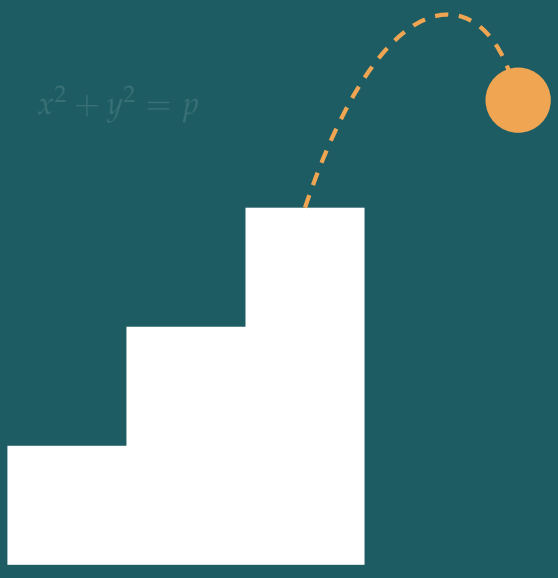
STEP 備試 · 大學數學

Beyond STEP

STEP I, II 及 III

2008 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

2008 年詳細解答

Keith McNulty

Beyond STEP：以歷屆 STEP 試題為門，走進大學數學

2008 年詳細解答

2008 年 STEP 試卷（STEP I, II 及 III）全部 39 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 反證法 純數學 . . . 3

第 2 題 代換積分法 純數學 . . . 7

第 3 題 以重排法證明不等式 純數學 . . . 10

第 4 題 凸函數 純數學 . . . 13

第 5 題 多項式的界限 純數學 . . . 17

第 6 題 反函數與面積 純數學 . . . 21

第 7 題 旋轉與距離 純數學 . . . 25

第 8 題 克萊羅微分方程 純數學 . . . 29

第 9 題 滾動圓環的動量 力學 . . . 33

第 10 題 拋體運動 力學 . . . 36

第 11 題 極限摩擦力平衡 力學 . . . 39

第 12 題 離散均勻分佈變量的最大值 概率與統計 . . . 43

第 13 題 環形排列 概率與統計 . . . 46

STEP II

第 1 題 複數迭代 純數學 . . . 50

第 2 題 部分分式與級數 純數學 . . . 54

第 3 題 不等式與反證法 純數學 . . . 58

第 4 題 隱函數微分法 純數學 . . . 62

第 5 題 代換積分法 純數學 . . . 66

第 6 題 和化積公式 純數學 . . . 70

第 7 題 一階微分方程 純數學 . . . 74

第 8 題 向量與角平分線 純數學 . . . 78

第 9 題 拋體運動 力學 . . . 81

第 10 題 牆壁的斜碰撞 力學 85
第 11 題 移動楔上的質點 力學 89
第 12 題 獨立事件 概率與統計 93
第 13 題 取放回抽取 概率與統計 96

STEP III

第 1 題 非線性聯立方程 純數學 100
第 2 題 冪和 純數學 103
第 3 題 橢圓的切線與焦點 純數學 107
第 4 題 反雙曲函數不等式 純數學 111
第 5 題 遞推關係 純數學 115
第 6 題 一階微分方程 純數學 118
第 7 題 複數與幾何 純數學 121
第 8 題 級數與遞推關係 純數學 125
第 9 題 摩擦力與變力 力學 129
第 10 題 彈性繩與能量 力學 133
第 11 題 轉動動力學 力學 137
第 12 題 生成函數 概率與統計 140
第 13 題 指示變量之和 概率與統計 144

詳細解答

STEP I 2008，第 1 題

純數學 • 反證法

> **STEP I 2008，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/08-s1-q1

關鍵思路

有理數（rational number）在加法、減法、乘法以及除以非零數的運算下是封閉（closed）的。本題幾乎所有結論都源自這一個事實：如果某些數是有理數，那麼用這些運算由它們構造出來的任何數也是有理數。在最後部分，假設四個數之中有兩個是有理數，把它們巧妙地組合起來，從而得出一個題目已告訴我們是無理數（irrational number）的數。

解答

定義

若實數 x 可以寫成 $x = \frac{m}{n}$，其中 m 和 n 為整數且 $n \neq 0$，則稱它為*有理數*。若一個實數不是有理數，即不存在整數 m 和 $n \neq 0$ 使得 $x = \frac{m}{n}$，則稱它為*無理數*。

命題 A

用反證法（proof by contradiction）：假設 pq 是無理數，但 p 和 q 都不是無理數。那麼兩者都是有理數，所以可以寫成 $p = \frac{a}{b}$ 和 $q = \frac{c}{d}$，其中 a, b, c, d 為整數且 $b, d \neq 0$。於是

$$pq = \frac{ac}{bd},$$

其中 ac 和 bd 都是整數，而且 $bd \neq 0$，因為 b 和 d 都不是零。所以 pq 是有理數，這與我們的假設矛盾。因此 p 和 q 之中至少有一個是無理數。

命題 B

同樣地，假設 $p+q$ 是無理數，但 $p=\frac{a}{b}$ 和 $q=\frac{c}{d}$ 都是有理數，其中 $b,d\neq 0$。於是

$$p+q=\frac{ad+bc}{bd},$$

這是兩個整數之比，而且分母不為零，所以 $p+q$ 是有理數。這個矛盾說明 p 和 q 之中至少有一個是無理數。

命題 C

取 $p=\sqrt{2}$ 和 $q=-\sqrt{2}$。兩者都是無理數，但 $p+q=0=\frac{0}{1}$ 是有理數，所以 C 不成立。

為完整起見，以下說明為何 $\sqrt{2}$ 是無理數。若 $\sqrt{2}=\frac{m}{n}$ 為最簡分數（lowest terms），則 $m^2=2n^2$，所以 m^2 是偶數，因而 m 也是偶數（奇數的平方是奇數）。設 $m=2k$，便得 $n^2=2k^2$，所以 n 也是偶數，這與最簡分數的假設矛盾。至於 $-\sqrt{2}$ 也是無理數，因為若 $-\sqrt{2}=\frac{m}{n}$，則 $\sqrt{2}=\frac{-m}{n}$。

最後部分

題目告訴我們 π、π^2 和 $\mathrm{e}\pi$ 都是無理數（我們只需要這三個事實）。以下兩個簡單的觀察會反覆用到：

- 若 r 和 s 是有理數，則 $r+s$、$r-s$、rs 也是有理數，而 r/s 在 $s\neq 0$ 時亦然（論證與 A 和 B 所用的完全相同）；
- 因此，若 r 是有理數，則 $\frac{1}{2}r$ 也是有理數，因為 $\frac{1}{2}$ 是有理數。

另外留意 $\pi+\mathrm{e}>0$ 和 $\pi-\mathrm{e}>0$，因為 $\pi>3>\mathrm{e}>0$；特別地，兩者都不是零。

用反證法：假設以下四個數

$$\pi+\mathrm{e},\quad \pi-\mathrm{e},\quad \pi^2-\mathrm{e}^2,\quad \pi^2+\mathrm{e}^2$$

之中有兩個是有理數。共有六種可能的配對，我們逐一排除。

1. $\pi+\mathrm{e}$ 和 $\pi-\mathrm{e}$ 爲有理數。兩者相加，2π 是有理數，所以 π 是有理數。矛盾。
2. $\pi^2-\mathrm{e}^2$ 和 $\pi^2+\mathrm{e}^2$ 爲有理數。兩者相加，$2\pi^2$ 是有理數，所以 π^2 是有理數。矛盾。
3. $\pi+\mathrm{e}$ 和 $\pi^2+\mathrm{e}^2$ 爲有理數。把前者平方，$(\pi+\mathrm{e})^2$ 是有理數，而

   $$2\mathrm{e}\pi=(\pi+\mathrm{e})^2-(\pi^2+\mathrm{e}^2),$$

   所以 $\mathrm{e}\pi$ 是有理數。矛盾。

4. $\pi - e$ 和 $\pi^2 + e^2$ 爲有理數。同理，

$$2e\pi = (\pi^2 + e^2) - (\pi - e)^2$$

是有理數，所以 $e\pi$ 是有理數。矛盾。

5. $\pi + e$ 和 $\pi^2 - e^2$ 爲有理數。由於 $\pi + e \neq 0$，

$$\pi - e = \frac{\pi^2 - e^2}{\pi + e}$$

是兩個有理數之商，而且分母不為零，所以它是有理數。這樣 $\pi + e$ 和 $\pi - e$ 都是有理數，而根據情況 1，這是不可能的。

6. $\pi - e$ 和 $\pi^2 - e^2$ 爲有理數。由於 $\pi - e \neq 0$，

$$\pi + e = \frac{\pi^2 - e^2}{\pi - e}$$

是有理數，於是再次由情況 1 得出矛盾。

想一想 為甚麼在情況 6 中需要 $\pi - e \neq 0$？

因為不可以除以零：「有理數除以有理數仍是有理數」這一步，只在分母不為零時才成立。假如 $\pi = e$，則 $\pi^2 - e^2 = 0$ 是有理數，卻不能告訴我們任何關於 $\pi + e$ 的資訊。這裏 $\pi - e \approx 0.42$，所以沒有問題。

每一種配對都導致矛盾，所以四個數之中不可能有兩個同時是有理數。因此其中至多有一個是有理數，這正是所要證明的。

驗算

所用的恆等式很容易用數值驗證：取 $\pi \approx 3.14159$ 和 $e \approx 2.71828$，可得 $(\pi + e)(\pi - e) \approx 5.85987 \times 0.42331 \approx 2.48055 \approx \pi^2 - e^2$，以及 $(\pi + e)^2 - (\pi^2 + e^2) \approx 17.0795 \approx 2e\pi$。也數一數配對的數目：四個數共有 $\binom{4}{2} = 6$ 種配對，而六種都已處理。留意我們從未用過 e 和 e^2 是無理數這一點；根據對稱性，在情況 1 和 2 中本來也可以用它們來代替 π 和 π^2。

常見錯誤

- 誤證了 A 或 B 的逆命題（converse）。A 說的是「乘積為無理數 $\Rightarrow$ 其中一個因數為無理數」；它的逆命題並不成立（$\sqrt{2} \cdot \sqrt{2} = 2$）。
- 除以 $\pi \pm e$ 時沒有指出它不為零。

- 把命題 C 當作成立來使用，例如聲稱「π 和 e 是無理數，所以 $\pi + e$ 是無理數」。事實上，$\pi + e$ 是否無理數，至今仍是一個未解決的問題（open problem）。

STEP 之外

這裏用到的封閉性質，說明有理數構成一個域（field）：在這樣的集合裏，你可以進行加、減、乘、除（除以零除外）而不會離開這個集合。域是大學一年級代數課程的核心對象，而「$\pi + e$ 是否無理數？」這類問題則通向超越數論（transcendence theory），這個領域中已知的結果少得驚人。代數方面，可以讀 Allenby 的 *Rings, Fields and Groups*（環、域與羣）；至於有理數與無理數的數論，可以讀 Davenport 的 *The Higher Arithmetic*（高等算術）。

詳細解答

STEP I 2008，第 2 題

純數學 • 代換積分法

> **STEP I 2008，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/08-s1-q2

關鍵思路

這個代換（substitution）的用意，是令平方根變成簡單的 $t - x$。我們不必先解出 x 再對一個繁複的商求導，而是先平方消去根號，再對所得的關係式作隱函數求導（implicit differentiation）。所得的導數（derivative）恰好含有因子 $t - x$，可與被積函數（integrand）中的平方根互相約去，於是積分簡化成 $\int \frac{1}{t+b}\,\mathrm{d}t$。最後一部分則是考驗這個約簡和除法在甚麼情況下才算合理。

解答

導數

設 $R = \sqrt{x^2 + 2bx + c}$，則 $t - x = R$。兩邊平方，

$$t^2 - 2tx + x^2 = x^2 + 2bx + c \quad \Longrightarrow \quad t^2 - 2tx - 2bx - c = 0.$$

把上式對 t 求導（視 x 為 t 的函數），

$$2t - 2x - 2t\frac{\mathrm{d}x}{\mathrm{d}t} - 2b\frac{\mathrm{d}x}{\mathrm{d}t} = 0, \qquad \text{所以} \qquad (t+b)\frac{\mathrm{d}x}{\mathrm{d}t} = t - x.$$

要除以 $t + b$，必須確定它不等於零。由於 $b^2 < c$，

$$x^2 + 2bx + c = (x+b)^2 + (c - b^2) > (x+b)^2,$$

所以 $R > |x + b|$，因此

$$t + b = (x + b) + R > (x + b) + |x + b| \geq 0.$$

故此對任何 x 都有 $t+b>0$，而且

$$\frac{\mathrm{d}x}{\mathrm{d}t}=\frac{t-x}{t+b},$$

這正是所求的結果。（平方後的關係式也可以直接解出：$x=\dfrac{t^2-c}{2(t+b)}$，對此式求導，經過一些代數運算後會得到相同的答案。）

積分

二次式 $x^2+2bx+c$ 對所有 x 都是正數，所以被積函數處處有定義。作代換，並利用 $R=t-x$，

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

其中因為 $t+b>0$，所以不需要加上絕對值（modulus）符號。約去 $t-x=R$ 是可以的，因為 $R>0$。換回 x，

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

你可以用求導來驗證：$x+b+R$ 的導數是 $1+\dfrac{x+b}{R}=\dfrac{R+x+b}{R}$，再除以 $x+b+R$ 便剩下 $\dfrac{1}{R}$。

$b^2=c$ 的情況

這時 $x^2+2bx+c=(x+b)^2$，所以 $\sqrt{x^2+2bx+c}=|x+b|$，被積函數是 $\dfrac{1}{|x+b|}$（在 $x\neq -b$ 時有定義）。

? **想一想** 為甚麼 $\sqrt{(x+b)^2}$ 不能簡單地寫成 $x+b$？

平方根符號所指的是非負的那個根，所以 $\sqrt{(x+b)^2}=|x+b|$，當 $x+b<0$ 時它等於 $-(x+b)$。這個正負號正正決定了公式是否仍然成立。

若 $x+b>0$：直接計算，

$$\int\frac{1}{x+b}\,\mathrm{d}x=\ln(x+b)+C.$$

公式給出

$$\ln\big(x+b+(x+b)\big)=\ln 2+\ln(x+b),$$

它與直接計算的答案只相差常數 $\ln 2$。所以結果仍然成立。

若 $x+b<0$：直接計算，

$$\int \frac{1}{-(x+b)}\,\mathrm{d}x = -\ln\big(-(x+b)\big)+C = -\ln|x+b|+C.$$

但公式給出 $\ln\big(x+b+|x+b|\big) = \ln\big((x+b)-(x+b)\big) = \ln 0$，這是沒有定義的。所以結果在這裏不成立。

值得看看推導是在哪一步出錯的。當 $x+b<0$ 時，對任何 x 都有 $t = x-(x+b) = -b$：t 是常數，根本不能用作新的變量，而 $t+b=0$ 正是我們用來作除數的那個量。嚴格不等式（strict inequality）$b^2<c$ 正是保證 $t+b>0$ 的條件。當 $b^2=c$ 時，只剩下非嚴格的不等式 $t+b\geq 0$，而等號恰好在 $x+b\leq 0$ 時成立。

驗算

- 取 $b=0$、$c=1$，結果變成 $\int \frac{1}{\sqrt{x^2+1}}\,\mathrm{d}x = \ln\big(x+\sqrt{x^2+1}\big)+C = \operatorname{arsinh} x + C$，即標準結果。
- 數值上，隨機選取 b 和 $c>b^2$，在隨機選取的上下限之間以辛普森法則（Simpson's rule）計算積分，所得的值與對數公式之差吻合，誤差小於 10^{-7}。至於 $b^2=c$ 而 $b=1.3$ 的情況，$\int_{0.5}^{4}\frac{\mathrm{d}x}{|x+b|}\approx 1.07992$，與公式給出的 $\ln\frac{5.3}{1.8}$ 相符。

常見錯誤

- 寫成 $\ln|t+b|$ 而沒有留意到 $t+b>0$，因而看不出 $b^2=c$ 時出了甚麼問題。
- 在最後一部分把 $\sqrt{(x+b)^2}=x+b$ 當作成立，以致完全掩蓋了 $x+b<0$ 時公式失效的情況。

STEP 之外

代換 $t=x+\sqrt{x^2+2bx+c}$ 是歐拉代換（*Euler's substitutions*）的其中一種，這類代換能把任何由 x 和 $\sqrt{ax^2+bx+c}$ 組成的有理函數（rational function）的積分，化為 t 的有理函數的積分。最後一部分說明了為甚麼變量代換（change of variable）需要新變量平滑地、而且一對一（one-to-one）地依賴於舊變量：當 t 是常數時，這個方法便完全失效。Bartle 與 Sherbert 所著的 *Introduction to Real Analysis*（實分析入門）對變量代換定理有嚴謹的表述，並精確地列明了這些假設。

詳細解答

STEP I 2008，第 3 題

純數學 • 以重排法證明不等式

> **STEP I 2008，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/08-s1-q3

關鍵思路

不等式（inequality）$(*)$ 其實只是兩個非負因子之積的另一種寫法。一旦證明了它，其後每一個不等式都只在於如何選取 a、b、c、d，使 $(*)$ 的兩邊變成我們想要的兩邊，同時確保條件 $c \geq a$ 和 $d \geq b$ 確實成立。由於要證明的不等式都是對稱（symmetric）的，我們可以先把各變數按大小排好次序，而正是這個次序提供了上述條件。

解答

證明 $(*)$

把所有項移到同一邊，然後因式分解（factorise）：

$$ab + cd - bc - ad = a(b-d) - c(b-d) = (c-a)(d-b).$$

若 $c \geq a$ 且 $d \geq b$，兩個因子皆為非負，故其積亦為非負。因此 $ab + cd \geq bc + ad$，而等號成立當且僅當 $c = a$ 或 $d = b$。

第 (i) 部分

兩個變數。設 $x \geq y$。在 $(*)$ 中取

$$a = y, \quad b = y, \quad c = x, \quad d = x.$$

條件 $c \geq a$ 和 $d \geq b$ 所說的都是 $x \geq y$，這正是已知條件。於是 $ab + cd = y^2 + x^2$ 而 $bc + ad = xy + xy$，所以由 $(*)$ 得

$$x^2 + y^2 \geq 2xy.$$

三個變數。現在再假設 $x \geq z$ 且 $y \geq z$。取

$$a = z, \quad b = z, \quad c = x, \quad d = y.$$

這裏 $c \geq a$ 即 $x \geq z$，$d \geq b$ 即 $y \geq z$，兩者皆為已知。此時 $ab + cd = z^2 + xy$ 而 $bc + ad = zx + zy$，所以

$$z^2 + xy \geq xz + yz.$$

把這個不等式與 $x^2 + y^2 \geq 2xy$ 相加（由於 $x \geq y$，後者可以使用）：

$$x^2 + y^2 + z^2 + xy \geq 2xy + xz + yz,$$

再從兩邊減去 xy，得

$$x^2 + y^2 + z^2 \geq xy + yz + zx.$$

到目前為止，這個結果只在 $x \geq y \geq z$ 的情況下得到證明。

想一想 一個只在 $x \geq y \geq z$ 時證明了的結果，怎樣才能推出對所有 x、y、z 都成立的結果？

$x^2 + y^2 + z^2 \geq xy + yz + zx$ 的兩邊都是對稱的：把任意兩個變數對調，每一邊都保持不變。給定任意實數 x、y、z，把它們由大至小排列為 $X \geq Y \geq Z$。根據剛才所證，$X^2 + Y^2 + Z^2 \geq XY + YZ + ZX$。但這與關於 x、y、z 的不等式是同一個不等式，只是各項的次序不同而已。所以這個不等式對所有實數 x、y 和 z 都成立。

（等號成立需要我們相加的兩個不等式同時取等號，這就迫使 $x = y = z$。）

第 (ii) 部分

這個式子是三對數之和，每一對都是一個數加上它的倒數（reciprocal）：

$$\left(\frac{s}{t} + \frac{t}{s}\right) + \left(\frac{t}{r} + \frac{r}{t}\right) + \left(\frac{r}{s} + \frac{s}{r}\right).$$

因此只需證明對任意正數 p、q 都有 $\dfrac{p}{q} + \dfrac{q}{p} \geq 2$，這正是 $x^2 + y^2 \geq 2xy$ 的類比。

式子 $\frac{p}{q} + \frac{q}{p}$ 對 p 和 q 是對稱的，所以我們可以適當命名，使得 $p \geq q > 0$。在 $(*)$ 中取

$$a = q, \quad b = \frac{1}{p}, \quad c = p, \quad d = \frac{1}{q}.$$

於是 $c \geq a$ 即 $p \geq q$，而 $d \geq b$ 即 $\frac{1}{q} \geq \frac{1}{p}$；由於 p 和 q 都是正數，後者同樣可由 $p \geq q$ 推出。現在

$$ab + cd = \frac{q}{p} + \frac{p}{q}, \qquad bc + ad = \frac{p}{p} + \frac{q}{q} = 2,$$

所以由 $(*)$ 得 $\dfrac{p}{q} + \dfrac{q}{p} \geq 2$。

把這個結果應用於 (s,t)、(t,r) 和 (r,s) 三對數，再把三個結果相加，便得

$$\frac{s}{t}+\frac{t}{r}+\frac{r}{s}+\frac{t}{s}+\frac{r}{t}+\frac{s}{r}\geq 2+2+2=6$$

對所有正數 r、s 和 t 都成立，這正是所要證明的。等號成立當且僅當 $r=s=t$，因為每一對都需要其中兩數相等。

驗算

- 在第 (i) 部分試代入 $x=3$、$y=1$、$z=-2$：左邊是 14，右邊是 $3-2-6=-5$。把這三個值任意置換，兩邊都保持不變，正如對稱性論證所需。
- 第 (i) 部分的另一個獨立證明：$2(x^2+y^2+z^2-xy-yz-zx)=(x-y)^2+(y-z)^2+(z-x)^2\geq 0$。
- 取 $r=2$、$s=1$、$t=1$，第 (ii) 部分的式子等於 $1+\frac{1}{2}+2+1+2+\frac{1}{2}=7\geq 6$。用電腦對 20,000 個隨機的精確有理數例子逐步檢查，沒有發現任何反例；而 $r=1.001$、$s=1$、$t=0.999$ 之類的數值所得的和僅僅略大於 6。

常見錯誤

- 使用 $(*)$ 時沒有檢查 $c\geq a$ 和 $d\geq b$。若代入方式不當，不等號的方向可能會相反。
- 在第 (ii) 部分，忘記了 $\frac{1}{q}\geq\frac{1}{p}$ 用到正數這個條件：對於負數，取倒數並不會以同樣的方式把大小次序倒轉。
- 只在 $x\geq y\geq z$ 的情況下證明第 (i) 部分，卻沒有解釋為何憑對稱性可得出一般情況。

STEP 之外

不等式 $(*)$ 是*排序不等式*（rearrangement inequality）在兩項時的情況：若兩列數都已按大小排好，則按相同次序配對時乘積之和最大，按相反次序配對時乘積之和最小。本題的兩個結果，以及兩項的算術—幾何平均不等式（AM–GM），都可以由它直接推出；它亦與切比雪夫（Chebyshev）總和不等式密切相關。Siklos 的 *Advanced Problems in Mathematics* 有更多 STEP 風格的不等式練習，而 Bartle 與 Sherbert 的 *Introduction to Real Analysis* 則展示了這類關於大小次序的論證如何成為分析學（analysis）的基礎。

詳細解答

STEP I 2008，第 4 題

純數學 • 凸函數

> **STEP I 2008，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/08-s1-q4

關鍵思路

兩個函數都是指數函數（exponential），形如 $\mathrm{e}^{h(x)}$，求導兩次便得 $\big(h''+(h')^2\big)\mathrm{e}^h$。指數因子恆為正，所以凸性（convexity）問題歸結為 $h''+(h')^2$ 的正負，這是一個純粹的三角不等式。在 (i) 中，這個不等式比較的恰好就是題目要求我們繪畫的兩條曲線。在 (ii) 中，它化簡為 $\sin 2x$ 與 $\sin 2\alpha$ 的比較。

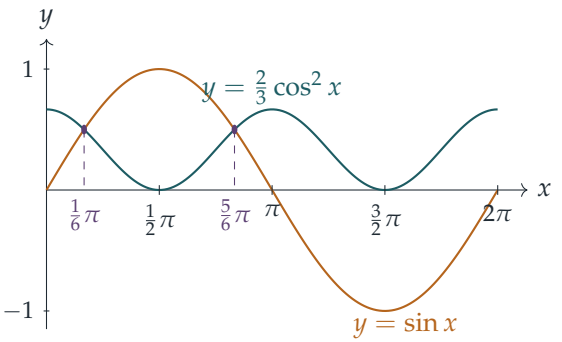
解答

第 (i) 部分

草圖。曲線 $y=\frac{2}{3}\cos^2 x$ 介乎 0 與 $\frac{2}{3}$ 之間。它在 $x=0,\pi,2\pi$ 處取值 $\frac{2}{3}$，並在 $x=\frac{1}{2}\pi,\frac{3}{2}\pi$ 處與橫軸相切，週期為 π。曲線 $y=\sin x$ 是常見的正弦波。要找出兩者的交點，設 $s=\sin x$，並用 $\cos^2 x=1-s^2$：

$$\tfrac{2}{3}(1-s^2)=s \iff 2s^2+3s-2=0 \iff (2s-1)(s+2)=0.$$

由於 $s=-2$ 不可能成立，故 $\sin x=\frac{1}{2}$，所以兩條曲線只在 $x=\frac{1}{6}\pi$ 和 $x=\frac{5}{6}\pi$ 相交。在 $(\pi,2\pi)$ 上，正弦曲線為負，而另一條曲線非負，所以該處沒有交點。



f 的凸性。設 $f(x) = \mathrm{e}^{\frac{2}{3}\sin x}$，由鏈式法則（chain rule）得 $f'(x) = \frac{2}{3}\cos x\, f(x)$，再用乘積法則（product rule），

$$\begin{aligned} f''(x) &= -\tfrac{2}{3}\sin x\, f(x) + \tfrac{2}{3}\cos x\, f'(x) \\ &= f(x)\left(\tfrac{4}{9}\cos^2 x - \tfrac{2}{3}\sin x\right) \\ &= \tfrac{2}{3}f(x)\left(\tfrac{2}{3}\cos^2 x - \sin x\right). \end{aligned}$$

想一想 為甚麼括號內的式子就能決定一切？

由於 $f(x)$ 是指數函數，它嚴格為正，所以 $f''(x) \geq 0$ 當且僅當 $\frac{2}{3}\cos^2 x \geq \sin x$。換言之，$f$ 為凸函數之處，恰好就是青綠色曲線位於琥珀色曲線之上或與之重合的地方。這就是題目先要求繪畫草圖的原因。

由草圖及上面求得的交點可知，當 $0 < x \leq \frac{1}{6}\pi$ 以及 $\frac{5}{6}\pi \leq x < 2\pi$ 時，$\frac{2}{3}\cos^2 x \geq \sin x$。在 $\frac{1}{6}\pi$ 與 $\frac{5}{6}\pi$ 之間，正弦曲線在上方（例如在 $x = \frac{1}{2}\pi$ 處，兩個值分別為 0 和 1）。所以 f 在

$$0 < x \leq \tfrac{1}{6}\pi \qquad 及 \qquad \tfrac{5}{6}\pi \leq x < 2\pi,$$

上是凸的，而且不能在更大的區間上是凸的。（端點 $\frac{1}{6}\pi$ 和 $\frac{5}{6}\pi$ 處 $f'' = 0$，是否把它們包括在內，視乎個人習慣。）

第 (ii) 部分

設 $g(x) = \mathrm{e}^{-k\tan x}$，則 $g'(x) = -k\sec^2 x\, g(x)$。利用 $\frac{\mathrm{d}}{\mathrm{d}x}\sec^2 x = 2\sec^2 x \tan x$ 及乘積法則，

$$\begin{aligned} g''(x) &= -2k\sec^2 x \tan x\, g(x) - k\sec^2 x\, g'(x) \\ &= g(x)\left(k^2\sec^4 x - 2k\sec^2 x \tan x\right) \\ &= k\sec^2 x\, g(x)\left(k\sec^2 x - 2\tan x\right). \end{aligned}$$

這裏 $k = \sin 2\alpha > 0$，因為 $0 < 2\alpha < \frac{1}{2}\pi$；另外 $\sec^2 x > 0$，$g(x) > 0$。所以 $g''(x) \geq 0$ 當且僅當 $k \sec^2 x \geq 2\tan x$。兩邊乘以 $\cos^2 x > 0$，這就變成

$$k \geq 2\tan x \cos^2 x = 2\sin x \cos x = \sin 2x.$$

所以 g 為凸函數之處，恰好就是滿足

$$\sin 2x \leq \sin 2\alpha.$$

的地方。

想一想 當 $0 < x < \frac{1}{2}\pi$ 時，$\sin 2x \leq \sin 2\alpha$ 在哪裏成立？

當 x 取遍 $(0, \frac{1}{2}\pi)$ 時，$2x$ 取遍 $(0, \pi)$。在這個區間上，sin 先遞增至 $\frac{1}{2}\pi$，然後遞減，並且關於 $\frac{1}{2}\pi$ 對稱。$\sin 2\alpha$ 這個值（其中 $0 < 2\alpha < \frac{1}{2}\pi$）在 $2x = 2\alpha$ 和 $2x = \pi - 2\alpha$ 處取得。所以 $\sin 2x \leq \sin 2\alpha$ 當且僅當 $2x \leq 2\alpha$ 或 $2x \geq \pi - 2\alpha$。即是

$$0 < x \leq \alpha \qquad \text{或} \qquad \tfrac{1}{2}\pi - \alpha \leq x < \tfrac{1}{2}\pi.$$

特別地，當 $0 < x < \alpha$ 時，我們有 $0 < 2x < 2\alpha < \frac{1}{2}\pi$，而 sin 在該處遞增，所以 $\sin 2x < \sin 2\alpha = k$，且 $g''(x) > 0$。因此 g 在 $0 < x < \alpha$ 上是凸的，正如題目所求。

另一個區間是

$$\tfrac{1}{2}\pi - \alpha < x < \tfrac{1}{2}\pi.$$

它與第一個區間確實不同，因為由 $\alpha < \frac{1}{4}\pi$ 可得 $\frac{1}{2}\pi - \alpha > \frac{1}{4}\pi > \alpha$。在中間一段 $\alpha < x < \frac{1}{2}\pi - \alpha$ 上，g 不是凸的。

驗算

- 在 $x = \frac{1}{6}\pi$ 處：$\frac{2}{3}\cos^2 x = \frac{2}{3} \cdot \frac{3}{4} = \frac{1}{2} = \sin x$，所以該處 $f'' = 0$，這正是邊界點應有的情況。
- 在 (ii) 中，問題在 $x \mapsto \frac{1}{2}\pi - x$ 之下是對稱的，因為 $\sin 2x$ 保持不變。這解釋了為甚麼第二個區間是第一個區間的鏡像。
- 數值上，取 $\alpha = 0.5$ ($k \approx 0.841$)，用有限差分（finite difference）計算 g 的二階導數（second derivative），在 $x = 0.3$ 和 $x = 1.3$ 處為正，在 $x = 0.8$ 處為負。由於 $\frac{1}{2}\pi - \alpha \approx 1.071$，這與求得的區間吻合。用電腦測試大量隨機的 α 和 x，結果與兩部分的答案都一致。

常見錯誤

- 對 $e^{h(x)}$ 求導兩次時遺漏了 $(h')^2$ 一項。二階導數並不只是 $h''e^h$。

- 在 (i) 中漏掉了整個 $(\pi, 2\pi)$。在那裏 $\sin x < 0 \leq \frac{2}{3}\cos^2 x$，所以 f 由 $\frac{5}{6}\pi$ 一直到 2π 都是凸的。
- 做到 $\sin 2x \leq \sin 2\alpha \Rightarrow x \leq \alpha$ 便停下來。由於 sin 在 $(0, \pi)$ 上並非單調 (monotonic)，這樣會漏掉第二個區間。

STEP 之外

二階導數判別法只是凸性的其中一面。真正的定義是幾何的：每一條弦（chord）都位於圖像之上或與圖像重合。對可導函數而言，這等價於 f' 遞增，並由此引出琴生不等式（Jensen's inequality），它是算術幾何平均不等式（AM–GM）以及概率和最優化中許多結果的基礎。Bartle 與 Sherbert 的 *Introduction to Real Analysis* 對凸函數有嚴謹的論述。Siklos 的 *Advanced Problems in Mathematics* 則提供更多 STEP 風格的練習。

詳細解答

STEP I 2008，第 5 題

純數學 • 多項式的界限

> **STEP I 2008，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/08-s1-q5

關鍵思路

當 $n = 1$ 時，多項式（polynomial）是一條直線，它在區間兩個端點的值互相牽制：兩者恰好相差 2，所以不可能同時都很小。第 (ii) 部分的竅門是先除以首項係數（leading coefficient），使切比雪夫（Chebyshev）定理可以直接應用於該曲線本身。定理於是迫使 $|y|$ 在 $[-1,1]$ 內某處必須很大，而端點處的值卻很小。一條曲線若升得比兩個端點值都高（或降得比兩者都低），就必定在中間某處轉向。

解答

第 (i) 部分

$n = 1$ 的情況。這時 $\mathrm{p}(x) = x + a_0$，我們需要證明 $M \geq 2^0 = 1$。由於 M 不小於 $|\mathrm{p}|$ 在 $[-1,1]$ 內任何一點的值，它自然不小於端點處的值：

$$M \geq |\mathrm{p}(1)| = |1 + a_0| \quad \text{及} \quad M \geq |\mathrm{p}(-1)| = |a_0 - 1|.$$

兩式相加，並利用三角不等式（triangle inequality）$|u| + |v| \geq |u - v|$，

$$2M \geq |1 + a_0| + |a_0 - 1| \geq \big|(1 + a_0) - (a_0 - 1)\big| = 2,$$

所以 $M \geq 1 = 2^{1-n}$，這就證明了定理在 $n = 1$ 時成立。（當 $\mathrm{p}(x) = x$ 時等號成立，所以這個界不能再改進。）

(a) 對於 $\mathrm{p}(x) = x^2 - \frac{1}{2}$ 及 $|x| \leq 1$，我們有 $0 \leq x^2 \leq 1$，所以 $-\frac{1}{2} \leq \mathrm{p}(x) \leq \frac{1}{2}$。在 $x = \pm 1$ 處取得 $\frac{1}{2}$ 這個值（而在 $x = 0$ 處取得 $-\frac{1}{2}$），所以

$$M = \tfrac{1}{2} = 2^{1-2}.$$

定理成立，而且等號成立。

(b) 對於 $\mathrm{p}(x) = x^3 - x$，我們需要 $M \geq 2^{1-3} = \frac{1}{4}$。只要找到一個點使得 $|\mathrm{p}| \geq \frac{1}{4}$ 就足夠了。駐點 (stationary point) 出現在 $3x^2 - 1 = 0$ 之處，即 $x = \pm\frac{1}{\sqrt{3}}$，兩者都在 $[-1,1]$ 內，而且

$$\mathrm{p}\left(-\tfrac{1}{\sqrt{3}}\right) = -\frac{1}{3\sqrt{3}} + \frac{1}{\sqrt{3}} = \frac{2}{3\sqrt{3}}.$$

由於 $\mathrm{p}(\pm 1) = 0$，這其實就是最大值：$M = \frac{2}{3\sqrt{3}} = \frac{2\sqrt{3}}{9} \approx 0.385$。若要不用計算機而與 $\frac{1}{4}$ 比較，可把兩邊平方：

$$\left(\frac{2}{3\sqrt{3}}\right)^2 = \frac{4}{27} > \frac{4}{64} = \frac{1}{16} = \left(\frac{1}{4}\right)^2,$$

所以 $M > \frac{1}{4}$，定理成立（這次還綽綽有餘）。

第 (ii) 部分

設 $f(x) = 64x^5 + 25x^4 - 66x^3 - 24x^2 + 3x + 1$。切比雪夫定理針對的是首項係數為 1 的多項式，所以我們把它應用於

$$\mathrm{p}(x) = \tfrac{1}{64}f(x) = x^5 + \tfrac{25}{64}x^4 - \tfrac{66}{64}x^3 - \tfrac{24}{64}x^2 + \tfrac{3}{64}x + \tfrac{1}{64},$$

它具有所需的形式，其中 $n = 5$。定理給出 $\max_{|x|\leq 1} |\mathrm{p}(x)| \geq 2^{-4} = \frac{1}{16}$，再乘以 64，

$$\max_{|x|\leq 1} |f(x)| \geq 4.$$

所以存在某個 $x_0 \in [-1,1]$ 使得 $|f(x_0)| \geq 4$。

現在看看端點：

$$f(1) = 64 + 25 - 66 - 24 + 3 + 1 = 3,$$
$$f(-1) = -64 + 25 + 66 - 24 - 3 + 1 = 1.$$

兩者的模（modulus）都小於 4，所以 x_0 不是端點：$-1 < x_0 < 1$。

想一想 我們知道 $|f|$ 在區間內部會變得很大。為甚麼這樣就必定有轉向點？

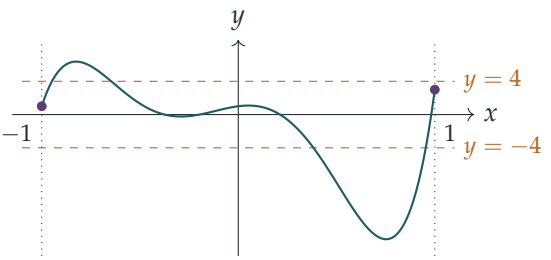
這裏有兩種情況。

- 若 $f(x_0) \geq 4$，則 $f(x_0)$ 比 $f(-1) = 1$ 和 $f(1) = 3$ 都大。連續函數（continuous function）f 在閉區間 (closed interval) $[-1,1]$ 上取得最大值；這個值至少是 $f(x_0) \geq 4$，所以它不是在任何一個端點取得的。因此它是在一個內點取得的，而該點就是 f 的局部極大點（local maximum）。

- 若 $f(x_0) \le -4$，則 $f(x_0)$ 低於兩個端點值，同理 f 在 $[-1,1]$ 上的最小值是在一個內點取得的，而該點就是局部極小點（local minimum）。

無論哪種情況，f 都有一個局部極大點或極小點嚴格位於 $(-1,1)$ 之內。在該點 $f' = 0$，而由於 f 是非常數多項式，f' 只有有限個零點。所以 f' 在該點左邊附近保持同一個符號，在右邊附近也保持同一個符號。因為該點是局部極大點或極小點，這兩個符號必定相反，所以它是一個真正的轉向點（turning point）。因此曲線在 $-1 \le x \le 1$ 內至少有一個轉向點，這正是所要證明的。

留意我們從頭到尾都不必解四次方程 $f'(x) = 0$。這正是本題的用意。



草圖顯示曲線在 $[-1,1]$ 上的形狀。兩個端點值都位於虛線 $y = \pm 4$ 之間，但曲線卻走出了這條帶狀區域。事實上它在區間內有四個轉向點，最深的一個在 $x \approx 0.75$ 附近，該處 $y \approx -15$。

