質點，為甚麼質點的速率仍會保持等於 v？

速度平行於 $(\cos\theta,\sin\theta)$，而繩子（因而張力）則沿 $(-\sin\theta,\cos\theta)$ 方向。兩者互相垂直，所以張力不做功。桌面光滑，張力是唯一的水平力。因此動能（kinetic energy）保持不變，速率也隨之恆等於 v。張力只會改變速度的方向。

所以，在質點到達圓柱之前的所有時刻 t，

$$(L - a\theta)\,\frac{\mathrm{d}\theta}{\mathrm{d}t} = v, \qquad \theta(0) = 0.$$

分離變量，並由 0 積分至 t，

$$L\theta - \tfrac{1}{2}a\theta^2 = vt.$$

當整條繩子都纏上圓柱時，質點便碰到圓柱，即 $a\theta = L$，所以 $\theta = L/a$。於是

$$vt = \frac{L^2}{a} - \frac{L^2}{2a} = \frac{L^2}{2a}, \qquad \text{所以} \qquad t = \frac{L^2}{2av},$$

這就是所求的時間。

驗算

- 當 $\theta = 0$ 時，公式給出 $\mathbf{r} = (a, L)$ 及 $\dot{\mathbf{r}} = (-L\dot{\theta}, 0)$，與初始的繩子垂直，理應如此。
- 單位：$L^2/(av)$ 的單位是 $\mathrm{m}^2/(\mathrm{m}\cdot\mathrm{m\,s^{-1}}) = \mathrm{s}$。
- 另一個理解答案的方法：質點整條路徑的長度是

  $$\int_0^{L/a} (L - a\theta)\,\mathrm{d}\theta = \frac{L^2}{2a},$$

  並以恆定速率 v 走完。
- 我們用數值方法模擬了這個運動：一條非常硬的繩子纏繞在圓柱上，取 $a = 1$、$L = 3$、$v = 2$。模擬所得的時間與 $L\theta - \frac{1}{2}a\theta^2 = vt$ 相符，誤差在 10^{-5} 以內；而把 $\dot{\theta} = v/(L - a\theta)$ 積分至 $a\theta = L$，得到 2.25，即 $L^2/(2av)$。

常見錯誤

- 使用關於 O 的角動量守恆（conservation of angular momentum）。張力並不通過 O（它沿切線作用），所以關於 O 的角動量並不守恆。能量才是合適的工具。
- 把未纏上的繩子寫成沿 $(\cos\theta, \sin\theta)$ 方向，即沿徑向，而不是沿切線方向。
- 忘記速率公式需要 $\dot{\theta} > 0$。該向量的絕對值是 $(L - a\theta)|\dot{\theta}|$。

STEP 之外

質點所描出的路徑是圓的*漸伸線*(involute)，即從線軸上拉緊展開的一條線，其末端所描出的曲線。齒輪的齒之所以造成那種形狀，正是因為漸伸線；它也是惠更斯（Huygens）

研究擺鐘時的核心。它的弧長 $\int (L - a\theta)\, \mathrm{d}\theta$ 是平面曲線微分幾何（differential geometry，包括漸屈線（evolute）、漸伸線和曲率（curvature））的一個入門例子。至於力學方面，包括能量與角動量在甚麼情況下守恆、甚麼情況下不守恆，可閱讀 Kleppner and Kolenkow, *An Introduction to Mechanics*（一本大學力學入門教科書）。

詳細解答

STEP II 1988，第 12 題

力學 • 下落的鏈條

STEP II 1988，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：

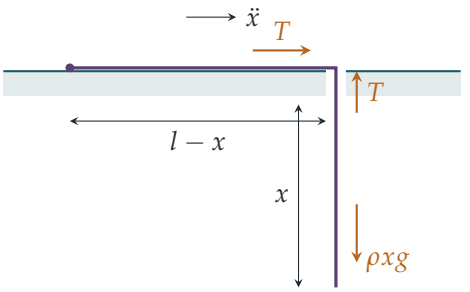
step.maths.org/questions/88-s2-q12

關鍵思路

由於繩子不可伸長（inextensible），繩上每一小段的速率都相同，所以整條繩子就像一個物體，而推動它的力就是垂在桌面以下那一段的重量。隨着愈來愈多繩子穿過小孔，這個重量不斷增大，於是得到一條以下垂長度為未知量的簡單線性微分方程（linear differential equation）。它的解是雙曲餘弦（hyperbolic cosine），而所求的時間則由反解這個函數得出。

解答

整個系統完全沒有摩擦：桌面光滑（smooth），小孔既小且光滑（所以它只改變繩子的方向，而不會沿繩子方向施加任何力），而繩子每單位長度的質量為 ρ。設在時間 t 時垂在桌面以下的長度為 x，則留在桌面上的長度為 $l - x$。起初 $x = l - d$ 且 $\dot{x} = 0$，而當 $x = l$ 時繩端到達小孔頂部。由於桌子的高度大於 l，在此之前下垂部分不會碰到地面。



設 T 為繩子在小孔處的張力（tension）。由於小孔光滑，緊貼小孔上方和下方的張力同為 T。

桌面上的部分。它的質量為 $\rho(l-x)$，以加速度 $\ddot{x}$ 沿水平方向移向小孔（因為繩子不可伸長，它與下垂部分移動得一樣快）。由於桌面光滑，作用在它上面的水平力只有張力 T。所以

$$\rho(l-x)\,\ddot{x} = T.$$

下垂的部分。它的質量為 ρx，以加速度 $\ddot{x}$ 向下運動，所受的力是向下的重量 $\rho x g$ 和向上的張力 T：

$$\rho x\,\ddot{x} = \rho x g - T.$$

把兩條方程相加，便可消去 T：

$$\rho l\,\ddot{x} = \rho g x, \qquad \text{即} \qquad \ddot{x} = n^2 x, \quad n = \sqrt{\frac{g}{l}}.$$

想一想 小孔兩邊的繩子長度一直在變。對每一段分別使用 $F = ma$，真的沒有問題嗎？

沒有問題。每條方程都是在某一瞬間，應用於該瞬間位於小孔那一邊的那一段繩子；整條繩子以同一速率 $\dot{x}$ 一起運動，所以整條繩子的動量（momentum）變化率是 $\rho l\ddot{x}$，而沿運動方向有分量的外力，就只有下垂部分的重量 $\rho x g$。（桌面的反作用力抵銷了水平部分的重量，而小孔只是令繩子轉向。）因此合併後的方程 $\rho l\ddot{x} = \rho g x$ 正是整條繩子的牛頓第二定律（Newton's second law）。下文的能量驗算亦印證了這一點。

$\ddot{x} = n^2 x$ 的通解（general solution）是 $x = A\cosh nt + B\sinh nt$。由 $t = 0$ 時 $\dot{x} = 0$ 得 $B = 0$，再由 $t = 0$ 時 $x = l - d$ 得 $A = l - d$：

$$x = (l-d)\cosh nt.$$

繩端在 $x = l$ 時到達小孔頂部，即在滿足下式的時間 t_1：

$$\cosh nt_1 = \frac{l}{l-d}.$$

由於 $l/(l-d) > 1$，此方程恰有一個正數解。利用 $\operatorname{arcosh} u = \ln\left(u + \sqrt{u^2-1}\right)$，有

$$\sqrt{u^2-1} = \frac{\sqrt{l^2-(l-d)^2}}{l-d} = \frac{\sqrt{2ld-d^2}}{l-d},$$

所以所需時間為

$$\boxed{t_1 = \sqrt{\frac{l}{g}}\,\operatorname{arcosh}\!\left(\frac{l}{l-d}\right) = \sqrt{\frac{l}{g}}\,\ln\left(\frac{l+\sqrt{2ld-d^2}}{l-d}\right).}$$

驗算

- **能量**。取桌面為勢能（potential energy）的零點。下垂部分的質心（centre of mass）位於深度 $x/2$ 處，所以它的勢能是 $-\frac{1}{2}\rho g x^2$。由靜止開始，能量守恆（conservation of energy）給出 $\frac{1}{2}\rho l\dot{x}^2 = \frac{1}{2}\rho g\left(x^2 - (l-d)^2\right)$，對 t 求導便再次得到 $l\ddot{x} = gx$。
- d **很小**。如果 d 極小，幾乎整條繩子都垂着，它應以接近 g 的加速度下落，所需時間約為 $\sqrt{2d/g}$。事實上，當 d 很小時，由 $\cosh nt_1 = 1 + \frac{d}{l} + \cdots$ 和 $\cosh nt_1 \approx 1 + \frac{1}{2}n^2t_1^2$ 可得 $t_1^2 \approx 2d/(n^2 l) = 2d/g$。
- $d \to l$。如果幾乎沒有繩子垂下，繩子幾乎不會動，於是 $t_1 \to \infty$，正如對數式所顯示的那樣。
- **數值**。取 $l = 2\,\mathrm{m}$、$d = 1.5\,\mathrm{m}$ 及 $g = 9.8\,\mathrm{m\,s^{-2}}$，公式給出 $t_1 \approx 0.932\,\mathrm{s}$，與對這個運動逐步進行數值積分所得的結果吻合。

常見錯誤

- 只對下垂部分寫出 $\rho x\ddot{x} = \rho x g$，忘記了它還要拖動桌面上的繩子（換句話說，是用錯了質量：被加速的是整條繩子，質量為 ρl）。
- 嘗試用 $x = A\cos nt$ 形式的解：方程是 $\ddot{x} = +n^2x$，所以運動是按指數增長，而不是振動。
- 把結束條件定為 $x = d$ 而不是 $x = l$：繩端到達小孔時，整段長度 l 都垂在下面。

STEP 之外

這是一個質量分佈隨時間變化的問題：總質量固定不變，但受重力驅動的那一部分隨位置而改變，由此得到一條線性方程，其解為指數函數。假如繩子是從無摩擦的桌邊滑落，同時有一堆原本靜止的繩子不斷被扯動起來，動量便會在非彈性（inelastic）的「猛扯」中損失，能量亦不再守恆；分辨這兩類情況，是大學力學課程中的經典課題。Kleppner 與 Kolenkow 所著的 *An Introduction to Mechanics*（力學入門教科書）對質量流動的系統以及鏈條和繩索的問題有細緻的討論。

詳細解答

STEP II 1988，第 13 題

力學 • 摩擦力與平衡

STEP II 1988，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：

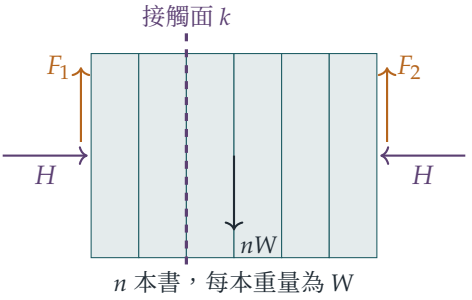
step.maths.org/questions/88-s2-q13

關鍵思路

托住書本的每一個垂直力都是摩擦力（friction），而任何接觸面上的摩擦力都不能超過該處法向反作用力（normal reaction）的 μ 倍。因此我們需要兩樣東西：每個接觸面上的法向反作用力（由水平方向的平衡（equilibrium）很容易求得），以及每個接觸面必須提供的摩擦力（考慮適當選取的一組書本在垂直方向的平衡）。然後我們找出哪一個接觸面最接近滑動。把相鄰的一組書本視為一個整體，可以令代數運算保持簡短。

解答

模型。 設有 n 本相同的書，每本重量為 W，書的封面保持垂直。每隻手以水平力 $H \leq Q$ 壓向書本，並提供向上的摩擦力；把這兩個摩擦力記作 F_1（左手）和 F_2（右手）。由於每隻手的垂直力正正就是這個摩擦力，所以我們需要 $F_1, F_2 \leq P$。題目指明可以忽略轉動效應，因此我們只需考慮每本書或每組書本的力平衡。



法向反作用力。 把書本由左至右編號為 $1, 2, \ldots, n$，並設接觸面 k 為第 k 本書與第 $k+1$ 本書之間的接觸面（$1 \le k \le n-1$）。作用於第 $1, \ldots, k$ 本書這一組書本上的水平力，只有左手的推力 H 和來自第 $k+1$ 本書的法向反作用力。所以該反作用力同樣是 H：每一個接觸面，不論是手與書之間還是書與書之間，法向反作用力都是 H，因此每個接觸面所能提供的摩擦力大小最多為 μH。

雙手。 在垂直方向上，對整排書本而言，

$$F_1 + F_2 = nW.$$

F_1, F_2 各自最多為 P（臂力），亦最多為 $\mu H \le \mu Q$（摩擦力）。記 $L = \min(P, \mu Q)$。於是 $nW = F_1 + F_2 \le 2L$，所以

$$n \le \frac{2L}{W} = \frac{2\min(P, \mu Q)}{W}.$$

這是必要條件。若此條件成立，最佳的做法是盡力壓緊，即 $H = Q$（H 愈大，只會提高每一個摩擦力的上限），並把負荷平均分配，即 $F_1 = F_2 = \frac{1}{2}nW \le L$。這樣雙手都在各自的限度之內。

想一想 即使雙手托得住，書本會不會在中間某處互相滑過？

書本之間的接觸面。 我們直接驗證這一點。對第 $1, \ldots, k$ 本書這一組書本應用垂直方向的平衡。它的重量 kW 由手提供的 $F_1 = \frac{1}{2}nW$ 以及來自第 $k+1$ 本書的摩擦力 f_k（以向上為正）所承托：

$$f_k = kW - \tfrac{1}{2}nW = \left(k - \tfrac{n}{2}\right) W.$$

對於 $1 \le k \le n-1$，我們有 $\left|k - \frac{n}{2}\right| \le \frac{n}{2} - 1$，等號在 $k = 1$ 和 $k = n-1$ 時成立，即緊貼最外側兩本書的接觸面。所以任何接觸面所需的最大摩擦力為 $\left(\frac{n}{2} - 1\right)W$，比每隻手所提供的 $\frac{1}{2}nW$ 少 W。由於每個接觸面的上限與手的一樣，都是 μQ，所以只要雙手不滑動，就沒有任何兩本書會互相滑動。（從物理上看：最外側的每本書由手托住，而它向內傳遞的摩擦力，就是手的摩擦力減去它本身的重量。）

答案。 這排書本能被提起，當且僅當 $nW \le 2\min(P, \mu Q)$。所以書本數目的最大值為

$$\boxed{n_{\max} = \left\lfloor \frac{2\min(P,\, \mu Q)}{W} \right\rfloor}$$

即同時不超過 $2P/W$ 和 $2\mu Q/W$ 的最大整數。哪一個上限起作用，便說明了是甚麼限制着圖書館管理員：當 $P < \mu Q$ 時是提舉的臂力，當 $\mu Q < P$ 時則是握力。

驗算

- 取 $W = 10\,\mathrm{N}$、$P = 60\,\mathrm{N}$、$Q = 100\,\mathrm{N}$ 及 $\mu = 0.4$。則 $\min(60, 40) = 40$，而 $n_{\max} = 8$。有 8 本書時，每隻手提供 40 N 的摩擦力，剛好達到其上限，而最吃力的接觸面需要 $30\,\mathrm{N} < 40\,\mathrm{N}$。若加上第九本書，每隻手便需要提供 45 N。
- 只有一本書（$n = 1$）時沒有書本之間的接觸面，我們只需要 $W \le 2\min(P, \mu Q)$，這顯然是正確的。
- 我們用電腦對 W, P, Q, μ 的 3000 組隨機取值求解了整組不等式（任意 $H \le Q$、任意的 F_1, F_2 分配、每一個接觸面），所得的最大可行 n 每次都與公式吻合。

常見錯誤

- 忘記了手臂的垂直力就是摩擦力，所以 P 和 μQ 這兩個限制是作用於同一個力之上。
- 未經驗證便假設中間的書本是最薄弱之處。書本之間的力比雙手處的力小，而不是大。
- 只給出 $2\min(P, \mu Q)/W$ 而沒有取下整（floor）：答案必須是整數本書。

STEP 之外

這裏的摩擦力是靜不定的（statically indeterminate）：平衡條件只確定了 $F_1 + F_2$，而我們問的是有沒有某一種分配方式能令每個接觸面都不超出其限度。這類問題其實是小型的線性規劃（linear programme），而「檢查是否存在某個可容許的平衡狀態」這個想法，同樣是結構極限分析（limit analysis）的基礎。Kleppner 與 Kolenkow 所著的 *An Introduction to Mechanics*（力學入門教科書）是一個好去處，可以看到摩擦力與平衡如何得到嚴謹的處理。

詳細解答

STEP II 1988，第 14 題

力學 • 轉動的剛體

> **STEP II 1988，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/88-s2-q14

關鍵思路

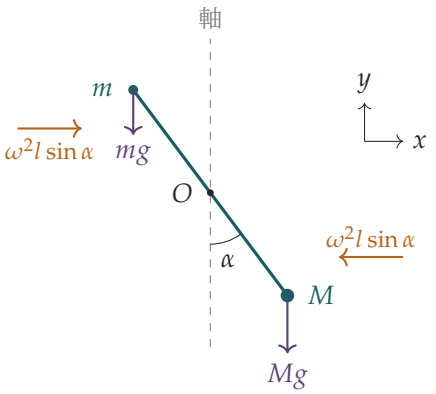
在穩定狀態（steady state）下，相對於包含桿的那個旋轉鉛垂平面，一切都靜止不動，所以每個質點（particle）只是以恆定速率沿水平圓周運動。這樣每個質點的加速度就完全確定了。接着由牛頓第二定律（Newton's second law）可知桿對每個質點所施的力，再加上關於輕桿（light rod）的兩個事實，問題便迎刃而解：作用於輕桿的各力必須平衡，而且（因為桿可繞軸自由擺動）這些力對軸的力矩（moment）也必須平衡。

解答

模型。 把兩個質量視為裝在輕質剛性桿上的質點，桿在 O 處光滑地樞接，因此軸不會沿其自身方向施加力偶（couple）。我們在包含桿的鉛垂平面內討論，這個平面以角速率（angular speed）ω 繞通過 O 的鉛垂線轉動。取 x 為此平面內的水平方向，y 為鉛垂向上，原點為 O。

設 M 位於 $(l\sin\alpha,\ -l\cos\alpha)$，$m$ 位於 $(-l\sin\alpha,\ l\cos\alpha)$；稍後便會看到，較重的質點必定是在下方的那個。每個質點都以恆定速率，沿半徑為 $l\sin\alpha$ 的水平圓周繞鉛垂軸運動，所以其加速度大小為 $\omega^2 l\sin\alpha$，方向水平指向軸：

$$\mathbf{a}_M = (-\omega^2 l\sin\alpha,\ 0), \qquad \mathbf{a}_m = (\omega^2 l\sin\alpha,\ 0).$$



（琥珀色箭頭表示加速度，紫紅色箭頭表示重量。）

桿所施的力。 設 $\mathbf{R}_M$ 和 $\mathbf{R}_m$ 分別為桿對 M 和 m 所施的力。這些力不一定沿着桿的方向：剛性桿也可以向側面施力。由牛頓第二定律得 $\mathbf{R}_M - Mg\,\mathbf{j} = M\mathbf{a}_M$，所以

$$\mathbf{R}_M = \left(-M\omega^2 l\sin\alpha,\ Mg\right), \qquad \mathbf{R}_m = \left(m\omega^2 l\sin\alpha,\ mg\right).$$

根據牛頓第三定律（Newton's third law），兩個質點對桿所施的力，在 M 的一端為 $-\mathbf{R}_M$，在 m 的一端為 $-\mathbf{R}_m$。

? **想一想** 軸在 O 處對桿施加一個未知的力。在不知道這個力的情況下，我們怎樣取力矩？

對軸取力矩。 軸的力作用於 O，所以它對 O 沒有力矩；而桿可自由擺動，所以軸沿其自身方向（即通過 O 並垂直於我們這個平面的直線）不提供力偶。輕桿本身沒有角動量（angular momentum），所以兩端的力對這條直線的力矩必須互相抵消。作用於 (x,y) 的力 (F_x,F_y)，其力矩為 $xF_y - yF_x$（以逆時針方向為正）。對於作用於 $(l\sin\alpha, -l\cos\alpha)$ 的 $-\mathbf{R}_M$：

$$l\sin\alpha\,(-Mg) - (-l\cos\alpha)(M\omega^2 l\sin\alpha) = -Mgl\sin\alpha + M\omega^2 l^2\sin\alpha\cos\alpha,$$

而對於作用於 $(-l\sin\alpha,\, l\cos\alpha)$ 的 $-\mathbf{R}_m$：

$$(-l\sin\alpha)(-mg) - l\cos\alpha\,(-m\omega^2 l\sin\alpha) = mgl\sin\alpha + m\omega^2 l^2\sin\alpha\cos\alpha.$$

