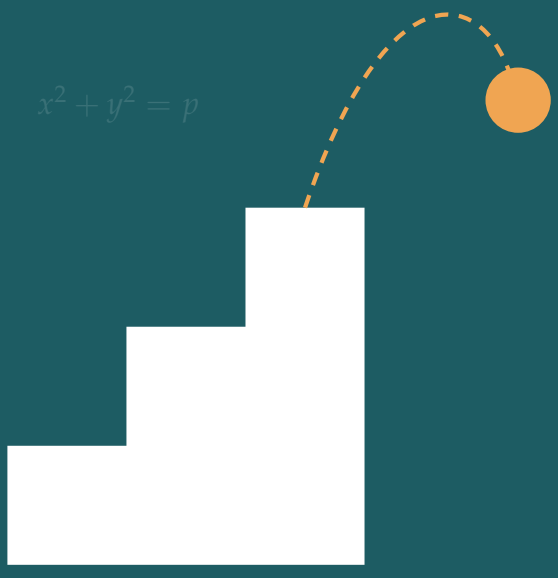
STEP 備試 · 大學數學

Beyond STEP

STEP I, II 及 III

1995 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

1995 年詳細解答

Keith McNulty

Beyond STEP：以歷屆 STEP 試題為門，走進大學數學

1995 年詳細解答

1995 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 多項式不等式 純數學 . . . 3

第 2 題 代換積分法 純數學 . . . 7

第 3 題 裂項相消級數 純數學 . . . 10

第 4 題 棣美弗定理 純數學 . . . 13

第 5 題 二項式級數與積分 純數學 . . . 16

第 6 題 伯努利微分方程 純數學 . . . 20

第 7 題 三角形的向量幾何 純數學 . . . 23

第 8 題 線性微分方程 純數學 . . . 27

第 9 題 拋體軌道的包絡線 力學 . . . 30

第 10 題 彈性繩與反彈 力學 . . . 33

第 11 題 圓柱之間的摩擦力 力學 . . . 37

第 12 題 排列與計數 概率與統計 . . . 41

第 13 題 鬆弛稀疏化 概率與統計 . . . 44

第 14 題 比例的估計 概率與統計 . . . 47

STEP II

第 1 題 等比級數的求導 純數學 . . . 51

第 2 題 以遞推法計算着色數目 純數學 . . . 55

第 3 題 邏輯推理與計數 純數學 . . . 58

第 4 題 沃利斯歸約公式 純數學 . . . 62

第 5 題 最優化與凹性 純數學 . . . 66

第 6 題 單位根 純數學 . . . 69

第 7 題 希羅公式 純數學 . . . 73

第 8 題 邏輯斯諦微分方程 純數學 . . . 77

第 9 題 摩擦力與平衡 力學 . . . 81
第 10 題 碰撞與恢復係數 力學 . . . 84
第 11 題 線性空氣阻力 力學 . . . 87
第 12 題 連續決鬥 概率與統計 . . . 90
第 13 題 最松分佈的最優化 概率與統計 . . . 94
第 14 題 假設檢驗 概率與統計 . . . 98

STEP III

第 1 題 聯立線性方程 純數學 . . . 102
第 2 題 歸約公式 純數學 . . . 106
第 3 題 阻尼諧運動 純數學 . . . 109
第 4 題 三角級數求和 純數學 . . . 113
第 5 題 麥克勞林級數 純數學 . . . 116
第 6 題 軌跡與反演 純數學 . . . 120
第 7 題 群與同構 純數學 . . . 124
第 8 題 平面與球面 純數學 . . . 128
第 9 題 擺擺 力學 . . . 132
第 10 題 斜面上的拋體 力學 . . . 136
第 11 題 相對速度 力學 . . . 139
第 12 題 徑向對稱密度函數 概率與統計 . . . 143
第 13 題 有噪聲的二元信道 概率與統計 . . . 147
第 14 題 矩生成函數 概率與統計 . . . 151

詳細解答

STEP I 1995，第 1 題

純數學 • 多項式不等式

> **STEP I 1995，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/95-s1-q1

關鍵思路

第 (i) 部分的三次式（cubic）可用分組法因式分解（factorise）；一旦寫成三個線性因式（linear factor）的乘積，它的正負號便可從數線上直接讀出。第 (ii) 部分其實是同一個三次式的化身：每一項的總次數（degree）都是 3，所以除以 y^3 之後，它就變成以 x/y 為變量的第 (i) 部分的三次式。同樣的因式分解於是給出三條通過原點的直線，而第 (iii) 部分就是一個正負號表，只不過這次是圍繞原點而畫，而不是沿着一條直線而畫。

解答

第 (i) 部分

把各項兩兩分組，

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

根（root）是 -1、1 和 4。每個因式都是線性的，只會在自己的根處變號，所以乘積在三個根處都會變號。當 x 是很大的正數時，三個因式全是正數，因此從右邊讀起，正負號的分佈是

x	$x < -1$	$-1 < x < 1$	$1 < x < 4$	$x > 4$
正負號	$-$	$+$	$-$	$+$

把三次式等於零的各個根也包括在內，不等式（inequality）成立的充要條件是

$$-1 \le x \le 1 \quad \text{或} \quad x \ge 4.$$

第 (ii) 部分

設 $f(x,y) = x^3 - 4x^2y - xy^2 + 4y^3$。它就是第 (i) 部分的三次式，只是每一項都用 y 的冪補足至 3 次，而同樣的分組法仍然適用：

$$\begin{aligned} f(x,y) &= x^2(x-4y) - y^2(x-4y) \\ &= (x-4y)(x-y)(x+y). \end{aligned}$$

乘積等於零，當且僅當其中一個因式等於零，所以 $f(x,y) = 0$ 在以下三條直線上成立：

$$x = 4y \text{ (即} y = \tfrac{1}{4}x), \qquad y = x, \qquad y = -x,$$

三條直線全都通過原點。

想一想 為甚麼 (ii) 的答案會與 (i) 有關？

當 $y \neq 0$ 時，設 $t = x/y$。那麼 $f(x,y) = y^3\,(t^3 - 4t^2 - t + 4)$，也就是 y^3 乘以第 (i) 部分的三次式。該三次式在 $t = -1, 1, 4$ 時等於零，即是在 $x = -y$、$x = y$ 和 $x = 4y$ 上等於零。（當 $y = 0$ 時，$f = x^3$ 只在原點等於零，而原點本來就在三條直線上。）

第 (iii) 部分

三條直線把平面分成六個在原點相接的扇形區域（sector）。由正 x 軸起逆時針量度，六條半直線所指的方向是

$$\arctan\tfrac{1}{4} \approx 14^\circ, \quad 45^\circ, \quad 135^\circ, \quad 194^\circ, \quad 225^\circ, \quad 315^\circ.$$

在一個扇形區域之內，f 沒有任何因式等於零，所以 f 在該區域內的正負號保持不變。每越過其中一條直線，恰好有一個因式變號（三條直線互不相同，而且在原點以外每次只會越過一條），所以 f 的正負號由一個扇形區域到下一個扇形區域交替變化。

要確定正負號，可取 $(1,0)$：$f(1,0) = 1 > 0$，而 $(1,0)$ 位於 315° 與 14° 之間的扇形區域內。繞一圈下去，正負號因此是

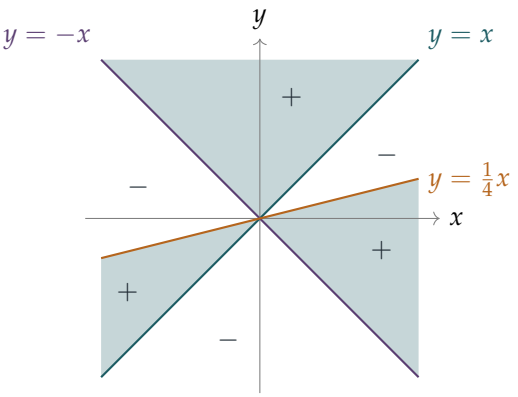
扇形區域（度）	$-45, 14$	$14, 45$	$45, 135$	$135, 194$	$194, 225$	$225, 315$
正負號：f	$+$	$-$	$+$	$-$	$+$	$-$

抽樣檢查一下：$f(0,1) = 4 > 0$ 屬於第三欄，$f(-3,-1) = 8 > 0$ 屬於第五欄。

所以 $f(x,y) \geq 0$ 在下圖三個有陰影的扇形區域上成立，另外還包括全部三條直線（在直線上 $f = 0$）：

- 包含正 x 軸的扇形區域，介乎 $y = -x$（下方）與 $y = \frac{1}{4}x$（上方）之間，其中 $x > 0$；
- 包含正 y 軸的扇形區域，介乎 $y = x$ 與 $y = -x$ 之間，其中 $y > 0$；

- 第三象限內介乎 $y = x$ 與 $y = \frac{1}{4}x$ 之間的狹窄扇形區域。



想一想 為甚麼把圖形繞原點旋轉半圈之後，圖形不變，只是陰影部分與空白部分對調？

f 的每一項都是奇數次 3 次，所以 $f(-x,-y) = -f(x,y)$。把一點繞原點旋轉 180°，會令 f 的正負號反轉，所以每個有陰影的扇形區域的對面都是一個沒有陰影的扇形區域。這是檢查任何草圖的快捷方法：有陰影和沒有陰影的扇形區域必須成對地相對，而圖中確實如此。

驗算

- 在直線 $y = 1$ 上，$f(x,1)$ 正好是第 (i) 部分的三次式，所以 $y = 1$ 上有陰影的部分應該是 $-1 \le x \le 1$ 和 $x \ge 4$。在草圖中，$y = 1$ 在 $y = -x$ 與 $y = x$ 之間（由 $x = -1$ 到 $x = 1$）有陰影，並且在它與 $y = \frac{1}{4}x$ 相交之處，即 $x = 4$，再次進入有陰影的扇形區域。
- 在 x 軸上 $f = x^3$，它非負的充要條件是 $x \ge 0$：正 x 軸有陰影，負的一半則沒有。
- 數值上，在 2×10^5 個隨機點計算 f 的值，其正負號每次都與六個扇形區域的正負號表吻合；例如 $f(1,0) = 1$、$f(0,-1) = -4$、$f(-3,-1) = 8$。

常見錯誤

- 除以 y^3 之後，在 $y < 0$ 時仍然保留原來的不等號。由於該處 $y^3 < 0$，條件變成 $t^3 - 4t^2 - t + 4 \le 0$，這正是下半平面看起來與上半平面不同的原因。
- 只在 $y > 0$ 時用第 (i) 部分的答案來畫陰影，然後完全忘記了 $y < 0$ 的情況。
- 漏掉直線本身：不等式並非嚴格不等式，所以直線也屬於該區域。

STEP 之外

各項次數全部相同的多項式稱為齊次（homogeneous）多項式。它的零點集是若干條通過原點的直線的併集，而它的正負號只取決於方向。齊次多項式是射影幾何（projective geometry）的起點，在射影幾何中，一個點就是一整條通過原點的直線；當你在線性代數（linear algebra）課程中為圓錐曲線（conics）和二次型（quadratic forms）分類時，它們又會再次出現。Strang 的 *Introduction to Linear Algebra* 涵蓋二次型，而 Siklos 的 *Advanced Problems in Mathematics* 則有更多這種風格的 STEP 題目。

詳細解答

STEP I 1995，第 2 題

純數學 • 代換積分法

> **STEP I 1995，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/95-s1-q2

關鍵思路

在第 (i) 部分，兩個積分（integral）單獨來看都不易處理，但兩者卻是天作之合：相加後得到的分式可約簡為 1，相減後得到的分子恰好是分母的導數（derivative）。兩個容易的積分，給出關於 S 和 T 的兩條聯立方程（simultaneous equations）。在第 (ii) 部分，選用代換（substitution）$x = \sin^2 t$，是因為它把 $\frac{1}{x} - 1$ 化為完全平方 $\cot^2 t$，令平方根消失；其後全靠倍角公式（double-angle formulae）便可完成。

解答

第 (i) 部分

把兩個被積函數（integrand）相加，

$$S + T = \int \frac{\cos x + \sin x}{\cos x + \sin x}\,\mathrm{d}x = \int 1\,\mathrm{d}x = x + c_1.$$

把兩者相減，

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

分子正是分母的導數，因為 $\frac{\mathrm{d}}{\mathrm{d}x}(\cos x + \sin x) = -\sin x + \cos x$。所以這個積分屬於 $\int \frac{f'(x)}{f(x)}\,\mathrm{d}x$ 的形式，而

$$S - T = \ln|\cos x + \sin x| + c_2.$$

把這兩條方程相加、相減，再除以二，

$$S = \tfrac{1}{2}(x + \ln|\cos x + \sin x|) + C,$$
$$T = \tfrac{1}{2}(x - \ln|\cos x + \sin x|) + C',$$

其中 C 和 C' 為任意常數（arbitrary constants）。在任何滿足 $\cos x + \sin x \neq 0$ 的區間上，被積函數均有定義，以上結果亦在該區間上成立。

第 (ii) 部分

設 I 為所求的積分。取 $x = \sin^2 t$，則 $\mathrm{d}x = 2\sin t \cos t \,\mathrm{d}t$。至於積分上下限（limits），$x = \frac{1}{4}$ 給出 $\sin t = \frac{1}{2}$，所以 $t = \frac{\pi}{6}$；而 $x = \frac{1}{2}$ 給出 $\sin t = \frac{1}{\sqrt{2}}$，所以 $t = \frac{\pi}{4}$。（我們取 t 在 $[\frac{\pi}{6}, \frac{\pi}{4}]$ 內；當 t 走遍這個區間時，x 由 $\frac{1}{4}$ 遞增至 $\frac{1}{2}$。）

接着處理平方根：

$$\frac{1}{x} - 1 = \frac{1 - \sin^2 t}{\sin^2 t} = \frac{\cos^2 t}{\sin^2 t}, \qquad \text{所以} \qquad \sqrt{\frac{1}{x} - 1} = \frac{\cos t}{\sin t}.$$

想一想 為甚麼平方根是 $\cos t / \sin t$，而不是 $|\cos t / \sin t|$？

一般而言，$\sqrt{\cot^2 t} = |\cot t|$。但當 $\frac{\pi}{6} \leq t \leq \frac{\pi}{4}$ 時，$\sin t$ 和 $\cos t$ 皆為正數，所以 $\cot t > 0$，絕對值（modulus）符號可以略去。這正是我們選取 t 在第一象限（first quadrant）的原因。

代入後，$\sin t$ 互相約去：

$$I = \int_{\pi/6}^{\pi/4} (1 - 4\sin^2 t) \frac{\cos t}{\sin t} 2\sin t \cos t \,\mathrm{d}t = \int_{\pi/6}^{\pi/4} 2\cos^2 t \,(1 - 4\sin^2 t) \,\mathrm{d}t.$$

現在利用倍角公式，令這個積分變得容易計算。由於 $2\cos^2 t = 1 + \cos 2t$ 及 $8\sin^2 t \cos^2 t = 2\sin^2 2t = 1 - \cos 4t$，

$$2\cos^2 t - 8\sin^2 t \cos^2 t = (1 + \cos 2t) - (1 - \cos 4t) = \cos 2t + \cos 4t.$$

因此

$$\begin{aligned} I &= \int_{\pi/6}^{\pi/4} (\cos 2t + \cos 4t) \,\mathrm{d}t = \left[\tfrac{1}{2}\sin 2t + \tfrac{1}{4}\sin 4t\right]_{\pi/6}^{\pi/4} \\ &= \left(\tfrac{1}{2}\sin\tfrac{\pi}{2} + \tfrac{1}{4}\sin\pi\right) - \left(\tfrac{1}{2}\sin\tfrac{\pi}{3} + \tfrac{1}{4}\sin\tfrac{2\pi}{3}\right) \\ &= \tfrac{1}{2} - \left(\tfrac{\sqrt{3}}{4} + \tfrac{\sqrt{3}}{8}\right), \end{aligned}$$

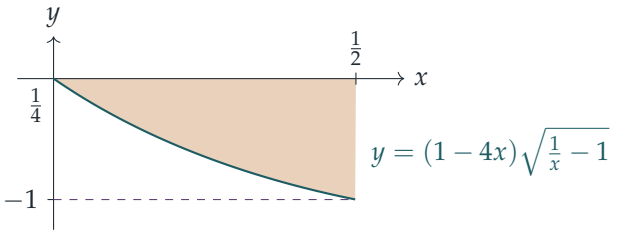
所以

$$\int_{1/4}^{1/2} (1 - 4x)\sqrt{\frac{1}{x} - 1} \,\mathrm{d}x = \frac{1}{2} - \frac{3\sqrt{3}}{8} = \frac{4 - 3\sqrt{3}}{8}.$$

驗算

- 正負號：在 $(\frac{1}{4}, \frac{1}{2}]$ 上，$1 - 4x < 0$ 而平方根為正，所以被積函數為負，I 必為負數。果然 $4 - 3\sqrt{3} \approx -1.196$，所以 $I \approx -0.1495$。下圖顯示被積函數的圖像。

- 把辛普森法則（Simpson's rule）直接應用於原來的積分，得出 $-0.149519\ldots$，與 $\frac{1}{2}-\frac{3\sqrt{3}}{8}$ 吻合至小數點後十個位。
- 至於第 (i) 部分，求導驗證：$\frac{\mathrm{d}}{\mathrm{d}x}\frac{1}{2}\big(x+\ln(\cos x+\sin x)\big) = \frac{1}{2}\left(1+\frac{\cos x-\sin x}{\cos x+\sin x}\right) = \frac{\cos x}{\cos x+\sin x}$。例如在 $[0,1]$ 上，S 和 T 的數值均與數值積分（quadrature）的結果相符（0.66168 和 0.33832）。



