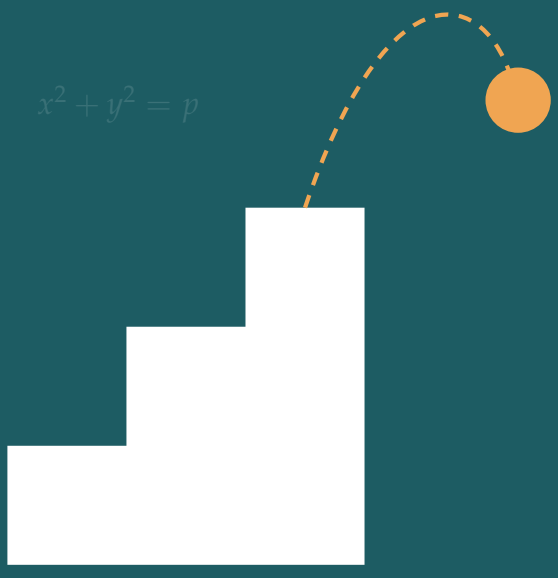
STEP 備試 · 大學數學

Beyond STEP

STEP I, II 及 III

2001 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

2001 年詳細解答

Keith McNulty

Beyond STEP：以歷屆 STEP 試題為門，走進大學數學

2001 年詳細解答

2001 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 幾何不等式 純數學 . . . 3

第 2 題 代數不等式 純數學 . . . 6

第 3 題 三次方程與判別式 純數學 . . . 9

第 4 題 反三角函數方程 純數學 . . . 13

第 5 題 積分號下求導 純數學 . . . 17

第 6 題 旋轉曲面 純數學 . . . 20

第 7 題 凹性與微分方程 純數學 . . . 23

第 8 題 二階線性方程 純數學 . . . 27

第 9 題 攔截與相對速度 力學 . . . 30

第 10 題 拋體的飛行時間 力學 . . . 33

第 11 題 彈性繩與能量 力學 . . . 36

第 12 題 欄形圖 概率與統計 . . . 40

第 13 題 條件泊松分佈 概率與統計 . . . 43

第 14 題 隨機排序 概率與統計 . . . 46

STEP II

第 1 題 二項式近似 純數學 . . . 49

第 2 題 整數部分的求和 純數學 . . . 52

第 3 題 歐幾里得幾何 純數學 . . . 55

第 4 題 三角方程 純數學 . . . 59

第 5 題 駐點與積分 純數學 . . . 62

第 6 題 代換積分法 純數學 . . . 66

第 7 題 複數與旋轉 純數學 . . . 69

第 8 題 周期函數與常微分方程 純數學 . . . 72

第 9 題 二次空氣阻力 力學 . . . 75
第 10 題 碰撞與恢復係數 力學 . . . 78
第 11 題 拋體運動 力學 . . . 82
第 12 題 泊松近似 概率與統計 . . . 86
第 13 題 連續分佈與估計 概率與統計 . . . 89
第 14 題 最小方差估計量 概率與統計 . . . 93

STEP III

第 1 題 麥克勞林級數 純數學 . . . 96
第 2 題 雙曲代換與面積 純數學 . . . 100
第 3 題 二次方程的根 純數學 . . . 104
第 4 題 反三角函數 純數學 . . . 108
第 5 題 拋物線的法線 純數學 . . . 112
第 6 題 向量與三角形的心 純數學 . . . 116
第 7 題 正交軌線 純數學 . . . 120
第 8 題 複數軌跡 純數學 . . . 124
第 9 題 摩擦力與簡諧運動 力學 . . . 128
第 10 題 相對速度 力學 . . . 132
第 11 題 轉動能量守恆 力學 . . . 136
第 12 題 示性指示變量求期望值 概率與統計 . . . 140
第 13 題 硬幣序列遊戲 概率與統計 . . . 143
第 14 題 方差的無偏估計量 概率與統計 . . . 147

詳細解答

STEP I 2001，第 1 題

純數學 • 幾何不等式

> **STEP I 2001，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/01-s1-q1

關鍵思路

三點分別位於正方形的三條不同邊上，其中必有兩點落在一對對邊上，而第三條邊則連接這兩條對邊。為正方形設立坐標（coordinates），使每一點只有一個介乎 0 與 1 之間的自由參數（parameter）。然後用反證法（proof by contradiction）：假如三角形的三條邊都比該上界長，每一個長度條件都會收緊其中一個參數的範圍，而這些限制合起來，便會把其中一點推出正方形之外。

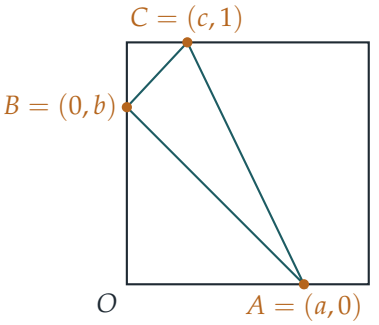
解答

以厘米為單位。正方形有四條邊，而三點位於不同的邊上，即這些點恰好不在正方形的其中一條邊上，所以其中兩點位於一對對邊上，而第三點則位於與這兩條邊都相鄰的一條邊上。選取坐標軸，使正方形為 $0 \le x \le 1$，$0 \le y \le 1$，並且

$$A = (a, 0) \text{ 在底邊}, \qquad B = (0, b) \text{ 在左邊}, \qquad C = (c, 1) \text{ 在頂邊},$$

其中 $a, b, c \in [0, 1]$。（把正方形旋轉或反射並不會改變任何長度，所以這樣做不失一般性（without loss of generality）。）於是

$$AB^2 = a^2 + b^2, \qquad BC^2 = c^2 + (1 - b)^2, \qquad CA^2 = (c - a)^2 + 1.$$



設 $k = \sqrt{6} - \sqrt{2}$ 及 $t = 2 - \sqrt{3}$。以下兩個小結果能讓數字變得容易處理：

$$k^2 = 8 - 2\sqrt{12} = 8 - 4\sqrt{3}, \qquad t^2 = 7 - 4\sqrt{3}, \qquad \text{所以} \qquad k^2 - 1 = t^2.$$

此外 $2t = 4 - 2\sqrt{3} = (\sqrt{3} - 1)^2$，所以 $\sqrt{2t} = \sqrt{3} - 1$，因此

$$\sqrt{2t} + t = (\sqrt{3} - 1) + (2 - \sqrt{3}) = 1.$$

假設三角形的三條邊都比 k 長，以求得出矛盾。把正方形沿直線 $y = \frac{1}{2}$ 反射（這會把 A 與 C 的角色對調，亦把 a 與 c 對調），由這個對稱性（symmetry），我們可以假設 $c \geq a$。

由 $BC > k$ 出發。由於 $c \leq 1$，

$$k^2 < BC^2 = c^2 + (1 - b)^2 \leq 1 + (1 - b)^2,$$

所以 $(1 - b)^2 > k^2 - 1 = t^2$。因為 $1 - b \geq 0$ 及 $t > 0$，由此得 $1 - b > t$，即 $0 \leq b < 1 - t$。

由 $AB > k$ 出發。由於 $0 \leq b < 1 - t$，我們有 $b^2 < (1 - t)^2$，所以

$$a^2 > k^2 - b^2 > k^2 - (1 - t)^2 = (1 + t^2) - (1 - 2t + t^2) = 2t.$$

因為 $a \geq 0$，由此得 $a > \sqrt{2t} = \sqrt{3} - 1$。

想一想 我們已經限定了 a 的範圍；第三條邊 CA 又會帶來甚麼？

邊 CA 由底至頂橫越整個正方形，所以它在垂直方向的跨度為 1，其餘的長度必須靠水平方向補足。

由 $CA > k$ 出發。我們有 $(c - a)^2 + 1 > k^2$，所以 $(c - a)^2 > t^2$，而由於 $c \geq a$，這表示 $c - a > t$。因此

$$c > a + t > \sqrt{2t} + t = 1.$$

但 C 位於頂邊上，所以 $c \leq 1$。這個矛盾說明三條邊不可能全都比 k 長：三角形 ABC 至少有一條邊的長度不超過 $(\sqrt{6} - \sqrt{2})$ cm，這正是所要證明的。

這個上界可以達到

當 $c = 1$ 及 $a = b = \sqrt{3} - 1$，即 $A = (\sqrt{3} - 1, 0)$、$B = (0, \sqrt{3} - 1)$、$C = (1, 1)$ 時，上面每一個不等式（inequality）都取到臨界情況。這時

$$AB^2 = 2(\sqrt{3} - 1)^2 = 8 - 4\sqrt{3}, \qquad BC^2 = CA^2 = 1 + (2 - \sqrt{3})^2 = 8 - 4\sqrt{3},$$

所以 ABC 是等邊三角形（equilateral triangle），邊長恰好為 k（C 位於角上，算作在頂邊上）。因此 $\sqrt{6} - \sqrt{2}$ 是最佳的常數。用三角學的說法，$k = \sec 15°$：這是能放進單位正方形內的最大等邊三角形，與各邊傾斜成 15°。

驗算

- $k = \sqrt{6} - \sqrt{2} \approx 1.0353$ 比正方形的邊長稍大，這是合理的：若三角形的三個頂點分別在三條邊上，它的三條邊可以全都比 1 長，但只能長一點點。
- 在 2×10^6 組隨機選取的 $(a, b, c) \in [0, 1]^3$ 中搜尋，最短邊從未超過 1.0337，低於 k；而上述等邊三角形的例子，三條邊在機器精度內均等於 1.03528。
- $\sec 15° = 4/(\sqrt{6} + \sqrt{2}) = \sqrt{6} - \sqrt{2}$，與等邊三角形的情況吻合。

常見錯誤

- 未經證明便假設三角形是等邊的，或假設其中一個頂點在角上。極端情況能提示答案，卻不能證明不等式對所有情況都成立。
- 忘記說明設定的理由：關鍵的結構性事實是，三條不同的邊之中必定包含一對對邊。
- 對 $(1 - b)^2 > t^2$ 或 $(c - a)^2 > t^2$ 開平方時，沒有說明所取的是哪一個正負號。

STEP 之外

這是一個小型的極大極小（max–min）問題：我們要在一個由各種情況組成的緊緻集（compact set）上，使最短邊盡量長，而極端情況會令所有約束條件同時取等號，這種模式在最優化（optimisation）和博弈論（game theory）中隨處可見。由於 $[0, 1]^3$ 是有界閉集，而最短邊是一個連續函數（continuous function），最大值必定存在，這個事實會在大學第一門分析課程中證明（極值定理（extreme value theorem））。Abbott 的 *Understanding Analysis*（大學分析入門教材）對此有很好的講解，而 Siklos 的 *Advanced Problems in Mathematics*（STEP 難題選講）則有更多同類風格的 STEP 幾何題。

詳細解答

STEP I 2001，第 2 題

純數學 • 代數不等式

> **STEP I 2001，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/01-s1-q2

關鍵思路

處理不等式（inequality）時，不能把兩邊同乘一個正負未知的量，所以 (i) 的穩妥做法是把所有項移到同一邊，然後因式分解（factorise）：這樣答案便可從符號圖（sign diagram）直接讀出。在 (ii) 中，平方根會引誘我們把兩邊平方，但這只有在已知兩邊皆非負時才是合理的。全題的關鍵就在於時刻留意這個條件。

解答

第 (i) 部分

我們不能乘以 x，因為 x 可能是負數。我們改為乘以 x^2，由於 $x \neq 0$，它必定嚴格為正，所以不等號方向不變：

$$\begin{aligned} 1+2x-x^2 > \frac{2}{x} &\iff x^2+2x^3-x^4 > 2x \\ &\iff 0 > x^4-2x^3-x^2+2x \\ &\iff x\left(x^3-2x^2-x+2\right) < 0. \end{aligned}$$

這個三次式（cubic）可用分組法因式分解：

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

所以我們需要

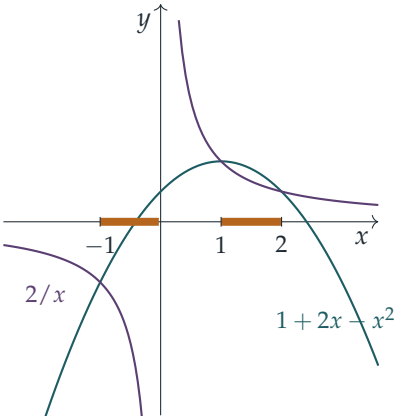
$$(x+1)\,x\,(x-1)(x-2) < 0.$$

這是一個首項係數（leading coefficient）為正的四次多項式（quartic），有四個單根（simple root）$-1, 0, 1, 2$，所以每當 x 越過一個根，它的正負號便交替一次：當 $x > 2$ 時為正，

在 $(1,2)$ 上為負，在 $(0,1)$ 上為正，在 $(-1,0)$ 上為負，當 $x<-1$ 時為正。在各個根上，其值為 0，並不 <0（而且 $x=0$ 本來就已被排除）。因此

$$\boxed{-1<x<0 \quad \text{或} \quad 1<x<2.}$$

圖像與此吻合：拋物線（parabola）$y=1+2x-x^2$ 位於雙曲線（hyperbola）$y=2/x$ 上方的範圍，恰好由它們在 $x=-1$、1、2 的交點劃分，而在負的一側，則只到雙曲線在 $x=0$ 處急墜至 $-\infty$ 為止。



第 (ii) 部分

首先留意，$x\geq -\frac{10}{3}$ 除了給出 $3x+10\geq 0$ 之外，還給出 $x+4\geq \frac{2}{3}>0$，所以兩個平方根在整個範圍內都有定義。

兩邊皆非負，而對於 $a,b\geq 0$，我們有 $a>b \iff a^2>b^2$。所以我們可以把兩邊平方，既不會多出解，也不會丟失解：

$$\begin{aligned}\sqrt{3x+10}>2+\sqrt{x+4} &\iff 3x+10>4+4\sqrt{x+4}+(x+4)\\ &\iff 2x+2>4\sqrt{x+4}\\ &\iff x+1>2\sqrt{x+4}.\end{aligned}$$

想一想 我們可以直接再平方一次嗎？

不能馬上這樣做。右邊的 $2\sqrt{x+4}$ 是正數，但 $x+1$ 可能是負數，這時平方便可能把一個不成立的不等式變成成立的不等式。所以我們要分情況討論。

- 若 $x+1\leq 0$，左邊至多為 0，而右邊嚴格為正。不等式不成立。

- 若 $x+1>0$，兩邊皆為正，所以平方這一步同樣是可逆的：

$$\begin{aligned} x+1>2\sqrt{x+4} &\iff x^2+2x+1>4x+16 \\ &\iff x^2-2x-15>0 \\ &\iff (x-5)(x+3)>0, \end{aligned}$$

  這在 $x>5$ 或 $x<-3$ 時成立。結合 $x>-1$，只餘下 $x>5$。

因此

$$\boxed{x>5.}$$

驗算

- 在 $x=5$：$\sqrt{25}=5$ 而 $2+\sqrt{9}=5$，所以該處等號成立，這正是邊界上應有的情況。在 $x=12$：$\sqrt{46}\approx 6.78>2+4=6$。在 $x=0$：$\sqrt{10}\approx 3.16<4$。
- 在 (i) 中，$x=1.5$ 給出 $1.75>1.33$，$x=-0.5$ 給出 $-0.25>-4$，兩者皆成立；$x=0.5$ 給出 $1.75<4$，不成立。
- 我們用電腦對兩個不等式各自在 2×10^5 個隨機點上作了測試，另加上與每個邊界相距不超過 10^{-6} 的點，結果與上述答案完全一致。

常見錯誤

- 把 (i) 乘以 x，卻沒有分開 $x>0$ 和 $x<0$ 兩種情況。這樣會丟失區間 $(-1,0)$，或者把該處的答案弄反。
- 把 $x+1>2\sqrt{x+4}$ 平方，卻沒有要求 $x+1>0$。這樣會得出多餘的區間 $-\frac{10}{3}\le x<-3$；試把 $x=-3.2$ 代入原式，便可看到它並不成立（$\sqrt{0.4}\approx 0.63<2+\sqrt{0.8}$）。

STEP 之外

(ii) 中的原則是：一個運算只有在所涉及的數值上是遞增函數（increasing function）時，才會保持不等式成立。這正是分析學家處理不等式的起點：$t\mapsto t^2$ 在 $[0,\infty)$ 上遞增，但在 $\mathbb{R}$ 上則不然。乘積的符號圖可推廣為以下事實：連續函數（continuous function）只有經過零才能改變正負號，這就是介值定理（intermediate value theorem）。Abbott 的 *Understanding Analysis*（一本分析學入門書）對這兩點都有細緻的論述。

詳細解答

STEP I 2001，第 3 題

純數學 • 三次方程與判別式

STEP I 2001，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/01-s1-q3

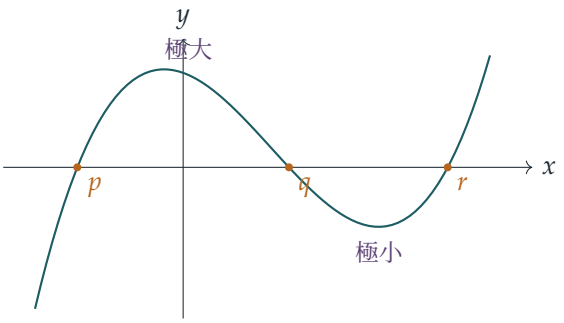
關鍵思路

一個有三個相異實根的三次多項式（cubic），在每兩個相鄰的根之間必定要轉向，所以它的導數（derivative）——一個二次式——有兩個相異實根。把「有兩個相異實根」寫成「判別式（discriminant）為正」，就得到第一個不等式。至於第二部分，題目給出的乘積是一個三次多項式，其係數正好就是目標不等式中的那些量，所以用第一部分的結果（或者配方（completing the square））便可完成。

解答

略圖

展開得 $f(x) = x^3 - (p+q+r)x^2 + \ldots$，所以當 $x \to \infty$ 時 $f(x) \to +\infty$，而當 $x \to -\infty$ 時 $f(x) \to -\infty$。圖像只在 p、q 和 r 三處與 x-軸相交，而且由於每個根都是單根（simple root），f 在每個根處都會變號：當 $x < p$ 時 $f < 0$，在 (p,q) 上 $f > 0$，在 (q,r) 上 $f < 0$，當 $x > r$ 時 $f > 0$。因此在 p 與 q 之間某處有一個局部極大值（local maximum），在 q 與 r 之間某處有一個局部極小值（local minimum）。y-截距為 $f(0) = -pqr$。



（圖中是按 $p < 0 < q < r$ 的情況繪畫的；無論各根位於何處，圖像的形狀都一樣，只是 y-軸的位置有所不同。）

第一個不等式

完全展開，得

$$f(x) = x^3 - (p+q+r)x^2 + (qr+rp+pq)x - pqr,$$

所以

$$f'(x) = 3x^2 - 2(p+q+r)x + (qr+rp+pq).$$

由於 $f(p) = f(q) = 0$，根據羅爾定理（Rolle's theorem），f' 在 p 與 q 之間（不含端點）有一個根；同理，在 q 與 r 之間（不含端點）也有一個根。這兩個區間互不重疊，所以二次方程 $f'(x) = 0$ 有兩個相異實根。（不用羅爾定理的做法：$f'(p) = (p-q)(p-r) > 0$、$f'(q) = (q-p)(q-r) < 0$ 及 $f'(r) = (r-p)(r-q) > 0$，所以 f' 在每個區間上都會變號。）

一個二次式有兩個相異實根，當且僅當它的判別式為正，所以

$$4(p+q+r)^2 - 12(qr+rp+pq) > 0, \qquad \text{即} \qquad (p+q+r)^2 > 3(qr+rp+pq),$$

這正是所要證明的。

第二個不等式

展開得

$$(x^2+gx+h)(x-k) = x^3 + (g-k)x^2 + (h-gk)x - hk.$$

把它與 f 的展開式比較：如果這個三次多項式的根是 p, q, r，則

$$p+q+r = -(g-k), \qquad qr+rp+pq = h-gk,$$

所以目標不等式 $(g-k)^2 > 3(h-gk)$ 正好就是把第一個不等式應用於這個三次多項式的根。

充分（*sufficient*）。假設 $g^2 > 4h$。那麼 $x^2 + gx + h$ 有兩個相異實根 $\alpha < \beta$，而這個三次多項式有實根 α、β、k。如果 k 與兩者都不相同，三個根便互不相同，由第一部分立即得出 $(g-k)^2 > 3(h-gk)$。

想一想 如果 k 剛好等於 α 或 β，以致各根並非互不相同，那又如何？

第一部分假設了 $p < q < r$，所以我們還需要多做一點工夫。恆等式（identity）

$$(p+q+r)^2 - 3(qr+rp+pq) = \tfrac{1}{2}\left[(p-q)^2 + (q-r)^2 + (r-p)^2\right]$$

顯示，除非 $p = q = r$，否則左邊必為正。由於 $\alpha \neq \beta$，三個根絕不會全部相等，所以不等式在這個情況下同樣成立。

另有一個直接的做法，可以完全避開根。乘開得

$$(g-k)^2 - 3(h-gk) = g^2 + gk + k^2 - 3h = \left(k + \tfrac{1}{2}g\right)^2 + \tfrac{3}{4}\left(g^2 - 4h\right),$$

當 $g^2 > 4h$ 時，這是一個平方加上一個正數，因此為正。無論用哪一個方法，$g^2 > 4h$ 都是充分條件。

非必要（*not necessary*）。我們需要找出滿足 $g^2 \leq 4h$ 而不等式仍然成立的數值。取 $g = 0$、$h = 1$、$k = 2$。那麼 $g^2 = 0 \leq 4 = 4h$，但是

$$(g-k)^2 = 4 > 3 = 3(h-gk).$$

所以即使沒有 $g^2 > 4h$，不等式也可以成立，這個條件並非必要。

想一想 第一部分明明需要三個實根，為甚麼會這樣？

第一部分其實只需要 $f'(x) = 0$ 有兩個相異實根，這個要求較弱。這裏的三次多項式 $(x^2+1)(x-2) = x^3 - 2x^2 + x - 2$ 只有一個實根，但 $f'(x) = 3x^2 - 4x + 1 = (3x-1)(x-1)$，所以它仍然有兩個駐點（stationary point）。不等式 $(g-k)^2 > 3(h-gk)$ 所說的，正正就是這個三次多項式有一個局部極大值和一個與之不同的局部極小值，不論它的圖像是否與橫軸相交三次。

驗算

- 取 $p, q, r = 0, 1, 2$：$(p+q+r)^2 = 9 > 6 = 3(qr+rp+pq)$，而 $f'(x) = 3x^2 - 6x + 2$ 的根是 $1 \pm 1/\sqrt{3}$，一個在 $(0,1)$ 內，一個在 $(1,2)$ 內。
- 配方所得的公式顯示，這個不等式的界限是最佳的，不能再改進：取 $g = 2$、$h = 1$（所以 $g^2 = 4h$）及 $k = -1$，得 $(g-k)^2 = 9 = 3(h-gk)$，是等式而不是嚴格不等式。這就是三次多項式 $(x+1)^3$，三個根全部相等。

- 我們用電腦以精確運算核對了兩個恆等式和兩個不等式，每個各用了 20 000 個隨機有理數例子，另加例子 $g = 0$、$h = 1$、$k = 2$（得出 $4 > 3$）。

常見錯誤

- 把三次曲線上下顛倒地畫：x^3 的係數是 $+1$，所以圖像向右上升。
- 忘記了 k 是 $x^2 + gx + h$ 的根這個情況，這時第一部分不能照原樣套用。
- 只用一個滿足 $g^2 \leq 4h$ 的例子說明不等式成立，就稱之為「充分」。一個例子證明的是*非必要*；要證明充分性，則須對所有 g、h、k 給出證明。

STEP 之外

$(p + q + r)^2 - 3(qr + rp + pq)$ 這個量，除了相差一個常數因子外，就是導數的判別式，而且它是各根的*對稱函數*（*symmetric function*）：任何這類函數都可以用係數來表示，這就是為甚麼這個不等式可以從根的形式轉寫成 g、h、k 的形式。三次多項式本身的完整判別式 $\prod(p - q)^2$，決定了三個根是否全為實數；本題用的是它較弱的近親。對稱函數和判別式再往前走，就通往伽羅瓦理論（Galois theory）；Allenby 的 *Rings, Fields and Groups*（環、域與羣）是一個很好的起點。

詳細解答

STEP I 2001，第 4 題

純數學 • 反三角函數方程

STEP I 2001，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/01-s1-q4

關鍵思路

三倍角公式（triple-angle formula）可直接由 tan 的複角公式（addition formula）推出。之後，一切都取決於一個事實：tan 以 π 為週期（period）重複，所以若 $\tan 3\varphi = \tan 3\theta$，則 φ 與 θ 可以相差 $\pi/3$ 的任何整數倍。(i) 和 (ii) 真正要下工夫的地方，是留意 $\cos^{-1}$ 和 $\tan^{-1}$ 的值域（range），因為這些值域決定了那無窮多個角之中哪些才是容許的。

解答

三倍角公式

設 $t = \tan\theta$。由複角公式，$\tan 2\theta = \dfrac{2t}{1-t^2}$，於是

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

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

其中我們把分子和分母同乘以 $1 - t^2$。這就是所求的公式。（只要等號兩邊都有定義，公式便成立。）

$\tan 3\theta$ 的值

由於 $\cos\theta = 2/\sqrt{5}$ 且 $0 < \theta < \pi/2$，我們有 $\sin\theta > 0$，所以

$$\sin\theta = \sqrt{1-\tfrac{4}{5}} = \frac{1}{\sqrt{5}}, \qquad \tan\theta = \frac{\sin\theta}{\cos\theta} = \frac{1}{2}.$$

（可以想像一個邊長為 1、2、$\sqrt{5}$ 的直角三角形。）代入 $t = \frac{1}{2}$：

$$\tan 3\theta = \frac{\frac{3}{2} - \frac{1}{8}}{1 - \frac{3}{4}} = \frac{11/8}{1/4} = \frac{11}{2},$$

正如所求。

稍後會用到 θ 相當小這一點：$\tan\theta = \frac{1}{2} < \frac{1}{\sqrt{3}} = \tan\frac{\pi}{6}$，所以 $0 < \theta < \pi/6$，因而 $0 < 3\theta < \pi/2$。

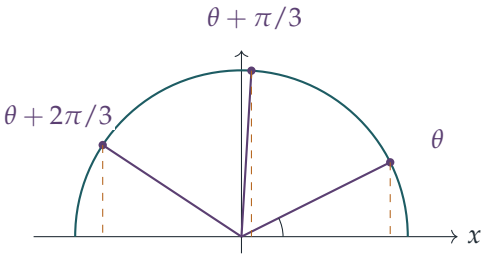
第 (i) 部分

設 $\varphi = \cos^{-1} x$，則 $0 \le \varphi \le \pi$ 且 $x = \cos\varphi$。方程即是說 $\tan 3\varphi = \frac{11}{2} = \tan 3\theta$。因為 tan 的週期為 π，而且在每個長度為 π 的區間上每個值都恰好取一次，

$$3\varphi = 3\theta + k\pi \quad \Longleftrightarrow \quad \varphi = \theta + \frac{k\pi}{3}, \qquad k \in \mathbb{Z}.$$

想一想 這些角之中，實際上有多少個是容許的？

我們需要 $0 \le \varphi \le \pi$。由於 $0 < \theta < \pi/6$，取 $k = 0, 1, 2$ 得 $\varphi = \theta,\ \theta + \frac{\pi}{3},\ \theta + \frac{2\pi}{3}$，全部都嚴格落在 $(0, \pi)$ 之內；$k = 3$ 給出 $\varphi > \pi$，而 $k = -1$ 給出 $\varphi < 0$。所以恰好有三個解。



x 的三個值就是這些角的餘弦（即各條虛線的垂足）。利用 $\cos\theta = 2/\sqrt{5}$、$\sin\theta = 1/\sqrt{5}$ 和複角公式：

$$x_1 = \cos\theta = \frac{2}{\sqrt{5}},$$

$$x_2 = \cos\theta\cos\tfrac{\pi}{3} - \sin\theta\sin\tfrac{\pi}{3} = \frac{2}{\sqrt{5}}\cdot\frac{1}{2} - \frac{1}{\sqrt{5}}\cdot\frac{\sqrt{3}}{2} = \frac{2-\sqrt{3}}{2\sqrt{5}},$$

$$x_3 = \cos\theta\cos\tfrac{2\pi}{3} - \sin\theta\sin\tfrac{2\pi}{3} = -\frac{2}{\sqrt{5}}\cdot\frac{1}{2} - \frac{1}{\sqrt{5}}\cdot\frac{\sqrt{3}}{2} = -\frac{2+\sqrt{3}}{2\sqrt{5}}.$$

這些角都不會令 3φ 成為 $\pi/2$ 的奇數倍（每個 3φ 都是 $3\theta + k\pi$，其中 $0 < 3\theta < \pi/2$），所以 $\tan 3\varphi$ 在每個情況下都有定義。(i) 的解是

$$x = \frac{2}{\sqrt{5}}, \qquad x = \frac{2-\sqrt{3}}{2\sqrt{5}}, \qquad x = -\frac{2+\sqrt{3}}{2\sqrt{5}}.$$

（有理化（rationalising）後，它們是 $\frac{2\sqrt{5}}{5}$、$\frac{(2-\sqrt{3})\sqrt{5}}{10}$ 和 $-\frac{(2+\sqrt{3})\sqrt{5}}{10}$。）

另一途徑。把 $T = \tan\varphi$ 代入三倍角公式，得 $2(3T - T^3) = 11(1 - 3T^2)$，即

$$2T^3 - 33T^2 - 6T + 11 = 0.$$

我們知道 $T = \frac{1}{2}$ 是一個根（root），因式分解（factorising）得 $(2T-1)(T^2 - 16T - 11) = 0$，所以 $T = \frac{1}{2}$ 或 $T = 8 \pm 5\sqrt{3}$。於是 $x = \cos\varphi = \pm 1/\sqrt{1+T^2}$，其正負號與 T 相同（T 為正表示 φ 是銳角；為負則表示是鈍角）。這樣得出同樣的三個答案，而且 $1 + (8+5\sqrt{3})^2 = 140 + 80\sqrt{3} = 1/x_2^2$ 也是一個不錯的驗算。

第 (ii) 部分

設 $\psi = \frac{1}{3}\tan^{-1} y$。由於 $-\frac{\pi}{2} < \tan^{-1} y < \frac{\pi}{2}$，我們有 $-\frac{\pi}{6} < \psi < \frac{\pi}{6}$。

> ? **想一想** 條件 $\cos\psi = 2/\sqrt{5}$ 給出一個角還是兩個角？

在 $(-\pi/6, \pi/6)$ 上，cos 是偶函數（even function），所以 $\cos\psi = 2/\sqrt{5} = \cos\theta$ 給出 $\psi = \theta$ 或 $\psi = -\theta$，而且因為 $0 < \theta < \pi/6$，兩者都在容許的範圍內。（再沒有其他解：在 $[0, \pi/6)$ 上，函數 cos 嚴格遞減（strictly decreasing），所以它在該處只取 $2/\sqrt{5}$ 這個值一次，在 $(-\pi/6, 0]$ 上亦然。）

所以 $\tan^{-1} y = \pm 3\theta$。由於 $0 < 3\theta < \pi/2$，$\pm 3\theta$ 兩者確實都是 $\tan^{-1}$ 的值，而

$$y = \tan(\pm 3\theta) = \pm\frac{11}{2}.$$

(ii) 的解是 $y = \frac{11}{2}$ 和 $y = -\frac{11}{2}$。

驗算

- $\theta = \cos^{-1}(2/\sqrt{5}) \approx 0.4636$，所以 $3\theta \approx 1.3909$，而 $\tan 3\theta \approx 5.5000$。
- $x_2 = \frac{2-\sqrt{3}}{2\sqrt{5}} \approx 0.05992$ 而 $x_3 \approx -0.83451$；用計算機算得兩者的 $\tan(3\cos^{-1} x) \approx 5.5$，而在 $[-1, 1]$ 上作細密的數值掃描，亦恰好找到三個解。
- $\cos\left(\frac{1}{3}\tan^{-1}(\pm 5.5)\right) = \cos(\pm 0.4636) \approx 0.8944 = 2/\sqrt{5}$。

常見錯誤

- 在 (i) 中只給出 $x = 2/\sqrt{5}$，忘記了 tan 每隔 π 便重複一次，所以 3φ 在 $[0, 3\pi]$ 內有三個可能的值。
- 在 (ii) 中只給出 $y = 11/2$：cos 是偶函數，所以 $\psi = -\theta$ 同樣成立。
- 沒有檢查範圍：答案取決於 $\theta < \pi/6$，正因如此，$\theta + 2\pi/3$ 在 (i) 中才是容許的，而 3θ 在 (ii) 中才落在 $\tan^{-1}$ 的值域之內。

STEP 之外

反三角函數（inverse trigonometric functions）讓我們初次體會到，要為一個並非一對一（one-to-one）的函數求反函數是怎樣的難題：我們選定一個*主支*（*principal branch*），而這類方程其實是在問我們身處哪一個分支。在複分析（complex analysis）中，$\log z$ 和 $z^{1/3}$ 也會出現同樣的問題：繞原點走一圈，便會由一個分支走到下一個分支，正如這裏把 θ 加上 $2\pi/3$ 一樣。Needham 的 *Visual Complex Analysis* 以圖像把分支解釋得十分精彩。

詳細解答

STEP I 2001，第 5 題

純數學 • 積分號下求導

STEP I 2001，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/01-s1-q5

關鍵思路

(i) 和 (ii) 的兩個積分都可以直接積出，第二個要先作代換（substitution）$u = 1 + tx$。真正的重點在於題末的那句話：(ii) 的被積函數（integrand）是 (i) 的被積函數對 t 的導數（derivative），而兩個答案之間也有同樣的關係。由此看來，在積分號內對參數（parameter）t 求導，似乎可以與對 x 積分交換次序。再做一次，然後代入 $t = 1$，便得到所要的猜想（conjecture）。

解答

全題中 $t > 0$ 且 $0 \le x \le 1$，所以 $1 + tx \ge 1 > 0$，各個被積函數都不會趨向無窮大。

第 (i) 部分

由於 $\dfrac{\mathrm{d}}{\mathrm{d}x}(1+tx)^{-1} = -t(1+tx)^{-2}$，$(1+tx)^{-2}$ 的一個原函數（antiderivative）是 $-\dfrac{1}{t(1+tx)}$（這裏用到 $t \neq 0$）。因此

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

這正是所要證明的結果。

第 (ii) 部分

分子中的 x 令直接積分變得不便，所以我們設 $u = 1 + tx$。於是 $x = \dfrac{u-1}{t}$，$\mathrm{d}x = \dfrac{\mathrm{d}u}{t}$，而積分上下限 $x = 0, 1$ 變成 $u = 1, 1+t$：

$$\begin{aligned}\int_0^1 \frac{-2x}{(1+tx)^3}\,\mathrm{d}x &= -\frac{2}{t^2}\int_1^{1+t}\frac{u-1}{u^3}\,\mathrm{d}u = -\frac{2}{t^2}\int_1^{1+t}\left(u^{-2}-u^{-3}\right)\mathrm{d}u \\ &= -\frac{2}{t^2}\left[-\frac{1}{u}+\frac{1}{2u^2}\right]_1^{1+t} = -\frac{2}{t^2}\left(\frac{1}{2}-\frac{1}{1+t}+\frac{1}{2(1+t)^2}\right).\end{aligned}$$

想一想 怎樣才能不必把所有項通分展開，就把括號內的式子化簡？

設 $w = \dfrac{1}{1+t}$，括號內的式子就是 $\frac{1}{2} - w + \frac{1}{2}w^2 = \frac{1}{2}(1-w)^2$，是一個完全平方（perfect square）。又因為 $1 - w = \dfrac{t}{1+t}$，所以

$$\int_0^1 \frac{-2x}{(1+tx)^3}\,\mathrm{d}x = -\frac{2}{t^2}\cdot\frac{1}{2}\cdot\frac{t^2}{(1+t)^2} = -\frac{1}{(1+t)^2},$$

這就是所求的結果。

猜想

把 x 固定，將被積函數看成 t 的函數：

$$\frac{\partial}{\partial t}(1+tx)^{-2} = -2x(1+tx)^{-3}.$$

所以 (ii) 的被積函數是 (i) 的被積函數對 t 的導數，正如 (ii) 的答案是 (i) 的答案的導數一樣。這提示了以下法則：

$$\frac{\mathrm{d}}{\mathrm{d}t}\int_0^1 f(x,t)\,\mathrm{d}x = \int_0^1 \frac{\partial f}{\partial t}\,\mathrm{d}x.$$

再應用一次，

$$\frac{\partial}{\partial t}\left(-2x(1+tx)^{-3}\right) = 6x^2(1+tx)^{-4}, \qquad \frac{\mathrm{d}}{\mathrm{d}t}\left(-(1+t)^{-2}\right) = 2(1+t)^{-3},$$

於是我們猜想，對於 $t > 0$，

$$\int_0^1 \frac{6x^2}{(1+tx)^4}\,\mathrm{d}x = \frac{2}{(1+t)^3}.$$

代入 $t = 1$，猜想的值為

$$\int_0^1 \frac{6x^2}{(1+x)^4}\,\mathrm{d}x = \frac{2}{8} = \frac{1}{4}.$$

驗算

- 當 $t \to 0$ 時，(i) 的被積函數趨向 1，而答案 $\frac{1}{1+t}$ 也趨向 1；在 (ii) 中，被積函數趨向 $-2x$，其積分為 -1，與 $-\frac{1}{(1+t)^2} \to -1$ 吻合。
- 題目沒有要求我們驗證這個猜想，但驗證起來很快：設 $u = 1 + x$，積分變成 $6\int_1^2 (u^{-2} - 2u^{-3} + u^{-4})\,\mathrm{d}u = 6\left[-\frac{1}{u} + \frac{1}{u^2} - \frac{1}{3u^3}\right]_1^2 = 6\left(-\frac{7}{24} + \frac{1}{3}\right) = \frac{1}{4}$。
- 用辛普森法則（Simpson's rule）計算最後一個積分，得 0.250000；另外取 25 個不同的 t 值，結果都與 $\frac{2}{(1+t)^3}$ 吻合至小數點後九位。

常見錯誤

- 在 (ii) 中漏掉由 $\mathrm{d}x = \mathrm{d}u/t$ 而來的因子 $\frac{1}{t}$：這裏共有兩個 $\frac{1}{t}$ 因子，一個來自 x，一個來自 $\mathrm{d}x$。
- 代換後沒有把積分上下限改為 $u = 1$ 和 $u = 1 + t$。
- 對 $-(1+t)^{-2}$ 求導時得出 $-2(1+t)^{-3}$：兩個負號互相抵消，所以猜想的值是正數（這也是必然的，因為被積函數是正的）。

STEP 之外

我們猜想出來的法則，就是積分號下求導（differentiation under the integral sign），又稱萊布尼茲法則（Leibniz's rule）。實分析（real analysis）中有一條定理指出：只要 f 和 $\partial f/\partial t$ 在一個閉矩形上連續（continuous），這條法則便成立；本題中，對於 $0 \le x \le 1$ 以及任何區間 $[a, b]$（其中 $a > 0$）內的 t，這個條件都成立。重複同樣的論證，可得對每一個 n 都有 $\int_0^1 \frac{x^n}{(1+tx)^{n+2}}\,\mathrm{d}x = \frac{1}{(n+1)(1+t)^{n+1}}$。理查德·費曼（Richard Feynman）曾用這一招計算其他方法都束手無策的積分，並因此為人津津樂道。至於交換極限（limit）、導數與積分次序背後的理論，可參閱 Bartle 與 Sherbert 的 *Introduction to Real Analysis*（實分析入門教材）。

詳細解答

STEP I 2001，第 6 題

純數學 • 旋轉曲面

> **STEP I 2001，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/01-s1-q6

本題在 *Beyond STEP* 第 4.8 節切球形麵包中有完整的引導式講解：https://beyondstep.keithmcnulty.org/zh-hant/ch04/spherical-loaf.html

關鍵思路

把半圓 $y = \sqrt{a^2 - x^2}$ 繞 x 軸旋轉，便得到球體。在表面積（surface area）公式中，被積函數（integrand）$y\sqrt{1 + (\mathrm{d}y/\mathrm{d}x)^2}$ 原來是常數，所以任何一片麵包上的麵包皮只取決於該片的厚度，而與它所在的位置無關。至於第二部分，我們從一端開始逐片吃掉（正如提示中的積分上下限 $-a$ 至 $-a + t$ 所暗示），因此剩下的是一個球冠（spherical cap）；比值 V/A 成為球冠厚度的二次函數，我們要求出它的最大值。

解答

麵包皮與片數成正比

設麵包的半徑為 a，中心位於原點，則其表面可由

$$y = \sqrt{a^2 - x^2}, \qquad -a \le x \le a,$$

繞 x 軸旋轉而得。求導後，當 $|x| < a$ 時 $\dfrac{\mathrm{d}y}{\mathrm{d}x} = -\dfrac{x}{\sqrt{a^2 - x^2}} = -\dfrac{x}{y}$，所以

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

其中用到 $x^2+y^2=a^2$ 及 $y>0$。因此，球體在平面 $x=x_1$ 與 $x=x_2$ 之間（其中 $-a\le x_1<x_2\le a$）的曲面面積為

$$2\pi\int_{x_1}^{x_2} a\,\mathrm{d}x = 2\pi a\,(x_2-x_1).$$

（被積函數只在端點 $x=\pm a$ 處沒有定義，但在該處仍然有界，所以這個積分沒有問題。）

現在假設麵包被切成 N 片，每片厚度為 $h=2a/N$。每一片無論位於何處，所帶的麵包皮面積都恰好是 $2\pi ah$。所以，如果剩下 n 片，不論它們在甚麼位置，剩餘的麵包皮面積都是

$$A = n\cdot 2\pi ah,$$

它與剩餘的片數 n 成正比。這正是我們要證明的結果。（這是阿基米德（Archimedes）的古老結果：球體與其外切圓柱（circumscribing cylinder）在任意兩個平行平面之間的面積相等。）