令兩者之和為零，並除以 $l\sin\alpha \neq 0$（因為 $\alpha > 0$）：

$$(M+m)\,\omega^2 l\cos\alpha = (M-m)\,g \quad\Longrightarrow\quad \omega^2 = \left(\frac{M-m}{M+m}\right)\frac{g}{l\cos\alpha},$$

其中除以了 $\cos\alpha > 0$，這正是所需的結果。

假如 M 反而是在上方的質點，同樣的計算（把 M 和 m 對調）會得出 $(M+m)\omega^2 l\cos\alpha = (m-M)g < 0$，這是不可能的。所以較重的質點確實如所假設的那樣，位於 O 的下方。

軸所受的反作用力。 桿是輕的，所以作用於它的合力為零：軸對桿所施的力為 $\mathbf{R}_M + \mathbf{R}_m$，因此桿對軸所施的力為

$$\mathbf{F} = -(\mathbf{R}_M + \mathbf{R}_m) = \Big((M-m)\,\omega^2 l\sin\alpha,\ -(M+m)\,g\Big).$$

這個力指向下方，其水平分量則在 M 的一側指向外。若 ϕ 是它與鉛垂向下方向的夾角，則代入 ω^2 的值，可得

$$\begin{aligned}\tan\phi &= \frac{(M-m)\,\omega^2 l\sin\alpha}{(M+m)\,g} = \frac{M-m}{M+m}\cdot\frac{l\sin\alpha}{g}\cdot\frac{M-m}{M+m}\cdot\frac{g}{l\cos\alpha}\\ &= \left(\frac{M-m}{M+m}\right)^2\tan\alpha.\end{aligned}$$

由於 $0 \le \phi < \frac{1}{2}\pi$，$\phi = \tan^{-1}\left[\left(\frac{M-m}{M+m}\right)^2\tan\alpha\right]$，這就是所需的結果。

驗算

- 若 $m = 0$，這個系統就是一個長度為 l 的圓錐擺（conical pendulum），而公式給出我們熟悉的 $\omega^2 = g/(l\cos\alpha)$。這時反作用力的角度為 α：桿沿着自身方向拉，就像一條繩一樣。
- 若 $M = m$，桿在 O 處平衡，所以 $\omega = 0$，反作用力是鉛垂的，大小為 $2Mg$；兩條公式都與此相符。
- 數值上，取 $M = 3$、$m = 1$、$l = 0.7$、$\alpha = 0.6$ 和 $g = 9.8$，可得 $\omega^2 \approx 8.48$，反作用力的角度約為 $9.71°$。把桿的角度所滿足的運動方程（equation of motion）由靜止於 α 開始逐步數值計算，角度一直保持在 α；而在三維空間中由兩個質點的位置計算它們的加速度，亦得出相同的角度。

常見錯誤

- 假設桿對質點所施的力沿着桿的方向。事實並非如此：剛性桿可以傳遞側向的力，這正是我們要取力矩的原因。
- 忘記了 m 的加速度同樣指向軸，因此與 M 的加速度位於相反的一側。反作用力的水平分量會部分抵消，得出的是 $M-m$ 而不是 $M+m$。
- 把較重的質點放在上方。力矩方程顯示這種情況不可能發生。

STEP 之外

桿對 O 的角動量並不平行於鉛垂的自轉軸，所以它會隨着系統旋轉而改變方向。要改變角動量就需要力偶，而這個力偶由軸提供，其方向垂直於我們剛才所用的那個方向。這就是運用慣性張量（inertia tensor）的剛體動力學（rigid-body dynamics）的起點，也解釋了為甚麼平衡不良的車輪會在軸承上搖晃。Kleppner 與 Kolenkow 的 *An Introduction to Mechanics*（力學入門教科書）對此有仔細的論述。

詳細解答

STEP II 1988，第 15 題

概率與統計 • 離散均勻分佈變量之和

> **STEP II 1988，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/88-s2-q15

關鍵思路

每份試卷的分數在 $\{0, 1, \ldots, m\}$ 上均勻分佈（uniform distribution），所以每一組可能的分數出現的機會均等，而概率（probability）就是符合條件的分數組合數目除以所有組合的總數。兩份試卷的數目只須簡單列舉便可得到；三份試卷的數目，則可先以第三份試卷的分數為條件，再把兩份試卷的數目相加而得（這就是卷積（convolution））。之後利用對稱性（symmetry）（把每份試卷的分數作 $k \mapsto m - k$ 的變換），最高的那個範圍幾乎不費工夫便可得出。

動筆之前先提醒一句：試卷上印出的三份試卷公式有一處印刷錯誤。下文會推導出正確的公式，並指出印刷版本錯在哪裏。

解答

設 X_1, X_2, X_3 為各份試卷的分數。由於各份試卷獨立評分，它們是獨立（independent）的，而且各自在 $\{0, 1, \ldots, m\}$ 上均勻分佈。因此每個有序對 (X_1, X_2) 的概率是 $(m+1)^{-2}$，每個有序三元組的概率是 $(m+1)^{-3}$。記 $N_2(n)$ 和 $N_3(n)$ 分別為分數取自 $\{0, \ldots, m\}$ 而總和為 n 的有序對和有序三元組的數目。

兩份試卷

我們需要 $X_1 = k$ 及 $X_2 = n - k$，而且兩者都在 $[0, m]$ 之內，所以

$$\max(0,\, n - m) \le k \le \min(m,\, n).$$

- 若 $0 \le n \le m$：k 由 0 取到 n，所以 $N_2(n) = n + 1$。

▶ 若 $m < n \le 2m$：k 由 $n-m$ 取到 m，所以 $N_2(n) = m - (n-m) + 1 = 2m - n + 1$。

因此

$$P(X_1 + X_2 = n) = \begin{cases} \dfrac{n+1}{(m+1)^2}, & 0 \le n \le m, \\ \dfrac{2m-n+1}{(m+1)^2}, & m < n \le 2m, \end{cases}$$

這就是題目要求的 $m < n \le 2m$ 的結果，連同 $0 \le n \le m$ 的結果。這個分佈形如「帳篷」：每一步上升一，直到 $n = m$，然後逐步下降。

三份試卷

以第三份試卷的分數為條件：若 $X_3 = k$，則需要 $X_1 + X_2 = n - k$，所以

$$N_3(n) = \sum_{k=0}^{m} N_2(n-k) = \sum_{j=n-m}^{n} N_2(j),$$

其中 $j = n - k$，而當 $j < 0$ 或 $j > 2m$ 時 $N_2(j) = 0$。

情況 $0 \le n \le m$。 只有 $j = 0, 1, \ldots, n$ 有貢獻（負的 j 給出零），而且它們全都落在帳篷上升的一段：

$$N_3(n) = \sum_{j=0}^{n} (j+1) = \frac{(n+1)(n+2)}{2}, \qquad P(X_1 + X_2 + X_3 = n) = \frac{(n+1)(n+2)}{2(m+1)^3}.$$

情況 $m < n \le 2m$。 這時 $1 \le n - m \le m < n \le 2m$，所以 $j = n-m, \ldots, n$ 這個範圍橫跨帳篷的頂點。我們在頂點處把它分成兩段：

$$N_3(n) = \sum_{j=n-m}^{m} (j+1) + \sum_{j=m+1}^{n} (2m - j + 1).$$

第一個和是把由 $n - m + 1$ 到 $m + 1$ 的整數相加：

$$\sum_{j=n-m}^{m} (j+1) = \frac{(m+1)(m+2) - (n-m)(n-m+1)}{2}.$$

在第二個和中，設 $i = 2m - j + 1$，它由 $2m - n + 1$ 取到 m：

$$\sum_{j=m+1}^{n} (2m - j + 1) = \frac{m(m+1) - (2m-n)(2m-n+1)}{2}.$$

把四個乘積展開，

$$\begin{aligned} 2N_3(n) &= (m^2 + 3m + 2) - (n^2 - 2mn + m^2 + n - m) \\ &\quad + (m^2 + m) - (4m^2 - 4mn + n^2 + 2m - n) \\ &= 6mn - 3m^2 - 2n^2 + 3m + 2. \end{aligned}$$

除以 $(m+1)^3$ 個機會均等的三元組，

$$P(X_1+X_2+X_3=n)=\frac{6mn-3m^2-2n^2+3m+2}{2(m+1)^3},\qquad m<n\le 2m.$$

想一想 這並不是題目印出的公式。究竟誰對？

印刷版本以 $-4m^2$ 代替了 $-3m^2$，又以 $(m+1)^2$ 代替了 $(m+1)^3$；兩處都是印刷錯誤。只須看一個例子便可定奪：取 $m=1$ 及 $n=2$，由 $\{0,1\}$ 取出的三個分數，在 8 種機會均等的情形中恰好有 3 種總和為 2，所以概率是 $\frac{3}{8}$。我們的公式給出 $\frac{12-3-8+3+2}{16}=\frac{3}{8}$，而印刷的公式卻給出 $\frac{12-4-8+3+2}{8}=\frac{5}{8}$。無論如何，分母必定是 $(m+1)^3$，因為這正是機會均等的結果的總數。再多驗算一次：當 $n=m$ 時，我們中間範圍的公式給出 $\frac{(m+1)(m+2)}{2(m+1)^3}$，與第一種情況吻合；這是理所當然的，因為兩個和在該處正好相接。

情況 $2m<n\le 3m$。 把每個分數 X_i 換成 $m-X_i$。新的分數同樣在 $\{0,\dots,m\}$ 上均勻分佈，而三個新分數仍然獨立，所以

$$\begin{aligned}P(X_1+X_2+X_3=n)&=P\big((m-X_1)+(m-X_2)+(m-X_3)=n\big)\\&=P(X_1+X_2+X_3=3m-n).\end{aligned}$$

當 $2m<n\le 3m$ 時，有 $0\le 3m-n<m$，所以可把第一種情況中的 n 換成 $3m-n$ 來應用：

$$P(X_1+X_2+X_3=n)=\frac{(3m-n+1)(3m-n+2)}{2(m+1)^3},\qquad 2m<n\le 3m.$$

（同樣的對稱性，以 $2m-n$ 代替 $3m-n$，可把兩份試卷在 $n\le m$ 時的結果變成 $n>m$ 時的結果。）

驗算

- 小例子：當 $m=1$ 時，三條公式對 $n=0,1,2,3$ 分別給出 $\frac{1}{8},\frac{3}{8},\frac{3}{8},\frac{1}{8}$，它們的總和是 1，並且與擲三次公平硬幣的二項分佈（binomial distribution）一致。
- 中間範圍的公式在 $n\mapsto 3m-n$ 之下對稱：在 $6mn-3m^2-2n^2$ 中把 n 換成 $3m-n$，式子保持不變，正如對稱性論證所要求的。
- 對由 1 到 8 的每一個 m 作完全列舉，五條公式全部與實際數目完全相符；例如 $m=10$ 及 $n=15$ 時，兩種方法都給出 $\frac{91}{1331}$。

常見錯誤

- 忘記 $N_2(j)$ 有兩條不同的公式，在中間範圍的求和中由頭到尾都用 $j+1$。

- 一心想湊出印刷的中間公式，於是「硬砌」代數運算。如果仔細計算的結果與題目給出的答案不符，先用一個極小的例子檢驗，不要急於懷疑自己。

STEP 之外

獨立隨機變量（random variable）之和的分佈，就是它們各自分佈的卷積，而我們的和 $\sum_k N_2(n-k)$ 正是如此。概率母函數（probability generating function）可把這個過程變成機械式的運算：數目 $N_3(n)$ 就是 $(1+t+\cdots+t^m)^3$ 中 t^n 的係數。隨着試卷的數目愈來愈多，分佈的形狀會由帳篷變成平滑的隆起，並趨近正態曲線（這就是中心極限定理（central limit theorem））。Grimmett 與 Welsh 合著的 *Probability: An Introduction*（概率論入門）以合適的程度講解母函數和中心極限定理。

詳細解答

STEP II 1988，第 16 題

概率與統計 • 正態分佈

STEP II 1988，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/88-s2-q16

關鍵思路

關於兩根的一切，都由單一個隨機變量（random variable）B 決定，所以每一個關於 X_1 和 X_2 的問題，都應該轉化為「B 落在哪個區間」的問題。判別式（discriminant）告訴我們兩根何時為實數，而韋達定理（Vieta's formulas）（兩根之和為 $-2B$，兩根之積為 1）令 (i) 和 (ii) 的條件很容易轉換。之後，每個答案都是標準正態變量（standard normal variable）的條件概率（conditional probability）或條件期望值（conditional expectation），而求期望值只需計算一個簡短的積分。

解答

全題設 $B \sim \mathrm{N}(0,1)$，其密度函數（density）為 $\phi(z) = \frac{1}{\sqrt{2\pi}}\mathrm{e}^{-z^2/2}$，分佈函數（distribution function）為 Φ。設 R 為「兩根皆為實數」這個事件。

實根

$X^2 + 2BX + 1$ 的判別式為 $4B^2 - 4$，所以兩根為實數當且僅當 $B^2 \geq 1$，即 $|B| \geq 1$。由正態分佈（normal distribution）的對稱性，

$$\mathrm{P}(R) = \mathrm{P}(|B| \geq 1) = 2\big(1 - \Phi(1)\big) = 2(1 - 0.84134) = 0.31731\ldots \approx 0.317.$$

（是否包括等根的情況 $|B| = 1$ 並無分別，因為它的概率為零。）

當兩根為實數時，它們是

$$X_{1,2} = -B \pm \sqrt{B^2 - 1}, \qquad X_1 + X_2 = -2B, \qquad X_1 X_2 = 1.$$

第 (i) 部分

想一想 怎樣把「兩根都大於 $\frac{1}{5}$」轉化為關於 B 的條件？

首先，兩根必須都是正數。兩根之積為 $1 > 0$，所以它們同號，而這個符號就是兩根之和 $-2B$ 的符號。因此我們需要 $B < 0$，連同 R 即表示 $B \le -1$。記 $b = -B \ge 1$；這時兩根為 $b \pm \sqrt{b^2-1}$，都是正數。

由於兩根之積為 1，較小的根是較大的根的倒數：

$$b - \sqrt{b^2-1} = \frac{1}{b+\sqrt{b^2-1}}.$$

所以較小的根大於 $\frac{1}{5}$ 當且僅當較大的根小於 5，而這時較大的根自然也大於 $\frac{1}{5}$。因此條件就是

$$\sqrt{b^2-1} < 5 - b.$$

這需要 $b < 5$，而當 $1 \le b < 5$ 時兩邊皆為非負數，所以可以把兩邊平方而不改變不等式：

$$b^2 - 1 < 25 - 10b + b^2 \iff 10b < 26 \iff b < 2.6.$$

（$b \ge 5$ 的值無論如何都不符合，因為 $2.6 < 5$。）因此，在兩根為實數的條件下，兩根都大於 $\frac{1}{5}$ 當且僅當 $-2.6 < B \le -1$，於是

$$\begin{aligned}\mathrm{P}\left(X_1, X_2 > \tfrac{1}{5} \,\middle|\, R\right) &= \frac{\mathrm{P}(-2.6 < B \le -1)}{\mathrm{P}(|B| \ge 1)} = \frac{\Phi(2.6) - \Phi(1)}{2\big(1-\Phi(1)\big)} \\ &= \frac{0.99534 - 0.84134}{0.31731} = 0.4853\ldots \approx 0.485.\end{aligned}$$

第 (ii) 部分

由於 $|X_1 + X_2| = 2|B|$，我們要求的是 $\mathrm{E}\big(2|B| \,\big|\, |B| \ge 1\big)$。要計算給定某事件下的條件期望值，我們在該事件上積分，再除以該事件的概率：

$$\mathrm{E}\big(|B| \,\big|\, |B| \ge 1\big) = \frac{1}{\mathrm{P}(|B| \ge 1)} \int_{|z|\ge 1} |z|\, \phi(z)\, \mathrm{d}z = \frac{2\int_1^\infty z\,\phi(z)\,\mathrm{d}z}{2\big(1-\Phi(1)\big)},$$

這裏用到了 ϕ 的對稱性。這個積分很容易計算，因為 $z\,\phi(z)$ 正是 $\phi(z)$ 的導數的負值：

$$\int_1^\infty z\,\phi(z)\,\mathrm{d}z = \Big[-\phi(z)\Big]_1^\infty = \phi(1) = \frac{\mathrm{e}^{-1/2}}{\sqrt{2\pi}} = 0.24197\ldots$$

因此

$$\mathrm{E}\big(|X_1 + X_2| \,\big|\, R\big) = \frac{2\phi(1)}{1-\Phi(1)} = \frac{2 \times 0.24197}{0.15866} = 3.050\ldots \approx 3.05.$$

驗算

- 試取 $B = -2.6$：方程為 $X^2 - 5.2X + 1 = 0$，兩根恰好是 5 和 $\frac{1}{5}$，正是 (i) 的邊界情況。
- (ii) 的答案必須大於 2，因為在 R 上 $2|B| \geq 2$。答案確實如此，而且合理：正態分佈在 1 以外的尾部很薄，所以 $|B|$ 通常只比 1 大一點（它的條件平均值約為 1.53）。
- 以 2×10^6 個標準正態的 B 值進行模擬，得出 0.3177、0.4855 和 3.049，與 0.317、0.485 和 3.05 吻合。

常見錯誤

- 忘記取條件：在 (i) 和 (ii) 中，我們必須除以 $\mathrm{P}(R) = 0.317$。無條件概率 $\Phi(2.6) - \Phi(1) = 0.154$ 並不是答案。
- 在 (i) 中把 $B \geq 1$ 也計算在內。這時兩根都是負數，所以沒有一個大於 $\frac{1}{5}$。
- 把 $\mathrm{E}|X_1 + X_2|$ 當作 $|\mathrm{E}(X_1 + X_2)|$ 來計算，而後者由對稱性可知等於 0。

STEP 之外

$\phi(1)/(1 - \Phi(1))$ 這個量是截尾（truncated）正態分佈的平均值，而比值 $\phi(a)/(1 - \Phi(a))$ 稱為逆米爾斯比率（inverse Mills ratio）；它出現於統計學（例如設限數據（censored data）的模型），也出現於可靠性理論（reliability theory），在那裏它是正態分佈的風險率（hazard rate）。關於隨機多項式（random polynomials）的問題，例如係數隨機的多項式平均有多少個實根，構成了概率論中一個豐富得出人意表的領域。關於條件期望值和連續分佈，可參閱 Grimmett 與 Welsh 的 *Probability: An Introduction*（概率論入門），或 Blitzstein 與 Hwang 的 *Introduction to Probability*（概率論導引）。

詳細解答

STEP III 1988，第 1 題

純數學 • 曲線描繪與界限

STEP III 1988，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/88-s3-q1

關鍵思路

繪畫草圖時，因子 $x^2\mathrm{e}^{-x}$ 永不為負，所以 y 的正負完全由 $1+x$ 決定。有了這一點，再加上一次求導，以及曲線在兩端和 $x=-1$ 附近的表現，整條曲線的形狀便清楚了。至於積分，最乾淨利落的做法是把 x^2 除以 $1+x$：這樣被積函數（integrand）的大部分積分後恰好互相抵消，餘下的部分則很容易夾在兩個簡單的界之間。

解答

圖像

設 $f(x)=\dfrac{x^2\mathrm{e}^{-x}}{1+x}$，定義於所有 $x\neq -1$。

正負與零點。由於 $x^2\geq 0$ 而 $\mathrm{e}^{-x}>0$，當 $x<-1$ 時 $f(x)<0$，當 $x>-1$ 時 $f(x)\geq 0$，而且只有在 $x=0$ 時才有 $f(x)=0$。因子 x^2 表示曲線在原點只是與 x 軸相切，而不是穿過它。

漸近線（asymptote）與兩端的表現。

- 當 $x\to -1$ 時，分子趨向 $\mathrm{e}>0$，分母趨向 0，所以 $x=-1$ 是一條鉛垂漸近線；當 $x\to -1^-$ 時 $f\to -\infty$，當 $x\to -1^+$ 時 $f\to +\infty$。
- 當 $x\to +\infty$ 時，$f(x)\approx x\mathrm{e}^{-x}\to 0^+$，因為指數函數的增長勝過冪函數。所以 $y=0$ 是右方的漸近線，曲線從上方趨近它。
- 當 $x\to -\infty$ 時，$\dfrac{x^2}{1+x}\approx x\to -\infty$ 而 $\mathrm{e}^{-x}\to +\infty$，所以 $f(x)\to -\infty$，而且下降得非常快。

駐點（stationary point）。根據商法則（quotient rule），

$$f'(x) = \frac{(2x - x^2)\mathrm{e}^{-x}(1+x) - x^2\mathrm{e}^{-x}}{(1+x)^2} = \frac{\mathrm{e}^{-x}\left(2x + x^2 - x^3 - x^2\right)}{(1+x)^2}$$
$$= \frac{x\left(2 - x^2\right)\mathrm{e}^{-x}}{(1+x)^2}.$$

