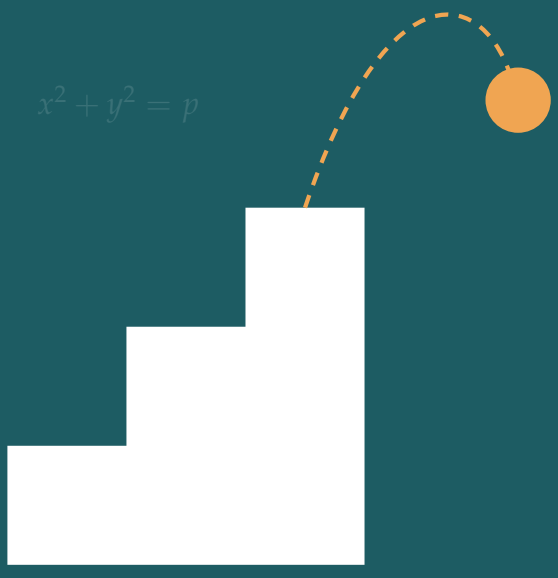
STEP 備試 · 大學數學

Beyond STEP

STEP I, II 及 III

2007 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

2007 年詳細解答

Keith McNulty

Beyond STEP：以歷屆 STEP 試題為門，走進大學數學

2007 年詳細解答

2007 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 數字排列的計數 純數學 . . . 3

第 2 題 反正切恆等式 純數學 . . . 6

第 3 題 三角函數積分 純數學 . . . 10

第 4 題 二次方程的公共根 純數學 . . . 13

第 5 題 多面體的幾何 純數學 . . . 16

第 6 題 麥克勞林級數 純數學 . . . 20

第 7 題 直線之間的距離 純數學 . . . 23

第 8 題 三次曲線的相切 純數學 . . . 27

第 9 題 斜坡上的極限摩擦力 力學 . . . 33

第 10 題 位移–時間圖 力學 . . . 37

第 11 題 拋體運動 力學 . . . 41

第 12 題 條件概率 概率與統計 . . . 45

第 13 題 條件概率 概率與統計 . . . 49

第 14 題 泊松參數的唯一性 概率與統計 . . . 53

STEP II

第 1 題 二項式近似 純數學 . . . 58

第 2 題 三次曲線與面積 純數學 . . . 61

第 3 題 代換積分法 純數學 . . . 65

第 4 題 三角方程 純數學 . . . 68

第 5 題 迭代函數 純數學 . . . 72

第 6 題 一階微分方程 純數學 . . . 76

第 7 題 琴生不等式 純數學 . . . 79

第 8 題 以向量證明塞瓦定理 純數學 . . . 82

第 9 題 斜碰撞與恢復係數 力學 . . . 86
第 10 題 滑及摩擦力的靜力學 力學 . . . 90
第 11 題 三維空間中的拋體 力學 . . . 93
第 12 題 幾何分佈的最大值 概率與統計 . . . 97
第 13 題 生日問題 概率與統計 . . . 100
第 14 題 分段概率密度函數 概率與統計 . . . 104

STEP III

第 1 題 正切和與韋達定理 純數學 . . . 107
第 2 題 二項式冪級數 純數學 . . . 110
第 3 題 柯西不等式 純數學 . . . 114
第 4 題 參數曲線與曲率 純數學 . . . 117
第 5 題 高階導數與數學歸納法 純數學 . . . 121
第 6 題 複數與圓 純數學 . . . 125
第 7 題 代換積分法 純數學 . . . 129
第 8 題 里卡蒂方程 純數學 . . . 133
第 9 題 簡諧運動 力學 . . . 136
第 10 題 斜面上的拋體 力學 . . . 140
第 11 題 擺過邊緣 力學 . . . 144
第 12 題 隨機和 概率與統計 . . . 148
第 13 題 隨機漫步的期望值 概率與統計 . . . 151
第 14 題 幾何概率 概率與統計 . . . 155

詳細解答

STEP I 2007，第 1 題

純數學 • 數字排列的計數

> **STEP I 2007，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/07-s1-q1

關鍵思路

不要嘗試把平衡數逐一列出。相反，應按兩組數字之和的共同值 s 把它們分類。對於每個 s，前面一對數字和後面一對數字可以獨立地選取，因此數字和為該值的平衡數的數目，就是兩個計數的乘積。於是整個問題歸結為計算把 s 寫成兩個數字之和的方法數目，而這等同於計算正方形網格某條對角線上的格點（lattice point）數目。

解答

把一個四位平衡數寫成 $abcd$，其中每個數字都屬於 $\{0, 1, \ldots, k\}$，$a \geq 1$（首位數字不能是 0），並且

$$a + b = c + d = s$$

其中 s 滿足 $0 \leq s \leq 2k$。對於每個 s，設

- $N(s)$ 為滿足 $c + d = s$ 的數字有序對（ordered pair）(c, d) 的數目；
- $M(s)$ 為滿足 $a \geq 1$ 及 $a + b = s$ 的數字有序對 (a, b) 的數目。

一旦固定了 s，任何一個合乎條件的前數對都可以配上任何一個合乎條件的後數對，所以數字和為 s 的平衡數共有 $M(s)N(s)$ 個。不同的 s 值給出不同的數，因此

$$\text{總數} = \sum_{s=0}^{2k} M(s)\, N(s).$$

計算數對的數目

$N(s)$。若 $0 \le s \le k$，則 c 可以是 $0, 1, \ldots, s$ 中的任何一個，而 $d = s - c$ 自動是一個數字。所以 $N(s) = s + 1$。若 $k < s \le 2k$，則 $d = s - c \le k$ 迫使 $c \ge s - k$，所以 c 由 $s - k$ 取到 k，而 $N(s) = 2k + 1 - s$。

$M(s)$。這個計數與上面相同，只是要減去 $a = 0$ 的那一對（如果它存在的話）。那一對是 $(0, s)$，而它只在 $s \le k$ 時才是容許的。所以

$$M(s) = \begin{cases} s, & 0 \le s \le k, \\ 2k + 1 - s, & k < s \le 2k. \end{cases}$$

特別地，$M(0) = 0$：沒有平衡數的數字和是 0，這是對的，因為 $a \ge 1$。

> **想一想** 為甚麼我們不能只計算所有滿足 $a + b = c + d$ 的四位數字串，然後就此作結？

因為像 0312 這樣的數字串並不是四位數。條件 $a \ge 1$ 只影響前數對，而且只影響較小的和，這正正就是 $M(s)$ 與 $N(s)$ 只在 $s \le k$ 時才有分別的原因。

第 (i) 部分

這裏 $k = 4$，而和的取值由 0 到 8：

s	0	1	2	3	4	5	6	7	8
$M(s)$	0	1	2	3	4	4	3	2	1
$N(s)$	1	2	3	4	5	4	3	2	1
$M(s)N(s)$	0	2	6	12	20	16	9	4	1

最底一行加起來是 $(2 + 6 + 12 + 20) + (16 + 9 + 4 + 1) = 40 + 30 = 70$。所以用數字 0 至 4 可以組成七十個四位平衡數，正如題目所求。

第 (ii) 部分

現在利用 M 和 N 的公式，在 $s = k$ 處把和式拆開：

$$\text{總數} = \sum_{s=1}^{k} s(s+1) \; + \; \sum_{s=k+1}^{2k} (2k + 1 - s)^2.$$

（$s=0$ 的一項是零。）在第二個和式中設 $r=2k+1-s$。當 s 由 $k+1$ 取到 $2k$ 時，r 由 k 遞減至 1，所以第二個和式就是 $\sum_{r=1}^{k} r^2$。因此

$$\begin{aligned}\text{總數} &= \sum_{r=1}^{k}(r^2+r)+\sum_{r=1}^{k} r^2 = 2\sum_{r=1}^{k} r^2+\sum_{r=1}^{k} r \\ &= \tfrac{1}{3}k(k+1)(2k+1)+\tfrac{1}{2}k(k+1) \\ &= \tfrac{1}{6}k(k+1)\big(2(2k+1)+3\big) \\ &= \tfrac{1}{6}k(k+1)(4k+5),\end{aligned}$$

其中用到題目給出的 $\sum r^2$ 恆等式（$r=0$ 的一項沒有貢獻）以及 $\sum_{r=1}^{k} r=\frac{1}{2}k(k+1)$。這就是所要求的結果。

驗算

- 把 $k=4$ 代入公式，得 $\frac{1}{6}\cdot 4\cdot 5\cdot 21=70$，與第 (i) 部分吻合。
- 當 $k=1$ 時，公式給出 3。直接列舉，只用 0 和 1 的四位平衡數是 1001、1010 和 1111。
- 對所有滿足 $a\geq 1$ 的數字串 $abcd$ 作窮舉（brute-force）計數，結果在 k 由 1 到 15 的每個情況下都與公式相符（例如 $k=9$，即用上全部十進制數字時，結果是 615）。

常見錯誤

- 忘記首位數字的條件，結果得出的是 $\sum N(s)^2=\frac{1}{3}(k+1)(2k^2+4k+3)$，當 $k=4$ 時這是 85。
- 在 $s>k$ 時算錯數對的數目：它是 $2k+1-s$，而不是 $s+1$，因為兩個數字都不能超過 k。
- 把中間的和 $s=k$ 在拆開後的兩半中都計算了，或者兩半都沒有計算。

STEP 之外

計數 $N(s)$ 是 $(1+x+\cdots+x^k)^2$ 中 x^s 的係數，而平衡數的總數則是若干個這類多項式之積的某個係數。這種生成函數（generating function）方法同樣能處理 $2n$ 位的平衡數，而在那種情況下直接計數很快便會變得難以應付。它是大學本科組合數學（combinatorics）的核心工具。想做更多同類的計數問題，可試試 Siklos 的 *Advanced Problems in Mathematics*（進階數學問題集）。

詳細解答

STEP I 2007，第 2 題

純數學 • 反正切恆等式

> **STEP I 2007，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/07-s1-q2

關鍵思路

只要對等式兩邊取正切，再運用 $\tan(A+B)$ 的加法公式（addition formula），一條關於反正切（inverse tangent）的等式就會變成一條關於普通分數的等式。代價是 tan 每隔 π 便重複一次，所以 $\tan(A+B)=1$ 只告訴我們 $A+B=\frac{1}{4}\pi+k\pi$；我們還需要交代一下各個角的大小，才能確定 k。在此之後，每一部分都是一小段數論（number theory）：作因式分解（factorise），或者運用整除性（divisibility）論證。

解答

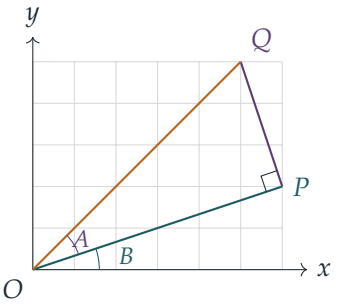
第 (i) 部分

由於 $\tan A=\frac{1}{2}$ 及 $\tan B=\frac{1}{3}$，

$$\tan(A+B)=\frac{\tan A+\tan B}{1-\tan A\tan B}=\frac{\frac{1}{2}+\frac{1}{3}}{1-\frac{1}{6}}=\frac{\frac{5}{6}}{\frac{5}{6}}=1.$$

現在 A 和 B 都是銳角（acute），所以 $0<A+B<\pi$，而在 $(0,\pi)$ 內正切為 1 的角只有 $\frac{1}{4}\pi$。因此 $A+B=\frac{1}{4}\pi$，正如所求。

下圖在方格紙上展示同一事實。取 $O=(0,0)$、$P=(6,2)$ 及 $Q=(5,5)$，直線 OP 與 x 軸的夾角為 $\arctan\frac{1}{3}$；$PQ=(-1,3)$ 垂直於 $OP=(6,2)$，而且長度是它的一半，所以角 POQ 是 $\arctan\frac{1}{2}$；而 Q 位於直線 $y=x$ 上，即在 $\frac{1}{4}\pi$ 的方向。



第 (i) 部分：整數 p 和 q

首先排除 $pq = 1$ 的情況，因為這時加法公式會出現除以零。此時 $p = q = 1$ 或 $p = q = -1$，左邊分別是 $\frac{1}{2}\pi$ 和 $-\frac{1}{2}\pi$，兩者都不是 $\frac{1}{4}\pi$。所以 $pq \neq 1$，對題目所給等式的兩邊取正切，得

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

分子和分母同乘以 pq，得 $\dfrac{p+q}{pq-1} = 1$，所以

$$p + q = pq - 1 \quad \Longrightarrow \quad pq - p - q + 1 = 2 \quad \Longrightarrow \quad (p-1)(q-1) = 2,$$

這就是所要求的結果。

由於 $p-1$ 和 $q-1$ 是乘積為 2 的整數，數對 $(p-1, q-1)$ 必為以下其中之一：

$$(1,2), \quad (2,1), \quad (-1,-2), \quad (-2,-1),$$

由此得 $(p,q) = (2,3)$、$(3,2)$、$(0,-1)$ 或 $(-1,0)$。最後兩個要捨去，因為 p 和 q 都不是零。所以 $\{p,q\} = \{2,3\}$，而根據第 (i) 部分的前半，兩種次序確實都滿足該等式。

? **想一想** 我們只證明了解必須滿足 $(p-1)(q-1) = 2$。為甚麼事後還要檢查這些候選答案？

取正切會失去資訊：$(p-1)(q-1) = 2$ 等價於 $\tan(\arctan\frac{1}{p} + \arctan\frac{1}{q}) = 1$，而後者也容許和為 $-\frac{3}{4}\pi$ 之類的值。這裏的候選數對要麼含有 0（於是 $\frac{1}{p}$ 沒有意義），要麼是 2 和 3，而我們已知這一對是成立的。

第 (ii) 部分

設 $\theta = \arctan\frac{1}{r}$ 及 $\phi = \arctan\dfrac{s}{s+t}$。對題目所給的等式取正切，

$$\frac{\frac{1}{r} + \frac{s}{s+t}}{1 - \frac{s}{r(s+t)}} = 1.$$

分母不為零：由於 $r \geq 1$ 及 $t \geq 1$，我們有 $r(s+t) > s$。分子和分母同乘以 $r(s+t)$，

$$\frac{(s+t)+rs}{r(s+t)-s} = 1 \quad \Longrightarrow \quad s+t+rs = rs+rt-s \quad \Longrightarrow \quad t(r-1) = 2s.$$

想一想 關於 s 和 t 的最大公因數（highest common factor）的條件，在哪裏派上用場？

等式 $t(r-1) = 2s$ 表示 t 整除 $2s$。由於 $\text{hcf}(s,t) = 1$，t 與 s 沒有公共的質因數（prime factor），所以 t 的每個質因數都必須來自那個 2：即 t 整除 2。由於 t 是正整數，可能的值只有

$$t = 1 \quad \text{或} \quad t = 2.$$

- 若 $t = 1$：$r - 1 = 2s$，所以 $r = 2s + 1$。
- 若 $t = 2$：$2(r-1) = 2s$，所以 $r = s + 1$。（這裏 s 必須是奇數，因為 $\text{hcf}(s,2) = 1$。）

兩種情況都確實會出現。事實上 $0 < \frac{1}{r} \leq 1$ 及 $0 < \frac{s}{s+t} < 1$，所以 θ 和 ϕ 都落在 $(0, \frac{1}{4}\pi]$ 內，並且 $0 < \theta + \phi < \frac{1}{2}\pi$。把上面的代數運算倒推回去，$t(r-1) = 2s$ 給出 $\tan(\theta+\phi) = 1$，於是 $\theta + \phi = \frac{1}{4}\pi$ 準確成立。例如，$t = 1$, $s = 1$ 給出 $\arctan\frac{1}{3} + \arctan\frac{1}{2}$，而 $t = 2$, $s = 1$ 給出 $\arctan\frac{1}{2} + \arctan\frac{1}{3}$：正是第 (i) 部分的恆等式，兩種次序各出現一次。

驗算

- 用計算機計算，$\arctan 0.5 + \arctan(1/3) = 0.4636 + 0.3218 = 0.7854 = \frac{1}{4}\pi$。
- 用電腦搜尋所有滿足 $|p|, |q| \leq 200$ 的非零 p, q，只找到 $(2,3)$ 和 $(3,2)$。
- 在 $r < 400$ 及 $s, t < 150$ 且 $\text{hcf}(s,t) = 1$ 的範圍內搜尋，只找到 $t = 1$, $r = 2s+1$ 和 $t = 2$, $r = s+1$；例如 $\arctan\frac{1}{8} + \arctan\frac{7}{9} = \frac{1}{4}\pi$（$s = 7$, $t = 2$），在機器精度內成立。

常見錯誤

- 由 $\tan(A+B) = 1$ 直接推斷 $A + B = \frac{1}{4}\pi$，卻沒有說明為何 $A + B$ 不可能是 $\frac{5}{4}\pi$。
- 忘記了 2 的負因數分解 $(-1)(-2)$ 和 $(-2)(-1)$，或者忘記因為它們會給出零而要把它們捨去。
- 寫「t 整除 $2s$，所以 t 整除 2」而沒有引用 $\text{hcf}(s,t) = 1$：這正是需要用到該假設的地方。

STEP 之外

tan 的加法公式其實就是複數（complex numbers）的乘法：$\arctan\frac{1}{p}$ 是 $p+\mathrm{i}$ 的輻角（argument），所以第 (i) 部分所說的，就是 $(2+\mathrm{i})(3+\mathrm{i}) = 5+5\mathrm{i}$ 的輻角為 $\frac{1}{4}\pi$。像梅欽（Machin）公式 $4\arctan\frac{1}{5} - \arctan\frac{1}{239} = \frac{1}{4}\pi$ 這類恆等式，曾被用來把 π 計算到數百位，它們正是以完全相同的方式，由高斯整數（Gaussian integers）的因數分解得來。Needham 的 *Visual Complex Analysis* 把複數乘法的幾何闡述得十分精彩，而 Davenport 的 *The Higher Arithmetic* 則涵蓋了第 (ii) 部分所用的整除性論證。

詳細解答

STEP I 2007，第 3 題

純數學 • 三角函數積分

> **STEP I 2007，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/07-s1-q3

關鍵思路

$\cos^4\theta$ 和 $\sin^4\theta$ 單獨來看都不易積分，但只要利用 $\cos^2\theta+\sin^2\theta=1$ 把兩者的差與和化為 2θ 的函數，這兩個式子就都容易積分（integrate）了。把和與差分別積分，便得到關於所求兩個積分的兩條簡單線性方程。同樣的技巧再升一級，就能處理六次方的情況。

解答

恆等式

把平方差（difference of two squares）因式分解，

$$\cos^4\theta-\sin^4\theta=(\cos^2\theta-\sin^2\theta)(\cos^2\theta+\sin^2\theta)=\cos 2\theta\cdot 1=\cos 2\theta.$$

至於兩者之和，我們用配方法（completing the square）：

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

其中用到 $\sin 2\theta=2\sin\theta\cos\theta$。兩個恆等式（identity）均已得證。

四次方

記 $C_4=\int_0^{\pi/2}\cos^4\theta\,\mathrm{d}\theta$ 及 $S_4=\int_0^{\pi/2}\sin^4\theta\,\mathrm{d}\theta$。把第一個恆等式積分，

$$C_4-S_4=\int_0^{\pi/2}\cos 2\theta\,\mathrm{d}\theta=\left[\tfrac{1}{2}\sin 2\theta\right]_0^{\pi/2}=0.$$

對於第二個恆等式，我們需要 $\int_0^{\pi/2} \sin^2 2\theta \,\mathrm{d}\theta$。利用 $\sin^2 2\theta = \frac{1}{2}(1-\cos 4\theta)$，

$$\int_0^{\pi/2} \sin^2 2\theta \,\mathrm{d}\theta = \left[\tfrac{1}{2}\theta - \tfrac{1}{8}\sin 4\theta\right]_0^{\pi/2} = \frac{\pi}{4},$$

因為 $\sin 2\pi = 0$。因此

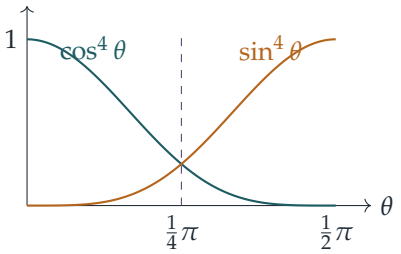
$$C_4 + S_4 = \int_0^{\pi/2} \left(1 - \tfrac{1}{2}\sin^2 2\theta\right)\mathrm{d}\theta = \frac{\pi}{2} - \frac{1}{2}\cdot\frac{\pi}{4} = \frac{3\pi}{8}.$$

把兩條方程相加和相減，

$$\int_0^{\pi/2} \cos^4\theta \,\mathrm{d}\theta = \int_0^{\pi/2} \sin^4\theta \,\mathrm{d}\theta = \frac{3\pi}{16}.$$

想一想 我們能否不經任何計算就看出 $C_4 = S_4$？

可以。代換（substitution）$\phi = \frac{1}{2}\pi - \theta$ 把 $\cos\theta$ 變成 $\sin\phi$，並把上下限 0 和 $\frac{1}{2}\pi$ 對調，所以對任何 n 都有 $\int_0^{\pi/2} \cos^n\theta \,\mathrm{d}\theta = \int_0^{\pi/2} \sin^n\phi \,\mathrm{d}\phi$。$\cos^4\theta$ 與 $\sin^4\theta$ 的圖像關於直線 $\theta = \frac{1}{4}\pi$ 互為鏡像，因此兩者所圍的面積相等：



所以第一個恆等式其實只是為了求和才用得上；差的積分等於 0，不過是這種對稱性（symmetry）的另一種面貌。

六次方

我們重複同一方法。設 $c = \cos\theta$、$s = \sin\theta$，並有 $c^2 + s^2 = 1$：

$$\begin{aligned} c^6 - s^6 &= (c^2 - s^2)(c^4 + c^2 s^2 + s^4) = \cos 2\theta\left(1 - \tfrac{1}{2}\sin^2 2\theta + \tfrac{1}{4}\sin^2 2\theta\right) \\ &= \cos 2\theta\left(1 - \tfrac{1}{4}\sin^2 2\theta\right), \end{aligned}$$

其中用到已證明的恆等式以及 $c^2 s^2 = \frac{1}{4}\sin^2 2\theta$。至於兩者之和，我們利用 $c^2 + s^2$ 的立方：

$$c^6 + s^6 = (c^2 + s^2)^3 - 3c^2 s^2 (c^2 + s^2) = 1 - \tfrac{3}{4}\sin^2 2\theta.$$

以 C_6 和 S_6 表示這兩個積分。由於 $\frac{\mathrm{d}}{\mathrm{d}\theta}\sin^3 2\theta = 6\sin^2 2\theta\cos 2\theta$，

$$C_6 - S_6 = \left[\tfrac{1}{2}\sin 2\theta - \tfrac{1}{24}\sin^3 2\theta\right]_0^{\pi/2} = 0,$$

這與對稱性論證所預料的結果一致。再次利用 $\int_0^{\pi/2} \sin^2 2\theta \,\mathrm{d}\theta = \frac{1}{4}\pi$，

$$C_6 + S_6 = \frac{\pi}{2} - \frac{3}{4}\cdot\frac{\pi}{4} = \frac{5\pi}{16}.$$

因此

$$\int_0^{\pi/2} \cos^6\theta \,\mathrm{d}\theta = \int_0^{\pi/2} \sin^6\theta \,\mathrm{d}\theta = \frac{5\pi}{32}.$$

驗算

- 冪次愈高，答案愈小：$n = 0$ 時為 $\frac{1}{2}\pi$，$n = 2$ 時為 $\frac{1}{4}\pi$，然後是 $\frac{3}{16}\pi$ 和 $\frac{5}{32}\pi$。這是必然的，因為在該區間上 $0 \le \cos\theta \le 1$。
- 數值上，用辛普森法則（Simpson's rule）可得 $\int_0^{\pi/2} \cos^4\theta \,\mathrm{d}\theta \approx 0.589049$ 及 $\int_0^{\pi/2} \cos^6\theta \,\mathrm{d}\theta \approx 0.490874$，與 $3\pi/16$ 和 $5\pi/32$ 吻合；正弦的積分在同樣的準確度下亦與這些數值一致。
- 把 $\theta = \frac{1}{4}\pi$ 代入兩個關於和的恆等式，得 $\frac{1}{4} + \frac{1}{4} = 1 - \frac{1}{2}$ 及 $\frac{1}{8} + \frac{1}{8} = 1 - \frac{3}{4}$，兩者皆成立。

常見錯誤

- 仿照 $\cos^2\theta + \sin^2\theta = 1$ 而寫出 $\cos^4\theta + \sin^4\theta = 1$：這樣就漏掉了交叉項 $2\sin^2\theta\cos^2\theta$。
- 對 $\cos 2\theta$ 或 $\cos 4\theta$ 積分時，忘記由鏈式法則（chain rule）而來的因子 $\frac{1}{2}$；在本題中，這對 $\cos 2\theta$ 沒有影響（反正積分是 0），但對 $\sin^2 2\theta$ 就有影響。

STEP 之外

這些都是沃利斯（Wallis）積分 $I_n = \int_0^{\pi/2} \sin^n\theta \,\mathrm{d}\theta$ 的特例，它們滿足遞推公式（reduction formula）$I_n = \frac{n-1}{n} I_{n-2}$（用分部積分法（integration by parts）即可得到）。你可以驗證 $I_4 = \frac{3}{4}\cdot\frac{1}{2}\cdot\frac{\pi}{2}$ 和 $I_6 = \frac{5}{6} I_4$ 正好給出上面的答案。比較 I_{2n} 與 I_{2n+1}，便可推導出沃利斯著名的 π 的無窮乘積（infinite product）。Siklos 的 *Advanced Problems in Mathematics* 有更多 STEP 風格的遞推公式練習，而 Bartle 與 Sherbert 的 *Introduction to Real Analysis* 則把當中的極限論證嚴格化。

詳細解答

STEP I 2007，第 4 題

純數學 • 二次方程的公共根

> **STEP I 2007，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/07-s1-q4

關鍵思路

第一部分是 $x^3+y^3+z^3-3xyz$ 的經典因式分解（factorisation）：一旦知道 $x+b+c$ 是因式，比較係數便可得出那個二次式。第二部分有兩個二次方程共用一個根。標準的做法是消去（eliminate）：把兩條方程組合起來，使 k 的其中一個冪消去。用兩種不同的方法這樣做，便可把 k 和 k^2 寫成比的形式，然後由 $k^2=(k)^2$ 得出一個只涉及係數的關係式。這個關係式原來就是第一部分取 $x=a$ 的情況。

解答

因式分解

把 $x=-(b+c)$ 代入 $P(x)=x^3-3xbc+b^3+c^3$：

$$\begin{aligned}P(-b-c)&=-(b+c)^3+3bc(b+c)+b^3+c^3\\&=-(b^3+3b^2c+3bc^2+c^3)+3b^2c+3bc^2+b^3+c^3=0,\end{aligned}$$

所以根據因式定理（factor theorem），$x+b+c$ 是因式。比較係數（P 的 x^2 項為零，常數項為 b^3+c^3）可得

$$x^3-3xbc+b^3+c^3=(x+b+c)\,\mathrm{Q}(x),\qquad \mathrm{Q}(x)=x^2+b^2+c^2-xb-bc-cx.$$

把右邊展開即可確認：x^2 項是 $x^2(b+c)-x^2(b+c)=0$，x 項合併後是 $-3bcx$，而常數項是 b^3+c^3。

至於平方和，x^2、b^2、c^2 各在 $2\mathrm{Q}(x)$ 中出現兩次，而每個交叉項則以 $-2xb$ 等形式出現，所以

$$2\mathrm{Q}(x)=(x-b)^2+(b-c)^2+(c-x)^2.$$

公共根

由於 k 是兩條方程的根，

$$ak^2 + bk + c = 0, \tag{1}$$
$$bk^2 + ck + a = 0. \tag{2}$$

要消去 k^2，取 $b \times (1) - a \times (2)$：

$$(b^2 - ac)k + bc - a^2 = 0, \qquad \text{所以} \qquad (ac - b^2)\,k = bc - a^2.$$

若改為消去 k 項，則取 $c \times (1) - b \times (2)$：

$$(ac - b^2)k^2 + c^2 - ab = 0, \qquad \text{所以} \qquad (ac - b^2)\,k^2 = ab - c^2.$$

這就是涉及 k^2 的類似式子。

現在利用 $k^2 = k \cdot k$。由於 $ac \neq b^2$，我們可以除以 $ac - b^2$，得 $k = \dfrac{bc - a^2}{ac - b^2}$，於是

$$(ac - b^2)\,\frac{(bc - a^2)^2}{(ac - b^2)^2} = ab - c^2 \quad \Longrightarrow \quad (ac - b^2)(ab - c^2) = (bc - a^2)^2,$$

其中最後一步是把全式乘以 $ac - b^2$。這就是所求的結果。

把兩邊展開，

$$a^2bc - ac^3 - ab^3 + b^2c^2 = b^2c^2 - 2a^2bc + a^4,$$
$$\text{所以} \quad 0 = a^4 + ab^3 + ac^3 - 3a^2bc = a\left(a^3 - 3abc + b^3 + c^3\right).$$

由於 $a \neq 0$，我們得到 $a^3 - 3abc + b^3 + c^3 = 0$，正如所求。

想一想 我們在哪裏見過 $a^3 - 3abc + b^3 + c^3$？

它就是第一部分取 $x = a$ 的情況。所以

$$(a + b + c)\,\mathrm{Q}(a) = 0, \qquad 2\mathrm{Q}(a) = (a - b)^2 + (b - c)^2 + (c - a)^2.$$

因此，要麼 $a + b + c = 0$，要麼 $\mathrm{Q}(a) = 0$。

情況 $\mathrm{Q}(a) = 0$。因為 a、b、c 是實數，三個平方都是非負的，而它們的和為零，只可能是每一個都為零。所以 $a = b = c$，兩條方程都是 $a(y^2 + y + 1) = 0$：它們完全相同。

情況 $a + b + c = 0$。把 $c = -a - b$ 代入 k 的公式：

$$bc - a^2 = -ab - b^2 - a^2, \qquad ac - b^2 = -a^2 - ab - b^2,$$

所以 $bc - a^2 = ac - b^2$，而由於這個值不為零，$k = 1$。（事實上 $y = 1$ 滿足兩條方程，因為這時兩條方程都變成 $a + b + c = 0$。）

所以要麼 $k=1$，要麼兩條方程完全相同，全題至此完成。

留意在方程相同的情況下，$a=b=c$ 會得出 $ac=b^2$，而這已被題目的假設排除。所以在題設條件下，公共根其實必定是 $k=1$；「相同」這個可能性，是把條件 $ac\neq b^2$ 拿走之後才會出現的，這時公共根是複數的單位立方根（cube roots of unity）。

驗算

取 $a=1$、$b=2$、$c=-3$，則 $a+b+c=0$。兩條方程是 $y^2+2y-3=(y-1)(y+3)=0$ 和 $2y^2-3y+1=(2y-1)(y-1)=0$，它們共用 $k=1$。這裏 $ac-b^2=-7$、$bc-a^2=-7$ 及 $ab-c^2=-7$，所以 $(ac-b^2)k=bc-a^2$、$(ac-b^2)k^2=ab-c^2$ 和 $(-7)(-7)=(-7)^2$ 全部成立，而且 $a^3-3abc+b^3+c^3=1+18+8-27=0$。我們亦以精確算術，用一千組隨機的三元數組檢驗過第一部分的兩個恆等式，以及展開式 $(ac-b^2)(ab-c^2)-(bc-a^2)^2=-a(a^3-3abc+b^3+c^3)$。

常見錯誤

- 消去常數項（$a\times(1)-c\times(2)$）會得出 $k\big((a^2-bc)k+ab-c^2\big)=0$，還需要 $k\neq 0$ 才用得上。$c\times(1)-b\times(2)$ 這個組合則可避免這個問題。
- 除以 $ac-b^2$ 時，沒有說明題目已保證它不為零。
- 由 $\mathrm{Q}(a)=0$ 推出 $a=b=c$ 時，沒有用到 a、b、c 是實數：平方和的論證對複數並不成立。

STEP 之外

兩個多項式有公共根的條件，一般由結式（resultant）給出，那是一個由它們的係數構成的行列式（determinant）；關係式 $(ac-b^2)(ab-c^2)=(bc-a^2)^2$ 正正就是令這兩個二次式的結式等於零。$a^3+b^3+c^3-3abc$ 這個式子也是一個循環矩陣（circulant matrix）的行列式，而且它可以完全分解為 $(a+b+c)(a+\omega b+\omega^2 c)(a+\omega^2 b+\omega c)$，其中 ω 是一個複數的單位立方根。想進一步了解多項式及其根，可參閱 Allenby 的 *Rings, Fields and Groups*（環、域與羣）。

詳細解答

STEP I 2007，第 5 題

純數學 • 多面體的幾何

STEP I 2007，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/07-s1-q5

關鍵思路

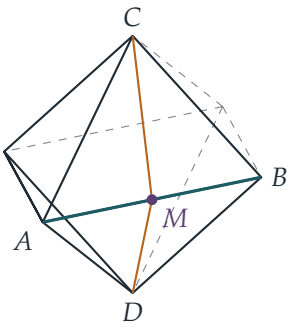
正八面體（regular octahedron）可以非常整齊地放進坐標系：它的六個頂點可取為各坐標軸正、負方向上距離原點一個單位的點。用了這個模型，兩部分都只是向量（vector）的運算。兩個面之間的夾角，是在垂直於它們公共棱的平面上量度的，所以我們取兩條直線：它們都由公共棱的中點出發，分別橫越兩個面，到達對面的頂點。至於各面的中心，只是三個頂點的平均。

解答

任何兩個正八面體都只相差一個縮放比例，而角度和體積之比在縮放之下都不會改變。因此我們不妨考慮頂點為

$$(\pm 1, 0, 0), \qquad (0, \pm 1, 0), \qquad (0, 0, \pm 1).$$

的正八面體。這些點之中，任何兩個不是相對的點都相距 $\sqrt{2}$，由此得出 12 條棱；而各個面就是每條坐標軸上各取一個頂點所成的 8 個三角形，每一種正負號的選擇對應一個面。每個面都是邊長為 $\sqrt{2}$ 的等邊三角形，所以這確實是一個正八面體。



第 (i) 部分

取兩個有公共棱的面；根據對稱性，每條棱的情況都一樣，所以只考慮其中一條便足夠。設 $A=(1,0,0)$、$B=(0,1,0)$、$C=(0,0,1)$ 及 $D=(0,0,-1)$，則面 ABC 與面 ABD 相交於 AB。設 $M=(\frac{1}{2},\frac{1}{2},0)$ 為 AB 的中點。

想一想 兩個平面相交於一條直線，它們之間的夾角應怎樣量度？

我們在公共直線上取一點，然後在每個平面上，作通過該點並垂直於公共直線的直線；這兩條直線之間的夾角，就是兩個面之間的夾角。由於每個面都是等邊三角形，三角形 ABC 的中線（median）MC 垂直於 AB，三角形 ABD 的中線 MD 亦然。所以我們要求的角就是 $\angle CMD$。

現在

$$\overrightarrow{MC}=\left(-\tfrac{1}{2},-\tfrac{1}{2},1\right),\qquad \overrightarrow{MD}=\left(-\tfrac{1}{2},-\tfrac{1}{2},-1\right),$$

而兩者都垂直於 $\overrightarrow{AB}=(-1,1,0)$，正如所料。兩個向量長度的平方都是 $\frac{1}{4}+\frac{1}{4}+1=\frac{3}{2}$，而

$$\overrightarrow{MC}\cdot\overrightarrow{MD}=\tfrac{1}{4}+\tfrac{1}{4}-1=-\tfrac{1}{2}.$$

因此

$$\cos\angle CMD=\frac{-\frac{1}{2}}{\frac{3}{2}}=-\frac{1}{3},$$

所以兩個面之間的夾角是 $\arccos\left(-\frac{1}{3}\right)$，約為 $109.5°$，正是所求。餘弦值是負數，所以這個角是鈍角：正八面體沿棱的地方不及正方體那麼尖銳。

若不用坐標，同一個計算就是在三角形 CMD 中運用餘弦定理（cosine rule）。設棱長為 a，則中線 $MC=MD=\frac{\sqrt{3}}{2}a$，而 C、D 是相對的頂點：A、C、$(-1,0,0)$、D 四點構成一個邊長為 a 的正方形，所以 $CD=\sqrt{2}\,a$。於是

$$2a^2=\tfrac{3}{4}a^2+\tfrac{3}{4}a^2-2\cdot\tfrac{3}{4}a^2\cos\angle CMD\quad\Longrightarrow\quad\cos\angle CMD=-\tfrac{1}{3}.$$

第 (ii) 部分

一個面的中心是它三個頂點的平均。每個面都有一個頂點是 $(\pm 1, 0, 0)$，一個是 $(0, \pm 1, 0)$，一個是 $(0, 0, \pm 1)$，所以它的中心是

$$\left(\pm\tfrac{1}{3}, \pm\tfrac{1}{3}, \pm\tfrac{1}{3}\right),$$

八種正負號的選擇給出八個中心。這些點正好是一個邊長為 $\frac{2}{3}$ 的正方體的頂點，所以

$$V_{\text{cube}} = \left(\tfrac{2}{3}\right)^3 = \tfrac{8}{27}.$$

正八面體是由兩個以正方形為底的棱錐（pyramid），沿着頂點為 $(\pm 1, 0, 0), (0, \pm 1, 0)$ 的正方形黏合而成。這個正方形的對角線長度為 2，所以它的面積是 $\frac{1}{2} \cdot 2 \cdot 2 = 2$，而每個棱錐的高是 1。因此

$$V_{\text{oct}} = 2 \times \tfrac{1}{3} \times 2 \times 1 = \tfrac{4}{3}.$$

兩者之比為

$$\frac{V_{\text{oct}}}{V_{\text{cube}}} = \frac{4/3}{8/27} = \frac{9}{2},$$

所以正八面體的體積與正方體的體積之比是 $9 : 2$。

驗算

- $\arccos(-\frac{1}{3}) \approx 109.47°$。對坐標模型的全部 12 條棱，以數值方式重複上述中點計算，每次都得出 $\cos = -\frac{1}{3}$。
- 正八面體就是 $|x| + |y| + |z| \le 1$。在 $[-1, 1]^3$ 內隨機取 10^6 個點，以蒙地卡羅（Monte Carlo）方法估算它的體積，結果是 1.330，而準確值是 $\frac{4}{3}$。
- 正方體位於正八面體之內（每個中心都滿足 $|x| + |y| + |z| = 1$），所以比值必定大於 1。而 $\frac{9}{2}$ 確實如此。

常見錯誤

- 取了向外法向量（outward normal）$(1, 1, 1)$ 與 $(1, 1, -1)$ 之間的夾角，得出 $\arccos \frac{1}{3}$。兩個面本身之間的夾角是它的補角。
- 沿着不垂直於公共棱的直線量度夾角，例如沿兩條棱 CA 和 DA。在這裏那個角是 $90°$，但它並不是兩個面之間的夾角。
- 以為正方體的邊長就是兩個相對的面的中心之間的距離。那個距離 $\frac{2}{\sqrt{3}}$ 其實是正方體的一條空間對角線。

STEP 之外

正方體與正八面體是互為對偶（dual）的多面體：把其中一個的各面中心連接起來，就得到另一個，面與頂點的角色互相對調。對偶性是第一個跡象，顯示這兩個立體的對稱羣（symmetry group）是同一個羣，其階為 48；你會在羣論的入門課程中遇到它。想了解柏拉圖立體（Platonic solids）的對稱羣，可參閱 Armstrong 的 *Groups and Symmetry*（一本關於羣與對稱的入門書）。

詳細解答

STEP I 2007，第 6 題

純數學 • 丟番圖方程

STEP I 2007，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/07-s1-q6

關鍵思路

兩部分都由冪的差出發，而 $x-y$ 是每一個這類式子的因式。約去這個因式（因為 $d \neq 0$，所以可以這樣做）之後，餘下的是一個關於 x 和 y 的對稱式（symmetric expression）。這個式子連同 $x-y=d$，就可以確定 x 和 y。要得到整數解，我們再選取適當的 d，使公式算出來的都是整數。在第 (ii) 部分，這就是要令一個平方根成為有理數（rational number）。

解答

第 (i) 部分

把左方因式分解（factorise），

$$(x-y)(x+y)=(x-y)^3.$$

由於 $x-y=d\neq 0$，我們可以除以 $x-y$，得 $x+y=(x-y)^2=d^2$。把這式與 $x-y=d$ 相加和相減，得

$$x=\frac{d^2+d}{2}=\frac{d(d+1)}{2}, \qquad y=\frac{d^2-d}{2}=\frac{d(d-1)}{2}.$$

要找整數對，設 $x=\sqrt{m}$ 及 $y=\sqrt{n}$。那麼 $x^2-y^2=m-n$，而所要求的條件正好是 $x^2-y^2=(x-y)^3$。取 $d=\sqrt{m}-\sqrt{n}$，由於 $m>n$，它不等於零，於是

$$m=\left(\frac{d(d+1)}{2}\right)^2, \qquad n=\left(\frac{d(d-1)}{2}\right)^2.$$

若 d 是正整數，則 $d(d\pm 1)$ 是兩個連續整數之積，所以是偶數。於是 $\sqrt{m}$ 和 $\sqrt{n}$ 都是整數，m 和 n 也是整數。此外，因為 $d>0$，所以 $m>n$。我們需要 $n>100$，即

$\frac{1}{2}d(d-1) > 10$。取 $d = 5$ 得 $\frac{1}{2}d(d-1) = 10$，並不足夠，所以我們取 $d = 6$：

$$\sqrt{m} = 21, \quad \sqrt{n} = 15, \qquad m = 441, \quad n = 225.$$

驗證：$m - n = 216 = 6^3 = (21-15)^3$。（其他答案同樣成立，例如 $d = 7$ 時 $m = 784$，$n = 441$。）

第 (ii) 部分

再次因式分解：$x^3 - y^3 = (x-y)(x^2+xy+y^2)$。除以 $x - y = d \neq 0$，得

$$x^2 + xy + y^2 = d^3.$$

我們還有 $(x-y)^2 = d^2$，即 $x^2 - 2xy + y^2 = d^2$。從上一條方程減去這一條，得 $3xy = d^3 - d^2$，正是所要證明的結果。

現在代入 $y = x - d$：

$$3x(x-d) = d^3 - d^2 \quad \Longrightarrow \quad 3x^2 - 3dx - (d^3 - d^2) = 0.$$

由二次公式（quadratic formula），

$$x = \frac{3d \pm \sqrt{9d^2 + 12(d^3 - d^2)}}{6} = \frac{3d \pm \sqrt{3d^2(4d-1)}}{6}.$$

