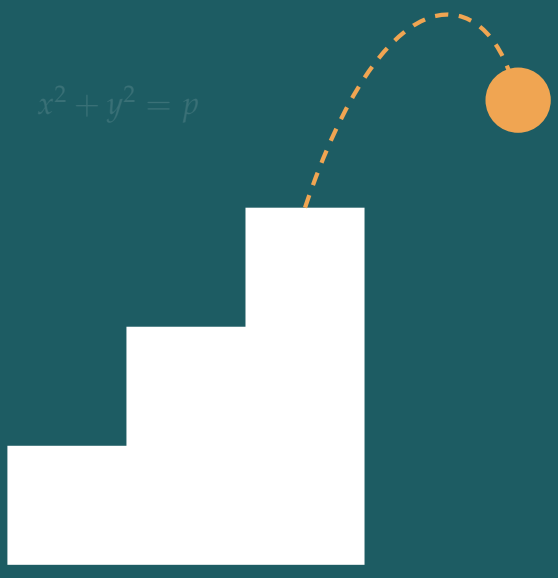
STEP 備試 · 大學數學

Beyond STEP

STEP I, II 及 III

2005 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

2005 年詳細解答

Keith McNulty

Beyond STEP：以歷屆 STEP 試題為門，走進大學數學

2005 年詳細解答

2005 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 數字和的計數 純數學 . . . 3

第 2 題 拋物線的切線與法線 純數學 . . . 6

第 3 題 分式方程 純數學 . . . 10

第 4 題 倍角公式 純數學 . . . 13

第 5 題 冪的積分 純數學 . . . 17

第 6 題 阿波羅尼奧斯圓 純數學 . . . 20

第 7 題 裂項相消乘積 純數學 . . . 24

第 8 題 一階微分方程 純數學 . . . 27

第 9 題 剛體平衡 力學 . . . 31

第 10 題 碰撞與恢復係數 力學 . . . 34

第 11 題 質量運動學 力學 . . . 38

第 12 題 韋氏圖與界限 概率與統計 . . . 42

第 13 題 對稱分佈 概率與統計 . . . 45

第 14 題 混合型隨機變量 概率與統計 . . . 48

STEP II

第 1 題 多項式與駐點 純數學 . . . 52

第 2 題 歐拉 φ 函數 純數學 . . . 55

第 3 題 分部積分法 純數學 . . . 59

第 4 題 反正切恆等式 純數學 . . . 63

第 5 題 內切圓與外接圓 純數學 . . . 67

第 6 題 二項式級數的和 純數學 . . . 71

第 7 題 向量與圓周運動 純數學 . . . 75

第 8 題 可分離變量微分方程 純數學 . . . 78

第 9 題 斜面上的摩擦力 力學 83
第 10 題 相撞的拋體 力學 87
第 11 題 鏈面上的摩擦力 力學 90
第 12 題 貝葉斯定理 概率與統計 94
第 13 題 泊松分佈與二項分佈 概率與統計 98
第 14 題 混合分佈 概率與統計 101

STEP III

第 1 題 三角方程與圖像 純數學 105
第 2 題 可分離變量微分方程 純數學 110
第 3 題 二次函數的複合 純數學 114
第 4 題 非線性遞推關係 純數學 118
第 5 題 拋物線的公切線 純數學 121
第 6 題 解三次方程 純數學 125
第 7 題 代換積分法 純數學 129
第 8 題 複數與圓 純數學 133
第 9 題 碰撞與恢復係數 力學 137
第 10 題 彈性帶與摩擦力 力學 141
第 11 題 剛體平衡 力學 145
第 12 題 無偏估計量的組合 概率與統計 149
第 13 題 幾何分佈 概率與統計 152
第 14 題 連續隨機變量 概率與統計 155

詳細解答

STEP I 2005，第 1 題

純數學 • 數字和的計數

> **STEP I 2005，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/05-s1-q1

關鍵思路

五位數的數字和（digit sum）最大可能是 $9 \times 5 = 45$，而題目要求的兩個數字和都很接近這個值。因此，與其直接數各個數字本身，不如數每個數字比 9 少了多少。這些差額（shortfall）都是很小的非負數，總和又是一個固定的小數目，數起來十分容易：既可以逐一列出各種情況，也可以用「星與棒」法（stars and bars）。

解答

把這個數寫成 $d_1d_2d_3d_4d_5$，其中 $d_1 \in \{1,\ldots,9\}$（首位不能是零），而 $d_2,\ldots,d_5 \in \{0,\ldots,9\}$。設

$$e_i = 9 - d_i \qquad (i = 1,\ldots,5),$$

即第 i 個數字的*差額*。每個 e_i 都是整數，且 $0 \le e_i \le 9$（並且 $e_1 \le 8$，因為 $d_1 \ge 1$），而

$$d_1 + d_2 + d_3 + d_4 + d_5 = 45 - (e_1 + e_2 + e_3 + e_4 + e_5).$$

不同的數給出不同的序列 $(e_1,\ldots,e_5)$，反之亦然，所以數算數字和為 S 的數，等同於數算總和為 $45 - S$ 的有序差額序列。

第 (i) 部分

數字和為 43，即差額的總和是 $45 - 43 = 2$。把 2 分拆成非負部分只有兩種方法：

- **一個數字少了 2，其餘少了 0**。一個數字是 7，其餘四個是 9。這個 7 可以放在 5 個位置中的任何一個：共 5 個數（$79999, 97999, \ldots, 99997$）。
- **兩個數字各少了 1，其餘少了 0**。兩個數字是 8，其餘三個是 9。我們從 5 個位置中選

出 2 個來放 8：共 $\binom{5}{2} = 10$ 個數（88999, 89899, . . . , 99988）。

每個數字都至少是 7，所以首位數字絕不會是 0，以上各數確實都是五位數。這兩種情況互不重疊，並且涵蓋了所有可能，所以共有

$$5 + 10 = 15$$

個五位數的數字和為 43，正如題目所求。

第 (ii) 部分

現在差額的總和必須是 $45 - 39 = 6$。由於 $6 \leq 8$，沒有任何一個差額會超出其上限：$e_1 \leq 6 \leq 8$ 保證 $d_1 \geq 3$，而 $e_i \leq 6 \leq 9$ 保證其餘數字都在 3 與 9 之間。所以我們只需求出下式的有序解的數目：

$$e_1 + e_2 + e_3 + e_4 + e_5 = 6, \qquad e_i \geq 0 \text{ 且為整數.}$$

? **想一想** 怎樣才可以不用逐一列出而數出這些解？

想像把 6 顆星排成一行，用 4 根棒把它們分成 5 組：

$$\star\star\,|\,|\,\star\,|\,\star\star\star\,| \quad \text{代表} \quad (e_1, e_2, e_3, e_4, e_5) = (2, 0, 1, 3, 0).$$

6 顆星和 4 根棒的每一種排列都恰好給出一個解，而每一個解也恰好來自一種排列。只要選定 10 個位置中哪 4 個放棒，排列便隨之確定，所以解的數目是

$$\binom{10}{4} = 210.$$

因此，共有 **210** 個五位數的數字和為 39。

如果你寧願像第 (i) 部分那樣分情況來數，可以列出把 6 分拆成最多五個正數部分的所有方法，再數出每種方法的排列數目：

差額	數的個數	差額	數的個數
6	5	3, 2, 1	$5 \cdot 4 \cdot 3 = 60$
5, 1	$5 \cdot 4 = 20$	3, 1, 1, 1	$5 \cdot 4 = 20$
4, 2	$5 \cdot 4 = 20$	2, 2, 2	$\binom{5}{2} = 10$
4, 1, 1	$5 \cdot \binom{4}{2} = 30$	2, 2, 1, 1	$\binom{5}{2}\binom{3}{2} = 30$
3, 3	$\binom{5}{2} = 10$	2, 1, 1, 1, 1	5

總數是 $5 + 20 + 20 + 30 + 10 + 60 + 20 + 10 + 30 + 5 = 210$，與星與棒法的結果一致。（6 的分拆（partition）1, 1, 1, 1, 1, 1 需要六個數字，所以不會出現。）

驗算

- 星與棒法可重現第 (i) 部分的結果：差額總和為 2 時，得 $\binom{6}{4} = 15$。
- 極端情況同樣成立：差額總和為 0 時，得 $\binom{4}{4} = 1$ 個數（99999）；差額總和為 1 時，得 $\binom{5}{4} = 5$。
- 用電腦搜尋全部 90 000 個五位數，結果恰好有 15 個數的數字和為 43，210 個數的數字和為 39，而各種情況的分佈亦與上表相符。

常見錯誤

- 把 $\{8,8,9,9,9\}$ 當作一個數，或當作 5 個數：數字的次序是有分別的，共有 $\binom{5}{2} = 10$ 種排列。
- 忘記檢查首位數字。本題的首位不可能是 0，因為每個數字都至少是 3；但如果數字和較小（例如 20），星與棒法便會多算，這時必須剔除 $d_1 = 0$ 或有數字大於 9 的情況。

STEP 之外

把 n 寫成 k 個非負整數的有序和，方法的數目是 $\binom{n+k-1}{k-1}$，這也是 $(1 + x + x^2 + \cdots)^k$ 中 x^n 的係數（coefficient）。若限制每個數字最大為 9，上式便要換成 $(1 + x + \cdots + x^9)^k$，而截斷所帶來的影響可用容斥原理（inclusion–exclusion principle）處理。這些生成函數（generating function）方法是大學一年級組合數學（combinatorics）的基本內容；Siklos 的 *Advanced Problems in Mathematics* 和 Davenport 的 *The Higher Arithmetic* 都是繼續探索數字與計數問題的好去處。

詳細解答

STEP I 2005，第 2 題

純數學 • 拋物線的切線與法線

> **STEP I 2005，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/05-s1-q2

關鍵思路

把拋物線 (parabola) $y^2 = 4x$ 用 $(t^2, 2t)$ 參數化 (parametrise) 之後，每條切線 (tangent) 和法線（normal）都是一條關於 x 和 y 的簡短線性方程，所以求交點不過是解兩條聯立線性方程。$(1, 0)$ 是焦點（focus），它位於 PQ 上這個條件，會化為 p 與 q 之間一條簡潔的關係式，而正正是這條關係式令 P 和 Q 處的切線互相垂直。得到這一點之後，那個長方形幾乎是自然而然地出現。

解答

切線與點 R

在點 $(t^2, 2t)$，我們有 $\dfrac{\mathrm{d}x}{\mathrm{d}t} = 2t$ 和 $\dfrac{\mathrm{d}y}{\mathrm{d}t} = 2$，所以 C 在該點的斜率（gradient）是 $\dfrac{2}{2t} = \dfrac{1}{t}$（這沒有問題，因為 $t \neq 0$）。因此 P 處的切線是

$$y - 2p = \frac{1}{p}\,(x - p^2) \quad \Longleftrightarrow \quad py = x + p^2,$$

同理，Q 處的切線是 $qy = x + q^2$。兩式相減，

$$(p - q)\,y = p^2 - q^2 = (p - q)(p + q),$$

由於 $p \neq q$，我們可以除以 $p - q$，得到 $y = p + q$。於是 $x = py - p^2 = p(p + q) - p^2 = pq$。所以

$$R = (pq,\ p + q),$$

正如題目所求。

法線與點 S

$(t^2, 2t)$ 處的法線與切線垂直，所以它的斜率是 $-t$。P 處的法線是

$$y - 2p = -p\,(x - p^2) \quad \Longleftrightarrow \quad px + y = 2p + p^3,$$

而 Q 處的法線是 $qx + y = 2q + q^3$。兩式相減，

$$(p - q)\,x = 2(p - q) + (p^3 - q^3) = (p - q)\big(2 + p^2 + pq + q^2\big),$$

所以，再次除以 $p - q \neq 0$，

$$x = p^2 + pq + q^2 + 2, \qquad y = 2p + p^3 - px = -p^2q - pq^2 = -pq\,(p + q).$$

以上結果對任何 p 和 q 都成立；現在我們運用那個額外條件。

運用 $(1, 0)$ 位於 PQ 上這個條件

弦（chord）PQ 的斜率是

$$\frac{2p - 2q}{p^2 - q^2} = \frac{2}{p + q}$$

（當 $p + q \neq 0$ 時），所以它的方程是 $y - 2p = \dfrac{2}{p + q}(x - p^2)$，化簡後得

$$(p + q)\,y = 2x + 2pq.$$

這個形式同時涵蓋 $p + q = 0$ 的情況，這時它給出通過 P 和 Q 的鉛垂線 $x = -pq = p^2$。代入 $(x, y) = (1, 0)$ 得 $0 = 2 + 2pq$，所以

$$pq = -1.$$

代入 S 的坐標：

$$S = \big(p^2 + q^2 + 2 + pq,\ -pq(p + q)\big) = \big(p^2 + q^2 + 1,\ p + q\big),$$

正如題目所求。同時留意，現在 $R = (-1,\ p + q)$，它位於直線 $x = -1$ 上，亦即準線（directrix）上。

$PSQR$ 是長方形

> ? **想一想** 要證明四個給定的點構成一個長方形，最乾淨利落的方法是甚麼？

一條不錯的路線是：先證明它是平行四邊形（parallelogram）（對角線互相平分），再證明其中一隻角是直角。

平行四邊形。在 $PSQR$ 中，對角線是 PQ 和 SR。利用 $pq = -1$，

$$R + S = \left(pq + p^2 + q^2 + 1,\ 2(p+q)\right) = \left(p^2 + q^2,\ 2p + 2q\right) = P + Q,$$

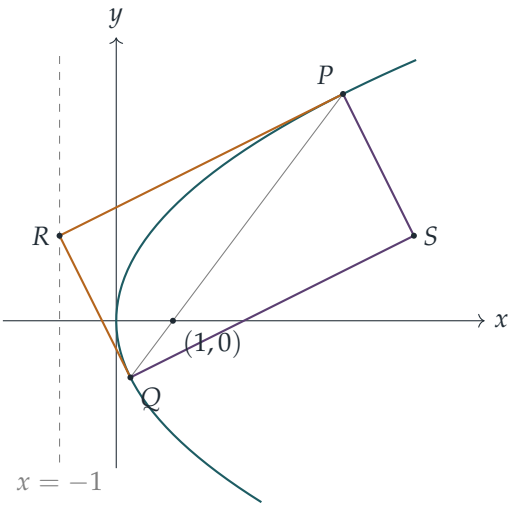
所以 RS 的中點等於 PQ 的中點。對角線互相平分，所以 $PSQR$ 是平行四邊形。它是一個真正的平行四邊形，而不是退化（degenerate）的：P、R 和 Q 並不共線（collinear），因為 R 同時位於兩條切線上，而拋物線上兩個不同點的切線只相交一次，並且任何一條都不會通過另一個切點。

一隻直角。邊 RP 和 RQ 分別沿着 P 和 Q 處的切線，斜率為 $\frac{1}{p}$ 和 $\frac{1}{q}$。它們的乘積是

$$\frac{1}{p} \cdot \frac{1}{q} = \frac{1}{pq} = -1,$$

所以兩條切線互相垂直，角 PRQ 是直角。

有一隻角是直角的平行四邊形就是長方形，所以 $PSQR$ 是長方形。（等價的說法是：P 和 Q 處的角都是直角，因為每條法線都與自己的切線垂直；再加上 R 處的直角，第四隻角也必定是直角。）



圖中所示為 $p = 2$、$q = -\frac{1}{2}$ 的情況：切線為琥珀色，法線為梅子色，焦點弦 PQ 為灰色。

驗算

取 $p = 2$、$q = -\frac{1}{2}$，於是 $pq = -1$。這時 $P = (4,4)$ 而 $Q = (\frac{1}{4}, -1)$，通過這兩點的直線 $y - 4 = \frac{4}{3}(x-4)$ 確實通過 $(1,0)$。由公式得 $R = (-1, \frac{3}{2})$ 和 $S = (\frac{21}{4}, \frac{3}{2})$。直接驗證：$\overrightarrow{RP} = (5, \frac{5}{2})$ 與 $\overrightarrow{QS} = (5, \frac{5}{2})$ 相等，而 $\overrightarrow{RQ} = (\frac{5}{4}, -\frac{5}{2})$ 與 $\overrightarrow{RP}$ 的點積（dot product）是

$\frac{25}{4} - \frac{25}{4} = 0$。所以 $PSQR$ 的確是長方形。我們亦以精確分數，就約二千組隨機選取的 p 和 q 核對了所有坐標公式。

常見錯誤

- 除以 $p - q$ 而沒有說明為何可以這樣做；題目給出 $p \neq q$，正是為了讓你可以這樣做。
- 只證明了某一個角上的兩條鄰邊互相垂直，單憑這一點並不足以構成長方形：你需要的是平行四邊形加上一隻直角，或者三個角都是直角。
- 忘記在 S 中使用 $pq = -1$：一般情況下法線的交點是 $\left(p^2 + pq + q^2 + 2, -pq(p + q)\right)$，只有在焦點弦上它才會化簡成題目給出的形式。

STEP 之外

通過焦點的弦稱為焦點弦（focal chord），而本題證明了關於它的兩個經典事實：弦兩端的切線互相垂直，並且相交於準線上。兩者都是圓錐曲線（conics）的極點與極線（poles and polars）理論的特殊情況，在這套理論中，準線就是焦點的極線；射影幾何（projective geometry）令這類結果幾乎不證自明。若想循這個方向探索處理圓錐曲線的參數方法，Siklos 的 *Advanced Problems in Mathematics*（數學進階難題集）是一個好去處。

詳細解答

STEP I 2005，第 3 題

純數學 • 分式方程

STEP I 2005，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/05-s1-q3

關鍵思路

把等式兩邊同乘 $(x-a)(x-b)$，將每條方程化為多項式方程（polynomial equation），然後數算它的實根（real root）數目。需要小心的地方是：去分母原則上有可能多出原方程不容許的「解」$x=a$ 或 $x=b$，所以我們必須驗證這種情況絕不會發生。在第 (ii) 部分，二次方程有重根（repeated root）當且僅當其判別式（discriminant）為零，由此便得出所求的條件。

解答

第 (i) 部分

當 $x \neq a, b$ 時，我們可以乘以 $(x-a)(x-b)$，而方程等價於

$$\begin{aligned} x(x-b)+x(x-a) &= (x-a)(x-b) \\ 2x^2-(a+b)x &= x^2-(a+b)x+ab \\ x^2 &= ab. \end{aligned}$$

由於 a 與 b 同號，$ab>0$，所以這方程有兩個相異實根 $x=\sqrt{ab}$ 和 $x=-\sqrt{ab}$。

我們仍須確認這兩個根是容許的值。假如 $x=a$ 是根，則 $a^2=ab$，而由於 $a\neq 0$，這給出 $a=b$，與事實不符。同理，$x=b$ 也不是根。因此兩個根都是真正的解，方程恰有兩個相異實數解 $x=\pm\sqrt{ab}$。

第 (ii) 部分

同樣，當 $x \neq a, b$ 時乘以 $(x-a)(x-b)$：

$$2x^2 - (a+b)x = (1+c)\big(x^2 - (a+b)x + ab\big),$$

整理後得

$$q(x) \equiv (1-c)x^2 + c(a+b)x - (1+c)ab = 0.$$

因為 $c \neq 1$，首項係數（leading coefficient）不為零，所以這確實是一條二次方程。

> **想一想** q 的其中一個根會不會是 a 或 b，以致必須捨去？

代入看看：

$$\begin{aligned} q(a) &= (1-c)a^2 + c(a+b)a - (1+c)ab = a^2 - ab = a(a-b), \\ q(b) &= b^2 - ab = b(b-a) \end{aligned}$$

（含 c 的項互相抵消）。由於 a、b 皆不為零且互不相等，兩者都不為零。所以 a 和 b 永遠不是 q 的根。這表示 q 的每個實根都是原方程的解，反之亦然。

q 的判別式為

$$\begin{aligned} \Delta &= c^2(a+b)^2 + 4(1-c)(1+c)ab \\ &= c^2\big[(a+b)^2 - 4ab\big] + 4ab \\ &= c^2(a-b)^2 + 4ab. \end{aligned}$$

若 $c^2 = -\dfrac{4ab}{(a-b)^2}$（而 $a \neq b$，所以可以除以 $(a-b)^2$），則 $\Delta = 0$。這時二次方程恰有一個（重）實根

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

而根據上面的驗證，這個根不是 a 或 b。所以方程恰有一個實數解，正如所求。（事實上逆命題（converse）也成立。若 $\Delta > 0$ 則有兩個解，若 $\Delta < 0$ 則無解，所以這個條件也是必要的。）

改寫條件。 通分後，

$$1 - \left(\frac{a+b}{a-b}\right)^2 = \frac{(a-b)^2 - (a+b)^2}{(a-b)^2} = \frac{-4ab}{(a-b)^2},$$

所以條件就是 $c^2 = 1 - \left(\dfrac{a+b}{a-b}\right)^2$，正如所求。

推導 $0 < c^2 \leq 1$。 平方數 $\left(\frac{a+b}{a-b}\right)^2$ 為非負，所以由第二種形式可得 $c^2 \leq 1$。第一種形式給出 $c^2 = -4ab/(a-b)^2$。由於 $ab \neq 0$，這個值不為零；而對於實數 c 有 $c^2 \geq 0$，所以 $c^2 > 0$。（這也告訴我們 $ab < 0$：a 與 b 異號，這與第 (i) 部分吻合。）因此，條件只有在 $0 < c^2 \leq 1$ 時才可能成立。

上限值是可以取到的。$c^2 = 1$ 當且僅當 $a + b = 0$，這時 $c = -1$（因為 $c \neq 1$）。

驗算

- 取 $a = 4$，$b = -1$。則 $c^2 = 16/25$；選 $c = \frac{4}{5}$。重根為 $x = -\frac{4}{5} \cdot 3/(2 \cdot \frac{1}{5}) = -6$，而確實有 $\frac{-6}{-10} + \frac{-6}{-5} = 0.6 + 1.2 = 1.8 = 1 + c$。
- 極端情況 $a = 1$，$b = -1$，$c = -1$ 給出 $\frac{x}{x-1} + \frac{x}{x+1} = 0$，即 $2x^2 = 0$：只有一個解 $x = 0$。
- 在第 (i) 部分，$a = 1$，$b = 4$ 給出 $x = \pm 2$：$\frac{2}{1} + \frac{2}{-2} = 1$ 及 $\frac{-2}{-3} + \frac{-2}{-6} = 1$。

常見錯誤

- 去分母之後忘記檢查所得的根不是 a 或 b。在本題中這一步結果並無影響，但仍必須加以證明。
- 展開判別式時沒有把含 c^2 的項歸併在一起。寫出 $(a+b)^2 - 4ab = (a-b)^2$，條件才會乾淨利落地得出。
- 聲稱 $c^2 \geq 0$ 而不是 $c^2 > 0$。嚴格不等式（strict inequality）來自 $ab \neq 0$。

STEP 之外

等式左邊是一個有理函數（rational function），而「它的圖像與水平線 $y = 1 + c$ 相交多少次？」其實是一個關於其值域（range）的問題。這裏找到的條件，挑出了該函數的轉向點（turning point）的值，也就是水平線剛好與圖像相切的高度。追問多項式何時有重根，會引出判別式，以及更一般的結式（resultant），這些都會在大學一年級的代數課程中出現。想做更多這類風格的題目，可參考 Siklos 的 *Advanced Problems in Mathematics*（《高等數學難題》）。

詳細解答

STEP I 2005，第 4 題

純數學 • 倍角公式

> **STEP I 2005，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/05-s1-q4

關鍵思路

兩部分都與倍角（multiple angles）有關。在 (i) 中，唯一真正要判斷的是 $\sin\theta$ 的正負號，而題目給定的範圍已把它確定下來；之後只需運用二倍角公式（double-angle formula）和三倍角公式（triple-angle formula）便可。在 (ii) 中，我們把 $\tan 3\theta$ 寫成 $\tan(2\theta+\theta)$ 來推導，然後「由此」（hence）一步便化為關於 $t=\tan\theta$ 的三次方程（cubic）。STEP 題目中的三次方程幾乎總有一個可以看出來的根（root），而 θ 的範圍則告訴我們應保留哪一個根。

解答

第 (i) 部分

由於 $\cos^2\theta+\sin^2\theta=1$，

$$\sin^2\theta = 1-\tfrac{9}{25} = \tfrac{16}{25}, \qquad \text{所以} \qquad \sin\theta = \pm\tfrac{4}{5}.$$

當 $\frac{3\pi}{2} \le \theta \le 2\pi$ 時，θ 位於第四象限（fourth quadrant），該處 $\sin\theta \le 0$。所以 $\sin\theta = -\frac{4}{5}$，而

$$\sin 2\theta = 2\sin\theta\cos\theta = 2\left(-\tfrac{4}{5}\right)\left(\tfrac{3}{5}\right) = -\frac{24}{25},$$

正是所求。

要求 $\cos 3\theta$，我們需要三倍角公式。記 $c = \cos\theta$ 及 $s = \sin\theta$，

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

代入 $c = \frac{3}{5}$，

$$\cos 3\theta = 4\cdot\frac{27}{125} - \frac{9}{5} = \frac{108-225}{125} = -\frac{117}{125}.$$

（這裏 $\sin\theta$ 的正負號並無影響，因為 $\cos 3\theta$ 只取決於 $\cos\theta$。）

第 (ii) 部分

恆等式（identity）。記 $t = \tan\theta$。由二倍角公式，$\tan 2\theta = \dfrac{2t}{1-t^2}$，再由和角公式（addition formula），

$$\tan 3\theta = \tan(2\theta+\theta) = \frac{\tan 2\theta + \tan\theta}{1 - \tan 2\theta\tan\theta} = \frac{\dfrac{2t}{1-t^2} + t}{1 - \dfrac{2t^2}{1-t^2}}.$$

把分子和分母同乘以 $1-t^2$，

$$\tan 3\theta = \frac{2t + t(1-t^2)}{(1-t^2) - 2t^2} = \frac{3t-t^3}{1-3t^2},$$

這就是所求的恆等式。以上推導需要 $\tan 2\theta$ 存在，即 $t^2 \neq 1$；當 $t = \pm 1$ 時，我們可以直接驗證：例如取 $\theta = \frac{\pi}{4}$，右方為 $\frac{3-1}{1-3} = -1 = \tan\frac{3\pi}{4}$，而 $t=-1$ 的情況相同，只是正負號互換。

解 $\tan 3\theta = \frac{11}{2}$。由恆等式，

$$\frac{3t-t^3}{1-3t^2} = \frac{11}{2} \quad\Longrightarrow\quad 6t - 2t^3 = 11 - 33t^2 \quad\Longrightarrow\quad 2t^3 - 33t^2 - 6t + 11 = 0.$$

這裏交叉相乘並無問題，因為 $1-3t^2 \neq 0$：假如它等於零，$\tan 3\theta$ 便沒有定義，而不會是 $\frac{11}{2}$。

> **想一想** 怎樣找出 $2t^3 - 33t^2 - 6t + 11$ 的一個根？

任何寫成最簡分數 p/q 的有理根（rational root），其 p 必整除 11，q 必整除 2，所以我們只需試 $\pm 1, \pm 11, \pm\frac{1}{2}, \pm\frac{11}{2}$。試 $t = \frac{1}{2}$：

$$2\cdot\tfrac{1}{8} - 33\cdot\tfrac{1}{4} - 3 + 11 = \tfrac{1}{4} - \tfrac{33}{4} + 8 = 0.$$

所以 $2t-1$ 是一個因式（factor），比較係數可得

$$2t^3-33t^2-6t+11=(2t-1)(t^2-16t-11).$$

該二次式的根為 $t=8\pm\sqrt{64+11}=8\pm5\sqrt{3}$。所以三個候選值是

$$t=\tfrac{1}{2},\qquad t=8+5\sqrt{3}\approx16.66,\qquad t=8-5\sqrt{3}\approx-0.66.$$

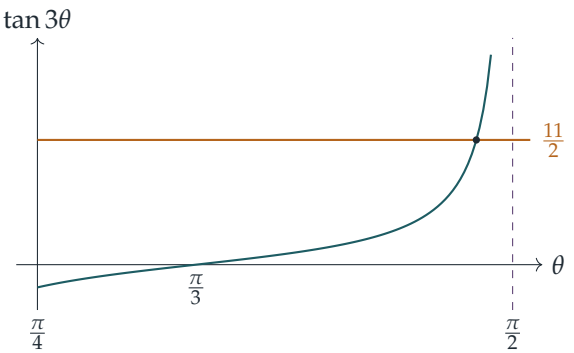
它們對應 $\left(-\frac{\pi}{2},\frac{\pi}{2}\right)$ 內的三個角，這些角的三倍的正切都是 $\frac{11}{2}$。

現在運用題目給定的範圍。由於 tan 在 $\left[\frac{\pi}{4},\frac{\pi}{2}\right)$ 上遞增，且 $\tan\frac{\pi}{4}=1$，我們需要 $t\geq1$（而 $\theta=\frac{\pi}{2}$ 本身須排除，因為 $\tan\theta$ 在該處沒有定義）。只有一個候選值符合要求：

$$\tan\theta=8+5\sqrt{3}.$$

另外兩個確實是三次方程的根，但所屬的角在範圍之外：$\frac{1}{2}<1$ 而 $8-5\sqrt{3}<0$。

圖像亦確認恰好只有一個解：當 θ 由 $\frac{\pi}{4}$ 變到 $\frac{\pi}{2}$ 時，3θ 由 $\frac{3\pi}{4}$ 變到 $\frac{3\pi}{2}$，而在該區間上 $\tan3\theta$ 由 -1 連續遞增至 $+\infty$，所以它恰好一次穿過 $\frac{11}{2}$。



驗算

- 在 (i) 中，$\theta=2\pi-\arccos\frac{3}{5}\approx5.3559$，用計算機可得 $\sin2\theta=-0.96$ 及 $\cos3\theta=-0.936=-\frac{117}{125}$。
- 在 (ii) 中，$\theta=\arctan(8+5\sqrt{3})\approx1.5108$，介乎 $\frac{\pi}{4}\approx0.785$ 與 $\frac{\pi}{2}\approx1.571$ 之間，而 $\tan3\theta\approx5.5000$。
- 快速檢驗恆等式：當 $\theta=\frac{\pi}{12}$ 時，$t=2-\sqrt{3}$，右方化簡後得 $1=\tan\frac{\pi}{4}$。

常見錯誤

- 在 (i) 中取 $\sin\theta=+\frac{4}{5}$，以致 $\sin2\theta$ 的正負號出錯。

▶ 在 (ii) 中因為 $t = \frac{1}{2}$ 是自己看出來的根，便就此停下。它確是三次方程的根，但 $\arctan\frac{1}{2} < \frac{\pi}{4}$。

STEP 之外

這個三次方程有三個根，是因為 $\tan 3\theta = \frac{11}{2}$ 在 $\left(-\frac{\pi}{2}, \frac{\pi}{2}\right)$ 內有三個 θ 的解，彼此相隔 $\frac{\pi}{3}$。同一機制反過來運作，便是三次方程的經典三角解法；更一般而言，棣美弗定理（de Moivre's theorem）把 $\tan n\theta$ 表示為 $\tan\theta$ 的兩個多項式（polynomial）之比，而這些多項式的根正是等距分佈的角的正切。Needham 的 *Visual Complex Analysis*（《複分析：可視化方法》）說明了為何倍角公式會如此自然地從 $(\cos\theta + \mathrm{i}\sin\theta)^n$ 中產生。

詳細解答

STEP I 2005，第 5 題

純數學 • 冪的積分

> **STEP I 2005，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/05-s1-q5

關鍵思路

$x+1$ 的冪可以用反向鏈式法則（reverse chain rule）來積分：$\int (x+1)^n \,\mathrm{d}x = \frac{(x+1)^{n+1}}{n+1}$，但有一個例外：當 $n=-1$ 時，答案是對數（logarithm）。因此，本題所有的分類討論都源於同一個問題：「指數何時會變成 -1？」在 (ii) 中，多出來的因子 x 可以透過寫成 $x=(x+1)-1$ 來消去，剩下兩個單純的冪，所以 m 有兩個例外值。

解答

第 (i) 部分

若 $k \neq 0$，指數 $k-1$ 不等於 -1，所以

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

若 $k=0$，被積函數（integrand）是 $\frac{1}{x+1}$，而由於在 $[0,1]$ 上 $x+1>0$，

$$\int_0^1 \frac{1}{x+1} \,\mathrm{d}x = \Big[\ln(x+1)\Big]_0^1 = \ln 2.$$

想一想 為甚麼 $k \neq 0$ 時的答案會與 $k=0$ 時的答案有關？

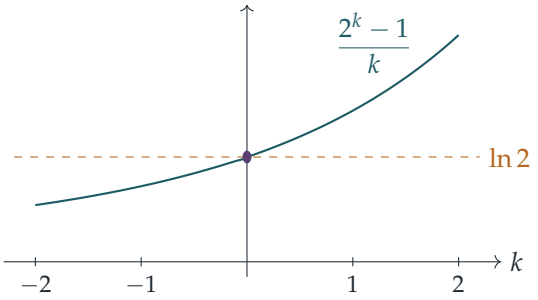
因為當 k 稍微偏離 0 時，被積函數幾乎沒有變化。對於 $0 \leq x \leq 1$，

$$\left|(x+1)^{k-1} - (x+1)^{-1}\right| = \frac{\left|(x+1)^k - 1\right|}{x+1} \leq \left|2^k - 1\right|,$$

這是因為 $1 \le x+1 \le 2$ 意味着 $(x+1)^k$ 介乎 1 與 2^k 之間。在長度為 1 的區間上積分，兩個積分之差至多為 $|2^k-1|$，而當 $k \to 0$ 時，此值趨向 0。因此，對於 $k \approx 0$（且 $k \neq 0$），

$$\frac{2^k-1}{k} = \int_0^1 (x+1)^{k-1}\,\mathrm{d}x \approx \int_0^1 (x+1)^{-1}\,\mathrm{d}x = \ln 2,$$

正如所求。換句話說，$k \neq 0$ 時的公式在 $k=0$ 處有一個「洞」，而 $k=0$ 時的積分恰好把它填上。



第 (ii) 部分

寫成 $x = (x+1)-1$，於是

$$x(x+1)^m = (x+1)^{m+1} - (x+1)^m.$$

每一項都是 $x+1$ 的冪，而對它積分時，恰好在其指數為 -1 時才會出現對數。這發生於 $m+1=-1$ 或 $m=-1$，所以要分的情況是 $m=-1$、$m=-2$，以及其餘所有情況。

一般情況， $m \neq -1, -2$。

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

通分之後，分子為

$$(m+1)\left(2^{m+2}-1\right) - (m+2)\left(2^{m+1}-1\right) = 2^{m+1}\left(2(m+1)-(m+2)\right)+1 = m\,2^{m+1}+1,$$

所以

$$\int_0^1 x(x+1)^m\,\mathrm{d}x = \frac{m\,2^{m+1}+1}{(m+1)(m+2)}, \qquad m \neq -1, -2.$$

情況 $m=-1$。

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

情況 $m=-2$。

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

想一想 一般公式在 $m=-1$ 和 $m=-2$ 處會出現 $0/0$。這是否矛盾？

不是：這其實是第 (i) 部分的重演。在 $m=-1$ 處，分子為 $-1\cdot 2^0+1=0$，而在 $m=-2$ 處，分子為 $-2\cdot 2^{-1}+1=0$，所以公式在這兩處只是沒有定義，而不是錯誤；而且與第 (i) 部分一樣，它的極限（limit）值正是特殊情況的答案（$1-\ln 2\approx 0.307$ 和 $\ln 2-\frac{1}{2}\approx 0.193$）。

驗算

- $m=0$：$\int_0^1 x\,\mathrm{d}x=\frac{1}{2}$，而公式給出 $\frac{0+1}{1\cdot 2}=\frac{1}{2}$。至於 $m=1$：$\int_0^1 (x^2+x)\,\mathrm{d}x=\frac{1}{3}+\frac{1}{2}=\frac{5}{6}$，而公式給出 $\frac{4+1}{6}$。
- 數值上，$k=10^{-4}$ 給出 $\frac{2^k-1}{k}\approx 0.693171$，對比 $\ln 2\approx 0.693147$。
- 對所有積分使用辛普森法則(Simpson's rule)，取多個 k 和 m 的值(包括 $m=-1,-2$)，結果與公式吻合至 10^{-9}。把 $m=-1\pm 10^{-5}$ 代入一般公式，得 $0.30685\ldots$，與 $1-\ln 2$ 相符。

常見錯誤

- 在 $k=0$ 時寫出 $\frac{(x+1)^0}{0}$，或沒有留意到 (ii) 中 m 有兩個例外值，而不是一個。
- 在 (ii) 中使用分部積分法（integration by parts），卻仍然除以 $m+1$，沒有察覺這在 $m=-1$ 時並不成立。
- 在 $m=-2$ 的情況中，弄錯 $\int (x+1)^{-2}\,\mathrm{d}x=-(x+1)^{-1}$ 的正負號。

STEP 之外

(i) 中的推論說明 $\lim_{k\to 0}\frac{2^k-1}{k}=\ln 2$，這正是 2^k 在 $k=0$ 處的導數（derivative）。當中用到的一般事實是：若被積函數一致收斂（converge uniformly），它們的積分也會收斂。這是大學一年級分析課程的一條核心定理，而缺少一致性時，結論便不成立。同樣的想法也解釋了為何當 $n\to -1$ 時，$\int x^n\,\mathrm{d}x$ 會「變成」$\ln x$。Abbott 的 *Understanding Analysis* 對一致收斂與積分有仔細的論述。

詳細解答

STEP I 2005，第 6 題

純數學 • 阿波羅尼奧斯圓

> **STEP I 2005，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/05-s1-q6

關鍵思路

像 $AP = 2BP$ 這樣的條件含有平方根，所以先把它平方：距離的平方是多項式（polynomial），而且兩邊 $x^2 + y^2$ 項的係數（coefficient）相同，這正是軌跡為圓的原因。在第 (ii) 部分，我們對 x-axis 上任意一對點作同樣的處理；而兩條形如 $x^2 + y^2 + px + q = 0$ 的方程描述同一個圓，當且僅當兩者的 p 和 q 分別相同。比較係數（comparing coefficients）可得 k^2 的兩個表達式，再消去 k 便得出兩個結果。

解答

第 (i) 部分

兩個長度都是非負的，所以 $AP = 2BP$ 等價於 $AP^2 = 4BP^2$：

$$(x-5)^2 + (y-16)^2 = 4\left[(x+4)^2 + (y-4)^2\right].$$

展開得

$$x^2 - 10x + y^2 - 32y + 281 = 4x^2 + 32x + 4y^2 - 32y + 128.$$

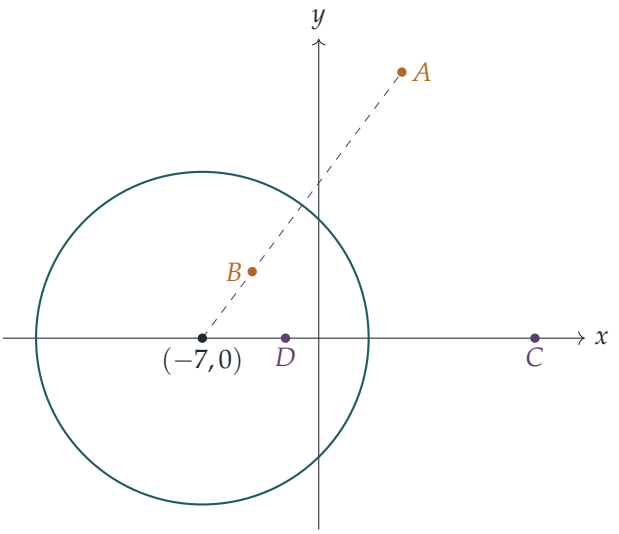
含 y 的一次項互相抵消，把所有項移到右邊，得

$$3x^2 + 42x + 3y^2 - 153 = 0, \qquad \text{即} \qquad x^2 + 14x + y^2 - 51 = 0.$$

配方（completing the square），$x^2 + 14x = (x+7)^2 - 49$，所以

$$(x+7)^2 + y^2 = 100,$$

這正是所要證明的。軌跡是圓心為 $(-7, 0)$、半徑為 10 的圓。



（點 $C = (13, 0)$ 和 $D = (-2, 0)$ 是第 (ii) 部分其中一對可能的點，此時 $k = 2$。留意 A、B 與圓心同樣共線（collinear）。）

第 (ii) 部分

與之前一樣，$QC = kQD$ 等價於 $QC^2 = k^2QD^2$。設 $Q = (x, y)$，則

$$(x-a)^2 + y^2 = k^2\left[(x-b)^2 + y^2\right].$$

把各項移到右邊並整理，

$$(k^2-1)(x^2+y^2) - 2(k^2b - a)x + (k^2b^2 - a^2) = 0.$$

由於 $k > 1$，我們有 $k^2 - 1 \neq 0$，故可以除以它：

$$x^2 + y^2 - \frac{2(k^2b-a)}{k^2-1}\,x + \frac{k^2b^2 - a^2}{k^2-1} = 0.$$

想一想 為甚麼我們可以與 $x^2 + 14x + y^2 - 51 = 0$ 比較係數？

方程 $x^2 + y^2 + px + q = 0$ 可以改寫成 $(x + \frac{1}{2}p)^2 + y^2 = \frac{1}{4}p^2 - q$。如果它是圓心為 $(-7, 0)$、半徑為 10 的圓，那麼圓心迫使 $p = 14$，半徑迫使 $\frac{1}{4}p^2 - q = 100$，所以 $q = -51$。一個圓只有一個圓心和一個半徑，所以兩條方程必須有相同的 p 和相同的 q。

比較 x 的係數：

$$\frac{k^2b - a}{k^2 - 1} = -7 \implies k^2b - a = -7k^2 + 7 \implies k^2(b+7) = a + 7.$$

比較常數項：

$$\frac{k^2b^2-a^2}{k^2-1}=-51 \implies k^2b^2-a^2=-51k^2+51 \implies k^2(b^2+51)=a^2+51.$$

現在 $b^2+51>0$，而且 $b+7\neq 0$：假如 $b=-7$，第一條方程便會給出 $a=-7=b$，而這是題目排除了的情況。因此我們可以分別除以這兩個量，得到 k^2 的兩個表達式：

$$\frac{a+7}{b+7}=k^2=\frac{a^2+51}{b^2+51},$$

這就是第一個結果。

至於第二個結果，交叉相乘：

$$(a+7)(b^2+51)=(b+7)(a^2+51).$$

展開後，常數 357 在兩邊互相抵消，餘下

$$ab^2+51a+7b^2=a^2b+51b+7a^2.$$

把各項分組，使每一組都出現 $a-b$：

$$\begin{aligned}0&=(a^2b-ab^2)-51(a-b)+7(a^2-b^2)\\&=(a-b)\big[ab-51+7(a+b)\big].\end{aligned}$$