所以在 $x = 0$ 和 $x = \pm\sqrt{2}$ 時 $f'(x) = 0$。因子 e^{-x} 和 $(1+x)^2$ 都是正數，所以 f' 的正負與 $x(2 - x^2)$ 的正負相同：

區間	$x < -\sqrt{2}$	$-\sqrt{2} < x < -1$	$-1 < x < 0$	$0 < x < \sqrt{2}$	$x > \sqrt{2}$
$f'(x)$	$+$	$-$	$-$	$+$	$-$

由此得出：

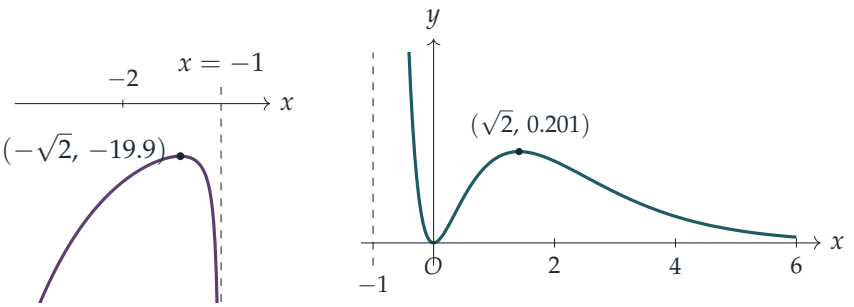
- 在 $x = \sqrt{2}$ 處有一個局部極大點（local maximum），該處 $y = \dfrac{2\mathrm{e}^{-\sqrt{2}}}{1+\sqrt{2}} = 2(\sqrt{2} - 1)\mathrm{e}^{-\sqrt{2}} \approx 0.201$；
- 在原點有一個局部極小點（local minimum）；
- 在左邊分支上的 $x = -\sqrt{2}$ 處有一個局部極大點，該處 $y = \dfrac{2\mathrm{e}^{\sqrt{2}}}{1-\sqrt{2}} = -2(\sqrt{2} + 1)\mathrm{e}^{\sqrt{2}} \approx -19.9$。

想一想 一條完全位於橫軸下方的分支，怎可能有極大點？

在 $x < -1$ 的範圍內，曲線隨着 $x \to -\infty$ 從 $-\infty$ 升上來，到了漸近線 $x = -1$ 又再跌回 $-\infty$，所以它必定在中間某處轉向。最高點位於 $x = -\sqrt{2}$，而這一點仍然遠在橫軸之下。這條分支是一個懸在橫軸下方的「小丘」，丘頂的高度約為 -19.9。

在右邊，曲線從緊貼 $x = -1$ 右側的 $+\infty$ 下降，到原點時降至 0，然後上升至 $x = \sqrt{2}$ 處那個小小的極大點，其後衰減至 0。曲線在原點附近向上凹，在極大點處向下凹，所以在 $x = \sqrt{2}$ 的兩側各有一個拐點（point of inflection）（數值上約在 $x \approx 0.35$ 和 $x \approx 2.5$ 附近）。繪畫草圖時，我們不需要知道它們的準確位置。

兩條分支在縱向上的尺度相差極大，所以我們把它們分開繪畫。在左邊，y 一直下降至約 -70。在右邊，我們把 y 軸的比例放大，讓 0.2 這個極大值清楚可見。



積分

設 $I=\int_0^\infty \frac{x^2\mathrm{e}^{-x}}{1+x}\,\mathrm{d}x$。把 x^2 除以 $1+x$，得

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

現在 $\int_0^\infty x\mathrm{e}^{-x}\,\mathrm{d}x=1$（取 $u=x$ 作分部積分（integration by parts)），而 $\int_0^\infty \mathrm{e}^{-x}\,\mathrm{d}x=1$，所以首兩項的貢獻是 $1-1=0$，於是

$$I=\int_0^\infty \frac{\mathrm{e}^{-x}}{1+x}\,\mathrm{d}x.$$

（三個積分全部收斂（converge），所以把積分拆開是合理的。）當 $x>0$ 時，我們有 $0<\frac{1}{1+x}<1$，所以

$$0<\frac{\mathrm{e}^{-x}}{1+x}<\mathrm{e}^{-x}\qquad(x>0).$$

這些不等式在整個區間上都是嚴格的，而且各函數均為連續，所以積分後不等式仍然嚴格成立：

$$0<I<\int_0^\infty \mathrm{e}^{-x}\,\mathrm{d}x=1.$$

所以 I 介乎 0 與 1 之間，正是題目所求。

另一種做法。我們可以略去除法這一步。當 $x>0$ 時，$\frac{x^2}{1+x}<\frac{x^2}{x}=x$，所以 $0<f(x)<x\mathrm{e}^{-x}$，從而 $0<I<\int_0^\infty x\mathrm{e}^{-x}\,\mathrm{d}x=1$。不過那個除法仍然值得認識，因為它清楚顯示了 I 究竟從何而來。

驗算

- 在一千個隨機選取的點上以數值方法對 f 求導，結果與 $f'(x)=x(2-x^2)\mathrm{e}^{-x}/(1+x)^2$ 吻合。

- 在 $[0, 60]$ 上使用辛普森法則（Simpson's rule）得出 $I \approx 0.59635$，而 $\int_0^\infty \frac{\mathrm{e}^{-x}}{1+x}\,\mathrm{d}x$ 亦給出相同的數值，穩穩地落在 0 與 1 之間。
- 從草圖作粗略估算：在 $x > 0$ 的範圍內，曲線是一個高度約為 0.2 的小丘，到了 $x = 6$ 已大致消退。以此為高和底的三角形，面積約為 0.6，與 0.596 相符。

常見錯誤

- 遺漏了 $x = -\sqrt{2}$ 處的駐點，因為它位於「另一條」分支上；又或者因為該處 y 是負數，便把它說成極小點。
- 說曲線在原點穿過橫軸。二重因子 x^2 令曲線只是與橫軸相切。
- 單憑 $\mathrm{e}^{-x} \leq 1$ 便聲稱 $I < 1$。這個界只會得出 $\int_0^\infty \frac{x^2}{1+x}\,\mathrm{d}x$，而這個積分是發散（diverge）的。

STEP 之外

I 的值是 $\mathrm{e}\,E_1(1) \approx 0.5963$，其中 $E_1(t) = \int_t^\infty \mathrm{e}^{-u}u^{-1}\,\mathrm{d}u$ 是指數積分（exponential integral）。這個數稱為 Gompertz 常數（Gompertz constant）。它不能用初等函數（elementary functions）表示，這正是題目只要求找出上下界的原因。對於無法求出確切值的積分，拿它與能夠求值的積分作比較，從而作出估計，是分析學（analysis）的一項核心技巧（即瑕積分（improper integral）的「比較判別法」（comparison test））。Abbott 的 *Understanding Analysis*（一本分析學入門教科書）對此有細緻的闡述。

詳細解答

STEP III 1988，第 2 題

純數學 • 線性遞推關係

STEP III 1988，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/88-s3-q2

關鍵思路

常係數線性遞歸關係（linear recurrence with constant coefficients）的表現就像一條二次方程：冪 x^n 是它的解，當且僅當 x 是輔助方程（auxiliary equation）的根。由於只要知道首兩項，遞歸關係便決定了其後每一項，所以只需配對 u_0 和 u_1，就能得到唯一的解。最後一部分的係數是可變的，但因子 $(n+1)$、$(n+2)$、$(n+3)$ 提示我們把全式除以它們的乘積，然後用一個簡單的代換（substitution），便可把它化為第一種的常係數遞歸關係。

解答

（試卷上把遞歸關係的首項係數印成 α。這顯然是 a 的誤植，應與二次式 ax^2+bx+c 一致，我們就按此理解。）由於這條二次方程有兩個根，它確實是二次的，所以 $a \neq 0$。

公式滿足遞歸關係

設 $u_n = A\alpha^n + B\beta^n$，其中 A、B 為任意常數。那麼

$$\begin{aligned} au_{n+2} + bu_{n+1} + cu_n &= A\alpha^n\left(a\alpha^2 + b\alpha + c\right) + B\beta^n\left(a\beta^2 + b\beta + c\right) \\ &= A\alpha^n \cdot 0 + B\beta^n \cdot 0 = 0, \end{aligned}$$

因為 α 和 β 是 $ax^2+bx+c=0$ 的根。所以遞歸關係對每個 $n \geq 0$ 都成立。

取題目所給的特定 A 和 B 時，公式亦能重現初始值。當 $n=0$，

$$A + B = \frac{(u_1 - \beta u_0) - (u_1 - \alpha u_0)}{\alpha - \beta} = \frac{(\alpha - \beta)u_0}{\alpha - \beta} = u_0,$$

而當 $n = 1$，

$$A\alpha + B\beta = \frac{\alpha u_1 - \alpha\beta u_0 - \beta u_1 + \alpha\beta u_0}{\alpha - \beta} = \frac{(\alpha - \beta)u_1}{\alpha - \beta} = u_1.$$

由於兩根相異，所以可以除以 $\alpha - \beta$。

它是唯一的解

想一想 我們究竟是對甚麼作歸納？為甚麼每次需要兩項？

假設 $u_0, u_1, u_2, \ldots$ 是任何一個滿足該遞歸關係的數列（sequence），並設 A、B 如上由它的 u_0 和 u_1 構造而成。記 $w_n = A\alpha^n + B\beta^n$。我們對 n 用歸納法（induction）證明以下命題：

$$P(n): \quad u_n = w_n \text{ 及 } u_{n+1} = w_{n+1}.$$

每個新的項都取決於它之前的兩項，所以命題必須同時包含連續兩項。

基礎步驟（*base case*）。$P(0)$ 是說 $u_0 = w_0$ 及 $u_1 = w_1$，這正是我們在上面驗證過的。

歸納步驟（*inductive step*）。假設 $P(n)$ 成立。(u_n) 和 (w_n) 都滿足遞歸關係，而且 $a \neq 0$，所以

$$u_{n+2} = -\frac{bu_{n+1} + cu_n}{a} = -\frac{bw_{n+1} + cw_n}{a} = w_{n+2}.$$

連同 $u_{n+1} = w_{n+1}$，這就是 $P(n+1)$。

所以 $P(n)$ 對所有 $n \geq 0$ 成立，即對每個 n 都有 $u_n = A\alpha^n + B\beta^n$。遞歸關係的任何解都具有這個形式，其中 A 和 B 由 u_0 和 u_1 決定。這條公式給出的就是唯一的解。

數列 v_n

每個係數都是 $(n+1)$、$(n+2)$、$(n+3)$ 這三個因子中其中兩個的乘積，而各自缺少的因子都不相同。除以 $(n+1)(n+2)(n+3)$（當 $n \geq 0$ 時它永不為零），得

$$8\,\frac{v_{n+2}}{n+3} - 2\,\frac{v_{n+1}}{n+2} - \frac{v_n}{n+1} = 0.$$

每個分數都具有 $v_k/(k+1)$ 的形式。因此設

$$w_n = \frac{v_n}{n+1},$$

遞歸關係便變成

$$8w_{n+2} - 2w_{n+1} - w_n = 0, \qquad w_0 = 0, \quad w_1 = \tfrac{1}{2}.$$

輔助方程是 $8x^2 - 2x - 1 = 0$，即 $(2x-1)(4x+1) = 0$。它的根 $\alpha = \frac{1}{2}$ 和 $\beta = -\frac{1}{4}$ 是相異實根，所以第一部分的結果適用，其中 $\alpha - \beta = \frac{3}{4}$：

$$A = \frac{w_1 - \beta w_0}{\alpha - \beta} = \frac{1/2}{3/4} = \frac{2}{3}, \qquad B = \frac{w_1 - \alpha w_0}{\beta - \alpha} = \frac{1/2}{-3/4} = -\frac{2}{3}.$$

因此 $w_n = \frac{2}{3}\left(\left(\frac{1}{2}\right)^n - \left(-\frac{1}{4}\right)^n\right)$。第一部分亦告訴我們，這是 w_n 唯一的可能，因而也是 v_n 唯一的可能。乘回 $n+1$，得

$$\boxed{v_n = \frac{2(n+1)}{3}\left(\frac{1}{2^n} - \left(-\frac{1}{4}\right)^n\right)} \qquad (n = 0, 1, 2, \ldots).$$

驗算

公式給出 $v_0 = \frac{2}{3}(1-1) = 0$ 和 $v_1 = \frac{4}{3}\left(\frac{1}{2} + \frac{1}{4}\right) = 1$，符合要求。以 $n = 0$ 和 $n = 1$ 用人手計算原來的遞歸關係，得 $16v_2 = 6v_1 + 6v_0$，所以 $v_2 = \frac{3}{8}$；然後 $48v_3 = 16v_2 + 12v_1$，所以 $v_3 = \frac{3}{8}$。公式的結果與此吻合：$v_2 = 2\left(\frac{1}{4} - \frac{1}{16}\right) = \frac{3}{8}$ 及 $v_3 = \frac{8}{3}\left(\frac{1}{8} + \frac{1}{64}\right) = \frac{3}{8}$。我們以有理數運算，對 n 直至 41 精確地核對過公式與遞歸關係相符，亦以數百組隨機的二次方程和初始值核對過第一部分。

常見錯誤

- 每次只對一項作歸納。單憑 $u_n = w_n$，無法推出任何關於 u_{n+1} 的結論：歸納假設必須涵蓋連續兩項。
- 忘記說明為何可以除以 a（因為方程有兩個根，所以它不為零）和除以 $\alpha - \beta$（兩根相異）。
- 直接把第一部分應用於 v_n。它的係數取決於 n，所以第一部分適用的，是經重新縮放的數列 $w_n = v_n/(n+1)$。

STEP 之外

二階線性遞歸關係的解構成一個二維向量空間（vector space），而 α^n、β^n 是它的一組基（basis）。這與微分方程（differential equation）$ay'' + by' + cy = 0$ 及其解 $e^{\alpha x}$ 和 $e^{\beta x}$ 的情況完全平行。把遞歸關係寫成 $\mathbf{u}_{n+1} = M\mathbf{u}_n$，其中 M 是一個 2×2 矩陣，便可看出 α 和 β 是 M 的特徵值（eigenvalue），而重根則對應於不能對角化（diagonalise）的矩陣。Strang 的 *Introduction to Linear Algebra*（線性代數入門）正是這樣透過特徵值來處理差分方程（difference equation），而 Boyce 與 DiPrima 的 *Elementary Differential Equations*（初等微分方程）則為微分方程建立了平行的理論。

詳細解答

STEP III 1988，第 3 題

純數學 • 複數軌跡

STEP III 1988，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/88-s3-q3

關鍵思路

映射 $w = (z+\mathrm{i})/(z-\mathrm{i})$ 可以倒過來用：解出 z 得 $z = \mathrm{i}(w+1)/(w-1)$，於是任何關於 z 的條件都變成關於 w 的條件。至於那個圓，則改用參數表示（parametrisation）來處理，因為這時 $z-\mathrm{i}$ 就只是 $2\mathrm{e}^{\mathrm{i}\theta}$。對於實數和虛數兩種情況，可從模（modulus）和輻角（argument）着眼：實數 z 與 i 和 $-\mathrm{i}$ 等距。

解答

參數形式

曲線 $|z-\mathrm{i}| = 2$ 是以 i 為圓心、半徑為 2 的圓，所以

$$z = \mathrm{i} + 2\mathrm{e}^{\mathrm{i}\theta}, \qquad -\pi < \theta \leq \pi,$$

或者，設 $z = x + \mathrm{i}y$，則 $x = 2\cos\theta$ 及 $y = 1 + 2\sin\theta$。

我們還會用到逆映射（inverse map）。由 $w(z-\mathrm{i}) = z + \mathrm{i}$ 得 $z(w-1) = \mathrm{i}(w+1)$，所以當 $w \neq 1$ 時

$$z = \frac{\mathrm{i}(w+1)}{w-1}.$$

沒有任何 z 會給出 $w = 1$，因為那需要 $z+\mathrm{i} = z-\mathrm{i}$。每個 $w \neq 1$ 都來自恰好一個 $z \neq \mathrm{i}$。

第 (i) 部分

代入參數形式，$z + \mathrm{i} = 2\mathrm{i} + 2\mathrm{e}^{\mathrm{i}\theta}$ 及 $z - \mathrm{i} = 2\mathrm{e}^{\mathrm{i}\theta}$，所以

$$w = \frac{2\mathrm{i} + 2\mathrm{e}^{\mathrm{i}\theta}}{2\mathrm{e}^{\mathrm{i}\theta}} = 1 + \mathrm{i}\,\mathrm{e}^{-\mathrm{i}\theta} = 1 + \mathrm{e}^{\mathrm{i}(\pi/2-\theta)}.$$

當 θ 繞行一周時，$\pi/2-\theta$ 也恰好取遍每個角度一次。所以 w 描出整個圓

$$|w-1|=1,$$

其圓心為 1，半徑為 1。例如，$z=3\mathrm{i}$ 給出 $w=2$，而 $z=-\mathrm{i}$ 給出 $w=0$。點 $w=1$ 不在這個圓上，這與它永遠取不到的事實一致。

第 (ii) 部分

若 $z=x$ 是實數，它與 i 的距離等於它與 $-\mathrm{i}$ 的距離，所以 $|x+\mathrm{i}|=|x-\mathrm{i}|$，而且

$$|w|=\frac{|x+\mathrm{i}|}{|x-\mathrm{i}|}=1.$$

具體地說，把分子和分母同乘以 $x+\mathrm{i}$，

$$w=\frac{(x+\mathrm{i})^2}{x^2+1}=\frac{x^2-1}{x^2+1}+\mathrm{i}\,\frac{2x}{x^2+1}.$$

想一想 單位圓上的每一點都能取到嗎？

不，還差一點。任取滿足 $|w|=1$ 且 $w\neq 1$ 的 w，並寫成 $w=\mathrm{e}^{\mathrm{i}\phi}$。那麼

$$z=\mathrm{i}\,\frac{\mathrm{e}^{\mathrm{i}\phi}+1}{\mathrm{e}^{\mathrm{i}\phi}-1}=\mathrm{i}\,\frac{2\cos(\phi/2)}{2\mathrm{i}\sin(\phi/2)}=\cot\tfrac{\phi}{2},$$

這是實數。所以每個這樣的點都能取到，但 $w=1$ 永遠取不到：它是 $x\to\pm\infty$ 時的極限。所求軌跡 (locus) 是單位圓 $|w|=1$ 除去 $w=1$ 這一點。例如，$x=0$ 給出 $w=-1$，而 $x=1$ 給出 $w=\mathrm{i}$。

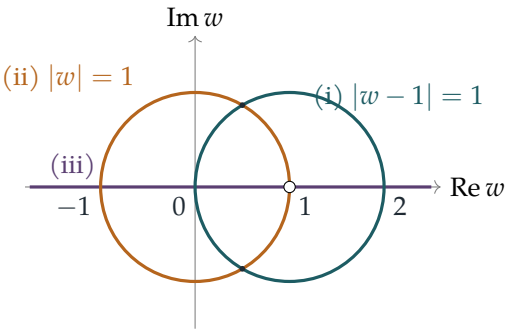
第 (iii) 部分

若 $z=\mathrm{i}y$，其中 y 為實數且 $y\neq 1$，則

$$w=\frac{\mathrm{i}(y+1)}{\mathrm{i}(y-1)}=\frac{y+1}{y-1}=1+\frac{2}{y-1},$$

這是實數。當 y 取遍 1 以外的所有實數時，分數 $2/(y-1)$ 取遍每個非零實數值，所以 w 取遍 1 以外的每個實數值。(反過來，若 $w=u$ 是實數且 $u\neq 1$，則 $z=\mathrm{i}(u+1)/(u-1)$ 是虛數。) 所求軌跡是實軸除去 $w=1$ 這一點。這裏 $z=0$ 給出 $w=-1$；如果你把「虛數」理解為不包括 $z=0$，那麼 $w=-1$ 也要除去。

略圖



藍綠色的圓是 (i) 的軌跡，琥珀色的圓（除去點 1）是 (ii) 的軌跡，紫色的軸（除去點 1）是 (iii) 的軌跡。空心圓點標示 $w = 1$，這一點永遠取不到。

驗算

- (i) 中的圓與實軸相交於 $z = \pm\sqrt{3}$。這兩點應該映射到同時位於藍綠色圓和琥珀色圓上的點，而事實上 $w(\sqrt{3}) = (\sqrt{3}+\mathrm{i})^2/4 = \frac{1}{2} + \frac{\sqrt{3}}{2}\mathrm{i} \approx 0.5 + 0.866\mathrm{i}$，它確實既在 $|w| = 1$ 上，也在 $|w-1| = 1$ 上。
- (i) 中的圓與虛軸相交於 $3\mathrm{i}$ 和 $-\mathrm{i}$，它們分別映射到 2 和 0：正好是藍綠色圓與實軸相交之處。
- 在每條軌跡上各取 10^4 個隨機點作數值檢驗，結果確認了 $|w-1| = 1$、$|w| = 1$ 和 $\operatorname{Im} w = 0$，並顯示逆映射能還原出 z。

