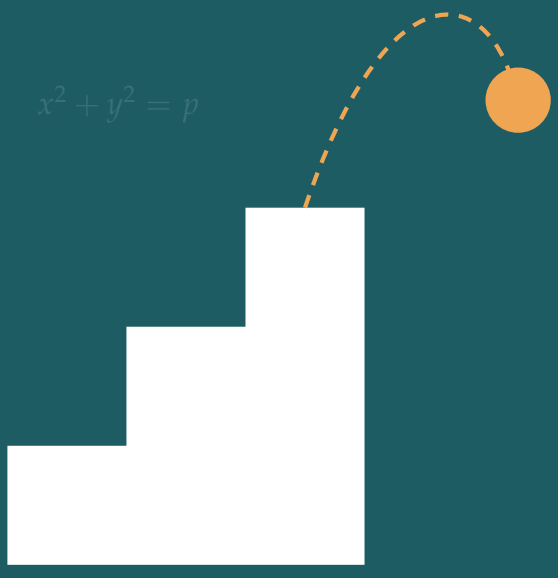
STEP 備試 · 大學數學

# Beyond STEP

## STEP I, II 及 III

2013 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

2013 年詳細解答

Keith McNulty

**Beyond STEP：以歷屆 STEP 試題為門，走進大學數學**

**2013 年詳細解答**

2013 年 STEP 試卷（STEP I, II 及 III）全部 39 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 整合平方根的方程 純數學 3

第 2 題 下取整函數方程 純數學 7

第 3 題 向量代數 純數學 12

第 4 題 代換積分法 純數學 16

第 5 題 圓錐曲線與旋轉 純數學 20

第 6 題 二項式係數與斐波那契數 純數學 25

第 7 題 一階微分方程 純數學 29

第 8 題 複合函數 純數學 33

第 9 題 拋體與碰撞 力學 37

第 10 題 恢復係數與摩擦力 力學 41

第 11 題 摩擦力與平衡 力學 45

第 12 題 不放回抽樣 概率與統計 49

第 13 題 隨機數與期望值 概率與統計 53


## STEP II


第 1 題 對數曲線的弦 純數學 56

第 2 題 歸納公式 純數學 60

第 3 題 三次方程的根 純數學 64

第 4 題 弦中點的軌跡 純數學 68

第 5 題 對稱性與曲線描繪 純數學 73

第 6 題 數列的收斂 純數學 77

第 7 題 佩爾方程的解 純數學 81

第 8 題 積分與不等式 純數學 84

第 9 題 平衡中的摩擦力 力學 88

第 10 題 拋體運動 力學 . . . . . 92
第 11 題 連續碰撞 力學 . . . . . 96
第 12 題 蓬松分佈的矩 概率與統計 . . . . . 100
第 13 題 硬幣中的游程 概率與統計 . . . . . 103

## STEP III

第 1 題 半角代換 純數學 . . . . . 107
第 2 題 麥克勞林級數 純數學 . . . . . 111
第 3 題 向量與球面 純數學 . . . . . 114
第 4 題 單位根的乘積 純數學 . . . . . 118
第 5 題 有理性與互質性 純數學 . . . . . 122
第 6 題 複數模的不等式 純數學 . . . . . 125
第 7 題 常微分方程的能量法 純數學 . . . . . 129
第 8 題 圓錐曲線的極坐標方程 純數學 . . . . . 132
第 9 題 浮力與簡諧運動 力學 . . . . . 135
第 10 題 轉動撞擊 力學 . . . . . 139
第 11 題 彈簧與能量 力學 . . . . . 143
第 12 題 指示隨機變量 概率與統計 . . . . . 147
第 13 題 矩與分佈函數 概率與統計 . . . . . 151

詳細解答

# STEP I 2013，第 1 題

純數學 • 含平方根的方程

> **STEP I 2013，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/13-s1-q1

## 關鍵思路

每條方程其實都是一條經過偽裝的二次方程（quadratic equation）。把平方根（square root）記作 $y$，再把其餘一切以 $y$ 表示，就得到一條關於 $y$ 的普通二次方程。唯一要留意的是平方根永不為負，所以 $y \geq 0$：每一次都正是這個條件把增根（spurious root）剔除。

## 解答

### 第 (i) 部分

設 $y = \sqrt{x}$，$y \geq 0$，則 $x = y^2$，方程變為

$$y^2 + 3y - \tfrac{1}{2} = 0 \quad \Longleftrightarrow \quad 2y^2 + 6y - 1 = 0.$$

由二次公式（quadratic formula），

$$y = \frac{-6 \pm \sqrt{36+8}}{4} = \frac{-3 \pm \sqrt{11}}{2}.$$

取負號會得到負值，而平方根不可能是負數。由於 $\sqrt{11} > 3$，取正號得 $y = \frac{1}{2}(\sqrt{11} - 3) > 0$，所以這是唯一的可能。於是

$$x = y^2 = \frac{11 - 6\sqrt{11} + 9}{4} = \frac{10 - 3\sqrt{11}}{2}.$$

一旦知道 $y \geq 0$，每一步都是可逆的（reversible），所以這確實是（唯一的）實根。

### 第 (ii)(a) 部分

這裏開平方的是 $x+2$，所以設 $y=\sqrt{x+2}$，$y\geq 0$，得 $x=y^2-2$。方程變為

$$y^2-2+10y-22=0 \quad\Longleftrightarrow\quad y^2+10y-24=0 \quad\Longleftrightarrow\quad (y+12)(y-2)=0.$$

由於 $y\geq 0$，我們捨棄 $y=-12$，故 $y=2$，而 $x=2^2-2=2$。

驗證：$2+10\sqrt{4}-22=0$。唯一的實根是 $x=2$。

> **想一想** 如果保留了 $y=-12$，會發生甚麼事？

這會得到 $x=144-2=142$，而 $142+10\sqrt{144}-22=240\neq 0$。$y=-12$ 是相關方程 $x-10\sqrt{x+2}-22=0$ 的根；如果你粗心地把式子平方，你所解的其實就是這條方程。條件 $y\geq 0$ 正是把兩者區分開來的關鍵。

### 第 (ii)(b) 部分

這次題目沒有提示代換（substitution），所以我們要自己找一個。根號內的式子是

$$2x^2-8x-3=2(x^2-4x)-3,$$

而 $x^2-4x$ 亦出現在根號外。這就是兩者之間的聯繫。

> **想一想** 這裏 $y$ 應該取甚麼：$x^2-4x$，還是整個平方根？

取 $y$ 為平方根，方程就會變成一條不含根號的二次方程。設 $y=\sqrt{2x^2-8x-3}$，其中 $y\geq 0$。則 $y^2=2(x^2-4x)-3$，所以

$$x^2-4x=\frac{y^2+3}{2}.$$

代入方程，得

$$\begin{aligned}\frac{y^2+3}{2}+y-9=0 \quad&\Longleftrightarrow\quad y^2+2y-15=0\\ &\Longleftrightarrow\quad (y+5)(y-3)=0.\end{aligned}$$

由於 $y\geq 0$，必須 $y=3$。於是

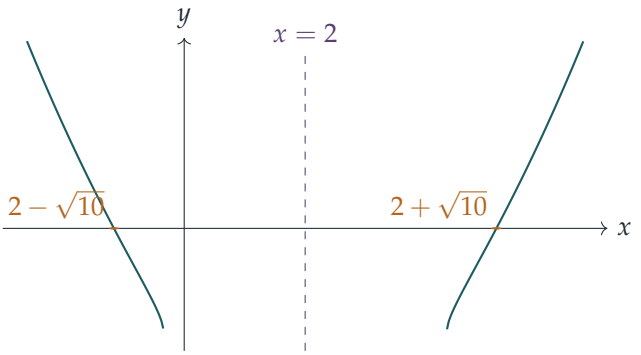
$$2x^2-8x-3=9 \quad\Longleftrightarrow\quad x^2-4x-6=0 \quad\Longleftrightarrow\quad x=2\pm\sqrt{10}.$$

這兩個值都使 $2x^2-8x-3$ 等於 9，是正數，所以平方根有定義且等於 3。因此兩者都是真正的根：實根為

$$x=2+\sqrt{10} \quad 及 \quad x=2-\sqrt{10}.$$

這兩個根關於 $x = 2$ 對稱（symmetric），這是理所當然的，因為方程只透過 $x^2 - 4x = (x-2)^2 - 4$ 依賴於 $x$。

左方的圖像顯示了這一點。它只在 $2x^2 - 8x - 3 \geq 0$ 時有定義，即 $x \leq 2 - \sqrt{5.5}$ 或 $x \geq 2 + \sqrt{5.5}$，而且在每一支上各穿過零一次。




## 驗算

- 第 (i) 部分：$x = \frac{1}{2}(10 - 3\sqrt{11}) \approx 0.02506$，所以 $\sqrt{x} \approx 0.1583$，而 $x + 3\sqrt{x} \approx 0.02506 + 0.47494 = 0.5$。
- 第 (ii)(b) 部分：$x = 2 + \sqrt{10} \approx 5.1623$ 給出 $x^2 - 4x = 6$ 及 $2x^2 - 8x - 3 = 9$，所以左方為 $6 + 3 - 9 = 0$。被捨棄的值 $y = -5$ 需要 $x^2 - 4x = 14$，此時左方為 $14 + 5 - 9 = 10 \neq 0$。
- 在每個左方式子的整個定義域（domain）上作數值掃描，所找到的根恰好是上述各根，別無其他。

## 常見錯誤

- 保留了關於 $y$ 的二次方程的負根。平方根永不為負，所以那個根不會給出解。
- 在 (ii)(b) 一開始就把兩邊平方。這會得到一條四次方程（quartic equation），難度大得多，而且會引入額外的根，之後還得逐一檢驗。
- 在 (ii)(a) 忘記了 $x = y^2 - 2$，而不是 $y^2$。

## STEP 之外

引入新變量，把一條困難的方程化為熟悉的方程，是數學中最有用的習慣之一。在 (ii)(b) 中，方程只依賴於 $x^2 - 4x$，所以它是 $(x-2)^2$ 的函數。看出這類對稱性（sym-

metry），並處理在該對稱下保持不變的量，這種做法在伽羅瓦（Galois）理論（Galois theory）中會再次出現：在那裏，人們透過找出根的、在對稱變換下保持不變的表達式來解多項式方程（polynomial equation）。Allenby 的 *Rings, Fields and Groups*（《環、域與羣》）是進入這些概念的一條平易途徑。若想做更多這類題目，Siklos 的 *Advanced Problems in Mathematics*（《高等數學難題》）非常出色（而且免費）。

詳細解答

# STEP I 2013，第 2 題

純數學 • 下取整函數方程

> **STEP I 2013，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/13-s1-q2

## 關鍵思路

函數 $\lfloor x \rfloor$ 在每個區間（interval）$[k, k+1)$ 上都是常數，所以在該區間上 $f(x) = k/x$ 只是一段直角雙曲線（rectangular hyperbola）。整個圖像就是由這些曲線段串連而成，每個整數 $k$ 對應一段。於是解 $f(x) = c$ 便歸結為：哪些曲線段能達到高度 $c$？在對應 $k$ 的那一段上，唯一的候選是 $x = k/c$，而它恰好在落入 $[k, k+1)$ 時才算數。

## 解答

### 第 (i) 部分

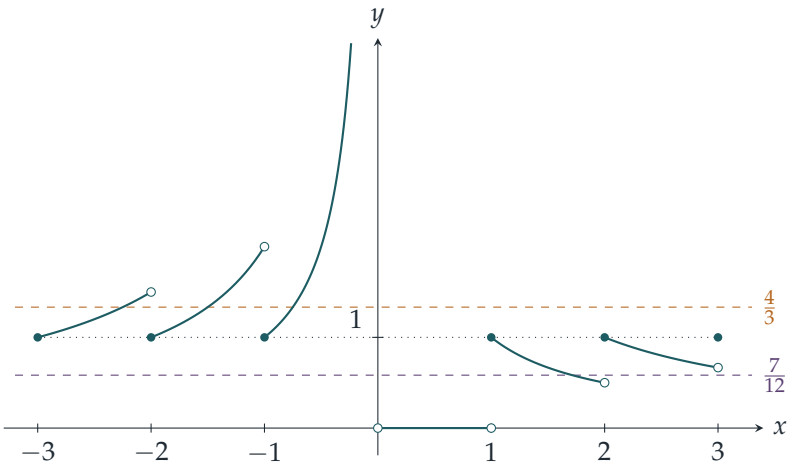
固定一個整數 $k$，考慮 $k \le x < k+1$（且 $x \neq 0$），此時 $f(x) = k/x$。

- $k = 0$，即 $0 < x < 1$：此時 $f(x) = 0$。
- $k \ge 1$：$f(x) = k/x$ 由 $f(k) = 1$ 遞減（decreasing），趨向 $k/(k+1)$，但永遠達不到。所以這一段的取值恰好是 $\left(\frac{k}{k+1}, 1\right]$。
- $k \le -1$：設 $k = -m$，其中 $m \ge 1$。此時 $x$ 和 $k$ 都是負數，所以 $f(x) = m/|x| > 0$。當 $x$ 由 $-m$ 增加至趨向 $-m+1$ 時，$|x|$ 遞減，所以 $f$ 由 $f(-m) = 1$ 起遞增。當 $m \ge 2$ 時，它趨向 $\frac{m}{m-1}$，但永遠達不到，所以這一段的取值恰好是 $\left[1, \frac{m}{m-1}\right)$。當 $m = 1$ 時，當 $x \to 0^-$ 時 $f(x) = -1/x \to +\infty$，所以這一段的取值是 $[1, \infty)$。

在 $-3 \le x \le 3$ 範圍內，各段如下：

| $x$ | $[-3,-2)$ | $[-2,-1)$ | $[-1,0)$ | $(0,1)$ | $[1,2)$ | $[2,3)$ |
|---|---|---|---|---|---|---|
| $f(x)$ | $-3/x$ | $-2/x$ | $-1/x$ | $0$ | $1/x$ | $2/x$ |
| 取值 | $[1, \frac{3}{2})$ | $[1,2)$ | $[1,\infty)$ | $0$ | $(\frac{1}{2}, 1]$ | $(\frac{2}{3}, 1]$ |

最後還有 $f(3)=1$。每一段都在其左端（整數端）從高度 1 開始，該點包含在內，而右端則是開端。




**第 (ii) 部分**

在 $\lfloor x\rfloor = k$ 的那一段上，方程 $f(x)=c$（其中 $c\neq 0$）即 $k/x=c$，所以 $x=k/c$。它恰好在 $k\le k/c<k+1$ 時才是真正的根（root）。因此對每個 $c$，我們只需利用第 (i) 部分得出的取值範圍，判斷哪些曲線段能達到高度 $c$。

$f(x)=\frac{7}{12}$。 由於 $0<\frac{7}{12}<1$，只有 $k\ge 1$ 的曲線段可能有根，而且需要 $\frac{k}{k+1}<\frac{7}{12}$，即 $12k<7k+7$，即 $k<\frac{7}{5}$。所以只有 $k=1$ 可行，得 $x=1\big/\frac{7}{12}$：

$$x=\tfrac{12}{7}.$$

從圖像上看，直線 $y=\frac{7}{12}$ 只與 $[1,2)$ 上的那一段相交，因為 $[2,3)$ 上的那一段一直高於 $\frac{2}{3}$。

$f(x)=\frac{17}{24}$。 現在需要 $\frac{k}{k+1}<\frac{17}{24}$，即 $24k<17k+17$，即 $k<\frac{17}{7}$，所以 $k=1$ 或 $k=2$。根為 $x=k\big/\frac{17}{24}$：

$$x=\tfrac{24}{17}\quad 及\quad x=\tfrac{48}{17}.$$

（確實 $2\le\frac{48}{17}\approx 2.82<3$。）

$f(x)=\frac{4}{3}$。 由於 $\frac{4}{3}>1$，只有負數那邊的曲線段可能有根。當 $\frac{4}{3}<\frac{m}{m-1}$（或 $m=1$）時，$k=-m$ 的那一段能達到 $\frac{4}{3}$，即 $4m-4<3m$，即 $m<4$。所以 $m=1,2,3$，根為 $x=-m\big/\frac{4}{3}=-\frac{3m}{4}$：

$$x=-\tfrac{3}{4},\quad x=-\tfrac{3}{2},\quad x=-\tfrac{9}{4}.$$

$m = 4$ 的情況剛好不行：$[-4, -3)$ 上的那一段只取小於 $\frac{4}{3}$ 的值。

### 第 (iii) 部分

這裏 $0 < \frac{9}{10} < 1$，所以根來自 $k \geq 1$ 且 $\frac{k}{k+1} < \frac{9}{10}$ 的曲線段，即 $10k < 9k + 9$，即 $k < 9$。所以 $k = 1, 2, \ldots, 8$ 各給出一個根 $x = \frac{10k}{9}$。最大的根對應 $k = 8$：

$$x = \frac{80}{9} \approx 8.89,$$

它確實落在 $[8, 9)$ 內。

### 最後部分

首先找出沒有根或有無限多個根的 $c$ 值。

- $c < 0$：沒有根，因為當 $x > 0$ 時 $f(x) \geq 0$，而當 $x < 0$ 時 $f(x) > 0$。
- $c = 0$：每個 $x \in (0, 1)$ 都是根，所以有無限多個根。
- $c = 1$：每個非零整數都是根，所以同樣有無限多個根。

> **想一想** 對於一般的 $c$，不畫圖的話，怎樣數根的個數？

利用第 (i) 部分得出的取值範圍：每一段最多貢獻一個根，所以只需數一數有多少段的取值範圍包含 $c$。

**情況** $0 < c < 1$。 能達到 $c$ 的曲線段是 $k \geq 1$ 且 $\frac{k}{k+1} < c$ 的那些。由於 $\frac{k}{k+1}$ 隨 $k$ 遞增，這些段就是 $k = 1, 2, \ldots$，直至最後一個滿足 $\frac{k}{k+1} < c$ 的 $k$。所以恰好有 $n$ 個根，當且僅當第 $n$ 段能達到 $c$ 而第 $(n+1)$ 段不能：

$$\frac{n}{n+1} < c \leq \frac{n+1}{n+2}.$$

（當 $0 < c \leq \frac{1}{2}$ 時，沒有任何一段能達到 $c$，因此沒有根。）

**情況** $c > 1$。 能達到 $c$ 的曲線段是 $k = -m$ 的那些，其中 $m = 1$ 或 $\frac{m}{m-1} > c$。由於 $\frac{m}{m-1} = 1 + \frac{1}{m-1}$ 隨 $m$ 增加而遞減，這些段就是 $m = 1, 2, \ldots$，直至最後一個滿足 $\frac{m}{m-1} > c$ 的 $m$。所以恰好有 $n$ 個根，當且僅當 $m = n$ 的那一段能達到 $c$ 而 $m = n + 1$ 的那一段不能：

$$\frac{n+1}{n} \leq c < \frac{n}{n-1} \qquad (n \geq 2),$$

而當 $n = 1$ 時，條件就是 $c \geq 2$（$[-1, 0)$ 上的那一段必定能達到 $c$，而當 $c \geq 2$ 時，$[-2, -1)$ 上的那一段不能）。

**答案。** 對 $n \geq 1$，方程 $f(x) = c$ 恰好有 $n$ 個根，當且僅當

$$\frac{n}{n+1} < c \leq \frac{n+1}{n+2} \qquad \text{或} \qquad \frac{n+1}{n} \leq c < \frac{n}{n-1},$$

其中當 $n = 1$ 時，第二個條件理解為 $c \geq 2$。留意等號放在哪裏。兩個等號都源自曲線段的開端：$c = \frac{n+1}{n+2}$ 不能由 $[n+1, n+2)$ 上的那一段達到，而 $c = \frac{n+1}{n}$ 不能由 $[-(n+1), -n)$ 上的那一段達到。所以這兩個邊界值仍然恰好給出 $n$ 個根。

## 驗算

- 代回原式：$\lfloor \frac{12}{7} \rfloor = 1$ 且 $1/\frac{12}{7} = \frac{7}{12}$；$\lfloor \frac{48}{17} \rfloor = 2$ 且 $2/\frac{48}{17} = \frac{17}{24}$；$\lfloor -\frac{9}{4} \rfloor = -3$ 且 $-3/(-\frac{9}{4}) = \frac{4}{3}$；$\lfloor \frac{80}{9} \rfloor = 8$ 且 $8/\frac{80}{9} = \frac{9}{10}$。
- 一般答案與個別情況吻合：$\frac{2}{3} < \frac{17}{24} \leq \frac{3}{4}$ 給出 $n = 2$；$\frac{4}{3} \leq \frac{4}{3} < \frac{3}{2}$ 給出 $n = 3$；$\frac{8}{9} < \frac{9}{10} \leq \frac{9}{10}$ 給出 $n = 8$。後兩個剛好落在範圍的端點上，正好檢驗 $\leq$ 和 $<$ 的位置是否正確。
- 一個邊界值：$c = 2$ 應該給出一個根。確實 $x = -\frac{1}{2}$ 可行，而 $x = -1$（來自 $k = -2$）有 $\lfloor -1 \rfloor = -1 \neq -2$。
- 試試 $c = \frac{1}{2}$：條件不給出任何 $n$；而確實對 $k \neq 0$，$x = k/c = 2k$ 永遠不滿足 $\lfloor 2k \rfloor = k$。

## 常見錯誤

- 忘記對負數 $x$ 而言，$\lfloor x \rfloor$ 是向下取整，所以 $\lfloor -2.25 \rfloor = -3$，而不是 $-2$。這樣會完全漏掉 $c > 1$ 時的負數分支。
- 端點搞錯：每一段包含其左端（整數端，高度為 1），但不包含右端。這決定了最終條件中出現的是 $\leq$ 還是 $<$。
- 漏掉 $c = 0$ 和 $c = 1$（有無限多個根），或漏掉 $c < 0$ 和 $0 < c \leq \frac{1}{2}$（沒有根）。

## STEP 之外

像 $\lfloor x \rfloor / x$ 這樣的函數，是「分段連續（continuous）但在孤立點上跳躍」的函數中最簡單的例子。小心分辨哪些端點能夠取到，正是大學第一個數學分析（analysis）課程處理上確界（supremum）和下確界（infimum）時所需的習慣。在這裏，$f$ 在 $x > 0$ 上的 $\sup f$ 是 1，而且能取到；而在 $x \geq 1$ 上的下確界是 $\frac{1}{2}$，卻取不到。同一個函數還說明了當 $x \to \infty$ 時 $\lfloor x \rfloor / x \to 1$，讓你初嘗解析數論（analytic number theory）中所用的估計方

法。若想進一步了解上確界、極限（limit）和連續性，Abbott 的 *Understanding Analysis*（《理解分析》）是很好的讀物。

詳細解答

# STEP I 2013，第 3 題

純數學 • 向量代數

> **STEP I 2013，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/13-s1-q3

## 關鍵思路

這裏的一切都是線性（linear）的，所以每次最乾淨利落的做法，都是把每一點寫成 $\mathbf{x}$、$\mathbf{y}$ 和 $\mathbf{z}$ 的組合，然後相減。兩點相異，當且僅當它們的位置向量（position vector）之差不為零；而這個差總會化成一個簡單的因子乘以像 $\mathbf{x}-\mathbf{y}$ 這樣的向量。從幾何上看，$X*Y$ 是直線 $XY$ 上的一點，而當 $0<\lambda<1$ 時，它位於兩點之間的線段（segment）上。

## 解答

在整個解答中，兩點相同當且僅當它們的位置向量相等。

### 第 (i) 部分

按定義，$Y*X$ 的位置向量是 $\lambda\mathbf{y}+(1-\lambda)\mathbf{x}$，所以

$$\begin{aligned}\big(\lambda\mathbf{x}+(1-\lambda)\mathbf{y}\big)-\big(\lambda\mathbf{y}+(1-\lambda)\mathbf{x}\big) &= (2\lambda-1)\mathbf{x}-(2\lambda-1)\mathbf{y}\\ &= (2\lambda-1)(\mathbf{x}-\mathbf{y}).\end{aligned}$$

由於 $X$ 與 $Y$ 相異，$\mathbf{x}-\mathbf{y}\neq\mathbf{0}$，因此這個差為零當且僅當 $2\lambda-1=0$。故此，除非 $\lambda=\frac{1}{2}$（此時兩者都是 $XY$ 的中點），否則 $X*Y$ 與 $Y*X$ 相異。

### 第 (ii) 部分

把兩邊分別展開：

$$\begin{aligned}(X*Y)*Z:&\quad \lambda\big(\lambda\mathbf{x}+(1-\lambda)\mathbf{y}\big)+(1-\lambda)\mathbf{z}=\lambda^2\mathbf{x}+\lambda(1-\lambda)\mathbf{y}+(1-\lambda)\mathbf{z},\\ X*(Y*Z):&\quad \lambda\mathbf{x}+(1-\lambda)\big(\lambda\mathbf{y}+(1-\lambda)\mathbf{z}\big)=\lambda\mathbf{x}+\lambda(1-\lambda)\mathbf{y}+(1-\lambda)^2\mathbf{z}.\end{aligned}$$

兩者的 $\mathbf{y}$ 項相同，相減得

$$\begin{aligned}\text{差} &= (\lambda^2-\lambda)\mathbf{x} + \big((1-\lambda)-(1-\lambda)^2\big)\mathbf{z}\\ &= -\lambda(1-\lambda)\mathbf{x} + \lambda(1-\lambda)\mathbf{z}\\ &= \lambda(1-\lambda)(\mathbf{z}-\mathbf{x}).\end{aligned}$$

這個差不為零，當且僅當每個因子都不為零。所以 $(X*Y)*Z$ 與 $X*(Y*Z)$ 相異，當且僅當

$$\lambda \neq 0, \qquad \lambda \neq 1 \qquad \text{且} \qquad Z \neq X.$$

留意 $Y$ 完全不起作用。被排除的值也合情合理：當 $\lambda = 1$ 時，這個運算總是給出第一點；當 $\lambda = 0$ 時，總是給出第二點，所以兩種加括號方式分別都給出 $X$ 或 $Z$。

### 第 (iii) 部分

為求簡潔，記 $\mu = 1-\lambda$。右方為

$$\begin{aligned}(X*Z)*(Y*Z): \quad &\lambda(\lambda\mathbf{x}+\mu\mathbf{z}) + \mu(\lambda\mathbf{y}+\mu\mathbf{z})\\ &= \lambda^2\mathbf{x} + \lambda\mu\,\mathbf{y} + \mu(\lambda+\mu)\mathbf{z}\\ &= \lambda^2\mathbf{x} + \lambda\mu\,\mathbf{y} + \mu\,\mathbf{z},\end{aligned}$$

因為 $\lambda+\mu = 1$。這正是第 (ii) 部分中 $(X*Y)*Z$ 的展開式，所以對所有點 $X$、$Y$、$Z$，

$$(X*Y)*Z = (X*Z)*(Y*Z)$$

得證。

> ? **想一想** $X*(Y*Z)$ 的「對應結果」應該是怎樣的？

在第 (iii) 部分中，外層的 $*Z$ 分配到左邊的括號上。其鏡像就是把 $X*$ 分配到右邊的括號上，這提示我們試試 $(X*Y)*(X*Z)$。果然，

$$\begin{aligned}(X*Y)*(X*Z): \quad &\lambda(\lambda\mathbf{x}+\mu\mathbf{y}) + \mu(\lambda\mathbf{x}+\mu\mathbf{z})\\ &= \lambda(\lambda+\mu)\mathbf{x} + \lambda\mu\,\mathbf{y} + \mu^2\mathbf{z}\\ &= \lambda\mathbf{x} + \lambda\mu\,\mathbf{y} + \mu^2\mathbf{z},\end{aligned}$$

這正是第 (ii) 部分中 $X*(Y*Z)$ 的展開式。所以對任意點 $X$、$Y$、$Z$，

$$X*(Y*Z) = (X*Y)*(X*Z).$$

### 第 (iv) 部分

我們斷言 $P_n$ 的位置向量為

$$\mathbf{p}_n = \lambda^n\mathbf{x} + (1-\lambda^n)\mathbf{y}.$$

當 $n = 1$ 時，這就是 $X * Y$ 的定義。若它對 $n-1$ 成立，則

$$\begin{aligned}\mathbf{p}_n = \lambda\mathbf{p}_{n-1} + (1-\lambda)\mathbf{y} &= \lambda^n\mathbf{x} + \big(\lambda - \lambda^n + 1 - \lambda\big)\mathbf{y} \\ &= \lambda^n\mathbf{x} + (1-\lambda^n)\mathbf{y},\end{aligned}$$

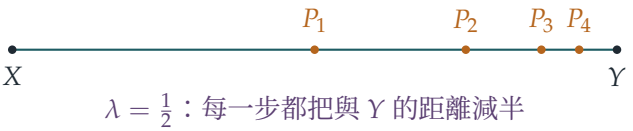
所以由歸納法（induction），此公式對所有 $n \geq 1$ 成立。

現在沿線段量度：

$$\mathbf{p}_n - \mathbf{x} = (1-\lambda^n)(\mathbf{y}-\mathbf{x}), \qquad \mathbf{y} - \mathbf{p}_n = \lambda^n(\mathbf{y}-\mathbf{x}).$$

兩者都是同一個非零向量 $\mathbf{y}-\mathbf{x}$ 的正倍數，因為 $0 < \lambda < 1$ 給出 $0 < \lambda^n < 1$。所以 $P_n$ 嚴格位於 $X$ 與 $Y$ 之間，而且

$$XP_n : P_nY = (1-\lambda^n) : \lambda^n.$$




$\lambda = \frac{1}{2}$：每一步都把與 $Y$ 的距離減半

每一步都把餘下與 $Y$ 的距離乘以 $\lambda$，所以這些點一步步向 $Y$ 推進，並且當 $n \to \infty$ 時 $P_n \to Y$。

## 驗算

- 在第 (i) 部分，沿一條直線取 $\mathbf{x} = \mathbf{0}$、$\mathbf{y} = \mathbf{1}$ 及 $\lambda = \frac{1}{3}$：$X * Y$ 位於 $\frac{2}{3}$，$Y * X$ 位於 $\frac{1}{3}$，兩者相差 $\frac{1}{3} = |2\lambda - 1|$。
- 在第 (ii) 部分，當 $Z = X$ 時，無論 $\lambda$ 取何值，兩種加括號方式都應相同；此時我們的差 $\lambda(1-\lambda)(\mathbf{z}-\mathbf{x})$ 確實為零，正如所料。
- 在第 (iv) 部分，取 $\lambda = \frac{1}{2}$ 及 $n = 3$，比為 $\frac{7}{8} : \frac{1}{8} = 7 : 1$，而 $P_3$ 確實位於由 $X$ 到 $Y$ 的 $\frac{7}{8}$ 處（距離減半三次）。取 $n = 1$ 時，比 $(1-\lambda) : \lambda$ 正是 $X * Y$ 的定義。

## 常見錯誤

- 在第 (ii) 部分忘記 $Z \neq X$：答案不只是關於 $\lambda$ 的條件。
- 在第 (iv) 部分把比的次序弄反。由於 $\mathbf{p}_n$ 在 $\mathbf{x}$ 上的權重是 $\lambda^n$，當 $\lambda^n$ 較大時它靠近 $X$，所以較短的一段 $XP_n$ 對應的是 $1-\lambda^n$。

## STEP 之外

點 $\lambda\mathbf{x} + (1-\lambda)\mathbf{y}$ 是一個凸組合（convex combination，或稱仿射組合（affine combination）），它是凸集（convex set）、重心坐標（barycentric coordinates）以及電腦圖像中線性插值（linear interpolation）的基本構件。第 (ii) 部分說明 $*$ 不滿足結合律（associative），而第 (iii) 部分則說明它滿足自分配律（self-distributive）；具有這種性質的運算在代數中以架（rack）、環結（quandle）和中介代數（medial algebra）等名稱來研究。若想從線性代數的角度理解，可參閱 Strang 的 *Introduction to Linear Algebra*（《線性代數導論》）。

詳細解答

# STEP I 2013，第 4 題

純數學 • 代換積分法

**STEP I 2013，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/13-s1-q4

## 關鍵思路

在第 (i) 部分，每個被積函數（integrand）都是「某個式子的函數，乘以該式子的導數（derivative）」：$\sec^2 x$ 是 $\tan x$ 的導數，而 $\sec x \tan x$ 是 $\sec x$ 的導數。在第 (ii) 部分，多出來的因子 $x$ 或 $x^2$ 提示我們要用分部積分法（integration by parts）。我們要積分的是三角函數那個因子，而第 (i) 部分剛好示範了該怎樣積。

## 解答

### 第 (i) 部分

對第一個積分，作代換（substitution）$u = \tan x$，於是 $\mathrm{d}u = \sec^2 x \,\mathrm{d}x$。當 $x = 0$ 時 $u = 0$，當 $x = \frac{1}{4}\pi$ 時 $u = 1$。因此

$$\int_0^{\frac{1}{4}\pi} \tan^n x \sec^2 x \,\mathrm{d}x = \int_0^1 u^n \,\mathrm{d}u = \left[\frac{u^{n+1}}{n+1}\right]_0^1 = \frac{1}{n+1},$$

這裏用了 $n > 0$，所以 $0^{n+1} = 0$。

? **想一想** 在第二個積分中，$\sec^n x \tan x$ 裏面藏着哪個函數的導數？

由於 $\dfrac{\mathrm{d}}{\mathrm{d}x} \sec x = \sec x \tan x$，我們先分出一個 $\sec x$ 因子：

$$\sec^n x \tan x = \sec^{n-1} x \,(\sec x \tan x).$$

現在代換 $u = \sec x$，於是 $\mathrm{d}u = \sec x \tan x \,\mathrm{d}x$，而 $x = 0$ 時 $u = 1$，$x = \frac{1}{4}\pi$ 時 $u = \sqrt{2}$。

由於 $n > 0$，我們可以除以 $n$，得

$$\int_0^{\frac{1}{4}\pi} \sec^n x \tan x \,\mathrm{d}x = \int_1^{\sqrt{2}} u^{n-1} \,\mathrm{d}u = \left[\frac{u^n}{n}\right]_1^{\sqrt{2}} = \frac{(\sqrt{2})^n - 1}{n},$$

即所求。

**第 (ii) 部分**

**第一個積分**。記之為 $I_1$。我們用分部積分法，對 $x$ 求導，並對 $\sec^4 x \tan x$ 積分。用第 (i) 部分的代換 $u = \sec x$，可得原函數（antiderivative）

$$\int \sec^4 x \tan x \,\mathrm{d}x = \tfrac{1}{4} \sec^4 x + C.$$

（它在 $[0, \frac{1}{4}\pi]$ 上的定積分（definite integral）是 $\frac{1}{4}(4-1) = \frac{3}{4}$，即第 (i) 部分中 $n = 4$ 的情況。）所以

$$I_1 = \left[\tfrac{1}{4} x \sec^4 x\right]_0^{\frac{1}{4}\pi} - \frac{1}{4}\int_0^{\frac{1}{4}\pi} \sec^4 x \,\mathrm{d}x = \frac{\pi}{4}\cdot\frac{4}{4} - \frac{1}{4}\int_0^{\frac{1}{4}\pi} \sec^4 x \,\mathrm{d}x,$$

因為 $\sec^4 \frac{1}{4}\pi = (\sqrt{2})^4 = 4$。對餘下的積分，寫 $\sec^2 x = 1 + \tan^2 x$：

$$\int_0^{\frac{1}{4}\pi} \sec^4 x \,\mathrm{d}x = \int_0^{\frac{1}{4}\pi} \sec^2 x \,\mathrm{d}x + \int_0^{\frac{1}{4}\pi} \tan^2 x \sec^2 x \,\mathrm{d}x = 1 + \frac{1}{3} = \frac{4}{3},$$

其中第一個積分是 $[\tan x]_0^{\frac{1}{4}\pi} = 1$，第二個是第 (i) 部分中 $n = 2$ 的情況。因此

$$\int_0^{\frac{1}{4}\pi} x \sec^4 x \tan x \,\mathrm{d}x = \frac{\pi}{4} - \frac{1}{3}.$$

**第二個積分**。記之為 $I_2$。同樣地，我們對 $x$ 的冪求導，並對 $\sec^2 x \tan x$ 積分。其中一個原函數是 $\frac{1}{2}\sec^2 x$，但 $\frac{1}{2}\tan^2 x$ 同樣可用（兩者相差常數 $\frac{1}{2}$），而且是更好的選擇。

> **想一想** 為甚麼 $\frac{1}{2}\tan^2 x$ 比 $\frac{1}{2}\sec^2 x$ 更好？

分部積分之後，我們會剩下 $x\tan^2 x$，而 $\tan^2 x = \sec^2 x - 1$ 可以拆成我們會積分的幾部分。此外，$\tan^2 x$ 在 0 處為零，令邊界項保持簡潔。（用 $\frac{1}{2}\sec^2 x$ 最終也會得到相同答案。）

所以，由 $\tan \frac{1}{4}\pi = 1$，

$$I_2 = \left[\tfrac{1}{2} x^2 \tan^2 x\right]_0^{\frac{1}{4}\pi} - \int_0^{\frac{1}{4}\pi} x \tan^2 x \,\mathrm{d}x = \frac{\pi^2}{32} - \int_0^{\frac{1}{4}\pi} x \tan^2 x \,\mathrm{d}x.$$

現在

$$\int_0^{\frac{1}{4}\pi} x \tan^2 x \,\mathrm{d}x = \int_0^{\frac{1}{4}\pi} x \sec^2 x \,\mathrm{d}x - \int_0^{\frac{1}{4}\pi} x \,\mathrm{d}x = \int_0^{\frac{1}{4}\pi} x \sec^2 x \,\mathrm{d}x - \frac{\pi^2}{32}.$$

再用一次分部積分法，以 $\tan x$ 作為 $\sec^2 x$ 的原函數，並利用 $\int \tan x \,\mathrm{d}x = \ln \sec x$（由於在此區間上 $\sec x > 0$，此式成立）：

$$\begin{aligned}\int_0^{\frac{1}{4}\pi} x \sec^2 x \,\mathrm{d}x &= \Big[x \tan x\Big]_0^{\frac{1}{4}\pi} - \int_0^{\frac{1}{4}\pi} \tan x \,\mathrm{d}x \\ &= \frac{\pi}{4} - \Big[\ln \sec x\Big]_0^{\frac{1}{4}\pi} = \frac{\pi}{4} - \ln\sqrt{2} = \frac{\pi}{4} - \frac{1}{2}\ln 2.\end{aligned}$$

把各部分合起來，

$$I_2 = \frac{\pi^2}{32} - \left(\frac{\pi}{4} - \frac{1}{2}\ln 2 - \frac{\pi^2}{32}\right),$$

即

$$\int_0^{\frac{1}{4}\pi} x^2 \sec^2 x \tan x \,\mathrm{d}x = \frac{\pi^2}{16} - \frac{\pi}{4} + \frac{1}{2}\ln 2.$$

## 驗算

- 第 (i) 部分取 $n = 1$：兩條公式分別給出 $\int_0^{\frac{1}{4}\pi} \tan x \sec^2 x \,\mathrm{d}x = \frac{1}{2}$ 和 $\int_0^{\frac{1}{4}\pi} \sec x \tan x \,\mathrm{d}x = \sqrt{2} - 1$，可以直接由 $\left[\frac{1}{2}\tan^2 x\right]$ 和 $[\sec x]$ 驗證。
- 正負與大小：第 (ii) 部分的兩個被積函數在 $(0, \frac{1}{4}\pi)$ 上都是正數，答案 $\frac{\pi}{4} - \frac{1}{3} \approx 0.4521$ 和 $\frac{\pi^2}{16} - \frac{\pi}{4} + \frac{1}{2}\ln 2 \approx 0.1780$ 亦然。第二個較小，理應如此，因為在該區間上 $x^2 \sec^2 x \le x \sec^4 x$（因為 $x < 1 < \sec^2 x$）。
- 粗略上界：$x \sec^4 x \tan x \le \frac{\pi}{4} \sec^4 x \tan x$，所以 $I_1 \le \frac{\pi}{4} \cdot \frac{3}{4} \approx 0.589$。我們得到的 0.452 與此相符。用計算機以辛普森法則（Simpson's rule）分幾個小區間計算，結果與兩個值都很接近。

## 常見錯誤

- 在第 (i) 部分的第二個積分中，未先分出 $\sec x \tan x$ 就代換 $u = \sec x$，結果冪次留在 $n$ 而非 $n - 1$。
- 忘記 $\sec \frac{1}{4}\pi = \sqrt{2}$，因此 $\sec^4 \frac{1}{4}\pi = 4$ 而 $\sec^2 \frac{1}{4}\pi = 2$。
- 漏掉邊界項（boundary term）$\left[\frac{1}{2}x^2 \tan^2 x\right]$，或把 $\int \tan x \,\mathrm{d}x$ 的正負號弄錯（它是 $\ln \sec x = -\ln \cos x$）。

## STEP 之外

第 (i) 部分的積分是積分 $\sec^m x \tan^n x$ 這套標準理論的起步，其中遞推公式（reduction formulae）讓你逐步降低冪次。像第 (ii) 部分那樣帶多項式因子的分部積分，可推廣為

「表格法」（tabular method），而在更後的階段，更發展成把導數從一個函數轉移到另一個函數的想法，這正是微分方程弱解（weak solutions）的基礎。若想仔細了解代換法和分部積分法為何成立，可參閱 Abbott 的 *Understanding Analysis*（《理解分析》）；若想做更多同類型的題目，可試試 Siklos 的 *Advanced Problems in Mathematics*（《數學進階難題》）。

詳細解答

# STEP I 2013，第 5 題

純數學 • 圓錐曲線與旋轉

**STEP I 2013，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/13-s1-q5

## 關鍵思路

唯一隨 $k$ 改變的是 $xy$ 項，而它決定了曲線的類型。若沒有 $xy$ 項，便可以對 $x$ 和 $y$ 分別配方（completing the square）。若二次部分可以因式分解（factorise），整條方程有可能分解成兩條直線。而若二次部分是完全平方（perfect square）$(x+y)^2$，把坐標軸旋轉（rotate）45° 就能把 $x+y$ 變成單一個坐標，剩下一條熟悉的拋物線（parabola）。

另外留意，設 $x=0$ 或 $y=0$ 都會令 $xy$ 項消失，所以對任何 $k$，與坐標軸的截距（intercept）都相同：$x^2+3x=0$ 給出 $x=0$ 或 $x=-3$，而 $y^2+y=0$ 給出 $y=0$ 或 $y=-1$。每一條軌跡（locus）都經過 $(0,0)$、$(-3,0)$ 和 $(0,-1)$。

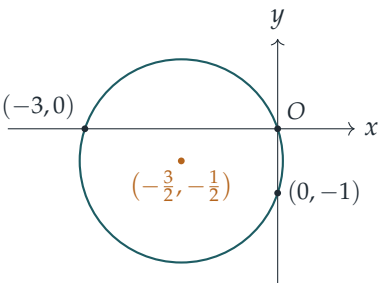
## 解答

### 第 (i) 部分

當 $k=0$ 時，對每個變量配方，得

$$\left(x+\tfrac{3}{2}\right)^2+\left(y+\tfrac{1}{2}\right)^2=\tfrac{9}{4}+\tfrac{1}{4}=\tfrac{5}{2}.$$

所以軌跡是圓心為 $\left(-\frac{3}{2},-\frac{1}{2}\right)$、半徑為 $\sqrt{5/2}\approx 1.58$ 的圓。它與 $x$ 軸相交於 $(0,0)$ 和 $(-3,0)$，與 $y$ 軸相交於 $(0,0)$ 和 $(0,-1)$。（原點和 $(-3,0)$ 是一條弦（chord）的兩個端點，所以圓心位於該弦的垂直平分線（perpendicular bisector）$x=-\frac{3}{2}$ 上，理應如此。）