由於 $a\neq b$，我們可以除以 $a-b$，所以

$$ab+7a+7b=51,\quad \text{所以}\quad ab+7a+7b+49=100,\quad \text{即}\quad (a+7)(b+7)=100,$$

這正是所要證明的。

驗算

- 試取 $b=-2$：這時 $b+7=5$，所以 $a+7=20$，$a=13$，得出 $k^2=20/5=4$。點 $(3,0)$ 在圓上，而 $QC=10=2\times 5=2\,QD$。點 $(-17,0)$ 也在圓上，此時 $QC=30=2\times 15$。
- 在第 (i) 部分，點 $(3,0)$ 給出 $AP=\sqrt{4+256}=\sqrt{260}$ 和 $BP=\sqrt{49+16}=\sqrt{65}$，而確實有 $\sqrt{260}=2\sqrt{65}$。
- 我們亦以數值方法驗證過：對於許多滿足 $(a+7)(b+7)=100$ 的點對，圓上隨機選取的點都滿足 $QC=kQD$；而且每一對滿足第一個結果且 $a\neq b$ 的數，也都滿足第二個結果。

常見錯誤

- 未除以 $k^2 - 1$ 便比較係數：兩條方程只需相差一個整體的常數倍便已描述同一個圓。
- 除以 $b + 7$ 之前沒有檢查它是否非零。
- 遺漏因子 $a - b$：它所對應的正是題目排除了的平凡解 $a = b$。

STEP 之外

與兩個定點的距離成固定比 $k \neq 1$ 的點的軌跡（locus），稱為阿波羅尼奧斯圓（*circle of Apollonius*）。結果 $(a+7)(b+7) = 100$ 說明 C 和 D 位於圓心的同一側，並且 $MC \cdot MD = r^2$，其中 M 是圓心，r 是半徑：C 和 D 是關於這個圓的反演點（*inverse points*）。（A 和 B 也是如此：$MA = 20$，$MB = 5$。）圓的反演（inversion）是幾何和複分析（complex analysis）中十分有力的工具，在複分析中它就是映射 $z \mapsto 1/\bar{z}$；Needham 的 *Visual Complex Analysis*（《複分析：可視化方法》）對此有精彩的闡述。

詳細解答

STEP I 2005，第 7 題

純數學 • 裂項相消乘積

> **STEP I 2005，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/05-s1-q7

關鍵思路

三個乘積全部都可以裂項相消（telescope）：把每個因子寫成比 $g(r+1)/g(r)$ 的形式，或寫成若干個這類比的乘積，幾乎所有項便會互相約去，只剩下首尾兩個值。第 (i) 及第 (ii) 部分是熱身，用來展示這個機制。第 (iii) 部分的工夫，在於看出那個三角因子其實暗地裏正是這樣的一個比。以 $\sin\frac{(2r-1)\pi}{n}$ 為公分母通分後，分子就變成一個複角（compound angle）的正弦。

解答

第 (i) 部分

把乘積逐項寫出，

$$\prod_{r=1}^{n}\frac{r+1}{r}=\frac{2}{1}\times\frac{3}{2}\times\frac{4}{3}\times\cdots\times\frac{n+1}{n}.$$

每個分子都與下一個因子的分母約去，所以只剩下最後一個分子和第一個分母：

$$\prod_{r=1}^{n}\frac{r+1}{r}=\frac{n+1}{1}=n+1.$$

第 (ii) 部分

把分子因式分解（factorise），$r^2-1=(r-1)(r+1)$，並把每個因子拆成兩個：

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

由於乘積的乘積就是各因子的乘積，我們可以把這兩組因子分開處理（這裏 $n \geq 2$）：

$$\prod_{r=2}^{n} \frac{r-1}{r} = \frac{1}{2} \times \frac{2}{3} \times \cdots \times \frac{n-1}{n} = \frac{1}{n},$$
$$\prod_{r=2}^{n} \frac{r+1}{r} = \frac{3}{2} \times \frac{4}{3} \times \cdots \times \frac{n+1}{n} = \frac{n+1}{2}.$$

兩者都與第 (i) 部分一模一樣地裂項相消。相乘得

$$\prod_{r=2}^{n} \frac{r^2-1}{r^2} = \frac{1}{n} \times \frac{n+1}{2} = \frac{n+1}{2n}.$$

第 (iii) 部分

設 $\alpha = \dfrac{2\pi}{n}$ 及 $\beta_r = \dfrac{(2r-1)\pi}{n}$，則第 r 個因子為 $\cos\alpha + \sin\alpha \cot\beta_r$。

> ? **想一想** $\cos\alpha + \sin\alpha \cot\beta$ 怎樣才能寫成兩個相似的量之比？

利用 $\cot\beta = \cos\beta / \sin\beta$ 並通分。這時分子正是一個複角的展開式：

$$\cos\alpha + \sin\alpha \frac{\cos\beta}{\sin\beta} = \frac{\sin\beta\cos\alpha + \cos\beta\sin\alpha}{\sin\beta} = \frac{\sin(\alpha+\beta)}{\sin\beta}.$$

這一步需要 $\sin\beta_r \neq 0$，而「n 是偶數」這個條件正是在這裏用上。只有當 $(2r-1)/n$ 是整數時，才會有 $\sin\beta_r = 0$。但 $2r-1$ 是奇數而 n 是偶數，奇數不可能是 n 的倍數。因此每個因子都有定義，而下文的每個分母都不為零。

現在 $\alpha + \beta_r = \dfrac{2\pi + (2r-1)\pi}{n} = \dfrac{(2r+1)\pi}{n} = \beta_{r+1}$，所以

$$\cos\alpha + \sin\alpha \cot\beta_r = \frac{\sin\beta_{r+1}}{\sin\beta_r}.$$

這正好是第 (i) 部分的形式，其中 $g(r) = \sin\beta_r$。乘積裂項相消：

$$\prod_{r=1}^{n} \frac{\sin\beta_{r+1}}{\sin\beta_r} = \frac{\sin\beta_2}{\sin\beta_1} \times \frac{\sin\beta_3}{\sin\beta_2} \times \cdots \times \frac{\sin\beta_{n+1}}{\sin\beta_n}$$
$$= \frac{\sin\beta_{n+1}}{\sin\beta_1} = \frac{\sin\left(\frac{(2n+1)\pi}{n}\right)}{\sin\frac{\pi}{n}} = \frac{\sin\left(2\pi + \frac{\pi}{n}\right)}{\sin\frac{\pi}{n}}.$$

由於 sin 的周期（period）為 2π，分子等於分母，故

$$\prod_{r=1}^{n} \left(\cos\tfrac{2\pi}{n} + \sin\tfrac{2\pi}{n} \cot\tfrac{(2r-1)\pi}{n}\right) = 1.$$

驗算

▶ 第 (i) 部分，取 $n = 3$：$2 \cdot \frac{3}{2} \cdot \frac{4}{3} = 4 = n+1$。

- 第 (ii) 部分，取 $n = 2$：唯一的因子是 $\frac{3}{4}$，而 $\frac{n+1}{2n} = \frac{3}{4}$。當 $n \to \infty$ 時，乘積趨向 $\frac{1}{2}$，這是一個著名的無窮乘積（infinite product）。
- 第 (iii) 部分，取 $n = 4$：$\cos\frac{\pi}{2} = 0$ 及 $\sin\frac{\pi}{2} = 1$，所以各因子為 $\cot\frac{\pi}{4}$、$\cot\frac{3\pi}{4}$、$\cot\frac{5\pi}{4}$ 和 $\cot\frac{7\pi}{4}$，即 1、-1、1 和 -1，乘積為 1。我們亦以數值計算驗證過，對 60 以內的每個偶數 n，乘積都等於 1（誤差在 10^{-8} 以內）。
- 若 n 是奇數，$r = \frac{n+1}{2}$ 的那個因子涉及 $\cot\pi$，而它沒有定義。這就解釋了題目為何要有這個條件。

常見錯誤

- 在第 (ii) 部分，由 $r = 1$ 開始相消：下限是 2，所以最先的兩個因子是 $\frac{1}{2}$ 和 $\frac{3}{2}$。
- 在第 (iii) 部分，停在 $\sin\frac{(2n+1)\pi}{n} / \sin\frac{\pi}{n}$ 就算，沒有利用周期性把它化簡為 1。
- 沒有解釋「n 是偶數」用在哪裏。正是這個條件，令每個餘切（cotangent）以及每個分母都有妥善的定義。

STEP 之外

乘積的裂項相消，與你在部分分式（partial fractions）中遇到的裂項求和，是乘法與加法上的一對近親。取對數（logarithm）就能把其中一個變成另一個。第 (ii) 部分是無窮乘積 $\prod_{r\geq 2}(1 - 1/r^2) = \frac{1}{2}$ 的有限版本。像這樣的無窮乘積，以及歐拉（Euler）乘積

$$\sin \pi x = \pi x \prod_{r\geq 1}\left(1 - \frac{x^2}{r^2}\right),$$

要到大學第一門分析（analysis）課程才會嚴謹地研究；在那裏，乘積的收斂（convergence）是透過對數之和的收斂來定義的。Abbott 的 *Understanding Analysis* 是建立所需收斂概念的好起點。

詳細解答

STEP I 2005，第 8 題

純數學 • 一階微分方程

> **STEP I 2005，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/05-s1-q8

關鍵思路

開頭的恆等式(identity)其實是一道配方，教我們辨認形如 $2xy\,\frac{\mathrm{d}y}{\mathrm{d}x} = ky^2 + (\text{只含}x\text{ 的式子})$ 的方程。方程一旦寫成這個形狀，代換（substitution）$y^2 = x^k f(x)$ 便會把 y^2 項消去，剩下 $f'(x)$ 等於一個 x 的函數，直接積分即可。真正要花工夫的，只有選取 k，以及在第 (ii) 部分把方程兩邊同除，使左邊變成 $2xy\,\frac{\mathrm{d}y}{\mathrm{d}x}$。

解答

恆等式

把 $y^2 = x^k f(x)$ 對 x 作隱函數求導（implicit differentiation），右邊用積法則（product rule）：

$$2y\frac{\mathrm{d}y}{\mathrm{d}x} = kx^{k-1}f(x) + x^k\frac{\mathrm{d}f}{\mathrm{d}x}.$$

兩邊乘以 x，

$$2xy\frac{\mathrm{d}y}{\mathrm{d}x} = k\,x^k f(x) + x^{k+1}\frac{\mathrm{d}f}{\mathrm{d}x} = ky^2 + x^{k+1}\frac{\mathrm{d}f}{\mathrm{d}x},$$

因為 $x^k f(x) = y^2$。這就是所求的結果。

第 (i) 部分

取 $k = 1$，寫成 $y^2 = xf(x)$，恆等式便給出 $2xy\,\frac{\mathrm{d}y}{\mathrm{d}x} = y^2 + x^2 f'(x)$。與題目給出的方程比較，$y^2$ 項互相吻合，所以方程成立當且僅當

$$x^2 f'(x) = x^2 - 1, \qquad \text{即} \qquad f'(x) = 1 - \frac{1}{x^2} \quad (x \neq 0).$$

積分得 $f(x) = x + \dfrac{1}{x} + c$，所以

$$y^2 = xf(x) = x^2 + cx + 1.$$

由條件「當 $x = 1$ 時 $y = 2$」得 $4 = 1 + c + 1$，所以 $c = 2$，而

$$y^2 = x^2 + 2x + 1 = (x+1)^2.$$

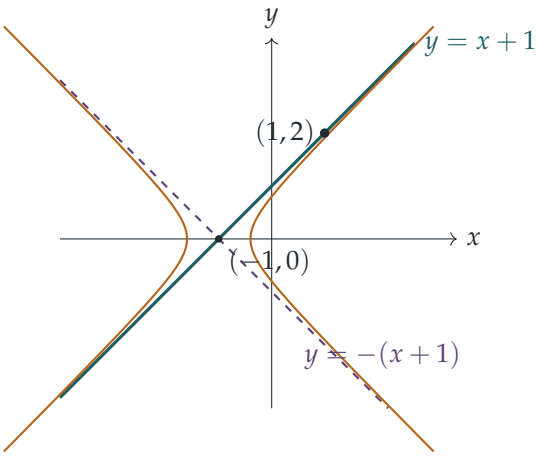
因此 $y = x + 1$ 或 $y = -(x+1)$；只有 $y = x + 1$ 通過 $(1, 2)$，所以解是

$$y = x + 1.$$

（代回原式：$2x(x+1) \cdot 1 = 2x^2 + 2x$，而 $(x+1)^2 + x^2 - 1 = 2x^2 + 2x$，所以它對每一個 x 都滿足方程，包括我們曾經用來作除數的 $x = 0$。）

想一想 答案只不過是 $y = x + 1$，那麼「從幾何角度描述」究竟想我們寫甚麼？

解曲線是一條直線：它的斜率（gradient）為 1，通過 $(-1, 0)$ 和 $(0, 1)$，與 x-axis 所在的橫軸成 45° 角。值得一提的是，這個結果其實有點出人意表。通解（general solution）$y^2 - x^2 - cx = 1$ 是一族圓錐曲線（conics）；配方（completing the square）後得 $y^2 - (x + \frac{c}{2})^2 = 1 - \frac{c^2}{4}$，這是一條等軸雙曲線（rectangular hyperbola），中心在 $(-\frac{c}{2}, 0)$，漸近線（asymptotes）為 $y = \pm(x + \frac{c}{2})$。初始條件（initial condition）選出了 $c = 2$，這正是右邊等於零的情況，雙曲線退化（degenerate）成它自己的一對漸近線 $y = \pm(x+1)$。我們的解就是這對直線中的 $y = x + 1$ 那一支。



實線是所求的解；虛線是 $y^2 = (x+1)^2$ 的另一半。琥珀色的曲線勾畫出這族曲線中非退化成員的形狀：一條等軸雙曲線，以類似這樣的一對直線為漸近線，並逐漸向它們靠近。

第 (ii) 部分

要使用恆等式，左邊必須是 $2xy\,\frac{\mathrm{d}y}{\mathrm{d}x}$，所以把方程除以 x（$\ln x$ 項已限定 $x>0$，所以這一步並無問題）：

$$2xy\frac{\mathrm{d}y}{\mathrm{d}x} = -y^2 + \frac{2\ln x}{x}.$$

y^2 的係數是 -1，所以取 $k=-1$，即 $y^2 = f(x)/x$。這時 $x^{k+1} = x^0 = 1$，恆等式變成 $2xy\,\frac{\mathrm{d}y}{\mathrm{d}x} = -y^2 + f'(x)$。比較兩式，

$$f'(x) = \frac{2\ln x}{x}.$$

由於 $\frac{\mathrm{d}}{\mathrm{d}x}(\ln x)^2 = \frac{2\ln x}{x}$，我們得到 $f(x) = (\ln x)^2 + c$，於是

$$y^2 = \frac{(\ln x)^2 + c}{x}.$$

當 $x=1$ 時 $y = 1 \div 1 = (0+c)/1$，所以 $c=1$。這裏取正平方根，因為在 $x=1$ 處 $y=1>0$，而且 y 永不為零（分子 $1+(\ln x)^2$ 最少是 1），所以解是

$$y = \sqrt{\frac{1+(\ln x)^2}{x}}, \qquad x>0.$$

驗算

- 在 (i) 中，直線 $y = x+1$ 可以如上文般直接代入方程驗證；由 $(1,2)$ 出發對方程作數值積分（龍格—庫塔法，Runge–Kutta），在 $x=2$ 得 $y=3.000$，在 $x=5$ 得 $y=6.000$。
- 在 (ii) 中，$y^2 = (1+(\ln x)^2)/x$ 給出 $y(3)\approx 0.8577$ 和 $y(0.2)\approx 4.2369$；由 $(1,1)$ 出發對微分方程作數值積分，結果在小數點後多個位都吻合。
- 恆等式本身：試取 $k=0$，它便簡化成 $2xy\,y' = xf'$，這只不過是 $(y^2)' = f'$ 乘以 x。

常見錯誤

- 在 (ii) 中出於習慣而用了 $k=1$。除以 x 之後，y^2 項的正負號才是決定 $k=-1$ 的關鍵。
- 在 (ii) 中忘記除以 x，結果嘗試拿 $2x^2y\,y'$ 去配對恆等式。
- 把 (i) 的答案停留在 $y^2 = (x+1)^2$，沒有選出通過 $(1,2)$ 的那一支。

STEP 之外

這裏的代換其實是一個標準技巧的變裝版本：設 $u = y^2$，便把 $2xy\,y' = ky^2 + g(x)$ 化成線性（linear）方程 $xu' - ku = g(x)$，而 x^{-k} 就是它的積分因子（integrating factor）。經冪代換後變成線性的方程稱為伯努利方程（Bernoulli equations），是大學常微分方程（ODE）課程中最早遇到的非線性方程之一。Boyce 與 DiPrima 的 *Elementary Differential Equations*（一本常微分方程入門教科書）把它們與積分因子和恰當方程（exact equations）放在一起討論。

詳細解答

STEP I 2005，第 9 題

力學 • 剛體平衡

> **STEP I 2005，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/05-s1-q9

關鍵思路

整道題的關鍵在於一個未知量：這根非均勻桿的質心（centre of mass）G 位於何處。頭兩個裝置各自給出以該位置表示的 T，把兩者等同起來，便可同時定出 G 和 T。之後，傾斜的裝置就是一道標準的剛體（rigid body）問題：先對未知接觸力所作用的那一點取力矩（moments），然後分解（resolve）各力。

解答

我們把桿看作剛體，其全部重量 W 作用於質心 G，並把繩看作輕而不可伸長（light and inextensible）。設 $AG = x$，則 $GB = 3l - x$。

求 G 的位置並證明 $5T = 2W$

兩端各繫一繩。對 B 取力矩，可消去 B 處未知的張力（tension）：

$$T \cdot 3l = W(3l - x).$$

支點距離 A 爲 l，繩繫於 B。設 P 為支點（pivot），則 $PB = 2l$。對 P 取力矩，可消去未知的支點反作用力。B 處的繩要把桿拉住，G 必須位於 P 靠 B 的一側，故 $PG = x - l$，而

$$T \cdot 2l = W(x - l).$$

消去 T（把第一條方程乘以 2，第二條乘以 3，然後比較）：

$$2W(3l - x) = 3W(x - l) \quad \Longrightarrow \quad 9l = 5x \quad \Longrightarrow \quad x = \tfrac{9}{5}l.$$

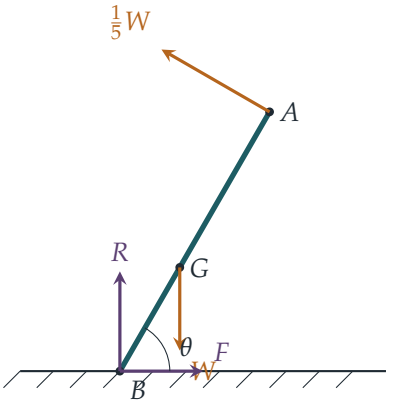
於是由第二條方程得 $2lT = W \cdot \frac{4}{5}l$，故 $T = \frac{2}{5}W$，即

$$5T = 2W,$$

正如所求。留意 $x = \frac{9}{5}l > l$，這與我們假設 G 在支點之外是一致的。因此 G 與 B 的距離為 $GB = 3l - \frac{9}{5}l = \frac{6}{5}l$。

桿擱在地面上

現在 B 在粗糙的地面上，桿與水平成 θ 角，而 A 處的繩垂直於桿，張力為 $\frac{1}{2}T = \frac{1}{5}W$。在 B 處有向上的法向反作用力（normal reaction）R 和沿地面的摩擦力（friction）F。



對 B 取力矩。這樣可消去 R 和 F。繩的拉力垂直於桿，作用點與 B 相距 $3l$，故其力矩為 $\frac{1}{5}W \cdot 3l$。重量沿桿相距 $\frac{6}{5}l$ 之處鉛垂作用，故其作用線與 B 的垂直距離為 $\frac{6}{5}l\cos\theta$。因此

$$\tfrac{1}{5}W \cdot 3l = W \cdot \tfrac{6}{5}l\cos\theta \quad\Longrightarrow\quad \cos\theta = \tfrac{1}{2},$$

所以 $\theta = 60°$（即 $\frac{\pi}{3}$）。

想一想 繩的張力的各個分量指向哪一方？

桿與水平成 θ 角，而繩與桿成直角，所以繩與鉛垂線成 θ 角。因此其張力 $\frac{1}{5}W$ 的鉛垂分量為 $\frac{1}{5}W\cos\theta$（向上），水平分量為 $\frac{1}{5}W\sin\theta$，指向越過 B 的後方，即背離 A 所在的一側。摩擦力必須抵銷這個水平拉力，所以它在 B 處指向 A 所在的一側，如圖所示。

沿鉛垂方向分解：

$$R + \tfrac{1}{5}W\cos 60° = W \quad\Longrightarrow\quad R = W - \tfrac{1}{10}W = \tfrac{9}{10}W.$$

沿水平方向分解：

$$F = \tfrac{1}{5}W\sin 60° = \tfrac{\sqrt{3}}{10}W.$$

只要 $F \le \mu R$，桿便不會滑動，即

$$\mu \ge \frac{F}{R} = \frac{\sqrt{3}/10}{9/10} = \frac{\sqrt{3}}{9}.$$

所以能防止滑動的最小摩擦系數（coefficient of friction）為 $\mu = \dfrac{\sqrt{3}}{9} = \dfrac{1}{3\sqrt{3}} \approx 0.192$（等號成立時，桿處於極限平衡（limiting equilibrium））。

驗算

- 取 $x = \frac{9}{5}l$，雙繩裝置給出 $T = W \cdot \frac{6/5}{3} = \frac{2}{5}W$，而支點裝置給出 $T = W \cdot \frac{4/5}{2} = \frac{2}{5}W$：兩者相同，理應如此。
- 在傾斜的位置，再對 A 取力矩（取 $W = l = 1$，$\theta = 60°$，$R = 0.9$，$F \approx 0.1732$），所得的合力矩同樣為零，這一併驗證了 R 和 F。
- 數值上 $\mu_{\min} = \sqrt{3}/9 \approx 0.19245$，這已由電腦建立各力的向量並檢查全部三條平衡方程加以確認。

常見錯誤

- 以 $\frac{6}{5}l\sin\theta$ 作為重量的力臂。由 B 至重量作用線的水平距離是 $\frac{6}{5}l\cos\theta$。
- 忘記垂直於桿的繩有鉛垂分量，而取 $R = W$。這樣會得出 $\mu = \sqrt{3}/10$，數值偏小。
- 假設桿是均勻的，以為 G 在中點。第一部分的全部目的正是要找出 G 的位置。

STEP 之外

利用平衡時的量度來確定質心的位置，是一個強大想法的靜力學版本：一個力系對某一點的力矩決定了該力系，只差一個合力。到了大學，這會發展成剛體靜力學以至動力學的研究，其中質心和轉動慣量（moment of inertia）就是支配物體運動的兩個數。Kleppner 與 Kolenkow 所著的 *An Introduction to Mechanics* 講述力矩（torque）與剛體，並附有大量這類風格而難度更高的題目。

詳細解答

STEP I 2005，第 10 題

力學 • 碰撞與恢復係數

STEP I 2005，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：

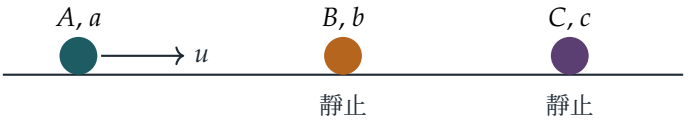
step.maths.org/questions/05-s1-q10

關鍵思路

每一次碰撞都是兩個光滑質點（particle）之間的直接碰撞（direct impact），因此可得兩條方程：沿直線方向的動量守恆（conservation of momentum），以及牛頓實驗定律（Newton's experimental law）（分離速率 $= e\times$ 接近速率）。只要對一般的一對質點解一次這兩條方程，所得公式便可在第二次碰撞中重複使用。至於各個特殊情況，真正的問題是：兩次碰撞之後，事情是否就此結束？題目要求的是最終速度，所以必須檢查沒有任何質點會追上前面的那一個。

解答

我們把 A、B、C 視為在光滑平面上沿同一直線運動的質點，因此沒有水平方向的外力作用，每次碰撞中沿直線方向的動量都守恆。取 u 的方向為正。設 p 為 A 與 B 碰撞後的速度，q 為 B 與 C 碰撞後的速度。



第一次碰撞：求 v

碰撞前 A 的速度為 u，而 B 靜止；碰撞後兩者的速度分別為 p 和 v。由動量守恆和牛頓定律得

$$au = ap + bv, \qquad v - p = eu.$$

把 $p = v - eu$ 代入第一條方程，得 $au = av - aeu + bv$，所以

$$v = \frac{a(1+e)u}{a+b}, \qquad p = v - eu = \frac{(a-eb)u}{a+b}.$$

留意 $v > 0$，所以 B 確實朝着 C 運動。

第二次碰撞：求 w

這與剛才的情況相同，只是換成 B（質量 b，速率 v）撞向 C（質量 c，靜止）。因此可以重用上面的公式，作代換 $a \to b$、$b \to c$、$u \to v$：

$$w = \frac{b(1+e)v}{b+c}, \qquad q = \frac{(b-ec)v}{b+c}.$$

代入 v 的表達式，

$$w = \frac{b(1+e)}{b+c} \cdot \frac{a(1+e)u}{a+b} = \frac{abu(1+e)^2}{(a+b)(b+c)},$$

正如所求。

這兩次碰撞之後，各質點的速度為 p、q、w。只有當 $p > q$（A 追上 B）或 $q > w$（B 追上 C）時，才會再有碰撞；由牛頓定律，$q - w = -ev < 0$，所以後者永不發生，我們只需比較 p 和 q。

> **想一想** 為甚麼不能在兩次碰撞之後就停下來，直接寫出 p、q、w 作為答案？

因為題目問的是最終速度。假如之後 A 仍然比 B 快，它便會再次撞上 B，一切都會隨之改變。所以在每種情況下，我們都必須檢查 $p \le q$。

第 (i) 部分

這裏 $a = eb$ 及 $b = ec$（其中 $e > 0$，因為質量為正數）。於是 $a - eb = 0$ 及 $b - ec = 0$，所以

$$p = 0, \qquad q = 0.$$

至於 w，利用 $a + b = b(1+e)$ 及 $b + c = b + b/e = b(1+e)/e$：

$$w = \frac{eb \cdot b\,u(1+e)^2}{b(1+e) \cdot b(1+e)/e} = e^2 u.$$

A 和 B 都靜止，而 C 向前離去，所以不會再有碰撞。最終速度為

$$A:\ 0, \qquad B:\ 0, \qquad C:\ e^2 u.$$

每次碰撞都把全部動量傳給下一個質點，就像一個牛頓擺（Newton's cradle），只是後面的質點較輕。

第 (ii) 部分

現在 $b = ea$ 及 $c = eb = e^2a$。於是

$$v = \frac{a(1+e)u}{a(1+e)} = u, \qquad p = \frac{(a-e^2a)u}{a(1+e)} = (1-e)u,$$

當中用了 $1-e^2 = (1-e)(1+e)$。在第二次碰撞中，質量 b、c 之比與 a、b 之比相同，而 B 以速率 u 到達，所以它完全是第一次碰撞的翻版：

$$w = u, \qquad q = (1-e)u.$$

（直接計算：$w = \dfrac{a \cdot ea \cdot u(1+e)^2}{a(1+e) \cdot ea(1+e)} = u$。）

現在 $p = q$，所以 A 永遠追不上 B；而當 $e > 0$ 時 $q = (1-e)u < u = w$，所以 B 永遠追不上 C。最終速度為

$$A:\ (1-e)u, \qquad B:\ (1-e)u, \qquad C:\ u,$$

全部沿 A 原來的運動方向。

驗算

- (ii) 的動量：$a(1-e)u + ea(1-e)u + e^2a\,u = au$。(i) 的動量：$c \cdot e^2u = e^2c\,u = au$。
- 在 (ii) 中取 $e = 1$（完全彈性，perfectly elastic），A 和 B 最終靜止，而 C 以 u 離去；動能（kinetic energy）守恆：$\frac{1}{2}au^2 = \frac{1}{2}cu^2$ 需要 $c = a$，這正是 $c = e^2a$ 在 $e = 1$ 時的情況。
- 一個數值例子：取 $e = \frac{1}{2}$，質量為 $1, 2, 4$ 及 $u = 1$（第 (i) 部分），最終速度為 $0, 0, \frac{1}{4}$；質量為 $4, 2, 1$（第 (ii) 部分）則得 $\frac{1}{2}, \frac{1}{2}, 1$。我們以精確算術，對數以千計隨機選取的質量和 e 值，並模擬其後可能出現的任何碰撞，驗證了這些結果以及 v 和 w 的一般公式。

常見錯誤

- 寫牛頓定律時把速度的次序弄反：正確的是碰撞後的（前方質點的速率 − 後方質點的速率）$= e \times$ 碰撞前的（後方 − 前方）。
- 直接寫出兩次碰撞後的速度，而沒有檢查 A 會否再次追上 B。

STEP 之外

連串碰撞是描述動量和能量如何在物料中傳遞的最簡單模型，而第 (i) 部分的質量比經過特意選取，使全部動量最終都落在最後一個質點上。若把質量的次序倒過來，重的在後、輕的在前，並且 e 接近 1，最後一個質點便會比第一個質點運動得更快：這正是那個玩具背後的原理——把一個小球疊在一個大球上，讓兩者一同落下，小球便會高高彈上半空。Kleppner 與 Kolenkow 的 *An Introduction to Mechanics*（力學入門教科書）以質心參考系（centre-of-mass frame）處理碰撞，令 v 和 w 這類公式變得幾乎一目了然。

詳細解答

STEP I 2005，第 11 題

力學 • 向量運動學

> **STEP I 2005，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/05-s1-q11

關鍵思路

這道題由始至終都在考兩個向量（vector）$\mathbf{r}$ 與 $\mathbf{v} = \dot{\mathbf{r}}$ 之間的相對位置。垂直（perpendicular）即純量積（scalar product）$\mathbf{r} \cdot \mathbf{v}$ 等於零；平行（parallel）即「交叉」組合 $xv_y - yv_x$ 等於零。只要寫出 $\sin 2t = 2\sin t \cos t$ 和 $\cos 2t = 1 - 2\sin^2 t$，這兩個式子都可以化簡得十分整齊。至於求距離和繪畫草圖，把所有式子都用 $\cos t$ 表示，或者乾脆消去 t，會方便得多。

解答

設 $\mathbf{r} = x\,\mathbf{i} + y\,\mathbf{j}$，其中

$$x = \sin 2t = 2\sin t \cos t, \qquad y = 2\cos t, \qquad 0 \le t < 2\pi.$$

求導後，得速度（velocity）為

$$\mathbf{v} = \dot{\mathbf{r}} = (2\cos 2t)\,\mathbf{i} - (2\sin t)\,\mathbf{j}.$$

第 (i) 部分

當 $x = 0$ 且 $y = 0$ 時，質點（particle）位於原點。條件 $y = 0$ 即 $\cos t = 0$，所以在給定範圍內 $t = \frac{\pi}{2}$ 或 $t = \frac{3\pi}{2}$。在這兩個時刻，$x = 2\sin t \cos t = 0$ 自動成立。因此質點恰好在 $t = \frac{\pi}{2}$ 和 $t = \frac{3\pi}{2}$ 這兩個時刻位於原點，其他時刻都不在原點，這正是要證明的結果。

第 (ii) 部分

當 $\mathbf{r}\cdot\mathbf{v}=0$ 時，速度與位移（displacement）垂直（兩個向量都須非零，「垂直」才有意義）。現在

$$\begin{aligned}\mathbf{r}\cdot\mathbf{v} &= 2\sin 2t\cos 2t - 4\sin t\cos t\\ &= 2\sin 2t\cos 2t - 2\sin 2t\\ &= 2\sin 2t\,(\cos 2t-1)\\ &= -4\sin 2t\,\sin^2 t,\end{aligned}$$

其中用了 $\cos 2t-1=-2\sin^2 t$。當 $\sin 2t=0$ 或 $\sin t=0$ 時，此式等於零，即

$$t=0,\ \tfrac{\pi}{2},\ \pi,\ \tfrac{3\pi}{2}.$$

想一想 $t=\frac{\pi}{2}$ 和 $t=\frac{3\pi}{2}$ 真的應該列入答案嗎？

不應該。由第 (i) 部分可知，在這兩個時刻 $\mathbf{r}=\mathbf{0}$，而零向量沒有方向，所以說它與任何向量垂直都沒有實質意義；純量積等於零，只是出於一個平凡的原因。在 $t=0$ 時，$\mathbf{r}=2\mathbf{j}$ 而 $\mathbf{v}=2\mathbf{i}$；在 $t=\pi$ 時，$\mathbf{r}=-2\mathbf{j}$ 而 $\mathbf{v}=2\mathbf{i}$：這些才是真正互相垂直的非零向量。因此，速度與位移垂直的時刻是

$$t=0 \quad 和 \quad t=\pi.$$

第 (iii) 部分

首先，$\mathbf{v}$ 永不為零：若 $\mathbf{v}=\mathbf{0}$，則必須 $\sin t=0$，但這時 $\cos 2t=1-2\sin^2 t=1\neq 0$。

兩個非零向量 (x,y) 與 (v_x,v_y) 平行，當且僅當 $xv_y-yv_x=0$。這裏

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

此式只在 $\cos t=0$ 時等於零，而由第 (i) 部分可知，這只在質點位於原點時才會發生。所以只要質點不在原點，就有 $\cos t\neq 0$，這個組合不等於零，速度亦不與位移平行。證明至此完成。

第 (iv) 部分

設 $c=\cos^2 t$，則 $0\le c\le 1$。於是

$$|\mathbf{r}|^2=\sin^2 2t+4\cos^2 t=4\sin^2 t\cos^2 t+4\cos^2 t=4c(1-c)+4c=4c(2-c).$$

二次式 $4c(2-c)$ 的頂點在 $c=1$，並在 $[0,1]$ 上遞增，所以它在這個區間上的最大值在 $c=1$ 處取得，這時 $|\mathbf{r}|^2=4$。因此質點與原點的最大距離是

$$|\mathbf{r}|_{\max}=2,$$

在 $\cos^2 t=1$ 時達到，即 $t=0$ 和 $t=\pi$，對應的點是 $(0,\pm2)$。這正好是第 (ii) 部分求得的時刻，令人安心：距離達到最大值時，速度的徑向分量（radial component）必定為零。

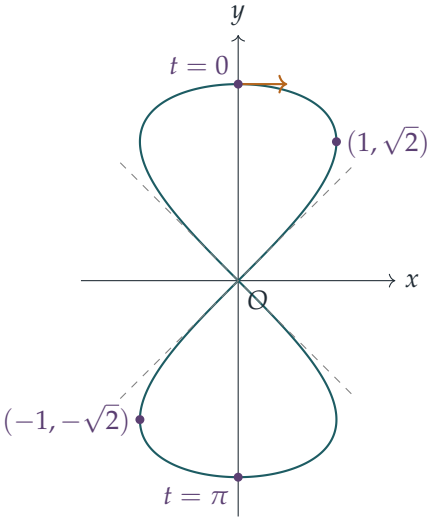
至於草圖，先消去 t。由於 $x=y\sin t$ 及 $\sin^2 t=1-\cos^2 t=1-\frac{1}{4}y^2$，

$$x^2=y^2\Big(1-\frac{y^2}{4}\Big),\qquad \text{即}\qquad 4x^2=y^2(4-y^2).$$

草圖上應顯示以下特徵：

- 曲線對兩條坐標軸都對稱，且 $-2\le y\le 2$。
- $|x|=|\sin 2t|\le 1$，當 $y=\pm\sqrt{2}$ 時 $|x|=1$。
- 在原點附近 $4x^2\approx 4y^2$，所以曲線在該處沿直線 $y=\pm x$ 自相交。這與 $t=\frac{\pi}{2}$ 時的速度 $-2\mathbf{i}-2\mathbf{j}$ 及 $t=\frac{3\pi}{2}$ 時的速度 $-2\mathbf{i}+2\mathbf{j}$ 吻合。
- 在 $(0,\pm2)$ 處路徑是水平的，正如第 (ii) 部分所示。

隨着 t 增加：質點從 $(0,2)$ 出發，沿 $+\mathbf{i}$ 方向運動，順着上方環圈的右邊下行至原點，再繞下方環圈一周（先走左邊）回到原點，最後沿上方環圈的左邊上行。整條路徑是一個 8 字形。



驗算

- 在 $t = \frac{\pi}{4}$ 時：$\mathbf{r} = \mathbf{i} + \sqrt{2}\,\mathbf{j}$，滿足 $4x^2 = 4 = y^2(4 - y^2) = 2 \cdot 2$，而 $|\mathbf{r}| = \sqrt{3} < 2$，與 $4c(2 - c) = 4 \cdot \frac{1}{2} \cdot \frac{3}{2} = 3$ 一致。
- 在 $[0, 2\pi)$ 上對 t 作細密取樣，得最大距離為 2.000，距離接近零的情況只出現在 $t \approx 1.57$ 和 4.71，並且在隨機選取的時刻驗證了恆等式 $\mathbf{r} \cdot \mathbf{v} = -4 \sin 2t \sin^2 t$ 和 $xv_y - yv_x = -4\cos^3 t$。最小速率約為 1.32，所以 $\mathbf{v}$ 永不為零。

常見錯誤

- 在第 (ii) 部分把 $t = \frac{\pi}{2}$ 和 $\frac{3\pi}{2}$ 列入答案：純量積在這兩個時刻等於零，只是因為 $\mathbf{r} = \mathbf{0}$。
- 在第 (iii) 部分忘記驗證 $\mathbf{v}$ 永不為零；零速度會與任何向量「平行」。
- 把兩個環圈畫成左右並排：這個 8 字是直立的，兩個環圈分別在原點的上方和下方，因為 y 的範圍是 $[-2, 2]$，而 x 的範圍只有 $[-1, 1]$。

STEP 之外

$\mathbf{r} \cdot \mathbf{v} = \frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}|\mathbf{r}|^2$ 和 $xv_y - yv_x$ 這兩個量，分別是運動的徑向部分和（每單位質量計）對原點的角動量（angular momentum）；在極坐標（polar coordinates）中，後者就是 $r^2\dot{\theta}$。第 (iii) 部分說明角動量只在原點處為零，所以質點從不沿直線直接趨向或遠離 O。這條曲線本身就是 Gerono 雙紐線（lemniscate of Gerono），是眾多經典 8 字形曲線之一。Kleppner 與 Kolenkow 合著的 *An Introduction to Mechanics*（力學入門）正是以這種思路闡述極坐標和角動量。

詳細解答

STEP I 2005，第 12 題

概率與統計 • 溫氏圖與界限

STEP I 2005，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/05-s1-q12

關鍵思路

凡是問「哪個範圍的數值才不會自相矛盾？」的題目，其實都是關於一幅溫氏圖（Venn diagram）：圖中各個區域的值未知，但必須全部非負。先用盡量少的自由未知數表示每一個區域；這樣，每個「此區域 ≥ 0」的條件就變成一個界限，把所有界限合起來便得到所求範圍。在第 (ii) 部分，我們用一個計數論證（把三個單項概率相加時，每名巫師被數了多少次），毋須知道個別區域的值，便可求出「恰好兩件」的概率。

解答

第 (i) 部分

設 S 為哈比人抽煙斗的事件（event），H 為他戴帽子的事件。令 $t = \mathrm{P}(S \cap H)$。溫氏圖的四個區域是

$$\mathrm{P}(S \cap H) = t, \qquad \mathrm{P}(S \cap H') = 0.7 - t,$$
$$\mathrm{P}(S' \cap H) = 0.4 - t, \qquad \mathrm{P}(S' \cap H') = 1 - (0.7 + 0.4 - t) = t - 0.1.$$

只要所選的 t 令四者皆為非負，就得到一個真正的概率分佈（probability distribution），反之亦然。真正起約束作用的條件是 $t \geq 0.1$（來自最後一個區域）和 $t \leq 0.4$（來自第三個區域）；其餘條件隨之成立。所以 $0.1 \leq t \leq 0.4$，而由於 $p = 0.7 - t$，

$$0.3 \leq p \leq 0.6.$$

兩個端點都可以取到：當抽煙斗而又戴帽子的哈比人盡量少時（每個哈比人至少做兩件事的其中一件），$p = 0.6$；當每個戴帽子的哈比人都抽煙斗時，$p = 0.3$。

第 (ii) 部分：恰好兩件

設 H、C、R 分別為巫師戴帽子、披斗篷和戴戒指的事件。設 a、b 分別為巫師在這三件物品中恰好穿戴一件和恰好穿戴兩件的概率。已知一件也沒有的概率是 0.05，三件齊備的概率是 0.1，所以

$$0.05 + a + b + 0.1 = 1 \quad \Longrightarrow \quad a + b = 0.85.$$

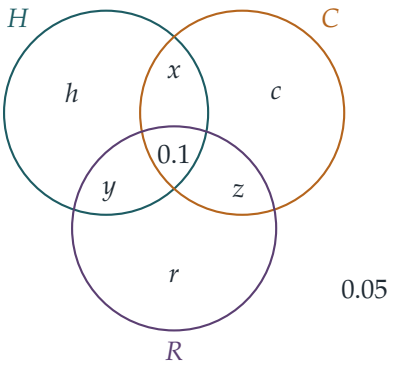
現在把三個已知概率相加：$\mathrm{P}(H) + \mathrm{P}(C) + \mathrm{P}(R) = 0.7 + 0.8 + 0.4 = 1.9$。在這個和之中，恰好穿戴 k 件物品的巫師被數了 k 次，所以

$$a + 2b + 3(0.1) = 1.9 \quad \Longrightarrow \quad a + 2b = 1.6.$$

兩式相減，得 $b = 0.75$（而 $a = 0.1$）。巫師恰好穿戴其中兩件物品的概率是 $\frac{3}{4}$。

第 (ii) 部分：q 的範圍

如圖所示，為圓內的七個區域標上記號：h、c、r 代表恰好一件，x、y、z 代表恰好兩件，中央是 0.1，圓外是 0.05。



由 H、C、R 各自的總和，以及第 (ii) 部分至今所得的結果，

$$h + x + y = 0.6, \qquad c + x + z = 0.7,$$
$$r + y + z = 0.3, \qquad x + y + z = 0.75.$$