常見錯誤

- 忘記 (ii) 和 (iii) 中缺少了 $w = 1$。它對應於「$z = \infty$」。
- 在 (i) 中展開成 x 和 y，結果迷失在代數運算之中。參數表示令 $z - \mathrm{i}$ 變成單一個指數式，餘下的只需一行。

STEP 之外

映射 $z \mapsto (z+\mathrm{i})/(z-\mathrm{i})$ 是一個莫比烏斯變換（Möbius transformation）。這類映射總是把圓和直線映射成圓和直線，這正是你在三個部分中所發現的。要理解它們，最好放在黎曼球面（Riemann sphere）上看：在那裏加入一個「無窮遠點」（point at infinity），便填補了 $w = 1$ 處的缺口。它的近親凱萊變換（Cayley transform）$z \mapsto (z-\mathrm{i})/(z+\mathrm{i})$

把上半平面（upper half-plane）映射到單位圓盤（unit disc）上。Needham 的 *Visual Complex Analysis*（《複分析：可視化方法》）對這一切有十分優美的幾何處理。

詳細解答

STEP III 1988，第 4 題

純數學 • 弧長的最優化

STEP III 1988，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/88-s3-q4

關鍵思路

把道路的總長度寫成函數 $L(b)$，然後求導（differentiate）。位於鉛垂平面內的兩段路必須貼着地面走，所以它們的長度是拋物線的弧長（arc length）；半圓的長度是 πb。把弧長積分對其下限求導，根本不需要做任何積分，而 $L'(b)$ 的正負號就決定了一切。

（題目印出的橫截面公式左邊是 $x=$；當然應該是 $z=$。）

解答

建立長度的表達式

設 $h(b)=(a^2-b^2)/(2a)$，即山丘在距離對稱軸 b 處的高度。道路必須鋪在地面上，所以在 (x,z) 平面內由 $(-R,0)$ 到 $(-b,h(b))$ 的路徑別無選擇：它先沿平地由 $x=-R$ 走到 $x=-a$，再沿拋物線由 $x=-a$ 上行至 $x=-b$。第 (iii) 段是它的鏡像。

在山丘上 $z=(a^2-x^2)/(2a)$，所以 $\dfrac{\mathrm{d}z}{\mathrm{d}x}=-\dfrac{x}{a}$，而由 $x=b$ 到 $x=a$ 的弧長是

$$\int_b^a \sqrt{1+\frac{x^2}{a^2}}\,\mathrm{d}x.$$

半圓 (ii) 是水平的，位於高度 $h(b)$，半徑為 b（它是山丘上通過 $(\pm b,h(b))$ 的緯圈），所以長度是 πb。因此

$$L(b)=2(R-a)+2\int_b^a \sqrt{1+\frac{x^2}{a^2}}\,\mathrm{d}x+\pi b,\qquad 0\le b\le a.$$

這個式子也涵蓋了 $b=0$（沒有半圓；道路直接越過山頂）和 $b=a$（道路繞着山腳走）兩種情況。

求最小值

根據微積分基本定理（fundamental theorem of calculus），對下限求導，所得的是被積函數（integrand）在該處的值再加上負號：

$$L'(b) = \pi - 2\sqrt{1 + \frac{b^2}{a^2}}.$$

> ? **想一想** 駐點（stationary point）會在哪裏？它在容許的範圍內嗎？

$L'(b) = 0$ 需要 $1 + b^2/a^2 = \pi^2/4$，即 $b = \frac{a}{2}\sqrt{\pi^2 - 4} \approx 1.21a$。這比 a 大，所以落在容許範圍 $0 \le b \le a$ 之外。事實上，當 $0 \le b \le a$ 時，我們有

$$2\sqrt{1 + \frac{b^2}{a^2}} \le 2\sqrt{2} \approx 2.83 < \pi,$$

所以在整個區間上 $L'(b) > 0$。L 嚴格遞增（strictly increasing），其最小值出現在左端點 $b = 0$。

建議：取 $b = 0$。道路應該從山丘的一邊直上，再從另一邊直下，不必繞半圓的彎路。

省下多少路程

順帶一提，原函數（antiderivative）是

$$\int \sqrt{1 + \frac{x^2}{a^2}}\,\mathrm{d}x = \frac{x}{2}\sqrt{1 + \frac{x^2}{a^2}} + \frac{a}{2}\sinh^{-1}\frac{x}{a} + C,$$

（代入 $x = a\sinh t$），所以

$$L(0) = 2(R - a) + a\big(\sqrt{2} + \ln(1 + \sqrt{2})\big) \approx 2(R - a) + 2.296a,$$

而繞山腳走則得 $L(a) = 2(R - a) + \pi a \approx 2(R - a) + 3.142a$。越過山頂大約可省下 $0.85a$。

驗算

- 答案與 R 無關，這是理所當然的：不論 b 是多少，平地部分的總長度都是 $2(R - a)$。
- 粗略檢查一下是否合理：山丘只有 $a/2$ 高、$2a$ 寬，所以越過它不可能比直線距離 $2a$ 多出很多；而繞山腳的半圓長 πa。果然 $2.296a$ 介乎 $2a$ 與 πa 之間。
- 數值上，取 $a = 1$ 和 $R = 3$，在 $[0, 1]$ 的 201 個點上以數值積分（quadrature）計算 $L(b)$，最小值出現在 $b = 0$；而 $L'(a) = \pi - 2\sqrt{2} \approx 0.313$ 仍然是正數。

常見錯誤

- 令 $L'(b) = 0$ 並寫下 $b = \frac{a}{2}\sqrt{\pi^2 - 4}$，卻沒有留意到它比 a 大。
- 第 (i) 段和第 (iii) 段用了直線（弦）的長度：那樣的路會穿山而過。
- 忘記了閉區間（closed interval）上的最小值可以出現在端點。

STEP 之外

在曲面上尋找最短路徑，就是測地線（geodesics）的問題，在變分法（calculus of variations）或微分幾何（differential geometry）的入門課程中會遇到。在旋轉曲面（surface of revolution）上，克萊羅關係式（Clairaut's relation）簡潔地刻畫了測地線；國王的這一族道路，只是該問題中一個小小的單參數切片。這裏用到的兩個分析學事實——對積分的上下限求導，以及閉區間上的連續函數（continuous function）必定取得最小值——在 Abbott 的 *Understanding Analysis* 中有嚴謹的證明。

詳細解答

STEP III 1988，第 5 題

純數學 • 遞推關係與貸款

STEP III 1988，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/88-s3-q5

關鍵思路

先寫下債務在一天之內的變化。每天債務先乘以 $1+k$（利息），然後加上（挖掘成本）或減去（收入）一個固定金額。形如 $S_n = (1+k)S_{n-1} +$ 常數 的遞推關係（recurrence）逐層展開後，就是一個等比級數（geometric series）。最後一部分則歸結為一個比較：隧道通車那一刻，每日收入是否多於債務的利息？

解答

挖掘隧道期間的債務

利率為 $100k\%$，即 £S 的債務每天的利息是 £kS。設 S_n 為第 n 天結束時的債務（以英鎊計），其中 $S_0 = 0$。在第 $n \le T$ 天結束時，公司要借入 c 支付挖掘費用，另借入 kS_{n-1} 支付利息，所以

$$S_n = S_{n-1} + kS_{n-1} + c = (1+k)S_{n-1} + c.$$

由 $S_0 = 0$ 開始，反覆運用這個關係，

$$\begin{aligned} S_n &= c + (1+k)S_{n-1} \\ &= c + c(1+k) + (1+k)^2 S_{n-2} \\ &= \cdots = c[1 + (1+k) + (1+k)^2 + \cdots + (1+k)^{n-1}]. \end{aligned}$$

這是一個共有 n 項、公比（common ratio）為 $1+k \neq 1$（因為 $k > 0$）的等比級數，所以

$$S_n = \frac{c[(1+k)^n - 1]}{k}, \qquad n \le T,$$

正是所求的結果。（為嚴謹起見，可用歸納法（induction）驗證：公式給出 $S_0 = 0$；若公式對 S_{n-1} 成立，則 $(1+k)S_{n-1} + c = \frac{c}{k}[(1+k)^n - (1+k) + k]$，這正是 S_n 的公式。）

隧道通車後的債務

由第 $T+1$ 天起，不再需要挖掘。每天債務增加利息 kS，而公司還款 d，所以只要債務仍然是正數，就有

$$S_{T+j} = (1+k)S_{T+j-1} - d.$$

這與之前的遞推關係相同，只是 c 換成了 $-d$，而且現在是由 S_T 而不是由 0 開始。展開 m 次，

$$\begin{aligned} S_{T+m} &= (1+k)^m S_T - d[1 + (1+k) + \cdots + (1+k)^{m-1}] \\ &= (1+k)^m S_T - \frac{d[(1+k)^m - 1]}{k}. \end{aligned}$$

代入第一部分求得的 S_T，

$$S_{T+m} = \frac{c[(1+k)^T - 1](1+k)^m - d[(1+k)^m - 1]}{k}.$$

我們需要 $S_{T+m} > 0$，因為這個遞推關係只在債務尚未還清時適用。到了還清債務的那一天，公司只需償還所欠的餘額，而不是整筆 d。

債務何時還清

把含 $(1+k)^m$ 的項合併：

$$S_{T+m} = \frac{d}{k} - \frac{(1+k)^m}{k}\Big(d - c[(1+k)^T - 1]\Big).$$

想一想 從金錢的角度看，條件 $d/c > (1+k)^T - 1$ 究竟是甚麼意思？

把它乘以 c：條件變成 $d > c[(1+k)^T - 1] = kS_T$。換句話說，第一天的收入多於隧道通車時債務的利息。因此債務在第 $T+1$ 天下降，利息亦隨之減少，於是債務下降得愈來愈快。以下的代數運算把這個想法嚴格地表達出來。

假設 $d > c[(1+k)^T - 1]$，並記 $\delta = d - c[(1+k)^T - 1] > 0$。那麼

$$S_{T+m} = \frac{d - \delta(1+k)^m}{k}.$$

由於 $1+k > 1$，當 $m \to \infty$ 時 $(1+k)^m \to \infty$，所以 $\delta(1+k)^m$ 終會超過 d。這條公式只在債務為正數時成立，因此不可能永遠成立：必定有第一天，餘下的債務加上當天的利息不超過 d，公司就在那一天把債務還清。事實上，債務在第 $T+m$ 天還清，其中 m 是滿足下式的最小值：

$$(1+k)^m \geq \frac{d}{\delta}, \qquad \text{即} \qquad m \geq \frac{\ln(d/\delta)}{\ln(1+k)}.$$

所以，若 $d/c > (1+k)^T - 1$，公司最終能還清債務，正是所求的結果。

這個條件也是必要的。若 $d \le kS_T$，則 $\delta \le 0$；又因為 $S_T = (d-\delta)/k$，

$$S_{T+m} - S_T = \frac{-\delta[(1+k)^m - 1]}{k} \ge 0$$

對所有 m 都成立。債務永遠不會低於 S_T，所以永遠無法還清。

驗算

- 代入 $n = 1$：$S_1 = c$。這是對的，因為第 1 天並無欠款，所以無須支付利息。
- 在 S_{T+m} 的公式中代入 $m = 0$：得出 S_T，與預期相符。
- 若剛好有 $d = kS_T$，則 $\delta = 0$，而且對所有 m 都有 $S_{T+m} = d/k = S_T$：收入僅僅足夠支付利息，債務始終不變。這是臨界情況。
- 取 $c = 100$、$k = 0.01$、$T = 30$ 逐日模擬，得出 $S_T \approx 3478.49$，與公式吻合。門檻是 $d > 34.78$。取 $d = 50$ 時，債務在 $m = 120$ 天後還清，與 $\ln(50/15.22) / \ln 1.01 \approx 119.6$ 相符。

常見錯誤

- 在第 1 天就計算利息，或者把級數數成 $n+1$ 項。由於 $S_0 = 0$，第一天的成本剛好是 c。
- 忘記第 T 天之後級數是由 S_T 而不是由 0 開始：$(1+k)^m S_T$ 這一項很容易漏掉。
- 說債務能還清是因為「d 大於 c」。真正要比較的，是 d 與通車時的利息 kS_T。

STEP 之外

遞推關係 $S_n = aS_{n-1} + b$ 是最簡單的一階線性差分方程（linear first-order difference equation）。它的解是不動點（fixed point）$b/(1-a)$ 加上 a^n 的某個倍數，這正是我們那條公式的形式：d/k 加上 $(1+k)^m$ 的某個倍數。這個倍數是正還是負，決定了債務是不斷增長還是得以還清；同樣的想法會在不動點的穩定性（stability）和線性常微分方程（linear ODE）中再次出現。Boyce 與 DiPrima 的 *Elementary Differential Equations*（微分方程入門教科書）把差分方程與微分方程並列討論，而 Siklos 的 *Advanced Problems in Mathematics*（STEP 進階題集）則有更多關於遞推關係的 STEP 練習。

詳細解答

STEP III 1988，第 6 題

純數學 • 逐次微分

> **STEP III 1988，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/88-s3-q6

關鍵思路

直接把一個乘積求導 1988 次是行不通的，但正弦與餘弦的乘積可以改寫成一個和。這樣每一項 $\sin kx$ 的導數（derivative）便以四步為一個循環，每求導一次就多出一個因子 k。至於零點（zero），$\sin 3x$ 前面那個龐大的因子 3^{1988} 意味着導數只可能在 $\sin 3x$ 幾乎為零的地方等於零，所以我們只需看 $\sin 3x$ 第一個零點稍後的位置，然後作線性化（linearise）。

解答

第 1988 階導數

由積化和差公式（product-to-sum formula）$2\sin A\cos B = \sin(A+B) + \sin(A-B)$，

$$\mathrm{f}(x) = \sin 2x \cos x = \tfrac{1}{2}\left(\sin 3x + \sin x\right).$$

把 $\sin kx$ 求導四次，得

$$\sin kx \;\to\; k\cos kx \;\to\; -k^2 \sin kx \;\to\; -k^3 \cos kx \;\to\; k^4 \sin kx,$$

所以每求導四次，效果只是把 $\sin kx$ 乘以 k^4。（等價地說，$\sin kx$ 的第 n 階導數是 $k^n \sin\left(kx + \tfrac{1}{2}n\pi\right)$。）由於 $1988 = 4 \times 497$，

$$\mathrm{f}^{(1988)}(x) = \tfrac{1}{2}\left(3^{1988}\sin 3x + \sin x\right).$$

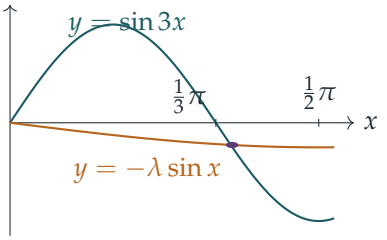
最小的正零點

我們要找出滿足下式的最小 $x > 0$：

$$3^{1988} \sin 3x + \sin x = 0. \qquad (*)$$

在 $(0, \frac{1}{3}\pi]$ 內沒有零點。當 $0 < x < \frac{1}{3}\pi$ 時，$0 < 3x < \pi$，所以 $\sin 3x$ 和 $\sin x$ 都是正數，$(*)$ 的左邊為正。在 $x = \frac{1}{3}\pi$ 處，左邊等於 $0 + \frac{\sqrt{3}}{2} > 0$。

緊接 $\frac{1}{3}\pi$ 之後有一個零點。在 $x = \frac{1}{2}\pi$ 處，左邊是 $3^{1988} \sin \frac{3}{2}\pi + 1 = 1 - 3^{1988} < 0$。左邊是連續（continuous）的，所以由介值定理（intermediate value theorem），在 $(\frac{1}{3}\pi, \frac{1}{2}\pi)$ 內必有零點，而最小的正零點就是其中的第一個。把它寫成 $x = \frac{1}{3}\pi + \epsilon$，其中 $0 < \epsilon < \frac{1}{6}\pi$。



草圖把 $(*)$ 表示為 $\sin 3x = -\lambda \sin x$，其中 $\lambda = 3^{-1988}$（圖中畫得遠比實際大）。兩條曲線首次相交之處，就在 $\sin 3x$ 於 $\frac{1}{3}\pi$ 穿過零之後不遠；對於真正的、極其微小的 λ，交點幾乎被推到 $\frac{1}{3}\pi$ 上。

由於 $\sin(\pi + 3\epsilon) = -\sin 3\epsilon$，方程 $(*)$ 變成

$$3^{1988} \sin 3\epsilon = \sin\left(\tfrac{1}{3}\pi + \epsilon\right). \qquad (**)$$

想一想 我們怎樣知道 ϵ 小得足以使用 $\sin 3\epsilon \approx 3\epsilon$？

$(**)$ 的右邊最多是 1，所以 $\sin 3\epsilon \le 3^{-1988}$。由於 $0 < 3\epsilon < \frac{1}{2}\pi$，而在 $[0, \frac{1}{2}\pi]$ 上有 $\sin t \ge \frac{2t}{\pi}$（正弦曲線在該區間內位於其弦（chord）的上方），我們得到 $3\epsilon \le \frac{\pi}{2} 3^{-1988}$。所以 ϵ 小得難以想像，這就說明了以下的近似是合理的。

ϵ 既然這樣小，便有 $\sin 3\epsilon = 3\epsilon\,(1 + O(\epsilon^2))$ 和 $\sin(\frac{1}{3}\pi + \epsilon) = \frac{\sqrt{3}}{2} + O(\epsilon)$。代入 $(**)$，

$$3^{1989}\epsilon \approx \frac{\sqrt{3}}{2}, \qquad \text{所以} \qquad \epsilon \approx \frac{3^{-1989}\sqrt{3}}{2} = \frac{1}{3} \cdot \frac{3^{-1988}\sqrt{3}}{2},$$

其相對誤差（relative error）的階為 3^{-1988}。

關於試卷上印出的數值。 試卷要求證明 $\epsilon \approx \frac{1}{2}3^{-1988}\sqrt{3}$，這是此處所得數值的三倍。因子 3 來自 $\sin 3\epsilon \approx 3\epsilon$，而下面的數值驗算也確認答案是 3^{-1989}，而不是 3^{-1988}。所以題目看來有一處筆誤：它原意要用的方法正是上面的方法，只是 3 的冪次差了一。如果你得到的是 3^{-1989}，請相信自己的計算。

驗算

- 同樣的論證適用於任何 4 的倍數 n：$3^n \sin 3x + \sin x$ 的第一個正零點位於 $\frac{1}{3}\pi + \epsilon_n$，其中 $\epsilon_n \approx \frac{1}{2}3^{-(n+1)}\sqrt{3}$。當 $n = 4$ 時，數值求解得 $\epsilon_4 = 3.5713 \times 10^{-3}$，而預測值是 $\frac{1}{2}\sqrt{3}/3^5 = 3.5640 \times 10^{-3}$（比值為 1.0021）。當 $n = 8$ 時，比值是 1.000025；到了 $n = 40$ 及以上，比值在機器精度內等於 1。它與 $\frac{1}{2}\sqrt{3}/3^n$ 的比值則趨向 $\frac{1}{3}$。
- 導數公式可以用最初幾個情況來檢驗：$\mathrm{f}^{(4)}(x) = \frac{1}{2}(81\sin 3x + \sin x)$，你可以把 $\frac{1}{2}(\sin 3x + \sin x)$ 求導四次來確認。

常見錯誤

- 弄錯正負號或相位：只有在求導次數是四的倍數之後，$\sin kx$ 才會回到 $+k^n \sin kx$。
- 忘記排除 $(0, \frac{1}{3}\pi)$ 內的零點；正是那裏的正負號論證，才使 $\frac{1}{3}\pi + \epsilon$ 成為最小的正零點。
- 寫成 $\sin(\pi + 3\epsilon) \approx \epsilon$ 而不是 -3ϵ，漏掉了正負號或因子 3。

STEP 之外

第二部分是攝動理論（perturbation theory）的一個簡單例子：一個很大的參數迫使解非常貼近某個已知的根，而在該根附近作線性化便得出首項修正，誤差的大小則由泰勒級數（Taylor series）的下一項控制。利用 $O(\cdot)$ 記號和誤差界，把「近似等於」說得精確，是大學第一門分析課程的核心主題；參見 Abbott 的 *Understanding Analysis*。

詳細解答

STEP III 1988，第 7 題

純數學 • 微分方程與正交性

> **STEP III 1988，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/88-s3-q7

本題在 *Beyond STEP* 第 8.6 節量子振子中有完整的引導式講解：https://beyondstep.keithmcnulty.org/zh-hant/ch08/quantum-oscillator.html