**第 (ii) 部分**

當 $k=\frac{10}{3}$ 時，兩邊乘以 3：

$$3x^2+3y^2+10xy+9x+3y=0.$$

二次部分可因式分解為

$$3x^2+10xy+3y^2=(3x+y)(x+3y),$$

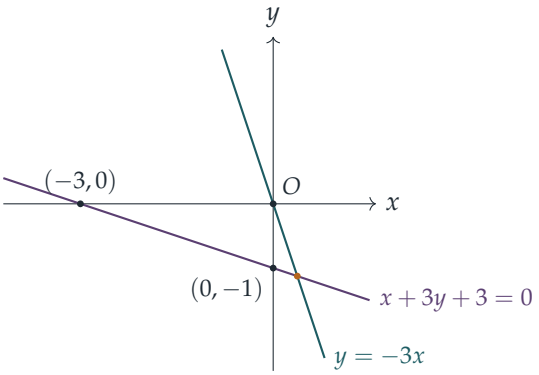
而一次部分是 $9x+3y=3(3x+y)$，同樣含有因式 $3x+y$。所以

$$(3x+y)(x+3y)+3(3x+y)=(3x+y)(x+3y+3)=0.$$

乘積為零當且僅當其中一個因式為零，所以軌跡是一對直線

$$y=-3x \qquad \text{和} \qquad x+3y+3=0.$$

第一條直線經過原點（這是它與兩條坐標軸唯一的交點）。第二條直線與 $x$ 軸相交於 $(-3,0)$，與 $y$ 軸相交於 $(0,-1)$。聯立兩條方程，得 $x-9x+3=0$，所以兩線相交於 $\left(\frac{3}{8},-\frac{9}{8}\right)$。



**第 (iii) 部分**

當 $k = 2$ 時，方程為

$$(x+y)^2 + 3x + y = 0.$$

當 $\theta = 45°$ 時，$\cos\theta = \sin\theta = \frac{1}{\sqrt{2}}$，所以

$$X = \frac{x+y}{\sqrt{2}}, \qquad Y = \frac{y-x}{\sqrt{2}}.$$

因此 $x + y = \sqrt{2}\,X$ 及 $y - x = \sqrt{2}\,Y$，把兩式相加和相減，得

$$x = \frac{X-Y}{\sqrt{2}}, \qquad y = \frac{X+Y}{\sqrt{2}}.$$

? **想一想** 為甚麼我們需要以 $X$ 和 $Y$ 表示 $x$ 和 $y$，而不是反過來？

我們知道 $P$ 在曲線上，而要求的是 $Q$ 所滿足的方程。所以必須把 $x$ 和 $y$ 代入 $P$ 的軌跡方程，這就需要以 $X$ 和 $Y$ 來表示它們。（由於這個映射是旋轉，其逆映射（inverse）就只是反方向的旋轉，所以這一步很容易。）

一次部分變成

$$3x + y = \frac{3X - 3Y + X + Y}{\sqrt{2}} = \frac{4X - 2Y}{\sqrt{2}} = 2\sqrt{2}\,X - \sqrt{2}\,Y,$$

而 $(x+y)^2 = 2X^2$。所以 $Q$ 的軌跡是

$$2X^2 + 2\sqrt{2}\,X - \sqrt{2}\,Y = 0 \quad\Longleftrightarrow\quad \sqrt{2}\,Y = 2X^2 + 2\sqrt{2}\,X = \left(\sqrt{2}\,X + 1\right)^2 - 1,$$

正如所求。每一步都是可逆的，所以這正正就是 $Q$ 的軌跡。

這是 $(X, Y)$ 平面上的一條拋物線，開口朝 $Y$ 增加的方向。它關於 $\sqrt{2}\,X + 1 = 0$ 即 $X = -\frac{1}{\sqrt{2}}$ 這條直線對稱，頂點（vertex）位於 $X = -\frac{1}{\sqrt{2}}$，$Y = -\frac{1}{\sqrt{2}}$。

現在把旋轉還原。$Q$ 是由 $P$ 順時針旋轉 45° 得到的，所以 $P$ 的軌跡就是把這條拋物線逆時針旋轉 45°。以 $x$ 和 $y$ 表示：

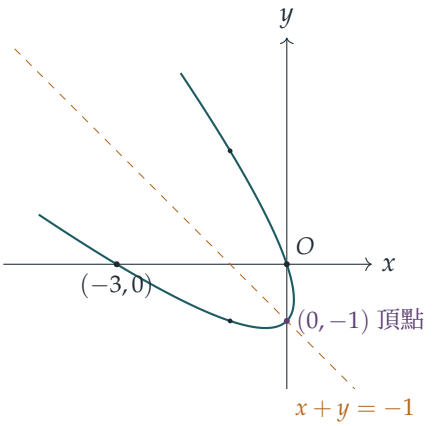
- 對稱軸（line of symmetry）$X = -\frac{1}{\sqrt{2}}$ 即 $\frac{x+y}{\sqrt{2}} = -\frac{1}{\sqrt{2}}$，也就是

$$x + y = -1;$$

- 頂點位於 $x = \frac{X-Y}{\sqrt{2}} = 0$，$y = \frac{X+Y}{\sqrt{2}} = -1$，即 $(0, -1)$；
- 拋物線開口朝 $Y = \frac{y-x}{\sqrt{2}}$ 增加的方向，即 $(-1, 1)$ 的方向：向左上方。

截距是 $(0,0)$、$(-3,0)$ 和 $(0,-1)$，正如開首所述；其中 $(0,-1)$ 是頂點。該處的切線（tangent）垂直於對稱軸，所以斜率（gradient）為 1：曲線在頂點處穿過 $y$ 軸，而不是與之相切。作為對稱性的驗證，關於 $x+y=-1$ 的反射（reflection）把 $(a,b)$ 映到 $(-1-b,-1-a)$，所以原點映到 $(-1,-1)$，而 $(-3,0)$ 映到 $(-1,2)$；兩點都滿足 $(x+y)^2+3x+y=0$。

繪圖時，設 $u=x+y$ 會有幫助：這樣 $3x+y=-u^2$，解得 $x=-\frac{1}{2}(u^2+u)$，$y=\frac{1}{2}(u^2+3u)$。




$P$ 的軌跡是一條頂點為 $(0,-1)$、對稱軸為 $x+y=-1$ 的拋物線。

## 驗算

- 把截距點 $(0,0)$、$(-3,0)$、$(0,-1)$ 代入全部三條方程：無論 $k$ 為何值，結果都是 0。
- 在 (ii) 中，展開 $(3x+y)(x+3y+3)$ 再除以 3，便得回 $k=\frac{10}{3}$ 時的原方程。交點 $\left(\frac{3}{8},-\frac{9}{8}\right)$ 同時滿足 $3x+y=0$ 和 $x+3y+3=0$。
- 在 (iii) 中，取 $u=1$：點 $(-1,2)$ 在 $P$ 的軌跡上。於是 $X=\frac{1}{\sqrt{2}}$，$Y=\frac{3}{\sqrt{2}}$，而 $\sqrt{2}\,Y=3=(1+1)^2-1$。

## 常見錯誤

- 在 (iii) 中，把 $X,Y$ 的公式代入方程，而不是以 $X,Y$ 表示 $x,y$。
- 把拋物線向錯誤方向旋轉回去，得出對稱軸為 $y-x=-1$，而非 $x+y=-1$。
- 在 (ii) 中，沒有發現一次項含有公因式 $3x+y$，因而把軌跡當作其他類型的圓錐曲線。

## STEP 之外

每條方程 $ax^2 + bxy + cy^2 + dx + ey + f = 0$ 都是一條圓錐曲線（conic），其類型由二次部分的判別式（discriminant）$b^2 - 4ac$ 決定：負數時是橢圓（ellipse）（這裏 $k = 0$ 時是圓），零時是拋物線（$k = 2$），正數時是雙曲線（hyperbola），而雙曲線可以退化（degenerate）成一對直線（$k = \frac{10}{3}$）。在大學一年級的線性代數（linear algebra）課程中，你會學到如何有系統地消去 $xy$ 項：把坐標軸旋轉到二次型（quadratic form）所對應的對稱矩陣（symmetric matrix）的特徵向量（eigenvectors）方向。斯特朗（Strang）的 *Introduction to Linear Algebra*（《線性代數導論》）對此有清晰的解釋。

詳細解答

# STEP I 2013，第 6 題

純數學 • 二項式係數與斐波那契數

> **STEP I 2013，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/13-s1-q6

## 關鍵思路

帕斯卡法則（Pascal's rule）就是整道題的引擎。每個 $B_n$ 都是若干二項式係數（binomial coefficients）之和；若用帕斯卡法則把 $B_{n+2}$ 的每一項拆開，拆出來的各部分正好可以重新組合成 $B_{n+1}$ 和 $B_n$ 的各項。唯一要小心的是和式的兩端：那裏帕斯卡法則並不適用，落單的 1 要靠人手配對。由於 $B_n$ 對偶數和奇數 $n$ 的定義不同，我們把兩種情況分開處理。

## 解答

### 帕斯卡法則

由於 $(1+x)^{n+1} = (1+x)(1+x)^n$，兩邊 $x^r$ 的係數相等。左邊的係數是 $\binom{n+1}{r}$。右邊則是

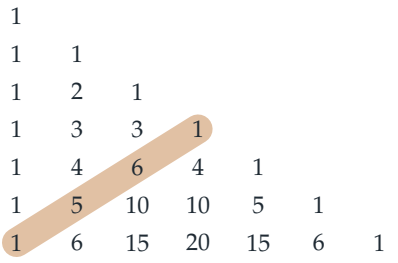
$$(1+x)\sum_{s=0}^{n}\binom{n}{s}x^s = \sum_{s=0}^{n}\binom{n}{s}x^s + \sum_{s=0}^{n}\binom{n}{s}x^{s+1},$$

當 $1 \le r \le n$ 時，第一個和式貢獻 $\binom{n}{r}$（來自 $s = r$），第二個和式貢獻 $\binom{n}{r-1}$（來自 $s = r-1$）。因此

$$\binom{n}{r} + \binom{n}{r-1} = \binom{n+1}{r} \qquad (1 \le r \le n).$$

### 遞推關係

先看清楚 $B_n$ 是甚麼會有幫助：它把帕斯卡三角形（Pascal's triangle）的一條「淺對角線」加起來，即 $\binom{n}{0} + \binom{n-1}{1} + \binom{n-2}{2} + \cdots$，一直加到上方數字不小於下方數字的最後一項為止。把三角形靠左排列時，組成 $B_6 = 1+5+6+1 = 13$ 的各項落在一條直線上：




（琥珀色帶標出了 $\binom{6}{0}, \binom{5}{1}, \binom{4}{2}, \binom{3}{3}$。）

**偶數情況，**$n = 2m$（$m \geq 0$）。我們需要證明 $B_{2m+2} = B_{2m+1} + B_{2m}$。按定義

$$B_{2m+2} = \sum_{j=0}^{m+1} \binom{2m+2-j}{j}.$$

兩端的項是 $j = 0$，給出 $\binom{2m+2}{0} = 1$；以及 $j = m+1$，給出 $\binom{m+1}{m+1} = 1$。對於 $1 \leq j \leq m$，我們以上方 $2m+1-j$、下方 $j$ 套用帕斯卡法則；這是容許的，因為 $j \leq m$ 推出 $j \leq 2m+1-j$。於是

$$B_{2m+2} = \underbrace{1 + \sum_{j=1}^{m} \binom{2m+1-j}{j}}_{\text{(a)}} + \underbrace{\sum_{j=1}^{m} \binom{2m+1-j}{j-1} + 1}_{\text{(b)}}.$$

在 (a) 中，那個 1 就是 $\binom{2m+1}{0}$，即缺少了的 $k = 0$ 項，所以 (a) 等於 $\sum_{k=0}^{m} \binom{2m+1-k}{k} = B_{2m+1}$。

**想一想** 第二個和式應該變成甚麼？多出來的那個 1 又該放到哪裏？

在 (b) 中令 $i = j - 1$：和式變成 $\sum_{i=0}^{m-1} \binom{2m-i}{i}$，這正是缺了最後一項 $i = m$ 的 $B_{2m}$。那最後一項是 $\binom{m}{m} = 1$，恰好就是多出來的 1。所以 (b) $= B_{2m}$，從而 $B_{2m+2} = B_{2m+1} + B_{2m}$。（當 $m = 0$ 時，和式都是空和，式子變成 $2 = 1 + 1$。）

**奇數情況，**$n = 2m+1$（$m \geq 0$）。我們需要證明 $B_{2m+3} = B_{2m+2} + B_{2m+1}$。此時

$$B_{2m+3} = \sum_{k=0}^{m+1} \binom{2m+3-k}{k}.$$

只有 $k = 0$ 項 $1 = \binom{2m+2}{0}$ 保持不動。對於 $1 \leq k \leq m+1$，以上方 $2m+2-k$、下方 $k$ 套

用帕斯卡法則是容許的，因為 $k \le m+1$ 推出 $k \le 2m+2-k$。因此

$$\begin{aligned} B_{2m+3} &= \binom{2m+2}{0} + \sum_{k=1}^{m+1} \binom{2m+2-k}{k} + \sum_{k=1}^{m+1} \binom{2m+2-k}{k-1} \\ &= \sum_{j=0}^{m+1} \binom{2m+2-j}{j} + \sum_{i=0}^{m} \binom{2m+1-i}{i} \qquad (i = k-1) \\ &= B_{2m+2} + B_{2m+1}. \end{aligned}$$

這一次沒有剩下任何端項。

每個 $n \ge 0$ 都可寫成 $2m$ 或 $2m+1$ 的形式，所以對所有 $n = 0, 1, 2, \ldots$ 都有 $B_{n+2} - B_{n+1} = B_n$，此即所求。

### 與斐波那契數的關係

直接計算可得 $B_0 = \binom{0}{0} = 1$ 及 $B_1 = \binom{1}{0} = 1$。斐波那契數（Fibonacci numbers）由 $F_1 = 1$、$F_2 = 1$ 開始。所以 $B_0 = F_1$、$B_1 = F_2$，而兩個數列都服從同一條遞推關係（recurrence）：「每一項等於前兩項之和」。由數學歸納法（induction）（若 $B_n = F_{n+1}$ 且 $B_{n+1} = F_{n+2}$，則 $B_{n+2} = F_{n+2} + F_{n+1} = F_{n+3}$），

$$B_n = F_{n+1} \qquad \text{對所有} n \ge 0.$$

數列 $B$ 就是把斐波那契數列移後一位：$1, 1, 2, 3, 5, 8, 13, \ldots$。

## 驗算

- 直接由定義計算幾項：$B_2 = 1+1 = 2$，$B_3 = 1+2 = 3$，$B_4 = 1+3+1 = 5$，$B_5 = 1+4+3 = 8$。這些就是 $F_3, \ldots, F_6$。
- 較大的一項：$B_9 = \binom{9}{0} + \binom{8}{1} + \binom{7}{2} + \binom{6}{3} + \binom{5}{4} = 1+8+21+20+5 = 55 = F_{10}$。
- 抽查帕斯卡法則：$\binom{5}{2} + \binom{5}{1} = 10+5 = 15 = \binom{6}{2}$。

## 常見錯誤

- 對 $\binom{m+1}{m+1}$ 這類端項套用帕斯卡法則，這樣會得出 $\binom{m}{m+1}$，超出了法則已獲證明的範圍 $1 \le r \le n$。應先把端項分離出來。
- 只處理一種奇偶性（parity）。$B_{2m}$ 和 $B_{2m+1}$ 的定義不同，而遞推關係把偶數下標的項與奇數下標的項聯繫起來，所以兩種情況都必須處理。
- 弄錯下標的位移：答案是 $B_n = F_{n+1}$，而不是 $B_n = F_n$（檢查一下 $B_0 = 1 \ne F_0$）。

## STEP 之外

恆等式 $F_{n+1} = \sum_k \binom{n-k}{k}$ 有一個漂亮的計數證明：兩邊都在計算用正方形和骨牌（dominoes）鋪滿一條 $1 \times n$ 長條的方法數（$k$ 是骨牌的數目，而在 $n-k$ 塊磚之中選出骨牌的位置，就得到 $\binom{n-k}{k}$）。以兩種方法計算同一個集合來證明恆等式（counting in two ways），是大學組合學（combinatorics）的核心技巧；生成函數（generating function）的觀點亦然，在這觀點下 $\sum F_n x^n = x/(1-x-x^2)$。如想深入了解，可以讀 Wilf 的 *generatingfunctionology*（《生成函數學》），或 Siklos 的 *Advanced Problems in Mathematics*（《高等數學問題》）。

詳細解答

# STEP I 2013，第 7 題

純數學 • 一階微分方程

> **STEP I 2013，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/13-s1-q7

## 關鍵思路

在 (i) 和 (ii) 中，右方只取決於比值 $y/x$，所以設 $y = ux$ 便可把方程化為可分離變數 (separation of variables) 的形式。第 (iii) 部分打破了這個模式，因為 $x^2/y$ 並非單單是 $y/x$ 的函數。解決方法是問：解「應該」表現得像 $x$ 的哪一個冪？平衡兩項後得出 $y = u\,x^{3/2}$，於是一切又可以分離變數。

## 解答

### 第 (i) 部分

設 $y = ux$，由乘積法則 (product rule) 得 $\dfrac{\mathrm{d}y}{\mathrm{d}x} = u + x\dfrac{\mathrm{d}u}{\mathrm{d}x}$。由於 $x/y = 1/u$ 及 $y/x = u$，方程變為

$$u + x\frac{\mathrm{d}u}{\mathrm{d}x} = \frac{1}{u} + u \quad\Longrightarrow\quad x\frac{\mathrm{d}u}{\mathrm{d}x} = \frac{1}{u}.$$

$u$ 項互相抵消，這正是這個代換 (substitution) 奏效的原因。分離變數（由於 $x > 0$，可以除以 $x$），

$$\int u\,\mathrm{d}u = \int \frac{1}{x}\,\mathrm{d}x \quad\Longrightarrow\quad \tfrac{1}{2}u^2 = \ln x + C,$$

因為 $x > 0$，所以寫 $\ln x$ 而非 $\ln|x|$。當 $x = 1$ 時 $y = 2$，所以 $u = y/x = 2$，得 $2 = 0 + C$。因此

$$u^2 = 4 + 2\ln x.$$

由於 $y > 0$ 及 $x > 0$，必有 $u > 0$，所以取正根 $u = \sqrt{4 + 2\ln x}$，從而

$$y = x\sqrt{4 + 2\ln x}.$$

這要求 $4+2\ln x>0$（我們需要 $y>0$，而不只是 $y\geq 0$），即 $\ln x>-2$，亦即 $x>\mathrm{e}^{-2}$。這就是所求的結果。

### 第 (ii) 部分

右方同樣是 $y/x$ 的函數，所以用同一個代換 $y=ux$：

$$u+x\frac{\mathrm{d}u}{\mathrm{d}x}=\frac{1}{u}+2u \quad\Longrightarrow\quad x\frac{\mathrm{d}u}{\mathrm{d}x}=\frac{1}{u}+u=\frac{1+u^2}{u}.$$

這次 $u$ 項沒有抵消，但變數仍然可以分離：

$$\int\frac{u}{1+u^2}\,\mathrm{d}u=\int\frac{1}{x}\,\mathrm{d}x \quad\Longrightarrow\quad \tfrac{1}{2}\ln(1+u^2)=\ln x+C,$$

其中左方的積分屬 $\frac{1}{2}\int f'/f$ 的形式，而 $1+u^2>0$，無需加絕對值（modulus）。在 $x=1$ 時 $u=2$，所以 $C=\frac{1}{2}\ln 5$。乘以 2 再取指數，

$$1+u^2=5x^2, \qquad \text{所以} \qquad u=\sqrt{5x^2-1}$$

（因 $u>0$，取正根）。因此

$$y=x\sqrt{5x^2-1} \qquad \left(x>\tfrac{1}{\sqrt{5}}\right),$$

這個限制來自 $5x^2-1>0$ 及 $x>0$。

### 第 (iii) 部分

現在 $y=ux$ 給出 $x^2/y=x/u$，當中仍單獨含有 $x$，所以方程不能分離變數。我們想找一個代換，使右方兩項以相同方式縮放。

> **想一想** 若 $y$ 表現得像冪 $x^k$，$k$ 取甚麼值時 $\dfrac{x^2}{y}$ 與 $\dfrac{2y}{x}$ 的大小相同？

我們需要 $x^{2-k}$ 與 $x^{k-1}$ 相配，所以 $2-k=k-1$，得 $k=\frac{3}{2}$。（此時左方的 $\mathrm{d}y/\mathrm{d}x$ 亦表現得像 $x^{1/2}$，所以三項一致。）這啟示我們用代換 $y=u\,x^{3/2}$，此時

$$\frac{\mathrm{d}y}{\mathrm{d}x}=x^{3/2}\frac{\mathrm{d}u}{\mathrm{d}x}+\tfrac{3}{2}u\,x^{1/2}, \qquad \frac{x^2}{y}=\frac{x^{1/2}}{u}, \qquad \frac{2y}{x}=2u\,x^{1/2}.$$

把整條方程除以 $x^{1/2}$（它是正數），

$$x\frac{\mathrm{d}u}{\mathrm{d}x}+\tfrac{3}{2}u=\frac{1}{u}+2u \quad\Longrightarrow\quad x\frac{\mathrm{d}u}{\mathrm{d}x}=\frac{1}{u}+\frac{u}{2}=\frac{2+u^2}{2u}.$$

分離變數，

$$\int\frac{2u}{2+u^2}\,\mathrm{d}u=\int\frac{1}{x}\,\mathrm{d}x \quad\Longrightarrow\quad \ln(2+u^2)=\ln x+C.$$

在 $x=1$ 時 $u=y/x^{3/2}=2$，所以 $C=\ln 6$，因此 $2+u^2=6x$。取正根 $u=\sqrt{6x-2}$，所以

$$y=x^{3/2}\sqrt{6x-2}=\sqrt{6x^4-2x^3}\qquad \left(x>\tfrac{1}{3}\right),$$

其中需要 $x>\frac{1}{3}$ 才能使 $6x-2>0$。

**(iii) 的另一解法**。把方程乘以 $2y$，並設 $z=y^2$，得

$$\frac{\mathrm{d}z}{\mathrm{d}x}-\frac{4}{x}z=2x^2,$$

這是一條線性方程（linear equation），積分因子（integrating factor）為 $x^{-4}$。於是 $\left(x^{-4}z\right)'=2x^{-2}$，利用 $z(1)=4$ 得 $x^{-4}z=-2/x+6$，所以 $y^2=6x^4-2x^3$，與前相同。同一技巧（$z=y^2$）亦可解 (i) 和 (ii)。

## 驗算

- 每個答案在 $x=1$ 時都給出 $y=2$：$1\cdot\sqrt{4}$、$1\cdot\sqrt{4}$ 及 $\sqrt{6-2}$。
- 把 (ii) 的答案求導後代回原方程。由 $y^2=5x^4-x^2$ 得 $2yy'=20x^3-2x$，所以 $y'=(10x^3-x)/y$。另一方面 $\dfrac{x}{y}+\dfrac{2y}{x}=\dfrac{x^2+2y^2}{xy}=\dfrac{10x^4-x^2}{xy}$，兩者相同。
- 在 $x=2$ 時，三個解分別給出 $2\sqrt{4+2\ln 2}\approx 4.642$、$2\sqrt{19}\approx 8.718$ 及 $4\sqrt{5}\approx 8.944$。正如預期，(ii) 中 $y/x$ 的係數較大，使 $y$ 增長得比 (i) 快。

## 常見錯誤

- 在 $\dfrac{\mathrm{d}y}{\mathrm{d}x}=u+x\dfrac{\mathrm{d}u}{\mathrm{d}x}$ 中漏掉 $u$，以致關於 $u$ 的方程出錯。
- 把初始條件（initial condition）套用在 $y$ 而非 $u$ 上：這裏在 $x=1$ 時兩者相同，但在 (iii) 中一般需要用 $u=y/x^{3/2}$。
- 遺漏定義域（domain）。平方根必須是實數且為正，因而得出 $x>\mathrm{e}^{-2}$、$x>1/\sqrt{5}$ 及 $x>\frac{1}{3}$。

## STEP 之外

右方只取決於 $y/x$ 的方程稱為齊次（homogeneous）方程，而代換 $y=ux$ 總能把它們分離變數。第 (iii) 部分其實是一種經過偽裝的縮放對稱性（scaling symmetry）：在 $x\mapsto\lambda x$、$y\mapsto\lambda^{3/2}y$ 之下方程不變，而這類縮放對稱性總會啟示一個使方程可分離變數的代換。形如 $y'+P(x)y=Q(x)y^n$ 的方程，在 $z=y^{1-n}$ 之下會變成線性方程，稱

為伯努利方程（Bernoulli equations）；這裏三部分都屬於此類，且 $n = -1$。Boyce 與 DiPrima 的 *Elementary Differential Equations*（《初等微分方程》）在一階方程一章中涵蓋了這兩類方程。

詳細解答

# STEP I 2013，第 8 題

純數學 • 複合函數

> **STEP I 2013，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/13-s1-q8

## 關鍵思路

合成函數（composite function）pq 的意思是「先做 q，再做 p」，而且只有在 q 的輸出落在 p 的定義域（domain）之內時才有定義。因此，定義域來自 q（有需要時再加以收窄），而值域（range）則要問 p 實際上接收到哪些值才能求得。代數運算並不繁重；得分關鍵在於小心處理 $\sqrt{x^2} = |x|$ 以及各個定義域。在第 (iii) 部分，恆等式 $(x + \sqrt{x^2-1})(x - \sqrt{x^2-1}) = 1$ 幫了大部分的忙。

## 解答

### 第 (i) 部分

- $\mathrm{cb}(x) = \mathrm{c}(\ln x) = 2\ln x$。定義域就是 b 的定義域，即 $x > 0$，而 c 接受任何實數。由於 $\ln x$ 取遍所有實數值，$2\ln x$ 亦然：值域是 $\mathbb{R}$。
- $\mathrm{ab}(x) = (\ln x)^2$，定義域為 $x > 0$。由於 $\ln x$ 取遍所有實數值，其平方便取遍 $[0, \infty)$ 中的每個值（在 $x = 1$ 時取 0）：值域是 $y \geq 0$。
- $\mathrm{da}(x) = \sqrt{x^2} = |x|$，而不是 $x$，因為平方根號表示非負根（non-negative root）。這裏 $x^2 \geq 0$ 恆成立，所以定義域是整個 $\mathbb{R}$，值域是 $y \geq 0$。
- $\mathrm{ad}(x) = (\sqrt{x})^2 = x$，但只限於 $x \geq 0$，因為必須先由 d 作用。所以定義域是 $x \geq 0$，值域是 $y \geq 0$。

留意 da 和 ad 的值域相同，但定義域和公式都不同：它們是不同的函數。

**第 (ii) 部分**

對於 fg：對每個實數 $x$，$\mathrm{g}(x) = \sqrt{x^2+1} \geq 1$，所以 $\mathrm{g}(x)$ 總落在 f 的定義域內。於是

$$\mathrm{fg}(x) = \sqrt{(x^2+1)-1} = \sqrt{x^2} = |x|, \qquad \text{定義域}\mathbb{R}, \quad \text{值域}y \geq 0.$$

對於 gf：f 先作用，所以需要 $|x| \geq 1$；而 g 接受任何輸入。於是

$$\mathrm{gf}(x) = \sqrt{(x^2-1)+1} = |x|, \qquad \text{定義域}|x| \geq 1, \quad \text{值域}y \geq 1.$$

兩條公式同樣一致，但函數卻不同：gf 的定義域不包括 $-1 < x < 1$，因此其值域由 1 開始。

**第 (iii) 部分**

基本事實是

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

所以當 $x \geq 1$ 時，$\mathrm{h}(x)\,\mathrm{k}(x) = 1$。

**h 的圖像（$x \geq 1$）。**

- $\mathrm{h}(1) = 1$。
- 當 $x > 1$ 時，$\mathrm{h}'(x) = 1 + \dfrac{x}{\sqrt{x^2-1}} > 0$，所以 h 是遞增（increasing）的；而當 $x \to 1^+$ 時 $\mathrm{h}'(x) \to \infty$，所以曲線以垂直切線（vertical tangent）離開 $(1,1)$。
- 當 $x \to \infty$ 時，$\mathrm{h}(x) \to \infty$。至於漸近線（asymptote），

  $$\mathrm{h}(x) - 2x = \sqrt{x^2-1} - x = \frac{-1}{\sqrt{x^2-1}+x} \to 0^-,$$

  所以 $y = 2x$ 是一條漸近線，曲線從下方趨近它。

因此 h 由 1 開始連續遞增，且無上界：其值域是 $[1,\infty)$。

**k 的圖像，$x \geq 1$ 的分支。**

- $\mathrm{k}(1) = 1$，而且由基本事實，$\mathrm{k}(x) = 1/\mathrm{h}(x)$。
- 由於 h 是正數並遞增至無限大，k 是正數並遞減，且 $\mathrm{k}(x) \to 0^+$。所以 $y = 0$ 是一條漸近線，曲線從上方趨近它。
- 當 $x \to 1^+$ 時，$\mathrm{k}'(x) = 1 - \dfrac{x}{\sqrt{x^2-1}} \to -\infty$：在 $(1,1)$ 處同樣有一條垂直切線，這分支就在該點與 h 的圖像相接。

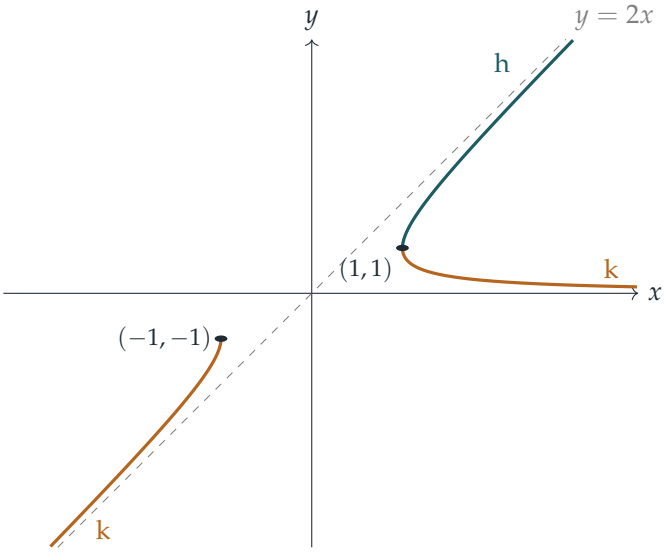
**k 的圖像，$x \le -1$ 的分支。**

- $\mathrm{k}(-1) = -1$。
- 當 $x < -1$ 時有 $x < 0$，所以 $\mathrm{k}'(x) = 1 - \dfrac{x}{\sqrt{x^2-1}} > 1 > 0$：k 是遞增的。當 $x \to -1^-$ 時，$\mathrm{k}'(x) \to +\infty$，所以在 $(-1,-1)$ 處有一條垂直切線。
- $x$ 和 $-\sqrt{x^2-1}$ 都是負數，所以當 $x \to -\infty$ 時 $\mathrm{k}(x) \to -\infty$。設 $x = -u$，其中 $u \ge 1$，

$$\mathrm{k}(x) - 2x = u - \sqrt{u^2-1} = \frac{1}{u + \sqrt{u^2-1}} \to 0^+,$$

所以 $y = 2x$ 亦是這分支的漸近線，曲線從上方趨近它。

這分支取遍 $(-\infty, -1]$ 中的每個值。




（有一個巧妙的方法可以一次看清這一切：設 $x = \cosh t$，其中 $t \ge 0$。那麼 $\sqrt{x^2-1} = \sinh t$，所以 $\mathrm{h}(x) = e^t$ 而 $\mathrm{k}(x) = e^{-t}$；對於 $x = -\cosh t$，則 $\mathrm{k}(x) = -e^t$。）

**合成函數 kh。** h 的定義域是 $x \ge 1$，而每個值 $\mathrm{h}(x)$ 都滿足 $\mathrm{h}(x) \ge 1$，落在 k 的定義域內。所以 kh 在整個 $x \ge 1$ 上都有定義。

**想一想** kh 實際上用到 k 圖像的哪一部分？

只用到 $x \ge 1$ 的分支，因為 h 從不輸出小於 1 的值。此外，h 取遍 $[1,\infty)$ 中的每一個值，所以 kh 所取的值正好就是 k 在 $[1,\infty)$ 上的值。在這分支上，k 由 $\mathrm{k}(1) = 1$ 開始連續遞減趨向 0，但永遠達不到 0（因為 $\mathrm{k} = 1/\mathrm{h} > 0$）。因此

$$\mathrm{kh}: \quad \text{定義域} x \ge 1, \qquad \text{值域} 0 < y \le 1.$$

從雙曲函數（hyperbolic function）的角度看，$\mathrm{kh}(\cosh t) = \mathrm{k}(e^t)$，這是 $t$ 的遞減函數，在 $t = 0$ 時由 1 開始。

## 驗算

- 代入數值：$\mathrm{h}(2) = 2 + \sqrt{3} \approx 3.732$，$\mathrm{k}(2) = 2 - \sqrt{3} \approx 0.268$；兩者之積為 1。另外 $\mathrm{k}(-2) = -2 - \sqrt{3} \approx -3.732$，略高於 $2x = -4$，正如漸近線分析所預測。
- $\mathrm{kh}(1) = \mathrm{k}(1) = 1$，而 $\mathrm{kh}(2) = \mathrm{k}(3.732\ldots) = 3.732 - \sqrt{3.732^2 - 1} \approx 0.136$，落在 $(0, 1]$ 之內。
- 當 $x = 1000$ 時，$\mathrm{h}(x) - 2x \approx -0.0005$，$\mathrm{k}(x) \approx 0.0005$，與 $-1/(2x)$ 和 $1/(2x)$ 吻合。

## 常見錯誤

- 寫成 $\sqrt{x^2} = x$。其實應是 $|x|$，這正是 da、fg 和 gf 都不是恆等函數（identity function）的原因。
- 因為最終公式 $|x|$ 處處有意義，便把 gf 的定義域寫成 $\mathbb{R}$。合成函數的定義域由最先作用的函數決定。
- 遺漏了 k 在 $x \leq -1$ 的第二個分支，或遺漏了它的漸近線 $y = 2x$。

## STEP 之外

在大學裏，函數是一個三元組：定義域、上域（codomain）和對應法則。兩個法則相同但定義域不同的函數，例如這裏的 fg 和 gf，確實是不同的函數；一旦討論反函數（inverse function）、單射性（injectivity）和滿射性（surjectivity），這種嚴謹就變得非常重要。函數 h 就是 $e^{\operatorname{arcosh} x}$，而 $\ln \mathrm{h}(x)$ 正是 $\operatorname{arcosh} x$ 的標準公式。Abbott 的 *Understanding Analysis*（《理解分析》）正是以這種細緻的態度處理函數、連續性（continuity）和單調性（monotonicity）。

詳細解答

# STEP I 2013，第 9 題

力學 • 拋體與碰撞

**STEP I 2013，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：

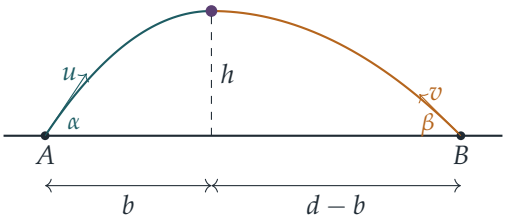
step.maths.org/questions/13-s1-q9

## 關鍵思路

拋射體（projectile）在飛行的最高點沒有鉛直速度，所以若在該處發生碰撞，之後落回地面所需的時間就等於上升所用的時間。要落回出發點，每個質點必須在同樣的時間內走回原來的水平路程，這就迫使它的水平速度剛好反向。把這一點代入動量守恆（conservation of momentum），再利用兩個質點在同一時刻到達最高點這個事實，其餘結果便順理成章。

## 解答

我們把 $A$ 和 $B$ 視為只受重力作用而自由運動的質點，不計空氣阻力（air resistance），而碰撞時的衝量（impulse）沿連心線（line of centres）作用。由於這是直接碰撞（direct collision），而碰撞一刻兩個質點都在水平運動，所以這個衝量是水平的。取 $x$ 軸沿平面、由 $A$ 的出發點指向 $B$ 的出發點，於是 $B$ 從 $x = d$ 出發。




### 在最高點相遇

$A$ 的鉛直速度是 $u\sin\alpha - gt$，所以它在時間 $t = u\sin\alpha/g$ 到達最大高度。同理，$B$ 在 $t = v\sin\beta/g$ 時最高。兩者同時拋出，並在各自處於最高點時相撞，因此這兩個時間相

等：

$$T = \frac{u\sin\alpha}{g} = \frac{v\sin\beta}{g}, \qquad \text{所以} \qquad u\sin\alpha = v\sin\beta. \tag{1}$$

碰撞前一刻，$A$ 的速度是沿正 $x$ 方向的 $u\cos\alpha$，$B$ 的速度是沿負 $x$ 方向的 $v\cos\beta$；兩者都沒有鉛直速度。

### 碰撞之後

**想一想** 我們對恢復係數（coefficient of restitution）一無所知，又怎能對碰撞後的速度說出甚麼來？

關鍵在於「返回原點」這個條件。衝量是水平的，所以碰撞後一刻 $A$ 的鉛直速度仍然是零，位置在平面上方高度 $h$ 處。（在鉛直方向上）由靜止下落 $h$ 所需的時間，恰好等於上升 $h$ 所用的時間，即 $T$。在這段時間內，$A$ 必須在水平方向由 $x = b$ 回到 $x = 0$，其中 $b = u\cos\alpha\, T$ 是它上升途中走過的水平距離。所以它新的水平速度是

$$-\frac{b}{T} = -u\cos\alpha:$$

$A$ 的水平速度剛好反向，同理 $B$ 的也是如此。（事實上，每個質點都沿自己原來的路徑倒退回去。）

現在，$x$ 方向的動量守恆給出

$$mu\cos\alpha - Mv\cos\beta = -mu\cos\alpha + Mv\cos\beta, \qquad \text{所以} \qquad mu\cos\alpha = Mv\cos\beta. \tag{2}$$

換言之，碰撞前後的總動量都是零。以 (1) 除 (2)——這是容許的，因為 $\alpha$ 和 $\beta$ 都嚴格介於 0 與 $\frac{1}{2}\pi$ 之間，所以 (1) 的兩邊都不為零——得

$$m\cot\alpha = M\cot\beta,$$

此即所求。

### 碰撞發生的位置

在時間 $T$ 內，$A$ 走過水平距離 $b = u\cos\alpha\, T$，$B$ 則走過 $v\cos\beta\, T$。兩者合起來剛好走完兩者之間的距離，所以

$$u\cos\alpha\, T + v\cos\beta\, T = d.$$

由 (2)，$v\cos\beta = \dfrac{m}{M}u\cos\alpha$，所以

$$b\left(1 + \frac{m}{M}\right) = d \qquad \Longrightarrow \qquad b = \frac{Md}{m+M},$$

此即所求。（$B$ 越重，它的水平速度就越慢，相遇點便離 $A$ 的出發點越遠。）

### 碰撞的高度

對 $A$ 在鉛直方向使用 $v^2 = u^2 + 2as$，得最大高度為

$$h = \frac{u^2 \sin^2 \alpha}{2g}, \qquad \text{而} \qquad b = u\cos\alpha \cdot \frac{u \sin\alpha}{g} = \frac{u^2 \sin\alpha\cos\alpha}{g}.$$

兩式相除，

$$\frac{h}{b} = \frac{\sin\alpha}{2\cos\alpha} = \tfrac{1}{2}\tan\alpha, \qquad \text{所以} \qquad h = \tfrac{1}{2} b \tan\alpha.$$

## 驗算

- 對稱性：當 $m = M$ 時，得 $\alpha = \beta$ 及 $b = \frac{1}{2}d$，所以兩個質點在正中間相遇，理應如此。
- 高度公式表明，由 $A$ 的出發點到其路徑最高點的連線，斜率是 $\frac{1}{2}\tan\alpha$，即拋射時斜率的一半：這正是一個熟悉的事實——拋物線（parabola）在起點的切線與對稱軸（axis of symmetry）相交之處，高度是頂點高度的兩倍。對 $B$ 作同樣的計算，得 $h = \frac{1}{2}(d - b)\tan\beta$；這與 $\frac{1}{2}b\tan\alpha$ 吻合，原因正是 $m\cot\alpha = M\cot\beta$。
- 數值驗證：取 $m = 1$、$M = 2$、$\alpha = 45°$、$u = 10\,\mathrm{m\,s^{-1}}$、$g = 9.8\,\mathrm{m\,s^{-2}}$。於是 $\cot\beta = \frac{1}{2}$，$v = \sqrt{50}/\sin\beta \approx 7.91\,\mathrm{m\,s^{-1}}$，$d \approx 7.65\,\mathrm{m}$，從而 $b = \frac{2}{3}d \approx 5.10\,\mathrm{m}$，$h = \frac{1}{2}b \approx 2.55\,\mathrm{m}$。用電腦把兩個運動逐步向前推算，並在碰撞時把水平速度反向，結果證實兩個質點確實在此點、各自處於最高點時相遇，而且都落回各自的出發點。

## 常見錯誤

- 試圖使用恢復係數：題目沒有給出，也根本不需要，因為返回原點的條件已經確定了碰撞後的速度。
- 忘記兩個質點必須同時到達最高點，而正是這一點給出 (1)；沒有它，就無法得出餘切（cotangent）。
- 動量方程中的正負號錯誤：碰撞前 $B$ 是沿負方向運動的。

## STEP 之外

返回原點的條件其實就是時間反演（time reversal）：在重力作用下，牛頓（Newton）的運動方程即使把時鐘倒轉，形式也不變，所以把質點的速度反向，它就會沿自己原來的路徑返回。這裏總動量為零，意味着我們身處質心參考系（centre-of-mass frame）之中；在這個參考系裏，彈性碰撞（elastic collision）只會把兩個速度反向，與這裏發生

的情況完全一樣（所以這次碰撞是完全彈性的）。質心參考系和守恆定律在 Kleppner 與 Kolenkow 合著的 *An Introduction to Mechanics*（《力學導論》）中有精彩的探討。

詳細解答

# STEP I 2013，第 10 題

力學 • 恢復係數與摩擦力

**STEP I 2013，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/13-s1-q10

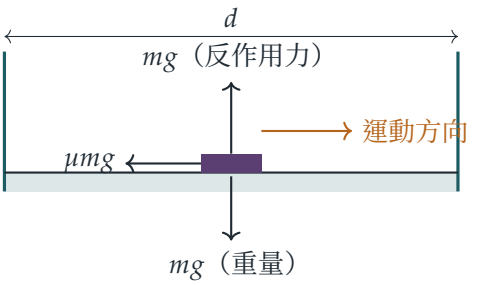
## 關鍵思路

在兩次碰撞之間，唯一的水平力是摩擦力（friction），所以冰球的減速度（deceleration）恆為 $\mu g$，而等加速度公式 $v^2 = u^2 + 2as$ 把冰球離開一道擋板時的速率與到達下一道擋板時的速率聯繫起來；然後牛頓（Newton）碰撞定律（law of restitution）把到達時的速率乘以 $r$。結果得到 $v_i^2$ 的一條線性遞推關係（linear recurrence）。要解線性遞推關係 $w_{i+1} = aw_i + b$，只需平移到它的不動點（fixed point），之後它就只是一個等比數列（geometric sequence）。

## 解答

### 遞推關係

把冰球視為一個質點，沿垂直於擋板的直線運動。在兩次碰撞之間冰球在冰面上滑行時，法向反作用力（normal reaction）為 $mg$（沒有鉛垂方向的加速度）；由於冰球正在滑動，摩擦力取其極限值（limiting value）$\mu mg$，方向與運動方向相反。