利用最後一條方程，從首三條方程中各消去一對未知數，

$$h = z - 0.15, \qquad c = y - 0.05, \qquad r = x - 0.45.$$

（驗算一下：$h + c + r = 0.75 - 0.65 = 0.1 = a$。）所以整個分佈完全由滿足 $x + y + z = 0.75$ 的 x、y、z 決定，而它是一個真正的分佈，當且僅當六個區域全部非負：

$$x \geq 0.45, \qquad y \geq 0.05, \qquad z \geq 0.15.$$

想一想 x 最大可以是多少？

由於 $x = 0.75 - y - z$，而 $y + z \geq 0.05 + 0.15 = 0.2$，所以 $x \leq 0.55$，等號在 $y = 0.05$、$z = 0.15$ 時成立。最小值 $x = 0.45$ 也可以取到，例如取 $y = 0.1$、$z = 0.2$。兩者之間的每一個值同樣可以取到，因為我們可以連續地改變 y 和 z。因此

$$0.45 \leq x \leq 0.55.$$

最後，披着斗篷時戴帽子但不戴戒指，正好就是區域 x，所以

$$q = \mathrm{P}(H \cap R' \mid C) = \frac{\mathrm{P}(H \cap C \cap R')}{\mathrm{P}(C)} = \frac{x}{0.8}, \qquad \text{由此得} \qquad \frac{9}{16} \leq q \leq \frac{11}{16}.$$

驗算

- 在極端情況 $x = 0.55$、$y = 0.05$、$z = 0.15$ 下，我們得到 $h = 0$、$c = 0$、$r = 0.1$；八個區域之和為 1，並給出 $\mathrm{P}(H) = 0.7$、$\mathrm{P}(C) = 0.8$、$\mathrm{P}(R) = 0.4$，符合要求。
- 我們以 0.01 為步長，列舉了八個區域所有與數據相符的賦值：每一個都滿足 $x + y + z = 0.75$，而 q 的取值由 0.5625 至 0.6875。同樣地，在 (i) 中，p 的取值由 0.3 至 0.6。

常見錯誤

- 在 (i) 中，忘記了「兩者皆非」這個區域，因而得到下界 $t \geq 0$ 而不是 $t \geq 0.1$，結果得出錯誤的上限 $p \leq 0.7$。
- 在 (ii) 中，除以 1 而不是除以 $\mathrm{P}(C) = 0.8$：「已知他披着斗篷」這個條件會改變分母。
- 把 H、C、R 當作獨立(independent)事件。題目從沒有說它們獨立，而事實上 $\mathrm{P}(H \cap C \cap R) = 0.1 \neq 0.7 \times 0.8 \times 0.4$。

STEP 之外

某個量是未知概率的線性函數，而這些概率須滿足若干線性方程及非負條件，要求這個量的取值範圍，其實就是一個小型的線性規劃 (linear programme) 問題：極值總是出現在可行域 (feasible region) 的角點，這正是我們的極值都來自把某些區域設為零的原因。同樣的想法可得出第 (i) 部分所用的弗雷歇（Fréchet）界限 $\max(0, \mathrm{P}(A) + \mathrm{P}(B) - 1) \leq \mathrm{P}(A \cap B) \leq \min(\mathrm{P}(A), \mathrm{P}(B))$。想進一步了解容斥原理（inclusion–exclusion）和條件概率（conditional probability），可閱讀 Blitzstein 與 Hwang 的 *Introduction to Probability*（概率論入門教材）。

詳細解答

STEP I 2005，第 13 題

概率與統計 • 對稱分佈

STEP I 2005，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/05-s1-q13

關鍵思路

題目除了給出兩個累積概率（cumulative probability）和對稱性（symmetry）之外，對分佈的形狀隻字未提，所以一切都要由這些資料推導出來。關於 μ 對稱，還附帶送來兩個事實：μ 的兩側各佔一半概率，而且 $\mu - k\sigma$ 以下的尾部與 $\mu + k\sigma$ 以上的尾部概率相同。第 (ii) 部分的訣竅，是留意到題中提及的每一個容量（495、500、505 ml），對其中一種牛奶而言，都正好落在 μ、$\mu \pm \frac{1}{2}\sigma$ 或 $\mu + \sigma$ 的位置。

解答

全題把 X 視為連續（continuous）隨機變量（它描述的是容量），因此單一數值的概率為零，寫 $<$ 還是 $\leq$ 並無分別。

由關於 μ 的對稱性，可得 $\mathrm{P}(X \leq \mu) = \frac{1}{2}$，並且對任意 k 都有 $\mathrm{P}(X \leq \mu - k\sigma) = \mathrm{P}(X \geq \mu + k\sigma)$。特別地，

$$\mathrm{P}\left(X \leq \mu - \tfrac{1}{2}\sigma\right) = 1 - b, \qquad \mathrm{P}\left(X \leq \mu - \sigma\right) = 1 - a.$$

不妨以 σ 為單位，從 μ 起量度，把每一段區間的概率記下來：



例如 $\mathrm{P}(\mu < X \leq \mu + \frac{1}{2}\sigma) = b - \frac{1}{2}$，而 $\mathrm{P}(\mu + \frac{1}{2}\sigma < X \leq \mu + \sigma) = a - b$；左邊各段的概率則由對稱性得出。

第 (i) 部分

把累積概率相減，

$$\mathrm{P}\big(\mu-\tfrac{1}{2}\sigma\le X\le\mu+\sigma\big)=\mathrm{P}(X\le\mu+\sigma)-\mathrm{P}\big(X<\mu-\tfrac{1}{2}\sigma\big)=a-(1-b)=a+b-1.$$

至於條件概率（conditional probability），分子所對應的事件是 $\{\mu-\frac{1}{2}\sigma\le X\le\mu+\frac{1}{2}\sigma\}$，其概率為 $b-(1-b)=2b-1$，而 $\mathrm{P}(X\ge\mu-\frac{1}{2}\sigma)=1-(1-b)=b$。因此

$$\mathrm{P}\big(X\le\mu+\tfrac{1}{2}\sigma\mid X\ge\mu-\tfrac{1}{2}\sigma\big)=\frac{2b-1}{b}.$$

第 (ii)(a) 部分

設 S 為該盒牛奶是脫脂奶的事件，F 為該盒牛奶是全脂奶的事件，則 $\mathrm{P}(S)=0.6$ 及 $\mathrm{P}(F)=0.4$。設 V 為以 ml 計的容量。

> **想一想** 以 σ 為單位，500 ml 和 505 ml 在每種牛奶的分佈中各落在甚麼位置？

對脫脂奶（$\mu=500$，$\sigma=10$）而言，500 是 μ，505 是 $\mu+\frac{1}{2}\sigma$。對全脂奶（$\mu=495$，$\sigma=10$）而言，500 是 $\mu+\frac{1}{2}\sigma$，505 是 $\mu+\sigma$。從上圖各段讀出：

$$\begin{aligned}&\mathrm{P}(V<505\mid S)=b,&&\mathrm{P}(500<V<505\mid S)=b-\tfrac{1}{2},\\&\mathrm{P}(V<505\mid F)=a,&&\mathrm{P}(500<V<505\mid F)=a-b.\end{aligned}$$

根據全概率公式（law of total probability），

$$\begin{aligned}\mathrm{P}(V<505)&=0.6b+0.4a,\\\mathrm{P}(500<V<505)&=0.6\big(b-\tfrac{1}{2}\big)+0.4(a-b)=0.4a+0.2b-0.3.\end{aligned}$$

「多於 500 ml 而少於 505 ml」這個事件正是我們所需的交集，所以

$$\mathrm{P}(V>500\mid V<505)=\frac{0.4a+0.2b-0.3}{0.4a+0.6b}=\frac{4a+2b-3}{4a+6b}.$$

第 (ii)(b) 部分

隨機抽取的牛奶盒數量非常大時，觀察所得的比例可用來估計相應的概率。第一項資料表示 $\mathrm{P}(V\le505)=0.7$，按 (a) 部分的計算，即

$$0.6b+0.4a=0.7,\quad\text{即}\quad 4a+6b=7.\tag{1}$$

> **想一想** 對每種牛奶而言，$\mathrm{P}(V\ge495)$ 是多少？

對脫脂奶而言，495 是 $\mu - \frac{1}{2}\sigma$，所以 $\mathrm{P}(V \geq 495 \mid S) = b$（這正是第 (i) 部分的分母）。對全脂奶而言，495 是 μ，所以由對稱性得 $\mathrm{P}(V \geq 495 \mid F) = \frac{1}{2}$。根據貝葉斯定理（Bayes' theorem），

$$\mathrm{P}(F \mid V \geq 495) = \frac{0.4 \times \frac{1}{2}}{0.6b + 0.4 \times \frac{1}{2}} = \frac{0.2}{0.6b + 0.2}.$$

令此式等於 $\frac{1}{3}$，得 $0.6b + 0.2 = 0.6$，所以 $b = \frac{2}{3}$。於是由 (1) 得 $4a = 7 - 4 = 3$，所以

$$a \approx \tfrac{3}{4}, \qquad b \approx \tfrac{2}{3}.$$

這些只是估計值，因為觀察所得的比例只是真實概率的近似。

驗算

- 估計值合情合理：$\frac{1}{2} < b < a < 1$，這是必然的，因為 $\mu < \mu + \frac{1}{2}\sigma < \mu + \sigma$。
- 代回檢驗：$0.6 \cdot \frac{2}{3} + 0.4 \cdot \frac{3}{4} = 0.4 + 0.3 = 0.7$，以及 $0.2/(0.4 + 0.2) = \frac{1}{3}$。
- 若分佈為正態分佈（Normal distribution），則 $a \approx 0.841$ 及 $b \approx 0.691$；以這些數值計算，第 (i) 部分得出 0.533 和 0.553。以（非正態的）拉普拉斯分佈（Laplace distribution）作模擬，樣本數目超過 10^5，結果與第 (i) 及 (ii)(a) 部分的所有公式吻合；例如 (ii)(a) 的模擬結果為 0.2518，而公式給出 0.2515。

常見錯誤

- 把 $\mathrm{P}(X \leq \mu - \frac{1}{2}\sigma) = b$ 當作正確，而其實應是 $1 - b$：對稱性反射的是*尾部*，而不是累積概率。
- 忘記兩種牛奶的平均值不同，所以同一個容量（例如 500 ml）對兩種牛奶而言，與平均值相距的標準差（standard deviation）數目並不相同。
- 對兩種牛奶都用 $\mathrm{P}(V \geq 495) = b$；對全脂奶而言，495 ml 正是平均值。

STEP 之外

這道題的要點在於：即使對分佈一無所知，單憑對稱性和全概率公式也可以走得很遠；這裏由兩個總體混合而成的情況，是一個簡單的*混合模型*（mixture model），而第 (ii)(b) 部分則是根據數據估計其參數的入門版本。另一個同樣不依賴分佈形狀的相關結果是切比雪夫（Chebyshev）不等式，它只用方差（variance）便能給出 $\mathrm{P}(|X - \mu| \geq k\sigma)$ 的上界。可參閱 Blitzstein and Hwang, *Introduction to Probability*（概率論入門），或 Grimmett and Welsh, *Probability: An Introduction*（概率論導引）。

詳細解答

STEP I 2005，第 14 題

概率與統計 • 混合型隨機變量

STEP I 2005，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/05-s1-q14

關鍵思路

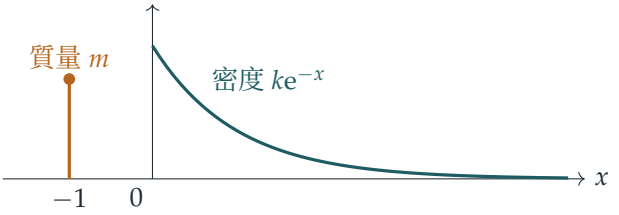
X 是一個混合型隨機變量（mixed random variable）：有一團概率集中在 -1 這一點上，其餘則連續地分佈在 $[0,\infty)$ 上。處理的方法是把每一項計算都拆成兩部分：來自 $X=-1$ 的離散項，加上在 $[0,\infty)$ 上對密度所作的積分。只要留意到連續部分不過是一個乘了常數的指數分佈（exponential distribution），大部分積分便都是我們早已認識的了。

解答

把題目給出的累積概率（cumulative probability）求導，可知 X 的連續部分的密度（density）為

$$f(x)=\frac{\mathrm{d}}{\mathrm{d}x}k(1-\mathrm{e}^{-x})=k\mathrm{e}^{-x},\qquad x\geq 0.$$

所以 X 的樣子是這樣的：在 -1 處有一個大小為 m 的點質量（point mass），在 $[0,\infty)$ 上則是曲線 $k\mathrm{e}^{-x}$。



第 (i) 部分

總概率為 1。令 $x \to \infty$，得 $\mathrm{P}(X \geq 0) = k(1-0) = k$，所以

$$m + k = 1, \qquad \text{即} \qquad k = 1 - m.$$

（由於 $m \geq 0$ 是一個概率，而且 $m < \frac{1}{2}$，我們有 $k > \frac{1}{2} > 0$，所以密度確實是正的。）

第 (ii) 部分

拆成兩部分，

$$\begin{aligned}\mathrm{E}(X) &= (-1)\cdot m + \int_0^\infty x\,(1-m)\mathrm{e}^{-x}\,\mathrm{d}x \\ &= -m + (1-m)\cdot 1 = 1 - 2m,\end{aligned}$$

其中用到 $\int_0^\infty x\mathrm{e}^{-x}\,\mathrm{d}x = 1$（可用分部積分法（integration by parts），或認出這是參數為 1 的指數分佈的平均值）。這正是所要求的結果。

第 (iii) 部分

方差（*variance*）。同樣地，利用 $\int_0^\infty x^2\mathrm{e}^{-x}\,\mathrm{d}x = 2$（分部積分兩次），

$$\mathrm{E}(X^2) = (-1)^2\cdot m + (1-m)\cdot 2 = 2 - m.$$

因此

$$\begin{aligned}\mathrm{Var}(X) &= \mathrm{E}(X^2) - \big(\mathrm{E}(X)\big)^2 = 2 - m - (1-2m)^2 \\ &= 1 + 3m - 4m^2 = (1-m)(1+4m).\end{aligned}$$

想一想 中位數（median）可以是 -1，或者落在 -1 與 0 之間嗎？

中位數。不可以：X 從不在 $(-1,0)$ 內取值，而且 $\mathrm{P}(X \leq -1) = m < \frac{1}{2}$，所以單靠點質量並不足以累積到一半的概率。條件 $m < \frac{1}{2}$ 正正是用在這裏。因此中位數 M 滿足 $M \geq 0$ 及 $\mathrm{P}(X \leq M) = \frac{1}{2}$：

$$\begin{aligned} m + (1-m)(1-\mathrm{e}^{-M}) &= \tfrac{1}{2} \\ 1 - (1-m)\mathrm{e}^{-M} &= \tfrac{1}{2} \\ \mathrm{e}^{-M} &= \frac{1}{2(1-m)}.\end{aligned}$$

於是

$$M = \ln\big(2(1-m)\big) = \ln(2-2m).$$

作為一致性檢查，由 $0 \le m < \frac{1}{2}$ 得 $1 < 2 - 2m \le 2$，所以 $M > 0$，與應有的情況相符。由於分佈在 $[0, \infty)$ 上是連續的，這個 M 同時滿足 $\mathrm{P}(X \ge M) = \frac{1}{2}$，所以中位數是毫不含糊的。

第 (iv) 部分

當 $X = -1$ 時，$|X|^{1/2} = 1$；當 $X \ge 0$ 時，$|X|^{1/2} = \sqrt{X}$。所以

$$\mathrm{E}\big(|X|^{1/2}\big) = 1 \cdot m + (1 - m) \int_0^\infty \sqrt{x}\, \mathrm{e}^{-x}\, \mathrm{d}x.$$

想一想 怎樣把 $\int_0^\infty \sqrt{x}\, \mathrm{e}^{-x}\, \mathrm{d}x$ 化成題目給出的那個積分？

題目給出的積分含有 e^{-y^2}，這提示我們代入 $x = y^2$（其中 $y \ge 0$），於是 $\sqrt{x} = y$ 及 $\mathrm{d}x = 2y\, \mathrm{d}y$；積分上下限 0 與 ∞ 保持不變：

$$\int_0^\infty \sqrt{x}\, \mathrm{e}^{-x}\, \mathrm{d}x = \int_0^\infty y\, \mathrm{e}^{-y^2}\, 2y\, \mathrm{d}y = 2 \int_0^\infty y^2 \mathrm{e}^{-y^2}\, \mathrm{d}y = 2 \cdot \tfrac{1}{4}\sqrt{\pi} = \tfrac{1}{2}\sqrt{\pi}.$$

因此

$$\mathrm{E}\big(|X|^{1/2}\big) = m + \tfrac{1}{2}\sqrt{\pi}\,(1 - m).$$

驗算

- 當 $m = 0$ 時，X 恰好是參數為 1 的指數分佈：我們的公式給出平均值 1、方差 1、中位數 $\ln 2$ 及 $\mathrm{E}(\sqrt{X}) = \frac{1}{2}\sqrt{\pi} = \Gamma(\frac{3}{2})$，全部都是標準數值。
- 當 $m \to \frac{1}{2}$ 時，中位數趨向 $\ln 1 = 0$，這是合理的：一半的概率位於 -1，所以中位數會滑落到連續部分的最底端。
- 取 $m = 0.3$ 並抽取 4×10^5 個樣本作模擬，得出平均值 0.402（公式為 0.4）、方差 1.551（公式為 1.54）、中位數 0.336（公式為 $\ln 1.4 \approx 0.336$）及 $\mathrm{E}(|X|^{1/2}) \approx 0.922$（公式為 0.920）。

常見錯誤

- 在 $\mathrm{E}(X^2)$ 或 $\mathrm{E}(|X|^{1/2})$ 中處理點質量時出錯：它的貢獻是 $+m$ 而不是 $-m$，因為 $(-1)^2 = |-1|^{1/2} = 1$；漏掉這一項同樣是錯的。
- 解 $(1 - m)(1 - \mathrm{e}^{-M}) = \frac{1}{2}$ 來求中位數，這樣便忽略了在 -1 處已經累積的概率 m。
- 換元時沒有把 $\mathrm{d}x$ 換成 $2y\, \mathrm{d}y$，結果得出 $\frac{1}{4}\sqrt{\pi}$ 而不是 $\frac{1}{2}\sqrt{\pi}$。

STEP 之外

同時包含點質量和密度的分佈在實際應用中十分常見（一宗保險索償有某個概率為零，否則便是連續的）。在大學裏，這類分佈會用統一的方法處理：把期望值（expectation）寫成對分佈函數（distribution function）的積分，$\mathrm{E}(g(X)) = \int g(x)\,\mathrm{d}F(x)$，一次過涵蓋離散、連續和混合三種情況。第 (iv) 部分的積分就是伽瑪函數（Gamma function）的值 $\Gamma(\frac{3}{2}) = \frac{1}{2}\sqrt{\pi}$。Grimmett 與 Welsh 合著的 *Probability: An Introduction*（概率論入門）是很好的下一步讀物。

詳細解答

STEP II 2005，第 1 題

純數學 • 多項式與駐點

STEP II 2005，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/05-s2-q1

關鍵思路

用積法則（product rule）對任何形如 $P(x)\mathrm{e}^{-x^2}$ 的式子求導，會得到 $\big(P'(x) - 2xP(x)\big)\mathrm{e}^{-x^2}$；由於 e^{-x^2} 永不為零，駐點（stationary point）正好就是多項式（polynomial）$P'(x) - 2xP(x)$ 的實根。於是第二部分變成一個逆向問題：先決定我們希望 $P' - 2xP$ 是哪一個多項式，然後透過比較係數（comparing coefficients）倒推出 P。

解答

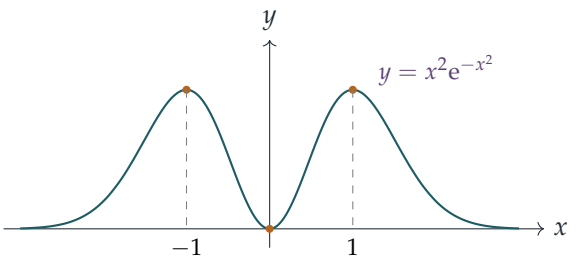
第一部分

由積法則，

$$\frac{\mathrm{d}}{\mathrm{d}x}\left(x^2\mathrm{e}^{-x^2}\right) = 2x\,\mathrm{e}^{-x^2} + x^2\cdot(-2x)\,\mathrm{e}^{-x^2} = 2x\big(1-x^2\big)\mathrm{e}^{-x^2}.$$

由於對所有 x 都有 $\mathrm{e}^{-x^2} > 0$，上式為零當且僅當 $x(1-x^2) = 0$，即

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



圖像在原點有一個極小點，並在 $x = \pm 1$ 處有高度為 e^{-1} 的極大點。

第二部分

對任何多項式 P，

$$\frac{\mathrm{d}}{\mathrm{d}x}\left(P(x)\mathrm{e}^{-x^2}\right) = \left(P'(x) - 2xP(x)\right)\mathrm{e}^{-x^2},$$

所以，同樣因為 $\mathrm{e}^{-x^2} > 0$，導數（derivative）為零之處正好就是 $P'(x) - 2xP(x) = 0$ 之處。第一部分正符合這個模式：$P(x) = x^2$ 給出 $P' - 2xP = -2x(x^2 - 1)$。因此我們的目標是

$$P'(x) - 2xP(x) = -2x\left(x^2 - a^2\right)\left(x^2 - b^2\right), \qquad (*)$$

它的實根只有 0、$\pm a$ 和 $\pm b$。這是五個相異的值，因為 a 和 b 是相異的正數。

> **想一想** P 的次數應該是多少？為甚麼它應該只含偶數次冪？

若 P 的次數為 n，首項係數（leading coefficient）為 c，則 P' 的次數為 $n-1$，所以 $P' - 2xP$ 的次數為 $n+1$，首項係數為 $-2c$。$(*)$ 的右方次數為 5，首項係數為 -2，所以我們需要 $n = 4$ 和 $c = 1$。此外，右方是一個奇函數（odd function）；若 P 是偶函數（even function），則 P' 和 $2xP$ 都是奇函數，所以很自然會嘗試

$$P(x) = x^4 + \beta x^2 + \gamma.$$

於是 $P'(x) = 4x^3 + 2\beta x$，而

$$\begin{aligned} P'(x) - 2xP(x) &= 4x^3 + 2\beta x - 2x^5 - 2\beta x^3 - 2\gamma x \\ &= -2x^5 + (4 - 2\beta)x^3 + (2\beta - 2\gamma)x, \end{aligned}$$

另一方面，$(*)$ 的右方展開後是

$$-2x^5 + 2\left(a^2 + b^2\right)x^3 - 2a^2b^2x.$$

比較 x^3 和 x 的係數：

$$\begin{aligned} 4 - 2\beta = 2\left(a^2 + b^2\right) \qquad &\Longrightarrow \quad \beta = 2 - a^2 - b^2, \\ 2\beta - 2\gamma = -2a^2b^2 \qquad &\Longrightarrow \quad \gamma = \beta + a^2b^2 = a^2b^2 - a^2 - b^2 + 2. \end{aligned}$$

所以其中一個合適的多項式是

$$\boxed{P(x) = x^4 + \left(2 - a^2 - b^2\right)x^2 + a^2b^2 - a^2 - b^2 + 2.}$$

這樣選取後，$(*)$ 恆成立，所以

$$\frac{\mathrm{d}}{\mathrm{d}x}\left(P(x)\mathrm{e}^{-x^2}\right) = -2x\left(x^2 - a^2\right)\left(x^2 - b^2\right)\mathrm{e}^{-x^2},$$

它在 $x = 0$、$x = \pm a$ 和 $x = \pm b$ 時為零，而且由於 $\mathrm{e}^{-x^2} \neq 0$，在其他 x 值都不為零。P 的任何非零常數倍同樣可行。

驗算

- 取 $a=1$、$b=2$：這時 $P(x)=x^4-3x^2+1$，而 $P'-2xP=4x^3-6x-2x^5+6x^3-2x=-2x^5+10x^3-8x=-2x(x^2-1)(x^2-4)$，符合要求。
- 數值上，當 $a=0.7$、$b=1.9$ 時，$P(x)\mathrm{e}^{-x^2}$ 的導數在 $[-5,5]$ 上只在 $x=0,\pm0.7,\pm1.9$ 附近變號；這條公式亦已用數千組隨機的 (a,b,x) 驗證過。
- 粗略地令 $b\to 0$：$P\to x^4+(2-a^2)x^2+2-a^2$，導數變成 $-2x^3(x^2-a^2)\mathrm{e}^{-x^2}$，位於 0 和 $\pm b$ 的根合而為一。

常見錯誤

- 漏掉鏈式法則（chain rule）帶來的因子：e^{-x^2} 的導數是 $-2x\,\mathrm{e}^{-x^2}$，而不是 e^{-x^2}。
- 為了配合五個根而猜 P 的次數是 5：對 $P\mathrm{e}^{-x^2}$ 求導會令次數增加一，所以 P 的次數是 4。
- 沒有說明為何沒有其他駐點：這需要用到 $\mathrm{e}^{-x^2}\neq 0$，以及該五次多項式只有這些根這個事實。

STEP 之外

映射 $P\mapsto P'-2xP$ 是多項式上的一個線性算子（linear operator），而解 $(*)$ 就是解一條一階線性微分方程（first-order linear differential equation），其積分因子（integrating factor）為 e^{-x^2}。形如 $H(x)\mathrm{e}^{-x^2}$（其中 H 為多項式）的函數引出埃爾米特多項式（Hermite polynomials），它們出現於概率（正態分佈，normal distribution）和量子諧振子（quantum harmonic oscillator）之中。Boyce 與 DiPrima 的 *Elementary Differential Equations*（初等微分方程教科書）涵蓋積分因子和埃爾米特方程（Hermite's equation）。

詳細解答

STEP II 2005，第 2 題

純數學 • 歐拉 φ 函數

STEP II 2005，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/05-s2-q2

關鍵思路

一切都源於把 N 寫成質因數分解（prime factorisation）的形式。每個因子 $\left(1-\frac{1}{p}\right)=\frac{p-1}{p}$ 的分母都有一個 p，而 N 必定含有至少一個 p 可以把它約去。一旦把 $\mathrm{f}(N)$ 寫成各質數冪（prime power）的乘積，便能清楚看出 $\mathrm{f}(m)\mathrm{f}(n)$ 與 $\mathrm{f}(mn)$ 何時相等：關鍵在於 m 與 n 有沒有公共的質因數。

解答

第 (i) 部分

(a) 整除 $12=2^2\cdot 3$ 的質數是 2 和 3，所以

$$\mathrm{f}(12)=12\left(1-\frac{1}{2}\right)\left(1-\frac{1}{3}\right)=12\cdot\frac{1}{2}\cdot\frac{2}{3}=4.$$

整除 $180=2^2\cdot 3^2\cdot 5$ 的質數是 2、3 和 5，所以

$$\mathrm{f}(180)=180\cdot\frac{1}{2}\cdot\frac{2}{3}\cdot\frac{4}{5}=48.$$

(b) 當 $N=1$ 時沒有質因數，乘積是空積（empty product），故 $\mathrm{f}(1)=1$。當 $N\geq 2$ 時，寫

$$N=p_1^{a_1}p_2^{a_2}\cdots p_k^{a_k},$$

其中 p_i 是整除 N 的相異質數，而每個 $a_i\geq 1$。於是

$$\begin{aligned}\mathrm{f}(N)&=p_1^{a_1}\cdots p_k^{a_k}\cdot\frac{p_1-1}{p_1}\cdots\frac{p_k-1}{p_k}\\&=p_1^{a_1-1}(p_1-1)\,p_2^{a_2-1}(p_2-1)\cdots p_k^{a_k-1}(p_k-1).\end{aligned}$$

由於每個 $a_i \geq 1$，每個指數 $a_i - 1$ 都是非負整數，所以每個因子 $p_i^{a_i-1}(p_i - 1)$ 都是正整數。因此對所有 N，$\mathrm{f}(N)$ 都是整數。本題餘下部分都會用到這條公式：

$$\mathrm{f}(N) = \prod_{i=1}^{k} p_i^{a_i-1}(p_i - 1), \qquad (*)$$

第 (ii) 部分

(a) 這是錯的。取 $m = n = 2$，則 $\mathrm{f}(2) = 2 \cdot \frac{1}{2} = 1$，所以 $\mathrm{f}(2)\mathrm{f}(2) = 1$，但 $\mathrm{f}(4) = 4 \cdot \frac{1}{2} = 2$。

問題在於因子 $\left(1 - \frac{1}{2}\right)$ 在左邊出現了兩次，在右邊卻只出現一次，因為 2 作為 4 的質因數只計算一次。

(b) 這是對的。若 p 和 q 是相異質數，則 $\mathrm{f}(p) = p\left(1 - \frac{1}{p}\right) = p - 1$，同理 $\mathrm{f}(q) = q - 1$。整除 pq 的質數恰好是 p 和 q，而且兩者不同，所以

$$\mathrm{f}(pq) = pq\left(1 - \frac{1}{p}\right)\left(1 - \frac{1}{q}\right) = (p-1)(q-1) = \mathrm{f}(p)\mathrm{f}(q).$$

(c) 這個命題是說：$\mathrm{f}(p)\mathrm{f}(q) = \mathrm{f}(pq)$ 只有在 p 和 q 是相異質數時才可能成立；換言之，這條等式迫使 p 和 q 必須是相異質數。這是錯的。取 $p = 3$ 和 $q = 4$；等式要有意義，兩者都不必是質數。於是

$$\mathrm{f}(3)\mathrm{f}(4) = 2 \cdot 2 = 4 = \mathrm{f}(12),$$

這裏用了第 (i)(a) 部分的結果，但 4 不是質數。（更簡單的例子是 $p = 1$，因為對任何 q 都有 $\mathrm{f}(1)\mathrm{f}(q) = \mathrm{f}(q)$。）

想一想 如果不是質數這個性質，那麼真正令等式成立的是甚麼？

真正重要的是 p 和 q 沒有公共的質因數。若 $\gcd(m, n) = 1$，則 m 與 n 各自所含的質數冪互不相同，所以 mn 的分解式不過是把兩個分解式並排寫在一起，而由 $(*)$ 可得 $\mathrm{f}(mn) = \mathrm{f}(m)\mathrm{f}(n)$。若有某個質數 r 同時整除兩者，則因子 $\left(1 - \frac{1}{r}\right)$ 在 $\mathrm{f}(m)\mathrm{f}(n)$ 中出現兩次，在 $\mathrm{f}(mn)$ 中卻只出現一次，所以 $\mathrm{f}(m)\mathrm{f}(n) < \mathrm{f}(mn)$。因此等式恰好在 $\gcd(m, n) = 1$ 時成立，而 3、4 只是其中一對互質（coprime）但並非兩個都是質數的數。

假如你把第 (ii) 部分中的 p 和 q 一律理解為質數，那麼 (c) 問的只是 $p = q$ 的情況會否出現，而答案是不會：$\mathrm{f}(p)^2 = (p-1)^2$，而 $\mathrm{f}(p^2) = p(p-1)$，由於 $p - 1 \neq p$，兩者永不相等。在這種理解下，(c) 是對的。不過，配合上面反例（counterexample）的一般理解，才是較自然的讀法。

第 (iii) 部分

由 $(*)$，質數冪滿足

$$\mathrm{f}(p^m) = p^{m-1}(p-1).$$

把目標數值分解，得 $146410 = 10 \times 14641 = 2 \cdot 5 \cdot 11^4$，因為 $14641 = 121^2 = 11^4$。這已經是 $p^{m-1}(p-1)$ 的形式，其中 $p = 11$：$11^4 \cdot 10 = 11^4(11-1)$，所以 $m - 1 = 4$。因此

$$p = 11, \quad m = 5: \qquad \mathrm{f}(11^5) = 11^4 \cdot 10 = 146410.$$

事實上這是唯一的答案。若 $m \geq 2$，則 p 整除 146410，所以 $p \in \{2, 5, 11\}$。取 $p = 2$ 會得到 $2^{m-1} = 146410$，這不可能，因為 146410 有奇質因數。取 $p = 5$ 會得到 $4 \cdot 5^{m-1} = 146410$，這也不可能，因為 146410 不能被 4 整除。餘下只有 $p = 11$，由此得 $m = 5$。若 $m = 1$，則需要 $p = 146411$ 是質數，但 $146411 = 41 \times 3571$。

驗算

- 這個函數就是歐拉函數（Euler's totient）：$\mathrm{f}(N)$ 計算 $\{1, \ldots, N\}$ 中與 N 互質的整數的個數。例如當 $N = 12$ 時，這樣的數有四個，就是 $1, 5, 7, 11$。經電腦驗證，對每個 $N < 2000$，乘積公式都與這個個數吻合，而且結果是整數。
- 題目中的示例給出 $\mathrm{f}(80) = 80 \cdot \frac{1}{2} \cdot \frac{4}{5} = 32 = 2^3(2-1) \cdot 5^0(5-1)$，與 $(*)$ 吻合。
- 電腦測試了 20 000 對隨機選取的 $m, n \leq 400$，發現 $\mathrm{f}(m)\mathrm{f}(n) = \mathrm{f}(mn)$ 恰好在 $\gcd(m, n) = 1$ 時成立。另外，搜尋所有質數 $p \leq 146411$ 後，發現 $(p, m) = (11, 5)$ 是 (iii) 的唯一解。

常見錯誤

- 在乘積中把同一個質數計算多於一次：對 $180 = 2^2 \cdot 3^2 \cdot 5$ 而言，因子有三個，不是五個。
- 把「僅當」（only if）當作「若」（if）。(b) 部分是「若」的方向，而 (c) 部分是它的逆命題（converse），兩者是不同的命題。
- 在 (iii) 中看不出 $14641 = 11^4$。留意到 $11^2 = 121$ 和 $121^2 = 14641$，或者認出帕斯卡三角形（Pascal's triangle）中 $1, 4, 6, 4, 1$ 這一行，便能很快找到。

STEP 之外

這個函數就是歐拉函數 φ，而第 (ii) 部分的結果說明它是積性（multiplicative）的：只要 $\gcd(m,n)=1$，便有 $\varphi(mn)=\varphi(m)\varphi(n)$，但一般情況下則不然。它出現在歐拉定理（Euler's theorem）之中：若 a 與 n 互質，則 $a^{\varphi(n)} \equiv 1 \pmod{n}$；這條定理推廣了費馬小定理（Fermat's little theorem），也是 RSA 加密法（RSA encryption）的基礎。在羣論（group theory）中，$\varphi(n)$ 是模 n 單位羣（group of units）的階（order）。Davenport 的 *The Higher Arithmetic*（一本經典的數論入門書）把這一切闡述得十分精彩。

詳細解答

STEP II 2005，第 3 題

純數學 • 分部積分法

STEP II 2005，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/05-s2-q3

關鍵思路

先求導：$\sin x - x\cos x$ 的導數（derivative）會化簡成一個正負號一目了然的式子，草圖和 y 的上下界便一併得出。兩個積分都是常規的分部積分法（integration by parts），平方項則借助二倍角公式（double-angle formulae）處理。最後的不等式來自兩者的比較：介乎 0 與 1 之間的數，不會小於它自己的平方。

解答

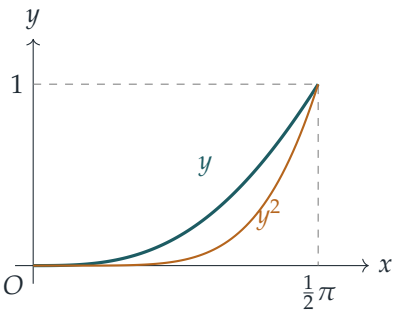
草圖與上下界

用積法則（product rule）求導，

$$\frac{\mathrm{d}y}{\mathrm{d}x} = \cos x - (\cos x - x\sin x) = x\sin x.$$

當 $0 < x \leq \frac{1}{2}\pi$ 時，x 和 $\sin x$ 均為正數，所以 y 在 $[0, \frac{1}{2}\pi]$ 上嚴格遞增（strictly increasing）。因此它的最小值是 $y(0) = 0 - 0 = 0$，最大值是 $y(\frac{1}{2}\pi) = 1 - \frac{1}{2}\pi \cdot 0 = 1$。故此 $0 \leq y \leq 1$，而且事實上當 $0 < x < \frac{1}{2}\pi$ 時有 $0 < y < 1$。

至於曲線的形狀，在 $(0, \frac{1}{2}\pi]$ 上 $\dfrac{\mathrm{d}^2y}{\mathrm{d}x^2} = \sin x + x\cos x > 0$，所以整條曲線都向上彎。在原點處斜率（gradient）為 0。利用馬克勞林級數（Maclaurin series），$y = \left(x - \frac{1}{6}x^3\right) - x\left(1 - \frac{1}{2}x^2\right) + \cdots = \frac{1}{3}x^3 + \cdots$，所以曲線離開原點時非常平坦，形狀像三次曲線。在 $x = \frac{1}{2}\pi$ 處斜率為 $\frac{1}{2}\pi$。



圖中也畫出了琥珀色的曲線 y^2，因為它是最後一部分的關鍵。

第 (i) 部分

我們需要求 $\int_0^{\pi/2} x\cos x\,\mathrm{d}x$。取 $u = x$、$\mathrm{d}v = \cos x\,\mathrm{d}x$ 作分部積分，

$$\int_0^{\pi/2} x\cos x\,\mathrm{d}x = \Big[x\sin x\Big]_0^{\pi/2} - \int_0^{\pi/2} \sin x\,\mathrm{d}x = \tfrac{1}{2}\pi - 1.$$

由於 $\int_0^{\pi/2} \sin x\,\mathrm{d}x = 1$，

$$\int_0^{\pi/2} y\,\mathrm{d}x = 1 - \left(\tfrac{1}{2}\pi - 1\right) = 2 - \frac{\pi}{2},$$

正是所求。

第 (ii) 部分

展開，

$$y^2 = \sin^2 x - 2x\sin x\cos x + x^2\cos^2 x = \sin^2 x - x\sin 2x + \tfrac{1}{2}x^2 + \tfrac{1}{2}x^2\cos 2x,$$

其中用到 $2\sin x\cos x = \sin 2x$ 和 $\cos^2 x = \frac{1}{2}(1+\cos 2x)$。我們逐一處理這四項。

第一，$\sin^2 x = \frac{1}{2}(1-\cos 2x)$，而 $\sin 2x$ 在上下限處均為零，所以

$$\int_0^{\pi/2} \sin^2 x\,\mathrm{d}x = \Big[\tfrac{1}{2}x - \tfrac{1}{4}\sin 2x\Big]_0^{\pi/2} = \frac{\pi}{4}.$$

其次，取 $u = x$、$\mathrm{d}v = \sin 2x\,\mathrm{d}x$ 作分部積分，並利用 $\cos\pi = -1$，

$$\int_0^{\pi/2} x\sin 2x\,\mathrm{d}x = \Big[-\tfrac{1}{2}x\cos 2x\Big]_0^{\pi/2} + \tfrac{1}{2}\int_0^{\pi/2} \cos 2x\,\mathrm{d}x = \frac{\pi}{4} + 0 = \frac{\pi}{4}.$$

第三，$\int_0^{\pi/2} \frac{1}{2}x^2\,\mathrm{d}x = \frac{1}{6}\left(\frac{1}{2}\pi\right)^3 = \dfrac{\pi^3}{48}$。

最後，取 $u = x^2$、$\mathrm{d}v = \cos 2x\,\mathrm{d}x$ 作分部積分；邊界項 $\left[\frac{1}{2}x^2 \sin 2x\right]_0^{\pi/2}$ 為零，而餘下的積分正是剛才求得的那一個：

$$\int_0^{\pi/2} x^2 \cos 2x\,\mathrm{d}x = 0 - \int_0^{\pi/2} x \sin 2x\,\mathrm{d}x = -\frac{\pi}{4},$$

所以最後一項的貢獻是 $\frac{1}{2}\cdot\left(-\frac{1}{4}\pi\right) = -\frac{1}{8}\pi$。

把四項相加，

$$\int_0^{\pi/2} y^2\,\mathrm{d}x = \frac{\pi}{4} - \frac{\pi}{4} + \frac{\pi^3}{48} - \frac{\pi}{8} = \frac{\pi^3}{48} - \frac{\pi}{8},$$

正是所求。

推論

想一想 兩個積分怎樣能夠得出一個關於 π 的不等式？

比較兩個被積函數(integrand)。由於 $0 \le y \le 1$，在整個區間上都有 $y - y^2 = y(1-y) \ge 0$。更進一步，我們已證明當 $0 < x < \frac{1}{2}\pi$ 時 $0 < y < 1$，所以在該處 $y(1-y) > 0$。一個非負的連續函數（continuous function），若在某個開區間上嚴格為正，它的積分必定嚴格為正。因此

$$\int_0^{\pi/2} y^2\,\mathrm{d}x < \int_0^{\pi/2} y\,\mathrm{d}x, \qquad \text{即} \qquad \frac{\pi^3}{48} - \frac{\pi}{8} < 2 - \frac{\pi}{2}.$$

兩邊乘以 48（它是正數，所以不等號方向不變），

$$\pi^3 - 6\pi < 96 - 24\pi, \qquad \text{所以} \qquad \pi^3 + 18\pi < 96.$$

這就是所要求的結果。

驗算

- 以數值計算，$2 - \frac{1}{2}\pi \approx 0.4292$，而 $\frac{1}{48}\pi^3 - \frac{1}{8}\pi \approx 0.2533$；對原來的被積函數使用辛普森法則（Simpson's rule），所得數值準確至小數點後十二位都相同。兩者都介乎 0 與 $\frac{1}{2}\pi \approx 1.571$ 之間，這是必然的，因為 $0 \le y \le 1$。
- 第二個積分比第一個小，正如圖中 y^2 位於 y 之下所顯示的那樣。
- 取 $\pi \approx 3.1416$，$\pi^3 + 18\pi \approx 31.006 + 56.549 = 87.55$，遠低於 96。

常見錯誤

- 在 $\int x \sin 2x\,\mathrm{d}x$ 中弄錯邊界項的正負號：在 $x = \frac{1}{2}\pi$ 處，$\cos 2x = \cos\pi = -1$，所以 $-\frac{1}{2}x\cos 2x$ 等於 $+\frac{1}{4}\pi$。

- 在最後一步只得出 $\leq$。要得到嚴格不等式，需要 $y(1-y) > 0$ 在某處成立，而不僅是 ≥ 0。
- 忘記說明為何 $y \leq 1$：只知道 $y(\frac{1}{2}\pi) = 1$ 並不足夠，還須用到 y 遞增這一事實。

STEP 之外

最後一步是一個重要想法的小小例子：函數之間的不等式，經積分後會變成數與數之間的不等式，而只要巧妙地選取函數，便可以把 π 這類常數困在某個範圍之內。同樣的策略可以由 $\int_0^1 \frac{x^4(1-x)^4}{1+x^2}\,\mathrm{d}x > 0$ 得出 $\frac{22}{7} > \pi$ 這個著名的證明，而更有分量的例子，是尼雲（Niven）對 π 是無理數（irrational）的證明。在大學第一門分析課程中，你會嚴格地證明：一個非負的連續函數，若在某處為正，它的積分必為正。Abbott 的 *Understanding Analysis*（《理解分析》）是一本平易近人的書，可以看到這個證明怎樣嚴謹地完成。

