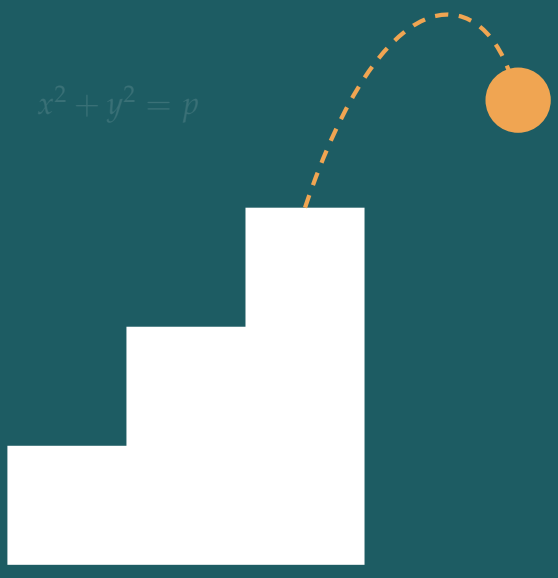
STEP 備試 · 大學數學

Beyond STEP

STEP I, II 及 III

2000 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

2000 年詳細解答

Keith McNulty

Beyond STEP：以歷屆 STEP 試題為門，走進大學數學

2000 年詳細解答

2000 年 STEP 試卷（STEP I, II 及 III）全部 42 題的詳細解答。自樣本試卷以來每一道 STEP 題目的詳細解答、*Beyond STEP* 全書以及各章的單行本，均可在 `beyondstep.keithmcnulty.org` 免費閱讀及下載。

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

由機器從英文翻譯而成；尚未經審閱。

2026 年 11 月首次於網上發表。

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

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

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

目錄

STEP I

第 1 題 對數與首位數字 純數學 . . . 3

第 2 題 二項展開式的係數 純數學 . . . 6

第 3 題 階梯函數的積分 純數學 . . . 10

第 4 題 積分的界限 純數學 . . . 14

第 5 題 三維向量 純數學 . . . 17

第 6 題 直線所圍成的區域 純數學 . . . 20

第 7 題 導數的界限 純數學 . . . 24

第 8 題 絕對值的積分 純數學 . . . 27

第 9 題 加速參考系中的拋體 力學 . . . 31

第 10 題 碰撞與恢復係數 力學 . . . 35

第 11 題 梯子的摩擦力平衡 力學 . . . 38

第 12 題 幾何分佈 概率與統計 . . . 42

第 13 題 條件概率 概率與統計 . . . 45

第 14 題 均勻分佈的矩 概率與統計 . . . 48

STEP II

第 1 題 歐幾里得及分數 純數學 . . . 51

第 2 題 多項式的重根 純數學 . . . 55

第 3 題 誤差傳遞 純數學 . . . 58

第 4 題 直角與反正切恆等式 純數學 . . . 62

第 5 題 最小二乘近似 純數學 . . . 66

第 6 題 半角代換 純數學 . . . 70

第 7 題 直線與圓錐 純數學 . . . 74

第 8 題 可分離變量微分方程 純數學 . . . 78

第 9 題 有空氣阻力的運動 力學 . . . 81
第 10 題 連接滑輪系統 力學 . . . 85
第 11 題 彈性繩與簡諧運動 力學 . . . 89
第 12 題 離態變量之和 概率與統計 . . . 93
第 13 題 捕捉再捕捉估計 概率與統計 . . . 97
第 14 題 中位數的分佈 概率與統計 . . . 101

STEP III

第 1 題 雙曲線的切線 純數學 . . . 105
第 2 題 代換積分法 純數學 . . . 110
第 3 題 複數與幾何 純數學 . . . 113
第 4 題 有理函數的描繪 純數學 . . . 117
第 5 題 線性變換 純數學 . . . 122
第 6 題 四次多項式的因式分解 純數學 . . . 126
第 7 題 不等式與極限 純數學 . . . 130
第 8 題 遞推關係 純數學 . . . 133
第 9 題 彈性斜碰撞 力學 . . . 137
第 10 題 極限摩擦力平衡 力學 . . . 141
第 11 題 粗糙支點的轉動 力學 . . . 145
第 12 題 泊松近似 概率與統計 . . . 149
第 13 題 概率生成函數 概率與統計 . . . 153
第 14 題 協方差與獨立性 概率與統計 . . . 157

詳細解答

STEP I 2000，第 1 題

純數學 • 對數與首位數字

STEP I 2000，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/00-s1-q1

關鍵思路

一個大數 N 的首位數字（leading digit）完全取決於 $\log_{10} N$ 的小數部分（fractional part）。若 $\log_{10} N = k + f$，其中 k 為整數而 $0 \le f < 1$，則 $N = 10^f \times 10^k$，而 N 的首位數字是 d，當且僅當 $\log_{10} d \le f < \log_{10}(d+1)$。因此我們只需要數字 1 至 9 的對數（logarithm），它們都可由 $\log_{10} 2$ 和 $\log_{10} 3$ 求得；此外還要留心我們能夠容許多大的捨入誤差（rounding error）。

解答

以下 log 一律表示 $\log_{10}$。區間 $[0,1)$ 分成九段，每段對應一個可能的首位數字：



小數部分 f 落在標示為 d 的一段內，首位數字就是 d。

第 (i) 部分

由於 $5 = 10/2$ 及 $6 = 2 \times 3$，

$$\log 5 = 1 - \log 2 = 1 - 0.301029996 = 0.698970004 \approx 0.699,$$
$$\log 6 = \log 2 + \log 3 = 0.301029996 + 0.477121255 = 0.778151251 \approx 0.778,$$

兩者均準確至小數點後三位。

現在

$$\log 3^{100} = 100 \log 3 = 47.7121255.$$

題目所給的 $\log 3$ 的值，誤差不超過 5×10^{-10}，所以上式的誤差不超過 5×10^{-8}；必定有 $47.712 < \log 3^{100} < 47.713$。與

$$\log\left(5 \times 10^{47}\right) = 47 + \log 5 \approx 47.699 \quad \text{及} \quad \log\left(6 \times 10^{47}\right) = 47 + \log 6 \approx 47.778,$$

比較，可得

$$\log\left(5 \times 10^{47}\right) < \log 3^{100} < \log\left(6 \times 10^{47}\right).$$

兩邊的餘裕（約 0.013 和 0.066）遠遠大於任何捨入誤差，而 log 是嚴格遞增（strictly increasing）的，所以

$$5 \times 10^{47} < 3^{100} < 6 \times 10^{47},$$

正如所求。因此 3^{100} 是 $5.\ldots \times 10^{47}$：它的首位數字是 $\boxed{5}$（而且它共有 48 位數字）。

第 (ii) 部分

我們沿用同一個想法，取 $\log 2 = 0.301029996$，其誤差至多為 5×10^{-10}。乘以 10^n 會把這個誤差也乘以 10^n，所以我們必須檢查小數部分是否仍然穩妥地留在數線上的同一段之內。

2^{1000}. $\log 2^{1000} = 301.029996$，誤差至多為 5×10^{-7}。小數部分約為 0.030，介乎 $\log 1 = 0$ 與 $\log 2 \approx 0.301$ 之間，所以 $2^{1000} = 1.\ldots \times 10^{301}$，其首位數字是 $\boxed{1}$。

$2^{10\,000}$. $\log 2^{10\,000} = 3010.29996$，誤差至多為 5×10^{-6}。小數部分介乎 0.299955 與 0.299965 之間。

> ? **想一想** 小數部分 0.29996 非常接近 $\log 2 = 0.30103$。我們能肯定首位數字不是 2 嗎？

能夠：兩者相差 $0.30103 - 0.29996 \approx 0.0011$，是最大可能誤差 5×10^{-6} 的二百多倍。所以小數部分確實小於 $\log 2$，而 $10^{3010} < 2^{10\,000} < 2 \times 10^{3010}$。（等價地，$2^{10\,000} < 2 \times 10^{3010}$ 即是說 $2^{9999} < 10^{3010}$，亦即 $\log 2 < 3010/9999 = 0.3010301\ldots$，而這是成立的。）首位數字是 $\boxed{1}$。

$2^{100\,000}$. $\log 2^{100\,000} = 30102.9996$，誤差至多為 5×10^{-5}，所以小數部分介乎 0.99955 與 0.99965 之間。我們需要知道它落在數線上的哪一段，而最頂的一段始於

$$\log 9 = 2 \log 3 = 0.954242510.$$

由於 $0.954 < 0.99955$ 及 $0.99965 < 1$，我們得到 $9 \times 10^{30102} < 2^{100\,000} < 10^{30103}$，所以首位數字是 $\boxed{9}$。

> **想一想** 誤差會不會把小數部分推過 1，令這個數變成 $1.\ldots \times 10^{30103}$？

除非 $100\,000 \log 2$ 的真確值至少是 30103 才會如此，而這需要它有 0.0004 的誤差，是最大誤差 5×10^{-5} 的八倍。所以首位數字是 9，而 $2^{100\,000}$ 共有 30 103 位數字。

驗算

用電腦作精確的整數運算，可證實以上所有結果：3^{100} 有 48 位數字，開首為 515…；2^{1000} 開首為 107…（共 302 位數字）；$2^{10\,000}$ 開首為 199…（共 3011 位數字）；而 $2^{100\,000}$ 開首為 999…（共 30 103 位數字）。再用一個簡單例子檢驗這個方法是否合理：$\log 2^{10} = 3.0103$ 的小數部分為 $0.0103 < \log 2$，與 $2^{10} = 1024$ 吻合。

常見錯誤

- 從對數的整數部分（integer part）讀出首位數字。整數部分只能告訴我們位數（位數比它多一）；首位數字藏在小數部分之中。
- 忘記乘以 100 000 時，$\log 2$ 的捨入誤差也會同樣被放大。在這裏九位小數僅僅足夠，而且值得說明原因。
- 看見 $f = 0.9996 \approx 1$ 便斷定首位數字是 1。小數部分略小於 1，表示這個數略小於十的某個冪，所以它以 9 開首。

STEP 之外

2^n 的首位數字由 $n \log 2$ 的小數部分決定。由於 $\log 2$ 是無理數（irrational），這些小數部分在 $[0, 1)$ 內是等分佈（equidistributed）的（外爾（Weyl）等分佈定理），所以在 2 的冪之中，以數字 d 開首的所佔比例趨向 $\log(1 + 1/d)$：這就是本福特定律（Benford's law），而大約 30% 的 2 的冪以 1 開首。至於無理性與逼近這方面的故事，達文波特（Davenport）的 *The Higher Arithmetic* 是一本十分值得一讀的好書。

詳細解答

STEP I 2000，第 2 題

純數學 • 二項展開式的係數

> **STEP I 2000，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/00-s1-q2

關鍵思路

把兩個帶分式項的二項式（binomial）直接展開，再逐一尋找相配的冪，過程相當繁瑣。較簡潔的做法是先把 x 的負數冪抽出來，令每個括號都變成只含 x 的單一種冪（x^6 或 x^2）的多項式（polynomial）。這樣一來，憑着因式分解（factorisation）$x^6-1=(x^2-1)(x^4+x^2+1)$，便會發現第二個乘積其實就是第一個乘積的化身，所以題目中「由此」（hence）的部分幾乎不費工夫。

解答

第一個展開式

從每個括號中抽出 x 的最低次冪，

$$\left(x^4-\frac{1}{x^2}\right)^5=\left(\frac{x^6-1}{x^2}\right)^5=x^{-10}\,(x^6-1)^5,$$
$$\left(x-\frac{1}{x}\right)^6=\left(\frac{x^2-1}{x}\right)^6=x^{-6}\,(x^2-1)^6.$$

因此整個算式等於 $x^{-16}P(x)$，其中

$$P(x)=(x^6-1)^5\,(x^2-1)^6.$$

所以原展開式中 x^n 的係數（coefficient），就是多項式 $P(x)$ 中 x^{n+16} 的係數。

根據二項式定理（binomial theorem），

$$(x^6-1)^5=\sum_{j=0}^{5}\binom{5}{j}(-1)^{5-j}x^{6j},\qquad (x^2-1)^6=\sum_{k=0}^{6}\binom{6}{k}(-1)^{6-k}x^{2k},$$

故此 $P(x)$ 中 x^m 的係數，就是

$$\binom{5}{j}(-1)^{5-j}\binom{6}{k}(-1)^{6-k}$$

對所有符合 $6j+2k=m$、$0\le j\le 5$、$0\le k\le 6$ 的數對求和的結果。

x^{-12} 的係數。 我們需要 $m=-12+16=4$，即 $6j+2k=4$。由於 $6j\le 4$，必有 $j=0$，所以唯一的數對是 $(j,k)=(0,2)$，得出

$$\binom{5}{0}(-1)^5\cdot\binom{6}{2}(-1)^4=(-1)(15)=-15,$$

正如題目所求。

x^2 的係數。 這次 $m=2+16=18$，所以 $6j+2k=18$，即 $3j+k=9$，而 $k\le 6$。這容許 $j=1,2,3$：

(j,k)	來自 $(x^6-1)^5$	來自 $(x^2-1)^6$	乘積
$(3,0)$	$\binom{5}{3}(-1)^2=10$	$\binom{6}{0}(-1)^6=1$	10
$(2,3)$	$\binom{5}{2}(-1)^3=-10$	$\binom{6}{3}(-1)^3=-20$	200
$(1,6)$	$\binom{5}{1}(-1)^4=5$	$\binom{6}{6}(-1)^0=1$	5

所以 x^2 的係數是 $10+200+5=215$。

> ? **想一想** 為甚麼 $j=0$ 或 $j=4$ 不可能對 x^{18} 有貢獻？

若 $j=0$，便需要 $k=9$；若 $j=4$，便需要 $k=-3$；兩者都不在 $0\le k\le 6$ 的範圍內。在這一步很容易漏掉某個數對（或者憑空多出一個），所以有系統地列出所有可能情況是值得的。

第二個展開式

> ? **想一想** x^4+x^2+1 與第一部分有甚麼關係？

把 x^6-1 看成以 x^2 表示的立方差（difference of cubes）時，它就是另一個因式：

$$x^6-1=(x^2)^3-1=(x^2-1)(x^4+x^2+1).$$

把 $(x^2-1)^{11}$ 拆成 $(x^2-1)^5(x^2-1)^6$，

$$(x^2-1)^{11}(x^4+x^2+1)^5=\left[(x^2-1)(x^4+x^2+1)\right]^5(x^2-1)^6=(x^6-1)^5(x^2-1)^6=P(x).$$

所以第二個展開式正正就是 $P(x)$，即第一個展開式乘以 x^{16}。

x^4 的係數。 這就是 $P(x)$ 中 x^4 的係數，我們在上面已經求得（它對應第一個展開式中的 x^{-12}）。答案是 -15。

x^{38} 的係數。 這對應 $6j+2k=38$，即 $3j+k=19$，而 $j\le 5$、$k\le 6$。由於 $k\le 6$，我們需要 $3j\ge 13$，所以 $j=5$、$k=4$，得出

$$\binom{5}{5}(-1)^0\cdot\binom{6}{4}(-1)^2=15.$$

所以 x^{38} 的係數是 15。

其實有一個巧妙的方法，不用點算也能看出這個結果。P 的次數（degree）是 $30+12=42$，而把 x 換成 $1/x$ 可得

$$x^{42}P(1/x)=(1-x^6)^5(1-x^2)^6=(-1)^{11}P(x)=-P(x).$$

比較係數可知，P 中 x^{42-m} 的係數等於 x^m 的係數的負數。取 $m=4$，x^{38} 的係數就是 $-(-15)=15$，與上面的結果吻合。

驗算

- 把 $x=1$ 代入 $P(x)$：其值為零，所以 P 的所有係數之和是 0。如果你把它完全展開，這是一個快捷的合理性檢查。
- 我們用電腦把兩個乘積精確地乘開（以係數列表的形式）。在第一個乘積中，x^{-12} 的係數是 -15，x^2 的係數是 215；在第二個乘積中，x^4 的係數是 -15，x^{38} 的係數是 15。我們亦確認了第二個乘積的次數是 42，並且等於第一個乘積乘以 x^{16}。
- 把隨機選取的有理數值代入原式中的 x，所得結果與計算出來的展開式完全吻合。

常見錯誤

- 正負號出錯：若選取的冪是 x^{6j} 和 x^{2k}，該項的正負號是 $(-1)^{5-j}(-1)^{6-k}=-(-1)^{j+k}$，而不是 $(-1)^{j+k}$。用錯了的話，每個答案的正負號都會顛倒。
- 計算 x^2 的係數時漏掉 $(1,6)$ 或 $(3,0)$ 這兩項，結果得出 200 或 205。
- 沒有留意到因式分解，直接展開 $(x^2-1)^{11}(x^4+x^2+1)^5$：這樣做是可行的，但冗長而且容易出錯。

STEP 之外

把多個和式的乘積寫成單一個和式，再從中抽出 x 的某一個冪，正是生成函數（generating function）背後的想法。乘積中的某個係數，數算的是由各部分湊出某個總數的方法數目，與我們在上面點算數對 (j,k) 的做法如出一轍。對稱關係 $x^{42}P(1/x) = -P(x)$ 說明 P 是一個（反）迴文多項式（(anti)palindromic polynomial）；這類多項式會在分圓多項式（cyclotomic polynomial）和單位根（roots of unity）的研究中出現，而 $x^6 - 1$ 正是其中最初步的例子。想做更多同類風格的題目，Siklos 的 *Advanced Problems in Mathematics* 是很好的下一步；而 Allenby 的 *Rings, Fields and Groups* 則對多項式的因式分解有完整而嚴謹的論述。

詳細解答

STEP I 2000，第 3 題

純數學 • 階梯函數的積分

STEP I 2000，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/00-s1-q3

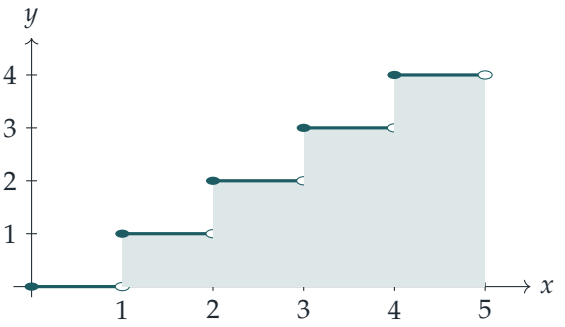
關鍵思路

任何量的整數部分（integer part）都是一個階梯函數（step function）：它在各個區間上保持不變，然後向上跳升一。因此這個積分不過是若干個長方形面積之和，真正要做的工夫只是找出每一級由哪裏開始、到哪裏結束。對於 $[\mathrm{e}^x]$，跳躍發生在 e^x 越過整數之處，即 $x = \ln k$，而由此得出的對數之和恰好可以逐項相消（telescope）。

解答

$y = [x]$ 的圖像

在每個區間 $k \le x < k+1$（其中 k 為整數）上，我們有 $[x] = k$。所以圖像是由水平線段組成的階梯，每段長度為 1，左端是實心點（函數取到該值），右端是空心點（函數在該處向上跳升 1）。



積分就是陰影部分的面積，由闊度為 1、高度分別為 $0, 1, 2, 3, 4$ 的長方形組成（函數在

$x = 1, 2, 3, 4$ 這幾個單獨的點上的取值對面積毫無影響）。因此

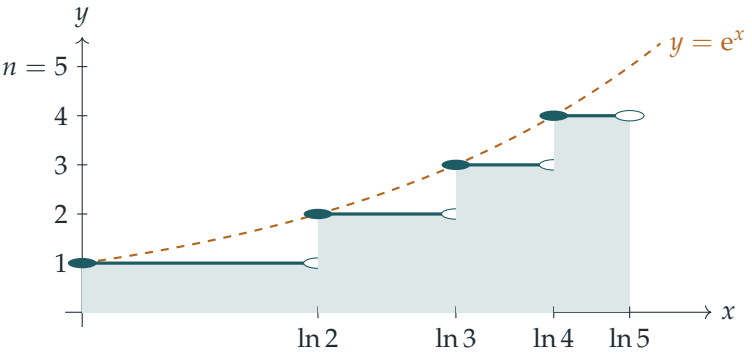
$$\int_0^5 [x]\,\mathrm{d}x = 0 + 1 + 2 + 3 + 4 = 10.$$

$y = [\mathrm{e}^x]$ 的圖像

當 $0 \leq x < \ln n$ 時，我們有 $1 \leq \mathrm{e}^x < n$。由於 e^x 是遞增的，

$$[\mathrm{e}^x] = k \iff k \leq \mathrm{e}^x < k+1 \iff \ln k \leq x < \ln(k+1),$$

其中 $k = 1, 2, \ldots, n-1$。所以圖像仍然是每級高度為 1 的階梯，但現在各級隨着 x 增大而愈來愈短：高度為 k 的那一級闊度為 $\ln(k+1) - \ln k = \ln\left(1 + \frac{1}{k}\right)$。下面的草圖所示為 $n = 5$ 的情況，曲線 $y = \mathrm{e}^x$ 以虛線表示；階梯在每一級的左端與曲線相接。



積分

把各長方形的面積相加，即高度 k 乘以闊度 $\ln(k+1) - \ln k$，

$$\int_0^{\ln n} [\mathrm{e}^x]\,\mathrm{d}x = \sum_{k=1}^{n-1} k\big(\ln(k+1) - \ln k\big) = \sum_{k=1}^{n-1} k\ln(k+1) - \sum_{k=1}^{n-1} k\ln k.$$

想一想 怎樣才能令這兩個和互相抵消？

每個 $\ln j$ 在兩個和中都會出現，只是係數不同。在第一個和中代入 $j = k+1$，便可把兩者對齊：

$$\begin{aligned}\sum_{k=1}^{n-1} k\ln(k+1) - \sum_{k=1}^{n-1} k\ln k &= \sum_{j=2}^{n}(j-1)\ln j - \sum_{j=1}^{n-1} j\ln j \\ &= (n-1)\ln n + \sum_{j=2}^{n-1}\big((j-1) - j\big)\ln j - 1\cdot\ln 1 \\ &= (n-1)\ln n - \sum_{j=2}^{n-1}\ln j.\end{aligned}$$

這裏我們從第一個和中抽出 $j = n$ 的一項，並從第二個和中抽出 $j = 1$ 的一項（該項為零，因為 $\ln 1 = 0$）。現在 $\sum_{j=2}^{n-1} \ln j = \ln\big((n-1)!\big)$，而 $(n-1)\ln n = n\ln n - \ln n$，所以

$$\int_0^{\ln n} [\mathrm{e}^x]\,\mathrm{d}x = n\ln n - \big(\ln n + \ln((n-1)!)\big) = n\ln n - \ln(n!),$$

正是所求。（當 $n = 1$ 時，積分區間的長度為零，等式兩邊都是 0。）

圖像證明（選讀）

闊度為 $\ln n$、高度為 n 的長方形，面積是 $n\ln n$。它位於階梯上方的部分，由高度為 1 的水平長條組成：在高度 $m-1$ 與 m 之間，其中 $m = 1, \ldots, n$，長條由 $x = 0$ 伸延至 $x = \ln m$，因為 $[\mathrm{e}^x] < m$ 恰好在 $x < \ln m$ 時成立。這長條的面積是 $\ln m$，所以階梯上方區域的面積為 $\ln 1 + \ln 2 + \cdots + \ln n = \ln(n!)$。從 $n\ln n$ 中減去它，便再次得到所求結果。這就是「改為對 y 而不是對 x 積分」的離散版本。

驗算

當 $n = 2$ 時，階梯只是一個高度為 1、闊度為 $\ln 2$ 的長方形，而公式給出 $2\ln 2 - \ln 2 = \ln 2$。當 $n = 4$ 時，公式給出 $4\ln 4 - \ln 24 \approx 2.3671$；以中點法則（midpoint rule）計算 $\int_0^{\ln 4} [\mathrm{e}^x]\,\mathrm{d}x$，結果與之吻合至小數點後五位，而對於直至 8 的每一個 n 亦然。答案還必須小於 $\int_0^{\ln n} \mathrm{e}^x\,\mathrm{d}x = n - 1$，因為階梯位於曲線之下：當 $n = 4$ 時，$2.37 < 3$。

常見錯誤

- 把高度為 k 的那一級的闊度當作 1（照搬第一部分），而不是 $\ln(k+1) - \ln k$。
- 讓階梯由高度 0 開始：在 $0 \le x < \ln n$ 上我們有 $\mathrm{e}^x \ge 1$，所以最低的一級位於高度 1。
- 移動求和指標時出錯，漏掉了 $(n-1)\ln n$ 這一項；或者沒有留意到它必須與 $\ln((n-1)!)$ 合併，才能得出 $\ln(n!)$。

STEP 之外

由於階梯位於 $y = \mathrm{e}^x$ 之下，積分小於 $n-1$，整理後即得 $\ln(n!) > n\ln n - n + 1$。以同樣的方法從上方給階梯定界，便會通往斯特靈公式（Stirling's formula）$\ln(n!) \approx n\ln n - n$，這是分析學、概率論和組合數學中的一個關鍵估計。由長方形之和轉換為長條之和，也是「分部求和法」（summation by parts）的一個例子，它是分部積分法（integration by parts）在離散情況下的近親。若想為階梯函數的積分建立嚴格的基礎，Abbott 的

Understanding Analysis（分析學入門讀本）是下一步的好選擇。

詳細解答

STEP I 2000，第 4 題

純數學 • 積分的界限

> **STEP I 2000，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/00-s1-q4

關鍵思路

用初等方法難以直接對 $1/(x^2+1)^4$ 積分，所以題目把它拆成兩部分：一部分恰好是某個函數的導數（exact derivative），積分時不費吹灰之力；另一部分是很小的餘項。第 (ii) 部分負責建立這個分拆，第 (i) 部分則提供餘項的上界。要估計一個積分的界（bound），通常的做法是先估計被積函數（integrand）的界，再乘以區間的長度。

解答

第 (i) 部分

設 $f(x)=\dfrac{x^6}{(x^2+1)^4}$。由於 f 只取決於 x^2，令 $t=x^2$，則 t 與 x 一樣在 $[0,1]$ 上取值，並且 $f=g(t)$，其中

$$g(t)=\frac{t^3}{(1+t)^4}.$$

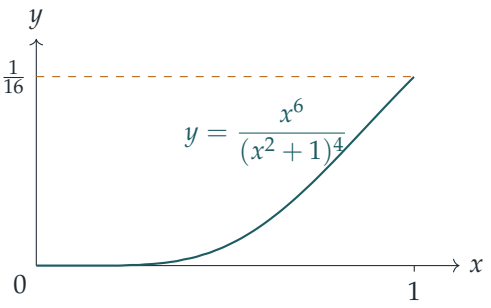
由商法則（quotient rule），

$$g'(t)=\frac{3t^2(1+t)^4-4t^3(1+t)^3}{(1+t)^8}=\frac{t^2\big(3(1+t)-4t\big)}{(1+t)^5}=\frac{t^2(3-t)}{(1+t)^5}.$$

當 $0<t\le 1$ 時，每個因子都是正數（特別地，$3-t\ge 2$），所以 $g'(t)>0$，而 g 在 $[0,1]$ 上嚴格遞增（strictly increasing）。因此它的最大值在 $t=1$ 處取得，即 $x=1$：

$$f(1)=\frac{1}{2^4}=\frac{1}{16},$$

這正是所要證明的。此外，處處都有 $f(x)\ge 0$，而最小值是 $f(0)=0$。



第 (ii) 部分

用商法則對括號內的式子求導。設 $N = Ax^5 + Bx^3 + Cx$，分母為 $(x^2+1)^3$，其導數為 $6x(x^2+1)^2$，於是可以約去一個因子 $(x^2+1)^2$：

$$\frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{N}{(x^2+1)^3}\right) = \frac{(5Ax^4+3Bx^2+C)(x^2+1) - 6x(Ax^5+Bx^3+Cx)}{(x^2+1)^4}.$$

展開分子，並按 x 的冪次整理：

$$\begin{aligned} &5Ax^6 + 5Ax^4 + 3Bx^4 + 3Bx^2 + Cx^2 + C - 6Ax^6 - 6Bx^4 - 6Cx^2 \\ &\quad = -Ax^6 + (5A-3B)x^4 + (3B-5C)x^2 + C. \end{aligned}$$

現在各項的分母都是 $(x^2+1)^4$，所以恆等式（identity）對所有 x 成立，當且僅當兩邊的分子作為多項式相等：

$$(D-A)x^6 + (5A-3B)x^4 + (3B-5C)x^2 + C = 1.$$

由低次項起逐項比較係數（comparing coefficients）：$C = 1$；然後 $3B = 5C = 5$；然後 $5A = 3B = 5$；最後 $D = A$。所以

$$A = 1, \qquad B = \tfrac{5}{3}, \qquad C = 1, \qquad D = 1.$$

第 (iii) 部分

把第 (ii) 部分的恆等式在 $[0,1]$ 上積分。右邊第一項恰好是一個導數，所以

$$\begin{aligned} \int_0^1 \frac{\mathrm{d}x}{(x^2+1)^4} &= \left[\frac{x^5 + \frac{5}{3}x^3 + x}{(x^2+1)^3}\right]_0^1 + \int_0^1 \frac{x^6}{(x^2+1)^4}\,\mathrm{d}x \\ &= \frac{1+\frac{5}{3}+1}{8} - 0 + J = \frac{11}{24} + J, \end{aligned}$$

其中 $J = \displaystyle\int_0^1 \frac{x^6}{(x^2+1)^4}\,\mathrm{d}x$。

想一想 怎樣才能不用算出 J 而估計它的界？

由第 (i) 部分，J 的被積函數滿足 $0 \le f(x) \le \frac{1}{16}$，對 $[0,1]$ 內的每個 x 都成立。把這些不等式在長度為 1 的區間上積分，得

$$0 \le J \le \frac{1}{16} \times 1 = \frac{1}{16}.$$

各項同時加上 $\frac{11}{24}$，

$$\frac{11}{24} \le \int_0^1 \frac{\mathrm{d}x}{(x^2+1)^4} \le \frac{11}{24} + \frac{1}{16},$$

這就是所求的結果。（其實兩個不等式都是嚴格的，因為 f 連續，而且並非在整個區間上都等於任何一個界，但題目只要求證明 $\le$。）

驗算

- 把 $x = 0$ 代入第 (ii) 部分的恆等式：導數項等於 $C = 1$，而含 D 的項為零，得 $1 = 1$。當 $x = 1$ 時，導數的分子是 $-1+0+0+1 = 0$，所以恆等式變成 $\frac{1}{16} = \frac{D}{16}$，與 $D = 1$ 吻合。
- 在數值上，用森遜法則（Simpson's rule）可得 $\int_0^1 \frac{\mathrm{d}x}{(x^2+1)^4} \approx 0.47460$，穩穩落在 $\frac{11}{24} \approx 0.45833$ 與 $\frac{11}{24} + \frac{1}{16} \approx 0.52083$ 之間。餘項 $J \approx 0.01627$ 遠小於 $\frac{1}{16}$，因為 f 除了在 $x = 1$ 附近之外都非常小。

常見錯誤

- 使用商法則時忘記約去公因子 $(x^2+1)^2$，結果分母變成 $(x^2+1)^6$，令比較係數的工作變得一團糟。
- 在第 (i) 部分，只檢查 $x = 1$ 是駐點（stationary point）或端點（endpoint）：你必須證明 f 是遞增的（或比較所有可能的點），才能斷定端點的值是最大值。

STEP 之外

代換(substitution)$x = \tan\theta$ 可把這個積分化為 $\int_0^{\pi/4} \cos^6\theta \,\mathrm{d}\theta$，再用歸約公式(reduction formula）便可求出準確值：其值為 $\frac{11}{48} + \frac{5\pi}{64}$，與我們的數值估計吻合。把被積函數拆成「一個導數加上一個很小的量」，正是歸約公式背後的想法；而透過估計被積函數的界來估計積分的界，則是實分析（real analysis）中最基本的估計方法。Bartle 與 Sherbert 合著的 *Introduction to Real Analysis*（實分析入門）對這些估計有嚴謹細緻的闡述。

詳細解答

STEP I 2000，第 5 題

純數學 • 三維向量

> **STEP I 2000，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/00-s1-q5

關鍵思路

一切都發生在斜面上，所以先在這個平面內設定兩個互相垂直的單位向量（unit vector）：一個沿最大斜率線（line of greatest slope）指向正上方，另一個是水平的。Arthur 和 Bertha 的位置就是這兩個向量的簡單組合，而和差化積公式（sum-to-product formulae）可以立刻把兩者之差整理好。至於與鉛垂線所成的角，我們只需要知道平面內這兩個方向各自與鉛垂方向的關係，θ 正是在這裏出現的。

解答

建立設定

設 $\mathbf{e}$ 為沿着通過 O 的最陡直線指向上方的單位向量，並設 $\mathbf{f}$ 為平面內與 $\mathbf{e}$ 垂直、指向兩人所在一側的單位向量。由於 $\mathbf{f}$ 與最大斜率線垂直，所以它是水平的。設 $\mathbf{k}$ 為鉛垂向上的單位向量。那麼

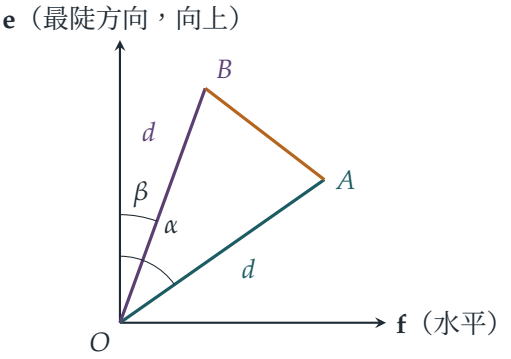
$$\mathbf{e}\cdot\mathbf{k}=\sin\theta,\qquad \mathbf{f}\cdot\mathbf{k}=0,\qquad \mathbf{e}\cdot\mathbf{f}=0.$$

（第一式成立，是因為 $\mathbf{e}$ 與水平面成角 θ，因而與 $\mathbf{k}$ 成角 $\frac{\pi}{2}-\theta$。）

兩人步行的速率相同，所以當 Arthur 走了距離 d 時，Bertha 也走了同樣的距離。他們相對於 O 的位置是

$$\mathbf{a}=d(\cos\alpha\,\mathbf{e}+\sin\alpha\,\mathbf{f}),\qquad \mathbf{b}=d(\cos\beta\,\mathbf{e}+\sin\beta\,\mathbf{f}).$$

兩人都在最陡直線的同一側向上走，即 $0\le\alpha,\beta<\frac{\pi}{2}$。



此圖畫的是斜面本身，即正對着斜面看過去的樣子。

兩人之間的距離

利用和差化積公式，

$$\begin{aligned}\mathbf{a}-\mathbf{b} &= d\big[(\cos\alpha-\cos\beta)\,\mathbf{e}+(\sin\alpha-\sin\beta)\,\mathbf{f}\big]\\ &= d\big[-2\sin\tfrac{1}{2}(\alpha+\beta)\sin\tfrac{1}{2}(\alpha-\beta)\,\mathbf{e}+2\cos\tfrac{1}{2}(\alpha+\beta)\sin\tfrac{1}{2}(\alpha-\beta)\,\mathbf{f}\big]\\ &= 2d\sin\tfrac{1}{2}(\alpha-\beta)\,\big[-\sin\tfrac{1}{2}(\alpha+\beta)\,\mathbf{e}+\cos\tfrac{1}{2}(\alpha+\beta)\,\mathbf{f}\big].\end{aligned}$$

記 $s=\frac{1}{2}(\alpha+\beta)$。向量 $\mathbf{u}=-\sin s\,\mathbf{e}+\cos s\,\mathbf{f}$ 是單位向量，因為 $\mathbf{e}$ 和 $\mathbf{f}$ 是互相垂直的單位向量，而且 $\sin^2 s+\cos^2 s=1$。因此

$$AB=|\mathbf{a}-\mathbf{b}|=2d\,\big|\sin\tfrac{1}{2}(\alpha-\beta)\big|,$$

正是所求。（從幾何上看：OAB 是等腰三角形（isosceles triangle），兩條等邊長 d，頂角為 $|\alpha-\beta|$；由 O 作垂線，可得相同的答案。）

與鉛垂線所成的角

當 $\alpha\neq\beta$ 時，由於 $0<|\alpha-\beta|<\frac{\pi}{2}$，我們有 $\sin\frac{1}{2}(\alpha-\beta)\neq 0$，所以 $A\neq B$，直線 AB 有明確定義。上面的計算顯示，這條直線平行於單位向量 $\mathbf{u}$，或者等價地說，平行於 $\mathbf{v}=-\mathbf{u}=\sin s\,\mathbf{e}-\cos s\,\mathbf{f}$。

? **想一想** 一條直線有兩個方向。我們應該拿哪一個來與鉛垂方向比較？

直線與鉛垂線所成的角是指銳角（或直角），所以我們取直線上指向上方的那個方向，或者至少是不指向下方的那個方向。現在

$$\mathbf{v}\cdot\mathbf{k}=\sin s\,(\mathbf{e}\cdot\mathbf{k})-\cos s\,(\mathbf{f}\cdot\mathbf{k})=\sin\theta\sin s.$$

由於 $0 \le s < \frac{\pi}{2}$ 及 $0 < \theta < \frac{\pi}{2}$，這個值非負，所以 $\mathbf{v}$ 就是向上的方向。因為 $\mathbf{v}$ 和 $\mathbf{k}$ 都是單位向量，兩者之間的夾角 ϕ 滿足

$$\cos\phi = \mathbf{v}\cdot\mathbf{k} = \sin\theta \sin\tfrac{1}{2}(\alpha+\beta),$$

這就是所求的結果。留意結果與 d 無關：兩人一路走，直線 AB 一路平行於自身移動。

驗算

- 若 Arthur 沿水平方向走（$\alpha = \frac{\pi}{2}$，即邊界情況），而 Bertha 筆直向上走（$\beta = 0$），則 $AB = 2d\sin\frac{\pi}{4} = \sqrt{2}\,d$，這正是把畢氏定理（Pythagoras' theorem）用於兩條互相垂直、長度同為 d 的路徑所得的結果。取 $\theta = 30°$ 及 $d = 1$，用坐標可得 $A = (1,0,0)$，$B = (0, \frac{\sqrt{3}}{2}, \frac{1}{2})$，所以由 B 到 A 的鉛垂落差是 $\frac{1}{2}$，而兩點相距 $\sqrt{2}$：$\cos\phi = \frac{1}{2\sqrt{2}} \approx 0.354$，與 $\sin 30° \sin 45°$ 吻合，所以 $\phi \approx 69.3°$。
- 若平面是水平的（$\theta = 0$），則 $\cos\phi = 0$：直線是水平的，理應如此。
- 用電腦在真正的三維坐標中隨機選取 10^5 組 θ、α、β、d 作檢驗，結果與兩條公式的誤差都在 10^{-12} 以內。

常見錯誤

- 量度了平面內 AB 與最陡直線之間的角，而不是空間中與鉛垂線所成的角；答案中必須出現 θ。
- 忘記加絕對值（modulus）：當 $\beta > \alpha$ 時，$\sin\frac{1}{2}(\alpha-\beta)$ 是負數，但距離不會是負數。
- 取了 AB 上向下的方向，得出 $\cos\phi = -\sin\theta\sin s$；直線與鉛垂線所成的角永遠不會是鈍角。

STEP 之外

這裏的方法，是選取一組切合問題的標準正交基（orthonormal basis）（先取平面內的兩個向量，再找出它們與周圍空間中鉛垂方向的關係），然後由點積（dot product）讀出長度和角度。在大學一年級的線性代數課程中，這會發展成基變換（change of basis）和正交投影（orthogonal projection）的概念：$\mathbf{v}\cdot\mathbf{k}$ 就是 $\mathbf{v}$ 在鉛垂方向上的投影的長度。Strang 的 *Introduction to Linear Algebra*（線性代數入門）細緻地闡述了這些概念，並配有大量幾何內容。

詳細解答

STEP I 2000，第 6 題

純數學 • 直線所圍成的區域

STEP I 2000，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/00-s1-q6

關鍵思路

一個關於 x 和 y 的二次式，若能因式分解（factorise）為兩個一次因式，所描述的就是一對直線。兩個因式的積為正，當且僅當兩個因式同號；兩者異號時積為負。因此每個不等式（inequality）所選出的，都是兩條相交直線之間一對相對的「楔形」區域。只要把兩個區域畫在同一組坐標軸上，要找一個公共點，一看便知。

解答

第一個因式分解與區域

展開右邊，

$$\begin{aligned}(x-y+2)(x+y-1) &= x^2+xy-x-xy-y^2+y+2x+2y-2\\ &= x^2-y^2+x+3y-2,\end{aligned}$$

正如所求。所以把第一個不等式兩邊同時減去 2 後，得

$$(x-y+2)(x+y-1)>0.$$

邊界由兩條直線組成：

$$L_1:\ y=x+2, \qquad L_2:\ y=1-x,$$

它們的斜率（gradient）分別為 1 和 -1，並在 $x+2=1-x$ 處相交，即交於 $\left(-\frac{1}{2},\frac{3}{2}\right)$。

兩個因式同號時，積為正：

- 兩者皆正：$y<x+2$ 且 $y>1-x$，即在 L_1 之下、L_2 之上。這是交點*右方*的楔形（例如它包含 $(5,0)$）；

- 兩者皆負：$y > x+2$ 且 $y < 1-x$，即在 L_1 之上、L_2 之下。這是交點左方的楔形（它包含 $(-5,0)$）。

所以這個區域就是直線 $y = x+2$ 與 $y = 1-x$ 之間左右相對的一對楔形，而直線本身並不包括在內（在直線上積為 0，而非 > 0）。簡單測試一下：在原點，該式的值為 $-2 < 0$，而原點位於下方的楔形內，這個楔形確實應被排除。

