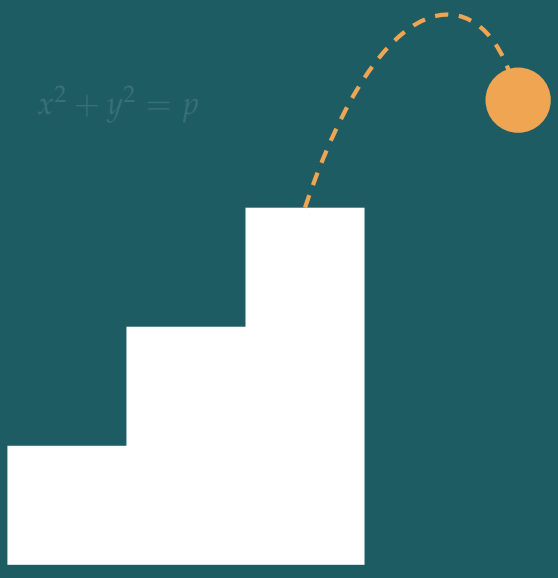
STEP 備試 · 大學數學

# Beyond STEP

STEP I, II 及 III

2006 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

2006 年詳細解答

Keith McNulty

**Beyond STEP：以歷屆 STEP 試題為門，走進大學數學**

**2006 年詳細解答**

2006 年 STEP 試卷（STEP I, II 及 III）全部 37 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 平方差 純數學 . . . 3

第 2 題 面積與最優化 純數學 . . . 6

第 3 題 多項式的根 純數學 . . . 9

第 4 題 正多邊形的面積 純數學 . . . 13

第 5 題 代換積分法 純數學 . . . 17

第 6 題 佩爾方程的變換 純數學 . . . 20

第 7 題 含絕對值符號的積分 純數學 . . . 24

第 8 題 三維向量 純數學 . . . 28

第 9 題 連接質點 力學 . . . 31

第 10 題 拋體運動 力學 . . . 35

第 11 題 彈性碰撞 力學 . . . 39

第 12 題 條件概率 概率與統計 . . . 43

第 13 題 超幾何混合分佈 概率與統計 . . . 46

第 14 題 不放回抽樣 概率與統計 . . . 49


## STEP II


第 1 題 周期性遞推關係 純數學 . . . 52

第 3 題 根式與近似 純數學 . . . 56

第 4 題 代換積分法 純數學 . . . 59

第 5 題 下取整函數與面積 純數學 . . . 63

第 7 題 橢圓的切線 純數學 . . . 67

第 8 題 直線的向量幾何 純數學 . . . 71

第 9 題 摩擦力與傾倒 力學 . . . 75

第 10 題 碰撞與恢復係數 力學 . . . 79

第 14 題 連續隨機變量 概率與統計 . . . . . 83

**STEP III**

第 1 題 曲線描繪 純數學 . . . . . 87
第 2 題 利用對稱性求定積分 純數學 . . . . . 92
第 3 題 冪級數的係數 純數學 . . . . . 96
第 4 題 函數方程 純數學 . . . . . 100
第 5 題 複數與幾何 純數學 . . . . . 104
第 6 題 極坐標曲線與反射 純數學 . . . . . 108
第 7 題 一階微分方程 純數學 . . . . . 112
第 8 題 公理化微分法 純數學 . . . . . 116
第 9 題 能量守恆 力學 . . . . . 120
第 10 題 動量守恆 力學 . . . . . 124
第 11 題 連接系統中的碰撞 力學 . . . . . 128
第 12 題 態近似與期望值 概率與統計 . . . . . 133
第 13 題 幾何概率 概率與統計 . . . . . 136
第 14 題 方差與獨立性 概率與統計 . . . . . 140

詳細解答

# STEP I 2006，第 1 題

純數學 • 平方差

> **STEP I 2006，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/06-s1-q1

## 關鍵思路

若 $(n+m)^2 - 33\,127 = k^2$，則 $33\,127 = (n+m)^2 - k^2$ 是一個平方差（difference of two squares），因此可分解為 $(n+m-k)(n+m+k)$。這正是費馬（Fermat）用來做因式分解（factorisation）的古老技巧：由剛好大於 $\sqrt{N}$ 的整數開始，逐步向上試，直至 $a^2 - N$ 是一個平方數為止。把這個論證倒過來看，33 127 的每一個分解都給出一個 $m$ 的值，所以要數算 $m$ 有多少個值，歸根究底就是數算有多少個分解。

## 解答

**求 $n$**

由於 $180^2 = 32\,400$，33 127 的平方根略大於 180。我們求得

$$182^2 = 33\,124 < 33\,127 < 33\,489 = 183^2,$$

所以 $n = 182$。留意 $33\,127 = 182^2 + 3$，所以它本身並不是完全平方數（perfect square）。

**一個較小的 $m$ 值**

我們依次試 $m = 1, 2, \ldots$：

$$183^2 - 33\,127 = 33\,489 - 33\,127 = 362, \quad \text{不是平方數}(19^2 = 361),$$
$$184^2 - 33\,127 = 33\,856 - 33\,127 = 729 = 27^2.$$

所以 $m = 2$ 可行，此時 $(n+m)^2 - 33\,127 = 27^2$。

**分解** 33 127

移項後運用平方差公式，

$$33\,127 = 184^2 - 27^2 = (184-27)(184+27) = 157 \times 211,$$

所以我們可取 $p = 157$ 和 $q = 211$，兩者均大於 1。

**恰好有兩個 $m$ 值**

兩個因數都是質數（prime）。對於 157：由於 $13^2 = 169 > 157$，我們只需測試質數 $2, 3, 5, 7, 11$，而當中沒有一個能整除 157。對於 211：由於 $15^2 = 225 > 211$，我們測試 $2, 3, 5, 7, 11, 13$，而當中沒有一個能整除 211。所以 $33\,127 = 157 \times 211$ 就是質因數分解（prime factorisation），而把 33 127 寫成正整數之積 $uv$（其中 $u \le v$）的方法只有

$$1 \times 33\,127 \qquad \text{和} \qquad 157 \times 211.$$

現在假設 $(n+m)^2 - 33\,127 = k^2$，其中 $k \ge 0$。我們取 $a = n + m > 0$，這正是題目所設想的情況（由 $n$ 開始向上走）。於是

$$(a-k)(a+k) = 33\,127.$$

這裏 $a + k > 0$，所以 $a - k > 0$ 亦成立，而且 $a - k \le a + k$。因此 $(a-k,\, a+k)$ 必定是上面兩對 $(u, v)$ 的其中之一，於是

$$a = \frac{u+v}{2}, \qquad k = \frac{v-u}{2}.$$

由於 $u$ 和 $v$ 都是奇數，這兩個數都是整數。

**想一想** 為甚麼每一個分解恰好給出一個 $m$，又為甚麼不會有其他 $m$ 混進來？

這個對應是雙向的。給定 $a$ 和 $k$，數對 $(a-k, a+k)$ 就是一個分解；給定一個分解 $(u, v)$，上面的公式可唯一地還原出 $a$ 和 $k$。所以 $a$ 的值（因而 $m = a - 182$ 的值）與因數對一一對應（one-to-one correspondence），而因數對恰好有兩個。

- $(u, v) = (157, 211)$ 給出 $a = 184$，$k = 27$，所以 $m = 184 - 182 = 2$，即我們之前找到的值。
- $(u, v) = (1, 33\,127)$ 給出 $a = 16\,564$，$k = 16\,563$，所以

$$m = 16\,564 - 182 = 16\,382.$$

因此恰好有兩個 $m$ 值，即 $m = 2$ 和 $m = 16\,382$。驗算第二個值：$16\,564^2 - 16\,563^2 = 16\,564 + 16\,563 = 33\,127$。

（假如我們也容許 $n + m$ 為負數，則 $a$ 可以是 $-184$ 或 $-16\,564$，給出鏡像的值 $m = -366$ 和 $m = -16\,746$。題目的原意顯然是 $n + m$ 為正數。）

## 驗算

用電腦直接搜尋由 0 至 16 564 的每一個 $m$，結果發現只有當 $m = 2$ 和 $m = 16\,382$ 時，$(182 + m)^2 - 33\,127$ 才是完全平方數，而且 33 127 的因數恰好是 $1, 157, 211, 33\,127$。用手計算，$157 \times 211 = 157 \times 200 + 157 \times 11 = 31\,400 + 1\,727 = 33\,127$。

## 常見錯誤

- 忘記了平凡的分解 $1 \times 33\,127$，因而只找到一個 $m$ 值。
- 未經檢查就聲稱 157 和 211 是質數：「恰好有兩個值」這個論證正是建基於此。
- 沒有檢查 $u$ 和 $v$ 的奇偶性（parity）是否相同。這裏兩者都是奇數，所以 $a$ 和 $k$ 都是整數；但若換成一個不是 4 的倍數的偶數，就沒有任何分解可行。

## STEP 之外

這就是費馬因式分解法（Fermat's factorisation method），當 $N$ 有兩個接近 $\sqrt{N}$ 的因數時（正如本題），這方法很快見效。同一個想法，即改為尋找 $a^2 \equiv k^2 \pmod{N}$，而非要求 $a^2 - N = k^2$ 嚴格成立，正是二次篩法（quadratic sieve）的基礎，而二次篩法是最快的通用因數分解演算法之一。分解與平方差表示法之間的聯繫，正好讓你初嘗 Davenport 所著 *The Higher Arithmetic*（一本初等數論入門書）中的初等數論（elementary number theory）。

詳細解答

# STEP I 2006，第 2 題

純數學 • 面積與最優化

> **STEP I 2006，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/06-s1-q2

## 關鍵思路

用一個數來確定繫繩點的位置，即它與其中一個角的距離 $x$，然後跟着繩子看它怎樣繞過穀倉。這樣，吃草區域便分成一個半圓（semicircle）和四個四分之一圓（quarter-circle），它們的半徑都是 $x$ 的簡單線性函數。因此 $A$ 是 $x$ 在一個閉區間（closed interval）上的二次函數（quadratic），題目的兩個部分都歸結為求它的最大值和最小值。

## 解答

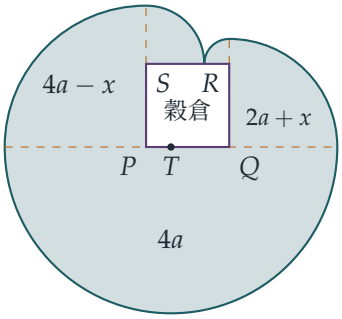
### 建立模型

把穀倉標記為 $PQRS$，山羊繫在邊 $PQ$ 上的點 $T$，其中 $PT = x$，$TQ = 2a - x$，$0 \le x \le 2a$。根據對稱性（symmetry），山羊繫在哪一條邊上並無分別。

若由 $T$ 出發、避開穀倉而到達某一點的最短路徑的長度不超過 $4a$，山羊便能到達該點。繩子拉緊時，這條路徑除了在穀倉的角上轉彎之外，都是直線。由 $T$ 起繞着穀倉看：

- $PQ$ **的前方**。直線 $PQ$ 另一側的每一點，都可以由 $T$ 直接看見，所以山羊可吃到一個半徑為 $4a$ 的半圓內的草。
- $PS$ **和** $QR$ **的旁邊**。繩子繞過 $P$ 之後，餘下的長度是 $4a - x$，它沿着邊 $PS$ 掃出一個以 $P$ 為圓心的四分之一圓。同樣，在 $Q$ 處餘下 $4a - (2a - x) = 2a + x$，於是沿着 $QR$ 得到一個以 $Q$ 為圓心的四分之一圓。
- **穀倉的後方**。由於 $x \le 2a$，我們有 $4a - x \ge 2a$，所以繩子可到達 $S$，而且還餘下 $(4a - x) - 2a = 2a - x$，這段繩子在邊 $SR$ 的後方掃出一個以 $S$ 為圓心的四分之一圓。同理，繩子到達 $R$ 時還餘下 $(2a + x) - 2a = x$，於是得到一個以 $R$ 為圓心、半

徑為 $x$ 的四分之一圓。




**想一想** 穀倉後方的兩塊區域會否重疊，以致我們把某些草地重複計算了兩次？

以 $S$ 和 $R$ 為圓心的兩個四分之一圓位於直線 $SR$ 的同一側。它們的半徑 $2a - x$ 和 $x$ 加起來等於 $2a = SR$，所以兩個圓剛好在 $SR$ 上與 $S$ 相距 $2a - x$ 的那一點相切。因此它們並不重疊，而兩者合起來，山羊可到達整條邊 $SR$。其他各塊區域也互不重疊：半圓位於直線 $PQ$ 的一側，四個四分之一圓則位於另一側，各自佔據穀倉旁邊或後方的一個區域。此外，穀倉後方的一點能夠到達，當且僅當它與 $S$ 或 $R$ 的距離不超過該處餘下的繩長，所以沒有遺漏任何地方。

### 面積

把半圓和四個四分之一圓加起來，

$$\begin{aligned} A &= \tfrac{1}{2}\pi(4a)^2 + \tfrac{1}{4}\pi\left[(4a-x)^2 + (2a+x)^2 + (2a-x)^2 + x^2\right] \\ &= 8\pi a^2 + \tfrac{1}{4}\pi\left(4x^2 - 8ax + 24a^2\right) \\ &= \pi\left(x^2 - 2ax + 14a^2\right) \\ &= \pi\left[(x-a)^2 + 13a^2\right]. \end{aligned}$$

這條公式對 $[0, 2a]$ 內的每一個 $x$ 都成立。在兩個端點（$x = 0$ 或 $x = 2a$，即山羊繫在角上），其中一個四分之一圓縮成一點，而半圓與半徑為 $4a$ 的四分之一圓則合併成四分之三個圓，正如我們所預期。

### 最大值與最小值

在 $0 \le x \le 2a$ 上，我們有 $-a \le x - a \le a$，所以 $0 \le (x-a)^2 \le a^2$。因此

$$13\pi a^2 \le A \le 14\pi a^2,$$

這就證明了 $A \leq 14\pi a^2$。右邊的等號恰好在 $x = 0$ 或 $x = 2a$ 時成立，即山羊繫在角上的時候。左邊的界在 $x = a$ 時取得，所以 $A$ 的最小值是

$$A_{\min} = 13\pi a^2,$$

在山羊繫於一條邊的中點時取得。

## 驗算

- 在角上（$x = 0$），我們得到四分之三個半徑為 $4a$ 的圓（$12\pi a^2$）、繞過鄰邊遠端那個角的一個半徑為 $2a$ 的四分之一圓（$\pi a^2$），以及穀倉後方一個半徑為 $2a$ 的四分之一圓（$\pi a^2$）。合共 $14\pi a^2$，與公式相符。
- 我們取 $a = 1$，在一個細密的網格上估算吃草面積，逐點按它繞過穀倉的真正最短距離（一條直線，或在角上轉彎的折線）來判斷。繫於中點時得 40.845，而 $13\pi \approx 40.841$；繫於角上時得 43.988，而 $14\pi \approx 43.982$；其他繫繩點的結果也同樣吻合。

## 常見錯誤

- 忘記了穀倉後方的區域：繩子的長度足以在每一邊繞過兩個角。
- 擔心穀倉後方的兩塊區域重疊，或在計算中硬加入重疊部分。它們的半徑之和永遠恰好是 $2a$，所以兩個圓只是相切。

## STEP 之外

拉緊的繩子所走的是一條測地線（geodesic），即在有障礙物的區域內的最短路徑。在這裏，它由若干直線段組成，只在穀倉的角上轉彎，所以山羊能到達的點集，是在一種非歐幾里得（non-Euclidean）的距離量度方式下的一個「球」。繞過障礙物的最短路徑，通向變分法（calculus of variations），也通向計算機科學中的運動規劃（motion planning）問題。Siklos 的 *Advanced Problems in Mathematics* 收錄了更多這類題目，當中只要仔細分情況討論，便可把幾何問題化為簡單的代數。

詳細解答

# STEP I 2006，第 3 題

純數學 • 多項式的根

**STEP I 2006，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/06-s1-q3

## 關鍵思路

着眼於圖像，而不是公式。一條曲線若起初在軸的下方，最後又必須到達軸的上方，就一定會穿過軸；而 $x = 0$ 處的值（即常數項）告訴你曲線在原點處位於軸的哪一側。對於三次函數（cubic），關鍵在於轉向點（turning point）：曲線與軸相交三次，當且僅當它的局部極大值（local maximum）在軸的上方，而局部極小值（local minimum）在軸的下方。

## 解答

### 第 (i) 部分

設 $f(x) = x^2 + bx + c$。其圖像是一條開口向上的拋物線（parabola），所以當 $x \to \pm\infty$ 時，$f(x) \to +\infty$。若 $c < 0$，則 $f(0) = c < 0$：曲線在 $x = 0$ 處位於軸的下方。

由於 $f$ 連續（continuous），當 $x$ 是絕對值很大的負數時取正值，而在 0 處取負值，所以它在 $(-\infty, 0)$ 內某處穿過軸。同理，它在 $(0, \infty)$ 內再穿過軸一次。這兩個根（root）分別位於 0 的兩側，所以它們並不相同。二次方程（quadratic）最多有兩個根，因此該方程恰好有兩個相異實根，一負一正。所以 $c < 0$ 是*充分*條件（sufficient condition）。

它並*非必要*條件（necessary condition）。例如，$x^2 - 3x + 2 = (x - 1)(x - 2)$ 有相異實根 1 和 2，但 $c = 2 > 0$。（既必要又充分的條件是 $b^2 > 4c$；當 $c < 0$ 時，這個條件自動成立，因為 $b^2 \geq 0 > 4c$。）

### 第 (ii) 部分

我們斷言：該方程有兩個相異正根，當且僅當

$$b^2 > 4c, \qquad b < 0, \qquad c > 0.$$

*必要性*。假設兩根為 $\alpha \neq \beta$，其中 $\alpha, \beta > 0$。那麼 $x^2 + bx + c = (x - \alpha)(x - \beta)$，所以

$$c = \alpha\beta > 0, \qquad b = -(\alpha + \beta) < 0, \qquad b^2 - 4c = (\alpha + \beta)^2 - 4\alpha\beta = (\alpha - \beta)^2 > 0.$$

*充分性*。若 $b^2 > 4c$，則兩根 $\alpha, \beta$ 是相異實數。兩根之積為 $c > 0$，所以它們同號。兩根之和為 $-b > 0$，所以它們同為正數。

**想一想** 為甚麼單靠 $b < 0$ 和 $c > 0$ 並不足夠？

考慮 $x^2 - x + 1$：這裏 $b = -1 < 0$ 而 $c = 1 > 0$，但 $b^2 - 4c = -3 < 0$，所以根本沒有實根。從圖像來看，$c > 0$ 使曲線在 $x = 0$ 處位於軸的上方，而 $b < 0$ 使頂點（vertex）位於 $x = -\frac{b}{2} > 0$。我們還需要頂點位於軸的下方：頂點的高度是 $c - \frac{b^2}{4}$，它為負數當且僅當 $b^2 > 4c$。三個條件缺一不可。

### 第 (iii) 部分

設 $g(x) = x^3 + px + q$，所以 $g'(x) = 3x^2 + p$。實係數三次方程連重數（multiplicity）計算共有三個根，而它的非實根以共軛複數（complex conjugate）成對出現。另外請注意，在兩種情況下都有 $g(0) = q < 0$。

**$p > 0$、$q < 0$ 的情況**。 這裏對所有 $x$ 都有 $g'(x) = 3x^2 + p > 0$，所以 $g$ 嚴格遞增（strictly increasing），恰好有一個實根。由於 $g(0) = q < 0$ 而 $g$ 遞增，這個根是正數。餘下的兩個根是一對非實數的共軛複數。

所以有一個實根，它是正數而且是單根（simple root），另有兩個非實數的共軛複根。

**$p < 0$、$q < 0$ 的情況**。 設 $p = -3a^2$，其中 $a > 0$，即 $a = \sqrt{-p/3}$。那麼 $g'(x) = 3(x^2 - a^2)$，所以 $g$ 在 $(-\infty, -a]$ 上遞增，在 $[-a, a]$ 上遞減，在 $[a, \infty)$ 上遞增。在 $x = -a$ 處有局部極大值，在 $x = a$ 處有局部極小值，並且

$$g(-a) = -a^3 + 3a^3 + q = 2a^3 + q, \qquad g(a) = a^3 - 3a^3 + q = q - 2a^3.$$

由於 $q < 0$，局部極小值 $g(a)$ 是負數。在 $[0, a]$ 上，$g$ 由 $g(0) = q < 0$ 開始遞減，所以在該區間內沒有根。在 $[a, \infty)$ 上，它由 $g(a) < 0$ 遞增至 $+\infty$，所以恰好有一個大於 $a$ 的根，這個根是正數。任何其他實根必定位於 $(-\infty, -a]$ 或 $(-a, 0)$ 內，所以是負數；而這樣的根是否存在，只取決於局部極大值 $g(-a)$ 的正負。

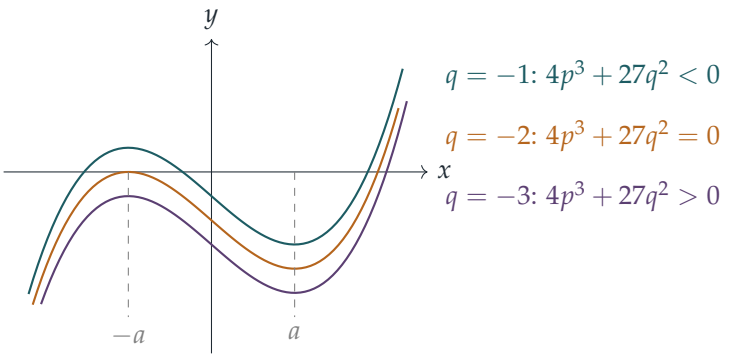
**想一想** $g(-a)$ 的正負與 $4p^3+27q^2$ 有甚麼關係？

把兩個轉向點的函數值相乘：

$$g(-a)\,g(a)=(q+2a^3)(q-2a^3)=q^2-4a^6=q^2-4\Big(-\frac{p}{3}\Big)^3=\frac{4p^3+27q^2}{27}.$$

由於 $g(a)<0$，$g(-a)$ 的正負與 $4p^3+27q^2$ 的正負相反。共有三種情況。

- $4p^3+27q^2>0$。此時 $g(-a)<0$，所以在整個 $(-\infty,a]$ 上都有 $g<0$（它先遞增至負值 $g(-a)$，然後遞減）。唯一的實根就是那個正根。與之前一樣，有一個正實根和兩個非實數的共軛複根。
- $4p^3+27q^2=0$。此時 $g(-a)=0$ 且 $g'(-a)=0$，所以 $x=-a$ 是重根（repeated root）。由於各根之和為 0（方程沒有 $x^2$ 項），第三個根是 $2a$。三個根是 $-a$、$-a$、$2a$：一個負的二重根和一個正的單根，全部都是實數。
- $4p^3+27q^2<0$。此時 $g(-a)>0$。由於當 $x\to-\infty$ 時 $g(x)\to-\infty$，在 $(-\infty,-a)$ 內有一個根。由於 $g(-a)>0>g(0)$，在 $(-a,0)$ 內還有一個根。連同大於 $a$ 的那個根，共有三個相異實根：兩負一正。




草圖顯示了 $q=-1$、$-2$、$-3$ 時的 $y=x^3-3x+q$（即 $p=-3$，$a=1$）：把局部極大值提升至越過軸，一個實根就變成三個。

在 $q<0$ 的每一種情況下，都恰好有一個正根；這與各根之積為 $-q>0$ 一致。

## 驗算

- 取 $p=-3$、$q=-1$，得 $4p^3+27q^2=-108+27=-81<0$，而數值計算所得的根約為 $-1.532$、$-0.347$ 和 $1.879$：兩負一正，與預測相符。
- 取 $p=-3$、$q=-2$：$4p^3+27q^2=0$，而 $x^3-3x-2=(x+1)^2(x-2)$，得出根 $-1,-1,2$。

- 在每種情況下，對數千組隨機的 $(p, q)$ 使用數值求根程式，並以隨機的 $(b, c)$ 檢驗 (ii) 中的條件，結果與上述每一項結論都吻合。

## 常見錯誤

- 在 (ii) 中只給出 $b < 0$ 和 $c > 0$（這兩個條件只保證如果根是實數，它們便是正數），而忘記了 $b^2 > 4c$。
- 在 (ii) 中只證明必要性，或只證明充分性。「必要且充分」需要兩個方向都證明。
- 在 (iii) 中忘記了邊界情況 $4p^3 + 27q^2 = 0$，此時有重根。

## STEP 之外

$-(4p^3 + 27q^2)$ 這個量是三次式 $x^3 + px + q$ 的判別式（discriminant），與 $b^2 - 4ac$ 完全對應。它等於各根兩兩之差的平方之積，所以它為零當且僅當有重根，而它為正數當且僅當三個根都是實數且互不相同。判別式，以及根與係數之間的關係，通向伽羅瓦理論（Galois theory）以及「哪些方程可用根式求解（solvable by radicals）」這個問題。Allenby 的 *Rings, Fields and Groups*（環、域與羣）是很好的入門讀物。

詳細解答

# STEP I 2006，第 4 題

純數學 • 正多邊形的面積

**STEP I 2006，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：

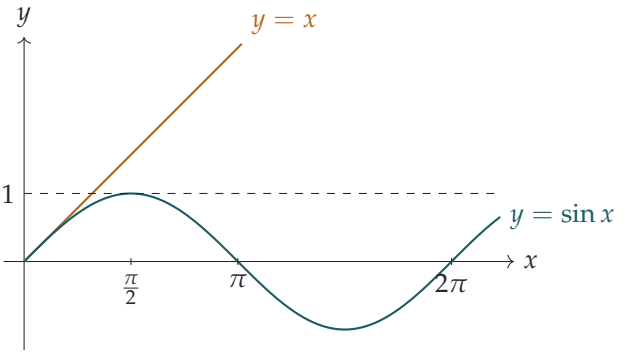
step.maths.org/questions/06-s1-q4

## 關鍵思路

直線 $y = x$ 是 $y = \sin x$ 在原點的切線（tangent），而正弦曲線向下彎離這條直線。單憑這一幅圖，就同時得到不等式和近似式。至於多邊形，把它分割成 $n$ 個在中心相接的全等（congruent）等腰三角形；所有的量於是都變成半角 $\pi/n$ 的三角運算，而不等式 $\sin x < x$ 原來正好決定了導數（derivative）的正負號。

## 解答

### 兩個圖像




兩條曲線都通過原點，而且在該處的斜率 (gradient) 都是 1，因為在 $x = 0$ 時 $\frac{\mathrm{d}}{\mathrm{d}x}\sin x = \cos x = 1$。所以 $y = x$ 是正弦曲線在 $O$ 的切線。

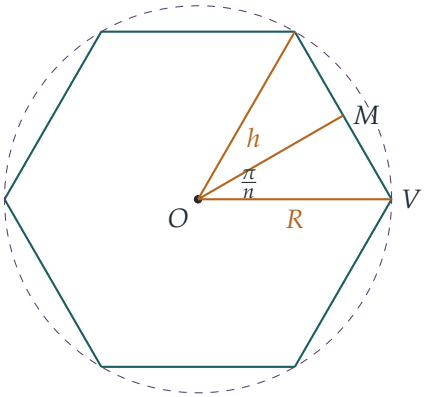
**(i)** 當 $0 < x \leq \frac{\pi}{2}$ 時，正弦曲線是凹的（concave）（$\frac{\mathrm{d}^2}{\mathrm{d}x^2}\sin x = -\sin x < 0$），所以它位於其在 $O$ 的切線之下，由此得 $\sin x < x$。（等價的說法：$g(x) = x - \sin x$ 滿足

$g(0) = 0$，且當 $0 < x < 2\pi$ 時 $g'(x) = 1 - \cos x > 0$，所以 $g$ 在該區間上嚴格遞增（strictly increasing）。當 $x > \frac{\pi}{2}$ 時，我們有 $\sin x \le 1 < \frac{\pi}{2} < x$。所以對所有 $x > 0$ 都有 $x > \sin x$。

**(ii)** 由於這條直線是在 $O$ 的切線，兩個圖像在原點附近幾乎無法分辨：當 $x$ 很小時，高度 $\sin x$ 與高度 $x$ 非常接近，所以 $\dfrac{\sin x}{x} \approx 1$。（更精確的說法：由於當 $x > 0$ 時 $\sin x \ge x - \frac{x^3}{6}$，我們有 $1 - \frac{x^2}{6} \le \frac{\sin x}{x} < 1$。）

### 多邊形的面積

把多邊形的中心 $O$ 與各頂點連起來。這樣就把多邊形分割成 $n$ 個全等的等腰三角形，每個三角形的底等於一條邊 $s = P/n$，在 $O$ 的頂角為 $2\pi/n$。從 $O$ 作垂線至一條邊的中點 $M$；這條垂線平分頂角。




在直角三角形 $OMV$ 中（其中 $V$ 是一個頂點），在 $O$ 的角是 $\pi/n$，且 $MV = \frac{1}{2}s$。所以由中心到一條邊的距離 $h = OM$ 是

$$h = \frac{s/2}{\tan(\pi/n)} = \frac{P}{2n\tan(\pi/n)}.$$

這 $n$ 個三角形每個的面積都是 $\frac{1}{2}sh$，所以多邊形的面積是

$$A = n \cdot \tfrac{1}{2} \cdot \frac{P}{n} \cdot \frac{P}{2n\tan(\pi/n)} = \frac{P^2}{4n\tan(\pi/n)},$$

正如所求。

### 面積隨 $n$ 增加

固定 $P$，把面積寫成 $A = \dfrac{P^2}{4f(n)}$，其中 $f(n) = n\tan\dfrac{\pi}{n}$。由於當 $n \ge 3$ 時 $f(n) > 0$，$A$ 遞增恰好等同於 $f$ 遞減。根據積法則（product rule）和鏈式法則（chain rule），

$$f'(n) = \tan\frac{\pi}{n} + n\sec^2\frac{\pi}{n} \cdot \left(-\frac{\pi}{n^2}\right) = \tan\theta - \theta\sec^2\theta, \qquad \text{其中} \theta = \frac{\pi}{n}.$$

**想一想** 我們怎樣判斷 $\tan\theta - \theta\sec^2\theta$ 的正負號？

抽出正因子 $\sec^2\theta$，並利用 $\sin\theta\cos\theta = \frac{1}{2}\sin 2\theta$：

$$f'(n) = \sec^2\theta\left(\sin\theta\cos\theta - \theta\right) = \tfrac{1}{2}\sec^2\theta\left(\sin 2\theta - 2\theta\right).$$

當 $n \geq 3$ 時，我們有 $0 < \theta \leq \frac{\pi}{3}$，所以 $2\theta > 0$，由第 (i) 部分得 $\sin 2\theta < 2\theta$。因此 $f'(n) < 0$，並且

$$\frac{\mathrm{d}A}{\mathrm{d}n} = -\frac{P^2 f'(n)}{4f(n)^2} > 0.$$

所以，在周界（perimeter）固定的情況下，面積隨 $n$ 增加而增加。

### 與外接圓比較

包含這個多邊形的最小的圓是外接圓（circumcircle），它以 $O$ 為圓心，並通過所有頂點。要看出更小的圓都不行，設各頂點為 $V_1, \ldots, V_n$，其中 $\overrightarrow{OV_k} = \mathbf{v}_k$，$|\mathbf{v}_k| = R$。由對稱性可知 $\sum \mathbf{v}_k = \mathbf{0}$，所以對任意一點 $C$，設 $\overrightarrow{OC} = \mathbf{c}$，則有

$$\frac{1}{n}\sum_{k=1}^{n}|\mathbf{v}_k - \mathbf{c}|^2 = R^2 - \frac{2}{n}\,\mathbf{c}\cdot\sum_{k=1}^{n}\mathbf{v}_k + |\mathbf{c}|^2 = R^2 + |\mathbf{c}|^2 \geq R^2.$$

所以必有某個頂點與 $C$ 的距離至少是 $R$，而以 $C$ 為圓心並包含這個多邊形的圓，半徑至少是 $R$。外接圓半徑（circumradius）是 $R = OV$，而由同一個直角三角形可得

$$\tfrac{1}{2}s = R\sin\theta \quad\Longrightarrow\quad R = \frac{P}{2n\sin\theta}.$$

所以兩個面積之比是

$$\frac{A}{\pi R^2} = \frac{P^2}{4n\tan\theta}\cdot\frac{4n^2\sin^2\theta}{\pi P^2} = \frac{n}{\pi}\cdot\frac{\sin^2\theta}{\tan\theta} = \frac{\sin\theta\cos\theta}{\theta} = \frac{\sin 2\theta}{2\theta},$$

當中用到 $n/\pi = 1/\theta$。當 $n$ 很大時，$2\theta = 2\pi/n$ 很小，所以由第 (ii) 部分可知這個比值約等於 1。

## 驗算

對於正方形（$n = 4$，$\theta = \pi/4$），公式給出 $A = P^2/16$，這是正確的，因為每條邊長 $P/4$；面積比是 $\sin(\pi/2)/(\pi/2) = 2/\pi \approx 0.637$，與單位圓內邊長為 $\sqrt{2}$ 的正方形吻合。對於六邊形，這個比值是 $3\sqrt{3}/(2\pi) \approx 0.827$。我們用鞋帶公式（shoelace formula）核對了 $n = 3, \ldots, 39$ 時的面積公式，用數值導數核對了 $f'(n) < 0$，並求得這個比值在 $n = 100$ 時是 0.99934，在 $n = 1000$ 時是 0.9999934。

## 常見錯誤

- 在直角三角形中用了整個頂角 $2\pi/n$，而不是半角 $\pi/n$。
- 對 $\tan(\pi/n)$ 求導時，漏掉鏈式法則帶來的因子 $-\pi/n^2$。
- 聲稱「因為 $\tan\theta \approx \theta$」，所以 $\tan\theta < \theta\sec^2\theta$：近似式不能決定正負號。能決定正負號的是精確的不等式 $\sin 2\theta < 2\theta$。

## STEP 之外

極限（limit）$\sin x/x \to 1$ 正是令 $\sin x$ 的導數等於 $\cos x$ 的事實，而大學的第一門分析（analysis）課程會從 sin 的定義出發嚴格地證明它，而不是靠一幅圖。這個單調性（monotonicity）結果，讓我們從離散的角度一窺等周不等式（isoperimetric inequality）：在所有長度給定的閉曲線之中，圓所圍的面積最大，而多邊形的面積會遞增並趨向圓的面積值 $P^2/(4\pi)$。如想看極限和導數的嚴格論述，可以讀 Abbott 的 *Understanding Analysis*（《理解分析》）。

詳細解答

# STEP I 2006，第 5 題

純數學 • 代換積分法

**STEP I 2006，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/06-s1-q5

## 關鍵思路

兩個被積函數(integrand)都含有一個棘手的平方根，而要消去線性式的平方根，最乾淨利落的方法就是把整個根式設為新變量。根式一旦消失，剩下的便是有理函數（rational function），而有理函數總可以用部分分式（partial fractions）來積分。在第 (ii) 部分，最好先把指數函數除去，使平方根變成某個線性式的根，然後同一招便能派上用場。

## 解答

### 第 (i) 部分

由於 $x > 4$，我們有 $2x + 1 > 9$，故可取 $u = \sqrt{2x+1} > 3$。由 $u^2 = 2x + 1$ 得

$$2u\,\mathrm{d}u = 2\,\mathrm{d}x, \qquad \text{所以} \qquad \mathrm{d}x = u\,\mathrm{d}u,$$

以及

$$x - 4 = \frac{u^2 - 1}{2} - 4 = \frac{u^2 - 9}{2}.$$

代入後，

$$\int \frac{3}{(x-4)\sqrt{2x+1}}\,\mathrm{d}x = \int \frac{3}{\frac{1}{2}(u^2-9)\,u}\,u\,\mathrm{d}u = \int \frac{6}{u^2-9}\,\mathrm{d}u.$$

來自 $\mathrm{d}x$ 的因子 $u$ 恰好把平方根約去，這正是這個代換（substitution）奏效的原因。現在 $u^2 - 9 = (u-3)(u+3)$，而

$$\frac{6}{(u-3)(u+3)} = \frac{1}{u-3} - \frac{1}{u+3},$$

你可以把右邊通分來驗證：分子是 $(u+3)-(u-3)=6$。因此

$$\int \frac{6}{u^2-9}\,\mathrm{d}u = \ln(u-3)-\ln(u+3)+K = \ln\left(\frac{u-3}{u+3}\right)+K,$$

這裏無需加上絕對值符號，因為 $u>3$ 令 $u-3$ 和 $u+3$ 均為正數。把 $u=\sqrt{2x+1}$ 代回，得

$$\int \frac{3}{(x-4)\sqrt{2x+1}}\,\mathrm{d}x = \ln\left(\frac{\sqrt{2x+1}-3}{\sqrt{2x+1}+3}\right)+K,$$

正是所求。

**第 (ii) 部分**

設 $I$ 為該積分。仿照第 (i) 部分，最自然的代換就是把平方根設為新變量：

$$w=\sqrt{\mathrm{e}^x+1}, \qquad w^2=\mathrm{e}^x+1.$$

於是 $2w\,\mathrm{d}w=\mathrm{e}^x\,\mathrm{d}x$，所以 $\mathrm{d}x=\dfrac{2w}{\mathrm{e}^x}\,\mathrm{d}w=\dfrac{2w}{w^2-1}\,\mathrm{d}w$。積分上下限變為

$$x=\ln 3:\ w=\sqrt{3+1}=2, \qquad x=\ln 8:\ w=\sqrt{8+1}=3.$$

因此

$$I=\int_2^3 \frac{2}{(w^2-1)\,w}\cdot\frac{2w}{w^2-1}\,\mathrm{d}w=\int_2^3\frac{4}{(w^2-1)^2}\,\mathrm{d}w.$$

? **想一想** 怎樣才能不經冗長的計算，就把 $\dfrac{4}{(w^2-1)^2}$ 分解成部分分式？

從簡單的分解式 $\dfrac{2}{w^2-1}=\dfrac{1}{w-1}-\dfrac{1}{w+1}$ 出發，把它平方：

$$\begin{aligned}\frac{4}{(w^2-1)^2} &= \frac{1}{(w-1)^2}-\frac{2}{(w-1)(w+1)}+\frac{1}{(w+1)^2}\\ &= \frac{1}{(w-1)^2}+\frac{1}{(w+1)^2}-\frac{1}{w-1}+\frac{1}{w+1},\end{aligned}$$

其中對中間一項再用了一次那個簡單的分解式。現在每一項都可以一眼看出積分：

$$\int\frac{4}{(w^2-1)^2}\,\mathrm{d}w=-\frac{1}{w-1}-\frac{1}{w+1}-\ln\left(\frac{w-1}{w+1}\right)+C,$$

由於在積分範圍內 $w\geq 2$，所以無需加上絕對值符號。代入上下限計算，

$$\begin{aligned}\text{當}w=3:&\quad -\tfrac{1}{2}-\tfrac{1}{4}-\ln\tfrac{2}{4}=-\tfrac{3}{4}+\ln 2,\\ \text{當}w=2:&\quad -1-\tfrac{1}{3}-\ln\tfrac{1}{3}=-\tfrac{4}{3}+\ln 3.\end{aligned}$$

所以

$$I = \left(-\tfrac{3}{4} + \ln 2\right) - \left(-\tfrac{4}{3} + \ln 3\right) = \frac{7}{12} + \ln\frac{2}{3},$$

這就是要證明的結果。

你同樣可以分兩步來做：先設 $v = \mathrm{e}^x$（得出 $\int_3^8 \frac{2}{v^2\sqrt{v+1}}\,\mathrm{d}v$），然後設 $w^2 = v+1$；上面的單一代換則一次過完成了這兩步。