驗算

- 當 $n = 1$ 時，試幾個 a_0 的值便可看到 $\max(|1 + a_0|, |1 - a_0|)$ 永遠不小於 1，而且只有在 $a_0 = 0$ 時才等於 1。
- 在 (b) 中，$\frac{2}{3\sqrt{3}} \approx 0.3849$，遠高於 0.25。在細密的格點上對 $x^3 - x$ 取樣，可確認這就是 $|x^3 - x|$ 的最大值。
- 在 (ii) 中，對 f 在 $[-1,1]$ 上作數值掃描，得到 $\max|f| \approx 15.0$（在 $x \approx 0.753$ 處，該處 $f \approx -15.0$），遠高於定理所保證的 4。掃描亦發現 f' 在 $x \approx -0.825,\ -0.292,\ 0.052,\ 0.753$ 處變號。

常見錯誤

- 直接把定理應用於 f。它的首項係數是 64 而不是 1，所以必須先除以 64。所得的界變成 $64 \cdot 2^{-4} = 4$。
- 停在「某處 $|f| \ge 4$」這一步。單憑這一點並不能得出轉向點：你還需要端點值 1 和

3，才能說明極值是在區間內部取得的。

► 在 (b) 中，草率地比較 $\frac{2}{3\sqrt{3}}$ 與 $\frac{1}{4}$。把兩個正數平方，便可作出準確的比較。

STEP 之外

切比雪夫定理中的等號恰好在 $2^{1-n}T_n(x)$ 時成立，其中 T_n 是由 $T_n(\cos\theta) = \cos n\theta$ 定義的切比雪夫多項式（Chebyshev polynomial）。(a) 部分就是 $\frac{1}{2}T_2$，而 $n = 1$ 的情況就是 $T_1(x) = x$。這些多項式是逼近理論（approximation theory）的起點：它們解釋了為甚麼在切比雪夫節點（Chebyshev nodes）上插值（interpolation）的效果最好，勝過在等距點上插值。第 (ii) 部分的論證用到極值定理（extreme value theorem，即閉區間上的連續函數必取得其上下界），這個定理會在大學第一門分析課中嚴格證明。Abbott 的 *Understanding Analysis*（一本分析學入門教科書）是閱讀這個定理的好去處。

詳細解答

STEP I 2008，第 6 題

純數學 • 反函數與面積

STEP I 2008，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/08-s1-q6

關鍵思路

反函數（inverse function）的圖像是原函數圖像沿直線 $y = x$ 的反射（reflection），所以把 g 對 y 積分，得到的是一塊橫向量度的面積，即曲線 $y = \mathrm{f}(x)$ 與 y 軸之間的面積。積分上限 k 並非隨意選取：它正好等於 $\mathrm{f}(\frac{1}{2})$。一旦看出這一點，兩塊面積便剛好拼成一個長方形。

解答

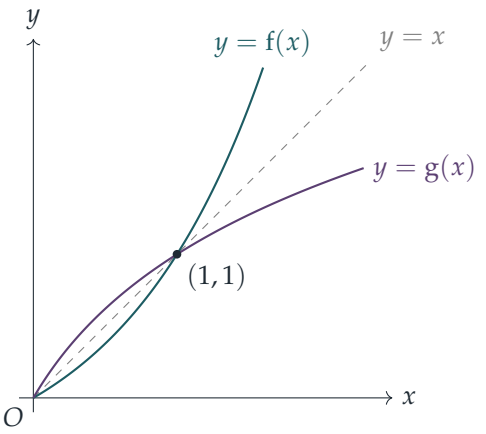
草圖

函數 f 是遞增（increasing）的，且 $\mathrm{f}(0) = 0$、$\mathrm{f}(1) = 1$。由於 $\mathrm{f}''(x) = \mathrm{e}^x/(\mathrm{e}-1) > 0$，它是凸（convex）的，而它在原點的斜率（gradient）為 $\mathrm{f}'(0) = 1/(\mathrm{e}-1) \approx 0.58 < 1$。因此在 $0 < x < 1$ 上，曲線位於直線 $y = x$ 之下，在 $(1,1)$ 與該直線相交，其後以指數方式增長。

要求出 g，可從 $y = \dfrac{\mathrm{e}^x - 1}{\mathrm{e}-1}$ 解出 x：

$$\mathrm{g}(x) = \ln\bigl(1 + (\mathrm{e}-1)x\bigr), \qquad x \geq 0,$$

這是因為 f 的值域（range）是 $[0, \infty)$。它的圖像是 $y = \mathrm{f}(x)$ 沿 $y = x$ 的反射：它通過 $(0,0)$，在該點的斜率為 $\mathrm{e} - 1 \approx 1.72$，是凹（concave）的，在 $(1,1)$ 與 $y = x$ 相交，其後只以對數方式增長。



第一個積分

$$\int_0^{1/2} \mathrm{f}(x)\,\mathrm{d}x = \frac{1}{\mathrm{e}-1}\Big[\mathrm{e}^x - x\Big]_0^{1/2} = \frac{1}{\mathrm{e}-1}\Big(\sqrt{\mathrm{e}} - \tfrac{1}{2} - 1\Big) = \frac{\sqrt{\mathrm{e}} - \frac{3}{2}}{\mathrm{e}-1}.$$

第二個積分

首先留意

$$\mathrm{f}(\tfrac{1}{2}) = \frac{\sqrt{\mathrm{e}}-1}{\mathrm{e}-1} = \frac{\sqrt{\mathrm{e}}-1}{(\sqrt{\mathrm{e}}-1)(\sqrt{\mathrm{e}}+1)} = \frac{1}{\sqrt{\mathrm{e}}+1} = k,$$

這裏用了平方差（difference of two squares）$\mathrm{e}-1 = (\sqrt{\mathrm{e}})^2 - 1$。換言之，$\mathrm{g}(k) = \frac{1}{2}$，而且 $1 + (\mathrm{e}-1)k = 1 + (\sqrt{\mathrm{e}}-1) = \sqrt{\mathrm{e}}$。

> **想一想** 怎樣對 $\ln\big(1 + (\mathrm{e}-1)x\big)$ 積分？

作代換（substitution）$u = 1 + (\mathrm{e}-1)x$，於是 $\mathrm{d}x = \mathrm{d}u/(\mathrm{e}-1)$，積分上下限變為 $u = 1$ 和 $u = \sqrt{\mathrm{e}}$。然後利用 $\int \ln u\,\mathrm{d}u = u\ln u - u$（把被積函數看成 $\ln u$ 乘以 1，再用分部積分法（integration by parts））：

$$\begin{aligned}\int_0^k \mathrm{g}(x)\,\mathrm{d}x &= \frac{1}{\mathrm{e}-1}\int_1^{\sqrt{\mathrm{e}}} \ln u\,\mathrm{d}u = \frac{1}{\mathrm{e}-1}\Big[u\ln u - u\Big]_1^{\sqrt{\mathrm{e}}} \\ &= \frac{1}{\mathrm{e}-1}\Big(\tfrac{1}{2}\sqrt{\mathrm{e}} - \sqrt{\mathrm{e}} - (0-1)\Big) = \frac{1 - \frac{1}{2}\sqrt{\mathrm{e}}}{\mathrm{e}-1},\end{aligned}$$

其中用了 $\ln\sqrt{\mathrm{e}} = \frac{1}{2}$。

相加

$$\int_0^{1/2} \mathrm{f}(x)\,\mathrm{d}x + \int_0^k \mathrm{g}(x)\,\mathrm{d}x = \frac{\sqrt{\mathrm{e}} - \frac{3}{2} + 1 - \frac{1}{2}\sqrt{\mathrm{e}}}{\mathrm{e}-1} = \frac{\frac{1}{2}(\sqrt{\mathrm{e}}-1)}{(\sqrt{\mathrm{e}}-1)(\sqrt{\mathrm{e}}+1)} = \frac{1}{2(\sqrt{\mathrm{e}}+1)},$$

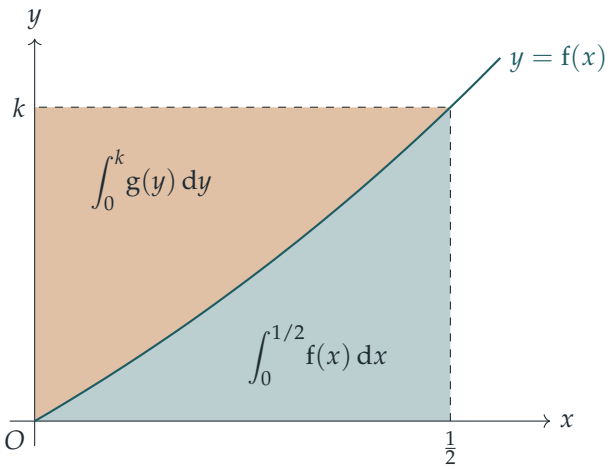
這正是所要求的結果。

解釋這個結果

答案是 $\frac{1}{2}\times k$，即以 $(0,0)$ 和 $(\frac{1}{2},k)$ 為對角頂點的長方形的面積。原因如下。

> **想一想** 在 f 的圖像上，$\int_0^k \mathrm{g}(x)\,\mathrm{d}x$ 在哪裏？

沿 $y=x$ 反射不會改變面積。$y=\mathrm{g}(x)$ 之下、$0\le x\le k$ 的區域，經反射後變成曲線 $y=\mathrm{f}(x)$ 與 y 軸之間、$0\le y\le k$ 的區域。換句話說，它就是沿曲線 $y=\mathrm{f}(x)$ 的 $\int_0^k x\,\mathrm{d}y$：曲線*左邊*的面積。由於 $\mathrm{f}(\frac{1}{2})=k$，曲線由長方形的頂點 $(0,0)$ 走到頂點 $(\frac{1}{2},k)$，把長方形分成兩塊：曲線之下的面積（第一個積分，藍綠色）和曲線左邊的面積（第二個積分，琥珀色）。



因此兩個積分合起來正好是整個長方形，面積為 $\frac{1}{2}k=\dfrac{1}{2(\sqrt{\mathrm{e}}+1)}$。同一幅圖還給出一個一般結論：對於滿足 $\mathrm{f}(0)=0$ 的遞增函數，

$$\int_0^a \mathrm{f}(x)\,\mathrm{d}x+\int_0^{\mathrm{f}(a)}\mathrm{f}^{-1}(y)\,\mathrm{d}y=a\,\mathrm{f}(a).$$

驗算

以數值計算，第一個積分約為 0.08655，第二個約為 0.10222；對兩者分別使用辛普森法則（Simpson's rule），結果與精確式子吻合至小數點後十二位，而兩者之和 0.18877 等於 $1/\big(2(\sqrt{\mathrm{e}}+1)\big)$。再用草圖作個合理性檢查：兩塊面積都是正數，而且藍綠色的一塊較小，這是理所當然的，因為凸曲線下垂至長方形對角線之下。

常見錯誤

- 沒有留意到 $k = \mathrm{f}(\frac{1}{2})$，因而錯過了 $1 + (\mathrm{e} - 1)k = \sqrt{\mathrm{e}}$ 這個化簡。
- 把 f 在原點附近畫在 $y = x$ 之上：它在該處的斜率是 $1/(\mathrm{e} - 1) < 1$。
- 作圖時，在另一幅圖上把 $y = \mathrm{g}(x)$ 之下的面積塗上陰影，卻沒有說明它與 $y = \mathrm{f}(x)$ 旁邊的區域有何關係；關鍵在於反射把其中一個區域變成另一個。

STEP 之外

恆等式 $\int_0^a \mathrm{f} + \int_0^{\mathrm{f}(a)} \mathrm{f}^{-1} = a\,\mathrm{f}(a)$ 是楊氏不等式（*Young's inequality*）中等號成立的情況：對任意 $b \geq 0$，$\int_0^a \mathrm{f} + \int_0^b \mathrm{f}^{-1} \geq ab$。取 $\mathrm{f}(x) = x^{p-1}$，便得到 $ab \leq a^p/p + b^q/q$，其中 $1/p + 1/q = 1$，這是分析學中赫爾德不等式（Hölder's inequality）的起點。想知道這類積分如何嚴格地建立起來，Abbott 的 *Understanding Analysis*（一本實分析入門書）是很好的讀物。

詳細解答

STEP I 2008，第 7 題

純數學 • 旋轉與距離

> **STEP I 2008，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/08-s1-q7

關鍵思路

要求一點到一條過原點直線的垂直距離（perpendicular distance），只要把平面旋轉（rotation），令該直線落在 x 軸上，問題就變得容易：所求距離正是新 y 坐標的大小。第一部分的旋轉公式，正好替三條邊中的兩條完成這件事。至於最後一部分，竅門是去掉絕對值符號（modulus signs）：三個有向距離（signed distance）之和永遠是同一個常數，而取絕對值只會令總和變大。

解答

旋轉公式

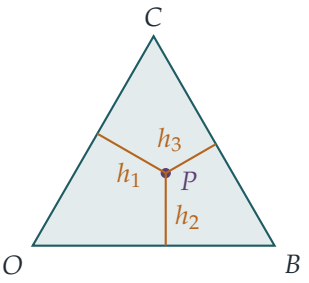
設 $x = r\cos\theta$，$y = r\sin\theta$，其中 $r = OP$，而 θ 是由正 x 軸量至 OP 的角。把 OP 順時針（clockwise）旋轉 $60°$，r 保持不變，角則變為 $\theta - 60°$，所以新的 y 坐標是

$$\begin{aligned} r\sin(\theta - 60°) &= r\sin\theta\cos 60° - r\cos\theta\sin 60° \\ &= \tfrac{1}{2}y - \tfrac{\sqrt{3}}{2}x = \tfrac{1}{2}\left(y - \sqrt{3}\,x\right), \end{aligned}$$

這正是所要證明的結果。若改為逆時針（anticlockwise）旋轉，角變為 $\theta + 60°$，同樣展開，只是 $\sin 60°$ 的符號相反，便得新的 y 坐標

$$\tfrac{1}{2}\left(y + \sqrt{3}\,x\right).$$

三個距離



h_1。直線 OC 與 x 軸的夾角為 60°。把整幅圖繞 O 順時針旋轉 60°，OC 便落在 x 軸上，而距離保持不變。在旋轉後的圖中，P 到 x 軸的距離就是其新 y 坐標的絕對值，所以

$$h_1 = \tfrac{1}{2}\left| y - \sqrt{3}\,x \right|.$$

h_2。直線 OB 就是 x 軸，所以 $h_2 = |y|$。

h_3。現在繞 O 逆時針旋轉 60°。這個旋轉把 $B = (1,0)$ 送到 $(\frac{1}{2}, \frac{1}{2}\sqrt{3})$，並把 C（位於角 60° 處）送到角 120° 處的點，即 $(-\frac{1}{2}, \frac{1}{2}\sqrt{3})$。因此直線 BC 變成水平線 $y = \frac{1}{2}\sqrt{3}$，而 P 的 y 坐標則變為 $\frac{1}{2}(y + \sqrt{3}\,x)$。於是

$$h_3 = \left| \tfrac{1}{2}(y + \sqrt{3}\,x) - \tfrac{1}{2}\sqrt{3} \right| = \tfrac{1}{2}\left| \sqrt{3}\,x + y - \sqrt{3} \right|.$$

（換個做法：BC 的方程是 $\sqrt{3}\,x + y = \sqrt{3}$，由於 $\sqrt{3+1} = 2$，用慣常的距離公式亦得出相同答案。）

最後一部分

定義以下帶正負號的量：

$$s_1 = \tfrac{1}{2}(\sqrt{3}\,x - y), \qquad s_2 = y, \qquad s_3 = \tfrac{1}{2}(\sqrt{3} - \sqrt{3}\,x - y),$$

使得對 $i = 1,2,3$ 均有 $h_i = |s_i|$。把三者相加，

$$s_1 + s_2 + s_3 = \tfrac{1}{2}\sqrt{3}\,x - \tfrac{1}{2}y + y + \tfrac{1}{2}\sqrt{3} - \tfrac{1}{2}\sqrt{3}\,x - \tfrac{1}{2}y = \tfrac{1}{2}\sqrt{3}$$

對每一點 (x,y) 都成立。

? **想一想** s_1、s_2、s_3 的正負號與三角形有甚麼關係？

每個 s_i 都是一個線性式，並且在三角形的其中一條邊上等於零。以位於三角形內部的形心（centroid）$(\frac{1}{2}, \frac{1}{6}\sqrt{3})$ 作測試，得 $s_1 = s_2 = s_3 = \frac{1}{6}\sqrt{3} > 0$。所以 $s_2 \geq 0$ 是 OB 包含 C

的一側，$s_1 \geq 0$ 是 OC 包含 B 的一側，而 $s_3 \geq 0$ 是 BC 包含 O 的一側。閉三角形 OBC 正是這三個半平面（half-plane）的交集，所以

$$P \text{ 落在（含邊界的）} OBC \iff s_1 \geq 0,\ s_2 \geq 0,\ s_3 \geq 0.$$

對任意實數 s，我們有 $|s| \geq s$，等號成立當且僅當（if and only if）$s \geq 0$。因此

$$h_1 + h_2 + h_3 = |s_1| + |s_2| + |s_3| \ \geq\ s_1 + s_2 + s_3 = \tfrac{1}{2}\sqrt{3},$$

等號成立當且僅當對每個 i 都有 $|s_i| = s_i$，亦即當且僅當三個 s_i 全部非負。根據上一段，這恰好在 P 位於三角形 OBC 的邊上或內部時發生。這樣便一次過證明了兩個方向：

$$h_1 + h_2 + h_3 = \tfrac{1}{2}\sqrt{3} \iff P \text{ 落在（含邊界的）三角形} OBC.$$

還要留意，在三角形外，這個和嚴格大於 $\frac{1}{2}\sqrt{3}$。

驗算

- 在頂點上，例如 $P = B = (1,0)$：$h_1 = \frac{1}{2}\sqrt{3}$，$h_2 = 0$，$h_3 = 0$，和為 $\frac{1}{2}\sqrt{3}$。正確。
- 在形心處，每個距離都是 $\frac{1}{6}\sqrt{3}$，三個加起來是 $\frac{1}{2}\sqrt{3} \approx 0.8660$，即三角形的高。
- 在三角形外，取 $P = (2,0)$：$h_1 = \sqrt{3}$，$h_2 = 0$，$h_3 = \frac{1}{2}\sqrt{3}$，和為 $\frac{3}{2}\sqrt{3} > \frac{1}{2}\sqrt{3}$。
- 我們在 10^5 個隨機點上，把三條公式與直接計算的點到直線距離作比較（吻合至 10^{-15}），並在 2×10^5 個點（包括許多位於邊上的點）上測試了「當且僅當」的結論，沒有發現任何例外。

常見錯誤

- 弄錯旋轉的正負號：順時針表示角*減小*，得 $\theta - 60°$。
- 只證明一個方向。驗證了在三角形內和為 $\frac{1}{2}\sqrt{3}$ 並不足夠；你還必須證明在三角形外和*不是* $\frac{1}{2}\sqrt{3}$。不等式 $|s| \geq s$ 可以一次過處理兩個方向。
- 忘記 h_1、h_2、h_3 中的絕對值符號：距離永遠不會是負數。

STEP 之外

這就是維維亞尼定理（Viviani's theorem）：在等邊三角形（equilateral triangle）中，內部任意一點到三邊的距離之和等於三角形的高。有向距離 s_1, s_2, s_3（相差一個比例常數）就是 P 的*重心坐標*（barycentric coordinates）或*三線坐標*（trilinear coordinates），它

們的和永遠是常數，並且恰好在三角形內全部非負；它們是幾何、電腦圖形學和凸分析中的標準工具。旋轉公式是 2×2 旋轉矩陣（rotation matrix）的第二行，而 Strang 的 *Introduction to Linear Algebra*（線性代數入門）是認識如何把旋轉看成線性映射（linear map）的好去處。

詳細解答

STEP I 2008，第 8 題

純數學 • 克萊羅微分方程

> **STEP I 2008，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/08-s1-q8

關鍵思路

兩條方程都含有 y' 而不含 y''，而且都不屬於可以直接分離變數（separate variables）或直接積分的類型。竅門是再求導一次：只含 y' 的項會互相抵消，餘下的式子可以因式分解（factorise），其中一個因式是 y''。每個因式對應一種情況，而在每種情況中，我們都代回原來的方程，而不是進行積分，這樣任意常數（arbitrary constant）便不費工夫地確定下來。

解答

以下一律記 $p = y'$，因此 $p' = y''$。

第 (i) 部分

把 $p^2 - xp + y = 0$ 對 x 求導，其中 xp 一項用積法則（product rule）：

$$2p\,y'' - \left(p + x\,y''\right) + p = 0 \quad \Longrightarrow \quad y''\,(2p - x) = 0.$$

兩個 p 項互相抵消，所以在每一點上，或者 $y'' = 0$，或者 $2y' = x$，正如題目所求。

情況 $y'' = 0$。這時 y' 是常數 m，而 $y = mx + c$。代入 $(*)$：

$$m^2 - xm + (mx + c) = 0 \quad \Longrightarrow \quad c = -m^2.$$

所以曲線是直線 $y = mx - m^2$。

情況 $2y' = x$。把 $p = \frac{1}{2}x$ 代入 $(*)$：

$$\tfrac{1}{4}x^2 - \tfrac{1}{2}x^2 + y = 0 \quad \Longrightarrow \quad 4y = x^2.$$

想一想 為甚麼不直接把 $y' = \frac{1}{2}x$ 積分，得出 $y = \frac{1}{4}x^2 + k$？

這樣做也可以，但對 $(*)$ 求導時丟失了一些資料：任何 $y = \frac{1}{4}x^2 + k$ 都滿足 $2y' = x$，卻只有 $k = 0$ 才滿足 $(*)$ 本身（例如 $k = 1$ 時，在 $x = 0$ 處得 $0 - 0 + 1 \neq 0$）。直接代回 $(*)$ 可一次過用盡所有資料，並且直接得出常數的值。第一種情況中 $c = -m^2$，也是同一道理。

所以曲線或者是直線 $y = mx + c$，其中 $c = -m^2$，或者是拋物線（parabola）$4y = x^2$。

這兩個答案互有關連：直線 $y = mx - m^2$ 與 $4y = x^2$ 相交於 $x^2 - 4mx + 4m^2 = 0$ 之處，即 $(x - 2m)^2 = 0$，所以每條直線都是拋物線在 $(2m, m^2)$ 的切線（tangent）。這條拋物線就是這族直線的包絡線（envelope）。

第 (ii) 部分

把 $(x^2 - 1)p^2 - 2xyp + y^2 - 1 = 0$ 對 x 求導：

$$2xp^2 + 2(x^2 - 1)p\,y'' - 2yp - 2xp^2 - 2xy\,y'' + 2yp = 0.$$

不含 y'' 的項兩兩抵消，餘下

$$2y''\big[(x^2 - 1)p - xy\big] = 0.$$

情況 $y'' = 0$。這時 $y = mx + c$。留意到方程可以改寫成

$$(y - xp)^2 = 1 + p^2,$$

會方便得多，因為 $(y - xp)^2 - p^2 - 1 = y^2 - 2xyp + (x^2 - 1)p^2 - 1$。對於這條直線，$y - xp = c$ 而 $p = m$，所以

$$c^2 = 1 + m^2, \qquad \text{即} \qquad y = mx \pm \sqrt{1 + m^2}.$$

情況 $(x^2 - 1)p = xy$。當 $x^2 \neq 1$ 時，有 $p = \dfrac{xy}{x^2 - 1}$，所以

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

代入 $(y - xp)^2 = 1 + p^2$，再乘以 $(x^2 - 1)^2$：

$$\begin{aligned} y^2 &= (x^2 - 1)^2 + x^2 y^2 \\ y^2(1 - x^2) &= (1 - x^2)^2 \\ y^2 &= 1 - x^2, \end{aligned}$$

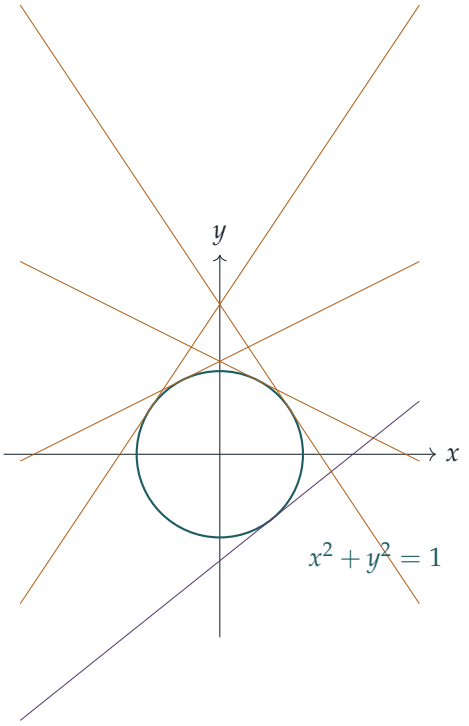
最後一步除以 $1-x^2 \neq 0$。所以曲線是圓 $x^2+y^2=1$。(在 $x=\pm 1$ 處，條件 $(x^2-1)p = xy$ 迫使 $y=0$，得出點 $(\pm 1, 0)$，它們也在同一個圓上。)

反過來，這個圓確實滿足方程：在圓上 $p=-x/y$，所以 $y-xp=(y^2+x^2)/y=1/y$，而 $1+p^2=(x^2+y^2)/y^2=1/y^2$，兩者吻合。

所以曲線或者是直線 $y=mx+c$，其中 $c^2=1+m^2$，或者是單位圓（unit circle）$x^2+y^2=1$。

> **想一想** 直線 $y=mx\pm\sqrt{1+m^2}$ 與這個圓有甚麼關係？

原點到 $y=mx+c$ 的距離是 $|c|/\sqrt{1+m^2}$，在這裏剛好等於 1。所以這些直線正好就是單位圓的所有（非鉛垂）切線，一如 (i) 中的直線是拋物線的切線。



驗算

- 在 (i) 中取 $m=1$：$y=x-1$ 有 $p=1$，而 $1-x+x-1=0$。它與 $4y=x^2$ 相切於 $(2,1)$。
- 在 (ii) 中，圓上的點 $(\frac{3}{5},\frac{4}{5})$ 有 $p=-\frac{3}{4}$；於是 $(x^2-1)p^2-2xyp+y^2-1 = -\frac{16}{25}\cdot\frac{9}{16}+\frac{18}{25}+\frac{16}{25}-1=0$。

- 數值方面，我們沿隨機選取的三次曲線檢驗了兩個求導後的因式分解，並確認直線、拋物線和圓所得的殘差（residual）均為 0（準確至 10^{-13}），而 $y = \frac{1}{4}x^2 + 1$ 和 $x^2 + y^2 = 4$ 則不然。

常見錯誤

- 把 $y' = \frac{1}{2}x$ 或 $y'' = 0$ 積分後留下一個任意常數：這個常數必須由原來的方程確定。
- 漏掉 $c = \pm\sqrt{1 + m^2}$ 中的 $\pm$：每個斜率（gradient）都有兩條互相平行的切線。

STEP 之外

第 (i) 部分是一條*克萊羅方程*（Clairaut equation）$y = xp - f(p)$：它的通解（general solution）是直線族 $y = mx - f(m)$，此外還有一個*奇解*（singular solution），即該直線族的包絡線，無論怎樣選取常數都得不出它。第 (ii) 部分是同一現象，$y = xp \pm \sqrt{1 + p^2}$，包絡線是那個圓。奇解說明，一階方程（first-order equation）經過每一點的解不一定唯一；唯一性何時成立，正是大學第一門常微分方程（ODE）課程中存在與唯一性定理（existence and uniqueness theorems）所探討的課題。Boyce 與 DiPrima 的 *Elementary Differential Equations*（常微分方程入門教科書）是閱讀這方面內容的好去處。

詳細解答

STEP I 2008，第 9 題

力學 • 滾動圓環的動量

STEP I 2008，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/08-s1-q9

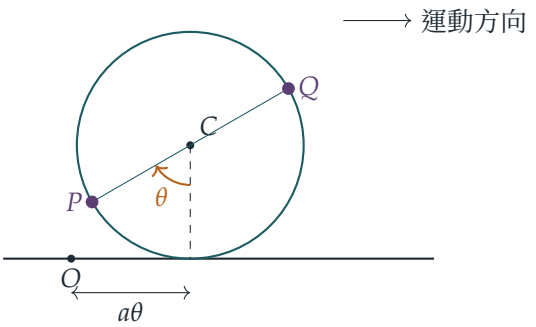
關鍵思路

純滾動（rolling without slipping）把平移與轉動連繫起來：當圓環轉過 θ 時，圓心移動的距離為 $a\theta$，即曾經接觸桌面的那段環邊的長度。因此每個質點的位置都是「圓心加上一條轉動的半徑」，求導便得出速度。至於最後一部分，留意外力沒有水平分量，所以總水平動量（momentum）不會改變，而這個動量原來是 $\dot{\theta}$ 的倍數。

解答

P 的位置與速率

把 P 和 Q 視為輕圓環上的質點（particle）。取 P 的起始點為原點，x 軸沿滾動的直線（指向圓環前進的方向），y 軸垂直向上。設圓環沿正 x 方向滾動，因此在下圖中它是順時針轉動的。



由於圓環沒有滑動，滾到桌面上的那段環邊弧長等於圓心移動的距離，所以圓心 C 位於

$(a\theta,\, a)$。起初 CP 指向正下方；順時針轉過 θ 之後，向量 $\overrightarrow{CP}$ 為 $(-a\sin\theta,\, -a\cos\theta)$。因此

$$\boxed{P = \big(a\theta - a\sin\theta,\; a - a\cos\theta\big)}$$

（這是一條擺線（cycloid））。利用鏈式法則（chain rule）對 t 求導，

$$\dot{x}_P = a\dot{\theta}(1-\cos\theta), \qquad \dot{y}_P = a\dot{\theta}\sin\theta,$$

所以

$$v_P^2 = a^2\dot{\theta}^2\big[(1-\cos\theta)^2 + \sin^2\theta\big] = 2a^2\dot{\theta}^2(1-\cos\theta) = 4a^2\dot{\theta}^2\sin^2\tfrac{1}{2}\theta,$$

其中用到 $1-\cos\theta = 2\sin^2\frac{1}{2}\theta$。因此 P 的速率為

$$v_P = 2a\left|\dot{\theta}\sin\tfrac{1}{2}\theta\right| \qquad \big(= 2a\dot{\theta}\sin\tfrac{1}{2}\theta\text{，當}0 \le \theta \le 2\pi,\ \dot{\theta} \ge 0\big).$$

當 θ 是 2π 的倍數時，這個速率為零，而這正好是 P 成為接觸點的時刻，與純滾動的要求一致。

動能

Q 位於直徑的另一端，所以 $\overrightarrow{CQ} = -\overrightarrow{CP}$，而

$$Q = \big(a\theta + a\sin\theta,\; a + a\cos\theta\big), \qquad \dot{x}_Q = a\dot{\theta}(1+\cos\theta), \quad \dot{y}_Q = -a\dot{\theta}\sin\theta.$$

同理，

$$v_Q^2 = a^2\dot{\theta}^2\big[(1+\cos\theta)^2 + \sin^2\theta\big] = 2a^2\dot{\theta}^2(1+\cos\theta) = 4a^2\dot{\theta}^2\cos^2\tfrac{1}{2}\theta.$$

圓環是輕的，所以動能（kinetic energy）只來自兩個質點：

$$T = \tfrac{1}{2}mv_P^2 + \tfrac{1}{2}mv_Q^2 = 2ma^2\dot{\theta}^2\big(\sin^2\tfrac{1}{2}\theta + \cos^2\tfrac{1}{2}\theta\big) = 2ma^2\dot{\theta}^2,$$

正如所求。

速率恆定

外力包括兩個質點的重量 mg 和桌面的反作用力（reaction），全部都是垂直的。因此系統所受的總水平力為零，而把牛頓第二定律（Newton's second law）應用於整個系統，可知其總水平動量恆定。（圓環是輕的，所以它不帶動量。）

想一想 質點也有垂直方向的運動，為甚麼只考慮水平動量就足夠？

因為水平動量恰好只取決於 $\dot{\theta}$。把水平速度相加，

$$m\dot{x}_P + m\dot{x}_Q = ma\dot{\theta}(1-\cos\theta) + ma\dot{\theta}(1+\cos\theta) = 2ma\dot{\theta}.$$

$\cos\theta$ 項互相抵消：系統的質心（centre of mass）是 C，它始終以速度 $a\dot{\theta}$ 水平移動。由於 $2ma\dot{\theta}$ 恆定，$\dot{\theta}$ 亦恆定，所以圓環的速率（即其圓心的速率 $a\dot{\theta}$）恆定。 □

或用能量方法。 P 和 Q 的高度之和為 $(a - a\cos\theta) + (a + a\cos\theta) = 2a$，所以勢能（potential energy）$2mga$ 永不改變。反作用力作用於接觸點，而純滾動時接觸點瞬時靜止，所以反作用力不作功。因此 $T = 2ma^2\dot{\theta}^2$ 恆定，即 $|\dot{\theta}|$ 恆定；又因為 $\dot{\theta}$ 連續地變化，不可能不經過 0 而由一個符號跳到另一個符號，所以 $\dot{\theta}$ 本身恆定。動量的論證省卻了這最後一步，這就是我們先用它的原因。

驗算

- 當 $\theta = \pi$ 時，P 在最高點：公式給出 $P = (a\pi, 2a)$，速率為 $2a\dot{\theta}$，是圓心速率的兩倍，這正是滾動車輪頂點的熟悉結果。
- 動能等於 $\frac{1}{2}(2m)(a\dot{\theta})^2 + \frac{1}{2}(2ma^2)\dot{\theta}^2$：質心的平移加上繞質心的轉動，其中轉動慣量（moment of inertia）為 $2ma^2$。兩項各為 $ma^2\dot{\theta}^2$。
- 數值上，取 $a = 1.3$、$\dot{\theta} = 2.1$ 和 $\theta = 1$，以有限差分（finite differences）對位置求導，得 $v_P^2 + v_Q^2 = 4a^2\dot{\theta}^2 \approx 29.8$；而當 $\dot{\theta}$ 恆定時，P 和 Q 的水平加速度互相抵消。

常見錯誤

- 把圓心放在 (a, a)，或者完全忘記圓心會移動：$a\theta$ 這一項正是滾動帶來的。
- 在 P 的位置中弄錯 $\sin\theta$ 的正負號；應檢查 P 起初的速度是否為零。
- 由動能恆定推斷速率恆定，卻沒有說明為甚麼 $\dot{\theta}$ 不能改變符號。

STEP 之外

這是一個雙質點的例子，說明了一個普遍的事實：系統的總動量只會因外力而改變，內力（這裏由圓環傳遞）則成對抵消。「驗算」中動能的分拆就是柯尼希定理（König's theorem）：動能等於質心的動能加上相對於質心運動的動能，這是剛體動力學（rigid-body dynamics）的起點。Kleppner 與 Kolenkow 的 *An Introduction to Mechanics*（力學入門教科書）對滾動運動和這些概念有細緻的論述。

詳細解答

STEP I 2008，第 10 題

力學 • 拋體運動

STEP I 2008，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/08-s1-q10

關鍵思路

重力只沿鉛垂方向作用，所以水平速度 $V\cos\alpha$ 始終不變，射程（range）就是 $V\cos\alpha$ 乘以總飛行時間（time of flight）。飛行時間完全由鉛垂方向的運動決定，而這段運動可分成三個階段：向上穿過底層、在高度 h 以上的一段往返，以及向下穿過底層返回地面。每個階段都是勻加速（constant acceleration）問題，而大部分工夫都可以交給對稱性（symmetry）代勞。

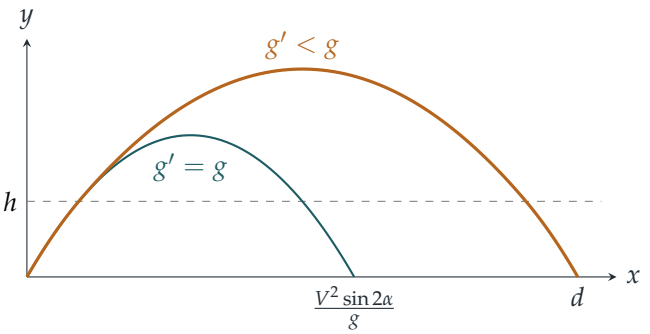
解答

我們把粒子（particle）看成一個質點，不計空氣阻力，在平地上方的一個鉛垂平面內運動。取 x 軸為水平方向，y 軸鉛垂向上，原點設在拋射點。

草圖

在高度 h 以下，兩條軌跡（trajectory）服從完全相同的方程，所以在該處兩者重合。條件 $V^2\sin^2\alpha > 2gh$ 表示初始的鉛垂動能（kinetic energy）足以在加速度 g 之下升過高度 h，所以粒子確實會越過 $y = h$。

在 h 以上，$g' < g$ 的情況重力較弱，鉛垂速率減少得較慢：粒子升得更高，在上面逗留得更久，描出一條較扁、較闊的拋物線（parabola）。速度在 $y = h$ 處連續（只有加速度出現跳躍），所以兩段弧線平滑相接，沒有尖角，但曲率（curvature）會突然改變。下降時，粒子重新進入底層，其速率和角度與離開時相同（由頂部弧線的對稱性可知），而最後一段弧線正是第一段的鏡像。因此 $g' < g$ 的路徑落在更遠的地方。



射程

初速度的鉛垂分量（vertical component）為 $V\sin\alpha$，而當 $y<h$ 時加速度為 $-g$。

第 1 階段：上升至高度 h。設 w 為粒子到達 $y=h$ 時的（向上）鉛垂速率。由 $v^2=u^2+2as$，

$$w^2=V^2\sin^2\alpha-2gh=\sin^2\alpha\left(V^2-2gh\operatorname{cosec}^2\alpha\right)=V'^2\sin^2\alpha,$$

所以 $w=V'\sin\alpha$；這裏取正根，因為 $\sin\alpha>0$，而且粒子仍在上升。（題目給出的條件正好保證 V' 是正實數。）由 $v=u+at$，所需時間為

$$t_1=\frac{V\sin\alpha-V'\sin\alpha}{g}=\frac{(V-V')\sin\alpha}{g}.$$

第 2 階段：在高度 h 以上。此時粒子從 $y=h$ 出發，以速率 w 向上運動，加速度為 $-g'$。當 $wt-\frac{1}{2}g't^2=0$ 時，它回到 $y=h$，即經過時間

$$t_2=\frac{2w}{g'}=\frac{2V'\sin\alpha}{g'},$$

之後，而這時它正以鉛垂速率 w 向下運動。

想一想 為甚麼最後一個階段所需的時間與第一個階段相同？

在第 3 階段，粒子從 $y=h$ 出發，以速率 w 向下運動，加速度為向下的 g，一直落到 $y=0$。這正是把第 1 階段在時間上倒過來播放：勻加速運動是可逆的，而第 1 階段是在同樣的加速度下，由 $y=0$、速率 $V\sin\alpha$ 運動到 $y=h$、速率 w。所以粒子經過時間 $t_3=t_1$ 後到達地面，鉛垂速率為 $V\sin\alpha$。（直接計算：鉛垂速率以 g 的變率由 w 增至 $V\sin\alpha$，需時 $(V\sin\alpha-w)/g$。）

因此總飛行時間為

$$T=2t_1+t_2=2\sin\alpha\left(\frac{V-V'}{g}+\frac{V'}{g'}\right).$$

由於沒有水平方向的力，水平速度自始至終保持為 $V\cos\alpha$，於是

$$\begin{aligned} d = V\cos\alpha \cdot T &= \left(\frac{V-V'}{g} + \frac{V'}{g'}\right) 2V\sin\alpha\cos\alpha \\ &= \left(\frac{V-V'}{g} + \frac{V'}{g'}\right) V\sin 2\alpha, \end{aligned}$$

