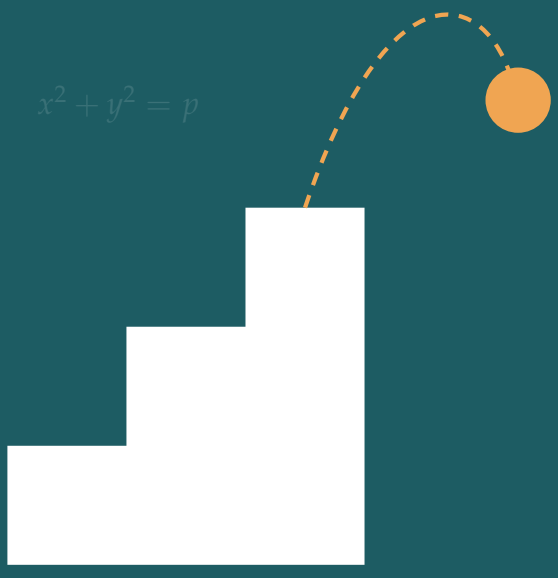
STEP 備試 · 大學數學

Beyond STEP

STEP I, II 及 III

2003 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

2003 年詳細解答

Keith McNulty

Beyond STEP：以歷屆 STEP 試題為門，走進大學數學

2003 年詳細解答

2003 年 STEP 試卷（STEP I, II 及 III）全部 42 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 冪和 純數學 . . . 3

第 2 題 代數恆等式 純數學 . . . 7

第 3 題 半角三角方程 純數學 . . . 10

第 4 題 三角不等式 純數學 . . . 14

第 5 題 二項式級數 純數學 . . . 17

第 6 題 以部分分式積分 純數學 . . . 20

第 7 題 數字與整除性 純數學 . . . 23

第 8 題 羅輯斯諦微分方程 純數學 . . . 26

第 9 題 通過定點的拋體 力學 . . . 29

第 10 題 薄片的質心 力學 . . . 33

第 11 題 鏡面上的碰撞 力學 . . . 37

第 12 題 乘積的期望值 概率與統計 . . . 41

第 13 題 幾何分佈的期望值 概率與統計 . . . 44

第 14 題 泊松分佈 概率與統計 . . . 47

STEP II

第 1 題 聯立線性方程 純數學 . . . 50

第 2 題 反三角函數 純數學 . . . 54

第 3 題 無理數與數學歸納法 純數學 . . . 57

第 4 題 弓形的面積 純數學 . . . 60

第 5 題 直線的向量幾何 純數學 . . . 64

第 6 題 迭代函數與積分 純數學 . . . 67

第 7 題 廣義積分與級數 純數學 . . . 70

第 8 題 可分離變量微分方程 純數學 . . . 73

第 9 題 極限摩擦力平衡 力學 77
第 10 題 滑及摩擦力的約束運動 力學 81
第 11 題 擺體運動 力學 85
第 12 題 分支過程 概率與統計 89
第 13 題 截尾泊松分佈 概率與統計 93
第 14 題 假設檢驗的錯誤 概率與統計 97

STEP III

第 1 題 反雙曲函數積分法 純數學 101
第 2 題 二項式級數 純數學 105
第 3 題 曲線描繪 純數學 108
第 4 題 參數曲線的切線 純數學 112
第 5 題 三次方程的根 純數學 116
第 6 題 三角函數裂項相消求和 純數學 121
第 7 題 西姆松線 純數學 125
第 8 題 一階微分方程 純數學 129
第 9 題 平衡的穩定性 力學 133
第 10 題 變加速度 力學 137
第 11 題 攔截運動學 力學 141
第 12 題 整數分佈的等候時間 概率與統計 145
第 13 題 吸收馬可夫鏈 概率與統計 149
第 14 題 概率生成函數 概率與統計 153

詳細解答

STEP I 2003，第 1 題

純數學 • 冪和

> **STEP I 2003，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/03-s1-q1

關鍵思路

題目已經告訴我們答案的形式：一個以 n 為變數、係數未知的多項式（polynomial），所以餘下的工作只是把這些係數確定下來。一個對所有 n 都成立的多項式恆等式（identity），對幾個方便計算的 n 值當然也成立，而每個值都給出一條線性方程（linear equation）。讓求和由負數下標開始是刻意安排的技巧：這樣我們便可以使用 $n = -1$ 和 $n = -2$ 等較小的值，令方程組呈現悅目的對稱性。

解答

題目中 r 既用作求和變數，又用作係數。為免混淆，我們把求和變數寫成 k，而 p, q, r, s 和 a, b, c, d, e 則保留作係數。

平方和

設 $S(n) = \sum_{k=-1}^{n} k^2$，於是

$$S(n) = pn^3 + qn^2 + rn + s.$$

由 $k = -1$ 開始的和對所有 $n \geq -1$ 都有意義，而且

$$S(-1) = 1, \qquad S(0) = 1 + 0 = 1, \qquad S(1) = 1 + 0 + 1 = 2, \qquad S(2) = 1 + 0 + 1 + 4 = 6.$$

把 $n=-1,0,1,2$ 代入題目給出的形式，得到四條方程

$$-p+q-r+s=1, \tag{1}$$
$$s=1, \tag{2}$$
$$p+q+r+s=2, \tag{3}$$
$$8p+4q+2r+s=6. \tag{4}$$

由 (2) 得 $s=1$，所以 (1) 和 (3) 變成 $-p+q-r=0$ 和 $p+q+r=1$。兩式相加得 $2q=1$，所以 $q=\frac{1}{2}$，從而 $p+r=\frac{1}{2}$。方程 (4) 變成

$$8p+2+2r+1=6 \quad\Longrightarrow\quad 8p+2r=3 \quad\Longrightarrow\quad 6p+2(p+r)=3,$$

所以 $6p=2$，得 $p=\frac{1}{3}$ 和 $r=\frac{1}{2}-\frac{1}{3}=\frac{1}{6}$。因此

$$\sum_{k=-1}^{n} k^2=\tfrac{1}{3}n^3+\tfrac{1}{2}n^2+\tfrac{1}{6}n+1.$$

想一想 題目要求的是由 $k=0$ 開始的和，但我們求得的是由 $k=-1$ 開始的和。兩者有何分別？

兩者只相差 $k=-1$ 那一項，即 $(-1)^2=1$。所以

$$\sum_{k=0}^{n} k^2=\tfrac{1}{3}n^3+\tfrac{1}{2}n^2+\tfrac{1}{6}n=\tfrac{1}{6}n\left(2n^2+3n+1\right)=\tfrac{1}{6}n(n+1)(2n+1),$$

這正是所要證明的結果。

立方和

現在設 $T(n)=\sum_{k=-2}^{n}k^3=an^4+bn^3+cn^2+dn+e$。這裏有五個未知數，所以我們取 $n=-2,-1,0,1,2$ 這五個值：

$$T(-2)=-8,\quad T(-1)=-9,\quad T(0)=-9,\quad T(1)=-8,\quad T(2)=0.$$

由此得到

$$16a-8b+4c-2d+e=-8, \tag{5}$$
$$a-b+c-d+e=-9, \tag{6}$$
$$e=-9, \tag{7}$$
$$a+b+c+d+e=-8, \tag{8}$$
$$16a+8b+4c+2d+e=0. \tag{9}$$

由 (7) 得 $e=-9$。這些方程按 $\pm n$ 成對出現，所以把每一對相加和相減最為划算：偶次冪和奇次冪便會分開。

利用 $e=-9$，(6) 和 (8) 變成 $a-b+c-d=0$ 和 $a+b+c+d=1$。相加和相減，得

$$a+c=\tfrac{1}{2}, \qquad b+d=\tfrac{1}{2}.$$

同樣地，(5) 和 (9) 變成 $16a-8b+4c-2d=1$ 和 $16a+8b+4c+2d=9$。相加和相減，得

$$32a+8c=10, \qquad 16b+4d=8,$$

即 $16a+4c=5$ 和 $4b+d=2$。代入 $c=\frac{1}{2}-a$ 得 $12a+2=5$，所以 $a=\frac{1}{4}$ 和 $c=\frac{1}{4}$。代入 $d=\frac{1}{2}-b$ 得 $3b+\frac{1}{2}=2$，所以 $b=\frac{1}{2}$ 和 $d=0$。因此

$$\sum_{k=-2}^{n} k^3=\tfrac{1}{4}n^4+\tfrac{1}{2}n^3+\tfrac{1}{4}n^2-9.$$

與由 $k=0$ 開始的和相比，由 $k=-2$ 開始的和多了 $(-2)^3+(-1)^3=-9$ 這些項，所以我們加上 9：

$$\sum_{k=0}^{n} k^3=\tfrac{1}{4}n^4+\tfrac{1}{2}n^3+\tfrac{1}{4}n^2=\tfrac{1}{4}n^2\left(n^2+2n+1\right)=\tfrac{1}{4}n^2(n+1)^2,$$

這正是所要證明的結果。

驗算

當 $n=10$ 時，兩條公式給出 $\frac{1}{6}\cdot 10\cdot 11\cdot 21=385$ 和 $\frac{1}{4}\cdot 100\cdot 121=3025$，而直接把 $0^2+1^2+\cdots+10^2$ 和 $0^3+1^3+\cdots+10^3$ 加起來，也得到相同的數字。還可留意 $3025=55^2=(0+1+\cdots+10)^2$：首 n 個立方數之和，等於首 n 個整數之和的平方。我們亦用電腦核對過，兩條公式對由 0 至 59 的每個 n 都準確成立。

常見錯誤

- 由我們算出的和轉到題目要求的和時，忘記了 $k=-1$（或 $k=-2$ 和 $k=-1$）的項。注意 $(-2)^3+(-1)^3=-9$，而不是 $+9$。
- 處理立方和時只打算用四個 n 值：五個未知數需要五條方程。
- 把係數 r 與題目中的求和變數 r 混淆；把其中一個改名便可避免出錯。

STEP 之外

為甚麼我們可以假設答案是一個次數高一次的多項式？因為一個 $m+1$ 次多項式的差分 $P(n)-P(n-1)$ 是 m 次的，所以總可以解出一個差分為 n^m 的多項式；這就是*有限差分演算*（calculus of finite differences）的起點，而一般情況下出現的係數由伯努利數（Bernoulli numbers）決定，即福爾哈伯公式（Faulhaber's formula）。這裏所用的方法，即憑多項式在若干點上的值把它確定下來，稱為*多項式插值*(polynomial interpolation)，它在數值分析（numerical analysis）中會以拉格朗日插值（Lagrange interpolation）的形式再度出現。想做更多同類風格的題目，可參閱 Siklos 的 *Advanced Problems in Mathematics*。

詳細解答

STEP I 2003，第 2 題

純數學 • 代數恆等式

> **STEP I 2003，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/03-s1-q2

關鍵思路

考生的答案正確，當且僅當它滿足方程，所以把它代入（substitute）方程，再消去分母。這樣每一部分都會化為一個關於已知數的多項式（polynomial）條件。在第一部分，這個條件是一個二次型（quadratic form），對非零實數而言不可能等於零。在第二部分，它是一個三次式，可以漂亮地因式分解（factorise），而所得的因式正正就是題目問及的那三條等式。

解答

第一問

在整個解答中，a 和 b 都是非零實數。考生的答案正確，恰好就是 $x = a + b$ 滿足方程的時候。要令這句話有意義，首先需要 $a + b \neq 0$，然後需要

$$\frac{1}{a+b} = \frac{1}{a} + \frac{1}{b} = \frac{a+b}{ab}.$$

兩邊同乘以非零的 $ab(a+b)$，這等價於

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

配方（completing the square）得

$$a^2 + ab + b^2 = \left(a + \tfrac{1}{2}b\right)^2 + \tfrac{3}{4}b^2.$$

兩項都是實數的平方，所以都 ≥ 0，而且因為 $b \neq 0$，第二項嚴格大於零。因此對任意實數 a 和任意非零實數 b，都有 $a^2 + ab + b^2 > 0$，所以這個條件永遠不可能成立。

（若 $a+b=0$，考生的「答案」$x=0$ 代入方程後根本沒有意義，所以這個情況也好不到哪裏去。）因此不存在任何 a 和 b 的值使 $x=a+b$ 正確，這正是所要證明的。

第二問

現在 a、b、c 都非零。答案 $x=a+b+c$ 正確，當且僅當（if and only if）$a+b+c\neq 0$ 並且

$$\frac{1}{a+b+c}=\frac{1}{a}+\frac{1}{b}+\frac{1}{c}=\frac{bc+ca+ab}{abc}.$$

兩邊同乘以 $abc(a+b+c)\neq 0$，這等價於

$$(a+b+c)(ab+bc+ca)-abc=0. \qquad (*)$$

想一想 怎樣才能看出 $(*)$ 的左邊可以因式分解？

要證明的結果提示我們，因式應該是 $a+b$、$b+c$ 和 $c+a$。先試其中一個：設 $b=-a$。這時 $a+b+c=c$，而 $ab+bc+ca=-a^2-ac+ca=-a^2$，所以左邊變成 $c(-a^2)-(-a^2c)=0$。因此 $a+b$ 是一個因式，而根據對稱性（symmetry），$b+c$ 和 $c+a$ 也是因式。左邊的次數是 3，所以它是 $(a+b)(b+c)(c+a)$ 的常數倍；比較 a^2b 的係數（兩邊都是 1），可知這個常數是 1。另一個做法是把兩邊都展開：

$$\begin{aligned}(a+b+c)(ab+bc+ca)-abc &= a^2b+a^2c+b^2a+b^2c+c^2a+c^2b+3abc-abc\\ &= a^2b+a^2c+b^2a+b^2c+c^2a+c^2b+2abc\\ &= (a+b)(b+c)(c+a).\end{aligned}$$

所以 $(*)$ 等同於

$$(a+b)(b+c)(c+a)=0.$$

現在我們證明兩個方向。

若答案正確。那麼 $a+b+c\neq 0$ 而且 $(*)$ 成立，所以 $(a+b)(b+c)(c+a)=0$。若干個數的乘積為零，其中必有一個因數為零，所以 $a+b=0$、$b+c=0$、$c+a=0$ 之中至少有一條成立。

反過來，假設這三條等式中至少有一條成立；根據對稱性，不妨設 $a+b=0$。那麼 $a+b+c=c\neq 0$，所以考生的答案是一個合理的非零數，而且 $(*)$ 的因式分解形式成立。由於從方程推到 $(*)$ 的每一步都是可逆的（我們只是乘以非零的數 $abc(a+b+c)$），所以 $x=a+b+c$ 滿足方程。這一點也可以直接看出：$\frac{1}{a}+\frac{1}{b}=\frac{1}{a}-\frac{1}{a}=0$，所以方程變成 $\frac{1}{x}=\frac{1}{c}$，它的解是 $x=c=a+b+c$。

因此，考生的答案正確，當且僅當 $a+b=0$、$b+c=0$、$c+a=0$ 之中至少有一條成立。 ■

留意這個論證從來沒有用到 a、b、c 是實數；這個因式分解是一個代數恆等式（identity），所以結果對非零複數（complex numbers）同樣成立。相比之下，第一部分確實需要 a 和 b 是實數：在複數範圍內，$a^2+ab+b^2=0$ 是有解的，例如 $b=a\,\mathrm{e}^{2\pi\mathrm{i}/3}$。

驗算

- 取 $a=2$、$b=-2$、$c=5$：這時 $\frac{1}{2}-\frac{1}{2}+\frac{1}{5}=\frac{1}{5}$，而 $a+b+c=5$，所以考生是對的。
- 取 $a=b=c=1$：正確答案是 $x=\frac{1}{3}$，但考生的答案是 3；而事實上 $(a+b)(b+c)(c+a)=8\neq 0$。
- 在 $a=1$、$b=2$、$c=3$ 時展開 $(a+b)(b+c)(c+a)$：得 $3\cdot 5\cdot 4=60$，而 $(a+b+c)(ab+bc+ca)-abc=6\cdot 11-6=60$。
- 我們用精確的分數運算，對全部 17 576 個由非零有理數 n/d（其中 $|n|\leq 6$ 且 $d\in\{1,2,3\}$）組成的三元組，窮舉檢驗了兩個方向，沒有發現任何例外。

常見錯誤

- 寫下 $a^2+ab+b^2=(a+b)^2-ab$ 便停筆：這並不能說明它是正數。配方才可以。
- 在第二部分只證明一個方向。「當且僅當」需要證明兩個方向，而且逆命題（converse）必須檢查 $a+b+c\neq 0$。
- 忘記在兩邊同乘之前，必須先確定 $abc(a+b+c)$ 非零；正是這一點令各步可逆。

STEP 之外

表達式 $(a+b)(b+c)(c+a)$ 是一個對稱多項式（symmetric polynomial），而恆等式 $(a+b)(b+c)(c+a)=e_1e_2-e_3$（其中 $e_1=a+b+c$、$e_2=ab+bc+ca$、$e_3=abc$）正是對稱多項式基本定理（fundamental theorem of symmetric polynomials）的一個例子：每個對稱多項式都可以寫成這些基本對稱多項式的多項式。除了正負號之外，這些 e_1、e_2、e_3 也就是以 a、b、c 為根的三次多項式的係數，這正是它們在方程論（theory of equations）中隨處可見的原因。Allenby 的 *Rings, Fields and Groups* 本着這種精神討論多項式環（polynomial rings），而 Siklos 的 *Advanced Problems in Mathematics* 則有更多辨認因式分解的 STEP 練習。

詳細解答

STEP I 2003，第 3 題

純數學 • 半角三角方程

STEP I 2003，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/03-s1-q3

關鍵思路

每一部分都是把 $\frac{1}{2}\theta$ 的某個函數與 θ 的同一個函數作比較，所以可利用倍角公式（double-angle formula），把所有式子都以半角（half-angle）表示。這樣，每條方程都會變成只含單一個量（$\sin\frac{1}{2}\theta$, $\tan\frac{1}{2}\theta$ 或 $\cos\frac{1}{2}\theta$）的多項式方程（polynomial equation），可以用因式分解（factorise）或當作二次方程（quadratic）來解。唯一要小心的，是留意半角函數實際上可以取哪些值，以及各函數在哪裏有定義。

解答

第 (i) 部分

利用 $\sin\theta = 2\sin\frac{1}{2}\theta\cos\frac{1}{2}\theta$，方程 $2\sin\frac{1}{2}\theta = \sin\theta$ 等價（equivalent）於

$$2\sin\tfrac{1}{2}\theta - 2\sin\tfrac{1}{2}\theta\cos\tfrac{1}{2}\theta = 0, \qquad \text{即} \qquad 2\sin\tfrac{1}{2}\theta\left(1-\cos\tfrac{1}{2}\theta\right) = 0.$$

所以方程成立，當且僅當（if and only if）$\sin\frac{1}{2}\theta = 0$ 或 $\cos\frac{1}{2}\theta = 1$。

但若 $\cos\frac{1}{2}\theta = 1$，則 $\sin^2\frac{1}{2}\theta = 1 - \cos^2\frac{1}{2}\theta = 0$，所以無論如何都有 $\sin\frac{1}{2}\theta = 0$。因此第二種可能已包含在第一種之內，於是

$$2\sin\tfrac{1}{2}\theta = \sin\theta \iff \sin\tfrac{1}{2}\theta = 0,$$

正如所求。（換句話說，方程的解恰好是 $\theta = 2n\pi$，其中 n 為整數。）

第 (ii) 部分

設 $t = \tan\frac{1}{2}\theta$。要令方程有意義，$\tan\frac{1}{2}\theta$ 和 $\tan\theta$ 都必須有定義，所以 $\cos\frac{1}{2}\theta \neq 0$ 且 $\cos\theta \neq 0$。由於 $\cos\theta = \dfrac{1-t^2}{1+t^2}$，第二個條件即 $t^2 \neq 1$。在這些限制之下，$\tan\theta = \dfrac{2t}{1-t^2}$，方程變成

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

兩邊乘以不等於零的 $1-t^2$，得 $2t - 2t^3 = 2t$，所以 $t^3 = 0$，即 $t = 0$。

> **想一想** 兩邊乘以 $1-t^2$ 是否穩妥？我們有沒有因此失去了解，或多出了解？

乘以一個非零的量是可逆的步驟，所以在兩個正切都有定義的集合上，這條關於 t 的方程與 $t = 0$ 完全等價。反過來，$t = 0$ 給出 $\tan\theta = 0$ 和 $\tan\frac{1}{2}\theta = 0$，兩邊都等於 0。

而 $\tan\frac{1}{2}\theta = 0$ 恰好在 $\frac{1}{2}\theta = n\pi$ 時成立。所以方程的解是

$$\theta = 2n\pi, \qquad n \in \mathbb{Z},$$

與第 (i) 部分的解集相同。

第 (iii) 部分

設 $c = \cos\frac{1}{2}\theta$。利用 $\cos\theta = 2\cos^2\frac{1}{2}\theta - 1 = 2c^2 - 1$，方程 $2\cos\frac{1}{2}\theta = \cos\theta$ 等價於

$$2c = 2c^2 - 1 \iff 2c^2 - 2c - 1 = 0 \iff c = \frac{2 \pm \sqrt{4+8}}{4} = \frac{1 \pm \sqrt{3}}{2}.$$

由於 c 是一個餘弦值，所以 $-1 \leq c \leq 1$。根 $\frac{1}{2}(1+\sqrt{3}) \approx 1.37$ 太大，而 $\frac{1}{2}(1-\sqrt{3}) \approx -0.37$ 則在容許範圍內。所以方程成立，當且僅當

$$\cos\tfrac{1}{2}\theta = \tfrac{1}{2}(1-\sqrt{3}) = -\cos\phi,$$

其中 ϕ 的定義與題目所給的相同。這個 ϕ 存在而且唯一，因為 $0 < \frac{1}{2}(\sqrt{3}-1) < 1$，而 cos 在 $[0, \frac{1}{2}\pi]$ 上由 1 遞減至 0。

> **想一想** 怎樣把 $\cos\frac{1}{2}\theta = -\cos\phi$ 化成題目所給的形式？

寫成 $-\cos\phi = \cos(\pi - \phi)$。方程 $\cos x = \cos\alpha$ 的通解 (general solution) 是 $x = 2m\pi \pm \alpha$，所以

$$\tfrac{1}{2}\theta = 2m\pi \pm (\pi - \phi) \qquad \text{即} \qquad \theta = 4m\pi \pm (2\pi - 2\phi), \quad m \in \mathbb{Z}.$$

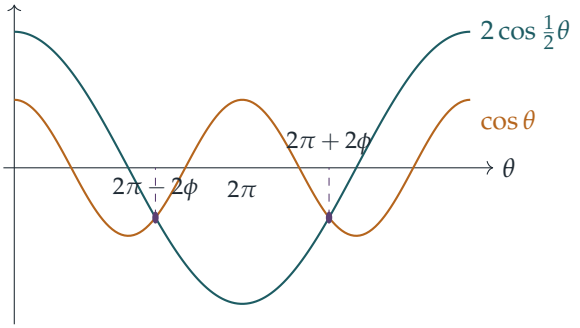
把兩個正負號分開處理。

- 取 $+$ 號：$\theta = 4m\pi + 2\pi - 2\phi = (4n+2)\pi - 2\phi$，其中 $n = m$。
- 取 $-$ 號：$\theta = 4m\pi - 2\pi + 2\phi = \big(4(m-1)+2\big)\pi + 2\phi = (4n+2)\pi + 2\phi$，其中 $n = m-1$。

當 m 取遍所有整數時，兩種情況下的 n 也取遍所有整數。因此

$$2\cos\tfrac{1}{2}\theta = \cos\theta \iff \theta = (4n+2)\pi \pm 2\phi\text{，對某個整數}n.$$

每一步都是等價關係，所以兩個方向都已證明。我們也可以直接驗證逆命題：若 $\theta = (4n+2)\pi \pm 2\phi$，則 $\cos\tfrac{1}{2}\theta = \cos\big((2n+1)\pi \pm \phi\big) = -\cos\phi = c$，而它滿足 $2c = 2c^2 - 1$。



在 $2\cos\tfrac{1}{2}\theta$ 的一個週期（period）$0 \le \theta \le 4\pi$ 內，兩條曲線相交兩次，交點關於 $\theta = 2\pi$ 對稱，而 $2\cos\tfrac{1}{2}\theta$ 正是在該處取得最小值 -2。

驗算

- 以數值計算，$\cos\phi \approx 0.3660$ 給出 $\phi \approx 1.1961$，所以 (iii) 的第一個正數解是 $2\pi - 2\phi \approx 3.8911$。在該處 $2\cos\tfrac{1}{2}\theta \approx -0.7321$ 而 $\cos\theta \approx -0.7321$，符合要求。在 $[-6\pi, 6\pi]$ 上對 $2\cos\tfrac{1}{2}\theta - \cos\theta$ 進行求根搜尋，找到的恰好是該範圍內形如 $(4n+2)\pi \pm 2\phi$ 的六個值。
- 對 (ii) 作類似的搜尋（避開極點（poles)），只在 2π 的倍數處找到根。
- 對 (i) 和 (ii) 的合理性檢查：$\theta = \pi$ 不滿足 (i)，因為 $2\sin\tfrac{1}{2}\pi = 2 \neq 0 = \sin\pi$；它亦不在 (ii) 的定義域（domain）之內，因為 $\tan\tfrac{1}{2}\pi$ 沒有定義。

常見錯誤

- 在 (i) 中，約去 $\sin\tfrac{1}{2}\theta$ 而沒有說明它可能等於零；或把 $\cos\tfrac{1}{2}\theta = 1$ 當作另一族獨立的解，而沒有留意到它已包含在前者之內。

- 在 (ii) 中，去分母之前忘記了 θ 必須避開 $\tan\theta$ 或 $\tan\frac{1}{2}\theta$ 沒有定義的點。
- 在 (iii) 中，保留了大於 1 的根 $\frac{1}{2}(1+\sqrt{3})$；或在把 $4m\pi \pm (2\pi - 2\phi)$ 化成 $(4n+2)\pi \pm 2\phi$ 的形式時，錯誤處理 $\pm$ 號。

STEP 之外

每一部分都是以半角的單一個函數作代換（substitution）來解方程，從而把三角方程化為代數方程。(ii) 中的代換 $t = \tan\frac{1}{2}\theta$ 尤其強大：它把 $\sin\theta$, $\cos\theta$ 和 $\tan\theta$ 都表示成 t 的有理函數（rational function），這正是大學一年級微積分用它來求 $\sin\theta$ 和 $\cos\theta$ 的任何有理函數的積分的原因。從幾何上看，它以有理點（rational points）把單位圓參數化（parametrise）。Siklos 的 *Advanced Problems in Mathematics* 一書中，有更多運用這些想法的 STEP 式練習。

詳細解答

STEP I 2003，第 4 題

純數學 • 三角不等式

> **STEP I 2003，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/03-s1-q4

關鍵思路

最容易想到的做法是把兩邊同乘 $\cos\theta$，但 $\cos\theta$ 在 $[0, 2\pi)$ 上會變號，而乘以負數會令不等式（inequality）的方向逆轉。因此我們分成 $\cos\theta > 0$ 和 $\cos\theta < 0$ 兩種情況討論。其中一種情況其實十分簡單，因為 $1 + \sin\theta$ 永不為負。另一種情況，只要把 $\sin\theta - \cos\theta$ 寫成 $R\sin(\theta - \alpha)$ 的形式，便可化為單一個正弦不等式。

解答

設 $f(\theta) = \dfrac{1 + \sin\theta}{\cos\theta}$。以下一律假設 $0 \le \theta < 2\pi$ 且 $\theta \ne \frac{\pi}{2}, \frac{3\pi}{2}$。

情況 1：$\cos\theta < 0$，即 $\frac{\pi}{2} < \theta < \frac{3\pi}{2}$

由於 $\sin\theta \ge -1$，分子 $1 + \sin\theta$ 最少為 0，而分母為負數。所以 $f(\theta) \le 0 < 1$。

$\left(\frac{\pi}{2}, \frac{3\pi}{2}\right)$ 內的每一個 θ 都滿足該不等式。

情況 2：$\cos\theta > 0$，即 $0 \le \theta < \frac{\pi}{2}$ 或 $\frac{3\pi}{2} < \theta < 2\pi$

乘以正數 $\cos\theta$ 不會改變不等式的方向，所以

$$f(\theta) \le 1 \iff 1 + \sin\theta \le \cos\theta \iff \sin\theta - \cos\theta \le -1.$$

現在 $\sin\theta - \cos\theta = \sqrt{2}\left(\sin\theta\cos\frac{\pi}{4} - \cos\theta\sin\frac{\pi}{4}\right) = \sqrt{2}\sin\left(\theta - \frac{\pi}{4}\right)$，於是條件變成

$$\sin\left(\theta - \tfrac{\pi}{4}\right) \le -\tfrac{1}{\sqrt{2}}.$$

想一想 我們需要考慮 $\phi = \theta - \frac{\pi}{4}$ 的哪些值？

我們只需考慮這個情況所容許的兩個範圍：

- 若 $0 \le \theta < \frac{\pi}{2}$，則 $-\frac{\pi}{4} \le \phi < \frac{\pi}{4}$。在這個區間（interval）上，$\sin\phi$ 是遞增的，且 $\sin\phi \ge \sin(-\frac{\pi}{4}) = -\frac{1}{\sqrt{2}}$，等號只在 $\phi = -\frac{\pi}{4}$ 時成立。所以這裏唯一的解是 $\theta = 0$。（驗算：$f(0) = 1/1 = 1$，等號成立。）
- 若 $\frac{3\pi}{2} < \theta < 2\pi$，則 $\frac{5\pi}{4} < \phi < \frac{7\pi}{4}$。在這個開區間（open interval）上 $\sin\phi < -\frac{1}{\sqrt{2}}$，因為 $\sin\phi$ 在兩個端點都等於 $-\frac{1}{\sqrt{2}}$，而在兩者之間的 $\phi = \frac{3\pi}{2}$ 處降至 -1。所以 $\left(\frac{3\pi}{2}, 2\pi\right)$ 內的每一個 θ 都是解。

綜合結果

$\theta = \frac{3\pi}{2}$ 這個值要剔除，因為該處 $\cos\theta = 0$。綜合各個情況，不等式成立，當且僅當

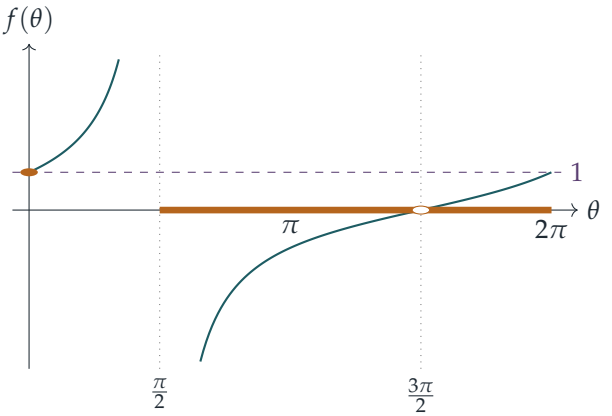
$$\theta = 0 \qquad \text{或} \qquad \frac{\pi}{2} < \theta < \frac{3\pi}{2} \qquad \text{或} \qquad \frac{3\pi}{2} < \theta < 2\pi.$$

更巧妙的看法

利用 $\sin\theta = 2\sin\frac{\theta}{2}\cos\frac{\theta}{2}$、$\cos\theta = \cos^2\frac{\theta}{2} - \sin^2\frac{\theta}{2}$ 和 $1 = \cos^2\frac{\theta}{2} + \sin^2\frac{\theta}{2}$，

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

這裏約去了公因式，而由於 $\cos\theta \ne 0$，該公因式不等於零。（當 $\theta = \pi$ 時，中間的式子沒有定義，但首尾兩式在該處仍然相等。）當 θ 取遍 $[0, 2\pi)$ 時，$\psi = \frac{\theta}{2} + \frac{\pi}{4}$ 取遍 $[\frac{\pi}{4}, \frac{5\pi}{4})$。從 $\tan\psi$ 的圖像一看便知，$\tan\psi \le 1$ 只在 $\psi = \frac{\pi}{4}$ 以及 $\frac{\pi}{2} < \psi < \frac{5\pi}{4}$ 時成立。只要剔除 $\psi = \pi$（即 $\theta = \frac{3\pi}{2}$），這就是同一個答案。



解集（琥珀色）就是曲線位於虛線 $f = 1$ 上或其下方的部分。留意 $\theta = \frac{3\pi}{2}$ 處的空心點：在該處 $f(\theta) \to 0$，但 f 沒有定義。

驗算

- $\theta = 0$：$f(0) = 1$，剛好在邊界上。在略大於 0 之處，分子大於 1 而分母小於 1，所以 $f > 1$。這確認了 $\theta = 0$ 是一個孤立的解。
- $\theta = \frac{7\pi}{4}$：$f = \left(1 - \frac{1}{\sqrt{2}}\right) / \frac{1}{\sqrt{2}} = \sqrt{2} - 1 \approx 0.414 \leq 1$。正確。
- $\theta = \pi$：$f = 1/(-1) = -1 \leq 1$。正確。
- 我們在 $[0, 2\pi)$ 內隨機抽取了 2×10^5 個 θ 作測試，結果與上述解集完全吻合。在 $\theta = \frac{3\pi}{2}$ 附近，我們得到 $f \approx 5 \times 10^{-5}$，與 $f \to 0$ 一致。

常見錯誤

- 沒有分情況便把整條式乘以 $\cos\theta$。這樣會在所有地方都得出 $1 + \sin\theta \leq \cos\theta$，錯誤地捨棄了整個區間 $\left(\frac{\pi}{2}, \frac{3\pi}{2}\right)$。
- 遺漏孤立的解 $\theta = 0$。題目容許 $\theta = 0$，而等號成立也算是解。
- 把 $\theta = \frac{3\pi}{2}$ 也包括在內。該處的式子是 $\frac{0}{0}$，所以沒有定義，儘管它的極限（limit）是 0。

STEP 之外

這裏的教訓是：只有遞增的運算才會保持不等式的方向，例如乘以正數、代入遞增函數（increasing function）等等。時刻留意這一點，正是實分析（real analysis）的日常功夫。$\theta = \frac{3\pi}{2}$ 處的可去奇點（removable singularity），即 f 有極限但沒有函數值的地方，是實分析的另一個主題：連續性（continuity）所講的，就是極限與函數值是否一致。Abbott 的 *Understanding Analysis* 把這兩個概念講解得既嚴謹又易讀。

詳細解答

STEP I 2003，第 5 題

純數學 • 二項式級數

> **STEP I 2003，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/03-s1-q5

關鍵思路

在任何二項展開式（binomial expansion）中，先寫出通項（general term），再追蹤它所帶的 x 的冪次；常數項（constant term）就是該冪次為零的那一項。第 (ii) 部分的展開式並非有限項，所以要先抽出主導因子：當 $|x|$ 很大時，主導因子是 x^6，剩下一個以 $3/x$ 為變量的級數（series）；當 $|x|$ 很小時，主導因子是 $3x^5$，剩下一個以 $x/3$ 為變量的級數。之後只要看清楚 x 的哪些冪次可能出現，問題便迎刃而解。

解答

第 (i) 部分

$(2x + x^{-2})^6$ 的通項是

$$\binom{6}{r}(2x)^{6-r}\left(x^{-2}\right)^r = \binom{6}{r}2^{6-r}x^{6-3r}, \qquad r = 0, 1, \ldots, 6.$$

這一項與 x 無關，當且僅當 $6 - 3r = 0$，即 $r = 2$，此時該項為

$$\binom{6}{2}2^4 = 15 \times 16 = 240,$$

正如題目所求。

同樣地，$(ax^3 + bx^{-2})^{5n}$ 的通項是

$$\binom{5n}{r}\left(ax^3\right)^{5n-r}\left(bx^{-2}\right)^r = \binom{5n}{r}a^{5n-r}b^r\,x^{15n-5r}.$$

當 $15n = 5r$ 時，x 的冪次為零，所以 $r = 3n$（它符合 $0 \le r \le 5n$）。因此與 x 無關的項

是

$$\binom{5n}{3n} a^{2n} b^{3n} = \frac{(5n)!}{(2n)!\,(3n)!}\, a^{2n} b^{3n}.$$

第 (ii) 部分

$|x|$ **很大時**。抽出因子 x^6，得

$$f(x) = \left(x^6(1+3/x)\right)^{1/2} = x^3\left(1+\frac{3}{x}\right)^{1/2},$$

這裏按題目的原意取 $(x^6)^{1/2} = x^3$（嚴格來說應是 $|x|^3$；見下文的說明）。當 $|x| > 3$ 時，$|3/x| < 1$，所以取 $u = 3/x$ 時，$(1+u)^{1/2}$ 的二項級數（binomial series）收斂（converge）：

$$(1+u)^{1/2} = 1 + \tfrac{1}{2}u - \tfrac{1}{8}u^2 + \tfrac{1}{16}u^3 - \tfrac{5}{128}u^4 + \cdots,$$

其中 u^3 的係數是

$$\binom{1/2}{3} = \frac{\frac{1}{2}\cdot\left(-\frac{1}{2}\right)\cdot\left(-\frac{3}{2}\right)}{3!} = \frac{3/8}{6} = \frac{1}{16}.$$

把級數乘以 x^3，項 $\binom{1/2}{k}3^k x^{-k}$ 便變成 $\binom{1/2}{k}3^k x^{3-k}$，所以只有 $k=3$ 會給出常數：

$$f(x) = x^3 + \tfrac{3}{2}x^2 - \tfrac{9}{8}x + \frac{1}{16}\cdot 27 - \frac{405}{128x} + \cdots,$$

而與 x 無關的項是 $\frac{27}{16}$，正如題目所求。

想一想 當 $x < -3$ 時會怎樣？

平方根指的是非負的那個根，而 $(x^6)^{1/2} = |x|^3$，當 $x < 0$ 時它等於 $-x^3$。所以當 $x < -3$ 時，展開式的每一項都會變號，常數項是 $-\frac{27}{16}$。題目所述的結果 $\frac{27}{16}$ 在 $x > 3$ 時完全正確；在 $x < -3$ 時，若把 $(x^6)^{1/2}$ 理解為 x^3，這個結果才成立，而這顯然就是題目的原意。

$|x|$ **很小時**。這時括號內的主導項是 $3x^5$，所以我們改為抽出它：

$$x^6 + 3x^5 = 3x^5\left(1+\frac{x}{3}\right).$$

要 $f(x)$ 是實數，必須有 $x^5(x+3) \geq 0$；由於 $|x| < 3$，因子 $x+3$ 是正數，所以必須有 $x \geq 0$。（當 $-3 < x < 0$ 時，f 根本不能定義為實函數。）當 $0 < x < 3$ 時，

$$f(x) = \sqrt{3}\,x^{5/2}\left(1+\frac{x}{3}\right)^{1/2} = \sqrt{3}\,x^{5/2}\sum_{k\geq 0}\binom{1/2}{k}\frac{x^k}{3^k},$$

而由於 $|x/3| < 1$，這個級數收斂。每一項都是

$$x^{5/2+k}, \qquad k = 0, 1, 2, \ldots,$$

的常數倍，而 $\frac{5}{2}+k$ 永遠不是整數，當然也就永遠不是 0。因此，當 $|x| < 3$ 時，$f(x)$ 按 x 的冪展開的式子中沒有與 x 無關的項。（在 $x = 0$ 這一點，$f(0) = 0$。）

驗算

- 當 x 是很大的正數時，$f(x) - \left(x^3 + \frac{3}{2}x^2 - \frac{9}{8}x\right)$ 應趨近 $\frac{27}{16} = 1.6875$。取 $x = 1000$，計算機給出 1.68434，而級數的下一項 $-\frac{405}{128x} \approx -0.00316$ 正好解釋了這個差距。
- 在 0 附近，$f(x) \approx \sqrt{3}\,x^{5/2}$：取 $x = 0.01$，$f(0.01) = \sqrt{10^{-12} + 3 \times 10^{-10}} \approx 1.734 \times 10^{-5}$，與 $\sqrt{3} \times 10^{-5}$ 相當接近，而且肯定不是趨向一個非零常數。
- 對於一般公式，取 $n = 1$、$a = b = 1$，得 $\binom{5}{3} = 10$，即 $(x^3 + x^{-2})^5$ 的常數項；直接展開便可確認。

常見錯誤

- 寫通項時把冪次的次序調轉了，例如寫成 $(2x)^r(x^{-2})^{6-r}$，然後又把 r 與二項式係數（binomial coefficient）配錯。兩種標記方式都可以，只要前後一致便行。
- 弄錯 $\binom{1/2}{3}$ 的正負號：兩個負因子 $(-\frac{1}{2})(-\frac{3}{2})$ 相乘，結果是正數。
- 在 x 很小時，以抽出 x^6 的方法展開 $(x^6 + 3x^5)^{1/2}$：這樣會得到一個以 $3/x$ 為變量的級數，而它在 $|x| < 3$ 時發散（diverge）。