## 驗算

- 對 (i) 的答案求導：設 $u = \sqrt{2x+1}$，則 $\frac{\mathrm{d}u}{\mathrm{d}x} = \frac{1}{u}$，而 $\frac{\mathrm{d}}{\mathrm{d}u}\ln\frac{u-3}{u+3} = \frac{6}{u^2-9}$，所以導數為 $\frac{6}{u(u^2-9)} = \frac{3}{(x-4)\sqrt{2x+1}}$。
- 數值上，$\frac{7}{12} + \ln\frac{2}{3} \approx 0.58333 - 0.40547 = 0.17787$，而在 $[\ln 3, \ln 8]$ 上對原來的被積函數應用辛普森法則（Simpson's rule），得出 0.177868，兩者有六位數字吻合。
- 就正負號和大小作一個合理性檢查：(ii) 的被積函數為正，並在長度為 $\ln\frac{8}{3} \approx 0.98$ 的區間上由約 0.33 下降至約 0.08，所以答案略小於 0.2 正合預期。

## 常見錯誤

- 在 (ii) 中變換變量時忘記更改積分上下限，或把 $x$ 的上下限代入以 $w$ 表示的式子。
- 把 $\frac{4}{(w^2-1)^2}$ 只寫成 $\frac{A}{(w-1)^2} + \frac{B}{(w+1)^2}$：重複因子還需要線性項 $\frac{1}{w-1}$ 和 $\frac{1}{w+1}$。
- 在 (i) 中得出 $\mathrm{d}x = \mathrm{d}u$ 而不是 $\mathrm{d}x = u\,\mathrm{d}u$，結果留下一個多餘的平方根。

## STEP 之外

兩部分的答案都由對數和有理項組成，原因在於一條一般性的定理：每個有理函數都有初等原函數（elementary antiderivative），可藉部分分式由有理函數、對數和反正切函數（inverse tangent）構成；而任何關於 $x$ 和 $\sqrt{ax+b}$ 的有理被積函數，經代換 $u^2 = ax+b$ 後都會變成有理函數。把這個想法推廣到二次式的根，便引出歐拉代換（Euler's substitutions）；而追問哪些積分不能以初等形式求出，則引出劉維爾（Liouville）的有限形式積分理論（integration in finite terms）。若想細緻地了解積分及其背後的道理，可閱讀 Abbott 的 *Understanding Analysis*（一本實分析入門書）。

詳細解答

# STEP I 2006，第 6 題

純數學 • 佩爾方程的變換

**STEP I 2006，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/06-s1-q6

## 關鍵思路

每一部分都在問：某個線性坐標變換（linear change of coordinates）會否保持一個二次式（quadratic expression）不變。直接的做法是代入、展開，然後比較 $a^2$、$ab$ 和 $b^2$ 的係數（coefficients）。背後其實有一幅更簡潔的圖像：$x^2 - 2y^2$ 是數 $x + y\sqrt{2}$ 的「大小」，而乘以 $3 + 2\sqrt{2}$（其大小為 $9 - 8 = 1$）不會改變這個大小。看通這一點，你就確切知道 (iii) 中的映射（map）該如何構造。

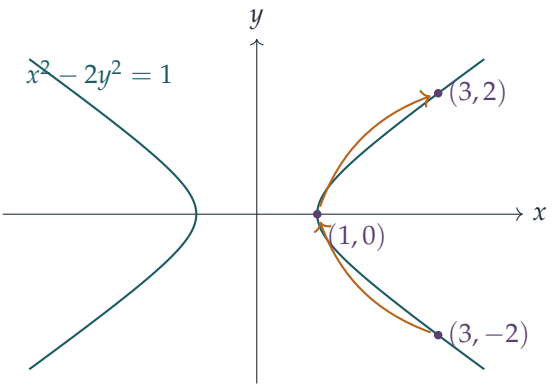
## 解答

### 第 (i) 部分

設 $a^2 - 2b^2 = 1$。則

$$\begin{aligned}(3a + 4b)^2 - 2(2a + 3b)^2 &= (9a^2 + 24ab + 16b^2) - 2(4a^2 + 12ab + 9b^2) \\ &= a^2 - 2b^2 \\ &= 1,\end{aligned}$$

所以 $(3a + 4b,\ 2a + 3b)$ 位於 $x^2 - 2y^2 = 1$ 上，正如所求。事實上，頭兩行是一條恆等式（identity），對任何 $a$ 和 $b$ 都成立：這個映射在每一處都保持 $x^2 - 2y^2$ 不變，而不僅是在曲線上。




這個映射把點沿着右邊的分支推移：$(3,-2) \mapsto (1,0) \mapsto (3,2) \mapsto (17,12) \mapsto \cdots$。

**第 (ii) 部分**

設 $Ma^2 - Nb^2 = 1$。展開得

$$M(5a+6b)^2 - N(4a+5b)^2 = (25M-16N)a^2 + (60M-40N)ab + (36M-25N)b^2.$$

像點位於曲線上，當且僅當上式等於 $1 = Ma^2 - Nb^2$。減去 $Ma^2 - Nb^2$ 後，條件變成

$$\begin{aligned} 0 &= (24M-16N)a^2 + (60M-40N)ab + (36M-24N)b^2 \\ &= 4(3M-2N)(2a^2+5ab+3b^2) \\ &= 4(3M-2N)(2a+3b)(a+b). \end{aligned}$$

**?** **想一想** 我們真的需要 $3M = 2N$ 嗎？會不會是其他因式等於零？

這個條件必須對曲線上每一點都成立。當 $M, N > 0$ 時，曲線 $Mx^2 - Ny^2 = 1$ 是一條雙曲線（hyperbola），上面有無窮多個點。直線 $2x + 3y = 0$ 和 $x + y = 0$ 各自與它最多相交於兩點，所以曲線上絕大部分的點都滿足 $(2a+3b)(a+b) \neq 0$。因此我們需要 $3M - 2N = 0$。反過來說，若 $3M = 2N$，條件在每一點都成立。

所以 $3M = 2N$。於是 $2 \mid 3M$，故 $M$ 是偶數。最小的選擇是 $M = 2$，由此得 $N = 3$。最後驗算一次：

$$2(5a+6b)^2 - 3(4a+5b)^2 = (50-48)a^2 + (120-120)ab + (72-75)b^2 = 2a^2 - 3b^2.$$

最小的正整數是 $M = 2$ 和 $N = 3$。

**第 (iii) 部分**

我們要找正整數，使得

$$(Pa+Qb)^2 - 3(Ra+Sb)^2 = a^2 - 3b^2 \quad \text{對任意} a, b,$$

這樣必定能令曲線上每一點的像都位於曲線上。比較係數：

$$P^2 - 3R^2 = 1, \qquad PQ = 3RS, \qquad Q^2 - 3S^2 = -3.$$

第一條方程要求在 $x^2 - 3y^2 = 1$ 上找一個正整數點 $(P, R)$。最小的是 $(2,1)$，所以試取 $P = 2$，$R = 1$。第二條方程隨即給出 $Q = \frac{3}{2}S$，而第三條變成 $\frac{9}{4}S^2 - 3S^2 = -3$，即 $S^2 = 4$。所以 $S = 2$ 而 $Q = 3$：

$$P = 2, \quad Q = 3, \quad R = 1, \quad S = 2.$$

驗算：$(2a + 3b)^2 - 3(a + 2b)^2 = 4a^2 + 12ab + 9b^2 - 3a^2 - 12ab - 12b^2 = a^2 - 3b^2 = 1$。所以 $(2a + 3b,\ a + 2b)$ 位於 $x^2 - 3y^2 = 1$ 上。

**想一想** 這樣的映射是從哪裏來的？

寫出 $(a + b\sqrt{3})(2 + \sqrt{3}) = (2a + 3b) + (a + 2b)\sqrt{3}$。$x + y\sqrt{3}$ 的範數（norm）$x^2 - 3y^2$ 具有積性（multiplicative），而 $2 + \sqrt{3}$ 的範數是 $4 - 3 = 1$，所以乘以它會保持範數不變。第 (i) 部分的原理相同，用的是 $3 + 2\sqrt{2}$。至於 (ii)，$(a\sqrt{2} + b\sqrt{3})(5 + 2\sqrt{6}) = (5a + 6b)\sqrt{2} + (4a + 5b)\sqrt{3}$，這就是 $2x^2 - 3y^2$ 這個形式在那裏出現的原因。(iii) 還有其他正確答案：乘以 $(2 + \sqrt{3})^2 = 7 + 4\sqrt{3}$ 會得出 $(7a + 12b,\ 4a + 7b)$。

## 驗算

- 從 $x^2 - 3y^2 = 1$ 上的 $(1,0)$ 出發。重複施行 (iii) 的映射得出 $(2,1)$、$(7,4)$、$(26,15)$、$(97,56)$，每一個都滿足 $x^2 - 3y^2 = 1$（例如 $26^2 - 3 \cdot 15^2 = 676 - 675 = 1$）。
- 對於 (ii)，$(1/\sqrt{2}, 0)$ 在 $2x^2 - 3y^2 = 1$ 上，並映射到 $(5/\sqrt{2},\ 4/\sqrt{2})$。於是 $2 \cdot \frac{25}{2} - 3 \cdot \frac{16}{2} = 25 - 24 = 1$。
- 在 $1 \le M, N < 30$ 的範圍內作窮舉搜尋，並以每條雙曲線上的實際點來測試，證實 $(2,3)$ 是最小的可行數對。

## 常見錯誤

- 在 (ii) 中，只令 $ab$ 的係數為零，卻沒有檢查 $a^2$ 和 $b^2$ 的係數是否隨之等於 $M$ 和 $-N$。在這題中它們確實如此，但這一點必須寫出來證明。
- 沒有給出理由就說「恆等式必須對所有 $a, b$ 成立」，而題目其實只給出曲線上的點。你需要一個論證，例如上面的因式分解（factorisation），來排除其他因式等於零的可能。

## STEP 之外

這些都是佩爾方程（Pell equations）$x^2 - dy^2 = 1$。它們的整數解由環（ring）$\mathbb{Z}[\sqrt{d}]$ 中一個基本單位（fundamental unit）$x_1 + y_1\sqrt{d}$ 的冪生成，而這裏的映射正是乘以這類單位。用矩陣的語言來說，每個映射都是一個保持某個二次型（quadratic form）不變的 $2 \times 2$ 整數矩陣，即某個「整正交羣」(integral orthogonal group) 的元素。Davenport 的 *The Higher Arithmetic*（《高等算術》）有精彩的一整章，以連分數（continued fractions）處理佩爾方程。

詳細解答

# STEP I 2006，第 7 題

純數學 • 含絕對值符號的積分

**STEP I 2006，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/06-s1-q7

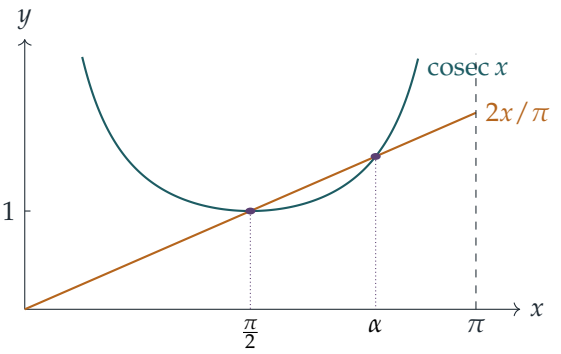
## 關鍵思路

把 $x\sin x = \pi/2$ 兩邊除以 $\frac{\pi}{2}\sin x$（它在 $(0,\pi)$ 上為正），方程便化為 $\operatorname{cosec} x = 2x/\pi$，所以方程的根（root）就是一條直線與一條 U 形曲線相交之處，而一條直線與一條凸（convex）曲線最多只能相交兩次。至於積分，處理絕對值（modulus）的全部竅門，就是找出絕對值符號內的式子在何處變號，在該處把積分範圍分開，然後以正確的正負號對每一段分別積分。

## 解答

### 第 (i) 部分：草圖與根

在 $0 < x < \pi$ 上，$\operatorname{cosec} x$ 在兩端均趨向 $+\infty$，關於 $x = \pi/2$ 對稱，並在該處取得最小值 1。直線 $y = 2x/\pi$ 由 $(0,0)$ 伸延至 $(\pi,2)$，並通過 cosec 曲線的最低點 $(\pi/2,1)$。



由於在 $(0,\pi)$ 上 $\sin x > 0$，

$$x\sin x = \frac{\pi}{2} \iff \frac{1}{\sin x} = \frac{2x}{\pi},$$

所以方程的根就是草圖中各交點的 $x$ 坐標。其中一個根一眼便可看出：$x = \pi/2$，因為 $\frac{\pi}{2}\sin\frac{\pi}{2} = \frac{\pi}{2}$。

為了確定恰好有兩個根，設 $h(x) = \operatorname{cosec} x - 2x/\pi$。那麼

$$h''(x) = \operatorname{cosec} x\,(\operatorname{cosec}^2 x + \cot^2 x) > 0 \quad \text{於}(0,\pi),$$

所以 $h$ 是嚴格凸（strictly convex）的，而嚴格凸函數最多只有兩個零點。（假如它有三個零點 $a < b < c$，那麼 $b$ 嚴格地位於 $a$ 與 $c$ 之間，但 $h(b)$ 卻落在連接 $(a,0)$ 與 $(c,0)$ 的弦（chord）上，而不是嚴格地在弦的下方。）現在 $h'(\pi/2) = -\cot\frac{\pi}{2}\operatorname{cosec}\frac{\pi}{2} - \frac{2}{\pi} = -\frac{2}{\pi} < 0$，所以 $h$ 在 $\pi/2$ 的右側附近為負，而當 $x \to \pi^-$ 時 $h(x) \to +\infty$。根據連續性（continuity），$h$ 有一個零點 $\alpha$，滿足 $\pi/2 < \alpha < \pi$。因此方程恰好有兩個根，即 $\pi/2$ 和 $\alpha$，而 $\alpha$ 是較大的一個。

### 第 (i) 部分：積分

由草圖（以及上述關於凸性的論證）可知，cosec 曲線在兩個根之間位於直線下方，而在 $\alpha$ 之後則位於直線上方。把 $\operatorname{cosec} x \leqslant 2x/\pi$ 乘以正數 $\frac{\pi}{2}\sin x$，即得

$$x\sin x - \frac{\pi}{2} \geq 0 \text{ 於 } \left[\tfrac{\pi}{2},\alpha\right], \qquad x\sin x - \frac{\pi}{2} \leq 0 \text{ 於} [\alpha,\pi].$$

**想一想** 我們應怎樣處理絕對值符號？

我們在變號之處 $\alpha$ 把積分分開，而在第二段上則對該式子的相反數積分。

運用分部積分法（integration by parts），取 $u = x$，$\mathrm{d}v = \sin x\,\mathrm{d}x$，

$$\int x\sin x\,\mathrm{d}x = -x\cos x + \int \cos x\,\mathrm{d}x = \sin x - x\cos x.$$

所以 $x\sin x - \pi/2$ 的一個原函數（antiderivative）是

$$F(x) = \sin x - x\cos x - \frac{\pi x}{2},$$

而所求的積分是

$$I = \big[F(x)\big]_{\pi/2}^{\alpha} - \big[F(x)\big]_{\alpha}^{\pi} = 2F(\alpha) - F\left(\tfrac{\pi}{2}\right) - F(\pi).$$

這三個值分別是

$$F(\alpha) = \sin\alpha - \alpha\cos\alpha - \frac{\alpha\pi}{2},$$

$$F\left(\tfrac{\pi}{2}\right) = 1 - 0 - \frac{\pi^2}{4},$$

$$F(\pi) = 0 - \pi(-1) - \frac{\pi^2}{2} = \pi - \frac{\pi^2}{2}.$$

因此

$$I = 2\sin\alpha - 2\alpha\cos\alpha - \alpha\pi - 1 + \frac{\pi^2}{4} - \pi + \frac{\pi^2}{2}$$
$$= 2\sin\alpha + \frac{3\pi^2}{4} - \alpha\pi - \pi - 2\alpha\cos\alpha - 1,$$

這正是所要求的結果。

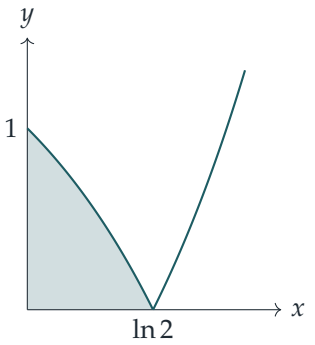
### 第 (ii) 部分

**想一想** 兩個絕對值符號之中，哪一個在這裏才真正起作用？

在正 $x$ 軸上，$x \geq 0$，所以 $\mathrm{e}^x \geq 1$，內層的絕對值符號並無影響：$|\mathrm{e}^x - 1| = \mathrm{e}^x - 1$。因此曲線就是

$$y = |\mathrm{e}^x - 2| \qquad (x \geq 0).$$

當 $0 \leq x \leq \ln 2$ 時，它等於 $2 - \mathrm{e}^x$，由 $x = 0$ 處的 $y = 1$ 下降至 $x = \ln 2$ 處的 $y = 0$；當 $x \geq \ln 2$ 時，它等於 $\mathrm{e}^x - 2$，無止境地上升。




所以曲線只在 $x = \ln 2$ 處觸及 $x$ 軸，而由正 $x$ 軸、$y$ 軸和曲線所圍成的區域，就只有 $0 \leq x \leq \ln 2$ 上的陰影部分（過了 $\ln 2$ 之後，曲線一直向上攀升，再圍不成任何封閉區域）。它的面積是

$$\int_0^{\ln 2} (2 - \mathrm{e}^x)\,\mathrm{d}x = \Big[2x - \mathrm{e}^x\Big]_0^{\ln 2} = (2\ln 2 - 2) - (0 - 1)$$
$$= 2\ln 2 - 1 = \ln 4 - 1,$$

正如所求。

## 驗算

用二分法（bisection）可得 $\alpha \approx 2.4433$，代入 (i) 的公式後得出約 0.61343；把森遜法則（Simpson's rule）直接應用於 $[\pi/2, \pi]$ 上的 $|x\sin x - \pi/2|$，所得的值在小數點後 14 位

以內完全相同。在 $(0, \pi)$ 上作格點搜尋（grid search），發現 $x \sin x - \pi/2$ 恰好變號兩次，分別在 1.5708 和 2.4433 附近。在 (ii) 中，$\ln 4 - 1 \approx 0.38629$，同樣與數值積分的結果吻合。作為對 (i) 中正負號的合理性檢查，答案必須為正，而它確實為正。

## 常見錯誤

- 沒有在 $\alpha$ 處分開，便直接在整個 $[\pi/2, \pi]$ 上對 $x \sin x - \pi/2$ 積分：這樣得出的是有向面積（signed area）（這裏是負數），而不是絕對值的積分。
- 遺漏了 $x = \pi/2$ 這個根，因為直線剛好通過 cosec 曲線的最低點，圖形看起來可能像是相切。其實並非相切：曲線在該處的斜率是 0，而直線的斜率是 $2/\pi$，所以兩者是相交而過的。
- 在 (ii) 中，把兩個絕對值符號都保留下來，結果分出過多的情況。應先利用 $x \geq 0$ 確定 $\mathrm{e}^x - 1$ 的正負號。

## STEP 之外

根 $\alpha$ 沒有閉合形式（closed form），但我們仍然可以用它把積分準確地表達出來：這是分析學（analysis）中常見的情況，即一個量是由一條方程而非一條公式所確定的。嚴格凸函數與一條直線最多相交兩次，這只是凸性眾多精妙推論之一；凸性也是延森不等式（Jensen's inequality）以及最優化（optimisation）中許多理論的基礎。Bartle 與 Sherbert 合著的 *Introduction to Real Analysis*（實分析入門）對連續性、介值定理（intermediate value theorem）和凸函數都有嚴謹細緻的論述。

詳細解答

# STEP I 2006，第 8 題

純數學 • 三維向量

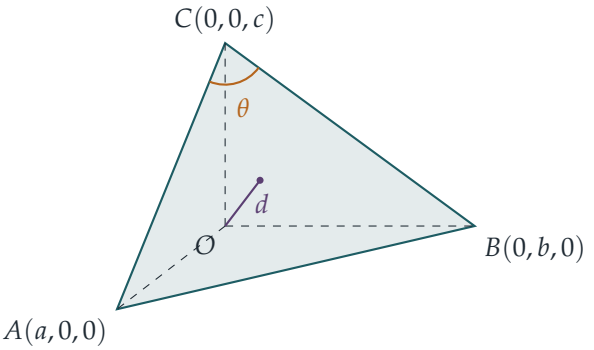
**STEP I 2006，第 8 題**

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/06-s1-q8

## 關鍵思路

四面體（tetrahedron）$OABC$ 在 $O$ 處有三條互相垂直的棱，所以若以 $OAB$ 為底，體積便很容易求出。同一個立體亦可以看成以 $ABC$ 為底，這時的高正好就是我們要求的距離 $d$。用兩種方法計算體積，便得到一條關於 $d$ 的方程，前提是我們知道 $ABC$ 的面積。這個面積可由 $C$ 處的角求得，而該角則由純量積（scalar product）給出。




## 解答

### 第 (i) 部分

取位於平面 $z = 0$ 上的三角形 $OAB$ 為底。由於 $OA$ 和 $OB$ 分別沿 $x$ 軸和 $y$ 軸，角 $AOB$ 是直角，所以

$$\text{面積}OAB = \tfrac{1}{2}ab.$$

高是 $C$ 到平面 $z=0$ 的距離，即 $c$，因為 $OC$ 沿 $z$ 軸，故與該平面垂直。因此

$$V = \tfrac{1}{3} \times \tfrac{1}{2}ab \times c = \frac{abc}{6}.$$

**第 (ii) 部分**

角 $ACB$ 是向量（vector）$\overrightarrow{CA}$ 與 $\overrightarrow{CB}$ 之間的夾角：

$$\overrightarrow{CA} = \begin{pmatrix} a \\ 0 \\ -c \end{pmatrix}, \qquad \overrightarrow{CB} = \begin{pmatrix} 0 \\ b \\ -c \end{pmatrix}.$$

它們的純量積是 $0+0+c^2 = c^2$，而長度分別是 $\sqrt{a^2+c^2}$ 和 $\sqrt{b^2+c^2}$。所以

$$\cos\theta = \frac{\overrightarrow{CA}\cdot\overrightarrow{CB}}{|\overrightarrow{CA}|\,|\overrightarrow{CB}|} = \frac{c^2}{\sqrt{(a^2+c^2)(b^2+c^2)}},$$

這就是所要求的結果。

要求面積，我們需要 $\sin\theta$。由於 $0<\theta<\pi$，$\sin\theta>0$，所以

$$\begin{aligned}\sin\theta = \sqrt{1-\cos^2\theta} &= \sqrt{\frac{(a^2+c^2)(b^2+c^2)-c^4}{(a^2+c^2)(b^2+c^2)}} \\ &= \frac{\sqrt{a^2b^2+b^2c^2+c^2a^2}}{\sqrt{(a^2+c^2)(b^2+c^2)}},\end{aligned}$$

因為展開 $(a^2+c^2)(b^2+c^2)$ 後，$c^4$ 項互相抵消。

**想一想** 哪一條三角形面積公式切合我們現有的資料？

我們知道兩條邊 $CA$ 和 $CB$，以及它們之間的夾角 $\theta$，所以使用 $\frac{1}{2}CA\cdot CB\sin\theta$。那些麻煩的平方根剛好約去：

$$\text{面積}ABC = \tfrac{1}{2}\sqrt{(a^2+c^2)(b^2+c^2)}\cdot\frac{\sqrt{a^2b^2+b^2c^2+c^2a^2}}{\sqrt{(a^2+c^2)(b^2+c^2)}} = \tfrac{1}{2}\sqrt{a^2b^2+b^2c^2+c^2a^2}.$$

令人欣喜的是，答案對 $a$、$b$、$c$ 是對稱的，而這也是必然的：三角形 $ABC$ 並不在乎我們把哪一個頂點稱為 $C$。

**距離 $d$**

現在把 $ABC$ 看成同一個四面體的底。它的高是對面的頂點 $O$ 到 $ABC$ 所在平面的垂直距離，即 $d$。由於體積與我們選哪一個面作底無關，由第 (i) 部分和第 (ii) 部分可得

$$\frac{abc}{6} = \tfrac{1}{3}\times\tfrac{1}{2}\sqrt{a^2b^2+b^2c^2+c^2a^2}\times d, \qquad \text{所以} \qquad d = \frac{abc}{\sqrt{a^2b^2+b^2c^2+c^2a^2}}.$$

兩邊平方後取倒數（各項均為正數），

$$\frac{1}{d^2} = \frac{a^2b^2 + b^2c^2 + c^2a^2}{a^2b^2c^2} = \frac{1}{c^2} + \frac{1}{a^2} + \frac{1}{b^2},$$

其中分子的每一項都除以 $a^2b^2c^2$。這就是所要求的結果。

## 驗算

- 取 $a = b = c = 1$。這時 $ABC$ 是邊長為 $\sqrt{2}$ 的等邊三角形，面積為 $\frac{\sqrt{3}}{4} \cdot 2 = \frac{\sqrt{3}}{2}$，與 $\frac{1}{2}\sqrt{3}$ 吻合。該平面是 $x + y + z = 1$，與 $O$ 的距離為 $1/\sqrt{3}$，而確實有 $1/d^2 = 3$。
- 取 $a = 1$、$b = c = 2$，由公式得體積為 $\frac{2}{3}$，面積為 $\sqrt{6} \approx 2.449$，以及 $d = 2/\sqrt{6} \approx 0.8165$，所以 $1/d^2 = \frac{3}{2} = 1 + \frac{1}{4} + \frac{1}{4}$。
- 數值驗證：隨機選取 10 000 組 $(a, b, c)$，體積（用純量三重積（scalar triple product）計算）、面積（用向量積（cross product）計算）和 $d$（直接找出垂足的位置來計算）全部與公式吻合，誤差在 $10^{-12}$ 以內。

## 常見錯誤

- 把 $\overrightarrow{AC}$ 與 $\overrightarrow{CB}$ 配搭使用：方向混亂會令純量積的正負號顛倒，得出的是 $\theta$ 的補角。
- 忘記了以 $ABC$ 為底時，高是由 $O$ 到該平面的垂線，而不是任何一條棱的長度。

## STEP 之外

通過 $A$、$B$、$C$ 的平面是 $\frac{x}{a} + \frac{y}{b} + \frac{z}{c} = 1$，其法向量(normal)為 $(\frac{1}{a}, \frac{1}{b}, \frac{1}{c})$，利用一般的點到平面距離公式，一行便可求出 $d$。面積的結果是三維的畢氏定理(Pythagoras' theorem)，即德古阿定理（de Gua's theorem）：$ABC$ 面積的平方，等於在 $O$ 處三個直角面的面積平方之和，因為 $\frac{1}{4}(a^2b^2 + b^2c^2 + c^2a^2)$ 正好就是這個和。這兩個想法都源自向量積和行列式（determinant），在大學第一門線性代數（linear algebra）課程的初段便會學到；Strang 的 *Introduction to Linear Algebra*（線性代數入門教科書）是認識它們的好地方。

詳細解答

# STEP I 2006，第 9 題

力學 • 連接質點

**STEP I 2006，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/06-s1-q9

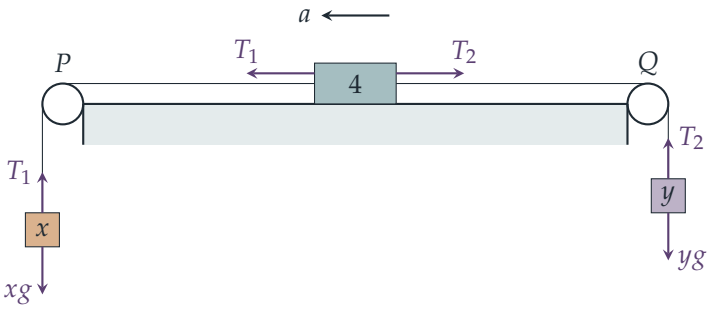
## 關鍵思路

整個運動分為兩個階段，每個階段的加速度都是常數。兩條繩子都連着的時候，三個物體組成的系統作為一個整體運動，由兩個懸掛重物的重量之差驅動。繩子剪斷之後，只有較輕的質點拉着方塊，而且拉力與運動方向相反，所以方塊作勻減速運動。剪斷一刻的速率把兩個階段連接起來，而約束條件（constraint）$x + y = 6$ 讓我們可以把一切都用 $y$ 表示。

## 解答

### 模型

我們把方塊和兩個懸掛的物體都視為質點（particle），繩子視為輕而不可伸長（light and inextensible），桌面和滑輪（pulley）視為光滑（smooth）。由於 $x > y$，方塊向 $P$ 移動，質量為 $x$ 的質點下降，質量為 $y$ 的質點上升。設跨過 $P$ 和 $Q$ 的繩子的張力（tension）分別為 $T_1$ 和 $T_2$。



**第一階段：兩條繩子都連着**

設共同的加速度為 $a_1$（方塊向 $P$，質點 $x$ 向下，質點 $y$ 向上）。對每個物體應用牛頓第二定律（Newton's second law），得

$$\begin{aligned} xg - T_1 &= xa_1, \\ T_1 - T_2 &= 4a_1, \\ T_2 - yg &= ya_1. \end{aligned}$$

三式相加便消去張力：

$$(x-y)g = (x+y+4)a_1 = 10a_1, \qquad \text{所以} \qquad a_1 = \frac{(x-y)g}{10} = \frac{(6-2y)g}{10},$$

其中用了 $x = 6-y$。這個值是正數，因為 $y < x$ 表示 $y < 3$。

方塊由靜止出發，以恆定加速度運動，在時間 $t_1$ 內走過距離 $d$，並達到速率 $V$，其中

$$d = \tfrac{1}{2}a_1 t_1^2, \qquad V^2 = 2a_1 d.$$

因此

$$t_1 = \sqrt{\frac{2d}{a_1}} = \sqrt{\frac{20d}{(6-2y)g}} = \sqrt{\frac{d}{5g}}\,\frac{10}{\sqrt{6-2y}}, \qquad V = \sqrt{\frac{(6-2y)gd}{5}}.$$

**第二階段：剪斷之後**

現在 $T_1 = 0$。方塊仍然以速率 $V$ 向 $P$ 移動，質量為 $y$ 的質點也仍然以速率 $V$ 上升。只有跨過 $Q$ 的繩子作用於方塊，並把它拉回 $Q$ 的方向。設減速度（deceleration）為 $a_2$。於是

$$\begin{aligned} T_2 &= 4a_2 \quad \text{（方塊）}, \\ yg - T_2 &= ya_2 \quad \text{（質點 } y\text{，上升時減速）}, \end{aligned}$$

所以

$$a_2 = \frac{yg}{4+y}.$$

? **想一想** 我們怎樣知道跨過 $Q$ 的繩子在剪斷之後仍然保持拉緊？

假如繩子鬆弛了，質點便會以 $g$ 減速，而方塊則會保持原有速率。繩子拉緊時，張力為 $T_2 = 4a_2 = 4yg/(4+y) > 0$，並無矛盾，而且共同的減速度 $a_2$ 小於 $g$。所以繩子確實保持拉緊，方塊和質點一起運動，直至停下為止。

方塊再經過以下時間便會靜止：

$$\begin{aligned} t_2 = \frac{V}{a_2} &= \sqrt{\frac{(6-2y)gd}{5}} \cdot \frac{4+y}{yg} \\ &= \sqrt{\frac{d}{5g}}\,\sqrt{6-2y}\left(1+\frac{4}{y}\right), \end{aligned}$$

其中我們用了 $\sqrt{(6-2y)gd/5}\,/\,g = \sqrt{d/(5g)}\sqrt{6-2y}$ 和 $(4+y)/y = 1+4/y$。

**總時間**

方塊在第二階段結束時首次回復靜止，所以總時間為

$$t_1 + t_2 = \sqrt{\frac{d}{5g}}\left[\frac{10}{\sqrt{6-2y}} + \left(1+\frac{4}{y}\right)\sqrt{6-2y}\right] = \sqrt{\frac{d}{5g}}\,\mathrm{f}(y),$$

正是所求。

**f 的駐值**

符合物理情況的範圍是 $0 < y < 3$。設 $u = 6-2y$，則 $\mathrm{d}u/\mathrm{d}y = -2$，而

$$\mathrm{f}(y) = 10u^{-1/2} + \left(1+\frac{4}{y}\right)u^{1/2}.$$

求導，並對第二項使用積法則（product rule），得

$$\begin{aligned} \mathrm{f}'(y) &= 10\cdot\left(-\tfrac{1}{2}\right)u^{-3/2}(-2) - \frac{4}{y^2}u^{1/2} + \left(1+\frac{4}{y}\right)\tfrac{1}{2}u^{-1/2}(-2) \\ &= \frac{10}{u^{3/2}} - \frac{4u^{1/2}}{y^2} - \frac{y+4}{y\,u^{1/2}}. \end{aligned}$$

兩邊乘以 $y^2u^{3/2}$（在 $0<y<3$ 時為正數），條件 $\mathrm{f}'(y)=0$ 便化為

$$10y^2 - 4u^2 - (y^2+4y)u = 0.$$

代入 $u = 6-2y$：

$$\begin{aligned} 10y^2 - 4(36-24y+4y^2) - (6y^2 - 2y^3 + 24y - 8y^2) &= 0 \\ 2y^3 - 4y^2 + 72y - 144 &= 0 \\ 2(y-2)(y^2+36) &= 0. \end{aligned}$$

由於 $y^2+36>0$，唯一的解是

$$y = 2,$$

它位於 $0<y<3$ 之內（並給出 $x=4$）。

因為當 $y\to 0^+$ 和 $y\to 3^-$ 時 $\mathrm{f}(y)\to\infty$，這個唯一的駐點（stationary point）就是極小值：$\mathrm{f}(2) = 10/\sqrt{2} + 3\sqrt{2} = 8\sqrt{2}$，所以回復靜止所需的最短時間是 $8\sqrt{2d/(5g)}$。

## 驗算

- 以數值方法逐步模擬運動（時間步長 $2\times10^{-6}\,\mathrm{s}$），隨機選取 $y\in(0,3)$ 和 $d$，所得結果與 $\sqrt{d/(5g)}\,\mathrm{f}(y)$ 吻合至五位有效數字；例如取 $y=1.671$、$d=0.581$，並用 $g=9.8$，兩種方法都得出 $1.27096\,\mathrm{s}$。
- f 的數值導數在 $(0,3)$ 上恰好變號一次，就在 $y=2$，而 $\mathrm{f}(2)=11.3137\ldots=8\sqrt{2}$。
- 極限情況合乎情理：當 $y\to3$ 時，兩邊的質量幾乎平衡，$a_1\to0$，第一階段需時無限長；當 $y\to0$ 時，剪斷之後沒有任何東西令方塊減速，所以第二階段需時無限長。

## 常見錯誤

- 在第二階段使用質量 $4+x+y$：繩子一經剪斷，質量為 $x$ 的質點便不再屬於這個系統。
- 忘記質量為 $y$ 的質點在剪斷之後仍以速率 $V$ 向上運動，因而把第二階段當作由靜止開始。
- 對 $\sqrt{6-2y}$ 求導時漏掉 $\mathrm{d}u/\mathrm{d}y=-2$。

## STEP 之外

把運動分成若干階段，再以速度的連續性把它們連接起來，正是分段定義的動力學（piecewise-defined dynamics）背後的基本想法；這個想法會在衝力（impulsive forces）以及控制問題中再次出現，後者研究的是某個力何時開啟、何時關閉。問哪一個 $y$ 值令時間最短，則是初嘗力學中的最優化（optimisation），日後會發展成變分法（calculus of variations），最速降線（brachistochrone）就是經典例子。Kleppner 與 Kolenkow 的 *An Introduction to Mechanics*（力學入門教科書）對連接體和約束力有仔細的討論。

詳細解答

# STEP I 2006，第 10 題

力學 • 拋體運動

**STEP I 2006，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/06-s1-q10

## 關鍵思路

軌跡（trajectory）是一條拋物線（parabola），屋頂是一條直線，所以「剛好碰到」的意思就是這條直線是一條切線（tangent）：把其中一條方程代入另一條，會得到一條關於 $x$ 的二次方程（quadratic），而且它有重根（repeated root）。至於正負號的問題，只要問這個重根位於何處便可解決。在最後一部分，兩個質點只可能在 $P$ 碰到屋頂的那一點相遇，所以 $Q$ 必須恰好在 $P$ 到達該點的那一刻到達；比較兩者在該時刻的水平位置，便得出結果。

## 解答

### 軌跡

把 $P$ 看成一個只受重力作用而自由運動的質點（particle），不計空氣阻力。它的初速度在水平和垂直方向的分量都是 $V\cos 45^\circ = V/\sqrt{2}$，所以在時間 $t$

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

利用 $t = \sqrt{2}\,x/V$ 消去時間，得 $\frac{1}{2}gt^2 = gx^2/V^2$，所以軌跡是

$$y = x - \frac{gx^2}{V^2}.$$

### 碰到屋頂

$P$ 與屋頂相遇之處滿足 $x - gx^2/V^2 = x\tan\alpha + b$，即

$$\frac{g}{V^2}\,x^2 + (\tan\alpha - 1)\,x + b = 0. \qquad (*)$$

相切（而不是穿過）表示 $(*)$ 有重根，所以它的判別式（discriminant）等於零：

$$(\tan\alpha - 1)^2 = \frac{4gb}{V^2} \quad \Longrightarrow \quad V(\tan\alpha - 1) = \pm 2\sqrt{bg}.$$

**想一想** 兩個正負號都滿足判別式的條件。怎樣分辨該取哪一個？

$(*)$ 的重根是

$$x_0 = -\frac{(\tan\alpha - 1)V^2}{2g} = \frac{(1 - \tan\alpha)V^2}{2g}.$$

質點沿 $x$ 遞增的方向運動（因為 $x = Vt/\sqrt{2}$），所以它只可能在 $x_0 > 0$ 之處碰到屋頂。這就迫使 $\tan\alpha - 1 < 0$，而由於 $V > 0$，左邊的 $V(\tan\alpha - 1)$ 是負數。因此

$$V(-1 + \tan\alpha) = -2\sqrt{bg},$$

這正是所要證明的。（若取正根，切點便會在 $x_0 < 0$ 之處，即位於拋物線向後延伸的部分，而 $P$ 從不經過那裏。）從幾何上看，屋頂的斜度必須小於初始運動方向的斜度，即 $\alpha < 45°$；否則屋頂一直在直線 $y = x$ 之上，而拋物線在 $x > 0$ 時位於該直線之下。

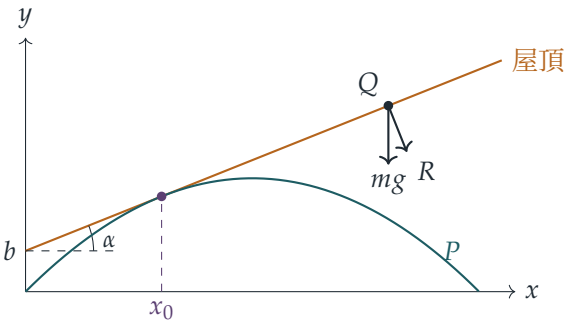
### 相切的時刻

由於 $x = Vt/\sqrt{2}$，相切的時刻是

$$T = \frac{\sqrt{2}\,x_0}{V} = \frac{\sqrt{2}\,(1 - \tan\alpha)V}{2g} = \frac{V(1 - \tan\alpha)}{\sqrt{2}\,g}.$$

### 路軌上的質點

把 $Q$ 看成一個在光滑（smooth）路軌上的質點，路軌沿直線 $y = x\tan\alpha + b$；它從 $(0, b)$ 沿路軌向 $x$ 遞增的方向射出（必須如此，才能到達 $P$，因為後者的 $x$ 永遠是正數）。作用於 $Q$ 的力有它的重量 $mg$，以及路軌對它的法向反作用力（normal reaction）$R$，後者垂直於路軌。下圖顯示了這些力，以及 $P$ 的路徑在 $x = x_0$ 處與屋頂相切的情況。