現在 $\sqrt{3d^2(4d-1)} = |d|\sqrt{3(4d-1)}$，而 $\pm$ 已經涵蓋正負兩種情況，所以可以把 $|d|$ 換成 d。乘以 2，

$$2x = d \pm \frac{d\sqrt{3(4d-1)}}{3} = d \pm d\sqrt{\frac{4d-1}{3}},$$

這就是所要求的結果。

想一想 怎樣選取 d，才能使 x 和 y 都是整數？

若 $m = x$ 和 $n = y$ 是整數，則 $d = m - n$ 是非零整數。要使 x 是有理數，$\sqrt{(4d-1)/3}$ 必須是有理數。把它記作 k。那麼 $k^2 = (4d-1)/3$ 是整數，而整數的平方根只有在該整數是完全平方數（perfect square）時才是有理數，所以 k 是整數。因此我們需要

$$4d - 1 = 3k^2 \quad \text{對某個整數} k \geq 0.$$

左方是奇數，所以 k 是奇數。於是 $2x = d(1 \pm k)$ 是偶數，所以 x 是整數，$y = x - d$ 也是整數。

- $k = 1$ 給出 $d = 1$，於是 $x = 1$，$y = 0$ 或 $x = 0$，$y = -1$。兩組都不是由兩個正整數組成。
- $k = 3$ 給出 $d = 7$，於是 $2x = 7 \pm 21$。取正號得 $x = 14$，$y = 7$。

所以 $m = 14$，$n = 7$ 符合要求：

$$m^3 - n^3 = 2744 - 343 = 2401 = 7^4 = (m-n)^4.$$

次序是重要的：若取 $m = 7$，$n = 14$，左方是 -2401，右方卻是 $+2401$。

驗算

- 第 (i) 部分：$21^2 - 15^2 = 441 - 225 = 216$，而 $(21-15)^3 = 216$。
- 第 (ii) 部分：取 $d = 7$，$3xy = 3 \cdot 14 \cdot 7 = 294$，而 $d^3 - d^2 = 343 - 49 = 294$。
- 用電腦在 $1 \le m, n < 200$ 的範圍內搜尋，恰好找到 $(14,7)$、$(57,38)$ 和 $(148,111)$。它們來自 $k = 3, 5, 7$（即 $d = 7, 19, 37$），與公式吻合。搜尋結果亦確認，在第 (i) 部分滿足 $n > 100$ 的解之中，$(441,225)$ 是最小的。

常見錯誤

- 忘記約去因式時需要 $x - y \neq 0$。這正是題目假設 $d \neq 0$ 的原因。
- 在第 (i) 部分取 $d = 5$，得 $n = 100$，不符合嚴格不等式（strict inequality）$n > 100$。
- 不加說明就寫 $\sqrt{d^2} = d$。它其實等於 $|d|$，是因為有 $\pm$，我們才可以去掉絕對值符號。

STEP 之外

第 (ii) 部分把丟番圖方程（Diophantine equation）$m^3 - n^3 = (m-n)^4$ 完全解出：對奇數 $k \ge 3$ 取 $d = (3k^2+1)/4$，解就是 $m = \frac{1}{2}d(k+1)$ 和 $n = \frac{1}{2}d(k-1)$。要求一個平方根是有理數，是把這類方程化為簡單得多的條件 $4d - 1 = 3k^2$ 的標準方法。這類方程，以及它們背後二次型（quadratic form）的算術，可參閱 Davenport 的 *The Higher Arithmetic*（一本數論入門經典）。

詳細解答

STEP I 2007，第 7 題

純數學 • 直線之間的距離

> **STEP I 2007，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/07-s1-q7

關鍵思路

在兩條直線上各寫出一個一般點，相減，再把長度平方：這樣 D^2 就是兩個參數（parameter）的二次式（quadratic）。竅門在於把這個二次式寫成平方和（sum of squares），因為當每個平方都取到最小時，平方和也就最小。在第 (ii) 部分，同樣的方法仍然適用，而對於不同的 k 值，這些平方如何拆開，正好告訴我們背後的幾何。

解答

第 (i) 部分

L_1 上的一般點是 $P = (1+2\lambda,\ 2\lambda,\ 2-3\lambda)$，$L_2$ 上的一般點是 $Q = (4+\mu,\ -2+2\mu,\ 9-2\mu)$。所以

$$\overrightarrow{PQ} = \begin{pmatrix} 3+\mu-2\lambda \\ -2+2\mu-2\lambda \\ 7-2\mu+3\lambda \end{pmatrix}$$

而

$$D^2 = (3+\mu-2\lambda)^2 + (-2+2\mu-2\lambda)^2 + (7-2\mu+3\lambda)^2.$$

逐個係數合併同類項：

$$\mu^2:\ \ 1+4+4=9, \qquad \lambda^2:\ \ 4+4+9=17, \qquad \lambda\mu:\ \ -4-8-12=-24,$$
$$\mu:\ \ 6-8-28=-30, \qquad \lambda:\ \ -12+8+42=38, \qquad \text{常數項}:\ \ 9+4+49=62,$$

所以 $D^2 = 9\mu^2 + 17\lambda^2 - 24\lambda\mu - 30\mu + 38\lambda + 62$。

現在把題目給出的形式展開：

$$\begin{aligned}(3\mu-4\lambda-5)^2+(\lambda-1)^2+36 &= \left(9\mu^2+16\lambda^2+25-24\lambda\mu-30\mu+40\lambda\right)\\ &\quad+\left(\lambda^2-2\lambda+1\right)+36\\ &= 9\mu^2+17\lambda^2-24\lambda\mu-30\mu+38\lambda+62.\end{aligned}$$

兩者相同，所以

$$D^2=(3\mu-4\lambda-5)^2+(\lambda-1)^2+36,$$

正如所求。

想一想 為甚麼我們可以令兩個平方同時為零？

每個平方都不小於 0，所以無論 λ 和 μ 怎樣選，都有 $D^2\geq 36$。等號成立需要 $\lambda=1$ 及 $3\mu=4\lambda+5=9$，即 $\mu=3$。這兩個條件互不衝突（第一個條件確定 λ，然後第二個條件確定 μ），所以 $D^2=36$ 確實可以取到。因此兩條直線之間的最短距離是

$$D_{\min}=6,$$

並在以下兩點取到：

$$\lambda=1:\ P=(3,\,2,\,-1)\quad \text{位於}L_1,\qquad \mu=3:\ Q=(7,\,4,\,3)\quad \text{位於}L_2.$$

作為驗算，$\overrightarrow{PQ}=(4,2,4)$ 的長度是 $\sqrt{16+4+16}=6$，而且它與兩個方向向量（direction vector）都垂直（perpendicular）：$(4,2,4)\cdot(2,2,-3)=8+4-12=0$ 及 $(4,2,4)\cdot(1,2,-2)=4+4-8=0$。連接兩條直線的最短線段，正應如此。

第 (ii) 部分

L_3 上的一般點是 $P=(2,\,3+\alpha,\,5)$，L_4 上的一般點是 $Q=\left(3+4k\beta,\,3+(1-k)\beta,\,-2-3k\beta\right)$。所以

$$\overrightarrow{PQ}=\begin{pmatrix}1+4k\beta\\(1-k)\beta-\alpha\\-7-3k\beta\end{pmatrix}$$

而

$$\begin{aligned}D^2&=(1+4k\beta)^2+\left((1-k)\beta-\alpha\right)^2+(7+3k\beta)^2\\&=\left((1-k)\beta-\alpha\right)^2+25k^2\beta^2+50k\beta+50\\&=\left((1-k)\beta-\alpha\right)^2+25(k\beta+1)^2+25,\end{aligned}$$

這是因為 $(1+4k\beta)^2+(7+3k\beta)^2=1+8k\beta+16k^2\beta^2+49+42k\beta+9k^2\beta^2$。

無論 β 是甚麼，我們都可以選 $\alpha=(1-k)\beta$ 令第一個平方為零，因為 α 沒有在其他地方出現。所以問題變成 $25(k\beta+1)^2$ 可以小到甚麼程度，而兩種情況正是在這裏分開。

想一想 當 $k=0$ 時，有甚麼不同？

若 $k\neq 0$，我們可以取 $\beta=-1/k$，令 $k\beta+1=0$。若 $k=0$，無論 β 是甚麼，$25(k\beta+1)^2$ 這一項都固定在 25。

情況 $k\neq 0$。每個平方都是非負的，所以 $D^2\geq 25$，等號在 $\beta=-1/k$、$\alpha=(1-k)\beta=(k-1)/k$ 時成立。最短距離是

$$D_{\min}=5,$$

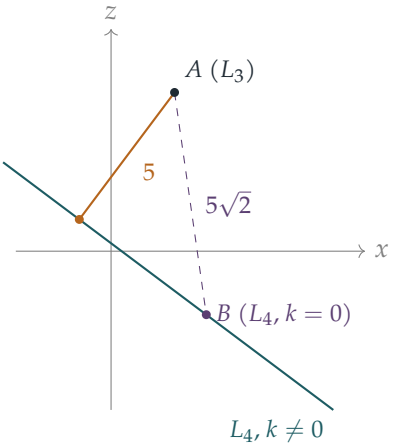
與 k 的（非零）值無關。最接近的兩點是 $P=\left(2,\,3+\frac{k-1}{k},\,5\right)$ 和 $Q=\left(-1,\,3+\frac{k-1}{k},\,1\right)$，其中 $\overrightarrow{PQ}=(-3,0,-4)$，長度為 5。由於 $D_{\min}>0$，兩條直線永不相交；而當 $k\neq 0$ 時，它們的方向 $(0,1,0)$ 與 $(4k,1-k,-3k)$ 並不平行（後者的第一和第三個分量不同時為零）。所以這兩條直線是*異面*（*skew*）直線，而 5 就是它們的公垂線（common perpendicular）的長度，這條公垂線是唯一的。

情況 $k=0$。這時 L_4 的方向是 $(0,1,0)$，與 L_3 相同：兩條直線*平行*（*parallel*）。此時 $D^2=(\beta-\alpha)^2+50$，只要 $\alpha=\beta$ 就取到最小值，所以

$$D_{\min}=\sqrt{50}=5\sqrt{2},$$

而且有無窮多對點可以取到這個值：其中一條直線上的任意一點，配上另一條直線上正對着它的那一點（y 坐標相同）。這其實就是兩條平行線之間的固定距離。

幾何意義。要看清這兩條直線，最容易的方法是沿着 y 軸直望過去，即投影（projection）到 xz 平面上。由於 L_3 與 y 軸平行，它投影成單獨一點 $A=(2,5)$，而與 L_3 垂直的距離在投影下保持不變。當 $k\neq 0$ 時，L_4 投影成通過 $(3,-2)$、方向為 $(4,-3)$ 的直線（$1-k$ 那個分量消失了），而 A 到這條直線的距離是 $\left|1\cdot(-3)-(-7)\cdot 4\right|/5=5$，與 k 無關。當 $k=0$ 時，L_4 投影成單獨一點 $B=(3,-2)$，距離是 $AB=\sqrt{1+49}=5\sqrt{2}$。



所以，當 $k \neq 0$ 時答案是 5（異面直線，只有一條最短線段，它與兩條直線都垂直）；當 $k = 0$ 時答案是 $5\sqrt{2}$（平行直線，距離處處相同）。

驗算

在第 (i) 部分，代入 $\lambda = 1$、$\mu = 3$ 所得的兩點相距 6，而且 $\overrightarrow{PQ}$ 與兩條直線都垂直。在 (λ, μ) 的網格上作數值搜尋，找不到低於 $D^2 = 36$ 的值。在第 (ii) 部分，對 $k = 0.001,\ 0.3,\ 1,\ -2,\ 5$ 以數值方法求 D^2 的最小值，每次都得到 25，而 $k = 0$ 時則得到 50。留意 $D_{\min}$ 在 $k = 0$ 處由 5 跳到 $5\sqrt{2}$。當 k 很小時，最接近的兩點位於很遠的地方，在 $\beta = -1/k$ 附近，所以極限不一定連續。

常見錯誤

- 忘記檢查兩個平方能否同時為零。這裏是可以的，但必須明確說明，不能想當然。
- 在 $k = 0$ 時除以 k：$\beta = -1/k$ 這個選擇只在 $k \neq 0$ 時可用，這正是要分兩種情況的原因。
- 把 $k \neq 0$ 時的兩條直線說成「相交」。它們不平行，但 $D_{\min} = 5 > 0$，所以它們是異面直線。

STEP 之外

多變數的配方法（completing the square）是二次型（quadratic forms）理論的起點：把 D^2 寫成平方和，就是把一個對稱矩陣（symmetric matrix）對角化（diagonalising），而求最小值則是一個最小平方法（least squares）問題。我們最後驗證的垂直關係，就是最小平方法中的「正規方程」（normal equations），而異面直線距離的一般公式 $|(\mathbf{b} - \mathbf{a}) \cdot (\mathbf{u} \times \mathbf{v})| / |\mathbf{u} \times \mathbf{v}|$ 也是來自同一套幾何。Strang 的 *Introduction to Linear Algebra*（線性代數入門）對投影和最小平方法有十分精彩的講解。

詳細解答

STEP I 2007，第 8 題

純數學 • 三次曲線的相切

> **STEP I 2007，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/07-s1-q8

關鍵思路

兩條曲線在某一點相切（touch），當且僅當求交點的方程在該處有重根（repeated root）。因此應考慮兩個三次式之差，並預期它含有因式 $(x-2)^2$。含 a 各項的係數 $1, -6, 12, -8$ 正是 $(x-2)^3$ 的二項式係數（binomial coefficients），看出這一點後，因式分解（factorisation）和繪畫略圖都會容易得多。

解答

在 $(2,8)$ 相切及另一個交點

把含 a 的各項歸在一起，

$$ax^3 - 6ax^2 + 12ax - 8a = a(x-2)^3, \qquad \text{所以} \qquad y = a(x-2)^3 + 12x - 16. \qquad (*)$$

它與 $y = x^3$ 的交點，就是下式的根：

$$a(x-2)^3 + 12x - 16 - x^3 = 0.$$

現在 $x^3 - 12x + 16$ 在 $x = 2$ 時為零，除去相應因式後得 $x^3 - 12x + 16 = (x-2)^2(x+4)$。因此

$$\begin{aligned} a(x-2)^3 + 12x - 16 - x^3 &= a(x-2)^3 - (x-2)^2(x+4) \\ &= (x-2)^2\big[(a-1)x - 2a - 4\big]. \end{aligned}$$

在 $x = 2$ 處，兩條曲線都通過 $(2,8)$，而因式 $(x-2)^2$ 表明 $x = 2$ 是交點方程的重根。等價地說，兩者的斜率（gradient）相同：由 $(*)$ 得 $y' = 3a(x-2)^2 + 12$，它在 $x = 2$ 時

等於 12，而 $y = x^3$ 在該處的斜率是 $3 \cdot 2^2 = 12$。所以對 a 的每一個值，兩條曲線都在 $(2,8)$ 相切。

其餘的交點都來自 $(a-1)x = 2a+4$。

- 若 $a \neq 1$，則 $x = \dfrac{2a+4}{a-1}$，而它永遠不會等於 2（否則需要 $2a+4 = 2a-2$）。所以恰好有另外一個交點，
$$\left(\frac{2a+4}{a-1}, \left(\frac{2a+4}{a-1}\right)^3\right).$$
- 若 $a = 1$，方括號內是常數 -6，所以沒有其他交點。

想一想 哪一條曲線在上方？在甚麼範圍內？

因式分解已順帶回答了這個問題。由於 $(x-2)^2 \geq 0$，$(*)$ 減去 x^3 的正負號就是線性因式 $(a-1)x - 2a - 4$ 的正負號。由此可知曲線 $(*)$ 在另一個交點的哪一側位於 $y = x^3$ 之上，而繪畫略圖所需的資料，大部分就在這裏。

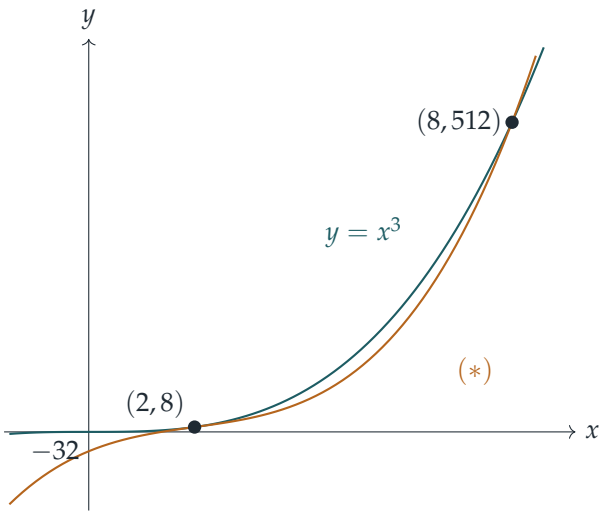
在全部三幅略圖中，$(*)$ 在 $(2,8)$ 都有一個拐點（point of inflection），該處斜率為 12，因為當 $a \neq 0$ 時 $y'' = 6a(x-2)$ 在該處變號。所以 $(*)$ 與 $y = x^3$ 在 $(2,8)$ 有共同的切線（tangent）$y = 12x - 16$；$y = x^3$ 在該處向上凹（concave up），而 $(*)$ 則不然。

第 (i) 部分：$a = 2$

這時 $y = 2(x-2)^3 + 12x - 16$，而

$$y - x^3 = (x-2)^2(x-8).$$

另一個交點是 $(8, 512)$。曲線 $(*)$ 在 $x < 8$ 時位於 $y = x^3$ 之下（在 $x = 2$ 處相切），在 $x > 8$ 時位於其上。它的斜率 $6(x-2)^2 + 12$ 恆為正數，所以它是遞增的，沒有轉向點（turning points）。它與 y 軸相交於 $(0, -32)$，並與 x 軸相交一次，位置在 $x \approx 1.37$ 附近。



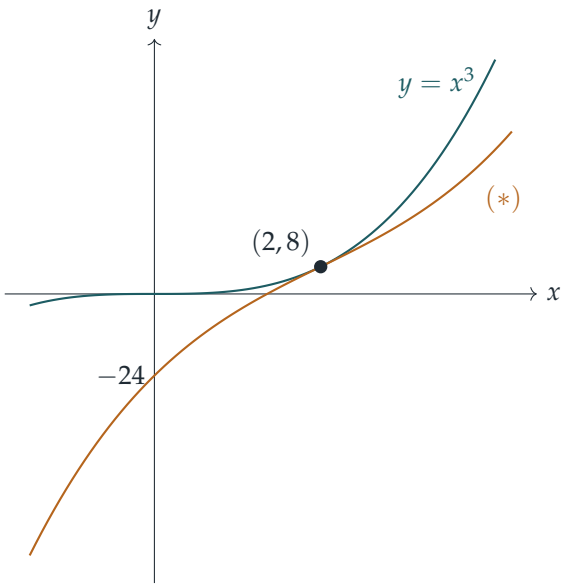
(縱向比例被大幅壓縮，所以在 $(2,8)$ 的相切看起來是平的；兩條曲線在該處有共同的切線。)

第 (ii) 部分：$a = 1$

這時 $y = (x-2)^3 + 12x - 16 = x^3 - 6x^2 + 24x - 24$，而

$$y - x^3 = -6(x-2)^2 \leq 0.$$

所以除了在 $(2,8)$ 相切之外，$(*)$ 處處位於 $y = x^3$ 之下；兩者沒有其他交點。無論向哪一個方向離開 $x = 2$，兩者的差距 $6(x-2)^2$ 都會愈來愈大。斜率 $3(x-2)^2 + 12$ 恆為正數，所以 $(*)$ 是遞增的。它與 y 軸相交於 $(0,-24)$，並與 x 軸相交一次，位置在 $x \approx 1.36$ 附近。



第 (iii) 部分：$a = -2$

現在另一個交點滿足 $x = \dfrac{-4+4}{-3} = 0$，所以它就是原點 $(0,0)$，而

$$y - x^3 = (x-2)^2(-3x) = -3x(x-2)^2.$$

所以 $(*)$ 在 $x < 0$ 時位於 $y = x^3$ 之上，在 $x > 0$ 時位於其下，並在 $(2,8)$ 相切。

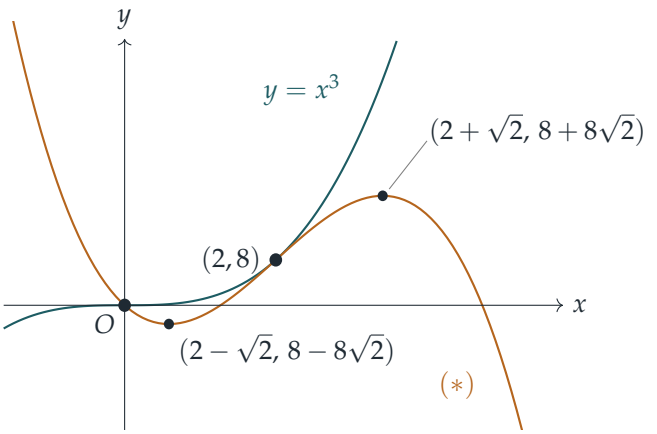
展開後，$y = -2(x-2)^3 + 12x - 16 = -2x^3 + 12x^2 - 12x = -2x(x^2 - 6x + 6)$，它與 x 軸相交於 $x = 0$ 和 $x = 3 \pm \sqrt{3}$（約為 1.27 和 4.73）。斜率是

$$y' = -6(x-2)^2 + 12,$$

它在 $(x-2)^2 = 2$ 時為零，即 $x = 2 \pm \sqrt{2}$。代入後，

$$x = 2 - \sqrt{2}: \quad y = 8 - 8\sqrt{2} \approx -3.31 \quad \text{(極小值)},$$
$$x = 2 + \sqrt{2}: \quad y = 8 + 8\sqrt{2} \approx 19.3 \quad \text{(極大值)}.$$

由於首項係數是負數，當 $x \to -\infty$ 時 $y \to +\infty$，當 $x \to +\infty$ 時 $y \to -\infty$。在原點，$(*)$ 的斜率是 -12，而 $y = x^3$ 的斜率是 0，所以兩條曲線在該處確實互相穿越。



驗算

- 在 $x = 2$ 處，對每一個 a，$(*)$ 都給出 $a \cdot 0 + 24 - 16 = 8$，且 $y' = 12$，與 $y = x^3$ 吻合。
- 取 $a = 2$ 和 $x = 8$，由簡潔的形式得 $2 \cdot 6^3 + 96 - 16 = 512 = 8^3$，確認了點 $(8, 512)$。
- 取 $a = 0$，$(*)$ 就是切線 $y = 12x - 16$，而公式給出另一個交點在 $x = -4$；事實上 $x^3 - 12x + 16 = (x-2)^2(x+4)$。
- 我們用數千組隨機有理數 a 和 x 精確地驗證了這個因式分解，並以數值方法核對了第 (iii) 部分的轉向值 $8 \pm 8\sqrt{2} \approx 19.31, -3.31$。

常見錯誤

- 遺漏 $a = 1$ 的情況：這時另一個交點的公式失效，因為根本沒有另一個交點。
- 只證明兩條曲線都通過 $(2, 8)$。相切還需要斜率相等，或等價地，需要有重根。
- 在第 (iii) 部分，把兩條曲線畫成在原點相切。它們在該處是互相穿越的，斜率並不相同。

STEP 之外

$f(x) - g(x)$ 的重根是相切在代數上的投影：兩條曲線在該點一階吻合。再推進一步，切觸階(order of contact)數算的是有多少階導數相同，而這正是泰勒多項式(Taylor polynomials）背後的想法：泰勒多項式就是在某一點與函數有最高可能切觸階的多項式。Bartle 與 Sherbert 的 *Introduction to Real Analysis*（實分析入門）嚴謹地建立了泰勒定理。

Siklos 的 *Advanced Problems in Mathematics*（高等數學難題集）有更多 STEP 風格的曲線描繪題目。

詳細解答

STEP I 2007，第 9 題

力學 • 斜坡上的極限摩擦力

STEP I 2007，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：

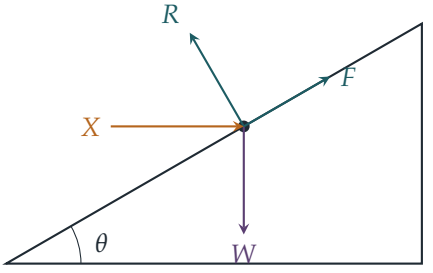
step.maths.org/questions/07-s1-q9

關鍵思路

這題是極限摩擦（limiting friction）的兩種情況，兩者的分別只在於摩擦力的方向。在每種情況下，我們沿斜面方向及垂直於斜面方向分解力，再消去法向反作用力（normal reaction），便可把每個水平力寫成 W 乘以一個三角函數式的比。將兩式相除可消去 W，剩下一條只涉及 k、μ 和 θ 的關係式。至於不等式，則來自 $\sin\theta\cos\theta = \frac{1}{2}\sin 2\theta$ 永遠不會超過 $\frac{1}{2}$ 這個事實。

解答

我們把物體模型化為質點(particle)，並假設兩種情況下摩擦力均達到極限，即 $F = \mu R$。為使代數運算簡潔，記 $s = \sin\theta$ 及 $c = \cos\theta$。



水平力把質點壓向斜面，所以它有一個沿斜面向上的分量 Xc，以及一個垂直壓向斜面的分量 Xs。重量則有沿斜面向下的分量 Ws 和垂直壓向斜面的分量 Wc。

即將下滑：力為 X

質點正處於即將向下滑動的狀態，所以摩擦力 μR 沿斜面向上（如圖所示）。先沿垂直於斜面的方向分解，再沿斜面方向分解，得

$$R = Wc + Xs, \qquad Xc + \mu R = Ws.$$

代入 R，

$$Xc + \mu Wc + \mu Xs = Ws \quad \Longrightarrow \quad X(c + \mu s) = W(s - \mu c). \tag{1}$$

即將上滑：力為 kX

現在質點正處於即將向上滑動的狀態，所以摩擦力沿斜面向下。同理，

$$R' = Wc + kXs, \qquad kXc = Ws + \mu R',$$

所以

$$kX(c - \mu s) = W(s + \mu c). \tag{2}$$

想一想 這裏眼前的每一項，我們都可以隨意拿來做除數嗎？

可以，這是由題目的物理情境保證的。題目說需要一個正的力 X 才能阻止質點下滑，即是說沒有這個力質點便會滑下，這表示 $\tan\theta > \mu$，即 $s - \mu c > 0$。於是由 (1) 得 $X > 0$，而 $c + \mu s > 0$ 本來就成立。水平推力若要能夠把質點推上斜面，(2) 便需要 $c - \mu s > 0$，即 $\tan\theta < 1/\mu$。兩者合起來必有 $\mu < \tan\theta < 1/\mu$，特別是 $\mu < 1$；這一點到最後會用得着。

關係式

將 (2) 除以 (1)，X 和 W 同時消去：

$$k\,\frac{c - \mu s}{c + \mu s} = \frac{s + \mu c}{s - \mu c} \quad \Longrightarrow \quad k(c - \mu s)(s - \mu c) = (s + \mu c)(c + \mu s).$$

把兩邊展開，並利用 $s^2 + c^2 = 1$：

$$(c - \mu s)(s - \mu c) = sc - \mu c^2 - \mu s^2 + \mu^2 sc = (1 + \mu^2)sc - \mu,$$
$$(s + \mu c)(c + \mu s) = sc + \mu s^2 + \mu c^2 + \mu^2 sc = (1 + \mu^2)sc + \mu.$$

所以 $k(1 + \mu^2)sc - k\mu = (1 + \mu^2)sc + \mu$，整理後得

$$(k - 1)(1 + \mu^2)\sin\theta\cos\theta = \mu(k + 1),$$

這正是所需的結果。

不等式

右邊是正數（因為 $\mu > 0$ 及 $k > 0$），而當 θ 為銳角時 $(1+\mu^2)\sin\theta\cos\theta$ 亦是正數；因此 $k > 1$。（這並不令人意外：把質點推上去所需的力，自然比僅僅把它頂住的大。）現在

$$\sin\theta\cos\theta = \tfrac{1}{2}\sin 2\theta \leq \tfrac{1}{2},$$

而由於 $(k-1)(1+\mu^2) > 0$，我們可以把這個不等式兩邊乘以它：

$$\mu(k+1) = (k-1)(1+\mu^2)\sin\theta\cos\theta \leq \tfrac{1}{2}(k-1)(1+\mu^2).$$

兩邊乘以 2，再把含 k 的項集合起來：

$$2\mu k + 2\mu \leq k(1+\mu^2) - (1+\mu^2) \quad\Longrightarrow\quad k(1-2\mu+\mu^2) \geq 1 + 2\mu + \mu^2,$$

即 $k(1-\mu)^2 \geq (1+\mu)^2$。上文已證明 $\mu < 1$，所以 $(1-\mu)^2 > 0$，我們可以除以它而不改變不等號的方向：

$$k \geq \frac{(1+\mu)^2}{(1-\mu)^2}.$$

等號成立當且僅當 $\sin 2\theta = 1$，即 $\theta = 45°$。

> **想一想** 為甚麼在最後一步之前，必須先知道 $\mu < 1$？

若 $\mu = 1$，我們便是除以零（而且上一行會變成 $0 \geq 4$，這是錯的）；若 $\mu > 1$，相除雖然仍然合法，但 θ 須滿足的條件早已不成立。從物理上看，當 $\mu \geq 1$ 時，$\tan\theta > \mu \geq 1$ 和 $\tan\theta < 1/\mu \leq 1$ 這兩個要求互相矛盾：要麼質點根本不需要外力頂住，要麼任何水平推力都無法令它開始向上滑。所以題目其實隱含了 $\mu < 1$。

驗算

- 取 $\mu = \frac{1}{2}$ 及 $\theta = 45°$，則 $s = c = 1/\sqrt{2}$。由 (1) 和 (2)，$X = W \cdot \frac{1/2}{3/2} = \frac{1}{3}W$ 及 $kX = W \cdot \frac{3/2}{1/2} = 3W$，所以 $k = 9 = (1.5/0.5)^2$：這正是等號成立的情況。若改取 $\theta = 60°$，同樣的公式給出 $k \approx 25.2 > 9$。
- 若 $\mu \to 0$，則 $X = kX = W\tan\theta$ 且 $k \to 1$，這與恆等式（兩邊皆為 0）及下界 1 都吻合。
- 以數值方法檢驗恆等式和不等式：隨機選取 10^5 組 $\mu \in (0,1)$ 和符合條件的 θ，沒有發現任何反例。

常見錯誤

- 取 $R = W\cos\theta$：水平力同樣會把質點壓向斜面，令反作用力增加 $X\sin\theta$。

- 兩種情況下都讓摩擦力指向同一方向。摩擦力總是與即將發生的運動方向相反。
- 除以 $(1-\mu)^2$ 而沒有說明它為何是正數（且不等於零）。

STEP 之外

要看清這個結果，有一個巧妙的方法，就是引入*摩擦角*（angle of friction）λ，其中 $\tan\lambda = \mu$：法向反作用力與極限摩擦力可合成為單一個力，它與法線成 λ 角，而 (1) 和 (2) 則變成 $X = W\tan(\theta-\lambda)$ 及 $kX = W\tan(\theta+\lambda)$。所以 $k = \tan(\theta+\lambda)/\tan(\theta-\lambda)$，而該下界所說的，就是這個比在 $\theta = 45°$ 時最小。這類運用摩擦角的論證，以及「摩擦錐」（friction cone）的概念，在大學的靜力學（statics）和接觸理論中會再次出現。Kleppner 與 Kolenkow 合著的 *An Introduction to Mechanics*（力學入門）對摩擦和受力分析的處理十分嚴謹，正合大學一年級課程的要求。

詳細解答

STEP I 2007，第 10 題

力學 • 位移–時間圖

> **STEP I 2007，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/07-s1-q10

關鍵思路

在位移—時間圖（displacement–time graph）上，每名騎兵和每支軍隊都是一條直線（或兩段直線），而「彼此相遇」不過是兩條線相交。因此我們以撒克遜軍隊為起點量度位移，寫出每條線的方程，再求它們的交點。題目中的不等式，正正就是各個交點按我們預期的次序出現的條件；而最後一部分所問的，是其中一個條件不成立時會發生甚麼事。

解答

建立圖像

我們把軍隊和騎兵看成沿一條直線運動的點。設 s 為由（靜止的）撒克遜軍隊起計、朝諾曼軍隊方向的位移，並在兩軍相距 d 時開始計時。於是：

- 撒克遜軍隊：$s = 0$。
- 諾曼軍隊：$s = d - ut$。
- 撒克遜騎兵，去程：$s = xt$。他到達諾曼軍隊時 $xt = d - ut$，即在 $t_S = \dfrac{d}{x+u}$，位置為 $s = \dfrac{xd}{x+u}$。他以速率 x 折返，

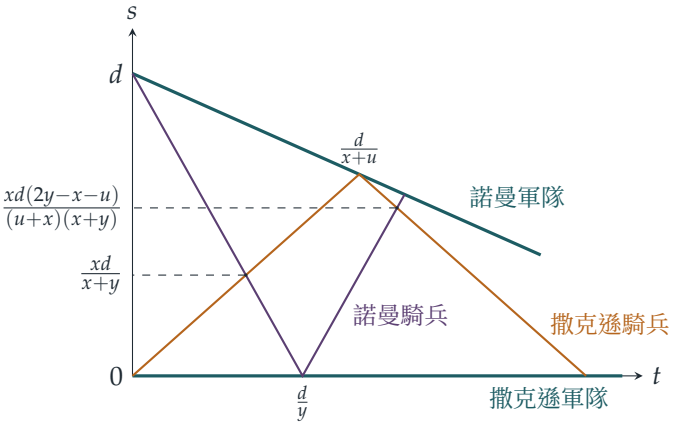
$$s = \frac{xd}{x+u} - x\left(t - \frac{d}{x+u}\right) = \frac{2xd}{x+u} - xt,$$

  並在 $t = \dfrac{2d}{x+u}$ 回到本陣。
- 諾曼騎兵，去程：$s = d - yt$。他在 $t_N = \dfrac{d}{y}$ 到達撒克遜軍隊，然後以速率 y 折返，

其直線為

$$s = y\left(t - \frac{d}{y}\right) = yt - d.$$

由於 $y > u$，他確實能追上自己的軍隊，此時 $yt - d = d - ut$，即在 $t = \dfrac{2d}{y+u}$。



（圖中取 $u = 1$、$x = 2$、$y = 4$、$d = 1$。）

第一次相遇

兩名騎兵都在去程，所以我們求 $s = xt$ 與 $s = d - yt$ 的交點：$(x+y)t = d$，得

$$t = \frac{d}{x+y}, \qquad s = \frac{xd}{x+y}.$$

這確實發生在兩人掉頭之前，因為 $\frac{d}{x+y}$ 比 $\frac{d}{y}$ 和 $\frac{d}{x+u}$ 都小（後者用到 $y > u$）。

第二次相遇

現在兩名騎兵都在回程：我們求 $s = \dfrac{2xd}{x+u} - xt$ 與 $s = yt - d$ 的交點。兩邊同時加上 $xt + d$，

$$(x+y)t = d + \frac{2xd}{x+u} = \frac{d(3x+u)}{x+u}, \qquad \text{所以} \qquad t_2 = \frac{d(3x+u)}{(x+u)(x+y)}.$$

於是

$$s = \frac{2xd}{x+u} - \frac{xd(3x+u)}{(x+u)(x+y)} = \frac{xd\big(2(x+y) - (3x+u)\big)}{(x+u)(x+y)} = \frac{xd(2y-x-u)}{(u+x)(x+y)},$$

這就是題目要求的第二個距離。

想一想 我們怎樣知道在時刻 t_2 兩名騎兵真的都在回程？

這正是那些不等式派上用場的地方。諾曼騎兵已經掉頭，若 $t_2 \geq \frac{d}{y}$：

$$y(3x+u) \geq (x+u)(x+y) \iff 3xy \geq x^2+xy+ux \iff u \leq 2y-x.$$

撒克遜騎兵已經掉頭，若 $t_2 \geq \frac{d}{x+u}$：

$$3x+u \geq x+y \iff u \geq y-2x.$$

所以條件 $y-2x < u < 2y-x$ 恰好保證 t_2 在兩個掉頭時刻之後。同一個條件 $u < 2y-x$ 亦使 $s>0$，即撒克遜騎兵尚未回到本陣。（還有一個題目沒有明言的要求：諾曼騎兵不可先回到自己的軍隊。把 $t_2 \leq \frac{2d}{y+u}$ 化簡，可得 $(x-u)(2x-y+u) \geq 0$；由於 $2x-y+u>0$，這表示 $x \geq u$，即撒克遜騎兵至少與諾曼軍隊一樣快。這肯定是題目的原意：否則他一掉頭，諾曼軍隊便會越過他。）

情況 (i)：$u > 2y-x$

此時 $t_2 < \frac{d}{y}$：撒克遜騎兵回程的直線與諾曼騎兵去程的直線先相交。由於 $y>u$，這個情況需要 $x>y$：撒克遜騎兵是較快的一方。他到達諾曼軍隊後掉頭，在兩人都朝撒克遜軍隊前進時從後追上諾曼騎兵。（求 $s=\frac{2xd}{x+u}-xt$ 與 $s=d-yt$ 的交點，可知這發生在距撒克遜軍隊 $\frac{xd(x+u-2y)}{(x+u)(x-y)}$ 處。）他在 $\frac{2d}{x+u}$ 回到本陣，早於諾曼騎兵到達的時刻 $\frac{d}{y}$，所以兩人不會第二次迎面相遇：諾曼騎兵只是到達撒克遜軍隊，而他的對手早已回來了。

情況 (ii)：$u < y-2x$

此時 $t_2 < \frac{d}{x+u}$：諾曼騎兵快得多。他到達撒克遜軍隊後掉頭，越過仍在朝諾曼軍隊前進的撒克遜騎兵。求 $s=yt-d$ 與 $s=xt$ 的交點，可知這發生在 $t=\frac{d}{y-x}$，距撒克遜軍隊 $\frac{xd}{y-x}$ 處。其後諾曼騎兵在 $\frac{2d}{y+u}$ 回到本陣，早於撒克遜騎兵到達諾曼軍隊的時刻 $\frac{d}{x+u}$（因為 $2(x+u)<y+u$），此後兩人在戰場上再無相遇。

驗算

- 取 $d=1$、$u=1$、$x=2$、$y=4$（如圖所示），公式給出相遇位置為 $\frac{1}{3}$ 和 $\frac{5}{9}$。以 10^{-5} 為時間步長用數值方法逐步模擬這個運動，得到的交點在 0.3333 和 0.5556。
- 若 $y=x+u$，兩個距離相等，都是 $\frac{x}{x+y}d$。這是一個很適合用來檢驗的特殊值。
- 當 $u \to 2y-x$ 時，第二個距離趨向 0：第二次相遇正好發生在撒克遜軍隊處，這就是與情況 (i) 的分界。

常見錯誤

- 把騎兵的回程寫成 $s = -xt$ 或 $s = yt$，忘記把時間原點移到掉頭的一刻。
- 求撒克遜騎兵的掉頭位置時，用了諾曼軍隊在 $t = 0$ 的位置，而不是他到達那一刻的位置。
- 在最後一部分只說「他們不會相遇」。在兩個情況中，他們其實都有第二次相遇，只是兩人同向而行：一人從後越過另一人。

STEP 之外

這類圖像稱為世界線（world lines），是狹義相對論（special relativity）的基本工具。在那裏，運動中的觀察者之間來回發送的光訊號，正是用這種方法分析的（雷達法），而直線的斜率則記錄了相對速度。這裏的代數運算，即把 u 加到一個速率上、又從另一個速率中減去，就是伽利略（Galileo）式的相對速度。Kleppner 與 Kolenkow 的 *An Introduction to Mechanics*（力學入門教科書）對伽利略圖景和相對論圖景都有論述。

詳細解答

STEP I 2007，第 11 題

力學 • 拋體運動

STEP I 2007，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：

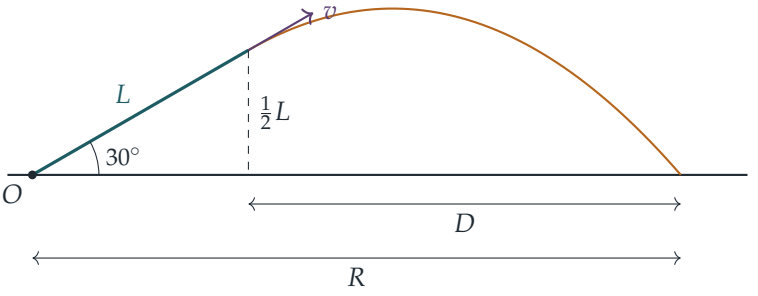
step.maths.org/questions/07-s1-q11

關鍵思路

把運動分成兩個階段。在光滑（smooth）管內，沿管方向唯一的力就是重力的分量，所以用能量（energy）（或等加速度公式（constant-acceleration formulae））便可求出離管時的速率。之後它就是一個普通的拋體（projectile），從高於終點水平 $\frac{1}{2}L$ 的一點以 30° 射出。消去飛行時間便得到關於 D 的二次方程；對它作隱函數微分，就可求得 $\mathrm{d}D/\mathrm{d}L$，完全不必解出 D。

解答

我們把粒子（particle）看成質點（point mass），把管看成光滑的（因此法向反作用力（normal reaction）不作功，沿管方向亦沒有分量），並忽略空氣阻力（air resistance）。



在管內

沿管方向分解力，唯一的分量是沿斜面向下的 $mg\sin 30^\circ = \frac{1}{2}mg$，所以減速度（deceleration）為 $\frac{1}{2}g$。經過距離 L 後，

$$v^2 = u^2 - 2\cdot\tfrac{1}{2}g\cdot L = u^2 - gL,$$

其中 v 是離管時的速率。（粒子要能離開管，必須有 $u^2 > gL$。）管的頂端比下端高 $L\sin 30^\circ = \frac{1}{2}L$，與下端的水平距離為 $L\cos 30^\circ = \frac{\sqrt{3}}{2}L$。

拋體階段

粒子沿管的方向離開，所以它的速度的水平分量為 $\frac{\sqrt{3}}{2}v$，垂直分量為向上的 $\frac{1}{2}v$。以離管點為原點量度，經過時間 t 後

$$x = \tfrac{\sqrt{3}}{2}vt, \qquad y = \tfrac{1}{2}vt - \tfrac{1}{2}gt^2.$$

當 $y = -\frac{1}{2}L$ 時，粒子回到下端的水平，此刻 $x = D$。由第一式得 $t = \dfrac{2D}{\sqrt{3}\,v}$，代入第二式，

$$\frac{D}{\sqrt{3}} - \frac{2gD^2}{3v^2} = -\frac{L}{2}.$$

全式乘以 $-6v^2$，

$$4gD^2 - 2\sqrt{3}\,v^2D - 3Lv^2 = 0,$$

再代入 $v^2 = u^2 - gL$，

$$4gD^2 - 2\sqrt{3}\,(u^2 - gL)D - 3L(u^2 - gL) = 0,$$

這就是所求的方程。（它的兩個根之中，D 是正的那一個：兩根之積為 $-3Lv^2/(4g) < 0$。）

導數

當 u 固定時，D 取決於 L，所以我們把方程對 L 求導，凡同時含有兩者的項都用積法則（product rule）：

$$8gD\frac{\mathrm{d}D}{\mathrm{d}L} - 2\sqrt{3}(u^2 - gL)\frac{\mathrm{d}D}{\mathrm{d}L} + 2\sqrt{3}\,gD - 3(u^2 - gL) + 3gL = 0.$$

最後兩項合併為 $-3(u^2 - 2gL)$。把含 $\mathrm{d}D/\mathrm{d}L$ 的項集中起來，

$$\big(8gD - 2\sqrt{3}(u^2 - gL)\big)\frac{\mathrm{d}D}{\mathrm{d}L} = -\big(2\sqrt{3}\,gD - 3(u^2 - 2gL)\big),$$