「脆皮」條件

麵包是從一端開始逐片吃掉的，所以吃掉若干片之後，剩下的麵包是由 $x=-a$ 至 $x=-a+t$ 的球冠，其中 t 滿足 $0<t\le 2a$。由第一部分可知 $A=2\pi at$。至於體積，作代換 $u=x+a$，使得 $y^2=a^2-x^2=(a-x)(a+x)=(2a-u)u$：

$$V=\pi\int_{-a}^{-a+t}(a^2-x^2)\,\mathrm{d}x=\pi\int_0^t(2au-u^2)\,\mathrm{d}u=\pi\Big(at^2-\frac{t^3}{3}\Big).$$

因此

$$\frac{V}{A}=\frac{\pi t^2\left(a-\frac{t}{3}\right)}{2\pi at}=\frac{t}{2}-\frac{t^2}{6a}=:f(t).$$

想一想 這個比值取決於已經吃掉多少。我們需要留意的是哪一個 t 值？

無論吃掉了多少片，條件都必須成立，所以我們需要 f 的最大值。利用配方法（completing the square），

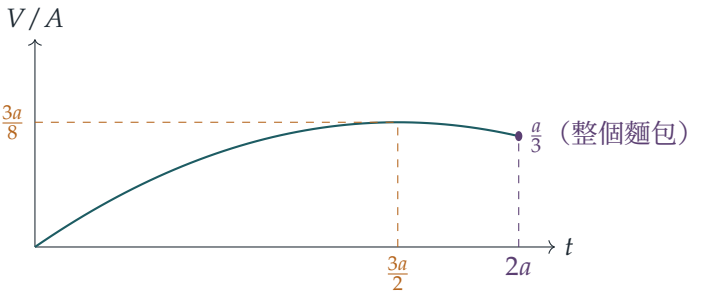
$$f(t)=\frac{3a}{8}-\frac{1}{6a}\Big(t-\frac{3a}{2}\Big)^2,$$

所以 f 在 $t=\frac{3a}{2}$ 時最大（這個值位於 $(0,2a]$ 之內），最大值為 $\frac{3a}{8}$。

$t=\frac{3}{2}a$ 是否真的會出現？會的，例如把麵包切成 4 片並吃掉一片（或切成 $4m$ 片並吃掉 m 片）。一個麵包若要作為平行切片麵包而稱得上「脆皮」，就必須在這樣的切法下符合規定，所以需要

$$\frac{3a}{8}<1,\qquad \text{即}\qquad a<\frac{8}{3}=2\tfrac{2}{3}\text{ 米},$$

正如題目所求。反過來說，若 $a<\frac{8}{3}$，則對每一個 t 都有 $f(t)\le\frac{3a}{8}<1$，所以無論怎樣切片，這個麵包都是「脆皮」的。



上圖顯示 $a = 2$ 時的 $f(t)$：當我們向麵包中央推進時，比值上升，其後回落，因為接近另一端時，我們加入了大量麵包皮，卻只加入少量麵包。

驗算

- 對整個麵包而言，$t = 2a$ 給出 $f(2a) = a - \frac{2a}{3} = \frac{a}{3}$，與 $\frac{4}{3}\pi a^3/(4\pi a^2)$ 吻合。
- 取 $a = \frac{8}{3}$，切成 4 片。剩下 1, 2, 3, 4 片時，比值分別為 $\frac{5}{9}, \frac{8}{9}, 1, \frac{8}{9}$，所以吃掉一片之後比值恰好是 1，剛好不滿足嚴格不等式（strict inequality）。
- 數值上，對面積公式作數值積分（quadrature），所得結果在隨機選取的切片上都與 $2\pi a(x_2 - x_1)$ 相符；而當 $a = 2.66$ 時，切成不多於 200 片的每一種切法，比值都保持在 1 以下。

常見錯誤

- 只考慮整個麵包：$a/3 < 1$ 給出 $a < 3$，這是必要條件，但弱得多。最壞的一刻是麵包寬度的四分之一被吃掉的時候。
- 把平坦的切面也算作麵包皮。麵包皮只是原來的曲面，這也正是題目所給公式所量度的。

STEP 之外

被積函數為常數，這就是阿基米德的帽盒定理（hat-box theorem）：把球面橫向投影到其外切圓柱上，面積保持不變。正因如此，在球面上均勻地選取一點，其高度服從均勻分佈（uniform distribution），這個事實在概率和物理中經常用到。像這樣的曲面積分（surface integral），以及映射在甚麼情況下保持面積，都會在向量微積分（vector calculus）的入門課程中詳細討論；Needham 的 *Visual Complex Analysis*（以幾何直觀講解複分析的著作）對球面的相關映射有一段優美的幾何討論。

詳細解答

STEP I 2001，第 7 題

純數學 • 凹性與微分方程

STEP I 2001，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/01-s1-q7

關鍵思路

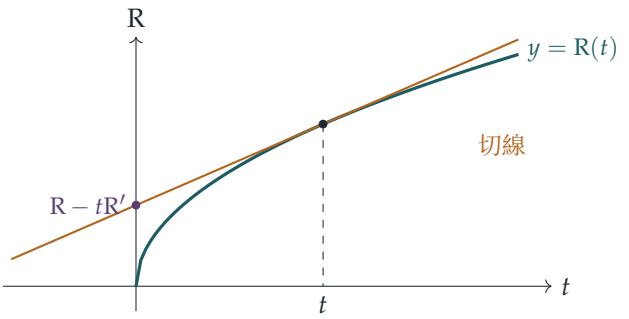
三個條件說明 R 由原點出發、一直上升，並且向下彎曲：它的圖像是凹（concave）的。凹的圖像位於它每一條切線（tangent）的下方，單憑這個幾何事實就可得出 (i) 的不等式。第 (ii) 和 (iii) 部分則是可分離變數的微分方程（separable differential equations），因為 $\mathrm{H} = \mathrm{R}'/\mathrm{R}$ 正是 $\ln \mathrm{R}$ 的導數。一旦明確求出 R，我們便把它逐一對照各個條件檢驗，而 (i) 的不等式提供了一個快捷的方法來排除某個模型。

解答

首先，有一個事實會貫穿全題：由於 $\mathrm{R}(0) = 0$，而且 R 在 $t > 0$ 時嚴格遞增（strictly increasing），所以當 $t > 0$ 時 $\mathrm{R}(t) > 0$。因此當 $t > 0$ 時，$\mathrm{H}(t)$ 有定義，且 $\mathrm{H}(t) > 0$。

第 (i) 部分

圖像通過原點，斜率（gradient）為正，而且斜率隨 t 增加而減少。



固定 $t > 0$。圖像在時間 t 處的切線斜率為 $\mathrm{R}'(t)$，所以它與縱軸相交於高度

$$\mathrm{R}(t) - t\,\mathrm{R}'(t).$$

從圖中可見這個截距（intercept）是正數，因為通過原點的曲線位於切線下方。要把這一點說得嚴謹，可對 R 在 $[0, t]$ 上應用中值定理（mean value theorem）：存在某個 c 滿足 $0 < c < t$，使得

$$\mathrm{R}(t) - \mathrm{R}(0) = t\,\mathrm{R}'(c), \qquad \text{所以} \qquad \mathrm{R}(t) = t\,\mathrm{R}'(c).$$

由於 $\mathrm{R}'' < 0$，函數 R' 嚴格遞減，而 $c < t$ 給出 $\mathrm{R}'(c) > \mathrm{R}'(t)$。因此

$$\mathrm{R}(t) > t\,\mathrm{R}'(t),$$

這正正就是「切線的截距在原點之上」這句話。兩邊除以 $\mathrm{R}'(t) > 0$，得

$$t < \frac{\mathrm{R}(t)}{\mathrm{R}'(t)} = \frac{1}{\mathrm{H}(t)},$$

這就是所要證明的結果。

（等價地說，切線與 t 軸相交於 $t - 1/\mathrm{H}(t)$，而這是負數。在宇宙學（cosmology）中，$1/\mathrm{H}$ 稱為「哈勃時間」（Hubble time），這個結果說明：在任何減速膨脹的模型中，宇宙的年齡都比它的哈勃時間短。）

第 (ii) 部分

當 $t > 0$ 時，$\mathrm{R} > 0$，所以可以寫成

$$\frac{\mathrm{d}}{\mathrm{d}t} \ln \mathrm{R}(t) = \frac{\mathrm{R}'(t)}{\mathrm{R}(t)} = \frac{a}{t}.$$

積分後得 $\ln \mathrm{R} = a \ln t + C$，所以

$$\mathrm{R}(t) = k\,t^a \qquad (t > 0),$$

其中 $k = \mathrm{e}^C > 0$。

現在利用 $k > 0$ 和 $t^a > 0$ 來檢驗各個條件：

- $\mathrm{R}'(t) = ka\,t^{a-1}$，它在 $t > 0$ 時為正，當且僅當 $a > 0$。
- $\mathrm{R}''(t) = ka(a-1)\,t^{a-2}$。已知 $a > 0$，它為負當且僅當 $a < 1$。
- 若 $a > 0$，則當 $t \to 0^+$ 時 $kt^a \to 0$，所以與 $\mathrm{R}(0) = 0$ 相容。（若 $a \le 0$，這個條件無論如何都不成立：kt^a 會趨向 $k \neq 0$，或者趨向無限大。）

所以條件 $(*)$ 成立，當且僅當

$$0 < a < 1.$$

這與第 (i) 部分吻合：在那裏 $t < 1/\mathrm{H} = t/a$，由此必有 $a < 1$。

第 (iii) 部分

現在 $\mathrm{H}(t) = b/t^2$。首先，當 $t > 0$ 時 $\mathrm{H}(t) > 0$（見開首的說明），所以必須有 $b > 0$。若 $b \le 0$，結論已經成立。

因此假設 $b > 0$。根據第 (i) 部分，每個 $t > 0$ 都必須滿足

$$t < \frac{1}{\mathrm{H}(t)} = \frac{t^2}{b}, \qquad \text{即} \qquad t > b$$

（兩邊除以 $t > 0$ 並乘以 $b > 0$）。但對於每個滿足 $0 < t \le b$ 的 t，這個不等式都不成立。所以沒有任何 b 的值能與 $(*)$ 相容。

想一想 這樣使用第 (i) 部分的結果是否合理？

合理。第 (i) 部分證明了*任何*滿足 $(*)$ 的函數，對所有 $t > 0$ 都有 $t < 1/\mathrm{H}(t)$。所以，假如某個建議的 H 哪怕只在一個 t 值上違反這個不等式，就沒有任何具有這個 H 的函數能滿足 $(*)$。這是一個反證法（proof by contradiction）。

若想直接驗證，可仿照 (ii) 解這個方程：$\ln \mathrm{R} = -b/t + C$，所以 $\mathrm{R}(t) = k\,\mathrm{e}^{-b/t}$，其中 $k > 0$。於是

$$\mathrm{R}'(t) = \frac{kb}{t^2}\,\mathrm{e}^{-b/t},$$
$$\mathrm{R}''(t) = kb\,\mathrm{e}^{-b/t}\left(\frac{b}{t^4} - \frac{2}{t^3}\right) = \frac{kb\,\mathrm{e}^{-b/t}}{t^4}\,(b - 2t).$$

當 $b > 0$ 時，$\mathrm{R}' > 0$，而且當 $t \to 0^+$ 時 $\mathrm{R}(t) \to 0$，所以頭兩個條件成立。但當 $0 < t < b/2$ 時 $\mathrm{R}''(t) > 0$，所以圖像在原點附近彎向了錯誤的方向，$(*)$ 不成立。

驗算

- 在 (ii) 中，取 $a = \frac{1}{2}$ 得 $\mathrm{R} = k\sqrt{t}$，正是草圖中的形狀：遞增而且是凹的，並有 $t < 1/\mathrm{H} = 2t$。取 $a = 1$ 則得一條通過原點的直線，此時恰好有 $t = 1/\mathrm{H}$：這是 (i) 的邊界情況。
- 在 (iii) 中取 $b = 1$、$k = 1$：數值上 $\mathrm{R}''(0.25) \approx 2.34 > 0$，而 $1/\mathrm{H}(0.5) = 0.25 < 0.5$，兩者都違反 $(*)$。

我們用電腦以幾個凹的測試函數檢驗了 (i)，以有限差分（finite differences）檢驗了 (ii) 和 (iii) 中的解及導數公式，並檢驗了 R'' 在 $t = b/2$ 處變號。

常見錯誤

- 在 (i) 中，只憑圖像作論證。「曲線位於切線下方」需要理由：中值定理，加上 R' 遞減。
- 在 (ii) 中，忘記了要有 $a > 0$ 才能保證 $\mathrm{R}' > 0$（以及 $\mathrm{R}(0) = 0$），結果只給出 $a < 1$。
- 在 (iii) 中，聲稱凹性在 t 很大時不成立。$b - 2t$ 的正負號顯示，出問題的地方是在 t 很小的時候。

STEP 之外

第 (i) 部分是關於凸（convex）函數和凹函數的一個核心事實的特殊情況：一個可微（differentiable）函數是凹的，當且僅當它的圖像位於所有切線的下方。這個事實是詹森不等式（Jensen's inequality）以及最優化（optimisation）中許多理論的基礎。這裏用中值定理的論證，是大學一年級分析學（analysis）的標準內容；可參閱 Abbott 的 *Understanding Analysis*。在真實的宇宙學中，$a = \frac{2}{3}$ 對應由物質主導的愛因斯坦–德西特（Einstein–de Sitter）宇宙，它的年齡是哈勃時間的三分之二。

詳細解答

STEP I 2001，第 8 題

純數學 • 二階線性方程

> **STEP I 2001，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/01-s1-q8

關鍵思路

把每個已知解代入方程，都會在每個 x 值處得到一條聯繫兩個未知函數 $p(x)$ 和 $q(x)$ 的方程。兩個解便給出兩條聯立線性方程（simultaneous linear equations），我們可以像解兩個未知數那樣把 p 和 q 解出來。在最後一部分，中間部分所證明的線性（linearity）讓我們可以把題目給出的兩個解換成它們較簡單的組合，而三角恆等式（trigonometric identities）令這些組合變得非常簡單。

解答

求 p 和 q

對於 $y = x$，有 $y' = 1$ 及 $y'' = 0$，所以代入 $(*)$ 得

$$p(x) + x\,q(x) = 0. \tag{1}$$

對於 $y = 1 - x^2$，有 $y' = -2x$ 及 $y'' = -2$，所以

$$-2 - 2x\,p(x) + (1 - x^2)\,q(x) = 0. \tag{2}$$

由 (1) 得 $p(x) = -x\,q(x)$。把它代入 (2)，

$$-2 + 2x^2 q(x) + (1 - x^2)q(x) = 0 \quad \Longrightarrow \quad (1 + x^2)\,q(x) = 2.$$

由於 $1 + x^2 > 0$，我們可以相除，所以 $q(x) = 2(1 + x^2)^{-1}$，然後由 (1) 得 $p(x) = -x\,q(x) = -2x(1 + x^2)^{-1}$，正如題目所求。（我們從沒有除以 x，所以結果對所有實數 x 都成立，包括 $x = 0$。）

任意組合 $ax + b(1 - x^2)$ 都是解

記 $L[y] = y'' + p\,y' + q\,y$。求導是線性的，所以若 $y = ay_1 + by_2$，其中 a, b 為常數，則

$$\begin{aligned} L[ay_1 + by_2] &= a\big(y_1'' + p\,y_1' + q\,y_1\big) + b\big(y_2'' + p\,y_2' + q\,y_2\big) \\ &= a\,L[y_1] + b\,L[y_2]. \end{aligned}$$

取 $y_1 = x$ 及 $y_2 = 1 - x^2$，則 $L[y_1]$ 和 $L[y_2]$ 都等於零（這正是它們滿足 $(*)$ 的意思），所以對所有常數 a 和 b，都有 $L[ax + b(1 - x^2)] = 0$。如果你想直接驗證：$y = ax + b(1 - x^2)$ 有 $y' = a - 2bx$ 及 $y'' = -2b$，而

$$\begin{aligned} y'' + p\,y' + q\,y &= -2b - \frac{2x(a - 2bx)}{1 + x^2} + \frac{2\big(ax + b - bx^2\big)}{1 + x^2} \\ &= \frac{-2b(1 + x^2) - 2ax + 4bx^2 + 2ax + 2b - 2bx^2}{1 + x^2} = 0. \end{aligned}$$

最後一部分

剛才的論證完全沒有用到 p 和 q 的具體形式：對於任何形如 $(*)$ 的方程，兩個解的組合 $ay_1 + by_2$ 仍然是解。所以，如果 $\cos^2(\frac{1}{2}x^2)$ 和 $\sin^2(\frac{1}{2}x^2)$ 都滿足 $(*)$，它們的和與差也同樣滿足，

$$\cos^2(\tfrac{1}{2}x^2) + \sin^2(\tfrac{1}{2}x^2) = 1, \qquad \cos^2(\tfrac{1}{2}x^2) - \sin^2(\tfrac{1}{2}x^2) = \cos(x^2),$$

這裏用了 $\cos^2\theta + \sin^2\theta = 1$ 和 $\cos^2\theta - \sin^2\theta = \cos 2\theta$，其中 $\theta = \frac{1}{2}x^2$。

> ? **想一想** 為甚麼可以用 1 和 $\cos(x^2)$ 來代替題目給出的函數？

因為這一步是可逆的：題目給出的函數是 $\frac{1}{2}\big(1 + \cos x^2\big)$ 和 $\frac{1}{2}\big(1 - \cos x^2\big)$，所以 p 和 q 使這兩個函數都成為解，當且僅當它們使 1 和 $\cos(x^2)$ 都成為解。我們沒有損失任何資訊，而導數卻容易計算得多。

解 $y = 1$。這時 $y' = y'' = 0$，所以 $(*)$ 簡化為 $q(x) \cdot 1 = 0$，得

$$q(x) = 0.$$

解 $y = \cos(x^2)$。由鏈式法則（chain rule）得 $y' = -2x\sin(x^2)$，再由積法則（product rule）得

$$y'' = -2\sin(x^2) - 4x^2\cos(x^2).$$

當 $q = 0$ 時，$(*)$ 變成

$$-2\sin(x^2) - 4x^2\cos(x^2) - 2x\sin(x^2)\,p(x) = 0.$$

凡是 $x\sin(x^2) \neq 0$ 之處，我們都可以相除，於是

$$p(x) = -\frac{\sin(x^2) + 2x^2\cos(x^2)}{x\sin(x^2)} = -\frac{1}{x} - 2x\cot(x^2).$$

所以 $p(x) = -\dfrac{1}{x} - 2x\cot(x^2)$ 及 $q(x) = 0$。

在 $x\sin(x^2) = 0$ 的各點（即 $x = 0$ 和 $x = \pm\sqrt{n\pi}$），p 的係數為零，所以方程對該處的 p 沒有任何限制；事實上，我們求得的 p 在這些點趨向無窮。這並不是出錯：在這些點上，題目給出的兩個函數的導數都是零，所以它們不可能是一條性質良好、形如 $(*)$ 的方程的兩個「獨立」解，而 p 在該處必定是奇異（singular）的。

驗算

- 作為第一部分的獨立驗算，兩個解的朗斯基行列式（*Wronskian*）$W = y_1y_2' - y_2y_1'$ 滿足 $W' = -pW$。取 $y_1 = x$，$y_2 = 1 - x^2$，得 $W = -2x^2 - (1 - x^2) = -(1 + x^2)$，所以 $-p = W'/W = 2x/(1 + x^2)$，與我們求得的 p 吻合。至於最後一部分，解 1 和 $\cos(x^2)$ 給出 $W = -2x\sin(x^2)$，而同樣有 $W'/W = 1/x + 2x\cot(x^2) = -p$。
- 數值上，在 $x = 0.9$ 處，把題目給出的兩個函數代入 $(*)$，並以有限差分（finite differences）計算導數，可迫使 $p(0.9) \approx -2.82465$ 及 $q(0.9) \approx 0$，與 $-1/0.9 - 1.8\cot(0.81) \approx -2.82465$ 相符。

常見錯誤

- 直接對 $\cos^2(\frac{1}{2}x^2)$ 求導兩次，結果迷失在繁複的代數運算之中。應先用倍角恆等式（double-angle identities），然後才求導。
- 把 (1) 除以 x 而得到 $q = -p/x$；這樣做是可行的，但之後必須另外處理 $x = 0$ 的情況。改為消去 p 便可避開這個問題。
- 忘記說明最後一部分中 p 的公式在哪些地方成立。

STEP 之外

這道題讓我們一窺二階線性微分方程（second-order linear differential equations）的理論：解構成一個二維向量空間（vector space），任意兩個獨立的解便可決定方程，而朗斯基行列式（連同阿貝爾（Abel）公式 $W' = -pW$）則用來衡量獨立性。朗斯基行列式為零的點，例如本題的 $x = \pm\sqrt{n\pi}$，正是係數必定變得奇異之處。Boyce 與 DiPrima 的 *Elementary Differential Equations*（微分方程入門教材）對這一切都有細緻的闡述。

詳細解答

STEP I 2001，第 9 題

力學 • 攔截與相對速度

STEP I 2001，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/01-s1-q9

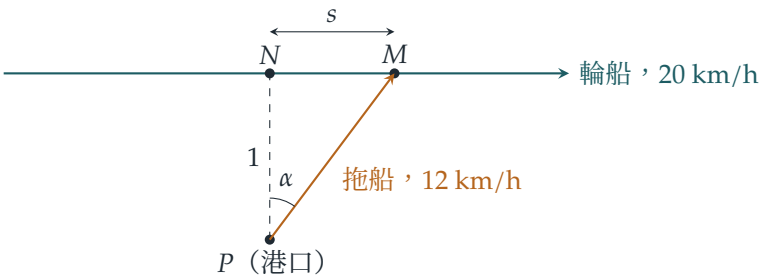
關鍵思路

要盡量遲出發，拖船就必須全速沿直線前進，所以真正要選擇的，只是在輪船航線上的哪一點與它會合。對每一個可能的會合點，先求出拖船最遲可以在甚麼時間出發而仍然及時趕到，然後對會合點求這個出發時間的最大值。只要把問題設定妥當，這就是一個單變量的最優化（optimisation）問題。

解答

建立模型

把輪船和拖船都視為質點（particle）。把港口 P 放在原點，並設輪船沿直線 $y = 1$ 向正 x 方向航行（距離以公里為單位，時間以小時為單位）。由 P 到航線的垂足（foot of the perpendicular）是 $N = (0, 1)$，即輪船最接近港口的一點。



假設兩船在點 $M = (s, 1)$ 會合。輪船沿航線的位置是時間的線性函數（linear function），

所以它到達 M 的時間是

$$t_M = \frac{s-c}{20},$$

其中 c 是一個常數，由輪船的起始位置決定。（輪船仍在駛近，所以 c 小於我們會用到的任何會合位置。）

想一想 為甚麼拖船必須全速沿直線前進？

如果拖船在時間 T 出發，並在時間 t_M 到達 M，它就以不超過 12 km/h 的速率走了至少等於直線距離 $PM = \sqrt{s^2+1}$ 的路程，所以 $t_M - T \geq \sqrt{s^2+1}/12$。以最高速率直駛往 M，等號便成立。任何較慢或較迂迴的走法，都會迫使拖船提早出發。

因此，能夠在 M 會合的最遲出發時間是

$$T(s) = \frac{s-c}{20} - \frac{\sqrt{s^2+1}}{12},$$

而我們要求的是 $T(s)$ 在所有 s 之中的最大值。

求出發時間的最大值

求導數（derivative），得

$$T'(s) = \frac{1}{20} - \frac{s}{12\sqrt{s^2+1}}.$$

令 $T'(s) = 0$，得

$$\frac{s}{\sqrt{s^2+1}} = \frac{12}{20} = \frac{3}{5}.$$

這要求 $s > 0$。兩邊平方，$25s^2 = 9(s^2+1)$，所以 $s^2 = \frac{9}{16}$，$s = \frac{3}{4}$。

這是極大值，而且是全局最大值（global maximum）：$\sqrt{s^2+1}$ 是凸函數（convex function），所以 T（一個線性函數減去一個凸函數）是凹函數（concave function）。換一個說法，$\dfrac{s}{\sqrt{s^2+1}}$ 隨 s 遞增，所以當 $s < \frac{3}{4}$ 時 $T'(s) > 0$，當 $s > \frac{3}{4}$ 時 $T'(s) < 0$。留意常數 c 已經消失：會合點與輪船從哪裏出發無關。

所以拖船在越過 N 之後 $\frac{3}{4}$ 公里處與輪船會合，而它所走的距離是

$$PM = \sqrt{\tfrac{9}{16}+1} = \sqrt{\tfrac{25}{16}} \quad \Longrightarrow \quad \boxed{PM = \tfrac{5}{4} \text{ 公里} = 1.25 \text{ 公里}.}$$

拖船在海上航行 $\frac{5/4}{12} = \frac{5}{48}$ 小時，即 $6\frac{1}{4}$ 分鐘。

幾何解讀

拖船的航向與 PN 成角 α，其中 $\sin\alpha = \frac{s}{\sqrt{s^2+1}} = \frac{3}{5}$，所以 $\sin\alpha$ 正是兩船速率之比 12/20。拖船的速度（velocity）是 $(7.2, 9.6)$，而它相對於輪船的速度是 $(7.2-20, 9.6) = (-12.8, 9.6)$。兩者的點積（dot product）是 $-92.16 + 92.16 = 0$。在臨界出發時間，拖

船的速度與它相對於輪船的速度互相垂直。這是較慢的追趕者僅僅來得及攔截時的標準條件。

驗算

- 以幾個不同的輪船起始位置，在一個細密的會合點網格上作窮舉搜尋，$T(s)$ 的最大值總是出現在 $s = 0.75$，所以拖船所走的距離是 1.25 公里。
- 由最遲出發時間開始，以航向 $(\frac{3}{5}, \frac{4}{5})$ 用數值方法逐步模擬運動，拖船會到達輪船所在之處。只要遲 0.002 小時出發，無論取甚麼航向，都不能駛到輪船 20 米以內。
- 合理性檢查：假如拖船與輪船一樣快，$\sin\alpha$ 就會是 1，會合點會跑到無窮遠處。這是合理的，因為速率相同的拖船可以無止境地追着輪船跑。

常見錯誤

- 在輪船最接近港口的一點 N 與它會合（拖船所走的距離為 1 公里）。這樣做是把拖船的距離減至最小，而不是把出發時間推到最遲：拖船向前多走一點，就能賺到時間，因為輪船還要航行一段才到那裏。
- 只把拖船的航行時間 $\sqrt{s^2+1}/12$ 減至最小，而漏掉 $s/20$ 一項。要求最大值的量是出發時間，它同時牽涉兩艘船。

STEP 之外

求到達時間之差的最大值，是最優控制（optimal control）或追逐問題（pursuit problem）的一個小例子。速率之比的條件 $\sin\alpha = 12/20$，與超音速飛機後方的馬赫錐（Mach cone）是同一種幾何：圓錐半頂角的正弦等於音速與飛機速率之比，而這裏那艘慢的拖船，就扮演着聲音追趕快速聲源的角色。Kleppner 與 Kolenkow 的 *An Introduction to Mechanics*（力學入門）仔細討論了相對速度（relative velocity）和參考系（frame of reference），而 Siklos 的 *Advanced Problems in Mathematics*（高等數學難題）則有更多 STEP 風格的力學最優化題目。

詳細解答

STEP I 2001，第 10 題

力學 • 拋體的飛行時間

> **STEP I 2001，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/01-s1-q10

關鍵思路

把選擇分成兩部分：仰角決定了飛行時間和水平射程，而方向只決定炮彈落在以炮為圓心的圓上的哪一點。因此，炮彈能擊中直線，當且僅當其射程至少為 d。飛行時間隨仰角增大而增加，但仰角超過 $45°$ 後，射程會縮短，所以我們把仰角一直調高，直到射程剛好降至 d 為止。

解答

給定仰角時的飛行時間與射程

把炮彈當作只受重力作用的質點（particle）（不計空氣阻力），由地面發射，並落在平原上。設發射的仰角（elevation）為 α，其中 $0 < \alpha \leq \frac{\pi}{2}$。它的垂直速度分量是 $v\sin\alpha$，水平分量是 $v\cos\alpha$。

在垂直方向，經過時間 t 後它的高度是 $vt\sin\alpha - \frac{1}{2}gt^2$，而高度再次為零的時刻是

$$T = \frac{2v\sin\alpha}{g}.$$

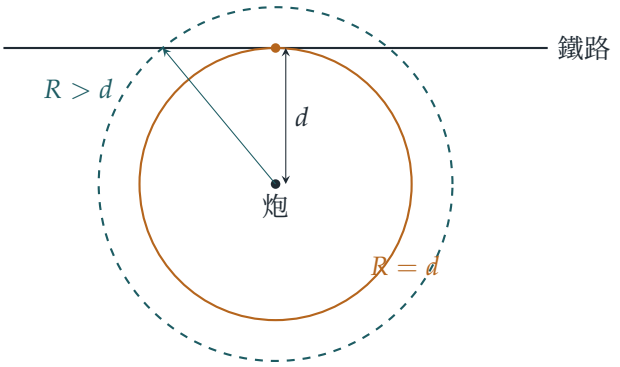
在這段時間內，它走過的水平距離是

$$R = v\cos\alpha \cdot \frac{2v\sin\alpha}{g} = \frac{v^2\sin 2\alpha}{g}.$$

炮彈何時能擊中直線？

仰角固定時，轉動炮身會使落點沿着一個以炮為圓心、半徑為 R 的圓移動。直線上最近的一點與炮的距離為 d，而直線上其他各點都更遠。所以這個圓與直線相交，當且僅當

$R \geq d$。當 $R = d$ 時，圓剛好與直線相切，炮必須正對着最近的那一點，即與直線垂直。



（俯視圖。虛線圓對應較低的仰角，它在兩點到達直線。實線圓對應最優的仰角，它剛好到達直線。）

所以炮手以仰角 α 發射時能擊中直線，當且僅當

$$\sin 2\alpha \geq \frac{gd}{v^2}.$$

記 $k = gd/v^2$。題目告訴我們 $v^2 > gd$，所以 $0 < k < 1$，確實有一些仰角可行。（最大射程是 v^2/g，在 45° 時取得，所以這個條件只是說直線在射程之內。）

使飛行時間最大

由於 $T = 2v \sin\alpha / g$，而 $\sin\alpha$ 在 $(0, \frac{\pi}{2}]$ 上是遞增的，要使 T 最大，就要選取容許範圍內最大的 α。

當 $0 < 2\alpha \leq \pi$ 時，條件 $\sin 2\alpha \geq k$ 成立，當且僅當

$$\arcsin k \leq 2\alpha \leq \pi - \arcsin k.$$

所以容許的最大仰角是 $\alpha = \frac{\pi}{2} - \frac{1}{2}\arcsin k$，它大於 45°。在這個仰角下 $\sin 2\alpha = k$，而由於 2α 位於 $(\frac{\pi}{2}, \pi)$ 之內，$\cos 2\alpha$ 是負數：

$$\cos 2\alpha = -\sqrt{1 - k^2}.$$

想一想 沿着直線瞄準遠處某一點，讓炮彈飛得更遠，不是更好嗎？

不是。飛得更遠需要更大的射程，而在 45° 以上，這意味着更低的仰角，飛行時間因而更短。飛行時間只取決於仰角，所以我們能做到的最好結果，就是仍能到達直線的最陡一擊。它的射程剛好是 d，所以炮彈落在由炮到直線的垂足上。

現在把 T 平方，並利用 $2\sin^2\alpha = 1 - \cos 2\alpha$：

$$\begin{aligned} g^2T^2 &= 4v^2\sin^2\alpha \\ &= 2v^2\left(1-\cos 2\alpha\right) \\ &= 2v^2 + 2v^2\sqrt{1-\frac{g^2d^2}{v^4}} \\ &= 2v^2 + 2\left(v^4 - g^2d^2\right)^{1/2}, \end{aligned}$$

其中把 v^2 移入了平方根之內（因為 $v^2 > 0$，這是容許的）。這就是所要求的結果。

驗算

- 當 $d \to 0$ 時，公式給出 $g^2T^2 \to 4v^2$，所以 $T \to 2v/g$。這相當於垂直向上發射，是最長的飛行時間；當直線緊貼着炮經過時，這正是應有的結果。
- 當 $v^2 = gd$ 時，直線僅僅在射程之內：唯一可行的是以 $45°$ 發射，而公式給出 $g^2T^2 = 2v^2$，與 $T = 2v\sin 45°/g = \sqrt{2}\,v/g$ 吻合。
- 以數值計算，取 $g = 9.8\,\mathrm{m\,s^{-2}}$、$d = 100\,\mathrm{m}$ 和 $v = 40\,\mathrm{m\,s^{-1}}$，最佳仰角約為 $71.1°$，公式給出 $T \approx 7.72\,\mathrm{s}$。以數值方法逐步模擬運動，得到相同的時間，炮彈亦落在直線上。對所有仰角作窮舉搜尋，在數百個隨機例子中都與公式一致。

常見錯誤

- 取了較小的根 $\alpha = \frac{1}{2}\arcsin k$。這樣發射同樣能擊中直線，但得到的是最短而不是最長的飛行時間：實際上就是取錯了 $\cos 2\alpha$ 的正負號。
- 一開始就假設炮彈必定落在直線上最近的一點。這個結論最終是對的，但需要上面的論證：其他任何一點都需要更大的射程，因而需要更低的仰角。

STEP 之外

這是一個小型的約束最優化（constrained optimisation）問題：在某個約束條件（到達直線）之下，使某個量（飛行時間）最大，而最優解位於容許範圍的邊界上。在大學一年級的力學和最優化課程中，你會遇到處理這類問題的一般工具，例如拉格朗日乘數（Lagrange multipliers），以及與之相關的概念：所有軌跡的包絡線（envelope）（即「安全拋物線」（parabola of safety）），它圍住了炮所能到達的每一點。Kleppner 與 Kolenkow 的 *An Introduction to Mechanics*（力學入門）是深入學習拋體與質點力學的好去處。

詳細解答

STEP I 2001，第 11 題

力學 • 彈性繩與能量

STEP I 2001，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：

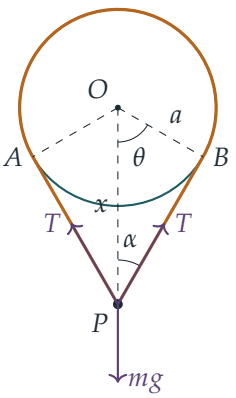
step.maths.org/questions/01-s1-q11

關鍵思路

質點把彈性帶向下拉之後，彈性帶大部分長度仍然貼在圓柱上，並在兩側沿切線方向離開圓柱，再筆直地連到質點。因此，整個問題歸結為由圓外一點向圓所作兩條切線（tangent）的幾何。由於圓柱是光滑的，彈性帶上各處的張力（tension）都相同，所以用總伸長量（extension）代入胡克定律（Hooke's law）便可求得張力。之後，第 (i) 部分是能量問題（質點在最低點瞬時靜止），第 (ii) 部分是力的問題（速率最大即加速度為零）。

解答

我們把質點視為點質量，把彈性帶視為輕而有彈性（遵從胡克定律），並把圓柱視為光滑。設 O 為橫截面的圓心，並設質點 P 位於 O 下方距離 x 處，其中 $x \geq a$。彈性帶在 A 和 B 兩點離開圓，其中 PA 和 PB 為切線。



幾何關係與張力

三角形 OAP 在 A 處為直角（切線垂直於半徑），所以

$$PA = PB = \sqrt{x^2 - a^2}, \qquad \cos\theta = \frac{a}{x}, \qquad \sin\alpha = \frac{a}{x},$$

其中 $\theta = \angle AOP$，而 $\alpha = \angle APO$，即每段筆直的彈性帶與鉛垂線所成的角。除了最接近 P 的一段圓心角為 2θ 的弧 AB 之外，彈性帶處處與圓接觸，所以它的長度是

$$L = (2\pi a - 2a\theta) + 2\sqrt{x^2 - a^2}.$$

自然長度（natural length）為 $2\pi a$，所以伸長量是

$$e(x) = 2\left(\sqrt{x^2 - a^2} - a\theta\right), \qquad \theta = \arccos\frac{a}{x}.$$

由於圓柱是光滑的，沿彈性帶方向沒有摩擦力，所以張力 T 處處相同，而根據胡克定律，$T = \dfrac{\lambda e}{2\pi a}$。

當 $x = 2a$ 時，我們有 $\cos\theta = \frac{1}{2}$，所以 $\theta = \frac{\pi}{3}$，$\alpha = \frac{\pi}{6}$，而 $\sqrt{x^2 - a^2} = \sqrt{3}\,a$。因此

$$e(2a) = 2\left(\sqrt{3}\,a - \frac{\pi a}{3}\right) = \frac{2a}{3}\left(3\sqrt{3} - \pi\right).$$

第 (i) 部分

釋放時，質點位於圓柱的底部，即 $x = a$，彈性帶未被拉長，質點處於靜止。到了 $x = 2a$，質點再次靜止（它不再往下掉），所以在這兩個時刻它都沒有動能（kinetic energy）。圓柱是光滑的，所以沒有能量損失，重力勢能（gravitational potential energy）的減少等於彈性勢能（elastic energy）的增加。質點下降了距離 $2a - a = a$，而自然長度為 $2\pi a$、伸長量為 e 的彈性帶，其彈性勢能為 $\dfrac{\lambda e^2}{2 \cdot 2\pi a}$。所以

$$mga = \frac{\lambda}{4\pi a} \cdot \frac{4a^2}{9}\left(3\sqrt{3} - \pi\right)^2 = \frac{\lambda a\left(3\sqrt{3} - \pi\right)^2}{9\pi},$$

由此得

$$\lambda = \frac{9\pi mg}{\left(3\sqrt{3} - \pi\right)^2},$$

正如所求。

> **想一想** 彈性帶有一部分繞在圓柱上，對整條彈性帶使用彈性勢能公式，這樣做合理嗎？

合理。由於圓柱是光滑的，無論是否繞在圓柱上，整條彈性帶的張力都相同，因而每單位長度的伸長也相同。彈性帶的表現，與一條自然長度為 $2\pi a$、總伸長量為 e 的彈性繩完全一樣，而這正是公式 $\lambda e^2/(2l)$ 所描述的情況。

第 (ii) 部分

加速度為零時，速率最大，也就是說，作用於質點的合力為零。在水平方向，兩個張力因對稱而互相抵消。在 P 處沿鉛垂方向分解，

$$2T\cos\alpha = mg.$$

當 $x = 2a$ 時，$\cos\alpha = \cos\frac{\pi}{6} = \frac{\sqrt{3}}{2}$，所以 $\sqrt{3}\,T = mg$，得 $T = \dfrac{mg}{\sqrt{3}}$。利用胡克定律和上面求得的伸長量，

$$T = \frac{\lambda}{2\pi a}\cdot\frac{2a}{3}\left(3\sqrt{3}-\pi\right) = \frac{\lambda\left(3\sqrt{3}-\pi\right)}{3\pi}.$$

令 T 的兩個表達式相等，

$$\lambda = \frac{3\pi mg}{\sqrt{3}\left(3\sqrt{3}-\pi\right)} = \frac{\sqrt{3}\,\pi mg}{3\sqrt{3}-\pi}.$$

? **想一想** 我們怎樣知道這是速率的極大值，而不僅僅是一個駐點？

當 x 增加時，e 會增加（每段筆直的彈性帶變長的速度，比弧長 $a\theta$ 縮短的速度快），而 $\cos\alpha = \sqrt{1-a^2/x^2}$ 也會增加，所以向上的力 $2T\cos\alpha$ 由 0 開始持續增加。質點在 $x = 2a$ 上方時，向下的合力為正，在其下方時則為負，所以質點一直加速至 $x = 2a$，之後便減速。

驗算

- 一個巧妙的一致性檢驗：對 L 求導，得 $\dfrac{\mathrm{d}L}{\mathrm{d}x} = \dfrac{2x}{\sqrt{x^2-a^2}} - \dfrac{2a^2}{x\sqrt{x^2-a^2}} = \dfrac{2\sqrt{x^2-a^2}}{x} = 2\cos\alpha$。所以彈性勢能的變化率 $T\dfrac{\mathrm{d}L}{\mathrm{d}x}$ 等於 $2T\cos\alpha$，即向上的力。(i) 的能量方法與 (ii) 的力學方法互相吻合。
- 以數值計算，(i) 中 $\lambda \approx 6.70\,mg$，(ii) 中 $\lambda \approx 2.65\,mg$。(ii) 中的彈性帶較弱，這是合理的：在那裏質點到達 $x = 2a$ 時仍在運動，並會繼續越過該點，而在 (i) 中，較硬的彈性帶已經令質點停下來。
- 由 $x = a$ 處的靜止狀態出發，以數值方法逐步求解運動方程 $m\ddot{x} = mg - 2T\cos\alpha$，取 λ 的第一個值時，最低點在 $x = 2.0000a$；取第二個值時，速率最大的位置在 $x = 2.0000a$。

常見錯誤

- 忘記了有一部分彈性帶離開了圓柱：伸長量是 $2\sqrt{x^2-a^2}$ 減去失去的弧長 $2a\theta$，而不僅僅是兩段筆直部分的長度。
- 由軸線而非釋放點起量度下降的距離：質點下降的距離是 a，而不是 $2a$。
- 沿鉛垂方向分解時，用了 $\sin\alpha$ 而不是 $\cos\alpha$；α 是由鉛垂線起量度的，而 $\sin\alpha = a/x$。

STEP 之外

驗算 d(彈性勢能)/dx 等於回復力，其實是虛功原理（principle of virtual work）的一個例子：在保守系統（conservative system）中，任何方向上的力，等於勢能沿該方向的變化率的負值。這種以一個坐標及其勢能來描述系統的觀點，正是拉格朗日力學（Lagrangian mechanics）的起點，你會在大學力學課程的第一或第二年遇到它。若想看對能量、力和勢的嚴謹論述，可參閱 Kleppner and Kolenkow, *An Introduction to Mechanics*（力學入門教科書）。

