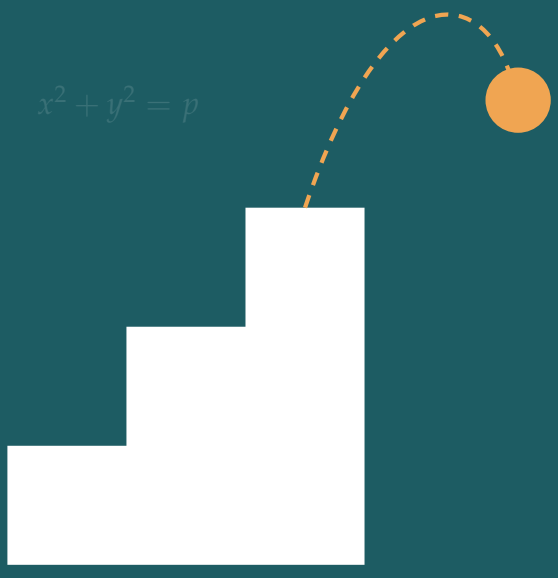
STEP 備試 · 大學數學

Beyond STEP

STEP I, II 及 III

2016 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

2016 年詳細解答

Keith McNulty

Beyond STEP：以歷屆 STEP 試題為門，走進大學數學

2016 年詳細解答

2016 年 STEP 試卷（STEP I, II 及 III）全部 37 題的詳細解答。自樣本試卷以來每一道 STEP 題目的詳細解答、*Beyond STEP* 全書以及各章的單行本，均可在 `beyondstep.keithmcnulty.org` 免費閱讀及下載。

閱讀解答之前，請先認真嘗試解題。每個答案均已用數值方法核對，但解答數量如此之多，難免會有疏漏：如果發現有可疑之處，請相信自己的演算，並仔細檢查。

由機器從英文翻譯而成；尚未經審閱。

2026 年 11 月首次於網上發表。

本 PDF 內凡屬作者本人創作的內容，包括正文、解答、插圖、練習及答案，均按照 Creative Commons CC0 1.0 Universal 授權條款貢獻予公有領域：你可以為任何目的複製、改編及分享，毋須徵求許可。詳見 `creativecommons.org/publicdomain/zero/1.0`。

本書討論的 STEP 試題，版權屬 Cambridge University Press & Assessment 所有。本書並無轉載這些試題：每一題均附連結，通往官方來源的題目，即 STEP Support Programme 的試題資料庫或 OCR 的歷屆試卷。

如有錯漏，歡迎到 github.com/keithmcnulty/step_book_mcnulty 提交 issue 指正。

目錄

STEP I

第 1 題 多項式恆等式與因式分解 純數學 . . . 3

第 4 題 曲線的曲率 純數學 . . . 6

第 5 題 相切圓 純數學 . . . 9

第 6 題 向量幾何 純數學 . . . 13

第 7 題 模 4 的因式分解 純數學 . . . 17

第 8 題 生成函數 純數學 . . . 21

第 9 題 極限摩擦力平衡 力學 . . . 25

第 10 題 碰撞與恢復係數 力學 . . . 29

第 11 題 拋體的最大射程 力學 . . . 33

第 12 題 項分佈的對稱性 概率與統計 . . . 37

第 13 題 整數分佈的順序統計量 概率與統計 . . . 40

STEP II

第 1 題 參數曲線與切線 純數學 . . . 43

第 2 題 因式定理 純數學 . . . 47

第 3 題 多項式的根 純數學 . . . 51

第 4 題 有理函數的值域 純數學 . . . 54

第 5 題 二項式係數恆等式 純數學 . . . 58

第 6 題 切比雪夫多項式方程 純數學 . . . 62

第 7 題 利用對稱性求定積分 純數學 . . . 66

第 8 題 以積分近似求和 純數學 . . . 70

第 9 題 動量與能量損失 力學 . . . 74

第 10 題 摩擦力與平衡 力學 . . . 77

第 11 題 拋體的碰撞 力學 . . . 80

第 12 題 容斥原理與錯排 概率與統計 84
第 13 題 正態近似 概率與統計 88

STEP III

第 1 題 歸約公式 純數學 92
第 2 題 拋物線的法線 純數學 96
第 3 題 積分與多項式 純數學 100
第 4 題 裂項相消級數 純數學 103
第 5 題 質數的乘積 純數學 106
第 6 題 雙曲函數 純數學 109
第 7 題 單位根 純數學 113
第 8 題 函數方程 純數學 117
第 9 題 彈簧與簡諧運動 力學 121
第 10 題 平面上的圓周運動 力學 125
第 11 題 功率與阻力 力學 129
第 12 題 切比雪夫不等式 概率與統計 133
第 13 題 種的峰度 概率與統計 136

詳細解答

STEP I 2016，第 1 題

純數學 • 多項式恆等式與因式分解

> **STEP I 2016，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/16-s1-q1

關鍵思路

商（quotient）$q_n(x)$ 其實是一個多項式（polynomial）：由於 $2n+1$ 是奇數，$x+1$ 整除 $x^{2n+1}+1$，而商是一個正負交替的等比數列之和。因此第 (i) 部分只需展開 p_n 並比較係數（coefficients）。第 (ii) 部分的關鍵在於 p_n 的形式：當 $2n+1$ 乘以 x 是完全平方數（perfect square）時，p_n 是兩個平方數之差（difference of two squares），因此可以因式分解（factorise）。

解答

第 (i) 部分

當 $x \neq -1$ 時，把首項為 1、公比為 $-x$ 的等比級數（geometric series）求和，得

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

這裏用到了 $(-x)^{2n+1} = -x^{2n+1}$，因為 $2n+1$ 是奇數。所以

$$q_n(x) = x^{2n} - x^{2n-1} + x^{2n-2} - \cdots - x + 1,$$

它的 $2n+1$ 個係數全部是 ± 1，而且正負交替。接下來我們逐一展開 p_n。

$n = 1$.

$$p_1(x) = (x+1)^2 - 3x = x^2 + 2x + 1 - 3x = x^2 - x + 1 = q_1(x).$$

$n = 2$.

$$\begin{aligned} p_2(x) &= (x+1)^4 - 5x(x^2+x+1) \\ &= (x^4+4x^3+6x^2+4x+1) - (5x^3+5x^2+5x) \\ &= x^4 - x^3 + x^2 - x + 1 = q_2(x). \end{aligned}$$

$n = 3$. 首先 $(x^2+x+1)^2 = x^4+2x^3+3x^2+2x+1$，所以

$$\begin{aligned} p_3(x) &= (x+1)^6 - 7x(x^2+x+1)^2 \\ &= (x^6+6x^5+15x^4+20x^3+15x^2+6x+1) - (7x^5+14x^4+21x^3+14x^2+7x) \\ &= x^6 - x^5 + x^4 - x^3 + x^2 - x + 1 = q_3(x). \end{aligned}$$

因此當 $n = 1, 2, 3$ 時，對 $x \neq -1$ 有 $p_n(x) \equiv q_n(x)$，正如所求。

$n = 4$. 只需一個 x 值便足以推翻一條恆等式（identity）。當 $x = 1$ 時，

$$p_4(1) = 2^8 - 9 \cdot 1 \cdot 3^3 = 256 - 243 = 13, \qquad q_4(1) = \frac{1+1}{1+1} = 1,$$

所以 $p_4 \not\equiv q_4$。（完全展開可得 $p_4(x) = x^8 - x^7 + x^6 + 2x^5 + 7x^4 + 2x^3 + x^2 - x + 1$，從 x^5 項起便出錯。）

第 (ii) 部分

想一想 應該用哪個 n 和哪個 x？p_n 又怎樣才可能因式分解？

把 $\dfrac{x^{2n+1}+1}{x+1}$ 與題目給出的數對應起來，便確定了 x 和 n。這時 p_n 的形式就派上用場：$(x+1)^{2n} = ((x+1)^n)^2$ 必定是平方數，而當 n 是奇數時，$(x^2+x+1)^{n-1}$ 也是平方數。因此，若 $(2n+1)x$ 同樣是完全平方數，$p_n(x)$ 便是兩個平方數之差。

(a) 取 $n = 1$ 及 $x = 300$，於是 $2n+1 = 3$，且 $3x = 900 = 30^2$。由第 (i) 部分，

$$\begin{aligned} \frac{300^3+1}{301} = q_1(300) = p_1(300) &= 301^2 - 3 \cdot 300 \\ &= 301^2 - 30^2 \\ &= (301-30)(301+30) = 271 \times 331. \end{aligned}$$

兩個因數都不是 1。（直接計算，$300^2 - 300 + 1 = 89\,701 = 271 \times 331$。）

(b) 由於 $7^{49} = (7^7)^7$，取 $n = 3$（於是 $2n+1 = 7$）及 $x = 7^7$。那麼 $(2n+1)x = 7 \cdot 7^7 =$

$7^8 = (7^4)^2$ 是完全平方數，而且 $n-1=2$ 是偶數。由第 (i) 部分，

$$\frac{7^{49}+1}{7^7+1} = q_3(7^7) = p_3(7^7) = (7^7+1)^6 - 7^8\,(7^{14}+7^7+1)^2$$
$$= [(7^7+1)^3]^2 - [7^4(7^{14}+7^7+1)]^2,$$

所以

$$\frac{7^{49}+1}{7^7+1} = \Big[(7^7+1)^3 - 7^4(7^{14}+7^7+1)\Big]\Big[(7^7+1)^3 + 7^4(7^{14}+7^7+1)\Big].$$

第二個因數顯然大於 1。至於第一個因數，$(7^7+1)^3 > 7^{21}$，而

$$7^4(7^{14}+7^7+1) = 7^{18}+7^{11}+7^4 < 7^{19},$$

所以第一個因數大於 $7^{21}-7^{19} = 48\cdot 7^{19} > 1$。兩個因數都不是 1，於是得到所求的因式分解。（把括號乘開，兩個因數亦可寫成 $7^{21}+3\cdot 7^{14}+3\cdot 7^7+1 \mp (7^{18}+7^{11}+7^4)$。）

驗算

- 在第 (i) 部分代入 $x=1$：$p_1(1) = 4-3 = 1$，$p_2(1) = 16-15 = 1$，$p_3(1) = 64-63 = 1$，全部等於 $q_n(1)=1$，而 $p_4(1) = 13$。
- 在 (a) 中，$271\times 331 = 89\,701 = 300^2 - 300 + 1$。
- 在 (b) 中，用電腦檢驗準確的整數值，證實兩個因數之積等於 $(7^{49}+1)/(7^7+1)$；兩個因數分別約為 5.57×10^{17} 和 5.60×10^{17}。

常見錯誤

- 忘記說明為何兩個因數都不是 1；在 (b) 中這需要簡短地估計大小。
- 在 (b) 中用 $n=24$ 或 $x=7$：指數 49 必須寫成 7×7，從而得到 $x=7^7$ 及 $2n+1=7$。

STEP 之外

這裏的技巧是奧里費伊因式分解（Aurifeuillean factorisation）的一個特例：例如 $x^4+4y^4 = (x^2+2y^2-2xy)(x^2+2y^2+2xy)$ 這類恆等式，能把形如 $a^k\pm 1$ 的數分解得比明顯的代數因式更細。它們源自分圓多項式（cyclotomic polynomials），這是大學一年級代數和數論（number theory）課程會出現的內容，而且曾被用於分解形如 $2^k\pm 1$ 的大數。Davenport 的 *The Higher Arithmetic*（《高等算術》）是進一步了解因式分解和質數（primes）的好讀物。

詳細解答

STEP I 2016，第 4 題

純數學 • 曲線的曲率

STEP I 2016，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/16-s1-q4

關鍵思路

κ 的表達式看似一條艱深的二階微分方程（second-order differential equation），但第 (i) 部分正是要說明它其實是某個函數的導數，可直接積分（exact derivative）。設 $p = \mathrm{f}'(x)$，曲率（curvature）就是 $p/\sqrt{1+p^2}$ 對 x 的導數，因此常數曲率一眼便可積分一次。第二次積分則屬例行工夫，而答案原來是一個熟悉的形狀。

解答

第 (i) 部分

對 $z(1+z^2)^{-1/2}$ 使用乘積法則（product rule），

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}z}\left(\frac{z}{(1+z^2)^{1/2}}\right) &= (1+z^2)^{-1/2} - z\cdot z(1+z^2)^{-3/2} \\ &= \frac{(1+z^2)-z^2}{(1+z^2)^{3/2}} = \frac{1}{(1+z^2)^{3/2}}.\end{aligned}$$

第 (ii) 部分

設 κ 為非零常數，並記 $p = \mathrm{f}'(x)$，於是 $\mathrm{f}''(x) = \dfrac{\mathrm{d}p}{\mathrm{d}x}$。由第 (i) 部分及鏈式法則（chain rule），

$$\frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{p}{\sqrt{1+p^2}}\right) = \frac{1}{(1+p^2)^{3/2}}\frac{\mathrm{d}p}{\mathrm{d}x} = \frac{\mathrm{f}''(x)}{\left(1+(\mathrm{f}'(x))^2\right)^{3/2}} = \kappa.$$

對 x 積分，存在一個常數，我們把它寫成 $-\kappa a$（因 $\kappa \neq 0$，這樣做是可行的），使得

$$\frac{p}{\sqrt{1+p^2}} = \kappa(x-a). \qquad (*)$$

想一想 如何從 $(*)$ 解出 p，而又不會弄錯正負號？

只要記住平方會丟失甚麼資訊，平方便是安全的。由於 $\sqrt{1+p^2} > 0$，$(*)$ 表示 p 與 $\kappa(x-a)$ 同號；又因 $|p| < \sqrt{1+p^2}$，它亦表示 $|\kappa(x-a)| < 1$。兩邊平方，得 $p^2 = \kappa^2(x-a)^2(1+p^2)$，所以

$$p^2 = \frac{\kappa^2(x-a)^2}{1-\kappa^2(x-a)^2}, \qquad \text{並按上述選取正負號，} \qquad p = \frac{\kappa(x-a)}{\sqrt{1-\kappa^2(x-a)^2}}.$$

反過來，這個 p 確實滿足 $(*)$，因為 $1+p^2 = 1/\big(1-\kappa^2(x-a)^2\big)$。

現在再積分一次。由於 $\dfrac{\mathrm{d}}{\mathrm{d}x}\sqrt{1-\kappa^2(x-a)^2} = \dfrac{-\kappa^2(x-a)}{\sqrt{1-\kappa^2(x-a)^2}}$，對某常數 b，我們得到

$$y = b - \frac{1}{\kappa}\sqrt{1-\kappa^2(x-a)^2}, \qquad |x-a| < \frac{1}{|\kappa|}.$$

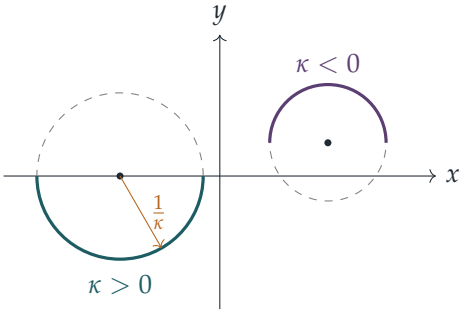
整理後得 $\kappa(y-b) = -\sqrt{1-\kappa^2(x-a)^2}$，再平方便得

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

因此，每條具有非零常數有號曲率（signed curvature）κ 的曲線，都是以 (a,b) 為圓心、半徑為 $1/|\kappa|$ 的圓的一部分；而且每段這樣的圓弧都會出現：把步驟倒過來，上述每個函數的有號曲率都恰好是 κ。更準確地說，$\kappa(y-b) \leq 0$，所以：

- 若 $\kappa > 0$，曲線是下半圓 $y = b - \sqrt{\kappa^{-2} - (x-a)^2}$，位於圓心下方，向上彎曲（$\mathrm{f}'' > 0$）；
- 若 $\kappa < 0$，曲線是上半圓 $y = b + \sqrt{\kappa^{-2} - (x-a)^2}$，位於圓心上方，向下彎曲（$\mathrm{f}'' < 0$）。

端點 $x = a \pm 1/|\kappa|$ 須排除，因為該處斜率為無限大；曲線 $y = \mathrm{f}(x)$ 只能是一個函數的圖像，所以它最多只能是一個開半圓（或其中一段）。



幾何意義

對這些曲線而言，$|\kappa|$ 是圓半徑的倒數，所以 $|\kappa|$ 越大，圓就越小、彎得越緊。κ 的正負號記錄曲線向哪一方彎曲：當圓心位於曲線上方（曲線上凹（concave up））時 $\kappa > 0$；當圓心位於曲線下方時 $\kappa < 0$。

驗算

- 取 $\kappa = \frac{1}{2}$，$a = b = 0$：曲線是 $y = -\sqrt{4 - x^2}$，即半徑為 2 的圓的下半部分。在 $x = 1$ 處，$y = -\sqrt{3}$，$y' = x/\sqrt{4 - x^2} = 1/\sqrt{3}$，而 $y'' = 4/(4 - x^2)^{3/2} = 4/3^{3/2}$，所以 $1 + y'^2 = \frac{4}{3}$，且 $\kappa = \dfrac{4/3^{3/2}}{(4/3)^{3/2}} = \frac{1}{2}$，與預期相符。
- 直線有 $\mathrm{f}'' = 0$，所以 $\kappa = 0$：這是半徑 $1/|\kappa| \to \infty$ 時圓的極限，亦是題目排除 $\kappa = 0$ 的原因。
- 用電腦從隨機初始值逐步數值求解 $\mathrm{f}'' = \kappa(1 + \mathrm{f}'^2)^{3/2}$，所得各點準確至小數點後六位都留在預測的圓上。

常見錯誤

- 第一次積分時忘記積分常數（constant of integration），因而失去把圓心左右移動的自由度。
- 為 p 選錯平方根，結果得到位於圓心錯誤一側的半圓，以及錯誤的 κ 正負號。
- 只說「κ 是半徑」：其實它是半徑的倒數，並帶有表示彎曲方向的正負號。

STEP 之外

曲率是曲線微分幾何（differential geometry）的起點。對以弧長（arc length）s 參數化的曲線，$\kappa = \mathrm{d}\psi/\mathrm{d}s$，其中 ψ 是切線（tangent）與 x 軸所成的角；這裏 $p/\sqrt{1 + p^2} = \sin\psi$，所以 $(*)$ 實際上是說 ψ 沿曲線以恆定速率變化。在三維空間中，曲率之外還有撓率（torsion），而 Frenet–Serret 公式（Frenet–Serret formulae）顯示，這兩個函數在相差一個剛體運動（rigid motion）的意義下決定了一條曲線。Pressley 的 *Elementary Differential Geometry*（《初等微分幾何》）是一本大學一年級程度、平易近人的入門書。

詳細解答

STEP I 2016，第 5 題

純數學 • 相切圓

STEP I 2016，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/16-s1-q5

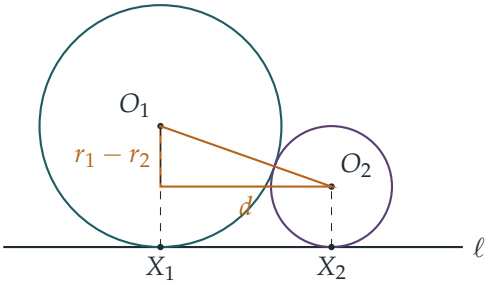
關鍵思路

一切都建基於一個事實：若兩個圓互相相切（touch），而且都與同一條直線相切，那麼它們與該直線的兩個切點之間的距離是 $2\sqrt{r_1 r_2}$。這只需要一個直角三角形（right-angled triangle）便可得出。之後，代換（substitution）$x = 1/\sqrt{a}$、$y = 1/\sqrt{b}$、$z = 1/\sqrt{c}$ 把 $(**)$ 化成一個對稱（symmetric）的四次式（quartic），而它可以因式分解（factorise）成四個一次因式（linear factor），其中一個因式恰好就是 $(*)$。

解答

第 (i) 部分

在直線上互相相切的兩個圓。取半徑分別為 r_1 和 r_2 的兩個圓，它們互相外切（touch externally），都位於直線 ℓ 的同一側，並分別在點 X_1 和 X_2 與 ℓ 相切。連到 X_1 和 X_2 的半徑都垂直於 ℓ，所以圓心 O_1、O_2 位於 ℓ 上方高度 r_1 和 r_2 之處。由於兩圓外切，$O_1O_2 = r_1 + r_2$。



從 O_2 向半徑 O_1X_1 作垂線（圖中 $r_1 \geq r_2$）。這樣便得到一個直角三角形，其斜邊（hypotenuse）為 $r_1 + r_2$，鉛垂邊為 $|r_1 - r_2|$，水平邊為 $d = X_1X_2$。由畢氏定理（Pythagoras' theorem），

$$d^2 = (r_1 + r_2)^2 - (r_1 - r_2)^2 = 4r_1r_2, \qquad \text{所以} \qquad X_1X_2 = 2\sqrt{r_1r_2}.$$

應用三次。在圖中，每兩個圓都互相相切，而三個圓都與直線相切，切點為 P、Q、R，其中 Q 位於 P 與 R 之間。所以

$$PQ = 2\sqrt{ab}, \qquad QR = 2\sqrt{bc}, \qquad PR = 2\sqrt{ac},$$

而由 $PQ + QR = PR$ 得 $\sqrt{ab} + \sqrt{bc} = \sqrt{ac}$。兩邊除以 $\sqrt{abc} > 0$，

$$\frac{1}{\sqrt{c}} + \frac{1}{\sqrt{a}} = \frac{1}{\sqrt{b}},$$

這正是 $(*)$。

推導 $(**)$。設 $x = \dfrac{1}{\sqrt{a}}$、$y = \dfrac{1}{\sqrt{b}}$、$z = \dfrac{1}{\sqrt{c}}$，於是 $\dfrac{1}{a} = x^2$，如此類推，而 $(*)$ 即是 $y = x + z$。那麼 $(**)$ 變成 $2(x^4 + y^4 + z^4) = (x^2 + y^2 + z^2)^2$；展開右方並移項，它等價於

$$F := x^4 + y^4 + z^4 - 2x^2y^2 - 2y^2z^2 - 2z^2x^2 = 0.$$

? **想一想** 這個四次式可以怎樣因式分解？

把它組合成平方差（difference of two squares）。由於 $(x^2 + y^2 - z^2)^2 = x^4 + y^4 + z^4 + 2x^2y^2 - 2y^2z^2 - 2z^2x^2$，

$$\begin{aligned} F &= (x^2 + y^2 - z^2)^2 - (2xy)^2 \\ &= \big(x^2 - 2xy + y^2 - z^2\big)\big(x^2 + 2xy + y^2 - z^2\big) \\ &= \big((x - y)^2 - z^2\big)\big((x + y)^2 - z^2\big) \\ &= (x - y - z)(x - y + z)(x + y - z)(x + y + z). \end{aligned}$$

由 $(*)$，$x - y + z = 0$，所以 $F = 0$，即 $(**)$ 成立。

第 (ii) 部分

現在 a、b、c 是任意滿足 $(**)$ 且 $b < c < a$ 的正數。如前定義 x、y、z；它們都是正數，而由於 $t \mapsto 1/\sqrt{t}$ 在 $t > 0$ 上是遞減（decreasing）的，條件 $b < c < a$ 給出

$$y > z > x > 0.$$

如第 (i) 部分所示，$(**)$ 等價於

$$(x - y - z)(x - y + z)(x + y - z)(x + y + z) = 0,$$

所以四個因式之中有一個為零。其中三個不可能為零：

- $x+y+z>0$，因為三者都是正數；
- $x+y-z>0$，因為 $y>z$ 且 $x>0$；
- $x-y-z<0$，因為 $x<y$ 且 $z>0$。

因此 $x-y+z=0$，即 $y=x+z$，亦即

$$\frac{1}{\sqrt{b}}=\frac{1}{\sqrt{a}}+\frac{1}{\sqrt{c}},$$

這正是 $(*)$。（其實只需要用到 b 是三者中最小這一點。）

驗算

- 取 $a=4$、$c=1$。那麼由 $(*)$ 得 $1/\sqrt{b}=\frac{1}{2}+1=\frac{3}{2}$，所以 $b=\frac{4}{9}$。此時 $2\left(\frac{1}{16}+\frac{81}{16}+1\right)=\frac{49}{4}$，而 $\left(\frac{1}{4}+\frac{9}{4}+1\right)^2=\left(\frac{7}{2}\right)^2=\frac{49}{4}$，正如 $(**)$ 所要求。
- 為甚麼 (ii) 需要大小次序的條件？取 $a=4$、$c=1$，關係式 $(**)$ 是 $1/b$ 的二次方程，有兩個根：$b=\frac{4}{9}$ 和 $b=4$（由 $1/\sqrt{b}=|x-z|=\frac{1}{2}$ 得出）。第二個根使 b 成為最大而非最小，它滿足 $(**)$ 卻不滿足 $(*)$：此時以 $1/\sqrt{c}=1/\sqrt{a}+1/\sqrt{b}$ 聯繫起來的是 a 和 b。對調字母的角色，對應於對調哪一個圓位於中間。
- 以電腦對數以千計隨機生成、在一條直線上互相相切的三圓構形作檢驗，$(*)$ 和 $(**)$ 均得到證實。

常見錯誤

- 把圓心之間的距離 r_1+r_2 當作切點之間的距離：兩個圓心的高度並不相同。
- 在第 (ii) 部分，未排除其餘三個因式便從因式分解得出 $(*)$；這正是需要用到 $b<c<a$ 的地方。

STEP 之外

關係式 $(**)$ 是笛卡兒圓定理（Descartes' circle theorem）的退化情況（degenerate case）：對於四個兩兩相切、曲率（curvature）為 $k_i=1/r_i$ 的圓，有 $2\sum k_i^2=\left(\sum k_i\right)^2$，而一條直線可視為曲率為 0 的圓。不斷在空隙中填入新的圓，便得到阿波羅尼斯墊片（Apollonian gasket），其中整數曲率引出令人驚訝的數論（number theory）結果。F 的因式分解也正是希羅公式（Heron's formula）背後的那一個：$16\Delta^2=-F$，其中 x,y,z 為三角形的

三邊。若想進一步了解圓的幾何以及能令這些結果一目了然的反演（inversion），可參閱 Needham 的 *Visual Complex Analysis*（《直觀複分析》）。

詳細解答

STEP I 2016，第 6 題

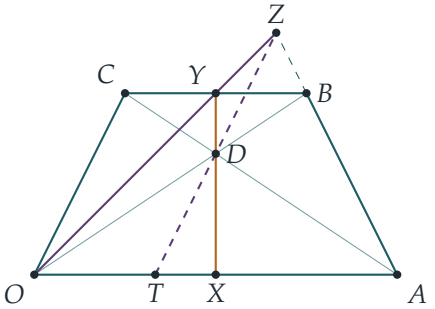
純數學 • 向量幾何

> **STEP I 2016，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/16-s1-q6

關鍵思路

圖中每一點都是兩條直線的交點，因此我們把每條直線寫成向量形式（vector form），然後令兩者相等。由於 **a** 與 **b** 不平行（parallel），任何向量都只能以唯一一種方式寫成 $\alpha\mathbf{a}+\beta\mathbf{b}$，所以把同一點的兩個表達式相等，就可以分別比較 **a** 與 **b** 的係數（coefficients）。全程以 **a**、**b** 為基（basis），每一步都只需解兩條簡短的線性方程（linear equations）。

解答



c 與 x 的形式

由於 CB 與 OA 平行，向量 $\overrightarrow{BC}=\mathbf{c}-\mathbf{b}$ 是 **a** 的純量倍數(scalar multiple)，設 $\mathbf{c}-\mathbf{b}=k\mathbf{a}$，於是 $\mathbf{c}=k\mathbf{a}+\mathbf{b}$。沿四邊形（quadrilateral）$O\to A\to B\to C$ 走一圈，邊 BC 的方向必須與 OA 相反（若方向相同，邊 AB 與 CO 便會相交，$OABC$ 就不是真正的四邊形），而且 $C\neq B$。所以

$$k<0.$$

任何負值都有可能；$k = -1$ 時就是平行四邊形（parallelogram）。

由於 X 在 OA 上，$\mathbf{x}$ 是 $\mathbf{a}$ 的倍數，設 $\mathbf{x} = m\mathbf{a}$；而 X 嚴格位於 O 與 A 之間，即

$$0 < m < 1.$$

此外，由於 O、A、B 不共線（collinear），$\mathbf{a}$ 與 $\mathbf{b}$ 不平行。

點 Z

求 D。OB 上的點為 $\lambda\mathbf{b}$；AC 上的點為

$$\mathbf{a} + \mu(\mathbf{c} - \mathbf{a}) = (1 - \mu + \mu k)\mathbf{a} + \mu\mathbf{b}.$$

比較係數，得 $\lambda = \mu$ 及 $1 - \mu(1 - k) = 0$。因為 $k < 0$，所以 $1 - k > 0$，故 $\mu = \frac{1}{1-k}$，且

$$\mathbf{d} = \frac{\mathbf{b}}{1-k}.$$

求 Y。BC 上的點為 $\mathbf{b} + t\mathbf{a}$。XD 上的點為

$$m\mathbf{a} + s(\mathbf{d} - m\mathbf{a}) = m(1-s)\mathbf{a} + \frac{s}{1-k}\,\mathbf{b}.$$

由 $\mathbf{b}$ 的係數得 $s = 1 - k$，於是 $\mathbf{a}$ 的係數為 $m(1 - (1-k)) = mk$。所以

$$\mathbf{y} = \mathbf{b} + mk\,\mathbf{a}.$$

求 Z。OY 上的點為 $\rho(\mathbf{b} + mk\mathbf{a})$；$AB$ 上的點為 $\mathbf{a} + \sigma(\mathbf{b} - \mathbf{a}) = (1-\sigma)\mathbf{a} + \sigma\mathbf{b}$。比較係數，得 $\rho = \sigma$ 及 $\rho mk = 1 - \sigma$，所以 $\rho(mk + 1) = 1$。只要 $mk + 1 \neq 0$，便有

$$\mathbf{z} = \frac{\mathbf{b} + mk\,\mathbf{a}}{mk+1},$$

此即所求。（若 $mk = -1$——這在 $k < -1$ 時可能發生——則 $\mathbf{y} = \mathbf{b} - \mathbf{a}$，所以 OY 與 AB 平行，Z 不存在。）

想一想 Z 是否在邊 AB 本身之上？

不是：因為 $-1 < mk < 0$，所以 $\sigma = \frac{1}{mk+1} > 1$，即 Z 位於 AB 越過 B 的延長線上，正如圖中所示。這並不要緊，因為題目討論的是直線。

點 T 與兩條關係式

DZ 上的點為 $\mathbf{d} + \tau(\mathbf{z} - \mathbf{d})$。其 $\mathbf{b}$ 的係數為

$$\frac{1}{1-k} + \tau\left(\frac{1}{mk+1} - \frac{1}{1-k}\right) = \frac{1}{1-k} - \tau\,\frac{k(1+m)}{(mk+1)(1-k)},$$

這裏用了 $(1-k)-(mk+1) = -k(1+m)$。要 T 在 OA 上，此係數必須為零，所以 $\tau = \dfrac{mk+1}{k(1+m)}$（留意 $k \neq 0$）。於是 $\mathbf{a}$ 的係數為

$$\tau \cdot \frac{mk}{mk+1} = \frac{mk+1}{k(1+m)} \cdot \frac{mk}{mk+1} = \frac{m}{1+m}.$$

因此

$$\mathbf{t} = \frac{m}{1+m}\,\mathbf{a}.$$

值得注意的是，k 消失了：T 只取決於 X 的位置。

現設 $a = |\mathbf{a}| = OA$。由於 $0 < \frac{m}{1+m} < 1$，T 位於 O 與 A 之間，所以

$$OT = \frac{ma}{1+m}, \qquad TA = a - \frac{ma}{1+m} = \frac{a}{1+m}, \qquad OX = ma.$$

因此

$$OT \times OA = \frac{ma^2}{1+m} = ma \times \frac{a}{1+m} = OX \times TA,$$

以及

$$\frac{1}{OX} + \frac{1}{OA} = \frac{1}{ma} + \frac{1}{a} = \frac{1+m}{ma} = \frac{1}{OT},$$

這兩條正是所要證明的結果。

驗算

試用圖中的梯形（trapezium）：$\mathbf{a} = (4,0)$，$\mathbf{b} = (3,2)$，$k = -\frac{1}{2}$，$m = \frac{1}{2}$。於是 $D = (2, \frac{4}{3})$，$Y = (2,2)$，$Z = (\frac{8}{3}, \frac{8}{3})$，$T = (\frac{4}{3}, 0)$；你可以驗證 T、D、Z 共線，而且 $\frac{1}{OT} = \frac{3}{4} = \frac{1}{2} + \frac{1}{4} = \frac{1}{OX} + \frac{1}{OA}$。以電腦對 20,000 個隨機梯形作檢驗，結果與上述每條公式的誤差都在 10^{-10} 以內。極限情況 $m \to 1$（$X \to A$）給出 $OT = \frac{1}{2}OA$，與 $\frac{1}{OT} = \frac{2}{OA}$ 一致。

常見錯誤

- 取 $k > 0$：若 $\mathbf{c} = \mathbf{b} + k\mathbf{a}$ 且 $k > 0$，這個「四邊形」會自我相交。
- 比較係數時沒有說明為何可以這樣做：這要依賴 $\mathbf{a}$ 與 $\mathbf{b}$ 不平行。
- 對兩條不同的直線使用同一個參數（parameter）字母，這會強加一個虛假的重合。

STEP 之外

T 不取決於 k（亦不取決於 $\mathbf{a}$ 與 $\mathbf{b}$ 之間的夾角），這說明它其實是仿射幾何（affine geometry），甚至是射影幾何（projective geometry）中的定理：要緊的只是點線的關聯

關係，以及同一直線上的比。關係式 $\frac{1}{OT} = \frac{1}{OX} + \frac{1}{OA}$ 正是交比（cross-ratio）與調和點列（harmonic range）能優雅處理的那類命題；而重複同一作圖，便可依次得到位於 $\frac{m}{1+2m}\mathbf{a}$、$\frac{m}{1+3m}\mathbf{a}$ 等處的點。如欲深入了解，可參閱考克斯特（Coxeter）與 Greitzer 合著的 *Geometry Revisited*（《幾何學的新探索》），或考克斯特的 *Projective Geometry*（《射影幾何》）。

詳細解答

STEP I 2016，第 7 題

純數學 • 模 4 的因式分解

> **STEP I 2016，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/16-s1-q7

本題在 *Beyond STEP* 第 1.7 節因數分解失效之時中有完整的引導式講解：https://beyondstep.keithmcnulty.org/zh-hant/ch01/factorisation-fails.html

關鍵思路

一切都源於模（modulo）4 運算。奇數必定 $\equiv 1$ 或 $\equiv 3 \equiv -1 \pmod 4$，因此一組奇數的乘積屬於 S 還是 T，取決於其中來自 T 的因數個數是偶數還是奇數。單憑這一條奇偶性（parity）規則，便可解決第 (ii)、(iii) 及 (iv)(a) 部分。第 (iv)(b) 部分則要求你看出 S-質數不一定是質數，而這就令唯一分解（unique factorisation）失效。

解答

第 (i) 部分

每個正奇數除以 4 的餘數都是 1 或 3，而偶數則不會有這樣的餘數。所以 $S \cup T$ 是所有正奇數組成的集合。任何整數都不會有兩個不同的餘數，所以 $S \cap T$ 是空集（empty set）。

第 (ii) 部分

取 S 中的 $4a+1$ 和 $4b+1$，其中 $a, b \geq 0$ 為整數。那麼

$$(4a+1)(4b+1) = 16ab + 4a + 4b + 1 = 4(4ab + a + b) + 1,$$

這是一個餘數為 1 的正整數，所以它屬於 S。

至於 T，答案是否。例如 $3 \in T$，但 $3 \times 3 = 9 = 4 \times 2 + 1 \in S$，而 $S \cap T$ 是空集。事實