常見錯誤

- 忘記把上下限改為 $t=\frac{\pi}{6}$ 和 $t=\frac{\pi}{4}$，或在代回 x 時出錯。
- 漏掉 $\mathrm{d}x$ 中的因子 $2\sin t\cos t$，或寫成 $\mathrm{d}x=2\sin t\,\mathrm{d}t$。
- 得出正數答案：被積函數在整個區間上都是負數，所以答案若為正值，即表示正負號出了錯。

STEP 之外

第 (i) 部分的技巧，是替一個棘手的積分找一個「拍檔」，使兩者之和與差都容易計算。這個技巧一再出現，例如用於 $\int \mathrm{e}^x\cos x\,\mathrm{d}x$ 和 $\int \mathrm{e}^x\sin x\,\mathrm{d}x$，而它的複數形式則是對 $\mathrm{e}^{(1+\mathrm{i})x}$ 積分。第 (ii) 部分是一個普遍原則的小例子：選得恰當的代換，可以把代數形式的被積函數化為三角多項式（trigonometric polynomial），而這正是傅立葉級數（Fourier series）的起點。想看更多這類風格的積分，並附有細緻的討論，可參閱 Siklos 的 *Advanced Problems in Mathematics*（《高等數學難題》）。

詳細解答

STEP I 1995，第 3 題

純數學 • 裂項相消級數

> **STEP I 1995，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/95-s1-q3

關鍵思路

當一個和式中相鄰各項有大部分內容相同時，把每一項寫成差 $\mathrm{f}(r)-\mathrm{f}(r-1)$ 的形式，幾乎所有東西都會互相抵消：這個和式會*裂項相消*（telescope）。功夫在於選取 f，使這個差恰好是你想要的那一項；本題已替我們選好了。至於交錯級數（alternating series），竅門是把它寫成全部立方之和減去偶數立方之和的兩倍，於是第 (ii) 部分的結果可以用上兩次。

解答

第 (i) 部分

把和式完整寫出：

$$\sum_{r=1}^{n}\{\mathrm{f}(r)-\mathrm{f}(r-1)\}=\{\mathrm{f}(1)-\mathrm{f}(0)\}+\{\mathrm{f}(2)-\mathrm{f}(1)\}+\cdots+\{\mathrm{f}(n)-\mathrm{f}(n-1)\}.$$

$\mathrm{f}(1),\mathrm{f}(2),\ldots,\mathrm{f}(n-1)$ 每一個都帶正號出現一次、帶負號出現一次，所以全部抵消。餘下的只有第一個括號中的 $-\mathrm{f}(0)$ 和最後一個括號中的 $\mathrm{f}(n)$，得出 $\mathrm{f}(n)-\mathrm{f}(0)$，正是所求。

如果你想要比「全部抵消」更嚴謹的說法，可以對 n 使用歸納法（induction）。當 $n=1$ 時，和為 $\mathrm{f}(1)-\mathrm{f}(0)$。若結果對 $n=k$ 成立，則加上下一項後得

$$\mathrm{f}(k)-\mathrm{f}(0)+\mathrm{f}(k+1)-\mathrm{f}(k)=\mathrm{f}(k+1)-\mathrm{f}(0),$$

這就是 $n=k+1$ 時的結果。因此結果對所有 $n\geq 1$ 成立。

第 (ii) 部分

取 $\mathrm{f}(r) = r^2(r+1)^2$，則 $\mathrm{f}(r-1) = (r-1)^2r^2$。兩者都含有因子 r^2，所以把它抽出，再運用平方差（difference of two squares）：

$$\begin{aligned}\mathrm{f}(r) - \mathrm{f}(r-1) &= r^2\left[(r+1)^2 - (r-1)^2\right] \\ &= r^2\left[(r+1) + (r-1)\right]\left[(r+1) - (r-1)\right] \\ &= r^2 \cdot 2r \cdot 2 = 4r^3.\end{aligned}$$

現在由 $r = 1$ 加至 n，並應用第 (i) 部分。由於 $\mathrm{f}(0) = 0$，

$$4\sum_{r=1}^{n} r^3 = \mathrm{f}(n) - \mathrm{f}(0) = n^2(n+1)^2, \qquad \text{所以} \qquad \boxed{\sum_{r=1}^{n} r^3 = \tfrac{1}{4}n^2(n+1)^2.}$$

第 (iii) 部分

這個級數（series）共有 $2n+1$ 項：奇數的立方帶正號，偶數的立方帶負號，最後一項是奇數立方 $(2n+1)^3$。

? **想一想** 只用 $\sum r^3$ 的公式，怎樣處理正負交錯的符號？

先考慮直到 $(2n+1)^3$ 為止的全部立方之和。當中每個偶數立方都帶正號，而我們想要的是負號，所以把每個偶數立方減去兩次：

$$1^3 - 2^3 + 3^3 - \cdots + (2n+1)^3 = \sum_{k=1}^{2n+1} k^3 - 2\sum_{m=1}^{n} (2m)^3.$$

偶數立方為 $(2m)^3 = 8m^3$，其中 $m = 1, \ldots, n$，最大的是 $(2n)^3$。運用第 (ii) 部分兩次，其中一次以 $2n+1$ 代替 n，

$$\begin{aligned}\sum_{k=1}^{2n+1} k^3 - 16\sum_{m=1}^{n} m^3 &= \tfrac{1}{4}(2n+1)^2(2n+2)^2 - 16 \cdot \tfrac{1}{4}n^2(n+1)^2 \\ &= (2n+1)^2(n+1)^2 - 4n^2(n+1)^2 \\ &= (n+1)^2\left[(2n+1)^2 - (2n)^2\right] \\ &= (n+1)^2(4n+1),\end{aligned}$$

其中第二行用了 $(2n+2)^2 = 4(n+1)^2$，最後一行則用了平方差 $(2n+1)^2 - (2n)^2 = (2n+1+2n) \cdot 1$。因此

$$\boxed{1^3 - 2^3 + 3^3 - 4^3 + \cdots + (2n+1)^3 = (n+1)^2(4n+1).}$$

驗算

- 當 $n = 0$ 時，級數只有 $1^3 = 1$，而 $(0+1)^2(0+1) = 1$。
- 當 $n = 1$ 時：$1 - 8 + 27 = 20$，而 $2^2 \times 5 = 20$。
- 當 $n = 2$ 時：$1 - 8 + 27 - 64 + 125 = 81$，而 $3^2 \times 9 = 81$。
- 答案的增長應與 $(2n+1)^3/2$ 相若，即大約是最後一項的一半，因為各項可配對成 $(2m+1)^3 - (2m)^3 \approx 12m^2$。事實上 $(n+1)^2(4n+1) \approx 4n^3$，與 $\sum 12m^2 \approx 4n^3$ 吻合。

我們亦以直接求和的方法，對由 0 至 199 的每一個 n 驗證了第 (ii) 和第 (iii) 部分。

常見錯誤

- 只把偶數立方減去一次，得出 $\sum k^3 - \sum (2m)^3$，這是奇數立方之和，而不是交錯和。
- 把偶數立方取到 $(2n+2)^3$，或在全部立方之和的公式中用了 n 而不是 $2n+1$：要仔細數清楚項數。

STEP 之外

第 (i) 部分是微積分基本定理（fundamental theorem of calculus）的離散版本：差 $\mathrm{f}(r) - \mathrm{f}(r-1)$ 扮演導數（derivative）的角色，把它求和便可還原 f，正如對 F' 積分會得出 $F(b) - F(a)$ 一樣。這套「有限差分演算」（calculus of finite differences）讓你只要找到合適的 f，就能求出任何多項式的和；而同樣的裂項相消思想，也是許多級數收斂（convergence）證明的基礎。Siklos 的 *Advanced Problems in Mathematics*（《高等數學難題》）有更多關於裂項相消和這類級數的練習。

詳細解答

STEP I 1995，第 4 題

純數學 • 棣美弗定理

> **STEP I 1995，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/95-s1-q4

關鍵思路

棣美弗定理（de Moivre's theorem）把 $\sin 5\theta$ 化為 $(\cos\theta + \mathrm{i}\sin\theta)^5$ 的虛部（imaginary part），而二項式定理（binomial theorem）則把這個冪展開成 $\cos\theta$ 與 $\sin\theta$ 的乘積。虛部的每一項都含有 $\sin\theta$ 的奇數次冪，所以我們可以抽出一個因子，再利用 $\sin^2\theta = 1 - \cos^2\theta$ 把餘下的部分化為餘弦。接着，角 $\frac{1}{5}\pi$ 之所以特別，是因為在該處 $\sin 5\theta = 0$，於是剩下一條關於 $\cos^2\theta$ 的二次方程（quadratic equation）。

解答

恆等式

設 $c = \cos\theta$ 及 $s = \sin\theta$。由棣美弗定理，

$$\cos 5\theta + \mathrm{i}\sin 5\theta = (c + \mathrm{i}s)^5.$$

由二項式定理，係數為 $1, 5, 10, 10, 5, 1$，

$$(c + \mathrm{i}s)^5 = c^5 + 5c^4(\mathrm{i}s) + 10c^3(\mathrm{i}s)^2 + 10c^2(\mathrm{i}s)^3 + 5c(\mathrm{i}s)^4 + (\mathrm{i}s)^5.$$

由於 $\mathrm{i}^2 = -1$、$\mathrm{i}^3 = -\mathrm{i}$、$\mathrm{i}^4 = 1$ 及 $\mathrm{i}^5 = \mathrm{i}$，虛數項就是含 $\mathrm{i}s$ 奇數次冪的各項：

$$\mathrm{Im}(c + \mathrm{i}s)^5 = 5c^4 s - 10c^2 s^3 + s^5.$$

c 和 s 都是實數，所以比較虛部可得

$$\begin{aligned}\sin 5\theta &= s\left(5c^4 - 10c^2s^2 + s^4\right)\\ &= s\left(5c^4 - 10c^2(1-c^2) + (1-c^2)^2\right)\\ &= s\left(5c^4 - 10c^2 + 10c^4 + 1 - 2c^2 + c^4\right)\\ &= s\left(16c^4 - 12c^2 + 1\right),\end{aligned}$$

其中第二行用了 $s^2 = 1 - c^2$。也就是說，

$$\sin 5\theta = \sin\theta\left(16\cos^4\theta - 12\cos^2\theta + 1\right),$$

這正是所要證明的恆等式（identity）。

求 $\cos^2\frac{1}{5}\pi$ 的值

取 $\theta = \frac{1}{5}\pi$。這時 $\sin 5\theta = \sin\pi = 0$，而 $\sin\frac{1}{5}\pi \neq 0$（因為 $0 < \frac{1}{5}\pi < \pi$）。所以我們可以把恆等式兩邊除以 $\sin\theta$，於是 $x = \cos^2\frac{1}{5}\pi$ 滿足

$$16x^2 - 12x + 1 = 0, \qquad \text{所以} \qquad x = \frac{12 \pm \sqrt{144-64}}{32} = \frac{12 \pm 4\sqrt{5}}{32} = \frac{3 \pm \sqrt{5}}{8}.$$

想一想 這條二次方程有兩個根。哪一個才是 $\cos^2\frac{1}{5}\pi$？

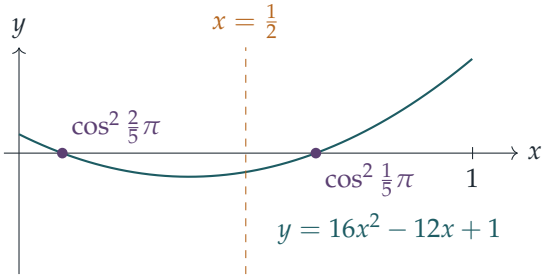
要選出正確的根，我們還需要多一項關於 $\cos\frac{1}{5}\pi$ 的事實。由於 $0 < \frac{1}{5}\pi < \frac{1}{4}\pi$，而 cos 在 $[0, \frac{1}{2}\pi]$ 上遞減，我們有 $\cos\frac{1}{5}\pi > \cos\frac{1}{4}\pi = \frac{1}{\sqrt{2}}$，所以

$$\cos^2\tfrac{1}{5}\pi > \tfrac{1}{2}.$$

現在由 $\sqrt{5} < 3$ 得 $\frac{3-\sqrt{5}}{8} < \frac{3}{8} < \frac{1}{2}$，而由 $\sqrt{5} > 1$ 得 $\frac{3+\sqrt{5}}{8} > \frac{1}{2}$。恰好只有一個根大於 $\frac{1}{2}$，所以

$$\cos^2\tfrac{1}{5}\pi = \frac{3+\sqrt{5}}{8}.$$

（另一個根並非多餘的假根：當 $\theta = \frac{2}{5}\pi$ 時同樣有 $\sin 5\theta = 0$ 及 $\sin\theta \neq 0$，而 $\cos^2\frac{2}{5}\pi = \frac{3-\sqrt{5}}{8}$。）



推導 $\cos\frac{1}{5}\pi$

由於 $0 < \frac{1}{5}\pi < \frac{1}{2}\pi$，$\cos\frac{1}{5}\pi > 0$，所以我們取正平方根：

$$\cos\tfrac{1}{5}\pi = \sqrt{\frac{3+\sqrt{5}}{8}} = \sqrt{\frac{6+2\sqrt{5}}{16}} = \frac{\sqrt{6+2\sqrt{5}}}{4}.$$

最後，$(1+\sqrt{5})^2 = 1+2\sqrt{5}+5 = 6+2\sqrt{5}$ 而且 $1+\sqrt{5} > 0$，所以 $\sqrt{6+2\sqrt{5}} = 1+\sqrt{5}$，於是

$$\cos\tfrac{1}{5}\pi = \tfrac{1}{4}\left(1+\sqrt{5}\right),$$

這正是所要求的結果。

驗算

- 當 $\theta = \frac{1}{2}\pi$ 時，恆等式給出 $\sin\frac{5}{2}\pi = 1\cdot(0-0+1) = 1$，結果正確。當 θ 很小時：兩邊都約等於 5θ，因為括號內的式子趨向 $16-12+1=5$。
- 以數值計算，$\cos 36^\circ \approx 0.809017$ 而 $\frac{1}{4}(1+\sqrt{5}) \approx 0.809017$；另外 $\cos^2 36^\circ \approx 0.654508 = \frac{3+\sqrt{5}}{8}$。我們亦在 10^4 個隨機角度上檢驗過這個恆等式。
- $\frac{1}{4}(1+\sqrt{5})$ 是黃金比（golden ratio）φ 的一半，而黃金比正是正五邊形的對角線與邊長之比，這與 36° 這個角所暗示的一致。

常見錯誤

- 忘記 i 的冪所帶來的正負號：s^3 項帶有 $\mathrm{i}^3 = -\mathrm{i}$，因而得到 $-10c^2s^3$。
- 毫無理由地選定二次方程的其中一個根。你需要一個不等式（inequality），例如 $\cos^2\frac{1}{5}\pi > \frac{1}{2}$，才能排除 $\frac{3-\sqrt{5}}{8}$。
- 除以 $\sin\theta$ 時沒有指出 $\sin\frac{1}{5}\pi \neq 0$。

STEP 之外

用同樣的方法，可以把 $\sin n\theta / \sin\theta$ 和 $\cos n\theta$ 寫成 $\cos\theta$ 的多項式，即切比雪夫多項式（Chebyshev polynomials），它們在數值分析（numerical analysis）中隨處可見。$\mathrm{e}^{2\pi\mathrm{i}k/5}$ 這些數是五次單位根（fifth roots of unity），而 $\cos\frac{1}{5}\pi$ 只涉及平方根這一事實，正是正五邊形可以用尺規作圖（ruler-and-compasses construction）的原因；高斯（Gauss）關於哪些正多邊形可以尺規作圖的定理，是大學本科伽羅瓦理論（Galois theory）的一大亮點。至於棣美弗定理背後的幾何，可參閱 Needham 的 *Visual Complex Analysis*（《複分析：可視化方法》）。

詳細解答

STEP I 1995，第 5 題

純數學 • 二項式級數與積分

> **STEP I 1995，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/95-s1-q5

關鍵思路

對多項式逐項求導（differentiating term by term），可以消去麻煩的 $1/r$ 因子，剩下的幾乎就是一個二項展開式（binomial expansion）：係數 $(-1)^{r+1}\binom{n}{r}$ 正好是 $-(1-x)^n$ 去掉常數項後的係數。一旦求得 f'，微積分基本定理（fundamental theorem of calculus）（連同 $\mathrm{f}(0)=0$）便把 f 表示成一個積分，再用代換（substitution）$y=1-t$ 就可化成所要求的形式。那裏的被積函數（integrand）是一個等比級數（geometric series），可以立即積分。

解答

導數

記

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

這是一個多項式，所以可以逐項求導：

$$\mathrm{f}'(x)=\sum_{r=1}^{n}(-1)^{r+1}\binom{n}{r}x^{r-1}.$$

當 $x \neq 0$ 時，乘以 x 並利用 $(-1)^{r+1}x^r = -(-x)^r$：

$$\begin{aligned} x\,\mathrm{f}'(x) &= -\sum_{r=1}^{n} \binom{n}{r}(-x)^r \\ &= -\Big(\sum_{r=0}^{n} \binom{n}{r}(-x)^r - 1\Big) \\ &= 1-(1-x)^n, \end{aligned}$$

這裏用了二項式定理（binomial theorem）。除以 x，得

$$\mathrm{f}'(x) = \frac{1-(1-x)^n}{x} \qquad (x \neq 0),$$

正是所求。（在 $x = 0$ 處，由級數得 $\mathrm{f}'(0) = n$，這也是右邊當 $x \to 0$ 時的極限，因為 $1-(1-x)^n = nx + O(x^2)$。）