第二個區域

想一想 我們怎樣知道這一個也能因式分解？它的因式又應該是甚麼樣子？

由二次項 $x^2 - 4y^2 = (x+2y)(x-2y)$ 可以想到嘗試

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

右邊展開後是 $x^2 - 4y^2 + (a+b)x + (2b-2a)y + ab$。比較係數（coefficient）得 $a+b=3$ 和 $b-a=-1$，所以 $a=2$，$b=1$，而常數項 $ab=2$ 亦同樣吻合。因此，把第二個不等式兩邊同時加上 2，得

$$x^2 - 4y^2 + 3x - 2y < -2 \iff (x+2y+2)(x-2y+1) < 0.$$

邊界直線為

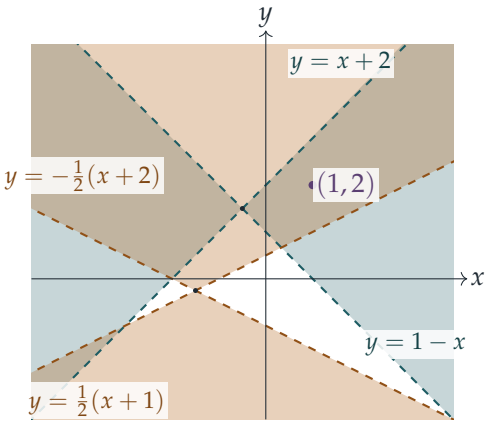
$$L_3:\ y = -\tfrac{1}{2}(x+2), \qquad L_4:\ y = \tfrac{1}{2}(x+1),$$

斜率分別為 $-\frac{1}{2}$ 和 $\frac{1}{2}$，並在 $-x-2 = x+1$ 處相交，即交於 $\left(-\frac{3}{2}, -\frac{1}{4}\right)$。

這次我們需要兩個因式異號：

- $x+2y+2 > 0$ 且 $x-2y+1 < 0$：$y > -\frac{1}{2}(x+2)$ 且 $y > \frac{1}{2}(x+1)$，即在兩條直線之上。這是上方的楔形（它包含 $(0,5)$）；
- $x+2y+2 < 0$ 且 $x-2y+1 > 0$：在兩條直線之下。這是下方的楔形（它包含 $(0,-5)$）。

所以第二個區域就是直線 L_3 與 L_4 之間上下相對的一對楔形，邊界同樣不包括在內。



在草圖中，第一個區域以藍綠色塗上陰影（左右兩個楔形），第二個區域則用琥珀色（上下兩個楔形），所以顏色較深的重疊部分就是兩個不等式同時成立之處；虛線並不包括在內。

同時位於兩個區域的一點

這兩對楔形必定有重疊：右邊的藍綠色楔形在斜率為 ± 1 的兩條直線之間向右張開，而上方的琥珀色楔形則包含了位於斜率為 $\pm\frac{1}{2}$ 的兩條較平緩直線之上的所有點。當 $x > -\frac{1}{2}$ 時，藍綠色楔形的上邊 $y = x + 2$ 比直線 $y = \frac{1}{2}(x+1)$ 上升得更快，所以兩者之間留有空間。

例如取 $(x, y) = (1, 2)$。則

$$x^2 - y^2 + x + 3y = 1 - 4 + 1 + 6 = 4 > 2$$

且

$$x^2 - 4y^2 + 3x - 2y = 1 - 16 + 3 - 4 = -16 < -2,$$

所以 $(1, 2)$ 同時滿足兩個不等式。（用因式核對：$(x - y + 2)(x + y - 1) = 1 \cdot 2 > 0$ 且 $(x + 2y + 2)(x - 2y + 1) = 7 \cdot (-2) < 0$。）

驗算

- 每個因式分解都可以隨意取一點來核對：在 $(3, 1)$，$x^2 - y^2 + x + 3y - 2 = 12$ 而 $(x - y + 2)(x + y - 1) = 4 \cdot 3 = 12$；在同一點，$x^2 - 4y^2 + 3x - 2y + 2 = 14$ 而 $(x + 2y + 2)(x - 2y + 1) = 7 \cdot 2 = 14$。
- 在每個楔形內各測試一點，例如 $(0, -5)$ 給出 $x^2 - 4y^2 + 3x - 2y = -90 < -2$，確認下方的琥珀色楔形屬於第二個區域。

- 在 $[-10,10]^2$ 內隨機抽取 200 000 個點作數值檢驗，結果顯示不等式與上述楔形的描述完全一致，沒有任何出入。

常見錯誤

- 把錯誤的一對楔形塗上陰影。務必測試一點（原點最方便），而不要單靠「之上」和「之下」來判斷。
- 把常數移項時忘記了 -2 和 $+2$：第一個不等式是 $(x-y+2)(x+y-1)>0$，第二個是 $(x+2y+2)(x-2y+1)<0$。
- 把邊界直線畫成實線：這些不等式是嚴格不等式，所以直線並不包括在內。

STEP 之外

一個二元二次式 $ax^2+bxy+cy^2+dx+ey+f$ 定義了一條圓錐曲線（conic）；當它能分解為兩個一次因式時，這條圓錐曲線是退化（degenerate）的，即一對直線，而這種情況恰好在由各係數構成的某個 3×3 行列式（determinant）等於零時出現。二次型（quadratic form）的正負號分佈（這裏的 x^2-y^2 和 x^2-4y^2 都是不定的（indefinite）），會在大學一年級的線性代數（linear algebra）中，透過對稱矩陣（symmetric matrix）及其特徵值（eigenvalue）來研究。Strang 的 *Introduction to Linear Algebra*（《線性代數導論》）是了解這個課題的好去處。

詳細解答

STEP I 2000，第 7 題

純數學 • 導數的界限

STEP I 2000，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/00-s1-q7

關鍵思路

求導之後，條件 $\mathrm{f}'(x) \geq 0$ 就變成「a 不小於某個關於 x 的函數」。因此一切取決於該函數的最大值。這個函數只與 x^2 有關，經代換（substitution）$u = 1 + x^2$ 後，它成為 $1/u$ 的二次式。再用配方法（completing the square）便可求得最大值，並清楚看出最大值在何處取得。

解答

導數

由商法則（quotient rule），

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

所以

$$\mathrm{f}'(x) = a - g(x), \qquad \text{其中} g(x) = \frac{x^2(x^2+3)}{(1+x^2)^2}.$$

因此，$\mathrm{f}'(x) \geq 0$ 對所有 x 成立，當且僅當 $a \geq g(x)$ 對所有 x 成立。我們需要求出 g 的最大值。

g 的最大值

設 $u = 1 + x^2$，則 $u \geq 1$ 且 $x^2 = u - 1$。於是

$$g = \frac{(u-1)(u+2)}{u^2} = \frac{u^2 + u - 2}{u^2} = 1 + \frac{1}{u} - \frac{2}{u^2}.$$

想一想 為甚麼用 $1/u$ 而不是 u 來表示這個式子會較好？

以 $s = 1/u$ 表示，$g = 1 + s - 2s^2$，這只是一個二次式，而二次式的最大值可由配方法求得，無須用到微積分：

$$1 + s - 2s^2 = \frac{9}{8} - 2\left(s - \frac{1}{4}\right)^2.$$

由於平方永不為負，對每個 x 都有 $g \le \frac{9}{8}$，等號僅在 $s = \frac{1}{4}$ 時成立，即 $u = 4$，$x = \pm\sqrt{3}$。

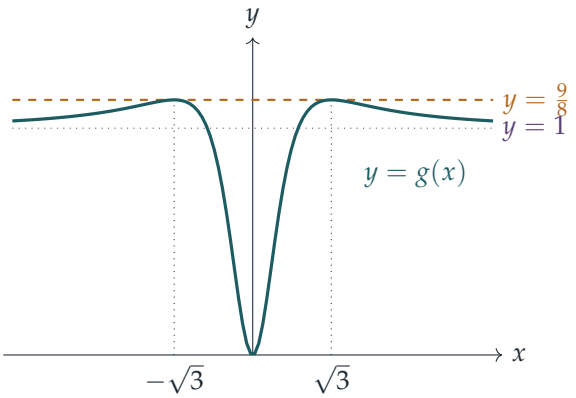
把結果寫回 x 的形式（利用 $s - \frac{1}{4} = \frac{3-x^2}{4(1+x^2)}$），便得到導數的一個簡潔恆等式（identity）：

$$\mathrm{f}'(x) = \left(a - \frac{9}{8}\right) + \frac{(3-x^2)^2}{8(1+x^2)^2}.$$

結論

若 $a \ge \frac{9}{8}$，第一個括號非負，而第二項是一個平方除以正數，所以同樣非負。因此 $\mathrm{f}'(x) \ge 0$ 對所有 x 成立，正如所求。

常數 $\frac{9}{8}$ 也是最佳的：若 $a < \frac{9}{8}$，則 $\mathrm{f}'(\sqrt{3}) = a - \frac{9}{8} < 0$。當 $a = \frac{9}{8}$ 時，f' 只在 $x = \pm\sqrt{3}$ 處為零，所以 f 仍然是嚴格遞增（strictly increasing）的，並有兩個駐拐點（stationary points of inflection）。



g 的圖像由 0 開始上升，越過其極限（limit）1，在 $x = \pm\sqrt{3}$ 時達到頂點 $\frac{9}{8}$，然後回落並趨向 1。導數 $\mathrm{f}'(x)$ 就是直線 $y = a$ 高出這條曲線的高度。

驗算

▶ 在 $x = \sqrt{3}$ 處：$g = \dfrac{3 \cdot 6}{4^2} = \dfrac{18}{16} = \dfrac{9}{8}$。✓

- 當 $x \to \infty$ 時，$g \to 1 < \frac{9}{8}$。這是合理的，因為當 x 很大時 $\frac{x^3}{1+x^2}$ 的表現與 x 相若，所以其斜率趨向 1。
- 數值上，取 $a = \frac{9}{8}$，在 $[-100, 100]$ 上以步長 0.001 的網格計算，$\mathrm{f}'(x)$ 的最小值約為 2×10^{-10}（基本上是 0，在 $x \approx \pm\sqrt{3}$ 處）。取 $a = 1.1$，則 $\mathrm{f}'(\sqrt{3}) = -0.025 < 0$。

常見錯誤

- 把 g 在無窮遠處的極限 1 當作它的最大值。曲線會越過這個極限，所以答案是 $\frac{9}{8}$ 而不是 1。
- 直接對 g 求導，結果迷失在繁複的代數運算之中。代換 $s = 1/(1+x^2)$ 可把問題化為一行就能完成的配方。

STEP 之外

在這裏，一個「對所有 x」的命題被化為求上*確界*（supremum）的問題：$\mathrm{f}' \geq 0$ 處處成立，當且僅當 $a \geq \sup g$。g 在無窮遠處有極限，卻在有限的點上取得最大值，這正好提醒我們：極限不一定是界（bound）。在分析學（analysis）中，你會證明導數非負的函數是遞增的（藉助中值定理（Mean Value Theorem）），這就是為甚麼這道題其實是在問 f 是否單調（monotonic）。Abbott 的 *Understanding Analysis* 對這些概念有細緻的闡述。

詳細解答

STEP I 2000，第 8 題

純數學 • 絕對值的積分

> **STEP I 2000，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/00-s1-q8

關鍵思路

處理積分（integral）內的絕對值（modulus）符號，方法是找出符號內的式子在何處變號。在相鄰兩個變號點之間的每個區間上，$|f|$ 不是 f 就是 $-f$，所以我們在這些點把積分範圍拆開，然後對普通的函數逐段積分。真正要花工夫的，是找出零點的位置：$x\mathrm{e}^x$ 可憑觀察得出，三次式要靠因式分解（factorising），而 $\sin x + \cos x$ 則要寫成單一個正弦波。

解答

第一個積分

由於對所有 x 都有 $\mathrm{e}^x > 0$，$x\mathrm{e}^x$ 的正負號與 x 的正負號相同。因此在 $[-1, 0]$ 上有 $x\mathrm{e}^x \le 0$ 及 $|x\mathrm{e}^x| = -x\mathrm{e}^x$，而在 $[0, 1]$ 上則有 $x\mathrm{e}^x \ge 0$ 及 $|x\mathrm{e}^x| = x\mathrm{e}^x$。在 0 處把積分範圍拆開，

$$\int_{-1}^{1} |x\mathrm{e}^x| \,\mathrm{d}x = \int_{-1}^{0} (-x\mathrm{e}^x) \,\mathrm{d}x + \int_{0}^{1} x\mathrm{e}^x \,\mathrm{d}x = -\int_{-1}^{0} x\mathrm{e}^x \,\mathrm{d}x + \int_{0}^{1} x\mathrm{e}^x \,\mathrm{d}x,$$

這正是所需的恆等式。

要計算其值，可用分部積分法（integration by parts），取 $u = x$ 及 $\mathrm{d}v/\mathrm{d}x = \mathrm{e}^x$：

$$\int x\mathrm{e}^x \,\mathrm{d}x = x\mathrm{e}^x - \int \mathrm{e}^x \,\mathrm{d}x = (x-1)\mathrm{e}^x + c.$$

因此

$$\int_{-1}^{0} x\mathrm{e}^x \,\mathrm{d}x = \left[(x-1)\mathrm{e}^x\right]_{-1}^{0} = -1 - (-2\mathrm{e}^{-1}) = \frac{2}{\mathrm{e}} - 1,$$

$$\int_{0}^{1} x\mathrm{e}^x \,\mathrm{d}x = \left[(x-1)\mathrm{e}^x\right]_{0}^{1} = 0 - (-1) = 1.$$

其中第一個積分是負數（這是必然的，因為被積函數（integrand）在該區間上是負的），於是

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

第 (i) 部分

設 $p(x) = x^3 - 2x^2 - x + 2$。把各項分組，

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

在 $[0,4]$ 內的零點只有 $x = 1$ 和 $x = 2$。首項係數為正、且有單根（simple root）$-1, 1, 2$ 的三次式，在 -1 與 1 之間為正，在 1 與 2 之間為負，在 2 之後為正。所以

$$|p(x)| = \begin{cases} p(x), & 0 \le x \le 1, \\ -p(x), & 1 \le x \le 2, \\ p(x), & 2 \le x \le 4. \end{cases}$$

取原函數（antiderivative）

$$F(x) = \tfrac{1}{4}x^4 - \tfrac{2}{3}x^3 - \tfrac{1}{2}x^2 + 2x,$$

則有

$$F(0) = 0, \qquad F(1) = \tfrac{1}{4} - \tfrac{2}{3} - \tfrac{1}{2} + 2 = \tfrac{13}{12},$$
$$F(2) = 4 - \tfrac{16}{3} - 2 + 4 = \tfrac{2}{3}, \qquad F(4) = 64 - \tfrac{128}{3} - 8 + 8 = \tfrac{64}{3}.$$

? **想一想** 為甚麼不能直接計算 $F(4) - F(0)$，然後取其絕對值？

因為絕對值必須作用在被積函數上，而不是作用在答案上。在 $[1,2]$ 上這個三次式是負的，$F(4) - F(0)$ 會把那塊面積減去，而不是加上。所以我們把三段分開處理：

$$\begin{aligned} \int_0^4 |p(x)|\,\mathrm{d}x &= \big(F(1) - F(0)\big) - \big(F(2) - F(1)\big) + \big(F(4) - F(2)\big) \\ &= \tfrac{13}{12} - \big(\tfrac{2}{3} - \tfrac{13}{12}\big) + \big(\tfrac{64}{3} - \tfrac{2}{3}\big) \\ &= \tfrac{13}{12} + \tfrac{5}{12} + \tfrac{62}{3} = \tfrac{3}{2} + \tfrac{62}{3} = \frac{133}{6}. \end{aligned}$$

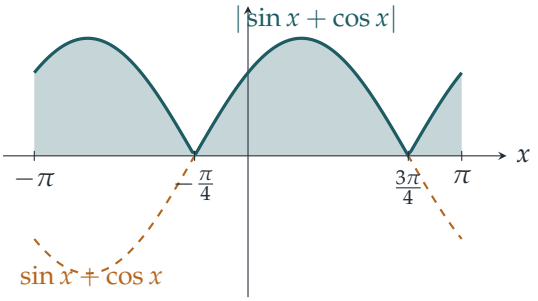
每一項都是正數，這是應有的結果：$\frac{13}{12}$、$\frac{5}{12}$ 和 $\frac{62}{3}$ 就是三塊各自的面積。（作為比較，$F(4) - F(0) = \frac{64}{3}$，比正確答案少了 $2 \times \frac{5}{12}$。）

第 (ii) 部分

利用 $\sin(x + \frac{\pi}{4}) = \frac{1}{\sqrt{2}}(\sin x + \cos x)$，

$$\sin x + \cos x = \sqrt{2}\,\sin\big(x + \tfrac{\pi}{4}\big).$$

在 $[-\pi, \pi]$ 內，此式在 $x = -\frac{\pi}{4}$ 及 $x = \frac{3\pi}{4}$ 處為零。它在 $[-\pi, -\frac{\pi}{4}]$ 和 $[\frac{3\pi}{4}, \pi]$ 上為負，在兩者之間為正。



我們可以像之前那樣拆成三段，但其實有一條更簡潔的路。作代換（substitution）$u = x + \frac{\pi}{4}$：

$$\int_{-\pi}^{\pi} |\sin x + \cos x| \, \mathrm{d}x = \sqrt{2} \int_{-3\pi/4}^{5\pi/4} |\sin u| \, \mathrm{d}u.$$

函數 $|\sin u|$ 的周期（period）為 π，而新的積分範圍長度為 2π，所以剛好涵蓋兩個周期。周期函數在整數個周期上的積分，與起點在哪裏無關，所以

$$\sqrt{2} \int_{-3\pi/4}^{5\pi/4} |\sin u| \, \mathrm{d}u = \sqrt{2} \times 2 \int_{0}^{\pi} \sin u \, \mathrm{d}u = \sqrt{2} \times 2 \times 2 = 4\sqrt{2}.$$

如果你較喜歡直接拆開來做，$G(x) = \sin x - \cos x$ 是 $\sin x + \cos x$ 的一個原函數，其中 $G(-\pi) = 1$、$G(-\frac{\pi}{4}) = -\sqrt{2}$、$G(\frac{3\pi}{4}) = \sqrt{2}$ 及 $G(\pi) = 1$。三塊面積分別是 $1 + \sqrt{2}$、$2\sqrt{2}$ 和 $\sqrt{2} - 1$，加起來是 $4\sqrt{2}$。

驗算

- 在每一段上（被積函數在該段上是光滑的）使用辛普森法則（Simpson's rule），第一個積分得 1.264241 . . .，與 $2 - 2/\mathrm{e} \approx 1.264241$ 吻合。
- 對於 (i)，數值積分（numerical integration）得 22.16667，與 $\frac{133}{6}$ 吻合。它也比 $F(4) - F(0) = \frac{64}{3} \approx 21.33$ 略大，這是必然的，因為絕對值把一小塊負面積變成了正面積。
- 對於 (ii)，數值積分得 5.656854，與 $4\sqrt{2}$ 吻合。再作一個合理性檢查：$|\sin x + \cos x|$ 的形狀與 $\sqrt{2}|\sin x|$ 相同，後者的平均值為 $\frac{2\sqrt{2}}{\pi}$，而 $2\pi \times \frac{2\sqrt{2}}{\pi} = 4\sqrt{2}$。

常見錯誤

- 直接跨過變號點積分，到最後才取絕對值。這樣會把每一個負的區域都少算了。

- 在 (i) 中遺漏了 $x = 1$ 這個根。務必徹底因式分解，或至少用因式定理（factor theorem）檢查細小的整數根。
- 計算 $\left[(x-1)\mathrm{e}^x\right]_{-1}^{0}$ 時弄錯正負號：在 $x = -1$ 處的值是 $-2\mathrm{e}^{-1}$，所以減去它即是加上 $2/\mathrm{e}$。

STEP 之外

積分 $\int_a^b |f(x)|\,\mathrm{d}x$ 是曲線與坐標軸之間的總面積。到了大學，它會以 f 的 L^1 範數（norm）的身分再次出現，這是量度一個函數的大小、以及兩個函數之間距離的一種方法。證明只要 f 可積（integrable）則 $|f|$ 亦可積，以及證明 $\left|\int f\right| \leq \int |f|$，是實分析（real analysis）入門階段的標準習題。Abbott 的 *Understanding Analysis*（實分析入門教材）對黎曼積分（Riemann integral）的建構相當嚴謹，足以證明這兩個結果。

詳細解答

STEP I 2000，第 9 題

力學 • 加速參考系中的拋體

> **STEP I 2000，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/00-s1-q9

關鍵思路

射程是在車廂內量度的，所以應在車廂的參考系（frame of reference）中考慮問題。由於車廂以加速度 a 向北加速，炮彈相對於地板的加速度是向下的 g 加上向南的 a，而向南正是大炮發射的方向。這仍然是一個恆加速度的拋體運動（projectile motion），只是多了一個水平項。只要把射程用 2θ 表示出來，求其最大值便只需一行微分。

解答

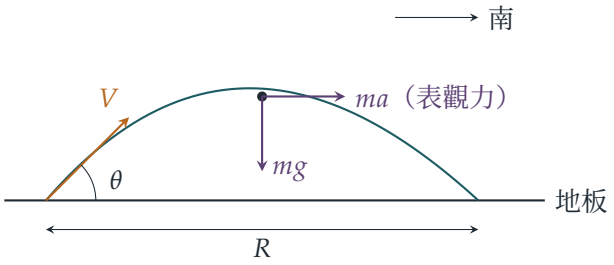
車廂內的射程

把炮彈當作質點（particle），並忽略空氣阻力。假設炮彈從地板高度發射並落在地板上，而且車廂夠長，炮彈會落在車廂之內。取 x 為水平方向並指向南（沿發射方向），y 為鉛垂向上，兩者均相對於車廂量度。

在地面參考系中，炮彈的加速度為向下的 g，並沒有水平加速度。車廂的加速度為向北的 a。因此炮彈相對於車廂的加速度為

$$\ddot{x} = 0 - (-a) = a, \qquad \ddot{y} = -g.$$

換一個說法，在車廂的參考系中，質量為 m 的炮彈除了受本身的重量作用外，還受到一個向南的表觀力（apparent force）ma。



相對於車廂的初速度為向南的 $V\cos\theta$ 和向上的 $V\sin\theta$，所以

$$x = Vt\cos\theta + \tfrac{1}{2}at^2, \qquad y = Vt\sin\theta - \tfrac{1}{2}gt^2.$$

鉛垂方向的運動不受 a 影響：炮彈在 $y = 0$ 且 $t > 0$ 時着地，即在 $T = \dfrac{2V\sin\theta}{g}$ 時着地。因此它在車廂內的射程為

$$\begin{aligned} R = x(T) &= \frac{2V^2\sin\theta\cos\theta}{g} + \frac{a}{2}\cdot\frac{4V^2\sin^2\theta}{g^2} \\ &= \frac{V^2}{g}\left(\sin 2\theta + \frac{a}{g}(1-\cos 2\theta)\right), \end{aligned}$$

當中用到 $2\sin\theta\cos\theta = \sin 2\theta$ 和 $2\sin^2\theta = 1 - \cos 2\theta$。

求射程的最大值

對 θ 求導，得

$$\frac{\mathrm{d}R}{\mathrm{d}\theta} = \frac{2V^2}{g}\left(\cos 2\theta + \frac{a}{g}\sin 2\theta\right),$$

它在以下情況等於零：

$$\boxed{\tan 2\theta = -\frac{g}{a}.}$$

想一想 正切是負數？這是否表示哪裏出了錯？

並沒有出錯。當 $0 < \theta < \frac{\pi}{2}$ 時，我們有 $0 < 2\theta < \pi$，而 $\tan 2\theta < 0$ 只不過表示 2θ 位於 $(\frac{\pi}{2}, \pi)$ 之內，所以 $\theta > \frac{\pi}{4}$。這是合理的：只要炮彈仍在空中，額外的推動 a 便一直起作用，所以發射角比 45° 稍高一點、換取更長的飛行時間是划算的。若 $a = 0$，便有 $\cos 2\theta = 0$，得出大家熟悉的 $\theta = \frac{\pi}{4}$。

要確認這是最大值，設 $\tan\varphi = \dfrac{a}{g}$，其中 $0 < \varphi < \frac{\pi}{2}$。於是

$$\sin 2\theta - \frac{a}{g}\cos 2\theta = \frac{\sin(2\theta - \varphi)}{\cos\varphi}, \qquad \text{所以} \qquad R = \frac{V^2}{g}\left(\frac{\sin(2\theta-\varphi)}{\cos\varphi} + \frac{a}{g}\right).$$

當 $2\theta-\varphi=\frac{\pi}{2}$ 時，此式取最大值，即 $2\theta=\frac{\pi}{2}+\varphi$，而這正是 $\tan 2\theta=-g/a$ 在 $(\frac{\pi}{2},\pi)$ 內的根。所以最佳發射角的準確值為

$$\theta=\frac{\pi}{4}+\frac{1}{2}\arctan\frac{a}{g},$$

又由於 $\cos\varphi=g/\sqrt{g^2+a^2}$，最大射程的準確值為

$$R_{\max}=\frac{V^2}{g}\left(\frac{\sqrt{g^2+a^2}}{g}+\frac{a}{g}\right)=\frac{V^2}{g^2}\left(a+\sqrt{g^2+a^2}\right).$$

a 很小的情況

當 u 很小時，$\arctan u=u+O(u^3)$。取 $u=a/g$，得

$$\theta=\frac{\pi}{4}+\frac{1}{2}\left(\frac{a}{g}+O\left(\frac{a^3}{g^3}\right)\right)\approx\frac{\pi}{4}+\frac{a}{2g},$$

這正是所需的結果。至於射程，$\sqrt{g^2+a^2}=g\sqrt{1+a^2/g^2}=g+O(a^2/g)$，所以

$$R_{\max}=\frac{V^2}{g^2}\left(a+g+O(a^2/g)\right)\approx\frac{V^2}{g}+\frac{V^2a}{g^2},$$

這就是第二個所需的結果。兩個近似式都準確至 a/g 的一階（first order）；略去的各項為 $(a/g)^2$ 階或更小。

如果你想繼續使用 $\tan 2\theta$，可設 $2\theta=\frac{\pi}{2}+\varepsilon$。於是 $\tan 2\theta=-\cot\varepsilon\approx-1/\varepsilon$，所以 $\varepsilon\approx a/g$。把 $\sin 2\theta=\cos\varepsilon\approx 1$ 和 $\cos 2\theta=-\sin\varepsilon\approx-\varepsilon$ 代入 R，可得準確至一階的 $R\approx\frac{V^2}{g}\left(1+\frac{a}{g}\right)$。

驗算

- $a=0$：公式化簡為 $\theta=\frac{\pi}{4}$ 和 $R_{\max}=V^2/g$，即水平地面上的標準結果。
- 量綱（dimensions）：a/g 是無量綱的，而 V^2/g 是長度，正如所料。
- 數值上，取 $V=7$、$g=9.8$ 和 $a=1.3$，在地面參考系中逐步計算運動，再減去車廂的位置，所得結果與 R 的公式吻合（例如在 $\theta=0.7$ 時 $R\approx 5.478$）。對若干個隨機選取的 a，以窮舉法搜尋 θ，發現最大值出現在 $\tan 2\theta=-g/a$ 之處，而 $R_{\max}$ 與 $\frac{V^2}{g^2}(a+\sqrt{g^2+a^2})$ 吻合至 10 位有效數字。當 $a=0.05$ 時，最佳發射角的準確值 0.787949… 與 $\frac{\pi}{4}+\frac{a}{2g}$ 吻合至小數點後七位。

常見錯誤

- 弄錯 a 的正負號。車廂加速的方向與炮彈的發射方向相反，所以相對於地板，炮彈向南的速率會增加：$\frac{1}{2}aT^2$ 這一項是加上去的，而不是減去的。

- 量度了地面上的射程，而不是車廂內的射程。題目問的是沿車廂地板的距離。
- 用計算機求得 $2\theta = \arctan(-g/a)$，這樣會得出一個負的角。正確的根位於 $(\frac{\pi}{2}, \pi)$ 之內。

STEP 之外

在加速參考系中考慮問題，並加上一個表觀力 $-m\mathbf{a}$，是*假想力*（fictitious force）最簡單的例子。旋轉參考系也以同樣的方式引入離心力（centrifugal force）和科里奧利力（Coriolis force）。在本題中，表觀力只不過令「等效重力」傾斜，使其大小變為 $\sqrt{g^2 + a^2}$，這是理解答案中為何出現 $\sqrt{g^2 + a^2}$ 的一種方法。Kleppner 與 Kolenkow 合著的 *An Introduction to Mechanics*（力學入門）對非慣性參考系（non-inertial frame）有清晰的講解。

詳細解答

STEP I 2000，第 10 題

力學 • 碰撞與恢復係數

STEP I 2000，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/00-s1-q10

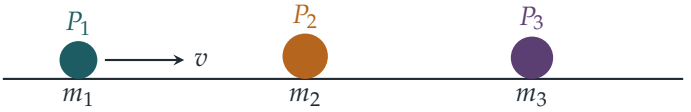
關鍵思路

每次碰撞都給我們兩條方程：動量守恆（conservation of momentum）和牛頓恢復定律（Newton's law of restitution）。這裏碰撞的結果已部分知道（P_1 停下；P_3 以速率 v 離開），所以我們不是求各個速度，而是求未知的係數。可行條件則來自任何恢復係數（coefficient of restitution）都必須符合的唯一物理限制 $0 \le e \le 1$，而最終能量成為唯一可自由變動的質量 m_2 的函數，我們在這些條件所容許的區間上求它的最大值和最小值。

解答

模型

我們把三個物體視為在光滑桌面上沿同一直線運動的質點（particle），因此沒有水平外力作用，每次碰撞中動量都守恆。取 P_1 初始運動的方向為正方向。



第一次碰撞

設 P_2 以速度 u 離開。由於 P_1 變為靜止，

$$\text{動量：} \quad m_1 v = m_2 u \quad \Longrightarrow \quad u = \frac{m_1 v}{m_2},$$

$$\text{恢復定律：} \quad u - 0 = e(v - 0) \quad \Longrightarrow \quad e = \frac{u}{v} = \frac{m_1}{m_2}.$$

第二次碰撞

P_2 現在以速率 u 撞向靜止的 P_3。設 P_2 碰撞後的速度為 w；題目告訴我們 P_3 的速度為 v。由動量守恆得

$$m_2u = m_2w + m_3v \quad\Longrightarrow\quad m_1v = m_2w + m_3v \quad\Longrightarrow\quad w = \frac{(m_1 - m_3)v}{m_2},$$

其中用到第一次碰撞所得的 $m_2u = m_1v$。由恢復定律（分離速率 $= e' \times$ 接近速率）得

$$v - w = e'u.$$

代入 w 和 u，並將全式乘以 m_2/v（它不等於零）：

$$m_2 - (m_1 - m_3) = e'm_1 \quad\Longrightarrow\quad e' = \frac{m_2 + m_3 - m_1}{m_1},$$

這就是所要求的結果。

這樣的碰撞何時可能發生？

任何恢復係數都介乎 0 與 1 之間。應用 $0 \le e' \le 1$，並乘以 $m_1 > 0$：

$$0 \le m_2 + m_3 - m_1 \le m_1 \quad\Longleftrightarrow\quad m_1 \le m_2 + m_3 \le 2m_1,$$

這正是題目要求的不等式。所以它是這些碰撞可能發生的必要條件。

> ? **想一想** 我們是否已用上每一個限制條件？e 又如何？

我們還需要 $0 \le e \le 1$。由於 $e = m_1/m_2$，這個條件成立當且僅當 $m_2 \ge m_1$。這個條件並不屬於題目要求我們證明的結果（事實上，有了它，左邊的不等式 $m_2 + m_3 \ge m_1$ 便自動成立），但它在最後一部分十分重要。綜合以上各點，這些碰撞可能發生當且僅當

$$m_1 \le m_2 \le 2m_1 - m_3.$$

要這個區間非空，我們需要 $m_3 \le m_1$。這時 $w = (m_1 - m_3)v/m_2 \ge 0$，所以第二次碰撞後 P_2 繼續向前運動（或停下），永遠不會回頭撞上靜止的 P_1。同時 $v - w = e'u \ge 0$，所以 P_2 不會再追上 P_3。因此總共只有這兩次碰撞。

最終能量

最後 P_1 靜止，P_2 的速度為 w，P_3 的速度為 v，所以最終動能（kinetic energy）為

$$E = \tfrac{1}{2}m_2w^2 + \tfrac{1}{2}m_3v^2 = \frac{(m_1 - m_3)^2v^2}{2m_2} + \tfrac{1}{2}m_3v^2.$$

當 m_1、m_3 和 v 固定時，只有 m_2 可以變動，範圍是 $m_1 \le m_2 \le 2m_1 - m_3$。第一項是一個非負常數除以 m_2，所以 E 是 m_2 的遞減函數（若 $m_1 = m_3$ 則為常數，不過這時區間本來就只有 $m_2 = m_1$ 這一點）。

最大能量在 $m_2 = m_1$ 時出現（第一次碰撞為完全彈性（perfectly elastic），$e = 1$）：

$$E_{\max} = \frac{(m_1 - m_3)^2 v^2}{2m_1} + \frac{m_3 v^2}{2} = \frac{(m_1^2 - m_1 m_3 + m_3^2)\, v^2}{2m_1}.$$

最小能量在 $m_2 = 2m_1 - m_3$ 時出現（第二次碰撞為完全彈性，$e' = 1$）：

$$E_{\min} = \frac{(m_1 - m_3)^2 v^2}{2(2m_1 - m_3)} + \frac{m_3 v^2}{2} = \frac{(m_1^2 - 2m_1 m_3 + m_3^2 + 2m_1 m_3 - m_3^2) v^2}{2(2m_1 - m_3)} = \frac{m_1^2 v^2}{2(2m_1 - m_3)}.$$

驗算

- 碰撞中能量永遠不會增加，所以我們需要 $E_{\max} \le \frac{1}{2} m_1 v^2$。這即是 $m_1^2 - m_1 m_3 + m_3^2 \le m_1^2$，亦即 $m_3(m_3 - m_1) \le 0$，由於 $m_3 \le m_1$，此式成立。
- 可以驗證 $E_{\max} - E_{\min} = \dfrac{(m_1 - m_3)^3 v^2}{2m_1(2m_1 - m_3)} \ge 0$，而且當 $m_1 = m_3$ 時兩個答案都等於 $\frac{1}{2} m_1 v^2$（這時 $m_2 = m_1$，兩次碰撞都是彈性碰撞，沒有能量損失）。
- 數值驗算：取 $m_1 = 2$、$m_3 = 1$、$v = 1$，初始能量為 1，$E_{\max} = \frac{3}{4}$，$E_{\min} = \frac{2}{3}$。以精確算術讓 m_2 逐步走遍 $[2, 3]$，可重現這兩個極值，並確認 $e, e' \in [0, 1]$ 在整個區間內成立，在區間外則不成立。

常見錯誤

- 在 $m_1 - m_3 \le m_2 \le 2m_1 - m_3$ 上求最大值和最小值（只考慮 e'）。這樣會得出「最大值」$\frac{1}{2} m_1 v^2$，即全部初始能量，但對應的 m_2 值使 $e = m_1 / m_2 > 1$，這是不可能的。
- 把恢復定律弄反了：是分離速率等於 e 乘以接近速率。

STEP 之外

這裏的模式是：每個限制條件為參數劃出一個容許的區間，然後我們在該區間上求一個單調函數的最大值和最小值。這是約束優化（constrained optimisation）最簡單的情形，由此可發展出線性規劃（linear programming）和拉格朗日乘數（Lagrange multipliers）。在物理方面，碰撞在質心參考系（centre-of-mass frame）中簡潔得多：在那裏，恢復係數只是把每個速度乘以 $-e$，而損失的能量為 $\frac{1}{2}\mu(1 - e^2) \times$（相對速率）^{2}，其中 μ 為約化質量（reduced mass）。Kleppner 與 Kolenkow 的 *An Introduction to Mechanics*（力學入門）正是以這個觀點闡述碰撞。

詳細解答

STEP I 2000，第 11 題

力學 • 梯子的摩擦力平衡

> **STEP I 2000，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/00-s1-q11

關鍵思路

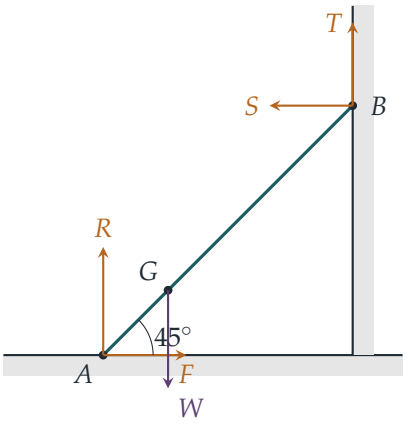
這是一道兩端都有摩擦力的梯子問題。「在 A 和 B 兩處同時達到極限」這個條件，令每個摩擦力（frictional force）都可以用對應的法向反作用力（normal reaction）表示，於是兩條分解力的方程便決定了全部四個接觸力。這樣一來，力矩（moment）方程就成了多出來的一個條件：它要麼碰巧成立，要麼不成立。在第一部分，我們證明它不成立。在第二部分，我們把額外質量的位置當作未知數，令它成立。

解答

模型與方程

我們把桿看成一根剛性桿（rigid rod），其重量 $W = 5g\,\mathrm{N}$ 作用於 G。A（地面）和 B（牆壁）兩處的接觸面都是粗糙的。各個力如下：

- 在 A：法向反作用力 R 向上，摩擦力 F 沿水平方向；
- 在 B：法向反作用力 S 沿水平方向離開牆壁，摩擦力 T 沿鉛垂方向（取向上為正，所以 $T < 0$ 即表示向下）；
- 重量 W 在 G 處向下。



水平方向的力只有 F 和 S，而 $S \geq 0$ 把桿推離牆壁。所以 F 必須指向牆壁，並且

$$F = S. \tag{1}$$

沿鉛垂方向分解，

$$R + T = W. \tag{2}$$

對 A 取力矩時，設 $AB = 0.81$，$c = \cos 45^\circ = \sin 45^\circ$。那麼 B 與 A 的水平距離是 $0.81c$，高度也是 $0.81c$，而 G 與 A 的水平距離是 $0.21c$。比較順時針和逆時針方向的力矩，

$$S\,(0.81c) + T\,(0.81c) = W\,(0.21c),$$

約去 c 後得

$$0.81\,(S + T) = 0.21\,W. \tag{3}$$

45° 的好處就在這裏：因子 c 可以從每一項中約去。

摩擦力不可能在兩端同時達到極限

假設摩擦力在兩端都達到極限。那麼 $F = 0.2R$，而 $|T| = 1.0\,S = S$。

> **想一想** 當 B 處的摩擦力達到極限時，它指向哪個方向？

最自然的情況是桿即將滑動，A 遠離牆壁而 B 向下滑，這時 $T = +S$。為求周全，我們把兩種符號都檢查一遍。

情況 $T = S$。由 (1)，$S = F = 0.2R$，而 (2) 給出 $R + 0.2R = W$。所以 $R = \frac{5}{6}W$，$S = T = \frac{1}{6}W$。於是 (3) 的左方是

$$0.81\left(\tfrac{1}{6}W + \tfrac{1}{6}W\right) = 0.27\,W \neq 0.21\,W.$$

所以力矩方程不成立。

情況 $T = -S$。(3) 的左方是 0，但右方是 $0.21W > 0$，這是不可能的。

在兩種情況下，桿都不可能處於平衡（equilibrium）。所以摩擦力不可能在 A 和 B 兩處同時達到極限，這正是所要證明的。

（在非極限的情況下，平衡當然是可能的：例如 $T = S = \frac{7}{54}W$ 滿足 (3)，並給出 $R = \frac{47}{54}W$，而且 $F = S \leq 0.2R$。）

在 P 加上質量

設 $AP = x\,\mathrm{m}$。額外的重量同樣是 $W = 5g\,\mathrm{N}$，作用於 P，所以總重量是 $2W$。方程 (1) 和 (2) 變成

$$F = S, \qquad R + T = 2W,$$

而對 A 的力矩方程多了一項 $W\,(xc)$。約去 c 後，它變成

$$0.81\,(S + T) = 0.21\,W + x\,W. \tag{4}$$

現在摩擦力在兩端都達到極限。$T = -S$ 的情況同樣不可能，因為這時 (4) 的左方是 0，右方卻是正數。所以 $T = S$ 且 $F = 0.2R$，因此

$$S = 0.2R, \qquad R + 0.2R = 2W \quad \Longrightarrow \quad R = \tfrac{5}{3}W, \quad S = T = \tfrac{1}{3}W.$$

代入 (4) 得

$$0.81 \times \tfrac{2}{3}\,W = 0.54\,W = 0.21\,W + x\,W,$$

所以 $x = 0.33$。因此

$$AP = 0.33\,\mathrm{m}.$$

這個值介乎 0 與 0.81 之間，所以 P 在桿上。它在 G 之外（也在中點之外），所以額外的質量把合重心（centre of gravity）沿桿向上推，而這往往會令梯子滑倒。

驗算

- 質量放在 P 時，以 W 為單位，各力為 $R = \frac{5}{3}$ 和 $F = S = T = \frac{1}{3}$。於是 $F/R = 0.2$，$T/S = 1$，所以兩處都達到極限。鉛垂方向有 $\frac{5}{3} + \frac{1}{3} = 2$，而力矩給出 $0.81 \times \frac{2}{3} = 0.54 = 0.21 + 0.33$。
- 合重心與 A 的距離是 $\frac{1}{2}(0.21 + 0.33) = 0.27\,\mathrm{m}$。這正是第一部分中，要令摩擦力在兩端同時達到極限所需的位置（$0.27W$ 對比 $0.21W$）。這是合理的，因為摩擦係數不變，只是總重量增加了一倍。
- 質量和 g 在整個計算中都會約去，所以答案與它們無關。以 $g = 9.8$ 和真正的 45° 力矩作數值代入，所得殘差的數量級為 10^{-15}。

常見錯誤

- 把 A 處摩擦力的方向取錯。由於 S 是另外唯一的水平力，F 必須指向牆壁。
- 忘記 B 處的摩擦力是沿鉛垂方向的，所以它會出現在鉛垂方向的分解式中，令 R 小於總重量。
- 用了桿的中點而不是 G。這根桿並不均勻。