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

驗算

- 若 $g' = g$，括號內的式子變成 V/g，於是 $d = V^2\sin 2\alpha/g$，即我們熟悉的地球上的射程。
- 若 $g' < g$，則 $V'/g' > V'/g$，所以 d 大於均勻重力下的射程，與草圖吻合。
- 若恰好有 $V^2\sin^2\alpha = 2gh$，則 $V' = 0$：粒子剛好觸及 $y = h$，完全感受不到 g'，於是同樣得到 $d = V^2\sin 2\alpha/g$。
- 數值上，取 $V = 10$、$\alpha = 60°$、$g = 10$、$h = 2$ 及 $g' = 5$，可得 $V' = \sqrt{140/3} \approx 6.83$ 及 $d \approx 14.576$；以數值方法逐步模擬這個運動，所得的落點相同，相對誤差在 10^{-5} 以內。（草圖所用的正是這些數值。）

常見錯誤

- 忘記了向下穿過底層的那一段，得出 $t_1 + t_2$ 而不是 $2t_1 + t_2$。
- 把高度 h 處的鉛垂速率當作 V' 而不是 $V'\sin\alpha$：V' 是按 $\operatorname{cosec}\alpha$ 放大了的速率，並不是鉛垂分量本身。
- 在兩段弧線相接處畫出尖角：速度是連續的，所以接合處是平滑的。

STEP 之外

重力隨位置而變，其實是常態而不是例外：真實的重力隨着與行星中心的距離按 $1/r^2$ 減弱，拋體的路徑也因而變成橢圓（ellipse）的弧，而不是拋物線。把一個運動分成若干個力的規律都很簡單的區域，再在邊界上令位置和速度互相銜接，是一種標準技巧；在粒子的折射（refraction），以及量子力學（quantum mechanics）中的階躍勢（step potential）問題裏，都會再次用到。Kleppner 與 Kolenkow 所著的 *An Introduction to Mechanics*（力學入門教科書）既講解勻加速運動的各種工具，也討論平方反比（inverse-square）重力下的軌道運動。

詳細解答

STEP I 2008，第 11 題

力學 • 極限摩擦力平衡

STEP I 2008，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/08-s1-q11

關鍵思路

由於 $\angle P_1OP_2$ 是直角，半徑 OP_2 平行於 P_1 處的切線（tangent），而半徑 OP_1 則平行於 P_2 處的切線。因此，只要沿這兩條半徑的方向分解（resolve）各力，每條方程便只涉及一個法向反作用力（normal reaction）和一個摩擦力（friction）。再加上一條對 O 取力矩（moment）的方程（兩個法向反作用力對該點都沒有力矩），並在兩端代入極限摩擦（limiting friction）的條件 $F = \mu R$，便得到三條方程，從中可以消去 α 以外的未知量。最後一部分其實是喬裝了的倍角公式（double-angle formulae）。

解答

模型與各力

把桿視為均勻（uniform）的剛性桿，質心（centre of mass）G 位於桿的中點；把兩個環視為輕環，並設金屬線圈的半徑為 a。由於環是輕的，線圈作用於每個環的力會原封不動地傳到桿上，所以可以當作桿在 P_1 和 P_2 兩點直接與線圈接觸。桿的每一端都受到一個沿半徑方向的法向反作用力和一個沿切線方向的摩擦力。

取 O 為原點，x 軸向右，y 軸向上，並設 P_1 在線圈最低點的左方。由於 $\angle P_1OP_2 = 90°$，而 P_1 與向下的鉛垂線成角 α，所以 OP_2 在另一側與該鉛垂線成角 $90° - \alpha$：

$$P_1 = a(-\sin\alpha, -\cos\alpha), \qquad P_2 = a(\cos\alpha, -\sin\alpha).$$

因為 $0 < \alpha < 45°$，$\cos\alpha > \sin\alpha$，所以 P_1 確實比 P_2 低，而且兩點都在 O 的下方。

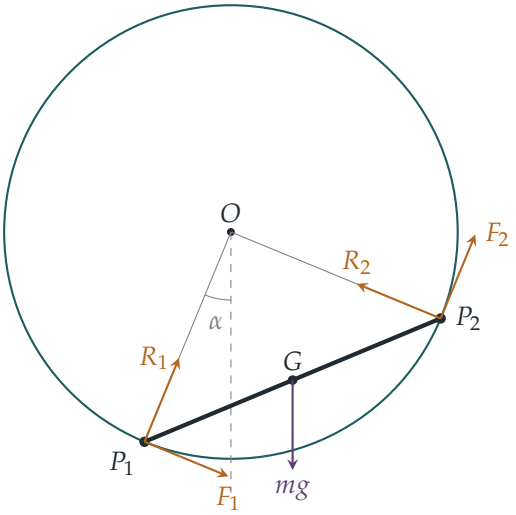
中點為 $G = \frac{a}{2}(\cos\alpha - \sin\alpha, -\cos\alpha - \sin\alpha)$，位於通過 O 的鉛垂線的*右方*。因此重量對 O 的力矩是順時針方向的，桿即將沿線圈順時針滑動（P_2 向下移，P_1 向上移）。摩擦力

阻礙這個運動，所以它在兩端都沿逆時針方向作用：在 P_1 處沿線圈向下，在 P_2 處沿線圈向上。

設 $\mathbf{u}_1$ 和 $\mathbf{u}_2$ 分別為由 P_1 和 P_2 指向 O 的單位向量（unit vector）：

$$\mathbf{u}_1 = (\sin\alpha, \cos\alpha), \qquad \mathbf{u}_2 = (-\cos\alpha, \sin\alpha), \qquad \mathbf{u}_1 \cdot \mathbf{u}_2 = 0.$$

P_1 處的逆時針切線方向是 $-\mathbf{u}_2$，而 P_2 處的則是 $\mathbf{u}_1$。設法向反作用力（指向 O）為 R_1, R_2，摩擦力為 F_1, F_2。



平衡方程

重量 mg 鉛垂向下作用，所以它沿 $\mathbf{u}_1$ 和 $\mathbf{u}_2$ 方向的分量分別是 $-mg\cos\alpha$ 和 $-mg\sin\alpha$。

沿 $\mathbf{u}_1$ 方向（即沿 P_1O）分解：R_1 和 F_2 有貢獻，而 R_2 和 F_1 則與這個方向垂直。所以

$$R_1 + F_2 = mg\cos\alpha. \tag{1}$$

沿 $\mathbf{u}_2$ 方向（沿 P_2O）分解：R_2 的貢獻是 $+R_2$，而 F_1 指向 $-\mathbf{u}_2$ 方向，貢獻是 $-F_1$。所以

$$R_2 - F_1 = mg\sin\alpha. \tag{2}$$

對 O 取力矩：兩個法向反作用力都通過 O。每個摩擦力都沿切線方向，與 O 的距離為 a，產生逆時針力矩 aF_1 和 aF_2。重量的作用線與 O 的水平距離為 $\frac{a}{2}(\cos\alpha - \sin\alpha)$，力矩為順時針方向。因此

$$F_1 + F_2 = \tfrac{1}{2}mg(\cos\alpha - \sin\alpha). \tag{3}$$

極限摩擦與 $\tan\alpha$

桿的兩端都處於即將滑動的狀態，所以 $F_1 = \mu R_1$ 及 $F_2 = \mu R_2$。記 $c = \cos\alpha$，$s = \sin\alpha$，方程 (1) 和 (2) 變成

$$R_1 + \mu R_2 = mgc, \qquad -\mu R_1 + R_2 = mgs,$$

其解為

$$R_1 = \frac{mg(c - \mu s)}{1 + \mu^2}, \qquad R_2 = \frac{mg(s + \mu c)}{1 + \mu^2}.$$

當 μ 很小時（這裏必然如此，見下文），兩者皆為正數，所以反作用力確實如圖所示指向 O。現在 (3) 表示 $\mu(R_1 + R_2) = \frac{1}{2}mg(c - s)$，即

$$\begin{aligned} 2\mu\big(c - \mu s + s + \mu c\big) &= (1 + \mu^2)(c - s) \\ c\,(2\mu + 2\mu^2 - 1 - \mu^2) + s\,(2\mu - 2\mu^2 + 1 + \mu^2) &= 0 \\ s\,(1 + 2\mu - \mu^2) &= c\,(1 - 2\mu - \mu^2). \end{aligned}$$

兩邊除以 $c\,(1 + 2\mu - \mu^2)$（它不等於零，因為 $c > 0$，而 $\tan\alpha > 0$ 迫使 μ 必須很小，這一點稍後便會看到），得

$$\tan\alpha = \frac{1 - 2\mu - \mu^2}{1 + 2\mu - \mu^2},$$

這正是所求的結果。

> **想一想** 我們怎樣知道摩擦力是沿逆時針方向而不是順時針方向作用？

重量對 O 的力矩是順時針方向的，因為 G 在 O 的右方，而法向反作用力對 O 沒有力矩。所以必須單靠摩擦力提供逆時針方向的力矩；若沒有摩擦力，桿便會順時針滑動。假如你把摩擦力畫成相反方向，整個運算過程會完全一樣，只是 μ 換成了 $-\mu$，結果得出 $\tan\alpha > 1$，與 $\alpha < 45^\circ$ 矛盾。

桿的傾角

由 P_1 到 P_2 的向量是 $a(\cos\alpha + \sin\alpha,\ \cos\alpha - \sin\alpha)$，它的兩個分量都是正數。所以桿與水平線所成的銳角 θ 滿足

$$\tan\theta = \frac{\cos\alpha - \sin\alpha}{\cos\alpha + \sin\alpha} = \frac{1 - \tan\alpha}{1 + \tan\alpha} = \tan(45^\circ - \alpha),$$

而由於 $0 < 45^\circ - \alpha < 45^\circ$，我們得到 $\theta = 45^\circ - \alpha$。

現在把 $\mu = \tan\lambda$ 代入 $\tan\alpha$ 的公式，並把分子和分母同乘以 $\cos^2\lambda$：

$$\tan\alpha = \frac{(\cos^2\lambda - \sin^2\lambda) - 2\sin\lambda\cos\lambda}{(\cos^2\lambda - \sin^2\lambda) + 2\sin\lambda\cos\lambda} = \frac{\cos 2\lambda - \sin 2\lambda}{\cos 2\lambda + \sin 2\lambda} = \frac{1 - \tan 2\lambda}{1 + \tan 2\lambda} = \tan(45^\circ - 2\lambda).$$

由於 $0 < \lambda < 22.5^\circ$，角 $45^\circ - 2\lambda$ 嚴格介乎 0° 與 45° 之間，α 亦然。正切函數在這個區間上是一對一（one-to-one）的，所以 $\alpha = 45^\circ - 2\lambda$。因此

$$\theta = 45^\circ - \alpha = 2\lambda,$$

這正是我們要證明的。（這也解釋了條件 $\lambda < 22.5^\circ$，即 $\mu < \sqrt{2} - 1$ 的由來：它恰好就是令 $\tan\alpha > 0$ 的條件。）

驗算

- 當 $\mu = 0$（光滑線圈）時，我們得到 $\tan\alpha = 1$，所以 $\alpha = 45^\circ$，桿呈水平，這是桿在光滑圓上能夠靜止的唯一位置。
- 當 $\mu = 0.2$ 時：$\tan\alpha = 0.56/1.36$，所以 $\alpha \approx 22.38^\circ$，而 $\theta = 45^\circ - \alpha \approx 22.62^\circ$，與 $2\arctan 0.2 \approx 22.62^\circ$ 吻合。
- 數值驗證：隨機選取 10^4 組 $\mu \in (0, \sqrt{2} - 1)$、$a$ 和 m 的值，把上述 R_1, R_2, F_1, F_2 代入完整的向量方程，所得的合力和力矩皆為零（誤差約為 10^{-14}），並且 $R_1, R_2 > 0$ 及 $\theta = 2\lambda$。

常見錯誤

- 沿水平和鉛垂方向分解。這樣做是可行的，但每條方程都會包含全部四個力；沿兩條互相垂直的半徑分解則簡潔得多。
- 把其中一端摩擦力的方向弄錯。兩個環沿圓周滑動的方向相同，所以兩端的摩擦力都是逆時針方向的。
- 忘記了 $\theta = 45^\circ - \alpha$ 需要略加說明：弦垂直於 $\angle P_1OP_2$ 的角平分線。

STEP 之外

這題還有一個更巧妙的幾何看法：在極限摩擦的情況下，每一端的總反作用力與半徑成角 λ，而一個在三個力作用下處於平衡（equilibrium）的物體，這三個力必須共點（concurrent），由此可直接推出 $\theta = 2\lambda$。這個「摩擦角」（angle of friction）的概念，以及摩擦力其實是不等式 $|F| \le \mu R$ 而不是等式這一事實，正是處理多個接觸點的靜力學問題的起點；在這類問題中，你必須問哪些接觸點會先滑動。Kleppner 與 Kolenkow 合著的 *An Introduction to Mechanics*（力學入門教科書）正是以這種嚴謹的態度處理摩擦力和剛體平衡。

詳細解答

STEP I 2008，第 12 題

概率與統計 • 離散均勻分佈變量的最大值

STEP I 2008，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/08-s1-q12

關鍵思路

X 其實就是兩個分數中較大的一個。若直接求「較大者等於 r」的概率，便要分幾種情況來數；但「較大者*不超過* r」卻很容易處理：這恰好在兩個分數都不超過 r 時發生。因此我們先求累積分佈（cumulative distribution）$\mathrm{P}(X \leq r)$，再取差。有了這個累積分佈，中位數（median）也幾乎立即可得。

解答

設 X_1 和 X_2 為兩個獨立（independent）的分數，並設 $X = \max(X_1, X_2)$（當 $X_1 = X_2$ 時，定義中的兩種情況給出相同的值，所以 X 是明確定義的）。

第 (i) 部分

對於 $r = 0, 1, \ldots, N$，事件 $\{X \leq r\}$ 就是事件 $\{X_1 \leq r$ 且 $X_2 \leq r\}$。這兩個事件各自的概率都是 r/N，而且由於 X_1 和 X_2 獨立，它們亦互相獨立，所以

$$\mathrm{P}(X \leq r) = \frac{r}{N} \cdot \frac{r}{N} = \frac{r^2}{N^2}.$$

這在 $r = 0$ 時同樣成立，所得的值為 0。因此，對於 $r = 1, 2, \ldots, N$，

$$\mathrm{P}(X = r) = \mathrm{P}(X \leq r) - \mathrm{P}(X \leq r - 1) = \frac{r^2 - (r-1)^2}{N^2} = \frac{2r - 1}{N^2},$$

這正是所要證明的結果。（簡單核對一下：$\sum_{r=1}^{N}(2r-1) = N^2$，所以這些概率之和為 1。）

第 (ii) 部分

利用題目給出的兩個求和公式，

$$\begin{aligned}\mu = \sum_{r=1}^{N} r\,\frac{2r-1}{N^2} &= \frac{1}{N^2}\left(2\sum_{r=1}^{N} r^2 - \sum_{r=1}^{N} r\right)\\ &= \frac{1}{N^2}\left(\frac{N(N+1)(2N+1)}{3} - \frac{N(N+1)}{2}\right)\\ &= \frac{N+1}{N}\cdot\frac{2(2N+1)-3}{6} = \frac{(N+1)(4N-1)}{6N}.\end{aligned}$$

當 $N = 100$ 時，

$$\mu = \frac{101\times 399}{600} = \frac{40\,299}{600} = 67.165,$$

與題目所求相符。

第 (iii) 部分

由第 (i) 部分，$\mathrm{P}(X \le m) = m^2/N^2$，而 $\mathrm{P}(X \ge m) = 1 - \mathrm{P}(X \le m-1) = 1 - (m-1)^2/N^2$。中位數須滿足的兩個條件於是變成

$$\mathrm{P}(X \le m) \ge \tfrac{1}{2} \iff m^2 \ge \tfrac{1}{2}N^2 \iff m \ge \frac{N}{\sqrt{2}},$$
$$\mathrm{P}(X \ge m) \ge \tfrac{1}{2} \iff (m-1)^2 \le \tfrac{1}{2}N^2 \iff m-1 \le \frac{N}{\sqrt{2}},$$

這裏取了正平方根；由於 $m \ge 1$，m 和 $m-1$ 都是非負數，所以這一步是成立的。因此 m 必須是滿足

$$\frac{N}{\sqrt{2}} \le m \le \frac{N}{\sqrt{2}} + 1.$$

的整數。

> **想一想** 長度為 1 的區間可以包含兩個整數。會不會有兩個中位數？

只有當區間的兩個端點都是整數時才會，而這需要 $N/\sqrt{2}$ 是整數。但若 $N/\sqrt{2} = k$，便有 $\sqrt{2} = N/k$，是一個有理數（rational number），這是不可能的。所以兩個不等式都是嚴格不等式，而長度為 1 的開區間（open interval）恰好包含一個整數，於是

$$m = \left\lceil \frac{N}{\sqrt{2}} \right\rceil,$$

即大於 $N/\sqrt{2}$ 的最小整數。

當 $N = 100$ 時，$N/\sqrt{2} = 50\sqrt{2} \approx 70.71$，所以 $m = 71$。直接驗證：$71^2 = 5041 \ge 5000 = \frac{1}{2}\times 100^2$，而 $70^2 = 4900 \le 5000$，所以兩個條件都成立。

第 (iv) 部分

由第 (iii) 部分，$m = N/\sqrt{2} + \theta$，其中 $0 < \theta < 1$，所以

$$\frac{m}{N} = \frac{1}{\sqrt{2}} + \frac{\theta}{N} \to \frac{1}{\sqrt{2}} \quad \text{當} N \to \infty.$$

由第 (ii) 部分，

$$\frac{\mu}{N} = \frac{(N+1)(4N-1)}{6N^2} = \frac{1}{6}\left(1+\frac{1}{N}\right)\left(4-\frac{1}{N}\right) \to \frac{4}{6} = \frac{2}{3}.$$

因此，當 N 很大時，

$$\frac{\mu}{m} = \frac{\mu/N}{m/N} \approx \frac{2/3}{1/\sqrt{2}} = \frac{2\sqrt{2}}{3}.$$

驗算

- 取 $N = 2$：四對等可能的結果中，$X = 1$ 出現一次，$X = 2$ 出現三次，與第 (i) 部分的 $\frac{1}{4}$ 和 $\frac{3}{4}$ 相符；而 $\mu = \frac{7}{4}$，與 $\frac{3\times 7}{12}$ 相符。
- 對 $N \le 40$，用電腦列出全部 N^2 對結果，可確認第 (i) 和第 (ii) 部分的結果完全正確；對 $N \le 2000$，把每個整數逐一代入中位數的定義檢驗，每次都恰好找到一個中位數，並且等於 $\lceil N/\sqrt{2} \rceil$。
- 當 $N = 10^5$ 時，$\mu/m \approx 0.942812$，而 $2\sqrt{2}/3 \approx 0.942809$。

常見錯誤

- 把 $\mathrm{P}(X = r)$ 算成 $2 \times \frac{r}{N} \cdot \frac{1}{N}$，忘記了這樣會把 $X_1 = X_2 = r$ 的情況數了兩次；正確的數目是 $2(r-1)+1 = 2r-1$。
- 寫成 $\mathrm{P}(X \ge m) = 1 - m^2/N^2$：這其實是 $\mathrm{P}(X > m)$，而不是 $\mathrm{P}(X \ge m)$。
- 在 $N = 100$ 時答 $m = 70$，原因是把 70.71 四捨五入至最接近的整數，而不是向上取整。

STEP 之外

X 是*順序統計量*（order statistic）最簡單的例子：樣本的最大值。當 $N \to \infty$ 時，X/N 趨近一個在 $[0,1]$ 上的連續隨機變量（continuous random variable），其密度函數為 $2x$，平均值是 $\frac{2}{3}$，中位數是 $1/\sqrt{2}$，這正是第 (iv) 部分出現該比值的原因。較大樣本的順序統計量會引出貝塔分佈（beta distribution）。Blitzstein 與 Hwang 合著的 *Introduction to Probability*（概率論入門）以淺白易明的方式講解順序統計量。

詳細解答

STEP I 2008，第 13 題

概率與統計 • 環形排列

> **STEP I 2008，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/08-s1-q13

關鍵思路

在圓桌上，只有各人之間的相對位置才重要，所以我們固定其中一人的座位，再把其餘五人排在他周圍。這樣共有 $5! = 120$ 種等可能的排列（arrangement），而每個概率（probability）都是在 120 種之中點算所得的比例。必須相鄰而坐的一對夫婦，可以黏合成一個「組塊」，組塊內部有兩種次序。至於「恰好」和「沒有」的情況，我們可以直接點算，也可以對「第 k 對夫婦相鄰而坐」這些事件（event）運用容斥原理（inclusion–exclusion）。

解答

把三對夫婦記為 A,a; B,b; C,c。所有人沿桌邊同時移動一個位置，各人之間的相對座次並沒有改變。因此我們固定 A 的座位，再按順時針方向把其餘五人依次安排在餘下的座位上。這樣共有 $5! = 120$ 種排列，全部等可能。

假如我們改為替座位編號，便會有 $6! = 720$ 種排列。這時每種相對排列都出現 6 次，分別對應 A 可以坐的每一個座位，所以概率不變。

一個有用的一般結果：把 n 件物件排成一個圓圈，共有 $(n-1)!$ 種排法（先固定一件物件，再排列其餘的）。

第 (i) 部分

如果每對夫婦都相鄰而坐，就把每對夫婦黏合成一個組塊。這樣圓桌旁共有 3 個組塊，排法有 $(3-1)! = 2$ 種。在每個組塊內，丈夫可以坐在妻子的任何一邊，所以有 $2^3 = 8$ 種選擇。因此

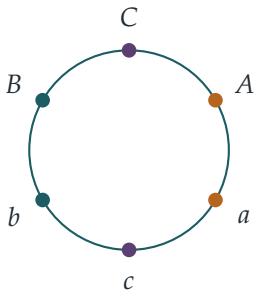
$$\mathrm{P}(\text{三對夫婦全部相鄰}) = \frac{2 \times 8}{120} = \frac{16}{120} = \frac{2}{15},$$

正如所求。

第 (ii) 部分

首先選出哪一對夫婦不相鄰，共有 3 種選擇。假設是 C, c，即 A, a 和 B, b 各自相鄰，但 C 和 c 不相鄰。

把 Aa 和 Bb 黏合成組塊。現在圓桌旁有 4 個單位：Aa, Bb, C, c，排法有 $3! = 6$ 種。固定 C，再看 c 可以落在這個 4 單位圓圈的其餘三個位置中的哪一個。其中兩個位置上 c 與 C 相鄰，另一個位置上 c 則在正對面。所以在這 6 種排列中，C 和 c 分開的恰好有 $6/3 = 2$ 種。在這些排列中，C 與 c 之間的兩個空隙各由一個組塊填滿，如圖所示：



想一想 為甚麼這比看起來簡單？

當 C 與 c 相對而坐時，餘下的自由度只有哪個組塊坐在哪一邊（2 種），以及每個組塊內部的次序（$2 \times 2 = 4$ 種）。這樣共有 $2 \times 4 = 8$ 種排列。把分開的那對夫婦的三種選擇全部計算在內，

$$\mathrm{P}(\text{恰好兩對夫婦相鄰}) = \frac{3 \times 8}{120} = \frac{24}{120} = \frac{1}{5}.$$

第 (iii) 部分

設 E_k 為第 k 對夫婦相鄰而坐的事件，其中 $k = 1, 2, 3$。我們在 120 種排列之中點算每一類排列的數目。

- $|E_1|$：把第 1 對夫婦黏合成一個組塊。圓桌旁便剩下 5 個單位，所以有 $4! = 24$ 種排列，每種有 2 種內部次序，合共 48 種。
- $|E_1 \cap E_2|$：兩個組塊加上兩個人，共 4 個單位。排列有 $3! = 6$ 種，內部次序有 $2^2 = 4$ 種，合共 24 種。
- $|E_1 \cap E_2 \cap E_3| = 16$，由第 (i) 部分得出。

根據對稱性（symmetry），無論選取哪幾對夫婦，以上數目都相同。由容斥原理，至少

有一對夫婦相鄰的排列數目是

$$|E_1 \cup E_2 \cup E_3| = 3(48) - 3(24) + 16 = 144 - 72 + 16 = 88.$$

所以有 $120 - 88 = 32$ 種排列沒有任何夫婦相鄰，而

$$\mathrm{P}(\text{沒有丈夫與自己的妻子相鄰}) = \frac{32}{120} = \frac{4}{15}.$$

想一想 第 (ii) 部分的答案可以從同一組數目直接讀出嗎？

可以。每個 $|E_i \cap E_j|$ 所點算的，是恰好只有 i, j 這兩對相鄰的排列，再加上三對全部相鄰的排列。所以恰好兩對夫婦相鄰的排列數目是 $3(24) - 3(16) = 24$，與第 (ii) 部分吻合。

驗算

- 四種情況必須涵蓋全部 120 種排列。恰好一對夫婦相鄰的排列數目是 $3(48) - 2 \cdot 3(24) + 3(16) = 48$，而 $32 + 48 + 24 + 16 = 120$。
- 用電腦窮舉全部 720 種有編號的座次，得出 $0, 1, 2, 3$ 對夫婦相鄰的概率分別為 $\frac{4}{15}, \frac{2}{5}, \frac{1}{5}, \frac{2}{15}$。這些結果與上面的答案吻合。
- 合理性檢查：某位妻子的兩個鄰座是從 5 個人中選出的，所以她丈夫坐在她旁邊的概率是 $\frac{2}{5}$。這與 $|E_1|/120 = 48/120$ 一致。

常見錯誤

- 總數用了 6! 種排列，卻把符合條件的座次當作環形排列來點算，或者反過來。分子和分母必須採用同一種點算方式。
- 忘記每個組塊內兩人的次序所帶來的因數 2。
- 在第 (ii) 部分，忘記剔除第三對夫婦也相鄰的排列。若不剔除，$\frac{3\times 24}{120} = \frac{3}{5}$ 就大得太多了。

STEP 之外

這是*夫婦入座問題*（*ménage problem*）的縮影。經典版本還要求男女相間而坐，而它的一般答案（Touchard 公式）正是來自這一類容斥原理的運用。在我們這個不設限制的版本中，若有 n 對夫婦，相鄰而坐的夫婦對數的期望值（expected value）是 $n \cdot \frac{2}{2n-1} \to 1$。

沒有任何夫婦相鄰的概率趨向 e^{-1}，與著名的錯排（derangement）極限一模一樣。在男女相間的 ménage 版本中，這個概率則趨向 e^{-2}。Grimmett 與 Welsh 所著的 *Probability: An Introduction*（概率入門教材）對容斥原理和這類極限有很好的論述。

詳細解答

STEP II 2008，第 1 題

純數學 • 複數迭代

> **STEP II 2008，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/08-s2-q1

關鍵思路

$x^2 - y^2$ 和 $2xy$ 這兩個式子，正是 $(x + \mathrm{i}y)^2$ 的實部（real part）和虛部（imaginary part）。因此，若把每個點寫成複數（complex number）$z_n = x_n + \mathrm{i}y_n$，這個含兩個分量的規則就化為單一方程 $z_{n+1} = z_n^2 + c$，當中只有一個複常數 c。這樣，「常數」和「週期為 2」兩個條件都變成只含一個複未知數的二次方程（quadratic equation），可以用二次公式（quadratic formula）求解。

解答

設 $z_n = x_n + \mathrm{i}y_n$ 及 $c = a + (b+2)\mathrm{i}$。由於

$$z_n^2 + c = (x_n^2 - y_n^2 + a) + \mathrm{i}\,(2x_n y_n + b + 2),$$

而兩個複數相等，當且僅當它們的實部和虛部分別相等，所以該規則等價於

$$z_{n+1} = z_n^2 + c.$$

因為 a、b、x_n 和 y_n 都是實數，每個點 (x, y) 恰好對應一個複數，反之亦然。所以改用 z 來計算，不會有任何損失。

第 (i) 部分

當 $a = 1$、$b = -1$ 時，$c = 1 + \mathrm{i}$。數列（sequence）為常數數列，當且僅當 $z_2 = z_1$，因為此後每一步都只是重複同一個計算。所以我們需要

$$z^2 - z + (1 + \mathrm{i}) = 0, \qquad z = \frac{1 \pm \sqrt{1 - 4(1+\mathrm{i})}}{2} = \frac{1 \pm \sqrt{-3 - 4\mathrm{i}}}{2}.$$

要求 $-3-4\mathrm{i}$ 的一個平方根，可設 $(p+q\mathrm{i})^2=-3-4\mathrm{i}$，其中 p、q 為實數：即 $p^2-q^2=-3$ 及 $pq=-2$。取 $p=1$、$q=-2$ 即可，因為 $(1-2\mathrm{i})^2=1-4\mathrm{i}-4=-3-4\mathrm{i}$。因此

$$z=\frac{1\pm(1-2\mathrm{i})}{2}=1-\mathrm{i}\quad\text{或}\quad\mathrm{i}.$$

所以數列為常數數列，當且僅當

$$(x_1,y_1)=(1,-1)\quad\text{或}\quad(x_1,y_1)=(0,1).$$

直接驗證：$(1,-1)\mapsto(1-1+1,\,-2+1)=(1,-1)$ 及 $(0,1)\mapsto(0-1+1,\,0+1)=(0,1)$。

第 (ii) 部分

現在 $z_1=-1+\mathrm{i}$，所以 $z_1^2=1-2\mathrm{i}-1=-2\mathrm{i}$，而

$$z_2=c-2\mathrm{i}.$$

記 $w=z_2$。當 $z_3=z_1$ 但 $z_2\neq z_1$ 時，數列的週期（period）為 2。一旦 $z_3=z_1$，按照規則，數列便會以 $z_1,z_2,z_1,z_2,\ldots$ 的形式不斷重複。利用 $c=w+2\mathrm{i}$，

$$z_3=w^2+c=w^2+w+2\mathrm{i},$$

所以 $z_3=z_1$ 變成

$$w^2+w+2\mathrm{i}=-1+\mathrm{i}\quad\Longleftrightarrow\quad w^2+w+(1+\mathrm{i})=0.$$

? **想一想** 為甚麼選 $w=z_2$ 作未知數，而不選 c？

無論用哪一個作未知數，條件 $z_3=z_1$ 都是一個關於 c 的二次方程；但若以 w 表示，常數數列 $w=z_1$ 必定是其中一個根（root）。所以我們已經知道這個二次方程的一個根，而另一個根正是我們要找的。事實上 $w=-1+\mathrm{i}$ 確實滿足方程：$(-1+\mathrm{i})^2+(-1+\mathrm{i})+1+\mathrm{i}=-2\mathrm{i}+2\mathrm{i}=0$。兩根之和為 -1，所以另一個根是 $w=-\mathrm{i}$。（用二次公式也得到相同結果：$w=\frac{1}{2}\big(-1\pm\sqrt{-3-4\mathrm{i}}\big)=\frac{1}{2}\big(-1\pm(1-2\mathrm{i})\big)$，所用的平方根與 (i) 相同。）

- $w=-1+\mathrm{i}=z_1$ 給出 $c=-1+3\mathrm{i}$，即 $a=-1$、$b=1$。這時 $z_2=z_1$，數列是常數數列，所以它的週期是 1，而不是 2。
- $w=-\mathrm{i}$ 給出 $c=\mathrm{i}$，即 $a=0$ 及 $b+2=1$。

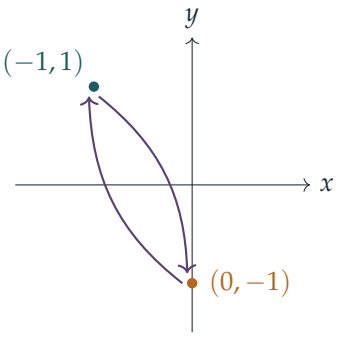
所以數列的週期為 2，當且僅當

$$a=0,\qquad b=-1.$$

用坐標驗算：當 $a = 0$、$b = -1$ 時，

$$(-1,1) \mapsto (1-1+0,\ -2+1) = (0,-1) \mapsto (0-1+0,\ 0+1) = (-1,1),$$

而這兩點並不相同，所以週期恰好是 2。



驗算

- 按原來的坐標規則迭代（iterate），可見當 $a = 1$、$b = -1$ 時，$(1,-1)$ 和 $(0,1)$ 都固定不動；而當 $a = 0$、$b = -1$ 時，$(-1,1) \to (0,-1) \to (-1,1)$。
- 平方根 $1 - 2\mathrm{i}$ 可以通過把它平方來驗證；因式分解（factorisation）$w^2 + w + 1 + \mathrm{i} = (w + 1 - \mathrm{i})(w + \mathrm{i})$ 同樣可以驗證：兩根之積為 $(-1+\mathrm{i})(-\mathrm{i}) = 1 + \mathrm{i}$。
- 以隨機起始值進行數值求根搜尋，在 (i) 中找不到其他不動點（fixed point），在 (ii) 中除了給出常數數列的 $(-1,1)$ 之外，也找不到其他滿足 $z_3 = z_1$ 的 (a,b)。

常見錯誤

- 忘記第二個分量中的「+2」，於是取了 $c = a + b\mathrm{i}$，而不是 $a + (b+2)\mathrm{i}$。
- 把 $a = -1$、$b = 1$ 當作 (ii) 的第二個答案。該數列是常數數列，所以它的週期是 1。
- 只用實坐標計算，結果在聯立二次方程中迷失方向。這樣做是可行的，但複數形式簡短得多。

STEP 之外

映射（map）$z \mapsto z^2 + c$ 是複動力系統（complex dynamics）中最著名的例子。曼德博集合（Mandelbrot set）是使 0 的軌道（orbit）保持有界的所有 c 所組成的集合；它最大的兩個部分，即主心臟線（main cardioid）和其左方的圓盤，分別由使該映射具有吸引

不動點（attracting fixed point）或吸引 2-循環（attracting 2-cycle）的 c 組成。在本題中，2-循環的方程同樣有一個來自不動點的因式：一般而言 $(z^2+c)^2+c-z=(z^2-z+c)(z^2+z+c+1)$。想了解這背後的幾何，可參閱 Needham 的 *Visual Complex Analysis*（《視覺複分析》），這是一本易讀的入門書。

詳細解答

STEP II 2008，第 2 題

純數學 • 部分分式與級數

STEP II 2008，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/08-s2-q2

關鍵思路

部分分式（partial fractions）把一個難以處理的有理函數（rational function）拆成三項，而這三項的展開式我們都爛熟於心：一個等比級數（geometric series）、它的導數，以及一個以 $-x^2$ 為公比的等比級數。頭兩項給出的係數大致隨 n 增長；第三項的貢獻永遠只是 $\pm\frac{1}{2}$，並以週期 4 重複，題目中「非此即彼」的兩種情況正是由此而來。至於最後一部分，要找一個 x 的值，使函數變成題目所給的分數。

解答

部分分式

由於 $(1-x)^2$ 是重複的一次因式，而 $1+x^2$ 是不可約二次式（irreducible quadratic），我們設

$$\frac{1+x}{(1-x)^2(1+x^2)} = \frac{A}{1-x} + \frac{B}{(1-x)^2} + \frac{Cx+D}{1+x^2},$$

於是對所有 x，

$$1+x = A(1-x)(1+x^2) + B(1+x^2) + (Cx+D)(1-x)^2.$$

- 代入 $x=1$：$2 = 2B$，所以 $B=1$。
- 代入 $x=0$：$1 = A+B+D$，所以 $A=-D$。
- 代入 $x=\mathrm{i}$（這是容許的，因為這是一個多項式恆等式）：$1+\mathrm{i} = (C\mathrm{i}+D)(1-\mathrm{i})^2 = (C\mathrm{i}+D)(-2\mathrm{i}) = 2C-2D\mathrm{i}$。比較實部和虛部，得 $C=\frac{1}{2}$ 和 $D=-\frac{1}{2}$，因此 $A=\frac{1}{2}$。

作為驗算，右邊 x^3 的係數是 $-A+C=0$，與左邊相符。所以

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

係數

當 $|x|<1$ 時，三個展開式全部收斂（converge），而且

$$\begin{aligned}\frac{1}{2(1-x)}&=\tfrac{1}{2}\sum_{n\geq 0}x^n, \qquad \frac{1}{(1-x)^2}=\sum_{n\geq 0}(n+1)x^n,\\ -\frac{1-x}{2(1+x^2)}&=-\tfrac{1}{2}(1-x)\left(1-x^2+x^4-x^6+\cdots\right)\\ &=-\tfrac{1}{2}\left(1-x-x^2+x^3+x^4-x^5-x^6+x^7+\cdots\right).\end{aligned}$$

最後一行括號內的係數是 $1,-1,-1,1$，以週期 4 重複：當 $n=2k$ 時，x^n 的係數是 $(-1)^k$；當 $n=2k+1$ 時，則是 $-(-1)^k$。所以，當 $n\equiv 0$ 或 $3\pmod 4$ 時，第三項對 x^n 的貢獻是 $-\frac{1}{2}$；當 $n\equiv 1$ 或 $2\pmod 4$ 時，貢獻是 $+\frac{1}{2}$。

> ? **想一想** 為甚麼週期為 4 的規律只產生兩種情況，而不是四種？

因為只有正負號才有影響：把三項的貢獻相加，

$$a_n=\tfrac{1}{2}+(n+1)\mp\tfrac{1}{2}=\begin{cases}n+1, & n\equiv 0\text{或}3 \pmod 4,\\ n+2, & n\equiv 1\text{或}2 \pmod 4,\end{cases}$$

這就是所要求的結果。這個數列的首幾項是 $1,3,4,4,5,7,8,8,9,11,12,12,\ldots$

小數近似值

> ? **想一想** x 取甚麼值時，函數會等於 $\frac{11\,000}{8181}$？

從分母可以猜到 $x=\frac{1}{10}$：這時 $(1-x)^2=0.81$ 而 $1+x^2=1.01$，所以 $(1-x)^2(1+x^2)=0.8181$，並且

$$\frac{1+\frac{1}{10}}{0.8181}=\frac{1.1}{0.8181}=\frac{11\,000}{8181}.$$

由於 $\frac{1}{10}<1$，級數展開成立，而且

$$\frac{11\,000}{8181}=\sum_{n\geq 0}\frac{a_n}{10^n}=1+0.3+0.04+0.004+0.0005+0.00007+\cdots.$$

只要 $a_n\leq 9$（即直至 $n=8$ 為止），各係數就正好是小數的各個數字：1.34457889。此後

的係數都是兩位數，所以必須進位：

$$\begin{aligned}\sum_{n=0}^{8}\frac{a_n}{10^n} &= 1.344\,578\,89\\ +11\times10^{-9} &\to 1.344\,578\,901\\ +12\times10^{-10} &\to 1.344\,578\,902\,2\\ +12\times10^{-11} &\to 1.344\,578\,902\,32\\ +13\times10^{-12} &\to 1.344\,578\,902\,333.\end{aligned}$$

其後各項只會影響第十一位有效數字（significant figure）及以後的數字，所以準確至九位有效數字，

$$\frac{11\,000}{8181}\approx 1.344\,578\,90.$$

驗算

- 在函數中代入 $x=0$ 得 $1=a_0$，而 $n=0$ 屬於 $n+1$ 的情況。把 $(1+x)(1+2x+\cdots)(1-x^2+\cdots)$ 乘開，得 $1+3x+\cdots$，所以 $a_1=3=1+2$。
- 用長除法（long division）得 $11\,000/8181=1.344\,578\,902\,33\ldots$，與級數的結果一致。我們亦以精確的冪級數（power series）除法核對了 $n<60$ 的 a_n，並在多個隨機選取的有理點上核對了部分分式恆等式。