詳細解答

STEP I 2001，第 12 題

概率與統計 • 樹形圖

STEP I 2001，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/01-s1-q12

關鍵思路

一切取決於仔細記錄誰知道甚麼。一名學生要知道某人已聽過謠言，只有兩種情況：那人把謠言告訴了他，或者他把謠言告訴了那人；沒有人會透露自己是從誰那裏聽來的。只要記下每名學生在每個階段所知道的事，樹形圖（tree diagram）其實很小：每次分支都是在兩個人之間選擇，所以只有八條路徑，而且結果它們的可能性全部相等。

解答

設定

以 A、B、C、D 表示四名學生。在任何階段，學生 X 知道已聽過謠言的人包括：X 自己、曾把謠言告訴 X 的人，以及 X 曾把謠言告訴的人。當 X 聽到謠言時，X 會在這個集合以外的人當中均勻隨機（uniformly at random）地選一人。如果這個集合已包含全部四名學生，謠言便在 X 處停止。

開始時，A 已告訴 C，所以 A 知道 $\{A, C\}$，而 C 知道 $\{A, C\}$。

樹形圖

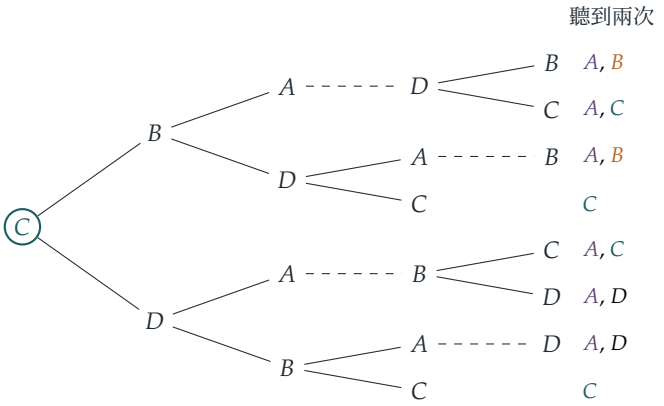
C 會略過 A，告訴 B 或 D，概率（probability）各為 $\frac{1}{2}$。先看分支 $C \to B$。這時 C 知道 $\{A, B, C\}$，而 B 知道 $\{B, C\}$（不包括 A，因為 C 沒有說謠言從何而來）。所以 B 會告訴 A 或 D，概率各為 $\frac{1}{2}$。

- $B \to A$。Arthur 現在知道 $\{A, B, C\}$，所以他必須告訴 D，之後 Arthur 便知道人人都已聽過。D 只知道 $\{A, D\}$，所以 D 會告訴 B 或 C，概率各為 $\frac{1}{2}$。無論是哪一個，

聽者的集合都會齊全：B 本來已知道 $\{A,B,C\}$，C 亦然，而兩人此時都得知 D 已聽過。於是謠言停止。

- $B \to D$。這時 B 知道 $\{B,C,D\}$，而 D 知道 $\{B,D\}$，所以 D 會告訴 A 或 C。
  - $D \to C$：C 本來已知道 $\{A,B,C\}$，現在又得知 D 已聽過，所以謠言停止。Arthur 始終沒有再次聽到謠言。
  - $D \to A$：Arthur 知道 $\{A,C,D\}$，所以他告訴 B，後者隨即知道四人都已聽過。謠言停止。

分支 $C \to D$ 的情況相同，只是 B 與 D 的角色互換。以下是整個樹形圖。除了以虛線表示的必然步驟（概率為 1）外，每條分支的概率都是 $\frac{1}{2}$。



每條通往樹葉的路徑都恰好包含三次概率為 $\frac{1}{2}$ 的選擇（必然步驟只貢獻因子 1），所以八個結果的概率各為 $\frac{1}{8}$。右方的標記顯示在該路徑上誰聽到謠言兩次。（在每條路徑上，一名學生最多聽到兩次。）

想一想 為甚麼 B 剛從 C 那裏聽到謠言，卻不知道 Arthur 已聽過？

因為 C 不會說是誰告訴自己的。B 所能推斷的，只是 C 已聽過謠言。正是這一點令謠言得以繼續流傳：在 $C \to B \to A$ 這條分支上，Arthur 被告知一件他早已知道的事。

Arthur 再次聽到謠言

除了兩條短路徑 $C \to B \to D \to C$ 和 $C \to D \to B \to C$ 之外，Arthur 在每條路徑上都會再次聽到謠言。所以

$$\mathrm{P}(\text{Arthur 再次聽到謠言}) = 1 - 2 \times \tfrac{1}{8} = \tfrac{3}{4},$$

正如所求。

Bertha 與 Chandra

Bertha 只在路徑 $C \to B \to A \to D \to B$ 和 $C \to B \to D \to A \to B$ 上聽到兩次，所以

$$\mathrm{P}(\text{Bertha 聽到兩次}) = \tfrac{2}{8} = \tfrac{1}{4}.$$

Chandra 在四條路徑上聽到兩次：樹形圖的上下兩半各有兩條（$C \to B \to A \to D \to C$、$C \to B \to D \to C$、$C \to D \to A \to B \to C$ 和 $C \to D \to B \to C$），所以

$$\mathrm{P}(\text{Chandra 聽到兩次}) = \tfrac{4}{8} = \tfrac{1}{2}.$$

Chandra 比 Bertha 更有可能再次聽到謠言，是因為 Chandra 很早便聽到：在傳話鏈中較後才被告知的人並不知道 Chandra 已聽過，所以直到最後 Chandra 仍然是可能被告知的對象。

驗算

根據 B 與 D 之間的對稱性（symmetry），Delilah 聽到兩次的概率也應是 $\frac{1}{4}$，而樹形圖中確有兩條這樣的路徑。各結果的概率之和正確：八片等可能的樹葉，總概率為 1。以電腦對整個過程作精確枚舉，得出 $\frac{3}{4}$、$\frac{1}{4}$ 和 $\frac{1}{2}$；而模擬（simulation）200 000 次謠言傳播，則得出 0.750、0.250 和 0.501。

常見錯誤

- 讓聽者知道傳話鏈中更早的每一個人。其實傳遞下去的只有告知者的身分；假如人人都掌握全部資訊，謠言便永遠不會回到 Arthur 那裏。
- 忘記告知者同樣會得知一些事情：告訴 Y 之後，X 便知道 Y 已聽過謠言，而 Arthur 的步驟之所以是必然的，正是這個緣故。
- Arthur 一再次聽到謠言便把樹形圖停下。其後的分支對 Bertha 和 Chandra 的計算是必需的。

STEP 之外

這個謠言是四名學生之間的一個隨機漫步（random walk），其規則取決於過往歷史，而這裏以每名學生所知道的事來記錄歷史。只要把狀態設定得足夠豐富，使下一步只取決於當前狀態，它便成為一條馬可夫鏈（Markov chain），亦即隨機過程（random processes）入門課程的核心對象；謠言模型和傳染病模型是它的經典應用。Grimmett and Welsh, *Probability: An Introduction*（概率論入門）對馬可夫鏈的講解，程度恰到好處。

詳細解答

STEP I 2001，第 13 題

概率與統計 • 條件泊松分佈

> **STEP I 2001，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/01-s1-q13

關鍵思路

以打破碟子的總數作為條件，便可把一道泊松（Poisson）問題化為二項（binomial）問題。假如四名學生各自獨立地以相同的比率打破碟子，那麼一旦知道總共打破了多少隻碟子，每一隻破碟由四人中任何一人打破的可能性都相等，而且與其他破碟互相獨立。因此，數學家在破碟中所佔的數目服從二項分佈，而未知的 λ 亦隨之消失。

解答

設定

設 M 為數學家打破的碟子數目，R 為其餘三名學生打破的碟子總數。題目指出每名學生打破的碟子數目均服從 $\mathrm{Po}(\lambda)$，且彼此獨立。所以 $M \sim \mathrm{Po}(\lambda)$，而 R 是三個獨立的 $\mathrm{Po}(\lambda)$ 變量之和，故服從 $\mathrm{Po}(3\lambda)$。此外，M 與 R 獨立（independent），因為 R 只取決於其餘三名學生。總數為 $T = M + R \sim \mathrm{Po}(4\lambda)$。

? **想一想** 為甚麼獨立泊松變量之和仍然服從泊松分佈？

對於獨立的 $X \sim \mathrm{Po}(\alpha)$ 和 $Y \sim \mathrm{Po}(\beta)$，就 n 的各種拆分方式求和：

$$\mathrm{P}(X+Y=n) = \sum_{k=0}^{n} \frac{\mathrm{e}^{-\alpha}\alpha^k}{k!} \cdot \frac{\mathrm{e}^{-\beta}\beta^{n-k}}{(n-k)!} = \frac{\mathrm{e}^{-(\alpha+\beta)}}{n!} \sum_{k=0}^{n} \binom{n}{k} \alpha^k \beta^{n-k} = \frac{\mathrm{e}^{-(\alpha+\beta)}(\alpha+\beta)^n}{n!},$$

最後一步用了二項式定理（binomial theorem）。所以 $X + Y \sim \mathrm{Po}(\alpha+\beta)$。應用這個結果兩次，得 $R \sim \mathrm{Po}(3\lambda)$；再應用一次，得 $T \sim \mathrm{Po}(4\lambda)$。

M 的條件分佈

對於 $m = 0, 1, \ldots, 5$，事件 $\{M = m,\ T = 5\}$ 與 $\{M = m,\ R = 5 - m\}$ 相同。由於 M 與 R 獨立，

$$\begin{aligned}\mathrm{P}(M = m \mid T = 5) &= \frac{\mathrm{P}(M = m)\,\mathrm{P}(R = 5 - m)}{\mathrm{P}(T = 5)} \\ &= \frac{\dfrac{\mathrm{e}^{-\lambda}\lambda^m}{m!} \cdot \dfrac{\mathrm{e}^{-3\lambda}(3\lambda)^{5-m}}{(5-m)!}}{\dfrac{\mathrm{e}^{-4\lambda}(4\lambda)^5}{5!}} \\ &= \frac{5!}{m!\,(5-m)!} \cdot \frac{3^{5-m}}{4^5} = \binom{5}{m}\left(\tfrac{1}{4}\right)^m \left(\tfrac{3}{4}\right)^{5-m}.\end{aligned}$$

指數項互相約去，λ 的所有冪次也一樣（分子和分母各有 λ^5）。所以，在 $T = 5$ 的條件下，不論 λ 取甚麼值，都有 $M \sim \mathrm{B}(5, \frac{1}{4})$。

答案

把 $m = 3, 4, 5$ 的二項概率相加：

$$\begin{aligned}\mathrm{P}(M \geq 3 \mid T = 5) &= \binom{5}{3}\frac{3^2}{4^5} + \binom{5}{4}\frac{3}{4^5} + \binom{5}{5}\frac{1}{4^5} \\ &= \frac{10 \times 9 + 5 \times 3 + 1}{1024} = \frac{106}{1024} = \frac{53}{512}.\end{aligned}$$

所以，五隻破碟之中有三隻或以上由數學家打破的概率是 $\dfrac{53}{512} \approx 0.104$。

驗算

- 六個條件概率（conditional probability）$\binom{5}{m}3^{5-m}/4^5$ 之和為 $(1+3)^5/4^5 = 1$，正如所料。
- 對稱性：按同樣的論證，每名學生在五隻破碟中所佔數目的期望值（expected value）為 $5 \times \frac{1}{4} = \frac{5}{4}$，而四人所佔數目之和為 5。
- 我們直接由四個泊松分佈計算 $\mathrm{P}(M \geq 3 \mid T = 5)$（就五隻破碟在四名學生之間的每一種分配方式求和），取 $\lambda = 0.3, 1, 2.7$ 和 5，每次都得到 $0.103515625 = 53/512$。另外模擬了 4×10^5 輪洗碗，在恰好打破五隻碟子的各輪之中，所得比例約為 0.105。

常見錯誤

- 答案中仍留有 λ。它必定會約去；如果約不去，就表示取條件的步驟出了錯。

- 除以 $\mathrm{P}(M+R=5)$ 時，把 R 當作 $\mathrm{Po}(\lambda)$ 而不是 $\mathrm{Po}(3\lambda)$，這等於只有兩名學生。
- 漏掉二項式係數，或把事件當作「恰好三隻」而不是「三隻或以上」。

STEP 之外

這是一個一般事實的小例子：若 $X_1,\ldots,X_k$ 是獨立的泊松變量，平均值分別為 $\lambda_1,\ldots,\lambda_k$，則在它們的和為 n 的條件下，向量 $(X_1,\ldots,X_k)$ 服從多項分佈（multinomial distribution），其概率為 $\lambda_i/\sum\lambda_j$。反過來看，這個事實說明：若事件的數目服從 $\mathrm{Po}(\mu)$，而每個事件各自獨立地標記為 k 種類型之一，則各類型的數目是互相獨立的泊松變量。這種「泊松分拆」（Poisson splitting）是研究泊松過程（Poisson process）的基本工具。Blitzstein 與 Hwang 合著的 *Introduction to Probability*（概率論入門教材）以大量例子討論了這兩個方向。

詳細解答

STEP I 2001，第 14 題

概率與統計 • 隨機排序

STEP I 2001，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/01-s1-q14

本題在 *Beyond STEP* 第 12.12 節何時停止尋找中有完整的引導式講解：https://beyondstep.keithmcnulty.org/zh-hant/ch12/secretary.html

關鍵思路

不要考慮整個面試次序，只需考慮個別候選人排在哪個位置。首 n 人中的最佳者就是全體最佳者，當且僅當最優秀的候選人在首 n 名接受面試的人之中；而根據對稱性（symmetry），最優秀的候選人落在 N 個位置中任何一個的可能性均等。第 (ii) 部分只需多考慮一種情況，或者巧妙地利用補事件（complement）。

解答

按名次為候選人編號：候選人 1 是全體最佳，候選人 2 是第二佳，如此類推。由於全部 $N!$ 種次序的可能性均等，每位候選人佔據任何指定位置的可能性均等，而首 n 場面試所見的候選人組成的集合，是 $\binom{N}{n}$ 個大小為 n 的子集（subset）中任何一個的可能性亦均等。以下一律假設 $1 \leq n \leq N$。

第 (i) 部分

設 A 為候選人 1 在首 n 名接受面試的人之中這一事件（event）。若 A 發生，候選人 1 必定是首 n 人中的最佳者；若不發生，首 n 人中的最佳者就是比候選人 1 遜色的人。因此我們所求的事件正正就是 A。

候選人 1 落在 N 個位置中每一個的可能性均等，而其中 n 個位置屬於首 n 個，所以

$$\mathrm{P}(\text{首}n\text{ 人中的最佳者是全體最佳}) = \frac{n}{N}.$$

第 (ii) 部分

現在取 $N \geq 2$，使第二佳的候選人存在。首 n 人中的最佳者是全體最佳或第二佳，當且僅當

- 候選人 1 在首 n 人之中（事件 A），或
- 候選人 1 不在首 n 人之中，但候選人 2 在。

> **想一想** 在第二種情況中，我們怎樣知道首 n 人中的最佳者是候選人 2，而不是更優秀的人？

比候選人 2 優秀的只有候選人 1，而在這種情況下他並不在較早接受面試的人之中。反過來說，若 1 和 2 都不在首 n 人之中，則首 n 人中的最佳者名次為 3 或以下。所以上述兩種情況合起來正正就是我們所求的事件，而且兩者互斥（mutually exclusive）。

由第 (i) 部分，第一種情況的概率為 $\frac{n}{N}$。至於第二種情況，$\mathrm{P}(A') = \frac{N-n}{N}$，而已知候選人 1 在最後 $N-n$ 個位置的其中之一時，候選人 2 落在餘下 $N-1$ 個位置中任何一個的可能性均等，其中 n 個屬於首 n 個位置。因此

$$\begin{aligned}\mathrm{P}(\text{首}n\text{ 人中的最佳者是全體第一或第二}) &= \frac{n}{N} + \frac{N-n}{N}\cdot\frac{n}{N-1}\\ &= \frac{n\big[(N-1)+(N-n)\big]}{N(N-1)}\\ &= \frac{n(2N-n-1)}{N(N-1)}.\end{aligned}$$

*另一方法*是利用補事件：所求事件不發生，當且僅當較早接受面試的 n 名候選人全部來自名次為 3 至 N 的 $N-2$ 名候選人，其概率為

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

而 $N(N-1)-(N-n)(N-n-1) = 2Nn-n^2-n = n(2N-n-1)$，得出相同的答案。

$N=4$、$n=2$ 的情況

按公式，(i) 的答案是 $\frac{2}{4} = \frac{1}{2}$，(ii) 的答案是 $\frac{2\cdot 5}{4\cdot 3} = \frac{5}{6}$。

要列出各種可能性，留意到重要的只是哪兩名候選人最先接受面試。$\binom{4}{2} = 6$ 個組合中

的每一個，都恰好由 $2! \times 2! = 4$ 種次序產生（全部共 24 種次序；先排好首兩人，再排好後兩人），所以這六個組合的可能性均等：

最先接受面試的兩人	$\{1,2\}$	$\{1,3\}$	$\{1,4\}$	$\{2,3\}$	$\{2,4\}$	$\{3,4\}$
其中的最佳者	1	1	1	2	2	3

首兩人中的最佳者是全體最佳的情況佔 3 種（共 6 種），概率為 $\frac{1}{2}$；是全體最佳或第二佳的情況佔 5 種（共 6 種），概率為 $\frac{5}{6}$。兩者都與公式吻合。

驗算

- $n = N$：我們見過所有人，所以兩個概率都應該是 1。果然 $\frac{N}{N} = 1$ 而 $\frac{N(N-1)}{N(N-1)} = 1$。
- $n = 1$：第一名候選人是均勻隨機選出的一人，所以 (ii) 的答案應該是 $\frac{2}{N}$；公式給出 $\frac{2N-2}{N(N-1)} = \frac{2}{N}$。
- (ii) 的答案不小於 (i) 的答案，因為 $2N - n - 1 \geq N - 1$。
- 把全部 $N!$ 種次序逐一列舉，對 $N \leq 7$ 及每個 n 的結果都與兩條公式完全吻合（例如 $N = 5$、$n = 2$ 給出 $\frac{2}{5}$ 和 $\frac{7}{10}$）。

常見錯誤

- 在 (ii) 中忘記了「全體第二佳」只在全體最佳者*不*在首 n 人之中時才算數；直接把 $\frac{n}{N} + \frac{n}{N}$ 相加會重複計算。
- 以候選人 1 的位置為條件（conditioning）之後，仍對候選人 2 使用 $\frac{n}{N}$：候選人 1 的位置一旦確定，就只剩下 $N - 1$ 個位置。

STEP 之外

這正是著名的*秘書問題*（secretary problem）的設定：假如你必須即場決定取錄或拒絕每名候選人，最佳策略是先觀察首 n 人，然後取錄其後第一名比他們全部都優秀的候選人；把第 (i) 部分的論證加以推廣，可以證明最優的 n 大約是 N/e，成功的概率約為 $1/e$。這裏所用的對稱性論證（「每名候選人落在任何位置的可能性均等」）是研究隨機排列（random permutation）的起點。Blitzstein and Hwang 的 *Introduction to Probability*（概率論入門教科書）對這類對稱性論證有十分精彩的闡述。

詳細解答

STEP II 2001，第 1 題

純數學 • 二項式近似

> **STEP II 2001，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/01-s2-q1

關鍵思路

當 x 很小時，$(1-x)^{1/2}$ 的二項級數（binomial series）收斂（converge）得很快，所以只取首幾項，便可得到極佳的有理近似（rational approximation）。訣竅在於如何選取 x：我們希望 $1-x$ 等於某個平方數乘以所求平方根的那個數，再除以十的某次冪。這樣，$1-x$ 的平方根就是所求平方根的簡單倍數。由於 $99 = 9 \times 11$ 和 $9999 = 9 \times 1111$，該怎樣選幾乎是不言而喻的。

解答

二次近似

由二項展開（binomial expansion），當 $|x| < 1$ 時，

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

所以當 x 很小時，

$$\sqrt{1-x} \approx 1 - \frac{x}{2} - \frac{x^2}{8}.$$

我們同時求得第一個被略去的項 $-\frac{1}{16}x^3$，稍後估計誤差（error）時會用到。事實上，首項之後的每一項都是負數（當 $n \geq 1$ 時，乘積 $\frac{1}{2}\cdot(-\frac{1}{2})\cdots(\frac{1}{2}-n+1)$ 有 $n-1$ 個負因子，而 $(-x)^n$ 則帶來 $(-1)^n$，合起來剛好剩下一個負號），所以當 $0 < x < 1$ 時，這個二次式是一個高估值（overestimate）。

第 (i) 部分

取 $x = \frac{1}{100}$，則

$$\sqrt{1-x} = \sqrt{\frac{99}{100}} = \frac{\sqrt{9\times 11}}{10} = \frac{3\sqrt{11}}{10}, \qquad \text{所以} \qquad \sqrt{11} = \frac{10}{3}\sqrt{1-\tfrac{1}{100}}.$$

利用二次近似，

$$\begin{aligned}\sqrt{11} &\approx \frac{10}{3}\left(1 - \frac{1}{200} - \frac{1}{80\,000}\right) = \frac{10}{3}\cdot\frac{80\,000 - 400 - 1}{80\,000}\\ &= \frac{10}{3}\cdot\frac{79\,599}{80\,000} = \frac{79\,599}{24\,000},\end{aligned}$$

這正是題目所要求的結果。

? **想一想** 第一個被略去的項是 $-\frac{1}{16}x^3$。這就是 $\sqrt{11}$ 的誤差嗎？

不完全是：那是 $\sqrt{1-x}$ 的近似值的誤差，而我們之後還乘上了 $\frac{10}{3}$。因此 $\sqrt{11}$ 的誤差約為

$$\frac{10}{3}\times\frac{1}{16}\left(\frac{1}{100}\right)^3 = \frac{10}{48\times 10^6} = \frac{1}{4\,800\,000} \approx 2.08\times 10^{-7}.$$

所以，準確至 1 位有效數字（significant figure），誤差為 2×10^{-7}，而 $\frac{79\,599}{24\,000}$ 略大於 $\sqrt{11}$。

第 (ii) 部分

我們照搬第 (i) 部分的技巧。由於 $1111\times 9 = 9999 = 10^4 - 1$，取 $x = \frac{1}{10\,000}$：

$$\sqrt{1-x} = \sqrt{\frac{9999}{10\,000}} = \frac{3\sqrt{1111}}{100}, \qquad \text{所以} \qquad \sqrt{1111} = \frac{100}{3}\sqrt{1-\tfrac{1}{10\,000}}.$$

於是

$$\begin{aligned}\sqrt{1111} &\approx \frac{100}{3}\left(1 - \frac{1}{20\,000} - \frac{1}{8\times 10^8}\right) = \frac{100}{3}\cdot\frac{800\,000\,000 - 40\,000 - 1}{8\times 10^8}\\ &= \frac{100}{3}\cdot\frac{799\,959\,999}{8\times 10^8} = \frac{799\,959\,999}{24\,000\,000} = \frac{266\,653\,333}{8\,000\,000},\end{aligned}$$

其中最後一步約去了因子 3（799 959 999 的數字之和是 75，可被 3 整除）。

誤差約為第一個被略去的項的 $\frac{100}{3}$ 倍：

$$\frac{100}{3}\times\frac{1}{16}\left(\frac{1}{10^4}\right)^3 = \frac{100}{48\times 10^{12}} \approx 2.08\times 10^{-12},$$

即大約 2×10^{-12}，符合題目的要求。所以一個合適的有理數是

$$\sqrt{1111} \approx \frac{266\,653\,333}{8\,000\,000}\left(= \frac{799\,959\,999}{24\,000\,000}\right).$$

驗算

- $79\,599/24\,000 = 3.316625$，而 $\sqrt{11} = 3.31662479\ldots$；兩者相差 2.096×10^{-7}，與估計值 2.08×10^{-7} 吻合，亦印證了這個近似值偏大。
- $266\,653\,333/8\,000\,000 = 33.331666625$，而 $\sqrt{1111} = 33.33166662291\ldots$；我們以高精度準確計算兩者之差，得出 2.0835×10^{-12}。
- 第 (ii) 部分的估計更為準確，因為當 x 較小時，下一個被略去的項 $-\frac{5}{128}x^4$ 相對於 x^3 項也較小。

常見錯誤

- 忘記把被略去的項乘以 $\frac{10}{3}$（或 $\frac{100}{3}$），以致得出的誤差小了十倍或一百倍。
- 弄錯 x^2 係數的正負號：由於是 $(-x)^2$，該項為 $\frac{\frac{1}{2}\cdot(-\frac{1}{2})}{2}x^2 = -\frac{1}{8}x^2$，而不是 $+\frac{1}{8}x^2$。
- 嘗試選取 x 的值，使 $1 - x$ 等於 1111 除以十的某次冪。這樣的值並不小，級數收斂得太慢，根本幫不上忙。

STEP 之外

「誤差大約等於第一個被略去的項」這個說法，可以用泰勒定理（Taylor's theorem）和拉格朗日餘項（Lagrange form of the remainder）嚴格地表述：它給出誤差的準確公式，當中含有一個位於 $(0, x)$ 內的未知點。由此我們可以得到確切的上界和下界（upper and lower bounds），而不僅僅是估計。把這些誤差界嚴格地建立起來，是大學第一門分析（analysis）課程的標準內容；參閱 Abbott 的 *Understanding Analysis*。

詳細解答

STEP II 2001，第 2 題

純數學 • 整數部分的求和

STEP II 2001，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/01-s2-q2

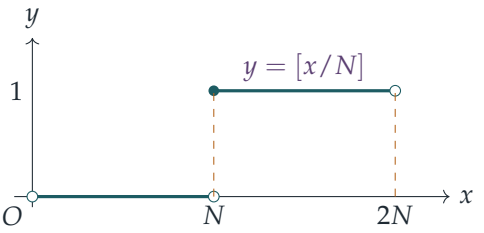
關鍵思路

指數 $[k/N]$ 看似棘手，但當 k 由 1 取到 $2N$ 時，它其實只有三個可能的值，而草圖正好顯示它在哪裏跳升。弄清這一點之後，正負號 $(-1)^{[k/N]}$ 在 k 的三個區段上各自保持不變，於是每個和式都可以按區段拆開：第 (i) 部分拆成等差級數（arithmetic series），第 (ii) 部分拆成等比級數（geometric series）。唯一要小心的，是各區段的端點。

解答

草圖

當 $0 < x < N$ 時，$0 < x/N < 1$，所以 $[x/N] = 0$。當 $N \le x < 2N$ 時，$1 \le x/N < 2$，所以 $[x/N] = 1$。圖像是一級階梯：跳升發生在 $x = N$，函數在該點已經取值 1；右端點 $x = 2N$ 並不包括在內（函數在該處的值會是 2）。



實心點表示包括該點，空心點表示不包括該點。

計算和式時，我們需要的是整數值 $k = 1, 2, \ldots, 2N$。從圖像讀出（並補上端點 $k = 2N$）：

$$[k/N] = \begin{cases} 0, & 1 \le k \le N-1, \\ 1, & N \le k \le 2N-1, \\ 2, & k = 2N. \end{cases}$$

因此，$(-1)^{[k/N]}$ 在 $k \le N-1$ 時是 $+1$，在 $N \le k \le 2N-1$ 時是 -1，而在 $k = 2N$ 這單獨一項又變回 $+1$。（當 $N = 1$ 時，第一個區段是空的；由於空和等於 0，以下所有公式仍然成立。）

> **想一想** 為甚麼不能直接說首 N 項取 $+$ 號、其後 N 項取 $-$ 號？

因為跳升恰好發生在 N 的倍數上：$k = N$ 這一項已經有 $[k/N] = 1$，而最後一項 $k = 2N$ 有 $[k/N] = 2$，所以它是正的。把這兩個端點處理正確，就是整道題的關鍵。

第 (i) 部分

按三個區段拆開，

$$\sum_{k=1}^{2N} (-1)^{[k/N]} k = \sum_{k=1}^{N-1} k \; - \sum_{k=N}^{2N-1} k \; + \; 2N.$$

第一個和是 $\frac{1}{2}(N-1)N$。第二個和共有 N 項，由 N 到 $2N-1$，所以它等於 N 乘以首項與末項的平均值：$\frac{1}{2}N(3N-1)$。因此

$$\begin{aligned} \sum_{k=1}^{2N} (-1)^{[k/N]} k &= \tfrac{1}{2}\left(N^2 - N\right) - \tfrac{1}{2}\left(3N^2 - N\right) + 2N \\ &= -N^2 + 2N = 2N - N^2, \end{aligned}$$

這正是所要證明的結果。

第 (ii) 部分

用同樣的拆法，得

$$S_N = \sum_{k=1}^{N-1} 2^{-k} \; - \sum_{k=N}^{2N-1} 2^{-k} \; + \; 2^{-2N}.$$

兩個和都是公比為 $\frac{1}{2}$ 的等比級數。利用 $\sum_{k=a}^{b} 2^{-k} = 2^{-a}\left(1 - 2^{-(b-a+1)}\right) \big/ \frac{1}{2} = 2^{1-a} - 2^{-b}$，

$$\sum_{k=1}^{N-1} 2^{-k} = 1 - 2^{1-N},$$

$$\sum_{k=N}^{2N-1} 2^{-k} = 2^{1-N} - 2^{1-2N}.$$

所以

$$\begin{aligned} S_N &= 1 - 2^{1-N} - 2^{1-N} + 2^{1-2N} + 2^{-2N} \\ &= 1 - 4\cdot 2^{-N} + 3\cdot 2^{-2N}. \end{aligned}$$

設 $t = 2^{-N}$，上式就是 $1 - 4t + 3t^2$，可以漂亮地因式分解：

$$\boxed{S_N = 1 - \frac{4}{2^N} + \frac{3}{4^N} = \left(1 - \frac{1}{2^N}\right)\left(1 - \frac{3}{2^N}\right).}$$

當 $N \to \infty$ 時，$2^{-N} \to 0$ 且 $4^{-N} \to 0$，所以

$$\lim_{N\to\infty} S_N = 1.$$

這個結果合情合理：隨着 N 增大，正號區段 $\frac{1}{2} + \frac{1}{4} + \cdots$ 愈來愈長，其和趨近 1；而負號區段的起點已經排得很後，它的總和大約是 2^{1-N}，可以忽略不計。

驗算

- $N = 1$：各項為 $k = 1$（此時 $[1] = 1$）和 $k = 2$（此時 $[2] = 2$）。第 (i) 部分給出 $-1 + 2 = 1 = 2 - 1$，第 (ii) 部分給出 $-\frac{1}{2} + \frac{1}{4} = -\frac{1}{4}$，與 $(1 - \frac{1}{2})(1 - \frac{3}{2}) = -\frac{1}{4}$ 相符。
- $N = 2$：對應 $k = 1, 2, 3, 4$ 的正負號依次為 $+, -, -, +$。於是 $1 - 2 - 3 + 4 = 0 = 4 - 4$，而 $\frac{1}{2} - \frac{1}{4} - \frac{1}{8} + \frac{1}{16} = \frac{3}{16} = \frac{3}{4}\cdot\frac{1}{4}$。
- 我們以分數作精確計算，對 N 由 1 到 59 的每一個值驗證了這兩條公式。例如 $S_{50} \approx 0.999\,999\,999\,999\,996$。

常見錯誤

- 把 $k = N$ 歸入正號區段，或者忘記最後一項 $k = 2N$ 的指數是 2，因而是正的。
- 數錯 $\sum_{k=N}^{2N-1}$ 的項數：它共有 N 項，而不是 $N - 1$ 項。

STEP 之外

整數部分函數（integer-part function，又稱下取整函數 floor function）在數論（number theory）中隨處可見：經典的例子是勒讓德公式（Legendre's formula），它指出整除 $n!$ 的質數 p 的最高次冪是 $\sum_{j\ge 1}[n/p^j]$。像第 (ii) 部分那樣正負號按區段變化的和式，也是一個淺易的入門例子，讓我們初步接觸級數各項的重排（rearranging）與分組（grouping），而分析（analysis）課程會把這些做法嚴格化。想看看下取整函數如何大派用場，Davenport 的 *The Higher Arithmetic*（一本數論入門名著）是個絕佳的起點。

詳細解答

STEP II 2001，第 3 題

純數學 • 歐幾里得幾何

STEP II 2001，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/01-s2-q3

關鍵思路

把長方體（cuboid）放進坐標系，令其中一個頂點位於原點，各棱沿坐標軸方向。這樣每個頂點的位置向量（position vector）都一目了然，對角線只是這些向量之差，而兩條對角線之間的夾角可直接由純量積（scalar product）求得。體積與表面積之比是 λ 的一個簡單有理函數（rational function），最後一部分則把關於 R 的條件轉化為關於 λ 的條件，從而把前兩個結果連繫起來。

解答

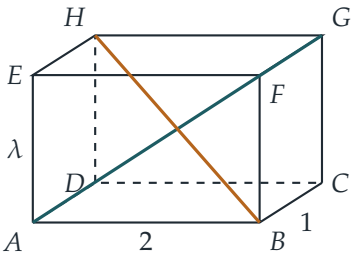
對角線之間的夾角

由於棱 AE、BF、CG、DH 垂直於各個面，各頂點處的角全是直角，因此可取坐標軸沿 AB、AD 和 AE 方向，並以 A 為原點：

$$A = (0,0,0), \quad B = (2,0,0), \quad C = (2,1,0), \quad D = (0,1,0),$$

而頂面就是把底面升高 λ：

$$E = (0,0,\lambda), \quad F = (2,0,\lambda), \quad G = (2,1,\lambda), \quad H = (0,1,\lambda).$$



兩條對角線的方向向量（direction vector）為

$$\overrightarrow{AG} = \begin{pmatrix} 2 \\ 1 \\ \lambda \end{pmatrix}, \qquad \overrightarrow{BH} = \begin{pmatrix} -2 \\ 1 \\ \lambda \end{pmatrix},$$

長度均為 $\sqrt{5+\lambda^2}$，而且

$$\overrightarrow{AG} \cdot \overrightarrow{BH} = -4 + 1 + \lambda^2 = \lambda^2 - 3.$$

設 θ 為這兩個向量之間的夾角，則

$$\cos\theta = \frac{\lambda^2 - 3}{5 + \lambda^2}.$$

? **想一想** 絕對值符號（modulus sign）是從哪裏來的？

兩條相交的直線形成兩個角 θ 和 $\pi - \theta$，它們的餘弦互為相反數。其中的銳角（當 $\lambda^2 = 3$ 時則為直角）就是餘弦非負的那一個，所以

$$\cos\alpha = |\cos\theta| = \left|\frac{3 - \lambda^2}{5 + \lambda^2}\right|,$$

正是所求。

比值 R

體積為 $2 \times 1 \times \lambda = 2\lambda$。表面由面積分別為 2×1、$2 \times \lambda$ 和 $1 \times \lambda$ 的面各兩個組成，所以表面積為 $2(2 + 3\lambda)$。因此

$$R = \frac{2\lambda}{2(2+3\lambda)} = \frac{\lambda}{2+3\lambda}.$$

由於 $\lambda > 0$，分母為正數，於是

$$R < \tfrac{1}{3} \iff 3\lambda < 2 + 3\lambda \iff 0 < 2,$$

而這顯然成立。所以對任何可能的 λ，均有 $R < \frac{1}{3}$。（當 $\lambda \to \infty$ 時，$R \to \frac{1}{3}$，所以這個界不能再改進。）

最後一部分

同樣因為 $2 + 3\lambda > 0$，可以把不等式兩邊同乘分母：

$$R \geq \tfrac{1}{4} \iff 4\lambda \geq 2 + 3\lambda \iff \lambda \geq 2.$$

現設 $\lambda \geq 2$。則 $\lambda^2 \geq 4 > 3$，所以 $\lambda^2 - 3 > 0$，絕對值符號可以去掉：

$$\cos\alpha = \frac{\lambda^2 - 3}{\lambda^2 + 5} = 1 - \frac{8}{\lambda^2 + 5}.$$

寫成這個形式，便可看出 $\cos\alpha$ 隨 λ^2 增大而增大。由 $\lambda^2 \geq 4$ 得 $\lambda^2 + 5 \geq 9$，所以

$$\cos\alpha \geq 1 - \frac{8}{9} = \frac{1}{9}.$$

α 和 $\arccos\frac{1}{9}$ 都在 $[0, \frac{\pi}{2}]$ 內，而餘弦在這個區間上嚴格遞減（strictly decreasing），所以由 $\cos\alpha \geq \frac{1}{9}$ 可得

$$\alpha \leq \arccos\tfrac{1}{9},$$

這正是要證明的結果。等號成立當且僅當 $\lambda = 2$，即 $R = \frac{1}{4}$。

驗算

- 當 $\lambda^2 = 3$ 時，公式給出 $\cos\alpha = 0$。驗算：$\overrightarrow{AG} \cdot \overrightarrow{BH} = -4 + 1 + 3 = 0$，所以兩條對角線確實互相垂直。
- 當 $\lambda \to \infty$ 時，長方體變成又高又窄的棱柱（prism），兩條對角線都變得接近垂直，所以 $\alpha \to 0$；公式給出 $\cos\alpha \to 1$，與此吻合。
- 當 $\lambda = 2$ 時：體積為 4，表面積為 $2(2 + 4 + 2) = 16$，所以 $R = \frac{1}{4}$，而 $\cos\alpha = \frac{1}{9}$，所以 $\alpha \approx 83.6°$。以 10^5 個隨機的 λ 值作數值檢驗，結果證實了夾角公式、$R < \frac{1}{3}$，以及每當 $R \geq \frac{1}{4}$ 時 $\alpha \leq \arccos\frac{1}{9}$。

常見錯誤

- 忘記加絕對值符號：當 $\lambda^2 < 3$ 時，向量 $\overrightarrow{AG}$ 與 $\overrightarrow{BH}$ 所成的是鈍角，而所求的銳角是它的補角。
- 把最後一個不等式的方向弄反：餘弦較大，角就較小，因為餘弦在 $[0, \frac{\pi}{2}]$ 上遞減。

STEP 之外

把物體放進坐標系，再用純量積求角度，是向量幾何的基本工具；在大學一年級的線性代數（linear algebra）課程中，它會推廣為任意維數的內積（inner product）：$\mathbb{R}^n$ 中兩個向量之間的夾角是用 $\cos\theta = \mathbf{u}\cdot\mathbf{v}/(|\mathbf{u}||\mathbf{v}|)$ 來定義的，而柯西－施瓦茨不等式（Cauchy–Schwarz inequality）保證了這個值落在 $[-1, 1]$ 之內。有一道不錯的練習：證明 n 維單位立方體的長對角線與每條棱所成的角的餘弦為 $1/\sqrt{n}$，所以當 n 增大時，這個角趨向

直角。Strang 的 *Introduction to Linear Algebra*（線性代數入門）把這些概念講解得十分清楚。

詳細解答

STEP II 2001，第 4 題

純數學 • 倍角方程

STEP II 2001，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/01-s2-q4

關鍵思路

f 的每一項都含有因子 $\sin x$：$\sin 2x = 2\sin x\cos x$，而題目的提示對 $\sin 3x$ 也有同樣的作用。抽出 $\sin x$ 之後，剩下的是一個關於 $\cos x$ 的二次式（quadratic），而 P、Q、R 所滿足的條件，正好就是「這個二次式的判別式（discriminant）為負」。第二部分是同一個函數，只是取 $P = 1$、$Q = 1$、$R = -1$，並把 x 換成 $2nx$。這時二次式確實有根，功夫就在於找出哪一個根最接近容許範圍的上端。

解答

第一個結果

設 $c = \cos x$。利用 $\sin 2x = 2\sin x\cos x$ 及題目給出的 $\sin 3x$ 恆等式（identity），

$$\begin{aligned} \mathrm{f}(x) &= P\sin x + 2Q\sin x\cos x + R\sin x\,(4\cos^2 x - 1) \\ &= \sin x\left(4Rc^2 + 2Qc + (P - R)\right). \end{aligned}$$

所以 $\mathrm{f}(x) = 0$ 當且僅當 $\sin x = 0$ 或 $4Rc^2 + 2Qc + (P - R) = 0$。

首先，$R \neq 0$：若 $R = 0$，條件便會變成 $Q^2 < 0$，這是不可能的。因此括號內確實是一個關於 c 的二次式，其判別式為

$$(2Q)^2 - 4\cdot 4R\,(P - R) = 4\big(Q^2 - 4R(P - R)\big) < 0$$

這由假設可得。判別式為負的實係數二次式沒有實根，所以對任何實數 c，括號內的式子都不會等於零，當 $c = \cos x$ 時自然也不例外。故此 $\mathrm{f}(x) = 0$ 必然導致 $\sin x = 0$，即 $x = m\pi$，其中 m 為整數。反過來，每一個 $x = m\pi$ 都是 f 的零點（zero），所以這些正好就是全部零點，結果得證。

g 的最大根

設 $y = 2nx$。那麼 $g(x) = \sin y + \sin 2y - \sin 3y$，即是取 $P = 1$、$Q = 1$、$R = -1$ 時的 $f(y)$。這裏 $Q^2 = 1$ 並不小於 $4R(P - R) = -8$，所以第一個結果並不適用，而二次式有實根。設 $c = \cos y$，由上面的因式分解（factorisation）可得

$$g(x) = \sin y\,(-4c^2 + 2c + 2) = -2\sin y\,(2c + 1)(c - 1).$$

所以 $g(x) = 0$ 當且僅當

$$\sin y = 0, \qquad \cos y = 1, \qquad 或 \qquad \cos y = -\tfrac{1}{2}.$$

$\cos y = 1$ 的情況已經包含在 $\sin y = 0$ 之內。因此以 y 表示的根為

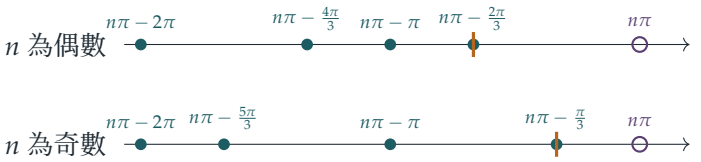
$$y = k\pi \quad 或 \quad y = 2k\pi \pm \tfrac{2\pi}{3} \qquad (k \in \mathbb{Z}).$$