關鍵思路

第一部分純粹是代入（substitution）：把試探函數（trial function）求導兩次，再比較 x 各次冪的係數，便可確定 λ。第二部分是經典的「朗斯基技巧」（Wronskian trick）：組合 $y_m y_n' - y_n y_m'$ 的導數中，麻煩的 $\omega^2 x^2$ 項會互相抵消，只剩下 $y_m y_n$ 的一個倍數。之後把這個導數在整條實數線上積分就很容易，因為在 $\pm\infty$ 處的條件令邊界項（boundary terms）全部消失。

解答

全篇以撇號表示 $\mathrm{d}/\mathrm{d}x$。把方程寫成以下形式會比較方便：

$$y_n'' = \left(\omega^2 x^2 - (2n+1)\omega\right) y_n. \tag{$*$}$$

$n = 0$ 的情況

取 $y_0 = \mathrm{e}^{-\lambda x^2}$，則

$$y_0' = -2\lambda x\,\mathrm{e}^{-\lambda x^2}, \qquad y_0'' = \left(4\lambda^2 x^2 - 2\lambda\right)\mathrm{e}^{-\lambda x^2}.$$

代入 $n = 0$ 時的 $(*)$，並約去非零因子 $\mathrm{e}^{-\lambda x^2}$，我們需要

$$4\lambda^2 x^2 - 2\lambda = \omega^2 x^2 - \omega \quad \text{對所有} x.$$

比較常數項得 $\lambda = \frac{1}{2}\omega$，這時 x^2 項亦相符，因為 $4\lambda^2 = \omega^2$。（$4\lambda^2 = \omega^2$ 的另一個根 $\lambda = -\frac{1}{2}\omega$ 不滿足常數項的要求。）

$n = 1$ 的情況

取 $y_1 = x\,\mathrm{e}^{-\lambda x^2}$，則

$$\begin{aligned} y_1' &= (1 - 2\lambda x^2)\,\mathrm{e}^{-\lambda x^2}, \\ y_1'' &= \big(-4\lambda x - 2\lambda x(1 - 2\lambda x^2)\big)\mathrm{e}^{-\lambda x^2} = (4\lambda^2 x^3 - 6\lambda x)\,\mathrm{e}^{-\lambda x^2}. \end{aligned}$$

$n = 1$ 時的方程 $(*)$ 要求 $y_1'' = (\omega^2 x^2 - 3\omega)\,x\,\mathrm{e}^{-\lambda x^2}$，即

$$4\lambda^2 x^3 - 6\lambda x = \omega^2 x^3 - 3\omega x \quad \text{對所有} x,$$

而這同樣恰好在 $\lambda = \frac{1}{2}\omega$ 時成立。因此同一個常數對兩個函數都適用：

$$\lambda = \tfrac{1}{2}\omega, \qquad y_0 = \mathrm{e}^{-\omega x^2/2}, \qquad y_1 = x\,\mathrm{e}^{-\omega x^2/2}.$$

無窮遠處的條件

由於 $\lambda = \frac{1}{2}\omega > 0$，凡是形如（多項式）$\times \mathrm{e}^{-\lambda x^2}$ 的函數，當 $x \to \pm\infty$ 時都趨向 0，因為 $\mathrm{e}^{-\lambda x^2}$ 衰減的速度比 x 的任何次冪增長的速度都要快。函數 y_0、y_0'、y_1 和 y_1' 全部屬於這種形式（對應的多項式分別是 1、$-2\lambda x$、x 和 $1 - 2\lambda x^2$），所以這四個函數在兩端都趨向 0，符合要求。

恆等式

根據積法則（product rule），$y_m' y_n'$ 項互相抵消：

$$\frac{\mathrm{d}}{\mathrm{d}x}\big(y_m y_n' - y_n y_m'\big) = y_m' y_n' + y_m y_n'' - y_n' y_m' - y_n y_m'' = y_m y_n'' - y_n y_m''.$$

現在對 y_n'' 和 y_m'' 都使用 $(*)$：

$$\begin{aligned} y_m y_n'' - y_n y_m'' &= \big(\omega^2 x^2 - (2n+1)\omega\big) y_m y_n - \big(\omega^2 x^2 - (2m+1)\omega\big) y_m y_n \\ &= \big((2m+1) - (2n+1)\big)\,\omega\, y_m y_n \\ &= 2(m-n)\,\omega\, y_m y_n, \end{aligned}$$

這就是所要求的結果。

正交性

想一想 在嘗試證明這個無窮積分等於零之前，我們怎樣知道它確實存在？

我們先在有限區間上處理，再讓兩個端點分別趨向無窮；這樣存在性便自然而然地得到。記 $W = y_m y_n' - y_n y_m'$。對任意 $S, R > 0$，把恆等式由 $-S$ 積分至 R，得

$$2(m-n)\,\omega \int_{-S}^{R} y_m y_n \,\mathrm{d}x = W(R) - W(-S).$$

當 $R \to \infty$ 時，根據題目所給的條件，y_m, y_n, y_m', y_n' 各自趨向 0，所以 $W(R) \to 0$。同理，當 $S \to \infty$ 時 $W(-S) \to 0$。因此右邊趨向 0，故積分 $\int_{-\infty}^{\infty} y_m y_n \,\mathrm{d}x$ 收斂(converges)，並且

$$2(m-n)\,\omega \int_{-\infty}^{\infty} y_m y_n \,\mathrm{d}x = 0.$$

若 $m \neq n$，則 $2(m-n)\omega \neq 0$，因為 $\omega > 0$。兩邊除以它，得

$$\int_{-\infty}^{\infty} y_m(x)\, y_n(x)\,\mathrm{d}x = 0,$$

正如所求。請留意，這裏只用到方程本身和兩端的條件：我們從來不需要 y_m 或 y_n 的具體公式。

驗算

- 當 $m = 0$、$n = 1$ 時，結果可以直接看出：$y_0 y_1 = x\,\mathrm{e}^{-\omega x^2}$ 是奇函數（odd function），所以它在整條實數線上的積分是 0。
- 一個沒那麼顯然的檢驗：$y_2 = (2\omega x^2 - 1)\mathrm{e}^{-\omega x^2/2}$ 滿足 $n = 2$ 時的方程（不妨自己試試）。利用 $\int_{-\infty}^{\infty} \mathrm{e}^{-\omega x^2}\,\mathrm{d}x = \sqrt{\pi/\omega}$ 和 $\int_{-\infty}^{\infty} x^2 \mathrm{e}^{-\omega x^2}\,\mathrm{d}x = \frac{1}{2\omega}\sqrt{\pi/\omega}$，可得 $\int y_0 y_2\,\mathrm{d}x = 2\omega \cdot \frac{1}{2\omega}\sqrt{\pi/\omega} - \sqrt{\pi/\omega} = 0$。
- 數值方面，取 $\lambda = \frac{1}{2}\omega$ 時，方程在隨機點上的殘差（residual）對 $n = 0, 1, 2$ 都小於 10^{-6}；以有限差分（finite differences）檢驗，導數恆等式準確至 10^{-4}；而在 $[-30, 30]$ 上使用辛普森法則（Simpson's rule），對 $\omega = 0.5, 1, 2.5$，$\int y_0 y_1$、$\int y_0 y_2$、$\int y_1 y_2$ 全部小於 10^{-15}。

常見錯誤

- 只比較 x^2 的係數而得出 $\lambda = \pm\frac{1}{2}\omega$：常數項排除了負號（無窮遠處的衰減條件亦然）。
- 忘記說明邊界項為何消失：這正是同時用到 y_n 和 y_n' 兩個條件的地方。
- 除以 $m - n$ 時，沒有指出它不等於零，亦沒有指出 $\omega > 0$。

STEP 之外

這就是量子諧振子（quantum harmonic oscillator）：y_n 是它的能量態，$(2n+1)\omega$ 是它的能級（相差一個比例常數），而這些函數是埃爾米特多項式（Hermite polynomials）乘以一個高斯函數（Gaussian）。這個正交性（orthogonality）論證是施圖姆－劉維爾理論（Sturm–Liouville theory）的一個特例；該理論證明，自伴（self-adjoint）微分算子對應於不同特徵值（eigenvalue）的特徵函數（eigenfunction）必定正交，正如對稱矩陣的特徵向量（eigenvector）一樣。Boyce 與 DiPrima 所著的 *Elementary Differential Equations*（微分方程入門教材）涵蓋邊值問題和施圖姆－劉維爾理論。

詳細解答

STEP III 1988，第 8 題

純數學 • 拋物線的切線與法線

> **STEP III 1988，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/88-s3-q8

關鍵思路

把每一條直線都用參數（parameter）t 表示，求出各對直線的交點，再把這些頂點代入行列式（determinant）。這些行列式看來很繁複，但它們對 t_1, t_2, t_3 是反對稱（antisymmetric）的，所以每一個都必定是 $(t_1 - t_2)(t_2 - t_3)(t_3 - t_1)$ 的倍數。只要用簡單的行運算和列運算（row and column operations），就可以把這個因子提出來，餘下的就是兩個面積之間的比較。

（印刷的題目要求的是 A_3；這是 A_2 的誤印，即法線三角形的面積。）

解答

切線與法線

沿着 $x = at^2, y = 2at$ 把 $y^2 = 4ax$ 對 t 求導，得 $\dfrac{\mathrm{d}y}{\mathrm{d}x} = \dfrac{2a}{2at} = \dfrac{1}{t}$。在 $(at^2, 2at)$ 的切線（tangent）是 $y - 2at = \frac{1}{t}(x - at^2)$，即

$$ty = x + at^2.$$

這個形式在 $t = 0$ 時也成立（在頂點的切線是 $x = 0$）。法線（normal）的斜率（gradient）是 $-t$：

$$y + tx = 2at + at^3.$$

我們全程假設 $a \neq 0$。

切線三角形

把在 t_i 和 t_j 的切線方程相減，得 $(t_i - t_j)y = a(t_i^2 - t_j^2)$。由於 $t_i \neq t_j$，我們可以相除，所以 $y = a(t_i + t_j)$，於是 $x = t_i y - a t_i^2 = a t_i t_j$。因此

$$\text{在} t_i, t_j \text{ 的兩條切線相交於} \big(a t_i t_j,\ a(t_i + t_j)\big).$$

各切線的斜率 $1/t_i$ 互不相同（或者其中一條是鉛垂的），所以沒有兩條是平行的，三角形確實存在。記 $s = t_1 + t_2 + t_3$。與 P_1 相對的頂點是 $(a t_2 t_3,\ a(s - t_1))$，其餘兩個頂點類似。從首兩列各提出因子 a，然後把第 3 列的 $-s$ 倍加到第 2 列，再把第 2 列變號，

$$\det \begin{pmatrix} a t_2 t_3 & a(s - t_1) & 1 \\ a t_1 t_3 & a(s - t_2) & 1 \\ a t_1 t_2 & a(s - t_3) & 1 \end{pmatrix} = -a^2 \det \begin{pmatrix} t_2 t_3 & t_1 & 1 \\ t_1 t_3 & t_2 & 1 \\ t_1 t_2 & t_3 & 1 \end{pmatrix} = a^2 D,$$

當中我們對調了首兩列，並設

$$D = \det \begin{pmatrix} t_1 & t_2 t_3 & 1 \\ t_2 & t_1 t_3 & 1 \\ t_3 & t_1 t_2 & 1 \end{pmatrix}.$$

要計算 D，把第 1 行換成第 1 行 − 第 2 行，把第 2 行換成第 2 行 − 第 3 行。第 1 行變成 $(t_1 - t_2,\ t_3(t_2 - t_1),\ 0) = (t_1 - t_2)(1,\ -t_3,\ 0)$，而第 2 行變成 $(t_2 - t_3)(1,\ -t_1,\ 0)$。沿最後一列展開，

$$D = (t_1 - t_2)(t_2 - t_3) \det \begin{pmatrix} 1 & -t_3 \\ 1 & -t_1 \end{pmatrix} = (t_1 - t_2)(t_2 - t_3)(t_3 - t_1).$$

因此

$$A_1 = \tfrac{1}{2} a^2 \left| (t_1 - t_2)(t_2 - t_3)(t_3 - t_1) \right|,$$

由於各 t_i 互不相同，這個面積不等於零。

法線三角形

把在 t_i 和 t_j 的法線方程相減，得 $(t_i - t_j)x = 2a(t_i - t_j) + a(t_i^3 - t_j^3)$，所以除以 $t_i - t_j \neq 0$ 後，

$$x = 2a + a(t_i^2 + t_i t_j + t_j^2), \qquad y = 2a t_i + a t_i^3 - t_i x = -a t_i t_j (t_i + t_j).$$

各法線的斜率 $-t_i$ 互不相同，所以同樣地每一對法線都相交。對於與 P_1 相對的頂點，利用 $t_2 + t_3 = s - t_1$，並記 $e_1 = t_2 t_3$（以及 $e_2 = t_1 t_3, e_3 = t_1 t_2$）。那麼

$$t_2^2 + t_2 t_3 + t_3^2 = (t_2 + t_3)^2 - t_2 t_3 = (s - t_1)^2 - e_1, \qquad -t_2 t_3 (t_2 + t_3) = -s e_1 + p,$$

其中 $p = t_1t_2t_3$。所以這個頂點是 $a\big(2 + (s - t_1)^2 - e_1,\ -se_1 + p\big)$，其餘兩個頂點類似。現在提出 a^2，並進行不改變行列式的值的列運算：從第 1 列減去第 3 列的 2 倍，從第 2 列減去第 3 列的 p 倍，然後從第 2 列提出因子 $-s$：

$$\begin{aligned}\frac{1}{a^2}\det(\text{法線三角形的頂點}) &= \det\Big((s-t_i)^2 - e_i \quad -se_i \quad 1\Big)_{i=1,2,3}\\ &= -s\det\Big((s-t_i)^2 - e_i \quad e_i \quad 1\Big)_{i=1,2,3},\end{aligned}$$

當中每個式子只顯示有代表性的第 i 行。接着把第 2 列加到第 1 列，這樣就消去了 $-e_i$，再展開 $(s - t_i)^2 = s^2 - 2st_i + t_i^2$。當中的 s^2 是第 3 列的倍數，可以消去，而行列式對第 1 列是線性的：

$$\det\Big(t_i^2 - 2st_i \quad e_i \quad 1\Big) = \det\Big(t_i^2 \quad e_i \quad 1\Big) - 2s\det\Big(t_i \quad e_i \quad 1\Big) = E - 2sD,$$

這裏我們如此記這兩個行列式。第二個行列式正好就是 D。

想一想 我們怎樣計算 E，即各行為 $(t_i^2,\ e_i,\ 1)$ 的行列式？

使用與計算 D 時相同的行運算。第 1 行 − 第 2 行是 $(t_1^2 - t_2^2,\ t_3(t_2 - t_1),\ 0) = (t_1 - t_2)(t_1 + t_2,\ -t_3,\ 0)$，而第 2 行 − 第 3 行是 $(t_2 - t_3)(t_2 + t_3,\ -t_1,\ 0)$。所以

$$\begin{aligned}E &= (t_1 - t_2)(t_2 - t_3)\big[-t_1(t_1 + t_2) + t_3(t_2 + t_3)\big]\\ &= (t_1 - t_2)(t_2 - t_3)\big[(t_3^2 - t_1^2) + t_2(t_3 - t_1)\big]\\ &= (t_1 - t_2)(t_2 - t_3)(t_3 - t_1)(t_1 + t_2 + t_3) = sD.\end{aligned}$$

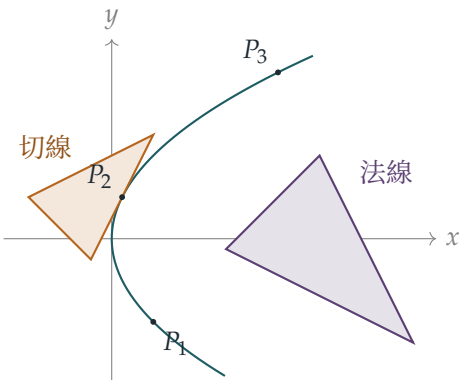
把以上各步合起來，

$$\det(\text{法線三角形的頂點}) = a^2 \cdot (-s)\,(sD - 2sD) = a^2s^2D.$$

取絕對值的一半，並與 $A_1 = \frac{1}{2}a^2|D|$ 比較，

$$A_2 = (t_1 + t_2 + t_3)^2A_1,$$

這就是所求的結果。



圖中取 $a = 1$ 及 $t_1, t_2, t_3 = -1, \frac{1}{2}, 2$，所以 $s = \frac{3}{2}$。切線三角形（琥珀色）的面積是 $\frac{27}{8}$，而法線三角形（紫紅色）的面積是 $\frac{243}{32} = \left(\frac{3}{2}\right)^2 \cdot \frac{27}{8}$。

法線何時共點？

沒有兩條法線是平行的，所以它們兩兩相交於三點 Q_{12}, Q_{13}, Q_{23}。如果三條法線共點（concurrent），這三點就會重合，三角形的面積為零。反過來，假設面積為零但三點並非全部相同，例如 $Q_{12} \neq Q_{13}$。這兩點都在 P_1 的法線上，所以通過它們的直線就是該法線。由於三點共線（collinear），Q_{23} 也在 P_1 的法線上。於是 Q_{23} 同時在三條法線上，所以 $Q_{23} = Q_{12}$（在 P_1 和 P_2 的兩條法線的唯一公共點），同理 $Q_{23} = Q_{13}$，這就產生矛盾。所以

$$\text{三條法線共點} \iff A_2 = 0 \iff (t_1 + t_2 + t_3)^2 A_1 = 0 \iff t_1 + t_2 + t_3 = 0,$$

這裏用到 $A_1 \neq 0$，因為各 t_i 互不相同。所求的條件是 $t_1 + t_2 + t_3 = 0$。

驗算

- 取 $t_1, t_2, t_3 = 0, 1, 2$ 及 $a = 1$，切線三角形的頂點是 $(2,3)$、$(0,2)$、$(0,1)$，面積為 1，而 $\frac{1}{2}|D| = \frac{1}{2}|(-1)(-1)(2)| = 1$。
- 在 $t = 1, 2, -3$（總和為零，$a = 1$）的法線全部通過 $(9, -6)$。例如，在 $t = 1$ 的法線是 $y + x = 3$。
- 我們用電腦以分數作精確計算，對大約 2,800 組隨機選取的 a 和互不相同的 t_i 求出各頂點，每一次都確認了 $A_2 = (t_1 + t_2 + t_3)^2 A_1$。

常見錯誤

- 忘記說明三條直線確實構成一個三角形。我們需要的是兩兩不平行，而這裏可由斜率互不相同得出。
- 未加理由就從 $A_2 = 0$ 斷定法線共點。面積為零只說明三個交點共線，上面那段簡短的論證才把它提升為「相同」。
- 把 3×3 行列式完全展開。這樣做是可行的，但幾乎肯定會出現正負號錯誤。應先用行運算提出因子 $(t_i - t_j)$。

STEP 之外

因子 $(t_1 - t_2)(t_2 - t_3)(t_3 - t_1)$ 是一個范德蒙行列式（Vandermonde determinant），而 $E = sD$ 中的餘因子 $t_1 + t_2 + t_3$ 是一個對稱多項式（symmetric polynomial）。任何反對稱多項式都是范德蒙乘積乘以一個對稱多項式，這個事實在行列式和對稱函數的理論中會再次出現。法線三角形與拋物線的漸屈線（evolute）有關，漸屈線是拋物線的曲率中心（centres of curvature）所描出的曲線，也就是各法線的包絡線（envelope）。Strang 的 *Introduction to Linear Algebra*（線性代數入門）對行列式及其行運算有很好的講解。

詳細解答

STEP III 1988，第 9 題

純數學 • 群元素的階

> **STEP III 1988，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/88-s3-q9

關鍵思路

第 (i) 部分其實是帶餘除法（division algorithm）的變相應用：把 N 寫成 n 的倍數加上餘數，而 n 的最小性迫使餘數為零。第 (ii) 部分之所以成立，是因為以 g 作共軛（conjugation）會保持乘積，所以在一串共軛元素的乘積中，夾在中間的每一對 $g^{-1}g$ 都會互相抵消。在第 (iii) 部分，我們反覆以 g 作共軛；由於 $g^5 = e$，做五次之後 h 便回到自身，從而確定 h 的階（order）。

解答

第 (i) 部分

根據帶餘除法，可寫 $N = qn + r$，其中 q 為整數，且 $0 \leq r < n$。（這對任何整數 N 都成立，包括零和負整數。）於是，由於 $g^n = e$，

$$e = g^N = (g^n)^q g^r = e^q g^r = g^r.$$

假如 r 是正數，它便是一個小於 n 而又滿足 $g^r = e$ 的正整數，這與 $n = o(g)$ 是符合此條件的最小正整數這個定義矛盾。因此 $r = 0$，即 $N = qn$，故 n 整除 N。