積分公式

f 的每一項都含有因子 x，所以 $\mathrm{f}(0) = 0$。由微積分基本定理，

$$\mathrm{f}(x) = \mathrm{f}(0) + \int_0^x \mathrm{f}'(t)\,\mathrm{d}t = \int_0^x \frac{1-(1-t)^n}{t}\,\mathrm{d}t.$$

想一想 公式的分母含有 t，積分時穿過 $t = 0$ 是否穩妥？

穩妥。被積函數其實就是多項式 $\mathrm{f}'(t)$；那個分數只是它在 $t \neq 0$ 時的一種簡潔寫法，而它在 $t = 0$ 處的值就是極限 n。所以這裏並沒有奇點（singularity），只是公式中有一個可去的缺口。

現在作代換 $y = 1 - t$，於是 $\mathrm{d}t = -\mathrm{d}y$；當 $t = 0$ 時 $y = 1$，當 $t = x$ 時 $y = 1 - x$。於是

$$\begin{aligned} \mathrm{f}(x) &= \int_1^{1-x} \frac{1-y^n}{1-y}\,(-\mathrm{d}y) \\ &= \int_{1-x}^{1} \frac{1-y^n}{1-y}\,\mathrm{d}y, \end{aligned}$$

其中把積分上下限對調，以抵消負號。這就是所要求的結果，而且對任何實數 x 都成立。

$\mathrm{f}(1)$ 的值

取 $x = 1$，則積分下限為 0。當 $y \neq 1$ 時，被積函數是一個等比級數之和，

$$\frac{1-y^n}{1-y} = 1 + y + y^2 + \cdots + y^{n-1},$$

而（一如前述）它在 $y = 1$ 處的值是極限 n，所以在整個 $[0,1]$ 上，被積函數就是這個多項式。因此

$$\begin{aligned} \mathrm{f}(1) &= \int_0^1 \left(1 + y + \cdots + y^{n-1}\right) \mathrm{d}y \\ &= \left[y + \frac{y^2}{2} + \cdots + \frac{y^n}{n} \right]_0^1 \\ &= 1 + \frac{1}{2} + \frac{1}{3} + \cdots + \frac{1}{n}, \end{aligned}$$

正是所求。

留意這個結果的含義：在原來的級數中代入 $x = 1$，便得到一個令人驚訝的恆等式

$$\sum_{r=1}^{n} \frac{(-1)^{r+1}}{r} \binom{n}{r} = 1 + \frac{1}{2} + \cdots + \frac{1}{n},$$

若要直接看出，絕不容易。

驗算

- $n = 1$：$\mathrm{f}(x) = x$，而 $\big(1 - (1 - x)\big)/x = 1 = \mathrm{f}'(x)$；同時 $\mathrm{f}(1) = 1$。
- $n = 4$：$\mathrm{f}(1) = 4 - \frac{6}{2} + \frac{4}{3} - \frac{1}{4} = \frac{25}{12} = 1 + \frac{1}{2} + \frac{1}{3} + \frac{1}{4}$。
- 我們在 300 個隨機有理點上精確地驗證了導數公式，又對隨機的 $x \in [-1.5, 1.5]$ 和 $n \le 15$ 把積分（用辛普森法則（Simpson's rule）計算）與級數比較（相對誤差小於 10^{-11}），並對每個 $n \le 40$ 精確地確認了關於 $\mathrm{f}(1)$ 的恆等式。

常見錯誤

- 與 $(1 - x)^n$ 的二項展開式對照時，忘記了 $r = 0$ 那一項：由 $r = 1$ 開始的和是 $(1 - x)^n - 1$，而不是 $(1 - x)^n$。
- 寫出 $\mathrm{f}(x) = \int_0^x \mathrm{f}'(t)\,\mathrm{d}t$ 而沒有指出 $\mathrm{f}(0) = 0$。
- 作代換 $y = 1 - t$ 之後把積分上下限的次序弄反：正是 $\mathrm{d}t = -\mathrm{d}y$ 的負號把上下限調回來。

STEP 之外

$H_n = 1 + \frac{1}{2} + \cdots + \frac{1}{n}$ 這些數稱為調和數（harmonic numbers）；它們的增長速度與 $\ln n$ 相若，而 $H_n - \ln n$ 趨向歐拉常數（Euler's constant）$\gamma \approx 0.5772$。這裏所用的技巧，即先求導以消去麻煩的分母，然後再積分回去，正是推導 $\ln(1 + x)$ 和 $\tan^{-1} x$ 的冪級數

（power series）時所用的方法；而它對無窮級數是否成立的論證（冪級數的逐項求導），是大學一年級分析學（analysis）的核心課題。Abbott 的 *Understanding Analysis* 對此有嚴謹的論述。

詳細解答

STEP I 1995，第 6 題

純數學 • 伯努利微分方程

STEP I 1995，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/95-s1-q6

關鍵思路

兩條方程對 y 而言都不是線性的，但各自對 $1/y$ 的某個適當冪次卻是線性的。左邊含有 y'/y^2，它（除正負號外）正是 $1/y$ 的導數（derivative）；因此設 $u = 1/y$，方程便化為關於 u 的一階線性方程（first-order linear equation），再用積分因子（integrating factor）即可解決。在第 (ii) 部分，y'/y^3 一項提示我們應改為考慮 $1/y^2$。

解答

第 (i) 部分

設 $u = 1/y$。由鏈式法則（chain rule），

$$\frac{\mathrm{d}u}{\mathrm{d}x} = -\frac{1}{y^2}\frac{\mathrm{d}y}{\mathrm{d}x},$$

所以方程的第一項恰好是 $-\mathrm{d}u/\mathrm{d}x$，而第二項是 u。方程變成

$$-\frac{\mathrm{d}u}{\mathrm{d}x} + u = \mathrm{e}^{2x}, \qquad \text{即} \qquad \frac{\mathrm{d}u}{\mathrm{d}x} - u = -\mathrm{e}^{2x}.$$

這對 u 而言是線性的。積分因子是 $\mathrm{e}^{\int -1\,\mathrm{d}x} = \mathrm{e}^{-x}$，全式乘以它便得

$$\frac{\mathrm{d}}{\mathrm{d}x}\left(u\,\mathrm{e}^{-x}\right) = \mathrm{e}^{-x}\frac{\mathrm{d}u}{\mathrm{d}x} - \mathrm{e}^{-x}u = -\mathrm{e}^{x}.$$

兩邊積分，其中 A 為任意常數（arbitrary constant），

$$u\,\mathrm{e}^{-x} = A - \mathrm{e}^{x}, \qquad \text{所以} \qquad u = A\mathrm{e}^{x} - \mathrm{e}^{2x}.$$

由於 $y = 1/u$，

$$y = \frac{1}{A\mathrm{e}^{x} - \mathrm{e}^{2x}},$$

這就是所求的通解（general solution）。（原方程含有 $1/y$，所以自始至終 $y \neq 0$，代換不會遺漏任何解；此公式在任何滿足 $Ae^x \neq e^{2x}$ 的區間上都成立，即當 $A > 0$ 時須避開 $x = \ln A$。）

第 (ii) 部分

> **想一想** 這裏哪一個代換能擔當 $u = 1/y$ 的角色？

我們希望新變量的導數能產生 y'/y^3。由於

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

正確的選擇是 $v = 1/y^2$。於是第一項是 $-\frac{1}{2}\,\mathrm{d}v/\mathrm{d}x$，第二項是 v，所以

$$-\frac{1}{2}\frac{\mathrm{d}v}{\mathrm{d}x} + v = e^{2x}, \qquad \text{即} \qquad \frac{\mathrm{d}v}{\mathrm{d}x} - 2v = -2e^{2x}.$$

這次積分因子是 e^{-2x}：

$$\frac{\mathrm{d}}{\mathrm{d}x}\left(v\,e^{-2x}\right) = -2e^{2x}\cdot e^{-2x} = -2.$$

兩邊積分，其中 B 為任意常數，

$$v\,e^{-2x} = B - 2x, \qquad \text{所以} \qquad \frac{1}{y^2} = (B - 2x)\,e^{2x}.$$

因此

$$y^2 = \frac{e^{-2x}}{B - 2x}, \qquad y = \pm\frac{e^{-x}}{\sqrt{B - 2x}}.$$

因為 $1/y^2 > 0$，我們需要 $B - 2x > 0$：解只在 $x < \frac{1}{2}B$ 時存在，而正負號則由 y 最初的符號決定（它不可能變號，因為 y 永不為零）。與第 (i) 部分不同，答案並非兩個指數函數（exponential）之差：這裏右邊的 e^{2x} 恰好與積分因子 e^{-2x} 互相抵消，剩下一個常數待積分，這就是線性項 $-2x$ 出現的原因。

驗算

- 代回 (ii)：由 $1/y^2 = (B - 2x)e^{2x}$，求導得 $-2y'/y^3 = -2e^{2x} + 2(B - 2x)e^{2x}$，所以 $y'/y^3 = e^{2x} - 1/y^2$，這正是原方程。
- 數值檢驗：在 (ii) 中取 $y(0) = 1$ 得 $B = 1$，於是公式預測 $y(0.4) = e^{-0.4}/\sqrt{0.2} \approx 1.49888$。由 $x = 0$ 開始以數值方法逐步求解該微分方程（differential equation），所得的值與此相同，準確至小數點後十位。兩個通解在許多隨機選取的點和常數下，亦都滿足各自的方程。

常見錯誤

- 在 $\mathrm{d}u/\mathrm{d}x$ 中漏掉負號（或第 (ii) 部分的因子 2），令整條線性方程出錯。
- 在第 (ii) 部分只寫正平方根，或沒有留意到解須滿足 $x < \frac{1}{2}B$。

STEP 之外

兩條方程都是伯努利方程（Bernoulli equation）$y' + P(x)y = Q(x)y^n$ 的例子：除以 y^n 並代入 $u = y^{1-n}$，總能得到一條線性方程。這裏第 (i) 部分是 $n = 2$，第 (ii) 部分是 $n = 3$。解可以在有限的 x 處「爆破」（blow up）（例如 $x \to \frac{1}{2}B$ 時），這是非線性方程（nonlinear equation）的典型特徵，在常微分方程（ODE）的入門課程中會學到；參見 Boyce and DiPrima, *Elementary Differential Equations*（常微分方程入門教科書）。

詳細解答

STEP I 1995，第 7 題

純數學 • 三角形的向量幾何

> **STEP I 1995，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/95-s1-q7

關鍵思路

把原點設在與 A、B、C 等距的一點，位置向量（position vector）$\mathbf{a}$、$\mathbf{b}$、$\mathbf{c}$ 的長度便全部相同。整道題就靠這一個事實，因為對於長度相等的向量，$(\mathbf{b}+\mathbf{c})\cdot(\mathbf{c}-\mathbf{b}) = |\mathbf{c}|^2 - |\mathbf{b}|^2 = 0$。一旦發現 Q 的位置向量是簡潔漂亮的 $\mathbf{a}+\mathbf{b}+\mathbf{c}$，兩部分都只需一兩行代數運算。

解答

選取原點

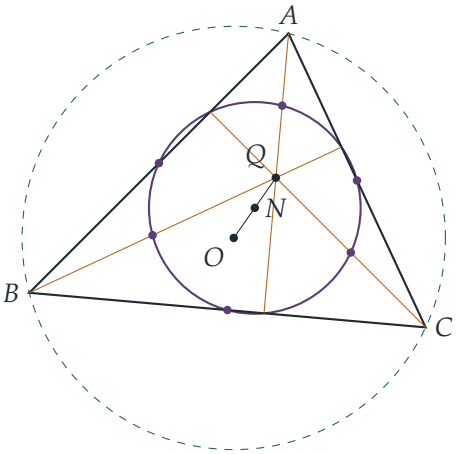
與 A 和 B 等距的點構成 AB 的垂直平分線（perpendicular bisector），這是一條垂直於 AB 的直線；B 和 C 的情況亦然。由於 A、B、C 不共線（collinear），AB 與 BC 不平行，所以它們的垂直平分線也不平行。因此這兩條直線恰好相交於一點，該點與 A 和 B 等距，也與 B 和 C 等距，故與三點皆等距。（這一點就是三角形的外心（circumcentre）。）

取這一點為 O，以 $\mathbf{a}$、$\mathbf{b}$、$\mathbf{c}$ 表示 A、B、C 的位置向量，並設 R 為這個共同的距離，於是

$$|\mathbf{a}| = |\mathbf{b}| = |\mathbf{c}| = R.$$

形心（centroid）的位置向量為 $\mathbf{g} = \frac{1}{3}(\mathbf{a}+\mathbf{b}+\mathbf{c})$。由於 $\overrightarrow{GQ} = 2\overrightarrow{OG}$ 即 $\mathbf{q}-\mathbf{g} = 2\mathbf{g}$，

$$\mathbf{q} = 3\mathbf{g} = \mathbf{a}+\mathbf{b}+\mathbf{c}, \qquad \mathbf{n} = \tfrac{1}{2}\mathbf{q} = \tfrac{1}{2}(\mathbf{a}+\mathbf{b}+\mathbf{c}).$$



圖中所示為以 O 為圓心、通過 A、B、C 的圓（虛線），相交於 Q 的三條高（altitude），以及以 N 為圓心、半徑為 $\frac{1}{2}R$、通過第 (ii) 部分六個中點的圓。

第 (i) 部分

我們有 $\overrightarrow{AQ} = \mathbf{q} - \mathbf{a} = \mathbf{b} + \mathbf{c}$ 及 $\overrightarrow{BC} = \mathbf{c} - \mathbf{b}$，所以

$$\overrightarrow{AQ} \cdot \overrightarrow{BC} = (\mathbf{c} + \mathbf{b}) \cdot (\mathbf{c} - \mathbf{b}) = |\mathbf{c}|^2 - |\mathbf{b}|^2 = R^2 - R^2 = 0,$$

其中用了 $\mathbf{b} \cdot \mathbf{c} = \mathbf{c} \cdot \mathbf{b}$ 來消去交叉項。因此 $\overrightarrow{AQ}$ 垂直於 $\overrightarrow{BC}$。

? **想一想** $\overrightarrow{AQ} \cdot \overrightarrow{BC} = 0$ 真的保證 Q 在由 A 作出的高上嗎？如果 $\overrightarrow{AQ}$ 是零向量（zero vector）又如何？

若 $Q \neq A$，直線 AQ 通過 A 且垂直於 BC，所以它正是由 A 作出的高，而 Q 在其上。若 $Q = A$（這恰好在 $\mathbf{b} + \mathbf{c} = \mathbf{0}$ 時發生，即 BC 為直徑而 A 處的角為直角），則 Q 顯然在由 A 作出的高上。無論哪種情況，Q 都在由 A 作出的高上。

以上論證並沒有對 A 另眼相看：表達式 $\mathbf{a} + \mathbf{b} + \mathbf{c}$ 是對稱的，所以同理可得

$$\overrightarrow{BQ} \cdot \overrightarrow{CA} = (\mathbf{c} + \mathbf{a}) \cdot (\mathbf{a} - \mathbf{c}) = |\mathbf{a}|^2 - |\mathbf{c}|^2 = 0,$$
$$\overrightarrow{CQ} \cdot \overrightarrow{AB} = (\mathbf{a} + \mathbf{b}) \cdot (\mathbf{b} - \mathbf{a}) = |\mathbf{b}|^2 - |\mathbf{a}|^2 = 0,$$

故 Q 也在由 B 和由 C 作出的高上。所以三條高都通過同一點 Q：它們共點（concurrent）。（Q 就是垂心（orthocentre）。）

第 (ii) 部分

AQ 的中點的位置向量為 $\frac{1}{2}(\mathbf{a} + \mathbf{q}) = \mathbf{a} + \frac{1}{2}(\mathbf{b} + \mathbf{c})$，所以它相對於 N 的位移為

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

BC 的中點的位置向量為 $\frac{1}{2}(\mathbf{b}+\mathbf{c})$，所以它相對於 N 的位移為

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

兩者的長度都是 $\frac{1}{2}|\mathbf{a}|=\frac{1}{2}R$。

根據同樣的對稱性，BQ 和 CA 的中點相對於 N 的位移為 $\pm\frac{1}{2}\mathbf{b}$，而 CQ 和 AB 的中點則為 $\pm\frac{1}{2}\mathbf{c}$。由於 $|\mathbf{a}|=|\mathbf{b}|=|\mathbf{c}|=R$，六個中點與 N 的距離全部是 $\frac{1}{2}R$，這正是我們要證明的。事實上我們還多證了一點：AQ 的中點和 BC 的中點位於這個圓某條直徑的兩端，其餘兩對中點亦然。

驗算

- 等邊三角形：此時 $\mathbf{a}+\mathbf{b}+\mathbf{c}=\mathbf{0}$，所以 O、G、Q 和 N 全部重合，而六個中點與中心的距離應該都是 $\frac{1}{2}R$。事實上，各邊中點與中心的距離為 $R\cos 60^\circ=\frac{1}{2}R$，而 AQ 等的中點則位於半徑的正中。
- A 處為直角：此時 $Q=A$，與上面論證所預測的一致，因為兩條直角邊本身就是通過 A 的高。
- 數值上，我們隨機取了 10^5 個三角形，直接計算外心，並確認了 $Q=\mathbf{a}+\mathbf{b}+\mathbf{c}$、三個點積（dot product）全部為零，以及六個中點與 N 的距離全部是 $\frac{1}{2}R$，最大相對誤差約為 10^{-13}。

常見錯誤

- 把 $\overrightarrow{GQ}=2\overrightarrow{OG}$ 讀成 $\mathbf{q}=2\mathbf{g}$。它說的是 $\mathbf{q}-\mathbf{g}=2\mathbf{g}$，所以 $\mathbf{q}=3\mathbf{g}$。
- 只證明了一個垂直關係便停筆。要證明共點，Q 必須在全部三條高上；應明確寫出同樣的計算對 B 和 C 也成立。
- 忘記說明原點為何可以這樣選：這一點之所以存在，是因為兩條不平行的邊的垂直平分線必定相交。