STEP 之外

f 的兩個展開式在不同的區域內成立，而抽出哪一個因子，就決定了你得到的是哪一個區域。這正是把函數圍繞不同的點展開的想法，它通往複分析（complex analysis）中的洛朗級數（Laurent series）和皮瑟級數（Puiseux series，即按分數冪展開的級數，就像這裏的 $x^{5/2+k}$）。在 $x < -3$ 時引起正負號問題的平方根分支（branch），也是複分析的一個核心課題。Needham 的 *Visual Complex Analysis*（《複分析：可視化方法》）以圖像方式解釋了這兩個概念。

詳細解答

STEP I 2003，第 6 題

純數學 • 以部分分式積分

> **STEP I 2003，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/03-s1-q6

關鍵思路

兩個二次式都可以因式分解，所以最自然的工具就是部分分式（partial fractions）。所謂「不同情況」，出現在兩個一次因式重合的那一刻：這時部分分式失效，我們改為對一個完全平方（perfect square）積分。動手做第 (ii) 部分之前，先仔細看清楚那兩個二次式：只要把變數簡單平移一下，其中一個就變成另一個。

解答

第 (i) 部分

由於 $2 \times a = 2a$ 及 $2 + a = a + 2$，這個二次式可以因式分解：

$$x^2 + (a+2)x + 2a = (x+2)(x+a).$$

因為 $a > 0$，當 $0 \le x \le 1$ 時兩個因式都是正數，所以被積函數（integrand）在整個區間上連續（continuous），而對每一個 $a > 0$，積分都存在。兩個因式相等，當且僅當 $a = 2$，這就是分界線。

情況 $a \neq 2$。兩個因式相異，且

$$\frac{1}{(x+2)(x+a)} = \frac{1}{a-2}\left(\frac{1}{x+2} - \frac{1}{x+a}\right),$$

這可以通分來驗證：分子變成 $(x+a) - (x+2) = a - 2$。由於 $x+2$ 和 $x+a$ 在 $[0,1]$ 上

都是正數，對數（logarithm）裏不需要加絕對值符號：

$$\begin{aligned}\int_0^1 \frac{\mathrm{d}x}{(x+2)(x+a)} &= \frac{1}{a-2}\Big[\ln(x+2)-\ln(x+a)\Big]_0^1\\ &= \frac{1}{a-2}\Big(\ln 3-\ln(1+a)-\ln 2+\ln a\Big)\\ &= \frac{1}{a-2}\ln\left(\frac{3a}{2(a+1)}\right).\end{aligned}$$

想一想 當 $0<a<2$ 時，因子 $\frac{1}{a-2}$ 是負數。被積函數是正的，答案必須是正數；它仍然是正數嗎？

是的。我們有 $\dfrac{3a}{2(a+1)}<1 \iff 3a<2a+2 \iff a<2$，所以當 $0<a<2$ 時，對數也是負數，兩者相除便得正數。同理，當 $a>2$ 時兩者皆為正數。如果你希望每一部分都是正數，可以把 $0<a<2$ 時的答案寫成 $\dfrac{1}{2-a}\ln\left(\dfrac{2(a+1)}{3a}\right)$；這是同一個數。

情況 $a=2$。這時被積函數是 $(x+2)^{-2}$，而

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

第 (ii) 部分

這裏的二次式是 $u^2+au+a-1=(u+1)(u+a-1)$，但在重複上面的工夫之前，先作代換（substitution）$u=x+1$：

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

這正正就是第 (i) 部分的二次式。此外 $\mathrm{d}u=\mathrm{d}x$，而 $u=1,2$ 對應 $x=0,1$。所以

$$\int_1^2 \frac{\mathrm{d}u}{u^2+au+a-1}=\int_0^1 \frac{\mathrm{d}x}{x^2+(a+2)x+2a},$$

答案亦與第 (i) 部分相同：

$$\int_1^2 \frac{\mathrm{d}u}{u^2+au+a-1}=\begin{cases}\dfrac{1}{a-2}\ln\left(\dfrac{3a}{2(a+1)}\right), & a\neq 2,\\[2ex] \dfrac{1}{6}, & a=2.\end{cases}$$

直接計算：在 $[1,2]$ 上有 $u+1>0$ 及 $u+a-1\geq a>0$，所以被積函數同樣連續，而兩個因式重合，當且僅當 $a-1=1$，即 $a=2$。部分分式給出 $\frac{1}{a-2}\big[\ln\frac{u+1}{u+a-1}\big]_1^2=\frac{1}{a-2}\big(\ln\frac{3}{a+1}-\ln\frac{2}{a}\big)$，這是同一個表達式。

驗算

- **在 $a = 2$ 處的連續性**。當 $a \to 2$ 時，對數趨向 $\ln 1 = 0$，所以公式呈 $0/0$ 的形式。它的極限（limit）就是 $\ln 3a - \ln(2a+2)$ 在 $a = 2$ 處的導數（derivative），即 $\frac{1}{2} - \frac{1}{3} = \frac{1}{6}$，與特殊情況的結果吻合。兩種情況順暢地銜接起來，本應如此。
- **一個特殊值**。取 $a = 1$，第 (i) 部分的被積函數是 $\frac{1}{(x+1)(x+2)}$，而公式給出 $-\ln\frac{3}{4} = \ln\frac{4}{3} \approx 0.2877$。
- **數值驗證**。對兩個積分使用辛普森法則（Simpson's rule），取 $a = 1$、$a = 2$、a 接近 0 和 2 的值，以及 200 個不超過 20 的隨機 a 值，結果與公式相符，誤差小於 10^{-10}。

常見錯誤

- 忘記了 $a = 2$ 的情況，這時部分分式的係數 $\frac{1}{a-2}$ 沒有定義。
- 把部分分式的正負號弄反了。通分檢查一下。
- 把 $0 < a < 2$ 和 $a > 2$ 當成真正不同的答案。只要留意正負號，一條公式已涵蓋兩者。

STEP 之外

在 $a = 2$ 處的表現，是可去奇點（removable singularity）的一個小例子：公式在某一點看似沒有定義，但它在該點的極限恰好就是積分的值。在分析學（analysis）中，你會學到一個依賴參數的積分在甚麼條件下對該參數連續（或可微），這就解釋了為何兩種情況必須吻合。Bartle 與 Sherbert 的 *Introduction to Real Analysis*（實分析入門）對極限、連續性和黎曼積分（Riemann integral）有嚴謹的論述。

詳細解答

STEP I 2003，第 7 題

純數學 • 數字與整除性

> **STEP I 2003，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/03-s1-q7

關鍵思路

設 $N = 100a + 10b + c$，並令 $S = N$，便得到一條關於三個數字的方程。把它整理成

$$99a + (10b - b^2) = c^3 - c,$$

題目第一段的兩個事實就可以直接派上用場：含 b 的一項被困在 0 與 25 之間，而 $99a$ 和 $c^3 - c$ 都是 3 的倍數。第一個事實表示 $c^3 - c$ 必定剛好略高於 99 的某個倍數；第二個事實則減少了 b 的可能取值。之後只剩下寥寥幾種情況，用手逐一檢查即可。

解答

兩個預備事實

配方（completing the square）得

$$10k - k^2 = 25 - (5 - k)^2.$$

由於 $(5-k)^2 \geq 0$，可得 $10k - k^2 \leq 25$。由於 $0 \leq k \leq 9$，我們有 $-4 \leq 5 - k \leq 5$，所以 $(5-k)^2 \leq 25$，因此 $10k - k^2 \geq 0$。（當 $k = 0$ 時取等號得 0，當 $k = 5$ 時取等號得 25。）

其次，$k-1$、k、$k+1$ 是三個連續整數（consecutive integers），所以當中恰有一個是 3 的倍數。因此它們的積 $k(k-1)(k+1)$ 可被 3 整除。（這對所有整數 k 都成立，不限於 $0 \leq k \leq 9$。）

找出這些數

設 $N = 100a + 10b + c$，其中各數字滿足 $1 \le a \le 9$（因為 $N \ge 100$）及 $0 \le b, c \le 9$。則 $S = a + b^2 + c^3$，而

$$\begin{aligned} S = N &\iff a + b^2 + c^3 = 100a + 10b + c \\ &\iff 99a + (10b - b^2) = c^3 - c = (c-1)c(c+1). \qquad (*) \end{aligned}$$

想一想 為甚麼要把 $10b - b^2$ 與 $99a$ 放在同一邊？

因為這正是我們剛剛學會控制的組合。設 $t = 10b - b^2$，第一個事實指出 $0 \le t \le 25$，而 $99a$ 是 99 的倍數，遠大於 25。所以 $(*)$ 表示 $c^3 - c$ 等於 99 的某個倍數加上一個細小的「餘數」t，這樣很快就能把情況鎖定。

限制 b。由第二個事實，3 整除 $c^3 - c = (c-1)c(c+1)$，而 3 亦整除 $99a$。所以由 $(*)$，3 整除 $t = b(10 - b)$。由於 $10 - b \equiv 1 - b \pmod 3$，這需要 $b \equiv 0$ 或 $b \equiv 1 \pmod 3$。可能的情況如下：

b	0	1	3	4	6	7	9
$t = 10b - b^2$	0	9	21	24	24	21	9

所以 $t \in \{0, 9, 21, 24\}$。（$b = 2, 5, 8$ 時 $t = 16, 25, 16$，全都不能被 3 整除。）

限制 c。由於 $a \ge 1$ 及 $t \ge 0$，$(*)$ 給出 $c^3 - c \ge 99$。$c^3 - c$ 的值如下：

c	0	1	2	3	4	5	6	7	8	9
$c^3 - c$	0	0	6	24	60	120	210	336	504	720

所以 $c \ge 5$。

各種情況。對每個 $c \ge 5$，我們需要一個倍數 $99a$（其中 $1 \le a \le 9$）落在範圍 $[c^3 - c - 25,\ c^3 - c]$ 之內，然後 $t = c^3 - c - 99a$ 必須是 $0, 9, 21, 24$ 其中之一。

- $c = 5$：$c^3 - c = 120$。在 $[95, 120]$ 內 99 的倍數只有 99，得 $a = 1$ 及 $t = 21$，所以 $b = 3$ 或 $b = 7$。由此得 $N = 135$ 和 $N = 175$。
- $c = 6$：$c^3 - c = 210$。在 $[185, 210]$ 內 99 的倍數只有 198，得 $a = 2$ 及 $t = 12$。但 12 並非 $0, 9, 21, 24$ 其中之一（事實上 $10b - b^2 = 12$ 沒有整數根）。無解。
- $c = 7$：$c^3 - c = 336$。附近的倍數是 297 和 396，兩者都不在 $[311, 336]$ 內。無解。
- $c = 8$：$c^3 - c = 504$。在 $[479, 504]$ 內 99 的倍數只有 495，得 $a = 5$ 及 $t = 9$，所以 $b = 1$ 或 $b = 9$。由此得 $N = 518$ 和 $N = 598$。
- $c = 9$：$c^3 - c = 720$。附近的倍數是 693 和 792，兩者都不在 $[695, 720]$ 內。無解。

以上每一步都可以逆推，所以這四個數確實滿足 $S = N$；直接驗證：

$$135 : 1 + 3^2 + 5^3 = 1 + 9 + 125 = 135, \quad 175 : 1 + 7^2 + 5^3 = 1 + 49 + 125 = 175,$$
$$518 : 5 + 1^2 + 8^3 = 5 + 1 + 512 = 518, \quad 598 : 5 + 9^2 + 8^3 = 5 + 81 + 512 = 598.$$

所以滿足 $S = N$ 的數恰好是

$$N = 135, \quad 175, \quad 518, \quad 598.$$

驗算

對全部 900 個三位數作窮舉搜尋（brute-force search），逐一計算 $a + b^2 + c^3$，找到的恰好是 135、175、518 和 598，與分情況討論的結果一致。還有一個巧妙的結構性檢驗：每一對解之中，b 的兩個值相加都等於 10（$3 + 7$ 和 $1 + 9$）。這是必然的，因為把 b 換成 $10 - b$ 時，$t = 10b - b^2$ 保持不變。

常見錯誤

- 忘記 $a \geq 1$，因而接受了 000 或 001 這類並非三位數的「解」。
- 對給定的 a 和 c 找到一個 b 之後便停下來：由於 $t(b) = t(10 - b)$，解是成對出現的。
- 只用被 3 整除這一性質，而沒有用上界限 $0 \leq t \leq 25$：正是這個界限令每個 c 最多只有一個可能的 a。

STEP 之外

等於自身各數字的冪之和的數，是趣味數學中的經典題材（其中最著名的一類是阿姆斯特朗數（Armstrong numbers），又稱「自戀數」（narcissistic numbers），例如 $153 = 1^3 + 5^3 + 3^3$）。背後真正重要的想法，正是這裏所用的方法：先估計兩邊的大小，使搜尋範圍變成有限，再利用同餘（congruence）來刪減。這種把大小估計與模算術（modular arithmetic）結合起來的手法，是解決許多丟番圖方程（Diophantine equations）的關鍵。Davenport 的 *The Higher Arithmetic*（一本數論入門著作）對同餘及其應用有極佳的介紹。

詳細解答

STEP I 2003，第 8 題

純數學 • 邏輯斯諦微分方程

> **STEP I 2003，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/03-s1-q8

關鍵思路

把「A 轉化為 B 的速率」翻譯成導數（derivative）：它就是 B 的體積增長的速率。由於 $x = 1 - y$，這樣便得到一條只含 y 的可分離變數（separable variables）微分方程（differential equation），再用部分分式（partial fractions）完成積分。最後一部分問的是 $x = 1 - y$ 的公式會否到達零，而答案取決於一個事實：指數函數（exponential）的值永遠不會是無限大。

解答

建立方程

B 的體積是 yV，而 V 是固定的。A 以速率 $kVxy$ 轉化為 B，所以 B 的體積以該速率增加：

$$\frac{\mathrm{d}}{\mathrm{d}t}(yV) = kVxy \quad \Longrightarrow \quad \frac{\mathrm{d}y}{\mathrm{d}t} = kxy = ky(1-y),$$

這裏除以 V 並用了 $x = 1 - y$。（等價地，A 的體積以這個速率減少：$V\,\mathrm{d}x/\mathrm{d}t = -kVxy$，由於 $\mathrm{d}x/\mathrm{d}t = -\mathrm{d}y/\mathrm{d}t$，這說的是同一件事。）

解方程

開始時 $x > 0$ 且 $y > 0$，所以 $0 < y < 1$。只要 $0 < y < 1$，我們便可以除以 $y(1-y)$，把變數分離：

$$\int \frac{\mathrm{d}y}{y(1-y)} = \int k\,\mathrm{d}t.$$

用部分分式，$\dfrac{1}{y(1-y)} = \dfrac{1}{y} + \dfrac{1}{1-y}$，所以

$$\ln y - \ln(1-y) = kt + c,$$
$$\ln \frac{y}{1-y} = kt + c,$$

這裏用到 y 和 $1-y$ 都是正數，所以不需要加絕對值符號。兩邊取指數，並記 $D = \mathrm{e}^c$，

$$\frac{y}{1-y} = D\mathrm{e}^{kt}.$$

於是 $y = D\mathrm{e}^{kt} - yD\mathrm{e}^{kt}$，所以 $y\left(1 + D\mathrm{e}^{kt}\right) = D\mathrm{e}^{kt}$，而

$$y = \frac{D\mathrm{e}^{kt}}{1 + D\mathrm{e}^{kt}},$$

正是所求。代入 $t = 0$ 可得 $D = \dfrac{y_0}{1-y_0} = \dfrac{y_0}{x_0}$，其中 x_0 和 y_0 是初始值；特別地，$D > 0$。

想一想 我們在積分時假設了 $0 < y < 1$。怎樣知道這個條件在任何時刻都一直成立？

假設不然，並設 T 為 y 首次到達 0 或 1 的時刻。在 $0 \le t < T$ 上，以上的推導是成立的，所以 y 在該處由公式給出。由於 y 是連續（continuous）的，$y(T)$ 就是公式在 $t \to T$ 時的極限（limit），即 $D\mathrm{e}^{kT}/(1 + D\mathrm{e}^{kT})$。因為 $D > 0$，這個值嚴格介乎 0 與 1 之間，於是出現矛盾。所以對所有 t 都有 $0 < y < 1$，而公式對所有 t 都成立。

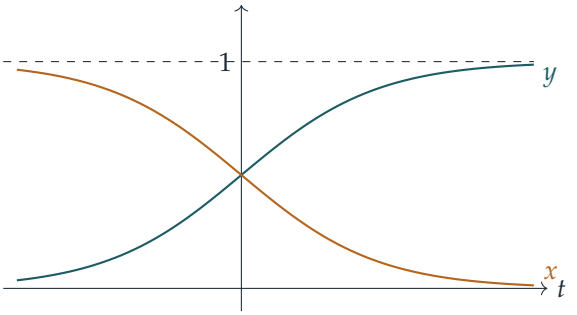
A 會否完全轉化？

不會。由公式，

$$x = 1 - y = \frac{1}{1 + D\mathrm{e}^{kt}},$$

而對每一個 t，$D\mathrm{e}^{kt}$ 都是有限的，所以在任何時刻都有 $x > 0$：A 總會剩下一些。不過，當 $t \to \infty$ 時，$D\mathrm{e}^{kt} \to \infty$（因為 $k > 0$ 且 $D > 0$），所以 $x \to 0$ 而 $y \to 1$。轉化過程愈來愈接近完成，但永遠不會在有限時間內結束。

這在物理上也說得通：反應速率 $kVxy$ 與剩餘的 A 的份量成正比，所以 A 快用完時，反應便會慢下來。



草圖（取 $D = k = 1$）顯示 y 的 S 形邏輯斯諦曲線（logistic curve）趨近 1，以及 x 的曲線趨近 0，兩者都永遠不會到達各自的極限。

驗算

- 對答案求導：設 $u = De^{kt}$，則 $y = u/(1+u)$，而 $\mathrm{d}y/\mathrm{d}t = ku/(1+u)^2 = ky(1-y)$，與預期相符。
- 取 $D = k = 1$（即開始時兩者體積相等），公式給出在 $t = \ln 3$ 時 $y = \frac{3}{4}$。我們亦隨機選取了 200 組 k 和初始值，用數值方法把方程逐步向前推算，結果與公式相差不超過 10^{-12}，而 x 自始至終都保持正值。

常見錯誤

- 把「在極限中 $y = 1$」當成「在某個時刻 $y = 1$」。$t \to \infty$ 時的極限，在任何有限的時刻都不會達到。
- 忘記 $\ln(y/(1-y))$ 這一步需要 $0 < y < 1$，或者沒有說明這個條件為何一直成立。

STEP 之外

方程 $\dot{y} = ky(1-y)$ 就是邏輯斯諦方程（logistic equation），是描述受資源限制的增長的標準模型。這裏 $y = 0$ 和 $y = 1$ 是平衡解（equilibrium solutions），而微分方程的唯一性定理（uniqueness theorem）指出，其他解永遠不能穿越或觸及它們，這是解決最後一部分的一個更簡潔的方法。Boyce 與 DiPrima 的 *Elementary Differential Equations*（初等微分方程教科書）詳細討論了邏輯斯諦方程和這類穩定性分析（stability analysis）。

詳細解答

STEP I 2003，第 9 題

力學 • 通過定點的拋體

STEP I 2003，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：

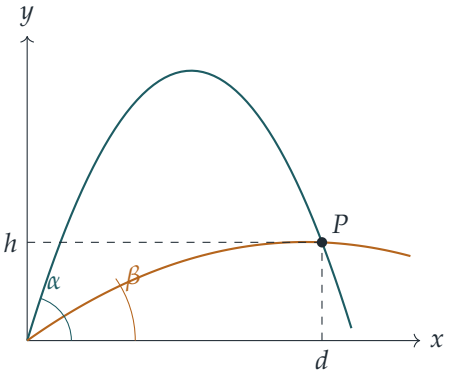
step.maths.org/questions/03-s1-q9

關鍵思路

寫出水平位移和垂直位移，消去時間，再把 $\sec^2\theta$ 換成 $1+\tan^2\theta$：軌跡方程（trajectory equation）就變成 $T = \tan\theta$ 的二次方程。此後一切都圍繞這個二次方程。有兩個角度，即有兩個相異實根，所以要看判別式（discriminant）。而 $\alpha+\beta$ 則透過 $\tan(\alpha+\beta)$ 的公式，由兩根之和與兩根之積決定。

解答

我們把質點（particle）視為只受重力作用而自由運動（不計空氣阻力），以拋射點為原點，x 軸水平指向 P，y 軸鉛垂向上。於是 P 是點 (d,h)，其中 $d>0$。我們取 $h>0$，即 P 在拋射點之上；$h\le 0$ 的情況留待最後再談。



T 的二次方程

水平速度恆為 $V\cos\theta$，而垂直方向的運動有恆定加速度 $-g$，所以在時間 t

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

質點在時間 $t = d/(V\cos\theta)$ 到達 $x = d$（這裏 $\cos\theta > 0$，因為質點朝 P 運動）。把它代入 y，並令 $y = h$，

$$h = d\tan\theta - \frac{gd^2}{2V^2\cos^2\theta} = dT - \frac{gd^2}{2V^2}\big(1+T^2\big),$$

其中用了 $\sec^2\theta = 1 + \tan^2\theta$。現在全式乘以 $\dfrac{2V^2}{gd^2} = \dfrac{2k}{d}$：

$$\frac{2kh}{d} = 2kT - \big(1+T^2\big), \qquad \text{即} \qquad T^2 - 2kT + \frac{2kh}{d} + 1 = 0,$$

這正是所求。

兩個相異的角度

對於給定的 V（因而給定的 k），可行的拋射角（angle of projection）就是那些 $\theta \in (-\frac{\pi}{2}, \frac{\pi}{2})$，其正切是這個二次方程的根。由於 tan 在該區間上嚴格遞增（strictly increasing），相異的實根 T 對應相異的角度。所以我們需要這個二次方程有兩個相異實根，而這恰好在判別式為正時成立：

$$4k^2 - 4\Big(\frac{2kh}{d} + 1\Big) > 0 \iff k^2d^2 - 2h(kd) - d^2 > 0 \iff (kd-h)^2 > h^2 + d^2,$$

這裏先乘以 $d^2 > 0$，再對 kd 配方（completing the square）。

> **想一想** 把不等式兩邊平方有時會出錯，那麼為甚麼由題目給出的條件可以推出 $(kd-h)^2 > h^2 + d^2$？

若 $kd > h + \sqrt{h^2+d^2}$，則 $kd - h > \sqrt{h^2+d^2} > 0$。兩邊皆為正數，所以平方後不等式仍然成立：$(kd-h)^2 > h^2 + d^2$。因此判別式為正，$T$ 有兩個相異的根，從而有兩個相異的拋射角。

（逆命題（converse）同樣成立。要令 $(kd-h)^2 > h^2+d^2$，另一個可能是 $kd < h - \sqrt{h^2+d^2}$，但由於 $\sqrt{h^2+d^2} > |h|$，右邊是負數，而 $kd = V^2/g > 0$。所以當且僅當 $kd > h + \sqrt{h^2+d^2}$ 時，才可以有兩個角度。）

兩個角度都確實在水平線之上。兩根之和為 $2k > 0$，兩根之積為 $1 + 2kh/d > 0$（因為 $h > 0$），所以兩根皆為正數。

兩角之和

設 $\tan\alpha$ 和 $\tan\beta$ 為兩根。由二次方程可得

$$\tan\alpha+\tan\beta=2k,\qquad \tan\alpha\tan\beta=1+\frac{2kh}{d}.$$

兩根之積不等於 1（因為 $kh/d\neq 0$），所以 $\alpha+\beta\neq\frac{\pi}{2}$，複角公式（compound-angle formula）適用：

$$\tan(\alpha+\beta)=\frac{\tan\alpha+\tan\beta}{1-\tan\alpha\tan\beta}=\frac{2k}{-2kh/d}=-\frac{d}{h}.$$

留意 k 已經約去：兩角之和與速率無關。

> **想一想** $\tan(\alpha+\beta)=-d/h$ 有無限多個解。哪一個才是 $\alpha+\beta$？

我們已證明 $\alpha,\beta\in(0,\frac{\pi}{2})$，所以 $0<\alpha+\beta<\pi$。此外 $\tan(\alpha+\beta)<0$，所以 $\alpha+\beta$ 位於 $(\frac{\pi}{2},\pi)$ 之內。在 $(\frac{\pi}{2},\pi)$ 內正切為 $-d/h$ 的角只有一個，就是 $\pi-\arctan(d/h)$，因為 $\arctan(d/h)\in(0,\frac{\pi}{2})$，而且 $\tan(\pi-u)=-\tan u$。因此

$$\alpha+\beta=\pi-\arctan\Big(\frac{d}{h}\Big),$$

這正是所求。

這個結果有一個漂亮的看法。$\arctan(d/h)$ 是 OP 與向上鉛垂線之間的夾角，所以 $\pi-\arctan(d/h)=\frac{\pi}{2}+\gamma$，其中 γ 是 P 的仰角（angle of elevation）。於是 $\frac{1}{2}(\alpha+\beta)=\frac{\pi}{4}+\frac{\gamma}{2}$：兩個拋射方向關於 OP 與鉛垂線夾角的角平分線對稱。

驗算

- 若 $h=0$（目標在同一水平），公式給出 $\alpha+\beta=\pi-\frac{\pi}{2}=\frac{\pi}{2}$：這就是大家熟悉的、射程（range）相同的一對餘角。
- 取 $d=3$、$h=4$、$k=4$。則 $kd=12>4+5$，兩根為 $T=4\pm\sqrt{13/3}$，得出 $\alpha\approx 1.4078$ 和 $\beta\approx 1.0903$。所以 $\alpha+\beta\approx 2.4981$，等於 $\pi-\arctan(0.75)\approx 2.4981$。
- 對數千組隨機的 d、h 和 k，分別以兩個角度對飛行過程作數值模擬，軌跡都通過 (d,h)，並印證了該條件和兩角之和。

常見錯誤

- 忘記 $\sec^2\theta=1+\tan^2\theta$，以致始終得不到 T 的二次方程。
- 寫成 $\alpha+\beta=\arctan(-d/h)=-\arctan(d/h)$，這是負數。反正切（inverse tangent）只會給出 $(-\frac{\pi}{2},\frac{\pi}{2})$ 內的角，所以必須利用 $\alpha+\beta$ 的範圍來選取正確的分支（branch）。

▶ 不加小心地取 $h < 0$（目標在拋射點之下）。這時其中一根可能是負數，而正確的一般結論是 $\alpha + \beta = \frac{\pi}{2} + \arctan(h/d)$。只有當 $h > 0$ 時，它才與題目中的形式一致。

STEP 之外

在臨界情況 $kd = h + \sqrt{h^2 + d^2}$，兩個角度合而為一，這個情況描繪出以速率 V 拋射的所有軌跡的包絡線（envelope）：「安全拋物線」（parabola of safety）$y = \dfrac{V^2}{2g} - \dfrac{gx^2}{2V^2}$。在它之內的點有兩種方法可以擊中，在它之上的點只有一種方法，在它之外的點則完全無法擊中。曲線族的包絡線，可以把一條方程與它對參數的導數聯立求解而得出，在微分方程（differential equations）和光學中還會再次出現。Kleppner 與 Kolenkow 的 *An Introduction to Mechanics*（力學導論）正是以這種嚴謹的態度講解拋體運動和約束運動。

詳細解答

STEP I 2003，第 10 題

力學 • 薄片的質心

> **STEP I 2003，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/03-s1-q10

關鍵思路

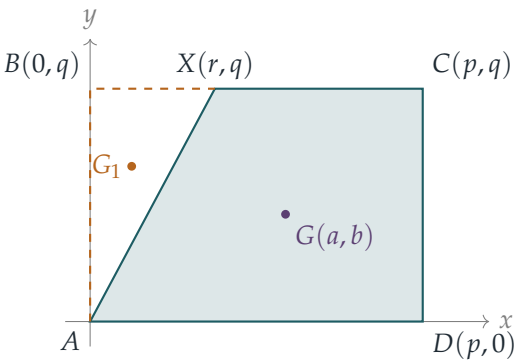
處理切去了一塊的圖形，最好的方法是相減：長方形對每條軸的矩（moment）等於梯形（trapezium）的矩加上三角形的矩，而各部分的質量與面積成正比。至於懸掛的部分，一塊薄片（lamina）從某一點自由懸掛，靜止時其重心（centre of gravity）必在該點的正下方，所以直線 AG 是鉛垂的。這樣，45° 這個條件就化為 a 與 b 之間一個非常簡單的關係。

解答

我們把薄片當作均勻而且很薄，因此任何一部分的質量都與其面積成正比，而三角形的重心位於其形心（centroid），即三個頂點的平均位置。

重心

按題目所給的坐標軸，$A=(0,0)$，$D=(p,0)$，$C=(p,q)$，因此 $B=(0,q)$。由於 X 在 BC 上，且 $BX=r$，所以 $X=(r,q)$，其中 $0\le r\le p$。



三個部分如下。

- 長方形 $ABCD$：面積為 pq，重心為 $(\frac{p}{2}, \frac{q}{2})$。
- 三角形 ABX（被切去的部分）：面積為 $\frac{1}{2}qr$，重心 G_1 是 $(0,0)$、$(0,q)$ 和 (r,q) 的平均位置，即 $(\frac{r}{3}, \frac{2q}{3})$。
- 餘下的梯形 $AXCD$：面積為 $pq - \frac{1}{2}qr = \frac{1}{2}q(2p-r)$，重心為 $G = (a,b)$。

長方形對每條軸的矩，等於它兩個部分的矩之和。對 y-軸取矩（即比較 x-坐標），

$$pq \cdot \frac{p}{2} = \frac{qr}{2} \cdot \frac{r}{3} + \frac{q(2p-r)}{2}\, a.$$

兩邊除以 $\frac{q}{2}$，得 $p^2 = \frac{r^2}{3} + (2p-r)a$；由於 $2p - r \geq p > 0$，我們可以除以它：

$$a = \frac{3p^2 - r^2}{3(2p-r)}.$$

同樣地，對 x-軸取矩，

$$pq \cdot \frac{q}{2} = \frac{qr}{2} \cdot \frac{2q}{3} + \frac{q(2p-r)}{2}\, b,$$

所以 $pq = \frac{2qr}{3} + (2p-r)b$，而

$$b = \frac{q(3p-2r)}{3(2p-r)}.$$

快速驗算一下：$r = 0$（即沒有切去任何部分）時，$a = \frac{p}{2}$，$b = \frac{q}{2}$，正是長方形的中心。

從 A 懸掛

當薄片從 A 自由懸掛並處於平衡（equilibrium）時，只有兩個力作用在它上面：通過 G 的重量，以及在 A 的反作用力。兩個力若互相平衡，必定沿同一直線作用，所以 G 在 A 的正下方。

想一想 「AD 在水平線以下 45°」怎樣變成一條關於 a 和 b 的方程？

由於 AG 是鉛垂的，而 AD 與水平線成 45°，所以 AD 與 AG 亦成 45°；即是說，角 GAD 是 45°。在我們的坐標軸中，AD 是正 x-軸，而 $G = (a, b)$，其中 $a, b > 0$（它位於梯形之內），所以 $\tan\angle GAD = b/a$。因此 $b/a = 1$，即

$$a = b.$$

代入上面的表達式，並約去共同的正分母 $3(2p - r)$，

$$\begin{aligned} & 3p^2 - r^2 = q(3p - 2r) \\ \iff \quad & r^2 - 2qr + 3pq - 3p^2 = 0. \end{aligned}$$

由二次公式（quadratic formula），

$$r = q \pm \sqrt{q^2 - (3pq - 3p^2)} = q \pm \sqrt{q^2 - 3pq + 3p^2}.$$

根號內的量是 $(q - \frac{3}{2}p)^2 + \frac{3}{4}p^2 > 0$，所以兩個根都是實數，而且互不相同。記 $s = \sqrt{q^2 - 3pq + 3p^2} > 0$。

選擇正負號。 由於 X 在邊 BC 上，而這條邊的長度是 p，我們需要 $0 \le r \le p$。現在證明根 $q + s$ 總是大於 p。

- 若 $q \ge p$，則 $q + s > q \ge p$，因為 $s > 0$。
- 若 $q < p$，則 $p - q > 0$，而且

  $$s^2 - (p - q)^2 = (q^2 - 3pq + 3p^2) - (p^2 - 2pq + q^2) = 2p^2 - pq = p(2p - q) > 0,$$

  因為 $q < p < 2p$。所以 $s > p - q$，從而 $q + s > p$。

無論哪種情況，取正號都會令 X 越過 C，落在薄片之外，所以不可能。因此

$$r = q - \sqrt{q^2 - 3pq + 3p^2},$$

這正是所要證明的結果。

為完整起見補充一點：這個根確實位於 $[0, p]$ 之內，當且僅當 $p \le q \le 2p$。事實上，$q - s \ge 0 \iff q^2 \ge s^2 \iff 3p(q - p) \ge 0 \iff q \ge p$，而且（當 $q > p$ 時）$q - s \le p \iff (q - p)^2 \le s^2 \iff p(2p - q) \ge 0 \iff q \le 2p$。所以題目描述的情況只會在這類形狀中出現，這與直覺吻合：未經切割的正方形（$q = p$，$r = 0$）從一角懸掛時，對角線是鉛垂的；而在另一個極端（$q = 2p$，$r = p$），圖形就是直角三角形 ACD，其形心 $(\frac{2p}{3}, \frac{q}{3})$ 同樣滿足 $a = b$。

驗算

- 取 $r = 0$ 便得回長方形的中心 $(\frac{p}{2}, \frac{q}{2})$，而 $r = p$ 則給出 $(\frac{2p}{3}, \frac{q}{3})$，即三角形 ACD 的形心。
- 取 $p = 1$，$q = 1.5$，公式給出 $r = 1.5 - \sqrt{0.75} \approx 0.634$，於是 $a = b \approx 0.634$，所以 AG 確實與 AD 成 45°。另一個根 ≈ 2.37，超過了 $p = 1$。
- 我們亦以一般的多邊形形心公式，用數千組隨機的 p, q, r 核對過重心公式。

常見錯誤

- 把三角形的形心寫成 $(\frac{r}{3}, \frac{q}{3})$：它有兩個頂點位於高度 q，所以形心的高度是 $\frac{2q}{3}$。
- 把三角形的矩加上去而不是減去：梯形是長方形減去三角形。
- 以一句含糊的「r 一定很小」來捨去正號。真正的理由是 $r \le p$，而這需要一段簡短的論證，按 $q \ge p$ 與否分開兩種情況處理。

STEP 之外

用相減法求質心（centre of mass），其實是計算 $\iint x \, \mathrm{d}A \big/ \iint \mathrm{d}A$ 的離散版本；到了大學，你會在二重積分（double integral）中再遇上同一個想法，用來處理密度不均勻的薄片以及立體。懸掛條件則是一個普遍原理的最簡單情況：一個以某一點為支點的剛體（rigid body），當其質心位於支點正下方時處於穩定平衡，此時它的勢能（potential energy）最小。Kleppner 與 Kolenkow 所著的 *An Introduction to Mechanics*（力學入門教科書）對質心和剛體平衡有仔細的論述。

詳細解答

STEP I 2003，第 11 題

力學 • 斜面上的碰撞

STEP I 2003，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/03-s1-q11

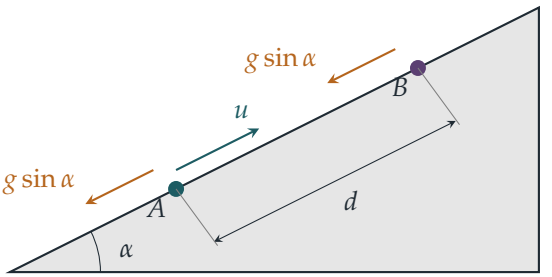
關鍵思路

在光滑（smooth）斜面上，無論速度如何，每個質點的加速度都相同，都是沿最大斜率線（line of greatest slope）向下的 $g\sin\alpha$。因此重力無法改變兩個質點之間的相對運動：從其中一個質點看，另一個質點以恆定速度運動。這樣一來，碰撞時間便很容易求出，而且不必經過繁瑣的運動學計算，就能知道碰撞一刻的相對速度（relative velocity）。碰撞中的能量損失只取決於這個相對速度，這正是 α、g 和 d 全都從答案中消失的原因。

解答

模型

我們把兩個質點都視為點質量，沿着通過 A 和 B 的最大斜率線運動。斜面是光滑的，所以每個質點所受的力只有重量和法向反作用力（normal reaction）。沿垂直於斜面的方向分解，反作用力與 $mg\cos\alpha$ 平衡。沿斜面方向分解，每個質點的加速度都是沿斜面向下的 $g\sin\alpha$。取 x 為由 A 起沿斜面向上的距離。



碰撞時間

設 $a = g\sin\alpha$。以沿斜面向上為正方向，從 A 出發的質點（稱為 P）和從 B 出發的質點（稱為 Q）的位置分別為

$$x_P = ut - \tfrac{1}{2}at^2,$$
$$x_Q = d - \tfrac{1}{2}at^2.$$

兩者之間的距離為

$$x_Q - x_P = d - ut,$$

因為 $\frac{1}{2}at^2$ 項互相抵消。這段距離由 d 開始均勻地縮短，並且恰好在

$$t = \frac{d}{u},$$

時變為零，這正是我們要證明的結果。這次抵消正是關鍵思路的體現：兩個質點以同樣的方式沿斜面下落，所以只有初始相對速率 u 才有影響。

動能損失

在 $t = d/u$ 時，兩個質點沿斜面向上的速度為

$$v_P = u - \frac{ad}{u}, \qquad v_Q = -\frac{ad}{u}.$$

為簡便起見，記 $w = ad/u$。於是 $v_P = u - w$ 及 $v_Q = -w$，而 P 以相對速率 $v_P - v_Q = u$ 接近 Q。

> **想一想** 重力正在作用，為甚麼我們仍可使用動量守恆（conservation of momentum）？

碰撞是瞬時發生的。在極短的碰撞時間內，重力和法向反作用力這些有限的力所產生的衝量（impulse），與質點之間的巨大衝量相比可以忽略不計。因此在碰撞中，沿斜面方向的動量守恆。

設碰撞後沿斜面向上的速度為 v'_P 和 v'_Q。由於兩個質點的質量相等，動量守恆給出

$$v'_P + v'_Q = v_P + v_Q = u - 2w. \tag{1}$$

牛頓實驗定律（Newton's experimental law）指出，分離速率等於接近速率的 e 倍：

$$v'_Q - v'_P = e(v_P - v_Q) = eu. \tag{2}$$

我們不分別解出 v'_P 和 v'_Q，而是利用恆等式

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

把它分別應用於碰撞前和碰撞後，總動能（kinetic energy，簡稱 KE）為

$$\text{碰撞前動能} = \tfrac{1}{2}m\left(v_P^2 + v_Q^2\right) = \tfrac{1}{4}m\left[(u-2w)^2 + u^2\right],$$
$$\text{碰撞後動能} = \tfrac{1}{2}m\left(v_P'^2 + v_Q'^2\right) = \tfrac{1}{4}m\left[(u-2w)^2 + e^2u^2\right],$$

其中碰撞後的和與差分別用了 (1) 和 (2)。兩式中「和」的一項相同，所以

$$\text{動能損失} = \tfrac{1}{4}mu^2 - \tfrac{1}{4}me^2u^2 = \tfrac{1}{4}mu^2\left(1-e^2\right),$$

正如所求。

想一想 $(u-2w)^2$ 這一項在物理上代表甚麼？為甚麼它不會改變？

$v_P + v_Q$ 的一半就是質心（centre of mass）的速度。$\frac{1}{4}m(u-2w)^2$ 這一項是總質量 $2m$ 以該速度運動時的動能，而動量守恆令它保持不變。只有相對於質心的運動的動能才可能損失，而這部分動能只取決於相對速率：碰撞前是 u，碰撞後是 eu。

驗算

- 若 $e = 1$（完全彈性，perfectly elastic），損失為零，理應如此。若 $e = 0$，兩個質點在碰撞後一起運動，損失為 $\frac{1}{4}mu^2$，即相對運動的全部動能。
- 若 $\alpha = 0$，斜面變成水平面，Q 保持靜止，而 P 經過時間 d/u 後以速率 u 到達它那裏，與結果吻合。
- 數值上，取 $\alpha = 30°$、$g = 9.8$、$d = 2$、$u = 4$、$e = \frac{1}{2}$ 及 $m = 1$：在 $t = 0.5$ 時，兩個速度為 1.55 和 -2.45。碰撞後它們變為 -1.45 和 0.55，動能由 4.2025 降至 1.2025，損失為 3，等於 $\frac{1}{4}\cdot 16\cdot\frac{3}{4}$。另外以數值方法逐步模擬運動，隨機測試了 1000 個情況，結果同樣吻合。