由於沒有摩擦力，沿路軌向斜上方分解，得

$$m\ddot{s} = -mg\sin\alpha,$$

其中 $s$ 是 $Q$ 沿路軌與 $(0,b)$ 的距離。以 $s=0$ 及 $\dot{s}=U$ 作為 $t=0$ 時的初始條件，

$$s = Ut - \tfrac{1}{2}gt^2\sin\alpha, \qquad x_Q = s\cos\alpha.$$

$P$ 除了在切點之外從不在屋頂上，而 $Q$ 則一直在屋頂上，所以兩個質點要相碰，唯一的可能是 $Q$ 在時間 $T$ 位於該點。那時兩個質點都在屋頂所在的直線上，所以只需它們的 $x$ 坐標在時間 $T$ 相等：

$$\left(UT - \tfrac{1}{2}gT^2\sin\alpha\right)\cos\alpha = \frac{VT}{\sqrt{2}}.$$

兩邊除以 $T>0$，並代入 $gT = V(1-\tan\alpha)/\sqrt{2}$：

$$U\cos\alpha = \frac{V}{\sqrt{2}} + \tfrac{1}{2}\sin\alpha\cos\alpha\cdot\frac{V(1-\tan\alpha)}{\sqrt{2}}.$$

兩邊乘以 $2\sqrt{2}$，並利用 $\sin\alpha\cos\alpha\tan\alpha = \sin^2\alpha$，

$$\begin{aligned} 2\sqrt{2}\,U\cos\alpha &= 2V + V\sin\alpha\cos\alpha\,(1-\tan\alpha) \\ &= V\left(2 + \sin\alpha\cos\alpha - \sin^2\alpha\right), \end{aligned}$$

這就是所要求的結果。

## 驗算

- 單位：$V(1-\tan\alpha)/g$ 是速率除以加速度，所以是時間；而 $\sqrt{bg}$ 是速率，與 $V(\tan\alpha - 1)$ 一樣。
- 極限情況 $\alpha \to 0$（屋頂水平）：$b = V^2/(4g)$ 正是大家熟悉的最大高度 $(V\sin 45^\circ)^2/(2g)$，在 $T = V/(\sqrt{2}g)$ 時到達；而 $Q$ 沿路軌方向不受力，所以它的速率必須與 $P$ 的水平速率相同，即 $U = V/\sqrt{2}$。公式給出的正是這個結果。
- 數值驗算：取 $g=9.8$、$V=10$ 及 $\alpha = 0.3$ 弧度，切點位於 $x_0 \approx 3.524$。讓 $Q$ 以推導所得的 $U$ 沿路軌逐步前進，對隨機選取的 $V$ 和 $\alpha < \pi/4$，它在時間 $T$ 的位置與 $P$ 的位置相差不超過 $10^{-6}$，而 $P$ 在其他所有時刻都在屋頂之下。

## 常見錯誤

- 把 $Q$ 的加速度寫成 $-g$ 或 $-g\cos\alpha$：只有重力沿路軌方向的分量 $g\sin\alpha$ 才沿路軌起作用。

- 把 $U$ 當作水平速率：$U$ 是沿路軌的速率，所以 $Q$ 的水平位移是 $s\cos\alpha$。
- 只用「速率是正數」來解釋為何取負根。真正起決定作用的是 $x_0$ 的正負號（等價於 $\tan\alpha < 1$）。

## STEP 之外

相切條件是一個普遍想法的特殊情況：一族曲線所能到達的區域，其邊界就是這族曲線的包絡線（envelope），而求包絡線，歸根究底就是問一條二次方程何時有重根。對於以固定速率向各個方向拋出的拋體（projectile），包絡線就是所謂的「安全拋物線」（parabola of safety）。$Q$ 的運動是最簡單的約束運動（constrained motion）；大學一年級的課程會用拉格朗日力學（Lagrangian mechanics）有系統地處理約束。Kleppner and Kolenkow, *An Introduction to Mechanics*（力學入門教科書）對拋體和約束運動都有細緻的講解。

詳細解答

# STEP I 2006，第 11 題

力學 • 彈性碰撞

**STEP I 2006，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/06-s1-q11

## 關鍵思路

我們無法逐次追蹤每一次碰撞：我們連 $A_1, \ldots, A_{n-2}$ 的質量也不知道。因此我們只考慮最終狀態。有兩項事實限制着最終狀態。第一，動量（momentum）和動能（kinetic energy）在結束時與開始時相同。第二，所有碰撞結束之後，沒有任何粒子能追上它前面的粒子，所以最終速度沿着槽（弱）遞增。於是每一部分都只需一段簡短的論證，說明這兩項事實容許哪些最終狀態。

## 解答

### 建立模型

我們把各粒子視為在一條直線上無摩擦地運動的質點，而每次碰撞都是完全彈性（perfectly elastic）的。取 $A_0$ 初始運動的方向為正方向，並設 $v_k$ 為 $A_k$ 的最終速度（$k = 0, 1, \ldots, n$）。以 $m_k$ 表示 $A_k$ 的質量，所以 $m_0 = m_{n-1} = m$ 而 $m_n = \lambda m$。由守恆（conservation）定律得

$$\sum_{k=0}^{n} m_k v_k = mu, \tag{M}$$

$$\sum_{k=0}^{n} m_k v_k^2 = mu^2. \tag{E}$$

粒子在槽內不能互相越過，所以它們一直保持 $A_0, A_1, \ldots, A_n$ 的次序。假如最終對某個 $k$ 有 $v_k > v_{k+1}$，那麼 $A_k$ 便會追上 $A_{k+1}$，於是還會再有一次碰撞。因此在所有碰撞結束之後

$$v_0 \leq v_1 \leq \cdots \leq v_n. \tag{O}$$

特別地，向前運動的粒子前面不能有靜止的粒子，而向後運動的粒子後面也不能有靜止的粒子。

**第 (i) 部分**

假設最終只有 $A_j$ 在運動，速度為 $v$。那麼由 (M) 和 (E) 得 $m_j v = mu$ 及 $m_j v^2 = mu^2$。兩式相除，得 $v = u$，從而 $m_j = m$。

由於 $v = u > 0$，次序條件 (O) 要求 $A_j$ 前面每一個粒子的速度都至少為 $u$。但它們全部靜止，所以前面根本沒有粒子，即 $A_j = A_n$。然而 $A_n$ 的質量是 $\lambda m \neq m$，因為 $\lambda > 1$。這是一個矛盾（contradiction），所以最終不可能恰好只有一個粒子在運動。

**第 (ii) 部分**

假設只有 $A_{n-1}$ 和 $A_n$ 在運動，速度分別為 $p$ 和 $q$。把 (M) 和 (E) 除以 $m$，

$$p + \lambda q = u, \qquad p^2 + \lambda q^2 = u^2.$$

把 $p = u - \lambda q$ 代入第二條方程，得 $-2\lambda uq + \lambda^2 q^2 + \lambda q^2 = 0$，即

$$\lambda q\big((1+\lambda)q - 2u\big) = 0.$$

因為 $A_n$ 在運動，我們需要 $q \neq 0$，所以

$$q = \frac{2u}{1+\lambda}, \qquad p = u - \frac{2\lambda u}{1+\lambda} = -\frac{(\lambda-1)u}{1+\lambda} < 0.$$

所以 $A_{n-1}$ 會向後運動。但 $n \geq 2$，所以它後面有一個粒子 $A_{n-2}$，而按假設該粒子是靜止的。於是 $v_{n-2} = 0 > p = v_{n-1}$，這與 (O) 矛盾：$A_{n-1}$ 會撞上 $A_{n-2}$。所以這個最終狀態同樣不可能出現。

**第 (iii) 部分**

這裏的主要工具是以下事實。

**引理（lemma）**。在運動結束時，必有某個粒子向後運動。

證明。假設相反，即對每個 $k$ 都有 $v_k \geq 0$。由 (O)，$v_n$ 是最大的速度。(E) 中的每一項都滿足 $m_k v_k^2 \leq v_n\, m_k v_k$，因為 $0 \leq v_k \leq v_n$ 而 $m_k > 0$。把這些不等式相加並利用 (M)，

$$mu^2 = \sum m_k v_k^2 \leq v_n \sum m_k v_k = v_n\, mu, \qquad \text{所以} \qquad v_n \geq u.$$

另一方面，(M) 中的每一項都是非負的，所以單獨一項 $\lambda m v_n$ 不會超過總和：

$$\lambda m v_n \leq mu, \qquad \text{所以} \qquad v_n \leq \frac{u}{\lambda} < u,$$

這裏用到 $\lambda > 1$ 和 $u > 0$。這兩個界互相矛盾。 □

**想一想** 題目從來沒有給出 $A_1, \ldots, A_{n-2}$ 的質量。為甚麼這並不要緊？

引理只用到 (M)、(E)、(O) 以及 $A_n$ 的質量，所以無論其他質量是多少，它都成立。正因如此，這道題才有可能解答。

現在假設最終只有 $A_{n-2}$、$A_{n-1}$ 和 $A_n$ 在運動。

若 $n \geq 3$，則 $A_{n-2}$ 後面有一個靜止的粒子 $A_{n-3}$。由 (O)，$v_{n-2} \geq v_{n-3} = 0$，從而 $v_{n-1}, v_n \geq v_{n-2} \geq 0$。每個粒子的速度都 $\geq 0$，而這是引理所不容許的。

若 $n = 2$，這三個粒子就是 $A_0$、$A_1$、$A_2$，亦即全部粒子，而整個運動是完全確定的。當 $A_0$ 撞上 $A_1$ 時（兩者質量相等），速度互相交換：$A_0$ 停下，而 $A_1$ 以速率 $u$ 繼續前進。然後 $A_1$ 撞上 $A_2$，按第 (ii) 部分的計算（碰撞前的速度為 $u$ 和 0），它以速度 $-\frac{(\lambda-1)u}{\lambda+1}$ 反彈，而 $A_2$ 則以 $\frac{2u}{\lambda+1}$ 離開。最後 $A_1$ 撞上靜止的 $A_0$，質量相等的兩者再次交換速度。此後，$A_0$ 向後運動，$A_1$ 靜止，$A_2$ 向前運動，所以 (O) 成立，不會再有碰撞。$A_1$ 最終靜止，所以三個粒子全部都在運動同樣是不可能的。

### 第 (iv) 部分

假設最終恰好有兩個粒子在運動。由 (M)，總動量為 $mu > 0$，所以其中至少一個向前運動。由引理，至少一個向後運動。所以其中一個，設為 $A_i$，有 $v_i < 0$，而另一個 $A_j$ 有 $v_j > 0$。

**想一想** 這兩個粒子可以是哪兩個？

由 (O)，$A_i$ 後面每一個粒子的速度都 $\leq v_i < 0$，所以它在運動。但另外在運動的只有 $A_j$，而它是向前運動的，所以 $A_i$ 後面沒有任何粒子：$A_i = A_0$。同樣地，$A_j$ 前面每一個粒子的速度都 $\geq v_j > 0$，所以前面沒有粒子，即 $A_j = A_n$。

所以在運動的是 $A_0$（質量 $m$）和 $A_n$（質量 $\lambda m$），速度分別為 $a < 0$ 和 $b > 0$。方程 (M) 和 (E) 成為

$$a + \lambda b = u, \qquad a^2 + \lambda b^2 = u^2,$$

與第 (ii) 部分的方程完全一樣。由於 $b \neq 0$，

$$b = \frac{2u}{\lambda+1}, \qquad a = -\frac{(\lambda-1)u}{\lambda+1}.$$

因此，在運動的兩個粒子是 $A_0$，以速率 $\dfrac{(\lambda-1)u}{\lambda+1}$ 向後運動，以及 $A_n$，以速率 $\dfrac{2u}{\lambda+1}$ 向前運動。

這種情況確實可以發生：第 (iii) 部分中 $n = 2$ 的情形正是以這個狀態告終。

## 驗算

- 取 $\lambda = 3$ 和 $u = 4$，兩個速率是 2 和 2。於是動量為 $m(-2) + 3m(2) = 4m$，能量為 $\frac{1}{2}m(4) + \frac{1}{2}(3m)(4) = 8m = \frac{1}{2}m(16)$，與應有的結果相符。
- 當 $\lambda \to 1$ 時，$A_0$ 的最終速率趨向 0，而 $A_n$ 的最終速率趨向 $u$：當所有質量都相等時，速度只是沿着整列粒子逐一傳遞下去，就像牛頓擺（Newton's cradle）一樣。
- 我們對 60,000 個隨機設定模擬了整個碰撞過程（$n$ 由 2 至 6，$A_1, \ldots, A_{n-2}$ 的質量隨機，$\lambda$ 和間距亦隨機）。$A_0$ 最終總是向後運動，(i)–(iii) 中的最終狀態從未出現，而每當恰好有兩個粒子在運動時，它們都是 $A_0$ 和 $A_n$，速率亦與上文所得相同。

## 常見錯誤

- 試圖追蹤每一次碰撞。中間各粒子的質量是未知的，所以只能分析最終狀態。
- 忘記次序條件 (O)。單憑守恆定律，第 (ii) 部分的狀態是容許的；把它排除的正是次序條件。
- 假定 $A_{n-2}$ 不是 $A_0$。當 $n = 2$ 時它正是，而第 (iii) 部分需要另作一個（簡短的）論證。

## STEP 之外

這個引理是兩個守恆量之間的一個不等式：當所有速度都非負時，能量不能超過（最大速度）$\times$（動量）。這類論證利用守恆量和次序關係來排除某些最終狀態，而無須解出整個動力學過程，在多粒子系統的研究中十分典型。下一步不妨閱讀 Kleppner 與 Kolenkow 的 *An Introduction to Mechanics*（力學導論），書中對碰撞、質心參考系（centre-of-mass frame）和守恆定律都有細緻的處理。

詳細解答

# STEP I 2006，第 12 題

概率與統計 • 條件概率

**STEP I 2006，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/06-s1-q12

## 關鍵思路

一條道路要暢通，必須路上每一段都暢通；而兩鎮要被隔絕，則必須每一條道路都受阻。因此我們逐條道路處理，利用獨立性（independence），並且每一步都借助補事件（complement）來計算。第二部分是一個條件概率（conditional probability）問題，當中選了哪條路是有影響的：只有一段的道路比有兩段的道路更可能暢通，所以發現自己所選的路暢通，其實透露了你選的是哪條路的一些訊息。我們用定義 $\mathrm{P}(B \mid A) = \mathrm{P}(A \cap B)/\mathrm{P}(A)$ 來處理，並按所選的道路分情況計算。

## 解答

### 被隔絕的概率

把兩條有兩段的道路稱為 $R_1$ 和 $R_2$，只有一段的道路稱為 $R_3$。設 $q = 1 - p$ 為單獨一段暢通的概率。

一條有兩段的道路暢通，當且僅當它的兩段都暢通；由獨立性，其概率為 $q^2 = (1-p)^2$。所以

$$\mathrm{P}(R_1 \text{ 受阻}) = \mathrm{P}(R_2 \text{ 受阻}) = 1 - (1-p)^2 = p(2-p),$$
$$\mathrm{P}(R_3 \text{ 受阻}) = p.$$

三條道路共用到五個不同的路段，所以「$R_1$ 受阻」、「$R_2$ 受阻」和「$R_3$ 受阻」這三個事件各自取決於互不相交的幾組獨立路段，因而互相獨立。Oxtown 與 Camville 被隔絕，當且僅當三條道路全部受阻，所以

$$\mathrm{P}(\text{被隔絕}) = p(2-p) \cdot p(2-p) \cdot p = p^3(2-p)^2,$$

正如所求。

**條件概率**

設 $C_k$ 為我選擇道路 $R_k$ 的事件，則對每個 $k$ 都有 $\mathrm{P}(C_k) = \frac{1}{3}$，而且我的選擇與水浸情況獨立。設 $O$ 為我所選的路暢通的事件，$B$ 為另外兩條路都受阻的事件。我們要求的是 $\mathrm{P}(B \mid O) = \mathrm{P}(B \cap O)/\mathrm{P}(O)$。

> **?** **想一想** 根據獨立性，答案不就是 P(另外兩條路受阻) 嗎？

假如我們知道自己選了哪條路，答案確是如此。但「另外兩條路」是哪兩條，要視乎我們的選擇而定；而且觀察到自己的路暢通之後，選中較短的 $R_3$ 的可能性就比原先高了。所以我們必須按所選的道路分情況，並配上正確的權重，而下面的公式做的正是這件事。

**分母**。對所選的道路運用全概率公式（law of total probability），

$$\begin{aligned}\mathrm{P}(O) &= \tfrac{1}{3}(1-p)^2 + \tfrac{1}{3}(1-p)^2 + \tfrac{1}{3}(1-p) \\ &= \tfrac{1}{3}(1-p)\big[2(1-p)+1\big] = \tfrac{1}{3}(1-p)(3-2p).\end{aligned}$$

**分子**。如果我選了 $R_1$，就需要 $R_1$ 暢通而 $R_2$、$R_3$ 受阻；由各條道路的獨立性，其概率為 $(1-p)^2 \cdot p(2-p) \cdot p$。選 $R_2$ 的結果相同。如果我選了 $R_3$，就需要 $R_3$ 暢通而兩條有兩段的道路都受阻：$(1-p) \cdot p^2(2-p)^2$。因此

$$\begin{aligned}\mathrm{P}(B \cap O) &= \tfrac{1}{3}\Big[2p^2(2-p)(1-p)^2 + p^2(2-p)^2(1-p)\Big] \\ &= \tfrac{1}{3}\,p^2(1-p)(2-p)\big[2(1-p)+(2-p)\big] \\ &= \tfrac{1}{3}\,p^2(1-p)(2-p)(4-3p).\end{aligned}$$

**答案**。兩式相除，並約去因子 $\frac{1}{3}(1-p)$（當 $p < 1$ 時它不等於零），得

$$\mathrm{P}(B \mid O) = \frac{p^2(2-p)(4-3p)}{3-2p}.$$

因子 $p$、$2-p$、$4-3p$ 和 $3-2p$ 全都是一次式，所以這已經是完全因式分解的形式。

**極限 $p \to 1$**

當 $p \to 1$ 時，每個因子都趨向 1，所以

$$\mathrm{P}(B \mid O) \to \frac{1 \cdot 1 \cdot 1}{1} = 1.$$

這是合理的。當水浸幾乎必然發生時，我自己的路暢通純屬僥倖；而其他道路是否水浸與我的路獨立，它們幾乎肯定沒有同樣的運氣：知道我的路暢通，並不能說明它們是否

暢通。所以我應該幾乎肯定，換了任何其他選擇都會行不通。（在 $p = 1$ 本身，作為條件的事件概率為零，條件概率沒有定義；只有極限才有意義。）

## 驗算

- 當 $p = 0$ 時，沒有任何路段會受阻，而公式給出 0，正該如此。
- 當 $p = \frac{1}{2}$ 時，公式給出 $\frac{\frac{1}{4} \cdot \frac{3}{2} \cdot \frac{5}{2}}{2} = \frac{15}{32}$。把路段受阻的全部 $2^5$ 種情況和三種道路選擇逐一精確列舉，得到的數值相同；對其他幾個 $p$ 值也是如此，例如 $p = \frac{2}{3}$ 時得 $\frac{32}{45}$。
- 對 $0 \leq p < 1$，答案總是介乎 0 與 1 之間：例如當 $p = 0.99$ 時，它約為 0.9996。

## 常見錯誤

- 把有兩段的道路受阻的概率當作 $p^2$（兩段都受阻），而不是 $1 - (1 - p)^2$（至少一段受阻）。
- 忘記取條件：另外兩條路受阻的概率，與已知我隨機選中的路暢通時它的條件概率，並不相同。
- 在觀察之後仍把三個選擇看成權重相等；相等的是先驗概率（prior probability）$\frac{1}{3}$，而不是後驗概率（posterior probability）。

## STEP 之外

多條道路並聯、每條道路又由若干路段串聯而成，是可靠性網絡（reliability network）最簡單的例子；一個網絡能把兩點連通的概率，其一般情況是滲流理論（percolation theory）的研究對象，而當 $p$ 變化時網絡行為出現的急劇轉變，正是該理論的核心課題。第二部分其實是喬裝了的貝葉斯定理（Bayes' theorem）：「我的路暢通」這項觀察，更新了選中各條道路的概率。Grimmett 與 Welsh 合著的 *Probability: An Introduction*（概率論入門）正是以這種思路講解條件概率和獨立性。

詳細解答

# STEP I 2006，第 13 題

概率與統計 • 泊松混合分佈

**STEP I 2006，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/06-s1-q13

## 關鍵思路

鑽石以固定的平均率獨立地散佈，這就是一個泊松過程（Poisson process），因此 $m$ 公斤朱古力中的鑽石數目服從平均值為 $m$ 的泊松分佈（Poisson distribution）。朱古力的份量本身也是隨機的，所以我們以它為條件：先就固定的份量求出答案，再利用全概率定理（law of total probability）（以及它在期望值方面的對應版本）對擲骰結果取平均。這樣每一部分都歸結為級數求和：(i) 是有限級數，(ii) 是幾何級數（geometric series）。

## 解答

設 $D$ 為我找到的鑽石數目。由於鑽石是以每公斤平均一顆的比率獨立地藏在朱古力中，質量為 $m$ kg 的一排朱古力所含的鑽石數目服從平均值為 $m$ 的泊松分佈。特別地，已知質量為 $m$ kg 時，

$$\mathrm{P}(D=0)=\mathrm{e}^{-m}, \qquad \mathrm{E}(D)=m.$$

以下一律把 $100k$ 克寫成 $0.1k$ 公斤。

### 第 (i) 部分

設 $N$ 為骰子擲出的點數，則對 $n=1,\ldots,6$ 有 $\mathrm{P}(N=n)=\frac{1}{6}$。已知 $N=n$ 時，我買 $0.1n$ kg，所以 $D \mid N=n$ 服從平均值為 $0.1n$ 的泊松分佈。由全概率定理，

$$\begin{aligned}\mathrm{P}(D=0) &= \sum_{n=1}^{6} \mathrm{P}(N=n)\,\mathrm{P}(D=0 \mid N=n) = \frac{1}{6}\sum_{n=1}^{6} \mathrm{e}^{-0.1n} \\ &= \frac{1}{6}\left(\mathrm{e}^{-0.1}+\mathrm{e}^{-0.2}+\cdots+\mathrm{e}^{-0.6}\right).\end{aligned}$$

這是一個幾何級數，首項為 $\mathrm{e}^{-0.1}$，公比（common ratio）為 $\mathrm{e}^{-0.1}$，共有六項，所以

$$\mathrm{P}(D=0)=\frac{\mathrm{e}^{-0.1}}{6}\left(\frac{1-\mathrm{e}^{-0.6}}{1-\mathrm{e}^{-0.1}}\right),$$

正是所求。

至於期望值（expectation），我們用同樣的方法取條件：

$$\mathrm{E}(D)=\sum_{n=1}^{6}\mathrm{P}(N=n)\,\mathrm{E}(D\mid N=n)=\frac{1}{6}\sum_{n=1}^{6}0.1n=\frac{0.1\times 21}{6}=0.35.$$

換言之，$\mathrm{E}(D)=0.1\,\mathrm{E}(N)=0.1\times 3.5$，這正是我們要證明的結果。

**第 (ii) 部分**

現在設 $T$ 為首次擲出 6 點時的擲骰次數。各次擲骰互相獨立，而 $T=t$ 表示先擲出 $t-1$ 次非 6 點，然後擲出一次 6 點，所以

$$\mathrm{P}(T=t)=\left(\tfrac{5}{6}\right)^{t-1}\tfrac{1}{6},\qquad t=1,2,3,\ldots$$

已知 $T=t$ 時，我買 $0.1t$ kg，所以 $\mathrm{P}(D=0\mid T=t)=\mathrm{e}^{-0.1t}$。因此

$$\mathrm{P}(D=0)=\sum_{t=1}^{\infty}\left(\tfrac{5}{6}\right)^{t-1}\tfrac{1}{6}\,\mathrm{e}^{-0.1t}=\frac{\mathrm{e}^{-0.1}}{6}\sum_{t=1}^{\infty}\left(\tfrac{5}{6}\mathrm{e}^{-0.1}\right)^{t-1}.$$

? **想一想** 為甚麼我們可以把這個級數求和至無窮？

這是一個公比為 $r=\frac{5}{6}\mathrm{e}^{-0.1}$ 的幾何級數，而 $0<r<1$，所以它收斂（converge），其和為 $1/(1-r)$。於是

$$\mathrm{P}(D=0)=\frac{\mathrm{e}^{-0.1}}{6}\cdot\frac{1}{1-\frac{5}{6}\mathrm{e}^{-0.1}}=\frac{\mathrm{e}^{-0.1}}{6-5\mathrm{e}^{-0.1}},$$

其中分子和分母同乘以 6。這就是所求的結果。

至於鑽石數目的期望值，

$$\mathrm{E}(D)=\sum_{t=1}^{\infty}\mathrm{P}(T=t)\,\mathrm{E}(D\mid T=t)=\sum_{t=1}^{\infty}\left(\tfrac{5}{6}\right)^{t-1}\tfrac{1}{6}\times 0.1t=\frac{0.1}{6}\sum_{t=1}^{\infty}t\left(\tfrac{5}{6}\right)^{t-1}.$$

題目的提示在這裏派上用場。當 $|x|<1$ 時，由二項展開式（binomial expansion）可得

$$(1-x)^{-2}=1+2x+3x^2+4x^3+\cdots=\sum_{t=1}^{\infty}tx^{t-1}.$$

取 $x=\frac{5}{6}$，這個和就是 $\left(1-\frac{5}{6}\right)^{-2}=36$，所以

$$\mathrm{E}(D)=\frac{0.1}{6}\times 36=0.6.$$

這同樣等於 $0.1\,\mathrm{E}(T)$，其中 $\mathrm{E}(T)=6$ 就是大家熟悉的、擲出一次 6 點所需擲骰次數的平均值。

## 驗算

- 按數值計算，(i) 的概率約為 0.7150，(ii) 的概率約為 0.6131。以泊松分佈產生鑽石數目，模擬 $2 \times 10^5$ 位顧客，所得的結果為 0.7151 和 0.6136，而鑽石數目的平均值則為 0.350 和 0.599。
- (ii) 的概率比 (i) 的小。這是合理的：我平均買 600 g，而不是 350 g。
- 兩個期望值都是那個隨機乘數的平均值的 0.1 倍。鑽石數目的期望值等於公斤數的期望值，因為已知質量時，鑽石數目的期望值是質量的線性函數。

## 常見錯誤

- 以 $\mathrm{e}^{-0.35}$（或 $\mathrm{e}^{-0.6}$）作為沒有鑽石的概率。這樣只用到平均質量，但 $\mathrm{e}^{-m}$ 並不是 $m$ 的線性函數，所以必須先取 $\mathrm{e}^{-m}$ 再求平均，而不是先求平均。期望值是線性的，所以這條捷徑對 $\mathrm{E}(D)$ 確實行得通。
- 忘記把克換算成公斤，結果得到平均值 $100n$ 而不是 $0.1n$。
- 寫成 $\mathrm{P}(T = t) = (\frac{5}{6})^t \frac{1}{6}$，這樣便多算了一次非 6 點。

## STEP 之外

這裏泊松分佈的參數本身是隨機的，這使 $D$ 成為一個混合（mixture）分佈。所用的工具，即全概率定理和重期望定律（tower law）$\mathrm{E}(D) = \mathrm{E}\big(\mathrm{E}(D \mid T)\big)$，都是大學一年級概率論的核心內容。概率母函數（probability generating function）可以俐落地處理這類問題：$\mathrm{P}(D = 0)$ 就是 $T$ 的母函數在 $\mathrm{e}^{-0.1}$ 處的值。先定出隨機的試驗次數，再在每次試驗中取一個隨機的數目，由此便可進一步引出複合分佈（compound distribution）和分支過程（branching process）。Grimmett 與 Welsh 合著的 *Probability: An Introduction*（概率論入門教材）對以上各項都有詳細闡述。

詳細解答

# STEP I 2006，第 14 題

概率與統計 • 不放回抽樣

**STEP I 2006，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/06-s1-q14

## 關鍵思路

在第 $r$ 次才取到紅色糖果，即是先取到 $r-1$ 顆藍色糖果，然後取到紅色那一顆。若每次抽取後放回（with replacement），每次抽取的情況都相同，所以概率只是幾個冪的簡單乘積；之後把 $n$ 當作連續變量，對其對數求導，便可求出使概率最大的 $n$。若抽取後不放回（without replacement），最簡潔的看法是想像把所有糖果都抽出來：紅色那一顆落在任何一個位置的可能性都相等。

## 解答

以下 $n \geq 0$ 均為整數，而 $R$ 表示首次取到紅色糖果時是第幾次抽取。

### 第 (i) 部分

由於糖果會放回袋中並搖勻，各次抽取互相獨立（independent），而每次抽取時，取到紅色糖果的概率是 $\frac{1}{n+1}$，取到藍色糖果的概率是 $\frac{n}{n+1}$。事件 $R=r$ 表示首 $r-1$ 次抽到藍色，而第 $r$ 次抽到紅色，所以根據獨立性，

$$P(R=r)=\left(\frac{n}{n+1}\right)^{r-1}\frac{1}{n+1}=\frac{n^{r-1}}{(n+1)^r}, \qquad r=1,2,3,\ldots$$

（這就是成功概率為 $\frac{1}{n+1}$ 的幾何分佈（geometric distribution）。）

現在固定 $r$，把上式看成 $n$ 的函數：

$$f(n)=\frac{n^{r-1}}{(n+1)^r}.$$

若 $r=1$，則 $f(n)=\frac{1}{n+1}$，它在 $n=0$ 時最大（沒有藍色糖果時，第一顆必定是紅色）。

因此假設 $r \geq 2$。這時 $f(0) = 0$，而當 $n > 0$ 時 $f(n) > 0$，所以我們可以在 $n > 0$ 上取對數，並把 $n$ 當作實變量：

$$\ln f(n) = (r-1)\ln n - r\ln(n+1), \qquad \frac{\mathrm{d}}{\mathrm{d}n}\ln f(n) = \frac{r-1}{n} - \frac{r}{n+1} = \frac{(r-1)-n}{n(n+1)}.$$

**想一想** $n$ 必須是整數。那麼求導是否合理？

合理，因為我們只是利用導數（derivative）來找出 $f$ 在哪裏遞增、在哪裏遞減。當 $n > 0$ 時，分母 $n(n+1)$ 是正數，所以導數在 $0 < n < r-1$ 時為正，在 $n > r-1$ 時為負。因此 $f$ 在 $[0, r-1]$ 上嚴格遞增（strictly increasing），在 $[r-1, \infty)$ 上嚴格遞減。由於 $r-1$ 本身是整數，它勝過其他所有整數，無須再比較相鄰的整數。

所以在所有情況下（包括 $r = 1$），使概率最大的是

$$n = r - 1,$$

最大值為 $\dfrac{(r-1)^{r-1}}{r^r}$（當 $r = 1$ 時理解為 1）。例如，當 $r = 3$ 時，最理想的袋中有 $n = 2$ 顆藍色糖果，所得概率為 $\frac{4}{27}$。

### 第 (ii) 部分

現在各次抽取不再獨立，但有兩個簡單的做法。

*直接相乘。*要 $R = r$，首 $r-1$ 次每次都要抽到藍色，然後抽到紅色。這只有在藍色糖果足夠多時才可能發生，即 $r - 1 \leq n$。已知之前抽到的全是藍色，第 $k$ 次抽取時袋中有 $n + 2 - k$ 顆糖果，其中一顆是紅色。所以

$$P(R = r) = \frac{n}{n+1} \cdot \frac{n-1}{n} \cdot \frac{n-2}{n-1} \cdots \frac{n-r+2}{n-r+3} \cdot \frac{1}{n-r+2}.$$

每個分子都與下一個分母約去，乘積逐項相消（telescope），得

$$P(R = r) = \begin{cases} \dfrac{1}{n+1}, & 1 \leq r \leq n+1, \\ 0, & r > n+1. \end{cases}$$

*對稱性。*抽到紅色糖果便停止，並不影響它所在的位置。所以不妨想像一直抽下去，直至袋中空無一物。這些（可區分的）糖果共有 $(n+1)!$ 種排列次序，每種的可能性都相等，因此紅色糖果落在 $n+1$ 個位置中任何一個的可能性都相等。這樣得出相同的答案。

至於求最大值，先固定 $r$。若 $n < r-1$，概率為 0。若 $n \geq r-1$，概率為 $\frac{1}{n+1}$，它隨 $n$ 增大而嚴格遞減。所以使概率最大的，是符合條件的最小的袋，

$$n = r - 1,$$

而最大概率為 $\frac{1}{r}$。（當 $n = r-1$ 時，紅色糖果是 $r$ 顆糖果中任何一顆的可能性都相等，而最後才抽到它，正是 $r$ 個等可能結果的其中一個。）

## 驗算

- 在 (i) 中，$\sum_{r\geq 1} P(R=r) = \frac{1}{n+1}\cdot\dfrac{1}{1-n/(n+1)} = 1$；在 (ii) 中，$n+1$ 個相等的值 $\frac{1}{n+1}$ 之和為 1。
- 在 (i) 中取 $r=3$：$n=1,2,3$ 分別給出 $\frac{1}{8}$、$\frac{4}{27}\approx 0.148$ 和 $\frac{9}{64}\approx 0.141$，所以 $n=2$ 勝出。
- 我們就 $n\leq 6$ 列出糖果的所有排列次序，驗證了 (ii)。就 (i) 取 $n=3$、$r=2$ 進行模擬，得出 0.188，而準確值為 $\frac{3}{16}=0.1875$。

## 常見錯誤

- 在 (i) 中寫成 $\left(\frac{n}{n+1}\right)^r\frac{1}{n+1}$：在紅色之前只有 $r-1$ 顆藍色。
- 在 (ii) 中，忘記了當 $r>n+1$ 時答案是 0。正是這一點，令「把 $n$ 取得愈小愈好」不致一直滑落到 $n=0$。

## STEP 之外

這兩個求最大值的問題，其實都是最大似然估計（maximum likelihood estimation）。假如你看到紅色糖果在第 $r$ 次首次出現，卻不知道 $n$ 是多少，那麼在兩個實驗中，令你的觀察結果最有可能出現的 $n$ 值都是 $n=r-1$。這是大學一年級統計學的基石之一。(ii) 中的對稱性論證是可交換性（exchangeability）的一個例子，它令許多不放回抽樣的問題變得容易。這兩個概念在 Blitzstein 與 Hwang 所著的 *Introduction to Probability*（概率論入門教科書）中都有詳細闡述。

詳細解答

# STEP II 2006，第 1 題

純數學 • 周期性遞推關係

> **STEP II 2006，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/06-s2-q1

## 關鍵思路

由於每一項都決定了下一項，只要 $u_{p+1} = u_1 = 2$，數列（sequence）便會以週期（period）$p$ 重複。因此 (i) 的每一部分，歸根究底都是把頭幾項寫成 $k$ 的函數，然後解一條多項式方程。週期為 4 的方程必定包含週期 1 和週期 2 的答案作為因式（週期為 1 或 2 的數列同樣滿足 $u_5 = u_1$），事先知道這一點，因式分解（factorisation）便輕而易舉。第 (ii) 部分是一行便可完成的歸納法（induction），之後再用慣常的「在遞推關係（recurrence）中令 $n \to \infty$」的論證。

## 解答

### 第 (i) 部分

設 $f(u) = k - \dfrac{36}{u}$，則 $u_{n+1} = f(u_n)$。若 $u_{p+1} = u_1$，則反覆應用 $f$ 可得 $u_{p+2} = u_2$、$u_{p+3} = u_3$，如此類推，所以對每個 $n$ 都有 $u_{n+p} = u_n$。反過來說，週期為 $p$ 當然需要 $u_{p+1} = u_1 = 2$。因此，數列的週期恰好為 $p$，當且僅當 $u_{p+1} = 2$（而且各項均有定義，即永不為 0），並且沒有更小的週期符合條件。

**(a)** 數列為常數數列，當且僅當 $u_2 = 2$，即 $k - 18 = 2$。所以 $k = 20$。

**(b)** 我們有 $u_2 = k - 18$，以及

$$u_3 = k - \frac{36}{k - 18} \qquad (k \neq 18).$$

令 $u_3 = 2$，得 $(k-2)(k-18) = 36$，所以 $k^2 - 20k = 0$，即 $k = 0$ 或 $k = 20$。$k = 20$ 給出常數數列，其週期為 1，而不是 2。當 $k = 0$ 時，數列為 2, −18, 2, −18, ...，週期確

實是 2。所以 $k = 0$。

**(c)** 我們還需要多計兩項。當 $k \neq 18$ 時，

$$u_3 = \frac{k(k-18)-36}{k-18} = \frac{k^2-18k-36}{k-18},$$

然後，只要 $k^2 - 18k - 36 \neq 0$，便有

$$u_4 = k - \frac{36(k-18)}{k^2-18k-36} = \frac{k^3-18k^2-72k+648}{k^2-18k-36}.$$

條件 $u_5 = 2$ 表示 $k - 36/u_4 = 2$，即 $u_4 = \dfrac{36}{k-2}$（這裏需要 $k \neq 2$；若 $k = 2$，則 $36/u_4 = 0$，這是不可能的）。交叉相乘，得

$$(k-2)(k^3-18k^2-72k+648) = 36(k^2-18k-36).$$

展開後把所有項移到左邊，

$$k^4-20k^3-36k^2+792k-1296 = 36k^2-648k-1296,$$
$$k^4-20k^3-72k^2+1440k = 0.$$

> **想一想** 怎樣才可以不必反覆試驗，就把這樣的四次多項式（quartic）分解因式？

動手之前，我們已經知道其中兩個根。任何週期為 1 或 2 的數列同樣滿足 $u_5 = u_1$，所以 $k = 20$ 和 $k = 0$ 必定是根。先抽出 $k$，再抽出 $k - 20$，

$$k^4-20k^3-72k^2+1440k = k(k-20)(k^2-72).$$

所以 $k \in \{0, 20, 6\sqrt{2}, -6\sqrt{2}\}$。0 和 20 分別給出週期 2 和 1，所以新的值是 $k = \pm 6\sqrt{2}$。最後，我們利用 $k^2 = 72$，驗證當 $k = \pm 6\sqrt{2}$ 時各項均有定義：

- $u_2 = k - 18 \neq 0$；
- $u_3$ 的分子是 $k^2 - 18k - 36 = 36 - 18k$，由於 $k \neq 2$，它不等於零；
- $u_4$ 的分子是 $k(k^2-72) - 18k^2 + 648 = -1296 + 648 = -648 \neq 0$；
- 按構造 $u_5 = 2$，而且 $k \neq 2$、$k \neq 18$，符合推導過程中的要求。

由於 $k \neq 0, 20$，週期不是 1 或 2；而一個滿足 $u_5 = u_1$ 的數列，若最小週期不是 1 或 2，其最小週期必定恰好是 4（最小週期必須整除 4）。所以數列的週期為 4，當且僅當 $k = \pm 6\sqrt{2}$。

### 第 (ii) 部分

取 $k = 37$。我們用歸納法證明 $u_n \geq 2$。當 $n = 1$ 時，由於 $u_1 = 2$，命題成立。若 $u_n \geq 2$，則 $u_n > 0$，所以 $0 < \dfrac{36}{u_n} \leq 18$，因此

$$u_{n+1} = 37 - \frac{36}{u_n} \geq 37 - 18 = 19 \geq 2.$$

所以對所有 $n$ 都有 $u_n \geq 2$，這正是我們要證明的。（事實上，對所有 $n \geq 2$ 都有 $u_n \geq 19$。）

> **想一想** 極限（limit）所滿足的方程有兩個根。哪一個才是正確的？又為甚麼可以在遞推關係中令 $n \to \infty$？

由於 $u_n \to \ell$ 而且每個 $u_n \geq 2$，我們有 $\ell \geq 2$；特別地，$\ell \neq 0$，所以 $36/u_n \to 36/\ell$。在 $u_{n+1} = 37 - 36/u_n$ 的兩邊令 $n \to \infty$（左邊同樣趨向 $\ell$），得

$$\ell = 37 - \frac{36}{\ell} \quad \Longrightarrow \quad \ell^2 - 37\ell + 36 = 0 \quad \Longrightarrow \quad (\ell - 1)(\ell - 36) = 0.$$