在每一個周期（period）$[2k\pi, 2k\pi + 2\pi)$ 之內，根位於 $2k\pi$ 加上 0、$\frac{2\pi}{3}$、π、$\frac{4\pi}{3}$ 之處。

範圍 $0 < x < \frac{1}{2}\pi$ 變成 $0 < y < n\pi$，而右端點 $y = n\pi$ 本身是一個根，卻被排除在外。所以我們要找的是嚴格小於 $n\pi$ 的最大根，而關鍵只在於 $n\pi$ 在其所屬周期內的位置。

> **想一想** 為甚麼 n 的奇偶性（parity）有影響？

根的分佈規律每隔 2π 重複一次，而不是每隔 π。若 n 為偶數，$n\pi$ 是 2π 的倍數；若 n 為奇數，$n\pi$ 是 π 的奇數倍。在這兩種情況下，緊接在 $n\pi$ 之下的根，其間距並不相同。



（g 以 y 表示的根，範圍為 $n\pi$ 之下的最後一段 2π，按比例繪畫。空心圓圈標示被排除的端點，琥珀色的短線標示最大根。）

n 為偶數。這時 $n\pi = 2k\pi$，其中 $k = \frac{n}{2} \geq 1$，而在它之下的最後一個周期內，以 $n\pi$ 為基準量度，根為

$$n\pi - 2\pi, \quad n\pi - \tfrac{4\pi}{3}, \quad n\pi - \pi, \quad n\pi - \tfrac{2\pi}{3}.$$

其中最後一個是 $2(k-1)\pi + \frac{4\pi}{3}$，它是根，因為其餘弦值為 $-\frac{1}{2}$。在 $\left(n\pi - \frac{2\pi}{3}, n\pi\right)$ 內沒有根：這個區間內沒有 π 的倍數，而 $\cos y = -\frac{1}{2}$ 則需要 $y \equiv \frac{2\pi}{3}$ 或 $\frac{4\pi}{3} \pmod{2\pi}$，但這裏 y 在模 2π 之下嚴格介乎 $\frac{4\pi}{3}$ 與 2π 之間。此外，由於 $n \geq 2$，有 $n\pi - \frac{2\pi}{3} > 0$。所以最大根對應 $y = n\pi - \frac{2\pi}{3}$，而

$$x = \frac{y}{2n} = \frac{(3n-2)\pi}{6n} \qquad (n \text{ 為偶數}).$$

n **為奇數**。這時 $n\pi = 2k\pi + \pi$，其中 $k = \frac{n-1}{2} \geq 0$。在由 $2k\pi = n\pi - \pi$ 開始的周期內量度，$n\pi$ 之下的根位於偏移量 0 和 $\frac{2\pi}{3}$ 之處，所以候選的根是 $n\pi - \pi$ 和 $n\pi - \frac{\pi}{3}$。區間 $\left(n\pi - \frac{\pi}{3}, n\pi\right)$ 對應於 y 在模 2π 之下嚴格介乎 $\frac{2\pi}{3}$ 與 π 之間，當中沒有根。而且 $n\pi - \frac{\pi}{3} \geq \frac{2\pi}{3} > 0$。所以最大根對應 $y = n\pi - \frac{\pi}{3}$，而

$$x = \frac{(3n-1)\pi}{6n} \qquad (n \text{ 為奇數}).$$

驗算

當 $n = 1$ 時，在 $(0, \frac{1}{2}\pi)$ 上 $\mathrm{g}(x) = \sin 2x + \sin 4x - \sin 6x$：公式給出 $x = \frac{\pi}{3}$，而確實有 $\sin\frac{2\pi}{3} + \sin\frac{4\pi}{3} - \sin 2\pi = 0$。當 $n = 2$ 時，公式同樣給出 $\frac{\pi}{3}$（$y = \frac{4\pi}{3}$）。當 $n \to \infty$ 時，兩條公式都趨向 $\frac{1}{2}\pi$，這是理所當然的，因為根會愈來愈密集。我們以細密的網格對 $n = 1, \ldots, 12$ 的 g 作了數值掃描。每一個預測值（例如 $n = 3$ 時的 $x = \frac{4\pi}{9} \approx 1.3963$）都是根，而 g 在該值與 $\frac{1}{2}\pi$ 之間保持同一符號。至於第一部分，隨機抽取 2000 組滿足 $Q^2 < 4R(P-R)$ 的三元組，所得的括號內式子在 $[-1, 1]$ 上全部沒有零點。

常見錯誤

- 答 $x = \frac{1}{2}\pi$ 本身：對每一個 n，它都是 g 的根，但它在開區間（open interval）之外。
- 忘記了由 $\sin y = 0$ 得出的根，只保留 $\cos y = -\frac{1}{2}$。當 n 為偶數時，$y = n\pi - \pi$ 是一個根，雖然並非最大的一個。
- 只寫 $\cos y = -\frac{1}{2} \Rightarrow y = \frac{2\pi}{3} + 2k\pi$，漏掉了 $\frac{4\pi}{3}$ 那一族解。這樣在 n 為偶數時會得出錯誤答案。

STEP 之外

這個因式分解之所以行得通，是因為 $\sin mx / \sin x$ 永遠是 $\cos x$ 的多項式（polynomial）：這些就是第二類切比雪夫多項式（Chebyshev polynomials of the second kind）$U_{m-1}(\cos x)$。所以任何有限正弦級數（sine series）$\sum a_m \sin mx$ 都等於 $\sin x$ 乘以一個 $\cos x$ 的多項式，而數算它的零點，就變成數算多項式在 $[-1, 1]$ 內的根。這個想法是逼近理論（approximation theory）和傅立葉分析（Fourier analysis）中許多內容的基礎。Needham 的 *Visual Complex Analysis*（《複分析：可視化方法》）透過 e^{ix}，為倍角公式（multiple-angle formulae）提供了一個優美的視角。

詳細解答

STEP II 2001，第 5 題

純數學 • 駐點與積分

STEP II 2001，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/01-s2-q5

關鍵思路

駐點（stationary point）可直接由斜率（gradient）求得：它們位於 $\frac{dy}{dx} = 0$ 之處，而其性質則取決於斜率的正負號如何變化。由於指數因子恆為正，斜率的正負號只取決於多項式因子 $x(1-x^2)$，而這個因子對兩條曲線都相同。至於高度，第一個斜率可以準確地積分（留意 $u = x^2$）；第二個則不能，但我們並不需要它的值，只需比較兩個積分。

解答

曲線 C_1

由於對所有 x 都有 $e^{-x^2} > 0$，斜率為零當且僅當 $x(1-x^2) = x(1-x)(1+x) = 0$，即 $x = -1$、0 及 1。此外，斜率的正負號與 $x(1-x)(1+x)$ 相同：

	$x < -1$	$-1 < x < 0$	$0 < x < 1$	$x > 1$
$\frac{dy}{dx}$ 的正負號	$+$	$-$	$+$	$-$

在 $x = 0$ 處，斜率由負變正，所以原點（它在曲線上）是極小點（minimum）。在 $x = 1$ 處，斜率由正變負，所以這是極大點（maximum）；同理，$x = -1$ 也是極大點。

要求出高度，我們進行積分。由於曲線通過原點，

$$y = \int_0^x t(1-t^2)\, e^{-t^2}\, dt.$$

? **想一想** 怎樣對 $t(1-t^2)e^{-t^2}$ 積分？

因子 t 旁邊是一個 t^2 的函數，這提示我們用代換（substitution）$u = t^2$，$\mathrm{d}u = 2t\,\mathrm{d}t$：

$$y = \tfrac{1}{2}\int_0^{x^2} (1-u)\,\mathrm{e}^{-u}\,\mathrm{d}u.$$

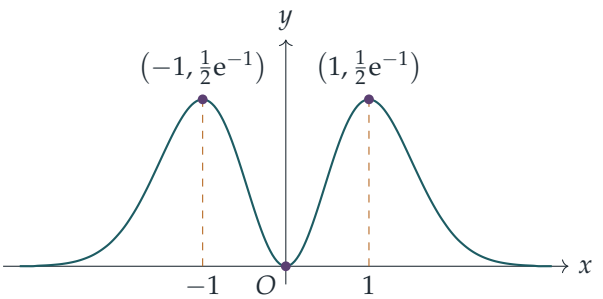
現在留意 $\frac{\mathrm{d}}{\mathrm{d}u}\left(u\mathrm{e}^{-u}\right) = \mathrm{e}^{-u} - u\mathrm{e}^{-u} = (1-u)\mathrm{e}^{-u}$（用分部積分法（integration by parts）也會得到同樣的結果）。因此

$$y = \tfrac{1}{2}\left[u\,\mathrm{e}^{-u}\right]_0^{x^2} = \tfrac{1}{2}\,x^2\mathrm{e}^{-x^2}.$$

在 $x = 1$ 處，這給出 $y = \frac{1}{2}\mathrm{e}^{-1}$，所以極大點位於 $\left(1, \frac{1}{2}\mathrm{e}^{-1}\right)$，正如題目所求。由於 y 是 x 的偶函數（even function），另一個駐點是極大點

$$\left(-1, \tfrac{1}{2}\mathrm{e}^{-1}\right).$$

草圖。 由 $y = \frac{1}{2}x^2\mathrm{e}^{-x^2}$ 可知：曲線關於 y 軸對稱；$y \geq 0$，且只在原點取等號；又因為指數函數衰減的速度比 x^2 增長的速度快，當 $x \to \pm\infty$ 時 $y \to 0$。所以這是一條「雙峯」曲線：它在原點與 x 軸相切，上升至 $x = \pm 1$ 處兩個等高的極大點（高度為 $\frac{1}{2}\mathrm{e}^{-1} \approx 0.18$），然後在兩邊逐漸下降，趨向 x 軸。（下圖的縱向比例經過放大。）



曲線 C_2

同樣，對所有 x 都有 $\mathrm{e}^{-x^3} > 0$，所以斜率恰好在 $x = -1, 0, 1$ 處為零，而且正負號與 $x(1-x)(1+x)$ 相同，得出的正負號表與之前完全一樣。所以原點（按題設在曲線上）是極小點，而 $x = 1$ 處的駐點是極大點。（$x = -1$ 處的駐點也是極大點，不過題目沒有問及。）

它的高度是

$$k = \int_0^1 x(1-x^2)\,\mathrm{e}^{-x^3}\,\mathrm{d}x.$$

這個積分沒有初等原函數（elementary antiderivative），所以我們改為把它與 C_1 的積分比較，而我們已知後者等於 $\frac{1}{2}\mathrm{e}^{-1}$。

想一想 怎樣可以不計算 k 的值，而把 k 與 $\frac{1}{2}\mathrm{e}^{-1}$ 比較？

比較被積函數（integrand）。當 $0 < x < 1$ 時，我們有 $x^3 < x^2$（把 $x < 1$ 兩邊乘以 $x^2 > 0$），所以 $-x^3 > -x^2$；又因為 e^t 是遞增函數，故 $\mathrm{e}^{-x^3} > \mathrm{e}^{-x^2}$。此外，在這個範圍內 $x(1-x^2) > 0$，所以乘上它之後不等式的方向不變：

$$x(1-x^2)\,\mathrm{e}^{-x^3} > x(1-x^2)\,\mathrm{e}^{-x^2} \qquad \text{當} 0 < x < 1.$$

兩個被積函數在 $[0,1]$ 上都連續，而不等式在開區間上嚴格成立，所以在 $[0,1]$ 上積分會得到嚴格不等式（strict inequality）：

$$k = \int_0^1 x(1-x^2)\,\mathrm{e}^{-x^3}\,\mathrm{d}x > \int_0^1 x(1-x^2)\,\mathrm{e}^{-x^2}\,\mathrm{d}x = \tfrac{1}{2}\mathrm{e}^{-1},$$

最後一個等式就是上面求得的 C_1 在 $x = 1$ 處的高度。所以極大點位於 $(1,k)$，其中 $k > \frac{1}{2}\mathrm{e}^{-1}$，正如題目所求。

驗算

用積法則（product rule）對 $\frac{1}{2}x^2\mathrm{e}^{-x^2}$ 求導，得 $x\mathrm{e}^{-x^2} - x^3\mathrm{e}^{-x^2} = x(1-x^2)\mathrm{e}^{-x^2}$，而它在 $x = 0$ 處的值為 0，所以 C_1 的結果正確。二階導數（second derivative）印證了駐點的分類：對兩條曲線都有 $y''(0) = 1 > 0$ 及 $y''(1) = -2\mathrm{e}^{-1} < 0$。以數值方法（森遜法則（Simpson's rule））計算，$k \approx 0.2034$，而 $\frac{1}{2}\mathrm{e}^{-1} \approx 0.1839$。

常見錯誤

- 忘記積分常數（constant of integration）：正是「通過原點」這個條件確定了它的值，而它對高度有影響。
- 只憑 $\mathrm{e}^{-x^3} > \mathrm{e}^{-x^2}$ 就聲稱 $k > \frac{1}{2}\mathrm{e}^{-1}$，卻沒有指出 $x(1-x^2)$ 在 $(0,1)$ 上為正；若乘上的是負因子，不等式的方向便會倒轉。
- 說 $x^3 < x^2$「對所有 x 成立」：它只在 $x < 1$、$x \neq 0$ 時成立，這正是為甚麼區間 $[0,1]$ 恰好符合我們所需。

STEP 之外

第二部分是一個比較論證（comparison argument）：對於一個寫不出來的解，我們把它與一個寫得出來的解比較，從而了解它的一些性質。這個想法貫穿整個微分方程（differential equations）理論，當中的比較定理（comparison theorem）以較簡單方程的解來界

定棘手方程的解的範圍；它也貫穿分析學（analysis），當中積分的單調性（monotonicity）是一件基本工具。Boyce 與 DiPrima 的 *Elementary Differential Equations*（微分方程入門教材）是個好去處，可以讓你看到即使沒有公式，也能從定性的角度對解作出多少論述。

詳細解答

STEP II 2001，第 6 題

純數學 • 代換積分法

STEP II 2001，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/01-s2-q6

關鍵思路

看似棘手的 $\tan^{-1}\left(\frac{1-x}{1+x}\right)$ 其實只是 $\frac{\pi}{4} - \tan^{-1} x$ 的另一種寫法：它就是把正切的差角公式（tangent subtraction formula）倒過來讀。一旦看穿這一點，第 (i) 部分便化為一次分部積分（integration by parts），而且剛好落在開首的積分上。第 (ii) 部分看來毫不相干，但映射 $y = \frac{1-x}{1+x}$ 把 0 和 1 互換，而且它是自身的逆映射，所以用它作代換（substitution），便可把 (ii) 變回 (i)。

解答

開首的積分

以 $1 + x^2$ 除 x^4（或寫成 $x^4 = (x^4 - 1) + 1 = (x^2 - 1)(x^2 + 1) + 1$），得

$$\frac{x^4}{1+x^2} = x^2 - 1 + \frac{1}{1+x^2}.$$

因此

$$\int_0^1 \frac{x^4}{1+x^2}\,\mathrm{d}x = \left[\tfrac{1}{3}x^3 - x + \tan^{-1} x\right]_0^1 = \frac{1}{3} - 1 + \frac{\pi}{4} = \frac{\pi}{4} - \frac{2}{3},$$

正是所求。

第 (i) 部分

設 $\theta = \tan^{-1} x$，則當 $0 \le x \le 1$ 時，有 $0 \le \theta \le \frac{\pi}{4}$。由差角公式，

$$\tan\left(\frac{\pi}{4} - \theta\right) = \frac{\tan\frac{\pi}{4} - \tan\theta}{1 + \tan\frac{\pi}{4}\tan\theta} = \frac{1-x}{1+x}.$$

想一想 有了 $\tan\left(\frac{\pi}{4}-\theta\right)=\frac{1-x}{1+x}$，是否真的就可以寫 $\tan^{-1}\left(\frac{1-x}{1+x}\right)=\frac{\pi}{4}-\theta$？

只有當 $\frac{\pi}{4}-\theta$ 落在 $\tan^{-1}$ 的值域（range）$\left(-\frac{\pi}{2},\frac{\pi}{2}\right)$ 之內才可以。這裏 $0\le\frac{\pi}{4}-\theta\le\frac{\pi}{4}$，所以條件成立，於是在整個積分區間上

$$\tan^{-1}\left(\frac{1-x}{1+x}\right)=\frac{\pi}{4}-\tan^{-1}x.$$

因此

$$\int_0^1 x^3\tan^{-1}\left(\frac{1-x}{1+x}\right)\mathrm{d}x=\frac{\pi}{4}\int_0^1 x^3\,\mathrm{d}x-\int_0^1 x^3\tan^{-1}x\,\mathrm{d}x=\frac{\pi}{16}-J,$$

其中 $J=\int_0^1 x^3\tan^{-1}x\,\mathrm{d}x$。取 $u=\tan^{-1}x$ 及 $\mathrm{d}v=x^3\,\mathrm{d}x$（於是 $v=\frac{1}{4}x^4$ 而 $\mathrm{d}u=\frac{\mathrm{d}x}{1+x^2}$）作分部積分：

$$\begin{aligned}J&=\left[\tfrac{1}{4}x^4\tan^{-1}x\right]_0^1-\frac{1}{4}\int_0^1\frac{x^4}{1+x^2}\,\mathrm{d}x\\&=\frac{\pi}{16}-\frac{1}{4}\left(\frac{\pi}{4}-\frac{2}{3}\right)=\frac{1}{6},\end{aligned}$$

當中用了開首的結果。因此

$$\int_0^1 x^3\tan^{-1}\left(\frac{1-x}{1+x}\right)\mathrm{d}x=\frac{\pi}{16}-\frac{1}{6}.$$

第 (ii) 部分

想一想 那個古怪的權重 $\frac{(1-y)^3}{(1+y)^5}$ 是從哪裏來的？

它正是 $x^3\,\mathrm{d}x$ 在同一個變換下所變成的樣子。代入 $y=\frac{1-x}{1+x}$，解出 $x=\frac{1-y}{1+y}$：這個映射是自身的逆映射，並把 $x=0$ 送到 $y=1$，把 $x=1$ 送到 $y=0$。於是

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

所以

$$\frac{(1-y)^3}{(1+y)^5}\,\mathrm{d}y=\frac{8x^3}{(1+x)^3}\cdot\frac{(1+x)^5}{32}\cdot\left(-\frac{2}{(1+x)^2}\right)\mathrm{d}x=-\frac{x^3}{2}\,\mathrm{d}x,$$

而 $\tan^{-1}y=\tan^{-1}\left(\frac{1-x}{1+x}\right)$。把積分上下限對調，負號便被抵銷（$y$ 由 0 到 1 即 x 由 1 到 0）：

$$\begin{aligned}\int_0^1\frac{(1-y)^3}{(1+y)^5}\tan^{-1}y\,\mathrm{d}y&=\int_1^0\left(-\frac{x^3}{2}\right)\tan^{-1}\left(\frac{1-x}{1+x}\right)\mathrm{d}x\\&=\frac{1}{2}\int_0^1 x^3\tan^{-1}\left(\frac{1-x}{1+x}\right)\mathrm{d}x=\frac{1}{2}\left(\frac{\pi}{16}-\frac{1}{6}\right)=\frac{\pi}{32}-\frac{1}{12}\end{aligned}$$

這裏用了第 (i) 部分的結果。

驗算

- 兩個答案都是正數，這是必然的（每個被積函數（integrand）在 $[0,1]$ 上均非負）：$\frac{\pi}{16} - \frac{1}{6} \approx 0.02968$ 及 $\frac{\pi}{32} - \frac{1}{12} \approx 0.01484$。對原來的被積函數使用辛普森法則（Simpson's rule），結果吻合至小數點後 12 位。
- (ii) 的答案恰好是 (i) 的一半，你可以在動手做任何代數運算之前，先用數值方法粗略核對一下。
- 在恆等式中代入 $x = 0$：$\tan^{-1} 1 = \frac{\pi}{4}$。代入 $x = 1$：$\tan^{-1} 0 = 0$。兩者都與 $\frac{\pi}{4} - \tan^{-1} x$ 相符。

常見錯誤

- 在 (ii) 中漏掉 $\mathrm{d}y$，或處理錯正負號：因子 $-\frac{2}{(1+x)^2}$ 與對調後的上下限，兩者的負號互相抵銷。
- 未檢查值域便寫下 $\tan^{-1}\left(\frac{1-x}{1+x}\right) = \frac{\pi}{4} - \tan^{-1} x$；當 $x < -1$ 時此式並不成立（兩邊相差 π）。

STEP 之外

代換 $y = \frac{1-x}{1+x}$ 是一個莫比烏斯變換（Möbius transformation），即形如 $\frac{ax+b}{cx+d}$ 的映射。這些映射在複合（composition）運算下構成一個羣（group），而這一個的階（order）為二。同樣的技巧，配合對稱性，可得出著名的結果 $\int_0^1 \frac{\ln(1+x)}{1+x^2}\,\mathrm{d}x = \frac{\pi}{8}\ln 2$，值得接着一試。莫比烏斯映射在複分析（complex analysis）中佔有核心地位；Needham 的 *Visual Complex Analysis*（《複分析：可視化方法》）對它們有一番優美的幾何闡述。

詳細解答

STEP II 2001，第 7 題

純數學 • 複數與旋轉

STEP II 2001，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/01-s2-q7

關鍵思路

沿着正多邊形（regular polygon）繞一圈，每條邊的長度都相同，而在每個頂點轉過的都是同一個外角（exterior angle）。在阿根圖（Argand diagram）上，轉過角 θ 相當於乘以 $\mathrm{e}^{\mathrm{i}\theta}$，所以每條邊都是前一條邊乘以一個固定的複數（complex number）。一旦求出各頂點，圓心便可由垂直平分線（perpendicular bisector）求得，而旋轉（rotation）只不過是再乘一次 $\mathrm{e}^{\mathrm{i}\theta}$。

解答

求 Q 和 S

每個頂點的內角（interior angle）是 $5\pi/6$，所以在每個頂點，前進方向轉過的是外角 $\pi - 5\pi/6 = \pi/6$。由於多邊形位於上半平面，而我們由 O 出發，沿正實軸走到 P，所以每次都必須向逆時針（anticlockwise）方向轉（若在 P 順時針轉，Q 便會落在實軸之下）。

設 $\omega = \mathrm{e}^{\mathrm{i}\pi/6} = \frac{\sqrt{3}}{2} + \frac{1}{2}\mathrm{i}$。邊 $\overrightarrow{OP}$ 就是複數 2，而其後每條邊都是前一條邊乘以 ω（長度不變，轉過 $\pi/6$）。因此

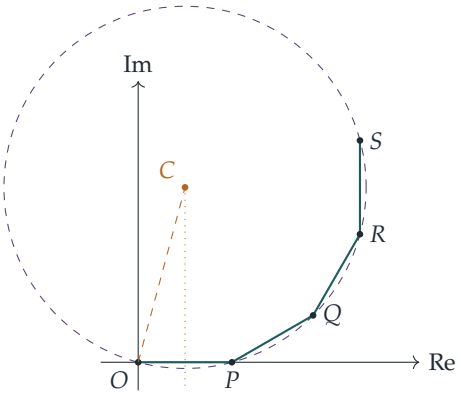
$$\begin{aligned}\overrightarrow{PQ} &= 2\omega = \sqrt{3} + \mathrm{i},\\ \overrightarrow{QR} &= 2\omega^2 = 2\mathrm{e}^{\mathrm{i}\pi/3} = 1 + \sqrt{3}\,\mathrm{i},\\ \overrightarrow{RS} &= 2\omega^3 = 2\mathrm{e}^{\mathrm{i}\pi/2} = 2\mathrm{i}.\end{aligned}$$

逐一相加，

$$\begin{aligned} Q &= 2 + (\sqrt{3} + \mathrm{i}) = 2 + \sqrt{3} + \mathrm{i}, \\ R &= Q + 1 + \sqrt{3}\,\mathrm{i} = (3 + \sqrt{3}) + (1 + \sqrt{3})\mathrm{i}, \\ S &= R + 2\mathrm{i} = (3 + \sqrt{3}) + (3 + \sqrt{3})\mathrm{i}. \end{aligned}$$

第一行就是題目要求的 Q 值，而

$$S = (3 + \sqrt{3})(1 + \mathrm{i}).$$



圓心 C

通過 O、P、Q 的圓就是整個正十二邊形的外接圓（circumcircle），所以 C 是這個多邊形的中心。由於 C 與 O 和 $P = 2$ 等距，它位於垂直平分線 $\operatorname{Re} z = 1$ 上；設 $C = 1 + h\mathrm{i}$。它與 O 和 Q 亦等距：

$$\begin{aligned} |C|^2 = |C - Q|^2 &\iff 1 + h^2 = (1 + \sqrt{3})^2 + (h - 1)^2 \\ &\iff 1 + h^2 = 4 + 2\sqrt{3} + h^2 - 2h + 1 \\ &\iff 2h = 4 + 2\sqrt{3}, \end{aligned}$$

所以 $h = 2 + \sqrt{3}$，而 $C = 1 + (2 + \sqrt{3})\mathrm{i}$，位於第一象限。

想一想 C 的輻角（argument）的準確值是甚麼？

我們需要 $\tan(\arg C) = 2 + \sqrt{3}$。現在

$$\tan \tfrac{5\pi}{12} = \tan\left(\tfrac{\pi}{4} + \tfrac{\pi}{6}\right) = \frac{1 + 1/\sqrt{3}}{1 - 1/\sqrt{3}} = \frac{\sqrt{3} + 1}{\sqrt{3} - 1} = \frac{(\sqrt{3} + 1)^2}{2} = 2 + \sqrt{3},$$

而 C 在第一象限，所以 $\arg C = 5\pi/12$。（從幾何角度看：三角形 OCP 是等腰三角形，在 C 的頂角為 $2\pi/12$，所以兩個底角都是 $\frac{1}{2}(\pi - \pi/6) = 5\pi/12$。）

旋轉

繞 O 逆時針旋轉角 θ，相當於把每一點乘以 $\mathrm{e}^{\mathrm{i}\theta}$，所以 C 會移到輻角為 $5\pi/12+\theta$ 的位置。要 C 落在實軸上，$5\pi/12+\theta$ 必須是 π 的倍數；最小的正數 θ 是

$$\theta=\pi-\tfrac{5\pi}{12}=\tfrac{7\pi}{12},$$

這時 C 落在負實軸上。

現在把 S 寫成模—輻角形式（modulus–argument form）：$1+\mathrm{i}=\sqrt{2}\,\mathrm{e}^{\mathrm{i}\pi/4}$，所以 $S=(3+\sqrt{3})\sqrt{2}\,\mathrm{e}^{\mathrm{i}\pi/4}=(3\sqrt{2}+\sqrt{6})\,\mathrm{e}^{\mathrm{i}\pi/4}$。$S$ 的新位置是

$$\begin{aligned}S'=S\,\mathrm{e}^{7\mathrm{i}\pi/12}&=(3\sqrt{2}+\sqrt{6})\,\mathrm{e}^{\mathrm{i}(3\pi/12+7\pi/12)}=(3\sqrt{2}+\sqrt{6})\,\mathrm{e}^{5\mathrm{i}\pi/6}\\&=(3\sqrt{2}+\sqrt{6})\left(-\tfrac{\sqrt{3}}{2}+\tfrac{1}{2}\mathrm{i}\right)=-\tfrac{1}{2}(3\sqrt{2}+\sqrt{6})(\sqrt{3}-\mathrm{i}),\end{aligned}$$

正是所求。

驗算

- $|C|^2=1+(2+\sqrt{3})^2=8+4\sqrt{3}=(\sqrt{6}+\sqrt{2})^2$，而 $|C-S|^2=(2+\sqrt{3})^2+1$ 亦相同，所以 S 確實在同一個圓上。
- $|S'|=3\sqrt{2}+\sqrt{6}=|S|$，旋轉理應保持模不變。
- 數值上 $S\approx 4.732+4.732\mathrm{i}$，而 $S'\approx -5.7956+3.3461\mathrm{i}$；電腦驗證了所有邊長、各個 $5\pi/6$ 的角，以及 $7\pi/12$ 確是逆時針旋轉時第一個令 C 落在實軸上的角。

常見錯誤

- 轉了內角 $5\pi/6$，而不是外角 $\pi/6$。
- 旋轉了 $-5\pi/12$（順時針），或旋轉了 $\pi/2-5\pi/12$：旋轉是逆時針的，而 C 首次碰到實軸是在負半軸上。

STEP 之外

把乘以 $\mathrm{e}^{\mathrm{i}\theta}$ 看成旋轉，是把複函數視為幾何變換（geometric transformation）的起點：$z\mapsto az+b$ 可給出平面上所有旋轉、放大（enlargement）和平移（translation），而這些變換構成一個羣（group）。正多邊形的頂點是等比級數的和（geometric sum）$2(1+\omega+\cdots+\omega^{k-1})$，與單位根（roots of unity）息息相關。Needham 的 *Visual Complex Analysis*（《複分析：可視化方法》）把這種幾何觀點發揮得淋漓盡致。

詳細解答

STEP II 2001，第 8 題

純數學 • 周期函數與常微分方程

STEP II 2001，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/01-s2-q8

關鍵思路

周期（period）為 1 的函數，在每一個長度為 1 的區間上的形狀都相同，所以它在任何這樣的區間上的積分都是同一個數。因此 F 每次增加的量都相同，都是 $F(1)$，只要 x 增加 1。最後一部分的微分方程（differential equation）是一階線性的，所以利用積分因子（integrating factor）便可把它的解化為 $-F(x)$ 的指數函數，其餘的結論則由 $F(1)$ 為正數這一事實推出。

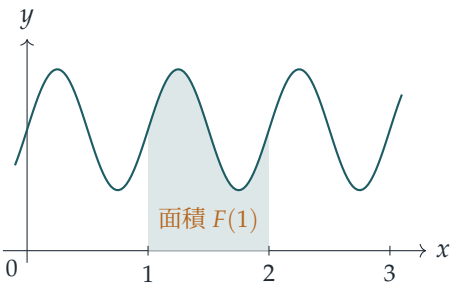
解答

第 (i) 部分

取

$$f(x) = 2 + \sin(2\pi x).$$

由於 sin 的周期為 2π，$\sin(2\pi(x+1)) = \sin(2\pi x + 2\pi) = \sin(2\pi x)$，所以 $f(x+1) = f(x)$。另外 $\sin(2\pi x) \geq -1$，故對每一個 x 都有 $f(x) \geq 1 > 0$。（常數函數如 $f(x) = 1$ 也符合要求，但沒那麼有趣。）



圖像下方每一條單位闊度的長條，面積都相同；第 (ii) 部分說的就是這回事。

第 (ii) 部分

定義 $G(x) = F(x+1) - F(x)$。求導，並對第一項使用鏈式法則（chain rule），得

$$G'(x) = f(x+1) - f(x) = 0$$

對所有 x 成立，這是由 f 的周期性得出的。所以 G 是常數，其值為 $G(0) = F(1) - F(0) = F(1)$，因為 $F(0) = 0$。也就是說，

$$F(x+1) = F(x) + F(1) \quad \text{對所有} x.$$

現在對 n 使用歸納法（induction）。當 $n = 1$ 時，命題 $F(1) = 1 \cdot F(1)$ 顯然成立。若對某個正整數 k 有 $F(k) = kF(1)$，則在上式中代入 $x = k$，得

$$F(k+1) = F(k) + F(1) = kF(1) + F(1) = (k+1)F(1).$$

因此對每一個正整數 n 都有 $F(n) = nF(1)$，命題得證。

（等價的做法是：若 f 連續，則 $F(x) = \int_0^x f(t)\,\mathrm{d}t$，而 $F(n)$ 是 n 個在 $[k, k+1]$ 上的積分之和，每一個積分經代換（substitution）$t \mapsto t + k$ 後都等於 $\int_0^1 f(t)\,\mathrm{d}t$。上面的論證則完全不必把 F 寫成積分。）

第 (iii) 部分

這個方程對 y 而言是線性的，積分因子為 $\mathrm{e}^{\int f\,\mathrm{d}x} = \mathrm{e}^{F(x)}$，$F$ 就是在這裏出場的。事實上，由乘積法則（product rule）及 $F' = f$，

$$\frac{\mathrm{d}}{\mathrm{d}x}\left(y\,\mathrm{e}^{F(x)}\right) = \mathrm{e}^{F(x)}\left(\frac{\mathrm{d}y}{\mathrm{d}x} + f(x)\,y\right) = 0.$$

所以 $y\,\mathrm{e}^{F(x)}$ 是常數。在 $x = 0$ 時它等於 $1 \cdot \mathrm{e}^{F(0)} = 1$，故

$$y = \mathrm{e}^{-F(x)}.$$

由第 (ii) 部分，對每一個正整數 n，

$$y(n) = \mathrm{e}^{-F(n)} = \mathrm{e}^{-nF(1)} = \left(\mathrm{e}^{-F(1)}\right)^n.$$

> **想一想** 我們怎樣知道 $F(1) > 0$？題目並沒有直接告訴我們任何關於 $F(1)$ 的資料。

由於對所有 x 都有 $F'(x) = f(x) > 0$，F 是嚴格遞增（strictly increasing）的，所以 $F(1) > F(0) = 0$。（如果你較喜歡用積分來想：$F(1)$ 是一條正值曲線在 0 與 1 之間下方的面積。）

因此 $0 < \mathrm{e}^{-F(1)} < 1$，而一個滿足 $0 < r < 1$ 的固定數 r 必有 $r^n \to 0$（當 $n \to \infty$ 時）。於是 $y(n) \to 0$（當 $n \to \infty$ 時），這正是我們要證明的。

留意我們只聲稱這個結論沿整數成立。在整數與整數之間，y 被夾住了：由於 F 遞增，y 便遞減，所以事實上 $y(x) \to 0$ 在 $x \to \infty$ 取遍所有實數值時同樣成立。

驗算

取 $f(x) = 2 + \sin(2\pi x)$，則 $F(1) = \int_0^1 (2 + \sin 2\pi t)\,\mathrm{d}t = 2$，因為正弦函數在一個完整周期上的積分為零。用辛普森法則（Simpson's rule）計算，$F(2), F(3), F(5), F(7)$ 分別等於 $4, 6, 10, 14$，誤差小於 10^{-12}。以數值方法（四階龍格—庫塔法，Runge–Kutta）逐步求解這個微分方程，由 $y(0) = 1$ 出發，得 $y(1) \approx 0.135335 = \mathrm{e}^{-2}$ 及 $y(4) \approx 3.3546 \times 10^{-4} = \mathrm{e}^{-8}$，與 $\mathrm{e}^{-nF(1)}$ 吻合。我們亦在非整數的 x 處驗證了 $F(x+1) - F(x) = F(1)$，所用的是對稱性較低的例子 $f(x) = \mathrm{e}^{\cos 2\pi x} + 0.1$。

常見錯誤

- 把 $F(n) = nF(1)$ 當作由周期性「顯然」得出的結論，卻沒有說明為何 $[k, k+1]$ 上的面積等於 $[0,1]$ 上的面積。這裏必須用到函數 $G(x) = F(x+1) - F(x)$ 或代換。
- 忘記說明為何 $F(1) > 0$。沒有這一點，$\mathrm{e}^{-nF(1)}$ 可能一直停留在 1，甚至趨向無窮大。
- 弄錯積分因子的正負號，最後得出 $y = \mathrm{e}^{F(x)}$，它會增長而不是衰減。

STEP 之外

這是弗洛凱理論（Floquet theory）最簡單的情形：對於係數具周期性的線性微分方程，經過一個周期之後的解，等於初始值乘以一個固定的因子（這裏是 $\mathrm{e}^{-F(1)}$），所以長遠的變化取決於這個因子是小於還是大於 1。同樣的想法，把數換成矩陣，便可解釋支點被上下搖動的單擺的穩定性。Boyce 與 DiPrima 合著的 *Elementary Differential Equations*（初等微分方程教科書）對線性方程和積分因子有詳盡的講解，很適合作為下一步的讀物。

詳細解答

STEP II 2001，第 9 題

力學 • 二次空氣阻力

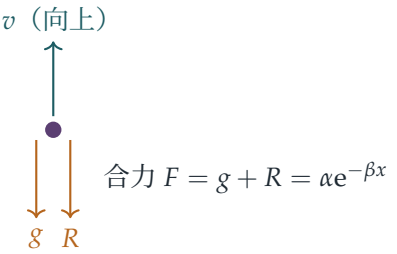
> **STEP II 2001，第 9 題**
>
> 繼續閱讀之前，請先看題目（原題為英文）。題目載於 STEP Support Programme 的試題資料庫：
>
> step.maths.org/questions/01-s2-q9

關鍵思路

題目給出的力是高度的函數，而不是時間的函數，所以最自然的工具是功能原理（work–energy principle）：質點在上升過程中克服 F 所作的功，恰好等於它損失的動能（kinetic energy）。這樣就把 $\mathrm{e}^{-\beta x}$ 與 v^2 聯繫起來，而所得結果顯示空氣阻力（air resistance）只取決於速率。一旦知道阻力與速率的關係，下落的過程就是一條關於 v^2 的標準微分方程（differential equation）。

解答

模型。 質點的質量為一個單位，沿一條鉛垂線運動。上升時，重力和空氣阻力都向下作用，與運動方向相反，所以合力 F（大小為 $\alpha\mathrm{e}^{-\beta x}$）向下作用。取 x 為由拋射點起向上量度的高度。



所耗的能量

到達高度 x 時所耗的能量，就是克服 F 所作的功：

$$\int_0^x \alpha \mathrm{e}^{-\beta s}\,\mathrm{d}s = \frac{\alpha}{\beta}\Big(1-\mathrm{e}^{-\beta x}\Big).$$

這些能量來自質點的動能，所以根據功能原理（單位質量）

$$\tfrac{1}{2}u^2 - \tfrac{1}{2}v^2 = \frac{\alpha}{\beta}\Big(1-\mathrm{e}^{-\beta x}\Big). \qquad (*)$$

（等價地，把牛頓第二定律（Newton's second law）寫成 $v\frac{\mathrm{d}v}{\mathrm{d}x} = -\alpha\mathrm{e}^{-\beta x}$ 的形式，再由 $x=0$、$v=u$ 起積分，也會得到同一條方程。）

以速率表示的力

把 $(*)$ 乘以 β 並整理：

$$\alpha - \alpha\mathrm{e}^{-\beta x} = \tfrac{1}{2}\beta u^2 - \tfrac{1}{2}\beta v^2 \quad\Longrightarrow\quad F = \alpha\mathrm{e}^{-\beta x} = \tfrac{1}{2}\beta v^2 + \alpha - \tfrac{1}{2}\beta u^2,$$

這正是所要求的結果。

想一想 為甚麼必定有 $\alpha = \frac{1}{2}\beta u^2 + g$？題目中似乎沒有任何條件迫使它成立。

把 F 分成兩部分：$F = g + R$，其中 R 是空氣阻力。於是

$$R = F - g = \tfrac{1}{2}\beta v^2 + \Big(\alpha - \tfrac{1}{2}\beta u^2 - g\Big).$$

所以阻力只是速率的函數：一個 v^2 項加上一個常數。空氣阻力是由物體在空氣中運動而產生的，所以當質點在最高點瞬間靜止時，阻力必定為零。在該處代入 $v=0$，得 $R=0$，所以 $F=g$（只有重力作用），而且

$$\alpha - \tfrac{1}{2}\beta u^2 = g, \qquad \text{即} \qquad \alpha = \tfrac{1}{2}\beta u^2 + g.$$

因此常數項為零，阻力定律就簡單地是

$$R = \tfrac{1}{2}\beta v^2.$$

（順帶一提，最高點就是 $\alpha\mathrm{e}^{-\beta x} = g$ 之處，高度為 $\frac{1}{\beta}\ln(\alpha/g)$。）

下落的過程

假設無論運動方向如何，同一條阻力定律 $R = \frac{1}{2}\beta v^2$ 都成立；下落時阻力向上作用，與運動方向相反，而重力仍然向下作用。設 y 為質點在最高點以下的距離，v 為向下的速率，並且在 $y=0$ 時 $v=0$。取向下為正，由牛頓第二定律（單位質量）得：

$$v\frac{\mathrm{d}v}{\mathrm{d}y} = g - \tfrac{1}{2}\beta v^2.$$

由於 $v\frac{\mathrm{d}v}{\mathrm{d}y} = \frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}y}(v^2)$，設 $w = v^2$：

$$\frac{\mathrm{d}w}{\mathrm{d}y} = 2g - \beta w, \qquad w = 0\text{，當}y = 0.$$

這是一條常係數線性方程（也可以用分離變數法）：其解為

$$w = \frac{2g}{\beta}\left(1 - \mathrm{e}^{-\beta y}\right),$$

它確實滿足 $w(0) = 0$ 和 $w' = 2g\mathrm{e}^{-\beta y} = 2g - \beta w$。因此，在最高點以下距離 y 處的空氣阻力為

$$R = \tfrac{1}{2}\beta v^2 = g\left(1 - \mathrm{e}^{-\beta y}\right).$$

驗算

當 $y \to \infty$ 時，$R \to g$：質點趨近終端速率（terminal speed）$\sqrt{2g/\beta}$，此時阻力與重力平衡。在 $y = 0$ 處，$R = 0$，與質點靜止時應有的情況相符。上升時，在 $x = 0$ 處公式給出 $F = \alpha = g + \frac{1}{2}\beta u^2$，即重力加上速率為 u 時的阻力。數值上，取 $g = 9.8$、$\beta = 0.1$、$u = 20$（所以 $\alpha = 29.8$），用龍格－庫塔法（Runge–Kutta method）逐步計算運動方程，所得結果與上升時的 $F = \alpha\mathrm{e}^{-\beta x}$ 和下落時的 $R = g(1 - \mathrm{e}^{-\beta y})$ 相符，誤差約為 10^{-12}，而質點到達的最高高度 ≈ 11.12，與 $\frac{1}{\beta}\ln(\alpha/g)$ 吻合。