由牛頓第二定律，減速度為常數 $\mu g$。因此，對於一趟長度為 $d$、起始速率為 $v_i$（即第 $i$ 次碰撞剛結束時的速率）的橫越（traverse），冰球到達另一道擋板時的速率 $u$ 滿足

$$u^2 = v_i^2 - 2\mu g d.$$

擋板是固定的，所以牛頓碰撞定律指出冰球離開時的速率是其接近速率的 $r$ 倍：$v_{i+1} = ru$。兩邊平方，

$$v_{i+1}^2 = r^2u^2 = r^2v_i^2 - 2r^2\mu g d, \qquad \text{所以} \qquad v_{i+1}^2 - r^2v_i^2 = -2r^2\mu g d,$$

這正是所求的結果。若把起始速率記作 $v_0 = v$，同樣的論證也適用於第一趟橫越，所以這條遞推關係對所有 $i \geq 0$ 都成立。

### 橫越的次數

兩邊除以 $2\mu g d$，並令 $w_i = \dfrac{v_i^2}{2\mu g d}$，於是 $w_0 = k$ 及

$$w_{i+1} = r^2(w_i - 1).$$

「經過 $n$ 趟橫越後在擋板旁靜止」的意思是：冰球在第 $n$ 趟橫越結束時以零速率到達擋板。這時它離開的速率也是零，所以 $v_n = 0$，即 $w_n = 0$（而對於 $i < n$，$w_i > 0$）。

**想一想** 怎樣解 $w_{i+1} = r^2(w_i - 1)$？

先找常數解 $w_i = c$：需要 $c = r^2(c-1)$，所以 $c = -\dfrac{r^2}{1-r^2}$；由於 $r \neq 1$，這是有意義的。把遞推關係減去 $c = r^2(c-1)$，得

$$w_{i+1} - c = r^2(w_i - c),$$

所以 $w_i - c$ 是公比為 $r^2$ 的等比數列：

$$w_i = c + r^{2i}(k - c).$$

令 $w_n = 0$，得 $r^{2n}(k-c) = -c$，即

$$r^{2n}\left(k + \frac{r^2}{1-r^2}\right) = \frac{r^2}{1-r^2} \quad \Longrightarrow \quad r^{2n-2}\left(k(1-r^2) + r^2\right) = 1.$$

兩邊乘以 $r^{-2n}$，整理成

$$r^{-2n} = 1 + k\left(r^{-2} - 1\right), \qquad \text{所以} \qquad \boxed{n = \frac{\ln\left(1 + k(r^{-2} - 1)\right)}{2\ln(1/r)}}$$

（當 $0 < r < 1$ 時，$\ln(1/r) > 0$；又因為 $k > 0$ 且 $r^{-2} > 1$，對數的自變量大於 1，所以 $n > 0$，理應如此。）

### $r = \mathrm{e}^{-1}$ 的情況

此時 $r^{-2} = \mathrm{e}^2$ 而 $\ln(1/r) = 1$，所以公式變成

$$n = \tfrac{1}{2}\ln\left(1 + k(\mathrm{e}^2 - 1)\right),$$

這正是所求的結果。

### $r = 1$ 的情況

遞推關係變成 $w_{i+1} = w_i - 1$：由於在擋板處沒有損失，每趟橫越恰好消耗 $v^2$ 中的 $2\mu g d$。所以 $w_i = k - i$，而由 $w_n = 0$ 得

$$n = k = \frac{v^2}{2\mu g d}.$$

這其實就是能量守恆的考慮：初始動能 $\frac{1}{2}mv^2$ 在總距離 $nd$ 上全部被摩擦力耗散，所以 $\frac{1}{2}mv^2 = \mu mg\, nd$。

## 驗算

- 取 $r = \mathrm{e}^{-1}$，從 $w_3 = 0$ 把遞推關係倒過來推：$w_2 = 1$，$w_1 = \mathrm{e}^2 + 1$，$w_0 = \mathrm{e}^4 + \mathrm{e}^2 + 1 \approx 62.99$。公式給出 $\frac{1}{2}\ln\left(1 + (\mathrm{e}^4 + \mathrm{e}^2 + 1)(\mathrm{e}^2 - 1)\right) = \frac{1}{2}\ln \mathrm{e}^6 = 3$，理應如此。
- 當 $r \to 1$ 時，$r^{-2} - 1 \approx 2(1 - r)$ 且 $\ln(1/r) \approx 1 - r$，所以一般公式趨向 $\dfrac{2k(1-r)}{2(1-r)} = k$，與 $r = 1$ 時的答案吻合。
- $k$ 是無量綱的（$v^2$ 和 $gd$ 都是速率的平方），而 $n$ 也必須是無量綱的。
- 對隨機選取的 $\mu$、$d$、$r$ 和 $v$ 以數值方法逐步模擬運動，結果以很高的精確度證實了這條遞推關係。

## 常見錯誤

- 對 $v_i^2$ 應用碰撞定律時用了因子 $r$ 而不是 $r^2$：碰撞定律講的是速率，所以速率的平方要乘以 $r^2$，摩擦力那一項也一樣。
- 忘記 $n$ 數的是橫越的趟數，而不是反彈速率非零的碰撞次數：條件是 $v_n = 0$，而 $v_0 = v$。

## STEP 之外

遞推關係 $w_{i+1} = aw_i + b$ 是最簡單的線性差分方程（linear difference equation），而減去不動點這個技巧，正是你將來在線性微分方程中會遇到的「餘函數加特解」（complementary function plus particular solution）方法。這類彈跳問題（經典例子是皮球在地面上彈跳）也是能量耗散的好例子；在沒有摩擦力的情況下，它們還展示了在有限時間內發生無限多次事件的現象。Kleppner 和 Kolenkow 的 *An Introduction to Mechanics*（《力學導論》）對碰撞和摩擦力有更深入的論述。

詳細解答

# STEP I 2013，第 11 題

力學 • 摩擦力與平衡

**STEP I 2013，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/13-s1-q11

## 關鍵思路

兩條繩在相對的兩側分別與水平成 $\frac{1}{2}\pi - \alpha$ 和 $\alpha$ 的角，因此兩繩互相垂直（perpendicular）。兩個互相垂直的力 $T\cos\beta$ 和 $T\sin\beta$ 合成一個大小為 $T$ 的單一力，方向由第一個力轉過 $\beta$。把兩條繩換成一個與水平成 $\alpha+\beta$ 角、大小為 $T$ 的拉力之後，兩部分都變成簡短的練習：第 (i) 部分是極限摩擦（limiting friction），第 (ii) 部分則要察覺物塊會被整個提離表面。

## 解答

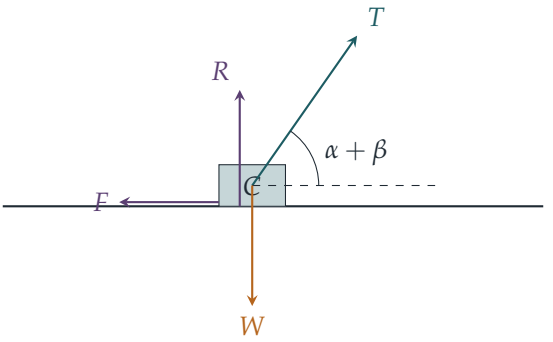
我們把 $C$ 視為質點（particle），並把繩視為輕繩。取 $x$ 軸為水平並指向 $B$，$y$ 軸為豎直向上。

### 張力的合力

繩 $BC$ 指向正 $x$ 軸上方 $\alpha$ 角的方向，所以其單位向量（unit vector）為 $(\cos\alpha, \sin\alpha)$。繩 $AC$ 與負 $x$ 軸成 $\frac{1}{2}\pi - \alpha$ 角，所以它與正 $x$ 軸成 $\pi - (\frac{1}{2}\pi - \alpha) = \frac{1}{2}\pi + \alpha$ 角，單位向量為 $(-\sin\alpha, \cos\alpha)$。兩者互相垂直。因此兩繩的總拉力為

$$\begin{aligned} T\cos\beta\,(\cos\alpha, \sin\alpha) + T\sin\beta\,(-\sin\alpha, \cos\alpha) &= T\big(\cos\alpha\cos\beta - \sin\alpha\sin\beta, \\ &\qquad \sin\alpha\cos\beta + \cos\alpha\sin\beta\big) \\ &= T\big(\cos(\alpha+\beta),\ \sin(\alpha+\beta)\big). \end{aligned}$$

所以兩繩的作用恰好等同一個大小為 $T$、在水平之上成 $\alpha+\beta$ 角、拉向 $B$ 一側的單一力。由於 $0 < \alpha + \beta < \frac{1}{2}\pi$，兩個分量都是正數。




**第 (i) 部分**

設 $R$ 為法向反作用力（normal reaction），$F$ 為摩擦力；若物塊靜止，摩擦力沿水平方向背離 $B$。平衡時，

$$\begin{aligned} &\text{豎直方向：} \quad R + T\sin(\alpha+\beta) = W, \qquad \text{所以} R = W - T\sin(\alpha+\beta),\\ &\text{水平方向：} \quad F = T\cos(\alpha+\beta). \end{aligned}$$

條件 $W > T\sin(\alpha+\beta)$ 使 $R > 0$，所以物塊確實被壓在表面上，可以產生摩擦。只要所需的摩擦力不超過其極限值，即 $F \le \mu R$，物塊便保持靜止：

$$T\cos(\alpha+\beta) \le \mu\big(W - T\sin(\alpha+\beta)\big).$$

現在代入 $\mu = \tan\lambda = \sin\lambda/\cos\lambda$，並在兩邊乘以 $\cos\lambda$；由於 $\lambda$ 是銳角，這個數是正數：

$$T\cos(\alpha+\beta)\cos\lambda \le W\sin\lambda - T\sin(\alpha+\beta)\sin\lambda.$$

移項並利用 $\cos(P-Q) = \cos P\cos Q + \sin P\sin Q$，得

$$W\sin\lambda \ge T\big(\cos(\alpha+\beta)\cos\lambda + \sin(\alpha+\beta)\sin\lambda\big) = T\cos(\alpha+\beta-\lambda),$$

這正是所求的條件。（每一步都可逆，所以事實上物塊保持靜止當且僅當此式成立。）

**想一想** 為甚麼我們需要條件 $W > T\sin(\alpha+\beta)$ 呢？

若沒有這個條件，豎直方向的方程會給出 $R \le 0$。表面只能推、不能拉，所以 $R$ 為負是不可能的：物塊會被提離表面（或僅僅剛好接觸），也就沒有摩擦力把它固定。這正是第 (ii) 部分的情況。

### 第 (ii) 部分

現在 $W = T\tan\phi$ 而 $\alpha+\beta = 2\phi$，所以 $0 < \phi < \frac{1}{4}\pi$。比較重量與兩繩向上的拉力：

$$\begin{aligned} T\sin(\alpha+\beta) - W &= T\big(2\sin\phi\cos\phi - \tan\phi\big) \\ &= T\tan\phi\left(2\cos^2\phi - 1\right) \\ &= T\tan\phi\cos 2\phi > 0, \end{aligned}$$

因為 $0 < \phi < \frac{1}{4}\pi$ 使 $\tan\phi$ 和 $\cos 2\phi$ 都是正數。所以兩繩向上的拉力大於重力向下的拉力。若物塊留在表面上，我們便需要 $R = W - T\sin(\alpha+\beta) < 0$，這是不可能的。因此物塊立即離開表面：$R = 0$，而既然沒有接觸，也就沒有摩擦力。

此時僅有的力是與水平成 $2\phi$ 角的拉力 $T$ 和重量。合力（resultant force）為

$$\big(T\cos 2\phi,\ T\sin 2\phi - T\tan\phi\big) = \big(T\cos 2\phi,\ T\tan\phi\cos 2\phi\big) = \frac{T\cos 2\phi}{\cos\phi}\,(\cos\phi,\ \sin\phi),$$

其中豎直分量用了上面的計算。其大小 $T\cos 2\phi/\cos\phi$ 是正數，所以不為零。物塊由靜止開始，所以其初始加速度，從而它開始運動的方向，都沿着合力的方向。該方向為 $(\cos\phi, \sin\phi)$，即在水平之上成 $\phi$ 角（指向 $B$），正如所求。

## 驗算

- 取 $T = 10\,\mathrm{N}$、$\alpha = 20°$、$\beta = 40°$，於是 $\phi = 30°$，$W = 10\tan 30° \approx 5.77\,\mathrm{N}$。向上的拉力為 $10\sin 60° \approx 8.66\,\mathrm{N} > W$，合力為 $(10\cos 60°,\ 8.66 - 5.77) = (5,\ 2.89)$，其角度為 $\arctan(2.89/5) = 30° = \phi$。
- 在第 (i) 部分中，$\lambda = \frac{1}{2}\pi$（無限粗糙）給出 $W \geq T\sin(\alpha+\beta)$，這本來已經成立，所以物塊永不滑動；$\lambda \to 0$（光滑）給出 $0 \geq T\cos(\alpha+\beta)$，這不成立，所以光滑表面無法使物塊靜止。兩者都符合我們的預期。
- 第 (i) 部分的條件亦已用電腦，就 $\alpha, \beta, T, W, \mu$ 的大量隨機取值，與直接檢驗 $F \leq \mu R$ 的結果互相核對。

## 常見錯誤

- 把每條繩分別分解，結果在代數中迷失，而沒有察覺兩繩互相垂直、應先把它們合成。
- 在第 (ii) 部分寫出 $R = W - T\sin(\alpha+\beta)$ 後繼續處理摩擦力，卻沒有察覺 $R$ 會是負數。
- 忘記說明由靜止開始的質點最初沿合力方向運動。

## STEP 之外

$\lambda$ 即摩擦角（angle of friction），其意思是法向反作用力與極限摩擦力合成一個與法線成 $\lambda$ 角的總反作用力；平衡可能出現，恰恰在總反作用力能落在這個「摩擦錐」（cone of friction）之內的時候。第 (i) 部分正是這個命題的變裝版本，而同一幅圖像亦能處理三維中的摩擦以及剛體問題。Kleppner 和 Kolenkow 的 *An Introduction to Mechanics*（《力學導論》）對摩擦、約束和接觸力有細緻的論述。

詳細解答

# STEP I 2013，第 12 題

概率與統計 • 不放回抽樣

**STEP I 2013，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/13-s1-q12

## 關鍵思路

「隨機」選取 $k$ 粒藥片，意思是容器中剩下的藥片裏，每一組 $k$ 粒藥片被選中的機會均等，所以每天的概率（probability）都是兩個計數之比：有利的選法數目除以 $\binom{\text{剩下的藥片}}{k}$。至於最後一天的問題，竅門在於對稱性（symmetry）：留到最後一天的藥片，與第一天選出的藥片同樣隨機。在 (iii) 中，各天之間環環相扣，十分順暢，因為在「好」的一天之後，容器的狀態恰好就像一個較短的假期。

## 解答

在整個解答中，把藥片視為可區分的（例如編上號碼），儘管它們看起來一模一樣。這樣做不會改變任何概率，而且令從 $N$ 粒中選 $k$ 粒的全部 $\binom{N}{k}$ 種選法機會均等。

### 第 (i) 部分

當 $k = 3$ 及 $n = 3$ 時，共有 9 粒藥片，每種 3 粒。從中選 3 粒的方法數為 $\binom{9}{3} = 84$。要每種各得一粒，就要從 3 粒第 1 種藥片中選一粒、從 3 粒第 2 種中選一粒、從 3 粒第 3 種中選一粒，所以共有 $3^3 = 27$ 種方法。因此

$$\mathrm{P}(\text{第 1 天每種各一粒}) = \frac{27}{84} = \frac{9}{28},$$

正如所求。

? **想一想** 為甚麼求最後一天的概率時，我不需要逐一分析第 1 天和第 2 天發生了甚麼？

在整個假期中，九粒藥片被隨機分成三組，每組三粒：第 1 天的一組、第 2 天的一組和

第 3 天的一組（即剩下的）。無論按 1、2、3 的次序還是按 3、2、1 的次序選出各組，所得的分佈（distribution）都相同：把藥片分成三個有標記的組，每一種分法的機會都均等。特別地，留給第 3 天的那組藥片是一組均勻隨機（uniformly random）的三粒藥片，與第 1 天選出的那組完全一樣。所以最後一天剩下每種各一粒藥片的概率同樣是

$$\frac{9}{28}.$$

**第 (ii) 部分**

現在共有 $3n$ 粒藥片，每種 $n$ 粒。共有 $\binom{3n}{3}$ 種機會均等的選法，其中 $n^3$ 種包含每種各一粒。所以

$$\begin{aligned}\mathrm{P}(\text{第 1 天每種各一粒}) &= \frac{n^3}{\binom{3n}{3}} = \frac{6n^3}{3n(3n-1)(3n-2)} \\ &= \frac{2n^2}{(3n-1)(3n-2)}.\end{aligned}$$

當 $n=3$ 時，此值為 $\frac{18}{8\cdot 7} = \frac{9}{28}$，與 (i) 吻合。

**第 (iii) 部分**

現在 $k=2$，共有 $2n$ 粒藥片，兩種藥片各 $n$ 粒。對 $m = n, n-1, \ldots, 1$，設 $G_m$ 為以下事件：在某天開始時容器中每種藥片各有 $m$ 粒，而我在當天選到每種各一粒（所以第 1 天時 $m=n$）。

若容器中每種各有 $m$ 粒，則選兩粒藥片共有 $\binom{2m}{2} = m(2m-1)$ 種選法，其中 $m\cdot m$ 種是每種各一粒。所以

$$\mathrm{P}(\text{每種各一粒} \mid m \text{ 為每種剩餘粒數}) = \frac{m^2}{m(2m-1)} = \frac{m}{2m-1}.$$

若到目前為止每一天都是「好」的，那麼每天都恰好取走了每種各一粒，所以下一天開始時每種各有 $m-1$ 粒。把每天的條件概率（conditional probability）逐一相乘，

$$\mathrm{P}(\text{每天都是每種各一粒}) = \prod_{m=1}^{n} \frac{m}{2m-1} = \frac{n!}{1\times 3\times 5\times\cdots\times(2n-1)}.$$

為了整理奇數之積，把分子和分母同乘以偶數之積 $2\times 4\times\cdots\times(2n) = 2^n\, n!$：

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

因此

$$\boxed{\mathrm{P}(\text{每天都是每種各一粒}) = \frac{2^n\,(n!)^2}{(2n)!}}$$

（驗算一下：$n=1$ 時得 1，這是正確的：每種只有一粒時，我必定兩粒都取。）

現在對 $n!$ 及 $(2n)!$ 使用斯特林（Stirling）近似（Stirling's approximation）：

$$(n!)^2 \approx 2n\pi\left(\frac{n}{\mathrm{e}}\right)^{2n},$$
$$(2n)! \approx \sqrt{4n\pi}\left(\frac{2n}{\mathrm{e}}\right)^{2n} = 2\sqrt{n\pi}\,2^{2n}\left(\frac{n}{\mathrm{e}}\right)^{2n}.$$

因子 $(n/\mathrm{e})^{2n}$ 互相抵消，剩下

$$\frac{2^n\,(n!)^2}{(2n)!} \approx \frac{2^n\cdot 2n\pi}{2\sqrt{n\pi}\,2^{2n}} = 2^{-n}\,\frac{n\pi}{\sqrt{n\pi}} = 2^{-n}\sqrt{n\pi},$$

這正是所求的近似式。

## 驗算

- 用電腦窮舉所有抽取藥片的方式，證實了 (i) 中第 1 天和第 3 天的概率都是 $\frac{9}{28}$，(ii) 的公式在 $n\le 7$ 時成立，以及 (iii) 的公式在 $n\le 4$ 時成立。
- 當 $n=10$ 時，(iii) 的準確概率為 $\frac{2^{10}(10!)^2}{20!}\approx 0.005542$，而 $2^{-10}\sqrt{10\pi}\approx 0.005474$：偏低約 1.3%。隨着 $n$ 增大，兩者之比趨向 1（$n=50$ 時約為 1.0025），這與斯特林近似的預期一致。
- (iii) 的另一種計數方法：把 $2n$ 粒藥片分成 $n$ 對並排成有序列，共有 $(2n)!/2^n$ 種機會均等的分法；好的分法是每天把第 1 種藥片與第 2 種藥片配對，共有 $n!\times n!$ 種。兩者之比同樣是 $2^n(n!)^2/(2n)!$。

## 常見錯誤

- 每次抽一粒藥片，寫出 $\frac{3}{9}\times\frac{3}{8}\times\frac{3}{7}$，但這是按此次序抽到第 1、2、3 種的機會。三種藥片出現的次序共有 3! 種，所以應得 $6\times\frac{27}{504}=\frac{9}{28}$。（又或者：第一粒可以是任何一種，所以是 $1\times\frac{6}{8}\times\frac{3}{7}=\frac{9}{28}$。）
- 求最後一天的概率時，嘗試對第 1 天發生的所有情況作條件分析；這方法可行，但十分冗長，而對稱性能立即給出答案。
- 對 $(2n)!$ 用斯特林近似時，要把所有 $n$ 都換成 $2n$，包括平方根之內：$\sqrt{2(2n)\pi}=2\sqrt{n\pi}$，而不是 $\sqrt{2n\pi}$。

## STEP 之外

(iii) 的答案是 $2^n/\binom{2n}{n}$，而估計式 $\binom{2n}{n} \approx 4^n/\sqrt{n\pi}$ 在概率論中隨處可見：它給出簡單隨機遊走（simple random walk）在 $2n$ 步後回到起點的機會，並且是證明一維或二維的這類遊走以概率 1 回到原點的關鍵。(i) 中的對稱性論證，就是可交換性（exchangeability）的概念：隨機抽取的次序不會改變它們的聯合分佈（joint distribution）。這兩個概念在 Grimmett 與 Welsh 的 *Probability: An Introduction*（《概率導論》）以及 Blitzstein 與 Hwang 的 *Introduction to Probability*（《概率導論》）中都有深入發展。

詳細解答

# STEP I 2013，第 13 題

概率與統計 • 計數與期望值

> **STEP I 2013，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/13-s1-q13

## 關鍵思路

數字 $1,\ldots,52$ 可分成 26 個互不相交的「伙伴對」$\{k, 53-k\}$，而被棄掉的一對，其實就是我們把兩個成員都選中了的伙伴對。因此，要數出恰好有 $k$ 個完整對的選法，我們先選定哪些對是完整的，再用來自不同對的數字填滿其餘位置，每次可自由選取該對中的哪一個成員。至於期望值（expectation），對這 26 個對運用線性性質（linearity）便可迅速驗證結果。

## 解答

### 數算選法

把這些整數寫成 26 個對 $\{1,52\},\{2,51\},\ldots,\{26,27\}$。由 1 至 52 中的兩個整數之和為 53，當且僅當它們組成其中一個對，而每個整數恰好屬於一個對。所以 $X$ 就是兩個成員都在所選七個數之中的對的數目。由於 $2X \le 7$，$X$ 的取值為 $0,1,2,3$。

所有 $\binom{52}{7}$ 種選法的可能性相等，因此我們數算 $X=k$ 的選法數目。

- 選出 $k$ 個完整對：共 $\binom{26}{k}$ 種方法。這些對用去七個數中的 $2k$ 個。
- 其餘 $7-2k$ 個數必須來自餘下 $26-k$ 個對中 $7-2k$ 個不同的對（若有兩個來自同一對，便會多出一個完整對）：共 $\binom{26-k}{7-2k}$ 種方法。
- 從每個這樣的對中取其兩個成員之一：共 $2^{7-2k}$ 種方法。

每個 $X=k$ 的選法以這種方式恰好出現一次，所以

$$\mathrm{P}(X=k)=\frac{\binom{26}{k}\binom{26-k}{7-2k}2^{7-2k}}{\binom{52}{7}},\qquad k=0,1,2,3.$$

**具體數值**

$$\begin{aligned}
X=0:&\quad \tbinom{26}{7}2^7 = 657\,800\times 128 = 84\,198\,400,\\
X=1:&\quad 26\tbinom{25}{5}2^5 = 26\times 53\,130\times 32 = 44\,204\,160,\\
X=2:&\quad \tbinom{26}{2}\tbinom{24}{3}2^3 = 325\times 2024\times 8 = 5\,262\,400,\\
X=3:&\quad \tbinom{26}{3}\tbinom{23}{1}2 = 2600\times 23\times 2 = 119\,600,
\end{aligned}$$

而 $\binom{52}{7} = 133\,784\,560$。這四個數目加起來確實等於 $\binom{52}{7}$，這是一個很好的驗證，說明沒有遺漏或重複計算。

**想一想** 分數中有九位數時，我們該如何化簡？

題目中的附註就是提示：$5593 = 7\times 17\times 47$ 應該是分母。事實上 $\binom{52}{7} = 23\,920\times 5593$，而這四個數目每一個都是 23 920 的倍數：

$$84\,198\,400 = 3520\times 23\,920, \qquad 44\,204\,160 = 1848\times 23\,920,$$
$$5\,262\,400 = 220\times 23\,920, \qquad 119\,600 = 5\times 23\,920.$$

所以 $X$ 的分佈（distribution）為

| $k$ | 0 | 1 | 2 | 3 |
|---|---|---|---|---|
| $\mathrm{P}(X=k)$ | $\frac{3520}{5593}$ | $\frac{1848}{5593}$ | $\frac{220}{5593}$ | $\frac{5}{5593}$ |

（第二個的最簡形式為 $\frac{264}{799}$），而 $3520+1848+220+5 = 5593$，理應如此。

**期望值**

$$\mathrm{E}(X) = \frac{0\times 3520 + 1\times 1848 + 2\times 220 + 3\times 5}{5593} = \frac{2303}{5593}.$$

由於 $2303 = 7^2\times 47$ 而 $5593 = 7\times 17\times 47$，所以

$$\mathrm{E}(X) = \frac{7^2\times 47}{7\times 17\times 47} = \frac{7}{17},$$

正如所求。

**求期望值的捷徑**

對於 $j = 1,\ldots,26$，若第 $j$ 個對的兩個成員都被選中，則設 $I_j = 1$，否則為 0，於是 $X = I_1 + \cdots + I_{26}$。選中某一指定對的兩個成員後，須從其餘 50 個數中再選 5 個，所以

$$\mathrm{P}(I_j = 1) = \frac{\binom{50}{5}}{\binom{52}{7}} = \frac{7\times 6}{52\times 51} = \frac{7}{442}.$$

由期望值的線性性質（這不需要獨立性（independence）；事實上這些 $I_j$ 並不獨立），

$$\mathrm{E}(X) = 26 \times \frac{7}{442} = \frac{7}{17}.$$

## 驗算

- 四個概率之和為 1，而兩種方法都給出相同的 $\mathrm{E}(X) = \frac{7}{17} \approx 0.412$。
- 以電腦模擬 300 000 次隨機選取，得到 $X = 0, 1, 2, 3$ 的相對頻率約為 0.630、0.329、0.039、0.001，而準確值為 0.6294、0.3304、0.0393、0.0009。
- 合理性檢查：$\mathrm{P}(X = 3)$ 非常小，這合乎預期，因為只有七個數卻要有三個完整對。

## 常見錯誤

- 把 $X = 0$ 的選法數目算成 $\binom{52}{7}$ 減去某個粗略的數，或只算 $\binom{26}{7}$，忘記了每個對取哪一個成員所帶來的因子 $2^7$。
- 對於 $X = 1$，把其餘五個數的選法算成 $\binom{50}{5}$：這容許出現第二個完整對，因而多算了。

## STEP 之外

把一個計數寫成指示變量（indicator variable）之和，再運用期望值的線性性質，是大學一年級概率論中最有用的技巧之一：即使分佈根本無法求出，它仍能給出期望值。本題的完整分佈尚可求得，但例如要求隨機排列（permutation）的不動點（fixed point）的期望數目，指示變量便是最自然的切入點。Blitzstein 與 Hwang 的 *Introduction to Probability*（《概率導論》）一書反覆運用這個想法。

詳細解答

# STEP II 2013，第 1 題

純數學 • 對數曲線的弦

**STEP II 2013，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/13-s2-q1

本題在 *Beyond STEP* 第 3.5 節 $\pi^e$ 還是 $e^\pi$？中有完整的引導式講解：https://beyondstep.keithmcnulty.org/zh-hant/ch03/pi-e-or-e-pi.html

## 關鍵思路

這裏每一條不等式，都是比較 $p^q$ 與 $q^p$。取對數（logarithm）再除以 $pq$，就變成比較 $\frac{\ln p}{p}$ 與 $\frac{\ln q}{q}$；而 $\frac{\ln x}{x}$ 正是由原點連到曲線上點 $(x, \ln x)$ 的直線的斜率（gradient）。因此整道題都圍繞與 $y = \ln x$ 相交的直線和弦（chord）。我們需要的事實是：這條曲線是凹（concave）的（即向下彎曲），而且由原點出發的切線（tangent）在 $x = \mathrm{e}$ 處與曲線相切。

## 解答

### 第 (i) 部分

設 $y = mx$ 在 $x = x_0$ 處與 $y = \ln x$ 相切。兩者斜率相同，而且曲線相交，所以

$$m = \frac{1}{x_0}, \qquad mx_0 = \ln x_0.$$

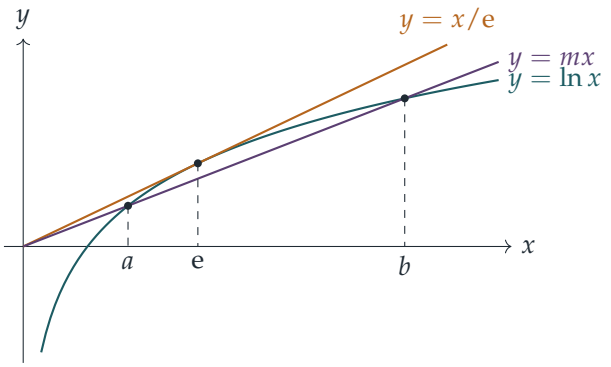
由第一條方程得 $mx_0 = 1$，所以 $\ln x_0 = 1$，$x_0 = \mathrm{e}$，以及

$$m = \frac{1}{\mathrm{e}}.$$

現在設 $y = mx$ 與曲線相交於 $x = a$ 及 $x = b$。那麼 $\ln a = ma$ 及 $\ln b = mb$，所以（因為 $a, b > 0$）

$$\frac{\ln a}{a} = m = \frac{\ln b}{b} \implies b \ln a = a \ln b \implies \ln\left(a^b\right) = \ln(b^a),$$

由於 ln 是一一對應（one-to-one）的，故 $a^b = b^a$，得證。




關於草圖：若 $m \le 0$，由於此時 $\ln x - mx$ 嚴格遞增，直線與曲線最多只相交一次。若 $m \ge \frac{1}{\mathrm{e}}$，則對 $x > 0$ 而言，直線就是那條切線或位於切線之上；而凹曲線位於其切線 $y = x/\mathrm{e}$ 之下，只在 e 處相切，所以同樣不會有兩個交點。因此 $0 < m < \frac{1}{\mathrm{e}}$，直線 $y = mx$ 位於切線之下，如圖所示。在 $x = \mathrm{e}$ 處，曲線的高度是 1，直線的高度是 $m\mathrm{e} < 1$，所以該處曲線在直線之上。由於曲線是凹的而直線是直的，$\ln x - mx$ 是凹函數。因此它取正值的集合是單一個區間，其兩端正是兩個交點 $a$ 和 $b$。這個區間包含 e，所以 $a < \mathrm{e} < b$。

**第 (ii) 部分**

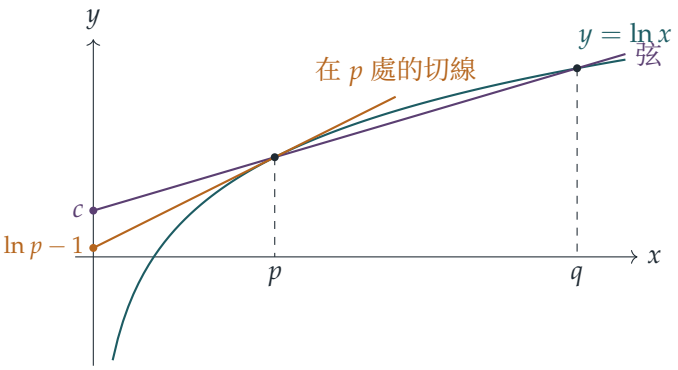
這裏 $\ln p = mp + c$ 及 $\ln q = mq + c$，其中 $0 < p < q$。分別除以 $p$ 和 $q$，得

$$\frac{\ln p}{p} = m + \frac{c}{p}, \qquad \frac{\ln q}{q} = m + \frac{c}{q}.$$

由於 $c > 0$ 且 $\frac{1}{p} > \frac{1}{q}$，我們有 $\frac{c}{p} > \frac{c}{q}$，所以 $\frac{\ln p}{p} > \frac{\ln q}{q}$。乘以 $pq > 0$ 得 $q \ln p > p \ln q$，即 $\ln(p^q) > \ln(q^p)$，因此 $p^q > q^p$。

從圖像上看，由原點連到 $(p, \ln p)$ 和 $(q, \ln q)$ 的直線，斜率分別是 $\frac{\ln p}{p}$ 和 $\frac{\ln q}{q}$。若弦在原點上方與 $y$ 軸相交，則較近的那一點與 $O$ 所連的直線較陡。（$c$ 的正負決定不等式的方向。我們在第 (iv) 部分會再用到這一點。）

### 第 (iii) 部分




設經過 $P = (p, \ln p)$ 和 $Q = (q, \ln q)$ 的弦的斜率為 $s$，$y$ 截距（intercept）為 $c$，所以 $c = \ln p - sp$。由於曲線是凹的，由 $P$ 連到更右方一點的弦，比 $P$ 處的切線平緩，所以

$$s < \frac{1}{p}.$$

（又或者，由中值定理（mean value theorem），對某個 $\xi \in (p, q)$ 有 $s = \frac{1}{\xi}$，而 $\frac{1}{\xi} < \frac{1}{p}$。）$P$ 處的切線是 $y = \ln p + \frac{1}{p}(x - p)$，它與 $y$ 軸相交於 $\ln p - 1$。弦和切線都經過 $P$，而弦較平緩。所以由 $P$ 向左移到 $x = 0$ 的過程中，弦一直在切線之上：

$$c = \ln p - sp > \ln p - 1 \geq 0,$$

最後一步用了 $p \geq \mathrm{e}$。所以弦在 $y$ 取正值之處與 $y$ 軸相交。

**想一想** 這樣如何判斷 $\pi^{\mathrm{e}}$ 與 $\mathrm{e}^{\pi}$ 的大小？

取 $p = \mathrm{e}$ 和 $q = \pi$，它們滿足 $\mathrm{e} \leq p < q$。這條弦的 $c > 0$，所以可應用第 (ii) 部分，得 $p^q > q^p$：

$$\mathrm{e}^{\pi} > \pi^{\mathrm{e}}.$$

### 第 (iv) 部分

現在經過 $(p, \ln p)$ 和 $(q, \ln q)$ 的弦的斜率是 $\frac{1}{\mathrm{e}}$，所以它是 $y = \frac{x}{\mathrm{e}} + c$，其中

$$c = \ln p - \frac{p}{\mathrm{e}}.$$

直線 $y = \frac{x}{\mathrm{e}}$ 就是第 (i) 部分中在 $x = \mathrm{e}$ 處的切線，而凹曲線位於它之上或之下——準確地說是在它上面或下方，即 $\ln x \leq \frac{x}{\mathrm{e}}$，等號只在 $x = \mathrm{e}$ 時成立。所以 $c \leq 0$，而 $c = 0$ 只在 $p = \mathrm{e}$ 時出現。但 $c = 0$ 會令弦就是切線本身，而切線只在 e 處與曲線相交，並非在兩點 $p < q$ 相交。因此

$$c < 0.$$

這條弦與切線平行，並位於切線之下。現在把第 (ii) 部分的論證反轉符號再做一次：

$$\frac{\ln p}{p} = \frac{1}{\mathrm{e}} + \frac{c}{p}, \qquad \frac{\ln q}{q} = \frac{1}{\mathrm{e}} + \frac{c}{q}.$$

由於 $c < 0$ 且 $\frac{1}{p} > \frac{1}{q}$，得 $\frac{c}{p} < \frac{c}{q}$，所以 $\frac{\ln p}{p} < \frac{\ln q}{q}$。乘以 $pq > 0$ 得 $q \ln p < p \ln q$，因此

$$q^p > p^q.$$

（與第 (i) 部分一樣，凹性在這裏也迫使 $p < \mathrm{e} < q$，但論證並不需要這一點。）

## 驗算

- $\mathrm{e}^\pi \approx 23.14$ 而 $\pi^\mathrm{e} \approx 22.46$，與第 (iii) 部分一致。
- 第 (i) 部分：$a = 2$、$b = 4$ 這一對滿足 $2^4 = 4^2$，而且確實 $\frac{\ln 2}{2} = \frac{\ln 4}{4}$ 及 $2 < \mathrm{e} < 4$。
- 第 (iv) 部分：取 $p = 1$，由弦的斜率 $\frac{\ln q}{q-1} = \frac{1}{\mathrm{e}}$ 得 $q \approx 5.76$。於是 $q^p \approx 5.76 > 1 = p^q$。
- 第 (ii) 和第 (iii) 部分亦已用電腦對大量隨機的弦作檢驗確認。

## 常見錯誤

- 把 $q \ln p > p \ln q$ 的方向弄反。乘以 $pq$ 之前，先清楚寫出 $\frac{\ln p}{p} > \frac{\ln q}{q}$。
- 在第 (iii) 部分，只寫到「曲線是凹的」便停下，沒有解釋為何截距是正數。$p$ 處的切線與軸相交於 $\ln p - 1 \geq 0$，而弦在它之上穿過。
- 在第 (iv) 部分，忘記排除 $c = 0$，而這是得到嚴格不等式所必需的。

## STEP 之外

函數 $\frac{\ln x}{x}$ 在 e 處取得最大值，這解釋了著名的一對 $2^4 = 4^2$。方程 $x^y = y^x$ 在 $x \neq y$ 時的正數解構成一條曲線，可以參數化（parametrise）為 $x = t^{1/(t-1)}$、$y = t^{t/(t-1)}$。更一般地說，「凹函數的弦位於曲線之下、切線之上」是凸性（convexity）的基本不等式。它是延森不等式（Jensen's inequality）和算術–幾何平均不等式（AM–GM inequality）的基礎。Bartle 與 Sherbert 的 *Introduction to Real Analysis*（《實分析導論》）對中值定理和凸函數有仔細的論述。

詳細解答

# STEP II 2013，第 2 題

純數學 • 歸約公式

> **STEP II 2013，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/13-s2-q2

## 關鍵思路

被積函數（integrand）$x^n(1-x)^n$ 在反射 $x \mapsto 1-x$ 下對稱（symmetric），而這種對稱性讓我們可以毫無代價地互換 $x$ 與 $1-x$ 的角色。接着用分部積分（integration by parts）把 $x$ 的一個冪換成 $1-x$ 的一個冪，再利用恆等式 $x+(1-x)=1$ 把乘積拆回我們已知的積分。兩者結合便得到一條遞推公式（reduction formula）；對整數 $n$，它可以逐步展開成階乘（factorial），而對半整數（half-integer）也同樣適用。

## 解答

### 第 (i) 部分

對 $n \geq 1$，記 $J_n = \int_0^1 x^{n-1}(1-x)^n \,\mathrm{d}x$。代換（substitute）$u = 1-x$，於是 $\mathrm{d}x = -\mathrm{d}u$，而積分限 $0, 1$ 變成 $1, 0$：

$$J_n = \int_1^0 (1-u)^{n-1}u^n\,(-\mathrm{d}u) = \int_0^1 u^n(1-u)^{n-1}\,\mathrm{d}u,$$

這就是所要求的等式（啞變量（dummy variable）用甚麼名字並不重要）。

把 $J_n$ 的兩個相等形式相加，

$$\begin{aligned} 2J_n &= \int_0^1 \left[x^{n-1}(1-x)^n + x^n(1-x)^{n-1}\right]\mathrm{d}x \\ &= \int_0^1 x^{n-1}(1-x)^{n-1}\left[(1-x)+x\right]\mathrm{d}x = I_{n-1}. \end{aligned}$$

接着對 $I_n$ 作分部積分，取 $u = x^n$ 及 $\mathrm{d}v = (1-x)^n\,\mathrm{d}x$，故 $v = -\dfrac{(1-x)^{n+1}}{n+1}$：

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

邊界項（boundary term）在兩端都等於零：在 $x = 1$ 處是因為有因子 $(1-x)^{n+1}$，在 $x = 0$ 處則是因為有因子 $x^n$ 且 $n \geq 1$。由此得到題目所給的公式

$$I_n = \frac{n}{n+1}\int_0^1 x^{n-1}(1-x)^{n+1}\,\mathrm{d}x.$$

**想一想** 怎樣才能從這個積分回到 $I_n$ 和 $I_{n-1}$？

分出一個因子 $1-x$，並照字面寫成 $1-x$：

$$x^{n-1}(1-x)^{n+1} = x^{n-1}(1-x)^n\,(1-x) = x^{n-1}(1-x)^n - x^n(1-x)^n.$$

積分後，利用上面的結果，右邊等於 $J_n - I_n = \frac{1}{2}I_{n-1} - I_n$。所以

$$I_n = \frac{n}{n+1}\left(\tfrac{1}{2}I_{n-1} - I_n\right) \quad\Longrightarrow\quad I_n \cdot \frac{2n+1}{n+1} = \frac{n}{2(n+1)}I_{n-1},$$

再乘以 $\dfrac{n+1}{2n+1}$，

$$I_n = \frac{n}{2(2n+1)}\,I_{n-1},$$

正是所求。留意這個論證完全沒有用到 $n$ 是整數：只需要 $n \geq 1$（為了處理邊界項，以及確保每個被積函數在 $[0,1]$ 上連續）。我們會在第 (iii) 部分用到這一點。

### 第 (ii) 部分

我們用歸納法（induction）證明 $I_n = \dfrac{(n!)^2}{(2n+1)!}$，從 $n = 0$ 開始：$I_0 = \int_0^1 1\,\mathrm{d}x = 1 = \dfrac{(0!)^2}{1!}$。假設公式對 $n-1$ 成立，其中 $n \geq 1$。那麼由第 (i) 部分，

$$\begin{aligned} I_n = \frac{n}{2(2n+1)} \cdot \frac{\big((n-1)!\big)^2}{(2n-1)!} &= \frac{n^2\big((n-1)!\big)^2}{2n(2n+1)(2n-1)!} \\ &= \frac{(n!)^2}{(2n+1)!}, \end{aligned}$$

這裏我們把分子和分母同乘以 $n$，並用了 $2n(2n+1)(2n-1)! = (2n+1)!$。歸納至此完成，所以公式對每個正整數 $n$（以及 $n = 0$）都成立。

### 第 (iii) 部分

取 $x = \sin^2\theta$，則 $\mathrm{d}x = 2\sin\theta\cos\theta\,\mathrm{d}\theta$，而 $\theta$ 由 0 變到 $\frac{\pi}{2}$。在這個區間上 $\sin\theta$ 和 $\cos\theta$ 都非負，所以

$$\big(x(1-x)\big)^{1/2} = \big(\sin^2\theta\cos^2\theta\big)^{1/2} = \sin\theta\cos\theta.$$

**想一想** 為甚麼這裏需要 $\sin\theta\cos\theta \geq 0$？

因為 $\sqrt{t^2} = |t|$，而不是 $t$；在 $[0, \frac{\pi}{2}]$ 上，絕對值（modulus）符號可以直接去掉。因此

$$\begin{aligned} I_{\frac{1}{2}} &= \int_0^{\pi/2} \sin\theta\cos\theta\cdot 2\sin\theta\cos\theta\,\mathrm{d}\theta = \int_0^{\pi/2} \tfrac{1}{2}\sin^2 2\theta\,\mathrm{d}\theta \\ &= \int_0^{\pi/2} \tfrac{1}{4}\big(1-\cos 4\theta\big)\mathrm{d}\theta = \tfrac{1}{4}\Big[\theta - \tfrac{1}{4}\sin 4\theta\Big]_0^{\pi/2} = \frac{\pi}{8}. \end{aligned}$$