常見錯誤

- 把 Q 在碰撞一刻的速度當作零。它已經沿斜面下滑了 d/u 這麼長的時間，正如 P 在這段時間內一直在減速。
- 在牛頓定律中弄錯正負號。應把兩個速度都按同一個正方向寫出，然後用分離速率 $= e \times$ 接近速率。
- 完整解出 v_P' 和 v_Q'，然後展開四個平方。這樣做是可行的，但利用和與差的恆等式可以省去幾乎所有代數運算。

STEP 之外

兩個部分都是轉換到質心參考系（centre-of-mass frame）的例子。均勻重力場令每個質點都有相同的加速度，所以它在相對運動中完全消失。此外，任何系統的動能都可以分拆為質心的動能，加上相對於質心的動能（柯尼希定理，König's theorem）。在二體問題中，第二部分是 $\frac{1}{2}\mu v_{\text{rel}}^2$，其中折合質量（reduced mass）為 $\mu = m_1 m_2/(m_1 + m_2)$，在這裏是 $\frac{1}{2}m$。Kleppner 與 Kolenkow 所著的 *An Introduction to Mechanics*（力學入門教科書）對這兩個想法都有仔細的闡述。

詳細解答

STEP I 2003，第 12 題

概率與統計 • 乘積的期望值

> **STEP I 2003，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/03-s1-q12

關鍵思路

期望值 (expected value) 就是對所有等可能的結果，把 (數值 × 概率) 加起來，所以兩部分都歸結為列出所有有序數對（ordered pairs），再把它們的乘積相加。二重和 $\sum_r \sum_s rs$ 可以分解為 $\left(\sum r\right)^2$，第 (i) 部分要做的就只有這一步。第 (ii) 部分唯一的分別，是 $r = s$ 的數對不再可能出現，所以我們取同一個二重和，再除去它的對角線項。

解答

設 X 為第一個球上的數字，Y 為第二個球上的數字。

第 (i) 部分

由於第一個球會放回袋中，第二次抽取時袋中仍是全部的球，結果與第一次無關。所以對任意 $r, s \in \{1, \ldots, n\}$，

$$\mathrm{P}(X = r,\ Y = s) = \mathrm{P}(X = r)\,\mathrm{P}(Y = s) = \frac{1}{n} \cdot \frac{1}{n} = \frac{1}{n^2}.$$

這樣的有序數對 (r, s) 共有 n^2 個，每個給出乘積 rs，所以按期望值的定義，

$$\mathrm{E}(XY) = \sum_{r=1}^{n} \sum_{s=1}^{n} rs\,\mathrm{P}(X = r,\ Y = s) = \frac{1}{n^2} \sum_{r=1}^{n} \sum_{s=1}^{n} rs,$$

這正是題目要求我們解釋的式子。

要化簡這個式子，留意 r 與 s 無關，所以可以把它提到內層求和之外：

$$\sum_{r=1}^{n} \sum_{s=1}^{n} rs = \sum_{r=1}^{n} r\left(\sum_{s=1}^{n} s\right) = \left(\sum_{r=1}^{n} r\right)^2 = \tfrac{1}{4} n^2 (n+1)^2.$$

因此

$$\mathrm{E}(XY) = \frac{(n+1)^2}{4}.$$

（等價地說，X 與 Y 獨立（independent），所以 $\mathrm{E}(XY) = \mathrm{E}(X)\,\mathrm{E}(Y) = \left(\frac{n+1}{2}\right)^2$。）

第 (ii) 部分

要抽出兩個球，須有 $n \geq 2$。現在第一個球不放回，所以可能的結果是滿足 $r \neq s$ 的有序數對 (r,s)。對這樣的一個數對，

$$\mathrm{P}(X=r,\ Y=s) = \mathrm{P}(X=r)\,\mathrm{P}(Y=s \mid X=r) = \frac{1}{n}\cdot\frac{1}{n-1},$$

因為取走 r 號球之後，第二個球是餘下 $n-1$ 個球中任何一個的可能性均等。這樣的數對共有 $n(n-1)$ 個，全部等可能，而 $r=s$ 的數對概率為 0。因此

$$\mathrm{E}(XY) = \frac{1}{n(n-1)}\sum_{r\neq s} rs.$$

想一想 怎樣只對 $r \neq s$ 的數對把 rs 加起來？

先取對所有數對的和（已在第 (i) 部分求出），再減去 $r=s$ 的對角線項，這些項各自貢獻 $r\cdot r = r^2$：

$$\sum_{r\neq s} rs = \sum_{r=1}^{n}\sum_{s=1}^{n} rs - \sum_{r=1}^{n} r^2 = \tfrac{1}{4}n^2(n+1)^2 - \tfrac{1}{6}n(n+1)(2n+1).$$

提出公因式 $\frac{1}{12}n(n+1)$，

$$\begin{aligned}\sum_{r\neq s} rs &= \tfrac{1}{12}n(n+1)\left[3n(n+1) - 2(2n+1)\right]\\ &= \tfrac{1}{12}n(n+1)\left(3n^2 - n - 2\right)\\ &= \tfrac{1}{12}n(n+1)(n-1)(3n+2).\end{aligned}$$

除以 $n(n-1)$（因為 $n \geq 2$，它不等於零），

$$\mathrm{E}(XY) = \frac{(n+1)(3n+2)}{12},$$

這就是所求的結果。

驗算

取 $n=3$。有放回時，九個乘積的平均值是 $\frac{1}{9}(1+2+3)^2 = 4$，而 $\frac{(3+1)^2}{4} = 4$。無放回時，無序數對 $\{1,2\},\{1,3\},\{2,3\}$ 給出乘積 $2,3,6$，平均值為 $\frac{11}{3}$，而 $\frac{4\cdot 11}{12} = \frac{11}{3}$。當 $n=2$

時，無放回的乘積只有 2 一個，而公式給出 $\frac{3\cdot 8}{12} = 2$。我們亦以逐一列舉的方式，對 12 或以下的每個 n 準確驗證了兩條公式。

比較兩個答案也是一個很好的合理性檢查：

$$\frac{(n+1)(3n+2)}{12} - \frac{(n+1)^2}{4} = -\frac{n+1}{12} < 0.$$

無放回時答案較小，這是合理的：我們失去了 $n \cdot n$ 這類數值大的對角線乘積。

常見錯誤

- 在第 (ii) 部分除以 n^2。可能的有序數對只有 $n(n-1)$ 個，每個的概率是 $\frac{1}{n(n-1)}$。
- 求和時用無序數對，計數時卻用有序數對（或者反過來）。對無序數對求和，會令總和與個數同時減半，所以比值不變；但把兩者混用，結果就會相差一個因子 2。
- 在第 (ii) 部分寫 $\mathrm{E}(XY) = \mathrm{E}(X)\,\mathrm{E}(Y)$。這需要獨立性，而球不放回時 X 與 Y 並不獨立。

STEP 之外

我們求得的差正好就是協方差（covariance）：$\mathrm{Cov}(X,Y) = \mathrm{E}(XY) - \mathrm{E}(X)\,\mathrm{E}(Y) = -\frac{n+1}{12}$，它是負數，因為第一個數字大，第二個數字也大的可能性就會稍低。在無放回抽樣（sampling without replacement）中求和的方差（variance）時，需要的正是這類協方差，由此可以得出超幾何分佈（hypergeometric distribution）的方差，以及抽樣調查中使用的「有限總體校正」（finite population correction）。Blitzstein 與 Hwang 的 *Introduction to Probability*（概率論入門教材）正是以這種方式發展協方差與指示變量（indicator）方法。

詳細解答

STEP I 2003，第 13 題

概率與統計 • 幾何分佈的期望值

STEP I 2003，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/03-s1-q13

關鍵思路

第一部分的做法，是數每隊射失的次數，而不是入球數：兩隊各射五球之後，比數相同的充要條件是兩隊射失的次數相同。至於期望值（expectation），留意突然死亡（sudden death）階段是以每回合兩球的方式進行的，而每一回合都是一次獨立試驗，以相同的概率結束這場互射十二碼。因此回合數服從幾何分佈（geometric distribution），而射球次數的期望值可以拆成「必定要射的十球」加上「額外的射球，並按需要加射的機會加權」。

解答

我們按照模型的原意，假設每一次射球都與其他各次射球獨立（independent），並且 A 隊每次射入的概率恆為 p_A，B 隊則恆為 p_B。記 $q_A = 1 - p_A$ 及 $q_B = 1 - p_B$。

射完十球後打和的概率 α

設 X 為 A 隊在首五球中射失的球數，Y 為 B 隊射失的球數。每隊的五次射球都是射失概率相同的獨立試驗，所以

$$X \sim \mathrm{B}(5, q_A), \qquad Y \sim \mathrm{B}(5, q_B),$$

而且由於兩隊的射球互相獨立，X 與 Y 亦互相獨立。

A 隊射入 $5 - X$ 球，B 隊射入 $5 - Y$ 球，所以比數相同的充要條件是 $X = Y$。對 $i =$

$0,1,\ldots,5$，事件 $\{X=Y=i\}$ 互不相交（disjoint），所以

$$\begin{aligned}\alpha = \mathrm{P}(X=Y) &= \sum_{i=0}^{5} \mathrm{P}(X=i)\,\mathrm{P}(Y=i)\\ &= \sum_{i=0}^{5} \binom{5}{i} q_A^i p_A^{5-i} \binom{5}{i} q_B^i p_B^{5-i}\\ &= \sum_{i=0}^{5} \binom{5}{i}^2 (1-p_A)^i (1-p_B)^i\, p_A^{5-i} p_B^{5-i},\end{aligned}$$

這就是所求的表達式。（若改為數入球數，即兩隊各入 $5-i$ 球，會得到相同的和，因為 $\binom{5}{i} = \binom{5}{5-i}$。）

射球次數的期望值

如果射完十球後比數仍然相同，突然死亡階段便以回合進行：每一回合中 A 隊射一球，B 隊射一球。若兩球之中恰好只有一球射入，該回合便分出勝負，而這個情況出現的概率是

$$p_A q_B + q_A p_B = p_A(1-p_B) + p_B(1-p_A) = p_A + p_B - 2p_A p_B = \beta.$$

否則（兩隊都射入或都射失）便要再進行一個回合。我們需要 $\beta > 0$，也就是說，我們排除 $p_A = p_B = 0$ 和 $p_A = p_B = 1$ 這兩個平凡情況，因為在這兩個情況下，任何回合都永遠無法分出勝負。

> ? **想一想** 為甚麼互射十二碼不可以在一個回合的中途結束？

規則訂明，當一隊射入而另一隊未能射入時，兩隊的射球次數必須相同。即使 A 隊射失，B 隊仍然要射自己的一球（B 隊也可能射失）。所以突然死亡階段的射球總是成對出現，射球總數是 $10+2R$，其中 R 是所進行的回合數（若比賽在首十球內已分出勝負，則 $R=0$）。

在已知進入突然死亡階段的條件下，各回合是獨立試驗，每個回合分出勝負的概率都是 β，所以 R 服從幾何分佈：

$$\mathrm{P}(R=k \mid \text{十球後打和}) = (1-\beta)^{k-1}\beta, \qquad k=1,2,3,\ldots$$

它的平均值是

$$\mathrm{E}(R \mid \text{十球後打和}) = \sum_{k=1}^{\infty} k\beta(1-\beta)^{k-1} = \beta \cdot \frac{1}{\big(1-(1-\beta)\big)^2} = \frac{1}{\beta},$$

這裏用到當 $|x|<1$ 時 $\sum_{k\geq 1} kx^{k-1} = (1-x)^{-2}$，而由於 $0<\beta\leq 1$，這條公式適用。

如果比賽在首十球內已分出勝負，則 $R=0$。按射完十球後比數是否相同來取條件，

$$\mathrm{E}(R) = \alpha \cdot \frac{1}{\beta} + (1-\alpha)\cdot 0 = \frac{\alpha}{\beta}.$$

因此射球次數的期望值是

$$\mathrm{E}(10+2R) = 10 + \frac{2\alpha}{\beta},$$

正如所求。

驗算

- 兩隊實力相同，$p_A = p_B = \frac{1}{2}$：由范德蒙恆等式（Vandermonde's identity）$\sum_i \binom{5}{i}^2 = \binom{10}{5} = 252$，所以 $\alpha = 252/2^{10} = \frac{63}{256}$，而 $\beta = \frac{1}{2}$。射球次數的期望值是 $10 + \frac{63}{64} = \frac{703}{64} \approx 10.98$。
- 如果一隊百發百中而另一隊一球不入（$p_A = 1$，$p_B = 0$），則 $\alpha = 0$（只有 $i = 0$ 的一項有可能不為零，但它含有 $p_B^5 = 0$），所以恰好射 10 球，與常理相符。
- 我們把 α 的公式與直接列舉首十球全部 2^{10} 個結果所得的答案作了比對，又把 $10 + 2\alpha/\beta$ 與模擬 2×10^5 場互射十二碼的結果作了比對；當 $p_A = \frac{3}{4}$、$p_B = \frac{2}{3}$ 時，公式給出 11.264，模擬則給出 11.25。

常見錯誤

- 把 β 當成 A 隊贏得一個回合的概率 $p_A(1 - p_B)$：其實任何一隊贏得一個回合，互射十二碼都會結束。
- 把答案寫成 $10 + 2/\beta$，忘記了突然死亡階段只以概率 α 出現。
- 把突然死亡階段的每一步算作一球，而不是兩球。

STEP 之外

$\mathrm{E}(R) = \alpha\,\mathrm{E}(R \mid \text{打和}) + (1-\alpha)\,\mathrm{E}(R \mid \text{未打和})$ 這一步就是全期望定律（law of total expectation），它是大學一年級概率論中最有用的工具之一。突然死亡階段也是一條帶有吸收態（absorbing state）的簡單馬可夫鏈（Markov chain），而這裏「到達吸收的期望時間」的計算，可以推廣到隨機漫步（random walk）和賭徒破產（gambler's ruin）問題。Grimmett 與 Welsh 合著的 *Probability: An Introduction*（概率論入門）對條件期望（conditional expectation）和這類馬可夫鏈都有論述。

詳細解答

STEP I 2003，第 14 題

概率與統計 • 泊松分佈

STEP I 2003，第 14 題

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

step.maths.org/questions/03-s1-q14

關鍵思路

Jane 見到的朋友數目並不是來訪的人數，而是這個人數以 2 為上限後的值。一個設有上限的變量只需考慮三種情況：無人來訪、恰好一人來訪、兩人或以上來訪。因此我們根本不必把整個泊松（Poisson）級數加起來，只需求出 $P(0)$ 和 $P(1)$，而「兩人或以上」的概率就是餘下的部分。至於新朋友，我們按其是否來訪作條件分析（conditioning）。

解答

原有的朋友

設 N 為某一天來訪的朋友數目，則 $N \sim \text{Po}(2)$，且

$$P(N=k) = \frac{\mathrm{e}^{-2}2^k}{k!}, \qquad k = 0,1,2,\ldots$$

Jane 見到 $S = \min(N,2)$ 位朋友。所以當 $N=0$ 時 $S=0$，當 $N=1$ 時 $S=1$，當 $N \geq 2$ 時 $S=2$。現在

$$P(N=0) = \mathrm{e}^{-2}, \qquad P(N=1) = 2\mathrm{e}^{-2}, \qquad P(N\geq 2) = 1-3\mathrm{e}^{-2},$$

因此

$$\mathrm{E}(S) = 0\cdot\mathrm{e}^{-2} + 1\cdot 2\mathrm{e}^{-2} + 2\left(1-3\mathrm{e}^{-2}\right) = 2-4\mathrm{e}^{-2},$$

這正是所要證明的結果。

同一個計算還有一個更簡潔的寫法，下文會再次用到。對任何取值於 $\{0,1,2,\ldots\}$ 的隨機變量（random variable）X，

$$\mathrm{E}\big(\min(X,2)\big) = 2 - 2P(X=0) - P(X=1),$$

因為 $\min(X,2)$ 與 2 的差距，在 $X=0$ 時是 2，在 $X=1$ 時是 1，其餘情況則為零。取 $X=N$，便再次得到 $2-2\mathrm{e}^{-2}-2\mathrm{e}^{-2}$。

加入新朋友之後

若新朋友來訪，設 $B=1$，否則設 $B=0$，於是 $P(B=1)=p$。題目指出 B 與 N 獨立（independent）。現在來訪的總人數是 $M=N+B$，而 Jane 見到 $\min(M,2)$ 位朋友。

> **想一想** 我們需要 M 的完整分佈嗎？

不需要。根據上面的恆等式，我們只需要 $P(M=0)$ 和 $P(M=1)$。利用獨立性把概率相乘：

$$\begin{aligned}P(M=0)&=P(N=0)\,P(B=0)=(1-p)\mathrm{e}^{-2},\\ P(M=1)&=P(N=1)\,P(B=0)+P(N=0)\,P(B=1)=2(1-p)\mathrm{e}^{-2}+p\,\mathrm{e}^{-2}.\end{aligned}$$

因此

$$\begin{aligned}\mathrm{E}\big(\min(M,2)\big)&=2-2(1-p)\mathrm{e}^{-2}-2(1-p)\mathrm{e}^{-2}-p\,\mathrm{e}^{-2}\\ &=2-(4-3p)\mathrm{e}^{-2}.\end{aligned}$$

另一個做法是以 B 為條件。若新朋友沒有來訪（概率為 $1-p$），平均值一如之前，是 $2-4\mathrm{e}^{-2}$。若新朋友來訪（概率為 p），Jane 見到 $\min(N+1,2)$ 位朋友，這個數在 $N=0$ 時是 1，其餘情況是 2，其平均值為 $\mathrm{e}^{-2}+2(1-\mathrm{e}^{-2})=2-\mathrm{e}^{-2}$。所以平均值是

$$(1-p)(2-4\mathrm{e}^{-2})+p(2-\mathrm{e}^{-2})=2-(4-3p)\mathrm{e}^{-2}.$$

所以 Jane 現在平均每天見到 $2-(4-3p)\mathrm{e}^{-2}$ 位朋友，比之前多了 $3p\,\mathrm{e}^{-2}$。

驗算

- 當 $p=0$ 時，新朋友永不來訪，我們重新得到 $2-4\mathrm{e}^{-2}\approx 1.4587$。
- 當 $p=1$ 時，我們得到 $2-\mathrm{e}^{-2}\approx 1.8647$，這正是 $\min(N+1,2)$ 的平均值，與預期相符。
- 答案隨 p 遞增，而且恆小於 2，這是合理的。
- 按精確分佈對 $\min(k+b,2)$ 求和，以及取 $p=0.3$ 模擬 4×10^5 天，兩者都得出約 1.580，與 $2-3.1\mathrm{e}^{-2}$ 吻合。

常見錯誤

- 忘記了最多只見兩位朋友的上限，答成 $\mathrm{E}(N) = 2$ 或 $2 + p$。
- 把 $N + B$ 當作泊松變量。伯努利（Bernoulli）變量並非泊松變量，所以兩者之和也不是泊松變量。
- 恰好一人來訪有兩種情況（一位舊朋友來訪而新朋友沒來，或沒有舊朋友來訪而新朋友來了），卻漏掉其中一種。

STEP 之外

恆等式 $\mathrm{E}(\min(X,2)) = P(X \geq 1) + P(X \geq 2)$ 是尾和公式（tail-sum formula）$\mathrm{E}(X) = \sum_{k\geq 1} P(X \geq k)$ 的特殊情況，該公式適用於取非負整數值的隨機變量。它是對生存函數（survival function）積分的離散版本。這類設有上限的平均值常見於排隊模型和存貨模型，在容量有限的情況下，它們給出獲得服務的顧客數目的期望值（expected value）。Grimmett 與 Welsh 合著的 *Probability: An Introduction*（概率論入門教材）涵蓋尾和公式和條件期望（conditional expectation），正是本題所用的兩件工具。

詳細解答

STEP II 2003，第 1 題

純數學 • 聯立線性方程

STEP II 2003，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/03-s2-q1

關鍵思路

未知數 x 只以 ax 的形式出現，而 y 在每條方程中的係數都是 -1。因此，把相鄰的方程相減，就可以完全消去 y，而且所得的兩個差式左邊相同。這樣一來，方程組只在 b 取某一個值時才相容（consistent）；一旦相容，就有一個自由參數（free parameter）。題目餘下的部分，就是看怎樣好好選取這個參數。

解答

把三條方程編號為 (1)、(2)、(3)。從 (2) 減去 (1)，再從 (3) 減去 (2)，得

$$ax - 2z = 4, \tag{A}$$
$$ax - 2z = b - 7. \tag{B}$$

方程組的任何解都同時滿足 (A) 和 (B)，所以只有當 $4 = b - 7$，即 $b = 11$ 時，才可能有解。這一點對 a 的每一個值都成立。

第 (i) 部分

當 $a = 0$ 時，方程變成 $-y - z = 3$、$-y - 3z = 7$、$-y - 5z = b$，x 完全沒有出現。

若方程組有解，則由上面的說明可知 $b = 11$。

反過來，假設 $b = 11$。此時方程 (A) 變成 $-2z = 4$，所以 $z = -2$，再由 (1) 得 $y = -3 - z = -1$。以 $y = -1$、$z = -2$ 驗算全部三條方程：

$$1 + 2 = 3, \qquad 1 + 6 = 7, \qquad 1 + 10 = 11 = b,$$

所以（不論 x 取甚麼值）我們都得到一個解。因此，方程組有解若且唯若（if and only if）$b = 11$。

第 (ii) 部分

現在 $a \neq 0$ 而 $b = 11$。設 $z = \lambda$。由方程 (A) 得 $ax = 4 + 2\lambda$，而由於 $a \neq 0$，我們可以兩邊相除：

$$x = \frac{4 + 2\lambda}{a}.$$

然後由 (1) 得 $y = ax - z - 3 = 4 + 2\lambda - \lambda - 3 = 1 + \lambda$。

我們只用了 (1) 和 (2) 來求出這些值，所以還必須驗證它們也滿足 (3)：

$$3ax - y - 5z = 3(4 + 2\lambda) - (1 + \lambda) - 5\lambda = 12 + 6\lambda - 1 - 6\lambda = 11 = b.$$

所以對每一個數 λ，

$$x = \frac{4 + 2\lambda}{a}, \qquad y = 1 + \lambda, \qquad z = \lambda$$

都是一個滿足 $z = \lambda$ 的解，正如題目所求。（事實上，這些就是全部的解：任何解的 z 都有某個值 λ，而上面的推導隨即確定了 x 和 y。）

想一想 為甚麼我們要另外驗證方程 (3)？

我們用了 (1) 和 (2)（以 (1) 和 (A) 的形式）來求 x 和 y。三條含三個未知數的方程不一定會自動互相吻合，而 (3) 正是 b 出場的地方。$b = 11$ 就是在這一步驗證中用上的。

第 (iii) 部分

當 $a = 2$ 時，由第 (ii) 部分可知所有解都是 $(x, y, z) = (2 + \lambda,\ 1 + \lambda,\ \lambda)$。於是

$$\begin{aligned} x^2 + y^2 + z^2 &= (2 + \lambda)^2 + (1 + \lambda)^2 + \lambda^2 \\ &= 3\lambda^2 + 6\lambda + 5 \\ &= 3(\lambda + 1)^2 + 2, \end{aligned}$$

它恰好在 $\lambda = -1$ 時取得最小值 2。所以所求的解是

$$x = 1, \qquad y = 0, \qquad z = -1.$$

從幾何角度看，所有解構成一條方向為 $(1, 1, 1)$ 的直線，而我們要找的是直線上最接近原點的一點。事實上 $(1, 0, -1) \cdot (1, 1, 1) = 0$，所以位置向量（position vector）與直線垂直（perpendicular），正如所料。

第 (iv) 部分

根據開首的說明，b 必須是 11。若 $a = 0$，由第 (i) 部分必有 $y = -1$、$z = -2$，於是 $y^2 + z^2 = 5$，太大了。所以 $a \neq 0$，而由第 (ii) 部分，解為 $x = (4 + 2\lambda)/a$、$y = 1 + \lambda$、$z = \lambda$。

> **想一想** x 怎樣才可以非常大，而 y 和 z 卻保持很小？

a 唯一出現的地方是 $x = (4 + 2\lambda)/a$。選定 λ 就確定了 y 和 z，然後只要 a 足夠小，x 就可以要多大有多大。

我們需要 $y^2 + z^2 = (1 + \lambda)^2 + \lambda^2 = 2\lambda^2 + 2\lambda + 1 < 1$，即 $2\lambda(\lambda + 1) < 0$，即 $-1 < \lambda < 0$。取 $\lambda = -\frac{1}{2}$：此時 $y = \frac{1}{2}$、$z = -\frac{1}{2}$ 而 $y^2 + z^2 = \frac{1}{2} < 1$，同時 $x = 3/a$。選取

$$a = 10^{-6}$$

便得 $x = 3 \times 10^6 > 10^6$。所以 $a = 10^{-6}$（連同 $b = 11$）符合要求。

更一般地說，當 $-1 < \lambda < 0$ 時，我們有 $2 < 4 + 2\lambda < 4$，所以恰好在 $0 < a < 4 \times 10^{-6}$ 時才存在合適的解。負的 a 不可行，因為那時 $x < 0$。

驗算

- 把 $a = 2$、$(x, y, z) = (1, 0, -1)$ 代入原來的方程：$2 - 0 + 1 = 3$、$4 - 0 + 3 = 7$、$6 - 0 + 5 = 11$。
- 對於第 (iv) 部分，取 $a = 10^{-6}$ 和 $(x, y, z) = (3 \times 10^6, \frac{1}{2}, -\frac{1}{2})$：$3 - \frac{1}{2} + \frac{1}{2} = 3$、$6 - \frac{1}{2} + \frac{3}{2} = 7$、$9 - \frac{1}{2} + \frac{5}{2} = 11$。
- 我們用精確分數驗證過，對數百組隨機選取的 a 和 λ，第 (ii) 部分的公式都滿足全部三條方程；另外，對 λ 作網格搜尋，也確認了第 (iii) 部分的最小值 2 在 $\lambda = -1$ 時取得。

常見錯誤

- 由其中兩條方程求出 x 和 y 後，忘記驗證第三條。
- 在第 (iv) 部分，以為 y 和 z「不可能同時很小」，就斷定題目所求不可能做到。其實是可以的：$y - z = 1$ 只能推出 $y^2 + z^2 \geq \frac{1}{2}$。

STEP 之外

這是對線性代數（linear algebra）中秩（rank）與相容性的初步認識。這裏的係數矩陣（coefficient matrix）是奇異（singular）的（它的各行滿足 $r_1 - 2r_2 + r_3 = 0$），所以方程組要麼無解，要麼有一整條直線的解，視乎右邊的常數項是否滿足同一個關係。第 (iv) 部分也說明了為甚麼接近奇異的方程組在數值計算中很危險：一個極小的係數 a 會產生一個極大的解。Strang 的 *Introduction to Linear Algebra*（線性代數入門教材）涵蓋了以上所有內容，包括像第 (iii) 部分那樣的最小範數解（least-norm solution）。

詳細解答

STEP II 2003，第 2 題

純數學 • 反三角函數

> **STEP II 2003，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/03-s2-q2

關鍵思路

$a\cos\theta + b\sin\theta$ 這類式子，總可以寫成單一個餘弦 $R\cos(\theta - \alpha)$。這樣一來，方程就變成一個關於 arccos 的命題，而 α 本身則是一個 arctan。因此，只要找到一個漂亮的解 θ，關於 π 的恆等式（identity）不過是把「$\theta = (\theta - \alpha) + \alpha$」改寫一下而已。唯一真正要小心的地方，是驗證每個角都落在所用反函數（inverse function）的值域（range）之內。

解答

第一個恆等式

先試 $(\frac{1}{4}\pi, \frac{1}{2}\pi)$ 內最容易想到的角。取 $\theta = \frac{1}{3}\pi$，

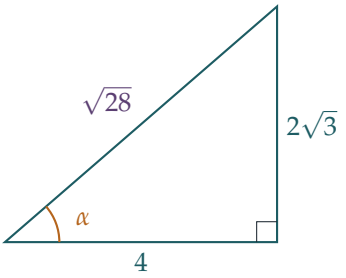
$$4\cos\tfrac{\pi}{3} + 2\sqrt{3}\,\sin\tfrac{\pi}{3} = 4\cdot\tfrac{1}{2} + 2\sqrt{3}\cdot\tfrac{\sqrt{3}}{2} = 2 + 3 = 5,$$

所以 $\theta = \frac{1}{3}\pi$ 滿足方程，而且 $\frac{1}{4}\pi < \frac{1}{3}\pi < \frac{1}{2}\pi$。

現在把左方寫成 $R\cos(\theta - \alpha) = R\cos\theta\cos\alpha + R\sin\theta\sin\alpha$ 的形式。比較係數，

$$R\cos\alpha = 4, \qquad R\sin\alpha = 2\sqrt{3},$$

所以 $R^2 = 16 + 12 = 28$，得 $R = \sqrt{28}$，而 $\tan\alpha = \frac{2\sqrt{3}}{4} = \frac{\sqrt{3}}{2}$。由於 $\cos\alpha$ 和 $\sin\alpha$ 皆為正數，我們可以取 α 為銳角，所以 $\alpha = \arctan(\sqrt{3}/2)$。



於是方程變成

$$\sqrt{28}\cos(\theta-\alpha)=5,\qquad \text{即}\qquad \cos(\theta-\alpha)=\frac{5}{\sqrt{28}},$$

而 $\theta=\frac{1}{3}\pi$ 滿足此方程。

> **想一想** 我們可以立即斷定 $\frac{1}{3}\pi-\alpha=\arccos(5/\sqrt{28})$ 嗎？

還不可以：$\cos x=c$ 有很多個解，而 $\arccos c$ 只是其中位於 $[0,\pi]$ 的那一個。所以我們必須驗證 $\frac{1}{3}\pi-\alpha$ 位於 $[0,\pi]$ 之內。由於 tan 在 $(0,\frac{1}{2}\pi)$ 上是遞增的，而 $\tan\alpha=\frac{\sqrt{3}}{2}<\sqrt{3}=\tan\frac{\pi}{3}$，故有 $0<\alpha<\frac{1}{3}\pi$。因此

$$0<\tfrac{1}{3}\pi-\alpha<\tfrac{1}{3}\pi\le\pi,$$

所以 $\frac{1}{3}\pi-\alpha=\arccos(5/\sqrt{28})$。移項後乘以 3，

$$\pi=3\arccos(5/\sqrt{28})+3\alpha=3\arccos(5/\sqrt{28})+3\arctan(\sqrt{3}/2),$$

這正是所要證明的結果。

第二個恆等式

沿用同一個想法：我們想得到 $\frac{1}{4}\pi=\arcsin(7\sqrt{2}/10)-\arctan(3/4)$，所以設 $\beta=\arctan(3/4)$，然後考慮 $\sin(\frac{1}{4}\pi+\beta)$。

由於 $0<\frac{3}{4}$，β 是銳角；再由 3–4–5 直角三角形，可得 $\sin\beta=\frac{3}{5}$ 及 $\cos\beta=\frac{4}{5}$。於是，根據複角公式（addition formula），

$$\begin{aligned}\sin\left(\tfrac{1}{4}\pi+\beta\right)&=\sin\tfrac{\pi}{4}\cos\beta+\cos\tfrac{\pi}{4}\sin\beta\\&=\frac{1}{\sqrt{2}}\left(\frac{4}{5}+\frac{3}{5}\right)=\frac{7}{5\sqrt{2}}=\frac{7\sqrt{2}}{10}.\end{aligned}$$

> **想一想** $\frac{1}{4}\pi+\beta$ 真的就是 arcsin 所給出的那個角嗎？

arcsin 的值域是 $[-\frac{1}{2}\pi, \frac{1}{2}\pi]$，所以我們需要 $\frac{1}{4}\pi + \beta$ 落在這個區間內。因為 $\tan\beta = \frac{3}{4} < 1 = \tan\frac{\pi}{4}$，而 tan 在 $(0, \frac{1}{2}\pi)$ 上是遞增的，故有 $0 < \beta < \frac{1}{4}\pi$，所以

$$\tfrac{1}{4}\pi < \tfrac{1}{4}\pi + \beta < \tfrac{1}{2}\pi.$$

因此 $\arcsin(7\sqrt{2}/10) = \frac{1}{4}\pi + \arctan(3/4)$，乘以 4 再移項，

$$\pi = 4\arcsin(7\sqrt{2}/10) - 4\arctan(3/4),$$

這就是所求的結果。

驗算

以數值計算，$\arctan(\sqrt{3}/2) \approx 0.713724$ 及 $\arccos(5/\sqrt{28}) \approx 0.333473$；兩者之和為 $1.047198 \approx \pi/3$，其三倍為 3.141593。至於第二個恆等式，$\arcsin(7\sqrt{2}/10) \approx 1.428899$ 及 $\arctan(3/4) \approx 0.643501$；兩者之差為 $0.785398 \approx \pi/4$。快速檢查一下亦可見，$\frac{1}{3}\pi$ 是方程在 $(\frac{1}{4}\pi, \frac{1}{2}\pi)$ 內的唯一解：附近的另一個解 $\alpha - \arccos(5/\sqrt{28}) \approx 0.380$ 小於 $\frac{1}{4}\pi$。

常見錯誤

- 沒有驗證 $x \in [0, \pi]$ 就寫下 $\arccos(\cos x) = x$（或沒有驗證 $x \in [-\frac{1}{2}\pi, \frac{1}{2}\pi]$ 就寫下 $\arcsin(\sin x) = x$）。嚴謹性的分數正正就在這裏。
- 把 $\tan\alpha$ 上下顛倒：由 $R\cos\alpha = 4$、$R\sin\alpha = 2\sqrt{3}$ 得到的是 $\tan\alpha = 2\sqrt{3}/4$，而不是 $4/(2\sqrt{3})$。

STEP 之外

把 π 表示成反正切（inverse tangent）組合的恆等式，是梅欽類公式（Machin-type formulae）的基礎，例如 $\frac{1}{4}\pi = 4\arctan\frac{1}{5} - \arctan\frac{1}{239}$。由於 $\arctan x$ 的級數（series）在 x 很小時收斂（converge）得很快，這類公式曾被用來計算 π 至許多個位數。要看清這些恆等式，最簡潔的方法是透過複數（complex numbers）：$\arctan(b/a)$ 是 $a + b\mathrm{i}$ 的輻角（argument），所以反正切相加就對應於複數相乘。Needham 的 *Visual Complex Analysis*（《複分析：可視化方法》）把這個幾何觀點闡述得十分精彩。

詳細解答

STEP II 2003，第 3 題

純數學 • 無理數與數學歸納法

STEP II 2003，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/03-s2-q3

關鍵思路

無理性（irrationality）是一種「不是」型的性質，所以最自然的工具是逆否命題（contra-positive）：與其證明無理數（irrational number）的立方根（cube root）是無理數，不如證明立方根若是有理數（rational number），原數本身就必定是有理數。之後的歸納法（induction）全靠這一個事實推動，因為每個 u_{n+1} 都是 u_n 的立方根。至於最後一部分，我們留意到 $u_n \to 1$，於是把整個數列（sequence）平移一個整數，這樣每一項仍然是無理數。

解答

無理數的立方根

設 x 為實數，並設 $y = \sqrt[3]{x}$ 為它的實立方根，即 $y^3 = x$。我們證明逆否命題：若 y 是有理數，則 x 是有理數。

假設 $y = \dfrac{p}{q}$，其中 p, q 為整數且 $q \neq 0$。那麼

$$x = y^3 = \frac{p^3}{q^3},$$

而 p^3 和 q^3 都是整數，且 $q^3 \neq 0$。所以 x 是兩個整數之比，也就是有理數。

因此，有理的立方根只能來自有理數。換言之，若 x 是無理數，則 $\sqrt[3]{x}$ 不可能是有理數，所以它是無理數。這正是我們要證明的結論。

? **想一想** 為甚麼不直接處理原命題？

直接證明要從「x 不是整數之比」出發，但這句話沒有提供任何具體的東西可供計算。逆否命題則從「$y = p/q$」出發，馬上給了我們可以運算的整數，於是一行就證完了。每當假設說某個數*不是*有理數時，都值得試試逆否命題或反證法（proof by contradiction）。

歸納法

由於 $3^{n+1} = 3 \cdot 3^n$，根據指數定律（laws of indices），

$$u_{n+1} = 5^{1/3^{n+1}} = \left(5^{1/3^n}\right)^{1/3} = \sqrt[3]{u_n}.$$

所以每一項都是前一項的立方根。

基礎情況。$u_1 = 5^{1/3} = \sqrt[3]{5}$，題目已告訴我們它是無理數。

歸納步驟。假設對某個正整數 k，u_k 是無理數。那麼 $u_{k+1} = \sqrt[3]{u_k}$ 是一個無理數的立方根，所以根據第一部分，它是無理數。

因此，根據歸納法，對所有正整數 n，u_n 都是無理數。

收斂至 m 的數列

首先求 u_n 的極限（limit）。取對數（logarithm），

$$\ln u_n = \frac{\ln 5}{3^n} \to 0 \quad \text{當} n \to \infty,$$

而由於指數函數是連續（continuous）的，$u_n = \mathrm{e}^{\ln u_n} \to \mathrm{e}^0 = 1$。此外 $\ln u_n > 0$，所以對所有 n 都有 $u_n > 1$；各項遞減並趨向 1：

$$u_1 \approx 1.710, \quad u_2 \approx 1.196, \quad u_3 \approx 1.061, \quad u_4 \approx 1.020, \ldots$$

想一想 怎樣把極限由 1 移到 m，而又不破壞無理性？

無理數加上一個整數，結果仍是無理數。事實上，若 a 是整數，而 $u + a = r$ 是有理數，那麼 $u = r - a$ 也會是有理數，因為兩個有理數之差是有理數。

所以，對給定的整數 m，定義

$$a_n = u_n + m - 1 = 5^{1/3^n} + m - 1.$$

由於 $m - 1$ 是整數而 u_n 是無理數，每個 a_n 都是無理數，而且

$$a_n \to 1 + m - 1 = m \quad \text{當} n \to \infty.$$

所以 (a_n) 是一個由無理數組成、收斂（converge）至 m 的無窮數列。這個構造對任何整數 m 都適用，包括 $m = 0$。

（另一個自然的選擇是 $m\,u_n$，它同樣是無理數並趨向 m，但只限於 $m \neq 0$ 的情況：當 $m = 0$ 時，每一項都是 0，是有理數。所以平移 $m - 1$ 才是較穩妥的答案。）

驗算

- 數值上，$u_2^3 \approx 1.19581^3 \approx 1.70998 \approx u_1$，印證了 $u_{n+1} = \sqrt[3]{u_n}$。
- 取 $m = 4$：$a_1 \approx 4.710$，$a_2 \approx 4.196$，$a_3 \approx 4.061$，而 a_{20} 與 4 相符至小數點後八位以上。
- 差 $a_n - m = u_n - 1$ 是正數，大小約為 $\ln 5/3^n$，所以每一步大約縮小至原來的 3 分之一。

常見錯誤

- 證明了逆命題（converse）（「無理數的立方是無理數」），而這個命題是錯的：$\left(\sqrt[3]{2}\right)^3 = 2$。
- 使用 $m\,u_n$，卻沒有留意到它在 $m = 0$ 時失效。
- 忘記說明為何 $u_n \to 1$；「指數趨向零」這句話還需要交代一下連續性，或者用上面的對數論證。

STEP 之外

這道題顯示無理數在實數線上是稠密（dense）的：每一個實數，而不僅是每一個整數，都是無理數的極限（試試 $q_n + \sqrt{2}/n$，其中有理數 $q_n \to x$）。有理數同樣稠密，但在某個精確的意義下，無理數遠比有理數多：有理數是可數（countable）的，而實數則不是。Abbott 的 *Understanding Analysis* 在開首幾章對稠密性、可數性和數列的極限有十分精彩的講解。

詳細解答

STEP II 2003，第 4 題

純數學 • 弓形的面積

STEP II 2003，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/03-s2-q4

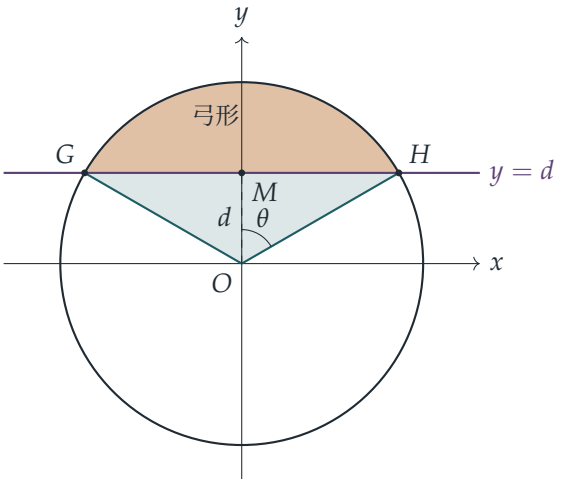
關鍵思路

劣弓形（minor segment）就是扇形（sector）減去一個三角形，所以第一個結果不過是「扇形減三角形」。至於其餘部分，留意在 $(*)$ 中，d 這個數其實是圓心到直線的垂直距離（*perpendicular distance*）。把圖形平移或旋轉都不會改變任何面積，所以在每一種情況下，我們都保留 $(*)$，只把 d 換成該距離。全部工夫都在於計算這個距離。