常見錯誤

- 把 F 當作只是空氣阻力：它是合力，所以必須減去重力才得到 R。
- 下落時仍然沿用相同的阻力符號：阻力總是與運動方向相反，所以此時它向上作用，方程變為 $v\frac{\mathrm{d}v}{\mathrm{d}y} = g - \frac{1}{2}\beta v^2$。
- 以 v 表示答案，或以 u 和 α 表示答案；題目要求只以 y、g 和 β 表示 R。

STEP 之外

二次阻力（quadratic drag）$R \propto v^2$ 是描述物體在空氣中高速運動的切合實際的模型，而這道題展示了一個巧妙的技巧：以距離為自變數時，v^2 所滿足的方程是線性的，儘管以時間為自變數時 v 所滿足的方程並非線性。大學本科的力學課程還會探討先上升後下落這個問題的不對稱性：質點返回起點時的速率小於 u，而下落所需的時間比上升長。Kleppner 與 Kolenkow 所著的 *An Introduction to Mechanics*（力學導論）詳細討論了取決於速度的力和終端速度。

詳細解答

STEP II 2001，第 10 題

力學 • 碰撞與恢復係數

STEP II 2001，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/01-s2-q10

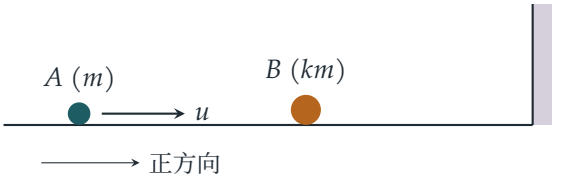
關鍵思路

這題是兩次碰撞加一場追逐。第一次碰撞用的是標準的兩條方程：動量守恆（conservation of momentum）和牛頓實驗定律（Newton's experimental law）。B 從牆上反彈之後，唯一的問題是：正在折返的 B 能否與 A 相遇。這恰好在 A 朝牆方向的速度大於 B 的速度時發生，由此便得出關於 k 的不等式。至於能量部分，我們把動能（kinetic energy）寫成 k 的函數，再與 $mu^2/27$ 比較。比較的結果原來是一個完全平方（perfect square）。

解答

第二次碰撞

把 A 和 B 當作光滑平面上的質點（particle），因此水平方向的衝量（impulse）只來自碰撞。取由 A 指向牆的方向為正方向。設 v_A 和 v_B 分別為 A 和 B 在第一次碰撞剛結束時的速度。



第一次碰撞。 動量守恆，而分離速率等於接近速率的 e_1 倍：

$$mv_A + km\,v_B = mu,$$
$$v_B - v_A = e_1 u.$$

把第二條方程乘以 k 再加到第一條方程，得 $(1+k)v_B = (1+e_1)u$。所以

$$v_B = \frac{(1+e_1)u}{1+k}, \qquad v_A = v_B - e_1 u = \frac{(1-ke_1)u}{1+k}.$$

由於 $v_B > 0$，B 繼續向牆移動。A 則可能向前移動、靜止或向後移動，視乎 $1-ke_1$ 的正負號而定。

牆。 B 以速率 v_B 撞牆，並以速率 $e_2 v_B$ 反彈，所以它的速度變為 $-e_2 v_B$。與此同時，A 繼續以恆定速度 v_A 移動。注意 A 不可能先到達牆，因為 $v_A < v_B$（兩者之差為 $e_1 u > 0$），而且 A 起初在 B 的後面。

想一想 反彈之後，A 和 B 可能朝同一方向移動。我們怎樣判斷它們會否相遇？

反彈之後，B 位於 A 與牆之間，而且再沒有其他作用施加於任何一個質點。所以兩者之間的距離會縮短，當且僅當 A 的速度大於 B 的速度：

$$v_A > -e_2 v_B \iff v_A + e_2 v_B > 0.$$

若 $v_A > 0$，兩者相向而行。若 $v_A \le 0$，則是 B 在追趕 A，而且只有在較快時才追得上。這一條不等式已涵蓋兩種情況。代入後，再乘以正數 $(1+k)/u$，得

$$1 - ke_1 + e_2(1+e_1) > 0 \iff k < \frac{1+e_2(1+e_1)}{e_1},$$

其中最後一步除以了 $e_1 > 0$。所以只要這條不等式成立，便會發生第二次碰撞，正如題目所求。（若等號成立，A 和 B 會以相同速度離開，永不相遇。）

動能

當 $e_1 = \frac{1}{3}$ 及 $e_2 = \frac{1}{2}$ 時，

$$v_A = \frac{(3-k)u}{3(1+k)}, \qquad e_2 v_B = \frac{1}{2}\cdot\frac{\frac{4}{3}u}{1+k} = \frac{2u}{3(1+k)}.$$

在 B 剛反彈之後，A 的速率仍是 $|v_A|$，而 B 的速率是 $e_2 v_B$。所以總動能為

$$\begin{aligned} T &= \tfrac{1}{2}mv_A^2 + \tfrac{1}{2}km\,(e_2 v_B)^2 \\ &= \frac{mu^2}{18(1+k)^2}\left[(3-k)^2 + 4k\right] = \frac{mu^2\,(k^2-2k+9)}{18(1+k)^2}. \end{aligned}$$

現在與 $mu^2/27$ 比較，取公分母 $54(1+k)^2$：

$$\begin{aligned} T - \frac{mu^2}{27} &= \frac{mu^2}{54(1+k)^2}\left[3(k^2-2k+9) - 2(1+k)^2\right] \\ &= \frac{mu^2}{54(1+k)^2}\left(k^2-10k+25\right) = \frac{mu^2\,(k-5)^2}{54(1+k)^2}. \end{aligned}$$

想一想 既然平方永不為負，條件 $k < 5$ 又有甚麼用？

右邊永不為負，而且只有在 $k = 5$ 時才等於零。所以當 $k < 5$ 時，我們有 $(k-5)^2 > 0$，因此

$$T > \frac{mu^2}{27},$$

正如題目所求。5 這個數值並非巧合。代入這些系數後，第一部分的上界是 $\left(1 + \frac{1}{2} \cdot \frac{4}{3}\right)\Big/\frac{1}{3} = 5$。所以 $k < 5$ 正是發生第二次碰撞的條件。當 $k = 5$ 時，兩個質點在反彈後都以 $-u/9$ 移動，而能量恰好是 $mu^2/27$。

驗算

- 取 $e_1 = 1$ 及 $k = 1$（質量相等，完全彈性）。這時 $v_A = 0$ 及 $v_B = u$，即大家熟悉的速度互換。上界變成 $k < 1 + 2e_2$，而只要 $e_2 > 0$，$k = 1$ 便滿足這條件：折返的 B 必定會撞上靜止的 A。
- 當 $k = 5$ 時，公式給出 $v_A = -\frac{u}{9}$ 及 $-e_2 v_B = -\frac{1}{2} \cdot \frac{2u}{9} = -\frac{u}{9}$，兩個速度相等，與上文所說一致。
- 數值方面，以精確分數計算 k 由 0.001 至 4.999 的情況，比值 $T/(mu^2/27)$ 的最小值為 1.00000001，出現在 $k = 5$ 附近，而且始終大於 1。另外隨機測試了 20 000 組 (k, e_1, e_2)，沒有發現不等式與直接比較反彈後速度所得結果不一致的情況。

常見錯誤

- 要求 $v_A > 0$。A 即使正在向後移動，仍然可以被 B 追上。正確的條件是比較反彈後兩者的速度。
- 忘記牆會把 B 的方向反轉。它的速度變為 $-e_2 v_B$，而不是 $e_2 v_B$。
- 計算能量時用了 B 撞牆前的速率。題目要求的是剛反彈後的能量，所以 B 的速率是 $e_2 v_B$。

STEP 之外

牛頓實驗定律是一個模型，用來描述撞擊期間那短暫而劇烈的形變過程。恢復系數（coefficient of restitution）概括了相對動能損失了多少。在大學一年級的力學課程中，碰撞通常會在質心參考系（centre-of-mass frame）中分析。在這個參考系中，碰撞只是把相對速度反轉並乘以 e，而損失的能量為 $\frac{1}{2}\mu(1-e^2)w^2$，其中 μ 是約化質量（reduced

mass），w 是接近速率。Kleppner 與 Kolenkow 所著的 *An Introduction to Mechanics* 對這些概念有細緻的闡述。

詳細解答

STEP II 2001，第 11 題

力學 • 拋體運動

STEP II 2001，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：

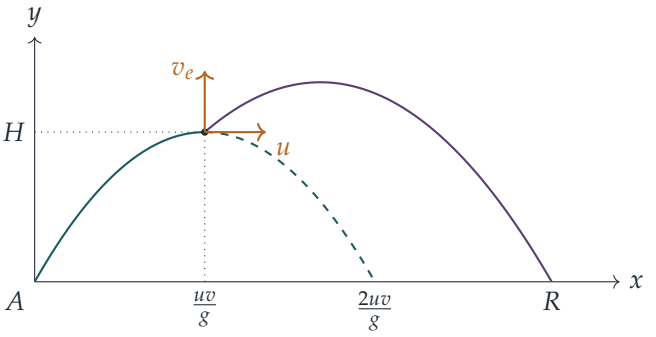
step.maths.org/questions/01-s2-q11

關鍵思路

以爆炸一刻為界，把整段飛行分成兩段。爆炸之前，導彈是普通的拋體（projectile），最高點就是垂直速度為零之處。爆炸之後，碎片 P 是一個新的拋體，從該高度以速度 (u, v_e) 拋出，因此落地所需時間可由一條二次方程（quadratic）求得。至於第二部分，射程（range）是 v_e 的遞增函數（increasing function），所以「$R > D$」就變成「v_e 大於恰好令 $R = D$ 的那個值」。

解答

我們把導彈和碎片 P 視為只受重力作用的質點（particle），並設地面為水平。取 A 為原點，x 軸沿水平方向，y 軸垂直向上。



虛線顯示假如沒有爆炸，導彈本來會落在哪裏；紫色曲線則是 P 的路徑。

P 的射程

爆炸之前。垂直速度為 $v-gt$，在 $t_1 = v/g$ 時變為零。此刻導彈的水平位移為

$$x_1 = ut_1 = \frac{uv}{g},$$

而它的高度，可由 $v^2 = 2gH$（末垂直速率為零）得出：

$$H = \frac{v^2}{2g}.$$

爆炸之後。碎片 P 從高度 H 出發，水平速度為 u，垂直速度為向上的 v_e。以爆炸一刻起計的時間為 t，它的高度是

$$y = H + v_e t - \tfrac{1}{2}gt^2.$$

當 $y = 0$ 時它落地：

$$\tfrac{1}{2}gt^2 - v_e t - \frac{v^2}{2g} = 0 \quad \Longleftrightarrow \quad g^2t^2 - 2gv_e t - v^2 = 0.$$

把它看成 gt 的二次方程，

$$gt = v_e \pm \sqrt{v_e^2 + v^2}.$$

兩根之積為 $-v^2 < 0$，所以其中一個根是負數。這個根代表在爆炸之前拋物線（parabola）與地面相交的時刻，與本題無關。因此爆炸之後的飛行時間是

$$t_2 = \frac{v_e + \sqrt{v_e^2 + v^2}}{g}.$$

水平速度自始至終都是 u，所以總水平距離是

$$R = \frac{uv}{g} + ut_2 = \frac{u}{g}\left(v + v_e + \sqrt{v_e^2 + v^2}\right),$$

亦即

$$\frac{gR}{u} = v + v_e + \sqrt{v_e^2 + v^2},$$

這正是所需的結果。

$R > D$ 的條件

設

$$f(v_e) = v_e + \sqrt{v_e^2 + v^2}, \qquad k = \frac{gD}{u} - v = \frac{gD - uv}{u}.$$

那麼 $R > D$ 等價於 $f(v_e) > k$。由於 $D > 2uv/g$，我們有 $gD > 2uv$，所以 $k > v > 0$。

? **想一想** 為甚麼只需找出令 $R = D$ 的那一個 v_e 值就足夠？

因為 f 是嚴格遞增（strictly increasing）的：

$$f'(v_e) = 1 + \frac{v_e}{\sqrt{v_e^2 + v^2}} > 0,$$

這是由於當 $v \neq 0$ 時 $|v_e| < \sqrt{v_e^2 + v^2}$。（從物理上看，向上的衝擊愈猛，$P$ 留在空中的時間就愈長。）所以若 $f(v_e^*) = k$，則 $f(v_e) > k$ 當且僅當 $v_e > v_e^*$。

要求出 v_e^*，解 $\sqrt{v_e^2 + v^2} = k - v_e$。兩邊平方得

$$v_e^2 + v^2 = k^2 - 2kv_e + v_e^2 \quad \Longrightarrow \quad v_e^* = \frac{k^2 - v^2}{2k} = \frac{(k-v)(k+v)}{2k}.$$

平方可能會引入增根（false root），所以我們要驗證 $k - v_e^* \geq 0$：事實上 $k - v_e^* = \frac{k^2 + v^2}{2k} > 0$，所以 v_e^* 確實滿足 $f(v_e^*) = k$。

現在把 k 代入：

$$k - v = \frac{gD - 2uv}{u}, \qquad k + v = \frac{gD}{u}, \qquad 2k = \frac{2(gD - uv)}{u},$$

所以

$$v_e^* = \frac{(gD - 2uv)\, gD/u^2}{2(gD - uv)/u} = \frac{gD}{2u}\left(\frac{gD - 2uv}{gD - uv}\right).$$

因此，由於 f 是遞增的，只要

$$v_e > \frac{gD}{2u}\left(\frac{gD - 2uv}{gD - uv}\right),$$

便有 $R > D$，這就是所需的條件。事實上，以上論證說明這個條件也是必要的。注意因為 $k > v$，所以 $v_e^* > 0$，即確實需要一個向上的衝擊。

驗算

- 若 $v_e = 0$，公式給出 $gR/u = 2v$，所以 $R = 2uv/g$：這正是沒有爆炸的拋體的一般射程，合乎預期。這也是題目假設 $D > 2uv/g$ 的原因，否則根本不需要任何衝擊。
- 取 $u = v = 10$、$g = 10$ 及 $D = 30$，臨界值為 $v_e^* = \frac{300}{20} \cdot \frac{100}{200} = 7.5$，於是 $gR/u = 10 + 7.5 + 12.5 = 30$，所以 $R = 30 = D$，恰好相等。
- 以數值方法逐步模擬這兩段飛行，所得結果與射程公式相差約 10^{-9}；另外 10^5 個隨機個案亦確認 $R > D$ 當且僅當 $v_e > v_e^*$。

常見錯誤

- 忘記爆炸之前已經走過的水平距離 uv/g，把 R 從爆炸點正下方起計，而不是從 A 起計。

- 把由最高點落地所需的時間又當作 v/g，忽略了新增的向上速度 v_e。
- 把 $\sqrt{v_e^2+v^2}=k-v_e$ 兩邊平方時，沒有驗證右邊為非負，或者沒有解釋為何一個使等式成立的值可以化為一個不等式。

STEP 之外

爆炸過程中整個系統的動量守恆（conservation of momentum），所以即使 P 偏離了原來的拋物線，兩塊碎片的質心（centre of mass）仍然沿着原來的拋物線繼續前進。內力不能改變質心的運動，這個原理正是大學一年級力學課程中火箭運動和碰撞的起點。Kleppner 與 Kolenkow 合著的 *An Introduction to Mechanics*（力學入門教科書）對這兩個課題都有細緻的論述。

詳細解答

STEP II 2001，第 12 題

概率與統計 • 泊松近似

> **STEP II 2001，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/01-s2-q12

關鍵思路

合夥集團（syndicate）必定恰好持有一張中獎彩票，所以一切取決於有多少其他投注者選中了中獎號碼。這個數目服從二項分佈（binomial distribution），試驗次數極多而成功概率極小，正是泊松分佈（Poisson distribution）接手的情況。於是所得獎金的期望值（expectation）就是 W 乘以 $1/(1+X)$ 的期望值，而這個和可以把 $k+1$ 併入階乘（factorial），從而求得閉合形式（closed form）。

解答

建立模型

合夥集團買了 N 張彩票，所以其他投注者買了 $3.8N - N = 2.8N$ 張。設 X 為這些非集團彩票中印有中獎號碼的張數。

每名非集團投注者都是均勻隨機地選一個號碼，與其他人的選擇及開獎結果獨立，所以無論中獎號碼是甚麼，他們每張彩票與之相符的概率都是 $1/N$，而且各張彩票互相獨立。因此

$$X \sim \mathrm{B}\left(2.8N,\ \tfrac{1}{N}\right).$$

由於 N 非常大，$2.8N$ 很大而 $1/N$ 很小，同時平均值 $2.8N \cdot \frac{1}{N} = 2.8$ 保持不變，所以 X 近似服從平均值為 $\lambda = 2.8$ 的泊松分佈：

$$\mathrm{P}(X = k) \approx \frac{\mathrm{e}^{-2.8}\,(2.8)^k}{k!}, \qquad k = 0, 1, 2, \ldots$$

合夥集團每個號碼都恰好持有一張彩票，所以它總是恰好持有一張中獎彩票，而中獎彩票的總數是 $T = 1 + X$。

最可能出現的中獎彩票數目

記 $p_k = \mathrm{P}(X = k)$。比較相鄰的概率，

$$\frac{p_k}{p_{k-1}} = \frac{\lambda}{k} = \frac{2.8}{k},$$

當 $k \le 2$ 時此比值大於 1，當 $k \ge 3$ 時小於 1。所以 $p_0 < p_1 < p_2$ 而 $p_2 > p_3 > p_4 > \cdots$，X 最可能取的值是 2。（對精確的二項分佈，結論相同：此時 $p_k/p_{k-1} = \frac{(2.8N-k+1)}{k(N-1)}$，它大於 1 當且僅當 $k < 2.8 + \frac{1}{N}$。）

因此最可能出現的中獎彩票數目是

$$T = 1 + 2 = 3.$$

淨損失的期望值

頭獎獎金由 $T = 1 + X$ 張中獎彩票平分，而合夥集團持有其中一張，所以它得到 $\dfrac{W}{1+X}$。它付出了 £N，所以淨損失是

$$L = N - \frac{W}{1+X}, \qquad \text{而} \qquad \mathrm{E}(L) = N - W\,\mathrm{E}\left(\frac{1}{1+X}\right).$$

> **想一想** 怎樣求 $\sum \frac{1}{k+1} \cdot \frac{\mathrm{e}^{-\lambda}\lambda^k}{k!}$ 這個和？

竅門在於 $(k+1)\cdot k! = (k+1)!$，所以那個礙事的因子 $\frac{1}{k+1}$ 可以併入階乘。同時乘以和除以 λ，這個和就變成缺了一項的泊松級數：

$$\begin{aligned}\mathrm{E}\left(\frac{1}{1+X}\right) &\approx \sum_{k=0}^{\infty} \frac{1}{k+1} \cdot \frac{\mathrm{e}^{-\lambda}\lambda^k}{k!} = \frac{\mathrm{e}^{-\lambda}}{\lambda} \sum_{k=0}^{\infty} \frac{\lambda^{k+1}}{(k+1)!} \\ &= \frac{\mathrm{e}^{-\lambda}}{\lambda} \sum_{j=1}^{\infty} \frac{\lambda^j}{j!} = \frac{\mathrm{e}^{-\lambda}}{\lambda}\left(\mathrm{e}^{\lambda} - 1\right) = \frac{1 - \mathrm{e}^{-\lambda}}{\lambda},\end{aligned}$$

其中我們令 $j = k+1$，並利用 $\sum_{j\ge 0} \lambda^j/j! = \mathrm{e}^{\lambda}$，再減去 $j = 0$ 的一項，即 1。

代入 $\lambda = 2.8 = \frac{14}{5}$，上式等於 $\frac{5}{14}\left(1 - \mathrm{e}^{-2.8}\right)$，於是

$$\mathrm{E}(L) \approx N - \frac{5\left(1 - \mathrm{e}^{-2.8}\right)}{14} W,$$

正是所求的結果。

驗算

- 同一竅門也適用於精確的二項分佈 $X \sim \mathrm{B}(n, p)$：$\mathrm{E}\left(\frac{1}{1+X}\right) = \dfrac{1 - (1-p)^{n+1}}{(n+1)p}$，當 $n \to \infty$ 而 $np = \lambda$ 時，它趨向 $(1 - \mathrm{e}^{-\lambda})/\lambda$。取 $N = 1000$ 時，精確值是 0.33536，而泊松

近似值是 $\frac{5}{14}(1-\mathrm{e}^{-2.8}) \approx 0.33542$。

- 以 $N = 200$ 模擬 10^5 次開獎，所得獎金的平均值為 $0.335W$，而最常出現的結果是 3 張中獎彩票。
- 極限情況：如果沒有其他人投注（$\lambda \to 0$），則 $(1-\mathrm{e}^{-\lambda})/\lambda \to 1$，合夥集團獨得全部頭獎獎金，與預期相符。

常見錯誤

- 把其他投注者的彩票數目當作 $3.8N$ 而不是 $2.8N$：合夥集團自己的彩票並非隨機，它們恰好貢獻一張中獎彩票。
- 寫成 $\mathrm{E}\left(\frac{W}{1+X}\right) = \frac{W}{1+\mathrm{E}(X)} = \frac{W}{3.8}$。期望值不能直接穿過非線性函數；事實上 $\frac{1}{3.8} \approx 0.263$ 明顯小於 0.335（延森不等式（Jensen's inequality）指出它必定不會較大，因為 $1/(1+x)$ 是凸函數（convex function））。

STEP 之外

以 $\mathrm{Po}(\lambda)$ 逼近 $\mathrm{B}(n, \lambda/n)$，就是所謂的「稀有事件定律」（law of rare events）；大學一年級的課程會證明各個概率收斂，或者證明概率生成函數（probability generating function）收斂，從而證明這個定律。計算 $\mathrm{E}\big(1/(1+X)\big)$ 其實就是對生成函數求積分 $\int_0^1 G_X(s)\,\mathrm{d}s$，這是生成函數在計算平均值和方差（variance）以外的一個巧妙用途。Grimmett 與 Welsh 合著的 *Probability: An Introduction*（概率論入門）涵蓋了這兩個概念。

詳細解答

STEP II 2001，第 13 題

概率與統計 • 連續分佈與估計

STEP II 2001，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/01-s2-q13

關鍵思路

題目直接給出了生存函數（survival function）$\mathrm{P}(T > t)$，所以令它等於 $\frac{1}{2}$ 便可求出中位數（median）；至於平均值（mean），則先求導得出密度函數，再積分。實驗提供了樣本中位數（九個由小至大排列的壽命中的第五個）和樣本平均值，把它們與公式對應起來，便得到關於 α 和 k 的兩條方程。兩式相除可消去 k，餘下一條只含 α 的方程。最後要求的是一個條件概率（conditional probability），而那裏正有一個令人驚喜的結果等着我們。

解答

中位數與平均值

設 $S(t) = \mathrm{P}(T > t) = (1+kt)^{-\alpha}$，其中 $t \geq 0$。要使它由 1 下降至 0，必須有 $k > 0$。

中位數。中位數 M 滿足 $\mathrm{P}(T > M) = \frac{1}{2}$，所以

$$(1+kM)^{\alpha} = 2 \quad \Longrightarrow \quad 1 + kM = 2^{1/\alpha} \quad \Longrightarrow \quad M = \frac{2^{1/\alpha} - 1}{k}.$$

平均值。累積分佈函數（cumulative distribution function）是 $F(t) = 1 - (1+kt)^{-\alpha}$，所以概率密度函數（probability density function）是

$$f(t) = F'(t) = \alpha k\,(1+kt)^{-\alpha-1}, \qquad t \geq 0.$$

然後用分部積分法（integration by parts），取 $u = t$ 和 $\mathrm{d}v/\mathrm{d}t = \alpha k(1+kt)^{-\alpha-1}$，於是

$v = -(1+kt)^{-\alpha}$，

$$m = \int_0^\infty t\,\alpha k(1+kt)^{-\alpha-1}\,\mathrm{d}t = \Big[-t(1+kt)^{-\alpha}\Big]_0^\infty + \int_0^\infty (1+kt)^{-\alpha}\,\mathrm{d}t$$
$$= 0 + \left[-\frac{(1+kt)^{1-\alpha}}{k(\alpha-1)}\right]_0^\infty = \frac{1}{k(\alpha-1)}.$$

兩個在無窮遠處的極限都用到 $\alpha > 1$：$t(1+kt)^{-\alpha}$ 的表現類似 $t^{1-\alpha}/k^\alpha \to 0$，而 $(1+kt)^{1-\alpha} \to 0$。（若 $\alpha \le 1$，平均值便會是無限大，這正是題目堅持要求 $\alpha > 1$ 的原因。）所以

$$M = \frac{2^{1/\alpha}-1}{k}, \qquad m = \frac{1}{k(\alpha-1)}.$$

估計 α 和 k

共有九個燈泡，第五個失效的燈泡正是九個順序排列的壽命中居中的一個：它就是樣本中位數（sample median）。所以我們估計 $M \approx 1000$，並由樣本平均值得 $m \approx 2400$。

想一想 我們有兩條方程和兩個未知數，但 k 在兩條方程中都出現。怎樣才能單獨求出 α？

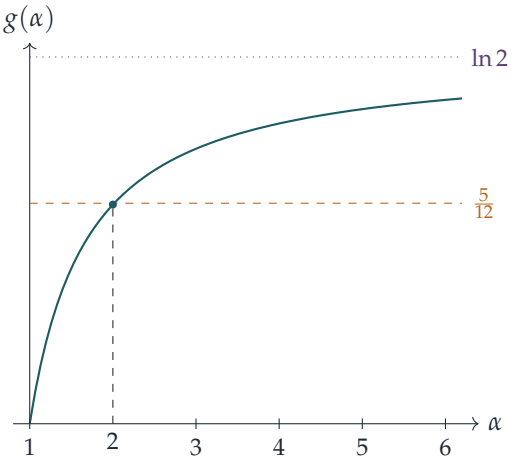
相除：由於 M 和 m 都與 $1/k$ 成正比，k 會被約去。由此得

$$\frac{M}{m} = \big(2^{1/\alpha}-1\big)(\alpha-1) \approx \frac{1000}{2400} = \frac{5}{12} \approx 0.4167.$$

把左方記作 $g(\alpha)$。當 $\alpha = 2$ 時，

$$g(2) = \big(\sqrt{2}-1\big)(2-1) = \sqrt{2}-1 \approx 0.4142,$$

這與 5/12 相差不到 0.003。要看出沒有其他數值同樣吻合，可留意 $g(1.9) \approx 0.396$ 及 $g(2.1) \approx 0.430$，所以 5/12 這個值介乎兩者之間。事實上，g 由 $g(1) = 0$ 開始穩步遞增，當 $\alpha \to \infty$ 時趨向 $\ln 2 \approx 0.693$（見圖），所以它只會取 5/12 這個值一次，而且是在略大於 2 的某個 α 處。因此 $\alpha \approx 2$。



現在把 $\alpha = 2$ 代回任何一條公式。由平均值得

$$k = \frac{1}{m(\alpha - 1)} \approx \frac{1}{2400},$$

而由中位數得 $k = (\sqrt{2} - 1)/1000 \approx 4.14 \times 10^{-4}$。兩者十分接近，於是

$$k \approx 4.2 \times 10^{-4} \text{ 每小時.}$$

條件概率

一個燈泡運作了 M 小時後再多運作 $m - M$ 小時，合共便運作了 m 小時。由於 $m > M$，事件 $\{T > m\}$ 包含於 $\{T > M\}$ 之內，所以

$$\mathrm{P}(T > m \mid T > M) = \frac{\mathrm{P}(T > m)}{\mathrm{P}(T > M)} = \frac{(1 + km)^{-\alpha}}{1/2} = \frac{2}{(1 + km)^{\alpha}}.$$

（嚴格來說，是在 M 小時那一刻發現燈泡仍在運作；對於連續的壽命，$T \geq M$ 與 $T > M$ 的概率相同。）驚喜來了：不論 k 是多少，都恰好有 $km = 1/(\alpha - 1)$，所以 $1 + km = \alpha/(\alpha - 1)$，而

$$\mathrm{P}(T > m \mid T > M) = 2\left(\frac{\alpha - 1}{\alpha}\right)^{\alpha}.$$

當 $\alpha = 2$ 時，這等於 $2 \cdot \left(\frac{1}{2}\right)^2 = \frac{1}{2}$。所以概率的估計值是

$$\mathrm{P}(\text{再多運作} m - M \text{ 小時}) \approx \tfrac{1}{2}.$$

最後這一步完全用不着 k 的估計值。

驗算

- 在中位數處，$S(M) = (2^{1/\alpha})^{-\alpha} = \frac{1}{2}$，與預期相符。隨機選取 $\alpha > 1$ 和 $k > 0$，對 $S(t)$ 在 $[0,\infty)$ 上作數值積分，結果與 $m = 1/(k(\alpha-1))$ 吻合（回想對於非負的變量，有 $\mathrm{E}(T) = \int_0^\infty \mathrm{P}(T > t)\,\mathrm{d}t$，這正是我們計算過程中的第二個積分）。
- 準確地解 $g(\alpha) = 5/12$ 可得 $\alpha \approx 2.015$，這時條件概率是 $2(1-1/\alpha)^\alpha \approx 0.502$，所以把它約為 $\alpha = 2$ 幾乎沒有任何損失。
- 以 $\alpha = 2$、$k = 1/2400$ 模擬 400 000 個壽命（取 $T = (U^{-1/2} - 1)/k$，其中 U 在 $(0,1)$ 上均勻分佈），得出 $\mathrm{P}(T > m \mid T > M) \approx 0.501$。

常見錯誤

- 把 $\mathrm{P}(T > t)$ 當作累積分佈函數，求導時漏掉負號，結果得出負的「密度」。
- 以 $\mathrm{P}(T > m - M)$ 作答，彷彿燈泡在時刻 M 重新開始計算壽命。這個分佈並不具有無記憶性（memoryless；只有指數分佈才有），所以條件 $T > M$ 是有影響的。
- 把第五個失效的時間當作平均值的估計，而不是中位數的估計。

STEP 之外

這是一種帕累托（Pareto）型分佈，稱為 Lomax 分佈，是描述壽命和收入的標準模型，帶有重尾（heavy tail）：$\mathrm{E}(T^r)$ 只在 $r < \alpha$ 時才是有限的。我們在這裏所用的是*矩估計法*（method of moments；嚴格來說，是配對一個分位數和一個矩）：選取參數，使模型的概括統計量等於樣本的概括統計量。大學一年級的統計學會把這個方法與最大似然估計（maximum likelihood）作比較，並探討只憑九個燈泡得出的估計究竟有多準確。比值 $\mathrm{P}(T > s+t \mid T > s)$ 是可靠性理論（reliability theory）中所用的生存函數：在這裏它隨 s *遞增*，所以舊燈泡比新燈泡更有可能繼續運作下去。Grimmett 與 Welsh 合著的 *Probability: An Introduction*（概率入門）是很好的下一步讀物。

詳細解答

STEP II 2001，第 14 題

概率與統計 • 最小方差估計量

> **STEP II 2001，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/01-s2-q14

關鍵思路

T 是兩個獨立（independent）伯努利變量（Bernoulli variable）的加權組合，因此它的平均值（mean）和方差（variance）可直接由期望值的線性性質（linearity of expectation），以及適用於獨立的 X、Y 的法則 $\operatorname{Var}(aX+cY) = a^2\operatorname{Var}(X) + c^2\operatorname{Var}(Y)$ 得出。權重經過特意選取，使得無論 λ 取何值，平均值都是 p：T 是 p 的無偏估計量（unbiased estimator），而題目要找的正是其中最好的一個。至於最後一部分，中心極限定理（central limit theorem）令 $\overline{T}$ 近似服從正態分佈（normal distribution），而平均值固定的正態變量，方差最小時最為集中。

解答

T 的平均值與方差

這裏 X 和 Y 都是伯努利變量，且

$$\mathrm{E}(X) = p, \quad \operatorname{Var}(X) = p(1-p), \qquad \mathrm{E}(Y) = 2p, \quad \operatorname{Var}(Y) = 2p(1-2p).$$

（若一個變量以概率 q 取值 1，否則取值 0，則 $\mathrm{E} = q$ 且 $\mathrm{E}(\cdot^2) = q$，所以方差是 $q - q^2 = q(1-q)$。）由期望值的線性性質，

$$\mathrm{E}(T) = \lambda p + \tfrac{1}{2}(1-\lambda)\cdot 2p = \lambda p + (1-\lambda)p = p,$$

正如所求。由於兩枚硬幣互相獨立，X 和 Y 亦獨立，所以沒有協方差（covariance）項，而

$$\begin{aligned}\operatorname{Var}(T) &= \lambda^2\operatorname{Var}(X) + \tfrac{1}{4}(1-\lambda)^2\operatorname{Var}(Y)\\ &= \lambda^2 p(1-p) + \tfrac{1}{2}(1-\lambda)^2 p(1-2p).\end{aligned}$$

求方差的最小值

記 $V(\lambda) = \mathrm{Var}(T)$。這是 λ 的二次函數，首項係數 $p(1-p) + \frac{1}{2}p(1-2p)$ 為正（因為 $0 < p < \frac{1}{2}$，兩項皆為正），所以它只有一個最小值，出現在 $V'(\lambda) = 0$ 之處：

$$V'(\lambda) = 2\lambda p(1-p) - (1-\lambda)p(1-2p) = 0.$$

兩邊除以 $p > 0$，再把含 λ 的項合併，

$$\lambda\big[2(1-p) + (1-2p)\big] = 1 - 2p \quad \Longrightarrow \quad \lambda = \frac{1-2p}{3-4p},$$

這正是題目給出的值。（這裏 $3 - 4p > 1 > 0$，所以不會出現除以零的情況。）

我們還需要最小值本身。設 $a = p(1-p)$ 及 $c = \frac{1}{2}p(1-2p)$，則 $V = a\lambda^2 + c(1-\lambda)^2$，在 $\lambda = c/(a+c)$（即上面的值）處取得最小值，由此得

$$V_{\min} = a\Big(\frac{c}{a+c}\Big)^2 + c\Big(\frac{a}{a+c}\Big)^2 = \frac{ac(a+c)}{(a+c)^2} = \frac{ac}{a+c}.$$

現在 $a + c = \frac{1}{2}p\big[2(1-p) + (1-2p)\big] = \frac{1}{2}p(3-4p)$ 而 $ac = \frac{1}{2}p^2(1-p)(1-2p)$，所以

$$V_{\min} = \frac{p(1-p)(1-2p)}{3-4p}.$$

n 次擲幣的平均值

設 $T_1, \ldots, T_n$ 為 n 次擲幣中 T 的取值；它們互相獨立，各自的平均值為 p，方差為 $V(\lambda)$。於是

$$\mathrm{E}(\overline{T}) = p, \qquad \mathrm{Var}(\overline{T}) = \frac{V(\lambda)}{n}.$$

想一想 T 只取四個值。為甚麼我們可以把 $\overline{T}$ 當作正態變量來處理？

因為 $\overline{T}$ 是 n 個獨立同分佈（independent, identically distributed）且方差有限的變量的平均值，而 $n > 30$ 夠大，所以根據中心極限定理，不論單獨一個 T 的分佈形狀如何，$\overline{T}$ 都近似服從 $\mathrm{N}\big(p, V(\lambda)/n\big)$。記 $\sigma_\lambda = \sqrt{V(\lambda)/n}$ 及 $Z = (\overline{T} - p)/\sigma_\lambda \approx \mathrm{N}(0,1)$，則

$$\mathrm{P}\big(|\overline{T} - p| < b\big) = \mathrm{P}\Big(|Z| < \frac{b}{\sigma_\lambda}\Big) \approx \Phi\Big(\frac{b}{\sigma_\lambda}\Big) - \Phi\Big(-\frac{b}{\sigma_\lambda}\Big) = 2\Phi\Big(\frac{b}{\sigma_\lambda}\Big) - 1,$$

其中用到對稱性 $\Phi(-z) = 1 - \Phi(z)$。

由於 $b > 0$ 固定，而 Φ 嚴格遞增（strictly increasing），上式在 b/σ_λ 最大時最大，亦即在 σ_λ 最小時最大，也就是 $V(\lambda)$ 取最小值之時。這發生在 $\lambda = (1-2p)/(3-4p)$，此時

$$\sigma_\lambda^2 = \frac{V_{\min}}{n} = \frac{p(1-p)(1-2p)}{(3-4p)n} = s^2.$$

所以當 λ 變動時，$\mathrm{P}\big(|\overline{T} - p| < b\big)$ 的最大值約為 $2\Phi(b/s) - 1$，正如所求。

驗算

- 當 $\lambda = 1$ 時，$T = X$，公式給出 $\operatorname{Var}(T) = p(1-p)$；當 $\lambda = 0$ 時，$T = \frac{1}{2}Y$，公式給出 $\frac{1}{4} \cdot 2p(1-2p)$。兩者都正確。
- $V_{\min}$ 比這兩個值都小：它等於 $p(1-p)$ 乘以 $\frac{1-2p}{3-4p} < 1$。把兩枚硬幣結合起來，勝過單獨使用其中任何一枚。
- 數值上，取 $p = 0.3$、$n = 100$、$b = 0.03$：最佳的 λ 是 $\frac{2}{9}$，$s \approx 0.0216$，而 $2\Phi(b/s) - 1 \approx 0.835$。模擬 40 000 輪、每輪擲幣 100 次，得出 0.836；而 $\lambda = 1$（只用硬幣 A）則只得約 0.487。

常見錯誤

- 寫成 $\operatorname{Var}(\frac{1}{2}(1-\lambda)Y) = \frac{1}{2}(1-\lambda)\operatorname{Var}(Y)$：常數從方差中提出來時要平方。
- 忘記說明爲甚麼把方差減至最小就能使概率最大：這需要 Φ 遞增，以及對每個 λ 而言 $\overline{T}$ 的平均值都是 p。

STEP 之外

這是估計理論（estimation theory）的初次體驗：在 p 的所有無偏估計量 $\lambda X + \frac{1}{2}(1-\lambda)Y$ 之中，我們找出了方差最小的一個。答案是按各個觀測值的方差成反比來加權，這與統計學中用來結合精確度不同的量度結果的法則相同（逆方差加權，inverse-variance weighting），而最小值 $ac/(a+c)$ 正是「並聯電阻」公式 $1/V_{\min} = 1/a + 1/c$。Grimmett 與 Welsh 的 *Probability: An Introduction*（概率論入門教材）涵蓋了最後一步背後的中心極限定理。

詳細解答

STEP III 2001，第 1 題

純數學 • 麥克勞林級數

STEP III 2001，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/01-s3-q1

關鍵思路

函數 $y = \ln\left(x + \sqrt{x^2+1}\right)$ 就是 $\sinh^{-1} x$，而它的一階導數（derivative）遠比函數本身簡單。把這個導數平方，便得到一個不含平方根的關係式 $(x^2+1)(y')^2 = 1$，再對它求導，就得到遞推關係（recurrence）中 $n = 0$ 的情況。每多求導一次，遞推關係便上移一步，這正是歸納法（induction）所需要的。在 $x = 0$ 處，含 x 的一項消失，遞推關係把 $y^{(n+2)}(0)$ 直接與 $y^{(n)}(0)$ 連繫起來，於是所有麥克勞林（Maclaurin）係數幾乎唾手可得。

解答

導數

設 $s = \sqrt{x^2+1}$，則 $\dfrac{\mathrm{d}s}{\mathrm{d}x} = \dfrac{x}{s}$。留意對任意實數 x 都有 $x + s > 0$，因為 $s > \sqrt{x^2} = |x|$，所以這個對數處處有定義。由鏈式法則（chain rule），

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

正如所求。

歸納法

設 $P(n)$ 為以下命題：

$$(x^2+1)\,y^{(n+2)} + (2n+1)\,x\,y^{(n+1)} + n^2\,y^{(n)} = 0 \quad \text{對所有實數} x.$$

基礎情況，$n = 0$。我們需要證明 $(x^2+1)y'' + xy' = 0$。由第一部分，$(x^2+1)(y')^2 =$

1。對 x 求導，

$$2x(y')^2 + (x^2+1)\cdot 2y'y'' = 0.$$

由於 $y' = 1/\sqrt{x^2+1}$ 永不為零，我們可以除以 $2y'$，得到 $(x^2+1)y'' + xy' = 0$。這就是 $P(0)$。（另一做法是直接對 $y' = (x^2+1)^{-1/2}$ 求導：$y'' = -x(x^2+1)^{-3/2}$，所以 $(x^2+1)y'' = -x(x^2+1)^{-1/2} = -xy'$。）

歸納步驟。假設 $P(k)$ 對某個 $k \geq 0$ 成立：

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

此式對所有 x 都成立，所以我們可以對它求導。對首兩項運用積法則（product rule），

$$2x\,y^{(k+2)} + (x^2+1)\,y^{(k+3)} + (2k+1)\,y^{(k+1)} + (2k+1)\,x\,y^{(k+2)} + k^2\,y^{(k+1)} = 0.$$

合併同類項：

$$(x^2+1)\,y^{(k+3)} + (2k+3)\,x\,y^{(k+2)} + (k^2+2k+1)\,y^{(k+1)} = 0,$$

而由於 $2k+3 = 2(k+1)+1$ 及 $k^2+2k+1 = (k+1)^2$，這正正就是 $P(k+1)$。

? 想一想 為甚麼 $y^{(k+2)}$ 的係數剛好增加 2？

對第一項中的 (x^2+1) 求導，會多出一個 $2x\,y^{(k+2)}$，它與第二項的 $(2k+1)x\,y^{(k+2)}$ 合併。同樣地，對第二項中的 x 求導所得的 $(2k+1)$ 與 k^2 合併，而 k^2+2k+1 是一個完全平方（perfect square）。這個遞推關係天生就會複製自己。