所以

$$\frac{\mathrm{d}D}{\mathrm{d}L} = -\frac{2\sqrt{3}\,gD - 3(u^2 - 2gL)}{8gD - 2\sqrt{3}(u^2 - gL)},$$

正如所求。

想一想 為甚麼 $-3L(u^2-gL)$ 這一項會產生兩部分？

因為 L 在其中出現了兩次：把 $-3Lu^2+3gL^2$ 求導得 $-3u^2+6gL$，即 $-3(u^2-gL)+3gL$。忘記括號內的 L，是漏掉答案中 $2gL$ 的最常見原因。

R 的駐值（stationary value）

粒子最終相對於下端的水平位移，等於管的水平跨度加上 D：

$$R=\tfrac{\sqrt3}{2}L+D,\qquad \text{所以}\qquad \frac{\mathrm{d}R}{\mathrm{d}L}=\frac{\sqrt3}{2}+\frac{\mathrm{d}D}{\mathrm{d}L}.$$

因此 $\mathrm{d}R/\mathrm{d}L=0$ 當且僅當 $\mathrm{d}D/\mathrm{d}L=-\frac{\sqrt3}{2}$，即（兩邊乘以分母，下文會驗證分母不為零）

$$2\sqrt3\,gD-3(u^2-2gL)=\tfrac{\sqrt3}{2}\big(8gD-2\sqrt3(u^2-gL)\big)=4\sqrt3\,gD-3(u^2-gL).$$

兩邊的 $3u^2$ 項互相抵消，餘下 $6gL-2\sqrt3\,gD=3gL$，所以 $2\sqrt3\,gD=3gL$，即

$$2D=L\sqrt3.$$

每一步都是可逆的，所以當 $2D=L\sqrt3$ 時 $\mathrm{d}R/\mathrm{d}L=0$，正如所求。

要求出 R，把 $D=\frac{\sqrt3}{2}L$ 代入二次方程：

$$4g\cdot\tfrac34L^2-2\sqrt3(u^2-gL)\cdot\tfrac{\sqrt3}{2}L-3L(u^2-gL)=3gL^2-6L(u^2-gL)=0.$$

由於 $L>0$，由此得 $gL=2(u^2-gL)$，所以 $L=\dfrac{2u^2}{3g}$。（這時 $v^2=u^2-gL=\frac13u^2>0$，所以粒子確實會離開管，而分母 $8gD-2\sqrt3v^2=\frac{8\sqrt3}{3}u^2-\frac{2\sqrt3}{3}u^2=2\sqrt3\,u^2$ 不為零。）最後

$$R=\tfrac{\sqrt3}{2}L+D=\sqrt3\,L=\frac{2\sqrt3\,u^2}{3g}=\frac{2u^2}{\sqrt3\,g}.$$

驗算

- 當 $L=0$ 時，二次方程給出 $D=\sqrt3u^2/(2g)$，正是以 30° 發射時大家熟悉的射程（range）$u^2\sin60^\circ/g$。
- 當 $L=u^2/g$ 時，粒子剛好到達頂端而 $v=0$，所以 $D=0$，同樣有 $R=\frac{\sqrt3}{2}u^2/g$。駐值 $2u^2/(\sqrt3g)\approx1.155\,u^2/g$ 出現在這兩個情況之間，而且比兩者都大，所以它是極大值。
- 數值上，取 $u=g=1$，讓 L 掃過 $[0,1]$，所得的最大 R 為 1.1547，出現於 $L\approx0.6667$，與 $2/\sqrt3$ 及 $L=\frac23$ 吻合；直接逐步模擬運動，亦能重現關於 D 的二次方程。

常見錯誤

- 把管內的減速度當作 g，而不是 $g\sin 30^\circ = \frac{1}{2}g$。
- 令 $y = 0$ 而不是 $y = -\frac{1}{2}L$：粒子的落點低於它的發射點。
- 寫成 $R = D$，忘記了管本身的水平長度 $\frac{\sqrt{3}}{2}L$。

STEP 之外

不解方程 $F(D, L) = 0$ 而直接對它求導，就是隱函數微分法（implicit differentiation）；到了大學一年級的分析課，它會發展成隱函數定理（implicit function theorem）：$\mathrm{d}D/\mathrm{d}L = -F_L/F_D$，凡 $F_D \neq 0$ 之處皆成立，而這正是我們驗證過的那個分母。力學方面，Kleppner 與 Kolenkow 的 *An Introduction to Mechanics*（力學入門）是很好的下一步。

詳細解答

STEP I 2007，第 12 題

概率與統計 • 條件概率

STEP I 2007，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/07-s1-q12

關鍵思路

抽第二粒糖果時，我們並不知道袋中的情況，因為那取決於之前發生了甚麼事。所以我們要考慮條件概率（conditional probability）：按第二次抽取之前發生的所有情況分類，求出每種情況下抽到紅色的機會，再用全概率定理（law of total probability）加起來。第 (ii) 部分有一條好用的捷徑：一個袋剛收到一粒隨機糖果後，袋中紅色糖果的期望數目（expected number）很容易寫出來，所以我們不必就轉移的那粒糖果的顏色重複分類。

解答

第 (i) 部分

設 R_1 為第一粒糖果是紅色的事件，R_2 為第二粒糖果是紅色的事件。每粒糖果被首先抽中的可能性均等，所以

$$\mathrm{P}(R_1) = \frac{a}{N}.$$

至於 R_2，按第一粒糖果的顏色分類。若第一粒是紅色，則餘下 $N-1$ 粒糖果，其中 $a-1$ 粒是紅色；若不是紅色，則餘下 $N-1$ 粒，其中 a 粒是紅色。（這裏 $N \geq 2$ 保證了第二粒糖果存在。）由全概率定理，

$$\begin{aligned}\mathrm{P}(R_2) &= \mathrm{P}(R_1)\,\mathrm{P}(R_2 \mid R_1) + \mathrm{P}(R_1')\,\mathrm{P}(R_2 \mid R_1') \\ &= \frac{a}{N}\cdot\frac{a-1}{N-1} + \frac{N-a}{N}\cdot\frac{a}{N-1} \\ &= \frac{a\big[(a-1)+(N-a)\big]}{N(N-1)} = \frac{a(N-1)}{N(N-1)} = \frac{a}{N}.\end{aligned}$$

所以 $\mathrm{P}(R_2) = \mathrm{P}(R_1)$，正是所求。（用對稱性（symmetry）也可得出同一結論：由兩粒不同糖果組成的每個有序對出現的可能性均等，所以第二粒糖果是某一粒指定糖果的機會，與第一粒是該糖果的機會相同。）

第 (ii) 部分

這裏的「第一粒糖果」是第一次擲硬幣後轉移的那一粒，「第二粒糖果」是第二次擲硬幣後選出的那一粒。設 R_1、R_2 分別為這兩粒糖果是紅色的事件。兩次擲硬幣互相獨立（independent），而每次擲硬幣的結果亦與選中哪粒糖果無關。

第一粒糖果。 按第一次擲硬幣的結果分類，

$$\mathrm{P}(R_1) = p \cdot \frac{a}{N} + q \cdot \frac{b}{N} = \frac{pa + qb}{N}.$$

第二粒糖果。 兩次擲硬幣共有四種情況。

正面、正面。先從第 1 袋取出一粒糖果，然後再從第 1 袋取出另一粒。這正是把第 (i) 部分的情況套用於第 1 袋（有 N 粒糖果，其中 a 粒紅色），所以第二粒糖果是紅色的概率為 a/N。

正面、反面。一粒糖果由第 1 袋移到第 2 袋，第 2 袋現在有 $N+1$ 粒糖果；然後我們從第 2 袋抽取。第 2 袋有 $b+1$ 粒紅色糖果的概率為 a/N，有 b 粒紅色糖果的概率為 $1 - a/N$，所以

$$\mathrm{P}(R_2 \mid \mathrm{HT}) = \frac{a}{N} \cdot \frac{b+1}{N+1} + \left(1 - \frac{a}{N}\right) \frac{b}{N+1} = \frac{b + \frac{a}{N}}{N+1} = \frac{Nb + a}{N(N+1)}.$$

? **想一想** 為甚麼第二行可以化簡得如此乾淨俐落？

分子就是轉移之後第 2 袋中紅色糖果的期望數目：原有的 b 粒，加上新來的糖果是紅色的機會 a/N。由於在這種情況下第 2 袋的大小固定為 $N+1$，抽到紅色的機會就是這個期望數目除以 $N+1$。

反面、正面。把兩個袋對調，用同樣的論證可知，第 1 袋現在有 $N+1$ 粒糖果，其中紅色糖果的期望數目為 $a + b/N$，所以

$$\mathrm{P}(R_2 \mid \mathrm{TH}) = \frac{Na + b}{N(N+1)}.$$

反面、反面。把第 (i) 部分套用於第 2 袋，得概率為 b/N。

現在把四種情況合併，它們的概率分別為 p^2、pq、qp、q^2：

$$\begin{aligned}\mathrm{P}(R_2) &= p^2\frac{a}{N} + pq\,\frac{(Nb+a)+(Na+b)}{N(N+1)} + q^2\frac{b}{N}\\ &= p^2\frac{a}{N} + pq\,\frac{(N+1)(a+b)}{N(N+1)} + q^2\frac{b}{N}\\ &= \frac{p^2a + pqa + pqb + q^2b}{N}\\ &= \frac{pa(p+q)+qb(p+q)}{N} = \frac{pa+qb}{N},\end{aligned}$$

最後一步用了 $p+q=1$。這等於 $\mathrm{P}(R_1)$，正是我們要證明的結果。

驗算

- 若 $p=1$，一切都在第 1 袋中發生，兩粒糖果的答案 a/N 正好就是第 (i) 部分。若 $p=0$，則得到第 2 袋的第 (i) 部分。
- 若 $a=b=0$，則沒有紅色糖果，兩個概率都是 0；若 $a=b=N$，則每粒糖果都是紅色，兩個概率都是 1。
- 取 $N=5$、$a=2$、$b=4$、$p=0.3$，公式給出 $(0.6+2.8)/5=0.68$。把整個過程模擬 200 000 次，第一粒糖果的結果為 0.679，第二粒為 0.681；而對所有較小的 N、a、b 逐一窮舉計算，結果亦吻合。

常見錯誤

- 忘記了轉移之後，收到糖果的袋有 $N+1$ 粒糖果而不是 N 粒，而送出糖果的袋則有 $N-1$ 粒。
- 在 HH 和 TT 兩種情況中，把第二次抽取當作與第一次獨立。其實兩者並不獨立，但第 (i) 部分告訴我們答案始終不變，所以應直接引用，而不必重新計算。

STEP 之外

這兩部分都是同一個深刻想法的例子：如果你從甕中抽取一件東西而不看抽到的是甚麼，你對甕的認識並沒有改變，所以下一次抽取的分佈（distribution）依然相同。嚴格來說，這一連串抽取是**可交換的**（exchangeable），而在第 (ii) 部分中，$p\times$(第 1 袋中紅色所佔比例) $+\,q\times$(第 2 袋中紅色所佔比例) 這個量，即下一粒抽出的糖果是紅色的機會，是一個**鞅**（martingale）：它未來的期望值等於它現在的值。鞅是現代概率論的核心，由賭博到金融都有它的身影。Grimmett 與 Welsh 合著的 *Probability:*

An Introduction（概率論入門教材）是認識條件期望（conditional expectation）和這些想法的好起點。

詳細解答

STEP I 2007，第 13 題

概率與統計 • 條件概率

STEP I 2007，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/07-s1-q13

本題在 *Beyond STEP* 第 12.1 節精確地取條件中有完整的引導式講解：https://beyondstep.keithmcnulty.org/zh-hant/ch12/conditioning.html

關鍵思路

雖然圓片是逐個取出的，但每一組四個圓片出現的可能性都相同，所以大部分工夫都花在數組合（combination）上。真正值得留意的是各個條件。知道第一個取出的是哪一個圓片，或者只知道某個指定的圓片確實被取出，餘下三個圓片都仍然是從剩下十個之中隨機選出的。只知道有某個有號碼的圓片被取出，則是性質完全不同的資訊，而最後兩部分正是為了顯示這個分別而設。

解答

設 N 為取出的四個圓片之中有號碼的圓片數目。由於袋子經過搖勻，每一個由四個不同圓片組成的有序序列出現的可能性都相同，因此每一個由四個圓片組成的無序集合出現的可能性也相同。這樣的集合共有

$$\binom{11}{4} = 330$$

個。要組成一個恰好有 k 個有號碼圓片的集合，須從 5 個有號碼的圓片中選 k 個，再從 6 個空白圓片中選 $4-k$ 個，所以

$$\mathrm{P}(N=k) = \frac{\binom{5}{k}\binom{6}{4-k}}{330}.$$

第 (i) 部分

四個全部有號碼，即 $k = 4$：

$$\mathrm{P}(N = 4) = \frac{\binom{5}{4}\binom{6}{0}}{330} = \frac{5}{330} = \frac{1}{66}.$$

換一個算法，按次序逐個抽取，$\frac{5}{11} \cdot \frac{4}{10} \cdot \frac{3}{9} \cdot \frac{2}{8} = \frac{1}{66}$。

第 (ii) 部分

設 T_1 為「3」號圓片最先被取出這個事件（event）。已知 T_1 發生，搖勻的袋子裏現在有 10 個圓片，其中 4 個有號碼、6 個空白，而接着取出的三個圓片是從這些圓片中隨機選出的。四個圓片全部有號碼，當且僅當這三個全部有號碼，所以

$$\mathrm{P}(N = 4 \mid T_1) = \frac{\binom{4}{3}}{\binom{10}{3}} = \frac{4}{120} = \frac{1}{30}.$$

第 (iii) 部分

仍然已知 T_1 發生，我們需要從餘下十個圓片抽出的三個之中，恰好再有一個有號碼的圓片：

$$\mathrm{P}(N = 2 \mid T_1) = \frac{\binom{4}{1}\binom{6}{2}}{\binom{10}{3}} = \frac{4 \times 15}{120} = \frac{1}{2}.$$

第 (iv) 部分

現在設 T 為「3」號圓片在某一次被取出這個事件。

> ? **想一想** 「3」號圓片是最先取出，還是在其他某一次取出，有沒有分別？

沒有分別。已知 T 發生，取出的四個圓片就是「3」號圓片連同另外三個，而根據對稱性（symmetry），從其餘十個圓片中選出三個的每一個集合，出現的可能性都相同。這正是第 (iii) 部分的情況，當中抽取的次序從來都無關重要。所以答案同樣是 $\frac{1}{2}$。

直接驗證如下：$\mathrm{P}(T) = \binom{10}{3}/\binom{11}{4} = \frac{120}{330} = \frac{4}{11}$，而包含「3」號圓片、另外一個有號碼圓片和兩個空白圓片的集合共有 $\binom{4}{1}\binom{6}{2} = 60$ 個，所以

$$\mathrm{P}(N = 2 \mid T) = \frac{60/330}{120/330} = \frac{1}{2}.$$

第 (v) 部分

設 F 為第一個圓片有號碼這個事件。按第一個圓片上的號碼把 F 分拆：F 是「j 號圓片最先被取出」這五個事件（$j = 1, \ldots, 5$）的不相交併集（disjoint union），每個事件的概率

都是 $\frac{1}{11}$。第 (iii) 部分的論證並沒有用到號碼是 3 這一點，所以對每個 j 都有

$$\mathrm{P}(N=2 \mid j \text{ 號圓片最先}) = \tfrac{1}{2}.$$

因此，根據全概率公式（law of total probability），

$$\mathrm{P}(N=2 \mid F) = \sum_{j=1}^{5} \frac{1/11}{5/11} \cdot \frac{1}{2} = \frac{1}{2}.$$

（直接的做法：已知 F 發生，其餘三個圓片是從 4 個有號碼和 6 個空白的圓片中隨機抽出的三個，這又回到第 (iii) 部分。）

第 (vi) 部分

設 A 為最少有一個有號碼的圓片被取出這個事件，即 A 就是事件 $N \geq 1$。這與之前的資訊並不屬於同一類：它沒有指明某一個圓片，再讓其餘的保持隨機。我們改為使用條件概率（conditional probability）的定義。由於 $N=2$ 蘊含 A，

$$\mathrm{P}(N=2 \mid A) = \frac{\mathrm{P}(N=2)}{\mathrm{P}(N \geq 1)}.$$

想一想 為甚麼不可以直接說「已知有一個有號碼的圓片，所以其餘三個是從剩下十個之中隨機抽出的」？

因為我們不知道已知的是哪一個圓片。一個有兩個有號碼圓片的集合，與一個有四個的集合，同樣滿足 A；並沒有哪一個圓片可以讓我們單獨抽出來放在一旁。所以這個條件是按 N 本身的值為各個集合重新加權，我們必須老老實實地數。

我們有

$$\mathrm{P}(N=2) = \frac{\binom{5}{2}\binom{6}{2}}{330} = \frac{10 \times 15}{330} = \frac{150}{330}, \qquad \mathrm{P}(N=0) = \frac{\binom{6}{4}}{330} = \frac{15}{330},$$

所以 $\mathrm{P}(N \geq 1) = \frac{315}{330}$，而

$$\mathrm{P}(N=2 \mid A) = \frac{150}{315} = \frac{10}{21}.$$

驗算

N 的分佈（distribution）加起來應該正確：當 $k = 0, \ldots, 4$ 時，$\binom{5}{k}\binom{6}{4-k}$ 的數目分別是 $15, 100, 150, 60, 5$，總和為 330。在第 (ii) 和第 (iii) 部分，其餘有號碼圓片數目的條件分佈，在 120 個之中的數目分別是 $20, 60, 36, 4$，總和同樣正確。我們亦以精確分數列出全部 $11 \times 10 \times 9 \times 8 = 7920$ 個有序抽取結果，核對了全部六個答案（$\frac{1}{66}$、$\frac{1}{30}$、$\frac{1}{2}$、$\frac{1}{2}$、$\frac{1}{2}$、$\frac{10}{21}$）。留意 $\frac{10}{21} \approx 0.476$ 略小於 $\frac{1}{2}$，這是合理的：以 $N \geq 1$ 為條件只剔除了 $N=0$ 的情況，所以與知道某個指定圓片被取出相比，它在 $N=1$ 上留下了較多的權重。

常見錯誤

- 把第 (iv) 和第 (vi) 部分當作同一類條件。「3 號圓片被取出」指明了一個圓片；「有一個有號碼的圓片被取出」則沒有，兩者的答案並不相同。
- 在第 (vi) 部分，忘記了條件是 $N \geq 1$，反而除以 $\mathrm{P}(N = 1)$ 或 $\frac{5}{11}$。
- 分子數的是有序抽取，分母數的卻是無序集合。兩者選定其一，然後貫徹始終。

STEP 之外

第 (v) 和第 (vi) 部分之間的對比，與著名的「男孩還是女孩」悖論（“boy or girl” paradox）是同一個陷阱：得知某個指定的孩子是男孩，與得知最少有一個是男孩，會得出不同的答案。在大學裏，這一點會有嚴格的表述：以一個事件為條件，就是把概率測度（probability measure）限制在該事件上再重新歸一化（renormalise），而這裏的數目正是超幾何分佈（hypergeometric distribution）。Grimmett 與 Welsh 的 *Probability: An Introduction*（概率入門），以及 Blitzstein 與 Hwang 的 *Introduction to Probability*（概率導論），都仔細討論了這些悖論。

詳細解答

STEP I 2007，第 14 題

概率與統計 • 泊松參數的唯一性

STEP I 2007，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/07-s1-q14

關鍵思路

每一項資料都給出一條形如 $h(\lambda) =$ 常數 的方程，其中 h 是平均值（mean）的某個函數。這條方程能否唯一地確定 λ，完全取決於 h 在 $\lambda > 0$ 上的形狀：嚴格單調（strictly monotonic）的 h 取每個值最多一次，而先升後降的 h 則會取大部分的值兩次，只有最大值才取一次。因此，整道題的重點就是找出每個函數在哪裏單調、在哪裏達到頂點。

解答

全題中，$\lambda > 0$（泊松分佈（Poisson distribution）的平均值必為正數），而且 $\mathrm{P}(X = k) = \mathrm{e}^{-\lambda}\lambda^k/k!$，其中 $k = 0, 1, 2, \ldots$

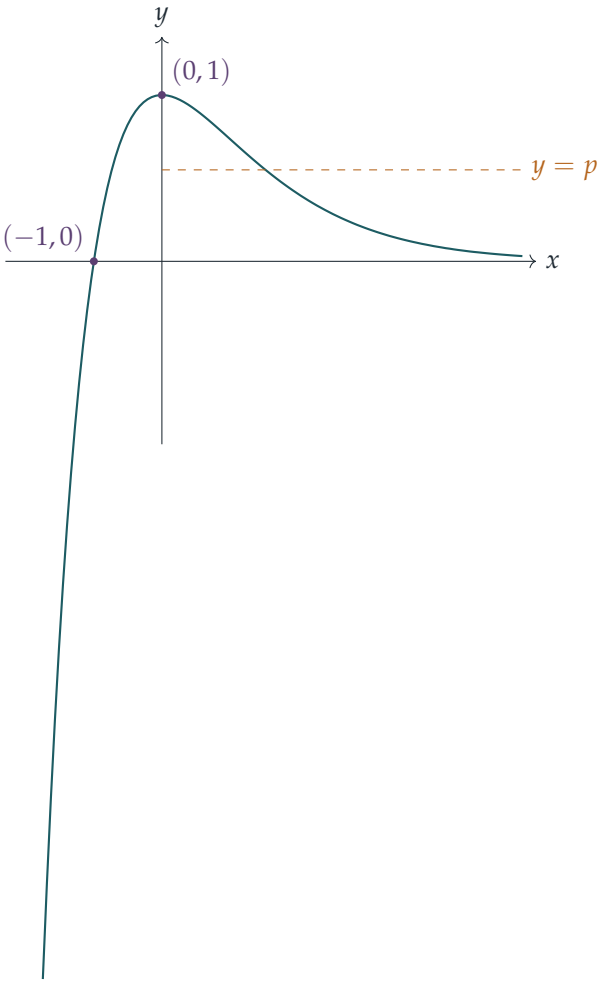
第 (i) 部分

設 $f(x) = (x+1)\mathrm{e}^{-x}$。由乘積法則（product rule），

$$f'(x) = \mathrm{e}^{-x} - (x+1)\mathrm{e}^{-x} = -x\,\mathrm{e}^{-x}.$$

由於 $\mathrm{e}^{-x} > 0$，$f'(x)$ 的正負號與 x 相反：f 在 $x < 0$ 時遞增，在 $x > 0$ 時遞減。所以唯一的轉向點（turning point）是位於 $(0, 1)$ 的極大點，這也正是曲線與 y 軸相交之處。曲線與 x 軸相交於 $x + 1 = 0$ 之處，即 $(-1, 0)$。

至於兩端的情況：當 $x \to \infty$ 時，e^{-x} 壓倒線性因子，$f(x) \to 0$，並且是從上方趨近（x 軸是一條漸近線（asymptote））；當 $x \to -\infty$ 時，兩個因子的大小都不斷增長，$f(x) \to -\infty$。



現在考慮概率。我們有

$$\mathrm{P}(X \geq 2) = 1 - \mathrm{P}(X = 0) - \mathrm{P}(X = 1) = 1 - \mathrm{e}^{-\lambda} - \lambda \mathrm{e}^{-\lambda} = 1 - f(\lambda),$$

所以題目所給的資料恰好等於說

$$f(\lambda) = (1 + \lambda)\mathrm{e}^{-\lambda} = p.$$

在 $\lambda > 0$ 上，函數 f 連續（continuous）而且嚴格遞減（因為 $f'(\lambda) = -\lambda \mathrm{e}^{-\lambda} < 0$），由 $f(0) = 1$ 開始下降，當 $\lambda \to \infty$ 時趨向 0。所以它取 $(0, 1)$ 內的每個值恰好一次：由中間值定理（intermediate value theorem）可知最少有一個解，而嚴格單調性則排除了第二個解。在略圖中，當 $0 < p < 1$ 時，水平線 $y = p$ 與右邊的分支恰好相交一次。因此 λ 被唯一確定。

第 (ii) 部分

這裏 $\mathrm{P}(X=1)=\lambda \mathrm{e}^{-\lambda}=q$。記 $g(\lambda)=\lambda \mathrm{e}^{-\lambda}$，則

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

所以 g 在 $(0,1]$ 上嚴格遞增，在 $[1,\infty)$ 上嚴格遞減，最大值為 $g(1)=\mathrm{e}^{-1}$。此外，當 $\lambda\to 0^+$ 及當 $\lambda\to\infty$ 時，都有 $g(\lambda)\to 0$。（事實上 $g(\lambda)=\mathrm{e}^{-1}f(\lambda-1)$，所以 g 的圖像就是第 (i) 部分的曲線向右平移一個單位，再沿垂直方向按 e^{-1} 的比例縮放。）

> ? **想一想** 一條水平線與一條先升至頂點、然後下降的曲線，可以相交多少次？

每一段單調的部分最多被擊中一次，所以：

- 若 $0<q<\mathrm{e}^{-1}$，則在 $(0,1)$ 內有一個解，在 $(1,\infty)$ 內有另一個解，所以 λ 有兩個可能的值；
- 若 $q>\mathrm{e}^{-1}$，則沒有解；
- 若 $q=\mathrm{e}^{-1}$，則唯一的解就是頂點本身，即 $\lambda=1$。

所以這項資料只對恰好一個 q 值確定唯一的 λ，即

$$q=\mathrm{e}^{-1},\qquad \lambda=1.$$

第 (iii) 部分

根據條件概率（conditional probability）的定義，並且由於事件 $\{X=1\}$ 包含於 $\{X\le 2\}$ 之內，

$$\mathrm{P}(X=1\mid X\le 2)=\frac{\mathrm{P}(X=1)}{\mathrm{P}(X\le 2)}=\frac{\lambda \mathrm{e}^{-\lambda}}{\mathrm{e}^{-\lambda}\left(1+\lambda+\frac{1}{2}\lambda^2\right)}=\frac{2\lambda}{2+2\lambda+\lambda^2}.$$

把它記作 $k(\lambda)$。將分子和分母同時除以 λ（由於 $\lambda>0$，這是容許的），

$$k(\lambda)=\frac{2}{\lambda+\dfrac{2}{\lambda}+2}.$$

> ? **想一想** 當 $\lambda>0$ 時，$\lambda+2/\lambda$ 的最小值是多少？

由算術平均—幾何平均不等式（AM–GM inequality），$\lambda+\dfrac{2}{\lambda}\ge 2\sqrt{\lambda\cdot\dfrac{2}{\lambda}}=2\sqrt{2}$，等號成立當且僅當 $\lambda=2/\lambda$，即 $\lambda=\sqrt{2}$。再者，$\lambda+2/\lambda$ 在 $(0,\sqrt{2}]$ 上嚴格遞減，在 $[\sqrt{2},\infty)$ 上嚴格遞增（它的導數 $1-2/\lambda^2$ 只在 $\sqrt{2}$ 處變號），而且在兩端都趨向 ∞。所以 k 與第 (ii) 部分的 g 一樣，具有先升後降的形狀，其最大值為

$$k(\sqrt{2})=\frac{2}{2\sqrt{2}+2}=\frac{1}{\sqrt{2}+1}=\sqrt{2}-1.$$

與之前完全一樣：當 $0 < r < \sqrt{2} - 1$ 時，λ 有兩個值；當 $r > \sqrt{2} - 1$ 時，沒有任何值；當 $r = \sqrt{2} - 1$ 時，恰好有一個值。所以 λ 只在以下情況被唯一確定：

$$r = \sqrt{2} - 1, \qquad \lambda = \sqrt{2}.$$

另一方法：在 $k(\lambda) = r$ 中消去分母，得到二次方程 $r\lambda^2 - 2(1-r)\lambda + 2r = 0$。它的兩根之積為 $2 > 0$，兩根之和為 $2(1-r)/r > 0$，所以若有實根，兩根必定都是正數；方程恰好有一個根，當且僅當判別式（discriminant）$4\left[(1-r)^2 - 2r^2\right] = 4(1 - 2r - r^2)$ 等於零，由此得 $r = \sqrt{2} - 1$（取正根），以及重根 $\lambda = (1-r)/r = \sqrt{2}$。

驗算

- 第 (ii) 部分：$\lambda = 1$ 給出 $\mathrm{P}(X = 1) = \mathrm{e}^{-1} \approx 0.3679$；以數值方法在 $0 < \lambda < 40$ 的範圍內掃描 $\lambda \mathrm{e}^{-\lambda}$，其值從不超過這個數，而 $q = 0.3$ 則給出兩個交點。
- 第 (iii) 部分：取 $\lambda = \sqrt{2}$，則 $\dfrac{2\sqrt{2}}{2 + 2\sqrt{2} + 2} = \dfrac{\sqrt{2}}{2 + \sqrt{2}} = \sqrt{2} - 1 \approx 0.4142$，數值掃描亦確認這就是最大值；$r = 0.35$ 給出 λ 的兩個值，$r = 0.45$ 則沒有任何值。
- 合理性檢查：當 $\lambda \to 0$ 時，幾乎全部概率質量都集中在 0；當 $\lambda \to \infty$ 時，在已知 $X \le 2$ 的條件下，概率質量堆積在 2。所以在這兩個極限情況下都有 $\mathrm{P}(X = 1 \mid X \le 2) \to 0$，與公式所示一致。

常見錯誤

- 忘記 $\lambda > 0$：在第 (i) 部分，曲線在 $x < 0$ 時同樣存在，但只有 $x > 0$ 的分支才與概率有關。
- 在第 (i) 部分聲稱解是唯一的，卻沒有說明原因：你既要說明解存在（連續性，以及函數值由 1 一直下降至 0），也要說明沒有第二個解（嚴格單調性）。
- 在第 (iii) 部分使用 $\mathrm{P}(X = 1 \text{ 及 } X \le 2) = \mathrm{P}(X = 1)\,\mathrm{P}(X \le 2)$；這兩個事件並非獨立（independent），而是 $\{X = 1\}$ 包含於 $\{X \le 2\}$ 之內。

STEP 之外

這是初嘗可識別性（identifiability）的滋味：在統計學中，我們會問一個觀察得到的量能否確定模型的參數，而答案是肯定的，當且僅當由參數到可觀察量的映射是一對一（one-to-one）的。在這裏，$\mathrm{P}(X \ge 2)$ 能識別 λ，但 $\mathrm{P}(X = 1)$ 和 $\mathrm{P}(X = 1 \mid X \le 2)$ 一般都不能。當你要求似然（likelihood）的最大值時，同樣的先升後降形狀會再次出現：$\lambda \mathrm{e}^{-\lambda}$

就是在看到單一觀察值 $X = 1$ 之後 λ 的似然，而它在 $\lambda = 1$ 處的頂點就是最大似然估計（maximum likelihood estimate）。Blitzstein 與 Hwang 合著的 *Introduction to Probability*（概率論入門）寫得生動活潑，是進一步了解泊松分佈和條件概率的好去處。

詳細解答

STEP II 2007，第 1 題

純數學 • 二項式近似

> **STEP II 2007，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/07-s2-q1

關鍵思路

二項式級數（binomial series）只有在那個「小量」真的很小時才會快速收斂（converge），所以整道題的關鍵，就是把根式（surd）改寫成 $\sqrt{1+\text{某個極小的數}}$ 的簡單有理數倍。為此，我們要找一個完全平方數（或完全立方數），使它乘以 3 或 6 之後剛好落在 10 的某個冪旁邊。這樣 k 相對於 100（或 1000）就很小，只需寥寥數項便能達到所需的準確度。

解答

第 (i) 部分

設 $x = \frac{k}{100}$，$(1+x)^{1/2}$ 的二項式級數的係數為

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

代入 $x = k/100$，得

$$\left(1+\frac{k}{100}\right)^{1/2} \approx 1 + \frac{k}{200} - \frac{k^2}{80\,000} + \frac{k^3}{16\,000\,000}.$$

(a) 取 $k = 8$，左方就是 $\sqrt{1.08}$，而關鍵在於 $108 = 3 \times 36$：

$$\sqrt{1.08} = \frac{\sqrt{108}}{10} = \frac{6\sqrt{3}}{10} = \frac{3\sqrt{3}}{5}, \qquad \text{所以} \qquad \sqrt{3} = \frac{5}{3}\sqrt{1.08}.$$

由級數可得

$$\sqrt{1.08} \approx 1 + 0.04 - 0.0008 + 0.000032 = 1.039232,$$

因此

$$\sqrt{3} \approx \frac{5}{3} \times 1.039232 = 1.7320533\ldots \approx 1.73205\text{(準確至小數點後五位)}。$$

(b)

? **想一想** k 取甚麼值，才會令 $1 + \frac{k}{100}$ 等於某個簡單平方數的 6 倍？

我們希望 $100 + k = 6m^2$，其中 m 為整數，而且 $6m^2$ 接近 100，這樣 k 才會小。由於 $6 \times 16 = 96$，取 $m = 4$ 及 $k = -4$：

$$\sqrt{0.96} = \frac{\sqrt{96}}{10} = \frac{4\sqrt{6}}{10} = \frac{2\sqrt{6}}{5}, \qquad \text{所以} \qquad \sqrt{6} = \frac{5}{2}\sqrt{0.96}.$$

當 $k = -4$ 時，首項之後的每一項都是負數：

$$\sqrt{0.96} \approx 1 - 0.02 - 0.0002 - 0.000004 = 0.979796,$$

於是

$$\sqrt{6} \approx \frac{5}{2} \times 0.979796 = 2.44949\text{(準確至小數點後五位)}。$$

這個 k 的大小只有 (a) 中的一半，所以這裏被略去的各項更加微小。(例如由 $1.5 = 6/4$ 得出的 $k = 50$，同樣可以準確地表示 $\sqrt{6}$，但當 $x = \frac{1}{2}$ 時，級數收斂得太慢，遠遠不足以準確至小數點後五位。)

第 (ii) 部分

取首兩項，得

$$\left(1 + \frac{k}{1000}\right)^{1/3} \approx 1 + \frac{k}{3000}.$$

? **想一想** 哪一個立方數乘以 3 之後，剛好略大於 1000？

我們需要 $1000 + k = 3m^3$，而且 k 要小。在 $1000/3 \approx 333$ 附近的立方數有 $6^3 = 216$ 和 $7^3 = 343$，而 $3 \times 343 = 1029$ 非常接近 1000。所以取 $k = 29$：

$$\sqrt[3]{1.029} = \frac{\sqrt[3]{3 \times 7^3}}{10} = \frac{7\sqrt[3]{3}}{10}, \qquad \text{所以} \qquad \sqrt[3]{3} = \frac{10}{7}\sqrt[3]{1.029}.$$

利用兩項的近似式，

$$\sqrt[3]{3} \approx \frac{10}{7}\left(1 + \frac{29}{3000}\right) = \frac{10}{7} \cdot \frac{3029}{3000} = \frac{3029}{2100},$$

這就是所求的近似值。

驗算

用計算機可得 $\sqrt{3} = 1.7320508\ldots$ 及 $\sqrt{6} = 2.4494897\ldots$，兩者都與上面的答案在小數點後五位之內吻合。至於第 (ii) 部分，$3029/2100 = 1.44238\ldots$，而 $\sqrt[3]{3} = 1.44225\ldots$。誤差約為 1.3×10^{-4}，與第一個被略去的項的大小 $\frac{10}{7} \cdot \frac{1}{9}\left(\frac{29}{1000}\right)^2 \approx 1.3 \times 10^{-4}$ 相符，所以這個近似值略為高估了真實值。

常見錯誤

- 算錯第三個係數：$\frac{1}{2} \cdot (-\frac{1}{2}) \cdot (-\frac{3}{2})/6 = +\frac{1}{16}$，而不是 $-\frac{1}{16}$。
- 計算級數之後，忘記乘回 $\frac{5}{3}$、$\frac{5}{2}$ 或 $\frac{10}{7}$。
- 選了一個很大的 k（例如 (b) 中的 $k = 50$），這時只取三四項根本遠遠不夠。

STEP 之外

究竟需要多少項，歸根究柢是要為泰勒級數（Taylor series）的餘項（remainder）定出上界。有了拉格朗日（Lagrange）餘項公式，你就可以證明（而不只是相信）這些近似值確實準確至小數點後五位。把問題重新縮放到級數快速收斂的範圍內，這個技巧正是計算機和電腦程式庫實際計算方根和對數的方法。關於泰勒級數及其餘項的理論，可參閱 Abbott 的 *Understanding Analysis*（一本實分析入門教材）。

詳細解答

STEP II 2007，第 2 題

純數學 • 三次曲線與面積

STEP II 2007，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/07-s2-q2

關鍵思路

三次曲線（cubic）轉向點（turning point）的一切資料，都藏在它的導數（derivative）裏。由於轉向點位於 $x = p$ 和 $x = q$，導數必定是 $(x-p)(x-q)$ 的倍數，比較係數（comparing coefficients）便立即得出 b 和 c。之後，求 $m-n$ 最簡潔的方法，是把導數在兩個轉向點之間積分（integrate）；而只要留意到 $y-n$ 在 $x = q$ 處有一個二重根（double root），面積也就容易計算了。

解答

第 (i) 部分

求導，得

$$\frac{\mathrm{d}y}{\mathrm{d}x} = 6x^2 - 2bx + c.$$

這個二次式在轉向點 $x = p$ 和 $x = q$ 處等於零，而它的首項係數是 6，所以

$$6x^2 - 2bx + c = 6(x-p)(x-q) = 6x^2 - 6(p+q)x + 6pq.$$

比較係數，得

$$b = 3(p+q), \qquad c = 6pq.$$

以下兩項事實稍後會用到。第一，由於 x^3 的係數為正，$\mathrm{d}y/\mathrm{d}x = 6(x-p)(x-q)$ 在兩根之外為正，在兩根之間為負，所以曲線先升、後降、再升：極大點（maximum）在先，故 $p < q$。

第二，條件 $n > 0$ 對 q 有所限制。代入 $x = q$，

$$n = 2q^3 - 3(p+q)q^2 + 6pq^2 = q^2(3p - q),$$

所以 $n > 0$ 即 $q < 3p$。綜合起來，$0 < p < q < 3p$。

第 (ii) 部分

曲線通過原點，而且

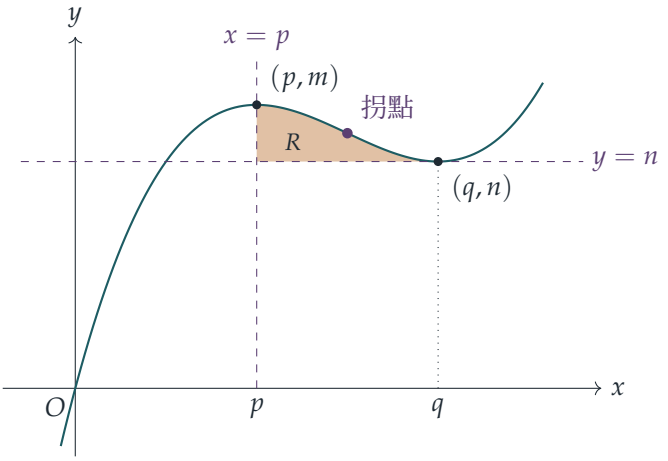
$$y = x\big(2x^2 - bx + c\big),$$

其中二次因式的判別式（discriminant）為

$$b^2 - 8c = 9(p+q)^2 - 48pq = 3(3p-q)(p-3q).$$

由於 $q < 3p$ 及 $p < q < 3q$，第一個括號為正，第二個括號為負，所以判別式為負，原點是曲線與 x 軸唯一的交點。因此曲線由 $-\infty$ 向上而來，在 O 穿過橫軸，升至極大點 (p, m)，再降至極小點（minimum）(q, n)（此時仍在橫軸上方），然後一直上升。

拐點（point of inflection）位於 $\mathrm{d}^2y/\mathrm{d}x^2 = 12x - 2b = 0$ 之處，即 $x = \frac{b}{6} = \frac{1}{2}(p+q)$，正好在兩個轉向點的正中間。



區域 R（陰影部分）位於 $x = p$ 與 $x = q$ 之間，在曲線之下、直線 $y = n$ 之上；曲線在極小點處與 $y = n$ 相切。

對稱性。設 $s = \frac{1}{2}(p+q)$，$h = q - p$，並令 $x = s + t$，即從拐點起量度。於是 $x - p = t + \frac{1}{2}h$ 及 $x - q = t - \frac{1}{2}h$，所以

$$\frac{\mathrm{d}y}{\mathrm{d}t} = 6\big(t^2 - \tfrac{1}{4}h^2\big),$$

這是 t 的偶函數（even function）。因此它由 0 起的積分，即 $y - y(s) = 2t^3 - \frac{3}{2}h^2t$，是 t 的奇函數（odd function）。所以 $y(s+t) - y(s) = -\big(y(s-t) - y(s)\big)$：曲線具有繞其拐點的 2 階旋轉對稱（*rotational symmetry*）（繞該點旋轉半周，曲線與自身重合）。特別地，這個半周旋轉把兩個轉向點互換，所以拐點是 (p, m) 與 (q, n) 的中點，即 $\big(\frac{1}{2}(p+q), \frac{1}{2}(m+n)\big)$。

第 (iii) 部分

由極大點到極小點的落差，就是斜率的積分：

$$m - n = y(p) - y(q) = -\int_p^q 6(x-p)(x-q)\,\mathrm{d}x.$$

> **想一想** 有沒有不必展開就能快速計算 $\int_p^q (x-p)(x-q)\,\mathrm{d}x$ 的方法？

代入 $x = p + u$，於是 $x - q = u - h$，其中 $h = q - p$，而 u 由 0 變到 h：

$$\int_0^h u(u-h)\,\mathrm{d}u = \frac{h^3}{3} - \frac{h^3}{2} = -\frac{h^3}{6}.$$

因此

$$m - n = -6\Big(-\frac{(q-p)^3}{6}\Big) = (q-p)^3,$$

正如所求。

第 (iv) 部分

在 $p \le x \le q$ 上，曲線由 m 遞減至 n，所以它位於 $y = n$ 之上或與之相接，而

$$\text{區域面積} R = \int_p^q (y-n)\,\mathrm{d}x.$$

三次式 $y - n = 2x^3 - bx^2 + cx - n$ 在 $x = q$ 處等於零，其導數在該處亦等於零，所以 $x = q$ 是二重根，並且

$$y - n = 2(x-q)^2(x-r)$$

對某個 r 成立。比較 x^2 的係數，$-2(2q + r) = -b = -3(p+q)$，所以 $r = \frac{1}{2}(3p - q)$。同樣設 $h = q - p$，並令 $u = x - q$（由 $-h$ 變到 0），

$$x - r = u + q - \tfrac{1}{2}(3p - q) = u + \tfrac{3}{2}h.$$

因此

$$\begin{aligned}\text{區域面積} R &= \int_{-h}^0 2u^2\big(u + \tfrac{3}{2}h\big)\,\mathrm{d}u = \Big[\tfrac{1}{2}u^4 + hu^3\Big]_{-h}^0 \\ &= -\big(\tfrac{1}{2}h^4 - h^4\big) = \tfrac{1}{2}h^4 = \tfrac{1}{2}(q-p)^4,\end{aligned}$$