上這種情況必定發生：

$$(4a+3)(4b+3) = 16ab+12a+12b+9 = 4(4ab+3a+3b+2)+1 \in S.$$

我們亦需要混合的乘積：

$$(4a+1)(4b+3) = 16ab+12a+4b+3 = 4(4ab+3a+b)+3 \in T.$$

簡而言之，把 T 中的數寫成 $\equiv -1 \pmod 4$：一組奇數的乘積 $\equiv (-1)^k \pmod 4$，其中 k 是來自 T 的因數個數。所以若 k 為偶數，乘積屬於 S；若 k 為奇數，乘積屬於 T。

第 (iii) 部分

設 $n \in T$ 為合數（composite），並把 $n = p_1 p_2 \cdots p_r$ 寫成質數的乘積。由於 n 是奇數，每個 p_i 都是奇數，所以由第 (i) 部分，每個 p_i 都屬於 S 或 T。

用反證法（proof by contradiction），假設每個 p_i 都屬於 S。反覆應用第 (ii) 部分，得 $p_1 p_2 \in S$，然後 $(p_1 p_2)p_3 \in S$，如此類推，所以 $n \in S$。但 $n \in T$ 而 $S \cap T = \varnothing$，這是矛盾。因此 n 至少有一個質因數屬於 T。（用奇偶性的說法：要令 $n \equiv -1$，需要奇數個、因而至少一個 $\equiv -1$ 的因數。）

第 (iv)(a) 部分

首先留意 $1 \notin T$，所以 T 中每個元素都至少是 3。

每個 $n \in T$ 要麼是 T-質數，要麼是若干個 T-質數的乘積。我們對 n 使用強歸納法（strong induction）。若 n 是 T-質數，則已完成。否則 $n = t_1 t_2 \cdots t_k$，其中 $k \geq 2$ 且每個 $t_i \in T$。每個 $t_i \geq 3$，所以每個 $t_i \leq n/3 < n$，由歸納假設，每個 t_i 要麼是 T-質數，要麼是若干個 T-質數的乘積。把這些代入 $n = t_1 \cdots t_k$，便把 n 寫成 T-質數的乘積。

? **想一想** 我們得到了一個 T-質數的乘積，但為甚麼它們的個數必定是奇數？

每個 T-質數都屬於 T，所以它 $\equiv -1 \pmod 4$。根據第 (ii) 部分末尾的規則，m 個這樣的數的乘積因此 $\equiv (-1)^m \pmod 4$。若 m 為偶數，乘積便 $\equiv 1$，因而屬於 S 而非 T。由於乘積是 $n \in T$，所以 m 是奇數。因此每個 $n \in T$ 要麼是 T-質數，要麼是奇數個 T-質數的乘積，得證。

留意 T-質數不一定是質數：$15 = 3 \times 5$ 是 T-質數，因為 $5 \notin T$，而分解成大於 1 的數的唯一方法是 3×5。

第 (iv)(b) 部分

與上面關於 15 的觀察相同的想法，在 S 中同樣適用：兩個來自 T 的質數的乘積屬於 S，但在 S 之內無法分解。取

$$9 = 3 \times 3, \qquad 21 = 3 \times 7, \qquad 49 = 7 \times 7.$$

每個都屬於 S。每個都是 S-質數：除了 1 和它本身以外，它的因數只有 3 或 7，而這些屬於 T 而非 S。（對於 21，因數是 $1, 3, 7, 21$。）現在

$$441 = 9 \times 49 = 21 \times 21,$$

而 $441 = 4 \times 110 + 1 \in S$。這是 441 分解成 S-質數的兩種真正不同的方法：第一種用了 9 和 49，第二種用了兩次 21。所以 441 是一個例子。

想一想 你會怎樣找到這個例子？

在普通整數中，$441 = 3^2 \times 7^2$，而質數 3, 7 屬於 T。在 S 之內，我們不能單獨分出一個 3 或 7，所以 T 中的質數必須兩兩配對。有兩個 3 和兩個 7 時，配對方法有兩種：$(3 \cdot 3)(7 \cdot 7)$ 或 $(3 \cdot 7)(3 \cdot 7)$。任何 p^2q^2（其中 $p, q \in T$ 為不同質數）都同樣可行，例如 $9 \times 121 = 33 \times 33$。

驗算

- 試試小的例子：$5 \times 9 = 45 \in S$，$3 \times 7 = 21 \in S$，$5 \times 7 = 35 \in T$。
- 對於 (iii)：$15 = 3 \times 5$ 和 $35 = 5 \times 7$ 都屬於 T，而且各自都有一個屬於 T 的質因數（3、7）。
- 用電腦驗證：對每個小於 3000 的 $n \in T$，每種分解成 T-質數的方法都有奇數個因數。T-質數的開頭是 $3, 7, 11, 15, 19, 23, 31, 35, \ldots$。留意 $27 = 3 \times 3 \times 3$ 不是 T-質數：它有三個因數，是奇數。

常見錯誤

- 沒有測試任何例子便聲稱 T 中兩個元素的乘積屬於 T：$3 \times 3 = 9$ 立即否定了這說法。
- 忘記 X-質數不一定是質數。這正是 (iv)(b) 的重點：9 和 21 都是 S-質數。
- 在 (iv)(a) 中只證明了因數個數是奇數，卻忘記證明分解成 T-質數的方法確實存在。

STEP 之外

集合 S 對乘法封閉，而它的「質數」並不能唯一分解。這是一個經典的玩具模型（有時稱為希爾伯特數（Hilbert numbers），以希爾伯特（Hilbert）命名），說明了在 $\mathbb{Z}[\sqrt{-5}]$ 這類環（ring）中出現的問題：在那裏 $6 = 2 \times 3 = (1 + \sqrt{-5})(1 - \sqrt{-5})$。為了在這類環中修補唯一分解，數學家發展出理想（ideals）和代數數論（algebraic number theory）。Davenport 的 *The Higher Arithmetic*（《高等算術》）解釋了為甚麼 $\mathbb{Z}$ 中的唯一分解是一條需要證明的定理，而 Allenby 的 *Rings, Fields and Groups*（《環、域與羣》）則討論了唯一分解失效的環。

詳細解答

STEP I 2016，第 8 題

純數學 • 生成函數

> **STEP I 2016，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/16-s1-q8

本題在 *Beyond STEP* 第 9.4 節把數列看成函數中有完整的引導式講解：https://beyondstep.keithmcnulty.org/zh-hant/ch09/generating-functions.html

關鍵思路

生成函數（generating function）把整個數列包裝成單一個函數，而對數列的運算就變成對這個函數的代數運算。在第 (i) 部分，$1-x$ 的負數次冪的二項級數（binomial series）給出我們想要的係數，然後我們用已知的數列組合出 n^2。在第 (ii) 部分，我們把遞推關係（recurrence）乘以 x^n 再求和，從而化成關於 f 的方程：把數列的下標平移，等同於把它的生成函數乘以 x。

解答

第 (i) 部分

由指數為 -2 的二項級數，

$$(1-x)^{-2} = \sum_{n\geq 0} \binom{-2}{n} (-x)^n = 1 + 2x + 3x^2 + 4x^3 + \cdots = \sum_{n\geq 0} (n+1)x^n,$$

因為 $\binom{-2}{n}(-1)^n = \frac{2\cdot 3\cdots(n+1)}{n!} = n+1$。乘以 x，

$$x(1-x)^{-2} = \sum_{n\geq 0} (n+1)x^{n+1} = \sum_{n\geq 0} nx^n,$$

這裏重新編排求和下標後，可以無礙地把 $n=0$ 也包括在內（該項為 0）。所以 $u_n = n$ 的生成函數是 $x(1-x)^{-2}$，如題所求。

同理，取指數 -3，

$$\binom{-3}{n}(-1)^n = \frac{3\cdot 4\cdots(n+2)}{n!} = \frac{(n+1)(n+2)}{2},$$

所以

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

乘以 x 並如前重新編排下標，

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

所以該數列是 $u_n = \frac{1}{2}n(n+1)$，即 $0, 1, 3, 6, 10, \ldots$，也就是三角數（triangular numbers）。

想一想 怎樣從我們現在已知的兩個數列得出 n^2？

生成函數是線性（linear）的：若 u_n 和 v_n 的生成函數分別是 f 和 g，則 $\alpha u_n + \beta v_n$ 的生成函數是 $\alpha \mathrm{f} + \beta \mathrm{g}$。由於

$$n^2 = 2\cdot\tfrac{1}{2}n(n+1) - n,$$

$u_n = n^2$ 的生成函數是

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

第 (ii)(a) 部分

對每個 $n \geq 1$，有 $u_n x^n = k u_{n-1} x^n$。對所有 $n \geq 1$ 求和：

$$\sum_{n\geq 1} u_n x^n = kx \sum_{n\geq 1} u_{n-1}x^{n-1}$$
$$\mathrm{f}(x) - u_0 = kx\,\mathrm{f}(x),$$

因為左邊的和是 $\mathrm{f}(x)$ 去掉常數項，而右邊的和在令 $m = n-1$ 後是 $\sum_{m\geq 0} u_m x^m = \mathrm{f}(x)$。由於 $u_0 = a$，這就是 $\mathrm{f}(x) = a + kx\,\mathrm{f}(x)$，如題所求。因此

$$\mathrm{f}(x) = \frac{a}{1-kx}.$$

（作為驗算，展開得 $a + akx + ak^2x^2 + \cdots$，而確實 $u_n = ak^n$。）

第 (ii)(b) 部分

把 $u_n = u_{n-1} + u_{n-2}$ 乘以 x^n，並對 $n \geq 2$ 求和（遞推關係只由 $n = 2$ 起成立，所以必須由這裏開始）：

$$\sum_{n\geq 2} u_n x^n = x\sum_{n\geq 2} u_{n-1}x^{n-1} + x^2\sum_{n\geq 2} u_{n-2}x^{n-2}.$$

想一想 這三個和各自缺少了 f 的哪些項？

左邊缺少 $u_0 + u_1x$ 這些項；右邊第一個和是 $u_1x + u_2x^2 + \cdots$，所以只缺少 u_0；第二個和就是整個 $\mathrm{f}(x)$。因此

$$\mathrm{f}(x) - u_0 - u_1x = x\big(\mathrm{f}(x) - u_0\big) + x^2\,\mathrm{f}(x).$$

代入 $u_0 = 0$ 和 $u_1 = 1$，這變成 $\mathrm{f}(x) - x = x\,\mathrm{f}(x) + x^2\,\mathrm{f}(x)$，所以

$$\mathrm{f}(x) = \frac{x}{1 - x - x^2}.$$

這是斐波那契數（Fibonacci numbers）$0, 1, 1, 2, 3, 5, 8, \ldots$ 的生成函數。

驗算

- 展開 $x(1+x)(1-x)^{-3} = (x+x^2)(1+3x+6x^2+10x^3+\cdots)$ 得 $x + 4x^2 + 9x^3 + 16x^4 + \cdots$，正是平方數，符合預期。
- 以長除法（long division）把 x 除以 $1 - x - x^2$，得 $x + x^2 + 2x^3 + 3x^4 + 5x^5 + \cdots$。
- 在 $x = 0.1$：和 $\sum n^2(0.1)^n$ 與 $0.11/0.9^3 = 0.150891\ldots$ 相符，而斐波那契數的 $\sum u_n(0.1)^n$ 與 $0.1/0.89 = 10/89 = 0.112359\ldots$ 吻合。這些結果，連同題目中每個生成函數的首 30 個係數，均已用電腦核實。

常見錯誤

- 弄錯負指數的二項式係數（binomial coefficients）：來自 $\binom{-r}{n}$ 和來自 $(-x)^n$ 的正負號互相抵消，所以每個係數都是正數。
- 在 (ii)(b) 中由 $n = 0$ 或 $n = 1$ 開始求和，而遞推關係在那裏並不成立；或者忘記減去首項 u_0 和 u_1x。
- 把三角數寫成 $\frac{1}{2}n(n-1)$：檢查一下 $u_1 = 1$。

STEP 之外

把 $1 - x - x^2 = (1 - \phi x)(1 - \psi x)$ 因式分解，其中 $\phi, \psi = \frac{1}{2}(1 \pm \sqrt{5})$，再運用部分分式（partial fractions）配合第 (ii)(a) 部分，便得到比奈公式（Binet's formula）$u_n = (\phi^n - \psi^n)/\sqrt{5}$。同一方法可以解任何常係數線性遞推關係，而生成函數是組合數學（combinatorics）和概率中的核心工具（在概率中它們成為概率生成函數（probability generating functions））。Grimmett 和 Welsh 的 *Probability: An Introduction*（《概率導論》）闡述概

率方面的內容；Wilf 的 *generatingfunctionology*（《生成函數學》）則是一本生動有趣、可免費取得的書，講述組合方面的內容。

詳細解答

STEP I 2016，第 9 題

力學 • 極限摩擦力平衡

> **STEP I 2016，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/16-s1-q9

關鍵思路

對 A 取力矩（moments）：牆上的兩個力因此消去，欄杆處的法向反作用力（normal reaction）可直接以 θ 表示，完全不涉及摩擦力（friction）。只要知道每個摩擦力的方向，再沿水平和鉛垂方向分解（resolving）便可完成。這正是本題的真正重點：摩擦力與桿即將發生的運動方向相反，所以我們必須判斷端點 A 是即將沿牆向下滑，還是向上滑。

解答

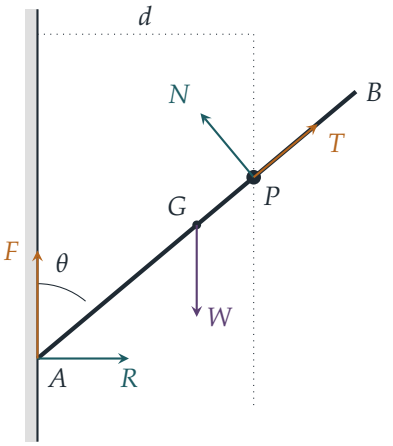
把桿模型化為重 W 的均勻剛性桿（uniform rigid rod），其中心 G 與 A 的距離為 a；把欄杆模型化為一條固定直線，與桿只相交於一點 P。由於 A 到欄杆的水平距離為 d，而桿與鉛垂線成角 θ，

$$AP = \frac{d}{\sin\theta}.$$

條件 $a\sin\theta < d$ 表示 G 位於 A 與 P 之間。

各力如下：

- 在 A：法向反作用力 R，方向水平並離開牆壁；以及沿牆的摩擦力 F；
- 在 P：垂直於桿的法向反作用力 N；以及沿桿的摩擦力 T；
- 作用於 G 的重量 W。



圖中 F 和 T 的箭頭方向對應下文的第一種情況。

對 A 取力矩

由於 N 垂直於桿，N 對 A 的力矩為 $N \cdot AP$；T 沿桿作用，對 A 沒有力矩，R 和 F 亦然。重量的力臂（lever arm）為 $a\sin\theta$。因此

$$N\frac{d}{\sin\theta} = Wa\sin\theta \quad\Longrightarrow\quad N = \frac{Wa\sin^2\theta}{d}. \tag{1}$$

無論桿即將向哪個方向滑動，此式均成立。

桿向哪個方向滑動？

想一想 A 是即將沿牆向下滑，還是向上滑？

先假設完全沒有摩擦力。沿鉛垂方向分解，需要 $N\sin\theta = W$，結合 (1) 即 $a\sin^3\theta = d$。但此處 $a\sin^3\theta < a\sin\theta < d$，故由 (1) 得 $N\sin\theta < W$：單靠法向反作用力無法撐住桿，其餘部分必須由摩擦力提供。桿即將向下滑動：A 沿牆下滑，θ 減小，$AP = d/\sin\theta$ 增大，於是桿穿過欄杆向牆壁方向滑下。摩擦力阻止這個運動，所以 F 沿牆向上，T 沿桿向上（指向 B），如圖所示。在兩個接觸點均處於極限平衡（limiting equilibrium）時，

$$F = \mu R, \qquad T = \lambda N.$$

分解各力

沿桿向上的單位向量（unit vector）為 $(\sin\theta, \cos\theta)$，而 N 的方向為 $(-\cos\theta, \sin\theta)$（指向牆壁並向上）。沿水平方向分解：

$$R - N\cos\theta + \lambda N\sin\theta = 0 \quad\Longrightarrow\quad R = N(\cos\theta - \lambda\sin\theta).$$

（牆壁只能推，故需 $R \geq 0$，所以此情況要求 $\tan\theta \leq 1/\lambda$。）沿鉛垂方向分解：

$$\begin{aligned} W &= \mu R + N\sin\theta + \lambda N\cos\theta \\ &= N\big(\mu\cos\theta - \lambda\mu\sin\theta + \sin\theta + \lambda\cos\theta\big) \\ &= N\big((\lambda+\mu)\cos\theta + (1-\lambda\mu)\sin\theta\big). \end{aligned}$$

代入 (1) 中的 N，再除以 $W\sin^2\theta$，得

$$d\operatorname{cosec}^2\theta = a\big((\lambda+\mu)\cot\theta + (1-\lambda\mu)\big),$$

這就是所求的結果。

$a\sin\theta > d$ 的情況

此時 G 位於欄杆之外。方程 (1) 和幾何關係都不變，所以唯一的問題仍然是滑動的方向。

> **想一想** G 移到欄杆之外後，有甚麼改變？

上面的論證顯示，當 $a\sin^3\theta < d$ 時桿向下滑，當 $a\sin^3\theta > d$ 時桿向上滑。在 $a\sin\theta < d$ 的條件下只可能出現前者。在 $a\sin\theta > d$ 的條件下，後者變得可能：伸出欄杆之外部分的重量使 $N\sin\theta > W$，於是桿傾向穿過欄杆向遠離牆壁的方向滑下，而 A 沿牆向上移動。這是新的情況，此時兩個摩擦力的方向都相反：$F = \mu R$ 沿牆向下，$T = \lambda N$ 沿桿向下。在上面的推導中把 λ 換成 $-\lambda$、μ 換成 $-\mu$，得

$$R = N(\cos\theta + \lambda\sin\theta), \qquad W = N\big((1-\lambda\mu)\sin\theta - (\lambda+\mu)\cos\theta\big),$$

再結合 (1)，

$$d\operatorname{cosec}^2\theta = a\big((1-\lambda\mu)\cot\theta - (\lambda+\mu)\big).$$

此時 $R > 0$ 自動成立。（若 $a\sin^3\theta < d < a\sin\theta$，桿仍然即將向下滑，第一個結果照樣適用，因為其推導從未用到 G 的位置。）

驗算

- 沒有摩擦力時（$\lambda = \mu = 0$），兩個結果都化為 $d = a\sin^3\theta$，即上文所得的無摩擦平衡。
- 設 $\lambda = \tan\alpha$ 及 $\mu = \tan\beta$（即摩擦角（angle of friction）），右方分別變為 $a\sin(\theta+\alpha+\beta)/(\cos\alpha\cos\beta)$ 和 $a\sin(\theta-\alpha-\beta)/(\cos\alpha\cos\beta)$：兩種情況只在摩擦角的正負號上有分別，理應如此。

- 取 $\lambda = \mu = \frac{1}{2}$ 及 $\theta = \frac{1}{3}\pi$，兩種情況分別給出 $d/a = \frac{3}{8} + \frac{9\sqrt{3}}{32} \approx 0.862$ 和 $d/a = \frac{9\sqrt{3}}{32} - \frac{3}{8} \approx 0.112$。按上述摩擦力方向，把每個值代回力和力矩的方程，均得合力為零及力矩為零。

常見錯誤

- 以為欄杆的法向反作用力是鉛垂的。事實上它垂直於桿，因為桿與欄杆只在一點接觸。
- 在力矩方程中，以長度 d 而非 $d/\sin\theta$ 作為 A 到欄杆的距離。
- 未先想清楚桿實際上會向哪個方向移動，便定下摩擦力的方向。

STEP 之外

在向下和向上兩個極限情況之間，桿對一整個區間的角度 θ 都能保持平衡：摩擦力帶來的是一系列平衡狀態，而不是單一個。描述這類集合（每個接觸點的摩擦錐（friction cone），以及各接觸點能夠維持的狀態集合）是接觸力學（contact mechanics）以及機械人抓取分析的起點。Kleppner 與 Kolenkow 的 *An Introduction to Mechanics*（《力學導論》）以大學一年級程度，仔細講解摩擦力與剛體靜力學（rigid-body statics）。

詳細解答

STEP I 2016，第 10 題

力學 • 碰撞與恢復係數

STEP I 2016，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/16-s1-q10

關鍵思路

題中共有三次碰撞（collision），每一次都用同樣的兩條方程處理：動量守恆（conservation of momentum）和牛頓恢復定律（Newton's law of restitution）。頭兩次碰撞同時發生，但彼此獨立，所以可以分開計算，求出 B 和 C 相遇前一刻的速度。再由 B 最終靜止這個條件，得到一條聯繫 λ 與 e 的方程。

解答

我們把四個物體視為光滑桌面上的質點（particle），因此除碰撞外沒有水平力作用，而且每次碰撞都是正碰（direct collision，即沿同一直線發生）。取由 A 指向 D 的方向為正方向，於是 A 的初速度為 u，D 的初速度為 $-u$。



第 (i) 部分

A **撞** B。設碰撞後 A 和 B 的速度分別為 v_A 和 v_B。則

$$\begin{aligned} \text{動量：} \quad & \lambda m u = \lambda m v_A + m v_B, \\ \text{恢復：} \quad & v_B - v_A = eu. \end{aligned}$$

把 $v_B = v_A + eu$ 代入第一條方程，得 $(\lambda+1)v_A = (\lambda - e)u$，所以

$$v_A = \frac{(\lambda-e)u}{\lambda+1}, \qquad v_B = v_A + eu = \frac{\lambda(1+e)u}{\lambda+1}.$$

D **撞** C。設碰撞後 C 和 D 的速度分別為 v_C 和 v_D。則

$$\begin{aligned} \text{動量：} \quad & -mu = mv_C + mv_D, \\ \text{恢復：} \quad & v_D - v_C = e\big(0-(-u)\big) = eu. \end{aligned}$$

（碰撞前 C 與 D 以速率 u 互相接近；碰撞後 D 在 C 的右方，所以兩者以 $v_D - v_C$ 的速率分離。）解這兩條方程：

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

因此 C 以速率 $\frac{1}{2}(1+e)u$ 向 B 移動，而 D 則以較小的速率 $\frac{1}{2}(1-e)u$ 跟在後面向左移動。

B **撞** C。此時 B 以 v_B 向右移動，C 向左移動，所以兩者的接近速率為 $v_B - v_C$。碰撞後 B 靜止；設 C 的速度為 w。則

$$\begin{aligned} \text{動量：} \quad & mv_B + mv_C = 0 + mw, \\ \text{恢復：} \quad & w - 0 = e(v_B - v_C). \end{aligned}$$

消去 w，得 $v_B + v_C = e(v_B - v_C)$，即

$$(1-e)v_B = -(1+e)v_C.$$

代入上面求得的速度：

$$(1-e)\cdot\frac{\lambda(1+e)u}{\lambda+1} = (1+e)\cdot\tfrac{1}{2}(1+e)u.$$

由於 $u \neq 0$ 且 $1+e>0$，可以兩邊除以 $(1+e)u$：

$$\begin{aligned} 2\lambda(1-e) &= (1+e)(\lambda+1) \\ 2\lambda - 2\lambda e &= \lambda + 1 + e\lambda + e \\ \lambda - 1 &= e(3\lambda+1), \end{aligned}$$

又因 $3\lambda+1>0$，

$$e = \frac{\lambda-1}{3\lambda+1},$$

即為所求。

至於推論部分，

$$\frac{1}{3} - e = \frac{(3\lambda+1)-3(\lambda-1)}{3(3\lambda+1)} = \frac{4}{3(3\lambda+1)} > 0,$$

因為 $\lambda > 1$。故 $e < \frac{1}{3}$。（另外，由於 $\lambda>1$，有 $e>0$，所以這些碰撞確實不是完全非彈性碰撞（perfectly inelastic collision）。）

第 (ii) 部分

第三次碰撞後，C 的速度為

$$w = v_B + v_C = (1+e)u\left(\frac{\lambda}{\lambda+1} - \frac{1}{2}\right) = \frac{(1+e)(\lambda-1)u}{2(\lambda+1)},$$

由於 $\lambda > 1$，此值為正：C 現在向右移動，折返向 D，而 D 仍以速率 $\frac{1}{2}(1-e)u$ 向左移動。

> **想一想** 「速率相同」給出哪一條方程？又怎樣消去 e？

兩者速率相等時，

$$\frac{(1+e)(\lambda-1)}{2(\lambda+1)} = \frac{1-e}{2}.$$

由第 (i) 部分，$1+e = \dfrac{4\lambda}{3\lambda+1}$，$1-e = \dfrac{2(\lambda+1)}{3\lambda+1}$。代入並約去公因子 $\dfrac{1}{3\lambda+1}$：

$$\frac{4\lambda(\lambda-1)}{\lambda+1} = 2(\lambda+1)$$

$$2\lambda^2 - 2\lambda = \lambda^2 + 2\lambda + 1$$

$$\lambda^2 - 4\lambda - 1 = 0,$$

所以 $\lambda = 2 \pm \sqrt{5}$。只有 $\lambda = 2+\sqrt{5}$ 滿足 $\lambda > 1$（另一個根是負數）。於是

$$e = \frac{1+\sqrt{5}}{7+3\sqrt{5}} = \frac{(1+\sqrt{5})(7-3\sqrt{5})}{49-45} = \frac{4\sqrt{5}-8}{4} = \sqrt{5}-2.$$

因此

$$\lambda = 2+\sqrt{5}, \qquad e = \sqrt{5}-2.$$

驗算

- $e = \sqrt{5}-2 \approx 0.236$，確實介乎 0 與 $\frac{1}{3}$ 之間，符合第 (i) 部分的要求。
- 取 $\lambda = 2+\sqrt{5} \approx 4.236$、$u = 1$、$e \approx 0.236$，以數值逐一計算三次碰撞，得第三次碰撞前 $v_B = 1$（恰巧是精確值）及 $v_C \approx -0.618$；碰撞後 B 靜止，C 以約 0.382 的速率向右移動，而 D 以約 0.382 的速率向左移動。共同速率為 $\frac{1}{2}(1-e)u = \frac{1}{2}(3-\sqrt{5})u$。
- 極限情況：當 $\lambda \to 1$ 時，$e \to 0$。質量相等且 $e = 0$ 時，B 與 C 以大小相等、方向相反的速度相遇，因此兩者黏合後靜止，結果合理。

常見錯誤

- 在 C 與 D 碰撞的恢復方程中弄錯正負號：所有速度都應以同一個正方向表示，並小心運用「分離速率 $= e\times$ 接近速率」。

- 忘記 C 到達第三次碰撞時是朝向 B 移動的，所以接近速率是 $v_B - v_C = v_B + \frac{1}{2}(1+e)u$，而不是 $v_B - \frac{1}{2}(1+e)u$。
- 保留了負數根 $\lambda = 2 - \sqrt{5}$。

STEP 之外

帶恢復係數的碰撞是能量損失最簡單的模型：正碰中損失的動能（kinetic energy）與 $1 - e^2$ 成正比，而 $e = 1$ 就是能量守恆的彈性碰撞（elastic collision）。在大學一年級的課程中，這些概念會引出質心參考系（centre-of-mass frame）：在這個參考系中，任何二體碰撞看起來都是對稱的，上面的代數運算幾乎變得微不足道。Kleppner 和 Kolenkow 的 *An Introduction to Mechanics*（《力學導論》）就是以這種方式講解碰撞的。

詳細解答

STEP I 2016，第 11 題

力學 • 拋體的最大射程

STEP I 2016，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：

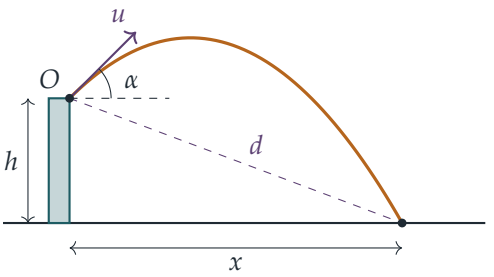
step.maths.org/questions/16-s1-q11

關鍵思路

寫出水平和鉛直方向的運動，消去時間，再把 $\sin\alpha$ 和 $\cos\alpha$ 的平方化成倍角（double angle）。至於最大射程，把落地方程視為以隱函數（implicit function）方式把 x 定義為 α 的函數，然後對它求導：在最大值處 $\mathrm{d}x/\mathrm{d}\alpha = 0$，餘下的就只是關於單一未知量 $\cos 2\alpha$ 的三角代數運算。

解答

我們把 P 模擬為只受重力作用的質點（particle）（不計空氣阻力（air resistance）），並假設平地是水平的。取坐標軸，原點在塔底，x 軸沿發射方向水平，y 軸鉛直向上，因此 O 為點 $(0, h)$。



落地方程

發射後時間 t 時，

$$X = ut\cos\alpha, \qquad Y = h + ut\sin\alpha - \tfrac{1}{2}gt^2.$$

塔是直立的，而 P 落在某個距離 $x > 0$ 處，所以 $\cos\alpha \neq 0$，且在落地時（即 $Y = 0$ 時）$t = \dfrac{x}{u\cos\alpha}$：

$$0 = h + x\tan\alpha - \frac{gx^2}{2u^2\cos^2\alpha}.$$

兩邊乘以 $2\cos^2\alpha$，並利用 $2\cos^2\alpha = 1 + \cos 2\alpha$ 及 $2\sin\alpha\cos\alpha = \sin 2\alpha$：

$$\frac{gx^2}{u^2} = 2h\cos^2\alpha + 2x\sin\alpha\cos\alpha = h(1+\cos 2\alpha) + x\sin 2\alpha,$$

這正是所求的方程。

最大水平距離

角 α 的取值範圍為 $-\frac{\pi}{2} < \alpha < \frac{\pi}{2}$（容許向下發射）。對每個這樣的 α，落地距離 x 是 α 的正值光滑函數，而且當 $\alpha \to \pm\frac{\pi}{2}$ 時 $x \to 0$，因為水平速率 $u\cos\alpha$ 趨於 0，而飛行時間保持有界。因此 x 在某個內部的 α 處取得最大值，而在該處 $\mathrm{d}x/\mathrm{d}\alpha = 0$。

把落地方程對 α 求導，記 $x' = \mathrm{d}x/\mathrm{d}\alpha$：

$$\frac{2gx}{u^2}\,x' = -2h\sin 2\alpha + x'\sin 2\alpha + 2x\cos 2\alpha.$$

令 $x' = 0$，得

$$x\cos 2\alpha = h\sin 2\alpha.$$

這裏 $\cos 2\alpha \neq 0$：否則由 $h\sin 2\alpha = 0$ 及 $h > 0$，便會迫使 $\sin 2\alpha = 0$ 同時成立，這是不可能的。所以在最大值處

$$x = h\tan 2\alpha,$$

即為所求。

> **想一想** 現在我們有兩條關於 x 和 α 的方程。怎樣才能乾淨俐落地消去 x？

把 $x = h\tan 2\alpha$ 代入落地方程，並記 $c = \cos 2\alpha$、$s = \sin 2\alpha$。右方變為

$$h(1+c) + h\frac{s^2}{c} = \frac{h\,(c + c^2 + s^2)}{c} = \frac{h(1+c)}{c},$$

而左方為

$$\frac{gh^2}{u^2}\cdot\frac{s^2}{c^2} = \frac{gh^2(1-c)(1+c)}{u^2c^2}.$$

令兩者相等，再除以 $h(1+c)/c$——它不為零，因為 $c \neq -1$（我們已證明 $c \neq 0$，而 $c = -1$ 會給出 $\alpha = \pm\frac{\pi}{2}$）：

$$\frac{gh(1-c)}{u^2c} = 1 \quad\Longrightarrow\quad gh - ghc = u^2c \quad\Longrightarrow\quad \cos 2\alpha = \frac{gh}{u^2+gh}.$$

由於這是駐點（stationary point）處唯一可能的值，而 $\tan 2\alpha = x/h > 0$ 確定了 $0 < 2\alpha < \frac{\pi}{2}$，所以恰好只有一個駐點，而它就是最大值。（順帶一提，$\tan 2\alpha = \dfrac{u\sqrt{u^2+2gh}}{gh}$，所以最大距離為 $x = \dfrac{u\sqrt{u^2+2gh}}{g}$。）

與 O 的最大距離

落地點在 $(x, 0)$，所以它與 $O = (0, h)$ 的距離為

$$d = \sqrt{x^2 + h^2}.$$

在 h 固定的情況下，d 隨 x 遞增，所以 d 恰好在 x 最大時取得最大值。在該處 $x = h\tan 2\alpha$，所以

$$d = \sqrt{h^2\tan^2 2\alpha + h^2} = h\sec 2\alpha = h \cdot \frac{u^2+gh}{gh} = \frac{u^2}{g} + h,$$

這裏用到 $\sec 2\alpha > 0$。這就是所求的最大距離。

驗算

- 當 $h = 0$ 時，得 $\cos 2\alpha = 0$，所以 $\alpha = 45°$，最大射程為 u^2/g：正是在平地上我們熟悉的結果。
- 取 $u = 10$、$g = 10$、$h = 5$（國際單位制）。那麼 $\cos 2\alpha = \frac{50}{150} = \frac{1}{3}$，最大射程為 $\sqrt{100\cdot 200}/10 = 10\sqrt{2} \approx 14.14$，而 $d = \sqrt{200+25} = 15 = \frac{100}{10} + 5$，與預測相符。用電腦對 α 進行搜索，亦證實 14.14 是最大的落地距離。
- 當 h 增大時，$\cos 2\alpha \to 1$，所以最佳角度趨於 0：從極高的塔上發射，應該差不多水平地射出。

常見錯誤

- 寫出 $x = h\tan 2\alpha$ 後，忘記它只在最大值處成立，而非對所有 α 都成立。
- 除以 $1 + \cos 2\alpha$ 或 $\cos 2\alpha$ 時，沒有說明它們為何不為零。
- 另行對 d 求最大值，而沒有留意到 $d^2 = x^2 + h^2$ 與 x 在同一角度取得最大值。

STEP 之外

答案 $d = u^2/g + h$ 有一個優美的幾何解釋。所有從 O 以速率 u 射出的軌跡，都與同一條包絡線（envelope）相切，這就是「安全拋物線」（parabola of safety），它的焦點

（focus）在 O，準線（directrix）位於 O 上方高度 u^2/g 處。最遠的落地點位於這條包絡線上，所以根據焦點—準線性質，它與 O 的距離等於它在準線下方的距離，即 $u^2/g + h$。曲線族的包絡線是大學一年級微積分和力學的標準課題。Kleppner 與 Kolenkow 的 *An Introduction to Mechanics*（《力學導論》）是下一步的好選擇，書中以同樣嚴謹的方式處理拋射體問題。

詳細解答

STEP I 2016，第 12 題

概率與統計 • 二項分佈的對稱性

STEP I 2016，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/16-s1-q12

本題在 *Beyond STEP* 第 12.6 節多一枚硬幣中有完整的引導式講解：https://beyondstep.keithmcnulty.org/zh-hant/ch12/one-more-coin.html

關鍵思路

第 (i) 部分和第 (ii) 部分是直接運用二項概率（binomial probabilities）的計算，而兩者都得出同一個簡單得驚人的答案。第 (iii) 部分則解釋箇中原因。把 Bob 的擲幣分成前 n 次和額外的一次：在前 n 次擲幣中，Bob 和 Alice 處於完全對等的地位，所以對稱性（symmetry）便能解決問題；而額外的一次擲幣則公平地打破平局。

解答

全文中，設 A 為 Alice 擲得正面的次數，B 為 Bob 擲得正面的次數。把一枚公正硬幣擲 m 次，正面次數 X 滿足

$$\mathrm{P}(X=k)=\binom{m}{k}\frac{1}{2^m},\qquad k=0,1,\ldots,m,$$

而 Alice 和 Bob 的擲幣互相獨立（independent），所以可以把 A 和 B 的概率相乘。

第 (i) 部分

Alice 擲兩次，Bob 擲三次，所以

k	0	1	2	3
$\mathrm{P}(A=k)$	$\frac{1}{4}$	$\frac{1}{2}$	$\frac{1}{4}$	
$\mathrm{P}(B=k)$	$\frac{1}{8}$	$\frac{3}{8}$	$\frac{3}{8}$	$\frac{1}{8}$