常見錯誤

- 忘記進位：直接把 a_n 當作各位數字讀出，會得到 $1.34457889\ldots$，第九位有效數字就錯了。
- 把 $\frac{Cx+D}{1+x^2}$ 寫成 $\frac{D}{1+x^2}$；不可約二次式的分子必須是一次式。
- 把 $1.344578902\ldots$ 捨入為 1.34457890 之後，卻略去末尾的零；這個零在這裏是一位有效數字。

STEP 之外

有理函數必定有冪級數展開式，其係數是 n 的多項式，再加上形如 r^n 的項和週期性的規律，分母的每個根各對應一項；這就是*線性遞推關係*（linear recurrences）和*生成函數*（generating functions）的理論。在這一題，根 1（二重根）和 $\pm\mathrm{i}$ 分別帶來線性增長和週期為 4 的起伏。這也解釋了為甚麼每個分數的小數展開最終都會循環。Wilf 的

generatingfunctionology 是一本生動有趣、可免費取得的入門書，而 Siklos 的 *Advanced Problems in Mathematics* 則有更多同類風格的題目。

詳細解答

STEP II 2008，第 3 題

純數學 • 不等式與反證法

STEP II 2008，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/08-s2-q3

關鍵思路

從草圖可見，當 $x \geq 0$ 時，這條三次曲線（cubic）永遠不會跌到 x 軸之下；把「$y \geq 0$」重新整理，便得到不等式 $x^2(1-x) \leq 4/27$。至於反證法（proof by contradiction）的部分，留意 $bc(1-a)$、$ca(1-b)$、$ab(1-c)$ 這三個數逐個來看都不好處理，但它們的乘積可以重新組合成 $a^2(1-a)\cdot b^2(1-b)\cdot c^2(1-c)$，而這正好受第一個結果控制。第 (ii) 部分用的是同一個技巧，只是把三次式換成二次式（quadratic）。

解答

第 (i) 部分：曲線

對 $y = 27x^3 - 27x^2 + 4$ 求導，

$$\frac{\mathrm{d}y}{\mathrm{d}x} = 81x^2 - 54x = 27x(3x-2),$$

它在 $x = 0$ 和 $x = \frac{2}{3}$ 時為零。當 $x = 0$ 時，$y = 4$；當 $x = \frac{2}{3}$ 時，

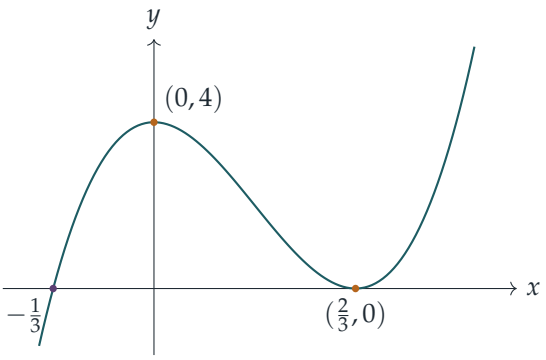
$$y = 27\cdot\tfrac{8}{27} - 27\cdot\tfrac{4}{9} + 4 = 8 - 12 + 4 = 0.$$

由於 $\mathrm{d}^2y/\mathrm{d}x^2 = 162x - 54$ 在 $x = 0$ 時為 $-54 < 0$，在 $x = \frac{2}{3}$ 時為 $54 > 0$，所以轉向點（turning point）是極大點（maximum）$(0, 4)$ 和極小點（minimum）$(\frac{2}{3}, 0)$。

極小點位於 x 軸上，所以曲線在該處與橫軸相切，而 $(3x-2)^2$ 是一個因式（factor）。比較 x^3 的係數和常數項，得

$$27x^3 - 27x^2 + 4 = (3x-2)^2(3x+1),$$

所以曲線與橫軸相交的另一處只有 $x = -\frac{1}{3}$。曲線從 $-\infty$ 上升，在 $x = -\frac{1}{3}$ 穿過橫軸，在 $(0,4)$ 達到頂點，然後下降至 $(\frac{2}{3},0)$ 與橫軸相切，之後再度上升。



第 (i) 部分：不等式

當 $x \geq 0$ 時，我們有 $3x+1 > 0$ 和 $(3x-2)^2 \geq 0$，所以由因式分解（factorisation）可得 $27x^3 - 27x^2 + 4 \geq 0$。（從草圖看也是一樣：在 $x \geq 0$ 的範圍內，曲線的最低點就是極小點 $(\frac{2}{3},0)$。）重新整理，

$$4 \geq 27x^2 - 27x^3 = 27x^2(1-x), \qquad \text{所以} \qquad x^2(1-x) \leq \frac{4}{27} \quad \text{對所有} x \geq 0,$$

而等號只在 $x = \frac{2}{3}$ 時成立。這就是所要求的結果。

第 (i) 部分：反證

用反證法，假設三個數全都大於 4/27：

$$bc(1-a) > \tfrac{4}{27}, \qquad ca(1-b) > \tfrac{4}{27}, \qquad ab(1-c) > \tfrac{4}{27}.$$

這裏六個量全是正數，所以我們可以把三條不等式相乘：

$$bc(1-a)\cdot ca(1-b)\cdot ab(1-c) > \left(\tfrac{4}{27}\right)^3.$$

? **想一想** 第一個結果怎樣才能對這些混合的乘積有所說明？

把左邊重新組合，讓每個字母與它自己的括號歸在一起：

$$bc(1-a)\cdot ca(1-b)\cdot ab(1-c) = \left[a^2(1-a)\right]\left[b^2(1-b)\right]\left[c^2(1-c)\right].$$

a,b,c 每個都是非負數，所以根據第一個結果，每個方括號的值都介乎 0（因為 $1-a \geq 0$，其餘類推）與 4/27 之間。把這三條非負數之間的不等式相乘，

$$\left[a^2(1-a)\right]\left[b^2(1-b)\right]\left[c^2(1-c)\right] \leq \left(\tfrac{4}{27}\right)^3.$$

這與上面的嚴格不等式矛盾。所以我們的假設不成立，即 $bc(1-a)$、$ca(1-b)$、$ab(1-c)$ 之中至少有一個小於或等於 $4/27$。

第 (ii) 部分

同一個想法對二次式也行得通。對任意實數 t，

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

這可由配方法(completing the square)得出。用反證法，假設 $p(1-q) > \frac{1}{4}$ 及 $q(1-p) > \frac{1}{4}$。兩邊都是正數，所以相乘可得

$$p(1-q)\cdot q(1-p) > \tfrac{1}{16}.$$

但左邊可以重新組合成 $\big[p(1-p)\big]\big[q(1-q)\big]$，而由於 $0 < p, q < 1$，每個方括號的值都介乎 0 與 $\frac{1}{4}$ 之間，所以乘積至多為 $\frac{1}{16}$。這是矛盾的，所以 $p(1-q)$ 和 $q(1-p)$ 之中至少有一個小於或等於 $\frac{1}{4}$。

驗算

- 等號成立的情況：$a = b = c = \frac{2}{3}$ 使 (i) 中的三個數全都等於 $\frac{4}{9}\cdot\frac{1}{3} = \frac{4}{27}$，而 $p = q = \frac{1}{2}$ 使 (ii) 中的兩個數都等於 $\frac{1}{4}$。所以 $4/27$ 和 $1/4$ 這兩個界不能再改進。
- 展開 $(3x-2)^2(3x+1) = (9x^2 - 12x + 4)(3x+1) = 27x^3 - 27x^2 + 4$，確認因式分解無誤。
- 以數值計算，$x^2(1-x)$ 在 $[0,3]$ 上的最大值是 $0.148148\ldots = 4/27$。在 $(0,1)$ 內隨機抽取一百萬組三元組 (a,b,c)，三個乘積中最小的一個從未超過 0.1468；隨機抽取一百萬對數值，$p(1-q)$、$q(1-p)$ 中較小的一個從未超過 0.24998。

常見錯誤

- 把不等式相乘，卻沒有說明所有量都是正數。「$A > \alpha$、$B > \beta$ 蘊含 $AB > \alpha\beta$」這一步對負數並不成立，所以 $0 < a, b, c < 1$ 這個假設確實是必需的。
- 在 $x < 0$ 時使用 $x^2(1-x) \leq 4/27$ 這個界，但它在該範圍並不成立（試試 $x = -1$）。
- 只證明某一個特定的數（例如 $bc(1-a)$）很小。它不一定小：當 a 接近 0 而 $b = c$ 接近 1 時，它接近 1。

STEP 之外

把若干個量相乘，使一個對稱的界得以派上用場，這種做法與算術—幾何平均不等式（AM–GM inequality）十分相近，而後者可以直接給出第一個結果：對 $0 \leq x \leq 1$，把算術—幾何平均不等式應用於 $\frac{x}{2}, \frac{x}{2}, 1-x$（三者之和為 1），便得 $\frac{x^2}{4}(1-x) \leq \left(\frac{1}{3}\right)^3$。這個論證也帶有鴿巢原理（pigeonhole principle）的味道：如果乘積很小，就不可能每個因子都很大。這些關於估界和求最優值的想法，會在大學一年級的分析（analysis）課程中進一步發展；Abbott 的 *Understanding Analysis*（分析學入門讀物）是個平易近人的起點，而 Siklos 的 *Advanced Problems in Mathematics*（進階數學難題集）則有更多這類風格的 STEP 題目。

詳細解答

STEP II 2008，第 4 題

純數學 • 隱函數微分法

> **STEP II 2008，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/08-s2-q4

關鍵思路

處理兩條直線之間的夾角，用向量（vector）最為方便：$F(x,y) = 1$ 的法線（normal）沿 ∇F 的方向，而對於平面上兩個向量 $\mathbf{p}$ 和 $\mathbf{n}$，兩者夾角的 tan 等於 |叉積| 除以點積（dot product）。在這一題，點積原來恰好是曲線方程的左方，即 1。之後，θ 取駐值（stationary value）當且僅當 $y^2 - x^2$ 取駐值，餘下的就是關於 $x^2 + y^2$ 和 xy 的代數運算。

解答

斜率

把 $x^2 + y^2 + 2axy = 1$ 對 x 作隱函數微分（implicit differentiation），

$$2x + 2y\frac{\mathrm{d}y}{\mathrm{d}x} + 2a\left(y + x\frac{\mathrm{d}y}{\mathrm{d}x}\right) = 0 \quad \Longrightarrow \quad (ax + y)\frac{\mathrm{d}y}{\mathrm{d}x} = -(x + ay).$$

因此，只要 $ax + y \neq 0$，斜率（gradient）就是 $-\dfrac{x + ay}{ax + y}$，正如題目所求。

角 θ

切線（tangent）的方向為 $(ax + y,\ -(x + ay))$，所以法線的方向為

$$\mathbf{n} = (x + ay,\ ax + y),$$

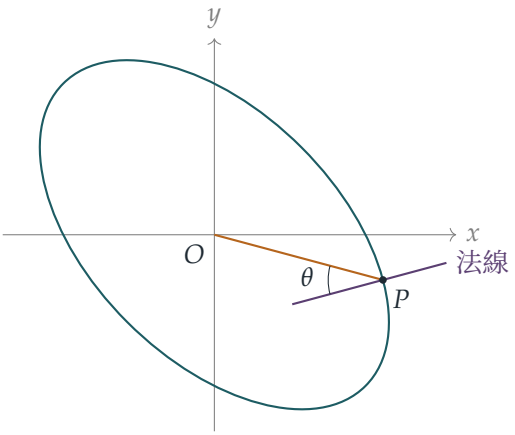
即與它垂直的向量。（這就是 $F = x^2 + y^2 + 2axy$ 的 $\frac{1}{2}\nabla F$，即使在 $ax + y = 0$ 的地方也有意義。）向量 $\overrightarrow{OP}$ 是 $\mathbf{p} = (x, y)$。設 ϕ 為 $\mathbf{p}$ 與 $\mathbf{n}$ 之間的夾角，則

$$|\mathbf{p}||\mathbf{n}|\cos\phi = \mathbf{p}\cdot\mathbf{n} = x(x+ay) + y(ax+y) = x^2 + y^2 + 2axy = 1,$$
$$|\mathbf{p}||\mathbf{n}|\sin\phi = |x(ax+y) - y(x+ay)| = |ax^2 - ay^2| = a|y^2 - x^2|,$$

其中第二行是二維叉積（cross product）的大小（即由 $\mathbf{p}$ 和 $\mathbf{n}$ 所張成的平行四邊形的面積），而 $a > 0$。

由於點積為 $+1 > 0$，ϕ 本身已是銳角（或零），所以 $\phi = \theta$。把兩行相除，得

$$\tan\theta = a\,|y^2 - x^2|.$$



上面的草圖（取 $a = \frac{1}{2}$）顯示了這條曲線：由於 $0 < a < 1$，它是一個橢圓（ellipse），兩軸分別沿 $y = \pm x$。圖中所畫的點 P 是 θ 最大的其中一點。

第 (i) 部分

函數 $t \mapsto \arctan(t)$ 嚴格遞增（strictly increasing），所以凡在 $f \neq 0$ 之處，θ 都是 $f = y^2 - x^2$ 的光滑遞增函數。

> **想一想** 在 $y^2 = x^2$ 的點，絕對值（modulus）出現尖角，那又如何？

在這些點 $\theta = 0$。假如 θ 在這樣的一點導數（derivative）為零，那麼 $a|f|$ 除以 x 的改變量會趨向 0，這迫使 $\mathrm{d}f/\mathrm{d}x = 0$ 同樣成立。所以在任何情況下都有

$$\frac{\mathrm{d}\theta}{\mathrm{d}x} = 0 \implies \frac{\mathrm{d}}{\mathrm{d}x}(y^2 - x^2) = 0 \implies y\frac{\mathrm{d}y}{\mathrm{d}x} = x.$$

這裏我們用到斜率公式，所以需要 $ax + y \neq 0$（下文會確認這個條件必然成立）。代入後再乘以 $ax + y$：

$$-y(x+ay) = x(ax+y) \implies -xy - ay^2 = ax^2 + xy \implies a(x^2 + y^2) + 2xy = 0.$$

這就是 (i)。

順帶一提，假如 $ax + y = 0$，則由 (i) 會得出 $ax^2(1+a^2) - 2ax^2 = -a(1-a^2)x^2 = 0$，於是 $x = y = 0$，但這一點不在曲線上。所以這類點永遠不會造成麻煩。

第 (ii) 部分

記 $S = x^2 + y^2$。由 (i)，$2xy = -aS$。把它代入曲線方程，

$$S + a(-aS) = 1 \quad \Longrightarrow \quad S = \frac{1}{1-a^2}.$$

因此

$$x^2 + y^2 + 2xy = S - aS = \frac{1-a}{1-a^2} = \frac{1}{1+a},$$

即 $(1+a)(x^2+y^2+2xy) = 1$，正如題目所求。

第 (iii) 部分

我們需要求 $|y^2 - x^2|$。利用 $(y^2-x^2)^2 = (x^2+y^2)^2 - 4x^2y^2 = S^2 - (2xy)^2$，

$$(y^2-x^2)^2 = S^2 - a^2S^2 = (1-a^2)\cdot\frac{1}{(1-a^2)^2} = \frac{1}{1-a^2}.$$

取正平方根（因 $0 < a < 1$，此舉成立），得 $|y^2 - x^2| = 1/\sqrt{1-a^2}$，所以

$$\tan\theta = \frac{a}{\sqrt{1-a^2}}.$$

驗算

- 旋轉坐標軸：設 $u = (x+y)/\sqrt{2}$ 及 $v = (x-y)/\sqrt{2}$，曲線便化為 $(1+a)u^2 + (1-a)v^2 = 1$，而 $y^2 - x^2 = -2uv$。由算術平均—幾何平均不等式（AM–GM），$1 \geq 2\sqrt{(1-a^2)}\,|uv|$，所以 $\tan\theta = 2a|uv| \leq a/\sqrt{1-a^2}$，等號在 $(1+a)u^2 = \frac{1}{2}$ 時成立。由於 $x^2+y^2+2xy = 2u^2$，這個等號條件正正就是 (ii)。所以該駐值是 θ 的最大值。
- 取 $a = \frac{1}{2}$：$P \approx (1.1154, -0.2989)$ 滿足曲線方程和 (i)，而 $\tan\theta = \frac{1}{2}|y^2 - x^2| \approx 0.5774 = 1/\sqrt{3}$，所以 $\theta = 30°$。對若干個 a 值，沿橢圓一周以數值方法掃描 $\tan\theta$，所得的最大值正是這些數值。
- 當 $a \to 0$ 時，曲線變成單位圓，其法線通過 O，所以 $\theta \to 0$，與 $\tan\theta \to 0$ 吻合。

常見錯誤

- 以 OP 和法線的斜率套用 $\tan(\alpha - \beta)$：這方法行得通，但要小心處理 $x = 0$ 的情況和正負號，而向量的做法可同時避開這兩個問題。

- 忘記了為何所求得的角是銳角：關鍵在於點積為 $+1$。
- 以平方求 $|y^2 - x^2|$，其後卻忘記取正根。

STEP 之外

把 $x^2 + y^2 + 2axy$ 寫成二次型（quadratic form）$\mathbf{x}^{\mathsf{T}} A\mathbf{x}$，其中 $A = \begin{pmatrix} 1 & a \\ a & 1 \end{pmatrix}$，則橢圓的兩軸就是 A 的特徵向量（eigenvector）$(1,1)$ 和 $(1,-1)$，而特徵值（eigenvalue）$1 \pm a$ 決定了半軸的長度。對稱矩陣（symmetric matrix）的這種對角化（diagonalisation）是大學一年級線性代數的核心概念；參看 Strang 的 *Introduction to Linear Algebra*（線性代數入門）。

詳細解答

STEP II 2008，第 5 題

純數學 • 代換積分法

STEP II 2008，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/08-s2-q5

關鍵思路

兩個積分(integral)都可以靠看出導數(derivative)來解決。第一個積分中，分子 $\sin 2x = 2\sin x\cos x$ 恰好是分母 $1+\sin^2 x$ 的導數，所以一眼就能看出積分結果是一個對數（logarithm）。第二個積分中，分子只有一個 $\sin x$，這提示我們設 $c = \cos x$，並先把分母寫成 $1+\sin^2 x = 2-\cos^2 x$，餘下的就是一個標準的部分分式（partial fractions）積分。於是，比較兩個答案就變成比較兩個含有 $\sqrt{2}$ 的數，而最後一部分的不等式（inequality）正是為了解決這個問題而設計的。

解答

兩個積分

把兩個積分分別記作 I_1（分子為 $\sin 2x$）和 I_2（分子為 $\sin x$）。

先看 I_1。由於 $\dfrac{\mathrm{d}}{\mathrm{d}x}\left(1+\sin^2 x\right) = 2\sin x\cos x = \sin 2x$，被積函數（integrand）具有 $f'(x)/f(x)$ 的形式，其中 $f(x) = 1+\sin^2 x > 0$。因此

$$I_1 = \Big[\ln\left(1+\sin^2 x\right)\Big]_0^{\pi/2} = \ln 2 - \ln 1 = \ln 2.$$

再看 I_2。寫成 $1+\sin^2 x = 2-\cos^2 x$，並作代換（substitution）$c = \cos x$，於是 $\mathrm{d}c = -\sin x\,\mathrm{d}x$；當 x 由 0 增至 $\frac{1}{2}\pi$ 時，c 由 1 減至 0。負號把積分上下限（limits）對調：

$$I_2 = \int_0^1 \frac{\mathrm{d}c}{2-c^2}.$$

利用部分分式，

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

而當 $0 \le c \le 1$ 時，兩個分母都保持為正，所以

$$I_2 = \frac{1}{2\sqrt{2}}\Big[\ln(\sqrt{2}+c) - \ln(\sqrt{2}-c)\Big]_0^1 = \frac{1}{2\sqrt{2}}\ln\frac{\sqrt{2}+1}{\sqrt{2}-1}.$$

這個結果可以化簡得很漂亮：$(\sqrt{2}+1)(\sqrt{2}-1) = 1$，所以 $\dfrac{\sqrt{2}+1}{\sqrt{2}-1} = (\sqrt{2}+1)^2$，因此

$$I_2 = \frac{1}{2\sqrt{2}} \cdot 2\ln(1+\sqrt{2}) = \frac{\ln(1+\sqrt{2})}{\sqrt{2}}.$$

不等式

根據二項式定理（binomial theorem），並利用 $(\sqrt{2})^2 = 2$、$(\sqrt{2})^3 = 2\sqrt{2}$、$(\sqrt{2})^4 = 4$ 和 $(\sqrt{2})^5 = 4\sqrt{2}$，

$$\begin{aligned}(1+\sqrt{2})^5 &= 1 + 5\sqrt{2} + 10\cdot 2 + 10\cdot 2\sqrt{2} + 5\cdot 4 + 4\sqrt{2} \\ &= 41 + 29\sqrt{2}.\end{aligned}$$

現在 $\sqrt{2} < 2$（因為 $2 < 4$），所以 $29\sqrt{2} < 58$，從而 $(1+\sqrt{2})^5 < 41 + 58 = 99$，這正是所要證明的。

其次，$1.4^2 = 1.96 < 2$，而 1.4 和 $\sqrt{2}$ 都是正數，平方不改變大小次序，所以 $\sqrt{2} > 1.4$。

想一想 $(1+\sqrt{2})^5 < 99$ 和 $\sqrt{2} > 1.4$ 怎麼可能告訴我們任何關於 $2^{\sqrt{2}}$ 的事情？

寫成 $1.4 = \frac{7}{5}$。那麼 $2^{7/5}$ 就是 $2^7 = 128$ 的五次方根，而 $1+\sqrt{2}$ 的五次方正是我們剛剛估計過上界的量。準確地說：由於 $t \mapsto 2^t$ 是遞增（increasing）函數，而且 $\sqrt{2} > \frac{7}{5}$，

$$2^{\sqrt{2}} > 2^{7/5}.$$

另外 $(1+\sqrt{2})^5 < 99 < 128 = \left(2^{7/5}\right)^5$，而 $y \mapsto y^5$ 在 $y > 0$ 時遞增，取正五次方根便得 $1+\sqrt{2} < 2^{7/5}$。兩者結合，

$$2^{\sqrt{2}} > 2^{7/5} > 1 + \sqrt{2}.$$

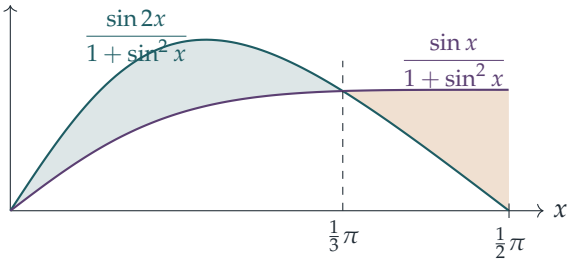
哪個積分較大？

函數 ln 是遞增的，所以由最後一個不等式可得 $\ln\left(2^{\sqrt{2}}\right) > \ln(1+\sqrt{2})$，即 $\sqrt{2}\ln 2 > \ln(1+\sqrt{2})$。兩邊除以 $\sqrt{2} > 0$，

$$I_1 = \ln 2 > \frac{\ln(1+\sqrt{2})}{\sqrt{2}} = I_2.$$

所以分子為 $\sin 2x$ 的那個積分較大。

這個大小關係並不能直接從被積函數看出來。由於 $\sin 2x - \sin x = \sin x\,(2\cos x - 1)$，第一個被積函數在 $0 < x < \frac{1}{3}\pi$ 時較大，而在 $\frac{1}{3}\pi < x < \frac{1}{2}\pi$ 時較小：



問題在於青綠色區域的面積是否大於琥珀色區域的面積。答案是肯定的，但只多出約 0.07。這就是為甚麼要對 $\sqrt{2}$ 作一番仔細的運算。

驗算

- 對原來的被積函數使用辛普森法則（Simpson's rule），得 $I_1 \approx 0.693147$ 和 $I_2 \approx 0.623225$，與 $\ln 2$ 和 $\ln(1+\sqrt{2})/\sqrt{2}$ 吻合至小數點後十位。
- $(1+\sqrt{2})^5 \approx 82.01$，而 $41 + 29\sqrt{2} \approx 82.01$；遠低於 99。
- $2^{\sqrt{2}} \approx 2.665$、$2^{1.4} \approx 2.639$ 和 $1+\sqrt{2} \approx 2.414$，次序與所斷言的一致。
- 對 I_2 作合理性檢查：被積函數介乎 $\frac{1}{2}\sin x$ 與 $\sin x$ 之間，所以 $\frac{1}{2} < I_2 < 1$，而 0.62 確實在此範圍內。

常見錯誤

- 忘記 $\mathrm{d}c = -\sin x\,\mathrm{d}x$ 會把積分上下限對調，結果為一個正的被積函數算出負值。
- 把 I_2 停留在 $\frac{1}{2\sqrt{2}}\ln\frac{\sqrt{2}+1}{\sqrt{2}-1}$ 的形式：若沒有看出對數內的分式就是 $(\sqrt{2}+1)^2$，它與最後一部分中 $1+\sqrt{2}$ 的聯繫便隱而不見。
- 單憑計算機就斷言 $2^{\sqrt{2}} > 1+\sqrt{2}$。題目要求的是經由 $2^{7/5}$ 的那一串不等式，並須說明 2^t 和 y^5 的單調性（monotonicity）。

STEP 之外

$2^{\sqrt{2}}$ 究竟是甚麼意思？在大學裏，它被定義為 $e^{\sqrt{2}\ln 2}$，而我們用過的那個事實，即 2^t 不論 t 是有理數還是無理數都隨 t 遞增，是由指數函數（exponential function）的性質推出的，這些性質會在大學的第一門分析課中嚴格證明。（蓋爾范德（Gelfond）與施奈德（Schneider）的一條著名定理表明，$2^{\sqrt{2}}$ 其實是超越數（transcendental number）。）至於 e^x 和 $\ln x$ 的嚴格構造，可參閱 Abbott 的 *Understanding Analysis*。

詳細解答

STEP II 2008，第 6 題

純數學 • 和化積公式

> **STEP II 2008，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/08-s2-q6

關鍵思路

一個餘弦與一個正弦相加，而兩者的輻角又不相同，處理起來並不方便；但兩個餘弦之和可以用和化積公式（sum-to-product formula）分解成乘積。先把正弦改寫成餘弦；這樣 f 的零點（zeros）就是兩個簡單因子的零點，而兩個因子同時為零的那一點，正是曲線與軸相切之處。至於最後一部分，要記得餘弦和正弦各自最大只能是 1。

解答

以下一律記

$$A = 2x + \frac{\pi}{3}, \qquad \sin\left(\frac{3x}{2} - \frac{\pi}{4}\right) = \cos\left(\frac{\pi}{2} - \frac{3x}{2} + \frac{\pi}{4}\right) = \cos B, \quad \text{其中} B = \frac{3\pi}{4} - \frac{3x}{2},$$

於是 $f(x) = \cos A + \cos B$。

第 (i) 部分

$\cos(2x + \frac{\pi}{3})$ 這一項的周期（period）是 π，而 $\sin(\frac{3x}{2} - \frac{\pi}{4})$ 的周期是 $\frac{2\pi}{3/2} = \frac{4\pi}{3}$。公共周期必須是兩者各自的整數倍：$p\pi = q \cdot \frac{4\pi}{3}$ 即 $3p = 4q$，其最小的正數解是 $p = 4, q = 3$。所以對所有 x 都有 $f(x + 4\pi) = f(x)$，而 f 的周期是

$$4\pi.$$

更小的數都不行：我們會在第 (iii) 部分看到，$f(x) = 2$ 當且僅當 $x = \frac{11\pi}{6} + 4t\pi$（$t$ 為整數），所以 f 取得最大值的各點彼此相隔 4π，任何周期都必定是 4π 的倍數。

第 (ii) 部分

利用 $\cos A + \cos B = 2\cos\frac{A+B}{2}\cos\frac{A-B}{2}$，其中

$$\frac{A+B}{2} = \frac{1}{2}\left(\frac{x}{2} + \frac{13\pi}{12}\right) = \frac{x}{4} + \frac{13\pi}{24},$$
$$\frac{A-B}{2} = \frac{1}{2}\left(\frac{7x}{2} - \frac{5\pi}{12}\right) = \frac{7x}{4} - \frac{5\pi}{24},$$

我們得到

$$f(x) = 2\cos\left(\frac{x}{4} + \frac{13\pi}{24}\right)\cos\left(\frac{7x}{4} - \frac{5\pi}{24}\right).$$

所以 $f(x) = 0$ 當且僅當兩個餘弦之中有一個為零。

第一個因子。由 $\frac{x}{4} + \frac{13\pi}{24} = \frac{\pi}{2} + k\pi$ 得 $x = -\frac{\pi}{6} + 4k\pi$。落在 $[-\pi, \pi]$ 內的只有 $x = -\frac{\pi}{6}$。

第二個因子。由 $\frac{7x}{4} - \frac{5\pi}{24} = \frac{\pi}{2} + k\pi$ 得 $\frac{7x}{4} = \frac{17\pi}{24} + k\pi$，所以

$$x = \frac{(17 + 24k)\pi}{42}.$$

我們需要 $-42 \le 17 + 24k \le 42$，所以 $k = -2, -1, 0, 1$，得出 $x = -\frac{31\pi}{42}$, $-\frac{7\pi}{42} = -\frac{\pi}{6}$, $\frac{17\pi}{42}$, $\frac{41\pi}{42}$。

綜合起來，f 在 $-\pi \le x \le \pi$ 內的零點是

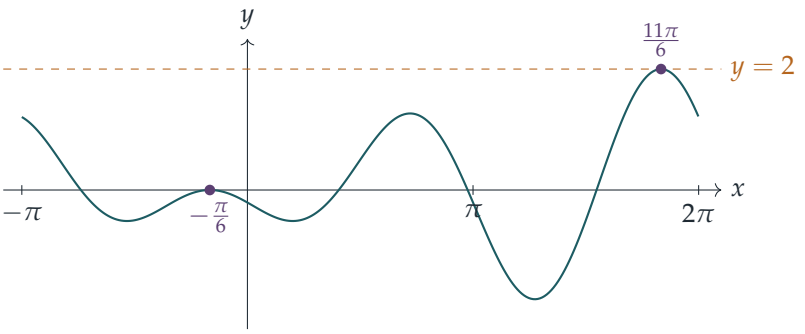
$$x = -\frac{31\pi}{42}, \quad -\frac{\pi}{6}, \quad \frac{17\pi}{42}, \quad \frac{41\pi}{42}.$$

想一想 為甚麼在兩張清單中都出現的值會與別不同？

在 $x = -\frac{\pi}{6}$ 處，兩個餘弦因子都是零。每個因子都是一個穿過零點的餘弦，所以各自在該處變號（其導數（derivative）不為零，因為在 cos 為 0 之處 sin 是 ± 1）。因此在 $x = -\frac{\pi}{6}$ 附近，f 的表現就像一個常數乘以 $(x + \frac{\pi}{6})^2$：兩次變號相乘，結果是沒有變號。事實上，當 x 遞增經過 $-\frac{\pi}{6}$ 時，第一個因子由正變負（其輻角向上穿過 $\frac{\pi}{2}$），第二個因子由負變正（其輻角向上穿過 $-\frac{\pi}{2}$），所以 f 在兩側都是負數。因此曲線與 x-軸相切而不穿過它之處是

$$x = -\frac{\pi}{6}.$$

在其餘三個零點，只有一個因子為零，所以 f 變號，曲線穿過軸。



第 (iii) 部分

由於 $\cos A \le 1$ 及 $\cos B \le 1$，我們有 $f(x) \le 2$，等號成立當且僅當 $\cos A = 1$ 並且 $\cos B = 1$。以原來的函數表示：

$$\cos\left(2x+\frac{\pi}{3}\right)=1 \iff 2x+\frac{\pi}{3}=2m\pi \iff x=-\frac{\pi}{6}+m\pi,$$
$$\sin\left(\frac{3x}{2}-\frac{\pi}{4}\right)=1 \iff \frac{3x}{2}-\frac{\pi}{4}=\frac{\pi}{2}+2n\pi \iff x=\frac{\pi}{2}+\frac{4n\pi}{3},$$

其中 m, n 為整數。兩者同時成立的條件是 $-\frac{\pi}{6}+m\pi=\frac{\pi}{2}+\frac{4n\pi}{3}$；兩邊乘以 $\frac{6}{\pi}$，即 $6m-1=3+8n$，也就是

$$3m-4n=2.$$

其中一個解是 $m=2, n=1$，而通解是 $m=2+4t, n=1+3t$（因為 3 與 4 互質 (coprime)）。所以 $f(x)=2$ 當且僅當

$$x=-\frac{\pi}{6}+(2+4t)\pi=\frac{11\pi}{6}+4t\pi, \qquad t\in\mathbb{Z}.$$

在 $0 \le x \le 2\pi$ 內，這樣的值只有

$$x=\frac{11\pi}{6}.$$

驗算：$2\cdot\frac{11\pi}{6}+\frac{\pi}{3}=4\pi$，而 $\frac{3}{2}\cdot\frac{11\pi}{6}-\frac{\pi}{4}=\frac{5\pi}{2}$，所以 $f=\cos 4\pi+\sin\frac{5\pi}{2}=1+1=2$。

驗算

- 在 10^4 個隨機點上，因式分解後的形式與原來的 f 一致（誤差在 10^{-12} 之內）。
- 以數值方法掃描 $[-\pi,\pi]$，只在 $-\frac{31\pi}{42}$、$\frac{17\pi}{42}$ 和 $\frac{41\pi}{42}$ 附近發現變號；在 $-\frac{\pi}{6}$ 處，我們發現兩側都有 $f(-\frac{\pi}{6}\pm 10^{-3})\approx -8.7\times 10^{-7}$，證實這是一個相切而不變號的零點。
- f 在 $[0,2\pi]$ 上的最大值是 2，在 $x\approx 1.8333\pi=\frac{11\pi}{6}$ 處取得。

常見錯誤

- 把 $-\frac{\pi}{6}$ 列出兩次，或者沒有留意到它同時來自兩個因子，而這恰恰就是找出相切點的線索。
- 分別解出 $\cos A = 1$ 和 $\cos B = 1$ 之後就停下來：兩個條件必須在同一個 x 上成立，這就需要整數方程 $3m - 4n = 2$。

STEP 之外

第 (iii) 部分其實是在問兩個周期運動何時會同步：條件 $3m - 4n = 2$ 是一條線性丟番圖方程（linear Diophantine equation），它有解是因為 $\gcd(3, 4)$ 整除 2。同一個想法若同時應用於多個頻率，就會引向傅立葉級數（Fourier series），以及這個問題：周期函數之和在甚麼情況下本身也是周期函數（只有當各周期之比為有理數時才是）。想了解背後的數論，可以讀 Davenport 的 *The Higher Arithmetic*（一本數論入門經典）。

詳細解答

STEP II 2008，第 7 題

純數學 • 一階微分方程

STEP II 2008，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/08-s2-q7

關鍵思路

左邊的 $\frac{1}{y}\frac{\mathrm{d}y}{\mathrm{d}x}$ 正是 $\ln y$ 的導數（derivative），所以處理這個代換（substitution）的最佳方法是取對數（logarithm）：乘積 $y = u\,(1+x^2)^{1/2}$ 會變成和，而右邊的 $\frac{x}{1+x^2}$ 一項恰好是 $\ln(1+x^2)^{1/2}$ 的導數。這一項消去後，剩下的是一條關於 u 的可分離變數方程（separable equation）。第 (ii) 部分是同一個想法，只是換成立方根；把兩個答案並排一看，第 (iii) 部分的規律便一目了然。

解答

第 (i) 部分

在 $x = 0$ 處有 $y = 1 \neq 0$，而方程中含有 $1/y$，所以我們在 $x = 0$ 附近一個使 $y \neq 0$ 的區間上進行討論。由於 $(1+x^2)^{1/2} > 0$，由代換可知在該區間上亦有 $u \neq 0$，並且 $u(0) = 1$。

對 $|y| = |u|\,(1+x^2)^{1/2}$ 取對數，然後求導，得

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

與原來的微分方程（differential equation）比較，$\frac{x}{1+x^2}$ 各項互相消去，剩下

$$\frac{1}{u}\frac{\mathrm{d}u}{\mathrm{d}x} = xy = x\,u\,(1+x^2)^{1/2}, \qquad \text{所以} \qquad \frac{1}{u^2}\frac{\mathrm{d}u}{\mathrm{d}x} = x(1+x^2)^{1/2},$$

這裏兩邊除以 u，由於 $u \neq 0$，這樣做是容許的。

想一想 怎樣才能乾淨利落地求 $x(1+x^2)^{1/2}$ 的積分？

因子 x 是 $1+x^2$ 的導數的一半，所以由逆向鏈式法則（reverse chain rule）得 $\int x(1+x^2)^{1/2}\,\mathrm{d}x = \frac{1}{2}\cdot\frac{2}{3}(1+x^2)^{3/2} = \frac{1}{3}(1+x^2)^{3/2}$，再加上一個常數。（不妨求導驗證一下。）

兩邊對 x 積分，得

$$-\frac{1}{u} = \tfrac{1}{3}(1+x^2)^{3/2} + C.$$

當 $x=0$ 時，$u=1$，所以 $-1 = \frac{1}{3}+C$，即 $C=-\frac{4}{3}$。因此

$$\frac{1}{u} = \frac{4-(1+x^2)^{3/2}}{3}, \qquad u = \frac{3}{4-(1+x^2)^{3/2}},$$

於是

$$\boxed{y = \frac{3\,(1+x^2)^{1/2}}{4-(1+x^2)^{3/2}}}$$

這個解在包含 0 而分母為正的區間上成立，亦即 $(1+x^2)^{3/2}<4$，即 $|x|<\sqrt{2^{4/3}-1}\approx 1.233$；當 $|x|$ 趨近這個值時，$y\to\infty$。

第 (ii) 部分

這次多出來的一項是 $\frac{x^2}{1+x^3}$，它是 $\frac{1}{3}\ln(1+x^3) = \ln(1+x^3)^{1/3}$ 的導數。這提示我們設 $y = u\,(1+x^3)^{1/3}$，並在 $x=0$ 附近（該處 $1+x^3>0$）進行討論。與上面完全一樣，

$$\frac{1}{y}\frac{\mathrm{d}y}{\mathrm{d}x} = \frac{1}{u}\frac{\mathrm{d}u}{\mathrm{d}x} + \frac{x^2}{1+x^3},$$

所以方程變成

$$\frac{1}{u}\frac{\mathrm{d}u}{\mathrm{d}x} = x^2 y = x^2 u\,(1+x^3)^{1/3}, \qquad \frac{1}{u^2}\frac{\mathrm{d}u}{\mathrm{d}x} = x^2(1+x^3)^{1/3}.$$

這次 x^2 是 $1+x^3$ 的導數的三分之一，所以

$$-\frac{1}{u} = \tfrac{1}{3}\cdot\tfrac{3}{4}(1+x^3)^{4/3} + C = \tfrac{1}{4}(1+x^3)^{4/3} + C.$$

由 $u=1$ 於 $x=0$，得 $-1=\frac{1}{4}+C$，所以 $C=-\frac{5}{4}$。於是 $\frac{1}{u} = \frac{5-(1+x^3)^{4/3}}{4}$，並且

$$\boxed{y = \frac{4\,(1+x^3)^{1/3}}{5-(1+x^3)^{4/3}}}$$

這個解在 0 附近成立，範圍是 $-1<x<\left(5^{3/4}-1\right)^{1/3}\approx 1.33$。

第 (iii) 部分

比較兩個答案：分子出現的是方根 $(1+x^n)^{1/n}$，分母出現的是冪 $(1+x^n)^{(n+1)/n}$，而常數分別是 $n+1$ 和 $n+2$。因此我們猜想（conjecture）