至於 $I_{\frac{3}{2}}$，用第 (i) 部分的遞推公式並取 $n = \frac{3}{2}$；這是容許的，因為 $\frac{3}{2} \geq 1$，而推導過程從來不需要 $n$ 是整數：

$$I_{\frac{3}{2}} = \frac{3/2}{2\cdot 4}\, I_{\frac{1}{2}} = \frac{3}{16}\cdot\frac{\pi}{8} = \frac{3\pi}{128}.$$

作為交叉驗證，同一代換可直接得到

$$I_{\frac{3}{2}} = \int_0^{\pi/2} 2\sin^4\theta\cos^4\theta\,\mathrm{d}\theta = \frac{1}{8}\int_0^{\pi/2} \sin^4 2\theta\,\mathrm{d}\theta = \frac{1}{8}\cdot\frac{3\pi}{16} = \frac{3\pi}{128},$$

其中用了 $\sin^4\phi = \frac{3}{8} - \frac{1}{2}\cos 2\phi + \frac{1}{8}\cos 4\phi$。

## 驗算

- $I_1 = \int_0^1 (x - x^2)\,\mathrm{d}x = \frac{1}{6}$，而公式給出 $\frac{1}{3!} = \frac{1}{6}$。同樣，$I_2 = \int_0^1 (x^2 - 2x^3 + x^4)\,\mathrm{d}x = \frac{1}{3} - \frac{1}{2} + \frac{1}{5} = \frac{1}{30} = \frac{4}{5!}$。
- $I_{\frac{1}{2}}$ 是半徑為 $\frac{1}{2}$ 的半圓（semicircle）$y = \sqrt{\frac{1}{4} - (x-\frac{1}{2})^2}$ 下方的面積，即 $\frac{1}{2}\pi\big(\frac{1}{2}\big)^2 = \frac{\pi}{8}$。
- 數值上，$I_{\frac{3}{2}} \approx 0.073631$，與 $3\pi/128$ 吻合；遞推公式亦已用電腦對非整數 $n$（例如 $n = 2.7$）驗證過。

## 常見錯誤

- 忘記檢查分部積分中的邊界項是否等於零，尤其是在 $x = 0$ 處。
- 寫 $\sqrt{\sin^2\theta\cos^2\theta} = \sin\theta\cos\theta$ 時沒有指出在 $[0, \frac{\pi}{2}]$ 上的正負號。
- 用第 (ii) 部分的階乘公式計算 $I_{\frac{3}{2}}$，但該公式只對整數證明過；用遞推公式才是穩妥的做法。

## STEP 之外

$I_n$ 是貝塔函數（beta function）$B(p,q) = \int_0^1 x^{p-1}(1-x)^{q-1}\,\mathrm{d}x$ 的特例，這裏 $I_n = B(n+1, n+1)$，而第 (ii) 部分正是 $B(p,q) = \Gamma(p)\Gamma(q)/\Gamma(p+q)$ 的整數情況。把 $n = \frac{1}{2}$ 代入這個恆等式，便可得出 $\Gamma(\frac{1}{2}) = \sqrt{\pi}$，這就是第 (iii) 部分出現 $\pi$ 的更深層原因。這類積分是統計學中貝塔分佈（beta distribution）以及沃利斯（Wallis）的 $\pi$ 乘積公式背後的基礎；Siklos 的 *Advanced Problems in Mathematics*（《高等數學難題》）收錄了更多同類風格的遞推公式題目。

詳細解答

# STEP II 2013，第 3 題

純數學 • 三次方程的根

> **STEP II 2013，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/13-s2-q3

## 關鍵思路

一切都源於首項係數為正、且有三個相異實根的三次函數（cubic）的形狀：曲線從 $-\infty$ 上升，向上穿過橫軸，回落，然後再次向上穿過橫軸。數值 $f(0) = c$ 告訴我們 0 相對於各根的位置，而轉向點（turning points）（即二次方程 $f'(x) = 0$ 的根）則可以透過它們的和與積來確定。最後一部分提醒我們：單靠轉向點的條件，並不能保證曲線真的與橫軸相交三次。

## 解答

全文中設

$$f(x) = x^3 + 3ax^2 + 3bx + c, \qquad f'(x) = 3\big(x^2 + 2ax + b\big).$$

當 $f(x) = 0$ 有三個相異實根 $\alpha < \beta < \gamma$ 時，我們可以寫成 $f(x) = (x-\alpha)(x-\beta)(x-\gamma)$，於是從三個因式的正負號可讀出

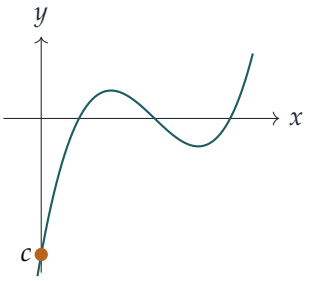
$$f(x) < 0 \text{ 於 } (-\infty, \alpha) \cup (\beta, \gamma), \qquad f(x) > 0 \text{ 於 } (\alpha, \beta) \cup (\gamma, \infty).$$

此外，由羅爾定理（Rolle's theorem），$f'$ 在 $(\alpha, \beta)$ 內有一個零點，在 $(\beta, \gamma)$ 內有另一個零點。由於 $f'$ 是二次函數，它再沒有其他零點，因此轉向點位於 $p < q$，且

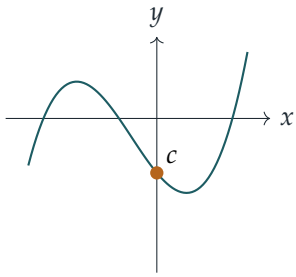
$$\alpha < p < \beta < q < \gamma.$$

### 第 (i) 部分

曲線 $y = f(x)$ 與 $y$ 軸相交於 $(0, c)$；由於 $c < 0$，這點位於 $x$ 軸下方。根據上面的正負號規律，$f(0) < 0$ 表示 $0 < \alpha$ 或 $\beta < 0 < \gamma$。這就是下面的兩幅圖。



三個根皆為正



恰有一個根為正

若 $0<\alpha$，則三個根皆為正。若 $\beta<0<\gamma$，則 $\alpha<\beta<0$，因此恰有一個根（$\gamma$）為正。由於 $f(0)=c\neq 0$，沒有根等於零。所以要麼恰有一個根為正，要麼三個根皆為正，得證。

**第 (ii) 部分**

現在 $0<\alpha<\beta<\gamma$。

$c$ 的正負號。由於 $0<\alpha$，點 0 位於 $(-\infty,\alpha)$ 內，而在該區間上 $f<0$。所以 $c=f(0)<0$。

轉向點。方程 $x^2+2ax+b=0$ 有兩個相異實根 $p$ 和 $q$，因此其判別式（discriminant）為正：

$$4a^2-4b>0 \quad\Longrightarrow\quad a^2>b.$$

由於 $p>\alpha>0$ 且 $q>p$，兩個轉向點皆為正。對於二次式 $x^2+2ax+b$，兩根之和與積分別為

$$p+q=-2a, \qquad pq=b.$$

由於 $p,q>0$，得 $-2a>0$，所以 $a<0$，且 $b=pq>0$。綜合以上，

$$a^2>b>0, \qquad a<0, \qquad c<0,$$

這就是 $(*)$。

（你也可以直接從根得出 $a<0$、$b>0$ 和 $c<0$：$\alpha+\beta+\gamma=-3a$、$\alpha\beta+\beta\gamma+\gamma\alpha=3b$ 和 $\alpha\beta\gamma=-c$ 皆為正。但 $a^2>b$ 確實需要用到轉向點。）

**第 (iii) 部分**

現在 $c=f(0)>0$，所以 0 位於 $(\alpha,\beta)$ 或 $(\gamma,\infty)$ 內，即 $f>0$ 的兩個區間。這給出兩幅可能的圖：要麼 $\alpha<0<\beta<\gamma$（一個負根、兩個正根），要麼 $\alpha<\beta<\gamma<0$（三個根皆為負）。

**想一想** 條件 $ab < 0$ 怎樣排除其中一種情況？

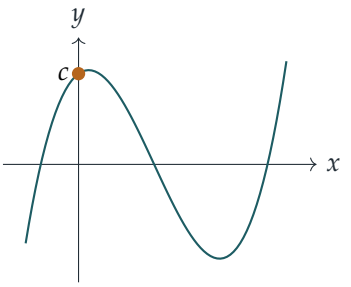
再看看轉向點。若三個根皆為負，則 $p < q < \gamma < 0$，所以兩個轉向點皆為負。於是

$$-2a = p + q < 0 \quad \text{及} \quad b = pq > 0,$$

所以 $a > 0$ 且 $b > 0$，得 $ab > 0$。這與 $ab < 0$ 矛盾。因此第二幅圖不可能出現，故

$$\alpha < 0 < \beta < \gamma :$$

一個根爲負，另外兩個根爲正。




這種情況確實可以出現。例如，根 $-1$、$2$ 和 $5$ 給出 $x^3 - 6x^2 + 3x + 10$，所以 $a = -2$、$b = 1$、$c = 10$，其中 $ab = -2 < 0$ 且 $c > 0$；這就是上面所繪的曲線。

### 第 (iv) 部分

**想一想** 除了 $(*)$ 之外，三次函數還需要甚麼額外條件才有三個實根？

條件 $(*)$ 只控制轉向點和 $f(0)$ 的正負號。要有三個實根，局部極大值（local maximum）還必須位於橫軸上方。所以我們應選取 $a$ 和 $b$ 使兩個轉向點皆為正，然後令 $c$ 足夠負，把局部極大值拉到橫軸下方。

取 $a = -1$ 及 $b = \frac{3}{4}$，使得

$$f'(x) = 3\left(x^2 - 2x + \tfrac{3}{4}\right) = 3\left(x - \tfrac{1}{2}\right)\left(x - \tfrac{3}{2}\right).$$

局部極大值位於 $x = \frac{1}{2}$，該處

$$f\left(\tfrac{1}{2}\right) = \tfrac{1}{8} - \tfrac{3}{4} + \tfrac{9}{8} + c = \tfrac{1}{2} + c.$$

選取 $c = -1$。方程於是為

$$x^3 - 3x^2 + \tfrac{9}{4}x - 1 = 0 \qquad \left(a = -1,\ b = \tfrac{3}{4},\ c = -1\right),$$

且 $a^2 = 1 > \frac{3}{4} = b > 0$、$a < 0$、$c < 0$，所以 $(*)$ 成立。

現在數一數根的個數。$f$ 在 $(-\infty, \frac{1}{2}]$ 上遞增，所以在該處 $f(x) \le f(\frac{1}{2}) = -\frac{1}{2} < 0$。在 $[\frac{1}{2}, \frac{3}{2}]$ 上，$f$ 遞減，所以一直低於 $-\frac{1}{2}$。在 $[\frac{3}{2}, \infty)$ 上，$f$ 從 $f(\frac{3}{2}) = -1$ 遞增至 $+\infty$，所以恰好穿過橫軸一次。因此恰有一個實根（由於 $f(2) = -\frac{1}{2}$ 及 $f(3) = \frac{23}{4}$，它位於 2 與 3 之間）。

## 驗算

- 三次式 $(x-1)(x-2)(x-3) = x^3 - 6x^2 + 11x - 6$ 的 $a = -2$、$b = \frac{11}{3}$、$c = -6$，而的確 $a^2 = 4 > \frac{11}{3} > 0$，正如第 (ii) 部分所預測。
- 把曲線反射，即 $x \mapsto -x$，會把根 $\alpha, \beta, \gamma$ 變為 $-\alpha, -\beta, -\gamma$，並改變 $a$ 和 $c$ 的正負號，但不改變 $b$。所以 $ab$ 和 $c$ 都改變正負號，而第 (i) 部分的鏡像版本說明：當 $c > 0$ 時，要麼恰有一個根為負，要麼三個根皆為負——這正是我們在第 (iii) 部分開始時的兩幅圖。
- 在第 (iv) 部分中，唯一的實根是 $x \approx 2.178$。以電腦檢查數以千計隨機生成、具三個相異實根的三次式，結果亦與第 (i)–(iii) 部分吻合。

## 常見錯誤

- 試圖單從根的正負號證明 $a^2 > b$。這個不等式其實來自轉向點是相異實數這一事實。
- 以為 $(*)$ 等價於有三個正根。第 (iv) 部分說明它只是必要條件（necessary），而非充分條件（sufficient）。
- 在第 (iv) 部分中，選取的 $c$ 使局部極小值為負，但局部極大值仍為正。這樣會得到三個根，而不是一個。

## STEP 之外

這裏的正負號論證是一套一般理論的起點：笛卡兒符號法則（Descartes' rule of signs）以係數中正負號改變的次數作為正根個數的上界，而斯圖姆定理（Sturm's theorem）則能準確計算某區間內實根的個數。對於三次函數，第 (iv) 部分所缺少的條件是判別式的正負號；判別式（相差一個正因子）就是兩個轉向點處函數值的乘積 $-f(p)f(q)$。Allenby 的 *Rings, Fields and Groups*（《環、域與羣》）以大學一年級程度討論多項式及其根。

詳細解答

# STEP II 2013，第 4 題

純數學 • 弦中點的軌跡

**STEP II 2013，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/13-s2-q4

## 關鍵思路

求弦（chord）的中點（midpoint）時，根本不必求出弦的兩個端點：$P$ 和 $Q$ 的 $x$ 坐標是某個二次方程（quadratic equation）的根，而它們的平均值就是首兩項係數之比的一半再取負號。之後，每個軌跡（locus）都只是消去一個參數（parameter）的問題。每個答案都可以用幾何解釋：由圓心到弦中點的直線與該弦垂直（perpendicular）。

## 解答

### $M$ 的坐標

直線為 $y = b(x - a)$。代入 $x^2 + y^2 = 1$，得

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

它的根是 $P$ 和 $Q$ 的 $x$ 坐標 $x_P$、$x_Q$，所以由根之和（sum of the roots）得

$$x_M = \tfrac{1}{2}(x_P + x_Q) = \frac{ab^2}{1 + b^2}.$$

由於 $M$ 在該直線上，$y_M = b(x_M - a) = b \cdot \dfrac{ab^2 - a - ab^2}{1 + b^2}$，所以

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

要 $P$ 和 $Q$ 存在且互不相同，判別式（discriminant）必須為正：

$$4a^2b^4 - 4(1 + b^2)(a^2b^2 - 1) = 4\left(1 + b^2 - a^2b^2\right) > 0 \iff a^2b^2 < 1 + b^2.$$

（這表示由 $O$ 到該直線的距離 $|ab|/\sqrt{1 + b^2}$ 小於 1。）每一部分都需要用到這個條件。

**第 (i) 部分**

固定 $b > 0$，則

$$-by_M = -b \cdot \left(-\frac{ab}{1+b^2}\right) = \frac{ab^2}{1+b^2} = x_M,$$

所以軌跡上每一點都滿足 $x = -by$，正如所求。這是通過 $O$、斜率（gradient）為 $-1/b$ 的直線，與弦垂直，與關鍵思路的預期完全一致。

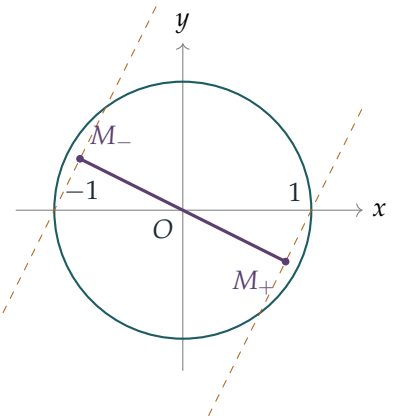
若 $-1 \le a \le 1$，則 $a^2b^2 \le b^2 < 1+b^2$，所以直線必與圓相交於兩點，對每個這樣的 $a$，$M$ 都有定義。$M$ 的兩個坐標都是 $a$ 乘以一個常數，所以當 $a$ 取遍 $[-1,1]$ 時，點 $M$ 沿直線 $x = -by$ 由 $a = -1$ 對應的點穩定地移動到 $a = 1$ 對應的點。因此軌跡是連接

$$M_\pm = \pm\left(\frac{b^2}{1+b^2}, -\frac{b}{1+b^2}\right),$$

的線段，其長度為

$$2\,\frac{\sqrt{b^4+b^2}}{1+b^2} = \frac{2b\sqrt{1+b^2}}{1+b^2} = \frac{2b}{(1+b^2)^{1/2}},$$

這裏利用了 $b > 0$ 來開 $b^2$ 的平方根。這就是所求的長度。線段以 $O$ 為中點；其端點 $M_+$ 是通過 $(1,0)$ 的弦的中點，而 $M_-$ 則是通過 $(-1,0)$ 的弦的中點。下圖顯示 $b = 2$ 的情況，此時端點為 $\pm(\frac{4}{5}, -\frac{2}{5})$。




虛線是兩條極端的弦，分別通過 $(\pm1,0)$，斜率為 2。

**第 (ii)(a) 部分**

現在固定 $a$，且 $0 < a < 1$。於是 $a^2b^2 < b^2 < 1+b^2$，所以每條通過 $A = (a,0)$ 的直線都與圓相交兩次。由 $M$ 的公式，

$$x^2 + y^2 = \frac{a^2b^4 + a^2b^2}{(1+b^2)^2} = \frac{a^2b^2}{1+b^2} = a \cdot \frac{ab^2}{1+b^2} = ax.$$

所以 $M$ 在

$$x^2+y^2-ax=0, \quad \text{即} \quad \left(x-\tfrac{a}{2}\right)^2+y^2=\left(\tfrac{a}{2}\right)^2,$$

上，這是圓心為 $(\frac{a}{2},0)$、半徑為 $\frac{a}{2}$ 的圓，以 $OA$ 為直徑（diameter）。

**想一想** 這個圓是否整個都被描出？

並不完全。由於 $x=ab^2/(1+b^2)=a-a/(1+b^2)$，總有 $0\le x<a$，所以點 $A=(a,0)$ 本身永遠不會到達；要到達它，需要通過 $A$ 的鉛垂線，而鉛垂線沒有斜率。其他每一點都能到達：設 $b=\tan\theta$，其中 $-\frac{\pi}{2}<\theta<\frac{\pi}{2}$，

$$M=\left(a\sin^2\theta,\ -a\sin\theta\cos\theta\right)=\left(\tfrac{a}{2}-\tfrac{a}{2}\cos 2\theta,\ -\tfrac{a}{2}\sin 2\theta\right),$$

當 $2\theta$ 取遍 $(-\pi,\pi)$ 時，這恰好走遍整個圓一次，只差 $2\theta=\pm\pi$ 對應的點，即 $A$。所以軌跡是以 $OA$ 為直徑的圓，但除去單獨一點 $A$。從幾何角度看：$OM$ 與弦 $AM$ 垂直，所以 $\angle OMA$ 是直角，$M$ 在以 $OA$ 為直徑的圓上（當 $b=0$ 時 $M=O$）。這個圓位於單位圓之內。

### 第 (ii)(b) 部分

現在 $a=1/b$，其中 $0<b\le 1$。條件 $a^2b^2<1+b^2$ 變成 $1<1+b^2$，此式成立。直線為 $y=b(x-1/b)=bx-1$，所以它必定通過單位圓上的點 $B=(0,-1)$；因此 $B$ 必定是 $P$、$Q$ 其中之一。公式給出

$$M=\left(\frac{b}{1+b^2},\ -\frac{1}{1+b^2}\right), \qquad x^2+y^2=\frac{b^2+1}{(1+b^2)^2}=\frac{1}{1+b^2}=-y.$$

所以 $M$ 在

$$x^2+y^2+y=0, \quad \text{即} \quad x^2+\left(y+\tfrac{1}{2}\right)^2=\left(\tfrac{1}{2}\right)^2,$$

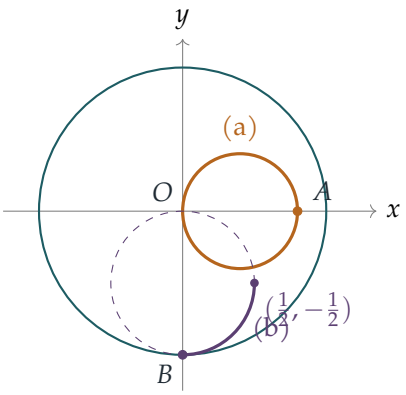
上，這是圓心為 $(0,-\frac{1}{2})$、半徑為 $\frac{1}{2}$ 的圓，以 $OB$ 為直徑（理由與 (a) 中的直角相同）。

這次只描出圓的一部分。設 $b=\tan\theta$，$0<\theta\le\frac{\pi}{4}$，

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

而 $2\theta$ 取遍 $(0,\frac{\pi}{2}]$。所以軌跡是四分之一個圓：由 $B=(0,-1)$（不包括在內，因為它需要 $b=0$）沿逆時針方向到 $(\frac{1}{2},-\frac{1}{2})$（包括在內，由 $b=1$、$a=1$ 得到）的圓弧。在這段弧上，$0<x\le\frac{1}{2}$ 且 $-1<y\le-\frac{1}{2}$。

下圖連同單位圓顯示兩個軌跡，(a) 中取 $a=0.8$。空心圓點標示 (a) 中缺少的點 $A$ 和 (b) 中缺少的端點 $B$。




在 (b) 中，實線弧是軌跡；小圓其餘的虛線部分不會到達。

## 驗算

- 當 $a = 0$ 時，直線通過 $O$，所以弦是直徑，$M$ 應該是 $O$：公式給出 $(0,0)$。
- 在 (i) 中取 $b = 1$：通過 $(1,0)$、斜率為 1 的弦連接 $(1,0)$ 和 $(0,-1)$，中點為 $(\frac{1}{2}, -\frac{1}{2})$，與 $M_+$ 吻合。此時線段長度為 $\sqrt{2} = 2/\sqrt{2}$，與公式一致。
- 同一條弦正是 (ii)(b) 中圓弧在 $b = 1$ 的端點，而 $(\frac{1}{2}, -\frac{1}{2})$ 滿足 $x^2 + y^2 + y = \frac{1}{4} + \frac{1}{4} - \frac{1}{2} = 0$。
- 在 (ii)(a) 中，$b = 1$ 給出 $M = (\frac{a}{2}, -\frac{a}{2})$，即以 $OA$ 為直徑的圓的最低點，對於傾角 45° 的弦，理應如此。

## 常見錯誤

- 用二次方程公式明確求出 $P$ 和 $Q$。這樣做可行，但用根之和一行就能得出 $M$。
- 在 (ii)(a) 或 (ii)(b) 中聲稱軌跡是整個圓。必須檢查參數實際取哪些值，以及這些值對應哪些點。
- 忘記檢查直線是否真的與圓相交；在 (i) 和 (ii) 的兩種情況中，它都必定相交，但這一點需要說明。

## STEP 之外

平行弦的中點位於一條通過中心的直線上，這個事實對任何圓錐曲線（conic）都成立：對橢圓而言，這條直線稱為*共軛直徑*（conjugate diameter），這也讓我們初步窺見射影幾

何中極點與極線（poles and polars）的概念。第 (ii) 部分中的圓，其實就是經典定理「半圓上的圓周角是直角」反過來運用。若想進一步了解軌跡和以參數化（parametrisation）作為解題工具，Siklos 的 *Advanced Problems in Mathematics*（《高等數學難題》）是很好的下一步。

詳細解答

# STEP II 2013，第 5 題

純數學 • 對稱性與曲線描繪

**STEP II 2013，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/13-s2-q5

## 關鍵思路

如果一個函數在 $x \mapsto 1-x$ 或 $x \mapsto 1/x$ 這類代換下保持不變，那麼用鏈式法則（chain rule）對該恆等式求導，便能把 $x$ 處的斜率與像點處的斜率聯繫起來。在代換的不動點（fixed point）上，斜率必須等於自身的相反數，因此為零。至於第 (iii) 部分的方程，對稱性提示我們把一切以 $u = x^2 - x$ 表示，而它本身在 $x \mapsto 1-x$ 下不變；這樣一來，處理關於 $u$ 的三次方程，遠比處理關於 $x$ 的六次方程容易。

## 解答

### 第 (i) 部分

把 $\mathrm{f}(x) = \mathrm{f}(1-x)$ 對 $x$ 求導，右邊使用鏈式法則，得

$$\mathrm{f}'(x) = -\mathrm{f}'(1-x) \qquad \text{對所有} x.$$

代入 $x = \frac{1}{2}$，得 $\mathrm{f}'(\frac{1}{2}) = -\mathrm{f}'(\frac{1}{2})$，因此 $\mathrm{f}'(\frac{1}{2}) = 0$。

現在再假設當 $x \neq 0$ 時 $\mathrm{f}(x) = \mathrm{f}(1/x)$。由於 $\frac{\mathrm{d}}{\mathrm{d}x}(\frac{1}{x}) = -\frac{1}{x^2}$，求導得

$$\mathrm{f}'(x) = -\frac{1}{x^2}\,\mathrm{f}'\left(\frac{1}{x}\right) \qquad (x \neq 0).$$

在 $x = -1$（它是自身的倒數）處，此式變為 $\mathrm{f}'(-1) = -\mathrm{f}'(-1)$，因此 $\mathrm{f}'(-1) = 0$。

在 $x = 2$ 處，由第一個結果，此式給出 $\mathrm{f}'(2) = -\frac{1}{4}\mathrm{f}'(\frac{1}{2}) = 0$。（同樣地，在第一個恆等式中取 $x = 2$，可得 $\mathrm{f}'(2) = -\mathrm{f}'(-1) = 0$。）

**第 (ii) 部分**

以 $1/x$ 代替 $x$。分子變為

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

而分母變為

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

相除，並利用 $x^2(1-x)^2=(x^2-x)^2$，得

$$\mathrm{f}\left(\frac{1}{x}\right)=\frac{(x^2-x+1)^3}{x^6}\cdot\frac{x^4}{(1-x)^2}=\frac{(x^2-x+1)^3}{x^2(x-1)^2}=\mathrm{f}(x).$$

以 $1-x$ 代替 $x$：由於 $(1-x)^2-(1-x)=(1-x)(-x)=x^2-x$，所以 $x^2-x$ 不變，從而 $x^2-x+1$ 亦不變。因此 $\mathrm{f}(1-x)=\mathrm{f}(x)$。

**草圖**。由第 (i) 部分，$x=-1$、$\frac{1}{2}$ 和 2 是駐點（stationary point）；由於駐點恰好有三個，這些就是全部。它們的函數值為

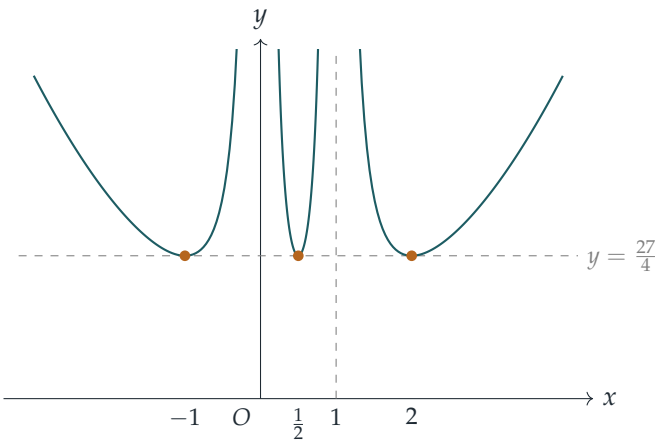
$$\mathrm{f}(-1)=\frac{3^3}{2^2}=\frac{27}{4},\qquad \mathrm{f}(\tfrac{1}{2})=\frac{(3/4)^3}{(1/4)^2}=\frac{27}{4},\qquad \mathrm{f}(2)=\frac{3^3}{2^2}=\frac{27}{4}.$$

三者相同，這是必然的：$\mathrm{f}(2)=\mathrm{f}(\frac{1}{2})$ 且 $\mathrm{f}(2)=\mathrm{f}(1-2)=\mathrm{f}(-1)$。

其他特徵：

- $x^2-x+1=(x-\frac{1}{2})^2+\frac{3}{4}>0$，所以處處有 $\mathrm{f}(x)>0$；曲線永不與 $x$ 軸相交。
- 當 $x\to 0$ 或 $x\to 1$ 時，分子趨向 1，分母趨向 $0^+$，所以從兩側皆有 $\mathrm{f}(x)\to+\infty$：$x=0$ 和 $x=1$ 是鉛直漸近線（vertical asymptote）。
- 當 $|x|$ 很大時，$\mathrm{f}(x)\approx x^6/x^4=x^2$，所以當 $x\to\pm\infty$ 時 $\mathrm{f}(x)\to+\infty$，大致像一條拋物線。

因此在 $(-\infty,0)$、$(0,1)$、$(1,\infty)$ 這三個區間的每一個上，函數都是連續的，並在兩端都趨向 $+\infty$。所以每個區間都必定包含一個最小值點，而它是駐點；每個區間恰好包含 $-1$、$\frac{1}{2}$、2 其中之一，所以該點就是最小值點，且沒有其他轉向點（turning point）。每一支都呈 U 形，最低值為 $\frac{27}{4}$。




**第 (iii) 部分**

**方程** $\mathrm{f}(x) = \frac{27}{4}$。從草圖可見，$\frac{27}{4}$ 是每一支上的最小值，而且只在三個駐點處取得。所以根為

$$x = -1, \qquad x = \tfrac{1}{2}, \qquad x = 2,$$

而對定義域內其他所有 $x$ 都有 $\mathrm{f}(x) > \frac{27}{4}$，即對於

$$x < -1, \quad -1 < x < 0, \quad 0 < x < \tfrac{1}{2}, \quad \tfrac{1}{2} < x < 1, \quad 1 < x < 2, \quad x > 2.$$

我們可以用代數方法確認這一點。設 $u = x^2 - x$（故在定義域內 $u \neq 0$）；則 $\mathrm{f}(x) = \dfrac{(u+1)^3}{u^2}$，且

$$\mathrm{f}(x) - \frac{27}{4} = \frac{4u^3 - 15u^2 + 12u + 4}{4u^2} = \frac{(u-2)^2(4u+1)}{4u^2}.$$

由於 $4u + 1 = 4x^2 - 4x + 1 = (2x-1)^2 \geq 0$，此式永不為負，而它為零當且僅當 $u = 2$（即 $x^2 - x - 2 = 0$，所以 $x = 2$ 或 $-1$）或 $x = \frac{1}{2}$。

? **想一想** 怎樣才能不解六次方程，就猜出 $\mathrm{f}(x) = \frac{343}{36}$ 的一個根？

**方程** $\mathrm{f}(x) = \frac{343}{36}$。注意 $\frac{343}{36} = \frac{7^3}{6^2}$，所以我們尋找滿足 $x^2 - x + 1 = 7$ 和 $x^2 - x = 6$ 的 $x$。$x = 3$ 兩者皆滿足，所以 $\mathrm{f}(3) = \frac{343}{36}$。對稱性現在免費送上更多的根：對 3 反覆施行 $x \mapsto 1 - x$ 和 $x \mapsto 1/x$，得到

$$3, \quad 1 - 3 = -2, \quad \tfrac{1}{3}, \quad 1 - \tfrac{1}{3} = \tfrac{2}{3}, \quad \tfrac{3}{2}, \quad 1 - \tfrac{3}{2} = -\tfrac{1}{2}.$$

這是六個不同的數，而去分母後，$\mathrm{f}(x) = \frac{343}{36}$ 會變成一個六次多項式方程，所以這些就是全部的根：

$$x = -2, -\tfrac{1}{2}, \tfrac{1}{3}, \tfrac{2}{3}, \tfrac{3}{2}, 3.$$

（以 $u$ 表示，方程為 $36u^3 - 235u^2 + 108u + 36 = 0$，可因式分解為 $(u-6)(9u+2)(4u-3) = 0$，得出數對 $\{3,-2\}$、$\{\frac{1}{3},\frac{2}{3}\}$ 和 $\{\frac{3}{2},-\frac{1}{2}\}$。）

由於 $\frac{343}{36} > \frac{27}{4}$，水平線 $y = \frac{343}{36}$ 與每一支 U 形曲線各相交兩次，在最小值兩側各一次，而在這兩個交點以外，曲線位於直線上方。$(-\infty, 0)$ 上的一支交於 $-2$ 和 $-\frac{1}{2}$，$(0,1)$ 上的一支交於 $\frac{1}{3}$ 和 $\frac{2}{3}$，$(1,\infty)$ 上的一支交於 $\frac{3}{2}$ 和 3。因此 $\mathrm{f}(x) > \frac{343}{36}$ 當且僅當

$$x < -2, \quad -\tfrac{1}{2} < x < 0, \quad 0 < x < \tfrac{1}{3}, \quad \tfrac{2}{3} < x < 1, \quad 1 < x < \tfrac{3}{2}, \quad x > 3.$$

## 驗算

- 試代一個值：$\mathrm{f}(-\frac{1}{2}) = \dfrac{(7/4)^3}{(3/4)^2} = \dfrac{343/64}{9/16} = \dfrac{343}{36}$，與所述相符。
- 取所述區間內的一點：$\mathrm{f}(4) = \frac{13^3}{12^2} \approx 15.3 > \frac{343}{36} \approx 9.53$，而 $\mathrm{f}(2) = 6.75$ 位於下方，與 $(\frac{3}{2}, 3)$ 被排除一致。
- 六個根既可配成和為 1 的數對，又可配成互為倒數的數對，正如兩個對稱性所要求；以電腦在 $10^5$ 個隨機點上檢驗，確認了這些不等式。

## 常見錯誤

- 對 $\mathrm{f}(1/x)$ 求導時，忘記鏈式法則帶來的因子 $-1/x^2$。這在 $x = -1$ 處無關重要，但要明白為何 2 處的斜率為零，這個因子不可或缺。
- 把範圍寫成 $x < -2$ 或 $x > 3$ 之類的單一區間，忘記了 $x = 0$ 和 $x = 1$ 處的漸近線把解集分割開來；必須排除 $x = 0$ 和 $x = 1$。
- 把駐點當作極大值點。每一支在兩端都趨向 $+\infty$，所以它們是極小值點。

## STEP 之外

由 $x \mapsto 1 - x$ 和 $x \mapsto 1/x$ 生成的六個映射構成一個羣（group）（同構（isomorphic）於三角形的對稱羣），而 f 在相差一個常數因子的意義下，就是交比（cross-ratio）的經典 $j$ 不變量（j-invariant）：即在全部六個代換下都保持不變的函數。在一組代換下不變的函數，是伽羅瓦理論（Galois theory）和橢圓曲線（elliptic curve）理論的核心主題。Allenby 的 *Rings, Fields and Groups*（《環、域與羣》）是認識這類變換羣的好起點。

詳細解答

# STEP II 2013，第 6 題

純數學 • 數列的收斂

**STEP II 2013，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/13-s2-q6

本題在 *Beyond STEP* 第 2.2 節作爲極限的黃金比例中有完整的引導式講解：https://beyondstep.keithmcnulty.org/zh-hant/ch02/golden-limit.html

## 關鍵思路

各項 $1, 2, \frac{3}{2}, \frac{5}{3}, \frac{8}{5}, \ldots$ 交替地跳到其最終極限（limit）的上方和下方，所以整個數列（sequence）並非單調（monotonic），不能直接套用該定理。然而，每隔一項取一項，便得到兩個單調數列。第 (i) 部分正是顯示這一點的工具：它說明由於分母為正，步長 $u_{n+2} - u_n$ 與上一步 $u_n - u_{n-2}$ 同號。

## 解答

首先注意每個 $u_n$ 都是正數：$u_1 = 1 > 0$，而若 $u_n > 0$，則 $u_{n+1} = 1 + 1/u_n > 1 > 0$。因此我們從不會除以零，而且像 $1 + u_n$ 這樣的因子都是正數。

### 第 (i) 部分

兩次使用 $(*)$，

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

（第二式需要 $u_{n-2}$ 存在，這就是 $n \geq 3$ 的原因。）把第二式代入第一式，兩個負號互相抵消：

$$u_{n+2} - u_n = \frac{u_n - u_{n-2}}{(u_n u_{n+1})(u_{n-2}u_{n-1})}.$$

最後，把 $(*)$ 乘以 $u_n$ 得 $u_n u_{n+1} = u_n + 1$，同理 $u_{n-2}u_{n-1} = u_{n-2} + 1$。因此

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

得證。

### 第 (ii) 部分

我們使用歸納法（induction）。由於 $u_1 = 1$，命題 $1 \le u_1 \le 2$ 成立。假設 $1 \le u_n \le 2$。對正數取倒數（reciprocal）會令不等式反向，所以 $\frac{1}{2} \le 1/u_n \le 1$，因此

$$\tfrac{3}{2} \le u_{n+1} = 1 + \frac{1}{u_n} \le 2,$$

這當然給出 $1 \le u_{n+1} \le 2$。由歸納法，對所有 $n$ 均有 $1 \le u_n \le 2$。

### 第 (iii) 部分

奇數項。我們有 $u_1 = 1$ 及 $u_3 = \frac{3}{2}$，所以 $u_3 - u_1 > 0$。對於奇數 $n \ge 3$，第 (i) 部分說明 $u_{n+2} - u_n$ 等於 $u_n - u_{n-2}$ 乘以正數 $1/\big((1+u_n)(1+u_{n-2})\big)$，所以兩者同號。由歸納法，每一步 $u_{n+2} - u_n$（$n$ 為奇數）都是正數：數列 $u_1, u_3, u_5, \ldots$ 是遞增（increasing）的。由 (ii)，它有上界（bounded above）2，所以由該定理，它趨向某個極限 $L$；由於其所有項都在 $[1,2]$ 內，故 $1 \le L \le 2$。

偶數項。這裏 $u_2 = 2$ 及 $u_4 = \frac{5}{3}$，所以 $u_4 - u_2 < 0$，同樣的論證顯示 $u_2, u_4, u_6, \ldots$ 是遞減（decreasing）的。它有下界 1。

> **想一想** 該定理是關於有上界的遞增數列的。我們怎樣把它用於遞減數列？

把它應用於各項的相反數：$-u_2, -u_4, -u_6, \ldots$ 是遞增的，且有上界 $-1$，所以它趨向某個極限，因此 $u_2, u_4, \ldots$ 亦然；記其極限為 $M$。同樣有 $1 \le M \le 2$。

求極限。使用 $(*)$ 兩步，便得到一條聯繫同奇偶性(parity)各項的遞推關係(recurrence)：

$$u_{n+2} = 1 + \frac{1}{1 + \dfrac{1}{u_n}} = 1 + \frac{u_n}{u_n + 1} = \frac{2u_n + 1}{u_n + 1}.$$

> **想一想** 為甚麼我們可以直接把 $u_{n+2}$ 和 $u_n$ 換成 $L$？

沿着奇數 $n$，$u_n$ 和 $u_{n+2}$ 都趨向 $L$。右方是 $u_n$ 的連續函數（continuous function）（其分母至少為 2），所以它趨向 $(2L+1)/(L+1)$。極限是唯一的，所以

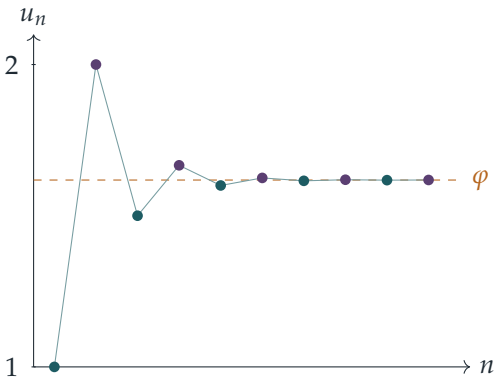
$$L = \frac{2L+1}{L+1} \quad \Longrightarrow \quad L^2 + L = 2L + 1 \quad \Longrightarrow \quad L^2 - L - 1 = 0,$$

於是 $L = \frac{1}{2}(1 \pm \sqrt{5})$。由於 $L \geq 1$，必須有 $L = \frac{1}{2}(1 + \sqrt{5})$。$M$ 滿足完全相同的方程，且 $M \geq 1$，所以

$$\lim_{k\to\infty} u_{2k-1} = \lim_{k\to\infty} u_{2k} = \tfrac{1}{2}(1 + \sqrt{5}).$$

這兩個極限本來有可能不同（那樣整個數列便會振盪），但該方程只有一個至少為 1 的根。

整個數列。記 $\varphi = \frac{1}{2}(1 + \sqrt{5})$，並設 $\varepsilon > 0$。存在 $K_1$ 使得對所有奇數 $n > K_1$ 均有 $|u_n - \varphi| < \varepsilon$，亦存在 $K_2$ 使得對所有偶數 $n > K_2$ 同樣成立。那麼對每個 $n > \max(K_1, K_2)$ 均有 $|u_n - \varphi| < \varepsilon$，所以 $u_n \to \varphi$。




奇數項（青綠色）向 $\varphi$ 攀升，偶數項（梅紫色）則向它下降。

**由 $u_1 = 3$ 開始**

結論不變。現在 $u_1 = 3$，$u_2 = \frac{4}{3}$，$u_3 = \frac{7}{4}$，$u_4 = \frac{11}{7}$。論證只需稍作調整。

- 第 (i) 部分只用到 $(*)$ 及各項為正，這仍然成立，所以該恆等式 (identity) 依然成立。
- 第 (ii) 部分在 $n = 1$ 時不成立，但 $1 \leq u_2 = \frac{4}{3} \leq 2$，而歸納步驟不變，所以對所有 $n \geq 2$ 均有 $1 \leq u_n \leq 2$。另外 $u_1 = 3 \geq 1$，所以每一項都至少為 1。
- 這次 $u_3 - u_1 = -\frac{5}{4} < 0$ 而 $u_4 - u_2 = \frac{5}{21} > 0$。所以奇數項遞減且有下界 1，偶數項遞增且有上界 2：角色互換了，但兩者仍然收斂 (converge)。
- 遞推關係 $u_{n+2} = (2u_n + 1)/(u_n + 1)$ 仍然成立，而兩個極限都至少為 1，所以兩者同樣都是 $\varphi$。

所以該數列仍然趨向 $\frac{1}{2}(1 + \sqrt{5})$。

## 驗算

當 $u_1 = 1$ 時，各項為 $1, 2, \frac{3}{2}, \frac{5}{3}, \frac{8}{5}, \frac{13}{8}, \ldots$，即相鄰斐波那契數（Fibonacci numbers）之比，而 $u_{10} = \frac{89}{55} \approx 1.61818$，已經非常接近 $\varphi \approx 1.61803$。當 $u_1 = 3$ 時，各項為 $3, \frac{4}{3}, \frac{7}{4}, \frac{11}{7}, \frac{18}{11}, \ldots$（盧卡斯數（Lucas numbers）之比），它們同樣趨近 1.61803。你亦可以用 $n = 3$ 驗算 (i) 中的恆等式：$u_5 - u_3 = \frac{1}{10}$，而 $\dfrac{u_3 - u_1}{(1+u_3)(1+u_1)} = \dfrac{1/2}{5} = \frac{1}{10}$。

## 常見錯誤

- 假設整個數列收斂，然後直接解 $L = 1 + 1/L$。這樣得出的數值是對的，但只有在證明了收斂之後才成立，而這正是本題的重點。
- 忘記了根 $\frac{1}{2}(1 - \sqrt{5})$：必須把它排除，這裏是利用 $L \geq 1$。
- 把該定理應用於遞減的偶數項子數列（subsequence），卻沒有說明怎樣做（例如透過 $-u_{2k}$）。