STEP 之外

平面上的剛體（rigid body）給出三條平衡方程，所以一道有四個未知接觸力的問題是靜不定(statically indeterminate)的。單靠剛體靜力學，無法得知摩擦力怎樣在兩端之間分配。加上「在兩端同時達到極限」這個條件，就多了兩條方程，令方程組變成超定（over-determined），而這正是第一部分無解的原因。在工程學上，靜不定的問題是透過為接觸點的彈性建立模型來解決的。Kleppner 與 Kolenkow 的 *An Introduction to Mechanics*（力學入門）以大學一年級的程度，仔細討論了剛體靜力學和摩擦力。

詳細解答

STEP I 2000，第 12 題

概率與統計 • 幾何分佈

STEP I 2000，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/00-s1-q12

關鍵思路

每一種情況都先問：已知之前每一次嘗試都失敗，下一次成功的機會是多少？然後把首 $n-1$ 次嘗試失敗的概率（probability）與第 n 次成功的概率相乘。三種規則得出的答案形態截然不同：一種是條件概率（conditional probability）不斷增大，一種是由一開始就保持不變，還有一種是第一次之後才保持不變。

全篇以 N 表示所需的嘗試次數，因此我們要求的是 $\mathrm{P}(N=n)$，其中 $n=1,2,3,\ldots$。我們把能開門的那條鑰匙稱為正確鑰匙。

解答

第 (i) 部分

這裏我絕不重複使用任何鑰匙，所以我是按一個均勻隨機的次序逐一試鑰匙。根據對稱性（symmetry），正確鑰匙落在這個次序中 k 個位置的任何一個，可能性都相同。因此

$$\mathrm{P}(N=n)=\frac{1}{k}\quad \text{其中} n=1,2,\ldots,k, \qquad \mathrm{P}(N=n)=0 \text{ 當} n>k.$$

如果覺得這個對稱性論證來得太取巧，可以改為把條件概率相乘。失敗 j 次之後，還有 $k-j$ 條未試過的鑰匙，其中一條是正確的，所以對於 $1\le n\le k$

$$\mathrm{P}(N=n)=\frac{k-1}{k}\cdot\frac{k-2}{k-1}\cdots\frac{k-n+1}{k-n+2}\cdot\frac{1}{k-n+1}=\frac{1}{k},$$

因為這個乘積會逐項相消（telescope）。

第 (ii) 部分

現在每一次嘗試都是從全部 k 條鑰匙中重新、獨立（independent）地選取，所以不論之前發生過甚麼，每一次嘗試成功的概率都是 $\frac{1}{k}$。要剛好嘗試 n 次，首 $n-1$ 次必須失敗，而第 n 次必須成功：

$$\mathrm{P}(N=n)=\left(\frac{k-1}{k}\right)^{n-1}\frac{1}{k},\qquad n=1,2,3,\ldots$$

這是參數為 $\frac{1}{k}$ 的幾何分佈（geometric distribution），而 n 並沒有上限。

第 (iii) 部分

第一次嘗試成功的概率是 $\frac{1}{k}$，所以 $\mathrm{P}(N=1)=\frac{1}{k}$。

現在假設第 m 次嘗試（對某個 $m\geq 1$）失敗了。那麼第 m 次所試的是一條錯的鑰匙。在第 $m+1$ 次嘗試，我從其餘 $k-1$ 條鑰匙中均勻地選取。

> **想一想** 第 m 次嘗試之前的經過，會否影響第 $m+1$ 次嘗試成功的機會？

不會。唯一有關係的是哪一條鑰匙被排除，而被排除的永遠是一條錯的鑰匙。所以可供選擇的 $k-1$ 條鑰匙永遠包括正確的那一條，而不論之前發生過甚麼，第 $m+1$ 次嘗試成功的概率都是 $\frac{1}{k-1}$，失敗的概率都是 $\frac{k-2}{k-1}$。

所以對於 $n\geq 2$：第一次嘗試失敗（概率為 $\frac{k-1}{k}$），第 $2,\ldots,n-1$ 次嘗試失敗（每次概率為 $\frac{k-2}{k-1}$），而第 n 次嘗試成功（概率為 $\frac{1}{k-1}$）。因此

$$\mathrm{P}(N=n)=\frac{k-1}{k}\left(\frac{k-2}{k-1}\right)^{n-2}\frac{1}{k-1}=\frac{1}{k}\left(\frac{k-2}{k-1}\right)^{n-2},\qquad n\geq 2,$$

另外還有 $\mathrm{P}(N=1)=\frac{1}{k}$。

有趣的是，同樣有 $\mathrm{P}(N=2)=\frac{1}{k}$，與第 (i) 部分的結果一樣。（當 $k=2$ 時，這個規則迫使我輪流使用兩條鑰匙，所以 N 是 1 或 2，概率各為 $\frac{1}{2}$；只要把 0^0 理解為 1，公式便給出這個結果。）

驗算

- 每個分佈的總和都是 1。在 (iii) 中，$\frac{1}{k}+\frac{1}{k}\sum_{j\geq 0}\left(\frac{k-2}{k-1}\right)^j=\frac{1}{k}+\frac{1}{k}(k-1)=1$。
- 各個平均值（mean）都合乎情理：(i) 是 $\frac{k+1}{2}$，(ii) 是 k，(iii) 是 $1+\frac{(k-1)^2}{k}$。與 (ii) 相比，不重複使用上一條鑰匙略有幫助，但遠遠不及完全不重複使用任何鑰匙。
- 我們對 $k=2,\ldots,5$ 和 $n\leq 7$，以分數列舉每一個選擇序列，精確地核對了全部三條公式。就 $k=5$ 的 (iii) 進行 2×10^5 次模擬（simulation）試驗，得出 $\mathrm{P}(N=3)\approx$

0.1496，而精確值是 $\frac{1}{5} \cdot \frac{3}{4} = 0.15$。

常見錯誤

- 在 (iii) 中，第一次嘗試也用了 $\frac{k-2}{k-1}$。第一次是從全部 k 條鑰匙中選取，所以第一個因子是 $\frac{k-1}{k}$。
- 在 (i) 中，忘記說明當 $n > k$ 時 $\mathrm{P}(N = n) = 0$。
- 在 (iii) 中，以為被排除的鑰匙可能是正確的那一條。其實被排除的永遠只是剛剛試過而失敗的鑰匙。

STEP 之外

第 (iii) 部分是一條小型的馬可夫鏈（*Markov chain*）。接下來發生甚麼只取決於當前的狀態（state）（這裏是指上一次試的是哪一條鑰匙），而與整段歷史無關。找出一個令過去變得無關重要的狀態，正是馬可夫鏈理論的核心技巧，它把繁複的條件概率化為簡單的乘積。Grimmett 與 Welsh 的 *Probability: An Introduction*（概率入門教材）正是以這個程度介紹馬可夫鏈。

詳細解答

STEP I 2000，第 13 題

概率與統計 • 條件概率

STEP I 2000，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/00-s1-q13

關鍵思路

整個問題歸根究柢就是追蹤父母各自傳下來的一個基因。在 (i) 中，兩個 AB 型父母所生的每個孩子，是 AA 或 BB 的概率（probability）同樣那麼小，而四個孩子互相獨立（independent），所以我們只需數出排列方式的數目，再把概率相乘。第 (ii) 部分是「逆向」的條件概率（conditional probability）：我們知道結果（孩子），要找的是原因（父親）。這正是貝葉斯定理（Bayes' theorem），而按父親三種可能的基因型（genotype）畫出樹形圖（tree diagram），就能把計算整理得井井有條。

解答

第 (i) 部分

父母各自都是 AB，所以各自以概率 $\frac{1}{2}$ 傳下 A 或 B，而且父母二人互相獨立。對單獨一個孩子而言，

$$\mathrm{P}(AA) = \tfrac{1}{2}\cdot\tfrac{1}{2} = \tfrac{1}{4}, \qquad \mathrm{P}(BB) = \tfrac{1}{4}, \qquad \mathrm{P}(AB) = \tfrac{1}{2},$$

其中最後一個已把兩種次序都計算在內（A 來自母親而 B 來自父親，或者反過來）。

不同孩子的遺傳互相獨立。四個孩子中兩個是 AA、兩個是 BB 的某一個特定次序，例如 AA, AA, BB, BB，其概率為 $\left(\frac{1}{4}\right)^4$。從四個孩子中選出哪兩個是 AA 的方法數目是 $\binom{4}{2} = 6$。所以

$$\mathrm{P}(\text{兩個}AA\text{ 和兩個}BB) = 6\times\left(\tfrac{1}{4}\right)^4 = \frac{6}{256} = \frac{3}{128},$$

正如題目所求。

第 (ii) 部分

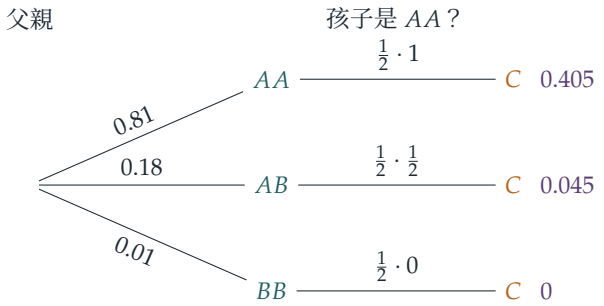
設 F_{AA}、F_{AB}、F_{BB} 分別為我父親是 AA、AB、BB 的事件，並設 C 為我是 AA 的事件。我們把父親視為從人口中隨機抽出的一員，而且這個選取與母親的基因型無關，所以

$$\mathrm{P}(F_{AA}) = 0.81, \qquad \mathrm{P}(F_{AB}) = 0.18, \qquad \mathrm{P}(F_{BB}) = 0.01.$$

> ? **想一想** 為甚麼 0.18 還不是答案？

因為我們有額外的資料：我是 AA。這完全排除了父親是 BB 的可能，同時令父親是 AA 的可能性增加，因為 AA 型的父親必定傳下 A。所以我們必須以 C 為條件。

我要成為 AA，就需要從母親得到一個 A，也從父親得到一個 A。我母親是 AB，所以不論父親是哪一種基因型，她傳下 A 的概率都是 $\frac{1}{2}$。父親傳下 A 的概率則視乎他的基因型，分別為 1、$\frac{1}{2}$ 或 0。樹形圖如下：



根據全概率公式（law of total probability），

$$\begin{aligned}\mathrm{P}(C) &= 0.81 \cdot \tfrac{1}{2} \cdot 1 + 0.18 \cdot \tfrac{1}{2} \cdot \tfrac{1}{2} + 0.01 \cdot \tfrac{1}{2} \cdot 0 \\ &= 0.405 + 0.045 + 0 = 0.45.\end{aligned}$$

再根據貝葉斯定理，

$$\mathrm{P}(F_{AB} \mid C) = \frac{\mathrm{P}(F_{AB})\,\mathrm{P}(C \mid F_{AB})}{\mathrm{P}(C)} = \frac{0.045}{0.45} = \frac{1}{10}.$$

所以我父親是 AB 的概率是 $\frac{1}{10}$。

驗算

- 在 (i) 中，四個孩子各種基因型數目組合的概率，總和應為 1；$\binom{4}{2}$ 這個因子，正是 $\left(\frac{1}{4} + \frac{1}{2} + \frac{1}{4}\right)^4$ 的多項式展開（multinomial expansion）中出現的那一個，而 $3/128 \approx 0.0234$ 就是其中一項。把八個基因的全部 2^8 種遺傳方式逐一列舉，所得的結果正是 $3/128$。

- 在 (ii) 中，答案由先驗概率（prior）0.18 降至 0.1，這是合理的：父親是 BB 的可能已被排除，而父親是 AA 的可能性則有所增加。模擬 10^6 個家庭，所得結果為 0.0998。

常見錯誤

- 在 (i) 中漏掉因子 $\binom{4}{2} = 6$，得出 $1/256$。
- 在 (ii) 中答 0.18，或者計算了 $\mathrm{P}(C \mid F_{AB}) = \frac{1}{4}$，而不是 $\mathrm{P}(F_{AB} \mid C)$。
- 只在部分分支中漏掉母親的因子 $\frac{1}{2}$：這個因子最終會約去，但前提是每一個分支都一致地使用它。

STEP 之外

人口中的比例 81%、18%、1% 正是 p^2、$2pq$、q^2，其中 $p = 0.9$、$q = 0.1$：這就是哈迪－溫伯格平衡（Hardy–Weinberg equilibrium），在隨機交配之下，它會一代一代地保持不變。從觀察到的結果推斷隱藏的原因，是貝葉斯統計學（Bayesian statistics）的核心，而同樣的計算也是遺傳諮詢和親子鑑定的基礎。Blitzstein and Hwang, *Introduction to Probability*（概率論入門）對條件概率和貝葉斯定理有詳盡的論述，並附有許多與本題類似的例子。

詳細解答

STEP I 2000，第 14 題

概率與統計 • 均勻分佈的矩

> **STEP I 2000，第 14 題**
>
> 繼續閱讀之前，請先看題目（原題為英文）。題目載於 STEP Support Programme 的試題資料庫：
>
> step.maths.org/questions/00-s1-q14

關鍵思路

這裏的一切都是均勻分佈（uniform distribution）的矩（moment），而 X 在 $[-1,1]$ 上的矩很容易計算：奇數階的矩因對稱性而等於零，偶數階的矩則為 $1/(k+1)$。至於 $Z = Y - X$，我們根本不需要知道 Z 的分佈。把 $(Y-X)^2$ 和 $(Y-X)^4$ 按二項式展開，再利用獨立性（independence），每個期望值（expectation）都化為我們已知的矩的乘積。

解答

X 的矩

X 在 $[-1,1]$ 上的密度（density）為 $\frac{1}{2}$，所以對非負整數 k，

$$\mathrm{E}(X^k) = \frac{1}{2}\int_{-1}^{1} x^k \,\mathrm{d}x = \begin{cases} 0, & k \text{ 為奇數（被積函數是奇函數）}, \\ \dfrac{1}{k+1}, & k \text{ 為偶數}. \end{cases}$$

特別地，$\mathrm{E}(X) = \mathrm{E}(X^3) = 0$，$\mathrm{E}(X^2) = \frac{1}{3}$，而 $\mathrm{E}(X^4) = \frac{1}{5}$。因此

$$\mathrm{E}(X^2) = \frac{1}{3}, \qquad \mathrm{Var}(X^2) = \mathrm{E}(X^4) - \big(\mathrm{E}(X^2)\big)^2 = \frac{1}{5} - \frac{1}{9} = \frac{4}{45}.$$

Y 的分佈相同，所以它的矩也相同。

Z^2 的平均值

展開得 $Z^2 = Y^2 - 2XY + X^2$。由於 X 和 Y 獨立，$\mathrm{E}(XY) = \mathrm{E}(X)\,\mathrm{E}(Y) = 0$，所以

$$\mathrm{E}(Z^2) = \frac{1}{3} - 0 + \frac{1}{3} = \frac{2}{3}.$$

Z^2 的方差

我們需要 $\mathrm{E}(Z^4)$。根據二項式定理（binomial theorem），

$$Z^4 = (Y-X)^4 = Y^4 - 4Y^3X + 6Y^2X^2 - 4YX^3 + X^4.$$

> **想一想** 我們真的可以把 Y^3X 的期望值寫成 $\mathrm{E}(Y^3)\,\mathrm{E}(X)$ 嗎？

可以：因為 X 和 Y 獨立，所以它們的任何函數（例如 Y^3 和 X）也獨立，而獨立隨機變量（random variable）之積的期望值等於各自期望值之積。中間兩個交叉項各含有 X 或 Y 的奇數次冪，所以它們的期望值都是零。因此

$$\begin{aligned}\mathrm{E}(Z^4) &= \mathrm{E}(Y^4) + 6\,\mathrm{E}(Y^2)\,\mathrm{E}(X^2) + \mathrm{E}(X^4)\\ &= \frac{1}{5} + 6\cdot\frac{1}{3}\cdot\frac{1}{3} + \frac{1}{5} = \frac{2}{5} + \frac{2}{3} = \frac{16}{15}.\end{aligned}$$

所以

$$\mathrm{Var}(Z^2) = \mathrm{E}(Z^4) - \big(\mathrm{E}(Z^2)\big)^2 = \frac{16}{15} - \frac{4}{9} = \frac{48-20}{45} = \frac{28}{45}.$$

由於 $\frac{28}{45} = 7\times\frac{4}{45}$，我們得到 $\mathrm{Var}(Z^2) = 7\,\mathrm{Var}(X^2)$，這正是所要證明的。

驗算

- Z 在 $[-2,2]$ 上有三角形密度 $f(z) = \frac{1}{4}(2-|z|)$。直接計算，$\mathrm{E}(Z^4) = 2\int_0^2 \frac{1}{4}z^4(2-z)\,\mathrm{d}z = \frac{1}{2}\left(\frac{64}{5} - \frac{32}{3}\right) = \frac{16}{15}$，與展開所得的結果一致。
- $\mathrm{E}(Z^2) = \mathrm{Var}(Y-X) = \mathrm{Var}(Y) + \mathrm{Var}(X) = \frac{1}{3} + \frac{1}{3}$，因為 $\mathrm{E}(Z) = 0$。
- 模擬一百萬對 (X,Y)，得出 $\mathrm{Var}(X^2) \approx 0.0889$ 和 $\mathrm{Var}(Z^2) \approx 0.623$，而理論值為 $4/45 \approx 0.0889$ 和 $28/45 \approx 0.622$。

常見錯誤

- 寫成 $\mathrm{Var}(Z^2) = \mathrm{Var}(Y^2) + \mathrm{Var}(X^2)$。這樣做忽略了 Z^2 內的 $-2XY$ 一項，它的平均值雖然是零，卻貢獻了 $4\,\mathrm{Var}(XY) = \frac{4}{9} = \frac{20}{45}$；這個錯誤會得出 $\frac{8}{45}$，而不是 $\frac{28}{45}$。
- 忘記了 Y^2X^2 一項的二項式係數 6。
- 在 $[-1,1]$ 上把密度當作 1 而不是 $\frac{1}{2}$，結果每個矩都變成兩倍。

STEP 之外

$\mathrm{E}(X^k)$ 這些量就是 X 的矩，而處理獨立變量之和的簡潔方法是矩母函數（moment generating function）$M_X(t) = \mathrm{E}(\mathrm{e}^{tX})$，它滿足 $M_{Y-X}(t) = M_Y(t)M_X(-t)$。把這個乘積展開成 t 的冪級數（power series），就可以一次過得到 Z 的所有矩。矩母函數和特徵函數（characteristic function）是大學第一門概率課程的核心內容：它們是證明中心極限定理（central limit theorem）的主要工具。Grimmett 與 Welsh 所著的 *Probability: An Introduction*（概率論入門教材）對這些內容有清晰的闡述。

詳細解答

STEP II 2000，第 1 題

純數學 • 埃及分數

> **STEP II 2000，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/00-s2-q1

關鍵思路

兩個例子有共同的規律：第一個分母比 N 大一，第二個分母則是 N 與它的乘積。至於唯一性的部分，先去分母，再「配成矩形」，便可把 $\frac{1}{a}+\frac{1}{b}=\frac{1}{N}$ 化為兩個整數之積等於 N^2 的形式。質數（prime）的平方只有極少數的因數分解（factorisation）方式，所以 a 和 b 的可能取值寥寥可數，我們只需逐一列出即可。

解答

一般結果

在 $\frac{1}{2}=\frac{1}{3}+\frac{1}{6}$ 和 $\frac{1}{3}=\frac{1}{4}+\frac{1}{12}$ 中，分母分別是 $N+1$ 和 $N(N+1)$（因為 $6=2\times 3$ 而 $12=3\times 4$）。於是我們猜想

$$\frac{1}{N}=\frac{1}{N+1}+\frac{1}{N(N+1)}.$$

要證明這條式，只需把右邊通分：

$$\frac{1}{N+1}+\frac{1}{N(N+1)}=\frac{N+1}{N(N+1)}=\frac{1}{N}.$$

兩個分母都是大於 1 的整數，所以兩項都是單位分數（unit fraction）。又因為當 $N>1$ 時 $N(N+1)>N+1$，所以兩者相異。因此，每個單位分數都可以寫成兩個相異單位分數之和。

N 為質數時的唯一性

把 $\frac{1}{N} = \frac{1}{a} + \frac{1}{b}$ 兩邊乘以 abN，得 $ab = Nb + Na$，即

$$ab - Na - Nb = 0 \quad \Longleftrightarrow \quad ab - Na - Nb + N^2 = N^2 \quad \Longleftrightarrow \quad (a-N)(b-N) = N^2.$$

每一步都可以逆推（因為 a、b、N 均不為零），所以 $(*)$ 的解恰好就是 $(a-N)(b-N) = N^2$ 的解。

> **想一想** $a - N$ 和 $b - N$ 有沒有可能同時為負數？

不可能。由於 $\frac{1}{b} > 0$，我們有 $\frac{1}{a} < \frac{1}{N}$，所以 $a > N$；同理 $b > N$。因此 $a - N$ 和 $b - N$ 是乘積為 N^2 的正整數。（這一點值得留意：若沒有這一步，$(-1, -N^2)$ 這一對便會是另一個需要排除的因數分解。）

現在設 $N = p$ 為質數。p^2 的正因數（divisor）只有 1、p 和 p^2，所以 $(a-p, b-p)$ 這一對必定是以下其中之一：

$$(1, p^2), \qquad (p, p), \qquad (p^2, 1).$$

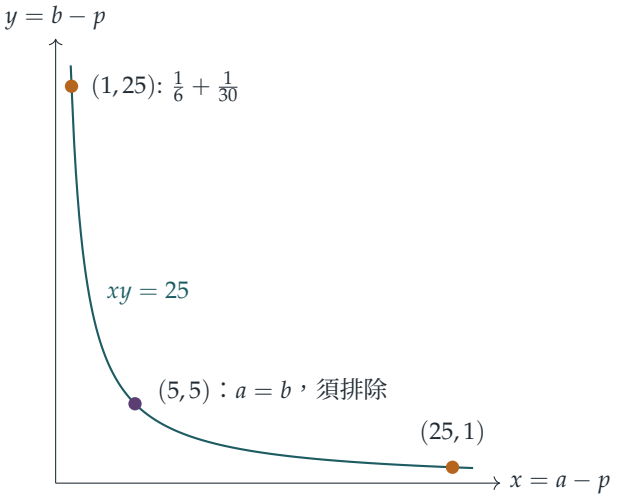
- (p, p) 給出 $a = b = 2p$，即 $\frac{1}{p} = \frac{1}{2p} + \frac{1}{2p}$：兩個單位分數並不相異，所以這個情況須排除。
- $(1, p^2)$ 給出 $a = p + 1$，$b = p^2 + p = p(p+1)$，這正是上面找到的表示式；$(p^2, 1)$ 則是同一對數，只是次序對調。

所以，若不計兩項的次序，唯一的寫法是

$$\frac{1}{p} = \frac{1}{p+1} + \frac{1}{p(p+1)},$$

這正是所要證明的。

下圖說明了箇中原因：這些解就是雙曲線（hyperbola）$xy = p^2$ 在正象限內的格點（lattice point），其中 $x = a - p$，$y = b - p$。圖中 $p = 5$。



形如 $2/N$ 的分數

現在假設 $\frac{2}{N} = \frac{1}{a} + \frac{1}{b}$，其中 $N = p$ 是奇質數。兩邊乘以 abp，得 $2ab = pa + pb$。要配成矩形，我們先乘以 2：

$$4ab - 2pa - 2pb = 0 \iff 4ab - 2pa - 2pb + p^2 = p^2 \iff (2a - p)(2b - p) = p^2.$$

> **想一想** 為甚麼要先乘以 2 才作因式分解？

若直接把 $2ab - pa - pb$ 因式分解，便會出現 $(a - \frac{p}{2})$ 這類項，而它們並不是整數。乘以 2 之後，一切都保持為整數，所以因式 $2a - p$ 和 $2b - p$ 都是整數，我們便可以運用整除性（divisibility）。

與之前一樣，$\frac{1}{a} < \frac{2}{p}$，所以 $2a > p$，即 $2a - p > 0$；同理 $2b - p > 0$。p^2 的正因數是 1、p、p^2，所以 $(2a - p,\ 2b - p)$ 是以下其中之一：

$$(1,\ p^2), \qquad (p,\ p), \qquad (p^2,\ 1).$$

- (p, p) 給出 $a = b = p$，即 $\frac{2}{p} = \frac{1}{p} + \frac{1}{p}$：兩者並不相異，須排除。
- $(1, p^2)$ 給出

  $$a = \frac{p+1}{2}, \qquad b = \frac{p^2 + p}{2} = \frac{p(p+1)}{2}.$$

  由於 p 是奇數，$p + 1$ 是偶數，所以兩者都是整數。由於 $p \geq 3$，$a \geq 2$，所以兩者都大於 1；而 $b = pa > a$，所以兩者相異。$(p^2, 1)$ 的情況是同一對數，只是次序對調。

直接驗證：

$$\frac{2}{p+1} + \frac{2}{p(p+1)} = \frac{2(p+1)}{p(p+1)} = \frac{2}{p}.$$

所以 $\frac{2}{p}$ 可以寫成兩個相異單位分數之和，而且（不計次序）唯一的寫法是

$$\frac{2}{p} = \frac{1}{\frac{1}{2}(p+1)} + \frac{1}{\frac{1}{2}p(p+1)},$$

這就證明了最後的結果。

驗算

- $p=3$：$\frac{2}{3} = \frac{1}{2} + \frac{1}{6}$；$p=5$：$\frac{2}{5} = \frac{1}{3} + \frac{1}{15}$；$p=7$：$\frac{2}{7} = \frac{1}{4} + \frac{1}{28}$。
- 以電腦搜尋所有 $N < 300$，結果發現：對每個質數 N，$\frac{1}{N}$ 恰好有一種表示法；對每個奇質數 N，$\frac{2}{N}$ 也恰好有一種表示法，與上面的公式吻合。
- 質數這個條件確實重要：$N = 6$ 有四種表示法，例如 $\frac{1}{6} = \frac{1}{7} + \frac{1}{42} = \frac{1}{8} + \frac{1}{24} = \frac{1}{9} + \frac{1}{18} = \frac{1}{10} + \frac{1}{15}$，每一種對應一個因數分解 $36 = 1\times 36 = 2\times 18 = 3\times 12 = 4\times 9$（$6\times 6$ 這一對則給出 $a=b$）。
- 為甚麼要 $p>2$：當 $p=2$ 時，$\frac{2}{2}=1$ 不能寫成 $\frac{1}{a}+\frac{1}{b}$ 的形式（其中 $a,b>1$ 且相異），而 $(p+1)/2$ 也確實不是整數。

常見錯誤

- 忘記證明 $a-N>0$ 和 $b-N>0$，因而忽略了 N^2 的負因數分解。
- 沒有排除 $N\times N$ 這個因數分解，它給出的是兩個相等的單位分數。
- 在 $2/N$ 的情況中，沒有檢查 $\frac{1}{2}(p+1)$ 是否整數（這正是用到 $p\neq 2$ 的地方）。

STEP 之外

把分數寫成相異單位分數之和，可追溯至古埃及人，而「埃及分數」（Egyptian fractions）至今仍有未解決的問題：艾狄胥–施特勞斯猜想（Erdős–Straus conjecture）指出，對所有 $N\geq 2$，$\frac{4}{N}$ 總可以寫成三個單位分數之和，但至今無人能夠證明。把 $xy - Nx - Ny = 0$ 化為 $(x-N)(y-N) = N^2$ 的技巧，是處理丟番圖方程（Diophantine equation）的標準工具，而數算解的個數就變成數算因數的個數。Davenport 的 *The Higher Arithmetic*（一本數論入門書）是進一步認識因數與質數的上佳讀物。

詳細解答

STEP II 2000，第 2 題

純數學 • 多項式的重根

STEP II 2000，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/00-s2-q2

關鍵思路

重複因式（repeated factor）經過求導後仍然保留，只是每求導一次，冪次便減一。因此，四次冪的因式 $(x-a)^4$ 會迫使 p 及其首三階導數（derivative）在 a 處全部等於零。對於這個六次多項式（sextic），三階導數只是一個三次式，它的根給出一張很短的 a 的候選名單。我們逐一檢驗這些候選值，最後才決定 k，因為一求導，k 便會消失。

解答

第一個結果

假設 $(x-a)^2$ 是因式，即對某個多項式（polynomial）q 有 $\mathrm{p}(x)=(x-a)^2\,\mathrm{q}(x)$。由積法則（product rule），

$$\mathrm{p}'(x)=2(x-a)\,\mathrm{q}(x)+(x-a)^2\,\mathrm{q}'(x)=(x-a)\big[\,2\mathrm{q}(x)+(x-a)\mathrm{q}'(x)\,\big].$$

代入 $x=a$ 便得 $\mathrm{p}'(a)=0$，正是所求。

相應的結果

同樣的計算適用於任何冪次。若 $\mathrm{f}(x)=(x-a)^m\,\mathrm{g}(x)$，其中 $m\geq 1$，而 g 是多項式，則

$$\mathrm{f}'(x)=(x-a)^{m-1}\big[\,m\,\mathrm{g}(x)+(x-a)\,\mathrm{g}'(x)\,\big],$$

而方括號內的式子仍然是多項式。所以若 $(x-a)^m$ 整除 f，則 $(x-a)^{m-1}$ 整除 f'。

現設 $(x-a)^4$ 是 $\mathrm{p}(x)$ 的因式。把上述事實連用三次，

$$(x-a)^3 \text{ 整除} \mathrm{p}'(x),\qquad (x-a)^2 \text{ 整除} \mathrm{p}''(x),\qquad (x-a) \text{ 整除} \mathrm{p}'''(x).$$

再加上 $(x-a)^4$ 整除 $\mathrm{p}(x)$ 本身，代入 $x=a$ 便得

$$\mathrm{p}(a)=\mathrm{p}'(a)=\mathrm{p}''(a)=\mathrm{p}'''(a)=0.$$

求 k

設 $\mathrm{p}(x)=x^6+4x^5-5x^4-40x^3-40x^2+32x+k$。求導三次，

$$\begin{aligned}\mathrm{p}'(x)&=6x^5+20x^4-20x^3-120x^2-80x+32,\\ \mathrm{p}''(x)&=30x^4+80x^3-60x^2-240x-80=10\,(3x^4+8x^3-6x^2-24x-8),\\ \mathrm{p}'''(x)&=120x^3+240x^2-120x-240=120\,(x^3+2x^2-x-2).\end{aligned}$$

想一想 四個條件之中，應該先用哪一個？

先用最容易解的那一個。常數 k 已從 p''' 中消失，而且 p''' 只是三次式，可用分組法因式分解（factorise）：

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

所以根據相應的結果，a 必定是 -2、1 或 -1。我們把它們逐一代入 p'' 檢驗：

$$\begin{aligned}\tfrac{1}{10}\mathrm{p}''(1)&=3+8-6-24-8=-27\neq 0,\\ \tfrac{1}{10}\mathrm{p}''(-1)&=3-8-6+24-8=5\neq 0,\\ \tfrac{1}{10}\mathrm{p}''(-2)&=48-64-24+48-8=0.\end{aligned}$$

只有 $a=-2$ 過關。它同樣通過 p' 的檢驗：

$$\mathrm{p}'(-2)=-192+320+160-480+160+32=0.$$

最後，由 $\mathrm{p}(-2)=0$ 便可確定 k：

$$\mathrm{p}(-2)=64-128-80+320-160-64+k=k-48=0,\qquad \text{所以}\qquad \boxed{k=48}.$$

想一想 我們只用了*必要*條件。怎樣知道當 $k=48$ 時，$(x+2)^4$ 真的是因式？

事實上，$\mathrm{p}(a)=\mathrm{p}'(a)=\mathrm{p}''(a)=\mathrm{p}'''(a)=0$ 這四個條件確實蘊含 $(x-a)^4$ 是因式（由多項式的泰勒定理（Taylor's theorem）可得）。不過，最快捷的確認方法是直接乘開。由於 $(x+2)^4=x^4+8x^3+24x^2+32x+16$，商式必定是首項係數為一（monic）的二次式，其常數項為 $48/16=3$。比較 x^5 的係數，得出 x^2-4x+3，而且確實有

$$(x+2)^4(x^2-4x+3)=x^6+4x^5-5x^4-40x^3-40x^2+32x+48.$$

所以，當 $k = 48$ 時，

$$x^6 + 4x^5 - 5x^4 - 40x^3 - 40x^2 + 32x + 48 = (x+2)^4(x-1)(x-3),$$

而這個因式符合所要求的形式，其中 $a = -2$。

驗算

- 把 $(x+2)^4(x-1)(x-3)$ 完全展開：每個係數都與題目給出的多項式相符，而常數項是 $16 \times 3 = 48$。
- 檢驗係數之和：$\mathrm{p}(1) = 1 + 4 - 5 - 40 - 40 + 32 + 48 = 0$，這是必然的，因為 $(x-1)$ 是因式。
- 我們用電腦在隨機選取的多項式 $(x-a)^4\mathrm{q}(x)$ 上驗證了一般結果，求得 p''' 的根為 $-2, \pm 1$，而 p'' 只在 -2 處等於零，並得出 $k = 48$。

常見錯誤

- 從 $\mathrm{p}(a) = 0$ 或 $\mathrm{p}'(a) = 0$ 入手。這兩個條件牽涉五次式或 k 本身，難以迅速得出結果。應從次數最低的導數開始。
- 找到 p''' 的根便停下來。候選值 $a = \pm 1$ 是 p''' 的根，卻不是 p'' 的根。
- 只把相應的結果寫成「$\mathrm{p}'''(a) = 0$」。應寫出完整的一組條件 $\mathrm{p}(a) = \mathrm{p}'(a) = \mathrm{p}''(a) = \mathrm{p}'''(a) = 0$，並加以證明。

STEP 之外

逆命題（converse）同樣成立：$(x-a)^m$ 整除 p，當且僅當 p 及其首 $m-1$ 階導數在 a 處都等於零。這是因為多項式等於它自身在 a 處的泰勒展開式。在大學一年級的代數課程中，同一想法用來定義根的重數（multiplicity）。它還提供了一個判別重根的方法：用歐幾里得算法（Euclid's algorithm）計算 $\gcd(\mathrm{p}, \mathrm{p}')$。在某些域（field）上，例如模質數的整數，形式導數（formal derivative）會有出人意表的表現。Allenby 的 *Rings, Fields and Groups*（環、域與羣）對此有詳細闡述。

詳細解答

STEP II 2000，第 3 題

純數學 • 誤差傳遞

> **STEP II 2000，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/00-s2-q3

關鍵思路

餘弦定理（cosine rule）把 a^2 準確地用 b、c 和 A 表示出來，所以整道題要問的，就是當這三個量各自稍作變動時，a 會怎樣隨之改變。把每個組成部分都按微小量展開至一階（first order），捨去微小量之間的乘積，最後利用 $\sqrt{1+x} \approx 1 + \frac{1}{2}x$ 開平方。至於取值範圍，留意 η 是若干個獨立誤差的線性組合（linear combination），所以當每一項都取到極端值，而且正負號都取得有利時，它便達到最大。

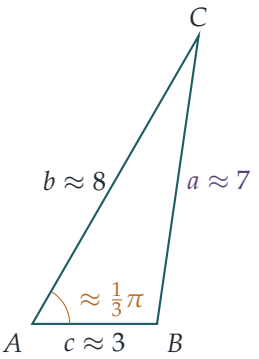
解答

a 的近似式

取未受擾動的數值 $b = 8$、$c = 3$、$A = \frac{1}{3}\pi$，餘弦定理給出

$$a^2 = 64 + 9 - 2 \cdot 8 \cdot 3 \cdot \tfrac{1}{2} = 49,$$

所以 $a = 7$。這就解釋了首項的來歷。



現在加入擾動。以下「≈」一律表示只保留 $\epsilon_1, \epsilon_2, \epsilon_3$ 的一階項，而捨去所有二次或更高次的項，例如 ϵ_1^2 或 $\epsilon_2\epsilon_3$。於是

$$\begin{aligned} b^2 &= 64 + 16\epsilon_1 + \epsilon_1^2 \approx 64 + 16\epsilon_1, \\ c^2 &= 9 + 6\epsilon_2 + \epsilon_2^2 \approx 9 + 6\epsilon_2, \\ bc &= 24 + 3\epsilon_1 + 8\epsilon_2 + \epsilon_1\epsilon_2 \approx 24 + 3\epsilon_1 + 8\epsilon_2. \end{aligned}$$

至於角度，由複角公式（compound-angle formula）和小角近似（small-angle approximation）$\cos\epsilon_3 \approx 1$、$\sin\epsilon_3 \approx \epsilon_3$，可得

$$\cos A = \cos\tfrac{1}{3}\pi\cos\epsilon_3 - \sin\tfrac{1}{3}\pi\sin\epsilon_3 \approx \tfrac{1}{2} - \tfrac{\sqrt{3}}{2}\epsilon_3.$$

（$\cos\epsilon_3 \approx 1$ 的誤差屬於 ϵ_3^2 的數量級，所以在這一階的精確度下確實可以忽略。）把括號乘開，並再次捨去微小量之間的乘積，

$$2bc\cos A \approx 2\big(24 + 3\epsilon_1 + 8\epsilon_2\big)\Big(\tfrac{1}{2} - \tfrac{\sqrt{3}}{2}\epsilon_3\Big) \approx 24 + 3\epsilon_1 + 8\epsilon_2 - 24\sqrt{3}\,\epsilon_3.$$

把各部分代入餘弦定理，

$$\begin{aligned} a^2 &\approx (64 + 16\epsilon_1) + (9 + 6\epsilon_2) - (24 + 3\epsilon_1 + 8\epsilon_2 - 24\sqrt{3}\,\epsilon_3) \\ &= 49 + 13\epsilon_1 - 2\epsilon_2 + 24\sqrt{3}\,\epsilon_3. \end{aligned}$$

設 $\delta = 13\epsilon_1 - 2\epsilon_2 + 24\sqrt{3}\,\epsilon_3$，這是一個微小量。由於 $a > 0$，

$$a = \sqrt{49 + \delta} = 7\Big(1 + \frac{\delta}{49}\Big)^{1/2} \approx 7\Big(1 + \frac{\delta}{98}\Big) = 7 + \frac{\delta}{14},$$

這裏用了二項式展開（binomial expansion）$(1+x)^{1/2} \approx 1 + \frac{1}{2}x$，適用於微小的 x。因此 $a \approx 7 + \eta$，其中

$$\eta = \frac{13\epsilon_1 - 2\epsilon_2 + 24\sqrt{3}\,\epsilon_3}{14},$$

這正是所求的結果。

想一想 我們能否避免逐項手算這些乘積的展開？

可以：對 $a^2 = b^2 + c^2 - 2bc\cos A$ 求微分，得到 $2a\,\delta a = (2b - 2c\cos A)\,\delta b + (2c - 2b\cos A)\,\delta c + 2bc\sin A\,\delta A$。在 $a = 7$、$b = 8$、$c = 3$、$A = \frac{1}{3}\pi$ 時，上式變成 $14\,\delta a = 13\,\delta b - 2\,\delta c + 24\sqrt{3}\,\delta A$，一行便得出相同的答案。兩種方法其實都只是一階泰勒展開（Taylor expansion）。

η 的取值範圍

試卷上 ϵ_2 的界限印作 $4 \cdot 9 \times 10^{-2}$，用的是舊式的居中小數點，所以它的意思是 $|\epsilon_2| \le 4.9 \times 10^{-2}$。

三個誤差各自獨立地變化，所以由三角不等式（triangle inequality），

$$|\eta| \le \frac{13|\epsilon_1| + 2|\epsilon_2| + 24\sqrt{3}\,|\epsilon_3|}{14} \le \frac{13 \times 2 \times 10^{-3} + 2 \times 4.9 \times 10^{-2} + 24\sqrt{3} \times \sqrt{3} \times 10^{-3}}{14}.$$

分子是 $0.026 + 0.098 + 0.072 = 0.196$，所以

$$|\eta| \le \frac{0.196}{14} = 0.014.$$

這個界限是可以達到的：取 $\epsilon_1 = 2 \times 10^{-3}$、$\epsilon_2 = -4.9 \times 10^{-2}$、$\epsilon_3 = \sqrt{3} \times 10^{-3}$，便令 η 的每一項都是正數，而且盡可能大，得出 $\eta = 0.014$；把三個正負號全部反過來，則得出 $\eta = -0.014$。由於 η 是各誤差的連續函數（continuous function）（其實更是線性函數），而容許的誤差所組成的集合是一個長方體，所以兩者之間的每一個值也都能取到。因此取值範圍是

$$-0.014 \le \eta \le 0.014,$$

也就是說，a 大約落在 7 ± 0.014 之內。

驗算

- 當所有誤差都是零時，三角形的三邊是 8、3、7，並有一個 60° 的角，而確實有 $64 + 9 - 24 = 49$。
- 取上述的極端誤差，準確的餘弦定理給出 $a \approx 7.0141$ 和 $a \approx 6.9861$，與 7 ± 0.014 非常接近；當中的微小差異來自被捨去的二階項。
- 以數值計算，a 對 A 的偏導數（partial derivative）在基準值處是 $2.9692\ldots = 24\sqrt{3}/14$，與預期相符。

常見錯誤

- 忘記最後要開平方：$a^2 \approx 49 + \delta$ 給出的是 $a \approx 7 + \delta/14$，而不是 $7 + \delta$。
- 求最大值時把三個誤差都取正數：ϵ_2 前面帶負號，所以它必須取負數才能把 η 推高。
- 把 $4 \cdot 9 \times 10^{-2}$ 讀成 36×10^{-2}；對於一條長度為 3 的邊來說，這根本算不上是微小的誤差。