由於 $P(0)$ 成立，且對每個 $k \geq 0$ 都有 $P(k) \Rightarrow P(k+1)$，根據歸納法，$P(n)$ 對所有 $n \geq 0$ 成立。

麥克勞林級數

在 $P(n)$ 中代入 $x = 0$。中間一項消失，剩下

$$y^{(n+2)}(0) = -n^2\,y^{(n)}(0) \qquad (n \geq 0).$$

我們需要兩個初始值：$y(0) = \ln(0+1) = 0$ 和 $y'(0) = 1/\sqrt{1} = 1$。

偶數階導數。由 $y(0) = 0$，遞推關係給出 $y''(0) = 0$，繼而 $y^{(4)}(0) = 0$，如此類推：每個偶數階導數在 0 處都等於零。（這亦可由 y 是奇函數（odd function）推出：$y(-x) = \ln\left(\sqrt{x^2+1} - x\right) = -\ln\left(\sqrt{x^2+1} + x\right)$，因為 $\left(\sqrt{x^2+1} - x\right)\left(\sqrt{x^2+1} + x\right) = 1$。）

奇數階導數。由 $y'(0) = 1$ 出發，

$$\begin{aligned} y'''(0) &= -1^2 \cdot y'(0) = -1, \\ y^{(5)}(0) &= -3^2 \cdot y'''(0) = 9, \\ y^{(7)}(0) &= -5^2 \cdot y^{(5)}(0) = -225. \end{aligned}$$

想一想 要由 $y'''(0)$ 得到 $y^{(5)}(0)$，n 應取甚麼值？

要得到 $n+2=5$ 的 $y^{(n+2)}$，我們需要 $n=3$，所以因子是 $-3^2=-9$，而不是 -5^2。這個因子用的是較低階導數的階數。

因此麥克勞林級數（Maclaurin series）$y=\sum y^{(k)}(0)\,x^k/k!$ 為

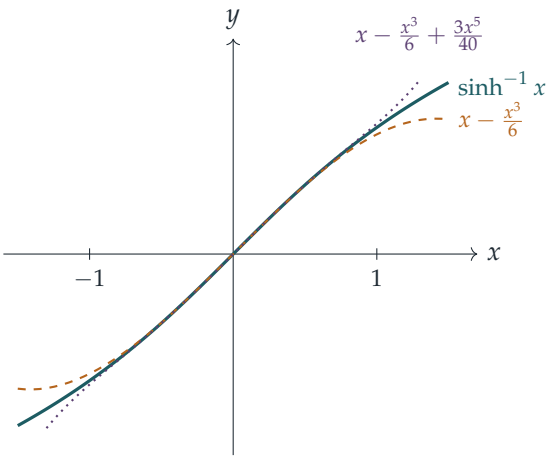
$$\begin{aligned} y &= x+\frac{-1}{3!}x^3+\frac{9}{5!}x^5+\frac{-225}{7!}x^7+\cdots \\ &= x-\frac{x^3}{6}+\frac{9}{120}x^5-\frac{225}{5040}x^7+\cdots \\ &= x-\frac{x^3}{6}+\frac{3x^5}{40}-\frac{5x^7}{112}+\cdots, \end{aligned}$$

其開首各項與題目所求相符，而下一個非零項是

$$-\frac{5x^7}{112}.$$

另一途徑。把二項式級數（binomial series）$(1+x^2)^{-1/2}=1-\frac{1}{2}x^2+\frac{3}{8}x^4-\frac{5}{16}x^6+\cdots$ 逐項積分（對 $|x|<1$ 成立），並利用 $y(0)=0$，便得到 $x-\frac{1}{6}x^3+\frac{3}{40}x^5-\frac{5}{112}x^7+\cdots$，結果一致。

下圖顯示部分和（partial sums）在原點附近如何緊貼 $y=\sinh^{-1}x$，而在較遠處則交替地偏高和偏低。



驗算

▶ 當 $x=0.1$ 時，$\ln\left(0.1+\sqrt{1.01}\right)=0.0998340789$，而 $x-\frac{x^3}{6}+\frac{3x^5}{40}-\frac{5x^7}{112}=0.0998340789$ 亦然；兩者之差約為 3×10^{-11}，即下一項（x^9 項）的大小。

- $n = 0$ 的情況很容易在 $x = 1$ 處檢驗：$y' = 1/\sqrt{2}$ 而 $y'' = -1/(2\sqrt{2})$，所以 $2y'' + y' = 0$。
- 在 0 處的奇數階導數為 $1, -1, 9, -225$，即 $(-1)^m\big(1 \cdot 3 \cdots (2m-1)\big)^2$：這是一個賞心悅目、值得留意的規律。

常見錯誤

- 在 $y^{(n+2)}(0) = -n^2 y^{(n)}(0)$ 中用錯了 n，得出因子 $-9, -25, -49$，而不是 $-1, -9, -25$。
- 忘記除以 $k!$：x^7 的係數是 $-225/5040$，而不是 -225。
- 在歸納步驟中，漏掉了對 $x^2 + 1$ 求導所產生的 $2x\,y^{(k+2)}$。

STEP 之外

泰勒級數（Taylor series）只在它收斂（converge）的地方才有用，而這個級數恰好在 $|x| \leq 1$ 時收斂，儘管 $\sinh^{-1} x$ 在整條實數線上都十分光滑。箇中原因藏在複平面（complex plane）上：$1/\sqrt{1+z^2}$ 在 $z = \pm\mathrm{i}$ 處趨於無窮，這兩點與原點的距離為 1，而這個距離就是收斂半徑（radius of convergence）。Needham 的 *Visual Complex Analysis*（視覺化複分析）對此有精彩的解釋；至於泰勒定理（Taylor's theorem）及其餘項在實變數方面的論述，可參閱 Abbott 的 *Understanding Analysis*（理解分析學）。

詳細解答

STEP III 2001，第 2 題

純數學 • 雙曲代換與面積

STEP III 2001，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/01-s3-q2

關鍵思路

兩個不等式說的都是「留在某條等軸型雙曲線（rectangular-type hyperbola）的其中一側」，所以第一步是畫出這兩條雙曲線的草圖，並找出它們的交點。結果發現，所求區域是一塊有界、具四重對稱性的範圍，而它的面積可寫成兩個形如 $\int \sqrt{x^2 \pm a^2}\,\mathrm{d}x$ 的積分。這類積分正好可用雙曲代換（hyperbolic substitution）來處理，而題目的第一部分，就是為了讓我們把所得的 $\cosh^{-1}$（以及 $\sinh^{-1}$）的值化為對數。

解答

$\cosh^{-1} x$ 的對數形式

設 $x \geq 1$，並設 $y = \cosh^{-1} x$；按定義，它是滿足 $\cosh y = x$ 的 非負數。於是

$$x = \frac{\mathrm{e}^{y} + \mathrm{e}^{-y}}{2} \quad \Longrightarrow \quad \mathrm{e}^{2y} - 2x\,\mathrm{e}^{y} + 1 = 0,$$

這是一條以 e^{y} 為未知數的二次方程（quadratic），所以

$$\mathrm{e}^{y} = x \pm \sqrt{x^2 - 1}.$$

? **想一想** 應該取哪一個符號？為甚麼？

兩個根的積是 1（即二次方程的常數項），而取正號的根至少是 1。因此取負號的根至多是 1，這會得出 $y \leq 0$。由於 $y \geq 0$，我們必須取正號（當 $x = 1$ 時，兩個根本來就相等），所以

$$\cosh^{-1} x = \ln\bigl(x + \sqrt{x^2 - 1}\bigr),$$

這正是所要證明的結果。以 $x = \sinh y$ 作完全相同的論證，可得 $\mathrm{e}^{2y} - 2x\,\mathrm{e}^{y} - 1 = 0$；這次兩個根的積是 -1，所以只有取正號的根是正數，於是 $\sinh^{-1} x = \ln\left(x + \sqrt{x^2+1}\right)$。這兩條公式稍後都會用到。

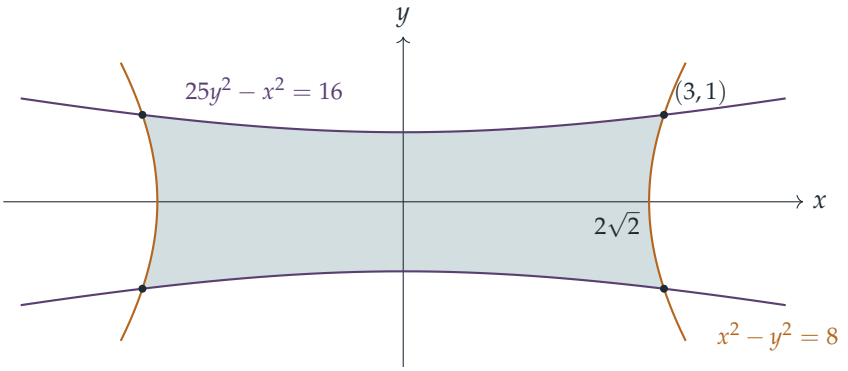
區域

把兩個不等式改寫為

$$x^2 - y^2 \le 8 \qquad \text{及} \qquad 25y^2 - x^2 \le 16.$$

曲線 $x^2 - y^2 = 8$ 是一條雙曲線，兩支分別向左、向右開口，並通過 $(\pm 2\sqrt{2}, 0)$；第一個不等式要求我們留在它的兩支之間，即 $|x| \le \sqrt{y^2+8}$。曲線 $25y^2 - x^2 = 16$ 是一條雙曲線，兩支分別向上、向下開口，並通過 $(0, \pm\frac{4}{5})$；第二個不等式要求我們留在這兩支之間，即 $|y| \le \frac{1}{5}\sqrt{x^2+16}$。

兩條曲線相交之處滿足 $x^2 = y^2 + 8$ 及 $25y^2 = x^2 + 16 = y^2 + 24$，所以 $y^2 = 1$ 及 $x^2 = 9$：交點就是 $(\pm 3, \pm 1)$ 這四點。



這個區域（圖中陰影部分）是有界的：若 (x, y) 在區域內，則 $x^2 \le y^2 + 8 \le \frac{1}{25}(x^2+16) + 8$，所以 $x^2 \le 9$，繼而 $y^2 \le \frac{25}{25} = 1$。兩個不等式在 $x \mapsto -x$ 及 $y \mapsto -y$ 之下都保持不變，所以區域對兩條坐標軸都對稱，其面積是第一象限（first quadrant）內那部分的面積 A 的四倍。

在第一象限內，我們沿垂直方向切割。

- 當 $0 \le x \le 2\sqrt{2}$ 時，我們有 $x^2 - 8 \le 0 \le y^2$，所以第一個不等式自動成立，而 y 由 0 上升至 $\frac{1}{5}\sqrt{x^2+16}$。
- 當 $2\sqrt{2} \le x \le 3$ 時，y 由 $\sqrt{x^2-8}$ 上升至 $\frac{1}{5}\sqrt{x^2+16}$（而根據交點的計算，下方的曲線位於上方的曲線之下，當且僅當 $x^2 \le 9$）。

因此

$$A = \frac{1}{5}\int_0^3 \sqrt{x^2+16}\,\mathrm{d}x \;-\; \int_{2\sqrt{2}}^3 \sqrt{x^2-8}\,\mathrm{d}x = \tfrac{1}{5}I_1 - I_2.$$

第一個積分

代入 $x = 4\sinh s$，於是 $\mathrm{d}x = 4\cosh s\,\mathrm{d}s$ 及 $\sqrt{x^2+16} = 4\cosh s$（是正數，因為 $\cosh s > 0$）。積分上下限是 $s = 0$ 及 $s = S = \sinh^{-1}\frac{3}{4} = \ln\left(\frac{3}{4}+\frac{5}{4}\right) = \ln 2$。利用 $\cosh^2 s = \frac{1}{2}(\cosh 2s + 1)$，

$$I_1 = \int_0^S 16\cosh^2 s\,\mathrm{d}s = \int_0^S 8(\cosh 2s + 1)\,\mathrm{d}s = \Big[8\sinh s\cosh s + 8s\Big]_0^S$$
$$= 8\cdot\tfrac{3}{4}\cdot\tfrac{5}{4} + 8\ln 2 = \tfrac{15}{2} + 8\ln 2,$$

因為 $\sinh S = \frac{3}{4}$ 及 $\cosh S = \sqrt{1+\frac{9}{16}} = \frac{5}{4}$。

第二個積分

代入 $x = 2\sqrt{2}\cosh t$，其中 $t \geq 0$，於是 $\mathrm{d}x = 2\sqrt{2}\sinh t\,\mathrm{d}t$ 及 $\sqrt{x^2-8} = 2\sqrt{2}\sinh t$（是正數，因為 $t \geq 0$）。積分上下限是 $t = 0$ 及

$$t = T = \cosh^{-1}\frac{3}{2\sqrt{2}} = \ln\left(\frac{3}{2\sqrt{2}} + \sqrt{\tfrac{9}{8}-1}\right) = \ln\frac{4}{2\sqrt{2}} = \ln\sqrt{2} = \tfrac{1}{2}\ln 2,$$

這是由第一部分得出的。利用 $\sinh^2 t = \frac{1}{2}(\cosh 2t - 1)$，

$$I_2 = \int_0^T 8\sinh^2 t\,\mathrm{d}t = \int_0^T 4(\cosh 2t - 1)\,\mathrm{d}t = \Big[4\sinh t\cosh t - 4t\Big]_0^T$$
$$= 4\cdot\frac{1}{2\sqrt{2}}\cdot\frac{3}{2\sqrt{2}} - 2\ln 2 = \tfrac{3}{2} - 2\ln 2,$$

因為 $\cosh T = \frac{3}{2\sqrt{2}}$ 及 $\sinh T = \sqrt{\frac{9}{8}-1} = \frac{1}{2\sqrt{2}}$。

綜合結果

$$A = \tfrac{1}{5}\left(\tfrac{15}{2} + 8\ln 2\right) - \left(\tfrac{3}{2} - 2\ln 2\right) = \tfrac{3}{2} + \tfrac{8}{5}\ln 2 - \tfrac{3}{2} + 2\ln 2 = \tfrac{18}{5}\ln 2.$$

代數部分剛好完全抵銷，而總面積是

$$4A = \frac{72}{5}\ln 2,$$

這正是所要證明的結果。

驗算

- $\frac{72}{5}\ln 2 \approx 9.981$。對兩個積分使用森遜法則（Simpson's rule），得出 $I_1 \approx 13.0452$ 及 $I_2 \approx 0.11371$，與 $\frac{15}{2} + 8\ln 2$ 及 $\frac{3}{2} - 2\ln 2$ 吻合；另外，直接在一個細密的網格上數算同時滿足兩個不等式的點，得出的面積約為 9.98。

- 檢查大小是否合理：這個區域可放進矩形 $[-3,3]\times[-1,1]$ 之內，該矩形的面積是 12，而區域只是欠缺了矩形四角幾塊細長的小片，所以面積略小於 10 是相當合理的。
- $\cosh^{-1}1=\ln 1=0$，與應有的結果一致。

常見錯誤

- 把不等式理解為雙曲線「之外」：$y^2\geq x^2-8$ 是 $x^2-y^2=8$ 兩支之間的區域，當中包含 y-軸。
- 忘記第一個不等式只在 $x>2\sqrt{2}$ 時才起作用，因而減去 $\int_0^3\sqrt{x^2-8}\,\mathrm{d}x$，而這個積分甚至不是實數。
- 在 $\mathrm{e}^y=x\pm\sqrt{x^2-1}$ 中取負號，這樣得出的是 $-\cosh^{-1}x$。

STEP 之外

$\int\sqrt{x^2+a^2}\,\mathrm{d}x$ 和 $\int\sqrt{x^2-a^2}\,\mathrm{d}x$ 這兩類積分可用雙曲代換處理，道理與 $\int\sqrt{a^2-x^2}\,\mathrm{d}x$ 可用 $x=a\sin\theta$ 處理一樣：恆等式 $\cosh^2-\sinh^2=1$ 能消去平方根。事實上，雙曲函數（hyperbolic functions）量度的是面積：$(\cosh t,\sinh t)$ 中的參數 t 是單位雙曲線上扇形面積的兩倍，正如 θ 是圓扇形面積的兩倍。Siklos 的 *Advanced Problems in Mathematics*（進階數學難題集）還有很多這類風格的題目，而 Bartle 與 Sherbert 的 *Introduction to Real Analysis*（實分析導論）則說明指數函數、對數函數以至雙曲函數，如何建立在嚴格的基礎之上。

詳細解答

STEP III 2001，第 3 題

純數學 • 二次方程的根

STEP III 2001，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/01-s3-q3

關鍵思路

把關於根（root）的敘述轉化為關於係數（coefficient）的敘述。設兩根為 α 和 β，則 $\alpha+\beta=b$ 且 $\alpha\beta=c$，而判別式（discriminant）可告訴我們兩根何時為實數。至於第 (ii) 部分，從 $f(x)=x^2-bx+c$ 的圖像入手會較為容易：兩根都嚴格介乎 -1 與 1 之間，當且僅當拋物線（parabola）在 ±1 處取正值，且其頂點（vertex）位於兩者之間。

解答

以下一律記 $f(x)=x^2-bx+c$；當兩根為實數時，稱之為 $\alpha\le\beta$，於是

$$\alpha+\beta=b,\qquad \alpha\beta=c,\qquad \alpha,\beta=\tfrac{1}{2}\bigl(b\mp\sqrt{b^2-4c}\bigr).$$

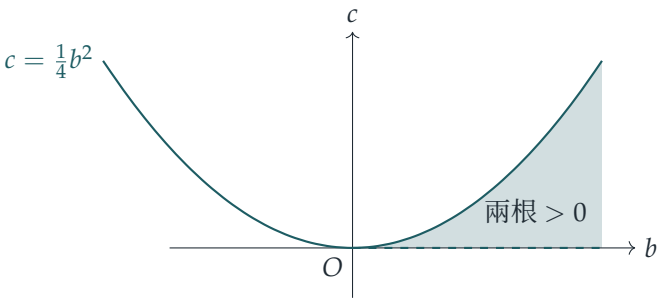
重根（repeated root）視作兩個相等的實根。

第 (i) 部分

($\Rightarrow$)假設兩根皆為正實數。有實根便須 $b^2-4c\ge0$。又因兩根皆為正數，故 $b=\alpha+\beta>0$ 且 $c=\alpha\beta>0$。合起來便得 $b>0$ 及 $b^2\ge4c>0$。

($\Leftarrow$) 假設 $b>0$ 且 $b^2\ge4c>0$。由於 $b^2-4c\ge0$，兩根為實數。由於 $\alpha\beta=c>0$，兩根均不為零且同號。由於 $\alpha+\beta=b>0$，這個共同的符號不可能是負號。所以兩根皆為正數。

這就證明了兩者等價。在 b-c 平面上，這些條件即 $b>0$ 及 $0<c\le\frac{1}{4}b^2$：也就是第一象限（first quadrant）內介乎正 b 軸與拋物線 $c=\frac{1}{4}b^2$ 之間的區域。拋物線包括在內（重根 $\frac{1}{2}b>0$），而 b 軸則不包括在內（該處 $c=0$，其中一根為 0）。



實線曲線（$b > 0$ 的部分）屬於該區域；b 軸上的虛線線段和原點則不屬於該區域。

第 (ii) 部分

我們需要兩根皆為實數，且 $-1 < \alpha \leq \beta < 1$。我們斷言：在兩根為實數的前提下，這情況成立當且僅當

$$f(1) = 1 - b + c > 0, \qquad f(-1) = 1 + b + c > 0, \qquad -1 < \tfrac{1}{2}b < 1.$$

($\Rightarrow$) 若 $-1 < \alpha \leq \beta < 1$，則 $f(1) = (1-\alpha)(1-\beta) > 0$ 且 $f(-1) = (1+\alpha)(1+\beta) > 0$，因為每個括號內的數都是正數。此外，$\frac{1}{2}b = \frac{1}{2}(\alpha + \beta)$ 是 $(-1, 1)$ 內兩個數的平均值，所以它也在 $(-1, 1)$ 內。

($\Leftarrow$) 因為 $f(x) = (x-\alpha)(x-\beta)$ 恰好在 $[\alpha, \beta]$ 上 ≤ 0，所以條件 $f(1) > 0$ 表示 1 不在 $[\alpha, \beta]$ 內。因此要麼 $1 < \alpha$，要麼 $1 > \beta$。頂點 $\frac{1}{2}b$ 位於 $[\alpha, \beta]$ 內，而 $\frac{1}{2}b < 1$，故 $\alpha \leq \frac{1}{2}b < 1$。這就排除了 $1 < \alpha$，所以 $\beta < 1$。同理，由 $f(-1) > 0$ 和 $\frac{1}{2}b > -1$ 可得 $\alpha > -1$。

? **想一想** 為甚麼不能只用 $f(1) > 0$ 和 $f(-1) > 0$ 這兩個條件？

因為兩根有可能都落在區間的同一側。例如 $x^2 - 5x + 6 = 0$ 的根是 2 和 3，但 $f(\pm 1) > 0$。正是頂點條件 $|b| < 2$ 令兩根落在 -1 與 1 之間，而不是同時落在某一端之外。

把這些條件與實根條件 $b^2 \geq 4c$ 合併，所求區域為

$$|b| < 2, \qquad c > b - 1, \qquad c > -b - 1, \qquad c \leq \tfrac{1}{4}b^2,$$

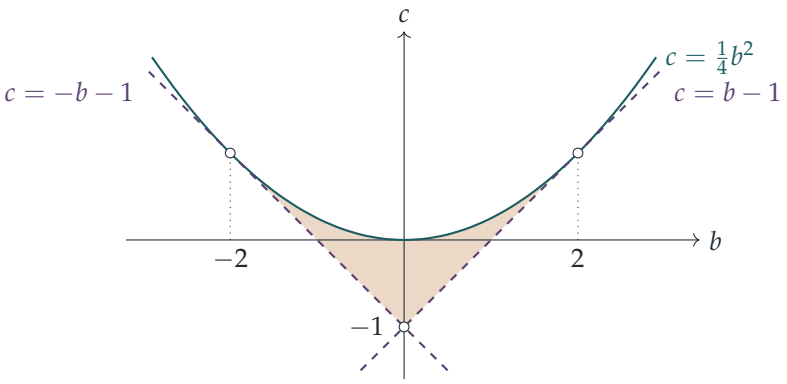
即 $|b| < 2$ 且 $|b| - 1 < c \leq \frac{1}{4}b^2$。

繪畫草圖時，要留意兩條直線與拋物線如何相交。由於

$$\tfrac{1}{4}b^2 - (b-1) = \tfrac{1}{4}(b-2)^2 \geq 0,$$

直線 $c = b - 1$ 位於拋物線下方，且只在 $(2, 1)$ 與它相接。所以它就是拋物線在該點的切線（tangent）。根據對稱性（$b \mapsto -b$），$c = -b - 1$ 是在 $(-2, 1)$ 的切線。兩條直線相交

於 $(0,-1)$。所求區域是一個曲邊三角形，由這兩條切線和拋物線在兩個切點之間的弧圍成。



拋物線上 $|b| < 2$ 的弧包括在內（重根 $\frac{1}{2}b$，且 $|\frac{1}{2}b| < 1$）。兩段虛線切線不包括在內，因為在它們上面有一根為 ± 1。三個角點 $(\pm 2, 1)$ 和 $(0,-1)$ 同樣不包括在內。

驗算

- $(b,c) = (0,-\frac{1}{4})$ 在第 (ii) 部分的區域之內：兩根為 $\pm\frac{1}{2}$。$(b,c) = (0,-1)$ 是被排除的角點：兩根恰好是 ± 1。
- $(b,c) = (2.7, 1.8)$ 滿足 $c > |b| - 1$ 和 $c \le \frac{1}{4}b^2$，但 $|b| > 2$。它的兩根是 1.2 和 1.5，所以頂點條件把它排除在外是正確的。
- 我們以數值方法檢驗了兩個區域：在超過 115 000 個有理點 (b,c) 上，把準確的根與所述條件互相比較，當中包括每條邊界曲線上的點。結果完全吻合。

常見錯誤

- 在第 (ii) 部分遺漏頂點條件 $|b| < 2$，以致錯誤地加入了 $|b| = 2$ 以外、夾在每條切線與拋物線之間的狹長部分。
- 弄錯邊界：兩部分的區域都包括拋物線，但第 (i) 部分的 b 軸和第 (ii) 部分的切線都不包括在內。

STEP 之外

關於 $f(\pm 1)$ 和頂點的條件，其實是一個一般問題的小特例：多項式（polynomial）的所有根何時都落在某個指定區域之內？對於單位圓盤（unit disc），答案由 Schur–Cohn 判別

法（Schur–Cohn test）給出。對於左半平面（left half-plane），答案則是 Routh–Hurwitz 判據（Routh–Hurwitz criterion），控制工程師用它來判斷一個系統是否穩定。第 (ii) 部分的區域，正是遞推關係（recurrence）$u_{n+2} = bu_{n+1} - cu_n$ 在特徵根（characteristic root）為實數時的穩定區域。Siklos 的 *Advanced Problems in Mathematics* 一書中有更多同類的題目。

詳細解答

STEP III 2001，第 4 題

純數學 • 反三角函數

STEP III 2001，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/01-s3-q4

關鍵思路

看到 $\dfrac{2x}{1+x^2}$ 這個式子，應該馬上有所聯想：當 $x = \tan\theta$ 時，它正是 $\sin 2\theta$ 的「t 公式」（t-formula）。代入 $x = \tan\theta$ 之後，這個式子的 $\sin^{-1}$ 就變成 $\sin^{-1}(\sin 2\theta)$，於是整道題就化為：$\sin^{-1}(\sin\phi)$ 在甚麼時候真的會還原為 ϕ？答案取決於 ϕ 是否落在主值範圍（principal range）$[-\frac{1}{2}\pi, \frac{1}{2}\pi]$ 之內，而在 $x = \pm 1$ 處要分段處理，原因正在於此。

解答

第 (i) 部分

先看 $g(x) = \dfrac{2x}{1+x^2}$ 的一些基本性質。

- 分母永不為零，所以 g 在整條實數線上都有定義，而且處處連續（continuous）。
- $g(-x) = -g(x)$，所以 g 是奇函數（odd function），其圖像關於原點旋轉對稱。
- $g(0) = 0$；$g(x)$ 與 x 同號；而且當 $x \to \pm\infty$ 時 $g(x) \to 0$，因為分母的增長與 x^2 同階，而分子只與 x 同階。所以 x 軸是漸近線（asymptote）：當 $x \to \infty$ 時圖像從上方趨近，當 $x \to -\infty$ 時從下方趨近。

由商法則（quotient rule），

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

它只在 $x = \pm 1$ 處為零。當 $|x| < 1$ 時它是正數，當 $|x| > 1$ 時是負數，所以 $x = 1$ 給出極大值 $g(1) = 1$，$x = -1$ 給出極小值 $g(-1) = -1$。在原點的斜率（gradient）是

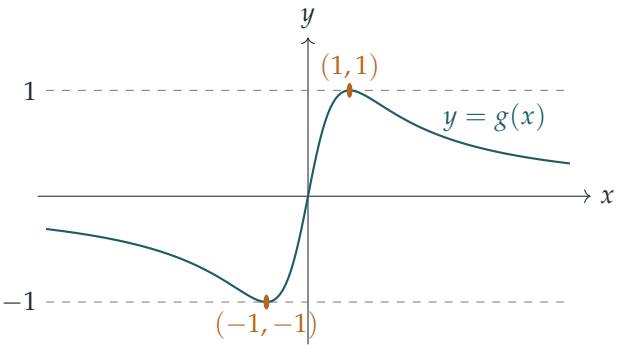
$g'(0) = 2$。

不用微積分，也可以確認這兩個轉向點（turning points）給出的是全局的最大值和最小值：

$$1 - g(x) = \frac{(1-x)^2}{1+x^2} \geq 0, \qquad 1 + g(x) = \frac{(1+x)^2}{1+x^2} \geq 0,$$

所以對所有 x 都有 $-1 \leq g(x) \leq 1$，等號只在 $x = \mp 1$ 時成立。由於 g 連續，並且取到 -1 和 1 這兩個值，所以它取遍兩者之間的每一個值（根據介值定理（intermediate value theorem））。因此

$$\text{函數} g \text{ 的值域（range）為 } -1 \leq g(x) \leq 1.$$



第 (ii) 部分

由第 (i) 部分，$\frac{2x}{1+x^2}$ 總是落在 $[-1, 1]$ 之內，而這正是 $\sin^{-1}$ 的定義域（domain），所以 f 對每個實數 x 都有定義。

設 $\theta = \tan^{-1} x$，即 $x = \tan\theta$，其中 $-\frac{1}{2}\pi < \theta < \frac{1}{2}\pi$。那麼

$$\frac{2x}{1+x^2} = \frac{2\tan\theta}{1+\tan^2\theta} = \frac{2\tan\theta}{\sec^2\theta} = 2\sin\theta\cos\theta = \sin 2\theta,$$

所以

$$f(x) = \sin^{-1}(\sin 2\theta).$$

想一想 $\sin^{-1}(\sin 2\theta)$ 不就是 2θ 嗎？

只有當 2θ 落在 $\sin^{-1}$ 的值域，即 $[-\frac{1}{2}\pi, \frac{1}{2}\pi]$ 之內時才是如此。按定義，$\sin^{-1} y$ 是該區間內正弦值等於 y 的唯一一個角。所以要求 $f(x)$，就必須在 $[-\frac{1}{2}\pi, \frac{1}{2}\pi]$ 內找出與 2θ 正弦值相同的那個角。

情況 $-1 \leq x \leq 1$。由於 tan 在 $(-\frac{1}{2}\pi, \frac{1}{2}\pi)$ 上遞增，這表示 $-\frac{1}{4}\pi \leq \theta \leq \frac{1}{4}\pi$，所以 $-\frac{1}{2}\pi \leq 2\theta \leq \frac{1}{2}\pi$。這時 2θ 本身就是主值範圍內正弦值為 $\sin 2\theta$ 的角，所以

$$f(x) = 2\theta = 2\tan^{-1} x \qquad (-1 \leq x \leq 1).$$

情況 $x \geq 1$。這時 $\frac{1}{4}\pi \leq \theta < \frac{1}{2}\pi$，所以 $\frac{1}{2}\pi \leq 2\theta < \pi$，已經超出主值範圍。不過

$$\sin(\pi - 2\theta) = \sin 2\theta \qquad 及 \qquad 0 < \pi - 2\theta \leq \tfrac{1}{2}\pi,$$

所以 $\pi - 2\theta$ 就是主值範圍內正弦值相符的那個角。因此

$$f(x) = \pi - 2\theta = \pi - 2\tan^{-1} x \qquad (x \geq 1).$$

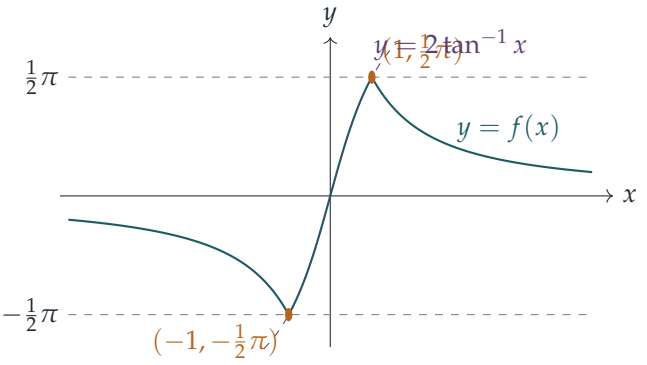
在 $x = 1$ 處，兩條公式都給出 $\frac{1}{2}\pi$，這是理所當然的。至此，題目中「證明」(show that) 的部分已經完成。

要畫出圖像，我們還需要知道 $x \leq -1$ 的情況。由於 g 是奇函數，$\sin^{-1}$ 也是奇函數，所以 f 是奇函數，於是當 $x \leq -1$ 時 $f(x) = -f(-x) = -\pi - 2\tan^{-1} x$。(也可以這樣看：這時 $-\pi < 2\theta \leq -\frac{1}{2}\pi$，而 $-\pi - 2\theta$ 就是 $[-\frac{1}{2}\pi, 0)$ 內正弦值相同的那個角。)

圖像的形狀。

- 在 $[-1, 1]$ 上，$f(x) = 2\tan^{-1} x$ 由 $-\frac{1}{2}\pi$ 遞增至 $\frac{1}{2}\pi$，並以斜率 2 通過原點。
- 當 $x \geq 1$ 時，$f(x) = \pi - 2\tan^{-1} x$ 由 $\frac{1}{2}\pi$ 開始遞減，並且當 $x \to \infty$ 時從上方趨向 $\pi - 2 \cdot \frac{1}{2}\pi = 0$。由奇函數的性質，當 $x \leq -1$ 時，隨着 $x \to -\infty$，它由 $-\frac{1}{2}\pi$ 上升，從下方趨向 0。所以 x 軸同樣是漸近線。
- 當 $|x| < 1$ 時斜率是 $\dfrac{2}{1+x^2}$，當 $|x| > 1$ 時斜率是 $-\dfrac{2}{1+x^2}$。在 $x = 1$ 處，斜率由 $+1$ 跳到 -1（在 $x = -1$ 處情況類似），所以圖像在 $(\pm 1, \pm\frac{1}{2}\pi)$ 處有尖角：f 在該處連續，但不可微（not differentiable）。

這些尖角可以直接從連鎖律（chain rule）看出來：$f'(x) = g'(x)/\sqrt{1 - g(x)^2}$，而第 (i) 部分已給出 $1 - g^2 = (1-x^2)^2/(1+x^2)^2$，它的平方根是 $|1-x^2|/(1+x^2)$。所以 $f'(x) = \dfrac{2(1-x^2)}{(1+x^2)\,|1-x^2|} = \pm\dfrac{2}{1+x^2}$，正負號與 $1 - x^2$ 相同。



虛線顯示 $2\tan^{-1} x$ 越過 $x = 1$ 之後繼續延伸的樣子：它會升到 $\frac{1}{2}\pi$ 以上，而 $\sin^{-1}$ 的值永遠不可能這樣，所以 f 改為向下「反射」回來。

驗算

- 在 $x=\sqrt{3}$ 處：$g(\sqrt{3})=\frac{2\sqrt{3}}{4}=\frac{\sqrt{3}}{2}$，所以 $f(\sqrt{3})=\frac{1}{3}\pi$；而公式給出 $\pi-2\cdot\frac{1}{3}\pi=\frac{1}{3}\pi$。兩者相符。
- 在 $x=10$ 處：$f(10)=\sin^{-1}(20/101)\approx 0.19934$，而 $\pi-2\tan^{-1}10\approx 0.19934$。用電腦在 $[-1,1]$、$[1,10^4]$ 和 $[-10^4,-1]$ 這三個區間內各隨機測試 10^5 個 x 值，結果兩者相符至 10^{-7} 以內，而在 $x=1$ 處的兩個單側斜率則分別算得 $+1$ 和 -1。
- 假如 f 處處都等於 $2\tan^{-1}x$，那麼當 x 很大時它會接近 π，而 $\sin^{-1}$ 的值不可能這樣；反射正好解決了這個問題。

常見錯誤

- 沒有檢查 2θ 是否在 $[-\frac{1}{2}\pi,\frac{1}{2}\pi]$ 之內，就寫下 $\sin^{-1}(\sin 2\theta)=2\theta$。
- 把 f 在 $x=1$ 處畫成平滑的圓頂。圖像在該處其實有一個尖角，因為 $\sqrt{1-g^2}$ 涉及的是 $|1-x^2|$，而不是 $1-x^2$。
- 忘記當 $x\to\pm\infty$ 時 f 趨向 0（而不是 $\pm\pi$）。

STEP 之外

反三角函數（inverse trigonometric functions）是選取分支（branch）的第一個例子：sin 不是一對一（one-to-one）的，所以在求反函數之前必須先限制它的定義域，而像 $\sin^{-1}(\sin\phi)=\phi$ 這樣的恆等式，也就只在該分支上成立。同樣的問題在複分析（complex analysis）中成為核心課題，在那裏 $\log z$ 和 $z^{1/2}$ 都需要分支切割（branch cuts）。$x=1$ 處的尖角也是一個精巧的例子：一個由平滑的片段拼合而成的函數，可以在某一點連續卻不可微。想了解分支和多值函數（multi-valued functions），並看到大量圖解，可參閱 Needham 的 *Visual Complex Analysis*（以圖像講解複分析）；至於連續性和可微性，可參閱 Abbott 的 *Understanding Analysis*（數學分析入門）。

詳細解答

STEP III 2001，第 5 題

純數學 • 拋物線的法線

> **STEP III 2001，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/01-s3-q5

關鍵思路

問有哪些法線（normal）通過定點 (h,k)，其實就是問一個以參數（parameter）t 為未知數的三次方程（cubic）有多少個實根。這個三次方程沒有 t^2 項，所以它正好具有題目開首部分的形式 t^3+pt+q。實根的數目只會在兩個根合併成重根（repeated root）時才改變，因此「一條法線」與「三條法線」之間的分界，就是該三次方程有重根之處。這條分界正是我們要找的軌跡（locus）。

解答

三次方程

設 $f(x)=x^3+px+q$。當 $p\geq 0$ 時，對所有 x 都有 $f'(x)=3x^2+p\geq 0$，而 $f'(x)=0$ 最多只會在 $x=0$ 這一點出現（且僅當 $p=0$ 時）。所以 f 在 $\mathbb{R}$ 上嚴格遞增（strictly increasing）：若 $x_1<x_2$，則 $f(x_2)-f(x_1)=\int_{x_1}^{x_2} f'(x)\,\mathrm{d}x>0$，因為被積函數（integrand）非負，而且最多只在一點為零。

此外，當 $x\to-\infty$ 時 $f(x)\to-\infty$，當 $x\to+\infty$ 時 $f(x)\to+\infty$。由於 f 連續（continuous），它必在某處取值 0；又因為它嚴格遞增，所以只會取這個值一次。因此方程恰有一個實數解。

通過 (h,k) 的法線

在 C 上，$\dfrac{\mathrm{d}y}{\mathrm{d}x}=\dfrac{\mathrm{d}y/\mathrm{d}t}{\mathrm{d}x/\mathrm{d}t}=\dfrac{2a}{2at}=\dfrac{1}{t}$，所以在參數為 t 的點上，法線的斜率（gradient）是 $-t$。（這在 $t=0$ 時同樣成立：該處的切線（tangent）是鉛垂的，而法線就是 x 軸。）法

線的方程是

$$y - 2at = -t\,(x - at^2), \qquad \text{即} \qquad y + tx = 2at + at^3.$$

它通過 (h,k)，當且僅當 $k + th = 2at + at^3$，即

$$\boxed{at^3 + (2a - h)\,t - k = 0} \qquad \text{或者，兩邊同除以} a > 0, \qquad t^3 + \frac{2a-h}{a}\,t - \frac{k}{a} = 0.$$

想一想 兩個不同的 t 值會否給出同一條法線，以致數根的數目時把法線多算了？

不會。在 t 處的法線斜率是 $-t$，所以不同的 t 值給出斜率不同的法線，它們是不同的直線。因此，通過 (h,k) 的法線數目，正好等於該三次方程相異實根的數目。

若 $h \le 2a$，則係數 $p = \dfrac{2a-h}{a}$ 為 ≥ 0，所以由第一部分可知，該三次方程恰有一個實根。因此恰有一條 C 的法線通過 (h,k)，這正是所要證明的。

恰有兩條法線通過的點的軌跡

把三次方程寫成 $t^3 + Pt + Q = 0$，其中 $P = \dfrac{2a-h}{a}$，$Q = -\dfrac{k}{a}$。實係數三次方程的實根，按重數（multiplicity）計算，不是一個就是三個。所以它恰有兩個相異實根，當且僅當它有一個二重根（double root）s 和另一個與之不同的單根（simple root）。

由於沒有 t^2 項，各根之和為 0，所以單根是 $-2s$。我們需要 $s \ne 0$，才能保證 $-2s \ne s$。比較係數，

$$t^3 + Pt + Q = (t-s)^2(t+2s) = t^3 - 3s^2t + 2s^3,$$

所以 $P = -3s^2$ 及 $Q = 2s^3$。換回原來的變數：

$$h = 2a + 3as^2, \qquad k = -2as^3, \qquad s \ne 0.$$

要消去 s，留意 $(h-2a)^3 = 27a^3s^6$ 及 $k^2 = 4a^2s^6$，所以 $4(h-2a)^3 = 108a^3s^6 = 27a\,k^2$。

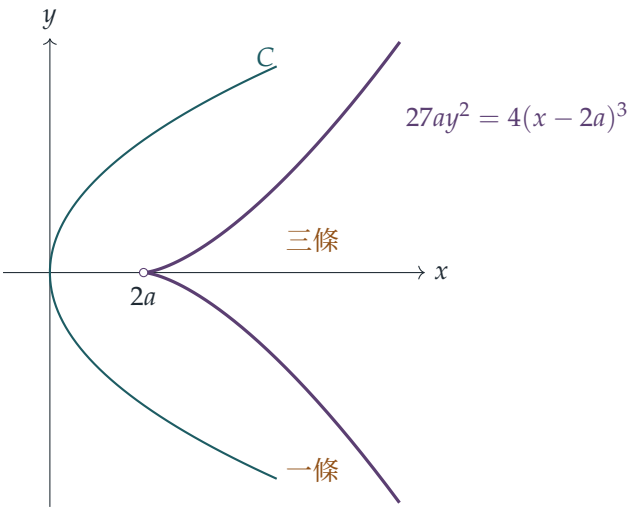
反過來，假設 $27ak^2 = 4(h-2a)^3$，且 $(h,k) \ne (2a,0)$。那麼 $h - 2a \ge 0$，而事實上 $h > 2a$（若 $h = 2a$ 則 $k = 0$）。設 s 為滿足 $2as^3 = -k$ 的實數。那麼 $4a^2s^6 = k^2 = \frac{4}{27a}(h-2a)^3$，所以 $(h-2a)^3 = (3as^2)^3$，因而 $h - 2a = 3as^2$。所以 (h,k) 具有上述形式，其中 $s \ne 0$，並且恰有兩條法線通過它。（這等價於重根條件 $4P^3 + 27Q^2 = 0$。）