以 Alice 的得分為條件（condition on）。給定 $A = a$，我們需要 $B > a$，因此需要尾概率（tail probabilities）$\mathrm{P}(B > 0) = \frac{7}{8}$、$\mathrm{P}(B > 1) = \frac{1}{2}$、$\mathrm{P}(B > 2) = \frac{1}{8}$。故此

$$\begin{aligned}\mathrm{P}(B > A) &= \tfrac{1}{4} \cdot \tfrac{7}{8} + \tfrac{1}{2} \cdot \tfrac{1}{2} + \tfrac{1}{4} \cdot \tfrac{1}{8} \\ &= \frac{7+8+1}{32} = \frac{1}{2}.\end{aligned}$$

第 (ii) 部分

現在對 $a = 0,1,2,3$，有 $\mathrm{P}(A = a) = \frac{1}{8}, \frac{3}{8}, \frac{3}{8}, \frac{1}{8}$；而對 $k = 0,\ldots,4$，有 $\mathrm{P}(B = k) = \frac{1}{16}, \frac{4}{16}, \frac{6}{16}, \frac{4}{16}, \frac{1}{16}$。尾概率為

$$\mathrm{P}(B > 0) = \tfrac{15}{16}, \quad \mathrm{P}(B > 1) = \tfrac{11}{16}, \quad \mathrm{P}(B > 2) = \tfrac{5}{16}, \quad \mathrm{P}(B > 3) = \tfrac{1}{16}.$$

所以

$$\begin{aligned}\mathrm{P}(B > A) &= \tfrac{1}{8} \cdot \tfrac{15}{16} + \tfrac{3}{8} \cdot \tfrac{11}{16} + \tfrac{3}{8} \cdot \tfrac{5}{16} + \tfrac{1}{8} \cdot \tfrac{1}{16} \\ &= \frac{15+33+15+1}{128} = \frac{64}{128} = \frac{1}{2}.\end{aligned}$$

答案再一次恰好是 $\frac{1}{2}$。

第 (iii) 部分

首先，看看擲幣次數相同的遊戲的一個事實。當 Alice 和 Bob 各擲 n 次時，「Bob 較多」、「相同」、「Alice 較多」三者中恰有一個發生。兩名玩家的處境是對稱的（他們進行完全相同的試驗），所以「Alice 較多」的概率也是 p_2。因此

$$2p_2 + p_1 = 1. \qquad (*)$$

現在讓 Alice 擲 n 次，Bob 擲 $n+1$ 次。設 B' 為 Bob 前 n 次擲幣中正面的次數，並設 H 為他最後一次擲得正面的事件。那麼若 H 發生，則 $B = B' + 1$；否則 $B = B'$。

> ? **想一想** 比較 B 與 A 的結果，怎樣取決於比較 B' 與 A 的結果？

由於 B' 和 A 都是整數，而 B 不是 B' 就是 $B' + 1$：

- 若 $B' > A$，則無論最後一次結果如何，必然有 $B > A$；
- 若 $B' = A$，則 $B > A$ 當且僅當最後一次擲得正面；
- 若 $B' < A$，則 $B' \leq A - 1$，所以 $B \leq A$，Bob 不會勝出。

Bob 的前 n 次擲幣對 Alice 的 n 次擲幣，正正就是擲幣次數相同的遊戲，所以 $\mathrm{P}(B' > A) = p_2$ 及 $\mathrm{P}(B' = A) = p_1$。最後一次擲幣與其他一切獨立，且 $\mathrm{P}(H) = \frac{1}{2}$。故此

$$\mathrm{P}(\text{Bob 擲得的正面比 Alice 多}) = p_2 + \tfrac{1}{2}p_1.$$

由 $(*)$，$p_2 + \frac{1}{2}p_1 = \frac{1}{2}(2p_2 + p_1) = \frac{1}{2}$。

這就是所求的推廣（generalisation）：*若 Alice 把一枚公正硬幣擲 n 次，而 Bob 把一枚公正硬幣擲 $n+1$ 次，則對每個 $n \geq 0$，Bob 擲得的正面比 Alice 多的概率都是 $\frac{1}{2}$*。第 (i) 部分和第 (ii) 部分分別是 $n = 2$ 和 $n = 3$ 的情況。

驗算

- $n = 0$：Alice 不擲，Bob 擲一次；Bob 勝出當且僅當他擲得正面，概率為 $\frac{1}{2}$。
- $n = 2$：用手計算，$p_1 = \frac{1}{16} + \frac{4}{16} + \frac{1}{16} = \frac{3}{8}$，再由 $(*)$ 得 $p_2 = \frac{5}{16}$。於是 $p_2 + \frac{1}{2}p_1 = \frac{5}{16} + \frac{3}{16} = \frac{1}{2}$，與第 (i) 部分吻合。
- 用電腦精確窮舉（exact enumeration）證實，對每個不超過 29 的 n，都有 $\mathrm{P}(B > A) = \frac{1}{2}$；而對第 (ii) 部分進行 200 000 次試驗的模擬（simulation），得出 0.4997。

常見錯誤

- 在擲幣次數相同的遊戲中假設 $p_2 = \frac{1}{2}$：平局的概率是正數，所以 $p_2 = \frac{1}{2}(1 - p_1) < \frac{1}{2}$。
- 在第 (iii) 部分漏掉 $B' = A$ 的情況，或把它整個計入，而沒有乘上額外一次擲幣的因子 $\frac{1}{2}$。
- 用了尾概率 $\mathrm{P}(B \geq a)$ 而不是 $\mathrm{P}(B > a)$。

STEP 之外

這個簡潔的答案來自對稱性論證，再加上以額外一步為條件——這種模式貫穿整個概率論，從隨機遊走（random walks）到選票問題（ballot problems）都可見到。還有一個更巧妙的版本：由於 Bob 多擲一次，他未能擲得比 Alice 多的正面，當且僅當他擲得比她多的反面。把正面和反面互換，可見「正面較多」與「反面較多」的概率相等，而兩者互補（complementary），所以各自的概率都是 $\frac{1}{2}$。Grimmett 和 Welsh 的 *Probability: An Introduction*（《概率導論》）一書，對這些條件化和對稱性技巧有深入的發展。

詳細解答

STEP I 2016，第 13 題

概率與統計 • 指數分佈的順序統計量

STEP I 2016，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/16-s1-q13

關鍵思路

對於第一封和第二封郵件的到達，要問的是它們何時尚未發生。第一封電子郵件在時間 t 之後才到達，當且僅當全部 n 封都在 t 之後到達；第二封在 t 之後才到達，當且僅當到那時至多只有一封已到達。利用獨立性（independence），這兩個事件都很容易計算，再把生存概率（survival probability）求導，便得到概率密度函數（density）。

解答

設 $X_1, \ldots, X_n$ 為各封郵件的到達時間。它們互相獨立，而且每一個都滿足

$$\mathrm{P}(X_i > t) = \int_t^\infty \lambda \mathrm{e}^{-\lambda s}\,\mathrm{d}s = \mathrm{e}^{-\lambda t}, \qquad \mathrm{P}(X_i \le t) = 1 - \mathrm{e}^{-\lambda t}$$

其中 $t \ge 0$。我們會用到標準積分 $\int_0^\infty t k \mathrm{e}^{-kt}\,\mathrm{d}t = \frac{1}{k}$（$k > 0$），這正是參數為 k 的指數分佈（exponential distribution）的平均值。（用分部積分法（integration by parts）：$\left[-t\mathrm{e}^{-kt}\right]_0^\infty + \int_0^\infty \mathrm{e}^{-kt}\,\mathrm{d}t = 0 + \frac{1}{k}$。）

第 (i) 部分

第一封郵件在 t 之後才到達，當且僅當每一封郵件都在 t 之後才到達。由獨立性，

$$\mathrm{P}(T > t) = \prod_{i=1}^{n} \mathrm{P}(X_i > t) = \left(\mathrm{e}^{-\lambda t}\right)^n = \mathrm{e}^{-n\lambda t}.$$

所以累積分佈函數（cumulative distribution function）為 $F(t) = 1 - \mathrm{e}^{-n\lambda t}$，其中 $t \ge 0$。求導便得到密度函數

$$f(t) = F'(t) = n\lambda\,\mathrm{e}^{-n\lambda t} \qquad (t \ge 0),$$

正如所求。這是參數為 $n\lambda$ 的指數密度函數，所以取 $k = n\lambda$ 代入標準積分，得

$$\mathrm{E}(T) = \frac{1}{n\lambda}.$$

第 (ii) 部分

設 S 為第二封郵件的到達時間（因此需要 $n \geq 2$）。第二封郵件在 t 之後才到達，當且僅當到時間 t 為止至多一封郵件已到達：要麼一封也沒有，要麼恰好一封。

> **想一想** 到時間 t 為止恰好有一封郵件已到達，有多少種可能情況？

n 封郵件中任何一封都可以是那封提早到達的。那一封已到達（概率為 $1 - \mathrm{e}^{-\lambda t}$），而其餘 $n-1$ 封尚未到達（每封概率為 $\mathrm{e}^{-\lambda t}$）。由獨立性，並且由於這 n 種情況互斥（mutually exclusive），

$$\mathrm{P}(S > t) = \mathrm{e}^{-n\lambda t} + n\left(1 - \mathrm{e}^{-\lambda t}\right)\mathrm{e}^{-(n-1)\lambda t}.$$

展開第二項得 $n\mathrm{e}^{-(n-1)\lambda t} - n\mathrm{e}^{-n\lambda t}$，所以

$$\mathrm{P}(S > t) = n\,\mathrm{e}^{-(n-1)\lambda t} - (n-1)\,\mathrm{e}^{-n\lambda t}.$$

簡單驗證一下：當 $t = 0$ 時它等於 1，而當 $t \to \infty$ 時它趨於 0。

S 的密度函數 g 是 $1 - \mathrm{P}(S > t)$ 的導數：

$$\begin{aligned} g(t) &= -\frac{\mathrm{d}}{\mathrm{d}t}\left(n\,\mathrm{e}^{-(n-1)\lambda t} - (n-1)\,\mathrm{e}^{-n\lambda t}\right) \\ &= n(n-1)\lambda\,\mathrm{e}^{-(n-1)\lambda t} - n(n-1)\lambda\,\mathrm{e}^{-n\lambda t} \\ &= n(n-1)\lambda\left(\mathrm{e}^{-(n-1)\lambda t} - \mathrm{e}^{-n\lambda t}\right) \qquad (t \geq 0). \end{aligned}$$

為求期望值（expectation），把 g 寫成指數密度函數的組合：

$$g(t) = n \cdot (n-1)\lambda\,\mathrm{e}^{-(n-1)\lambda t} \;-\; (n-1)\cdot n\lambda\,\mathrm{e}^{-n\lambda t}.$$

分別取 $k = (n-1)\lambda$ 和 $k = n\lambda$ 代入標準積分（這是成立的，因為 $n \geq 2$，所以兩種情況下都有 $k > 0$），

$$\begin{aligned} \mathrm{E}(S) = \int_0^\infty t\,g(t)\,\mathrm{d}t &= \frac{n}{(n-1)\lambda} - \frac{n-1}{n\lambda} \\ &= \frac{1}{\lambda}\cdot\frac{n^2 - (n-1)^2}{n(n-1)} = \frac{1}{\lambda}\cdot\frac{2n-1}{n(n-1)}. \end{aligned}$$

由於 $\dfrac{1}{n-1} + \dfrac{1}{n} = \dfrac{n + (n-1)}{n(n-1)} = \dfrac{2n-1}{n(n-1)}$，所以

$$\mathrm{E}(S) = \frac{1}{\lambda}\left(\frac{1}{n-1} + \frac{1}{n}\right),$$

這正是所求的結果。

想一想 為甚麼答案是兩個簡單項之和？

指數分佈具有無記憶性（memoryless）。第一封郵件在一段速率為 $n\lambda$ 的指數時間後到達，平均值為 $\frac{1}{n\lambda}$。之後，其餘 $n-1$ 封郵件重新開始，就好像時鐘剛被重設一樣，所以額外的等待時間是速率為 $(n-1)\lambda$ 的指數分佈，平均值為 $\frac{1}{(n-1)\lambda}$。把兩段等待時間相加，立即得到答案，這也是記住它的好方法。

驗算

- 當 $n=2$ 且 $\lambda=1$ 時，第二封郵件的到達時間是兩個獨立指數時間中較晚的一個。它的平均值應為 $1+\frac{1}{2}=1.5$，而以電腦模擬 10^5 次試驗得到 1.501。
- 密度函數 g 的積分為 1：$\int_0^\infty g = n-(n-1)=1$。
- 當 $n=5$ 且 $\lambda=1$ 時，公式給出 $\mathrm{E}(S)=\frac{1}{4}+\frac{1}{5}=0.45$，與 $\int_0^\infty t\,g(t)\,\mathrm{d}t$ 的數值計算結果吻合。

常見錯誤

- 在恰好一封郵件已到達的情況中漏掉因子 n。n 封中任何一封都可以是提早到達的那封。
- 把 $\mathrm{P}(S>t)$ 寫成恰好一封已到達的概率，這樣便漏掉了一封也未到達的情況。

STEP 之外

T 和 S 是一個指數樣本的首兩個順序統計量（order statistics）。上面的無記憶性論證可以推廣，證明第 k 封郵件的到達時間平均值為 $\frac{1}{\lambda}\left(\frac{1}{n}+\frac{1}{n-1}+\cdots+\frac{1}{n-k+1}\right)$，而這正是泊松過程（Poisson process）和排隊論（queueing theory）的基礎。Blitzstein 和 Hwang 的 *Introduction to Probability*（《概率導論》）對順序統計量和無記憶性都有清晰的講解。

詳細解答

STEP II 2016，第 1 題

純數學 • 參數曲線與切線

> **STEP II 2016，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/16-s2-q1

關鍵思路

用參數（parameter）而不是坐標來處理。O 處的直角會變成兩點的參數 p 和 q 之間一條非常簡單的關係，而一般點上的切線（tangent）用 t 表示時有一條簡潔的方程。求出兩條切線的交點後，交點的坐標會是 p 和 q 的對稱函數（symmetric function），而 p 與 q 之間的關係正好足以把它們消去（eliminate）。

解答

一般點上的切線

在參數為 t 的點上，

$$\frac{\mathrm{d}y}{\mathrm{d}x} = \frac{\mathrm{d}y/\mathrm{d}t}{\mathrm{d}x/\mathrm{d}t} = \frac{3t^2}{2t} = \frac{3t}{2} \qquad (t \neq 0).$$

所以 (t^2, t^3) 處的切線是 $y - t^3 = \frac{3t}{2}(x - t^2)$，即

$$2y = 3tx - t^3. \qquad (*)$$

（當 $t = 0$ 時，這給出 $y = 0$，即尖點（cusp）處的切線，但我們不需要用到該點。）

直角

設 $P = (p^2, p^3)$ 及 $Q = (q^2, q^3)$。要令 $\angle POQ$ 有意義，兩點都不能是 O，所以 $p \neq 0$ 且 $q \neq 0$。該角為直角當且僅當 $\overrightarrow{OP} \cdot \overrightarrow{OQ} = 0$：

$$p^2q^2 + p^3q^3 = p^2q^2(1 + pq) = 0.$$

由於 $p^2q^2 \neq 0$，這迫使

$$pq = -1.$$

特別地，$p \neq q$（它們的積是負數），所以 P 和 Q 處的切線斜率（gradient）不同，$\frac{3p}{2} \neq \frac{3q}{2}$，確實相交於唯一一點。

切線的交點

把 $(*)$ 分別取 $t = p$ 和 $t = q$ 的兩個版本相減：

$$3(p-q)x = p^3 - q^3 = (p-q)(p^2+pq+q^2).$$

除以 $p - q \neq 0$，並利用 $pq = -1$，

$$3x = p^2 + pq + q^2 = p^2 + q^2 - 1.$$

代回 $t = p$ 時的 $(*)$：

$$\begin{aligned} 2y &= p(p^2+pq+q^2) - p^3 \\ &= p^2q + pq^2 = pq(p+q) = -(p+q). \end{aligned}$$

> **想一想** 我們已用兩個參數表示 x 和 y。怎樣才能得到一條只含 x 和 y 的方程？

把 y 平方：出現的交叉項 $2pq$ 是已知常數，剩下的是 $p^2 + q^2$，而這正是 $3x$ 所含的部分。事實上，

$$4y^2 = (p+q)^2 = p^2 + q^2 + 2pq = p^2 + q^2 - 2 = (3x+1) - 2 = 3x - 1.$$

所以 P 和 Q 處的切線相交於 $C_2\colon 4y^2 = 3x - 1$ 上，得證。

（事實上 C_2 上每一點都會出現：當 p 取遍所有正實數時，$2y = -(p+q) = \frac{1}{p} - p$ 取遍所有實數值，而 x 則由 C_2 的方程確定。）

C_1 與 C_2 是否相交？

把 $x = t^2$、$y = t^3$ 代入 C_2 的方程：

$$4t^6 = 3t^2 - 1 \iff 4u^3 - 3u + 1 = 0, \quad \text{其中} u = t^2 \geq 0.$$

顯然 $u = -1$ 是一個根，除去因式 $u + 1$ 得

$$4u^3 - 3u + 1 = (u+1)(4u^2 - 4u + 1) = (u+1)(2u-1)^2.$$

由於 $u \geq 0$，唯一可能是 $u = \frac{1}{2}$，所以 $t = \pm\frac{1}{\sqrt{2}}$。因此兩曲線確實相交，而且恰好交於兩點

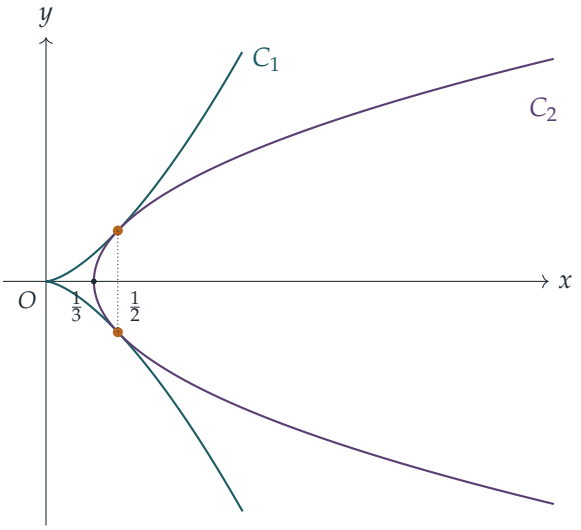
$$\left(\tfrac{1}{2}, \tfrac{1}{2\sqrt{2}}\right) \quad \text{及} \quad \left(\tfrac{1}{2}, -\tfrac{1}{2\sqrt{2}}\right).$$

重根（repeated root）告訴我們更多：兩曲線在該處相切，而不是相交穿過。我們可以直接驗證。在 $t = \frac{1}{\sqrt{2}}$ 處，C_1 的斜率是 $\frac{3t}{2} = \frac{3}{2\sqrt{2}}$。對 $4y^2 = 3x - 1$ 進行隱函數求導（implicit differentiation）得 $8y\frac{\mathrm{d}y}{\mathrm{d}x} = 3$，所以在 $y = \frac{1}{2\sqrt{2}}$ 處，C_2 的斜率是 $\frac{3}{8y} = \frac{3}{2\sqrt{2}}$，兩者相同。由於圖形關於 x 軸對稱，另一點亦然。

由於對 $u \geq 0$ 有 $4u^3 - 3u + 1 = (u+1)(2u-1)^2 \geq 0$，曲線 C_1 永遠不會走進拋物線（parabola）內部：除了兩個切點外，C_2 都位於 C_1 的兩條分支之間。

草圖

C_1 滿足 $y^2 = x^3$：它位於 $x \geq 0$ 的區域，關於 x 軸對稱，並在 O 處有一個尖點（兩條分支的斜率 $\frac{3t}{2} \to 0$，所以它們沿正 x 軸離開 O）。C_2 是拋物線 $x = \frac{1}{3}(4y^2+1)$，頂點（vertex）為 $(\frac{1}{3}, 0)$，開口向右，同樣關於 x 軸對稱。拋物線位於尖點之內、C_1 的兩條分支之間，並在 $x = \frac{1}{2}$ 處與每條分支各相切一次。



驗算

- 取 $p = 1$、$q = -1$：$P = (1,1)$，$Q = (1,-1)$，且 $\overrightarrow{OP} \cdot \overrightarrow{OQ} = 1 - 1 = 0$。切線為 $2y = 3x - 1$ 及 $2y = -3x + 1$，交於 $(\frac{1}{3}, 0)$，即 C_2 的頂點。
- 取 $p = 2$、$q = -\frac{1}{2}$：切線 $2y = 6x - 8$ 與 $2y = -\frac{3}{2}x + \frac{1}{8}$ 交於 $x = \frac{13}{12}$、$y = -\frac{3}{4}$，而 $4y^2 = \frac{9}{4} = 3 \cdot \frac{13}{12} - 1$。
- 切點 $(\frac{1}{2}, \frac{1}{2\sqrt{2}}) \approx (0.5, 0.35355)$ 滿足 $4y^2 = \frac{1}{2} = 3x - 1$，而該處兩條曲線的斜率都是 ≈ 1.06066。用電腦檢驗了 10^5 對滿足 $pq = -1$ 的隨機參數，所有交點在捨入誤差範圍內都落在 C_2 上。

常見錯誤

- 忘記排除 $p = 0$ 或 $q = 0$：在除去因式以得到 $pq = -1$ 之前，必須確定 p^2q^2 不為零。
- 除以 $p - q$ 時沒有說明它為何不為零。
- 停在「$4u^3 - 3u + 1 = 0$ 有根」，而沒有檢查它是否給出 $u = t^2 \geq 0$；根 $u = -1$ 並不對應 C_1 上的任何點。

STEP 之外

曲線 $y^2 = x^3$ 是經典的尖點三次曲線（cuspidal cubic），而把一條曲線與一條參數化曲線聯立、再從重根看出相切的技巧，正是代數幾何（algebraic geometry）的起點：相交重數（intersection multiplicity）計算曲線相交了多少次，而貝祖（Bézout）定理指出，一條三次曲線與一條二次曲線（conic）在正確計算重數下相交於六點。這裏的六點是二重根 $t = \pm\frac{1}{\sqrt{2}}$（共四個）和一對複數根 $t = \pm i$。一本淺白的入門書是 Kirwan 的 *Complex Algebraic Curves*（《複代數曲線》）。

詳細解答

STEP II 2016，第 2 題

純數學 • 因式定理

> **STEP II 2016，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/16-s2-q2

關鍵思路

把 $(*)$ 看成以 c 為變量的多項式（polynomial），其係數（coefficients）恰好涉及 a 和 b。因式定理（factor theorem）指出：$c-(a+b)$ 是因式（factor），當且僅當代入 $c=a+b$ 後結果為零。一旦找到一個因式，$(*)$ 對 a、b、c 的對稱性（symmetry）便免費送上另外兩個，再數一數次數（degree），因式分解（factorisation）就完成了。之後的兩條方程其實是經過偽裝的特殊情況：要做的就是選取 a、b、c（以及後來的 d、e）來配合它們。

解答

因式 $a+b-c$

記 $(*)$ 為 $F(a,b,c)$。視之為 c 的多項式，F 有因式 $c-(a+b)$（因而亦有因式 $a+b-c$），當且僅當 $F(a,b,a+b)=0$。設 $p=a+b$ 及 $c=p$。那麼 $a+b+c=2p$，並利用 $a^2+b^2=p^2-2ab$ 和 $a^3+b^3=p^3-3abp$，得

$$a^2+b^2+c^2=2p^2-2ab, \qquad a^3+b^3+c^3=2p^3-3abp.$$

所以

$$\begin{aligned} F(a,b,a+b) &= 8p^3-6(2p)(2p^2-2ab)+8(2p^3-3abp) \\ &= 8p^3-24p^3+24abp+16p^3-24abp=0, \end{aligned}$$

因此 $a+b-c$ 是 $(*)$ 的因式，正如所求。

將 $(*)$ 完全因式分解

無論怎樣重新排列 a、b、c，F 都不變。交換 b 與 c，因式 $a+b-c$ 變成 $a+c-b$；交換 a 與 c，它變成 $b+c-a$。所以以下三者

$$a+b-c, \qquad b+c-a, \qquad c+a-b$$

都是因式。它們是互不相同的線性（linear）多項式，沒有一個是另一個的常數倍，所以它們的乘積整除 F。F 的每一項次數都是 3，所以餘下的因式次數為 0：它是一個常數 k，而

$$F(a,b,c) = k(a+b-c)(b+c-a)(c+a-b).$$

為求 k，設 $a=b=c=1$：左方為 $27-6\cdot3\cdot3+8\cdot3=-3$，而乘積為 1。因此 $k=-3$，且

$$\begin{aligned}(a+b+c)^3 - 6(a+b+c)(a^2+b^2+c^2) + 8(a^3+b^3+c^3)\\ = -3(a+b-c)(b+c-a)(c+a-b).\end{aligned}$$

第 (i) 部分

把方程的中間項和最後一項除開，化成 $(*)$ 的形式：

$$(x+1)^3 - 6(x+1)\left(x^2+\tfrac{5}{2}\right) + 8\left(x^3+\tfrac{13}{4}\right) = 0.$$

想一想 a、b、c 取甚麼值，才能令這式子恰好就是 $(*)$？

試取 $a=x$，而 b、c 為常數。我們需要

$$b+c=1, \qquad b^2+c^2=\tfrac{5}{2}, \qquad b^3+c^3=\tfrac{13}{4}.$$

由前兩式得 $2bc=(b+c)^2-(b^2+c^2)=-\tfrac{3}{2}$，所以 $bc=-\tfrac{3}{4}$。於是

$$b^3+c^3=(b+c)^3-3bc(b+c)=1+\tfrac{9}{4}=\tfrac{13}{4},$$

所以第三個條件自動成立。現在 b、c 是 $t^2-t-\tfrac{3}{4}=0$ 的根，即 $(2t-3)(2t+1)=0$，所以 $b=\tfrac{3}{2}$，$c=-\tfrac{1}{2}$。

取 $a=x$、$b=\tfrac{3}{2}$、$c=-\tfrac{1}{2}$，因式分解給出

$$a+b-c=x+2, \qquad b+c-a=1-x, \qquad c+a-b=x-2,$$

所以方程為 $-3(x+2)(1-x)(x-2)=0$，而

$$x=-2, \quad x=1, \quad x=2.$$

第 (ii) 部分

記該四變量表達式為 $G(a,b,d,e)$。(設 $c=d+e$，則 $a+b-d-e$ 看起來就像 $a+b-c$，由此可猜到這個因式；不過直接驗證也一樣快。) 根據因式定理，把 G 視為 e 的多項式，只需證明當 $a+b=d+e$ 時 $G=0$。設 $p=a+b=d+e$。與之前完全一樣，$a^2+b^2=p^2-2ab$、$d^2+e^2=p^2-2de$、$a^3+b^3=p^3-3abp$ 及 $d^3+e^3=p^3-3dep$，所以設 $q=ab+de$，得

$$\begin{aligned} G &= (2p)^3-6(2p)(2p^2-2q)+8(2p^3-3pq)\\ &= 8p^3-24p^3+24pq+16p^3-24pq=0. \end{aligned}$$

所以 $a+b-d-e$ 是 G 的因式。

無論怎樣重新排列 a、b、d、e，G 都不變。因式 $a+b-d-e$ 來自把四個字母分成 $\{a,b\}$ 和 $\{d,e\}$ 兩對；另外兩種配對方法給出因式

$$a+d-b-e \quad (\text{交換}b,d), \qquad a+e-b-d \quad (\text{交換}b,e).$$

這三個線性因式互不相同，沒有一個是另一個的倍數，而 G 是 3 次齊次 (homogeneous) 多項式，所以對某常數 k，有 $G=k(a+b-d-e)(a+d-b-e)(a+e-b-d)$。設 $a=1$、$b=d=e=0$，左方得 $1-6+8=3$，右方得 k，所以

$$G=3(a+b-d-e)(a+d-b-e)(a+e-b-d).$$

至於方程，取 $a=x$，並取常數 b、d、e 使得

$$b+d+e=6, \qquad b^2+d^2+e^2=14, \qquad b^3+d^3+e^3=36.$$

數字 1、2、3 符合要求：$1+2+3=6$，$1+4+9=14$，$1+8+27=36$。所以取 $a=x$、$b=1$、$d=2$、$e=3$，方程恰好就是 $G=0$，其因式分解為

$$3(x+1-2-3)(x+2-1-3)(x+3-1-2)=3(x-4)(x-2)x=0.$$

因此

$$x=0, \quad x=2, \quad x=4.$$

驗算

- 把 $a=b=1$、$c=2$ 代入 $(*)$：$64-6\cdot4\cdot6+8\cdot10=0$，正如因式 $a+b-c$ 所預示。
- 在 (i) 中，$x=1$ 給出 $8-3\cdot2\cdot7+2\cdot17=8-42+34=0$。
- 在 (ii) 中，$x=0$ 給出 $216-6\cdot6\cdot14+8\cdot36=216-504+288=0$，而 $x=4$ 給出 $1000-6\cdot10\cdot30+8\cdot100=0$。
- 這些因式分解亦已用電腦在變量取幾千個隨機有理數值時驗證過。

常見錯誤

- 忘記常數：$(*)$ 是各因式乘積的 -3 倍，而不是乘積本身。務必用方便的代入來確定 k。
- 在 (ii) 中列出多於三個因式。例如交換 a 與 d，得到 $d+b-a-e$，這不過是 $-(a+e-b-d)$；三個真正不同的因式來自把四個字母兩兩配對的三種方法。
- 用蠻力把所有東西展開。這樣做行得通，但又長又容易出錯；把方程與恆等式配對才是這道題的重點。

STEP 之外

因式定理在任何域（field）上都成立；這裏我們把除一個以外的所有變量視為常數，從而把它用於多變量多項式。對稱性加上多項式的唯一因式分解（unique factorisation）是威力強大的組合：對稱多項式（symmetric polynomials）的理論指出，每個對稱多項式都可以用基本對稱多項式（elementary symmetric polynomials）$a+b+c$、$ab+bc+ca$、abc 表示，這正是 (i) 中關於 b、c 的條件背後隱藏的結構。順帶一提，乘積 $(a+b-c)(b+c-a)(c+a-b)$ 正是三角形面積的海倫公式（Heron's formula）中出現的式子。想進一步了解多項式環（polynomial rings）與因式分解，可參閱 Allenby 的 *Rings, Fields and Groups*（《環、域與羣》）。

詳細解答

STEP II 2016，第 3 題

純數學 • 多項式的根

> **STEP II 2016，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/16-s2-q3

關鍵思路

整道題建基於一條恆等式（identity）：$\mathrm{f}_n(x) = \mathrm{f}_{n-1}(x) + \dfrac{x^n}{n!}$。配合第 (i) 部分，它告訴我們：在 f_n 的根（root）a 處，斜率（gradient）為 $\mathrm{f}_n'(a) = -a^n/n!$。因此，即使不知道根是甚麼，我們也知道每個根處的斜率。由於所有根的正負號相同，這些斜率的正負號亦全部相同；而一條曲線不可能連續兩次向上穿過坐標軸，中間卻沒有一次向下穿過。

解答

全題中記 $\mathrm{f}_n(x) = \displaystyle\sum_{k=0}^{n} \frac{x^k}{k!}$。

第 (i) 部分

逐項微分（differentiating term by term），常數項 1 消失，而對 $1 \le k \le n$，

$$\frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{x^k}{k!}\right) = \frac{kx^{k-1}}{k!} = \frac{x^{k-1}}{(k-1)!}.$$

所以

$$\mathrm{f}_n'(x) = \sum_{k=1}^{n} \frac{x^{k-1}}{(k-1)!} = \sum_{j=0}^{n-1} \frac{x^j}{j!} = \mathrm{f}_{n-1}(x),$$

得證。

第 (ii) 部分

若 $x \geq 0$，則 $\mathrm{f}_n(x)$ 的每一項都非負，而首項是 1，所以 $\mathrm{f}_n(x) \geq 1 > 0$。因此沒有根能滿足 $a \geq 0$，任何實根都有 $a < 0$。（當 $n = 0$ 時，$\mathrm{f}_0 = 1$ 根本沒有根，與此一致。）

第 (iii) 部分

由定義，$\mathrm{f}_n(x) = \mathrm{f}_{n-1}(x) + \dfrac{x^n}{n!}$。若 $\mathrm{f}_n(a) = 0$，則由第 (i) 部分，

$$\mathrm{f}_n'(a) = \mathrm{f}_{n-1}(a) = \mathrm{f}_n(a) - \frac{a^n}{n!} = -\frac{a^n}{n!},$$

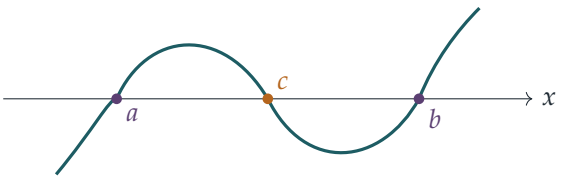
同理 $\mathrm{f}_n'(b) = -b^n/n!$。因此

$$\mathrm{f}_n'(a)\,\mathrm{f}_n'(b) = \frac{(ab)^n}{(n!)^2} > 0,$$

這是因為由第 (ii) 部分得 $a < 0$ 及 $b < 0$，所以 $ab > 0$。這正是所求的不等式。特別地，兩個斜率都不為零，而且正負號相同。

想一想 為甚麼兩個斜率正負號相同，就迫使兩者之間必有另一個根？

設 $a < b$，並假設兩個斜率皆為正（負的情況與之對稱）。由於 $\mathrm{f}_n'(a) > 0$，曲線在 a 處向上穿過坐標軸，所以當 x 略大於 a 時 $\mathrm{f}_n(x) > 0$。由於 $\mathrm{f}_n'(b) > 0$，曲線在 b 處亦向上穿過坐標軸，所以當 x 略小於 b 時 $\mathrm{f}_n(x) < 0$。一個連續函數若在某點為正、在較後的一點為負，則必在兩者之間取值零（介值定理（intermediate value theorem）），所以存在某個 c 滿足 $a < c < b$ 且 $\mathrm{f}_n(c) = 0$。



曲線上升穿過 a，又上升穿過 b，所以在兩者之間某處必定向下穿過坐標軸。

*最多只有一個實根。*假設 $(*)$ 有兩個或以上相異的實根。它是 n 次多項式（polynomial），所以只有有限個根，我們可以從中選出兩個*相鄰*（consecutive）的根 $a < b$，即兩者之間嚴格地沒有其他根。但上面的論證在 a 與 b 之間嚴格地找到一個根 c，矛盾。所以 $(*)$ 最多只有一個實根。（論證按題目所設使用了 $n \geq 2$；當 $n = 1$ 時唯一的根是 -1，當 $n = 0$ 時沒有根，所以結論對每個 n 都成立。）

*n 爲奇數。*此時 f_n 的次數為奇數，首項係數（leading coefficient）$1/n!$ 為正，所以當 $x \to -\infty$ 時 $\mathrm{f}_n(x) \to -\infty$，當 $x \to +\infty$ 時 $\mathrm{f}_n(x) \to +\infty$。由介值定理，它最少有一個實根，因此**恰好有一個**。

n 爲偶數。當 $n = 0$ 時沒有根。對偶數 $n \geq 2$，當 $x \to \pm\infty$ 時 $\mathrm{f}_n(x) \to +\infty$，所以 f_n 在某點 x_0 取得全域最小值（global minimum），而該處 $\mathrm{f}_n'(x_0) = \mathrm{f}_{n-1}(x_0) = 0$。由於 $\mathrm{f}_{n-1}(0) = 1 \neq 0$，故 $x_0 \neq 0$，於是

$$\mathrm{f}_n(x_0) = \mathrm{f}_{n-1}(x_0) + \frac{x_0^n}{n!} = \frac{x_0^n}{n!} > 0,$$

因為 n 是偶數。最小值為正，所以對所有 x 都有 $\mathrm{f}_n(x) > 0$，$(*)$ **沒有實根**。

（另一方法：由於 $\mathrm{f}_n'(a) \neq 0$，任何根 a 都會是單根（simple root），所以 f_n 在該處會變號；但一個首項係數為正的偶次多項式若恰好只變號一次，則它對所有很大的 x 或對所有很負的 x 必為負，這是不可能的。）