STEP 之外

這就是透過全微分（total differential）進行的誤差傳播（error propagation）：對於多變量函數 f，有 $\delta f \approx \sum (\partial f/\partial x_i)\,\delta x_i$，而最壞的情況是 $\sum |\partial f/\partial x_i|\,|\delta x_i|$。在大學一年級的課程中，這會發展成多變量的泰勒定理（Taylor's theorem），並附有二階餘項（remainder）的明確界限，讓你確切知道「$\approx$」到底有多準確。實驗科學家則往往改用適用於獨立隨機誤差的方和根法則（root-sum-of-squares rule）。Bartle 與 Sherbert 的 *Introduction to Real Analysis*（實分析入門）對帶餘項的泰勒定理有嚴謹細緻的講解。

詳細解答

STEP II 2000，第 4 題

純數學 • 輻角與反正切恆等式

> **STEP II 2000，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/00-s2-q4

關鍵思路

頭兩個結果說明：複數（complex number）相乘時，它們的輻角（argument）相加。因此，形如「若干個反正切（arctangent）之和 $= \pi/4$」的恆等式，其實是關於 $5-\mathrm{i}$ 和 $1+\mathrm{i}$ 這類高斯整數（Gaussian integer）之乘積的命題：每個滿足 $a>0$ 的因子 $a+b\mathrm{i}$ 貢獻一個角 $\arctan(b/a)$，而直線 $y=x$ 上任何一點的輻角都是 $\pi/4$。要做的工作是小心地把乘積展開，然後驗證兩個角確實相等，而不僅僅是相差 2π 的某個倍數。

解答

乘積公式與棣美弗定理

展開並利用 $\mathrm{i}^2=-1$，得

$$\begin{aligned}(\cos\theta+\mathrm{i}\sin\theta)(\cos\phi+\mathrm{i}\sin\phi) &= (\cos\theta\cos\phi-\sin\theta\sin\phi)+\mathrm{i}(\sin\theta\cos\phi+\cos\theta\sin\phi)\\ &= \cos(\theta+\phi)+\mathrm{i}\sin(\theta+\phi),\end{aligned}$$

其中用到 $\cos(A+B)$ 和 $\sin(A+B)$ 的複角公式（compound-angle formulae）。

第二個結果（即棣美弗定理，de Moivre's theorem）我們對 n 用數學歸納法（induction）證明。當 $n=1$ 時，兩邊都是 $\cos\theta+\mathrm{i}\sin\theta$。假設結果對 $n=k$ 成立，則

$$\begin{aligned}(\cos\theta+\mathrm{i}\sin\theta)^{k+1} &= (\cos k\theta+\mathrm{i}\sin k\theta)(\cos\theta+\mathrm{i}\sin\theta)\\ &= \cos(k\theta+\theta)+\mathrm{i}\sin(k\theta+\theta),\end{aligned}$$

這裏先用歸納假設（induction hypothesis），再用取 $\phi=k\theta$ 的乘積公式。所以結果對 $n=k+1$ 成立，因而對每個正整數 n 都成立。

為方便起見，我們把 $\cos\theta + \mathrm{i}\sin\theta$ 記作 $\operatorname{cis}\theta$。於是我們已證明 $\operatorname{cis}\theta\operatorname{cis}\phi = \operatorname{cis}(\theta+\phi)$ 和 $(\operatorname{cis}\theta)^n = \operatorname{cis} n\theta$。

關於角的一個事實

若 $a > 0$ 而 b 為實數，則

$$a + b\mathrm{i} = \sqrt{a^2+b^2}\,\operatorname{cis}\big(\arctan(b/a)\big),$$

這是因為角 $\alpha = \arctan(b/a)$ 位於 $(-\frac{\pi}{2}, \frac{\pi}{2})$ 之內，所以 $\cos\alpha > 0$，而且 $\tan\alpha = b/a$；這兩個條件決定了 $\cos\alpha = a/\sqrt{a^2+b^2}$ 及 $\sin\alpha = b/\sqrt{a^2+b^2}$。

此外，若 $R\operatorname{cis}S = R'\operatorname{cis}T$ 且 $R, R' > 0$，則取模（modulus）可得 $R = R'$，所以 $\cos S = \cos T$ 及 $\sin S = \sin T$，因此 $S - T$ 是 2π 的倍數。

第一個恆等式

設 $\alpha = \arctan\frac{1}{5}$。則 $5 - \mathrm{i} = \sqrt{26}\operatorname{cis}(-\alpha)$ 及 $1 + \mathrm{i} = \sqrt{2}\operatorname{cis}\frac{\pi}{4}$，所以由上面的結果得

$$(5-\mathrm{i})^2(1+\mathrm{i}) = 26\sqrt{2}\,\operatorname{cis}\big(\tfrac{\pi}{4} - 2\alpha\big).$$

另一方面，直接展開可得

$$(5-\mathrm{i})^2 = 24 - 10\mathrm{i}, \qquad (24-10\mathrm{i})(1+\mathrm{i}) = 24 + 24\mathrm{i} - 10\mathrm{i} + 10 = 34 + 14\mathrm{i},$$

而 $34 + 14\mathrm{i} = \sqrt{1352}\operatorname{cis}\beta$，其中 $\beta = \arctan\frac{14}{34} = \arctan\frac{7}{17}$。兩相比較，$\frac{\pi}{4} - 2\alpha$ 與 β 相差 2π 的某個倍數。

> **想一想** 我們怎樣知道這個 2π 的倍數是零？

我們把兩個角的範圍都確定下來。由於 $0 < \frac{1}{5} < 1$，我們有 $0 < \alpha < \frac{\pi}{4}$，所以 $-\frac{\pi}{4} < \frac{\pi}{4} - 2\alpha < \frac{\pi}{4}$。另外 $0 < \beta < \frac{\pi}{2}$。$(-\frac{\pi}{4}, \frac{\pi}{2})$ 內的兩個角相差必定小於 2π，所以該倍數是零，即 $\beta = \frac{\pi}{4} - 2\alpha$。也就是說，

$$\arctan\tfrac{7}{17} + 2\arctan\tfrac{1}{5} = \frac{\pi}{4},$$

這正是所要證明的。

第二個恆等式

同樣的想法提示我們考慮 $(4+\mathrm{i})^3(20+\mathrm{i})(1985+\mathrm{i})$，它的輻角應該就是等式的左邊。逐步相乘：

$$\begin{aligned}(4+\mathrm{i})^2 &= 15 + 8\mathrm{i},\\ (4+\mathrm{i})^3 &= (15+8\mathrm{i})(4+\mathrm{i}) = 52 + 47\mathrm{i},\\ (4+\mathrm{i})^3(20+\mathrm{i}) &= (52+47\mathrm{i})(20+\mathrm{i}) = 993 + 992\mathrm{i}.\end{aligned}$$

這已經非常接近直線 $y = x$，而最後一個因子正好完成餘下的工作：

$$\begin{aligned}(993 + 992\mathrm{i})(1985 + \mathrm{i}) &= (993 \times 1985 - 992) + (993 + 992 \times 1985)\mathrm{i} \\ &= 1970113 + 1970113\,\mathrm{i} = 1970113\sqrt{2}\,\operatorname{cis}\tfrac{\pi}{4}.\end{aligned}$$

（如果你好奇 1985 從何而來：要求 $(993 + 992\mathrm{i})(a + \mathrm{i})$ 的實部與虛部相等，便得 $993a - 992 = 992a + 993$，所以 $a = 1985$。）

現在設 $p = \arctan\frac{1}{4}$、$q = \arctan\frac{1}{20}$ 及 $r = \arctan\frac{1}{1985}$。根據上面關於角的事實，$4 + \mathrm{i}$、$20 + \mathrm{i}$、$1985 + \mathrm{i}$ 分別是 $\operatorname{cis} p$、$\operatorname{cis} q$、$\operatorname{cis} r$ 的正實數倍，所以它們的乘積是 $\operatorname{cis}(3p + q + r)$ 的正實數倍。與上一行比較，$3p + q + r - \frac{\pi}{4}$ 是 2π 的倍數。

p, q, r 各自位於 $(0, \frac{\pi}{4})$ 之內，因為它們的正切值都位於 $(0, 1)$ 之內，所以 $0 < 3p + q + r < \frac{5\pi}{4}$。因此 $3p + q + r$ 與 $\frac{\pi}{4}$ 相差小於 2π，該倍數是零，於是

$$3\arctan\tfrac{1}{4} + \arctan\tfrac{1}{20} + \arctan\tfrac{1}{1985} = \frac{\pi}{4}.$$

驗算

模必須吻合：$|34 + 14\mathrm{i}|^2 = 1156 + 196 = 1352 = 26^2 \times 2$，正如所料。至於第二個乘積，$17^3 \times 401 \times (1985^2 + 1)$ 應該等於 2×1970113^2，而事實確是如此。在數值上，用計算機計算，兩個等式的左邊都等於 $0.785398163\ldots = \pi/4$；例如 $\arctan\frac{7}{17} \approx 0.390607$ 而 $\frac{\pi}{4} - 2\arctan\frac{1}{5} \approx 0.390607$。

常見錯誤

- 由複數相等直接推斷角相等，而沒有排除兩者相差 $2\pi k$ 的可能。關於範圍的論證是證明中不可或缺的一環。
- 計算 $(5 - \mathrm{i})^2$ 時出現正負號錯誤：它等於 $24 - 10\mathrm{i}$，所以 $5 - \mathrm{i}$ 貢獻的是 $-\arctan\frac{1}{5}$，這就是為甚麼該反正切移到另一邊後帶着正號。
- 計算大數乘積時出現算術錯誤；可如上文所述，用模來驗算。

STEP 之外

這類恆等式稱為梅欽類公式（Machin-type formulae）；梅欽（John Machin）的 $\frac{\pi}{4} = 4\arctan\frac{1}{5} - \arctan\frac{1}{239}$ 是用同樣的方法由 $(5 + \mathrm{i})^4(239 - \mathrm{i})$ 得出的，並在 1706 年用來把 π 計算到小數點後一百位。其背後的結構是高斯整數 $\mathbb{Z}[\mathrm{i}]$ 的算術，當中分解為高斯質數（Gaussian primes）的唯一分解性（unique factorisation）解釋了哪些這類恆等式有可

能存在。想了解乘法作為旋轉的幾何意義，可閱讀 Needham 的 *Visual Complex Analysis*（以圖像方式講解複分析）；想了解高斯整數，可閱讀 Davenport 的 *The Higher Arithmetic*（數論入門）。

詳細解答

STEP II 2000，第 5 題

純數學 • 最小二乘近似

> **STEP II 2000，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/00-s2-q5

關鍵思路

把平方展開之後，要求最小值的量其實只是 λ 的二次式（quadratic），其係數都是固定的數（涉及 f 的積分）。只要配方（completing the square），就能一次過求出最佳的 λ 和最小的誤差。至於正弦函數的例子，要點在於當 n 很大時，$\sin(\pi x/n)$ 幾乎就是一條通過原點的直線，而小角近似（small-angle approximations）正好把這個說法精確化。

解答

λ 的最佳值

設

$$J = \int_0^1 x f(x)\,\mathrm{d}x, \qquad K = \int_0^1 \big(f(x)\big)^2\,\mathrm{d}x.$$

這兩個都是常數，與 λ 無關。把平方展開，並利用 $\int_0^1 x^2\,\mathrm{d}x = \frac{1}{3}$，得

$$\begin{aligned} I(\lambda) = \int_0^1 \big(f(x) - \lambda x\big)^2\,\mathrm{d}x &= K - 2\lambda J + \tfrac{1}{3}\lambda^2 \\ &= \tfrac{1}{3}\big(\lambda - 3J\big)^2 + K - 3J^2. \end{aligned}$$

平方項永不為負，而且只在 $\lambda = 3J$ 時等於零，所以 $I(\lambda)$ 恰好在

$$\lambda = 3\int_0^1 x f(x)\,\mathrm{d}x,$$

時取得最小值，這正是所求。（等價地，由 $I'(\lambda) = -2J + \frac{2}{3}\lambda = 0$ 可得 $\lambda = 3J$，而 $I''(\lambda) = \frac{2}{3} > 0$ 確認這是極小值。）

殘差

R^2 是 I 的最小值，即殘差（residual error）的平方；由配方後的式子可見，這個最小值是 $K - 3J^2$。由於 $J = \lambda/3$，我們有 $3J^2 = \frac{1}{3}\lambda^2$，所以

$$R^2 = \int_0^1 \big(f(x)\big)^2 \,\mathrm{d}x - \tfrac{1}{3}\lambda^2,$$

這正是所求。（如果你是用求導的方法找出 λ，可把 $J = \lambda/3$ 代入 $K - 2\lambda J + \frac{1}{3}\lambda^2$，得到 $K - \frac{2}{3}\lambda^2 + \frac{1}{3}\lambda^2$，結果相同。）

$f(x) = \sin(\pi x/n)$ 的情況

為了令代數運算較為簡潔，設 $a = \pi/n$，於是 $f(x) = \sin ax$，而且當 $n \to \infty$ 時 $a \to 0$。取 $u = x$ 和 $\mathrm{d}v = \sin ax\,\mathrm{d}x$ 作分部積分（integration by parts），

$$\begin{aligned}\int_0^1 x \sin ax \,\mathrm{d}x &= \Big[-\frac{x\cos ax}{a}\Big]_0^1 + \frac{1}{a}\int_0^1 \cos ax \,\mathrm{d}x \\ &= -\frac{\cos a}{a} + \frac{\sin a}{a^2},\end{aligned}$$

所以

$$\lambda = \frac{3(\sin a - a\cos a)}{a^2}.$$

? **想一想** 為甚麼 $\sin\theta$ 的近似式中需要保留 θ^3 項，而不能只用 $\sin\theta \approx \theta$？

如果只用 $\sin a \approx a$ 和 $\cos a \approx 1$，分子就變成 $a - a = 0$，這甚麼也沒有告訴我們：首項互相抵消了，答案藏在下一階的項之中。把提示中的兩個近似式都用上，

$$\sin a - a\cos a \approx \Big(a - \tfrac{1}{6}a^3\Big) - a\Big(1 - \tfrac{1}{2}a^2\Big) = \tfrac{1}{3}a^3,$$

因此當 n 很大時，

$$\lambda \approx \frac{3\cdot\frac{1}{3}a^3}{a^2} = a = \frac{\pi}{n}$$

這就是結果 (i)。

至於 (ii)，我們還需要 K。利用 $\sin^2 ax = \frac{1}{2}(1 - \cos 2ax)$，

$$K = \int_0^1 \sin^2 ax \,\mathrm{d}x = \frac{1}{2} - \frac{\sin 2a}{4a}.$$

當 $n \to \infty$ 時，$a \to 0$ 而 $\dfrac{\sin 2a}{2a} \to 1$，所以 $K \to \frac{1}{2} - \frac{1}{2} = 0$。同時 $\lambda \to 0$，因為 $\lambda \approx \pi/n$。故此

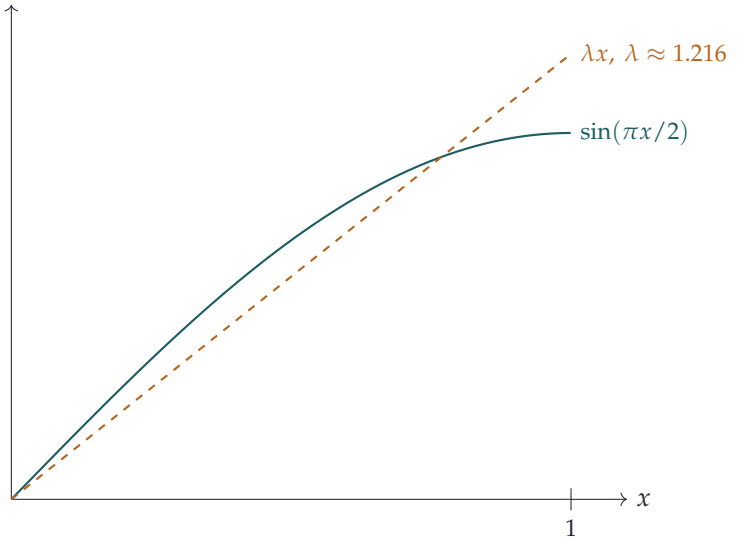
$$R^2 = K - \tfrac{1}{3}\lambda^2 \to 0 - 0 = 0,$$

又因為 $R \geq 0$，我們得到 $\lim_{n\to\infty} R = 0$。這就是結果 (ii)。

事實上，這裏的抵消遠比上述的更徹底。利用 $\sin 2a \approx 2a - \frac{4}{3}a^3$，我們得到 $K \approx \frac{1}{2} - \left(\frac{1}{2} - \frac{1}{3}a^2\right) = \frac{1}{3}a^2$，而 $\frac{1}{3}\lambda^2 \approx \frac{1}{3}a^2$ 亦然。所以 R^2 不單是小：與 K（即 f 本身的大小）相比，它也是小的。這條直線幾乎完全捕捉了 f。

為何這些結果在意料之中

在 $0 \le x \le 1$ 上，角 $\pi x/n$ 介乎 0 與 π/n 之間，當 n 很大時這個角很小。所以在整個區間上 $\sin(\pi x/n) \approx \pi x/n$：這個函數本身就非常接近線性函數 λx，其中 $\lambda = \pi/n$。因此，最佳擬合直線的斜率（gradient）應該接近 π/n，而擬合應該變得近乎完美，所以殘差應該趨向零。下圖顯示，即使 $n = 2$，曲線與直線也相差不遠。



驗算

- 如果本來就有 $f(x) = cx$，則 $\lambda = 3c\int_0^1 x^2\,\mathrm{d}x = c$ 而 $R^2 = \frac{1}{3}c^2 - \frac{1}{3}c^2 = 0$，正如所料。
- 以數值計算，當 $n = 10$ 時，精確公式給出 $\lambda = 0.31107$，而 $\pi/10 = 0.31416$；當 $n = 100$ 時，比值 $\lambda/(\pi/n)$ 為 0.99990。
- 當 $n = 10$ 時，數值積分（quadrature）給出 $R^2 \approx 6.0 \times 10^{-7}$，而 $K \approx 0.032$，所以與 f 相比，殘差確實微不足道。

常見錯誤

- 停在 $\sin a \approx a$、$\cos a \approx 1$ 而得到 $\lambda \approx 0$：這是對的，但毫無用處。要保留三次項。

▶ 忘記了 $\lambda = 3\int_0^1 xf(x)\,\mathrm{d}x$ 中的 3 是來自 $\int_0^1 x^2\,\mathrm{d}x = \frac{1}{3}$。

STEP 之外

這就是最小二乘逼近（least-squares approximation），而它本質上是幾何。把 $[0,1]$ 上的函數看成向量，並以 $\langle f,g\rangle = \int_0^1 fg\,\mathrm{d}x$ 作為「點積」（dot product）。那麼 λx 就是 f 在 x 所張成的直線上的正交投影（orthogonal projection），而 $R^2 = K - \frac{1}{3}\lambda^2$ 就是畢氏定理（Pythagoras' theorem）。把展開式再推進一步，可得 $R^2 \approx a^6/1575$，這來自把 x^3 項中沿 x 方向的分量投影掉。Strang 的 *Introduction to Linear Algebra*（線性代數入門）詳細闡述了投影與最小二乘法，而同樣的想法更可引伸至傅立葉級數（Fourier series）。

詳細解答

STEP II 2000，第 6 題

純數學 • 半角代換

STEP II 2000，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/00-s2-q6

關鍵思路

代換（substitution）$t = \tan\frac{1}{2}\theta$ 可以把任何由 $\sin\theta$ 和 $\cos\theta$ 組成的有理函數（rational function）化為 t 的有理函數，而在這一題，結果是 2 除以一個二次式。配方（completing the square）之後得到反正切（arctangent），而第一部分的第三個恆等式（identity）正好用來化簡上限處的反正切。第二個積分則可憑對稱性（symmetry）由第一個積分直接得出，毋須再做新的積分。

解答

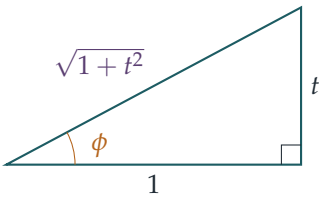
三個恆等式

設 $\phi = \frac{1}{2}\theta$，則 $t = \tan\phi$ 且 $1 + t^2 = \sec^2\phi$。由二倍角公式（double-angle formulae），

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

$$\cos\theta = \cos^2\phi - \sin^2\phi = \cos^2\phi\,(1 - \tan^2\phi) = \frac{1-t^2}{1+t^2}.$$

（只要 $\tan\frac{1}{2}\theta$ 有定義，即 $\cos\frac{1}{2}\theta \neq 0$，這兩條式子便成立。）下圖的直角三角形，兩條直角邊分別為 1 和 t，夾角為 ϕ，是記住這兩條公式的好方法。



至於第三個恆等式，當 $\sin\theta \neq 0$ 時，利用 $1+\cos\theta = 2\cos^2\phi$ 和 $\sin\theta = 2\sin\phi\cos\phi$：

$$\frac{1+\cos\theta}{\sin\theta} = \frac{2\cos^2\phi}{2\sin\phi\cos\phi} = \cot\phi = \tan\left(\tfrac{1}{2}\pi - \phi\right) = \tan\left(\tfrac{1}{2}\pi - \tfrac{1}{2}\theta\right),$$

因為對餘角（complementary angles）而言，$\cot\phi = \tan(\frac{1}{2}\pi - \phi)$。至此三個恆等式均已證明。

第一個積分

設 $t = \tan\frac{1}{2}\theta$，則 $\theta = 2\arctan t$，所以

$$\mathrm{d}\theta = \frac{2}{1+t^2}\,\mathrm{d}t.$$

當 $\theta = 0$ 時，$t = 0$；當 $\theta = \frac{1}{2}\pi$ 時，$t = \tan\frac{1}{4}\pi = 1$。設 $c = \cos\alpha$ 和 $s = \sin\alpha$；由於 $0 < \alpha < \frac{1}{2}\pi$，兩者皆為正數。於是

$$\int_0^{\frac{1}{2}\pi} \frac{\mathrm{d}\theta}{1+c\sin\theta} = \int_0^1 \frac{1}{1+\dfrac{2ct}{1+t^2}} \cdot \frac{2}{1+t^2}\,\mathrm{d}t = \int_0^1 \frac{2}{t^2+2ct+1}\,\mathrm{d}t.$$

配方，並利用 $1-c^2 = s^2$，得

$$t^2 + 2ct + 1 = (t+c)^2 + s^2.$$

由於 $s > 0$，由標準積分 $\int \frac{\mathrm{d}u}{u^2+s^2} = \frac{1}{s}\arctan\frac{u}{s}$ 可得

$$\int_0^1 \frac{2}{(t+c)^2+s^2}\,\mathrm{d}t = \frac{2}{s}\left[\arctan\frac{t+c}{s}\right]_0^1 = \frac{2}{s}\left(\arctan\frac{1+\cos\alpha}{\sin\alpha} - \arctan\frac{\cos\alpha}{\sin\alpha}\right).$$

想一想 怎樣化簡 $\arctan\dfrac{1+\cos\alpha}{\sin\alpha}$？

第一部分的功夫在這裏派上用場。由第三個恆等式（以 α 代替 θ），$\frac{1+\cos\alpha}{\sin\alpha} = \tan(\frac{1}{2}\pi - \frac{1}{2}\alpha)$。不過，在把 arctan 與 tan 互相抵消之前，必須先檢查這個角是否位於 arctan 的值域（range）$(-\frac{1}{2}\pi, \frac{1}{2}\pi)$ 之內。由於 $0 < \alpha < \frac{1}{2}\pi$，我們有 $\frac{1}{4}\pi < \frac{1}{2}\pi - \frac{1}{2}\alpha < \frac{1}{2}\pi$，所以確實有

$$\arctan\frac{1+\cos\alpha}{\sin\alpha} = \tfrac{1}{2}\pi - \tfrac{1}{2}\alpha.$$

同理，$\frac{\cos\alpha}{\sin\alpha} = \cot\alpha = \tan(\frac{1}{2}\pi - \alpha)$，而 $0 < \frac{1}{2}\pi - \alpha < \frac{1}{2}\pi$，所以 $\arctan\frac{\cos\alpha}{\sin\alpha} = \frac{1}{2}\pi - \alpha$。因此

$$\int_0^{\frac{1}{2}\pi} \frac{\mathrm{d}\theta}{1+\cos\alpha\sin\theta} = \frac{2}{\sin\alpha}\left(\left(\tfrac{1}{2}\pi - \tfrac{1}{2}\alpha\right) - \left(\tfrac{1}{2}\pi - \alpha\right)\right) = \frac{2}{\sin\alpha}\cdot\frac{\alpha}{2} = \frac{\alpha}{\sin\alpha},$$

此即所求。

第二個積分

代入 $\theta = \frac{1}{2}\pi - u$。於是 $\cos\theta = \sin u$，$\mathrm{d}\theta = -\mathrm{d}u$，而上下限 $0, \frac{1}{2}\pi$ 互換，剛好抵消負號：

$$\int_0^{\frac{1}{2}\pi} \frac{\mathrm{d}\theta}{1+\sin\alpha\cos\theta} = \int_0^{\frac{1}{2}\pi} \frac{\mathrm{d}u}{1+\sin\alpha\sin u}.$$

現在設 $\beta = \frac{1}{2}\pi - \alpha$。於是 $\sin\alpha = \cos\beta$，且 $0 < \beta < \frac{1}{2}\pi$，所以剛才證明的結果可以套用，只需以 β 代替 α：

$$\int_0^{\frac{1}{2}\pi} \frac{\mathrm{d}u}{1+\cos\beta\sin u} = \frac{\beta}{\sin\beta}.$$

由於 $\sin\beta = \cos\alpha$，我們得出結論：對於 $0 < \alpha < \frac{1}{2}\pi$，

$$\int_0^{\frac{1}{2}\pi} \frac{\mathrm{d}\theta}{1+\sin\alpha\cos\theta} = \frac{\frac{1}{2}\pi - \alpha}{\cos\alpha}.$$

驗算

- 當 $\alpha \to 0$ 時，$\cos\alpha \to 1$，第一個積分應趨向 $\int_0^{\pi/2} \frac{\mathrm{d}\theta}{1+\sin\theta} = 1$（作代換後得 $\int_0^1 \frac{2}{(1+t)^2}\,\mathrm{d}t = 1$）；而 $\alpha/\sin\alpha \to 1$，果然吻合。
- 當 $\alpha \to \frac{1}{2}\pi$ 時，被積函數（integrand）趨向 1，所以第一個積分應趨向 $\frac{1}{2}\pi$，與 $\frac{\pi/2}{1}$ 相符。同樣，當 $\alpha \to \frac{1}{2}\pi$ 時，第二個積分趨向 1，與 $\frac{\frac{1}{2}\pi-\alpha}{\cos\alpha} \to 1$ 一致。
- 數值上，當 $\alpha = \frac{1}{4}\pi$ 時，用辛普森法則（Simpson's rule）算得第一個積分為 1.110721，與 $\frac{\pi\sqrt{2}}{4} \approx 1.110721$ 相符；當 $\alpha = \frac{1}{3}\pi$ 時，第二個積分算得 $1.047198 \approx \frac{1}{3}\pi$，正如公式所預測。

常見錯誤

- 忘記 $\mathrm{d}\theta$ 所帶來的因子 $\frac{2}{1+t^2}$，或忘記把上下限改為 0 和 1。
- 未檢查 x 是否位於 $(-\frac{1}{2}\pi, \frac{1}{2}\pi)$ 之內，便寫下 $\arctan(\tan x) = x$；在這一題，正是限制條件 $0 < \alpha < \frac{1}{2}\pi$ 令這一步成立。
- 為第二個積分把整個積分過程重做一遍，而不利用 $\theta \mapsto \frac{1}{2}\pi - \theta$ 和 $\alpha \mapsto \frac{1}{2}\pi - \alpha$。

STEP 之外

代換 $t = \tan\frac{1}{2}\theta$（有時稱為魏爾斯特拉斯代換，Weierstrass substitution）之所以行得通，是因為點 $(\cos\theta, \sin\theta) = \left(\frac{1-t^2}{1+t^2}, \frac{2t}{1+t^2}\right)$ 是單位圓的一個有理參數化（rational parametrisation）；取 t 為有理數，便可得出所有畢氏三元數組（Pythagorean triple）。在大學裏，

形如 $\int_0^{2\pi} R(\cos\theta, \sin\theta)\,\mathrm{d}\theta$ 的積分，通常改為設 $z = e^{i\theta}$ 並運用留數定理（residue theorem）來計算，這樣就把問題轉化為找出單位圓內的極點（poles）。Needham 的 *Visual Complex Analysis*（《複分析：可視化方法》）是這個想法的一本精彩入門書。

詳細解答

STEP II 2000，第 7 題

純數學 • 直線與圓錐

STEP II 2000，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/00-s2-q7

關鍵思路

動手做任何事之前，先求出 $\mathbf{s}$ 的長度。$\cos\theta \pm \sqrt{3}$ 這些項是刻意安排的：交叉項會互相抵消，並出現 $\cos^2\theta + \sin^2\theta$，所以 $|\mathbf{s}|$ 完全與 θ 無關。之後的一切都來自純量積（scalar product）。$s_x - s_z$ 這個組合同樣不含 θ，這直接指出了 m 的方向，也解釋了為何最後一部分會出現平面 $x - z =$ 常數。

解答

l 與一般直線的夾角

為簡便起見，記 $c_\theta = \cos\theta$。$\mathbf{s}$ 的長度為

$$\begin{aligned} |\mathbf{s}|^2 &= (c_\theta + \sqrt{3})^2 + 2\sin^2\theta + (c_\theta - \sqrt{3})^2 \\ &= 2c_\theta^2 + 6 + 2\sin^2\theta = 8, \end{aligned}$$

因為交叉項 $\pm 2\sqrt{3}\,c_\theta$ 互相抵消。所以對任何 θ，都有 $|\mathbf{s}| = 2\sqrt{2}$。

設 $\mathbf{d} = a\mathbf{i} + b\mathbf{j} + c\mathbf{k}$。則

$$\mathbf{s}\cdot\mathbf{d} = a(c_\theta + \sqrt{3}) + \sqrt{2}\,b\sin\theta + c(c_\theta - \sqrt{3}) = (a+c)\cos\theta + \sqrt{2}\,b\sin\theta + \sqrt{3}\,(a-c).$$

因此兩條直線的夾角 ϕ 滿足

$$\cos\phi = \frac{(a+c)\cos\theta + \sqrt{2}\,b\sin\theta + \sqrt{3}\,(a-c)}{2\sqrt{2}\,\sqrt{a^2+b^2+c^2}}.$$

直線本身沒有指定的方向，所以如果你要的是銳角（acute angle），就取右邊的絕對值（modulus）。

直線 m

我們要找一個方向 $\mathbf{d}$，使 $\cos\phi$ 與 θ 無關。觀察分子可知，恰好當 $a+c=0$ 且 $b=0$ 時，對 θ 的依賴便會消失。所以試取 $\mathbf{d}=\mathbf{i}-\mathbf{k}$（$a=1$，$b=0$，$c=-1$）：

$$\cos\phi=\frac{\sqrt{3}\cdot 2}{2\sqrt{2}\cdot\sqrt{2}}=\frac{\sqrt{3}}{2}.$$

這個值是正數，所以 ϕ 本身已是銳角，而且對任何 θ 都有 $\phi=\pi/6$。因此，通過原點的直線

$$m:\quad \mathbf{r}=\mu(\mathbf{i}-\mathbf{k})$$

符合要求。當 θ 變化時，l 掃出一個圓錐（cone），其軸為 m，半頂角（semi-vertical angle）為 $\pi/6$。

P 描出的圓

在 l 上，$x-z=\lambda(s_x-s_z)=\lambda\big((c_\theta+\sqrt{3})-(c_\theta-\sqrt{3})\big)=2\sqrt{3}\,\lambda$。令此式等於 $4\sqrt{3}$，便得 $\lambda=2$，不論 θ 是甚麼，所以

$$\overrightarrow{OP}=2\mathbf{s}=(2\cos\theta+2\sqrt{3})\,\mathbf{i}+2\sqrt{2}\sin\theta\,\mathbf{j}+(2\cos\theta-2\sqrt{3})\,\mathbf{k}.$$

直線 m 與該平面相交於 $\mu-(-\mu)=4\sqrt{3}$ 之處，即 $\mu=2\sqrt{3}$，交點為 C，其中

$$\overrightarrow{OC}=2\sqrt{3}\,\mathbf{i}-2\sqrt{3}\,\mathbf{k}.$$

> ? **想一想** 為甚麼 C 是圓心的自然人選？

平面 $x-z=4\sqrt{3}$ 的法線（normal）是 $\mathbf{i}-\mathbf{k}$，正是 m 的方向。所以 m 是圓錐的軸，它垂直於該平面，而任何圓錐被垂直於其軸的平面所截，都會得到一個圓，圓心就在軸與該平面相交之處。現在我們直接驗證這一點：

$$\overrightarrow{CP}=2\cos\theta\,\mathbf{i}+2\sqrt{2}\sin\theta\,\mathbf{j}+2\cos\theta\,\mathbf{k},\qquad |\overrightarrow{CP}|^2=4\cos^2\theta+8\sin^2\theta+4\cos^2\theta=8.$$

所以 P 的每一個位置都在該平面上，並且與平面上的定點 C 相距 $2\sqrt{2}$。換句話說，P 位於該平面上以 C 為圓心、半徑為 $2\sqrt{2}$ 的圓上，而 C 在 m 上。

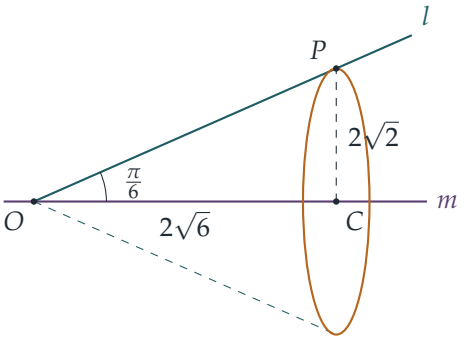
要看出 P 確實描出整個圓，可寫成

$$\overrightarrow{CP}=2\sqrt{2}\,(\cos\theta\,\mathbf{u}+\sin\theta\,\mathbf{v}),\qquad \mathbf{u}=\tfrac{1}{\sqrt{2}}(\mathbf{i}+\mathbf{k}),\quad \mathbf{v}=\mathbf{j}.$$

這裏 $\mathbf{u}$ 和 $\mathbf{v}$ 是互相垂直的單位向量（unit vector），兩者都垂直於 $\mathbf{i}-\mathbf{k}$，所以它們都平行於該平面。當 θ 由 0 變到 2π 時，P 沿圓走了剛好一圈。半徑為

$$\boxed{2\sqrt{2}}.$$

求半徑還有一個較快的方法：$|OP| = 2|\mathbf{s}| = 4\sqrt{2}$ 而 $|OC| = 2\sqrt{6}$，且 OC 垂直於該平面，所以由畢氏定理（Pythagoras）得 $CP = \sqrt{32-24} = 2\sqrt{2}$。等價地，$CP = OC\tan\frac{\pi}{6} = 2\sqrt{6}/\sqrt{3}$。



驗算

- 當 $\theta = 0$ 時，$\mathbf{s} = (1+\sqrt{3}, 0, 1-\sqrt{3})$，其中 $|\mathbf{s}|^2 = (4+2\sqrt{3}) + (4-2\sqrt{3}) = 8$，而 $\mathbf{s}\cdot(\mathbf{i}-\mathbf{k}) = 2\sqrt{3}$，得 $\cos\phi = \sqrt{3}/2$。
- 求半徑的兩種方法結果一致：$\sqrt{(4\sqrt{2})^2 - (2\sqrt{6})^2} = \sqrt{8}$ 及 $2\sqrt{6}\tan\frac{\pi}{6} = \sqrt{8} \approx 2.828$。
- 數值上，取 20 000 個隨機的 θ（一般公式則另取隨機的 a, b, c），與 m 的夾角等於 $\pi/6$，P 在該平面上，且 $|CP| = 2.8284\ldots$，誤差全部在 10^{-15} 以內。

常見錯誤

- 忘記除以 $|\mathbf{s}|$，或假設 $\mathbf{s}$ 是單位向量。它的長度是 $2\sqrt{2}$。
- 只證明了 P 在以 O 為球心的球面（sphere）上（即 $|OP| = 4\sqrt{2}$）。球面被平面所截固然得到一個圓，但你仍須說明為何圓心是 m 上的 C：這是因為 OC 垂直於該平面。

STEP 之外

直線 l 掃出一個圓錐，而用平面去截圓錐，正是圓錐曲線（conic sections）名稱的由來。垂直於軸的平面截出一個圓；把平面傾斜，便會得到橢圓（ellipse）、拋物線（parabola）或雙曲線（hyperbola）。在大學一年級的課程中，你會把圓錐描述為二次曲面（quadric surface）$(\mathbf{r}\cdot\hat{\mathbf{n}})^2 = |\mathbf{r}|^2\cos^2\alpha$，並利用對稱矩陣（symmetric matrix）的特徵值（eigenvalues）為這類曲面分類。Strang 的 *Introduction to Linear Algebra*（線性代數入門）涵蓋

了這方面的內容。

詳細解答

STEP II 2000，第 8 題

純數學 • 可分離變量微分方程

STEP II 2000，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/00-s2-q8

關鍵思路

兩條方程都是可分離變數（separable）的：含 x 的部分與含 y 的部分以乘積形式出現，所以可以把所有涉及 y 的項放在一邊，所有涉及 x 的項放在另一邊。兩題中 x 的一邊都具有 $x\,\mathrm{e}^{ax^2}$ 的形式，由於 x 是 x^2 的導數（derivative）（的倍數），所以可以立即積分。在第 (ii) 部分，$(y+3)$ 的冪次由我們自行選擇，而我們選擇的冪次要令 y 的增長恰好與 e^{3x^2} 相符。

解答

第 (i) 部分

由於在 $x=0$ 時 $y=6$，起初附近有 $y+3>0$，因此可以除以 $(y+3)^{1/2}$：

$$(y+3)^{-1/2}\frac{\mathrm{d}y}{\mathrm{d}x} = -4x\,\mathrm{e}^{-x^2}.$$

兩邊對 x 積分，並利用 $\frac{\mathrm{d}}{\mathrm{d}x}\mathrm{e}^{-x^2} = -2x\,\mathrm{e}^{-x^2}$，得

$$2(y+3)^{1/2} = 2\mathrm{e}^{-x^2} + C.$$

代入 $x=0$、$y=6$ 得 $2\cdot 3 = 2 + C$，所以 $C=4$，而

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

右邊恆不小於 2，所以它確實是一個合理的（正）平方根，而且 $y+3$ 永遠不會達到 0；因此上面的除法對所有 $x \geq 0$ 都成立。兩邊平方，得

$$\boxed{y = \left(2+\mathrm{e}^{-x^2}\right)^2 - 3 = 1 + 4\mathrm{e}^{-x^2} + \mathrm{e}^{-2x^2}}.$$

當 $x\to\infty$ 時，$\mathrm{e}^{-x^2}\to 0$ 且 $\mathrm{e}^{-2x^2}\to 0$，所以 $y\to 1$，正如題目所求。

第 (ii) 部分

這裏的方程是

$$(y+3)^{k-1}\frac{\mathrm{d}y}{\mathrm{d}x} = x\,\mathrm{e}^{6x^2}.$$

要令一般的（可能不是整數的）冪次有意義，我們需要 $y+3>0$；又由於當 $x>0$ 時 $\frac{\mathrm{d}y}{\mathrm{d}x}>0$，$y$ 是遞增的，所以只要 $y+3$ 起初是正數，它就會一直保持正數。當 $k\neq 0$ 時，積分得

$$\frac{(y+3)^k}{k} = \frac{\mathrm{e}^{6x^2}}{12} + C, \qquad \text{所以} \qquad (y+3)^k = \frac{k}{12}\,\mathrm{e}^{6x^2} + kC.$$

? **想一想** 我們應該嘗試哪一個 k？

若 $k>0$，則當 x 很大時，常數可以忽略，於是 $y+3\approx (k/12)^{1/k}\,\mathrm{e}^{6x^2/k}$。我們希望 y 的增長與 e^{3x^2} 相若，所以需要 $6/k=3$，即 $k=2$。（任何其他 $k>0$ 都會令 $\mathrm{e}^{-3x^2}y$ 趨向 0 或趨向 ∞。）

因此取 $k=2$。這時方程是 $\frac{\mathrm{d}y}{\mathrm{d}x} = x\,\mathrm{e}^{6x^2}/(y+3)$，其解為

$$(y+3)^2 = \frac{\mathrm{e}^{6x^2}}{6} + D,$$

其中 $D=(y_0+3)^2-\frac{1}{6}$，而 y_0 是 y 在 $x=0$ 時的值。右邊由 $(y_0+3)^2>0$ 開始遞增，所以一直保持正數，而解對每一個 $x\geq 0$ 都存在：

$$y = \sqrt{\tfrac{1}{6}\mathrm{e}^{6x^2}+D}\; - 3.$$

乘以 e^{-3x^2}，並把 $\mathrm{e}^{-3x^2}=\sqrt{\mathrm{e}^{-6x^2}}$ 放入根號內，得

$$\mathrm{e}^{-3x^2}y = \sqrt{\tfrac{1}{6}+D\,\mathrm{e}^{-6x^2}}\; - \; 3\mathrm{e}^{-3x^2} \longrightarrow \sqrt{\tfrac{1}{6}} = \frac{1}{\sqrt{6}}$$

當 $x\to\infty$ 時成立，因為兩個指數項都趨向 0。所以當 $k=2$ 時，極限（limit）是

$$\boxed{\lim_{x\to\infty}\mathrm{e}^{-3x^2}y = \frac{1}{\sqrt{6}}},$$

不論初始值是甚麼：常數 D 已從答案中消失，這正是題目可以說「任何解」的原因。

當 $k=2$ 時，冪 $(y+3)^{-1}$ 對 $y<-3$ 也有意義。在該分支上 $y+3=-\sqrt{\frac{1}{6}\mathrm{e}^{6x^2}+D}$，而同樣的論證給出極限 $-1/\sqrt{6}$。題目預期的答案是 $1/\sqrt{6}$，對應 $y>-3$ 的解。