這就是所求的結果。

驗算

取 $p = 1$，$q = 2$（符合條件，因為 $q < 3p$）。於是 $b = 9$，$c = 12$，而 $y = 2x^3 - 9x^2 + 12x$ 滿足 $y' = 6(x-1)(x-2)$，且 $m = y(1) = 5$ 及 $n = y(2) = 4 > 0$。所以 $m - n = 1 =$

$(q-p)^3$。拐點是 (1.5, 4.5)，即兩個轉向點的中點；把 $y-4$ 在 $[1,2]$ 上作數值積分，得 $0.5000 = \frac{1}{2}(q-p)^4$。我們亦以 2000 對滿足 $0 < p < q < 3p$ 的隨機數值驗證了全部四個結果。

常見錯誤

- 取 $q < p$：三次項係數為正時，極大點必定在極小點的左邊。
- 積分時用了 y 而不是 $y-n$，這樣會多算了直線 $y=n$ 之下、高度為 n 的長方形。
- 把對稱性說成是對直線 $x = \frac{1}{2}(p+q)$ 的反射；三次曲線具有的是繞一點的半周旋轉對稱，而不是鏡像對稱。

STEP 之外

每一條三次曲線都繞其拐點成半周旋轉對稱，因為平移 x 以消去 x^2 項（即卡爾達諾（Cardano）公式中所用的「缺項三次式」（depressed cubic）代換）之後，餘下的是一個奇函數加一個常數。第 (iv) 部分的二重根技巧，是相切與根的重數（multiplicity）理論的起點，這套理論在代數幾何（algebraic geometry）中佔有核心地位。Siklos 的 *Advanced Problems in Mathematics* 載有更多同類風格的題目。

詳細解答

STEP II 2007，第 3 題

純數學 • 代換積分法

STEP II 2007，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/07-s2-q3

關鍵思路

一切都建基於 $1/(a^2+x^2)$ 的標準積分（standard integral），所以整道題的重點，就是在代換（substitution）之後認出這個形狀。在 (i)(a) 中，分子（numerator）$\cos x$ 正是 $\sin x$ 的導數（derivative），這是個很強的提示。在 (i)(b) 中，半角代換（half-angle substitution）把 $1+\sin^2 x$ 變成 t 的四次式（quartic），而關鍵在於看出同樣的做法可以倒過來用：在 (ii) 中，我們要找出哪一個三角積分會產生新的四次式。

解答

標準積分

設 $a \neq 0$，並取 $x = a\tan\theta$，其中 $-\frac{1}{2}\pi < \theta < \frac{1}{2}\pi$。當 θ 走遍這個區間時，x 恰好取遍每個實數一次（不論 a 的正負），所以這個代換是有效的，而且由 $\tan\theta = x/a$ 可得 $\theta = \arctan(x/a)$。現在 $\mathrm{d}x = a\sec^2\theta\,\mathrm{d}\theta$，而

$$a^2 + x^2 = a^2(1+\tan^2\theta) = a^2\sec^2\theta,$$

所以

$$\int \frac{1}{a^2+x^2}\,\mathrm{d}x = \int \frac{a\sec^2\theta}{a^2\sec^2\theta}\,\mathrm{d}\theta = \int \frac{1}{a}\,\mathrm{d}\theta = \frac{\theta}{a} + \text{常數} = \frac{1}{a}\arctan\frac{x}{a} + \text{常數},$$

這正是所要證明的結果。

第 (i)(a) 部分

代入 $u = \sin x$，則 $\mathrm{d}u = \cos x \,\mathrm{d}x$，而當 x 由 0 變到 $\frac{1}{2}\pi$ 時，u 由 0 變到 1。利用標準積分並取 $a = 1$，

$$I = \int_0^1 \frac{1}{1+u^2}\,\mathrm{d}u = \Big[\arctan u\Big]_0^1 = \frac{\pi}{4}.$$

第 (i)(b) 部分

取 $t = \tan\frac{1}{2}x$，便有以下標準的半角結果

$$\sin x = \frac{2t}{1+t^2}, \qquad \cos x = \frac{1-t^2}{1+t^2}, \qquad \mathrm{d}x = \frac{2}{1+t^2}\,\mathrm{d}t,$$

最後一式是因為 $\mathrm{d}t/\mathrm{d}x = \frac{1}{2}\sec^2\frac{1}{2}x = \frac{1}{2}(1+t^2)$。積分上下限（limits）$x = 0$ 和 $x = \frac{1}{2}\pi$ 分別變成 $t = 0$ 和 $t = \tan\frac{1}{4}\pi = 1$。接着，

$$1 + \sin^2 x = \frac{(1+t^2)^2 + 4t^2}{(1+t^2)^2} = \frac{1+6t^2+t^4}{(1+t^2)^2}.$$

把各部分合起來，$1+t^2$ 的因子完全約去：

$$I = \int_0^1 \frac{1-t^2}{1+t^2}\cdot\frac{(1+t^2)^2}{1+6t^2+t^4}\cdot\frac{2}{1+t^2}\,\mathrm{d}t = 2\int_0^1 \frac{1-t^2}{1+6t^2+t^4}\,\mathrm{d}t.$$

兩邊除以 2，便得 $\displaystyle\int_0^1 \frac{1-t^2}{1+6t^2+t^4}\,\mathrm{d}t = \frac{1}{2}I$，這正是所要證明的結果。（由 (i)(a)，它的值是 $\pi/8$。）

第 (ii) 部分

> **想一想** 四次式 $1+14t^2+t^4$ 可能是從哪裏來的？

在 (i)(b) 中，$6t^2$ 的來源是：把 $1+t^2$ 平方得到的 $2t^2$，再加上由 $\sin^2 x$ 得到的 $4t^2 = (2t)^2$。要得到 $14t^2$，我們需要額外的 $12t^2 = 3(2t)^2$，也就是以 $3\sin^2 x$ 取代 $\sin^2 x$。事實上，取 $t = \tan\frac{1}{2}x$，

$$1 + 3\sin^2 x = \frac{(1+t^2)^2 + 12t^2}{(1+t^2)^2} = \frac{1+14t^2+t^4}{(1+t^2)^2}.$$

因此，把 (i)(b) 的計算原封不動地重做一次，只是以 $1+3\sin^2 x$ 取代 $1+\sin^2 x$，便得

$$\int_0^{\frac{1}{2}\pi} \frac{\cos x}{1+3\sin^2 x}\,\mathrm{d}x = 2\int_0^1 \frac{1-t^2}{1+14t^2+t^4}\,\mathrm{d}t.$$

仿照 (i)(a)，用 $u = \sin x$ 計算左邊：

$$\int_0^1 \frac{1}{1+3u^2}\,\mathrm{d}u = \frac{1}{3}\int_0^1 \frac{1}{\frac{1}{3}+u^2}\,\mathrm{d}u = \frac{1}{3}\cdot\sqrt{3}\Big[\arctan\left(\sqrt{3}\,u\right)\Big]_0^1 = \frac{1}{\sqrt{3}}\cdot\frac{\pi}{3},$$

這裏用了標準積分，並取 $a = 1/\sqrt{3}$。取其一半，

$$\int_0^1 \frac{1-t^2}{1+14t^2+t^4}\,\mathrm{d}t = \frac{\pi}{6\sqrt{3}} = \frac{\sqrt{3}\,\pi}{18}.$$

驗算

用辛普森法則（Simpson's rule）可得 $I \approx 0.785398$（即 $\pi/4$），(i)(b) 的積分 ≈ 0.392699（即 $\pi/8$），而 (ii) 的積分 ≈ 0.302300，與 $\sqrt{3}\,\pi/18 \approx 0.302300$ 吻合。再做一個簡單的合理性檢查：兩個四次式的被積函數（integrand）在 $t = 0$ 時都等於 1，在 $t = 1$ 時都等於 0，而分母含 $14t^2$ 的那個在整個區間上都較小，所以它的積分應該小於 $\pi/8 \approx 0.393$，事實也確是如此。

常見錯誤

- 忘記了來自 $\mathrm{d}x$ 的因子 $\frac{2}{1+t^2}$，以致漏掉 (i)(b) 中的因子 2。
- 寫成 $\int \frac{\mathrm{d}u}{1+3u^2} = \arctan(\sqrt{3}u)$ 而漏掉因子 $\frac{1}{\sqrt{3}}$；應先把 3 提出來，或者代入 $v = \sqrt{3}u$。
- 代換之後仍然沿用 x 的上下限：$x = \frac{1}{2}\pi$ 對應的是 $t = 1$。

STEP 之外

還有一條直接的路徑：把分子和分母同時除以 t^2，再取 $v = t + 1/t$，便可把 $\frac{1-t^2}{1+kt^2+t^4}$ 變成 $-\frac{1}{v^2+k-2}$，這又是一個 arctan 積分。這裏帶出的一般道理是：$\sin x$ 和 $\cos x$ 的有理函數（rational function），在代換 $t = \tan\frac{1}{2}x$ 之下總會變成 t 的有理函數，因而原則上總是可以積分的；這正是以初等函數表示積分的系統理論的起點。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 II 2007，第 4 題

純數學 • 三角方程

STEP II 2007，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/07-s2-q4

關鍵思路

用複角公式（compound-angle formulae）把所有項展開，再除以 $\cos A \cos B$：方程便化成關於 $t = \tan A$ 和 $u = \tan B$ 的關係式，其中有一個 tu 項、一個 t 項、一個 u 項和一個常數項。這樣的式子能因式分解（factorise）成兩個一次括號之積，當且僅當「配成矩形」之後沒有餘項，而這個餘項原來是 $\beta^2 + \gamma^2 - \alpha^2$ 的倍數。其後三條方程都有 $\alpha = 2$, $\beta = \sqrt{3}$, $\gamma = 1$，它們滿足這個條件；唯一要做的工夫就是正確地選取 A 和 B。

解答

因式分解

由複角公式，

$$\begin{aligned}&\alpha \sin(A-B) + \beta\cos(A+B) - \gamma \sin(A+B)\\ &\quad = \alpha(\sin A\cos B - \cos A \sin B) + \beta(\cos A\cos B - \sin A \sin B)\\ &\qquad - \gamma(\sin A\cos B + \cos A \sin B).\end{aligned}$$

由於 $\cos A \cos B \neq 0$，我們可以除以它。記 $t = \tan A$, $u = \tan B$，則方程等價於 $P(t,u) = 0$，其中

$$P(t,u) = \alpha(t-u) + \beta(1-tu) - \gamma(t+u) = -\beta tu + (\alpha-\gamma)t - (\alpha+\gamma)u + \beta.$$

由於 $\beta \neq 0$，設

$$m = -\frac{\alpha+\gamma}{\beta}, \qquad n = \frac{\alpha-\gamma}{\beta}.$$

於是 $-\beta(t-m)(u-n) = -\beta tu + \beta n\,t + \beta m\,u - \beta mn = -\beta tu + (\alpha-\gamma)t - (\alpha+\gamma)u + \dfrac{\alpha^2-\gamma^2}{\beta}$，除常數項外與 P 完全吻合。因此，對所有 t 和 u，

$$P(t,u) = -\beta(t-m)(u-n) + \frac{\beta^2+\gamma^2-\alpha^2}{\beta}. \qquad (*)$$

充分性。當 $\alpha^2 = \beta^2+\gamma^2$ 時，$(*)$ 中的常數項為零，所以（因為 $\beta \neq 0$）方程就是 $(\tan A - m)(\tan B - n) = 0$，其中 m 和 n 只取決於 α, β, γ。

必要性。假設方程可化為 $(t-m')(u-n') = 0$，其中 m', n' 為常數，即 $P(t,u) = 0$ 成立當且僅當 $t = m'$ 或 $u = n'$。記 $c = (\beta^2+\gamma^2-\alpha^2)/\beta$，並假設 $c \neq 0$。這時每一對 (m', u) 都是解，所以 $P(m', u) = -\beta(m'-m)(u-n) + c$ 對每個 u 都等於零。若 $m' \neq m$，這是 u 的一個非常數線性函數（linear function），不會恆等於零；若 $m' = m$，它等於 $c \neq 0$，永不為零。無論哪種情況都得出矛盾，所以 $c = 0$，即 $\alpha^2 = \beta^2+\gamma^2$。（若把「可化為」理解成「$P$ 是 $(t-m)(u-n)$ 的常數倍」，比較 tu 的係數便知這個倍數必為 $-\beta$，然後 $(*)$ 同樣迫使 $c = 0$。）

想一想 在下面各部分中，如果 $\cos x = 0$，即我們不能作除法的時候，會發生甚麼事？

把 $(*)$ 兩邊乘回 $\cos A \cos B$。當 $\alpha^2 = \beta^2+\gamma^2$ 時，我們得到

$$\alpha\sin(A-B) + \beta\cos(A+B) - \gamma\sin(A+B) = -\beta(\sin A - m\cos A)(\sin B - n\cos B), \qquad (**)$$

這時兩邊都只是正弦和餘弦的組合，所以這個恆等式（identity）對所有 A 和 B 都成立（兩邊都是連續的，而且只要 $\cos A\cos B \neq 0$ 便相等）。使用 $(**)$，我們就毋須另外處理 $\cos A = 0$ 的情況。

在全部三個部分中，$\alpha = 2, \beta = \sqrt{3}, \gamma = 1$，而 $\alpha^2 = 4 = \beta^2+\gamma^2$。所以

$$m = -\frac{3}{\sqrt{3}} = -\sqrt{3}, \qquad n = \frac{1}{\sqrt{3}},$$

方程就是 $(\sin A + \sqrt{3}\cos A)(\sin B - \frac{1}{\sqrt{3}}\cos B) = 0$。留意 $\sin A + \sqrt{3}\cos A = 0$ 必然令 $\cos A \neq 0$（否則 $\sin A = 0$ 亦成立），所以它等同於 $\tan A = -\sqrt{3}$；同樣地，第二個括號為零當且僅當 $\tan B = \frac{1}{\sqrt{3}}$。

第 (i) 部分

取 $A = x$ 及 $B = \frac{1}{4}\pi$，使得 $A - B = x - \frac{1}{4}\pi$ 及 $A + B = x + \frac{1}{4}\pi$。由於 $\tan\frac{1}{4}\pi = 1 \neq \frac{1}{\sqrt{3}}$，第二個括號不為零，所以我們需要 $\tan x = -\sqrt{3}$：

$$x = \tfrac{2}{3}\pi \quad 或 \quad x = \tfrac{5}{3}\pi.$$

第 (ii) 部分

現在 $A = x$ 及 $B = \frac{1}{6}\pi$。但 $\tan\frac{1}{6}\pi = \frac{1}{\sqrt{3}} = n$，所以第二個括號是 $\sin\frac{1}{6}\pi - \frac{1}{\sqrt{3}}\cos\frac{1}{6}\pi = \frac{1}{2} - \frac{1}{2} = 0$。由 $(**)$，左方減右方恆等於零：這條方程是一個恆等式，每一個滿足 $0 \le x < 2\pi$ 的 x 都是解。

第 (iii) 部分

我們需要 $A - B = x + \frac{1}{3}\pi$ 及 $A + B = 3x$。兩式相加和相減，得

$$A = 2x + \tfrac{1}{6}\pi, \qquad B = x - \tfrac{1}{6}\pi.$$

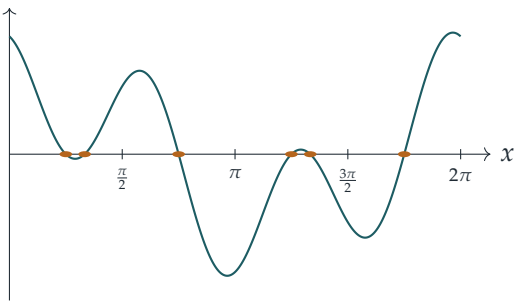
所以解由 $\tan(2x + \frac{1}{6}\pi) = -\sqrt{3}$ 或 $\tan(x - \frac{1}{6}\pi) = \frac{1}{\sqrt{3}}$ 給出。

對於前者，$2x + \frac{1}{6}\pi = \frac{2}{3}\pi + k\pi$，所以 $x = \frac{1}{4}\pi + \frac{1}{2}k\pi$，其中 k 為整數。在 $[0, 2\pi)$ 內，這給出 $x = \frac{1}{4}\pi,\ \frac{3}{4}\pi,\ \frac{5}{4}\pi,\ \frac{7}{4}\pi$。

對於後者，$x - \frac{1}{6}\pi = \frac{1}{6}\pi + k\pi$，所以 $x = \frac{1}{3}\pi + k\pi$，給出 $x = \frac{1}{3}\pi,\ \frac{4}{3}\pi$。

這六個值互不相同，所以解為

$$x = \tfrac{1}{4}\pi,\ \tfrac{1}{3}\pi,\ \tfrac{3}{4}\pi,\ \tfrac{5}{4}\pi,\ \tfrac{4}{3}\pi,\ \tfrac{7}{4}\pi.$$



(iii) 的左方減右方的圖像，恰好在標示的六個點穿過零。

驗算

- 在 (i) 中，當 $x = \frac{2}{3}\pi$ 時：$2\sin\frac{5}{12}\pi \approx 1.932$，$\sqrt{3}\cos\frac{11}{12}\pi \approx -1.673$，而 $\sin\frac{11}{12}\pi \approx 0.259$；的確 $1.932 - 1.673 = 0.259$。
- 在 (ii) 中，試 $x = 0$：$2\sin(-\frac{1}{6}\pi) + \sqrt{3}\cos\frac{1}{6}\pi = -1 + \frac{3}{2} = \frac{1}{2} = \sin\frac{1}{6}\pi$。其他任何 x 也同樣成立。
- 在 (iii) 中，當 $x = \frac{1}{4}\pi$ 時：$2\sin\frac{7}{12}\pi \approx 1.932$，而 $\sin\frac{3}{4}\pi - \sqrt{3}\cos\frac{3}{4}\pi \approx 0.707 + 1.225 = 1.932$。在 $[0, 2\pi)$ 上以數值方法搜尋根，找到的正好是上述六個根。

常見錯誤

- 在 (iii) 中取 $A = x + \frac{1}{3}\pi$，並由 $3x$ 得出 B：你必須聯立求解 $A - B = x + \frac{1}{3}\pi$, $A + B = 3x$。
- 在 (iii) 中只取到 $k = 0, 1$ 便停下而漏掉解：$\tan(2x + \frac{1}{6}\pi)$ 的週期是 $\frac{1}{2}\pi$，所以該括號給出四個解。
- 沒有察覺 (ii) 是恆等式，反而去找個別的解。

STEP 之外

式子 $-\beta tu + (\alpha - \gamma)t - (\alpha + \gamma)u + \beta$ 是圓錐曲線（conic）的一種退化情況：它的零點集是一條等軸雙曲線（rectangular hyperbola），除非某個類似行列式（determinant）的量為零，這時它便分裂成兩條直線。判斷一個二次型（quadratic form）何時可以因式分解，是用矩陣和特徵值（eigenvalues）為圓錐曲線分類的第一步，這是大學一年級線性代數（linear algebra）的課題；見 Strang, *Introduction to Linear Algebra*（線性代數入門教科書）。Siklos, *Advanced Problems in Mathematics*（進階數學難題集）中有更多同類風格的三角方程。

詳細解答

STEP II 2007，第 5 題

純數學 • 迭代函數

> **STEP II 2007，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/07-s2-q5

關鍵思路

$\dfrac{x+\sqrt{3}}{1-\sqrt{3}\,x}$ 這個形狀，正是 $\tan(A+B)$ 的複角公式（compound-angle formula），其中 $\tan B = \sqrt{3}$，即 $B = \frac{\pi}{3}$。所以 f 的作用不過是把角度加上 $\frac{\pi}{3}$；而 tan 的週期（period）是 π，因此走三步便回到起點。第 (iii) 部分可以用同一招，只是換成 sin 和 $\frac{\pi}{6}$，但當中有個陷阱：$\sqrt{1-t^2}$ 只有在角度位於 $[-\frac{\pi}{2}, \frac{\pi}{2}]$ 之內時才是餘弦，而這一點會改變答案。

解答

第 (i) 部分

設 $y = f(x) = \dfrac{x+\sqrt{3}}{1-\sqrt{3}\,x}$。那麼 $f^2(x) = \dfrac{y+\sqrt{3}}{1-\sqrt{3}\,y}$，把分子和分母同乘以 $1-\sqrt{3}\,x$，得

$$f^2(x) = \frac{(x+\sqrt{3})+\sqrt{3}(1-\sqrt{3}\,x)}{(1-\sqrt{3}\,x)-\sqrt{3}(x+\sqrt{3})} = \frac{2\sqrt{3}-2x}{-2-2\sqrt{3}\,x} = \frac{x-\sqrt{3}}{1+\sqrt{3}\,x}.$$

現在再應用 f 一次，這次取 $y = f^2(x)$，並把分子和分母同乘以 $1+\sqrt{3}\,x$：

$$f^3(x) = \frac{(x-\sqrt{3})+\sqrt{3}(1+\sqrt{3}\,x)}{(1+\sqrt{3}\,x)-\sqrt{3}(x-\sqrt{3})} = \frac{4x}{4} = x.$$

所以 f^3 是恆等映射（identity）。（題目排除的那兩個點，正好保證每一步都有定義：$f(x)$ 需要 $x \neq 1/\sqrt{3}$，$f^2(x)$ 需要 $x \neq -1/\sqrt{3}$，而 f 把 $\mathbb{R} \setminus \{\pm 1/\sqrt{3}\}$ 映滿到它自身，所以我們可以一直做下去。）

由於 $2007 = 3 \times 669$，

$$f^{2007}(x) = \left(f^3\right)^{669}(x) = x.$$

第 (ii) 部分

設 $x = \tan\theta$。對任何使 $\tan\alpha$ 有定義且 $1 - \sqrt{3}\tan\alpha \neq 0$ 的角 α，利用 $\tan\frac{\pi}{3} = \sqrt{3}$ 和複角公式可得

$$f(\tan\alpha) = \frac{\tan\alpha + \tan\frac{\pi}{3}}{1 - \tan\alpha\tan\frac{\pi}{3}} = \tan\big(\alpha + \tfrac{\pi}{3}\big).$$

取 $\alpha = \theta$，這就是 $n = 1$ 的情況。若 $f^k(x) = \tan(\theta + \frac{1}{3}k\pi)$，則以 $\alpha = \theta + \frac{1}{3}k\pi$ 代入上述公式，得

$$f^{k+1}(x) = f\big(\tan(\theta + \tfrac{1}{3}k\pi)\big) = \tan\big(\theta + \tfrac{1}{3}(k+1)\pi\big),$$

所以根據歸納法（induction），對每個正整數 n 都有 $f^n(x) = \tan(\theta + \frac{1}{3}n\pi)$。

整個過程不會出問題：公式只會在某個角 $\theta + \frac{1}{3}k\pi$ 與 $\frac{\pi}{2}$ 或 $\frac{\pi}{6}$ 模（modulo）π 同餘時失效，而這會迫使 $\theta = \pm\frac{\pi}{6}$ 或 $\pm\frac{\pi}{2}$，即 $x = \pm 1/\sqrt{3}$ 或 x 為無窮大，這些情況都已被排除。這也解釋了第 (i) 部分：加三次 $\frac{\pi}{3}$ 就是加上 π，即 tan 的週期。

第 (iii) 部分

由於 $|t| \leq 1$，我們可以寫 $t = \sin\varphi$，其中 $\varphi = \arcsin t \in [-\frac{\pi}{2}, \frac{\pi}{2}]$。

想一想 $\sqrt{1 - \sin^2\alpha}$ 是否總是等於 $\cos\alpha$？

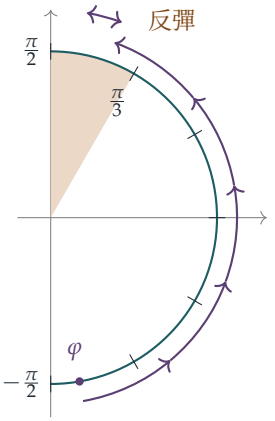
只有當 $\cos\alpha \geq 0$ 時才是。對於角 $\alpha \in [-\frac{\pi}{2}, \frac{\pi}{2}]$，這沒有問題，所以

$$g(\sin\alpha) = \cos\tfrac{\pi}{6}\sin\alpha + \sin\tfrac{\pi}{6}\cos\alpha = \sin\big(\alpha + \tfrac{\pi}{6}\big). \qquad (*)$$

要再次應用 $(*)$，我們必須把結果改寫成 $[-\frac{\pi}{2}, \frac{\pi}{2}]$ 內某個角的 sin。這裏分兩種情況。

- 若 $\alpha \leq \frac{\pi}{3}$，則 $\alpha + \frac{\pi}{6} \leq \frac{\pi}{2}$ 已經在範圍之內。
- 若 $\frac{\pi}{3} \leq \alpha \leq \frac{\pi}{2}$，則 $\alpha + \frac{\pi}{6}$ 可能會超過 $\frac{\pi}{2}$，但 $\sin(\alpha + \frac{\pi}{6}) = \sin\big(\pi - \alpha - \frac{\pi}{6}\big) = \sin\big(\frac{5\pi}{6} - \alpha\big)$，而且 $\frac{5\pi}{6} - \alpha \in [\frac{\pi}{3}, \frac{\pi}{2}]$。

（當 $\alpha = \frac{\pi}{3}$ 時，兩者都給出 $\frac{\pi}{2}$。）所以這個範圍內的角每次上升 $\frac{\pi}{6}$，直至進入 $[\frac{\pi}{3}, \frac{\pi}{2}]$；此後它便按 $\alpha \mapsto \frac{5\pi}{6} - \alpha$ 被反射，而再反射一次就會還原：這個角在兩個值之間來回反彈。



每應用一次 g，角就沿單位圓的右半邊移動 $\frac{\pi}{6}$；一旦進入陰影區域 $[\frac{\pi}{3}, \frac{\pi}{2}]$，它便來回翻轉。

現在設 m 為所需的上升次數，即滿足 $\varphi + \frac{m\pi}{6} \geq \frac{\pi}{3}$ 的最小整數 $m \geq 0$。由於 $\varphi \geq -\frac{\pi}{2}$，故 $m \leq 5$；以 t 表示則為：

t	$[\frac{\sqrt{3}}{2}, 1]$	$[\frac{1}{2}, \frac{\sqrt{3}}{2})$	$[0, \frac{1}{2})$	$[-\frac{1}{2}, 0)$	$[-\frac{\sqrt{3}}{2}, -\frac{1}{2})$	$[-1, -\frac{\sqrt{3}}{2})$
m	0	1	2	3	4	5

寫 $g^n(t) = \sin\varphi_n$，其中 $\varphi_n \in [-\frac{\pi}{2}, \frac{\pi}{2}]$，且 $\varphi_0 = \varphi$。當 $n \leq m$ 時，之前的每個角都小於 $\frac{\pi}{3}$，所以 $\varphi_n = \varphi + \frac{n\pi}{6}$。於是 $\varphi_m \in [\frac{\pi}{3}, \frac{\pi}{2}]$，所以 $\varphi_{m+1} = \frac{5\pi}{6} - \varphi_m$，$\varphi_{m+2} = \varphi_m$，如此類推，交替出現。由於 $\sin\varphi_{m+1} = \sin(\varphi_m + \frac{\pi}{6}) = \sin\big(\varphi + \frac{(m+1)\pi}{6}\big)$，我們得到（其中 $\varphi = \arcsin t$，m 取自上表）

$$g^n(t) = \begin{cases} \sin\big(\varphi + \frac{n\pi}{6}\big), & n \leq m+1, \\ \sin\big(\varphi + \frac{m\pi}{6}\big), & n > m+1,\ n-m \text{ 為偶數}, \\ \sin\big(\varphi + \frac{(m+1)\pi}{6}\big), & n > m+1,\ n-m \text{ 為奇數}. \end{cases}$$

特別地，對每個 t 和每個 $n \geq 5$，都有 $g^{n+2}(t) = g^n(t)$。

例如，$t = 0$ 時 $\varphi = 0$，$m = 2$，其軌道（orbit）為 $0 \mapsto \frac{1}{2} \mapsto \frac{\sqrt{3}}{2} \mapsto 1 \mapsto \frac{\sqrt{3}}{2} \mapsto 1 \mapsto \cdots$。

驗算

- 取 $x = 1$：$f(1) = \frac{1+\sqrt{3}}{1-\sqrt{3}} = -2 - \sqrt{3}$，$f^2(1) = \frac{1-\sqrt{3}}{1+\sqrt{3}} = \sqrt{3} - 2$，且 $f(\sqrt{3} - 2) = 1$，與 $\tan\frac{\pi}{4}$、$\tan\frac{7\pi}{12}$、$\tan\frac{11\pi}{12}$ 相符。
- 直接計算，$g(\frac{\sqrt{3}}{2}) = \frac{3}{4} + \frac{1}{4} = 1$ 及 $g(1) = \frac{\sqrt{3}}{2}$：反彈確實存在。想當然地猜 $g^5(0)$ 等於 $\sin\frac{5\pi}{6} = \frac{1}{2}$ 是錯的；逐次迭代得到的是 1。

- 數值上，從 5000 個隨機的 $t \in [-1,1]$ 出發，把 g 迭代十九次，結果與方框內的公式相差不超過 2×10^{-12}。

常見錯誤

- 在 (iii) 中對所有 n 都答 $\sin(\varphi + \frac{n\pi}{6})$。這忽略了 $\sqrt{1-t^2} \geq 0$，所以角永遠不能越過 $\frac{\pi}{2}$。
- 計算 f^2 時的運算失誤：應先把分子和分母同乘以 $1 - \sqrt{3}\,x$，然後才化簡。

STEP 之外

形如 $x \mapsto \frac{ax+b}{cx+d}$ 的映射稱為莫比烏斯變換（Möbius transformation），而把兩個這樣的變換複合，就相當於把它們的矩陣 $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$ 相乘。這裏 f 的矩陣是 $2\left(\begin{smallmatrix} \cos\frac{\pi}{3} & \sin\frac{\pi}{3} \\ -\sin\frac{\pi}{3} & \cos\frac{\pi}{3} \end{smallmatrix}\right)$，即一個經縮放、轉角為 $\frac{\pi}{3}$ 的旋轉，它的三次方是單位矩陣的倍數，這就是 $f^3(x) = x$ 的原因。第 (iii) 部分讓我們淺嘗動力系統（dynamical systems）的滋味：那些點是最終週期（eventually periodic）的，而不是週期的。Needham 的 *Visual Complex Analysis*（以幾何直觀講解複分析的著作）對莫比烏斯變換有十分優美的幾何處理。

詳細解答

STEP II 2007，第 6 題

純數學 • 一階微分方程

STEP II 2007，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/07-s2-q6

關鍵思路

第 (i) 部分給出兩個函數，它們的導數（derivative）的分母都是 $\sqrt{3+x^2}$；把兩者適當地組合起來，所得的導數正好是 $\sqrt{3+x^2}$ 本身的倍數，於是所求的積分（integral）便唾手可得。第 (ii) 部分的微分方程（differential equation）是 $\frac{\mathrm{d}y}{\mathrm{d}x}$ 的二次方程。解這個二次方程，便得到兩條普通的一階方程，而每一條所需的正是第 (i) 部分的積分。

解答

第 (i) 部分

設 $s=\sqrt{3+x^2}$，則對任何實數 x 都有 $s>0$，並且 $\dfrac{\mathrm{d}s}{\mathrm{d}x}=\dfrac{x}{s}$。另外，由於 $s>|x|$，所以 $x+s>0$，因此該對數對所有 x 都有定義。

根據鏈式法則（chain rule），

$$\frac{\mathrm{d}}{\mathrm{d}x}\ln(x+s)=\frac{1+\dfrac{x}{s}}{x+s}=\frac{\dfrac{s+x}{s}}{x+s}=\frac{1}{\sqrt{3+x^2}}.$$

根據積法則（product rule），

$$\frac{\mathrm{d}}{\mathrm{d}x}(xs)=s+x\cdot\frac{x}{s}=\frac{s^2+x^2}{s}=\frac{3+2x^2}{\sqrt{3+x^2}}.$$

? **想一想** 這兩個導數怎樣給出 $\int\sqrt{3+x^2}\,\mathrm{d}x$？

我們希望分母為 s 而分子為 $3+x^2$，因為 $\dfrac{3+x^2}{s}=s$。第二個導數的分子是 $3+2x^2$；加

上第一個導數的 3 倍，便得

$$\frac{\mathrm{d}}{\mathrm{d}x}\Big(xs+3\ln(x+s)\Big)=\frac{3+2x^2+3}{s}=\frac{2(3+x^2)}{s}=2\sqrt{3+x^2}.$$

因此

$$\int\sqrt{3+x^2}\,\mathrm{d}x=\tfrac{1}{2}x\sqrt{3+x^2}+\tfrac{3}{2}\ln\big(x+\sqrt{3+x^2}\,\big)+C.$$

第 (ii) 部分

設 $p=\dfrac{\mathrm{d}y}{\mathrm{d}x}$。原方程就是 p 的二次方程 $3p^2+2xp-1=0$，所以

$$p=\frac{-2x\pm\sqrt{4x^2+12}}{6}=\frac{-x\pm\sqrt{3+x^2}}{3}.$$

判別式（discriminant）$4x^2+12$ 恆為正數，所以在每一個 x 處都有兩個相異實根。兩根之差 $\frac{2}{3}\sqrt{3+x^2}$ 永不為零，所以一個解不可能中途由一個根轉到另一個根。（導數具有介值性質（intermediate value property），所以它不可能在兩個始終相差最少 $\frac{2}{3}\sqrt{3}$ 的值之間跳躍。）因此，每一個解都對所有 x 滿足

$$\frac{\mathrm{d}y}{\mathrm{d}x}=-\frac{x}{3}+\frac{1}{3}\sqrt{3+x^2}\qquad\text{或}\qquad\frac{\mathrm{d}y}{\mathrm{d}x}=-\frac{x}{3}-\frac{1}{3}\sqrt{3+x^2}$$

其中之一。利用第 (i) 部分的結果積分，

$$y=-\frac{x^2}{6}\pm\Big(\tfrac{1}{6}x\sqrt{3+x^2}+\tfrac{1}{2}\ln\big(x+\sqrt{3+x^2}\,\big)\Big)+C.$$

當 $x=1$ 時，$\sqrt{3+x^2}=2$，而 $\ln(x+\sqrt{3+x^2})=\ln 3$。

取 $+$ 號：由 $y(1)=0$ 得 $0=-\frac{1}{6}+\frac{1}{3}+\frac{1}{2}\ln 3+C$，所以 $C=-\frac{1}{6}-\frac{1}{2}\ln 3$。把對數項合併，

$$y=\frac{x\sqrt{3+x^2}-x^2-1}{6}+\frac{1}{2}\ln\left(\frac{x+\sqrt{3+x^2}}{3}\right).$$

取 $-$ 號：由 $y(1)=0$ 得 $0=-\frac{1}{6}-\frac{1}{3}-\frac{1}{2}\ln 3+C$，所以 $C=\frac{1}{2}+\frac{1}{2}\ln 3$，於是

$$y=\frac{3-x^2-x\sqrt{3+x^2}}{6}-\frac{1}{2}\ln\left(\frac{x+\sqrt{3+x^2}}{3}\right).$$

這就是通過 $(1,0)$ 的兩個解，兩者對所有實數 x 都成立。

驗算

- 把 $x=1$ 代入每個答案：分式部分是 $\frac{2-1-1}{6}=0$（或 $\frac{3-1-2}{6}=0$），而 $\ln\frac{1+2}{3}=0$。所以兩個答案都有 $y(1)=0$。

- p 的兩個根之和是 $-\frac{2x}{3}$，所以兩個解之和應為 $\frac{1-x^2}{3}$。平方根項和對數項確實互相抵消，剩下的正好是 $\frac{2-2x^2}{6} = \frac{1-x^2}{3}$。
- 數值上，用辛普森法則（Simpson's rule）可得 $\int_{-1}^{2} \sqrt{3+x^2}\,\mathrm{d}x \approx 5.949681$，與第 (i) 部分的公式吻合。用龍格－庫塔法（Runge–Kutta）把方程的每一個分支由 $(1,0)$ 逐步計算至 $x = 4$，得 $y \approx 0.584957$ 和 $y \approx -5.584957$，與兩條公式吻合。（兩者之和是 $-5 = \frac{1-16}{3}$。）

常見錯誤

- 把 $3p^2 + 2xp = 1$ 當作線性方程處理，或嘗試分離變數（separate variables），而不是把它當作 p 的二次方程來解。
- 兩個分支共用同一個常數 C。每個符號都需要各自的常數，並由 $y(1) = 0$ 確定。
- 漏掉 $\sqrt{3+x^2}$ 前面的因子 $\frac{1}{3}$，以致使用第 (i) 部分的積分時倍數出錯。

STEP 之外

形如 $F\left(x, y, \frac{\mathrm{d}y}{\mathrm{d}x}\right) = 0$ 而對導數並非線性的方程，稱為隱式（implicit）方程。通過同一點可以有多個解，方程中 p 的每一個根各對應一個。在兩個根重合之處，可能出現奇解（singular solutions），即一族曲線的包絡線（envelope），克萊羅方程（Clairaut's equation）便是一例。本題的兩個根永不重合，所以恰好有兩個解。第 (i) 部分的 $\ln(x + \sqrt{3+x^2})$ 等於 $\operatorname{arsinh}(x/\sqrt{3})$ 加上一個常數，而該積分其實就是標準的代換（substitution）$x = \sqrt{3}\sinh u$，只是換了個面貌。Boyce 與 DiPrima 合著的 *Elementary Differential Equations*（初等微分方程教科書）涵蓋一階方程，以及其解的存在性與唯一性（existence and uniqueness）。

詳細解答

STEP II 2007，第 7 題

純數學 • 琴生不等式

> **STEP II 2007，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/07-s2-q7

本題在 *Beyond STEP* 第 3.3 節凸性的威力中有完整的引導式講解：https://beyondstep.keithmcnulty.org/zh-hant/ch03/convexity.html

關鍵思路

延森不等式（Jensen's inequality）指出：對於凹函數（concave function），函數值的平均不超過平均值處的函數值。每一部分的關鍵，都在於看出該用哪一個凹函數，以及該對哪些數取平均。就三角形而言，各角之和固定，所以它們的平均值是已知的。至於算術—幾何平均不等式（AM–GM inequality），對數（logarithm）能把積化為和。(a) 和 (b) 兩部分則需要選對代入 AM–GM 的一組數，當中包括一些常數。

解答

凹性

若 $f(x) = \sin x$，則 $f''(x) = -\sin x$。當 $0 < x < \pi$ 時，$\sin x > 0$，所以 $f''(x) < 0$，故 $\sin x$ 在該區間上是凹的。

若 $f(x) = \ln x$，則 $f'(x) = 1/x$ 及 $f''(x) = -1/x^2$。對所有 $x > 0$，這都是負數，所以 $\ln x$ 在 $x > 0$ 時是凹的。

第 (i) 部分

三角形的各角滿足 $A, B, C > 0$ 及 $A + B + C = \pi$，所以每個角都在 $(0, \pi)$ 之內，而 sin 在此區間上是凹的。取 $n = 3$，由延森不等式得

$$\frac{\sin A + \sin B + \sin C}{3} \leq \sin\left(\frac{A + B + C}{3}\right) = \sin\frac{\pi}{3} = \frac{\sqrt{3}}{2}.$$

兩邊乘以 3，

$$\sin A + \sin B + \sin C \leq \frac{3\sqrt{3}}{2},$$

正是所求。等號成立當且僅當 $A = B = C = \pi/3$，即等邊三角形。

第 (ii) 部分

取 $f(x) = \ln x$，它在 $x > 0$ 上是凹的，並設 $x_k = t_k > 0$。由延森不等式得

$$\frac{1}{n}\sum_{k=1}^{n} \ln t_k \leq \ln\left(\frac{t_1 + t_2 + \cdots + t_n}{n}\right).$$

左邊等於 $\frac{1}{n}\ln(t_1 t_2 \cdots t_n) = \ln\sqrt[n]{t_1 t_2 \cdots t_n}$。由於 e^x 是遞增函數，我們可以對兩邊取指數而不改變不等號的方向：

$$\sqrt[n]{t_1 t_2 \cdots t_n} \leq \frac{t_1 + t_2 + \cdots + t_n}{n},$$

等號成立當且僅當 $t_1 = t_2 = \cdots = t_n$。（當 $n = 1$ 時，這顯然是等式，與上述結論一致。）

(a)

想一想 我們應該對哪四個數取平均？

左邊有四項，所以取 $n = 4$。我們需要幾何平均（geometric mean）得出 xyz，而 x^4, y^4, z^4 給出 $\sqrt[4]{x^4 y^4 z^4} = xyz$。常數 16 就是第四項，而 $\sqrt[4]{16} = 2$。因此取 $t_1 = x^4$，$t_2 = y^4$，$t_3 = z^4$，$t_4 = 16$，全部為正數：

$$\frac{x^4 + y^4 + z^4 + 16}{4} \geq \sqrt[4]{16\, x^4 y^4 z^4} = 2xyz,$$

其中取四次方根時用到 $x, y, z > 0$。兩邊乘以 4 得

$$x^4 + y^4 + z^4 + 16 \geq 8xyz,$$

正是所求。等號成立需要 $x^4 = y^4 = z^4 = 16$，即 $x = y = z = 2$。

(b) 現在出現的是五次方，所以我們取 $n = 5$，並讓 x^5, y^5, z^5 成為其中三個數。要令五次方根恰好得出 xyz，其餘兩個數的積必須是 1。最自然的選擇是 $t_4 = t_5 = 1$：

$$\frac{x^5 + y^5 + z^5 + 1 + 1}{5} \geq \sqrt[5]{x^5 y^5 z^5 \cdot 1 \cdot 1} = xyz.$$

移項整理後，

$$x^5 + y^5 + z^5 - 5xyz \geq -2$$

對所有正數 x, y, z 成立。

想一想 這是否已證明最小值是 -2？

單憑這一點並不足夠：下界（lower bound）必須能夠取到，才是最小值。AM–GM 的等號在五個數全部相等時成立，即 $x^5 = y^5 = z^5 = 1$，所以 $x = y = z = 1$。此時該式的值為 $1 + 1 + 1 - 5 = -2$。因此最小值是

$$\boxed{-2}, \quad \text{取得於} x = y = z = 1.$$

驗算

- 在 (i) 中，等腰直角三角形給出 $1 + 2 \cdot \frac{\sqrt{2}}{2} \approx 2.414$，小於 $3\sqrt{3}/2 \approx 2.598$。隨機抽取 200 000 個三角形，結果從未超過 2.598。
- 在 (a) 中，當 $x = y = z = 2$ 時，兩邊都等於 64。
- 在 (b) 中，從隨機起點出發進行數值最小化，結果總是收斂到 $(1, 1, 1)$，其值為 -2。此外，$x = y = z = 1.1$ 給出 $3(1.61051) - 5(1.331) \approx -1.82 > -2$。

常見錯誤

- 忘記了取指數之所以保持不等號方向，全因 e^x 是遞增函數；這一點值得寫上一句說明。
- 在 (b) 中，寫到「≥ -2」便停筆，沒有舉出令該式達到 -2 的 x、y、z 的值。
- 對 x^5, y^5, z^5 使用 $n = 3$ 的 AM–GM，這樣只會得到 $x^5 + y^5 + z^5 \geq 3(xyz)^{5/3}$，無法與 $5xyz$ 直接比較。