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

當 $n=2$ 和 $n=3$ 時，這個式子正好給出 (i) 和 (ii) 的答案。（題目沒有要求證明，但用同樣的代換 $y = u(1+x^n)^{1/n}$ 便可證明：$\int x^{n-1}(1+x^n)^{1/n}\,\mathrm{d}x = \frac{1}{n+1}(1+x^n)^{(n+1)/n}$，而 $x=0$ 處的條件給出常數 $-\frac{n+2}{n+1}$。）

驗算

- 當 $x=0$ 時，每條公式都給出 $\frac{n+1}{n+2-1} = 1$，符合要求。
- 在 (i) 中 $x=0$ 附近，由方程可得 $y'(0)=0$，而 $(1+x^2)^{1/2}$ 和 $(1+x^2)^{3/2}$ 在 0 處的斜率確實都是零。
- 以數值方法解這些方程（由 $x=0$ 出發，用龍格－庫塔（Runge–Kutta）法逐步計算），所得結果與公式相符，誤差小於 10^{-8}，已對 $n=2,3,4,5,7$ 在多個點上作過檢驗。例如在 (i) 中當 $x=1$ 時，$y = \frac{3\sqrt{2}}{4-2\sqrt{2}} \approx 3.6213$；在 (ii) 中當 $x=1$ 時，$y = \frac{4\cdot 2^{1/3}}{5-2^{4/3}} \approx 2.0320$。

常見錯誤

- 代入 $y = u(1+x^2)^{1/2}$ 後直接用積法則（product rule）求導：這樣做是可行的，但式子很繁瑣。先把全式除以 y（等價於取對數），相消的情況便一目了然。
- 忘記 $\int u^{-2}\,\mathrm{d}u = -u^{-1}$ 中的負號，以致常數出錯。
- 在錯誤的階段確定常數：應在移項整理之前代入 $u(0)=1$（而不是在關於 u 的方程中代入 $y(0)=1$）。

STEP 之外

這條方程是一條伯努利方程（*Bernoulli equation*）：$y' = P(x)y + Q(x)y^2$，通用的技巧是代換 $v = 1/y$，把它化為線性（linear）方程。本題所給的代換則分兩步完成同一件事：先用類似積分因子（integrating factor）的 $(1+x^n)^{1/n}$ 消去含 P 的一項，然後分離變數。還要留意，雖然方程看來平平無奇，它的解卻在有限的 x 處趨向無窮；這種「有限時間

爆破（finite-time blow-up）」是含有 y^2 項的方程的典型現象，在常微分方程（ODE）的入門課程中會學到。Boyce 與 DiPrima 的 *Elementary Differential Equations*（常微分方程入門教科書）涵蓋了伯努利方程和解的存在性。

詳細解答

STEP II 2008，第 8 題

純數學 • 向量與角平分線

STEP II 2008，第 8 題

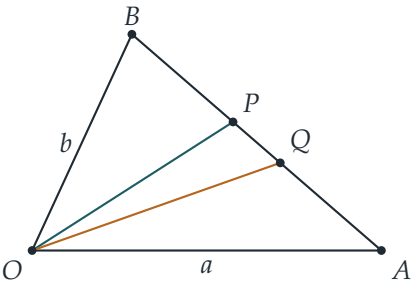
繼續閱讀之前，請先看題目（原題為英文）。題目載於 STEP Support Programme 的試題資料庫：

step.maths.org/questions/08-s2-q8

關鍵思路

沿 $\angle AOB$ 的角平分線（angle bisector）方向的向量，就是沿 OA 和 OB 的單位向量（unit vector）之和：把兩個長度相等的向量相加，得到的是菱形（rhombus）的對角線，而菱形的對角線平分其頂角。把這個向量與 P 的位置向量（position vector）對照，便可定出 λ。之後，Q 不過是把 A 與 B 的角色互換後的 P，而最後的恆等式只需用純量積（scalar product）直接計算，當中棘手的一項 $\mathbf{a} \cdot \mathbf{b}$ 會互相抵銷。

解答



圖中 $a > b$，所以 P（位於角平分線上，以藍綠色標示）較接近 B，而 Q 與 B 的距離，等於 P 與 A 的距離。

P 的位置向量

由於 $AP : PB = (1-\lambda) : \lambda$，而這兩段合起來就是 AB，所以 $AP = (1-\lambda)AB$，因此 $\overrightarrow{AP} = (1-\lambda)(\mathbf{b}-\mathbf{a})$，並且

$$\overrightarrow{OP} = \mathbf{a} + (1-\lambda)(\mathbf{b}-\mathbf{a}) = \lambda\mathbf{a} + (1-\lambda)\mathbf{b}.$$

（驗證：$\lambda = 1$ 時得 A，$\lambda = 0$ 時得 B，與該比例相符。）

求 λ

向量 $\mathbf{a}/a$ 和 $\mathbf{b}/b$ 分別是沿 OA 和 OB 的單位向量。兩者之和是一個菱形的對角線，該菱形的邊分別沿 OA 和 OB，所以它指向 $\angle AOB$ 的角平分線方向。由於 P 位於這條角平分線上，

$$\lambda\mathbf{a} + (1-\lambda)\mathbf{b} = k\left(\frac{\mathbf{a}}{a} + \frac{\mathbf{b}}{b}\right)$$

其中 k 為某個純量。

> **想一想** 為甚麼可以分別比較 **a** 和 **b** 的係數？

因為 O、A、B 不共線（collinear），**a** 和 **b** 均非零向量，而且互不平行，所以 **a** 和 **b** 線性獨立（linearly independent）：若 $\alpha\mathbf{a} + \beta\mathbf{b} = \mathbf{0}$，則 $\alpha = \beta = 0$。因此

$$\lambda = \frac{k}{a}, \qquad 1-\lambda = \frac{k}{b}.$$

兩式相加，得 $1 = k\left(\frac{1}{a} + \frac{1}{b}\right)$，所以 $k = \dfrac{ab}{a+b}$，並且

$$\boxed{\lambda = \frac{b}{a+b}}, \qquad 1-\lambda = \frac{a}{a+b}.$$

這其實就是角平分線定理（angle bisector theorem）的另一種面貌：$AP : PB = (1-\lambda) : \lambda = a : b = OA : OB$。所以

$$\overrightarrow{OP} = \frac{b\,\mathbf{a} + a\,\mathbf{b}}{a+b}.$$

點 Q 與恆等式

Q 在 AB 上，且 $BQ = AP = (1-\lambda)AB$，所以沿用求 P 時的推論，只是改由 B 而非 A 量起，

$$\overrightarrow{OQ} = \mathbf{b} + (1-\lambda)(\mathbf{a}-\mathbf{b}) = (1-\lambda)\mathbf{a} + \lambda\mathbf{b} = \frac{a\,\mathbf{a} + b\,\mathbf{b}}{a+b}.$$

設 $c = \mathbf{a}\cdot\mathbf{b}$，並利用 $\mathbf{a}\cdot\mathbf{a} = a^2$、$\mathbf{b}\cdot\mathbf{b} = b^2$。於是

$$(a+b)^2\,OQ^2 = (a\mathbf{a} + b\mathbf{b})\cdot(a\mathbf{a} + b\mathbf{b}) = a^2\cdot a^2 + 2ab\,c + b^2\cdot b^2 = a^4 + b^4 + 2abc,$$
$$(a+b)^2\,OP^2 = (b\mathbf{a} + a\mathbf{b})\cdot(b\mathbf{a} + a\mathbf{b}) = b^2a^2 + 2ab\,c + a^2b^2 = 2a^2b^2 + 2abc.$$

兩式相減，含 c 的項互相抵銷：

$$(a+b)^2\left(OQ^2-OP^2\right)=a^4-2a^2b^2+b^4=(a^2-b^2)^2=(a-b)^2(a+b)^2.$$

兩邊除以 $(a+b)^2\neq 0$，得

$$OQ^2-OP^2=(a-b)^2=(b-a)^2,$$

這正是所求。

驗算

取 $\mathbf{a}=(3,0)$ 和 $\mathbf{b}=(0,4)$，則 $a=3$、$b=4$，而 $\angle AOB$ 是直角。這時 $\lambda=\frac{4}{7}$，$\overrightarrow{OP}=\left(\frac{12}{7},\frac{12}{7}\right)$，它位於 $y=x$ 上，即該直角的角平分線。另外 $\overrightarrow{OQ}=\left(\frac{9}{7},\frac{16}{7}\right)$，所以 $OQ^2-OP^2=\frac{337}{49}-\frac{288}{49}=1=(4-3)^2$。若 $a=b$，三角形是等腰三角形，$P=Q$ 是 AB 的中點，兩者之差為 0，正如所料。以二維和三維空間中 20,000 個隨機三角形作數值測試，所有結果均吻合，誤差在 10^{-13} 以內。

常見錯誤

- 寫成 $\overrightarrow{OP}=(1-\lambda)\mathbf{a}+\lambda\mathbf{b}$：比例中較大的一份 $1-\lambda$ 是與 A 的距離，所以它應該配 $\mathbf{b}$。可用 $\lambda=1$ 檢驗。
- 用 $\mathbf{a}+\mathbf{b}$ 作為角平分線的方向。這只在 $a=b$ 時才成立；必須使用單位向量。
- 把 $|\mathbf{a}+\mathbf{b}|^2$ 展開成 a^2+b^2，漏掉了交叉項（cross term）$2\,\mathbf{a}\cdot\mathbf{b}$。本題中這一項會抵銷，但前提是兩個展開式中都有它。

STEP 之外

最後的計算，其實是長度在線性映射（linear map）下如何變換的一個特例：OP^2 和 OQ^2 是二次型（quadratic form）$\mathbf{x}\mapsto\mathbf{x}\cdot\mathbf{x}$ 的值，而交叉項則由 $\mathbf{a}$ 和 $\mathbf{b}$ 的格拉姆矩陣（Gram matrix）決定。單位向量相加得出角平分線這一事實亦可推廣：在任何內積空間（inner-product space）中，要找出兩個向量「正中間」的方向，做法就是先歸一化（normalise）再相加。Strang 的 *Introduction to Linear Algebra*（線性代數入門）正是本着這種精神闡述純量積、長度和角度。

詳細解答

STEP II 2008，第 9 題

力學 • 拋體運動

STEP II 2008，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：

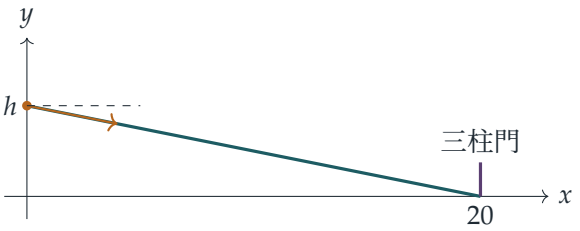
step.maths.org/questions/08-s2-q9

關鍵思路

水平方向的運動決定了以 α 表示的飛行時間（time of flight），把這個時間代入垂直方向的運動，便得到一條連繫 α 與 h 的方程。利用 $\sec^2\alpha = 1 + \tan^2\alpha$，可把它化為關於 $\tan\alpha$ 的二次方程（quadratic），這是處理「以哪個角度才能擊中目標？」一類問題的標準技巧。之後，憑着對板球的常識選出合適的根，再以小角近似（small-angle approximation）完成餘下的工作。

解答

我們把球看成一個質點（particle），只受重力作用而自由運動，不計空氣阻力，也不計旋轉。取拋射點正下方地面上的一點為原點，x 軸沿水平方向指向三柱門，y 軸垂直向上，於是球從 $(0, h)$ 出發，速率為 $40\,\mathrm{m\,s^{-1}}$，方向與水平線成仰角 α（α 為負數表示球是向下投出的）。



第 (i) 部分

水平方向沒有力的作用，所以 $x = 40t\cos\alpha$。球到達 $x = 20$ 的時刻為

$$t = \frac{20}{40\cos\alpha} = \frac{1}{2\cos\alpha}.$$

垂直方向的加速度為 $-10\,\mathrm{m\,s^{-2}}$，所以

$$y = h + 40t\sin\alpha - 5t^2.$$

此時球位於地面，即 $y = 0$，所以

$$0 = h + 40\sin\alpha\cdot\frac{1}{2\cos\alpha} - \frac{5}{4\cos^2\alpha} = h + 20\tan\alpha - \tfrac{5}{4}\sec^2\alpha.$$

現在利用 $\sec^2\alpha = 1 + \tan^2\alpha$，並記 $T = \tan\alpha$：

$$h + 20T - \tfrac{5}{4}(1 + T^2) = 0 \quad\Longleftrightarrow\quad 5T^2 - 80T + 5 - 4h = 0.$$

由二次公式（quadratic formula），

$$T = \frac{80 \pm \sqrt{6400 - 20(5-4h)}}{10} = \frac{80 \pm \sqrt{6300 + 80h}}{10},$$

所以兩個可能的值是

$$\tan\alpha = 8 \pm \sqrt{63 + \tfrac{4}{5}h}.$$

兩者皆為實數，因為 $h \geq 0$。

想一想 快速投球手實際上會用哪一個根？

由於 $\sqrt{63} \approx 7.9$，較大的根給出 $\tan\alpha \approx 16$，即 $\alpha \approx 86°$：這是一記沖天高吊球，會升高約 80 m，並要大約 8 s 才落到三柱門上。快速投球手不會這樣投球。較小的根，

$$\tan\alpha = 8 - \sqrt{63 + \tfrac{4}{5}h},$$

才是合適的選擇。對於合乎現實的出手高度（約兩米左右），$63 + \frac{4}{5}h$ 接近 64，所以 $\tan\alpha$ 接近 0。例如，只要 h 介乎 0 與 3 之間，便有 $|\tan\alpha| < 0.09$。因此 α 是一個小角，而 $\cos\alpha \approx 1$，球到達三柱門所需的時間為

$$t = \frac{1}{2\cos\alpha} \approx \tfrac{1}{2}\,\mathrm{s},$$

正如題目所求。（事實上 $t = \frac{1}{2}\sqrt{1 + \tan^2\alpha}$，只要 $|\tan\alpha| < 0.1$，它與 $\frac{1}{2}$ 的差距便不超過 0.5%。）

第 (ii) 部分

球向水平線以下投出，當且僅當 $\tan\alpha < 0$。較大的根恆為正數，所以我們採用較小的根：

$$8 - \sqrt{63 + \tfrac{4}{5}h} < 0 \iff \sqrt{63 + \tfrac{4}{5}h} > 8 \iff 63 + \tfrac{4}{5}h > 64 \iff h > \tfrac{5}{4},$$

其中平方這一步可逆，因為兩邊皆為正數。所以當 $h > 1.25$ 時，投球手是向水平線以下投出球的。當 $h = 1.25$ 恰好成立時，球沿水平方向投出。

第 (iii) 部分

當 $h = 2.5$ 時，由第 (ii) 部分可知 α 是一個小的負角。取 $\cos\alpha \approx 1$ 和 $\sin\alpha \approx \alpha$（以弧度（radians）計），便有 $\tan\alpha \approx \alpha$ 和 $\sec^2\alpha \approx 1$，所以第 (i) 部分的方程 $0 = h + 20\tan\alpha - \frac{5}{4}\sec^2\alpha$ 變成

$$0 \approx 2.5 + 20\alpha - 1.25, \qquad \text{所以} \qquad \alpha \approx -\frac{1.25}{20} = -0.0625 \text{ rad}.$$

換算成度數，

$$\alpha \approx -0.0625 \times \frac{180}{\pi} \approx -3.58^\circ,$$

所以 $\alpha \approx -3.6^\circ$。球出手時的方向在水平線以下約 3.6°。

另一個看法：利用二項式近似（binomial approximation），$\tan\alpha = 8 - \sqrt{65} = 8 - 8\sqrt{1 + \frac{1}{64}} \approx 8 - 8\left(1 + \frac{1}{128}\right) = -\frac{1}{16}$。這給出相同的答案。

驗算

- *精確值*。$\tan\alpha = 8 - \sqrt{65} = -0.06226$，所以 $\alpha = -3.563^\circ$，準確至兩位有效數字是 -3.6°，與近似結果吻合。
- *模擬*。從 $(0, 2.5)$ 出發，以 $40\,\text{m}\,\text{s}^{-1}$ 和 -3.563° 逐步作數值計算，球在 $0.501\,\text{s}$ 後到達 $x = 20$，此時高度為 $0.0000\,\text{m}$。
- *臨界情況*。當 $h = 1.25$ 時，較小的根是 $8 - \sqrt{64} = 0$：球沿水平方向投出，在半秒的飛行中下墜 $5(0.5)^2 = 1.25\,\text{m}$，剛好落在三柱門處。

常見錯誤

- 忘記 $\sec^2\alpha = 1 + \tan^2\alpha$，而試圖分別以 $\sin\alpha$ 和 $\cos\alpha$ 求解，結果變得十分繁複。
- 以弧度給出角度，或在度數制下使用小角近似：$\sin\theta \approx \theta$ 只在 θ 以弧度計時才成立。
- 在第 (iii) 部分漏掉負號。正如第 (ii) 部分所預示，這個角是在水平線以下的。

STEP 之外

兩個根的結構反映了一個普遍事實：以固定速率拋出的拋體，可沿兩條軌跡到達其「包絡線」（envelope，即安全拋物線（parabola of safety））以內的任何一點。在包絡線上，兩條軌跡合而為一，而包絡線以外的點則完全無法到達。真實的板球還會受到阻力（drag）以及由旋轉產生的馬格努斯效應（Magnus effect）影響，這會引出非線性微分方程（nonlinear differential equations），通常要用數值方法求解。Kleppner 與 Kolenkow 所著的 *An Introduction to Mechanics*（力學入門）以大學一年級的程度討論拋體運動和有阻力的運動。

詳細解答

STEP II 2008，第 10 題

力學 • 與牆壁的斜碰撞

STEP II 2008，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/08-s2-q10

關鍵思路

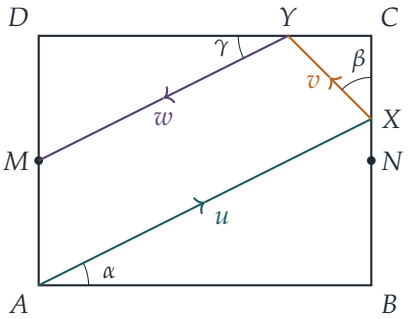
在光滑（smooth）的桌邊上，衝量（impulse）垂直於桌邊，所以沿桌邊方向的速度分量（component）不變，只有垂直撞向桌邊的分量會乘以 e。把速度分解成平行於 AB 和平行於 AD 的兩個分量，一切便一目了然：每次反彈只會改變兩個分量的其中一個。兩次反彈之後，兩個分量都各乘了 e，最終的方向和動能（kinetic energy）便同時確定下來。至於第 (ii) 部分的幾何，只需把各段水平距離加起來。

解答

我們把球看作質點（particle），把桌邊看作光滑的牆。取 A 為原點，x 軸沿 AB，y 軸沿 AD，於是

$$B = (a, 0), \quad C = (a, 2b), \quad D = (0, 2b), \quad M = (0, b), \quad N = (a, b).$$

為簡便起見，記 $t = \tan\alpha$。



第 (i) 部分

最初的速度是 $(u\cos\alpha,\ u\sin\alpha)$。

*在 X 的碰撞。*邊 BC 平行於 y 軸。由於它是光滑的，球所受的衝量沿 x 方向，所以 y 分量 $u\sin\alpha$ 不變。根據牛頓恢復定律（Newton's law of restitution），x 分量會反向並乘以 e。碰撞後的速度是

$$(-eu\cos\alpha,\ u\sin\alpha).$$

角 $\beta = \angle CXY$ 是從 XC 量起的，而這個方向指向 y 軸的正方向。所以

$$\tan\beta = \frac{eu\cos\alpha}{u\sin\alpha} = \frac{e}{\tan\alpha}, \qquad \text{即} \qquad \tan\alpha\tan\beta = e,$$

這正是所要證明的。

*在 Y 的碰撞。*邊 DC 平行於 x 軸，所以這次是 x 分量不變，而 y 分量反向並乘以 e。碰撞後的速度是

$$(-eu\cos\alpha,\ -eu\sin\alpha) = -e\,(u\cos\alpha,\ u\sin\alpha).$$

這是初速度的負數倍，所以路徑的最後一段平行於 AX，而且方向相反。角 $\gamma = \angle DYM$ 是從 YD 量起的，而這個方向指向 x 軸的負方向；方向 $(-\cos\alpha, -\sin\alpha)$ 與它的夾角為 α。因此

$$\gamma = \alpha.$$

第 (ii) 部分

我們沿着路徑追蹤 x 坐標。

*由 A 到 X。*球橫越寬度 a 的同時上升了 $a\tan\alpha$，所以 $X = (a,\ at)$。

*由 X 到 Y。*球再上升 $2b - at$，而在這一段，它每上升一個單位，就在 x 方向移動 $\tan\beta$ 個單位（因為 β 是從豎直的邊量起的）。利用第 (i) 部分的 $\tan\beta = e/t$，球向左移動的距離是 $(2b - at)\,e/t$，所以

$$x_Y = a - \frac{e(2b - at)}{t}.$$

*由 Y 到 M。*球由高度 $2b$ 下降到高度 b，即下降了 b，路徑在水平線之下成角 $\gamma = \alpha$。所以走過的水平距離是 $b/\tan\alpha$，而這段距離必須剛好把球帶到 $x = 0$：

$$x_Y = \frac{b}{t}.$$

想一想 為甚麼只追蹤 x 坐標就足夠？

每一段直線路徑都從上一段的終點開始，而它的方向已由第 (i) 部分得知。只要知道球在某一段上升或下降了多少，角度就決定了它橫向移動多遠。各段端點的高度都是固定

的（先是 o，然後在 DC 處是 $2b$，再在 M 處是 b），所以餘下的唯一條件，就是各段橫向距離加起來要吻合。

令 x_Y 的兩個表達式相等，並乘以 t：

$$\begin{aligned} b &= at - e(2b - at) \\ &= at(1+e) - 2be, \end{aligned}$$

所以 $at(1+e) = b(1+2e)$，由此得

$$\tan\alpha = \frac{(1+2e)b}{(1+e)a},$$

這正是所要證明的。

這一擊總是可行的。對任何 $e > 0$，這個 $\tan\alpha$ 的值都是正數，所以它確定了 $(0, \frac{\pi}{2})$ 內的一個角 α。我們還要確認各個反彈點確實落在應該落在的邊上。

- X 的高度是 $at = \dfrac{1+2e}{1+e}\,b$。由於當 $e > 0$ 時 $1+e < 1+2e < 2+2e$，這個高度嚴格介乎 b 與 $2b$ 之間。所以 X 確實在 BC 上，位於 C 之下，並且在球袋 N 之上，球不會在途中掉進 N。
- Y 位於 $x_Y = b/t = \dfrac{1+e}{1+2e}\,a$，當 $e > 0$ 時這個值嚴格介乎 0 與 a 之間。所以 Y 確實在 DC 上，位於 D 與 C 之間。

最後一段則由 Y 沿直線走到 M，全程都在球桌之內。因此，無論 e 取甚麼值，路徑的每一段都是可行的。

第 (iii) 部分

由第 (i) 部分可知，末速度是初速度的 $-e$ 倍，所以 $w = eu$。（這也可以直接看出：每個分量都恰好乘了一次 e。）動能由 $\frac{1}{2}mu^2$ 減至 $\frac{1}{2}me^2u^2$，其中 m 是球的質量。因此損失的比例是

$$\frac{\frac{1}{2}mu^2 - \frac{1}{2}me^2u^2}{\frac{1}{2}mu^2} = 1 - e^2.$$

這個結果與 a、b 和各個角度都無關。（若 $e > 1$，這個「損失」是負數，即桌邊反而給球增添能量。這並不符合物理實際，但代數運算依然成立。）

驗算

- 取 $a = 4$、$b = 1.5$ 和 $e = \frac{1}{2}$（即圖中的情況），公式給出 $\tan\alpha = \frac{1}{2}$。於是 $X = (4, 2)$，$Y = (3, 3)$，而由 Y 出發、斜率（gradient）為 $\frac{1}{2}$ 的直線在水平距離 3 之內下降 1.5，剛好落在 $M = (0, 1.5)$。損失的動能比例是 $1 - \frac{1}{4} = \frac{3}{4}$。

- 當 $e = 1$（完全彈性，perfectly elastic）時，$\tan\alpha = 3b/(2a)$。把球桌先對 BC 反射、再對 DC 反射而「展開」：路徑變成由 A 到 M 的像的一條直線，橫向走了 $2a$，向上走了 $3b$。這給出相同的斜率 $3b/(2a)$。沒有能量損失，與 $1 - e^2 = 0$ 吻合。
- 我們對 a、b、e 和 u 隨機取了 20,000 組值來模擬反彈，結果證實了 $\tan\alpha\tan\beta = e$、$\gamma = \alpha$、球到達 M，以及比例 $1 - e^2$，誤差全部在 10^{-13} 之內。

常見錯誤

- 在碰撞時把 e 乘到整個速率上。其實只有垂直於桌邊的分量會改變，沿光滑桌邊的分量絲毫不變。
- 從 BC 的法線（normal）而不是從 XC 量度 β，結果把 $\tan\beta$ 變成了它的倒數。
- 忘記驗證 X 在 N 之上，以及 Y 在 D 與 C 之間。這項驗證正是「這一擊可行」的內容。

STEP 之外

當 $e = 1$ 時，把球桌反射來「展開」桌球路徑，可以把路徑變成一條直線。這是數學桌球（mathematical billiards）的起點，它與動力系統（dynamical systems）相關，而對於有理球桌，更與曲面上的幾何相關。當 $e < 1$ 時，反射不再像鏡面反射，但把速度分解成分量後仍然容易處理。關於衝量與恢復的力學，可參閱 Kleppner 與 Kolenkow 的 *An Introduction to Mechanics*（力學導論）。想做更多這類題目，可以試試 Siklos 的 *Advanced Problems in Mathematics*（高等數學難題集）。

詳細解答

STEP II 2008，第 11 題

力學 • 移動楔上的質點

STEP II 2008，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：

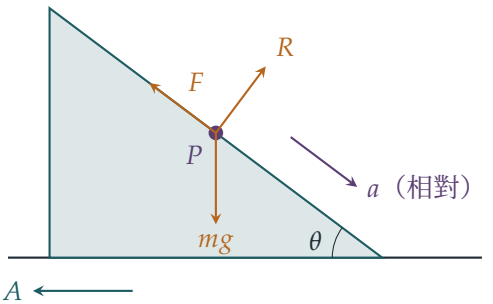
step.maths.org/questions/08-s2-q11

關鍵思路

系統以外沒有任何東西沿水平方向施力（地面是光滑的），所以質點（particle）與楔塊（wedge）的水平動量（momentum）一直保持為零；楔塊的加速度幾乎不費工夫就得到了。其餘一切，都來自把質點的眞實加速度寫成「它相對於楔塊的加速度，加上楔塊的加速度」，然後沿斜面方向和垂直於斜面方向分解（resolve）質點所受的力。靜止的觀察者所見的路徑，是一條沿真實加速度方向的直線，因為一切都由靜止開始，而且各加速度都是常數。

解答

我們把 P 視為質點，把楔塊視為剛體（rigid body）。地面光滑；斜面粗糙，摩擦系數（coefficient of friction）為 μ。把鉛垂面放在左邊，於是斜面向右下方傾斜。當 P 向右下方滑動時，楔塊被推向左方。設楔塊的加速度為向左的 A，並設 P 相對於楔塊沿斜面向下的加速度為 a。設 R 為法向反作用力（normal reaction），F 為作用於 P 的摩擦力（friction）；當 P 向下滑動時，摩擦力沿斜面向上。



取 x 軸向右、y 軸向上。沿斜面向下的單位向量（unit vector）是 $(\cos\theta, -\sin\theta)$，所以 P 的真實加速度是

$$\big(a\cos\theta - A,\ -a\sin\theta\big).$$

第 (i) 部分：楔塊的加速度

作用於「P 加楔塊」這個系統的外力，只有重力和光滑地面的法向反作用力，兩者都是鉛垂的。所以總水平動量是常數，而由於一切都由靜止開始，這個常數是零。因此 $m\times$（加速度）的水平分量之和為零：

$$m(a\cos\theta - A) + km(-A) = 0 \quad\Longrightarrow\quad A = \frac{a\cos\theta}{k+1},$$

這正是所要求的結果。（P 與楔塊之間的力 R 和 F 是內力，根據牛頓第三定律（Newton's third law）互相抵消。）

第 (i) 部分：45° 路徑

? **想一想** 為甚麼觀察到的路徑必定是一條沿真實加速度方向的直線？

下文的運動方程（equations of motion）確定了 a 和 A 都是常數，所以 P 的真實加速度是一個常向量。由靜止開始，它的速度總是該向量的 t 倍，位移則是該向量的 $\frac{1}{2}t^2$ 倍，所以它沿着該方向的一條固定直線運動。

P 的加速度的水平分量是

$$a\cos\theta - \frac{a\cos\theta}{k+1} = \frac{k}{k+1}\,a\cos\theta,$$

而鉛垂分量的大小是 $a\sin\theta$。路徑是一條 45° 的直線，即表示這兩者相等：

$$a\sin\theta = \frac{k}{k+1}\,a\cos\theta \quad\Longrightarrow\quad \tan\theta = \frac{k}{k+1},$$

其中兩邊除以了 $a\cos\theta \neq 0$（因為 P 確實在滑動，$a > 0$）。這正是所要求的結果。

第 (i) 部分：$k = 3$ 的情況

現在 $\tan\theta = \frac{3}{4}$，所以 $\sin\theta = \frac{3}{5}$，$\cos\theta = \frac{4}{5}$，而 $A = \frac{a\cdot 4/5}{4} = \frac{a}{5}$。

我們對 P 分解各力。它的真實加速度沿斜面向下的分量是 $a - A\cos\theta$，沿向外法線方向的分量是 $-A\sin\theta$（楔塊向左的加速度有這兩個分量；相對加速度則完全沿着斜面）。

垂直於斜面方向：

$$R - mg\cos\theta = -mA\sin\theta \quad\Longrightarrow\quad R = m\Big(\tfrac{4}{5}g - \tfrac{3}{25}a\Big).$$

留意 R 比 $mg\cos\theta$ 小：楔塊正在離 P 而去。

沿斜面向下方向，由於 P 正在滑動，所以 $F = \mu R$：

$$mg\sin\theta - \mu R = m(a - A\cos\theta) = m\left(a - \tfrac{4}{25}a\right) = \tfrac{21}{25}ma.$$

代入 R 並乘以 $25/m$：

$$15g - 20\mu g + 3\mu a = 21a \quad\Longrightarrow\quad a = \frac{5g(3-4\mu)}{3(7-\mu)}.$$

這個值恰好在 $\mu < \frac{3}{4} = \tan\theta$ 時為正，而 P 要滑動就必須如此；這正好帶我們進入第 (ii) 部分。

第 (ii) 部分

我們斷言：若 $\tan\theta \le \mu$，則 P 留在原處，楔塊也不會移動。

靜止是可能的。若兩個物體都不動，則對 P 而言，我們需要 $R = mg\cos\theta$，以及沿斜面向上的摩擦力 $F = mg\sin\theta$。只要 $F \le \mu R$，即 $\sin\theta \le \mu\cos\theta$，亦即 $\tan\theta \le \mu$，粗糙的接觸面就能提供這個力，而這正是我們的條件。楔塊所受的則是方向相反的 R 和 F；它們的水平分量分別是向左的 $R\sin\theta = mg\sin\theta\cos\theta$ 和向右的 $F\cos\theta = mg\sin\theta\cos\theta$，兩者互相抵消。所以楔塊也處於平衡（equilibrium），而地面只提供一個鉛垂的力。

想一想 即使靜止是可能的，P 會不會仍然滑動？

滑動是不可能的。假設 P 確實向下滑動，且 $F = \mu R$，其中 k 和 θ 為一般的值。設 $s = \sin\theta - \mu\cos\theta$。楔塊的水平方程是 $kmA = R\sin\theta - \mu R\cos\theta = Rs$。把它與法向方程 $R = m(g\cos\theta - A\sin\theta)$ 結合，得

$$R = \frac{kmg\cos\theta}{k + s\sin\theta}, \qquad a = \frac{(k+1)\,gs}{k + s\sin\theta}.$$

要保持接觸，我們需要 $R > 0$，這迫使 $k + s\sin\theta > 0$；於是 a 與 s 同號。但 $\tan\theta \le \mu$ 表示 $s \le 0$，所以 $a \le 0$，這與向下滑動矛盾。也沒有任何東西能使 P 沿斜面向上滑動，因為只有重力才能引起相對運動，而重力的分量是沿斜面向下的。所以當 $\tan\theta \le \mu$ 時，P 在楔塊上保持靜止，楔塊也在地面上保持靜止。（當 $\tan\theta = \mu$ 時，摩擦力達到極限值（limiting friction），但 P 仍然不動。）

驗算

- 取 $\mu = 0$ 和 $k = 3$，我們得到 $a = \frac{5g}{7}$ 和 $A = \frac{g}{7}$。於是 P 的真實加速度的水平分量是 $\frac{4}{5}\cdot\frac{5g}{7} - \frac{g}{7} = \frac{3g}{7}$，鉛垂分量是 $\frac{3}{5}\cdot\frac{5g}{7} = \frac{3g}{7}$，方向確實是 45°。

- 當 $k \to \infty$（楔塊無法移動）時，$A \to 0$，而一般公式給出 $a \to g(\sin\theta - \mu\cos\theta)$，即大家熟悉的固定斜面結果。
- 數值上，取 $g = 9.8$、$k = 3$ 和 $\mu = 0.2$，直接解三條運動方程得 $a \approx 5.284\ \mathrm{m\,s^{-2}}$，與 $\frac{5g(3-0.8)}{3(6.8)}$ 吻合。把運動隨時間逐步推算，水平動量一直保持為零，所得路徑的傾角恰好是 45°。

常見錯誤

- 取 $R = mg\cos\theta$。這只在楔塊固定時才成立；這裏楔塊加速離開，所以 R 較小。
- 分解時只用相對加速度 a，忘記了楔塊的加速度。牛頓第二定律需要的是真實加速度。
- 在第 (ii) 部分，即使 P 沒有滑動，仍假設摩擦力是 μR。摩擦力最多是 μR，實際取值視乎需要而定。

STEP 之外

第 (i) 部分最巧妙的一步是一條守恆定律（conservation law）：由於外力沒有水平分量，動量的其中一個分量守恆。在大學一年級的力學課程中，這會發展成一個把對稱性（symmetry）與守恆定律聯繫起來的一般原理：在拉格朗日力學（Lagrangian mechanics）中，若拉格朗日量（Lagrangian）不依賴某個坐標（這裏是整個系統的水平位置），就會自動得出一個守恆的動量。光滑楔塊與物塊正是那裏標準的入門例子。Kleppner 與 Kolenkow 的 *An Introduction to Mechanics*（力學入門教科書）詳細講解了正是這一類的問題。

詳細解答

STEP II 2008，第 12 題

概率與統計 • 獨立事件

> **STEP II 2008，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/08-s2-q12

關鍵思路

一個計數的期望值（expectation），等於構成它的各個事件的概率之和，所以第一部分只是一條關於 p 的二次方程。至於第二部分，法官們在心中擲硬幣時根本沒有看被告一眼，所以國王的宣判與真相獨立（independent）。把宣判正確的概率用 t 表示，再圍繞 $\frac{1}{2}$ 重新整理，答案一行便出來。

解答

求 p

設 A、B、C 分別為驢子、河狸和半人馬裁定有罪的事件，並設 I_A、I_B、I_C 為它們的指示變量（indicator variable）（事件發生時等於 1，否則等於 0）。那麼 $X = I_A + I_B + I_C$，而由於 $\mathrm{E}(I_A) = \mathrm{P}(A)$，其餘類推，由期望值的線性性質（linearity of expectation）可得

$$\mathrm{E}(X) = p + \frac{p}{3} + p^2.$$

（線性性質並不需要獨立性；我們稍後才會用到獨立性。）令上式等於 $\frac{4}{3}$，再乘以 3：

$$3p^2 + 4p - 4 = 0 \quad \Longrightarrow \quad (3p - 2)(p + 2) = 0.$$

概率不可能是負數，所以 $p + 2 \neq 0$，因此

$$p = \frac{2}{3}.$$

於是三位法官說「有罪」的概率分別為

$$\mathrm{P}(A) = \frac{2}{3}, \qquad \mathrm{P}(B) = \frac{2}{9}, \qquad \mathrm{P}(C) = \frac{4}{9},$$

全部都在 $[0, 1]$ 之內，正如所應有的。

國王說「有罪」的概率

設 K 為最少兩位法官裁定有罪的事件，即 K 是國王宣判被告有罪的事件。把上面三個概率記作 a、b、c。恰好兩個有罪裁決有三種情況，三個全是有罪則有一種情況；利用法官之間的獨立性，

$$\begin{aligned}\mathrm{P}(K) &= ab(1-c) + ac(1-b) + bc(1-a) + abc \\ &= ab + ac + bc - 2abc.\end{aligned}$$

這裏 $ab = \frac{4}{27}$，$ac = \frac{8}{27}$，$bc = \frac{8}{81}$，$abc = \frac{16}{243}$，所以

$$\mathrm{P}(K) = \frac{36}{243} + \frac{72}{243} + \frac{24}{243} - \frac{32}{243} = \frac{100}{243}.$$

為簡便起見，把它記作 g。

求 t

設 D 為被告確實有罪的事件，所以 $\mathrm{P}(D) = t$。當 D 與 K 同時發生，或兩者皆不發生時，國王的宣判便是正確的。

> ? **想一想** 為甚麼我們可以把 $\mathrm{P}(D)$ 乘以 $\mathrm{P}(K)$？

法官們的決定是「完全隨機」的：他們的裁決並不取決於證據，所以 K 與 D 獨立。因此

$$\mathrm{P}(\text{正確}) = \mathrm{P}(D)\,\mathrm{P}(K) + \mathrm{P}(D')\,\mathrm{P}(K') = tg + (1-t)(1-g).$$

現在減去 $\frac{1}{2}$ 並因式分解（factorise）：

$$\begin{aligned}tg + (1-t)(1-g) - \tfrac{1}{2} &= 2tg - t - g + \tfrac{1}{2} \\ &= (2t-1)\big(g - \tfrac{1}{2}\big).\end{aligned}$$

所以國王宣判正確的概率為 $\frac{1}{2}$，當且僅當 $(2t-1)(g-\frac{1}{2}) = 0$。由於 $g = \frac{100}{243} \neq \frac{1}{2}$，第二個因式不等於零，故必有

$$t = \frac{1}{2}.$$

不用因式分解也可以得到同樣的結果：由 $\frac{100}{243}t + \frac{143}{243}(1-t) = \frac{1}{2}$ 得 $143 - 43t = \frac{243}{2}$，所以 $43t = \frac{43}{2}$，即 $t = \frac{1}{2}$。

留意 p 的值在這裏幾乎無關重要：不論法官們的概率是多少，只要 $\mathrm{P}(K) \neq \frac{1}{2}$，一個無視證據的法庭有一半機會判對，當且僅當有罪與無辜的被告同樣常見。

驗算

► 取 $p = \frac{2}{3}$：$\frac{2}{3} + \frac{2}{9} + \frac{4}{9} = \frac{12}{9} = \frac{4}{3}$，符合要求。