## STEP 之外

這是大學一年級分析學（analysis）中一個標準論證的範本：單調收斂定理（monotone convergence theorem）、子數列，以及「若奇數項子數列與偶數項子數列收斂到同一極限，則整個數列亦收斂到該極限」這一事實。事實上，$u_n$ 是連分數（continued fraction）$1 + \cfrac{1}{1 + \cfrac{1}{1 + \cdots}}$ 的第 $n$ 個漸近分數（convergent），而映射 $x \mapsto 1 + 1/x$ 在 $\varphi$ 附近是一個壓縮映射（contraction），這就是任何正數起點都會趨向 $\varphi$ 的原因。Abbott 的 *Understanding Analysis*（《理解分析》）仔細討論了單調數列和子數列，而 Davenport 的 *The Higher Arithmetic*（《高等算術》）則涵蓋了連分數。

詳細解答

# STEP II 2013，第 7 題

純數學 • 佩爾方程的解

**STEP II 2013，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/13-s2-q7

## 關鍵思路

形如 $x^2 - 2y^2 = 1$ 的方程稱為佩爾方程（Pell equation）：只要找到一組解，一個固定的線性映射（linear map）便能產生無限多組解。第 (i) 和 (ii) 部分把這些解與同時是平方數的三角形數（triangular number）聯繫起來。第 (iii) 部分的技巧是把 $a^2$ 移到另一邊，再把 $c^4 - a^2$ 按平方差（difference of two squares）因式分解。由於 $b$ 是質數（prime），把 $b^3$ 拆成兩個因數的方法寥寥可數。最後，第 (ii) 部分為最後一部分提供一個非質數的 $b$。

## 解答

### 第 (i) 部分

顯然的解是 $x = 1$、$y = 0$，因為 $1^2 - 2 \cdot 0^2 = 1$。

現設 $p^2 - 2q^2 = 1$。那麼

$$\begin{aligned}(3p + 4q)^2 - 2(2p + 3q)^2 &= (9p^2 + 24pq + 16q^2) - 2(4p^2 + 12pq + 9q^2) \\ &= p^2 - 2q^2 = 1,\end{aligned}$$

因為含 $pq$ 的項互相抵消。所以 $x = 3p + 4q$、$y = 2p + 3q$ 也是一組解，正如所求。

把這個映射作用於 $(p, q) = (1, 0)$ 得 $(3, 2)$，再作用於 $(3, 2)$ 得 $(3 \cdot 3 + 4 \cdot 2,\ 2 \cdot 3 + 3 \cdot 2) = (17, 12)$。驗證：$9 - 8 = 1$ 及 $289 - 288 = 1$。因此，兩組 $x$ 為正奇數、$y$ 為正偶數的解是

$$(x, y) = (3, 2) \quad \text{及} \quad (x, y) = (17, 12).$$

（事實上，這個映射總會保持 $x$ 為奇數，因為 $3p + 4q \equiv p \pmod 2$；而一旦 $y$ 是偶數，

它亦會保持 $y$ 為偶數，因為 $2p+3q \equiv q \pmod 2$。所以這條鏈中往後的每一組解都具有正確的奇偶性（parity）：$(99,70)$、$(577,408)$，如此類推。）

### 第 (ii) 部分

設 $x$ 為奇數、$y$ 為偶數，並寫 $x=2m+1$ 及 $y=2n$，其中 $m$、$n$ 為整數。那麼 $(*)$ 變成

$$(2m+1)^2-2(2n)^2=1 \iff 4m^2+4m=8n^2 \iff n^2=\tfrac{1}{2}m(m+1),$$

最後一步是除以 8。這正是所求的形式。（每一步都是可逆的，所以對這樣的 $x$ 和 $y$，兩條方程是等價的。）

所以 $n^2$ 是三角形數 $1+2+\cdots+m$。由第 (i) 部分，$(3,2)$ 給出 $m=1$、$n=1$（即 $1=1$），而 $(17,12)$ 給出 $m=8$、$n=6$（即 $36=\frac{1}{2}\cdot 8\cdot 9$）。我們會在第 (iii) 部分用到後者。

### 第 (iii) 部分

移項並按平方差因式分解，

$$b^3=c^4-a^2=(c^2-a)(c^2+a).$$

由於 $b^3>0$ 及 $c^2+a>0$，因數 $c^2-a$ 也是正整數。此外 $a>0$，所以 $c^2-a<c^2+a$。

> **想一想** 把 $b^3$ 寫成兩個正整數之積，共有多少種方法？

因為 $b$ 是質數，$b^3$ 的正因數只有 1、$b$、$b^2$ 和 $b^3$。所以把 $b^3$ 分解成較小因數乘以較大因數的方式只有 $1\times b^3$ 和 $b\times b^2$。此外別無其他，而且兩個因數永遠不可能相等，因為 $b^3$ 不是完全平方數（perfect square）。

*情況 1*：$c^2-a=1$ 及 $c^2+a=b^3$。兩式相加及相減，得

$$c^2=\tfrac{1}{2}(b^3+1), \qquad a=\tfrac{1}{2}(b^3-1).$$

*情況 2*：$c^2-a=b$ 及 $c^2+a=b^2$。兩式相加及相減，得

$$c^2=\tfrac{1}{2}b(b+1), \qquad a=\tfrac{1}{2}b(b-1).$$

（題目只要求這些公式，但值得知道哪些情況真的會出現。情況 2 永遠不會出現。當 $b=2$ 時，它需要 $c^2=3$。若 $b$ 是奇質數，則 $b$ 與 $\frac{1}{2}(b+1)$ 互質（coprime），因為 $1\le\frac{1}{2}(b+1)<b$；而它們的積是平方數 $c^2$，所以兩者都必須是平方數。但質數永遠不是平方數。情況 1 則確實會出現：$b=23$ 給出 $\frac{1}{2}(23^3+1)=6084=78^2$，所以 $a=6083$、$b=23$、$c=78$ 是一組解。）

### 最後部分

情況 2 對質數不成立，但若我們放棄 $b$ 必須是質數的要求，它的公式仍然能給出解。每當 $\frac{1}{2}b(b+1)$ 是完全平方數 $c^2$ 時，令 $a=\frac{1}{2}b(b-1)$。那麼

$$a^2+b^3=(c^2-b)^2+b^3=c^4-2bc^2+b^2+b^3=c^4-b^2(b+1)+b^2+b^3=c^4,$$

這裏用了 $2c^2=b(b+1)$。第 (ii) 部分正好告訴我們 $\frac{1}{2}b(b+1)$ 何時是平方數：$b=m$，其中 $n^2=\frac{1}{2}m(m+1)$。解 $m=1$ 給出 $a=0$，這是不容許的，但 $m=8$、$n=6$ 可行。取 $b=8$、$c=6$、$a=\frac{1}{2}\cdot 8\cdot 7=28$：

$$28^2+8^3=784+512=1296=6^4.$$

所以 $a=28$、$b=8$、$c=6$ 是一組解，其中 $b=2^3$ 不是質數。

## 驗算

- 鏈中的下一組解 $(99,70)$ 滿足 $9801-2\cdot 4900=1$，它給出 $m=49$、$n=35$，且 $\frac{1}{2}\cdot 49\cdot 50=1225=35^2$。
- $6^4=1296$ 及 $28^2+8^3=784+512=1296$。
- 以電腦搜尋所有 $c<400$，發現 $a=6083$、$b=23$、$c=78$ 是 $b$ 為質數時的唯一解。它屬於情況 1 的解：$78^4-6083^2=(78^2-6083)(78^2+6083)=1\times 12167=23^3$。

## 常見錯誤

- 忘記 $c^2-a$ 必須是較小的因數，因而多出一些 $a$ 為負數的「情況」。
- 漏掉分解 $1\times b^3$，只找到 $b\times b^2$ 的情況，或者反過來。
- 在最後部分使用由 $(3,2)$ 得出的 $m=1$，這會給出 $a=0$；但題目要求 $a$ 為正數。

## STEP 之外

方程 $x^2-2y^2=1$ 就是佩爾方程。它的解恰好就是冪 $(3+2\sqrt{2})^k=x+y\sqrt{2}$，而第 (i) 部分的映射只不過是乘以 $3+2\sqrt{2}$。這是環（ring）如 $\mathbb{Z}[\sqrt{2}]$ 中的單位（unit）理論以及連分數（continued fraction）理論的起點。像 $a^2+b^3=c^4$ 這樣的方程屬於廣義費馬方程（generalised Fermat equation）的研究範疇。若想讀一本精彩的入門書，涵蓋以上種種，可閱讀 Davenport 的 *The Higher Arithmetic*（《高等算術》）。

詳細解答

# STEP II 2013，第 8 題

純數學 • 微積分與不等式

**STEP II 2013，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/13-s2-q8

## 關鍵思路

由於 f 是遞減（decreasing）函數，這個區域是一道越往右越低的「樓梯」，所以最佳的矩形會位於左下角：底邊在直線 $y = \mathrm{f}(t)$ 上，左邊貼着 $y$ 軸，而右上角落在曲線上。這樣第 (i) 部分就化為求單變量函數的最大值（maximum）。第 (ii) 部分比較兩個面積，即整個區域與其內的矩形；而第 (iii) 部分只是把 $\mathrm{f}(x) = 1/(1+x)$ 代入，把一切具體算出來。

## 解答

### 第 (i) 部分

該區域是 $\{(x,y) : 0 \le x \le t,\ \mathrm{f}(t) \le y \le \mathrm{f}(x)\}$：由於 f 遞減，當 $x < t$ 時曲線位於直線 $y = \mathrm{f}(t)$ 之上。

取區域內任意一個矩形 $[a,b] \times [c,d]$。它的底邊必須滿足 $c \ge \mathrm{f}(t)$，左邊滿足 $a \ge 0$，而右上角 $(b,d)$ 必須在區域內，所以 $b \le t$ 及 $d \le \mathrm{f}(b)$。因此其面積至多為

$$b\big(\mathrm{f}(b) - \mathrm{f}(t)\big),$$

而矩形 $[0,b] \times [\mathrm{f}(t), \mathrm{f}(b)]$ 正好達到這個面積，並且確實放得進區域：對於 $0 \le x \le b$，有 $\mathrm{f}(x) \ge \mathrm{f}(b)$，所以該矩形位於曲線之下。於是

$$A_0(t) = \max_{0 \le x \le t} R(x), \qquad R(x) = x\big(\mathrm{f}(x) - \mathrm{f}(t)\big).$$

由於 $R(0) = R(t) = 0$，而對於 $0 < x < t$ 有 $R(x) > 0$，所以當 $t > 0$ 時，最大值在某個內點 $x_0 \in (0,t)$ 取得，而該點滿足 $R'(x_0) = 0$：

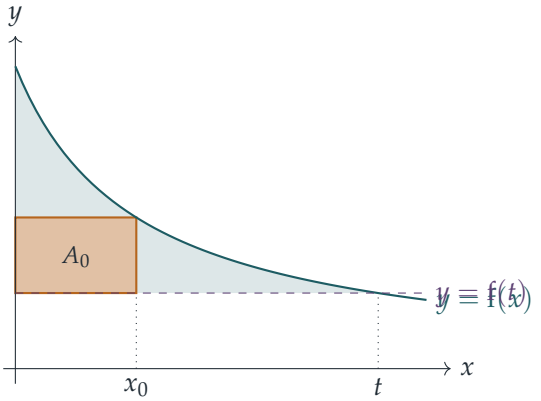
$$R'(x) = \mathrm{f}(x) + x\mathrm{f}'(x) - \mathrm{f}(t) = 0 \quad \Longrightarrow \quad x_0\mathrm{f}'(x_0) + \mathrm{f}(x_0) = \mathrm{f}(t).$$

因此 $A_0(t) = x_0\big(\mathrm{f}(x_0) - \mathrm{f}(t)\big)$，其中 $x_0$ 滿足上述條件，得證。（當 $t = 0$ 時，區域沒有寬度，$A_0(0) = 0$，而 $x_0 = 0$ 顯然滿足該條件。）

**第 (ii) 部分**

記 $t\,\mathrm{g}(t) = \displaystyle\int_0^t \mathrm{f}(x)\,\mathrm{d}x$。兩邊求導，左邊用乘積法則（product rule），右邊用微積分基本定理（fundamental theorem of calculus），得

$$\mathrm{g}(t) + t\,\mathrm{g}'(t) = \mathrm{f}(t) \quad \Longrightarrow \quad t\,\mathrm{g}'(t) = \mathrm{f}(t) - \mathrm{g}(t).$$




積分 $\int_0^t \big(\mathrm{f}(x) - \mathrm{f}(t)\big)\,\mathrm{d}x$ 正是陰影區域的面積，即 $0 \le x \le t$ 範圍內曲線與直線 $y = \mathrm{f}(t)$ 之間的面積。第 (i) 部分的最佳矩形位於這個區域之內，但並未填滿它：當 $x_0 < x < t$ 時，矩形沒有任何貢獻，但由於 f 嚴格遞減（strictly decreasing），區域在該處的高度為 $\mathrm{f}(x) - \mathrm{f}(t) > 0$。所以那一窄條的面積是正數，因此

$$\int_0^t \big(\mathrm{f}(x) - \mathrm{f}(t)\big)\,\mathrm{d}x > A_0(t).$$

**想一想** 怎樣把左邊化成與 $\mathrm{g}'$ 有關的式子？

把積分拆開，利用 g 的定義，再用剛證明的結果：

$$\begin{aligned}\int_0^t \big(\mathrm{f}(x) - \mathrm{f}(t)\big)\,\mathrm{d}x &= t\,\mathrm{g}(t) - t\,\mathrm{f}(t)\\ &= -t\big(\mathrm{f}(t) - \mathrm{g}(t)\big)\\ &= -t\cdot t\,\mathrm{g}'(t) = -t^2\mathrm{g}'(t).\end{aligned}$$

因此 $-t^2\mathrm{g}'(t) > A_0(t)$。

### 第 (iii) 部分

這裏 $\mathrm{f}(x) = \dfrac{1}{1+x}$ 在 $x \geq 0$ 時為正且嚴格遞減，所以以上所有結果都適用。

*左邊*。由第 (ii) 部分最後的計算，

$$-t^2\mathrm{g}'(t) = \int_0^t \frac{\mathrm{d}x}{1+x} - \frac{t}{1+t} = \ln(1+t) - \frac{t}{1+t}.$$

*矩形*。由 $\mathrm{f}'(x) = -\dfrac{1}{(1+x)^2}$，

$$x\mathrm{f}'(x) + \mathrm{f}(x) = \frac{-x + (1+x)}{(1+x)^2} = \frac{1}{(1+x)^2},$$

所以 $x_0$ 的條件是 $(1+x_0)^2 = 1+t$。記 $s = \sqrt{1+t}$，則 $s > 1$ 且 $x_0 = s-1$（取正根 (root)；留意 $0 < x_0 < t$）。於是 $1+t = s^2$，而

$$A_0(t) = (s-1)\left(\frac{1}{s} - \frac{1}{s^2}\right) = \frac{(s-1)^2}{s^2} = \left(1 - \frac{1}{s}\right)^2.$$

**想一想** 兩邊看起來都很繁複：自然的變量是甚麼？

以 $s$ 表示，一切都簡潔得多。我們有 $\ln(1+t) = 2\ln s$ 及 $\dfrac{t}{1+t} = 1 - \dfrac{1}{s^2}$，所以 $-t^2\mathrm{g}'(t) > A_0(t)$ 變成

$$\begin{aligned} 2\ln s - 1 + \frac{1}{s^2} &> 1 - \frac{2}{s} + \frac{1}{s^2} \\ \iff \quad 2\ln s &> 2 - \frac{2}{s} \\ \iff \quad \ln s &> 1 - \frac{1}{s}. \end{aligned}$$

由於 $s = \sqrt{1+t}$ 及 $\ln s = \ln\sqrt{1+t}$，上式即

$$\ln\sqrt{1+t} > 1 - \frac{1}{\sqrt{1+t}} \qquad (t > 0),$$

得證。

## 驗算

- 取 $t = 3$：$s = 2$，$x_0 = 1$，$A_0 = \frac{1}{4}$，而區域面積為 $\ln 4 - \frac{3}{4} \approx 0.636 > 0.25$。最終的不等式為 $\ln 2 \approx 0.693 > 0.5$。
- 以電腦對多個 $t$ 值直接搜尋矩形面積 $x\big(\mathrm{f}(x) - \mathrm{f}(t)\big)$，確認最大值為 $(1 - 1/\sqrt{1+t})^2$，在 $x = \sqrt{1+t} - 1$ 取得。

- 當 $t \to 0^+$ 時，最終不等式兩邊都趨向 0；展開可得 $\ln s \approx (s-1) - \frac{1}{2}(s-1)^2$ 及 $1 - \frac{1}{s} \approx (s-1) - (s-1)^2$，所以左邊大約多出 $\frac{1}{2}(s-1)^2$，正合預期。

## 常見錯誤

- 不加說明便假設最佳矩形的角落在曲線上：這需要用到遞減的性質，如第 (i) 部分所示。
- 忘記解釋為何面積不等式是嚴格的：矩形漏掉了區域中 $x_0 < x < t$ 的部分。
- 取了負根 $1 + x_0 = -\sqrt{1+t}$，這會得出 $x_0 < 0$。

## STEP 之外

最終的不等式，即對 $s > 1$ 有 $\ln s > 1 - 1/s$，是經典界限 $1 - \frac{1}{s} < \ln s < s - 1$ 的其中一半；這組界限來自把 $\int_1^s \mathrm{d}u/u$ 與矩形作比較，而本題則以一條更巧妙的幾何途徑得到它。把積分與矩形面積作比較，正是黎曼（Riemann）積分的核心；大學一年級的分析課程會由上和與下和（upper and lower sums）出發，仔細建立這套理論。Abbott 的 *Understanding Analysis*（《理解分析》）對此有非常易讀的闡述。

詳細解答

# STEP II 2013，第 9 題

力學 • 平衡中的摩擦力

> **STEP II 2013，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/13-s2-q9

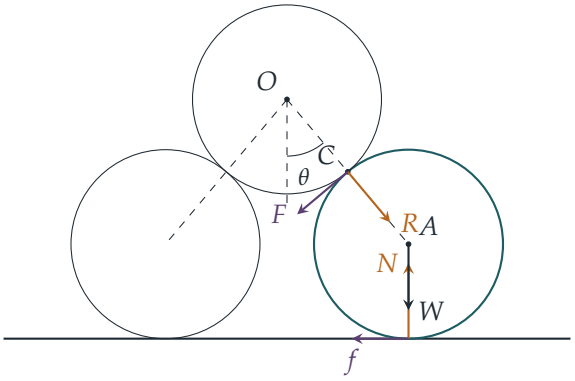
## 關鍵思路

大部分工作都由對稱性（symmetry）完成。把三個圓盤當作一個整體系統，立即就能得出地面的反作用力。然後對其中一個底部圓盤的圓心取力矩（moments）：地面的摩擦力（friction）和頂部圓盤施加的摩擦力都沿切線（tangent）方向作用，因此兩者的力臂（lever arm）相同（都是半徑），大小必然相等。知道這一點後，分別對底部圓盤和頂部圓盤分解力（resolving forces），便得到兩條方程，可解出兩個未知數。

## 解答

我們把各圓盤視為全同的均勻剛體（rigid body）圓盤，半徑為 $r$，重量各為 $W$，全部位於同一鉛直平面內。設 $O$ 為頂部圓盤的圓心，$A$ 為右邊底部圓盤的圓心，則 $OA = 2r$，且 $OA$ 與向下的鉛直線成角 $\theta$。這兩個圓盤的接觸點 $C$ 是 $OA$ 的中點。

由左右對稱可知，左邊底部圓盤所受的力與右邊圓盤所受的力互為鏡像，所以我們只需處理右邊的圓盤。在 $C$ 點，右邊圓盤受到沿 $OA$ 方向的法向反作用力（normal reaction）$R$（把它推離 $O$），以及沿公切線方向的摩擦力 $F$。在地面上，它受到法向反作用力 $N$ 和摩擦力 $f$。




圖中顯示右邊底部圓盤所受的力。根據牛頓（Newton）第三定律（Newton's third law），頂部圓盤在 $C$ 點受到方向相反的 $R$ 和 $F$，因此摩擦力沿切線把頂部圓盤向上及向外推。

### 第 (i) 部分

把三個圓盤合起來視為一個系統。圓盤之間的力是內力，會互相抵消，所以僅有的外力是三個重量和地面施加的力。由對稱性，兩個地面反作用力相等，而地面的摩擦力是水平的，因此鉛直方向分解力得

$$2N = 3W \quad \Longrightarrow \quad N = \tfrac{3}{2}W,$$

得證。

### 第 (ii) 部分

**想一想** 底部圓盤上似乎有三個未知數 $R$、$F$ 和 $f$。第三條方程從何而來？

來自力矩。對右邊底部圓盤關於 $A$ 取力矩。重量 $W$、法向反作用力 $N$ 和法向反作用力 $R$ 都通過 $A$，所以沒有力矩。摩擦力 $f$ 沿最低點的切線作用，$F$ 沿 $C$ 點的切線作用，所以兩者的力臂都是 $r$。它們必須使圓盤向相反方向轉動，因此

$$fr = Fr \quad \Longrightarrow \quad f = F.$$

（這也確定了圖中各力的方向。圓盤被向外推，所以 $f$ 指向內，於是底部圓盤上的 $F$ 必須沿切線指向下方及內側，才能平衡 $f$ 的力矩。）

由 $O$ 到 $A$ 的單位向量是 $(\sin\theta, -\cos\theta)$，而底部圓盤上 $F$ 的切線方向是 $(-\cos\theta, -\sin\theta)$。對底部圓盤水平方向分解力：

$$R\sin\theta - F\cos\theta - f = 0 \quad \Longrightarrow \quad R\sin\theta = F(1+\cos\theta),$$

其中用了 $f = F$。

對頂部圓盤而言，兩個接觸點提供相等的向上分量，各為 $R\cos\theta + F\sin\theta$。鉛直方向分解力：

$$2(R\cos\theta + F\sin\theta) = W.$$

代入 $F = \dfrac{R\sin\theta}{1+\cos\theta}$：

$$W = 2R\left(\cos\theta + \frac{\sin^2\theta}{1+\cos\theta}\right) = 2R\cdot\frac{\cos\theta+\cos^2\theta+\sin^2\theta}{1+\cos\theta} = 2R\cdot\frac{1+\cos\theta}{1+\cos\theta} = 2R.$$

所以兩個互相接觸的圓盤之間的法向反作用力是

$$R = \tfrac{1}{2}W,$$

且與 $\theta$ 無關。於是

$$F = \frac{R\sin\theta}{1+\cos\theta} = \frac{W\sin\theta}{2(1+\cos\theta)},$$

這就是所求的摩擦力。

作為一致性檢查，對底部圓盤鉛直方向分解力得 $N = W + R\cos\theta + F\sin\theta = W + \frac{1}{2}W = \frac{3}{2}W$，與第 (i) 部分吻合。

**第 (iii) 部分**

利用半角公式（half-angle formulae）$\sin\theta = 2\sin\frac{\theta}{2}\cos\frac{\theta}{2}$ 及 $1+\cos\theta = 2\cos^2\frac{\theta}{2}$，

$$\frac{F}{R} = \frac{\sin\theta}{1+\cos\theta} = \tan\tfrac{\theta}{2}.$$

要達到平衡 (equilibrium)，每個接觸點的摩擦力都不能超過其極限值 (limiting value)。在圓盤之間，這表示 $F \le \mu R$，即

$$\tan\tfrac{\theta}{2} \le \mu.$$

(在地面上，我們需要 $f \le \mu N$，即 $\frac{1}{2}W\tan\frac{\theta}{2} \le \frac{3}{2}\mu W$。這個條件較弱，所以起決定作用的是圓盤之間的接觸。)

? **想一想** 是甚麼阻止 $\theta$ 任意地小，從而令 $\tan\frac{\theta}{2}$ 變得極小？

兩個底部圓盤並不接觸。它們的圓心相距 $2\cdot 2r\sin\theta = 4r\sin\theta$，而這必須大於 $2r$，所以 $\sin\theta > \frac{1}{2}$。由於 $0 < \theta < \frac{\pi}{2}$，這表示 $\theta > \frac{\pi}{6}$，所以 $\frac{\theta}{2} > \frac{\pi}{12}$。由於 tan 在 $(0, \frac{\pi}{2})$ 上遞增，

$$\tan\tfrac{\theta}{2} > \tan\tfrac{\pi}{12} = \tan(45^\circ - 30^\circ) = \frac{1-\frac{1}{\sqrt{3}}}{1+\frac{1}{\sqrt{3}}} = \frac{\sqrt{3}-1}{\sqrt{3}+1} = \frac{(\sqrt{3}-1)^2}{2} = 2-\sqrt{3}.$$

所以平衡要求 $\mu \ge \tan\frac{\theta}{2} > 2-\sqrt{3}$。若 $\mu < 2-\sqrt{3}$，則沒有任何容許的 $\theta$ 值能滿足此條件，平衡不可能出現。

## 驗算

- 當 $\theta \to 0$ 時，頂部圓盤正正坐在其他圓盤之上，而 $F \to 0$：不需要摩擦力，與預期一致。
- 取任何樣本角度，例如 $\theta = 1$ 弧度，代入 $R = \frac{1}{2}W$、$F = f = \frac{1}{2}W\tan\frac{1}{2}$ 及 $N = \frac{3}{2}W$，每個圓盤的兩個力分量和力矩都平衡。
- $2 - \sqrt{3} \approx 0.2679$，與 $\tan 15^\circ \approx 0.2679$ 吻合。對略大於 $30^\circ$ 的 $\theta$ 作數值掃描，證實 $\mu = 0.2679$ 時平衡不可能，而 $\mu = 0.268$ 時平衡可能。

## 常見錯誤

- 假設圓盤之間的摩擦力為零，或完全忘記它。沒有它的話，底部圓盤的力矩方程會迫使 $f = 0$，那就沒有東西能阻止底部圓盤向兩邊滑開。
- 遺漏了底部圓盤互不接觸這個條件。它是唯一給 $\theta$ 設下限的條件，因此 $2 - \sqrt{3}$ 正是由此而來。
- 弄錯 $F$ 的方向。應利用關於 $A$ 的力矩方程來確定方向，而不是靠猜。

## STEP 之外

這個問題之所以是靜定(statically determinate)的，全靠對稱性和力矩方程；在許多由粗糙物體堆疊而成的情況中，單憑平衡條件無法求出摩擦力，工程師需要為物體的形變建立模型。簡潔的結果 $F/R = \tan\frac{\theta}{2}$ 是摩擦角（friction angle）概念的一個例子：當且僅當總接觸力與法線所成的角不超過 $\arctan\mu$ 時，平衡才可能出現。Kleppner 和 Kolenkow 的 *An Introduction to Mechanics*（《力學導論》）在大學一年級程度上仔細討論了摩擦力和剛體靜力學。

詳細解答

# STEP II 2013，第 10 題

力學 • 拋體運動

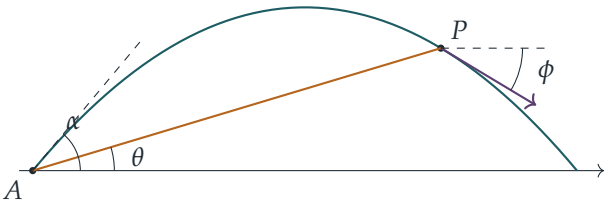
**STEP II 2013，第 10 題**

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/13-s2-q10

## 關鍵思路

寫出時間 $t$ 時的位置和速度，便會發現 $\tan\theta$ 和 $\tan\phi$ 都是 $t$ 的線性（linear）函數。(i) 中的恆等式（identity）隨即得出，而它把 $B$ 和 $C$ 都轉化為關於角的正切的陳述。由於兩個角都隨時間穩定地減小，「哪一點先到」不過是比較兩個數的大小，而 (i) 讓我們只需在 $B$ 這一點上作比較。




在一般的點 $P$，$\theta$ 是從 $A$ 看 $P$ 的仰角（angle of elevation），而 $\phi$ 是速度與水平方向所成的角（圖中為負值，因為質點正在下落）。

## 解答

我們把質點模型化為只受重力 $g$ 作用而自由運動，不計空氣阻力（air resistance）。設質點在時間 $t = 0$ 以速率 $u$ 拋出，其中 $0 < \alpha < 90°$。以 $A$ 為原點，在時間 $t$ 時

$$x = ut\cos\alpha, \qquad y = ut\sin\alpha - \tfrac{1}{2}gt^2, \qquad \dot{x} = u\cos\alpha, \qquad \dot{y} = u\sin\alpha - gt.$$

當 $y = 0$ 再次成立時質點落地，此時 $t = 2u\sin\alpha/g$。

**第 (i) 部分**

當 $t > 0$ 時，從 $A$ 看質點的仰角滿足

$$\tan\theta = \frac{y}{x} = \frac{ut\sin\alpha - \frac{1}{2}gt^2}{ut\cos\alpha} = \tan\alpha - \frac{gt}{2u\cos\alpha},$$

而運動方向滿足

$$\tan\phi = \frac{\dot{y}}{\dot{x}} = \frac{u\sin\alpha - gt}{u\cos\alpha} = \tan\alpha - \frac{gt}{u\cos\alpha}.$$

把第一式乘以 2，再利用第二式，

$$2\tan\theta = 2\tan\alpha - \frac{gt}{u\cos\alpha} = \tan\alpha + \left(\tan\alpha - \frac{gt}{u\cos\alpha}\right) = \tan\alpha + \tan\phi,$$

得證。

值得記下稍後會用到的結果：記 $k = \dfrac{g}{u\cos\alpha} > 0$，

$$\tan\theta = \tan\alpha - \tfrac{1}{2}kt, \qquad \tan\phi = \tan\alpha - kt, \tag{$*$}$$

因此 $\tan\theta$ 和 $\tan\phi$ 都隨時間嚴格遞減。在整個飛行過程中，$\theta$ 由 $\alpha$ 減至 0，$\phi$ 由 $\alpha$ 減至 $-\alpha$，所以 $\theta = \frac{1}{2}\alpha$ 和 $\phi = -\frac{1}{2}\alpha$ 各恰好出現一次：$B$ 和 $C$ 的定義是明確的。

**第 (ii) 部分**

記 $T = \tan\frac{1}{2}\alpha$；由於 $0 < \frac{1}{2}\alpha < 45°$，我們有 $0 < T < 1$。在 $B$，$\tan\theta = T$。在 $C$，$\tan\phi = \tan(-\frac{1}{2}\alpha) = -T$。若 $B$ 和 $C$ 是同一點，則兩式在該點同時成立，由 (i) 得

$$2T = \tan\alpha - T, \qquad \text{即} \qquad \tan\alpha = 3T.$$

由二倍角公式（double-angle formula）$\tan\alpha = \dfrac{2T}{1-T^2}$，故

$$\frac{2T}{1-T^2} = 3T \implies 2 = 3(1-T^2) \implies T^2 = \tfrac{1}{3},$$

其中我們除以了 $T$，這是容許的，因為 $T > 0$。取正根，$\tan\frac{1}{2}\alpha = \frac{1}{\sqrt{3}}$，所以 $\frac{1}{2}\alpha = 30°$，且

$$\alpha = 60°.$$

**第 (iii) 部分**

> ? **想一想** 不計算兩個時間，怎樣判斷哪一點先到達？

由 $(*)$，$\phi$ 在整個飛行過程中遞減。因此 $C$ 在 $B$ 之後到達，當且僅當質點位於 $B$ 的那一刻，$\phi$ 尚未減至 $-\frac{1}{2}\alpha$；即當 $\tan\phi_B > -T$ 時，其中 $\phi_B$ 是質點在 $B$ 的運動方向。

在 $B$，由第 (i) 部分得 $\tan\phi_B = 2\tan\theta - \tan\alpha = 2T - \tan\alpha$。因此

$$B \text{ 先到達} \iff 2T - \tan\alpha > -T \iff \tan\alpha < 3T.$$

再次利用 $\tan\alpha = \dfrac{2T}{1-T^2}$，並除以 $T > 0$、乘以 $1 - T^2 > 0$（這樣不等式的方向不變），

$$\frac{2T}{1-T^2} < 3T \iff 2 < 3 - 3T^2 \iff T^2 < \tfrac{1}{3} \iff \tfrac{1}{2}\alpha < 30^\circ,$$

最後一步成立是因為 tan 在 $(0^\circ, 45^\circ)$ 上遞增。每一步都可逆，所以

$$\alpha < 60^\circ \iff B \text{ 比 } C \text{ 先到達.}$$

所以，既然 $\alpha < 60^\circ$，質點會先到達 $B$。

若你偏好明確寫出時間：由 $(*)$，$kt_B = 2(\tan\alpha - T)$ 及 $kt_C = \tan\alpha + T$，而 $t_B < t_C$ 同樣等價於 $\tan\alpha < 3T$。

## 驗算

- 在 $t = 0$ 時，$(*)$ 中兩條公式都給出 $\tan\alpha$，理應如此：剛拋出時質點沿拋射方向的直線運動。落地時 $kt = 2\tan\alpha$，得 $\tan\theta = 0$ 及 $\tan\phi = -\tan\alpha$，這正是軌跡為人熟知的對稱性。
- 當 $\alpha = 60^\circ$ 時，$\tan\alpha = \sqrt{3} = 3\tan 30^\circ$，與 (ii) 一致。
- 取 $u = 10\,\mathrm{m\,s^{-1}}$、$g = 9.8\,\mathrm{m\,s^{-2}}$ 及 $\alpha = 45^\circ$，公式給出 $t_B \approx 0.845\,\mathrm{s}$ 及 $t_C \approx 1.020\,\mathrm{s}$，所以 $B$ 先到；以電腦逐步數值模擬運動，得出相同的時間，並確認當 $\alpha = 61^\circ$、$75^\circ$ 及 $85^\circ$ 時 $C$ 先到。

## 常見錯誤

- 把 $\theta$（質點*相對於* $A$ 的方向）當作 $\phi$（質點*速度*的方向）。兩者只在拋出的一刻相同。
- 在 $C$ 弄錯符號：該處 $\tan\phi = -\tan\frac{1}{2}\alpha$，而不是 $+\tan\frac{1}{2}\alpha$。
- 在 (iii) 中直接斷言答案，卻沒有解釋為何「$\tan\phi_B > -T$」表示 $B$ 先到。支持這一點的正是 $\phi$ 的單調遞減（monotonic decrease）。

## STEP 之外

(i) 中的恆等式說明：從 $A$ 到拋物線（parabola）上任一點的弦（chord），其斜率等於兩端點切線斜率的平均值，這是二次函數的特殊性質（相當於均值定理（mean value theorem），而且中間點被準確定在正中）。在大學一年級的力學課程中，研究斜面（inclined

plane）上的拋射體（projectile）問題時會遇到同樣的想法，那裏從拋出點出發的直線所成的角，所扮演的角色正與這裏的 $\theta$ 完全相同。Kleppner 和 Kolenkow 的 *An Introduction to Mechanics*（《力學導論》）是很好的參考書，可看到拋射體問題以這種嚴謹方式處理。

詳細解答

# STEP II 2013，第 11 題

力學 • 連續碰撞

**STEP II 2013，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：

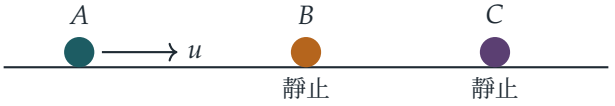
step.maths.org/questions/13-s2-q11

## 關鍵思路

這裏每一次碰撞（collision）都發生在兩個質量相等的質點之間，所以每次都是同一個小計算：動量守恆（conservation of momentum）加上牛頓恢復定律（Newton's law of restitution）。先做一次一般情況，之後重複套用即可。會否再發生碰撞，只取決於後面的質點是否比前面的質點移動得快，因此整道題歸結為比較速度。

## 解答

把這些質點視為在光滑（smooth）平面上、質量均為 $m$ 的相同質點，因此在兩次碰撞之間沒有水平力作用，每個質點都以等速移動。沿直線依次稱它們為 $A$、$B$、$C$，其中 $A$ 是被射向 $B$ 和 $C$ 的那個質點。取 $A$ 初始運動的方向為正方向。




### 一般的碰撞

假設一個速度為 $p$ 的質點追上前方一個速度為 $q < p$ 的相同質點，碰撞後兩者速度分別為 $p'$ 和 $q'$。由動量守恆（約去 $m$）及恢復定律得

$$p' + q' = p + q, \qquad q' - p' = e(p - q).$$

兩式相加及相減，

$$p' = \tfrac{1}{2}\big[(p+q) - e(p-q)\big], \qquad q' = \tfrac{1}{2}\big[(p+q) + e(p-q)\big]. \tag{$*$}$$

由於 $e > 0$ 且 $p > q$，我們總有 $q' > p'$：碰撞後前方的質點比後方的移動得快，所以除非有其他因素改變其中一個的速度，這兩個質點不可能再相遇。

**第 (i) 部分**

**第一次碰撞**（$A$ 撞 $B$）：在 $(*)$ 中代入 $p = u$、$q = 0$：

$$v_A = \tfrac{1}{2}u(1-e), \qquad v_B = \tfrac{1}{2}u(1+e).$$

此時 $v_B > v_A$，所以 $A$ 追不上 $B$；但 $B$ 正向靜止的 $C$ 移動，所以第二次碰撞是 $B$ 撞 $C$。

**第二次碰撞**（$B$ 撞 $C$）：代入 $p = \tfrac{1}{2}u(1+e)$、$q = 0$：

$$v_B = \tfrac{1}{2}\cdot\tfrac{1}{2}u(1+e)(1-e) = \tfrac{1}{4}u(1-e^2), \qquad v_C = \tfrac{1}{4}u(1+e)^2.$$

$A$ 沒有受到影響，所以仍然是 $v_A = \tfrac{1}{2}u(1-e)$。三個速度都是正數（因為 $0 < e < 1$），所以速率分別為 $\tfrac{1}{2}u(1-e)$、$\tfrac{1}{4}u(1-e^2)$、$\tfrac{1}{4}u(1+e)^2$，正如所求。

**第三次碰撞**。$C$ 比 $B$ 快（由 $(*)$），所以 $B$ 和 $C$ 會分開。唯一可能的下一次碰撞是 $A$ 追上 $B$，而這恰好在 $v_A > v_B$ 時發生：

$$v_A - v_B = \tfrac{1}{4}u(1-e)\big[2-(1+e)\big] = \tfrac{1}{4}u(1-e)^2 > 0,$$

因為 $e \neq 1$。所以 $A$ 在 $B$ 後方且移動得比 $B$ 快，對於 $0 < e < 1$ 中的每一個 $e$，第三次碰撞（$A$ 撞 $B$）都會發生。

**第 (ii) 部分**

**第三次碰撞**（$A$ 撞 $B$）：此時 $p = \tfrac{1}{2}u(1-e)$、$q = \tfrac{1}{4}u(1-e^2)$，所以

$$p+q = \tfrac{1}{4}u(1-e)\big[2+(1+e)\big] = \tfrac{1}{4}u(1-e)(3+e), \qquad p-q = \tfrac{1}{4}u(1-e)^2.$$

由 $(*)$，$B$ 的新速度為

$$\begin{aligned} v_B &= \tfrac{1}{2}\Big[\tfrac{1}{4}u(1-e)(3+e) + \tfrac{1}{4}eu(1-e)^2\Big] \\ &= \tfrac{1}{8}u(1-e)\big[3+e+e-e^2\big] \\ &= \tfrac{1}{8}u(1-e)(3+2e-e^2) = \tfrac{1}{8}u(1-e)(1+e)(3-e), \end{aligned}$$

而（為求完整）$v_A = \tfrac{1}{8}u(1-e)\big[3+e-e(1-e)\big] = \tfrac{1}{8}u(1-e)(3+e^2)$，它小於 $v_B$。

**想一想** 第三次碰撞之後，哪些質點對有可能再次碰撞？

$A$ 和 $B$ 剛剛碰撞過，所以由 $(*)$，$B$ 現在比 $A$ 快，兩者正在分開。$C$ 仍以 $\tfrac{1}{4}u(1+e)^2$ 移動，而 $C$ 前方再沒有其他質點。所以第四次碰撞發生若且唯若 $B$ 追上 $C$，即若且唯若

$v_B > v_C$：兩邊除以 $\frac{1}{8}u(1+e) > 0$，

$$\begin{aligned}\tfrac{1}{8}u(1-e)(1+e)(3-e) > \tfrac{1}{4}u(1+e)^2 &\iff (1-e)(3-e) > 2(1+e)\\ &\iff e^2 - 6e + 1 > 0.\end{aligned}$$

二次式 $e^2 - 6e + 1$ 的根為 $e = 3 \pm 2\sqrt{2}$，它在兩根以外取正值。較大的根 $3 + 2\sqrt{2}$ 大於 1，而 $3 - 2\sqrt{2} \approx 0.172$ 位於 $(0, 1)$ 之內。所以，對於 $0 < e < 1$，

$$\text{發生第四次碰撞} \iff e < 3 - 2\sqrt{2}.$$

當 $e = 3 - 2\sqrt{2}$ 時，$B$ 和 $C$ 以相同速度移動，永不相遇，所以不等式（inequality）是嚴格的。

## 驗算

- 動量：第二次碰撞後總動量為 $\frac{1}{4}u\left[2(1-e) + (1-e^2) + (1+e)^2\right] = \frac{1}{4}u \cdot 4 = u$，與預期相符。
- 極限情況：當 $e \to 1$ 時，兩次碰撞後的速度趨向 0、0、$u$（即大家熟悉的「牛頓擺（Newton's cradle）」式速度交換），而 $v_A - v_B \to 0$，所以第三次碰撞僅僅勉強發生；臨界值 $3 - 2\sqrt{2}$ 遠小於 1，與此一致。
- 數值上，取 $e = 0.1$，第三次碰撞後的速度為 $v_B \approx 0.3589u$ 和 $v_C = 0.3025u$，所以 $B$ 確實追上 $C$；取 $e = 0.2$，$v_B = 0.336u < v_C = 0.36u$，則追不上。以電腦模擬這個運動，證實了對每個 $e > 0.1716$ 恰好有三次碰撞，而低於此值則至少有四次。

## 常見錯誤

- 忘記說明下一次碰撞的是哪一對質點。每次都要檢查剛碰撞過的兩個質點正在分開，而餘下的一對正在接近。
- 恢復定律中的符號弄錯：應為（分離速率）$= e \times$（接近速率），所以 $q' - p' = e(p - q)$，左邊先寫前方質點的速度。
- 保留了根 $3 + 2\sqrt{2}$，但它在容許範圍 $0 < e < 1$ 之外。

## STEP 之外

規則 $(*)$ 是一個線性映射（linear map），把碰撞前的一對速度映射到碰撞後的一對速度，所以一連串碰撞就是矩陣（matrix）的乘積；問總共會發生多少次碰撞是一個經典問題，對於彈性（elastic）桌球系統，把運動「展開」成楔形區域中的一條直線，便可得

到一個優美的幾何答案。若想看以大學程度的嚴謹處理碰撞、動量及參考系（frames of reference），可參閱 Kleppner 與 Kolenkow 的 *An Introduction to Mechanics*（《力學導論》）。

詳細解答

# STEP II 2013，第 12 題

概率與統計 • 泊松分佈的矩

> **STEP II 2013，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/13-s2-q12

## 關鍵思路

$X$ 保留 $U$ 的奇數值，$Y$ 則保留偶數值，因此每個期望值（expectation）都是只取奇數項或只取偶數項的泊松（Poisson）和。求 $\mathrm{E}(U)$ 和 $\mathrm{E}(U(U-1))$ 的技巧依然適用：把 $k$ 或 $k(k-1)$ 與 $k!$ 約去，再提出 $\lambda$ 或 $\lambda^2$。$k$ 每減少一，奇偶便互換一次，這正是 $\alpha$ 與 $\beta$ 互換位置的原因。另外留意 $\alpha+\beta=\mathrm{e}^{\lambda}$，所以因子 $\mathrm{e}^{-\lambda}$ 就是 $1/(\alpha+\beta)$。