第 (ii) 部分

$m = 0$：兩邊都等於 e。

$m \geq 1$：我們對 m 使用歸納法（induction）。$m = 1$ 的情況顯然成立。若 $gh^m g^{-1} =$

$(ghg^{-1})^m$，則

$$(ghg^{-1})^{m+1} = (ghg^{-1})^m\,(ghg^{-1}) = (gh^mg^{-1})(ghg^{-1})$$
$$= gh^m(g^{-1}g)hg^{-1} = gh^{m+1}g^{-1},$$

歸納法至此完成。（把乘積 $(ghg^{-1})(ghg^{-1})\cdots(ghg^{-1})$ 全部寫出來，便會逐項相消：中間的每一個 $g^{-1}g$ 都抵消掉。）

$m \le -1$：首先，ghg^{-1} 的逆元（inverse）是 $gh^{-1}g^{-1}$，因為

$$(ghg^{-1})(gh^{-1}g^{-1}) = gh(g^{-1}g)h^{-1}g^{-1} = ghh^{-1}g^{-1} = e.$$

現在寫 $m = -k$，其中 $k \ge 1$。以元素 h^{-1} 代替 h，套用正數情況的結果，得

$$gh^mg^{-1} = g(h^{-1})^kg^{-1} = (gh^{-1}g^{-1})^k = \left((ghg^{-1})^{-1}\right)^k = (ghg^{-1})^{-k} = (ghg^{-1})^m.$$

所以這個恆等式對每一個整數 m 都成立。

第 (iii) 部分

先用 $ghg^{-1} = h^2$，再用第 (ii) 部分並取 $m = 2$，

$$g^2hg^{-2} = g\,(ghg^{-1})\,g^{-1} = gh^2g^{-1} = (ghg^{-1})^2 = (h^2)^2 = h^4,$$

正如所求。

同樣的步驟在每一個階段都適用。我們斷言：對所有 $k \ge 0$，$g^khg^{-k} = h^{2^k}$。$k = 0$ 時這顯然成立；若它對 k 成立，則由第 (ii) 部分並取 $m = 2^k$，

$$g^{k+1}hg^{-(k+1)} = g\,(g^khg^{-k})\,g^{-1} = gh^{2^k}g^{-1} = (ghg^{-1})^{2^k} = (h^2)^{2^k} = h^{2^{k+1}}.$$

取 $k = 5$，並利用 $g^5 = e$（因此 $g^{-5} = e$ 亦成立），得

$$h = g^5hg^{-5} = h^{32}, \qquad \text{所以} \qquad h^{31} = e$$

（兩邊乘以 h^{-1} 即得）。

> **想一想** 我們知道 $h^{31} = e$。為甚麼由此可知 $o(h)$ 恰好是 31，而不僅僅是不超過 31？

由第 (i) 部分，$o(h)$ 整除 31。由於 31 是質數（prime），$o(h)$ 只能是 1 或 31。但 $o(h) = 1$ 即表示 $h = e$，而題目已說明並非如此。因此

$$o(h) = 31.$$

這樣的羣（group）確實存在，所以題目的條件並無矛盾。考慮 $\{0, 1, \ldots, 30\}$ 上的兩個置換（permutation）：$h(x) = x + 1$ 和 $g(x) = 2x$，運算皆以 31 為模。那麼 $g^5(x) = 32x = x$，而且 $ghg^{-1}(x) = 2\left(\frac{x}{2} + 1\right) = x + 2 = h^2(x)$，同時 h 的階確實是 31。

驗算

- 出現質數 31 是合理的：以 g 作共軛，對 h 的各次冪的作用就是「把指數乘以 2」，而加倍五次後必須回到起點，所以 $2^5 \equiv 1$（模 $o(h)$）。換句話說，$o(h)$ 整除 $2^5 - 1 = 31$。
- 我們以置換 $g(x) = 2x$ 和 $h(x) = x + 1$（模 31）作電腦驗算，確認了 $g^5 = e$、$ghg^{-1} = h^2$、$g^k h g^{-k} = h^{2^k}$（$k = 0, \ldots, 5$）以及 $o(h) = 31$。此外，亦在數百個隨機置換上，就 $-12 \le m \le 12$ 檢驗了 (i) 和 (ii)。

常見錯誤

- 在 (ii) 中只證明 $m \ge 1$ 的情況。題目說的是任何整數，所以 $m = 0$ 和負數 m 都需要各自的（簡短）論證。
- 直接由 $h^{31} = e$ 斷定 $o(h) = 31$。這只能得出 $o(h) \mid 31$；還需要用到 31 是質數以及 $h \neq e$ 這個條件。

STEP 之外

映射 $x \mapsto gxg^{-1}$ 是 G 的一個自同構（automorphism），即保持結構的雙射（bijection），而第 (ii) 部分正是這個事實的一個推論。第 (iii) 部分是半直積（semidirect product）的一個小例子：由 g 和 h 生成的羣的階是 $5 \times 31 = 155$，而它之所以存在，全賴 2 在模 31 之下的階是 5。這種構造在小階羣的分類中佔有核心地位。若想進一步了解階、共軛和拉格朗日定理（Lagrange's theorem），Allenby 的 *Rings, Fields and Groups*（環、域與羣）是很好的讀物。

詳細解答

STEP III 1988，第 10 題

純數學 • 追蹤曲線

> **STEP III 1988，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/88-s3-q10

本題在 *Beyond STEP* 第 8.5 節格力犬追逐格力犬中有完整的引導式講解：https://beyondstep.keithmcnulty.org/zh-hant/ch08/pursuit.html

關鍵思路

不要嘗試分別追蹤四隻移動中的狗。整個設定繞中心旋轉四分之一圈後維持不變，所以四隻狗始終位於一個以 O 為中心的正方形的四個頂點上：這個正方形會縮小、會旋轉，但始終是正方形。這就確定了 A 的速度與直線 AO 之間的夾角，而速度與半徑之間的夾角固定，正正就是產生等角螺線（equiangular spiral）$r = \lambda \mathrm{e}^{-k\theta}$ 的條件。

解答

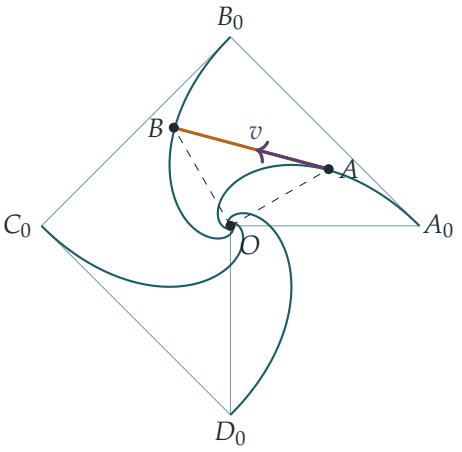
四隻狗始終位於正方形上

取正方形的中心為原點 O，並按逆時針方向為四隻狗編號，使得 B 位於 A 逆時針旋轉四分之一圈之處，C 位於 B 逆時針旋轉四分之一圈之處，如此類推。

把整幅圖繞 O 旋轉 $\frac{\pi}{2}$。A, B, C, D 的起始位置分別變成 B, C, D, A 的起始位置，而「以速率 v 直奔下一隻狗」這條規則對每隻狗都是一樣的。所以旋轉後的運動，同樣是由相同的起始佈局出發、遵從相同規則的運動。由給定起點出發的運動是唯一的（只要四隻狗尚未相遇，速度就是位置的光滑函數），所以旋轉後的運動就是實際的運動。換言之，在每一瞬間，B 的位置就是 A 的位置繞 O 旋轉 $\frac{\pi}{2}$ 所得的位置。C 和 D 亦然。所以 $ABCD$ 始終是一個以 O 為中心的正方形。

速度與半徑之間的夾角

設 A 在極坐標（polar coordinates）中位於 (r,θ)。那麼 B 位於 $(r,\theta+\frac{\pi}{2})$，所以 $OA=OB=r$，而 $\angle AOB=\frac{\pi}{2}$。三角形 OAB 是等腰直角三角形，所以 $\angle OAB=\frac{\pi}{4}$。



圖中顯示四條路徑由起始頂點 A_0,B_0,C_0,D_0 向內盤旋，並標出某一瞬間 A 和 B 兩隻狗的位置。A 的速度（紫紅色）沿 AB 方向。

A 的速度大小為 v，方向沿 AB，與向內的半徑 AO 成 $\frac{\pi}{4}$ 角。把它分解為徑向分量(radial component，以向外為正）和橫向分量（transverse component，沿 θ 增加的方向）：

$$\dot r=-v\cos\frac{\pi}{4},\qquad r\dot\theta=v\sin\frac{\pi}{4}.$$

橫向分量為正，是因為 B 位於 A 的逆時針方向，所以 A 被拉着逆時針轉動。

想一想 為甚麼可以把這兩條方程相除？

由於四隻狗尚未相遇時 $r\dot\theta=v/\sqrt{2}>0$，θ 隨時間嚴格遞增。所以我們可以用 θ 代替 t 作為路徑上的參數，再由鏈式法則（chain rule）得

$$\frac{\mathrm{d}r}{\mathrm{d}\theta}=\frac{\dot r}{\dot\theta}=\frac{-v/\sqrt{2}}{v/(\sqrt{2}r)}=-r.$$

求出路徑

由於 $r>0$，我們可以分離變數（separate variables）：$\int\frac{\mathrm{d}r}{r}=-\int\mathrm{d}\theta$，所以 $\ln r=-\theta+c$，即

$$r=\lambda\mathrm{e}^{-\theta}$$

其中 $\lambda=\mathrm{e}^c$ 是常數。這就是所求的結果。如果我們選取 OA_0 為始線，使得開始時 $\theta=0$，那麼 λ 就是 A 與 O 的初始距離，對於邊長為 a 的正方形，即 $a/\sqrt{2}$。若選取其他始

線，λ 只會相差一個常數因子。（如果我們按順時針方向為四隻狗編號，則以順時針方向量度 θ 便會得到相同的方程。）

推廣至 n 隻狗

對於位於正 n 邊形各頂點上的 n 隻狗（按逆時針方向編號，每隻追逐下一隻），把旋轉角改為 $\frac{2\pi}{n}$，同樣的論證依然成立。這些狗始終構成一個以 O 為中心的正 n 邊形。所以若 A 位於 (r,θ)，下一隻狗 B 就位於 $(r,\theta+\frac{2\pi}{n})$。

三角形 OAB 是等腰三角形，頂角為 $\frac{2\pi}{n}$，所以每個底角為

$$\angle OAB = \frac{1}{2}\left(\pi - \frac{2\pi}{n}\right) = \frac{\pi}{2} - \frac{\pi}{n}.$$

如前分解 A 的速度，

$$\dot{r} = -v\cos\left(\frac{\pi}{2}-\frac{\pi}{n}\right) = -v\sin\frac{\pi}{n}, \qquad r\dot{\theta} = v\sin\left(\frac{\pi}{2}-\frac{\pi}{n}\right) = v\cos\frac{\pi}{n}.$$

由於 $n \geq 3$，$\cos\frac{\pi}{n} > 0$，所以 θ 同樣穩步遞增，並且

$$\frac{\mathrm{d}r}{\mathrm{d}\theta} = -r\tan\frac{\pi}{n} \qquad \Longrightarrow \qquad r = \lambda\,\mathrm{e}^{-\theta\tan(\pi/n)}.$$

當 $n=4$ 時，$\tan\frac{\pi}{4}=1$，我們重新得到 $r=\lambda\mathrm{e}^{-\theta}$。每隻狗都沿着一條等角螺線（又稱對數螺線，logarithmic spiral）前進，以固定的角 $\frac{\pi}{2}-\frac{\pi}{n}$ 與每條半徑相交。

一個有趣的副產品：$\dot{r} = -v\sin\frac{\pi}{n}$ 是常數，所以這些狗在時間 $r_0/(v\sin\frac{\pi}{n})$ 之後於 O 相遇，其中 r_0 是與 O 的初始距離。在這段時間內，每隻狗跑過的距離是有限的 $r_0/\sin\frac{\pi}{n}$，儘管 $\theta\to\infty$，牠繞着 O 轉了無限多圈。（對於正方形，這個距離是 a，即邊長。）

驗算

- 當 $n=4$ 時，結果就是一般公式在 $\tan\frac{\pi}{4}=1$ 時的特殊情況。
- 當 $n\to\infty$ 時，$\tan\frac{\pi}{n}\to 0$，路徑趨向一個圓：每隻狗所追逐的鄰居，在圓上幾乎與牠並排。
- 我們以數值方法模擬了這場追逐（用 RK4 方法，取 $n=3,4,5,6,8$ 隻狗，由半徑 1 出發）。算出的 r 與 $\mathrm{e}^{-\theta\tan(\pi/n)}$ 吻合，相對誤差為 10^{-11}。當 $n=4$ 時，到達 $r=0.02$ 所需的時間和路徑長度均算得 $1.386 = 0.98\sqrt{2}$，與 $(1-0.02)/\sin\frac{\pi}{4}$ 所預測的一致。

常見錯誤

- 未先論證這些狗始終位於正多邊形上，就直接寫下 $\dot{r}$ 和 $r\dot{\theta}$。正是靠對稱性論證，我們才知道 A 處的角。

- 用了多邊形的內角（正方形為 $\frac{\pi}{2}$），而不是 AB 與半徑 AO 之間的夾角（$\frac{\pi}{4}$）。
- 忘記了橫向速度分量是 $r\dot{\theta}$，而不是 $\dot{\theta}$。

STEP 之外

這是一道經典的追逐問題（pursuit problem）。對稱性那一步，是微分方程（differential equations）中一個普遍原理的例子：如果方程和初始數據在某個對稱變換下都保持不變，那麼解的唯一性就迫使解本身也保持不變。與每條半徑以固定角相交的曲線，就是等角螺線。以複數形式表示，它們是 $z = \mathrm{e}^{(-k+\mathrm{i})\theta}$，所以整個運動是平面上一個連續的「旋轉相似變換」（spiral similarity）。Needham 的 *Visual Complex Analysis*（以幾何觀點講解複分析）闡述了這種幾何觀點，而 Boyce 與 DiPrima 的 *Elementary Differential Equations*（初等微分方程教科書）則涵蓋了解的存在性與唯一性。

詳細解答

STEP III 1988，第 11 題

力學 • 涉及摩擦力的梯子平衡

STEP III 1988，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/88-s3-q11

關鍵思路

先弄清幾何關係：地面、斜坡和梯子構成一個三角形，它的各個角你其實已經知道，由此便可判斷光滑斜坡對梯頂的推力朝哪個方向。之後就是處理梯子問題的標準步驟：分解（resolve）兩次，對梯腳取力矩（moments），再把所需的摩擦力（friction）與 μ 乘以法向反作用力（normal reaction）作比較。由於工人的位置只出現在力矩方程之中，一條通用公式便可同時回答兩部分。

解答

準備工作

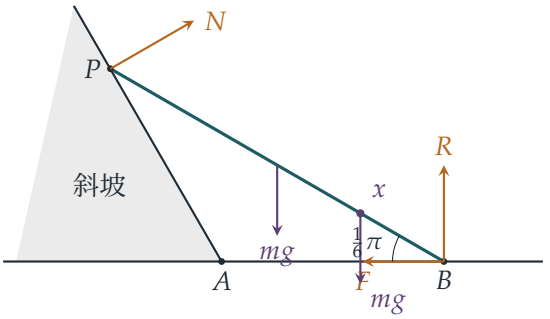
把梯子視為均勻桿（uniform rod），把工人視為質點（particle）。設 A 為斜坡的底端，B 為梯腳，P 為梯頂，梯頂位於斜坡上。

*幾何關係。*在三角形 ABP 中，B 處的角為 $\frac{1}{6}\pi$（即梯子的傾角）。斜坡與鉛垂線成 $\frac{1}{6}\pi$，所以它在 A 的另一側與水平線成 $\frac{1}{3}\pi$。因此三角形內 A 處的角為 $\pi - \frac{1}{3}\pi = \frac{2}{3}\pi$，餘下 $\frac{1}{6}\pi$ 便是 P 處的角。所以這是一個等腰三角形（isosceles triangle），$AP = AB$，而由正弦定理（sine rule）可得

$$AB = \frac{l \sin \frac{1}{6}\pi}{\sin \frac{2}{3}\pi} = \frac{l}{\sqrt{3}},$$

與題目給出的距離吻合。題目的數據互相一致，P 的確在斜坡上。

*斜坡反作用力的方向。*斜坡是光滑（smooth）的，所以它對梯子的反作用力 N 垂直於斜坡。斜坡與水平線成 $\frac{1}{3}\pi$，故其法線與水平線成 $\frac{1}{6}\pi$，指向斜上方並離開斜坡，朝着 B 的一方。因此 N 的水平分量為 $\frac{\sqrt{3}}{2}N$，把梯子推離斜坡；鉛垂分量為 $\frac{1}{2}N$，方向向上。



方程。設工人沿梯子向上距離梯腳 x。在 B 處，地面施加向上的法向反作用力 R，以及沿水平方向指向斜坡的摩擦力 F（除此之外唯一的水平力是 N 的分量，它把梯腳推離斜坡，所以摩擦力必須與之相反）。這裏的距離是由 B 起計。

沿水平和鉛垂方向分解：

$$F = \tfrac{\sqrt{3}}{2}N, \qquad\qquad R + \tfrac{1}{2}N = 2mg.$$

對 B 取力矩，這樣 R 和 F 便不會出現。由 B 到梯子上距離梯腳 s 的一點，水平距離為 $s\cos\frac{1}{6}\pi = \frac{\sqrt{3}}{2}s$。$N$ 的作用線與水平線成 $\frac{1}{6}\pi$，而 BP 則在水平線的另一側與之成 $\frac{1}{6}\pi$，所以 N 與梯子的夾角為 $\frac{1}{3}\pi$。它對 B 的力矩是 $Nl\sin\frac{1}{3}\pi = \frac{\sqrt{3}}{2}Nl$。因此

$$\tfrac{\sqrt{3}}{2}Nl = mg\cdot\tfrac{\sqrt{3}}{2}\cdot\tfrac{l}{2} + mg\cdot\tfrac{\sqrt{3}}{2}x \quad\Longrightarrow\quad N = mg\left(\frac{1}{2}+\frac{x}{l}\right).$$

想一想 為甚麼 N 的力矩會這樣簡單？

由 B 到 N 的作用線的垂直距離是 $l\sin\frac{1}{3}\pi$，因為 N 與梯子相交的角度是 $\frac{1}{3}\pi$。又或者把 N 分解成分量：水平部分 $\frac{\sqrt{3}}{2}N$ 作用於高度 $\frac{1}{2}l$ 處，鉛垂部分 $\frac{1}{2}N$ 作用於水平距離 $\frac{\sqrt{3}}{2}l$ 處。兩者的力矩都是 $\frac{\sqrt{3}}{4}Nl$，加起來就是 $\frac{\sqrt{3}}{2}Nl$。

於是當 $R = 2mg - \frac{1}{2}N = mg\left(\frac{7}{4} - \frac{x}{2l}\right) > 0$ 對 $0 \le x \le l$ 成立，梯腳便一直留在地面上。只要 $F \le \mu R$，梯子就不會滑動，即

$$\mu \ge \frac{F}{R} = \frac{\sqrt{3}\,N}{4mg - N} = \frac{\sqrt{3}\,(l+2x)}{7l-2x}.$$

這個比值隨 x 增加而增加：分子增加，而恆為正數的分母減少。所以工人愈往上爬，所需的摩擦力便愈大。

第 (i) 部分

工人一到達梯子的中點，梯子便滑動。所以極限摩擦（limiting friction）恰好在 $x = \frac{1}{2}l$ 時出現，而由於 F/R 隨 x 遞增，在此之前不會出現。這時 $N = mg$、$F = \frac{\sqrt{3}}{2}mg$ 及

$R = \frac{3}{2}mg$，所以

$$\mu = \frac{F}{R} = \frac{\sqrt{3}/2}{3/2} = \frac{1}{\sqrt{3}}.$$

第 (ii) 部分

若 $F \le \mu R$ 對每一個 $x \in [0, l]$ 都成立，工人便能到達梯頂。由於 F/R 隨 x 遞增，最壞的情況是 $x = l$。這時 $N = \frac{3}{2}mg$、$F = \frac{3\sqrt{3}}{4}mg$ 及 $R = \frac{5}{4}mg$，所以摩擦系數（coefficient of friction）的最小值為

$$\mu = \frac{3\sqrt{3}/4}{5/4} = \frac{3\sqrt{3}}{5} \approx 1.04.$$

驗算

- 梯上無人時（$x = 0$），公式給出 $\mu \ge \sqrt{3}/7 \approx 0.247$。這比兩個答案都小，正合預期。
- 把 $l = 1$、$m = 1$、$g = 9.8$ 代入三條平衡方程作數值計算，並把各個力寫成向量，可重現 $N/mg = 1$ 和 $\frac{3}{2}$，以及比值 $F/R = 0.57735\ldots$ 和 $1.03923\ldots$。計算亦確認，在梯子上的 101 個位置，F/R 都是遞增的。