詳細解答

STEP II 2005，第 4 題

純數學 • 反正切恆等式

> **STEP II 2005，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/05-s2-q4

關鍵思路

處理反正切（inverse tangent）之和，方法是對等式兩邊取正切，再運用 $\tan(A+B) = \dfrac{\tan A + \tan B}{1 - \tan A \tan B}$。條件 $bc = a^2 + 1$ 恰好令分母可以因式分解（factorise），並與分子約去。其後的問題就在於結構：第一個結果是一條把一個反正切拆開成兩個的法則，而四項恆等式（identity）就是把這條法則用兩次，在樹的每一層各用一次。至於數值部分，則要倒推回去，找出是哪一種拆法產生了那四個數。

解答

第一個恆等式

設 $A = \arctan\frac{1}{a+b}$ 及 $B = \arctan\frac{1}{a+c}$。由於 $a, b, c > 0$，$\frac{1}{a+b}$ 與 $\frac{1}{a+c}$ 均為正數，所以 $A, B \in (0, \frac{\pi}{2})$，而 $A + B \in (0, \pi)$。利用 $bc = a^2 + 1$，

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

因此，根據複角公式（addition formula），

$$\begin{aligned}\tan(A+B) &= \frac{\dfrac{1}{a+b} + \dfrac{1}{a+c}}{1 - \dfrac{1}{(a+b)(a+c)}} = \frac{(a+c)+(a+b)}{(a+b)(a+c) - 1} \\ &= \frac{2a+b+c}{a(2a+b+c)} = \frac{1}{a},\end{aligned}$$

其中因為 $2a + b + c > 0$，所以可以約去。

想一想 由 $\tan(A+B)=\frac{1}{a}$ 真的可以得出 $A+B=\arctan\frac{1}{a}$ 嗎？

不是必然的：tan 的週期（period）為 π，所以我們需要 $A+B$ 位於 $(-\frac{\pi}{2},\frac{\pi}{2})$ 之內，即 arctan 的值域（range）。我們已知 $A+B\in(0,\pi)$。此外 $(a+b)(a+c)-1=a(2a+b+c)>0$，所以 $\tan A\tan B<1$，複角公式中的分母為正數；加上分子為正數，便得 $\tan(A+B)>0$，這就排除了 $A+B\in[\frac{\pi}{2},\pi)$ 的可能。（事實上，$\tan(A+B)$ 在 $\frac{\pi}{2}$ 處無定義，而在 $(\frac{\pi}{2},\pi)$ 上為負數。）所以 $A+B\in(0,\frac{\pi}{2})$，而 $\arctan\frac{1}{a}$ 是該區間內唯一一個正切值為 $\frac{1}{a}$ 的角，故

$$\arctan\frac{1}{a+b}+\arctan\frac{1}{a+c}=\arctan\frac{1}{a}. \qquad \blacksquare$$

四項恆等式

把第一個恆等式應用三次；每次所涉及的三個數均為正數。

- 取 $a=p+q$、$b=s$、$c=t$：條件為 $st=(p+q)^2+1$，這是已知的，所以

$$\arctan\frac{1}{p+q+s}+\arctan\frac{1}{p+q+t}=\arctan\frac{1}{p+q}.$$

- 取 $a=p+r$、$b=u$、$c=v$：條件 $uv=(p+r)^2+1$ 是已知的，所以

$$\arctan\frac{1}{p+r+u}+\arctan\frac{1}{p+r+v}=\arctan\frac{1}{p+r}.$$

- 取 $a=p$、$b=q$、$c=r$：條件 $qr=p^2+1$ 是已知的，所以

$$\arctan\frac{1}{p+q}+\arctan\frac{1}{p+r}=\arctan\frac{1}{p}.$$

把首兩式相加，再利用第三式，可知四個反正切之和為 $\arctan\frac{1}{p}$，正是所求。$\blacksquare$

數值恆等式

我們需要 $p=7$，所以 $qr=50$。與其猜測 q 和 r，不如由那四個數倒推回去。若 13 和 21 是某次使用第一個恆等式時的一對 $a+b$、$a+c$，則 $b=13-a$、$c=21-a$，而 $bc=a^2+1$ 給出

$$(13-a)(21-a)=a^2+1 \iff 273-34a=1 \iff a=8.$$

同樣地，若 82 和 187 是一對，則

$$(82-a)(187-a)=a^2+1 \iff 15334-269a=1 \iff a=57.$$

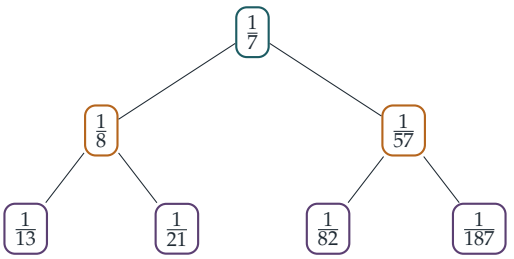
於是嘗試取 $p+q=8$ 及 $p+r=57$，即 $q=1$、$r=50$；然後 $s=13-8=5$、$t=21-8=13$、$u=82-57=25$、$v=187-57=130$。七個數全為正數，而三個條件均成立：

$$qr=50=7^2+1,\qquad st=65=8^2+1,\qquad uv=3250=57^2+1.$$

取這些值時，$p+q+s=13$、$p+q+t=21$、$p+r+u=82$ 及 $p+r+v=187$，所以由四項恆等式可得

$$\arctan\frac{1}{13}+\arctan\frac{1}{21}+\arctan\frac{1}{82}+\arctan\frac{1}{187}=\arctan\frac{1}{7}.\qquad\blacksquare$$

下面的樹顯示 $\arctan\frac{1}{7}$ 如何拆開。



每個節點的反正切都等於其子節點的反正切之和：$1\cdot 50=7^2+1$，$5\cdot 13=8^2+1$，$25\cdot 130=57^2+1$。

驗算

- 以數值計算，最後那個恆等式的兩邊均為 0.1418970546…。
- 以分數作精確計算：對於 $\frac{1}{13},\frac{1}{21}$，$\tan(A+B)$ 為 $\frac{34}{272}=\frac{1}{8}$；對於 $\frac{1}{82},\frac{1}{187}$，其值為 $\frac{269}{15333}=\frac{1}{57}$；而對於 $\frac{1}{8},\frac{1}{57}$，其值為 $\frac{65}{455}=\frac{1}{7}$。
- 第一個恆等式的一個特殊情況：取 $a=1$、$b=1$、$c=2$，便得著名的 $\arctan\frac{1}{2}+\arctan\frac{1}{3}=\arctan 1=\frac{\pi}{4}$。

常見錯誤

- 由 $\tan(A+B)=\frac{1}{a}$ 直接斷定 $A+B=\arctan\frac{1}{a}$，而沒有檢查 $A+B$ 是否位於 $(-\frac{\pi}{2},\frac{\pi}{2})$ 之內。
- 盲目地逐一嘗試 50、65 和 3250 的各種因數分解；其實由 13, 21 這一對倒推回去，一行便可求得 $a=8$。

STEP 之外

這類恆等式正是梅欽類公式（Machin-type formulae）背後的原理，例如 $\frac{\pi}{4} = 4\arctan\frac{1}{5} - \arctan\frac{1}{239}$；數百年來，人們一直用這些公式以人手計算 π，因為當 x 很小時，$\arctan x$ 的級數（series）收斂（converge）得很快。要看清這些恆等式，最簡潔的途徑是透過複數（complex numbers）：$\arctan\frac{1}{n}$ 是 $n+\mathrm{i}$ 的輻角（argument），所以反正切相加就對應於高斯整數（Gaussian integers）相乘，而條件 $bc = a^2+1$ 其實是 $a^2+1 = (a+\mathrm{i})(a-\mathrm{i})$ 的一個因式分解，只是換了個模樣。Needham 的 *Visual Complex Analysis* 深入闡述了這種從幾何角度看乘法的觀點，而 Davenport 的 *The Higher Arithmetic* 則討論了二平方和（sums of two squares）。

詳細解答

STEP II 2005，第 5 題

純數學 • 內切圓與外接圓

STEP II 2005，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/05-s2-q5

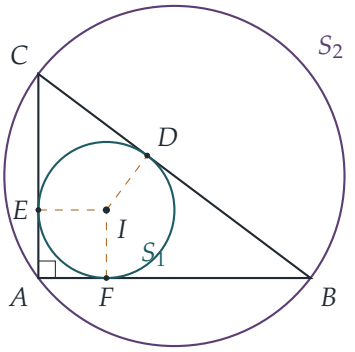
關鍵思路

由圓外一點向圓所作的兩條切線（tangent）長度相等。把這個性質用於每個頂點，三角形的每條邊便分成兩段切線長，而在直角那個頂點，兩段切線長剛好都是 r。這樣幾乎立刻就得到 r 的公式。至於面積比，直角告訴我們斜邊（hypotenuse）是 S_2 的直徑，而畢氏定理（Pythagoras' theorem）讓我們把一切只用 $b+c$ 和 a 表示，這就是 q 出現的原因。

解答

內切圓半徑

設 S_1 的圓心為 I，並分別與 BC、CA、AB 相切於 D、E、F。



半徑與切線互相垂直，所以 $IE \perp AC$ 及 $IF \perp AB$。連同 A 處的直角，可知 $AEIF$ 是長

方形，而且它的一對鄰邊 $IE = IF = r$，所以它是正方形，並且

$$AE = AF = r.$$

由圓外一點所作的切線長度相等，所以 $BF = BD$ 及 $CE = CD$。因此

$$c = AF + FB = r + BD, \qquad b = AE + EC = r + CD, \qquad a = BD + DC.$$

把首兩式相加，再減去第三式，

$$b + c - a = 2r + BD + CD - (BD + DC) = 2r,$$

這正是所需的結果。

面積比

由於 $\angle BAC = 90°$，根據半圓上的圓周角定理（angle in a semicircle theorem）的逆定理，BC 是外接圓（circumcircle）S_2 的直徑。所以 S_2 的半徑是 $\frac{1}{2}a$，面積是 $\frac{1}{4}\pi a^2$。

S_1 與三角形之間的區域，就是三角形除去 S_1 的圓盤後餘下的部分，所以它的面積是 $\frac{1}{2}bc - \pi r^2$。因此

$$R = \frac{\frac{1}{2}bc - \pi r^2}{\frac{1}{4}\pi a^2}, \qquad \text{所以} \qquad \pi R = \frac{2bc - 4\pi r^2}{a^2} = \frac{2bc}{a^2} - \pi\,\frac{(b + c - a)^2}{a^2},$$

這裏用了 $2r = b + c - a$。

想一想 怎樣把 $2bc/a^2$ 化成只含 q 的式子？

由畢氏定理得 $b^2 + c^2 = a^2$，所以 $2bc = (b + c)^2 - (b^2 + c^2) = (b + c)^2 - a^2$。除以 a^2，

$$\frac{2bc}{a^2} = q^2 - 1, \qquad \frac{(b + c - a)^2}{a^2} = \left(\frac{b + c}{a} - 1\right)^2 = (q - 1)^2.$$

因此

$$\begin{aligned} \pi R &= q^2 - 1 - \pi(q^2 - 2q + 1) \\ &= -(\pi - 1)q^2 + 2\pi q - (\pi + 1), \end{aligned}$$

這就是所求的結果。

R 的上界

右邊是 q 的二次式，首項係數 $-(\pi-1)$ 為負數，所以我們用配方法（completing the square）：

$$\begin{aligned}\pi R &= -(\pi-1)\Big(q^2 - \frac{2\pi}{\pi-1}\,q\Big) - (\pi+1)\\ &= -(\pi-1)\Big(q - \frac{\pi}{\pi-1}\Big)^2 + \frac{\pi^2}{\pi-1} - (\pi+1)\\ &= -(\pi-1)\Big(q - \frac{\pi}{\pi-1}\Big)^2 + \frac{\pi^2-(\pi^2-1)}{\pi-1}\\ &= \frac{1}{\pi-1} - (\pi-1)\Big(q - \frac{\pi}{\pi-1}\Big)^2.\end{aligned}$$

由於 $\pi-1>0$，平方項是被減去的，所以對任何實數 q 都有 $\pi R \le \dfrac{1}{\pi-1}$；再除以 $\pi>0$，

$$R \le \frac{1}{\pi(\pi-1)}.$$

這就是題目要求我們推導的結論。

值得留意的是，這個上界其實並不能達到。等號成立需要 $q=\pi/(\pi-1)\approx 1.467$，但對直角三角形而言，$(b+c)^2 = a^2+2bc \le a^2+(b^2+c^2) = 2a^2$（由算術平均—幾何平均不等式（AM–GM），$2bc \le b^2+c^2$），所以 $q \le \sqrt{2} \approx 1.414$。$R$ 的最大可能值來自等腰三角形，即 $q=\sqrt{2}$，此時 $R=(2\sqrt{2}\,\pi-3\pi+1)/\pi \approx 0.1467$，僅略低於上界 $1/(\pi(\pi-1))\approx 0.1486$。

驗算

對於 3–4–5 三角形（$b=4$，$c=3$，$a=5$），公式給出 $r=1$；而面積 6 除以半周長（semi-perimeter）6 的確是 1，這正是標準公式 $r=$ 面積$/s$。於是 $R=(6-\pi)/(\frac{25}{4}\pi)\approx 0.14558$，而把 $q=1.4$ 代入二次式也得到相同的值。隨機測試了 10^5 個直角三角形，每一次都驗證了 πR 的恆等式，而且沒有發現 R 超過 0.14674，與上界 0.14863 吻合。

常見錯誤

- 忘記 S_2 的半徑是 $\frac{1}{2}a$ 而不是 a：斜邊是直徑。
- 試圖直接展開 $2bc/a^2$，而沒有利用 $b^2+c^2=a^2$ 把它寫成 $2bc=(b+c)^2-a^2$。
- 沒有檢查 $q=\pi/(\pi-1)$ 是否可以取到，就聲稱上界可以達到；其實並不能達到，不過題目只要求證明不等式。

STEP 之外

一般三角形的內切圓（incircle）與外接圓，由歐拉（Euler）公式 $d^2 = R_2(R_2 - 2r)$ 聯繫起來，其中 d 是兩圓圓心之間的距離，R_2 是外接圓半徑（circumradius）；由此可得 $R_2 \geq 2r$，而等號只在等邊三角形時成立。這類結果，以及在一族圖形中求某個比值的最優值這種想法，都通往幾何不等式（geometric inequalities）的研究。Coxeter 與 Greitzer 合著的 *Geometry Revisited*（《幾何學的再探索》）是很優美的下一步讀物，而 Siklos 的 *Advanced Problems in Mathematics*（《高等數學難題》，免費，採用 CC BY 授權）則有更多 STEP 風格的練習。

詳細解答

STEP II 2005，第 6 題

純數學 • 二項式級數的和

> **STEP II 2005，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/05-s2-q6

關鍵思路

這裏每一個和都是某個冪級數(power series)在特定的 x 處的取值，所以要做的就是認出它來自哪一個函數。在第 (i) 部分，竅門是把 n 和 n^2 改寫成 1、$n+1$ 和 $(n+1)(n+2)$ 的組合，而這些正是我們剛剛寫下的係數。在第 (ii) 部分，只要留意到 $(1-x)^{-1/2}$ 和 $(1-x)^{-3/2}$ 的係數之間的關係，同樣的想法就行得通。

解答

第 (i) 部分

根據負整數指數的二項式定理（binomial theorem）（當 $|x|<1$ 時成立），對於 $n \geq 0$，x^n 的係數（coefficient）為

$$(1-x)^{-1}: \quad 1, \qquad (1-x)^{-2}: \quad n+1, \qquad (1-x)^{-3}: \quad \tfrac{1}{2}(n+1)(n+2).$$

例如，$(1-x)^{-3}$ 中的係數是 $\dfrac{(-3)(-4)\cdots(-3-n+1)}{n!}(-1)^n = \dfrac{3\cdot 4\cdots(n+2)}{n!} = \dfrac{(n+1)(n+2)}{2}$。

第一個和。由於 $n=(n+1)-1$，當 $|x|<1$ 時

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

（把和拆開是可以的，因為右邊兩個級數都收斂（converge）。）$n=0$ 的一項為零，所以代入 $x=\frac{1}{2}$，

$$\sum_{n=1}^{\infty} n\,2^{-n} = 4-2 = 2.$$

第二個和。我們想把 n^2 寫成 $(n+1)(n+2)$、$n+1$ 和 1 的組合。由於 $(n+1)(n+2) = n^2+3n+2$，

$$n^2 = (n+1)(n+2) - 3(n+1) + 1.$$

因此，當 $|x| < 1$ 時，

$$\sum_{n=0}^{\infty} n^2 x^n = \frac{2}{(1-x)^3} - \frac{3}{(1-x)^2} + \frac{1}{1-x},$$

其中出現因子 2，是因為 $(1-x)^{-3}$ 中的係數是 $(n+1)(n+2)$ 的一半。同樣，$n=0$ 的一項為零，而在 $x = \frac{1}{2}$ 處

$$\sum_{n=1}^{\infty} n^2 2^{-n} = 2\cdot 8 - 3\cdot 4 + 2 = 6.$$

第 (ii) 部分

根據指數為 $-\frac{1}{2}$ 的二項式定理，$(1-x)^{-1/2}$ 中 x^n 的係數是

$$\frac{(-\frac{1}{2})(-\frac{3}{2})\cdots(-\frac{2n-1}{2})}{n!}(-1)^n = \frac{1\cdot 3\cdot 5\cdots(2n-1)}{2^n\, n!}.$$

想一想 怎樣把奇數的連乘積化成階乘（factorial）？

把缺少的偶數補上：$(2n)! = \big(1\cdot 3\cdots(2n-1)\big)\big(2\cdot 4\cdots 2n\big)$，而 $2\cdot 4\cdots 2n = 2^n n!$。所以 $1\cdot 3\cdots(2n-1) = \dfrac{(2n)!}{2^n n!}$，而係數就是

$$a_n = \frac{(2n)!}{2^n n!\cdot 2^n n!} = \frac{(2n)!}{(n!)^2}\frac{1}{2^{2n}}.$$

這對 $n=0$ 也成立（空積，$a_0 = 1$），所以當 $|x| < 1$ 時

$$(1-x)^{-1/2} = \sum_{n=0}^{\infty} \frac{(2n)!}{(n!)^2}\frac{x^n}{2^{2n}},$$

正如所求。

第一個和。題目所求的和是 $\sum a_n x^n$ 在 $x = \frac{1}{3}$ 處的值，而這一點位於展開式成立的區間之內。所以它等於

$$\left(1-\tfrac{1}{3}\right)^{-1/2} = \sqrt{\tfrac{3}{2}} = \frac{\sqrt{6}}{2}.$$

第二個和。現在我們需要 $\sum n a_n x^n$。按照第 (i) 部分的思路，把 na_n 與一個我們已知的級數聯繫起來。與上面完全一樣，$(1-x)^{-3/2}$ 中 x^m 的係數是

$$b_m = \frac{3\cdot 5\cdots(2m+1)}{2^m\, m!} = (2m+1)\, a_m.$$

由 a_n 的公式，相鄰兩個係數之比是 $a_n/a_{n-1} = \dfrac{(2n)(2n-1)}{n^2 \cdot 4} = \dfrac{2n-1}{2n}$，所以對於 $n \geq 1$

$$n\,a_n = \tfrac{1}{2}(2n-1)\,a_{n-1} = \tfrac{1}{2}\,b_{n-1}.$$

因此，當 $|x| < 1$ 時，

$$\sum_{n=1}^{\infty} n a_n x^n = \frac{x}{2}\sum_{n=1}^{\infty} b_{n-1}x^{n-1} = \frac{x}{2}(1-x)^{-3/2}.$$

（等價的做法是：把 $(1-x)^{-1/2}$ 的級數逐項求導（differentiate term by term），這在它的收斂區間（interval of convergence）內是容許的，然後乘以 x。）在 $x = \frac{1}{3}$ 處：

$$\sum_{n=1}^{\infty} \frac{n\,(2n)!}{(n!)^2\,2^{2n}\,3^n} = \frac{1}{6}\left(\frac{3}{2}\right)^{3/2} = \frac{1}{6}\cdot\frac{3}{2}\sqrt{\frac{3}{2}} = \frac{\sqrt{6}}{8}.$$

驗算

- $\sum n2^{-n}$ 的首幾項是 $\frac{1}{2}+\frac{1}{2}+\frac{3}{8}+\frac{1}{4}+\cdots$，加到第四項已經是 1.625，並慢慢向 2 靠近；$\sum n^2 2^{-n}$ 加到 200 項的部分和（partial sum）與 6 吻合至小數點後很多個位。
- 數值上，$\sqrt{6}/2 \approx 1.224745$ 而 $\sqrt{6}/8 \approx 0.306186$；第 (ii) 部分兩個級數的部分和與這兩個值吻合至小數點後 12 個位。
- 對通項作一個簡單的合理性檢查：$n = 1$ 給出 $a_1 = \frac{2}{4} = \frac{1}{2}$，正是 $(1-x)^{-1/2} = 1 + \frac{1}{2}x + \frac{3}{8}x^2 + \cdots$ 中那個熟悉的係數。

常見錯誤

- 忘記了 $(1-x)^{-3}$ 的係數中的 $\frac{1}{2}$，結果得出 $8 - 12 + 2 = -2$ 作為 $\sum n^2 2^{-n}$ 的值，而正項之和不可能是負數。
- 把 $(1-x)^{-2}$ 的級數的係數寫成 n 而不是 $n+1$：這個級數是 $1 + 2x + 3x^2 + \cdots$。
- 在最後一個和中，漏掉一個因子 x：$\sum n a_n x^{n-1}$ 是導數，但題目所求的和是 $\sum n a_n x^n$。

STEP 之外

這裏用到的手法（把級數拆成幾部分、逐項求導）之所以合理，是因為冪級數在其收斂半徑（radius of convergence）之內絕對收斂（converges absolutely），並且可以在那裏逐項求導。嚴謹地證明這一點，是大學第一門分析（analysis）課程的標準內容。係數 $\binom{2n}{n}/4^n$ 也會在概率中出現，它就是簡單隨機漫步（simple random walk）在走了 $2n$ 步之後回到

起點的概率。關於分析方面的內容，可參閱 Abbott 的 *Understanding Analysis*（一本分析入門書）。

詳細解答

STEP II 2005，第 7 題

純數學 • 向量與圓周運動

> **STEP II 2005，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/05-s2-q7

關鍵思路

每個位置向量（position vector）都具有 cos(角) $\mathbf{a}$ + sin(角) $\mathbf{b}$ 的形式；當 $\mathbf{a}$ 與 $\mathbf{b}$ 互相垂直且長度相等時，它在 $\mathbf{a}$ 與 $\mathbf{b}$ 所在的平面上描出一個圓。一旦知道 $|OP|$ 和 $|OQ|$ 是常數，角 POQ 便可直接由純量積（scalar product）求得，再用積化和差公式（product-to-sum formulae）把它整理成單一個餘弦。最後一部分就變成在兩個完整周期內關於一個餘弦的不等式。

解答

第 (i) 部分

以 $\mathbf{p}$ 和 $\mathbf{q}$ 表示 P 和 Q 的位置向量。

對於 P，$\mathbf{p} = \cos t\,\mathbf{i} + \sin t\,\mathbf{j}$，所以 $|\mathbf{p}| = 1$，而且 $\mathbf{p}$ 位於 xy 平面上。因此 P 在平面 $z = 0$ 上沿着以 O 為圓心、半徑為 1 的圓運動；從上方（即從正 z 軸方向）看是逆時針方向，角速率（angular speed）恆為 1（所以速率為 1）。它由 $(1, 0, 0)$ 出發，當 t 由 0 變到 2π 時剛好轉一圈。

對於 Q，設 $s = t + \frac{1}{4}\pi$，並設

$$\mathbf{a} = \tfrac{3}{2}\mathbf{i} + \tfrac{3\sqrt{3}}{2}\mathbf{k}, \qquad \mathbf{b} = 3\mathbf{j}, \qquad \text{於是} \qquad \mathbf{q} = \cos s\,\mathbf{a} + \sin s\,\mathbf{b}.$$

現在 $|\mathbf{a}|^2 = \frac{9}{4} + \frac{27}{4} = 9$，$|\mathbf{b}|^2 = 9$，且 $\mathbf{a} \cdot \mathbf{b} = 0$。因此

$$|\mathbf{q}|^2 = 9\cos^2 s + 9\sin^2 s + 2\cos s \sin s\,(\mathbf{a} \cdot \mathbf{b}) = 9,$$

而且 $\mathbf{q}$ 始終位於通過 O、由 $\mathbf{a}$ 和 $\mathbf{j}$ 張成的平面上。該平面包含 y 軸和方向 $\mathbf{i} + \sqrt{3}\,\mathbf{k}$，所以它的方程是 $z = \sqrt{3}\,x$；它的法線（normal）$-\sqrt{3}\,\mathbf{i} + \mathbf{k}$ 與 $\mathbf{k}$ 的夾角為 $\arccos \frac{1}{2} = \frac{1}{3}\pi$。

所以 Q 在平面 $z = \sqrt{3}\,x$ 上沿着以 O 為圓心、半徑為 3 的圓運動；這個平面包含 y 軸，並與 P 的圓所在的平面成 $60°$ 傾角。它的角速率同樣是 1（速率為 3），也是轉一圈，由 **a** 的方向轉向 **j**，而它的相位（phase）始終比 P 的領先 $\frac{1}{4}\pi$：在 $t = 0$ 時，它位於與 **a** 成 $\frac{1}{4}\pi$ 角之處，即 $\frac{3}{2\sqrt{2}}(1,\, 2,\, \sqrt{3})$。

第 (ii) 部分

由於 $|\mathbf{p}| = 1$ 及 $|\mathbf{q}| = 3$，

$$\cos\theta = \frac{\mathbf{p}\cdot\mathbf{q}}{3} = \frac{\frac{3}{2}\cos t\cos s + 3\sin t\sin s}{3} = \tfrac{1}{2}\cos t\cos s + \sin t\sin s,$$

其中 $s = t + \frac{1}{4}\pi$。利用積化和差公式，並注意 $t + s = 2t + \frac{1}{4}\pi$ 及 $s - t = \frac{1}{4}\pi$，可得

$$\cos t\cos s = \tfrac{1}{2}\big[\cos(2t + \tfrac{1}{4}\pi) + \cos\tfrac{1}{4}\pi\big],$$
$$\sin t\sin s = \tfrac{1}{2}\big[\cos\tfrac{1}{4}\pi - \cos(2t + \tfrac{1}{4}\pi)\big].$$

記 $C = \cos(2t + \frac{1}{4}\pi)$，並利用 $\cos\frac{1}{4}\pi = \frac{\sqrt{2}}{2}$，

$$\begin{aligned}\cos\theta &= \tfrac{1}{4}\big(C + \tfrac{\sqrt{2}}{2}\big) + \tfrac{1}{2}\big(\tfrac{\sqrt{2}}{2} - C\big)\\ &= \tfrac{\sqrt{2}}{8} + \tfrac{\sqrt{2}}{4} - \tfrac{1}{4}C = \tfrac{3\sqrt{2}}{8} - \tfrac{1}{4}\cos(2t + \tfrac{1}{4}\pi),\end{aligned}$$

這正是所要證明的結果。

第 (iii) 部分

想一想 怎樣把 $\theta \geq \frac{1}{4}\pi$ 轉化為關於 $\cos\theta$ 的條件？

餘弦在 $[0, \pi]$ 上嚴格遞減（strictly decreasing），而 θ 和 $\frac{1}{4}\pi$ 都在這個區間內，所以

$$\theta \geq \tfrac{1}{4}\pi \iff \cos\theta \leq \cos\tfrac{1}{4}\pi = \tfrac{\sqrt{2}}{2} = \tfrac{4\sqrt{2}}{8}.$$

由第 (ii) 部分，這等價於

$$\tfrac{3\sqrt{2}}{8} - \tfrac{1}{4}\cos(2t + \tfrac{1}{4}\pi) \leq \tfrac{4\sqrt{2}}{8} \iff \cos(2t + \tfrac{1}{4}\pi) \geq -\tfrac{\sqrt{2}}{2}.$$

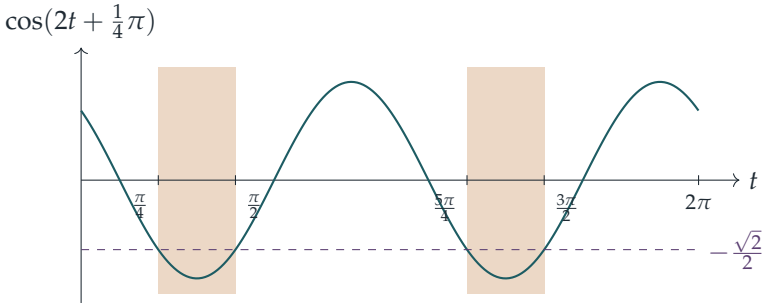
設 $u = 2t + \frac{1}{4}\pi$。當 t 取遍 $[0, 2\pi]$ 時，u 取遍 $[\frac{1}{4}\pi, \frac{17}{4}\pi]$，這個區間的長度是 4π：剛好是 $\cos u$ 的兩個周期。在任何一個周期內，$\cos u < -\frac{\sqrt{2}}{2}$ 恰好在一個長度為 $\frac{1}{2}\pi$ 的開區間上成立（即 $\frac{3}{4}\pi$ 與 $\frac{5}{4}\pi$ 之間，模 2π）。在這裏，不符合條件的值是

$$u \in (\tfrac{3}{4}\pi, \tfrac{5}{4}\pi) \cup (\tfrac{11}{4}\pi, \tfrac{13}{4}\pi) \iff t \in (\tfrac{1}{4}\pi, \tfrac{1}{2}\pi) \cup (\tfrac{5}{4}\pi, \tfrac{3}{2}\pi),$$

以 t 計算，每段的長度為 $\frac{1}{4}\pi$。所以 $\theta < \frac{1}{4}\pi$ 的總時間是 $\frac{1}{2}\pi$，而 $\theta \geq \frac{1}{4}\pi$ 的總時間是

$$2\pi - \tfrac{1}{2}\pi = \tfrac{3}{2}\pi,$$

這正是所要證明的結果。



陰影部分是曲線降到 $-\frac{\sqrt{2}}{2}$ 以下的地方，也就是 $\theta < \frac{1}{4}\pi$ 的地方。

驗算

- 在 $t = 0$ 時：$\mathbf{p} = \mathbf{i}$ 及 $\mathbf{q} = \frac{3}{2\sqrt{2}}(\mathbf{i} + 2\mathbf{j} + \sqrt{3}\,\mathbf{k})$，所以直接計算得 $\cos\theta = \frac{3}{2\sqrt{2}} \cdot \frac{1}{3} = \frac{\sqrt{2}}{4}$。公式則給出 $\frac{3\sqrt{2}}{8} - \frac{1}{4} \cdot \frac{\sqrt{2}}{2} = \frac{\sqrt{2}}{4}$。兩者吻合。
- $\cos\theta$ 介乎 $\frac{3\sqrt{2}}{8} - \frac{1}{4} \approx 0.280$ 與 $\frac{3\sqrt{2}}{8} + \frac{1}{4} \approx 0.780$ 之間，所以 θ 永遠不會是 0。這是合理的：$\mathbf{p}$ 與 $\mathbf{q}$ 只可能沿着兩個平面相交的 y 軸平行，但 P 在 $t = \frac{1}{2}\pi, \frac{3}{2}\pi$ 時位於該軸上，而 Q 則在 $t = \frac{1}{4}\pi, \frac{5}{4}\pi$ 時位於該軸上。
- 在 $[0, 2\pi]$ 內對 t 取二百萬個樣本點，並直接由向量計算 θ，得出 $\theta \geq \frac{1}{4}\pi$ 的總時間為 4.712389，與 $\frac{3}{2}\pi \approx 4.712389$ 相符。

常見錯誤

- 忘記除以 $|\mathbf{p}||\mathbf{q}| = 3$，或因為 $\mathbf{i}$ 和 $\mathbf{j}$ 的係數不同而以為 $|\mathbf{q}|$ 會變化。
- 把不等號的方向弄反：$\theta \geq \frac{1}{4}\pi$ 表示 $\cos\theta \leq \frac{\sqrt{2}}{2}$，因為餘弦在 $[0, \pi]$ 上遞減。
- 只計算了 $\cos(2t + \frac{1}{4}\pi)$ 的一個周期：$2t$ 表示當 t 轉一圈時，它轉了兩圈。

STEP 之外

任何以原點為圓心的圓都可以寫成 $\cos s\,\mathbf{a} + \sin s\,\mathbf{b}$，其中 $\mathbf{a} \perp \mathbf{b}$ 且 $|\mathbf{a}| = |\mathbf{b}|$；如果去掉這些條件，便會得到一個橢圓（ellipse），而一般來說，得到的是單位圓在線性映射（linear map）下的像。以這種方式理解曲線，並用純量積量度運動中的向量之間的夾角，正是大學一年級所學的線性代數（linear algebra）和向量微積分（vector calculus）的起點。Strang 的 *Introduction to Linear Algebra*（線性代數入門）把這些概念闡述得十分清晰。

詳細解答

STEP II 2005，第 8 題

純數學 • 可分離變量微分方程

STEP II 2005，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/05-s2-q8

關鍵思路

這條方程是可分離變數(separable)的：與 y 有關的部分只有 y^2 這一個因子，所以除以它之後便得到 $\frac{1}{y^2}\frac{\mathrm{d}y}{\mathrm{d}x}$，也就是 $-1/y$ 的導數。對 x 的積分可用代換 (substitution) $u = 1 + x^2$ 來處理，因為奇數次冪 x^3 正好提供了 u 所需要的 $x\,\mathrm{d}x$。至於第二條曲線，留意右邊是 z^3 而不是 y^2：這時應該考慮的量是 $1/z^2$，而它原來與 $1/y$ 滿足完全相同的方程。

解答

求 y

在 C 上有 $y(0) = 1 > 0$，而當 $x \geq 0$ 時方程的右邊永不為負，所以 y 永不遞減。因此 $y \geq 1$ 恆成立，我們可以除以 y^2：

$$\frac{1}{y^2}\frac{\mathrm{d}y}{\mathrm{d}x} = \frac{x^3}{(1+x^2)^{5/2}} \quad\Longrightarrow\quad -\frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{1}{y}\right) = \frac{x^3}{(1+x^2)^{5/2}}.$$

計算積分時，設 $u = 1 + x^2$，於是 $\mathrm{d}u = 2x\,\mathrm{d}x$，而 $x^3\,\mathrm{d}x = x^2 \cdot x\,\mathrm{d}x = \frac{1}{2}(u-1)\,\mathrm{d}u$：

$$\begin{aligned}\int \frac{x^3}{(1+x^2)^{5/2}}\,\mathrm{d}x &= \frac{1}{2}\int (u^{-3/2} - u^{-5/2})\,\mathrm{d}u = -u^{-1/2} + \tfrac{1}{3}u^{-3/2} + \text{常數} \\ &= -\frac{3u-1}{3u^{3/2}} + \text{常數} = -\frac{2+3x^2}{3(1+x^2)^{3/2}} + \text{常數}.\end{aligned}$$

因此

$$\frac{1}{y} = \frac{2+3x^2}{3(1+x^2)^{3/2}} + k$$

其中 k 為某個常數。在 $x=0$ 處，上式變成 $1=\frac{2}{3}+k$，所以 $k=\frac{1}{3}$，這就是所要求的結果：

$$\frac{1}{y}=\frac{2+3x^2}{3(1+x^2)^{3/2}}+\frac{1}{3}.$$

x 很大時

把主導的冪次提出來：從分子提出 $3x^2$，從 $(1+x^2)^{3/2}$ 提出 x^3：

$$\frac{2+3x^2}{3(1+x^2)^{3/2}}=\frac{3x^2\left(1+\frac{2}{3x^2}\right)}{3x^3\left(1+\frac{1}{x^2}\right)^{3/2}}=\frac{1}{x}\left(1+\frac{2}{3x^2}\right)\left(1-\frac{3}{2x^2}+\cdots\right)=\frac{1}{x}+O\left(\frac{1}{x^3}\right),$$

這裏用了二項展開（binomial expansion）；由於 $1/x^2<1$ 對很大的 x 成立，展開式有效。所以

$$\frac{1}{y}=\frac{1}{3}+\frac{1}{x}+O\left(\frac{1}{x^3}\right)=\frac{1}{3}\left(1+\frac{3}{x}+O\left(\frac{1}{x^3}\right)\right),$$

再取倒數，並利用 $(1+t)^{-1}=1-t+O(t^2)$，得

$$y=3\left(1-\frac{3}{x}+O\left(\frac{1}{x^2}\right)\right)\approx 3-\frac{9}{x},$$

正如題目所求。

> ? **想一想** 為甚麼可以捨去 $\frac{2}{3x^2}$ 和 $-\frac{3}{2x^2}$ 這兩項，卻不能捨去 $\frac{1}{x}$？

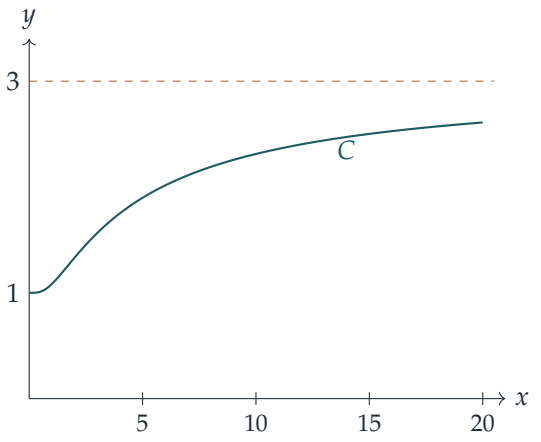
目標 $3-\frac{9}{x}$ 準確至 $\frac{1}{x}$ 階。在 $1/y$ 中，各修正項比 $\frac{1}{x}$ 還要小一個 $\frac{1}{x^2}$ 的因子，所以它們只會在 $\frac{1}{x^2}$ 階或更高階影響 y。而 $\frac{1}{x}$ 這一項本身，正是產生 $-\frac{9}{x}$ 的來源。

C 的略圖

現在把曲線的特徵整理出來。

- $y(0)=1$，而且 $\mathrm{d}y/\mathrm{d}x=0$ 於 $x=0$ 處成立。在原點附近 $\mathrm{d}y/\mathrm{d}x\approx x^3$，所以 $y\approx 1+\frac{1}{4}x^4$：曲線起初非常平坦。
- 當 $x>0$ 時斜率嚴格為正，所以 y 遞增。
- $1/y$ 的第一項為正，所以 $1/y>\frac{1}{3}$，由此得 $1\le y<3$ 對所有 x 成立。曲線永不會趨向無窮大。
- 當 $x\to\infty$ 時，$y\to 3$ 從下方趨近，形如 $3-\frac{9}{x}$。這個趨近過程很慢：$y(10)\approx 2.31$。

由於曲線起初向上凹，其後漸趨平緩並靠近 $y=3$，它必定有一個拐點（point of inflection）。



z 的曲線

只要 $z \neq 0$，我們就可以除以 z^3。由於 $\dfrac{\mathrm{d}}{\mathrm{d}x}\left(z^{-2}\right) = -2z^{-3}\dfrac{\mathrm{d}z}{\mathrm{d}x}$，

$$\frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{1}{z^2}\right) = -\frac{2}{z^3}\cdot\frac{x^3z^3}{2(1+x^2)^{5/2}} = -\frac{x^3}{(1+x^2)^{5/2}},$$

這正是 $1/y$ 所滿足的方程。在兩條曲線上，$1/z^2 = 1$ 於 $x = 0$ 處成立，與 $1/y$ 的值相同。所以

$$\frac{1}{z^2} = \frac{1}{y}, \qquad \text{即} \qquad z^2 = y.$$

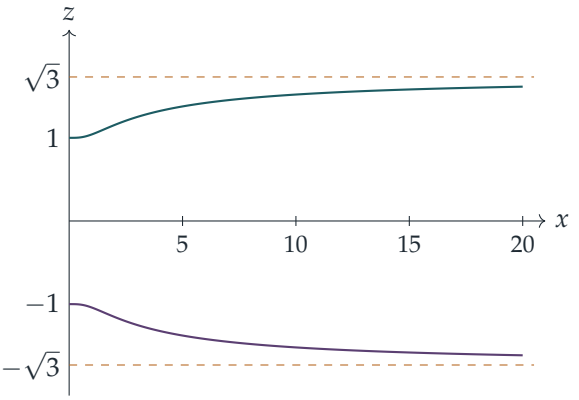
這同時確認了 z 永不會到達 0，從而說明剛才的除法是合理的。只要 $z \neq 0$，便有 $z^2 = y \geq 1$，所以 z 不可能接近 0。根據連續性，z 會一直保持起始時的正負號，於是兩條曲線分別是

$$z = \sqrt{y} \quad (z(0) = 1), \qquad z = -\sqrt{y} \quad (z(0) = -1).$$

上方的曲線由 1 上升並趨向 $\sqrt{3}$，同樣是起初很平坦（這次是 $z \approx 1 + \frac{1}{8}x^4$），而當 x 很大時

$$z = \sqrt{3}\left(1 - \frac{3}{x} + \cdots\right)^{1/2} \approx \sqrt{3} - \frac{3\sqrt{3}}{2x}.$$

下方的曲線是它對 x 軸的反射：它由 -1 下降，趨向漸近線（asymptote）$z = -\sqrt{3}$。



想一想 為甚麼 $z(0) = -1$ 的那條曲線，不只是把另一條曲線平移或作其他變換而得的複製品？

這條方程對 z 而言是奇的：把 z 換成 $-z$，兩邊都會乘上因子 -1。所以若 $z(x)$ 是解，則 $-z(x)$ 也是解，而由 -1 出發的解就是由 $+1$ 出發的解的反射。這與 $z = \pm\sqrt{y}$ 吻合。

驗算

- 對答案求導：由 $1/y = u^{-1/2} - \frac{1}{3}u^{-3/2} + \frac{1}{3}$，可得 $\mathrm{d}(1/y)/\mathrm{d}x = 2x\left(-\frac{1}{2}u^{-3/2} + \frac{1}{2}u^{-5/2}\right) = -x^3u^{-5/2}$，正合所需。
- 以數值方法（龍格－庫塔法，Runge–Kutta）解原方程，得 $y(1) = 1.08391$ 和 $y(5) = 1.89777$，與公式吻合至 10 位或以上有效數字。以數值方法解 z 的方程，得 $z(5) = \pm 1.37759 = \pm\sqrt{y(5)}$。
- 在 $x = 100$ 處，公式給出 $y = 2.9126$，而 $3 - \frac{9}{x} = 2.91$。兩者相差約 $27/x^2$，與 $O(1/x^2)$ 的誤差一致。