解答

基本結果

由於直線與圓相交於兩點，故 $0 < d < R$。把 $y = d$ 代入 $x^2 + y^2 = R^2$ 得 $x = \pm\sqrt{R^2 - d^2}$，所以 $G = \left(-\sqrt{R^2 - d^2}, d\right)$ 及 $H = \left(\sqrt{R^2 - d^2}, d\right)$。設 $M = (0, d)$ 為 GH 的中點，並設 $\theta = \angle MOH$。



在直角三角形 OMH 中，斜邊（hypotenuse）為 $OH = R$，鄰邊為 $OM = d$，所以 $\cos\theta = d/R$。由於 $0 < d/R < 1$，我們有 $\theta = \arccos(d/R)$，其中 $0 < \theta < \frac{\pi}{2}$，而根據對稱性，$\angle GOH = 2\theta < \pi$。

圓心角為 2θ 的扇形 OGH，面積為

$$\tfrac{1}{2}R^2(2\theta) = R^2 \arccos\left(\frac{d}{R}\right).$$

三角形 OGH 的底為 $GH = 2\sqrt{R^2 - d^2}$，高為 $OM = d$，所以其面積為 $d\sqrt{R^2 - d^2}$。因為 $2\theta < \pi$，圓心 O 與劣弧（minor arc）分別位於 GH 的兩側，所以扇形恰好就是三角形加上劣弓形。因此

$$\text{劣弓形的面積} = R^2 \arccos\left(\frac{d}{R}\right) - d\sqrt{R^2 - d^2},$$

這就是 $(*)$。

想一想 從幾何的角度看，$(*)$ 中的 d 究竟代表甚麼？

上述論證完全沒有用到坐標軸：它只用到半徑 R，以及圓心到弦（chord）的距離 $OM = d$。把整個圖形平移（translation）或旋轉（rotation），這兩個量都保持不變，面積亦保持不變。所以對於任何半徑為 R 的圓，以及任何與圓心的垂直距離為 p 的直線，其中 $0 \le p < R$，劣弓形的面積為

$$R^2 \arccos\left(\frac{p}{R}\right) - p\sqrt{R^2 - p^2}. \qquad (\dagger)$$

（若 $p = 0$，直線就是直徑，兩部分是相等的兩半，而 $(\dagger)$ 正確地給出 $\frac{1}{2}\pi R^2$。）在每一部分中，我們求出 p，再代入 $(\dagger)$。

第 (i) 部分

圓心為 (a,b)，而直線 $y=c$ 是水平線，所以圓心到直線的距離為 $p=|c-b|$。把 d 換成 $|c-b|$：

$$R^2\arccos\left(\frac{|c-b|}{R}\right)-|c-b|\sqrt{R^2-(c-b)^2}.$$

這裏的絕對值（modulus）不可省略：c 可能小於 b。

第 (ii) 部分

這裏圓心為 O。我們需要求 O 到 $y=mx+c$ 的距離。設直線與 x 軸的夾角為 α，故 $\tan\alpha=m$，並設直線與 y 軸相交於 $C=(0,c)$。由 O 向直線作垂線，垂足為 F。OC（鉛垂）與 OF（垂直於直線）之間的夾角，等於水平線與該直線之間的夾角，即 α。所以在直角三角形 OFC 中，

$$p=OF=|c|\,|\cos\alpha|=\frac{|c|}{\sqrt{1+\tan^2\alpha}}=\frac{|c|}{\sqrt{1+m^2}}.$$

（這就是常用公式 $|Ax_0+By_0+C|/\sqrt{A^2+B^2}$ 應用於直線 $mx-y+c=0$ 和點 $(0,0)$ 的結果。）於是

$$\sqrt{R^2-p^2}=\sqrt{R^2-\frac{c^2}{1+m^2}}=\frac{\sqrt{R^2(1+m^2)-c^2}}{\sqrt{1+m^2}},$$

所以劣弓形的面積為

$$R^2\arccos\left(\frac{|c|}{R\sqrt{1+m^2}}\right)-\frac{|c|\sqrt{R^2(1+m^2)-c^2}}{1+m^2}.$$

直線與圓相交，當且僅當 $c^2<R^2(1+m^2)$，所以平方根是實數。

第 (iii) 部分

設 $X=x-a$、$Y=y-b$，把原點移到圓心。圓變成 $X^2+Y^2=R^2$，而直線變成

$$Y+b=m(X+a)+c \quad\Longleftrightarrow\quad Y=mX+k, \qquad k=c+ma-b.$$

這就是把 c 換成 k 的第 (ii) 部分，而平移不會改變面積。所以 $p=|c+ma-b|/\sqrt{1+m^2}$，劣弓形的面積為

$$R^2\arccos\left(\frac{|c+ma-b|}{R\sqrt{1+m^2}}\right)-\frac{|c+ma-b|\sqrt{R^2(1+m^2)-(c+ma-b)^2}}{1+m^2}.$$

代入 $m=0$ 便得回第 (i) 部分，代入 $a=b=0$ 便得回第 (ii) 部分。

想一想 為甚麼在第 (i) 部分不能直接寫 $c-b$ 而不加絕對值？

若 $c < b$，則 $(c-b)/R$ 為負數，arccos 給出的角是大於 $\frac{\pi}{2}$ 的 $\pi - \theta$，而第二項亦會變號。該式變成 $R^2(\pi-\theta) + p\sqrt{R^2-p^2}$，這是優弓形（major segment）的面積，而不是劣弓形的面積。

驗算

- 取 $R = 2$、$d = 1$，得 $\theta = \frac{\pi}{3}$，而 $(*)$ 給出 $\frac{4\pi}{3} - \sqrt{3} \approx 2.4567$。這小於圓盤面積的一半（$2\pi \approx 6.28$），與預期相符。
- 當 $d \to R$ 時，面積趨向 0；而當 $d = 0$ 時，面積等於 $\frac{1}{2}\pi R^2$。
- 我們隨機選取了數百組圓和直線，把全部四條公式與直接對弦長作數值積分（numerical integration）的結果比較，兩者相差不超過 2×10^{-6}。

常見錯誤

- 漏掉絕對值。每當圓心位於直線「錯」的一側，帶正負號的距離便會給出優弓形的面積。
- 在第 (ii) 和第 (iii) 部分，用了鉛垂距離 $|c|$，而不是垂直距離 $|c|/\sqrt{1+m^2}$。
- 用 $\frac{1}{2}R^2\theta$ 計算扇形面積，忘記了圓心角是 2θ 而不是 θ。

STEP 之外

這道題建基於一個事實：面積在等距變換（isometry）之下保持不變，等距變換即平面上的平移、旋轉和反射（reflection）。在大學一年級的課程中，這會變成以下的說法：等距變換的雅可比行列式（Jacobian determinant）的絕對值為 1，所以二重積分（double integral）的變數代換公式不會改變面積。把等距變換分類，並找出它們所保持的量，是進入幾何學和羣論（group theory）的一條好途徑。Strang 的 *Introduction to Linear Algebra*（線性代數入門）講述了旋轉矩陣，以及行列式作為面積比例因子的角色。

詳細解答

STEP II 2003，第 5 題

純數學 • 直線的向量幾何

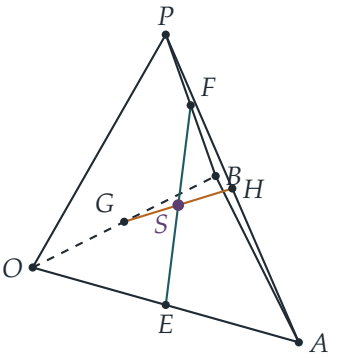
STEP II 2003，第 5 題

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/03-s2-q5

關鍵思路

先寫出四個中點，把每條直線寫成「點 + 參數 × 方向」的形式，然後令兩者相等。三維空間中的兩條直線通常不會相交，因此真正要做的，是證明這三條含兩個未知參數的分量方程是相容（consistent）的。**k** 分量很快就能解決問題。從幾何上看，O、A、B、P 構成一個四面體（tetrahedron），而 EF 和 GH 各自連接一對對稜的中點。由此可以猜想，它們相交於四個頂點的「平均」位置。



解答

直線 EF 與 GH 相交

設 $\mathbf{p} = l\mathbf{i} + m\mathbf{j} + n\mathbf{k}$。各中點的位置向量（position vector）為

$$\mathbf{e} = \tfrac{1}{2}a\mathbf{i}, \qquad \mathbf{f} = \tfrac{1}{2}\big(l\mathbf{i} + (b+m)\mathbf{j} + n\mathbf{k}\big),$$
$$\mathbf{g} = \tfrac{1}{2}b\mathbf{j}, \qquad \mathbf{h} = \tfrac{1}{2}\big((a+l)\mathbf{i} + m\mathbf{j} + n\mathbf{k}\big).$$

EF 上的一般點是 $\mathbf{e} + s(\mathbf{f} - \mathbf{e})$，而 GH 上的一般點是 $\mathbf{g} + t(\mathbf{h} - \mathbf{g})$。把兩者都乘以二以消去二分之一，得

$$2\big(\mathbf{e} + s(\mathbf{f} - \mathbf{e})\big) = \big(a + s(l-a)\big)\mathbf{i} + s(b+m)\mathbf{j} + sn\,\mathbf{k},$$
$$2\big(\mathbf{g} + t(\mathbf{h} - \mathbf{g})\big) = t(a+l)\mathbf{i} + \big(b + t(m-b)\big)\mathbf{j} + tn\,\mathbf{k}.$$

兩條直線相交，當且僅當存在 s 和 t 的值使以上兩式相等，因此以下三式必須全部成立：

$$\mathbf{i}: \quad a + s(l-a) = t(a+l),$$
$$\mathbf{j}: \quad s(b+m) = b + t(m-b),$$
$$\mathbf{k}: \quad sn = tn.$$

想一想 三條方程、兩個未知數：為甚麼它們竟然會有解？

一般而言是沒有解的，這正是空間中兩條直線通常不會相交的原因。這裏我們先解其中兩條方程，再驗證第三條也成立。由於 $n \neq 0$，$\mathbf{k}$ 方程給出 $s = t$。於是 $\mathbf{i}$ 方程變成 $a + sl - sa = sa + sl$，即 $a = 2sa$，而由於 $a \neq 0$，我們得到 $s = \frac{1}{2}$。最後，當 $s = t = \frac{1}{2}$ 時，$\mathbf{j}$ 方程成為

$$\tfrac{1}{2}(b+m) = b + \tfrac{1}{2}(m-b) = \tfrac{1}{2}(b+m),$$

這是成立的。所以當 $s = t = \frac{1}{2}$ 時三條方程全部成立，直線 EF 與 GH 相交，正如所求。

交點是 EF 的中點（也是 GH 的中點）：

$$\mathbf{s} = \tfrac{1}{4}\big((a+l)\mathbf{i} + (b+m)\mathbf{j} + n\mathbf{k}\big) = \tfrac{1}{4}\big(\mathbf{0} + a\mathbf{i} + b\mathbf{j} + \mathbf{p}\big),$$

即 O、A、B 和 P 的位置向量的平均。

點 T

由於 S 是 DT 的中點，我們有 $\mathbf{s} = \frac{1}{2}(\mathbf{d} + \mathbf{t})$，所以 $\mathbf{t} = 2\mathbf{s} - \mathbf{d}$：

$$\mathbf{t} = \tfrac{1}{2}(a+l)\,\mathbf{i} + \tfrac{1}{2}(b+m)\,\mathbf{j} + \big(\tfrac{1}{2}n - d\big)\mathbf{k}.$$

平面 OAB 包含 O 以及向量 $a\mathbf{i}$ 和 $b\mathbf{j}$。由於 a 和 b 均不為零，這兩個向量分別沿 x 軸和 y 軸的方向，所以平面 OAB 就是平面 $z = 0$。因此 T 位於該平面上，當且僅當它的 $\mathbf{k}$ 分

量為零，即 $\frac{1}{2}n - d = 0$。故此

$$d = \tfrac{1}{2}n.$$

由於 $n \neq 0$，這個值不為零，所以與條件 $d \neq 0$ 並無矛盾。

驗算

- 取 $a = 2$，$b = 4$，$\mathbf{p} = (2,0,4)$。則 $E = (1,0,0)$，$F = (1,2,2)$，$G = (0,2,0)$，$H = (2,0,2)$。EF 和 GH 的中點都是 $(1,1,1)$，亦即 $\frac{1}{4}\big((0,0,0) + (2,0,0) + (0,4,0) + (2,0,4)\big)$。
- 取 $d = 2 = \frac{1}{2}n$，則 $T = 2(1,1,1) - (0,0,2) = (2,2,0)$，它位於 $z = 0$ 上。
- 我們以電腦用精確的有理數運算，對 20 000 組隨機選取的 a, b, l, m, n, d 作了檢驗，確認兩條直線在 $s = t = \frac{1}{2}$ 時相交於上述的點，並且 T 位於平面 OAB 上（純量三重積（scalar triple product）為零）當且僅當 $d = \frac{1}{2}n$。

常見錯誤

- 一開始就用同一個參數來令兩條直線相等。這樣做等於預先假設了答案 $s = t$，而在這裏它是由 **k** 方程推出來的。
- 解了其中兩條分量方程就停下來。要證明空間中兩條直線相交，還必須驗證第三條方程。
- 忘記說明在哪裏用到 $a \neq 0$ 和 $n \neq 0$：正是有了這兩個條件，我們才可以作除法求出 s 和 t。

STEP 之外

任何四面體中，連接對稜中點的三條直線（稱為雙中線（bimedians））全都通過它的形心（centroid），即四個頂點的平均，並且各自在該點被平分。這是「三角形的中線（median）交於形心」在三維空間中的近親，也讓我們初次體會到把點看成加權平均的想法，而這正是重心坐標（barycentric coordinates）和仿射坐標（affine coordinates）背後的概念。Strang 的 *Introduction to Linear Algebra*（線性代數入門）詳細闡述了這裏所用的向量與線性方程觀點，包括方程數目多於未知數的方程組在甚麼情況下相容。

詳細解答

STEP II 2003，第 6 題

純數學 • 迭代函數與積分

> **STEP II 2003，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/03-s2-q6

關鍵思路

再施行一次 f，意思就是「先減去 1，再把負的部分反射回橫軸上方」。若某一段圖像的值介乎 0 與 1 之間，這樣做只會把該段在帶狀區域 $0 \le y \le 1$ 內上下倒轉，而兩條長長的直線臂則各多出一個鋸齒。於是圖像變成鋸齒形（三角波，triangle wave）；看出這一點之後，最後一部分的兩個積分都只是若干個相同週期上的積分之和。

解答

略圖

整個解答只靠以下兩項事實。

遞推關係（*recurrence*）。由於 $g_{n+1} = f \circ g_n$，我們有 $g_{n+1}(x) = |g_n(x) - 1|$：把 g_n 的圖像向下平移 1，再把位於 x 軸下方的部分反射上去。

對稱性（*symmetry*）。由於 $f(2-x) = |1-x| = f(x)$，當 $n \ge 2$ 時可得 $g_n(2-x) = g_{n-1}\big(f(2-x)\big) = g_{n-1}\big(f(x)\big) = g_n(x)$，而 $n = 1$ 時則可直接驗證。所以每個 g_n 都關於直線 $x = 1$ 對稱，只需弄清楚 $x \ge 1$ 的情況便可。

先從 $g_1(x) = |x-1|$ 開始，這是一個頂點在 $(1,0)$ 的 V 形。

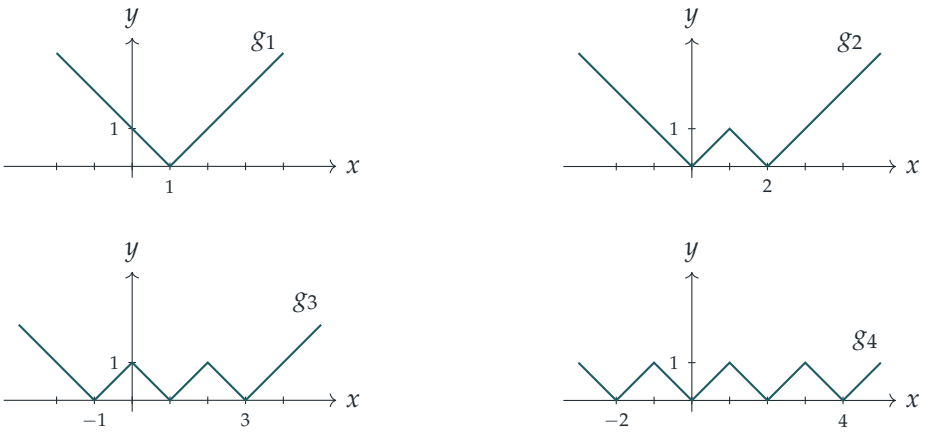
至於 g_2：在 $[0,2]$ 上有 $0 \le g_1 \le 1$，所以 $g_2 = 1 - |x-1|$，是一個帳篷形，在 $x = 1$ 處升至 1，在 $x = 0$ 和 $x = 2$ 處降至 0。當 $x \ge 2$ 時，$g_2 = x - 2$；由對稱性，當 $x \le 0$ 時 $g_2 = -x$。所以 g_2 呈 W 形。

之後每一步都是做同樣的事：鋸齒部分（取值在 $[0,1]$ 內）上下倒轉，而直線臂 $x - n$ 變成 $|x-n-1|$，在 $[n, n+1]$ 上多添一個鋸齒（由對稱性，在 $[1-n, 2-n]$ 上亦然）。以

$d_{\text{even}}(x)$ 和 $d_{\text{odd}}(x)$ 分別表示 x 與最接近的偶數及最接近的奇數之間的距離，便得

$$g_n(x) = \begin{cases} 2-n-x, & x \le 2-n, \\ d_{\text{even}}(x)\text{（若 } n \text{ 為偶數），} d_{\text{odd}}(x)\text{（若 } n \text{ 為奇數），} & 2-n \le x \le n, \\ x-n, & x \ge n. \end{cases}$$

這個結果可由遞推關係用歸納法（induction）證明：對任意實數 x 都有 $d_{\text{even}} + d_{\text{odd}} = 1$（在兩個相鄰整數之間，這兩個距離分別是 t 和 $1-t$），所以「1 減去鋸齒」會把奇偶性對調；而在 $[n, n+1]$ 上，新增的一段 $n+1-x$ 正好是到 $n+1$ 的距離，而 $n+1$ 就是具有新奇偶性的最接近整數。左邊的部分則由關於 $x=1$ 的對稱性得出。



用文字描述：g_3 的零點在 $-1, 1, 3$，在 0 和 2 處有高度為 1 的尖峯，兩臂為 $y = -1-x$ 和 $y = x-3$；g_4 的零點在 $-2, 0, 2, 4$，在 $-1, 1, 3$ 處有高度為 1 的尖峯，兩臂為 $y = -2-x$ 和 $y = x-4$。四個圖像都關於 $x=1$ 對稱。

積分

設 n 為偶數。由於 $n \ge 2$，區間 $[0, n]$ 包含在 $[2-n, n]$ 之內，所以在這個區間上 $g_n(x) = d_{\text{even}}(x)$。

> **想一想** 為甚麼只需在 $[0, 2]$ 上積分便足夠？

兩個函數的週期（period）都是 2：$d_{\text{even}}(x+2) = d_{\text{even}}(x)$，因為平移 2 會把偶數映到偶數；而 $\left|\sin\frac{\pi(x+2)}{2}\right| = \left|-\sin\frac{\pi x}{2}\right| = h(x)$。由於 n 是偶數，$[0, n]$ 正好由 $\frac{n}{2}$ 個完整週期 $[0,2], [2,4], \ldots, [n-2, n]$ 組成，所以在 $[0, n]$ 上的每個積分都是在 $[0, 2]$ 上的積分的 $\frac{n}{2}$ 倍。

在 $[0,2]$ 上，g_n 就是帳篷形 $1-|x-1|$：一個底為 2、高為 1 的三角形，所以

$$\int_0^2 g_n(x)\,\mathrm{d}x = \tfrac{1}{2}\cdot 2\cdot 1 = 1.$$

在 $[0,2]$ 上，$\frac{\pi x}{2}\in[0,\pi]$，所以 $\sin\frac{\pi x}{2}\geq 0$，絕對值（modulus）符號可以略去：

$$\int_0^2 h(x)\,\mathrm{d}x = \int_0^2 \sin\frac{\pi x}{2}\,\mathrm{d}x = \Big[-\frac{2}{\pi}\cos\frac{\pi x}{2}\Big]_0^2 = \frac{2}{\pi}+\frac{2}{\pi} = \frac{4}{\pi}.$$

因此

$$\int_0^n \big(h(x)-g_n(x)\big)\,\mathrm{d}x = \frac{n}{2}\Big(\frac{4}{\pi}-1\Big) = \frac{2n}{\pi}-\frac{n}{2},$$

這正是所求。

驗算

- 取一個數值，親手迭代一次：$g_4(-1)$：$-1\to 2\to 1\to 0\to 1$，所以 $g_4(-1)=1$，是一個尖峯，與略圖相符。
- 答案是正數，因為 $4/\pi\approx 1.27>1$：正弦拱形比同底同高的直邊帳篷形向外鼓出，這與兩個圖像的形狀吻合。
- 數值方面，取 $n=2$ 時用辛普森法則（Simpson's rule）計算得 0.273240，與 $\frac{4}{\pi}-1\approx 0.273240$ 一致；當 $n=4,6,8,10$ 時同樣吻合。

常見錯誤

- 忘記左邊的部分：當 $n\geq 3$ 時，在 $x<0$ 的範圍內圖像並不是直線。利用對稱性 $g_n(2-x)=g_n(x)$ 是把這部分畫對的最快方法。
- 在 $[0,n]$ 上積分 $\sin\frac{\pi x}{2}$ 時漏掉絕對值，這樣對偶數 n 會得出 0，而不是 $\frac{2n}{\pi}$。

STEP 之外

把一個簡單的映射反覆迭代，觀察圖像如何變化，正是離散動力系統（discrete dynamical systems）的起點。這裏的鋸齒最終穩定成一個三角波，而與之密切相關的「帳篷映射」（tent map）則是混沌（chaos）的經典例子。把 h 與三角波作比較，也引向傅立葉級數（Fourier series）：在那裏，三角波會寫成正弦與餘弦之和，而 $|\sin|$ 是最早遇到的例子之一。Boyce 與 DiPrima 合著的 *Elementary Differential Equations*（初等微分方程）對傅立葉級數有一段易讀的入門介紹。

詳細解答

STEP II 2003，第 7 題

純數學 • 廣義積分與級數

STEP II 2003，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/03-s2-q7

關鍵思路

形如 $\ln x \times$ (冪，底數為x) 的被積函數(integrand)，最適合用分部積分法(integration by parts）處理：對對數求導，它便會消失。第一個積分就這樣解決了。題目餘下的部分把精確值 $2/(n^2\mathrm{e})$ 反過來用：它讓我們可以把 $1/n^2$ 寫成一個積分。然後我們把所有積分下限移到同一位置，這樣只會令每個積分變大，再在積分號內把一個等比級數（geometric series）加起來。

解答

積分

對 $a > 1$ 及 $n > 0$，設

$$I_n(a) = \int_a^\infty \frac{\ln x}{x^{n+1}}\,\mathrm{d}x.$$

取 $u = \ln x$ 及 $\mathrm{d}v = x^{-n-1}\,\mathrm{d}x$ 作分部積分，於是 $v = -\frac{1}{n}x^{-n}$（這裏用到 $n \neq 0$）。在 $[a, X]$ 上，

$$\begin{aligned}\int_a^X \frac{\ln x}{x^{n+1}}\,\mathrm{d}x &= \left[-\frac{\ln x}{nx^n}\right]_a^X + \frac{1}{n}\int_a^X x^{-n-1}\,\mathrm{d}x \\ &= \frac{\ln a}{na^n} - \frac{\ln X}{nX^n} + \frac{1}{n^2}\left(\frac{1}{a^n} - \frac{1}{X^n}\right).\end{aligned}$$

現在令 $X \to \infty$。由於 $n > 0$，$X^{-n} \to 0$。至於另一項，設 $Y = X^n$，則 $\ln X = \frac{1}{n}\ln Y$，並且

$$\frac{\ln X}{X^n} = \frac{1}{n}\cdot\frac{\ln Y}{Y} \to 0,$$

因為當 $X \to \infty$ 時 $Y \to \infty$，而題目容許我們假設 $(\ln Y)/Y \to 0$。（這裏不能直接引用 $(\ln X)/X \to 0$：當 $0 < n < 1$ 時，X^n 增長得比 X 慢。）所以這個積分收斂（converge），並且

$$I_n(a) = \frac{\ln a}{na^n} + \frac{1}{n^2 a^n}.$$

當 $a = \mathrm{e}^{1/n}$ 時，$a^n = \mathrm{e}$ 且 $\ln a = 1/n$，所以

$$I_n\left(\mathrm{e}^{1/n}\right) = \frac{1}{n^2\mathrm{e}} + \frac{1}{n^2\mathrm{e}} = \frac{2}{n^2\mathrm{e}},$$

這正是所求的結果。

比較兩個積分

若 $1 < a < b$，則

$$I_n(a) - I_n(b) = \int_a^b \frac{\ln x}{x^{n+1}}\,\mathrm{d}x.$$

對 $x > 1$，有 $\ln x > 0$ 及 $x^{n+1} > 0$，所以被積函數在 $[a, b]$ 上嚴格為正。它同時是連續的，而且 $b > a$，所以這個積分嚴格為正。因此 $I_n(b) < I_n(a)$：把積分下限向右移，等於捨棄了一塊正的面積。

不等式

由第一部分，對每個正整數 n，

$$\frac{1}{n^2} = \frac{\mathrm{e}}{2} I_n\left(\mathrm{e}^{1/n}\right).$$

> **想一想** 和式中每一項的積分下限 $\mathrm{e}^{1/n}$ 各不相同，但右邊只有單一個下限 $\mathrm{e}^{1/N}$。我們怎樣由前者過渡到後者？

取 $1 \le n \le N$。那麼 $1/N \le 1/n$，所以 $1 < \mathrm{e}^{1/N} \le \mathrm{e}^{1/n}$。當 $n < N$ 時不等式是嚴格的，而第二部分（取 $a = \mathrm{e}^{1/N}$，$b = \mathrm{e}^{1/n}$）給出

$$I_n\left(\mathrm{e}^{1/n}\right) < I_n\left(\mathrm{e}^{1/N}\right);$$

當 $n = N$ 時兩邊相等。對 $n = 1, \ldots, N$ 求和，所得的不等式是嚴格的，因為 $N > 1$ 意味着最少有一項（例如 $n = 1$）滿足 $n < N$：

$$\sum_{n=1}^{N} \frac{1}{n^2} = \frac{\mathrm{e}}{2} \sum_{n=1}^{N} I_n\left(\mathrm{e}^{1/n}\right) < \frac{\mathrm{e}}{2} \sum_{n=1}^{N} I_n\left(\mathrm{e}^{1/N}\right).$$

這正是需要用到 $N > 1$ 的地方。若 $N = 1$，兩邊便會相等。

餘下的工作是把這些積分加起來。有限個在同一範圍上收斂的積分之和，等於各被積函數之和的積分，所以

$$\sum_{n=1}^{N} I_n\left(\mathrm{e}^{1/N}\right) = \int_{\mathrm{e}^{1/N}}^{\infty} \left(\sum_{n=1}^{N} x^{-(n+1)} \right) \ln x\,\mathrm{d}x.$$

括號內是一個等比級數，首項為 x^{-2}，公比為 $x^{-1} \neq 1$（因為 $x > 1$），共 N 項：

$$\sum_{n=1}^{N} x^{-(n+1)} = \frac{x^{-2}\left(1 - x^{-N}\right)}{1 - x^{-1}} = \frac{1 - x^{-N}}{x^2 - x},$$

這裏分子和分母同乘以 x^2。因此

$$\sum_{n=1}^{N} \frac{1}{n^2} < \frac{\mathrm{e}}{2} \int_{\mathrm{e}^{1/N}}^{\infty} \left(\frac{1 - x^{-N}}{x^2 - x}\right) \ln x \, \mathrm{d}x,$$

正如所求。

驗算

- 把 $I_n(a)$ 的公式對 a 求導：你應該得到 $-\ln a / a^{n+1}$，即被積函數的相反數，正如微積分基本定理（fundamental theorem of calculus）所要求的。
- 數值上，利用代換（substitution）$x = \mathrm{e}^u$，最後一部分的積分變成 $\int_{1/N}^{\infty} u(1 - \mathrm{e}^{-Nu})/(\mathrm{e}^u - 1)\,\mathrm{d}u$。當 $N = 2$ 時，右邊約為 1.4865，而左邊是 $1 + \frac{1}{4} = 1.25$。當 $N = 10$ 時，右邊約為 2.0579，而左邊是 1.5498。
- 當 $N \to \infty$ 時，右邊趨向 $\frac{\mathrm{e}}{2} \int_1^{\infty} \frac{\ln x}{x^2 - x}\,\mathrm{d}x$，而代換 $x = 1/t$ 把它化為 $\frac{\mathrm{e}}{2} \int_0^1 \frac{-\ln t}{1-t}\,\mathrm{d}t = \frac{\mathrm{e}}{2} \cdot \frac{\pi^2}{6} \approx 2.236$。這個值明顯大於 $\pi^2/6 \approx 1.645$，而這是必然的。

常見錯誤

- 直接用 $(\ln X)/X \to 0$ 來消去 $(\ln X)/X^n$。這對 $n \geq 1$ 行得通，但題目的結果是對所有 $n > 0$ 而言的。
- 把積分下限移錯方向。由於 $\mathrm{e}^{1/N}$ 是各個 $\mathrm{e}^{1/n}$ 之中最小的，每個積分都會變大。
- 聲稱不等式是嚴格的，卻沒有說明原因。$n = N$ 的一項給出等式，所以嚴格性來自 $n < N$ 的各項，而這些項之所以存在，是因為 $N > 1$。

STEP 之外

把一個數寫成積分，然後在積分號內求和，是分析(analysis)中的標準手法。把這個做法推到極限，便得到 $\sum 1/n^2 = \int_0^1 \frac{-\ln t}{1-t}\,\mathrm{d}t$，這是通往 $\pi^2/6$ 的其中一條路徑。令 $N \to \infty$，就是把無窮和與反常積分（improper integral）交換次序。這個交換需要論證，要用到單調收斂定理或控制收斂定理（monotone or dominated convergence theorems），你會在大學的第一門分析或測度論(measure theory)課程中遇到它們。Abbott 的 *Understanding Analysis* 寫得平易近人，很適合用來了解這類交換為何可能失敗，以及何時才是安全的。

詳細解答

STEP II 2003，第 8 題

純數學 • 可分離變量微分方程

STEP II 2003，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/03-s2-q8

關鍵思路

這是一階可分離變數（separable）的微分方程：$\frac{1}{y}\frac{\mathrm{d}y}{\mathrm{d}t}$ 等於 t 的一個有理函數（rational function）。把那個二次式除以 $t+1$，這個函數便分拆成一個多項式加一個簡單分式，兩者都可以立即積分。至於駐點（stationary point），直接由微分方程本身便可得出，無須對解求導：$\frac{\mathrm{d}y}{\mathrm{d}t}=0$ 恰好在二次式 t^2-3t+2 等於零之處成立。

解答

求 y

由多項式除法（polynomial division）得

$$\frac{t^2-3t+2}{t+1}=t-4+\frac{6}{t+1},$$

因為 $(t+1)(t-4)+6=t^2-3t-4+6=t^2-3t+2$。

解必須由 $y=A>0$ 出發。假如 y 在某處等於零，則根據這個線性方程的解的唯一性（uniqueness），它必定是零解，這與 $y(0)=A$ 矛盾。所以自始至終 $y>0$（在 $t>-1$ 上，即係數有定義的範圍內），因此可以除以 y：

$$\frac{1}{y}\frac{\mathrm{d}y}{\mathrm{d}t}=-k\left(t-4+\frac{6}{t+1}\right).$$

由 0 積分至 t，並利用 $\ln 1=0$，

$$\ln\frac{y}{A}=-k\left(\frac{t^2}{2}-4t+6\ln(t+1)\right),$$

所以

$$\boxed{y = A\,(t+1)^{-6k}\,\mathrm{e}^{-k(t^2/2-4t)} = A\,(t+1)^{-6k}\,\mathrm{e}^{k(8t-t^2)/2}}.$$

駐點值

由微分方程得

$$\frac{\mathrm{d}y}{\mathrm{d}t} = -k\,\frac{(t-1)(t-2)}{t+1}\,y.$$

我們假設 $k \neq 0$（否則 $y \equiv A$，每一點都是駐點）。由於 $y > 0$ 及 $t+1 > 0$，導數只在 $t = 1$ 和 $t = 2$ 處等於零，而且因為兩者都是單根（simple root），導數在這兩處都會變號。所以恰好有兩個駐點值（stationary value）：

$$y(1) = A \cdot 2^{-6k}\,\mathrm{e}^{-k(1/2-4)} = A \cdot 2^{-6k}\,\mathrm{e}^{7k/2},$$
$$y(2) = A \cdot 3^{-6k}\,\mathrm{e}^{-k(2-8)} = A \cdot 3^{-6k}\,\mathrm{e}^{6k}.$$

兩者之比為

$$\frac{y(1)}{y(2)} = \frac{3^{6k}}{2^{6k}}\,\mathrm{e}^{7k/2-6k} = \left(\frac{3}{2}\right)^{6k}\mathrm{e}^{-5k/2},$$

正是所求。

? **想一想** 兩個駐點值之中，哪一個較大？

取對數：

$$\ln\frac{y(1)}{y(2)} = k\left(6\ln\tfrac{3}{2} - \tfrac{5}{2}\right),$$

而 $6\ln 1.5 \approx 2.433 < 2.5$，所以括號內的數是負數（約為 -0.067）。因此當 $k > 0$ 時 $y(1) < y(2)$，當 $k < 0$ 時 $y(1) > y(2)$。兩者相差很小，所以圖像上的起伏十分平緩。同樣地，

$$\ln\frac{y(1)}{A} = k\left(\tfrac{7}{2} - 6\ln 2\right) \approx -0.659k, \qquad \ln\frac{y(2)}{A} = k(6 - 6\ln 3) \approx -0.592k,$$

所以當 $k > 0$ 時兩個駐點值都小於 A，當 $k < 0$ 時兩個都大於 A。

當 $t \to +\infty$ 時的表現

寫成

$$\ln\frac{y}{A} = -k\left(\frac{t^2}{2} - 4t + 6\ln(t+1)\right).$$

當 $t \to \infty$ 時，括號內的式子趨向 $+\infty$，並由 $\frac{1}{2}t^2$ 主導（ln 項的增長慢得多）。所以：

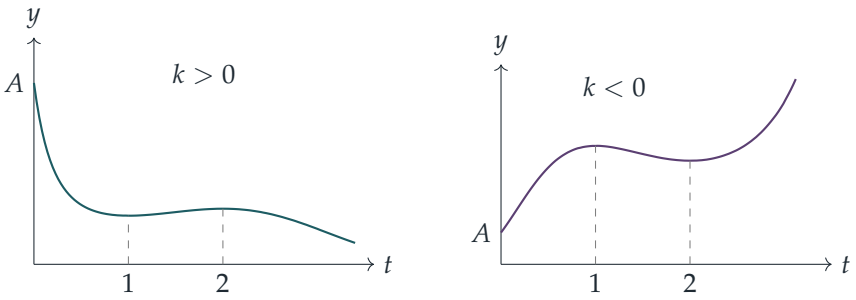
- 若 $k > 0$，則 $\ln(y/A) \to -\infty$，所以 $y \to 0$（速度極快，有如 $\mathrm{e}^{-kt^2/2}$）；
- 若 $k < 0$，則 $\ln(y/A) \to +\infty$，所以 $y \to \infty$（有如 $\mathrm{e}^{|k|t^2/2}$）。

草圖

當 $t > 0$ 時，$\frac{\mathrm{d}y}{\mathrm{d}t}$ 的正負號與 $-k(t-1)(t-2)$ 的正負號相同。

- $k > 0$：y 在 $(0,1)$ 上由 A 遞減，在 $t = 1$ 處達到局部極小值（local minimum），然後上升，在 $t = 2$ 處達到局部極大值（local maximum）（仍然低於 A），其後遞減並趨向 0。
- $k < 0$：y 由 A 遞增，在 $t = 1$ 處達到局部極大值，然後下降，在 $t = 2$ 處達到局部極小值（仍然高於 A），其後無止境地遞增。

以下圖像是取 $A = 1$ 及 $k = \pm 2$ 準確繪製的。



驗算

- 在 $t = 0$ 處，公式給出 $A \cdot 1 \cdot \mathrm{e}^0 = A$。對 $\ln y = \ln A - 6k\ln(t+1) - k(t^2/2 - 4t)$ 求導，得 $\frac{y'}{y} = -\frac{6k}{t+1} - k(t-4)$，合併後便回到原來的方程。
- 數值方面，取 $A = 1.5$ 及 $k = \pm 2$，以龍格－庫塔法（Runge–Kutta）逐步對微分方程積分，在 $t = 3$ 處所得結果與公式吻合至八位有效數字。取 $k = 2$ 時，比值 $y(1)/y(2)$ 為 0.8742，與 $(1.5)^{12}\mathrm{e}^{-5}$ 相符。

常見錯誤

- 除法出錯：餘數是 6，不是 2。可乘回 $t+1$ 來檢查。
- 把比值倒轉：題目要求的是 $y(1)/y(2)$，而指數中的正負號很容易令人誤得 $y(2)/y(1)$。
- 草圖中把兩個轉向點（turning point）畫在 A 的錯誤一側：先前的計算表明，當 $k > 0$ 時兩者都低於 A，當 $k < 0$ 時兩者都高於 A。

STEP 之外

這個方程是線性的，形如 $y' + p(t)y = 0$，而它的解 $A\exp\left(-\int_0^t p\right)$ 正是積分因子法（integrating-factor method）最簡單的情形。它也說明了你在第一門常微分方程（ODE）課程中會遇到的一個普遍原則：很多時候無須解出方程，便可直接從方程本身看出解的定性特徵（駐點、解的正負號、增長或衰減）。Boyce 與 DiPrima 合著的 *Elementary Differential Equations*（常微分方程入門教科書）對這兩個想法都有詳細闡述，當中包括我們用來證明 y 永不為零的唯一性定理（uniqueness theorem）。

詳細解答

STEP II 2003，第 9 題

力學 • 極限摩擦力平衡

STEP II 2003，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/03-s2-q9

關鍵思路

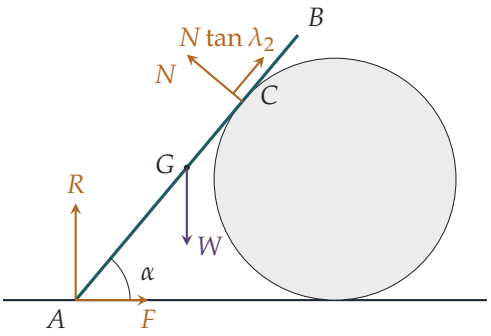
由於摩擦系數（coefficient of friction）是以 $\tan\lambda_1$ 和 $\tan\lambda_2$ 的形式給出，最自然的工具就是總反作用力（total reaction，即把法向反作用力（normal reaction）與摩擦力合成為一個力）。當摩擦達到極限（limiting）時，總反作用力與法線的夾角就是摩擦角（angle of friction）λ。這樣一來，桿只受三個力作用，而且三個力的方向都已知；沿水平和垂直方向分解（resolve），便得到兩條方程，其中只有兩個未知的大小。真正要花工夫的，只是判斷每個摩擦力指向哪一邊。

解答

模型與摩擦力的方向

我們把桿視為均勻（uniform）的剛體（rigid body）桿，並把圓柱視為固定不動。桿沿切線（tangent）與圓柱接觸，所以在 C 處的法向反作用力垂直於桿。設圓柱位於 A 的水平正方向一側。於是作用在桿上的力有：

- 桿的重量 W，作用於中點 G，垂直向下；
- 在 A 處：垂直向上的法向反作用力 R，以及沿地面的摩擦力 F；
- 在 C 處：垂直於桿的法向反作用力 N，指向上方並向後偏往 A 的一邊，另有一個沿桿方向的摩擦力。