## 解答

全文中，對 $k=0,1,2,\ldots$，$\mathrm{P}(U=k)=\mathrm{e}^{-\lambda}\lambda^k/k!$，其中 $\lambda>0$。把兩個級數（series）逐項相加，得

$$\alpha+\beta=\sum_{k\geq 0}\frac{\lambda^k}{k!}=\mathrm{e}^{\lambda},\qquad \text{所以}\qquad \mathrm{e}^{-\lambda}=\frac{1}{\alpha+\beta}.$$

（事實上 $\alpha=\cosh\lambda$，而 $\beta=\sinh\lambda$。）

### 第 (i) 部分

除非 $U$ 是奇數，否則 $X$ 為 0，所以只有奇數 $k$ 有貢獻：

$$\mathrm{E}(X)=\sum_{k\text{ 為奇數}}k\frac{\mathrm{e}^{-\lambda}\lambda^k}{k!}=\mathrm{e}^{-\lambda}\lambda\sum_{k\text{ 為奇數}}\frac{\lambda^{k-1}}{(k-1)!}=\mathrm{e}^{-\lambda}\lambda\left(1+\frac{\lambda^2}{2!}+\frac{\lambda^4}{4!}+\cdots\right),$$

因為 $k-1$ 取遍 $0,2,4,\ldots$。所以

$$\mathrm{E}(X)=\frac{\lambda\alpha}{\alpha+\beta}.$$

同理，$Y$ 取的是偶數 $k \geq 2$（$k = 0$ 那一項本來就是零），此時 $k-1$ 取遍 $1, 3, 5, \ldots$：

$$\mathrm{E}(Y) = \mathrm{e}^{-\lambda}\lambda \sum_{k \text{ 為偶數}, k\geq 2} \frac{\lambda^{k-1}}{(k-1)!} = \frac{\lambda\beta}{\alpha+\beta}.$$

驗證一下：$\mathrm{E}(X) + \mathrm{E}(Y) = \lambda = \mathrm{E}(U)$，理應如此，因為 $X + Y = U$。

**第 (ii) 部分**

**想一想** 當 $k^2$ 不能與 $k!$ 乾淨地約去時，怎樣求 $k^2\lambda^k/k!$ 的和？

把 $k^2 = k(k-1) + k$ 寫出來。$k(k-1)$ 部分約去 $k!$ 的兩個因子，而 $k$ 部分約去一個。對奇數 $k$，$k(k-1)$ 的和由 $k = 3$ 開始（在 $k = 1$ 時該項為零），而 $k-2$ 取遍 $1, 3, 5, \ldots$：

$$\sum_{k \text{ 為奇數}} k(k-1)\frac{\lambda^k}{k!} = \lambda^2 \sum_{k \text{ 為奇數}, k\geq 3} \frac{\lambda^{k-2}}{(k-2)!} = \lambda^2\beta, \qquad \sum_{k \text{ 為奇數}} k\frac{\lambda^k}{k!} = \lambda\alpha$$

（第二個和就是第 (i) 部分中的那個）。因此

$$\mathrm{E}(X^2) = \mathrm{e}^{-\lambda}\left(\lambda\alpha + \lambda^2\beta\right) = \frac{\lambda\alpha + \lambda^2\beta}{\alpha+\beta},$$

再減去 $\mathrm{E}(X)^2$，得

$$\mathrm{Var}(X) = \frac{\lambda\alpha + \lambda^2\beta}{\alpha+\beta} - \frac{\lambda^2\alpha^2}{(\alpha+\beta)^2},$$

正如所求。

對 $Y$ 而言，奇偶性（parity）互換。對偶數 $k$，$k(k-1)$ 的和由 $k = 2$ 開始，$k-2$ 取遍 $0, 2, 4, \ldots$，結果是 $\lambda^2\alpha$。$k$ 的和則如第 (i) 部分一樣給出 $\lambda\beta$。所以

$$\mathrm{Var}(Y) = \frac{\lambda\beta + \lambda^2\alpha}{\alpha+\beta} - \frac{\lambda^2\beta^2}{(\alpha+\beta)^2}.$$

**最後一問**。 由於 $X + Y = U$，我們有 $\mathrm{Var}(X+Y) = \mathrm{Var}(U) = \lambda$。把兩個方差（variance）相加，

$$\begin{aligned}\mathrm{Var}(X) + \mathrm{Var}(Y) &= \frac{\lambda(\alpha+\beta) + \lambda^2(\alpha+\beta)}{\alpha+\beta} - \frac{\lambda^2(\alpha^2+\beta^2)}{(\alpha+\beta)^2} \\ &= \lambda + \lambda^2\frac{(\alpha+\beta)^2 - \alpha^2 - \beta^2}{(\alpha+\beta)^2} = \lambda + \frac{2\lambda^2\alpha\beta}{(\alpha+\beta)^2}.\end{aligned}$$

**想一想** 為甚麼我們應該預期兩邊不相等？

一般而言，$\mathrm{Var}(X+Y) = \mathrm{Var}(X) + \mathrm{Var}(Y) + 2\,\mathrm{Cov}(X,Y)$。此處 $XY = 0$ 恆成立，因為 $X$、$Y$ 之中最多只有一個非零。所以 $\mathrm{Cov}(X,Y) = -\mathrm{E}(X)\mathrm{E}(Y) = -\lambda^2\alpha\beta/(\alpha+\beta)^2$，與

上面的計算吻合。要等式成立，便需要 $\lambda^2\alpha\beta = 0$。然而對 $\lambda > 0$，有 $\alpha \geq 1$ 及 $\beta \geq \lambda > 0$，所以 $\lambda^2\alpha\beta > 0$。因此**不存在非零的 $\lambda$ 值**使 $\mathrm{Var}(X) + \mathrm{Var}(Y) = \mathrm{Var}(X+Y)$：左邊總是嚴格較大。事實上 $\alpha = \cosh\lambda$，$\beta = \sinh\lambda$，所以多出的部分是 $\lambda^2 \sinh 2\lambda / \mathrm{e}^{2\lambda} = \frac{1}{2}\lambda^2\left(1 - \mathrm{e}^{-4\lambda}\right)$。

## 驗算

- 當 $\lambda \to 0$ 時，$U$ 幾乎總是 0 或 1，所以 $Y$ 幾乎總是 0。的確，$\mathrm{E}(Y) \approx \lambda^2 \to 0$，而 $\mathrm{E}(X) \approx \lambda$。
- 當 $\lambda$ 很大時，$\alpha \approx \beta \approx \frac{1}{2}\mathrm{e}^{\lambda}$，所以 $\mathrm{E}(X) \approx \mathrm{E}(Y) \approx \frac{1}{2}\lambda$：奇數值與偶數值變得同樣可能出現。
- 在 $\lambda = 1$ 時，直接把泊松概率相加，得 $\mathrm{Var}(X) \approx 0.6778$ 及 $\mathrm{Var}(Y) \approx 0.8131$，所以兩者之和比 $\mathrm{Var}(U) = 1$ 多出約 $0.4908 = \frac{1}{2}(1 - \mathrm{e}^{-4})$。兩條公式都吻合。

## 常見錯誤

- 約去之後弄錯奇偶：若 $k$ 是奇數，則 $k-1$ 是偶數（給出 $\alpha$），而 $k-2$ 是奇數（給出 $\beta$）。
- 忘記 $\mathrm{e}^{-\lambda} = 1/(\alpha+\beta)$，以致答案不是題目要求的形式。
- 假設方差可以直接相加。$X$ 和 $Y$ 遠非獨立(independent)：知道 $X \neq 0$ 便迫使 $Y = 0$。

## STEP 之外

把一個級數分拆成偶數部分和奇數部分，就是分解式 $\mathrm{e}^{\lambda} = \cosh\lambda + \sinh\lambda$。借助概率母函數(probability generating functions)，這變成例行手續：$G(s)$ 的偶數部分是 $\frac{1}{2}\big(G(s) + G(-s)\big)$。最後用到的協方差（covariance）恆等式，是大學概率課程最早接觸的工具之一。Grimmett 與 Welsh 合著的 *Probability: An Introduction*（《概率導論》）同時涵蓋母函數和協方差。

詳細解答

# STEP II 2013，第 13 題

概率與統計 • 擲硬幣中的游程

> **STEP II 2013，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/13-s2-q13

## 關鍵思路

事件 $S = n$ 完全確定了首 $n+1$ 次擲幣的結果：先是 $n$ 次相同的結果，然後一次打斷這段連串（run）的結果。$A = n$ 亦是如此：先是 $n$ 次交替（alternating）出現的結果，然後一次重複。因此每個概率（probability）都是兩個乘積之和，各對應第一次擲幣的一種可能結果。比較 $S$ 與 $A$ 便歸結為將四項之差因式分解（factorisation）；若我們把 $p$ 和 $q$ 當作兩個獨立的字母，而不是以 $1-p$ 代替 $q$，因式分解會較容易看出來。

## 解答

全文中，各次擲幣互相獨立（independent），每次得正面的概率為 $p$，得反面的概率為 $q$，且 $p, q > 0$，$p \neq q$。

### 一般公式

$S = n$ 發生，當且僅當第 $1, \ldots, n$ 次擲幣全部與第一次相同，而第 $n+1$ 次不同。若第一次是正面，結果便是 $\mathrm{H}^n\mathrm{T}$，概率為 $p^n q$；若第一次是反面，結果便是 $\mathrm{T}^n\mathrm{H}$，概率為 $q^n p$。這兩種情況互斥（disjoint），所以

$$\mathrm{P}(S = n) = p^n q + q^n p = pq\left(p^{n-1} + q^{n-1}\right).$$

$A = n$ 發生，當且僅當第 $1, \ldots, n$ 次擲幣交替出現，而第 $n+1$ 次與第 $n$ 次相同，從而結束交替。

- 當 $n = 2m$（偶數）：連串 $\mathrm{HTHT}\cdots\mathrm{HT}$ 有 $m$ 個正面和 $m$ 個反面，並以 T 結尾，所以

下一次必須是 T，得 $p^m q^m \cdot q$。以反面開始則得 $q^m p^m \cdot p$。因此

$$\mathrm{P}(A = 2m) = p^m q^m (p + q).$$

- 當 $n = 2m + 1$（奇數）：連串 HTH$\cdots$H 有 $m + 1$ 個正面和 $m$ 個反面，並以 H 結尾，所以下一次必須是 H，得 $p^{m+2} q^m$。以反面開始則得 $q^{m+2} p^m$。因此

$$\mathrm{P}(A = 2m + 1) = p^m q^m \left(p^2 + q^2\right).$$

**第 (i) 部分**

$A = 1$ 表示第二次擲幣重複第一次的結果：HH 或 TT。所以 $\mathrm{P}(A = 1) = p^2 + q^2$（即上面 $m = 0$ 的情況）。$S = 1$ 表示第二次擲幣與第一次不同：HT 或 TH，所以

$$\mathrm{P}(S = 1) = pq + qp = 2pq.$$

於是

$$\mathrm{P}(A = 1) - \mathrm{P}(S = 1) = p^2 - 2pq + q^2 = (p - q)^2 > 0,$$

由於 $p \neq q$，不等號是嚴格的。所以 $\mathrm{P}(S = 1) < \mathrm{P}(A = 1)$。

**第 (ii) 部分**

由公式，

$$\mathrm{P}(S = 2) = p^2 q + q^2 p = pq(p + q) = \mathrm{P}(A = 2),$$

此即所求。至於長度 3，

$$\mathrm{P}(S = 3) = p^3 q + q^3 p = pq\left(p^2 + q^2\right) = \mathrm{P}(A = 3),$$

所以這兩者亦相等。

**第 (iii) 部分**

**偶數長度**。設 $n > 1$。將兩個概率都除以 $pq > 0$，

$$\frac{\mathrm{P}(S = 2n)}{pq} = p^{2n-1} + q^{2n-1}, \qquad \frac{\mathrm{P}(A = 2n)}{pq} = p^{n-1} q^{n-1}(p + q) = p^n q^{n-1} + p^{n-1} q^n.$$

? **想一想** 怎樣才能判斷 $p^{2n-1} + q^{2n-1} - p^n q^{n-1} - p^{n-1} q^n$ 的正負號？

將每個純冪與含同一字母最高次冪的交叉項配對：

$$\begin{aligned} p^{2n-1} + q^{2n-1} - p^n q^{n-1} - p^{n-1} q^n &= p^n\left(p^{n-1} - q^{n-1}\right) - q^n\left(p^{n-1} - q^{n-1}\right) \\ &= \left(p^n - q^n\right)\left(p^{n-1} - q^{n-1}\right). \end{aligned}$$

由於 $p,q>0$ 且 $p\neq q$，對任何 $k\geq 1$，$p^k-q^k$ 均非零，且與 $p-q$ 同號（因為 $x\mapsto x^k$ 在 $x>0$ 上嚴格遞增（strictly increasing））。這裏 $n\geq 2$，所以 $n$ 和 $n-1$ 都至少是 1，兩個括號均非零且同號。它們的乘積為正，再乘回 $pq>0$，得

$$\mathrm{P}(S=2n)>\mathrm{P}(A=2n)\qquad (n>1).$$

（當 $n=1$ 時，第二個括號是 $p^0-q^0=0$，這正是第 (ii) 部分的等式。）

**奇數長度**。同樣地，仍設 $n>1$，

$$\frac{\mathrm{P}(S=2n+1)}{pq}=p^{2n}+q^{2n},$$

$$\frac{\mathrm{P}(A=2n+1)}{pq}=p^{n-1}q^{n-1}\left(p^2+q^2\right)=p^{n+1}q^{n-1}+p^{n-1}q^{n+1},$$

而

$$\begin{aligned}p^{2n}+q^{2n}-p^{n+1}q^{n-1}-p^{n-1}q^{n+1}&=p^{n+1}\left(p^{n-1}-q^{n-1}\right)-q^{n+1}\left(p^{n-1}-q^{n-1}\right)\\&=\left(p^{n+1}-q^{n+1}\right)\left(p^{n-1}-q^{n-1}\right).\end{aligned}$$

根據同樣的論證，兩個括號均非零且與 $p-q$ 同號，所以

$$\mathrm{P}(S=2n+1)>\mathrm{P}(A=2n+1)\qquad (n>1).$$

同樣，因子 $p^{n-1}-q^{n-1}$ 在 $n=1$ 時為零，與 $\mathrm{P}(S=3)=\mathrm{P}(A=3)$ 吻合。

總括而言：長度為 1 的相同連串比交替連串出現的機會小，長度 2 和 3 的機會相等，而由 4 起的每個長度，相同連串出現的機會都較大。

## 驗算

- 取 $p=\frac{1}{3}$，$q=\frac{2}{3}$。則 $\mathrm{P}(S=4)=pq(p^3+q^3)=\frac{2}{9}\cdot\frac{1}{3}=\frac{2}{27}\approx 0.074$，而 $\mathrm{P}(A=4)=p^2q^2=\frac{4}{81}\approx 0.049$，所以 $S=4$ 的機會確實較大。
- 將 $\mathrm{P}(S=n)$ 對所有 $n\geq 1$ 求和，得 $q\cdot\frac{p}{1-p}+p\cdot\frac{q}{1-q}=p+q=1$，理應如此。
- 每條公式都已用電腦核實：對多個 $p$ 值，以精確分數列舉所有 14 次擲幣的序列，驗證長度至 13 的情況。

## 常見錯誤

- 忘記那次結束連串的擲幣。$S=n$ 需要第 $n+1$ 次擲幣不同；缺了這一步，你計算的其實是 $\mathrm{P}(S\geq n)$。

- 交替連串的結尾字母弄錯：奇數長度的連串以開頭的同一字母結尾，偶數長度的則以另一字母結尾。
- 在第 (iii) 部分代入 $q = 1 - p$：這樣各式會變成 $p$ 的多項式，因式分解難看得多。

## STEP 之外

$S$ 是一個幾何型隨機變量（geometric-type random variable）：給定第一次擲幣的結果，連串長度就是等到第一次改變的等候時間（waiting time）。這類連串是統計學中遊程檢定（runs test）的基礎，該檢定用來判斷一串結果看來是否隨機；而生成函數（generating functions）則提供一個簡潔的方法，求出最長連串的完整分佈。Grimmett 和 Welsh 的 *Probability: An Introduction*（《概率導論》）以合適的程度講解生成函數和等候時間。

詳細解答

# STEP III 2013，第 1 題

純數學 • 半角代換

**STEP III 2013，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/13-s3-q1

## 關鍵思路

半角代換（half-angle substitution）$t = \tan\frac{1}{2}x$ 可把任何 $\sin x$ 的有理函數（rational function）化為 $t$ 的有理函數，因此第一個積分會變成 $\int \mathrm{d}t/(\text{二次式})$，再用配方法（completing the square)即可完成。要把答案寫成題目要求的形式，則需要用到 $\tan(A-B)$ 公式。至於 $I_n$，提示的設計正是要令那個麻煩的分母約去：$\sin^{n+1}x + 2\sin^n x = \sin^n x\,(2+\sin x)$。由此得到一條遞推關係（recurrence），從 $I_0$ 開始逐級往上推；而 $I_0$ 是唯一真正需要用到第一部分結果的積分。

## 解答

### 代換

設 $t = \tan\frac{1}{2}x$，由連鎖法則（chain rule）得

$$\frac{\mathrm{d}t}{\mathrm{d}x} = \tfrac{1}{2}\sec^2\tfrac{1}{2}x = \tfrac{1}{2}\left(1+\tan^2\tfrac{1}{2}x\right) = \tfrac{1}{2}(1+t^2).$$

至於正弦，先用二倍角公式（double-angle formula），再寫成 $\cos^2\frac{1}{2}x = 1/\sec^2\frac{1}{2}x$：

$$\sin x = 2\sin\tfrac{1}{2}x\cos\tfrac{1}{2}x = 2\tan\tfrac{1}{2}x\cos^2\tfrac{1}{2}x = \frac{2\tan\frac{1}{2}x}{\sec^2\frac{1}{2}x} = \frac{2t}{1+t^2}.$$

兩個結果均已證明。

### 第一個積分

當 $x$ 由 0 變到 $\frac{1}{2}\pi$ 時，$t$ 由 $\tan 0 = 0$ 變到 $\tan\frac{1}{4}\pi = 1$，而且 $\mathrm{d}x = \dfrac{2\,\mathrm{d}t}{1+t^2}$。此外

$$1 + a\sin x = \frac{1 + t^2 + 2at}{1 + t^2},$$

因此因子 $1 + t^2$ 互相約去，得

$$\int_0^{\frac{1}{2}\pi} \frac{\mathrm{d}x}{1 + a\sin x} = \int_0^1 \frac{2\,\mathrm{d}t}{t^2 + 2at + 1} = \int_0^1 \frac{2\,\mathrm{d}t}{(t+a)^2 + (1-a^2)}.$$

由於 $0 < a < 1$，常數 $1 - a^2$ 為正；記 $r = \sqrt{1-a^2}$。利用 $\int \mathrm{d}u/(u^2 + r^2) = \frac{1}{r}\arctan\frac{u}{r}$，

$$\int_0^{\frac{1}{2}\pi} \frac{\mathrm{d}x}{1 + a\sin x} = \frac{2}{r}\Big[\arctan\frac{t+a}{r}\Big]_0^1 = \frac{2}{r}(\alpha - \beta),$$

其中

$$\tan\alpha = \frac{1+a}{r} = \frac{1+a}{\sqrt{1-a}\sqrt{1+a}} = \frac{\sqrt{1+a}}{\sqrt{1-a}}, \qquad \tan\beta = \frac{a}{r},$$

且 $\alpha, \beta \in (0, \frac{1}{2}\pi)$。

> **想一想** 這看來不像我們想要的答案。怎樣才能把兩個反正切之差化為一個反正切？

取差的正切。我們有

$$\tan\alpha - \tan\beta = \frac{(1+a) - a}{r} = \frac{1}{r}, \qquad 1 + \tan\alpha\tan\beta = 1 + \frac{a(1+a)}{1-a^2} = 1 + \frac{a}{1-a} = \frac{1}{1-a},$$

所以

$$\tan(\alpha - \beta) = \frac{\tan\alpha - \tan\beta}{1 + \tan\alpha\tan\beta} = \frac{1-a}{r} = \frac{1-a}{\sqrt{1-a}\sqrt{1+a}} = \frac{\sqrt{1-a}}{\sqrt{1+a}}.$$

在取反函數之前，我們需要 $\alpha - \beta$ 位於 arctan 的值域 $(-\frac{1}{2}\pi, \frac{1}{2}\pi)$ 之內。事實確是如此：兩個角都在 $(0, \frac{1}{2}\pi)$ 內，而且 $\tan\alpha > \tan\beta$（兩者之差為 $1/r > 0$），所以 $0 < \alpha - \beta < \frac{1}{2}\pi$。因此 $\alpha - \beta = \arctan\dfrac{\sqrt{1-a}}{\sqrt{1+a}}$，並且

$$\int_0^{\frac{1}{2}\pi} \frac{\mathrm{d}x}{1 + a\sin x} = \frac{2}{\sqrt{1-a^2}}\arctan\frac{\sqrt{1-a}}{\sqrt{1+a}},$$

此即所求。

### 計算 $I_3$

按照提示，

$$I_{n+1} + 2I_n = \int_0^{\frac{1}{2}\pi} \frac{\sin^n x\,(\sin x + 2)}{2 + \sin x}\,\mathrm{d}x = \int_0^{\frac{1}{2}\pi} \sin^n x\,\mathrm{d}x.$$

我們需要右方在 $n = 0, 1, 2$ 時的值：

$$\int_0^{\frac{1}{2}\pi} 1\,\mathrm{d}x = \tfrac{1}{2}\pi, \qquad \int_0^{\frac{1}{2}\pi} \sin x\,\mathrm{d}x = 1, \qquad \int_0^{\frac{1}{2}\pi} \sin^2 x\,\mathrm{d}x = \int_0^{\frac{1}{2}\pi} \tfrac{1}{2}(1-\cos 2x)\,\mathrm{d}x = \tfrac{1}{4}\pi.$$

**想一想** 遞推關係把每個 $I_n$ 與下一個聯繫起來，但它從哪裏開始？

我們需要另行求出 $I_0$，這正是第一部分派上用場的地方。提出因子 2 並取 $a = \frac{1}{2}$，

$$\begin{aligned} I_0 = \int_0^{\frac{1}{2}\pi} \frac{\mathrm{d}x}{2+\sin x} &= \frac{1}{2}\int_0^{\frac{1}{2}\pi} \frac{\mathrm{d}x}{1+\frac{1}{2}\sin x} = \frac{1}{2}\cdot\frac{2}{\sqrt{3/4}} \arctan\frac{\sqrt{1/2}}{\sqrt{3/2}} \\ &= \frac{2}{\sqrt{3}} \arctan\frac{1}{\sqrt{3}} = \frac{2}{\sqrt{3}}\cdot\frac{\pi}{6} = \frac{\pi}{3\sqrt{3}}. \end{aligned}$$

現在沿遞推關係 $I_{n+1} = \int_0^{\frac{1}{2}\pi} \sin^n x\,\mathrm{d}x - 2I_n$ 逐級往上推：

$$\begin{aligned} I_1 &= \tfrac{1}{2}\pi - 2I_0 = \frac{\pi}{2} - \frac{2\pi}{3\sqrt{3}}, \\ I_2 &= 1 - 2I_1 = 1 - \pi + \frac{4\pi}{3\sqrt{3}}, \\ I_3 &= \tfrac{1}{4}\pi - 2I_2 = \frac{\pi}{4} - 2 + 2\pi - \frac{8\pi}{3\sqrt{3}}. \end{aligned}$$

因此

$$I_3 = \frac{9\pi}{4} - \frac{8\sqrt{3}\,\pi}{9} - 2.$$

## 驗算

- 當 $a \to 0$ 時，第一個結果應給出 $\int_0^{\frac{1}{2}\pi} 1\,\mathrm{d}x = \frac{1}{2}\pi$；它的確趨向 $2\arctan 1 = \frac{1}{2}\pi$。
- 由於在 $[0, \frac{1}{2}\pi]$ 上 $0 \le \sin^3 x/(2+\sin x) \le \frac{1}{3}$，我們需要 $0 < I_3 < \frac{1}{6}\pi \approx 0.52$。我們的答案是 $I_3 \approx 0.2318$，而 $I_3$ 的數值積分結果與之在許多位小數上吻合。同樣，$I_0 = \pi/(3\sqrt{3}) \approx 0.6046$，它理應介乎 $\frac{\pi}{6}$ 與 $\frac{\pi}{4}$ 之間，事實亦然。
- 各值遞減，$I_0 > I_1 > I_2 > I_3 > 0$，這是應有的結果，因為在該區間上 $0 \le \sin x \le 1$。

## 常見錯誤

- 忘記把積分上下限改為 $t = 0$ 和 $t = 1$，或忘記 $\mathrm{d}x$ 帶來的因子 $2/(1+t^2)$。
- 用 $\tan(\alpha - \beta)$ 推斷 $\alpha - \beta = \arctan(\dots)$，卻沒有檢查 $\alpha - \beta$ 是否位於 $(-\frac{1}{2}\pi, \frac{1}{2}\pi)$ 之內。

▶ 以 $a = 2$ 把第一個結果套用於 $I_0$：該結果只在 $0 < a < 1$ 時成立，所以必須先提出因子 2。

## STEP 之外

$t = \tan\frac{1}{2}x$ 代換（常稱為魏爾斯特拉斯代換（Weierstrass substitution））說明了每個 $\sin x$ 與 $\cos x$ 的有理函數都有初等的原函數（antiderivative），因為部分分式（partial fractions）可以處理任何 $t$ 的有理函數。從幾何角度看，它就是單位圓的球極投影參數化（stereographic parametrisation），$(\cos x, \sin x) = \left(\frac{1-t^2}{1+t^2}, \frac{2t}{1+t^2}\right)$，而由有理數 $t$ 出發，它亦能產生所有畢氏三元數組（Pythagorean triple）。像 $\int_0^{2\pi} \mathrm{d}x/(1 + a\sin x)$ 這類積分，亦是複分析（complex analysis）中圍道積分（contour integration）的經典入門應用；Needham 的 *Visual Complex Analysis*（《直觀複分析》）是理解箇中原因的絕佳讀物。

詳細解答

# STEP III 2013，第 2 題

純數學 • 麥克勞林級數

> **STEP III 2013，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/13-s3-q2

## 關鍵思路

先寫成 $y\sqrt{1-x^2} = \arcsin x$ 以消去平方根，一階導數（derivative）便輕而易舉。再用萊布尼茲法則（Leibniz's rule）對這條一階方程反覆求導，就得到高階關係式。在 $x = 0$ 處大部分項都消失，剩下一條隔兩步的導數遞推關係（recurrence relation），由此逐一構造出麥克勞林（Maclaurin）係數。最後的和式看起來正是該級數在一個精心選取的 $x$ 值處的取值。

## 解答

### 一階方程

兩邊乘開得 $y\sqrt{1-x^2} = \arcsin x$，然後兩邊對 $x$ 求導，左邊用乘積法則（product rule）：

$$\sqrt{1-x^2}\,\frac{\mathrm{d}y}{\mathrm{d}x} - \frac{xy}{\sqrt{1-x^2}} = \frac{1}{\sqrt{1-x^2}}.$$

全式乘以 $\sqrt{1-x^2}$（當 $|x| < 1$ 時不為零），得

$$(1-x^2)\frac{\mathrm{d}y}{\mathrm{d}x} - xy - 1 = 0,$$

即為所求。

### 高階關係式

以 $y^{(k)}$ 表示 $\dfrac{\mathrm{d}^k y}{\mathrm{d}x^k}$。將一階方程求導 $n+1$ 次。常數 1 消失，而兩個乘積則由萊布尼茲法則處理。由於 $1-x^2$ 的各階導數依次為 $-2x$、$-2$，之後是 0，而 $x$ 的導數依次為 1，之

後是 0，

$$\begin{aligned}\frac{\mathrm{d}^{n+1}}{\mathrm{d}x^{n+1}}\Big[(1-x^2)y'\Big] &= (1-x^2)y^{(n+2)} + \binom{n+1}{1}(-2x)\,y^{(n+1)} + \binom{n+1}{2}(-2)\,y^{(n)} \\ &= (1-x^2)y^{(n+2)} - 2(n+1)x\,y^{(n+1)} - n(n+1)\,y^{(n)},\end{aligned}$$

$$\frac{\mathrm{d}^{n+1}}{\mathrm{d}x^{n+1}}\Big[xy\Big] = x\,y^{(n+1)} + (n+1)\,y^{(n)}.$$

第一式減去第二式並合併同類項，

$$(1-x^2)y^{(n+2)} - \big(2(n+1)+1\big)x\,y^{(n+1)} - \big(n(n+1)+(n+1)\big)y^{(n)} = 0,$$

即

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

以上推導並不需要 $n \geq 1$：只求導一次（$n = 0$）也得到同樣的模式 $(1-x^2)y'' - 3xy' - y = 0$，而我們稍後也會用到這個情況。

### 麥克勞林級數

代入 $x = 0$。中間一項帶有因子 $x$，因此消失，得

$$y^{(n+2)}(0) = (n+1)^2\,y^{(n)}(0) \qquad (n \geq 0).$$

這條遞推關係每次跨兩步，所以需要兩個初始值。由定義，$y(0) = \arcsin 0 = 0$；而一階方程在 $x = 0$ 處給出 $y'(0) - 0 - 1 = 0$，所以 $y'(0) = 1$。

**想一想** 怎樣只憑這兩個值得出每一階導數？

遞推關係只把 $y^{(n)}(0)$ 與 $y^{(n+2)}(0)$ 聯繫起來，所以偶數階導數組成一條由 $y(0)$ 開始的鏈，奇數階導數則組成另一條由 $y'(0)$ 開始的鏈。

*偶次冪。*由於 $y(0) = 0$，遞推關係給出 $y''(0) = 1^2 \cdot 0 = 0$，然後 $y^{(4)}(0) = 0$，如此類推：對每個 $k$ 都有 $y^{(2k)}(0) = 0$。因此 $x$ 的每個偶次冪的係數都是零。（這與 $y$ 是奇函數（odd function）相符，因為它是奇函數 $\arcsin x$ 除以偶函數 $\sqrt{1-x^2}$ 的商。）

*奇次冪。*以 $n = 1, 3, \ldots, 2k-1$ 代入遞推關係，

$$y^{(2k+1)}(0) = (2k)^2(2k-2)^2 \cdots 4^2 \cdot 2^2 \cdot y'(0) = \big(2 \cdot 4 \cdots 2k\big)^2 = \big(2^k\,k!\big)^2,$$

這對 $k = 0$ 亦成立。除以 $(2k+1)!$ 便得到麥克勞林係數：

$$\frac{\arcsin x}{\sqrt{1-x^2}} = \sum_{k=0}^{\infty} \frac{2^{2k}(k!)^2}{(2k+1)!}\,x^{2k+1} = x + \frac{2^2}{3!}x^3 + \frac{2^2 \cdot 4^2}{5!}x^5 + \cdots,$$

所以一般奇次項為 $\dfrac{2^{2k}(k!)^2}{(2k+1)!}x^{2k+1}$，而一般偶次項為 $0 \cdot x^{2k}$。

### 無窮級數之和

和式的第 $n$ 項是 $\dfrac{(n!)^2}{(2n+1)!}$，其中 $n=0$ 給出首項 1。與級數比較，我們想消去因子 $2^{2n}$，所以選取 $x$ 使 $2^{2n}x^{2n+1}$ 為 1 的常數倍：$x=\frac{1}{2}$ 給出 $2^{2n}\left(\frac{1}{2}\right)^{2n+1}=\frac{1}{2}$。因此

$$\frac{\arcsin\frac{1}{2}}{\sqrt{1-\frac{1}{4}}}=\frac{1}{2}\sum_{n=0}^{\infty}\frac{(n!)^2}{(2n+1)!}.$$

左邊等於 $\dfrac{\pi/6}{\sqrt{3}/2}=\dfrac{\pi}{3\sqrt{3}}$，所以

$$1+\frac{1}{3!}+\frac{2^2}{5!}+\frac{2^2\times 3^2}{7!}+\cdots=\frac{2\pi}{3\sqrt{3}}=\frac{2\sqrt{3}\,\pi}{9}.$$

## 驗算

- 將已知級數 $\arcsin x=x+\frac{1}{6}x^3+\cdots$ 乘以 $(1-x^2)^{-1/2}=1+\frac{1}{2}x^2+\cdots$，得 $x+\left(\frac{1}{6}+\frac{1}{2}\right)x^3+\cdots=x+\frac{2}{3}x^3+\cdots$，而 $\frac{2^2}{3!}=\frac{2}{3}$ 果然吻合。
- 和式的首五項為 $1+0.16667+0.03333+0.00714+0.00159\approx 1.2087$，已經接近 $\frac{2\pi}{3\sqrt{3}}\approx 1.20920$；各項每次大約縮小 4 倍，所以餘項很小。

## 常見錯誤

- 在萊布尼茲法則中漏掉 $y^{(n)}$ 項上的二項式係數（binomial coefficient）$\binom{n+1}{2}$，以致得不出 $(n+1)^2$。
- 只對 $n\geq 1$ 使用遞推關係，因而始終沒有證明 $y''(0)=0$；偶數鏈需要 $n=0$（求導一次）的情況。
- 在最後一部分代入 $x=1$ 或 $x=\frac{1}{4}$：因子 $2^{2n}$ 必須由 $x^{2n}$ 抵消，這迫使 $x=\frac{1}{2}$。

## STEP 之外

這裏我們「忽略了收斂性的問題」，但大學第一門分析課會把這點嚴格化：由比值審斂法（ratio test）可知該級數的收斂半徑（radius of convergence）為 1，而帶餘項的泰勒定理（Taylor's theorem with remainder）則證明在該區間內級數確實等於該函數。係數 $\frac{2^{2k}(k!)^2}{(2k+1)!}$ 與中心二項式係數（central binomial coefficients）關係密切，而同一個函數積分一次後便可生成 $(\arcsin x)^2$ 的級數。關於冪級數背後的分析，可參閱 Abbott 的 *Understanding Analysis*（《理解分析》）。

詳細解答

# STEP III 2013，第 3 題

純數學 • 向量與球面

> **STEP III 2013，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/13-s3-q3

## 關鍵思路

一切都源於兩個事實：每個 $\mathbf{p}_i$ 都是單位向量（unit vector），而且四者之和為 $\mathbf{0}$。把長度的平方展開為點積（dot product），就能把 $(XP_i)^2$ 化成關於 $\mathbf{x}\cdot\mathbf{p}_i$ 的線性（linear）式子；由於 $\sum \mathbf{p}_i = \mathbf{0}$，線性項之和便會全部消去。至於四次方，會剩下一個二次項，因此需要實際的坐標，而這正是第 (ii) 部分所提供的。

## 解答

### 點積 $\mathbf{p}_i \cdot \mathbf{p}_j$

由於這是正四面體（regular tetrahedron），六條棱的長度全部相等。對 $i \neq j$，

$$(P_iP_j)^2 = (\mathbf{p}_i - \mathbf{p}_j)\cdot(\mathbf{p}_i - \mathbf{p}_j) = |\mathbf{p}_i|^2 + |\mathbf{p}_j|^2 - 2\,\mathbf{p}_i\cdot\mathbf{p}_j = 2 - 2\,\mathbf{p}_i\cdot\mathbf{p}_j,$$

這裏用了 $|\mathbf{p}_i| = 1$（每個頂點都在單位球面（unit sphere）上）。因此，棱長相等迫使每個 $\mathbf{p}_i\cdot\mathbf{p}_j$（$i \neq j$）都取同一個值；記這個值為 $c$。

現在把 $\mathbf{p}_1 + \mathbf{p}_2 + \mathbf{p}_3 + \mathbf{p}_4 = \mathbf{0}$ 與 $\mathbf{p}_1$ 作點積：

$$\mathbf{p}_1\cdot\mathbf{p}_1 + \mathbf{p}_1\cdot\mathbf{p}_2 + \mathbf{p}_1\cdot\mathbf{p}_3 + \mathbf{p}_1\cdot\mathbf{p}_4 = 0 \quad\Longrightarrow\quad 1 + 3c = 0,$$

所以 $c = -\frac{1}{3}$。即對所有 $i \neq j$，都有 $\mathbf{p}_i\cdot\mathbf{p}_j = -\frac{1}{3}$，得證。

### 第 (i) 部分

由於 $|\mathbf{x}| = 1$ 亦成立，

$$(XP_i)^2 = (\mathbf{p}_i - \mathbf{x})\cdot(\mathbf{p}_i - \mathbf{x}) = |\mathbf{p}_i|^2 - 2\,\mathbf{x}\cdot\mathbf{p}_i + |\mathbf{x}|^2 = 2 - 2\,\mathbf{x}\cdot\mathbf{p}_i.$$

對 $i$ 求和，並利用點積的分配律（distributive law），

$$\sum_{i=1}^{4}(XP_i)^2 = 8 - 2\,\mathbf{x}\cdot\left(\mathbf{p}_1+\mathbf{p}_2+\mathbf{p}_3+\mathbf{p}_4\right) = 8 - 2\,\mathbf{x}\cdot\mathbf{0} = 8,$$

得證。

**第 (ii) 部分**

設 $\mathbf{p}_1 = (0,0,1)$ 及 $\mathbf{p}_2 = (a,0,b)$，條件 $\mathbf{p}_1\cdot\mathbf{p}_2 = -\frac{1}{3}$ 立即給出 $b = -\frac{1}{3}$。再由 $|\mathbf{p}_2| = 1$ 得

$$a^2 = 1 - b^2 = \tfrac{8}{9}, \qquad \text{所以} \qquad a = \frac{2\sqrt{2}}{3},$$

由於 $a > 0$，取正根。

設 $\mathbf{p}_3 = (u,v,w)$。三個條件把它確定下來：

$$\mathbf{p}_1\cdot\mathbf{p}_3 = -\tfrac{1}{3} \implies w = -\tfrac{1}{3},$$

$$\mathbf{p}_2\cdot\mathbf{p}_3 = -\tfrac{1}{3} \implies \tfrac{2\sqrt{2}}{3}u + \tfrac{1}{9} = -\tfrac{1}{3} \implies u = -\tfrac{4}{9}\cdot\tfrac{3}{2\sqrt{2}} = -\frac{\sqrt{2}}{3},$$

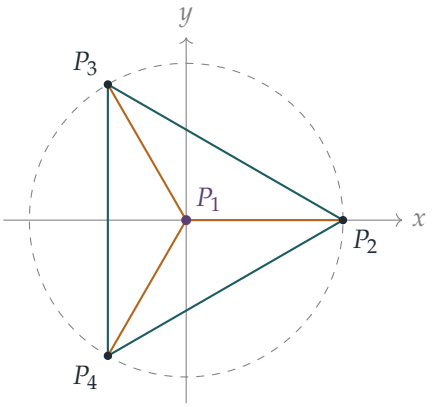
$$|\mathbf{p}_3| = 1 \implies v^2 = 1 - \tfrac{2}{9} - \tfrac{1}{9} = \tfrac{6}{9} \implies v = \pm\frac{\sqrt{6}}{3}.$$

$\mathbf{p}_4$ 滿足完全相同的方程，所以 $P_3$ 和 $P_4$ 就是由此得出的兩個點（它們必須是不同的點，所以各取一個符號）。按其中一種方式標記，

$$P_3 = \left(-\frac{\sqrt{2}}{3}, \frac{\sqrt{6}}{3}, -\frac{1}{3}\right), \qquad P_4 = \left(-\frac{\sqrt{2}}{3}, -\frac{\sqrt{6}}{3}, -\frac{1}{3}\right).$$

（$P_3$ 和 $P_4$ 的標記可以互換。）驗證一下：四個位置向量（position vector）之和確實為 $\mathbf{0}$，而且 $\mathbf{p}_3\cdot\mathbf{p}_4 = \frac{2}{9} - \frac{6}{9} + \frac{1}{9} = -\frac{1}{3}$。

從上方俯視，$P_2$、$P_3$、$P_4$ 在平面 $z = -\frac{1}{3}$ 上組成一個等邊三角形（equilateral triangle），而 $P_1$ 位於其中心的正上方：



**第 (iii) 部分**

由第 (i) 部分，$(XP_i)^2 = 2(1 - \mathbf{x}\cdot\mathbf{p}_i)$，兩邊平方，

$$(XP_i)^4 = 4(1-\mathbf{x}\cdot\mathbf{p}_i)^2 \qquad \Longrightarrow \qquad \sum_{i=1}^{4}(XP_i)^4 = 4\sum_{i=1}^{4}(1-\mathbf{x}\cdot\mathbf{p}_i)^2,$$

得證。展開括號，並再次利用 $\sum \mathbf{x}\cdot\mathbf{p}_i = 0$，

$$\sum_{i=1}^{4}(XP_i)^4 = 4\Big(4 - 2\sum_{i=1}^{4}\mathbf{x}\cdot\mathbf{p}_i + \sum_{i=1}^{4}(\mathbf{x}\cdot\mathbf{p}_i)^2\Big) = 16 + 4\sum_{i=1}^{4}(\mathbf{x}\cdot\mathbf{p}_i)^2.$$

? **想一想** 線性項又一次消失了。為甚麼同一招不能把 $\sum(\mathbf{x}\cdot\mathbf{p}_i)^2$ 也消掉？

平方和對 $\mathbf{p}_i$ 而言並非線性，所以知道 $\sum \mathbf{p}_i = \mathbf{0}$ 並不能直接告訴我們它的值。這正是需要用到坐標的地方。設 $\mathbf{x} = (x, y, z)$，並用第 (ii) 部分的頂點，

$$\mathbf{x}\cdot\mathbf{p}_1 = z, \qquad \mathbf{x}\cdot\mathbf{p}_2 = \tfrac{1}{3}(2\sqrt{2}\,x - z),$$
$$\mathbf{x}\cdot\mathbf{p}_3 = \tfrac{1}{3}(-\sqrt{2}\,x + \sqrt{6}\,y - z), \qquad \mathbf{x}\cdot\mathbf{p}_4 = \tfrac{1}{3}(-\sqrt{2}\,x - \sqrt{6}\,y - z).$$

平方後相加，合併同類項：

- $x^2$：$\frac{1}{9}(8+2+2) = \frac{4}{3}$；　$y^2$：$\frac{1}{9}(6+6) = \frac{4}{3}$；　$z^2$：$1 + \frac{1}{9}(1+1+1) = \frac{4}{3}$。
- $xy$ 和 $yz$：這兩項只出現在 $\mathbf{x}\cdot\mathbf{p}_3$ 和 $\mathbf{x}\cdot\mathbf{p}_4$ 之中，而這兩者只在 $y$ 項的符號上有分別，所以互相抵消。
- $xz$：$\frac{1}{9}\big(-4\sqrt{2} + 2\sqrt{2} + 2\sqrt{2}\big) = 0$。

因此

$$\sum_{i=1}^{4}(\mathbf{x}\cdot\mathbf{p}_i)^2 = \tfrac{4}{3}(x^2+y^2+z^2) = \tfrac{4}{3},$$

因為 $X$ 在單位球面上。所以

$$\sum_{i=1}^{4}(XP_i)^4 = 16 + 4\cdot\tfrac{4}{3} = \frac{64}{3},$$

這與 $X$ 的位置無關，得證。

? **想一想** 我們用了特定的坐標。這個結論對內接於球面的每一個正四面體都成立嗎？

成立。任何這樣的四面體都可以繞 $O$ 旋轉（rotation），使 $P_1$ 轉到 $(0,0,1)$，而 $P_2$ 轉到半平面 $y = 0$、$x > 0$ 之內；旋轉保持長度不變，而當 $X$ 走遍整個球面時，它旋轉後的像亦走遍整個球面。所以第 (ii) 部分的坐標並不失一般性。