常見錯誤

- 漏掉負號：$\int y^{-2}\,\mathrm{d}y = -y^{-1}$。
- 斷定 $y \to 3$ 的速度很快，理由是 $3 - \frac{9}{x}$「看來很接近」3。其實在 $x = 10$ 處，曲線仍只到達約 2.3。
- 兩條曲線都寫成 $z = \sqrt{y}$，或者沒有解釋為甚麼 z 不能變號。

STEP 之外

兩條方程都具有 $\mathrm{d}y/\mathrm{d}x = f(x)\,y^n$ 的形式，這是 *伯努利方程*（*Bernoulli equation*）的特殊情況；代換 $w = y^{1-n}$ 總能把這類方程化為線性方程。我們用的正是這個代換：$w = 1/y$ 和 $w = 1/z^2$。「方程相同，初值相同，所以函數相同」這一步，是一階方程的存在唯一性定理（existence and uniqueness theorem，即 Picard–Lindelöf 定理）的一個例子。Boyce 與 DiPrima 合著的 *Elementary Differential Equations*（初等微分方程教科書）對兩者都有論述。

詳細解答

STEP II 2005，第 9 題

力學 • 斜面上的摩擦力

> **STEP II 2005，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/05-s2-q9

關鍵思路

把兩個質點（particle）分開處理。A 沒有受到外加的力，所以由它的方程便可確定繩子必須提供的張力（tension）。所有可供選擇的餘地，都在於 P 作用於 B 的方向。把 P 傾斜會改變 B 所受的法向反作用力（normal reaction），從而改變 B 所受的摩擦力（friction）。令 P 偏離斜面，會浪費它沿斜面方向的部分分量，卻同時改變了摩擦力；最佳的傾斜角度正好令這兩種效應取得平衡。最後問題歸結為求形如 $a\cos\theta + b\sin\theta$ 的式子的最大值。

解答

我們把 A 和 B 視為質點，並把繩子視為輕而不可伸長（light and inextensible），且沿 ℓ 放置。我們取 P 作用於包含 ℓ 的鉛垂平面內。P 橫過斜面的分量毫無幫助。注意

$$\sin 30^\circ = \tfrac{1}{2}, \qquad \cos 30^\circ = \tfrac{\sqrt{3}}{2}, \qquad \tan 30^\circ = \tfrac{1}{\sqrt{3}} = \tfrac{1}{3}\sqrt{3}.$$

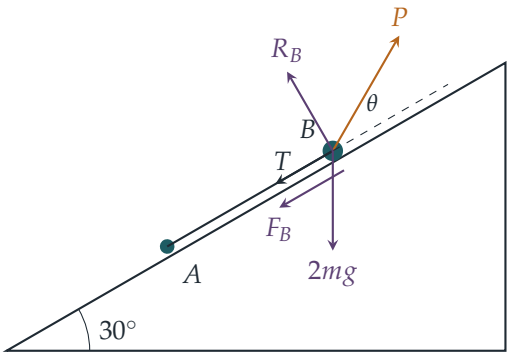
所以 $\mu_B = \tan 30^\circ$：若只有 B 本身，它恰好處於即將滑動的狀態。另外 $\mu_A = \frac{1}{2}\tan 30^\circ$，所以若只有 A 本身，它會滑下。

質點 A。沿垂直於斜面的方向分解，得 $R_A = mg\cos 30^\circ = \frac{\sqrt{3}}{2}mg$，所以 A 所受的最大摩擦力為

$$\mu_A R_A = \tfrac{\sqrt{3}}{6}\cdot\tfrac{\sqrt{3}}{2}mg = \tfrac{1}{4}mg.$$

重量沿斜面向下的分量為 $mg\sin 30^\circ = \frac{1}{2}mg$。

質點 B。設 P 與 ℓ 向上的方向成角 θ，以離開斜面的一方為正。（θ 為負表示 P 把質點壓向斜面。）設 T 為張力。



圖中所示的摩擦力 F_B 沿斜面向下，即第 (i) 部分的情況。沿垂直於斜面的方向，

$$R_B + P\sin\theta = 2mg\cos 30^\circ \quad \Longrightarrow \quad R_B = \sqrt{3}\,mg - P\sin\theta.$$

所以 B 處可提供的最大摩擦力為

$$\mu_B R_B = \tfrac{1}{\sqrt{3}}\left(\sqrt{3}\,mg - P\sin\theta\right) = mg - \tfrac{1}{\sqrt{3}}P\sin\theta.$$

B 的重量沿斜面向下的分量為 $2mg\sin 30^\circ = mg$。

第 (i) 部分

要令兩個質點沿 ℓ 向上移動，每個質點所受的摩擦力都沿斜面向下。當 P 取最小值時，兩者都處於即將移動的狀態，所以兩者所受的摩擦力都是極限摩擦力(limiting friction)。

要令 A 向上移動，繩子必須克服重量的分量和摩擦力：

$$T = \tfrac{1}{2}mg + \tfrac{1}{4}mg = \tfrac{3}{4}mg.$$

對於 B，沿斜面方向，P 的分量必須克服 T、重量的分量和極限摩擦力：

$$P\cos\theta = \tfrac{3}{4}mg + mg + \left(mg - \tfrac{1}{\sqrt{3}}P\sin\theta\right),$$

$$P\left(\cos\theta + \tfrac{1}{\sqrt{3}}\sin\theta\right) = \tfrac{11}{4}mg.$$

想一想 把 P 傾斜離開斜面，明明會減少 P 沿斜面的分量，為甚麼反而有幫助？

把 B 稍微向上提，會減少 R_B，從而減少 P 需要克服的摩擦力。當傾斜角 θ 很小時，$\cos\theta$ 只按 $1 - \frac{1}{2}\theta^2$ 的方式下降，而 $\sin\theta$ 則按 θ 的方式增長。所以起初是摩擦力減少所帶來的好處佔優。要求出最佳角度，可寫成

$$\cos\theta + \tfrac{1}{\sqrt{3}}\sin\theta = \tfrac{2}{\sqrt{3}}\left(\tfrac{\sqrt{3}}{2}\cos\theta + \tfrac{1}{2}\sin\theta\right) = \tfrac{2}{\sqrt{3}}\cos(\theta - 30^\circ).$$

此式最大為 $\frac{2}{\sqrt{3}}$，而等號恰好在 $\theta = 30^\circ$ 時成立。由於 P 乘以這個因子必須等於 $\frac{11}{4}mg$，所以當這個因子最大時 P 最小：

$$P_{\min} = \tfrac{11}{4}mg \cdot \tfrac{\sqrt{3}}{2} = \tfrac{11}{8}\sqrt{3}\,mg,$$

這正是所要求的結果。我們必須檢查 B 仍然留在斜面上：

$$R_B = \sqrt{3}\,mg - \tfrac{11}{8}\sqrt{3}\,mg \cdot \tfrac{1}{2} = \tfrac{5}{16}\sqrt{3}\,mg > 0.$$

所以 P 作用於通過 ℓ 的鉛垂平面內，與 ℓ 成 30°，並偏向離開斜面的一方。換言之，它**與水平成** 60° **向上**，把質點沿斜面向上拉。（留意與斜面所成的最佳角度是 $\arctan \mu_B$，即 B 的摩擦角（angle of friction）。）

第 (ii) 部分

現在 P 只需阻止系統向下滑，所以每個質點所受的摩擦力都沿斜面向上。若沒有 P，A 不能獨自保持靜止，因為它的重量分量 $\frac{1}{2}mg$ 大於最大摩擦力 $\frac{1}{4}mg$。所以繩子必須拉住 A：

$$T \geq \tfrac{1}{2}mg - \tfrac{1}{4}mg = \tfrac{1}{4}mg.$$

特別地，繩子必須拉緊。要令 B 不向下滑，沿斜面向上的力必須足以平衡 T 和重量的分量：

$$P\cos\theta + \left(mg - \tfrac{1}{\sqrt{3}}P\sin\theta\right) \geq T + mg.$$

左邊與 T 無關，所以最小的 P 對應最小的張力 $T = \frac{1}{4}mg$。於是

$$P\left(\cos\theta - \tfrac{1}{\sqrt{3}}\sin\theta\right) \geq \tfrac{1}{4}mg.$$

而 $\cos\theta - \frac{1}{\sqrt{3}}\sin\theta = \frac{2}{\sqrt{3}}\cos(\theta + 30^\circ) \leq \frac{2}{\sqrt{3}}$，等號在 $\theta = -30^\circ$ 時成立。這一次，最佳方向是把 B 壓向斜面，這樣會增大 R_B，從而增大有助穩住 B 的摩擦力。因此力的最小量值為

$$P_{\min} = \tfrac{1}{4}mg \cdot \tfrac{\sqrt{3}}{2} = \tfrac{1}{8}\sqrt{3}\,mg.$$

它的方向在 ℓ 向上方向之下 30°。由於 ℓ 以 30° 向上傾斜，這個方向正是水平方向：即沿水平方向把 B 推向斜面。取這個值時，A 和 B 都處於極限平衡（limiting equilibrium），而 $T = \frac{1}{4}mg$。

驗算

- 我們取 $m = g = 1$ 作數值檢驗。在鉛垂平面內掃描 P 的每一個方向，並以二分法（bisection）求出每個方向所需的最小 P。在 (i) 中，最小值為 $2.3816 \approx \frac{11}{8}\sqrt{3}$，出現於與斜面成 30.0° 之處。在 (ii) 中，最小值為 $0.2165 \approx \frac{1}{8}\sqrt{3}$，出現於 -30.0°，即水平方向。我們亦確認了在 (ii) 的最小值處確實存在完整的平衡狀態。

- 就 (i) 而言，當 P 沿斜面方向（$\theta = 0$）時，把各項加起來：所需的 $P = \frac{1}{2}mg + \frac{1}{4}mg + mg + mg = \frac{11}{4}mg \approx 2.75mg$。這比 $2.38mg$ 大，正如所料。
- 就 (ii) 而言，當 P 沿斜面方向時，系統需要 $\frac{3}{2}mg - \frac{1}{4}mg - mg = \frac{1}{4}mg$。水平推力的效果更佳，只需 $\frac{\sqrt{3}}{8}mg \approx 0.217mg$。

常見錯誤

- 假設 P 沿斜面方向作用。題目問的是最小的量值，所以方向可由你自行選擇。
- 忘記 $P\sin\theta$ 會改變 R_B，因而改變 B 所受的極限摩擦力。
- 在 (ii) 中沿用 (i) 的摩擦力方向。當要防止的是向下滑動時，摩擦力沿斜面向上。

STEP 之外

兩部分都建基於同一原理：對抗摩擦力時最有效率的力，與運動方向所成的角正是摩擦角 $\lambda = \arctan\mu$。在 (i) 中它偏離斜面，在 (ii) 中則偏向斜面。原因是法向反作用力與極限摩擦力可合成為單一個「總反作用力」（total reaction），它與法線成角 λ。以這個總反作用力來思考，許多摩擦力問題都可化為力三角形（triangle of forces）的幾何問題。Kleppner 與 Kolenkow 所著的 *An Introduction to Mechanics*（力學入門）以大學一年級的程度，仔細地處理了摩擦力和受約束運動。

詳細解答

STEP II 2005，第 10 題

力學 • 相撞的拋體

STEP II 2005，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/05-s2-q10

關鍵思路

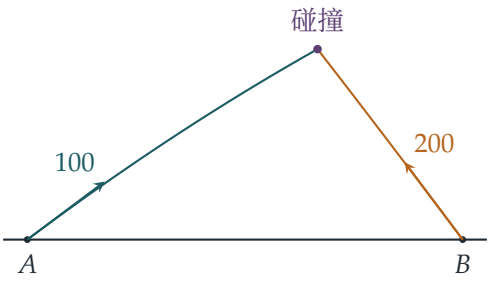
兩枚導彈都是拋體（projectile），所以我們可以寫出每一枚在任何時刻的位置。碰撞（collision）的意思是兩者在同一瞬間位置相同，由此得到兩條方程（水平和垂直方向各一），含兩個未知數：碰撞時間和 T。水平方向的方程是線性的，可以用來消去碰撞時間，垂直方向的方程隨之變成 T 的二次方程（quadratic）。之後我們要判斷哪一個根在物理上合理。

解答

建立模型

我們把兩枚導彈都視為質點（particle），只受重力作用而自由運動，不計空氣阻力。取 A 為原點，x 軸沿 AB 方向（所以 B 位於 $x = 180$），y 軸垂直向上。時間 t 以秒為單位，由第一枚導彈發射的一刻起計。

對第一枚導彈，由 $\sin\alpha = \frac{3}{5}$ 得 $\cos\alpha = \frac{4}{5}$，所以它的初速度分量是 $100 \cdot \frac{4}{5} = 80$ 和 $100 \cdot \frac{3}{5} = 60$。對第二枚導彈，由 $\sin\beta = \frac{4}{5}$ 得 $\cos\beta = \frac{3}{5}$。它朝着 A 飛行，所以它的分量是 $-200 \cdot \frac{3}{5} = -120$ 和 $200 \cdot \frac{4}{5} = 160$。



設 $s = t - T$ 為第二枚導彈發射後所經過的時間。那麼，當兩者都在飛行時，

$$\begin{aligned} &\text{第一枚導彈：} \quad x = 80t, && y = 60t - 5t^2, \\ &\text{第二枚導彈：} \quad x = 180 - 120s, && y = 160s - 5s^2. \end{aligned}$$

碰撞

碰撞時兩者的 x 坐標相同：

$$80t = 180 - 120(t - T) \quad \Longrightarrow \quad 200t = 180 + 120T \quad \Longrightarrow \quad t = 0.9 + 0.6T,$$

因此 $s = t - T = 0.9 - 0.4T$。

y 坐標亦相同：$60t - 5t^2 = 160s - 5s^2$，即

$$60t - 160s = 5(t^2 - s^2) = 5(t - s)(t + s) = 5T(t + s),$$

這裏用了 $t - s = T$。代入 t 和 s 的表達式：

$$\begin{aligned} 60t - 160s &= 60(0.9 + 0.6T) - 160(0.9 - 0.4T) = 100T - 90, \\ 5T(t + s) &= 5T(1.8 + 0.2T) = 9T + T^2. \end{aligned}$$

令兩式相等，

$$T^2 - 91T + 90 = 0 \quad \Longrightarrow \quad (T - 1)(T - 90) = 0,$$

所以 $T = 1$ 或 $T = 90$。

? **想一想** 兩個根都滿足方程。為甚麼 $T = 90$ 不可能是答案？

當 $T = 90$ 時，我們得到 $s = 0.9 - 36 = -35.1$，即是說這次「碰撞」會在第二枚導彈發射之前 35.1 秒發生。（碰撞點亦遠低於地面：這些方程所描述的，是延伸至實際飛行範圍以外的拋物線。）所以這個根沒有物理意義。況且第一枚導彈只在空中停留 $2 \cdot 60/10 = 12$ 秒，早就已經落地了。

當 $T = 1$ 時，我們得到 $t = 1.5$ 和 $s = 0.5 > 0$，所以兩枚導彈都在飛行中。高度分別是 $60(1.5) - 5(1.5)^2 = 78.75$ 和 $160(0.5) - 5(0.5)^2 = 78.75$，兩者皆為正數，而且兩枚導彈

都位於 $x = 120$。因此

$$T = 1 \text{ 秒}.$$

驗算

把 $T = 1$ 代回去：在 $t = 1.5\,\text{s}$ 時，兩枚導彈都位於 $(120, 78.75)$ 米，這一點我們已用精確分數核實。以發射延遲 $1\,\text{s}$ 用數值方法逐步計算兩者的運動，兩枚導彈之間的距離在 $t = 1.5\,\text{s}$ 時降至幾乎為零。再作一個合理性檢查：碰撞發生在 $x = 120$，介乎 A 與 B 之間，而第一枚導彈仍在上升，因為它在該處的垂直速度是 $60 - 15 = 45\,\text{m}\,\text{s}^{-1}$。

常見錯誤

- 兩枚導彈用了同一個時間變量。第二枚導彈的飛行時間是 $t - T$，而不是 t。
- 忘記第二枚導彈是沿 x 軸的負方向運動。
- 保留 $T = 90$，而沒有檢查是否 $t - T > 0$。

STEP 之外

這個問題有一個更巧妙的看法。兩枚導彈的加速度同為 g，所以一旦兩者都在飛行，它們的相對速度（relative velocity）便是常數，各自看到對方沿直線運動。碰撞發生的充要條件，是這個相對速度的方向正好沿着時刻 T 時連接兩者的直線，把這個條件寫出來便得到同一條二次方程。像這樣轉換到自由下落的參考系（frame），正是通往愛因斯坦（Einstein）等效原理（equivalence principle）的第一步。Kleppner 與 Kolenkow 的 *An Introduction to Mechanics*（力學入門教科書）對相對運動和非慣性參考系（non-inertial frame）有嚴謹的論述。

詳細解答

STEP II 2005，第 11 題

力學 • 斜面上的摩擦力

STEP II 2005，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：

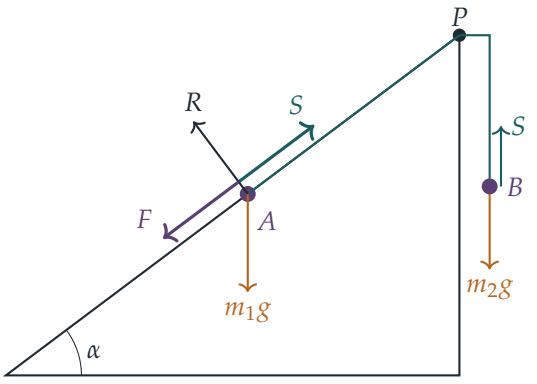
step.maths.org/questions/05-s2-q11

關鍵思路

由於 $\tan\alpha = \frac{3}{4}$，我們得到一個 3–4–5 三角形，所以 $\sin\alpha = \frac{3}{5}$ 及 $\cos\alpha = \frac{4}{5}$，而作用於 A 的每一個力都成為 $m_1 g$ 的簡潔倍數。A 的運動於是分成三個階段，每個階段的加速度（acceleration）都是常數：被繩子拉上斜面、自由地向上滑行，以及向下滑回。摩擦力（friction）的方向在最後兩個階段之間改變，而這正是本題暗藏玄機的地方。

解答

我們把 A 和 B 視為質點（particle），繩子視為輕而不可伸長（light and inextensible），滑輪（pulley）視為光滑，因此兩邊的張力（tension）S 相同。設 $\alpha = \arctan\frac{3}{4}$，使得 $\sin\alpha = \frac{3}{5}$ 及 $\cos\alpha = \frac{4}{5}$。



繩子拉緊時的加速度

在垂直於斜面的方向上，A 沒有加速度，所以

$$R = m_1 g\cos\alpha = \tfrac{4}{5}m_1 g, \qquad \text{而極限摩擦力（limiting friction）為} \qquad \mu R = \tfrac{1}{2}\cdot\tfrac{4}{5}m_1 g = \tfrac{2}{5}m_1 g.$$

首先我們檢查系統確實會運動。假如 B 下降，A 便會沿斜面向上移動，阻礙它的是其重量的分量 $\frac{3}{5}m_1 g$ 以及最多為 $\frac{2}{5}m_1 g$ 的摩擦力：合共最多為 $m_1 g$。向另一方拉動的是 B 的重量 $m_2 g$，而 $m_2 > m_1$，所以摩擦力無法令系統保持靜止：B 下降，A 沿斜面向上移動，而作用於 A 的摩擦力達到極限值，方向沿斜面向下。

設 a 為兩者共同的加速度。對 A（沿斜面向上）和對 B（向下）應用牛頓第二定律（Newton's second law），得

$$\begin{aligned} m_1 a &= S - \tfrac{3}{5}m_1 g - \tfrac{2}{5}m_1 g = S - m_1 g, \\ m_2 a &= m_2 g - S. \end{aligned}$$

兩式相加，$(m_1 + m_2)a = (m_2 - m_1)g$，所以

$$a = \frac{m_2 - m_1}{m_2 + m_1}\, g,$$

正如所求。為方便起見，我們寫成 $a = kg$，其中 $k = \dfrac{m_2 - m_1}{m_2 + m_1}$ 且 $0 < k < 1$。

到達最高點的時間

在時間 T，A 正以速率 $u = aT = kgT$ 沿斜面向上移動，並已走了 $\frac{1}{2}kgT^2$ 的距離。

繩子斷開時，S 消失。A 仍在沿斜面向上移動，所以摩擦力仍然沿斜面向下，而沿斜面向下的力合共為 $\frac{3}{5}m_1 g + \frac{2}{5}m_1 g = m_1 g$。因此 A 的減速度恰好是 g，並在再過一段時間

$$\frac{u}{g} = kT.$$

之後靜止下來。故此 A 在釋放後的時間

$$T + kT = (1+k)T = \frac{2m_2}{m_1 + m_2}\, T$$

到達最高點。（題目告訴我們 A 永遠不會到達滑輪，所以這個階段不會被提早中斷。）在這個階段中，A 再走了 $\dfrac{u^2}{2g} = \frac{1}{2}k^2 gT^2$，所以它由釋放點起沿斜面向上所走的總距離為

$$d = \tfrac{1}{2}kgT^2 + \tfrac{1}{2}k^2 gT^2 = \tfrac{1}{2}k(1+k)gT^2.$$

想一想 A 停下來之後，會留在原處嗎？

只有當摩擦力足以支撐它時，它才會留在原處。重量沿斜面向下的分量是 $\frac{3}{5}m_1g$，而摩擦力最多只能提供 $\frac{2}{5}m_1g$。換句話說，$\tan\alpha = \frac{3}{4} > \frac{1}{2} = \mu$。所以 A 會向下滑回，而這時摩擦力沿斜面向上。沿斜面向下的淨力為 $\frac{3}{5}m_1g - \frac{2}{5}m_1g = \frac{1}{5}m_1g$，所以 A 以 $\frac{1}{5}g$ 的加速度沿斜面向下加速，比它上行時的減速度緩和得多。

質量之比

A 由最高點從靜止出發，需時 t_d 滑回距離 d 而返抵釋放點，其中

$$\tfrac{1}{2}\cdot\tfrac{1}{5}g\,t_d^2 = d = \tfrac{1}{2}k(1+k)gT^2, \qquad \text{所以} \qquad t_d^2 = 5k(1+k)T^2.$$

題目告訴我們 t_d 等於到達最高點的時間 $(1+k)T$。兩邊平方，這表示

$$(1+k)^2T^2 = 5k(1+k)T^2.$$

兩邊除以正數 $(1+k)T^2$，得 $1+k=5k$，所以 $k=\frac{1}{4}$。留意 T 已經消去：這就是為何無論繩子何時斷開，這個等式都成立，而它亦唯一地確定了 k 的值。於是

$$\frac{m_2-m_1}{m_2+m_1} = \frac{1}{4} \iff 4m_2-4m_1 = m_2+m_1 \iff 3m_2 = 5m_1,$$

所以

$$m_1 : m_2 = 3 : 5.$$

驗算

- 極限情況：若 $m_2 \to m_1$，則 $k \to 0$，而到達最高點的時間趨向 T（幾乎沒有速率可以失去）；若 $m_2 \gg m_1$，則 $k \to 1$，而 A 靠慣性滑行的時間與它被拉動的時間一樣長，這正如我們所料，因為它達到速率 gT 後以 g 減速。
- 取 $m_1 = 3$、$m_2 = 5$ 及 $T = 1\,\text{s}$，我們有 $k = \frac{1}{4}$；到達最高點的時間是 $1.25\,\text{s}$，$d = \frac{5}{32}g$，而下滑需時 $\sqrt{2d/(g/5)} = \sqrt{25/16} = 1.25\,\text{s}$。對全部三個階段逐步進行數值模擬，在多個 T 值之下，兩個時間都得出 $1.250\,\text{s}$。

常見錯誤

- 忘記了 A 開始滑回時摩擦力會反向，因而以 g（或 $\frac{3}{5}g$）而非 $\frac{1}{5}g$ 作為下滑的加速度。
- 只由繩子斷開的位置起計算下滑的距離，而不是由釋放點起計算。
- 沒有檢查 A 是否真的會滑回：這需要 $\tan\alpha > \mu$。

STEP 之外

摩擦力是取決於運動狀態的力的經典例子：它的方向隨速度而轉換，而當物體靜止時，它的大小有一個上限。這使得運動成為分段的，每個階段各有不同的運動方程(equation of motion)，而同一個想法亦支配着黏滑（stick–slip）現象，例如吱吱作響的門鉸和地震斷層。Kleppner 與 Kolenkow 的 *An Introduction to Mechanics*（力學入門）以大學一年級物理所要求的嚴謹程度，處理摩擦力和受約束的運動。

詳細解答

STEP II 2005，第 12 題

概率與統計 • 貝葉斯定理

STEP II 2005，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/05-s2-q12

關鍵思路

這是以公平硬幣作為先驗（prior）的貝葉斯定理（Bayes' theorem）。由於硬幣是公平的，在已知回答的條件下，擲出正面的概率為 $\frac{1}{2}$，當且僅當無論硬幣擲出正面還是反面，得到這些回答的可能性都相等。因此在每一部分，我們只需比較兩個似然（likelihood）$\mathrm{P}(\text{回答} \mid H)$ 和 $\mathrm{P}(\text{回答} \mid T)$，並以哪一位孿生姊妹在線作為條件來計算它們。

解答

設 H 和 T 分別為硬幣擲出正面和反面的事件，故 $\mathrm{P}(H) = \mathrm{P}(T) = \frac{1}{2}$。設 A 為 Anna 在線的事件，B 為 Bella 在線的事件，其中 $\mathrm{P}(A) = p$，$\mathrm{P}(B) = q = 1 - p$。誰在線與硬幣毫無關係，所以 A 和 B 均與 H 獨立（independent）。在第 (ii) 和第 (iii) 部分，我們假設硬幣只擲一次，而兩封電郵都由同一位孿生姊妹回覆，即我們發信時在線的那一位。（另一種理解方式見文末的備註。）

記 $t = ap + bq$，即單獨一個回答屬實的整體概率。

對於任何一組回答紀錄 E，由貝葉斯定理及 $\mathrm{P}(H) = \mathrm{P}(T)$ 可得

$$\mathrm{P}(H \mid E) = \frac{\frac{1}{2}\mathrm{P}(E \mid H)}{\frac{1}{2}\mathrm{P}(E \mid H) + \frac{1}{2}\mathrm{P}(E \mid T)} = \frac{\mathrm{P}(E \mid H)}{\mathrm{P}(E \mid H) + \mathrm{P}(E \mid T)}.$$

此式等於 $\frac{1}{2}$，當且僅當 $\mathrm{P}(E \mid H) = \mathrm{P}(E \mid T)$。單憑這一點觀察，便已完成了大部分工作。

第 (i) 部分

設 Y 為回答是「是」的事件。若硬幣擲出正面，「是」便是真話。以誰在線作為條件，

$$\mathrm{P}(Y \mid H) = pa + qb = t.$$

若硬幣擲出反面，「是」便是謊話，故

$$\mathrm{P}(Y \mid T) = p(1-a) + q(1-b) = (p+q) - (ap+bq) = 1 - t.$$

因此

$$\mathrm{P}(H \mid Y) = \frac{t}{t + (1-t)} = t = ap + bq.$$

所以 $\mathrm{P}(H \mid Y) = \frac{1}{2}$ 當且僅當 $ap + bq = \frac{1}{2}$，亦即當且僅當 $2(ap+bq) = 1$，正如所求。

第 (ii) 部分

設 E 為先答「是」、後答「否」的事件。假設 Anna 在線。她的兩個回答互相獨立，而且無論硬幣擲出哪一面，兩個回答中都恰好有一個是真話：

$$\begin{aligned} \mathrm{P}(E \mid H \cap A) &= a(1-a) && \text{（如實答「是」，然後說謊），} \\ \mathrm{P}(E \mid T \cap A) &= (1-a)a && \text{（說謊，然後如實答「否」）．} \end{aligned}$$

兩者相等。同理，$\mathrm{P}(E \mid H \cap B) = b(1-b) = \mathrm{P}(E \mid T \cap B)$。因此

$$\mathrm{P}(E \mid H) = pa(1-a) + qb(1-b) = \mathrm{P}(E \mid T),$$

再由上面的觀察可知 $\mathrm{P}(H \mid E) = \frac{1}{2}$，不論 p、a 和 b 取何值。

> ? **想一想** 為甚麼這裏完全不需要任何條件，就恰好得到 $\frac{1}{2}$？

回答互相矛盾，意味着無論硬幣擲出哪一面，這位孿生姊妹都是說了一次真話、說了一次謊話。這個組合在正面和反面的情況下出現的可能性相等，所以它不帶有任何關於硬幣的資訊，我們得到的仍然是先驗概率 $\frac{1}{2}$。

第 (iii) 部分

現在設 E 為兩次都答「是」的事件。已知擲出正面時，兩個回答都是真話；已知擲出反面時，兩個回答都是謊話。以誰在線作為條件，

$$\begin{aligned} \mathrm{P}(E \mid H) &= pa^2 + qb^2, \\ \mathrm{P}(E \mid T) &= p(1-a)^2 + q(1-b)^2. \end{aligned}$$

兩式相減，並利用 $(1-a)^2 = 1 - 2a + a^2$，

$$\begin{aligned} \mathrm{P}(E \mid H) - \mathrm{P}(E \mid T) &= p\big(a^2 - (1-a)^2\big) + q\big(b^2 - (1-b)^2\big) \\ &= p(2a-1) + q(2b-1) \\ &= 2(ap+bq) - (p+q) \\ &= 2(ap+bq) - 1. \end{aligned}$$

所以若 $2(ap+bq) = 1$，兩個似然便相等，而 $\mathrm{P}(H \mid E) = \frac{1}{2}$，正如所求。

（事實上，這個計算所顯示的不止於此：兩個似然之差恰好是 $2(ap+bq) - 1$，所以 $\mathrm{P}(H \mid E) = \frac{1}{2}$ 成立當且僅當 $2(ap+bq) = 1$，與第 (i) 部分的條件相同。）

備註：另一種理解方式

如果改為每封電郵都重新、獨立地決定是哪一位孿生姊妹剛好在線，那麼每個回答屬實的概率都是 t，而且互相獨立。這樣，在 (ii) 中 $\mathrm{P}(E \mid H) = t(1-t) = \mathrm{P}(E \mid T)$，而在 (iii) 中 $\mathrm{P}(E \mid H) = t^2$ 及 $\mathrm{P}(E \mid T) = (1-t)^2$，兩者在 $t = \frac{1}{2}$ 時相等。結論並無分別。

驗算

- 若兩位孿生姊妹永遠說真話（$a = b = 1$），則 $t = 1 \neq \frac{1}{2}$，而在 (i) 中，一個「是」的回答令正面成為必然：$\mathrm{P}(H \mid Y) = t = 1$。這時 (ii) 中互相矛盾的回答不可能出現，這與 $a(1-a) = b(1-b) = 0$ 一致。
- 取 $p = 0.3$、$a = 0.8$ 及 $b = \frac{0.5 - 0.24}{0.7} \approx 0.371$（故 $2(ap+bq) = 1$），在同一位孿生姊妹回覆的模型中模擬 2×10^5 輪，得出 $\mathrm{P}(H \mid \text{是、是}) \approx 0.496$。以分數對數千組隨機的 (p, a, b) 作精確枚舉，確認了全部三個結果在兩種理解方式下均成立。
- 若 $p = \frac{1}{2}$ 及 $a = b = 0.9$，條件不成立，而兩個「是」的回答給出 $\mathrm{P}(H \mid E) = \frac{81}{82}$，與 $\frac{1}{2}$ 相去甚遠，可見 (iii) 中的條件確實重要。

常見錯誤

- 未加說明便寫下 $\mathrm{P}(Y \mid T) = 1 - \mathrm{P}(Y \mid H)$。此式在這裏碰巧成立，但那只是因為在反面的情況下答「是」屬於說謊；作答時應說明這一點。
- 在 (iii) 中，把兩個回答當作各自以概率 t 屬實的獨立回答，同時卻又假設兩個回答出自同一位孿生姊妹。在已知是哪一位孿生姊妹的條件下，兩個回答是獨立的，但在無條件的情況下並不獨立，因為兩者都取決於誰在線。

STEP 之外

這裏貫穿始終的原理只有一個：後驗（posterior）賠率（odds）等於先驗賠率乘以似然比（likelihood ratio）；在兩個假設下出現的可能性相等的證據，不會告訴你任何事情。同一位孿生姊妹回覆的模型是一個簡單的混合模型（*mixture*）：在已知某個隱變量（hidden variable，即誰在線）的條件下，各個回答互相獨立，但整體而言卻是相依的，這個想法正是隱變量模型和貝葉斯模型的核心。Blitzstein 與 Hwang 合著的 *Introduction to Probability*（概率論入門）正是以這種精神，對條件概率和貝葉斯法則作了十分出色的闡述。

詳細解答

STEP II 2005，第 13 題

概率與統計 • 泊松分佈與二項分佈

STEP II 2005，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/05-s2-q13

關鍵思路

每一頁抽樣頁要麼錯誤少於兩個，要麼不是，而且各頁互相獨立，所以 Y 數算的是 n 次獨立試驗（independent trials）中成功的次數：它服從二項分佈（binomial distribution）。其後的一切都在探討 λ 很小時會發生甚麼事，而所需的工具就是 $\mathrm{e}^{-\lambda}$（或 $\ln(1+\lambda)$）的麥克勞林級數（Maclaurin series），只保留第一個沒有被抵消的項。

解答

Y 的分佈

設 X 為某一頁上的錯誤數目，則 $X \sim \mathrm{Po}(\lambda)$。若 $X \leq 1$，便稱該頁為良好頁。於是

$$\mathrm{P}(\text{良好}) = \mathrm{P}(X=0) + \mathrm{P}(X=1) = \mathrm{e}^{-\lambda} + \lambda\mathrm{e}^{-\lambda} = (1+\lambda)\mathrm{e}^{-\lambda} = p.$$

樣本由 n 頁不同的頁組成。不同頁上的錯誤數目互相獨立，而且全都服從同一個 $\mathrm{Po}(\lambda)$ 分佈，所以無論選中哪 n 頁，「第 j 頁是良好頁」這些事件都互相獨立，各自的概率均為 p。（條件 $n \ll N$ 只是讓我們放心，有足夠多的頁可供抽樣；獨立性並不需要這個條件，它來自模型本身。）

所以 Y 是 n 次獨立試驗中成功的次數，每次成功的概率為 p。某一組指定的 k 頁良好頁和 $n-k$ 頁不良頁出現的概率是 $p^k q^{n-k}$，而選出哪 k 頁為良好頁共有 $\binom{n}{k}$ 種方法，所以

$$\mathrm{P}(Y=k) = \binom{n}{k} p^k q^{n-k}, \qquad k = 0, 1, \ldots, n,$$

其中 $p = (1+\lambda)\mathrm{e}^{-\lambda}$ 及 $q = 1-p$，正如題目所求。換言之，$Y \sim \mathrm{B}(n, p)$。

第 (i) 部分

展開 $\mathrm{e}^{-\lambda} = 1 - \lambda + \frac{1}{2}\lambda^2 - \frac{1}{6}\lambda^3 + \cdots$，再乘以 $1+\lambda$：

$$\begin{aligned} p = (1+\lambda)\mathrm{e}^{-\lambda} &= \left(1 - \lambda + \tfrac{1}{2}\lambda^2 - \tfrac{1}{6}\lambda^3 + \cdots\right) + \left(\lambda - \lambda^2 + \tfrac{1}{2}\lambda^3 - \cdots\right) \\ &= 1 - \tfrac{1}{2}\lambda^2 + \tfrac{1}{3}\lambda^3 - \cdots . \end{aligned}$$

λ 的一次項互相抵消，所以當 λ 很小時，

$$q = 1 - p = \tfrac{1}{2}\lambda^2 - \tfrac{1}{3}\lambda^3 + \cdots \approx \tfrac{1}{2}\lambda^2$$

因為被略去的各項比 $\frac{1}{2}\lambda^2$ 還要小，相差一個 λ 階的因子。

> **想一想** 為甚麼不是 $q \approx \lambda$，像 $\mathrm{e}^{-\lambda}$ 的首項所暗示的那樣？

因為來自 $\mathrm{e}^{-\lambda}$ 的 $-\lambda$ 剛好被乘以 $1+\lambda$ 所產生的 $+\lambda$ 完全抵消。一頁只有在至少有兩個錯誤時才是不良頁，而 $\mathrm{P}(X \geq 2) \approx \mathrm{P}(X = 2) = \frac{1}{2}\lambda^2\mathrm{e}^{-\lambda} \approx \frac{1}{2}\lambda^2$。這正好從另一個角度獨立地印證了 (i)。

第 (ii) 部分

$\mathrm{P}(Y = n) = p^n$，而我們要找滿足 $p^n \geq 1 - \lambda$ 的最大 n。兩邊都是正數(對於 $0 < \lambda < 1$)，而 ln 是遞增函數，所以這等價於

$$n \ln p \geq \ln(1 - \lambda).$$

現在 $\ln p = \ln(1+\lambda) - \lambda = \left(\lambda - \frac{1}{2}\lambda^2 + \cdots\right) - \lambda \approx -\frac{1}{2}\lambda^2$，而 $\ln(1-\lambda) \approx -\lambda$。由於 $\ln p < 0$，除以它時不等式（inequality）的方向會倒轉：

$$n \leq \frac{\ln(1-\lambda)}{\ln p} \approx \frac{-\lambda}{-\frac{1}{2}\lambda^2} = \frac{2}{\lambda}.$$

所以滿足條件的最大 n 約為 $2/\lambda$。

等價地，利用 (i)：$p^n = (1-q)^n \approx 1 - nq \approx 1 - \frac{1}{2}n\lambda^2$，而 $1 - \frac{1}{2}n\lambda^2 \geq 1 - \lambda$ 成立，當且僅當 $n \leq 2/\lambda$。$(1-q)^n \approx 1 - nq$ 這一步需要 nq 很小；這裏 $nq \approx \frac{2}{\lambda} \cdot \frac{1}{2}\lambda^2 = \lambda$，確實很小，所以這個近似是自洽的。

第 (iii) 部分

由於 $Y > 1$ 蘊含 $Y > 0$，

$$\mathrm{P}(Y > 1 \mid Y > 0) = \frac{\mathrm{P}(Y > 1)}{\mathrm{P}(Y > 0)} = \frac{\mathrm{P}(Y > 0) - \mathrm{P}(Y = 1)}{\mathrm{P}(Y > 0)} = 1 - \frac{npq^{n-1}}{1 - q^n}.$$

（這裏 $n \geq 2$ 是有作用的：它令 $Y > 1$ 成為可能發生的事件，並令冪 q^{n-1} 的次數至少與 q 相同。）現在固定 n，並設 λ 很小。那麼 $q \approx \frac{1}{2}\lambda^2$ 很小，所以 $p \approx 1$ 及 $1 - q^n \approx 1$，由此得

$$\mathrm{P}(Y > 1 \mid Y > 0) \approx 1 - nq^{n-1} \approx 1 - n\left(\frac{\lambda^2}{2}\right)^{n-1},$$

最後一步用了 (i)。這就是所求的結果。

驗算

- 以 $\lambda = 0.8$ 模擬 200,000 個樣本，每個樣本有 $n = 5$ 頁泊松（Poisson）頁，所得 $Y = 0, \ldots, 5$ 的相對頻數與 $\mathrm{B}(5, p)$ 相符，誤差在 0.002 以內。
- 當 $\lambda = 0.01$ 時，$q = 4.967 \times 10^{-5}$，而 $\frac{1}{2}\lambda^2 = 5 \times 10^{-5}$。
- 當 $\lambda = 0.01$ 時，滿足 $p^n \geq 0.99$ 的最大 n 的精確值是 202，而 $2/\lambda = 200$。（在每個級數中多保留一項，會得到 $2/\lambda + \frac{7}{3}$，所以當 $2/\lambda$ 增大時，誤差仍然有界。）
- 當 $\lambda = 0.01$，$n = 3$ 時：精確值為 $1 - \mathrm{P}(Y > 1 \mid Y > 0) = 7.40 \times 10^{-9}$，而 $3(\lambda^2/2)^2 = 7.5 \times 10^{-9}$。

常見錯誤

- 把「少於兩個錯誤」理解為只有 $X = 0$，或理解為 $X \leq 2$。
- 把 $(1 + \lambda)\mathrm{e}^{-\lambda}$ 的展開式停在一次項，然後斷定 $q \approx 0$：第一個沒有被抵消的項其實是二次項。
- 在 (ii) 中除以 $\ln p < 0$ 時，忘記把不等式的方向倒轉。

STEP 之外

第 (ii) 部分是漸近分析（asymptotic analysis）的一個例子：條件由乘積 $n\lambda^2$ 所控制，所以 n 的自然尺度是 $1/\lambda$，而像 p^n 這樣的量，其表現就如 $\mathrm{e}^{-n\lambda^2/2}$。二項分佈的泊松近似，以及概率論中許多「稀有事件」的極限，背後都是同一套推理。Grimmett 與 Welsh 所著的 *Probability: An Introduction*（概率論入門教材）對這些極限定理有嚴謹細緻的闡述。

詳細解答

STEP II 2005，第 14 題

概率與統計 • 混合分佈

> **STEP II 2005，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/05-s2-q14

關鍵思路

這個密度是兩個我們早已熟悉的密度的加權混合：一個是標準正態分佈（standard normal distribution），另一個是 $[0,\lambda]$ 上的均勻分佈（uniform distribution）。因此，f 的每個積分都可拆成正態部分和均勻部分，而每一部分都可以直接由標準事實（總概率為 1、均值、$\mathrm{E}(X^2)$）讀出，完全不必做任何艱深的積分。方差（variance）則由 $\mathrm{E}(X^2)-\mu^2$ 求得。

解答

常數 k、均值與方差

由於 ϕ 和 g 都是概率密度函數（probability density function），兩者在整條實數線上的積分都是 1。所以

$$1=\int_{-\infty}^{\infty} f(x)\,\mathrm{d}x = k\,[1+\lambda\cdot 1] \quad\Longrightarrow\quad k=\frac{1}{1+\lambda}.$$