想一想 在 A 處和在 C 處，摩擦力分別指向哪一邊？桿會不會是即將向相反方向滑動？

桿要同時保持兩個接觸點，就只有一種運動方式：當 A 沿地面滑動時，角 α 隨之改變，桿亦在圓柱上滑過。若 A 遠離圓柱，桿便會下降（即 α 減小），而桿上位於 C 的那一點相對於圓柱沿桿向下滑動。若 A 移向圓柱，一切就反過來。

在第二種情況下，A 處的摩擦力會指向遠離圓柱的方向，而 C 處的摩擦力會沿桿向下。這樣三個水平方向的力（A 處的摩擦力、N 的水平分量，以及 C 處摩擦力的水平分量）全都指向遠離圓柱的一邊，水平方向便不可能平衡。所以在極限平衡（limiting equilibrium）時，桿是即將向下滑：桿在 A 處所受的摩擦力指向圓柱，而 C 處的摩擦力沿桿向上，正如圖中所示。

總反作用力

設 S 為 A 處的總反作用力，T 為 C 處的總反作用力。由於兩個接觸點的摩擦都達到極限：

- S 與向上的垂直線成 λ_1 角，並向圓柱一方傾斜，因為 $F = R\tan\lambda_1$。
- C 處的法線與向上的垂直線成 α 角，並向後偏往 A 的一邊，因為桿與水平線成 α 角。沿桿向上的摩擦力使 T 朝垂直線轉回 λ_2。所以 T 與向上的垂直線成 $\alpha - \lambda_2$ 角，這個角是在遠離圓柱的一側量度的。

對桿分解各力，

$$\begin{aligned} &\text{水平方向：} && S\sin\lambda_1 = T\sin(\alpha - \lambda_2), \\ &\text{垂直方向：} && S\cos\lambda_1 + T\cos(\alpha - \lambda_2) = W. \end{aligned}$$

為了消去 T，我們把第一條方程乘以 $\cos(\alpha-\lambda_2)$，把第二條乘以 $\sin(\alpha-\lambda_2)$，再從第二條減去第一條：

$$\begin{aligned} S\big[\cos\lambda_1\sin(\alpha-\lambda_2)+\sin\lambda_1\cos(\alpha-\lambda_2)\big] &= W\sin(\alpha-\lambda_2),\\ S\sin(\alpha+\lambda_1-\lambda_2) &= W\sin(\alpha-\lambda_2), \end{aligned}$$

這裏用了 sin 的複角公式（addition formula）。題目告訴我們 $\alpha\neq\lambda_2-\lambda_1$，所以 $\sin(\alpha+\lambda_1-\lambda_2)\neq 0$（所有角都是銳角，所以 $\alpha+\lambda_1-\lambda_2$ 嚴格介乎 $-\frac{\pi}{2}$ 與 π 之間）。因此可以相除：

$$S=\frac{W\sin(\alpha-\lambda_2)}{\sin(\alpha+\lambda_1-\lambda_2)}.$$

A 處的摩擦力就是 S 的水平分量，即 $F=S\sin\lambda_1$，所以

$$F=\frac{W\sin\lambda_1\,\sin(\alpha-\lambda_2)}{\sin(\alpha+\lambda_1-\lambda_2)},$$

這正是所求。同樣地，$T=W\sin\lambda_1/\sin(\alpha+\lambda_1-\lambda_2)$。

想一想 我們從頭到尾都沒有取力矩（moment）。這不是很可疑嗎？

在這裏並不可疑。沿兩個方向分解，已足以求出兩個未知的大小。對 A 取力矩（S 對該點沒有力矩，C 處的摩擦力沿桿作用，同樣沒有力矩），得 $N\cdot AC=W\cdot AG\cos\alpha$，其中 $N=T\cos\lambda_2$。這只告訴我們，要達到極限平衡，C 必須位於哪裏：

$$\frac{AC}{AB}=\frac{\cos\alpha\,\sin(\alpha+\lambda_1-\lambda_2)}{2\sin\lambda_1\cos\lambda_2}.$$

求摩擦力時，並不需要用到題目所給的條件 $AC>CB$。從物理上說，我們還需要 $S>0$ 和 $T>0$，而這恰好在 $\alpha>\lambda_2$ 時成立：桿若以小於 λ_2 的傾角擱在圓柱上，就不可能處於即將沿圓柱滑下的狀態。

驗算

- 地面光滑，$\lambda_1=0$：公式給出 $F=0$，理應如此。
- $\alpha=\lambda_2$：此時 T 是垂直的，單靠它就把桿托住，而 $F=0$。這是桿剛好能擱在圓柱上的臨界角度。
- 數值驗算：取 $\lambda_1=0.3$、$\lambda_2=0.5$、$\alpha=1.0$（弧度），直接從分量方程解出 R 和 N，得 $F\approx 0.1975W$，與公式吻合。在這個例子中，再對 A 取力矩，可知 C 大約位於桿長的 0.75 處，與 $AC>CB$ 相符。

常見錯誤

- 把 C 處的反作用力當成垂直的，或當成垂直於圓柱的「頂部」。反作用力應垂直於公切線，而公切線就是桿本身。
- 選定摩擦力的方向卻不加檢驗。選錯了那一對方向，水平方向便不可能平衡。
- 把 C 處的總反作用力向錯誤的一邊傾斜，結果得出 $\alpha + \lambda_2$ 而不是 $\alpha - \lambda_2$。

STEP 之外

取總反作用力與法線成 λ 角，正是摩擦錐（cone of friction）背後的想法：只要總反作用力留在以法線為軸、半頂角為 λ 的圓錐之內，接觸點就不會滑動。若只有三個力，它們的作用線（line of action）還必須交於一點，這就把許多靜力學（statics）問題變成幾何問題。有多個粗糙接觸點、而且各接觸點滑動的先後次序有影響的問題，是剛體靜力學的上佳入門。Kleppner 與 Kolenkow 的 *An Introduction to Mechanics*（力學入門教科書）對摩擦和剛體平衡有仔細的論述。

詳細解答

STEP II 2003，第 10 題

力學 • 涉及摩擦力的約束運動

STEP II 2003，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/03-s2-q10

關鍵思路

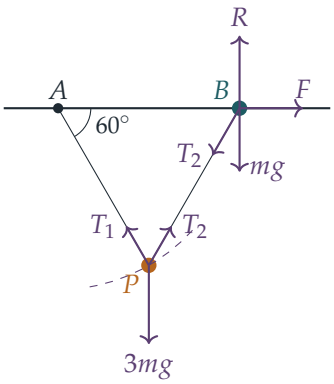
開始時，$AP = PB = AB = \ell$，所以 APB 是等邊三角形（equilateral triangle）。由於 AP 一直拉緊，P 沿着以 A 為圓心的圓運動。整個系統由靜止開始，所以沒有向心（centripetal）項，P 的初始加速度（acceleration）必定沿着該圓的切線（tangent）方向，這就確定了它的方向。其次，P 始終位於以 AB 為底的等腰三角形（isosceles triangle）的頂點，所以 P 的水平位置永遠是 B 的一半。至於最後一部分，分別對 B 和 P 寫出牛頓第二定律（Newton's second law）。對 P 沿其加速度方向（垂直於 AP）分解，AP 中的張力（tension）便完全不會出現。

解答

我們把 B 和 P 視為質點（particle），並把繩視為輕而不可伸長（light and inextensible）。繩的兩半都保持拉緊。取 A 為原點，金屬線沿 x 軸，B 最初位於 $(\ell, 0)$。設 θ 為角 PAB。由於 $AP = PB = \ell$，

$$B = (2\ell\cos\theta,\, 0), \qquad P = (\ell\cos\theta,\, -\ell\sin\theta),$$

而開始時 $\theta = \frac{\pi}{3}$。



P 的加速度方向

由於 $AP = \ell$ 固定不變，P 沿着以 A 為圓心、半徑為 ℓ 的圓運動（圖中以虛線表示）。圓上一點的加速度有一個指向圓心的分量 $\ell\dot{\theta}^2$，以及一個沿切線的分量 $\ell\ddot{\theta}$。開始時 $\dot{\theta} = 0$，因為系統是由靜止釋放的。所以 P 的初始加速度沿着切線方向，即垂直於 AP。

AP 與水平線成 $60°$ 角，所以 AP 的垂直線與水平線成 $30°$ 角。若加速度的大小為 a，則

$$a_1 = a\cos 30° = \tfrac{\sqrt{3}}{2}a, \qquad a_2 = a\sin 30° = \tfrac{1}{2}a,$$

因此 $a_1 = \sqrt{3}\,a_2$，正如所求。

以 θ 表示：對 P 的位置求導，並令 $\dot{\theta} = 0$，可得初始加速度為 $\ell\ddot{\theta}(-\sin\theta, -\cos\theta)$。當 $\ddot{\theta} > 0$ 時，P 向下並朝 A 的方向運動。其大小為 $a = \ell\ddot{\theta}$。

B 的加速度

B 的 x 坐標永遠恰好是 P 的兩倍，因為 P 位於 AB 的垂直平分線（perpendicular bisector）上。求導兩次可知，B 的加速度永遠是 P 的水平加速度的兩倍。所以 B 的初始加速度大小為 $2a_1$，方向指向 A，正如所求。

想一想 為甚麼可以說 B 朝 A 移動？摩擦力又朝哪個方向作用？

P 下落時，θ 增大，所以 $x_B = 2\ell\cos\theta$ 減小。因此 B 朝 A 滑動，而作用於 B 的摩擦力（friction）則背離 A（沿正 x 方向）。這裏我們假設 B 確實會移動。題目給出摩擦力為 $\frac{\sqrt{3}}{6}R$，這正是滑動時的數值，從而確認了這一點。

初始加速度的大小

設 T_1 和 T_2 分別為 AP 和 PB 中的張力。設 R 為作用於 B 的法向反作用力（normal reaction），$F = \frac{\sqrt{3}}{6}R$ 為摩擦力。以 a 表示 P 的初始加速度的大小，故 $a_1 = \frac{\sqrt{3}}{2}a$，而 B

的加速度為 $2a_1 = \sqrt{3}\,a$，方向指向 A。

珠子 B。繩 BP 以水平線以下 60° 的方向把 B 拉向 A 的一方。B 沒有垂直方向的加速度，所以沿垂直方向分解可得

$$R = mg + \tfrac{\sqrt{3}}{2}T_2.$$

沿水平方向分解，取指向 A 的方向為正：

$$m(\sqrt{3}\,a) = \tfrac{1}{2}T_2 - \tfrac{\sqrt{3}}{6}\left(mg + \tfrac{\sqrt{3}}{2}T_2\right) = \tfrac{1}{4}T_2 - \tfrac{\sqrt{3}}{6}mg,$$

所以

$$T_2 = 4\sqrt{3}\,ma + \tfrac{2\sqrt{3}}{3}mg. \tag{1}$$

質點 P。沿 P 的加速度方向分解，即向下並朝向 A、在水平線以下 30° 的方向。這個方向垂直於 AP，所以 T_1 沒有任何貢獻。重量 $3mg$ 與這個方向成 60° 角，所以它的貢獻是 $3mg\cos 60^\circ = \tfrac{3}{2}mg$。張力 T_2 沿 PB 作用，即向上並背離 A、在水平線以上 60° 的方向。它與加速度方向成 150° 角，所以它的貢獻是 $T_2\cos 150^\circ = -\tfrac{\sqrt{3}}{2}T_2$。因此

$$3ma = \tfrac{3}{2}mg - \tfrac{\sqrt{3}}{2}T_2. \tag{2}$$

把 (1) 代入 (2)：

$$3ma = \tfrac{3}{2}mg - \tfrac{\sqrt{3}}{2}\left(4\sqrt{3}\,ma + \tfrac{2\sqrt{3}}{3}mg\right) = \tfrac{3}{2}mg - 6ma - mg,$$

$$9ma = \tfrac{1}{2}mg,$$

所以 P 的初始加速度的大小為

$$a = \frac{g}{18},$$

正如所求。

這個答案前後一致：由 (1) 得 $T_2 = \tfrac{8\sqrt{3}}{9}mg > 0$，所以繩 PB 是拉緊的。於是 $R = \tfrac{7}{3}mg > 0$。對 P 沿水平方向分解，得 $T_1 = T_2 + 6ma_1 = \tfrac{19\sqrt{3}}{18}mg > 0$，所以 AP 也是拉緊的。

驗算

- P 的垂直方向方程，我們之前沒有用到：$\tfrac{\sqrt{3}}{2}(T_1 + T_2) - 3mg = \tfrac{35}{12}mg - 3mg = -\tfrac{1}{12}mg$。這等於 $-3ma_2 = -3m\cdot\tfrac{g}{36}$，與預期相符。
- 數值上，我們對 B 和 P 的位置求導，得到 $a_1/a_2 \approx 1.7320508$ 及 $a_B/a_1 = 2$。然後取 $g = 9.81$，從三條力的方程解出 T_1、T_2 和 $\ddot{\theta}$。結果是 $|a| = 0.545 = g/18$，並有 $T_2/mg \approx 1.5396$ 及 $R/mg \approx 2.3333$。
- 若沒有摩擦力，(1) 會變成 $T_2 = 2\sqrt{3}\,ma$，而 (2) 會給出 $6ma = \tfrac{3}{2}mg$，所以 $a = g/4$。摩擦力理應令運動減慢，而的確 $g/18 < g/4$。

常見錯誤

- 加入了向心項。在釋放的一刻 $\dot{\theta} = 0$，所以加速度純粹沿切線方向。
- 忘記了 T_2 會把 B 向下拉，令 R 增大，摩擦力亦隨之增大。取 $R = mg$ 會得出錯誤答案。
- 對 P 沿水平和垂直方向分解。這樣做是可行的，但會引入 T_1，並需要多一條方程。沿垂直於 AP 的方向分解便可把它消去。

STEP 之外

這是一個只有一個自由度（degree of freedom）的約束系統（constrained system）：一切都由單一個角 θ 決定。把位置以 θ 表示然後求導，正是拉格朗日力學（Lagrangian mechanics）的第一步。在拉格朗日力學中，像 T_1 這類約束力（constraint force）是以有系統的方法消去的，而不是靠巧妙地選擇方向。摩擦力並不能妥貼地納入這個框架，這也是這類題目適合用來練習牛頓力學方法的原因之一。Kleppner 與 Kolenkow 所著的 *An Introduction to Mechanics*（力學入門）對約束運動和非勻速圓周運動（non-uniform circular motion）有細緻的講解。

詳細解答

STEP II 2003，第 11 題

力學 • 拋體運動

STEP II 2003，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：

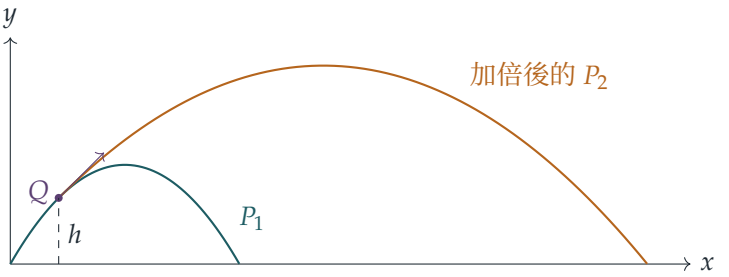
step.maths.org/questions/03-s2-q11

關鍵思路

把速度的水平分量和垂直分量分開處理。運動方向與水平線成 45°，正正就是兩個分量相等的時候，由此可直接求出第一個時刻。速率加倍時方向不變，所以兩個分量都加倍，而餘下的飛行過程就是一個從已知高度出發的新拋體問題。

解答

我們把每個質點（particle）都當作只受重力作用的點狀拋體（projectile），不計空氣阻力。記 $s = \sin\alpha$ 及 $c = \cos\alpha$。由於 $\alpha > 45°$，我們有 $s > c > 0$。



此草圖取 $\alpha = 60°$。兩個質點都沿青綠色的路徑運動至 Q，運動方向在該處首次與水平線成 45°。其後 P_2 沿琥珀色的路徑運動。

P_1 的飛行時間

在時間 t 的垂直位移（displacement）是 $y = Vs\,t - \frac{1}{2}gt^2 = t\left(Vs - \frac{1}{2}gt\right)$。它再次為零的時刻是

$$T_1 = \frac{2V\sin\alpha}{g}.$$

首個 45° 時刻

在時間 t，速度的分量（component）是

$$\dot{x} = Vc, \qquad \dot{y} = Vs - gt.$$

水平分量始終不變。發射時運動方向比 45° 更陡，因為 $\dot{y} = Vs > Vc = \dot{x}$。隨着 $\dot{y}$ 減小，當 $\dot{y} = \dot{x}$ 時，角度首次降至 45°，此時質點仍在上升：

$$Vs - gt_1 = Vc \quad\Longrightarrow\quad t_1 = \frac{V(s-c)}{g}.$$

（下降途中當 $\dot{y} = -Vc$ 時，運動方向會再次成 45°，但那是較後的事。）由於 $\frac{1}{2}T_1 = Vs/g$，

$$t_1 = \frac{Vs}{g}\left(1 - \frac{c}{s}\right) = \tfrac{1}{2}(1 - \cot\alpha)\,T_1,$$

正如所求。

P_2 的飛行

在時間 t_1 之前，P_2 的運動與 P_1 完全相同。在 t_1 時，速度的兩個分量都等於 Vc，而高度是

$$\begin{aligned} h &= Vs\,t_1 - \tfrac{1}{2}gt_1^2 = \frac{V^2}{g}\left[s(s-c) - \tfrac{1}{2}(s-c)^2\right] \\ &= \frac{V^2}{2g}(s-c)\left[2s - (s-c)\right] = \frac{V^2}{2g}\,(s^2 - c^2). \end{aligned}$$

由於 $s > c$，這是正數。

> **想一想** 「速率加倍」對速度的每個分量有甚麼影響？

速率改變了，但方向不變，所以整個速度向量（velocity vector）乘以 2。兩個分量都變成 $2Vc$：質點從高度 h 出發，方向在水平線之上 45°，速率為 $2\sqrt{2}\,Vc$。

設 τ 為由速率加倍至 P_2 着地所經過的時間。以 Q 為起點、向上為正量度垂直位移，着地即位移為 $-h$：

$$2Vc\,\tau - \tfrac{1}{2}g\tau^2 = -\frac{V^2}{2g}(s^2 - c^2).$$

兩邊乘以 $-2/g$，並記 $\tau = (V/g)u$，其中 u 是無量綱（dimensionless）的量：

$$u^2 - 4cu - (s^2 - c^2) = 0 \quad \Longrightarrow \quad u = 2c \pm \sqrt{4c^2 + s^2 - c^2} = 2c \pm \sqrt{s^2 + 3c^2}.$$

兩根之積是 $-(s^2 - c^2) < 0$，所以其中一個根是負數。我們需要 $\tau > 0$，故取正號：

$$\tau = \frac{V}{g}\left(2c + \sqrt{s^2 + 3c^2}\right).$$

因此

$$T_2 = t_1 + \tau = \frac{V}{g}\left(s - c + 2c + \sqrt{s^2 + 3c^2}\right) = \frac{V}{g}\left(s + c + \sqrt{s^2 + 3c^2}\right).$$

除以 $\frac{1}{2}T_1 = Vs/g$，並利用 $\sqrt{s^2 + 3c^2}/s = \sqrt{1 + 3c^2/s^2}$（因 $s > 0$ 而成立）：

$$\frac{2T_2}{T_1} = \frac{s + c + \sqrt{s^2 + 3c^2}}{s} = 1 + \cot\alpha + \sqrt{1 + 3\cot^2\alpha},$$

這就是所要求的結果。

驗算

- 當 $\alpha \to 90^\circ$ 時，$\cot\alpha \to 0$，比值趨向 2。這合乎常理：45° 時刻此時就是飛行的最高點，該處速率為零，所以加倍毫無作用，$T_2 = T_1$。
- 當 $\alpha = 45^\circ$（邊界情況）時，加倍在發射一刻發生，所以 P_2 其實就是以 $2V$ 拋出，而 $T_2 = 2T_1$。公式給出 $1 + 1 + 2 = 4$，兩者吻合。
- 當 $\alpha = 60^\circ$ 時，$t_1/T_1 = \frac{1}{2}(1 - 1/\sqrt{3}) \approx 0.2113$，而 $2T_2/T_1 = 1 + 1/\sqrt{3} + \sqrt{2} \approx 2.9916$。對 P_2 的飛行以數值方法逐步模擬，隨機選取 2000 組 α 和 V，結果均與 T_2 的公式相符。

常見錯誤

- 只把垂直分量加倍，或只把水平分量加倍。方向不變，所以兩個分量都加倍。
- 讓 P_2 從地面重新開始運動，而不是從高度 h 開始。
- 沒有檢查正負號就取了 τ 的負根。

STEP 之外

速度的瞬時改變是一個衝量（impulse）：一個極大的力作用了極短的時間。大學一年級的力學正是用衝量來處理碰撞、火箭點火和踢球等情況。把一條軌跡分成若干段，各段

之間以衝量相接，正是規劃軌道機動（orbital manoeuvre）的方法：引擎短暫點火，就能把太空船從一條圓錐曲線（conic）軌道轉移到另一條。Kleppner 與 Kolenkow 的 *An Introduction to Mechanics*（力學入門）對衝量和動量（momentum）有嚴謹的闡述，並收錄了許多這類題目。

詳細解答

STEP II 2003，第 12 題

概率與統計 • 分支過程

STEP II 2003，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/03-s2-q12

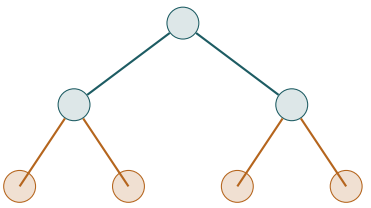
本題在 *Beyond STEP* 第 12.15 節滅絕還是爆發中有完整的引導式講解：https://beyondstep.keithmcnulty.org/zh-hant/ch12/branching.html

關鍵思路

這是一個分支過程（branching process）：每一代都由上一代產生，而每顆粒子各自獨立地分裂。整道題只靠一個事實：若 m 顆粒子各自獨立地產生後代，而每顆的後代數目服從 $\mathrm{B}(2, p)$，則後代總數服從 $\mathrm{B}(2m, p)$。因此，無論是求概率還是求期望值（expectation），我們都以上一代的大小作為條件。

解答

我們按照模型的原意，假設不同粒子的衰變彼此獨立，亦與過去的情況獨立。記 $X_0 = 1$，代表在 $0 < t < T$ 期間存活的那一顆粒子。在 $t = nT$ 時，存活的 X_{n-1} 顆粒子各自被自己的後代取代，所以 X_n 就是第 $n-1$ 代的後代總數。



$X_0 = 1$

$X_1 = 2$

$X_2 \leq 4$

圖中所示為 $X_1 = 2$ 的情況：這兩顆粒子合共最多有四個後代，而四個可能的後代各自

以概率 p 獨立地誕生。

聯合概率

由於 $X_1 \sim \mathrm{B}(2,p)$，我們有 $\mathrm{P}(X_1=2)=p^2$。已知 $X_1=2$ 時，兩顆粒子各自獨立地產生後代，每顆的後代數目服從 $\mathrm{B}(2,p)$，所以 X_2 就是 4 次獨立試驗中成功的次數，而每次成功的概率為 p；即 $X_2 \mid X_1=2 \sim \mathrm{B}(4,p)$。因此

$$\mathrm{P}(X_1=2 \text{ 及} X_2=2)=\mathrm{P}(X_1=2)\,\mathrm{P}(X_2=2 \mid X_1=2)=p^2\cdot\binom{4}{2}p^2q^2=6p^4q^2,$$

正如題目所求。

條件概率

要運用條件概率（conditional probability）的定義，我們需要 $\mathrm{P}(X_2=2)$，這可以透過以 X_1 為條件求得：

- 若 $X_1=0$，整個羣體已經滅絕，所以 $X_2=0$；
- 若 $X_1=1$（概率為 $2pq$），則 $X_2\sim\mathrm{B}(2,p)$，而 $\mathrm{P}(X_2=2 \mid X_1=1)=p^2$；
- 若 $X_1=2$，其貢獻就是剛才求得的 $6p^4q^2$。

所以

$$\mathrm{P}(X_2=2)=2pq\cdot p^2+6p^4q^2=2p^3q\,(1+3pq),$$

因此，當 $0<p<1$ 時，

$$\mathrm{P}(X_1=2 \mid X_2=2)=\frac{6p^4q^2}{2p^3q(1+3pq)}=\frac{3pq}{1+3pq}.$$

令上式等於 $\frac{9}{25}$，得 $75pq=9+27pq$，所以 $48pq=9$，即 $pq=\frac{3}{16}$。代入 $q=1-p$，這就是

$$p^2-p+\tfrac{3}{16}=0 \quad\Longleftrightarrow\quad (p-\tfrac{1}{4})(p-\tfrac{3}{4})=0,$$

所以 $p=\frac{1}{4}$ 或 $p=\frac{3}{4}$。（$p=0$ 和 $p=1$ 這兩種情況須排除：當 $p=0$ 時，事件 $X_2=2$ 不可能發生；當 $p=1$ 時，條件概率為 0。）

> **想一想** 為甚麼 p 的兩個值會給出相同的答案？

條件概率只透過 $pq=p(1-p)$ 依賴於 p，而把 p 換成 $1-p$ 時，這個量保持不變。所以方程的根必定成對出現，即 p 與 $1-p$，在這裏就是 $\frac{1}{4}$ 和 $\frac{3}{4}$。

期望值

設 $X_{n-1} = m$。這 m 顆粒子每顆平均產生 $\mathrm{E}(X) = 2p$ 個後代，所以下一代大小的期望值為 $2pm$：

$$\mathrm{E}(X_n \mid X_{n-1} = m) = 2pm.$$

對 m 的所有可能取值求平均（全期望值定律，law of total expectation），

$$\mathrm{E}(X_n) = \sum_m \mathrm{P}(X_{n-1} = m)\, 2pm = 2p\,\mathrm{E}(X_{n-1}).$$

用文字來說：在某一段時間內存活的每顆粒子，平均為下一段時間貢獻 $2p$ 顆粒子。由於 $\mathrm{E}(X_1) = 2p$（或 $\mathrm{E}(X_0) = 1$），這個遞推關係（recurrence）給出一個等比數列（geometric sequence）：

$$\mathrm{E}(X_n) = (2p)^n.$$

極限

當 $p = \frac{1}{4}$ 時，我們得到 $\mathrm{E}(X_n) = \left(\frac{1}{2}\right)^n \to 0$（當 $n \to \infty$），因為 $0 < \frac{1}{2} < 1$。

當 $p = \frac{3}{4}$ 時，我們得到 $\mathrm{E}(X_n) = \left(\frac{3}{2}\right)^n \to +\infty$，因為 $\frac{3}{2} > 1$。

由此可見，p 的這兩個值在條件概率中看似毫無分別，長遠表現卻截然不同：平均而言，當 $p = \frac{1}{4}$ 時羣體逐漸消亡，而當 $p = \frac{3}{4}$ 時羣體則急劇膨脹。

驗算

- 取 $p = \frac{3}{4}$：$6p^4q^2 = \frac{243}{2048} \approx 0.1187$，而模擬 10^6 個兩代羣體所得的結果是 0.1186。
- 當 $p = \frac{1}{4}$ 時，$pq = \frac{3}{16}$，所以 $3pq = \frac{9}{16}$，而 $\dfrac{9/16}{25/16} = \frac{9}{25}$，與預期相符。
- 對若干個 p 值計算 $X_1, \ldots, X_5$ 的精確分佈（已知上一代時，每一代服從 $\mathrm{B}(2m, p)$），所得的平均值恰好是 $(2p)^n$。

常見錯誤

- 在求 $\mathrm{P}(X_2 = 2)$ 時，遺漏了 $X_1 = 1$、$X_2 = 2$ 這條路徑；至於經過 $X_1 = 0$ 的路徑，其貢獻確實為零。
- 把 $X_2 \mid X_1 = 2$ 當作 $2 \times \mathrm{B}(2, p)$（即一顆粒子後代數目的兩倍），而不是兩個獨立副本之和，後者才是 $\mathrm{B}(4, p)$。

STEP 之外

這是一個高爾頓–沃森（Galton–Watson）分支過程，最初是為了解釋姓氏為何會消失而提出的。後代數目的平均值 $\mu = 2p$ 決定了一切：若 $\mu \leq 1$，羣體以概率 1 滅絕；若 $\mu > 1$，羣體則有正概率永遠存活下去。即使 $p = \frac{3}{4}$，羣體仍然有 $\frac{1}{9}$ 的概率滅絕，這個數值是 $s = (q + ps)^2$ 的最小根，而這條方程來自概率母函數（probability generating function）。Grimmett 與 Welsh 所著的 *Probability: An Introduction*（概率論入門教材）以母函數處理分支過程，其講法正好可以直接承接本題。

詳細解答

STEP II 2003，第 13 題

概率與統計 • 截尾泊松分佈

STEP II 2003，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/03-s2-q13

關鍵思路

這是一個剔除了 0 這個值的泊松分佈（Poisson distribution），餘下的概率按比例放大，使它們的總和仍然是 1。因此，我們需要的每一個和，都是缺了 $k = 0$ 一項的泊松和；而在計算平均值（mean）和方差（variance）時，$k = 0$ 一項本來就沒有任何貢獻。所以我們可以借用熟悉的泊松分佈結果 $\sum k\,p_k = \lambda$ 和 $\sum k^2 p_k = \lambda + \lambda^2$，然後乘以 A 就可以了。

解答

常數 A

記 $p_k = \dfrac{\lambda^k \mathrm{e}^{-\lambda}}{k!}$ 為普通 $\mathrm{Po}(\lambda)$ 的概率。由於 $\sum_{k\ge 0} \lambda^k/k! = \mathrm{e}^{\lambda}$，

$$\sum_{k=1}^{\infty} p_k = \mathrm{e}^{-\lambda}\left(\mathrm{e}^{\lambda} - 1\right) = 1 - \mathrm{e}^{-\lambda}.$$

概率 $\mathrm{P}(X = k) = A p_k$ 的總和必須是 1，所以 $A(1 - \mathrm{e}^{-\lambda}) = 1$，即

$$A = (1 - \mathrm{e}^{-\lambda})^{-1},$$

這正是題目要求的結果。留意當 $\lambda > 0$ 時 $0 < 1 - \mathrm{e}^{-\lambda} < 1$，所以 $A > 1$：剩下來的概率 p_0 已經分攤給 $k \ge 1$ 的各個值。

平均值

$\sum k\,p_k$ 中 $k=0$ 的一項是零，所以

$$\sum_{k=1}^{\infty} k\,p_k = \sum_{k=0}^{\infty} k\,p_k = \lambda,$$

即 $\mathrm{Po}(\lambda)$ 的平均值。（直接計算：$\sum_{k\geq 1} k\lambda^k \mathrm{e}^{-\lambda}/k! = \lambda \sum_{k\geq 1} \lambda^{k-1}\mathrm{e}^{-\lambda}/(k-1)! = \lambda$。）因此

$$\mu = \mathrm{E}(X) = A\lambda = \frac{\lambda}{1-\mathrm{e}^{-\lambda}}.$$

方差

同樣地，$\sum k^2 p_k$ 中 $k=0$ 的一項是零，而對於 $\mathrm{Po}(\lambda)$，我們有 $\mathrm{E}(Y^2) = \mathrm{Var}(Y) + (\mathrm{E}\,Y)^2 = \lambda + \lambda^2$。如果你不想直接引用這個結果，可以寫 $k^2 = k(k-1)+k$：

$$\sum_{k\geq 2} k(k-1)\frac{\lambda^k \mathrm{e}^{-\lambda}}{k!} = \lambda^2 \sum_{k\geq 2} \frac{\lambda^{k-2}\mathrm{e}^{-\lambda}}{(k-2)!} = \lambda^2, \qquad \text{所以} \qquad \sum_{k\geq 1} k^2 p_k = \lambda^2 + \lambda.$$

於是

$$\begin{aligned} \mathrm{E}(X^2) &= A\lambda(1+\lambda) = \mu(1+\lambda), \\ \mathrm{Var}(X) &= \mathrm{E}(X^2) - \mu^2 = \mu(1+\lambda) - \mu^2 = \mu(1-\mu+\lambda), \end{aligned}$$

這正是題目要求的結果。

不等式

> **想一想** 嚴格不等式 $\mu < 1+\lambda$ 可以從哪裏來？

從方差而來。X 取 $1, 2, 3, \ldots$ 中每一個值的概率都是正數，所以它不是常數，而一個非常數的隨機變量（random variable）的方差必定嚴格大於零（因為 $\mathrm{Var}(X) = \mathrm{E}\big((X-\mu)^2\big)$，而 $(X-\mu)^2 > 0$ 的概率是正數）。所以

$$\mu(1-\mu+\lambda) > 0.$$

由於 $\mu > 0$（它是一個取正值的隨機變量的平均值），我們可以兩邊除以 μ，得到 $1-\mu+\lambda > 0$，即 $\mu < 1+\lambda$。

至於下界，$\mu = A\lambda$，其中 $A > 1$ 而 $\lambda > 0$，所以 $\mu > \lambda$。兩者合起來，

$$\lambda < \mu < 1+\lambda.$$

$\lambda = 100$ 時的正態近似

e^{-100} 小得驚人（大約是 4×10^{-44}），所以在任何合理的準確度下

$$\mu = \frac{100}{1 - e^{-100}} \approx 100, \qquad \mathrm{Var}(X) = \mu(101 - \mu) \approx 100 \times 1 = 100.$$

所以我們用 $Y \sim N(100, 10^2)$ 來近似 X，這就是正態近似（normal approximation）。由於 X 只取整數值，我們要使用連續性校正（continuity correction）：事件 $X = 100$ 對應 $99.5 < Y < 100.5$。以 $Z = (Y - 100)/10$ 標準化，

$$\begin{aligned} P(X = 100) &\approx P(-0.05 < Z < 0.05) \\ &= 2\Phi(0.05) - 1 \\ &\approx 2(0.51994) - 1 = 0.0399, \end{aligned}$$

所以 $P(X = 100) \approx 0.04$（準確至小數點後 2 位）。

驗算

- 當 $\lambda \to 0$ 時，$\mu = \lambda/(1 - e^{-\lambda}) \to 1$，而方差 $\to 0$：幾乎所有概率都集中在 $k = 1$，這是理所當然的。這也與不等式 $\lambda < \mu < 1 + \lambda$ 把 μ 夾向 1 的情況吻合。
- 對許多隨機選取的 λ 以數值方法求級數的和，所得結果與 A、μ 和 $\mu(1 - \mu + \lambda)$ 一致，準確至機器精度。
- 當 $\lambda = 100$ 時，準確的概率是 $\dfrac{100^{100} e^{-100}}{100!\,(1 - e^{-100})} \approx 0.03986$，而正態近似給出 0.03988。兩者四捨五入後都是 0.04。

常見錯誤

- 忘記連續性校正，算出 $P(Y = 100) = 0$，或者未加解釋就使用密度函數的值。
- 只證明了 $\mu \leq 1 + \lambda$：要得到嚴格不等式，必須指出 X 不是常數。
- 除以 μ 時沒有說明 $\mu > 0$。

STEP 之外

X 服從零截尾泊松（zero-truncated Poisson）分佈，統計學家會在無法觀察到計數為零的情況下使用它（例如經過檢查站的汽車的乘載人數：沒有一輛車的乘載人數是零）。正態近似之所以行得通，是因為一個 Po(100) 變量是 100 個獨立的 Po(1) 變量之和，所

以中心極限定理（central limit theorem）適用；在這個尺度下，截尾的影響根本看不出來。Grimmett and Welsh, *Probability: An Introduction*（概率論入門教材）講解了母函數（generating functions）和中心極限定理，它們正是這道題前後兩部分背後的理論。

詳細解答

STEP II 2003，第 14 題

概率與統計 • 假設檢驗的錯誤

> **STEP II 2003，第 14 題**
>
> 繼續閱讀之前，請先看題目（原題為英文）。題目載於 STEP Support Programme 的試題資料庫：
>
> step.maths.org/questions/03-s2-q14

關鍵思路

擲出第一個 6 所需的投擲次數服從幾何分佈（geometric distribution），而它的尾部概率形式極為簡潔：事件「需要多於 M 次投擲」其實就是「首 M 次投擲全部失敗」。因此兩個錯誤概率都純粹是 M 次冪，而要在兩者之間消去 M，只需取對數（logarithm）即可。

解答

兩個錯誤概率

把各次投擲視為互相獨立（independent），每次擲出 6 的概率同為 p。設 X 為直至首次擲出 6 為止（包括該次）的投擲次數，則

$$X > M \iff \text{首 } M \text{ 次投擲中沒有一次是 6}, \qquad \text{所以} \qquad \mathrm{P}(X > M) = (1-p)^M,$$

這裏用到了各次投擲的獨立性。（毋須求幾何級數（geometric series）$\sum p(1-p)^{k-1}$ 的和：直接考慮事件本身更快捷。）

這個檢驗恰好在 $X > M$ 時接受 H_1。因此 α，即實際上 $p = B$ 時接受 H_1 的概率，是

$$\alpha = (1-B)^M,$$

而 β，即實際上 $p = A$ 時接受 H_0（亦即 $X \leq M$）的概率，是

$$\beta = 1 - \mathrm{P}(X > M \mid p = A) = 1 - (1-A)^M.$$

消去 M

我們想把 $(1-A)^M$ 寫成 $\alpha = (1-B)^M$ 的冪。由於 $0 < B < 1$，$\ln(1-B)$ 有定義且不等於零，所以對 α 取對數可得

$$M = \frac{\ln \alpha}{\ln(1-B)}.$$

於是

$$(1-A)^M = \mathrm{e}^{M\ln(1-A)} = \exp\left(\frac{\ln(1-A)}{\ln(1-B)}\ln\alpha\right) = \alpha^K, \qquad \text{其中} \quad K = \frac{\ln(1-A)}{\ln(1-B)}.$$

因此

$$\beta = 1-\alpha^K, \qquad K = \frac{\ln(1-A)}{\ln(1-B)},$$

而 K 只取決於 A 和 B，與 M 無關，正如題目所求。

? **想一想** 關於 K 的大小，我們可以說些甚麼？

由於 $0 < A < B < 1$，我們有 $1 > 1-A > 1-B > 0$，所以兩個對數都是負數，且 $\ln(1-B) < \ln(1-A) < 0$。兩者相除，可知 K 是正數，並且

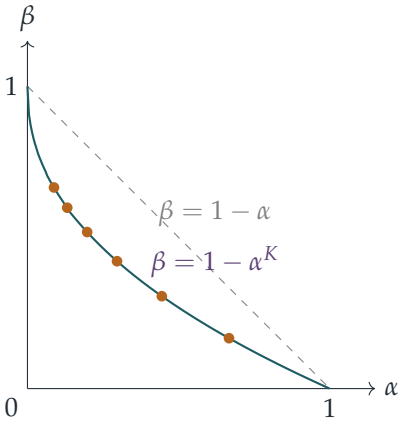
$$0 < K < 1.$$

圖像的形狀正是由此決定。

β 對 α 的圖像

當 M 依次取 $1, 2, 3, \ldots$ 時，$\alpha = (1-B)^M$ 在 $(0,1)$ 內取值，並遞減趨向 0；嚴格來說，圖像是一列點，全部位於曲線 $\beta = 1-\alpha^K$，$0 \le \alpha \le 1$ 之上。對於這條曲線：

- 它通過 $(0,1)$ 和 $(1,0)$；
- $\dfrac{\mathrm{d}\beta}{\mathrm{d}\alpha} = -K\alpha^{K-1} < 0$，所以 β 隨 α 增加而減少：令其中一種錯誤較難發生，另一種錯誤就會較易發生；
- 由於 $K - 1 < 0$，當 $\alpha \to 0^+$ 時斜率（gradient）趨向 $-\infty$（曲線從 $(0,1)$ 垂直出發），而在 $\alpha = 1$ 處斜率等於 $-K$，介乎 -1 與 0 之間；
- $\dfrac{\mathrm{d}^2\beta}{\mathrm{d}\alpha^2} = K(1-K)\alpha^{K-2} > 0$，所以曲線是凸的（convex）；
- 對於 $0 < \alpha < 1$ 和 $0 < K < 1$，我們有 $\alpha^K > \alpha$，所以 $\beta < 1-\alpha$：曲線位於連接 $(0,1)$ 與 $(1,0)$ 的直線下方。



圖中的曲線是按 $A = \frac{1}{6}$、$B = \frac{1}{3}$（所以 $K \approx 0.45$）繪畫的，並標出了 $M = 1, \ldots, 6$ 所對應的點；M 增加時，點會向左上方移動。