- 用分數精確地列舉全部八種裁決組合，得 $\mathrm{P}(K) = \frac{100}{243} \approx 0.412$ 及 $\mathrm{E}(X) = \frac{4}{3}$。
- 取 $t = \frac{1}{2}$ 進行 200 000 次模擬試驗，國王宣判正確的個案所佔比例為 0.5007。
- 檢查公式是否合理：若 $t = 1$（人人有罪），國王判對的概率為 g；若 $t = 0$，則為 $1 - g$；這個概率是 t 的線性函數，所以它在中點處經過 $\frac{1}{2}$。

常見錯誤

- 把「最少兩個」當作 $ab + ac + bc$。這樣會把三個全中的情況數了三次；正確的表達式要減去 $2abc$（或者直接把四個互斥的情況相加）。
- 忘記捨去 $p = -2$，或者沒有檢查 p^2 和 $p/3$ 是否有效的概率。
- 把「國王宣判正確」只當作 $\mathrm{P}(K)$，忘記了那些獲正確判定無罪的無辜被告。

STEP 之外

第一部分是一個強大技巧的小例子：把一個計數寫成指示變量之和，於是它的期望值就只是概率之和，即使各事件並不獨立亦然。大學本科的概率論經常使用這個技巧，例如用來求隨機排列（random permutation）的不動點（fixed point）的期望數目。第二部分講的是一個輸出與真相獨立的分類器（classifier），這正是思考檢測中的靈敏度（sensitivity）、特異度（specificity）和貝葉斯定理（Bayes' theorem）的起點。Blitzstein 與 Hwang 的 *Introduction to Probability*（概率論入門）大量運用指示變量，是很好的下一步讀物。

詳細解答

STEP II 2008，第 13 題

概率與統計 • 不放回抽取

STEP II 2008，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/08-s2-q13

關鍵思路

不必追蹤每一枚籌碼：只追蹤黑色的就夠了。每枚黑色籌碼的去向由兩次抽取決定，而每次抽取時，某一枚指定籌碼被抽中的機會就是（抽出的數目）/（袋中的數目）。按第一次抽取的結果分情況討論，兩部分都會化成簡短的和，而且化簡的程度遠超你的預期。

解答

在整個解答中，「隨機抽出」是指袋中每一組 k 枚籌碼被抽出的可能性均等。我們反覆用到的事實是：若從一個裝有 m 枚籌碼的袋中隨機抽出 k 枚，則根據對稱性（symmetry），某一枚指定的籌碼在其中的概率為 k/m（這 m 枚籌碼各自被抽中的可能性均等，而平均有 k 枚被抽中）。

第一次轉移之後，袋 P 有 $n-k$ 枚籌碼，袋 Q 有 $n+k$ 枚。第二次轉移時，從這 $n+k$ 枚中抽出 k 枚，移到 P。

第 (i) 部分

設 B 為那枚黑色籌碼，它起初在 P 中。按 B 在第一次轉移時有沒有移動來分情況。

- B 沒有從 P 中被抽出：概率為 $\dfrac{n-k}{n}$。這時 B 自始至終留在 P 中，因為第二次轉移只會從 Q 取出籌碼。
- B 從 P 中被抽出：概率為 $\dfrac{k}{n}$。這時 B 是 Q 中 $n+k$ 枚籌碼之一，而它回到 P 當且僅當它在被抽回的 k 枚之中，其概率為 $\dfrac{k}{n+k}$。

這兩種情況互斥（disjoint），並且涵蓋了所有可能，所以

$$\begin{aligned}\mathrm{P}(B\text{ 最終在}P) &= \frac{n-k}{n} + \frac{k}{n}\cdot\frac{k}{n+k} = \frac{(n-k)(n+k)+k^2}{n(n+k)}\\ &= \frac{n^2}{n(n+k)} = \frac{n}{n+k}.\end{aligned}$$

這個概率隨 k 嚴格遞減，所以在 $0 \le k \le n$ 的範圍內，它只在 $k = 0$ 時取得最大值，此時其值為 1。（這並不意外：如果甚麼都不移動，黑色籌碼必定留在原處。）

第 (ii) 部分

現在設 B_1 為起初在 P 中的黑色籌碼，B_2 為起初在 Q 中的那一枚。設 S 為兩者最終在同一個袋中的事件。同樣按 B_1 在第一次轉移時有沒有移動來分情況。

情況 1：B_1 留在 P 中（概率為 $\frac{n-k}{n}$）。這時無論之後發生甚麼，B_1 最終都在 P 中，所以 S 發生當且僅當 B_2 在從 $n+k$ 枚籌碼中抽出的 k 枚之中，這些籌碼都在 Q 內：

$$\mathrm{P}(S \mid \text{情況 1}) = \frac{k}{n+k}.$$

情況 2：B_1 移到 Q（概率為 $\frac{k}{n}$）。這時兩枚黑色籌碼都在 Q 的 $n+k$ 枚籌碼之中，而 S 發生的條件是第二次抽取把它們兩枚都取走，或者兩枚都不取走。

> **想一想** 從 $n+k$ 枚籌碼中隨機抽出一個 k 元子集，它同時包含兩枚指定籌碼，或者兩枚都不包含的概率是多少？

點算子集的數目。共有 $\binom{n+k}{k}$ 種可能性均等的選法。同時包含兩枚黑色籌碼的選法，是從餘下的 $n+k-2$ 枚中選出另外 $k-2$ 枚；兩枚都不包含的選法，則是從 $n+k-2$ 枚白色籌碼中選出全部 k 枚。相除可得

$$\begin{aligned}\mathrm{P}(\text{兩枚都有}) &= \binom{n+k-2}{k-2}\Big/\binom{n+k}{k} = \frac{k(k-1)}{(n+k)(n+k-1)},\\ \mathrm{P}(\text{兩枚都沒有}) &= \binom{n+k-2}{k}\Big/\binom{n+k}{k} = \frac{n(n-1)}{(n+k)(n+k-1)}.\end{aligned}$$

（這些公式在 $k = 0, 1$ 時同樣成立，此時 $\mathrm{P}(\text{兩枚都有}) = 0$。）所以

$$\mathrm{P}(S \mid \text{情況 2}) = \frac{k(k-1)+n(n-1)}{(n+k)(n+k-1)}.$$

把各情況合併。根據全概率定律（law of total probability），

$$\begin{aligned}\mathrm{P}(S) &= \frac{n-k}{n}\cdot\frac{k}{n+k} + \frac{k}{n}\cdot\frac{k(k-1)+n(n-1)}{(n+k)(n+k-1)}\\ &= \frac{k\big[(n-k)(n+k-1)+k^2-k+n^2-n\big]}{n(n+k)(n+k-1)}.\end{aligned}$$

在方括號內，$(n-k)(n+k-1) = n^2 - k^2 - n + k$，所以方括號等於 $2n^2 - 2n = 2n(n-1)$。因此

$$\boxed{\mathrm{P}(S) = \frac{2k(n-1)}{(n+k)(n+k-1)}}$$

求最大值。當 $k = 0$ 時概率為 0，所以取 $1 \le k \le n$，並求其倒數的最小值。撇開正常數 $2(n-1)$ 不計，

$$\frac{(n+k)(n+k-1)}{k} = k + (2n-1) + \frac{n(n-1)}{k},$$

所以我們需要在整數範圍內求 $g(k) = k + \dfrac{n(n-1)}{k}$ 的最小值。比較相鄰兩項：

$$g(k+1) - g(k) = 1 - \frac{n(n-1)}{k(k+1)}.$$

當 $k(k+1) < n(n-1)$，即 $k < n-1$ 時，這個差為負數；而當 $k = n-1$ 時，它恰好為零，因為此時 $k(k+1) = (n-1)n$。所以 g 由 $k = 1$ 到 $k = n-1$ 嚴格遞減，並且 $g(n) = g(n-1)$。因此概率取得最大值之處為

$$k = n-1 \quad \text{及} \quad k = n,$$

由於 $n \ge 2$，兩者都在容許的範圍內。當 $k = n-1$ 時，最大值為

$$\frac{2(n-1)^2}{(2n-1)(2n-2)} = \frac{n-1}{2n-1},$$

而 $k = n$ 給出相同的值，$\dfrac{2n(n-1)}{2n(2n-1)} = \dfrac{n-1}{2n-1}$。（把 g 看成連續函數時，其最小值點為 $\sqrt{n(n-1)}$，嚴格位於 $n-1$ 與 n 之間；這兩個整數處的值是完全相等，而非近似相等。）

驗算

- 第 (i) 部分取 $k = n$：得 $\frac{1}{2}$。這時全部 $2n$ 枚籌碼都經過 Q，其中 n 枚返回，所以黑色籌碼最終落在兩個袋中的可能性均等。吻合。
- 第 (ii) 部分取 $k = 0$ 得 0，這是正確的，因為兩枚黑色籌碼永遠不會相遇。取 $n = 2$、$k = 1$ 得 $\frac{2}{3 \cdot 2} = \frac{1}{3}$，你可以親手驗證。
- 我們對 $2 \le n \le 7$ 及每一個 k，以分數精確地列舉了每一對可能的抽取結果；兩條公式在所有情況下都吻合，而 (ii) 中取得最大值之處總是 $\{n-1, n\}$，例如最大值為 $\frac{3}{7} \approx 0.4286$，對應 $n = 4$。

常見錯誤

- 在第二次抽取時用了 k/n 而不是 $k/(n+k)$：袋 Q 已增至 $n+k$ 枚籌碼。

- 在第 (ii) 部分的情況 2 中，忘記了「兩枚都沒有被抽出」的可能：如果兩枚黑色籌碼都留在 Q 中，它們仍然在一起。
- 用微積分求最大值，再把 $\sqrt{n(n-1)}$ 捨入為單一個整數。離散的比較顯示，有兩處同樣取得最大值。

STEP 之外

兩部分背後的工具是超幾何分佈（hypergeometric distribution）：在不放回的隨機抽樣中，樣本內帶記號物件的數目。先以第一階段的結果為條件（conditioning），再把它用於第二階段，這就是全概率定律；同樣的兩階段推理，也是容器之間混合與擴散模型的基礎，例如埃倫費斯特甕模型（Ehrenfest urn）。Blitzstein and Hwang, *Introduction to Probability*（概率論入門教材）正是以這種方式處理超幾何分佈與條件概率。

詳細解答

STEP III 2008，第 1 題

純數學 • 非線性聯立方程

STEP III 2008，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/08-s3-q1

關鍵思路

設 $s_k = ax^k + by^k$，則題目的方程給出了 s_0, s_1, s_2, s_3。由於 x 和 y 是 $t^2 - pt + q = 0$ 的兩個根，其中 $p = x + y$、$q = xy$，這些冪和(power sums)滿足一個遞推關係(recurrence)：每個 s_k 等於 p 乘以前一項，再減去 q 乘以再前一項。這正是提示中「乘以 $x + y$」所得到的結果。於是四條非線性方程化為關於 p 和 q 的兩條線性方程。之後解一條二次方程便得 x 和 y，再由首兩條方程求出 a 和 b。

解答

設 $s_k = ax^k + by^k$，其中 $k = 0, 1, 2, 3$，則 $s_0 = 1$、$s_1 = \frac{1}{3}$、$s_2 = \frac{1}{5}$、$s_3 = \frac{1}{7}$。令 $p = x + y$ 及 $q = xy$。

冪和的遞推關係

對任意 a, b, x, y 及任意 $k \geq 1$，

$$(x + y)(ax^k + by^k) = ax^{k+1} + by^{k+1} + xy\,(ax^{k-1} + by^{k-1}),$$

因為交叉項是 $axyx^{k-1}$ 和 $bxy\,y^{k-1}$。用我們的記號，即

$$s_{k+1} = p\,s_k - q\,s_{k-1}.$$

這是一條恆等式（identity），所以對方程組的每一個解都成立。

求 $x+y$ 和 xy

取 $k=1$ 和 $k=2$，並代入已知數值：

$$\begin{aligned}\tfrac{1}{5} &= \tfrac{1}{3}p - q,\\ \tfrac{1}{7} &= \tfrac{1}{5}p - \tfrac{1}{3}q.\end{aligned}$$

由第一式得 $q = \frac{1}{3}p - \frac{1}{5}$。代入第二式，

$$\tfrac{1}{7} = \tfrac{1}{5}p - \tfrac{1}{9}p + \tfrac{1}{15} \quad\Longrightarrow\quad \tfrac{4}{45}p = \tfrac{1}{7} - \tfrac{1}{15} = \tfrac{8}{105} \quad\Longrightarrow\quad p = \tfrac{6}{7},$$

於是 $q = \frac{2}{7} - \frac{1}{5} = \frac{3}{35}$。這個線性方程組的行列式（determinant）不為零，所以這些數值是唯一確定的：每一個解都滿足 $x+y=\frac{6}{7}$ 及 $xy=\frac{3}{35}$。

求 x 和 y

因此 x 和 y 是以下方程的兩個根：

$$t^2 - \tfrac{6}{7}t + \tfrac{3}{35} = 0, \qquad \text{即} \qquad 35t^2 - 30t + 3 = 0,$$

由此得

$$t = \frac{30 \pm \sqrt{900-420}}{70} = \frac{30 \pm 4\sqrt{30}}{70} = \frac{15 \pm 2\sqrt{30}}{35}.$$

判別式（discriminant）為正，所以 $x \neq y$，而 $\{x, y\}$ 就是這兩個數，次序可以任選其一。

> ? **想一想** 知道了 x 和 y 之後，能否肯定 a 和 b 也隨之確定？

可以。x 和 y 已知之後，首兩條方程 $a+b=1$ 和 $ax+by=\frac{1}{3}$ 是關於 a 和 b 的線性方程，其行列式為 $y-x$。由於 $x \neq y$，這個行列式不為零，所以恰好有一對 (a,b)。

求 a 和 b

取 $x = \frac{15+2\sqrt{30}}{35}$ 及 $y = \frac{15-2\sqrt{30}}{35}$，則 $x-y = \frac{4\sqrt{30}}{35}$。從 $a+b=1$ 和 $ax+by=\frac{1}{3}$ 消去 b，得 $a(x-y) = \frac{1}{3} - y$。現在

$$\tfrac{1}{3} - y = \frac{35 - 3(15-2\sqrt{30})}{105} = \frac{6\sqrt{30}-10}{105},$$

所以

$$a = \frac{6\sqrt{30}-10}{105} \cdot \frac{35}{4\sqrt{30}} = \frac{6\sqrt{30}-10}{12\sqrt{30}} = \frac{1}{2} - \frac{5}{6\sqrt{30}} = \frac{1}{2} - \frac{\sqrt{30}}{36},$$

而 $b = 1 - a = \frac{1}{2} + \frac{\sqrt{30}}{36}$。

這些確實是解

我們至此只證明了任何解都必定是這個樣子，所以仍須驗證它確實成立。按構造方式，$s_0 = 1$ 及 $s_1 = \frac{1}{3}$。由於 p 和 q 滿足上面兩條線性方程，遞推關係給出 $s_2 = ps_1 - qs_0 = \frac{1}{5}$，繼而 $s_3 = ps_2 - qs_1 = \frac{1}{7}$。所以四條方程全部成立。

把 x 和 y 的角色互換，a 和 b 也隨之互換。因此恰好有兩個解：

$$(a, b, x, y) = \left(\frac{18 - \sqrt{30}}{36}, \frac{18 + \sqrt{30}}{36}, \frac{15 + 2\sqrt{30}}{35}, \frac{15 - 2\sqrt{30}}{35} \right)$$

以及把 $a \leftrightarrow b$ 和 $x \leftrightarrow y$ 同時互換所得的解。在兩個解中，x, y 之中較大者都與較小的係數配對。

驗算

數值上，$x \approx 0.741556$、$y \approx 0.115587$、$a \approx 0.347855$ 及 $b \approx 0.652145$。於是 $ax + by \approx 0.333333$、$ax^2 + by^2 \approx 0.200000$ 及 $ax^3 + by^3 \approx 0.142857$，與 $\frac{1}{3}$、$\frac{1}{5}$ 和 $\frac{1}{7}$ 吻合至機器精度。再作一個簡單的合理性檢查：x 和 y 都在 $(0, 1)$ 之內，而 a 和 b 都是正數。這與右邊各數是 t^{2k} 在 $[0, 1]$ 上的平均值一致（見下文）。

常見錯誤

- 解出了 x 和 y，卻忘記配對是有關係的：a 與 x 相配，所以不能只互換 x 和 y 而不同時互換 a 和 b。
- 找到必要條件（necessary conditions）便停下來。題目說的是「求所有解」，所以既要證明每一個解都是這個形式，也要證明這些數值確實滿足全部四條方程。

STEP 之外

$1, \frac{1}{3}, \frac{1}{5}, \frac{1}{7}$ 這些數就是 $\frac{1}{2}\int_{-1}^{1} t^{2k}\,\mathrm{d}t$。所以這個方程組其實是在找一條法則 $\frac{1}{2}\int_{-1}^{1} f \approx \frac{a}{2}\big(f(\sqrt{x}) + f(-\sqrt{x})\big) + \frac{b}{2}\big(f(\sqrt{y}) + f(-\sqrt{y})\big)$，使它對次數不超過 7 的多項式都精確成立。這就是四點高斯—勒讓德求積法（*Gauss–Legendre quadrature*）：它的節點（nodes）是 $\pm\sqrt{x}$ 和 $\pm\sqrt{y}$，而 $a \approx 0.3479$ 和 $b \approx 0.6521$ 就是它著名的權（weights）。像我們這裏所做的那樣，由冪和找出遞推關係，也是信號處理中 Prony 方法（Prony's method）背後的想法。想進一步了解，可參閱 Süli 與 Mayers 的 *An Introduction to Numerical Analysis*（數值分析入門）。

詳細解答

STEP III 2008，第 2 題

純數學 • 冪和

> **STEP III 2008，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/08-s3-q2

關鍵思路

$(r+1)^k - r^k$ 的和可以裂項相消（telescoping），最後只剩下一項。但我們也可以用二項式定理（binomial theorem）把每個括號展開，這樣同一個和就變成較低次冪和 $S_j(n)$ 的組合。令兩者相等，便得到一條遞推關係（recurrence），可由前面各個冪和逐一建立 S_{k-1}；而要看出第 (ii) 部分的多項式結構，所需的也正是這條遞推關係。

解答

第 (i) 部分

首先是裂項相消。把和式逐項寫出，

$$\sum_{r=0}^{n}\left[(r+1)^k - r^k\right] = (1^k - 0^k) + (2^k - 1^k) + \cdots + \left((n+1)^k - n^k\right) = (n+1)^k,$$

因為中間各項全部互相抵消，而且 $0^k = 0$（因為 $k \geq 1$）。

其次是二項式定理。對每個 r，

$$(r+1)^k - r^k = \sum_{j=0}^{k-1}\binom{k}{j}r^j = 1 + \binom{k}{1}r + \binom{k}{2}r^2 + \cdots + \binom{k}{k-1}r^{k-1},$$

因為 r^k 項互相抵消。對 $r = 0, 1, \ldots, n$ 求和，常數 1 貢獻 $n+1$（r 共有 $n+1$ 個值），而每個 r^j（$j \geq 1$）則貢獻 $S_j(n)$。所以

$$(n+1)^k = (n+1) + \binom{k}{1}S_1(n) + \binom{k}{2}S_2(n) + \cdots + \binom{k}{k-2}S_{k-2}(n) + \binom{k}{k-1}S_{k-1}(n).$$

最後一個係數是 $\binom{k}{k-1} = k$。至於其餘各項，我們利用對稱性 $\binom{k}{j} = \binom{k}{k-j}$，可知 $S_{k-j}(n)$ 的係數是 $\binom{k}{k-j} = \binom{k}{j}$；特別地，$S_1(n)$ 的係數是 $\binom{k}{1} = \binom{k}{k-1}$。把 $kS_{k-1}(n)$ 留在一邊，其餘各項全部移到另一邊，得

$$kS_{k-1}(n) = (n+1)^k - (n+1) - \binom{k}{2}S_{k-2}(n) - \binom{k}{3}S_{k-3}(n) - \cdots - \binom{k}{k-1}S_1(n),$$

這正是所求的 $(*)$。（這裏需要 $k \geq 2$，S_{k-1} 才是我們所定義的和之一。）

想一想 為甚麼我們把 $j = 0$ 的一項分開處理，而不把它寫成 $S_0(n)$？

S_k 只對正數 k 有定義，而且在 $r = 0$ 時「0^0」的意義並不明確。把常數抽出來，直接數出它的總和是 $n+1$，就能同時避開這兩個問題，也解釋了 $(*)$ 中那個單獨出現的 $-(n+1)$。

$S_3(n)$。在 $(*)$ 中取 $k = 4$，這時 $\binom{4}{2} = 6$ 及 $\binom{4}{3} = 4$：

$$\begin{aligned} 4S_3(n) &= (n+1)^4 - (n+1) - 6S_2(n) - 4S_1(n) \\ &= (n+1)^4 - (n+1) - n(n+1)(2n+1) - 2n(n+1). \end{aligned}$$

每一項都有因式 $n+1$，所以

$$\begin{aligned} 4S_3(n) &= (n+1)\left[(n+1)^3 - 1 - n(2n+1) - 2n\right] \\ &= (n+1)\left[n^3 + 3n^2 + 3n - 2n^2 - n - 2n\right] \\ &= (n+1)(n^3 + n^2) = n^2(n+1)^2. \end{aligned}$$

因此

$$S_3(n) = \tfrac{1}{4}n^2(n+1)^2 = \big(S_1(n)\big)^2.$$

$S_4(n)$。取 $k = 5$，這時 $\binom{5}{2} = \binom{5}{3} = 10$ 及 $\binom{5}{4} = 5$：

$$5S_4(n) = (n+1)^5 - (n+1) - 10S_3(n) - 10S_2(n) - 5S_1(n).$$

我們再次抽出因式 $n+1$，並代入 S_3、S_2 和 S_1 的表達式：

$$\begin{aligned} 5S_4(n) &= (n+1)\left[(n+1)^4 - 1 - \tfrac{5}{2}n^2(n+1) - \tfrac{5}{3}n(2n+1) - \tfrac{5}{2}n\right] \\ &= (n+1)\Big[n^4 + 4n^3 + 6n^2 + 4n - \tfrac{5}{2}n^3 - \tfrac{5}{2}n^2 \\ &\qquad - \tfrac{10}{3}n^2 - \tfrac{5}{3}n - \tfrac{5}{2}n\Big] \\ &= (n+1)\left[n^4 + \tfrac{3}{2}n^3 + \tfrac{1}{6}n^2 - \tfrac{1}{6}n\right] \\ &= \tfrac{1}{6}n(n+1)\left(6n^3 + 9n^2 + n - 1\right). \end{aligned}$$

由於 S_4 應該像 S_2 一樣含有因式 $2n+1$，我們把 $n = -\frac{1}{2}$ 代入這個三次式試一試：$-\frac{3}{4} + \frac{9}{4} - \frac{1}{2} - 1 = 0$。所以 $2n+1$ 是因式，而 $6n^3 + 9n^2 + n - 1 = (2n+1)(3n^2 + 3n - 1)$。於是

$$S_4(n) = \tfrac{1}{30}\,n(n+1)(2n+1)(3n^2 + 3n - 1).$$

第 (ii) 部分

我們對 k 使用強歸納法（strong induction）。命題在 $k=1$ 時成立，因為 $S_1(n)=\frac{1}{2}n^2+\frac{1}{2}n$ 的次數是 2（在 $k=2$ 時亦成立）。

假設對某個 $k\geq 2$，$S_1(n),\ldots,S_{k-1}(n)$ 每一個都是 n 的多項式（polynomial），其中 S_j 的次數是 $j+1$。在 $(*)$ 中以 $k+1$ 代替 k：

$$(k+1)\,S_k(n)=(n+1)^{k+1}-(n+1)-\binom{k+1}{2}S_{k-1}(n)-\cdots-\binom{k+1}{k}S_1(n).$$

在右邊：

- $(n+1)^{k+1}$ 是次數為 $k+1$ 的多項式，首項係數（leading coefficient）是 1；
- $n+1$ 的次數是 1；
- 根據歸納假設，每個 $S_j(n)$（$1\leq j\leq k-1$）都是次數為 $j+1\leq k$ 的多項式。

所以右邊是一個多項式，它的 n^{k+1} 項只來自 $(n+1)^{k+1}$，不可能被抵消。除以 $k+1\neq 0$ 後，可知 $S_k(n)$ 是次數恰好為 $k+1$ 的多項式（首項係數是 $\frac{1}{k+1}$）。歸納至此完成。

> **想一想** 多項式 $P(n)$ 的「常數項」和「各係數之和」分別對應甚麼？

它們就是 $P(0)$ 和 $P(1)$。設 $S_k(n)=P(n)=a_{k+1}n^{k+1}+\cdots+a_1n+a_0$；對每個整數 $n\geq 0$，它都與和式 $\sum_{r=0}^{n}r^k$ 相等，所以我們可以代入 $n=0$ 和 $n=1$ 求值。

- 常數項：$a_0=P(0)=S_k(0)=0^k=0$，因為 $k\geq 1$。
- 各係數之和：$a_{k+1}+\cdots+a_1+a_0=P(1)=S_k(1)=0^k+1^k=1$。

兩個結果均已證明。

驗算

- 較小的數值：$S_3(3)=1+8+27=36=\frac{1}{4}\cdot 9\cdot 16$，以及 $S_4(3)=1+16+81=98=\frac{1}{30}\cdot 3\cdot 4\cdot 7\cdot 35$。
- 用我們的答案檢驗第 (ii) 部分：$S_4(n)=\frac{1}{5}n^5+\frac{1}{2}n^4+\frac{1}{3}n^3-\frac{1}{30}n$ 的常數項是 0，各係數之和是 $\frac{6+15+10-1}{30}=1$；它的首項係數是 $\frac{1}{5}=\frac{1}{k+1}$，與歸納法所預測的一致。
- 我們以直接求和的方法，對 $2\leq k\leq 9$ 及 $0\leq n\leq 30$ 準確地驗證了 $(*)$，並對 $0\leq n<200$ 驗證了兩個閉合形式（closed form）。

常見錯誤

- 忘記了 $\sum 1$ 中 $r = 0$ 的一項：它的貢獻是 $n + 1$，而不是 n。
- 沒有先抽出公因式 $n + 1$，而是把所有項全部展開，令 S_4 的代數運算繁複得多。
- 在第 (ii) 部分只說次數「至多是 $k + 1$」。題目要求的是恰好 $k + 1$，所以要指出 n^{k+1} 項不可能被抵消。

STEP 之外

這些多項式的係數由伯努利數（Bernoulli numbers）決定，而福爾哈伯公式（Faulhaber's formula）把 $S_k(n)$ 寫成 n 的各次冪的組合，其係數是二項式係數（binomial coefficients）乘以伯努利數。$S_k(n)$ 是次數為 $k + 1$ 的多項式，首項是 $\frac{n^{k+1}}{k+1}$，這一事實正是 $\int_0^x t^k \, \mathrm{d}t = \frac{x^{k+1}}{k+1}$ 的離散類比，而有限差分學（calculus of finite differences）把這個想法嚴格化。想做更多同類題目，可試試 Siklos 的 *Advanced Problems in Mathematics*（《高等數學難題》）。

詳細解答

STEP III 2008，第 3 題

純數學 • 橢圓的切線與焦點

STEP III 2008，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：

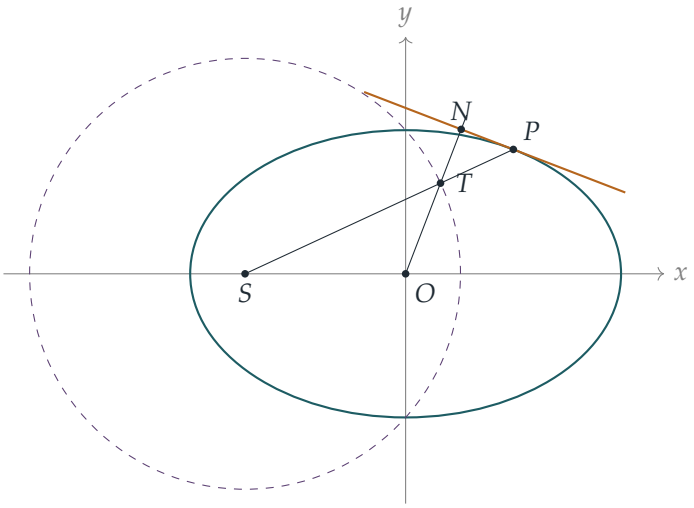
step.maths.org/questions/08-s3-q3

關鍵思路

整道題只涉及兩條直線，所以要做的就是把它們的方程乾淨利落地寫出來，然後求交點。直線 ON 垂直於切線（tangent），所以它的方向就是切線的法向量（normal vector），可以直接從切線方程讀出。由 S 出發把 SP 參數化（parametrise），求交點便只需一行計算；而求得的參數原來正是處理圓時所需要的：它告訴我們點 T 位於 SP 上幾分之幾的位置。

解答

全題中 $0 < e < 1$：我們有 $e^2 = 1 - b^2/a^2$，由於 $a > b > 0$，此值嚴格介乎 0 與 1 之間。我們亦假設 $\sin\theta \neq 0$。若 $\sin\theta = 0$，則 $P = (\pm a, 0)$，切線為 $x = \pm a$，$N = P$，而 ON 和 SP 都是 x 軸，所以 T 沒有定義。



（圖中取 $a = 3$，$b = 2$，$\theta = 60°$。橢圓（ellipse）為藍綠色，在 P 的切線為琥珀色，虛線圓的圓心為 S，半徑為 a。）

T 的 y 坐標

對 $x^2/a^2 + y^2/b^2 = 1$ 作隱函數微分（implicit differentiation），或直接引用標準結果，可知在 (x_1, y_1) 的切線為 $xx_1/a^2 + yy_1/b^2 = 1$。在 P 處，這就是

$$\frac{x\cos\theta}{a} + \frac{y\sin\theta}{b} = 1.$$

直線 $px + qy = 1$ 的法向量為 (p, q)，所以切線的法向量為 $\left(\frac{\cos\theta}{a}, \frac{\sin\theta}{b}\right)$，或者乘以 ab 後，寫成 $(b\cos\theta,\, a\sin\theta)$。直線 ON 沿這個方向通過 O（它垂直於切線），所以線上的點為 $\mu(b\cos\theta, a\sin\theta)$，其方程為

$$a\sin\theta\, x = b\cos\theta\, y. \tag{1}$$

現在由 S 出發把 SP 參數化：線上的點為

$$\big(-ea + \lambda a(\cos\theta + e),\ \lambda b\sin\theta\big),$$

其中 $\lambda = 0$ 對應 S，$\lambda = 1$ 對應 P。代入 (1)，並約去因子 $\sin\theta \neq 0$，得

$$\begin{aligned} a\big(-ea + \lambda a(\cos\theta + e)\big) &= \lambda b^2\cos\theta \\ \lambda\big(a^2\cos\theta + a^2e - b^2\cos\theta\big) &= ea^2. \end{aligned}$$

焦點(focus)條件正是在這裏派上用場：$a^2 - b^2 = a^2e^2$，所以括號內的式子是 $a^2e^2\cos\theta + a^2e = a^2e(1 + e\cos\theta)$。除以 $a^2e \neq 0$，得

$$\lambda = \frac{1}{1 + e\cos\theta},$$

這一步是合理的，因為 $1+e\cos\theta \geq 1-e > 0$。因此 T 的 y 坐標為

$$y_T = \lambda b \sin\theta = \frac{b\sin\theta}{1+e\cos\theta},$$

正如所求。（留待稍後參考：由 (1) 可得 $x_T = \dfrac{b\cos\theta}{a\sin\theta}\, y_T = \dfrac{a(1-e^2)\cos\theta}{1+e\cos\theta}$。）

T 位於以 S 為圓心、半徑為 a 的圓上

? **想一想** 我們已經知道線段 ST 佔 SP 的幾分之幾。還需要知道甚麼？

由於 $T = S + \lambda(P-S)$ 且 $\lambda > 0$，我們有 $ST = \lambda\, SP$。所以只需求出長度 SP，即 P 的焦半徑（focal distance）：

$$\begin{aligned} SP^2 &= (a\cos\theta + ea)^2 + b^2\sin^2\theta \\ &= a^2\left[\cos^2\theta + 2e\cos\theta + e^2 + (1-e^2)\sin^2\theta\right] \\ &= a^2\left[(\cos^2\theta+\sin^2\theta) + 2e\cos\theta + e^2(1-\sin^2\theta)\right] \\ &= a^2\left[1 + 2e\cos\theta + e^2\cos^2\theta\right] = a^2(1+e\cos\theta)^2, \end{aligned}$$

其中第二行用了 $b^2 = a^2(1-e^2)$。由於 $a > 0$ 且 $1+e\cos\theta > 0$，由此得 $SP = a(1+e\cos\theta)$。於是

$$ST = \lambda\, SP = \frac{a(1+e\cos\theta)}{1+e\cos\theta} = a,$$

所以 T 位於以 S 為圓心、半徑為 a 的圓上。

（如果你寧願直接用坐標計算：$x_T + ea = \dfrac{a(\cos\theta+e)}{1+e\cos\theta}$，而 $(x_T+ea)^2 + y_T^2$ 可化簡為 a^2，所用的代數運算與上面計算 SP^2 時完全相同。）

驗算

- 當 $\theta = \frac{\pi}{2}$ 時，$P = (0, b)$，切線為 $y = b$，所以 ON 是 y 軸，而 $T = (0, y_T)$，由公式得 $y_T = b$。於是 $ST^2 = e^2a^2 + b^2 = a^2$，與預期相符。
- 取 $a = 3$，$b = 2$，$\theta = 60°$（即圖中的情況），以數值方法求兩條直線的交點，得 $T \approx (0.4857,\ 1.2618)$；公式則給出 $y_T = \sqrt{3}/(1+\sqrt{5}/6) \approx 1.2618$，而 $ST \approx 3.0000$。
- 在趨向圓的極限 $e \to 0$ 之下，$S \to O$ 且 $T \to P$，而後者確實位於以 O 為圓心、半徑為 a 的圓上。

常見錯誤

- 把 ON 當成平行於切線，而不是垂直於切線：ON 的方向是法向量 $(b\cos\theta, a\sin\theta)$，而不是 $(-a\sin\theta, b\cos\theta)$。
- 忘記把 $a^2 - b^2$ 換成 a^2e^2，結果得到一個無法化簡的 λ 表達式。
- 除以 $\sin\theta$ 而不加說明；$\sin\theta = 0$ 的情況確實不成立。

STEP 之外

結果 $SP = a(1 + e\cos\theta) = a + ex_P$ 其實是焦點—準線性質（focus–directrix property）的另一種面貌：到焦點的距離等於到準線 $x = -a/e$ 的距離的 e 倍。以焦點為極點，用極坐標（polar coordinates）把圓錐曲線（conic）寫成 $r = \ell/(1 + e\cos\phi)$，正是大學一年級力學課描述開普勒（Kepler）軌道的方式，而同一個分母 $1 + e\cos\theta$ 亦在這裏出現了。關於圓錐曲線的幾何，以及更多 STEP 風格的題目，可參閱 Siklos 的 *Advanced Problems in Mathematics*；關於軌道，可參閱 Kleppner 與 Kolenkow 的 *An Introduction to Mechanics*。

詳細解答

STEP III 2008，第 4 題

純數學 • 反雙曲函數不等式

STEP III 2008，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/08-s3-q4

關鍵思路

設 $x = \cosh y$。半角公式（half-angle formula）把 $\tanh(y/2)$ 化為 x 的代數函數，於是一個關於 tanh 的顯淺不等式（inequality），就變成一個關於 arcosh 的不顯淺不等式。之後，題目其實是在運轉一部小機器：把不等式由 1 積分至 x，整理一下，便得到一個更強的不等式。把這部機器運轉兩次，先得出 (ii)，再得出 (iii)。

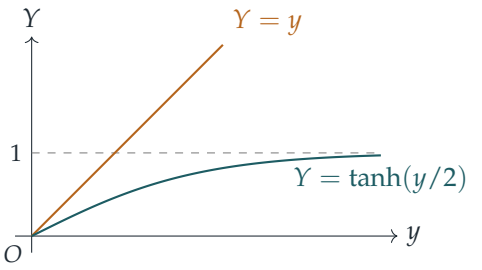
解答

以下一律記 $A(x) = \operatorname{arcosh} x$ 及 $S(x) = \sqrt{x^2-1}$，其中 $x \geq 1$。我們會用到以下兩個導數（derivative）：

$$\frac{\mathrm{d}}{\mathrm{d}x}\operatorname{arcosh} x = \frac{1}{\sqrt{x^2-1}}, \qquad \frac{\mathrm{d}}{\mathrm{d}x}\sqrt{x^2-1} = \frac{x}{\sqrt{x^2-1}} \qquad (x > 1).$$

第 (i) 部分

曲線 $Y = \tanh(y/2)$ 通過原點，在原點的斜率(gradient)為 $\frac{1}{2}$，在 $y > 0$ 時是凹(concave)的，而且永遠達不到 1；而直線 $Y = y$ 的斜率為 1：



對所有 $y > 0$，直線都在曲線之上。要嚴格說明這一點，設 $g(t) = t - \tanh t$。則 $g(0) = 0$，而且

$$g'(t) = 1 - \operatorname{sech}^2 t = \tanh^2 t > 0 \quad \text{對於} t > 0,$$

所以 g 嚴格遞增（strictly increasing），而當 $t > 0$ 時 $\tanh t < t$。取 $t = y/2$，得

$$\tanh(y/2) < \tfrac{1}{2}y < y \qquad (y > 0).$$

現在設 $x > 1$，並令 $y = \operatorname{arcosh} x$，於是 $y > 0$、$\cosh y = x$ 及 $\sinh y = +\sqrt{x^2 - 1}$（因為 $y > 0$，所以取正值）。利用 $\sinh y = 2\sinh \frac{y}{2}\cosh \frac{y}{2}$ 及 $1 + \cosh y = 2\cosh^2 \frac{y}{2}$，

$$\tanh \frac{y}{2} = \frac{\sinh y}{1 + \cosh y} = \frac{\sqrt{x^2 - 1}}{x + 1} = \frac{x^2 - 1}{(x + 1)\sqrt{x^2 - 1}} = \frac{x - 1}{\sqrt{x^2 - 1}}.$$

因此 $y > \tanh(y/2)$ 變成

$$\operatorname{arcosh} x > \frac{x - 1}{\sqrt{x^2 - 1}} \qquad (x > 1),$$

這就是 $(*)$。

第 (ii) 部分

固定 $x > 1$，以 t 為變數，把 $(*)$ 在 $[1, x]$ 上積分。

> ? **想一想** $(*)$ 的右邊，分母含有 $\sqrt{t^2 - 1}$。由 $t = 1$ 開始積分是否穩妥？積分之後，不等式是否仍然是嚴格的？

兩個問題的答案都是肯定的。當 $t > 1$ 時，$\dfrac{t - 1}{\sqrt{t^2 - 1}} = \sqrt{\dfrac{t - 1}{t + 1}}$，而當 $t \to 1$ 時它趨向 0，所以兩邊在 $[1, x]$ 上都是連續（continuous）的。兩者之差是連續的，並且在 $(1, x]$ 上嚴格為正，所以它在 $[1, x]$ 上的積分嚴格為正。

左邊用分部積分法（integration by parts），取 $u = \operatorname{arcosh} t$：

$$\int_1^x \operatorname{arcosh} t \, \mathrm{d}t = \Big[t \operatorname{arcosh} t\Big]_1^x - \int_1^x \frac{t}{\sqrt{t^2 - 1}} \, \mathrm{d}t = x\,A(x) - S(x),$$