## 驗算

- 取 $X = P_1$。各距離為 0 及三條棱，每條長 $\sqrt{2-2(-\frac{1}{3})} = \sqrt{8/3}$。於是 $\sum(XP_i)^2 = 3\cdot\frac{8}{3} = 8$，而 $\sum(XP_i)^4 = 3\cdot\frac{64}{9} = \frac{64}{3}$，與兩個結果都吻合。
- 取 $X = (0,0,-1)$，即 $P_1$ 的對徑點。此時 $XP_1 = 2$，而其餘三個頂點有 $\mathbf{x}\cdot\mathbf{p}_i = \frac{1}{3}$，所以 $(XP_i)^2 = \frac{4}{3}$。這給出 $4 + 3\cdot\frac{4}{3} = 8$ 及 $16 + 3\cdot\frac{16}{9} = \frac{64}{3} \approx 21.33$。
- 用電腦在球面上數以千計的隨機點上驗算，每次都得到 8 和 21.333…。

## 常見錯誤

- 斷言所有 $\mathbf{p}_i\cdot\mathbf{p}_j$ 都相等，卻沒有說明原因：原因是棱長相等，透過 $(P_iP_j)^2 = 2 - 2\,\mathbf{p}_i\cdot\mathbf{p}_j$ 得出。
- 忘記 $|\mathbf{x}| = 1$；推導 $(XP_i)^2 = 2 - 2\,\mathbf{x}\cdot\mathbf{p}_i$ 和最後一步 $x^2+y^2+z^2 = 1$ 都需要用到這一點。
- 第 (iii) 部分交叉項（cross terms）的計算錯誤：那裏一個符號出錯，就會留下對 $X$ 的虛假依賴。

## STEP 之外

第 (iii) 部分的核心是恆等式 $\sum_i(\mathbf{x}\cdot\mathbf{p}_i)^2 = \frac{4}{3}|\mathbf{x}|^2$，它對每一個向量 $\mathbf{x}$ 都成立。用矩陣（matrix）的語言來說，就是 $\sum_i \mathbf{p}_i\mathbf{p}_i^{\mathsf{T}} = \frac{4}{3}I$：這四個頂點向量的表現猶如一組超完備的標準正交基（over-complete orthonormal basis），稱為緊框架（tight frame）。這一點由對稱性必然成立：該矩陣與四面體的每一個旋轉都可交換，因此必定是 $I$ 的倍數，而它的跡（trace）是 4。球面上的點集，若其低次平均值與球面本身的平均值相同，便稱為球面設計（spherical designs）；它們應用於數值積分（numerical integration）和編碼理論（coding theory）。關於背後的線性代數（外積（outer product）、對稱矩陣（symmetric matrix）和特徵值（eigenvalue）），可參閱斯特朗（Strang）的 *Introduction to Linear Algebra*（《線性代數導論》）。

詳細解答

# STEP III 2013，第 4 題

純數學 • 單位根的乘積

**STEP III 2013，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/13-s3-q4

本題在 *Beyond STEP* 第 5.3 節因式分解 $z^{2n}+1$ 中有完整的引導式講解：https://beyondstep.keithmcnulty.org/zh-hant/ch05/z2n-plus-one.html

## 關鍵思路

$z^{2n}=-1$ 的根以共軛（conjugate）對 $\mathrm{e}^{\pm \mathrm{i}\theta}$ 的形式出現，而每一對相乘都得出一個實係數二次式，由此便得到所需的因式分解（factorisation）。接着，只要選取一個恰當的 $z$ 值，就能把這個因式分解變成一個關於餘弦的命題：當 $z=\mathrm{i}$ 時，每個二次式中 $z^2+1$ 的部分都等於零，只剩下 $-2z\cos\theta$ 這一項。在第 (ii) 部分，同樣的代入會令兩邊都等於零，所以我們要先約去在 $z=\mathrm{i}$ 時等於零的那個因式。

## 解答

### 因式分解

展開，

$$\begin{aligned}(z-\mathrm{e}^{\mathrm{i}\theta})(z-\mathrm{e}^{-\mathrm{i}\theta}) &= z^2 - z\left(\mathrm{e}^{\mathrm{i}\theta}+\mathrm{e}^{-\mathrm{i}\theta}\right)+\mathrm{e}^{\mathrm{i}\theta}\mathrm{e}^{-\mathrm{i}\theta}\\ &= z^2-2z\cos\theta+1,\end{aligned}$$

因為 $\mathrm{e}^{\mathrm{i}\theta}+\mathrm{e}^{-\mathrm{i}\theta}=2\cos\theta$。

現在，$\mathrm{e}^{\mathrm{i}\theta}$ 滿足 $z^{2n}=-1=\mathrm{e}^{\mathrm{i}\pi}$ 當且僅當 $2n\theta$ 是 $\pi$ 的奇數倍，即對某整數 $k$，$\theta=\frac{(2k-1)\pi}{2n}$。在 $(-\pi,\pi]$ 內的值是

$$\theta=\pm\frac{\pi}{2n},\ \pm\frac{3\pi}{2n},\ \ldots,\ \pm\frac{(2n-1)\pi}{2n},$$

因為 $\frac{(2n-1)\pi}{2n} < \pi$，而 $\frac{(2n+1)\pi}{2n} > \pi$。這些是 $2n$ 個互不相同的根，所以它們就是 $-1$ 的全部 $(2n)$ 次方根：

$$\mathrm{e}^{\pm \mathrm{i}\theta_k}, \qquad \theta_k = \frac{(2k-1)\pi}{2n}, \quad k = 1, 2, \ldots, n.$$

由於 $z^{2n}+1$ 是 $2n$ 次的首一（monic）多項式（polynomial），而且有這 $2n$ 個互不相同的根，

$$z^{2n}+1 = \prod_{k=1}^{n} (z - \mathrm{e}^{\mathrm{i}\theta_k})(z - \mathrm{e}^{-\mathrm{i}\theta_k}) = \prod_{k=1}^{n} \left(z^2 - 2z\cos\theta_k + 1\right),$$

這裏對每一對共軛根都用了第一個結果。這就是所求的恆等式。

### 第 (i) 部分

代入 $z = \mathrm{i}$。每個因式變成

$$\mathrm{i}^2 - 2\mathrm{i}\cos\theta_k + 1 = -2\mathrm{i}\cos\theta_k,$$

而由於 $n$ 是偶數，左邊是 $\mathrm{i}^{2n}+1 = (-1)^n + 1 = 2$。因此

$$2 = (-2\mathrm{i})^n \prod_{k=1}^{n} \cos\theta_k.$$

由於 $n$ 是偶數，$(-2)^n = 2^n$ 而 $\mathrm{i}^n = (\mathrm{i}^2)^{n/2} = (-1)^{\frac{1}{2}n}$，所以 $(-2\mathrm{i})^n = 2^n(-1)^{\frac{1}{2}n}$。因此

$$\prod_{k=1}^{n} \cos\theta_k = \frac{2}{2^n(-1)^{\frac{1}{2}n}} = (-1)^{\frac{1}{2}n} 2^{1-n},$$

這裏用了 $1/(-1)^{\frac{1}{2}n} = (-1)^{\frac{1}{2}n}$。左邊的乘積正是 $\cos\frac{\pi}{2n}\cos\frac{3\pi}{2n}\cdots\cos\frac{(2n-1)\pi}{2n}$，得證。

### 第 (ii) 部分

> **?** **想一想** 為甚麼不直接把 $z = \mathrm{i}$ 再代入因式分解？

當 $n$ 是奇數時，$\mathrm{i}^{2n}+1 = (-1)^n + 1 = 0$，所以左邊等於零。在右邊，中間的值 $k = m$（其中 $m = \frac{1}{2}(n+1)$）給出 $\theta_m = \frac{n\pi}{2n} = \frac{\pi}{2}$，所以該因式是 $z^2+1$，它在 $z = \mathrm{i}$ 時同樣等於零。我們只得到 $0 = 0$，甚麼也得不到。解決辦法是先從兩邊約去因式 $z^2+1$。

利用題目給出的事實和因式分解，

$$(1+z^2)\left(1 - z^2 + z^4 - \cdots + z^{2n-2}\right) = (z^2+1)\prod_{k\neq m}\left(z^2 - 2z\cos\theta_k + 1\right).$$

這些都是多項式，而 $z^2+1$ 不是零多項式，所以我們可以把它約去（多項式除法的結果是唯一的）：

$$1 - z^2 + z^4 - \cdots + z^{2n-2} = \prod_{k\neq m}\left(z^2 - 2z\cos\theta_k + 1\right),$$

這是一個多項式恆等式，對每個 $z$ 都成立，包括 $z = \mathrm{i}$。

現在代入 $z = \mathrm{i}$。在左邊，$(-1)^j z^{2j}$ 這一項變成 $(-1)^j(-1)^j = 1$，而共有 $n$ 項（$j = 0, 1, \ldots, n-1$），所以左邊等於 $n$。右邊有 $n-1$ 個因式，每個都等於 $-2\mathrm{i}\cos\theta_k$：

$$n = (-2\mathrm{i})^{n-1} \prod_{k \neq m} \cos\theta_k.$$

由於 $n-1$ 是偶數，與第 (i) 部分完全一樣，$(-2\mathrm{i})^{n-1} = 2^{n-1}(-1)^{\frac{1}{2}(n-1)}$。

要把乘積化成平方，可把 $k$ 與 $n+1-k$ 配對。由於

$$\theta_{n+1-k} = \frac{(2n+1-2k)\pi}{2n} = \pi - \theta_k,$$

我們有 $\cos\theta_{n+1-k} = -\cos\theta_k$。所有 $k \neq m$ 的值可分成 $\frac{1}{2}(n-1)$ 對 $\{k, n+1-k\}$，其中 $1 \leq k \leq \frac{1}{2}(n-1)$，所以

$$\prod_{k \neq m} \cos\theta_k = \prod_{k=1}^{(n-1)/2} \left(-\cos^2\theta_k\right) = (-1)^{\frac{1}{2}(n-1)} P, \qquad P = \prod_{k=1}^{(n-1)/2} \cos^2\theta_k.$$

綜合起來，兩個符號 $(-1)^{\frac{1}{2}(n-1)}$ 相乘得 $+1$：

$$n = 2^{n-1}(-1)^{\frac{1}{2}(n-1)}(-1)^{\frac{1}{2}(n-1)} P = 2^{n-1} P,$$

所以 $P = n\,2^{1-n}$。$P$ 中最後一個因式對應 $k = \frac{1}{2}(n-1)$，即 $\theta_k = \frac{(n-2)\pi}{2n}$，所以 $P$ 正是

$$\cos^2\frac{\pi}{2n}\cos^2\frac{3\pi}{2n}\cdots\cos^2\frac{(n-2)\pi}{2n} = n\,2^{1-n},$$

得證。（當 $n = 1$ 時，乘積是空乘積，等於 $1 = 1 \cdot 2^0$，結果一致。）

## 驗算

- 第 (i) 部分取 $n = 2$：$\cos\frac{\pi}{4}\cos\frac{3\pi}{4} = -\frac{1}{2} = (-1)^1 2^{-1}$。當 $n = 4$ 時，乘積是 $\frac{1}{8}$，與 $(-1)^2 2^{-3}$ 吻合。
- 第 (ii) 部分取 $n = 3$：$\cos^2\frac{\pi}{6} = \frac{3}{4} = 3 \cdot 2^{-2}$。當 $n = 5$ 時，$\cos^2\frac{\pi}{10}\cos^2\frac{3\pi}{10} = 0.3125 = \frac{5}{16}$。
- 兩條恆等式亦已用電腦對所有不超過 39 的 $n$ 驗證過，而因式分解也在隨機選取的複數 $z$ 上測試過。

## 常見錯誤

- 在第 (ii) 部分直接把 $z = \mathrm{i}$ 代入因式分解，結果只得到 $0 = 0$。

- 弄錯符號：在第 (i) 部分，當 $n \equiv 2 \pmod 4$ 時答案確實是負數，因為對於大於 $\frac{\pi}{2}$ 的那 $\frac{1}{2}n$ 個角，$\cos\theta_k < 0$。
- 忘記 $\cos\theta_{n+1-k} = -\cos\theta_k$，因而在組成平方時弄錯符號。

## STEP 之外

同樣的技巧——把 $z^N \pm 1$ 按其單位根（roots of unity）作因式分解，再在一個巧妙的點上求值——可以得出一整族乘積公式，例如 $\prod_{k=1}^{n-1} \sin\frac{k\pi}{n} = \frac{n}{2^{n-1}}$（試在 $\frac{z^n-1}{z-1}$ 中代入 $z=1$）。在這類有限乘積中令 $n \to \infty$，便會得到歐拉（Euler）的無窮乘積（infinite product）$\sin\pi x = \pi x \prod_{k\geq 1}\left(1-\frac{x^2}{k^2}\right)$，讓你初嘗複分析（complex analysis）如何由函數的零點把函數建構出來。Needham 的 *Visual Complex Analysis*（《直觀複分析》）是從幾何角度欣賞單位根及這些概念的絕佳讀物。

詳細解答

# STEP III 2013，第 5 題

純數學 • 有理性與互質性

**STEP III 2013，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/13-s3-q5

## 關鍵思路

每一步都只靠一件工具：如果一個數整除一個乘積，而且與其中一個因數互質（co-prime），那麼它必定整除另一個因數。把有理數（rational number）寫成最簡分數（lowest terms），消去分母，然後留意每個因數必須落在哪一邊。在第 (ii) 部分，這個方法把分母 $b$ 夾在兩個事實之間，而除非 $b$ 很小，否則這兩個事實不可能同時成立。

我們會反覆用到一個小事實：*若 $p$ 與 $q$ 互質，則對任何正整數 $s$、$t$，$p^s$ 與 $q^t$ 亦互質*。事實上，若有一個質數（prime）同時整除 $p^s$ 和 $q^t$，它就會同時整除 $p$ 和 $q$（整除一個乘積的質數必整除其中一個因數，這正是題目所給結果在 $a$ 為質數時的情況），與互質的假設矛盾。

## 解答

### 第 (i) 部分

由 $q^nN = p^n$ 可見 $p^n$ 整除 $q^nN$。由於 $p$ 與 $q$ 互質，$p^n$ 與 $q^n$ 亦互質，於是由題目所給的結果（取 $a = p^n$、$b = q^n$、$c = N$）得知 $p^n$ 整除 $N$。因此對某個正整數 $k$，有 $N = kp^n$，即為所求。

代回原式，得 $q^nkp^n = p^n$，所以 $q^nk = 1$。兩者皆為正整數，故 $k = 1$ 且 $q^n = 1$，從而

$$q = 1 \quad (\text{且} N = p^n).$$

現在取任意正整數 $n$ 和 $N$，並假設 $\sqrt[n]{N}$ 是有理數。它是正數，所以可以寫成 $\sqrt[n]{N} = p/q$，其中 $p$、$q$ 為正整數，且分數為最簡形式，即 $p$ 與 $q$ 互質。兩邊取 $n$ 次方，得 $q^nN = p^n$，而上面的推導給出 $q = 1$。因此 $\sqrt[n]{N} = p$ 是正整數。所以 $\sqrt[n]{N}$ 要麼是正

整數，要麼是無理數（irrational）。

### 第 (ii) 部分

已知 $a^a d^b = b^a c^b$，且 $\gcd(a,b) = \gcd(c,d) = 1$。

**證明** $d^b = b^a$。由右邊可見 $b^a$ 整除 $a^a d^b$。由於 $a$ 與 $b$ 互質，$b^a$ 與 $a^a$ 互質，所以 $b^a$ 整除 $d^b$。同理，$d^b$ 整除 $b^a c^b$，而 $d^b$ 與 $c^b$ 互質，所以 $d^b$ 整除 $b^a$。兩個互相整除的正整數必定相等：

$$d^b = b^a,$$

再在原方程中約去，亦得 $a^a = c^b$。

若 $p$ 是 $d$ 的一個質因數（prime factor），則 $p$ 整除 $d^b = b^a$，而整除 $b$ 的某個冪的質數必整除 $b$。所以 $p$ 是 $b$ 的一個質因數。

**$p$ 的冪**。設 $p^m$ 和 $p^n$ 分別是整除 $d$ 和 $b$ 的 $p$ 的最高次冪，由上一段可知 $m, n \geq 1$。由質因數分解唯一性（unique factorisation），整除 $d^b$ 的 $p$ 的最高次冪是 $p^{mb}$，而整除 $b^a$ 的則是 $p^{na}$。由於 $d^b = b^a$，兩者相等，所以 $mb = na$，且

$$b = \frac{na}{m}.$$

> **想一想** $b = na/m$ 怎樣給出 $p^n$ 的一個上界？

把它改寫成 $mb = na$：於是 $b$ 整除 $na$。但 $b$ 與 $a$ 互質，所以由題目所給的結果，$b$ 整除 $n$，特別地 $b \leq n$。另外 $p^n$ 整除 $b$，所以 $p^n \leq b$。合起來，

$$p^n \leq b \leq n.$$

但 $p \geq 2$ 且 $n > 0$，所以由題目給出的事實 $y^x > x$（取 $y = p$、$x = n$）得 $p^n > n$。這個矛盾說明我們的假設不可能成立：$d$ *沒有*任何質因數。因此 $d = 1$，於是 $b^a = d^b = 1$，從而

$$b = 1.$$

### 最後部分

設 $r$ 是正有理數，且 $r^r$ 是有理數。把 $r = a/b$ 寫成最簡形式（$a, b$ 為正整數且互質），並把 $r^r = c/d$ 寫成最簡形式（由於 $r^r > 0$，$c, d$ 為正整數，且互質）。那麼

$$\left(\frac{a}{b}\right)^{a/b} = \frac{c}{d} \quad \Longrightarrow \quad \left(\frac{a}{b}\right)^{a} = \left(\frac{c}{d}\right)^{b} \quad \Longrightarrow \quad a^a d^b = b^a c^b,$$

這裏把兩邊取 $b$ 次方，再消去分母。這正好是第 (ii) 部分的情況，所以 $b = 1$，而 $r = a$ 是正整數。

## 驗算

- 整數 $r$ 確實符合條件：$r=2$ 給出 $r^r=4$；而結論 $c^b=a^a$ 在 $b=1$ 時就是 $c=a^a$，理應如此。
- 一個反面例子：$r=\frac{1}{2}$ 給出 $r^r=1/\sqrt{2}$，是無理數，與結果一致。在第 (i) 部分，$\sqrt[3]{27}=3$ 而 $\sqrt[3]{24}$ 是無理數：兩種情況都會出現。
- 用電腦搜尋所有互質的 $a,b\le 24$、$c<200$、$d<60$，只找到 $b=d=1$ 時 $a^ad^b=b^ac^b$ 的解（即 $a=1,2,3$，且 $c=a^a$）。

## 常見錯誤

- 忘記交代爲何 $p^n$ 與 $q^n$ 互質：題目所給的結果是關於 $a$ 與 $b$ 互質的，所以你需要那些冪互質。
- 由 $b=na/m$ 直接得出 $p^n\le n$，卻漏了 $b\mid n$ 這一步：正是 $a$ 與 $b$ 互質，才把 $mb=na$ 變成 $b\mid n$。
- 寫了 $(a/b)^{a/b}=c/d$，卻沒有堅持 $c/d$ 是最簡形式；沒有 $\gcd(c,d)=1$，$d^b\mid b^a$ 這一步就不成立。

## STEP 之外

題目所給的「互質因數」結果就是歐幾里得（Euclid）引理（Euclid's lemma），通常由貝祖（Bézout）等式（Bézout's identity）證明，它也是質因數分解唯一性背後的推動力。第一部分是 $\sqrt{2}$ 為無理數這個結果的一般化版本；第二部分則讓你初嘗超越數論（transcendence theory）的滋味：在這個理論中，格爾豐德–施奈德定理（Gelfond–Schneider theorem）證明了，例如 $\sqrt{2}^{\sqrt{2}}$ 甚至不是代數數（algebraic number）。Davenport 的 *The Higher Arithmetic*（《高等算術》）把這些基礎發展得非常優美。

詳細解答

# STEP III 2013，第 6 題

純數學 • 複數模的不等式

**STEP III 2013，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/13-s3-q6

本題在 *Beyond STEP* 第 5.8 節圓盤中隱藏的距離中有完整的引導式講解：https://beyondstep.keithmcnulty.org/zh-hant/ch05/poincare-disc.html

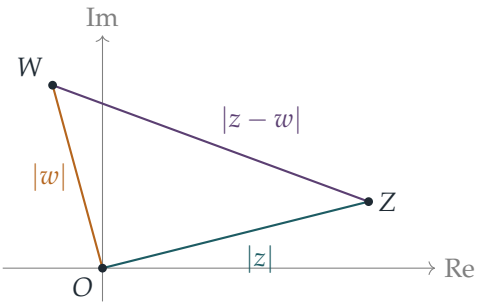
## 關鍵思路

一切都建基於把模的平方寫成一個數乘以它的共軛（conjugate），即 $|u|^2 = uu^*$。用這個方法展開 $|z-w|^2$ 和 $|1-zw^*|^2$，兩者都會出現相同的交叉項（cross terms）$zw^* + z^*w$，這正是同一個量 $E$ 在兩式中都出現的原因。一旦把兩個模（modulus）都寫成「某數的平方減去 $E$」的形式，最後的不等式就變成比較 $\frac{A-E}{B-E}$ 與 $\frac{A}{B}$ 這兩個分數，而這歸結為 $B-A$ 的正負號。

## 解答

### 三角不等式

在阿岡圖（Argand diagram）中標出原點 $O$，以及代表 $z$ 和 $w$ 的點 $Z$ 和 $W$。距離 $OZ$ 和 $OW$ 分別是 $|z|$ 和 $|w|$，而距離 $WZ$ 是 $|z-w|$。




線段 $WZ$ 是由 $W$ 到 $Z$ 的最短路徑，所以它不會比路徑 $W \to O \to Z$ 長：

$$|z-w| = WZ \le WO + OZ = |w| + |z|.$$

等號成立若且唯若 $O$ 位於線段 $WZ$ 上，即 $z$ 和 $w$ 指向相反方向（或其中一個為 0）。若 $O$、$Z$、$W$ 共線（collinear）但 $O$ 不在兩者之間，則這個「三角形」是扁平的，但最短路徑的論證仍然適用。

**第 (i) 部分**

利用 $|u|^2 = uu^*$ 及 $(z-w)^* = z^* - w^*$，

$$\begin{aligned} |z-w|^2 &= (z-w)(z^*-w^*) \\ &= zz^* + ww^* - zw^* - z^*w \\ &= |z|^2 + |w|^2 - (zw^* + z^*w). \end{aligned}$$

由於 $|zw| = |z||w|$，我們有 $(|z|+|w|)^2 = |z|^2 + |w|^2 + 2|zw|$，所以

$$(|z|+|w|)^2 - E = |z|^2 + |w|^2 + 2|zw| - zw^* - z^*w - 2|zw| = |z-w|^2,$$

得證。移項可得

$$E = (|z|+|w|)^2 - |z-w|^2.$$

右方是兩個實數的平方之差，所以 $E$ 是實數。根據三角不等式（triangle inequality），$0 \le |z-w| \le |z| + |w|$，而把非負的量平方不會改變不等號方向，所以 $E \ge 0$。

（直接計算：$z^*w$ 是 $zw^*$ 的共軛，所以 $E = 2\,\mathrm{Re}(zw^*) + 2|zw^*|$，這是實數，而且由於 $\mathrm{Re}\,u \ge -|u|$，它至少為 0。）

### 第 (ii) 部分

同樣地，利用 $(zw^*)^* = z^*w$，

$$\begin{aligned}|1 - zw^*|^2 &= (1 - zw^*)(1 - z^*w)\\ &= 1 - zw^* - z^*w + zz^*ww^*\\ &= 1 + |zw|^2 - (zw^* + z^*w)\\ &= \left(1 + 2|zw| + |zw|^2\right) - (zw^* + z^*w + 2|zw|)\\ &= (1 + |zw|)^2 - E,\end{aligned}$$

得證。

### 最後的不等式

設

$$A = (|z| + |w|)^2, \qquad B = (1 + |zw|)^2,$$

於是第 (i) 及 (ii) 部分表明 $|z - w|^2 = A - E$ 及 $|1 - zw^*|^2 = B - E$。

首先，左方是有意義的：若 $|z| > 1$ 且 $|w| > 1$，則 $|zw^*| = |z||w| > 1$，所以 $zw^* \neq 1$，且 $B - E = |1 - zw^*|^2 > 0$。此外 $B > 0$。

所求不等式的兩邊都是非負的，所以它等價（equivalent）於兩邊平方後的不等式

$$\frac{A - E}{B - E} \leq \frac{A}{B}.$$

兩邊同乘以正數 $B(B - E)$，這等價於

$$(A - E)B \leq A(B - E) \iff -EB \leq -EA \iff E(B - A) \geq 0.$$

> **想一想** $B - A$ 的正負號是甚麼？關於 $|z|$ 和 $|w|$ 的條件在哪裏發揮作用？

設 $a = |z|$ 及 $b = |w|$，

$$B - A = (1 + ab)^2 - (a + b)^2 = 1 + a^2b^2 - a^2 - b^2 = (1 - a^2)(1 - b^2).$$

若 $a > 1$ 且 $b > 1$，兩個括號都是負數，所以 $B - A > 0$。由於第 (i) 部分已證 $E \geq 0$，故 $E(B - A) \geq 0$。上面每一步都是等價的，所以

$$\frac{|z - w|}{|1 - zw^*|} \leq \frac{|z| + |w|}{1 + |zw|}.$$

等號成立若且唯若 $E = 0$，即 $z$ 和 $w$ 指向相反方向。

**當 $|z|<1$ 且 $|w|<1$ 時**

是的，不等式仍然成立。此時 $|zw^*| = ab < 1$，所以同樣有 $zw^* \neq 1$ 及 $B - E > 0$，整條等價推導鏈原封不動地成立。這次 $(1-a^2)(1-b^2)$ 中兩個括號都是正數，所以同樣有 $B - A > 0$，得到 $E(B-A) \geq 0$，從而得出該不等式。

關鍵只在於 $1-a^2$ 與 $1-b^2$ 同號。若一個模大於 1 而另一個小於 1，則 $B - A < 0$，不等式可能不成立：取 $z = 2$、$w = 0.4$，左方為 $1.6/0.2 = 8$，而右方為 $2.4/1.8 = \frac{4}{3}$。

## 驗算

- 取 $z = 2$、$w = 2\mathrm{i}$。則 $|z-w| = 2\sqrt{2}$，而 $1 - zw^* = 1 - 2(-2\mathrm{i}) = 1 + 4\mathrm{i}$，所以左方為 $2\sqrt{2}/\sqrt{17} \approx 0.686$，右方為 $4/5 = 0.8$。把兩者都乘以 $\frac{1}{4}$（$z = \frac{1}{2}$、$w = \frac{1}{2}\mathrm{i}$）會得到同樣的兩個數，$0.686 \leq 0.8$，這說明了 $|z|, |w| < 1$ 的情況。
- 等號情況：$z = 2 + \mathrm{i}$、$w = -1.5z$ 給出 $E = 0$，兩邊準確至四位小數都等於 0.6577。
- 用電腦對數十萬組隨機數對進行檢驗，證實了兩條恆等式、$E \geq 0$，以及兩種情況下的不等式。

## 常見錯誤

- 寫成 $|z-w|^2 = |z|^2 + |w|^2 - 2zw^*$：交叉項是 $zw^* + z^*w$，兩者互為共軛，但一般而言並不相等。
- 交叉相乘前沒有檢查 $|1 - zw^*| \neq 0$；這裏 $|zw| \neq 1$ 保證了這一點。
- 因為推導時用了 $|z|, |w| > 1$ 而在最後一部分回答「否」：因式分解（factorisation）$(1-a^2)(1-b^2)$ 顯示，關鍵只在於正負號是否一致。

## STEP 之外

由於 $|1 - zw^*| = |1 - z^*w|$，左方是 $\dfrac{z-w}{1-z^*w}$ 的模。對於固定的 $z$（其中 $|z| < 1$），映射 $w \mapsto \dfrac{z-w}{1-z^*w}$ 是一個把單位圓盤（unit disc）映到自身的莫比烏斯變換（Möbius transformation），而 $\left|\frac{z-w}{1-z^*w}\right|$ 是 $z$ 與 $w$ 之間的偽雙曲距離（pseudo-hyperbolic distance）。本題的不等式表明，這個距離不超過 $z$ 和 $w$ 各自到 0 的距離的「雙曲和」：這是三角不等式的一種加強形式，也是雙曲幾何（hyperbolic geometry）的基礎之一。Needham 的 *Visual Complex Analysis*（《直觀複分析》）以精美的圖像解說莫比烏斯變換和雙曲圓盤。

詳細解答

# STEP III 2013，第 7 題

純數學 • 常微分方程的能量法

> **STEP III 2013，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/13-s3-q7

本題在 *Beyond STEP* 第 8.2 節*不求解而估界*中有完整的引導式講解：https://beyondstep.keithmcnulty.org/zh-hant/ch08/bounding-without-solving.html

## 關鍵思路

每條方程都像一粒繫在彈簧上的質點：一個只依賴於未知函數的「力」項，再加上（在 (ii) 和 (iii) 中）一個與導數（derivative）成正比的摩擦項。將方程乘以導數，就可以把它化成某個「能量」的導數。沒有摩擦時，能量守恆；若摩擦從不朝錯誤方向推動，能量就只會減少，因此永遠不會超過其初始值。第 (iii) 部分要求我們自行構造能量，並證明摩擦係數從不為負。

## 解答

### 第 (i) 部分

記 $y' = \frac{\mathrm{d}y}{\mathrm{d}x}$。由鏈式法則（chain rule），

$$\begin{aligned}\frac{\mathrm{dE}}{\mathrm{d}x} &= 2y'y'' + 2y^3y' \\ &= 2y'\left(y'' + y^3\right) = 0,\end{aligned}$$

這裏用了微分方程（differential equation）。因此 E 是常數，其值為 $\mathrm{E}(0) = 0^2 + \frac{1}{2}\cdot 1^4 = \frac{1}{2}$。故對每個 $x$，

$$\tfrac{1}{2}y^4 = \tfrac{1}{2} - (y')^2 \le \tfrac{1}{2},$$

因為 $(y')^2 \ge 0$。所以 $y^4 \le 1$，即對所有 $x$ 都有 $|y(x)| \le 1$。

**第 (ii) 部分**

用同樣方法求導，再把方程寫成 $v'' + \sinh v = -xv'$ 的形式代入，

$$\begin{aligned}\frac{\mathrm{dE}}{\mathrm{d}x} &= 2v'v'' + 2\sinh v\, v' \\ &= 2v'\big(v'' + \sinh v\big) \\ &= -2x\,(v')^2.\end{aligned}$$

當 $x \geq 0$ 時，$x$ 與 $(v')^2$ 均非負，所以 $\frac{\mathrm{dE}}{\mathrm{d}x} \leq 0$，正如所求。

因此 E 在 $x \geq 0$ 上非遞增（non-increasing），並在該範圍內有 $\mathrm{E}(x) \leq \mathrm{E}(0)$。在 $x = 0$ 處，$\cosh(\ln 3) = \frac{1}{2}\big(3 + \frac{1}{3}\big) = \frac{5}{3}$，所以 $\mathrm{E}(0) = 0 + 2 \cdot \frac{5}{3} = \frac{10}{3}$。因此，當 $x \geq 0$ 時，

$$2\cosh v(x) = \mathrm{E}(x) - (v')^2 \leq \mathrm{E}(x) \leq \tfrac{10}{3},$$

故 $\cosh v(x) \leq \frac{5}{3}$。

**第 (iii) 部分**

仿照 (ii)，我們尋找形如 $\mathrm{E} = (w')^2 + F(w)$ 的能量。於是

$$\frac{\mathrm{dE}}{\mathrm{d}x} = 2w'w'' + F'(w)\,w' = w'\big(2w'' + F'(w)\big),$$

我們希望借助方程，使括號內變成 $w'$ 的倍數。只要 $F'(w) = 2(w\cosh w + 2\sinh w)$，就能做到這一點。

> **想一想** 哪個 $w$ 的函數的導數是 $w\cosh w + 2\sinh w$？

用分部積分（integration by parts），$\int w\cosh w\,\mathrm{d}w = w\sinh w - \cosh w$，而 $\int 2\sinh w\,\mathrm{d}w = 2\cosh w$。兩者相加，

$$\frac{\mathrm{d}}{\mathrm{d}w}\big(w\sinh w + \cosh w\big) = \sinh w + w\cosh w + \sinh w = w\cosh w + 2\sinh w.$$

所以我們取

$$\mathrm{E}(x) = \left(\frac{\mathrm{d}w}{\mathrm{d}x}\right)^2 + 2\big(w\sinh w + \cosh w\big).$$

記 $g(x) = 5\cosh x - 4\sinh x - 3$，方程即為 $w'' + (w\cosh w + 2\sinh w) = -g(x)\,w'$，所以

$$\frac{\mathrm{dE}}{\mathrm{d}x} = 2w'\big(w'' + w\cosh w + 2\sinh w\big) = -2g(x)\,(w')^2.$$

> **想一想** $g(x)$ 會不會是負數？

用指數函數表示，

$$\begin{aligned} g(x) &= \tfrac{5}{2}\left(\mathrm{e}^{x}+\mathrm{e}^{-x}\right)-2\left(\mathrm{e}^{x}-\mathrm{e}^{-x}\right)-3 \\ &= \tfrac{1}{2}\mathrm{e}^{x}-3+\tfrac{9}{2}\mathrm{e}^{-x} \\ &= \tfrac{1}{2}\mathrm{e}^{-x}\left(\mathrm{e}^{2x}-6\mathrm{e}^{x}+9\right)=\tfrac{1}{2}\mathrm{e}^{-x}\left(\mathrm{e}^{x}-3\right)^{2} \;\geq\; 0 \end{aligned}$$

對每個 $x$ 都成立（只在 $x=\ln 3$ 處等於零）。因此 $\frac{\mathrm{dE}}{\mathrm{d}x}\leq 0$，並且當 $x\geq 0$ 時 $\mathrm{E}(x)\leq \mathrm{E}(0)$。在 $x=0$ 處，$w=0$ 且 $w'=\frac{1}{\sqrt{2}}$，所以 $\mathrm{E}(0)=\frac{1}{2}+2(0+1)=\frac{5}{2}$。因此，當 $x\geq 0$ 時，

$$2\left(w\sinh w+\cosh w\right)\leq \mathrm{E}(x)\leq \tfrac{5}{2}, \qquad \text{所以} \qquad w\sinh w+\cosh w\leq \tfrac{5}{4}.$$

最後，對每個實數 $w$ 都有 $w\sinh w\geq 0$，因為 $w$ 與 $\sinh w$ 總是同號。略去這一項，便得到當 $x\geq 0$ 時 $\cosh w(x)\leq \frac{5}{4}$，正如所求。

## 驗算

- 在 (ii) 中，這個界在起點處是緊的（sharp）：$\cosh v(0)=\cosh(\ln 3)=\frac{5}{3}$。
- 在 (iii) 中，$g(0)=5-0-3=2$，與因式分解形式所得的 $\frac{1}{2}(1-3)^2=2$ 一致。
- 對每條方程求數值解，都印證了上述圖像：在 (i) 中 E 一直保持 0.5；在 (ii) 中 E 從 $\frac{10}{3}$ 開始下降；在 (iii) 中 E 從 2.5 開始下降，而 $\cosh w$ 從不超過約 1.03，遠在界 1.25 之內。

## 常見錯誤

- 把 $F(w)$ 對 $x$ 求導時，忘記鏈式法則帶來的因子 $w'$。
- 在 (iii) 中，停在 $w\sinh w+\cosh w\leq\frac{5}{4}$，卻沒有解釋為何 $w\sinh w\geq 0$ 使我們可以略去該項。
- 聲稱 $g(x)>0$：它在 $x=\ln 3$ 處等於零，所以只有 $\leq$ 成立（而這已足夠我們所需）。

## STEP 之外

這些能量函數就是李雅普諾夫函數（*Lyapunov functions*）：沿着解從不增加的量，可用來在不求解的情況下證明一個系統保持有界或逐漸趨於穩定。同一個想法——能量守恆加上阻尼造成的耗散——正是非線性振子（nonlinear oscillators）穩定性理論的基礎。Boyce 與 DiPrima 的 *Elementary Differential Equations*（《初等微分方程》）在講述非線性系統與穩定性的一章中介紹了李雅普諾夫方法。

詳細解答

# STEP III 2013，第 8 題

純數學 • 圓錐曲線的極坐標方程

**STEP III 2013，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/13-s3-q8

## 關鍵思路

關係式 $r = ks$（其中 $k < 1$）表明 $E$ 是一個以 $O$ 為焦點（focus）、以 $D$ 為準線（directrix）的橢圓（ellipse）；若採用以 $O$ 為中心的極坐標（polar coordinates），它便化為一條以 $\theta$ 表示 $r$ 的非常簡單的方程。一條經過 $O$ 的弦（chord）其實就是兩條方向相反的半徑，因此 $1/l$ 會化為一個關於 $\cos^2\theta$ 的簡單式子。接着，$n$ 條等距分佈的直線把總和轉化為一些 $\cos 2\theta$ 項之和，而這正是開首那個求和所處理的。

## 解答

### 開首的求和

各項組成一個等比數列（geometric progression），首項為 $\mathrm{e}^{2\mathrm{i}\alpha}$，公比（common ratio）為 $w = \mathrm{e}^{2\pi\mathrm{i}/n}$。由於 $n \geq 2$，角 $2\pi/n$ 位於 $(0, \pi]$ 內，所以 $w \neq 1$，我們可以使用一般的公式：

$$\sum_{r=0}^{n-1} \mathrm{e}^{2\mathrm{i}(\alpha + r\pi/n)} = \mathrm{e}^{2\mathrm{i}\alpha}\,\frac{1-w^n}{1-w} = \mathrm{e}^{2\mathrm{i}\alpha}\,\frac{1-\mathrm{e}^{2\pi\mathrm{i}}}{1-w} = 0.$$

取實部（real part），得 $\sum_{r=0}^{n-1} \cos\bigl(2\alpha + 2r\pi/n\bigr) = 0$；我們將會用到的正是這個形式。

### 距離 $s$

如圖所示，從 $O$ 到 $D$ 的垂線起量度 $\theta$。該垂線的垂足與 $O$ 的距離為 $d$，而 $P$ 在同一方向上與 $O$ 的距離為 $r\cos\theta$，因此當 $P$ 與 $O$ 位於 $D$ 的同一側時，

$$s = d - r\cos\theta.$$

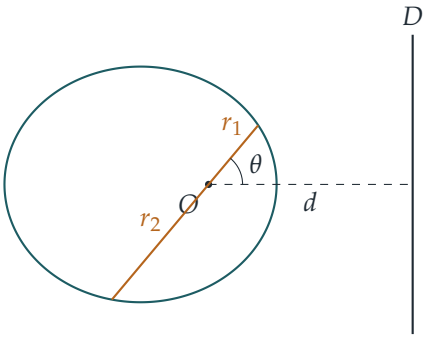
（這對所有 $\theta$ 均成立：當 $\cos\theta < 0$ 時，$P$ 比 $O$ 離 $D$ 更遠，而公式亦正確地給出 $s > d$。）

### $E$ 的極坐標方程

$E$ 上每一點都與 $O$ 位於 $D$ 的同一側。事實上，若 $P$ 位於另一側，則有 $s = r\cos\theta - d$，而 $r = ks$ 會給出 $r(1 - k\cos\theta) = -kd < 0$；這是不可能的，因為當 $0 < k < 1$ 時 $1 - k\cos\theta > 0$。因此我們可以用 $s = d - r\cos\theta$，於是 $r = ks$ 變為

$$r = k(d - r\cos\theta) \quad \Longrightarrow \quad r = \frac{kd}{1 + k\cos\theta}.$$

由於 $k < 1$，分母對所有 $\theta$ 均為正，所以從 $O$ 出發的每條射線與 $E$ 恰好相交一次。




### 經過 $O$ 的弦的長度

一條經過 $O$、與垂線成角 $\theta$ 的直線，與 $E$ 相交於方向 $\theta$ 及 $\theta + \pi$ 上的兩點，而這兩點位於 $O$ 的兩側。由於 $\cos(\theta + \pi) = -\cos\theta$，弦長為

$$l = \frac{kd}{1 + k\cos\theta} + \frac{kd}{1 - k\cos\theta} = \frac{2kd}{1 - k^2\cos^2\theta}, \qquad \text{因此} \qquad \frac{1}{l} = \frac{1 - k^2\cos^2\theta}{2kd}.$$

留意這只取決於直線本身，而與我們把它的哪一個方向稱為 $\theta$ 無關。

**想一想** 我們應如何把 $\cos^2\theta$ 對 $n$ 條直線求和？

利用倍角公式（double-angle formula）$\cos^2\theta = \frac{1}{2}(1 + \cos 2\theta)$：那些 $\cos 2\theta$ 項正好就是開首求和的實部。

### 對 $n$ 條直線求和

由於相鄰直線之間的夾角為 $\pi/n$，對某個固定的 $\alpha$，這 $n$ 條直線所成的角為 $\theta_j = \alpha + j\pi/n$（$j = 1, \ldots, n$）。這 $n$ 個 $2\theta_j$ 的值是 $2\alpha + 2j\pi/n$；把 $j = n$ 換成 $j = 0$（這只會令 $2\theta_j$ 改變 $2\pi$）後，它們正是開首求和中的各個角。因此

$$\sum_{j=1}^{n} \cos^2\theta_j = \frac{n}{2} + \frac{1}{2}\sum_{j=1}^{n} \cos 2\theta_j = \frac{n}{2} + 0 = \frac{n}{2},$$

從而

$$\sum_{j=1}^{n} \frac{1}{l_j} = \frac{1}{2kd}\left(n - k^2 \cdot \frac{n}{2}\right) = \frac{(2-k^2)n}{4kd},$$

這就是所求的結果。條件 $n \geq 2$ 只用過一次，就是用來確保開首求和中 $w \neq 1$。

## 驗算

- 取 $n = 2$、$k = \frac{1}{2}$、$d = 1$，兩條直線分別沿着 $OD$ 及垂直於 $OD$。沿 $OD$ 的弦是長軸（major axis），長度為 $\frac{1/2}{3/2} + \frac{1/2}{1/2} = \frac{4}{3}$；垂直於它的弦（即正焦弦（latus rectum））長度為 $2kd = 1$。所以總和為 $\frac{3}{4} + 1 = \frac{7}{4}$，而公式給出 $\frac{(7/4)\cdot 2}{2} = \frac{7}{4}$。
- 當 $n = 1$ 時，結果並不成立，這是理所當然的：單一條直線的 $1/l$ 取決於它的方向。而 $n = 1$ 正是 $w = 1$ 的情況。
- 我們用電腦對隨機的 $d$、$k$、$n$ 及 $\alpha$ 直接從 $r = ks$ 求出每條弦作驗算，結果與公式吻合；例如 $d = 1$、$k = \frac{1}{2}$、$n = 3$ 時，對任何 $\alpha$ 都得到 $2.625 = \frac{21}{8}$。

## 常見錯誤

- 寫成 $s = d + r\cos\theta$：在圖中 $\theta$ 是朝 $D$ 的方向量度的，所以沿 $OP$ 向外移動會令你更接近 $D$。
- 把弦的另一半再次當作 $kd/(1 + k\cos\theta)$，忘記了相反方向的角是 $\theta + \pi$。
- 直接把角 $\alpha + j\pi/n$ 求和，而不是對 $2\theta_j$ 求和：均勻地繞圓周一圈的是倍角，這正是開首求和中出現因子 2 的原因。