驗算

- $n = 1$：$1 + x = 0$ 只有一個根 -1。$n = 2$：$1 + x + \frac{1}{2}x^2 = \frac{1}{2}\big((x+1)^2 + 1\big) > 0$，沒有根。兩者皆吻合。
- $n = 3$：根約為 -1.596，而該處 $\mathrm{f}_3'(a) = \mathrm{f}_2(a) \approx 0.678 = -a^3/3!$，與預測相符。
- 用電腦檢查 f_n 對每個 $n \leq 25$ 的變號情況，發現 n 為奇數時恰有一個根，n 為偶數時沒有根。

常見錯誤

- 證明斜率之積為正時，沒有說明爲甚麼 $ab > 0$：這需要用到第 (ii) 部分。
- 由「任意兩根之間必有一根」推出「最多只有一個根」，卻沒有選取兩個相鄰的根；正是這個選取才導致矛盾。
- n 為偶數時，忘記檢查 $x_0 \neq 0$，而這是 $x_0^n/n! > 0$ 嚴格成立所必需的。

STEP 之外

多項式 f_n 是 e^x 的泰勒多項式（Taylor polynomials），而第 (i) 部分正是 $(\mathrm{e}^x)' = \mathrm{e}^x$ 在有限情況下的影子。由於 e^x 永不為零，f_n 只有很少實根也就順理成章；它們的複數根（complex roots）按 $1/n$ 縮放後，當 $n \to \infty$ 時會聚集在一條稱為塞格曲線（Szegő curve）的曲線上。這裏用到的工具——介值定理，以及連續函數必有最小值——會在第一門分析（analysis）課程中嚴謹地證明；可參閱 Abbott 的 *Understanding Analysis*（《理解分析》）。

詳細解答

STEP II 2016，第 4 題

純數學 • 有理函數的值域

> **STEP II 2016，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/16-s2-q4

關鍵思路

去分母後，y 的定義便化為一條關於 x 的二次方程（quadratic equation），而「x 是實數」就變成「判別式（discriminant）非負」。要消去 θ，我們需要一個對所有 θ 都成立的 $y\cos\theta - \sin\theta$ 的上界（upper bound）；把它寫成 $(y,-1)$ 與單位向量（unit vector）$(\cos\theta, \sin\theta)$ 的點積（dot product），便恰好得到 $\sqrt{y^2+1}$。在第 (ii) 部分，不等式鏈中的每一個不等式都必須取等號，而這正好把 x 和 θ 都確定下來。

解答

第 (i) 部分

首先留意分母永不為零：把 $x^2 + x\cos\theta + 1$ 看作 x 的二次式，其判別式為 $\cos^2\theta - 4 < 0$，所以它恆為正。因此可以放心地兩邊相乘：

$$y\left(x^2 + x\cos\theta + 1\right) = x^2 + x\sin\theta + 1 \iff (y-1)x^2 + (y\cos\theta - \sin\theta)\,x + (y-1) = 0.$$

若 $y \neq 1$，這是一條真正關於 x 的二次方程，而它有實根 x 當且僅當其判別式非負：

$$(y\cos\theta - \sin\theta)^2 - 4(y-1)^2 \geq 0.$$

若 $y = 1$，所求的不等式變為 $(\cos\theta - \sin\theta)^2 \geq 0$，這本來就成立。所以對實數 x，

$$(y\cos\theta - \sin\theta)^2 \geq 4(y-1)^2,$$

即為所求。

想一想 怎樣才能消去 θ？我們需要一個比 $(y\cos\theta-\sin\theta)^2$ 大、而又不依賴 θ 的式子。

式子 $y\cos\theta-\sin\theta$ 是 $(y,-1)$ 與單位向量 $(\cos\theta,\sin\theta)$ 的純量積（scalar product），所以其大小至多為 $|(y,-1)|=\sqrt{y^2+1}$。等價地，由柯西－施瓦茨不等式（Cauchy–Schwarz inequality），或直接由

$$(y^2+1)(\cos^2\theta+\sin^2\theta)-(y\cos\theta-\sin\theta)^2=(y\sin\theta+\cos\theta)^2\geq 0,$$

可得 $(y\cos\theta-\sin\theta)^2\leq y^2+1$。把這與第一個結果結合，

$$y^2+1\geq(y\cos\theta-\sin\theta)^2\geq 4(y-1)^2,$$

這就是第二個不等式。

展開得 $y^2+1\geq 4y^2-8y+4$，即

$$3y^2-8y+3\leq 0.$$

$3y^2-8y+3=0$ 的根為 $y=\dfrac{8\pm\sqrt{64-36}}{6}=\dfrac{4\pm\sqrt{7}}{3}$，而由於這個二次式的首項係數（leading coefficient）為正，它恰好在兩根之間取非正值。因此

$$\frac{4-\sqrt{7}}{3}\leq y\leq\frac{4+\sqrt{7}}{3}.$$

第 (ii) 部分

設 $y=\frac{1}{3}(4+\sqrt{7})$，則 $y-1=\frac{1}{3}(1+\sqrt{7})>0$。於是

$$y^2+1=\frac{16+8\sqrt{7}+7}{9}+1=\frac{32+8\sqrt{7}}{9},$$
$$4(y-1)^2=\frac{4\left(1+2\sqrt{7}+7\right)}{9}=\frac{32+8\sqrt{7}}{9}.$$

兩者相等，而且 $2(y-1)>0$，所以取正平方根便得 $\sqrt{y^2+1}=2(y-1)$，即為所求。（這是理所當然的：y 是 $3y^2-8y+3=0$ 的根，而這條方程正是 $y^2+1=4(y-1)^2$。）

想一想 這個等式對第 (i) 部分的不等式鏈說明了甚麼？

不等式鏈 $y^2+1\geq(y\cos\theta-\sin\theta)^2\geq 4(y-1)^2$ 的兩端現在相等，所以兩個不等式都取等號。

判別式爲零，所以關於 x 的二次方程有重根（repeated root）

$$x=-\frac{y\cos\theta-\sin\theta}{2(y-1)}.$$

關於 θ 的上界被達到，所以 $y\sin\theta+\cos\theta=0$，從而

$$\tan\theta=-\frac{1}{y}=-\frac{3}{4+\sqrt{7}}=-\frac{3(4-\sqrt{7})}{16-7}=\frac{\sqrt{7}-4}{3}.$$

（此處 $\cos\theta\neq 0$，否則 $\sin\theta=0$ 亦成立，這是不可能的。）此外 $(y\cos\theta-\sin\theta)^2=y^2+1=4(y-1)^2$，所以 $y\cos\theta-\sin\theta=\pm 2(y-1)$，因此

$$x=\mp 1.$$

具體來說，$(\cos\theta,\sin\theta)$ 與 $(y,-1)$ 平行。若兩者方向相同，即 $\cos\theta>0$ 且 $\sin\theta<0$，則 $y\cos\theta-\sin\theta=+2(y-1)$，$x=-1$。若兩者方向相反，即 $\cos\theta<0$ 且 $\sin\theta>0$，則 $x=1$。兩種情況下都有

$$x=\pm 1,\qquad \tan\theta=\frac{\sqrt{7}-4}{3}.$$

這兩種情況其實是同一回事：把 x 換成 $-x$、θ 換成 $\theta+\pi$，y 的值不變。

驗算

- 當 $x=1$ 時，公式為 $y=\dfrac{2+\sin\theta}{2+\cos\theta}$。取第二象限（quadrant）中滿足 $\tan\theta=\frac{1}{3}(\sqrt{7}-4)\approx -0.4514$ 的 θ，得 $\theta\approx 2.7176$，於是 $y\approx 2.21525$，與 $\frac{1}{3}(4+\sqrt{7})\approx 2.21525$ 吻合。
- 用電腦隨機抽取大量實數 x 和 θ，所得 y 值都落在 0.45142 與 2.21524 之間，剛好在界限 $\frac{1}{3}(4\mp\sqrt{7})\approx 0.451416$ 和 2.215250 之內。
- 一個值得留意的巧合：$\tan\theta=-\frac{1}{3}(4-\sqrt{7})$ 正是下界的相反數，因為兩個界互為倒數（reciprocal）（它們的積是 $\frac{1}{9}(16-7)=1$）。

常見錯誤

- 忘記 $y=1$ 的情況：此時關於 x 的方程不是二次方程，判別式的論證不能直接套用。
- 取 $\sqrt{y^2+1}=-2(y-1)$，或沒有檢查正負號：這裏取正根之所以正確，只是因為 $y>1$。
- 只給出一個 x 值；其實 $x=1$ 和 $x=-1$ 都會出現，對應的 θ 位於不同象限。

STEP 之外

步驟 $(y\cos\theta-\sin\theta)^2\leq(y^2+1)(\cos^2\theta+\sin^2\theta)$ 是柯西－施瓦茨不等式 $|\mathbf{u}\cdot\mathbf{v}|\leq|\mathbf{u}|\,|\mathbf{v}|$ 的二維情況，等號恰好在兩向量平行時成立。這個不等式在所有內積空間（inner-

product space）中都成立，是大學一年級分析（analysis）與線性代數（linear algebra）中最常用的工具之一。Strang 的 *Introduction to Linear Algebra*（線性代數導論）從幾何角度發展這個不等式；Abbott 的 *Understanding Analysis*（理解分析）則展示它在分析中的應用。

詳細解答

STEP II 2016，第 5 題

純數學 • 二項式係數恆等式

STEP II 2016，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/16-s2-q5

關鍵思路

三條恆等式（identity）都源自同一個技巧：把一個函數用兩種方式寫成乘積，然後比較 x 某一個冪的係數（coefficient）。若一個和式由二項式係數（binomial coefficient）的乘積組成，而各下標之和固定，那正正就是把兩個冪級數（power series）相乘後讀出某一項係數所得的結果。題目約定在 $p \geq q \geq 0$ 以外 $\binom{p}{q} = 0$，正因如此，和式可以在一個整齊的 j 範圍內求和，而不必費神理會哪些項存在。

解答

第 (i) 部分

當 $|x| < 1$ 時，由二項式級數（binomial series）可得 $(1-x)^{-N}$ 中 x^n 的係數為

$$\frac{(-N)(-N-1)\cdots(-N-n+1)}{n!}(-1)^n = \frac{N(N+1)\cdots(N+n-1)}{n!}$$
$$= \frac{(N+n-1)!}{n!\,(N-1)!} = \binom{N+n-1}{n}.$$

各因子帶來的 n 個負號，與 $(-x)^n$ 帶來的 n 個負號互相抵消。因此

$$(1-x)^{-N} = \sum_{n=0}^{\infty} \binom{N+n-1}{n} x^n \qquad (|x| < 1).$$

（當 $n = 0$ 時，係數為 $\binom{N-1}{0} = 1$。由於 $N \geq 1$，這沒有問題。）

現在 $(1-x)^{-1}(1-x)^{-N} = (1-x)^{-(N+1)}$。把上述展開式中的 N 換成 $N+1$，可知右邊 x^n 的係數為 $\binom{N+n}{n}$。

在左邊，$(1-x)^{-1} = \sum_{i\geq 0} x^i$ 的每一項係數都等於 1。$(1-x)^{-N}$ 中的 x^j 項必須與 $(1-x)^{-1}$ 中的 x^{n-j} 配對，而 j 由 0 取到 n。所以乘積中 x^n 的係數為

$$\sum_{j=0}^{n} 1 \cdot \binom{N+j-1}{j}.$$

兩個冪級數在 $|x|<1$ 時都收斂（converge），並在該處表示同一個函數，因此它們的係數相同。於是得到

$$\sum_{j=0}^{n} \binom{N+j-1}{j} = \binom{N+n}{n},$$

即為所求。

第 (ii) 部分

利用 $(1+x)^{m+n} = (1+x)^m(1+x)^n$，三者皆為多項式（polynomial）。按題目的約定，可以寫成

$$(1+x)^m = \sum_{j\geq 0} \binom{m}{j} x^j, \qquad (1+x)^n = \sum_{k\geq 0} \binom{n}{k} x^k,$$

因為 $j > m$ 或 $k > n$ 的係數均為零。乘積中 x^r 的係數來自滿足 $j+k=r$ 的配對，即 $k = r-j$ 且 $0 \leq j \leq r$：

$$\sum_{j=0}^{r} \binom{m}{j}\binom{n}{r-j}.$$

? **想一想** 那麼 $j > m$ 或 $r-j > n$ 的項又如何？這些冪並不出現在多項式之中。

這些項含有因子 $\binom{m}{j}$（其中 $j > m$）或 $\binom{n}{r-j}$（其中 $r-j > n$），按約定它們都是零。它們沒有任何貢獻，同時也不對應任何實際的項的乘積，所以在 $0 \leq j \leq r$ 上求和恰好正確。$(1+x)^{m+n}$ 中 x^r 的係數為 $\binom{m+n}{r}$。相等的多項式有相同的係數，因此

$$\binom{m+n}{r} = \sum_{j=0}^{r} \binom{m}{j}\binom{n}{r-j}.$$

第 (iii) 部分

因子 $(-1)^j\binom{m+j-1}{j}$ 提示我們用第 (i) 部分，並以 $-x$ 代替 x。在第 (i) 部分的展開式中以 $-x$ 代替 x，並以 m 代替 N，可得

$$(1+x)^{-m} = \sum_{j=0}^{\infty} (-1)^j \binom{m+j-1}{j} x^j \qquad (|x|<1).$$

? **想一想** 應該用哪一個乘積？

另一個因子 $\binom{N+m}{n-j}$ 是 $(1+x)^{N+m}$ 的係數，而答案 $\binom{N}{n}$ 是 $(1+x)^N$ 的係數。由於 m 是正整數，當 $|x|<1$ 時有

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

在左邊，$(1+x)^{N+m}=\sum_{k\geq 0}\binom{N+m}{k}x^k$（對 $k>N+m$ 使用上述約定）。x^n 的係數來自 $(1+x)^{-m}$ 中的 x^j 乘以 $(1+x)^{N+m}$ 中的 x^{n-j}，其中 $0\leq j\leq n$：

$$\sum_{j=0}^{n}\binom{N+m}{n-j}(-1)^j\binom{m+j-1}{j}.$$

在右邊，$(1+x)^N$ 中 x^n 的係數為 $\binom{N}{n}$。當 $n>N$ 時，按約定它等於 0，這也是正確的，因為 $(1+x)^N$ 的次數是 N。兩邊在 $|x|<1$ 上相等，所以它們的冪級數係數相同：

$$\sum_{j=0}^{n}(-1)^j\binom{N+m}{n-j}\binom{m+j-1}{j}=\binom{N}{n}.$$

驗算

- 第 (i) 部分，取 $N=2$、$n=3$：$1+2+3+4=10=\binom{5}{3}$。
- 第 (ii) 部分，取 $m=n=2$、$r=2$：$1\cdot 1+2\cdot 2+1\cdot 1=6=\binom{4}{2}$。
- 第 (iii) 部分，取 $m=2$、$N=3$、$n=2$：$\binom{5}{2}\binom{1}{0}-\binom{5}{1}\binom{2}{1}+\binom{5}{0}\binom{3}{2}=10-10+3=3=\binom{3}{2}$。當 $n>N$ 時，這個正負交替的和應該等於 0。例如取 $m=1$、$N=1$、$n=2$，得 $\binom{2}{2}-\binom{2}{1}+\binom{2}{0}=1-2+1=0$。

三條恆等式亦已用電腦對所有 $m,N\leq 9$ 及 $n\leq 24$ 驗證過。

常見錯誤

- 在 $(1-x)^{-N}$ 的係數中弄亂了正負號。那裏有兩組 $(-1)^n$，兩者互相抵消。
- 在第 (iii) 部分，忘記 $(1+x)^{-m}$ 的級數是無窮級數，只在 $|x|<1$ 時成立。比較係數之所以成立，是因為在某區間上相等的兩個冪級數，其係數相同。

STEP 之外

這就是*生成函數*（generating functions）的方法：把一個數列編碼為某冪級數的係數，於是級數相乘便化為這裏所見的卷積（convolution）和。第 (ii) 部分是范德蒙恆等式（Vandermonde's identity），而第 (i) 部分常被稱為曲棍球棒恆等式（hockey-stick identity）。

兩者都有漂亮的組合計數證明（counting proof），值得自己找找看。Wilf 的 *generating-functionology*（可在網上免費閱讀），以及 Graham、高德納（Knuth）與 Patashnik 合著的 *Concrete Mathematics*（《具體數學》），都把這套工具發展得更深入。

詳細解答

STEP II 2016，第 6 題

純數學 • 切比雪夫多項式方程

STEP II 2016，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/16-s2-q6

關鍵思路

每一部分都只需把一個候選函數代入微分方程(differential equation)，並檢查它在 $x = 1$ 處的條件；其餘的就交給唯一性（uniqueness）假設。第 (iii) 及 (iv) 部分的竅門是：永遠不要展開 y_n（我們並不知道它的明確表達式），而是直接利用方程 $(1 - x^2)y_n'^2 = n^2(1 - y_n^2)$ 本身來消去導數（derivative）。

解答

在整個解答中，y_k 滿足

$$(1 - x^2)\big(y_k'(x)\big)^2 = k^2\big(1 - y_k(x)^2\big), \qquad y_k(1) = 1,$$

這不過是 $(*)$ 移項後的形式。要證明某函數等於 y_k，(根據所假設的唯一性) 只需證明它在該 k 值下滿足 $(*)$，並且在 $x = 1$ 處取值 1。

第 (i) 部分

當 $k = 1$ 時，方程為

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

對於 $y = x$，我們有 $y' = 1$，而左方為 $(1 - x^2) + x^2 = 1$，符合要求。此外 $y(1) = 1$，所以 $y_1(x) = x$。

第 (ii) 部分

當 $k = 2$ 時，方程為

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

對於 $y = 2x^2 - 1$，我們有 $y' = 4x$，所以

$$\begin{aligned}(1 - x^2)(4x)^2 + 4(2x^2 - 1)^2 &= 16x^2 - 16x^4 + 16x^4 - 16x^2 + 4\\ &= 4.\end{aligned}$$

由於 $y(1) = 2 - 1 = 1$，我們得到 $y_2(x) = 2x^2 - 1$。

第 (iii) 部分

為簡便起見，記 $y = y_n(x)$。以鏈式法則 (chain rule) 對 $z = 2y^2 - 1$ 求導，得 $z' = 4yy'$，所以

$$(1 - x^2)z'^2 = 16y^2\,(1 - x^2)y'^2 = 16y^2 \cdot n^2(1 - y^2),$$

這裏用到了 y_n 所滿足的方程。此外

$$4n^2z^2 = 4n^2(2y^2 - 1)^2 = 4n^2(4y^4 - 4y^2 + 1).$$

兩式相加，

$$(1 - x^2)z'^2 + 4n^2z^2 = 16n^2y^2 - 16n^2y^4 + 16n^2y^4 - 16n^2y^2 + 4n^2 = 4n^2,$$

這正是所求的方程。

? **想一想** 這條方程與 y_{2n} 有甚麼關係？

由於 $(2n)^2 = 4n^2$，它恰好就是 $k = 2n$ 時的 $(*)$。而且 $z(1) = 2y_n(1)^2 - 1 = 2 - 1 = 1$。所以 z 是 $k = 2n$ 時 $(*)$ 的一個解，並且在 $x = 1$ 處取值 1；根據唯一性，

$$y_{2n}(x) = 2\big(y_n(x)\big)^2 - 1.$$

作為驗證，$n = 1$ 給出 $y_2 = 2x^2 - 1$，與 (ii) 一致；而 $n = 2$ 給出 $y_4 = 2(2x^2 - 1)^2 - 1 = 8x^4 - 8x^2 + 1$。

第 (iv) 部分

記 $u = y_m(x)$，於是 $v = y_n(u)$。根據鏈式法則，

$$v'(x) = y_n'(u)\,y_m'(x).$$

因此，先用 y_m 的方程（在點 x），再用 y_n 的方程（在點 u），得

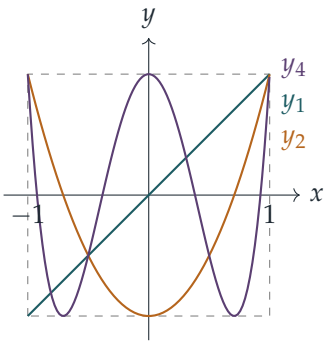
$$\begin{aligned}(1-x^2)v'^2 &= y_n'(u)^2\,(1-x^2)y_m'(x)^2\\ &= y_n'(u)^2\cdot m^2(1-u^2)\\ &= m^2\,(1-u^2)y_n'(u)^2\\ &= m^2\cdot n^2\big(1-y_n(u)^2\big) = m^2n^2(1-v^2).\end{aligned}$$

所以 $(1-x^2)v'^2+(mn)^2v^2=(mn)^2$，即 $k=mn$ 時的 $(*)$。最後

$$v(1)=y_n\big(y_m(1)\big)=y_n(1)=1,$$

所以根據唯一性，$v(x)=y_{mn}(x)$，得證。

y_n 的方程是在點 u 而非點 x 應用的：該方程對其變量的每一個值都成立，所以我們可以用 u 代替 x。正是這一步令 (iv) 得以成立。



三條曲線都通過 $(1,1)$，並且在 $[-1,1]$ 上都介乎 -1 與 1 之間，這是方程所迫使的：在該區間上 $(1-x^2)y'^2\geq 0$，所以 $k^2y^2\leq k^2$。

驗算

- 第 (iii) 部分取 $n=2$，預測 $y_4(x)=8x^4-8x^2+1$。直接代入，$y_4'=32x^3-16x$，而 $(1-x^2)(32x^3-16x)^2+16(8x^4-8x^2+1)^2$ 化簡後等於 16。在 $x=0.3$ 處，$y_4=0.3448$。
- 第 (iv) 部分取 $m=n=2$：$y_2(y_2(x))=2(2x^2-1)^2-1$，與 (iii) 得出的 y_4 相同，理應如此。
- 以電腦用精確算術驗證過：以這種方法構造的多項式對每個不超過 24 的 k 都滿足 $(*)$，而且對 $m,n\leq 4$ 有 $y_n(y_m(x))=y_{mn}(x)$。

常見錯誤

- 忘記 $x = 1$ 處的條件。單單滿足 $(*)$ 並不足夠：$-y_k$ 同樣滿足它。
- 在 (iii) 或 (iv) 中試圖明確展開 y_n。我們對 y_n 並沒有任何明確的認識；應利用它的微分方程來取代 $(1-x^2)y_n'^2$。
- 在 (iv) 中，把 y_n 的方程用在 x 上，而不是 $u = y_m(x)$ 上，結果令因子 $(1-x^2)$ 出現在錯誤的位置。

STEP 之外

函數 y_k 就是切比雪夫多項式（Chebyshev polynomials）T_k：設 $x = \cos\theta$，$(*)$ 便化為 $(\mathrm{d}y/\mathrm{d}\theta)^2 = k^2(1-y^2)$，其解為 $y = \cos k\theta$，因此 $T_k(\cos\theta) = \cos k\theta$。於是第 (iii) 部分就是二倍角公式，而第 (iv) 部分說的是 $\cos\big(n(m\theta)\big) = \cos(mn\theta)$。不過要提醒一句：題目容許你假設的唯一性其實並不完全成立，因為常數函數 $y = 1$ 同樣滿足 $(*)$ 及 $y(1) = 1$。像這樣導數只能確定至正負號的方程，正是使用唯一性定理時須格外小心的地方；Boyce 與 DiPrima 合著的 *Elementary Differential Equations*（《初等微分方程》）討論了解的存在性與唯一性。

詳細解答

STEP II 2016，第 7 題

純數學 • 利用對稱性求定積分

> **STEP II 2016，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/16-s2-q7

關鍵思路

把區間 $[0, a]$ 關於其中點作反射（reflection），即 $x \mapsto a - x$，積分的值不會改變，但被積函數（integrand）卻可能變成一個能與原式漂亮地結合的形式。對於 $[0, \frac{1}{2}\pi]$ 上的三角被積函數，這個反射會把 sin 與 cos 互換；在 $[0, \frac{1}{4}\pi]$ 上，則靠 $\tan(\frac{1}{4}\pi - x)$ 的複角公式（compound-angle formula）來完成工作。最後一部分看似全新，但一次分部積分（integration by parts）就能把它化為第 (iii) 部分。

解答

結果 $(*)$

在 $\int_0^a f(a-x)\,\mathrm{d}x$ 中作代換（substitution）$u = a - x$，於是 $\mathrm{d}u = -\mathrm{d}x$；當 $x = 0$ 時 $u = a$，當 $x = a$ 時 $u = 0$。因此

$$\int_0^a f(a-x)\,\mathrm{d}x = \int_a^0 f(u)\,(-\mathrm{d}u) = \int_0^a f(u)\,\mathrm{d}u = \int_0^a f(x)\,\mathrm{d}x,$$

因為積分變量用甚麼名稱並不重要。這就是 $(*)$。

第 (i) 部分

設該積分為 I。取 $a = \frac{1}{2}\pi$，有 $\sin(\frac{1}{2}\pi - x) = \cos x$ 及 $\cos(\frac{1}{2}\pi - x) = \sin x$，所以由 $(*)$ 得

$$I = \int_0^{\frac{1}{2}\pi} \frac{\cos x}{\sin x + \cos x}\,\mathrm{d}x.$$

把它與 I 的原來形式相加，

$$2I = \int_0^{\frac{1}{2}\pi} \frac{\sin x + \cos x}{\cos x + \sin x}\,\mathrm{d}x = \int_0^{\frac{1}{2}\pi} 1\,\mathrm{d}x = \frac{\pi}{2},$$

所以

$$I = \frac{\pi}{4}.$$

（分母 $\cos x + \sin x$ 在 $[0, \frac{1}{2}\pi]$ 上為正，因此不會出問題。）

第 (ii) 部分

現在 $a = \frac{1}{4}\pi$，第 (i) 部分的技巧不再能互換 sin 與 cos。改為把分子分母同除以 $\cos x$（在此區間上為正），被積函數變成

$$\frac{\sin x}{\cos x + \sin x} = \frac{\tan x}{1 + \tan x}.$$

設 $t = \tan x$。由於 $\tan(\frac{1}{4}\pi - x) = \dfrac{1-t}{1+t}$，

$$\frac{\tan(\frac{1}{4}\pi - x)}{1 + \tan(\frac{1}{4}\pi - x)} = \frac{(1-t)/(1+t)}{2/(1+t)} = \frac{1-t}{2} = \frac{1}{2} - \frac{1}{2}\tan x.$$

於是設該積分為 J，由 $(*)$ 得

$$\begin{aligned} J &= \int_0^{\frac{1}{4}\pi} \Big(\frac{1}{2} - \frac{1}{2}\tan x\Big)\mathrm{d}x = \frac{\pi}{8} - \frac{1}{2}\Big[-\ln\cos x\Big]_0^{\frac{1}{4}\pi} \\ &= \frac{\pi}{8} + \frac{1}{2}\ln\frac{1}{\sqrt{2}} = \frac{\pi}{8} - \frac{1}{4}\ln 2. \end{aligned}$$

想一想 有沒有方法完全不用 $(*)$ 就能做第 (ii) 部分？

有。設 $K = \int_0^{\frac{1}{4}\pi} \frac{\cos x}{\cos x + \sin x}\,\mathrm{d}x$。那麼 $J + K = \frac{1}{4}\pi$；又因為分子 $\cos x - \sin x$ 正是分母的導數，

$$K - J = \Big[\ln(\cos x + \sin x)\Big]_0^{\frac{1}{4}\pi} = \ln\sqrt{2} = \tfrac{1}{2}\ln 2.$$

兩式相減，得 $2J = \frac{1}{4}\pi - \frac{1}{2}\ln 2$，與上面的答案一致。

第 (iii) 部分

設該積分為 L。同一條複角公式給出

$$1 + \tan(\tfrac{1}{4}\pi - x) = 1 + \frac{1 - \tan x}{1 + \tan x} = \frac{2}{1 + \tan x}.$$

所以由 $(*)$，

$$L = \int_0^{\frac{1}{4}\pi} \ln \frac{2}{1+\tan x}\,\mathrm{d}x = \int_0^{\frac{1}{4}\pi} \big(\ln 2 - \ln(1+\tan x)\big)\,\mathrm{d}x = \frac{\pi}{4}\ln 2 - L.$$

因此 $2L = \frac{1}{4}\pi \ln 2$，而

$$L = \frac{\pi}{8}\ln 2.$$

第 (iv) 部分

把分子分母同除以 $\cos^2 x$：

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

? **想一想** $\dfrac{\sec^2 x}{1+\tan x}$ 跟第 (iii) 部分有甚麼關係？

它恰好是 $\ln(1+\tan x)$ 的導數，因為 $\frac{\mathrm{d}}{\mathrm{d}x}(1+\tan x) = \sec^2 x$；而且在 $[0, \frac{1}{4}\pi]$ 上 $1+\tan x > 0$。於是取 $u = x$ 及 $\mathrm{d}v = \dfrac{\sec^2 x}{1+\tan x}\,\mathrm{d}x$、$v = \ln(1+\tan x)$ 作分部積分：

$$\begin{aligned}\int_0^{\frac{1}{4}\pi} \frac{x\sec^2 x}{1+\tan x}\,\mathrm{d}x &= \Big[x\ln(1+\tan x)\Big]_0^{\frac{1}{4}\pi} - \int_0^{\frac{1}{4}\pi} \ln(1+\tan x)\,\mathrm{d}x \\ &= \frac{\pi}{4}\ln 2 - \frac{\pi}{8}\ln 2 \qquad \text{（利用第 (iii) 部分）} \\ &= \frac{\pi}{8}\ln 2.\end{aligned}$$

所以第 (iv) 部分的積分與第 (iii) 部分的積分數值相同。

驗算

- 第 (i) 部分：被積函數由 0 升至 1，並關於點 $(\frac{1}{4}\pi, \frac{1}{2})$ 對稱，所以其平均值為 $\frac{1}{2}$，積分等於區間長度的一半，即 $\frac{1}{4}\pi$。
- 第 (ii) 部分應小於第 (i) 部分的一半，因為被積函數在 $[0, \frac{1}{4}\pi]$ 上低於 $\frac{1}{2}$：的確 $\frac{1}{8}\pi - \frac{1}{4}\ln 2 \approx 0.2194 < 0.3927$。
- 數值積分（numerical quadrature）（辛普森法則（Simpson's rule））給出四個部分的值分別為 0.785398、0.219412、0.272198 及 0.272198，與 $\frac{1}{4}\pi$、$\frac{1}{8}\pi - \frac{1}{4}\ln 2$ 及 $\frac{1}{8}\pi \ln 2$（兩次）吻合。

常見錯誤

- 忘記代換 $u = a - x$ 會把積分上下限對調；而 $\mathrm{d}u = -\mathrm{d}x$ 中的負號會把它們調回來。

- 在第 (ii) 部分仿照第 (i) 部分試圖互換 sin 與 cos：在 $[0, \frac{1}{4}\pi]$ 上，反射是 $x \mapsto \frac{1}{4}\pi - x$，而不是 $\frac{1}{2}\pi - x$。
- 寫成 $\int \tan x \, \mathrm{d}x = \ln \cos x$；正確應為 $-\ln \cos x$（或 $\ln \sec x$）。

STEP 之外

第 (iii) 部分是經典題：同一個反射技巧可以求出像 $\int_0^{\frac{1}{2}\pi} \ln \sin x \, \mathrm{d}x = -\frac{1}{2}\pi \ln 2$ 這類積分，而這些積分無法透過初等函數（elementary functions）的原函數（antiderivative）求得。到了大學，這類對稱性論證與圍道積分（contour integration）及積分號下求導（differentiation under the integral sign）並列，都是處理沒有初等原函數的定積分（definite integral）的工具。想看更多富 STEP 風格的例子，可參閱 Siklos 的 *Advanced Problems in Mathematics*（《數學進階難題》）；至於為代換法與分部積分提供嚴格依據的分析（analysis），可參閱 Abbott 的 *Understanding Analysis*（《理解分析》）。

詳細解答

STEP II 2016，第 8 題

純數學 • 以積分近似求和

> **STEP II 2016，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/16-s2-q8

關鍵思路

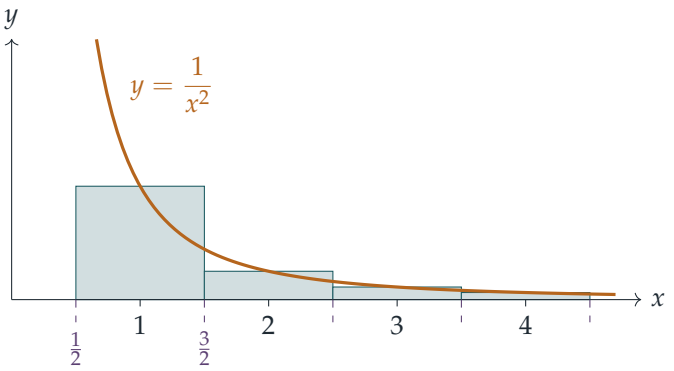
項 $\frac{1}{r^2}$ 是一個寬 1、高 $\frac{1}{r^2}$ 的長方形的面積。若這個長方形以 $x = r$ 為中心，它的面積就非常接近 $y = \frac{1}{x^2}$ 在 $r - \frac{1}{2}$ 與 $r + \frac{1}{2}$ 之間曲線下的面積，這就是中點法則（midpoint rule）。把這些窄條拼接起來，一個和式就變成一個積分。接着可以準確算出每一窄條的誤差（error），而把這些誤差加起來，幾乎不費吹灰之力就得到第二個級數（series）$\sum \frac{1}{r^4}$。

解答

積分與草圖

當 $m > \frac{1}{2}$ 時，積分下限為正，所以被積函數（integrand）在積分範圍內連續，並且

$$\int_{m-\frac{1}{2}}^{\infty} \frac{1}{x^2}\,\mathrm{d}x = \lim_{X\to\infty}\left[-\frac{1}{x}\right]_{m-\frac{1}{2}}^{X} = \frac{1}{m-\frac{1}{2}} = \frac{2}{2m-1}.$$



以 $x = r$ 為中心的長方形寬 1、高 $\frac{1}{r^2}$，所以面積是 $\frac{1}{r^2}$。在窄條 $r - \frac{1}{2} \le x \le r + \frac{1}{2}$ 上，曲線在左半邊位於長方形頂邊之上，在右半邊則位於其下，而這兩小塊幾乎互相抵消。因此

$$\frac{1}{r^2} \approx \int_{r-\frac{1}{2}}^{r+\frac{1}{2}} \frac{1}{x^2}\,\mathrm{d}x.$$

對 $r = m, m+1, \ldots, n$ 把這些式子相加，各窄條首尾相接，剛好覆蓋 $[m - \frac{1}{2}, n + \frac{1}{2}]$，於是得到 $(*)$。

第 (i) 部分

在 $(*)$ 中令 $n \to \infty$。右邊趨向我們剛才算出的積分，所以

$$\sum_{r=m}^{\infty} \frac{1}{r^2} \approx \frac{2}{2m-1}.$$

我們可以把這個近似用於整個級數，也可以先準確地加上頭幾項，只對尾部（tail）作近似：

$$\begin{aligned}
m = 1: &\quad E \approx \frac{2}{1} = 2,\\
m = 2: &\quad E = 1 + \sum_{r=2}^{\infty} \frac{1}{r^2} \approx 1 + \frac{2}{3} = \frac{5}{3},\\
m = 3: &\quad E = 1 + \frac{1}{4} + \sum_{r=3}^{\infty} \frac{1}{r^2} \approx \frac{5}{4} + \frac{2}{5} = \frac{33}{20}.
\end{aligned}$$

這就是題目要求的三個近似值。隨着 m 增大，近似應該愈來愈好，因為曲線在愈遠處愈平坦，中點近似在那裏也愈準確。

第 (ii) 部分

在一個窄條上的積分是

$$\int_{r-\frac{1}{2}}^{r+\frac{1}{2}} \frac{1}{x^2}\,\mathrm{d}x = \frac{1}{r-\frac{1}{2}} - \frac{1}{r+\frac{1}{2}} = \frac{1}{r^2 - \frac{1}{4}} = \frac{4}{4r^2-1}.$$