STEP 之外

延森不等式是關於凸（convex）函數或凹函數的命題，它真正的歸宿是分析學和概率論：用期望值（expectation）的語言來說，它就是對凹函數 f 有 $\mathrm{E}[f(X)] \leq f(\mathrm{E}[X])$。(a) 和 (b) 中的技巧，即補上常數令幾何平均恰好符合所需，正是加權 AM–GM 不等式和赫爾德（Hölder）不等式理論的起點。想進一步了解不等式和凸性，可參閱 Bartle 與 Sherbert 的 *Introduction to Real Analysis*（實分析入門）；至於概率版本，可參閱 Grimmett 與 Welsh 的 *Probability: An Introduction*（概率論入門）。

詳細解答

STEP II 2007，第 8 題

純數學 • 以向量證明塞瓦定理

> **STEP II 2007，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/07-s2-q8

關鍵思路

平面上每一點都可以寫成 $\beta\mathbf{b} + \gamma\mathbf{c}$，而第 (i) 部分正好告訴我們這樣的點何時位於 BC 上：兩個係數之和為 1。因此，要找出一條直線與 BC（或 CA、AB）的交點，我們只需把該直線上的一般點用 $\mathbf{b}$ 和 $\mathbf{c}$ 表示，再對係數加上適當的條件。這樣每個比都可以用 β 和 γ 表示，而三者的乘積會全部約簡。

解答

在整個解答中，像 $BD : DC$ 這樣的比是指有向比（signed ratio）：若 $\overrightarrow{BD} = k\overrightarrow{DC}$，則 $BD/DC = k$；當 D 位於線段 BC 之外時，這個比是負數。由於 A、B、C 不共線（collinear），$\mathbf{b}$ 和 $\mathbf{c}$ 互不平行且均非零，所以我們可以比較係數：由 $p\mathbf{b} + q\mathbf{c} = p'\mathbf{b} + q'\mathbf{c}$ 必可得出 $p = p'$ 及 $q = q'$。

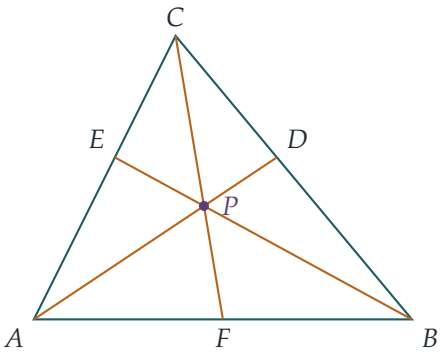
第 (i) 部分

若 $s + t = 1$，則

$$s\mathbf{b} + t\mathbf{c} = (1 - t)\mathbf{b} + t\mathbf{c} = \mathbf{b} + t(\mathbf{c} - \mathbf{b}),$$

這是通過 B 且方向為 $\overrightarrow{BC}$ 的直線的向量方程（vector equation），其中 t 為參數。當 t 取遍所有實數時，X 走遍整條直線。因此 X 的軌跡（locus）是通過 B 和 C 的（無限延伸的）直線；線段 BC 則是其中 $0 \le t \le 1$ 的部分。

反過來，比較係數可知，點 $\beta\mathbf{b} + \gamma\mathbf{c}$ 位於直線 BC 上當且僅當 $\beta + \gamma = 1$。這就是我們從現在起使用的判別方法。特別地，$\beta + \gamma \neq 1$ 表示 P 不在 BC 上，而 $\beta, \gamma \neq 0$ 表示 P 不在 CA 或 AB 上。



第 (ii) 部分

直線 AP 上的一般點是 $\lambda(\beta\mathbf{b}+\gamma\mathbf{c}) = \lambda\beta\,\mathbf{b} + \lambda\gamma\,\mathbf{c}$。根據第 (i) 部分，它位於 BC 上正好是當 $\lambda\beta + \lambda\gamma = 1$，即

$$\lambda = \frac{1}{\beta+\gamma}.$$

（這需要 $\beta+\gamma \neq 0$。若 $\beta+\gamma = 0$，則 $\overrightarrow{AP} = \beta(\mathbf{b}-\mathbf{c})$ 與 BC 平行，D 並不存在，所以題目其實隱含了 $\beta+\gamma \neq 0$ 的假設。）因此

$$\mathbf{d} = \frac{\beta\mathbf{b}+\gamma\mathbf{c}}{\beta+\gamma}.$$

現在

$$\overrightarrow{BD} = \mathbf{d}-\mathbf{b} = \frac{\beta\mathbf{b}+\gamma\mathbf{c}-(\beta+\gamma)\mathbf{b}}{\beta+\gamma} = \frac{\gamma}{\beta+\gamma}(\mathbf{c}-\mathbf{b}),$$
$$\overrightarrow{DC} = \mathbf{c}-\mathbf{d} = \frac{(\beta+\gamma)\mathbf{c}-\beta\mathbf{b}-\gamma\mathbf{c}}{\beta+\gamma} = \frac{\beta}{\beta+\gamma}(\mathbf{c}-\mathbf{b}).$$

兩者都是同一個向量 $\mathbf{c}-\mathbf{b}$ 的倍數，所以 $\beta\overrightarrow{BD} = \gamma\overrightarrow{DC}$，又因為 $\beta \neq 0$，

$$BD : DC = \gamma : \beta,$$

這正是所需的結果。

第 (iii) 部分

我們用同樣的方法求 E 和 F，不過這次的目標直線是 CA（$\mathbf{c}$ 的倍數）和 AB（$\mathbf{b}$ 的倍數）。

> ? **想一想** 怎樣判斷一點 $p\mathbf{b}+q\mathbf{c}$ 是否在直線 CA 上？

直線 CA 通過原點 A，方向為 $\mathbf{c}$，所以線上的點正好就是 $\mathbf{c}$ 的所有倍數：$\mathbf{b}$ 的係數必須為零。同樣地，一點在 AB 上，正好是當它的 $\mathbf{c}$ 係數為零。

求 E。直線 BP 上的一般點是

$$\mathbf{b}+\mu(\beta\mathbf{b}+\gamma\mathbf{c}-\mathbf{b})=\big(1-\mu(1-\beta)\big)\mathbf{b}+\mu\gamma\,\mathbf{c}.$$

令 $\mathbf{b}$ 的係數為零，得 $\mu=1/(1-\beta)$（我們需要 $\beta\neq1$；若 $\beta=1$，則 $\overrightarrow{BP}=\gamma\mathbf{c}$ 與 CA 平行，E 並不存在）。所以

$$\mathbf{e}=e\,\mathbf{c},\qquad e=\frac{\gamma}{1-\beta}.$$

於是 $\overrightarrow{CE}=(e-1)\mathbf{c}$ 及 $\overrightarrow{EA}=-e\,\mathbf{c}$，而且因為 $\gamma\neq0$，所以 $e\neq0$，因此

$$\frac{CE}{EA}=\frac{e-1}{-e}=\frac{1-e}{e}=\frac{(1-\beta-\gamma)/(1-\beta)}{\gamma/(1-\beta)}=\frac{1-\beta-\gamma}{\gamma}.$$

求 F。把 B 和 C 的角色互換（即 $\mathbf{b}\leftrightarrow\mathbf{c}$ 及 $\beta\leftrightarrow\gamma$），在假設 $\gamma\neq1$ 之下，直線 CP 與 AB 相交於 $\mathbf{f}=f\,\mathbf{b}$，其中 $f=\beta/(1-\gamma)$。於是 $\overrightarrow{AF}=f\,\mathbf{b}$ 及 $\overrightarrow{FB}=(1-f)\mathbf{b}$，其中 $1-f=(1-\beta-\gamma)/(1-\gamma)\neq0$，因為 $\beta+\gamma\neq1$。所以

$$\frac{AF}{FB}=\frac{f}{1-f}=\frac{\beta}{1-\beta-\gamma}.$$

乘積。中間的因子用第 (ii) 部分的結果，得

$$\frac{AF}{FB}\times\frac{BD}{DC}\times\frac{CE}{EA}=\frac{\beta}{1-\beta-\gamma}\times\frac{\gamma}{\beta}\times\frac{1-\beta-\gamma}{\gamma}=1,$$

其中每個被約去的因子都不是零，因為 $\beta\neq0$、$\gamma\neq0$ 及 $\beta+\gamma\neq1$。這就是所需的結果。（它對有向比成立，因此對長度本身也成立。）

驗算

- 形心（centroid）：取 $\beta=\gamma=\frac{1}{3}$，P 便是形心。這時 $BD:DC=1:1$、$CE/EA=\frac{1/3}{1/3}=1$ 及 $AF/FB=1$，所以三條塞瓦線（cevian）都是中線（median），正如所料。
- 取 $A=(0,0)$、$B=(4,0)$、$C=(1.5,3)$、$\beta=0.3$、$\gamma=0.4$（即圖中的情況），我們得到 $D=(2.571,1.714)$、$E=\frac{4}{7}\mathbf{c}$ 及 $F=(2,0)$，從而 $\frac{AF}{FB}=1$、$\frac{BD}{DC}=\frac{4}{3}$、$\frac{CE}{EA}=\frac{3}{4}$，乘積為 1。
- 我們以超過 1700 個隨機三角形和隨機的 β,γ（包括 P 在三角形之外的情況），用分數精確地核對了 D、E、F 的公式、全部三個有向比以及乘積。

常見錯誤

- 把 (i) 中的軌跡說成線段 BC：s 和 t 並沒有限定為正數，所以軌跡是整條直線。

- 把比倒轉了：$\overrightarrow{BD}$ 與 $\mathbf{c}$ 的係數 γ 成正比，所以 γ 大的時候，D 會被拉向 C。
- 忘記了沿 BP 和 CP 的參數是從 B 和 C 起計，而不是從原點起計。

STEP 之外

第 (iii) 部分就是塞瓦定理（Ceva's theorem），而 $(1-\beta-\gamma, \beta, \gamma)$ 是 P 相對於三角形 ABC 的重心坐標（barycentric coordinates）：它們與三角形 PBC、PCA、PAB 的面積成正比，而那三個比不過是這些面積之間的商。重心坐標是仿射幾何（affine geometry）最自然的框架：在仿射幾何中，直線以及沿直線的比都有意義，但長度和角度則沒有；重心坐標亦會在電腦圖像（computer graphics）和有限元法（finite-element method）中再次出現。Strang 的 *Introduction to Linear Algebra*（線性代數入門）有系統地處理相對於一組基（basis）的坐標，是很好的入門讀物。

詳細解答

STEP II 2007，第 9 題

力學 • 斜碰撞與恢復係數

STEP II 2007，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/07-s2-q9

關鍵思路

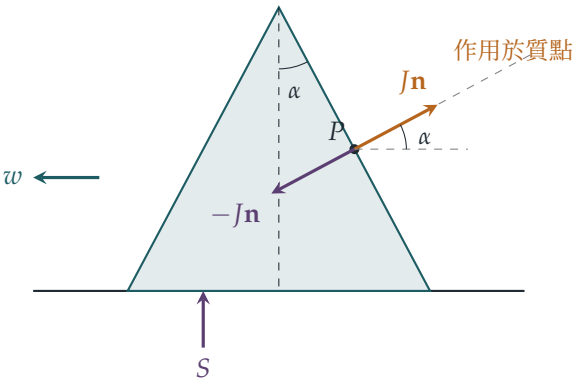
所有接觸面都是光滑的（smooth），所以質點與圓錐之間唯一的衝量（impulse）沿碰撞點的公法線（common normal）方向作用，而桌面只能沿鉛垂方向施力。由此我們同時得知兩件事：質點沿錐面方向的速度不變，而質點與圓錐合起來的水平動量（momentum）守恆。把這些與沿法線方向的牛頓恢復定律（Newton's law of restitution）結合起來，便可求出圓錐的速率。第 (ii) 部分其實就是把第 (i) 部分應用於下落質點速度的法向分量。

解答

把質點視為一點，把圓錐視為剛體（rigid body）；按題目所述，圓錐既不轉動，亦不作鉛垂方向的運動，所以剛碰撞後它以某速率 w 沿水平方向滑動。我們在通過圓錐的軸和碰撞點 P 的鉛垂平面內考慮。由於通過 P 的母線（generator）與鉛垂的軸成 α 角，在 P 的向外法線與水平方向成 α 角。取單位向量

$$\mathbf{n} = (\cos\alpha,\ \sin\alpha)\text{（向外法線）}, \qquad \mathbf{t} = (\sin\alpha,\ -\cos\alpha)\text{（沿斜面向下）},$$

其中第一個坐標為水平方向（遠離軸），第二個坐標為鉛垂向上。



設作用於質點的衝量為 $J\mathbf{n}$（$J > 0$），則圓錐所受的衝量為 $-J\mathbf{n}$，其水平部分的大小為 $J\cos\alpha$，方向遠離 P。桌面提供向上的鉛垂衝量 $S = J\sin\alpha$，剛好抵消 $-J\mathbf{n}$ 的向下部分；這就是圓錐沒有鉛垂速度的原因。所以

$$Mw = J\cos\alpha, \tag{1}$$

而圓錐沿水平方向遠離 P 運動。

第 (i) 部分

質點到達時的速度為 $-u\mathbf{n}$，它沒有沿 $\mathbf{t}$ 方向的分量。由於接觸面光滑，作用於質點的衝量完全沿 $\mathbf{n}$ 方向，不能改變質點沿 $\mathbf{t}$ 方向的分量。因此碰撞後該分量仍然是零，質點的速度是 $\mathbf{n}$ 的倍數：它沿 P 處的法線運動（向外運動；又或者，若圓錐退得夠快，則仍然向內運動，只是較慢）。

碰撞後質點的速度為 $(-u + J/m)\,\mathbf{n}$。圓錐以速率 w 沿水平方向遠離 P 運動（即圖中向左），所以它的速度是 $(-w, 0)$，其沿 $\mathbf{n}$ 方向的分量為 $-w\cos\alpha$。

> **想一想** 牛頓恢復定律比較的是沿哪個方向的相對速度？圓錐在當中有甚麼貢獻？

恢復定律沿公法線 $\mathbf{n}$ 方向起作用。碰撞前，沿 $\mathbf{n}$ 方向的接近速率為 u（圓錐靜止）。碰撞後的分離速率，等於質點的 $\mathbf{n}$ 分量減去圓錐的相應分量：

$$\left(-u + \frac{J}{m}\right) - (-w\cos\alpha) = eu \quad\Longrightarrow\quad J = m\big((1+e)u - w\cos\alpha\big).$$

代入 (1)，得

$$Mw = m\cos\alpha\,\big((1+e)u - w\cos\alpha\big) \quad\Longrightarrow\quad \boxed{w = \frac{m(1+e)u\cos\alpha}{M + m\cos^2\alpha}}.$$

換一個說法，水平動量守恆（桌面的衝量是鉛垂的）：$m(-u\cos\alpha) = m\left(-u + \frac{J}{m}\right)\cos\alpha - Mw$，這又是 (1)。

第 (ii) 部分

現在質點到達時的速度為 $(0,-u)$。把它分解：

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

切向部分 $u\cos\alpha$ 不受碰撞影響（同樣因為接觸面光滑），之後亦再無作用：衝量 $J\mathbf{n}$、方程 (1) 和恢復定律的方程都只涉及法向分量。所以第 (i) 部分的推導完全適用，只需把法向接近速率 u 換成 $u\sin\alpha$：

$$w = \frac{m(1+e)(u\sin\alpha)\cos\alpha}{M+m\cos^2\alpha} = \frac{mu(1+e)\sin\alpha\cos\alpha}{M+m\cos^2\alpha},$$

正是所求。

對 α 求最大值。固定 M, m, e, u，我們需要在 $0<\alpha<\frac{\pi}{2}$ 的範圍內求 $f(\alpha)=\dfrac{\sin\alpha\cos\alpha}{M+m\cos^2\alpha}$ 的最大值。

? **想一想** 直接對這個商求導相當繁複。有沒有一個代換可以令它變得簡潔？

把分子和分母同除以 $\cos^2\alpha$，並設 $T=\tan\alpha>0$，利用 $\sec^2\alpha = 1+T^2$：

$$f = \frac{\tan\alpha}{M\sec^2\alpha+m} = \frac{T}{MT^2+(M+m)} = \frac{1}{MT+\dfrac{M+m}{T}}.$$

根據算術—幾何平均不等式（AM–GM inequality）（兩項皆為正數），

$$MT+\frac{M+m}{T} \geq 2\sqrt{M(M+m)},$$

等號成立當且僅當 $MT=(M+m)/T$，即 $T^2=(M+m)/M$。所以 f 在該處取得最大值，w 亦然。由於 $T\mapsto\alpha$ 是由 $(0,\infty)$ 到 $(0,\frac{\pi}{2})$ 的雙射（bijection），這是唯一的最大值，而且

$$\cos^2\alpha = \frac{1}{1+\tan^2\alpha} = \frac{1}{1+\frac{M+m}{M}} = \frac{M}{2M+m}, \qquad \text{所以} \qquad \cos\alpha = \sqrt{\frac{M}{2M+m}}$$

（因為 α 是銳角，故取正根）。最大速率為 $w_{\max} = \dfrac{mu(1+e)}{2\sqrt{M(M+m)}}$。

驗算

- 若 $\alpha\to\frac{\pi}{2}$，圓錐變成一塊平的圓盤：鉛垂下落不會產生水平推力，故 $w\to 0$。在第 (i) 部分中，當 $\alpha\to\frac{\pi}{2}$ 時法線是鉛垂的，同樣有 $w\to 0$。吻合。

- 在第 (i) 部分中，當 $\alpha \to 0$（一面鉛垂的牆）時，我們得到 $w = \dfrac{m(1+e)u}{M+m}$，這正是與一個質量為 M 的自由方塊正碰時大家熟悉的答案。
- 若 $m \ll M$，最佳角度滿足 $\cos^2\alpha \to \frac{1}{2}$，即 $\alpha = 45°$，這正是 $\sin\alpha\cos\alpha$ 取最大值之處。
- 數值上，取 $M = 3, m = 1, e = \frac{1}{2}, u = 4, \alpha = \frac{\pi}{6}$，第 (ii) 部分給出 $w = \frac{2\sqrt{3}}{5} \approx 0.693$。以隨機選取的參數直接計算衝量（由恢復定律解出 J），結果與兩條公式都一致；對 α 作網格搜尋，亦確認了取得最大值的角度。

常見錯誤

- 忘記圓錐會移動，寫恢復定律的方程時，把分離速率當成只是質點的速率 eu。圓錐的速度沿法線方向有分量 $w\cos\alpha$。
- 沿法線方向應用動量守恆。桌面的衝量沿 $\mathbf{n}$ 方向有分量，所以只有水平動量守恆。
- 以為法線與鉛垂方向成 α 角。與鉛垂方向成 α 角的是錐面，所以法線是與水平方向成 α 角。

STEP 之外

這裏的桌面起着約束（constraint）的作用：它令圓錐失去一個自由度（degree of freedom），而它的衝量則是為了維持這個約束所需的任何值。到了大學，這會發展成衝擊約束（impulsive constraints）的理論，在拉格朗日力學（Lagrangian mechanics）中以廣義坐標（generalised coordinates）處理十分簡潔；而向法線方向投影的做法，與受約束物體的有效質量（effective mass）背後是同一個想法。Kleppner 與 Kolenkow 的 *An Introduction to Mechanics*（力學入門教科書）對二維的碰撞和衝量有仔細的論述。

詳細解答

STEP II 2007，第 10 題

力學 • 涉及摩擦力的靜力學

STEP II 2007，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/07-s2-q10

關鍵思路

這是一道靜力學（statics）問題，唯一棘手的地方是質心（centre of mass）的位置。一旦知道 G 在軸上的位置，對接觸點取力矩（taking moments）便可同時消去法向反作用力（normal reaction）和摩擦力（friction），直接求出 P。然後沿水平方向分解力，可知摩擦力的大小等於 P，餘下的交給摩擦定律 $|F| \leq \mu R$ 便可。式中出現絕對值（modulus），是因為視乎 α 的大小，P 可能要推，也可能要拉。

解答

模型

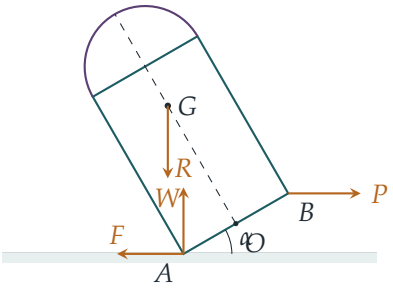
我們把這個立體視為剛體（rigid body），它與平面只在一點 A 接觸，這一點是平底面（即圓柱體不與半球相連的一端）邊緣上的最低點。設 B 為該底面的最高點，則 AB 是一條長度為 $2r$ 的直徑，與水平方向成 α 角。作用力有：作用於質心 G 的重量 W；作用於 A 的法向反作用力 R 和摩擦力 F；以及作用於 B 的水平力 P。

質心

沿軸由底面的圓心 O 起量度距離。圓柱體的體積為 $3\pi r^3$，故質量為 $3\pi r^3\rho$，其質心位於 $\frac{3}{2}r$ 處。半球的體積為 $\frac{2}{3}\pi r^3$，密度為 3ρ，故質量為 $2\pi r^3\rho$；它的質心在其平面之外 $\frac{3}{8}r$ 處，而該平面位於 $3r$ 處，所以質心位於 $\frac{27}{8}r$ 處。總質量為 $5\pi r^3\rho$，而

$$OG = \frac{3\cdot\frac{3}{2}r + 2\cdot\frac{27}{8}r}{5} = \frac{\frac{9}{2}r + \frac{27}{4}r}{5} = \frac{9}{4}r.$$

因此 G 在軸上，距離底面 $\frac{9}{4}r$，並且 $W = 5\pi r^3\rho g$。



平衡

取 A 為原點，x 軸沿水平方向（由 A 指向 B 的垂足），y 軸鉛垂向上。設 P 在 x 方向的水平分量為 X，故 X 可正可負；同樣地，設摩擦力在 x 方向為 F。

底面由 A 出發，沿方向 $(\cos\alpha, \sin\alpha)$ 伸展，所以 $B = (2r\cos\alpha,\, 2r\sin\alpha)$ 而 $O = (r\cos\alpha,\, r\sin\alpha)$。軸垂直於底面，指向立體內部，方向為 $(-\sin\alpha, \cos\alpha)$。因此 G 的水平坐標為

$$x_G = r\cos\alpha - \tfrac{9}{4}r\sin\alpha.$$

沿鉛垂方向分解，得 $R = W$。沿水平方向分解，得 $F + X = 0$，故 $|F| = |X|$。

想一想 我們應該對哪一點取力矩？

對 A 取力矩：這樣 R 和 F 都沒有力矩，只剩下 W 和 P。重量沿鉛垂方向作用並通過 G，所以它對 A 的力矩為 Wx_G（當 $x_G > 0$ 時為順時針方向）。作用於 B 的水平力位於高度 $2r\sin\alpha$ 處，所以它的力矩為 $X \cdot 2r\sin\alpha$（當 $X < 0$ 時為逆時針方向，當 $X > 0$ 時為順時針方向）。取順時針方向為正，

$$Wx_G + 2r\sin\alpha\, X = 0 \quad\Longrightarrow\quad X = -\frac{W\left(r\cos\alpha - \frac{9}{4}r\sin\alpha\right)}{2r\sin\alpha} = W\left(\tfrac{9}{8} - \tfrac{1}{2}\cot\alpha\right),$$

其中我們除以了 $\sin\alpha$，由於 $0 < \alpha < \frac{1}{2}\pi$，它是正數。

摩擦條件

由於平面是粗糙的，而立體在 A 處不滑動，我們需要 $|F| \le \mu R$。利用 $|F| = |X|$ 和 $R = W > 0$，

$$W\left|\tfrac{9}{8} - \tfrac{1}{2}\cot\alpha\right| \le \mu W \quad\Longrightarrow\quad \mu \ge \left|\tfrac{9}{8} - \tfrac{1}{2}\cot\alpha\right|,$$

這正是所要求的結果。

X 的正負號告訴我們 P 的作用方向。若 $\tan\alpha > \frac{4}{9}$，則 G 越過了 A，位於遠離 B 的一側，立體本身會向後翻倒，因此 P 必須把 B 向外拉離立體；若 $\tan\alpha < \frac{4}{9}$，則 P 必須向內推。無論哪種情況，摩擦力都與 P 方向相反，而絕對值正好涵蓋這兩種情況。

驗算

- 若 $\tan\alpha = \frac{4}{9}$，則 $x_G = 0$：G 在 A 的正上方，立體自行平衡，$P = 0$，無需摩擦力。所得的界給出 $\mu \geq |\frac{9}{8} - \frac{9}{8}| = 0$，與預期相符。
- 當 $\alpha = 45°$ 時，所得的界為 $\mu \geq \frac{5}{8}$。
- 以數值方法驗證：把立體切成薄圓片，得總質量為 $5\pi\rho r^3$ 及 $OG = 2.2500r$；再對 10 000 個隨機的 α 值直接解三條平衡方程，每次所得的 $|F|/R$ 都等於 $|\frac{9}{8} - \frac{1}{2}\cot\alpha|$。

常見錯誤

- 忘記半球的密度是三倍，因而把它的質量當作 $\frac{2}{3}\pi r^3\rho$；這會令 G 的位置改變，常數 $\frac{9}{8}$ 也隨之改變。
- 把半球質心的 $\frac{3}{8}r$ 由圓柱體的底面量起，而不是由半球本身位於高度 $3r$ 的平面量起。
- 預先假定 P（或摩擦力）的方向，因而漏掉絕對值：無論 P 向哪一方作用，結果都必須成立。

STEP 之外

求組合體的質心，其實是體積分（volume integral）的離散版本：$\bar{\mathbf{r}} = \int \rho\,\mathbf{r}\,\mathrm{d}V \big/ \int \rho\,\mathrm{d}V$，而半球的 $\frac{3}{8}r$ 本來就是這樣推導出來的。在大學一年級的力學課程中，這會發展成剛體理論，其中把同樣的積分裏的 $\mathbf{r}$ 換成 r^2，便得到轉動慣量（moments of inertia）。Kleppner 與 Kolenkow 合著的 *An Introduction to Mechanics*（力學導論）對這兩個概念都有細緻的闡述。

詳細解答

STEP II 2007，第 11 題

力學 • 三維空間中的拋體

> **STEP II 2007，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/07-s2-q11

關鍵思路

把整個運動寫成一個位置向量（position vector）：水平速度保持不變，只有垂直分量受重力影響。一旦把 $\mathbf{r}(t)$ 寫在紙上，與 O 的距離就是 $|\mathbf{r}|$，而最短距離只需求一個二次式的最小值便可得到。最後一部分之所以行得通，是因為子彈極快，在它飛行期間質點幾乎沒有移動。

解答

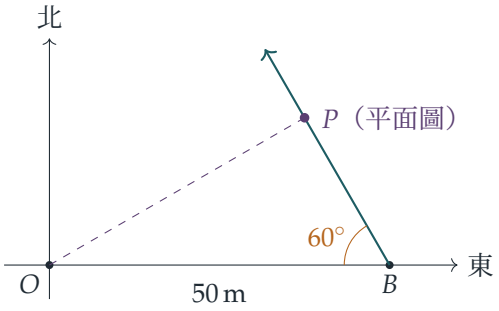
我們把質點（particle）視為只在重力作用下運動的一點（$g = 10\,\mathrm{m\,s^{-2}}$，不計空氣阻力），並取 O 為原點，於是 B 的位置向量為 $50\mathbf{i}$。

第 (i) 部分

初速度的分量。若 $\tan\alpha = \frac{1}{2}$，則 $\cos\alpha = \frac{2}{\sqrt{5}}$ 及 $\sin\alpha = \frac{1}{\sqrt{5}}$（邊長為 1、2、$\sqrt{5}$ 的三角形）。所以初速度的

$$\text{水平速率為} 25\cdot\frac{2}{\sqrt{5}} = 10\sqrt{5}, \qquad \text{垂直速率為} 25\cdot\frac{1}{\sqrt{5}} = 5\sqrt{5}.$$

水平方向是哪一邊？該方向與直線 OB 成 60° 角。假如質點大致朝東、遠離 O 而去，它根本不可能經過 O 附近；要從 O 的北面經過，它必須折返向西並偏向北方。所以水平方向與 $\overrightarrow{BO}$（向西）成 60° 角，並在其北側，由此得單位向量（unit vector）$-\frac{1}{2}\mathbf{i} + \frac{\sqrt{3}}{2}\mathbf{j}$。



因此水平速度為

$$10\sqrt{5}\left(-\tfrac{1}{2}\mathbf{i}+\tfrac{\sqrt{3}}{2}\mathbf{j}\right)=-5\sqrt{5}\,\mathbf{i}+5\sqrt{15}\,\mathbf{j},$$

而且它保持不變，因為重力只沿垂直方向作用。垂直方向上，$z=5\sqrt{5}\,t-\frac{1}{2}(10)t^2$。加上起點 $50\mathbf{i}$，得

$$\mathbf{r}(t)=\left(50-5\sqrt{5}\,t\right)\mathbf{i}+5\sqrt{15}\,t\,\mathbf{j}+\left(5\sqrt{5}\,t-5t^2\right)\mathbf{k}.$$

與 O 的距離。把各分量平方後相加，

$$\begin{aligned}|\mathbf{r}|^2&=\left(2500-500\sqrt{5}\,t+125t^2\right)+375t^2+\left(125t^2-50\sqrt{5}\,t^3+25t^4\right)\\&=25t^4-50\sqrt{5}\,t^3+625t^2-500\sqrt{5}\,t+2500.\end{aligned}$$

把它與下式比較：

$$\begin{aligned}25\left(t^2-\sqrt{5}\,t+10\right)^2&=25\left(t^4+5t^2+100-2\sqrt{5}\,t^3+20t^2-20\sqrt{5}\,t\right)\\&=25t^4-50\sqrt{5}\,t^3+625t^2-500\sqrt{5}\,t+2500,\end{aligned}$$

兩者相同。所以 $|\mathbf{r}|=5\left|t^2-\sqrt{5}\,t+10\right|$。二次式 $t^2-\sqrt{5}\,t+10$ 的判別式（discriminant）為 $5-40<0$，故它恆為正數，絕對值符號可以除去：

$$|\mathbf{r}|=5\left(t^2-\sqrt{5}\,t+10\right)\text{ m},$$

正是所求。

點 P。配方（completing the square）得

$$t^2-\sqrt{5}\,t+10=\left(t-\tfrac{\sqrt{5}}{2}\right)^2+\tfrac{35}{4},$$

所以當 $t=\frac{\sqrt{5}}{2}$ 時距離最小，而最短距離為 $5\cdot\frac{35}{4}=\frac{175}{4}=43.75\,\text{m}$。代入 $t=\frac{\sqrt{5}}{2}$，此時 $\sqrt{5}\,t=\frac{5}{2}$ 及 $t^2=\frac{5}{4}$：

$$\overrightarrow{OP}=\tfrac{75}{2}\,\mathbf{i}+\tfrac{25\sqrt{3}}{2}\,\mathbf{j}+\tfrac{25}{4}\,\mathbf{k}\ \text{（單位為米）}.$$

$\overrightarrow{OP}$ 的水平部分，向東分量為 $\frac{75}{2}$，向北分量為 $\frac{25\sqrt{3}}{2}$。方位角（bearing）θ 是由正北起順時針量度的，所以

$$\tan\theta = \frac{\text{東}}{\text{北}} = \frac{75/2}{25\sqrt{3}/2} = \sqrt{3},$$

又因為兩個分量均為正數，故 $\theta = 60°$：由 O 測得 P 的方位角為 060°。

第 (ii) 部分

速度的垂直分量為

$$\dot{z} = 5\sqrt{5} - 10t,$$

它恰好在 $t = \frac{\sqrt{5}}{2}$ 時為零。這正是質點位於 P 的時刻，所以質點在 P 處達到最高點（高度為 $\frac{25}{4}$ m）。

第 (iii) 部分

子彈以 $350\,\mathrm{m\,s^{-1}}$ 沿直線由 O 飛往 P（我們忽略重力對它的影響），所走的距離為 $OP = \frac{175}{4}$ m。所需時間為

$$T = \frac{175/4}{350} = \frac{1}{8}\,\mathrm{s}.$$

> **想一想** 多過了這 $\frac{1}{8}$ 秒之後，質點在哪裏？

由第 (ii) 部分，質點在 P 處的垂直速度為零，所以它在該處的速度純粹是水平的，大小為 $10\sqrt{5}\,\mathrm{m\,s^{-1}}$。在其後的 T 秒內，它水平移動了 $10\sqrt{5}\,T$，並下墜了 $\frac{1}{2}gT^2$。取 $T = \frac{1}{8}$，

$$\text{水平位移} = \frac{10\sqrt{5}}{8} \approx 2.80\,\mathrm{m}, \qquad \text{下墜距離} = 5 \cdot \frac{1}{64} \approx 0.08\,\mathrm{m}.$$

這兩段位移互相垂直，所以質點與 P 的距離為

$$\sqrt{\frac{500}{64} + \frac{25}{4096}} = \frac{5\sqrt{1281}}{64} \approx 2.80\,\mathrm{m},$$

即約為 3 m，正是所求。（下墜的影響微乎其微：答案基本上就是 $10\sqrt{5}/8$。）神槍手射失了。

驗算

- 當 $t = 0$ 時，公式給出 $5 \cdot 10 = 50\,\mathrm{m}$，即距離 OB。吻合。
- $|\overrightarrow{OP}|^2 = 37.5^2 + 21.65^2 + 6.25^2 = 1406.25 + 468.75 + 39.06 = 1914.06 = 43.75^2$，所以 P 的位置向量與最短距離一致。

- 以數值方法掃描 $|\mathbf{r}(t)|$，得最小值 43.75 m，出現於 $t \approx 1.118 = \frac{\sqrt{5}}{2}$；再讓質點由 P 繼續運動 0.125 s，得相距 2.796 m。

常見錯誤

- 把水平方向取為 $\frac{1}{2}\mathbf{i} + \frac{\sqrt{3}}{2}\mathbf{j}$（遠離 O）。這樣質點永遠不會從 O 的北面經過，距離也算不出題目所給的式子。
- 把方位角寫成 30°：那是由正東量起的角，但方位角是由正北起順時針量度的。

STEP 之外

第 (iii) 部分是攔截問題（interception problem）的入門版本：要擊中移動的目標，必須瞄準它將會到達的位置，這就要解 $|\mathbf{r}_{\text{目標}}(t) - \mathbf{r}_{\text{子彈}}(t)| = 0$，同時求出發射方向和時間。求 $|\mathbf{r}(t)|$ 的最小值，也正是軌道問題和碰撞問題中「最接近距離」（distance of closest approach）背後的想法；在這類問題中，把距離平方以得到多項式，是標準的第一步。Kleppner 與 Kolenkow 合著的 *An Introduction to Mechanics*（力學入門）對這類向量運動學（vector kinematics）有嚴謹而完整的闡述。

詳細解答

STEP II 2007，第 12 題

概率與統計 • 幾何分佈的最大值

STEP II 2007，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/07-s2-q12

關鍵思路

先撇開題目對程序那段頗為繁瑣的描述，逐顆骰子單獨來看。每顆骰子都一直擲到首次出現六點為止，而整個實驗在兩顆骰子中較慢的那一顆擲出六點時結束。因此，擲骰次數就是兩個獨立（independent）幾何隨機變數（geometric random variable）的最大值，而處理最大值最方便的途徑，是考慮累積概率（cumulative probability）$\mathrm{P}(\max \le r)$。

解答

第 (i) 部分

設 X 為第一顆骰子首次出現六點時的擲骰次序，並以同樣方式為第二顆骰子定義 Y。骰子的每一次投擲都與其他各次投擲獨立，所以 X 與 Y 獨立，而且兩者都服從幾何分佈：

$$\mathrm{P}(X = k) = q^{k-1}p, \qquad \mathrm{P}(X \le r) = 1 - \mathrm{P}(\text{首 } r \text{ 次投擲全非六點}) = 1 - q^r.$$

題目的程序怎樣與此吻合？當兩顆骰子都未出現過六點時，我們兩顆一起擲；一旦其中一顆出現了六點，便只繼續擲另一顆。無論哪種情況，每顆骰子都恰好擲到它首次出現六點為止（之後不再擲），而當兩顆都出現過六點，我們便立即停止。如果兩顆骰子在同一次投擲中首次出現六點，那就是兩顆同時擲出六點的情況。因此擲骰次數為

$$N = \max(X, Y).$$

? **想一想** 為甚麼要處理 $\mathrm{P}(N \le r)$，而不直接處理 $\mathrm{P}(N = r)$？

事件 $N \le r$ 就是「$X \le r$ 而且 $Y \le r$」，而獨立性可把它化為乘積。事件 $N = r$ 則要分成幾種麻煩的情況（一顆骰子在第 r 次完成而另一顆較早完成，或兩顆都在第 r 次完成）。

所以

$$\mathrm{P}(N \le r) = \mathrm{P}(X \le r)\,\mathrm{P}(Y \le r) = (1-q^r)^2,$$

而對於 $r \ge 1$，

$$\begin{aligned}\mathrm{P}(N = r) &= \mathrm{P}(N \le r) - \mathrm{P}(N \le r-1)\\ &= (1-q^r)^2 - (1-q^{r-1})^2\\ &= \left[(1-q^r) - (1-q^{r-1})\right]\left[(1-q^r) + (1-q^{r-1})\right]\\ &= (q^{r-1}-q^r)(2-q^{r-1}-q^r)\\ &= pq^{r-1}\,(2-q^{r-1}-q^r),\end{aligned}$$

當中先用了平方差(difference of two squares)，再用了 $q^{r-1}-q^r = q^{r-1}(1-q) = pq^{r-1}$。這就是所要求的結果。

至於期望值（expectation），把這個概率拆成兩部分：

$$\mathrm{P}(N = r) = 2pq^{r-1} - pq^{2r-2}(1+q).$$

把題目給出的結果求導（或把它除以 x），可得對於 $|x| < 1$ 有 $\sum_{r\ge1} rx^{r-1} = (1-x)^{-2}$。分別取 $x = q$ 和 $x = q^2$ 代入（當 $0 < p \le 1$ 時兩者都小於 1）：

$$\begin{aligned}\mathrm{E}(N) &= 2p\sum_{r=1}^{\infty} rq^{r-1} - p(1+q)\sum_{r=1}^{\infty} r(q^2)^{r-1}\\ &= \frac{2p}{(1-q)^2} - \frac{p(1+q)}{(1-q^2)^2}\\ &= \frac{2}{p} - \frac{p(1+q)}{p^2(1+q)^2}\\ &= \frac{2}{p} - \frac{1}{p(1+q)}\\ &= \frac{1+2q}{p(1+q)},\end{aligned}$$

其中用了 $1-q = p$ 和 $1-q^2 = (1-q)(1+q) = p(1+q)$。若只以 p 表示（代入 $q = 1-p$），

$$\mathrm{E}(N) = \frac{3-2p}{p(2-p)}.$$

第 (ii) 部分

進行大量實驗後，擲骰次數的平均值應接近 $\mathrm{E}(N)$，所以我們解 $\mathrm{E}(N) = m$ 來估計 p：

$$m\,p(2-p) = 3-2p \quad\Longleftrightarrow\quad mp^2 - 2(m+1)p + 3 = 0,$$

所以

$$p = \frac{(m+1) \pm \sqrt{(m+1)^2 - 3m}}{m} = \frac{m+1 \pm \sqrt{m^2-m+1}}{m}.$$

想一想 我們要取哪一個根？

每次實驗最少要擲一次，所以 $m \geq 1$。於是 $\sqrt{m^2-m+1} \geq 1$（因為 $m^2-m = m(m-1) \geq 0$），所以取正號的根至少是 $(m+2)/m > 1$，這對概率來說是不可能的。根據同一個不等式，取負號的根至多是 $m/m = 1$，而且它是正數，因為 $(m+1)^2 > m^2-m+1$。所以估計值為

$$\hat{p} = \frac{m+1-\sqrt{m^2-m+1}}{m}.$$

驗算

- 若 $p = 1$，每次實驗都只需擲一次：$\mathrm{E}(N) = \frac{1}{1\cdot 1} = 1$，而 $m = 1$ 給出 $\hat{p} = (2-1)/1 = 1$。
- 各概率之和為 1：$\sum_r \mathrm{P}(N = r)$ 逐項相消（telescope）後得 $\lim_{r\to\infty}(1-q^r)^2 = 1$。
- 對於公平的骰子，$p = \frac{1}{6}$ 給出 $\mathrm{E}(N) = \frac{8/3}{11/36} = \frac{96}{11} \approx 8.73$，比一顆骰子平均所需的 6 次稍多，這正合預期。一次包含 200,000 次實驗的模擬得出平均值 8.71，而把 $m = 96/11$ 代入 $\hat{p}$ 的公式，則得回 $\frac{1}{6}$。

常見錯誤

- 把每顆骰子的投擲分開計算：兩顆骰子同時擲一次只算作一次投擲，這正是 $N = \max(X, Y)$ 的原因。
- 忘記 $pq^{2r-2}(1+q)$ 這一項所需的求和是取 $x = q^2$，而不是 $x = q$。
- 在第 (ii) 部分保留兩個根，而沒有檢查哪一個才可以是概率。

STEP 之外

$\mathrm{P}(\max \leq r) = \prod \mathrm{P}(X_i \leq r)$ 這個技巧，是處理獨立變數的最大值（把符號換成 $\geq$ 則可處理最小值）的標準方法，也是順序統計量（order statistics）的起點。第 (ii) 部分用的是矩估計法（method of moments）：令樣本平均值等於理論平均值，再解出參數。大學一年級的統計學會把它與最大似然法（maximum likelihood）比較，並探討這類估計量（estimator）有多準確。至於概率方面，Grimmett 與 Welsh 的 *Probability: An Introduction*（概率論入門教材）是很好的下一步。

詳細解答

STEP II 2007，第 13 題

概率與統計 • 生日問題

> **STEP II 2007，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/07-s2-q13

本題在 *Beyond STEP* 第 9.9 節生日問題中有完整的引導式講解：https://beyondstep.keithmcnulty.org/zh-hant/ch09/birthday.html

關鍵思路

沒有任何兩位賓客生日相同的概率（probability），是若干個形如 $\frac{365-r}{365}$ 的因子的乘積，而乘積並不好處理。第一部分把每個因子化成指數函數（exponential），而指數函數的乘積等於以和為指數的指數函數，這個和可以準確地求出。至於主人的問題，各位賓客的事件互相獨立（independent），所以互補事件（complement）的概率只是單獨一個冪，同一個近似方法照樣適用。