點 $(2a,0)$ 本身要排除。在該處 $P = Q = 0$，三次方程是 $t^3 = 0$，有一個三重根（triple root），只有一條法線（即 x 軸）通過該點。

所以所求的軌跡是

$$\boxed{27a\,y^2 = 4\,(x-2a)^3}, \qquad \text{不包括點} (2a,0).$$

這是一條半立方拋物線（semicubical parabola），在 $(2a, 0)$ 處有一個尖點（cusp），關於 x 軸對稱，並位於 $x \geq 2a$ 的範圍內，其中 $y = \pm\dfrac{2}{3\sqrt{3a}}(x-2a)^{3/2}$。在它的內部（即包含正 x 軸上 $2a$ 以外部分的區域），三次方程有三個相異實根，所以每一點都有三條法線通過。在它的外部，則只有一條法線通過。



此草圖取 $a = 1$。空心圓圈標示被排除的尖點，而圖中的標籤則表示通過各區域內的點的法線數目。

驗算

- 取 $a = 1$ 及 $s = 1$，得出點 $(5, -2)$。三次方程是 $t^3 - 3t + 2 = (t-1)^2(t+2)$，所以在 $(1, 2)$ 和 $(4, -4)$ 處的法線都通過 $(5, -2)$。事實上，在 $t = 1$ 處的法線是 $y + x = 3$，而在 $t = -2$ 處的法線是 $y - 2x = -12$。兩者都包含 $(5, -2)$，而且 $27 \cdot 4 = 108 = 4 \cdot 27$。
- 在 $(5, 0.1)$（取 $a = 1$），$4P^3 + 27Q^2 = -108 + 0.27 < 0$，得出三條法線。在 $(5, 4)$，其值為 $-108 + 432 > 0$，得出一條法線。這與草圖吻合。
- 我們以數百個隨機例子，用數值方法檢驗了法線方程以及各個「只有一個根」的結論。

常見錯誤

- 忘記排除尖點 $(2a, 0)$。該點滿足方程，但只有一條法線通過。
- 聲稱當 $p \geq 0$ 時 $f'(x) > 0$。當 $p = 0$ 時，$f'(0) = 0$，所以需要用稍強一點的論證，說明 f 仍然嚴格遞增。

► 把法線的斜率寫成 $1/t$（那是切線的斜率），而不是 $-t$。

STEP 之外

這條軌跡是拋物線的*漸屈線*（evolute）：即其曲率中心（centre of curvature）所描出的曲線，同時也是其法線的包絡線（envelope）。越過漸屈線時，通過一點的法線數目會改變兩條，這是突變理論（catastrophe theory）和三次方程判別式（discriminant）背後思想的一個入門例子。Siklos 的 *Advanced Problems in Mathematics* 收錄了更多同類的 STEP 題目。Bruce 與 Giblin 的 *Curves and Singularities* 則嚴謹而有系統地闡述漸屈線和包絡線。

詳細解答

STEP III 2001，第 6 題

純數學 • 向量與三角形的心

> **STEP III 2001，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/01-s3-q6

關鍵思路

一切都由平面的法向量（normal vector）$\mathbf{n} = \left(\frac{1}{a}, \frac{1}{b}, \frac{1}{c}\right)$ 決定：點 $\mathbf{r}$ 位於平面上，當且僅當 $\mathbf{r}\cdot\mathbf{n} = 1$，而由 O 和 M 所作的垂線都沿 $\mathbf{n}$ 的方向。垂心（orthocentre）和外心（circumcentre）來自兩項立體幾何事實：OA、OB、OC 兩兩互相垂直；而 M 是以 OA、OB、OC 為棱的長方體的中心，所以它與 O、A、B、C 等距。至於最後一部分，兩條垂線互相平行，因此只需一對相似三角形（或一行向量代數）便可完成餘下的工作。

解答

全題中 a、b、c 均不為零（否則平面不會與三條坐標軸都相交），所以

$$A = (a,0,0), \qquad B = (0,b,0), \qquad C = (0,0,c),$$

並記 $S = \dfrac{1}{a^2} + \dfrac{1}{b^2} + \dfrac{1}{c^2} = |\mathbf{n}|^2 > 0$。

形心

直線 OM 上的點是 $t\left(\frac{1}{2}a, \frac{1}{2}b, \frac{1}{2}c\right)$。代入平面方程得 $\frac{1}{2}t + \frac{1}{2}t + \frac{1}{2}t = 1$，所以 $t = \frac{2}{3}$，交點為

$$G = \left(\tfrac{1}{3}a, \tfrac{1}{3}b, \tfrac{1}{3}c\right) = \tfrac{1}{3}(A + B + C),$$

這正是三角形 ABC 的形心（centroid）。留意 $\overrightarrow{OG} = \frac{2}{3}\overrightarrow{OM}$，稍後會用到。

由 O 所作的垂線：垂心

由 O 所作的垂線是直線 $\lambda\mathbf{n}$。它與平面相交於 $\lambda\,\mathbf{n}\cdot\mathbf{n}=1$ 之處，即 $\lambda=1/S$，交點為

$$H=\frac{1}{S}\left(\frac{1}{a},\frac{1}{b},\frac{1}{c}\right).$$

我們證明 $AH\perp BC$。利用 $\overrightarrow{BC}=(0,-b,c)$，

$$\overrightarrow{AH}\cdot\overrightarrow{BC}=\left(\frac{1}{aS}-a\right)\cdot 0+\frac{1}{bS}\cdot(-b)+\frac{1}{cS}\cdot c=-\frac{1}{S}+\frac{1}{S}=0.$$

想一想 這個點積（dot product）是否有理由必定為零，而不只是碰巧如此？

有的。BC 位於該平面上，所以它垂直於法線 OH；而 BC 又位於 yz 平面上，所以它垂直於沿 x 軸的 OA。一條直線若垂直於兩條不平行的直線 OA 和 OH，便垂直於整個平面 OAH，當然也垂直於 AH。

把字母輪換，用同樣的論證可得 $BH\perp CA$ 和 $CH\perp AB$。由於 H 位於三角形所在的平面上，$AH\perp BC$ 即表示 H 在由 A 所作的高（altitude）上，另外兩條高亦同理。所以 H 是三角形 ABC 的垂心。

由 M 所作的垂線：外心

由 M 所作的垂線是 $M+\mu\mathbf{n}$。它與平面相交於 $\frac{3}{2}+\mu S=1$ 之處，所以 $\mu=-\dfrac{1}{2S}$，交點為

$$P=M-\frac{1}{2S}\,\mathbf{n}.$$

現在，M 與 O、A、B、C 四點全部等距：例如

$$MA^2=\left(\tfrac{1}{2}a-a\right)^2+\left(\tfrac{1}{2}b\right)^2+\left(\tfrac{1}{2}c\right)^2=\tfrac{1}{4}\left(a^2+b^2+c^2\right),$$

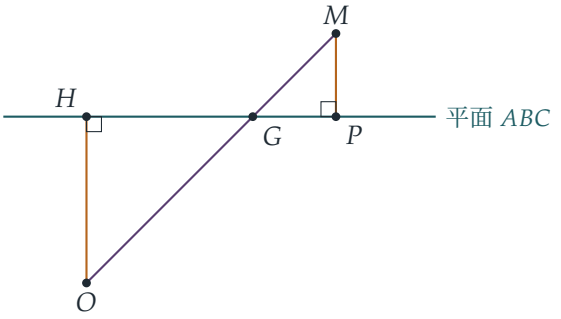
而根據對稱性，MB^2、MC^2（以及 MO^2）都得出相同的值。由於 MP 垂直於平面而 PA 位於平面上，三角形 MPA 在 P 處有一直角，所以根據畢氏定理（Pythagoras' theorem）

$$PA^2=MA^2-MP^2,$$

同樣地 $PB^2=MB^2-MP^2$ 及 $PC^2=MC^2-MP^2$。各式右方相等，所以 $PA=PB=PC$。平面上與三個頂點等距的點只有外心，所以 P 是三角形 ABC 的外心。

最後一部分

垂線 OH 和 MP 都平行於 $\mathbf{n}$，所以 O、H、M、P 位於同一平面上，而直線 OM 在 G 處穿過三角形所在的平面。此外，O 和 M 分別位於三角形所在平面的兩側，因為 $O\cdot\mathbf{n}=0<1$ 而 $M\cdot\mathbf{n}=\frac{3}{2}>1$。以下是通過這四點的截面：



三角形 OHG 和 MPG 分別在 H 和 P 處有直角，在 G 處的對頂角亦相等，所以兩者相似（similar），比為 $OG : GM = 2 : 1$。因此 $HG : GP = 2 : 1$，而且 G 位於線段 HP 上。同一結論也可直接由向量得出。以 $H = \frac{1}{S}\mathbf{n}$、$P = M - \frac{1}{2S}\mathbf{n}$ 及 $G = \frac{2}{3}M$ 代入，

$$\tfrac{1}{3}(H + 2P) = \tfrac{1}{3}\left(\tfrac{1}{S}\mathbf{n} + 2M - \tfrac{1}{S}\mathbf{n}\right) = \tfrac{2}{3}M = G.$$

所以 G 位於由垂心 H 至外心 P 的線段上，並把它按 $HG : GP = 2 : 1$ 的比分割，正如題目所求。（若三角形是等邊三角形，則 $H = P = G$，無需證明。）

想一想 這是否已為每一個三角形證明了這個結果？

並不盡然。要令這個論證適用，三角形必須與某個上述類型的三角形 ABC 全等（congruent）。這時它的邊滿足 $BC^2 = b^2 + c^2$、$CA^2 = c^2 + a^2$ 及 $AB^2 = a^2 + b^2$，所以

$$AB^2 + AC^2 - BC^2 = 2a^2 > 0,$$

在 B 和 C 處亦同理。根據餘弦定理（cosine rule），每個角都是銳角。反過來說，任何銳角三角形都可以這樣實現：由這三條方程解出 a^2、b^2、c^2，它們全是正數。所以這個方法為銳角三角形證明了結果。至於直角或鈍角三角形，以下是一個簡短的平面論證，可涵蓋所有情況。把外心置於原點，則頂點 $\mathbf{a}$、$\mathbf{b}$、$\mathbf{c}$ 的長度全部相同。設 $\mathbf{h} = \mathbf{a} + \mathbf{b} + \mathbf{c}$。那麼

$$(\mathbf{h} - \mathbf{a}) \cdot (\mathbf{b} - \mathbf{c}) = (\mathbf{b} + \mathbf{c}) \cdot (\mathbf{b} - \mathbf{c}) = |\mathbf{b}|^2 - |\mathbf{c}|^2 = 0,$$

所以 $\mathbf{h}$ 在由 $\mathbf{a}$ 所作的高上，根據對稱性亦在另外兩條高上：它就是垂心。形心是 $\frac{1}{3}\mathbf{h} = \frac{1}{3}(\mathbf{h} + 2 \cdot \mathbf{0})$，這正是同一個 $2 : 1$ 的結論。

驗算

取 $a = 1$、$b = c = 2$，則 $S = \frac{3}{2}$。於是 $H = \left(\frac{2}{3}, \frac{1}{3}, \frac{1}{3}\right)$、$P = \left(\frac{1}{6}, \frac{5}{6}, \frac{5}{6}\right)$ 及 $G = \left(\frac{1}{3}, \frac{2}{3}, \frac{2}{3}\right)$。各點均滿足 $x + \frac{1}{2}y + \frac{1}{2}z = 1$。你可以驗證 $\overrightarrow{AH} \cdot \overrightarrow{BC} = 0$ 和 $PA^2 = PB^2 = PC^2 = \frac{25}{12}$，以及 $\frac{1}{3}(H + 2P) = \left(\frac{1}{3}, \frac{2}{3}, \frac{2}{3}\right) = G$。我們亦曾隨機選取 3000 組正負皆有的 a、b、c，以精確算術驗證了以上所有事實。

常見錯誤

- 只證明 $AH \perp BC$ 便停下來。你需要證明全部三條高，或清楚寫明「根據對稱性」。你還需要指出 H 位於三角形所在的平面上：平面以外的點也可以滿足 $AH \perp BC$，卻不在高上。
- 把比的次序弄反。形心與垂心的距離，是它與外心的距離的兩倍。
- 聲稱「由此」（hence）的部分涵蓋每一個三角形。如上文所示，它只適用於銳角三角形。

STEP 之外

通過垂心、形心和外心的直線稱為歐拉線（Euler line）。$\mathbf{h} = \mathbf{a} + \mathbf{b} + \mathbf{c}$（以外心為原點）這個事實，是研究九點圓（nine-point circle）的起點；九點圓的圓心同樣在歐拉線上，位於 H 與外心的正中間。把平面圖形看成一個簡單立體圖形的截面，這種技巧在幾何中屢見不鮮。考克斯特（Coxeter）的 *Introduction to Geometry*（《幾何學導論》），以及考克斯特與 Greitzer 合著的 *Geometry Revisited*（《幾何學重溫》），對歐拉線和九點圓都有精彩的論述。

詳細解答

STEP III 2001，第 7 題

純數學 • 正交軌線

> **STEP III 2001，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/01-s3-q7

關鍵思路

這條微分方程（differential equation）是可分離變數（separable）的，所以只要把它解出來，就能直接看出這族曲線是甚麼。至於正交（orthogonal）曲線，切線（tangent）互相垂直即表示兩個斜率（gradient）的乘積為 -1。由此得到第二條可分離變數的方程，而對它積分時，開首所描繪的那個函數便會出現。題目先要我們畫第一幅略圖，正是為了讓我們幾乎不費工夫就能畫出最後的曲線。

解答

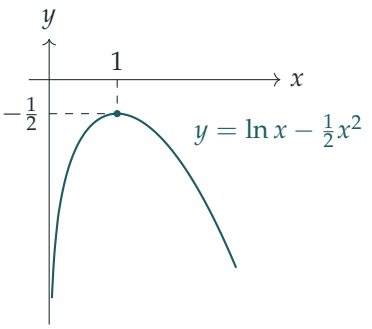
$\ln x - \frac{1}{2}x^2$ 的圖像

設 $f(x) = \ln x - \frac{1}{2}x^2$，定義域為 $x > 0$。則

$$f'(x) = \frac{1}{x} - x = \frac{1 - x^2}{x}, \qquad f''(x) = -\frac{1}{x^2} - 1 < 0.$$

當 $x > 0$ 時，唯一的駐點（stationary point）位於 $x = 1$，該處 $f(1) = -\frac{1}{2}$。由於 $f'' < 0$ 處處成立，這是一個極大值（maximum），而且整個圖像都是向下凹（concave）的。所以對任何 $x > 0$ 都有 $f(x) \le -\frac{1}{2} < 0$，圖像永不與 x 軸相交。

在定義域的兩端，當 $x \to 0^+$ 時 $f(x) \to -\infty$（因為 $\ln x \to -\infty$），所以 y 軸是一條垂直漸近線（vertical asymptote）。另外，當 $x \to \infty$ 時也有 $f(x) \to -\infty$，因為 $\frac{1}{2}x^2$ 的增長遠比 $\ln x$ 快。



拋物線族

常數函數 $y = 0$ 是一個解。在其他情況下，我們分離變數，並避開 $x = \pm 1$，因為方程在該處沒有定義：

$$\int \frac{\mathrm{d}y}{y} = \int \frac{2x}{x^2-1}\,\mathrm{d}x \quad \Longrightarrow \quad \ln|y| = \ln|x^2-1| + c.$$

所以 $|y| = \mathrm{e}^c\,|x^2-1|$。把正負號併入常數，再連同 $y = 0$ 一併考慮，方程的解為

$$y = A(x^2-1), \qquad A \text{ 為常數.}$$

反過來，每一條這樣的曲線都滿足該方程，因為

$$y' = 2Ax = \frac{2x \cdot A(x^2-1)}{x^2-1} = \frac{2xy}{x^2-1}.$$

當 $A \neq 0$ 時，每條曲線都是拋物線（parabola）。它通過 $(1,0)$ 和 $(-1,0)$，因為在該兩點 $x^2-1=0$。它的斜率 $2Ax$ 只在 $x=0$ 時為零，所以它的頂點（vertex）是 $(0,-A)$，位於 y 軸上。這就是題目所求的曲線族。

正交曲線

取一點 (x,y)，其中 $x \neq \pm 1$ 且 $y \neq 0$。曲線族中恰好有一條拋物線通過這一點，即 $A = y/(x^2-1)$ 的那一條，而它在該點的斜率為 $\dfrac{2xy}{x^2-1}$。所求的曲線若要與這條拋物線成直角相交，兩個斜率的乘積必須為 -1，所以該曲線必須滿足

$$\frac{\mathrm{d}y}{\mathrm{d}x} = -\frac{x^2-1}{2xy} = \frac{1-x^2}{2xy}.$$

? **想一想** 為甚麼可以用拋物線的微分方程，而不用公式 $y = A(x^2-1)$ 來求斜率？

微分方程只用 x 和 y 就表示出通過 (x,y) 的那條族中曲線的斜率，完全不涉及 A。這正是我們所需要的，因為正交曲線在它的每一點上所遇到的都是不同的拋物線。

再次分離變數，

$$2y\frac{\mathrm{d}y}{\mathrm{d}x} = \frac{1}{x} - x \quad \Longrightarrow \quad y^2 = \ln x - \tfrac{1}{2}x^2 + C,$$

這裏取 $x > 0$，因為曲線通過 $(1,1)$。右邊是 $f(x) + C$，即我們在開首所描繪的函數。在 $(1,1)$ 處須有 $1 = 0 - \frac{1}{2} + C$，所以 $C = \frac{3}{2}$，曲線為

$$\boxed{y^2 = \ln x - \tfrac{1}{2}x^2 + \tfrac{3}{2}} \qquad \text{或等價地} \qquad y^2 = f(x) + \tfrac{3}{2}.$$

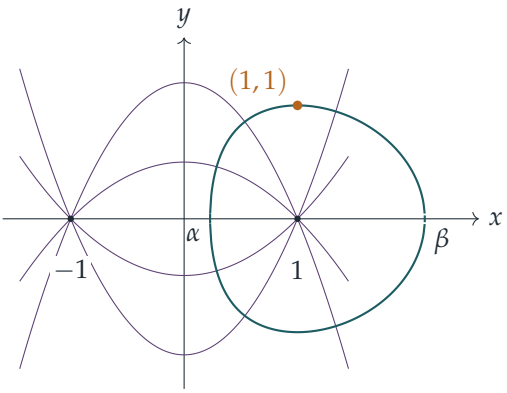
嚴格來說，微分方程通過 $(1,1)$ 的解是上半部分 $y = \sqrt{f(x) + \frac{3}{2}}$。它對 x 軸的反射同樣與每條拋物線正交，所以自然的答案是整條閉曲線。

描繪曲線

實數點須滿足 $f(x) \geq -\frac{3}{2}$。我們的第一幅略圖顯示，f 由 $-\infty$ 上升至 $x = 1$ 處的極大值 $-\frac{1}{2}$，然後再下降回 $-\infty$。所以 $f(x) = -\frac{3}{2}$ 恰好有一個根 $\alpha \in (0,1)$ 和一個根 $\beta > 1$。數值上，$\alpha \approx 0.229$，$\beta \approx 2.123$。因此：

- 曲線位於帶狀區域 $\alpha \leq x \leq \beta$ 之內，並且對 x 軸對稱；
- y^2 在 $x = 1$ 處最大，其值為 1，由此得到兩點 $(1, \pm 1)$，該處的切線是水平的；
- 在 $(\alpha, 0)$ 和 $(\beta, 0)$ 處，我們有 $y^2 = 0$，而 $\mathrm{d}(y^2)/\mathrm{d}x = f'(x) \neq 0$，所以 $\mathrm{d}y/\mathrm{d}x$ 為無窮大，切線是垂直的。

所以這條曲線是一個不對稱的閉合卵形：左邊陡峭上升，右邊則較平緩地下降。圖中也畫出了幾條拋物線以作比較。



假如我們容許 $x < 0$，並把 $\ln x$ 換成 $\ln|x|$，同樣的運算還會在左半平面得出一個鏡像的卵形。那個卵形並不通過 $(1,1)$，所以不屬於答案的一部分。

驗算

- 在 $(1,1)$ 處，曲線的斜率為 $\frac{1-1}{2}=0$。沒有任何拋物線通過 $(1,1)$ 這一點本身，因為每條拋物線與直線 $x=1$ 只相交於 $(1,0)$。在 $x=1$ 附近，拋物線幾乎是垂直的，這與曲線上的水平切線吻合。
- 數值上，在 $x=0.5$ 處，曲線上 $y=\sqrt{\ln 0.5-0.125+1.5}\approx 0.826$，斜率為 $\frac{0.75}{2(0.5)(0.826)}\approx 0.908$。通過該點的拋物線的斜率為 $\frac{2(0.5)(0.826)}{-0.75}\approx -1.101$，而 $0.908\times(-1.101)\approx -1$。
- 我們在卵形上數千個隨機點驗證了正交性，並以二分法（bisection）計算出兩個根 $\alpha\approx 0.2291$ 和 $\beta\approx 2.1226$。

常見錯誤

- 使用某一條固定拋物線的斜率，例如取特定 A 值的 $2Ax$，而不是通過一般點 (x,y) 的那條族中曲線的斜率。
- 忘記 $y^2=f(x)+\frac{3}{2}$ 需要 $f(x)\geq -\frac{3}{2}$，因而畫出一條無界的曲線，而不是閉合的卵形。

STEP 之外

求一族曲線，使它與給定的曲線族成直角相交，這就是正交軌線（orthogonal trajectories）問題。在物理學中，它們以場線和等勢線的形式出現，例如在靜電學或二維流體流動之中。解析（analytic）複變函數的實部和虛部總會給出這樣的兩族曲線，這是通往保角映射（conformal mapping）的其中一條途徑。Boyce 與 DiPrima 的 *Elementary Differential Equations* 涵蓋可分離變數方程和正交軌線。Needham 的 *Visual Complex Analysis* 則展示了複變數的圖景。

詳細解答

STEP III 2001，第 8 題

純數學 • 複數軌跡

> **STEP III 2001，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/01-s3-q8

關鍵思路

設 $z = x + \mathrm{i}y$ 並展開：(i) 中的兩條方程都會化簡成同一個圓。至於 (ii) 中的輻角（argument）條件，可直接計算 $\frac{z-2}{z}$；輻角為 $\frac{\pi}{4}$ 的意思是「實部（real part）等於虛部（imaginary part），而且兩者皆為正數」，這就是同一個圓再加上一個不等式。在 (iii) 中，把 $z = 2/w$ 代入圓的兩種形式之中運算最簡潔的那一種。(i) 中的第二種形式正是為此而設。

解答

第 (i) 部分

設 $z = x + \mathrm{i}y$。第一條方程是

$$(x-1)^2 + (y-1)^2 = 2 \iff x^2 + y^2 - 2x - 2y = 0.$$

第二條方程是

$$\begin{aligned} & (x-1)^2 + (y+1)^2 = 2\left[(x-1)^2 + y^2\right] \\ \iff\ & x^2 + y^2 - 2x + 2y + 2 = 2x^2 + 2y^2 - 4x + 2 \\ \iff\ & 0 = x^2 + y^2 - 2x - 2y. \end{aligned}$$

每一步都可逆，所以兩條方程都等價於 $x^2 + y^2 - 2x - 2y = 0$，描述的是同一個軌跡（locus）。這個軌跡是以 $1 + \mathrm{i}$ 為圓心、半徑為 $\sqrt{2}$ 的圓。它通過 0、2 和 2i（以及 $2 + 2\mathrm{i}$），最低點是 $1 + (1 - \sqrt{2})\mathrm{i}$，剛好在實軸（real axis）之下。

第 (ii) 部分

當 $z \neq 0$ 時，

$$\frac{z-2}{z} = \frac{(z-2)\bar{z}}{|z|^2} = \frac{x^2+y^2-2x+2\mathrm{i}y}{x^2+y^2},$$

因為 $(z-2)\bar{z} = |z|^2 - 2\bar{z} = x^2+y^2-2x+2\mathrm{i}y$。

> **想一想** $\arg(\text{某數}) = \frac{\pi}{4}$ 究竟說明了它的實部和虛部有甚麼關係？

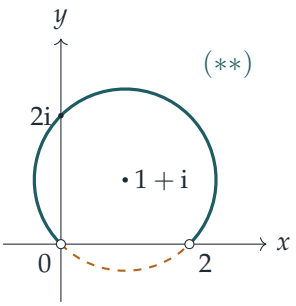
一個非零複數（complex number）$a+\mathrm{i}b$ 的輻角為 $\frac{\pi}{4}$，當且僅當它位於射線 $a = b > 0$ 上。單有相等並不足夠：$a = b < 0$ 時輻角是 $-\frac{3\pi}{4}$。分母 x^2+y^2 為正數，所以 $(**)$ 成立當且僅當

$$x^2+y^2-2x = 2y \quad \text{及} \quad 2y > 0,$$

即當且僅當 z 位於 (i) 的圓上而且 $y > 0$。條件 $y > 0$ 自動排除了 $z = 0$（該處商沒有定義）和 $z = 2$（該處商為零，沒有輻角）。

因此 $(**)$ 描述的是圓在實軸以上的部分：由 0 出發，繞經 2i 和 $2+2\mathrm{i}$ 到達 2 的優弧（major arc），兩個端點都不包括在內。實軸以下的短弧所對應的輻角則是 $-\frac{3\pi}{4}$。

這與幾何上的理解吻合。$\arg\frac{z-2}{z}$ 是在 z 處、連往 0 和連往 2 的兩條弦之間的夾角。由 0 到 2 的弦在圓心 $1+\mathrm{i}$ 所對的角是直角，所以根據圓心角定理（angle-at-the-centre theorem），它在優弧上每一點所對的角都是 $\frac{\pi}{4}$。



整個圓是 $(*)$；實線弧（不包括端點 0 和 2）是 $(**)$，虛線弧則不屬於 $(**)$。

第 (iii) 部分

由於 $w = 2/z$，我們有 $z \neq 0$ 及 $z = 2/w$。採用圓的第二種形式 $|z-(1-\mathrm{i})|^2 = 2|z-1|^2$。代入後得

$$z-1 = \frac{2-w}{w}, \qquad z-(1-\mathrm{i}) = \frac{2-(1-\mathrm{i})w}{w} = \frac{(1-\mathrm{i})\big[(1+\mathrm{i})-w\big]}{w},$$

其中用到 $\frac{2}{1-\mathrm{i}} = 1+\mathrm{i}$。由於 $|1-\mathrm{i}|^2 = 2$，方程變成

$$\frac{2\,|w-(1+\mathrm{i})|^2}{|w|^2} = \frac{2\,|w-2|^2}{|w|^2} \iff |w-(1+\mathrm{i})| = |w-2|.$$

所以 w 與 $1+\mathrm{i}$ 和 2 等距：軌跡是連接這兩點的線段的垂直平分線（perpendicular bisector）。中點是 $\frac{3}{2}+\frac{1}{2}\mathrm{i}$，線段的方向是 $1-\mathrm{i}$，所以平分線的方向是 $1+\mathrm{i}$。設 $w = u+\mathrm{i}v$，軌跡就是直線

$$v = u - 1.$$

這條直線上的每一點都會出現。線上每個 w 都非零，因為 0 不在這條直線上，而 $z = 2/w$ 則由上面各個可逆的步驟可知滿足 $(*)$。圓上唯一沒有像（image）的點是 $z = 0$，它被「送到無窮遠」。一如所料，通過原點的圓在 $w = 2/z$ 之下映射成一條直線。

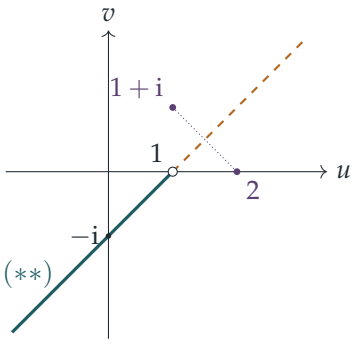
至於對應 $(**)$ 的部分，留意

$$\frac{z-2}{z} = 1 - \frac{2}{z} = 1 - w,$$

所以 $(**)$ 變成 $\arg(1-w) = \frac{\pi}{4}$，即 $1-w = r\,\mathrm{e}^{\mathrm{i}\pi/4}$，其中 $r > 0$。因此

$$w = 1 - \frac{r}{\sqrt{2}}(1+\mathrm{i}), \qquad r > 0,$$

這是直線 $v = u-1$ 上由 $w = 1$（不包括在內）出發、向左下方延伸的半直線（half-line）：即 $u < 1$ 的點，等價地說是 $v < 0$ 的點。這與 (ii) 吻合：$\operatorname{Im} w = -2y/|z|^2$，所以 $y > 0$ 對應 $v < 0$。例如，$z = 2\mathrm{i}$ 映射到 $w = -\mathrm{i}$，而 $z = 2$ 則映射到被排除的端點 $w = 1$。



整條直線 $v = u-1$ 是 $(*)$ 的像；實線半直線（$u < 1$，不包括 $w = 1$）是 $(**)$ 的像。點線線段連接的是被這條直線平分的兩點。

驗算

- $z = 2\mathrm{i}$ 在圓上，而且在實軸之上：$\frac{2\mathrm{i}-2}{2\mathrm{i}} = 1+\mathrm{i}$，其輻角為 $\frac{\pi}{4}$。它的像 $w = -\mathrm{i}$ 滿足 $v = u-1$，且 $u < 1$。

- 圓的最高點 $z = 1 + (1+\sqrt{2})\mathrm{i}$ 映射到 $w \approx 0.2929 - 0.7071\mathrm{i}$，而確實有 $0.2929 - 1 = -0.7071$。
- 實軸以下的一點 $z = 1 + (1-\sqrt{2})\mathrm{i}$ 給出 $\arg\frac{z-2}{z} = -\frac{3\pi}{4}$。它的像有 $u > 1$，位於直線的另一半。

常見錯誤

- 把 $\arg Z = \frac{\pi}{4}$ 僅僅當作 $\operatorname{Re} Z = \operatorname{Im} Z$，於是聲稱答案是整個圓。實部和虛部還須同為正數。
- 在 (ii) 中忘記排除 $z = 0$ 和 $z = 2$，或在 (iii) 中忘記排除 $w = 1$。
- 以為圓的像必定是另一個圓。在 $w = 2/z$ 之下，通過原點的圓會變成直線。

STEP 之外

映射 $w = 2/z$ 是一個莫比烏斯變換（Möbius transformation）。這類映射把圓和直線映射成圓和直線，而一個圓變成直線，當且僅當它通過那個被送到無窮遠的點。這裏弧的兩個端點 0 和 2 分別去到 ∞ 和 1。商 $\frac{z-2}{z-0}$ 量度的是這兩個端點所對的角，它變成了 $1 - w$：由像點指向 1 的方向。這就是 (ii) 中的弧在 (iii) 中變成半直線的原因。Needham 的 *Visual Complex Analysis*（《複分析：可視化方法》）把這種幾何觀點闡述得十分精彩。

詳細解答

STEP III 2001，第 9 題

力學 • 摩擦力與簡諧運動

STEP III 2001，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/01-s3-q9

關鍵思路

兩條橫樑沿一條與水平成 α 角的直線相距 d，因此木板的傾角就是 α。木板滑動時，它的重量由 B_1 逐漸轉移到 B_2，而兩個法向反作用力（normal reaction）隨重心（centre of gravity）的位置線性地變化。由於兩個摩擦系數（coefficient of friction）不同，總摩擦力也隨之線性變化；回復力（restoring force）與位移成正比，這就是簡諧運動（simple harmonic motion）。條件 $\mu_1 + \mu_2 = 2\tan\alpha$ 的作用，正是令這個運動的中心剛好落在兩條橫樑的正中間。

解答

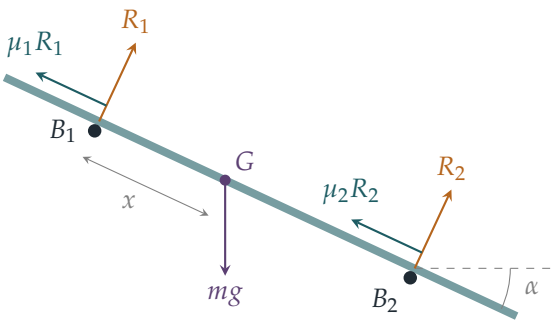
模型

把木板視為質量為 m 的剛性桿（rigid rod），它始終與兩條橫樑保持接觸，而橫樑則視為點。由 B_1 到 B_2 的直線長度為

$$\sqrt{d^2\sin^2\alpha + d^2\cos^2\alpha} = d$$

並與水平成 α 角，所以木板沿這條直線擺放，由 B_1 向下傾斜至 B_2。設 x 為重心 G 與 B_1 的距離，沿木板向斜下方量度，於是釋放時 $x = 0$，而當 $x = d$ 時 G 位於 B_2 之上。

當 $0 \le x \le d$ 時，G 位於兩條橫樑之間。作用於木板的力有：作用於 G 的重量 mg；在 B_1、B_2 處垂直於木板的法向反作用力 R_1、R_2；以及當木板沿斜面向下運動時，沿斜面向上的摩擦力 $\mu_1 R_1$ 和 $\mu_2 R_2$。



法向反作用力

木板在垂直於自身的方向上沒有運動，所以沿垂直於木板的方向分解力，得

$$R_1 + R_2 = mg\cos\alpha.$$

木板也沒有轉動，所以對 G 取力矩（moment）（重量和兩個摩擦力的作用線或通過 G，或沿着木板，因此它們沒有力矩），得

$$R_1 x = R_2(d - x).$$

解這兩條方程，得

$$R_1 = \frac{mg\cos\alpha\,(d-x)}{d}, \qquad R_2 = \frac{mg\cos\alpha\,x}{d}.$$

當 $0 \le x \le d$ 時兩者皆非負，所以木板確實一直擱在兩條橫樑上。

運動方程

當木板沿斜面向下運動時，沿斜面向下分解力，得

$$\begin{aligned} m\ddot{x} &= mg\sin\alpha - \mu_1 R_1 - \mu_2 R_2 \\ &= mg\cos\alpha\left[\tan\alpha - \mu_1 - \frac{(\mu_2-\mu_1)x}{d}\right]. \end{aligned}$$

現在使用題目給出的條件：由 $\mu_1 + \mu_2 = 2\tan\alpha$ 得 $\tan\alpha - \mu_1 = \frac{1}{2}(\mu_2 - \mu_1)$。因此

$$\ddot{x} = -\omega^2\left(x - \tfrac{1}{2}d\right), \qquad \text{其中} \qquad \omega^2 = \frac{g(\mu_2-\mu_1)\cos\alpha}{d} > 0,$$

不等式成立是因為 $\mu_2 > \mu_1$ 且 $\cos\alpha > 0$。

? **想一想** 木板真的會開始運動嗎？摩擦力是否一直指向斜面上方？

在 $x = 0$ 處，$\ddot{x} = \frac{1}{2}\omega^2 d > 0$，所以木板確實會開始向下滑。上面的方程描述的是以 $x = \frac{1}{2}d$ 為中心的簡諧運動，而由於 $t = 0$ 時 $x = 0$、$\dot{x} = 0$，它的解為

$$x = \tfrac{1}{2}d\,(1 - \cos\omega t), \qquad \dot{x} = \tfrac{1}{2}d\,\omega \sin\omega t.$$

當 $0 < t < \pi/\omega$ 時有 $\dot{x} > 0$，所以木板在整段時間內的確一直沿斜面向下運動，假設摩擦力沿斜面向上是正確的。此外，整段時間內 $0 \le x \le d$，所以 G 一直位於兩條橫樑之間，R_1、R_2 的表達式仍然有效。

靜止下來

速度第一次再度為零是在 $\omega t = \pi$ 之時，這時 $x = \frac{1}{2}d(1 - \cos\pi) = d$：重心正好位於 B_2 之上。在這一刻 $R_1 = 0$ 而 $R_2 = mg\cos\alpha$，所以木板只擱在 B_2 上。它會停留在那裏，因為 B_2 處的摩擦力最大可達 $\mu_2 mg\cos\alpha$，而重量沿斜面向下的分量是 $mg\sin\alpha$，並且

$$\mu_2 = \tfrac{1}{2}(\mu_1 + \mu_2) + \tfrac{1}{2}(\mu_2 - \mu_1) > \tfrac{1}{2}(\mu_1 + \mu_2) = \tan\alpha.$$

所以木板靜止下來時，重心位於 B_2 之上，所需時間為

$$\frac{\pi}{\omega} = \pi\left(\frac{d}{g(\mu_2 - \mu_1)\cos\alpha}\right)^{\frac{1}{2}},$$

這正是所要證明的結果。

驗算

- 若 $\mu_2 - \mu_1$ 很小，木板幾乎不減速，所需時間很長，與公式相符。這時每個 μ 都接近 $\tan\alpha$，所以合力在各處都接近零。
- 我們隨機選取 α、d、μ_1（並取 $\mu_2 = 2\tan\alpha - \mu_1$），以數值方法逐步求解運動方程，直至速度第一次回到零為止。例如，取 $\alpha = 0.7$、$d = 2$、$\mu_1 = 0.3$ 及 $g = 9.8$，預測的時間約為 1.56 秒。在每一個情況中，停止的時間和位置都與 π/ω 和 $x = d$ 吻合，（相對）誤差在 10^{-5} 以內。

常見錯誤

- 取 $R_1 = R_2 = \frac{1}{2}mg\cos\alpha$。反作用力取決於 G 的位置，而正是這種依賴關係產生了振盪。
- 只做到「速度為零」便停筆，沒有檢查木板其後是否保持靜止：你需要 $\mu_2 \ge \tan\alpha$。

STEP 之外

這個運動是一個振子（oscillator）的半個週期，而它的「彈簧」來自隨負載而變的摩擦力。摩擦力取決於位置或速度的系統會產生豐富多變的現象，包括黏滑振盪（stick–slip oscillation），門鉸吱吱作響和琴弓拉動小提琴弦線，背後都是這種振盪。Kleppner 與 Kolenkow 所著的 *An Introduction to Mechanics*（力學入門）對摩擦、剛體平衡和簡諧運動都有細緻的論述。

詳細解答

STEP III 2001，第 10 題

力學 • 相對速度

> **STEP III 2001，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/01-s3-q10

關鍵思路

相對速度（relative velocity）就是兩個速度之差：A 相對於 C 的速度是 $\mathbf{v}_1 - \mathbf{u}$。只要把兩個條件都寫成點積（dot product），一切便可從中點 $\frac{1}{2}(\mathbf{v}_1 + \mathbf{v}_2)$ 量起，而兩個條件就化為平方差的形式。從幾何角度看，這兩條方程描述的是一個圓和一條直線，這正是「兩個速度」這個數目的來由。

解答

兩個條件

我們把船隻視為在海面這個水平面上運動的質點（particle），因此所有速度都是二維向量。A 和 B 相對於 C 的速度分別是

$$\mathbf{a} = \mathbf{v}_1 - \mathbf{u}, \qquad \mathbf{b} = \mathbf{v}_2 - \mathbf{u}.$$

大小（magnitude）相等和互相垂直（perpendicular）即是

$$|\mathbf{v}_1 - \mathbf{u}|^2 = |\mathbf{v}_2 - \mathbf{u}|^2 \qquad 及 \qquad (\mathbf{v}_1 - \mathbf{u}) \cdot (\mathbf{v}_2 - \mathbf{u}) = 0.$$

兩條方程

設

$$\mathbf{w} = \mathbf{u} - \tfrac{1}{2}(\mathbf{v}_1 + \mathbf{v}_2), \qquad \mathbf{d} = \tfrac{1}{2}(\mathbf{v}_1 - \mathbf{v}_2).$$

於是 $\mathbf{v}_1 = \frac{1}{2}(\mathbf{v}_1 + \mathbf{v}_2) + \mathbf{d}$ 及 $\mathbf{v}_2 = \frac{1}{2}(\mathbf{v}_1 + \mathbf{v}_2) - \mathbf{d}$，所以

$$\mathbf{a} = \mathbf{d} - \mathbf{w}, \qquad \mathbf{b} = -\mathbf{d} - \mathbf{w}.$$

現在利用 $\mathbf{w}\cdot\mathbf{d}=\mathbf{d}\cdot\mathbf{w}$，把兩個條件逐一展開：

$$\mathbf{a}\cdot\mathbf{b}=-(\mathbf{d}-\mathbf{w})\cdot(\mathbf{d}+\mathbf{w})=|\mathbf{w}|^2-|\mathbf{d}|^2,$$
$$|\mathbf{a}|^2-|\mathbf{b}|^2=(|\mathbf{d}|^2-2\mathbf{d}\cdot\mathbf{w}+|\mathbf{w}|^2)-(|\mathbf{d}|^2+2\mathbf{d}\cdot\mathbf{w}+|\mathbf{w}|^2)=-4\,\mathbf{w}\cdot\mathbf{d}.$$

因此，垂直條件 $\mathbf{a}\cdot\mathbf{b}=0$ 恰好等價於 $|\mathbf{w}|^2=|\mathbf{d}|^2$，即

$$\left|\mathbf{u}-\tfrac{1}{2}(\mathbf{v}_1+\mathbf{v}_2)\right|^2=\left|\tfrac{1}{2}(\mathbf{v}_1-\mathbf{v}_2)\right|^2,$$

而大小相等的條件 $|\mathbf{a}|^2=|\mathbf{b}|^2$ 恰好等價於 $\mathbf{w}\cdot\mathbf{d}=0$，（乘以 2 後）即

$$\Big(\mathbf{u}-\tfrac{1}{2}(\mathbf{v}_1+\mathbf{v}_2)\Big)\cdot(\mathbf{v}_1-\mathbf{v}_2)=0.$$