驗算

- 在 (i) 中，$x=0$ 給出 $y=(2+1)^2-3=6$，而對 $y=(2+\mathrm{e}^{-x^2})^2-3$ 求導得 $-4x\mathrm{e}^{-x^2}(2+\mathrm{e}^{-x^2}) = -4x\mathrm{e}^{-x^2}(y+3)^{1/2}$。

- 以數值方法（龍格－庫塔法，Runge–Kutta）逐步求解兩條方程，結果與公式吻合；例如在 (i) 中 $y(2) \approx 1.07360$，而在 (ii) 中，由 $y_0 = 6$ 開始，$\mathrm{e}^{-12} y(2) \approx 0.408230$，已經很接近 $1/\sqrt{6} \approx 0.408248$。

常見錯誤

- 到平方之後才記起要加積分常數（constant of integration），或者把 $y = 6$ 代入了錯誤的式子。
- 把 $x\,\mathrm{e}^{6x^2}$ 積分成 $\mathrm{e}^{6x^2}/6$ 而不是 $\mathrm{e}^{6x^2}/12$；這不會改變 k，但會改變極限。

STEP 之外

第 (ii) 部分是漸近分析（asymptotic analysis）的一個小例子：我們無須理會初始值，就找出了解在 x 很大時的表現，因為積分常數被不斷增長的項所淹沒。探究解的哪些特徵與初始條件（initial conditions）無關，正是微分方程（differential equations）定性理論的核心問題。Boyce 與 DiPrima 的 *Elementary Differential Equations*（微分方程入門教科書）涵蓋可分離變數方程以及這類長期表現的討論。

詳細解答

STEP II 2000，第 9 題

力學 • 有空氣阻力的運動

STEP II 2000，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：

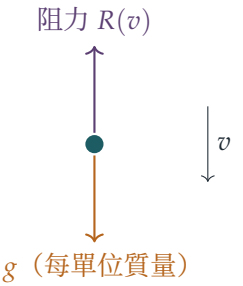
step.maths.org/questions/00-s2-q9

關鍵思路

每位跳傘者都遵從牛頓第二定律（Newton's second law），而阻力只取決於速率，因此我們得到一條關於 v 和 t 的一階微分方程（first-order differential equation），而且可以分離變數（separation of variables）。「速率永遠達不到」這類結果源自終端速率（terminal speed）：即阻力剛好與重力平衡時的速率。解會慢慢向它靠近，但在有限時間內永遠到不了。至於 Karen，只要用無因次（dimensionless）速率 $u = kv/g$ 來表示，那個二次式便可以乾淨俐落地因式分解。

解答

我們把每位跳傘者看成一個由靜止開始垂直下落的質點（particle），重力加速度 g 為常數，方向向下。取 $v \geq 0$ 為向下的速率，並由起跳一刻開始量度 t。以每單位質量計，所受的力是向下的重力 g 和向上的空氣阻力（air resistance），所以由牛頓第二定律得 $\dfrac{\mathrm{d}v}{\mathrm{d}t} = g -$ (每單位質量的阻力)。



Jane 的速率上界

Jane 的運動方程（equation of motion）是

$$\frac{\mathrm{d}v}{\mathrm{d}t} = g - kv = k\Big(\frac{g}{k} - v\Big).$$

起初 $v = 0 < g/k$。只要 $v < g/k$，我們便可以除以 $g/k - v > 0$，然後分離變數：

$$\int_0^v \frac{\mathrm{d}w}{g/k - w} = \int_0^t k\,\mathrm{d}s \quad\Longrightarrow\quad -\ln\Big(1 - \frac{kv}{g}\Big) = kt,$$

所以

$$v = \frac{g}{k}\big(1 - \mathrm{e}^{-kt}\big).$$

這個解對所有 $t \geq 0$ 都成立，因為對任何有限的 t，右邊都小於 g/k，而這正是我們做除法時所用的條件。由於 $\mathrm{e}^{-kt} > 0$，對所有 t 都有 $v < g/k$。所以當 $t \to \infty$ 時，Jane 的速率趨向 g/k，但永遠達不到這個值。

Karen 的速率上界

Karen 的運動方程是

$$\frac{\mathrm{d}v}{\mathrm{d}t} = g - kv - \frac{2k^2}{g}v^2.$$

設 $u = kv/g$，這是一個無因次的速率。於是右邊變成

$$g\big(1 - u - 2u^2\big) = g(1 - 2u)(1 + u), \qquad \text{及} \qquad \frac{\mathrm{d}u}{\mathrm{d}t} = \frac{k}{g}\frac{\mathrm{d}v}{\mathrm{d}t} = k(1 - 2u)(1 + u).$$

當 $u = \frac{1}{2}$ 時加速度為零（另一個根 $u = -1$ 並不是物理上合理的速率），即 $v = g/(2k)$。這就是 Karen 的終端速率。

起初 $u = 0 < \frac{1}{2}$。只要 $0 \leq u < \frac{1}{2}$，兩個因式都是正數，所以可以分離變數。化成部分分式（partial fractions），

$$\frac{1}{(1 - 2u)(1 + u)} = \frac{2/3}{1 - 2u} + \frac{1/3}{1 + u},$$

這可以在 $1 = A(1 + u) + B(1 - 2u)$ 中代入 $u = -1$ 和 $u = \frac{1}{2}$ 來驗證。因此

$$kt = \int_0^u \Big(\frac{2/3}{1 - 2w} + \frac{1/3}{1 + w}\Big)\mathrm{d}w = \frac{1}{3}\Big[-\ln(1 - 2w) + \ln(1 + w)\Big]_0^u = \frac{1}{3}\ln\frac{1 + u}{1 - 2u}.$$

把 $\dfrac{1 + u}{1 - 2u} = \mathrm{e}^{3kt}$ 重新整理，得

$$u = \frac{\mathrm{e}^{3kt} - 1}{2\mathrm{e}^{3kt} + 1} = \frac{1}{2} - \frac{3/2}{2\mathrm{e}^{3kt} + 1} < \frac{1}{2}.$$

基於與之前相同的理由，這個解對所有 $t \geq 0$ 都成立，所以 Karen 的速率滿足

$$v < \frac{g}{2k} \quad \text{對所有} t.$$

這就是對應的結果：她的速率趨近 $g/(2k)$，但永遠達不到這個值。

Jane 達到 $g/(3k)$ 所需的時間

在 Jane 的解中令 $v = g/(3k)$，得 $1 - \mathrm{e}^{-kt} = \frac{1}{3}$，所以 $\mathrm{e}^{-kt} = \frac{2}{3}$，而

$$t = \frac{1}{k} \ln \frac{3}{2},$$

正如題目所求。

Karen 達到 $g/(3k)$ 所需的時間

想一想 我們需要以 t 表示 u 的公式嗎？還是有更快的途徑？

更快的途徑是利用以 u 表示 t 的公式，而這個公式我們已經有了。速率為 $g/(3k)$ 即 $u = \frac{1}{3}$，它小於 $\frac{1}{2}$，所以她確實會達到這個速率。於是

$$\frac{1+u}{1-2u} = \frac{4/3}{1/3} = 4, \qquad \text{所以} \qquad t = \frac{1}{3k} \ln 4 = \frac{2}{3k} \ln 2.$$

驗算

- 在任何正的速率下，Karen 所受的阻力都比 Jane 大，所以她應該要用更長時間才達到相同的速率。果然，$\frac{2}{3} \ln 2 \approx 0.462$ 大於 $\ln \frac{3}{2} \approx 0.405$。
- 當 t 很小時，兩條公式都給出 $v \approx gt$，即沒有阻力的自由下落（free fall），這正是應有的結果。把 $u = (\mathrm{e}^{3kt} - 1)/(2\mathrm{e}^{3kt} + 1) \approx 3kt/3 = kt$ 展開，便得 $v \approx gt$。
- 我們由靜止開始，用數值方法逐步求解兩條方程（龍格－庫塔法（Runge–Kutta），取 $g = 9.8$，$k = 0.7$）。速率一直保持在 14 和 7 以下，這兩個數正是 g/k 和 $g/(2k)$；而速率 $g/(3k)$ 分別在 $t \approx 0.5792$ 和 $t \approx 0.6601$ 時達到。這些數值與 $k^{-1} \ln \frac{3}{2}$ 和 $\frac{2}{3k} \ln 2$ 相符，誤差約為 10^{-9}。

常見錯誤

- 單憑加速度為正便斷言「速率永遠達不到 g/k」。這只說明速率在增加，卻不能說明它一直低於 g/k。你需要明確的解，或者論證達到終端速率所需的時間是無限的。
- 弄錯 $\ln(1-2u)$ 一項的正負號。對 $\frac{1}{1-2u}$ 積分，得 $-\frac{1}{2} \ln(1-2u)$。

STEP 之外

兩條方程都是自治（autonomous）方程：$\mathrm{d}v/\mathrm{d}t = f(v)$。終端速率是這類方程的穩定平衡點（stable equilibrium）。常微分方程（ODE）的唯一性定理（uniqueness theorem）表

明，當 f 光滑時，解永遠不能在有限時間內到達平衡點，這就是兩個上界背後的一般原因。Karen 的方程是一條里卡蒂方程（Riccati equation），而同樣的邏輯斯蒂型（logistic-type）行為亦見於人口模型。Boyce 與 DiPrima 的 *Elementary Differential Equations* 涵蓋自治方程及其平衡點，而 Kleppner 與 Kolenkow 的 *An Introduction to Mechanics* 則討論有阻力的運動。

詳細解答

STEP II 2000，第 10 題

力學 • 連接滑輪系統

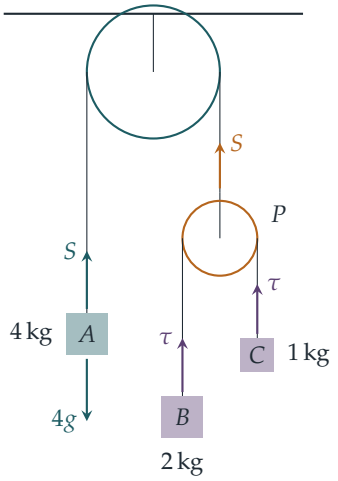
> **STEP II 2000，第 10 題**
>
> 繼續閱讀之前，請先看題目（原題為英文）。題目載於 STEP Support Programme 的試題資料庫：
>
> step.maths.org/questions/00-s2-q10

關鍵思路

這是一部雙重阿特伍德機（double Atwood machine）。當 A 被固定時，第二個滑輪（pulley）不動，B、C 組成一部普通的阿特伍德機。一旦放開 A，第二個滑輪也會移動，所以最乾淨利落的入手方法，是量度 B 和 C 相對於該滑輪的運動，並利用「輕滑輪所受合力（net force）為零」這個事實：滑輪上方的繩的拉力，恰好是下方那條繩的張力（tension）的兩倍。

解答

我們把各質點（particle）視為點質量，把繩視為輕（light）而不可伸長（inextensible），並把兩個滑輪都視為光滑（smooth）而輕。稱第二個滑輪為 P。所有加速度（acceleration）一律取向下為正，以 S 表示穿過 A 的那條繩的張力，以 τ 表示 BC 的張力，而 g 的單位為 $\mathrm{m\,s^{-2}}$。



由於 P 是輕的，它所受的合力必定為零（否則它的加速度會是無限大）。它被 S 向上拉，兩邊又各被 τ 向下拉，所以在任何時刻都有

$$S = 2\tau$$

。

第一階段：A 被固定

A 被固定不動時，繞過固定滑輪的繩不會移動，所以 P 固定不動。這時 B 和 C 就是一部普通的阿特伍德機：若 B 以 a 向下加速，則 C 以 a 向上加速。由牛頓第二定律(Newton's second law）得

$$2a = 2g - \tau, \qquad 1 \cdot a = \tau - g.$$

兩式相加，得 $3a = g$，所以

$$a = \tfrac{1}{3}g, \qquad \tau = \tfrac{4}{3}g.$$

因此 B 向下加速，C 向上加速，大小各為 $\frac{1}{3}g\ \mathrm{m\,s^{-2}}$。

在這個階段結束時（$t = T$），B 正向下運動，C 正向上運動，速率各為 $\frac{1}{3}gT$，而 B 已下降了 $\frac{1}{2} \cdot \frac{1}{3}g\,T^2 = \frac{1}{6}gT^2$。在整個階段中，$A$ 和 P 都靜止不動。

第二階段：全部自由運動

設 A 的加速度為 a_A（向下）。繞過固定滑輪的繩長度不變，所以 P 的加速度為 $-a_A$：A 下降得多快，它就上升得多快。

> **想一想** 當 B 和 C 所懸掛的滑輪本身也在加速時，我們該怎樣描述它們的加速度？

相對於 P 而言，繩 BC 的表現與固定的阿特伍德機完全一樣：若 B 相對於 P 以加速度 r 向下運動，則 C 相對於 P 以加速度 r 向上運動。再加上 P 本身的加速度，得

$$a_B = -a_A + r, \qquad a_C = -a_A - r.$$

現在對每個質點應用牛頓第二定律（向下為正），並利用 $S = 2\tau$：

$$\begin{aligned} A:&\quad 4a_A = 4g - 2\tau,\\ B:&\quad 2(-a_A + r) = 2g - \tau,\\ C:&\quad 1(-a_A - r) = g - \tau. \end{aligned}$$

由 A 的方程得 $\tau = 2g - 2a_A$。把它代入 B 的方程，得 $-2a_A + 2r = 2a_A$，所以 $r = 2a_A$。把兩者都代入 C 的方程：

$$-3a_A = g - (2g - 2a_A) = -g + 2a_A \quad\Longrightarrow\quad a_A = \tfrac{1}{5}g.$$

於是 $r = \frac{2}{5}g$，$\tau = \frac{8}{5}g$，$S = \frac{16}{5}g$，並且

$$a_A = \tfrac{1}{5}g\,(\text{向下}), \qquad a_B = -\tfrac{1}{5}g + \tfrac{2}{5}g = \tfrac{1}{5}g\,(\text{向下}), \qquad a_C = -\tfrac{1}{5}g - \tfrac{2}{5}g = -\tfrac{3}{5}g,$$

所以 C 以 $\frac{3}{5}g\ \mathrm{m\,s^{-2}}$ 向上加速。

這些加速度都是常數，而且在 A 被釋放的一刻，沒有任何量突然改變：A 和 P 由靜止開始運動，而 B、C 則保持原有的速率 $\frac{1}{3}gT$。

A 的速率

在第一階段中 B 移動了 $\frac{1}{6}gT^2$，所以它還要再走

$$0.6\,gT^2 - \tfrac{1}{6}gT^2 = \tfrac{3}{5}gT^2 - \tfrac{1}{6}gT^2 = \tfrac{13}{30}gT^2$$

。在第二階段中，B 開始時的向下速度為 $\frac{1}{3}gT$，向下加速度為 $\frac{1}{5}g$，所以它的速度一直為正，所走的距離等於它的位移。若這段路程需時 t，則由 $s = ut + \frac{1}{2}at^2$ 得

$$\tfrac{1}{3}gT\,t + \tfrac{1}{10}gt^2 = \tfrac{13}{30}gT^2.$$

設 $t = kT$，並把全式乘以 $30/(gT^2)$：

$$3k^2 + 10k - 13 = 0 \quad\Longrightarrow\quad (k-1)(3k+13) = 0.$$

由於 $t > 0$，故 $k = 1$：第二階段恰好歷時 T 秒。A 在這個階段由靜止開始，加速度為 $\frac{1}{5}g$，所以它這時的速率為

$$v_A = \tfrac{1}{5}g \cdot T = \tfrac{gT}{5}\ \mathrm{m\,s^{-1}}.$$

驗算

- 把求得的數值代入，每條運動方程都成立：對 A 而言，$4 \cdot \frac{1}{5}g = 4g - \frac{16}{5}g$；對 C 而言，$-\frac{3}{5}g = g - \frac{8}{5}g$。
- 一個獨立的驗算：兩條繩迫使 $y_C = \text{常數} - 2y_A - y_B$（各位置均以向下量度），而以兩個自由坐標 y_A、y_B 運用能量方法，同樣得出 $a_A = a_B = \frac{1}{5}g$，$a_C = -\frac{3}{5}g$。
- 取 $g = 9.8$，$T = 1.7$，以數值方法逐步計算運動，B 在 $t = 3.400$（即 $2T$）時走了 $0.6gT^2$，此時 A 的速率為 $3.332 = gT/5$。
- 總重量為 $7g$，但固定滑輪只承受 $2S = \frac{32}{5}g$，小於 $7g$，所以在第二階段中，系統的質心（centre of mass）向下加速，理應如此。

常見錯誤

- 在第二階段仍把 P 當作固定，並對 B 和 C 沿用 $\pm\frac{1}{3}g$。
- 忘記了輕滑輪有 $S = 2\tau$，或者為滑輪加上質量項。
- 把整段 $0.6gT^2$ 當作第二階段的距離，而沒有減去已經走過的 $\frac{1}{6}gT^2$。

STEP 之外

像第二項驗算那樣，選取能自動滿足繩的約束條件的位置坐標，正是拉格朗日力學（Lagrangian mechanics）的起點：把動能（kinetic energy）和勢能（potential energy）用自由坐標表示出來，運動方程便隨之而得，完全不必理會張力。Kleppner 與 Kolenkow 所著的 *An Introduction to Mechanics*（力學入門教材）對滑輪這類受約束系統有細緻的處理，是很好的下一步讀物。

詳細解答

STEP II 2000，第 11 題

力學 • 彈性繩與簡諧運動

STEP II 2000，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/00-s2-q11

關鍵思路

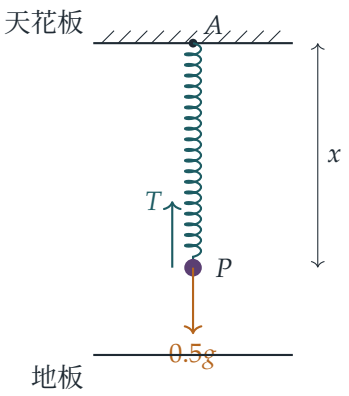
用能量可以很快解決兩個定性的結論：把釋放時儲存的彈性勢能（elastic potential energy），與到達自然長度（natural length）及到達天花板所需的重力勢能作比較。至於時間，則把運動分成兩個階段：彈性繩拉緊時，質點繞其平衡位置（equilibrium position）作簡諧運動（simple harmonic motion）；彈性繩一旦鬆弛，質點便只在重力作用下自由運動。回程就是把上升的過程倒過來走一遍，所以我們把上升所需的時間加倍。

解答

我們把 P 視為繫於一條輕彈性繩上的質點（particle），彈性繩服從胡克定律（Hooke's law），並且沒有空氣阻力。設 x 為 P 在 A 之下的距離，以向下為正，單位為米。當 $x > 1.5$ 時彈性繩拉緊，此時其伸長量（extension）為 $x - 1.5$，張力（tension）為

$$T = \frac{\lambda}{l}(x - 1.5) = \frac{5g}{1.5}(x - 1.5) = \frac{10g}{3}(x - 1.5).$$

房間高 2.5 m，所以 P 由 $x = 2$ 出發，伸長量為 0.5 m。



彈性繩變得鬆弛

釋放時，儲存的彈性勢能為

$$\frac{\lambda e^2}{2l} = \frac{5g\,(0.5)^2}{2(1.5)} = \frac{5g}{12}\,\text{J}.$$

彈性繩要回復到自然長度，P 必須上升 0.5 m，所增加的勢能為 $0.5g \times 0.5 = \frac{g}{4}$ J。由於 $\frac{5g}{12} > \frac{g}{4}$，$P$ 到達 $x = 1.5$ 時仍在向上運動，其動能（kinetic energy）為

$$\tfrac{1}{2}(0.5)v^2 = \frac{5g}{12} - \frac{g}{4} = \frac{g}{6}, \qquad \text{所以} \qquad v^2 = \frac{2g}{3}.$$

因此它會越過彈性繩處於自然長度的那一點，繼續向上，彈性繩於是變得鬆弛。

P 不會到達天花板

要到達天花板，P 須由起點上升 2 m，勢能要增加 $0.5g \times 2 = g$ J。可用的能量就只有釋放時儲存在彈性繩中的 $\frac{5g}{12}$ J（彈性繩一旦鬆弛便不儲存任何能量），而 $\frac{5g}{12} < g$。所以 P 不可能到達天花板。

更準確地說，彈性繩鬆弛之後，P 是一個以速率 v 沿鉛垂方向運動的拋體（projectile），所以它會再上升 $\dfrac{v^2}{2g} = \dfrac{1}{3}$ m。它的最高點位於 $x = 1.5 - \frac{1}{3} = \frac{7}{6}$，仍在天花板之下 $\frac{7}{6}$ m。

彈性繩拉緊時的簡諧運動

當 $x > 1.5$ 時，由牛頓第二定律（Newton's second law）（以向下為正）得

$$0.5\,\ddot{x} = 0.5g - \frac{10g}{3}(x - 1.5).$$

平衡位置就是右邊等於零之處：$x - 1.5 = 0.15$，所以 $x = 1.65$。設 $y = x - 1.65$，即在平衡位置之下的位移（displacement）。於是

$$\ddot{y} = 2g - \frac{20g}{3}(y + 0.15) = -\frac{20g}{3}\,y,$$

這是繞 $y = 0$ 的簡諧運動，其中 $\omega^2 = \dfrac{20g}{3}$。只要彈性繩處於拉緊狀態，這結果便成立，正如題目所要求。

返回所需的時間

P 由靜止出發，起點為 $y = 2 - 1.65 = 0.35$，所以在彈性繩拉緊期間

$$y = 0.35\cos\omega t.$$

彈性繩在 $x = 1.5$ 時變得鬆弛，即 $y = -0.15$。所以首次出現這情況的時刻 t_1 滿足

$$\cos\omega t_1 = -\frac{0.15}{0.35} = -\frac{3}{7}.$$

想一想 $\cos\omega t_1 = -\frac{3}{7}$ 的各個解之中，我們要的是哪一個？又怎樣用 arccos 把它寫出來？

當 t 由 0 開始增加時，ωt 亦由 0 開始增加，而 $\cos\omega t$ 首次達到 $-\frac{3}{7}$ 時，ωt 位於 $(\frac{\pi}{2}, \pi)$ 之內。由於 $\arccos\frac{3}{7} \in (0, \frac{\pi}{2})$ 且 $\cos(\pi - \theta) = -\cos\theta$，這個值就是 $\omega t_1 = \pi - \arccos\frac{3}{7}$。因此

$$t_1 = \frac{1}{\omega}\left(\pi - \arccos\tfrac{3}{7}\right) = \left(\frac{3}{20g}\right)^{1/2}\left(\pi - \arccos\tfrac{3}{7}\right).$$

（作為驗算，此時的速率為 $\omega\sqrt{0.35^2 - 0.15^2} = \sqrt{\frac{20g}{3} \times 0.1} = \sqrt{\frac{2g}{3}}$，與能量計算的結果吻合。）

彈性繩鬆弛期間，P 只在重力作用下運動。它以速率 $v = \sqrt{2g/3}$ 向上離開 $x = 1.5$，然後以相同的速率向下回到 $x = 1.5$，所需時間為

$$t_2 = \frac{2v}{g} = \frac{2}{g}\sqrt{\frac{2g}{3}} = \left(\frac{8}{3g}\right)^{1/2}.$$

若作用力只取決於位置，運動便是可逆的。P 在 $x = 1.5$ 重新進入拉緊區域時，速率與離開時相同，於是它把簡諧運動階段倒過來重走一遍，再用 t_1 的時間在 $y = 0.35$（即原來的位置）靜止下來。總時間為

$$\begin{aligned} 2t_1 + t_2 &= 2\left(\frac{3}{20g}\right)^{1/2}\left(\pi - \arccos\tfrac{3}{7}\right) + \left(\frac{8}{3g}\right)^{1/2} \\ &= \left(\frac{8}{3g}\right)^{1/2} + \left(\frac{3}{5g}\right)^{1/2}\left(\pi - \arccos\tfrac{3}{7}\right), \end{aligned}$$

其中用到 $2\sqrt{3/(20g)} = \sqrt{12/(20g)} = \sqrt{3/(5g)}$。這就是所要求的結果。

驗算

- 求自然長度處速率的兩種方法（能量法，以及簡諧運動的 $v^2 = \omega^2(a^2 - y^2)$）同樣得出 $v^2 = \frac{2g}{3}$。
- 取 $g = 9.8$，公式給出約 1.0199 s。以數值方法逐步求解運動方程（當 $x < 1.5$ 時把張力關掉），得出返回起點並靜止所需的時間相同，而最高點亦在天花板之下 $\frac{7}{6}$ m。
- 若沒有鬆弛階段，簡諧運動的週期（period）會是 $2\pi/\omega \approx 0.77$ s。我們的 $2t_1 \approx 0.50$ s 比這個值小，這是應有的結果，因為只完成了振動的一部分。

常見錯誤

- 把簡諧運動的中心寫成自然長度處，而不是平衡位置。重力把中心移到伸長量為 0.15 m 之處。
- 取 $\omega t_1 = \arccos(-\frac{3}{7})$，卻沒有認出這就是 $\pi - \arccos\frac{3}{7}$。
- 忘記回程：所求的時間包括上升以及返回，所以兩個階段都要加倍。

STEP 之外

這是一個分段定義的，或稱「非光滑」（non-smooth）的動力系統（dynamical system）：不同的區域適用不同的運動方程，而各段的解則在邊界處透過位置和速度的配對拼接起來。同樣的想法也支配着彈跳的球、碰撞和開關電路，而我們用過的可逆性論證，正是初次窺見保守系統（conservative system）中的時間反演對稱性（time-reversal symmetry）。想了解更多，可參閱 Kleppner 與 Kolenkow 的 *An Introduction to Mechanics*（力學入門教科書）。

詳細解答

STEP II 2000，第 12 題

概率與統計 • 正態變量之和

> **STEP II 2000，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/00-s2-q12

關鍵思路

Henry 遲到，正正就是他的總行程時間超過他在 0900 之前所餘時間的時候，所以對每條路線，我們把兩段獨立的正態行程時間相加，得出單一個正態變量(normal variable)。標準化（standardising）之後，「遲到」就變成「一個標準正態變量超過某個數」，而由於 Φ 是遞增的，兩條路線毋須查表便可比較。餘下的工作就是全概率定律（law of total probability）和貝葉斯定理（Bayes' theorem）。

解答

比較兩條路線

我們假設在每條路線上，火車行程與其後的行程互相獨立（independent）。

King's Cross。設 $X_1 \sim N(55, 25)$ 為火車行程時間，$Y_1 \sim N(30, 144)$ 為其後的行程時間。由於它們是獨立的正態變量，它們的和也是正態的，而平均值（mean）與方差（variance）各自相加：

$$T_1 = X_1 + Y_1 \sim N(85,\, 169), \qquad \text{標準差} 13.$$

火車在 0715 開出，所以 Henry 有 105 分鐘，而當 $T_1 > 105$ 時他便會遲到。設 $Z \sim N(0,1)$，則

$$\mathrm{P}(\text{遲到} \mid \text{King's Cross}) = \mathrm{P}\left(Z > \frac{105 - 85}{13}\right) = 1 - \Phi\left(\tfrac{20}{13}\right) = A.$$

Liverpool Street。同樣地，$T_2 = X_2 + Y_2$，其中 $X_2 \sim N(65, 16)$ 與 $Y_2 \sim N(25, 9)$ 互相

獨立，所以

$$T_2 \sim N(90, 25), \qquad \text{標準差5.}$$

這班火車在 0720 開出，餘下 100 分鐘，所以

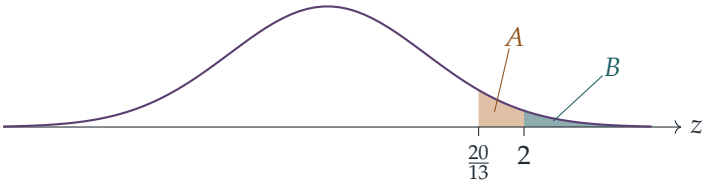
$$\mathrm{P}(\text{遲到} \mid \text{Liverpool Street}) = \mathrm{P}\left(Z > \frac{100-90}{5}\right) = 1 - \Phi(2) = B.$$

想一想 不查表，怎樣比較 A 與 B？

Φ 是一個累積分佈函數（cumulative distribution function），其密度為正，所以它是嚴格遞增的。由於 $\frac{20}{13} < \frac{26}{13} = 2$，我們有 $\Phi(\frac{20}{13}) < \Phi(2)$，因此

$$A = 1 - \Phi\left(\tfrac{20}{13}\right) > 1 - \Phi(2) = B.$$

所以 Henry 若乘搭往 King's Cross 的火車，遲到的機會較大，這正是所要證明的。（有趣的是，這條路線平均而言反而較快；它之所以落敗，是因為由 King's Cross 出發的那段地鐵行程變化太大。）



標準正態曲線下塗上陰影的尾部：A 是整個陰影部分的面積，B 是 2 以外顏色較深的部分。

遲到次數的期望值

在單一次行程中，設 K 為 Henry 乘搭往 King's Cross 的火車這一事件，故 $\mathrm{P}(K) = p$，並設 D 為他遲到這一事件。根據全概率定律，

$$\mathrm{P}(D) = \mathrm{P}(D \mid K)\,\mathrm{P}(K) + \mathrm{P}(D \mid K')\,\mathrm{P}(K') = pA + (1-p)B.$$

假設 M 次行程互相獨立，遲到次數便服從參數為 M 和 $\mathrm{P}(D)$ 的二項分佈（binomial distribution），所以

$$L = M(pA + (1-p)B) = M(B + p(A-B)).$$

由於 $A - B > 0$ 且 $0 \le p \le 1$，表達式 $B + p(A-B)$ 隨 p 遞增，由 $p = 0$ 時的 B 增至 $p = 1$ 時的 A。因此對 p 的每一個可能值，

$$BM \le L \le AM,$$

左邊的等號僅在 $p = 0$ 時成立，右邊的等號僅在 $p = 1$ 時成立。用文字來說：L/M 是 A 與 B 的加權平均（weighted average），所以它必定介乎兩者之間。

估計 p

在遲到的日子當中，他乘搭了往 King's Cross 火車的比例，是 $\mathrm{P}(K \mid D)$ 的估計值；而因為 M 很大，我們把觀察所得的 $\frac{3}{5}$ 視為等於這個概率。根據貝葉斯定理，

$$\mathrm{P}(K \mid D) = \frac{\mathrm{P}(D \mid K)\,\mathrm{P}(K)}{\mathrm{P}(D)} = \frac{pA}{pA + (1-p)B}.$$

令此式等於 $\frac{3}{5}$：

$$5pA = 3pA + 3(1-p)B$$
$$2pA + 3pB = 3B,$$

所以估計值是

$$\hat{p} = \frac{3B}{2A + 3B}.$$

驗算

- 以數值計算，$A \approx 0.0620$ 而 $B \approx 0.0228$，所以經 King's Cross 約有 6% 的時候遲到，經 Liverpool Street 則約 2%。對每條路線各模擬 4×10^5 次行程，結果吻合至小數點後三位。
- 代入這些數值，$\hat{p} \approx 0.355$。代回原式，$pA/(pA + (1-p)B) = 0.6$ 準確成立；而以這個 p 模擬整個過程，所得的遲到個案中約有 60% 來自 King's Cross。
- 對最後答案作合理性檢查：假如兩條路線同樣可靠（$A = B$），遲到的情況便只反映他選乘哪班火車，所以當中有 $\frac{3}{5}$ 經 King's Cross，就表示 $p = \frac{3}{5}$。公式給出 $3B/5B = \frac{3}{5}$。由於 King's Cross 是較不可靠的路線，它在遲到的日子中所佔的比例偏高，所以 p 必定小於 $\frac{3}{5}$，而事實正是如此。

常見錯誤

- 把標準差（standard deviation）相加，而不是把方差相加：King's Cross 路線的總時間方差為 $25 + 144 = 169$，所以其標準差是 13，而不是 $5 + 12 = 17$。
- 兩班火車都用同一個 105 分鐘的時限：往 Liverpool Street 的火車遲五分鐘開出。
- 直接令 $p = \frac{3}{5}$。這個 $\frac{3}{5}$ 是 $\mathrm{P}(K \mid D)$，而不是 $\mathrm{P}(K)$；兩者不同，是因為兩條路線並非同樣可靠。

STEP 之外

這裏有兩個想法，會在大學第一個概率課程中反覆出現。獨立正態變量之和仍是正態的這個事實，用矩母函數（moment generating function）來證明最為簡潔；它也是正態分佈族的穩定性（stability）的一個特例，而這種穩定性正是中心極限定理（central limit theorem）的基礎。最後一部分則是統計推論（statistical inference）的縮影：我們觀察到一個結果（遲到），然後運用貝葉斯定理推斷隱藏的原因（選乘哪班火車）。Grimmett 與 Welsh 的 *Probability: An Introduction*（概率入門）涵蓋了這兩個課題，而 Blitzstein 與 Hwang 的 *Introduction to Probability*（概率導論）則有更多正是這種風格的貝葉斯題目。

詳細解答

STEP II 2000，第 13 題

概率與統計 • 捕捉再捕捉估計

STEP II 2000，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/00-s2-q13

本題在 *Beyond STEP* 第 13.2 節點算田鼠中有完整的引導式講解：https://beyondstep.keithmcnulty.org/zh-hant/ch13/voles.html

關鍵思路

這是捉放法（capture–recapture）：第二次抽樣是從 N 隻田鼠中隨機選出 200 隻，而當中有 200 隻帶有標記，所以樣本中帶標記田鼠的數目服從超幾何分佈（hypergeometric distribution）。要找出 p_N 在何處最大，不要求導（N 的階乘（factorial）並沒有簡單意義下的導數）；應該改為透過比值 p_N / p_{N-1} 比較相鄰兩項，這時幾乎所有東西都會約去。最後的恆等式則不費吹灰之力：第三次抽樣的各個概率之和必定是 1。

解答

概率 p_N

把第二次抽樣看成從 N 隻田鼠中選出 200 隻，每個大小為 200 的子集出現的機會均等。這樣的子集共有 $\binom{N}{200}$ 個。在 N 隻田鼠中，200 隻帶有標記，$N-200$ 隻沒有。一個恰好含 11 隻帶標記田鼠的樣本，就是選出 11 隻帶標記的和 189 隻沒有標記的，所以

$$p_N = \frac{\binom{200}{11}\binom{N-200}{189}}{\binom{N}{200}}.$$

把含有 N 的兩個二項式係數（binomial coefficient）寫成階乘，

$$\begin{aligned} p_N &= \binom{200}{11} \cdot \frac{(N-200)!}{189!\,(N-389)!} \cdot \frac{200!\,(N-200)!}{N!} \\ &= k\,\frac{((N-200)!)^2}{N!\,(N-389)!}, \qquad \text{其中} k = \binom{200}{11}\frac{200!}{189!}, \end{aligned}$$

而這個常數與 N 無關。這正是所要求的形式。（這裏需要 $N \geq 389$，因為兩次抽樣合共包含 $200 + 189 = 389$ 隻不同的田鼠。）

求 p_N 的最大值

對於 $N \geq 390$，把 p_N 除以 p_{N-1}。利用 $(N-200)! = (N-200)\,(N-201)!$、$N! = N\,(N-1)!$ 和 $(N-389)! = (N-389)\,(N-390)!$，所有東西都會約去，只餘下

$$\frac{p_N}{p_{N-1}} = \frac{(N-200)^2}{N(N-389)}.$$

分子和分母都是正數，所以

$$\begin{aligned} p_N > p_{N-1} &\iff (N-200)^2 > N(N-389) \\ &\iff N^2 - 400N + 40000 > N^2 - 389N \\ &\iff 40000 > 11N \\ &\iff N < 3636.\overline{36}. \end{aligned}$$

想一想 這告訴我們數列（sequence）$p_{389}, p_{390}, \ldots$ 的形狀是怎樣的？

當 $390 \leq N \leq 3636$ 時，比值大於 1，所以數列嚴格遞增，直至 p_{3636}。當 $N \geq 3637$ 時，比值小於 1（等號成立需要 $11N = 40000$，而這沒有整數解），所以數列自此嚴格遞減。因此 p_N 有唯一的最大值，位於

$$N = 3636.$$

這就是 $\frac{200\times200}{11}$ 的整數部分。這也正是你會猜到的答案：只要認為樣本中帶標記的比例 $\frac{11}{200}$ 應該等於總體中帶標記的比例 $\frac{200}{N}$ 便可。

第三次抽樣

取總體大小為 3636，即剛才求得的估計值。第二次放回之後，帶標記的田鼠是原來的 200 隻加上新標記的 189 隻，所以有 389 隻帶有標記，$3636 - 389 = 3247$ 隻沒有標記。按照與之前完全相同的推理，一個大小為 200 的隨機樣本中恰好有 j 隻帶標記田鼠的概

率是

$$\frac{\binom{389}{j}\binom{3247}{200-j}}{\binom{3636}{200}}, \qquad j = 0, 1, \ldots, 200.$$

由於 $389 \geq 200$ 且 $3247 \geq 200$，由 0 至 200 的每一個 j 都有可能出現，而這些就是全部可能的結果。它們互斥（mutually exclusive），所以它們的概率之和是 1：

$$\sum_{j=0}^{200} \frac{\binom{389}{j}\binom{3247}{200-j}}{\binom{3636}{200}} = 1.$$

乘以 $\binom{3636}{200}$ 便得

$$\sum_{j=0}^{200} \binom{389}{j}\binom{3247}{200-j} = \binom{3636}{200},$$

這正是所要證明的。

驗算

- 用整數二項式係數準確計算 p_N：對於若干個 N 的值（例如 389、1000 和 5000），結果與階乘公式一致；而在 $389 \leq N < 20000$ 的範圍內搜尋，最大值位於 $N = 3636$。
- 在轉折點處，$p_{3636}/p_{3635} = 3436^2/(3636 \times 3246) \approx 1.0003 > 1$，但 $p_{3637}/p_{3636} < 1$。
- 最後的恆等式兩邊以整數準確計算，結果相等。留意 $389 + 3247 = 3636$，這是必然的。

常見錯誤

- 把 N 當作連續變量來求導。應該改用相鄰兩項的比值，並準確地判定不等式。
- 忘記了第二次放回之後，帶標記的田鼠是 389 隻，而不是 200 隻。
- 說「最大值位於 $N = 3636.36$」：總體大小必須是整數，而且你應該檢查相鄰的兩個整數中哪一個勝出。

STEP 之外

選取使觀測數據出現的概率最大的參數值，就是*最大似然法*（maximum likelihood），這是大學一年級統計學的核心思想。在這裏，它給出捉放法的 Lincoln–Petersen 估計量

（estimator）。最後的恆等式是范德蒙德恆等式（Vandermonde's identity），而所用的論證方法，即用兩種方法數算同一樣東西，是組合數學（combinatorics）中一種強而有力的技巧。Grimmett and Welsh 的 *Probability: An Introduction*（概率入門教材）涵蓋了超幾何分佈和這類論證方法。

詳細解答

STEP II 2000，第 14 題

概率與統計 • 中位數的分佈

STEP II 2000，第 14 題

繼續閱讀之前，請先看題目（原題為英文）。題目載於 STEP Support Programme 的試題資料庫：

step.maths.org/questions/00-s2-q14

本題在 *Beyond STEP* 第 13.4 節平均數還是中位數？中有完整的引導式講解：https://beyondstep.keithmcnulty.org/zh-hant/ch13/mean-or-median.html

關鍵思路

中位數（median）的概率密度函數（probability density function）是形如 $y^r(1-y)^s$ 的多項式，所以題目給出的積分已包辦所有工作：先用一次求歸一化常數（normalising constant），再把 y 的冪次分別提高一和提高二，求出首兩個矩（moments）。至於最後一部分，Y 和 $\bar{X}$ 都以 $\frac{1}{2}$ 為中心，而對正態分佈（normal distribution）來說，落在距離中心某固定距離之內的概率只取決於標準差（standard deviation）。離散程度較小的那個變量勝出。

解答

常數 k

密度函數的積分必須等於 1。在題目給出的結果中取 $r = s = n$，

$$1 = \int_0^1 k y^n (1-y)^n \,\mathrm{d}y = k\,\frac{n!\,n!}{(2n+1)!}, \qquad \text{所以} \qquad k = \frac{(2n+1)!}{(n!)^2},$$

正如所求。

平均值

密度函數關於 $y = \frac{1}{2}$ 對稱（把 y 換成 $1-y$，函數不變），所以 $\mathrm{E}(Y) = \frac{1}{2}$。要用積分來確認這一點，取 $r = n+1$、$s = n$：

$$\mathrm{E}(Y) = k\int_0^1 y^{n+1}(1-y)^n \,\mathrm{d}y = \frac{(2n+1)!}{(n!)^2}\cdot\frac{(n+1)!\,n!}{(2n+2)!} = \frac{n+1}{2n+2} = \frac{1}{2}.$$

方差

現在取 $r = n+2$、$s = n$：

$$\mathrm{E}(Y^2) = \frac{(2n+1)!}{(n!)^2}\cdot\frac{(n+2)!\,n!}{(2n+3)!} = \frac{(n+2)(n+1)}{(2n+3)(2n+2)} = \frac{n+2}{2(2n+3)}.$$

因此

$$\mathrm{Var}(Y) = \frac{n+2}{2(2n+3)} - \frac{1}{4} = \frac{2(n+2)-(2n+3)}{4(2n+3)} = \frac{1}{4(2n+3)}.$$

比較中位數與平均值

每個 X_i 都在 $[0,1]$ 上服從均勻分佈（uniform distribution），所以 $\mathrm{E}(X_i) = \frac{1}{2}$，而 $\mathrm{Var}(X_i) = \frac{1}{12}$。由於各 X_i 互相獨立（independent），

$$\mathrm{E}(\bar{X}) = \tfrac{1}{2}, \qquad \mathrm{Var}(\bar{X}) = \frac{1}{(2n+1)^2}\cdot(2n+1)\cdot\frac{1}{12} = \frac{1}{12(2n+1)}.$$