至於均值（mean），我們利用兩個成分各自的均值：標準正態分佈的均值為 0，而 $[0,\lambda]$ 上的均勻密度 g 的均值為 $\frac{1}{2}\lambda$。因此

$$\mu=\int_{-\infty}^{\infty} xf(x)\,\mathrm{d}x = k\left[\int x\phi(x)\,\mathrm{d}x+\lambda\int xg(x)\,\mathrm{d}x\right]=\frac{1}{1+\lambda}\left[0+\lambda\cdot\frac{\lambda}{2}\right]=\frac{\lambda^2}{2(1+\lambda)}.$$

同樣地，標準正態分佈滿足 $\int x^2\phi(x)\,\mathrm{d}x=1$（由於其均值為 0，這就是它的方差），而

$$\int x^2 g(x)\,\mathrm{d}x=\frac{1}{\lambda}\int_0^{\lambda} x^2\,\mathrm{d}x=\frac{\lambda^2}{3}.$$

所以

$$\mathrm{E}(X^2) = \frac{1}{1+\lambda}\left[1 + \lambda \cdot \frac{\lambda^2}{3}\right] = \frac{3+\lambda^3}{3(1+\lambda)}.$$

想一想 為甚麼 $\mathrm{E}(X^2)$ 可以這樣拆開，方差卻不可以？

期望值（expectation）是積分，而積分是線性的，所以 $\mathrm{E}(X^2)$ 是各成分二階矩（second moment）的加權組合，權重與 f 由各成分密度組合而成時所用的完全相同。方差並非線性（它涉及 μ^2），所以我們必須先把二階矩組合起來，然後才減去 μ^2。這樣做便得

$$\begin{aligned}\sigma^2 = \mathrm{E}(X^2) - \mu^2 &= \frac{3+\lambda^3}{3(1+\lambda)} - \frac{\lambda^4}{4(1+\lambda)^2} \\ &= \frac{4(3+\lambda^3)(1+\lambda) - 3\lambda^4}{12(1+\lambda)^2} \\ &= \frac{12+12\lambda+4\lambda^3+4\lambda^4-3\lambda^4}{12(1+\lambda)^2} = \frac{\lambda^4+4\lambda^3+12\lambda+12}{12(1+\lambda)^2},\end{aligned}$$

這正是所要求的結果。

$\lambda = 2$ 的情況

此時 $k = \frac{1}{3}$，而在 $[0,2]$ 上 $\lambda g(x) = 1$，所以

$$f(x) = \begin{cases} \frac{1}{3}\phi(x) & x < 0 \text{ 或} x > 2, \\ \frac{1}{3}\big(\phi(x)+1\big) & 0 \le x \le 2. \end{cases}$$

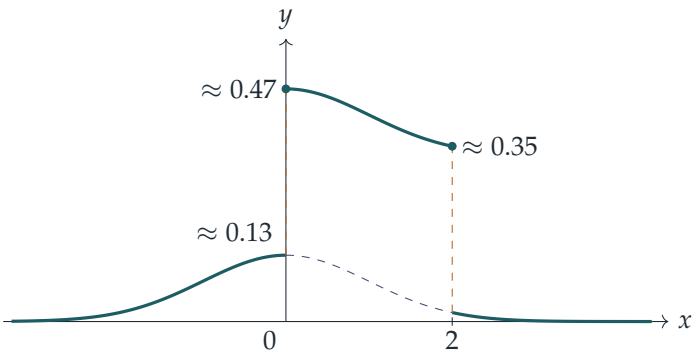
此外

$$\mu = \frac{4}{6} = \frac{2}{3}, \qquad \sigma^2 = \frac{16+32+24+12}{12\cdot 9} = \frac{84}{108} = \frac{7}{9}, \qquad \sigma = \frac{\sqrt{7}}{3}.$$

第 (i) 部分

這條曲線是標準正態鐘形曲線的三分之一，並在區間 $[0,2]$ 上向上提升了 $\frac{1}{3}$。所以它在 $x = 0$ 處向上跳躍，在 $x = 2$ 處向下跳躍，跳躍的高度都是 $\frac{1}{3}$。利用 $\phi(0) = 1/\sqrt{2\pi} \approx 0.399$ 和 $\phi(2) \approx 0.054$：

- 在 0 的左邊緊貼處，$f \approx 0.133$；在 0 的右邊緊貼處，$f \approx 0.466$（最大值）；
- 在 2 的左邊緊貼處，$f \approx 0.351$；在 2 的右邊緊貼處，$f \approx 0.018$；
- 在 $[0,2]$ 上曲線遞減，形狀與正態曲線一致；
- 當 $x \to \pm\infty$ 時 $f \to 0$，而且處處 $f > 0$。



$[0,2]$ 上的紫紅色虛線是 $\frac{1}{3}\phi(x)$，畫出來以作比較；實線則是 $y = f(x)$。（縱軸的比例已被拉長。）

第 (ii) 部分

設 $F(x) = \mathrm{P}(X \le x) = \int_{-\infty}^{x} f(t)\,\mathrm{d}t$。對 $\frac{1}{3}\phi$ 積分得 $\frac{1}{3}\Phi$，而 $[0,2]$ 上額外的 $\frac{1}{3}$ 所貢獻的，是 $\frac{1}{3}$ 乘以 $[0,2]$ 中位於 x 左邊那一段的長度。所以

$$F(x) = \begin{cases} \frac{1}{3}\Phi(x) & x < 0, \\ \frac{1}{3}\big(\Phi(x) + x\big) & 0 \le x \le 2, \\ \frac{1}{3}\big(\Phi(x) + 2\big) & x > 2. \end{cases}$$

各段在 $x = 0$ 和 $x = 2$ 處互相吻合，這是必然的（沒有點質量（point mass）時，累積分佈函數（CDF）是連續的），而且當 $x \to \infty$ 時 $F(x) \to \frac{1}{3}(1+2) = 1$。

第 (iii) 部分

這裏 $\mu + 2\sigma = \frac{2}{3} + \frac{2}{3}\sqrt{7} \approx 2.43$，大於 2，所以在該處要用 F 的最後一個分支。由於 X 是連續的，

$$\begin{aligned} \mathrm{P}(0 < X < \mu + 2\sigma) &= F\big(\tfrac{2}{3} + \tfrac{2}{3}\sqrt{7}\big) - F(0) \\ &= \tfrac{1}{3}\big(\Phi(\tfrac{2}{3} + \tfrac{2}{3}\sqrt{7}) + 2\big) - \tfrac{1}{3}\Phi(0) \\ &= \tfrac{1}{3}\big(0.9921 + 2 - 0.5\big) = \frac{2.4921}{3} = 0.8307. \end{aligned}$$

驗算

- 當 $\lambda \to 0$ 時，均勻部分消失：公式給出 $\mu \to 0$ 和 $\sigma^2 \to \frac{12}{12} = 1$，正是標準正態分佈的數值，合乎預期。

- 對於 $\lambda = 2$，全方差定律（law of total variance，見下文）以另一種方法給出相同的答案。各成分的權重為 $\frac{1}{3}, \frac{2}{3}$，方差為 $1, \frac{1}{3}$，均值為 $0, 1$。方差的平均值是 $\frac{1}{3} + \frac{2}{9} = \frac{5}{9}$，均值的方差是 $\frac{2}{3} - \frac{4}{9} = \frac{2}{9}$，總和是 $\frac{7}{9}$。
- 以數值方法對 f 積分（中點法則，在 0 和 λ 處分段），取 $\lambda = 0.5, 1, 2, 3.7$，所得的總概率 1、均值和方差公式全部吻合，誤差小於 10^{-6}；對於 $\lambda = 2$，結果是 $\mu = 0.6667$ 和 $\sigma^2 = 0.7778$。
- 直接計算 $\Phi(2.4305)$ 得 0.99246，而非題目所引用的 0.9921；概率則是 0.8308，而非 0.8307。這微小的差異源自題目所給的數值，而我們應該採用題目的數值。

常見錯誤

- 忘記了因子 $k = \frac{1}{3}$，以致總概率變成 3。
- 把方差算成 $k[1 + \lambda \cdot \frac{\lambda^2}{12}]$，即把各成分的方差組合起來。混合分佈的方差不能線性地組合；二階矩才可以。
- 把 f 畫成連續曲線：它在 $x = 0$ 和 $x = \lambda$ 處有跳躍。

STEP 之外

形如 $\sum_i w_i f_i(x)$ 的密度，其中權重 $w_i \geq 0$ 且總和為 1，稱為**混合分佈**（mixture distribution）：以概率 w_i 選出成分 i，然後從 f_i 中抽樣。這裏的權重是 $\frac{1}{1+\lambda}$ 和 $\frac{\lambda}{1+\lambda}$。混合分佈的方差其實就是全方差定律的化身：各成分方差的平均值，加上各成分均值的方差。混合分佈是統計學和機器學習中聚類（clustering）方法的基礎。Blitzstein 與 Hwang 的 *Introduction to Probability*（概率論入門教材）對條件期望值（conditional expectation）和全方差定律有清晰的講解。

詳細解答

STEP III 2005，第 1 題

純數學 • 三角方程與圖像

> **STEP III 2005，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/05-s3-q1

關鍵思路

凡是形如 sin(某式) = cos(另一式) 的方程，都可以利用 $\cos B = \sin\left(\frac{\pi}{2} - B\right)$ 轉化為關於兩個自變量（argument）的命題。把這個做法用於 $\sin(\sin x) = \cos(\cos x)$，便得出 $\sin x \mp \cos x$ 必須等於 $\frac{\pi}{2}$ 的奇數倍，而 $|\sin x \pm \cos x|$ 的上界表明它根本太小，無法達到這樣的值。至於草圖，只要想清楚當內層自變量只在一個很短的區間內變動時，外層函數會對它產生甚麼作用，便可畫出。

解答

方程 $\sin A = \cos B$

由於 $\cos B = \sin\left(\frac{\pi}{2} - B\right)$，我們需要求 $\sin A = \sin C$ 的通解（general solution），其中 $C = \frac{\pi}{2} - B$。正弦函數在每個長度為 2π 的週期（period）內，對 $[-1, 1]$ 中的每個值都恰好取到兩次，分別在 C 和 $\pi - C$（當該值為 ± 1 時，兩者重合），所以

$$\sin A = \sin C \iff A = C + 2k\pi \text{ 或 } A = \pi - C + 2k\pi \quad \text{對某個整數} k.$$

代入 $C = \frac{\pi}{2} - B$：

$$\begin{aligned} A &= \tfrac{\pi}{2} - B + 2k\pi = (4k+1)\tfrac{\pi}{2} - B, \\ \text{或} \quad A &= \pi - \tfrac{\pi}{2} + B + 2k\pi = (4k+1)\tfrac{\pi}{2} + B. \end{aligned}$$

因此 $\sin A = \cos B$ 當且僅當存在整數 n 使得 $A = (4n+1)\frac{\pi}{2} \pm B$。兩個方向都已涵蓋，因為上面每一步都是「當且僅當」（if and only if）。（若要直接驗證「若」的方向：$\sin\left((4n+1)\frac{\pi}{2} \pm B\right) = \sin\left(\frac{\pi}{2} \pm B\right) = \cos B$，這裏用到週期為 2π。）

上界與方程 $\sin(\sin x) = \cos(\cos x)$

利用複角公式（compound angle formula），

$$\sin x \pm \cos x = \sqrt{2}\left(\sin x \cos \tfrac{\pi}{4} \pm \cos x \sin \tfrac{\pi}{4}\right) = \sqrt{2}\sin\left(x \pm \tfrac{\pi}{4}\right),$$

又因為對任何 θ 都有 $|\sin\theta| \le 1$，所以對所有 x 都有 $|\sin x \pm \cos x| \le \sqrt{2}$，這正是所要證明的。

現在假設對某個 x 有 $\sin(\sin x) = \cos(\cos x)$。在第一個結果中取 $A = \sin x$ 和 $B = \cos x$，可知存在整數 n 使得

$$\sin x = (4n+1)\tfrac{\pi}{2} \pm \cos x, \qquad \text{即} \qquad \sin x \mp \cos x = (4n+1)\tfrac{\pi}{2}.$$

想一想 為甚麼右邊不可能很小？

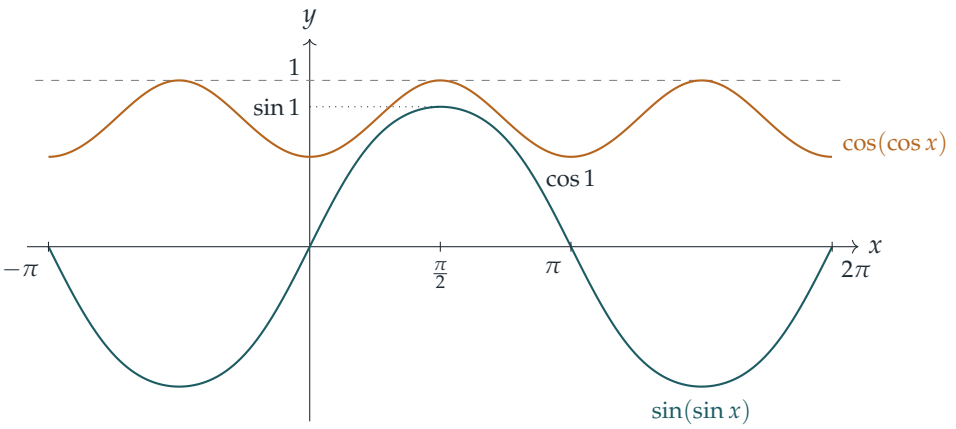
因為 $4n+1$ 是奇數，它永遠不等於 0，所以 $|4n+1| \ge 1$，右邊的絕對值（modulus）最少是 $\frac{\pi}{2} \approx 1.571$。而左邊的絕對值最多是 $\sqrt{2} \approx 1.414$。由於 $\sqrt{2} < \frac{\pi}{2}$（兩邊平方：$2 < \frac{\pi^2}{4}$，因為 $\pi^2 > 9 > 8$），這是不可能的。因此 $\sin(\sin x) = \cos(\cos x)$ 無解。

繪畫 $y = \sin(\sin x)$ 和 $y = \cos(\cos x)$ 的草圖

$y = \sin(\sin x)$。內層函數 $\sin x$ 的取值範圍是 $[-1, 1]$，而在這個區間上 sin 是遞增的（因為 $[-1,1] \subset [-\frac{\pi}{2}, \frac{\pi}{2}]$）。所以 $\sin(\sin x)$ 的升降與 $\sin x$ 完全一致：它的週期為 2π，是奇函數（odd function），在 $x = n\pi$ 處為零，在 $x = \frac{\pi}{2} + 2n\pi$ 處取極大值 $\sin 1 \approx 0.841$，在 $x = -\frac{\pi}{2} + 2n\pi$ 處取極小值 $-\sin 1$。它的斜率（gradient）$\cos x \cos(\sin x)$ 在原點等於 1。它看起來像一條稍為壓扁了的正弦曲線。

$y = \cos(\cos x)$。由於 $\cos(-u) = \cos u$ 及 $\cos(x + \pi) = -\cos x$，我們有 $\cos(\cos(x + \pi)) = \cos(\cos x)$：這個圖像的週期為 π，而且是偶函數（even function）。在 $[-1, 1]$ 上，$\cos u$ 在 $u = 0$ 處最大（$= 1$），在 $u = \pm 1$ 處最小（$= \cos 1 \approx 0.540$）。所以 $\cos(\cos x)$ 在 $\cos x = 0$ 之處，即 $x = \frac{\pi}{2} + n\pi$，取極大值 1；在 $\cos x = \pm 1$ 之處，即 $x = n\pi$，取極小值 $\cos 1$。它永遠介乎 0.54 與 1 之間。

根據上一部分，兩條曲線永不相交；又因為 $\cos(\cos 0) = \cos 1 > 0 = \sin(\sin 0)$，由連續性（continuity）可知餘弦曲線處處嚴格位於正弦曲線之上。兩者在 $x \approx 0.88$ 附近最為接近（相距約 0.107），那裏 $\sin(\sin x)$ 幾乎到達頂峯，而 $\cos(\cos x)$ 仍在上升。



繪畫 $y = \sin(2\sin x)$ 的草圖

這個函數同樣以 2π 為週期，而且是奇函數，零點在 $x = n\pi$，所以只需弄清楚 $0 \leq x \leq \pi$ 的情況；在這個區間上，圖像關於 $x = \frac{\pi}{2}$ 對稱（因為 $\sin(\pi - x) = \sin x$）。

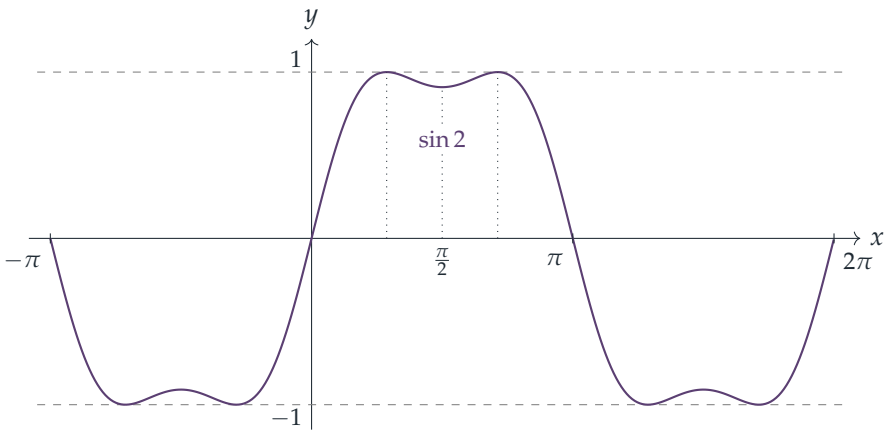
現在的分別在於，內層自變量 $u = 2\sin x$ 在這個區間上的取值範圍是 $[0, 2]$，而 $2 > \frac{\pi}{2}$。所以當 x 由 0 增加到 $\frac{\pi}{2}$ 時：

- u 由 0 上升到 $\frac{\pi}{2}$，同時 $\sin x$ 上升到 $\frac{\pi}{4}$，此時 $x = \arcsin\frac{\pi}{4} \approx 0.903$；期間 $y = \sin u$ 由 0 上升到它的極大值 1；
- u 繼續由 $\frac{\pi}{2}$ 增加到 2，在這一段 $\sin u$ 是遞減的，所以 y 回落，在 $x = \frac{\pi}{2}$ 處降至 $\sin 2 \approx 0.909$。

根據對稱性，y 在 $x = \pi - \arcsin\frac{\pi}{4} \approx 2.238$ 處再次升至 1，然後在 $x = \pi$ 處降至 0。用導數（derivative）驗證，

$$\frac{\mathrm{d}y}{\mathrm{d}x} = 2\cos x \cos(2\sin x),$$

它等於零，恰好是當 $\cos x = 0$（在 $\frac{\pi}{2}$ 處的局部極小值）或 $2\sin x = \pm\frac{\pi}{2}$（值為 1 的極大值）之時；在原點的斜率為 2。在 $[\pi, 2\pi]$ 上，圖像是上述情況的反射：兩個值為 -1 的極小值，中間在 $x = \frac{3\pi}{2}$ 處有一個值為 $-\sin 2$ 的局部極大值。正弦曲線的每一個峯都變成了帶有淺淺凹陷的駝峯。



驗算

- 取 $B = 0.3$ 和 $n = 1$：$A = \frac{5\pi}{2} - 0.3$ 給出 $\sin A = \cos 0.3 \approx 0.9553$，與預期相符。
- 在一個週期內用細密的格點計算 $\cos(\cos x) - \sin(\sin x)$，得到最小值約為 0.107（在 $x \approx 0.878$ 附近），從不為零，這印證了兩條曲線並不相交。
- 對於 $\sin(2\sin x)$，在 $[0, \pi]$ 上作數值搜尋，恰好找到三個轉向點（turning points）：在 $x \approx 0.903$ 和 2.238 處的極大值 1，以及在 $\frac{\pi}{2}$ 處的極小值 $\sin 2 \approx 0.909$。

常見錯誤

- 只寫出 $A = \frac{\pi}{2} - B + 2n\pi$，而遺漏了第二組解 $A = \frac{\pi}{2} + B + 2n\pi$，這組解來自 $\sin(\pi - C) = \sin C$。
- 忘記了 $n = 0$ 也是容許的，只檢查 $|4n + 1| \geq 5$ 的情況：最關鍵的比較，正是與 $\frac{\pi}{2}$ 本身的比較。
- 把 $\sin(2\sin x)$ 畫成一條高度為 1 的普通正弦曲線：由於 $2 > \frac{\pi}{2}$，內層自變量會越過正弦函數的頂峯，因而在 $x = \frac{\pi}{2}$ 處出現凹陷。

STEP 之外

這裏的草圖全都關乎函數的複合（composition）：$f(g(x))$ 的形狀取決於 g 的值域（range）覆蓋了 f 的圖像的哪一部分。把這個想法反覆迭代（$\sin(\sin(\sin x))$，如此類推），便會引向不動點（fixed points）和迭代映射（iterated maps）的動力學，當中要問的是 $f(f(\cdots f(x)))$ 何時會穩定下來。至於那個不等式論證，即比較等式兩邊各自可能的取

值範圍，則是分析學（analysis）中隨處可見的一種技巧的簡單例子；Abbott 的 *Understanding Analysis* 是了解如何把這類論證嚴格化的好去處。

詳細解答

STEP III 2005，第 2 題

純數學 • 可分離變量微分方程

STEP III 2005，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/05-s3-q2

關鍵思路

這條微分方程（differential equation）是可分離變數（separable）的，積分後得到的對數可以合併成題目所述的形式。至於距離的問題，應求距離的平方 $D = x^2 + y^2$ 的最小值，這樣便可避開平方根。在曲線上，y^2 是 x 的顯函數，故 D 變成只含 x 的函數，而其導數可以俐落地因式分解。會出現哪些駐點（stationary point），取決於 c 與 a^2 的大小關係。

解答

通解

分離變數（設 $y \neq 0$），

$$\int \frac{1}{y}\,\mathrm{d}y = -\int \frac{x}{x^2+a^2}\,\mathrm{d}x \quad\Longrightarrow\quad \ln|y| = -\tfrac{1}{2}\ln(x^2+a^2) + k,$$

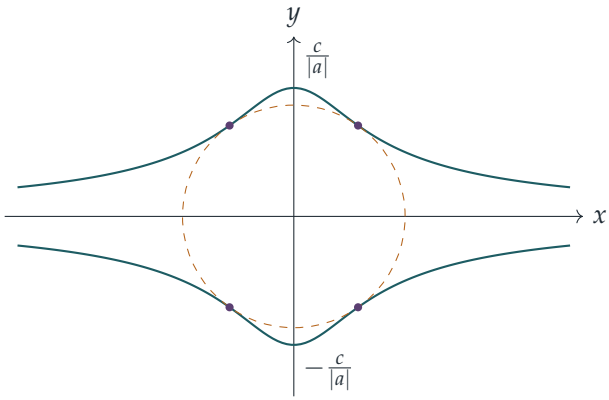
其中 k 為常數。注意 $x^2 + a^2 > 0$ 對任何 x 均成立，因為 $a \neq 0$，所以右邊的對數處處有定義。兩邊乘以二再取指數，

$$y^2 = \frac{\mathrm{e}^{2k}}{x^2+a^2}, \qquad \text{故} \qquad y^2(x^2+a^2) = c^2,$$

其中 $c^2 = \mathrm{e}^{2k}$。解 $y = 0$ 在我們除以 y 時遺失了，它正是 $c = 0$ 的情況，所以 c 確實是任意常數。明確地寫出來，通解（general solution）是 $y = c/\sqrt{x^2+a^2}$，其中 c 可正可負。由於曲線的方程中只出現 c^2，以下我們取 $c > 0$。

繪畫曲線

曲線 $y^2(x^2+a^2)=c^2$ 由 $y=\pm c/\sqrt{x^2+a^2}$ 兩支組成。它對兩條坐標軸都對稱（方程中只出現 x^2 和 y^2），而且永不與 x 軸相交。在上面的一支，y 在 x^2+a^2 最小時最大，於是在 $x=0$ 處取得極大值 $c/|a|$。當 $|x|\to\infty$ 時，我們有 $y\approx\pm c/|x|\to 0$，所以 x 軸是漸近線（asymptote）。下圖所示為 $a=1$、$c=2$ 的情況，並畫出了我們將在第 (ii) 部分求得的圓。



x^2+y^2 的導數

設 $D=x^2+y^2$。先利用微分方程，再代入 $y^2=c^2/(x^2+a^2)$，

$$\frac{\mathrm{d}D}{\mathrm{d}x}=2x+2y\frac{\mathrm{d}y}{\mathrm{d}x}=2x-\frac{2xy^2}{x^2+a^2}=2x\left(1-\frac{c^2}{(x^2+a^2)^2}\right).$$

此式在兩支上都成立，因為它只涉及 y^2。用積法則（product rule）再求導一次，並利用 $\frac{\mathrm{d}}{\mathrm{d}x}(x^2+a^2)^{-2}=-4x(x^2+a^2)^{-3}$，

$$\frac{\mathrm{d}^2D}{\mathrm{d}x^2}=2\left(1-\frac{c^2}{(x^2+a^2)^2}\right)+2x\cdot\frac{4c^2x}{(x^2+a^2)^3}=2\left(1-\frac{c^2}{(x^2+a^2)^2}\right)+\frac{8c^2x^2}{(x^2+a^2)^3},$$

即為所求。

D 的駐點出現在 $x=0$ 或 $(x^2+a^2)^2=c^2$ 之處。由於 $x^2+a^2>0$ 及 $c>0$，第二個條件即 $x^2+a^2=c$，亦即 $x^2=c-a^2$。

第 (i) 部分

設 $0<c<a^2$。對任何 x，我們有 $x^2+a^2\geq a^2>c$，所以 $(x^2+a^2)^2>c^2$，而 $\mathrm{d}D/\mathrm{d}x$ 中的括號嚴格為正。因此 $\mathrm{d}D/\mathrm{d}x$ 的正負號與 x 相同：D 在 $x<0$ 時遞減，在 $x>0$ 時遞增。所以 D 在 $x=0$ 處取得全局最小值（global minimum），而且別無他處。（與此吻合的是，該處的二階導數為 $2(1-c^2/a^4)>0$。）

在 $x = 0$ 處，$y^2 = c^2/a^2$，所以最接近原點的點是

$$\left(0,\ \pm\frac{c}{a}\right),$$

與原點的距離為 $c/|a|$。這就是所要求的結果。

第 (ii) 部分

現在設 $c > a^2$。這時有三個駐點：$x = 0$ 和 $x = \pm\sqrt{c - a^2}$。

> **想一想** $x = 0$ 仍然是最近的點嗎？

不是。在 $x = 0$ 處，二階導數為 $2(1 - c^2/a^4)$，它是負數，因為由 $c > a^2 > 0$ 可得 $c^2 > a^4$。所以 $x = 0$ 現在是 D 的局部極大值（local maximum）。在 $x^2 = c - a^2$ 處，我們有 $x^2 + a^2 = c$，所以括號等於零，而

$$\frac{\mathrm{d}^2 D}{\mathrm{d}x^2} = \frac{8c^2(c - a^2)}{c^3} = \frac{8(c - a^2)}{c} > 0,$$

得到局部極小值（local minimum）。要知道它們是全局最小值，可以考察 $\mathrm{d}D/\mathrm{d}x$ 的正負號。當 $x > 0$ 時，括號在 $x^2 < c - a^2$ 時為負，在 $x^2 > c - a^2$ 時為正，所以在 $x \geq 0$ 上，函數 D 一直遞減至 $x = \sqrt{c - a^2}$，其後遞增。因此 D 在 $x \geq 0$ 上的最小值出現在 $x = \sqrt{c - a^2}$，而根據對稱性，它在 $x \leq 0$ 上的最小值出現在 $x = -\sqrt{c - a^2}$。

在這些點上，$y^2 = c^2/(x^2 + a^2) = c^2/c = c$。所以曲線上最接近原點的點是以下四點

$$\left(\pm\sqrt{c - a^2},\ \pm\sqrt{c}\right)$$

（正負號的全部四種組合），與原點的距離為 $\sqrt{2c - a^2}$。

作為覆核，在 $x = 0$ 處距離的平方是 c^2/a^2，而

$$\frac{c^2}{a^2} - (2c - a^2) = \frac{(c - a^2)^2}{a^2} > 0,$$

所以 y 軸上的點確實較遠。

驗算

- 把 $y = c/\sqrt{x^2 + a^2}$ 代回微分方程，結果成立：$y' = -cx(x^2 + a^2)^{-3/2} = -xy/(x^2 + a^2)$。
- 取 $a = 1$、$c = 2$，以窮舉方式搜尋 $x^2 + 4/(x^2 + 1)$ 的值，可找到最小值 $3 = 2c - a^2$，出現在 $x = \pm 1 = \pm\sqrt{c - a^2}$，而 $y^2 = 2 = c$。這就是圖中的四個圓點，即半徑為 $\sqrt{3}$ 的虛線圓與曲線相切之處。
- 當 $c = a^2$ 時，兩個答案一致：$\sqrt{c - a^2} = 0$ 且 $\sqrt{c} = c/|a| = |a|$，所以四個點合併成 $(0, \pm|a|)$。

常見錯誤

- 由「$\mathrm{d}D/\mathrm{d}x = 0$ 且 $\mathrm{d}^2D/\mathrm{d}x^2 > 0$」便斷定某點的距離最短。這只能說明它是局部極小值；全局的問題要靠 $\mathrm{d}D/\mathrm{d}x$ 的正負號來解決。
- 忘記了下面的一支，以致在第 (ii) 部分只給出兩個點而不是四個。
- 寫成 $x^2 + a^2 = \pm c$：負號的情況不可能發生，因為 $x^2 + a^2 > 0$。

STEP 之外

在 $y^2(x^2 + a^2) = c^2$ 的條件下求 $x^2 + y^2$ 的最小值，是一個約束優化（constrained optimisation）問題。在大學裏，你會學到拉格朗日乘數（Lagrange multipliers），它毋須先解出其中一個變數便能處理這類問題：在最近的點上，曲線的法線（normal）沿半徑方向，這正是虛線圓與曲線相切的原因。當 c 越過 a^2 時，原本在 $x = 0$ 的單一最小值一分為二，這是叉式分岔（pitchfork bifurcation）的一個簡單例子。想進一步了解可分離變數方程及其解曲線，可參閱 Boyce and DiPrima, *Elementary Differential Equations*（常微分方程入門教科書）。

詳細解答

STEP III 2005，第 3 題

純數學 • 二次函數的複合

STEP III 2005，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/05-s3-q3

關鍵思路

把兩個首一（monic）二次多項式複合（compose），會得到一個首一四次多項式（quartic），而它的係數是互相牽連的：四個參數 p,q,r,s 要產生四個係數 a,b,c,d，但在其中兩個係數裏，p 和 s 只會以 $2s+p$ 這個組合出現，所以必定有一條關係式要成立。第二種形式其實是同一回事，只是換了個樣子：對外層的二次式配方（completing the square），就把 $\mathrm{f}(\mathrm{g}(x))$ 變成（一個平方）$-$（一個常數）。正是這個形式，讓我們只用二次公式（quadratic formula）就能解出最後的四次方程。

解答

把二次式複合

把 f 中的 x 換成 $\mathrm{g}(x)$，

$$\begin{aligned}\mathrm{f}(\mathrm{g}(x)) &= (x^2+rx+s)^2 + p(x^2+rx+s) + q \\ &= x^4 + 2rx^3 + (r^2+2s+p)x^2 + r(2s+p)x + (s^2+ps+q).\end{aligned}$$

所以 $x^4+ax^3+bx^2+cx+d = \mathrm{f}(\mathrm{g}(x))$ 對所有 x 成立，當且僅當

$$a = 2r, \qquad b = r^2+(2s+p), \qquad c = r(2s+p), \qquad d = s^2+ps+q.$$

必要性（necessity）。若這樣的 p,q,r,s 存在，則 $r=\frac{1}{2}a$ 且 $2s+p = b-r^2 = b-\frac{1}{4}a^2$，所以

$$c = \tfrac{1}{2}a\big(b-\tfrac{1}{4}a^2\big), \qquad \text{即} \qquad a^3-4ab+8c=0.$$

充分性（sufficiency）。反過來，假設 $a^3-4ab+8c=0$。取 $r=\frac{1}{2}a$，s 可取任意值（例

如 $s=0$），然後令 $p=b-\frac{1}{4}a^2-2s$ 及 $q=d-s^2-ps$。第一、第二和第四條等式按構造自然成立，而第三條就是 $c=\frac{1}{2}a(b-\frac{1}{4}a^2)$，正是我們的條件移項後的樣子。留意 d 完全不受限制，因為 q 只在常數項中出現。

所以所求的充要條件（necessary and sufficient condition）是

$$\boxed{a^3-4ab+8c=0}$$

（d 可為任意值）。

$(x^2+vx+w)^2-k$ 這個形式

想一想 像 $\mathrm{g}^2+p\,\mathrm{g}+q$ 這樣的和，怎樣才能改寫成一個平方減去一個常數？

把 $\mathrm{g}(x)$ 當作單一個變數，對 g 配方：

$$\mathrm{f}(\mathrm{g}(x))=\left(\mathrm{g}(x)+\tfrac{1}{2}p\right)^2+q-\tfrac{1}{4}p^2=\left(x^2+rx+s+\tfrac{1}{2}p\right)^2-\left(\tfrac{1}{4}p^2-q\right).$$

現在兩個方向都可以證明了。

若條件成立，由第一部分可得 p,q,r,s 使該四次式等於 $\mathrm{f}(\mathrm{g}(x))$，而上面的式子顯示它等於 $(x^2+vx+w)^2-k$，其中

$$v=r,\qquad w=s+\tfrac{1}{2}p,\qquad k=\tfrac{1}{4}p^2-q.$$

反過來，$(x^2+vx+w)^2-k$ 本身就是 $\mathrm{f}(\mathrm{g}(x))$，其中 $\mathrm{g}(x)=x^2+vx+w$ 而 $\mathrm{f}(x)=x^2-k$（即 $p=0,q=-k,r=v,s=w$）。根據第一部分的必要性論證，條件 $a^3-4ab+8c=0$ 必定成立。

這樣就完成了「當且僅當」（if and only if）的證明。實際運用時，把各個值明確寫出來會方便得多：展開得

$$(x^2+vx+w)^2-k=x^4+2vx^3+(v^2+2w)x^2+2vw\,x+(w^2-k),$$

所以我們需要

$$v=\tfrac{1}{2}a,\qquad w=\tfrac{1}{2}\left(b-\tfrac{1}{4}a^2\right)=\tfrac{1}{8}(4b-a^2),\qquad k=w^2-d,$$

而這時 x 的係數是 $2vw=\frac{1}{2}a(b-\frac{1}{4}a^2)$，它恰好在條件成立時等於 c。

解四次方程

這裏 $a=-4,b=10,c=-12,d=4$，而且

$$a^3-4ab+8c=-64+160-96=0,$$

所以這個方法適用。由公式得

$$v = -2, \qquad w = \tfrac{1}{8}(40 - 16) = 3, \qquad k = 3^2 - 4 = 5,$$

而確實有 $(x^2 - 2x + 3)^2 - 5 = x^4 - 4x^3 + 10x^2 - 12x + 4$（值得親手展開一次，以策萬全）。方程變成

$$(x^2 - 2x + 3)^2 = 5 \quad \Longleftrightarrow \quad x^2 - 2x + 3 = \sqrt{5} \ \text{ 或 } \ x^2 - 2x + 3 = -\sqrt{5}.$$

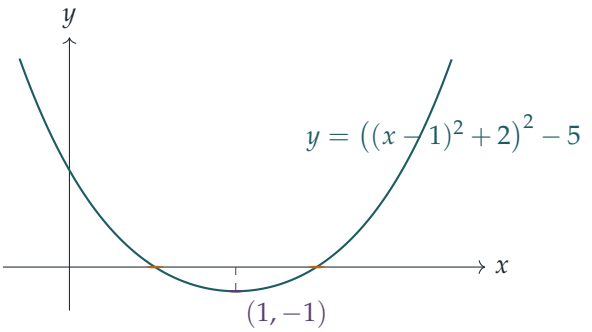
寫成 $x^2 - 2x + 3 = (x-1)^2 + 2$：

- $(x-1)^2 = \sqrt{5} - 2$。由於 $\sqrt{5} > 2$，右邊是正數，得出兩個實根 $x = 1 \pm \sqrt{\sqrt{5} - 2}$。
- $(x-1)^2 = -2 - \sqrt{5} < 0$，得出一對複數根 $x = 1 \pm \mathrm{i}\sqrt{2 + \sqrt{5}}$。

所以方程的根是

$$\boxed{x = 1 \pm \sqrt{\sqrt{5} - 2}, \qquad x = 1 \pm \mathrm{i}\sqrt{2 + \sqrt{5}}}$$

其中只有第一對是實數（如果把題目理解為只求實根，答案就是這兩個）。



圖像關於 $x = 1$ 對稱（它只取決於 $(x-1)^2$），在 $x = 1$ 處取得最小值 $2^2 - 5 = -1$，所以它與橫軸恰好相交兩次。

驗算

- 數值上，$1 \pm \sqrt{\sqrt{5} - 2} \approx 1.4859$ 和 0.5141；把其中任何一個代入 $x^4 - 4x^3 + 10x^2 - 12x + 4$，結果與零相差不超過 10^{-14}，代入 $1 \pm 2.0582\,\mathrm{i}$ 亦然。
- 根之和：四個根加起來是 $4 = -a$，正如所料。
- 試試自己構造一個複合式，例如 $\mathrm{f}(x) = x^2 + x$, $\mathrm{g}(x) = x^2 + 2x$：會得到 $x^4 + 4x^3 + 5x^2 + 2x$，而 $a^3 - 4ab + 8c = 64 - 80 + 16 = 0$。

常見錯誤

- 對 d 加上條件：常數項總可以用 q（或 k）來配合，所以 d 是自由的。
- 只證明了「當且僅當」的其中一個方向；只要留意到 $(x^2+vx+w)^2-k$ 本身就是 $\mathrm{f}(\mathrm{g}(x))$ 的形式，一行就能得出逆命題。
- 求出實根後就停筆，沒有留意另一個因式會給出一對複數根。

STEP 之外

條件 $a^3-4ab+8c=0$ 有一個簡潔的意義：代入 $x=y-\frac{1}{4}a$ 可消去三次項，而這個條件正好表示所得的降次（depressed）四次式沒有 y 項，因此它是 y^2 的二次式。這正是費拉里（Ferrari）四次方程一般解法的起點；更廣泛地說，它也是「多項式何時能分解為複合式」這個問題的起點，由此可通往伽羅瓦理論（Galois theory）。Allenby 的 *Rings, Fields and Groups*（環、域與羣）很適合用來了解哪些多項式方程可用根式求解（solved by radicals）背後的代數。

詳細解答

STEP III 2005，第 4 題

純數學 • 非線性遞推關係

> **STEP III 2005，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/05-s3-q4

關鍵思路

這個遞推關係（recurrence relation）並非線性，但它其實是關於相鄰兩項之比的。若寫成 $u_{n+1} = \rho\, u_n$，遞推關係告訴我們下一個比就是 $k - \rho$。因此，這些比在兩個和為 k 的數值之間來回交替，其餘一切皆由此而來。題目的三個特殊情況，也就化為對這兩個比的條件。

解答

歸納法

設 $r = \frac{b}{a}$。把遞推關係兩邊除以 u_{n+1}（當它不等於零時），得

$$\frac{u_{n+2}}{u_{n+1}} = k - \frac{u_{n+1}}{u_n},$$

所以每一個相鄰兩項之比，都等於 k 減去前一個比。第一個比是 $u_2/u_1 = r$，故下一個比應為

$$c = k - \frac{b}{a} = \frac{ka - b}{a}.$$

注意 $c \neq 0$，因為 $b \neq ka$。另外 $r \neq 0$，因為 $b \neq 0$。這兩點合起來保證每一項都不等於零，所以我們永遠不會除以零。

設 $P(n)$ 為命題：$u_{2n-1} \neq 0$ 且 $u_{2n} = r\, u_{2n-1}$。$P(1)$ 成立，因為 $u_1 = a \neq 0$ 且 $u_2 = b = ra$。

假設 $P(n)$ 對某個 $n \geq 1$ 成立。那麼 $u_{2n} = r u_{2n-1} \neq 0$，而由遞推關係得

$$u_{2n+1} = \frac{u_{2n}}{u_{2n-1}}\left(k u_{2n-1} - u_{2n}\right) = r\,(k - r)\, u_{2n-1} = (k - r)\, u_{2n} = c\, u_{2n},$$

這正是題目要求的第二條關係式。因此 $u_{2n+1} \neq 0$，並且

$$u_{2n+2} = \frac{u_{2n+1}}{u_{2n}}\left(ku_{2n} - u_{2n+1}\right) = c\,(k-c)\,u_{2n} = (k-c)\,u_{2n+1} = r\,u_{2n+1},$$

因為 $k - c = r$。這就是 $P(n+1)$。根據數學歸納法（induction），對所有 $n \geq 1$，

$$u_{2n} = \frac{b}{a}\,u_{2n-1}, \qquad u_{2n+1} = c\,u_{2n}, \qquad c = k - \frac{b}{a}.$$

想一想 為甚麼要把「$u_{2n-1} \neq 0$」寫進歸納命題之中？

遞推關係要除以 u_n，所以只有當沒有任何一項等於零時，這個數列才有意義。把「不等於零」這一點帶進歸納過程，便可以在證明公式的同時，一併證明各項均有定義。$b \neq 0$ 和 $b \neq ka$ 這兩個條件，正是在這裏用上的。

通項公式

每走兩步，先乘以 c，再乘以 $\frac{b}{a}$，所以 $u_{2n+1} = \frac{bc}{a}\,u_{2n-1}$。奇數項構成一個等比數列（geometric progression），首項為 a，公比為 $\frac{bc}{a}$；而每個偶數項都是它前面那個奇數項的 $\frac{b}{a}$ 倍：

$$u_{2n-1} = a\left(\frac{bc}{a}\right)^{n-1} = \frac{b^{n-1}c^{n-1}}{a^{n-2}}, \qquad u_{2n} = b\left(\frac{bc}{a}\right)^{n-1} = \frac{b^n c^{n-1}}{a^{n-1}}.$$

第 (i) 部分

相鄰兩項之比依次為 $\frac{b}{a}, c, \frac{b}{a}, c, \ldots$，交替出現，所以數列是等比數列，當且僅當 $c = \frac{b}{a}$，即 $k - \frac{b}{a} = \frac{b}{a}$：

$$ka = 2b.$$

此時公比為 $\frac{b}{a}$。這與 $b \neq ka$ 並無矛盾，因為當 $b \neq 0$ 時，$ka = 2b \neq b$。

第 (ii) 部分

由通項公式可知，對所有 n 都有 $u_{n+2} = \frac{bc}{a}\,u_n$。所以 $u_{n+2} = u_n$ 對所有 n 成立，當且僅當 $\frac{bc}{a} = 1$，即 $b(ka - b) = a^2$：

$$kab = a^2 + b^2.$$

此時數列為 $a, b, a, b, \ldots$。若要求週期（period）恰好是 2（而不是常數數列），還需要 $a \neq b$。若 $a = b$，條件變成 $k = 2$，而每一項都等於 a。