常見錯誤

- 把斜坡的反作用力當作水平方向，就像處理鉛垂牆壁那樣。這裏它向水平線上方傾斜了 $\frac{1}{6}\pi$，而它向上的分量會令 R 減小。
- 沿鉛垂方向分解時忘記了工人的重量，結果寫成 $R + \frac{1}{2}N = mg$ 而不是 $2mg$。

STEP 之外

第 (ii) 部分其實是對各種配置所作的一個小型最優化（optimisation）：你需要找出負載在所有位置之中 F/R 的最壞情況。工程師正是這樣找出結構的臨界負載情況。要想像條件 $F \le \mu R$，一個有用的方法是摩擦錐（cone of friction）：地面的總反作用力必須落在與法線成 $\arctan\mu$ 的角度之內，而當所需的反作用力離開這個錐的一刻，梯子便會滑動。Kleppner 與 Kolenkow 的 *An Introduction to Mechanics*（力學入門教科書）對剛體靜力學和摩擦有仔細的論述。

詳細解答

STEP III 1988，第 12 題

力學 • 斜碰撞

STEP III 1988，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/88-s3-q12

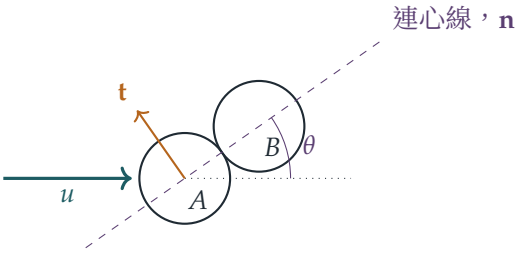
關鍵思路

對光滑球體而言，碰撞瞬間的衝量（impulse）只沿連心線（line of centres）作用。因此我們把每個速度分解為沿連心線的分量和垂直於連心線的分量：垂直分量完全不受影響，而沿連心線方向則是一個普通的一維碰撞。於是一切都取決於一個未知量，即連心線與第一個球原來速度之間的夾角。

解答

建立模型

把兩球視為質量同為 m 的光滑球體，在光滑桌面上滑動，因此碰撞期間唯一的水平力就是沿連心線的接觸衝量。設第一個球 A 以速率 u 運動，並設碰撞時連心線與 A 的速度成角 θ，其中 $0 \le \theta < \frac{1}{2}\pi$（若 $\theta = \frac{1}{2}\pi$，兩球只是擦過，甚麼也不會發生）。取單位向量（unit vector）$\mathbf{n}$ 沿連心線（由 A 指向 B），$\mathbf{t}$ 與之垂直。



碰撞前，A 沿 $\mathbf{n}$ 的分量為 $u\cos\theta$，沿 $\mathbf{t}$ 的分量為 $u\sin\theta$；B 靜止。

沿 $\mathbf{t}$ 方向：兩球在這個方向上都沒有受到衝量，所以 A 保持 $u\sin\theta$，B 保持 0。

*沿 **n** 方向*：設碰撞後的分量為 v_A 和 v_B。由動量守恆（conservation of momentum）（質量互相約去）和牛頓恢復定律（Newton's law of restitution）得

$$v_A + v_B = u\cos\theta, \qquad v_B - v_A = e\,u\cos\theta,$$

所以

$$v_A = \tfrac{1}{2}(1-e)\,u\cos\theta, \qquad v_B = \tfrac{1}{2}(1+e)\,u\cos\theta.$$

因此碰撞後

$$\mathbf{v}_A = \tfrac{1}{2}(1-e)u\cos\theta\,\mathbf{n} + u\sin\theta\,\mathbf{t}, \qquad \mathbf{v}_B = \tfrac{1}{2}(1+e)u\cos\theta\,\mathbf{n}.$$

兩個速度之間的夾角

B 沿連心線運動，而兩個速度之間的夾角為銳角，當且僅當它們的純量積（scalar product）為正。這裏

$$\mathbf{v}_A \cdot \mathbf{v}_B = \tfrac{1}{4}(1-e)(1+e)\,u^2\cos^2\theta > 0,$$

因為 $e < 1$、$u \neq 0$，而且當 $0 \le \theta < \frac{1}{2}\pi$ 時 $\cos\theta > 0$。故此兩個速度之間的夾角小於 $\frac{1}{2}\pi$，這正是所要證明的。（對於正面碰撞，即 $\theta = 0$，兩球沿同一直線運動，夾角為 0。）

留意 $e < 1$ 在哪裏用到：若 $e = 1$，則 $v_A = 0$，兩球恰好成直角分開，這就是大家熟悉的桌球結果。

最大偏轉角

設 ϕ 為 A 的偏轉角（angle of deflection），即 $\mathbf{v}_A$ 與原來方向 $\cos\theta\,\mathbf{n} + \sin\theta\,\mathbf{t}$ 之間的夾角。記 $k = \frac{1}{2}(1-e)$，則 $0 < k \le \frac{1}{2}$，且 $\mathbf{v}_A = u(k\cos\theta\,\mathbf{n} + \sin\theta\,\mathbf{t})$。對方向向量 $(\cos\theta, \sin\theta)$ 和 $(k\cos\theta, \sin\theta)$ 使用純量積和二維叉積（cross product），得

$$|\mathbf{v}_A|\cos\phi = u\,(k\cos^2\theta + \sin^2\theta),$$
$$|\mathbf{v}_A|\sin\phi = u\,(\cos\theta\sin\theta - k\cos\theta\sin\theta) = u\,(1-k)\sin\theta\cos\theta.$$

第一式為正，所以 ϕ 是銳角；兩式相除（再把分子和分母同除以 $\cos^2\theta$），並記 $t = \tan\theta \ge 0$，得

$$\tan\phi = \frac{(1-k)t}{k+t^2}.$$

想一想 怎樣才能不經繁瑣的求導而求出這個式子的最大值？

由於 ϕ 是銳角，求 ϕ 的最大值等同於求 $\tan\phi$ 的最大值。當 $t > 0$ 時，

$$\tan\phi = \frac{1-k}{k/t + t} \quad \text{而} \quad \frac{k}{t} + t \ge 2\sqrt{k}$$

這是由算術平均—幾何平均不等式（AM–GM inequality）得出的，等號成立當且僅當 $t = \sqrt{k}$。（當 $t = 0$ 時，$\phi = 0°$）所以最大值在 $\tan\theta = \sqrt{k}$ 時取得，且

$$\tan\phi_{\max} = \frac{1-k}{2\sqrt{k}}.$$

要轉換成正弦，留意 $(1-k)^2 + (2\sqrt{k})^2 = 1 - 2k + k^2 + 4k = (1+k)^2$，所以在兩條直角邊為 $1-k$ 和 $2\sqrt{k}$ 的直角三角形中，斜邊為 $1+k$，於是

$$\sin\phi_{\max} = \frac{1-k}{1+k} = \frac{\frac{1}{2}(1+e)}{\frac{1}{2}(3-e)} = \frac{1+e}{3-e},$$

當中用到 $1-k = \frac{1}{2}(1+e)$ 和 $1+k = \frac{1}{2}(3-e)$。由於 $\phi_{\max}$ 是銳角，

$$\phi_{\max} = \sin^{-1}\left(\frac{1+e}{3-e}\right),$$

這就是所求的結果。它出現於連心線與原來運動方向成角 $\tan^{-1}\sqrt{\frac{1}{2}(1-e)}$ 之時。

驗算

- $e \to 1$：公式給出 $\sin^{-1}1 = \frac{1}{2}\pi$，與完全彈性（perfectly elastic）的情況一致：此時 A 失去沿連心線的全部速度，所以它總是沿垂直於連心線的方向運動，而一次接近正面的撞擊會使它以接近直角的方向偏離原來的路徑。
- $e = 0$：$\sin\phi_{\max} = \frac{1}{3}$，所以 $\phi_{\max} \approx 0.3398$ rad（19.5°）。
- $e = \frac{1}{2}$：$\sin\phi_{\max} = \frac{3}{5}$，$\phi_{\max} \approx 0.6435$ rad，出現於 $\theta = \tan^{-1}\frac{1}{2} \approx 0.4636$。以數值方法對 200 000 個 θ 值逐一計算偏轉角，結果與這兩個數字吻合；而 200 000 次隨機碰撞中，兩個速度之間的夾角從未達到或超過 $\frac{1}{2}\pi$。

常見錯誤

- 把恢復定律套用於整個速率，而不是沿連心線的分量。
- 忘記 A 的速度的垂直分量保持不變；正是這一點才令 A 發生偏轉。
- 對反正切函數求導來直接求 ϕ 的最大值；改用 $\tan\phi$（因為 ϕ 是銳角，這樣做是合理的）會簡潔得多。

STEP 之外

這裏的分解方法，其實就是牛頓第二定律（Newton's second law）的衝量形式：衝量只會在其作用的方向上改變動量。在彈性碰撞的情況下，直角這個結果是一個巧妙的參考

系（frame of reference）論證的特例：在質心參考系（centre-of-mass frame）中，彈性碰撞保持每個球的速率不變，只改變其方向，而變換回實驗室參考系（lab frame）後，便一次過得出所有可能的結果。Kleppner 與 Kolenkow 所著的 *An Introduction to Mechanics*（力學導論）正是這樣在質心參考系中處理碰撞的。

詳細解答

STEP III 1988，第 13 題

力學 • 有空氣阻力的拋體

STEP III 1988，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/88-s3-q13

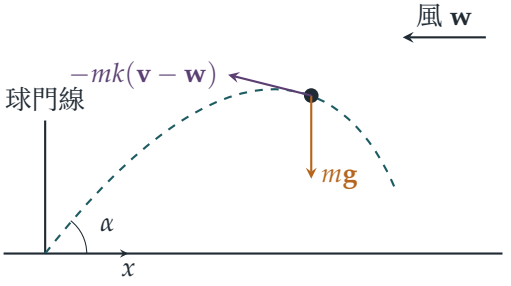
關鍵思路

空氣阻力與速度成正比時，運動方程（equation of motion）是線性的，因此每個分量都可以用積分因子（integrating factor）求解。要看清最後一部分，最簡潔的方法是問：終端速度（terminal velocity）是甚麼？速度會由初始值按指數規律趨向終端速度，所以若這兩個向量平行，速度便永遠不會離開通過原點的同一條直線。於是皮球沿着同一條直線路徑飛出去，再飛回來。

解答

運動方程

我們把皮球看成質量為 m 的質點（particle）。作用於它的力有兩個：重量 $m\mathbf{g}$，以及空氣阻力。皮球相對於空氣的速度是 $\mathbf{v}-\mathbf{w}$，而阻力與這個相對速度方向相反，所以阻力是 $-mk(\mathbf{v}-\mathbf{w})$。



由牛頓第二定律（Newton's second law）得

$$m\frac{\mathrm{d}\mathbf{v}}{\mathrm{d}t} = m\mathbf{g} - mk(\mathbf{v} - \mathbf{w}).$$

兩邊除以 m 並移項，

$$\frac{\mathrm{d}\mathbf{v}}{\mathrm{d}t} + k\mathbf{v} = \mathbf{g} + k\mathbf{w},$$

這正是所要證明的結果。

時間 t 時的位置

取開球點為原點，x 軸為水平方向並指向開球的方向，y 軸垂直向上。記 $V = |\mathbf{v}_0|$、$W = |\mathbf{w}|$ 及 $g = |\mathbf{g}|$。皮球是迎着風踢出的，所以

$$\mathbf{w} = (-W, 0), \qquad \mathbf{g} = (0, -g), \qquad \mathbf{v}_0 = (V\cos\alpha, V\sin\alpha).$$

設 $\mathbf{v} = (u, v)$，運動方程的兩個分量便是

$$\frac{\mathrm{d}u}{\mathrm{d}t} + ku = -kW, \qquad \frac{\mathrm{d}v}{\mathrm{d}t} + kv = -g.$$

把第一式乘以積分因子 e^{kt}，得 $\frac{\mathrm{d}}{\mathrm{d}t}\left(u\,\mathrm{e}^{kt}\right) = -kW\mathrm{e}^{kt}$，所以 $u\,\mathrm{e}^{kt} = -W\mathrm{e}^{kt} + C$。由於 $t = 0$ 時 $u = V\cos\alpha$，故 $C = V\cos\alpha + W$，於是

$$u = -W + (V\cos\alpha + W)\,\mathrm{e}^{-kt}.$$

用完全相同的方法，

$$v = -\frac{g}{k} + \left(V\sin\alpha + \frac{g}{k}\right)\mathrm{e}^{-kt}.$$

再由 0 至 t 積分一次，並利用 $t = 0$ 時 $x = y = 0$ 以及 $\int_0^t \mathrm{e}^{-ks}\,\mathrm{d}s = \frac{1}{k}\left(1 - \mathrm{e}^{-kt}\right)$，得

$$x = -Wt + \frac{V\cos\alpha + W}{k}\left(1 - \mathrm{e}^{-kt}\right),$$

$$y = -\frac{g}{k}\,t + \frac{1}{k}\left(V\sin\alpha + \frac{g}{k}\right)\left(1 - \mathrm{e}^{-kt}\right).$$

寫成向量形式，設 $\mathbf{u}_\infty = \mathbf{w} + \mathbf{g}/k$（即終端速度，亦即運動方程的常數解），上式便是

$$\mathbf{r} = \mathbf{u}_\infty t + \frac{1}{k}\left(\mathbf{v}_0 - \mathbf{u}_\infty\right)\left(1 - \mathrm{e}^{-kt}\right).$$

烏龍球

想一想 $\tan\alpha = g/(kW)$ 這個條件，說明了終端速度的甚麼性質？

終端速度是 $\mathbf{u}_\infty = (-W, -g/k) = -W\left(1, g/(kW)\right)$。條件給出 $g/(kW) = \tan\alpha$，所以 $\mathbf{u}_\infty = -W(1, \tan\alpha)$，它與 $\mathbf{v}_0$ 平行（方向相反）。由向量公式可見，$\mathbf{r}$ 此時始終是兩個沿 $(\cos\alpha, \sin\alpha)$ 方向的向量的組合，所以皮球沿直線 $y = x\tan\alpha$ 運動。

若要直接從分量看出這一點，可利用 $g/k = W\tan\alpha$：

$$V\sin\alpha + \frac{g}{k} = V\sin\alpha + W\tan\alpha = \tan\alpha\,(V\cos\alpha + W),$$

所以

$$y = \tan\alpha\left[-Wt + \frac{V\cos\alpha + W}{k}\left(1 - \mathrm{e}^{-kt}\right)\right] = x\tan\alpha \quad \text{對所有} t \geq 0.$$

因此，皮球位於地面（即 $y = 0$）當且僅當 $x = 0$（因為 $\tan\alpha \neq 0$）。餘下要檢查的，是這情況確實在某個 $t > 0$ 發生。設 $f(t) = x(t)$。則 $f(0) = 0$ 而 $f'(0) = V\cos\alpha > 0$，所以剛踢出之後 $x > 0$。此外

$$f''(t) = -k(V\cos\alpha + W)\,\mathrm{e}^{-kt} < 0,$$

所以 f 是嚴格凹函數（strictly concave），而且由於有 $-Wt$ 一項，當 $t \to \infty$ 時 $f(t) \to -\infty$。因此 f 恰好有一個正零點 T，並且在該處 $f'(T) < 0$（一個由 0 升起的凹函數，在跌回 0 時必定正在遞減）。

在時間 T，皮球回到原點，即球門線上開球的那一點，而且水平速度為負，也就是說正朝着球門移動。它沿着飛出去時的那條直線飛回來，並在地面高度越過球門線，遠低於橫楣。既然守門員不在位，這便是一個烏龍球。

驗算

- 當 $k \to 0$ 時，$\frac{1}{k}(1 - \mathrm{e}^{-kt}) \to t$ 而 $\frac{1}{k}\left(t - \frac{1}{k}(1 - \mathrm{e}^{-kt})\right) \to \frac{1}{2}t^2$，所以 $x \to Vt\cos\alpha$ 而 $y \to Vt\sin\alpha - \frac{1}{2}gt^2$：這正是普通的拋體運動（projectile motion）。
- 當 $t \to \infty$ 時，速度趨向 $(-W, -g/k)$，即運動方程的常數解：皮球最終隨風飄移，同時以終端速率下墜。
- 以龍格－庫塔法（Runge–Kutta）對運動方程作數值積分，在隨機選取多組 V, α, W, k 的情況下，結果與 x 和 y 的公式相符，誤差約為 10^{-13}。當 $\tan\alpha = g/(kW)$ 時，計算所得的路徑一直留在 $y = x\tan\alpha$ 上，並在 $x = 0$ 處着地，而且 $u < 0$；若取例如 $V = 20$、$\alpha = 0.6$、$W = 5$、$k = 0.5$，皮球則落在前場 17.1 m 處。

常見錯誤

- 把阻力取為 $-mk\mathbf{v}$（相對於地面），而不是 $-mk(\mathbf{v} - \mathbf{w})$。
- 弄錯 $\mathbf{w}$ 的正負號：皮球是迎着風踢出的，所以 $\mathbf{w}$ 指向球門的方向。
- 得出 $y = x\tan\alpha$ 便停下來，沒有檢查皮球是否真的回到 $x = 0$，以及回到該處時是否正朝着球門移動。

STEP 之外

方程 $\dot{\mathbf{v}} + k\mathbf{v} = \mathbf{F}$（其中 $\mathbf{F}$ 為常數）是線性系統鬆弛至平衡狀態（equilibrium）的最簡單例子：偏差 $\mathbf{v} - \mathbf{F}/k$ 按 e^{-kt} 的方式衰減。線性阻力只對細小而緩慢的物體才符合實際；足球實際所受的阻力大致與其速率的平方成正比，這令方程變成非線性，並使水平與垂直方向的運動互相耦合。Kleppner 與 Kolenkow 的 *An Introduction to Mechanics*（力學入門教科書）對這兩種阻力都有討論。

詳細解答

STEP III 1988，第 14 題

力學 • 光滑金屬線上的運動

STEP III 1988，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：

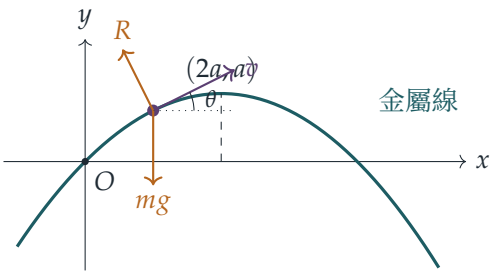
step.maths.org/questions/88-s3-q14

關鍵思路

如果金屬線對小珠完全不施力，小珠就只是一個拋體（projectile），僅在重力作用下運動。所以第 (i) 部分問的是：一個沿金屬線方向拋出的自由拋體，能否剛好沿着金屬線的拋物線（parabola）運動？把兩條拋物線對照起來即可。至於第 (ii) 部分，金屬線是光滑的，所以能量守恆（conservation of energy）。真正要做的功夫，是利用 $\tan\theta = \mathrm{d}y/\mathrm{d}x$，把速率和高度用斜角 θ 表示出來。

解答

我們把小珠視為一個質量為 m 的質點（particle），套在一條固定的光滑金屬線上。作用在它身上的力只有兩個：垂直向下的重量 mg，以及金屬線的法向反作用力（normal reaction）R，後者與金屬線垂直。



以下會反覆用到關於金屬線的兩個事實。在一般的點上，

$$\frac{\mathrm{d}y}{\mathrm{d}x} = 1 - \frac{x}{2a},$$

所以在原點處斜率為 1，小珠出發時的方向與 x-軸成 45° 仰角。由於金屬線是一條開口向下的拋物線，它的最高點位於 $x = 2a, y = a$。

第 (i) 部分

假設在整個運動過程中 $R = 0$。那麼重力就是唯一的力，小珠的運動便與一個以初速率 V、仰角 45° 拋出的拋體完全相同。它的水平速度保持不變，恆為 $V/\sqrt{2}$，所以 $x = Vt/\sqrt{2}$。同時

$$y = \frac{V}{\sqrt{2}}\,t - \tfrac{1}{2}gt^2.$$

以 $t = \sqrt{2}\,x/V$ 消去時間，得到軌跡（trajectory）

$$y = x - \frac{gx^2}{V^2}.$$

只有當這條曲線在小珠所到達的每一個 x 處都與金屬線 $y = x - x^2/(4a)$ 重合，它才是小珠的路徑。這就需要

$$\frac{g}{V^2} = \frac{1}{4a}, \qquad \text{即} \qquad V = 2\sqrt{ag}.$$

反過來，取 $V = 2\sqrt{ag}$。一個從 O 以這個速率、仰角 45° 拋出的自由質點，會沿着曲線 $y = x - x^2/(4a)$ 運動，而這正是金屬線本身。因此它沿着金屬線運動，每一刻都有 $R = 0$，而這個運動在反作用力為零的情況下滿足小珠的運動方程(equation of motion)。由於 $x = Vt/\sqrt{2}$ 穩定地增加，小珠會越過位於 $(2a, a)$ 的頂點，並沿另一邊一直向下運動，永不停止。在整個過程中，它對金屬線不施加任何壓力，符合題目要求，而所求的速率是 $V = 2\sqrt{ag}$。