解答

近似式

對任意實數 x，

$$\mathrm{e}^{-x} = 1 - x + \frac{x^2}{2!} - \frac{x^3}{3!} + \cdots.$$

設 $x = r/n$。由於 r 遠小於 n，x 很小，x^2 及其後各項與 x 相比可以忽略不計。所以

$$\mathrm{e}^{-r/n} \approx 1 - \frac{r}{n} = \frac{n-r}{n},$$

這正是所求。

最少兩位賓客生日相同

設 A 為 k 位賓客中最少有兩位生日相同的事件。求 $\mathrm{P}(A')$，即 k 個生日全部不同的概率，會比較容易。逐一考慮各位賓客。第一位的生日可以是任何一天。已知首 r 位賓客有 r 個不同的生日，第 $(r+1)$ 位賓客必須避開這 r 天，其概率為 $\frac{365-r}{365}$，因為各人的生日互相獨立，並且在 365 天上均勻分佈（uniform）。把這些概率相乘，

$$\mathrm{P}(A') = \prod_{r=0}^{k-1} \frac{365-r}{365}.$$

若派對的人數不多，這裏每個 r 都遠小於 365，所以由第一部分，取 $n = 365$，

$$\begin{aligned}\mathrm{P}(A') &\approx \prod_{r=0}^{k-1} \mathrm{e}^{-r/365} = \exp\Big(-\frac{1}{365}\sum_{r=0}^{k-1} r\Big) \\ &= \exp\Big(-\frac{1}{365}\cdot\frac{k(k-1)}{2}\Big) = \mathrm{e}^{-k(k-1)/730}.\end{aligned}$$

（$r = 0$ 的因子恰好是 $1 = \mathrm{e}^0$，所以把它包括在內並無影響。）因此

$$\mathrm{P}(A) = 1 - \mathrm{P}(A') \approx 1 - \mathrm{e}^{-k(k-1)/730},$$

這就是所求的結果。

> **想一想** 一次過對這麼多個因子取近似值，難道不令人擔心嗎？誤差可能會累積起來。

每個因子的誤差屬 $(r/365)^2$ 的數量級，而指數本身則屬 $k^2/730$ 的數量級。對於我們所需的 k 值（20 左右），最大的 $r/365$ 約為 0.06，指數的總相對誤差只有百分之幾。當 $k = 23$ 時，近似式給出 0.500，而真確值為 0.507。

機會過半所需的最少人數

我們需要

$$1 - \mathrm{e}^{-k(k-1)/730} \geq \tfrac{1}{2} \iff \mathrm{e}^{-k(k-1)/730} \leq \tfrac{1}{2} \iff \frac{k(k-1)}{730} \geq \ln 2,$$

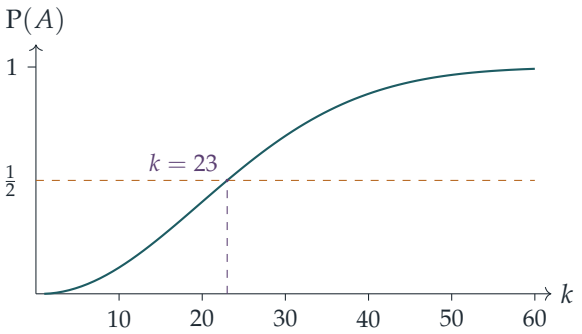
其中先取對數（對數函數是遞增的，所以不等號方向不變），再乘以 -1（不等號方向因而倒轉）。利用 $\ln 2 \approx \frac{253}{365}$，條件變成

$$k(k-1) \geq 730 \times \frac{253}{365} = 2 \times 253 = 506.$$

由於當 $k \geq 1$ 時 $k(k-1)$ 隨 k 遞增，而且

$$22 \times 21 = 462 < 506, \qquad 23 \times 22 = 506,$$

所以最小的值是 $k = 23$。



與主人生日相同

現在設 m 為賓客人數（主人不計算在內）。設 B 為最少有一位賓客與主人生日相同的事件。無論主人的生日是哪一天，每位賓客的生日不是這一天的概率都是 $\frac{364}{365}$，而且各人互相獨立，所以

$$\mathrm{P}(B') = \left(\frac{364}{365}\right)^m \approx \left(\mathrm{e}^{-1/365}\right)^m = \mathrm{e}^{-m/365},$$

這裏用了第一部分的結果，取 $n = 365$ 和 $r = 1$。這一次，無論 m 有多大，$r/n = 1/365$ 都確實很小。

我們需要 $\mathrm{P}(B) \geq \frac{1}{2}$，即 $\mathrm{e}^{-m/365} \leq \frac{1}{2}$。與之前完全一樣，這等價於

$$\frac{m}{365} \geq \ln 2 \approx \frac{253}{365} \iff m \geq 253.$$

所以最少要有 253 位賓客。

? **想一想** 為甚麼答案比 23 大這麼多？

在第一個問題中，任何一對賓客都可以生日相同，而 23 位賓客可組成 $\binom{23}{2} = 253$ 對。在第二個問題中，只有 m 對「賓客與主人」的配對才算數。要有同樣多的配對機會，便需要 $m = 253$，這正是 253 在兩個答案中都出現的原因。

驗算

- 準確計算乘積，可以確認兩個答案。有人生日相同的真確概率，22 位賓客時為 0.4757，23 位時為 0.5073。$1 - (364/365)^m$ 的真確值在 $m = 252$ 時小於 $\frac{1}{2}$，在 $m = 253$ 時為 0.5005。
- $\frac{253}{365} = 0.69315\ldots$，而 $\ln 2 = 0.69315\ldots$，兩者準確至小數點後五位都相同，所以題目給出的近似值不影響任何一個答案。
- 模擬 200 000 場各有 23 人的派對，結果有 50.7% 的派對出現生日相同的情況。

常見錯誤

- 計算了兩位指定賓客生日相同的概率，或者為每一對賓客各加一次 $\frac{1}{365}$。互相重疊的事件，其概率不能直接相加；改為考慮互補事件便可避開這個問題。
- 在第一個乘積中每個因子都用 $\frac{364}{365}$。第 $(r+1)$ 位賓客必須避開 r 天，而不只是一天。
- 在最後一部分把主人也算進賓客人數內。主人是另外一個人，所以需要 253 位賓客。

STEP 之外

這裏的經驗法則（heuristic）是：把 k 件物件放進 N 個盒子，一旦 k 達到大約 $\sqrt{2N\ln 2} \approx 1.18\sqrt{N}$，便很可能出現重複。密碼學（cryptography）中的「生日攻擊」（birthday attack）正是以此為基礎：要在一個 N 位元的雜湊（hash）中找出碰撞（collision），只需大約 $2^{N/2}$ 次嘗試，而不是 2^N 次。這個近似式的背後是泊松極限（Poisson limit）：生日相同的配對數目近似服從平均值為 $\binom{k}{2}/365$ 的泊松分佈，而 $\mathrm{P}(\text{沒有配對相同}) \approx \mathrm{e}^{-k(k-1)/730}$。Blitzstein 與 Hwang 合著的 *Introduction to Probability*（概率論入門教材）詳細討論了生日問題及其泊松近似。

詳細解答

STEP II 2007，第 14 題

概率與統計 • 分段概率密度函數

STEP II 2007，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/07-s2-q14

關鍵思路

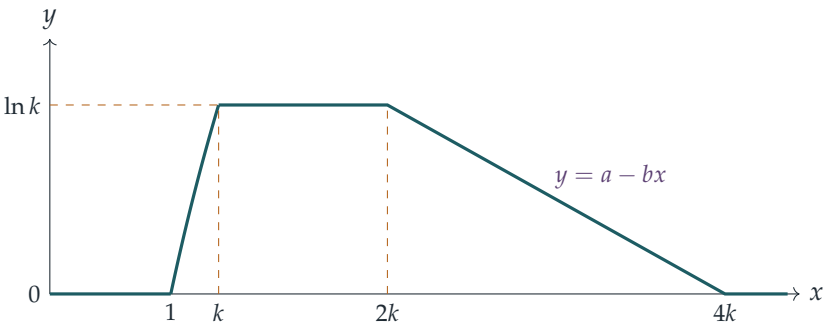
「連續」這兩個字舉足輕重。分段定義的密度函數（density function）在每個分界點都必須銜接起來，由此我們馬上得到關於 a 和 b 的兩條方程。餘下的一個條件，即圖像下的總面積為 1，便可確定 k。至於中位數（median），我們逐段把面積累加，直至超過二分之一為止。

解答

全題之中 $k > 1$：否則區間 $1 \le x \le k$ 便是空的，而且無論如何我們都需要 $\ln k > 0$，密度函數才會是非負，並且在 $x = 1$ 之後不恆等於零。

第 (i) 部分

在 $[1, k]$ 上，圖像是常見的對數曲線，在 $x = 1$ 時由 0 開始（因此與取值為零的一段連續地銜接），遞增而且是凹的（concave）。它在 $x = k$ 時到達高度 $\ln k$，保持在 $\ln k$ 的水平直至 $x = 2k$，然後沿直線下降，在 $x = 4k$ 時降至 0（下文會證明連續性恰好迫使如此），其後一直為零。草圖採用了我們在第 (ii) 部分求得的數值：$k = \mathrm{e}^{1/3} \approx 1.40$，因此平台的高度為 $\frac{1}{3}$。



第 (ii) 部分

連續性（*continuity*）。在 $x = k$ 處，兩條公式都給出 $\ln k$，所以這裏沒有新的資料。在 $x = 2k$ 和 $x = 4k$ 處，連續性要求

$$a - 2bk = \ln k, \qquad a - 4bk = 0.$$

兩式相減，得 $2bk = \ln k$，所以

$$b = \frac{\ln k}{2k}, \qquad a = 4bk = 2\ln k.$$

總概率。我們把面積分成三部分。用分部積分法（integration by parts）（取 $u = \ln x$，$\mathrm{d}v = \mathrm{d}x$），

$$\int_1^k \ln x \,\mathrm{d}x = \Big[x\ln x - x\Big]_1^k = k\ln k - k + 1.$$

平台是一個闊 k、高 $\ln k$ 的長方形，面積為 $k\ln k$。最後一部分是一個三角形，底為 $4k - 2k = 2k$，高為 $\ln k$，面積為 $\frac{1}{2}\cdot 2k\cdot \ln k = k\ln k$。所以

$$1 = (k\ln k - k + 1) + k\ln k + k\ln k = 3k\ln k - k + 1,$$

化簡後得

$$k(3\ln k - 1) = 0.$$

由於 $k > 1$，不可能有 $k = 0$，所以 $\ln k = \frac{1}{3}$，而

$$k = \mathrm{e}^{1/3} \approx 1.396, \qquad a = 2\ln k = \tfrac{2}{3}, \qquad b = \frac{\ln k}{2k} = \frac{1}{6\,\mathrm{e}^{1/3}} \approx 0.119.$$

留意 $b > 0$，所以最後一段確實如圖所示向下傾斜。

第 (iii) 部分

我們把概率逐段累加，找出中位數 m 落在哪一段。利用 $\ln k = \frac{1}{3}$，

$$\mathrm{P}(X \le k) = k\ln k - k + 1 = 1 - \tfrac{2k}{3} \approx 0.070,$$

而平台再加上 $k\ln k = \frac{k}{3}$，所以

$$\mathrm{P}(X \le 2k) = 1 - \tfrac{2k}{3} + \tfrac{k}{3} = 1 - \tfrac{k}{3} \approx 0.535.$$

想一想 中位數落在圖像的哪一段上？

由於 $0.070 < \frac{1}{2} < 0.535$，中位數落在平台上，即 $k \le m \le 2k$。在這一段，密度函數恆等於 $\frac{1}{3}$，所以

$$\mathrm{P}(X \le m) = 1 - \tfrac{2k}{3} + \tfrac{1}{3}(m-k) = \tfrac{1}{2}.$$

乘以 3：$3 - 2k + m - k = \frac{3}{2}$，所以

$$m = 3k - \tfrac{3}{2} = 3\mathrm{e}^{1/3} - \tfrac{3}{2} \approx 2.69.$$

這個值確實介乎 $k \approx 1.40$ 與 $2k \approx 2.79$ 之間，正如所應有的那樣。

驗算

- 取 $k = \mathrm{e}^{1/3}$、$a = \frac{2}{3}$、$b = 1/(6\mathrm{e}^{1/3})$，對密度函數作數值積分（中點法則），得總面積為 1.0000000，而 $\int_1^{2.68684} f(x)\,\mathrm{d}x = 0.5000000$。
- 中位數的合理性檢查：略多於一半的概率落在 $2k$ 的左邊（約 0.535），所以中位數應該在 $2k \approx 2.79$ 稍左的位置；果然 $m \approx 2.69$。

常見錯誤

- 忘記了 $\int_1^k \ln x\,\mathrm{d}x$ 的下限：$-[x\ln x - x]$ 在 $x = 1$ 處給出的 $+1$，正是令關於 k 的方程可以如此俐落地因式分解的原因。
- 因為中位數在支集（support）$[1, 4k]$ 的「中間」，便假設它落在傾斜的一段上；務必先把面積累加。

STEP 之外

本題的密度函數由幾段拼合而成，而各常數則由連續性和歸一化（normalisation）確定。到了大學，你會遇到*累積分佈函數*（cumulative distribution function）$F(x) = \mathrm{P}(X \le x)$，並以它作為最基本的對象：每一個隨機變量（random variable），不論是離散的還是連續的，都有累積分佈函數，而中位數這類量不過是 $F(m) = \frac{1}{2}$ 的解。Grimmett and Welsh, *Probability: An Introduction*（概率論入門教材）清楚地闡述了這個觀點。

詳細解答

STEP III 2007，第 1 題

純數學 • 正切和與韋達定理

> **STEP III 2007，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/07-s3-q1

關鍵思路

多個角之和的正切，是各個角的正切的對稱函數（symmetric function）之比；而多項式的根的對稱函數，正正就是韋達定理（Vieta's formulae）以係數表示出來的東西。至於最後一部分，竅門是把那條三角方程化為 $t = \tan\theta$ 的四次方程（quartic）；其係數的對稱性便會迫使角之和的正切等於零。

解答

四個角之和的正切

把和拆成 $A + B$，其中 $A = \theta_1 + \theta_2$，$B = \theta_3 + \theta_4$，於是

$$\tan A = \frac{t_1 + t_2}{1 - t_1 t_2}, \qquad \tan B = \frac{t_3 + t_4}{1 - t_3 t_4}.$$

然後，把 $\tan(A + B) = \dfrac{\tan A + \tan B}{1 - \tan A \tan B}$ 的分子和分母同乘以 $(1 - t_1 t_2)(1 - t_3 t_4)$，得

$$\tan(A + B) = \frac{(t_1 + t_2)(1 - t_3 t_4) + (t_3 + t_4)(1 - t_1 t_2)}{(1 - t_1 t_2)(1 - t_3 t_4) - (t_1 + t_2)(t_3 + t_4)}.$$

分子是 $t_1 + t_2 + t_3 + t_4 - (t_1 t_3 t_4 + t_2 t_3 t_4 + t_1 t_2 t_3 + t_1 t_2 t_4)$，而分母是 $1 - t_1 t_2 - t_3 t_4 + t_1 t_2 t_3 t_4 - (t_1 t_3 + t_1 t_4 + t_2 t_3 + t_2 t_4)$。以

$$\sigma_1 = \sum t_i, \quad \sigma_2 = \sum_{i<j} t_i t_j, \quad \sigma_3 = \sum_{i<j<k} t_i t_j t_k, \quad \sigma_4 = t_1 t_2 t_3 t_4$$

表示各個初等對稱函數（elementary symmetric functions），上式便是

$$\boxed{\tan(\theta_1 + \theta_2 + \theta_3 + \theta_4) = \frac{\sigma_1 - \sigma_3}{1 - \sigma_2 + \sigma_4}.}$$

（規律與兩個角或三個角的情況相同：奇數次的對稱函數正負號交替地放在分子，偶數次的則放在分母。）

以係數表示

若 $t_1, \ldots, t_4$ 是 $at^4 + bt^3 + ct^2 + dt + e = 0$ 的根，則 $at^4 + bt^3 + ct^2 + dt + e = a(t - t_1)(t - t_2)(t - t_3)(t - t_4)$，比較係數（韋達定理）得

$$\sigma_1 = -\frac{b}{a}, \qquad \sigma_2 = \frac{c}{a}, \qquad \sigma_3 = -\frac{d}{a}, \qquad \sigma_4 = \frac{e}{a}.$$

代入後，分子和分母同乘以 $a \neq 0$，得

$$\boxed{\tan(\theta_1 + \theta_2 + \theta_3 + \theta_4) = \frac{d - b}{a - c + e}.}$$

最後一部分

由於 $1 + \cos 2\theta = 2\cos^2\theta$，方程即為

$$2p\cos^2\theta + \cos\theta\cos\alpha + \sin\theta\sin\alpha = 0.$$

除以 $\cos^2\theta$（我們不理會 $\cos\theta = 0$ 這個特殊情況），得

$$2p = -(\cos\alpha + t\sin\alpha)\sec\theta, \qquad t = \tan\theta.$$

$\sec\theta$ 並不是 t 的有理函數（rational function），所以我們把兩邊平方，並利用 $\sec^2\theta = 1 + t^2$：

$$4p^2 = (\cos\alpha + t\sin\alpha)^2(1 + t^2).$$

展開後，

$$\sin^2\alpha\, t^4 + 2\sin\alpha\cos\alpha\, t^3 + t^2 + 2\sin\alpha\cos\alpha\, t + (\cos^2\alpha - 4p^2) = 0, \qquad (*)$$

其中 t^2 的係數用了 $\sin^2\alpha + \cos^2\alpha = 1$。

? **想一想** 平方可能會引入額外的解。各個 $\tan\theta_i$ 真的就是 $(*)$ 的四個根嗎？

平方之後，$p\cos 2\theta - \cos(\theta - \alpha) + p = 0$ 的解也同樣滿足 $(*)$。不過，把 θ 換成 $\theta + \pi$，$t = \tan\theta$ 和 $\cos^2\theta$ 都保持不變，而 $\cos(\theta - \alpha)$ 則會變號。因此，對於 $(*)$ 的每一個實根 t，在 $[0, 2\pi)$ 內正切等於 t 的兩個角之中，恰好有一個滿足原來的方程（不會兩個都滿足，因為那樣便需要 $\cos(\theta - \alpha) = 0$，從而 $p\cos^2\theta = 0$，屬於特殊情況）。所以，$[0, 2\pi)$ 內的解與 $(*)$ 的實根一一對應（one-to-one correspondence），而題目所給的四個角的正切，正是 $(*)$ 的四個根。

現在應用前面的結果，取 $a = \sin^2\alpha$，$b = d = 2\sin\alpha\cos\alpha$，$c = 1$，$e = \cos^2\alpha - 4p^2$（我們需要 $a \neq 0$，即 $\sin\alpha \neq 0$；否則 $(*)$ 便不是四次方程，這又是一個特殊情況）。分子是 $d - b = 0$，而分母是

$$a - c + e = \sin^2\alpha - 1 + \cos^2\alpha - 4p^2 = -4p^2,$$

它不等於零（若 $p = 0$，方程只是 $\cos(\theta - \alpha) = 0$，只有兩個解）。因此

$$\tan(\theta_1 + \theta_2 + \theta_3 + \theta_4) = 0,$$

所以對某個整數 n 有 $\theta_1 + \theta_2 + \theta_3 + \theta_4 = n\pi$，這正是所要證明的。

驗算

- 在四個角的公式中取 $t_3 = t_4 = 0$：公式便簡化為 $(t_1 + t_2)/(1 - t_1t_2)$，即我們熟悉的兩角公式。
- 取 $p = \frac{1}{2}$ 和 $\alpha = 3$，在 $[0, 2\pi)$ 上以數值方法解方程，得到四個根，它們的和除以 π 後是一個整數，誤差在 10^{-9} 以內。（並非每一組 p 和 α 都給出四個解：$p = \frac{1}{2}$、$\alpha = 1$ 便只有兩個解。）我們隨機試過的每一組 (p, α)，只要給出四個解，結果都一樣；我們亦用數以千計的隨機角和隨機四次方程驗證過兩條公式。

常見錯誤

- 韋達定理中的正負號出錯：相間的係數 b 和 d 是帶負號的。
- 試圖不經平方就直接以 $\tan\theta$ 表示 $\cos(\theta - \alpha)$，又或者平方之後沒有檢查是否混入了不想要的角。

STEP 之外

以初等對稱函數表示角之和的 tan，是一條重要定理的其中一個例子：根的每一個對稱多項式（symmetric polynomial），都是係數的多項式。這個結果也可以用複數（complex numbers）漂亮地解釋，因為 $\prod(1 + \mathrm{i}t_k) = \prod \sec\theta_k \, \mathrm{e}^{\mathrm{i}\sum\theta_k}$，而它的實部和虛部分別是 $1 - \sigma_2 + \sigma_4$ 和 $\sigma_1 - \sigma_3$。Needham 的 *Visual Complex Analysis* 深入發展了這種幾何的思考方式，而 Siklos 的 *Advanced Problems in Mathematics* 則收錄了更多同類風格的 STEP 題目。

詳細解答

STEP III 2007，第 2 題

純數學 • 二項式冪級數

STEP III 2007，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/07-s3-q2

關鍵思路

一切都源自同一個二項級數（binomial series）。只要知道 $(1-4x)^{-1/2}$ 的係數，對級數求導便會帶出一個因子 n（把 $(n!)^2$ 變成 $n!\,(n-1)!$），而對它積分則會除以 $n+1$（把 $(n!)^2$ 變成 $n!\,(n+1)!$）。因此，後面每一個和都是第一個級數的導數或積分，在 $|x|<\frac{1}{4}$ 內一個精心挑選的點上取值。

解答

第 (i) 部分

把奇數的乘積同時乘以和除以所欠缺的偶數 $2\cdot4\cdot6\cdots(2n)$：

$$1\cdot3\cdot5\cdots(2n-1)=\frac{1\cdot2\cdot3\cdot4\cdots(2n-1)(2n)}{2\cdot4\cdots(2n)}=\frac{(2n)!}{2^n\,(1\cdot2\cdots n)}=\frac{(2n)!}{2^n\,n!},$$

因為 n 個偶數 $2k$ 各自提供一個因子 2 和一個因子 k。

當 $|4x|<1$，即 $|x|<\frac{1}{4}$ 時，由二項級數得

$$(1-4x)^{-1/2}=1+\sum_{n=1}^{\infty}\frac{\left(-\frac{1}{2}\right)\left(-\frac{3}{2}\right)\cdots\left(-\frac{2n-1}{2}\right)}{n!}(-4x)^n.$$

分子有 n 個因子，每個都是負數，所以它等於 $(-1)^n\,\dfrac{1\cdot3\cdots(2n-1)}{2^n}$。這個 $(-1)^n$ 與 $(-4)^n$ 中的符號互相抵消，而 $4^n/2^n=2^n$，所以 x^n 的係數是

$$\frac{2^n\cdot1\cdot3\cdots(2n-1)}{n!}=\frac{2^n}{n!}\cdot\frac{(2n)!}{2^n\,n!}=\frac{(2n)!}{(n!)^2},$$

這裏用了第一個結果。因此，當 $|x| < \frac{1}{4}$ 時，

$$\frac{1}{\sqrt{1-4x}} = 1 + \sum_{n=1}^{\infty} \frac{(2n)!}{(n!)^2}\, x^n,$$

正是所求。

第 (ii) 部分

冪級數（power series）在其收斂區間（interval of convergence）之內可以逐項求導（differentiate term by term），這裏的收斂區間是 $|x| < \frac{1}{4}$。對兩邊求導，

$$\frac{\mathrm{d}}{\mathrm{d}x}(1-4x)^{-1/2} = -\tfrac{1}{2}(1-4x)^{-3/2} \cdot (-4) = 2(1-4x)^{-3/2},$$

而右邊的常數項消失，x^n 變成 nx^{n-1}。由於 $n/(n!)^2 = 1/\big(n!\,(n-1)!\big)$，

$$2(1-4x)^{-3/2} = \sum_{n=1}^{\infty} \frac{(2n)!}{n!\,(n-1)!}\, x^{n-1}.$$

乘以 x 便得到我們想要的冪 x^n：

$$\sum_{n=1}^{\infty} \frac{(2n)!}{n!\,(n-1)!}\, x^n = \frac{2x}{(1-4x)^{3/2}} \qquad (|x| < \tfrac{1}{4}).$$

現在代入 $x = \frac{6}{25}$，這是容許的，因為 $\frac{6}{25} < \frac{6}{24} = \frac{1}{4}$。這時 $1-4x = \frac{1}{25}$，所以 $(1-4x)^{3/2} = \frac{1}{125}$，於是

$$\sum_{n=1}^{\infty} \frac{(2n)!}{n!\,(n-1)!} \left(\frac{6}{25}\right)^n = 2 \cdot \frac{6}{25} \cdot 125 = 60.$$

第 (iii) 部分

> ? **想一想** 現在階乘是 $(n+1)!\,n!$ 而不是 $(n!)^2$。對級數做甚麼運算才會得出這個結果？

我們需要把第 n 個係數除以 $n+1$，因為 $(n+1)! = (n+1)\,n!$。對 x^n 積分正好做到這一點：$\int_0^x t^n\,\mathrm{d}t = \dfrac{x^{n+1}}{n+1}$。冪級數在其收斂區間之內同樣可以逐項積分（term-by-term integration），所以當 $|x| < \frac{1}{4}$ 時，我們把第 (i) 部分的級數由 0 積分至 x：

$$\int_0^x (1-4t)^{-1/2}\,\mathrm{d}t = \Big[-\tfrac{1}{2}(1-4t)^{1/2}\Big]_0^x = \frac{1-\sqrt{1-4x}}{2},$$

於是

$$\frac{1-\sqrt{1-4x}}{2} = x + \sum_{n=1}^{\infty} \frac{(2n)!}{n!\,(n+1)!}\, x^{n+1}.$$

接着我們要對上目標和式中的冪。由於 $3^{2n} = 9^n = 9^{n+1}/9$，

$$\frac{2^{n+1}}{3^{2n}} = 9\left(\frac{2}{9}\right)^{n+1},$$

所以目標和式等於我們的級數在 $x = \frac{2}{9}$ 處的值的 9 倍。這個值是合理的，因為 $\frac{2}{9} < \frac{2}{8} = \frac{1}{4}$。在該處，$1 - 4x = \frac{1}{9}$ 而 $\sqrt{1-4x} = \frac{1}{3}$，由此得

$$\begin{aligned}\sum_{n=1}^{\infty} \frac{2^{n+1}(2n)!}{3^{2n}(n+1)!\,n!} &= 9\sum_{n=1}^{\infty} \frac{(2n)!}{n!\,(n+1)!}\left(\frac{2}{9}\right)^{n+1} \\ &= 9\left(\frac{1-\frac{1}{3}}{2} - \frac{2}{9}\right) = 9\left(\frac{1}{3} - \frac{2}{9}\right) = 1,\end{aligned}$$

正是所求。留意首項 x（即常數 1 的積分）必須減去，因為和式由 $n = 1$ 開始。

驗算

- 在 (i) 中，$n = 2$ 給出 $1 \cdot 3 = 3$ 和 $\frac{4!}{2^2\,2!} = 3$；級數的開首是 $1 + 2x + 6x^2 + 20x^3 + \cdots$，與平方後的結果吻合：$(1 + 2x + 6x^2)^2(1 - 4x) = 1 + O(x^3)$。
- 以精確分數把 (ii) 的首 3000 項相加，得 60.000…；而 (iii) 的首 400 項之和為 1.000…，準確至機器精度。
- 在 (iii) 中，首項（$n = 1$）是 $\frac{4\cdot 2}{9\cdot 2} = \frac{4}{9}$，已佔總和 1 的相當大部分；其後各項大致按 $(8/9)^n$ 的速度遞減。

常見錯誤

- 弄錯二項式係數（binomial coefficient）的正負號：分子有 n 個負因子，$(-4)^n$ 中也有 n 個，所以它們剛好互相抵消。
- 忘記檢查 $\frac{6}{25}$ 和 $\frac{2}{9}$ 是否位於 $|x| < \frac{1}{4}$ 之內；在這個範圍之外，級數會發散（diverge）。
- 在 (iii) 中，忘記了對常數 1 積分所得的 x 項，這樣會得出答案 3 而不是 1。

STEP 之外

第 (iii) 部分出現的數 $\frac{(2n)!}{n!\,(n+1)!} = \frac{1}{n+1}\binom{2n}{n}$ 就是卡塔蘭數(Catalan numbers)1, 1, 2, 5, 14, …，它們可用來點算加括號的方式、二元樹（binary trees）以及許多其他結構的數目，而 $\frac{1-\sqrt{1-4x}}{2x}$ 是它們的母函數（generating function）。我們未經證明便使用的事實，即冪級數在嚴格位於其收斂半徑（radius of convergence）之內的範圍可以逐項求導和逐

項積分，是大學一年級分析學（analysis）的標準定理；參見 Abbott 的 *Understanding Analysis*（分析學入門教材）。

詳細解答

STEP III 2007，第 3 題

純數學 • 斐波那契恆等式

STEP III 2007，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/07-s3-q3

關鍵思路

這裏每一條恆等式的證法都一樣：取下標最大的那一項，用遞推關係（recurrence）$F_m = F_{m-1} + F_{m-2}$ 把它拆開。這樣做一次，(ii) 中的式子就變成所有下標都減一的式子；做兩次，便回到同一形狀而下標全部減二，這正是歸納法（induction）所需要的。至於 (iv) 的整除性（divisibility），留意 $F_{2n-1} + F_{2n} = F_{2n+1}$，所以 F_{2n-1}^2 與 F_{2n}^2 除以 F_{2n+1} 時餘數相同。

解答

第 (i) 部分

由 $F_1 = F_2 = 1$ 開始，把相鄰兩項相加：

$$F_3 = 2, \quad F_4 = 3, \quad F_5 = 5, \quad F_6 = 8, \quad F_7 = 13, \quad F_8 = 21.$$

第 (ii) 部分

設 $k \geq 1$，使每一項都有定義。把下標最大的項 F_{2k+3} 換成 $F_{2k+2} + F_{2k+1}$：

$$\begin{aligned} F_{2k+3}F_{2k+1} - F_{2k+2}^2 &= (F_{2k+2} + F_{2k+1})F_{2k+1} - F_{2k+2}^2 \\ &= F_{2k+1}^2 + F_{2k+2}\big(F_{2k+1} - F_{2k+2}\big) \\ &= F_{2k+1}^2 - F_{2k+2}F_{2k}, \end{aligned}$$

因為由遞推關係可知 $F_{2k+2} - F_{2k+1} = F_{2k}$。這就是 $-F_{2k+2}F_{2k} + F_{2k+1}^2$，正如所求。

第 (iii) 部分

以 $P(n)$ 表示命題 $F_{2n+1}F_{2n-1} - F_{2n}^2 = 1$。

基礎情況（*base case*）。當 $n = 1$ 時：$F_3F_1 - F_2^2 = 2 \cdot 1 - 1^2 = 1$，所以 $P(1)$ 成立。

歸納步驟（*inductive step*）。假設 $P(k)$ 對某個 $k \geq 1$ 成立。由第 (ii) 部分，

$$F_{2k+3}F_{2k+1} - F_{2k+2}^2 = F_{2k+1}^2 - F_{2k+2}F_{2k}.$$

想一想 第 (ii) 部分只把下標減了一。怎樣才能一路退回到 $P(k)$？

再用一次同樣的技巧，這次拆開下標最大的項 $F_{2k+2} = F_{2k+1} + F_{2k}$：

$$\begin{aligned} F_{2k+1}^2 - F_{2k+2}F_{2k} &= F_{2k+1}^2 - (F_{2k+1} + F_{2k})F_{2k} \\ &= F_{2k+1}\big(F_{2k+1} - F_{2k}\big) - F_{2k}^2 \\ &= F_{2k+1}F_{2k-1} - F_{2k}^2, \end{aligned}$$

當中用到 $F_{2k+1} - F_{2k} = F_{2k-1}$（因為 $2k+1 \geq 3$，所以成立）。根據歸納假設（hypothesis）$P(k)$，這個式子等於 1。把以上兩組等式連起來，

$$F_{2k+3}F_{2k+1} - F_{2k+2}^2 = 1,$$

這就是 $P(k+1)$。由歸納法，$P(n)$ 對所有 $n \geq 1$ 成立。

推論。把 $P(n)$ 移項，

$$F_{2n}^2 + 1 = F_{2n+1}F_{2n-1},$$

而 F_{2n-1} 是整數（由遞推關係可知每個 F_m 都是整數），所以 $F_{2n}^2 + 1$ 可被 F_{2n+1} 整除。

第 (iv) 部分

由遞推關係，$F_{2n-1} = F_{2n+1} - F_{2n}$。兩邊平方，

$$F_{2n-1}^2 = F_{2n+1}^2 - 2F_{2n+1}F_{2n} + F_{2n}^2,$$

所以

$$F_{2n-1}^2 + 1 = F_{2n+1}\big(F_{2n+1} - 2F_{2n}\big) + \big(F_{2n}^2 + 1\big).$$

第一個括號所在的一項是 F_{2n+1} 的倍數（它的餘因子（cofactor）是整數），而由第 (iii) 部分可知 $F_{2n}^2 + 1$ 也是 F_{2n+1} 的倍數。因此 $F_{2n-1}^2 + 1$ 可被 F_{2n+1} 整除。

事實上我們還可以求出商（quotient）。利用 $F_{2n}^2 + 1 = F_{2n+1}F_{2n-1}$，

$$F_{2n-1}^2 + 1 = F_{2n+1}\big(F_{2n+1} - 2F_{2n} + F_{2n-1}\big) = F_{2n+1}\big(2F_{2n-1} - F_{2n}\big),$$

而當 $n \geq 2$ 時，括號內的式子是 $F_{2n-1} - F_{2n-2} = F_{2n-3}$。所以 $F_{2n-1}^2 + 1 = F_{2n+1}F_{2n-3}$，與 (iii) 中的恆等式（identity）十分相似。

驗算

取 $n = 3$：$F_6^2 + 1 = 65 = 13 \times 5 = F_7F_5$，而 $F_5^2 + 1 = 26 = 13 \times 2 = F_7F_3$。取 $n = 4$：$F_7^2 + 1 = 170 = 34 \times 5 = F_9F_5$。我們亦用電腦驗證過，對於直至 37 的每一個 n，(ii) 和 (iii) 中的恆等式、兩個整除性命題以及商 F_{2n-3} 全部成立。

常見錯誤

- 只用 (ii) 來做歸納：它只把下標減一，所得的式子 $F_{2k+1}^2 - F_{2k+2}F_{2k}$ 還不是歸納假設，必須再用一次遞推關係。
- 忘記基礎情況，或驗證時用錯下標（$n = 1$ 涉及 F_1、F_2、F_3）。

STEP 之外

第 (iii) 部分是卡西尼恆等式（Cassini's identity）$F_{m+1}F_{m-1} - F_m^2 = (-1)^m$ 的一半，而這條恆等式用行列式（determinant）即可立刻得出：矩陣 $\begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}^m$ 的元素是 F_{m+1}、F_m、F_m、F_{m-1}，而它的行列式是 $(-1)^m$。像 (iv) 這類整除性的結果，是線性遞推數列（linear recurrence）算術的起點，當中大部分工作都靠同餘（congruence）完成。想進一步學習數論（number theory），Davenport 的 *The Higher Arithmetic*（高等算術）是很好的下一步；想了解矩陣的觀點，則可讀 Strang 的 *Introduction to Linear Algebra*（線性代數導論）。

詳細解答

STEP III 2007，第 4 題

純數學 • 參數曲線與曲率

> **STEP III 2007，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/07-s3-q4

關鍵思路

一切都建基於 $\dot{x}$ 的一條簡潔公式。看似棘手的一項 $\ln\tan\frac{1}{2}t$ 求導後得 $1/\sin t$，於是 $\dot{x}$ 便化簡為 $a\cos^2 t/\sin t$。一旦求得 $\dot{x}$ 和 $\dot{y}$，斜率（gradient）、切線（tangent）、曲率（curvature）和法線（normal）都可以用直接的坐標幾何逐一得出。把單位切向量（unit tangent）寫成 $(\cos t, \sin t)$，可令最後一部分十分簡短。

解答

斜率

由於 $\dfrac{\mathrm{d}}{\mathrm{d}t}\ln\tan\frac{1}{2}t = \dfrac{\frac{1}{2}\sec^2\frac{1}{2}t}{\tan\frac{1}{2}t} = \dfrac{1}{2\sin\frac{1}{2}t\cos\frac{1}{2}t} = \dfrac{1}{\sin t}$，我們得到

$$\dot{x} = a\left(-\sin t + \frac{1}{\sin t}\right) = a\,\frac{1-\sin^2 t}{\sin t} = \frac{a\cos^2 t}{\sin t},$$
$$\dot{y} = a\cos t.$$

當 $0 < t < \frac{1}{2}\pi$ 時兩者皆為正數，所以

$$\frac{\mathrm{d}y}{\mathrm{d}x} = \frac{\dot{y}}{\dot{x}} = a\cos t\cdot\frac{\sin t}{a\cos^2 t} = \tan t,$$

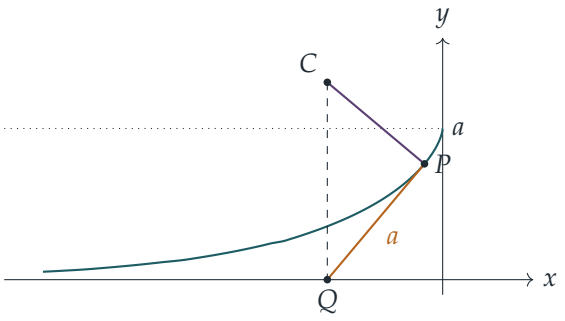
正如題目所求。

草圖

- 當 $t \to 0^+$ 時，$\tan\frac{1}{2}t \to 0^+$，所以 $x \to -\infty$ 而 $y \to 0$。x 軸是一條漸近線（asymptote）。

- 當 $t \to \frac{1}{2}\pi^-$ 時，$x \to a(0+\ln 1) = 0$ 而 $y \to a$。曲線終止於 $(0,a)$，該處的斜率 $\tan t \to \infty$，所以該處的切線是垂直的。
- 由於 $\dot{x} > 0$，x 隨 t 遞增，斜率 $\tan t$ 亦然。所以曲線穩步上升，並且向上凹（concave upwards），而 y 介乎 0 與 a 之間。

這條曲線就是曳物線（tractrix）。圖中亦標示了後面各部分所用的點 P、Q 和 C。



切線：$PQ = a$

在 $P = (x,y)$ 的切線斜率為 $\tan t$。由 P 沿水平方向往回走 $y/\tan t$ 的距離，切線便與 $y = 0$ 相交，所以

$$x_Q = x - \frac{y}{\tan t} = x - \frac{a\sin t\cos t}{\sin t} = x - a\cos t.$$

因此 $\overrightarrow{QP} = (a\cos t,\ a\sin t)$，並且

$$PQ = \sqrt{a^2\cos^2 t + a^2\sin^2 t} = a.$$

所以由曲線至 x 軸的切線長度恆為 a。（留待後用：$x_Q = a\ln\tan\frac{1}{2}t$。）

曲率半徑

把 $\dot{x}$ 寫成 $\dot{x} = a\Big(\dfrac{1}{\sin t} - \sin t\Big)$，求導便容易得多：

$$\ddot{x} = a\Big(-\frac{\cos t}{\sin^2 t} - \cos t\Big), \qquad \ddot{y} = -a\sin t.$$

於是

$$\begin{aligned}\dot{x}\ddot{y} - \dot{y}\ddot{x} &= \frac{a\cos^2 t}{\sin t}(-a\sin t) + a\cos t\cdot a\Big(\frac{\cos t}{\sin^2 t} + \cos t\Big)\\ &= a^2\Big(-\cos^2 t + \frac{\cos^2 t}{\sin^2 t} + \cos^2 t\Big) = a^2\cot^2 t,\end{aligned}$$

以及

$$\dot{x}^2+\dot{y}^2 = a^2\cos^2 t\left(\frac{\cos^2 t}{\sin^2 t}+1\right) = \frac{a^2\cos^2 t}{\sin^2 t} = a^2\cot^2 t.$$

由於在我們的範圍內 $\cot t > 0$，故 $(\dot{x}^2+\dot{y}^2)^{3/2} = a^3\cot^3 t$，因此

$$\rho = \frac{a^3\cot^3 t}{a^2\cot^2 t} = a\cot t,$$

正如題目所求。

曲率中心：CQ 是垂直的

在 P 的切線沿單位向量 $(\cos t, \sin t)$ 的方向，所以法線沿 $(-\sin t, \cos t)$ 或其反方向。

> ? **想一想** 兩個法線方向之中，哪一個是「在曲線上方」？

由於 $\cos t > 0$，向量 $(-\sin t, \cos t)$ 的 y 分量為正數。由 P 沿這個方向走一小段，便會到達曲線上方，因為曲線向上凹（而且無論如何，另一個方向只會把我們帶往下方）。所以

$$\overrightarrow{OC} = \overrightarrow{OP} + \rho(-\sin t, \cos t) = \big(x - a\cot t\sin t,\ a\sin t + a\cot t\cos t\big).$$

根據切線部分的結果，其 x 坐標為

$$x_C = x - a\cos t = x_Q$$

。所以 C 和 Q 的 x 坐標相同，CQ 平行於 y 軸，正如題目所求。（y 坐標可化簡為 $a\sin t + \dfrac{a\cos^2 t}{\sin t} = \dfrac{a}{\sin t}$，所以 $C = \big(a\ln\tan\tfrac{1}{2}t,\ a\operatorname{cosec} t\big)$。）

驗算

- 取 $a = 1$ 和 $t = \frac{5\pi}{18}$（50°），得 $P \approx (-0.1202, 0.7660)$ 和 $Q \approx (-0.7630, 0)$，由此 $PQ = 1.0000$。公式給出 $C \approx (-0.7630, 1.3054)$，正好在 Q 的正上方。
- 以 2000 組隨機的 a 和 t 值作數值檢驗，其中 $\dot{x}$、$\ddot{x}$ 以有限差分（finite differences）求得，結果與 $\mathrm{d}y/\mathrm{d}x = \tan t$、$PQ = a$、$\rho = a\cot t$ 和 $x_C = x_Q$ 相符，誤差只在捨入誤差範圍之內。
- 當 $t \to \frac{1}{2}\pi$ 時，$\rho \to 0$：曲線趨近 $(0, a)$ 時彎得愈來愈急（它對 y 軸的鏡像與它在該處相接，形成一個尖點（cusp））。當 $t \to 0$ 時，$\rho \to \infty$，與曲線逐漸貼平於漸近線的情況吻合。

常見錯誤

- 對 $\ln\tan\frac{1}{2}t$ 求導時，漏掉由鏈式法則(chain rule)而來的因子 $\frac{1}{2}$。答案必須是 $1/\sin t$。
- 把 $\dot{x}$ 留作 $a(-\sin t+\operatorname{cosec} t)$ 而不加化簡。正是 $a\cos^2 t/\sin t$ 這個形式，才令其後每一部分都整齊利落。
- 採用法線方向 $(\sin t, -\cos t)$，這個方向指向曲線下方，會把 C 放在錯誤的位置。

STEP 之外

設有人沿 x 軸行走，用一條長度為 a 的繩拖着一件重物，重物所走的路徑就是曳物線，而 $PQ = a$ 正是這個性質。點 C 描出曲線的*漸屈線*（evolute），即其曲率中心的軌跡(locus)。這裏 $C = (a\ln\tan\frac{1}{2}t,\ a\operatorname{cosec} t)$ 原來位於一條懸鏈線（catenary）上。把曳物線繞其漸近線旋轉，便得到偽球面（pseudosphere），這是一個具有恆定負曲率的曲面，也是雙曲幾何（hyperbolic geometry）的一個模型。Needham 的 *Visual Complex Analysis* 對偽球面有一段優美的幾何論述。