根 $\ell = 1$ 因 $\ell \geq 2$ 而被排除，所以

$$\ell = 36.$$

## 驗算

- 當 $k = 6\sqrt{2} \approx 8.485$ 時，各項為 2, −9.515, 12.269, 5.551, 2, . . . ；當 $k = -6\sqrt{2}$ 時，各項為 2, −26.485, −7.126, −3.433, 2, . . . 。兩者都在四步之後重複。
- 因式 $k^2 - 72$ 有一個簡潔的含義：映射 $u \mapsto (ku - 36)/u$ 對應矩陣 $\begin{pmatrix} k & -36 \\ 1 & 0 \end{pmatrix}$，其跡（trace）為 $k$，行列式（determinant）為 36。它的特徵值（eigenvalues）是 $6e^{\pm i\theta}$，其中 $k = 12\cos\theta$；當 $\theta = \pi/4$ 或 $3\pi/4$ 時，它的四次方是單位矩陣（identity）的倍數。這給出 $k = \pm 6\sqrt{2}$，而 $\theta = \pi/2$ 則給出週期為 2 的 $k = 0$。
- 當 $k = 37$ 時：$u_2 = 19$、$u_3 \approx 35.105$、$u_4 \approx 35.975$、$u_5 \approx 35.9993$，各項明顯趨近 36，而不是 1。

## 常見錯誤

- 在 (b) 中答 $k = 0, 20$，或在 (c) 中把四個根全部寫上。題目問的是週期 2 和週期 4，而按這個意思，常數數列的週期並不是 2。
- 除了 $\ell = 36$ 之外，把 $\ell = 1$ 也當作答案。(ii) 前半部分得出的界 $u_n \geq 2$，正是用來排除它的。

► 令 $n \to \infty$ 時沒有指出 $\ell \neq 0$，而這是 $36/u_n \to 36/\ell$ 成立所必需的。

## STEP 之外

映射 $u \mapsto (ku - 36)/u$ 是一個莫比烏斯變換（Möbius transformation），而這類映射的複合，對應 $2 \times 2$ 矩陣的乘法。這就是為甚麼週期性歸根究底取決於兩個特徵值之比是否單位根（root of unity）。當 $k = 37$ 時，兩個不動點（fixed points）1 和 36 的表現並不相同：36 會吸引附近的項，而 1 則會排斥它們，這一點可以從 $|f'(\ell)| = 36/\ell^2$ 看出。這是不動點迭代（fixed-point iteration）和數列收斂（convergence）理論的起點，Abbott 的 *Understanding Analysis*（分析學入門教材）對此有嚴格的論述。Needham 的 *Visual Complex Analysis*（以幾何直觀講解複分析）中，有一章精彩地講述莫比烏斯變換及其分類。

詳細解答

# STEP II 2006，第 3 題

純數學 • 根式與近似

> **STEP II 2006，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/06-s2-q3

## 關鍵思路

數 $5+\sqrt{24}$ 有一個「共軛數」(conjugate) $5-\sqrt{24}$，兩者相乘等於 1。把一個數與其共軛數的同次冪相加，所有根式 (surd) 都會互相抵消，所以和是整數。共軛數很小，它的冪就更加微小，因此那個大的冪必定恰好略小於一個整數。第 (ii) 部分是同一個論證，只是把 5 換成 $N$。

## 解答

### 第 (i) 部分

設 $a=5+\sqrt{24}$ 及 $b=5-\sqrt{24}$。則

$$a+b=10, \qquad ab=25-24=1,$$

所以 $b=1/a$。因此我們要求的式子就是 $a^4+b^4$。兩次運用 $a^2+b^2=(a+b)^2-2ab$：

$$\begin{aligned} a^2+b^2 &= 10^2-2=98, \\ a^4+b^4 &= (a^2+b^2)^2-2(ab)^2=98^2-2=9602. \end{aligned}$$

所以 $a^4+\dfrac{1}{a^4}=9602$，是一個整數。

至於各個不等式（inequality），記得 $\dfrac{1}{a}=b=5-\sqrt{24}$。

- $b>0.1 \iff \sqrt{24}<4.9 \iff 24<24.01$，這是成立的。（這裏取平方是可逆的，因為兩邊都是正數。）

- $b < \frac{2}{19} \iff \sqrt{24} > 5 - \frac{2}{19} = \frac{93}{19} \iff 24 \cdot 19^2 > 93^2 \iff 8664 > 8649$，這是成立的。
- $\frac{2}{19} < 0.11 \iff 2 < 2.09$，這是成立的。

所以 $0.1 < \dfrac{1}{5+\sqrt{24}} < \dfrac{2}{19} < 0.11$。

**想一想** 這些上下界怎樣把 $a^4$ 確定至小數點後四位？

由第一個結果，$a^4 = 9602 - b^4$。把各項皆為正數的不等式 $0.1 < b < 0.11$ 取四次方，

$$0.0001 < b^4 < 0.11^4 = 0.00014641,$$

因此

$$9601.99985359 < a^4 < 9601.9999.$$

這個區間內的每一個數都介乎 9601.99985 與 9601.9999 之間，所以捨入後得

$$\left(5+\sqrt{24}\right)^4 = 9601.9999 \text{ （準確至小數點後四位）。}$$

（較精確的上界 $b < \frac{2}{19}$ 給出 $b^4 < \frac{16}{130321} \approx 0.000123$，這裏用不着，但它說明了 $\frac{2}{19}$ 的來歷：見「驗算」。）

**第 (ii) 部分**

設 $A = N + \sqrt{N^2-1}$ 及 $B = N - \sqrt{N^2-1}$，由於 $N \geq 2$，兩者皆為實數。與之前一樣，

$$A + B = 2N, \qquad AB = N^2 - (N^2-1) = 1.$$

*第一步：$A^k + B^k$ 是整數。*用二項式定理(binomial theorem)展開這兩個冪。含 $\sqrt{N^2-1}$ 偶數次冪的項都是整數，而且在兩個展開式中完全相同；含奇數次冪的項則正負號相反，互相抵消。所以對每個正整數 $k$，$S_k = A^k + B^k$ 都是整數。（另一做法：$A$ 和 $B$ 是 $x^2 - 2Nx + 1 = 0$ 的根，所以 $S_k = 2NS_{k-1} - S_{k-2}$，其中 $S_0 = 2$，$S_1 = 2N$，再用歸納法（induction）便知它們全是整數。）

*第二步：$B$ 很小。*我們斷言 $0 < B < \dfrac{1}{2N - \frac{1}{2}}$。由於 $B = 1/A$ 且 $A > 0$，這等價於

$$A > 2N - \tfrac{1}{2} \iff \sqrt{N^2-1} > N - \tfrac{1}{2}.$$

兩邊都是正數（因為 $N \geq 2$），所以可以取平方：

$$N^2 - 1 > N^2 - N + \tfrac{1}{4} \iff N > \tfrac{5}{4},$$

而這是成立的，因為 $N \geq 2$。所以 $0 < B < \left(2N - \frac{1}{2}\right)^{-1}$，再取 $k$ 次方（各項皆為正數），

$$0 < B^k < \left(2N - \tfrac{1}{2}\right)^{-k}.$$

第三步：作結。我們有 $A^k = S_k - B^k$，所以 $A^k$ 與整數 $S_k$ 的距離恰好是 $B^k$。最接近 $A^k$ 的整數，與它的距離不會比 $S_k$ 更遠（事實上這個整數就是 $S_k$，因為 $B^k < \frac{1}{2}$）。因此

$$\left|A^k - (\text{最接近的整數})\right| \le B^k < \left(2N - \tfrac{1}{2}\right)^{-k},$$

這就是所要證明的結果。

## 驗算

- 以數值計算，$(5+\sqrt{24})^4 \approx 9601.999896$，落在我們的區間之內，捨入後是 9601.9999。
- 取 $N=5$，第 (ii) 部分給出 $(2N-\frac{1}{2})^{-1} = \frac{1}{9.5} = \frac{2}{19}$：正正是第 (i) 部分的上界。所以第 (i) 部分其實就是 $N=5$，$k=4$ 的情況，只是換了個面貌。
- 當 $N=2$，$k=1$：$2+\sqrt{3} \approx 3.732$ 與 4 相距約 0.268，而 $(3.5)^{-1} \approx 0.286$，所以上界成立。我們亦用電腦核對過 $N$ 至 1000、$k$ 至 24 的所有情況。

## 常見錯誤

- 把 $\sqrt{N^2-1} > N - \frac{1}{2}$ 兩邊取平方，卻沒有指出兩邊都是正數。
- 寫成 $a^4 \approx 9602.0000$：這個值比 9602 小，相差介乎 0.0001 與 0.00015 之間，而這已足以改變小數點後第四位。
- 忘記了排除 $N=1$ 的情況是有原因的：那時 $B=1$，整個論證便不成立。

## STEP 之外

像 $5+\sqrt{24}$ 這樣的數是環(ring)$\mathbb{Z}[\sqrt{6}]$ 中的單位(unit)：它們有同一形式的逆元 $5-\sqrt{24}$，而它們的冪給出佩爾方程（Pell equation）$x^2 - 24y^2 = 1$ 的所有解。這類單位的冪會越來越接近整數，這也是通往皮索數（Pisot number）理論的其中一條途徑，皮索數的冪同樣有這種接近整數的性質。Davenport 的 *The Higher Arithmetic*（一本初等數論名著）對佩爾方程和連分數（continued fraction）有十分精彩的論述。

詳細解答

# STEP II 2006，第 4 題

純數學 • 代換積分法

**STEP II 2006，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/06-s2-q4

## 關鍵思路

代換（substitution）$x = \pi - t$ 不會改變 $\sin x$，卻會把因子 $x$ 變成 $\pi - t$，於是原來的積分在等式右邊再次出現，可以像解方程一樣把它解出來。這樣就完全消去了礙事的因子 $x$。此後，題目中每個積分都是 $\sin x$ 的函數乘以 $\sin x$（或 $\cos x$），再用代換 $u = \cos x$ 或 $u = \sin x$ 便可完成。真正要花心思的，是看出怎樣把第 (ii) 和第 (iii) 部分化成這條恆等式（identity）適用的形式。

## 解答

### 恆等式

設 $I = \int_0^\pi x\, f(\sin x)\, \mathrm{d}x$。令 $x = \pi - t$，則 $\mathrm{d}x = -\mathrm{d}t$；當 $x = 0$ 時，$t = \pi$，而當 $x = \pi$ 時，$t = 0$。由於 $\sin(\pi - t) = \sin t$，

$$\begin{aligned} I &= \int_\pi^0 (\pi - t)\, f(\sin t)\, (-\mathrm{d}t) = \int_0^\pi (\pi - t)\, f(\sin t)\, \mathrm{d}t \\ &= \pi \int_0^\pi f(\sin t)\, \mathrm{d}t - \int_0^\pi t\, f(\sin t)\, \mathrm{d}t = \pi \int_0^\pi f(\sin x)\, \mathrm{d}x - I, \end{aligned}$$

因為積分變量叫甚麼名字並無關係。因此 $2I = \pi \int_0^\pi f(\sin x)\, \mathrm{d}x$，這正是我們所需的結果：

$$\int_0^\pi x\, f(\sin x)\, \mathrm{d}x = \tfrac{1}{2}\pi \int_0^\pi f(\sin x)\, \mathrm{d}x.$$

**第 (i) 部分**

這裏 $f(\sin x) = \dfrac{\sin x}{3+\sin^2 x}$，所以由恆等式可得

$$\int_0^{\pi} \frac{x\sin x}{3+\sin^2 x}\,\mathrm{d}x = \tfrac{1}{2}\pi J, \qquad J = \int_0^{\pi} \frac{\sin x}{3+\sin^2 x}\,\mathrm{d}x.$$

寫成 $3+\sin^2 x = 4-\cos^2 x$，並代入 $u = \cos x$，$\mathrm{d}u = -\sin x\,\mathrm{d}x$，其中 $u$ 由 1 遞減至 $-1$：

$$J = \int_{-1}^{1} \frac{\mathrm{d}u}{4-u^2}.$$

利用部分分式（partial fractions），$\dfrac{1}{4-u^2} = \dfrac{1}{4}\left(\dfrac{1}{2-u} + \dfrac{1}{2+u}\right)$，所以

$$J = \frac{1}{4}\left[\ln\frac{2+u}{2-u}\right]_{-1}^{1} = \frac{1}{4}\left(\ln 3 - \ln\tfrac{1}{3}\right) = \tfrac{1}{2}\ln 3.$$

（在 $[-1,1]$ 上，$2\pm u$ 兩者皆為正數，所以無須加上絕對值符號。）因此

$$\int_0^{\pi} \frac{x\sin x}{3+\sin^2 x}\,\mathrm{d}x = \frac{\pi}{4}\ln 3.$$

**第 (ii) 部分**

**想一想** 恆等式只適用於 $[0,\pi]$。我們該怎樣處理 $[\pi, 2\pi]$？

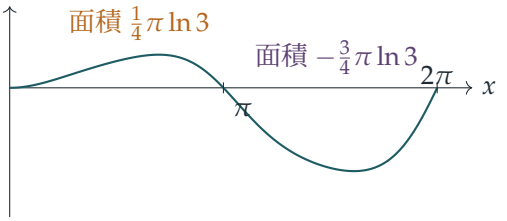
在 $\pi$ 處把積分範圍分成兩段，再用 $x = \pi + t$ 把第二段平移回 $[0,\pi]$。這時 $\sin x = -\sin t$ 而 $\sin^2 x = \sin^2 t$，所以

$$\begin{aligned}\int_{\pi}^{2\pi} \frac{x\sin x}{3+\sin^2 x}\,\mathrm{d}x &= -\int_0^{\pi} \frac{(\pi+t)\sin t}{3+\sin^2 t}\,\mathrm{d}t = -\pi J - \int_0^{\pi} \frac{t\sin t}{3+\sin^2 t}\,\mathrm{d}t \\ &= -\tfrac{1}{2}\pi\ln 3 - \tfrac{1}{4}\pi\ln 3 = -\tfrac{3}{4}\pi\ln 3,\end{aligned}$$

當中用到 $J = \frac{1}{2}\ln 3$ 和第 (i) 部分的結果。加上第 (i) 部分在 $[0,\pi]$ 上的積分，

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

下面的草圖說明了答案為何是負數：被積函數（integrand）在兩半區間上的形狀相同，但在 $[\pi, 2\pi]$ 上，它所乘的 $x$ 值較大。



### 第 (iii) 部分

由於 $\sin 2x = 2\sin x\cos x$，而且在 $[0,\pi]$ 上 $\sin x \geq 0$，

$$|\sin 2x| = 2\sin x\,|\cos x| = 2\sin x\sqrt{1-\sin^2 x},$$

這純粹是 $\sin x$ 的函數。

> **想一想** 為甚麼這裏出現的是 $|\cos x|$ 而不是 $\cos x$，這一點如此重要？

在 $[0,\pi]$ 上，$\cos x$ 並不是 $\sin x$ 的函數：$\cos(\pi - x) = -\cos x$，而 $\sin(\pi - x) = \sin x$。但 $|\cos x| = \sqrt{1-\sin^2 x}$ 卻是，所以正是有了絕對值，我們才可以使用恆等式：

$$\int_0^\pi \frac{x|\sin 2x|}{3+\sin^2 x}\,\mathrm{d}x = \tfrac{1}{2}\pi\int_0^\pi \frac{|\sin 2x|}{3+\sin^2 x}\,\mathrm{d}x.$$

新的被積函數在 $x \mapsto \pi - x$ 之下保持不變，所以在 $[0,\pi]$ 上的積分是在 $[0,\frac{1}{2}\pi]$ 上的積分的兩倍，而在後者的區間上 $\cos x \geq 0$，絕對值符號可以去掉：

$$\int_0^\pi \frac{|\sin 2x|}{3+\sin^2 x}\,\mathrm{d}x = 2\int_0^{\pi/2} \frac{2\sin x\cos x}{3+\sin^2 x}\,\mathrm{d}x = 2\Big[\ln(3+\sin^2 x)\Big]_0^{\pi/2} = 2\ln\tfrac{4}{3},$$

因為分子正是分母的導數。因此

$$\int_0^\pi \frac{x|\sin 2x|}{3+\sin^2 x}\,\mathrm{d}x = \tfrac{1}{2}\pi\cdot 2\ln\tfrac{4}{3} = \pi\ln\frac{4}{3}.$$

## 驗算

- 用辛普森法則（Simpson's rule）計算，(i) 得 0.86285，(ii) 得 $-1.72570$，(iii) 得 0.90378，與 $\frac{1}{4}\pi\ln 3$、$-\frac{1}{2}\pi\ln 3$ 和 $\pi\ln\frac{4}{3}$ 吻合。
- (ii) 的答案恰好是 (i) 的答案的 $-2$ 倍。在 $\pi$ 處分段便可解釋這一點：兩半分別是 $\frac{1}{4}\pi\ln 3$ 和 $-\frac{3}{4}\pi\ln 3$。
- 快速檢驗一下恆等式：取 $f \equiv 1$，兩邊都等於 $\frac{1}{2}\pi^2$。

## 常見錯誤

- 在 $[0,2\pi]$ 上套用類似的「恆等式」，這樣會得出 $\pi\int_0^{2\pi} f(\sin x)\,\mathrm{d}x = 0$。證明需要用到 $\sin(2\pi - t) = \sin t$，而這是錯的。
- 在 (iii) 中，在整個 $[0,\pi]$ 上都寫 $|\sin 2x| = 2\sin x\cos x$，這樣裏面的積分會得出 0。

## STEP 之外

讓積分再次出現然後把它解出來的技巧，以及運用 $x \mapsto \pi - x$ 這類對稱性（symmetry），都是分析學（analysis）中的標準工具，也常用於計算沒有初等原函數（elementary antiderivative）的定積分。它們通向傅立葉級數（Fourier series），在那裏，$x$ 乘以三角函數在一個周期上的積分正是最基本的研究對象。想多了解這類論證手法，可參閱 Siklos 的 *Advanced Problems in Mathematics*；至於積分的嚴格理論，可參閱 Abbott 的 *Understanding Analysis*。

詳細解答

# STEP II 2006，第 5 題

純數學 • 下取整函數與面積

**STEP II 2006，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/06-s2-q5

## 關鍵思路

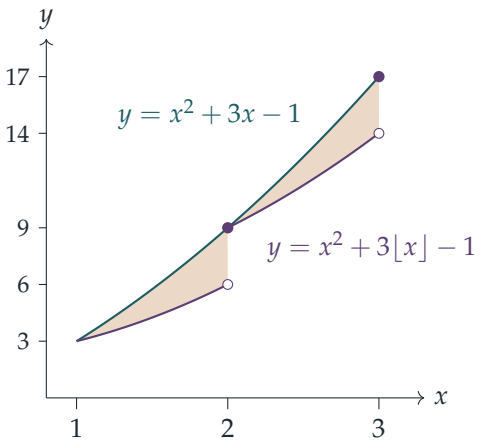
在兩個相鄰整數之間，即 $k \le x < k+1$ 時，下取整函數（floor function）是常數，其值為 $k$。因此在每一個這樣的單位區間上，那條鋸齒狀的曲線其實只是一條普通的光滑曲線，而所求面積可以分拆成若干長條的面積之和，每個單位區間對應一條。先把 $[k, k+1]$ 上那一條的面積寫成 $k$ 的函數，再把 $k = 1, 2, \ldots, n-1$ 的各條加起來。

## 解答

以下一律記 $f(x) = x^2 + 3x - 1$。在 $1 \le x \le 3$ 上，它由 $f(1) = 3$ 經過 $f(2) = 9$ 升至 $f(3) = 17$；而且對所有 $x \ge 1$ 它都是遞增的，因為在該範圍內 $f'(x) = 2x + 3 > 0$。

### 第 (i) 部分

第二條曲線是 $y = x^2 + 3\lfloor x \rfloor - 1$。在 $[1, 2)$ 上有 $\lfloor x \rfloor = 1$，所以它是 $y = x^2 + 2$，由 3 升向 6；在 $[2, 3)$ 上有 $\lfloor x \rfloor = 2$，所以它是 $y = x^2 + 5$，由 9 升向 14；而在 $x = 3$ 處它等於 17。每一段都是拋物線 $y = f(x)$ 向下平移所得，而兩條曲線在每個整數處相交。




兩條曲線之間的垂直距離是

$$(x^2 + 3x - 1) - (x^2 + 3\lfloor x \rfloor - 1) = 3(x - \lfloor x \rfloor) \geq 0,$$

所以第一條曲線永遠不會在第二條之下。在區間 $[k, k+1]$ 上（$k$ 為整數），這個距離是 $3(x-k)$，只有 $x = k+1$ 這一點例外，而單獨一點並不影響積分。因此 $[k, k+1]$ 上的長條面積為

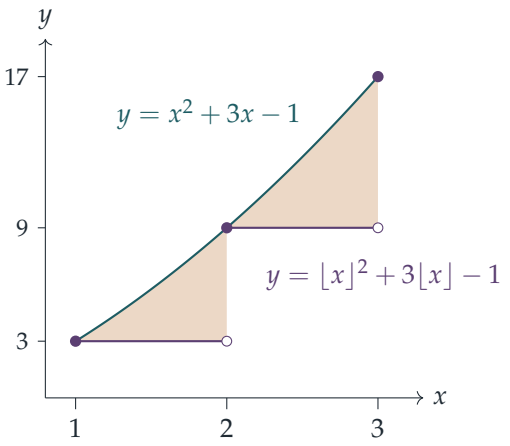
$$\int_k^{k+1} 3(x-k)\,\mathrm{d}x = \left[\tfrac{3}{2}(x-k)^2\right]_k^{k+1} = \tfrac{3}{2},$$

對每個 $k$ 都相同。區間 $1 \leq x \leq n$ 由 $n-1$ 個單位區間 $[1,2], [2,3], \ldots, [n-1, n]$ 組成，所以兩條曲線之間的面積是

$$\tfrac{3}{2}(n-1).$$

**第 (ii) 部分**

現在第二條曲線是 $y = f(\lfloor x \rfloor)$，這是一個階梯函數（step function）：在 $[1,2)$ 上 $y = f(1) = 3$，在 $[2,3)$ 上 $y = f(2) = 9$，而在 $x = 3$ 處 $y = f(3) = 17$。每一級階梯都由拋物線上某個整數點出發，水平向右延伸。




**想一想** 哪一條曲線在上方？它會一直在上方嗎？

由於 $f$ 在 $x \geq 1$ 時遞增，而且 $\lfloor x \rfloor \leq x$，所以在整個 $1 \leq x \leq n$ 上都有 $f(\lfloor x \rfloor) \leq f(x)$。因此拋物線始終在上方，所求面積就是 $f(x) - f(\lfloor x \rfloor)$ 的積分，無須擔心正負號改變。

在 $[k, k+1]$ 上（$k \geq 1$ 為整數），階梯的高度是 $f(k)$（同樣不理會 $x = k+1$ 這一點）。代入 $x = k + t$，其中 $0 \leq t \leq 1$，

$$\begin{aligned} f(k+t) - f(k) &= (k+t)^2 + 3(k+t) - 1 - (k^2 + 3k - 1) \\ &= 2kt + t^2 + 3t, \end{aligned}$$

所以 $[k, k+1]$ 上的長條面積為

$$\int_0^1 (2kt + t^2 + 3t)\,\mathrm{d}t = k + \tfrac{1}{3} + \tfrac{3}{2} = k + \tfrac{11}{6}.$$

與第 (i) 部分不同，這次長條的面積隨 $k$ 增大，因為拋物線愈來愈陡。把 $k = 1, 2, \ldots, n-1$ 的長條加起來，並利用 $\sum_{k=1}^{n-1} k = \frac{1}{2}n(n-1)$，

$$\begin{aligned} \text{面積} &= \sum_{k=1}^{n-1} \left(k + \tfrac{11}{6}\right) = \tfrac{1}{2}n(n-1) + \tfrac{11}{6}(n-1) \\ &= \tfrac{1}{6}(n-1)(3n+11), \end{aligned}$$

這正是所要求的結果。

## 驗算

- 當 $n = 1$ 時根本沒有區域，而兩個答案都給出 0。
- 當 $n = 2$ 時，第 (ii) 部分給出 $\frac{1}{6} \cdot 1 \cdot 17 = \frac{17}{6}$。直接計算，$\int_1^2 (x^2 + 3x - 4)\,\mathrm{d}x = \frac{7}{3} + \frac{9}{2} - 4 = \frac{17}{6}$。

- 以中點法則（midpoint rule）對 $n = 2, \ldots, 7$ 的差距作數值積分，結果與 $\frac{3}{2}(n-1)$ 及 $\frac{1}{6}(n-1)(3n+11)$ 吻合至小數點後四位；例如 $n = 4$ 時第 (ii) 部分給出 11.5。

## 常見錯誤

- 數成 $n$ 條長條而不是 $n-1$ 條：區間 $[1, n]$ 只包含 $n-1$ 個單位區間。
- 在第 (ii) 部分，誤用 $\int_0^1 2kt \, \mathrm{d}t = 2k$ 而不是 $k$。
- 在整個區間上把 $\lfloor x \rfloor$ 當作 $x$ 來積分（或乾脆忽略它），而沒有先在各整數處把區間分開。

## STEP 之外

第 (ii) 部分比較的是曲線下的面積與曲線下方一列「階梯」狀矩形的面積，這正正是步長為 1 的黎曼下和（lower Riemann sum）。在大學第一門分析（analysis）課程中，積分的定義就是讓階梯愈來愈窄，把這樣的下和與上和夾在一起；而像 $\lfloor x \rfloor$ 這類有跳躍的函數仍然可積（integrable），因為它們的跳躍點都是孤立點。Abbott 的 *Understanding Analysis* 對黎曼積分（Riemann integral）有非常易讀的講解。

詳細解答

# STEP II 2006，第 7 題

純數學 • 橢圓的切線

> **STEP II 2006，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/06-s2-q7

## 關鍵思路

求切線（tangent）最快的方法，是用「把 $x^2$ 換成 $xx_0$、把 $y^2$ 換成 $yy_0$」的形式，用隱函數微分法（implicit differentiation）也同樣可以。至於那個作圖，只需求出 $Q$ 的坐標和 $PQ$ 的方程，再與一般點上的切線比較。題目中 $\tan(\alpha/2) = k$ 這幾個字是很強的提示：它叫你用 $t = \tan(\alpha/2)$ 來表示 $\cot\alpha$ 和 $\operatorname{cosec}\alpha$。

## 解答

### 切線

把 $\dfrac{x^2}{a^2} + \dfrac{y^2}{b^2} = 1$ 對 $x$ 作隱函數微分，得

$$\frac{2x}{a^2} + \frac{2y}{b^2}\frac{\mathrm{d}y}{\mathrm{d}x} = 0, \qquad \text{所以} \qquad \frac{\mathrm{d}y}{\mathrm{d}x} = -\frac{b^2x}{a^2y}.$$

在 $(a\cos\alpha, b\sin\alpha)$ 處，設 $\sin\alpha \neq 0$（否則切線是鉛垂線，題目所給的方程便沒有意義），斜率（gradient）為

$$-\frac{b^2 a\cos\alpha}{a^2 b\sin\alpha} = -\frac{b\cot\alpha}{a}.$$

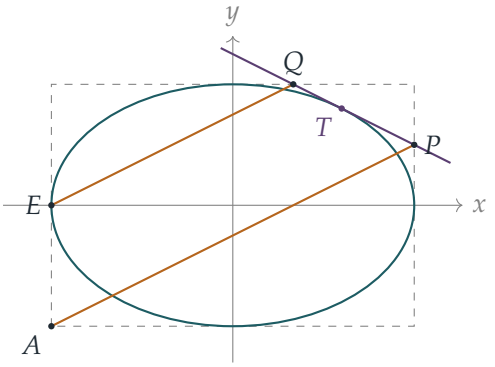
所以切線是

$$\begin{aligned} y - b\sin\alpha &= -\frac{b\cot\alpha}{a}(x - a\cos\alpha) \\ \implies \quad y &= -\frac{b\cot\alpha}{a}x + b\frac{\cos^2\alpha}{\sin\alpha} + b\sin\alpha = -\frac{b\cot\alpha}{a}x + \frac{b(\cos^2\alpha + \sin^2\alpha)}{\sin\alpha}, \end{aligned}$$

即 $y = -\dfrac{b\cot\alpha}{a}x + b\operatorname{cosec}\alpha$，這正是所求。

**作圖**

下圖取 $a=3$、$b=2$、$k=\frac{1}{2}$。橢圓（ellipse）位於以 $(\pm a,\pm b)$ 為頂點的長方形之內；$A$ 是左下角的頂點，$E$ 是左邊的中點，$P$ 在右邊上，而 $Q$ 在上邊上。




直線 $AP$ 連接 $(-a,-b)$ 與 $(a,kb)$，所以它的斜率是

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

通過 $E(-a,0)$ 的平行線是 $y=\dfrac{(1+k)b}{2a}(x+a)$。代入 $y=b$ 得 $x+a=\dfrac{2a}{1+k}$，所以

$$Q=\left(\frac{a(1-k)}{1+k},\ b\right).$$

於是 $PQ$ 的斜率是

$$\frac{b-kb}{\dfrac{a(1-k)}{1+k}-a}=\frac{b(1-k)}{\dfrac{-2ak}{1+k}}=-\frac{b(1-k^2)}{2ak},$$

由於 $k\neq 0$，這個式子沒有問題。利用點 $P(a,kb)$，直線 $PQ$ 是

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

所以

$$PQ:\qquad y=-\frac{b}{a}\cdot\frac{1-k^2}{2k}\,x+b\cdot\frac{1+k^2}{2k}.$$

**想一想** 我們怎樣看出這是橢圓上某一點的切線？

把它與上面求得在 $(a\cos\alpha,b\sin\alpha)$ 處的切線比較。我們需要找一個 $\alpha$，使得 $\cot\alpha=\dfrac{1-k^2}{2k}$ 及 $\operatorname{cosec}\alpha=\dfrac{1+k^2}{2k}$。這正是 $t$ 公式（t-formulae）：設 $t=\tan(\alpha/2)$，則

$$\tan\alpha=\frac{2t}{1-t^2},\qquad \sin\alpha=\frac{2t}{1+t^2},\qquad \text{所以}\qquad \cot\alpha=\frac{1-t^2}{2t},\qquad \operatorname{cosec}\alpha=\frac{1+t^2}{2t}.$$

由於 $0 < k < 1$，我們可以取 $\alpha = 2\arctan k$，它位於 $(0, \frac{\pi}{2})$ 之內，所以 $\sin\alpha \neq 0$，而且 $\tan(\alpha/2) = k$。取 $t = k$，這兩個表達式便與 $PQ$ 的斜率和截距（intercept）都吻合。所以 $PQ$ 就是直線 $y = -\frac{b\cot\alpha}{a}x + b\operatorname{cosec}\alpha$，亦即橢圓在 $(a\cos\alpha, b\sin\alpha)$ 處的切線，其中 $\tan(\alpha/2) = k$。這正是我們要證明的。

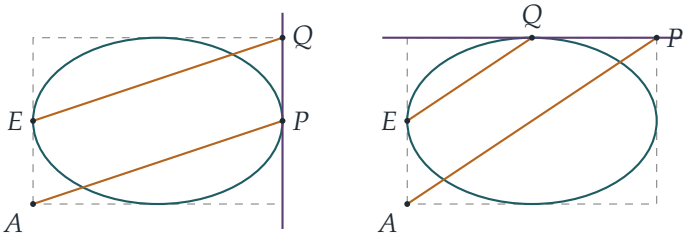
用坐標表示，切點（point of contact）是 $T = \left(\frac{a(1-k^2)}{1+k^2}, \frac{2kb}{1+k^2}\right)$。

### $k = 0$ 與 $k = 1$ 的情況

上面的代數運算曾經除以 $k$，而切線公式又需要 $\sin\alpha \neq 0$，所以我們直接把作圖重做一次。

$k = 0$。這時 $P = (a, 0)$，而 $AP$ 的斜率是 $\frac{b}{2a}$。通過 $E$ 的平行線是 $y = \frac{b}{2a}(x + a)$，它與 $y = b$ 相交於 $x = a$，所以 $Q = (a, b)$。於是 $PQ$ 是鉛垂線 $x = a$。這是橢圓在 $(a, 0)$ 處的切線，該點對應 $\alpha = 0$，而 $\tan(0/2) = 0 = k$。所以即使這條切線不能寫成 $y = mx + c$ 的形式，結果*確實*仍然成立。

$k = 1$。這時 $P = (a, b)$，而 $AP$ 是長方形的對角線，斜率為 $\frac{b}{a}$。通過 $E$ 的平行線是 $y = \frac{b}{a}(x + a)$，它與 $y = b$ 相交於 $x = 0$，所以 $Q = (0, b)$。於是 $PQ$ 是水平線 $y = b$，即橢圓最高點 $(0, b)$ 處的切線。這一點對應 $\alpha = \frac{\pi}{2}$，而 $\tan(\pi/4) = 1 = k$。所以結果在這裏同樣成立（而且公式也適用：由 $\cot\frac{\pi}{2} = 0$ 和 $\operatorname{cosec}\frac{\pi}{2} = 1$ 得出 $y = b$）。




## 驗算

- 取 $a = 3$、$b = 2$、$k = \frac{1}{2}$。這時 $Q = (1, 2)$，而 $PQ$ 是 $y = -\frac{1}{2}x + \frac{5}{2}$。另外 $\tan\alpha = \frac{4}{3}$，所以 $\cos\alpha = \frac{3}{5}$、$\sin\alpha = \frac{4}{5}$，而 $T = (\frac{9}{5}, \frac{8}{5})$。這一點在 $PQ$ 上（$-\frac{9}{10} + \frac{5}{2} = \frac{8}{5}$），而該處的斜率是 $-\frac{b^2x}{a^2y} = -\frac{4\cdot 9/5}{9\cdot 8/5} = -\frac{1}{2}$。
- 把 $PQ$ 代入橢圓方程，應該得到一條有重根（repeated root）的二次方程。我們已用一千組隨機的 $a$、$b$、$k$ 以數值方法驗證了這一點，以及 $Q$ 和 $PQ$ 的公式。
- 當 $k \to 0^+$ 時，$T \to (a, 0)$；當 $k \to 1^-$ 時，$T \to (0, b)$。這與兩個端點情況一致。

## 常見錯誤

- 弄錯 $PQ$ 斜率的正負號：分母 $\frac{a(1-k)}{1+k} - a$ 是負數。
- 因為 $\operatorname{cosec} 0$ 沒有定義，便說結果在 $k = 0$ 時不成立。失效的只是公式，幾何上並沒有問題：$PQ$ 仍然是 $\alpha = 0$ 處的切線。

## STEP 之外

代換 $t = \tan(\alpha/2)$ 給出橢圓的一個有理參數化（rational parametrisation）：$\left(a\frac{1-t^2}{1+t^2}, b\frac{2t}{1+t^2}\right)$。把三角函數積分化為有理函數積分，以及求出所有畢氏三元數（Pythagorean triples），用的都是同一個想法。每一條圓錐曲線（conic）都可以這樣作有理參數化，而像本題這種只靠平行線和交點完成的作圖，正是射影幾何（projective geometry）研究的對象。想初步了解這與數論（number theory）的關係，可以讀 Davenport 的 *The Higher Arithmetic*。

詳細解答

# STEP II 2006，第 8 題

純數學 • 直線的向量幾何

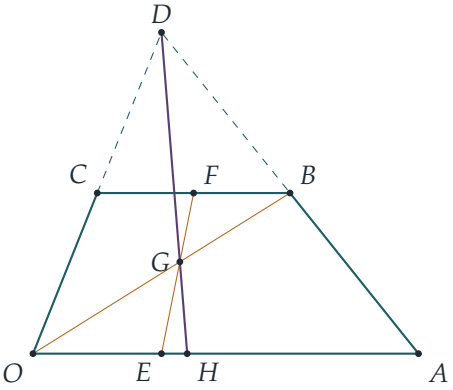
**STEP II 2006，第 8 題**

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/06-s2-q8

## 關鍵思路

題目中每一點都定義為兩條直線的交點，所以整條題目只靠一套固定做法：把每條直線寫成第一部分的形式，令兩個表達式相等，然後比較 **a** 和 **c** 的係數（comparing coefficients）。這樣比較是合理的，因為 **a** 與 **c** 不平行，所以一個向量寫成 $p\mathbf{a} + q\mathbf{c}$ 的方式只有一種。




## 解答

### 直線的方程

通過位置向量（position vector）為 **x** 和 **y** 的兩點的直線，經過 **x**，方向為 $\mathbf{y} - \mathbf{x}$，所以線上的點是

$$\mathbf{r} = \mathbf{x} + \alpha(\mathbf{y} - \mathbf{x}) = (1 - \alpha)\mathbf{x} + \alpha\mathbf{y},$$

其中 $\alpha$ 取遍所有實數。留意 $\alpha = 0$ 給出 $\mathbf{x}$，$\alpha = 1$ 給出 $\mathbf{y}$，而兩個係數之和恆為 1。這就是所要求的形式。

在整個解答中，$\mathbf{a}$ 與 $\mathbf{c}$ 都不平行（否則 $O$、$A$、$C$ 三點共線（collinear），便不會有梯形（trapezium）），所以若 $p\mathbf{a} + q\mathbf{c} = p'\mathbf{a} + q'\mathbf{c}$，則 $p = p'$ 且 $q = q'$。

**第 (i) 部分**

由於 $CB$ 平行於 $OA$，向量 $\overrightarrow{CB}$ 是 $\overrightarrow{OA} = \mathbf{a}$ 的純量倍數（scalar multiple）；記它為 $\lambda\mathbf{a}$。於是

$$\mathbf{b} = \overrightarrow{OC} + \overrightarrow{CB} = \mathbf{c} + \lambda\mathbf{a}.$$

按次序繞 $OABC$ 一周，$\overrightarrow{CB}$ 與 $\overrightarrow{OA}$ 方向相同，所以 $\lambda > 0$；而 $CB < OA$ 給出 $\lambda < 1$。因此 $0 < \lambda < 1$，特別地 $\lambda \neq 1$，這一點稍後會用到。

**第 (ii) 部分**

*E* **和** *F*。這兩點可直接由比例得出：

$$\mathbf{e} = \tfrac{1}{3}\mathbf{a}, \qquad \mathbf{f} = \tfrac{1}{2}(\mathbf{c} + \mathbf{b}) = \tfrac{1}{2}\lambda\mathbf{a} + \mathbf{c}.$$

$D$。直線 $OC$ 上的點是 $\mu\mathbf{c}$。直線 $AB$ 上的點是

$$(1 - \beta)\mathbf{a} + \beta(\lambda\mathbf{a} + \mathbf{c}) = (1 - \beta + \beta\lambda)\mathbf{a} + \beta\mathbf{c}.$$