> **想一想** 套在金屬線上的小珠，怎可能在完全沒有反作用力的情況下運動？金屬線不是要把它留在線上嗎？

只有當單靠重力會令小珠偏離曲線時，金屬線才需要施力。換了任何其他速率，情況正是如此。V 較大時，拋體的拋物線較平坦，所以金屬線必須把小珠向下拉。V 較小時，拋物線彎曲得較厲害，所以金屬線必須把小珠托住。而剛好在 $V = 2\sqrt{ag}$ 時，重力已經提供了曲線所需的加速度，金屬線便無事可做。

第 (ii) 部分

現在 V 是任意的。速度沿着金屬線的切線（tangent）方向，所以

$$\tan\theta = \frac{\mathrm{d}y}{\mathrm{d}x} = 1 - \frac{x}{2a}, \qquad \text{所以} \qquad x = 2a(1 - \tan\theta).$$

(如果小珠向後滑回，它的速度便沿切線指向相反方向，角度改變 π。這時 $\tan\theta$ 和 $\cos^6\theta$ 都不變，所以下文不受任何影響。)

以 θ 表示速率。把 $x = 2a(1 - \tan\theta)$ 對 t 求導，得

$$\dot{x} = -2a\sec^2\theta\,\dot{\theta}.$$

由於 $\dot{y} = \tan\theta\,\dot{x}$，速率 v 滿足

$$v^2 = \dot{x}^2 + \dot{y}^2 = \dot{x}^2(1 + \tan^2\theta) = \dot{x}^2\sec^2\theta = \frac{4a^2\dot{\theta}^2}{\cos^6\theta}.$$

以 θ 表示高度。記 $T = \tan\theta$，於是 $x = 2a(1 - T)$，

$$\begin{aligned} y = x - \frac{x^2}{4a} &= 2a(1-T) - a(1-T)^2 \\ &= a(1-T)\big(2 - (1-T)\big) \\ &= a(1-T)(1+T) = a(1 - \tan^2\theta). \end{aligned}$$

能量。金屬線是光滑的，而 R 與速度垂直，所以 R 不作功，機械能（mechanical energy）守恆。小珠從 O 以速率 V 出發，所以

$$\tfrac{1}{2}mv^2 + mgy = \tfrac{1}{2}mV^2, \qquad \text{即} \qquad v^2 + 2gy = V^2.$$

代入 v^2 和 y 的表達式，得

$$\frac{4a^2\dot{\theta}^2}{\cos^6\theta} + 2ga(1 - \tan^2\theta) = V^2,$$

這就是所要證明的結果。

驗算

- 在 $t = 0$ 時，我們有 $\theta = 45°$，所以 $\tan^2\theta = 1$，第二項消失。這時第一項就是 $v^2 = V^2$，與預期相符。
- 取第 (i) 部分的 $V^2 = 4ag$，則 $\dot{x} = V/\sqrt{2}$ 為常數，所以 $\dot{\theta} = -\dot{x}\cos^2\theta/(2a)$。第一項變成 $2ag\sec^2\theta = 2ag(1 + \tan^2\theta)$。加上 $2ga(1 - \tan^2\theta)$ 便得 $4ag = V^2$。兩部分的結果互相吻合。
- 數值方面，我們用龍格－庫塔法（Runge–Kutta method）沿金屬線逐步計算小珠的運動方程。取 $a = 1$、$g = 9.8$ 和 $V = 2\sqrt{9.8}$，算出的法向反作用力（以每單位質量計）一直到 $x \approx 13$ 都保持在 10^{-15} 以下。把 V 增加 30% 後，它在起點處約為 4.8。對於隨機選取的 V 值，包括一些令小珠向後滑回的情況，第 (ii) 部分的恆等式都成立，相對誤差約為 10^{-13}。

常見錯誤

- 只證明了在拋出的一刻 $R = 0$。題目要求的是一個自始至終都沒有反作用力的運動，這正是我們要把整條軌跡與金屬線對照的原因。
- 對 $\tan\theta$ 求導時，忘記了鏈式法則（chain rule）所帶來的因子 $\sec^2\theta$。這樣第一項中便會出現 $\cos^2\theta$，而不是 $\cos^6\theta$。

STEP 之外

第 (ii) 部分是這個運動的一個首次積分（first integral）：能量守恆把一條二階方程化為只含單一坐標（這裏是 θ）的一階方程。在大學一年級的課程中，這會發展成處理約束運動（constrained motion）的拉格朗日（Lagrangian）方法：你只需選取一個能描述約束的坐標，完全不必計算反作用力。Kleppner 與 Kolenkow 合著的 *An Introduction to Mechanics*（力學入門教科書）對約束運動和能量方法有細緻的處理，很值得一讀。

詳細解答

STEP III 1988，第 15 題

概率與統計 • 期望值的遞推關係

STEP III 1988，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/88-s3-q15

關鍵思路

對當天的第一本書取條件（conditioning）。這本書一旦上了架，當天餘下的過程就與書架較短的全新一天完全一樣，所以長度為 n 的書架的期望（expected）本數，與長度為 $n-1$ 和 $n-2$ 時的期望本數有關。由此得到一條線性差分方程（difference equation），解法與二階遞推關係（recurrence）一樣：齊次（homogeneous）部分用輔助二次方程（auxiliary equation）處理，再為右邊的常數 1 找一個特解（particular solution）。

解答

以下一律設 N_n 為書架長度為 n 的一天內放上書架的書本數目，因此 $M(n) = \mathrm{E}(N_n)$。各本書是厚是薄互相獨立（independent），每本為厚書的概率是 p。題目說明書本數量總是多於書架所能容納的，所以每天最終必定有某本書放不下。

第 (i) 部分

當 $n = 0$ 時，書架完全沒有空位。第一本書不論厚薄都放不下，所以它和其後每一本書都要放到手推車上。因此 $N_0 = 0$ 恆成立，而 $M(0) = 0$。

第 (ii) 部分

當 $n = 1$ 時，考慮第一本書。

- 若它是薄書（概率為 q），它放得下，並剛好放滿書架。下一本書至少需要一個單位，所以放不下，自此所有書都放到手推車上。故 $N_1 = 1$。

- 若它是厚書（概率為 p），它需要兩個單位，放不下，故 $N_1 = 0$。

因此 $M(1) = q \cdot 1 + p \cdot 0 = q$。

第 (iii) 部分

設 $n \geq 2$，同樣對第一本書取條件。不論它是哪一種書都放得下，因為它最多只需要 $2 \leq n$ 個單位。

- 若它是薄書（概率為 q），它上架後餘下 $n-1$ 個單位。
- 若它是厚書（概率為 p），它上架後餘下 $n-2$ 個單位。

> **想一想** 為甚麼在第一種情況下，其後上架書本數目的期望值恰好是 $M(n-1)$？

第一本書之後，其餘的書依次到來，它們是厚是薄與第一本書獨立，概率亦相同，而且按同一規則擺放：一直上架，直到有一本放不下為止。第一本書唯一改變了的是剩餘空間，現在是 $n-1$ 個單位。所以當天餘下的過程，就是書架長度為 $n-1$ 的完整一天的翻版，其後上架書本數目的期望值是 $M(n-1)$。同理，若第一本是厚書，這個期望值就是 $M(n-2)$。（根據第 (i) 部分，當 $n-2=0$ 時這也成立。）

把第一本書也計算在內，由全期望定律（law of total expectation）得

$$\begin{aligned} M(n) &= q\big(1 + M(n-1)\big) + p\big(1 + M(n-2)\big) \\ &= (p+q) + qM(n-1) + pM(n-2), \end{aligned}$$

而由於 $p+q=1$，

$$M(n) - qM(n-1) - pM(n-2) = 1 \qquad (n \geq 2),$$

這正是所要證明的。

閉合公式

先考慮齊次方程 $M(n) - qM(n-1) - pM(n-2) = 0$。代入 $M(n) = \lambda^n$ 試解，得輔助方程

$$\lambda^2 - q\lambda - p = 0.$$

由於 $q = 1-p$，$\lambda = 1$ 是一個根（$1 - q - p = 0$），而兩根之積為 $-p$，所以

$$\lambda^2 - q\lambda - p = (\lambda - 1)(\lambda + p).$$

因此對任意 A 和 B，$A(-p)^n$ 和常數 B 都滿足齊次方程。

因為 1 是一個根，常數無法產生右邊的 1，所以我們改試 Cn。代入得

$$\begin{aligned} Cn - qC(n-1) - pC(n-2) &= C\big[(1-q-p)n + q + 2p\big] \\ &= C(q+2p) = C(1+p). \end{aligned}$$

當 $C = \dfrac{1}{1+p}$ 時，上式等於 1；由於 $1+p > 0$，這個取值沒有問題。於是由線性性質（linearity），不論 A 和 B 取何值，$M(n) = A(-p)^n + B + Cn$ 對所有 $n \geq 2$ 都滿足 (iii) 中的方程。

現在代入初始值。由第 (i) 部分得

$$M(0) = A + B = 0, \qquad \text{所以} B = -A,$$

而由第 (ii) 部分得

$$M(1) = -Ap + B + C = -A(1+p) + \frac{1}{1+p} = 1 - p.$$

因此

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

所以

$$A = \frac{p^2}{(1+p)^2}, \qquad B = -\frac{p^2}{(1+p)^2}, \qquad C = \frac{1}{1+p}.$$

即

$$M(n) = \frac{n}{1+p} - \frac{p^2}{(1+p)^2}\big(1 - (-p)^n\big).$$

由於 (i)、(ii) 和 (iii) 確定了 $M(0)$、$M(1)$，繼而逐一確定其後每一個值，所以這不僅是一個可能的解，更是唯一的解。

驗算

- $p = 0$（全是薄書）：$M(n) = n$，理應如此。
- $p = 1$（全是厚書）：$M(n) = \frac{n}{2} - \frac{1}{4}\big(1 - (-1)^n\big)$，當 n 為偶數時是 $\frac{n}{2}$，當 n 為奇數時是 $\frac{n-1}{2}$，即 $\lfloor n/2 \rfloor$。正確。
- 取 $p = q = \frac{1}{2}$，直接由遞推關係計算：$M(2) = 1 + \frac{1}{4} = \frac{5}{4}$，而公式給出 $\frac{2}{3} - \frac{1}{9} \cdot \frac{3}{4} = \frac{5}{4}$。
- 當 n 很大時，每本書平均佔用 $1+p$ 個單位，所以我們預期大約有 $n/(1+p)$ 本書；公式與此吻合，另加一個有界的修正項，來自書架末端浪費掉的空間。

我們已把這個閉合公式與上架過程的精確枚舉結果核對（$n \leq 12$，並取多個 p 值），亦以模擬作了核對；例如當 $p = \frac{1}{2}$ 及 $n = 5$ 時，兩者都給出 $M(5) = \frac{103}{32} \approx 3.219$。

常見錯誤

- 以為厚書放不下之後，薄書仍然可以上架。題目說明其後所有書都放到手推車上，這正是 $M(1) = q$ 而不是更大的數值的原因。
- 以常數作為特解來試：因為 $\lambda = 1$ 是輔助方程的解，常數會被齊次部分吸收，所以需要用 Cn。

STEP 之外

對第一步取條件，並留意到餘下的部分是同一問題的較小翻版，這種方法稱為*首步分析*（first-step analysis）。它是在馬可夫鏈（Markov chain）中求期望首達時間（hitting time）和吸收概率（absorption probability）的標準方法。我們的答案亦說明了*更新定理*（renewal theorem）：若把大小隨機的物件裝進長度 n 之內，能裝下的數目的增長速度與 n 除以平均大小相若。Grimmett 與 Welsh 的 *Probability: An Introduction*（概率入門教科書）既涵蓋隨機漫步（random walk）的差分方程，亦有馬可夫鏈的入門介紹。

詳細解答

STEP III 1988，第 16 題

概率與統計 • 超幾何分佈

STEP III 1988，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/88-s3-q16

關鍵思路

不放回抽取（drawing without replacement）令各次抽取互相依賴，逐次分析會十分繁瑣。有兩個想法可以繞過這個困難。第一，數算不計次序的選法，從而得到紅球數目的分佈。第二，利用對稱性（symmetry）：假想把整個甕以隨機次序全部抽空，則每個球出現在任何指定位置的機會均等。把 X 寫成指示變量（indicator）X_i 之和，便可求得期望值（mean）和方差（variance），完全不需要知道分佈。我們只需要各對變量之間的協方差（covariance）。

解答

恰好有 k 個紅球的概率

把各個球視為可區分的。抽出的 n 個球組成一個大小為 n 的集合，而每個這樣的集合出現的機會均等，所以共有 $\binom{M}{n}$ 個等可能的結果。要組成一個恰好含 k 個紅球的集合，須從 m 個紅球中選 k 個，並從 $M-m$ 個綠球中選 $n-k$ 個。因此

$$\mathrm{P}(\text{恰好}k\text{ 個紅球}) = \frac{\binom{m}{k}\binom{M-m}{n-k}}{\binom{M}{n}},$$

其中沿用慣常的約定：若 $b<0$ 或 $b>a$，則 $\binom{a}{b}=0$。（這已涵蓋不可能發生的情況，例如 $n-k>M-m$。）這就是超幾何（hypergeometric）分佈。

第 (i) 部分

假想一直抽下去，直至甕中再沒有球為止。M 個（可區分的）球被抽出的次序是一個均勻隨機的排列：全部 $M!$ 種次序出現的機會均等。抽了 n 次便停下來並不會改變首 n 個位置，所以我們可以用完整的排列來處理。

固定一個位置 i。某個指定的球位於位置 i 的排列共有 $(M-1)!$ 個，這個數目對每個球都相同。所以位置 i 上的球是 M 個球中任何一個的機會均等，而

$$\mathrm{P}(X_i = 1) = \frac{m\,(M-1)!}{M!} = \frac{m}{M},$$

正是所求。

想一想 第二個球的顏色，不是明明取決於第一個球是甚麼顏色嗎？

在已知第一個球的情況下，確實如此。但我們要求的是無條件（unconditional）概率。當 $i = 2$ 時，由全概率定律（law of total probability）可得

$$\mathrm{P}(X_2 = 1) = \frac{m}{M}\cdot\frac{m-1}{M-1} + \frac{M-m}{M}\cdot\frac{m}{M-1} = \frac{m(M-1)}{M(M-1)} = \frac{m}{M}.$$

對稱性論證則一次過為每個 i 完成了同樣的計算。

第 (ii) 部分

固定兩個位置 $i \neq j$（所以 $M \geq 2$）。數算這兩個位置上都是紅球的排列。位置 i 上的球有 m 個選擇，然後位置 j 有 $m-1$ 個選擇，其餘的球則有 $(M-2)!$ 種排法。因此

$$\mathrm{P}(X_i = 1 \text{ 及} X_j = 1) = \frac{m(m-1)(M-2)!}{M!} = \frac{m(m-1)}{M(M-1)},$$

正是所求。換句話說，位置 (i, j) 上的有序球對，是 $M(M-1)$ 個由相異的球組成的有序對中任何一個的機會均等。

X 的期望值和方差

X 數算首 n 個抽出的球當中紅球的數目。根據期望值的線性性質（linearity of expectation）——這並不需要獨立性（independence）——再加上 $\mathrm{E}(X_i) = \mathrm{P}(X_i = 1)$，

$$\mathrm{E}(X) = \sum_{i=1}^{n} \mathrm{E}(X_i) = \frac{nm}{M}.$$

至於方差，可用 $\mathrm{Var}(X) = \mathrm{E}(X^2) - [\mathrm{E}(X)]^2$。把和的平方展開，並利用 $X_i^2 = X_i$（因為 X_i 只取 0 或 1），

$$X^2 = \sum_{i=1}^{n} X_i^2 + \sum_{i\neq j} X_i X_j = X + \sum_{i\neq j} X_i X_j,$$

其中第二個和遍歷滿足 $i \neq j$ 的 $n(n-1)$ 個有序對 (i,j)。而 $X_iX_j = 1$ 當且僅當兩個球都是紅球，所以由第 (ii) 部分得 $\mathrm{E}(X_iX_j) = \dfrac{m(m-1)}{M(M-1)}$。因此，當 $M \geq 2$ 時，

$$\begin{aligned}\mathrm{Var}(X) &= \frac{nm}{M} + \frac{n(n-1)m(m-1)}{M(M-1)} - \frac{n^2m^2}{M^2} \\ &= \frac{nm}{M^2(M-1)}\Big[M(M-1) + (n-1)(m-1)M - nm(M-1)\Big].\end{aligned}$$

方括號內的式子展開後為

$$\begin{aligned}&M^2 - M + nmM - nM - mM + M - nmM + nm \\ &\quad = M^2 - nM - mM + nm = (M-m)(M-n).\end{aligned}$$

因此

$$\mathrm{Var}(X) = \frac{nm(M-m)(M-n)}{M^2(M-1)}.$$

（若 $M = 1$，則 $n \leq 1$，根本沒有球對，而 X 是常數，所以 $\mathrm{Var}(X) = 0$。）

設 $p = m/M$ 會令結果更具啟發性。這時

$$\mathrm{E}(X) = np, \qquad \mathrm{Var}(X) = np(1-p)\cdot\frac{M-n}{M-1}.$$

這正是二項分佈（binomial distribution）的方差乘以一個修正因子，而當 $n \geq 2$ 時，這個因子小於 1。

驗算

- $n = M$：我們把所有球都抽出，所以必定有 $X = m$。公式給出期望值 m 和方差 0，後者源於因子 $M - n$。
- $n = 1$：X 是單獨一個伯努利（Bernoulli）變量，其 $p = m/M$，而方差化簡為 $p(1-p)$。
- $M = 6$、$m = 2$、$n = 3$：公式給出 $\mathrm{Var}(X) = \frac{3\cdot2\cdot4\cdot3}{36\cdot5} = \frac{2}{5}$。把全部 6! 種次序逐一列舉，結果完全吻合；$M \leq 6$ 的每一個情況亦然。

常見錯誤

- 把各個 X_i 當作互相獨立，於是得出二項分佈的方差 $np(1-p)$。其實它們是負相關（negatively correlated）的：早段抽到紅球，會令其後抽到紅球的機會降低。
- 忘記了當球對是有序的時候，二重和 $\sum_{i\neq j}$ 共有 $n(n-1)$ 項，而不是 $\binom{n}{2}$ 項。

STEP 之外

把一個計數寫成指示變量之和，是概率論中最強大的技巧之一。即使分佈本身根本無從寫出，這個方法仍能給出期望值和方差。因子 $\frac{M-n}{M-1}$ 就是抽樣調查中所用的有限總體修正（finite population correction）。它說明了為何一個訪問 1000 人的民意調查，對一個國家所能提供的資訊，與對一個無限總體所能提供的幾乎一樣多。Blitzstein 與 Hwang 所著的 *Introduction to Probability*（概率論入門教材）全書貫徹使用指示變量，並詳細討論超幾何分佈。

Beyond STEP

STEP 是大多數中學生一生中應考過最艱深的數學考試，也是最出色的考試之一：題目簡短、陌生而深刻，當中最精彩的，其實是喬裝了的大學數學小品。

本書精選最有趣的 STEP 試題，由 1987 年首次開考之前的樣本試卷（specimen papers），到 2025 年的試卷，並以每一道題目作為一扇門。每道題目都會仔細解答，沿途提出優秀學生會問的問題，然後帶你穿過這扇門：這個想法從何而來，到了大學會演變成甚麼，以及接下來該讀些甚麼。

本書為正在準備 STEP、有天分又有志投身數學的學生而寫。他們想要的不止是模擬試卷：他們想開始像數學系本科生那樣思考，並一探前方有甚麼正等着自己。

- 共 13 章，涵蓋純數學（pure mathematics）、力學（mechanics）、概率（probability）與統計（statistics）
- 136 道 STEP 歷屆試題，每題均附引導式解答
- 每道題目背後的大學數學
- 附答案的延伸練習，以及延伸閱讀

1988 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。