詳細解答

STEP III 2007，第 5 題

純數學 • 高階導數與數學歸納法

STEP III 2007，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/07-s3-q5

關鍵思路

只要把 r 和 $\mathrm{d}\theta/\mathrm{d}x$ 都單用 θ 表示，一切便變得簡單。由 $\coth\theta = x$ 可得 $r = 1/\sinh\theta$，繼而有 $\mathrm{d}r/\mathrm{d}x = \cosh\theta$ 及 $\mathrm{d}\theta/\mathrm{d}x = -\sinh^2\theta$。對 $\cosh n\theta / r^n$ 求導，所得的恰好就是 $\cosh(n+1)\theta$ 的複角公式（compound-angle formula），於是便可用歸納法（induction）證明。

解答

準備工作

由於 $x > 1$，方程 $\coth\theta = x$ 有唯一滿足 $\theta > 0$ 的解，因此 $\sinh\theta > 0$。利用 $\coth^2\theta - 1 = \operatorname{cosech}^2\theta$，

$$r^2 = x^2 - 1 = \coth^2\theta - 1 = \frac{1}{\sinh^2\theta}, \qquad \text{所以} \qquad r = \frac{1}{\sinh\theta},$$

這裏取正根，因為 r 和 $\sinh\theta$ 都是正數。

整個解答只靠兩個導數（derivative）。把 $\coth\theta = x$ 對 x 求導，

$$-\operatorname{cosech}^2\theta\,\frac{\mathrm{d}\theta}{\mathrm{d}x} = 1 \quad \Longrightarrow \quad \frac{\mathrm{d}\theta}{\mathrm{d}x} = -\sinh^2\theta = -\frac{1}{r^2}.$$

把 $r^2 = x^2 - 1$ 求導，得 $2r\,\dfrac{\mathrm{d}r}{\mathrm{d}x} = 2x$，所以

$$\frac{\mathrm{d}r}{\mathrm{d}x} = \frac{x}{r} = \coth\theta\,\sinh\theta = \cosh\theta.$$

首兩個導數

由於 $y = \ln r^2 = 2\ln r$，

$$\frac{\mathrm{d}y}{\mathrm{d}x} = \frac{2}{r}\frac{\mathrm{d}r}{\mathrm{d}x} = \frac{2\cosh\theta}{r},$$

正是所求。至於二階導數，最簡潔的做法是寫成 $1/r = \sinh\theta$，於是 $\dfrac{\mathrm{d}y}{\mathrm{d}x} = 2\cosh\theta\sinh\theta = \sinh 2\theta$。由鏈式法則（chain rule），

$$\frac{\mathrm{d}^2y}{\mathrm{d}x^2} = 2\cosh 2\theta\,\frac{\mathrm{d}\theta}{\mathrm{d}x} = -\frac{2\cosh 2\theta}{r^2},$$

正是所求。

三階導數

現在把二階導數寫成 $\dfrac{\mathrm{d}^2y}{\mathrm{d}x^2} = -2\cosh 2\theta\sinh^2\theta$，並運用積法則（product rule），其中 $\mathrm{d}\theta/\mathrm{d}x = -\sinh^2\theta$：

$$\begin{aligned}\frac{\mathrm{d}^3y}{\mathrm{d}x^3} &= -2\big(2\sinh 2\theta\sinh^2\theta + 2\cosh 2\theta\sinh\theta\cosh\theta\big)\frac{\mathrm{d}\theta}{\mathrm{d}x}\\ &= 4\sinh^3\theta\big(\cosh 2\theta\cosh\theta + \sinh 2\theta\sinh\theta\big)\\ &= 4\sinh^3\theta\cosh 3\theta,\end{aligned}$$

最後一步用了恆等式 $\cosh(A+B) = \cosh A\cosh B + \sinh A\sinh B$。所以

$$\frac{\mathrm{d}^3y}{\mathrm{d}x^3} = \frac{4\cosh 3\theta}{r^3}.$$

一般公式

從 $2\cosh\theta/r$、$-2\cosh 2\theta/r^2$、$4\cosh 3\theta/r^3$ 這個規律可見，係數應為 $2\cdot 0!$、$-2\cdot 1!$、$2\cdot 2!$，因此我們猜想

$$\frac{\mathrm{d}^ny}{\mathrm{d}x^n} = (-1)^{n-1}\,\frac{2\,(n-1)!\,\cosh n\theta}{r^n} \qquad (n \geq 1).$$

想一想 剛才求三階導數的那個計算，推廣到一般情況會是怎樣的？

那就是求 $\cosh n\theta/r^n = \cosh n\theta\,\sinh^n\theta$ 的導數。由積法則和鏈式法則，並利用 $\mathrm{d}\theta/\mathrm{d}x = -\sinh^2\theta$，

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}x}\big(\cosh n\theta\,\sinh^n\theta\big) &= \big(n\sinh n\theta\,\sinh^n\theta + n\cosh n\theta\,\sinh^{n-1}\theta\cosh\theta\big)\big(-\sinh^2\theta\big)\\ &= -n\sinh^{n+1}\theta\,\big(\cosh n\theta\cosh\theta + \sinh n\theta\sinh\theta\big)\\ &= -n\sinh^{n+1}\theta\,\cosh(n+1)\theta,\end{aligned}$$

即

$$\frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{\cosh n\theta}{r^n}\right) = -\frac{n\cosh(n+1)\theta}{r^{n+1}}.$$

現在進行歸納。公式對 $n = 1$ 成立（它就是上面求得的一階導數）。假設公式對某個 $n = k \geq 1$ 成立。那麼，把上面的計算用於 $n = k$，

$$\frac{\mathrm{d}^{k+1}y}{\mathrm{d}x^{k+1}} = (-1)^{k-1}\,2\,(k-1)!\cdot\left(-\frac{k\cosh(k+1)\theta}{r^{k+1}}\right) = (-1)^k\,\frac{2\,k!\,\cosh(k+1)\theta}{r^{k+1}},$$

這正是 $n = k+1$ 時的公式。由歸納法，

$$\boxed{\frac{\mathrm{d}^n y}{\mathrm{d}x^n} = (-1)^{n-1}\,\frac{2\,(n-1)!\,\cosh n\theta}{r^n} \quad \text{對所有} n \geq 1.}$$

（$n = 2$ 和 $n = 3$ 的情況與我們直接求得的結果一致。）

另一途徑

部分分式（partial fractions）提供了一個獨立的推導。由於 $y = \ln(x-1) + \ln(x+1)$，

$$\frac{\mathrm{d}^n y}{\mathrm{d}x^n} = (-1)^{n-1}(n-1)!\left[(x-1)^{-n} + (x+1)^{-n}\right].$$

但 $x \pm 1 = \dfrac{\cosh\theta \pm \sinh\theta}{\sinh\theta} = r\,\mathrm{e}^{\pm\theta}$，所以方括號內的式子等於 $r^{-n}\left(\mathrm{e}^{-n\theta} + \mathrm{e}^{n\theta}\right) = 2\cosh n\theta / r^n$，得出相同的公式。這也解釋了這個代換的來歷：$\theta$ 就是 $\ln\dfrac{x+1}{x-1}$ 的一半。

驗算

- 當 $x = 2$ 時：由部分分式得 $\mathrm{d}^3y/\mathrm{d}x^3 = 2\left(1 + \frac{1}{27}\right) = \frac{56}{27} \approx 2.0741$，而以 $\theta = \tanh^{-1}\frac{1}{2}$、$r = \sqrt{3}$ 代入 $4\cosh 3\theta / r^3$ 亦得相同的值。
- 我們在 200 個滿足 $x > 1$ 的隨機點上，把首三個導數與 $\ln(x^2-1)$ 的有限差分（finite differences）作了比較，又就 $n \leq 14$ 把一般公式與部分分式的形式作了比較（兩者相符至約 10^{-14}）。
- 正負號：當 $x > 1$ 時，$\ln(x^2-1)$ 遞增而且是凹的（concave），與 $y' > 0$ 及 $y'' < 0$ 吻合。

常見錯誤

- 寫成 $\mathrm{d}\theta/\mathrm{d}x = +\sinh^2\theta$：$\coth\theta$ 的導數是 $-\operatorname{cosech}^2\theta$，而正是這個負號令各階導數的正負號交替出現。

- 階乘（factorial）差了一位：第 n 階導數所帶的是 $(n-1)!$，因為一階導數的係數是 $2 = 2 \cdot 0!$。
- 只驗證幾個情況便停下來：題目要求「with proof」（附證明），就必須寫出一般的歸納步驟。

STEP 之外

每一階導數都有如此簡潔的閉合形式（closed form），讓我們一窺複分析（complex analysis）的面貌：$\ln(x^2-1)$ 在 $x = \pm 1$ 處有奇點（singularities），而它的高階導數正是由到這些奇點的距離所支配，在這裏就是 $x \mp 1 = re^{\mp\theta}$。在大學一年級的分析課程中，要估計泰勒級數（Taylor series）餘項的界，並求出其收斂半徑（radius of convergence），所需要的正是這類公式。Needham 的 *Visual Complex Analysis* 把奇點為何支配冪級數（power series）解釋得十分精彩。

詳細解答

STEP III 2007，第 6 題

純數學 • 複數與圓

STEP III 2007，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/07-s3-q6

關鍵思路

在以原點為圓心、半徑為 a 的圓上，取共軛（conjugation）與反演（inversion）是同一回事：$pp^* = a^2$，所以 $p^* = a^2/p$。單憑這一點，題目中每一個共軛都可化為普通的分式，恆等式（identity）隨即得出。至於垂直，判別方法是兩條弦所對應的向量之比為純虛數（purely imaginary）。利用恆等式，這個判別法便化為一個簡單的乘積條件。對於多邊形，這個乘積條件就是一個遞推關係（recurrence）：一旦選定 B_1，其餘每一個 B_k 都隨之確定，唯一的問題是這條鏈繞回 B_1 時能否閉合。

解答

恆等式

由於 P 在圓上，$|p| = a$，所以 $pp^* = a^2$，即 $p^* = a^2/p$；同理 $q^* = a^2/q$。因此

$$p^* - q^* = \frac{a^2}{p} - \frac{a^2}{q} = \frac{a^2(q-p)}{pq} = -\frac{a^2(p-q)}{pq}.$$

因為 $p \neq q$（而且由 $a > 0$ 知 $p, q \neq 0$），上式不等於零，所以可以移項得

$$pq = -a^2 \frac{p-q}{p^* - q^*},$$

這正是所求。

互相垂直的弦

弦（chord）PQ 的方向由複數 $p-q$ 給出，而 RS 的方向由 $r-s$ 給出。兩個非零複數互相垂直，當且僅當它們的比

$$w=\frac{p-q}{r-s}$$

是純虛數，因為這個比的輻角（argument）就是兩者之間的夾角。

> **想一想** 怎樣把「w 是純虛數」化為一條可以運用的方程？

非零的 w 是純虛數，當且僅當 $w^*=-w$，即 $w/w^*=-1$。對兩條弦分別使用恆等式（第二條以 r,s 代替 p,q），

$$\frac{w}{w^*}=\frac{p-q}{p^*-q^*}\cdot\frac{r^*-s^*}{r-s}=\frac{-pq/a^2}{-rs/a^2}=\frac{pq}{rs}.$$

所以 $PQ\perp RS$ 當且僅當 $pq/rs=-1$。由於 $rs\neq 0$，這就是說

$$pq+rs=0.$$

這裏每一步都可逆，所以我們證明的比題目要求的稍多：對於這個圓的弦，$PQ\perp RS\iff pq+rs=0$。下文需要用到逆命題（converse）這個方向。

多邊形問題

設點 A_k 和 B_k 分別由 α_k 和 β_k 表示。根據上一部分（及其逆命題），那 n 個垂直條件恰好就是

$$\beta_k\beta_{k+1}=-\alpha_k\alpha_{k+1}\quad(k=1,\dots,n-1),\qquad \beta_n\beta_1=-\alpha_n\alpha_1,$$

前提是相鄰的 B 互不相同，使每條 B_kB_{k+1} 確實是一條弦。

選定 β_1。由於 $\beta_{k+1}=-\alpha_k\alpha_{k+1}/\beta_k$，首 $n-1$ 個條件便依次確定 $\beta_2,\beta_3,\dots,\beta_n$，而且每個 β_{k+1} 的模（modulus）自動等於 a。算出最初幾個：

$$\beta_2=-\frac{\alpha_1\alpha_2}{\beta_1},\qquad \beta_3=-\frac{\alpha_2\alpha_3}{\beta_2}=\frac{\alpha_3\beta_1}{\alpha_1},\qquad \beta_4=-\frac{\alpha_3\alpha_4}{\beta_3}=-\frac{\alpha_1\alpha_4}{\beta_1},$$

再用簡單的歸納法（induction）可得

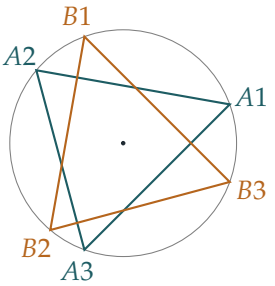
$$\beta_k=\frac{\alpha_k\beta_1}{\alpha_1}\ (k\text{ 為奇數}),\qquad \beta_k=-\frac{\alpha_1\alpha_k}{\beta_1}\ (k\text{ 為偶數}).$$

（若公式對 k 成立，則由 $\beta_{k+1}=-\alpha_k\alpha_{k+1}/\beta_k$ 可得 $k+1$ 時的另一種形式。）現在一切取決於最後一個條件 $\beta_n\beta_1=-\alpha_n\alpha_1$ 是否成立。

$n=3$。這時 $\beta_3=\alpha_3\beta_1/\alpha_1$，所以閉合條件是

$$\frac{\alpha_3\beta_1^2}{\alpha_1}=-\alpha_3\alpha_1\iff\beta_1^2=-\alpha_1^2\iff\beta_1=\pm\mathrm{i}\,\alpha_1.$$

所以 B_1 只有兩個選擇。取 $\beta_1 = \mathrm{i}\alpha_1$ 時，由公式得 $\beta_2 = \mathrm{i}\alpha_2$ 和 $\beta_3 = \mathrm{i}\alpha_3$；取 $\beta_1 = -\mathrm{i}\alpha_1$ 時，每個 $\beta_k = -\mathrm{i}\alpha_k$。兩種情況下，三角形 $B_1B_2B_3$ 都是三角形 $A_1A_2A_3$ 繞圓心旋轉一個直角而成。它的頂點互不相同，所以兩個選擇都確實可行；而把一條弦旋轉 $90°$，所得的弦當然與原來的弦垂直。



圖中所示的是 $\beta_1 = \mathrm{i}\alpha_1$ 這個選擇：琥珀色三角形的每一邊，都是藍綠色三角形的對應邊轉過 $90°$ 而成。

$n = 4$。這時 $\beta_4 = -\alpha_1\alpha_4/\beta_1$，所以不論 β_1 是甚麼，

$$\beta_4\beta_1 = -\alpha_1\alpha_4$$

都自動成立。因此 B_1 可以選在圓上*任何位置*，有無窮多個選擇。唯一的例外，是會令兩個相鄰的 B 重合的有限多個位置。例如，$\beta_1 = \beta_2$ 需要 $\beta_1^2 = -\alpha_1\alpha_2$，這排除了兩個點。四種重合情況各自最多排除兩個點，所以最多有八個位置被排除。

$n > 4$。同樣的計算對每個 n 都適用，而且只有 n 的奇偶性（parity）起作用：

- 若 n 是奇數，閉合條件是 $\beta_1^2 = -\alpha_1^2$。恰好有兩個選擇 $\beta_1 = \pm\mathrm{i}\alpha_1$，而 $B_1B_2\cdots B_n$ 就是多邊形 $A_1A_2\cdots A_n$ 繞圓心旋轉 $\pm 90°$ 而成；
- 若 n 是偶數，閉合條件自動成立。除了最多 $2n$ 個會令相鄰的 B 重合的位置外，B_1 可以是圓上任何一點，所以有無窮多個選擇。

驗算

- 試取 $p = a$，$q = \mathrm{i}a$：這時 $pq = \mathrm{i}a^2$，而 $-a^2(a - \mathrm{i}a)/(a + \mathrm{i}a) = -a^2 \cdot (-\mathrm{i}) = \mathrm{i}a^2$，與預期相符。
- 試取直徑 $p = a$，$q = -a$ 和 $r = \mathrm{i}a$，$s = -\mathrm{i}a$，兩者互相垂直：$pq + rs = -a^2 + a^2 = 0$。
- 數值方面，我們以 1000 個隨機構形測試了恆等式和判別準則 $PQ \perp RS \iff pq + rs = 0$。對 $n = 3, \ldots, 9$，我們由隨機的起點出發執行遞推。當 n 是奇數時，讓 B_1 以

0.1° 為步長繞圓掃描一周，這條鏈只在 $\beta_1 = \pm i\alpha_1$ 附近閉合。當 n 是偶數時，它每次都閉合，而且所有弦都互相垂直，誤差僅為捨入誤差。

常見錯誤

- 只證明「垂直 $\Rightarrow pq + rs = 0$」。處理多邊形時還需要逆命題，才能肯定我們構造出來的 β 確實給出互相垂直的弦。
- 把「垂直即 w 是虛數」寫成 $w + w^* = 0$，卻沒有指出 $w \neq 0$。這裏由於各點互不相同，$w \neq 0$ 是有保證的。
- 當 n 是偶數時，說「每個 B_1 都可行」，卻沒有留意到有少數位置會令兩個相鄰的 B 重合，以致那條「弦」沒有方向。

STEP 之外

映射 $z \mapsto a^2/z^*$ 是關於圓 $|z| = a$ 的*反演*，而圓上的點在這個映射下保持不動（$p = a^2/p^*$），正是整道題目的動力所在。反演是莫比烏斯變換（Möbius transformation）的一種，而莫比烏斯變換在複分析與幾何的入門課程中佔有核心地位。多邊形的論證帶有*單值性*（monodromy）的味道：沿着一連串別無選擇的步驟繞一圈，然後問是否回到起點，而答案取決於 n 的奇偶性。Needham 的 *Visual Complex Analysis*（《可視化複分析》）把反演和這些幾何思想闡述得十分精彩。

詳細解答

STEP III 2007，第 7 題

純數學 • 代換積分法

STEP III 2007，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/07-s3-q7

關鍵思路

題目不容許我們提及「arctan」或「π」，所以每一項事實都必須靠積分內的變數代換（substitution）得出。每個代換都經過挑選，令被積函數（integrand）$1/(1+u^2)$（或 $1/\sqrt{1-u^2}$）變回自己，或變成另一個，只有積分上下限（limits）改變。比較上下限，便得到 t 與 s 各個值之間的等式；而第 (i) 部分的對稱性（symmetry）$u \mapsto 1/u$，正是把這些等式化為 p 的倍數的關鍵。

解答

第 (i) 部分

取 $x > 0$，並設 $u = v^{-1}$，於是 $\mathrm{d}u = -v^{-2}\,\mathrm{d}v$。當 u 由 0 增至 x 時，v 由 ∞ 減至 $1/x$。另外

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

因此

$$t(x) = \int_{\infty}^{1/x} \frac{v^2}{1+v^2}\left(-\frac{1}{v^2}\right)\mathrm{d}v = \int_{1/x}^{\infty} \frac{1}{1+v^2}\,\mathrm{d}v,$$

其中負號用來對調上下限。這就是所求的結果。（這個廣義積分（improper integral）收斂（converges），因為被積函數不超過 $1/v^2$。）

以 $1/x$ 代替 x 套用上式，得 $t(1/x) = \int_x^{\infty} \frac{1}{1+v^2}\,\mathrm{d}v$。加上 $t(x) = \int_0^x \frac{1}{1+v^2}\,\mathrm{d}v$，兩個積分範圍便合併成 $[0,\infty)$：

$$t(x) + t(1/x) = \int_0^{\infty} \frac{1}{1+v^2}\,\mathrm{d}v = \tfrac{1}{2}p \qquad (x > 0).$$

代入 $x = 1$，得 $2t(1) = \frac{1}{2}p$。

第 (ii) 部分

當 $u \geq 0$ 時，有 $0 \leq y < 1$。兩邊平方，$y^2(1+u^2) = u^2$，所以 $u^2(1-y^2) = y^2$，而

$$u = \frac{y}{\sqrt{1-y^2}},$$

這裏取正根，因為 u 與 y 同號。由商法則（quotient rule），

$$\frac{\mathrm{d}u}{\mathrm{d}y} = \frac{\sqrt{1-y^2} + \dfrac{y^2}{\sqrt{1-y^2}}}{1-y^2} = \frac{(1-y^2)+y^2}{(1-y^2)^{3/2}} = \frac{1}{\sqrt{(1-y^2)^3}},$$

正如所求。

現在代入 $t(x)$。由 $u^2 = y^2/(1-y^2)$ 得

$$1 + u^2 = \frac{1}{1-y^2},$$

所以

$$\frac{1}{1+u^2}\frac{\mathrm{d}u}{\mathrm{d}y} = (1-y^2)\cdot\frac{1}{(1-y^2)^{3/2}} = \frac{1}{\sqrt{1-y^2}}.$$

映射 $u \mapsto y$ 是遞增（increasing）的（它的反函數的導數為正），而且 $u = 0 \mapsto y = 0$，$u = x \mapsto y = x/\sqrt{1+x^2}$，後者小於 1。所以

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

取 $x = 1$，並利用第 (i) 部分，上式即 $s\left(\frac{1}{\sqrt{2}}\right) = t(1) = \frac{1}{4}p$。

第 (iii) 部分

記 $a = \frac{1}{\sqrt{3}}$，於是 $z = \dfrac{u+a}{1-au}$。在 $0 \leq u \leq a$ 上有 $au \leq \frac{1}{3}$，所以分母保持為正，z 在該範圍內是 u 的光滑函數（smooth function）。求導得

$$\frac{\mathrm{d}z}{\mathrm{d}u} = \frac{(1-au)+a(u+a)}{(1-au)^2} = \frac{1+a^2}{(1-au)^2}.$$

想一想 為甚麼這個代換會令 $1/(1+u^2)$ 保持不變？

通分計算 $1+z^2$。交叉項 $\mp 2au$ 互相抵消：

$$1 + z^2 = \frac{(1-au)^2 + (u+a)^2}{(1-au)^2} = \frac{1 + a^2u^2 + u^2 + a^2}{(1-au)^2} = \frac{(1+a^2)(1+u^2)}{(1-au)^2}.$$

兩式相除，

$$\frac{1}{1+z^2}\frac{\mathrm{d}z}{\mathrm{d}u} = \frac{1}{1+u^2},$$

所以被積函數不變，只有上下限移動。（這背後其實是正切的複角公式（tangent addition formula），但我們用不着它。）當 $u = 0$ 時，$z = a = \frac{1}{\sqrt{3}}$。當 $u = a$ 時，

$$z = \frac{2/\sqrt{3}}{1 - \frac{1}{3}} = \frac{2}{\sqrt{3}} \cdot \frac{3}{2} = \sqrt{3}.$$

因此

$$t\left(\tfrac{1}{\sqrt{3}}\right) = \int_0^{1/\sqrt{3}} \frac{\mathrm{d}u}{1+u^2} = \int_{1/\sqrt{3}}^{\sqrt{3}} \frac{\mathrm{d}z}{1+z^2}.$$

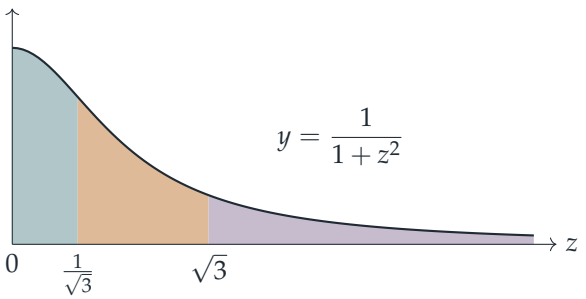
最後一步，把範圍 $[0, \infty)$ 在 $\frac{1}{\sqrt{3}}$ 和 $\sqrt{3}$ 處分開：

$$\tfrac{1}{2}p = \int_0^{1/\sqrt{3}} \frac{\mathrm{d}z}{1+z^2} + \int_{1/\sqrt{3}}^{\sqrt{3}} \frac{\mathrm{d}z}{1+z^2} + \int_{\sqrt{3}}^{\infty} \frac{\mathrm{d}z}{1+z^2}.$$

第一段按定義就是 $t(\frac{1}{\sqrt{3}})$，第二段按剛才證明的結果也是 $t(\frac{1}{\sqrt{3}})$。由第 (i) 部分取 $x = \frac{1}{\sqrt{3}}$，第三段是 $\int_{1/x}^{\infty} \frac{\mathrm{d}z}{1+z^2} = t(x) = t(\frac{1}{\sqrt{3}})$，同樣如此。因此

$$3\,t\left(\tfrac{1}{\sqrt{3}}\right) = \tfrac{1}{2}p.$$

下圖顯示這三段：$1/(1+z^2)$ 之下由 0 至 ∞ 的面積，分成三個面積相等的區域。



驗算

以數值方法（辛普森法則（Simpson's rule））計算，$t(\frac{1}{\sqrt{3}}) \approx 0.5235988$，$\int_{1/\sqrt{3}}^{\sqrt{3}} \frac{\mathrm{d}z}{1+z^2} \approx 0.5235988$，而 $p \approx 3.1415927$，所以 $3t(\frac{1}{\sqrt{3}}) \approx 1.5707963 \approx \frac{1}{2}p$。同樣，$s(\frac{1}{\sqrt{2}}) \approx t(1) \approx 0.7853982 \approx \frac{1}{4}p$。當然，$p$ 其實就是 π，而這些結果說的是 $\arctan \frac{1}{\sqrt{3}} = \frac{\pi}{6}$，$\arcsin \frac{1}{\sqrt{2}} = \arctan 1 = \frac{\pi}{4}$；你可以用這些來驗算，但不可寫在答案裏。

常見錯誤

- 忘記 $u = 1/v$ 會把積分方向倒轉；$\mathrm{d}u$ 中的負號正好把上下限的次序還原。

- 在第 (iii) 部分，做完代換便停下來：由 $\frac{1}{\sqrt{3}}$ 至 $\sqrt{3}$ 的積分只是三段相等部分的其中一段，而 $\sqrt{3}$ 之後的那一段需要用到第 (i) 部分。
- 沒有檢查在積分範圍內 $1 - au \neq 0$；這裏 $au \leq \frac{1}{3}$，所以代換成立。

STEP 之外

這道題讓我們一窺分析學家如何嚴格地定義三角函數（trigonometric functions）：先把 $\arctan x$ 定義為 $\int_0^x \frac{\mathrm{d}u}{1+u^2}$，再把 π 定義為它在無窮遠處的極限（limit）的兩倍，然後靠這類代換推導出所有性質，完全不用幾何。第 (iii) 部分的代換其實是 tan 的複角公式的化身，用同樣的方法可以證明：當 $xy < 1$ 時，$t(x) + t(y) = t\left(\frac{x+y}{1-xy}\right)$。Abbott 的 *Understanding Analysis* 說明了如何利用積分和冪級數（power series）為這類定義奠定穩固的基礎。

詳細解答

STEP III 2007，第 8 題

純數學 • 里卡蒂方程

> **STEP III 2007，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/07-s3-q8

關鍵思路

非線性方程 $(*)$ 是一條里卡蒂方程（Riccati equation）：y^2 項破壞了線性。把 y 寫成 u'/u 的倍數，$y' + y^2$ 便會變成 u''/u，平方項隨之消失，剩下一條關於 u 的線性二階方程。第 (i) 部分的前半已經給了我們這條線性方程的兩個解，所以餘下的工作只是定出常數。至於第 (ii) 部分，我們需要一條新線性方程的兩個解，而線索在於：把 $1 + x$ 中的 x 換成 $-x$，就得到 $1 - x$。

解答

第 (i) 部分

求 a 和 b。當 $u = x$ 時，$u' = 1$，$u'' = 0$，所以需要

$$\mathrm{a}(x) + x\,\mathrm{b}(x) = 0.$$

當 $u = \mathrm{e}^{-x}$ 時，$u' = -\mathrm{e}^{-x}$，$u'' = \mathrm{e}^{-x}$；除以 $\mathrm{e}^{-x} \neq 0$ 得

$$1 - \mathrm{a}(x) + \mathrm{b}(x) = 0.$$

由第一條方程得 $\mathrm{b} = -\mathrm{a}/x$（其中 $x \neq 0$），再由第二條方程得 $\mathrm{a}\left(1 + \frac{1}{x}\right) = 1$。所以

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

（這兩個函數在 $x = 0$ 時亦滿足上述兩個條件。）這條方程是線性而且齊次（homogeneous）的，而 x 與 e^{-x} 並非彼此的倍數，所以通解（general solution）是

$$u = Ax + B\mathrm{e}^{-x},$$

其中 A 和 B 為任意常數。

代換。設 $y = \dfrac{u'}{3u}$，其中 $u \neq 0$。由商法則（quotient rule），

$$\frac{\mathrm{d}y}{\mathrm{d}x} = \frac{u''}{3u} - \frac{(u')^2}{3u^2} = \frac{u''}{3u} - 3\left(\frac{u'}{3u}\right)^2 = \frac{u''}{3u} - 3y^2.$$

所以 $y' + 3y^2 = \dfrac{u''}{3u}$：平方項剛好完全抵銷，這正是代換中要放入因子 3 的原因。方程 $(*)$ 變成

$$\frac{u''}{3u} + \frac{x}{1+x}\cdot\frac{u'}{3u} = \frac{1}{3(1+x)},$$

乘以 $3u$ 得

$$\frac{\mathrm{d}^2u}{\mathrm{d}x^2} + \frac{x}{1+x}\frac{\mathrm{d}u}{\mathrm{d}x} - \frac{1}{1+x}u = 0,$$

正如所求。這就是本部分開首的方程，其中的 a 和 b 正是我們剛才求得的。

特解（particular solution）。因此 $u = Ax + B\mathrm{e}^{-x}$，而

$$y = \frac{u'}{3u} = \frac{A - B\mathrm{e}^{-x}}{3(Ax + B\mathrm{e}^{-x})}.$$

想一想 常數有兩個，條件卻只有一個。這會有問題嗎？

不會。把 u 乘以一個非零常數並不會改變 u'/u，所以只有比 $A : B$ 才有影響，而一階方程本來就應該只有一個自由常數。由條件「當 $x = 0$ 時 $y = 0$」得 $A - B = 0$。於是 $B = A \neq 0$（否則 $u \equiv 0$），約去 A 後，

$$y = \frac{1 - \mathrm{e}^{-x}}{3(x + \mathrm{e}^{-x})}.$$

分母永不為零，因為由 $\mathrm{e}^{-x} \geq 1 - x$ 可得 $x + \mathrm{e}^{-x} \geq 1$。第 (i) 部分至此完成。

第 (ii) 部分

現在 y^2 項的係數是 1，所以我們用 $y = \dfrac{u'}{u}$。與之前完全一樣，

$$y' = \frac{u''}{u} - \left(\frac{u'}{u}\right)^2 = \frac{u''}{u} - y^2,$$

所以 $y' + y^2 = u''/u$。把方程乘以 u，便得到線性方程

$$u'' + \frac{x}{1-x}u' - \frac{1}{1-x}u = 0.$$

這就是把第 (i) 部分的方程中的 x 換成 $-x$ 所得的方程。事實上，若 $v(x) = w(-x)$，其中 w 是第 (i) 部分方程的解，則 $v'(x) = -w'(-x)$ 及 $v''(x) = w''(-x)$，而第 (i) 部分的方程在點 $-x$ 處就是

$$v''(x) + \frac{-x}{1-x}\big(-v'(x)\big) - \frac{1}{1-x}v(x) = 0,$$

這正是新的方程。所以由 x 和 e^{-x} 可得出解 $-x$（或乾脆取 x）和 e^{x}。直接驗證：當 $u=x$ 時，$0+\frac{x}{1-x}-\frac{x}{1-x}=0$；當 $u=\mathrm{e}^{x}$ 時，$\mathrm{e}^{x}\left(1+\frac{x}{1-x}-\frac{1}{1-x}\right)=\mathrm{e}^{x}(1-1)=0$。因此

$$u=Ax+B\mathrm{e}^{x},\qquad y=\frac{u'}{u}=\frac{A+B\mathrm{e}^{x}}{Ax+B\mathrm{e}^{x}}.$$

當 $x=0$ 時，上式等於 $(A+B)/B$，而它必須等於 2，所以 $A=B$（且 $B\neq 0$）。因此

$$y=\frac{1+\mathrm{e}^{x}}{x+\mathrm{e}^{x}}.$$

這個解在包含 0、滿足 $x+\mathrm{e}^{x}>0$ 而且方程有定義的區間上成立，即 $x_0<x<1$，其中 $x_0\approx -0.567$ 是 $x+\mathrm{e}^{x}=0$ 的根。

驗算

- 當 $x=0$ 時，第 (ii) 部分的答案給出 $2/1=2$；對它求導得 $y'(0)=\frac{1\cdot 1-2\cdot 2}{1}=-3$，與方程吻合：$1-4-0=-3$。
- 當 $x\to\infty$ 時，第 (i) 部分的答案趨向 0，其表現有如 $\frac{1}{3x}$。這時 $y'\approx -\frac{1}{3x^2}$ 與 $3y^2\approx\frac{1}{3x^2}$ 互相抵銷，而 $\frac{x}{1+x}y\approx\frac{1}{3(1+x)}$，與 $(*)$ 一致。
- 我們以數值方法解了第 (ii) 部分的方程（由 $y(0)=2$ 出發，使用龍格－庫塔法（Runge–Kutta)），得到 $y(0.9)=1.0297654\ldots$，與公式的結果在小數點後 10 位之內相符。

常見錯誤

- 在第 (ii) 部分仍然保留因子 $\frac{1}{3}$：乘數必須與 y^2 的係數配合，所以這裏應該用 $y=u'/u$。
- 試圖只憑一個條件同時定出 A 和 B。其實只有兩者之比才會影響 y。
- 在第 (ii) 部分未經驗證便再次取 e^{-x}：$1-x$ 中的正負號改變，令它變成 e^{x}。

STEP 之外

形如 $y'+P(x)y^2+Q(x)y=R(x)$ 的方程稱為里卡蒂方程，而代換 $y=u'/(Pu)$ 總能把它化為一條線性二階方程。反過來說，任何二階線性方程的比 u'/u 都滿足一條里卡蒂方程，這就是為甚麼這類非線性方程在物理學中不斷出現，由量子力學（quantum mechanics）以至控制理論（control theory）都有它們的蹤影。Boyce 與 DiPrima 的 *Elementary Differential Equations*（常微分方程入門教材）涵蓋了這個課題，以及線性二階方程的理論（為甚麼兩個獨立的解就能給出通解）。

詳細解答

STEP III 2007，第 9 題

力學 • 簡諧運動

STEP III 2007，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/07-s3-q9

關鍵思路

彈簧對兩顆小珠的拉力或推力大小相等、方向相反，而且沿同一直線作用；圓環的反作用力（reaction）則沿徑向，所以對圓心的總角動量（angular momentum）一直保持為零：兩顆小珠的運動互為鏡像。這樣整個問題就簡化為只有一個坐標，再由能量守恆（conservation of energy）便得出關於 θ 的簡諧運動（simple harmonic motion，簡稱 SHM）。要留神的是，θ 是銳角的半角：如果兩顆小珠分開得夠遠，越過了直徑的兩端，能量的公式就會改變形式，這時的運動必須由兩段 SHM 拼合而成。

解答

建立模型

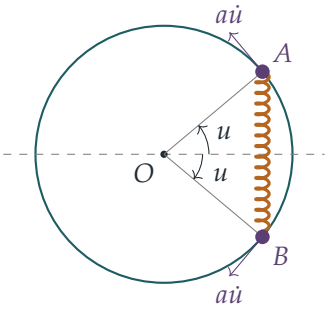
把小珠看成質點（particle），圓環固定而光滑（smooth），彈簧為輕（light）彈簧。作用於小珠的力有：重力和圓環反作用力的垂直分量（兩者互相抵消）、反作用力的水平分量（由於圓環光滑，它沿半徑方向），以及彈簧的力。彈簧作用於 A 的力和彈簧作用於 B 的力大小相等、方向相反，並且沿同一直線 AB 作用，所以它們對通過 O 的垂直軸的總力矩（moment）為零；徑向的力對該軸則完全沒有力矩。因此，對該軸的總角動量守恆，而由於小珠從靜止開始，總角動量為零。

從一條固定的半徑起沿圓環量度角度，設 A 和 B 所在的角度分別為 ϕ_A 和 ϕ_B。角動量為零即 $ma^2(\dot{\phi}_A + \dot{\phi}_B) = 0$，所以 $\phi_A + \phi_B$ 是常數。選取參考半徑，使它平分初始的角 AOB；那麼在任何時刻都有

$$\phi_A = u, \qquad \phi_B = -u,$$

其中只有一個坐標 u，而開始時 $u = \beta$。每顆小珠的速率都是 $a\dot{u}$，所以動能（kinetic

energy）是 $2\cdot\frac{1}{2}ma^2\dot{u}^2 = ma^2\dot{u}^2$。



當 $0 < u \le \frac{\pi}{2}$ 時，角 AOB 等於 $2u$；但如果 u 超過 $\frac{\pi}{2}$，小珠便已越過直徑的兩端，這時角 AOB 等於 $2\pi - 2u$。所以

$$\theta = u \quad (0 < u \le \tfrac{\pi}{2}), \qquad \theta = \pi - u \quad (\tfrac{\pi}{2} \le u < \pi).$$

小珠在 $u = \beta$ 處由靜止釋放，由能量守恆得

$$ma^2\dot{u}^2 + mk^2a^2(\theta-\alpha)^2 = mk^2a^2(\beta-\alpha)^2, \qquad \text{所以} \qquad \dot{u}^2 = k^2\left[(\alpha-\beta)^2 - (\theta-\alpha)^2\right].$$

「處於壓縮狀態」(in compression）是指彈簧比它的自然長度（natural length）短。能量在 $\theta = \alpha$ 時最小，所以那就是自然長度的位置；而長度 $2a\sin\theta$ 在 $(0, \frac{\pi}{2})$ 上隨 θ 增加而增加，因此 $0 < \beta < \alpha$。

$u \le \frac{\pi}{2}$ 時的運動

這時 $\theta = u$，把能量方程對 t 求導，再除以 $2\dot{u}$（只要小珠在運動，這一步就成立，其餘時刻則由連續性補足），得

$$\ddot{\theta} = -k^2(\theta - \alpha).$$

這是以 $\theta = \alpha$ 為中心、角頻率（angular frequency）為 k 的 SHM。由於小珠在 $\theta = \beta$ 處從靜止出發，

$$\theta - \alpha = -(\alpha - \beta)\cos kt,$$

所以兩顆小珠互相遠離，如果中途沒有任何事情打斷，便會在 $t = \frac{\pi}{k}$ 時到達 $\theta = 2\alpha - \beta$。一切都取決於 $2\alpha - \beta$ 是否小於 $\frac{\pi}{2}$，也就是取決於 β 與 $2\alpha - \frac{\pi}{2}$ 的大小關係。這個臨界值落在 $(0, \alpha)$ 之內，正是因為 $\frac{\pi}{4} < \alpha < \frac{\pi}{2}$，所以以下三種情況都有可能出現。

情況 1： $2\alpha - \frac{\pi}{2} < \beta < \alpha$

這時 $2\alpha - \beta < \frac{\pi}{2}$，所以小珠永遠不會到達直徑的兩端，全程都有 $\theta = u$，運動是在 $\theta = \beta$ 與 $\theta = 2\alpha - \beta$ 之間的純粹 SHM。週期（period）為

$$T = \frac{2\pi}{k}.$$

情況 2： $0 < \beta < 2\alpha - \frac{\pi}{2}$

這時 θ 在 SHM 掉頭之前就到達 $\frac{\pi}{2}$，所需時間 t_1 滿足

$$\tfrac{\pi}{2} - \alpha = -(\alpha - \beta)\cos kt_1, \qquad \text{所以} \qquad \cos kt_1 = -\lambda, \quad \text{其中}\lambda = \frac{\frac{\pi}{2} - \alpha}{\alpha - \beta}.$$

由於 $\alpha < \frac{\pi}{2}$ 且 $\beta < 2\alpha - \frac{\pi}{2}$，我們有 $0 < \lambda < 1$。因此

$$kt_1 = \pi - \arccos\lambda = \tfrac{\pi}{2} + \arcsin\lambda.$$

在這一刻，能量方程給出 $\dot{u}^2 = k^2[(\alpha - \beta)^2 - (\frac{\pi}{2} - \alpha)^2] > 0$，所以兩顆小珠仍在互相遠離，並會繼續越過直徑的兩端。

> **想一想** 小珠越過直徑的兩端之後，會發生甚麼事？

當 $u \geq \frac{\pi}{2}$ 時，我們有 $\theta = \pi - u$，所以能量方程變成

$$\dot{u}^2 = k^2\left[(\alpha - \beta)^2 - (u - (\pi - \alpha))^2\right],$$

這是 u 以 $u = \pi - \alpha$ 為中心的 SHM，角頻率同樣是 k。它是到達直徑之前那段運動關於 $u = \frac{\pi}{2}$ 的鏡像。所以小珠再用一段時間 t_1，便在 $u = \pi - \beta$ 處靜止（那裏再次有 $\theta = \beta$，只是小珠現在位於 O 的另一側），然後沿原路走完全程回到 $u = \beta$。經過四個這樣的階段後，小珠回到起始位置，所以運動的週期是

$$T = 4t_1 = \frac{4}{k}\left(\frac{\pi}{2} + \arcsin\frac{\frac{\pi}{2} - \alpha}{\alpha - \beta}\right) = \frac{2}{k}\left(\pi + 2\arcsin\frac{\pi - 2\alpha}{2(\alpha - \beta)}\right).$$

（至於角 θ 本身，它並不記錄小珠在 O 的哪一側，所以每隔 $2t_1$ 就重複一次。）

情況 3： $\beta = 2\alpha - \frac{\pi}{2}$

這時 θ 在 $t = \frac{\pi}{k}$ 到達 $2\alpha - \beta = \frac{\pi}{2}$，而 SHM 恰好在這一刻令小珠靜止。這時兩顆小珠位於直徑的兩端，所以彈簧沿 AOB 放置，它作用於每顆小珠的力都沿半徑方向：這個力沒有沿圓環方向的分量，而沿圓環方向也沒有其他力作用。小珠處於平衡（equilibrium）而且靜止，所以會一直留在那裏。因此

$$\beta = 2\alpha - \frac{\pi}{2}$$

就是不會出現振動的那個值。這與兩個週期的結果吻合：當 β 從下方趨近這個值時，$\lambda \to 1$，情況 2 的週期趨向 $\frac{4\pi}{k}$，即情況 1 週期的兩倍，因為小珠在直徑附近逗留的時間愈來愈長。

驗算

- 取 $k = 1$，$\alpha = \frac{3\pi}{8}$，$\beta = \frac{\pi}{8}$，由於 $2\alpha - \frac{\pi}{2} = \frac{\pi}{4}$，這屬於情況 2。於是 $\lambda = \frac{\pi/8}{\pi/4} = \frac{1}{2}$，$\arcsin\lambda = \frac{\pi}{6}$，而 $T = 2(\pi + \frac{\pi}{3}) = \frac{8\pi}{3} \approx 8.378$。