在 $D$ 點，兩者相等。比較係數得 $\beta = \mu$ 及 $1 - \beta(1 - \lambda) = 0$，所以 $\beta = \dfrac{1}{1 - \lambda}$（沒有問題，因為 $\lambda \neq 1$）。因此

$$\mathbf{d} = \frac{1}{1 - \lambda}\,\mathbf{c}.$$

由於 $\frac{1}{1-\lambda} > 1$，$D$ 確實位於 $OC$ 越過 $C$ 的延長線上，與預期相符。

$G$。$OB$ 上的點是 $t(\lambda\mathbf{a} + \mathbf{c})$。$EF$ 上的點是

$$(1 - s)\tfrac{1}{3}\mathbf{a} + s\big(\tfrac{1}{2}\lambda\mathbf{a} + \mathbf{c}\big) = \big(\tfrac{1}{3} - \tfrac{1}{3}s + \tfrac{1}{2}\lambda s\big)\mathbf{a} + s\mathbf{c}.$$

比較 $\mathbf{c}$ 的係數得 $s = t$；再比較 $\mathbf{a}$ 的係數，得

$$\lambda t = \tfrac{1}{3} - \tfrac{1}{3}t + \tfrac{1}{2}\lambda t \quad\Longrightarrow\quad t\Big(\frac{3\lambda + 2}{6}\Big) = \frac{1}{3} \quad\Longrightarrow\quad t = \frac{2}{3\lambda + 2}.$$

所以

$$\mathbf{g} = \frac{2}{3\lambda + 2}\,(\lambda\mathbf{a} + \mathbf{c}).$$

$H$。直線 $DG$ 上的點是

$$(1 - k)\mathbf{d} + k\mathbf{g} = \frac{2k\lambda}{3\lambda + 2}\,\mathbf{a} + \Big(\frac{1 - k}{1 - \lambda} + \frac{2k}{3\lambda + 2}\Big)\mathbf{c}.$$

**想一想** 怎樣從這條直線上找出位於 $OA$ 上的那一點？

直線 $OA$ 上的點只是 $\mathbf{a}$ 的倍數，所以我們需要 $\mathbf{c}$ 的係數等於零。兩邊乘以 $(1-\lambda)(3\lambda+2)$，

$$\begin{aligned}(1-k)(3\lambda+2)+2k(1-\lambda)&=0\\(3\lambda+2)-k\big(3\lambda+2-2+2\lambda\big)&=0\\k&=\frac{3\lambda+2}{5\lambda}.\end{aligned}$$

這時 $\mathbf{a}$ 的係數是

$$\frac{2\lambda}{3\lambda+2}\cdot\frac{3\lambda+2}{5\lambda}=\frac{2}{5},\qquad \text{所以}\qquad \mathbf{h}=\tfrac{2}{5}\,\mathbf{a}.$$

把答案整理如下：

$$\mathbf{d}=\frac{\mathbf{c}}{1-\lambda},\quad \mathbf{e}=\tfrac{1}{3}\mathbf{a},\quad \mathbf{f}=\tfrac{1}{2}\lambda\mathbf{a}+\mathbf{c},\quad \mathbf{g}=\frac{2(\lambda\mathbf{a}+\mathbf{c})}{3\lambda+2},\quad \mathbf{h}=\tfrac{2}{5}\mathbf{a}.$$

由於 $OH=\frac{2}{5}OA$ 及 $HA=\frac{3}{5}OA$，

$$OH:HA=2:3,$$

不論梯形是甚麼形狀，結果都一樣。

## 驗算

- 取 $\mathbf{a}=(6,0)$、$\mathbf{c}=(1,2.5)$ 及 $\lambda=\frac{1}{2}$（即圖中的情況）。由公式得 $D=(2,5)$、$G\approx(2.286,1.429)$ 及 $H=(2.4,0)$；由 $D$ 到 $G$ 的斜率是 $-3.571/0.286=-12.5$，由 $D$ 到 $H$ 的斜率也是這個值，即 $-5/0.4$。所以 $D$、$G$、$H$ 確實共線。
- $H$ 的答案與 $\lambda$、$\mathbf{a}$ 和 $\mathbf{c}$ 都無關。我們用電腦以精確的有理數運算，對 2000 個隨機梯形重複整個作圖過程，每次都得到 $OH:HA=2:3$。
- 當 $\lambda\to1$ 時，邊 $OC$ 與 $AB$ 趨於平行，而 $\mathbf{d}=\mathbf{c}/(1-\lambda)$ 跑向無窮遠，與預期相符。

## 常見錯誤

- 兩條不同的直線用了同一個參數 (parameter) 字母。每條直線都要有自己的參數，然後才可令兩式相等。
- 比較係數卻沒有說明為何可以這樣做：這方法行得通，是因為 $\mathbf{a}$ 與 $\mathbf{c}$ 不平行。
- 寫成 $\mathbf{b}=\mathbf{a}+\lambda\mathbf{c}$。與 $\mathbf{a}$ 平行的是 $\overrightarrow{CB}$，而不是 $\overrightarrow{AB}$。

## STEP 之外

答案 2 : 3 與梯形的形狀無關，是因為每一步只用到直線、交點和直線上的比例，而仿射變換（affine map，即線性變換（linear map）之後再作平移）會保持這一切不變。所以你可以把梯形變換成一個方便計算的梯形，例如令 $OABC$ 中的 $C$ 位於 $O$ 的正上方，然後在那裏計算。這就是仿射幾何（affine geometry）的想法；係數形式 $(1-\alpha)\mathbf{x}+\alpha\mathbf{y}$ 正是重心坐標（barycentric coordinates）的起點。Strang 的 *Introduction to Linear Algebra* 講解線性變換和坐標，而 Siklos 的 *Advanced Problems in Mathematics* 則有更多這類風格的向量幾何題目。

詳細解答

# STEP II 2006，第 9 題

力學 • 摩擦力與傾倒

**STEP II 2006，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：

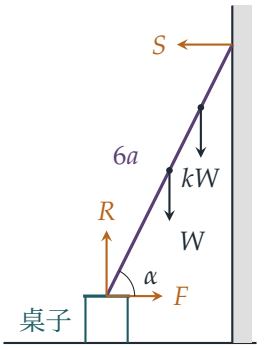
step.maths.org/questions/06-s2-q9

## 關鍵思路

把梯子和桌子視為兩個獨立的物體，兩者靠梯腳處的接觸連繫起來。單看梯子，便可求出梯腳處的反作用力和摩擦力，而這兩個力取決於油漆工站得多高。然後根據牛頓第三定律（Newton's third law），同樣的力以相反方向作用於桌子。對於桌子，我們比較兩個臨界值：一個是令桌子滑動的推力，另一個是對遠端桌腳的力矩（moment）大於桌上重量的力矩的推力。

## 解答

我們把梯子模擬為均勻桿（uniform rod），油漆工模擬為質點（particle），牆壁視為光滑（smooth），桌子則視為剛體（rigid body），桌面高度為 $a$（忽略桌面的厚度）。設 $\alpha = \arctan 2$，則 $\tan\alpha = 2$、$\sin\alpha = \frac{2}{\sqrt{5}}$ 及 $\cos\alpha = \frac{1}{\sqrt{5}}$。設油漆工站在沿梯子距梯腳 $x$ 之處，其中 $0 \le x \le 6a$。



**第 (i) 部分**

作用於梯子的力有：作用於梯子中點的重量 $W$、油漆工的重量 $kW$、光滑牆壁給予的水平反作用力 $S$，以及梯腳處向上的法向反作用力（normal reaction）$R$ 和指向牆壁的摩擦力（friction）$F$。（摩擦力必須指向牆壁，因為它是唯一能與 $S$ 平衡的水平力。）

沿鉛垂和水平方向分解力，得

$$R = (k+1)W, \qquad F = S.$$

對梯腳取力矩：梯頂比梯腳高 $6a\sin\alpha$，而兩個重量的作用線與梯腳的水平距離分別為 $3a\cos\alpha$ 和 $x\cos\alpha$，所以

$$S \cdot 6a\sin\alpha = W \cdot 3a\cos\alpha + kW \cdot x\cos\alpha.$$

兩邊除以 $6a\sin\alpha$，並利用 $\cot\alpha = \frac{1}{2}$，得

$$F = S = \frac{W}{12}\left(3 + \frac{kx}{a}\right).$$

此式隨 $x$ 遞增，所以當油漆工站在梯子最頂端，即 $x = 6a$ 時，所需的摩擦力最大：

$$F_{\max} = \frac{(3+6k)W}{12} = \frac{(2k+1)W}{4}.$$

可提供的最大摩擦力為

$$\tfrac{1}{2}R = \frac{(k+1)W}{2} = \frac{(2k+2)W}{4}.$$

由於 $2k+1 < 2k+2$，無論油漆工站在哪個位置，也無論 $k$ 取何值，都有 $F \le F_{\max} < \frac{1}{2}R$。因此，在桌子固定的情況下，不論油漆工站得多高，梯子都不會滑動，這正是所要證明的。

**第 (ii) 部分**

現在 $k = 9$，所以 $R = 10W$，而

$$F = \frac{W}{4}\left(1 + \frac{3x}{a}\right),$$

它由 $x = 0$ 時的 $\frac{1}{4}W$ 增至 $x = 6a$ 時的 $\frac{19}{4}W$。第 (i) 部分的結論仍然適用，所以梯子不會在桌面上滑動。

根據牛頓第三定律，梯子在桌面中心對桌面施力：一個向下的力 $R = 10W$，以及一個背離牆壁的水平力 $F$。桌子本身的重量 $2W$ 亦作用於其中心，而地面則以法向反作用力和摩擦力作用於各桌腳。

? **想一想** 作用於桌面中心的水平推力，怎樣能令桌子傾側？

推力 $F$ 作用於高度 $a$ 處，所以它對離牆最遠的兩條桌腳的底部有力矩。如果桌子傾側，它會繞着連接這兩條桌腳底部的直線轉動。各桌腳位於邊長為 $\frac{1}{2}a$ 的正方形的四個頂點，所以該直線與桌子中心的水平距離為 $\frac{1}{4}a$。

**傾側**。對連接遠端桌腳底部的直線而言，地面作用於遠端桌腳的力沒有力矩，地面作用於任何一條桌腳的摩擦力也沒有力矩（因為它作用於地面的高度）。作用於中心的向下的力 $10W + 2W = 12W$ 有恢復力矩 $12W \cdot \frac{1}{4}a = 3Wa$。推力 $F$ 則有傾覆力矩 $Fa$。對這條直線取力矩，離牆最近的兩條桌腳所受的法向反作用力（合計）$N_1$ 滿足

$$N_1 \cdot \tfrac{1}{2}a = 3Wa - Fa,$$

所以 $N_1 \geq 0$ 當且僅當 $F \leq 3W$。當 $N_1 = 0$ 時，桌子即將傾側，也就是當

$$\frac{W}{4}\left(1 + \frac{3x}{a}\right) = 3W \quad \Longleftrightarrow \quad x = \frac{11a}{3}.$$

**滑動**。對桌子沿鉛垂方向分解力，地面給予的法向反作用力合共為 $12W$，所以可提供的最大總摩擦力為 $\frac{1}{3} \cdot 12W = 4W$。沿水平方向分解力，所需的摩擦力為 $F$。所以桌子即將滑動的條件是

$$\frac{W}{4}\left(1 + \frac{3x}{a}\right) = 4W \quad \Longleftrightarrow \quad x = 5a.$$

這兩個值都在梯子的長度 $6a$ 之內，而 $F$ 隨着油漆工向上爬而不斷增加。它先達到 $3W$（在 $x = \frac{11}{3}a$ 時），然後才達到 $4W$（在 $x = 5a$ 時）。所以**桌子先傾側**。這發生在油漆工沿梯子走了 $\frac{11}{3}a$ 之時，即在桌面之上 $\frac{11}{3}a \sin\alpha = \frac{22a}{3\sqrt{5}} \approx 3.28a$ 的高度。在那一刻，地面所需提供的摩擦力只是 $3W$，小於可提供的 $4W$，所以桌子並未滑動。

## 驗算

- 梯上沒有油漆工時（$x = 0$），$F = \frac{1}{4}W$，近牆的桌腳承受 $N_1 = \frac{11}{2}W$，而遠端的桌腳承受 $\frac{13}{2}W$。遠端桌腳承受的力稍大，這與推力背離牆壁的方向相符。
- 用電腦對多個 $k$ 和 $x$ 的值計算各個力。在 (i) 中，所需摩擦力與可提供摩擦力之比的最大值僅略小於 1（在梯頂時為 $\frac{2k+1}{2k+2}$）。在 (ii) 中，電腦求得近牆桌腳的反作用力在 $x \approx 3.667a$ 時變為零，而地面摩擦力在 $x = 5a$ 時達到極限。

## 常見錯誤

- 把傾側的轉軸取在距離中心 $\frac{1}{2}a$ 的桌邊：桌面的邊長是 $\frac{1}{2}a$，所以半闊度是 $\frac{1}{4}a$。
- 忘記了作用於桌子的推力位於高度 $a$ 處，即桌腳的頂端，而不是在地面的高度。正是這個高度令推力產生傾覆力矩。

- 計算地面對桌子的法向反作用力時，漏掉了梯子和油漆工的 10*W*。只計桌子本身的重量 2*W* 並不足夠。

## STEP 之外

這道題探討的是由多個剛體組成的系統怎樣以互相競爭的方式失效：滑動，或是翻倒。方法是求出每種方式的臨界值，再看負載先達到哪一個。工程師把同樣的想法應用於擋土牆、起重機和堆疊的貨物；在靜力學（statics）課程中，它會進一步引申到虛功原理（principle of virtual work）。Kleppner 與 Kolenkow 的 *An Introduction to Mechanics*（力學入門教科書）對剛體平衡和摩擦力有細緻的論述。

詳細解答

# STEP II 2006，第 10 題

力學 • 碰撞與恢復係數

> **STEP II 2006，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/06-s2-q10

## 關鍵思路

每一次碰撞都是一道兩條方程的問題：沿直線方向的動量守恆（conservation of momentum），以及牛頓恢復定律（Newton's law of restitution）（分離速率等於 $e$ 乘以接近速率）。先解 $A$–$B$ 碰撞，再把 $B$ 的新速度代入 $B$–$C$ 碰撞。第二次 $A$–$B$ 碰撞發生，當且僅當位於 $B$ 後方的 $A$，其向前的速度最終大於 $B$ 在 $B$ 撞上 $C$ 之後的速度。第 (ii) 部分於是成為一道匀速的追及問題。

## 解答

我們把 $A$、$B$ 和 $C$ 視為在光滑（smooth）平面上沿一條直線運動的質點（particle），因此除碰撞期間外，並無水平力作用。取 $A$ 被拋射的方向為正方向。以下所有速度均帶正負號。

### 第 (i) 部分

**$A$ 與 $B$ 的碰撞**。設 $A$ 和 $B$ 在碰撞剛結束時的速度分別為 $v_A$ 和 $v_B$。由動量守恆和恢復定律（其中 $e = \frac{1}{2}$）得

$$mv_A + km\,v_B = mu, \qquad v_B - v_A = \tfrac{1}{2}u.$$

把 $v_A = v_B - \frac{1}{2}u$ 代入第一條方程，得 $(1+k)v_B = \frac{3}{2}u$，所以

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

由於 $v_B > 0$，$B$ 確實會繼續向 $C$ 移動；又由於 $v_A < v_B$，$A$ 不可能在 $B$ 到達 $C$ 之前追上 $B$。

**$B$ 與 $C$ 的碰撞**。設 $B$ 和 $C$ 在碰撞剛結束時的速度分別為 $w_B$ 和 $w_C$。取 $e = \frac{1}{4}$，

$$km\,w_B + 3m\,w_C = km\,v_B, \qquad w_C - w_B = \tfrac{1}{4}v_B.$$

代入 $w_C = w_B + \frac{1}{4}v_B$，得 $(k+3)w_B = \left(k - \frac{3}{4}\right)v_B$，所以

$$w_B = \frac{(4k-3)\,v_B}{4(k+3)} = \frac{3(4k-3)\,u}{8(1+k)(k+3)}.$$

此外 $w_C > w_B$，所以 $B$ 和 $C$ 互相分開，此後不再相遇。這次碰撞之後，在 $A$ 或 $B$ 身上再無其他作用，直至兩者相遇為止（假如真的會相遇）。

**想一想** $A$ 與 $B$ 再次碰撞的確切條件是甚麼？

一旦 $B$ 撞上 $C$，$A$ 和 $B$ 都以恆定速度運動，而 $A$ 在 $B$ 的後方（即位於 $B$ 的負方向一側）。兩者之間的距離縮短，當且僅當 $v_A > w_B$，而這時距離會在有限時間內縮減至零。若 $v_A \le w_B$，距離永不減少。這一條不等式已涵蓋所有正負號的情況：例如，若 $A$ 正在向後運動，但 $B$ 向後運動得更快，$B$ 仍會追上 $A$，而這同樣是 $v_A > w_B$。

所以我們需要

$$\frac{(2-k)u}{2(1+k)} > \frac{3(4k-3)u}{8(1+k)(k+3)}.$$

兩邊乘以 $8(1+k)(k+3)/u$（由於 $k>0$，此數為正），得到等價的不等式

$$\begin{aligned} 4(2-k)(k+3) &> 3(4k-3) \\ 24 - 4k - 4k^2 &> 12k - 9 \\ 4k^2 + 16k - 33 &< 0 \\ (2k-3)(2k+11) &< 0. \end{aligned}$$

由於 $2k+11 > 0$，上式成立當且僅當 $2k-3<0$。因為 $k$ 是質量之比，$k>0$。因此 $A$ 與 $B$ 發生第二次碰撞，當且僅當

$$0 < k < \tfrac{3}{2}.$$

當 $k = \frac{3}{2}$ 時，我們得到 $v_A = w_B = \frac{u}{10}$：兩者以相同速率一同前進，實際上永不再相遇，所以端點不包括在內。

### 第 (ii) 部分

取 $k=1$，由上面的公式得

$$v_A = \frac{u}{4}, \qquad v_B = \frac{3u}{4}, \qquad w_B = \frac{3u}{64}, \qquad w_C = \frac{15u}{64}.$$

把原點設在 $A$ 與 $B$ 首次碰撞之處（即 $B$ 的起始位置），並由該次碰撞起計算時間 $t$。

**直至 $B$–$C$ 碰撞**。$B$ 以速率 $\frac{3}{4}u$ 走過與 $C$ 相隔的距離 $d$，所需時間為

$$t_1 = \frac{d}{3u/4} = \frac{4d}{3u}.$$

在這段時間內，$A$ 移動了 $\frac{u}{4} \cdot \frac{4d}{3u} = \frac{d}{3}$。所以 $A$ 位於 $\frac{d}{3}$，而 $B$ 位於 $d$：兩者相距 $\frac{2d}{3}$。

**追趕**。現在 $A$ 以 $\frac{u}{4} = \frac{16u}{64}$ 運動，而 $B$ 以 $\frac{3u}{64}$ 運動，所以 $A$ 以相對速率 $\frac{13u}{64}$ 拉近距離。所需時間為

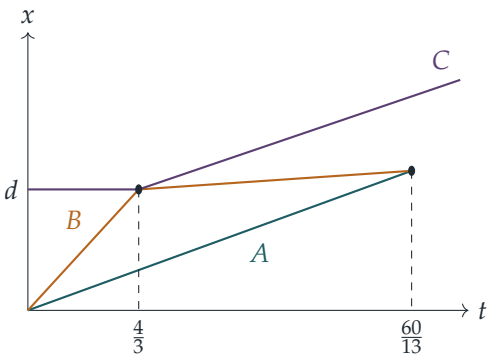
$$t_2 = \frac{2d/3}{13u/64} = \frac{128d}{39u}.$$

**總時間**。兩次 $A$–$B$ 碰撞之間相隔的時間為

$$t_1 + t_2 = \frac{4d}{3u} + \frac{128d}{39u} = \frac{52d + 128d}{39u} = \frac{180d}{39u} = \frac{60d}{13u},$$

正如所求。

下面的距離–時間圖顯示了整個運動過程，其中 $u = d = 1$。代表 $A$ 和 $B$ 的兩條線在 $t = \frac{60}{13}$ 再次相交。




## 驗算

- 動量：取 $k = 1$，第一次碰撞後 $m \cdot \frac{u}{4} + m \cdot \frac{3u}{4} = mu$，第二次碰撞後 $m \cdot \frac{3u}{64} + 3m \cdot \frac{15u}{64} = \frac{48}{64}mu = m \cdot \frac{3u}{4}$。兩者均平衡。
- 邊界 $k = \frac{3}{2}$：$v_A$ 和 $w_B$ 都等於 $\frac{u}{10}$，這印證了臨界值。
- 合理性檢查：當 $k$ 很大時，$A$ 向後反彈（$v_A < 0$），而 $B$ 繼續向前，所以顯然沒有第二次碰撞。這與所得範圍吻合。
- 我們以數值方法驗證過：對 $10^5$ 個隨機有理數 $k$，$v_A > w_B$ 成立當且僅當 $k < \frac{3}{2}$；另外，取 $u = d = 1$ 作逐步時間模擬，得出第二次碰撞發生於 $t \approx 4.6154$，與 $\frac{60}{13} \approx 4.6154$ 相符。

## 常見錯誤

- 在恢復定律中使用速率而非帶正負號的速度。使用帶正負號的速度時，分離速率就是「前者減後者」，而 $v_A$ 的正負號自然會得到妥善處理。
- 把 $v_A > 0$ 當作條件的一部分。原則上 $A$ 可以向後運動而仍被 $B$ 追上（若 $w_B < v_A$）。唯一的條件 $v_A > w_B$ 才是正確的條件。
- 在第 (ii) 部分，由第一次碰撞起計算追趕時間，而不是由 $B$ 的速度改變的一刻起計算，因而漏掉了 $\frac{d}{3}$ 這段 $A$ 早已走過的領先距離。

## STEP 之外

動量守恆與恢復定律合起來，使每一次碰撞都成為由碰撞前速度到碰撞後速度的線性映射（linear map）。一連串的碰撞，例如牛頓擺（Newton's cradle），或是計算兩個方塊與一面牆之間碰撞次數的著名問題，便可透過複合這些映射來研究。在大學一年級的力學中，這會引伸至質心參考系（centre-of-mass frame），在其中恢復定律變成相對速度的簡單縮放，並進一步推展至二維和三維的碰撞。Kleppner 與 Kolenkow 的 *An Introduction to Mechanics*（力學導論）對碰撞和質心參考系有清晰的論述。

詳細解答

# STEP II 2006，第 14 題

概率與統計 • 連續隨機變量

**STEP II 2006，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/06-s2-q14

## 關鍵思路

整道題都建基於一個原函數（antiderivative）。分母 $x\ln x$ 是 $\ln x$ 乘以 $x$，而 $1/x$ 正是 $\ln x$ 的導數，所以被積函數具有 $g'(x)/g(x)$ 的形式，其中 $g(x)=\ln x$。因此

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

此後，每一部分都只需運用「總概率為 1」，並小心留意正負號，以及密度（density）實際上在哪裏不為零。

## 解答

### 草圖

設 $y=\dfrac{1}{x\ln x}$，其中 $x>0$，$x\neq 1$。

正負號。由於 $x>0$，$y$ 的正負號與 $\ln x$ 相同：在 $0<x<1$ 上 $y<0$，在 $x>1$ 上 $y>0$。曲線與兩條坐標軸都不相交。

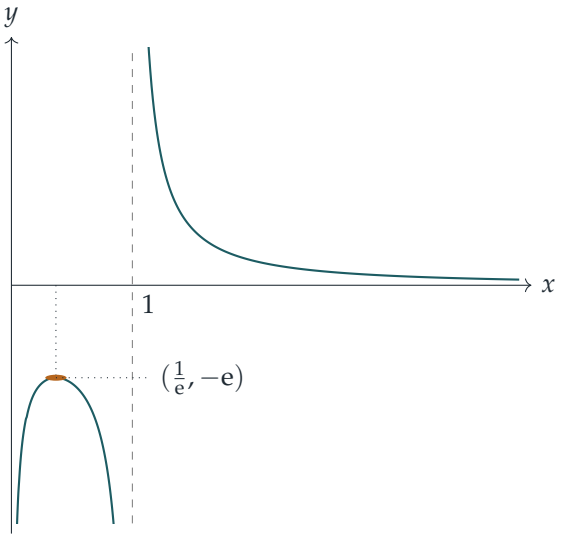
兩端的變化。

- 當 $x\to 0^+$ 時：題目告訴我們 $x\ln x\to 0$，而且在該處 $x\ln x<0$，所以 $y\to-\infty$。
- 當 $x\to 1$ 時：$x\ln x\to 0$，從左邊趨近時為負，從右邊趨近時為正，所以當 $x\to 1^-$ 時 $y\to-\infty$，當 $x\to 1^+$ 時 $y\to+\infty$。直線 $x=1$ 是垂直漸近線(vertical asymptote)。
- 當 $x\to\infty$ 時：$x\ln x\to\infty$，所以 $y\to 0^+$，而 $x$ 軸是漸近線。

轉向點（*turning points*）。由商法則（quotient rule）或鏈式法則（chain rule），

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

它只在 $\ln x = -1$ 時為零，即 $x = 1/\mathrm{e}$，此時 $y = \dfrac{1}{(1/\mathrm{e})(-1)} = -\mathrm{e}$。當 $0 < x < 1/\mathrm{e}$ 時，我們有 $\ln x + 1 < 0$，所以 $y$ 遞增；當 $x > 1/\mathrm{e}$（且 $x \neq 1$）時，函數遞減。因此 $(1/\mathrm{e}, -\mathrm{e})$ 是左邊分支的局部極大值（local maximum）點，而右邊分支則由 $+\infty$ 一直遞減至 0。




**第 (i) 部分**

這裏 $a = \frac{1}{4}$ 而 $b = \frac{1}{2}$，兩者都在 $(0,1)$ 內，而在該區間上 $x\ln x < 0$。利用上面的原函數，

$$\begin{aligned}\int_{1/4}^{1/2} \frac{\mathrm{d}x}{x\ln x} &= \Big[\ln|\ln x|\Big]_{1/4}^{1/2} = \ln|\ln\tfrac{1}{2}| - \ln|\ln\tfrac{1}{4}| \\ &= \ln(\ln 2) - \ln(2\ln 2) = -\ln 2.\end{aligned}$$

總概率必須為 1，所以 $\lambda\cdot(-\ln 2) = 1$，由此得

$$\lambda = -\frac{1}{\ln 2}.$$

這個負號正是我們所需要的：在 $[\frac{1}{4}, \frac{1}{2}]$ 上 $x\ln x < 0$，所以 $\lambda$ 必須為負，f 在該處才會非負。

### 第 (ii) 部分

當 $\lambda = 1$ 時，密度必須非負，所以區間 $[a,b]$ 必須位於 $x\ln x > 0$ 的範圍內；由於 $b > a > 1$，這一點確實成立。於是 $\ln a, \ln b > 0$，並且

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

所以 $\dfrac{\ln b}{\ln a} = \mathrm{e}$，即 $\ln b = \mathrm{e}\ln a = \ln\left(a^{\mathrm{e}}\right)$；又因為 ln 是一對一（one-to-one）的，故 $b = a^{\mathrm{e}}$，正如所求。

### 第 (iii) 部分

取 $a = \mathrm{e}$，由第 (ii) 部分得 $b = \mathrm{e}^{\mathrm{e}}$。由於 $1 < \frac{3}{2} < 2 < \mathrm{e}$，整個區間 $[\mathrm{e}^{3/2}, \mathrm{e}^2]$ 都位於 $[a,b] = [\mathrm{e}, \mathrm{e}^{\mathrm{e}}]$ 之內，所以

$$\mathrm{P}(\mathrm{e}^{3/2} \le X \le \mathrm{e}^2) = \Big[\ln(\ln x)\Big]_{\mathrm{e}^{3/2}}^{\mathrm{e}^2} = \ln 2 - \ln\tfrac{3}{2} = \ln\tfrac{4}{3}.$$

要把這個結果化成分數，可使用麥克勞林級數（Maclaurin series）$\ln(1+t) = t - \frac{t^2}{2} + \frac{t^3}{3} - \frac{t^4}{4} + \cdots$，取 $t = \frac{1}{3}$，並保留四項：

$$\ln\tfrac{4}{3} \approx \frac{1}{3} - \frac{1}{18} + \frac{1}{81} - \frac{1}{324} = \frac{108 - 18 + 4 - 1}{324} = \frac{93}{324} = \frac{31}{108}.$$

**想一想** 我們怎樣知道取到第四項就停下來是合理的？

這個級數正負項交替出現，而且各項的大小遞減，所以取四項之後的誤差小於第五項，即 $\frac{1}{5\cdot 3^5} = \frac{1}{1215} < 0.001$。因此 $\mathrm{P} = \ln\frac{4}{3} \approx \frac{31}{108}$，正如所求。

### 第 (iv) 部分

現在 $a = \mathrm{e}^{1/2}$，所以由第 (ii) 部分得

$$b = a^{\mathrm{e}} = \mathrm{e}^{\mathrm{e}/2}.$$

**想一想** 在積分之前，先想想：密度實際上在哪裏不為零？

由於 $\mathrm{e} < 3$，我們有 $\mathrm{e}/2 < \frac{3}{2}$，所以 $b = \mathrm{e}^{\mathrm{e}/2} < \mathrm{e}^{3/2}$。區間 $[\mathrm{e}^{3/2}, \mathrm{e}^2]$ 完全位於 $b$ 的右邊，而在那裏 $\mathrm{f}(x) = 0$。因此

$$\mathrm{P}(\mathrm{e}^{3/2} \le X \le \mathrm{e}^2) = 0.$$

（若不加思索地把 $1/(x\ln x)$ 由 $\mathrm{e}^{3/2}$ 積分至 $\mathrm{e}^2$，便會再次得到 $\ln\frac{4}{3}$，而這是錯的。）

## 驗算

- 以數值計算，$e^{-1} \approx 0.368$，而在 $(0,1)$ 上逐點檢視 $y$ 的值，可見它在該區間上的最大值 $\approx -2.718 = -e$。
- 辛普森法則(Simpson's rule)給出 $\int_{1/4}^{1/2} \frac{dx}{x \ln x} \approx -0.6931 = -\ln 2$，所以 $\lambda \approx -1.4427$。
- 對於若干個 $a > 1$ 的值（例如 $a = 2$、$a = 7.3$），$\int_a^{a^e} \frac{dx}{x \ln x}$ 的結果都是 1.0000。
- $\ln \frac{4}{3} \approx 0.28768$ 而 $\frac{31}{108} \approx 0.28704$，兩者相差不超過 0.0007。
- 在 (iv) 中，$b = e^{e/2} \approx 3.89$，而 $e^{3/2} \approx 4.48$。

## 常見錯誤

- 在第 (i) 部分寫成 $\ln(\ln x)$，但那裏 $\ln x < 0$：原函數需要加上絕對值（modulus），即 $\ln|\ln x|$。
- 答 $\lambda = \frac{1}{\ln 2}$：正的 $\lambda$ 會令「密度」變成負數。
- 在第 (iv) 部分，沒有先檢查 $[e^{3/2}, e^2]$ 是否位於支撐集（support）$[a,b]$ 之內，便直接在這個區間上積分。

## STEP 之外

代換(substitution)$u = \ln x$ 把這個密度變成 $[\ln a, \ln b]$ 上的 $1/u$，而條件 $b = a^e$ 其實就是 $\int_{\ln a}^{\ln b} \frac{du}{u} = 1$。像這樣對隨機變量（random variable）作變換（這裏是 $U = \ln X$），是大學一年級概率論的核心工具；在那裏，隨機變量的函數的密度是由鏈式法則求出的。同一個函數 $1/(x \ln x)$ 在分析學中也十分有名：$\sum 1/(n \ln n)$ 發散（diverges），這可以透過與這個積分比較而得出，儘管它的各項趨向零。關於隨機變量的變換，可參閱 Grimmett 與 Welsh 的 *Probability: An Introduction*。

詳細解答

# STEP III 2006，第 1 題

純數學 • 曲線描繪

> **STEP III 2006，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/06-s3-q1

## 關鍵思路

先做除法：把曲線寫成一條直線加上一個細小的修正項，便可得出斜漸近線（oblique asymptote），知道曲線位於漸近線的哪一側，而且求導數（derivative）也變得容易。結果導數處處大於 2，而單憑這一個事實，便足以完成數根部分的大半工作：每條方程都可以改寫成「曲線 = 某個簡單的式子」，然後比較斜率。

## 解答

### 曲線

由於 $2x^3 - 10x = 2x(x^2 - 4) - 2x$，

$$y = \frac{2x(x^2-5)}{x^2-4} = 2x - \frac{2x}{x^2-4}.$$

有些特徵可以即時看出：

- $y$ 是奇函數（odd function），所以曲線關於原點有 2 重旋轉對稱（rotational symmetry）。
- 曲線在 $x = 0$ 和 $x = \pm\sqrt{5}$ 處與 $x$ 軸相交。
- **漸近線**。分母在 $x = \pm 2$ 處為零（該處分子為 $\pm 4 \cdot (-1) \cdot 2 \neq 0$），所以 $x = 2$ 和 $x = -2$ 是鉛垂漸近線（vertical asymptote）。當 $|x| \to \infty$ 時，修正項 $2x/(x^2-4) \to 0$，所以 $y = 2x$ 是斜漸近線。
- **在漸近線的哪一側**。$y - 2x = -2x/(x^2-4)$。當 $x > 2$ 時，此式為負（曲線略低於 $y = 2x$）；當 $x < -2$ 時為正（略高於它）。當 $0 < x < 2$ 時為正，當 $-2 < x < 0$ 時

為負。

- **在 $x = \pm 2$ 附近的表現**。當 $x \to 2^+$ 時，$-2x/(x^2-4) \to -\infty$，而當 $x \to 2^-$ 時它趨向 $+\infty$；由奇函數的性質，當 $x \to -2^-$ 時 $y \to +\infty$，當 $x \to -2^+$ 時 $y \to -\infty$。

對除法後的形式求導，

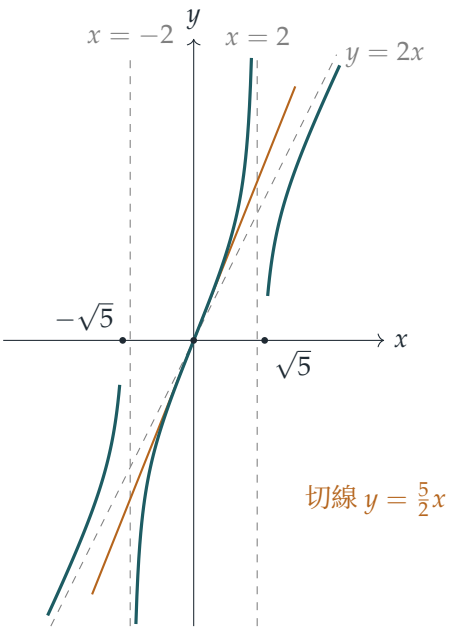
$$\frac{\mathrm{d}y}{\mathrm{d}x} = 2 - \frac{2(x^2-4) - 2x \cdot 2x}{(x^2-4)^2} = 2 + \frac{2(x^2+4)}{(x^2-4)^2}.$$

所以對任何 $x \neq \pm 2$ 都有 $\mathrm{d}y/\mathrm{d}x > 2$：曲線沒有駐點（stationary point），而且在 $x < -2$、$-2 < x < 2$、$x > 2$ 這三個區間的每一個上，曲線都嚴格遞增（strictly increasing），由 $-\infty$ 升至 $+\infty$。

**原點處的切線（tangent）**。在 $x = 0$ 處斜率為 $2 + 8/16 = \frac{5}{2}$，所以切線是 $y = \frac{5}{2}x$。再求導一次，

$$\frac{\mathrm{d}^2y}{\mathrm{d}x^2} = -\frac{4x(x^2+12)}{(x^2-4)^3},$$

它只在 $x = 0$ 處為零，並且在該處變號：原點是拐點（point of inflection）。在 $0 < x < 2$ 上二階導數（second derivative）為正（分子和分母皆為負），所以中間的分支在該處是凸（convex）的，位於其切線 $y = \frac{5}{2}x$ 的上方；在 $-2 < x < 0$ 上它是凹（concave）的，位於切線的下方。




**小結**。漸近線為 $x = 2$、$x = -2$ 和 $y = 2x$；原點處的切線為 $y = \frac{5}{2}x$。

在以下每一部分中，$x = \pm 2$ 都不會是根（該處左方不為零而右方為零），所以我們可以除以 $x^2 - 4$，然後數算曲線與另一個圖像的交點。

### 第 (i) 部分

除以 $\frac{3}{2}(x^2-4)$，方程變成

$$\frac{2x(x^2-5)}{x^2-4} = \tfrac{2}{3}(x+3),$$

所以我們要找曲線與直線 $y = \frac{2}{3}x + 2$ 的交點，該直線的斜率為 $\frac{2}{3}$。

> **想一想** 為甚麼斜率小於 2 的直線與每個分支都恰好相交一次？

設 $h(x)$ 為曲線的 $y$ 值減去直線的對應值。在每個分支上 $h'(x) > 2 - \frac{2}{3} > 0$，所以 $h$ 嚴格遞增，最多只有一個零點。在靠近鉛垂漸近線的兩端，$h \to \pm\infty$，與 $y$ 的情況一樣（直線保持有限）。在外側的兩端，$h \approx 2x - \frac{2}{3}x = \frac{4}{3}x$，當 $x \to -\infty$ 時趨向 $-\infty$，當 $x \to +\infty$ 時趨向 $+\infty$。所以在三個分支的每一個上，$h$ 都由 $-\infty$ 走到 $+\infty$，根據中間值定理（intermediate value theorem），它恰好有一個零點。

因此方程 (i) 有三個實根。（事實上它就是三次方程 $2x^3 - 3x^2 - 11x + 12 = 0$，其中一個根是 $x = 1$。）

### 第 (ii) 部分

除以 $2(x^2-4)$，我們需要找曲線與 $y = \frac{5}{2}x - 1$ 的交點：這是把原點處的切線向下平移 1。它的斜率現在大於 2，所以 (i) 的論證不再適用，我們要逐個分支考察。

- $x > 2$：曲線位於 $y = 2x$ 的下方，而直線滿足 $\frac{5}{2}x - 1 - 2x = \frac{1}{2}x - 1 > 0$，所以它位於 $y = 2x$ 的上方。沒有交點。
- $x < -2$：曲線位於 $y = 2x$ 的上方，而由於 $\frac{1}{2}x - 1 < 0$，直線位於其下方。沒有交點。
- $0 \le x < 2$：曲線位於其切線 $y = \frac{5}{2}x$ 之上或恰在其上（凸性），而切線又在直線的上方。沒有交點。
- $-2 < x < 0$：當 $x \to -2^+$ 時，曲線趨向 $-\infty$，而直線趨向 $-6$；在 $x = 0$ 處曲線的值為 0，高於直線的 $-1$。所以至少有一個交點。因為這裏 $y'' < 0$，$y'$ 在 $(-2, 0)$ 上遞減，所以 $y' > y'(0) = \frac{5}{2}$；「曲線減直線」這個差嚴格遞增，故交點唯一。

因此方程 (ii) 恰好有一個實根，位於 $-2$ 與 0 之間。（寫成三次方程就是 $x^3 - 2x^2 + 8 = 0$，它的駐點值 8 和 $\frac{184}{27}$ 都是正數，由此確認只有一個實根，約在 $x \approx -1.51$ 附近。）

### 第 (iii) 部分

除以 $(x^2-4)^2$ 得

$$\left(\frac{2x(x^2-5)}{x^2-4}\right)^2 = x^2+1,$$

所以我們需要找曲線與等軸雙曲線（rectangular hyperbola）$y^2-x^2=1$ 的交點，即與 $y=\sqrt{x^2+1}$（上支）或 $y=-\sqrt{x^2+1}$（下支）的交點。

> **想一想** 可以利用對稱性把工作減半嗎？

可以。映射 $(x,y)\mapsto(-x,-y)$ 把曲線映回自身（因為它是奇函數），並把雙曲線的兩支互換。所以曲線與下支的交點和它與上支的交點一一對應，我們只需數算後者。