所以用這個積分近似 $\frac{1}{r^2}$ 的誤差是

$$\frac{4}{4r^2-1}-\frac{1}{r^2}=\frac{4r^2-(4r^2-1)}{r^2(4r^2-1)}=\frac{1}{r^2(4r^2-1)}=\frac{1}{4r^4}\cdot\frac{1}{1-\frac{1}{4r^2}}.$$

當 r 很大時，$\frac{1}{4r^2}$ 很小，所以最後一個因子接近 1，誤差約為 $\frac{1}{4r^4}$，正如題目所要求。（積分值總是較大，這與圖像吻合：$\frac{1}{x^2}$ 是凸函數（convex），所以每個窄條左邊多出的面積勝過右邊不足的面積。）

至於最後的結果，把由 $r=m$ 起的誤差全部加起來。各窄條的積分加起來正是我們最先算出的積分，所以

$$\frac{2}{2m-1}-\sum_{r=m}^{\infty}\frac{1}{r^2}=\sum_{r=m}^{\infty}\frac{1}{r^2(4r^2-1)}\approx\frac{1}{4}\sum_{r=m}^{\infty}\frac{1}{r^4}.$$

想一想 為甚麼不取看來最簡單的 $m=1$？

取 $m=1$ 時，左邊是 $2-1.645=0.355$，得出 $\sum\frac{1}{r^4}\approx 1.42$，這大得太多了。問題在於「$\approx\frac{1}{4r^4}$」只在 r 很大時才準確：在 $r=1$ 時，真正的誤差是 $\frac{1}{3}$ 而非 $\frac{1}{4}$，而這第一項主導了整個和。所以我們應該把頭幾項準確保留，只對尾部使用近似。取 $m=3$（即得出 $\frac{33}{20}$ 的那個選擇），最差的因子 $\frac{1}{1-1/(4r^2)}$ 是 $\frac{36}{35}$，與 1 相差不足 3%：

$$\frac{1}{4}\sum_{r=3}^{\infty}\frac{1}{r^4}\approx\frac{2}{5}-\left(E-1-\frac{1}{4}\right)\approx 0.4-0.395=0.005,$$

所以 $\sum_{r=3}^{\infty}\frac{1}{r^4}\approx 0.02$，而

$$\sum_{r=1}^{\infty}\frac{1}{r^4}\approx 1+\frac{1}{16}+0.02=1.0825\approx 1.08.$$

這就是所要求的結果。由於 E 只給到三位小數，E 中最多 0.0005 的誤差在這裏會變成最多 0.002，結果仍然是 1.08。

驗算

- 近似值 2、$\frac{5}{3}\approx 1.667$、$\frac{33}{20}=1.65$ 逐步遞減，趨向 $E=\frac{\pi^2}{6}\approx 1.6449$，而且每一個都偏大，正如凸性論證所預測。
- 在 $r=10$ 時，準確誤差 $\frac{1}{100\times 399}\approx 2.506\times 10^{-5}$ 非常接近 $\frac{1}{4\times 10^4}=2.5\times 10^{-5}$。
- 準確值是 $\sum\frac{1}{r^4}=\frac{\pi^4}{90}\approx 1.0823$，與我們的 1.0825 吻合。若改用 $m=2$，得 $1+4\left(\frac{2}{3}-0.645\right)\approx 1.087$，明顯較差；$m=1$ 則得 1.42。

常見錯誤

- 在最後一部分使用 $m = 1$，而大 r 近似 $\frac{1}{4r^4}$ 在那裏並不準確，結果得出 1.42。
- 把誤差的正負號弄反：積分大於 $\frac{1}{r^2}$，所以 $\frac{2}{2m-1}$ 是尾部的高估值。
- 忘記把準確計算的頭幾項（$1 + \frac{1}{4}$ 及 $1 + \frac{1}{16}$）加回去。

STEP 之外

這就是中點法則，而誤差項 $\frac{1}{4r^4}$ 讓我們窺見歐拉–麥克勞林公式（Euler–Maclaurin formula）：它以一個積分加上一連串導數項來修正一個和式，而像 $\sum \frac{1}{r^2}$ 這樣的和，正是這樣計算到許多位小數的。準確值 $\frac{\pi^2}{6}$ 和 $\frac{\pi^4}{90}$ 來自傅里葉級數（Fourier series），或來自歐拉（Euler）對 $\sin x$ 的乘積展開。Abbott 的 *Understanding Analysis*（《理解分析》）仔細地處理級數與積分，而 Siklos 的 *Advanced Problems in Mathematics*（《高等數學問題》）則有更多同類精神的題目。

詳細解答

STEP II 2016，第 9 題

力學 • 動量與能量損失

STEP II 2016，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：

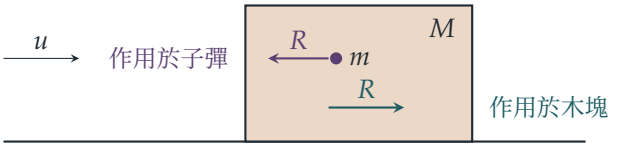
step.maths.org/questions/16-s2-q9

關鍵思路

阻力 R 是子彈與木塊之間的內力（internal force）：它令子彈減速，而根據牛頓第三定律（Newton's third law），它同時以相同大小把木塊向前推。因此兩者組成的系統動量守恆（conservation of momentum），但動能（kinetic energy）並不守恆。損失的能量恰好等於 R 沿子彈在木塊內移動的相對距離所做的功。對木塊單獨應用功能原理（work–energy principle），便可求出木塊移動的距離。

解答

我們把子彈和木塊視為沿同一條水平直線運動的質點，並假設子彈相對木塊運動期間阻力 R 保持不變。取子彈的初始運動方向為正方向。



鉛垂方向的力（重力及桌面的法向反作用力）互相平衡，不起任何作用。在水平方向上，子彈受向後的力 R，木塊則受向前的力 R。

第 (i) 部分

木塊固定時，子彈的減速度恆為 R/m。利用 $v^2 = u^2 + 2\alpha s$，代入末速 0 及 $\alpha = -R/m$：

$$0 = u^2 - \frac{2R}{m}a \quad \Longrightarrow \quad \boxed{a = \frac{mu^2}{2R}.}$$

（等價地說，克服阻力所做的功 Ra 等於子彈的初始動能 $\frac{1}{2}mu^2$。）

第 (ii) 部分

共同末速度。由於桌面光滑，子彈和木塊組成的系統不受任何水平外力。當子彈相對木塊靜止時，兩者以相同速度 V 運動，由動量守恆得

$$mu = (M+m)V \quad \Longrightarrow \quad V = \frac{mu}{M+m}.$$

> **想一想** 阻力在子彈於木塊內的路徑上作用了距離 b，但子彈本身相對桌面移動了 $b+c$。能量方程中應該用哪一個距離？

求 b。相對桌面而言，作用於子彈的力 R 所做的功為 $-R(b+c)$，因為子彈移動了 $b+c$；作用於木塊的力 R 所做的功為 $+Rc$，因為木塊移動了 c。因此對整個系統所做的淨功為 $-Rb$：只有相對位移才算數。由功能原理，

$$\begin{aligned} Rb &= \tfrac{1}{2}mu^2 - \tfrac{1}{2}(M+m)V^2 \\ &= \tfrac{1}{2}mu^2 - \frac{m^2u^2}{2(M+m)} \\ &= \tfrac{1}{2}mu^2\left(1 - \frac{m}{M+m}\right) = \tfrac{1}{2}mu^2 \cdot \frac{M}{M+m}. \end{aligned}$$

由於第 (i) 部分得 $\frac{1}{2}mu^2 = Ra$，兩邊除以 R 得

$$\boxed{b = \frac{M}{M+m}\,a.}$$

求 c。現在對木塊單獨應用功能原理。木塊由靜止開始，最終速率為 V，而它所受的唯一水平力是向前的 R，作用距離為 c：

$$Rc = \tfrac{1}{2}MV^2 = \frac{Mm^2u^2}{2(M+m)^2} = \tfrac{1}{2}mu^2 \cdot \frac{mM}{(M+m)^2}.$$

再次利用 $\frac{1}{2}mu^2 = Ra$，

$$\boxed{c = \frac{mM}{(M+m)^2}\,a.}$$

另一方法：運動學。子彈以 R/m 減速，木塊以 R/M 加速。當 $u - Rt/m = Rt/M$ 時兩者速度相等，所以

$$t = \frac{mMu}{R(M+m)}, \qquad c = \frac{1}{2}\cdot\frac{R}{M}\,t^2 = \frac{m^2Mu^2}{2R(M+m)^2},$$

與上面的結果一致；而 b 則等於子彈的位移減去 c。

驗算

- 當 $M \to \infty$ 時，木塊的表現猶如固定不動：$b \to a$ 且 $c \to 0$，正合所料。
- 由於 $\frac{M}{M+m} < 1$，故 $b < a$：可移動的木塊把較少的子彈能量轉化為熱，所以子彈穿入得沒有那麼深。
- 取 $M = 4m$，得 $b = \frac{4}{5}a$ 及 $c = \frac{4}{25}a$。以細小的時間步長逐步數值模擬運動，並隨機選取多組 m、M、u 和 R，所得的 a、b 和 c 均準確至四位有效數字。

常見錯誤

- 在能量損失方程中用了子彈相對桌面的距離 $b + c$，而非相對距離 b。
- 假設碰撞中動能守恆；事實並非如此，損失的 Rb 就是所產生的熱。
- 忘記木塊同樣受到力 R 的作用（牛頓第三定律），以致無法解釋木塊的運動。

STEP 之外

內力只有在沒有相對運動時才做零淨功，這個事實正是嚴謹處理質點系能量的起點。在質心參考系（centre-of-mass frame）中處理，問題就變得簡潔：可供耗散的動能為 $\frac{1}{2}\mu u^2$，其中約化質量（reduced mass）$\mu = mM/(M+m)$，而這恰好就是本題中的 Rb。Kleppner 與 Kolenkow 的 *An Introduction to Mechanics*（《力學導論》）對這些概念有詳盡的闡述。

詳細解答

STEP II 2016，第 10 題

力學 • 摩擦力與平衡

> **STEP II 2016，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/16-s2-q10

關鍵思路

作用在三角形上的只有兩個力：作用於質心（centre of mass）G 的重量，以及釘在接觸點 P 給予的單一接觸力（contact force）。兩個力處於平衡（equilibrium）時，必須大小相等、方向相反，並且作用在同一直線上，所以三角形會擺動，直至 G 位於 P 的正下方。於是問題變成：釘的反作用力要偏離 BC 的法線多少，而摩擦力（friction）是否足以承受？因此第一步是找出 G 的位置。

解答

模型與質心

把金屬線視為三根單位長度質量相同的均勻桿（uniform rod），並把釘視為單一接觸點，提供一個法向反作用力（normal reaction）R（垂直於 BC）和一個摩擦力 F（沿 BC），且 $F \le \mu R$。

設 $AB = AC = b$，並設 M 為 BC 的中點。由於 A 處的角是 2θ，直線 AM 平分此角並垂直於 BC，所以

$$BM = MC = b\sin\theta, \qquad AM = b\cos\theta, \qquad BC = 2b\sin\theta.$$

由對稱性，G 位於 AM 上。從 BC 向 A 的方向量度距離。桿 AB 和 AC 的質量各與 b 成正比，其中心位於各自的中點，與 BC 的距離為 $\frac{1}{2}b\cos\theta$；桿 BC 的質量與 $2b\sin\theta$ 成正比，其中心位於 BC 上。所以距離 $d = GM$ 為

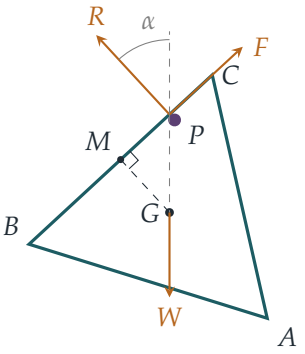
$$d = \frac{b\cdot\frac{1}{2}b\cos\theta + b\cdot\frac{1}{2}b\cos\theta + 2b\sin\theta\cdot 0}{b+b+2b\sin\theta} = \frac{b\cos\theta}{2(1+\sin\theta)}.$$

釘位於一般位置時的平衡

假設釘在 P 點接觸 BC，其中 $MP = x$，且 $0 \le x \le b\sin\theta$（由對稱性，可取 P 位於 M 與 C 之間）。三角形懸掛在釘的下方，所以釘位於三角形內部，而 BC 擱在釘上。

想一想 三角形可以以任何角度懸掛。我們怎樣知道它會取哪一個角度？

作用在三角形上的力是作用於 G 的重量 W，以及釘在 P 的總接觸力。只有兩個力時，平衡要求它們大小相等、方向相反（所以接觸力是鉛直向上的 W），並且作用在同一直線上（否則它們會構成一個力偶（couple））。所以 PG 必須是鉛直的：三角形會穩定在 G 位於 P 正下方的位置。



設 α 為 PG（鉛直線）與 BC 的法線 GM 之間的夾角。在直角三角形 PMG 中，

$$\tan\alpha = \frac{MP}{GM} = \frac{x}{d}.$$

BC 的法線與鉛直線成 α 角，而 BC 與水平線成 α 角。沿垂直於 BC 及沿 BC 的方向分解（resolve）：

$$R = W\cos\alpha, \qquad F = W\sin\alpha.$$

由於 $0 \le \alpha < \frac{1}{2}\pi$，兩者皆非負，且 $R > 0$。只要 $F \le \mu R$，三角形便能保持不動而不滑動，即

$$\mu \ge \tan\alpha = \frac{x}{d}.$$

BC 上的每一點

右方的 x/d 隨 x 遞增，所以當 P 位於端點，即 $x = b\sin\theta$ 時，條件最難滿足。當且僅當在該處可以平衡時，在 BC 上每一點都可以平衡：

$$\mu \ge \frac{b\sin\theta}{d} = b\sin\theta \cdot \frac{2(1+\sin\theta)}{b\cos\theta} = 2\tan\theta\,(1+\sin\theta),$$

這就是所求的條件。若此條件成立，則對 $[0, b\sin\theta]$ 中每一個 x 都有 $\mu \ge x/d$，所以無論釘位於 BC 上哪一點，三角形都可以靜止。

驗算

- 當 $x = 0$（釘位於 M）時，得 $\alpha = 0$：三角形懸掛時 BC 保持水平，完全不需要摩擦力，正如對稱性所預期。
- 當 $\theta = 30^\circ$ 時，三角形是等邊三角形，$d = \frac{1}{3}b\cos 30^\circ$，而 G 應為形心（centroid），即從 BC 沿高 $b\cos 30^\circ$ 向上三分之一處。結果確實如此。條件變成 $\mu \geq 2 \cdot \frac{1}{\sqrt{3}} \cdot \frac{3}{2} = \sqrt{3} \approx 1.732$：在角位時，$BC$ 須與水平線成 60°。
- 當 $\theta \to 0$ 時，三角形變成一根細長的尖刺，BC 縮小至零，界限趨於 0；當 $\theta \to \frac{1}{2}\pi$ 時，三角形變成一條又長又扁的窄條，界限趨於無窮大。兩者都符合直覺。
- 以電腦驗算：把三根桿分割成許多小段來求 G，並對隨機的 θ 和 b 在 BC 上 201 個點分解力，結果與 $2\tan\theta(1+\sin\theta)$ 的差在 10^{-11} 以內。

常見錯誤

- 用了三角形薄板（lamina）的形心（AM 向上三分之一處），而不是金屬線的質心。金屬線框的質量分佈在三邊上，所以 BC 的權重是 $2b\sin\theta$。
- 假設 BC 保持水平，然後嘗試取力矩。三角形會傾斜，直至 G 位於釘的下方；摩擦力要應付的正是這個傾斜角。
- 只檢查 BC 上的一點。最壞的情況在 BC 的端點，而正是這一點決定了界限。

STEP 之外

寫出 $\tan\alpha \leq \mu$，即表示總接觸力必須位於摩擦錐（cone of friction）之內，這是一個以法線為軸、半頂角為 $\arctan\mu$ 的圓錐。這個圖像令許多摩擦問題（梯子、楔子、粗糙的釘）幾乎變成幾何問題，並且可以推廣至有多個接觸點的剛體（rigid body）：問題變成能否在每個錐內選取某些力，使之與負載平衡。Kleppner 與 Kolenkow 的 *An Introduction to Mechanics*（《力學導論》）便以這種嚴謹的態度處理摩擦與剛體平衡。

詳細解答

STEP II 2016，第 11 題

力學 • 拋體的碰撞

> **STEP II 2016，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/16-s2-q11

關鍵思路

兩個質點碰撞（collide），即是在同一時刻兩個坐標都相等，因此我們得到兩條共用同一個未知數 t 的方程；把其中一條除以另一條，便可消去 t。第 (ii) 部分的竅門在於留意到：在每一瞬間，重力把子彈和靶子往下拉的距離完全相同，都是 $\frac{1}{2}gt^2$。所以兩者能否相遇，取決於第 (i) 部分那種沒有重力的運動；重力只在問「相遇時靶子在哪裏」時才有關係。

解答

第 (i) 部分

把第一個質點記作 P（從 $(a,0)$ 出發），第二個記作 Q（從 $(0,b)$ 出發）。設兩者在時刻 t 碰撞。令 x 坐標和 y 坐標分別相等，

$$\begin{aligned} a + ut\cos\alpha &= vt\cos\beta, & \text{所以} \quad t(v\cos\beta - u\cos\alpha) &= a, \\ ut\sin\alpha &= b + vt\sin\beta, & \text{所以} \quad t(u\sin\alpha - v\sin\beta) &= b. \end{aligned}$$

由於 $a > 0$，第一條方程顯示 $t \neq 0$（兩個質點從不同的點出發），因此 $t > 0$，我們可以把第二條方程除以第一條：

$$\frac{u\sin\alpha - v\sin\beta}{v\cos\beta - u\cos\alpha} = \frac{b}{a} = \tan\theta = \frac{\sin\theta}{\cos\theta}.$$

（分母等於 $a/t \neq 0$。）交叉相乘，

$$\cos\theta\,(u\sin\alpha - v\sin\beta) = \sin\theta\,(v\cos\beta - u\cos\alpha),$$

再把含 u 的項移到左邊、含 v 的項移到右邊，

$$u(\sin\alpha\cos\theta+\cos\alpha\sin\theta)=v(\sin\beta\cos\theta+\cos\beta\sin\theta).$$

由 $\sin(A+B)$ 的複角公式（compound-angle formula），上式即是

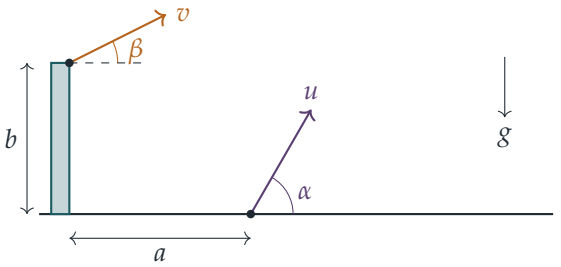
$$u\sin(\theta+\alpha)=v\sin(\theta+\beta),$$

此即所求。我們另外記下兩個稍後會用到的事實：由於 $a,b,t>0$，

$$v\cos\beta>u\cos\alpha \qquad 及 \qquad u\sin\alpha>v\sin\beta. \tag{$*$}$$

第 (ii) 部分

把子彈和靶子視為只受重力作用、自由運動的質點，不計空氣阻力（air resistance）。取塔底為原點，x 軸為通過靶子發射點的水平線（方向為遠離塔的一方），y 軸鉛直向上。子彈必須朝靶子的方向發射才有可能命中，所以兩者都在這個鉛直平面內運動。



在時刻 t（兩者落地之前），靶子和子彈的位置分別是

$$T=\left(a+ut\cos\alpha,\ ut\sin\alpha-\tfrac{1}{2}gt^2\right), \qquad B=\left(vt\cos\beta,\ b+vt\sin\beta-\tfrac{1}{2}gt^2\right).$$

這恰好是第 (i) 部分中的位置，各自向下平移了相同的距離 $\frac{1}{2}gt^2$。

? **想一想** 第 (i) 部分沒有重力，為甚麼我們仍可以用第 (i) 部分的結果？

命中發生在 $T=B$ 的時候。在 y 方程中，$-\frac{1}{2}gt^2$ 這一項同時出現在兩邊，可以互相抵消，所以碰撞條件與第 (i) 部分一字不差：

$$t\,(v\cos\beta-u\cos\alpha)=a, \qquad t\,(u\sin\alpha-v\sin\beta)=b.$$

特別地，第 (i) 部分及 $(*)$ 都適用，而命中的時刻是

$$t=\frac{b}{u\sin\alpha-v\sin\beta},$$

其中由 $(*)$ 知分母為正。

此刻靶子仍在空中，當且僅當它的高度為正：

$$ut\sin\alpha - \tfrac{1}{2}gt^2 > 0 \iff t < \frac{2u\sin\alpha}{g},$$

這裏除以了 $\frac{1}{2}t > 0$。代入命中的時刻，

$$\frac{b}{u\sin\alpha - v\sin\beta} < \frac{2u\sin\alpha}{g}.$$

兩個分母都是正數，所以兩邊同乘 $g(u\sin\alpha - v\sin\beta)$ 不會改變不等式（inequality）的方向，於是得到

$$2u\sin\alpha\,(u\sin\alpha - v\sin\beta) > bg,$$

此即所求。（命中時子彈與靶子處於同一高度，所以子彈也在地面之上。）

為甚麼需要 $\alpha > \beta$

我們已經看到，要命中，兩個質點必須按第 (i) 部分的意義碰撞，所以

$$u\sin(\theta+\alpha) = v\sin(\theta+\beta),$$

其中 $\tan\theta = b/a$；而且由 $(*)$，還有 $u\sin\alpha > v\sin\beta$。這裏四個正弦值都是正數，因為 θ、α、β 都是銳角，所以 $\theta+\alpha$ 和 $\theta+\beta$ 嚴格介乎 0 與 π 之間。把等式除以這個不等式，不等號方向會倒轉：

$$\frac{\sin(\theta+\alpha)}{\sin\alpha} < \frac{\sin(\theta+\beta)}{\sin\beta}.$$

展開得 $\dfrac{\sin(\theta+\alpha)}{\sin\alpha} = \sin\theta\cot\alpha + \cos\theta$，$\beta$ 的情況同理，所以

$$\sin\theta\cot\alpha < \sin\theta\cot\beta \iff \cot\alpha < \cot\beta,$$

因為 $\sin\theta > 0$。在 $(0, \frac{\pi}{2})$ 上餘切函數（cotangent）嚴格遞減（strictly decreasing），所以 $\alpha > \beta$。

從物理角度看：相對於子彈，靶子以恆定速度沿直線運動，因為重力已經抵消了。這個相對速度（relative velocity）必須由靶子的起點 $(a, 0)$ 指向位於 $(0, b)$ 的槍，所以它必須指向後方及上方；由 $(*)$，靶子在水平方向必須比子彈慢，在鉛直方向卻比子彈快，這迫使它的路徑較陡，即 $\tan\alpha > \tan\beta$。（把 $(*)$ 的兩部分相除，也可以直接得到同一結論。）

驗算

試取 $u = 20$、$\alpha = 60°$、$v = 15$、$\beta = 30°$，並設在 $t = 1$ 時命中，$g = 9.8$。碰撞方程給出 $a \approx 2.990$ 及 $b \approx 9.821$，所以 $\theta \approx 73.06°$。於是 $u\sin(\theta+\alpha) \approx 14.612$，而

$v\sin(\theta+\beta) \approx 14.612$，兩者吻合。命中時靶子的高度為 $17.32 - 4.9 \approx 12.42 > 0$，而確實 $2u\sin\alpha(u\sin\alpha - v\sin\beta) \approx 340.2 > bg \approx 96.2$。我們用電腦檢驗了數以千計的隨機發射情況，證實該不等式成立當且僅當命中時靶子仍在空中，而且每一次命中都有 $\alpha > \beta$。

常見錯誤

- 忘記兩個質點必須在同一時刻處於同一位置，結果解出兩個不同的時刻。
- 未先確定 $u\sin\alpha - v\sin\beta$ 為正，便把兩邊同乘以它；它確實為正，因為 $t > 0$ 且 $b > 0$。
- 單憑 $u\sin\alpha > v\sin\beta$ 推出 $\alpha > \beta$；除非同時比較速率，否則這條不等式對角度毫無啟示。

STEP 之外

重力互相抵消，是伽利略相對性（Galilean relativity）的一個例子：在隨子彈一同自由下落的參考系中，靶子以恆定速度沿直線運動，所以「直接瞄準它」就行。這正是經典的「猴子與獵人」（monkey and hunter）示範實驗；而慣性參考系（inertial frame）（或自由下落參考系）的概念，是大學一年級力學的起點，日後亦是廣義相對論（general relativity）中等效原理（equivalence principle）的起點。Kleppner 和 Kolenkow 的 *An Introduction to Mechanics*（《力學導論》）對相對運動和參考系有細緻的闡述。

詳細解答

STEP II 2016，第 12 題

概率與統計 • 容斥原理與錯排

STEP II 2016，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/16-s2-q12

本題在 *Beyond STEP* 第 9.8 節沒有人拿到自己的恤衫中有完整的引導式講解：https://beyondstep.keithmcnulty.org/zh-hant/ch09/derangements.html

關鍵思路

對於三個事件（event），把 $A \cup B$ 看作單一事件，然後兩次運用兩個事件的公式。同樣的正負交替規律（加上單個事件、減去兩兩之交、加上三個之交，...）對任意數目的事件都成立。把這規律應用於「第 i 張卡在第 i 個位置」這些事件時，每個由 r 個事件組成的項都有相同的概率，所以和式的每一層都會化簡成一個分數。至於「恰好一個」，先固定哪一張卡是配對的那張，再要求其餘的卡完全沒有配對。

解答

三個事件

寫成 $A \cup B \cup C = (A \cup B) \cup C$，然後以 $A \cup B$ 代替 A、以 C 代替 B，運用兩個事件的公式：

$$\mathrm{P}(A \cup B \cup C) = \mathrm{P}(A \cup B) + \mathrm{P}(C) - \mathrm{P}\big((A \cup B) \cap C\big).$$

第一項是 $\mathrm{P}(A) + \mathrm{P}(B) - \mathrm{P}(A \cap B)$。至於最後一項，由分配律（distributive law）得 $(A \cup B) \cap C = (A \cap C) \cup (B \cap C)$，再運用一次兩個事件的公式，

$$\begin{aligned}\mathrm{P}\big((A \cap C) \cup (B \cap C)\big) &= \mathrm{P}(A \cap C) + \mathrm{P}(B \cap C) - \mathrm{P}\big((A \cap C) \cap (B \cap C)\big) \\ &= \mathrm{P}(A \cap C) + \mathrm{P}(B \cap C) - \mathrm{P}(A \cap B \cap C).\end{aligned}$$

把兩者代入，

$$\mathrm{P}(A\cup B\cup C)=\mathrm{P}(A)+\mathrm{P}(B)+\mathrm{P}(C)-\mathrm{P}(A\cap B)-\mathrm{P}(B\cap C)-\mathrm{P}(C\cap A)+\mathrm{P}(A\cap B\cap C),$$

得證。

四個事件

加上四個單一事件的概率，減去六個兩兩交集（intersection）的概率，加上四個三重交集的概率，再減去四重交集的概率：

$$\begin{aligned}\mathrm{P}(A\cup B\cup C\cup D)=\ &\mathrm{P}(A)+\mathrm{P}(B)+\mathrm{P}(C)+\mathrm{P}(D)\\&-\mathrm{P}(A\cap B)-\mathrm{P}(A\cap C)-\mathrm{P}(A\cap D)\\&-\mathrm{P}(B\cap C)-\mathrm{P}(B\cap D)-\mathrm{P}(C\cap D)\\&+\mathrm{P}(A\cap B\cap C)+\mathrm{P}(A\cap B\cap D)+\mathrm{P}(A\cap C\cap D)+\mathrm{P}(B\cap C\cap D)\\&-\mathrm{P}(A\cap B\cap C\cap D).\end{aligned}$$

第 (i)–(iii) 部分

全部 $n!$ 種排列的可能性相等。若要求某 r 張指定的卡各自在自己的位置上，餘下的 $n-r$ 張卡可以有 $(n-r)!$ 種排法，所以概率是 $(n-r)!/n!$。因此

$$\text{(i) }\mathrm{P}(E_i)=\frac{1}{n},\qquad \text{(ii) }\mathrm{P}(E_i\cap E_j)=\frac{1}{n(n-1)},\qquad \text{(iii) }\mathrm{P}(E_i\cap E_j\cap E_k)=\frac{1}{n(n-1)(n-2)}.$$

（第 (ii) 部分亦可這樣看：第 i 張卡落在第 i 個位置的概率是 $\frac{1}{n}$，然後第 j 張卡在餘下的 $n-1$ 個位置中落在第 j 個位置的概率是 $\frac{1}{n-1}$。）

至少一個配對

至少有一張卡配對，正正就是 $E_1\cup E_2\cup\cdots\cup E_n$ 發生。上述規律可推廣至 n 個事件：第 r 層是 $(-1)^{r+1}$ 乘以所有由 r 個不同事件組成的交集的概率之和。

> **想一想** 在不知道是哪些事件的情況下，怎樣計算一個遍及所有 r 個事件選法的和？

由對稱性（symmetry），每個由 r 個不同事件組成的交集都有相同的概率 $(n-r)!/n!$，而選出 r 個事件的方法有 $\binom{n}{r}$ 種。所以第 r 層貢獻

$$(-1)^{r+1}\binom{n}{r}\frac{(n-r)!}{n!}=(-1)^{r+1}\frac{n!}{r!\,(n-r)!}\cdot\frac{(n-r)!}{n!}=\frac{(-1)^{r+1}}{r!}.$$

對 $r=1,2,\ldots,n$ 求和，

$$\mathrm{P}(\text{至少一個配對})=1-\frac{1}{2!}+\frac{1}{3!}-\cdots+(-1)^{n+1}\frac{1}{n!},$$

得證。（當 $r = 1,2,3$ 時，這不過是第 (i)–(iii) 部分得出的 $n \cdot \frac{1}{n}$、$\binom{n}{2} \cdot \frac{1}{n(n-1)}$ 和 $\binom{n}{3} \cdot \frac{1}{n(n-1)(n-2)}$。）

恰好一個配對

以 p_m 表示 m 張洗亂了的卡中沒有任何一張在自己位置上的概率。由上一個結果，

$$p_m = 1 - \left(1 - \frac{1}{2!} + \cdots + (-1)^{m+1}\frac{1}{m!}\right) = \frac{1}{2!} - \frac{1}{3!} + \cdots + (-1)^m \frac{1}{m!}.$$

「恰好一個配對」這事件是各事件「第 i 張卡配對而其他卡都不配對」對 i 的不相交併集（disjoint union）。在 E_i 發生的條件下，其餘 $n-1$ 張卡在其餘 $n-1$ 個位置上的任何排列都有相等的可能性，所以這等同於重新洗一副 $n-1$ 張的卡，而其中沒有任何一張配對的概率是 p_{n-1}。因此

$$\mathrm{P}(\text{恰好一個配對}) = \sum_{i=1}^{n} \mathrm{P}(E_i)\, p_{n-1} = n \cdot \frac{1}{n} \cdot p_{n-1} = p_{n-1},$$

即是

$$\mathrm{P}(\text{恰好一個配對}) = \frac{1}{2!} - \frac{1}{3!} + \cdots + (-1)^{n-1}\frac{1}{(n-1)!} \qquad (n \geq 2),$$

而當 $n = 1$ 時概率為 1。等價地，它等於 $\sum_{r=0}^{n-1}(-1)^r/r!$，這個寫法連 $n = 1$ 的情況也包括在內。

驗算

當 $n = 3$ 時，排列有 123、132、213、231、312、321。其中四個有配對，得 $\frac{4}{6} = \frac{2}{3} = 1 - \frac{1}{2} + \frac{1}{6}$；而三個（即對換 132、213、321）恰好有一個配對，得 $\frac{1}{2}$，與公式相符。當 $n = 2$ 時，恰好一個配對是不可能的（若一張卡在其位置上，另一張也必然如此），而公式給出 0。以電腦列舉 $n \leq 7$ 的所有排列，結果與兩條公式都相符；例如當 $n = 7$ 時，恰好一個配對的概率是 $\frac{53}{144} \approx 0.368$。當 $n \to \infty$ 時，兩個答案都趨向熟悉的極限：至少一個配對趨向 $1 - \mathrm{e}^{-1} \approx 0.632$，而恰好一個配對趨向 e^{-1}。

常見錯誤

- 寫成 $\mathrm{P}(E_i \cap E_j) = \frac{1}{n^2}$：這些事件並不獨立（independent），因為一旦第 i 張卡在第 i 個位置，第 j 張卡便只剩下 $n-1$ 個位置。
- 把「恰好一個」答成 $n \cdot \frac{1}{n} = 1$ 或 $\mathrm{P}(E_1)$：我們需要第 i 張卡在其位置上，*而且*沒有其他卡在其位置上。
- 忘記「至少一個」的和式中 $r = 1$ 的項是 $n \cdot \frac{1}{n} = 1$，這正是級數以 1 開始的原因。

STEP 之外

這是經典的錯排（derangement）問題（即 *problème des rencontres*，「配對問題」），而一般的容斥（inclusion–exclusion）公式可以用歸納法（induction）或指示隨機變量（indicator random variables）簡潔地證明。原來當 n 很大時，配對的數目近似服從平均值為 1 的泊松分佈（Poisson distribution）：恰好有 k 個配對的概率趨向 $\mathrm{e}^{-1}/k!$。Grimmett 與 Welsh 的 *Probability: An Introduction*（《概率導論》），以及 Blitzstein 與 Hwang 的 *Introduction to Probability*（《概率導論》），都有討論容斥原理和配對問題。

詳細解答

STEP II 2016，第 13 題

概率與統計 • 正態近似

STEP II 2016，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/16-s2-q13

本題在 *Beyond STEP* 第 12.23 節由概率得出斯特靈公式中有完整的引導式講解：https://beyondstep.keithmcnulty.org/zh-hant/ch12/stirling-probability.html

關鍵思路

離散分佈（discrete distribution）的單一概率 $\mathrm{P}(X = k)$，是一條以 k 為中心、寬度為 1 的長條的面積。若該分佈接近正態分佈（normal distribution），這個面積大約等於 k 處的正態概率密度（density）乘以寬度 1。在平均值（mean）處應用這一點（標準正態分佈在該處的密度為 $\frac{1}{\sqrt{2\pi}}$），便能把一個已知的精確概率化為一個近似式。只要巧妙地選取分佈，就能令精確概率包含我們想要的階乘（factorial）。

解答

全題通用的近似

設 W 取整數值，平均值為 μ，方差（variance）為 σ^2，並且可用 $V \sim \mathrm{N}(\mu, \sigma^2)$ 很好地近似。V 的概率密度為

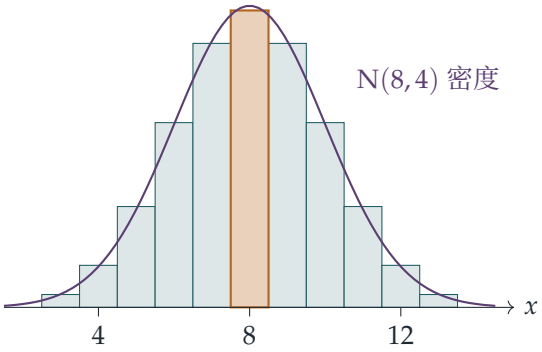
$$f(v) = \frac{1}{\sigma} \cdot \frac{1}{\sqrt{2\pi}} \, \mathrm{e}^{-\frac{1}{2}\left(\frac{v-\mu}{\sigma}\right)^2}.$$

加上連續性校正（continuity correction），

$$\mathrm{P}(W = k) \approx \mathrm{P}\left(k - \tfrac{1}{2} < V < k + \tfrac{1}{2}\right) = \int_{k-1/2}^{k+1/2} f(v)\,\mathrm{d}v \approx 1 \times f(k),$$