> **想一想** 兩個變量都以 $\frac{1}{2}$ 為中心。哪一個數值決定了誰較有可能落在距離中心 $d/\sqrt{n}$ 之內？

答案是標準差。當 n 很大時，我們可以取 $Y \sim \mathrm{N}(\frac{1}{2}, \sigma_Y^2)$ 和 $\bar{X} \sim \mathrm{N}(\frac{1}{2}, \sigma_X^2)$。若 $W \sim \mathrm{N}(\frac{1}{2}, \sigma^2)$ 且 $c > 0$，則以 Φ 表示標準正態分佈函數（standard normal distribution function），有

$$\mathrm{P}\big(|W - \tfrac{1}{2}| < c\big) = \mathrm{P}\Big(|Z| < \frac{c}{\sigma}\Big) = 2\Phi\Big(\frac{c}{\sigma}\Big) - 1,$$

其中 $Z \sim \mathrm{N}(0,1)$。這是 σ 的嚴格遞減函數，因為 Φ 嚴格遞增。所以取 $c = d/\sqrt{n}$ 時，所求的不等式成立，當且僅當 $\sigma_Y > \sigma_X$，即 $\mathrm{Var}(Y) > \mathrm{Var}(\bar{X})$：

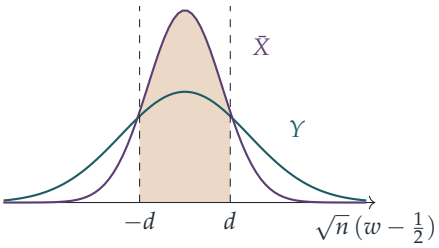
$$\begin{aligned}\frac{1}{4(2n+3)} > \frac{1}{12(2n+1)} &\iff 12(2n+1) > 4(2n+3)\\ &\iff 24n+12 > 8n+12\\ &\iff n > 0.\end{aligned}$$

這對每個 $n \geq 1$ 都成立。所以只要 n 大得足以令正態近似（normal approximation）有效，則不論正數 d 是甚麼，Y 的概率都嚴格小於 $\bar{X}$ 的概率。證明至此完成。

要看看差距有多大，可以把變量標準化：

$$\frac{d/\sqrt{n}}{\sigma_Y} = 2d\sqrt{\frac{2n+3}{n}} \to 2\sqrt{2}\,d, \qquad \frac{d/\sqrt{n}}{\sigma_X} = d\sqrt{\frac{12(2n+1)}{n}} \to 2\sqrt{6}\,d.$$

漸近而言，平均值的標準差是中位數的 $\sqrt{3}$ 分之一。



圖中以 $\sqrt{n}\,(w-\frac{1}{2})$ 為標度，畫出兩個極限密度函數。較窄的曲線（平均值）在任何一條中央帶內都有較大的概率。

驗算

- $n=0$：一個均勻變量的「中位數」就是 X_1 本身，而我們的公式給出 $k=1$ 和 $\mathrm{Var}(Y)=\frac{1}{12}$，正是均勻分佈的方差，與預期相符。
- 當 $n=0$ 時兩個方差都等於 $\frac{1}{12}$，而當 $n\geq 1$ 時兩者便分開，這與上面的不等式 $n>0$ 吻合。
- 以 $n=50$（即 101 個均勻變量）進行模擬，得到 $\mathrm{Var}(Y)\approx 0.00243$，而 $1/412\approx 0.00243$。取 $d=0.3$ 時，模擬得到 $\mathrm{P}(|Y-\frac{1}{2}|<d/\sqrt{n})\approx 0.607$ 和 $\mathrm{P}(|\bar{X}-\frac{1}{2}|<d/\sqrt{n})\approx 0.859$，與正態分佈給出的數值 0.611 和 0.860 相當接近。

常見錯誤

- 寫成 $\mathrm{Var}(\bar{X})=\frac{1}{12}$，或者只除以 $2n+1$ 一次，而不是先在分母得到 $(2n+1)^2$，再乘以 $2n+1$。
- 弄錯方向：方差較大，落在中心附近的概率反而較小。
- 忘記了計算 $\mathrm{E}(Y^2)$ 時用的是 $r=n+2$，而不是 $r=2n$。

STEP 之外

中位數的密度函數是一個貝塔分佈（Beta distribution）$\mathrm{Beta}(n+1,n+1)$，而題目給出的積分就是貝塔函數（Beta function）。你剛才所作的比較，正是*漸近相對效率*（asymptotic

relative efficiency）的概念：對均勻分佈的數據，用樣本平均值（sample mean）估計中心，其方差只有中位數的三分之一；但對厚尾（heavy-tailed）的數據，中位數卻可以勝出。順序統計量（order statistics）和貝塔函數在 Blitzstein 與 Hwang 合著的 *Introduction to Probability*（概率論入門）中有很好的講解。

詳細解答

STEP III 2000，第 1 題

純數學 • 雙曲線的切線

> **STEP III 2000，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/00-s3-q1

關鍵思路

兩條曲線都是以原點為中心的等軸雙曲線（rectangular hyperbola），而且在半周旋轉（half-turn）$(x, y) \mapsto (-x, -y)$ 之下都保持不變。這個對稱性（symmetry）讓我們不費吹灰之力就得到 Q。至於切線（tangent），巧妙的做法不是硬解聯立方程，而是利用 $ab = 1$ 和 $a^2 - b^2 = 2$ 這兩個事實，驗證題目給出的點同時位於兩條直線上。求得四個點之後，繞原點旋轉四分之一周就能解釋為何得出正方形。

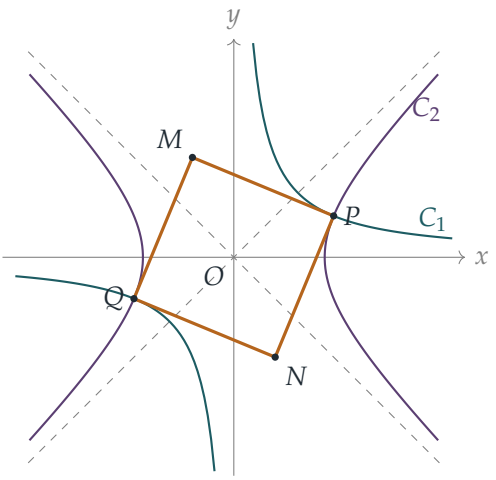
解答

略圖

C_1，即 $xy = 1$，是大家熟悉的等軸雙曲線，位於第一和第三象限，以兩條坐標軸為漸近線（asymptote）。

C_2，即 $x^2 - y^2 = 2$，與 x 軸相交於 $(\pm\sqrt{2}, 0)$，且永不與 y 軸相交（否則需要 $-y^2 = 2$）。寫成 $y^2 = x^2 - 2$，可見 $|y| < |x|$，而且當 $|x|$ 很大時 $y \approx \pm x$，所以它的漸近線是 $y = \pm x$。它就是把 C_1 繞原點旋轉 45°（並按 $\sqrt{2}$ 的比例放大）所得的曲線，兩支分別向左、右開口。

兩條曲線在哪裏相交？由於 $xy = 1 > 0$，任何公共點都位於第一或第三象限。把 $y = 1/x$ 代入 C_2，得 $x^4 - 2x^2 - 1 = 0$，所以 $x^2 = 1 + \sqrt{2}$（另一個根 $1 - \sqrt{2}$ 是負數）。因此恰好有兩個交點，在這兩個象限中各有一個。



圖中藍綠色的是 C_1，紫紅色的是 C_2，虛線是 C_2 的漸近線 $y = \pm x$。琥珀色的正方形 $PMQN$ 就是我們即將求出的圖形。

點 Q

把 (x, y) 換成 $(-x, -y)$，xy 和 $x^2 - y^2$ 都保持不變。所以若 (a, b) 同時位於兩條曲線上，則 $(-a, -b)$ 亦然。由於只有兩個交點，

$$Q = (-a, -b).$$

以下我們會一直用到 P 同時位於兩條曲線上所帶來的兩個事實：

$$ab = 1, \qquad a^2 - b^2 = 2.$$

切線

對於 C_1，把 $xy = 1$ 作隱函數求導（implicit differentiation），得 $y + x\frac{\mathrm{d}y}{\mathrm{d}x} = 0$，所以 C_1 在其上一點 (p, q) 的斜率（gradient）是 $-q/p$。該點的切線是 $p(y - q) = -q(x - p)$，即

$$qx + py = 2pq = 2.$$

對於 C_2，把 $x^2 - y^2 = 2$ 求導，得 $2x - 2y\frac{\mathrm{d}y}{\mathrm{d}x} = 0$，所以在 (p, q) 的斜率是 p/q。切線是 $q(y - q) = p(x - p)$，即

$$px - qy = p^2 - q^2 = 2.$$

（即使在切線垂直的點，這兩個形式仍然成立；它們其實就是曲線的梯度向量（gradient vector）與 (x,y) 的點積。）代入 $P=(a,b)$ 和 $Q=(-a,-b)$：

$$\begin{aligned}
\text{曲線}C_1\text{ 的切線，切點為}P:&\quad bx+ay=2,\\
\text{曲線}C_2\text{ 的切線，切點為}Q:&\quad -ax+by=2,\\
\text{曲線}C_2\text{ 的切線，切點為}P:&\quad ax-by=2,\\
\text{曲線}C_1\text{ 的切線，切點為}Q:&\quad -bx-ay=2.
\end{aligned}$$

點 M

M 是頭兩條直線的交點。試代入 $(x,y)=(-b,a)$：

$$\begin{aligned}
bx+ay&=-b^2+a^2=2,\\
-ax+by&=ab+ab=2ab=2.
\end{aligned}$$

所以 $(-b,a)$ 同時位於兩條直線上。這兩條直線並不平行，因為係數的行列式（determinant）是 $b\cdot b-a\cdot(-a)=a^2+b^2\neq 0$，所以它們恰好相交於一點。因此

$$M=(-b,a),$$

正如題目所求。

想一想 驗證一個已給出的答案，不是比直接求解來得弱嗎？

在這裏並非如此。兩條不平行的直線恰好相交於一點，所以任何同時位於兩條直線上的點就是交點。直接求解當然也可以，但到頭來我們還是要用 $ab=1$ 和 $a^2-b^2=2$ 來整理答案。

點 N

同樣地，N 是 $ax-by=2$ 與 $-bx-ay=2$ 的交點。代入點 $(b,-a)$ 得

$$ab+ab=2,\qquad -b^2+a^2=2,$$

而行列式是 $a(-a)-(-b)(-b)=-(a^2+b^2)\neq 0$。所以

$$N=(b,-a).$$

$PMQN$ 是正方形

看看這四個位置向量（position vector）：

$$P=(a,b),\quad M=(-b,a),\quad Q=(-a,-b),\quad N=(b,-a).$$

繞 O 逆時針旋轉 $90°$ 會把 (x, y) 送到 $(-y, x)$。所以它把 $P \to M \to Q \to N \to P$。這已經暗示了答案；以下是一個直接的論證。

- 對角線 PQ 和 MN 的中點都是 O，所以它們互相平分，$PMQN$ 是平行四邊形（parallelogram）。
- 兩條對角線的長度都是 $2\sqrt{a^2 + b^2}$，所以它們相等，這個平行四邊形是長方形。
- $\overrightarrow{OP} \cdot \overrightarrow{OM} = a(-b) + b(a) = 0$，所以對角線互相垂直，這個長方形是菱形（rhombus）。

既是長方形又是菱形的四邊形就是正方形。頂點的次序也正確：P、M、Q、N 繞 O 依次相隔四分之一周，所以 $PMQN$ 是沿着正方形繞一圈，而不是自我交叉。因此 $PMQN$ 是正方形。 ■

它的邊長的平方是 $|PM|^2 = (a + b)^2 + (b - a)^2 = 2(a^2 + b^2)$。由於 $a^2 = 1 + \sqrt{2}$ 及 $b^2 = 1/a^2 = \sqrt{2} - 1$，這個值是 $4\sqrt{2}$。

驗算

- 以數值計算，$a = \sqrt{1 + \sqrt{2}} \approx 1.5538$ 及 $b \approx 0.6436$。用數值方法求切線的交點，得 $M \approx (-0.6436, 1.5538)$ 和 $N \approx (0.6436, -1.5538)$，與所述相符。四條邊的長度都是 $2.3784 \approx \sqrt{4\sqrt{2}}$，而且相鄰的邊互相垂直。
- P 的方向角是 $\arctan(b/a) = \arctan(\sqrt{2} - 1) = 22.5°$，剛好在 C_1 的漸近線 $y = 0$ 與 C_2 的漸近線 $y = x$ 的正中間，所以正方形傾斜了 $22.5°$，正如略圖所示。

常見錯誤

- 把 C_2 的切線寫成 $px + qy = 2$。負號十分重要，正是它令整個運算行得通。
- 憑空猜測有反射對稱，於是取 $Q = (b, a)$ 或 $(-b, -a)$。對 $y = x$ 的反射會保持 C_1 不變，但不會保持 C_2 不變；只有繞 O 的半周旋轉才同時保持兩者不變。
- 只證明四邊相等。這樣只得到菱形；你還需要一個直角（或對角線相等）。

STEP 之外

我們用過的切線公式 $qx + py = 2$ 和 $px - qy = 2$，都是以下一般法則的例子：圓錐曲線（conic）$\mathbf{x}^\mathsf{T} A \mathbf{x} = c$ 在 $\mathbf{p}$ 的切線是 $\mathbf{p}^\mathsf{T} A \mathbf{x} = c$。在大學一年級關於二次型（quadratic form）的課程中，你會學到如何用旋轉把 A 對角化（diagonalise），這正是 $xy = 1$ 和

$x^2 - y^2 = 2$ 其實是同一條曲線轉了 45° 的原因。Strang 的 *Introduction to Linear Algebra* 清楚地講解了二次型和旋轉，而 Siklos 的 *Advanced Problems in Mathematics* 則有更多 STEP 風格的坐標幾何題目。

詳細解答

STEP III 2000，第 2 題

純數學 • 代換積分法

STEP III 2000，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/00-s3-q2

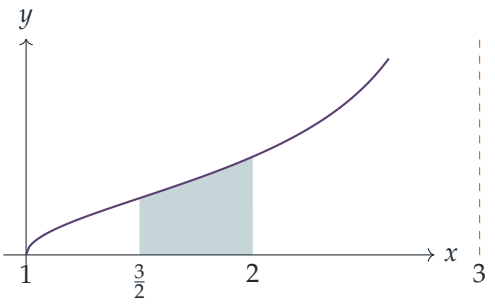
關鍵思路

像 $(x-1)/(3-x)$ 這類商的平方根，最適合用半角公式（half-angle formulae）處理：設 $x = 2 - \cos\theta$，分子變成 $1 - \cos\theta$，分母變成 $1 + \cos\theta$，而兩者之比是一個完全平方，即 $\tan^2\frac{\theta}{2}$。至於一般情況的結果，我們不必把運算重做一遍。只要作一個線性變數代換（linear change of variable），把由 a 到 b 的區間映到由 1 到 3 的區間，一般的積分就會變成第一個積分的倍數。

解答

第一個積分

被積函數（integrand）在 $1 \le x < 3$ 上有定義，而我們的積分範圍 $\frac{3}{2} \le x \le 2$ 正好在其中。下圖的陰影面積就是所求的積分。



設 $x = 2 - \cos\theta$，則 $\mathrm{d}x = \sin\theta\,\mathrm{d}\theta$。當 $x = \frac{3}{2}$ 時，$\cos\theta = \frac{1}{2}$；當 $x = 2$ 時，$\cos\theta = 0$，

所以我們取 θ 由 $\frac{\pi}{3}$ 到 $\frac{\pi}{2}$。於是

$$x-1=1-\cos\theta=2\sin^2\tfrac{\theta}{2},\qquad 3-x=1+\cos\theta=2\cos^2\tfrac{\theta}{2},$$

所以

$$\left(\frac{x-1}{3-x}\right)^{\frac12}=\left|\tan\tfrac{\theta}{2}\right|=\tan\tfrac{\theta}{2},$$

這是因為在我們的範圍內 $0<\frac{\theta}{2}<\frac{\pi}{2}$，所以正切值為正。

想一想 當被積函數變成 $\tan\frac{\theta}{2}\sin\theta$ 之後，怎樣才可以避免對正切函數求積分？

寫出 $\sin\theta=2\sin\frac{\theta}{2}\cos\frac{\theta}{2}$。餘弦項會互相抵消：

$$\tan\tfrac{\theta}{2}\,\sin\theta=\frac{\sin\frac{\theta}{2}}{\cos\frac{\theta}{2}}\cdot 2\sin\tfrac{\theta}{2}\cos\tfrac{\theta}{2}=2\sin^2\tfrac{\theta}{2}=1-\cos\theta.$$

因此

$$\begin{aligned}\int_{3/2}^{2}\left(\frac{x-1}{3-x}\right)^{\frac12}\mathrm{d}x&=\int_{\pi/3}^{\pi/2}(1-\cos\theta)\,\mathrm{d}\theta=\Big[\theta-\sin\theta\Big]_{\pi/3}^{\pi/2}\\&=\left(\frac{\pi}{2}-1\right)-\left(\frac{\pi}{3}-\frac{\sqrt3}{2}\right)=\frac{\pi}{6}-1+\frac{\sqrt3}{2}\\&=\frac{\pi+3\sqrt3-6}{6}.\end{aligned}$$

一般的積分

第一個積分就是 $a=1$、$b=3$ 的情況，此時 $p=\frac{3+3}{4}=\frac32$，$q=\frac{1+3}{2}=2$。所以我們要找一個線性代換，把 $u=1$ 送到 $x=a$，把 $u=3$ 送到 $x=b$：

$$x=a+\tfrac12(b-a)(u-1),\qquad \mathrm{d}x=\tfrac12(b-a)\,\mathrm{d}u.$$

於是

$$x-a=\tfrac12(b-a)(u-1),\qquad b-x=(b-a)-\tfrac12(b-a)(u-1)=\tfrac12(b-a)(3-u),$$

又因為 $b-a\neq0$，公因子可以約去：

$$\frac{x-a}{b-x}=\frac{u-1}{3-u}.$$

至於積分上下限：$u=\frac32$ 給出 $x=a+\frac14(b-a)=\frac{3a+b}{4}=p$，而 $u=2$ 給出 $x=a+\frac12(b-a)=\frac{a+b}{2}=q$。由於 $b-a>0$，這個代換是遞增的，所以上下限的次序保持不變。因此

$$\begin{aligned}\int_p^q\left(\frac{x-a}{b-x}\right)^{\frac12}\mathrm{d}x&=\frac{b-a}{2}\int_{3/2}^{2}\left(\frac{u-1}{3-u}\right)^{\frac12}\mathrm{d}u\\&=\frac{b-a}{2}\cdot\frac{\pi+3\sqrt3-6}{6}=\frac{(b-a)(\pi+3\sqrt3-6)}{12},\end{aligned}$$

其中用到了第一個結果。這正是我們要證明的。

（另一個等價的做法，是直接代入 $x = \frac{a+b}{2} - \frac{b-a}{2}\cos\theta$。上下限同樣是 $\frac{\pi}{3}$ 和 $\frac{\pi}{2}$，運算過程也一樣，只是多了一個因子 $\frac{b-a}{2}$。）

驗算

- 數值上，$\frac{1}{6}(\pi + 3\sqrt{3} - 6) \approx 0.38962$。把辛普森法則（Simpson's rule）用於原來的被積函數，在 $[\frac{3}{2}, 2]$ 上計算，所得的值準確至小數點後十位都相同。
- 合理性：被積函數在長度為 $\frac{1}{2}$ 的區間上由 $\sqrt{1/3} \approx 0.577$ 上升至 1，所以面積必定介乎約 0.29 與 0.5 之間。結果確實如此。
- 隨機選取 20 組 $a < b$（包括 a 為負數的情況），一般公式都與辛普森法則的結果吻合。
- 縮放：把區間 $[a, b]$ 按因子 k 拉伸，被積函數的取值不變，但 $[p, q]$ 會拉伸 k 倍。所以答案必定與 $b - a$ 成正比，而事實正是這樣。

常見錯誤

- 忘記了 $\mathrm{d}x = \sin\theta\,\mathrm{d}\theta$，試圖只對 $\tan\frac{\theta}{2}$ 關於 θ 求積分。
- 把上下限取為 $\theta = -\frac{\pi}{3}$ 或 $\frac{5\pi}{3}$。無論怎樣選取，都必須使 θ 由 $\frac{\pi}{3}$ 遞增至 $\frac{\pi}{2}$，並且途經的範圍內 $\sin\theta > 0$ 和 $\tan\frac{\theta}{2} > 0$ 都成立。
- 在一般情況中漏掉 $\mathrm{d}x$ 帶來的因子 $\frac{1}{2}(b-a)$，結果分母變成 6 而不是 12。

STEP 之外

第二部分說明，定積分（definite integral）在仿射變數代換（affine change of variable）之下的變化是有規律可循的。這就是為甚麼像這樣的量只透過長度 $b - a$ 依賴於 a 和 b。被積函數本身屬於貝塔函數（beta function）類型：在整個區間上，$\int_a^b \left(\frac{x-a}{b-x}\right)^{1/2}\mathrm{d}x = \frac{\pi}{2}(b-a)$，而函數 $\int_0^1 t^{\alpha-1}(1-t)^{\beta-1}\,\mathrm{d}t$ 在概率和分析中反覆出現。Blitzstein 與 Hwang 的 *Introduction to Probability*（概率導論）把它們作為貝塔分佈（Beta distribution）的歸一化常數（normalising constants）來介紹。Bartle 與 Sherbert 的 *Introduction to Real Analysis*（實分析導論）則嚴謹地證明了代換法則（substitution rule）。

詳細解答

STEP III 2000，第 3 題

純數學 • 複數與幾何

> **STEP III 2000，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/00-s3-q3

本題在 *Beyond STEP* 第 5.6 節費馬點中有完整的引導式講解：https://beyondstep.keithmcnulty.org/zh-hant/ch05/fermat-point.html

關鍵思路

在阿根圖（Argand plane）上，把一個向量旋轉 60° 只不過是乘以 $\alpha = e^{i\pi/3}$，所以在給定邊上作等邊三角形（equilateral triangle），其第三個頂點一行就可以寫出來。其後的一切都來自恆等式（identity）$1 + \alpha^2 = \alpha$ 及其相關的式子，例如 $\alpha^3 = -1$。它們把各頂點的表達式化成非常簡單的形式，一旦看出這些形式，最後一部分幾乎立即可得。

解答

恆等式

由於 $\alpha^2 = e^{2i\pi/3} = -\frac{1}{2} + \frac{\sqrt{3}}{2}i$，

$$1 + \alpha^2 = \tfrac{1}{2} + \tfrac{\sqrt{3}}{2}i = \cos\tfrac{\pi}{3} + i\sin\tfrac{\pi}{3} = \alpha.$$

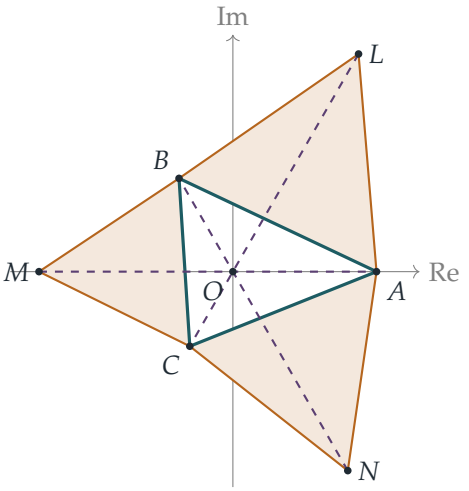
我們還會用到以下幾個推論：

- $\alpha^3 = e^{i\pi} = -1$；
- $1 - \alpha = -\alpha^2$（這是把恆等式移項所得）；
- $\alpha^2 - \alpha = -1$（同樣是移項所得）；
- $1 - \alpha = \bar{\alpha} = e^{-i\pi/3}$，因為 $\alpha + \bar{\alpha} = 2\cos\frac{\pi}{3} = 1$。

最後一項說明，乘以 $1-\alpha$ 相當於順時針旋轉 $60°$。

三角形 ABC

$A = p$ 位於正實軸上，$B = q\alpha^2 = q\,\mathrm{e}^{2\mathrm{i}\pi/3}$ 位於輻角（argument）為 $120°$ 的射線上，而 $C = -r\alpha = r\,\mathrm{e}^{4\mathrm{i}\pi/3}$ 位於輻角為 $240°$ 的射線上。因此三個頂點分別落在由 O 出發、彼此相隔 $120°$ 的三條射線上，與 O 的距離分別為 p, q, r。原點在三角形之內，而 $A \to B \to C$ 是逆時針方向。下圖取 $p = 2$、$q = 1.5$、$r = 1.2$，並同時畫出在各邊上所作的三角形。



頂點 N、L 和 M

若 XYZ 是一個按順時針方向標記頂點的等邊三角形，那麼從 X 出發來看，邊 XZ 就是邊 XY 繞 X 順時針旋轉 $60°$ 所得。因此

$$Z - X = (Y - X)\,\mathrm{e}^{-\mathrm{i}\pi/3} = (Y - X)(1 - \alpha), \qquad \text{即} \qquad Z = \alpha X + (1 - \alpha)Y.$$

? **想一想** 為甚麼「順時針」會令新作的三角形落在 ABC 的外面？

由於 $A \to B \to C$ 是逆時針方向，C 位於有向邊 AB 的左方。把 AB 繞 A 順時針旋轉，它便擺向右方，所以 L 與 C 分別位於 AB 的兩側。其餘兩條邊的情況相同，這與圖中所見吻合。

N（三角形 CAN，故 $X = C$、$Y = A$）：

$$N = \alpha(-r\alpha) + (1 - \alpha)p = (1 - \alpha)p - \alpha^2 r,$$

正如所求。利用 $1 - \alpha = -\alpha^2$，可把它整理成 $N = -\alpha^2(p + r)$。

L（三角形 ABL，故 $X=A$、$Y=B$）：

$$L=\alpha p+(1-\alpha)q\alpha^2=\alpha p+q(\alpha^2-\alpha^3)=\alpha p+q(1+\alpha^2)=(p+q)\alpha,$$

當中先用 $\alpha^3=-1$，再用 $1+\alpha^2=\alpha$。

M（三角形 BCM，故 $X=B$、$Y=C$）：

$$M=\alpha\cdot q\alpha^2+(1-\alpha)(-r\alpha)=q\alpha^3+r(\alpha^2-\alpha)=-q-r=-(q+r),$$

當中用到 $\alpha^3=-1$ 和 $\alpha^2-\alpha=-1$。

因此，仿照題目所給 N 的形式，

$$L=\alpha p+\alpha q,\qquad M=-q-r,\qquad N=(1-\alpha)p-\alpha^2 r.$$

三條直線

現在把每一對點都看成某個單位複數的正數倍：

$$\begin{aligned} L&=(p+q)\,\mathrm{e}^{\mathrm{i}\pi/3}, & C&=-r\alpha=r\,\mathrm{e}^{\mathrm{i}\pi/3}\cdot(-1),\\ M&=(q+r)\cdot(-1), & A&=p\cdot 1,\\ N&=(p+r)\cdot(-\alpha^2), & B&=q\,\alpha^2. \end{aligned}$$

在每一對之中，兩點分別是 u 和 $-u$ 的正實數倍，其中 u 是一個單位複數(分別為 $u=\alpha$、$u=1$ 和 $u=\alpha^2$)。所以這兩點位於同一條通過 O 的直線上，並且分處 O 的兩側。由此可知，直線 LC、MA、NB 各自都通過原點，而且原點其實位於每一條線段上。又因為兩點分處 O 的兩側，每條線段的長度就是兩點與 O 的距離之和：

$$LC=(p+q)+r,\qquad MA=(q+r)+p,\qquad NB=(p+r)+q,$$

全部等於 $p+q+r$，正如所求。

驗算

取 $p=1$、$q=2$、$r=3$，由公式得 $L=3\alpha\approx 1.5+2.598\mathrm{i}$、$M=-5$ 和 $N=-4\alpha^2\approx 2-3.464\mathrm{i}$。計算得 $|L-A|=|L-B|=|A-B|=\sqrt{7}$，確認 ABL 是等邊三角形。我們又用電腦直接由 $\pm 60^\circ$ 的兩個旋轉作出每個三角形，保留順時針的那一個，並就 10 000 組隨機選取的 p,q,r 確認了三條公式、三線共點（concurrence）於 O，以及共同長度 $p+q+r$。一個漂亮的特殊情況：當 $p=q=r$ 時，三角形 ABC 是以 O 為中心的等邊三角形，而整個圖形具有預期中的三重旋轉對稱。

常見錯誤

- 旋轉方向弄錯，乘了 α 而不是 $\bar{\alpha} = 1 - \alpha$。這樣 L、M、N 會落在三角形之內，題目所給 N 的形式也推不出來。
- 只證明了 O 在每條直線上，卻忘記要得出長度，還需要 O 位於兩個端點之間。各實數倍的正負號正好解決這一點。

STEP 之外

這是拿破崙構形（Napoleon's configuration）的一個特殊情況：對任意三角形，在各邊向外作等邊三角形，直線 AM、BN、CL 都交於同一點（費馬點，Fermat point），而且三條線段長度相等。費馬點使它到三個頂點的距離總和達到最小；在本題中它就是 O，從這一點看每條邊所張的角都是 120°。複數把旋轉變成乘法，這正是此類證明可以化為簡短代數運算的原因。Needham 的 *Visual Complex Analysis*（《複分析：可視化方法》）把這種幾何觀點發揮得十分精彩。

詳細解答

STEP III 2000，第 4 題

純數學 • 有理函數的描繪

STEP III 2000，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/00-s3-q4

關鍵思路

令 $f(x) = x + c$ 並消去分母後，三次項互相抵消，餘下的是一個關於 x 的二次方程（quadratic）。於是直線與曲線是否相交，就取決於一個判別式（discriminant）的正負；而這個判別式本身又是 c 的二次式，所以我們要求「判別式的判別式」。至於描繪圖像，把 f 拆成線性部分加上部分分式（partial fractions），就能得知一切：各條漸近線（asymptote）、曲線位於漸近線的哪一側，以及曲線在何處穿過斜漸近線（oblique asymptote）。

解答

直線與曲線在何處相交

當 $x \neq \pm 1$ 時，方程 $f(x) = x + c$ 等價於

$$x(x-2)(x-a) = (x+c)(x^2-1).$$

展開得 $x^3 - (a+2)x^2 + 2ax = x^3 + cx^2 - x - c$。兩邊的 x^3 項互相抵消，餘下

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

因此，直線與曲線相交，當且僅當 $(*)$ 有一個不等於 $x = \pm 1$ 的實根（real root）。

? **想一想** $(*)$ 的根會不會恰好是被排除的值 $x = \pm 1$ 之一？

代入可得 $q(1) = (a+2+c) - (2a+1) - c = 1 - a$ 及 $q(-1) = (a+2+c) + (2a+1) - c = 3(a+1)$。兩者都與 c 無關，所以只有當 $a = 1$ 時 $x = 1$ 才是根，只有當 $a = -1$ 時 $x = -1$ 才是根。當 $|a| \neq 1$ 時，$(*)$ 的每一個實根都對應一個真正的交點。

$a+2+c=0$ 的情況。這時 $(*)$ 是線性方程：$-(2a+1)x+(a+2)=0$。若 $|a|>1$，則 $2a+1\neq 0$，所以有根 $x=\frac{a+2}{2a+1}$。

$a+2+c\neq 0$ 的情況。這時 $(*)$ 是真正的二次方程，其判別式為

$$D(c)=(2a+1)^2+4c(a+2+c)=4c^2+4(a+2)c+(2a+1)^2.$$

把 D 看成 c 的二次式，其首項係數（leading coefficient）為正。它自身的判別式是

$$\begin{aligned}16(a+2)^2-16(2a+1)^2&=16\big[(a+2)-(2a+1)\big]\big[(a+2)+(2a+1)\big]\\&=16(1-a)\cdot 3(1+a)=48(1-a^2).\end{aligned}$$

若 $|a|>1$，上式為負，所以對任何 c 都有 $D(c)>0$；於是 $(*)$ 有兩個相異實根，而且都不等於 ±1。連同線性的情況，可知對任何 c，直線 $y=x+c$ 都與曲線相交。

若 $|a|=1$，對任何 c 都有 $D(c)\geq 0$，所以 $(*)$ 仍然總有實根，這就是該命題在代數上的內容。不過這時曲線本身會退化：分子和分母約去一個公因式，曲線變成一條缺了一點的雙曲線（hyperbola）。例如，取 $a=1$ 時，我們得到 $f(x)=x-3+\frac{3}{x+1}$ （$x\neq 1$），而直線 $y=x-3$（即它的漸近線）與它並不相交；在 $(*)$ 中這對應 $c=-3$，此時唯一的根正是被排除的 $x=1$。所以當 $|a|=1$ 時，除了這些例外的 c 值之外，命題都成立；而當 $|a|>1$ 時，命題毫無例外地成立。

若 $|a|<1$，二次式 $D(c)$ 有兩個實根，並且在兩根之間嚴格為負。一個明確的選擇是 $c=-\frac{1}{2}(a+2)$，即兩根的中點：

$$D\big(-\tfrac{1}{2}(a+2)\big)=(a+2)^2-2(a+2)^2+(2a+1)^2=3(a^2-1)<0.$$

對於這個 c，$a+2+c=\frac{1}{2}(a+2)\neq 0$（因為 $a>-1$），所以 $(*)$ 是一個判別式為負的真正二次方程，沒有實根。因此直線 $y=x-\frac{1}{2}(a+2)$ 與曲線不相交，這正是所要證明的。

斜漸近線

以 x^2-1 除 $x^3-(a+2)x^2+2ax$，得

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

當 $x\to\pm\infty$ 時，該分式趨向 0，所以斜漸近線是

$$y=x-a-2.$$

為了描繪圖像，把該分式拆成部分分式會很有幫助：

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

由此我們得知一切：

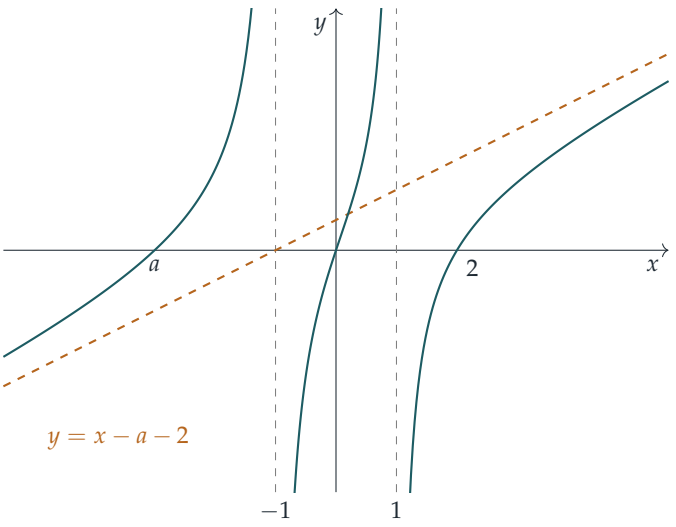
- 零點位於 $x = 0,\ 2,\ a$，鉛垂漸近線（vertical asymptote）位於 $x = \pm 1$；
- 曲線只在 $(2a+1)x = a+2$ 之處與斜漸近線相交，即在 $x_0 = \frac{a+2}{2a+1}$（這又是 $(*)$ 的線性情況）；
- 當 $|x|$ 很大時，差距 $f(x) - (x-a-2)$ 的表現有如 $\frac{2a+1}{x}$。在兩種情況中都有 $2a+1 < 0$，所以當 $x \to +\infty$ 時曲線位於漸近線下方，而當 $x \to -\infty$ 時則位於其上方；
- 係數 $\frac{a-1}{2}$ 和 $\frac{3(a+1)}{2}$ 決定了曲線在 $x = 1$ 和 $x = -1$ 兩側的表現。

情況 (i)：$a < -1$

兩個係數都是負數：$\frac{a-1}{2} < 0$ 及 $\frac{3(a+1)}{2} < 0$。所以當 $x \to 1^-$ 或 $x \to -1^-$ 時 $f(x) \to +\infty$，而當 $x \to 1^+$ 或 $x \to -1^+$ 時 $f(x) \to -\infty$。此外

$$f'(x) = 1 - \frac{a-1}{2(x-1)^2} - \frac{3(a+1)}{2(x+1)^2} > 0,$$

因為被減去的兩項都是負數，所以 f 在它的三個分支上都是遞增的，完全沒有轉向點（turning point）。交點滿足 $x_0 + 1 = \frac{3(a+1)}{2a+1} > 0$ 及 $x_0 - 1 = \frac{1-a}{2a+1} < 0$，所以 $-1 < x_0 < 1$：曲線在中間的分支上穿過它的漸近線。零點分別位於 $a < -1$（左分支）、0（中間分支）和 2（右分支）。



（圖中取 $a = -3$，這時漸近線是 $y = x + 1$，而 $x_0 = \frac{1}{5}$。）

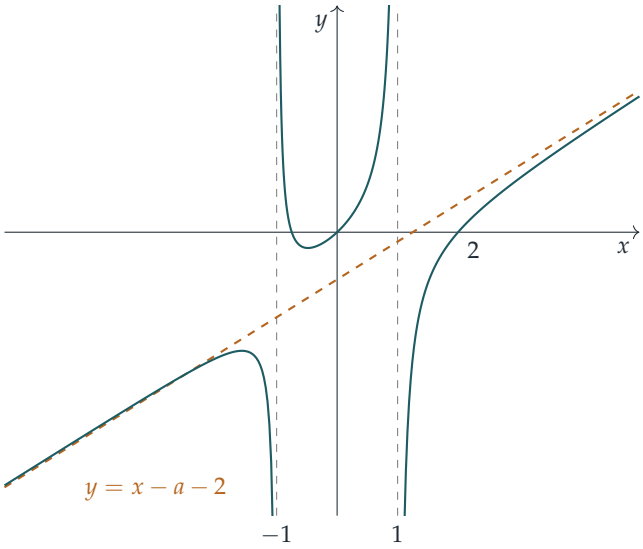
情況 (ii)：$-1 < a < -\frac{1}{2}$

現在 $\frac{a-1}{2} < 0$，但 $\frac{3(a+1)}{2} > 0$。在 $x = 1$ 附近，表現與之前一樣（從左邊趨向 $+\infty$，從右邊趨向 $-\infty$），但在 $x = -1$ 附近則反轉過來：當 $x \to -1^-$ 時 $f(x) \to -\infty$，而當

$x \to -1^+$ 時 $f(x) \to +\infty$。現在 $x_0 + 1 = \frac{3(a+1)}{2a+1} < 0$，所以 $x_0 < -1$：交點在左分支上。零點 a 和 0 都位於 $(-1, 0)$ 之內。

想一想 雖然題目沒有要求我們求出轉向點，但它們必定在哪裏？

在左分支上，曲線在漸近線上方從 $-\infty$ 升上來，在 x_0 處穿過漸近線，然後當 $x \to -1^-$ 時跌向 $-\infty$；所以它有一個極大點（maximum）。曲線在這個分支上沒有零點（三個零點全都 > -1），所以整個分支都位於 x 軸下方。在中間分支上，$f \to +\infty$ 於兩端成立，曲線在 a 與 0 之間降到軸的下方，在其他地方則為正，所以在 a 與 0 之間有一個極小點（minimum）。在右分支上，當 $x > 1$ 時 $\frac{3(a+1)}{2(x+1)^2} < \frac{3}{4} \cdot \frac{1}{4} < 1$，所以 $f'(x) > 0$：曲線從 $-\infty$ 上升，經過 $x = 2$，並從下方趨近漸近線。



（圖中取 $a = -\frac{3}{4}$：漸近線為 $y = x - \frac{5}{4}$，交點在 $x_0 = -\frac{5}{2}$，而曲線在 $-\frac{3}{4}$ 與 0 之間略為降到軸的下方。）

驗算

- 取 $a = 0$ 及 $c = -1$（即我們所選的 $c = -\frac{1}{2}(a+2)$），$(*)$ 變成 $x^2 - x + 1 = 0$，其判別式為 -3：沒有交點，與預期相符。
- 取 $a = -3$：$f(-1.1) \approx 30.9$ 及 $f(-0.9) \approx -28.8$，與情況 (i) 中就 $x \to -1^{\mp}$ 所得的正負號吻合。取 $a = -\frac{3}{4}$，則 $f(-1.1) \approx -5.7$ 及 $f(-0.9) \approx 2.1$，正是情況 (ii) 中反轉了的表現。
- 在每個範圍內取若干個 a 值，以數值方法數算 f' 的變號次數，可以確認情況 (i) 中沒

有轉向點，而情況 (ii) 中左分支和中間分支各恰有一個轉向點。

常見錯誤

- 忘記 $(*)$ 的根 $x = \pm 1$ 並不是交點；當 $|a| = 1$ 時出錯的正是這一點。
- 忽略 $a + 2 + c = 0$ 的情況，這時 $(*)$ 是線性方程，判別式的論證並不適用。
- 在情況 (ii) 中弄錯曲線在鉛垂漸近線 $x = -1$ 兩側的走向：$a + 1$ 的正負號會改變那裏的圖像。

STEP 之外

當 $|a| < 1$ 時，$D(c)$ 的兩個根給出與曲線相切的直線 $y = x + c$，即切線 (tangent)，而介乎兩者之間的直線則與曲線不相交。當 a 越過 ± 1 時，這些切線會出現或消失，曲線的形狀也隨之改變：一個零點穿過一條鉛垂漸近線，漸近線旁邊的表現隨即反轉。研究一族曲線的定性形狀如何隨參數 (parameter) 而改變，正是分歧理論 (bifurcation theory) 的起點，在大學一年級的微分方程 (differential equations) 課程中就會遇到；可參閱 Boyce 與 DiPrima 的 *Elementary Differential Equations*（微分方程入門教科書）。若想多練習這類圖像描繪，Siklos 的 *Advanced Problems in Mathematics*（進階數學難題集）是極佳的選擇。