驗算

- 取 $A = \frac{1}{6}$、$B = \frac{1}{3}$、$M = 4$。直接計算，$\alpha = (\frac{2}{3})^4 \approx 0.1975$，$\beta = 1 - (\frac{5}{6})^4 \approx 0.5177$。以 $K = \ln \frac{5}{6} / \ln \frac{2}{3} \approx 0.4497$ 計算，$1 - \alpha^K \approx 0.5177$，結果相同。把這個檢驗模擬 2×10^5 次，所得結果為 0.198 和 0.518。
- 極端情況的表現亦合乎常理：若 A 接近 B，則 $K \approx 1$ 而 $\beta \approx 1 - \alpha$，這個檢驗比起不理會投擲結果、隨機作決定好不了多少；若 A 極小，則 $K \approx 0$，對於任何合理的 α，都可以令 β 變得很小。

常見錯誤

- 混淆兩個假設（hypothesis）：這裏 H_0 是較大的值 B，所以等了很久才擲出 6，是不利於 H_0 的證據。
- 把 K 上下顛倒，寫成 $\ln(1 - B) / \ln(1 - A)$，其值大於 1；檢查 $\beta < 1 - \alpha$（任何合理的檢驗都勝過亂猜）便可發現這個錯誤。

STEP 之外

β 對 α 的曲線，就是兩類錯誤之間的取捨曲線（trade-off curve），與統計學和機器學習中廣泛使用的 ROC 曲線（ROC curve）關係密切。奈曼–皮爾遜引理（Neyman–Pearson lemma）表明，對於像這樣的兩個簡單假設（simple hypotheses），基於似然比（likelihood

ratio）的檢驗能就每個 α 給出最小可能的 β；這裏似然比對 X 是單調的（monotonic），所以「當 X 很大時拒絕 H_0」這個檢驗其實是最優的。想進一步了解這些概念，可先閱讀 Grimmett 與 Welsh 的 *Probability: An Introduction*（概率論入門）打好基礎，然後再讀任何一門數理統計（mathematical statistics）的入門課程。

詳細解答

STEP III 2003，第 1 題

純數學 • 反雙曲函數積分法

STEP III 2003，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/03-s3-q1

關鍵思路

頭兩個結果是細心運用鏈式法則（chain rule）的計算。它們真正有用之處，是提供了一個範本：每個反函數（inverse function）的導數（derivative）都具有 $\dfrac{\text{常數}}{(\text{一次式})\sqrt{\text{一次式}}}$ 的形式，而這正正是兩個被積函數（integrand）的形式。因此，對每個積分，我們選取 a 和 b 去配合這個範本。根號外的括號告訴我們該用哪一條公式。若它是 $x+b$，即常數較大的那個，便用 arcsin；若它是 $x+a$，即常數較小的那個，便用 arcosh。

解答

兩個導數

設 $y = \dfrac{x+a}{x+b}$。由商法則（quotient rule），

$$\frac{\mathrm{d}y}{\mathrm{d}x} = \frac{(x+b)-(x+a)}{(x+b)^2} = \frac{b-a}{(x+b)^2}.$$

由於 $x+a>0$ 及 $x+b>0$，可知 $y>0$；又由於 $b>a$，分子小於分母，所以 $0<y<1$，而 $\arcsin y$ 在該處可微（differentiable）。接着，

$$1-y^2 = \frac{(x+b)^2-(x+a)^2}{(x+b)^2} = \frac{(b-a)(2x+a+b)}{(x+b)^2},$$

這裏用了平方差（difference of two squares）公式。由於 $x+b>0$，$(x+b)^2$ 的平方根就是 $x+b$ 本身，所以

$$\sqrt{1-y^2} = \frac{\sqrt{b-a}\,\sqrt{a+b+2x}}{x+b}.$$

由鏈式法則，

$$\frac{\mathrm{d}}{\mathrm{d}x}\arcsin\left(\frac{x+a}{x+b}\right) = \frac{1}{\sqrt{1-y^2}}\cdot\frac{\mathrm{d}y}{\mathrm{d}x} = \frac{x+b}{\sqrt{b-a}\,\sqrt{a+b+2x}}\cdot\frac{b-a}{(x+b)^2}$$
$$= \frac{\sqrt{b-a}}{(x+b)\sqrt{a+b+2x}},$$

因為 $(b-a)/\sqrt{b-a} = \sqrt{b-a}$。這就是所求的結果。

至於 arcosh，設 $z = \dfrac{x+b}{x+a}$。這裏 $z > 1$，因為 $x+b > x+a > 0$。於是

$$\frac{\mathrm{d}z}{\mathrm{d}x} = \frac{(x+a)-(x+b)}{(x+a)^2} = -\frac{b-a}{(x+a)^2}, \qquad z^2-1 = \frac{(b-a)(2x+a+b)}{(x+a)^2}.$$

由於 $x+a>0$，$\sqrt{z^2-1} = \sqrt{b-a}\,\sqrt{a+b+2x}\,/(x+a)$，因此

$$\frac{\mathrm{d}}{\mathrm{d}x}\operatorname{arcosh}\left(\frac{x+b}{x+a}\right) = \frac{x+a}{\sqrt{b-a}\,\sqrt{a+b+2x}}\cdot\left(-\frac{b-a}{(x+a)^2}\right) = -\frac{\sqrt{b-a}}{(x+a)\sqrt{a+b+2x}}.$$

兩個結果看來幾乎一樣。分別只在於正負號，以及根號外是哪一個括號。

想一想 arcsin 的計算實際上用到了哪些假設？

我們需要 $x+b>0$ 來化簡 $\sqrt{(x+b)^2}$，也需要 $1-y^2>0$。從因式分解後的形式可見，在 $b>a$ 的前提下，$1-y^2>0$ 恰好在 $a+b+2x>0$ 時成立。我們從沒有用過 $x+a>0$ 本身。所以，只要 $x+b>0$ 及 $a+b+2x>0$，arcsin 的公式便成立，即使 $x+a$ 是負數亦然，此時 y 位於 $(-1,0]$ 之內。這一點在 (ii) 中至關重要。

第 (i) 部分

根號外的括號是 $x+1$，所以我們嘗試把它與 arcosh 公式中的 $x+a$ 配對。寫成 $\sqrt{x+3} = \sqrt{2x+6}\,/\sqrt{2}$，於是除了 $a=1$ 之外，還需要 $a+b=6$。由此得 $b=5$，故 $b-a=4$。取這些數值時，假設 $x+a>0$、$x+b>0$ 及 $b>a$ 在 $x>-1$ 時全部成立，而且

$$\frac{\mathrm{d}}{\mathrm{d}x}\operatorname{arcosh}\left(\frac{x+5}{x+1}\right) = -\frac{2}{(x+1)\sqrt{2x+6}} = -\frac{\sqrt{2}}{(x+1)\sqrt{x+3}}.$$

兩邊除以 $-\sqrt{2}$，

$$\int \frac{1}{(x+1)\sqrt{x+3}}\,\mathrm{d}x = -\frac{1}{\sqrt{2}}\operatorname{arcosh}\left(\frac{x+5}{x+1}\right) + C.$$

第 (ii) 部分

現在根號外的括號是 $x+3$，所以我們把它與 arcsin 公式中的 $x+b$ 配對：$b=3$。根號內需要 $a+b+2x=2(x+1)$，所以 $a+b=2$，$a=-1$。於是 $b-a=4$，並且

$$\frac{\mathrm{d}}{\mathrm{d}x}\arcsin\left(\frac{x-1}{x+3}\right)=\frac{2}{(x+3)\sqrt{2x+2}}=\frac{\sqrt{2}}{(x+3)\sqrt{x+1}}.$$

想一想 當 $a=-1$ 時，假設 $x+a>0$ 要求 $x>1$。我們的答案在整個範圍 $x>-1$ 上仍然成立嗎？

成立。正如上文所見，arcsin 的公式只需要 $x+b>0$ 及 $a+b+2x>0$。在這裏，這兩個條件就是 $x+3>0$ 及 $2x+2>0$，對所有 $x>-1$ 都成立。在這個範圍內，$(x-1)/(x+3)$ 由 -1 向上趨近 1，所以 arcsin 有定義。因此，對 $x>-1$，

$$\int\frac{1}{(x+3)\sqrt{x+1}}\,\mathrm{d}x=\frac{1}{\sqrt{2}}\arcsin\left(\frac{x-1}{x+3}\right)+C.$$

另一做法：代換

兩個積分也可以用代換（substitution）$u=\sqrt{x+c}$ 解決，其中 $x+c$ 是根號下的式子，因為這時 $\mathrm{d}x=2u\,\mathrm{d}u$。在 (ii) 中，取 $u=\sqrt{x+1}$，積分變成 $\int\frac{2}{u^2+2}\,\mathrm{d}u=\sqrt{2}\arctan\left(\sqrt{(x+1)/2}\right)+C$。這與我們的答案只相差常數 $-\pi/(2\sqrt{2})$。在 (i) 中，取 $u=\sqrt{x+3}$，得到 $\int\frac{2}{u^2-2}\,\mathrm{d}u=\frac{1}{\sqrt{2}}\ln\left|\frac{\sqrt{x+3}-\sqrt{2}}{\sqrt{x+3}+\sqrt{2}}\right|+C$。這正是以對數（logarithm）寫出的 arcosh 答案。

驗算

- 把答案求導還原。例如，在 $x=0$ 處，$\frac{1}{\sqrt{2}}\arcsin\frac{x-1}{x+3}$ 的導數應為 $\frac{1}{3}$，而上面的公式給出 $\frac{\sqrt{2}}{\sqrt{2}\cdot3\cdot1}=\frac{1}{3}$。
- 數值上，用辛普森法則（Simpson's rule）可得 $\int_0^1\frac{\mathrm{d}x}{(x+1)\sqrt{x+3}}\approx0.374543$ 及 $\int_0^1\frac{\mathrm{d}x}{(x+3)\sqrt{x+1}}\approx0.240301$。兩個數值與原函數（antiderivative）算出的結果吻合，誤差小於 10^{-12}。在下限低至 $x=-0.9$ 的區間上，結果同樣吻合，而這些區間已伸入 $-1<x<1$ 的範圍，即 (ii) 中 $x+a<0$ 之處。
- 正負號：(i) 的被積函數是正的，所以它的積分應該遞增。我們的答案確是如此，因為 $\frac{x+5}{x+1}$ 隨 x 增加而遞減，所以它的 arcosh 的負值遞增。

常見錯誤

- 把括號配錯公式：在 (i) 中，根號外的 $x+1$ 是常數較小的那個，所以它必定是 $x+a$，即要用 arcosh，而不是 arcsin。
- 忘記了把 $\sqrt{x+3}$ 寫成 $\sqrt{2x+6}/\sqrt{2}$ 時所產生的因子 $\sqrt{2}$。
- 漏掉 arcosh 導數中的負號，這個負號來自 $\mathrm{d}z/\mathrm{d}x < 0$。

STEP 之外

(ii) 的兩個答案，一個用 arcsin，一個用 arctan，兩者相差一個常數。這是必然的，因為在一個區間上導數相同的兩個函數只相差一個常數，這是中值定理（mean value theorem）的推論。若集合由兩段分開的部分組成，各段上的常數便可以不同。(i) 中出現對數也並非偶然：反雙曲函數（inverse hyperbolic functions）其實是喬裝了的對數，這正是 $1/\sqrt{\text{二次式}}$ 的積分不是得出反三角函數（inverse trigonometric functions）便是得出對數的原因。Abbott 的 *Understanding Analysis* 嚴謹地證明了中值定理，而 Siklos 的 *Advanced Problems in Mathematics* 則收錄了更多這類積分。

詳細解答

STEP III 2003，第 2 題

純數學 • 二項式級數

> **STEP III 2003，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/03-s3-q2

關鍵思路

中心二項式係數（central binomial coefficient）$\binom{2r}{r}$ 就藏在負半整數冪的二項級數（binomial series）之中。只要用開首的恆等式把 $(2r)!$ 拆成奇數因子和偶數因子，便會發現 $(1-p)^{-1/2}$ 中 p^r 的係數恰好是 $\binom{2r}{r}/4^r$。於是每個和其實都是喬裝了的二項級數，餘下的工作只是選取 p（以及在 (ii) 中選對冪次）。

解答

恆等式

把 $(2r)!$ 的因子分成奇數和偶數兩組：

$$(2r)! = \big[1 \times 3 \times \cdots \times (2r-1)\big] \times \big[2 \times 4 \times \cdots \times 2r\big].$$

r 個偶數因子各自是 $1, 2, \ldots, r$ 其中一個數的 2 倍，所以第二個括號等於 $2^r\, r!$。因此

$$\binom{2r}{r} = \frac{(2r)!}{r!\,r!} = \frac{1 \times 3 \times \cdots \times (2r-1) \times 2^r\, r!}{r!\,r!} = \frac{1 \times 3 \times \cdots \times (2r-1)}{r!} \times 2^r,$$

這正是所要證明的結果。

第 (i) 部分

當 $|p|<1$ 時，由二項級數可得

$$\begin{aligned}(1-p)^{-\frac{1}{2}} &= 1+\left(-\tfrac{1}{2}\right)(-p)+\frac{\left(-\frac{1}{2}\right)\left(-\frac{3}{2}\right)}{2!}(-p)^2\\ &\quad+\frac{\left(-\frac{1}{2}\right)\left(-\frac{3}{2}\right)\left(-\frac{5}{2}\right)}{3!}(-p)^3+\cdots\\ &= 1+\tfrac{1}{2}p+\tfrac{3}{8}p^2+\tfrac{5}{16}p^3+\cdots.\end{aligned}$$

通項也依循同樣的規律。在 p^r 的係數中，因子 $-\frac{1}{2},-\frac{3}{2},\ldots$ 帶來 r 個負號，$(-p)^r$ 又帶來 r 個，兩者互相抵消，剩下

$$\frac{\frac{1}{2}\times\frac{3}{2}\times\cdots\times\frac{2r-1}{2}}{r!}\,p^r=\frac{1\times3\times\cdots\times(2r-1)}{2^r\,r!}\,p^r.$$

想一想 這與 $\binom{2r}{r}$ 有甚麼關係？

由恆等式可知 $\dfrac{1\times3\times\cdots\times(2r-1)}{r!}=\dfrac{\binom{2r}{r}}{2^r}$。所以 p^r 的係數是 $\binom{2r}{r}/4^r$；當 $r=0$ 時這也成立，因為兩邊都等於 1。因此

$$(1-p)^{-\frac{1}{2}}=\sum_{r=0}^{\infty}\binom{2r}{r}\left(\frac{p}{4}\right)^r\qquad(|p|<1).$$

我們要求的和中 $(p/4)^r=1/8^r$，所以取 $p=\frac{1}{2}$，這個值在收斂區間（interval of convergence）之內。於是

$$\sum_{r=0}^{\infty}\frac{\binom{2r}{r}}{8^r}=\left(1-\tfrac{1}{2}\right)^{-\frac{1}{2}}=\left(\tfrac{1}{2}\right)^{-\frac{1}{2}}=\sqrt{2}.$$

第 (ii) 部分

多出來的因子 $2r+1$ 提示我們把冪次再推前一步，考慮 $(1-p)^{-3/2}$。當 $|p|<1$ 時，其二項級數中 p^r 的係數是（正負號如前一樣互相抵消）

$$\begin{aligned}\frac{\frac{3}{2}\times\frac{5}{2}\times\cdots\times\frac{2r+1}{2}}{r!} &= \frac{3\times5\times\cdots\times(2r+1)}{2^r\,r!}\\ &= (2r+1)\,\frac{1\times3\times\cdots\times(2r-1)}{2^r\,r!}=\frac{(2r+1)\binom{2r}{r}}{4^r},\end{aligned}$$

這是因為 $3\times5\times\cdots\times(2r+1)=(2r+1)\times\left[1\times3\times\cdots\times(2r-1)\right]$，然後再像 (i) 那樣運用恆等式。（當 $r=0$ 時，首尾兩端都等於 1。）所以

$$(1-p)^{-\frac{3}{2}}=\sum_{r=0}^{\infty}(2r+1)\binom{2r}{r}\left(\frac{p}{4}\right)^r\qquad(|p|<1).$$

現在需要 $p/4 = \frac{1}{5}$，即 $p = \frac{4}{5}$，同樣滿足 $|p| < 1$。因此

$$\sum_{r=0}^{\infty} \frac{(2r+1)\binom{2r}{r}}{5^r} = \left(1 - \tfrac{4}{5}\right)^{-\frac{3}{2}} = \left(\tfrac{1}{5}\right)^{-\frac{3}{2}} = 5^{\frac{3}{2}} = (\sqrt{5})^3,$$

這正是所要證明的結果。

另一做法：當 $|x| < \frac{1}{4}$ 時，記 $f(x) = \sum \binom{2r}{r} x^r = (1-4x)^{-1/2}$。在收斂區間內對冪級數（power series）逐項求導，得 $\sum r\binom{2r}{r}x^r = xf'(x) = 2x(1-4x)^{-3/2}$，所以

$$\sum (2r+1)\binom{2r}{r}x^r = 4x(1-4x)^{-\frac{3}{2}} + (1-4x)^{-\frac{1}{2}} = (1-4x)^{-\frac{3}{2}},$$

代入 $x = \frac{1}{5}$ 便得到相同的答案。

驗算

- (i) 的首幾項是 $1 + \frac{2}{8} + \frac{6}{64} + \frac{20}{512} \approx 1.371$，已經接近 $\sqrt{2} \approx 1.414$；把 200 項加起來得 1.41421356…。
- (ii) 的級數收斂得較慢（相鄰項之比趨向 $\frac{4}{5}$），但 400 項之和是 11.18033988…，與 $5\sqrt{5} \approx 11.1803$ 吻合。
- 當 $r = 0, 1, 2, 3$ 時，係數 $\binom{2r}{r}/4^r$ 分別是 $1, \frac{1}{2}, \frac{3}{8}, \frac{5}{16}$，與直接求得的四項一致。

常見錯誤

- 弄錯正負號：係數中的 r 個負因子與 $(-p)^r$ 所帶來的 $(-1)^r$ 必定互相抵消，所以每個係數都是正數。
- 忘記檢查所選的 p 是否滿足 $|p| < 1$；若不滿足，級數未必收斂至該函數。
- 在 (ii) 中換一個 p 再用 $(1-p)^{-1/2}$：單憑這個級數，不可能得出因子 $2r+1$。

STEP 之外

函數 $\sum \binom{2r}{r}x^r = (1-4x)^{-1/2}$ 是中心二項式係數的生成函數（generating function），是組合數學（combinatorics）中的標準工具；例如把它平方便得 $\sum_r 4^r x^r$，從而證明 $\sum_k \binom{2k}{k}\binom{2r-2k}{r-k} = 4^r$。逐項求導為何成立，以及冪級數確切在哪裏收斂，都是大學一年級分析學（analysis）的課題。關於冪級數，可參閱 Abbott 的 *Understanding Analysis*；想做更多與級數有關的 STEP 練習，可參閱 Siklos 的 *Advanced Problems in Mathematics*。

詳細解答

STEP III 2003，第 3 題

純數學 • 曲線描繪

> **STEP III 2003，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/03-s3-q3

關鍵思路

設 $u = (1+x)^m$ 及 $v = (1-x)^m$。於是 $\mathrm{f} = (u-v)/(u+v)$，而商法則（quotient rule）的結果會化簡得非常漂亮，因為 $u'v - uv'$ 有公因式 $(1-x^2)^{m-1}$。這個導數（derivative）永不為負，單憑這一點，我們幾乎已掌握 $y = \mathrm{f}(x)$ 的整個形狀。至於倒數（reciprocal）曲線，則藏着一種對稱（symmetry）：把 x 換成 $1/x$，f 就變成 $1/\mathrm{f}$。

解答

分母永不為零

假設對某個實數 x 有 $(1+x)^m + (1-x)^m = 0$。

若 m 是偶數，兩項皆為非負，故兩項必須同時為零。這就需要同時有 $x = -1$ 和 $x = 1$，並不可能。

若 m 是奇數，則 $(1+x)^m = -(1-x)^m = (x-1)^m$。當 m 是奇數時，映射 $t \mapsto t^m$ 在 $\mathbb{R}$ 上嚴格遞增（strictly increasing），因此是一對一（one-to-one）的，於是 $1+x = x-1$，即 $1 = -1$。這是矛盾（contradiction）。

所以對每個實數 x 都有 $(1+x)^m + (1-x)^m \neq 0$，而 f 在整個 $\mathbb{R}$ 上有定義並且可微（differentiable）。

導數

設 $u=(1+x)^m$ 及 $v=(1-x)^m$，則 $u'=m(1+x)^{m-1}$ 及 $v'=-m(1-x)^{m-1}$。根據商法則，

$$\mathrm{f}'(x)=\frac{(u'-v')(u+v)-(u-v)(u'+v')}{(u+v)^2}=\frac{2(u'v-uv')}{(u+v)^2}.$$

分子可以化簡：

$$\begin{aligned}u'v-uv' &= m(1+x)^{m-1}(1-x)^m+m(1+x)^m(1-x)^{m-1}\\ &= m(1+x)^{m-1}(1-x)^{m-1}\big[(1-x)+(1+x)\big]\\ &= 2m(1-x^2)^{m-1}.\end{aligned}$$

因此

$$\mathrm{f}'(x)=\frac{4m\,(1-x^2)^{m-1}}{\big[(1+x)^m+(1-x)^m\big]^2}.$$

$m=5$ 的情況：曲線 $y=\mathrm{f}(x)$

用二項式定理（binomial theorem）展開，分子中的偶次冪互相抵消，分母中的奇次冪互相抵消：

$$\mathrm{f}(x)=\frac{2(5x+10x^3+x^5)}{2(1+10x^2+5x^4)}=\frac{x(x^4+10x^2+5)}{5x^4+10x^2+1},\qquad \mathrm{f}'(x)=\frac{20(1-x^2)^4}{\big[(1+x)^5+(1-x)^5\big]^2}.$$

現在逐一讀出曲線的特徵。

- 對稱性。$\mathrm{f}(-x)=-\mathrm{f}(x)$，所以曲線繞原點有旋轉對稱（rotational symmetry）。曲線通過 $(0,0)$，該處的斜率（gradient）為 $\mathrm{f}'(0)=20/2^2=5$。
- 單調性（*monotonic*）。對所有 x 都有 $\mathrm{f}'(x)\geq 0$，等號只在 $x=\pm1$ 時成立。所以 f 處處遞增。在 $x=\pm1$ 處曲線有駐點（stationary point），而由於 $(1-x^2)^4$ 不變號，這些駐點是拐點（point of inflection），而非轉向點（turning point）。它們位於 $(1,1)$ 和 $(-1,-1)$：$(1-x)^5$ 或 $(1+x)^5$ 一項變為零，餘下 $\mathrm{f}=\pm1$。
- $|x|$ 很大時。$\mathrm{f}(x)\approx x/5$。更準確地說，

  $$\mathrm{f}(x)-\frac{x}{5}=\frac{x(40x^2+24)}{5(5x^4+10x^2+1)},$$

  當 $|x|\to\infty$ 時，上式趨向 0。所以 $y=x/5$ 是斜漸近線（oblique asymptote）。曲線在 $x>0$ 時位於漸近線之上，在 $x<0$ 時位於漸近線之下。

想一想 $\mathrm{f}(x)$ 的絕對值在哪裏大於 1，又在哪裏小於 1？

由於 f 遞增且 $f(\pm1) = \pm1$，當 $|x| < 1$ 時有 $|f(x)| < 1$，當 $|x| > 1$ 時有 $|f(x)| > 1$。這一點對倒數曲線很重要。（當 $|x| < 1$ 時，代入 $x = \tanh t$ 可得 $(1+x)/(1-x) = e^{2t}$，再把 f 的分子和分母同除以 $(1-x)^5$，便得 $f(\tanh t) = \tanh 5t$。這是理解曲線在 -1 與 1 之間呈 S 形的另一個角度。）

曲線 $y = 1/f(x)$

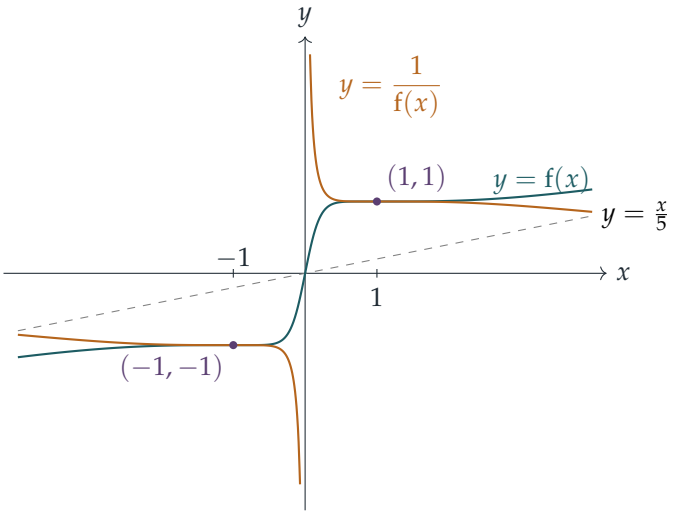
把 x 換成 $1/x$，再把分子和分母同乘以 x^5，可得

$$f(1/x) = \frac{(x+1)^5 - (x-1)^5}{(x+1)^5 + (x-1)^5} = \frac{(1+x)^5 + (1-x)^5}{(1+x)^5 - (1-x)^5} = \frac{1}{f(x)},$$

當中用了 $(x-1)^5 = -(1-x)^5$。所以 $y = 1/f(x)$ 就是曲線 $y = f(1/x)$，它的特徵可由 f 的特徵推出：

- 它是奇函數（odd function），在 $x = 0$ 處沒有定義，而這是 f 唯一的零點。在 0 附近，$1/f(x) \approx 1/(5x)$，所以 y 軸是鉛垂漸近線（vertical asymptote），當 $x \to 0^+$ 時 $y \to +\infty$，當 $x \to 0^-$ 時 $y \to -\infty$。
- 它的導數是 $-f'(x)/f(x)^2 \leq 0$，所以它在 $x > 0$ 上遞減，在 $x < 0$ 上亦遞減。它在 $(1,1)$ 和 $(-1,-1)$ 有屬於駐點的拐點。
- 當 $|x| \to \infty$ 時，$1/f(x) \approx 5/x \to 0$，所以 x 軸是漸近線。曲線在 $x > 0$ 時位於其上，在 $x < 0$ 時位於其下。

兩條曲線只在 $f(x)^2 = 1$ 之處相遇，即 $(1,1)$ 和 $(-1,-1)$。兩條曲線在該處都是水平的，所以它們是相切而非相交穿過。



驗算

- 取 $m = 1$：公式給出 $f'(x) = 4/2^2 = 1$，而事實上 $f(x) = 2x/2 = x$。
- 以數值計算，$f(1000)/1000 \approx 0.2000016$，與漸近線 $y = x/5$ 吻合。對 $m = 1, \ldots, 8$，在隨機選取的點上，以中心差分（central difference）求得的導數與 f' 的公式一致。
- $f(\tanh t) = \tanh 5t$ 和 $f(1/x) = 1/f(x)$ 在隨機選取的點上均準確至 10^{-9}。

常見錯誤

- 把 $x = \pm 1$ 稱為極大點或極小點。導數在該處觸及零但不變號，所以這些是屬於駐點的拐點。
- 遺漏斜漸近線：f 是五次多項式除以四次多項式，所以它的增長有如 $x/5$，而不會趨於平緩。
- 把 $y = 1/f(x)$ 畫成在 $(\pm 1, \pm 1)$ 穿過 $y = f(x)$。兩條曲線在該處都是平的，所以它們只是相切。

STEP 之外

代換 $x = \tanh t$ 顯示，f 正是以 $\tanh t$ 表示 $\tanh 5t$ 的公式，也就是 $\tan 5\theta$ 的倍角公式（multiple-angle formula）的雙曲（hyperbolic）版本。由此觀之，f 是一個有理映射（rational map），它共軛（conjugate）於乘以 5 的運算。這類映射，以及它們在迭代（iteration）下的表現，正是複動力系統（complex dynamics）的起點之一。Needham 的 *Visual Complex Analysis*（《複分析：可視化方法》）是一本優美的書，可從中看到這類有理函數如何作用於整個複平面。

詳細解答

STEP III 2003，第 4 題

純數學 • 參數曲線的切線

STEP III 2003，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/03-s3-q4

關鍵思路

直線與這條曲線相交之處，正是某個以參數為變量的三次式（cubic）等於零之處。切線（tangent）會在 t 處產生一個二重根（double root），所以由根之和便可立即求出第三個根 T。T 的公式看起來很像 tan 的倍角公式（double-angle formula），這提示我們設 $t = \tan\theta$：映射 $t \mapsto T$ 於是變成角度上的一個簡單線性映射，要把它迭代（iterate）便很容易。

解答

T 的公式

由於 $y = t + t^3$ 對 t 嚴格遞增（strictly increasing），不同的參數（parameter）給出不同的點，所以曲線上的點與參數的值一一對應（one-to-one）。

求導可得 $\dfrac{\mathrm{d}x}{\mathrm{d}t} = 2t$ 和 $\dfrac{\mathrm{d}y}{\mathrm{d}t} = 1 + 3t^2$，所以當 $t \neq 0$ 時，在 t 處的切線的斜率（gradient）為

$$m = \frac{1 + 3t^2}{2t}.$$

（因為 $t \neq 0$，切線不是垂直的。）切線的方程是 $y - (t + t^3) = m(x - t^2)$。參數為 s 的點在這條切線上，當且僅當

$$s + s^3 - (t + t^3) = m(s^2 - t^2), \qquad \text{即} \qquad f(s) \equiv s^3 - ms^2 + s - \left(t + t^3 - mt^2\right) = 0.$$

顯然 $f(t) = 0$，而根據 m 的選取，$f'(t) = 3t^2 - 2mt + 1 = 0$。所以 $s = t$ 是這個三次式的（至少）二重根，而各根為 t, t, T，其中 T 是某個實數。$s^3 - ms^2 + \ldots$ 的根之和是

m，所以

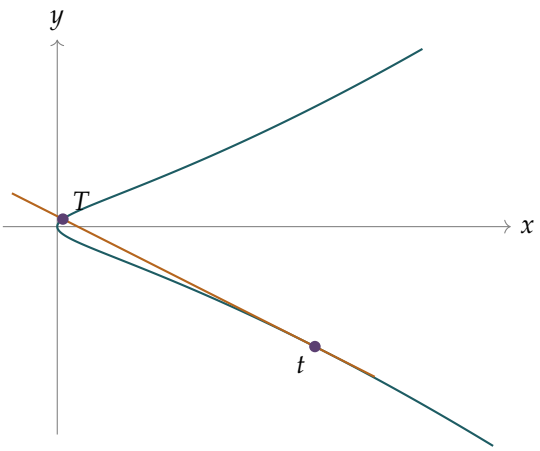
$$2t + T = \frac{1+3t^2}{2t}$$
$$T = \frac{1+3t^2-4t^2}{2t} = \frac{1-t^2}{2t}.$$

這個根 T 是該直線與曲線唯一可能相交的另一處。

只有當 $T \neq t$ 時，切線才會與曲線再次相交。現在

$$T = t \iff 1-t^2 = 2t^2 \iff 3t^2 = 1,$$

所以我們需要 $3t^2 \neq 1$。（當 $3t^2 = 1$ 時，t 是三重根，切線只與曲線相接一次：這些正是拐點（points of inflection）。）這就證明了兩個結果。



以 $t_0 = \tan\frac{7}{18}\pi$ 開始的數列

設 $t = \tan\theta$。由倍角公式，

$$T = \frac{1-\tan^2\theta}{2\tan\theta} = \frac{1}{\tan 2\theta} = \cot 2\theta = \tan\left(\frac{\pi}{2} - 2\theta\right).$$

所以若 $t_i = \tan\theta_i$，我們可以取 $\theta_{i+1} = \frac{\pi}{2} - 2\theta_i$。由於 tan 的週期（period）為 π，而且在任何長度為 π 的區間上（除去 $\frac{\pi}{2}$ 的奇數倍）都是一對一的，兩個參數相等，當且僅當它們的角相差 π 的倍數。

想一想 在迭代之前，我們必須檢查每個 t_i 的甚麼性質，才能確保下一個點存在？

這個構作在每一步都需要 $t_i \neq 0$ 和 $3t_i^2 \neq 1$，即 θ_i 不是 $\frac{\pi}{2}$ 的倍數，也不是 $\pm\frac{\pi}{6}$ 加上 π 的倍數。以 $\frac{\pi}{18}$ 的倍數計，我們必須避開模（modulo）18 下的 $0, \pm 3, \pm 9$。

由 $\theta_0 = \frac{7\pi}{18}$ 開始：

$$\begin{aligned}\theta_1 &= \frac{9\pi}{18} - \frac{14\pi}{18} = -\frac{5\pi}{18},\\ \theta_2 &= \frac{9\pi}{18} + \frac{10\pi}{18} = \frac{19\pi}{18} \equiv \frac{\pi}{18} \pmod{\pi},\\ \theta_3 &= \frac{9\pi}{18} - \frac{2\pi}{18} = \frac{7\pi}{18}.\end{aligned}$$

$7, -5, 1$ 都不在禁用之列，所以每一步都成立。因此 $t_3 = \tan\frac{7\pi}{18} = t_0$，即 $P_3 = P_0$。但 $\theta_1 - \theta_0 = -\frac{12\pi}{18}$ 不是 π 的倍數，所以 $t_1 = \tan(-\frac{5\pi}{18}) < 0 < t_0$，而 $P_1 \neq P_0$。這三點的參數分別是 $\tan 70^\circ$、$-\tan 50^\circ$ 和 $\tan 10^\circ$。

第二個起始值

設 $g(\theta) = \frac{\pi}{2} - 2\theta$。迭代可得

$$g^2(\theta) = -\frac{\pi}{2} + 4\theta, \qquad g^3(\theta) = \frac{3\pi}{2} - 8\theta.$$

所以 $P_3 = P_0$ 當且僅當 $\frac{3\pi}{2} - 8\theta \equiv \theta \pmod{\pi}$，即 $9\theta \equiv \frac{3\pi}{2} \pmod{\pi}$，由此得

$$\theta = \frac{\pi}{6} + \frac{k\pi}{9} = \frac{(3+2k)\pi}{18}, \qquad k \in \mathbb{Z}.$$

同理，$P_1 = P_0$ 當且僅當 $3\theta \equiv \frac{\pi}{2} \pmod{\pi}$，即 $\theta = \frac{(3+6j)\pi}{18}$。這些正是 $3 + 2k \equiv 3, 9, 15 \pmod{18}$ 的情況：它們給出 $t = \pm\frac{1}{\sqrt{3}}$，或者根本沒有有限的 t，本來就已被排除。

在模 π 之下，餘下的是 $\theta \in \{\frac{\pi}{18}, \frac{5\pi}{18}, \frac{7\pi}{18}, \frac{11\pi}{18}, \frac{13\pi}{18}, \frac{17\pi}{18}\}$。要 $t_0 > 0$，我們需要 $0 < \theta < \frac{\pi}{2}$，由此得 $\tan\frac{\pi}{18}$、$\tan\frac{5\pi}{18}$ 和題目給出的 $\tan\frac{7\pi}{18}$。$\tan\frac{\pi}{18}$ 這個值只是上述循環（cycle）中的 P_2。給出一個真正全新循環的值是

$$t_0 = \tan\frac{5\pi}{18} \ (= \tan 50^\circ),$$

其中 $\theta_0 = \frac{5\pi}{18}$，$\theta_1 = -\frac{\pi}{18}$，$\theta_2 = \frac{11\pi}{18}$，$\theta_3 = \frac{9\pi}{18} - \frac{22\pi}{18} = -\frac{13\pi}{18} \equiv \frac{5\pi}{18}$。$5, -1, 11$ 都不在禁用之列，而且 $t_1 = -\tan 10^\circ \neq t_0$。$\tan\frac{5\pi}{18}$ 或 $\tan\frac{\pi}{18}$ 都可以作為本題的答案。

> ? **想一想** 為甚麼可以「模 π」地處理角度？

因為我們唯一用到的就是 $t = \tan\theta$，而 $\tan(\theta + \pi) = \tan\theta$。把 θ_i 加上 π，$g(\theta_i)$ 便改變 2π，這並不改變 t_{i+1}。所以整個參數數列只取決於 θ_0 模 π 的值。

驗算

▶ 取 $t = -1$ 可得 $T = 0$：在 $(1, -2)$ 處的切線斜率為 -2，而它確實通過原點。

- 以數值計算，$t_0 = \tan 70^\circ \approx 2.747477$ 給出 $t_1 \approx -1.191754$、$t_2 \approx 0.176327$ 和 $t_3 \approx 2.747477$。同樣地，$\tan 50^\circ \approx 1.191754 \to -0.176327 \to -2.747477 \to 1.191754$。
- 隨機選取 t，以數值方法解三次方程求另一個交點，結果與 $T = (1 - t^2)/(2t)$ 吻合。

常見錯誤

- 先求弦（chord）的斜率再令兩者相等，卻忘記約去因子 $T - t$。利用關於根之和的韋達公式（Vieta's formula），便可完全避開這個問題。
- 直接用小數對 t 迭代這個映射，結果看不出精確的循環。正是代換 $t = \tan\theta$ 才令 $P_3 = P_0$ 精確成立。
- 把 $\tan\frac{\pi}{6}$ 當作第二個值：這時 $T = t$，所以根本沒有下一個點。

STEP 之外

弦切構作法（chord-and-tangent construction）是橢圓曲線（elliptic curves）理論的核心：在該理論中，同一條「第三個交點」規則把三次曲線上的點變成一個羣（group）。這裏的曲線 $y^2 = x(1+x)^2$ 是一條奇異三次曲線（singular cubic）（它在 $(-1, 0)$ 有一個孤立點），而代換 $t = \tan\theta$ 表明，切線映射其實是喬裝了的角度加倍。這就是為甚麼這些循環來自把圓分成九等份。Silverman 與 Tate 合著的 *Rational Points on Elliptic Curves*（《橢圓曲線上的有理點》）以本科程度把這個課題闡述得十分精彩。

詳細解答

STEP III 2003，第 5 題

純數學 • 三次方程的根

> **STEP III 2003，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/03-s3-q5

關鍵思路

從圖像入手數算根（root）的個數。首項係數（leading coefficient）為正的二次函數，其圖像與 x 軸相交兩次，當且僅當其極小值位於該軸下方。首項係數為正的三次函數，其圖像與該軸相交三次，當且僅當其極大值在軸的上方而極小值在軸的下方。要令所有根皆為正數，就要看轉向點（turning point）位於 y 軸的哪一側，以及曲線與 y 軸相交處的正負號。最後一條方程其實是經過平移的三次曲線的一個喬裝例子：配立方（complete the cube）便可把它化成該形式。

解答

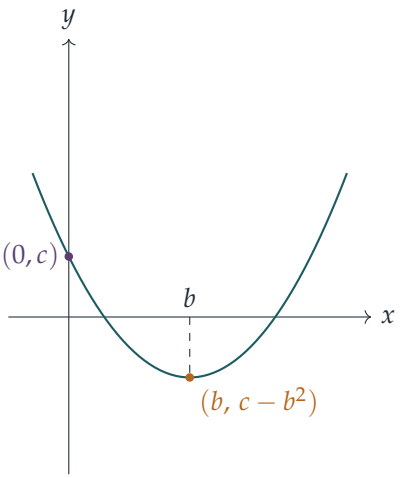
二次函數

配方（completing the square）得

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

所以轉向點（極小點）位於 $(b,\ c - b^2)$。

當有兩個相異實根時，曲線是一條 $\cup$ 形的拋物線（parabola），對稱於 $x = b$，其頂點（vertex）位於 x 軸下方。它與 y 軸相交於 $(0, c)$。



兩個相異實根。拋物線的首項係數為正，所以當 $x \to \pm\infty$ 時 $y \to +\infty$，而它的最小值是 $c - b^2$。若 $c - b^2 < 0$，曲線會降至軸的下方，然後在 $x = b$ 的兩側各自回升，因此與軸相交兩次。若 $c - b^2 = 0$，曲線僅與軸相切（重根（repeated root））；若 $c - b^2 > 0$，則曲線與軸永不相遇。所以條件是

$$b^2 > c.$$

兩個相異正根。我們斷言條件是

$$b^2 > c, \qquad b > 0, \qquad c > 0.$$

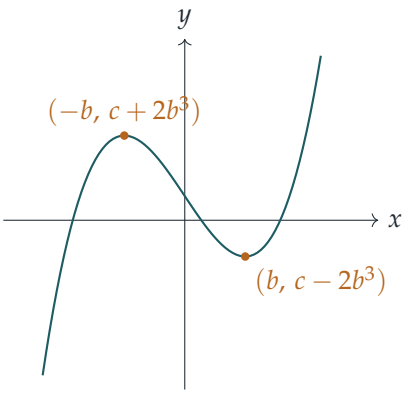
必要性（necessary）：兩根 α, β 滿足 $\alpha + \beta = 2b$ 及 $\alpha\beta = c$，若兩根皆為正數，則這兩個量亦皆為正數。**充分性（sufficient）：**若 $b^2 > c$，則有兩個根，分別位於 $x = b$ 的兩側。由於 $b > 0$，右邊的根是正數。左邊的根位於曲線遞減的區間內；曲線在 $x = 0$ 處為正（高度為 $c > 0$），在 $x = b > 0$ 處為負，所以這個根嚴格介乎 0 與 b 之間，因而也是正數。