最後一步成立，是因為區間寬度為 1，而當 σ 不太小時，f 在區間上幾乎不變。在 $k=\mu$ 處，指數部分為 $\mathrm{e}^0=1$，所以

$$\mathrm{P}(W=\mu)\approx\frac{1}{\sigma\sqrt{2\pi}}. \qquad (*)$$



第 (i) 部分

對於 $X\sim\mathrm{B}(16,\frac{1}{2})$，

$$\mathrm{E}(X)=np=8,\qquad \mathrm{Var}(X)=np(1-p)=16\cdot\tfrac{1}{2}\cdot\tfrac{1}{2}=4,$$

所以 $\sigma=2$。由於 $p=\frac{1}{2}$，分佈關於 8 對稱，而且 n 相當大，因此 $\mathrm{N}(8,4)$ 是很好的近似（圖中把各長條與正態曲線作了比較）。由 $(*)$，陰影長條的面積為

$$\mathrm{P}(X=8)\approx\frac{1}{2\sqrt{2\pi}},$$

即為所求。

第 (ii) 部分

設 $Y\sim\mathrm{B}(2n,\frac{1}{2})$。則 $\mathrm{E}(Y)=n$，$\mathrm{Var}(Y)=2n\cdot\frac{1}{2}\cdot\frac{1}{2}=\frac{n}{2}$，所以 $\sigma=\sqrt{n/2}$。精確地，

$$\mathrm{P}(Y=n)=\binom{2n}{n}\left(\tfrac{1}{2}\right)^n\left(\tfrac{1}{2}\right)^n=\frac{(2n)!}{(n!)^2\,2^{2n}}.$$

當 n 很大時，正態近似適用，由 $(*)$ 得

$$\mathrm{P}(Y=n)\approx\frac{1}{\sqrt{n/2}\,\sqrt{2\pi}}=\frac{1}{\sqrt{n\pi}}.$$

令兩個表達式相等，再乘以 $2^{2n}(n!)^2$：

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

這正是所求的結果。（第 (i) 部分就是 $n=8$ 的情況：$\sqrt{8\pi}=2\sqrt{2\pi}$。）

第 (iii) 部分

設 $Z \sim \mathrm{Po}(n)$，於是 $\mathrm{E}(Z) = \mathrm{Var}(Z) = n$，$\sigma = \sqrt{n}$。精確地，

$$\mathrm{P}(Z = n) = \frac{\mathrm{e}^{-n} n^n}{n!}.$$

想一想 為甚麼這裏可以用正態分佈來近似泊松分佈（Poisson distribution）？

一個 $\mathrm{Po}(n)$ 隨機變量與 n 個獨立的 $\mathrm{Po}(1)$ 隨機變量之和同分佈，所以由中心極限定理（central limit theorem），當 n 很大時，它近似服從 $\mathrm{N}(n, n)$。這就是「當 λ 很大時 $\mathrm{Po}(\lambda) \approx \mathrm{N}(\lambda, \lambda)$」這條標準法則。

因此由 $(*)$，

$$\frac{\mathrm{e}^{-n} n^n}{n!} \approx \frac{1}{\sqrt{n}\sqrt{2\pi}} = \frac{1}{\sqrt{2\pi n}},$$

整理後得

$$n! \approx \sqrt{2\pi n}\, \mathrm{e}^{-n}\, n^n,$$

即為所求。

驗算

- 第 (i) 部分：精確值為 $\mathrm{P}(X = 8) = \binom{16}{8}/2^{16} = 12870/65536 \approx 0.1964$，而 $\frac{1}{2\sqrt{2\pi}} \approx 0.1995$，相差約 1.6% 以內。
- 第 (ii) 部分取 $n = 10$：$\binom{20}{10}/2^{20} \approx 0.1762$，對比 $1/\sqrt{10\pi} \approx 0.1784$。隨着 n 增大，階乘近似式兩邊之比趨向 1（在 $n = 100$ 時約為 0.9988）。
- 第 (iii) 部分取 $n = 10$：$10! = 3\,628\,800$，而 $\sqrt{20\pi}\,\mathrm{e}^{-10}\,10^{10} \approx 3\,598\,696$，偏小不足 1%。
- 一致性：把 (iii) 的結果同時代入 (ii) 中的 $(2n)!$ 和 $n!$，兩邊都得到 $\sqrt{4\pi n}\,\mathrm{e}^{-2n}(2n)^{2n}$，所以兩個近似式互相吻合。

常見錯誤

- 直接使用標準正態密度 $\frac{1}{\sqrt{2\pi}}$ 而沒有除以 σ。$\mathrm{N}(\mu, \sigma^2)$ 在平均值處的密度是 $\frac{1}{\sigma\sqrt{2\pi}}$。
- 在第 (ii) 部分中把標準差（standard deviation）誤用為 $\frac{n}{2}$ 而不是 $\sqrt{n/2}$，或把 $\mathrm{B}(2n, \frac{1}{2})$ 的方差算錯。
- 在第 (iii) 部分中忘記說明為甚麼泊松分佈可以用正態分佈近似。

STEP 之外

第 (iii) 部分的結果就是斯特林公式（Stirling's formula），而上述論證是中心極限定理在單一點上應用的一種啟發式（heuristic）形式，即所謂「局部極限定理」（local limit theorem）。要把它嚴格化需要小心：知道一個分佈函數接近正態分佈函數，本身並不能推出個別概率接近密度。斯特林公式也可以直接證明，例如把 $\ln n! = \sum \ln k$ 與 $\int \ln x \, \mathrm{d}x$ 作比較，或從伽瑪函數(gamma function)出發，運用拉普拉斯(Laplace)方法。Blitzstein 與 Hwang 的 *Introduction to Probability*（《概率導論》）講解了中心極限定理及泊松分佈與正態分佈之間的聯繫；Grimmett 與 Welsh 的 *Probability: An Introduction*（《概率入門》）是很好的輔助讀物。

詳細解答

STEP III 2016，第 1 題

純數學 • 歸約公式

STEP III 2016，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/16-s3-q1

關鍵思路

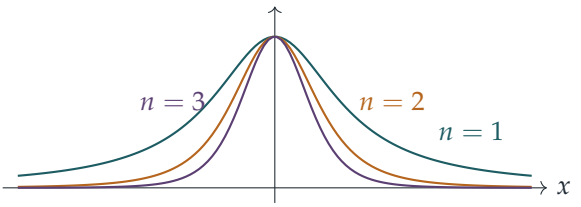
配方（completing the square）把二次式化成 $(x+a)^2+c^2$，其中 $c^2=b-a^2>0$，之後用正切代換（tangent substitution）便可立即求出 I_1。至於遞推公式（reduction formula），竅門在於對 1 進行分部積分（integration by parts）：對 $(x^2+2ax+b)^{-n}$ 求導會把冪次提升到 $n+1$，這正是我們把 I_n 與 I_{n+1} 聯繫起來所需要的。其後的歸納法（induction）只不過是整理二項式係數（binomial coefficient）而已。

解答

全文記

$$Q(x)=x^2+2ax+b=(x+a)^2+c^2, \qquad c=\sqrt{b-a^2}>0.$$

由於對所有 x 都有 $Q(x)\geq c^2>0$，被積函數（integrand）是連續的；又由於 $Q(x)^n$ 的增長速度與 x^{2n} 相若，且 $2n\geq 2$，所以每個 I_n 都收斂（converge）。



當 $a=0$、$b=1$ 時的被積函數 $1/Q^n$：隨着 n 增大，曲線變得愈來愈窄，因此我們預期當 $c=1$ 時 I_n 會逐漸變小。

第 (i) 部分

設 $x + a = c\tan u$，則 $\mathrm{d}x = c\sec^2 u\,\mathrm{d}u$。當 x 由 $-\infty$ 變到 ∞ 時，u 由 $-\frac{\pi}{2}$ 變到 $\frac{\pi}{2}$，而且

$$Q(x) = c^2\tan^2 u + c^2 = c^2\sec^2 u.$$

因此

$$I_1 = \int_{-\pi/2}^{\pi/2} \frac{c\sec^2 u}{c^2\sec^2 u}\,\mathrm{d}u = \frac{1}{c}\int_{-\pi/2}^{\pi/2}\mathrm{d}u = \frac{\pi}{c} = \frac{\pi}{\sqrt{b-a^2}},$$

即為所求。

第 (ii) 部分

對 I_n 進行分部積分，取 $u = Q^{-n}$ 及 $\mathrm{d}v = \mathrm{d}x$。取 $v = x + a$ 比取 $v = x$ 更巧妙（任何原函數（antiderivative）皆可），因為 $Q'(x) = 2(x+a)$。於是

$$\frac{\mathrm{d}}{\mathrm{d}x}Q^{-n} = -2n(x+a)\,Q^{-n-1},$$

所以

$$I_n = \Big[\frac{x+a}{Q^n}\Big]_{-\infty}^{\infty} + 2n\int_{-\infty}^{\infty}\frac{(x+a)^2}{Q^{n+1}}\,\mathrm{d}x.$$

> **想一想** 為甚麼邊界項（boundary term）會等於零？

分子的增長速度與 $|x|$ 相若，而分母的增長速度與 x^{2n} 相若，又因為 $n \geq 1$，所以 $2n \geq 2$。因此當 $x \to \pm\infty$ 時 $(x+a)/Q^n \to 0$，方括號內的項為零。

接着利用 $(x+a)^2 = Q - c^2$ 把餘下的積分拆開：

$$\begin{aligned} I_n &= 2n\int_{-\infty}^{\infty}\frac{Q-c^2}{Q^{n+1}}\,\mathrm{d}x \\ &= 2n\,I_n - 2nc^2\,I_{n+1}. \end{aligned}$$

移項得 $2nc^2 I_{n+1} = (2n-1)I_n$，即

$$2n(b-a^2)\,I_{n+1} = (2n-1)\,I_n,$$

即為所求。（由於 I_n 與 I_{n+1} 均收斂，這樣拆開積分是合理的。）

第 (iii) 部分

設 $P(n)$ 為命題

$$I_n = \frac{\pi}{2^{2n-2}\,c^{2n-1}}\binom{2n-2}{n-1},$$

其中 $c^{2n-1} = (b-a^2)^{n-\frac{1}{2}}$。

基本情況。當 $n=1$ 時，右方為 $\dfrac{\pi}{2^0 c^1}\binom{0}{0}=\dfrac{\pi}{c}$，由第 (i) 部分可知它等於 I_1。所以 $P(1)$ 成立。

歸納步驟。假設對某個 $n\geq 1$，$P(n)$ 成立。由第 (ii) 部分，

$$I_{n+1}=\frac{2n-1}{2nc^2}\,I_n=\frac{2n-1}{2n}\cdot\frac{\pi}{2^{2n-2}\,c^{2n+1}}\binom{2n-2}{n-1}.$$

我們要證明它等於 $\dfrac{\pi}{2^{2n}c^{2n+1}}\binom{2n}{n}$，這就是 $P(n+1)$。兩個二項式係數之間的關係是

$$\binom{2n}{n}=\frac{(2n)!}{n!\,n!}=\frac{(2n)(2n-1)}{n\cdot n}\cdot\frac{(2n-2)!}{(n-1)!\,(n-1)!}=\frac{2(2n-1)}{n}\binom{2n-2}{n-1}.$$

> **想一想** 怎樣才能令 2 的冪次吻合？

寫成 $\dfrac{2n-1}{2n}=\dfrac{4}{2^2}\cdot\dfrac{2n-1}{2n}=\dfrac{1}{4}\cdot\dfrac{2(2n-1)}{n}$。於是

$$\begin{aligned}I_{n+1}&=\frac{1}{4}\cdot\frac{2(2n-1)}{n}\cdot\frac{\pi}{2^{2n-2}\,c^{2n+1}}\binom{2n-2}{n-1}\\&=\frac{\pi}{2^{2n}\,c^{2n+1}}\cdot\frac{2(2n-1)}{n}\binom{2n-2}{n-1}=\frac{\pi}{2^{2n}\,c^{2n+1}}\binom{2n}{n}.\end{aligned}$$

由於 $2n=2(n+1)-2$、$n=(n+1)-1$ 及 $c^{2n+1}=(b-a^2)^{(n+1)-\frac12}$，這正是 $P(n+1)$。由歸納法，$P(n)$ 對所有正整數 n 成立：

$$I_n=\frac{\pi}{2^{2n-2}(b-a^2)^{n-\frac12}}\binom{2n-2}{n-1}.$$

驗算

- 取 $a=0$、$b=1$ 及 $n=2$，公式給出 $I_2=\frac{\pi}{4}\binom{2}{1}=\frac{\pi}{2}\approx 1.5708$。以數值方法計算 $\int_{-\infty}^{\infty}(1+x^2)^{-2}\,\mathrm{d}x$，結果準確至小數點後十位都吻合。
- 正切代換亦給出 $I_n=c^{1-2n}\int_{-\pi/2}^{\pi/2}\cos^{2n-2}u\,\mathrm{d}u$，這是標準的沃利斯積分（Wallis integral），其已知值與公式相符。
- 對多組滿足 $b>a^2$ 的隨機 a 和 b，數值積分證實公式在 $n=1,\dots,5$ 時成立，因此第 (ii) 部分的遞推關係（recurrence）亦獲證實。

常見錯誤

- 忘記說明第 (ii) 部分的邊界項為何等於零；這需要 $n\geq 1$。

- 在分部積分中取 $v = x$ 而不取 $v = x + a$。這樣做仍然可行，但會多出一項 $a\int (x + a)Q^{-n-1}\,\mathrm{d}x$ 需要處理。
- 在歸納步驟中弄錯 $b - a^2$ 的冪次：每一步都乘以 $1/c^2$，所以 c 的指數剛好增加 2，由 $2n - 1$ 變為 $2n + 1$。

STEP 之外

有理函數在整條實數線上的積分，是展示複分析（complex analysis）威力的絕佳例子：留數定理（residue theorem）只需考慮 $1/Q^n$ 在上半平面（upper half-plane）內位於 $z = -a + ic$ 的唯一極點（pole），寥寥數行便能求出 I_n。這裏出現的中心二項式係數（central binomial coefficient）亦見於 π 的沃利斯乘積（Wallis product）及隨機漫步（random walk）之中。Needham 的 *Visual Complex Analysis*（《直觀複分析》）以優美的幾何方式闡述留數。

詳細解答

STEP III 2016，第 2 題

純數學 • 拋物線的法線

STEP III 2016，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/16-s3-q2

關鍵思路

先寫出一般點 $(at^2, 2at)$ 處的法線（normal），再問哪些 t 值使法線經過 P。這個條件是一個關於 t 的三次方程（cubic），而且其中一個根 $t = p$ 必定存在（P 處的法線當然經過 P）。把這個根除去，剩下一個二次方程（quadratic），它的兩個根正好是 q 和 r。此後的一切都由兩根之和與兩根之積，即 $q + r$ 和 qr，來推動。

解答

第 (i) 部分

在 $y^2 = 4ax$ 上取 $x = at^2$、$y = 2at$，得 $\dfrac{\mathrm{d}y}{\mathrm{d}x} = \dfrac{2a}{2at} = \dfrac{1}{t}$，所以參數為 t 的點處的法線斜率（gradient）為 $-t$：

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

（這個形式同樣適用於頂點（vertex）$t = 0$，該處的法線就是 x 軸。）

t 處的法線經過 $P(ap^2, 2ap)$ 當且僅當

$$2ap + tap^2 = 2at + at^3.$$

除以 $a > 0$ 並整理，得

$$\begin{aligned} 0 &= 2(p - t) + t(p^2 - t^2) \\ &= (p - t)\big(2 + t(p + t)\big) \\ &= (p - t)\,(t^2 + pt + 2). \end{aligned}$$

想一想 為甚麼可以捨去因子 $p-t$？

根 $t=p$ 是平凡的那一個：P 本身的法線經過 P。題目說明 $Q\neq P$，所以 $q\neq p$，因子 $p-q$ 不為零。因此

$$q^2+qp+2=0,$$

正是所求。同理 $r^2+rp+2=0$。

第 (ii) 部分

於是 q 和 r 都是 $t^2+pt+2=0$ 的根。由於 $Q\neq R$，有 $q\neq r$，所以它們就是這個方程的兩個根，從而

$$q+r=-p,\qquad qr=2.$$

（特別地，判別式（discriminant）為正：$p^2>8$。另外留意，$q=p$ 需要 $2p^2+2=0$，這是不可能的，所以 Q 和 R 自動與 P 不同。）

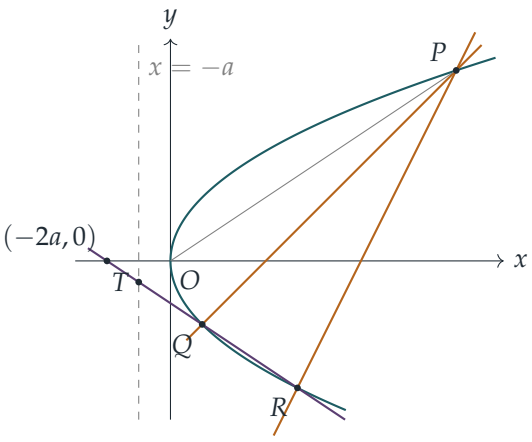
經過 $(aq^2,2aq)$ 和 $(ar^2,2ar)$ 的弦（chord）的斜率為 $\dfrac{2a(q-r)}{a(q^2-r^2)}=\dfrac{2}{q+r}$，其方程為

$$(q+r)\,y=2x+2aqr.$$

（驗證：在 Q 處左邊是 $2aq^2+2aqr$，與右邊相等；在 R 處同理。）代入 $q+r=-p$ 和 $qr=2$：

$$-p\,y=2x+4a.$$

無論 p 取何值，點 $(-2a,0)$ 都滿足這個方程。所以 QR 必定經過 $(-2a,0)$，而這一點與 P 的選取無關。



上圖顯示 $a=1$、$p=3$ 的情況，此時 $q=-1$、$r=-2$：兩條法線（琥珀色）相交於 P，弦 QR（紫紅色）經過 $(-2,0)$，而 T 位於虛線 $x=-1$ 上。

第 (iii) 部分

由於 $p^2 > 8$，$p \neq 0$，所以 $P \neq O$，直線 OP 為 $y = \dfrac{2ap}{ap^2}\,x = \dfrac{2}{p}\,x$，即 $x = \frac{1}{2}py$。（它與 QR 不平行，因為後者的斜率是 $-2/p$。）

把 $x = \frac{1}{2}py$ 代入 QR 的方程：

$$-py = py + 4a \quad\Longrightarrow\quad y = -\frac{2a}{p}, \qquad x = \tfrac{1}{2}p\Big(-\frac{2a}{p}\Big) = -a.$$

所以 $T = \left(-a, -\frac{2a}{p}\right)$，它必定位於直線 $x = -a$（拋物線的準線（directrix））上，與 P 的選取無關。

T 與 x 軸的距離是 $\dfrac{2a}{|p|}$。由第 (ii) 部分可知，經過 P 存在兩條不同的法線，這迫使 $p^2 > 8$，即 $|p| > 2\sqrt{2}$。由於 $a > 0$，

$$\frac{2a}{|p|} < \frac{2a}{2\sqrt{2}} = \frac{a}{\sqrt{2}},$$

證明完畢。不等式是嚴格的，因為 $p^2 = 8$ 會使 $q = r$。

驗算

- 在圖中 $a = 1$、$p = 3$ 的情況下：$t^2 + 3t + 2$ 的根是 -1 和 -2；法線 $y = x - 3$ 和 $y = 2x - 12$ 都經過 $P(9, 6)$；經過 $(1, -2)$ 和 $(4, -4)$ 的弦是 $-3y = 2x + 4$，它經過 $(-2, 0)$；而 $T = (-1, -\frac{2}{3})$，且 $\frac{2}{3} < \frac{1}{\sqrt{2}} \approx 0.707$。
- 令 $|p| \to 2\sqrt{2}$ 從上方趨近，Q 和 R 會合而為一，T 的高度趨於 $a/\sqrt{2}$，所以這個界是最佳的。以電腦對 20,000 組隨機選取的 a 和 p 進行檢驗，所有結果均得到確認，T 的高度最高達到 $a/\sqrt{2}$ 的 0.9996 倍，但從未超過。

常見錯誤

- 忘記說明為何可以除以 $p - q$：正是條件 $Q \neq P$ 容許這樣做。
- 沒有利用 $q \neq r$ 來說明 q 和 r 都是同一個二次方程的根，而這正是得出 $q + r = -p$ 和 $qr = 2$ 的依據。
- 證明最後的界時沒有交代 $p^2 > 8$ 從何而來：它來自二次方程必須有兩個相異實根這一要求。

STEP 之外

第 (i) 部分背後關於 t 的三次方程，是一個經典事實的特例：對一般點 (h,k)，條件是 $at^3+(2a-h)\,t-k=0$，所以從平面上一點最多可以作三條拋物線的法線，而「一條法線」與「三條法線」兩個區域之間的分界是一條稱為漸屈線（evolute）的曲線（這裏是一條半立方拋物線（semicubical parabola））。漸屈線、包絡線（envelope）和曲率（curvature）都是曲線微分幾何（differential geometry）入門課程的標準課題。Siklos 的 *Advanced Problems in Mathematics*（《數學進階題》）收錄了更多這種坐標幾何風格的 STEP 題目。

詳細解答

STEP III 2016，第 3 題

純數學 • 積分與多項式

STEP III 2016，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/16-s3-q3

本題在 *Beyond STEP* 第 4.6 節沒有公式的積分中有完整的引導式講解：https://beyondstep.keithmcnulty.org/zh-hant/ch04/no-formula.html

關鍵思路

形如 $\int g(x)\,\mathrm{e}^x\,\mathrm{d}x = \frac{P}{Q}\,\mathrm{e}^x + \text{常數}$ 的等式，其實只是一個關於導數(derivative)的陳述，所以把右邊求導，再消去分母即可。這樣問題便化為一個多項式恆等式（polynomial identity），而多項式恆等式有兩種處理方法：代入一個精心挑選的值（這裏是 $x = -1$），以及比較次數（degree）。

解答

以下把 $P(x)$、$Q(x)$ 簡寫為 P、Q。由商法則（quotient rule）和積法則（product rule），

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

因此，積分等式成立當且僅當此式等於被積函數（integrand）（其中 Q 不是零多項式）。

第 (i) 部分

令導數等於 $\frac{x^3-2}{(x+1)^2}\,\mathrm{e}^x$，約去 $\mathrm{e}^x \neq 0$ 並去分母，得多項式恆等式

$$(x+1)^2\left(PQ + P'Q - PQ'\right) = (x^3 - 2)\,Q^2. \qquad (*)$$

Q 有因式 $x+1$。在 $(*)$ 中代入 $x = -1$。左邊為 0，右邊為 $-3\,Q(-1)^2$。因此 $Q(-1) = 0$，由因式定理（factor theorem）知 $x+1$ 整除 Q。

次數。P 不是零多項式，否則導數會是 0，而不是被積函數。設 $\deg P = p$、$\deg Q = q$，首項係數（leading coefficient）分別為 $a \neq 0$ 和 $b \neq 0$。在 $PQ + P'Q - PQ'$ 中，PQ 一項的次數為 $p+q$，首項係數為 ab；而 $P'Q$ 和 PQ' 的次數各自至多為 $p+q-1$。所以 $PQ + P'Q - PQ'$ 的次數恰好是 $p+q$，而 $(*)$ 左邊的次數是 $p+q+2$。右邊的次數是 $3+2q$。因此

$$p+q+2 = 2q+3 \quad \Longrightarrow \quad p = q+1,$$

即 P 的次數恰好比 Q 的次數多一。

$Q = x+1$ 的情況。此時 P 的次數為 2，設 $P = ax^2 + bx + c$。由於 $Q = x+1$、$Q' = 1$，$(*)$ 變為

$$(x+1)^2\big((x+1)P + (x+1)P' - P\big) = (x^3-2)(x+1)^2,$$

約去 $(x+1)^2$ 得

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

由於 $xP = ax^3 + bx^2 + cx$，而 $(x+1)P' = (x+1)(2ax+b) = 2ax^2 + (2a+b)x + b$，

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

比較係數：$a = 1$；由常數項得 $b = -2$；於是 x^2 項的條件 $2a+b = 0$ 亦相符；而由 x 項得 $c = -(2a+b) = 0$。因此

$$P(x) = x^2 - 2x, \qquad \int \frac{x^3-2}{(x+1)^2}\,\mathrm{e}^x\,\mathrm{d}x = \frac{x^2-2x}{x+1}\,\mathrm{e}^x + \text{常數}.$$

第 (ii) 部分

用反證法（proof by contradiction），假設存在這樣的 P 和 Q，且兩者沒有公因式。和之前一樣，求導並去分母，得

$$(x+1)\big(PQ + P'Q - PQ'\big) = Q^2. \qquad (**)$$

代入 $x = -1$ 得 $Q(-1)^2 = 0$，所以 $x+1$ 整除 Q。記

$$Q = (x+1)^n R, \qquad n \geq 1, \quad R(-1) \neq 0,$$

使 n 恰好是 Q 中 $x+1$ 的冪次。於是 $Q' = n(x+1)^{n-1}R + (x+1)^n R'$，而

$$\begin{aligned} PQ + P'Q - PQ' &= (x+1)^n\big(PR + P'R - PR'\big) - n(x+1)^{n-1}PR \\ &= (x+1)^{n-1}\Big[(x+1)\big(PR + P'R - PR'\big) - nPR\Big]. \end{aligned}$$

代入 $(**)$，並從兩邊約去公因式 $(x+1)^n$（右邊是 $(x+1)^{2n}R^2$）：

$$(x+1)\big(PR + P'R - PR'\big) - nPR = (x+1)^n R^2.$$

想一想 為甚麼要費功夫把 $x+1$ 的確切冪次抽出來，而不是直接再把 $x=-1$ 代入 $(**)$？

直接把 $x=-1$ 代入 $(**)$，只能告訴我們 $Q(-1)=0$；之後每一項都等於零，我們對 P 一無所知。把 Q 中所有 $x+1$ 的因式移除後，便剩下一項 $-nPR$，它在 $x=-1$ 時不會自動等於零。

現在代入 $x=-1$。由於 $n\geq 1$，右邊為 0，左邊第一項亦為 0。因此

$$-n\,P(-1)\,R(-1)=0.$$

但 $n\neq 0$ 且 $R(-1)\neq 0$，所以 $P(-1)=0$。於是 $x+1$ 既是 Q 的因式，也是 P 的因式，與兩者沒有公因式的假設矛盾。所以這樣的多項式不存在。 ■

（限定 P、Q 互質（coprime）並無損一般性：任何公因式都可以從 P/Q 中約去，而不改變該函數。）

驗算

- 直接對 (i) 的答案求導：$\frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{x^2-2x}{x+1}\right)=\frac{x^2+2x-2}{(x+1)^2}$，再加上 $\frac{x^2-2x}{x+1}=\frac{x^3-x^2-2x}{(x+1)^2}$，得 $\frac{x^3-2}{(x+1)^2}$，正如所需。
- 數值上，$\int_0^1 \frac{x^3-2}{(x+1)^2}\,\mathrm{e}^x\,\mathrm{d}x\approx -1.359141$，與 $\left[\frac{x^2-2x}{x+1}\mathrm{e}^x\right]_0^1=-\frac{\mathrm{e}}{2}$ 相符。
- 在 (i) 中次數吻合：P 的次數為 2，比 $Q=x+1$ 多一。

常見錯誤

- 忘記 $P'Q-PQ'$ 的次數小於 $p+q$，因為首項互相抵消；決定次數的只有 PQ。
- 在 (ii) 中，只把 $x=-1$ 代入 $(**)$，求得 $Q(-1)=0$ 便停下來。要得出矛盾，必須用到 Q 中 $x+1$ 的確切冪次。

STEP 之外

第 (ii) 部分說明，$\int \frac{\mathrm{e}^x}{x+1}\,\mathrm{d}x$（本質上就是指數積分（exponential integral）Ei）在這個受限的意義下並不是「初等」（elementary）的。判定一個積分究竟能否用初等函數表示的完整理論，是劉維爾（Liouville）定理以及微分代數（differential algebra）中的 Risch 算法；這裏的論證只是其中一個小例子。若想仔細了解上面用到的多項式代數（因式定理、整除性、互質），可參閱 Allenby 的 *Rings, Fields and Groups*（《環、域與羣》）。

詳細解答

STEP III 2016，第 4 題

純數學 • 裂項相消級數

STEP III 2016，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/16-s3-q4

關鍵思路

題目建議的兩個分數之差，原來是通項（general term）的常數倍，因此這個和會伸縮相消（telescoping）：幾乎所有項都互相抵消，只剩下首尾兩部分。第 (ii) 部分中，把 sech 用指數函數表示並提出一個因子後，每一項看起來就與第 (i) 部分取 $x = \mathrm{e}^{-2y}$ 時的項完全一樣。由於 sech 是偶函數（even function），雙向無限和便可以摺回單向的和。

解答

第 (i) 部分

把建議的差化為公分母（common denominator），

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

由於 $|x| \neq 1$，故 $x \neq 1$，所以可以除以 $x-1$。（這個條件亦保證 $x \neq -1$，因此沒有分母 $1+x^r$ 會等於零：若 $x^r = -1$，則必有 $|x| = 1$。）因此

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

由 $r = 1$ 加到 N，每個括號中的第二個分數都與下一個括號的第一個分數抵消，剩下

$$\begin{aligned}\sum_{r=1}^{N} \frac{x^r}{(1+x^r)(1+x^{r+1})} &= \frac{1}{x-1}\left(\frac{1}{1+x} - \frac{1}{1+x^{N+1}}\right)\\ &= \frac{x\,(x^N-1)}{(x-1)(1+x)(1+x^{N+1})},\end{aligned}$$

其中第二個形式是把兩個分數合併而得：分子為 $x^{N+1} - x = x(x^N - 1)$。

現設 $|x|<1$。當 $N\to\infty$ 時 $x^{N+1}\to 0$，因此部分和（partial sums）趨向

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

這正是所求的無窮級數之和。

第 (ii) 部分

對任何實數 t，

$$\operatorname{sech} t=\frac{2}{\mathrm{e}^{t}+\mathrm{e}^{-t}}=\frac{2\mathrm{e}^{-t}}{1+\mathrm{e}^{-2t}}.$$

設 $x=\mathrm{e}^{-2y}$；由於 $y>0$，故 $0<x<1$。於是 $\operatorname{sech}(ry)=2\mathrm{e}^{-ry}/(1+x^r)$，而

$$\operatorname{sech}(ry)\operatorname{sech}\big((r+1)y\big)=\frac{4\,\mathrm{e}^{-(2r+1)y}}{(1+x^r)(1+x^{r+1})}=4\mathrm{e}^{-y}\,\frac{x^r}{(1+x^r)(1+x^{r+1})},$$

因為 $\mathrm{e}^{-2ry}=x^r$。

> **想一想** 為甚麼要把 $\mathrm{e}^{-(2r+1)y}$ 拆成 $\mathrm{e}^{-y}\cdot\mathrm{e}^{-2ry}$？

我們希望分子是 x^r，即分母中出現的同一個 x 的冪，這樣第 (i) 部分便可以逐項套用；剩下的因子 e^{-y} 與 r 無關，可以提到和號之外。

由於 $|x|<1$，由第 (i) 部分得

$$\sum_{r=1}^{\infty}\operatorname{sech}(ry)\operatorname{sech}\big((r+1)y\big)=4\mathrm{e}^{-y}\,\frac{x}{1-x^2}=\frac{4\mathrm{e}^{-y}}{x^{-1}-x}=\frac{4\mathrm{e}^{-y}}{\mathrm{e}^{2y}-\mathrm{e}^{-2y}}=\frac{4\mathrm{e}^{-y}}{2\sinh 2y},$$

即 $2\mathrm{e}^{-y}\operatorname{cosech}(2y)$，與題目所求相同。把這個和記作 S。

雙向無限和

把對所有整數 r 求的和分成三部分：$r\ge 1$、$r\in\{-1,0\}$ 及 $r\le -2$。

> **想一想** r 為負數的項是怎樣的？

對 $r\le -2$，設 $s=-r-1$，則 $s\ge 1$，且 $r+1=-s$、$r=-(s+1)$。由於 sech 是偶函數，

$$\operatorname{sech}(ry)\operatorname{sech}\big((r+1)y\big)=\operatorname{sech}\big((s+1)y\big)\operatorname{sech}(sy),$$

而當 r 取遍 $-2,-3,\ldots$ 時，s 取遍 $1,2,\ldots$。所以 $r\le -2$ 的各項之和亦為 S。（兩個單向的和都收斂（converge），因此這樣拆分雙向無限和是合理的。）

中間的兩項為

$$r=-1:\ \operatorname{sech}(-y)\operatorname{sech}0=\operatorname{sech}y,\qquad r=0:\ \operatorname{sech}0\operatorname{sech}y=\operatorname{sech}y.$$

因此，利用 $\sinh 2y = 2\sinh y\cosh y$，

$$\sum_{r=-\infty}^{\infty} \operatorname{sech}(ry)\operatorname{sech}\big((r+1)y\big) = 2S + 2\operatorname{sech} y = \frac{2\mathrm{e}^{-y}}{\sinh y\cosh y} + \frac{2\sinh y}{\sinh y\cosh y}$$
$$= \frac{2(\mathrm{e}^{-y}+\sinh y)}{\sinh y\cosh y}.$$

但 $\mathrm{e}^{-y} + \sinh y = \mathrm{e}^{-y} + \frac{1}{2}(\mathrm{e}^{y} - \mathrm{e}^{-y}) = \frac{1}{2}(\mathrm{e}^{y} + \mathrm{e}^{-y}) = \cosh y$，所以 $\cosh y$ 互相抵消，得

$$\sum_{r=-\infty}^{\infty} \operatorname{sech}(ry)\operatorname{sech}\big((r+1)y\big) = \frac{2}{\sinh y} = 2\operatorname{cosech} y.$$

驗算

- 取 $N = 1$，有限和公式給出 $\dfrac{x(x-1)}{(x-1)(1+x)(1+x^2)} = \dfrac{x}{(1+x)(1+x^2)}$，正是 $r = 1$ 的單一項。
- 取 $x = \frac{1}{2}$：無窮級數之和應為 $\frac{1/2}{3/4} = \frac{2}{3}$，而部分和確實趨近 0.6667。
- 取 $y = 1$：$2\mathrm{e}^{-1}\operatorname{cosech} 2 \approx 0.202864$，$2\operatorname{cosech} 1 \approx 1.701836$，兩者都與直接對級數求和的結果在多個小數位上吻合（已用電腦對多個 y 值驗證）。
- 當 $y \to \infty$ 時，只有 $r = -1, 0$ 兩項不會消失，每項約為 $2\mathrm{e}^{-y}$，合共約 $4\mathrm{e}^{-y}$；而 $2\operatorname{cosech} y \approx 4\mathrm{e}^{-y}$，亦相符。

常見錯誤

- 忘記因子 $\frac{1}{x-1}$，或弄錯其正負號：該差等於項的 $(x-1)$ 倍，而不是 $(1-x)$ 倍。
- 在雙向無限和中，只把 S 乘以二，卻忘記 $r = -1$ 及 $r = 0$ 兩項；這兩項涉及 $\operatorname{sech} 0 = 1$，並不屬於任何一個單向的和。

STEP 之外

在這裏重排及拆分無窮級數是安全的，因為每一項都是正數（級數絕對收斂（converges absolutely））；對於條件收斂（conditionally convergent）的級數，黎曼（Riemann）重排定理（rearrangement theorem）指出，重新排列各項可以令和變成任何你想要的數值——這是大學第一個分析（analysis）課程中會遇到的驚人結果。像 $\operatorname{sech}(ry)\operatorname{sech}\big((r+1)y\big)$ 這類乘積對所有整數求和，亦會在橢圓函數（elliptic functions）理論中出現，而這類雙邊級數（bilateral series）往往有優美的閉合形式。關於級數的分析，可參閱 Abbott，*Understanding Analysis*（《理解分析》）。