詳細解答

STEP III 2000，第 5 題

純數學 • 線性變換

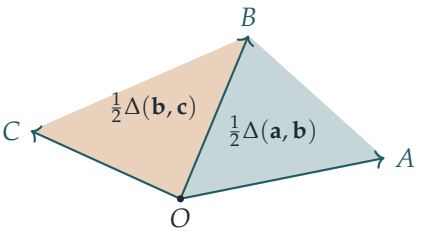
STEP III 2000，第 5 題

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/00-s3-q5

關鍵思路

$\Delta(\mathbf{a},\mathbf{b})$ 是以 $\mathbf{a}$ 和 $\mathbf{b}$ 為列的矩陣（matrix）的行列式（determinant），所以它是三角形 OAB 有向面積（signed area）的兩倍，而且恰好在 $\mathbf{a}$ 與 $\mathbf{b}$ 不平行時才不為零。線性映射（linear map）會把所有這類面積乘以同一個因子，即它的行列式。單憑這一點，必要性便立刻得出。至於充分性，(ii) 中的恆等式（identity）只用各個 Δ 就把 $\mathbf{c}$ 以 $\mathbf{a}$ 和 $\mathbf{b}$ 表示出來，所以一旦 $\mathbf{M}$ 由 A 和 B 確定下來，C 被送到哪裏便再無選擇餘地。



解答

以下一律以 $[\mathbf{u}\ \mathbf{v}]$ 表示以 $\mathbf{u}$ 和 $\mathbf{v}$ 為列的 2×2 矩陣。於是

$$\Delta(\mathbf{u},\mathbf{v}) = u_1v_2 - u_2v_1 = \det[\mathbf{u}\ \mathbf{v}].$$

對於非零的 $\mathbf{u}$、$\mathbf{v}$，$\Delta(\mathbf{u},\mathbf{v}) = 0$ 恰好在 $\mathbf{u}$ 與 $\mathbf{v}$ 平行時成立：行列式為零，當且僅當其中一列是另一列的倍數。因此 $\Delta(\mathbf{a},\mathbf{b})$、$\Delta(\mathbf{b},\mathbf{c})$、$\Delta(\mathbf{c},\mathbf{a})$、$\Delta(\mathbf{p},\mathbf{q})$、$\Delta(\mathbf{q},\mathbf{r})$、$\Delta(\mathbf{r},\mathbf{p})$ 這六個數全部不為零。

還有兩個事實會反覆用到。由於 $\mathbf{M}[\mathbf{u}\ \mathbf{v}] = [\mathbf{Mu}\ \ \mathbf{Mv}]$（逐列計算），而且乘積的行列式

等於行列式的乘積，

$$\Delta(\mathbf{Mu},\mathbf{Mv}) = \det\mathbf{M}\,\Delta(\mathbf{u},\mathbf{v}). \qquad (*)$$

另外，$\Delta(\mathbf{v},\mathbf{u}) = -\Delta(\mathbf{u},\mathbf{v})$。

第 (i) 部分

我們需要 $\mathbf{Ma} = \mathbf{p}$ 和 $\mathbf{Mb} = \mathbf{q}$，兩者合起來就是

$$\mathbf{M}\,[\mathbf{a}\ \mathbf{b}] = [\mathbf{p}\ \mathbf{q}].$$

由於 $\mathbf{a}$ 與 $\mathbf{b}$ 不平行，$\det[\mathbf{a}\ \mathbf{b}] = \Delta(\mathbf{a},\mathbf{b}) \neq 0$，所以 $[\mathbf{a}\ \mathbf{b}]$ 可逆（invertible），而

$$\mathbf{M} = [\mathbf{p}\ \mathbf{q}]\,[\mathbf{a}\ \mathbf{b}]^{-1} = \frac{1}{\Delta(\mathbf{a},\mathbf{b})}\begin{pmatrix} p_1 & q_1 \\ p_2 & q_2 \end{pmatrix}\begin{pmatrix} b_2 & -b_1 \\ -a_2 & a_1 \end{pmatrix}$$

正符合要求。事實上，這個 $\mathbf{M}$ 是唯一符合要求的矩陣，因為我們可以在右邊乘以逆矩陣。取行列式，得 $\det\mathbf{M} = \Delta(\mathbf{p},\mathbf{q})/\Delta(\mathbf{a},\mathbf{b})$，此數不為零。

第 (ii) 部分：恆等式

取左方的第一個分量：

$$\begin{aligned}&(a_1b_2 - a_2b_1)c_1 + (c_1a_2 - c_2a_1)b_1 + (b_1c_2 - b_2c_1)a_1\\ &\quad = a_1b_2c_1 - a_2b_1c_1 + a_2b_1c_1 - a_1b_1c_2 + a_1b_1c_2 - a_1b_2c_1 = 0,\end{aligned}$$

因為六項兩兩相消。第二個分量的計算完全相同，只是把各個 Δ 以外的 $\mathbf{a}$、$\mathbf{b}$、$\mathbf{c}$ 的下標全部由 1 改為 2：

$$\begin{aligned}&(a_1b_2 - a_2b_1)c_2 + (c_1a_2 - c_2a_1)b_2 + (b_1c_2 - b_2c_1)a_2\\ &\quad = a_1b_2c_2 - a_2b_1c_2 + a_2b_2c_1 - a_1b_2c_2 + a_2b_1c_2 - a_2b_2c_1 = 0.\end{aligned}$$

所以

$$\Delta(\mathbf{a},\mathbf{b})\,\mathbf{c} + \Delta(\mathbf{c},\mathbf{a})\,\mathbf{b} + \Delta(\mathbf{b},\mathbf{c})\,\mathbf{a} = \mathbf{0}, \qquad (1)$$

這正是所要證明的。這個恆等式對任意三個向量（vector）都成立，所以對 $\mathbf{p}$、$\mathbf{q}$、$\mathbf{r}$ 同樣成立：

$$\Delta(\mathbf{p},\mathbf{q})\,\mathbf{r} + \Delta(\mathbf{r},\mathbf{p})\,\mathbf{q} + \Delta(\mathbf{q},\mathbf{r})\,\mathbf{p} = \mathbf{0}. \qquad (2)$$

想一想 這個比例條件寫成等式，確切是在說甚麼？

所有 Δ 都不為零，所以條件 $\Delta(\mathbf{a},\mathbf{b}) : \Delta(\mathbf{b},\mathbf{c}) : \Delta(\mathbf{c},\mathbf{a}) = \Delta(\mathbf{p},\mathbf{q}) : \Delta(\mathbf{q},\mathbf{r}) : \Delta(\mathbf{r},\mathbf{p})$ 的意思是：存在同一個數 $k \neq 0$，使得

$$\Delta(\mathbf{p},\mathbf{q}) = k\,\Delta(\mathbf{a},\mathbf{b}), \qquad \Delta(\mathbf{q},\mathbf{r}) = k\,\Delta(\mathbf{b},\mathbf{c}), \qquad \Delta(\mathbf{r},\mathbf{p}) = k\,\Delta(\mathbf{c},\mathbf{a}). \qquad (3)$$

用文字來說：三角形 OPQ、OQR、ORP 的三個有向面積，就是 OAB、OBC、OCA 的面積全部按同一個因子縮放的結果。

第 (ii) 部分：必要性

假設 $\mathbf{Ma}=\mathbf{p}$、$\mathbf{Mb}=\mathbf{q}$ 及 $\mathbf{Mc}=\mathbf{r}$。由 $(*)$，

$$\Delta(\mathbf{p},\mathbf{q})=\det\mathbf{M}\,\Delta(\mathbf{a},\mathbf{b}),\quad \Delta(\mathbf{q},\mathbf{r})=\det\mathbf{M}\,\Delta(\mathbf{b},\mathbf{c}),\quad \Delta(\mathbf{r},\mathbf{p})=\det\mathbf{M}\,\Delta(\mathbf{c},\mathbf{a}).$$

這就是 $k=\det\mathbf{M}$ 時的 (3)，而且因為 $\Delta(\mathbf{p},\mathbf{q})\neq 0$，所以 $k\neq 0$。因此兩組比相等。

第 (ii) 部分：充分性

現在假設 (3) 對某個 $k\neq 0$ 成立。由第 (i) 部分，存在矩陣 $\mathbf{M}$ 使得 $\mathbf{Ma}=\mathbf{p}$ 及 $\mathbf{Mb}=\mathbf{q}$；我們必須證明這同一個 $\mathbf{M}$ 也把 $\mathbf{c}$ 送到 $\mathbf{r}$。

把 $\mathbf{M}$ 作用於 (1)。矩陣乘法是線性的，而各個 Δ 只是數，所以

$$\Delta(\mathbf{a},\mathbf{b})\,\mathbf{Mc}=-\Delta(\mathbf{c},\mathbf{a})\,\mathbf{q}-\Delta(\mathbf{b},\mathbf{c})\,\mathbf{p}.$$

兩邊乘以 k，並對每一項使用 (3)：

$$\Delta(\mathbf{p},\mathbf{q})\,\mathbf{Mc}=-\Delta(\mathbf{r},\mathbf{p})\,\mathbf{q}-\Delta(\mathbf{q},\mathbf{r})\,\mathbf{p}.$$

由 (2)，右方等於 $\Delta(\mathbf{p},\mathbf{q})\,\mathbf{r}$。所以

$$\Delta(\mathbf{p},\mathbf{q})\,\mathbf{Mc}=\Delta(\mathbf{p},\mathbf{q})\,\mathbf{r},$$

又因為 $\Delta(\mathbf{p},\mathbf{q})\neq 0$（$\mathbf{p}$、$\mathbf{q}$ 不平行），故 $\mathbf{Mc}=\mathbf{r}$。因此 P、Q、R 就是 A、B、C 在 $\mathbf{M}$ 之下的像（image）。

必要性與充分性合起來，便證明了這個比例條件是充分必要條件（necessary and sufficient）。

驗算

- 試取 $\mathbf{a}=\binom{1}{0}$、$\mathbf{b}=\binom{0}{1}$、$\mathbf{c}=\binom{1}{1}$。這時 $\Delta(\mathbf{a},\mathbf{b})=1$、$\Delta(\mathbf{b},\mathbf{c})=-1$、$\Delta(\mathbf{c},\mathbf{a})=-1$，而 (1) 就是 $\binom{1}{1}-\binom{0}{1}-\binom{1}{0}=\mathbf{0}$。取 $\mathbf{M}=\begin{pmatrix}2&1\\0&3\end{pmatrix}$，得 $\mathbf{p}=\binom{2}{0}$、$\mathbf{q}=\binom{1}{3}$、$\mathbf{r}=\binom{3}{3}$，三個 Δ 分別是 6、-6、-6：比例仍然是 $1:-1:-1$，其中 $k=6=\det\mathbf{M}$。
- 在數值方面，我們用精確分數運算，以 20,000 組隨機整數向量三元組檢驗了這個恆等式。在其中約 17,700 組沒有任何兩個向量平行的三元組中，我們按第 (i) 部分構造 $\mathbf{M}$，確認了必要性，並確認 $\mathbf{Mc}=\mathbf{r}$ 恰好在 (3) 成立時才成立。

常見錯誤

- 把比例條件寫成一條分數等式，卻沒有說明各個 Δ 不為零。題目假設任何兩個向量都不平行，正是為了保證每一次除法都合法。
- 只證明必要性。真正的內容在於充分性：矩陣由它對兩個線性獨立（independent）向量的作用所確定，所以這個條件必須能迫使第三個像符合要求。
- 因 Δ 內的次序而弄錯正負號：循環次序 $\mathbf{b},\mathbf{c}$ 然後 $\mathbf{c},\mathbf{a}$ 是有關係的，因為 $\Delta(\mathbf{a},\mathbf{c}) = -\Delta(\mathbf{c},\mathbf{a})$。

STEP 之外

恆等式 (1) 其實是克萊姆法則（Cramer's rule）的另一種面貌：它以行列式之比給出 $\mathbf{c}$ 在基（basis）$\mathbf{a}$、$\mathbf{b}$ 之下的坐標。事實 $(*)$，即線性映射把每個有向面積都按 $\det \mathbf{M}$ 縮放，正是行列式的幾何定義，而二重積分（double integral）換元時出現的雅可比（Jacobian）因子，背後也是這個道理。Strang 的 *Introduction to Linear Algebra*（《線性代數導論》）對這兩個想法都有詳細闡述。

詳細解答

STEP III 2000，第 6 題

純數學 • 四次多項式的因式分解

> **STEP III 2000，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/00-s3-q6

關鍵思路

比較係數（comparing coefficients）可把這個因式分解（factorisation）化為三條把 p, q, r 與 a, b, c 聯繫起來的方程。要得到只含 a 的方程，就必須消去 b 和 c，而恆等式（identity）$4bc = (b+c)^2 - (b-c)^2$ 正好做到這一點，因為 $b+c$ 和 $b-c$ 正是自然出現的組合。至於最後的四次式（quartic），須先消去 y^3 項，才能用上那條三次方程（cubic），而平移 $y = x + k$ 便可做到。

解答

求 p、q 和 r

展開右邊，得

$$(x^2 - ax + b)(x^2 + ax + c) = x^4 + (b + c - a^2)x^2 + a(b - c)x + bc,$$

其中 x^3 項 $ax^3 - ax^3$ 互相抵消。比較係數，得

$$p = b + c - a^2, \qquad q = a(b - c), \qquad r = bc.$$

a^2 所滿足的三次方程

設 $u = a^2$。由第一條方程，$b + c = p + u$。

若 $a \neq 0$，第二條方程給出 $b - c = q/a$，所以 $(b-c)^2 = q^2/u$。於是

$$4bc = (b+c)^2 - (b-c)^2 \quad \Longrightarrow \quad 4r = (p+u)^2 - \frac{q^2}{u}.$$

兩邊乘以 $u \neq 0$，

$$4ru = u(u^2 + 2pu + p^2) - q^2 \quad \Longrightarrow \quad u^3 + 2pu^2 + (p^2 - 4r)u - q^2 = 0,$$

所以 $u = a^2$ 是一個根（root）。

若 $a = 0$，則 $q = a(b - c) = 0$，而 $u = a^2 = 0$ 也是一個根，因為這時三次式的每一項在 $u = 0$ 時都等於零。因此在任何情況下 a^2 都是一個根，正如題目所求。

想一想 為甚麼不論 p、q 和 r 是甚麼，都必定有一個 $u \geq 0$ 的根？

把這個三次式記為 $f(u)$。在 $u = 0$ 處有 $f(0) = -q^2 \leq 0$，而由於首項是 u^3，當 $u \to +\infty$ 時 $f(u) \to +\infty$。若 $q = 0$，則 $u = 0$ 本身就是一個根。否則 $f(0) < 0$，而對某個足夠大的 U 有 $f(U) > 0$；由於 f 是連續（continuous）的，根據介值定理（intermediate value theorem），在 0 與 U 之間（不含兩端）必有一個根。無論哪種情況，都有一個非負根。

這一點很重要，因為我們需要 $a = \sqrt{u}$ 是實數。非負根必定給出實數 a，所以每個不含 x^3 項的實係數四次式，都可以分解為兩個符合題目形式的實係數二次式（quadratic）。（若 $u > 0$，我們可以由 $b + c = p + u$ 和 $b - c = q/\sqrt{u}$ 求回 b 和 c。）

$p = -1$、$q = -6$、$r = 15$ 的情況

這裏 $p^2 - 4r = 1 - 60 = -59$，所以三次方程是

$$u^3 - 2u^2 - 59u - 36 = 0.$$

在 $u = 9$ 處：$729 - 162 - 531 - 36 = 0$，所以 $u = 9$ 是一個根。（除去這個因式後得 $(u - 9)(u^2 + 7u + 4)$，其中二次因式有兩個負根，所以 $u = 9$ 是唯一的非負根。）

分解四次式

題目給出的四次式含有 y^3 項，所以我們先作平移把它消去。各根之和為 8，故其平均值為 2，於是設 $y = x + 2$。逐項展開：

$$\begin{aligned}
(x+2)^4 &= x^4 + 8x^3 + 24x^2 + 32x + 16,\\
-8(x+2)^3 &= -8x^3 - 48x^2 - 96x - 64,\\
23(x+2)^2 &= 23x^2 + 92x + 92,\\
-34(x+2) &= -34x - 68,
\end{aligned}$$

把以上各式與 39 相加，得

$$y^4 - 8y^3 + 23y^2 - 34y + 39 = x^4 - x^2 - 6x + 15.$$

這正是 $p = -1$、$q = -6$、$r = 15$ 的情況，這就是題目先考慮該情況的原因。

取 $u = a^2 = 9$，$a = 3$。則

$$b + c = p + a^2 = 8, \qquad b - c = \frac{q}{a} = -2,$$

所以 $b = 3$，$c = 5$。驗算一下：$bc = 15 = r$。因此

$$x^4 - x^2 - 6x + 15 = (x^2 - 3x + 3)(x^2 + 3x + 5).$$

最後代入 $x = y - 2$：

$$x^2 - 3x + 3 = (y^2 - 4y + 4) - (3y - 6) + 3 = y^2 - 7y + 13,$$
$$x^2 + 3x + 5 = (y^2 - 4y + 4) + (3y - 6) + 5 = y^2 - y + 3.$$

所以

$$y^4 - 8y^3 + 23y^2 - 34y + 39 = (y^2 - 7y + 13)(y^2 - y + 3).$$

若改取 $a = -3$，則 b 和 c 互換，得出同樣的兩個因式，只是次序相反。

驗算

- 展開答案：$(y^2 - 7y + 13)(y^2 - y + 3) = y^4 - 8y^3 + (3 + 7 + 13)y^2 - (21 + 13)y + 39$，與原式相符。
- 代入 $y = 1$：四次式的值是 $1 - 8 + 23 - 34 + 39 = 21$，而兩個因式給出 $7 \times 3 = 21$。
- 兩個二次因式的判別式（discriminant）都是負數（分別為 -3 和 -11），所以兩者在實數範圍內都不能再分解，而這個四次式沒有實根。這與 $y = 1$ 時取正值相符。

我們也隨機選取了 2000 組有理數 a, b, c 來檢驗 p, q, r 的公式和那條三次方程，並對 2000 組隨機的 (p, q, r) 用二分法（bisection）找到了非負根。

常見錯誤

- 除以 a 得出 $b - c = q/a$，卻沒有處理 $a = 0$ 的情況。
- 以「三次方程必有實根」作為論據：這句話雖然正確，卻沒有證明該根是非負的。關鍵在於 $f(0) = -q^2$ 的正負號。
- 最後忘記由 x 代回 y。

STEP 之外

這基本上就是笛卡兒（Descartes）（以及費拉里（Ferrari））解一般四次方程的方法：先用平移消去三次項，再透過解一條輔助的三次方程，把四次式分解為兩個二次式；這條

輔助方程稱為預解三次方程（resolvent cubic）。配合卡爾達諾（Cardano）的三次方程公式，這個方法說明了每一條四次方程都可以用根式求解（solvable by radicals）。為甚麼這種做法到四次為止，一般的五次方程（quintic）再沒有這樣的公式，正是伽羅瓦理論（Galois theory）所研究的課題。Allenby 的 *Rings, Fields and Groups*（環、域與羣）是很好的入門讀物。

詳細解答

STEP III 2000，第 7 題

純數學 • 不等式與極限

STEP III 2000，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/00-s3-q7

關鍵思路

兩個不等式都是把一個乘積與我們已經熟悉的東西作比較而得出的。用二項式定理（binomial theorem）展開 $\left(1+\frac{1}{n}\right)^n$，所得各項與 $\frac{1}{r!}$ 十分相似，只是略小一點。至於乘積 $\mathrm{P}(n)$，算術平均–幾何平均不等式（AM–GM inequality）可以把 n 個因子的乘積化為它們平均值的 n 次方。這些因子是 $1+\frac{1}{2^k}$，它們的平均值剛好略小於 $1+\frac{1}{n}$，於是便直接連回第一部分。

解答

上界 $\left(1+\frac{1}{n}\right)^n < \mathrm{e}$

由二項式定理，

$$\left(1+\frac{1}{n}\right)^n = \sum_{r=0}^{n}\binom{n}{r}\frac{1}{n^r} = \sum_{r=0}^{n}\frac{1}{r!}\cdot\frac{n(n-1)(n-2)\cdots(n-r+1)}{n^r}.$$

第二個分數是 r 個因子的乘積，

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

由於 $0 \le j \le r-1 \le n-1$，每個因子都在 $(0,1]$ 之內。因此這個乘積不超過 1，而展開式中每一項都不超過 $\frac{1}{r!}$。所以

$$\left(1+\frac{1}{n}\right)^n \le \sum_{r=0}^{n}\frac{1}{r!} < \sum_{r=0}^{\infty}\frac{1}{r!} = \mathrm{e},$$

其中最後一個不等式是嚴格的，因為被略去的各項 $\frac{1}{(n+1)!}, \frac{1}{(n+2)!}, \ldots$ 全部為正數。這對每一個正整數 n 都成立。

上界 $\mathrm{P}(n) < \mathrm{e}$

把乘積寫成 $\mathrm{P}(n) = a_1 a_2 \cdots a_n$，其中 $a_k = \dfrac{2^k+1}{2^k} = 1 + \dfrac{1}{2^k}$，它們全部為正數，所以 AM–GM 不等式適用。算術平均（arithmetic mean）是

$$\frac{a_1 + \cdots + a_n}{n} = \frac{1}{n}\left(n + \frac{1}{2} + \frac{1}{4} + \cdots + \frac{1}{2^n}\right) = 1 + \frac{1-2^{-n}}{n},$$

這裏用了等比級數（geometric series）求和 $\frac{1}{2} + \cdots + \frac{1}{2^n} = 1 - 2^{-n}$。由於 $1 - 2^{-n} < 1$，這個平均值嚴格小於 $1 + \frac{1}{n}$。於是 AM–GM 不等式給出

$$\mathrm{P}(n)^{1/n} \le 1 + \frac{1-2^{-n}}{n} < 1 + \frac{1}{n}.$$

兩邊都是正數，所以取 n 次方後不等式仍然成立：

$$\mathrm{P}(n) < \left(1 + \frac{1}{n}\right)^n < \mathrm{e},$$

其中第二個不等式用了第一部分的結果。這對每一個正整數 n 都成立。

> **想一想** 嚴格不等式是從哪裏來的？AM–GM 不等式只給出 $\le$。

嚴格性來自 $1 - 2^{-n} < 1$，而不是來自 AM–GM 不等式。事實上，當 $n \ge 2$ 時，各因子 a_k 互不相同，所以 AM–GM 不等式本身也是嚴格的。但當 $n = 1$ 時只有一個因子，AM–GM 不等式成為等式（$\frac{3}{2} = \frac{3}{2}$）。而 $1 + \frac{1-2^{-n}}{n} < 1 + \frac{1}{n}$ 這一步在任何情況下都是嚴格的，所以這個論證涵蓋所有 n。

極限

每個因子 $a_{n+1} = 1 + 2^{-(n+1)}$ 都大於 1，所以 $\mathrm{P}(n+1) = a_{n+1}\mathrm{P}(n) > \mathrm{P}(n)$。因此數列 $\mathrm{P}(n)$ 嚴格遞增（strictly increasing），並且有上界（bounded above）e。有上界的遞增數列必定收斂（converge），所以 $\mathrm{P}(n)$ 趨向某個極限（limit）L。

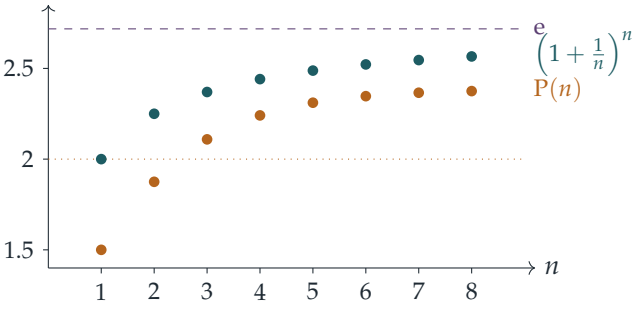
先看上界：每一個 $\mathrm{P}(n) < \mathrm{e}$，所以極限滿足 $L \le \mathrm{e}$。（取極限只能保留非嚴格不等式，這就是為甚麼我們得到的是 $\le$ 而不是 $<$。）

再看下界，先計算幾項：

$$\mathrm{P}(1) = \frac{3}{2}, \qquad \mathrm{P}(2) = \frac{3}{2}\cdot\frac{5}{4} = \frac{15}{8}, \qquad \mathrm{P}(3) = \frac{15}{8}\cdot\frac{9}{8} = \frac{135}{64} > \frac{128}{64} = 2.$$

由於數列遞增，$L \ge \mathrm{P}(4) > \mathrm{P}(3) > 2$。綜合起來，

$$2 < L \le \mathrm{e}.$$



驗算

- $n=1$：$(1+1)^1=2<\mathrm{e}$，而 $\mathrm{P}(1)=1.5<2$。當 $n=4$：$\mathrm{P}(4)=\frac{2295}{1024}\approx 2.241<\left(\frac{5}{4}\right)^4\approx 2.441$。
- 以數值方法計算這個乘積，$\mathrm{P}(n)$ 穩定在 $L\approx 2.3842$（到 $n=60$ 時，它在小數點後很多個位都已不再變動），這個數值穩妥地落在 2 與 $\mathrm{e}\approx 2.7183$ 之間。

常見錯誤

- 僅僅因為每個 $\mathrm{P}(n)<\mathrm{e}$ 就聲稱 $L<\mathrm{e}$。嚴格不等式在取極限後可能變成等式；能夠推出的只有 $L\le \mathrm{e}$。（這裏 $L<\mathrm{e}$ 其實是對的，但需要另作論證。）
- 證明了 $\mathrm{P}(n)>2$ 對某個 n 成立，卻忘記說明數列是遞增的，而正是這一點才把這個界傳給 L。
- 使用 AM–GM 不等式時沒有檢查各 a_k 是否為正數。

STEP 之外

「遞增且有上界，所以收斂」這一步就是單調收斂定理（monotone convergence theorem），它與實數的完備性（completeness）等價。從基本原理出發證明這個定理，是大學分析課程最先做的事情之一。取對數後，$\mathrm{P}(n)$ 變成級數 $\sum \ln\left(1+2^{-k}\right)$，把它與 $\sum 2^{-k}$ 比較便可知它收斂。這是研究無窮乘積（infinite product）的標準方法。Abbott 的 *Understanding Analysis*（一本分析學入門書）把這兩個想法都講得十分精彩。

詳細解答

STEP III 2000，第 8 題

純數學 • 遞推關係

> **STEP III 2000，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/00-s3-q8

關鍵思路

定義數列的規則是非線性的，而且涉及除法，直接處理並不方便。竅門是把它改寫成乘積的形式：$a_n a_{n-2} - a_{n-1}^2 = 1$。這是說，連續三項的某個組合恆等於 1。把同一事實往後推一步再寫一次，然後將兩個式子相減，非線性的規則便化為線性的規則。遞推關係（recurrence）一旦變成線性，求通項公式（closed form）就只需用上二階遞推關係的標準方法：先求輔助二次方程（auxiliary quadratic）的根，再配合兩個初始值。

解答

預備工作：數列有明確定義

每一步都要除以 a_{n-2}，所以我們應先確定沒有任何一項是零。事實上每一項都是正數。我們有 $a_0 = a_1 = 1 > 0$。若 a_{n-1} 和 a_{n-2} 均為正數，則 $1 + a_{n-1}^2 > 0$ 且 $a_{n-2} > 0$，所以 $a_n > 0$。根據歸納法（induction），對所有 n 均有 $a_n > 0$。最初幾項是

$$a_0 = 1, \quad a_1 = 1, \quad a_2 = \frac{1+1}{1} = 2, \quad a_3 = \frac{1+4}{1} = 5, \quad a_4 = \frac{1+25}{2} = 13.$$

線性遞推關係

把定義式乘以 a_{n-2}，可得對每個 $n \geq 2$，

$$a_n a_{n-2} = 1 + a_{n-1}^2. \tag{$*$}$$

設 $P(k)$ 為命題 $a_k = 3a_{k-1} - a_{k-2}$。我們對 k 用歸納法，證明 $P(k)$ 對所有 $k \geq 2$ 成立。

基礎步驟。$a_2 = 2$，而 $3a_1 - a_0 = 3 - 1 = 2$，所以 $P(2)$ 成立。

歸納步驟。假設 $P(k)$ 對某個 $k \geq 2$ 成立，即

$$a_k + a_{k-2} = 3a_{k-1}.$$

我們需要證明 $a_{k+1} = 3a_k - a_{k-1}$。根據定義，$a_{k+1} = (1 + a_k^2)/a_{k-1}$。做法是利用 $n = k$ 時的 $(*)$，把分子中那個礙事的 1 換掉；該式說明 $1 = a_k a_{k-2} - a_{k-1}^2$：

$$\begin{aligned} a_{k+1} &= \frac{a_k a_{k-2} - a_{k-1}^2 + a_k^2}{a_{k-1}} \\ &= \frac{a_k\,(a_k + a_{k-2}) - a_{k-1}^2}{a_{k-1}} \\ &= \frac{a_k \cdot 3a_{k-1} - a_{k-1}^2}{a_{k-1}} && \text{根據} P(k) \\ &= 3a_k - a_{k-1} && \text{因為} a_{k-1} > 0. \end{aligned}$$

所以 $P(k+1)$ 成立。根據歸納法，對所有 $n \geq 2$ 均有 $a_n = 3a_{n-1} - a_{n-2}$。

> **想一想** 歸納假設（inductive hypothesis）在哪裏用上了？為甚麼可以放心使用 $n = k$ 時的 $(*)$？

歸納假設只用了一次，就是把 $a_k + a_{k-2}$ 換成 $3a_{k-1}$。這個括號之所以出現，全因我們把 1 換成了 $a_k a_{k-2} - a_{k-1}^2$。使用 $(*)$ 則完全不需要任何假設：它不過是把 a_k 的定義乘開，對每個 $k \geq 2$ 都成立。換句話說，$a_n a_{n-2} - a_{n-1}^2$ 這個量在每一步都相同，正是它把線性關係一步一步傳遞下去。

通項公式

記 $\beta = \alpha^{-1}$。由於 $\alpha = \frac{1+\sqrt{5}}{2}$，

$$\beta = \frac{2}{1+\sqrt{5}} = \frac{2(\sqrt{5}-1)}{5-1} = \frac{\sqrt{5}-1}{2}, \qquad \text{所以} \qquad \alpha + \beta = \sqrt{5}, \quad \alpha\beta = 1.$$

因此

$$\alpha^2 + \beta^2 = (\alpha+\beta)^2 - 2\alpha\beta = 5 - 2 = 3, \qquad \alpha^2\beta^2 = 1,$$

所以 α^2 和 β^2 就是我們這個遞推關係的輔助方程 $x^2 - 3x + 1 = 0$ 的兩個根。（具體來說，它們是 $\frac{3\pm\sqrt{5}}{2}$。）

現在對 $n \geq 1$ 定義

$$b_n = \frac{\alpha^{2n-1} + \beta^{2n-1}}{\sqrt{5}}.$$

我們只要驗證 b_n 與 a_n 的首兩項相同，並且滿足同一個遞推關係，便可證明 $b_n = a_n$。

它滿足遞推關係。若 x 是 $x^2 = 3x - 1$ 的根，則乘以 x^{n-2} 可得 $x^n = 3x^{n-1} - x^{n-2}$。由於 $\alpha^{2n-1} = \beta\,(\alpha^2)^n$（這裏用了 $\alpha\beta = 1$），取 $x = \alpha^2$ 並乘以 β，得

$$\alpha^{2n-1} = 3\alpha^{2n-3} - \alpha^{2n-5}.$$

同樣地，取 $x = \beta^2$ 並乘以 α，可得 $\beta^{2n-1} = 3\beta^{2n-3} - \beta^{2n-5}$。兩式相加再除以 $\sqrt{5}$：

$$b_n = 3b_{n-1} - b_{n-2} \qquad (n \geq 3).$$

初始值相符。

$$b_1 = \frac{\alpha + \beta}{\sqrt{5}} = \frac{\sqrt{5}}{\sqrt{5}} = 1 = a_1,$$
$$b_2 = \frac{\alpha^3 + \beta^3}{\sqrt{5}} = \frac{(\alpha+\beta)^3 - 3\alpha\beta(\alpha+\beta)}{\sqrt{5}} = \frac{5\sqrt{5} - 3\sqrt{5}}{\sqrt{5}} = 2 = a_2.$$

結論。(a_n) 和 (b_n) 都對 $n \geq 3$ 滿足 $u_n = 3u_{n-1} - u_{n-2}$，而且在 $n = 1$ 和 $n = 2$ 時相等。這類遞推關係由前兩項決定每一項，所以根據（強）歸納法（strong induction），對每個 $n \geq 1$ 均有 $a_n = b_n$。也就是說，

$$a_n = \frac{\alpha^{2n-1} + \alpha^{-(2n-1)}}{\sqrt{5}} \qquad (n \geq 1),$$

這正是所要證明的。

? **想一想** 為甚麼從 $n = 1$ 而不是 $n = 0$ 開始比較？

a_n 的線性遞推關係在 $n = 3$ 時涉及 a_2 和 a_1，所以要處理 $n \geq 1$ 的情況，只需核對 b_1 和 b_2 便已足夠。其實公式在 $n = 0$ 時同樣成立：$b_0 = (\alpha^{-1} + \alpha)/\sqrt{5} = 1 = a_0$。所以無論從哪裏開始，都沒有損失。

驗算

- 各項為 $1, 1, 2, 5, 13, 34, 89, 233, \ldots$，即斐波那契數（Fibonacci numbers）$F_1, F_3, F_5, \ldots$（其中 $a_0 = 1 = F_{-1}$）。每一項都等於前一項的三倍減去再前一項：$3 \cdot 13 - 5 = 34$。
- 直接驗證 $(*)$：$34 \cdot 5 - 13^2 = 170 - 169 = 1$。
- 數值上，$a_{10} = 4181$，而 $(\alpha^{19} + \alpha^{-19})/\sqrt{5} \approx 4181.000$。$\alpha^{-19}/\sqrt{5}$ 這一項極小，所以 a_n 就是最接近 $\alpha^{2n-1}/\sqrt{5}$ 的整數。
- 按原來的規則精確計算數列直至 $n = 39$，所得各項都是整數，每次都滿足線性遞推關係，並與通項公式吻合。

常見錯誤

- 試圖把 $a_k = 3a_{k-1} - a_{k-2}$ 代入 $(1 + a_k^2)/a_{k-1}$ 再展開，以證明 $a_{k+1} = 3a_k - a_{k-1}$。這樣只會弄得一團糟。那個 1 必須用前一步的定義來替換。
- 忘記定義中要除以 a_{n-2}。完整的答案應指出各項均為正數，因此這個除法以及最後約去 a_{k-1} 的一步都是合理的。
- 只驗證 b_n 滿足遞推關係，卻沒有核對兩個初始值。二階遞推關係需要兩個初始值才能確定。

STEP 之外

常數 $a_n a_{n-2} - a_{n-1}^2$ 是斐波那契數的卡西尼恆等式（Cassini's identity）$F_{m+1}F_{m-1} - F_m^2 = (-1)^m$ 的一個例子。這裏只出現奇數下標，所以符號永遠是 $+1$。更一般地，對任何二階線性遞推關係，$u_{n+1}u_{n-1} - u_n^2$ 每一步都乘以一個固定的因子（它是一個 2×2 行列式（determinant），而行列式具有可乘性）。像本題這樣的有理遞推關係，竟然出人意料地總是產生整數，這與「Somos 數列」（Somos sequences）和洛朗現象（Laurent phenomenon）的研究有關。至於斐波那契型數列和連分數（continued fractions）的數論（留意 α 就是黃金比（golden ratio)），Davenport 的 *The Higher Arithmetic*（《高等算術》）是一本十分值得細讀的好書。

詳細解答

STEP III 2000，第 9 題

力學 • 彈性斜碰撞

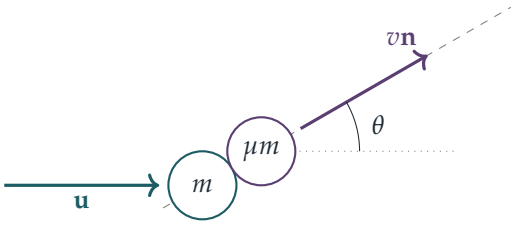
> **STEP III 2000，第 9 題**
>
> 繼續閱讀之前，請先看題目（原題為英文）。題目載於 STEP Support Programme 的試題資料庫：
>
> step.maths.org/questions/00-s3-q9

關鍵思路

兩個圓盤都是光滑的，所以碰撞期間它們之間唯一的力沿連心線（line of centres）作用。原本靜止的圓盤因此會沿這條線離開，所以 $\mathbf{n}$ 就是連心線的方向。把它的速度寫成 $v\mathbf{n}$，用動量守恆（conservation of momentum，這是一條向量方程）求出另一個圓盤的速度，再讓動能守恆（conservation of kinetic energy）定出 v。第 (ii) 部分只是多用一條能量方程，而不等式則來自餘弦值不能超過 1 這個要求。

解答

我們把兩個圓盤當作光滑水平面上的光滑質點（particle），所以圓盤之間以及圓盤與平面之間都沒有摩擦力。設 θ 為 $\mathbf{u}$ 與 $\mathbf{n}$ 之間的夾角，並記 $u = |\mathbf{u}|$，於是 $\mathbf{u} \cdot \mathbf{n} = u\cos\theta$。



在碰撞的一刻，連心線（虛線）指向 $\mathbf{n}$ 的方向，與 $\mathbf{u}$ 的方向（點線）成 θ 角。這幅圖只是示意圖，代數運算並不需要用到它。

第 (i) 部分

質量為 μm 的圓盤原本靜止，碰撞期間它所受的唯一水平力，是另一個圓盤施加的接觸力，而這個力沿連心線作用。由於它沿 $\mathbf{n}$ 的方向離開，它碰撞後的速度是 $v\mathbf{n}$，其中 $v > 0$ 是某個速率。

設質量為 m 的圓盤碰撞後的速度為 $\mathbf{w}$。由動量守恆得

$$m\mathbf{u} = m\mathbf{w} + \mu m v\,\mathbf{n} \qquad \Longrightarrow \qquad \mathbf{w} = \mathbf{u} - \mu v\,\mathbf{n}.$$

這次碰撞是完全彈性（perfectly elastic）的，所以動能守恆：

$$\tfrac{1}{2}m\,\mathbf{u}\cdot\mathbf{u} = \tfrac{1}{2}m\,\mathbf{w}\cdot\mathbf{w} + \tfrac{1}{2}\mu m v^2.$$

利用 $\mathbf{n}\cdot\mathbf{n} = 1$，

$$\mathbf{w}\cdot\mathbf{w} = (\mathbf{u} - \mu v\mathbf{n})\cdot(\mathbf{u} - \mu v\mathbf{n}) = u^2 - 2\mu v\,\mathbf{u}\cdot\mathbf{n} + \mu^2 v^2.$$

代入並消去 $\frac{1}{2}m$ 和 u^2，

$$0 = -2\mu v\,\mathbf{u}\cdot\mathbf{n} + \mu^2 v^2 + \mu v^2 = \mu v\big((1+\mu)v - 2\,\mathbf{u}\cdot\mathbf{n}\big).$$

現在 $\mu > 0$，而且 $v \neq 0$，因為被撞的圓盤確實動了。所以我們可以除以 μv，得到

$$v = \frac{2\,\mathbf{u}\cdot\mathbf{n}}{1+\mu},$$

這就是所求的速率。順帶一提，$v > 0$ 告訴我們 $\mathbf{u}\cdot\mathbf{n} > 0$，所以 $0 \le \theta < \frac{1}{2}\pi$。

> **想一想** 為甚麼這裏不需要用牛頓實驗定律（Newton's experimental law）？

其實也可以用：取 $e = 1$，相對速度沿 $\mathbf{n}$ 的分量會反轉，而垂直於 $\mathbf{n}$ 的分量則因為圓盤光滑而保持不變。但「完全彈性」的意思就是動能守恆，而這一條純量方程（scalar equation）連同動量守恆，已經足以確定 v。上面的做法與之完全等價，而且不必分解成分量。

第 (ii) 部分

碰撞後的總動能等於最初的 $\frac{1}{2}mu^2$。如果兩個圓盤平分這些動能，每個就有 $\frac{1}{4}mu^2$，所以對被撞的圓盤而言

$$\tfrac{1}{2}\mu m v^2 = \tfrac{1}{4}mu^2 \qquad \Longrightarrow \qquad \mu v^2 = \tfrac{1}{2}u^2.$$

利用第 (i) 部分的結果和 $\mathbf{u}\cdot\mathbf{n} = u\cos\theta$，我們有 $v = \dfrac{2u\cos\theta}{1+\mu}$，所以

$$\frac{4\mu u^2\cos^2\theta}{(1+\mu)^2} = \frac{u^2}{2} \qquad \Longrightarrow \qquad \cos^2\theta = \frac{(1+\mu)^2}{8\mu}.$$

由於 $\cos\theta > 0$（見第 (i) 部分末段），我們取正根：

$$\cos\theta = \frac{1+\mu}{2\sqrt{2\mu}}.$$

要這樣的角存在，我們需要 $\cos\theta \le 1$，即 $\cos^2\theta \le 1$：

$$\frac{(1+\mu)^2}{8\mu} \le 1 \iff (1+\mu)^2 \le 8\mu \qquad \text{（兩邊乘以}8\mu > 0\text{）}$$
$$\iff \mu^2 - 6\mu + 1 \le 0.$$

二次式 $\mu^2 - 6\mu + 1$ 的根是 $\mu = 3 \pm \sqrt{9-1} = 3 \pm \sqrt{8}$，而它恰好在兩根之間（包括兩根）為負數或零。因此

$$3 - \sqrt{8} \le \mu \le 3 + \sqrt{8},$$

這正是所要證明的結果。（下界是正數，$3 - \sqrt{8} \approx 0.172$，所以與 $\mu > 0$ 並無矛盾。）反過來說，對這個範圍內的每一個 μ，公式都給出一個在 $(0, 1]$ 內的有效餘弦值，所以平分動能確實可以做到；在範圍的兩端 $\theta = 0$，即正面碰撞（head-on collision）。