設 $u(x)=\sqrt{x^2+1}$。那麼 $|u'(x)| = |x|/\sqrt{x^2+1} < 1$，所以在曲線的每個分支上，差 $y-u$ 的導數都大於 $2-1=1>0$，每個分支最多給出一個交點。現在檢查各端：

- $-2<x<2$：$y-u$ 由 $-\infty$ 走到 $+\infty$。一個根。
- $x>2$：當 $x\to 2^+$ 時 $y-u\to-\infty$，而當 $x\to\infty$ 時 $y-u\approx 2x-x\to+\infty$。一個根。
- $x<-2$：當 $x\to-\infty$ 時 $y-u\approx 2x-|x|=3x\to-\infty$，而當 $x\to-2^-$ 時 $y-u\to+\infty$。一個根。

所以曲線與上支相交三次，而根據對稱性，與下支也相交三次。兩支沒有公共點（$u\geq 1$），所以這六個點有六個不同的 $x$ 值，方程 (iii) 有**六**個實根。由於它是一條 6 次多項式方程（即 $3x^6-33x^4+92x^2-16=0$），它的所有根都是實數。

## 驗算

- 在 $x=0$ 附近，$y\approx 2x(-5)/(-4)=\frac{5}{2}x$，與切線吻合。
- 在 $x=1000$ 處，$y-2x\approx -0.002$：數值細小而且為負，與預期相符。
- 對原來的方程以數值方法搜尋變號位置，可找到 (i) 的根 $x\approx -2.212,\ 1,\ 2.713$；(ii) 只有 $x\approx -1.510$；而 (iii) 的根為 $x\approx \pm 0.431,\ \pm 2.155,\ \pm 2.484$。

## 常見錯誤

- 在除以 $x^2-4$ 之前，忘記檢查 $x=\pm 2$ 不是根。
- 在 (iii) 中，只數算與 $y=\sqrt{x^2+1}$ 的交點，遺漏了負平方根。

- 在 (ii) 中，像 (i) 那樣假設「直線夠斜，所以每個分支一個根」：當斜率大於 2 時，外側的分支需要另作論證。

## STEP 之外

這裏的數根方法，是中間值定理與單調性（monotonicity）的結合：一個連續、嚴格遞增而且由 $-\infty$ 走到 $+\infty$ 的函數，恰好有一個零點。在大學一年級的分析課程中，這兩個事實都會從實數的完備性（completeness）出發嚴格證明，而同樣的推理也是「每個奇數次實係數多項式都有一個實根」這類結果的基礎。Abbott 的 *Understanding Analysis* 對此有非常精彩的闡述。

詳細解答

# STEP III 2006，第 2 題

純數學 • 利用對稱性求定積分

**STEP III 2006，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/06-s3-q2

## 關鍵思路

區間 $[-\frac{1}{2}\pi, \frac{1}{2}\pi]$ 是對稱的，所以把 $\theta$ 換成 $-\theta$ 便得到 $I$ 的第二個表達式；兩式相加，分母中那個礙事的 $\sin\theta$ 就會消失。一旦把所有式子都用 $\tan\theta$ 表示，積分（integral）$I$ 和 $J$ 便有相同的分母，而兩者經適當選取的組合，其被積函數（integrand）恰好是 1。整個過程中，要時刻留意 $\cos 2\alpha$ 的正負號：最後一部分的變化正在於此。

## 解答

設 $s = \sin 2\alpha$ 及 $c = \cos 2\alpha$，則 $s^2 + c^2 = 1$。在所考慮的兩個 $\alpha$ 範圍內（$\alpha \neq \frac{1}{4}\pi$），都有 $0 < s < 1$，所以 $1 \mp s\sin\theta \geq 1 - s > 0$：$I$ 的被積函數在整個區間上連續（continuous）。

### 第 (i) 部分

作代換（substitution）$\theta = -\phi$，則 $\mathrm{d}\theta = -\mathrm{d}\phi$，而上下限互換。由於 $\cos^2(-\phi) = \cos^2\phi$ 及 $\sin(-\phi) = -\sin\phi$，

$$I = \int_{\frac{1}{2}\pi}^{-\frac{1}{2}\pi} \frac{\cos^2\phi}{1 + \sin\phi\, s}\,(-\mathrm{d}\phi) = \int_{-\frac{1}{2}\pi}^{\frac{1}{2}\pi} \frac{\cos^2\phi}{1 + s\sin\phi}\,\mathrm{d}\phi,$$

這就是第一個結果（把 $\phi$ 改寫為 $\theta$ 即可）。

把 $I$ 的兩個表達式相加，

$$2I = \int_{-\frac{1}{2}\pi}^{\frac{1}{2}\pi} \cos^2\theta \left(\frac{1}{1 - s\sin\theta} + \frac{1}{1 + s\sin\theta}\right)\mathrm{d}\theta = \int_{-\frac{1}{2}\pi}^{\frac{1}{2}\pi} \frac{2\cos^2\theta}{1 - s^2\sin^2\theta}\,\mathrm{d}\theta.$$

現在把分子和分母同時除以 $\cos^2\theta$（在區間內部不為零）：

$$\frac{2\cos^2\theta}{1-s^2\sin^2\theta}=\frac{2}{\sec^2\theta-s^2\tan^2\theta}=\frac{2}{1+\tan^2\theta\,(1-s^2)}=\frac{2}{1+c^2\tan^2\theta},$$

其中用到 $\sec^2\theta=1+\tan^2\theta$ 及 $1-\sin^2 2\alpha=\cos^2 2\alpha$。因此

$$2I=\int_{-\frac12\pi}^{\frac12\pi}\frac{2}{1+\tan^2\theta\cos^2 2\alpha}\,\mathrm{d}\theta,$$

正如所求。（在 $\theta=\pm\frac12\pi$ 處，這個被積函數趨向 0，所以在兩個端點不會出問題。）

**第 (ii) 部分**

分子中的 $\sec^2\theta$ 提示我們採用 $t=\tan\theta$，$\mathrm{d}t=\sec^2\theta\,\mathrm{d}\theta$。當 $\theta$ 由 $-\frac12\pi$ 變到 $\frac12\pi$ 時，$t$ 走遍整條實數線：

$$J=\int_{-\infty}^{\infty}\frac{\mathrm{d}t}{1+c^2t^2}.$$

由於 $0<\alpha<\frac14\pi$，我們有 $0<2\alpha<\frac12\pi$，所以 $c>0$。設 $u=ct$，

$$J=\frac1c\int_{-\infty}^{\infty}\frac{\mathrm{d}u}{1+u^2}=\frac1c\Big[\arctan u\Big]_{-\infty}^{\infty}=\frac{\pi}{c}=\pi\sec 2\alpha.$$

**第 (iii) 部分**

由第 (i) 部分，$I=\displaystyle\int_{-\frac12\pi}^{\frac12\pi}\frac{\mathrm{d}\theta}{1+c^2\tan^2\theta}$，它與 $J$ 有相同的分母。所以

$$s^2I+c^2J=\int_{-\frac12\pi}^{\frac12\pi}\frac{s^2+c^2\sec^2\theta}{1+c^2\tan^2\theta}\,\mathrm{d}\theta.$$

**?** **想一想** 為甚麼偏偏這個組合會比較容易處理？

看看分子：$s^2+c^2\sec^2\theta=s^2+c^2+c^2\tan^2\theta=1+c^2\tan^2\theta$。它恰好就是分母，所以被積函數是 1，而

$$s^2I+c^2J=\pi.$$

利用第 (ii) 部分的 $J=\pi/c$，

$$s^2I=\pi-\pi c=\pi(1-c)\quad\Longrightarrow\quad I=\frac{\pi(1-c)}{1-c^2}=\frac{\pi}{1+c},$$

其中用到 $s^2=1-c^2=(1-c)(1+c)$，並約去了 $1-c\neq0$（因為 $c<1$）。最後，$1+\cos 2\alpha=2\cos^2\alpha$，所以

$$I=\frac{\pi}{2\cos^2\alpha}=\tfrac12\pi\sec^2\alpha.\qquad\checkmark$$

### 第 (iv) 部分

現在 $\frac{1}{4}\pi < \alpha < \frac{1}{2}\pi$，所以 $\frac{1}{2}\pi < 2\alpha < \pi$：仍然有 $0 < s < 1$，但這次 $c < 0$。第 (i) 部分只用到 $s^2 + c^2 = 1$，所以依然成立；第 (iii) 部分的恆等式（identity）$s^2 I + c^2 J = \pi$ 亦然。唯一改變的是第 (ii) 部分。

> **想一想** 第 (ii) 部分究竟在哪裏用到了 $\alpha < \frac{1}{4}\pi$？

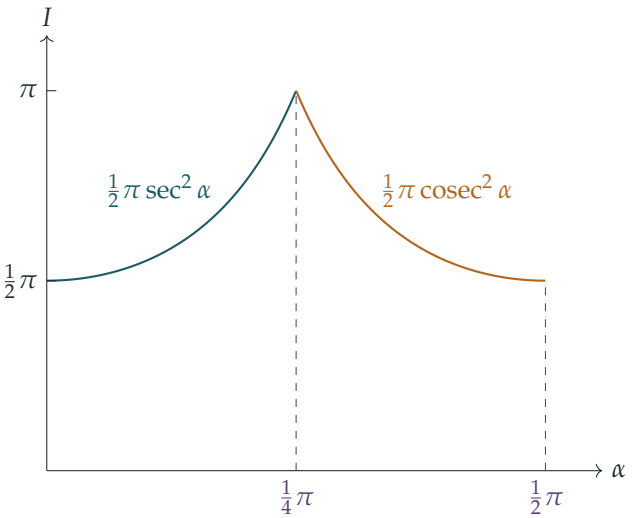
就在代換 $u = ct$ 之中：若 $c < 0$，$u$ 的上下限會變成由 $+\infty$ 到 $-\infty$，把它們換回來便多出一個負號。換句話說，$J = \pi/|c| = -\pi/c$。（這是必然的：$J$ 是一個正函數的積分，所以不可能等於負數 $\pi \sec 2\alpha$。）於是

$$s^2 I = \pi - c^2\Big(-\frac{\pi}{c}\Big) = \pi(1+c) \quad\Longrightarrow\quad I = \frac{\pi(1+c)}{(1-c)(1+c)} = \frac{\pi}{1-c},$$

其中約去了 $1 + c \neq 0$（因為 $c > -1$）。由於 $1 - \cos 2\alpha = 2\sin^2\alpha$，

$$I = \frac{\pi}{2\sin^2\alpha} = \tfrac{1}{2}\pi \operatorname{cosec}^2\alpha \qquad \big(\tfrac{1}{4}\pi < \alpha < \tfrac{1}{2}\pi\big).$$

有一個快捷的方法可以核實：$I$ 只透過 $\sin 2\alpha$ 依賴於 $\alpha$，而 $\sin 2(\frac{1}{2}\pi - \alpha) = \sin 2\alpha$。所以 $I(\alpha) = I(\frac{1}{2}\pi - \alpha) = \frac{1}{2}\pi \sec^2(\frac{1}{2}\pi - \alpha) = \frac{1}{2}\pi \operatorname{cosec}^2\alpha$。$I$ 的圖像關於 $\alpha = \frac{1}{4}\pi$ 對稱，而兩條公式在該處都給出 $\pi$：




## 驗算

- 在 $\alpha = \frac{1}{4}\pi$（$c = 0$）時，第 (i) 部分的 2I 形式的被積函數是 2，得出 $I = \pi$，與兩條公式都吻合。

- 當 $\alpha \to 0$ 時，$I \to \int \cos^2\theta \,\mathrm{d}\theta = \frac{1}{2}\pi$，與 $\frac{1}{2}\pi \sec^2 0$ 吻合。
- 數值計算方面（對原來的積分使用辛普森法則（Simpson's rule）），$\alpha = 0.5$ 給出 $I \approx 2.039595$，等於 $\frac{1}{2}\pi \sec^2 0.5$；而 $\alpha = 1.1$ 給出 $I \approx 1.977709 = \frac{1}{2}\pi \operatorname{cosec}^2 1.1$。在所有測試過的情況中，$s^2 I + c^2 J$ 的結果都是 3.141593。

## 常見錯誤

- 在第 (iv) 部分仍然寫 $J = \pi \sec 2\alpha$：這個值在該範圍內是負數，而正函數的積分不可能是負數。
- 忘記了代換 $t = \tan\theta$ 會把有限區間變成整條實數線，所以上下限變成 $\pm\infty$。

## STEP 之外

第 (i) 部分的技巧，即取 $f(\theta)$ 與 $f(-\theta)$ 的平均，正是把函數分解為偶部和奇部（even and odd parts）的起點：在對稱區間上積分，只有偶部會留下來。$\sin\theta$ 和 $\cos\theta$ 的有理函數（rational function）在一個完整週期上的積分，是大學第一門複分析（complex analysis）課程中圍道積分（contour integration，即留數定理 the residue theorem）的經典應用，像本題這樣的結果在那裏寥寥數行便可得出；參閱 Needham 的 *Visual Complex Analysis*（《可視化複分析》）。

詳細解答

# STEP III 2006，第 3 題

純數學 • 冪級數的係數

**STEP III 2006，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/06-s3-q3

## 關鍵思路

我們完全不必反覆對 $\tan x$ 求導。每一條三角恆等式（trigonometric identity）都把其中兩個函數連繫起來，因而化為它們的係數（coefficient）之間的一條關係式。只要比較 $x^n$ 的係數，三個級數就全由 $b_n$ 這一個數列決定。接着，畢氏恆等式（Pythagorean identity）$\cot^2 x + 1 = \operatorname{cosec}^2 x$ 給出 $b_n$ 必須滿足的方程，我們可以逐個解出。

## 解答

### 第 (i) 部分

若一個冪級數（power series）在 0 附近收斂（converge）至某函數，它就是該函數的麥克勞林級數（Maclaurin series），所以 $a_n = \tan^{(n)}(0)/n!$。而 tan 是奇函數（odd function）：$\tan(-x) = -\tan x$。對奇函數求導得到偶函數（even function），對偶函數求導則得到奇函數。因此 tan 的每個偶數階導數都是奇函數，而奇函數在 0 處的值為零。故此當 $n$ 為偶數時 $\tan^{(n)}(0) = 0$，於是當 $n$ 為偶數時 $a_n = 0$。換言之，這個級數只含 $x$ 的奇數次冪。同樣的論證適用於 $\cot x - \frac{1}{x} = \sum b_n x^n$，它也是奇函數，所以當 $n$ 為偶數時 $b_n = 0$。我們稍後亦會從下面的關係式得出這個結論。

至於恆等式，把各項通分，以 $\sin x \cos x$ 為分母：

$$\cot x - \tan x = \frac{\cos^2 x - \sin^2 x}{\sin x \cos x} = \frac{\cos 2x}{\frac{1}{2}\sin 2x} = 2\cot 2x.$$

現在取 $0 < x < \frac{1}{4}\pi$，使 $2x$ 同樣落在 $(0, \frac{1}{2}\pi)$ 之內，於是可以使用 $\cot 2x$ 的級數，即以

$2x$ 代替 $x$：

$$2\cot 2x = 2\left(\frac{1}{2x} + \sum_{n=0}^{\infty} b_n 2^n x^n\right) = \frac{1}{x} + \sum_{n=0}^{\infty} 2^{n+1} b_n x^n.$$

恆等式的左方是

$$\cot x - \tan x = \frac{1}{x} + \sum_{n=0}^{\infty} (b_n - a_n) x^n.$$

兩邊的 $\frac{1}{x}$ 項互相抵消，餘下兩個冪級數，它們對 $(0, \frac{1}{4}\pi)$ 內的每個 $x$ 都相等。

**想一想** 為甚麼我們可以比較兩個無窮級數的係數？

若兩個冪級數在整個區間上有相同的和，它們的係數便一一相等。把兩者相減：所得的差是一個在整個區間上恆等於零的冪級數。於是它的各階導數在該區間上都是零，再令 $x \to 0^+$，便知每個係數都是零。所以

$$b_n - a_n = 2^{n+1} b_n, \qquad \text{即} \qquad a_n = (1 - 2^{n+1})\, b_n,$$

這正是所要證明的。由於 $1 - 2^{n+1} \neq 0$，我們有 $b_n = a_n/(1 - 2^{n+1})$，所以每當 $a_n = 0$ 時 $b_n = 0$，也就是當 $n$ 為偶數時。

### 第 (ii) 部分

同理，

$$\cot x + \tan x = \frac{\cos^2 x + \sin^2 x}{\sin x \cos x} = \frac{1}{\frac{1}{2}\sin 2x} = 2\operatorname{cosec} 2x.$$

當 $0 < x < \frac{1}{4}\pi$ 時，在 cosec 的級數中以 $2x$ 代替 $x$，得

$$2\operatorname{cosec} 2x = \frac{1}{x} + \sum_{n=0}^{\infty} 2^{n+1} c_n x^n, \qquad \cot x + \tan x = \frac{1}{x} + \sum_{n=0}^{\infty} (b_n + a_n) x^n.$$

比較係數得 $b_n + a_n = 2^{n+1} c_n$。代入第 (i) 部分的 $a_n = (1 - 2^{n+1}) b_n$，

$$2^{n+1} c_n = (2 - 2^{n+1})\, b_n \quad \Longrightarrow \quad c_n = \frac{2 - 2^{n+1}}{2^{n+1}}\, b_n = (2^{-n} - 1)\, b_n,$$

這就是所求的結果。特別地，當 $n$ 為偶數時同樣有 $c_n = 0$。

### 第 (iii) 部分

把 $\cot x$ 和 $\operatorname{cosec} x$ 的級數乘以 $x$，

$$1 + x\sum_{n=0}^{\infty} b_n x^n = x\cot x, \qquad 1 + x\sum_{n=0}^{\infty} c_n x^n = x\operatorname{cosec} x.$$

所以題目所述的等式就是 $x^2\cot^2 x + x^2 = x^2\operatorname{cosec}^2 x$。這是恆等式 $\cot^2 x + 1 = \operatorname{cosec}^2 x$ 乘以 $x^2$ 的結果，而該恆等式則由 $\cos^2 x + \sin^2 x = 1$ 除以 $\sin^2 x$ 得出。等式由此得證。

現在把它展開。下標為偶數的 $b_n$ 和 $c_n$ 都是零，所以

$$x\cot x = 1 + b_1x^2 + b_3x^4 + \cdots, \qquad x\operatorname{cosec} x = 1 + c_1x^2 + c_3x^4 + \cdots,$$

而等式變成

$$\left(1 + b_1x^2 + b_3x^4 + \cdots\right)^2 + x^2 = \left(1 + c_1x^2 + c_3x^4 + \cdots\right)^2.$$

把兩邊平方展開，

$$1 + (2b_1 + 1)x^2 + (2b_3 + b_1^2)x^4 + \cdots = 1 + 2c_1x^2 + (2c_3 + c_1^2)x^4 + \cdots.$$

由第 (ii) 部分，$c_1 = (\frac{1}{2} - 1)b_1 = -\frac{1}{2}b_1$ 及 $c_3 = (\frac{1}{8} - 1)b_3 = -\frac{7}{8}b_3$。

$x^2$ 的係數：

$$2b_1 + 1 = 2c_1 = -b_1 \quad\Longrightarrow\quad b_1 = -\tfrac{1}{3}.$$

於是由第 (i) 部分得 $a_1 = (1 - 2^2)b_1 = (-3)(-\frac{1}{3}) = 1$，正如所求。

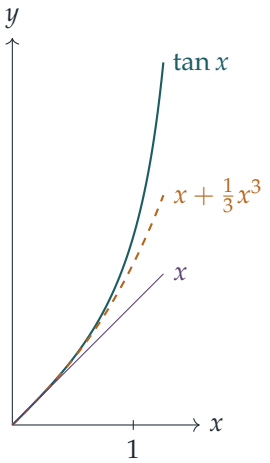
$x^4$ 的係數：代入 $b_1 = -\frac{1}{3}$ 及 $c_1 = \frac{1}{6}$，

$$2b_3 + \tfrac{1}{9} = -\tfrac{7}{4}b_3 + \tfrac{1}{36} \quad\Longrightarrow\quad \tfrac{15}{4}\,b_3 = -\tfrac{3}{36} = -\tfrac{1}{12} \quad\Longrightarrow\quad b_3 = -\tfrac{1}{45}.$$

因此

$$a_3 = (1 - 2^4)\,b_3 = (-15)\left(-\tfrac{1}{45}\right) = \tfrac{1}{3}.$$

所以 $\tan x = x + \frac{1}{3}x^3 + \cdots$。




## 驗算

- 在 $x = 0.3$ 處，$\tan 0.3 \approx 0.309336$，而 $0.3 + \frac{1}{3}(0.3)^3 = 0.309$。兩者相差約 $3 \times 10^{-4}$，正好是下一項 $\frac{2}{15}x^5$ 的大小。

- $b_1 = -\frac{1}{3}$ 這個值與熟悉的 $\cot x \approx \frac{1}{x} - \frac{x}{3}$ 吻合。此外 $c_1 = \frac{1}{6}$ 與 $\operatorname{cosec} x \approx \frac{1}{x} + \frac{x}{6}$ 吻合，後者可由 $\sin x \approx x - \frac{x^3}{6}$ 得出。
- 我們把 sin 和 cos 的級數相除，以有理數運算精確地驗證了三條關係式對 48 以內的每個 $n$ 都成立。我們亦驗證了在 $(0,1)$ 內隨機選取的點上，部分和（partial sum）與 tan、cot 和 cosec 的值相符。

## 常見錯誤

- 比較係數時忘記了 $\frac{1}{x}$ 項。這些項確實會抵消，但全靠 $2\cot 2x$ 和 $2\operatorname{cosec} 2x$ 中的因子 2。
- 在第 (iii) 部分展開平方時，忘記了 $b_0 = b_2 = 0$。若保留 $b_0$，便會多出本不存在的 $x$ 的奇數次冪，你還得另行證明它們的係數為零。
- 計算 $c_3 = -\frac{7}{8}b_3$ 時弄錯正負號或 2 的冪次。留意 $2^{-3} - 1 = -\frac{7}{8}$。

## STEP 之外

把這個方法推演到更高階，就會發現這些係數原來是伯努利數（Bernoulli numbers）$B_{2k}$：$x\cot x = \sum_k (-1)^k 2^{2k} B_{2k}\, x^{2k}/(2k)!$。第 (i) 部分的因子 $1 - 2^{n+1}$ 隨即給出 $\tan x$ 各係數的一般公式。為甚麼 $\tan x$ 的級數到了 $|x| = \frac{1}{2}\pi$ 就不再收斂？複分析（complex analysis）回答了這個問題：冪級數的收斂範圍一直延伸到函數趨向無窮的最近一點，對 tan 而言，就是位於 $\pm\frac{1}{2}\pi$ 的極點（pole）。Needham 的 *Visual Complex Analysis* 對此有精彩的解說，而 Abbott 的 *Understanding Analysis* 則為冪級數的逐項比較提供了嚴格的論證。

詳細解答

# STEP III 2006，第 4 題

純數學 • 函數方程

> **STEP III 2006，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/06-s3-q4

## 關鍵思路

開首的恆等式(identity)就是柯西函數方程(Cauchy's functional equation)。代入 $y = x$，便得到 f 在 $x$ 與在 $2x$ 的值之間的關係；把這個關係求導 $n$ 次，再令 $x = 0$，麥克勞林級數（Maclaurin series）的係數便幾乎全部消失。其後每一部分都是同一條方程的變奏：只要作一個變數代換（change of variable）（對數、指數、正切），新的恆等式就會變回 $(*)$，於是可以直接套用已經得到的答案。

## 解答

### 基本方程

在 $(*)$ 中代入 $y = x$：$\mathrm{f}(x) + \mathrm{f}(x) = \mathrm{f}(2x)$，即

$$2\mathrm{f}(x) \equiv \mathrm{f}(2x).$$

這是關於 $x$ 的恆等式，所以兩邊可以求導。根據鏈式法則（chain rule），每對 $\mathrm{f}(2x)$ 求導一次，就多出一個因子 2，因此求導 $n$ 次之後得

$$2\mathrm{f}^{(n)}(x) \equiv 2^n \mathrm{f}^{(n)}(2x).$$

令 $x = 0$，得 $(2 - 2^n)\,\mathrm{f}^{(n)}(0) = 0$。取 $n = 2$，此式即 $-2\mathrm{f}''(0) = 0$，所以 $\mathrm{f}''(0) = 0$，正是所求。

同樣的論證對每個 $n \neq 1$ 都成立：因子 $2 - 2^n$ 只在 $n = 1$ 時才等於零，所以

$$\mathrm{f}(0) = 0 \quad \text{及} \quad \mathrm{f}^{(n)}(0) = 0 \ \text{對所有} n \geq 2.$$

（$n = 0$ 的情況就是 $2f(0) = f(0)$。）麥克勞林級數中只有一次項留下來：

$$f(x) = f(0) + f'(0)\,x + \tfrac{1}{2!}f''(0)\,x^2 + \cdots = kx, \qquad k = f'(0).$$

反過來，$kx + ky = k(x+y)$，所以每個這樣的函數都滿足方程。$(*)$ 的通解是 $f(x) = kx$，其中 $k$ 是任意常數。

**第 (i) 部分**

要取對數，必須有 $g > 0$，所以先處理正負號的問題。代入 $x = y = \frac{1}{2}s$，得 $g(s) = \left(g(\frac{1}{2}s)\right)^2 \geq 0$。若對某個 $a$ 有 $g(a) = 0$，則對每個 $x$，

$$g(x) = g(a)\,g(x-a) = 0,$$

所以 $g \equiv 0$，而它確實滿足恆等式。否則對所有 $x$ 都有 $g(x) > 0$，於是 $G = \ln g$ 處處有定義。對 $g(x)g(y) = g(x+y)$ 取對數，得

$$G(x) + G(y) = G(x+y),$$

這正是 $(*)$。所以 $G(x) = kx$，而

$$g(x) = e^{kx}$$

其中 $k$ 為常數；它顯然滿足恆等式，因為 $e^{kx}e^{ky} = e^{k(x+y)}$。除了平凡解 $g \equiv 0$ 之外，這些就是全部的解。

**第 (ii) 部分**

每個正數 $x$ 都可以唯一地寫成 $e^u$，其中 $u$ 是實數，即 $u = \ln x$。所以我們可以寫 $x = e^u$ 和 $y = e^v$，其中 $u$ 和 $v$ 取遍所有實數。這時 $xy = e^{u+v}$，恆等式變成

$$H(u) + H(v) = h(e^u) + h(e^v) = h(e^{u+v}) = H(u+v)$$

對所有實數 $u$、$v$ 成立。這又是 $(*)$，所以 $H(u) = ku$。由於 $u = \ln x$，

$$h(x) = k\ln x \qquad (x > 0),$$

而它滿足恆等式，因為 $\ln x + \ln y = \ln(xy)$。

**第 (iii) 部分**

式子 $\dfrac{x+y}{1-xy}$ 其實是喬裝了的正切加法公式（tangent addition formula）。設 $x = \tan u$ 和 $y = \tan v$，其中 $u, v \in (-\frac{\pi}{2}, \frac{\pi}{2})$，於是 $u = \arctan x$ 及 $v = \arctan y$。那麼

$$z = \frac{\tan u + \tan v}{1 - \tan u \tan v} = \tan(u+v).$$

在 $(-\frac{\pi}{2}, \frac{\pi}{2})$ 上定義 $\mathrm{T}(u) = \mathrm{t}(\tan u)$。只要 $u+v$ 也落在 $(-\frac{\pi}{2}, \frac{\pi}{2})$ 之內，恆等式就給出

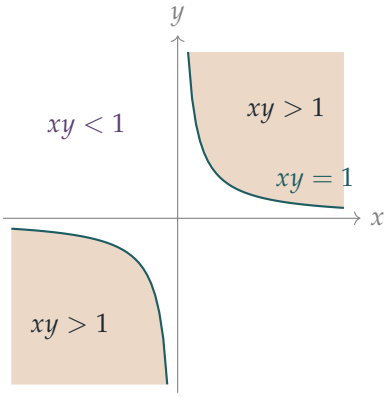
$$\mathrm{T}(u) + \mathrm{T}(v) = \mathrm{t}(x) + \mathrm{t}(y) = \mathrm{t}\big(\tan(u+v)\big) = \mathrm{T}(u+v).$$

**想一想** 處理 $(*)$ 的論證用到恆等式對*所有* $x$ 和 $y$ 成立。如果我們只知道它在 $u$、$v$ 和 $u+v$ 都屬於 $(-\frac{\pi}{2}, \frac{\pi}{2})$ 時成立，論證還行得通嗎？

行得通。麥克勞林級數的論證只涉及 0 附近的情況。當 $|u| < \frac{\pi}{4}$ 時可以取 $v = u$（這時 $|2u| < \frac{\pi}{2}$），於是在 0 附近有 $2\mathrm{T}(u) = \mathrm{T}(2u)$。在 $u = 0$ 處求導 $n$ 次，與之前完全一樣得到 $(2 - 2^n)\mathrm{T}^{(n)}(0) = 0$，所以 $\mathrm{T}(u) = ku$。由於 $u = \arctan x$，

$$\mathrm{t}(x) = k \arctan x.$$

究竟對哪些 $x$ 和 $y$ 這個答案才真正成立？若 $xy < 1$，則 $u + v \in (-\frac{\pi}{2}, \frac{\pi}{2})$。當 $u$ 和 $v$ 異號時，這是必然的。當兩者皆為正數時，$\tan u \tan v < 1$ 表示 $\tan v < \cot u = \tan(\frac{\pi}{2} - u)$，所以 $u + v < \frac{\pi}{2}$；兩者皆為負數的情況由對稱性同理可得。因此 $\arctan x + \arctan y = \arctan z$，恆等式成立。然而若 $xy > 1$，$u + v$ 便落在 $(-\frac{\pi}{2}, \frac{\pi}{2})$ 之外，這時 $\arctan z = u + v \mp \pi$。例如 $x = y = 2$ 給出 $z = -\frac{4}{3}$，而 $2\arctan 2 - \arctan(-\frac{4}{3}) = \pi$。




所以 $\mathrm{t}(x) = k \arctan x$ 對每一對滿足 $xy < 1$ 的數都符合恆等式，也就是圖中包含原點的區域。這就是題目想要的答案。在 $xy > 1$ 的陰影區域內，等式兩邊相差 $\pm k\pi$。如果我們堅持恆等式要對每一對滿足 $xy \neq 1$ 的數都毫無例外地成立，那就只剩下 $k = 0$，唯一的解是 $\mathrm{t} \equiv 0$。

## 驗算

- 對於第 (iii) 部分，把恆等式對 $y$ 求導，然後代入 $y = 0$。由於 $\partial z/\partial y = (1+x^2)/(1-xy)^2$，在 $y = 0$ 時等於 $1 + x^2$，我們得到 $\mathrm{t}'(0) = \mathrm{t}'(x)(1 + x^2)$。所以 $\mathrm{t}'(x) = k/(1+$

$x^2)$，再配合 $\mathrm{t}(0) = 0$，積分便得 $k \arctan x$。對 $(*)$ 用同樣的技巧，立即得到 $\mathrm{f}'(x) = \mathrm{f}'(0)$。

- 我們對全部四個答案各用 $10^4$ 對隨機數作了數值檢驗。在第 (iii) 部分，$k \arctan x$ 在約 $5 \times 10^4$ 對滿足 $xy < 1$ 的隨機數上全部符合恆等式，而在每一對滿足 $xy > 1$ 的數上，誤差的大小都恰好是 $k\pi$。

## 常見錯誤

- 忘記處理常數項：$\mathrm{f}(0) = 0$ 來自 $n = 0$ 的情況，所以答案是 $kx$，而不是 $kx + c$。
- 沒有檢查 $\mathrm{g} > 0$ 就取 $\ln \mathrm{g}$。零函數是第 (i) 部分名副其實的一個解。
- 以為 $\arctan x + \arctan y = \arctan \dfrac{x+y}{1-xy}$ 總是成立。它需要 $xy < 1$。

## STEP 之外

即使不假設麥克勞林級數存在，柯西方程 $(*)$ 所需的條件其實少得多：只要 f 連續（continuous）（甚至只需在某個區間上有界（bounded）），仍然可以推出 $\mathrm{f}(x) = kx$，出發點是對有理數 $q$ 有 $\mathrm{f}(q) = q\,\mathrm{f}(1)$。不過，利用選擇公理（axiom of choice），可以構造出並非線性的怪異解，它們的圖像在平面上稠密（dense）。第 (i) 部分還說明，指數函數本質上是由加法到乘法的唯一連續同態（homomorphism），這個想法在羣論（group theory）中反覆出現。想看這類連續性與稠密性論證的嚴謹處理，Abbott 的 *Understanding Analysis* 是很好的入門讀物。

詳細解答

# STEP III 2006，第 5 題

純數學 • 複數與幾何

**STEP III 2006，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/06-s3-q5

## 關鍵思路

一個三角形是等邊三角形（equilateral triangle），當且僅當從某個頂點看，其中一條邊是另一條邊旋轉 $60^\circ$ 的結果。在複數（complex number）中，這個旋轉就是乘以 $\mathrm{e}^{\pm \mathrm{i}\pi/3}$，而這兩個數正好是 $t^2 - t + 1 = 0$ 的根。把原點移到其中一個頂點，那個對稱式便恰好化為這個二次式。此後，三次方程的部分用韋達公式（Vieta's formulae）即可，而最後一部分只需留意 $z \mapsto pw + q$ 只是把圖形旋轉、縮放和平移。

## 解答

### 等邊三角形的條件

設 $u = \beta - \alpha$ 和 $v = \gamma - \alpha$，即由頂點 $\alpha$ 出發的兩條邊。由於各點互不相同，$u \neq 0$、$v \neq 0$ 且 $u \neq v$。首先，

$$\alpha^2 + \beta^2 + \gamma^2 - \beta\gamma - \gamma\alpha - \alpha\beta = \tfrac{1}{2}\left[(\alpha-\beta)^2 + (\beta-\gamma)^2 + (\gamma-\alpha)^2\right],$$

把右邊展開即可驗證。現在 $\alpha - \beta = -u$、$\gamma - \alpha = v$ 及 $\beta - \gamma = u - v$，所以方括號內的式子是 $u^2 + (u-v)^2 + v^2 = 2(u^2 - uv + v^2)$。因此

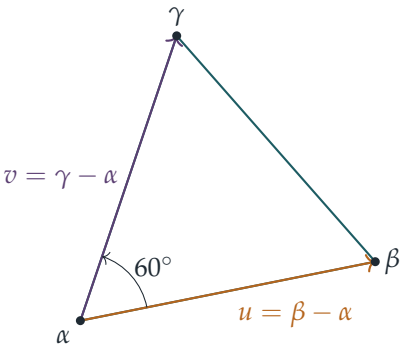
$$\alpha^2 + \beta^2 + \gamma^2 - \beta\gamma - \gamma\alpha - \alpha\beta = u^2 - uv + v^2 = u^2\left(t^2 - t + 1\right), \qquad t = \frac{v}{u}.$$

由於 $u \neq 0$，這個式子等於零當且僅當 $t^2 - t + 1 = 0$，即

$$t = \frac{1 \pm \mathrm{i}\sqrt{3}}{2} = \mathrm{e}^{\pm \mathrm{i}\pi/3}.$$

**想一想** 為甚麼 $\dfrac{\gamma-\alpha}{\beta-\alpha}=\mathrm{e}^{\pm \mathrm{i}\pi/3}$ 恰好表示這個三角形是等邊三角形？

$v/u$ 的模（modulus）是 $|\gamma-\alpha|/|\beta-\alpha|$，而它的輻角（argument）是在 $\alpha$ 處由邊 $\alpha\beta$ 轉到邊 $\alpha\gamma$ 的角。所以 $v/u=\mathrm{e}^{\pm \mathrm{i}\pi/3}$ 的意思是：由 $\alpha$ 出發的兩條邊相等，而它們之間的夾角是 $60°$。這樣的等腰三角形，底角是 $\frac{1}{2}(180°-60°)=60°$，所以它是等邊三角形。反過來，如果三角形是等邊的，則 $|v|=|u|$，而 $\alpha$ 處的角是 $60°$，至於向哪一邊轉則視乎三角形的定向（orientation），所以 $v/u=\mathrm{e}^{\mathrm{i}\pi/3}$ 或 $\mathrm{e}^{-\mathrm{i}\pi/3}$。




因此，這些互不相同的點（不論次序）構成等邊三角形，當且僅當該式等於零，正如題目所要求。

### 三次方程

設 $(*)$ 的根為 $\alpha,\beta,\gamma$。根據韋達公式，$\alpha+\beta+\gamma=-a$ 及 $\alpha\beta+\beta\gamma+\gamma\alpha=b$。所以

$$\begin{aligned}\alpha^2+\beta^2+\gamma^2-\beta\gamma-\gamma\alpha-\alpha\beta&=(\alpha+\beta+\gamma)^2-3(\alpha\beta+\beta\gamma+\gamma\alpha)\\&=a^2-3b.\end{aligned}$$

根據第一個結果，這些（互不相同的）根構成等邊三角形，當且僅當 $a^2-3b=0$，即 $a^2=3b$。

有一點要留意：第一個結果需要各點互不相同。如果有兩個根重合，例如 $\alpha=\beta$，該式便化為 $(\gamma-\alpha)^2$，所以只有當三個根全部相等時它才等於零。因此，在有重根（repeated root）的情況下，$a^2=3b$ 只會在三重根時出現，即 $z^3+az^2+bz+c=(z+\frac{a}{3})^3$，亦即同時有 $c=a^3/27$。除了這個只有一點的退化「三角形」之外，$a^2=3b$ 正是所求的條件。

### 變換

把 $z = pw + q$ 代入 $(*)$，再除以 $p^3 \neq 0$：

$$w^3 + \frac{3q + a}{p}\, w^2 + \frac{3q^2 + 2aq + b}{p^2}\, w + \frac{q^3 + aq^2 + bq + c}{p^3} = 0.$$

所以

$$A = \frac{3q + a}{p}, \qquad B = \frac{3q^2 + 2aq + b}{p^2},$$

因而

$$A^2 - 3B = \frac{(9q^2 + 6aq + a^2) - (9q^2 + 6aq + 3b)}{p^2} = \frac{a^2 - 3b}{p^2}.$$

如果 $(*)$ 的根構成等邊三角形，則它們互不相同，而且 $a^2 = 3b$，所以 $A^2 = 3B$。$(**)$ 的根是 $w = (z - q)/p$，其中 $z$ 是 $(*)$ 的根；由於 $p \neq 0$，這個映射是一對一（one-to-one）的，所以這些根同樣互不相同。根據上一個結果，$(**)$ 的根構成等邊三角形。

從幾何角度看，這並不意外：$w = (z - q)/p$ 是平移 $-q$、旋轉 $-\arg p$ 再縮放 $1/|p|$ 倍，而這些運算都不會破壞等邊三角形。用第一部分的說法，每條邊都除以 $p$，所以比 $v/u$ 不變。

## 驗算

- 單位立方根（cube roots of unity）$1, \omega, \omega^2$ 構成等邊三角形，而且是 $z^3 - 1 = 0$ 的根，其中 $a = b = 0$，所以 $a^2 = 3b$ 成立。
- 取 $\alpha = 0$、$\beta = 1$ 及 $\gamma = \frac{1}{2} + \frac{\sqrt{3}}{2}\mathrm{i}$，該式為 $1 + \gamma^2 - \gamma$，它等於零，因為 $\gamma = \mathrm{e}^{\mathrm{i}\pi/3}$ 是 $t^2 - t + 1$ 的根。
- 以 1000 組隨機三元組作數值驗算，證實了對於以這些數為根的三次方程，$\sum \alpha^2 - \sum \alpha\beta = a^2 - 3b$ 成立。同一驗算亦證實了 $A^2 - 3B = (a^2 - 3b)/p^2$，以及在隨機選取的 $p$ 和 $q$ 之下，構成等邊三角形的根仍然構成等邊三角形。