- 我們用數值方法逐步求解兩顆小珠的運動方程（能量寫成兩顆小珠位置的函數，並不假設任何對稱性）。取 $k = 1$，$\alpha = 1.2$，$\beta = 0.5$ 時，小珠在 8.5162 之後回到起點並靜止，而公式給出的是 8.51622；取 $\beta = 1.0$（情況 1）時，週期是 $6.2832 \approx 2\pi$。
- 量綱：k 的單位是 s^{-1}（能量等於 ma^2k^2 乘以角度的平方），所以 $\frac{2\pi}{k}$ 是時間。

常見錯誤

- 把動能寫成 $\frac{1}{2}ma^2\dot{\theta}^2$：兩顆小珠都在運動，所以動能是 $ma^2\dot{\theta}^2$。漏掉因子 2 會得出錯誤的角頻率 $\sqrt{2}k$。
- 毫無根據地假設小珠對稱地運動：這個結論來自對 O 的角動量守恆，而這需要彈簧作用於每顆小珠的力都沿同一直線 AB。
- 忘記 θ 是銳角，於是在所有情況下都答 $\frac{2\pi}{k}$。

STEP 之外

利用守恆量把兩個質點的系統簡化為一個坐標，正是拉格朗日力學（Lagrangian mechanics）的起點；在那裏，對稱性（在這題中，是把整個位形繞 O 旋轉）會帶來守恆定律，而諾特定理（Noether's theorem）把這個想法說得精確。情況 2 的運動是在「雙井」（double-well）位勢中的振動，它的週期取決於振幅，並且在能量等於中央勢壘高度之處出現跳躍。想進一步了解守恆定律和振動，可參閱 Kleppner 與 Kolenkow 的 *An Introduction to Mechanics*（力學入門教科書）。

詳細解答

STEP III 2007，第 10 題

力學 • 斜面上的拋體

STEP III 2007，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/07-s3-q10

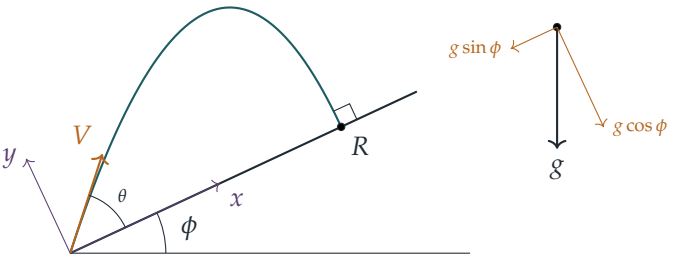
關鍵思路

取沿斜面及垂直於斜面的坐標軸，重力便化為兩個恆定的加速度：沿斜面向下的 $g\sin\phi$ 和指向斜面的 $g\cos\phi$。最特別的條件是質點沿原路折返。光滑平面只能沿垂直於自身的方向施力，所以只有當質點到達時正正垂直撞向平面，平面才能把速度完全反轉。兩個結果都由這一點得出；餘下的只是代數運算，以及一次利用算術平均—幾何平均不等式（AM–GM inequality）求最小值。

解答

運動方程

我們把質點（particle）視為只受重力作用的點質量，不計空氣阻力。重量 mg 鉛垂向下。沿斜面及垂直於斜面分解（resolve）：



加速度在 x 方向為 $-g\sin\phi$，在 y 方向為 $-g\cos\phi$，兩者皆為常數。以題目給定的初速度及起點 $(0,0)$ 積分兩次，得

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

沿原路折返的條件

質點首次回到平面，是在 $y = 0$ 且 $t > 0$ 之時，即

$$t = T = \frac{2V \sin\theta}{g \cos\phi}.$$

想一想 「沿同一路徑返回」說明了碰撞的甚麼性質？

反彈之後，質點仍然在恆定的重力下運動。若它要沿同一條拋物線（parabola）返回原點，反彈後一刻的速度必須與反彈前一刻的速度剛好相反。（不妨想像把影片倒過來播放。）我們假設平面是光滑（smooth）的，所以它給予質點的衝量（impulse）垂直於平面：沿斜面的分量 $\dot{x}$ 不會因反彈而改變。這個分量要被反轉，唯一的可能就是它本身為零。因此質點必須垂直撞擊平面，即在 $t = T$ 時 $\dot{x} = 0$。（我們還需要反彈把 $\dot{y}$ 反轉而不損失速率，所以碰撞必須是完全彈性（perfectly elastic）的。題目把這一點視為理所當然。）

現在 $\dot{x} = V\cos\theta - gt\sin\phi$，所以由 $\dot{x}(T) = 0$ 得

$$V\cos\theta = gT\sin\phi = \frac{2V\sin\theta\sin\phi}{\cos\phi}.$$

兩邊除以 $V\cos\theta$（由於 $0 < \theta < \pi/2$，此數為正），

$$2\tan\phi\tan\theta = 1,$$

即為所求。

射程

由 $\dot{x}(T) = 0$ 可得較簡潔的形式 $T = \dfrac{V\cos\theta}{g\sin\phi}$。於是

$$\begin{aligned} R = x(T) &= V\cos\theta \cdot \frac{V\cos\theta}{g\sin\phi} - \frac{1}{2}g\sin\phi \cdot \frac{V^2\cos^2\theta}{g^2\sin^2\phi} \\ &= \frac{V^2\cos^2\theta}{g\sin\phi} - \frac{V^2\cos^2\theta}{2g\sin\phi} = \frac{V^2\cos^2\theta}{2g\sin\phi}, \end{aligned}$$

這就是所求的結果。（換一個說法：x 在 $\dot{x} = 0$ 時最大，所以落點就是這個運動沿斜面向上所能到達的最高點。）

關於 $2V^2/(gR)$ 的恆等式

由射程公式，

$$\frac{2V^2}{gR} = \frac{4\sin\phi}{\cos^2\theta} = 4\sin\phi\,(1 + \tan^2\theta).$$

利用 $\tan\theta = \dfrac{1}{2\tan\phi}$，即 $\tan^2\theta = \dfrac{\cos^2\phi}{4\sin^2\phi}$，

$$\frac{2V^2}{gR} = 4\sin\phi + \frac{\cos^2\phi}{\sin\phi} = 4\sin\phi + \frac{1-\sin^2\phi}{\sin\phi} = 3\sin\phi + \operatorname{cosec}\phi.$$

最大射程

對給定的 V，R 在 $3\sin\phi + \operatorname{cosec}\phi$ 最小時最大。設 $s = \sin\phi$；由於 $0 < \phi < \pi/2$，故 $0 < s < 1$。對正數 $3s$ 和 $1/s$ 應用算術平均—幾何平均不等式，

$$3s + \frac{1}{s} \geq 2\sqrt{3s \cdot \frac{1}{s}} = 2\sqrt{3},$$

等號成立當且僅當 $3s = 1/s$，即 $s = 1/\sqrt{3}$。這個值確實落在 $(0,1)$ 之內，所以這個界可由一個真實存在的斜面達到。因此

$$\frac{2V^2}{gR} \geq 2\sqrt{3} \quad\Longrightarrow\quad R \leq \frac{V^2}{\sqrt{3}\,g},$$

等號在 $\sin\phi = 1/\sqrt{3}$ 時成立。所以最大可能射程為 $V^2/(\sqrt{3}\,g)$。

> ? **想一想** 在最大值處，條件 $\phi + \theta < \pi/2$ 是否成立？

成立，而且事實上任何時候都成立。由 $\tan\theta\tan\phi = \frac{1}{2}$ 及兩角皆為銳角，

$$\tan(\phi+\theta) = \frac{\tan\phi + \tan\theta}{1 - \tan\phi\tan\theta} = 2(\tan\phi + \tan\theta) > 0,$$

而由於 $0 < \phi + \theta < \pi$，這就迫使 $\phi + \theta < \pi/2$。在最大值處，$\sin\phi = 1/\sqrt{3}$ 給出 $\tan\phi = 1/\sqrt{2}$，所以 $\tan\theta = 1/\sqrt{2}$ 亦然：最佳情況是 $\theta = \phi \approx 35.3°$。

驗算

- 單位：V^2/g 是長度，與射程應有的量綱相符。
- 極限情況：當 $\phi \to 0$ 時，$\operatorname{cosec}\phi \to \infty$ 而 $R \to 0$。在近乎水平的平面上，要垂直落在平面上，唯一的方法是幾乎筆直向上發射，射程自然接近零。當 $\phi \to \pi/2$ 時，平面成為一幅鉛垂的牆，$\theta \to 0$，且 $2V^2/(gR) \to 4$，所以 $R \to V^2/(2g)$：這正是質點貼着牆筆直向上拋出所達到的高度。兩者都合理。
- 數值驗算：取 $g = 9.8$、$V = 10$ 及 $\phi = 0.7$，逐步向前推算運動，落點為 $x \approx 5.8561$，與公式吻合。落地時的速度有 $\dot{x} \approx 0$，把它反轉後質點便回到原點。對 ϕ 作網格搜尋，在 $V = 1$ 時得最大值 $R \approx 0.05891$，等於 $1/(9.8\sqrt{3})$。

常見錯誤

- 把重力的兩個分量弄反了。沿斜面的分量是 $g\sin\phi$，垂直於斜面的分量是 $g\cos\phi$。可用 $\phi = 0$ 檢查。
- 說質點必須「水平地」落下，或在「最高點」落下。正確的條件是它垂直撞擊平面，即在傾斜坐標軸中 $\dot{x} = 0$。
- 忘記檢查 $\sin\phi = 1/\sqrt{3}$ 能否達到。算術平均—幾何平均不等式只給出一個界，只有當等號確實可以成立時，它才是最大值。

STEP 之外

質點能沿原路折返，靠的是牛頓力學的時間可逆性（time-reversibility）。若力只取決於位置，把所有速度反轉，系統便會沿着自己的歷史倒退回去。選取配合幾何形狀的坐標軸，而不是慣用的水平和鉛垂方向，是你在大學力學中會經常用到的習慣。Kleppner 與 Kolenkow 的 *An Introduction to Mechanics*（力學入門）對這兩個想法都有詳細闡述。

詳細解答

STEP III 2007，第 11 題

力學 • 滾過邊緣

STEP III 2007，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/07-s3-q11

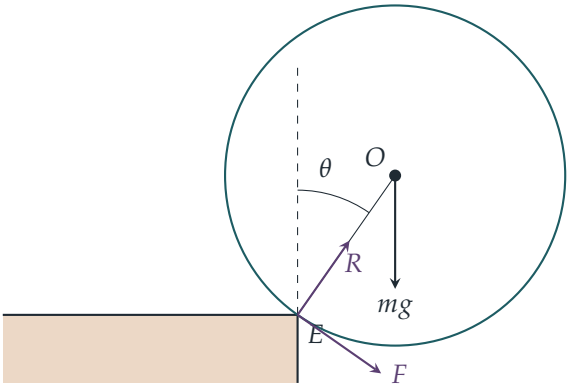
關鍵思路

輪緣一到達桌邊，桌邊便成為支點（pivot）：輪緣上與桌邊接觸的那一點是靜止的，所以輪心沿着以桌邊為圓心、半徑為 a 的圓運動。我們把桌邊的接觸力分解為沿半徑方向的分量 R（法向反作用力，normal reaction）和沿輪緣方向的分量 F（摩擦力，friction），再與質心（centre of mass）的運動作比較。整道題的關鍵在於質量位於何處。在 (i) 中，全部質量都在輪心，所以輪子對其中心的轉動慣量（moment of inertia）為零。在 (ii) 中，質量分佈在輪緣上，摩擦力正是由此而生。

解答

模型與記號。輪緣（rim）、輪輻（spokes）和圓環（hoop）都是輕而剛硬的；輪轂（hub）和兩個質點（particles）都視為質點。設 E 為桌邊上與輪緣接觸的點，O 為輪心，θ 為 EO 與豎直向上方向所成的角，所以在輪子到達桌邊的一刻 $\theta = 0$。只要輪緣與 E 保持接觸而不滑動，輪緣上位於 E 的那個物質點便是靜止的。因此這個剛體（rigid body）繞 E 轉動，角速率（angular speed）為 $\dot{\theta}$，而 O 沿着以 E 為圓心、半徑為 a 的圓運動，速率為 $a\dot{\theta}$。

在 E 的接觸力有一個沿 EO 方向（背離 E）的分量 R，以及一個垂直於 EO 的分量 F，後者以 θ 增加的方向為正。由於 E 靜止，接觸力不作功，所以能量守恆（conservation of energy）。



圖中 R 沿 EO 作用，而 F 則按其正方向繪出。

質心 G 與 E 的距離為 a，其加速度沿運動方向的分量為 $a\ddot{\theta}$，指向 E 的分量為 $a\dot{\theta}^2$。分解作用於整個物體的外力（重力 mg 和接觸力）：

$$\text{沿運動方向：}\quad ma\ddot{\theta} = mg\sin\theta + F, \tag{1}$$

$$\text{指向}E\text{：}\quad ma\dot{\theta}^2 = mg\cos\theta - R. \tag{2}$$

在兩部分中，質心都位於 O（即輪轂；亦即兩個質點的中點），所以 (1) 和 (2) 在兩部分都成立。

第 (i) 部分

輪子在桌面上以恆定速率 u 運動，所以在 $\theta = 0$ 時有 $\dot{\theta} = u/a$。唯一的動能是輪轂的動能 $\frac{1}{2}ma^2\dot{\theta}^2$，而輪轂下降了 $a(1-\cos\theta)$，所以

$$a^2\dot{\theta}^2 = u^2 + 2ga(1-\cos\theta).$$

對時間求導，並約去 $2a\dot{\theta}$（它不為零，因為 $\dot{\theta} \geq u/a > 0$）：

$$a\ddot{\theta} = g\sin\theta.$$

代入 (1) 得 $F = 0$：**摩擦力為零**，正如所求。

這個結果還有一個直接的理由。輪緣和輪輻都是輕的，所以作用於它們的各力對 O 的總力矩（moment）必須為零。這些力是在 E 的接觸力，以及輪轂在 O 施加的力，而後者對 O 沒有力矩。接觸力對 O 的力矩是 Fa，所以 $F = 0$。

由 (2)，法向反作用力為

$$R = mg\cos\theta - \frac{m}{a}\left(u^2 + 2ga(1-\cos\theta)\right) = mg(3\cos\theta - 2) - \frac{mu^2}{a}.$$

在 $\theta = 0$ 時，這個值是 $m(g - u^2/a) > 0$，因為 $u^2 < ag$，所以輪子不會立即離開桌邊。其後 R 隨 θ 增加而減少，當 $R = 0$ 時輪緣便失去接觸，即當

$$\cos\theta = \frac{2ag + u^2}{3ag},$$

這個值小於 1，因為 $u^2 < ag$。此時輪轂已下降了

$$a(1 - \cos\theta) = a \cdot \frac{3ag - 2ag - u^2}{3ag} = \frac{ag - u^2}{3g},$$

正如所求。由於自始至終 $F = 0$，不滑動的假設根本無須依靠摩擦力來維持，所以以上分析一直到這一點都成立。

第 (ii) 部分

想一想 圓環繞 E 轉動時，兩個質點在哪裏？

在 $\theta = 0$ 時，通過兩個質點的直徑是豎直的，所以下方的質點位於接觸點 E。其後圓環作為剛體繞着位於 E 的物質點轉動，而這個物質點正是該質點，所以該質點一直留在 E。另一個質點則一直留在直徑的另一端，沿 EO 方向與 E 相距 $2a$。

因此對 E 的轉動慣量是 $\frac{m}{2}(2a)^2 = 2ma^2$。圓環在桌面上以恆定速率 u 滾動（沒有任何力作功），所以在 $\theta = 0$ 時 $\dot\theta = u/a$。運動中的質點比 E 高出 $2a\cos\theta$，由能量守恆得

$$\tfrac{1}{2}(2ma^2)\dot\theta^2 + \tfrac{m}{2}g(2a\cos\theta) = \tfrac{1}{2}(2ma^2)\frac{u^2}{a^2} + \tfrac{m}{2}g(2a), \qquad \text{所以} \qquad a^2\dot\theta^2 = u^2 + ga(1 - \cos\theta).$$

求導並約去 $2a\dot\theta \neq 0$，得 $a\ddot\theta = \frac{1}{2}g\sin\theta$。於是由 (1) 得

$$F = ma\ddot\theta - mg\sin\theta = -\tfrac{1}{2}mg\sin\theta.$$

所以對每一個 $\theta > 0$，**摩擦力都不為零**。（它的大小是 $\frac{1}{2}mg\sin\theta$，方向與 O 的運動方向相反。）與 (i) 不同，現在輪緣帶有質量，所以圓環需要一個對 O 的力矩（torque），才能使它的轉動跟得上 O 的運動，而只有摩擦力能夠提供這個力矩。

由 (2)，

$$R = mg\cos\theta - \frac{m}{a}\left(u^2 + ga(1 - \cos\theta)\right) = mg(2\cos\theta - 1) - \frac{mu^2}{a}.$$

在 $\theta = 0$ 時，$R = m(g - u^2/a) > 0$ 而 $F = 0$，所以越過桌邊的運動確實是在不滑動的情況下開始的。假如圓環一直保持接觸而不滑動，R 便會在以下角度降至零：

$$\cos\theta_1 = \frac{ag + u^2}{2ag} < 1,$$

這裏用到了 $u^2 < ag$。所以 $\theta_1 > 0$，而在該處 $|F| = \frac{1}{2}mg\sin\theta_1 > 0$。

想一想 為甚麼這樣就必定會在失去接觸之前滑動？

設摩擦系數（coefficient of friction）為 μ，不滑動的條件是 $|F| \le \mu R$。當 $\theta \to \theta_1$（從下方趨近）時，$R \to 0^+$ 但 $|F| \to \frac{1}{2}mg\sin\theta_1 > 0$，所以 $|F|/R \to \infty$。因此無論 μ 取甚麼（有限）值，總有一個角 $\theta < \theta_1$，在 R 仍然為正的時候，$|F|$ 就必須超過 μR。所以圓環**在與桌子失去接觸之前便會滑動**。

驗算

- 在 (i) 中，若 $u^2 \to ag$，下降的距離趨向 0：此時輪轂在 $\theta = 0$ 已處於即將飛離的臨界狀態，該處 $g = u^2/a$。若 $u \to 0$，下降的距離是 $a/3$，這正是質點從光滑球面頂點由靜止開始滑下的那個熟悉結果。
- 取 $a = 1\,\mathrm{m}$、$u = 2\,\mathrm{m\,s^{-1}}$ 和 $g = 9.81\,\mathrm{m\,s^{-2}}$，以數值方法逐步求解方程，在 (i) 中得出下降距離為 $0.1974\,\mathrm{m}$，與 $(ag - u^2)/(3g)$ 吻合。在 (ii) 中我們得出 $\cos\theta_1 = 0.7039$，而且 $|F|/R$ 在臨近 $R = 0$ 之前已超過 10^4。
- 在 (ii) 中直接由兩個質點的加速度計算接觸力（$m\ddot{\mathbf{r}}_G$ 減去重力），所得的 F 和 R 都與上面相同。

常見錯誤

- 在 (ii) 中使用對 O 的滾動動能 $\frac{1}{2}mv^2 + \frac{1}{2}I\omega^2$，卻忘記了一旦開始繞支點轉動，其中一個質點便固定在 E。對 E 來計算會簡潔得多。
- 在桌邊沿水平和豎直方向分解力。在邊角處，「法向」是沿半徑 EO 的方向，而不是豎直方向。
- 僅僅因為 $F \neq 0$ 便斷定「它會滑動」。論證需要的是 $R \to 0$ 而 $|F|$ 始終不趨近零。

STEP 之外

兩部分的對比在於質量的分佈：同樣的質心運動，會因轉動慣量不同而需要不同的接觸力，因為角動量（angular momentum）方程和線動量（linear momentum）方程必須同時滿足。在大學第一門力學課程中，這就發展成剛體動力學（rigid-body dynamics）：物體的運動由 $\mathbf{F} = m\ddot{\mathbf{r}}_G$ 連同對 G 的力矩方程決定，而接觸約束（滾動、滑動、失去接觸）則要用不等式 $R \ge 0$ 和 $|F| \le \mu R$ 來檢驗。Kleppner 與 Kolenkow 的 *An Introduction to Mechanics*（力學入門教材）對這兩條方程都有透徹的講解。

詳細解答

STEP III 2007，第 12 題

概率與統計 • 隨機和

> **STEP III 2007，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/07-s3-q12

關鍵思路

Y 是一個項數隨機的和，所以不能直接套用固定項數的變量之和的法則。竅門是以 N 為條件（condition）：一旦知道 $N = k$，Y 就是 k 個獨立（independent）變量的普通和，其平均值（mean）和方差（variance）我們都知道。然後再對 N 的分佈取平均，這正是題目第一部分要求 $\mathrm{E}(N)$ 和 $\mathrm{E}(N^2)$ 的原因。

解答

N 的分佈

我們假設該數是均勻地隨機選取的，因此對 $k = 1, 2, \ldots, 2n-1$ 有 $\mathrm{P}(N = k) = \frac{1}{2n-1}$。利用 $\sum_{k=1}^{m} k = \frac{1}{2}m(m+1)$ 和 $\sum_{k=1}^{m} k^2 = \frac{1}{6}m(m+1)(2m+1)$，並取 $m = 2n-1$：

$$\mathrm{E}(N) = \frac{1}{2n-1} \cdot \frac{(2n-1)(2n)}{2} = n,$$
$$\mathrm{E}(N^2) = \frac{1}{2n-1} \cdot \frac{(2n-1)(2n)(4n-1)}{6} = \frac{n(4n-1)}{3}.$$

（平均值 n 亦可由對稱性直接看出：n 正是 $1, \ldots, 2n-1$ 的中間一項。）以下會用到

$$\mathrm{Var}(N) = \frac{n(4n-1)}{3} - n^2 = \frac{n^2 - n}{3} = \frac{n(n-1)}{3}.$$

以 N 為條件

固定 k，並假設 $N = k$。由於各 X_i 與 N 獨立，以 $N = k$ 為條件並不會改變它們的分佈，所以在已知 $N = k$ 的情況下，$Y = X_1 + \cdots + X_k$，即 k 個獨立變量之和，每個變量的平

均值為 μ，方差為 σ^2。因此

$$\mathrm{E}(Y \mid N=k) = k\mu, \qquad \mathrm{Var}(Y \mid N=k) = k\sigma^2,$$

其中第二式用到各 X_i 互相獨立（方差可以相加）。於是

$$\mathrm{E}(Y^2 \mid N=k) = \mathrm{Var}(Y \mid N=k) + \big(\mathrm{E}(Y \mid N=k)\big)^2 = k\sigma^2 + k^2\mu^2.$$

想一想 怎樣由條件期望值（conditional expectation）回到普通的期望值？

根據全期望定律（law of total expectation，又稱分割定理 partition theorem）：$\mathrm{E}(g(Y,N)) = \sum_k \mathrm{P}(N=k)\,\mathrm{E}(g(Y,N) \mid N=k)$。在每一項之內，$N$ 只不過是數字 k。

Y 的平均值

$$\mathrm{E}(Y) = \sum_k \mathrm{P}(N=k)\,k\mu = \mu\,\mathrm{E}(N) = n\mu,$$

正是所求。

協方差

已知 $N=k$ 時，$YN = kY$，所以 $\mathrm{E}(YN \mid N=k) = k\cdot k\mu = k^2\mu$。因此 $\mathrm{E}(YN) = \mu\,\mathrm{E}(N^2)$，而協方差（covariance）為

$$\begin{aligned}\mathrm{Cov}(Y,N) &= \mathrm{E}(YN) - \mathrm{E}(Y)\,\mathrm{E}(N) = \mu\,\mathrm{E}(N^2) - n\mu\cdot n \\ &= \mu\,\mathrm{Var}(N) = \tfrac{1}{3}n(n-1)\mu,\end{aligned}$$

這就是所要求的結果。

Y 的方差

把 $\mathrm{E}(Y^2 \mid N=k) = k\sigma^2 + k^2\mu^2$ 對 N 取平均：

$$\mathrm{E}(Y^2) = \sigma^2\,\mathrm{E}(N) + \mu^2\,\mathrm{E}(N^2) = n\sigma^2 + \frac{n(4n-1)}{3}\mu^2.$$

因此

$$\mathrm{Var}(Y) = n\sigma^2 + \frac{n(4n-1)}{3}\mu^2 - n^2\mu^2 = n\sigma^2 + \frac{n(n-1)}{3}\mu^2.$$

這兩項各有自然的含義：$n\sigma^2$ 是試驗次數固定時 Y 的平均離散程度，而 $\frac{1}{3}n(n-1)\mu^2 = \mu^2\,\mathrm{Var}(N)$ 則是因為不知道試驗次數而額外產生的離散程度。

驗算

- 若 $n = 1$，則必有 $N = 1$，各公式化為 $\mathrm{E}(Y) = \mu$、$\mathrm{Cov}(Y, N) = 0$ 和 $\mathrm{Var}(Y) = \sigma^2$，與單獨一個 X_1 的情況相符。
- 若 $\mu = 0$，則額外的一項消失：項數不再影響平均值，所以它既不會增加離散程度，也不會與 Y 相關。
- 取 $n = 2$，並設每個 X_i 各以概率 $\frac{1}{2}$ 取 0 或 1。那麼 $N \in \{1, 2, 3\}$，而 $Y \mid N \sim \mathrm{B}(N, \frac{1}{2})$；公式給出 $\mathrm{Var}(Y) = 2 \cdot \frac{1}{4} + \frac{2}{3} \cdot \frac{1}{4} = \frac{2}{3}$。我們已逐一列舉 (N, Y) 的聯合分佈（joint distribution），就 $n = 1, \ldots, 6$ 以及 X_i 的幾種不同分佈，核實了這個數值和所有結果。

常見錯誤

- 只寫 $\mathrm{Var}(Y) = \mathrm{E}(N)\sigma^2$，忘記了 N 本身也會變動。第二項才是整題的重點。
- 把 Y 和 N 當作獨立，得出 $\mathrm{Cov}(Y, N) = 0$。當 $\mu > 0$ 時，試驗次數越多，和就越大，所以兩者是正相關的。

STEP 之外

Y 是一個*隨機和*（random sum，即複合分佈 compound distribution），而方差的結果是全方差定律（law of total variance）的一個特例：$\mathrm{Var}(Y) = \mathrm{E}(\mathrm{Var}(Y \mid N)) + \mathrm{Var}(\mathrm{E}(Y \mid N))$，在這裏即 $\sigma^2\mathrm{E}(N) + \mu^2\mathrm{Var}(N)$。隨機和可用作保險索償和分支過程（branching processes）的模型，而用概率生成函數（probability generating functions）處理則十分簡潔：$G_Y(s) = G_N(G_X(s))$。Grimmett 與 Welsh 的 *Probability: An Introduction*（概率論入門）正是以這種方法處理隨機和與分支過程。

詳細解答

STEP III 2007，第 13 題

概率與統計 • 隨機漫步的期望值

STEP III 2007，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：

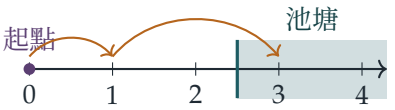
step.maths.org/questions/07-s3-q13

關鍵思路

那半米的作用，是消除「剛好落在池邊」的任何疑問：青蛙由距離池邊 $(n-\frac{1}{2})$ m 處出發，牠身處池中，當且僅當牠已跳的總距離至少為 n。因此整個問題的核心，就是由 1 和 2 組成的累計總和首次達到 n 的時刻。當 n 較小時，我們可以直接列出所有可能情況；至於一般公式，u_n 的三個值可給出三條方程，用以求出三個常數。

解答

設 S_j 為首 j 次跳躍所跳的總距離（以米為單位），並設 T 為到達池塘所需的跳躍次數。青蛙由距離池邊 $(n-\frac{1}{2})$ m 處出發，跳了 j 次後身處池中，當且僅當 $S_j > n-\frac{1}{2}$，亦即 $S_j \geq n$，因為 S_j 是整數。所以 $p_n(j) = \mathrm{P}(T=j)$，其中 T 是首個滿足 $S_j \geq n$ 的 j。



圖中所示為 $n=3$ 的情況（池邊距離起點 2.5 m），以及「短跳、長跳」這個跳躍序列，青蛙跳兩次便到達池塘。

第 (i) 部分

當 $n=2$ 時，池邊距離起點 1.5 m。若第一跳是長跳（概率（probability）為 q），青蛙便直接跳進池中，所以 $p_2(1)=q$。否則第一跳是短跳（概率為 p），青蛙距離池邊 0.5 m，

而第二跳無論長短都會把牠帶進池中。各次跳躍互相獨立（independent），所以

$$p_2(2) = p \times 1 = p,$$

這正是所要證明的結果。（此外，當 $j \geq 3$ 時 $p_2(j) = 0$，而 $p + q = 1$ 則確認了這個分佈（distribution）是完整的。）

第 (ii) 部分

由距離池邊 $\frac{1}{2}$ m 處出發，任何一跳都會落入池中，所以 $u_1 = 1$。

利用第 (i) 部分，

$$u_2 = 1 \cdot p_2(1) + 2 \cdot p_2(2) = q + 2p = 2 - q.$$

當 $n = 3$ 時，我們需要 $S_j \geq 3$。

- 只跳一次永遠不夠，因為 $S_1 \leq 2$：$p_3(1) = 0$。
- 跳兩次仍不夠，只會在兩次都是短跳時發生（$S_2 = 2$）；其餘每一種組合都有 $S_2 \geq 3$。所以 $p_3(2) = 1 - p^2$。
- 兩次短跳之後，青蛙距離池邊 0.5 m，第三跳必定落入池中：$p_3(3) = p^2$。

因此

$$\begin{aligned} u_3 &= 2(1 - p^2) + 3p^2 = 2 + p^2 \\ &= 2 + (1 - q)^2 = 3 - 2q + q^2, \end{aligned}$$

這正是所要證明的結果。

想一想 有沒有更快的方法看出這些值？這個方法能否推廣至所有 n？

有：以第一跳的結果作條件（conditioning）來考慮。第一跳之後，青蛙的處境與原來相同，只是 n 減少了 1 或 2，所以當 $n \geq 2$ 時

$$u_n = 1 + p\,u_{n-1} + q\,u_{n-2}, \qquad u_0 = 0,$$

其中 $u_0 = 0$ 對應由 1.5 m 處出發而第一跳是長跳的情況。由此再次得出 $u_2 = 1 + p$ 和 $u_3 = 1 + p(1 + p) + q = 2 + p^2$。這也解釋了第 (iii) 部分的形式為何正確：輔助方程（auxiliary equation）$\lambda^2 - p\lambda - q = 0$ 可因式分解為 $(\lambda - 1)(\lambda + q) = 0$，由此得出 $(-q)^{n-1}$ 這一項。而重根 1（源自那個「+1」）則解釋了 $B + Cn$ 的出現。

第 (iii) 部分

把 $n = 1, 2, 3$ 代入 $u_n = A(-q)^{n-1} + B + Cn$，並利用第 (ii) 部分：

$$A + B + C = 1, \tag{1}$$

$$-qA + B + 2C = 2 - q, \tag{2}$$

$$q^2A + B + 3C = 3 - 2q + q^2. \tag{3}$$

兩式相減以消去 B：

$$(2) - (1): \quad -(1+q)A + C = 1 - q, \tag{4}$$

$$(3) - (2): \quad q(1+q)A + C = 1 - q + q^2. \tag{5}$$

現在 $q \times (4) + (5)$ 可消去 A：

$$(1+q)C = q - q^2 + 1 - q + q^2 = 1, \qquad \text{所以} \qquad C = \frac{1}{1+q},$$

這正是所要證明的結果。然後由 (4) 得

$$(1+q)A = \frac{1}{1+q} - (1-q) = \frac{1-(1-q^2)}{1+q} = \frac{q^2}{1+q}, \qquad \text{所以} \qquad A = \frac{q^2}{(1+q)^2},$$

再由 (1) 得

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

所以

$$u_n = \frac{q^2(-q)^{n-1} + q + (1+q)n}{(1+q)^2}.$$

n 很大時。 由於 $0 \le q \le 1$，$A(-q)^{n-1}$ 和 B 這兩項都是有界（bounded）的（各自的絕對值最多為 $\frac{1}{4}$），而 Cn 則無止境地增大。因此當 n 很大時

$$u_n \approx Cn = \frac{n}{1+q} = \frac{n}{p+2q},$$

這裏用到 $1 = p + q$，故 $1 + q = p + 2q$。

為何這在意料之中。 每一跳的平均長度是 $1 \cdot p + 2 \cdot q = p + 2q$ 米。青蛙須跳過大約 n 米，而根據大數定律（law of large numbers），當 N 很大時，N 次跳躍的總長度接近 $N(p+2q)$。所以所需的跳躍次數接近 $n/(p+2q)$。這個近似相當準確，因為起點那半米的偏移，以及終點處多跳出的距離（青蛙落點最多超出池邊 1.5 m）都是有界的。它們只會令次數改變一個有界的量，與 n 相比微不足道。這個有界的修正，正正就是 $A(-q)^{n-1} + B$ 這部分。

驗算

- 在公式中取 $n = 0$，得 $\left(q^2 \cdot (-1/q) + q\right)/(1+q)^2 = 0 = u_0$，與應有的結果相符，而且公式滿足上面的遞推關係（recurrence）。
- $q = 0$（所有跳躍都是短跳）：$u_n = n$，這是正確的，因為青蛙恰好需要 n 次單位長度的跳躍。
- 以分數精確枚舉所有跳躍序列，結果在 $n = 1, \ldots, 12$ 及若干個 p 值下均與公式吻合。以 $q = 0.7$、$n = 200$ 進行模擬，得出平均跳躍次數為 117.90，而公式給出 117.89，$n/(p+2q) \approx 117.65$。

常見錯誤

- 把「落入池中」當作恰好 $S_j = n$。由於有那半米，青蛙可以跳過頭，凡是滿足 $S_j \geq n$ 的序列都要計算在內。
- 寫成 $p_3(2) = 2pq + q^2$ 時，忘記了長跳在先的情況。事實上 $pq + qp + q^2 = 1 - p^2$ 已涵蓋兩種次序。

STEP 之外

u_n 的遞推關係是一條更新方程（renewal equation）。結果 $u_n \sim n/\mu$（其中 μ 為每步的平均長度）就是基本更新定理（elementary renewal theorem）。由沃爾德恆等式（Wald's identity）$\mathrm{E}(S_T) = \mu\,\mathrm{E}(T)$ 可得出一個精確的版本：這裏 S_T 不是 n 就是 $n+1$，所以 $(1+q)u_n = n + \mathrm{P}(\text{跳過頭})$。Grimmett 與 Welsh 合著的 *Probability: An Introduction*（概率論入門）以同樣的思路講述隨機漫步（random walk）和母函數（generating function）。

詳細解答

STEP III 2007，第 14 題

概率與統計 • 幾何概率

STEP III 2007，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/07-s3-q14

關鍵思路

包圍所有洞孔的最小同心圓（或正方形），由最遠的那個洞孔決定，所以我們要找的是最大值的分佈，而 $P(\max \le r)$ 不過是單次投擲的概率（probability）的 n 次方。在圓盤上，一個洞孔與 O 的距離的平方 R^2 原來在 $[0,1]$ 上服從均勻分佈（uniform distribution），於是一切都變得簡潔。在正方形上，合適的「距離」是 $\max(|x|,|y|)$，它的表現與圓盤上的半徑完全一樣。這個巧合也令第 (iii) 部分幾乎不費吹灰之力就解決了。

解答

第 (i) 部分

設 n 個洞孔與 O 的距離為 $R_1,\dots,R_n$。由於各次投擲互相獨立，這些距離亦互相獨立（independent）。對單次投擲及 $0 \le r \le 1$，洞孔與 O 的距離不超過 r，當且僅當它落在半徑為 r 的圓盤內，所以

$$P(R_i \le r) = \frac{\pi r^2}{\pi \cdot 1^2} = r^2.$$

設 $U_i = R_i^2$。那麼對 $0 \le u \le 1$，$P(U_i \le u) = P(R_i \le \sqrt{u}) = u$，所以每個 U_i 都在 $[0,1]$ 上服從均勻分佈，而且各 U_i 互相獨立。

以 O 為圓心並包圍所有洞孔的最小圓，半徑為 $M = \max_i R_i$，面積為 $\pi M^2 = \pi \max_i U_i$。記 $V = \max_i U_i$。由獨立性，

$$P(V \le v) = P(U_1 \le v) \cdots P(U_n \le v) = v^n \qquad (0 \le v \le 1),$$

所以 V 的概率密度函數（density）為 nv^{n-1}，而

$$E(V) = \int_0^1 v \cdot nv^{n-1}\,\mathrm{d}v = \frac{n}{n+1}.$$

因此，包圍全部 n 個洞孔的最小圓的期望（expected）面積為

$$\boxed{\frac{n\pi}{n+1}}.$$

至於包圍最接近 O 的 $n-1$ 個洞孔的圓，其半徑是第二大的距離，所以面積為 πW，其中 W 是 $U_1,\ldots,U_n$ 之中第二大的值。

想一想 怎樣求出第二大的值的分佈？

$W \le w$ 當且僅當各 U_i 之中最多只有一個超過 w。要麼沒有任何一個超過（概率為 w^n），要麼恰好有一個超過：選出是哪一個有 n 種方法，得出概率 $n(1-w)w^{n-1}$。所以

$$P(W \le w) = w^n + n(1-w)w^{n-1} = nw^{n-1} - (n-1)w^n.$$

我們不必求導，而可以利用 $E(W) = \int_0^1 P(W > w)\,\mathrm{d}w$，這條公式對任何取值於 $[0,1]$ 的隨機變量（random variable）都成立：

$$E(W) = \int_0^1 \left(1 - nw^{n-1} + (n-1)w^n\right)\mathrm{d}w = 1 - 1 + \frac{n-1}{n+1} = \frac{n-1}{n+1}.$$

所以，包圍最近的 $n-1$ 個洞孔的圓的期望面積為

$$\boxed{\frac{(n-1)\pi}{n+1}}.$$

（這符合一個漂亮的規律：n 個獨立的均勻分佈點把 $[0,1]$ 分成 $n+1$ 段間隙，每段的期望長度同為 $\frac{1}{n+1}$，所以第 k 小的點的平均值為 $\frac{k}{n+1}$。）

第 (ii) 部分

取通過 Q 且與鏢靶各邊平行的坐標軸，於是鏢靶就是 $-1 \le x, y \le 1$。我們取包圍洞孔的正方形，使其各邊與鏢靶的邊平行。以 Q 為中心、半邊長為 s 的正方形是 $\{|x| \le s,\ |y| \le s\}$，所以它包含洞孔 (x,y) 當且僅當 $\max(|x|,|y|) \le s$。

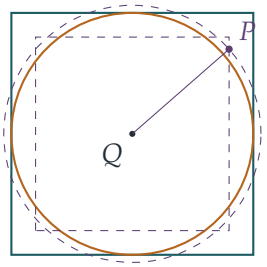
對第 i 個洞孔，設 $S_i = \max(|x_i|,|y_i|)$。對 $0 \le s \le 1$，

$$P(S_i \le s) = \frac{(2s)^2}{2^2} = s^2,$$

與第 (i) 部分的 $P(R_i \le r) = r^2$ 相同。包圍所有洞孔的最小正方形，半邊長為 $\max_i S_i$，面積為 $(2\max_i S_i)^2 = 4\max_i S_i^2$。與第 (i) 部分完全一樣，各 S_i^2 互相獨立，並在 $[0,1]$ 上服從均勻分佈，所以它們的最大值的期望值為 $\frac{n}{n+1}$。包圍所有洞孔的最小正方形的期望面積為

$$\boxed{\frac{4n}{n+1}}.$$

第 (iii) 部分



現在鏢靶是正方形，而我們要找以 Q 為圓心並包圍所有洞孔的最小圓。洞孔 $P=(x,y)$ 與 Q 的距離為 $\sqrt{x^2+y^2}$，而

$$\sqrt{x^2+y^2} \;\geq\; \max(|x|,|y|) = S,$$

等號僅在 P 位於坐標軸上時成立。（在圖中，通過 P 的虛線圓永遠不會小於通過 P 的虛線正方形的內切圓。）所以，若 D 是洞孔與 Q 的最大距離，則 $D \geq \max_i S_i$，而包圍所有洞孔的圓的面積滿足

$$\pi D^2 \;\geq\; \pi \max_i S_i^2.$$

由第 (ii) 部分，$\max_i S_i^2$ 的分佈與圓形鏢靶上 $\max_i R_i^2$ 的分佈完全相同，所以

$$E(\pi D^2) \;\geq\; \pi\, E\big(\max_i S_i^2\big) = \frac{n\pi}{n+1},$$

這正是第 (i) 部分圓形鏢靶的答案。這個不等式（inequality）是嚴格的：沒有任何洞孔位於坐標軸上的概率為正，而這時 $D > \max_i S_i$。所以正方形鏢靶上的圓，其期望面積**較大**。

直觀地看：正方形鏢靶包含了圓形鏢靶（圓盤就是它的內切圓），而飛鏢會散落到四個角落，所以最遠的洞孔往往離中心更遠。

驗算

- 當 $n=1$ 時，公式給出 $\frac{\pi}{2}$ 和 2：均勻分佈的 R^2 的平均值為 $\frac{1}{2}$，所以結果正確。當 $n\to\infty$ 時，面積趨向 π 和 4，即整個鏢靶，這正合預期。
- 以蒙地卡羅模擬（Monte Carlo simulation）對 $n=2,3,5$ 各進行 2×10^5 次試驗，結果與 $\frac{n\pi}{n+1}$、$\frac{(n-1)\pi}{n+1}$ 和 $\frac{4n}{n+1}$ 吻合至約三位有效數字；例如 $n=3$ 時得出 2.3561，而 $\frac{3\pi}{4}\approx 2.3562$。
- 對第 (iii) 部分，以數值積分（quadrature）計算精確分佈，在 $n=3$ 時得出正方形鏢靶上的期望面積約為 3.257，而圓盤上則為 2.356：一如所言，前者較大。

常見錯誤

- 先求 $E(M)$ 然後平方：期望面積是 $\pi E(M^2)$，而不是 $\pi\big(E(M)\big)^2$。
- 以為與中心的距離在 $[0,1]$ 上服從均勻分佈。命中的概率與面積成正比，因此服從均勻分佈的是 R^2 而不是 R；外圍的環面積較大。
- 處理第二大的值時，忘記了因子 n，即哪一個洞孔落在圓外的選法數目。

STEP 之外

樣本的最大值和第二大的值都是順序統計量（order statistics）。對 $[0,1]$ 上 n 個獨立的均勻分佈變量，第 k 小的值服從貝塔分佈（Beta distribution），平均值為 $\frac{k}{n+1}$。第 (iii) 部分是隨機優勢（stochastic dominance）（耦合，coupling）的一個例子：若一個隨機變量總是不小於另一個，則它的期望值亦然，而這個比較無須任何細緻的計算。Grimmett 與 Welsh 的 *Probability: An Introduction*（概率入門）以及 Blitzstein 與 Hwang 的 *Introduction to Probability*（概率導論）都有討論順序統計量和隨機變量的變換。

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

2007 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。