驗算

- 正面碰撞的情況：如果 **n** 與 **u** 平行，第 (i) 部分給出 $v = 2u/(1+\mu)$，這是大家熟悉的一維彈性碰撞結果。當 $\mu = 1$ 時就是 $v = u$：運動中的圓盤停下來，把全部運動傳給對方，就像桌球中的情形一樣。
- 質量相等，$\mu = 1$：第 (ii) 部分給出 $\cos\theta = 2/(2\sqrt{2}) = 1/\sqrt{2}$，所以 $\theta = 45°$。這符合一個經典事實：兩個質量相等的光滑圓盤作彈性碰撞後，會沿互相垂直的方向離開；動能相等又迫使兩者速率相等，所以它們各與 **u** 成 45° 角。
- 數值方面，我們模擬了 10 000 次隨機的光滑彈性碰撞（衝量沿 **n** 方向），確認了動量守恆、能量守恆和第 (i) 部分的速率，並確認了當 μ 介乎 0.172 與 5.83 之間時，在第 (ii) 部分求得的角度下兩者動能相等。舉例來說，$\mu = 0.1$ 給出 $\cos^2\theta = 1.51 > 1$，這是不可能的。

常見錯誤

- 把 **u** 和 **n** 當作平行。這是斜碰（oblique collision）；只有沿連心線的分量 $\mathbf{u}\cdot\mathbf{n}$ 會傳遞過去。
- 除以 v 卻沒有說明它為何不等於零（$v = 0$ 是能量方程中「沒有發生碰撞」的那個解）。
- 忘記排除 $\cos\theta$ 的負平方根。

STEP 之外

把被撞質點的速度寫成沿衝量（impulse）方向，然後運用動量和能量，這個技巧正是物理學中彈性散射（elastic scattering）的核心。在質心參考系（centre-of-mass frame）中，情況最為簡潔：彈性碰撞只是把兩個速度旋轉，而不改變它們的大小；第 (ii) 部分的角度範圍，於是變成一個關於實驗室參考系（lab frame）中哪些角度可以達到的命題。Kleppner 與 Kolenkow 合著的 *An Introduction to Mechanics*（力學入門教材）把兩個參考系中的碰撞都講解得非常清楚。

詳細解答

STEP III 2000，第 10 題

力學 • 極限摩擦力平衡

STEP III 2000，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/00-s3-q10

關鍵思路

先從幾何入手：題目選用 $\sqrt{3}a$ 和 a 這兩個距離，是為了令樑與地面成一個特殊角，而這個角決定了題中所有方向。然後對兩個能令未知力消失的點取力矩（moment）。對鉸鏈（hinge）X 取力矩時，樑的方程只含球體給予的法向反作用力（normal reaction），因為球體給予的摩擦力（friction）正好沿著樑本身的方向。對球心取力矩時，除了兩個摩擦力之外，其餘各力都通過球心，所以兩個摩擦力必定相等。

解答

模型與幾何

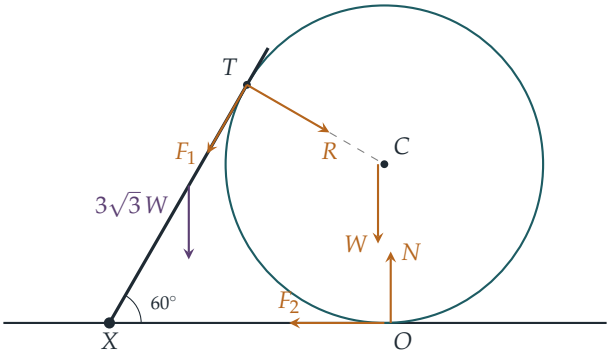
我們把樑視為均勻（uniform）桿，把球體視為均勻剛體（rigid body），其重量作用於球心 C。兩個接觸點都是粗糙的。設 O 為球體與地面的接觸點，則 $OX = \sqrt{3}a$，$OC = a$，而 OC 是鉛垂的。

樑與球體在點 T 接觸，並在該處與球體相切。地面則是由 X 所作的另一條切線（tangent）。在直角三角形 XOC 中，

$$\tan\angle OXC = \frac{a}{\sqrt{3}a} = \frac{1}{\sqrt{3}}, \qquad \text{所以} \qquad \angle OXC = 30^\circ.$$

由 X 所作的兩條切線關於直線 XC 對稱，所以樑與地面的夾角是 $2 \times 30^\circ = 60^\circ$。由同一點所作的切線長度相等，所以 $XT = XO = \sqrt{3}a$。由於 $\sqrt{3}a < 2a$，接觸點確實在樑上，位置稍稍超過樑的中點。

半徑 CT 垂直於樑。因此，在 T 的法向反作用力與水平方向成 30°，而在 T 的摩擦力沿著樑的方向，與水平方向成 60°。



琥珀色的箭頭表示作用於球體的力。樑所受的 R 和 F_1 方向與此相反，此外還有作用於樑中點的重量 $3\sqrt{3}\,W$（與 X 相距 a），以及鉸鏈處一個未知的反作用力。

摩擦力的方向是我們猜測的：對球體而言，F_1 沿著樑向下指向 X，而 F_2 沿著地面指向 X。假如猜錯了，該力算出來便會是負數。

樑：對 X 取力矩

鉸鏈的反作用力作用於 X，所以它對 X 沒有力矩。球體給予的摩擦力沿著樑的方向，而樑所在的直線通過 X，所以它也沒有力矩。法向反作用力 R 垂直於樑，作用點的距離為 $XT = \sqrt{3}\,a$。重量的作用線與 X 的水平距離為 $a\cos 60° = \frac{1}{2}a$。因此

$$R \cdot \sqrt{3}\,a = 3\sqrt{3}\,W \cdot \tfrac{1}{2}a, \qquad \text{所以} \qquad R = \tfrac{3}{2}W.$$

球體：對 C 取力矩

R、W 和 N 這三個力都通過 C。（法向反作用力沿半徑方向作用。）兩個摩擦力都沿切線方向，各與 C 相距 a。在 T 的 F_1 使球體繞 C 向一個方向轉動，而在 O 的 F_2 則使它向相反方向轉動，所以

$$F_1 a = F_2 a, \qquad \text{所以} \qquad F_1 = F_2.$$

把這個共同的值記為 F。

> ? **想一想** 在對兩個摩擦力還一無所知的時候，為甚麼我們已經可以肯定它們相等？

球體處於平衡（equilibrium）狀態，所以對任何一點的總力矩都是零。選球心最為巧妙：重量和兩個法向反作用力都通過球心，於是只剩下兩個摩擦力，而它們的力臂（lever arm）同樣是 a。

球體：分解力

水平方向（resolving，即把力分解），取遠離 X 的方向為正：R 的分量是 $R\cos 30^\circ$，F_1 的分量是 $-F\cos 60^\circ$，而 F_2 的貢獻是 $-F$。所以

$$\frac{\sqrt{3}}{2}R = \frac{1}{2}F + F = \frac{3}{2}F, \qquad \text{所以} \qquad F = \frac{R}{\sqrt{3}} = \frac{\sqrt{3}}{2}W.$$

這個值是正數，可見先前猜測的摩擦力方向是對的。鉛垂方向，R 向下推，分量為 $R\sin 30^\circ$；F_1 向下拉，分量為 $F\sin 60^\circ$。所以

$$N = W + \tfrac{1}{2}R + \tfrac{\sqrt{3}}{2}F = W + \tfrac{3}{4}W + \tfrac{3}{4}W = \tfrac{5}{2}W.$$

摩擦係數

球體在兩個接觸點都即將滑動，所以兩處的摩擦都達到極限（limiting friction）：$F_1 = \mu_1 R$ 及 $F_2 = \mu_2 N$。因此，兩個摩擦係數（coefficient of friction）為

$$\mu_1 = \frac{F}{R} = \frac{\frac{\sqrt{3}}{2}W}{\frac{3}{2}W} = \frac{1}{\sqrt{3}}, \qquad \mu_2 = \frac{F}{N} = \frac{\frac{\sqrt{3}}{2}W}{\frac{5}{2}W} = \frac{\sqrt{3}}{5}.$$

驗算

- 用坐標作數值驗算（$a = 1.3$，$W = 2$），確認了 T 在球面上，而且 CT 垂直於樑。直接解球體的三條平衡方程，得出 $R = 1.5W$、$F_1 = F_2 \approx 0.8660W$ 和 $N = 2.5W$，所以 $\mu_1 \approx 0.5774 = 1/\sqrt{3}$，$\mu_2 \approx 0.3464 = \sqrt{3}/5$。
- 一條獨立的方程：對球體以 O 為支點取力矩。這時 N、F_2 和 W 都沒有力矩。R 的作用線通過 C，所以它的力矩等於其水平分量在高度 a 處的力矩，即 $\frac{\sqrt{3}}{2}R\,a$。F_1 對 O 的力矩算出來是 $\frac{3}{2}aF$。令兩者相等，便再次得到 $F = R/\sqrt{3}$，而無須用到對 C 的力矩方程。
- 兩個係數都小於 1，對一般的粗糙接觸面來說，這在物理上是合理的。

常見錯誤

- 把樑的傾角當作 30°（即 XC 的傾角），而不是 60°。
- 求 N 時忘記了 F_1 的鉛垂分量。這樣會得出 $N = \frac{7}{4}W$，以及錯誤的 μ_2。
- 在樑對 X 的力矩方程中計入球體給予的摩擦力。這個力沿著樑的方向，所以它對 X 的力矩是零。

STEP 之外

平面上的剛體有三個自由度（degrees of freedom），所以可給出三條獨立的平衡方程。本題中，球體有五個未知的接觸力，而樑提供了 R，減少了未知數的數目。再加上兩個「極限摩擦」條件，問題便成為一個可解的方程組。像這樣把方程的數目與未知數的數目互相比較，並辨認出一個系統何時屬於靜不定（statically indeterminate），正是結構力學（structural mechanics）的起點。Kleppner 與 Kolenkow 所著的 *An Introduction to Mechanics*（力學入門教科書）對剛體靜力學和力矩（torque）有細緻的論述。

詳細解答

STEP III 2000，第 11 題

力學 • 繞粗糙支點的轉動

STEP III 2000，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/00-s3-q11

關鍵思路

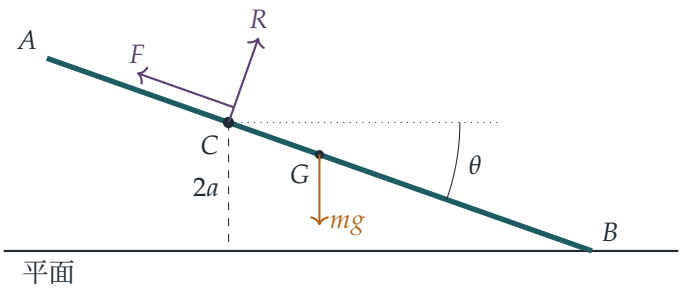
只要沒有滑動，棒就只是繞着橫樑上的點 C 轉動，因此它是一個繞固定軸轉動的剛體（rigid body）。由能量可求得角速率（angular speed），由力矩方程可求得角加速度（angular acceleration）。然後對質心（centre of mass）應用牛頓第二定律（Newton's second law），分別沿棒的方向和垂直於棒的方向分解，便可知道橫樑必須提供的摩擦力（friction）和法向反作用力（normal reaction）。只要摩擦力不超過法向反作用力的 μ 倍，棒就不會滑動，而我們需要這個條件一直成立，直至 B 到達平面為止。

解答

模型與幾何

我們把棒視為均勻剛棒，把橫樑視為一條固定而粗糙的細線，因此橫樑只在 C 處施加單一個接觸力：垂直於棒的法向反作用力 R 和沿棒方向的摩擦力 F。

質心 G 是棒的中點，與 A 相距 $4.5a$。由於 $AC = 3a$，G 位於 C 靠 B 的一側，且 $CG = \frac{3}{2}a$，$CB = 6a$。所以重力對 C 有力矩（moment），棒會傾側，B 端向下。設 θ 為棒轉過的角，即棒與水平線所成的角為 θ。



當 B 下降了 $2a$ 時便到達平面，即當 $6a\sin\theta = 2a$ 時，亦即

$$\sin\theta = \tfrac{1}{3}, \qquad \tan\theta = \frac{1}{2\sqrt{2}}.$$

繞 C 的轉動

根據平行軸定理（parallel axes theorem），對 C 的轉動慣量（moment of inertia）為

$$I_C = \tfrac{1}{12}m(9a)^2 + m\left(\tfrac{3}{2}a\right)^2 = \tfrac{27}{4}ma^2 + \tfrac{9}{4}ma^2 = 9ma^2.$$

在沒有滑動的期間，C 固定不動，接觸力不作功。G 下降了 $\frac{3}{2}a\sin\theta$，所以由能量守恆（conservation of energy）（從靜止開始）得

$$\tfrac{1}{2}(9ma^2)\dot\theta^2 = mg\cdot\tfrac{3}{2}a\sin\theta \quad\Longrightarrow\quad \dot\theta^2 = \frac{g\sin\theta}{3a}.$$

對 C 取力矩（只有重力有力矩）：

$$9ma^2\,\ddot\theta = mg\cdot\tfrac{3}{2}a\cos\theta \quad\Longrightarrow\quad \ddot\theta = \frac{g\cos\theta}{6a}.$$

（把能量方程對 t 求導，會得到相同的結果，這是一個有用的驗算。）

接觸力

G 在以 C 為圓心、半徑為 $\frac{3}{2}a$ 的圓上運動，所以它的加速度有一個沿棒指向 C 的分量 $\frac{3}{2}a\dot\theta^2$，以及一個垂直於棒的分量 $\frac{3}{2}a\ddot\theta$，方向為 θ 增加的方向（即垂直於棒向下）。

> ? **想一想** 摩擦力的方向是哪一邊？為甚麼它不單單是 $mg\sin\theta$？

重力沿棒向下的分量 $mg\sin\theta$ 會把棒朝 B 的方向拉離橫樑；此外，G 還需要一個指向 C 的向心力（centripetal force）。兩者都必須由摩擦力提供，因此摩擦力沿棒向上，指向 A。沿棒的方向分解，並以指向 C 的方向作為 G 的加速度的正方向：

$$F - mg\sin\theta = m\cdot\tfrac{3}{2}a\dot\theta^2 = m\cdot\tfrac{3}{2}a\cdot\frac{g\sin\theta}{3a} = \tfrac{1}{2}mg\sin\theta,$$

所以

$$F = \tfrac{3}{2} mg \sin\theta.$$

垂直於棒的方向分解，取 G 加速的方向：

$$mg\cos\theta - R = m \cdot \tfrac{3}{2} a\ddot{\theta} = m \cdot \tfrac{3}{2} a \cdot \frac{g\cos\theta}{6a} = \tfrac{1}{4} mg\cos\theta,$$

所以

$$R = \tfrac{3}{4} mg\cos\theta,$$

這在整個過程中都是正數，所以棒一直與橫樑保持接觸。

不滑動的條件

只要 $F \le \mu R$，棒就不會在 C 處滑動，即

$$\frac{F}{R} = \frac{\frac{3}{2} mg\sin\theta}{\frac{3}{4} mg\cos\theta} = 2\tan\theta \le \mu.$$

比值 $2\tan\theta$ 隨 θ 遞增，在釋放時由 0 開始，所以這個條件在最後一刻最難滿足，亦即 B 剛到達平面之時。該處 $\tan\theta = 1/(2\sqrt{2})$，所以我們需要

$$\mu \ge 2 \cdot \frac{1}{2\sqrt{2}} = \frac{1}{\sqrt{2}}.$$

因此 μ 的最小值為

$$\boxed{\mu = \frac{1}{\sqrt{2}} = \frac{\sqrt{2}}{2} \approx 0.707.}$$

若恰好取這個值，摩擦力只在 B 撞到平面的一瞬間才達到極限；若 μ 再小一點，棒便會在此之前開始滑動。

驗算

- 釋放時 $\dot{\theta} = 0$，所以 $F = 0$ 而 $R = \frac{3}{4} mg$：橫樑所承受的力小於整個重量，因為 G 開始以 $\frac{3}{2} a\ddot{\theta} = \frac{1}{4} g$ 向下加速。這就是常見的「放手後的木板感覺變輕」效應。
- 單位：F 和 R 都是 mg 的倍數，而 μ 是無量綱（dimensionless）的，理應如此。
- 我們由靜止開始，用數值方法逐步求解方程 $\ddot{\theta} = g\cos\theta/(6a)$，以有限差分（finite differences）計算 G 的加速度，再由牛頓第二定律還原出 F 和 R。結果與 $\frac{3}{2} mg\sin\theta$ 和 $\frac{3}{4} mg\cos\theta$ 吻合，而 F/R 穩步上升，在 $\sin\theta = \frac{1}{3}$ 時達到 $0.7071\ldots = 1/\sqrt{2}$。

常見錯誤

- 對棒的中心取力矩，或忘記平行軸定理中的附加項：$I_C = 9ma^2$，而不是 $\frac{27}{4}ma^2$。
- 把問題當作靜力學（statics）來處理，取 $F = mg\sin\theta$、$R = mg\cos\theta$，從而得出 $\mu \geq \tan\theta$。其實 G 的向心加速度和切向加速度都會改變這些力。
- 只在釋放的一刻檢查摩擦條件（那時 $F = 0$），而沒有在條件最嚴苛的位置檢查。

STEP 之外

這道題讓我們初嘗帶有約束力（constraint forces）的剛體動力學：我們先由能量和力矩求出運動，然後才對質心應用牛頓定律，倒推出約束必須提供的力。同樣的模式（先解出運動，再檢查約束是否真的能提供所需的力）可以用來判斷梯子何時開始滑動、傾倒的鉛筆何時離開桌面，或者珠子何時離開鐵線。Kleppner 與 Kolenkow 合著的 *An Introduction to Mechanics*（力學入門）仔細講解了定軸轉動，以及力矩與質心運動之間的關係。

詳細解答

STEP III 2000，第 12 題

概率與統計 • 泊松近似

> **STEP III 2000，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/00-s3-q12

關鍵思路

一旦知道自己的彩票中獎，餘下的隨機性就只在其他玩家的彩票上。他們每人各自獨立地以概率 $\frac{1}{N}$ 選中中獎號碼，所以其他中獎者的人數服從二項分佈（binomial distribution）。當玩家眾多而成功概率很小時，這個二項分佈接近泊松分佈（Poisson distribution），而題目給出的極限 $\left(1-\frac{a}{n}\right)^n \approx \mathrm{e}^{-a}$ 正是令這個近似成立的關鍵。

解答

精確概率

假設你持有一張中獎彩票。設 X 為其餘 $2N-1$ 位玩家中同樣持有中獎號碼的人數，則中獎者總數為 $W = 1 + X$。

機器為每位玩家選號時，是均勻（uniform）地抽選，與其他人的號碼互相獨立（independent），亦與開獎結果獨立。因此，給定中獎號碼（不論是哪一個）及你的彩票後，其餘 $2N-1$ 位玩家每人以概率 $\frac{1}{N}$ 選中該號碼，且彼此獨立。所以

$$X \sim \mathrm{B}\left(2N-1,\ \tfrac{1}{N}\right).$$

中獎者不多於三人即 $X \le 2$，故精確概率為

$$\mathrm{P}(W \le 3) = \sum_{k=0}^{2} \binom{2N-1}{k} \left(\frac{1}{N}\right)^k \left(1-\frac{1}{N}\right)^{2N-1-k}.$$

完整寫出來就是

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

近似值

對每個固定的 $k \in \{0,1,2\}$，把冪拆開：

$$\left(1-\frac{1}{N}\right)^{2N-1-k} = \left[\left(1-\frac{1}{N}\right)^{N}\right]^{2}\left(1-\frac{1}{N}\right)^{-(1+k)}.$$

在題目給出的結果中取 $a=1$ 及 $n=N$，方括號內的式子約為 e^{-1}。最後一個因子是一個非常接近 1 的數的固定小次冪，所以約為 1。因此每個冪都約為 e^{-2}。

至於二項式係數（binomial coefficient），k 最大為 2 而 N 很大，所以

$$\binom{2N-1}{k}\frac{1}{N^k} = \frac{(2N-1)(2N-2)\cdots(2N-k)}{k!\,N^k} \approx \frac{(2N)^k}{k!\,N^k} = \frac{2^k}{k!}.$$

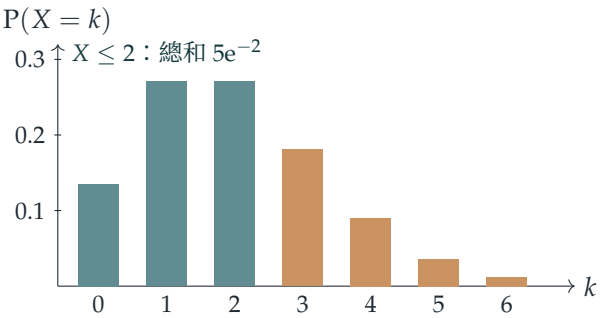
綜合以上各項，$k=0,1,2$ 的項分別約為 e^{-2}、$2\mathrm{e}^{-2}$ 和 $2\mathrm{e}^{-2}$，所以

$$\mathrm{P}(W \le 3) \approx \mathrm{e}^{-2}\left(1+2+\frac{2^2}{2!}\right) = 5\mathrm{e}^{-2} \approx 0.677 \approx \frac{2}{3},$$

其中用了 $\mathrm{e}^2 \approx 7.39$。這就是所求的結果。

> **想一想** 為甚麼 X 的表現有如一個平均值為 2 的泊松變量？

上面的計算就是二項分佈的標準泊松極限。試驗次數 $2N-1$ 很大，成功概率 $\frac{1}{N}$ 很小，而兩者之積始終接近 2。在這種情況下，$\binom{n}{k}p^k(1-p)^{n-k} \to \mathrm{e}^{-\lambda}\lambda^k/k!$，其中 $\lambda = np$，而這裏 $\lambda \to 2$。



X 的極限分佈是平均值為 2 的泊松分佈。藍綠色的長條（$X \le 2$，即總共最多三位中獎者）約佔總和的三分之二。

玩家自行選號

如果由玩家自行選號，號碼就不再均勻分佈：人們偏愛生日（不超過 31 的號碼）、幸運號碼和有規律的號碼。假設每位玩家以概率 p_i 選擇號碼 i，其中 $\sum p_i = 1$，而中獎號碼

仍是均勻抽出。那麼，在你中獎的條件下，你的號碼是 i 的概率與 $p_i \cdot \frac{1}{N}$ 成正比，即概率為 p_i。所以熱門號碼在中獎彩票中所佔的比例偏高。若你的號碼是 i，其他中獎者的人數服從 $\mathrm{B}(2N-1, p_i)$，平均值為 $(2N-1)p_i$。對 i 取平均，其他中獎者人數的期望值（expectation）為

$$(2N-1)\sum_i p_i^2 \;\geq\; (2N-1)\cdot\frac{1}{N},$$

這是根據柯西－施瓦茲不等式（Cauchy–Schwarz inequality）$(\sum p_i)^2 \leq N\sum p_i^2$ 得出的，等號僅在每個 $p_i = \frac{1}{N}$ 時成立。所以選號不均勻時，平均而言會有更多人與你分享獎金，而中獎者不多於三人的概率通常會下降。（個別玩家若刻意選擇冷門號碼，表現會比平均好，但玩家整體而言則較差。）

中獎者人數的期望值

由於 $X \sim \mathrm{B}\left(2N-1, \frac{1}{N}\right)$，其平均值為 $np = (2N-1)\cdot\frac{1}{N}$。（等價地，可把 X 寫成 $2N-1$ 個指示變量（indicator variable）之和，每位其他玩家對應一個，各自的平均值為 $\frac{1}{N}$。）因此

$$\mathrm{E}(W) = 1 + \mathrm{E}(X) = 1 + \frac{2N-1}{N} = 3 - \frac{1}{N} = 3 - N^{-1},$$

正如所求。注意這是精確結果，並非近似值。

驗算

- 當 $N = 2$ 和 $N = 3$ 時，列舉 $2N$ 位玩家所有可能的號碼分配及所有可能的中獎號碼，所得結果與二項式表達式完全吻合，期望值亦恰為 $3 - \frac{1}{N}$。
- 精確概率在 $N = 10$ 時為 0.7054，$N = 100$ 時為 0.6794，$N = 10\,000$ 時為 0.6767，趨近 $5\mathrm{e}^{-2} = 0.67668$。
- 以 $N = 50$ 進行模擬（10^5 個星期），得出 0.683，而精確值為 0.682；中獎者人數的平均值為 2.981，而 $3 - \frac{1}{50} = 2.98$。
- 當 $N = 10$ 而其中一半號碼的受歡迎程度是另一半的兩倍時，概率由 0.705 降至 0.641，中獎者人數的期望值則由 2.9 升至 3.11。

常見錯誤

- 用了 $2N$ 次試驗而非 $2N-1$ 次：你自己的彩票已知中獎，所以只有其他玩家是隨機的。若用 $2N$ 次試驗，期望值會算成 $1+2=3$，而不是 $3 - \frac{1}{N}$。
- 用了無條件平均值：就所有星期而言，中獎者人數的期望值是 $2N\cdot\frac{1}{N} = 2$，但知道你中獎後，便多了一位必然的中獎者，餘下 $2N-1$ 位其他玩家。

► 遺漏了 $k = 2$ 的項，或把它寫成 2^2e^{-2} 而漏掉 $\frac{1}{2!}$。

STEP 之外

這裏的近似就是稀有事件定律（law of rare events）：試驗次數多而成功概率小的二項分佈會收斂（converge）至泊松分佈。最後一部分觸及檢查悖論（inspection paradox），又稱大小偏差抽樣（size-biased sampling）。以你的彩票中獎為條件，會令多人選擇的號碼更有可能出現，這正是條件平均值是 3 而不是 2 的原因。這兩個概念在 Grimmett 與 Welsh 的 *Probability: An Introduction*（概率入門教材）以及 Blitzstein 與 Hwang 的 *Introduction to Probability*（概率導論教材）中都有很好的論述。

詳細解答

STEP III 2000，第 13 題

概率與統計 • 概率生成函數

> **STEP III 2000，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/00-s3-q13

關鍵思路

處理這兩個隨機變量（random variable），最方便的做法是考慮「尾部」或「累積」事件。第一個六點在第 r 次擲骰之後才出現，當且僅當每一顆骰子在首 r 次擲骰中全部落空；最後一個六點在第 r 次擲骰或之前已經出現，當且僅當每一顆骰子都已經擲出過六點。這兩個事件各自都是若干獨立事件（independent events）的交集，所以其概率是一個乘積，而把相鄰兩個值相減便得出分佈。一旦認出「第一個六點」這個變量服從幾何分佈（geometric distribution），與它有關的一切都可以很快求得。

解答

以下一律假設各顆骰子之間互相獨立，而每次擲骰之間亦互相獨立。對單獨一顆骰子，設 X_i 為第 i 顆骰子首次擲出六點時的擲骰次數，則

$$\mathrm{P}(X_i > r) = q^r \qquad (r = 0, 1, 2, \dots),$$

因為 $X_i > r$ 表示連續 r 次都擲不出六點。於是 $R = \min_i X_i$，$S = \max_i X_i$。

第一個六點

事件 $R > r$ 表示 n 顆骰子在首 r 次擲骰中都沒有擲出六點，即全部 nr 次個別投擲的結果都不是六點。由獨立性，

$$\mathrm{P}(R > r) = q^{nr}.$$

因此，對 $r \geq 1$，

$$\mathrm{P}(R = r) = \mathrm{P}(R > r-1) - \mathrm{P}(R > r) = q^{n(r-1)} - q^{nr} = q^{nr}\left(q^{-n} - 1\right),$$

這正是所求的結果。

記 $Q = q^n$，上式即 $\mathrm{P}(R = r) = Q^{r-1}(1-Q)$：$R$ 服從幾何分佈，每次把整組骰子擲一次，「成功」（至少出現一個六點）的概率為 $1 - Q$。所以它的概率生成函數（probability generating function，簡稱 PGF）是一個等比級數（geometric series）：

$$\begin{aligned} G_R(t) = \sum_{r=1}^{\infty} t^r Q^{r-1}(1-Q) &= (1-Q)\,t \sum_{r=1}^{\infty} (Qt)^{r-1} \\ &= \frac{(1-q^n)\,t}{1-q^n t}, \end{aligned}$$

對 $|t| < q^{-n}$ 成立（特別是在 $t = 1$ 附近）。

要求平均值，用商法則（quotient rule）求導：

$$G_R'(t) = (1-q^n)\,\frac{(1-q^n t) + q^n t}{(1-q^n t)^2} = \frac{1-q^n}{(1-q^n t)^2},$$

所以

$$\mathrm{E}(R) = G_R'(1) = \frac{1}{1-q^n}.$$

當 $n = 2$ 且 $p = \frac{1}{6}$ 時，$q^2 = \frac{25}{36}$，故 $\mathrm{E}(R) = 1 \big/ \frac{11}{36} = \frac{36}{11}$，正如所求。

最後一個六點

事件 $S \le r$ 表示每一顆骰子都在首 r 次擲骰中的某一次擲出過六點。對一顆骰子而言，這個概率是 $1 - q^r$，而各顆骰子互相獨立，所以

$$\mathrm{P}(S \le r) = (1-q^r)^n.$$

把相鄰兩個值相減，對 $r \ge 1$，

$$\mathrm{P}(S = r) = \mathrm{P}(S \le r) - \mathrm{P}(S \le r-1) = \left(1-q^r\right)^n - \left(1-q^{r-1}\right)^n,$$

正如所求。（當 $r = 1$ 時，第二項是 $0^n = 0$，結果正確。）

> **想一想** 要對 $t^r\left[(1-q^r)^2 - (1-q^{r-1})^2\right]$ 求和，看來相當繁複。怎樣才可以把它化成等比級數？

當 $n = 2$ 時，把平方展開：

$$\begin{aligned} \mathrm{P}(S = r) &= \left(1 - 2q^r + q^{2r}\right) - \left(1 - 2q^{r-1} + q^{2r-2}\right) \\ &= 2q^{r-1}(1-q) - q^{2(r-1)}(1-q^2) \\ &= 2p\,q^{r-1} - (1-q^2)\,(q^2)^{r-1}. \end{aligned}$$

每一項都是一個幾何分佈（的倍數），所以

$$G_S(t) = \sum_{r=1}^{\infty} t^r \,\mathrm{P}(S = r) = \frac{2pt}{1-qt} - \frac{(1-q^2)\,t}{1-q^2t}.$$

這個結果合乎情理：由於 $S = X_1 + X_2 - R$，它的分佈就是「參數為 p 的幾何分佈的兩倍，減去 R 的分佈」。要化簡，可把各項通分，以 $(1-qt)(1-q^2t)$ 為公分母，並利用 $1-q^2 = p(1+q)$：

$$\begin{aligned} 2p(1-q^2t) - p(1+q)(1-qt) &= p\big[\,2 - 2q^2t - 1 - q + qt + q^2t\,\big] \\ &= p\big[(1-q) + qt(1-q)\big] = p^2(1+qt). \end{aligned}$$

因此

$$G_S(t) = \frac{p^2\,t\,(1+qt)}{(1-qt)(1-q^2t)}.$$

作為簡單的核對，$G_S(1) = \dfrac{p^2(1+q)}{p\cdot p(1+q)} = 1$。

要求平均值，以拆開的形式求導最為容易，因為每一部分都是一個幾何分佈的 PGF，它在 1 處的導數就是其平均值乘以相應的係數。對 $\dfrac{(1-a)t}{1-at}$ 而言，在 $t = 1$ 處的導數是 $\dfrac{1}{1-a}$（與上面 G_R 的情況完全一樣），所以

$$\mathrm{E}(S) = G_S'(1) = 2\cdot\frac{1}{p} - \frac{1}{1-q^2} = \frac{2(1+q)-1}{p(1+q)} = \frac{1+2q}{p(1+q)}.$$

當 $p = \frac{1}{6}$，$q = \frac{5}{6}$ 時：

$$\mathrm{E}(S) = 12 - \frac{36}{11} = \frac{96}{11} \approx 8.73.$$

驗算

- $\mathrm{E}(R) + \mathrm{E}(S) = \frac{36}{11} + \frac{96}{11} = 12 = \mathrm{E}(X_1) + \mathrm{E}(X_2)$，這是必然的，因為 $\min + \max = X_1 + X_2$。
- 當 $n = 1$ 時，第一個六點與最後一個六點重合，而 G_R 化為單顆骰子的幾何分佈 PGF $pt/(1-qt)$。
- 以蒙地卡羅模擬（Monte Carlo simulation）擲 2×10^5 對骰子，所得樣本平均值為 3.274 和 8.709，而理論值為 $\frac{36}{11}\approx 3.273$ 和 $\frac{96}{11}\approx 8.727$；另外，以數值方法對 G_S 的級數求和，在 t 的若干個取值上都與閉合形式相符。

常見錯誤

- 把 $\mathrm{P}(S = r)$ 寫成「某一顆骰子在第 r 次擲骰時首次擲出六點，而其餘骰子在此之前已擲出」，卻忘記了兩顆骰子在同一次擲骰中首次擲出六點的情況。改用 $\mathrm{P}(S \leq r)$ 來處理，便可完全避免這個問題。
- 對化簡後的 G_S 求導，而不是對拆開的形式求導：做法正確，但代數運算多得多，出錯的機會也多。

STEP 之外

R 和 S 分別是若干獨立隨機變量的最小值和最大值，是順序統計量（order statistics）最簡單的例子；利用 $\mathrm{P}(\min > r)$ 和 $\mathrm{P}(\max \leq r)$ 這個技巧，對任何分佈都適用。具有無記憶性（memoryless）的幾何分佈，是指數分佈（exponential distribution）在離散情況下的近親；在連續極限之下，n 個指數分佈變量的最大值的期望值變成調和和（harmonic sum）$1 + \frac{1}{2} + \cdots + \frac{1}{n}$（再乘以平均值），這是優惠券收集問題（coupon collector's problem）中的關鍵一步。Grimmett 與 Welsh 所著的 *Probability: An Introduction*（概率論入門）對生成函數及上述各種想法有深入的論述。

詳細解答

STEP III 2000，第 14 題

概率與統計 • 協方差與獨立性

STEP III 2000，第 14 題

繼續閱讀之前，請先看題目（原題為英文）。題目載於 STEP Support Programme 的試題資料庫：

step.maths.org/questions/00-s3-q14

關鍵思路

先補上欠缺的第四個概率（probability），然後把 $\mathrm{E}(XY) - \mathrm{E}(X)\mathrm{E}(Y)$ 逐項寫出。所有項會合併成單一因子 $a - pq$ 乘以一個不可能為零的量，而 $a = pq$ 正好就是說左上角那一格可以分解成邊際概率之積。在 2×2 的表中，只要有一格可以這樣分解，邊際概率就會迫使其餘三格也一樣。若變量可取三個值，迴旋的空間就多了，一個對稱的表便足以打破這種聯繫。

解答

恆等式

四個概率之和為 1，所以

$$\mathrm{P}(X = x_2,\, Y = y_2) = 1 - a - (q - a) - (p - a) = 1 - p - q + a.$$

沿各行各列相加，便得到邊際分佈（marginal distribution）：

$$\mathrm{P}(X = x_1) = a + (q - a) = q, \qquad \mathrm{P}(Y = y_1) = a + (p - a) = p,$$

所以 $\mathrm{P}(X = x_2) = 1 - q$ 及 $\mathrm{P}(Y = y_2) = 1 - p$。這就是 p 和 q 的真正意義。

	y_1	y_2	
x_1	a	$q - a$	q
x_2	$p - a$	$1 - p - q + a$	$1 - q$
	p	$1 - p$	

現在

$$\begin{aligned}\mathrm{E}(XY) &= a\,x_1y_1 + (q-a)\,x_1y_2 + (p-a)\,x_2y_1 + (1-p-q+a)\,x_2y_2,\\ \mathrm{E}(X)\mathrm{E}(Y) &= \big(qx_1 + (1-q)x_2\big)\big(py_1 + (1-p)y_2\big)\\ &= pq\,x_1y_1 + q(1-p)\,x_1y_2 + p(1-q)\,x_2y_1 + (1-p)(1-q)\,x_2y_2.\end{aligned}$$

逐個係數相減：

$$\begin{aligned}x_1y_1:&\quad a - pq,\\ x_1y_2:&\quad (q-a) - (q-pq) = -(a-pq),\\ x_2y_1:&\quad (p-a) - (p-pq) = -(a-pq),\\ x_2y_2:&\quad (1-p-q+a) - (1-p-q+pq) = a-pq.\end{aligned}$$

因此

$$\mathrm{E}(XY) - \mathrm{E}(X)\mathrm{E}(Y) = (a-pq)(x_1y_1 - x_1y_2 - x_2y_1 + x_2y_2) = (a-pq)(x_1-x_2)(y_1-y_2).$$

所以若 $\mathrm{E}(XY) = \mathrm{E}(X)\mathrm{E}(Y)$，則左方為零，即 $(a-pq)(x_1-x_2)(y_1-y_2) = 0$，這正是所求。

獨立性

任何一對各取兩個相異值的隨機變量（random variable），其聯合概率表都具有上述形式，只要選取適當的 a、p、q 便可（替各格命名即可）。若 $\mathrm{E}(XY) = \mathrm{E}(X)\mathrm{E}(Y)$，則上面的乘積為零。由於 $x_1 \neq x_2$ 及 $y_1 \neq y_2$，最後兩個因子均不為零，所以 $a = pq$。

想一想 我們現在知道有一格可以分解。為甚麼這樣就能確定其餘三格？

因為其餘每一格都等於一個邊際概率減去一個已知的格：

$$\begin{aligned}&\mathrm{P}(X = x_1, Y = y_2) = q - pq = q(1-p) = \mathrm{P}(X = x_1)\,\mathrm{P}(Y = y_2),\\ &\mathrm{P}(X = x_2, Y = y_1) = p - pq = p(1-q) = \mathrm{P}(X = x_2)\,\mathrm{P}(Y = y_1),\\ &\mathrm{P}(X = x_2, Y = y_2) = 1 - p - q + pq = (1-p)(1-q) = \mathrm{P}(X = x_2)\,\mathrm{P}(Y = y_2).\end{aligned}$$

每一格都是其邊際概率之積，所以 X 與 Y 獨立（independent）。（反過來，獨立性可推出 $\mathrm{E}(XY) = \mathrm{E}(X)\mathrm{E}(Y)$，這是標準結果，所以對於只取兩個值的變量，這兩個條件是等價的。）

取三個值的例子

由於 A 和 B 各自在 $\{-1, 0, 1\}$ 上均勻分佈，我們有 $\mathrm{E}(A) = \mathrm{E}(B) = 0$，所以我們需要一個表，其每行每列之和均為 $\frac{1}{3}$，且 $\mathrm{E}(AB) = 0$，但並非每一格都等於 $\frac{1}{9}$。一個對稱的選

擇便可做到：

	$B=-1$	$B=0$	$B=1$
$A=-1$	$\frac{1}{6}$	0	$\frac{1}{6}$
$A=0$	0	$\frac{1}{3}$	0
$A=1$	$\frac{1}{6}$	0	$\frac{1}{6}$

每行及每列之和均為 $\frac{1}{3}$，所以 A 和 B 各自取 -1、0、1 的概率均為 $\frac{1}{3}$。只有四個角落的格才給出非零的乘積 ab：

$$\mathrm{E}(AB)=\tfrac{1}{6}(1)+\tfrac{1}{6}(-1)+\tfrac{1}{6}(-1)+\tfrac{1}{6}(1)=0=\mathrm{E}(A)\mathrm{E}(B).$$

但 $\mathrm{P}(A=0,B=0)=\frac{1}{3}\neq\frac{1}{9}=\mathrm{P}(A=0)\,\mathrm{P}(B=0)$，所以 A 與 B 並不獨立。用文字來說：$B=0$ 當且僅當 $A=0$，否則無論 A 是甚麼，B 取 ±1 的機會均等。

驗算

- 試取 $X=Y$，其中 $x_1=y_1=1$，$x_2=y_2=0$：此時 $a=p=q$，而恆等式給出 $\mathrm{Var}(X)=p-p^2$，這對伯努利變量（Bernoulli variable）而言是正確的。
- 這個恆等式是一個多項式恆等式：我們隨機選取了 2000 組 x_i、y_i、a、p、q，以有理數運算精確地驗證了它。同一次驗算亦檢查了取三個值的表：邊際概率全為 $\frac{1}{3}$，$\mathrm{E}(AB)=0$，以及 $\mathrm{P}(A=0,B=0)=\frac{1}{3}$。

常見錯誤

- 證出 $a=pq$ 便停筆。獨立性要求每一格都可以分解，所以其餘三格也需要檢查（很快便可完成）。
- 忘記說明為何可以約去 $(x_1-x_2)(y_1-y_2)$：因為所取的值相異，所以它不為零。
- 在最後一部分提出 $B=A^2$ 或 $B=|A|$。這些例子確實是不獨立而協方差為零，但 B 永不等於 -1，所以它們沒有題目要求的邊際分佈。

STEP 之外

$\mathrm{E}(XY)-\mathrm{E}(X)\mathrm{E}(Y)$ 這個量就是協方差(covariance)，而本題說明「不相關」(uncorrelated)比「獨立」為弱，只有取兩個值的（類似指示變量的）變量例外：對這類變量，恆等式把一切歸結為兩個指示變量（indicator）的協方差。你日後會遇到一個相關的事實：對於聯合正態（jointly normal）的隨機變量，協方差為零確實可推出獨立。想進一步了解，可參閱 Grimmett 與 Welsh 的 *Probability: An Introduction*（概率論入門），或 Blitzstein 與 Hwang 的 *Introduction to Probability*（概率論導論）。

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

2000 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。