## 常見錯誤

- 只檢查一種定向：$\mathrm{e}^{\mathrm{i}\pi/3}$ 和 $\mathrm{e}^{-\mathrm{i}\pi/3}$ 都會出現，而「當且僅當」需要兩者兼顧。
- 忘記了正是因為 $u \neq 0$（各點互不相同），我們才可以除以 $u^2$。
- 代入時把 $B$ 算錯：$w$ 的係數包括來自立方項的 $3q^2$、來自平方項的 $2aq$，以及 $b$，全部再除以 $p^2$。

## STEP 之外

映射 $z \mapsto pz + q$ 就是平面上的正向相似變換（direct similarity）。在這些映射下保持不變的量，例如比 $(\gamma - \alpha)/(\beta - \alpha)$，或者 $a^2 - 3b$ 是否等於零，稱為不變量（invariant），這是幾何和代數中的核心概念。$a^2 - 3b$ 這個式子（除了一個常數因子外）就是 $f'$ 的判別式（discriminant），其中 $f(z) = z^3 + az^2 + bz + c$。這又連繫到一個漂亮的事實——Marden 定理（Marden's theorem）：以 $f$ 的根為頂點的三角形，有一個橢圓與它的三邊相切於各邊中點，而 $f'$ 的根正是這個橢圓的焦點。對於等邊三角形，這個橢圓是一個圓，它的兩個焦點合而為一，成為 $f'$ 的二重根。Needham 的 *Visual Complex Analysis*（《複分析：可視化方法》）正是以這種精神來闡述複數幾何。

詳細解答

# STEP III 2006，第 6 題

純數學 • 極坐標曲線與反射

**STEP III 2006，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/06-s3-q6

## 關鍵思路

斜率（gradient）公式不過是把鏈式法則（chain rule）應用於 $x = r\cos\theta, y = r\sin\theta$ 而已。至於鏡面，把極點（pole）設在 $O$，並令始線（initial line）沿着 $L$：這樣，在點 $P$ 反射的光線便沿着徑向直接返回，因此其方向只由 $\theta$ 決定。反射定律（law of reflection）指出法線（normal）平分兩條光線之間的夾角，於是法線（以至切線（tangent））便可用 $\theta$ 確定下來。這樣一來，幾何問題就化為關於 $r(\theta)$ 的可分離變數微分方程（separable differential equation）。

## 解答

### 極坐標下的斜率

由 $r = r(\theta)$ 給出的曲線，其直角坐標（Cartesian coordinates）為 $x = r\cos\theta, y = r\sin\theta$，其中 $r$ 是 $\theta$ 的函數。利用積法則（product rule）對 $\theta$ 求導，

$$\frac{\mathrm{d}x}{\mathrm{d}\theta} = \frac{\mathrm{d}r}{\mathrm{d}\theta}\cos\theta - r\sin\theta, \qquad \frac{\mathrm{d}y}{\mathrm{d}\theta} = \frac{\mathrm{d}r}{\mathrm{d}\theta}\sin\theta + r\cos\theta.$$

根據鏈式法則，凡是 $\mathrm{d}x/\mathrm{d}\theta \neq 0$ 之處，

$$\frac{\mathrm{d}y}{\mathrm{d}x} = \frac{\mathrm{d}y/\mathrm{d}\theta}{\mathrm{d}x/\mathrm{d}\theta} = \frac{\dfrac{\mathrm{d}r}{\mathrm{d}\theta}\sin\theta + r\cos\theta}{\dfrac{\mathrm{d}r}{\mathrm{d}\theta}\cos\theta - r\sin\theta} = \frac{\dfrac{\mathrm{d}r}{\mathrm{d}\theta}\tan\theta + r}{\dfrac{\mathrm{d}r}{\mathrm{d}\theta} - r\tan\theta},$$

這裏把分子和分母同除以 $\cos\theta$（當 $\cos\theta \neq 0$ 時可以這樣做）。這就是所求的斜率。

未經相除的形式也值得保留：向量

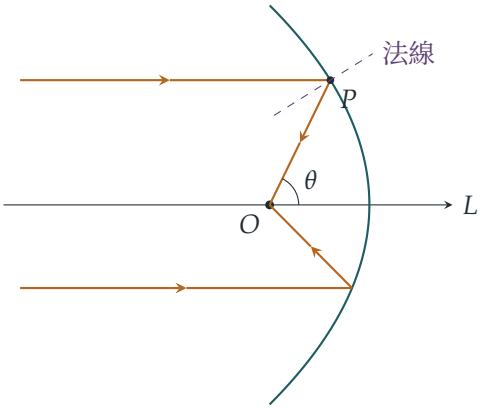
$$\mathbf{t} = \left(\tfrac{\mathrm{d}r}{\mathrm{d}\theta}\cos\theta - r\sin\theta,\ \tfrac{\mathrm{d}r}{\mathrm{d}\theta}\sin\theta + r\cos\theta\right)$$

在曲線上每一點都沿着切線方向，包括 $\tan\theta$ 或斜率無定義的地方。

## 鏡面

任取一個包含 $L$ 的平面 $\Pi$，並設 $P$ 為鏡面與 $\Pi$ 相交的一點。在 $P$ 的入射光線（incident ray）平行於 $L$，所以它位於 $\Pi$ 之內；反射光線（reflected ray）通過 $O$，而該點在 $L$ 上，所以它同樣位於 $\Pi$ 之內。反射定律亦要求在 $P$ 的法線位於兩條光線所在的平面內，因此法線位於 $\Pi$ 之內，而交線在 $P$ 的切線就是 $\Pi$ 內與法線垂直的那條直線。所以一切都發生在平面 $\Pi$ 之內，我們可以在二維中處理問題。

在 $\Pi$ 內採用極坐標（polar coordinates），以 $O$ 為極點，始線沿着 $L$，並選定方向使入射光沿 $\theta = 0$ 的方向行進（即圖中由左至右）。設 $P$ 為 $(r,\theta)$，其中 $-\pi < \theta < \pi$。（$\theta = \pi$ 所對應的點位於 $L$ 上、$O$ 的後方，光線在那裏會直接穿過 $O$，根本不需要鏡面。）




**法線的方向**。 在 $P$ 取兩個沿着光線、背向 $P$ 的單位向量（unit vectors）：

$$\mathbf{a} = (-1,\, 0)\ \text{（沿入射光線往回走）}, \qquad \mathbf{b} = (-\cos\theta,\, -\sin\theta)\ \text{（指向 } O\text{）}.$$

法線與 $\mathbf{a}$ 和 $\mathbf{b}$ 所成的角相等，所以它沿着兩者夾角的平分線，亦即 $\mathbf{a}+\mathbf{b}$ 的方向（兩個單位向量所構成的菱形，其對角線平分它們的夾角）。現在

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

這裏用了 $1+\cos\theta = 2\cos^2\frac{\theta}{2}$ 和 $\sin\theta = 2\sin\frac{\theta}{2}\cos\frac{\theta}{2}$。由於 $-\pi < \theta < \pi$，$\cos\frac{\theta}{2} > 0$，所以這個向量非零，而在 $P$ 的法線沿着

$$\mathbf{n} = \left(\cos\tfrac{\theta}{2},\, \sin\tfrac{\theta}{2}\right):$$

法線與 $L$ 所成的角為 $\frac{\theta}{2}$。

**想一想** 為甚麼不直接令斜率公式等於切線的斜率？

這樣做是可以的：切線垂直於 $\mathbf{n}$，所以其斜率為 $-\cot\frac{\theta}{2}$，令它等於極坐標公式並去分母，便會得到與下文相同的方程。不過這條路要除以 $\tan\theta$、$\cot\frac{\theta}{2}$ 以及公式中的分母，所以在 $\theta = 0$ 和 $\theta = \pm\frac{\pi}{2}$ 處行不通。採用切向量 $\mathbf{t}$ 則可避開這一切麻煩。

**微分方程**。 切線垂直於法線，所以 $\mathbf{t}\cdot\mathbf{n} = 0$：

$$\frac{\mathrm{d}r}{\mathrm{d}\theta}\left(\cos\theta\cos\tfrac{\theta}{2} + \sin\theta\sin\tfrac{\theta}{2}\right) + r\left(\cos\theta\sin\tfrac{\theta}{2} - \sin\theta\cos\tfrac{\theta}{2}\right) = 0.$$

根據複角公式（compound-angle formulae），兩個括號分別等於 $\cos\left(\theta - \frac{\theta}{2}\right) = \cos\frac{\theta}{2}$ 和 $\sin\left(\frac{\theta}{2} - \theta\right) = -\sin\frac{\theta}{2}$。因此

$$\frac{\mathrm{d}r}{\mathrm{d}\theta}\cos\tfrac{\theta}{2} = r\sin\tfrac{\theta}{2}, \qquad \text{即} \qquad \frac{1}{r}\frac{\mathrm{d}r}{\mathrm{d}\theta} = \tan\tfrac{\theta}{2},$$

這裏除以了 $r\cos\frac{\theta}{2}$，由於 $P \neq O$ 且 $-\pi < \theta < \pi$，它是正數。

**求解**。 這個方程是可分離變數的。對 $\theta$ 積分，並注意到 $\int \tan\frac{\theta}{2}\,\mathrm{d}\theta = -2\ln\cos\frac{\theta}{2}$（因為 $\cos\frac{\theta}{2} > 0$，此式成立），

$$\ln r = -2\ln\cos\tfrac{\theta}{2} + C \quad\Longrightarrow\quad r\cos^2\tfrac{\theta}{2} = k,$$

其中 $k = \mathrm{e}^C > 0$ 沿着整條曲線是常數（曲線是區間 $-\pi < \theta < \pi$ 上單獨一段連通的 $r = r(\theta)$）。利用 $2\cos^2\frac{\theta}{2} = 1 + \cos\theta$，

$$r(1 + \cos\theta) = 2k.$$

**轉換為直角坐標**。 以 $x = r\cos\theta$ 和 $r = \sqrt{x^2 + y^2}$ 代入，上式變成 $\sqrt{x^2 + y^2} = 2k - x$。兩邊平方（兩邊均為非負，因為 $2k - x = r \geq 0$），

$$x^2 + y^2 = 4k^2 - 4kx + x^2 \qquad\Longrightarrow\qquad y^2 = 4k(k - x).$$

這是一條拋物線（parabola），以 $L$ 為軸，頂點（vertex）在 $(k, 0)$，焦點（focus）在原點 $O$，開口朝向入射光的來向。由於 $\Pi$ 是任意一個包含 $L$ 的平面，鏡面與每一個這樣的平面的交線都是拋物線，這正是所要證明的。■

## 驗算

- 倒過來檢驗：在 $y^2 = 4k(k - x)$ 上，法線沿着 $(4k, 2y)$；把方向 $(1, 0)$ 對這條法線作反射，便得到一條通過 $O$ 的光線。我們以數值方法在 $10^4$ 個隨機點上驗證了這一點

（取 $k = 1.3$），並且用有限差分（finite differences）驗證了 $r = 2k/(1+\cos\theta)$ 滿足 $r'/r = \tan\frac{\theta}{2}$。

- 對於隨機選取的曲線 $r(\theta)$，斜率公式與有限差分所得的斜率相符，誤差約為 $10^{-8}$。
- 在 $\theta = 0$ 處，法線沿着 $L$，而切線垂直於 $L$：這就是頂點 $(k, 0)$，沿 $L$ 射來的光線在那裏會直接反射回來並通過 $O$。

## 常見錯誤

- 以外角平分線（external bisector）$\mathbf{a} - \mathbf{b}$ 作為法線。這個方向與兩條光線所成的角是互補而非相等，並會導致 $r'/r = -\cot\frac{\theta}{2}$，其解 $r(1-\cos\theta) = 2k$ 是一條開口朝相反方向的拋物線，所聚焦的是從相反方向射來的光。
- 忘記解釋為何問題是二維的：法線必須位於包含 $L$ 和 $P$ 的平面內，否則交線的切線未必就是我們所用的那一條。
- 把 $\sqrt{x^2+y^2} = 2k - x$ 兩邊平方時，沒有指出兩邊均為非負。

## STEP 之外

這塊鏡面是一個旋轉拋物面（paraboloid of revolution），而這裏所證明的性質其實就是焦點－準線（focus–directrix）性質的另一種面貌：$r(1+\cos\theta) = 2k$ 是說，與 $O$ 的距離等於與直線 $x = 2k$ 的距離。這正是衛星碟形天線、汽車車頭燈和反射望遠鏡都採用拋物面的原因。更一般地說，所有圓錐曲線（conics）都具有極坐標形式 $r(1+e\cos\theta) = \ell$，而這恰恰就是牛頓（Newton）求解行星軌道時所出現的式子；Kleppner 與 Kolenkow 的 *An Introduction to Mechanics*（力學入門教科書）由平方反比定律（inverse-square law）把它推導出來。

詳細解答

# STEP III 2006，第 7 題

純數學 • 一階微分方程

**STEP III 2006，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/06-s3-q7

## 關鍵思路

兩條微分方程（differential equation）都是關於 $\mathrm{d}y/\mathrm{d}x$ 的二次方程。我們不急於把它們當作微分方程來處理，而是先就導數（derivative）解出每條二次方程，這樣原方程便分拆成兩條普通的一階方程。利用雙曲恆等式（hyperbolic identities），兩個根會化簡成很簡單的形式，而 $x = 0$ 處的條件則告訴我們應該跟隨哪一個根。

## 解答

### 第 (i) 部分

由二次公式（quadratic formula），

$$u = -\sinh x \pm \sqrt{\sinh^2 x + 1} = -\sinh x \pm \cosh x,$$

因為 $\sinh^2 x + 1 = \cosh^2 x$ 而且 $\cosh x > 0$。利用 $\cosh x + \sinh x = \mathrm{e}^x$ 和 $\cosh x - \sinh x = \mathrm{e}^{-x}$，兩個根是

$$u = \mathrm{e}^{-x} \qquad 或 \qquad u = -\mathrm{e}^x.$$

微分方程就是這條二次方程，其中 $u = \mathrm{d}y/\mathrm{d}x$，所以在每一個 $x$ 處必有

$$\frac{\mathrm{d}y}{\mathrm{d}x} = \mathrm{e}^{-x} \qquad 或 \qquad \frac{\mathrm{d}y}{\mathrm{d}x} = -\mathrm{e}^x.$$

第一個根恆為正，第二個根恆為負，所以兩者永不相等。因此，導數連續（continuous）的解不可能由一個根跳到另一個根，而 $x = 0$ 處 $\mathrm{d}y/\mathrm{d}x > 0$ 這個條件，就決定了自始至終都取第一個根。積分後得 $y = -\mathrm{e}^{-x} + c$，再由 $x = 0$ 時 $y = 0$ 得 $c = 1$：

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

**第 (ii) 部分**

設 $p = \mathrm{d}y/\mathrm{d}x$。若 $\sinh y = 0$，即 $y = 0$，方程便化為 $2p = 0$。$y \equiv 0$ 就是這樣成為一個解的，而題目正是叫我們把這個解撇開不理。在 $y \neq 0$ 之處，由二次公式得

$$p = \frac{-1 \pm \sqrt{1 + \sinh^2 y}}{\sinh y} = \frac{-1 \pm \cosh y}{\sinh y}.$$

利用半角形式（half-angle forms）$\cosh y - 1 = 2\sinh^2 \frac{y}{2}$、$\cosh y + 1 = 2\cosh^2 \frac{y}{2}$ 和 $\sinh y = 2\sinh \frac{y}{2} \cosh \frac{y}{2}$，兩個根是

$$\frac{\mathrm{d}y}{\mathrm{d}x} = \tanh \frac{y}{2} \qquad 或 \qquad \frac{\mathrm{d}y}{\mathrm{d}x} = -\coth \frac{y}{2}.$$

這兩個根永不相等，因為它們的積是 $-1$，而若兩根相等，其積應為 $\tanh^2 \frac{y}{2} \geq 0$。

**想一想** 兩條方程之中，哪一條可以給出一個通過原點而又不恆等於零的解？

第一條方程 $y' = \tanh \frac{y}{2}$ 有解 $y \equiv 0$。它的右邊是 $y$ 的光滑函數（smooth function），所以通過 $(0,0)$ 的解就只有這一個：這又是那個平凡解（trivial solution）。第二條方程在 $y = 0$ 處是奇異（singular）的，該處 $|\mathrm{d}y/\mathrm{d}x| \to \infty$，所以它的解曲線可以帶着鉛垂切線與 $y = 0$ 相交。這正是我們要找的解。分離變數（separating the variables），

$$\int \tanh \frac{y}{2} \, \mathrm{d}y = -\int \mathrm{d}x \quad \Longrightarrow \quad 2\ln\cosh \frac{y}{2} = -x + C.$$

代入 $x = 0$ 時 $y = 0$，得 $C = 0$，所以 $\cosh^2 \frac{y}{2} = \mathrm{e}^{-x}$。由於 $\cosh y = 2\cosh^2 \frac{y}{2} - 1$，

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

我們可以直接驗證這個結果。求導得 $\sinh y \, y' = -2\mathrm{e}^{-x} = -(\cosh y + 1)$，這正是 $y' = -\coth \frac{y}{2}$。無論 $y$ 是正是負，這都成立，所以整條曲線，包括上半和下半，都是解。由於 $\cosh y \geq 1$，曲線只在 $x \leq 0$ 時存在。它關於 $x$ 軸對稱。在原點附近 $1 + \frac{1}{2}y^2 \approx 1 - 2x$，所以曲線在那裏看起來像拋物線 $y^2 = -4x$。

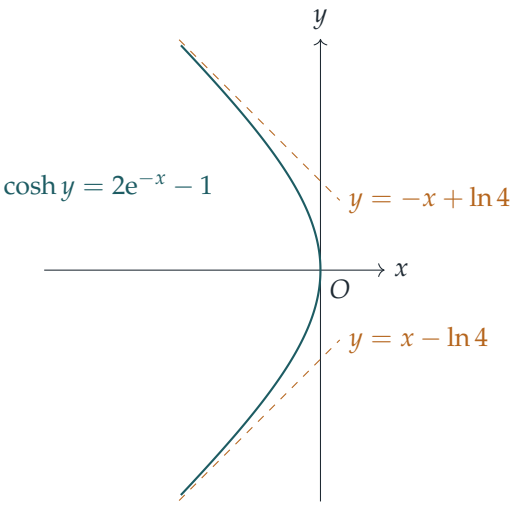
**漸近線(asymptotes)**。 當 $x \to -\infty$ 時，$\cosh y \to \infty$，所以 $|y| \to \infty$。設 $w = 2\mathrm{e}^{-x} - 1$，則

$$|y| = \operatorname{arcosh} w = \ln\bigl(w + \sqrt{w^2 - 1}\bigr),$$

所以

$$|y| - (-x + \ln 4) = \ln \frac{w + \sqrt{w^2 - 1}}{4\mathrm{e}^{-x}} = \ln\Bigl(\frac{w}{4\mathrm{e}^{-x}}\Bigl(1 + \sqrt{1 - w^{-2}}\Bigr)\Bigr).$$

當 $x \to -\infty$ 時，$w/(4\mathrm{e}^{-x}) = \frac{1}{2} - \frac{1}{4}\mathrm{e}^{x} \to \frac{1}{2}$，而 $1 + \sqrt{1 - w^{-2}} \to 2$，所以對數內的式子趨向 1，上述差值趨向 0。因此曲線的上半趨近 $y = -x + \ln 4$，而由對稱性，下半趨近 $y = -(-x + \ln 4)$。漸近線是 $y = \pm(-x + \ln 4)$，正如題目所求。（當 $x \to 0^-$ 時曲線保持有界，所以漸近線就只有這兩條。）




## 驗算

- 在 (i) 中，$y = 1 - \mathrm{e}^{-x}$ 的導數是 $y' = \mathrm{e}^{-x}$，而 $\mathrm{e}^{-2x} + 2\mathrm{e}^{-x}\sinh x - 1 = \mathrm{e}^{-2x} + (1 - \mathrm{e}^{-2x}) - 1 = 0$。
- 在 (ii) 中，由原點附近開始，以數值方法逐步求解 $y' = -\coth\frac{y}{2}$ 直到 $x = -6$，得 $y \approx 7.38505$，與 $\operatorname{arcosh}(2\mathrm{e}^6 - 1)$ 吻合。在 $x = -10$ 處，曲線與 $-x + \ln 4$ 之間的差距約為 $2 \times 10^{-5}$。

## 常見錯誤

- 在 (i) 中取了 $-\sinh x - \cosh x = -\mathrm{e}^x$ 這個根：這樣導數在 $x = 0$ 處是負數。
- 在 (ii) 中因為 $y' = \tanh\frac{y}{2}$ 在 $y = 0$ 處「性質良好」而選了它。通過原點時，它只給出 $y \equiv 0$。
- 忘記了曲線的下半；這一半有 $y < 0$，第二條漸近線正是由它而來。

## STEP 之外

像這兩條方程一樣，對 $y'$ 而言並非線性的方程，稱為一階隱式（first-order implicit）方程。就 $y'$ 求解，會把它分拆成若干個分支。第 (ii) 部分亦顯示解的唯一性（uniqueness）可以不成立：在 $y = 0$ 處，右邊的 $-\coth\frac{y}{2}$ 不滿足利普希茨（Lipschitz）條件，所以一個非平凡解才可以與 $y \equiv 0$ 從同一點出發。Boyce 與 DiPrima 的 *Elementary Differential Equations*（初等微分方程教科書）講述了存在性與唯一性定理，以及這些定理在甚麼情

況下適用。

詳細解答

# STEP III 2006，第 8 題

純數學 • 公理化微分法

> **STEP III 2006，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/06-s3-q8

## 關鍵思路

這四條規則正是你日常使用的求導（differentiation）性質：線性（linearity）和積法則（product rule），再加上一個事實，就是 $x$「求導後得」1。每個多項式（polynomial）都是由數字 1 和字母 $x$ 經過相乘和相加構成的，所以只要知道 $\triangle$ 對 1 和對 $x$ 的作用，這些規則就令它在其他地方再無任何自由。把積法則應用於 $x \cdot x^n$，便可對冪次作歸納法（induction），最後再由線性完成餘下的工作。

## 解答

以下「規則 (i)」至「規則 (iv)」均指題目所給的四條定義規則，次序與題目相同。

### 常數

先考慮常數多項式 1。由於 $1 = 1 \cdot 1$，在規則 (iv) 中取 $f(x) = g(x) = 1$，得

$$\triangle 1 = \triangle(1 \cdot 1) = 1 \cdot \triangle 1 + 1 \cdot \triangle 1 = 2\,\triangle 1.$$

兩邊同時減去 $\triangle 1$，得 $\triangle 1 = 0$（零多項式）。

> ? **想一想** 為甚麼不直接說「$\triangle$ 作用於常數得零，就像導數一樣」？

因為我們還不知道 $\triangle$ 就是導數（derivative）；這正是最後一部分要我們證明的。我們能用的只有那四條規則，而其中沒有一條直接提及常數。竅門在於 1 等於它自己的平方，所以積法則迫使 $\triangle 1$ 等於它自身的兩倍。

現設 $f(x) = \lambda$ 為任意常數。則 $\lambda = \lambda \cdot 1$，故由規則 (iii)

$$\triangle\lambda = \triangle(\lambda \cdot 1) = \lambda \triangle 1 = \lambda \cdot 0 = 0.$$

所以對每個次數（degree）為零的多項式，都有 $\triangle f(x) = 0$，正如所求。

### 平方與立方

把 $x^2 = x \cdot x$ 寫出，先用規則 (iv)，再用規則 (i)，

$$\triangle x^2 = x \triangle x + x \triangle x = x \cdot 1 + x \cdot 1 = 2x.$$

把 $x^3 = x \cdot x^2$ 寫出，先用規則 (iv)，再用剛得到的結果和規則 (i)，

$$\triangle x^3 = x \triangle x^2 + x^2 \triangle x = x \cdot 2x + x^2 \cdot 1 = 3x^2.$$

### 一般多項式

*步驟 1：$x$ 的冪。*我們用歸納法證明

$$\triangle x^n = n\, x^{n-1} \qquad \text{對每個整數} n \geq 1.$$

當 $n = 1$ 時，這就是 $\triangle x = 1$，即規則 (i)。假設結果對某個 $n = k \geq 1$ 成立。由於 $x^{k+1} = x \cdot x^k$，規則 (iv) 給出

$$\begin{aligned} \triangle x^{k+1} &= x \triangle x^k + x^k \triangle x \\ &= x \cdot k\, x^{k-1} + x^k \cdot 1 \\ &= (k+1)\, x^k, \end{aligned}$$

其中第二行用了歸納假設（induction hypothesis）和規則 (i)。所以結果對 $n = k+1$ 成立，由歸納法可知它對所有 $n \geq 1$ 成立。當 $n = 0$ 時，$x^0 = 1$，而我們已證明 $\triangle 1 = 0$，這與 $\dfrac{\mathrm{d}}{\mathrm{d}x} 1 = 0$ 一致。所以對每個 $n \geq 0$，

$$\triangle x^n = \frac{\mathrm{d}}{\mathrm{d}x}\, x^n.$$

*步驟 2：多項之和。*規則 (ii) 只是就兩個多項式而言，但它可以推廣至任意有限個。若 $\triangle(f_1 + \cdots + f_m) = \triangle f_1 + \cdots + \triangle f_m$，則把規則 (ii) 應用於 $f_1 + \cdots + f_m$ 和 $f_{m+1}$ 這兩個多項式，得

$$\triangle(f_1 + \cdots + f_{m+1}) = \triangle(f_1 + \cdots + f_m) + \triangle f_{m+1} = \triangle f_1 + \cdots + \triangle f_{m+1},$$

所以對 $m$ 作歸納法，可知它對任意多項都成立。

步驟 3：綜合起來。任何多項式都可以寫成 $h(x) = \sum_{k=0}^{n} a_k x^k$，其中 $a_k$ 為常數。於是

$$\begin{aligned}
\triangle h(x) &= \sum_{k=0}^{n} \triangle\left(a_k x^k\right) && \text{根據規則 (ii)，並按步驟 2 推廣}\\
&= \sum_{k=0}^{n} a_k \triangle x^k && \text{根據規則 (iii)}\\
&= \sum_{k=1}^{n} k\, a_k\, x^{k-1} && \text{根據步驟 1（}k = 0\text{ 的項為 0）}\\
&= \frac{\mathrm{d}h(x)}{\mathrm{d}x},
\end{aligned}$$

最後一行就是多項式通常的導數。這對每個多項式 $h(x)$ 都成立，所以 $\triangle h(x) \equiv \dfrac{\mathrm{d}h(x)}{\mathrm{d}x}$，正如所求。

## 驗算

令人安心的是，普通的求導確實符合全部四條規則，所以這些規則並不互相矛盾，而 $\mathrm{d}/\mathrm{d}x$ 的確是題目所描述的那類運算。你也可以用公式來檢驗積法則：取 $f = x^2$ 和 $g = x^3$，規則 (iv) 給出 $x^2 \cdot 3x^2 + x^3 \cdot 2x = 5x^4$，這正是 $\triangle x^5$。我們在電腦上只用這些規則（如步驟 1 至 3）構造出 $\triangle$，並把它與導數作精確比較，共測試了 2000 個次數不超過 9 的隨機多項式；例如 $\triangle\left(\frac{1}{2}x^4 - 3x + 7\right) = 2x^3 - 3$，這也與數值差商（difference quotient）吻合。

## 常見錯誤

- 斷言 $\triangle(\text{常數}) = 0$，「因為它是導數」。這是預先假設了最終答案；這個結果必須由規則推導出來。
- 不加說明就把規則 (ii) 用於一長串的和。規則只是就兩個多項式而言，所以項數更多時，需要用一句話交代歸納法。
- 沒有說明哪一步用了哪條規則：題目明確要求這樣做。

## STEP 之外

符合線性和積法則的運算稱為**導子**（derivation），而這裏的論證顯示，在多項式上，導子完全由它在 $x$ 上的值決定。假如改為 $\triangle x = p(x)$，其中 $p$ 為某個多項式，同樣的證明會給出 $\triangle h = p(x)\, h'(x)$。求導可以純粹用代數方法定義，完全不涉及極限，這個想法在抽象代數（abstract algebra）和微分幾何（differential geometry）中佔有核心地位；在

微分幾何中，向量場（vector field）正正就是函數上的導子。Allenby 的 *Rings, Fields and Groups*（環、域與羣）是認識以代數方式處理多項式環（polynomial ring）的好去處。

詳細解答

# STEP III 2006，第 9 題

力學 • 能量守恆

**STEP III 2006，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：

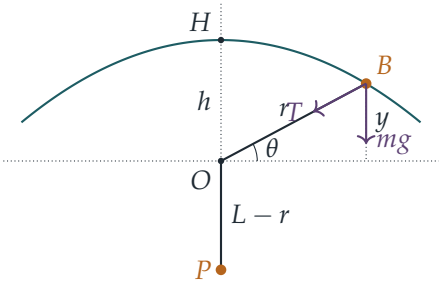
step.maths.org/questions/06-s3-q9

## 關鍵思路

繩的長度固定，所以小珠與 $O$ 的距離增加多少，質點 $P$ 就上升多少。這樣便把兩者的勢能（potential energy）連繫起來：總勢能只取決於 $y + r$。「在任何位置都處於平衡（equilibrium）」表示勢能並不偏好任何位置，所以它沿着金屬線必定是常數。既然勢能是常數，能量守恆（conservation of energy）便告訴我們總動能（kinetic energy）永不改變，餘下的工作只是用極坐標（polar coordinates）寫出小珠的速率。

## 解答

**模型**。$P$ 和 $B$ 是質量為 $m$ 的質點；繩是輕而不可伸長（light and inextensible）的；小環和金屬線都是光滑（smooth）的，所以繩的張力（tension）所作的淨功為零，而金屬線的法向反作用力（normal reaction）不作功。取 $O$ 為原點，$x$ 軸水平，$y$ 軸垂直向上，並設繩長為 $L$。



### 金屬線的形狀

由 $O$ 到 $B$ 的一段繩長為 $r$，所以 $OP$ 的長度是 $L-r$，即 $P$ 位於 $O$ 之下 $L-r$ 的距離。取通過 $O$ 的水平面為勢能的零點，總勢能為

$$U = mgy - mg(L-r) = mg(y+r) - mgL.$$

這個系統只有一個自由度（degree of freedom），即小珠在金屬線上的位置，而平衡位置就是 $U$ 沿金屬線的駐點（stationary point）。如果每一個位置都是平衡位置，那麼沿着金屬線處處都有 $\mathrm{d}U/\mathrm{d}s = 0$（其中 $s$ 是弧長（arc length)），所以 $U$ 沿金屬線是常數。因此 $y+r$ 是常數。

在 $H$ 處，小珠位於 $O$ 的正上方，高度為 $h$，所以該處 $y = h$ 且 $r = h$，這個常數就是 $2h$。於是金屬線的方程是

$$y + r = 2h,$$

正如題目所求。（寫成直角坐標形式，由 $x^2 + y^2 = (2h-y)^2$ 得 $y = h - x^2/(4h)$：這是一條以 $O$ 為焦點（focus）的拋物線（parabola），如圖所示。你也可以直接驗證平衡：由於 $P$ 處於平衡，張力是 $mg$；沿 $BO$ 方向的 $mg$ 與垂直向下的 $mg$ 的合力（resultant）平分這兩個方向之間的夾角，而根據拋物線的反射性質（reflection property），這正好是拋物線在該點的法線（normal）方向。）

### $\theta$ 的變化率

在金屬線上 $y = 2h - r$，所以

$$\sin\theta = \frac{y}{r} = \frac{2h}{r} - 1.$$

對時間求導，

$$\cos\theta\,\dot{\theta} = -\frac{2h\dot{r}}{r^2}.$$

由於 $\theta = \arcsin(y/r)$ 位於 $[-\frac{\pi}{2}, \frac{\pi}{2}]$ 之內，我們有 $\cos\theta \geq 0$，所以

$$\cos\theta = \sqrt{1 - \left(\frac{2h-r}{r}\right)^2} = \frac{\sqrt{r^2 - (2h-r)^2}}{r} = \frac{\sqrt{4rh - 4h^2}}{r} = \frac{2\sqrt{rh-h^2}}{r}.$$

小珠離開 $H$ 之後，我們有 $r > h$（因為 $r = 2h - y$，而在 $H$ 以外 $y < h$），所以 $\cos\theta \neq 0$，可以放心相除：

$$\dot{\theta} = -\frac{2h\dot{r}}{r^2} \cdot \frac{r}{2\sqrt{rh-h^2}} = -\frac{h\dot{r}}{r\sqrt{rh-h^2}},$$

這就是所求的結果。

### $P$ 的速率

**想一想** 為甚麼總動能必須一直等於 $\frac{1}{2}mV^2$？

對系統作功的力只有重力和繩的張力（屬內力）。張力所作的淨功為零：繩不可伸長，小環又光滑，所以張力對 $B$ 所作的功與對 $P$ 所作的功互相抵銷。金屬線是光滑的，所以它的反作用力不作功。因此動能與勢能之和守恆。但我們剛剛證明了勢能 $U$ 沿這條金屬線是常數，所以動能也是常數。

開始時小珠在 $H$，該處 $r$ 取最小值 $h$（因為 $y$ 在該處最大）。所以最初 $\dot{r}=0$，而 $P$ 的速率是 $|\dot{r}|$，故它由靜止開始運動。因此初始動能是 $\frac{1}{2}mV^2$。

在以 $O$ 為極點的極坐標中，小珠速率的平方是 $\dot{r}^2+r^2\dot{\theta}^2$。利用 $\dot{\theta}$ 的結果，

$$r^2\dot{\theta}^2=\frac{h^2\dot{r}^2}{rh-h^2}=\frac{h\dot{r}^2}{r-h},\qquad \text{所以}\qquad v_B^2=\dot{r}^2\Big(1+\frac{h}{r-h}\Big)=\frac{r\dot{r}^2}{r-h}.$$

由動能守恆，約去 $\frac{1}{2}m$ 之後得

$$\frac{r\dot{r}^2}{r-h}+\dot{r}^2=V^2\quad\Longrightarrow\quad \dot{r}^2\,\frac{2r-h}{r-h}=V^2.$$

由於 $r>h$，$r-h$ 和 $2r-h$ 都是正數，所以 $P$ 的速率是

$$|\dot{r}|=V\Big(\frac{r-h}{2r-h}\Big)^{\frac{1}{2}},$$

正如題目所求。

## 驗算

- 在 $H$ 處，$r=h$，公式給出 0：正確，因為 $P$ 由靜止開始。
- 當 $r\to\infty$ 時，$P$ 的速率趨向 $V/\sqrt{2}$。在很遠的地方，小珠幾乎沿徑向運動，所以 $B$ 和 $P$ 的速率幾乎相等，兩者平分動能。
- 數值驗證：取 $h=1$ 和 $V=2$，我們沿 $y=1-\frac{1}{4}x^2$ 逐步計算小珠的運動方程（用 RK4），並把 $\dot{\theta}$ 和 $\dot{r}$ 的有限差分值與公式比較。在 $r\approx 4.043$ 時，$P$ 的速率是 1.31064，與 $2\sqrt{3.043/7.087}$ 吻合至六位有效數字。

## 常見錯誤

- 把 $P$ 的勢能取為 $+mg(L-r)$：$P$ 在 $O$ 的下方，所以它的高度是 $-(L-r)$。
- 忘記 $P$ 也在運動，於是寫成 $\frac{1}{2}mv_B^2=\frac{1}{2}mV^2$。
- 除以 $\cos\theta$ 時，沒有指出它只在 $H$ 處才等於零。

## STEP 之外

尋找一條曲線，使系統在其上處處都處於「隨遇」平衡（neutral equilibrium），是「平衡位置即勢能的駐點」這個原理的一個小例子。這個原理通往拉格朗日力學（Lagrangian mechanics），在那裏，像繩和金屬線這樣的約束（constraints）是透過選取合適的坐標而直接納入其中的。在這一題，單一個坐標 $r$ 已經足夠，而能量守恆就是首次積分（first integral）。Kleppner 與 Kolenkow 所著的 *An Introduction to Mechanics*（力學入門）正是以這種精神講解能量方法和極坐標。

詳細解答

# STEP III 2006，第 10 題

力學 • 角動量守恆

> **STEP III 2006，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/06-s3-q10

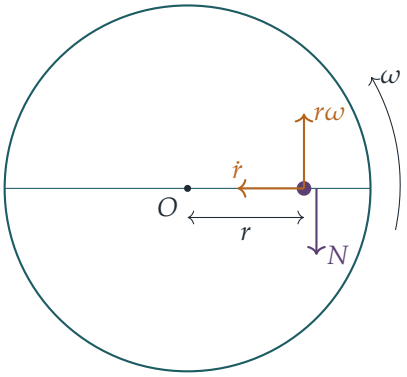
## 關鍵思路

沒有任何外來作用令圓盤與質點組成的系統繞軸轉動，而光滑凹槽所作的淨功為零，所以有兩條守恆定律：繞軸的角動量（angular momentum）守恆和動能（kinetic energy）守恆。由此可得 $\omega$ 和 $\dot{r}$ 均為只含 $r$ 的函數，而那個把 $V$、$\Omega$、$a$ 和 $k$ 連繫起來、看似古怪的條件，正是為了令能量的表達式大幅化簡。把 $\dot{r}$ 除以 $\dot{\theta} = \omega$ 便可消去時間，剩下一條描述路徑的可分離變量（separable）方程。

## 解答

### 模型

我們把質點視為質量為 $m$ 的點質量，並把凹槽視為光滑，所以質點所受的唯一水平力是垂直於凹槽的反作用力（reaction）$N$。圓盤繞固定的鉛垂軸自由轉動，所以軸承對該軸不施加力矩（torque）。重力和鉛垂方向的反作用力都是鉛垂的，不起任何作用。設 $\theta$ 為圓盤已轉過的角度，故 $\omega = \dot{\theta}$，並設 $r$ 為質點與軸的距離。由於質點一直留在凹槽內，它的速度有一個沿凹槽的分量 $\dot{r}$ 和一個垂直於凹槽的分量 $r\omega$。




### 角速率

反作用力 $N$ 作用於質點，而 $-N$ 則作用於圓盤上的同一點，所以兩者繞軸的力矩互相抵銷。既然沒有繞軸的外力矩，繞軸的總角動量便是常數。圓盤的貢獻是 $mk^2\omega$，質點的貢獻是 $mr\cdot r\omega = mr^2\omega$（只有其速度的橫向分量才有矩）。在 $t=0$ 時，總角動量為 $m(k^2+a^2)\Omega$，所以

$$(k^2+r^2)\,\omega = (k^2+a^2)\,\Omega, \qquad \text{即} \qquad \omega = \frac{\Omega(k^2+a^2)}{k^2+r^2},$$

這就是第一個結果。

### 徑向速率

**想一想** 質點運動時，凹槽從側面推着它。為甚麼能量仍然守恆？

反作用力 $N$ 垂直於凹槽。質點與圓盤上跟它接觸的那一點，在垂直於凹槽方向上的速度分量相同，都是 $r\omega$；兩者的差別只在於沿凹槽的滑動速度 $\dot r$，而這個速度垂直於 $N$。所以 $N$ 對質點所作的功，恰好被 $-N$ 對圓盤所作的功抵銷。軸上的軸承不作功，所以總動能是常數。

約去公因子 $\frac{1}{2}m$ 後，能量守恆給出

$$k^2\omega^2 + \dot r^2 + r^2\omega^2 = k^2\Omega^2 + V^2 + a^2\Omega^2.$$

現在把題目所給的條件 $V^2 = \Omega^2a^2(k^2+a^2)/k^2$ 用於右邊：

$$(k^2+a^2)\Omega^2 + \frac{\Omega^2a^2(k^2+a^2)}{k^2} = \frac{\Omega^2(k^2+a^2)(k^2+a^2)}{k^2} = \frac{\Omega^2(k^2+a^2)^2}{k^2}.$$