第 (iii) 部分

同理，$u_{n+4} = \left(\frac{bc}{a}\right)^2 u_n$，所以 $u_{n+4} = u_n$ 對所有 n 成立，當且僅當 $\left(\frac{bc}{a}\right)^2 = 1$，即 $\frac{bc}{a} = \pm 1$。取值 $+1$ 就是第 (ii) 部分的情況，週期為 2（或 1）。若要週期恰好是 4，則需要

$\frac{bc}{a} = -1$，即 $b(ka - b) = -a^2$：

$$kab = b^2 - a^2.$$

此時數列為 $a, b, -a, -b, a, b, \ldots$。由於 $a \neq 0$，我們有 $-a \neq a$，所以它的週期確實是 4，而不會更小。

在 (ii) 和 (iii) 兩部分中，條件 $b \neq ka$ 都自動成立。把 $b = ka$ 代入任何一個條件，都會得出 $a^2 = 0$，而這是不可能的。

驗算

試取 $a = 1$，$b = 2$。當 $k = 4$（第 (i) 部分），遞推關係給出 $1, 2, 4, 8, 16, \ldots$。當 $k = \frac{5}{2}$（第 (ii) 部分），給出 $1, 2, 1, 2, \ldots$。當 $k = \frac{3}{2}$（第 (iii) 部分），給出 $1, 2, -1, -2, 1, 2, \ldots$。我們亦以精確的有理數運算，對大約 2000 組隨機選取的 a、b 和 k 執行了這個遞推關係。每一組的結果都與通項公式吻合，而且數列是等比數列、週期為 2 或週期為 4，恰好就是上述條件成立的時候。

常見錯誤

- 遞推關係要除以 u_n，卻忘記說明為甚麼沒有任何一項會等於零。
- 把 $\frac{bc}{a} = \pm 1$ 當作 (iii) 的答案。$+1$ 的情況週期是 2。
- 在 (ii) 中，沒有留意到 $a = b$、$k = 2$ 會得出常數數列。

STEP 之外

代換 $\rho_n = u_{n+1}/u_n$ 把這個非線性遞推關係化為一階映射 $\rho \mapsto k - \rho$。這個映射是一個對合（involution）：連續施行兩次便會回到起點，這就是為甚麼各個比的週期是 2。透過變數代換來簡化非線性的遞推關係或微分方程（differential equation），是大學一年級的標準技巧；而研究映射的週期軌道（periodic orbit），正是離散動力系統（discrete dynamical systems）的起點。想找更多同類型的題目，可參閱 Siklos 的 *Advanced Problems in Mathematics*。

詳細解答

STEP III 2005，第 5 題

純數學 • 拋物線的公切線

STEP III 2005，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/05-s3-q5

關鍵思路

與其用切點來標記切線（tangent），不如用斜率（gradient）來標記。對每一個斜率 m，拋物線（parabola）都恰好有一條該斜率的切線，而第一部分告訴我們它的 y 軸截距（intercept）。兩條拋物線有一條斜率為 m 的公切線（common tangent），當且僅當這兩個截距相等；於是數公切線的數目，就變成數一條方程的根 m 的數目。這條方程在 $a_1 \neq a_2$ 時是二次方程，在 $a_1 = a_2$ 時是一次方程。

解答

斜率為 m 的切線

$y = ax^2 + bx + c$ 的斜率是 $2ax + b$。由於 $a \neq 0$，它恰好在一點等於 m，即 $x_0 = \dfrac{m-b}{2a}$，而 $y_0 = ax_0^2 + bx_0 + c$。該點的切線是 $y - y_0 = m(x - x_0)$，即 $y - mx = y_0 - mx_0$。利用 $m = 2ax_0 + b$，

$$\begin{aligned} y_0 - mx_0 &= ax_0^2 + bx_0 + c - (2ax_0 + b)x_0 \\ &= c - ax_0^2 = c - a \cdot \frac{(m-b)^2}{4a^2} = c - \frac{(m-b)^2}{4a}, \end{aligned}$$

所以在 P 點的切線是 $y - mx = c - \dfrac{(m-b)^2}{4a}$，正如所求。

有公切線的條件

由第一部分，每條曲線都恰好有一條斜率為 m 的切線，分別是

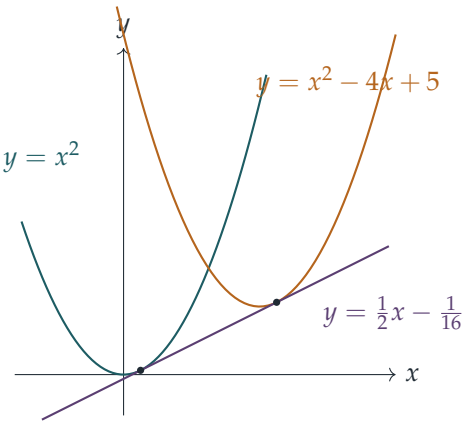
$$y = mx + k_1, \quad k_1 = c_1 - \frac{(m-b_1)^2}{4a_1}, \qquad \text{和} \qquad y = mx + k_2, \quad k_2 = c_2 - \frac{(m-b_2)^2}{4a_2}.$$

兩條斜率相同的直線重合，當且僅當它們的截距相等，所以兩條曲線有一條斜率為 m 的公切線，當且僅當 $k_1 = k_2$。乘以 $4a_1a_2 \neq 0$（這不會改變等式是否成立），

$$\begin{aligned} 4a_1a_2(k_2 - k_1) &= 4a_1a_2(c_2 - c_1) - a_1(m-b_2)^2 + a_2(m-b_1)^2 \\ &= (a_2 - a_1)m^2 + 2(a_1b_2 - a_2b_1)m + 4a_1a_2(c_2 - c_1) + a_2b_1^2 - a_1b_2^2. \end{aligned}$$

因此，有一條斜率為 m 的公切線，當且僅當這個式子等於零，這就是所求的條件。把該條件的左方記作 $Q(m)$。

由於 $Q(m) = 0$ 的每一個根 m 都恰好給出一條公切線，而不同的根給出不同的直線（它們的斜率不同），所以*公切線的數目等於 $Q(m) = 0$ 的相異實根的數目*。還要留意，公切線不可能是鉛垂的，因為兩條曲線都沒有鉛垂的切線。



兩條滿足 $a_1 = a_2$ 及 $b_1 \neq b_2$ 的拋物線恰好有一條公切線；圖中這條公切線的斜率是 $\frac{1}{2}$。

$a_1 \neq a_2$ 的情況

這時 Q 是一個真正的 m 的二次式，所以它恰好有一個相異實根，當且僅當它的判別式（discriminant）等於零。判別式是

$$\Delta = 4(a_1b_2 - a_2b_1)^2 - 4(a_2 - a_1)\big(4a_1a_2(c_2 - c_1) + a_2b_1^2 - a_1b_2^2\big).$$

不含 c 的各項可以化簡得很漂亮：

$$\begin{aligned}&(a_1b_2-a_2b_1)^2-(a_2-a_1)(a_2b_1^2-a_1b_2^2)\\&=\left(a_1^2b_2^2-2a_1a_2b_1b_2+a_2^2b_1^2\right)-\left(a_2^2b_1^2-a_1a_2b_2^2-a_1a_2b_1^2+a_1^2b_2^2\right)\\&=a_1a_2\left(b_1^2-2b_1b_2+b_2^2\right)=a_1a_2(b_1-b_2)^2.\end{aligned}$$

因此

$$\Delta=4a_1a_2\Big[(b_1-b_2)^2-4(a_1-a_2)(c_1-c_2)\Big],$$

而由於 $a_1a_2\neq0$，恰好有一條公切線，當且僅當

$$(b_1-b_2)^2=4(a_1-a_2)(c_1-c_2). \qquad (*)$$

? **想一想** 兩條曲線相切（touch）究竟是甚麼意思？怎樣把它化成一條方程？

兩條曲線相切，是指它們在某一點相交，而且在該點的斜率相同。設

$$d(x)=(a_1-a_2)x^2+(b_1-b_2)x+(c_1-c_2),$$

即兩個二次式之差。兩條曲線在 x_0 相交，即 $d(x_0)=0$；在該處斜率相等，即 $d'(x_0)=0$。所以它們相切，當且僅當 d 有一個根 x_0 滿足 $d'(x_0)=0$，也就是有重根（repeated root）。由於 $a_1\neq a_2$，d 是二次式，而它有重根，正正是當它的判別式 $(b_1-b_2)^2-4(a_1-a_2)(c_1-c_2)$ 等於零，這就是條件 $(*)$。

所以，當 $a_1\neq a_2$ 時，兩條曲線恰好有一條公切線，當且僅當 $(*)$ 成立，亦即當且僅當它們相切。　　（在這個情況下，唯一的公切線自然就是切點處的切線。）

$a_1=a_2$ 的情況

把共同的值記作 a。Q 的 m^2 項消失，再除以 $a\neq0$，條件變成

$$2(b_2-b_1)m+4a(c_2-c_1)+(b_1-b_2)(b_1+b_2)=0.$$

- 若 $b_1\neq b_2$，這條一次方程恰好有一個根，

  $$m=\frac{b_1+b_2}{2}-\frac{2a(c_2-c_1)}{b_2-b_1},$$

  所以恰好有一條公切線。

- 若 $b_1=b_2$，條件化為 $4a(c_2-c_1)=0$。若 $c_1\neq c_2$，則完全沒有公切線（其中一條曲線是另一條沿鉛垂方向平移而成，所以它們每個斜率的切線都互相平行而不重合）；若 $c_1=c_2$，則兩條曲線完全相同，每一條切線都是公切線，共有無限多條。

因此，當 $a_1=a_2$ 時，兩條曲線恰好有一條公切線，當且僅當 $b_1\neq b_2$。

驗算

對於 $y = x^2$ 和 $y = -x^2 + 2x - \frac{1}{2}$，兩者之差是 $2(x - \frac{1}{2})^2$，所以它們在 $x = \frac{1}{2}$ 相切；這時 $Q(m) = -2m^2 + 4m - 2 = -2(m-1)^2$，只有一個根 $m = 1$，正是 $y = x^2$ 在 $x = \frac{1}{2}$ 的斜率。對於圖中的一對曲線 $y = x^2$ 和 $y = x^2 - 4x + 5$，公式給出 $m = -2 + \frac{5}{2} = \frac{1}{2}$，而兩條切線的截距都是 $-\frac{1}{16}$。我們還用精確運算，以數千組隨機係數驗證了恆等式 $Q(m) = 4a_1a_2(k_2 - k_1)$ 和判別式的化簡。

常見錯誤

- 聲稱 $Q = 0$ 有重根就「表示」兩條曲線相切，卻沒有解釋原因：兩者的關聯要經過另一個二次式 $d(x)$，而把它們連繫起來的正是判別式的恆等式。
- 在 $a_1 = a_2$ 的情況下，遺漏了 $b_1 = b_2$、$c_1 = c_2$ 這個子情況（兩條曲線完全相同，有無限多條公切線），或者答成「$b_1 \neq b_2$ 或 $c_1 \neq c_2$」。

STEP 之外

不用曲線上的點，而用它的一族切線來描述曲線，這就是射影幾何（projective geometry）中對偶性（duality）的想法：圓錐曲線（conic）的切線 $y = mx + k$ 滿足 m 與 k 之間自成一格的二次關係，而兩條處於一般位置的圓錐曲線有四條公切線，正如它們相交於四點一樣。對於兩條對稱軸為鉛垂的拋物線，無窮遠線（line at infinity）與兩者都相切，並佔去了這四條中的兩條，所以餘下的普通公切線最多只有兩條。想做更多同類的題目，可以試試 Siklos 的 *Advanced Problems in Mathematics*（進階數學難題集）。

詳細解答

STEP III 2005，第 6 題

純數學 • 解三次方程

> **STEP III 2005，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/05-s3-q6

關鍵思路

恆等式（identity）$4\cosh^3 y - 3\cosh y = \cosh 3y$ 的形狀，與「缺項」三次式（depressed cubic）$x^3 - 3bx$ 完全一樣，只要把 x 乘上適當的常數便可。因此，沒有 x^2 項的三次方程，可以透過與三倍角公式（triple-angle formula）對照來求解，而 arcosh 的對數（logarithm）表達式則把雙曲函數（hyperbolic function）形式的答案變回立方根。求得一個根之後，因式分解（factorisation）便給出另外兩個根所滿足的二次方程，再用單位立方根（cube roots of unity）把它們整理妥當。

解答

雙曲函數形式的根

設 $x = 2a\cosh\frac{1}{3}T$。則

$$\begin{aligned} x^3 - 3a^2x &= 8a^3\cosh^3\tfrac{1}{3}T - 6a^3\cosh\tfrac{1}{3}T \\ &= 2a^3\left(4\cosh^3\tfrac{1}{3}T - 3\cosh\tfrac{1}{3}T\right) \\ &= 2a^3\cosh T, \end{aligned}$$

其中用了題目給出的恆等式，並取 $y = \frac{1}{3}T$。所以 $2a\cosh\frac{1}{3}T$ 滿足該方程，這正是所要證明的。

$x^3 - 3bx = 2c$ 的一個根

先假設 $c > 0$。由於 $b > 0$，可取 $a = \sqrt{b} > 0$，使得 $3a^2 = 3b$，而我們需要 $2a^3 \cosh T = 2c$，即

$$\cosh T = \frac{c}{b^{3/2}}.$$

這需要 $c/b^{3/2} \geq 1$，而題目的條件正正用在這裏：$c^2 \geq b^3$ 且 $c > 0$，故 $c \geq b^{3/2}$。所以可取

$$T = \operatorname{arcosh} \frac{c}{b^{3/2}} = \ln\left(\frac{c}{b^{3/2}} + \sqrt{\frac{c^2}{b^3} - 1} \right) = \ln \frac{c + \sqrt{c^2 - b^3}}{b^{3/2}}.$$

因此

$$e^{T/3} = \frac{(c + \sqrt{c^2 - b^3})^{1/3}}{b^{1/2}} = \frac{u}{\sqrt{b}}, \qquad e^{-T/3} = \frac{\sqrt{b}}{u},$$

而根據第一部分，其中一個根是

$$2\sqrt{b} \cosh \tfrac{1}{3}T = \sqrt{b}\left(\frac{u}{\sqrt{b}} + \frac{\sqrt{b}}{u} \right) = u + \frac{b}{u}.$$

想一想 如果 $c < 0$ 又如何？這時 $c/b^{3/2} \leq -1$，沒有任何實數 T 的 cosh 會等於這個值。

把 x 換成 $-x$：方程變成 $x^3 - 3bx = -2c$，其中 $-c > 0$，所以根據剛才處理過的情況，它有根 $u' + b/u'$，其中 $u' = (-c + \sqrt{c^2 - b^3})^{1/3}$，而 $-(u' + b/u')$ 就是原方程的一個根。現在

$$u^3 u'^3 = \left(\sqrt{c^2 - b^3} + c\right)\left(\sqrt{c^2 - b^3} - c\right) = (c^2 - b^3) - c^2 = -b^3,$$

所以，取實立方根，得 $u' = -b/u$ 及 $-(u' + b/u') = u + b/u$。因此無論 c 是正是負，公式都成立。（留意在兩種情況下均有 $u \neq 0$，因為當 $b > 0$ 時 $\sqrt{c^2 - b^3} < |c|$。）

另外兩個根

記 $r = u + \frac{b}{u}$。由於 r 是一個根，$x - r$ 是因式，在

$$x^3 - 3bx - 2c = (x - r)(x^2 + rx + q) = x^3 + (q - r^2)x - rq$$

中比較係數（comparing coefficients）（x^2 項自動抵消），必有 $q = r^2 - 3b$。展開得

$$q = u^2 + 2b + \frac{b^2}{u^2} - 3b = u^2 + \frac{b^2}{u^2} - b,$$

所以另外兩個根滿足

$$x^2 + \left(u + \frac{b}{u}\right)x + u^2 + \frac{b^2}{u^2} - b = 0,$$

這正是所要證明的。（作為一致性的檢查，常數項亦吻合：$rq = u^3 + b^3/u^3$，而有理化後 $b^3/u^3 = b^3/(c+\sqrt{c^2-b^3}) = c - \sqrt{c^2-b^3}$，所以 $rq = 2c$。）

要解這個二次方程，記 $v = b/u$，於是 $uv = b$，而二次方程成為

$$x^2 + (u+v)x + (u^2 - uv + v^2) = 0.$$

現在 $\omega^3 = 1$ 且 $1 + \omega + \omega^2 = 0$。試取 $x_1 = \omega u + \omega^2 v$ 和 $x_2 = \omega^2 u + \omega v$：

$$x_1 + x_2 = (\omega + \omega^2)(u+v) = -(u+v),$$
$$x_1 x_2 = \omega^3 u^2 + \omega^2 uv + \omega^4 uv + \omega^3 v^2 = u^2 + (\omega^2+\omega)uv + v^2 = u^2 - uv + v^2.$$

這正是該二次方程的兩根之和與兩根之積，所以它的根是

$$x = \omega u + \omega^2 \frac{b}{u} \qquad 及 \qquad x = \omega^2 u + \omega \frac{b}{u}.$$

由於 $\omega = -\frac{1}{2} + \frac{\sqrt{3}}{2}\mathrm{i}$ 及 $\omega^2 = -\frac{1}{2} - \frac{\sqrt{3}}{2}\mathrm{i}$，這兩個根亦可寫成 $-\frac{1}{2}\left(u + \frac{b}{u}\right) \pm \frac{\sqrt{3}}{2}\left(u - \frac{b}{u}\right)\mathrm{i}$。

解 $x^3 - 6x = 6$

這是 $x^3 - 3bx = 2c$，其中 $b = 2$，$c = 3$，而 $c^2 = 9 \geq 8 = b^3 > 0$，所以上述結果適用。這裏

$$u = \left(3 + \sqrt{9-8}\right)^{1/3} = 4^{1/3} = 2^{2/3}, \qquad \frac{b}{u} = \frac{2}{2^{2/3}} = 2^{1/3}.$$

所以各根為

$$x = 2^{2/3} + 2^{1/3}, \qquad x = -\tfrac{1}{2}\left(2^{2/3} + 2^{1/3}\right) \pm \tfrac{\sqrt{3}}{2}\left(2^{2/3} - 2^{1/3}\right)\mathrm{i},$$

其中那對複數根（complex roots）就是 $\omega\, 2^{2/3} + \omega^2\, 2^{1/3}$ 及其共軛（conjugate）。

驗算

- 實根是 $2^{2/3} + 2^{1/3} \approx 2.847322$，而 $2.847322^3 - 6(2.847322) \approx 23.083930 - 17.083932 \approx 6.0000$。
- 三個根之和應為 0（沒有 x^2 項）：$(u+v) - (u+v) = 0$。它們的積應為 6：$r(u^2 - uv + v^2) = u^3 + v^3 = 4 + 2 = 6$。
- 若 $c^2 = b^3$，則 $u = \sqrt{b} = b/u$，兩個複數根都變成 $-\sqrt{b}$，即一個重根（repeated root），這可由 $x^3 - 3bx - 2b^{3/2} = (x - 2\sqrt{b})(x + \sqrt{b})^2$ 印證。

常見錯誤

- 忘記了 $\cosh T \geq 1$，所以雙曲函數的論證只能直接用於 $c > 0$；$c < 0$ 的情況需要代換 $x \to -x$。

- 把複數根寫成 $\omega u + \omega \frac{b}{u}$：兩個立方根的積必須保持等於 b，所以 ω 的冪相乘必須等於 1。

STEP 之外

這就是三次方程的卡爾達諾公式（Cardano's formula），而 $u + b/u$ 這個形式（其中 u^3 是某個二次方程的根），正是十六世紀最初發現它時的模樣。當 $c^2 < b^3$ 時，雖然三個根全是實數，平方根卻是虛數（即所謂「不可約情況」（irreducible case）），這時便由三角形式 $\cos 3\theta = 4\cos^3\theta - 3\cos\theta$ 接手。至於一般的五次方程為何沒有這類公式，那就是伽羅瓦理論（Galois theory）的故事了；Allenby 的 *Rings, Fields and Groups*（環、域與羣）是很好的入門讀物。

詳細解答

STEP III 2005，第 7 題

純數學 • 代換積分法

STEP III 2005，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/05-s3-q7

關鍵思路

題目開首的結果告訴我們：只要被積函數（integrand）外面另有一個單獨的因子 $\frac{1}{x}$，藏在被積函數裏的 x^m 就可以當作一個新變數 $u = x^m$ 來處理。因此，對每一個積分，我們先從分母抽出一個因子 x，看看剩下的是哪一個冪 x^m、函數 f 又是甚麼，然後改為計算容易得多的 u 積分。

解答

一般結果

已知 $\mathrm{F}'(u) = \dfrac{1}{u\,\mathrm{f}(u)}$。用鏈式法則（chain rule）對 $\mathrm{F}(x^m)$ 求導（取 $x > 0$，使 x^m 對任何實數 m 都有意義），

$$\frac{\mathrm{d}}{\mathrm{d}x}\mathrm{F}(x^m) = \mathrm{F}'(x^m)\cdot mx^{m-1} = \frac{mx^{m-1}}{x^m\,\mathrm{f}(x^m)} = \frac{m}{x\,\mathrm{f}(x^m)}.$$

所以 $\mathrm{F}(x^m)$ 是 $\dfrac{m}{x\,\mathrm{f}(x^m)}$ 的一個原函數（antiderivative），這正是所要證明的結果：

$$\int \frac{m}{x\,\mathrm{f}(x^m)}\,\mathrm{d}x = \mathrm{F}(x^m) + c.$$

換句話說，這就是把代換（substitution）$u = x^m$、$\mathrm{d}u = mx^{m-1}\,\mathrm{d}x$ 倒過來看。條件 $m \neq 0$ 在下文十分重要：我們將會除以 m。

實際運用時，我們把它寫成

$$\int \frac{1}{x\,\mathrm{f}(x^m)}\,\mathrm{d}x = \frac{1}{m}\mathrm{F}(x^m) + c. \qquad (*)$$

第 (i) 部分

抽出一個因子 x，

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

這符合 $(*)$ 的形式，其中 $m = n - 1$，$\mathrm{f}(u) = u - 1$。我們需要 $n \neq 1$；當 $n = 1$ 時，分母 $x^n - x$ 恆等於零，根本沒有積分可求。

接着要求 F，即 $\dfrac{1}{u(u-1)}$ 的一個原函數。由部分分式（partial fractions）得

$$\frac{1}{u(u-1)} = \frac{1}{u-1} - \frac{1}{u}, \qquad \text{所以} \qquad \mathrm{F}(u) = \ln|u-1| - \ln|u| = \ln\left|\frac{u-1}{u}\right|.$$

把 $u = x^{n-1}$ 代入 $(*)$：

$$\boxed{\int \frac{1}{x^n - x}\,\mathrm{d}x = \frac{1}{n-1}\ln\left|\frac{x^{n-1}-1}{x^{n-1}}\right| + c, \qquad n \neq 1.}$$

有了絕對值（modulus）符號，這條公式在被積函數有定義的任何區間上都成立（例如在 $0 < x < 1$ 上和在 $x > 1$ 上都成立；當 n 是整數時，在 $x < 0$ 上也成立）。

第 (ii) 部分

這裏把 x 從平方根中抽出來時要小心一點。我們只考慮 $x > 0$，這時 x^n 對每個 n 都有定義，而且 $\sqrt{x^2} = x$。於是

$$\sqrt{x^n + x^2} = \sqrt{x^2\,(x^{n-2} + 1)} = x\sqrt{1 + x^{n-2}},$$

所以被積函數是

$$\frac{1}{x\sqrt{1 + x^{n-2}}},$$

這符合 $(*)$ 的形式，其中 $m = n - 2$，$\mathrm{f}(u) = \sqrt{1+u}$。

? **想一想** 當 $n = 2$ 時會怎樣？

這時 $m = 0$，一般結果並不適用。不過這個積分可以直接計算，十分容易：被積函數是 $\dfrac{1}{\sqrt{2x^2}} = \dfrac{1}{\sqrt{2}\,x}$，所以

$$\int \frac{1}{\sqrt{x^2 + x^2}}\,\mathrm{d}x = \frac{1}{\sqrt{2}}\ln x + c \qquad (n = 2).$$

當 $n \neq 2$ 時，我們需要 $\mathrm{F}(u) = \displaystyle\int \frac{1}{u\sqrt{1+u}}\,\mathrm{d}u$，其中 $u = x^{n-2} > 0$。代入 $t = \sqrt{1+u}$，則 $u = t^2 - 1$，$\mathrm{d}u = 2t\,\mathrm{d}t$，且 $t > 1$：

$$\int \frac{1}{u\sqrt{1+u}}\,\mathrm{d}u = \int \frac{2t}{(t^2-1)\,t}\,\mathrm{d}t = \int\left(\frac{1}{t-1} - \frac{1}{t+1}\right)\mathrm{d}t = \ln\left(\frac{t-1}{t+1}\right),$$

由於 $t > 1$，這裏不需要絕對值符號。這個結果可以化簡得很整齊：把分子和分母同乘以 $t-1$，並利用 $t^2-1=u$，

$$\frac{t-1}{t+1}=\frac{(t-1)^2}{t^2-1}=\frac{\left(\sqrt{1+u}-1\right)^2}{u}, \qquad \text{所以} \qquad \mathrm{F}(u)=2\ln\left(\sqrt{1+u}-1\right)-\ln u.$$

現在把 $u=x^{n-2}$ 代入 $(*)$，其中 $m=n-2$。$-\ln u$ 一項變成 $-(n-2)\ln x$，再除以 $m=n-2$ 便成為 $-\ln x$：

$$\boxed{\int \frac{1}{\sqrt{x^n+x^2}}\,\mathrm{d}x=\frac{2}{n-2}\ln\left(\sqrt{1+x^{n-2}}-1\right)-\ln x+c, \qquad n\neq 2,\ x>0.}$$

答案也可以等價地寫成 $\dfrac{1}{n-2}\ln\left(\dfrac{\sqrt{1+x^{n-2}}-1}{\sqrt{1+x^{n-2}}+1}\right)+c$；兩種形式都可以。

驗算

- 把答案求導，看看能否還原。對於 (i)，取 $n=2$，公式給出 $\ln|x-1|-\ln|x|$，其導數是 $\frac{1}{x-1}-\frac{1}{x}=\frac{1}{x^2-x}$，正如所料。
- 對於 (i)，取 $n=3$，公式給出 $\displaystyle\int_2^3 \frac{\mathrm{d}x}{x^3-x}=\frac{1}{2}\ln\frac{32}{27}\approx 0.084950$，與辛普森法則（Simpson's rule）的結果吻合。
- 對於 (ii)，取 $n=3$，公式給出 $\displaystyle\int_1^3 \frac{\mathrm{d}x}{\sqrt{x^3+x^2}}=\ln\frac{3+2\sqrt{2}}{3}\approx 0.664135$，同樣與數值積分（numerical integration）的結果吻合。我們還用許多不同的 n 值檢驗過這兩條公式，包括負數和非整數的值。

常見錯誤

- 漏掉因子 $\frac{1}{m}$：一般結果的分子有 m，而我們的被積函數沒有，所以必須除以 m。
- 不加說明就寫 $\sqrt{x^2}=x$。這只在 $x\geq 0$ 時成立；當 $x<0$ 時，它等於 $-x$（或寫成 $|x|$）。
- 遺漏例外的值：(i) 中的 $n=1$ 和 (ii) 中的 $n=2$，這時 $m=0$。

STEP 之外

題目開首的結果，其實是從另一面看代換法則：代換不過是把鏈式法則倒過來用，而在大學第一門分析（analysis）課程中，代換法則正是這樣證明的，當中會對所涉及的函數仔細列明假設。像 (ii) 這類含有多項式平方根的積分，會引出一個問題：到底哪些積分

可以用初等函數（elementary functions）表示？例如 $\int \mathrm{d}x/\sqrt{1+x^4}$ 就不可以。如想看微積分基本定理（fundamental theorem of calculus）和代換法則的嚴謹論述，可參閱 Abbott 的 *Understanding Analysis*。

詳細解答

STEP III 2005，第 8 題

純數學 • 複數與圓

> **STEP III 2005，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/05-s3-q8

關鍵思路

阿根圖（Argand diagram）上的每一個長度，都可以用 $|z|^2 = zz^*$ 化為代數式。一旦得到第一條恆等式，直角就變成畢氏定理（Pythagoras' theorem），而切線（tangent）就變成切線與半徑之間的直角。於是「OA 與圓相切」就化為單獨一條方程，由此定出半徑的平方為 $cc^* - aa^*$。此後，「P 在圓上」與「P' 在圓上」都歸結為同一條方程。

解答

恆等式與直角

由於 $|z|^2 = zz^*$，而且取共軛（conjugate）與減法可以交換次序，

$$\begin{aligned}|a-c|^2 &= (a-c)(a-c)^* = (a-c)(a^*-c^*)\\ &= aa^* + cc^* - ac^* - ca^*,\end{aligned}$$

這正是所要證明的。

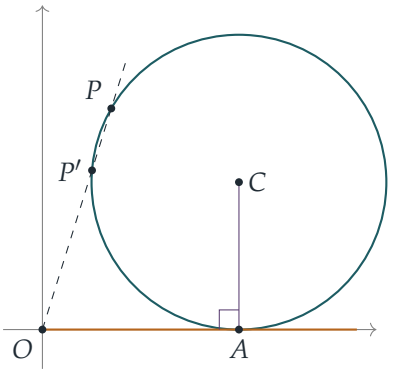
三角形 OAC 的邊長為 $OA = |a|$、$OC = |c|$ 和 $AC = |a-c|$，由於 a, c 非零且互不相同，三者皆不為零。根據畢氏定理及其逆定理（converse），在 A 的角是直角，當且僅當 $OC^2 = OA^2 + AC^2$，即

$$cc^* = aa^* + \left(aa^* + cc^* - ac^* - ca^*\right) \iff 2aa^* = ac^* + ca^*.$$

這裏每一步都可逆，所以「當且僅當」成立。

（畢氏定理的逆定理同時排除了退化的「三角形」。假如 O、A、C 共線（collinear），則三個長度之中有一個等於另外兩個之和；當三個長度皆為正數時，這不可能與 $OC^2 = OA^2 + AC^2$ 同時成立。）

相切的圓



切線在切點處與半徑垂直。因此 OA 在 A 與圓相切，即表示角 OAC 是直角，而根據第一部分，

$$2aa^* = ac^* + ca^*. \tag{$*$}$$

半徑 r 滿足 $r^2 = |a-c|^2 = aa^* + cc^* - (ac^* + ca^*) = cc^* - aa^*$，當中用到 $(*)$。(這其實又是畢氏定理：$r^2 = OC^2 - OA^2$。)

設 $k = bb^* > 0$ 及 $s = abc^* + a^*b^*c$。則

$$|ab-c|^2 = (ab-c)(a^*b^* - c^*) = aa^*k - s + cc^*,$$
$$\left|\frac{a}{b^*} - c\right|^2 = \left(\frac{a}{b^*} - c\right)\left(\frac{a^*}{b} - c^*\right) = \frac{aa^*}{k} - \frac{ac^*}{b^*} - \frac{a^*c}{b} + cc^*.$$

把第二式乘以 $k = bb^*$，得

$$k\left|\frac{a}{b^*} - c\right|^2 = aa^* - s + k\,cc^*. \tag{$**$}$$

P **在圓上**。$|ab-c|^2 = cc^* - aa^*$ 成立，當且僅當

$$aa^*k - s + cc^* = cc^* - aa^* \iff aa^*(k+1) = s.$$

P' **在圓上**。$|a/b^* - c|^2 = cc^* - aa^*$ 成立，當且僅當（乘以 $k \neq 0$ 並利用 $(**)$）

$$aa^* - s + k\,cc^* = k\,cc^* - k\,aa^* \iff aa^*(k+1) = s.$$

兩個條件完全相同，所以 P 在圓上當且僅當 P' 在圓上。

逆命題

現在不再假設相切。設圓的圓心為 C 並通過 A，於是 $r^2 = |a-c|^2 = aa^* + cc^* - ac^* - ca^*$。再設 ab 和 a/b^* 都在圓上，其中 $k = bb^* \neq 1$。利用同樣的展開式，

$$aa^*k - s + cc^* = r^2,$$
$$aa^* - s + k\,cc^* = kr^2 \qquad (由\ (**))\ .$$

從第一條方程減去第二條，便可消去 s：

$$(k-1)\,aa^* - (k-1)\,cc^* = (1-k)\,r^2.$$

想一想 為甚麼題目堅持要 $bb^* \neq 1$？

若 $bb^* = 1$，則 $b^* = 1/b$，於是 $a/b^* = ab$，兩點重合。「兩點都在圓上」這時只是單獨一個條件，當然不足以推出相切。在我們的代數運算中，這正正就是除以 $k-1 \neq 0$ 的那一步：

$$r^2 = cc^* - aa^*.$$

與 $r^2 = aa^* + cc^* - ac^* - ca^*$ 比較，得 $2aa^* = ac^* + ca^*$。根據第一部分，角 OAC 是直角。因此 OA 在圓上的點 A 處與半徑 CA 垂直，即 OA 是圓的切線。

驗算

- 這個結果背後有一個簡潔的幾何原因。由於 $\dfrac{a}{b^*} = \dfrac{ab}{bb^*}$，點 P' 就是把 P 按正實數 $1/|b|^2$ 縮放而得。所以 O、P、P' 共線，且 $OP \cdot OP' = |a|^2 = OA^2$，這正是切割線定理（tangent–secant theorem）（見圖）。
- 數值上，取 $a = 2$、$c = 2 + 1.5\mathrm{i}$（於是 $r^2 = 2.25 = |c|^2 - |a|^2$），以及圓上的 $P = 0.701 + 2.25\mathrm{i}$。則 $b = P/2$ 及 $P' = a/b^* \approx 0.505 + 1.620\mathrm{i}$，而 $|P' - c|^2 \approx 2.250$。
- 我們用電腦以隨機的 a, b, c 驗證了恆等式和相切的情況。另外又取了 1000 個隨機的非相切圓，並設 $|b| \neq 1$：每一次，只要 P 在圓上，P' 就必定不在圓上。

常見錯誤

- 正確寫出 $(a-c)^* = a^* - c^*$，其後卻把 $|a-c|^2$ 展開成 $(a-c)^2$。模（modulus）的平方必須用到共軛。
- 在逆命題中，一開始就假設相切（從而假設 $r^2 = cc^* - aa^*$）。這正是你要證明的結論。只可從 $r = |a-c|$ 出發。
- 除以 $1 - bb^*$ 時沒有說明它為何不等於零。

STEP 之外

由於 $|a|^2/(ab)^* = aa^*/(a^*b^*) = a/b^*$，點 P' 是 P 在 $z \mapsto |a|^2/z^*$ 之下的像。這個映射就是關於圓 $|z| = |a|$ 的反演（inversion）：它把每一點送到由 O 出發的同一條射線上、距離為 $|a|^2/OP$ 的那一點。本題說明，與反演圓正交（orthogonal）的圓會被映射到自身。反演以及由它生成的莫比烏斯變換（Möbius transformations），是大學一年級複分析（complex analysis）和幾何的核心內容。Needham 的 *Visual Complex Analysis* 用了整整一章的篇幅作出精彩的論述。

詳細解答

STEP III 2005，第 9 題

力學 • 碰撞與恢復係數

STEP III 2005，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：

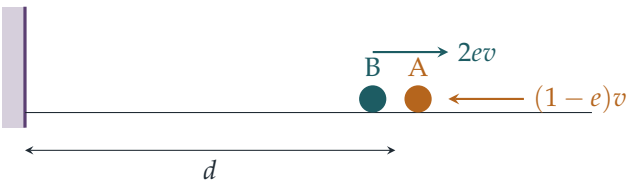
step.maths.org/questions/05-s3-q9

關鍵思路

這兩個質量看似古怪，其實是特意選定的，好讓第一次碰撞得出非常整齊的答案。求出第二次碰撞之後，把它發生前一刻的速度與第一次碰撞前一刻的速度比較：兩者相同，只是各自乘以 e。碰撞定律對速度而言是線性（linear）的，所以整個過程會以較小的尺度重演，其後的每一次碰撞也就一併得出。

解答

我們把 A 和 B 視為光滑水平直線上的質點（particle），每次碰撞均應用牛頓實驗定律（Newton's experimental law）。取朝向牆壁的方向為正，距離則由牆壁量起。



（圖中所示為第一次碰撞前的一刻，兩個質點畫得稍為分開。）

第一次碰撞

設碰撞後 A 和 B 的速度（朝向牆壁）分別為 v_A 和 v_B。碰撞前，A 的速度為 $(1-e)v$，B 的速度為 $-2ev$。

由動量守恆（conservation of momentum）得

$$\begin{aligned}4em\,v_A + (1-e)^2 m\,v_B &= 4em(1-e)v - (1-e)^2 m\cdot 2ev \\ &= 2e(1-e)v\big[2-(1-e)\big]m = 2e(1-e)(1+e)v\,m.\end{aligned}$$

接近速率（speed of approach）為 $(1-e)v + 2ev = (1+e)v$，故由牛頓定律得

$$v_B - v_A = e(1+e)v.$$

把 $v_B = v_A + e(1+e)v$ 代入動量方程，再除以 m：

$$\big[4e+(1-e)^2\big]v_A = 2e(1-e)(1+e)v - (1-e)^2 e(1+e)v = e(1-e)(1+e)v\big[2-(1-e)\big].$$

由於 $4e+(1-e)^2 = (1+e)^2$，右方等於 $e(1-e)(1+e)^2 v$，因此

$$v_A = e(1-e)v, \qquad v_B = e(1-e)v + e(1+e)v = 2ev.$$

兩者皆為正數：碰撞後，**A 以速率 $e(1-e)v$ 朝向牆壁運動，B 以速率 $2ev$ 朝向牆壁運動**。

第二次碰撞

B 較快，所以會拋離 A，並在時間 $t_1 = d/(2ev)$ 之後到達牆壁。在這段時間內，A 移動了 $e(1-e)v\cdot\frac{d}{2ev} = \frac{1}{2}(1-e)d$，所以 A 與牆壁的距離為

$$d - \tfrac{1}{2}(1-e)d = \tfrac{1}{2}(1+e)d$$

。B 以速率 $e\cdot 2ev = 2e^2 v$ 反彈，此時正遠離牆壁，而 A 仍以速率 $e(1-e)v$ 朝向牆壁運動。兩者互相接近，接近速率為 $e(1-e)v + 2e^2 v = e(1+e)v > 0$，所以它們必定再次相遇（而且是在 A 到達牆壁之前，因為 B 位於 A 與牆壁之間）。所需時間為

$$t_2 = \frac{\frac{1}{2}(1+e)d}{e(1+e)v} = \frac{d}{2ev},$$

其間 B 由牆壁起計移動了 $2e^2 v\cdot\frac{d}{2ev} = ed$。所以**第二次碰撞確會發生，位置距離牆壁 de**，正如題目所求。

其後的碰撞

想一想 怎樣才能不必把所有功夫重做一遍，便推斷出還有更多碰撞？

考察第二次碰撞前一刻的狀態：A 以速率 $e(1-e)v$ 朝向牆壁運動，B 以速率 $2e^2 v = 2e(ev)$ 遠離牆壁。這正好是把 v 換成 ev、把 d 換成 de 之後的初始狀態（質量和 e 均不

變）。上文所做的一切對任何 $v > 0$ 和 $d > 0$ 都成立，所以把 (v, d) 換成 (ev, de) 後，同一論證依然適用：第三次碰撞會發生，位置距離牆壁 $(de)e = de^2$，而它發生前一刻的狀態就是取 (e^2v, de^2) 的初始狀態。由歸納法（induction）可知，對每個 n 都有第 $(n+1)$ 次碰撞，所以碰撞會永遠持續下去。

第 n 次碰撞前一刻的狀態，就是把 v 和 d 分別換成 $e^{n-1}v$ 和 de^{n-1} 的初始狀態，所以

$$\text{第 } n \text{ 次碰撞發生於距離牆壁} de^{n-1} \text{ 之處。}$$

由第二部分可知，由一次碰撞到下一次碰撞所需的時間為 $t_1 + t_2 = d/(ev)$。若由第 n 次碰撞起計，這個時間變成

$$\frac{de^{n-1}}{e \cdot e^{n-1}v} = \frac{d}{ev},$$

與 n 無關。所以相繼兩次碰撞之間的時間全部相等，均為 $d/(ev)$。每一個循環中，距離和速率都按因子 e 縮小，這兩個效應在時間上互相抵銷。

驗算

- 第一次碰撞的動量：取 $e = \frac{1}{2}$、$m = v = 1$，兩個質量分別為 2 和 $\frac{1}{4}$；碰撞前，$2 \cdot \frac{1}{2} - \frac{1}{4} \cdot 1 = \frac{3}{4}$；碰撞後，$2 \cdot \frac{1}{4} + \frac{1}{4} \cdot 1 = \frac{3}{4}$。相對速率由 $\frac{3}{2}$ 變為 $\frac{3}{4}$，即乘以因子 e。
- 我們以分數作精確運算，就 200 組隨機選取的 e、v、d、m 模擬了這個運動，每組計算八次碰撞：每次碰撞的位置都是 de^{n-1}，而相隔的時間全部是 $d/(ev)$。當 $e = \frac{1}{2}$、$v = d = 1$ 時，碰撞發生於時間 $0, 2, 4, 6$，距離為 $1, \frac{1}{2}, \frac{1}{4}, \frac{1}{8}$。

常見錯誤

- 混淆正負號：B 起初是遠離牆壁運動的，所以按我們的約定，它的初速度是負數。
- 忘記牆壁的恢復係數（coefficient of restitution）同樣是 e，所以 B 離開牆壁時的速率是 $2e^2v$，而不是 $2ev$。
- 用蠻力把第二次碰撞重新計算一遍，而沒有留意到當中的縮放關係；這樣做也行得通，但冗長得多，而且容易看漏當中的規律。

STEP 之外

這裏的推論用到線性和自相似性（self-similarity）：若把每個速度都乘以同一個常數，碰撞的方程保持不變，所以一個除尺度外能重現自身的狀態會永遠重複下去。同一想法也是分析彈跳無限多次的球的基礎，更廣泛而言，亦是量綱分析（dimensional analysis）

的基礎。Kleppner 與 Kolenkow 所著的 *An Introduction to Mechanics*（力學入門）正是以這種嚴謹的方式處理碰撞和動量。