STEP 之外

由於 $\mathbf{q}=3\mathbf{g}$ 及 $\mathbf{n}=\frac{3}{2}\mathbf{g}$，$O$、$G$、$N$、$Q$ 四點全在同一條直線上，這條直線稱為歐拉線（*Euler line*），且 $OG:GQ=1:2$。第 (ii) 部分的圓就是九點圓（*nine-point circle*）：它還通過三條高的垂足（圖中可見），因為每個垂足對某條直徑（例如由 AQ 的中點至 BC 的中點的那一條）所張的角都是直角。巧妙地選取原點，讓一個對稱的表達式代勞，是向量幾何和線性代數（linear algebra）中反覆出現的技巧；Strang 的 *Introduction to Linear*

Algebra（線性代數入門教材）進一步發展了向量和點積的工具，而 Siklos 的 *Advanced Problems in Mathematics*（STEP 進階題集）則有更多同類風格的 STEP 式幾何題。

詳細解答

STEP I 1995，第 8 題

純數學 • 線性微分方程

> **STEP I 1995，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/95-s1-q8

關鍵思路

未知的是三個函數（function），但在每個固定的 x 處，它們只不過是三個數 $\mathrm{f}(x)$、$\mathrm{g}(x)$、$\mathrm{h}(x)$。把題目給出的三個解代入 $(*)$，便得到關於這三個數的三條線性方程（linear equation），我們把它們當作聯立方程（simultaneous equations）來解。至於第二段，$(*)$ 的左邊對 y 而言是線性的，所以 x、1 和 x^{-1} 的組合所得的結果，可由每一項各自產生的結果推出。

解答

求 f、g 和 h

把左邊 $y'' + \mathrm{f}(x)y' + \mathrm{g}(x)y$ 記作 Ly，於是 $(*)$ 就是 $Ly = \mathrm{h}(x)$。我們把各個解逐一代入。

- $y = 1$：這時 $y' = y'' = 0$，所以 $\mathrm{g}(x) = \mathrm{h}(x)$。
- $y = x$：這時 $y' = 1$，$y'' = 0$，所以 $\mathrm{f}(x) + x\,\mathrm{g}(x) = \mathrm{h}(x)$。
- $y = x^{-1}$：這時 $y' = -x^{-2}$，$y'' = 2x^{-3}$，所以

$$\frac{2}{x^3} - \frac{\mathrm{f}(x)}{x^2} + \frac{\mathrm{g}(x)}{x} = \mathrm{h}(x).$$

在第二條方程中使用 $\mathrm{h} = \mathrm{g}$，得 $\mathrm{f}(x) = (1-x)\,\mathrm{g}(x)$。把這兩個結果代入第三條方程，並在兩邊同乘 x^3（因為 $x > 0$，這樣做是容許的）：

$$\begin{aligned} 2 - x(1-x)\,\mathrm{g}(x) + x^2\,\mathrm{g}(x) &= x^3\,\mathrm{g}(x) \\ 2 &= \mathrm{g}(x)\left(x^3 - 2x^2 + x\right) = \mathrm{g}(x)\,x(1-x)^2. \end{aligned}$$

由於 $0 < x < 1$，因子 $x(1-x)^2$ 不等於零，所以可以用它來除。因此

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

在每個 x 處，我們都是一步一步解出這三條方程，過程中沒有任何自由選擇的餘地，所以符合條件的函數就只有這一組。限制 $0 < x < 1$ 使我們避開 $x = 0$ 和 $x = 1$，這些函數在該兩處會趨向無窮大。

所以這條方程是

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

$ax + b + c/x$ 何時是解？

算子（operator）L 是線性的：對常數 α, β，有 $L(\alpha u + \beta v) = \alpha\,Lu + \beta\,Lv$，因為求導（differentiation）是線性的，而 f、g 只是分別乘以 y' 和 y。由第一段可知

$$L(x) = L(1) = L(x^{-1}) = \mathrm{h}(x).$$

因此，對於 $y = ax + b + c/x$，

$$Ly = a\,L(x) + b\,L(1) + c\,L(x^{-1}) = (a+b+c)\,\mathrm{h}(x).$$

想一想 為甚麼我們可以斷定 $a + b + c = 1$，而不僅僅是「在某個 x 處 $a + b + c = 1$」？

我們需要 $(a+b+c)\,\mathrm{h}(x) = \mathrm{h}(x)$ 對所有 $0 < x < 1$ 成立，即 $(a+b+c-1)\,\mathrm{h}(x) = 0$。但在這個區間上 $\mathrm{h}(x) = 2/\big(x(1-x)^2\big) > 0$，所以可以用它來除。因此 $y = ax + b + c/x$ 是 $(*)$ 的解，當且僅當（if and only if）

$$a + b + c = 1.$$

驗算

- 取 $(a,b,c) = (1,0,0)$、$(0,1,0)$ 或 $(0,0,1)$，便得回題目給出的三個解，而每一個都滿足 $a + b + c = 1$。
- 直接驗證：取 $x = \frac{1}{2}$，得 $\mathrm{f} = 8$，$\mathrm{g} = \mathrm{h} = 16$。對於 $y = 1/x$，有 $y = 2$，$y' = -4$，$y'' = 16$，而 $16 - 32 + 32 = 16$，符合要求。
- 電腦在 $(0,1)$ 內的 49 個有理數點上精確地檢驗了全部三個解，並對 2000 組隨機的 (a,b,c,x) 確認了 $Ly = (a+b+c)\mathrm{h}$。

常見錯誤

- 對 $y = x^{-1}$ 寫成 $y'' = -2x^{-3}$：兩個負號相乘，結果應是 $+2x^{-3}$。
- 答成 $a + b + c = 0$。那是 $ax + b + c/x$ 滿足齊次（homogeneous）方程 $Ly = 0$ 的條件；這裏右邊是 h，不是 0。

STEP 之外

這正是線性微分方程（linear differential equation）的解的結構：$Ly = \mathrm{h}$ 的任意兩個解之差都是 $Ly = 0$ 的解。這裏 $x - 1$ 和 $x^{-1} - 1$ 是齊次方程的解，而對於二階方程，兩個線性無關（independent）的解便可給出全部的解。所以 $1 + \alpha(x - 1) + \beta(x^{-1} - 1)$ 就是通解（general solution），而這恰好就是滿足 $a + b + c = 1$ 的那一族解。朗斯基行列式（Wronskian）可以把「線性無關」的意思說得精確。Boyce 與 DiPrima 的 *Elementary Differential Equations*（初等微分方程教科書）涵蓋了以上所有內容。

詳細解答

STEP I 1995，第 9 題

力學 • 拋體軌道的包絡線

> **STEP I 1995，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/95-s1-q9

關鍵思路

固定一點 (x, y)，然後問：是否存在某個拋射角（angle of projection），使軌跡（trajectory）經過這一點？把軌跡方程用 $\tan\theta$ 而不是 θ 來表示，它就是一個關於 $\tan\theta$ 的二次方程（quadratic）。一點可以到達，當且僅當這個二次方程有實根（real root），所以整個問題歸結為判別式（discriminant）。

解答

模型

我們把粒子視為只受重力作用的質點（point mass）：沒有空氣阻力（air resistance），g 的大小和方向均不變，路徑上亦沒有任何障礙。這樣，粒子便在包含其初速度的鉛垂平面內運動；又因為只有與 O 的水平距離才有關係，由對稱性可取 $x \geq 0$，並設粒子朝 x 增加的方向拋出，與水平線所成的仰角為 θ，其中 $-\frac{\pi}{2} < \theta < \frac{\pi}{2}$。（垂直向上拋出，即 $\theta = \frac{\pi}{2}$，留待最後處理。）

軌跡

設 $u = \sqrt{2gh}$，故 $u^2 = 2gh$。加速度恆為向下的 g，經過時間 t 後

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

利用 $t = x/(u\cos\theta)$ 消去時間，並運用 $\sec^2\theta = 1 + \tan^2\theta$，得

$$y = x\tan\theta - \frac{gx^2}{2u^2}\,(1 + \tan^2\theta) = x\tan\theta - \frac{x^2}{4h}\,(1 + \tan^2\theta),$$

因為 $2u^2 = 4gh$。設 $T = \tan\theta$，並把全式乘以 $4h$，則軌跡經過 (x, y) 當且僅當

$$x^2T^2 - 4hx\,T + \left(x^2 + 4hy\right) = 0. \qquad (*)$$

當 θ 取遍 $\left(-\frac{\pi}{2}, \frac{\pi}{2}\right)$ 時，$T = \tan\theta$ 恰好取每個實數值一次，所以選擇一個角度等同於選擇一個實數 T。

想一想 為甚麼要用 $\tan\theta$ 而不用 θ 來表示軌跡？

若以 θ 表示，方程中混雜了 sin、cos 和 $\sec^2$，處理起來很不方便。利用 $\sec^2\theta = 1 + \tan^2\theta$，一切都變成單一變量 T 的多項式，而「是否存在這樣的角度？」就變成「這個二次方程有沒有實根？」，這個問題我們可以立即回答。

哪些點可以到達

取 $x > 0$。此時 $(*)$ 是一個真正關於 T 的二次方程，它有實根當且僅當其判別式非負：

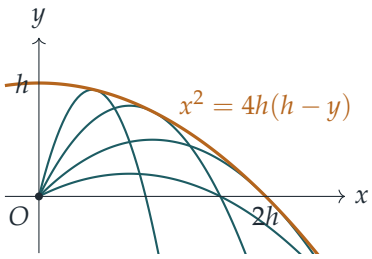
$$\begin{aligned}(4hx)^2 - 4x^2\left(x^2 + 4hy\right) \geq 0 &\iff 4x^2\left(4h^2 - x^2 - 4hy\right) \geq 0\\ &\iff x^2 \leq 4h^2 - 4hy = 4h(h - y),\end{aligned}$$

其中我們除以了 $4x^2 > 0$。因此，若 (x, y) 位於拋物線*上方*，即 $x^2 > 4h(h - y)$，或等價地 $y > h - \dfrac{x^2}{4h}$，則 $(*)$ 沒有實根，沒有任何拋射角能使粒子經過 (x, y)。

餘下要處理的是 $x = 0$，即通過 O 的鉛垂線。這條線上 $y \neq 0$ 的點，只有垂直拋射才能到達（任何其他角度都會令 $x > 0$ 對所有 $t > 0$ 成立）。以速率 u 垂直向上拋出時，粒子上升至高度 $u^2/(2g) = h$ 便不再上升，而這正是拋物線與 $x = 0$ 相交的一點。所以在這條線上，拋物線上方的點同樣無法到達。

因此，無論拋射角是多少，粒子永遠不能到達拋物線 $x^2 = 4h(h - y)$ 上方的任何一點。
■

這個界限是最佳的：當 $x^2 = 4h(h - y)$ 時，判別式為零，恰好有一個角度 $\tan\theta = 2h/x$，其路徑剛好與拋物線相切。這條拋物線是所有軌跡的*包絡線*（envelope），常稱為安全拋物線（parabola of safety）。



四條速率為 $\sqrt{2gh}$ 的軌跡（青綠色），每條都與邊界拋物線（琥珀色）相切一次。

簡化假設

- 粒子是一個質點（沒有大小，沒有自轉）。
- 忽略空氣阻力（以及重力以外的任何其他力）。
- 重力場是均勻的：g 在整個運動範圍內為常數，這亦意味着把地球表面視為平面，並忽略地球的自轉。
- 路徑不受阻擋（例如不被地面阻擋），所以粒子確實能夠到達所討論的點。

驗算

- 在 $x = 0$ 處，拋物線給出 $y = h = u^2/(2g)$，即垂直上拋的最大高度。在 $y = 0$ 上，它給出 $x = 2h = u^2/g$，即大家熟悉的水平地面上的最大射程（maximum range）（在 45° 時取得，此時 $\tan\theta = 2h/x = 1$，與預測相符）。
- 數值上，取 $g = 9.8$ 和 $h = 1.7$，我們對 399 個拋射角（由幾乎垂直向下到幾乎垂直向上）逐步模擬運動；在所有路徑上，高出拋物線的最大高度約為 -5×10^{-13}，即在捨入誤差範圍內為零；而在固定 x 之下對角度求 y 的最大值，每次都得回 $h - x^2/(4h)$。

常見錯誤

- 只考慮 θ 在 $\left[0, \frac{\pi}{2}\right]$ 範圍內的情況。位於 O 下方的點需要向下拋射（$T < 0$），論證必須涵蓋這些情況。
- 除以 x（或 x^2）時，沒有另外處理 $x = 0$ 的情況。
- 只證明了某一條特定軌跡位於拋物線之下；命題所說的是每一個角度，這正是我們固定一點而讓角度變化的原因。

STEP 之外

這裏的拋物線是單參數軌跡族的*包絡線*：即與族中每一條曲線都相切的曲線。對於曲線族 $F(x, y, T) = 0$，包絡線可由 $F = 0$ 與 $\partial F/\partial T = 0$ 消去 T 而求得；對於關於 T 的二次方程，這恰恰就是判別式等於零。包絡線還會在光學的焦散線（caustics）、一階微分方程理論（奇解，singular solutions），以及給定速率的所有拋射體的「安全區域」中再度出現。Kleppner 與 Kolenkow 所著的 *An Introduction to Mechanics*（力學入門）是進一步學習有空氣阻力和無空氣阻力的拋體運動的好選擇。

詳細解答

STEP I 1995，第 10 題

力學 • 彈性繩與反彈

STEP I 1995，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/95-s1-q10

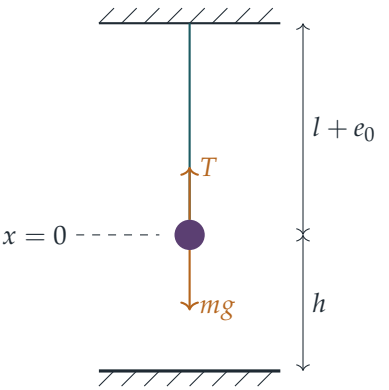
關鍵思路

若從平衡位置(equilibrium position)量起，重量與張力(tension)中屬於平衡狀態的那一部分恰好互相抵消，因此唯一未被平衡的力，就是與位移 x 成正比的胡克定律（Hooke's law）回復力（restoring force）。由此可知運動是簡諧運動（simple harmonic motion），而它的能量方程正是題目所求的速率公式。其後的一切只是逐項記帳：能量方程負責把球在 $x = 0$ 與表面之間的運動聯繫起來，而牛頓實驗定律（Newton's experimental law）則在每次反彈時把速度反向並按比例縮小。

解答

我們把球當作質點（particle），把繩當作輕而有彈性的繩，並假設它在拉長時服從胡克定律。我們忽略空氣阻力，並假設反彈在瞬間完成。全題記

$$\omega^2 = \frac{\lambda}{lm}.$$



平衡時的長度

設平衡時的伸長量（extension）為 e_0，則張力為 $T_0 = \lambda e_0/l$。球處於平衡時，沿鉛垂方向分解力，得 $T_0 = mg$，所以

$$\frac{\lambda e_0}{l} = mg \quad \Longrightarrow \quad e_0 = \frac{mgl}{\lambda},$$

而總長度為 $l + e_0 = l(1 + mg/\lambda)$，正如題目所求。

速率公式

設 x 為球在平衡位置以下的距離，並取向下為正。當 $x \geq 0$ 時，伸長量為 $e_0 + x > 0$，所以繩是拉緊的，張力為 $\lambda(e_0 + x)/l$。沿向下方向應用牛頓第二定律（Newton's second law），得

$$m\ddot{x} = mg - \frac{\lambda(e_0 + x)}{l} = -\frac{\lambda x}{l},$$

因為 $\lambda e_0/l = mg$。所以 $\ddot{x} = -\omega^2 x$：這是以 $x = 0$ 為中心的簡諧運動。

寫成 $\ddot{x} = v\,\dfrac{\mathrm{d}v}{\mathrm{d}x}$ 並積分，

$$\int v\,\mathrm{d}v = -\omega^2 \int x\,\mathrm{d}x \quad \Longrightarrow \quad v^2 = -\omega^2 x^2 + C.$$

在 $x = 0$ 時速率為 u_0，所以 $C = u_0^2$，而

$$v^2 + \frac{\lambda}{lm}x^2 = u_0^2,$$

正如題目所求。（這等價於能量守恆（conservation of energy）：重力勢能與彈性勢能合起來就是 $\frac{1}{2} \cdot \frac{\lambda}{l}x^2$ 加一個常數。）

第一次反彈

球能到達表面，當且僅當在 $x = h$ 時 $v^2 > 0$，即 $u_0^2 - \omega^2 h^2 > 0$。這正是題目給出的條件 $\lambda h^2 < lmu_0^2$。所以球以速率 w_0 撞擊表面，其中

$$w_0^2 = u_0^2 - \omega^2 h^2.$$

根據牛頓實驗定律，球離開表面時向上運動，速率為 ew_0。

球由 $x = h$ 上升到 $x = 0$ 的過程中，繩仍然處於拉長狀態，所以同一條運動方程依然成立，而能量積分為 $v^2 + \omega^2 x^2 = $ 常數。唯一不同的是這個常數的值。在 $x = h$ 時 $v^2 = e^2 w_0^2$，所以常數為 $e^2 w_0^2 + \omega^2 h^2$，而在 $x = 0$ 時

$$\begin{aligned} u_1^2 &= e^2 w_0^2 + \omega^2 h^2 \\ &= e^2\left(u_0^2 - \omega^2 h^2\right) + \omega^2 h^2 \\ &= e^2 u_0^2 + \frac{\lambda}{lm}h^2\left(1 - e^2\right), \end{aligned}$$