## STEP 之外

焦點－準線定義使每一條圓錐曲線（conic）都有極坐標方程 $r = \ell/(1 + e\cos\theta)$，其中 $e < 1$、$e = 1$ 及 $e > 1$ 分別對應橢圓、拋物線（parabola）及雙曲線（hyperbola）；當我們解牛頓（Newton）的行星運動方程時，開普勒（Kepler）的軌道正是以這個形式出現。任意兩條互相垂直的焦點弦（focal chord）的 $1/l_1 + 1/l_2$ 都相同，這是一個經典性質，而本題的結果正是它的推廣。Kleppner 與 Kolenkow 所著的 *An Introduction to Mechanics*（《力學導論》）便是用這種方法推導軌道方程的。

詳細解答

# STEP III 2013，第 9 題

力學 • 浮力與簡諧運動

**STEP III 2013，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/13-s3-q9

## 關鍵思路

球體浸沒在液體中的部分是一個球冠（spherical cap），其體積可以把它切成一片片圓盤求得，這就是旋轉體體積（volume of revolution）的計算。之後，題目就是應用牛頓第二定律（Newton's second law），除了重力之外，唯一的力就是浮力（buoyancy）。對於小幅振動，我們設 $x = x_0 + y$，把右邊的三次式在平衡位置附近線性化（linearise），再從形如 $\ddot{y} = -\omega^2 y$ 的方程讀出 $\omega^2$。

## 解答

### 排開的體積

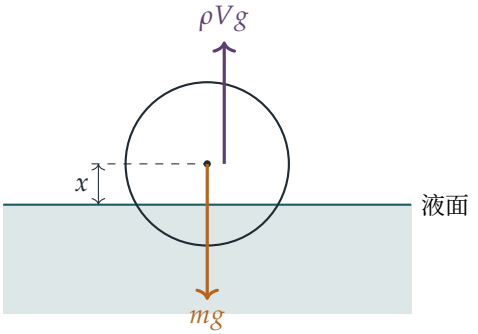
取球心為原點，$z$ 向上量度。液面就是平面 $z = -x$，浸沒部分就是球體中 $-R \le z \le -x$ 的部分。這對 $-R < x < R$ 有意義，因此涵蓋了題目所述的情況。在高度 $z$ 處，截面是半徑為 $\sqrt{R^2 - z^2}$ 的圓盤，所以

$$\begin{aligned} V &= \int_{-R}^{-x} \pi\left(R^2 - z^2\right) \mathrm{d}z = \pi\left[R^2 z - \tfrac{1}{3}z^3\right]_{-R}^{-x} \\ &= \pi\left(-R^2 x + \tfrac{1}{3}x^3 + R^3 - \tfrac{1}{3}R^3\right) = \frac{\pi}{3}\left(2R^3 - 3R^2 x + x^3\right), \end{aligned}$$

這就是所求的體積。

### 運動方程

把球體視為作鉛直運動的剛體（rigid body）。由於水缸很大，球體上下浮動時液面不會移動，所以 $\ddot{x}$ 就是球體向上的加速度。球體的質量為 $m = \frac{4}{3}\pi R^3 \rho_s$。作用在球體上的力有兩個：向下的重量 $mg$，以及向上的浮力 $\rho V g$。




沿鉛直方向分解，以向上為正，

$$m\ddot{x} = \rho Vg - mg \quad \Longrightarrow \quad m\,(g + \ddot{x}) = \rho g V.$$

代入 $m$ 和 $V$：

$$\tfrac{4}{3}\pi R^3 \rho_s\,(g + \ddot{x}) = \tfrac{\pi}{3}\left(2R^3 - 3R^2 x + x^3\right)\rho g,$$

兩邊乘以 $3/\pi$，得

$$4R^3 \rho_s\,(g + \ddot{x}) = \left(2R^3 - 3R^2 x + x^3\right)\rho g,$$

此即所求。

### 平衡時的密度

在平衡（equilibrium）時 $\ddot{x} = 0$。代入 $x = \frac{1}{2}R$：

$$4R^3 \rho_s g = \left(2R^3 - \tfrac{3}{2}R^3 + \tfrac{1}{8}R^3\right)\rho g = \tfrac{5}{8}R^3 \rho g,$$

所以

$$\rho_s = \frac{5}{32}\,\rho.$$

這個值小於 $\frac{1}{2}\rho$，理應如此：球心在液面之上，所以浸沒在液體中的部分少於球體的一半。

### 小幅振動

設 $f(x) = 2R^3 - 3R^2 x + x^3$，並設 $x = \frac{1}{2}R + y$，其中 $y$ 很小。於是 $\ddot{x} = \ddot{y}$。

> ? **想一想** 怎樣把這個三次式變成關於 $y$ 的線性式？

我們把 $f$ 在平衡位置附近展開，只保留一階項：$f(\frac{1}{2}R + y) = f(\frac{1}{2}R) + f'(\frac{1}{2}R)\,y + O(y^2)$。這裏

$$f'(x) = -3R^2 + 3x^2, \qquad f'(\tfrac{1}{2}R) = -3R^2 + \tfrac{3}{4}R^2 = -\tfrac{9}{4}R^2,$$

而由前面的計算，$f(\frac{1}{2}R) = \frac{5}{8}R^3$。利用 $\rho_s = \frac{5}{32}\rho$，運動方程的左邊為 $4R^3 \cdot \frac{5}{32}\rho\,(g+\ddot{y}) = \frac{5}{8}R^3\rho\,(g+\ddot{y})$。因此，準確至一階，

$$\tfrac{5}{8}R^3\rho\,(g+\ddot{y}) = \left(\tfrac{5}{8}R^3 - \tfrac{9}{4}R^2 y\right)\rho g.$$

兩邊的 $\frac{5}{8}R^3\rho g$ 項互相抵消（這正是平衡條件），剩下

$$\tfrac{5}{8}R^3\ddot{y} = -\tfrac{9}{4}R^2 g\, y \quad\Longrightarrow\quad \ddot{y} = -\frac{18g}{5R}\,y.$$

這是簡諧運動（simple harmonic motion），其中 $\omega^2 = \dfrac{18g}{5R}$，所以小幅振動的週期（period）為

$$T = \frac{2\pi}{\omega} = 2\pi\sqrt{\frac{5R}{18g}}.$$

## 驗算

- 體積：當 $x = R$ 時，公式給出 0（球體剛好接觸液面）；當 $x = 0$ 時給出 $\frac{2}{3}\pi R^3$（半個球體）；當 $x = -R$ 時給出 $\frac{4}{3}\pi R^3$（完全浸沒）。
- 量綱（dimensions）：$\omega^2 = 18g/(5R)$ 的單位是 $\mathrm{s}^{-2}$，理應如此。
- 數值：取 $R = 1\,\mathrm{m}$ 及 $g = 9.81\,\mathrm{m\,s^{-2}}$，週期約為 $1.06\,\mathrm{s}$。用電腦從一個小位移出發，逐步推進完整的非線性方程，得到的週期與此相同，誤差小於萬分之一（$10^4$ 分之一）。

## 常見錯誤

- 把浸沒球冠的高度當作 $x$ 而不是 $R - x$：$x$ 是球心在液面之上的高度。
- 忘記回復力來自 $f'(\frac{1}{2}R)$，而它是負數。這裏出現正負號錯誤會得到 $\ddot{y} = +\omega^2 y$，那根本不是振動。
- 漏掉左邊球體本身的質量：與 $\ddot{x}$ 相乘的是 $\rho_s$，而不是 $\rho$。（現實中的浮體還會帶動周圍部分液體一起運動，即所謂「附加質量」（added mass），但題目叫我們忽略這一點。）

## STEP 之外

這裏所用的方法——把力在平衡位置附近展開並保留線性項——是求任何系統小幅振動的標準方法。線性項的正負號決定穩定性，其大小則決定頻率。在大學一年級的課程中，

這會發展為 $\ddot{x} = F(x)$ 的線性化，然後在有多個自由度（degrees of freedom）時，進一步研究簡正模態（normal modes）。Kleppner 與 Kolenkow 合著的 *An Introduction to Mechanics*（《力學導論》）以這個程度講解了平衡位置附近的小幅振動及流體浮力。

詳細解答

# STEP III 2013，第 10 題

力學 • 轉動撞擊

**STEP III 2013，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：

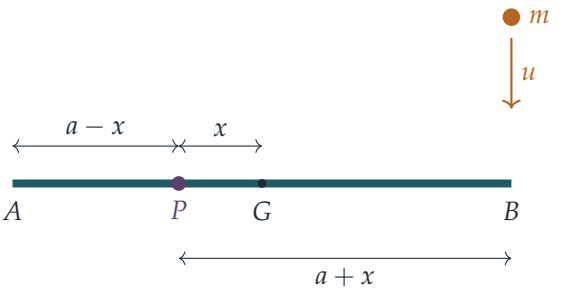
step.maths.org/questions/13-s3-q10

## 關鍵思路

只要留意到桿的中點與 $P$ 相距多遠，轉動慣量（moment of inertia）便可直接由平行軸定理（parallel axis theorem）得出。至於撞擊，轉軸會對桿施加一個大小未知的衝量（impulse），因此線動量（linear momentum）並不守恆；但該衝量作用於 $P$，所以它對 $P$ 的力矩為零。對 $P$ 的角動量守恆（conservation of angular momentum），再加上在 $B$ 處應用牛頓（Newton）實驗定律（Newton's experimental law），便得到兩條方程，足以求出兩個未知量。

## 解答

我們把桿視為均勻剛性桿，把質點視為質點，並假設撞擊在瞬間完成，因此唯一的水平衝量來自 $B$ 處和轉軸處（重力等有限大小的力在撞擊的一瞬間不產生衝量）。



### 轉動慣量

設 $G$ 為桿的中點，即其質心（centre of mass）。由於 $AG = a$ 及 $AP = a - x$，距離 $PG$ 為 $x$（當 $x > 0$ 時，$G$ 位於 $P$ 靠 $B$ 的一側）。均勻桿對通過其中心的垂直軸的轉動慣量為 $\frac{1}{3}Ma^2$（桿長為 $2a$）。由平行軸定理，

$$I_P = \tfrac{1}{3}Ma^2 + Mx^2 = \tfrac{1}{3}M(a^2 + 3x^2),$$

與所求相符。

如果你喜歡從基本原理出發：線密度（line density）為 $M/(2a)$，以 $s$ 表示沿桿從 $P$ 量起的距離，則桿由 $s = -(a - x)$ 延伸至 $s = a + x$，而

$$I_P = \frac{M}{2a}\int_{-(a-x)}^{a+x} s^2\,\mathrm{d}s = \frac{M}{6a}\left[(a+x)^3 + (a-x)^3\right] = \frac{M}{6a}(2a^3 + 6ax^2),$$

結果相同。

### 撞擊

由 $P$ 到 $B$ 的距離為 $2a - (a - x) = a + x$。設 $v$ 為質點在撞擊後瞬間的速度（以原來的運動方向為正），$\omega$ 為桿的角速度（angular velocity），於是 $B$ 此時以速率 $(a + x)\omega$ 沿同一方向運動。

> **想一想** 為甚麼這裏不能用線動量守恆？

因為轉軸是固定的，它在碰撞期間對桿施加一個大小未知的衝量。該衝量作用於 $P$，所以它對 $P$ 的力矩為零。質點與桿之間的衝量是桿與質點所組成系統的內部衝量。因此，系統對 $P$ 的總角動量守恆：

$$mu(a + x) = mv(a + x) + I_P\,\omega. \tag{1}$$

在 $B$ 處應用牛頓實驗定律（分離速率等於 $e$ 乘以接近速率），得

$$(a + x)\omega - v = eu. \tag{2}$$

由 (2)，$v = (a + x)\omega - eu$。代入 (1)，

$$\begin{aligned} mu(a + x) &= m(a + x)\left[(a + x)\omega - eu\right] + I_P\,\omega \\ mu(1 + e)(a + x) &= \left[m(a + x)^2 + \tfrac{1}{3}M(a^2 + 3x^2)\right]\omega. \end{aligned}$$

兩邊乘以 3，

$$\omega = \frac{3mu(1 + e)(a + x)}{M(a^2 + 3x^2) + 3m(a + x)^2},$$

即題目所給的結果。

**$m = 2M$ 的情況**

代入 $m = 2M$ 並約去 $M$：

$$\omega = \frac{6u(1+e)(a+x)}{a^2+3x^2+6(a+x)^2}.$$

自然出現的變量是力臂（lever arm）$s = a + x$，即距離 $PB$，它的取值範圍是 $[0, 2a]$。以 $x = s - a$ 代入，分母為

$$a^2 + 3(s-a)^2 + 6s^2 = 9s^2 - 6as + 4a^2,$$

所以

$$\omega = \frac{6u(1+e)\,s}{9s^2 - 6as + 4a^2}.$$

當 $s = 0$（轉軸位於 $B$）時 $\omega = 0$，所以最大值出現在 $s > 0$ 處，此時可把分子分母同除以 $s$：

$$\omega = \frac{6u(1+e)}{9s + \dfrac{4a^2}{s} - 6a}.$$

我們要令分母盡可能小。由於 $s > 0$，算術—幾何平均不等式（AM–GM inequality）給出

$$9s + \frac{4a^2}{s} \;\geq\; 2\sqrt{9s \cdot \frac{4a^2}{s}} = 12a,$$

等號成立當且僅當 $9s = 4a^2/s$，即 $s = \frac{2}{3}a$，而這確實落在 $(0, 2a]$ 之內。因此分母至少為 $12a - 6a = 6a$，且只在 $s = \frac{2}{3}a$ 時取等號，故

$$\omega \leq \frac{6u(1+e)}{6a} = \frac{u(1+e)}{a}.$$

所以最大角速度為 $u(1+e)/a$，在 $a + x = \frac{2}{3}a$ 時達到，即

$$x = -\tfrac{1}{3}a.$$

負號只是表示 $P$ 位於中點靠 $B$ 的一側：$P$ 與 $A$ 相距 $\frac{4}{3}a$，與 $B$ 相距 $\frac{2}{3}a$。

## 驗算

- 當 $x = a$（轉軸位於 $A$）且 $e = 1$ 時，公式給出 $\omega = 3mu/\big(a(M+3m)\big)$；令 $M \to 0$，得 $\omega = u/a$，於是 $B$ 以 $2u$ 運動，這正是極輕的「靶子」受彈性撞擊時大家熟悉的結果。
- 在最佳情況下取 $a = 1$、$u = 1$、$e = \frac{1}{2}$，桿以 $\omega = 1.5$ 轉動，所以 $B$ 以 1 運動，而質點在撞擊後的速度為 $v = 1 - \frac{1}{2} = \frac{1}{2}$。此時對 $P$ 的角動量為 $2M \cdot \frac{1}{2} \cdot \frac{2}{3} + \frac{1}{3}M\big(1 + \frac{1}{3}\big) \cdot 1.5 = \frac{4}{3}M$，與撞擊前的 $2M \cdot 1 \cdot \frac{2}{3}$ 相符。
- 以電腦在 $x \in [-a, a]$ 上作細密搜尋，確認最大值出現在 $x = -a/3$。

## 常見錯誤

- 使用線動量守恆：固定的轉軸會產生衝量，所以只有對 $P$ 的角動量守恆。
- 把距離 $PB$ 當作 $a - x$ 或 $2a - x$，而不是 $a + x$。
- 因 $x = -\frac{1}{3}a$ 是負數而捨棄它：$x$ 是帶正負號的位置，任何 $[-a, a]$ 內的 $x$ 都會使 $P$ 落在桿上。

## STEP 之外

選擇適當的點來取力矩，使未知的反作用力消去，是剛體動力學（rigid-body dynamics）中最有用的技巧之一；它在「撞擊中心」（centre of percussion）中再次出現——那是球棒或桿上的一點，打擊該點時，支點處不會產生衝量。大學一年級的力學會把這發展成角動量和慣性張量（inertia tensor）的一般理論。Kleppner 與 Kolenkow 的 *An Introduction to Mechanics*（《力學導論》）對衝擊性轉動和撞擊中心有非常清晰的論述。

詳細解答

# STEP III 2013，第 11 題

力學 • 彈簧與能量

**STEP III 2013，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/13-s3-q11

## 關鍵思路

一切都由對稱性（symmetry）決定：三角形保持水平，其中心一直位於懸掛點的正下方，而且只能上下移動。因此只需考慮一個質點和一條彈簧，便能掌握整個情況。經過一條彈簧的鉛垂平面上的幾何關係，給出彈簧長度關於 $\theta$ 的表達式。沿鉛垂方向分解力可得平衡（equilibrium）條件，而能量守恆（conservation of energy）（彈性勢能、重力勢能和動能）則給出速率。

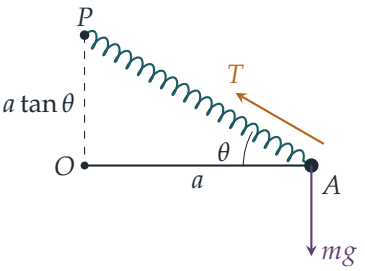
## 解答

### 模型與幾何

把各質量視為質點（particle），各桿視為輕而剛性，各彈簧視為輕彈簧並遵守胡克定律（Hooke's law）。設 $P$ 為懸掛點，$O$ 為三角形的中心，位於 $P$ 的正下方。邊長為 $s$ 的等邊三角形，其中心到頂點的距離為 $s/\sqrt{3}$，所以每個頂點與 $O$ 的水平距離為

$$OA = \frac{\sqrt{3}a}{\sqrt{3}} = a$$

取一個頂點 $A$，考慮經過 $P$、$O$ 和 $A$ 的鉛垂平面：




彈簧 $PA$ 與水平線 $OA$ 成角 $\theta$，所以彈簧長度為 $a/\cos\theta$，而 $A$ 位於 $P$ 下方 $a\tan\theta$ 的深度。

### 張力

每條彈簧的伸長量（extension）為

$$x = \frac{a}{\cos\theta} - a = \frac{a(1-\cos\theta)}{\cos\theta},$$

因此由胡克定律，配合彈性模量（modulus）$kmg$ 和自然長度（natural length）$a$，得

$$T = \frac{kmg}{a}\,x = \frac{kmg(1-\cos\theta)}{\cos\theta},$$

即所求結果。

### $\theta = \frac{1}{6}\pi$ 時的平衡

作用於 $A$ 處質點的力包括其重量、沿 $AP$ 方向的彈簧張力（tension），以及兩條桿施加的力。各桿位於水平面內，所以桿的作用力是水平的。

> ? **想一想** 為甚麼可以不理會桿的作用力？

我們沿鉛垂方向分解，而桿的作用力沒有鉛垂分量。（在水平方向上，張力分量 $T\cos\theta$ 由桿的作用力平衡，這正是需要桿的原因，但我們用不着這條方程。）在 $A$ 處沿鉛垂方向分解：

$$T\sin\theta = mg \quad\Longrightarrow\quad k(1-\cos\theta)\tan\theta = 1.$$

當 $\theta = \frac{1}{6}\pi$ 時，$\cos\theta = \frac{\sqrt{3}}{2}$，$\tan\theta = \frac{1}{\sqrt{3}}$，所以

$$k\cdot\frac{2-\sqrt{3}}{2}\cdot\frac{1}{\sqrt{3}} = 1 \quad\Longrightarrow\quad k = \frac{2\sqrt{3}}{2-\sqrt{3}} = \frac{2\sqrt{3}(2+\sqrt{3})}{(2-\sqrt{3})(2+\sqrt{3})} = 4\sqrt{3}+6,$$

因為 $(2-\sqrt{3})(2+\sqrt{3}) = 1$。這就是所求的值。

### 在平衡位置時的速率

由對稱性，三角形保持水平，中心位於 $P$ 的正下方，所以三個質點都以相同速率作鉛垂運動。除重力和彈簧外，沒有其他外力做功，所以能量守恆。伸長量為 $x$ 的彈簧所儲存的彈性勢能為 $kmg\,x^2/(2a)$。

起始位置（$\theta = \frac{1}{3}\pi$，靜止）。此時 $\cos\theta = \frac{1}{2}$，所以伸長量為 $x_1 = a$，而在 $P$ 下方的深度為 $a\tan\frac{1}{3}\pi = \sqrt{3}a$。

平衡位置（$\theta = \frac{1}{6}\pi$）。伸長量為

$$x_2 = a\left(\frac{2}{\sqrt{3}} - 1\right) = \frac{a(2-\sqrt{3})}{\sqrt{3}}, \qquad x_2^2 = \frac{a^2(7-4\sqrt{3})}{3},$$

深度為 $a\tan\frac{1}{6}\pi = \frac{a}{\sqrt{3}}$。所以每個質點上升了

$$\sqrt{3}a - \frac{a}{\sqrt{3}} = \frac{2a}{\sqrt{3}}.$$

三條彈簧釋放的彈性勢能：

$$\begin{aligned} 3\cdot\frac{kmg}{2a}\left(x_1^2 - x_2^2\right) &= \frac{3kmg}{2a}\cdot\frac{a^2\left(3-7+4\sqrt{3}\right)}{3} \\ &= 2kmga(\sqrt{3}-1). \end{aligned}$$

由於 $k = 4\sqrt{3}+6$，$k(\sqrt{3}-1) = 12-4\sqrt{3}+6\sqrt{3}-6 = 6+2\sqrt{3}$，所以釋放的彈性勢能為 $4mga(3+\sqrt{3})$。

三個質點增加的重力勢能：$3mg\cdot\dfrac{2a}{\sqrt{3}} = 2\sqrt{3}\,mga$。

兩者之差轉化為三個質點的動能，每個質點的速率均為 $V$：

$$\begin{aligned} \tfrac{3}{2}mV^2 &= 4mga(3+\sqrt{3}) - 2\sqrt{3}\,mga = mga\left(12+2\sqrt{3}\right), \\ V^2 &= \frac{2ag}{3}\left(12+2\sqrt{3}\right) = \frac{4ag}{3}\left(6+\sqrt{3}\right), \end{aligned}$$

即所求結果。

## 驗算

- 正負號：在 $\theta = \frac{1}{3}\pi$ 時，每條彈簧的鉛垂拉力為 $T\sin\theta = kmg\cdot\frac{\sqrt{3}}{2} \approx 11.2\,mg > mg$，所以三角形起初確實是向上朝平衡位置移動，與能量方程的假設一致。
- 答案為正數，而且具有速率平方的單位（$ag$）。
- 取 $a = 1.3\,\mathrm{m}$ 及 $g = 9.81\,\mathrm{m\,s^{-2}}$，由 $\theta = \frac{1}{3}\pi$ 的靜止狀態出發，以數值方法逐步推進運動方程 $\ddot{y} = g - \frac{T}{m}\sin\theta$，在 $\theta = \frac{1}{6}\pi$ 時得 $V^2 \approx 131.48\,\mathrm{m^2\,s^{-2}}$，與 $\frac{4ag}{3}(6+\sqrt{3})$ 吻合。

## 常見錯誤

- 把中心到頂點的距離誤當作 $\sqrt{3}a$ 或 $\frac{\sqrt{3}}{2}a$，而不是 $a$。
- 忘記共有三條彈簧和三個質點：因數 3 同時出現在彈性勢能、重力勢能和動能各項中。
- 弄錯運動方向：質點在兩個位置之間是上升的，所以重力勢能是增加，而不是減少。

## STEP 之外

在這裏，三角形只有一個自由度(degree of freedom)，因為對稱性排除了所有其他運動。選取單一坐標（這裏是 $\theta$，或深度 $y$），並以它表示能量，正是拉格朗日力學（Lagrangian mechanics）的起點。把勢能在平衡位置附近展開，還可以求得微小鉛垂振動的週期。Kleppner 和 Kolenkow 所著的 *An Introduction to Mechanics*（《力學導論》）便是本着這種精神，發展能量方法和小振動（small oscillations）的理論。

詳細解答

# STEP III 2013，第 12 題

概率與統計 • 指示隨機變量

**STEP III 2013，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/13-s3-q12

## 關鍵思路

每個 $X_k$ 恰好在一段新的 $A$ 從位置 $k$ 開始時等於 1，所以 $S$ 數的是這一行中 $A$ 的段數（連串，runs）。這些變量都是指示變量（indicator variables），所以每個期望值（expectation）都只是若干指定位置放着指定字母的概率，而這個概率只取決於涉及多少個位置，與它們在哪裏無關。至於方差（variance），我們把 $S^2$ 展開，並把乘積 $X_iX_j$ 分類：下標相等、下標相鄰（結果是不可能發生的）以及其餘情況。

## 解答

全文中，所有 $\binom{n}{a}$ 種排列的可能性均等。由此得到一個有用的結論：對任意 $m$ 個不同的位置，它們放着某個指定圖樣（其中含 $p$ 個 $A$ 和 $q$ 個 $B$，$p+q=m$）的概率，可以逐個位置填入來求得，例如

$$\Pr(\text{位置}r,s\text{ 分別為}B,A)=\frac{b}{n}\cdot\frac{a}{n-1},$$

不論不同的位置 $r$ 和 $s$ 是哪兩個。由於每個 $X_k$ 只取值 0 和 1，所以 $\mathrm{E}(X_k)=\Pr(X_k=1)$；同理 $\mathrm{E}(X_iX_j)=\Pr(X_i=1\text{ 且}X_j=1)$。

### 第 (i) 部分

當 $i=1$ 時，我們需要第一個字母是 $A$，所以

$$\mathrm{E}(X_1)=\frac{a}{n}.$$

當 $2\le i\le n$ 時，我們需要位置 $i-1$ 是 $B$，而位置 $i$ 是 $A$，所以

$$\mathrm{E}(X_i)=\frac{b}{n}\cdot\frac{a}{n-1}=\frac{ab}{n(n-1)}.$$

根據期望值的線性性質（linearity of expectation），對 $n-1$ 個下標 $i \geq 2$ 求和，

$$\mathrm{E}(S) = \frac{a}{n} + (n-1)\cdot\frac{ab}{n(n-1)} = \frac{a}{n} + \frac{ab}{n} = \frac{a(b+1)}{n},$$

即為所求。

**第 (ii) 部分**

**(a)** 設 $j \geq 3$。那麼 $X_1X_j = 1$ 恰好在位置 1 是 $A$、位置 $j-1$ 是 $B$ 且位置 $j$ 是 $A$ 時成立。由於 $j-1 \geq 2$，這是三個不同的位置，放着兩個 $A$ 和一個 $B$。逐個填入，

$$\mathrm{E}(X_1X_j) = \frac{a}{n}\cdot\frac{b}{n-1}\cdot\frac{a-1}{n-2} = \frac{a(a-1)b}{n(n-1)(n-2)}.$$

**(b)** 現在取 $2 \leq i$ 及 $j \geq i+2$。那麼 $X_iX_j = 1$ 恰好在位置 $i-1$, $i$, $j-1$, $j$ 分別是 $B, A, B, A$ 時成立。由於 $j-1 \geq i+1 > i$，這四個位置互不相同，放着兩個 $A$ 和兩個 $B$，所以

$$\mathrm{E}(X_iX_j) = \frac{b}{n}\cdot\frac{a}{n-1}\cdot\frac{b-1}{n-2}\cdot\frac{a-1}{n-3} = \frac{a(a-1)b(b-1)}{n(n-1)(n-2)(n-3)}.$$

（這裏因為 $a, b \geq 2$，所以 $n \geq 4$，分母不會為零。）這個值對雙重和中的每一對都相同，所以我們只需數出有多少對。對給定的 $i$，$j$ 由 $i+2$ 取到 $n$，共 $n-i-1$ 個值；當 $i$ 由 2 取到 $n-2$ 時，這個數依次取值 $n-3$, $n-4$, …, 1。所以對數為

$$1 + 2 + \cdots + (n-3) = \tfrac{1}{2}(n-3)(n-2),$$

而

$$\sum_{i=2}^{n-2}\left(\sum_{j=i+2}^{n}\mathrm{E}(X_iX_j)\right) = \frac{(n-2)(n-3)}{2}\cdot\frac{a(a-1)b(b-1)}{n(n-1)(n-2)(n-3)} = \frac{a(a-1)b(b-1)}{2n(n-1)}.$$

**(c)** 我們利用 $\mathrm{Var}(S) = \mathrm{E}(S^2) - \big(\mathrm{E}(S)\big)^2$，並展開

$$S^2 = \sum_{i=1}^{n} X_i^2 + 2\sum_{i<j} X_iX_j.$$

由於 $X_i$ 取值 0 或 1，所以 $X_i^2 = X_i$，因此第一個和的期望值是 $\mathrm{E}(S)$。

**想一想** 那麼 (b) 部分略去的 $j = i+1$ 的那些對又如何？

它們沒有任何貢獻：$X_iX_{i+1} = 0$ 恆成立。當 $i = 1$ 時，$X_1 = 1$ 要求位置 1 是 $A$，但 $X_2 = 1$ 要求該位置是 $B$。當 $i \geq 2$ 時，$X_i = 1$ 要求位置 $i$ 是 $A$，但 $X_{i+1} = 1$ 要求該位置是 $B$。（兩段 $A$ 不可能在相鄰的位置開始。）

所以真正有關的 $i<j$ 對，就是那些 $j\geq i+2$ 的。其中 $i=1$ 的有 $j=3,\ldots,n$，共 $n-2$ 對，每一對都由 (a) 部分給出；$i\geq 2$ 的恰好就是 (b) 部分的雙重和。因此

$$\begin{aligned}\sum_{i<j}\mathrm{E}(X_iX_j) &= (n-2)\cdot\frac{a(a-1)b}{n(n-1)(n-2)}+\frac{a(a-1)b(b-1)}{2n(n-1)}\\ &= \frac{a(a-1)b}{n(n-1)}\left(1+\frac{b-1}{2}\right)=\frac{a(a-1)b(b+1)}{2n(n-1)},\end{aligned}$$

於是

$$\mathrm{E}(S^2)=\frac{a(b+1)}{n}+\frac{a(a-1)b(b+1)}{n(n-1)}.$$

減去 $\big(\mathrm{E}(S)\big)^2=a^2(b+1)^2/n^2$，並抽出公因式 $\dfrac{a(b+1)}{n^2(n-1)}$，

$$\mathrm{Var}(S)=\frac{a(b+1)}{n^2(n-1)}\Big[n(n-1)+n(a-1)b-a(b+1)(n-1)\Big].$$

展開括號，$nab$ 與 $-nab$ 兩項互相抵消，剩下

$$\begin{aligned}n^2-n-nb-na+ab+a &= n(n-a-b)-n+ab+a\\ &= -n+ab+a \qquad \text{因為} n=a+b\\ &= ab-b=b(a-1).\end{aligned}$$

所以

$$\mathrm{Var}(S)=\frac{a(b+1)}{n^2(n-1)}\cdot b(a-1)=\frac{a(a-1)b(b+1)}{n^2(n-1)},$$

即為所求。

## 驗算

取 $a=3$、$b=2$，於是 $n=5$，共有 10 種排列。數出 $A$ 的段數：排列 $AAABB$、$BAAAB$、$BBAAA$ 給出 $S=1$；$ABABA$ 給出 $S=3$；其餘六種給出 $S=2$。所以 $\mathrm{E}(S)=\frac{18}{10}=\frac{9}{5}$，$\mathrm{E}(S^2)=\frac{36}{10}$，從而 $\mathrm{Var}(S)=\frac{18}{5}-\frac{81}{25}=\frac{9}{25}$。公式給出 $\frac{3\cdot3}{5}=\frac{9}{5}$ 及 $\frac{3\cdot2\cdot2\cdot3}{25\cdot4}=\frac{9}{25}$。以電腦窮舉驗證，對所有 $2\leq a,b\leq 6$，題中每條公式都完全正確。

這些公式在題目所述範圍以外也表現合理。當 $a=2$、$b=1$ 時，排列 $AAB$、$ABA$、$BAA$ 給出 $S=1,2,1$，所以 $\mathrm{E}(S)=\frac{4}{3}$，$\mathrm{Var}(S)=2-\frac{16}{9}=\frac{2}{9}$，與 $\frac{2\cdot2}{3}$ 及 $\frac{2\cdot1\cdot1\cdot2}{9\cdot2}$ 吻合。而若 $a=1$，方差公式給出 0，這理應如此，因為此時永遠恰好只有一段。

## 常見錯誤

- 把 $X_1$ 當作與其他變量一樣處理：並沒有位置 0，所以 $X_1$ 只涉及一個字母，$\mathrm{E}(X_1)=a/n$，而不是 $ab/(n(n-1))$。

- 忘記 $X_i$ 與 $X_{i+1}$ 不可能同時為 1，或用一條非零的公式去計算 $\mathrm{E}(X_iX_{i+1})$。
- 數錯 (b) 部分的對數：應有 $\frac{1}{2}(n-2)(n-3)$ 對，而不是 $\frac{1}{2}n(n-1)$ 對。

## STEP 之外

把一個計數寫成指示變量之和，再用線性性質求平均值，並由兩兩乘積求方差，是大學概率論中最有用的技巧之一；分析連串、配對、隨機排列（random permutations）的不動點（fixed points）以及隨機圖（random graphs），都是用這個方法。這裏的 $S$ 本質上就是 Wald–Wolfowitz 連串檢定（runs test）的統計量，用來檢測一個序列是否並非隨機。Blitzstein 與 Hwang 的 *Introduction to Probability*（《概率導論》）載有許多運用指示變量方法的例子。

詳細解答

# STEP III 2013，第 13 題

概率與統計 • 矩與分佈函數

**STEP III 2013，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/13-s3-q13

## 關鍵思路

由於 $\mathrm{F}' = \mathrm{f}$，表達式 $2\mathrm{F}(x)\mathrm{f}(x)$ 正是 $\big(\mathrm{F}(x)\big)^2$ 的導數（derivative）。因此第 (i)(b) 部分只需一次分部積分（integration by parts）；而只要適當選取 g，關於 $Y$ 的每個問題都能化為關於 $X$ 的問題。至於各個界（bound），則來自把 f 或 F 換成第 (i)(a) 部分中較大的量 $M$ 或 $Mx$，而每次替換都朝着令積分變大的方向進行。

## 解答

全文中，$\mathrm{F}(x) = \int_0^x \mathrm{f}(t)\,\mathrm{d}t$，故 $\mathrm{F}'(x) = \mathrm{f}(x)$；又由於 $X$ 取值於 $[0,1]$，有 $\mathrm{F}(0) = 0$ 及 $\mathrm{F}(1) = 1$。

### 第 (i) 部分

**(a)** 對 $0 \le x \le 1$，由於對每個 $t$ 都有 $0 \le \mathrm{f}(t) \le M$，

$$0 \le \int_0^x \mathrm{f}(t)\,\mathrm{d}t \le \int_0^x M\,\mathrm{d}t = Mx,$$

即 $0 \le \mathrm{F}(x) \le Mx$。

**(b)** 由鏈式法則（chain rule），$\dfrac{\mathrm{d}}{\mathrm{d}x}\big(\mathrm{F}(x)\big)^2 = 2\mathrm{F}(x)\mathrm{f}(x)$。取 $u = \mathrm{g}(x)$ 及 $\mathrm{d}v = 2\mathrm{F}(x)\mathrm{f}(x)\,\mathrm{d}x$ 進行分部積分，於是 $v = \big(\mathrm{F}(x)\big)^2$，

$$\begin{aligned}\int_0^1 2\mathrm{g}(x)\mathrm{F}(x)\mathrm{f}(x)\,\mathrm{d}x &= \Big[\mathrm{g}(x)\big(\mathrm{F}(x)\big)^2\Big]_0^1 - \int_0^1 \mathrm{g}'(x)\big(\mathrm{F}(x)\big)^2\,\mathrm{d}x \\ &= \mathrm{g}(1) - \int_0^1 \mathrm{g}'(x)\big(\mathrm{F}(x)\big)^2\,\mathrm{d}x,\end{aligned}$$

因為 $\mathrm{F}(1) = 1$ 且 $\mathrm{F}(0) = 0$。（這裏需要 g 可微（differentiable），題目已隱含此假設。）

**第 (ii) 部分**

**(a)** $Y$ 的概率密度函數（probability density function）的積分必須等於 1。在 (i)(b) 中取 $\mathrm{g}(x) = 1$，使 $\mathrm{g}' = 0$，得

$$\int_0^1 2\mathrm{F}(y)\mathrm{f}(y)\,\mathrm{d}y = 1,$$

故 $k = 2$。（亦可直接計算：$\int_0^1 k\mathrm{F}\mathrm{f}\,\mathrm{d}y = \frac{k}{2}\big[\mathrm{F}^2\big]_0^1 = \frac{k}{2}$。）另請注意 $k\mathrm{F}\mathrm{f} \geq 0$，所以這確實是一個概率密度函數。

**(b)** 我們有

$$\mathrm{E}(Y^n) = \int_0^1 2y^n\,\mathrm{F}(y)\mathrm{f}(y)\,\mathrm{d}y.$$

上界。由 (i)(a)，$\mathrm{F}(y) \leq My$，而其餘因子 $2y^n\mathrm{f}(y)$ 均非負，所以

$$\mathrm{E}(Y^n) \leq \int_0^1 2y^n \cdot My \cdot \mathrm{f}(y)\,\mathrm{d}y = 2M\int_0^1 y^{n+1}\mathrm{f}(y)\,\mathrm{d}y = 2M\mu_{n+1}.$$

下界。現在在 (i)(b) 中取 $\mathrm{g}(y) = y^n$，於是 $\mathrm{g}(1) = 1$ 及 $\mathrm{g}'(y) = ny^{n-1}$：

$$\mathrm{E}(Y^n) = 1 - \int_0^1 ny^{n-1}\big(\mathrm{F}(y)\big)^2\,\mathrm{d}y.$$

（當 $n = 0$ 時，該積分就是 0。）要求 $\mathrm{E}(Y^n)$ 的下界，我們需要求這個積分的上界。

**?** **想一想** $\big(\mathrm{F}(y)\big)^2$ 的兩個因子中，應該把哪一個換成 $My$？

只換其中一個。由於 $0 \leq \mathrm{F}(y) \leq My$，有 $\big(\mathrm{F}(y)\big)^2 \leq My\,\mathrm{F}(y)$，加上 $ny^{n-1} \geq 0$，所以

$$\int_0^1 ny^{n-1}\big(\mathrm{F}(y)\big)^2\,\mathrm{d}y \leq nM\int_0^1 y^n\mathrm{F}(y)\,\mathrm{d}y.$$

若把兩個因子都換掉，就只剩下 $\int y^{n+1}\,\mathrm{d}y$，對 $X$ 毫無資訊；保留一個 F，便可再做一次分部積分，得出 $X$ 的一個矩（moment）。取 $u = \mathrm{F}(y)$ 及 $\mathrm{d}v = y^n\,\mathrm{d}y$ 進行分部積分，

$$\int_0^1 y^n\mathrm{F}(y)\,\mathrm{d}y = \Big[\frac{y^{n+1}}{n+1}\mathrm{F}(y)\Big]_0^1 - \int_0^1 \frac{y^{n+1}}{n+1}\mathrm{f}(y)\,\mathrm{d}y = \frac{1-\mu_{n+1}}{n+1}.$$

綜合以上，

$$\mathrm{E}(Y^n) \geq 1 - \frac{nM}{n+1}\big(1-\mu_{n+1}\big) = 1 + \frac{nM}{n+1}\mu_{n+1} - \frac{nM}{n+1},$$

連同上界，便是所求的一對不等式（inequality）。

**(c)** 把 (b) 的兩端結合起來，對 $n \geq 0$，

$$1 + \frac{nM}{n+1}\mu_{n+1} - \frac{nM}{n+1} \leq 2M\mu_{n+1},$$

再把含 $\mu_{n+1}$ 的項移到右邊，

$$1 - \frac{nM}{n+1} \leq \Big(2M - \frac{nM}{n+1}\Big)\mu_{n+1} = \frac{(n+2)M}{n+1}\,\mu_{n+1}.$$

係數 $\frac{(n+2)M}{n+1}$ 為正（$M > 0$，因為 f 的積分等於 1），所以可以除以它而不改變不等號方向：

$$\mu_{n+1} \geq \frac{n+1}{(n+2)M} - \frac{n}{n+2}.$$

此式對每個 $n \geq 0$ 成立。令 $m = n+1$，則 $m \geq 1$，上式變為

$$\mu_m \geq \frac{m}{(m+1)M} - \frac{m-1}{m+1} \qquad (m \geq 1),$$

這正是所求結果（只需把 $m$ 改名為 $n$）。

## 驗算

- 均勻分佈（uniform distribution）的 $X$：$\mathrm{f} = 1$，$M = 1$，$\mu_n = \frac{1}{n+1}$。(c) 中的界為 $\frac{n}{n+1} - \frac{n-1}{n+1} = \frac{1}{n+1}$，所以等號成立。此時 $Y$ 的概率密度函數為 $2y$，且 $\mathrm{E}(Y^n) = \frac{2}{n+2}$，而 (b) 中的兩個界亦都等於 $\frac{2}{n+2}$。這是合理的：我們用過的每一個不等式 $\mathrm{F}(y) \leq My$，對均勻分佈而言都是等式。
- $\mathrm{f}(x) = 3x^2$，$M = 3$，$n = 2$：$\mu_2 = \frac{3}{5}$，而界為 $\frac{2}{9} - \frac{1}{3} < 0$，輕鬆成立。對這個概率密度函數，$\mathrm{E}(Y^2) = \frac{3}{4}$ 及 $\mu_3 = \frac{1}{2}$，所以 (b) 中的界為 $1 + 2\cdot\frac{1}{2} - 2 = 0$ 及 $2\cdot 3\cdot\frac{1}{2} = 3$，而 $\frac{3}{4}$ 確實介乎兩者之間。
- 我們亦以數值積分，對多個概率密度函數及多種 g 的選擇，驗證了 (i)(b) 中的恆等式及所有不等式。

## 常見錯誤

- 以 $M^2y^2$ 作為 $\big(\mathrm{F}\big)^2$ 的上界：這雖然正確，卻毫無用處，因為它丟失了關於 $X$ 的所有資訊。

- 忘記 $\mathrm{E}(Y^n)$ 的表達式中積分前面帶有負號，所以要求 $\mathrm{E}(Y^n)$ 的下界，就需要積分的上界。
- 得出 $\mu_{n+1}$ 的界後便停下，沒有移動下標，以得出題目所述對 $n \geq 1$ 的形式。

## STEP 之外

概率密度函數 $2\mathrm{F}(y)\mathrm{f}(y)$ 並非偶然：它是 $X$ 的兩個獨立副本的 $\max(X_1, X_2)$ 的概率密度函數，因為 $\mathrm{P}(\max \leq y) = \big(\mathrm{F}(y)\big)^2$。更一般地，$n$ 個獨立觀測值中第 $r$ 小者的概率密度函數，由 F 和 $1 - \mathrm{F}$ 的冪組成；這些順序統計量（order statistics）是無母數統計（non-parametric statistics）的核心。Blitzstein 與 Hwang 的 *Introduction to Probability*（《概率導論》）對順序統計量有清晰的講解。

# Beyond STEP

STEP 是大多數中學生一生中應考過最艱深的數學考試，也是最出色的考試之一：題目簡短、陌生而深刻，當中最精彩的，其實是喬裝了的大學數學小品。

本書精選最有趣的 STEP 試題，由 1987 年首次開考之前的樣本試卷（specimen papers），到 2025 年的試卷，並以每一道題目作為一扇門。每道題目都會仔細解答，沿途提出優秀學生會問的問題，然後帶你穿過這扇門：這個想法從何而來，到了大學會演變成甚麼，以及接下來該讀些甚麼。

本書為正在準備 STEP、有天分又有志投身數學的學生而寫。他們想要的不止是模擬試卷：他們想開始像數學系本科生那樣思考，並一探前方有甚麼正等着自己。

- 共 13 章，涵蓋純數學（pure mathematics）、力學（mechanics）、概率（probability）與統計（statistics）
- 136 道 STEP 歷屆試題，每題均附引導式解答
- 每道題目背後的大學數學
- 附答案的延伸練習，以及延伸閱讀

**2013 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。