因為 $\operatorname{arcosh} 1 = 0$ 及 $\sqrt{1^2 - 1} = 0$。

右邊可拆成兩個標準積分：

$$\int_1^x \frac{t - 1}{\sqrt{t^2 - 1}} \, \mathrm{d}t = \int_1^x \frac{t}{\sqrt{t^2 - 1}} \, \mathrm{d}t - \int_1^x \frac{1}{\sqrt{t^2 - 1}} \, \mathrm{d}t = S(x) - A(x).$$

因此

$$x\,A(x) - S(x) > S(x) - A(x) \quad \Longrightarrow \quad (x + 1)A(x) > 2S(x).$$

兩邊除以 $x+1>0$，並如第 (i) 部分般化簡，

$$\operatorname{arcosh} x > \frac{2\sqrt{x^2-1}}{x+1} = 2\,\frac{x-1}{\sqrt{x^2-1}} \qquad (x>1),$$

正是所求。這個下界剛好是 $(*)$ 中下界的兩倍。

第 (iii) 部分

第 (ii) 部分說明，當 $t>1$ 時 $\operatorname{arcosh} t > 2\cdot\dfrac{t-1}{\sqrt{t^2-1}}$：右邊與 $(*)$ 的右邊相同，只是乘大了一倍。所以我們把同一部機器再運轉一次。在 $[1,x]$ 上積分（理據與之前相同），並利用已求得的兩個積分，

$$x\,A(x)-S(x) > 2\big(S(x)-A(x)\big) \quad\Longrightarrow\quad (x+2)A(x) > 3S(x).$$

由於 $x+2>0$，

$$\operatorname{arcosh} x > 3\,\frac{\sqrt{x^2-1}}{x+2} \qquad (x>1),$$

這就是所求的結果。

> **想一想** 為甚麼到此為止？可以再積分一次嗎？

這部機器要求右邊是 $\dfrac{t-1}{\sqrt{t^2-1}}$ 的倍數，因為我們知道它的積分。新的下界 $3\sqrt{t^2-1}/(t+2)$ 並非這個形式，所以這個技巧到此便用盡了。

作為獨立的驗證(即題目所謂「或用其他方法」的途徑)，設 $f(x)=A(x)-3S(x)/(x+2)$。則 $f(1)=0$，而根據商法則（quotient rule），

$$f'(x)=\frac{(x+2)^2-3x(x+2)+3(x^2-1)}{(x+2)^2\sqrt{x^2-1}}=\frac{(x-1)^2}{(x+2)^2\sqrt{x^2-1}}>0 \quad (x>1),$$

所以 f 嚴格遞增，而當 $x>1$ 時 $f(x)>0$，與之前的結果一致。

驗算

在 $x=2$：$\operatorname{arcosh}2=\ln(2+\sqrt{3})\approx 1.3170$，而三個下界分別是 $1/\sqrt{3}\approx 0.577$、$2/\sqrt{3}\approx 1.155$ 及 $3\sqrt{3}/4\approx 1.299$。每個下界都比前一個好，而且全部低於 1.3170。在 $x=1$ 附近，這些下界變得非常精確：在 $x=1.001$，$\operatorname{arcosh} x\approx 0.04471763$，(ii) 的下界只差約 7×10^{-6}，而 (iii) 的下界更只差約 10^{-9}，因為每個下界與 $\operatorname{arcosh} x$ 吻合至 $x-1$ 的更高階。我們亦以數值方法檢驗了那兩個積分（在 $x=1.5,\,2,\,5$ 用中點求積法（midpoint quadrature)），並在 10 000 個隨機選取的 $x>1$ 檢驗了全部三個不等式，均沒有發現任何出入。

常見錯誤

- 取 $\sinh(\operatorname{arcosh} x) = \pm\sqrt{x^2-1}$ 而沒有確定正負號。因為 $\operatorname{arcosh} x > 0$，所以應取正根。
- 由 0 而不是由 1 開始積分：arcosh 只在 $x \geq 1$ 時有定義，而正是下限 1 令邊界項（boundary terms）消失。
- 忘記說明為何積分後的不等式仍然是嚴格的。

STEP 之外

下界愈來愈好，是因為每積分一次，誤差在 $x = 1$ 處消失的階數便高一階；這是一系列愈來愈精確的有理型近似的開端，它們是帕德逼近（Padé approximants）的近親。全題貫徹使用的原理（若 $f(a) = 0$ 且 $f' > 0$，或其積分形式成立，則 $f > 0$）其實是中值定理（mean value theorem）的化身，也是大量分析學結果的基礎。Abbott 的 *Understanding Analysis* 對中值定理和積分有嚴謹的論述。

詳細解答

STEP III 2008，第 5 題

純數學 • 遞推關係

> **STEP III 2008，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/08-s3-q5

關鍵思路

$\mathrm{T}_n^2 - \mathrm{T}_{n-1}\mathrm{T}_{n+1}$ 這個量是遞推關係（recurrence）的一個不變量（invariant）：當 n 增加一時，前後兩項之差會互相抵銷，因為遞推關係容許我們把 $\mathrm{T}_{n+1} + \mathrm{T}_{n-1}$ 換成 $2x\mathrm{T}_n$。當不變量為零時，相鄰兩項之比全部相等，所以數列是等比數列（geometric sequence）。再把等比數列代入遞推關係，便可確定公比（common ratio）是某個二次方程（quadratic）的根。

解答

以下 x 始終固定，並把 $\mathrm{T}_n(x)$ 記作 T_n，把 $\mathrm{f}(x)$ 記作 f。遞推關係 $(*)$ 即是

$$\mathrm{T}_{n+1} + \mathrm{T}_{n-1} = 2x\,\mathrm{T}_n \qquad (n \geq 1).$$

不變量

對 $n \geq 1$，設 $D_n = \mathrm{T}_n^2 - \mathrm{T}_{n-1}\mathrm{T}_{n+1}$。我們用數學歸納法（induction）證明對所有 $n \geq 1$ 都有 $D_n = f$。

基本情況。$D_1 = \mathrm{T}_1^2 - \mathrm{T}_0\mathrm{T}_2$，按定義這就是 f。

歸納步驟。假設對某個 $k \geq 1$ 有 $D_k = f$。則

$$\begin{aligned} D_{k+1} - D_k &= \mathrm{T}_{k+1}^2 - \mathrm{T}_k\mathrm{T}_{k+2} - \mathrm{T}_k^2 + \mathrm{T}_{k-1}\mathrm{T}_{k+1} \\ &= \mathrm{T}_{k+1}\big(\mathrm{T}_{k+1} + \mathrm{T}_{k-1}\big) - \mathrm{T}_k\big(\mathrm{T}_{k+2} + \mathrm{T}_k\big) \\ &= \mathrm{T}_{k+1} \cdot 2x\mathrm{T}_k - \mathrm{T}_k \cdot 2x\mathrm{T}_{k+1} \\ &= 0, \end{aligned}$$

其中第三行對第一個括號用了 $(*)$ 取 $n = k$，對第二個括號則取 $n = k+1$（兩者皆可，因為 $k \geq 1$）。因此 $D_{k+1} = D_k = f$，由歸納法可知

$$\mathrm{T}_n^2 - \mathrm{T}_{n-1}\mathrm{T}_{n+1} = f \quad \text{對所有} n \geq 1,$$

這正是所要證明的。

$f \equiv 0$ 的情形

首先看看 $f = 0$ 對 $r = \mathrm{T}_1/\mathrm{T}_0$ 有甚麼含意。由 $(*)$ 取 $n = 1$，得 $\mathrm{T}_2 = 2x\mathrm{T}_1 - \mathrm{T}_0$，所以

$$f = \mathrm{T}_1^2 - \mathrm{T}_0(2x\mathrm{T}_1 - \mathrm{T}_0) = \mathrm{T}_0^2\big(r^2 - 2xr + 1\big),$$

這裏用了 $\mathrm{T}_1 = r\mathrm{T}_0$。於是，由於 $\mathrm{T}_0 \neq 0$，

$$f = 0 \iff r^2 = 2xr - 1. \tag{†}$$

> **想一想** 不變量告訴我們 $\mathrm{T}_n^2 = \mathrm{T}_{n-1}\mathrm{T}_{n+1}$，所以各個比值相等。為甚麼不直接相除？

要除以 T_{n-1}，就必須先確定它不是零，這需要額外的論證。較簡潔的做法是從「比值相等」這個想法猜出答案，然後利用 (†)，直接以遞推關係本身來證明。

斷言。對所有 $n \geq 0$，$\mathrm{T}_n = \mathrm{T}_0\, r^n$。

證明。我們使用強歸納法（strong induction），每次用上相鄰的兩個情況。斷言在 $n = 0$ 時顯然成立；在 $n = 1$ 時亦成立，因為按 r 的定義有 $\mathrm{T}_1 = r\mathrm{T}_0$。假設斷言對 $n = k-1$ 和 $n = k$ 成立，其中 $k \geq 1$。則由 $(*)$，

$$\begin{aligned}\mathrm{T}_{k+1} &= 2x\mathrm{T}_k - \mathrm{T}_{k-1} = 2x\mathrm{T}_0 r^k - \mathrm{T}_0 r^{k-1} \\ &= \mathrm{T}_0 r^{k-1}\big(2xr - 1\big) = \mathrm{T}_0 r^{k-1} \cdot r^2 = \mathrm{T}_0 r^{k+1},\end{aligned}$$

其中第四步用了 (†)。所以斷言對 $n = k+1$ 成立，由歸納法可知它對所有 $n \geq 0$ 成立：

$$\boxed{\mathrm{T}_n(x) = \mathrm{T}_0(x)\,\big(\mathrm{r}(x)\big)^n}$$

$\mathrm{r}(x)$ 的可能形式

由 (†)，r 滿足 $r^2 - 2xr + 1 = 0$，所以由二次公式（quadratic formula）得

$$\mathrm{r}(x) = x + \sqrt{x^2 - 1} \qquad \text{或} \qquad \mathrm{r}(x) = x - \sqrt{x^2 - 1}.$$

反過來，無論取哪一個都滿足 (†)，所以 $f \equiv 0$，而數列 $\mathrm{T}_0 r^n$ 確實符合 $(*)$。兩根之積為 1，所以它們互為倒數。當 $|x| < 1$ 時兩根是複數，即 $x \pm \mathrm{i}\sqrt{1 - x^2}$；設 $x = \cos\theta$，它們便成為 $\mathrm{e}^{\pm \mathrm{i}\theta}$，而 $\mathrm{T}_n = \mathrm{T}_0 \mathrm{e}^{\pm \mathrm{i}n\theta}$。

驗算

- 切比雪夫多項式（Chebyshev polynomials）$\mathrm{T}_0 = 1$、$\mathrm{T}_1 = x$、$\mathrm{T}_2 = 2x^2 - 1$ 給出 $f = x^2 - (2x^2 - 1) = 1 - x^2$，而事實上對每個 n 都有 $\cos^2 n\theta - \cos(n-1)\theta\cos(n+1)\theta = \sin^2\theta$。取 $x = \frac{3}{7}$，D_n 的首十四個值全部等於 $\frac{40}{49}$。
- 取 $r = 2$，則 $x = \frac{1}{2}(2 + \frac{1}{2}) = \frac{5}{4}$，並取 $\mathrm{T}_0 = 1$、$\mathrm{T}_1 = 2$。由遞推關係得 $\mathrm{T}_2 = \frac{5}{2}\cdot 2 - 1 = 4$ 及 $\mathrm{T}_3 = \frac{5}{2}\cdot 4 - 2 = 8$：正如所料，都是 2 的冪。此外，$x - \sqrt{x^2-1} = \frac{5}{4} - \frac{3}{4} = \frac{1}{2}$ 是另一個根。

常見錯誤

- 把 $(*)$ 用於 $n = 0$：遞推關係由 $n = 1$ 開始，所以基本情況必須是 D_1，而不是 D_0（後者根本沒有定義）。
- 把 $\mathrm{T}_n^2 = \mathrm{T}_{n-1}\mathrm{T}_{n+1}$ 除以 T_{n-1}，卻沒有檢查它是否非零。

STEP 之外

常係數線性遞推關係的解具有 $\alpha r_1^n + \beta r_2^n$ 的形式，其中 r_1, r_2 是其特徵方程（characteristic equation）的根，在這裏特徵方程是 $r^2 - 2xr + 1 = 0$。條件 $f = 0$ 所挑出的，恰好是只剩下一個等比項的情況；一般而言，f 是 $\alpha\beta$ 的倍數。不變量 D_n 可說是二階微分方程中朗斯基行列式（Wronskian）的離散版本，而對於 $y'' + q(x)y = 0$，朗斯基行列式是常數。Boyce 與 DiPrima 的 *Elementary Differential Equations* 詳細闡述了朗斯基行列式，而 Siklos 的 *Advanced Problems in Mathematics* 則有更多 STEP 風格的遞推關係題目。

詳細解答

STEP III 2008，第 6 題

純數學 • 一階微分方程

> **STEP III 2008，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/08-s3-q6

關鍵思路

每條方程都直接以 x 和斜率 p 表示 y，但並未解出 p，所以不能直接積分。對 x 求導可以消去 y（因為 $\mathrm{d}y/\mathrm{d}x = p$），餘下一條關於 x、p 和 $\mathrm{d}p/\mathrm{d}x$ 的關係式。把這條關係式上下倒轉，便得到一條以 x 為 p 的函數的線性一階微分方程（linear first-order differential equation），可用積分因子（integrating factor）求解。然後把 $x(p)$ 代回原方程，便可求得 y。

解答

第 (i) 部分

把 $y = p^2 + 2xp$ 對 x 求導，對 p^2 用鏈式法則（chain rule），對 $2xp$ 用積法則（product rule），並以 p 取代 $\mathrm{d}y/\mathrm{d}x$：

$$p = 2p\frac{\mathrm{d}p}{\mathrm{d}x} + 2p + 2x\frac{\mathrm{d}p}{\mathrm{d}x} \quad \Longrightarrow \quad -p = (2p + 2x)\frac{\mathrm{d}p}{\mathrm{d}x}.$$

> ? **想一想** 我們可以把 $\mathrm{d}p/\mathrm{d}x$ 倒轉而得出 $\mathrm{d}x/\mathrm{d}p$ 嗎？

我們需要 $\mathrm{d}p/\mathrm{d}x \neq 0$。上面的方程顯示，凡是 $p \neq 0$ 之處，$\mathrm{d}p/\mathrm{d}x$ 和 $2p + 2x$ 都不等於零。（反正我們要除以 p，所以整個過程都需要 $p \neq 0$。）因此可以寫成

$$\frac{\mathrm{d}x}{\mathrm{d}p} = -\frac{2p + 2x}{p} = -2 - \frac{2x}{p},$$

正是題目所求。

整理後，這是一條對 x 為線性的方程：

$$\frac{\mathrm{d}x}{\mathrm{d}p} + \frac{2}{p}x = -2.$$

積分因子是 $\mathrm{e}^{\int 2/p\,\mathrm{d}p} = p^2$，全式乘以它，得

$$\frac{\mathrm{d}}{\mathrm{d}p}(p^2x) = -2p^2 \quad\Longrightarrow\quad p^2x = -\tfrac{2}{3}p^3 + A \quad\Longrightarrow\quad x = -\tfrac{2}{3}p + Ap^{-2},$$

其中 A 為任意常數。這就是題目給出的結果。

現在利用當 $x = 2$ 時 $p = -3$：

$$2 = -\tfrac{2}{3}(-3) + \frac{A}{9} = 2 + \frac{A}{9},$$

所以 $A = 0$，而 $x = -\frac{2}{3}p$，即 $p = -\frac{3}{2}x$。代入原方程，

$$y = p^2 + 2xp = \tfrac{9}{4}x^2 + 2x\left(-\tfrac{3}{2}x\right) = \tfrac{9}{4}x^2 - 3x^2,$$

所以

$$y = -\tfrac{3}{4}x^2.$$

驗算一下：$\mathrm{d}y/\mathrm{d}x = -\frac{3}{2}x$，確實就是 p，而且在 $x = 2$ 時等於 -3。

第 (ii) 部分

這裏出現了 $\ln p$，所以 $p > 0$。把 $y = 2xp + p\ln p$ 對 x 求導，並利用 $\frac{\mathrm{d}}{\mathrm{d}p}(p\ln p) = \ln p + 1$：

$$p = 2p + 2x\frac{\mathrm{d}p}{\mathrm{d}x} + (\ln p + 1)\frac{\mathrm{d}p}{\mathrm{d}x} \quad\Longrightarrow\quad -p = (2x + \ln p + 1)\frac{\mathrm{d}p}{\mathrm{d}x}.$$

與之前一樣，$p \neq 0$ 令 $\mathrm{d}p/\mathrm{d}x \neq 0$，所以

$$\frac{\mathrm{d}x}{\mathrm{d}p} = -\frac{2x + \ln p + 1}{p} \quad\Longrightarrow\quad \frac{\mathrm{d}x}{\mathrm{d}p} + \frac{2}{p}x = -\frac{\ln p + 1}{p}.$$

積分因子同樣是 p^2：

$$\frac{\mathrm{d}}{\mathrm{d}p}(p^2x) = -p\ln p - p.$$

用分部積分法（integration by parts），取 $u = \ln p$，$\mathrm{d}v = p\,\mathrm{d}p$，

$$\int p\ln p\,\mathrm{d}p = \tfrac{1}{2}p^2\ln p - \int \tfrac{1}{2}p\,\mathrm{d}p = \tfrac{1}{2}p^2\ln p - \tfrac{1}{4}p^2,$$

所以

$$\begin{aligned} p^2x &= -\tfrac{1}{2}p^2\ln p + \tfrac{1}{4}p^2 - \tfrac{1}{2}p^2 + B \\ &= -\tfrac{1}{2}p^2\ln p - \tfrac{1}{4}p^2 + B, \end{aligned}$$

而

$$x = -\tfrac{1}{2}\ln p - \tfrac{1}{4} + Bp^{-2}.$$

利用當 $x = -\frac{1}{4}$ 時 $p = 1$：$-\frac{1}{4} = 0 - \frac{1}{4} + B$，所以 $B = 0$，而

$$x = -\tfrac{1}{2}\ln p - \tfrac{1}{4},$$

這正是題目要求我們證明的結果。

要求出 y，留意上式即 $\ln p = -2x - \frac{1}{2}$。把它代入原方程，比起把每一個 p 都代換掉要快：

$$y = 2xp + p\left(-2x - \tfrac{1}{2}\right) = -\tfrac{1}{2}p,$$

而 $p = \mathrm{e}^{-2x-1/2}$，所以

$$y = -\tfrac{1}{2}\mathrm{e}^{-2x-\frac{1}{2}}.$$

驗算：$\mathrm{d}y/\mathrm{d}x = \mathrm{e}^{-2x-\frac{1}{2}} = p$，理應如此；而在 $x = -\frac{1}{4}$ 時，其值為 $\mathrm{e}^0 = 1$。

驗算

在每一部分，把最後求得的 y 求導，確認導數確實就是你所用的 p：$y = -\frac{3}{4}x^2$ 給出 $p = -\frac{3}{2}x$，而 $p^2 + 2xp = \frac{9}{4}x^2 - 3x^2 = y$。至於第 (ii) 部分，在 $x = 0$ 時，$p = \mathrm{e}^{-1/2} \approx 0.6065$，而 $2xp + p\ln p = 0.6065 \times (-0.5) \approx -0.3033 = y(0)$。數值驗算亦確認了，對隨機選取的 A 和 B 值，一般的參數曲線 $(x(p), y(p))$ 的斜率 $\mathrm{d}y/\mathrm{d}x$ 都等於 p。

常見錯誤

- 忘記左邊的 $\mathrm{d}y/\mathrm{d}x$ 就是 p，因而看不出相消之後餘下 $-p$。
- 未用初始條件（initial condition）便把積分常數丟掉。這裏的常數碰巧等於零，但你要把它求出來之後才知道。
- 在第 (ii) 部分，把 $p\ln p$ 對 x 求導時只寫成 $\ln p$，遺漏了因子 $\mathrm{d}p/\mathrm{d}x$ 和額外的 $+1$。

STEP 之外

形如 $y = xf(p) + g(p)$ 的方程稱為達朗貝爾（d'Alembert）方程，又稱拉格朗日（Lagrange）方程；其特殊情況 $y = xp + g(p)$ 就是克萊羅（Clairaut）方程。把 p 當作自變量（independent variable）的技巧對這些方程全部適用。它們還可能有奇異解（singular solutions），出現在我們曾除以 $\mathrm{d}p/\mathrm{d}x$ 的地方，是通解（general solution）族的包絡線（envelope）。Boyce 與 DiPrima 所著的 *Elementary Differential Equations*（初等微分方程）涵蓋一階方程及其解法。

詳細解答

STEP III 2008，第 7 題

純數學 • 複數與幾何

> **STEP III 2008，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/08-s3-q7

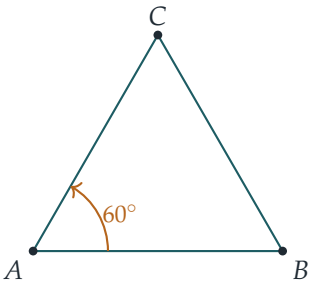
關鍵思路

在阿根圖（Argand diagram）上，乘以 $e^{i\pi/3}$ 相當於逆時針旋轉 60°。等邊三角形（equilateral triangle）正好就是把一條邊繞着一個頂點旋轉 60° 而成，所以開首的關係式把「等邊」這個條件化成了 a、b、c 的線性方程。之後兩部分都只是按部就班的整理功夫：把每個新的點用原有的點表示出來，再驗證我們想要的條件——若要證明是平行四邊形（parallelogram），就驗證 $p + r = q + s$；若要證明是等邊三角形，就驗證開首的關係式本身。

解答

開首的關係式

設 $\omega = e^{i\pi/3} = \frac{1}{2} + \frac{\sqrt{3}}{2}i$。由於 ABC 是等邊三角形，AC 與 AB 等長，並與它成 60° 角。因為三角形的頂點是按逆時針方向排列的，所以由 $\overrightarrow{AB}$ 轉到 $\overrightarrow{AC}$ 是逆時針方向。因此 $\overrightarrow{AC}$ 就是 $\overrightarrow{AB}$ 旋轉 $+60°$ 所得：



$$c-a=\omega(b-a)=\tfrac{1}{2}(b-a)+\tfrac{\sqrt{3}}{2}\mathrm{i}(b-a),$$
$$2c=2a+(b-a)+\mathrm{i}\sqrt{3}(b-a)=(a+b)+\mathrm{i}\sqrt{3}(b-a),$$

這就是所要求的結果。

若三角形的頂點按順時針方向排列，$\overrightarrow{AC}$ 就是 $\overrightarrow{AB}$ 旋轉 -60° 所得，所以 $c-a=\bar{\omega}(b-a)$，其中 $\bar{\omega}=\frac{1}{2}-\frac{\sqrt{3}}{2}\mathrm{i}$。同樣的代數運算給出

$$2c=(a+b)-\mathrm{i}\sqrt{3}(b-a).$$

反過來說，任何滿足其中一條關係式的 c 都會令 ABC 成為等邊三角形，因為每條關係式都等價於 $c-a=\omega^{\pm1}(b-a)$。所以要證明一個三角形是等邊三角形，只需驗證其中一條關係式便可。

第 (i) 部分

四個頂點由我們自行選擇，而每一個都可以放在所屬邊的任何一側。利用開首的結果，並讓這些頂點擔當 c 的角色，取

$$\begin{aligned}
2p&=(d+e)+\mathrm{i}\sqrt{3}(e-d) && (D,E,P\text{ 逆時針}),\\
2q&=(e+f)-\mathrm{i}\sqrt{3}(f-e) && (E,F,Q\text{ 順時針}),\\
2r&=(f+g)+\mathrm{i}\sqrt{3}(g-f) && (F,G,R\text{ 逆時針}),\\
2s&=(g+d)-\mathrm{i}\sqrt{3}(d-g) && (G,D,S\text{ 順時針}).
\end{aligned}$$

根據開首的結果，PDE、QEF、RFG、SGD 於是全都是等邊三角形。

> **想一想** 為甚麼要讓方向交替變換，而不是把每個三角形都放在所屬邊的同一側？

不妨試試看。$PQRS$ 是平行四邊形，當且僅當 $\overrightarrow{PQ}=\overrightarrow{SR}$，即 $q-p=r-s$，亦即 $p+r=q+s$。（這表示對角線 PR 與 QS 有相同的中點。）按上面的選擇，

$$2(p+r)=(d+e+f+g)+\mathrm{i}\sqrt{3}\,(e-d+g-f),$$
$$2(q+s)=(d+e+f+g)-\mathrm{i}\sqrt{3}\,(f-e+d-g)=(d+e+f+g)+\mathrm{i}\sqrt{3}\,(e-d+g-f).$$

兩者相等，所以 $p+r=q+s$，$PQRS$ 是平行四邊形。假如我們全部都用「+」號，虛部就會是 $\mathrm{i}\sqrt{3}(e-d+g-f)$ 和 $\mathrm{i}\sqrt{3}(f-e+d-g)$。這兩者只在 $d+f=e+g$ 時才相等，也就是只在 $DEFG$ 本身已經是平行四邊形時才相等。正是正負號交替，才令含 $\sqrt{3}$ 的項對任何四邊形都能吻合。

（對某些四邊形來說，這四點可能剛好共線，得出一個退化的平行四邊形。在這種情況下，條件 $\overrightarrow{PQ}=\overrightarrow{SR}$ 仍然成立。）

第 (ii) 部分

若 LMN 的頂點按順時針方向排列，可把整個圖形對實軸作反射（即取共軛複數（complex conjugate））。反射保持等邊三角形和形心（centroid）不變，也保持邊上的三角形是否向外作這一性質，所以只需處理 LMN 按逆時針方向排列的情況。

這時 N 位於有向邊 $L \to M$ 的左方，而 LM 上向外作的三角形 LMX 的頂點 X 則在右方，所以 L, M, X 是按順時針方向排列的：

$$2x = (l+m) - \mathrm{i}\sqrt{3}(m-l).$$

LMX 的形心是

$$u = \tfrac{1}{3}(l+m+x) = \tfrac{1}{3}(l+m) + \tfrac{1}{6}(l+m) - \tfrac{\mathrm{i}\sqrt{3}}{6}(m-l) = \frac{l+m}{2} + \frac{\mathrm{i}(l-m)}{2\sqrt{3}},$$

其中用到 $\frac{\sqrt{3}}{6} = \frac{1}{2\sqrt{3}}$。作輪換 $l \to m \to n \to l$（這樣會保持逆時針的次序），得

$$v = \frac{m+n}{2} + \frac{\mathrm{i}(m-n)}{2\sqrt{3}}, \qquad w = \frac{n+l}{2} + \frac{\mathrm{i}(n-l)}{2\sqrt{3}}.$$

我們現在證明 U, V, W 滿足逆時針的關係式 $2w = (u+v) + \mathrm{i}\sqrt{3}(v-u)$。首先，

$$u+v = \tfrac{1}{2}l + m + \tfrac{1}{2}n + \frac{\mathrm{i}(l-n)}{2\sqrt{3}},$$
$$v-u = \tfrac{1}{2}(n-l) + \frac{\mathrm{i}(2m-l-n)}{2\sqrt{3}},$$
$$\mathrm{i}\sqrt{3}(v-u) = \frac{\mathrm{i}\sqrt{3}}{2}(n-l) - \tfrac{1}{2}(2m-l-n).$$

想一想 m 是怎樣消失的？

把第一行和第三行相加。實部給出 $\frac{1}{2}l + m + \frac{1}{2}n - m + \frac{1}{2}l + \frac{1}{2}n = l + n$，所以 m 在實部中互相抵消，而它從來沒有在虛部中出現過。虛部合併成

$$(n-l)\Big(\frac{\sqrt{3}}{2} - \frac{1}{2\sqrt{3}}\Big)\mathrm{i} = (n-l)\,\frac{3-1}{2\sqrt{3}}\,\mathrm{i} = \frac{\mathrm{i}(n-l)}{\sqrt{3}}.$$

因此

$$(u+v) + \mathrm{i}\sqrt{3}(v-u) = (n+l) + \frac{\mathrm{i}(n-l)}{\sqrt{3}} = 2w.$$

根據開首的結果（及其逆命題），UVW 是一個頂點按逆時針方向排列的等邊三角形。這就是拿破崙定理（Napoleon's theorem）。

驗算

- 開首的關係式：$a = 0$，$b = 2$ 給出 $c = 1 + \mathrm{i}\sqrt{3}$，頂點在實軸上方，與逆時針排列的三角形應有的情況相符。
- 第 (ii) 部分，取 $l = 0$，$m = 6$，$n = 6\mathrm{i}$：由公式得 $u = 3 - \mathrm{i}\sqrt{3}$，$v = (3 + \sqrt{3})(1 + \mathrm{i})$ 和 $w = -\sqrt{3} + 3\mathrm{i}$。於是 UVW 三條邊的長度平方都是 $24 + 12\sqrt{3}$。
- 我們用 10^4 個隨機構形，以數值方法檢驗了開首的關係式、平行四邊形條件和拿破崙定理。我們亦確認了，在第 (i) 部分若每條邊都用同一方向，一般而言並不成立。

常見錯誤

- 在第 (ii) 部分弄錯向外的方向。對於按逆時針方向排列的 LMN，LM 上向外的頂點相對於 L, M 是順時針方向的。用錯了正負號，得到的就會是內拿破崙三角形。（它同樣是等邊三角形，但並不是題目所問的。）
- 在第 (i) 部分，試圖令四個三角形全部指向同一方。這只在 $DEFG$ 本身是平行四邊形時才行得通。
- 用 $p + q = r + s$ 來檢驗平行四邊形。應該是相對的頂點配成一對：條件是 $p + r = q + s$。

STEP 之外

背後的想法是：映射 $z \mapsto \alpha z + \beta$ 正好就是平面上的旋轉、放大和平移，所以幾何條件會變成 $\mathbb{C}$ 上的線性方程。開首的關係式有一個簡潔的改寫方式：ABC 是等邊三角形，當且僅當 $a + \omega^2 b + \omega^4 c = 0$ 或 $a + \omega^4 b + \omega^2 c = 0$，而拿破崙定理只需幾行代數運算便隨之而得。Needham 的 *Visual Complex Analysis*（《複分析：可視化方法》）把複數乘法的這種幾何觀點發揮得淋漓盡致。

詳細解答

STEP III 2008，第 8 題

純數學 • 級數與遞推關係

STEP III 2008，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/08-s3-q8

關鍵思路

係數滿足線性遞推關係（linear recurrence），意味着只要把級數乘以適當的多項式，幾乎所有項都會互相抵消。這個多項式可以直接從遞推關係讀出來：$a_{r+1} + pa_r = 0$ 提示我們用 $1 + px$，而 $a_{r+2} + pa_{r+1} + qa_r = 0$ 則提示我們用 $1 + px + qx^2$。餘下的只是一個很短的多項式，所以級數之和是一個有理函數（rational function）。對於有限項之和，同樣的技巧會在末端留下幾個額外的項。另一個做法是用部分分式（partial fractions）把級數拆成幾個等比級數（geometric series），而等比級數的和我們是懂得求的。

解答

第 (i) 部分

這裏級數由 x 項開始，所以 $a_1 = \frac{1}{3}$，$a_2 = \frac{1}{6}$，$a_3 = \frac{1}{12}$。每個係數都是前一個的一半，所以 $a_{r+1} - \frac{1}{2}a_r = 0$，即

$$p = -\tfrac{1}{2}.$$

（這裏沒有常數項。遞推關係只把 $a_1, a_2, a_3, \ldots$ 聯繫起來，對 $r = 0$ 的情況並無任何說明。）解這個遞推關係，得 $a_r = \frac{1}{3}\left(\frac{1}{2}\right)^{r-1}$。

無窮項之和（sum to infinity）。 乘以 $1 - \frac{1}{2}x$：

$$\left(1 - \tfrac{1}{2}x\right) S = a_1 x + \sum_{r \geq 1} \left(a_{r+1} - \tfrac{1}{2}a_r\right) x^{r+1} = \tfrac{1}{3}x,$$

因為根據遞推關係，每個括號都等於零。因此

$$S = \frac{\frac{1}{3}x}{1-\frac{1}{2}x} = \frac{2x}{3(2-x)}.$$

這是一個公比為 $\frac{1}{2}x$ 的等比級數，所以無窮項之和恰好在 $|x| < 2$ 時存在。

首 $n+1$ 項。 這些就是 $x, x^2, \ldots, x^{n+1}$ 各項。設 $S_{n+1} = \sum_{r=1}^{n+1} a_r x^r$。重複同樣的乘法，中間各項仍然全部抵消，但最高次的一項 $-\frac{1}{2}a_{n+1}x^{n+2}$ 卻沒有對應的項與它抵消：

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

所以，當 $x \neq 2$ 時，

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

當 $x = 2$ 時，每一項都等於 $\frac{2}{3}$，所以 $S_{n+1} = \frac{2}{3}(n+1)$。

第 (ii) 部分

現在 $a_0 = 2$，$a_1 = 8$，$a_2 = 18$，$a_3 = 37$。把 $r = 0$ 和 $r = 1$ 代入遞推關係，得到兩條方程：

$$18 + 8p + 2q = 0, \qquad 37 + 18p + 8q = 0.$$

由第一條方程得 $q = -9 - 4p$。代入第二條方程，得 $37 + 18p - 72 - 32p = 0$，所以 $p = -\frac{5}{2}$，從而 $q = 1$。遞推關係是

$$a_{r+2} - \tfrac{5}{2}a_{r+1} + a_r = 0.$$

無窮項之和。 乘以 $1 + px + qx^2 = 1 - \frac{5}{2}x + x^2$。對每個 $r \geq 0$，乘積中 x^{r+2} 的係數是 $a_{r+2} - \frac{5}{2}a_{r+1} + a_r = 0$。只有 x^0 和 x^1 的係數留下來：

$$\left(1 - \tfrac{5}{2}x + x^2\right) T = a_0 + \left(a_1 - \tfrac{5}{2}a_0\right) x = 2 + 3x.$$

把分子和分母同時加倍，並把 $2 - 5x + 2x^2 = (1-2x)(2-x)$ 因式分解，得

$$T = \frac{4+6x}{(1-2x)(2-x)}.$$

部分分式。 設 $\dfrac{4+6x}{(1-2x)(2-x)} = \dfrac{A}{1-2x} + \dfrac{B}{2-x}$。用遮蓋法（cover-up），在 $x = \frac{1}{2}$ 處得 $A = \dfrac{7}{3/2} = \frac{14}{3}$，在 $x = 2$ 處得 $B = \dfrac{16}{-3} = -\frac{16}{3}$。所以

$$T = \frac{14}{3}\cdot\frac{1}{1-2x} - \frac{8}{3}\cdot\frac{1}{1-\frac{1}{2}x}.$$

把每一項展開成等比級數，得

$$a_r = \tfrac{14}{3}\,2^r - \tfrac{8}{3}\left(\tfrac{1}{2}\right)^r.$$

驗算：$r = 0, 1, 2, 3$ 分別給出 2、$\frac{28}{3} - \frac{4}{3} = 8$、$\frac{56}{3} - \frac{2}{3} = 18$ 和 $\frac{112}{3} - \frac{1}{3} = 37$，與已知的值相符。

想一想 無窮項之和究竟對哪些 x 才存在？

兩個等比級數都必須收斂（converge），所以需要 $|2x| < 1$ 和 $|\frac{1}{2}x| < 1$。第一個條件較為嚴格，所以我們需要 $|x| < \frac{1}{2}$。

首 $n+1$ 項。 這些就是由 x^0 到 x^n 的各項。把每個等比級數求和至 $n+1$ 項，當 $x \neq \frac{1}{2}, 2$ 時，得

$$\sum_{r=0}^{n} a_r x^r = \frac{14}{3} \cdot \frac{1-(2x)^{n+1}}{1-2x} - \frac{16}{3} \cdot \frac{1-\left(\frac{x}{2}\right)^{n+1}}{2-x}.$$

（第二個分式就是 $\frac{8}{3} \cdot \frac{1-(x/2)^{n+1}}{1-x/2}$，只是分子和分母都乘以 2。）若 $x = \frac{1}{2}$ 或 $x = 2$，對應的等比級數之和就只是 $n+1$ 個 1 相加。

驗算

- 在 $|x| < \frac{1}{2}$ 的情況下令 $n \to \infty$：冪 $(2x)^{n+1}$ 和 $(x/2)^{n+1}$ 趨向零，於是我們重新得到 T 的部分分式形式。在第 (i) 部分，當 $|x| < 2$ 時情況相同。
- 取 $n = 0$，第 (i) 部分的公式給出 $\frac{2x}{3(2-x)} \cdot \frac{2-x}{2} = \frac{1}{3}x$。第 (ii) 部分的公式給出 $\frac{14}{3} - \frac{16}{3} \cdot \frac{1}{2} = 2$。兩者都是首項，正如所料。
- 數值上，在 $x = 0.3$ 處有 $T = \dfrac{5.8}{0.4 \times 1.7} \approx 8.5294$，而 $2 + 8x + 18x^2 + 37x^3 + \cdots$ 的部分和（partial sums）亦趨近這個值。

常見錯誤

- 在第 (i) 部分，把首項當作 $a_0 x^0$。級數由 x^1 開始，所以首 $n+1$ 項一直去到 x^{n+1}。
- 把有限項之和乘以 $1 + px$ 時，忘記了末端剩下的項。正是這些項產生了 $(\frac{x}{2})^{n+1}$ 這個因子。
- 對 T 寫出 $|x| < 2$。決定收斂半徑（radius of convergence）的，是較接近原點的根 $x = \frac{1}{2}$。

STEP 之外

冪級數（power series）$\sum a_r x^r$ 是數列 (a_r) 的生成函數（generating function）。常係數線性遞推關係必定給出有理的生成函數，而這個函數的極點（poles）既決定 a_r 的閉合形式（closed form），也決定收斂半徑。生成函數是組合數學（combinatorics）和概率論中的標準工具，而冪級數的收斂半徑則會在大學一年級的分析課程中嚴謹地建立起來。Abbott 的 *Understanding Analysis*（一本分析學入門教科書）對收斂這方面有很好的講解。

詳細解答

STEP III 2008，第 9 題

力學 • 摩擦力與變力

STEP III 2008，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：

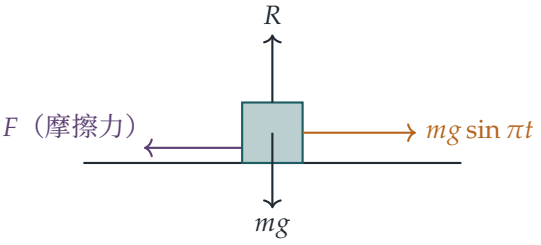
step.maths.org/questions/08-s3-q9

關鍵思路

摩擦力有兩種狀態。質點（particle）靜止時，靜摩擦力（static friction）恰好與外加的力相等，上限為 μmg，因此加速度為零。質點一旦滑動，摩擦力的大小固定為 μmg，方向與速度相反。所以加速度的圖像由若干段平坦的零值線段和平移後的正弦曲線連接而成。「靜止下來」的意思是：速度（即自運動開始起加速度圖像下的面積）回到零。

解答

我們把物體看成粗糙水平面上的質點。法向反作用力（normal reaction）為 $R = mg$，所以靜止時摩擦力的大小至多為 μmg，滑動時則恰好為 μmg。取 $\sin \pi t > 0$ 時外力的方向為正方向，並設 a 和 v 分別為加速度和速度。



開始運動

沿鉛垂方向分解，得 $R = mg$。質點靜止時，水平方向的平衡需要大小為 $mg|\sin \pi t|$ 的摩擦力，而只要 $mg|\sin \pi t| \leq \mu mg$，這是可以做到的。當外力首次超過極限摩擦力