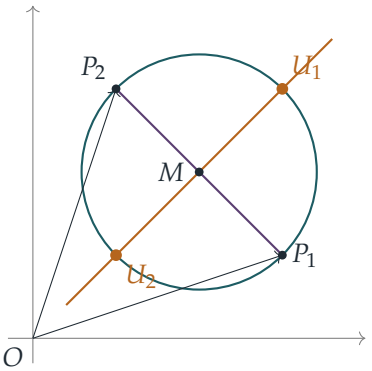
這就是所求的兩條方程。以上每一步都是恆等式（identity），所以論證也可以倒過來走：兩條方程成立，當且僅當兩個條件成立。

為何 C 有兩個可能的速度

把各個速度畫成從同一個原點出發的位置向量（position vector），設 P_1、P_2、U 分別為 $\mathbf{v}_1$、$\mathbf{v}_2$、$\mathbf{u}$ 的終點，並設 M 為 P_1P_2 的中點。

? **想一想** 每條方程分別說明點 U 可以在甚麼位置？

第一條方程說明 U 與 M 的距離是 $\frac{1}{2}|\mathbf{v}_1-\mathbf{v}_2|$：它位於以 P_1P_2 為直徑的圓上。第二條方程說明 MU 垂直於 P_1P_2（或 $U=M$）：U 位於 P_1P_2 的垂直平分線（perpendicular bisector）上。



在平面上，一條通過圓心的直線與一個半徑為正數的圓恰好相交於兩點，這兩點位於一條直徑的兩端。這裏的半徑是 $\frac{1}{2}|\mathbf{v}_1-\mathbf{v}_2|$，它是正數，正正因為 $\mathbf{v}_1\neq\mathbf{v}_2$。所以 C 恰好有兩個可能的速度：

$$\mathbf{u}=\tfrac{1}{2}(\mathbf{v}_1+\mathbf{v}_2)\pm\tfrac{1}{2}\,\mathbf{d}^{\perp},$$

其中 $\mathbf{d}^{\perp}$ 是把 $\mathbf{v}_1 - \mathbf{v}_2$ 旋轉一個直角所得的向量。（若 $\mathbf{v}_1 = \mathbf{v}_2$，圓便收縮成單獨一點 M，這時兩個相對速度都是零，兩個條件就沒有甚麼實質內容。）在圖中，直角 P_1UP_2 就是半圓上的圓周角，而 $UP_1 = UP_2$ 是因為 U 在垂直平分線上。

最後一部分

現在 $|\mathbf{v}_1| = |\mathbf{v}_2| = k$ 且 $\mathbf{v}_1 \cdot \mathbf{v}_2 = 0$，其中 $k > 0$（只有非零向量才談得上垂直）。於是 $\mathbf{v}_1$ 和 $\mathbf{v}_2$ 互相垂直而且非零，所以它們張成（span）整個平面，我們可以寫 $\mathbf{u} = \alpha\mathbf{v}_1 + \beta\mathbf{v}_2$。

利用 $\mathbf{v}_1 \cdot \mathbf{v}_2 = 0$ 和 $|\mathbf{v}_1|^2 = |\mathbf{v}_2|^2$，第二條方程給出

$$\mathbf{u} \cdot (\mathbf{v}_1 - \mathbf{v}_2) = \tfrac{1}{2}\left(|\mathbf{v}_1|^2 - |\mathbf{v}_2|^2\right) = 0, \qquad \text{所以} \qquad \alpha k^2 - \beta k^2 = 0,$$

因此 $\beta = \alpha$。至於第一條方程，把它展開：

$$|\mathbf{u}|^2 - \mathbf{u} \cdot (\mathbf{v}_1 + \mathbf{v}_2) + \tfrac{1}{4}|\mathbf{v}_1 + \mathbf{v}_2|^2 = \tfrac{1}{4}|\mathbf{v}_1 - \mathbf{v}_2|^2.$$

由於 $\mathbf{v}_1 \cdot \mathbf{v}_2 = 0$，$|\mathbf{v}_1 + \mathbf{v}_2|^2$ 和 $|\mathbf{v}_1 - \mathbf{v}_2|^2$ 都等於 $2k^2$，所以這兩項互相抵消，得 $|\mathbf{u}|^2 = \mathbf{u} \cdot (\mathbf{v}_1 + \mathbf{v}_2)$。代入 $\mathbf{u} = \alpha(\mathbf{v}_1 + \mathbf{v}_2)$，這就是

$$2\alpha^2 k^2 = 2\alpha k^2 \quad \Longrightarrow \quad \alpha(\alpha - 1) = 0.$$

所以 $\alpha = 0$，得 $\mathbf{u} = \mathbf{0}$；或 $\alpha = 1$，得 $\mathbf{u} = \mathbf{v}_1 + \mathbf{v}_2$。因此若 $\mathbf{u} \neq \mathbf{0}$，則 $\mathbf{u} = \mathbf{v}_1 + \mathbf{v}_2$，正如所求。

這在物理上也合情合理。若 C 靜止，相對速度就是 $\mathbf{v}_1$ 和 $\mathbf{v}_2$ 本身，它們本來就滿足兩個條件。若 $\mathbf{u} = \mathbf{v}_1 + \mathbf{v}_2$，相對速度是 $-\mathbf{v}_2$ 和 $-\mathbf{v}_1$，它們同樣長度相等而且互相垂直。

驗算

取 $\mathbf{v}_1 = (3,1)$，$\mathbf{v}_2 = (1,3)$，與圖中相同。公式給出 $\mathbf{u} = (2,2) \pm (1,1)$，即 $(3,3)$ 或 $(1,1)$。當 $\mathbf{u} = (3,3)$ 時，相對速度是 $(0,-2)$ 和 $(-2,0)$；當取 $(1,1)$ 時，相對速度是 $(2,0)$ 和 $(0,2)$。兩對向量都長度相等而且互相垂直。用電腦測試 10^4 對隨機的 $\mathbf{v}_1, \mathbf{v}_2$，結果確認了兩個恆等式和兩個解；另外對 10^3 對隨機的等長垂直向量，電腦求得的解恰好是 $\mathbf{0}$ 和 $\mathbf{v}_1 + \mathbf{v}_2$。

常見錯誤

- 把 A 相對於 C 的速度寫成 $\mathbf{u} - \mathbf{v}_1$。正負號不影響本題的兩個條件，但在大部分相對速度的題目中卻至關重要。
- 忘記說明為何恰好有兩個速度。這個數目取決於我們是在平面上討論（在三維空間中，所有解構成一整個圓），也取決於半徑不為零。

STEP 之外

由 $\mathbf{u}$ 轉到 $\mathbf{w} = \mathbf{u} - \frac{1}{2}(\mathbf{v}_1 + \mathbf{v}_2)$，其實是轉換參考系（reference frame）；而所有速度都平移同一個向量時相對速度不變，這個事實就是伽利略不變性（Galilean invariance），它是經典力學（classical mechanics）的起點。這裏所用的「速度圖」，即把速度畫成點的做法，也能把碰撞和追逐問題化為幾何問題。Kleppner 與 Kolenkow 合著的 *An Introduction to Mechanics*（力學入門）對相對運動和參考系有仔細的闡述。

詳細解答

STEP III 2001，第 11 題

力學 • 轉動能量守恆

> **STEP III 2001，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/01-s3-q11

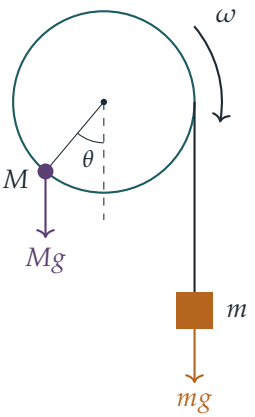
關鍵思路

整個系統都是光滑的（輪軸無摩擦，繩子是輕繩），所以機械能守恆（conservation of mechanical energy），只需一條能量方程便可把 ω 表示成 θ 的函數，完全不必考慮張力或其他力。圓柱體恰好在 ω^2 降至零時停下。於是第二部分變成關於單一個 θ 的函數的問題：懸掛質量所失去的勢能（potential energy），必須始終大於質點所增加的勢能。整理之後，這就是比較 m/M 與 $(1-\cos\theta)/\theta$ 的最大值，而求這個最大值並不困難。

解答

角速度

我們把 m 和 M 視為質點（particle），繩子視為不可伸長的輕繩（light and inextensible），且不在圓柱體上打滑，輪軸則視為光滑。設圓柱體沿着令 m 下降的方向轉過角度 θ，並設 $\omega = \dot{\theta}$。



當圓柱體轉過 θ 時，有長度為 $a\theta$ 的繩子鬆開，所以 m 下降了 $a\theta$，並以速率 $a\omega$ 運動。質點 M 位於圓柱體的邊緣，所以它的速率也是 $a\omega$；它起初位於軸線下方深度 a 處，現在位於深度 $a\cos\theta$ 處，所以它上升了 $a(1-\cos\theta)$。

動能（kinetic energy）為

$$\tfrac{1}{2}I\omega^2+\tfrac{1}{2}M(a\omega)^2+\tfrac{1}{2}m(a\omega)^2=\tfrac{1}{2}\big(I+(M+m)a^2\big)\omega^2.$$

光滑的輪軸不作功，而繩子既輕且不可伸長，傳遞能量時並無損耗。所以動能的增加等於勢能的淨減少：

$$\tfrac{1}{2}\big(I+(M+m)a^2\big)\omega^2=mga\theta-Mga(1-\cos\theta),$$

由此得

$$\boxed{\omega^2=\frac{2ga\big(m\theta-M(1-\cos\theta)\big)}{I+(M+m)a^2}}\qquad\text{及}\qquad\omega=\sqrt{\frac{2ga\big(m\theta-M(1-\cos\theta)\big)}{I+(M+m)a^2}}.$$

建模上有一點值得驗證：繩子必須保持拉緊。把能量方程對時間求導，得 $\big(I+(M+m)a^2\big)\ddot\theta=ga(m-M\sin\theta)$，而張力（tension）為 $T=m(g-a\ddot\theta)$。現在 $a\ddot\theta<g$ 等價於 $a^2(m-M\sin\theta)<I+(M+m)a^2$，即 $-Ma^2\sin\theta<I+Ma^2$，而這恆成立。所以在整個運動過程中 $T>0$，能量方程對所有 θ 都有效。

永不停下

開始時，$\ddot\theta=gam/(I+(M+m)a^2)>0$，所以圓柱體確實沿我們所選的方向開始轉動。由於 ω 連續，它只有經過零才能改變正負號，所以只要對每個 $\theta>0$ 都有 $\omega^2>0$，圓柱體便會永遠轉動下去。設 $k=m/M$，由 ω^2 的表達式可見，這種情況恰好在以下條件成立時出現：

$$k\theta>1-\cos\theta\quad\text{對所有}\theta>0,\qquad\text{即}\qquad k>h(\theta)=\frac{1-\cos\theta}{\theta}\quad\text{對所有}\theta>0.$$

所以我們需要求出 h 在 $\theta > 0$ 上的最大值。

駐點（stationary points）。我們有

$$h'(\theta) = \frac{\theta\sin\theta - (1-\cos\theta)}{\theta^2}.$$

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

$$h'(\theta) = \frac{2\sin\frac{1}{2}\theta\left(\theta\cos\frac{1}{2}\theta - \sin\frac{1}{2}\theta\right)}{\theta^2}.$$

在 $0 < \theta < \pi$ 上，$\sin\frac{1}{2}\theta$ 和 $\cos\frac{1}{2}\theta$ 皆為正，所以 $h'(\theta)$ 的正負號與 $\theta - \tan\frac{1}{2}\theta$ 相同。設 $u(\theta) = \tan\frac{1}{2}\theta - \theta$。那麼 $u(0) = 0$，$u'(\theta) = \frac{1}{2}\sec^2\frac{1}{2}\theta - 1$ 在 0 附近為負，而 $u''(\theta) = \frac{1}{2}\sec^2\frac{1}{2}\theta\tan\frac{1}{2}\theta > 0$，所以 u 在 $[0,\pi)$ 上是凸函數（convex function）。它先降至零以下，然後當 $\theta \to \pi$ 時趨向 $+\infty$。凸函數最多只能穿過零兩次，所以 u 在 $(0,\pi)$ 內恰好有一個零點，那就是 α。因此 h 在 $(0,\alpha)$ 上遞增，在 (α,π) 上遞減。

? **想一想** h 在最大值處的值是多少？對於更大的 θ 又如何？

在 $\theta = \alpha$ 處，h' 的分子為零，所以 $1-\cos\alpha = \alpha\sin\alpha$，因此

$$h(\alpha) = \frac{1-\cos\alpha}{\alpha} = \sin\alpha.$$

對於 $\pi \le \theta \le 2\pi$，我們有 $\sin\frac{1}{2}\theta \ge 0$ 和 $\cos\frac{1}{2}\theta \le 0$，所以 $\theta\cos\frac{1}{2}\theta - \sin\frac{1}{2}\theta \le 0$，從而 $h' \le 0$：h 一直遞減至 2π，所以在 $(\alpha, 2\pi]$ 上 $h < \sin\alpha$。對於 $\theta > 2\pi$，$h(\theta) \le 2/\theta < 1/\pi$。最後，$u(\frac{3}{4}\pi) = \tan\frac{3}{8}\pi - \frac{3}{4}\pi \approx 2.414 - 2.356 > 0$ 而 $u(\frac{2}{3}\pi) = \sqrt{3} - \frac{2}{3}\pi < 0$，所以 $\frac{2}{3}\pi < \alpha < \frac{3}{4}\pi$，由此得 $\sin\alpha > \sin\frac{3}{4}\pi = \frac{1}{\sqrt{2}} > \frac{1}{\pi}$。所以 $\sin\alpha$ 是 h 在 $\theta > 0$ 上的最大值，而且只在 $\theta = \alpha$ 處取得。

結論。若 $m/M > \sin\alpha$，則對每個 $\theta > 0$，

$$m\theta - M(1-\cos\theta) = M\theta\big(k - h(\theta)\big) \ge M\theta(k - \sin\alpha) > 0,$$

所以對所有 $\theta > 0$ 都有 $\omega^2 > 0$，圓柱體會一直轉動而不會停下，這正是所要證明的。

（反過來說，若 $m/M < \sin\alpha$，則在 $\theta = \alpha$ 處 $\omega^2 < 0$，所以 ω 在此之前已降至零，圓柱體會停下並往回擺動。在臨界情況 $m/M = \sin\alpha$ 下，系統會緩緩趨近平衡位置 $\theta = \alpha$，卻永遠不會真正到達。）

驗算

- 當 $M = 0$ 時，公式給出 $\omega^2 = 2gma\theta/(I + ma^2)$，這是質量從滑輪上鬆開繩子下降的標準結果：$\omega^2$ 線性增長，與恆定角加速度的情況一致。

- 用二分法（bisection）解 $\alpha = \tan\frac{1}{2}\alpha$，得 $\alpha \approx 2.3311$（約 133.6°）和 $\sin\alpha \approx 0.7246$；在 $0 < \theta < 100$ 上對 $(1-\cos\theta)/\theta$ 作數值搜尋，得出相同的最大值 0.7246。
- 對隨機選取的 I, a, M, m 以數值方法逐步求解運動方程，結果與 ω^2 的公式吻合；而且當 $m/M = 0.735$ 時圓柱體永不停下，當 $m/M = 0.715$ 時則會停下。

常見錯誤

- 遺漏 M 或 m 的動能：兩者都以速率 $a\omega$ 運動，所以有效轉動慣量（moment of inertia）是 $I + (M+m)a^2$。
- 弄錯 M 的高度：它上升了 $a(1-\cos\theta)$，而不是 $a\sin\theta$。
- 只證明 h 在 α 處有局部極大值，卻沒有檢查對所有更大的 θ，h 的值都較小。

STEP 之外

這是在勢場中的單自由度（one degree of freedom）運動：$\frac{1}{2}J\omega^2 + V(\theta) = 0$，其中 $V(\theta) = Mga(1-\cos\theta) - mga\theta$，是一個傾斜的「洗衣板」勢（washboard potential）。從靜止開始，運動會一去不返，當且僅當對每個 $\theta > 0$，V 都保持低於其初始值；同樣的圖像也可描述物理學中的受驅擺（driven pendulum）和約瑟夫森結（Josephson junction）。Kleppner 與 Kolenkow 所著的 *An Introduction to Mechanics*（力學入門）有討論這類能量圖。

詳細解答

STEP III 2001，第 12 題

概率與統計 • 以指示變量求期望值

STEP III 2001，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/01-s3-q12

關鍵思路

在第 (i) 部分，不要嘗試求出黑球數目的分佈：把這個數目拆開，每個黑球對應一個指示變數（indicator），然後運用期望值的線性性質（linearity of expectation）。不放回地抽球直至袋空，其實就是把所有球隨機排成一列，所以每個指示變數都歸結為一個對稱性論證（symmetry argument）。在第 (ii) 部分，兩次抽到白球之間袋中的球並無改變，所以等候每個白球出現的次數服從幾何分佈（geometric distribution）；而題目中的「hence」，則來自這個觀察：取走一個白球之後，我們回到同樣的處境，只是白球少了一個。

解答

第 (i) 部分

設想一直抽下去，直至袋空為止。由於每次都是隨機抽取，而且抽出的球不放回，這 $b+w$ 個球被抽出的次序，就是它們的一個均勻隨機排列（uniformly random ordering）。到最後一個白球被取走時已抽出的黑球，正好就是在這個排列中位於最後一個白球之前的黑球。

把黑球編號為 $1, 2, \ldots, b$；若黑球 j 在最後一個白球之前被抽出，令 $I_j = 1$，否則令 $I_j = 0$。設 N 為最後一個白球被取走時已抽出的黑球數目，則

$$N = I_1 + I_2 + \cdots + I_b, \qquad \text{所以} \qquad \mathrm{E}(N) = \sum_{j=1}^{b} \mathrm{P}(I_j = 1)$$

這是根據期望值的線性性質（它並不需要各個 I_j 之間互相獨立）。

想一想 黑球 j 排在所有白球之後的概率是多少？

只看由黑球 j 和 w 個白球組成的這 $w+1$ 個球。在所有球的均勻隨機排列中，這 $w+1$ 個球的相對次序同樣是均勻隨機的，所以黑球 j 處於它們當中 $w+1$ 個位置的任何一個，機會均等。它排在這組球最後的概率是 $\frac{1}{w+1}$，而這正是事件 $I_j = 0$。因此

$$\mathrm{P}(I_j = 1) = 1 - \frac{1}{w+1} = \frac{w}{w+1},$$

從而

$$\boxed{\mathrm{E}(N) = \frac{bw}{w+1}.}$$

第 (ii) 部分

現在黑球會放回袋中，所以在抽到白球之前，袋中始終有 b 個黑球和 w 個白球。因此每次抽到黑球的概率是 $p = \frac{b}{b+w}$，抽到白球的概率是 $q = \frac{w}{b+w}$，而且與之前各次抽取獨立，因為每次袋中的球都相同。

設 X 為抽到第一個白球之前抽到黑球的次數。那麼 $X = n$ 表示先抽到 n 次黑球，然後抽到一個白球，所以

$$\mathrm{P}(X = n) = p^n q, \qquad n = 0, 1, 2, \ldots$$

若 $b = 0$，則 $X = 0$，結果顯然成立，故取 $b \geq 1$；這時 $0 < p < 1$，利用當 $|x| < 1$ 時 $\sum_{n\geq 1} nx^{n-1} = \frac{1}{(1-x)^2}$（把等比級數（geometric series）求導即得），

$$\mathrm{E}(X) = \sum_{n=0}^{\infty} np^n q = pq \sum_{n=1}^{\infty} np^{n-1} = \frac{pq}{(1-p)^2} = \frac{p}{q} = \frac{b/(b+w)}{w/(b+w)} = \frac{b}{w},$$

因為 $1 - p = q$。這就是所要證明的結果。

至於最後一個白球，把抽到黑球的總次數寫成 $X_w + X_{w-1} + \cdots + X_1$，其中 X_k 是袋中恰好有 k 個白球期間抽到黑球的次數。當袋中尚餘 k 個白球時，袋中有 b 個黑球和 k 個白球，所以這個階段正是把 w 換成 k 的「第一個白球」問題，而剛才證明的結果給出 $\mathrm{E}(X_k) = \frac{b}{k}$。根據期望值的線性性質，最後一個白球被取走時抽到黑球的期望次數是

$$\boxed{\sum_{k=1}^{w} \frac{b}{k} = b\left(1 + \frac{1}{2} + \frac{1}{3} + \cdots + \frac{1}{w}\right).}$$

驗算

當 $w = 1$ 時，兩部分的結果都很容易直接看出。在 (i) 中，每一個黑球排在白球之前的概率都是 $\frac{1}{2}$，得出 $\frac{b}{2}$，與 $\frac{bw}{w+1}$ 相符。在 (ii) 中，只有一個階段，而總和就是 $\frac{b}{1} = b$。在 (i) 中當 $w \to \infty$ 時，幾乎每個黑球都在最後一個白球之前被抽出，而確實有 $\frac{bw}{w+1} \to b$。

我們對所有 $b \le 4$、$w \le 4$ 的情況逐一列出每個排列來驗算 (i)；當 $b = 3$、$w = 4$ 時，結果恰好是 $\frac{12}{5}$。至於 (ii)，以 $b = 3$、$w = 4$ 進行 2×10^5 次模擬，所得平均值約為 6.245，而 $3\left(1 + \frac{1}{2} + \frac{1}{3} + \frac{1}{4}\right) = \frac{25}{4} = 6.25$。

常見錯誤

- 在 (i) 中，假設各個黑球是否被抽出是互相獨立的。期望值的線性性質並不需要獨立性，這正是指示變數如此好用的原因。
- 在 (ii) 中，採用了把抽到白球那一次也計算在內的幾何分佈，結果得出 $\frac{1}{q} = \frac{b+w}{w}$，比答案多了一。
- 在最後一部分的每個階段都以 $b + w$ 作分母：袋中白球的數目每過一個階段就減少一個。

STEP 之外

第二部分其實是喬裝了的優惠券收集問題（coupon collector's problem）：把各個階段的幾何等候時間加起來，便得到一個調和和（harmonic sum），其增長速度與 $\ln w$ 相若。像第 (i) 部分那樣把一個計數寫成指示變數之和，是概率論中最強而有力的技巧之一；日後計算隨機置換（random permutation）中不動點（fixed point）、紀錄（record）和循環（cycle）的期望數目時，你會再次遇上它。Blitzstein 與 Hwang 合著的 *Introduction to Probability*（概率論入門）中，有許多關於這兩個想法的例子。

詳細解答

STEP III 2001，第 13 題

概率與統計 • 硬幣序列遊戲

> **STEP III 2001，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/01-s3-q13

本題在 *Beyond STEP* 第 12.8 節 *Penney* 遊戲中有完整的引導式講解：https://beyondstep.keithmcnulty.org/zh-hant/ch12/penney.html

關鍵思路

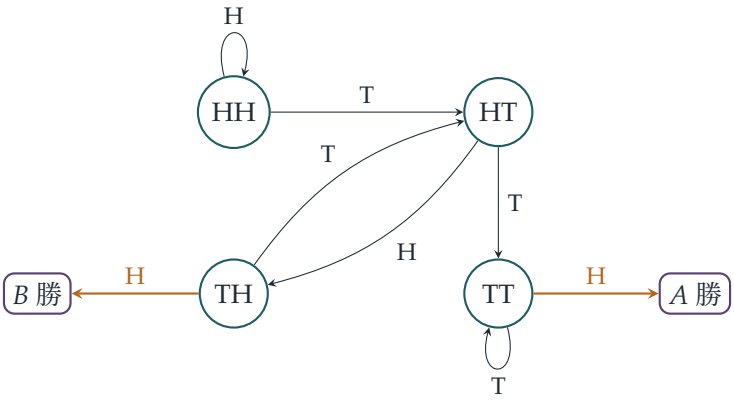
兩個目標序列的長度都是三，所以沒有人能在第三次擲幣之前獲勝；而從那時起，對往後有影響的就只有最後兩次擲幣的結果。這樣便只有四個「狀態」（state）（HH、HT、TH、TT），而以下一次擲幣的結果作條件（conditioning），就可把這個遊戲化為一個關於各狀態下獲勝概率的小型線性方程組。第 (ii) 和第 (iii) 部分也可以用同樣的方法處理，不過在四場遊戲之中，有兩場用簡短的直接論證會更快。

解答

在整個解答中，各次擲幣互相獨立（independent），每次出現 H 或 T 的概率均為 $\frac{1}{2}$。只要已經擲了最少兩次而仍未有人獲勝，下一次擲幣連同最後兩次的結果，就決定了此刻是否有人獲勝；若沒有，也決定了新的最後兩次結果。更早的擲幣結果並不相干，因為每個目標序列只有三個字母長。所以，從某一刻起獲勝的概率只取決於最後兩次擲幣的結果，而且與把這兩次當作一場新遊戲的頭兩次擲幣時的概率相同。

第 (i) 部分

設 p_{XY} 為已知頭兩次擲幣結果是 XY 時，A（TTH）擊敗 B（THH）的概率。下圖顯示下一次擲幣如何使我們在各狀態之間轉移。



爲何 $p_{\mathrm{TT}} = 1$。假設最後兩次擲幣的結果是 TT。正面終會出現的概率為 1（連續出現 n 次反面的機會是 $2^{-n} \to 0$）。在正面出現之前，擲出的全是反面，所以 THH 不可能出現；而第一個正面之前最少有兩個反面，所以它完成了 TTH。因此 A 必定獲勝：$p_{\mathrm{TT}} = 1$。

爲何 $p_{\mathrm{HT}} = \frac{1}{2}p_{\mathrm{TH}} + \frac{1}{2}p_{\mathrm{TT}}$。在 HT 之後，HTH 和 HTT 都不是目標序列，所以沒有人會在第三次擲幣時獲勝。若第三次是 H（概率為 $\frac{1}{2}$），最後兩次的結果便是 TH；若是 T，則是 TT。以第三次擲幣的結果作條件，並運用上文「只有最後兩次擲幣才有影響」這一點，

$$p_{\mathrm{HT}} = \tfrac{1}{2}p_{\mathrm{TH}} + \tfrac{1}{2}p_{\mathrm{TT}}.$$

其餘兩個狀態。在 TH 之後，擲出正面便得 THH，B 獲勝；擲出反面則變成 HT。在 HH 之後，擲出正面仍是 HH，擲出反面則變成 HT（而 HHH、HHT 在這裏都不是目標序列）。所以

$$p_{\mathrm{TH}} = \tfrac{1}{2}\cdot 0 + \tfrac{1}{2}p_{\mathrm{HT}}, \qquad p_{\mathrm{HH}} = \tfrac{1}{2}p_{\mathrm{HH}} + \tfrac{1}{2}p_{\mathrm{HT}}.$$

由第二條方程得 $p_{\mathrm{HH}} = p_{\mathrm{HT}}$。把 $p_{\mathrm{TH}} = \frac{1}{2}p_{\mathrm{HT}}$ 和 $p_{\mathrm{TT}} = 1$ 代入 p_{HT} 的方程，

$$p_{\mathrm{HT}} = \tfrac{1}{4}p_{\mathrm{HT}} + \tfrac{1}{2} \quad\Longrightarrow\quad \tfrac{3}{4}p_{\mathrm{HT}} = \tfrac{1}{2} \quad\Longrightarrow\quad p_{\mathrm{HT}} = \tfrac{2}{3}.$$

因此 $p_{\mathrm{HH}} = p_{\mathrm{HT}} = \frac{2}{3}$，而 $p_{\mathrm{TH}} = \frac{1}{2}\cdot\frac{2}{3} = \frac{1}{3}$，正如題目所求。

想一想 假如 TH 其實是在第 5 和第 6 次擲幣時出現的，為何可以把遊戲從 TH「重新開始」，當作 TH 就是頭兩次擲幣的結果？

因為遊戲往後的發展只取決於往後的擲幣結果和最後兩次擲幣的結果：任何尚未完成的 TTH 或 THH，最多只能用到已擲出的最後兩次結果。更早的歷史對兩個序列都沒有幫助，而往後的擲幣結果亦與它獨立。

整體概率。沒有人能在頭兩次擲幣內獲勝，而四種可能的組合各有 $\frac{1}{4}$ 的概率。根據全概率定理（law of total probability），

$$\mathrm{P}(A \text{ 勝}) = \tfrac{1}{4}\left(\tfrac{2}{3} + \tfrac{2}{3} + \tfrac{1}{3} + 1\right) = \tfrac{1}{4}\cdot\tfrac{8}{3} = \tfrac{2}{3}.$$

第 (ii) 部分

現在 B 的序列是 THH，C 的是 HHT。這裏用直接論證最為簡潔。

若頭兩次擲幣的結果是 HH，則 C 獲勝：其後出現的第一個反面（它出現的概率為 1）之前是 HH，於是完成了 HHT；而在它之前每次擲出的都是正面，所以 THH 不可能出現過。

否則頭兩次擲幣之中有一個反面。HH 這個模式終會出現的概率為 1。考慮它第一次出現的位置。它並非由第 1 次擲幣開始，所以緊接在它之前的一次擲幣是存在的，而且結果不是 H（否則便有一個更早的 HH），所以是 T。於是第一個 HH 完成了一個 THH。同時，HHT 不可能在此之前出現，因為它需要一個更早的 HH。所以 B 獲勝。

因此，除非頭兩次擲幣的結果是 HH，否則 B 獲勝：

$$\mathrm{P}(B \text{ 勝}) = 1 - \tfrac{1}{4} = \tfrac{3}{4}.$$

第 (iii) 部分

C（HHT）對 D（HTT）。設 q_{XY} 為最後兩次擲幣結果是 XY 時 C 獲勝的概率。由 HH 出發，C 必定獲勝，理由與 (ii) 相同：第一個反面完成了 HHT，而 HTT 需要在 HT 之後再有一個反面，這不可能先出現。所以 $q_{\mathrm{HH}} = 1$。在其餘各狀態下以下一次擲幣的結果作條件：

$$\begin{aligned}
&\mathrm{HT}\text{：}\quad \text{T 得 HTT（}D\text{ 勝），H 得 TH：} && q_{\mathrm{HT}} = \tfrac{1}{2}q_{\mathrm{TH}},\\
&\mathrm{TH}\text{：}\quad \text{H 得 HH，T 得 HT：} && q_{\mathrm{TH}} = \tfrac{1}{2} + \tfrac{1}{2}q_{\mathrm{HT}},\\
&\mathrm{TT}\text{：}\quad \text{H 得 TH，T 得 TT：} && q_{\mathrm{TT}} = \tfrac{1}{2}q_{\mathrm{TH}} + \tfrac{1}{2}q_{\mathrm{TT}}.
\end{aligned}$$

由最後一條得 $q_{\mathrm{TT}} = q_{\mathrm{TH}}$。把第一條代入第二條，得 $q_{\mathrm{TH}} = \frac{1}{2} + \frac{1}{4}q_{\mathrm{TH}}$，所以 $q_{\mathrm{TH}} = \frac{2}{3}$、$q_{\mathrm{TT}} = \frac{2}{3}$ 及 $q_{\mathrm{HT}} = \frac{1}{3}$。因此

$$\mathrm{P}(C \text{ 勝}) = \tfrac{1}{4}\left(1 + \tfrac{1}{3} + \tfrac{2}{3} + \tfrac{2}{3}\right) = \tfrac{1}{4} \cdot \tfrac{8}{3} = \tfrac{2}{3} > \tfrac{1}{2},$$

所以 C 比 D 更有可能獲勝。

D（HTT）對 A（TTH）。這與第 (ii) 部分互為鏡像，只是 H 和 T 對調了。若頭兩次擲幣的結果是 TT，則 A 獲勝：第一個正面完成了 TTH，而 HTT 需要一個正面之後接兩個反面，這不可能在那第一個正面之前發生。否則頭兩次擲幣之中有一個正面。TT 這個模式終會出現的概率為 1；它第一次出現時並非由第 1 次擲幣開始，而它之前的一次擲幣不可能是 T（否則便有一個更早的 TT），所以是 H，於是第一個 TT 完成了 HTT。TTH 不可能在更早時出現，因為它需要一個更早的 TT。所以 D 獲勝，而

$$\mathrm{P}(D \text{ 勝}) = 1 - \tfrac{1}{4} = \tfrac{3}{4} > \tfrac{1}{2}.$$

所以 D 比 A 更有可能獲勝。

想一想 *A* 勝 *B*，*B* 勝 *C*，*C* 勝 *D*，而 *D* 又勝 *A*。這怎麼可能？

「較有可能擊敗」並不是一個傳遞關係（transitive relation），所以並無矛盾：四位玩家形成一個循環，*A*、*B*、*C*、*D* 在各自的對局中獲勝的概率分別是 $\frac{2}{3}$、$\frac{3}{4}$、$\frac{2}{3}$、$\frac{3}{4}$。後選的一方總可以挑一個能擊敗先選者的序列。

驗算

- 在第 (i) 部分，代回驗證：$\frac{1}{2}\cdot\frac{1}{3}+\frac{1}{2}\cdot 1=\frac{2}{3}=p_{\mathrm{HT}}$，以及 $\frac{1}{2}\cdot\frac{2}{3}=\frac{1}{3}=p_{\mathrm{TH}}$。
- 狀態法同樣適用於第 (ii) 部分和 *D*–*A* 這場遊戲：在 *B* 對 *C* 的遊戲中，由 HH、HT、TH、TT 出發時 *B* 獲勝的概率分別是 $0, 1, 1, 1$，平均值為 $\frac{3}{4}$。
- 為每組對局模擬 200 000 場遊戲，所得的獲勝頻率為 0.664、0.752、0.666 和 0.749，而精確值是 $\frac{2}{3}$、$\frac{3}{4}$、$\frac{2}{3}$、$\frac{3}{4}$。

常見錯誤

- 忘記遊戲可以在第三次擲幣時結束，例如寫成 $p_{\mathrm{TH}}=\frac{1}{2}p_{\mathrm{HH}}+\frac{1}{2}p_{\mathrm{HT}}$，而沒有留意到 TH 之後擲出 H 就是 *B* 即時獲勝。
- 把四個狀態的概率取平均，卻沒有說明頭兩次擲幣的每種組合概率均為 $\frac{1}{4}$，以及沒有人能在第三次擲幣之前獲勝。

STEP 之外

這就是 Penney 遊戲（Penney's game），是非傳遞概率（non-transitive probabilities）的著名例子。狀態法其實就是吸收馬可夫鏈（absorbing Markov chains）的理論，當中的吸收概率所滿足的，正是這裏所解的那類線性方程。任意兩個序列之間的勝負比（odds）甚至有一條閉合公式，即康威（Conway）的「領先數」（leading numbers）演算法，可以用鞅（martingales）來證明。關於馬可夫鏈和首步分析（first-step analysis），可參閱 Grimmett 與 Welsh 的 *Probability: An Introduction*，或 Blitzstein 與 Hwang 的 *Introduction to Probability*。

詳細解答

STEP III 2001，第 14 題

概率與統計 • 方差的無偏估計量

STEP III 2001，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/01-s3-q14

關鍵思路

整道題的關鍵只有一點：「樣本的方差」所用的除數是 $n = 2$，因此會低估總體方差（population variance），而把它乘以二便可修正過來。只有兩個值時，樣本平均值（sample mean）就是兩者的中點，所以每個值與它相距都是兩值差距的一半。於是 V 只是單一個差 $X_1 - X_2$ 的函數，而最後要求的概率就變成正方形 $[0, a]^2$ 內的一塊面積。

解答

X 的方差

X 的密度函數（density）在 $[0, a]$ 上是 $f(x) = \frac{1}{a}$。由對稱性可知 $\mathrm{E}(X) = \frac{1}{2}a$，而

$$\mathrm{E}(X^2) = \int_0^a \frac{x^2}{a}\,\mathrm{d}x = \frac{a^2}{3}.$$

所以

$$\mathrm{Var}(X) = \mathrm{E}(X^2) - \big(\mathrm{E}(X)\big)^2 = \frac{a^2}{3} - \frac{a^2}{4} = \frac{a^2}{12},$$

正如題目所求。

樣本方差

對於樣本 X_1, X_2，樣本平均值是 $\bar{X} = \frac{1}{2}(X_1 + X_2)$，而樣本的方差就是各值與它的偏差平方的平均值：

$$V = \tfrac{1}{2}\big[(X_1 - \bar{X})^2 + (X_2 - \bar{X})^2\big].$$

由於 $X_1 - \bar{X} = \frac{1}{2}(X_1 - X_2)$ 及 $X_2 - \bar{X} = -\frac{1}{2}(X_1 - X_2)$，兩個偏差平方都等於 $\frac{1}{4}(X_1 - X_2)^2$。因此

$$V = \tfrac{1}{2} \cdot 2 \cdot \tfrac{1}{4}(X_1 - X_2)^2 = \tfrac{1}{4}(X_1 - X_2)^2,$$

這正是題目給出的形式。

> ? **想一想** 怎樣不用二重積分（double integral）而求出 $\mathrm{E}\big((X_1 - X_2)^2\big)$？

由於 X_1 和 X_2 的分佈相同，$\mathrm{E}(X_1 - X_2) = 0$，所以 $\mathrm{E}\big((X_1 - X_2)^2\big) = \mathrm{Var}(X_1 - X_2)$。因為 X_1 和 X_2 獨立（independent），它們的方差可以相加（負號並無影響，因為 $\mathrm{Var}(-X_2) = \mathrm{Var}(X_2)$）：

$$\mathrm{Var}(X_1 - X_2) = \mathrm{Var}(X_1) + \mathrm{Var}(X_2) = \frac{a^2}{12} + \frac{a^2}{12} = \frac{a^2}{6}.$$

因此

$$\mathrm{E}(2V) = \tfrac{1}{2}\,\mathrm{E}\big((X_1 - X_2)^2\big) = \tfrac{1}{2} \cdot \frac{a^2}{6} = \frac{a^2}{12} = \mathrm{Var}(X).$$

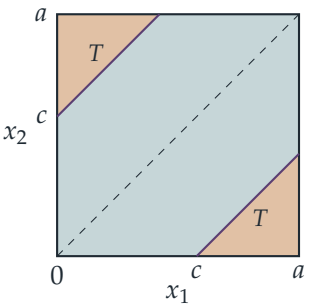
無論 a 取甚麼值，$2V$ 的期望值（expected value）都等於它所估計的量，所以 $2V$ 是 X 的方差的無偏估計量（unbiased estimator）。（相比之下，$\mathrm{E}(V) = \frac{a^2}{24}$，只有真正方差的一半。）

概率

由於 $V \geq 0$ 而 $\frac{1}{12}a^2 > 0$，我們可以取平方根：

$$V < \frac{a^2}{12} \iff (X_1 - X_2)^2 < \frac{a^2}{3} \iff |X_1 - X_2| < \frac{a}{\sqrt{3}}.$$

因為 X_1 和 X_2 獨立，而且各自服從均勻分佈（uniform distribution），所以數對 (X_1, X_2) 均勻分佈於正方形 $[0, a]^2$ 上。因此所求概率就是正方形面積中落在帶狀區域 $|x_1 - x_2| < c$ 內的比例，其中 $c = a/\sqrt{3} < a$。



餘下的部分由角落兩個着色的三角形 T 組成，每個都是等腰直角三角形，兩條直角邊的

長度均為 $a-c$。它們的總面積是 $2\cdot\frac{1}{2}(a-c)^2=(a-c)^2$，所以

$$\mathrm{P}\left(V<\frac{a^2}{12}\right)=1-\frac{(a-c)^2}{a^2}=1-\left(1-\frac{1}{\sqrt{3}}\right)^2=1-\left(\frac{4}{3}-\frac{2}{\sqrt{3}}\right)=\frac{2}{\sqrt{3}}-\frac{1}{3}.$$

也就是說，概率的準確值是

$$\mathrm{P}\left(V<\frac{a^2}{12}\right)=\frac{2\sqrt{3}-1}{3}.$$

取 $\sqrt{3}\approx 1.732$，這個值是 $\frac{2.464}{3}\approx 0.821$，所以準確至一位有效數字，概率是 0.8。

驗算

- 答案與 a 無關，這是理所當然的：把 X 按某個倍數縮放，V 和 $\frac{1}{12}a^2$ 都會按該倍數的平方縮放。
- 取 $a=3$ 模擬 4×10^5 個數對，得出 $2V$ 的平均值為 0.751（目標值 0.75），而 $V<\frac{1}{12}a^2$ 所佔的比例為 0.821，與 $\frac{1}{3}(2\sqrt{3}-1)=0.8214$ 吻合。
- 合理性檢查：概率落在 $(0,1)$ 之內；若把 $\frac{1}{12}a^2$ 換成 $\frac{1}{4}a^2$（V 可取的最大值），則 $c=a$，而公式 $1-(1-c/a)^2$ 給出 1。

常見錯誤

- 計算「樣本的方差」時用了除數 $n-1$：這樣會得到 $\frac{1}{2}(X_1-X_2)^2$，而不是題目所述的 $\frac{1}{4}(X_1-X_2)^2$。題目要的是除數 n。
- 忘記了絕對值，結果只計算了兩個三角形的其中一個（這樣得出的概率會小於 $\frac{1}{2}$）。
- 寫成 $\mathrm{Var}(X_1-X_2)=\mathrm{Var}(X_1)-\mathrm{Var}(X_2)=0$。

STEP 之外

這裏的結果是一條一般定理在 $n=2$ 時的情況：從方差為 σ^2 的分佈中抽取任何大小為 n 的樣本，$\sum(X_i-\bar{X})^2$ 的平均值都是 $(n-1)\sigma^2$。這就是為甚麼用除數 $n-1$（貝塞爾校正，Bessel's correction）會得到無偏估計量。大學一年級的統計學會進一步探討：在各個無偏估計量之中，哪一個的方差最小。Grimmett 與 Welsh 的 *Probability: An Introduction*，以及 Blitzstein 與 Hwang 的 *Introduction to Probability*，都以本文所用的方式講解和的期望值與方差。

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

2001 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。