正如題目所求。這裏 u_1 的方向向上。

第二次反彈

> **想一想** 球經過 $x = 0$ 之後會升到平衡位置之上，繩可能會變鬆。球回到 $x = 0$ 時，速率還是 u_1 嗎？

是的。在 $x = 0$ 之上，球在重量和繩的張力作用下運動；若繩變鬆，則只受重力作用。這兩個力都是保守力（conservative force）：繩鬆弛時，它的彈性勢能就是零。因此在 $x = 0$ 之上的整段運動中，總機械能守恆。球落回 $x = 0$ 時，勢能與離開時相同，所以速率同樣是 u_1，只是方向變為向下。

所以第二次下降與第一次完全一樣，只是以 u_1 取代 u_0。我們仍須確定球能到達表面，而它確實可以，因為

$$u_1^2 - \omega^2 h^2 = e^2\left(u_0^2 - \omega^2 h^2\right) > 0.$$

把 u_0 換成 u_1，重複第一次反彈的論證，

$$\begin{aligned} u_2^2 &= e^2 u_1^2 + \omega^2 h^2\left(1 - e^2\right) \\ &= e^2\left[e^2 u_0^2 + \omega^2 h^2(1 - e^2)\right] + \omega^2 h^2(1 - e^2) \\ &= e^4 u_0^2 + \omega^2 h^2(1 - e^2)(1 + e^2) \\ &= e^4 u_0^2 + \frac{\lambda}{lm} h^2\left(1 - e^4\right), \end{aligned}$$

全題至此完成。

驗算

- 若 $e = 1$，則沒有能量損失，兩條公式都給出 $u_1 = u_2 = u_0$，與預期相符。
- 若 $e = 0$，球離開表面時處於靜止，於是 $u_1^2 = \omega^2 h^2$：球由振幅（amplitude）h 處從靜止開始上升，並以速率 ωh 通過中心，正如簡諧運動所預測。
- 有一個簡潔的方法看出規律：$u_n^2 - \omega^2 h^2 = e^{2n}\left(u_0^2 - \omega^2 h^2\right)$，因為每次反彈都把撞擊速率的平方乘以 e^2。
- 我們以數值方法逐步計算這個運動（龍格—庫塔法（Runge–Kutta），並容許繩變鬆）。例如，取 $m = 1$、$l = 1$、$\lambda = 20$、$g = 9.8$、$u_0 = 6$、$h = 0.3$ 及 $e = 0.9$，模擬結果為 $u_1 = 5.4316$ 和 $u_2 = 4.9233$，與公式吻合。在這個例子中，繩在兩次反彈之間確實會變鬆。

常見錯誤

- 從自然長度（natural length）而非平衡位置量度 x。這樣重量就不會抵消，公式中會多出一次項。
- 把恢復係數（coefficient of restitution）e 用於 $x = 0$ 處的速率，而不是用於 $x = h$ 處的撞擊速率。
- 假設繩在平衡位置之上仍然保持拉緊。其實不一定，但由於能量守恆，這一點並不影響球在 $x = 0$ 處的速率。

STEP 之外

重力在平衡點附近抵消，是一個普遍原理的最初例子：在平衡點附近作線性化（linearise），則小幅振動是簡諧運動，其 ω^2 等於「勁度」除以質量。彈跳的球是分段光滑振子（piecewise-smooth oscillator），又稱*碰撞*振子（impact oscillator）的一個例子。當表面本身也在運動時，這類系統的行為可以異常豐富（倍週期（period doubling），甚至混沌（chaos））。Kleppner 與 Kolenkow 的 *An Introduction to Mechanics*（力學導論）以大學一年級的程度講解振子和能量方法。

詳細解答

STEP I 1995，第 11 題

力學 • 圓柱之間的摩擦力

> **STEP I 1995，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/95-s1-q11

關鍵思路

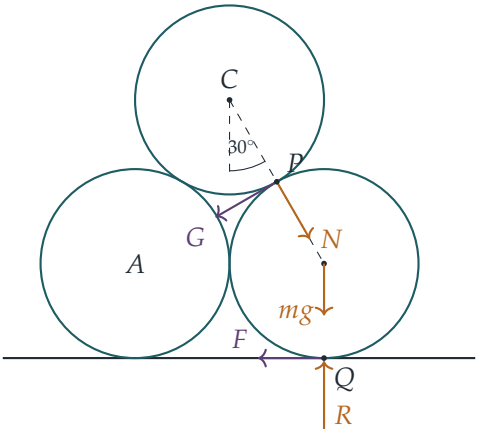
先看底下其中一個圓柱。它的重量和來自頂部圓柱的法向反作用力（normal reaction）都通過它的中心，所以對中心取力矩（moments）時只涉及兩個摩擦力（friction），而這兩個摩擦力的大小必定相等。接着只要沿水平方向分解（resolve）一次，就能定出圓柱之間接觸點上摩擦力與法向反作用力之比，餘下的工作就交給三個互相接觸的圓的幾何（30° 和 60° 的角）來完成。

解答

模型

我們把這些圓柱視為半徑為 r 的均勻剛體（rigid body）圓柱，並在垂直於各軸的橫截面上考慮問題，因此每個圓柱都是一個圓，其重量作用於圓心。每個接觸點的摩擦系數（coefficient of friction）同為 μ。把底下的兩個圓柱稱為 A（左）和 B（右），頂部的一個稱為 C。A 和 B 的中心相距 $2r$，高度同為 r；C 的中心與兩者各相距 $2r$，所以三個中心所成的三角形是等邊三角形，而由 C 的中心至 B 的中心的直線與鉛垂線成 30° 角。

我們暫且假設 A 和 B 之間互不施力；這一點留待最後再作討論。根據對稱性（symmetry），只需考慮 B 以及 C 右邊的接觸點。



作用於 B 的力有：

- 作用於其中心的重量 mg；
- 在與地面的接觸點 Q：向上的法向反作用力 R，以及沿水平方向指向 A 的摩擦力 F；
- 在與 C 的接觸點 P：沿連心線的法向反作用力 N（向下向外，與鉛垂線成 30° 角），以及沿公切線（common tangent）的摩擦力 G，我們設它的方向為向下並指向 A，即在水平線以下 30°。

假如其中一個摩擦力實際上是朝相反方向作用，算出來的值自然會是負數。

對 B 的中心取力矩

重量通過 B 的中心，N 亦然，因為兩圓之間的法向反作用力沿它們的連心線作用。兩個摩擦力都沿切線作用，與中心的垂直距離為 r。F 位於底部並指向 A，在圖中使 B 順時針轉動；G 位於 P 並沿切線向下指向 A，使它逆時針轉動。因此

$$Fr = Gr, \qquad \text{所以} \qquad F = G.$$

對 B 沿水平方向分解

N 有一個向外的水平分量 $N\sin 30^\circ = \frac{1}{2}N$。$F$ 以及 G 的水平分量 $G\cos 30^\circ = \frac{\sqrt{3}}{2}G$ 都指向內。所以

$$\tfrac{1}{2}N = F + \tfrac{\sqrt{3}}{2}G = G\left(1 + \tfrac{\sqrt{3}}{2}\right),$$

由此得

$$\frac{G}{N} = \frac{1}{2+\sqrt{3}} = \frac{2-\sqrt{3}}{(2+\sqrt{3})(2-\sqrt{3})} = 2-\sqrt{3}.$$

這是正數，所以 G 確實沿我們所設的方向作用，$F = G$ 亦然。

要令 P 處的接觸點不發生滑動，我們需要 $G \le \mu N$，即

$$\mu \ge 2 - \sqrt{3}.$$

與地面的接觸點

我們還須檢查與地面的接觸點不會要求更大的摩擦力。對 C 而言，來自 B 的力有：N，方向為向上偏左，與鉛垂線成 30° 角；以及 G，沿切線向上，與水平線成 30° 角。它們的鉛垂分量分別為 $\frac{\sqrt{3}}{2}N$ 和 $\frac{1}{2}G$，而根據對稱性，A 的貢獻亦相同。對 C 沿鉛垂方向分解，

$$2\left(\tfrac{\sqrt{3}}{2}N + \tfrac{1}{2}G\right) = mg \implies N\left(\sqrt{3} + 2 - \sqrt{3}\right) = mg \implies N = \tfrac{1}{2}mg.$$

對整個系統沿鉛垂方向分解，系統的重量為 $3mg$，根據對稱性由兩邊平均承擔，故 $R = \frac{3}{2}mg$。於是

$$\frac{F}{R} = \frac{(2-\sqrt{3})\cdot\frac{1}{2}mg}{\frac{3}{2}mg} = \frac{2-\sqrt{3}}{3} < 2 - \sqrt{3},$$

所以只要 $\mu \ge 2 - \sqrt{3}$，與地面的接觸點同樣不會滑動。圓柱之間的接觸點才是關鍵所在。

> **想一想** A 和 B 互相接觸。為甚麼我們可以忽略它們之間的任何力？

它們只能把對方推開，所以假設 A 以大小為 $S \ge 0$ 的力沿連心線把 B 向外推。這個力通過 B 的中心，所以力矩方程仍然給出 $F = G$，但水平方向的分解變成 $\frac{1}{2}N + S = G\left(1 + \frac{\sqrt{3}}{2}\right)$。與此同時，$C$ 的鉛垂方向方程仍然成立，把兩者聯立求解，可知 G/N 隨 S 增加而增加（當 $S = 0$ 時其值為 $2 - \sqrt{3}$）。這個問題是靜不定（statically indeterminate）的，因為 S 不能由這些方程確定，而當 $S = 0$ 時所需的摩擦力最小。所以摩擦系數的最小可能值是 $2 - \sqrt{3}$，正如題目所求。

驗算

- $2 - \sqrt{3} = \tan 15° \approx 0.268$。所以在圓柱之間的接觸點，總反作用力與法線恰好成 15° 角，正是幾何中 30° 角的一半。這是個令人欣喜的跡象，顯示代數運算無誤。
- 地面的比值 $(2 - \sqrt{3})/3 \approx 0.089$ 是圓柱之間比值的三分之一，這很合理：地面承受 $\frac{3}{2}mg$，是 N 的三倍，但摩擦力的大小相同。
- 作數值驗算，把數值代入 B 和 C 的所有力方程和力矩方程，並對未知的推力 $S \ge 0$ 求最小值，得出最小的 μ 為 0.267949…，與 $2 - \sqrt{3}$ 吻合。

常見錯誤

- 對地面接觸點或對 P 取力矩，而不是對中心取力矩。這樣做也行得通，但會引入 N，令代數運算較為繁複；對中心取力矩則可同時消去 mg 和 N。
- 只檢查地面。地面接觸點所需的摩擦力小得多，所以會得出錯誤的臨界值。
- 對 C 沿鉛垂方向分解時，忘記了根據牛頓第三定律（Newton's third law），作用於 C 的摩擦力是沿切線向上的。

STEP 之外

底下兩個圓柱之間的力 S 無法由平衡方程求出：這個結構是靜不定的，而題目很合理地要求在 S 的所有可能取值之中求最小的摩擦。在工程學上，這類系統要靠建立彈性形變的模型來解決，而同樣的問題亦見於顆粒物料的力學，例如疊起的喉管或沙堆。若想細讀有關靜力學、摩擦和剛體平衡的嚴謹論述，可參閱 Kleppner and Kolenkow, *An Introduction to Mechanics*（力學入門教科書）。

詳細解答

STEP I 1995，第 12 題

概率與統計 • 排列與計數

STEP I 1995，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/95-s1-q12

關鍵思路

n 名學生的每一種排列（arrangement）出現的機會均等，所以每個概率（probability）都等於有利排列的數目除以 $n!$。更好的做法是利用對稱性（symmetry）：我們只需關心曲棍球員佔據了哪些位置，而 $\binom{n}{r}$ 個由 r 個位置組成的集合，每一個出現的機會都相同。於是每一部分都變成一個關於選取位置的計數問題。至於最後一部分，題目中關於空隙的提示正是標準的「隔板法」（stars and bars）技巧：先把非曲棍球員排成一列，曲棍球員就必須分別放進他們之間不同的空隙。

解答

以下一律把 r 名曲棍球員記作 H，其餘 $n-r$ 名學生記作 N。把學生視為可區分的，則全部 $n!$ 種排序出現的機會均等。由於 H 所佔的每一個 r 個位置的集合，都對應相同數目的排序，即 $r!\,(n-r)!$ 種，所以 H 所佔位置的集合，等可能地是 $\binom{n}{r}$ 個可能集合中的任何一個。以下哪一種觀點較方便，我們就用哪一種。

第 (i) 部分

先填兩端的位置。左端的學生是曲棍球員的概率為 $\frac{r}{n}$。在此條件下，右端的學生是餘下 $n-1$ 名學生的其中一人，各人機會均等，而當中有 $r-1$ 人打曲棍球。所以

$$\mathrm{P}(\text{兩端都是曲棍球員}) = \frac{r}{n}\cdot\frac{r-1}{n-1} = \frac{r(r-1)}{n(n-1)}.$$

（要這個概率不為零，需要 $r \geq 2$，而題目已保證這一點。）

第 (ii) 部分

曲棍球員全部站在一起，當且僅當他們的位置構成連續 r 個位置的一段。這樣的一段可以由第 $1, 2, \ldots, n-r+1$ 個位置開始，所以在 $\binom{n}{r}$ 個位置集合之中，有 $n-r+1$ 個是有利的：

$$\mathrm{P}(\text{全部站在一起}) = \frac{n-r+1}{\binom{n}{r}} = \frac{r!\,(n-r+1)!}{n!}.$$

另一個等價的做法，是把 r 名曲棍球員黏合成一個「超級學生」：$n-r+1$ 個單位共有 $(n-r+1)!$ 種排序，而該段內部有 $r!$ 種排序，得出相同的答案。作為驗算，當 $r=n$ 時這個概率為 1，正如所料。

第 (iii) 部分

先讓 $n-r$ 名非曲棍球員排成一列。他們造成 $n-r+1$ 個空隙：第一人之前有一個，每對相鄰的人之間各有一個，最後一人之後有一個。



沒有兩名曲棍球員相鄰，當且僅當每個空隙最多放入一名曲棍球員。（若兩個 H 在同一個空隙，他們便會相鄰；若他們在不同的空隙，則至少有一個 N 把他們隔開。）因此，有利的排列對應於從 $n-r+1$ 個空隙中選出 r 個，每個選中的空隙放一個 H。

以位置來計數：一個有利的 H 位置集合，完全取決於用了哪 r 個空隙，所以共有 $\binom{n-r+1}{r}$ 個。因此

$$\mathrm{P}(\text{沒有兩人相鄰}) = \frac{\binom{n-r+1}{r}}{\binom{n}{r}} = \frac{(n-r+1)!\,(n-r)!}{(n-2r+1)!\,n!}.$$

（直接數學生的排序也得出同樣的結果：N 的排序有 $(n-r)!$ 種，然後把 r 個互不相同的 H 放進不同的空隙，有 $(n-r+1)(n-r)\cdots(n-2r+2) = \frac{(n-r+1)!}{(n-2r+1)!}$ 種方法，再全部除以 $n!$。）

想一想 在甚麼情況下，不可能把所有曲棍球員都分隔開？

我們需要 r 個不同的空隙，而空隙只有 $n-r+1$ 個。所以：

- 若 $r \le n-r+1$，即 $n \ge 2r-1$，概率就是上面那個非零的值；
- 若 $r > n-r+1$，即 $n \le 2r-2$（非曲棍球員太少，不足以把他們隔開），概率為 0。

在臨界情況 $n = 2r-1$，排列必然是 $HNHN\cdots NH$ 的模式，而公式給出 $\frac{r!\,(r-1)!}{(2r-1)!} = 1\big/\binom{2r-1}{r}$，即 H 恰好落在那 r 個相間位置的機會。

驗算

- 取 $n = 10$，$r = 3$。則 (i) 給出 $\frac{6}{90} = \frac{1}{15}$；(ii) 給出 $\frac{8}{120} = \frac{1}{15}$；(iii) 給出 $\binom{8}{3}/\binom{10}{3} = \frac{56}{120} = \frac{7}{15}$。三者都與電腦枚舉 120 個可能位置集合所得的結果吻合，而且各公式已用這個方法對所有 $2 \le r \le n \le 14$ 驗證過（當 $n \le 7$ 時，更與全部 $n!$ 種排序對照過）。
- 當 $r = 2$ 時，第 (ii) 和第 (iii) 部分是互補事件（complementary events）（兩名球員要麼相鄰，要麼不相鄰），而的確有 $\frac{2}{n} + \frac{(n-1)(n-2)}{n(n-1)} = \frac{2}{n} + \frac{n-2}{n} = 1$。

常見錯誤

- 只數出 $n - r - 1$ 個空隙（即嚴格位於非曲棍球員之間的那些），而忘記了隊列的兩端。
- 混用模型：分子數的是 H 的位置，分母卻用學生的全部排序（$n!$）。選定一種觀點，然後貫徹到底。
- 把答案非零的條件寫成 $n \ge 2r$；其實 $n = 2r - 1$ 的情況（交替排列，首尾都是 H）同樣可行。

STEP 之外

空隙論證就是「隔板法」的想法，是計數組合學（enumerative combinatorics）中的常用工具；同一個數目 $\binom{n-r+1}{r}$ 也回答了「$\{1, \ldots, n\}$ 有多少個 r 元子集不含兩個連續整數」這個問題，而把它對 r 求和就得到斐波那契數（Fibonacci numbers）。把人的排序換成位置的集合這一步，讓人初嘗利用對稱性來縮小樣本空間（sample space）的滋味，這種手法在大學概率論中經常用到。Blitzstein 與 Hwang 合著的 *Introduction to Probability*（概率論入門）中，關於計數和概率的樸素定義（naive definition of probability）的章節寫得十分出色。