（limiting friction）時，質點便首次開始運動，所以（當 $\mu < 1$ 時）

$$\sin \pi T_0 = \mu, \qquad \text{即} \qquad T_0 = \frac{1}{\pi} \arcsin \mu, \quad 0 \le T_0 < \tfrac{1}{2}.$$

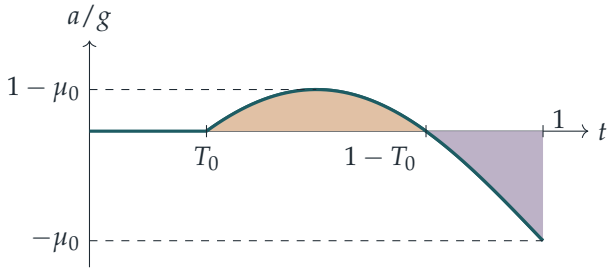
（若 $\mu \ge 1$，質點永遠不會運動。）

第 (i) 部分

當 $0 \le t \le T_0$ 時質點靜止，所以 $a = 0$。當 $t > T_0$ 時，質點沿正方向滑動，摩擦力 $\mu_0 mg$ 指向後方，由牛頓（Newton）第二定律得

$$a = g(\sin \pi t - \mu_0) \qquad (\text{只要} v > 0).$$

這個加速度在 $t = T_0$ 時由 0 開始，在 $t = \frac{1}{2}$ 時升至 $g(1 - \mu_0)$，在 $t = 1 - T_0$ 時回落至 0（由對稱性 $\sin \pi(1 - t) = \sin \pi t$ 可知），此後一直為負，在 $t = 1$ 時達到 $-\mu_0 g$。



速度就是這個圖像由 T_0 起計的面積：曲線在橫軸之上時（琥珀色區域）速度增加，過了 $1 - T_0$ 之後（紫紅色區域）速度減少。在兩塊面積互相抵銷之前，速度一直為正，而抵銷的一刻就是質點停下來的時刻。所以我們需要

$$v(1) = \int_{T_0}^{1} g(\sin \pi t - \mu_0)\, \mathrm{d}t = 0.$$

現在

$$\begin{aligned} \int_{T_0}^{1} (\sin \pi t - \mu_0)\, \mathrm{d}t &= \Big[-\frac{\cos \pi t}{\pi} \Big]_{T_0}^{1} - \mu_0 (1 - T_0) \\ &= \frac{1 + \cos \pi T_0}{\pi} - \mu_0 + \mu_0 T_0. \end{aligned}$$

令上式等於零，再乘以 π：

$$1 + \cos \pi T_0 - \mu_0 \pi + \mu_0 \pi T_0 = 0.$$

想一想 用 μ_0 表示 $\cos \pi T_0$ 時，它應取哪個符號？

由於 $0 \le \pi T_0 < \frac{\pi}{2}$，餘弦為正，所以 $\cos \pi T_0 = \sqrt{1 - \sin^2 \pi T_0} = (1 - \mu_0^2)^{1/2}$。另外 $\pi T_0 = \arcsin \mu_0$。把兩者代入，得

$$1 + \left(1 - \mu_0^2\right)^{\frac{1}{2}} - \mu_0 \pi + \mu_0 \arcsin \mu_0 = 0,$$

這正是所要證明的。（數值上 $\mu_0 \approx 0.7246$，$T_0 \approx 0.258$。）

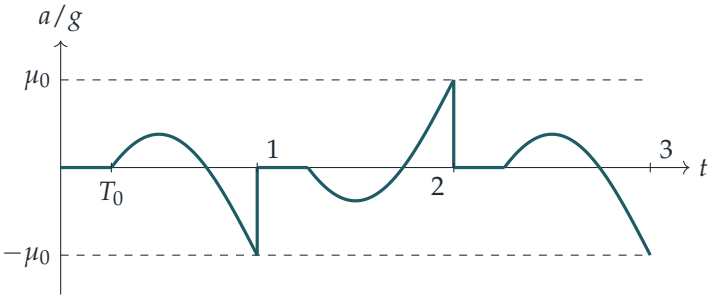
第 (ii) 部分

在 $t = 1$ 時，質點靜止，外力為 $mg \sin \pi = 0$。當 $1 < t < 2$ 時，外力指向負方向，大小為 $mg|\sin \pi t|$。在 $|\sin \pi t| = \mu_0$ 之前，這個力小於 $\mu_0 mg$，而該等式首次成立是在 $t = 1 + T_0$（因為 $\sin \pi(1 + T_0) = -\sin \pi T_0$）。所以當 $1 \le t \le 1 + T_0$ 時，質點保持靜止，$a = 0$。

此後質點沿負方向滑動，摩擦力指向正方向，於是

$$a = g(\sin \pi t + \mu_0).$$

由於 $\sin \pi(t + 1) = -\sin \pi t$，這恰好是第 (i) 部分中時刻 $t - 1$ 的加速度的相反數。因此 $[1, 2]$ 上的運動是 $[0, 1]$ 上運動的鏡像：速度是先前速度的相反數，所以質點在 $t = 2$ 時再次靜止，並且往回走了相同的距離。這時質點靜止而外力為零，與 $t = 0$ 時的情況完全相同，所以在 $[2, 3]$ 上，$[0, 1]$ 的運動會重演一次。



在 $t = 1$ 和 $t = 2$ 處的鉛垂線段表示加速度的跳躍：在質點停下的一瞬間，動摩擦力（kinetic friction，大小為 $\mu_0 mg$）被靜摩擦力取代，而靜摩擦力只需平衡大小為零的外力。

$\mu = \mu_0$ 時的運動。質點一直靜止至 $t = T_0$，然後向前滑動，在 $t = 1$ 時停下，等到 $t = 1 + T_0$，再向後滑動，在 $t = 2$ 時停在出發點，如此類推。它在同一段地面上來回往返，週期為 2，在兩端各停留 T_0 的時間。（再積分一次，可得這段路的長度為 $g\left[\mu_0/\pi^2 - (1 - T_0)(1 - \sqrt{1 - \mu_0^2})/(2\pi)\right] \approx 0.0367g$，不過題目並沒有要求這個結果。）

$\mu = 0$ 時的運動。這時 $T_0 = 0$，而且在任何時刻都有 $a = g \sin \pi t$，所以

$$v = \frac{g}{\pi}(1 - \cos \pi t) \ge 0, \qquad x = \frac{g}{\pi}\left(t - \frac{\sin \pi t}{\pi}\right).$$

速度永不為負，並且只在 $t = 0, 2, 4, \ldots$ 時為零。所以質點永不回頭：它一陣一陣地向前衝，在偶數時刻瞬間靜止，每個長度為 2 的週期前進 $2g/\pi$，平均速率為 g/π。在粗糙的情況下，令質點在 $t = 1$ 停下來的正是摩擦力；沒有摩擦力時，外力向後的那一半只會把質點減速至速度為零。

驗算

用二分法（bisection）解第 (i) 部分的方程，得 $\mu_0 \approx 0.72461$，因而 $T_0 \approx 0.25798$。取 $g = 9.8$，對 $0 \le t \le 3$ 的運動作數值模擬，並在靜摩擦力與動摩擦力之間切換，結果顯示質點在 $t \approx 0.258$ 時起動，走了 0.3598 之後在 $t = 1.000$ 時停下，在 $t \approx 1.258$ 時開始往回走，並在 $t = 2.000$ 時停在出發點，與 $0.0367g$ 吻合。當 $\mu = 0$ 時，模擬結果與 $x = g(t - \sin \pi t/\pi)/\pi$ 一致。對方程作一個簡單的合理性檢查：在 $\mu_0 = 0$ 時左邊為 $2 > 0$，在 $\mu_0 = 1$ 時左邊為 $1 - \pi/2 < 0$，所以在 $(0, 1)$ 內有根。

常見錯誤

- 由 $t = 0$ 起就取 $a = g(\sin \pi t - \mu)$：在 T_0 之前質點靜止，摩擦力是靜摩擦力，所以 $a = 0$。
- 在第 (ii) 部分中，質點反向之後仍讓摩擦力保持原來的方向。摩擦力總是與速度方向相反。
- 以為質點在 $t = 1 - T_0$ 時停下，但那時為零的是加速度，而不是速度。

STEP 之外

摩擦力在「黏住」與「滑動」之間切換，令運動方程變成分段（piecewise）的形式，加速度亦變得不連續。這就是門鉸吱吱作響和小提琴弓拉琴弦背後的黏滑（stick–slip）現象，也是非光滑動力系統（non-smooth dynamical system）的一個入門例子。Kleppner 與 Kolenkow 合著的 *An Introduction to Mechanics*（力學入門教科書）以合適的程度討論了摩擦力模型和受迫運動。

詳細解答

STEP III 2008，第 10 題

力學 • 彈性繩與能量

STEP III 2008，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/08-s3-q10

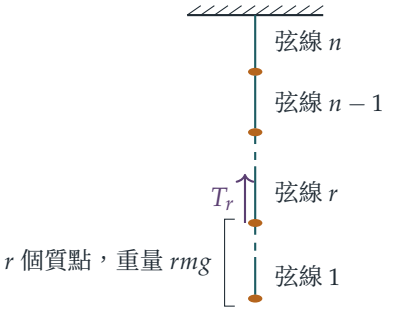
關鍵思路

每條短弦線都要承托懸掛在它下方的一切，所以它的張力（tension）很容易寫出；再由胡克定律（Hooke's law）得出它的伸長量（extension）和所儲存的能量，而總和就是 r 與 r^2 的標準求和。繩子的情況其實是同一幅圖畫，只是質量分散開來：把它切成 n 等份，把每一份的質量當作位於其下端的一個質點（particle），再讓 $n \to \infty$。求和變成積分，而長度中的因子 $\frac{1}{2}$ 之所以出現，是因為平均而言，繩子的每一小段承托著繩子一半的重量。

解答

掛著質點的弦線

把各條短弦線視為遵從胡克定律的輕彈性弦線（light elastic string），並把各個質量視為質點。按題目所述，由底部起為弦線編號，於是弦線 r 下方有 r 個質點：它自己的質點，以及弦線 $1, \ldots, r-1$ 的質點。



整個系統處於靜止狀態。考慮弦線 r 頂端以下的一切：弦線 r、弦線 $1,\dots,r-1$，以及掛在它們上面的 r 個質點。作用於這部分的外力只有向下的重量 rmg 和弦線 r 中向上的張力 T_r（弦線是輕的）。所以

$$T_r = rmg.$$

根據胡克定律，弦線 r 的伸長量為

$$x_r = \frac{T_r\,\ell}{\lambda} = \frac{rmg\ell}{\lambda}.$$

所以長弦線的長度為

$$\sum_{r=1}^{n}(\ell + x_r) = n\ell + \frac{mg\ell}{\lambda}\sum_{r=1}^{n} r = n\ell + \frac{n(n+1)mg\ell}{2\lambda}.$$

自然長度(natural length)為 ℓ 的弦線伸長了 x 時，其彈性能(elastic energy)為 $\lambda x^2/(2\ell)$。對弦線 r 而言，這等於

$$\frac{\lambda}{2\ell}\Big(\frac{rmg\ell}{\lambda}\Big)^2 = \frac{r^2m^2g^2\ell}{2\lambda},$$

再利用 $\sum_{r=1}^{n} r^2 = \frac{1}{6}n(n+1)(2n+1)$，總彈性能為

$$\frac{m^2g^2\ell}{2\lambda}\cdot\frac{n(n+1)(2n+1)}{6} = \frac{n(n+1)(2n+1)\,m^2g^2\ell}{12\lambda}.$$

有重量的繩子

想一想 對於質量連續分佈的繩子，第一部分能幫上甚麼忙？

把繩子切成 n 等份。每一份的自然長度為 $\ell = L_0/n$，質量為 $m = M/n$，而且每一份的彈性模量（modulus）都與整條繩子相同，同為 λ（模量與小段的長度無關）。如果把每一份的質量集中在其下端，得到的正正就是上面那條長弦線。當 $n\to\infty$ 時，各小段愈來愈短，這個集中質量的模型便趨近真實的繩子。

代入長度的公式，

$$n\ell + \frac{n(n+1)mg\ell}{2\lambda} = L_0 + \frac{n(n+1)}{2\lambda}\cdot\frac{Mg}{n}\cdot\frac{L_0}{n} = L_0 + \frac{MgL_0}{2\lambda}\Big(1+\frac{1}{n}\Big) \longrightarrow L_0\Big(1+\frac{Mg}{2\lambda}\Big)$$

當 $n\to\infty$。所以

$$L = L_0\Big(1+\frac{Mg}{2\lambda}\Big),$$

正如所求。

你也可以直接計算。設 s 為某一點下方繩子的自然長度，則該點所承托的重量為 Mgs/L_0。在該處取自然長度為 δs 的一小段，其張力為 $T = Mgs/L_0$，所以它伸長了 $T\,\delta s/\lambda$。全部加起來，

$$L - L_0 = \int_0^{L_0}\frac{Mgs}{\lambda L_0}\,\mathrm{d}s = \frac{MgL_0}{2\lambda}.$$

至於能量，作同樣的代入可得

$$\frac{n(n+1)(2n+1)}{12\lambda}\cdot\frac{M^2g^2}{n^2}\cdot\frac{L_0}{n}=\frac{M^2g^2L_0}{12\lambda}\left(1+\frac{1}{n}\right)\left(2+\frac{1}{n}\right)\longrightarrow\frac{M^2g^2L_0}{6\lambda}.$$

（直接計算：該小段儲存的能量為 $T^2\,\delta s/(2\lambda)$，而 $\int_0^{L_0}\frac{M^2g^2s^2}{2\lambda L_0^2}\,\mathrm{d}s=\frac{M^2g^2L_0}{6\lambda}$。）

要把結果以 L、L_0 和 λ 表示，可利用長度的公式消去 Mg：

$$L-L_0=\frac{MgL_0}{2\lambda}\quad\Longrightarrow\quad Mg=\frac{2\lambda(L-L_0)}{L_0}.$$

因此繩子所儲存的彈性能為

$$\frac{L_0}{6\lambda}\cdot\frac{4\lambda^2(L-L_0)^2}{L_0^2}=\frac{2\lambda(L-L_0)^2}{3L_0}.$$

驗算

- 取 $n=1$ 時，第一部分給出長度 $\ell+mg\ell/\lambda$ 和能量 $m^2g^2\ell/(2\lambda)$：這正是一條弦線掛著一個質點的情況，理應如此。
- 試比較一條總伸長量同為 $e=L-L_0$、但張力均勻的繩子：它所儲存的能量是 $\lambda e^2/(2L_0)$。我們的答案是這個值的 $\frac{4}{3}$ 倍，這是合理的，因為拉伸集中在頂部附近，而能量隨局部應變（strain）的平方增長。
- 數值上，取 $M=3$、$L_0=2$、$\lambda=50$、$g=9.8$：公式給出 $L=2.588$ 和能量 5.7624；取 $n=100\,000$ 的離散鏈則給出 2.58801 和 5.76249，而 $2\lambda(L-L_0)^2/(3L_0)=5.7624$。

常見錯誤

- 把弦線 r 的張力當作 $(n-r+1)mg$：按題目的編號，弦線 n 在最頂，所以弦線 r 承托的是 r 個質點。
- 把每一小段繩子的模量取為 λ/n。模量是材料和橫截面的性質，與長度無關；隨長度改變的是勁度（stiffness）λ/ℓ。
- 題目要求以 L、L_0 和 λ 表示能量，卻把答案留在以 M 和 g 表示的形式。

STEP 之外

由 n 個集中質量過渡到連續的繩子，是連續介質力學（continuum mechanics）最基本的一步：張力變成函數 $T(s)$，胡克定律變成局部應力（stress）與應變之間的關係，而能量變成能量密度的積分。同樣是這條由質量和彈簧組成的鏈，一旦讓它振動起來，取

極限後便得到弦線的波動方程（wave equation）。Kleppner 與 Kolenkow 所著的 *An Introduction to Mechanics*（力學入門教科書）是了解這些想法如何展開的好去處。

詳細解答

STEP III 2008，第 11 題

力學 • 轉動動力學

STEP III 2008，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：

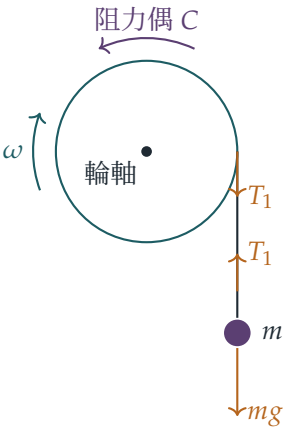
step.maths.org/questions/08-s3-q11

關鍵思路

力偶（couple）恆定，即角減速度（angular deceleration）恆定，所以繩子脫離之後，輪子服從轉動版本的「suvat」方程，而第一個結果不過是「距離 = 平均速率 × 時間」。同一階段亦讓我們以 I 表示阻力偶的大小。至於下降階段，能量方程是最自然的工具：我們知道各部分移動了多遠，而克服恆定力偶所作的功就是力偶 × 角度。

解答

模型。輪子是繞固定光滑輪軸轉動的剛體（rigid body），唯一例外是受到一個大小恆為 C 的阻力偶（resisting couple），它在整個運動過程中都起作用（質點下降期間如是，其後亦如是）。繩子是輕而不可伸長的，並且不打滑，所以繩子仍然連着時，質點的速率為 $v = r\omega$，其中 ω 是輪子的角速率（angular speed）。



繩子從輪子的右邊離開，所以質點下落時輪子順時針轉動，而力偶 C 則逆時針作用，阻礙運動。（我們其實用不着繩子的張力（tension）T_1，因為我們用的是能量方法。）

繩子脫離之後

此時輪子所受的力偶只有 C，所以對輪軸取矩（taking moments），

$$I\dot{\omega} = -C,$$

而角減速度 C/I 是常數。因此角速率在時間 T 內由 ω_0 線性地下降至 0，所以

$$\omega_0 = \frac{C}{I}\,T, \qquad \text{即} \qquad C = \frac{I\omega_0}{T}. \tag{1}$$

角速率既然線性下降，平均角速率就是 $\frac{1}{2}(\omega_0 + 0)$，而轉過的角度等於平均角速率乘以時間。這個角度是 n_2 轉，即 $2\pi n_2$ 弧度（radians），所以

$$\tfrac{1}{2}\omega_0 T = 2\pi n_2,$$

這就是第一個結果。（它是 $s = \frac{1}{2}(u+v)t$ 的轉動版本。）由此亦得

$$\omega_0 = \frac{4\pi n_2}{T}, \qquad \text{於是由 (1) 得} \qquad C = \frac{4\pi n_2 I}{T^2}. \tag{2}$$

質點下降期間

輪子轉過 n_1 轉（即 $2\pi n_1$ 弧度）期間，繩子鬆開的長度為 $2\pi n_1 r$，所以質點下落的距離為 $2\pi n_1 r$。在脫離的一刻，輪子的角速率為 ω_0，而由於繩子沒有打滑，質點的速率為 $r\omega_0$。

想一想 既然有阻力，為甚麼這裏仍可以用能量？

能量並不守恆，但功能原理（work–energy principle）依然成立：勢能（potential energy）的損失等於動能（kinetic energy）的增加，再加上克服阻力所作的功。轉過角度 θ 時，克服恆定力偶 C 所作的功為 $C\theta$，正如克服恆力所作的功是力 $\times$ 距離。張力對整個系統不作淨功，因為繩子輕而不可伸長，並且不打滑，所以它對輪子所作的功恰好等於它從質點取走的功。

所以，由釋放至脫離，

$$mg\,(2\pi n_1 r) = \tfrac{1}{2}m(r\omega_0)^2 + \tfrac{1}{2}I\omega_0^2 + C\,(2\pi n_1).$$

現在把 (2) 中的 ω_0 和 C 代入：

$$2\pi mgrn_1 = \tfrac{1}{2}\left(mr^2 + I\right)\frac{16\pi^2 n_2^2}{T^2} + 2\pi n_1 \cdot \frac{4\pi n_2 I}{T^2}.$$

全式乘以 $T^2/(2\pi)$，

$$\begin{aligned} mgrn_1T^2 &= 4\pi n_2^2\,(mr^2+I) + 4\pi n_1 n_2 I \\ &= 4\pi mr^2 n_2^2 + 4\pi n_2(n_1+n_2)\,I. \end{aligned}$$

移項，再除以 $4\pi n_2(n_1+n_2)$（此數為正），

$$I = \frac{mgrn_1T^2 - 4\pi mr^2 n_2^2}{4\pi n_2(n_1+n_2)},$$

正是所求。

驗算

- 量綱。分子的兩項 $mgrT^2$ 和 mr^2 都具有轉動慣量(moment of inertia)的單位($\mathrm{kg\,m^2}$)，而 n_1、n_2 是純數。
- 數值驗算。我們取 $I = 0.5\,\mathrm{kg\,m^2}$、$m = 1\,\mathrm{kg}$、$r = 0.2\,\mathrm{m}$、$C = 0.5\,\mathrm{N\,m}$、$g = 9.8\,\mathrm{m\,s^{-2}}$ 和 $n_1 = 5$，以數值方法逐步計算這個運動。結果是 $\omega_0 \approx 13.03\,\mathrm{rad\,s^{-1}}$、$T \approx 13.03\,\mathrm{s}$ 和 $n_2 \approx 13.52$。於是 $\frac{1}{2}\omega_0 T \approx 84.94 \approx 2\pi n_2$，而公式給出 $I = 0.5000$。
- 沒有阻力。若 $C \to 0$，輪子永不停下，所以 T 和 n_2 都趨向無限大；公式的表現仍然合理，因為分子的增長如 T^2，而分母的增長如 $n_2^2 \propto T^2$。

常見錯誤

- 漏掉下降期間克服阻力所作的功。題目說明力偶是恆定的，而且它自始至終都起作用，並非只在繩子脫離之後。少了這一項，公式就會出錯：在上面的例子中，會得出 $I \approx 0.39$ 而不是 0.5。
- 忘記質點的動能 $\frac{1}{2}mr^2\omega_0^2$；$-4\pi mr^2 n_2^2$ 這一項正是由此而來。
- 混淆轉數與弧度：n 轉等於 $2\pi n$ 弧度。

STEP 之外

這是一個巧妙的實驗方法：只要量度輪子由轉動至停下所需的時間，並數算轉數，就可以同時測出它的轉動慣量和軸承的摩擦，即使事前兩者皆不知道。大學本科的力學會把這些內容發展成完整的剛體轉動理論，其中角動量(angular momentum)和力矩(torque)分別扮演動量和力的角色，而對於繞任意軸旋轉的物體，轉動慣量則變成一個張量（tensor）。Kleppner 與 Kolenkow 合著的 *An Introduction to Mechanics*（力學導論）對此有細緻的闡述，並附有許多正是這種風格的習題。

詳細解答

STEP III 2008，第 12 題

概率與統計 • 矩生成函數

STEP III 2008，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/08-s3-q12

本題在 *Beyond STEP* 第 12.24 節中心極限定理中有完整的引導式講解：https://beyondstep.keithmcnulty.org/zh-hant/ch12/clt.html

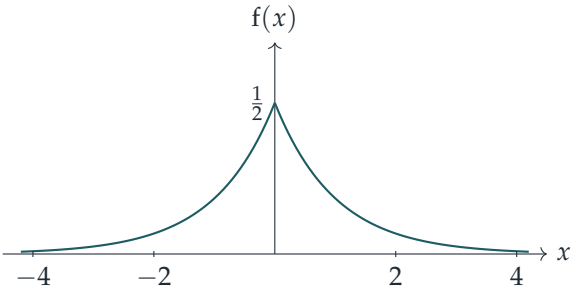
關鍵思路

密度函數中的絕對值（absolute value）把每個積分分成 $x < 0$ 和 $x > 0$ 兩半，每一半都是普通的指數積分。此後整道題只靠矩母函數（moment generating function，簡稱 m.g.f.）的兩個性質：獨立（independent）變量之和的矩母函數，等於各自矩母函數之積；以及 $\mathrm{M}_{aY}(\theta) = \mathrm{M}_Y(a\theta)$。取對數後，$n$ 次方就變成 n 乘以一個很小的量，這正是麥克勞林級數（Maclaurin series）最擅長處理的情況。這就是中心極限定理（central limit theorem），而且在這個例子中，每一步都可以親手算出來。

解答

略圖

密度函數是偶函數，在 $x = 0$ 處取得最大值 $\frac{1}{2}$，該處有一個尖角（左、右導數分別為 $+\frac{1}{2}$ 和 $-\frac{1}{2}$），並在兩側以指數形式遞減至 0。



矩母函數

根據定義，$\mathrm{M}_X(\theta) = \mathrm{E}(\mathrm{e}^{\theta X})$。在 0 處把積分分開，

$$\begin{aligned}\mathrm{M}_X(\theta) &= \tfrac{1}{2}\int_{-\infty}^{0} \mathrm{e}^{(1+\theta)x}\,\mathrm{d}x + \tfrac{1}{2}\int_{0}^{\infty} \mathrm{e}^{-(1-\theta)x}\,\mathrm{d}x \\ &= \frac{1}{2(1+\theta)} + \frac{1}{2(1-\theta)} = \frac{(1-\theta)+(1+\theta)}{2(1-\theta^2)} = \frac{1}{1-\theta^2}.\end{aligned}$$

第一個積分只在 $1+\theta > 0$ 時收斂（converge），第二個只在 $1-\theta > 0$ 時收斂，所以上式對 $-1 < \theta < 1$ 成立，而我們只需要 $\theta = 0$ 附近的情況，這已經足夠。這就是所求的結果。

至於方差（variance），當 $|\theta| < 1$ 時，把它展開為等比級數（geometric series）：

$$\mathrm{M}_X(\theta) = 1 + \theta^2 + \theta^4 + \cdots.$$

與 $\mathrm{M}_X(\theta) = 1 + \mathrm{E}(X)\,\theta + \frac{1}{2}\mathrm{E}(X^2)\,\theta^2 + \cdots$ 比較，可得 $\mathrm{E}(X) = 0$ 和 $\frac{1}{2}\mathrm{E}(X^2) = 1$。因此

$$\mathrm{Var}(X) = \mathrm{E}(X^2) - \big(\mathrm{E}(X)\big)^2 = 2.$$

（求導兩次再代入 $\theta = 0$ 也得到相同結果，但用級數較快。）

青蛙：$Y/\sqrt{2n}$ 的矩母函數

由於各次跳躍 $X_1, \ldots, X_n$ 互相獨立，$\mathrm{e}^{\theta X_1}, \ldots, \mathrm{e}^{\theta X_n}$ 也互相獨立，所以它們乘積的期望值（expectation）等於各期望值之積：

$$\mathrm{M}_Y(\theta) = \mathrm{E}\big(\mathrm{e}^{\theta X_1}\cdots\mathrm{e}^{\theta X_n}\big) = \prod_{i=1}^{n}\mathrm{M}_{X_i}(\theta) = (1-\theta^2)^{-n}.$$

設 $Z = Y/\sqrt{2n}$，則 $\mathrm{E}(\mathrm{e}^{\theta Z}) = \mathrm{E}\big(\mathrm{e}^{(\theta/\sqrt{2n})Y}\big) = \mathrm{M}_Y\big(\theta/\sqrt{2n}\big)$，所以

$$\mathrm{M}_Z(\theta) = \left(1 - \frac{\theta^2}{2n}\right)^{-n}, \qquad |\theta| < \sqrt{2n}.$$

以 $\sqrt{2n}$ 作縮放是很自然的：$\mathrm{Var}(Y) = 2n$，所以 Z 的平均值為 0，方差為 1。

極限

固定 θ。只要 $n > \frac{1}{2}\theta^2$，便有 $0 \le \theta^2/(2n) < 1$，於是可以取對數，並使用對 $|u| < 1$ 成立的 $\ln(1-u) = -u - \frac{1}{2}u^2 - \frac{1}{3}u^3 - \cdots$：

$$\begin{aligned}\ln \mathrm{M}_Z(\theta) &= -n \ln\left(1 - \frac{\theta^2}{2n}\right) = n\left(\frac{\theta^2}{2n} + \frac{\theta^4}{8n^2} + \frac{\theta^6}{24n^3} + \cdots\right) \\ &= \frac{\theta^2}{2} + \frac{\theta^4}{8n} + \frac{\theta^6}{24n^2} + \cdots .\end{aligned}$$

? **想一想** 我們怎樣知道，當 $n \to \infty$ 時，這個級數的尾部真的會消失？

設 $u = \theta^2/(2n)$，首項之後的所有項之和為 $n\left(\frac{1}{2}u^2 + \frac{1}{3}u^3 + \cdots\right) \le n \cdot \frac{1}{2}u^2(1 + u + u^2 + \cdots) = \dfrac{nu^2}{2(1-u)} = \dfrac{\theta^4}{8n(1-u)}$，由於 $u \to 0$，這個量趨向 0。所以 $\ln \mathrm{M}_Z(\theta) \to \frac{1}{2}\theta^2$，而 exp 是連續（continuous）函數，故

$$\mathrm{M}_{Y/\sqrt{2n}}(\theta) \to \exp\left(\tfrac{1}{2}\theta^2\right) \quad \text{當} n \to \infty,$$

這正是所要證明的。

最少跳躍次數

由於這個極限正是標準正態分佈（standard Normal）的矩母函數，當 n 很大時，我們把 $Z = Y/\sqrt{2n}$ 視為近似服從 $\mathrm{N}(0,1)$。「落在距離起點 25 cm 或以上的位置」是指 $|Y| \ge 25$，兩個方向都計算在內，所以

$$\mathrm{P}(|Y| \ge 25) = \mathrm{P}\left(|Z| \ge \frac{25}{\sqrt{2n}}\right) \approx 2\left(1 - \Phi\left(\frac{25}{\sqrt{2n}}\right)\right).$$

這個概率隨 n 增加而上升（青蛙散佈得更遠），所以我們要找使它達到 5% 的最小 n。由於 $\Phi(1.96) = 0.975$，我們需要

$$\frac{25}{\sqrt{2n}} \le 1.96 \iff 2n \ge \left(\frac{25}{1.96}\right)^2 \approx 162.7 \iff n \ge 81.3.$$

所以最少跳躍次數是 $n = 82$。（當 $n = 81$ 時，近似概率約為 4.95%；當 $n = 82$ 時，約為 5.09%。）

驗算

- $\mathrm{M}_X(0) = 1$，任何矩母函數都必須如此；而且 M_X 是偶函數，與略圖的對稱性吻合。
- 以數值積分（numerical quadrature）計算 $\int x^2 \mathrm{f}(x)\,\mathrm{d}x$ 得 2.0000，計算 $\int \mathrm{e}^{0.3x}\mathrm{f}(x)\,\mathrm{d}x$ 得 $1.0989 = 1/0.91$。

- 取 $\theta = 1$，$(1 - \frac{1}{2n})^{-n}$ 在 $n = 10$ 時為 1.6702，在 $n = 100$ 時為 1.6508，逐漸逼近 $e^{1/2} = 1.6487$。
- 模擬 10^5 隻青蛙，每隻作 82 次拉普拉斯（Laplace）跳躍，結果約有 5.1% 的青蛙最後落在距離起點至少 25 cm 的位置。

常見錯誤

- 使用單尾（one-tailed）數值 1.645：青蛙可以落在任何一側 25 cm 以外，所以兩邊尾部都要計算。
- 寫成 $M_{Y/\sqrt{2n}}(\theta) = M_Y(\theta)/\sqrt{2n}$。把變量縮放，縮放的是其矩母函數的自變量。
- 把 81.3 四捨五入至最接近的整數而答 $n = 81$：當 $n = 81$ 時，概率僅略低於 5%。

STEP 之外

你剛才證明的，就是步長服從拉普拉斯分佈時的中心極限定理，而一般情況的證明也循同一路線：把 $\ln M(\theta/\sqrt{n})$ 展開至二階，再證明餘下部分會消失。在大學一年級的課程中，你會遇到「連續性定理」（continuity theorem），它說明為何可以由矩母函數（或特徵函數（characteristic function），它必定存在）的收斂推出概率的收斂，這一步我們在這裏只是姑且相信。Grimmett 與 Welsh 的 *Probability: An Introduction*（概率論入門）對兩者都有論述。

詳細解答

STEP III 2008，第 13 題

概率與統計 • 指示變量之和

STEP III 2008，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/08-s3-q13

本題在 *Beyond STEP* 第 12.10 節把繩頭繫起來中有完整的引導式講解：https://beyondstep.keithmcnulty.org/zh-hant/ch12/loose-ends.html

關鍵思路

無論某一步發生甚麼事，繩子的段數總是剛好減少一：要麼有一個繩圈退出遊戲，要麼兩段繩子合成一段。因此每一步之後，不管之前發生過甚麼，我們都回到同樣的處境，只是少了一段繩子。這樣就可以把繩圈的總數寫成一些獨立（independent）指示變量（indicator variable）之和，於是平均值和方差（variance）都只是簡單的求和。

解答

第一步

有 n 段繩子時，共有 $2n$ 個自由端，所以選取兩個端點的方法有

$$\binom{2n}{2} = n(2n-1)$$

種，而且各種選法的可能性相等。當且僅當兩個端點屬於同一段繩子時，才會形成繩圈，而這樣的配對共有 n 個（每段繩子一個）。因此

$$\mathrm{P}(\text{第一步形成繩圈}) = \frac{n}{n(2n-1)} = \frac{1}{2n-1}.$$

換個說法：無論我們先選哪一個端點，餘下的 $2n-1$ 個端點之中，恰好有一個是它的同伴。設 X_n 在第一步形成繩圈時取值 1，否則取值 0，則第一步所形成繩圈數目的期望值

（expectation）是

$$\mathrm{E}(X_n) = \frac{1}{2n-1}.$$

繩圈數目的期望值

第一步之後，剩下 $n-1$ 段繩子，每段各有兩個自由端。兩種情況下都是如此：形成繩圈會令一段繩子退出遊戲，而把兩段不同的繩子接起來則令兩段變成一段。繩子的長短無關重要，要緊的只是它們的端點。

> **想一想** 現在有些繩子比其他的長，又或者之前有沒有形成過繩圈，這些有影響嗎？

沒有影響。唯一會影響選中哪一對端點的，就是還剩下多少段繩子（因而還剩下多少個端點）。所以當剩下 k 段繩子時，不論之前各步的結果如何，這一步形成繩圈的概率都是 $\frac{1}{2k-1}$。

以每一步開始前的繩子段數 $k = n, n-1, \ldots, 1$ 來標記各步，並設 X_k 為該步形成繩圈的指示變量。於是 $\mathrm{P}(X_k = 1) = \frac{1}{2k-1}$，而整個過程在 $k = 1$ 那一步之後結束（這一步必定形成繩圈，即 $X_1 = 1$）。繩圈的總數是

$$R = X_1 + X_2 + \cdots + X_n,$$

所以由期望值的線性性質（linearity of expectation）可得

$$\mathrm{E}(R) = \sum_{k=1}^{n} \frac{1}{2k-1} = 1 + \frac{1}{3} + \frac{1}{5} + \cdots + \frac{1}{2n-1}.$$

方差

求方差時，單靠線性性質並不足夠：我們還需要各個 X_k 互相獨立。它們確實如此。無論之前各步的結果是甚麼，有 k 段繩子的那一步形成繩圈的概率仍然是 $\frac{1}{2k-1}$。所以在已知之前所有 X 的條件下，X_k 的條件分佈（conditional distribution）與它的無條件分佈相同，因此各個 X_k 互相獨立。

每個 X_k 都服從伯努利分佈（Bernoulli distribution），其中 $p_k = \frac{1}{2k-1}$，所以 $\mathrm{Var}(X_k) = p_k(1-p_k)$；而對於獨立變量，方差可以相加：

$$\begin{aligned}\mathrm{Var}(R) &= \sum_{k=1}^{n} \frac{1}{2k-1}\left(1 - \frac{1}{2k-1}\right) = \sum_{k=1}^{n} \frac{2k-2}{(2k-1)^2} \\ &= \sum_{k=1}^{n} \frac{1}{2k-1} - \sum_{k=1}^{n} \frac{1}{(2k-1)^2}.\end{aligned}$$

$n = 40\,000$ 的情況

記 $H_m = 1 + \frac{1}{2} + \cdots + \frac{1}{m}$。直至 $2n-1$ 為止的奇數倒數之和，等於直至 $2n$ 為止的所有倒數之和，減去其中偶數的倒數：

$$\sum_{k=1}^{n} \frac{1}{2k-1} = H_{2n} - \left(\frac{1}{2} + \frac{1}{4} + \cdots + \frac{1}{2n}\right) = H_{2n} - \tfrac{1}{2}H_n.$$

利用 m 很大時 $H_m \approx \ln m$，

$$\mathrm{E}(R) \approx \ln(2n) - \tfrac{1}{2}\ln n = \tfrac{1}{2}\left(2\ln 2n - \ln n\right) = \tfrac{1}{2}\ln(4n).$$

當 $n = 40\,000$ 時，$4n = 160\,000 = 20^4$，所以

$$\mathrm{E}(R) \approx \tfrac{1}{2}\ln\left(20^4\right) = 2\ln 20 \approx 6.$$

因此，由 40 000 段繩子開始，我們預期只會得到大約六個繩圈。

驗算

- $n = 1$：只有一段繩子，必定得到一個繩圈。公式給出 $\mathrm{E}(R) = 1$ 和 $\mathrm{Var}(R) = 0$，正如所料。
- $n = 2$：第一步形成繩圈的概率是 $\frac{1}{3}$，而最後一步必定形成繩圈，所以 R 取值 2 的概率是 $\frac{1}{3}$，否則取值 1。由此得平均值為 $\frac{4}{3}$，方差為 $\frac{2}{9}$，與公式給出的 $1 + \frac{1}{3}$ 和 $\frac{2}{9}$ 吻合。
- 對 $n \le 5$ 精確地列舉所有可能的選擇序列，並對 $n = 6$ 進行模擬，結果都與兩條公式一致。當 $n = 40\,000$ 時，精確的和是 $\mathrm{E}(R) \approx 6.28$。比我們的估計多出的 0.29，來自 $H_m \approx \ln m$ 這個粗略的近似（見下文）。

常見錯誤

- 把形成繩圈的概率當作 $\frac{1}{n}$（選兩段繩子），而不是 $\frac{1}{2n-1}$（選兩個端點）。
- 沒有先說明各個 X_k 互相獨立，就把方差相加。求平均值時有線性性質已經足夠，求方差則不然。
- 沒有把偶數項分拆出來，就以 $\ln n$ 或 $\frac{1}{2}\ln n$ 來近似 $\sum \frac{1}{2k-1}$。

STEP 之外

更精確的估計是 $H_m = \ln m + \gamma + o(1)$，其中 $\gamma \approx 0.5772$ 是歐拉－馬斯刻若尼常數（Euler–Mascheroni constant）。由此可得 $\mathrm{E}(R) \approx \frac{1}{2}\ln(4n) + \frac{1}{2}\gamma$，以及

$\mathrm{Var}(R) \approx \frac{1}{2}\ln(4n) + \frac{1}{2}\gamma - \frac{\pi^2}{8}$，因為 $\sum 1/(2k-1)^2 = \pi^2/8$。同樣的「概率為 $1/k$ 的獨立指示變量之和」結構，也可以用來數算隨機置換（random permutation）的循環（cycle）數目，以及隨機數列中破紀錄（record）的次數；兩種情況下，這個數目的增長都與 $\ln n$ 相若。想進一步了解指示變量和這些隨機結構，Grimmett 與 Welsh 的 *Probability: An Introduction*（概率論入門教材）是個很好的起點。

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

2008 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。