詳細解答

STEP III 2005，第 10 題

力學 • 彈性帶與摩擦力

STEP III 2005，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/05-s3-q10

關鍵思路

一切都取決於一個變量：兩圓盤圓心之間的距離 d。橡皮圈的長度由 d 決定，因此它的張力（tension）和所儲存的能量都是 d 的簡單函數。由此出發，能量的考慮（彈性勢能（elastic energy）對比克服摩擦力（friction）所作的功）告訴我們圓盤在哪裏停下，而比較橡皮圈的拉力與極限摩擦力（limiting friction），則告訴我們圓盤停下後會否保持靜止。

解答

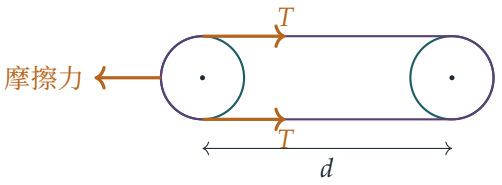
我們把兩個圓盤視為完全相同，橡皮圈視為輕而有彈性，且與圓盤的接觸是光滑（smooth）的，並假設圓盤一旦滑動，摩擦力的大小就是 μmg。由對稱性，兩個圓盤總是沿相反方向移動相等的距離，所以橡皮圈的兩段直線部分始終保持平行。

把橡皮圈表示為 d 的函數。 當兩圓心相距 d 時，橡皮圈繞過每個圓盤半周，再沿着兩段長度為 d 的直線部分伸展。所以它的長度是 $2\pi r + 2d$，伸長量（extension）是 $x = 2d$。根據虎克定律（Hooke's law），取彈性模量（modulus of elasticity）$\lambda = \frac{\pi mg}{12}$ 及自然長度（natural length）$2\pi r$，

$$T = \frac{\lambda x}{2\pi r} = \frac{\pi mg}{12} \cdot \frac{2d}{2\pi r} = \frac{mg\,d}{12r}, \qquad E = \frac{\lambda x^2}{2 \cdot 2\pi r} = \frac{mg\,d^2}{12r}.$$

由於接觸是光滑的，整條橡皮圈上的張力處處相同。每個圓盤都受兩段直線部分拉扯，兩者都平行於圓心連線，並指向另一個圓盤。所以每個圓盤所受的合拉力是

$$P = 2T = \frac{mg\,d}{6r}.$$



上圖為俯視圖，顯示左邊圓盤向另一個圓盤滑動時所受的水平力。它的重量與法向反作用力（normal reaction）都是鉛垂的，並且互相抵消。

釋放時 $d = 6r$，所以 $E_0 = 3mgr$ 而 $P = mg$。接觸時 $d = 2r$，所以 $E = \frac{1}{3}mgr$。

圓盤互相靠近時的能量。 當圓盤從 $d = 6r$ 向內移動時，每個圓盤都要克服摩擦力 μmg 移動 $\frac{1}{2}(6r - d)$ 的距離。所以克服摩擦力所作的總功是 $\mu mg(6r - d)$。因此總動能（kinetic energy）為

$$\begin{aligned} K(d) &= 3mgr - \frac{mg\,d^2}{12r} - \mu mg(6r - d) \\ &= mg\,(6r - d)\left(\frac{6r + d}{12r} - \mu\right), \end{aligned}$$

其中用到 $3r - \frac{d^2}{12r} = \frac{(6r-d)(6r+d)}{12r}$。

第 (i) 部分

*起動。*釋放時每個圓盤所受的拉力是 $P = mg$，而摩擦力最多只能提供 μmg。所以圓盤開始滑動，當且僅當 $mg > \mu mg$，即 $\mu < 1$。

*停下。*假設 $\mu < 1$。對於 $d < 6r$，$K(d)$ 只會在 $\frac{6r+d}{12r} = \mu$ 之處為零，即在

$$d^* = 12\mu r - 6r.$$

當 d 由 $6r$ 開始減少時，第二個因子 $\frac{6r+d}{12r} - \mu$ 由 $1 - \mu > 0$ 開始遞減，所以在 d 到達 d^* 之前 $K > 0$。圓盤在相撞之前停下，當且僅當在接觸之前已到達 d^*，即 $2r < d^* < 6r$：

$$8r < 12\mu r < 12r \iff \tfrac{2}{3} < \mu < 1.$$

若 $\mu \le \frac{2}{3}$，則 $K(2r) \ge 0$，圓盤會先碰上對方。（當 $\mu = \frac{2}{3}$ 時，它們剛好以零速率互相接觸，這不算是「之前」。）

> **想一想** 圓盤在 d^* 停下之後，橡皮圈會否令它們再次移動？

在 d^* 處，拉力是 $P = \frac{mg\,d^*}{6r} = (2\mu - 1)mg$。由於 $\mu < 1$，這小於極限摩擦力 μmg，所以靜摩擦力（static friction）能把圓盤固定，它們保持靜止。結合兩個條件，圓盤開始滑動但在相撞之前停下，當且僅當 $\frac{2}{3} < \mu < 1$。 ■

第 (ii) 部分

若圓盤至少相撞一次，它們首次到達 $d = 2r$ 時必定是向內移動，而靠近階段的公式仍然有效。所以緊接首次碰撞之前的總動能是

$$K(2r) = mg\,(4r)\left(\frac{8r}{12r} - \mu\right) = 4mgr\left(\tfrac{2}{3} - \mu\right) = \tfrac{4}{3}mgr(2 - 3\mu),$$

正如所求。（此值為正，恰好當 $\mu < \frac{2}{3}$，與第 (i) 部分一致。）

第 (iii) 部分

首先，由 $\mu^2 < \frac{4}{9}$ 得 $\mu < \frac{2}{3}$，所以根據第 (i) 和第 (ii) 部分，圓盤確實會相撞。碰撞中損失一半動能，所以它們從 $d = 2r$ 反彈時的總動能是

$$K_1 = \tfrac{2}{3}mgr(2 - 3\mu).$$

現在圓盤互相遠離。橡皮圈的拉力和摩擦力都指向另一個圓盤，所以每個圓盤都在減速。當間距由 $2r$ 增至 d 時，彈性勢能增加了 $\frac{mg(d^2-4r^2)}{12r}$，而克服摩擦力所作的功是 $\mu mg(d - 2r)$，所以

$$K(d) = K_1 - \frac{mg(d^2 - 4r^2)}{12r} - \mu mg(d - 2r).$$

這是 d 的嚴格遞減函數，所以它恰好有一個零點，就是圓盤停下時的間距 d_1。設 $d_1 = 2r + u$，並乘以 $\frac{12r}{mg}$：

$$u^2 + 4r(1 + 3\mu)\,u - 8r^2(2 - 3\mu) = 0.$$

取正根，並化簡判別式（discriminant），$(1 + 3\mu)^2 + 2(2 - 3\mu) = 9\mu^2 + 5$，得

$$u = 2r\left(\sqrt{9\mu^2 + 5} - (1 + 3\mu)\right), \qquad d_1 = 2r\left(\sqrt{9\mu^2 + 5} - 3\mu\right).$$

（可以驗證 $2r < d_1 \le 6r$，所以橡皮圈保持繃緊，其能量公式仍然適用。）

? **想一想** 圓盤在 d_1 處瞬間靜止。甚麼因素決定它們會否留在那裏？

若橡皮圈的拉力不超過極限摩擦力，它們便保持靜止：

$$\frac{mg\,d_1}{6r} \le \mu mg \iff \sqrt{9\mu^2 + 5} - 3\mu \le 3\mu \iff 9\mu^2 + 5 \le 36\mu^2 \iff \mu^2 \ge \tfrac{5}{27},$$

其中兩邊皆為正數，所以可以取平方。因此若 $\frac{5}{27} < \mu^2 < \frac{4}{9}$，圓盤會相撞、反彈、單調地減速至在 d_1 處靜止，然後被靜摩擦力永久固定在那裏。它們在首次碰撞之後恰好停下一次。 ■

驗算

- 當 $\mu = \frac{2}{3}$ 時，第 (i) 部分的停止點 $d^* = 12\mu r - 6r$ 和第 (ii) 部分的動能都給出「以零速率到達 $d = 2r$」，所以兩部分在邊界上吻合。
- 取 $\mu = 0.5$（在第 (iii) 部分的範圍之內）：$K(2r) = \frac{2}{3}mgr$，$K_1 = \frac{1}{3}mgr$，而 $d_1 = 2r(\sqrt{7.25} - 1.5) \approx 2.385r$。該處的拉力 $\approx 0.397mg < 0.5mg$，所以圓盤留在原處。以逐步數值模擬這個運動，重現了以上所有結果，亦重現了多個 $\mu \in (\frac{2}{3}, 1)$ 所對應的停止點 $d^* = 12\mu r - 6r$。對於 $\mu^2 < \frac{5}{27}$（例如 $\mu = 0.42$），模擬顯示圓盤停下後再次移動，與預測相符。

常見錯誤

- 把伸長量當作 d 而不是 $2d$：兩段直線部分都被拉長。
- 每個圓盤只計算一個張力：橡皮圈以兩股拉着每個圓盤，合共 $2T$。
- 忘記檢查圓盤停下後是否保持靜止。「速率為零」與「永久靜止」是兩個不同的條件。

STEP 之外

摩擦力令這成為一個分段系統。每當速度改變正負號，運動方程就會改變；而當外加的力低於摩擦力的臨界值時，便會出現「卡住」的狀態。這類系統（黏滑運動，stick–slip motion）在大學一年級的動力學課程中會以相平面（phase-plane）方法研究。當 μ 很小時，帶有能量損失的反覆碰撞會在有限時間內產生無限多次反彈，這值得細想。關於能量方法和摩擦力的更多內容，可參閱 Kleppner 與 Kolenkow 的 *An Introduction to Mechanics*（力學入門教科書）。

詳細解答

STEP III 2005，第 11 題

力學 • 剛體平衡

> **STEP III 2005，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/05-s3-q11

關鍵思路

轉軸與各桿組成一個繞固定水平軸轉動的剛體（rigid body），所以只需一個角 θ 便能描述整個系統，而能量自然是最合適的工具。勢能（potential energy）是三個餘弦之和，它們的角度相差 $\frac{2}{3}\pi$；把它們合併，便得到單一的表達式 $A\cos\theta + B\sin\theta$，這其實是一個喬裝了的餘弦波。一旦寫成這個形式，平衡位置、它們的穩定性，以及越過最高點所需的能量，全都可以直接讀出。

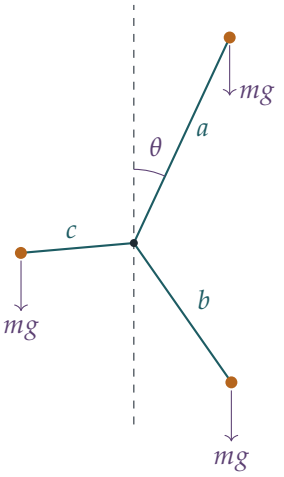
解答

能量

我們把各粒子視為輕質剛桿上的質點，並假設軸承光滑，因此沒有能量損失。如圖所示，所有角度均由向上的鉛垂線起按順時針方向量度。於是長度為 a、b、c 的桿與向上的鉛垂線所成的角分別是 θ、$\theta + \frac{2}{3}\pi$ 和 $\theta + \frac{4}{3}\pi$，所以各粒子在轉軸之上的高度為

$$a\cos\theta, \qquad b\cos\left(\theta + \tfrac{2}{3}\pi\right), \qquad c\cos\left(\theta + \tfrac{4}{3}\pi\right)$$

。



利用 $\cos\frac{2}{3}\pi = \cos\frac{4}{3}\pi = -\frac{1}{2}$、$\sin\frac{2}{3}\pi = \frac{\sqrt{3}}{2}$ 和 $\sin\frac{4}{3}\pi = -\frac{\sqrt{3}}{2}$，可得

$$\cos\left(\theta + \tfrac{2}{3}\pi\right) = -\tfrac{1}{2}\cos\theta - \tfrac{\sqrt{3}}{2}\sin\theta, \qquad \cos\left(\theta + \tfrac{4}{3}\pi\right) = -\tfrac{1}{2}\cos\theta + \tfrac{\sqrt{3}}{2}\sin\theta.$$

取轉軸為勢能的零點，

$$\begin{aligned} V &= mg\left[a\cos\theta + b\cos\left(\theta + \tfrac{2}{3}\pi\right) + c\cos\left(\theta + \tfrac{4}{3}\pi\right)\right] \\ &= mg\left[\tfrac{1}{2}(2a - b - c)\cos\theta - \tfrac{\sqrt{3}}{2}(b - c)\sin\theta\right]. \end{aligned}$$

每個粒子都以角速度（angular velocity）$\dot\theta$ 繞轉軸作圓周運動，所以它們的速率分別是 $a\dot\theta$、$b\dot\theta$、$c\dot\theta$。各桿為輕桿，因此總能量為

$$E = \tfrac{1}{2}m\left(a^2 + b^2 + c^2\right)\dot\theta^2 + \tfrac{1}{2}mg\left[(2a - b - c)\cos\theta - \sqrt{3}\,(b - c)\sin\theta\right].$$

平衡

系統在 $\mathrm{d}V/\mathrm{d}\theta = 0$ 之處處於平衡（equilibrium），即

$$-\tfrac{1}{2}mg\left[(2a - b - c)\sin\theta + \sqrt{3}\,(b - c)\cos\theta\right] = 0.$$

由於 $a > b$ 且 $a > c$，我們有 $2a - b - c > 0$。所以在平衡位置 $\cos\theta \neq 0$：若 $\cos\theta = 0$，則 $\sin\theta = \pm 1$，方括號內的值便是 $\pm(2a - b - c) \neq 0$。兩邊除以 $\cos\theta$，得

$$\tan\theta = -\frac{(b - c)\sqrt{3}}{2a - b - c},$$

這正是所求。這個論證也可以倒過來：任何滿足此方程的 θ 都使 $\mathrm{d}V/\mathrm{d}\theta = 0$。所以平衡位置恰好就是此方程的解。

恰好兩個平衡位置，以及哪一個穩定

正切函數的週期為 π，並且在每個有定義而長度為 π 的區間上，每個實數值都恰好取一次。所以方程在 $0 \le \theta < 2\pi$ 內恰好有兩個解，兩者相差 π。這就是兩個平衡位置。

要判斷穩定性（stability），可把 V 寫成單一餘弦。設

$$A = \tfrac{1}{2}(2a - b - c), \qquad B = -\tfrac{\sqrt{3}}{2}(b - c),$$

則 $V = mg(A\cos\theta + B\sin\theta)$。於是

$$\begin{aligned} 4(A^2 + B^2) &= (2a - b - c)^2 + 3(b - c)^2 \\ &= 4a^2 + 4b^2 + 4c^2 - 4ab - 4bc - 4ca \\ &= 2\left[(a - b)^2 + (b - c)^2 + (c - a)^2\right] = 4R^2, \end{aligned}$$

所以 $A^2 + B^2 = R^2$，而由於各長度互不相同，$R > 0$。選取 α 使 $\cos\alpha = A/R$ 及 $\sin\alpha = B/R$。由於 $A > 0$ 且 $B < 0$，故 $-\frac{1}{2}\pi < \alpha < 0$。於是

$$V = mgR\cos(\theta - \alpha).$$

兩個平衡位置是 $\theta = \alpha$，該處 $V = mgR$ 為極大值；以及 $\theta = \alpha + \pi$，該處 $V = -mgR$ 為極小值。

> **想一想** 為甚麼 V 取極小值就代表穩定？

這裏 $V''(\theta) = -mgR\cos(\theta - \alpha)$。在 $\theta = \alpha + \pi$ 時，它等於 $mgR > 0$，所以微小的位移會令勢能上升。總能量既然固定，系統便不能走得太遠，只會在這個位置附近振動。在 $\theta = \alpha$ 時，$V'' = -mgR < 0$：微小的位移令 V 下降，動能（kinetic energy）增加，系統便會遠離該位置。所以穩定平衡位置是 $\theta = \alpha + \pi$，即滿足 $\frac{1}{2}\pi < \theta < \pi$ 的那一個，該處 $\cos\theta < 0$ 而 $\sin\theta > 0$。在這個位置，最長的桿 a 指向水平線以下，與向下的鉛垂線成角 $|\alpha|$。另一個平衡位置滿足 $-\frac{1}{2}\pi < \theta < 0$，桿 a 在水平線以上，是不穩定的。

完整轉圈

由於軸承光滑，而重力是保守力（conservative force），所以能量守恆。記 $I = m(a^2 + b^2 + c^2)$，則

$$\tfrac{1}{2}I\dot{\theta}^2 + mgR\cos(\theta - \alpha) = E.$$

要完整地轉圈，系統必須越過每一個角度，尤其要越過 $\theta = \alpha$，該處 V 取其最大值 mgR。由於處處都有 $\frac{1}{2}I\dot{\theta}^2 \ge 0$，這就需要 $E \ge mgR$。設 ω 為系統經過穩定位置 $\theta = \alpha + \pi$ 時的角速度，該處 $V = -mgR$。於是 $E = \frac{1}{2}I\omega^2 - mgR$，條件變成

$$\tfrac{1}{2}I\omega^2 - mgR \ge mgR \quad\Longleftrightarrow\quad \omega^2 \ge \frac{4mgR}{I} = \frac{4gR}{a^2 + b^2 + c^2}.$$

所以系統經過穩定平衡位置時，角速度必須至少為 $\sqrt{4gR/(a^2+b^2+c^2)}$，這正是所求。

（在等號成立的臨界情況下，系統會緩緩趨近不穩定平衡位置，要經過無限長的時間才到達。嚴格來說，要完整轉一圈，ω 必須嚴格大於這個值；題目只要求必要的下界。）

驗算

- 若 $a=b=c$，則 $R=0$ 且 $V\equiv 0$：三個相等的質量在任何位置都平衡（它們的質心在軸上），而且無論轉得多慢都能轉過一圈。這與下界趨向零吻合。
- 若 $b=c$，正切條件給出 $\tan\theta=0$。根據對稱性，長桿指向正上方（不穩定）或正下方（穩定）。
- 以隨機長度作數值檢驗，可確認 V 的形式、恆等式 $A^2+B^2=R^2$，以及恰好有兩個平衡位置，其穩定性亦如上所述。取 $a=3$、$b=2$、$c=0.5$ 及 $g=9.81$，臨界角速度為 $2.5406\ \mathrm{rad\,s^{-1}}$。對運動方程逐步作數值計算可見，比這個值低 2% 時系統會擺回來，而高 2% 時則會轉過一整圈。

常見錯誤

- 把位於 $\theta-\frac{2}{3}\pi$ 和 $\theta+\frac{2}{3}\pi$ 的兩根桿與圖中的位置對調，會令 $\sin\theta$ 項的符號反轉，正切條件中的符號亦隨之反轉。
- 斷言「有兩個解」，卻沒有指出 $\cos\theta=0$ 不可能是平衡位置，因而除以 $\cos\theta$ 不會漏掉任何解。
- 把動能寫成 $\frac{1}{2}m(a+b+c)^2\dot\theta^2$ 或類似的式子。每個粒子各有自己的速率 $r\dot\theta$，所以轉動慣量（moment of inertia）是 $m(a^2+b^2+c^2)$。

STEP 之外

這個系統是一個複擺（compound pendulum），即繞固定軸轉動的剛體。它的能量只取決於轉動慣量和質心（centre of mass）的位置，而這裏質心與軸的距離為 $R/3$。把 $\frac{1}{2}I\dot\theta^2+V(\theta)=E$ 看成 $(\theta,\dot\theta)$ 相平面（phase plane）上的一條曲線，振動在分界線（separatrix）之內，轉動在分界線之外，這個圖像是大學一年級動力學的核心概念。Kleppner 與 Kolenkow 所著的 *An Introduction to Mechanics*（力學導論）涵蓋了剛體轉動，以及處理這類問題的能量方法。

詳細解答

STEP III 2005，第 12 題

概率與統計 • 無偏估計量的組合

> **STEP III 2005，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/05-s3-q12

關鍵思路

整道題都建基於獨立（independent）隨機變量（random variable）的線性組合（linear combination）的兩個事實：期望值（expectation）可以相加，而方差（variance）則要乘上係數的平方後相加。要求某個期望值無論真實的單圈時間是多少都符合要求，就等於要比較係數，於是每個條件都化成一條關於常數的方程。這樣，X 就是比賽時間的一個合理估計，而 Y 則利用兩種計時方法之間的差異來估計未知的 σ。

解答

設真實的單圈時間為 $\mu_1, \ldots, \mu_4$，真實的比賽時間為 $\mu = \mu_1 + \mu_2 + \mu_3 + \mu_4$。記

$$S = T_1 + T_2 + T_3 + T_4.$$

那麼 $\mathrm{E}(S) = \mu_1 + \cdots + \mu_4 = \mu$，而且由於 $T_1, \ldots, T_4$ 互相獨立（各計時員互相獨立），$\mathrm{Var}(S) = 4\sigma^2$。另外 $\mathrm{E}(T) = \mu$，$\mathrm{Var}(T) = \sigma^2$，而 T 與 S 獨立，因為 S 只取決於另外四位計時員。

求 X

對於 $X = aT + bS$，

$$\mathrm{E}(X) = a\mu + b\mu = (a+b)\mu.$$

無論真實的單圈時間是多少，這個值都必須等於 μ，特別是對某個 $\mu \neq 0$ 也要成立；因此

$$a + b = 1.$$

由 T 與 S 的獨立性，

$$\operatorname{Var}(X) = a^2 \operatorname{Var}(T) + b^2 \operatorname{Var}(S) = \sigma^2\left(a^2 + 4b^2\right) = \sigma^2\left(a^2 + 4(1-a)^2\right).$$

二次式 $a^2 + 4(1-a)^2 = 5a^2 - 8a + 4$ 的首項係數為正，所以它在導數 $10a - 8$ 等於零之處取得最小值，即 $a = \frac{4}{5}$。（用配方法（completing the square）：$5a^2 - 8a + 4 = 5\left(a - \frac{4}{5}\right)^2 + \frac{4}{5}$。）因此 $a = \frac{4}{5}$，$b = \frac{1}{5}$，而

$$X = \tfrac{4}{5}T + \tfrac{1}{5}(T_1 + T_2 + T_3 + T_4), \qquad \operatorname{Var}(X) = \tfrac{4}{5}\sigma^2.$$

這個結果合情合理：T 的可靠程度是 S 的四倍（方差是 σ^2 對 $4\sigma^2$），所以它得到四倍的權重。

求 Y

對於 $Y = cT + dS$，我們利用 $\mathrm{E}(Y^2) = \operatorname{Var}(Y) + \big(\mathrm{E}(Y)\big)^2$：

$$\mathrm{E}(Y^2) = \sigma^2(c^2 + 4d^2) + (c+d)^2\mu^2.$$

想一想 無論真實的單圈時間是多少，這個式子怎樣才能都等於 σ^2？

常數 c、d 不可以依賴 μ，但 μ 可以是任何正數。這個式子是 $\sigma^2(c^2 + 4d^2)$ 加上 μ^2 的某個倍數，所以要它在 μ 變動時仍然等於 σ^2，μ^2 的係數就必須為零：$(c+d)^2 = 0$，所以 $d = -c$。接着我們需要 $\sigma^2(c^2 + 4c^2) = \sigma^2$，即 $c^2 = \frac{1}{5}$。取 $c = 1/\sqrt{5}$，

$$Y = \frac{T - (T_1 + T_2 + T_3 + T_4)}{\sqrt{5}}$$

（取另一個符號會得到 $-Y$，同樣可行，因為要緊的只是 Y^2）。所以 Y 量度的是兩種計時結果之間的差異；它的期望值是 0，而它的平方是 σ^2 的無偏估計（unbiased estimate）。

這一場比賽

當 $T = 220$ 及 $S = 220.5$ 時，

$$X = \tfrac{4}{5}(220) + \tfrac{1}{5}(220.5) = 176 + 44.1 = 220.1,$$
$$Y = \frac{220 - 220.5}{\sqrt{5}} = -\frac{0.5}{\sqrt{5}}, \qquad Y^2 = \frac{0.25}{5} = 0.05.$$

所以我們估計 $\sigma^2 \approx 0.05$，因而

$\operatorname{Var}(X) = \frac{4}{5}\sigma^2 \approx \frac{4}{5}(0.05) = 0.04$，　就標準差（standard deviation）而言，$X \approx 0.2$ 秒.

因此我們對真實比賽時間的估計是 220.1 ± 0.2 秒：真實時間大致落在區間

$$[219.9,\ 220.3] \text{ 秒},$$

之內，這裏在兩邊各取一個標準差。如果想要更大的把握，可以在兩邊各取兩個標準差，得到 [219.7, 220.5] 秒；若誤差服從正態分佈（normal distribution），這大約是一個 95% 的區間。要留意 σ^2 只是由 Y^2 的單一個值估計出來的，所以區間的寬度本身也只是一個粗略的估計。

驗算

- 當 $a = \frac{4}{5}$ 時，$\mathrm{Var}(X) = \frac{4}{5}\sigma^2$ 比 $\mathrm{Var}(T) = \sigma^2$ 和 $\mathrm{Var}(S) = 4\sigma^2$ 都要小，這正是結合兩個獨立量度所應有的效果。
- $X = 220.1$ 介乎兩個量度值之間，而且遠較接近 $T = 220$，即較可靠的那一個。
- 以 200 000 場比賽（$\sigma = 0.7$）進行模擬，所得 X 的平均值與真實時間相差不超過 0.001，$\mathrm{Var}(X) \approx 0.393$，對比 $\frac{4}{5}(0.49) = 0.392$；而 Y^2 的平均值為 0.491，對比 $\sigma^2 = 0.49$。

常見錯誤

- 把方差取作 $\mathrm{Var}(S) = \sigma^2$ 或 $16\sigma^2$：正確的是 $4\sigma^2$，即四個獨立方差之和。
- 忘記了 $\mathrm{E}(Y^2)$ 中的 $(\mathrm{E}\,Y)^2$ 一項，而正是這一項迫使 $d = -c$。
- 把 $Y^2 = 0.05$ 當作 X 的方差，而沒有乘上 $\frac{4}{5}$。

STEP 之外

X 是兩個估計的最小方差無偏線性組合：一般的法則是以每個獨立估計的方差的倒數作為其權重，這是薈萃分析（meta-analysis）中逆方差加權（inverse-variance weighting）的基礎，也是最小二乘法（least squares）的高斯–馬可夫定理（Gauss–Markov theorem）的基礎。Y^2 是 σ^2 的無偏估計量（estimator），就像以 $n - 1$ 作除數的樣本方差（sample variance）一樣。Grimmett 與 Welsh 的 *Probability: An Introduction*（概率論入門）仔細講解了和的期望值與方差，而大學的第一門統計課程則會正式探討估計量與置信區間（confidence interval）。

詳細解答

STEP III 2005，第 13 題

概率與統計 • 幾何分佈

> **STEP III 2005，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/05-s3-q13

關鍵思路

由 $w+1$ 至 n 的牌完全不起作用：它們既不派錢，也不會令遊戲結束。因此在第 (i) 部分，我們大可不理會它們，只看抽出 $0, 1, \ldots, w$ 的那些抽牌；每一次這樣的抽牌，都是從 $w+1$ 個等可能的數字中公平地抽出一個，於是總獎金就是一個幾何隨機變量（geometric random variable）。在第 (ii) 部分，各次抽牌不再獨立（independent），我們把總獎金拆成若干個指示變量（indicator）之和，每張有獎金的牌對應一個，然後利用期望值的線性性質（linearity of expectation）。

解答

若某次抽牌抽出 $0, 1, \ldots, w$ 其中之一，稱它為有關的抽牌；若抽出 $w+1, \ldots, n$ 其中之一，則稱它為無關的抽牌。

第 (i) 部分

為何兩個事件的概率相同。無關的抽牌不會改變任何事情：玩家甚麼也得不到，遊戲繼續進行。因此玩家的總獎金只取決於有關抽牌的序列：每次非零的有關抽牌可得 £1，而第一次抽出 0 的有關抽牌令遊戲結束。所以，玩家恰好贏得 £3，當且僅當在有關的抽牌之中，首三次抽出非零的數字，而第四次抽出 0。這兩者是同一個事件，故概率相同。

（每次抽牌都獨立地以概率 $\frac{w+1}{n+1} > 0$ 為有關的抽牌，所以有關的抽牌以概率 1 有無限多次，當中自然有第一、第二、第三和第四次；遊戲以概率 1 結束。）

想一想 為甚麼不論之前發生過甚麼，每次有關的抽牌抽出 $0,1,\ldots,w$ 中任何一個數的可能性都相同？

由於每次抽出的牌都會放回，各次抽牌互相獨立，而且每次都在 $\{0,1,\ldots,n\}$ 上均勻分佈（uniform）。已知某一次抽牌是有關的抽牌，它就在 $\{0,1,\ldots,w\}$ 上均勻分佈，因為這 $w+1$ 個結果本來就是等可能的。各次抽牌互相獨立，意味着先後各次有關抽牌的數值也互相獨立。所以每次有關的抽牌以概率 $\frac{1}{w+1}$ 抽出 0，以概率 $\frac{w}{w+1}$ 抽出非零的數，並且與其他各次獨立。

因此

$$P(\text{恰好贏得£3}) = \left(\frac{w}{w+1}\right)^3 \cdot \frac{1}{w+1} = \frac{w^3}{(w+1)^4},$$

正如所求。

一般的 r。把 3 換成 r，用完全相同的論證（首 r 次有關的抽牌非零，第 $(r+1)$ 次為零），可得

$$P(\text{恰好贏得£}r) = \frac{w^r}{(w+1)^{r+1}}, \qquad r = 0,1,2,\ldots$$

設 $q = \frac{w}{w+1}$，上式就是 $(1-q)q^r$，即幾何分佈（geometric distribution）；這些概率之和為 1，理應如此。

總獎金的期望值。設 X 為以英鎊計的總獎金。當 $|q|<1$ 時，對 $\sum_{r\geq 0} q^r = \frac{1}{1-q}$ 求導，再乘以 q，得 $\sum_{r\geq 0} rq^r = \frac{q}{(1-q)^2}$。所以

$$E(X) = \sum_{r=0}^{\infty} r(1-q)q^r = (1-q)\cdot\frac{q}{(1-q)^2} = \frac{q}{1-q} = \frac{w/(w+1)}{1/(w+1)} = w.$$

總獎金的期望值是 £w。

第 (ii) 部分

現在抽出的牌不再放回，所以這個遊戲等同於把整疊牌洗成一個均勻隨機的次序，然後逐張翻開，直至 0 出現為止。在這個次序中，牌 $1,\ldots,w$ 之中每有一張排在牌 0 之前，玩家就得到 £1（排在 0 之前的每張牌都會被抽出，排在它之後的則一張也不會）。

對 $k = 1,\ldots,w$，若牌 k 排在牌 0 之前，設 $I_k = 1$，否則設 $I_k = 0$。總獎金為 $Y = I_1 + I_2 + \cdots + I_w$。

想一想 各個 I_k 並不獨立。這有沒有影響？

沒有：不論各變量之間如何相依，和的期望值總是等於期望值之和。所以我們只需求出 $P(I_k = 1)$。在均勻隨機的次序中，把牌 k 與牌 0 的位置對調，就是「k 在 0 之前」的次序與「0 在 k 之前」的次序之間的一個雙射（bijection），所以兩者等可能，$P(I_k = 1) = \frac{1}{2}$。

因此

$$E(Y) = \sum_{k=1}^{w} E(I_k) = \sum_{k=1}^{w} \tfrac{1}{2} = \frac{w}{2},$$

正是第 (i) 部分獎金期望值 w 的一半，正如所求。

驗算

- 取 $w = n$（每張非零的牌都有獎金）而且不放回，牌 0 的位置在 $1, \ldots, n+1$ 上均勻分佈，所以排在它之前的牌數的平均值是 $\frac{n}{2}$，與 $\frac{w}{2}$ 吻合。
- 放回而 $w = 1$ 時，每次有關的抽牌相當於在 0 與 1 之間擲一枚公平硬幣，所以 $P(X = r) = 2^{-(r+1)}$ 而 $E(X) = 1$，與 $w^r/(w+1)^{r+1}$ 及 $E(X) = w$ 吻合。
- 取 $n = 5$、$w = 2$，模擬 2×10^5 局遊戲，得 $P(X = 3) \approx 0.0995$，對比 $\frac{8}{81} \approx 0.0988$；平均獎金為 1.995，對比 2。對細小的牌疊（例如 $n = 5$、$w = 3$）窮舉所有次序，得出不放回時的平均值為 $\frac{3}{2} = \frac{w}{2}$。

常見錯誤

- 以 $\frac{1}{n+1}$ 作為某次抽牌抽出 0 的機會。這是每次抽牌的機會，但其間可以出現任意多次無關的抽牌；以有關的抽牌為條件，才得出簡潔的 $\frac{1}{w+1}$。
- 遺漏 $r = 0$：第一次有關的抽牌可能就是 0，所以有 $\frac{1}{w+1}$ 的概率一無所獲。
- 在第 (ii) 部分試圖求出獎金的完整分佈。這並無必要；用指示變量和線性性質，兩行就可完成。

STEP 之外

不理會無關的抽牌，是*稀疏化*（thinning）的一個例子，即只在有趣的事情發生的時刻觀察一個過程；這個想法會在馬可夫鏈（Markov chain）（所謂「跳躍鏈」（jump chain））和泊松過程（Poisson process）中再次出現。把一個計數寫成指示變量之和，再利用期望值的線性性質（即使各指示變量並不獨立），是大學一年級概率論與組合數學中最有用的技巧之一。這兩個想法在 Blitzstein 與 Hwang 的 *Introduction to Probability*（概率論入門教材）以及 Grimmett 與 Welsh 的 *Probability: An Introduction*（概率論導論）中都有詳細闡述。

詳細解答

STEP III 2005，第 14 題

概率與統計 • 連續隨機變量

STEP III 2005，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/05-s3-q14

關鍵思路

這一題由始至終只靠題目容許我們直接引用的那一個積分，所以功夫全在於看出每個積分需要取哪個 m 和 n。代換（substitution）$x = k^2/u$ 把區間 $[0, k]$ 與 $[k, \infty)$ 對調，而被積函數（integrand）的形狀保持不變，這種對稱性馬上確定了中位數（median）。至於時間 $T = s/V$，同樣的倒數技巧說明 T 的概率密度（density）與 V 的屬於完全相同的族，只是 k 換成了 s/k。之後，中位數和平均值（mean）都可以從 V 的結果直接讀出。

解答

常數 C

在已知結果中取 $m = a$ 和 $n = 2a$。這是容許的，因為 a 是正整數，而且 $2a \geq a$。這時 $n + 2 = 2a + 2$，$n - m + 1 = a + 1$，所以

$$\int_0^\infty \frac{x^a}{(x+k)^{2a+2}}\,\mathrm{d}x = \frac{a!\,a!}{(2a+1)!\,k^{a+1}}.$$

密度函數在其取值範圍上的積分為 1，所以

$$1 = Ck^{a+1} \cdot \frac{a!\,a!}{(2a+1)!\,k^{a+1}} \qquad \Longrightarrow \qquad C = \frac{(2a+1)!}{a!\,a!},$$

正如所求。

代換與中位數

在左邊的積分中，設 $x = k^2/u$，於是 $\mathrm{d}x = -\dfrac{k^2}{u^2}\,\mathrm{d}u$。當 x 由 0 增至 v 時，u 由 ∞ 減至 k^2/v。此外

$$x + k = \frac{k^2}{u} + k = \frac{k(u+k)}{u},$$

所以

$$\frac{x^a}{(x+k)^{2a+2}} = \frac{k^{2a}/u^a}{k^{2a+2}(u+k)^{2a+2}/u^{2a+2}} = \frac{u^{a+2}}{k^2\,(u+k)^{2a+2}}.$$

乘以 $\mathrm{d}x$ 後，k^2/u^2 剛好約去，而負號則把積分上下限對調：

$$\int_0^v \frac{x^a}{(x+k)^{2a+2}}\,\mathrm{d}x = \int_\infty^{k^2/v} \frac{u^a}{(u+k)^{2a+2}}\,(-\mathrm{d}u) = \int_{k^2/v}^\infty \frac{u^a}{(u+k)^{2a+2}}\,\mathrm{d}u,$$

這正是所要證明的恆等式。

想一想 v 取哪個值時，這條恆等式會告訴我們關於中位數的事情？

取 $v = k$，於是 $k^2/v = k$ 也成立。把兩邊同乘 Ck^{a+1} 後，恆等式變成

$$\begin{aligned}\mathrm{P}(V < k) &= \int_0^k \mathrm{f}(x)\,\mathrm{d}x \\ &= \int_k^\infty \mathrm{f}(u)\,\mathrm{d}u = \mathrm{P}(V > k).\end{aligned}$$

V 是連續（continuous）隨機變量，所以 $\mathrm{P}(V = k) = 0$，這兩個概率之和為 1。因此兩者各為 $\frac{1}{2}$，V 的中位數就是 k。

V 的期望值

期望值（expected value）為

$$\mathrm{E}(V) = Ck^{a+1}\int_0^\infty \frac{x^{a+1}}{(x+k)^{2a+2}}\,\mathrm{d}x.$$

現在取 $m = a+1$ 和 $n = 2a$。這需要 $2a \geq a+1$，而由於 $a \geq 1$，這是成立的。這時 $n - m + 1 = a$，並且

$$\mathrm{E}(V) = \frac{(2a+1)!}{a!\,a!}\,k^{a+1}\cdot\frac{(a+1)!\,(a-1)!}{(2a+1)!\,k^a} = k\cdot\frac{(a+1)!}{a!}\cdot\frac{(a-1)!}{a!} = \frac{(a+1)k}{a}.$$

T 的分佈

速率是正數，所以對於 $t > 0$，我們可以把不等式兩邊同乘 $V/t > 0$ 而不必改變不等號的方向：

$$T < t \iff \frac{s}{V} < t \iff V > \frac{s}{t}.$$

因此

$$\mathrm{P}(T<t)=\mathrm{P}\Big(V>\frac{s}{t}\Big)=\int_{s/t}^{\infty}\frac{Ck^{a+1}x^a}{(x+k)^{2a+2}}\,\mathrm{d}x, \qquad (*)$$

正如所求。

現在在 $(*)$ 中作代換 $u=s/x$，於是 $x=s/u$，$\mathrm{d}x=-\dfrac{s}{u^2}\,\mathrm{d}u$。當 x 由 s/t 增至 ∞ 時，u 由 t 減至 0。由於 $x+k=(s+ku)/u$，

$$\mathrm{P}(T<t)=\int_0^t Ck^{a+1}\,\frac{s^a/u^a}{(s+ku)^{2a+2}/u^{2a+2}}\cdot\frac{s}{u^2}\,\mathrm{d}u=\int_0^t\frac{Ck^{a+1}s^{a+1}\,u^a}{(s+ku)^{2a+2}}\,\mathrm{d}u.$$

這對每個 $t>0$ 都成立，而當 $t\le 0$ 時 $\mathrm{P}(T<t)=0$。所以對 t 求導，便得到 T 的密度函數：

$$\mathrm{g}(t)=\frac{Ck^{a+1}s^{a+1}\,t^a}{(s+kt)^{2a+2}}\qquad(t>0).$$

為了與 f 比較，從分母抽出 k^{2a+2}，並記 $K=s/k$：

$$\mathrm{g}(t)=\frac{C\,(s/k)^{a+1}\,t^a}{(t+s/k)^{2a+2}}=\frac{CK^{a+1}t^a}{(t+K)^{2a+2}}.$$

這正正是把 k 換成 $K=s/k$ 之後的密度函數 f。所以 T 與 V 屬於同一個分佈（distribution）族。

中位數與平均值

由於 T 的分佈形式相同，只是參數（parameter）為 $K=s/k$，之前的結果可以直接套用：

$$\text{就中位數而言，}T=\frac{s}{k},\qquad \mathrm{E}(T)=\frac{(a+1)}{a}\cdot\frac{s}{k}.$$

因此，時間的中位數與速率的中位數之積為

$$\frac{s}{k}\cdot k=s,$$

這就是距離，與題目所述相符。時間的期望值與速率的期望值之積為

$$\mathrm{E}(T)\,\mathrm{E}(V)=\frac{(a+1)s}{ak}\cdot\frac{(a+1)k}{a}=s\Big(1+\frac{1}{a}\Big)^2>s,$$

因為 a 是正整數，所以 $1+1/a>1$。至此全題完成。

驗算

- 當 $a=1$，$k=1$ 時，我們得到 $C=3!/1=6$，而對 $6x/(x+1)^4$ 在 $[0,\infty)$ 上作數值積分（numerical integration）得 1.000000。在 $[0,1]$ 上則得 0.500000，印證了中位數。

- 對若干組數值作數值求積，例如 $a = 2$，$k = 0.7$，得 $\mathrm{E}(V) = 1.05000 = \frac{3}{2}(0.7)$。取 $s = 3.2$ 時，得 $\mathrm{E}(T)\mathrm{E}(V) = 7.20000 = 3.2 \times \frac{9}{4}$。
- 以 $a = 2$，$k = s = 1$ 模擬（simulation）2×10^5 個速率，結果是速率中位數 $\times$ 時間中位數 ≈ 1.000001，但平均速率 ≈ 1.504（理論值：1.5）。
- 中位數之積等於 s 這一點，不用任何公式也看得出來。$T = s/V$ 是 V 的遞減函數，所以 T 的中位數就是 s 除以 V 的中位數。

常見錯誤

- 引用已知積分時忘記檢查 $n \geq m$。對 $\mathrm{E}(V)$ 而言，這需要 $a \geq 1$，這正是 a 必須為正整數的原因。對於更高階的矩（moment），公式最終會失效，事實上 $\mathrm{E}(V^{a+1})$ 是無限大。
- 對 $(*)$ 求導以求 T 的密度函數時，漏掉因子 s/t^2（鏈式法則，chain rule）。
- 以為 $\mathrm{E}(s/V) = s/\mathrm{E}(V)$。最後一部分的重點，正是要指出這是錯的。

STEP 之外

最後的不等式是琴生不等式（Jensen's inequality）的一個例子：當 $v > 0$ 時 $1/v$ 是凸（convex）函數，所以 $\mathrm{E}(1/V) \geq 1/\mathrm{E}(V)$，等號僅在 V 為常數時成立。相比之下，中位數卻可以直接穿過任何單調（monotonic）函數。代換 $V/(V+k)$ 把 V 變成一個 $\mathrm{Beta}(a+1, a+1)$ 變量，它關於 $\frac{1}{2}$ 對稱，這是理解中位數為何是 k 的另一個途徑。Blitzstein 與 Hwang 的 *Introduction to Probability*（概率論入門）對琴生不等式、變量代換和 Beta 分佈族都有十分易讀的講解。

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

2005 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。