詳細解答

STEP I 1995，第 13 題

概率與統計 • 泊松稀疏化

STEP I 1995，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/95-s1-q13

關鍵思路

第 (i) 部分是泊松分佈（Poisson distribution）的稀疏化（thinning）性質：先以癌細胞的真實數目為條件，則被標記的細胞數目服從二項分佈（binomial distribution）。再對未知的真實數目求和，結果恰好化簡成一個指數級數（exponential series）。至於第 (ii) 部分，「第二個標記落在第 k 塊玻片上」這個事件，取決於頭 $k-1$ 塊玻片上合共有多少個標記（零個或一個），而第 (i) 部分連同獨立泊松變量的可加性（additivity），正好告訴我們這個總數的分佈。

解答

全文記 $q = 1-p$ 及 $\lambda = \mu p$。我們假設不同玻片上的癌細胞數目互相獨立（independent）（這是對「隨機」一詞最自然的理解），並假設科學家每次能否察覺細胞亦互相獨立，且與細胞數目獨立。

第 (i) 部分

在一塊玻片上，設 N 為癌細胞的數目，M 為她標記的數目，故 $N \sim \mathrm{Po}(\mu)$。給定 $N = n$，這 n 個細胞各自獨立地以概率 p 被發現，所以 $M \mid N = n \sim \mathrm{B}(n,p)$。對於 $m = 0,1,2,\ldots$，由全概率公式（law of total probability）（只有 $n \geq m$ 的項有貢獻），

$$\begin{aligned}\mathrm{P}(M=m) &= \sum_{n=m}^{\infty} \mathrm{e}^{-\mu}\frac{\mu^n}{n!}\binom{n}{m}p^m q^{n-m} = \mathrm{e}^{-\mu}\frac{(\mu p)^m}{m!}\sum_{n=m}^{\infty}\frac{(\mu q)^{n-m}}{(n-m)!}\\ &= \mathrm{e}^{-\mu}\frac{(\mu p)^m}{m!}\mathrm{e}^{\mu q} = \mathrm{e}^{-\mu p}\frac{(\mu p)^m}{m!},\end{aligned}$$

其中先用 $\binom{n}{m}/n! = 1/\big(m!\,(n-m)!\big)$，再以 $j = n-m$ 套用指數級數，最後用 $-\mu+\mu q = -\mu p$。因此 $M \sim \mathrm{Po}(\mu p)$，正如所求。

第 (ii) 部分

設 M_i 為第 i 塊玻片上被標記的細胞數目。由第 (i) 部分及各玻片之間的獨立性，$M_1, M_2, \ldots$ 是互相獨立的 $\mathrm{Po}(\lambda)$ 變量。設 $S = M_1 + \cdots + M_{k-1}$ 為頭 $k-1$ 塊玻片上的標記總數；獨立泊松變量之和仍服從泊松分佈，其均值為各均值之和，所以 $S \sim \mathrm{Po}\big((k-1)\lambda\big)$，而且 S 與 M_k 獨立。

? **想一想** 要令第二個標記落在第 k 塊玻片上，究竟必須發生甚麼事？

在第 k 塊玻片之前，她最多只作了一個標記；而到第 k 塊玻片完結時，她至少已作了兩個標記。所以有兩種互斥（disjoint）的情況：



$S = 0$ 且 $M_k \geq 2$，或 $S = 1$ 且 $M_k \geq 1$

（當 $k = 1$ 時，前面沒有玻片，$S = 0$ 自動成立，以下公式依然適用。）利用 S 與 M_k 的獨立性：

$$\mathrm{P}(S=0,\ M_k \geq 2) = \mathrm{e}^{-(k-1)\lambda}\big(1-\mathrm{e}^{-\lambda}-\lambda\mathrm{e}^{-\lambda}\big),$$
$$\mathrm{P}(S=1,\ M_k \geq 1) = (k-1)\lambda\,\mathrm{e}^{-(k-1)\lambda}\big(1-\mathrm{e}^{-\lambda}\big).$$

兩式相加，

$$Q = \mathrm{e}^{-(k-1)\lambda}\Big\{\big(1+(k-1)\lambda\big)\big(1-\mathrm{e}^{-\lambda}\big)-\lambda\mathrm{e}^{-\lambda}\Big\}.$$

為了化成題目要求的形式，寫 $1+(k-1)\lambda = (1+k\lambda)-\lambda$：

$$\begin{aligned}\big(1+(k-1)\lambda\big)\big(1-\mathrm{e}^{-\lambda}\big)-\lambda\mathrm{e}^{-\lambda} &= (1+k\lambda)\big(1-\mathrm{e}^{-\lambda}\big)-\lambda+\lambda\mathrm{e}^{-\lambda}-\lambda\mathrm{e}^{-\lambda}\\ &= (1+k\lambda)\big(1-\mathrm{e}^{-\lambda}\big)-\lambda.\end{aligned}$$

代入 $\lambda = \mu p$ 即得

$$Q = \mathrm{e}^{-\mu p(k-1)}\Big\{(1+k\mu p)\big(1-\mathrm{e}^{-\mu p}\big)-\mu p\Big\},$$

這正是題目所述的結果。

驗算

▶ 當 $k = 1$ 時，公式給出 $1-\mathrm{e}^{-\lambda}-\lambda\mathrm{e}^{-\lambda} = \mathrm{P}(M_1 \geq 2)$，理應如此。

- 對 $k = 1, 2, \ldots$ 求和，各概率之和應為 1（她最終會標記第二個細胞的概率為 1）；取若干個 λ 的值作數值求和，結果確認了這一點。
- 以 $\mu = 1.6$、$p = 0.45$（故 $\lambda = 0.72$）進行模擬，在 200 000 次試驗中，第二個標記落在第 2 塊玻片上的約佔 26.0%；公式則給出 $Q \approx 0.2591$。

常見錯誤

- 遺漏了「第 k 塊玻片之前有一個標記，第 k 塊玻片上至少有一個」這種情況，或誤把 $S = 1, M_k = 0$ 也計算在內。
- 逐塊玻片地以「第一塊有標記的玻片」來論證，這樣做會變得十分繁複；頭 $k - 1$ 塊玻片上的總數 S 才是最有效率的切入對象。
- 沒有說明在哪裏用到獨立性（各玻片之間的獨立性，以及 S 與 M_k 之間的獨立性）。

STEP 之外

稀疏化是泊松過程（Poisson process）的其中一面：若事件按泊松過程發生，而每個事件各自獨立地以概率 p 被保留，則被保留的事件仍然構成一個泊松過程，其發生率（rate）乘以 p。這樣看來，第 (ii) 部分就是「等待第二個事件發生所需時間」的離散版本，而在連續時間下，這個等待時間服從伽瑪分佈（gamma distribution）。Grimmett 與 Welsh 所著的 *Probability: An Introduction*（概率論入門教材）詳細講述了泊松過程及這些等待時間的分佈。

詳細解答

STEP I 1995，第 14 題

概率與統計 • 比例的估計

STEP I 1995，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/95-s1-q14

關鍵思路

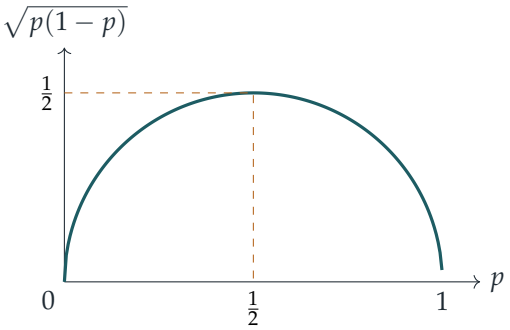
樣本比例（sample proportion）就是一個二項（binomial）計數除以 n。當 n 很大時，它近似服從正態分佈（normal distribution），平均值為 p，方差（variance）為 $p(1-p)/n$。麻煩在於，這個方差偏偏取決於我們不知道的那個 p。第 (i) 部分正好解決這個問題：它給出標準差（standard deviation）最壞情況下的值，所以按最壞情況選定的樣本大小，無論 p 是多少都適用。第 (iii) 部分接着問：如果我們對 p 已經有粗略的認識，情況會有甚麼不同。

解答

第 (i) 部分

用配方法（completing the square），設 $f(p) = p(1-p) = \frac{1}{4} - \left(p - \frac{1}{2}\right)^2$。所以在 $[0,1]$ 上 $f(p) \le \frac{1}{4}$，等號僅在 $p = \frac{1}{2}$ 時成立。平方根是遞增函數，因此

$$\max_{0 \le p \le 1} \sqrt{p(1-p)} = \sqrt{\tfrac{1}{4}} = \tfrac{1}{2}, \qquad \text{取於} p = \tfrac{1}{2}.$$



第 (ii) 部分

設 X 為樣本中女性的人數。每個被抽中的人是女性的概率為 p。我們假設總體（population）很大，因此這 n 次抽選實際上可視為互相獨立（從龐大的總體中不放回抽樣，與放回抽樣幾乎沒有分別）。於是

$$X \sim \mathrm{B}(n,p), \qquad \hat{p} = \frac{X}{n}, \qquad \mathrm{E}(\hat{p}) = p, \qquad \mathrm{Var}(\hat{p}) = \frac{np(1-p)}{n^2} = \frac{p(1-p)}{n}.$$

當 n 很大時，由中心極限定理（central limit theorem，即二項分佈的正態近似）可得

$$\hat{p} \approx \mathrm{N}\left(p,\, \frac{p(1-p)}{n}\right), \qquad \text{所以} \qquad Z = \frac{\hat{p}-p}{\sqrt{p(1-p)/n}} \approx \mathrm{N}(0,1).$$

我們要求 $\mathrm{P}\big(|\hat{p}-p| > 0.01\big) < 0.01$，即

$$\mathrm{P}\left(|Z| > \frac{0.01\sqrt{n}}{\sqrt{p(1-p)}}\right) < 0.01.$$

由對稱性，這 1% 平均分到兩邊尾部，每邊 0.5%。由於 $\Phi(2.576) = 0.995$，我們需要

$$\frac{0.01\sqrt{n}}{\sqrt{p(1-p)}} > 2.576 \quad \Longleftrightarrow \quad \sqrt{n} > 257.6\sqrt{p(1-p)}.$$

? **想一想** 我們不知道 p，那麼怎樣從一個含有它的不等式去選定 n 呢？

這正是第 (i) 部分派上用場的地方。無論 p 是多少，都有 $\sqrt{p(1-p)} \le \frac{1}{2}$。所以只要 $\sqrt{n} > 257.6 \times \frac{1}{2} = 128.8$，不等式便對任何 p 都成立。這個條件也是必要的，因為當 $p = \frac{1}{2}$ 時（而女性所佔的比例確實接近這個值），上界可以取到。因此我們需要

$$n > 128.8^2 = 16\,589.44,$$

所求的樣本大小為

$$\boxed{n = 16\,590} \qquad (\text{約}16\,600).$$

若採用更精確的百分點 $z = 2.5758$，則得 $n > 16\,587.2$。無論用哪個值，答案都是「約 16 600」，而第三位有效數字則視乎所用的統計表而定。

第 (iii) 部分

一般而言，所需的條件是

$$n > \left(\frac{2.576}{0.01}\right)^2 p(1-p) \approx 66\,360\, p(1-p).$$

16 590 這個數值來自最壞情況 $p = \frac{1}{2}$。如果我們知道 p 離 $\frac{1}{2}$ 相當遠，便可以為 $p(1-p)$ 採用一個較小的上界。

(a) 左撇子。大約 10% 的人是左撇子。如果可以穩妥地假設 $p \le 0.15$，那麼 $p(1-p) \le 0.15 \times 0.85 = 0.1275$，因為 $p(1-p)$ 在 $[0, \frac{1}{2}]$ 上遞增。由此得 $n > 8461$。如果恰好 $p = 0.1$，則 $n > 66\,360 \times 0.09 \approx 5970$，即大約 6000。無論哪種情況，所需樣本都遠小於一無所知時所需的 16 590。正態近似仍然相當準確，因為 np 數以百計。

(b) 百萬富翁。這時 p 非常小，也許介乎 0.001 至 0.01 之間。按字面計算，$p \le 0.01$ 給出 $p(1-p) \le 0.0099$ 及 $n > 657$。但這樣做並沒有抓住重點：

- 當 p 本身只有大約 0.001 時，0.01 的誤差毫無用處：估計值 $\hat{p} = 0$ 與真實值的差距已經在 0.01 之內。我們真正想要的是相對準確度，例如誤差不超過 1% 的 p。這樣便需要 $n > 2.576^2(1-p)/(p \times 0.01^2)$，當 $p = 0.001$ 時約為 6.6×10^7。這比先前的數字大得多。
- 當 p 極小而 n 只有數百時，np 小於 1。這時二項分佈非常偏斜，正態近似不再有效；以泊松近似（Poisson approximation）作為模型會更合適。
- 實際上，百萬富翁很難接觸得到，而且人們未必會如實說出自己的財富，所以能否做到隨機抽樣、受訪者是否誠實作答，兩者都成疑問。

因此，就左撇子而言，粗略知道 p 的大小可以讓我們採用較小的樣本。就百萬富翁而言，任何合理的準確度要求都需要大得多的樣本，或者要改用完全不同的抽樣方法。

驗算

- 當 $p = \frac{1}{2}$ 而 $n = 16\,590$ 時，正態近似給出 $\mathrm{P}(|\hat{p} - p| > 0.01) = 2\big(1 - \Phi(2.5762)\big) \approx 0.00999$，剛好略低於 1%。
- 把 $n = 16\,590$、$p = \frac{1}{2}$ 時的精確二項概率加起來，結果約為 0.0102。稍微超出的部分源於二項分佈的離散性（這裏的正態近似沒有作連續性修正（continuity correction））。當 $p = 0.3$ 時，精確值約為 0.005，綽綽有餘地符合要求，正如第 (i) 部分所預示。
- 當 $p = 0.1$ 而 $n = 6000$ 時，精確的二項概率約為 0.0098，印證了 (iii)(a) 中的數字。

常見錯誤

- 用了 1.96 或 2.326 而不是 2.576。這 1% 是分到兩邊尾部的，所以我們需要 $\Phi(z) = 0.995$。
- 把 n 留作以 p 表示的式子。題目要的是一個對任何 p 都適用的確定數值，這正是設置第 (i) 部分的原因。
- 把「$\sqrt{p(1-p)} \leq \frac{1}{2}$」理解成得出 $n > 128.8$。必須超過 128.8 的是 $\sqrt{n}$，而不是 n。

STEP 之外

民意調查中所報的「誤差範圍」（margin of error），背後就是這個標準計算。一項訪問約 1000 人的民調，在 95% 置信水平下的誤差大約為 ± 3 個百分點，用的正是第 (i) 部分的最壞情況論證。到了大學，這會引伸出置信區間（confidence interval），以及如何估計罕見比例的問題；在那種情況下，泊松方法和精確二項方法會取代正態近似。Blitzstein 與 Hwang 合著的 *Introduction to Probability*（概率論入門）對中心極限定理和這些近似方法有十分清晰易讀的講解。

詳細解答

STEP II 1995，第 1 題

純數學 • 等比級數的求導

> **STEP II 1995，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/95-s2-q1

關鍵思路

有限等比級數（geometric series）有閉合形式（closed form），而兩個 x 的函數之間的恆等式（identity）可以求導，從而得出新的恆等式。對 $\sum x^k$ 求導會帶出一個因子 k，這正好把等比級數變成 $\sum kx^{k-1}$；再代入 $x = -1$，便得到正負相間的符號。至於平方的情況，最簡潔的做法是讓答案的形式告訴我們 A 和 B 必須是甚麼，然後用歸納法（induction）證明。

解答

第 (i) 部分

設 $S = 1 + x + x^2 + \cdots + x^n$。則

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

因為 $x, x^2, \ldots, x^n$ 每一個冪都以正號和負號各出現一次，互相抵消。若 $x \neq 1$，則 $1 - x \neq 0$，所以可以除以它：

$$1 + x + x^2 + \cdots + x^n = \frac{1 - x^{n+1}}{1 - x},$$

這正是所求。

第 (ii) 部分

(i) 的兩邊都是在 $x \neq 1$ 上可導（differentiable）的 x 的函數，而且在該處相等，所以它們的導數也相等。左邊是

$$\frac{\mathrm{d}}{\mathrm{d}x}\left(1 + x + \cdots + x^n\right) = 1 + 2x + 3x^2 + \cdots + nx^{n-1}.$$

右邊則用商法則（quotient rule），得

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

現在取 $x = -1$，這是容許的，因為 $-1 \neq 1$。左邊變成 $1 - 2 + 3 - 4 + \cdots + (-1)^{n-1}n$，因為 kx^{k-1} 這一項變成 $(-1)^{k-1}k$。右邊變成

$$\frac{1-(n+1)(-1)^n + n(-1)^{n+1}}{4} = \frac{1-(-1)^n(2n+1)}{4},$$

當中用到 $(-1)^{n+1} = -(-1)^n$。

- 若 n 是偶數，則 $(-1)^n = 1$，其值為 $\dfrac{1-2n-1}{4} = -\dfrac{n}{2}$。
- 若 n 是奇數，則 $(-1)^n = -1$，其值為 $\dfrac{1+2n+1}{4} = \dfrac{n+1}{2}$。

這就是我們要證明的結果。

第 (iii) 部分

設 $T_n = 1^2 - 2^2 + 3^2 - \cdots + (-1)^{n-1}n^2$，並假設 $T_n = (-1)^{n-1}f(n)$，其中 $f(n) = An^2 + Bn$。