詳細解答

STEP III 2016，第 5 題

純數學 • 質數的乘積

STEP III 2016，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/16-s3-q5

本題在 *Beyond STEP* 第 1.2 節質數分佈得有多稀疏？中有完整的引導式講解：https://beyondstep.keithmcnulty.org/zh-hant/ch01/prime-distribution.html

關鍵思路

中間的二項式係數（binomial coefficient）$\binom{2m+1}{m}$ 並不大（它和它的「孿生」項 $\binom{2m+1}{m+1}$ 與其他各項共同分享總和 2^{2m+1}），但它可被 $m+1$ 與 $2m+1$ 之間的每一個質數（prime）整除。因此這些質數的乘積也不大。把不超過 $2m+1$ 的質數分成「不超過 $m+1$」和「大於 $m+1$」兩組，便可以建立一個強歸納法（strong induction）；而偶數完全不會帶來額外麻煩，因為它們（除了 2 以外）都不是質數。

解答

第 (i) 部分

在 $(1+x)^{2m+1}$ 的二項展開式（binomial expansion）中代入 $x=1$：

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

左邊每一項都是正數。由對稱性，$\binom{2m+1}{m}=\binom{2m+1}{m+1}$，而這是和式中兩個不同的項。由於 $m\geq 1$，項 $\binom{2m+1}{0}=1$ 是第三個不同的項。因此

$$2\binom{2m+1}{m}=\binom{2m+1}{m}+\binom{2m+1}{m+1}<2^{2m+1},$$

所以 $\binom{2m+1}{m}<2^{2m}$，得證。

第 (ii) 部分

寫成

$$\binom{2m+1}{m} = \frac{(2m+1)!}{m!\,(m+1)!}, \qquad \text{所以} \qquad m!\,(m+1)!\binom{2m+1}{m} = (2m+1)!.$$

設 p 為滿足 $m+1 < p \leq 2m+1$ 的質數。那麼 p 是 $(2m+1)!$ 的其中一個因數，所以 p 整除右邊。但 $m!\,(m+1)!$ 的每個因數都不超過 $m+1 < p$，所以質數 p 不整除其中任何一個，因此 $p \nmid m!\,(m+1)!$。若一個質數整除乘積 ab，它必定整除 a 或 b，所以 $p \mid \binom{2m+1}{m}$。

想一想 每個這樣的質數都整除該係數。為甚麼它們的乘積也整除它？

這些質數互不相同，所以它們兩兩互質（pairwise coprime），而一個能被若干個兩兩互質的數整除的數，必能被它們的乘積整除。（等價地，由唯一分解定理（unique factorisation），這些質數每一個都出現在 $\binom{2m+1}{m}$ 的質因數分解中。）所以 $P_{m+1,2m+1}$ 整除 $\binom{2m+1}{m}$。若不存在這樣的質數，則 $P_{m+1,2m+1} = 1$，它整除一切整數。

正整數的正因數不會大於該數本身，所以由第 (i) 部分，

$$P_{m+1,2m+1} \leq \binom{2m+1}{m} < 2^{2m}.$$

第 (iii) 部分

由於 $m \geq 1$，我們有 $1 < m+1 < 2m+1$。每個質數 $p \leq 2m+1$ 要麼滿足 $p \leq m+1$，要麼滿足 $m+1 < p \leq 2m+1$，但不會兩者同時成立，所以

$$P_{1,2m+1} = P_{1,m+1}\,P_{m+1,2m+1}.$$

此外 $2 \leq m+1 \leq 2m$（因為 $m \geq 1$），所以 $k = m+1$ 是假設所涵蓋的值之一，從而得到 $P_{1,m+1} < 4^{m+1}$。結合第 (ii) 部分，

$$P_{1,2m+1} < 4^{m+1} \cdot 2^{2m} = 4^{m+1} \cdot 4^m = 4^{2m+1}.$$

第 (iv) 部分

我們對 n 使用強歸納法。設 $S(n)$ 為命題 $P_{1,n} < 4^n$。

基本情況。 $P_{1,2} = 2 < 16 = 4^2$。

歸納步驟。設 $n \geq 3$，並假設 $S(k)$ 對 $k = 2, 3, \ldots, n-1$ 均成立。

- 若 n 是奇數，寫成 $n = 2m+1$，其中 $m \geq 1$。那麼 $n-1 = 2m$，所以 $S(k)$ 對 $k = 2, \ldots, 2m$ 成立。由第 (iii) 部分得 $P_{1,2m+1} < 4^{2m+1}$，這正是 $S(n)$。

- 若 n 是偶數，則 $n \geq 4$ 是大於 2 的偶數，所以它不是質數。因此不超過 n 的質數恰好就是不超過 $n-1$ 的質數，於是

$$P_{1,n} = P_{1,n-1} < 4^{n-1} < 4^n.$$

想一想 為甚麼我們不能對每個 n 都直接使用第 (iii) 部分？

第 (iii) 部分只處理奇數 $2m+1$。對於偶數 n，我們需要另一個論證；而上面的論證很簡單，因為大於 2 的偶數不會為乘積添加新的質數。

所以 $S(n)$ 在兩種情況下都成立，由強歸納法，$P_{1,n} < 4^n$ 對所有 $n \geq 2$ 成立。

驗算

- $m = 3$：$\binom{7}{3} = 35 < 64 = 2^6$。在 $(4, 7]$ 中的質數是 5 和 7，而 $P_{4,7} = 35$ 恰好整除 35。
- $m = 5$：$\binom{11}{5} = 462 = 2 \cdot 3 \cdot 7 \cdot 11$，它可被 $P_{6,11} = 77$ 整除。此外 $462 < 1024$。
- $P_{1,10} = 2 \cdot 3 \cdot 5 \cdot 7 = 210$，遠小於 $4^{10} = 1\,048\,576$。電腦驗證確認 $P_{1,n} < 4^n$ 對 2 至 1500 的每個 n 都成立。

常見錯誤

- 在第 (i) 部分把 $<$ 寫成 $\leq$：要得到嚴格不等式，除了兩個中間項之外，還需要第三個正項，例如 $\binom{2m+1}{0}$。
- 在第 (iii) 部分，忘記檢查 $m+1$ 是否位於假設適用的範圍 $2, \ldots, 2m$ 之內。
- 在第 (iv) 部分，只處理奇數 n，或者忘記基本情況 $n = 2$。

STEP 之外

這是艾狄胥（Erdős）對上界 $\prod_{p \leq n} p < 4^n$ 的初等證明。取對數後，它說明切比雪夫（Chebyshev）函數 $\theta(n) = \sum_{p \leq n} \ln p$ 小於 $n \ln 4$。這是切比雪夫定理的其中一半；該定理指出不超過 n 的質數個數的數量級為 $n / \ln n$。艾狄胥用同類的二項式係數論證證明了伯特蘭公設（Bertrand's postulate）：在 n 與 $2n$ 之間必定存在一個質數。Davenport 的 *The Higher Arithmetic*（《高等算術》）是進一步學習初等數論的好選擇。

詳細解答

STEP III 2016，第 6 題

純數學 • 雙曲函數

STEP III 2016，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/16-s3-q6

關鍵思路

把 $\operatorname{sech} x = a\tanh x + b$ 兩邊乘以 $\cosh x$，得 $a\sinh x + b\cosh x = 1$，這正是開首部分所處理的組合。因此第一步就是把大家熟悉的「$R\cos(x+\alpha)$」技巧搬到雙曲函數（hyperbolic functions）上。與三角函數的情況不同，所得的形式取決於 B 與 $\pm A$ 的大小關係，而每種形式都能讓你立即看出解的個數。

解答

$A\sinh x + B\cosh x$ 的各種形式

展開，

$$R\cosh(x+\gamma) = R\sinh\gamma\,\sinh x + R\cosh\gamma\,\cosh x,$$

所以我們需要 $R\sinh\gamma = A$ 及 $R\cosh\gamma = B$。由於 $\cosh^2\gamma - \sinh^2\gamma = 1$，這就迫使 $R^2 = B^2 - A^2$ 及 $\tanh\gamma = A/B$。當 $B > A > 0$ 時，兩者都可以滿足：取

$$R = \sqrt{B^2 - A^2} > 0, \qquad \gamma = \operatorname{artanh}\frac{A}{B},$$

由於 $0 < A/B < 1$，γ 是存在的。於是 $\cosh\gamma = 1/\sqrt{1 - A^2/B^2} = B/R$，且 $\sinh\gamma = \tanh\gamma\cosh\gamma = A/R$，所以 $R\cosh\gamma = B$ 及 $R\sinh\gamma = A$，符合要求。因此 $A\sinh x + B\cosh x = R\cosh(x+\gamma)$。

其他情況（設 $A > 0$）：

- $-A < B < A$：改為比較 $R\sinh(x+\gamma) = R\cosh\gamma\sinh x + R\sinh\gamma\cosh x$ 的係數，得

$$A\sinh x + B\cosh x = \sqrt{A^2 - B^2}\,\sinh(x+\gamma), \qquad \gamma = \operatorname{artanh}\frac{B}{A}.$$

- $B = A$：$A(\sinh x + \cosh x) = A\mathrm{e}^{x}$。
- $B = -A$：$A(\sinh x - \cosh x) = -A\mathrm{e}^{-x}$。
- $B < -A$：把第一個結果應用於 $-A\sinh x - B\cosh x$（此時 $-B > A$ 而 $-A < 0$；上面的推導其實只需要 $|A| < B$），得

$$A\sinh x + B\cosh x = -\sqrt{B^2 - A^2}\,\cosh(x + \gamma), \qquad \gamma = \operatorname{artanh}\frac{A}{B}.$$

交點方程

由於 $\cosh x > 0$，兩曲線相交於

$$\operatorname{sech} x = a\tanh x + b \iff 1 = a\sinh x + b\cosh x. \tag{$*$}$$

第 (i) 部分

當 $b > a > 0$ 時，記 $R = \sqrt{b^2 - a^2}$；於是 $(*)$ 變成

$$\cosh\left(x + \operatorname{artanh}\tfrac{a}{b}\right) = \frac{1}{\sqrt{b^2 - a^2}}.$$

若兩曲線相交，則此方程有解，所以右方至少為 1。當 $c \geq 1$ 時，$\cosh u = c$ 的解為 $u = \pm\operatorname{arcosh} c$，所以

$$x = \pm\operatorname{arcosh}\left(\frac{1}{\sqrt{b^2 - a^2}}\right) - \operatorname{artanh}\frac{a}{b},$$

這正是所要求的形式。

第 (ii) 部分

當 $a > b > 0$ 時，屬於 sinh 的情況：$(*)$ 變成 $\sqrt{a^2 - b^2}\,\sinh\left(x + \operatorname{artanh}\frac{b}{a}\right) = 1$。由於 sinh 是從 $\mathbb{R}$ 到 $\mathbb{R}$ 的雙射（bijection），兩曲線必定恰有一個交點，位於

$$x = \operatorname{arsinh}\left(\frac{1}{\sqrt{a^2 - b^2}}\right) - \operatorname{artanh}\frac{b}{a}.$$

第 (iii) 部分

按開首部分的每種情況（取 $A = a > 0$，$B = b$）數算 $(*)$ 的解：

- $b < -a$：左方 $-\sqrt{b^2 - a^2}\cosh(x + \gamma)$ 是負數，所以無解。
- $b = -a$：$-a\mathrm{e}^{-x} = 1$ 無解。
- $-a < b < a$：恰有一個解，與 (ii) 相同；那裏的論證對此範圍內任何 b 都成立。

- $b = a$：$ae^x = 1$ 只有一個解 $x = -\ln a$。
- $b > a$：若 $1/\sqrt{b^2 - a^2} > 1$，則 $\cosh(x + \gamma) = 1/\sqrt{b^2 - a^2}$ 有兩個解；若等於 1，則有一個解；若小於 1，則無解。

> ? **想一想** 為甚麼 $b \le a$ 的情況永遠不可能給出兩個交點？

在這些情況下，$(*)$ 的左方要麼是 x 的嚴格單調（strictly monotonic）函數（$\sinh(x+\gamma)$ 或 $e^{\pm x}$ 的倍數），要麼恆為負數，所以它最多只取值 1 一次。只有先下降再上升的 cosh 形式才能兩次達到 1。

所以兩曲線相交於兩個不同點，當且僅當 $b > a$ 且 $b^2 - a^2 < 1$，即

$$a < b < \sqrt{a^2 + 1}.$$

第 (iv) 部分

由於 $\operatorname{sech} x - (a\tanh x + b) = \operatorname{sech} x\,(1 - a\sinh x - b\cosh x)$ 且 $\operatorname{sech} x \neq 0$，兩曲線相切（touch）恰好發生在 $g(x) = a\sinh x + b\cosh x - 1$ 有重根（repeated zero）之處，即 $g(x) = g'(x) = 0$。在 sinh 及指數的情況下，$g'(x)$ 是 $\cosh(x+\gamma)$ 或 e^x 的倍數，永不為零，所以沒有重根。在 $b > a$ 的情況下，$g(x) = R\cosh(x+\gamma) - 1$ 而 $g'(x) = R\sinh(x+\gamma)$，兩者只在 $x + \gamma = 0$ 且 $R = 1$ 時同時為零。因此兩曲線相切，當且僅當

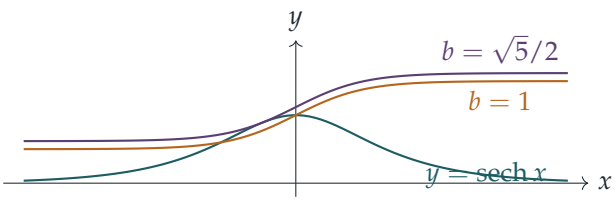
$$b > a \quad \text{且} \quad b^2 - a^2 = 1, \qquad \text{即} \quad b = \sqrt{1 + a^2}.$$

切點的 $x = -\gamma = -\operatorname{artanh}(a/b)$，所以

$$y = \operatorname{sech}\left(\operatorname{artanh}\tfrac{a}{b}\right) = \sqrt{1 - \tfrac{a^2}{b^2}} = \frac{\sqrt{b^2 - a^2}}{b} = \frac{1}{b} = \frac{1}{\sqrt{1 + a^2}},$$

這裏用了 $\operatorname{sech}^2 u = 1 - \tanh^2 u$ 及 $\operatorname{sech} u > 0$。作為驗證，另一條曲線同樣給出 $y = -a \cdot \frac{a}{b} + b = \frac{b^2 - a^2}{b} = \frac{1}{b}$。

下圖顯示 $a = \frac{1}{2}$ 的情況：當 $b = 1$（琥珀色）時有兩個交點；當 $b = \sqrt{5}/2$（梅紅色）時，兩曲線在高度 $2/\sqrt{5}$ 處相切。



驗算

取 $a=\frac{3}{4}$，$b=1$，則 $b^2-a^2=\frac{7}{16}<1$，按 (iii) 應有兩個交點。此時 $\operatorname{arcosh}(4/\sqrt{7})=\ln\bigl((4+3)/\sqrt{7}\bigr)=\ln\sqrt{7}$，而 $\operatorname{artanh}\frac{3}{4}=\frac{1}{2}\ln 7=\ln\sqrt{7}$，所以 (i) 的公式給出 $x=0$ 及 $x=-\ln 7\approx -1.946$。在 $x=0$ 處，兩曲線都經過 $(0,1)$，與預期相符。另外，以電腦對大量隨機的 a 和 b 檢驗 $(*)$，亦確認了每種情況下根的個數，以及相切條件確實給出重根。

常見錯誤

- 忘記邊界情況 $b=\pm a$：此時 cosh 和 sinh 形式都不存在，但該式化為單一個指數函數。
- 在 (iii) 中只寫出 $b^2-a^2<1$ 而漏掉 $b>a$：當 $b<a$ 時兩曲線仍然相交，但只有一個交點。
- 弄錯 γ 的正負號：解是 $\pm\operatorname{arcosh}(\cdot)$ *減去* $\operatorname{artanh}(a/b)$。

STEP 之外

情況的劃分由 B^2-A^2 的正負決定：點 (A,B) 位於由直線 $B=\pm A$ 所分隔的某個區域之內，或位於這些直線上。這正是狹義相對論（special relativity）的幾何：勞侖茲推動（Lorentz boost）是保持 t^2-x^2 不變的雙曲「旋轉」，正如普通旋轉保持 x^2+y^2 不變；而 $B=\pm A$ 擔當光錐（light cone）的角色——那是推動永遠無法令你越過的唯一邊界。想看更多雙曲函數及這類題目，可參閱 Siklos 的 *Advanced Problems in Mathematics*。

詳細解答

STEP III 2016，第 7 題

純數學 • 單位根

> **STEP III 2016，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/16-s3-q7

本題在 *Beyond STEP* 第 5.4 節正多邊形中的距離中有完整的引導式講解：https://beyondstep.keithmcnulty.org/zh-hant/ch05/polygon-distances.html

關鍵思路

距離就是模（modulus），而模的乘積等於乘積的模。若把正多邊形放在阿岡圖（Argand diagram）上，使其頂點為經旋轉的 n 次單位根（roots of unity），那麼根據開首的恆等式，距離的乘積就變成某個適當的 w 所對應的 $|w^n - 1|$。唯一真正要花心思的，是選好坐標：把 O 放在原點，並令 X_0X_1 的垂直平分線（perpendicular bisector）為實軸，這樣 P 就是一個實數。

解答

因式分解

對 $k = 0, 1, \ldots, n-1$，我們有 $(\omega^k)^n = \mathrm{e}^{2\pi i k} = 1$，因此每個 ω^k 都是 $z^n - 1$ 的根。這 n 個數互不相同，因為它們的輻角（argument）$2\pi k/n$ 是 $[0, 2\pi)$ 中互不相同的數。恆等式兩邊都是 n 次的首一多項式（monic polynomial），所以它們的差

$$D(z) = (z-1)(z-\omega)\cdots(z-\omega^{n-1}) - (z^n - 1)$$

的次數（degree）至多為 $n-1$。但 D 在 n 個不同的點 $1, \omega, \ldots, \omega^{n-1}$ 處取值為零，而次數至多為 $n-1$ 的非零多項式最多只有 $n-1$ 個根。因此 $D \equiv 0$，即

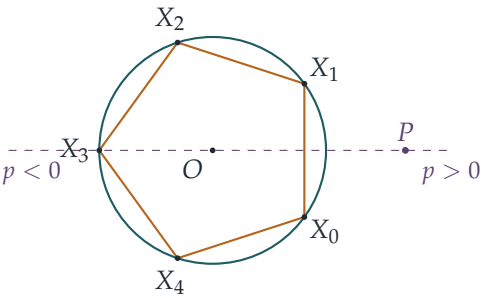
$$\prod_{k=0}^{n-1}(z - \omega^k) = z^n - 1 \quad \text{對所有複數} z.$$

第 (i) 部分

我們假設頂點沿圓周依次標記，因此 X_0 與 X_1 相鄰，且 $\angle X_0OX_1 = 2\pi/n$。旋轉或反射圖形都不會改變任何距離，所以可以把 O 放在原點，並令

$$X_k = \mathrm{e}^{\mathrm{i}(2k-1)\pi/n} = \mathrm{e}^{-\mathrm{i}\pi/n}\,\omega^k, \qquad k = 0, 1, \ldots, n-1.$$

於是 $X_0 = \mathrm{e}^{-\mathrm{i}\pi/n}$ 與 $X_1 = \mathrm{e}^{\mathrm{i}\pi/n}$ 互為關於實軸的反射，所以 X_0X_1 的垂直平分線就是實軸。與 X_0 和 X_1 等距的點恰好就是這條直線上的點，所以 P 是一個實數 p，且 $|OP| = |p|$。X_0X_1 的中點是 $\cos(\pi/n) > 0$，因此 $p > 0$ 表示 P 與邊 X_0X_1 位於 O 的同一側。



（圖中所示為 $n = 5$；當 n 為奇數時，通過 O 的這條直線亦會經過對面的頂點。）

現在，利用 $|ab| = |a||b|$ 及 $|\mathrm{e}^{-\mathrm{i}\pi/n}| = 1$，

$$\begin{aligned}|PX_0|\,|PX_1|\cdots|PX_{n-1}| &= \prod_{k=0}^{n-1}|p - \mathrm{e}^{-\mathrm{i}\pi/n}\omega^k| = \prod_{k=0}^{n-1}|\mathrm{e}^{-\mathrm{i}\pi/n}|\,|p\,\mathrm{e}^{\mathrm{i}\pi/n} - \omega^k| \\ &= \left|\prod_{k=0}^{n-1}\left(p\,\mathrm{e}^{\mathrm{i}\pi/n} - \omega^k\right)\right| = \left|\left(p\,\mathrm{e}^{\mathrm{i}\pi/n}\right)^n - 1\right|,\end{aligned}$$

這裏用了因式分解（factorisation），並取 $z = p\,\mathrm{e}^{\mathrm{i}\pi/n}$。由於 $\mathrm{e}^{\mathrm{i}\pi} = -1$，

$$|PX_0|\,|PX_1|\cdots|PX_{n-1}| = |-p^n - 1| = |p^n + 1|.$$

? **想一想** 為甚麼 p 的正負號對奇數 n 有影響，對偶數 n 卻沒有影響？

若 n 為偶數，則無論 p 的正負號如何，都有 $p^n = |p|^n \geq 0$，所以 $|p^n + 1| = |p|^n + 1$，從而

$$|PX_0| \times |PX_1| \times \cdots \times |PX_{n-1}| = |OP|^n + 1,$$

這正是所求的結果。若 n 為奇數，則 p^n 與 p 同號，需分兩種情況。

- 若 P 位於由 O 出發、經過 X_0X_1 中點的射線（ray）上（$p \geq 0$），則 $p^n = |OP|^n$，乘積為 $|OP|^n + 1$。

- 若 P 位於相反方向的射線上（$p < 0$），對奇數 n 而言，這是由 O 出發、經過頂點 $X_{(n+1)/2}$（位於 $\mathrm{e}^{\mathrm{i}\pi} = -1$）的射線，則 $p^n = -|OP|^n$，乘積為

$$\big|\,|OP|^n - 1\big|.$$

在第二種情況中，模的符號是不可或缺的：當 P 在圓內時，其值為 $1 - |OP|^n$；當 $P = X_{(n+1)/2}$ 時，其值為 0，這是理所當然的，因為此時其中一段距離為零。

第 (ii) 部分

這次旋轉圖形，使 $X_k = \omega^k$，因此 $X_0 = 1$。對 $z \neq 1$，因式分解給出

$$(z-\omega)(z-\omega^2)\cdots(z-\omega^{n-1}) = \frac{z^n - 1}{z - 1} = 1 + z + z^2 + \cdots + z^{n-1}.$$

兩端都是多項式，而且在無限多個點（每個 $z \neq 1$）上取值相同，所以它們對所有 z 都相等，包括 $z = 1$。代入 $z = 1$：

$$(1-\omega)(1-\omega^2)\cdots(1-\omega^{n-1}) = \underbrace{1 + 1 + \cdots + 1}_{n} = n.$$

取模，並利用 $|1 - \omega^k| = |X_0X_k|$，

$$|X_0X_1| \times |X_0X_2| \times \cdots \times |X_0X_{n-1}| = n,$$

這正是所求的結果。

驗算

- 正方形（$n = 4$），取 $P = O$：每段距離都是 1，所以乘積為 $1 = 0^4 + 1$。當 P 很遠時，乘積的表現與 $|OP|^n$ 相若，與公式相符。
- 等邊三角形（$n = 3$），P 在遠側且與 O 相距 2：各距離為 1（到位於 -1 的對面頂點）及兩次 $\sqrt{(-2 - \frac{1}{2})^2 + \frac{3}{4}} = \sqrt{7}$，得出 $7 = 2^3 - 1$。
- 正方形的第 (ii) 部分：$|X_0X_1|\,|X_0X_2|\,|X_0X_3| = \sqrt{2} \times 2 \times \sqrt{2} = 4$。三個結果亦已用電腦就隨機的 n、旋轉及 P 的位置加以驗證。

常見錯誤

- 在第 (i) 部分把 X_0 放在 1：這樣 X_0X_1 的垂直平分線便是傾角為 π/n 的直線，而不是實軸，代數運算會更繁複。應按 P 的位置來選取坐標。
- 在奇數情況且 P 在遠側時忘記取模：當 P 在圓內時，$|OP|^n - 1$ 是負數。
- 直接把 $z = 1$ 代入 $\dfrac{z^n - 1}{z - 1}$；能夠證明其值為 n 的，是那條多項式恆等式。

STEP 之外

第 (ii) 部分其實是說 $z^n - 1$ 在 $z = 1$ 處的導數（derivative）為 n：一般而言，若 $f(z) = \prod(z - a_k)$，則 $f'(a_0) = \prod_{k \neq 0}(a_0 - a_k)$。這類對單位根所取的乘積，出現於分圓多項式（cyclotomic polynomials）的理論及判別式（discriminant）的計算之中；而第 (i) 部分本質上是羅傑·科茨（Roger Cotes）的一個經典結果，涉及正多邊形對稱軸上一點到各頂點的距離。Needham 的 *Visual Complex Analysis*（《直觀複分析》）是欣賞複數乘法這種幾何意義如何發展的絕佳讀物。

詳細解答

STEP III 2016，第 8 題

純數學 • 函數方程

> **STEP III 2016，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/16-s3-q8

關鍵思路

每條方程都把未知函數在 x 的值，與它在某個經變換後的點的值聯繫起來。如果再施行一次（或第三次）該變換就會回到 x，那麼把變換後的點代入方程，便會得到同一組未知量的第二條方程。這樣你便有了聯立線性方程（simultaneous linear equations），可以像處理普通未知數一樣，消去不需要的值。

解答

第 (i) 部分

方程對所有 x 都成立，因此把 x 換成 $-x$ 後仍然成立：

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

這正是所求的結果。

現在我們有兩條方程，含兩個未知量 $\mathrm{f}(x)$ 和 $\mathrm{f}(-x)$。由第二條得 $\mathrm{f}(-x) = x^2 - (1+x)\mathrm{f}(x)$，代入原方程，得

$$\begin{aligned} \mathrm{f}(x) + (1-x)\big(x^2 - (1+x)\mathrm{f}(x)\big) &= x^2 \\ \mathrm{f}(x)\big(1 - (1-x^2)\big) &= x^2 - (1-x)x^2 \\ x^2\,\mathrm{f}(x) &= x^3. \end{aligned}$$

當 $x \neq 0$ 時，可以除以 x^2，得 $\mathrm{f}(x) = x$。當 $x = 0$ 時，原方程變為 $\mathrm{f}(0) + \mathrm{f}(0) = 0$，所以同樣有 $\mathrm{f}(0) = 0$。因此

$$\mathrm{f}(x) = x \quad \text{對所有} x.$$

驗證：$x+(1-x)(-x)=x-x+x^2=x^2$，符合要求。

想一想 為甚麼題目堅持要我們驗證答案？

我們的論證只說明了：如果有函數滿足該方程，它必定是 $\mathrm{f}(x)=x$。它並沒有說明這樣的函數存在：運算過程中可能丟失了資訊，那樣的話便沒有任何函數可行。代回原方程便填補了這個缺口。

第 (ii) 部分

首先注意，對每個 $x\neq 1$ 都有 $\mathrm{K}(x)\neq 1$，因為 $\frac{x+1}{x-1}=1$ 需要 $x+1=x-1$。所以 $\mathrm{K}(\mathrm{K}(x))$ 有定義；把分子分母同乘以 $x-1$，

$$\mathrm{K}(\mathrm{K}(x))=\frac{\dfrac{x+1}{x-1}+1}{\dfrac{x+1}{x-1}-1}=\frac{(x+1)+(x-1)}{(x+1)-(x-1)}=\frac{2x}{2}=x.$$

g 的方程是 $\mathrm{g}(x)+x\,\mathrm{g}(\mathrm{K}(x))=x$。由於 $\mathrm{K}(x)\neq 1$，我們可以把 x 換成 $\mathrm{K}(x)$；又因 $\mathrm{K}(\mathrm{K}(x))=x$，得

$$\mathrm{g}(\mathrm{K}(x))+\mathrm{K}(x)\,\mathrm{g}(x)=\mathrm{K}(x).$$

以 K 表示 $\mathrm{K}(x)$。於是 $\mathrm{g}(K)=K-K\,\mathrm{g}(x)$，代入原方程，

$$\mathrm{g}(x)+x\big(K-K\,\mathrm{g}(x)\big)=x \quad\Longrightarrow\quad (1-xK)\,\mathrm{g}(x)=x-xK.$$

而

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

由於 $x^2+1>0$，因子 $1-xK$ 永不為零，所以可以相除：

$$\mathrm{g}(x)=\frac{-2x/(x-1)}{-(x^2+1)/(x-1)}=\frac{2x}{x^2+1}\qquad (x\neq 1),$$

符合所求。（可以驗證這個函數確實滿足方程；例如在 $x=2$ 時，$\mathrm{K}(2)=3$，而 $\frac{4}{5}+2\cdot\frac{6}{10}=2$。）

第 (iii) 部分

設 $\mathrm{M}(x)=\dfrac{1}{1-x}$。方程為

$$\mathrm{h}(x)+\mathrm{h}(\mathrm{M}(x))=1-x-\mathrm{M}(x),$$

而右方只用 x 和 $\mathrm{M}(x)$ 表示，這是個好兆頭。

這次 M 不是自身的逆（inverse），所以要看重複施行的結果。對 $x \neq 0, 1$：

$$\mathrm{M}(\mathrm{M}(x)) = \frac{1}{1-\frac{1}{1-x}} = \frac{1-x}{-x} = \frac{x-1}{x}, \qquad \mathrm{M}(\mathrm{M}(\mathrm{M}(x))) = \frac{1}{1-\frac{x-1}{x}} = \frac{x}{1} = x.$$

所以 M 的週期（period）為三。我們還需要 M 讓我們留在定義域（domain）之內：$\mathrm{M}(x)$ 永不為 0，而 $\mathrm{M}(x) = 1$ 只在 $x = 0$ 時發生，因此若 $x \neq 0, 1$，則 $\mathrm{M}(x) \neq 0, 1$ 亦成立。

想一想 施行兩次 M 不會回到 x。我們怎樣才能得到足夠的方程？

利用循環中的全部三個點。記 $a = x$，$b = \mathrm{M}(x)$，$c = \mathrm{M}(b) = \frac{x-1}{x}$，於是 $\mathrm{M}(c) = a$。依次在 a、b、c 應用方程：

$$\begin{aligned} \mathrm{h}(a) + \mathrm{h}(b) &= 1 - a - b, \\ \mathrm{h}(b) + \mathrm{h}(c) &= 1 - b - c, \\ \mathrm{h}(c) + \mathrm{h}(a) &= 1 - c - a. \end{aligned}$$

三式相加再除以二，

$$\mathrm{h}(a) + \mathrm{h}(b) + \mathrm{h}(c) = \tfrac{3}{2} - (a + b + c).$$

減去第二條方程，得

$$\mathrm{h}(a) = \tfrac{3}{2} - a - b - c - (1 - b - c) = \tfrac{1}{2} - a.$$

因此

$$\mathrm{h}(x) = \tfrac{1}{2} - x \qquad (x \neq 0,\ x \neq 1).$$

驗證：$\left(\frac{1}{2} - x\right) + \left(\frac{1}{2} - \frac{1}{1-x}\right) = 1 - x - \frac{1}{1-x}$，符合要求。

驗算

- 在 (i) 中，試 $x = 3$：$3 + (-2)(-3) = 9 = 3^2$。
- 在 (ii) 中，試 $x = -1$：$\mathrm{K}(-1) = 0$，而 $\mathrm{g}(-1) + (-1)\,\mathrm{g}(0) = -1 - 0 = -1$。此外 $\mathrm{K}(0) = -1$，證實了週期為二。
- 在 (iii) 中，試 $x = 2$：循環為 $2 \to -1 \to \frac{1}{2} \to 2$，而 $\mathrm{h}(2) + \mathrm{h}(-1) = -\frac{3}{2} + \frac{3}{2} = 0$，同時 $1 - 2 - (-1) = 0$。

三個函數亦已經由電腦以精確算術，在兩千個隨機有理數值上逐一檢驗每條方程，全部確認無誤。

常見錯誤

- 在 (i) 中除以 x^2 時忘記了 $x = 0$，這情況需要另作（只需一行的）論證。
- 把 $\mathrm{K}(x)$ 或 $\mathrm{M}(x)$ 代入 x 時，沒有檢查它是否仍在方程成立的定義域之內。
- 在 (iii) 中只代入一次，結果卡在三個未知量上：你需要整個循環。

STEP 之外

映射 $x \mapsto -x$、K 和 M 都是莫比烏斯變換（Möbius transformations）$x \mapsto \frac{ax+b}{cx+d}$，它們在複合（composition）運算下構成一個羣（group）。有限階（finite order）的變換生成一個有限循環羣（cyclic group），而這裏的方法就是對該羣求和（或消元），這個想法在羣表示論（group representation theory）中會再次出現。莫比烏斯映射與 2×2 矩陣相對應，你可以透過把一個矩陣自乘三次，得出 $\mathrm{M}^3 = $ 恆等變換。Needham 的 *Visual Complex Analysis*（《直觀複分析》）對莫比烏斯變換有精彩的論述。

詳細解答

STEP III 2016，第 9 題

力學 • 彈簧與簡諧運動

> **STEP III 2016，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/16-s3-q9

關鍵思路

把原點放在三角形的中心，並利用對稱性（symmetry）。平衡位置（equilibrium position）就是中心，在那裏三條相等的張力（tension）互相抵消。當質點沿着直線向 P 移動時，彈簧 QX 和 RX 始終互為鏡像，因此它們的側向拉力互相抵消，運動保持在該直線上。求 RX 需要用到餘弦定律（cosine rule）。至於週期，我們先寫出準確的恢復力（restoring force），然後只保留對小位移為線性的項。

解答

我們把 X 視為光滑桌面上的質點，把彈簧視為遵守虎克定律（Hooke's law）的輕彈簧，因此長度為 r 的彈簧的張力為 $\frac{\lambda}{l}(r-l)$。

平衡

設 O 為三角形的中心。每個頂點與 O 的距離為

$$c = \frac{2}{3} \times (\text{高}) = \frac{2}{3} \cdot \sqrt{3}\, a = \frac{2a}{\sqrt{3}}$$

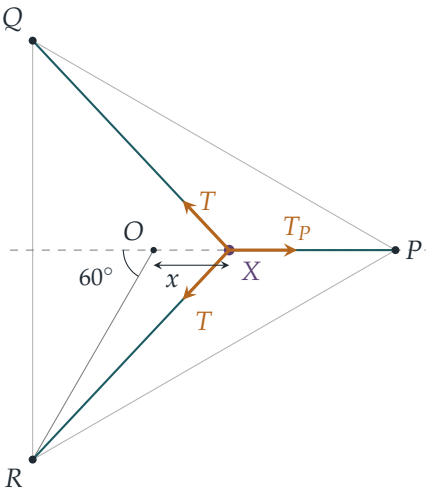
而直線 OP、OQ、OR 兩兩之間的夾角為 120°。當 X 在 O 時，每條彈簧的長度都是 c，所以三條張力相等。三個大小相等、兩兩夾角為 120° 的力的合力為零，所以 O 是一個平衡位置。我們使用的正是這個平衡位置。由於 $l < \frac{2}{\sqrt{3}}a$，每條彈簧都被拉伸，伸長量（extension）為

$$\frac{2}{\sqrt{3}}\, a - l,$$

得證。

RX 中的張力

取 x 軸沿 OP，以 P 的方向為正方向。設 X 在由 O 向 P 方向位移 x 之處。那麼 $OX = x$、$OR = c$，而 $\angle XOR = 180° - 60° = 120°$。



在三角形 XOR 中應用餘弦定律，

$$RX^2 = c^2 + x^2 - 2cx\cos 120° = c^2 + cx + x^2 = \frac{4a^2}{3} + \frac{2ax}{\sqrt{3}} + x^2.$$

因此由虎克定律，RX 中的張力為

$$T = \frac{\lambda}{l}\left(\sqrt{\frac{4a^2}{3} + \frac{2ax}{\sqrt{3}} + x^2} - l\right),$$

得證。由關於軸 OP 的對稱性，$QX = RX$，所以 QX 的張力同樣是 T。

運動方程

記 $r = RX$。彈簧 PX 的長度為 $c - x$，所以它的張力為 $T_P = \frac{\lambda}{l}(c - x - l)$，方向指向 P（即 $+x$ 方向）。