三次函數

對於 $y = x^3 - 3b^2x + c$，我們有 $\dfrac{\mathrm{d}y}{\mathrm{d}x} = 3x^2 - 3b^2 = 3(x - b)(x + b)$，它在 $x = \pm b$ 時等於零。代入得

$$x = -b:\ y = -b^3 + 3b^3 + c = c + 2b^3, \qquad x = b:\ y = b^3 - 3b^3 + c = c - 2b^3.$$

由於 $b > 0$，轉向點分別是位於 $(-b,\ c + 2b^3)$ 的極大點和位於 $(b,\ c - 2b^3)$ 的極小點：導數（derivative）在 $x < -b$ 時為正，在兩者之間為負，在 $x > b$ 時為正。



曲線由 $-\infty$ 上升至極大點，下降至極小點，然後上升至 $+\infty$，而且在這三段上各自嚴格單調（strictly monotonic）。所以曲線在每一段上各有一個根，從而共有三個相異的根，當且僅當極大值在軸的上方而極小值在軸的下方：

$$c+2b^3>0 \text{ 及 } c-2b^3<0, \qquad \text{即} \qquad -2b^3<c<2b^3.$$

若其中一個轉向點的函數值為零，便會得到重根；若兩者同為嚴格的正數或同為嚴格的負數，則只有一個實根。所以 $|c|<2b^3$ 是充分必要條件。

平移後的三次曲線的三個相異正根

曲線 $y=(x-a)^3-3b^2(x-a)+c$ 是把上一條曲線向右平移（translate）a 而得（我們仍假設 $b>0$）。所以它有三個相異實根，當且僅當 $|c|<2b^3$；這時它的極大點在 $x=a-b$，極小點在 $x=a+b$，而三個根 $x_1<x_2<x_3$ 滿足

$$x_1<a-b<x_2<a+b<x_3.$$

三個根全為正數，當且僅當最小的根 x_1 為正數。記

$$y(0)=(-a)^3-3b^2(-a)+c=c-a^3+3ab^2$$

為曲線與 y 軸相交處的高度。我們斷言條件是

$$|c|<2b^3, \qquad a>b, \qquad c<a^3-3ab^2.$$

? **想一想** 為甚麼單憑「$y(0)<0$」並不足以保證三個根全都在原點的右方？

因為 y 在 x_2 與 x_3 之間同樣是負數。在 $x=0$ 處取負值，有可能表示原點位於中間的根與最大的根之間。我們還須知道 0 位於曲線第一段上升的部分，即極大點的左方。這正是 $a>b$ 的作用。

必要性。假設有三個相異正根。那麼由上一部分可知 $|c| < 2b^3$。此外 $0 < x_1 < a - b$，所以 $a > b$。最後，y 在 $(-\infty, a-b]$ 上嚴格遞增，而 $y(x_1) = 0$，所以 $y(0) < 0$，因為 $0 < x_1$。由此得 $c < a^3 - 3ab^2$。

充分性。假設三個條件全部成立。這時有三個相異實根，而且 $0 < a - b$，所以 0 位於 $(-\infty, a-b]$ 內，而 y 在此區間上嚴格遞增。由於 $y(0) < 0 = y(x_1)$，我們得到 $0 < x_1$，因此三個根全為正數。

所以該方程有三個相異正根，當且僅當

$$a > b \quad \text{及} \quad -2b^3 < c < \min\bigl(2b^3,\ a^3 - 3ab^2\bigr).$$

（由於當 $a > b > 0$ 時 $a^3 - 3ab^2 + 2b^3 = (a-b)^2(a+2b) > 0$，上述對 c 的上下界並無矛盾。）

最後的方程

先除以 2，然後配立方。$(x-a)^3$ 中的 x^2 項是 $-3ax^2$，所以我們需要 $3a = \frac{9}{2}$，即 $a = \frac{3}{2}$。設 $x = t + \frac{3}{2}$，

$$\begin{aligned} x^3 - \tfrac{9}{2}x^2 + \tfrac{7}{2}x - \tfrac{1}{2} &= t^3 + \bigl(\tfrac{27}{4} - \tfrac{27}{2} + \tfrac{7}{2}\bigr)t + \bigl(\tfrac{27}{8} - \tfrac{81}{8} + \tfrac{21}{4} - \tfrac{1}{2}\bigr) \\ &= t^3 - \tfrac{13}{4}t - 2, \end{aligned}$$

其中 t^2 項因 a 的選取而互相抵消。所以

$$2x^3 - 9x^2 + 7x - 1 = 0 \iff \bigl(x - \tfrac{3}{2}\bigr)^3 - 3b^2\bigl(x - \tfrac{3}{2}\bigr) + c = 0, \qquad a = \tfrac{3}{2},\ b^2 = \tfrac{13}{12},\ c = -2,$$

其中 $b = \sqrt{13/12} > 0$。現在逐一檢查三個條件。

- $|c| < 2b^3$：這即是說 $2 < 2b^3$，亦即 $b^2 > 1$，而這是成立的，因為 $b^2 = \frac{13}{12} > 1$。
- $a > b$：兩者皆為正數，而且 $a^2 = \frac{9}{4} = \frac{27}{12} > \frac{13}{12} = b^2$。
- $c < a^3 - 3ab^2$：這裏 $a^3 - 3ab^2 = a(a^2 - 3b^2) = \frac{3}{2}\bigl(\frac{9}{4} - \frac{13}{4}\bigr) = -\frac{3}{2}$，而 $-2 < -\frac{3}{2}$。

三個條件全部成立，所以 $2x^3 - 9x^2 + 7x - 1 = 0$ 有三個相異正根，正如所求。

驗算

- 正負號的變化可直接確認最後的答案。設 $p(x) = 2x^3 - 9x^2 + 7x - 1$，我們有 $p(0) = -1$、$p(\frac{1}{2}) = \frac{1}{2}$、$p(1) = -1$、$p(3) = -7$ 及 $p(4) = 11$。所以在 $(0, \frac{1}{2})$、$(\frac{1}{2}, 1)$ 及 $(3,4)$ 內各有一個根。它們的數值分別是 0.185、0.760 及 3.555。
- 數值 $y(0) = c - a^3 + 3ab^2 = -2 + \frac{3}{2} = -\frac{1}{2}$ 與首一（monic）三次多項式 $x^3 - \frac{9}{2}x^2 + \frac{7}{2}x - \frac{1}{2}$ 的常數項吻合，這是理所當然的。

- 以數千組隨機選取的 (a, b, c) 作測試，結果每一次都與上述條件相符。

常見錯誤

- 只用 $y(0) < 0$（或只用 $a > b$）來判斷根是否為正。兩者缺一不可：當原點位於較大的兩個根之間時，$y(0) < 0$ 同樣成立。
- 忘記了 $|c| \leq 2b^3$ 中等號成立時會得到重根，因此不算「三個相異的根」。
- 把平移的方向弄反：以 $x - a$ 表示的曲線，是把原來的曲線向*右*移動 a。

STEP 之外

任何三次多項式都可以像這裏一樣，藉着平移化成「缺項」(depressed) 形式 $t^3 + pt + q$。這時它有三個相異實根，當且僅當 $4p^3 + 27q^2 < 0$；取 $p = -3b^2$、$q = c$，這就是上文求得的條件 $c^2 < 4b^6$。$-(4p^3 + 27q^2)$ 這個量稱為*判別式*（discriminant），它是三次方程如何得以求解（卡爾達諾（Cardano）公式）這段故事的開端，最終更引出伽羅瓦理論（Galois theory）。Allenby 的 *Rings, Fields and Groups*（環、域與羣）是邁向該理論的理想下一步。

詳細解答

STEP III 2003，第 6 題

純數學 • 三角函數裂項相消求和

> **STEP III 2003，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/03-s3-q6

關鍵思路

各項角度以相等間隔遞增的餘弦之和，照原樣處理並不方便；但只要把每一項都乘以 $2\sin\frac{1}{2}\theta$，每個餘弦就會變成兩個正弦之*差*，整個和式隨即可以裂項相消（telescoping）。餘下的是兩個正弦之差，再用和差化積公式（factor formula）便可化為乘積。令一個乘積等於零，是很容易處理的事。唯一要留神的是：我們乘上了 $\sin\frac{1}{2}\theta$，而它本身也可能等於零。

解答

恆等式

利用 $\sin(A+B)-\sin(A-B)=2\cos A\sin B$，取 $A=r\theta$ 及 $B=\frac{1}{2}\theta$：

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

這就是所需的結果。

級數求和

記

$$S(\theta)=\sum_{r=a}^{b-1}\cos r\theta.$$

乘以 $2\sin\frac{1}{2}\theta$，並對由 a 至 $b-1$ 的每個 r 使用上述恆等式（identity），

$$\begin{aligned} 2\sin\tfrac{1}{2}\theta\, S(\theta) &= \sum_{r=a}^{b-1}\left[\sin\left(r+\tfrac{1}{2}\right)\theta - \sin\left(r-\tfrac{1}{2}\right)\theta\right] \\ &= \sin\left(b-\tfrac{1}{2}\right)\theta - \sin\left(a-\tfrac{1}{2}\right)\theta, \end{aligned}$$

因為每個 $\sin\left(r+\frac{1}{2}\right)\theta$ 都與下一項的 $-\sin\left((r+1)-\frac{1}{2}\right)\theta$ 互相抵消，最後只剩下末項的上端和首項的下端。

現在使用 $\sin P - \sin Q = 2\cos\frac{P+Q}{2}\sin\frac{P-Q}{2}$，其中 $\frac{P+Q}{2} = \frac{1}{2}(a+b-1)\theta$ 及 $\frac{P-Q}{2} = \frac{1}{2}(b-a)\theta$：

$$\sin\tfrac{1}{2}\theta\, S(\theta) = \cos\tfrac{1}{2}(a+b-1)\theta\ \sin\tfrac{1}{2}(b-a)\theta. \qquad (*)$$

解 $S(\theta) = 0$

> **想一想** 我們可以直接把 $(*)$ 除以 $\sin\frac{1}{2}\theta$ 嗎？

只有當 $\sin\frac{1}{2}\theta \neq 0$，即 θ 不是 2π 的倍數時才可以。所以我們先處理這些值。若 $\theta = 2m\pi$，其中 m 為整數，則每一項 $\cos r\theta$ 都等於 1，故 $S(\theta) = b-a \geq 2$。這些值永遠不是解。

對於其餘所有 θ，我們有 $\sin\frac{1}{2}\theta \neq 0$，而由 $(*)$ 可得

$$S(\theta) = 0 \iff \cos\tfrac{1}{2}(a+b-1)\theta = 0 \text{ 或 } \sin\tfrac{1}{2}(b-a)\theta = 0.$$

- $\cos\frac{1}{2}(a+b-1)\theta = 0$ 當且僅當 $\frac{1}{2}(a+b-1)\theta = \left(k+\frac{1}{2}\right)\pi$，即

$$\theta = \frac{(2k+1)\pi}{a+b-1}, \qquad k \in \mathbb{Z}.$$

  （注意 $a+b-1 \geq 3$，所以我們並沒有除以零。）這些值全都不是 2π 的倍數：否則須有 $2k+1 = 2m(a+b-1)$，即一個奇數等於一個偶數。所以它們全部都是解。

- $\sin\frac{1}{2}(b-a)\theta = 0$ 當且僅當 $\frac{1}{2}(b-a)\theta = k\pi$，即

$$\theta = \frac{2k\pi}{b-a}, \qquad k \in \mathbb{Z}.$$

  這個值是 2π 的倍數，當且僅當 $b-a$ 整除 k，而我們已經知道這些值並不是解。所以必須把它們剔除。

> **想一想** 在 2π 的倍數處，$(*)$ 看來是成立的，為甚麼仍要把它們從第二族解中剔除？

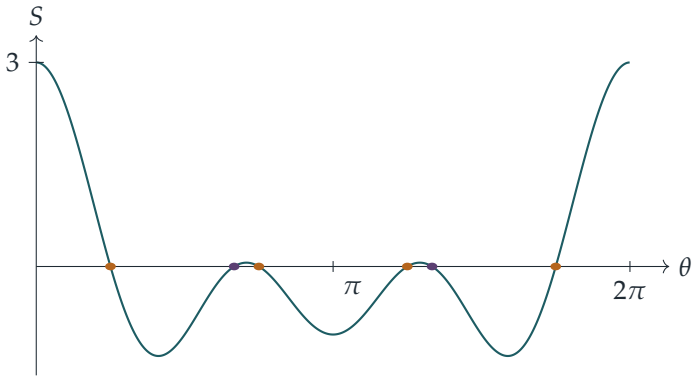
在 $\theta = 2m\pi$ 處，不論 S 取甚麼值，$(*)$ 的兩邊都是零，所以 $(*)$ 在這些位置完全沒有告訴我們關於 S 的任何資料。這只是全式乘以 $\sin\frac{1}{2}\theta$ 所造成的假象。由於 $b-a \geq 2$，第二族之中仍然有真正的解，例如 $\theta = 2\pi/(b-a)$。

綜合以上所得，全部的解正好是

$$\boxed{\theta = \frac{(2k+1)\pi}{a+b-1}\ (k \in \mathbb{Z}) \qquad \text{或} \qquad \theta = \frac{2k\pi}{b-a}\ (k \in \mathbb{Z},\ b-a \nmid k).}$$

這兩族解可能有重疊。這並無大礙，因為同時屬於兩族的值仍然只算一個解。

舉例來說，當 $a = 1$ 及 $b = 4$ 時，方程是 $\cos\theta + \cos 2\theta + \cos 3\theta = 0$。在 $(0, 2\pi)$ 內，它的根有來自第一族的 $\frac{\pi}{4}, \frac{3\pi}{4}, \frac{5\pi}{4}, \frac{7\pi}{4}$，以及來自第二族的 $\frac{2\pi}{3}, \frac{4\pi}{3}$。注意 $\theta = 0$ 和 $\theta = 2\pi$ 並不包括在內，該處 $S = 3$。



曲線 $S = \cos\theta + \cos 2\theta + \cos 3\theta$：琥珀色的根來自餘弦因子，紫紅色的根來自正弦因子。

驗算

- 取 $a = 1$、$b = 4$：在 $\theta = \frac{\pi}{4}$ 處，$\frac{\sqrt{2}}{2} + 0 - \frac{\sqrt{2}}{2} = 0$；在 $\theta = \frac{2\pi}{3}$ 處，$-\frac{1}{2} - \frac{1}{2} + 1 = 0$。
- 以數值方法檢驗，對多組 a, b 隨機選取 θ，$(*)$ 均成立。對所有 $1 \leq a \leq 6$、$a+2 \leq b \leq 11$，在 $(0, 6\pi)$ 上對 S 作細密掃描，找到的根與公式所給出的完全一致，再無其他。例如當 $a = 1$、$b = 3$ 時，根 $\theta = 13\pi/3 \approx 13.6136$ 來自第一族。

常見錯誤

- 未經檢查便把 $(*)$ 除以 $\sin\frac{1}{2}\theta$，因而把 $\theta = 2m\pi$ 也算作解；在這些位置 $S = b - a \neq 0$。

- 弄錯裂項相消後剩下的首末兩項：它們是 $\sin\left(b-\frac{1}{2}\right)\theta$ 和 $\sin\left(a-\frac{1}{2}\right)\theta$，因為最後一個餘弦是 $\cos(b-1)\theta$。
- 只給出某一區間（例如 $[0,2\pi)$）內的解，但題目要求的是全部的解。

STEP 之外

$\sum\cos r\theta$ 的閉合形式（closed form），其實是以 $e^{i\theta}$ 為公比的等比級數（geometric series）的實部。取 $a=-n$ 及 $b=n+1$，便得到狄利克雷核（Dirichlet kernel）$\sin\left(n+\frac{1}{2}\right)\theta/\sin\frac{1}{2}\theta$，它是傅立葉級數（Fourier series）理論以及其收斂（converge）問題的核心。想明白為何這類三角和一旦改用 $e^{i\theta}$ 來處理就會變得簡單，Needham 的 *Visual Complex Analysis* 是很好的入門讀物。

詳細解答

STEP III 2003，第 7 題

純數學 • 西姆松線

> **STEP III 2003，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/03-s3-q7

關鍵思路

要找垂足（foot of the perpendicular），可從 P 出發，沿直線的法線（normal）方向一直走到碰上該直線為止，所以寫 $N = P + \lambda(\text{法向量})$，再解出 λ。至於第二部分，關於 s 和 t 的條件是兩個斜率（gradient）之積等於 -1：它說明 PA 與 PB 互相垂直。把它化成圓的方程後，正好令共線（collinear）的驗證迎刃而解。

解答

N 的坐標

直線 AB 的方程為 $\dfrac{x}{a} + \dfrac{y}{b} = 1$，即

$$bx + ay = ab.$$

這條直線的一個法向量（normal vector）是 (b, a)，所以由 P 所作垂線的垂足 N 具有以下形式：

$$N = (s + \lambda b,\ t + \lambda a)$$

其中 λ 為某個數。由於 N 在 AB 上，

$$b(s + \lambda b) + a(t + \lambda a) = ab \quad \Longrightarrow \quad \lambda = \frac{ab - bs - at}{a^2 + b^2}.$$

因此

$$x = s + \lambda b = \frac{s(a^2 + b^2) + ab^2 - b^2 s - abt}{a^2 + b^2} = \frac{ab^2 - abt + a^2 s}{a^2 + b^2} = \frac{ab^2 - a(bt - as)}{a^2 + b^2},$$

$$y = t + \lambda a = \frac{t(a^2 + b^2) + a^2 b - abs - a^2 t}{a^2 + b^2} = \frac{a^2 b + b^2 t - abs}{a^2 + b^2} = \frac{a^2 b + b(bt - as)}{a^2 + b^2},$$

這正是所需的結果。

N 在 XY 上

首先看看這個條件說的是甚麼。要令它有意義，必須有 $s \neq 0$ 及 $s \neq a$；又因為乘積是 -1 而不是 0，所以還有 $t \neq 0$ 及 $t \neq b$。展開後得

$$t(t-b) = -s(s-a) \quad \Longleftrightarrow \quad s^2 + t^2 = as + bt. \tag{$*$}$$

現在 $X = (s, 0)$，$Y = (0, t)$。由於 $s \neq 0$ 及 $t \neq 0$，這是兩個不在原點的相異點，而直線 XY 為

$$\frac{x}{s} + \frac{y}{t} = 1, \qquad \text{即} \qquad tx + sy = st.$$

所以我們要證明 N 的坐標滿足 $tx + sy = st$。

> **想一想** 我們應該直接代入然後聽天由命，還是有更整潔的方法去整理這些代數運算？

為求簡潔，設 $k = bt - as$。利用上面的公式，

$$\begin{aligned}(a^2+b^2)(tx+sy) &= t(ab^2 - ak) + s(a^2b + bk) \\ &= ab^2t + a^2bs + k(bs - at).\end{aligned}$$

最後那個乘積展開後十分整齊：

$$k(bs-at) = (bt-as)(bs-at) = b^2st - abt^2 - abs^2 + a^2st = (a^2+b^2)st - ab(s^2+t^2).$$

因此

$$(a^2+b^2)(tx+sy) = (a^2+b^2)st + ab\big(as + bt - s^2 - t^2\big).$$

這對任何 P 都成立。根據 $(*)$，括號內的式子等於零，所以 $tx + sy = st$，即 N 在直線 XY 上。

幾何意義

PB 的斜率是 $\dfrac{t-b}{s}$，PA 的斜率是 $\dfrac{t}{s-a}$，所以該條件說的是 $PA \perp PB$：角 APB 是直角。等價地，$(*)$ 可寫成

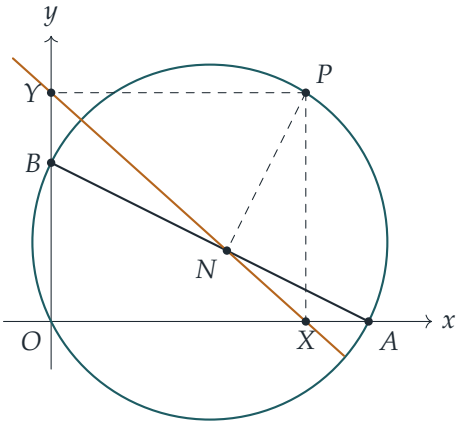
$$\left(s - \tfrac{a}{2}\right)^2 + \left(t - \tfrac{b}{2}\right)^2 = \frac{a^2+b^2}{4},$$

所以 P 在以 AB 為直徑的圓上。由於角 AOB 也是直角，這個圓通過 O：它就是三角形 OAB 的外接圓（*circumcircle*）。

X、Y 和 N 三點，分別是由 P 到三角形三邊 OA、OB 和 AB 所作垂線的垂足。所以這個結果說的是：

若 P 在一個三角形的外接圓上，則由 P 到三邊（有需要時延長）所作垂線的垂足共線。

這就是西姆松線（*Simson line*）定理，這裏證明的是直角三角形的情況，而直角頂點放在原點。



取 $a = 4$，$b = 2$：P 在通過 O、A、B 的圓上；直線 XY（琥珀色）通過 N。

驗算

- 取 $P = O$，即 $s = t = 0$：公式給出 $N = \left(\frac{ab^2}{a^2+b^2}, \frac{a^2b}{a^2+b^2}\right)$，它既在 $bx + ay = ab$ 上，也在通過 O 且垂直於 AB 的直線 $y = \frac{a}{b}x$ 上，與預期相符。
- 取 $a = 4$，$b = 2$，以及圓上的點 $P = (2 + \sqrt{5}\cos 1,\ 1 + \sqrt{5}\sin 1) \approx (3.208, 2.882)$，公式給出 $N \approx (2.214, 0.893)$，而 $\frac{2.214}{3.208} + \frac{0.893}{2.882} \approx 0.690 + 0.310 = 1$，所以 N 在 XY 上。
- 恆等式 $(a^2 + b^2)(tx + sy - st) = ab(as + bt - s^2 - t^2)$ 同時給出逆命題（converse）：若 P 不在坐標軸上，則僅當 P 在圓上時，N 才在 XY 上。

常見錯誤

- 把 (a, b) 當作 AB 的法向量。AB 的方向是 $(-a, b)$，所以法向量是 (b, a)。
- 一次過把所有東西展開，結果淹沒在算式之中。正如題目給出的答案所提示，把 $bt - as$ 當作一個符號來處理，代數運算只需寥寥數行。

- 寫到「角 APB 是直角」便停筆。題目想要的幾何意義是西姆松線：X、Y、N 是到三角形 OAB 三邊所作垂線的垂足。

STEP 之外

西姆松線對任何三角形都成立，其綜合幾何證明（synthetic proof）用到圓內接四邊形（cyclic quadrilateral）和同弓形內的圓周角相等。它的逆命題也成立，而當 P 沿外接圓移動時，各條西姆松線的包絡（envelope）是一條三尖瓣線（deltoid，即施泰納內擺線，Steiner's hypocycloid）。坐標法和複數法令這類結果變得按部就班便可推出；Needham 的 *Visual Complex Analysis*（《複分析：可視化方法》）展示了複數如何優雅地處理圓的幾何，而 Coxeter 與 Greitzer 的 *Geometry Revisited*（《幾何學的新探索》）則涵蓋西姆松線以及更多經典的三角形幾何。

詳細解答

STEP III 2003，第 8 題

純數學 • 一階微分方程

STEP III 2003，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/03-s3-q8

關鍵思路

第 (i) 部分是隱函數求導（implicit differentiation），但它其實是一個技巧的示範：曲線族 $(y+ax)^n(y+bx)=c$ 的斜率（gradient）是 x 和 y 的兩個線性式之比。第 (ii) 部分則把這個技巧倒過來用。我們對一般的曲線族求導，讀出以 a、b 和 n 表示的斜率，再選取這些數值，使它與題目給出的方程吻合。一旦吻合，這個式子沿每一條解曲線（solution curve）都是常數，而這個常數關係就是通解（general solution）。

解答

第 (i) 部分

利用乘積法則（product rule）和鏈式法則（chain rule），把 $(y+2x)^3(y-4x)=c$ 對 x 求導，並以 y' 表示 $\frac{\mathrm{d}y}{\mathrm{d}x}$：

$$3(y+2x)^2(y'+2)(y-4x)+(y+2x)^3(y'-4)=0.$$

在 $y+2x\neq 0$ 的點上，我們可以除以 $(y+2x)^2$：

$$3(y'+2)(y-4x)+(y+2x)(y'-4)=0.$$

把含 y' 的項合併，

$$y'\left[3(y-4x)+(y+2x)\right]=-6(y-4x)+4(y+2x),$$
$$y'(4y-10x)=32x-2y,$$

再把分子和分母同除以 2，

$$\frac{\mathrm{d}y}{\mathrm{d}x}=\frac{16x-y}{2y-5x},$$

這正是所求。（在直線 $y = -2x$ 上，曲線的 $c = 0$，而公式在那裏仍然給出正確的斜率 -2。）

第 (ii) 部分

設 $F = (y+ax)^n(y+bx)$，其中 y 是 x 的函數。與第 (i) 部分完全一樣，

$$\begin{aligned}\frac{\mathrm{d}F}{\mathrm{d}x} &= n(y+ax)^{n-1}(y'+a)(y+bx) + (y+ax)^n(y'+b)\\ &= (y+ax)^{n-1}\Big[y'\big((n+1)y+(nb+a)x\big) + (na+b)y + (n+1)ab\,x\Big].\end{aligned}$$

因此，在 $y = -ax$ 以外，F 沿一條曲線為常數，當且僅當

$$y' = -\frac{(na+b)y + (n+1)ab\,x}{(n+1)y + (nb+a)x}. \qquad (*)$$

（第 (i) 部分就是 $a = 2$、$b = -4$、$n = 3$ 的情況，得出 $(32x-2y)/(4y-10x)$。）

> **想一想** 一個分數可以有很多種寫法，那麼我們怎樣把 $(*)$ 與 $\dfrac{10x-4y}{3x-y}$ 配對起來？

分子和分母同乘一個數，分數的值不變，所以我們只能在相差一個公因子的意義下把兩者配對。把目標寫成 $-(4y-10x)/(y-3x)$，再把分子和分母同乘 $k = n+1$，使兩個分母中 y 的係數相同。於是我們需要

$$na+b = -4(n+1), \qquad (n+1)ab = 10(n+1), \qquad nb+a = -3(n+1).$$

由於 $n+1 \neq 0$（若 $n = -1$，$(*)$ 的分母便沒有 y 項），中間的方程給出 $ab = 10$。把首尾兩個方程相加，得 $(n+1)(a+b) = -7(n+1)$，所以 $a+b = -7$。因此 a 和 b 是 $t^2+7t+10 = 0$ 的根，即 -2 和 -5。把首尾兩個方程相減，得

$$(n-1)(a-b) = -(n+1).$$

- 取 $a = -5$、$b = -2$：$-3(n-1) = -(n+1)$，所以 $n = 2$。
- 取 $a = -2$、$b = -5$：$3(n-1) = -(n+1)$，所以 $n = \frac{1}{2}$。

兩種選擇描述的是同一批曲線，因為把 $(y-2x)^{1/2}(y-5x) = c$ 兩邊平方便得 $(y-5x)^2(y-2x) = c^2$。我們取 $n = 2$，這樣可以避開平方根。

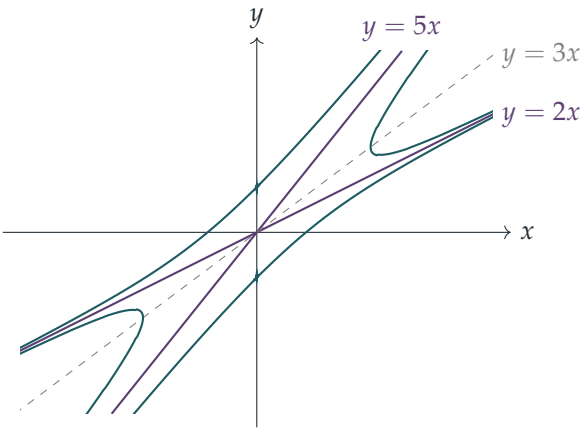
為穩妥起見，我們直接驗證。設 $G = (y-5x)^2(y-2x)$，

$$\begin{aligned}\frac{\mathrm{d}G}{\mathrm{d}x} &= (y-5x)\big[2(y'-5)(y-2x) + (y-5x)(y'-2)\big]\\ &= (y-5x)\big[3y'(y-3x) - 12y + 30x\big]\\ &= 3(y-5x)\big[y'(y-3x) - (4y-10x)\big].\end{aligned}$$

若 y 滿足該微分方程（differential equation），則 $y'(y-3x)=4y-10x$，所以方括號內的式子等於零，$\mathrm{d}G/\mathrm{d}x=0$。因此 G 在每一條解曲線上都是常數，通解為

$$\boxed{(y-5x)^2(y-2x)=C}$$

其中 C 為任意常數。$C=0$ 給出直線 $y=5x$ 和 $y=2x$。兩者確實都是解：在 $y=2x$ 上，右邊是 $2x/x=2$；在 $y=5x$ 上，右邊是 $(-10x)/(-2x)=5$。所有曲線在與直線 $y=3x$ 相交之處都有鉛垂切線，而方程在該直線上沒有定義。



曲線 $(y-5x)^2(y-2x)=\pm1$（藍綠色）和 $C=0$ 的直線解（紫紅色）。位於兩條直線之間的分支，在與虛線 $y=3x$ 相交之處有鉛垂切線。

第 (ii) 部分的另一種解法

右邊只取決於 y/x，所以代換 $y=vx$ 同樣可行。它給出

$$x\frac{\mathrm{d}v}{\mathrm{d}x}=\frac{(v-2)(v-5)}{3-v},$$

分離變數（separating the variables）並運用部分分式（partial fractions），可得 $\ln|x|=-\frac{1}{3}\ln|v-2|-\frac{2}{3}\ln|v-5|+$ 常數，整理後得到相同的答案。題目所用的方法則無須用到部分分式。

驗算

- 第 (i) 部分本身就是對 $(*)$ 的一次驗算：把 $a=2$、$b=-4$、$n=3$ 代入，便重現該處求得的斜率。
- 在點 $(1,0)$，方程給出斜率 $10/3$。在該點對 $(y-5x)^2(y-2x)=C$ 作隱函數求導，得 $(-5)\big[2(y'-5)(-2)+(-5)(y'-2)\big]=0$，所以 $-9y'+30=0$，$y'=10/3$。

- 數值方面，我們用龍格—庫塔法（Runge–Kutta method）從五十個隨機起點逐步求解該微分方程。沿每一條路徑，$(y-5x)^2(y-2x)$ 的值都保持不變，誤差約為 10^{-13}。

常見錯誤

- 不考慮公因子，直接把 $(*)$ 的係數與 $10x-4y$ 和 $3x-y$ 配對，結果得出互相矛盾的方程。
- 把答案留作 $(y-2x)^{1/2}(y-5x)=c$，而沒有留意到它其實是同一個曲線族；而且除非小心處理正負號，否則還會漏掉 $y<2x$ 的解。
- 忘記 $C=0$ 是容許的，它給出直線解 $y=2x$ 和 $y=5x$。

STEP 之外

這裏的方程是*齊次*（homogeneous）的：右邊只取決於 y/x。尋找一個沿着解保持不變的函數 $G(x,y)$，正是*恰當方程*（exact equations）和*積分因子*（integrating factors）背後的想法。在力學中，同樣的想法就是尋找能量之類的守恆量（conserved quantities）。Boyce 與 DiPrima 的 *Elementary Differential Equations*（初等微分方程）在開首幾章便涵蓋了這兩個課題。

詳細解答

STEP III 2003，第 9 題

力學 • 平衡的穩定性

STEP III 2003，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/03-s3-q9

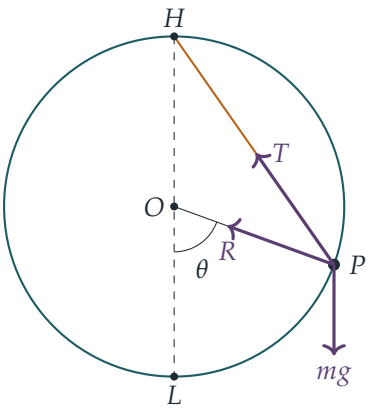
關鍵思路

金屬線是光滑的，所以它的反作用力（reaction）總是沿半徑方向，並不作功。因此一切都發生在切向（tangential）方向：把張力（tension）和重量沿切線分解，便得到一條關於 $\ddot{\theta}$ 的方程，平衡（equilibrium）位置和微小振動（small oscillations）都可由它求出。等腰三角形 OHP 的幾何性質提供了所需的全部角度。

解答

建立模型

把 P 看成光滑金屬線上的質點（particle），所以金屬線的反作用力 R 沿 PO 作用。繩子是輕繩，並服從胡克定律（Hooke's law）。取 $0 < \theta < \pi$，即 P 位於垂直直徑 LH 的其中一側。



由於 $OH = OP = a$，三角形 OHP 是等腰三角形，且 $\angle HOP = \pi - \theta$，所以它的兩個底角為

$$\angle OHP = \angle OPH = \tfrac{1}{2}\theta, \qquad HP = 2a\cos\tfrac{1}{2}\theta.$$

當 $HP > a$，即 $\cos\frac{1}{2}\theta > \frac{1}{2}$ 時，繩子處於繃緊（taut）狀態，這時由胡克定律得

$$T = \frac{\alpha mg}{a}\left(2a\cos\tfrac{1}{2}\theta - a\right) = \alpha mg\left(2\cos\tfrac{1}{2}\theta - 1\right).$$

現在沿 P 點的切線分解，取 θ 增加的方向（即朝向 H）為正。繩子與 PO 成 $\frac{1}{2}\theta$ 角，所以它與這條切線成 $\frac{1}{2}\pi - \frac{1}{2}\theta$ 角，其切向分量為 $T\sin\frac{1}{2}\theta$。重量的切向分量為 $-mg\sin\theta$，而 R 沒有切向分量。切向加速度為 $a\ddot{\theta}$，故

$$\begin{aligned} ma\ddot{\theta} &= T\sin\tfrac{1}{2}\theta - mg\sin\theta \\ &= mg\left[\alpha\left(2\cos\tfrac{1}{2}\theta - 1\right)\sin\tfrac{1}{2}\theta - 2\sin\tfrac{1}{2}\theta\cos\tfrac{1}{2}\theta\right], \end{aligned}$$

所以，只要繩子繃緊，

$$\ddot{\theta} = \frac{g}{a}\sin\tfrac{1}{2}\theta\left[2(\alpha-1)\cos\tfrac{1}{2}\theta - \alpha\right]. \qquad (*)$$

（把能量 $\frac{1}{2}ma^2\dot{\theta}^2 - mga\cos\theta + \frac{1}{2}\alpha mga(2\cos\frac{1}{2}\theta - 1)^2$ 求導，也會得到同一條方程。）

平衡位置

平衡時 $\ddot{\theta} = 0$。當 $0 < \theta < \pi$ 時，$\sin\frac{1}{2}\theta \neq 0$，所以我們需要

$$\cos\tfrac{1}{2}\theta = \frac{\alpha}{2(\alpha-1)},$$

由於 $\alpha > 1$，這個式子是有意義的。要由此得到 $0 < \theta < \pi$，我們需要 $0 < \cos\frac{1}{2}\theta < 1$。這個值是正數，而

$$\frac{\alpha}{2(\alpha-1)} < 1 \iff \alpha < 2\alpha - 2 \iff \alpha > 2.$$

我們還需要繩子在該處繃緊，因為 $(*)$ 是在這個假設下得出的：$\dfrac{\alpha}{2(\alpha-1)} > \dfrac{1}{2} \iff \alpha > \alpha - 1$，這恆成立。所以若 $\alpha > 2$，便存在一個滿足 $0 < \theta < \pi$ 的平衡位置，正如題目所求。

想一想 假如繩子在平衡位置是鬆弛（slack）的，會怎樣？

那樣就只有重力有切向分量，而 $-mg\sin\theta = 0$ 在 $0 < \theta < \pi$ 內無解。這就是為甚麼我們求得的平衡位置自動滿足 $\cos\frac{1}{2}\theta > \frac{1}{2}$ 這一點如此重要。

當 $\alpha = 2+\sqrt{2}$ 時的微小振動

這時

$$\cos\tfrac{1}{2}\theta = \frac{2+\sqrt{2}}{2(1+\sqrt{2})} = \frac{\sqrt{2}(\sqrt{2}+1)}{2(\sqrt{2}+1)} = \frac{1}{\sqrt{2}},$$

所以平衡位置在 $\frac{1}{2}\theta = \frac{1}{4}\pi$，即 $\theta = \frac{1}{2}\pi$，此時 P 與 O 處於同一水平。設 $\theta = \frac{1}{2}\pi + \phi$，則 $\ddot{\theta} = \ddot{\phi}$。把 $(*)$ 寫成 $\ddot{\theta} = \frac{g}{a} f(\theta)$，其中 $f(\frac{1}{2}\pi) = 0$。當 ϕ 很小時，

$$f(\tfrac{1}{2}\pi + \phi) \approx f'(\tfrac{1}{2}\pi)\,\phi.$$

由於 $(*)$ 中的方括號在 $\theta = \frac{1}{2}\pi$ 時等於零，使用乘積法則（product rule）後只剩下一項：

$$f'(\tfrac{1}{2}\pi) = \sin\tfrac{1}{4}\pi \cdot \left[-2(\alpha-1)\cdot\tfrac{1}{2}\sin\tfrac{1}{4}\pi\right] = -(\alpha-1)\sin^2\tfrac{1}{4}\pi = -\frac{\alpha-1}{2} = -\frac{\sqrt{2}+1}{2}.$$

因此，當 ϕ 很小時，

$$\ddot{\phi} \approx -\frac{g(\sqrt{2}+1)}{2a}\,\phi$$

正如題目所求。

最後部分

ϕ 的係數是負數，所以質點受到微小位移後，會圍繞 $\theta = \frac{1}{2}\pi$ 作（近似的）簡諧運動（simple harmonic motion）：這個平衡是穩定（stable）的。角頻率（angular frequency）為 $\omega = \sqrt{g(\sqrt{2}+1)/(2a)}$，所以週期（period）為

$$\frac{2\pi}{\omega} = 2\pi\sqrt{\frac{2a}{(\sqrt{2}+1)g}} = 2\pi\sqrt{\frac{2(\sqrt{2}-1)a}{g}},$$

而振幅（amplitude）會一直保持很小，因為能量守恆（金屬線光滑，繩子具彈性，所以沒有任何東西耗散能量）。

驗算

- 當 $\alpha = 2$ 時，平衡位置的公式給出 $\cos\frac{1}{2}\theta = 1$，即 $\theta = 0$：平衡位置與最低點重合，這與條件 $\alpha > 2$ 吻合。
- 當 $\alpha \to \infty$ 時，$\cos\frac{1}{2}\theta \to \frac{1}{2}$，所以 $\theta \to \frac{2}{3}\pi$ 而 $HP \to a$：非常硬的繩子幾乎不會伸長。
- 取 $a = 1.3\,\mathrm{m}$ 和 $g = 9.81\,\mathrm{m\,s^{-2}}$，由 $\phi = 0.01$ 開始對完整的方程 $(*)$ 作數值逐步計算，得到週期約為 $2.0819\,\mathrm{s}$，與 $2\pi\sqrt{2(\sqrt{2}-1)a/g} \approx 2.0818\,\mathrm{s}$ 相符。

常見錯誤

- 使用 $HP = 2a\sin\frac{1}{2}\theta$：那是弦 LP 的長度，而不是 HP。三角形 OHP 中位於 O 的角是 $\pi - \theta$。
- 忘記檢查繩子在平衡位置是否繃緊，因為鬆弛的繩子沒有張力。
- 沿徑向而不是沿切向分解：徑向方程會引入未知的反作用力 R，而且完全沒有提供關於 $\ddot{\theta}$ 的資訊。

STEP 之外

在平衡位置附近作線性化（linearising），$\ddot{\phi} \approx f'(\theta_0)\,\phi$，是研究任何動力系統（dynamical system）穩定性的標準第一步：$f'(\theta_0)$ 的正負號（等價地，就勢能（potential energy）而言，是 $V''(\theta_0)$ 的正負號）決定了系統是振動還是一去不返。在大學一年級的課程中，這會發展成非線性常微分方程（nonlinear ODEs）的相平面（phase plane）分析。Kleppner 與 Kolenkow 的 *An Introduction to Mechanics*（力學入門教科書）透過勢能來處理穩定平衡附近的微小振動，講解得十分清楚。