> **想一想** 在未知道公式是否成立之前，我們怎樣才能求出 A 和 B？

利用我們對 T_n 唯一確知的事實：相鄰兩個和之差就是最後一項，即 $T_n - T_{n-1} = (-1)^{n-1}n^2$。由於 $T_{n-1} = (-1)^{n-2}f(n-1) = -(-1)^{n-1}f(n-1)$，上式可寫成

$$(-1)^{n-1}\big(f(n) + f(n-1)\big) = (-1)^{n-1}n^2, \qquad \text{所以} \qquad f(n) + f(n-1) = n^2.$$

代入 $f(n) = An^2 + Bn$，得

$$\begin{aligned} f(n) + f(n-1) &= An^2 + Bn + A(n-1)^2 + B(n-1) \\ &= 2An^2 + (2B - 2A)n + (A - B). \end{aligned}$$

要使這個式子對每一個 n 都等於 n^2，我們需要 $2A = 1$ 和 $B = A$，即

$$A = B = \tfrac{1}{2}, \qquad T_n = (-1)^{n-1}\,\frac{n(n+1)}{2}.$$

（這些值是唯一可能的：$n=1$ 給出 $A+B=1$，而 $n=2$ 給出 $-(4A+2B)=-3$，其唯一解是 $A=B=\frac{1}{2}$。）

用歸納法證明。當 $n=1$ 時，$T_1=1=(-1)^0\cdot\frac{1\cdot 2}{2}$。假設對某個 $n\geq 2$ 有 $T_{n-1}=(-1)^{n-2}\frac{(n-1)n}{2}$。則

$$\begin{aligned}T_n=T_{n-1}+(-1)^{n-1}n^2&=(-1)^{n-1}\left(n^2-\frac{n(n-1)}{2}\right)\\&=(-1)^{n-1}\,\frac{n(n+1)}{2},\end{aligned}$$

這正是 n 時的公式。因此公式對所有 $n\geq 1$ 都成立，其中 $A=B=\frac{1}{2}$。

另一做法是把各項配對：$(2m-1)^2-(2m)^2=-(4m-1)$。當 $n=2m$ 時，由此得 $T_n=-\sum_{j=1}^{m}(4j-1)=-m(2m+1)=-\frac{n(n+1)}{2}$；當 $n=2m+1$ 時，再加上 $(2m+1)^2$，便得 $(2m+1)(m+1)=\frac{n(n+1)}{2}$。（也可以把 (ii) 的恆等式乘以 x 後再求導一次，但代數運算會繁複得多。）

驗算

當 $n=4$ 時：$1-2+3-4=-2=-\frac{4}{2}$，以及 $1-4+9-16=-10=-\frac{4\cdot 5}{2}$。當 $n=5$ 時：$1-2+3-4+5=3=\frac{6}{2}$，以及 $-10+25=15=\frac{5\cdot 6}{2}$。用電腦對 n 直至 500 驗算全部三個恆等式（例如 $T_{10}=-55$ 和 $T_{11}=66$），結果完全吻合。

還可留意 $|T_n|=1+2+\cdots+n$：平方的交錯和（alternating sum），其大小竟與首 n 個正整數的普通和相同。這個巧合令人意外，而配對的論證正好解釋了它。

常見錯誤

- 忘記了 $x=-1$ 之所以是合法的取值，全因 (i) 的恆等式對所有 $x\neq 1$ 成立；在 $x=1$ 時，右邊沒有定義。
- 處理 $(-1)^{n+1}$ 時弄錯符號：它等於 $-(-1)^n$，而不是 $(-1)^n$。
- 由 $n=1$ 和 $n=2$ 求出 A 和 B 後便停下來。這只說明了這些是唯一可能的值；要靠歸納法（或配對）才能說明公式確實對每一個 n 都成立。

STEP 之外

對 $\sum x^k$ 逐項求導，是一個威力強大的想法的有限版本：冪級數（power series）在其收斂半徑（radius of convergence）之內可以逐項求導，由此得出當 $|x|<1$ 時 $\sum kx^{k-1}=(1-x)^{-2}$。在 $x=-1$ 時，無窮級數 $1-2+3-4+\cdots$ 發散（diverges），但公式卻暗示

其值為 $\frac{1}{4}$；事實上，(ii) 的公式可以寫成 $\frac{1}{4} - (-1)^n \frac{2n+1}{4}$，所以部分和（partial sums）以 $\frac{1}{4}$ 為中心對稱地擺動。為這類發散級數賦予一個值，正是阿貝爾求和法（Abel summation）背後的想法。Abbott 的 *Understanding Analysis*（實分析入門教材）對冪級數的逐項求導有嚴謹的論述。

詳細解答

STEP II 1995，第 2 題

純數學 • 以遞推法計算着色數目

STEP II 1995，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/95-s2-q2

本題在 *Beyond STEP* 第 9.3 節爲籬笆樁着色中有完整的引導式講解：https://beyondstep.keithmcnulty.org/zh-hant/ch09/fence-posts.html

關鍵思路

把以紅色開頭的塗色方法，按最後一根柱子的顏色分類。直線上的每一種塗色方法都可以逐根柱子加上去而得出，而每根新柱子恰好有兩種可用的顏色，所以總數很容易計算；要用遞推關係（recurrence）處理的，是「以紅色結尾」與「不以紅色結尾」之間的劃分。至於圓形，不過是一條兩端顏色也必須不同的直線而已。

解答

為何 $r_{n+1} = s_n$

取 $n+1$ 根柱子的一種合規塗色方法，其中第一根和最後一根柱子都是紅色。第 n 根柱子與第 $n+1$ 根柱子相鄰，而後者是紅色，所以第 n 根柱子不是紅色。因此，刪去第 $n+1$ 根柱子後，餘下的是 n 根柱子的一種合規塗色方法，其中第一根柱子是紅色而最後一根不是紅色：這正是 s_n 種塗色方法之一。

反過來，給定 n 根柱子的 s_n 種塗色方法之一，由於第 n 根柱子不是紅色，我們可以在末端加上一根紅色柱子，從而得到一種計入 r_{n+1} 的塗色方法。這兩個操作互為逆操作，所以它們給出一個一一對應（one-to-one correspondence），於是

$$r_{n+1} = s_n.$$

求 $r_n + s_n$

每一種第一根柱子為紅色的合規塗色方法，最後一根柱子不是紅色就是另一種顏色，所以 $r_n + s_n$ 就是第一根柱子為紅色的所有合規塗色方法的總數。按次序為柱子塗色。第一根是紅色（一種選擇），其後每根柱子可以塗上除前一根柱子的顏色以外的任何顏色，因此不論之前怎樣塗，每次都有 2 種選擇。故此

$$r_n + s_n = 2^{n-1}.$$

求 $r_{n+1} + r_n$

代入 $s_n = r_{n+1}$，得

$$r_{n+1} + r_n = 2^{n-1} \qquad (n \geq 1).$$

歸納法

設 $P(n)$ 為命題 $r_n = \frac{1}{3}\left(2^{n-1} + 2(-1)^{n-1}\right)$。（這裏用的是數學歸納法（induction）。）

基礎情況（*base case*）。只有一根柱子並塗上紅色時，這根柱子既是第一根也是最後一根，所以 $r_1 = 1$。公式給出 $\frac{1}{3}(1+2) = 1$，所以 $P(1)$ 成立。

歸納步驟（*inductive step*）。假設 $P(n)$ 對某個 $n \geq 1$ 成立。由上一個結果，

$$\begin{aligned} r_{n+1} &= 2^{n-1} - r_n \\ &= \frac{3 \cdot 2^{n-1} - 2^{n-1} - 2(-1)^{n-1}}{3} \\ &= \frac{2^n + 2(-1)^n}{3}, \end{aligned}$$

其中用到 $2 \cdot 2^{n-1} = 2^n$ 和 $-(-1)^{n-1} = (-1)^n$。這正是 $P(n+1)$。根據歸納法，$P(n)$ 對所有 $n \geq 1$ 成立。

> ? **想一想** 為甚麼不由 $n = 2$ 開始歸納，或者使用兩個基礎情況？

遞推關係 $r_{n+1} = 2^{n-1} - r_n$ 只用前一項來表示每一項，所以一個基礎情況已經足夠。若是涉及前兩項的遞推關係，就需要兩個。

圍成圓形的柱子

把柱子沿圓周編號為 $1, 2, \ldots, n$，於是第 n 根柱子現在也與第 1 根柱子相鄰。一種合規的圓形塗色方法，正正就是直線 $1, 2, \ldots, n$ 的一種合規塗色方法，並且額外要求第 n 根柱子與第 1 根柱子顏色不同。

> ? **想一想** 如果第一根柱子不固定為紅色，我們怎樣數算塗色方法？

利用對稱性（symmetry）。把兩種顏色的名稱對調，會把合規塗色方法變成合規塗色方法，所以第 1 根柱子不論選哪一種顏色，塗色方法的數目都相同。當第 1 根柱子為紅色時，最後一根柱子與第一根顏色不同的直線塗色方法，正是 s_n 所計算的那些。所以圓形塗色方法的數目是

$$\begin{aligned} 3s_n &= 3\left(2^{n-1} - r_n\right) \\ &= 3 \cdot 2^{n-1} - 2^{n-1} - 2(-1)^{n-1} \\ &= 2^n + 2(-1)^n. \end{aligned}$$

這裏的柱子是可以區分的（它們豎立在固定的位置），所以兩種只相差一個旋轉的塗色方法會分開計算。

驗算

- 用手計算小的情況：$r_2 = 0$（兩根相鄰的柱子不能同為紅色）和 $r_3 = 2$（RWR、RBR），與 $\frac{1}{3}(2-2) = 0$ 和 $\frac{1}{3}(4+2) = 2$ 吻合。
- 三根柱子圍成三角形時必須用齊三種顏色，共有 $3! = 6$ 種方法，而 $2^3 + 2(-1)^3 = 6$。四根柱子圍成正方形則有 $16 + 2 = 18$ 種。
- 以窮舉法（brute force）數算 $n \leq 12$ 的所有塗色方法，結果與上面每一條公式都一致；例如 $r_{10} = 170$、$s_{10} = 342$，而 10 根柱子的圓形塗色方法有 1026 種。

常見錯誤

- 以 r_n 或 s_n 作為圓形的答案，忘記了第 1 根柱子不一定是紅色；因子 3 正是用來補回另外兩種顏色。
- 除以 n 來消去旋轉。柱子的位置是固定的，所以旋轉後的塗色方法是不同的塗法。

STEP 之外

圓形的答案 $2^n + 2(-1)^n$ 是圈圖（cycle graph）C_n 的色多項式（chromatic polynomial）在 $k = 3$ 時的值；一般而言，使用 k 種顏色時，這個多項式是 $(k-1)^n + (-1)^n(k-1)$，而同樣的「先排成直線，再把兩端接合」論證就可以證明它。色多項式以及像 $r_{n+1} + r_n = 2^{n-1}$ 這樣的遞推關係，都是組合數學（combinatorics）或圖論（graph theory）入門課程的標準內容。想做更多同類題目，可參看 Siklos 的 *Advanced Problems in Mathematics*（進階數學難題集）。

詳細解答

STEP II 1995，第 3 題

純數學 • 邏輯推理與計數

> **STEP II 1995，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/95-s2-q3

關鍵思路

這其實是一道喬裝了的邏輯謎題，而入手的方法是計數（counting）。每個賽段都派出 $5+4+3+2+1=15$ 分，所以五位車手的得分之和是一個已知的總數。Arouet 的得分非常接近滿分，幾乎足以確定他所有賽段的成績。其後，Eluard 的賽果把 Eluard 的得分從上下兩面夾住，而 Camus「四次名次相同」這個條件，只剩下幾種情況需要檢查。

解答

我們以 A, B, C, D, E 表示各車手的總得分。已知 $A > B > C > D > E$ 及 $A = 24$。

總分與 Arouet 的各賽段成績

每個賽段派出 15 分，所以五個賽段合共

$$A+B+C+D+E = 75, \qquad \text{因此} \qquad B+C+D+E = 51.$$

Arouet 最多可得 $5 \times 5 = 25$ 分。得到 24 分即是剛好失去一分，所以 Arouet 在四個賽段得第一名，在一個賽段得第二名。Eluard 贏了第 5 賽段，所以 Arouet 的第二名必定出現在第 5 賽段，而 Arouet 贏了第 1–4 賽段。

為何 Eluard 得 11 分

下界（*lower bound*）。Eluard 在第 4 賽段得 3 分，在第 5 賽段得 5 分，而在其餘三個賽段每段最少得 1 分。所以 $E \geq 3+5+3 = 11$。

上界（*upper bound*）。各人得分是互不相同的整數，且 $B > C > D > E$，所以 $D \geq$

$E+1$、$C \geq E+2$ 及 $B \geq E+3$。因此

$$51 = B+C+D+E \geq 4E+6, \qquad \text{所以} \qquad E \leq \tfrac{45}{4} < 12.$$

結合兩個界，得 $E = 11$。下界中等號成立，亦告訴我們 Eluard 在第 1、2 及 3 賽段每段都是最後一名。

留給 Barthes、Camus 和 Diderot 的名次

我們現在知道 Arouet 和 Eluard 在每個賽段的名次，所以餘下的名次由 Barthes、Camus 和 Diderot 三人瓜分：

賽段	Arouet、Eluard	留給 B、C、D 的名次	
1, 2, 3	第 1、第 5	第 2、第 3、第 4	（4、3、2 分）
4	第 1、第 3	第 2、第 4、第 5	（4、2、1 分）
5	第 2、第 1	第 3、第 4、第 5	（3、2、1 分）

此外 $B+C+D = 51-11 = 40$，且 $B > C > D > 11$，所以 $D \geq 12$ 及 $C \geq 13$。

想一想 Camus 有四次取得相同的名次。根據上表，那可以是哪些名次？

答案是第 2、第 3 或第 4 名。Camus 只能在第 4 和第 5 賽段取得第五名，那只有兩個賽段，不是四個。我們逐一考慮這三種可能。

Camus 四次得第 2 名。第二名只在第 1–4 賽段可供取得，所以 Camus 在這四個賽段全部得第二名。於是 Barthes 在第 1–3 賽段每段最多得 3 分，在第 4 賽段最多得 2 分，在第 5 賽段最多得 3 分。所以 $B \leq 9+2+3 = 14$。但 $C \geq 4 \times 4 = 16 > B$，出現矛盾（contradiction）。

Camus 四次得第 4 名。第四名在每個賽段都值 2 分，所以 Camus 的最佳總分是 $4 \times 2 + 4 = 12$（餘下的一個賽段得第二名）。這與 $C \geq 13$ 矛盾。

Camus 四次得第 3 名。第三名在第 1、2、3 及 5 賽段可供取得，每次值 3 分，所以這就是 Camus 的四次第三名。在第 4 賽段 Camus 得 $x \in \{4,2,1\}$ 分，而 $C = 12+x$。

- $x = 4$：$C = 16$，所以 $B \geq 17$ 及 $D = 40-B-C \leq 7$。這與 $D \geq 12$ 矛盾。
- $x = 2$：$C = 14$，所以 $B \geq 15$ 及 $D \leq 11$。同樣出現矛盾。
- $x = 1$：$C = 13$，所以 $B+D = 27$，其中 $B \geq 14$ 及 $12 \leq D \leq 12$。因此 $D = 12$ 及 $B = 15$。

所以各人得分為

$$\boxed{B = 15, \quad C = 13, \quad D = 12} \qquad (\text{其中} A = 24,\ E = 11).$$

Barthes 在最後賽段的名次

> **想一想** 我們知道總分，但不知道每個賽段的賽果。現有的資料足以確定第 5 賽段的結果嗎？

足夠，方法是用奇偶性論證（parity argument）。Camus 在第 1、2、3 及 5 賽段得第三名，在第 4 賽段得第五名，於是 Barthes 和 Diderot 在第 1–4 賽段每段都瓜分值 4 分和 2 分的名次。在第 5 賽段，他們瓜分值 2 分和 1 分的名次（第 4 及第 5 名）。假設 Diderot 在第 1–4 賽段中有 k 段取得值 4 分的名次，並在第 5 賽段得 $s \in \{1, 2\}$ 分。那麼

$$D = 4 \times 2 + 2k + s = 8 + 2k + s = 12.$$

所以 $s = 4 - 2k$ 是偶數，由此必有 $s = 2$（及 $k = 1$）。Diderot 在最後賽段得第四名，所以 **Barthes 在最後賽段得最後一名（第五名）**。

這樣的賽事確實存在。例如，取第 1 賽段：A、D、C、B、E；第 2 及第 3 賽段：A、B、C、D、E；第 4 賽段：A、B、E、D、C；第 5 賽段：E、A、C、D、B。這樣得出 $24, 15, 13, 12, 11$，符合要求。

驗算

- 總分吻合：$24 + 15 + 13 + 12 + 11 = 75 = 5 \times 15$。
- 在上述例子的賽事中，Camus 在四個賽段得第三名，而 Eluard 在第 4 賽段得第三名，在第 5 賽段得第一名。
- 用電腦搜尋全部 120^5 組可能的賽段賽果，只保留符合所有條件的，結果只找到得分 $(24, 15, 13, 12, 11)$。在每一種情況下，Barthes 在第 5 賽段都是第五名。（符合條件的賽事恰好有四個，視乎 Diderot 在第 1–4 賽段中哪一段得第二名而定。）

常見錯誤

- 只證明了 $E \geq 11$。題目問的是為何 Eluard 必定得 11 分，所以由 $B + C + D + E = 51$ 及得分互不相同而得出的上界是必不可少的。
- 假設 Camus 是在四個連續賽段取得相同名次，或者遺漏了第四名的情況。
- 求出了得分，卻沒有證明最後賽段的名次是必然的。只舉出一個賽事例子並不足夠；奇偶性論證說明了在每一個可能的賽事中，Barthes 在第 5 賽段都是最後一名。