點 R 的 x 坐標為 $-c\cos 60° = -\frac{c}{2}$，所以由 X 到 R 的位移的 x 分量為 $-\left(x + \frac{c}{2}\right)$。因此沿 XR 的力 T 的 x 分量為 $-T\,\dfrac{x + c/2}{r}$。來自 QX 的力有相同的 x 分量，而兩個垂直分量由對稱性互相抵消。所以沒有橫越直線 OP 的力，X 沿該直線運動。

? **想一想** 怎樣從這個平方根得出一條線性方程？

先寫出準確的力，再對小 x 作展開。沿 OP 分解，

$$m\ddot{x} = T_P - 2T\,\frac{x+c/2}{r} = \frac{\lambda}{l}\left[(c-x-l) - (2x+c)\left(1-\frac{l}{r}\right)\right]$$
$$= \frac{\lambda}{l}\left[-3x - l + \frac{l(2x+c)}{r}\right].$$

這條方程是準確的。現在 $r^2 = c^2\left(1+\frac{x}{c}+\frac{x^2}{c^2}\right)$。利用二項式展開（binomial expansion），並捨去 x^2 及更高次的項，

$$\frac{1}{r} \approx \frac{1}{c}\left(1-\frac{x}{2c}\right), \qquad \frac{2x+c}{r} \approx \left(1+\frac{2x}{c}\right)\left(1-\frac{x}{2c}\right) \approx 1+\frac{3x}{2c}.$$

代入，

$$m\ddot{x} \approx \frac{\lambda}{l}\left[-3x - l + l + \frac{3lx}{2c}\right] = -\frac{3\lambda}{l}\left(1-\frac{l}{2c}\right)x.$$

由於 $\frac{l}{2c} = \frac{\sqrt{3}l}{4a}$，

$$\ddot{x} \approx -\omega^2 x, \qquad \omega^2 = \frac{3\lambda(4a-\sqrt{3}\,l)}{4mla}.$$

由於 $\sqrt{3}\,l < 2a < 4a$，這個 ω^2 是正數。所以這個力是恢復力，而對於小 x，質點繞 O 作近似的簡諧運動（simple harmonic motion），週期為

$$\frac{2\pi}{\omega} = 2\pi\sqrt{\frac{4mla}{3(4a-\sqrt{3}\,l)\lambda}},$$

得證。

驗算

- 當 $x = 0$ 時，準確的力為 $\frac{\lambda}{l}\left[-l+\frac{lc}{c}\right] = 0$，與最初求得的平衡位置一致。
- 單位：λ 是力，l、a 是長度，所以 $\frac{mla}{a\lambda}$ 的單位為 $\mathrm{kg\,m/N} = \mathrm{s}^2$，而週期是時間。
- 求 ω^2 的另一途徑。對於小位移，一條彈簧因拉伸而貢獻勁度（stiffness）$\frac{\lambda}{l}\cos^2\phi$，因轉向而貢獻 $\frac{T_0}{c}\sin^2\phi$，其中 ϕ 是彈簧與 OP 的夾角，$T_0 = \frac{\lambda}{l}(c-l)$ 是平衡時的張力。取 $\phi = 0°, 60°, 60°$，總勁度為 $\frac{3}{2}\cdot\frac{\lambda}{l} + \frac{3}{2}\cdot\frac{T_0}{c} = \frac{3\lambda}{l}\left(1-\frac{l}{2c}\right)$，與之前相同。
- 取 $a = \lambda = m = 1$ 及 $l = 0.8$，公式給出的週期約為 2.838。以電腦模擬完整的二維運動，從向 P 方向的小位移開始，得出的數值在四位有效數字內相同。

常見錯誤

- 用了 $\cos 120° = +\frac{1}{2}$，或在餘弦定律中用錯角。這會在平方根內得出 $-\frac{2ax}{\sqrt{3}}$。

- 把 QX 和 RX 中的張力當作常數。它們方向的改變和大小的改變都會貢獻對 x 為線性的項。
- 求中心到頂點的距離時，把三角形的邊長寫成 a 而不是 $2a$。

STEP 之外

在任何穩定平衡位置附近，力對位移都近似為線性。這裏的恢復力是彈性勢能的梯度（gradient），而 $m\omega^2$ 是它的二階導數（second derivative）。在二維情況下，這個二階導數變成一個 2×2 矩陣（黑森矩陣（Hessian）），而它的特徵向量（eigenvector）就是小振動的簡正模（normal modes）。由於三重對稱性，這個系統繞 O 在每個方向的振動週期都相同，而不只是沿 OP 的方向。Kleppner 與 Kolenkow 的 *An Introduction to Mechanics*（《力學導論》）講述了小振動和耦合系統。

詳細解答

STEP III 2016，第 10 題

力學 • 平面上的圓周運動

> **STEP III 2016，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/16-s3-q10

關鍵思路

只要繩子保持拉緊而質點留在斜面上，P 與 A 的距離恆為 a，且 P 位於斜面上，因此它在斜面內的一個圓上運動，圓心是由 A 向斜面所作垂線的垂足。在斜面之內，這不過是在減弱了的重力 $g\sin\alpha$ 下的鉛垂圓周運動；在垂直於斜面的方向，繩子把 P 拉離斜面，而法向反作用力（normal reaction）必須保持為正。所以我們先求張力（tension），再求反作用力，然後問兩者各自在何處最小。

解答

建立模型

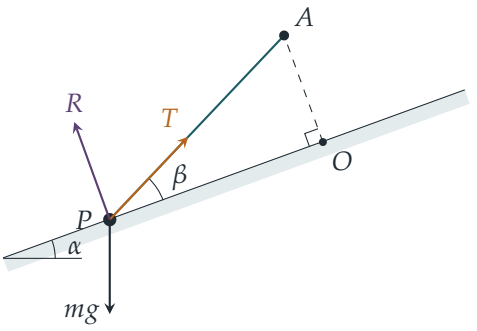
把 P 視為質點，繩子視為輕而不可伸長（inextensible），斜面視為光滑（smooth）。設 O 為由 A 向斜面所作垂線的垂足。由於繩長為 a，且與斜面成角 β，

$$AO = a\sin\beta, \qquad OP = a\cos\beta.$$

在平衡（equilibrium）狀態下，除重力外，唯一有沿斜面分量的力是張力，因此繩子必須沿着一條最大斜率線（line of greatest slope），P 位於 O 沿斜面正下方。沿斜面及垂直於斜面方向分解力（張力 T_0，反作用力 R_0）：

$$T_0\cos\beta = mg\sin\alpha, \qquad R_0 + T_0\sin\beta = mg\cos\alpha,$$

所以 $R_0 = mg(\cos\alpha - \sin\alpha\tan\beta) = mg\cos(\alpha+\beta)/\cos\beta$。因此，平衡狀態要存在，就必須有 $\cos(\alpha+\beta) > 0$。



上圖是開始時經過一條最大斜率線的橫切面。AO 垂直於斜面。

施加衝量（impulse）之後，只要繩子保持拉緊而 P 留在斜面上，就有 $AP = a$ 及 $AO = a\sin\beta$，所以由畢氏定理（Pythagoras' theorem），$OP = a\cos\beta$ 恆成立：P 在斜面內一個以 O 為圓心、半徑為

$$r = a\cos\beta$$

的圓上運動。繩子與斜面所成的角恆為 β，而繩子在斜面上的投影由 P 指向 O。衝量是水平的，且平行於斜面，因此它在這個圓的最低點與圓相切。

設 θ 為 OP 與經過 O 向下的最大斜率線所成的角，於是開始時 $\theta = 0$；又設 v 為速率。重量沿斜面向下（在斜面內）的分量為 $mg\sin\alpha$，垂直壓入斜面的分量為 $mg\cos\alpha$。

垂直於斜面方向，P 沒有加速度，所以

$$R + T\sin\beta = mg\cos\alpha. \tag{1}$$

在斜面內指向 O 的方向，加速度為 v^2/r。張力貢獻 $T\cos\beta$，而沿斜面向下的重量分量貢獻 $-mg\sin\alpha\cos\theta$：

$$T\cos\beta - mg\sin\alpha\cos\theta = \frac{mv^2}{a\cos\beta}. \tag{2}$$

能量：反作用力和張力都不做功（兩者各自垂直於速度），而 P 沿斜面上升的距離為 $r(1-\cos\theta)$，即上升高度為 $r(1-\cos\theta)\sin\alpha$。因此

$$v^2 = u^2 - 2ga\cos\beta\sin\alpha\,(1-\cos\theta). \tag{3}$$

不會立即離開斜面

剛施加衝量後，$\theta = 0$ 且 $v = u$。由 (2) 得 $T\cos\beta = mg\sin\alpha + \dfrac{mu^2}{a\cos\beta}$，再由 (1) 得

$$\begin{aligned} R &= mg\cos\alpha - \tan\beta\left(mg\sin\alpha + \frac{mu^2}{a\cos\beta}\right) \\ &= \frac{m}{\cos\beta}\left(g\cos\alpha\cos\beta - g\sin\alpha\sin\beta\right) - \frac{mu^2\tan\beta}{a\cos\beta} \\ &= \frac{m}{a\cos\beta}\Big(ag\cos(\alpha+\beta) - u^2\tan\beta\Big). \end{aligned}$$

由於 $\cos\beta > 0$，若

$$ag\cos(\alpha+\beta) > u^2\tan\beta,$$

則反作用力為正，質點不會立即離開斜面，此即所求。

完整的圓周

現在把 (3) 代入 (2)，得出一般位置的張力：

$$T\cos\beta = \frac{mu^2}{a\cos\beta} - 2mg\sin\alpha + 3mg\sin\alpha\cos\theta,$$

於是由 (1)，

$$R = mg\cos\alpha - \tan\beta\left(\frac{mu^2}{a\cos\beta} - 2mg\sin\alpha + 3mg\sin\alpha\cos\theta\right).$$

想一想 質點最有可能在圓上哪處離開斜面：頂部還是底部？

從鉛垂圓周運動得來的直覺會說「頂部」，但在這裏，頂部是繩子的危險位置，而不是斜面的。繩子以 $T\sin\beta$ 的力把 P 拉離斜面，而張力在底部最大，因為 P 在那裏速度最快，而斜面內的重力分量也把它拉離 O。事實上，R 在 $\cos\theta$ 最大時最小，即在 $\theta = 0$；而 T 在 $\cos\theta = -1$，即頂部時最小。所以要保持接觸地走完整個圓周，我們需要兩個條件。

- **底部保持接觸**。在 $\theta = 0$ 時 $R > 0$，這正是第一部分的條件：$u^2\tan\beta < ag\cos(\alpha+\beta)$。
- **頂部繩子拉緊**。在 $\theta = \pi$ 時 $T \geq 0$：$\dfrac{u^2}{a\cos\beta} - 5g\sin\alpha \geq 0$，即 $u^2 \geq 5ag\sin\alpha\cos\beta$。(由 (3)，這同時保證頂部有 $v^2 > 0$。)

把第二個條件乘以 $\tan\beta > 0$，可見兩者能同時成立，僅當

$$5ag\sin\alpha\cos\beta\tan\beta \leq u^2\tan\beta < ag\cos(\alpha+\beta),$$

即

$$5\sin\alpha\sin\beta < \cos\alpha\cos\beta - \sin\alpha\sin\beta \iff 6\sin\alpha\sin\beta < \cos\alpha\cos\beta.$$

兩邊除以 $\cos\alpha\cos\beta > 0$，得到必要條件

$$6\tan\alpha\tan\beta < 1.$$

（這條件也是充分的：若它成立，則任何滿足 $5ag\sin\alpha\cos\beta \le u^2 < ag\cos(\alpha+\beta)/\tan\beta$ 的 u 都可行，因為此時 R 和 T 在圓上每一處都為正。）

驗算

- 當 $u = 0$ 時，第一部分的反作用力變成 $mg\cos(\alpha+\beta)/\cos\beta$，與平衡值 R_0 吻合。
- 若 $\beta \to 0$（繩子幾乎沿着斜面），繩子無法把質點提起，而條件 $6\tan\alpha\tan\beta < 1$ 自動成立，理應如此。
- 取 $\alpha = 30°$、$\beta = 5°$（於是 $6\tan\alpha\tan\beta \approx 0.30$）、$a = 1\,\mathrm{m}$ 及 $g = 9.8\,\mathrm{m\,s^{-2}}$，容許範圍是 $24.4 \le u^2 < 91.8$（單位為 $\mathrm{m^2\,s^{-2}}$）。取 $u^2 = 50$，則最小反作用力約為 $0.37mg$，出現在底部；最小張力約為 $2.63mg$，出現在頂部。我們用電腦在三維空間中沿整個圓周求解牛頓第二定律（Newton's second law），證實 R 和 T 保持為正，數值與上述相符。

常見錯誤

- 把整條繩長 a 當作圓的半徑：半徑是 $a\cos\beta$。
- 以為質點最有可能在圓的頂部離開斜面。反作用力其實在底部最小。
- 在能量方程中用 g 而不是 $g\sin\alpha$：質點在鉛垂方向只上升 $r(1-\cos\theta)\sin\alpha$。

STEP 之外

這是約束（constraints）下的運動：斜面和繩子各自提供一個反作用力，其大小恰好是維持約束所需的大小，而運動只有在這些力的正負號正確時才成立。在大學一年級的課程中，你會學到如何有系統地處理這類問題，就是利用拉格朗日力學（Lagrangian mechanics）消去約束，之後再求回反作用力。Kleppner 與 Kolenkow 的 *An Introduction to Mechanics*（《力學導論》）收錄了許多正是這類「何時失去接觸？」的問題。

詳細解答

STEP III 2016，第 11 題

力學 • 功率與阻力

STEP III 2016，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：

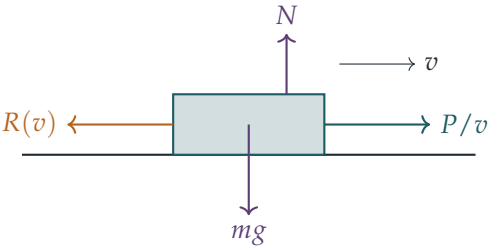
step.maths.org/questions/16-s3-q11

關鍵思路

以恆定功率（power）P 運作的引擎，在速率 v 時產生的牽引力為 P/v。由於我們要求的是以速率表示的距離，而非以時間表示，所以把加速度寫成 $v\frac{\mathrm{d}v}{\mathrm{d}x}$。這樣，牛頓第二定律（Newton's second law）就變成一條關於 x 與 v 的可分離變數微分方程（separable equation）。在 $4U$ 時加速度為零這個條件，可以把阻力常數以 P 和 U 表示出來，之後每段距離都歸結為一個定積分。

解答

我們把汽車模擬為一個質量為 m 的質點，沿水平直線運動，向前的牽引力為 P/v，向後的阻力為 $R(v)$。在鉛垂方向，重量與法向反作用力互相抵消，此後不再起任何作用。



利用 $a = v\frac{\mathrm{d}v}{\mathrm{d}x}$，水平方向的運動方程為

$$mv\frac{\mathrm{d}v}{\mathrm{d}x} = \frac{P}{v} - R(v), \qquad \text{所以} \qquad \frac{\mathrm{d}x}{\mathrm{d}v} = \frac{mv^2}{P - vR(v)}.$$

在兩個模型中，$vR(v)$ 都隨 v 遞增，並在 $v = 4U$ 時等於 P。因此當 $U \le v \le 2U$ 時分母

為正，汽車在整個過程中確實一直在加速，即 x 隨 v 遞增，而所求距離為

$$X = \int_U^{2U} \frac{mv^2}{P - vR(v)}\,\mathrm{d}v.$$

第 (i) 部分

設 $R = kv$。在 $4U$ 時加速度為零，得 $\dfrac{P}{4U} = 4kU$，所以 $k = \dfrac{P}{16U^2}$。於是

$$P - kv^2 = k\,(16U^2 - v^2), \qquad \frac{m}{k} = \frac{16mU^2}{P} = \frac{1}{\lambda U}.$$

因此

$$\lambda X_1 = \frac{1}{U}\int_U^{2U} \frac{v^2}{16U^2 - v^2}\,\mathrm{d}v = \int_1^2 \frac{w^2}{16 - w^2}\,\mathrm{d}w,$$

其中代入了 $v = Uw$。把被積函數寫成 $\dfrac{w^2}{16 - w^2} = -1 + \dfrac{16}{16 - w^2} = -1 + \dfrac{2}{4 - w} + \dfrac{2}{4 + w}$，得

$$\begin{aligned}\lambda X_1 &= \Big[-w + 2\ln\frac{4 + w}{4 - w}\Big]_1^2 \\ &= -1 + 2\ln 3 - 2\ln\tfrac{5}{3} = 2\ln\tfrac{9}{5} - 1,\end{aligned}$$

此即所求。

第 (ii) 部分

現設 $R = cv^2$。在 $4U$ 時加速度為零，得 $\dfrac{P}{4U} = 16cU^2$，所以 $c = \dfrac{P}{64U^3}$。於是

$$P - cv^3 = c\,(64U^3 - v^3), \qquad \frac{m}{c} = \frac{64mU^3}{P} = \frac{4}{\lambda}.$$

想一想 這裏有沒有簡潔的原函數？還是需要再用一次部分分式？

分子 v^2 除了相差一個常數倍外，正是分母 $64U^3 - v^3$ 的導數。所以這個積分可以直接得出對數：

$$\begin{aligned}\lambda X_2 &= 4\int_U^{2U} \frac{v^2}{64U^3 - v^3}\,\mathrm{d}v = 4\Big[-\tfrac{1}{3}\ln\big(64U^3 - v^3\big)\Big]_U^{2U} \\ &= \tfrac{4}{3}\ln\frac{63U^3}{56U^3} = \tfrac{4}{3}\ln\tfrac{9}{8},\end{aligned}$$

此即所求。

第 (iii) 部分

由於 $\lambda > 0$，$X_1 > X_2$ 當且僅當 $\lambda X_1 - \lambda X_2 > 0$。題目只給出 $\ln 24$ 和 $\ln 5$ 的界，所以我們把差改寫成以這兩者表示：

$$\begin{aligned}\lambda(X_1 - X_2) &= 2\ln 9 - 2\ln 5 - 1 - \tfrac{4}{3}\ln 9 + \tfrac{4}{3}\ln 8 \\ &= \tfrac{4}{3}\ln 3 + 4\ln 2 - 2\ln 5 - 1 \\ &= \tfrac{4}{3}(\ln 3 + 3\ln 2) - 2\ln 5 - 1 = \tfrac{4}{3}\ln 24 - 2\ln 5 - 1.\end{aligned}$$

? **想一想** 每個界應該怎樣用，是用上界還是下界？

要證明差為正，我們需要一個下界，所以對 $\ln 24$（帶正號）用其下界，對 $\ln 5$（帶負號）用其上界：

$$\lambda(X_1 - X_2) > \tfrac{4}{3}(3.17) - 2(1.61) - 1 = 4.22\dot{6} - 3.22 - 1 = 0.00\dot{6} > 0.$$

因此 $X_1 > X_2$：汽車由 U 加速至 $2U$ 期間，阻力與速率成正比時所走的距離，比阻力與速率平方成正比時更遠。

這在物理上是合理的。兩種阻力在 $v = 4U$ 時相等，但當 $v < 4U$ 時，有 $cv^2 = kv \cdot \frac{v}{4U} < kv$。所以在所考慮的速率範圍內，平方定律的阻力較小，汽車加速得更猛，因而更快到達 $2U$。

驗算

- 數值上，$\lambda X_1 = 2\ln 1.8 - 1 \approx 0.1756$，$\lambda X_2 = \frac{4}{3}\ln 1.125 \approx 0.1570$，所以差約為 0.0185，是正數，與所得結果一致。用電腦由 $v = U$ 開始按時間逐步模擬運動方程，可把兩個數值重現至小數點後五位。
- 單位：$P/(mU^3)$ 的單位是 $\mathrm{W}/(\mathrm{kg\,m^3\,s^{-3}}) = \mathrm{m^{-1}}$，所以 λX 是無量綱的，理應如此。
- 第 (iii) 部分末尾的物理論證，獨立地預測了同樣的大小次序。

常見錯誤

- 使用 $a = \frac{\mathrm{d}v}{\mathrm{d}t}$，這樣得到的是時間而不是距離。這裏需要的是 $v\,\frac{\mathrm{d}v}{\mathrm{d}x}$。
- 在第 (iii) 部分把界用反了。用 3.18 和 1.60 只會得到一個上界，對差的正負號毫無證明作用。
- 忘記餘量極小（$0.00\dot{6}$），以致把 $\frac{4}{3}(3.17)$ 四捨五入為 4.2，結果便證不出來。

STEP 之外

把加速度寫成 $a = v\,\mathrm{d}v/\mathrm{d}x$，是邁向大學力學中能量方法（energy method）的第一步。把運動方程乘以 v，便得到動能的變化率，它等於輸入功率減去耗散功率。像 $4U$ 這種兩者平衡的速率，是一階微分方程的平衡點（equilibria），而研究解如何趨近平衡點，正是動力系統（dynamical systems）的起點。Kleppner 與 Kolenkow 的 *An Introduction to Mechanics*（《力學導論》）便以這種風格討論功、功率及與速度相關的力。

詳細解答

STEP III 2016，第 12 題

概率與統計 • 切比雪夫不等式

> **STEP III 2016，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/16-s3-q12

關鍵思路

切比雪夫不等式（Chebyshev's inequality）給出的是遠離均值（mean）的概率的上界，所以要證明某個量接近均值，就要從補事件（complement）入手。每一部分的工作都是：找出合適的隨機變量（random variable），求出它的均值和標準差（standard deviation），再選取 k，使 $k\sigma$ 恰好等於我們關心的距離。在第 (ii) 部分，和式 $\sum n^r/r!$ 與 e^n 並列出現，正是泊松分佈（Poisson distribution）的標誌。

解答

第 (i) 部分

設 X 為 $100n$ 次拋擲中得到正面的次數。假設各次拋擲互相獨立（independent），而每次得到正面的概率為 0.2，則 $X \sim \mathrm{B}(100n, 0.2)$，因此

$$\mu = 100n \times 0.2 = 20n, \qquad \sigma^2 = 100n \times 0.2 \times 0.8 = 16n, \qquad \sigma = 4\sqrt{n}.$$

事件 $16n \le X \le 24n$ 正好就是 $|X - 20n| \le 4n$，所以

$$\alpha = \mathrm{P}\big(|X-\mu| \le 4n\big) = 1 - \mathrm{P}\big(|X-\mu| > 4n\big).$$

我們希望 $k\sigma = 4n$，即 $k \cdot 4\sqrt{n} = 4n$，故 $k = \sqrt{n}$，這是正數。由切比雪夫不等式得

$$\mathrm{P}\big(|X-\mu| > 4n\big) = \mathrm{P}\big(|X-\mu| > \sqrt{n}\,\sigma\big) \le \frac{1}{(\sqrt{n})^2} = \frac{1}{n},$$

因此

$$\alpha \ge 1 - \frac{1}{n},$$

得證。（條件 $n > 1$ 只是為了讓這個界有實質內容：當 $n = 1$ 時，它只說 $\alpha \ge 0$。）

第 (ii) 部分

想一想 哪一個隨機變量的概率形如 $n^r/r!$？

若 $Y \sim \mathrm{Po}(n)$，則對 $r = 0, 1, 2, \ldots$ 有 $\mathrm{P}(Y = r) = \mathrm{e}^{-n} n^r / r!$，所以

$$\mathrm{e}^{-n}\left(1 + n + \frac{n^2}{2!} + \cdots + \frac{n^{2n}}{(2n)!}\right) = \mathrm{P}(Y \le 2n).$$

所求不等式兩邊除以 $\mathrm{e}^n > 0$ 後，便等價於

$$\mathrm{P}(Y \le 2n) \ge 1 - \frac{1}{n}.$$

泊松分佈的均值和方差（variance）都是 n，所以 $\mu = n$，$\sigma = \sqrt{n}$。

想一想 切比雪夫不等式是雙側的，但 $Y \le 2n$ 是單側事件。兩者如何對應起來？

由於 $Y \ge 0$，必有 $Y - n \ge -n$，所以 $Y - n < -n$ 不可能發生。因此

$$|Y - n| > n \iff Y - n > n \iff Y > 2n.$$

取 $k = \sqrt{n}$，使 $k\sigma = n$，由切比雪夫不等式得

$$\mathrm{P}(Y > 2n) = \mathrm{P}\big(|Y - n| > \sqrt{n}\,\sigma\big) \le \frac{1}{n},$$

所以 $\mathrm{P}(Y \le 2n) = 1 - \mathrm{P}(Y > 2n) \ge 1 - \frac{1}{n}$。兩邊乘以 e^n，

$$1 + n + \frac{n^2}{2!} + \cdots + \frac{n^{2n}}{(2n)!} \ge \left(1 - \frac{1}{n}\right)\mathrm{e}^n,$$

這正是所求結果。

驗算

這些界是正確的，但相當寬鬆。當 $n = 2$ 時，第 (i) 部分的準確二項分佈（binomial distribution）概率為 $\alpha \approx 0.868$，而界只是 0.5；當 $n = 10$ 時，$\alpha \approx 0.9986$，而界是 0.9。在第 (ii) 部分取 $n = 2$，左邊為 $1 + 2 + 2 + \frac{4}{3} + \frac{2}{3} = 7$，而 $\frac{1}{2}\mathrm{e}^2 \approx 3.69$。我們已就第 (i) 部分中 n 由 2 至 30、第 (ii) 部分中 n 由 1 至 59 的每一個值，準確驗證了兩個不等式。

常見錯誤

- 選取 k 時，誤用方差 $16n$ 代替標準差 $4\sqrt{n}$。

- 忘記切比雪夫不等式給出的是落在區間以外的概率的上界，因此要用補事件才能得到 α。
- 在第 (ii) 部分沒有解釋為何下尾 $Y < 0$ 沒有任何貢獻。

STEP 之外

切比雪夫不等式是集中不等式（concentration inequality）最簡單的例子；把它應用於 m 次獨立試驗的平均值，便得到弱大數定律（weak law of large numbers）：離散程度以 $1/\sqrt{m}$ 的速度收窄。第 (ii) 部分還告訴我們，指數級數（exponential series）在 $n^{2n}/(2n)!$ 之後的尾項之和至多為 e^n/n——這是一個用概率證明的分析學事實。更精細的工具，例如切爾諾夫界（Chernoff bounds），可以證明真正的概率與 1 的差距是指數級地小。如欲深入了解，可參閱 Grimmett 與 Welsh 的 *Probability: An Introduction*（《概率論導引》）。

詳細解答

STEP III 2016，第 13 題

概率與統計 • 和的峰度

STEP III 2016，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/16-s3-q13

關鍵思路

峯度（kurtosis）只由偏離平均值的差構成，所以任何平移變量的操作都不會改變它；這讓我們可以把每個變量都中心化至 0，然後直接處理普通的矩（moments）。對於正態分佈（Normal distribution），四階矩可用分部積分法（integration by parts）求得，方法是分出一個 x 因子與指數函數配對。對於若干個互相獨立（independent）且平均值為零的變量之和，展開 T^4 再取期望值，凡是含有某個變量的一次方的項都會變成零。這就是第 (ii) 部分的全部內容，而它會直接用於第 (iii) 部分。

解答

平移。 設 $W = X - a$。由期望值的線性性質（linearity of expectation），$\mathrm{E}(W) = \mu - a$，所以

$$W - \mathrm{E}(W) = (X - a) - (\mu - a) = X - \mu.$$

W 偏離其平均值的差，正好等於 X 偏離其平均值的差。因此 W 有相同的方差 $\mathrm{E}\big((X-\mu)^2\big) = \sigma^2$、相同的四階中心矩（central moment）$\mathrm{E}\big((X-\mu)^4\big)$，從而有相同的峯度。特別地，取 $a = \mu$，則每個隨機變量都與其中心化版本 $X - \mu$ 有相同的峯度，而後者的平均值為 0。

第 (i) 部分

設 $X \sim \mathrm{N}(0, \sigma^2)$，其概率密度函數（probability density function）為

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}}\,\mathrm{e}^{-x^2/(2\sigma^2)}, \qquad \text{因此} \qquad f'(x) = -\frac{x}{\sigma^2}\,f(x).$$

由於 $\mu = 0$，我們需要求 $\mathrm{E}(X^4) = \int_{-\infty}^{\infty} x^4 f(x)\,\mathrm{d}x$。

想一想 既然 $\mathrm{e}^{-x^2/(2\sigma^2)}$ 沒有初等的原函數，我們怎樣積分 $x^4\mathrm{e}^{-x^2/(2\sigma^2)}$？

函數 $xf(x)$ 確實有原函數（antiderivative），就是 $-\sigma^2 f(x)$。所以寫成 $x^4 f(x) = x^3 \cdot xf(x)$，並取 $u = x^3$ 及 $\mathrm{d}v = xf(x)\,\mathrm{d}x$、$v = -\sigma^2 f(x)$ 進行分部積分：

$$\int_{-\infty}^{\infty} x^4 f(x)\,\mathrm{d}x = \Big[-\sigma^2 x^3 f(x)\Big]_{-\infty}^{\infty} + 3\sigma^2 \int_{-\infty}^{\infty} x^2 f(x)\,\mathrm{d}x = 3\sigma^2 \int_{-\infty}^{\infty} x^2 f(x)\,\mathrm{d}x.$$

邊界項為零，因為當 $x \to \pm\infty$ 時 $x^3\mathrm{e}^{-x^2/(2\sigma^2)} \to 0$（指數函數衰減的速度比任何冪的增長都快）。以 $u = x$ 重複同一步驟，得

$$\int_{-\infty}^{\infty} x^2 f(x)\,\mathrm{d}x = \Big[-\sigma^2 x f(x)\Big]_{-\infty}^{\infty} + \sigma^2 \int_{-\infty}^{\infty} f(x)\,\mathrm{d}x = \sigma^2,$$

因為 f 是概率密度函數，其積分為 1。（這也證實了方差是 σ^2。）因此 $\mathrm{E}(X^4) = 3\sigma^2 \cdot \sigma^2 = 3\sigma^4$，而

$$\kappa = \frac{3\sigma^4}{\sigma^4} - 3 = 0.$$

由平移的結果，每個正態分佈，無論其平均值是多少，峯度都是 0。

第 (ii) 部分

用多項式定理（multinomial theorem）展開 $T^4 = (Y_1 + \cdots + Y_n)^4$。每一項都是四個因子 Y_r 的乘積，而按照這四個因子如何分配到不同的下標，各項可分為五類：

$$Y_r^4, \qquad 4\,Y_r^3 Y_s, \qquad 6\,Y_r^2 Y_s^2, \qquad 12\,Y_r^2 Y_s Y_t, \qquad 24\,Y_r Y_s Y_t Y_u,$$

其中不同的字母代表不同的下標。係數計算的是排列方式的數目：例如，$Y_r^2 Y_s^2$ 來自從 4 個括號中選出哪 2 個提供 Y_r，共有 $\binom{4}{2} = 6$ 種方法。每個無序對 $\{r, s\}$ 在 $Y_r^2 Y_s^2$ 這一類中只出現一次。

現在取期望值。各個 Y 互相獨立，所以不同 Y 的冪的乘積的期望值，等於各自期望值的乘積。在 $Y_r^3 Y_s$、$Y_r^2 Y_s Y_t$ 和 $Y_r Y_s Y_t Y_u$ 這幾類中，總有某個變量只以一次方出現，而它的因子 $\mathrm{E}(Y_s) = 0$ 令整個期望值為零。例如，

$$\mathrm{E}(Y_r^3 Y_s) = \mathrm{E}(Y_r^3)\,\mathrm{E}(Y_s) = 0.$$

只有第一類和第三類保留下來，而同樣由獨立性，當 $r \neq s$ 時 $\mathrm{E}(Y_r^2 Y_s^2) = \mathrm{E}(Y_r^2)\,\mathrm{E}(Y_s^2)$。把各對寫成 $r < s$，得

$$\mathrm{E}(T^4) = \sum_{r=1}^{n} \mathrm{E}(Y_r^4) + 6 \sum_{r=1}^{n-1} \sum_{s=r+1}^{n} \mathrm{E}(Y_r^2)\,\mathrm{E}(Y_s^2),$$

此即所求。（注意這部分用到了獨立性和平均值為零，但沒有用到各個 Y 同分佈（identically distributed）。）

第 (iii) 部分

設每個 X_r 的平均值為 μ、標準差為 σ，並設 $S = X_1 + \cdots + X_n$。令 $Y_r = X_r - \mu$。各個 Y_r 互相獨立且同分佈，平均值為 0，而

$$\mathrm{E}(Y_r^2) = \sigma^2, \qquad \mathrm{E}(Y_r^4) = (\kappa + 3)\sigma^4,$$

第二式由移項 κ 的定義得出。於是 $T = \sum Y_r = S - n\mu$。由平移的結果，S 與 T 有相同的峯度，所以我們處理平均值為 0 的 T。

由獨立性，$\mathrm{Var}(T) = \sum \mathrm{Var}(Y_r) = n\sigma^2$。第 (ii) 部分中的雙重和共有 $\binom{n}{2} = \frac{1}{2}n(n-1)$ 項，每項都等於 σ^4，所以

$$\begin{aligned}\mathrm{E}(T^4) &= n(\kappa+3)\sigma^4 + 6\cdot\tfrac{1}{2}n(n-1)\sigma^4 \\ &= n\sigma^4\big(\kappa + 3 + 3(n-1)\big) = n\sigma^4(\kappa + 3n).\end{aligned}$$

因此 T 的峯度，亦即 S 的峯度，為

$$\frac{\mathrm{E}(T^4)}{(\mathrm{Var}\,T)^2} - 3 = \frac{n\sigma^4(\kappa+3n)}{n^2\sigma^4} - 3 = \frac{\kappa}{n} + 3 - 3 = \frac{\kappa}{n},$$

此即所求。

驗算

- 取 $Y = \pm 1$，概率各為 $\frac{1}{2}$。則 $\mathrm{E}(Y^4) = \sigma^4 = 1$，所以 $\kappa = 1 - 3 = -2$。當 $n = 2$ 時，和的取值為 $-2, 0, 2$，概率分別為 $\frac{1}{4}, \frac{1}{2}, \frac{1}{4}$，所以 $\mathrm{E}(T^4) = 8$ 及 $\mathrm{Var}\,T = 2$。由此得峯度為 $8/4 - 3 = -1 = \kappa/2$，與預測相符。
- 當 $n = 1$ 時，第 (iii) 部分表示甚麼都不變，理應如此。
- 由於互相獨立的正態變量之和仍是正態變量，第 (iii) 部分取 $\kappa = 0$ 時必須再次得出 0，而事實確是如此。
- 我們用電腦以精確分數驗證了：一個取三個值的變量，其最多六個獨立副本之和的峯度恰好是 κ/n。我們亦以數值方法驗證了，對多個 σ 值，$\int x^4 f(x)\,\mathrm{d}x = 3\sigma^4$。

常見錯誤

- 弄錯 $Y_r^2 Y_s^2$ 的係數：每個無序對的係數是 $\binom{4}{2} = 6$（或每個有序對 3），不是 4 或 12。
- 在使用第 (ii) 部分之前忘記把 X_r 中心化，而第 (ii) 部分要求平均值為 0。
- 除以 σ^4 而不是 $(n\sigma^2)^2$：和的方差是 $n\sigma^2$。

STEP 之外

這裏定義的峯度，是四階累積量（cumulant）除以 σ^4。累積量是矩母函數（moment generating function）的對數展開式中的係數，而對於獨立變量之和，它們直接相加。這就是為甚麼和的標準化四階累積量會像 $1/n$ 那樣縮小，而三階累積量（偏度（skewness））則像 $1/\sqrt{n}$ 那樣縮小。這讓我們定量地窺見中心極限定理（central limit theorem）：和看起來愈來愈像正態分佈，而正態分佈二階以上的累積量全部為零。Grimmett 和 Welsh 的 *Probability: An Introduction*（《概率導論》）在這個程度上講解生成函數和中心極限定理。

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

2016 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。