詳細解答

STEP III 2003，第 10 題

力學 • 變加速度

STEP III 2003，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/03-s3-q10

關鍵思路

加速度 $kx\dot{x}$ 是一個對時間的恰當導數（exact derivative）：它就是 $\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{1}{2}kx^2\right)$。因此我們可以立即積分一次，得到一階方程（first-order equation）$\dot{x} = \frac{1}{2}kx^2 + C$，而這個方程可以分離變量。常數 C 的正負決定一切：常數為正時得到反正切（arctangent）（解會爆破），常數為負時得到部分分式（partial fractions）和對數（解會穩定下來）。

解答

我們把質點看成在 x 軸上運動，並且 $\ddot{x} = kx\dot{x}$。由於 $\frac{\mathrm{d}}{\mathrm{d}t}(x^2) = 2x\dot{x}$，運動方程為

$$\frac{\mathrm{d}\dot{x}}{\mathrm{d}t} = \frac{\mathrm{d}}{\mathrm{d}t}\left(\tfrac{1}{2}kx^2\right), \qquad \text{所以} \qquad \dot{x} = \tfrac{1}{2}kx^2 + C$$

其中 C 為常數。當 $t = 0$ 時，$x = d$ 且 $\dot{x} = U$，所以 $C = U - \frac{1}{2}kd^2$，於是

$$\dot{x} = \tfrac{1}{2}k\left(x^2 - d^2\right) + U. \qquad (*)$$

第 (i) 部分

當 $U = kd^2$ 時，方程 $(*)$ 變成

$$\dot{x} = \tfrac{1}{2}k\left(x^2 + d^2\right).$$

右邊恆為正，所以我們可以分離變量（separate variables）：

$$\int_d^x \frac{\mathrm{d}u}{u^2 + d^2} = \int_0^t \tfrac{1}{2}k\,\mathrm{d}s \quad \Longrightarrow \quad \frac{1}{d}\left[\arctan\frac{x}{d} - \arctan 1\right] = \tfrac{1}{2}kt.$$

由於 $\arctan 1 = \frac{\pi}{4}$，由此得 $\arctan(x/d) = \frac{1}{2}dkt + \frac{\pi}{4}$，所以

$$x = d\tan\left(\frac{dkt}{2} + \frac{\pi}{4}\right).$$

只要 tan 的輻角保持小於 $\frac{\pi}{2}$，這個式子就成立（反正切的取值範圍是 $(-\frac{\pi}{2}, \frac{\pi}{2})$）。當 t 由 0 開始增加時，輻角由 $\frac{\pi}{4}$ 開始增加，並在以下時刻達到 $\frac{\pi}{2}$：

$$\frac{dkt}{2} = \frac{\pi}{4}, \qquad \text{即} \qquad t = \frac{\pi}{2dk}.$$

由於 $\tan\theta \to +\infty$ 當 $\theta \to \frac{\pi}{2}$ 從下方趨近時成立，而且 $d > 0$，所以當 $t \to \frac{\pi}{2dk}$ 時，$x \to \infty$，這正是所要證明的。質點在有限時間內到達無窮遠。

第 (ii) 部分

現在 $U < 0$，所以 $C = U - \frac{1}{2}kd^2 < 0$。設 $C = -\frac{1}{2}ka^2$，其中 $a > 0$，即

$$a = \sqrt{d^2 - \frac{2U}{k}}, \qquad \text{使得} \qquad \dot{x} = \tfrac{1}{2}k\left(x^2 - a^2\right).$$

因為 $U < 0$ 且 $k > 0$，所以 $a > d > 0$。

> ? **想一想** 部分分式所得的對數，應該用哪一個：$\ln\frac{x-a}{x+a}$ 還是 $\ln\frac{a-x}{a+x}$？

這取決於 x 落在哪個範圍。常數解 $x = \pm a$ 正是 $\dot{x} = 0$ 之處。起初 $x = d$ 嚴格地位於 $-a$ 與 a 之間，而由該處出發的解不可能越過常數解 $x = \pm a$（這類一階方程的兩個解不可能相遇），所以在任何時刻都有 $-a < x < a$。在這個範圍內 $\dot{x} < 0$，所以 x 遞減，而 $a - x$ 和 $a + x$ 這兩個量皆為正。

利用 $\dfrac{1}{x^2 - a^2} = \dfrac{1}{2a}\left(\dfrac{1}{x-a} - \dfrac{1}{x+a}\right)$ 並分離變量，

$$\frac{1}{2a}\left[\ln\frac{a-u}{a+u}\right]_d^x = \tfrac{1}{2}kt \quad \Longrightarrow \quad \frac{a-x}{a+x} = A\,\mathrm{e}^{akt}, \qquad \text{其中} A = \frac{a-d}{a+d}.$$

注意 $0 < A < 1$，因為 $a > d > 0$。解出 x，

$$x = a\,\frac{1 - A\mathrm{e}^{akt}}{1 + A\mathrm{e}^{akt}}, \qquad a = \sqrt{d^2 - \frac{2U}{k}}, \quad A = \frac{a-d}{a+d}.$$

這個式子對所有 $t \geq 0$ 都有定義（分母為正）。等價地，由於 $\ln A = -2\,\mathrm{artanh}(d/a)$，它可以寫成 $x = a\tanh\left(\mathrm{artanh}(d/a) - \frac{1}{2}akt\right)$。

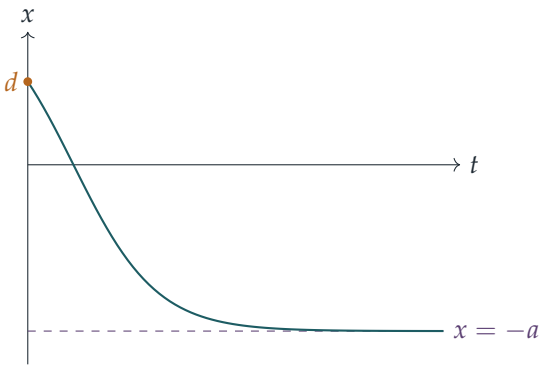
當 $t \to \infty$ 時，$A\mathrm{e}^{akt} \to \infty$，因為 $A > 0$ 且 $ak > 0$。把分子和分母同時除以 $A\mathrm{e}^{akt}$，

$$x = a\,\frac{A^{-1}\mathrm{e}^{-akt} - 1}{A^{-1}\mathrm{e}^{-akt} + 1} \longrightarrow -a.$$

所以 x 趨向極限（limit）

$$-a = -\sqrt{d^2 - \frac{2U}{k}}.$$

（這個極限也與 k 有關；它不能只用 d 和 U 來表示，因為固定 d 和 U 而改變 k，極限也會隨之改變。）



圖像（以 $d=1$、$a=2$、$k=1$ 繪畫）顯示質點由 d 出發，初速度為負，經過原點後越來越慢地趨近 $-a$。

驗算

- 在第 (i) 部分，$t=0$ 給出 $x=d\tan\frac{\pi}{4}=d$，而求導得 $\dot{x}=\frac{1}{2}kd^2\sec^2\frac{\pi}{4}=kd^2=U$。
- 在第 (ii) 部分，$t=0$ 給出 $x=a\frac{1-A}{1+A}=d$。當 $U\to 0^-$ 時，$a\to d$ 且 $A\to 0$，所以 x 一直停留在 d：這是正確的，因為 $x=d$、$\dot{x}=0$ 會令加速度為零。
- 數值上，以龍格－庫塔法（Runge–Kutta method）對隨機選取的 k、d 和 U 逐步求解 $\ddot{x}=kx\dot{x}$，結果與兩條公式吻合，誤差小於 10^{-6}。例如，取 $k=1$、$d=1$、$U=-1.5$ 時有 $a=2$，而數值計算所得的 x 穩定於 -2.000000。

常見錯誤

- 使用 $\ln\frac{x-a}{x+a}$ 而不加絕對值：當 $|x|<a$ 時，這是負數的對數。
- 求極限時取了正平方根：x 由 d 開始遞減，所以極限是 $-a$，而不是 $+a$。
- 把加速度寫成 $v\frac{\mathrm{d}v}{\mathrm{d}x}$，得到 $v\frac{\mathrm{d}v}{\mathrm{d}x}=kxv$，然後除以 v 而沒有指出 $v\neq 0$。這在本題行得通，但直接對 $\frac{\mathrm{d}}{\mathrm{d}t}$ 積分便可避開這個問題。

STEP 之外

第 (i) 部分是有限時間爆破（finite-time blow-up）的一個小例子：對於方程 $\dot{x}=f(x)$，若 f 的增長快於線性（這裏是二次增長），解便可能在有限時間內趨向無窮，所以解不一定對所有 t 都存在。第 (ii) 部分展示了另一種情況：常數解 $x=\pm a$ 是平衡點（equilibria），$-a$ 是穩定的而 $+a$ 是不穩定的，而且相線（phase line）圖在未作任何積分之前已能預

測極限。這兩個概念，連同阻止解互相穿越的唯一性定理（uniqueness theorem），在 Boyce and DiPrima, *Elementary Differential Equations*（常微分方程入門教科書）中都有論述。

詳細解答

STEP III 2003，第 11 題

力學 • 攔截運動學

> **STEP III 2003，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/03-s3-q11

關鍵思路

暫且放下運動過程，只考慮碰撞會發生的那一點。在那一刻，A、B 與碰撞點構成一個三角形，三邊分別是 d、P 走過的距離和 Q 走過的距離，而 B 處的角由方位角（bearing）決定。於是餘弦定理（cosine rule）給出一條方程，再把兩段走過的距離都用時間表示，便得到一條二次方程（quadratic）。能否發生碰撞，就變成這條二次方程有沒有正根（positive root）的問題。

解答

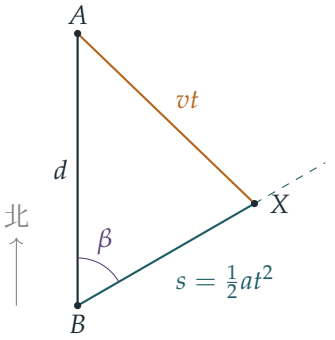
我們把 P 和 Q 視為在水平面上運動的質點（particle）。方位角由正北起順時針量度，且 $v > 0$。

碰撞方程

假設 Q 在時間 $t > 0$ 於點 X 擊中 P。到那時，P 已由靜止出發，以恆定加速度 a 沿一條固定直線運動，所以

$$BX = s = \tfrac{1}{2}at^2,$$

而 Q 以恆定速率沿直線運動，所以 $AX = vt$。



由 B 到 A 的方向是正北，而由 B 到 X 的方向的方位角為 β，所以角 ABX 等於 β（若 $\beta > 180°$，則為 $360° - \beta$，兩者的餘弦相同）。在三角形 ABX 中應用餘弦定理，得

$$v^2t^2 = d^2 + s^2 - 2ds\cos\beta.$$

（若 $\beta = 0$ 或 $180°$，這個「三角形」會退化成一條線段，但方程依然成立：它變成 $(vt)^2 = (d \mp s)^2$。）由於 $t^2 = 2s/a$，左邊等於 $2v^2s/a$，整理後得

$$s^2 - 2\Big(d\cos\beta + \frac{v^2}{a}\Big)s + d^2 = 0. \qquad (*)$$

反過來，若 $(*)$ 有一個根 $s > 0$，設 $t = \sqrt{2s/a}$，並設 X 為 P 在該時刻的位置。那麼 $(*)$ 所說的正是 $AX = vt$，而且 $AX > 0$。因此，若把 Q 由 A 朝着 X 射出，它會在時間 t 到達 X，與 P 同時到達，並擊中它。故此

$$\text{存在合適的方向} \iff (*)\text{ 有正根}s.$$

必要條件（necessary condition）

設 $k = d\cos\beta + v^2/a$，則 $(*)$ 就是 $s^2 - 2ks + d^2 = 0$。

> **想一想** 這條二次方程當然需要有實根，但這樣就足夠了嗎？關於這些根，我們還要知道甚麼？

這些根還必須是正數，因為 s 是走過的距離。兩根之積是 $d^2 > 0$，所以兩根同號，而兩根之和是 $2k$。因此，方程有正根，當且僅當判別式（discriminant）非負且 $k > 0$：

$$k^2 \geq d^2 \quad \text{且} \quad k > 0 \iff k \geq d.$$

（若 $k > 0$，則 $k^2 \geq d^2$ 等同於 $k \geq d$，因為 $d > 0$；而 $k \geq d$ 必然導致 $k > 0$。）

所以，若 Q 能擊中 P，則

$$d\cos\beta + \frac{v^2}{a} \geq d, \qquad \text{即} \qquad v^2 \geq ad(1 - \cos\beta),$$

這正是所要求的結果。

在 P 移動距離 d 之前擊中

現在假設 $v^2 > ad(1-\cos\beta)$，於是 $k > d > 0$。這時判別式 $4(k^2-d^2)$ 嚴格為正，$(*)$ 有兩個相異的根 $s_1 < s_2$。兩根皆為正數，因為兩根之和 $2k$ 與兩根之積 d^2 都是正數。由於兩根相異，且

$$s_1 s_2 = d^2,$$

它們不可能同時不小於 d，也不可能同時不大於 d，所以

$$s_1 < d < s_2.$$

明確地寫出來，$s_1 = k - \sqrt{k^2-d^2}$。

根據上文的逆向論證（converse），把 Q 朝着點 X_1 射出，其中該點與 B 相距 s_1 並位於 P 的路徑上，則 Q 會在時間 $t_1 = \sqrt{2s_1/a}$ 擊中 P，而在那一刻 P 只移動了 $s_1 < d$。所以 Q 可以瞄準某個方向，在 P 移動距離 d 之前擊中 P，這就是所要證明的。

驗算

- 取 $\beta = 0$：P 筆直地朝 A 走去，所以無論 v 的值是多少，兩者必定相遇。條件給出 $v^2 \geq 0$，這恆成立。這時 $(*)$ 變成 $s^2 - 2(d+v^2/a)s + d^2 = 0$，而較小的根小於 d：兩者在 P 到達 A 之前相遇，與常理相符。
- 取 $a = d = 1$、$\beta = 90^\circ$（即 P 向正東走），以及 $v^2 = 2 > 1 = ad(1-\cos\beta)$。這時 $(*)$ 是 $s^2 - 4s + 1 = 0$，所以 $s_1 = 2 - \sqrt{3} \approx 0.268 < 1$。在 $t_1 = \sqrt{2s_1}$ 時，我們有 $s_1^2 + 1 = 2t_1^2$，所以碰撞點與 A 的距離確實是 vt_1。我們又對 a、d、β 和 v 的大量隨機取值作了數值檢驗，既確認了在 $s_1 < d$ 處擊中，也確認了當 $v^2 < ad(1-\cos\beta)$ 時，P 與 A 的距離始終大於 vt。
- 在臨界情況 $v^2 = ad(1-\cos\beta)$ 下，我們得到 $k = d$ 和重根（double root）$s = d$：唯一可能的擊中時刻，恰好是 P 走了 d 的時候，這就是最後一部分需要嚴格不等式的原因。

常見錯誤

- 只要求判別式非負。這樣得到的是 $|k| \geq d$，它同時容許 $k \leq -d$，而這時兩根皆為負數，不會發生碰撞。
- 在餘弦定理中用錯了角。B 處的角是 BA（正北）與 BX（方位角 β）之間的夾角，所以它就是 β 本身，而不是 $180^\circ - \beta$。
- 在最後一部分只證明了必要條件。你還必須證明朝着 X_1 瞄準確實可行，這正是那個逆向論證。

STEP 之外

這是一道追蹤與攔截問題。這裏目標的路徑是預先知道的，而攔截者沿直線運動，所以問題歸結為一條關於相遇時間的方程。導彈制導和「追逐曲線」(pursuit curves) 則不同：追逐者不斷轉向目標，結果引出的是微分方程（differential equations）。先定下相遇點，再問誰能及時趕到，這種習慣是大學一年級力學的標準工具。Kleppner 與 Kolenkow 的 *An Introduction to Mechanics*（力學導論）對相對運動的運動學（kinematics）有細緻的闡述。

詳細解答

STEP III 2003，第 12 題

概率與統計 • 指數分佈的等候時間

STEP III 2003，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/03-s3-q12

關鍵思路

如果任何一段時間內的打擾次數都服從泊松分佈（Poisson distribution），那麼「下一次打擾在 t 分鐘之後才出現」與「在接下來的 t 分鐘內完全沒有打擾」是同一個事件，這樣就把離散分佈（discrete distribution）變成連續分佈（continuous distribution）：等候時間服從指數分佈（exponential distribution）。其後，相繼兩次打擾之間的各個間隔互相獨立（independent），所以尋找一個長間隔的過程，就是一連串獨立試驗，每次試驗不是「成功」（間隔達半小時或以上）就是「失敗」。餘下的工作交給幾何分佈（geometric distribution）和條件期望（conditional expectation）便可。

解答

T 的分佈

打擾平均每十分鐘出現一次，所以發生率是每分鐘 $\frac{1}{10}$ 次，而在長 t 分鐘的時段內，打擾次數服從平均值為 $\frac{t}{10}$ 的泊松分佈。

從剛剛發生的那次打擾開始計時。到下一次打擾的時間 T 不少於 t，當且僅當在接下來的 t 分鐘內沒有打擾，所以

$$\mathrm{P}(T \geq t) = \mathrm{P}(\text{沒有任何打擾，為時} t) = \mathrm{e}^{-t/10}.$$

因此，對於 $t \geq 0$，工作時間少於 t 分鐘的概率是

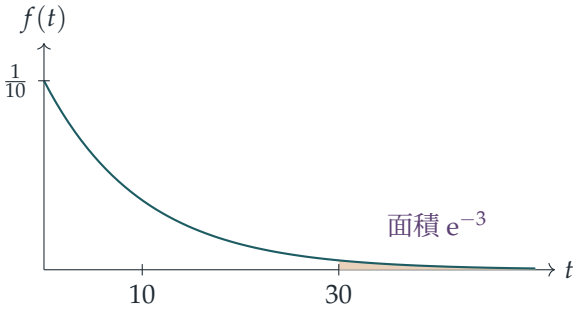
$$\mathrm{P}(T < t) = 1 - \mathrm{e}^{-t/10}.$$

這就是 T 的累積分佈函數（cumulative distribution function）$F(t)$（當 $t < 0$ 時 $F(t) = 0$）。

求導，得

$$f(t) = F'(t) = \tfrac{1}{10}\,\mathrm{e}^{-t/10} \quad (t \geq 0), \qquad f(t) = 0 \quad (t < 0).$$

所以 T 服從平均值為 10 分鐘的指數分佈。



打擾次數的期望值

由於這個過程沒有記憶（互不重疊的時段內的泊松計數互相獨立），每當一次打擾發生，到下一次打擾的時間都重新開始計算，其分佈與 T 相同，並且獨立於之前發生的一切。所以相繼的間隔 $T_1, T_2, \ldots$ 互相獨立，各自的分佈都與 T 相同。

如果一個間隔不少於 30 分鐘，就稱它為長間隔。每個間隔是長間隔的概率為

$$p = \mathrm{P}(T \geq 30) = \mathrm{e}^{-30/10} = \mathrm{e}^{-3},$$

而且各個間隔互相獨立。設 N 為第一個長間隔出現之前的打擾次數（剛剛發生的那一次不計算在內）。每個短間隔都以一次打擾告終，而第一個長間隔就是我那不受打擾的半小時，所以 N 就是第一個長間隔之前的短間隔數目。因此

$$\mathrm{P}(N = n) = (1-p)^n p, \qquad n = 0, 1, 2, \ldots$$

? **想一想** 這個幾何分佈由 0 而不是由 1 開始，我們怎樣求它的平均值？

對幾何級數（geometric series）$\sum_{n\geq 0} x^n = \frac{1}{1-x}$ 求導，可得當 $|x| < 1$ 時 $\sum_{n\geq 1} n x^{n-1} = \frac{1}{(1-x)^2}$。取 $x = 1-p$，

$$\mathrm{E}(N) = \sum_{n=1}^{\infty} n(1-p)^n p = p(1-p)\cdot\frac{1}{p^2} = \frac{1-p}{p}.$$

（另一個做法：$N+1$ 是直至首次成功為止（包括該次）的試驗次數，其平均值為 $\frac{1}{p}$。）代入 $p = \mathrm{e}^{-3}$，

$$\mathrm{E}(N) = \frac{1-\mathrm{e}^{-3}}{\mathrm{e}^{-3}} = \mathrm{e}^3 - 1,$$

正是所要求的結果。

短間隔長度的期望值

我們要求 $\mathrm{E}(T \mid T < 30)$。已知 $T < 30$ 時，T 的條件概率密度函數（conditional density）在 $0 \le t < 30$ 上是 $f(t)/\mathrm{P}(T < 30)$，所以

$$\mathrm{E}(T \mid T < 30) = \frac{1}{1 - \mathrm{e}^{-3}} \int_0^{30} \frac{t}{10}\, \mathrm{e}^{-t/10}\, \mathrm{d}t.$$

用分部積分法（integration by parts），取 $u = t$ 和 $\mathrm{d}v = \frac{1}{10}\mathrm{e}^{-t/10}\,\mathrm{d}t$，於是 $v = -\mathrm{e}^{-t/10}$，

$$\begin{aligned}\int_0^{30} \frac{t}{10}\, \mathrm{e}^{-t/10}\, \mathrm{d}t &= \Big[-t\, \mathrm{e}^{-t/10} \Big]_0^{30} + \int_0^{30} \mathrm{e}^{-t/10}\, \mathrm{d}t \\ &= -30\, \mathrm{e}^{-3} + 10\big(1 - \mathrm{e}^{-3}\big) \\ &= 10 - 40\, \mathrm{e}^{-3}.\end{aligned}$$

因此，把分子和分母同乘以 e^3，

$$\mathrm{E}(T \mid T < 30) = \frac{10 - 40\, \mathrm{e}^{-3}}{1 - \mathrm{e}^{-3}} = \frac{10\, \mathrm{e}^3 - 40}{\mathrm{e}^3 - 1} \text{ 分鐘} \approx 8.43 \text{ 分鐘}.$$

這個數值略小於 10，是合理的：我們已經把長間隔剔除了。

等候時間的期望值

在不受打擾的半小時開始之前，等候時間就是 N 個短間隔之和。已知 $N = n$ 時，這是 n 個獨立的間隔，每個都以短於 30 分鐘為條件，所以已知 $N = n$ 時等候時間的期望值是 $n\,\mathrm{E}(T \mid T < 30)$。再對 N 取平均，

$$\mathrm{E}(\text{等候時間}) = \mathrm{E}(N)\, \mathrm{E}(T \mid T < 30) = (\mathrm{e}^3 - 1) \cdot \frac{10\, \mathrm{e}^3 - 40}{\mathrm{e}^3 - 1} = 10\, \mathrm{e}^3 - 40 \text{ 分鐘},$$

即大約 161 分鐘，或 2 小時 41 分鐘。（如果想求直至試卷完成為止的時間，就要加上那半小時本身：$10\, \mathrm{e}^3 - 10 \approx 191$ 分鐘。）

驗算

- $\int_0^{\infty} \frac{1}{10}\mathrm{e}^{-t/10}\,\mathrm{d}t = 1$，所以 f 確實是一個概率密度函數，而它的平均值是 10，與「平均每十分鐘一次」吻合。
- 在沒有條件的情況下，$\mathrm{E}(T) = 10$ 應該可以分拆為 $\mathrm{P}(T < 30)\,\mathrm{E}(T \mid T < 30) + \mathrm{P}(T \ge 30)\,\mathrm{E}(T \mid T \ge 30)$。由無記憶性（memorylessness）可知 $\mathrm{E}(T \mid T \ge 30) = 40$，而果然有 $(10 - 40\mathrm{e}^{-3}) + 40\mathrm{e}^{-3} = 10$。
- 把這個過程模擬運行 200 000 次，得出平均打擾次數為 19.1（對照 $\mathrm{e}^3 - 1 \approx 19.09$），平均等候時間為 161.1 分鐘（對照 $10\mathrm{e}^3 - 40 \approx 160.86$）。

常見錯誤

- 使用 $\mathrm{E}(N) = \frac{1}{p} = \mathrm{e}^3$：這樣連長間隔也計算在內，但長間隔結束時，那半小時之前並沒有再出現打擾。
- 把 $\mathrm{E}(N)$ 乘以無條件平均值 10 來求等候時間。組成等候時間的間隔已知是短間隔，所以它們的平均值是 $\mathrm{E}(T \mid T < 30)$，而不是 10。

STEP 之外

這些打擾構成一個泊松過程（*Poisson process*），而各個間隔是互相獨立的指數分佈變量這個結果，正是泊松過程的基本性質之一；最後一步 $\mathrm{E}(\text{共}N\text{ 項之和}) = \mathrm{E}(N) \times \mathrm{E}(\text{每項})$ 是一個取條件的論證，與沃爾德恆等式（*Wald's identity*）關係密切。兩者合起來，通向更新理論（renewal theory）和排隊模型（queueing models），這些都是大學一、二年級概率課程的內容。Grimmett 與 Welsh 合著的 *Probability: An Introduction*（概率論入門）以合適的程度講解泊松過程和條件期望。

詳細解答

STEP III 2003，第 13 題

概率與統計 • 吸收馬可夫鏈

STEP III 2003，第 13 題

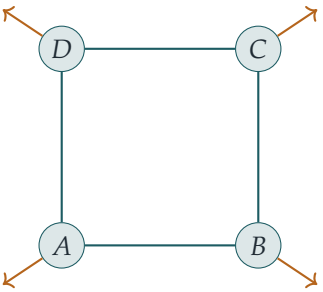
繼續閱讀之前，請先看題目（原題為英文）。題目載於 STEP Support Programme 的試題資料庫：

step.maths.org/questions/03-s3-q13

關鍵思路

這是一種「首步分析」(first-step analysis)：以兔子下一步的行動作為條件，把從某個洞室出發的概率，用牠能到達的各個洞室出發的概率來表示。這樣便得到四條線性方程，而正方形的對稱性（對對角線 AC 作反射，B 與 D 互換）可以減省工夫。最後一部分，我們把同一技巧用於一個修改過的兔穴，當中進入 C 即算作失敗，然後應用條件概率 (conditional probability) 的定義。

解答



通往地面的隧道

在整個解答中，每一步都以概率 $\frac{1}{3}$ 選出，並且與兔子過去的行蹤獨立（對於一隻「發現自己身處」某個洞室的兔子，我們就是這樣理解「可能性均等」的）。以 E_A 表示兔子最終經由 A 的隧道離開這一事件 (event)。

第 (i) 部分

假設兔子身處 A，並以牠的下一步作為條件。

- 牠有 $\frac{1}{3}$ 的概率立即從 A 的隧道往上走，此時 E_A 已經發生。
- 牠有 $\frac{1}{3}$ 的概率跑到 B。這時牠身處 B，而由於牠往後的行動與牠如何到達該處無關，牠最終從 A 離開的概率是 p_B。
- 牠有 $\frac{1}{3}$ 的概率跑到 D，同理，此時的概率是 p_D。

根據全概率公式（law of total probability），

$$p_A = \tfrac{1}{3}\cdot 1 + \tfrac{1}{3}p_B + \tfrac{1}{3}p_D,$$

正如所求。

第 (ii) 部分

從其餘洞室出發，用同樣的論證可再得三條方程。唯一的分別是：經由 B、C 或 D 的隧道離開，即表示 E_A *不發生*，所以該分支的貢獻是 $\frac{1}{3}\cdot 0$：

$$\begin{aligned} p_B &= \tfrac{1}{3}p_A + \tfrac{1}{3}p_C, \\ p_C &= \tfrac{1}{3}p_B + \tfrac{1}{3}p_D, \\ p_D &= \tfrac{1}{3}p_A + \tfrac{1}{3}p_C. \end{aligned}$$

p_B 與 p_D 的方程右邊相同，所以 $p_B = p_D$；把這個共同的值記作 b。（這就是對 AC 的反射對稱。）於是

$$p_A = \tfrac{1}{3} + \tfrac{2}{3}b, \qquad p_C = \tfrac{2}{3}b,$$

代入 p_B 的方程，得

$$b = \tfrac{1}{3}\left(\tfrac{1}{3} + \tfrac{2}{3}b + \tfrac{2}{3}b\right) = \tfrac{1}{9} + \tfrac{4}{9}b \quad\Longrightarrow\quad \tfrac{5}{9}b = \tfrac{1}{9} \quad\Longrightarrow\quad b = \tfrac{1}{5}.$$

因此

$$p_A = \tfrac{1}{3} + \tfrac{2}{15} = \frac{7}{15}, \qquad \text{並且} \qquad p_B = p_D = \tfrac{1}{5},\ p_C = \tfrac{2}{15}.$$

想一想 我們怎樣知道兔子一定會離開，從而確定這些方程真的能把各個概率定下來？

每一步兔子都有 $\frac{1}{3}$ 的概率離開，所以經過 n 步後牠仍在地底的機會是 $(\frac{2}{3})^n \to 0$：牠以概率 1 離開。此外，這四條線性方程只有一個解（我們解方程時並無任何選擇餘地），所以上面的數值就是所求的概率。

第 (iii) 部分

設 N 為兔子從未到過 C 這一事件。我們要求的是

$$\mathrm{P}(N \mid E_A) = \frac{\mathrm{P}(N \cap E_A)}{\mathrm{P}(E_A)},$$

其中兔子由 A 出發，所以分母是 $p_A = \frac{7}{15}$。

至於分子，設 q_A、q_B、q_D 分別為兔子現時身處 A、B、D（且至今未到過 C）的條件下，從未進入 C 而經由 A 的隧道離開的概率。首步分析的做法與之前完全一樣，只是現在走進 C 也算作失敗：

$$q_A = \tfrac{1}{3} + \tfrac{1}{3}q_B + \tfrac{1}{3}q_D,$$
$$q_B = \tfrac{1}{3}q_A + \tfrac{1}{3}\cdot 0 + \tfrac{1}{3}\cdot 0,$$

因為兔子從 B 出發，要麼跑回 A，要麼跑進 C（失敗），要麼經由錯的隧道離開（失敗）。根據對稱性，$q_D = q_B = \frac{1}{3}q_A$，所以

$$q_A = \tfrac{1}{3} + \tfrac{2}{9}q_A \quad\Longrightarrow\quad \tfrac{7}{9}q_A = \tfrac{1}{3} \quad\Longrightarrow\quad q_A = \tfrac{3}{7}.$$

因此

$$\mathrm{P}(N \mid E_A) = \frac{3/7}{7/15} = \frac{45}{49}.$$

驗算

- p_A, p_B, p_C, p_D 這四個值是從各個洞室出發而由 A 離開的概率。根據對稱性，它們也分別是從 A 出發而由 A、B、C、D 離開的概率，所以它們的和必須是 1：果然 $\frac{7}{15} + \frac{3}{15} + \frac{2}{15} + \frac{3}{15} = 1$。
- $p_A > p_B = p_D > p_C$，這合乎預期：兔子的起點離 A 越遠，從那裏離開的可能性就越低。
- (iii) 的答案接近 1：從 A 離開的兔子，通常只走了很短的路程便離開。模擬 4×10^5 隻兔子，得出 p_A 為 0.469，條件概率為 0.918，而理論值是 $\frac{7}{15} \approx 0.467$ 和 $\frac{45}{49} \approx 0.918$。

常見錯誤

- 在 p_B、p_C、p_D 的方程中也加入「立即離開」的 $\frac{1}{3}$ 一項：經由 A 以外任何一條隧道離開，貢獻都是 0。
- 以 $q_A = \frac{3}{7}$ 作為 (iii) 的答案。那是 $\mathrm{P}(N \cap E_A)$；題目問的是條件概率，所以我們還要除以 p_A。

STEP 之外

兔子的行動構成一條馬可夫鏈（Markov chain）：牠的下一步只取決於牠現時身處何處。地面是一個吸收態（absorbing state），而 p_A, p_B, p_C, p_D 是吸收概率（absorption probabilities），首步分析總能把它們化為一個線性方程組。同一想法還可以求出被吸收前的期望時間（expected time）、賭徒破產（gambler's ruin）概率等等。Grimmett 與 Welsh 合著的 *Probability: An Introduction*（概率入門）對馬可夫鏈有很好的初步介紹。

詳細解答

STEP III 2003，第 14 題

概率與統計 • 概率生成函數

STEP III 2003，第 14 題

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

step.maths.org/questions/03-s3-q14

關鍵思路

獨立得分之和的概率生成函數（probability generating function，簡稱 pgf），等於各個得分的 pgf 的乘積。因此，「哪些骰子會給出相同的總點數分佈」這個問題，就變成「一個多項式可以怎樣分拆成兩個因式」的問題。把所有式子分解成不可約（irreducible）的因式，再把這些因式分配給兩顆骰子，便可以讀出所有可能的重新標號方式。

解答

第一部分

設 X 為一顆標準公平骰子的得分。$1,\ldots,6$ 每個點數的概率都是 $\frac{1}{6}$，所以

$$G_X(t) = \mathrm{E}\big(t^X\big) = \tfrac{1}{6}\big(t + t^2 + t^3 + t^4 + t^5 + t^6\big) = \tfrac{1}{6}\,t\,\big(1 + t + t^2 + t^3 + t^4 + t^5\big).$$

要把括號內的式子因式分解，先把各項兩兩配對，然後再配成平方差：

$$1 + t + t^2 + t^3 + t^4 + t^5 = (1+t) + t^2(1+t) + t^4(1+t) = (1+t)(1+t^2+t^4),$$
$$1 + t^2 + t^4 = (1+t^2)^2 - t^2 = (1+t+t^2)(1-t+t^2).$$

因此

$$G_X(t) = \tfrac{1}{6}\,t\,(1+t)(1-t+t^2)(1+t+t^2).$$

若 X_1 和 X_2 是獨立擲出的兩顆骰子的得分，則 t^{X_1} 和 t^{X_2} 互相獨立（independent），所以 $\mathrm{E}\big(t^{X_1+X_2}\big) = \mathrm{E}\big(t^{X_1}\big)\mathrm{E}\big(t^{X_2}\big)$，而且

$$G_{X_1+X_2}(t) = G_X(t)^2 = \frac{1}{36}\,t^2(1+t)^2(1-t+t^2)^2(1+t+t^2)^2,$$

這正是題目要求的形式。

第二部分

設 Y 和 Z 分別為標號是 $1,2,2,3,3,4$ 和 $1,3,4,5,6,8$ 的兩顆骰子的得分。數一數各面的點數，得

$$G_Y(t) = \tfrac{1}{6}\left(t + 2t^2 + 2t^3 + t^4\right),$$
$$G_Z(t) = \tfrac{1}{6}\left(t + t^3 + t^4 + t^5 + t^6 + t^8\right).$$

對於 Y：$1 + 2t + 2t^2 + t^3 = (1+t)(1+t+t^2)$，把右邊展開即可驗證。所以

$$G_Y(t) = \tfrac{1}{6}\, t\,(1+t)(1+t+t^2).$$

對於 Z，第一部分中有用的乘積是 $(1+t)(1-t+t^2) = 1+t^3$ 和 $(1+t+t^2)(1-t+t^2) = 1+t^2+t^4$。於是

$$(1+t^3)(1+t^2+t^4) = 1+t^2+t^3+t^4+t^5+t^7,$$

再乘以 t，剛好得到 $t+t^3+t^4+t^5+t^6+t^8$。因此

$$G_Z(t) = \tfrac{1}{6}\, t\,(1+t)(1+t+t^2)(1-t+t^2)^2.$$

兩顆骰子獨立擲出，故

$$G_{Y+Z}(t) = G_Y(t)G_Z(t) = \frac{1}{36}\, t^2(1+t)^2(1-t+t^2)^2(1+t+t^2)^2 = G_{X_1+X_2}(t).$$

兩個 pgf 若作為多項式相等，它們的係數便一一相等，而 t^n 的係數就是總點數為 n 的概率。所以用這兩顆骰子擲出每個總點數的概率，都與用兩顆標準骰子時相同。

最後部分

一顆標準四面骰子的 pgf 是

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

所以兩顆這樣的骰子的總點數，其目標 pgf 是

$$\frac{1}{16}\, t^2(1+t)^2(1+t^2)^2.$$

假設兩顆新骰子的 pgf 分別是 $\frac{1}{4}f(t)$ 和 $\frac{1}{4}g(t)$。f 中 t^n 的係數是標號為 n 的面的數目，所以 f 和 g 都是係數為非負整數的多項式，並且 $f(1) = g(1) = 4$。由於每個標號都是正整數，所以 $f(0) = g(0) = 0$。我們需要

$$f(t)\, g(t) = t^2(1+t)^2(1+t^2)^2.$$

想一想 為甚麼 f 必定只由 t、$1+t$、$1+t^2$ 這幾個因式組成，而不會有其他因式？

因式 t 和 $1+t$ 是一次式，而 $1+t^2$ 沒有實根，所以三者在實數範圍內都是不可約的。多項式分解成不可約因式的方法是唯一的（不計常數倍），所以右邊的任何因式都必定是

$$f(t) = c\,t^k(1+t)^a(1+t^2)^b, \qquad g(t) = \tfrac{1}{c}\,t^{2-k}(1+t)^{2-a}(1+t^2)^{2-b},$$

其中 $0 \le k, a, b \le 2$，而常數 $c > 0$。

現在逐一運用 f 和 g 須符合的條件。

- *常數 c。* f 的最低次項係數是 c，而 g 的是 $1/c$。兩者都是面的數目，所以都是正整數，由此必有 $c = 1$。
- *t 的冪次。* $f(0) = 0$ 需要 $k \ge 1$，而 $g(0) = 0$ 需要 $2-k \ge 1$。所以 $k = 1$：每顆骰子恰好有一個因式 t。
- *四個面。* 由 $f(1) = 2^a \cdot 2^b = 4$ 得 $a+b = 2$，這時 $g(1) = 2^{2-a}2^{2-b} = 4$ 自動成立。

所以 (a,b) 是 $(2,0)$、$(1,1)$ 或 $(0,2)$，而 g 則取與之互補的選擇。由於把兩顆骰子對調會得出同一對骰子，所以只有兩種情況。

- 兩者皆為 $(a,b) = (1,1)$：$f = g = t(1+t)(1+t^2) = t + t^2 + t^3 + t^4$，即標準的一對 $1,2,3,4$ 和 $1,2,3,4$。
- $(a,b) = (2,0)$ 和 $(0,2)$：

  $$f(t) = t(1+t)^2 = t + 2t^2 + t^3, \qquad g(t) = t(1+t^2)^2 = t + 2t^3 + t^5,$$

  即標號為 $1,2,2,3$ 和 $1,3,3,5$ 的兩顆骰子。

在每種情況中，兩個多項式的係數都是非負整數，而且總和為 4，所以兩者確實都是骰子；而按照構造方法，它們會給出正確的 pgf。

因此，僅有的標號方式就是標準的一對，以及 $\{1,2,2,3\}$、$\{1,3,3,5\}$ 這一對。

驗算

對於那對非標準的四面骰子，把總點數逐一列出：16 個等可能的和是 2, 4, 4, 6, 3, 5, 5, 7, 3, 5, 5, 7, 4, 6, 6, 得出總點數 $2,\ldots,8$ 的頻數為 $1,2,3,4,3,2,1$，與兩顆標準骰子完全相同。把 $t = 1$ 代入上面任何一個 pgf 都會得到 1，這是檢查每個因式分解的快捷方法。用電腦搜尋所有標號由 1 至 7 的四面骰子對，找到的恰好就是上述兩對；而 $1,2,2,3,3,4$ 和 $1,3,4,5,6,8$ 這兩顆骰子的總點數，亦與標準分佈完全吻合（例如兩種情況下都有 P(總點數 $= 7$) $= \frac{6}{36}$）。

常見錯誤

- 忘記了常數 c：你需要說明為甚麼不可以一顆骰子帶有因子 2，而另一顆帶有因子 $\frac{1}{2}$。理由就是兩顆骰子的最低次項係數都必須是正整數。
- 容許 $k = 0$ 或 $k = 2$。這樣其中一顆骰子便會有一面標號為 0，而這並不是正整數。
- 把 $(1 - t + t^2)$ 分解成複數一次因式。在實數範圍內它是不可約的，而這正是分配因式時所需要的形式。

STEP 之外

這裏的一對六面骰子稱為 *Sicherman* 骰子（Sicherman dice），用同樣的方法可以證明，它們是唯一一對標號為正整數的非標準六面骰子。這個論證建基於多項式的唯一分解（unique factorisation）；你會在大學第一個環與域（rings and fields）的課程中遇到它，它是算術基本定理（fundamental theorem of arithmetic）的多項式版本。生成函數亦能把許多困難的計數問題化為代數問題。關於這兩個課題，可參閱 Allenby 的 *Rings, Fields and Groups*（環、域與羣），以及 Grimmett 與 Welsh 的 *Probability: An Introduction*（概率論入門）。

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

2003 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。