至於左邊，由第一個結果可得 $(k^2+r^2)\omega^2 = \Omega^2(k^2+a^2)^2/(k^2+r^2)$。因此

$$\begin{aligned}\dot{r}^2 &= \Omega^2(k^2+a^2)^2\left(\frac{1}{k^2}-\frac{1}{k^2+r^2}\right)\\ &= \frac{\Omega^2 r^2(k^2+a^2)^2}{k^2(k^2+r^2)},\end{aligned}$$

正是所求。

### 路徑的方程

質點向軸運動時，$\dot{r}<0$，所以取負平方根，

$$\dot{r} = -\frac{\Omega r(k^2+a^2)}{k\sqrt{k^2+r^2}}.$$

由於 $\omega = \dot{\theta} > 0$，我們可以使用鏈式法則（chain rule），$\dfrac{\mathrm{d}r}{\mathrm{d}\theta} = \dot{r}\Big/\dot{\theta} = \dot{r}/\omega$：

$$\frac{\mathrm{d}r}{\mathrm{d}\theta} = -\frac{\Omega r(k^2+a^2)}{k\sqrt{k^2+r^2}}\cdot\frac{k^2+r^2}{\Omega(k^2+a^2)} = -\frac{r\sqrt{k^2+r^2}}{k},$$

所以 $k\dfrac{\mathrm{d}r}{\mathrm{d}\theta} = -r(k^2+r^2)^{1/2}$，正是所求。

### 求解 $r$

設 $u = k/r$，故 $u>0$。於是

$$\frac{\mathrm{d}u}{\mathrm{d}\theta} = -\frac{k}{r^2}\frac{\mathrm{d}r}{\mathrm{d}\theta} = \frac{k}{r^2}\cdot\frac{r\sqrt{k^2+r^2}}{k} = \sqrt{\frac{k^2}{r^2}+1} = \sqrt{1+u^2}.$$

這條方程可分離變量：

$$\int\frac{\mathrm{d}u}{\sqrt{1+u^2}} = \int\mathrm{d}\theta \quad\Longrightarrow\quad \operatorname{arsinh}u = \theta+\alpha$$

其中 $\alpha$ 為某個常數。所以 $k/r = \sinh(\theta+\alpha)$，即

$$r\sinh(\theta+\alpha) = k.$$

在 $t=0$ 時，$\theta=0$ 而 $r=a$，所以 $\sinh\alpha = k/a$，正是所求。

### 質點永遠不會到達軸

**想一想** $r = k/\sinh(\theta+\alpha)$ 永不為零，但 $\theta$ 會否在有限時間內變成無窮大？

不會。由於 $r^2 \geq 0$，第一個結果給出 $\omega \leq \Omega(k^2+a^2)/k^2$，所以 $\theta \leq \Omega(k^2+a^2)t/k^2$ 在每一個有限的時刻 $t$ 都是有限的。於是在任何時刻都有 $r = k/\sinh(\theta+\alpha) > 0$。

我們還應檢查這些公式是否一直成立，即質點是否一直向軸運動。根據 $\dot{r}^2$ 的公式，$\dot{r}$ 只有在 $r = 0$ 時才可能等於零。所以只要 $r > 0$，$\dot{r}$ 便保持最初的負號，而上述論證對所有 $t > 0$ 都適用。質點沿螺旋線不斷向軸靠近，只有當 $\theta \to \infty$ 時才有 $r \to 0$，但它永遠不會到達軸。

## 驗算

- 在 $r = a$ 時，$\dot{r}^2$ 的公式給出 $\Omega^2 a^2(k^2+a^2)/k^2 = V^2$，與初始條件吻合。
- 取 $k = a = \Omega = 1$（故 $V = \sqrt{2}$），我們以數值方法對運動方程 $\ddot{r} = r\omega^2$ 和 $(k^2+r^2)\omega =$ 常數作積分。當圓盤轉過 $\theta \approx 4.4170$ 後，質點到達 $r = 0.01$，而 $\operatorname{arsinh} 100 - \operatorname{arsinh} 1 \approx 5.2983 - 0.8814 = 4.4169$ 與此相符。在整個計算過程中，能量、$\dot{r}^2$ 和 $r\sinh(\theta+\alpha) = k$ 全都成立，誤差約為 $10^{-11}$。

## 常見錯誤

- 忘記質點本身的角動量 $mr^2\omega$，或它的橫向動能 $\frac{1}{2}mr^2\omega^2$。
- 對 $\dot{r}$ 取正平方根，令 $\mathrm{d}r/\mathrm{d}\theta$ 的正負號出錯。
- 單憑 $r \neq 0$ 便斷定「永遠不會到達軸」，而沒有指出 $\theta$ 在有限時間內保持有限。

## STEP 之外

質點不受任何徑向力，但在圓盤的旋轉參考系（rotating frame）中，它卻似乎被向外推。這就是 $\ddot{r} = r\omega^2$ 中的離心（centrifugal）項，而側向的反作用力 $N$ 則與科里奧利力（Coriolis force）平衡。大學力學會用拉格朗日（Lagrange）方法處理這類問題：角度 $\theta$ 不出現在能量之中，所以它的共軛動量（conjugate momentum），在這裏即角動量，便自動守恆。Kleppner 與 Kolenkow 所著的 *An Introduction to Mechanics*（力學入門）以恰當的程度講解旋轉參考系和角動量。

詳細解答

# STEP III 2006，第 11 題

力學 • 連接系統中的碰撞

> **STEP III 2006，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/06-s3-q11

## 關鍵思路

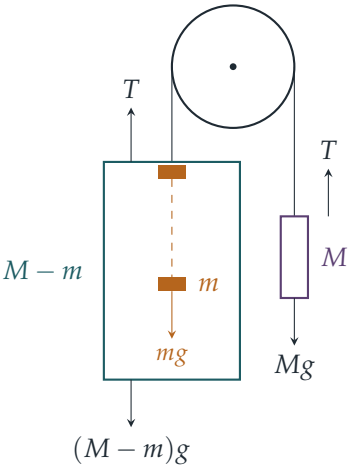
瓷磚一旦鬆脫，升降機便比平衡錘輕，於是升降機開始上升，而瓷磚則下墜：瓷磚要走過的，是相對於升降機的高度 $h$。至於碰撞，升降機與平衡錘可視為一個質量為 $2M - m$ 的單一物體，經地板與瓷磚相連；題目對纜索衝量所給的條件，正好讓我們為這個物體和瓷磚寫出一條動量式的守恆定律。這個「動量」在碰撞前一刻原來剛好是零，令能量的計算變得十分簡潔。

## 解答

我們把升降機、平衡錘和瓷磚都當作質點（particle）處理（只考慮垂直方向的運動），纜索為輕而不可伸長（light and inextensible），滑輪為光滑（smooth）。瓷磚脫落之前，系統處於靜止狀態，所以連同瓷磚的升降機與平衡錘互相平衡：升降機連瓷磚的質量為 $M$，不連瓷磚則為 $M - m$。

### 下墜

對升降機和瓷磚，取向上為正。由於纜索不可伸長，當升降機以速率 $u$、加速度 $a$ 向上運動時，平衡錘便以相同的速率和加速度向下運動。設張力（tension）為 $T$。




對升降機（向上）和平衡錘（向下）應用牛頓第二定律（Newton's second law），得

$$(M-m)a = T-(M-m)g, \qquad Ma = Mg - T.$$

兩式相加，得 $(2M-m)a = mg$，所以

$$a = \frac{mg}{2M-m}.$$

瓷磚作自由落體運動（free fall），加速度為向下的 $g$（導軌光滑，亦沒有其他東西碰到它）。因此，相對於升降機，瓷磚向下的加速度為

$$g + a = g + \frac{mg}{2M-m} = \frac{2Mg}{2M-m},$$

而它相對於升降機是由靜止開始運動的。設它經過時間 $t$ 後碰到地板，則 $\frac{1}{2}(g+a)t^2 = h$，所以

$$t^2 = \frac{(2M-m)h}{Mg}.$$

在那一刻，瓷磚的速率為

$$v = gt = \sqrt{g^2 \cdot \frac{(2M-m)h}{Mg}} = \sqrt{\frac{(2M-m)gh}{M}},$$

正如題目所求。此時升降機（和平衡錘）的速率為

$$U = at = \frac{m}{2M-m}gt = \frac{mv}{2M-m}.$$

在這段時間內，升降機上升了 $\frac{1}{2}at^2 = \frac{mh}{2M} < h$，所以題目對升降機與滑輪距離所設的條件，保證了升降機不會先碰到滑輪。

### 碰撞

設 $J$ 為地板作用於瓷磚的衝量（impulse）（向上），又設 $I$ 為纜索張力的衝量；根據題目所給的條件，這個衝量作用於升降機和作用於平衡錘的大小相同（兩者都是向上）。碰撞在瞬間完成，重力在此期間不產生衝量。設碰撞後一刻瓷磚的速度為 $w$（向上），升降機的速度為 $U'$（向上），則平衡錘向下的速度同樣是 $U'$。於是：

$$\begin{aligned} \text{瓷磚：}\quad & m\big(w-(-v)\big) = J, \\ \text{升降機：}\quad & (M-m)(U'-U) = I - J, \\ \text{平衡錘（向下）：}\quad & M(U'-U) = -I. \end{aligned}$$

三式相加，便消去 $J$ 和 $I$：

$$mw + (2M-m)U' = -mv + (2M-m)U.$$

**想一想** 平衡錘的方程以「向下」為正，為甚麼可以把它與另外兩條方程相加？

因為有纜索相連，「平衡錘向下」與「升降機向上」其實是同一個坐標。我們所證明的是：$mw + (2M-m)U'$ 這個組合（瓷磚向上的動量，加上升降機與平衡錘沿纜索方向量度的動量）在碰撞前後保持不變。衝量相等這個條件正是用在這裏：沒有它，$I$ 便不會抵消。

現在是關鍵的觀察：由於 $U = \dfrac{mv}{2M-m}$，右邊等於 $-mv + mv = 0$。所以

$$mw + (2M-m)U' = 0 = -mv + (2M-m)U.$$

在地板處應用牛頓實驗定律（Newton's experimental law）：分離速率等於接近速率的 $e$ 倍，即

$$w - U' = e\big(U - (-v)\big) = e(U+v).$$

碰撞前和碰撞後的速度都滿足同一條「零動量」方程，所以 $(w, U')$ 和 $(-v, U)$ 這兩對數各自都是 $(2M-m,\ -m)$ 的倍數。恢復方程隨即決定了這個倍數：$w = ev$ 及 $U' = -eU$。（驗算：$mev - (2M-m)eU = e(mv - mv) = 0$，且 $w - U' = e(v+U)$。）

### 能量損失

碰撞前一刻，動能（kinetic energy）為

$$\tfrac{1}{2}mv^2 + \tfrac{1}{2}(2M-m)U^2 = \tfrac{1}{2}mv^2\left(1 + \frac{m}{2M-m}\right) = \frac{Mmv^2}{2M-m} = mgh,$$

其中先用了 $U = \frac{mv}{2M-m}$，再用了 $v^2 = \frac{(2M-m)gh}{M}$。碰撞後一刻，每個速度都乘上了 $\pm e$，所以動能為 $e^2\,mgh$。碰撞在瞬間完成，期間勢能（potential energy）不變，因此系統的能量損失為

$$mgh - e^2 mgh = mgh(1-e^2),$$

正如題目所求。

### 評註

答案完全與 $M$ 無關：它正好等於瓷磚在一部固定不動的升降機內下墜高度 $h$、再以恢復系數（coefficient of restitution）$e$ 反彈時的能量損失。這是合理的。碰撞前一刻，系統的動能為 $mgh$，這正是所釋放的勢能（瓷磚與升降機之間的距離縮短了 $h$，而升降機上升的高度恰好被平衡錘下降的高度抵消）。又因為沿纜索方向的「動量」為零，這次碰撞就如同在零動量參考系（zero-momentum frame）中觀察的碰撞，而在這個參考系中，損失的動能所佔的比例永遠是 $1-e^2$。特別地，若 $e=0$，全部動能都會損失，整個系統靜止下來；若 $e=1$，則毫無損失。

## 驗算

- 取 $M=5$、$m=2$、$h=3$、$g=10$、$e=\frac{1}{3}$。則 $v^2=48$ 及 $U^2=3$，所以碰撞前的動能為 $\frac{1}{2}\cdot 2\cdot 48+\frac{1}{2}\cdot 8\cdot 3=60=mgh$，而損失為 $60\cdot\frac{8}{9}=\frac{160}{3}$。
- 以數值方法逐步模擬這個運動（瓷磚自由下墜，升降機以 $\frac{mg}{2M-m}$ 向上加速），隨機選取 200 組 $M$、$m$、$h$、$e$ 的值，所得的碰撞速率 $v$、升降機速率 $U$ 和能量損失 $mgh(1-e^2)$ 均與上述結果吻合。
- 當 $m\to 0$ 時，升降機幾乎不動，而 $v\to\sqrt{2gh}$，即自由落體的速率，與預期相符。
- 張力的衝量 $I=M(U-U')=M(1+e)U$ 為正數，所以纜索在碰撞期間保持拉緊。

## 常見錯誤

- 瓷磚脫落之後，仍把升降機的質量當作 $M$。這樣的話，甚麼都不會加速，瓷磚的速率便會是 $\sqrt{2gh}$。
- 忘記地板會上升迎向瓷磚，所以瓷磚相對於地面下墜的距離小於 $h$。
- 只對瓷磚和升降機應用動量守恆（conservation of momentum），忘記了經纜索與升降機相連的平衡錘。

## STEP 之外

碰撞中守恆的量並不是普通的動量，而是與描述纜索位置的那個單一坐標共軛的動量：這是初嘗廣義坐標（generalised coordinates）和拉格朗日力學（Lagrangian mechanics）的滋味。總動量為零的碰撞恰好損失動能的 $1-e^2$ 這個比例，這個事實正是貫穿整個碰撞理論的質心參考系（centre-of-mass frame）思想。Kleppner 與 Kolenkow 合著的 *An Introduction to Mechanics*（力學入門）對質心參考系中的動量和連結體系統都有細緻的論

述。

詳細解答

# STEP III 2006，第 12 題

概率與統計 • 正態近似與期望值

**STEP III 2006，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/06-s3-q12

## 關鍵思路

多加一輛巴士，其價值只在於它平均能賺到的錢，所以公司應考慮的牌照費，最多只能是每年額外利潤的*期望值*（expected value）。在現有車隊已足以應付的日子，新增的巴士一分錢也賺不到；而它載客再多，也不會超過車上的 32 個座位。因此我們要求的，是一個「被截斷」的隨機變量（random variable）的期望值。題目的數字經過刻意挑選：在正態近似（normal approximation）之下，新舊載客量剛好落在分佈上非常整齊的位置。

## 解答

### 建立模型

考慮單獨一次抵達，設 $X$ 為 1024 名遊客中被指示前往巴士站的人數。每名遊客前往巴士站的概率為 $\frac{1}{2}$，並假設各遊客獲得的指示彼此獨立（independent）。於是

$$X \sim \mathrm{B}\left(1024, \tfrac{1}{2}\right), \qquad \mathrm{E}(X) = 512, \qquad \mathrm{Var}(X) = 1024 \cdot \tfrac{1}{2} \cdot \tfrac{1}{2} = 256.$$

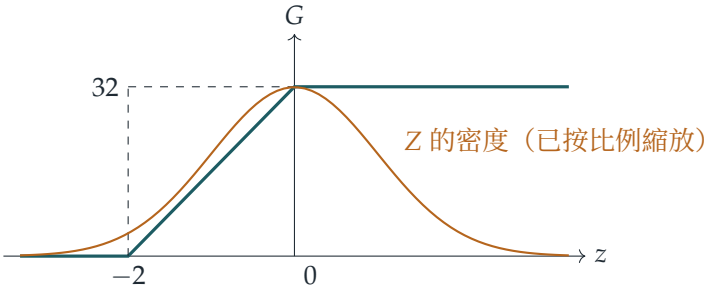
由於 $n$ 很大而 $p = \frac{1}{2}$，我們採用正態近似 $X \approx \mathrm{N}(512, 16^2)$，並寫成 $X = 512 + 16Z$，其中 $Z \sim \mathrm{N}(0,1)$。按題目指示，我們不作連續性校正（continuity correction）。

15 輛巴士可載 $15 \times 32 = 480$ 名乘客，所以現時每次抵達的利潤是 $£\min(X, 480)$。加上第 16 輛巴士後，載客量是 512，利潤是 $£\min(X, 512)$。因此每次抵達的額外利潤是

$$G = \min(X,512) - \min(X,480) = \begin{cases} 0, & X \le 480, \\ X - 480, & 480 \le X \le 512, \\ 32, & X \ge 512. \end{cases}$$

以 $Z$ 表示，兩個分界點分別是 $X = 480 \iff Z = -2$ 和 $X = 512 \iff Z = 0$，而在中間的範圍內 $X - 480 = 32 + 16Z$。所以

$$G = \begin{cases} 0, & Z \le -2, \\ 32 + 16Z, & -2 \le Z \le 0, \\ 32, & Z \ge 0. \end{cases}$$




**額外利潤的期望值**

設 $\phi(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2}$ 為標準正態密度函數（standard normal density）。把期望值按三個範圍拆開，得

$$\mathrm{E}(G) = 32\,\mathrm{P}(Z \ge 0) + \int_{-2}^{0} (32 + 16z)\,\phi(z)\,\mathrm{d}z.$$

第一項是 $32 \times \frac{1}{2} = 16$。至於積分，常數部分給出

$$32 \int_{-2}^{0} \phi(z)\,\mathrm{d}z = 32\big(\Phi(0) - \Phi(-2)\big) = 32\big(\tfrac{1}{2} - (1 - \Phi(2))\big) = 32\Phi(2) - 16,$$

其中用到了對稱性 $\Phi(-2) = 1 - \Phi(2)$。

? **想一想** 沒有數表，怎樣求 $z\,\phi(z)$ 的積分？

對密度函數求導：$\phi'(z) = -z\,\phi(z)$。所以 $z\,\phi(z)$ 有精確的原函數(antiderivative) $-\phi(z)$，於是

$$16 \int_{-2}^{0} z\,\phi(z)\,\mathrm{d}z = 16\big[-\phi(z)\big]_{-2}^{0} = -16\big(\phi(0) - \phi(2)\big) = -\frac{16}{\sqrt{2\pi}}\big(1 - \mathrm{e}^{-2}\big).$$

(結果是負數，這正合道理：在整個範圍內 $z < 0$。)

把三部分相加，兩個 16 互相抵消：

$$\mathrm{E}(G) = 16 + \big(32\Phi(2) - 16\big) - \frac{16}{\sqrt{2\pi}}\big(1 - \mathrm{e}^{-2}\big) = 32\Phi(2) - \frac{16}{\sqrt{2\pi}}\big(1 - \mathrm{e}^{-2}\big).$$

### 牌照費

每年有 50 次抵達，所以根據期望值的線性性質（linearity of expectation），第 16 輛巴士每年帶來的額外利潤期望值是

$$50\,\mathrm{E}(G) = 1600\,\Phi(2) - \frac{800}{\sqrt{2\pi}}\left(1 - \mathrm{e}^{-2}\right) \text{ 英鎊.}$$

若所付的牌照費高於此數，公司平均而言會虧本；而任何較低的費用都會留下期望收益。所以這就是值得考慮的最高年費（只是近似值，因為用了正態近似），亦即所求的結果。其數值約為 £1288。

## 驗算

- 合理範圍：新增的巴士每程賺取的金額介乎 0 與 32 之間，所以全年的數字必定介乎 0 與 1600 之間。我們的數值 $1600 \times 0.97725 - 319.15 \times 0.86466 \approx 1287.6$ 確實如此。
- 精確計算：以精確的 $\mathrm{B}(1024, \frac{1}{2})$ 概率對 $\min(k, 512) - \min(k, 480)$ 求和，再乘以 50，得 1287.69，與正態近似所得的 1287.64 相差不超過 6 便士。

## 常見錯誤

- 把額外利潤當作 $32\,\mathrm{P}(X > 480)$，彷彿只要用得着新巴士，它就一定滿座。其實在巴士乘客人數介乎 480 與 512 之間的日子，它只坐滿了一部分。
- 計算 $\int_{-2}^{0} z\phi(z)\,\mathrm{d}z$ 時弄錯正負號：被積函數是負的，所以答案必定是負數。

## STEP 之外

$\mathrm{E}\big(\min(X, c)\big)$ 這個量，就是存貨為 $c$ 的報販的期望銷量；而如何選擇 $c$ 的「報販問題」（newsvendor problem），是運籌學（operations research）的經典課題。同樣的截斷收益 $\min(\max(X - a, 0), b - a)$，在金融數學中稱為「看漲期權價差」（call spread）。$\phi'(z) = -z\phi(z)$ 這個技巧，是截斷正態分佈的矩（truncated normal moments）理論的起點。Grimmett 與 Welsh 合著的 *Probability: An Introduction*（概率論入門）涵蓋了二項分佈（binomial distribution）的正態近似，以及隨機變量函數的期望值。

詳細解答

# STEP III 2006，第 13 題

概率與統計 • 幾何概率

> **STEP III 2006，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/06-s3-q13

## 關鍵思路

周界由兩段性質迥異的部分組成：一段長度為 $\pi$ 的弧和一條長度為 2 的直徑，所以我們按每一點落在哪一段來取條件（condition）。在三種情況的每一種之中，三角形的面積都有簡單的公式 $\frac{1}{2}ab\sin C$，其中頂點 $C$ 位於圓心，而各個條件期望值（conditional expectation）都只是簡短的積分。最後用全期望定律（law of total expectation）把各部分重新合併起來。

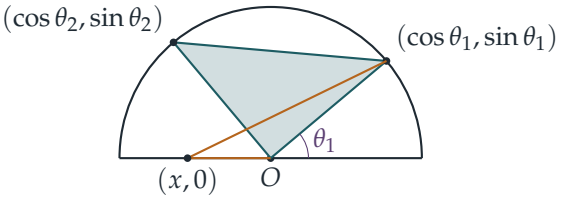
## 解答

### 設定

把直徑的中點 $O$ 放在原點，直徑沿 $x$ 軸由 $(-1,0)$ 伸展至 $(1,0)$，而弧則位於上半平面。「在周界上隨機選取」的意思是沿整條邊界按長度均勻分佈（uniformly distributed），邊界的總長度為 $\pi+2$。因此每一點各自獨立地

- 以概率 $\dfrac{\pi}{\pi+2}$ 落在弧上，這時它是 $(\cos\theta,\sin\theta)$，其中 $\theta$ 在 $[0,\pi]$ 上均勻分佈；
- 以概率 $\dfrac{2}{\pi+2}$ 落在直徑上，這時它是 $(x,0)$，其中 $x$ 在 $[-1,1]$ 上均勻分佈。

（在弧上按長度均勻分佈，等同於按角度均勻分佈，因為單位圓上的弧長等於所對的角。）




設 $D$ 為兩點之中落在直徑上的點的數目。那麼 $D = 2$、$1$、$0$ 的概率分別為

$$\frac{4}{(\pi+2)^2}, \qquad \frac{2\cdot 2\pi}{(\pi+2)^2} = \frac{4\pi}{(\pi+2)^2}, \qquad \frac{\pi^2}{(\pi+2)^2},$$

這裏用到兩點互相獨立（independent）　（中間一種情況的因子 2 是計算哪一點落在直徑上）　。在每一種情況下，三角形都有一個頂點在 $O$，所以它的面積是 $\frac{1}{2}|OP|\,|OQ|\sin\angle POQ$。

### 兩點都在直徑上

三個頂點全在同一直線上，所以三角形是退化（degenerate）的，$A = 0$。因此 $\mathrm{E}(A \mid D = 2) = 0$。

### 一點在弧上，一點在直徑上

設這兩點為 $P = (\cos\theta, \sin\theta)$ 和 $R = (x, 0)$。那麼 $|OP| = 1$，$|OR| = |x|$，而底邊 $OR$ 沿着直徑，$P$ 則在它上方高度 $\sin\theta \geq 0$ 之處，所以

$$A = \tfrac{1}{2}|x|\sin\theta.$$

由於 $x$ 與 $\theta$ 互相獨立，

$$\mathrm{E}(A \mid D = 1) = \tfrac{1}{2}\,\mathrm{E}|x|\,\mathrm{E}(\sin\theta) = \tfrac{1}{2}\cdot\frac{1}{2}\int_{-1}^{1}|x|\,\mathrm{d}x\cdot\frac{1}{\pi}\int_0^{\pi}\sin\theta\,\mathrm{d}\theta = \tfrac{1}{2}\cdot\tfrac{1}{2}\cdot\frac{2}{\pi} = \frac{1}{2\pi}.$$

### 兩點都在弧上

現在 $P = (\cos\theta_1, \sin\theta_1)$ 和 $Q = (\cos\theta_2, \sin\theta_2)$，其中 $\theta_1, \theta_2$ 互相獨立，並在 $[0, \pi]$ 上均勻分佈。兩條半徑的長度都是 1，而角 $POQ$ 等於 $|\theta_1 - \theta_2|$，所以

$$A = \tfrac{1}{2}\,|\sin(\theta_1 - \theta_2)|.$$

? **想一想** 要求出 $\mathrm{E}|\sin(\theta_1 - \theta_2)|$，我們是否需要知道 $\theta_1 - \theta_2$ 的分佈？

不需要：先對 $\theta_2$ 取條件。對於固定的 $\theta_2$，當 $\theta_1$ 走遍 $[0, \pi]$ 時，角 $\theta_1 - \theta_2$ 走遍一個長度為 $\pi$ 的區間。函數 $|\sin u|$ 的週期（period）為 $\pi$，所以它在任何長度為 $\pi$ 的區間上的積

分，都與在 $[0,\pi]$ 上的積分相同：

$$\int_0^{\pi} |\sin(\theta_1-\theta_2)|\,\mathrm{d}\theta_1 = \int_0^{\pi} \sin u\,\mathrm{d}u = 2.$$

所以無論 $\theta_2$ 是甚麼，都有 $\mathrm{E}\big(|\sin(\theta_1-\theta_2)| \,\big|\, \theta_2\big) = 2/\pi$，因此

$$\mathrm{E}(A \mid D=0) = \tfrac{1}{2}\cdot\frac{2}{\pi} = \frac{1}{\pi}.$$

### 合併結果

根據全期望定律，

$$\begin{aligned}\mathrm{E}(A) &= \frac{4}{(\pi+2)^2}\cdot 0 + \frac{4\pi}{(\pi+2)^2}\cdot\frac{1}{2\pi} + \frac{\pi^2}{(\pi+2)^2}\cdot\frac{1}{\pi}\\ &= \frac{2+\pi}{(\pi+2)^2} = \frac{1}{2+\pi},\end{aligned}$$

這就是 $(2+\pi)^{-1}$，正如題目所求。

## 驗算

- 三種情況的概率之和 $\dfrac{4+4\pi+\pi^2}{(\pi+2)^2}$ 等於 1。
- 每個條件期望值都不超過該情況下可能出現的最大面積：$\frac{1}{\pi}\approx 0.318 \le \frac{1}{2}$（兩條半徑互相垂直）以及 $\frac{1}{2\pi}\approx 0.159 \le \frac{1}{2}$。
- 用 $10^6$ 對點進行蒙地卡羅模擬（Monte Carlo simulation），得出 $\mathrm{E}(A)\approx 0.1943$，而 $1/(2+\pi)\approx 0.1945$；數值積分（quadrature）亦證實了條件期望值 $1/\pi$ 和 $1/(2\pi)$。

## 常見錯誤

- 以各 $\frac{1}{2}$ 的概率選擇每一點落在弧上或直徑上。「在周界上隨機選取」是按每一段的長度來加權的。
- 在 $\mathrm{P}(D=1)$ 中漏掉因子 2：兩點中任何一點都可以是落在直徑上的那一點。
- 寫成 $A=\frac{1}{2}\sin(\theta_1-\theta_2)$ 而漏掉絕對值（modulus），這樣由對稱性得出的期望值會是 0。

## STEP 之外

這是幾何概率（geometric probability）的一個小例子，這門學問的經典例子包括蒲豐投針問題（Buffon's needle）和貝特朗悖論（Bertrand's paradox）。貝特朗悖論說明了為何「隨機」一詞必須界定清楚，正如我們以沿周界的長度作為測度（measure）那樣。這裏真正發揮作用的工具，即對一個劃分（partition）取條件然後求平均（全期望定律），是大學一年級概率論的核心內容。Grimmett 與 Welsh 合著的 *Probability: An Introduction*（概率論入門教材）對此有很好的講解。

詳細解答

# STEP III 2006，第 14 題

概率與統計 • 協方差與獨立性

> **STEP III 2006，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/06-s3-q14

## 關鍵思路

這裏的一切都建基於兩個事實：期望值（expectation）是線性的；而對於獨立（independent)的變量，乘積的期望值等於期望值的乘積。方差(variance)可由 $\mathrm{Var}(Y) = \mathrm{E}(Y^2) - [\mathrm{E}(Y)]^2$ 求得，所以我們自始至終只需要計算期望值。要證明兩個變量並非獨立，只須找出獨立性的一個推論不成立便已足夠，而最容易檢驗的就是 $\mathrm{E}(UV) = \mathrm{E}(U)\,\mathrm{E}(V)$。

## 解答

### 兩個事實

對於任意隨機變量（random variable）及常數 $a$、$b$，

$$\mathrm{E}(aX_1 + bX_2) = a\,\mathrm{E}(X_1) + b\,\mathrm{E}(X_2),$$

而若 $X_1$ 與 $X_2$ 獨立，則

$$\mathrm{E}(X_1X_2) = \mathrm{E}(X_1)\,\mathrm{E}(X_2).$$

第二個事實同樣適用於這些變量的函數：若 $X_1$、$X_2$ 獨立，則 $X_1^2$ 與 $X_2$、$X_1$ 與 $X_2^2$，以及 $X_1^2$ 與 $X_2^2$ 亦各自獨立。我們還會用到 $\mathrm{E}(X_i^2) = \sigma_i^2 + \mu_i^2$，這不過是把 $\mathrm{Var}(X_i) = \mathrm{E}(X_i^2) - \mu_i^2$ 移項而已。

### 周界與面積

周界為 $P = 2X_1 + 2X_2$，面積為 $A = X_1X_2$。由線性及獨立性，

$$\mathrm{E}(P) = 2(\mu_1 + \mu_2), \qquad \mathrm{E}(A) = \mu_1\mu_2.$$

要求 $P$ 的方差，我們從 $\mathrm{E}(P^2)$ 入手：

$$\begin{aligned}\mathrm{E}(P^2) &= 4\,\mathrm{E}(X_1^2 + 2X_1X_2 + X_2^2)\\ &= 4(\sigma_1^2 + \mu_1^2 + 2\mu_1\mu_2 + \sigma_2^2 + \mu_2^2)\\ &= 4(\sigma_1^2 + \sigma_2^2) + 4(\mu_1 + \mu_2)^2,\end{aligned}$$

其中用到線性及 $\mathrm{E}(X_1X_2) = \mu_1\mu_2$。減去 $[\mathrm{E}(P)]^2 = 4(\mu_1 + \mu_2)^2$，得

$$\mathrm{Var}(P) = 4(\sigma_1^2 + \sigma_2^2), \qquad \text{所以}P\text{ 的標準差（standard deviation）為}2\sqrt{\sigma_1^2 + \sigma_2^2}.$$

至於 $A$，由於 $X_1^2$ 與 $X_2^2$ 獨立，

$$\mathrm{E}(A^2) = \mathrm{E}(X_1^2)\,\mathrm{E}(X_2^2) = (\sigma_1^2 + \mu_1^2)(\sigma_2^2 + \mu_2^2),$$

減去 $\mu_1^2\mu_2^2$ 後餘下

$$\mathrm{Var}(A) = \sigma_1^2\sigma_2^2 + \mu_1^2\sigma_2^2 + \mu_2^2\sigma_1^2,$$

所以 $A$ 的標準差為

$$\sqrt{\sigma_1^2\sigma_2^2 + \mu_1^2\sigma_2^2 + \mu_2^2\sigma_1^2}.$$

**$P$ 與 $A$ 並非獨立**

假如 $P$ 與 $A$ 獨立，便會有 $\mathrm{E}(PA) = \mathrm{E}(P)\,\mathrm{E}(A)$。現在 $PA = 2X_1^2X_2 + 2X_1X_2^2$，所以由線性以及 $X_1$ 與 $X_2$ 的獨立性，

$$\begin{aligned}\mathrm{E}(PA) &= 2(\sigma_1^2 + \mu_1^2)\mu_2 + 2\mu_1(\sigma_2^2 + \mu_2^2)\\ &= 2(\mu_1 + \mu_2)\mu_1\mu_2 + 2(\mu_2\sigma_1^2 + \mu_1\sigma_2^2)\\ &= \mathrm{E}(P)\,\mathrm{E}(A) + 2(\mu_2\sigma_1^2 + \mu_1\sigma_2^2).\end{aligned}$$

長度和闊度都是正數，所以 $\mu_1, \mu_2 > 0$，而且 $\sigma_1, \sigma_2 \neq 0$，因此多出來的一項 $2(\mu_2\sigma_1^2 + \mu_1\sigma_2^2)$ 嚴格為正。故此 $\mathrm{E}(PA) \neq \mathrm{E}(P)\,\mathrm{E}(A)$，即 $P$ 與 $A$ 並非獨立。

**變量 $Z = P - \alpha A$**

同樣地，若 $Z$ 與 $A$ 獨立，則必有 $\mathrm{E}(ZA) = \mathrm{E}(Z)\,\mathrm{E}(A)$。由線性，

$$\begin{aligned}\mathrm{E}(ZA) - \mathrm{E}(Z)\,\mathrm{E}(A) &= \big[\mathrm{E}(PA) - \mathrm{E}(P)\,\mathrm{E}(A)\big] - \alpha\big[\mathrm{E}(A^2) - [\mathrm{E}(A)]^2\big]\\ &= 2(\mu_2\sigma_1^2 + \mu_1\sigma_2^2) - \alpha(\mu_1^2\sigma_2^2 + \mu_2^2\sigma_1^2 + \sigma_1^2\sigma_2^2),\end{aligned}$$

其中用到上面的兩個結果。與 $\alpha$ 相乘的括號就是 $\mathrm{Var}(A)$，它是正數（當中含有 $\sigma_1^2\sigma_2^2 > 0$），所以這個式子恰好只在 $\alpha$ 的一個值處等於零：

$$\alpha_0 = \frac{2(\mu_1\sigma_2^2 + \mu_2\sigma_1^2)}{\mu_1^2\sigma_2^2 + \mu_2^2\sigma_1^2 + \sigma_1^2\sigma_2^2}.$$

若 $\alpha \neq \alpha_0$，則 $\mathrm{E}(ZA) \neq \mathrm{E}(Z)\,\mathrm{E}(A)$，所以 $Z$ 與 $A$ 並非獨立，這正是所要證明的。

**想一想** $\alpha = \alpha_0$ 是否表示 $Z$ 與 $A$ 確實獨立？

不是。協方差（covariance）為零是獨立性的推論，卻不能保證獨立。當 $\alpha = \alpha_0$ 時，我們的檢驗根本提供不了任何資訊，必須另覓論證。最後一部分說明，在一個具體例子中，獨立性在該處同樣不成立。

### 最後一部分

這裏 $\mu_1 = \mu_2 = 2$ 而 $\sigma_1^2 = \sigma_2^2 = 1$，所以

$$\alpha_0 = \frac{2(2+2)}{4+4+1} = \frac{8}{9}.$$

由上一部分，對任何 $\alpha \neq \frac{8}{9}$，$Z$ 與 $A$ 都並非獨立。若要用一個論證涵蓋所有 $\alpha$，可列出四個等可能的結果：

| $(X_1, X_2)$ | $A$ | $P$ | $Z = P - \alpha A$ |
|---|---|---|---|
| $(1,1)$ | 1 | 4 | $4 - \alpha$ |
| $(1,3), (3,1)$ | 3 | 8 | $8 - 3\alpha$ |
| $(3,3)$ | 9 | 12 | $12 - 9\alpha$ |

關鍵的觀察是：$P$ 完全由 $A$ 決定，因而 $Z$ 也是如此。設 $z = 4 - \alpha$。知道 $A = 1$ 便必有 $Z = z$，所以

$$\mathrm{P}(Z = z,\ A = 1) = \mathrm{P}(A = 1) = \tfrac{1}{4}.$$

若要獨立，這個概率便須等於 $\mathrm{P}(Z = z)\,\mathrm{P}(A = 1)$，即 $\mathrm{P}(Z = z) = 1$，於是 $Z$ 在每一個結果上都要取值 $4 - \alpha$。這需要 $8 - 3\alpha = 4 - \alpha$，即 $\alpha = 2$，同時又需要 $12 - 9\alpha = 4 - \alpha$，即 $\alpha = 1$。兩者不可能同時成立，所以 $\mathrm{P}(Z = z) < 1$，獨立性不成立。

因此，無論 $\alpha$ 取甚麼值，$Z$ 與 $A$ 都並非獨立。例如當 $\alpha = \frac{8}{9}$ 時，$Z$ 的取值為 $\frac{28}{9}$、$\frac{16}{3}$ 和 4，而 $\mathrm{P}(Z = \frac{28}{9}, A = 1) = \frac{1}{4}$，但 $\mathrm{P}(Z = \frac{28}{9})\,\mathrm{P}(A = 1) = \frac{1}{16}$。

## 驗算

- 若完全沒有變異（$\sigma_1 = \sigma_2 = 0$），兩個標準差都化為 0，理應如此。
- 在最後一部分，按公式得 $\mathrm{Var}(A) = 1 + 4 + 4 = 9$；直接計算，$A$ 取值 $1, 3, 9$ 的概率分別為 $\frac{1}{4}, \frac{1}{2}, \frac{1}{4}$，得出 $\mathrm{E}(A) = 4$、$\mathrm{E}(A^2) = 25$，方差為 9。
- 我們以分數作精確計算，在 200 個隨機的離散分佈上驗證了平均值、方差和協方差的一般公式。我們亦確認了在 $\alpha_0$ 處 $\mathrm{E}(ZA) = \mathrm{E}(Z)\mathrm{E}(A)$，確認了最後一部分中 $\alpha_0 = \frac{8}{9}$，並確認了在該部分中，對於 $\alpha = \frac{8}{9}$ 及另外 500 個 $\alpha$ 值，$Z$ 與 $A$ 均不獨立。

## 常見錯誤

- 寫成 $\mathrm{E}(X_1^2) = \mu_1^2$。正是方差項 $\sigma_1^2$ 令 $\mathrm{Var}(A)$ 和協方差不為零。
- 當 $\alpha = \alpha_0$ 時斷定 $Z$ 與 $A$ 獨立。協方差為零並不蘊涵獨立。
- 忘記說明為何 $\mu_2\sigma_1^2 + \mu_1\sigma_2^2 \neq 0$。正是因為長度為正數，各項才不可能互相抵銷。

## STEP 之外

$\mathrm{E}(UV) - \mathrm{E}(U)\mathrm{E}(V)$ 這個量就是協方差，而特殊值 $\alpha_0$ 等於 $\mathrm{Cov}(P, A)/\mathrm{Var}(A)$，即 $P$ 對 $A$ 的最小二乘迴歸（least-squares regression）直線的斜率。取 $\alpha = \alpha_0$ 會令殘差（residual）$Z$ 與 $A$ 不相關（uncorrelated），但並不令兩者獨立。對於聯合正態（jointly normal）的變量，這兩個概念是一致的，這也是正態分佈如此方便的原因之一。Grimmett 與 Welsh 所著的 *Probability: An Introduction*（概率論入門）對協方差、相關與獨立性有細緻的論述。

# 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

**2006 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。