STEP 之外

這個論證有三步：守恆定律（conservation law，即總分固定）、極端推理（extremal reasoning，24 分幾乎是最高分），以及奇偶性論證。這三者在組合數學（combinatorics）中一再出現。從上下兩面夾住一個量，直至它的值被迫確定下來，也是分析學（analysis）和數學奧林匹克解題的基本技巧。想做更多同類的謎題，可試試 Siklos 的 *Advanced Problems in Mathematics*。Engel 的 *Problem-Solving Strategies* 則有整章篇幅討論不變量原理（invariance principle）和極端原理（extremal principle）。

詳細解答

STEP II 1995，第 4 題

純數學 • 沃利斯歸約公式

> **STEP II 1995，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/95-s2-q4

本題在 *Beyond STEP* 第 4.4 節沃利斯、*Beta* 與二分之一的階乘中有完整的引導式講解：https://beyondstep.keithmcnulty.org/zh-hant/ch04/wallis-beta.html

關鍵思路

利用分部積分法（integration by parts），可把 $\cos^2 t$ 化為 $1-\sin^2 t$，從而得到一條把 u_n 與 u_{n-2} 聯繫起來的歸約公式（reduction formula）。把這條公式乘以 u_{n-1}，可見 nu_nu_{n-1} 不隨 n 增大而改變，所以它等於 $n=1$ 時的值。餘下的就是夾逼（squeeze）：由於各 u_n 遞減，乘積 u_nu_{n-1} 介乎 u_n^2 與 u_{n-1}^2 之間。

解答

歸約公式

設 $n \geq 2$。我們把積分寫成 $u_n = \int_0^{\pi/2} \sin t \, \sin^{n-1} t \, \mathrm{d}t$，然後分部積分：對 $\sin^{n-1} t$ 求導，並把 $\sin t$ 積分為 $-\cos t$：

$$\begin{aligned} u_n &= \Big[-\cos t \, \sin^{n-1} t\Big]_0^{\pi/2} + (n-1)\int_0^{\pi/2} \cos^2 t \, \sin^{n-2} t \, \mathrm{d}t \\ &= 0 + (n-1)\int_0^{\pi/2} \big(1-\sin^2 t\big)\sin^{n-2} t \, \mathrm{d}t \\ &= (n-1)\big(u_{n-2} - u_n\big). \end{aligned}$$

邊界項等於零，因為在上限有 $\cos\frac{\pi}{2} = 0$，在下限有 $\sin 0 = 0$（這裏用到 $n-1 \geq 1$）。把含 u_n 的項合併，得

$$nu_n = (n-1)u_{n-2} \qquad (n \geq 2).$$

兩邊同乘以 u_{n-1}，得

$$nu_nu_{n-1} = (n-1)u_{n-1}u_{n-2} \qquad (n \geq 2),$$

這正是所需的結果。

恆定的乘積

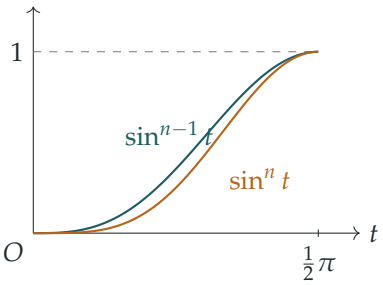
對 $n \geq 1$，記 $p_n = nu_nu_{n-1}$。上面的結果說明，對每個 $n \geq 2$ 都有 $p_n = p_{n-1}$，所以對每個 $n \geq 1$ 都有 $p_n = p_1$。現在

$$u_0 = \int_0^{\pi/2} 1\,\mathrm{d}t = \tfrac{1}{2}\pi, \qquad u_1 = \int_0^{\pi/2} \sin t\,\mathrm{d}t = \Big[-\cos t\Big]_0^{\pi/2} = 1,$$

所以 $p_1 = 1 \cdot u_1u_0 = \frac{1}{2}\pi$。因此

$$nu_nu_{n-1} = \tfrac{1}{2}\pi \qquad \text{對所有} n \geq 1.$$

草圖與 $0 < u_n < u_{n-1}$



兩條曲線在 $t = 0$ 時都由 0 出發，一直遞增，並在 $t = \frac{1}{2}\pi$ 處達到 1（草圖所示為 $n = 4$ 的情況；當 $n = 1$ 時，上方的曲線是直線 $y = 1$）。對 $0 < t < \frac{1}{2}\pi$，我們有 $0 < \sin t < 1$，所以

$$0 < \sin^n t = \sin t \cdot \sin^{n-1} t < \sin^{n-1} t.$$

因此，除了兩個端點之外，$\sin^n t$ 的圖像嚴格位於 $\sin^{n-1} t$ 的圖像之下。兩個函數都是連續（continuous）的，所以它們之下的面積滿足

$$0 < u_n < u_{n-1} \qquad (n \geq 1).$$

想一想 兩條曲線在端點處相接，為甚麼不等式仍然是嚴格的？

差 $\sin^{n-1} t - \sin^n t$ 是連續的，永不為負，而且在（例如）$t = \frac{1}{4}\pi$ 處為正。一個連續函數若在某一點為正，則在該點附近的一個小區間上也為正，所以它圍出的面積為正。在兩個孤立點處相接，並不會改變這個結論。同樣的論證可證明 $u_n > 0$。

雙邊界限

設 $n \geq 1$。由於 $u_n > 0$，把 $u_n < u_{n-1}$ 乘以 nu_n，得 $nu_n^2 < nu_n u_{n-1}$。由於 $u_{n-1} > 0$，乘以 nu_{n-1}，得 $nu_n u_{n-1} < nu_{n-1}^2$。中間的量是 $\frac{1}{2}\pi$，所以

$$nu_n^2 < \tfrac{1}{2}\pi < nu_{n-1}^2 \qquad (n \geq 1).$$

夾逼與極限

右邊的不等式對每個 $n \geq 1$ 都成立，所以我們可以用 $n+1$ 代替 n：

$$\tfrac{1}{2}\pi < (n+1)u_n^2, \qquad \text{所以} \qquad nu_n^2 > \frac{n}{n+1}\cdot\tfrac{1}{2}\pi.$$

把它與左邊的不等式結合，得

$$\left(\frac{n}{n+1}\right)\tfrac{1}{2}\pi < nu_n^2 < \tfrac{1}{2}\pi.$$

當 $n \to \infty$ 時，$\frac{n}{n+1} = 1 - \frac{1}{n+1} \to 1$，所以下界趨向 $\frac{1}{2}\pi$。上界本身就是 $\frac{1}{2}\pi$。根據夾逼定理（squeeze theorem，又稱 sandwich theorem），

$$nu_n^2 \to \tfrac{1}{2}\pi \quad \text{當} n \to \infty.$$

驗算

- 歸約公式給出 $u_2 = \frac{1}{2}u_0 = \frac{1}{4}\pi$，這與 $\int_0^{\pi/2} \sin^2 t\,\mathrm{d}t = \frac{1}{4}\pi$ 吻合（$\sin^2$ 的平均值是 $\frac{1}{2}$）。於是 $2u_2u_1 = \frac{1}{2}\pi$，與預期一致。
- 以數值方法（辛普森法則，Simpson's rule）計算，$u_{10} \approx 0.38656$，即 $\frac{63}{512}\pi$。所以 $10u_{10}^2 \approx 1.4943$，介乎 $\frac{10}{11}\cdot\frac{1}{2}\pi \approx 1.4280$ 與 $\frac{1}{2}\pi \approx 1.5708$ 之間。
- 把遞推關係（recursion）一直算到 $n = 10^5$，得 $nu_n^2 \approx 1.570788$，已經非常接近 $\frac{1}{2}\pi$。

常見錯誤

- 忘記歸約公式需要 $n \geq 2$，所以常數必須用 u_0 和 u_1 來確定，而不能只用 u_1 和 u_2。
- 單憑 $\sin^n t \leq \sin^{n-1} t$ 就斷言 $u_n < u_{n-1}$。這樣只能得到 $\leq$；要得到嚴格不等式，被積函數的不等式必須在某個區間上嚴格成立。
- 試圖直接由 $\frac{1}{2}\pi < nu_{n-1}^2$ 得出下界。這條不等式限制的是 u_{n-1}，而不是 u_n，所以要先把下標移位。

STEP 之外

這些積分稱為沃利斯（Wallis）積分。把 u_{2m} 和 u_{2m+1} 寫成乘積，再利用 $nu_n^2 \to \frac{1}{2}\pi$，便可得出沃利斯乘積 $\frac{\pi}{2} = \frac{2}{1}\cdot\frac{2}{3}\cdot\frac{4}{3}\cdot\frac{4}{5}\cdots$。同一個估計也確定了斯特靈（Stirling）公式 $n! \sim \sqrt{2\pi n}\,(n/\mathrm{e})^n$ 中的常數 $\sqrt{2\pi}$。這裏所用的夾逼論證，是大學一年級分析學的標準工具。Abbott 的 *Understanding Analysis* 對此有細緻的闡述。

詳細解答

STEP II 1995，第 5 題

純數學 • 最優化與凹性

STEP II 1995，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：

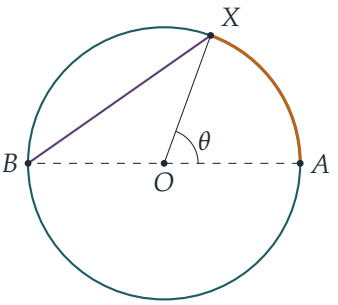
step.maths.org/questions/95-s2-q5

關鍵思路

選擇 (i) 和 (iii) 正是 (ii) 這一族路線的兩個極端：$\theta \to 0$ 表示直接游過對岸，$\theta \to \pi$ 表示全程跑步。因此我們把總時間寫成 θ 在 $[0, \pi]$ 上的函數，再問它的最小值在哪裏。出人意表的是，這個函數是凹函數（concave function），所以它唯一的駐點（stationary point）是極大點，而最佳路線永遠在其中一端。

解答

取泳池的半徑為 1 個長度單位，她的游泳速率為 1，於是她跑步的速率為 k。這樣時間便以（半徑）/（游泳速率）為單位，這並不影響任何比較。



三個時間

(i) 她游過直徑 AB，長度為 2，需時 2。

(iii) 她跑過一個半圓，長度為 π，需時 π/k。

(ii) 她跑過弧 AX，長度為 θ（半徑為 1），需時 θ/k。由於 $\angle AOB = \pi$，我們有 $\angle XOB = \pi - \theta$。三角形 XOB 是等腰三角形，其中 $OX = OB = 1$，所以弦（chord）長為

$$XB = 2\sin\tfrac{\pi-\theta}{2} = 2\cos\tfrac{\theta}{2}.$$

（另一個等價的做法：因為 AB 是直徑，所以 $\angle AXB = \frac{\pi}{2}$；又由等腰三角形 AOX 得 $\angle XAB = \frac{\pi-\theta}{2}$，所以 $XB = 2\sin\angle XAB$。）她的總時間為

$$T(\theta) = \frac{\theta}{k} + 2\cos\frac{\theta}{2}, \qquad 0 < \theta < \pi.$$

這條公式在兩端同樣成立：$T(0) = 2$ 是 (i) 的時間，而 $T(\pi) = \pi/k$ 是 (iii) 的時間。因此，三個選擇全部可以用 $[0, \pi]$ 上的同一個連續函數來描述。

T 在哪裏最小？

求導數，得

$$T'(\theta) = \frac{1}{k} - \sin\frac{\theta}{2}, \qquad T''(\theta) = -\tfrac{1}{2}\cos\frac{\theta}{2}.$$

當 $0 < \theta < \pi$ 時，我們有 $0 < \frac{\theta}{2} < \frac{\pi}{2}$，所以 $\cos\frac{\theta}{2} > 0$，從而 $T''(\theta) < 0$：函數 T 在 $[0, \pi]$ 上是嚴格凹（strictly concave）的。

想一想 若 $k > 1$，在 $\sin\frac{\theta}{2} = \frac{1}{k}$ 處有一個駐點。那不就是最佳的下水位置嗎？

不是：由於在該處 $T'' < 0$，它是一個極大點，是最差的下水位置。更一般地說，閉區間上的嚴格凹函數，其圖像嚴格位於連接兩端函數值的弦之上，所以對每一個滿足 $0 < \theta < \pi$ 的 θ，

$$T(\theta) > \Big(1 - \frac{\theta}{\pi}\Big)T(0) + \frac{\theta}{\pi}\,T(\pi) \geq \min\big(T(0),\, T(\pi)\big).$$

所以無論 k 取甚麼值，把 X 選在中途的任何位置，都及不上 (i) 和 (iii) 之中較快的那一個。（若 $k \leq 1$，我們可以更直接地看出這一點：這時在 $(0, \pi)$ 上 $T'(\theta) > 0$，因為 $\sin\frac{\theta}{2} < 1 \leq \frac{1}{k}$，所以 T 是遞增的，立即下水游泳便是最佳選擇。）

比較 (i) 和 (iii)

餘下只需比較 $T(0) = 2$ 與 $T(\pi) = \pi/k$：

- 若 $k < \frac{\pi}{2}$，則 $\pi/k > 2$，所以她應該**直接游過去**，即選擇 (i)；
- 若 $k > \frac{\pi}{2}$，則 $\pi/k < 2$，所以她應該**全程繞着池邊跑**，即選擇 (iii)；
- 若 $k = \frac{\pi}{2}$，(i) 和 (iii) 所需時間相同，兩者皆為最佳。

背後的物理原因很簡單：半圓的長度是直徑的 $\frac{\pi}{2}$ 倍。

最後，選擇 (ii) 是她永遠不會採用的：對每一個 $k > 0$ 和每一個 $\theta \in (0, \pi)$，它都嚴格地慢於 (i) 和 (iii) 之中較快的那一個。

驗算

取 $k = 3$，跑步需時 $\pi/3 \approx 1.047$，勝過游泳的 2；駐點位於 $\theta = 2\arcsin\frac{1}{3} \approx 0.680$，該處 $T \approx 2.112$，比兩端都慢，正是極大點應有的表現。用電腦在 e^{-4} 與 e^{4} 之間隨機抽取數千個 k，並在細密的 θ 網格上計算 T，從未發現區間內部有任何值低於 $\min(2, \pi/k)$，同時確認了分界點在 $k = \pi/2 \approx 1.571$。

常見錯誤

- 令 $T'(\theta) = 0$ 後，沒有檢查駐點的性質便宣稱所得結果是最優解：這裏它其實是極大點。
- 用了弧 XB 而不是弦，或者把弦長算成 $2\sin\frac{\theta}{2}$（那是 AX，不是 XB）。
- 忘記了 $k \le 1$ 的情況，這時根本沒有駐點。

STEP 之外

這裏起作用的原理是：區間上的凹函數必在端點取得最小值；在更高維度，這變成以下事實：線性（或凹）目標函數在多邊形上的最小值必在頂點取得，而這正是線性規劃（linear programming）的基礎。這道題也是斯涅爾定律（Snell's law）和費馬（Fermat）最短時間原理的近親，不過在那些情況下，答案通常是內點，因為時間函數反而是凸（convex）的。想做更多這類風格的題目，可以試試 Siklos 的 *Advanced Problems in Mathematics*（進階數學難題集）。

詳細解答

STEP II 1995，第 6 題

純數學 • 單位根

> **STEP II 1995，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/95-s2-q6

關鍵思路

七次單位根（roots of unity）是 $1, \alpha, \alpha^2, \ldots, \alpha^6$，它們的和為零。把六個非實數的根分成兩組，每組三個：$\{\alpha, \alpha^2, \alpha^4\}$ 和 $\{\alpha^3, \alpha^5, \alpha^6\}$。第二組正好是第一組各數的共軛複數（complex conjugate），所以兩組的和互為共軛。因此它們的和與積都是實數，而第一部分的結果把這兩個數變成二次式的係數。再取該根的實部和虛部，便得到所求的三角和。

解答

根與係數

若 u 和 v 是 $z^2 + az + b = 0$ 的根，則首一（monic）二次式 $z^2 + az + b$ 有因式 $z - u$ 和 $z - v$，所以對所有 z 都有

$$z^2 + az + b = (z - u)(z - v) = z^2 - (u + v)z + uv$$

。比較 z 的係數和常數項，得 $a = -u - v$ 及 $b = uv$，這正是所要證明的。

單位根

先提醒一點。試卷上印的是 $z^6 - 1 = 0$，但這是 $z^7 - 1 = 0$ 的誤印。事實上，$\alpha^6 = \cos(12\pi/7) + \mathrm{i}\sin(12\pi/7) \neq 1$，而題目其餘部分顯然都與七次有關。我們以 $z^7 - 1 = 0$ 來解題。

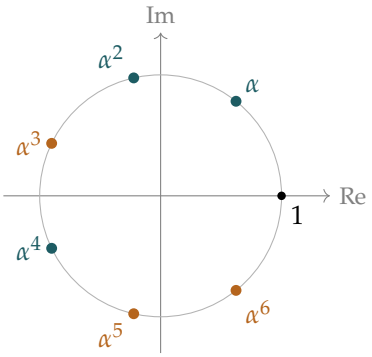
由棣美弗定理（de Moivre's theorem），

$$\alpha^7 = \cos 2\pi + \mathrm{i}\sin 2\pi = 1,$$

所以 α 是 $z^7-1=0$ 的根。更一般地，$\alpha^k=\cos(2k\pi/7)+\mathrm{i}\sin(2k\pi/7)$，而且 $(\alpha^k)^7=(\alpha^7)^k=1$。以下七個數

$$1,\ \alpha,\ \alpha^2,\ \alpha^3,\ \alpha^4,\ \alpha^5,\ \alpha^6$$

互不相同，因為它們的輻角（argument）$2k\pi/7$（$k=0,\dots,6$）在 $[0,2\pi)$ 內互不相同。7 次多項式最多有 7 個根，所以這些就是 $z^7-1=0$ 的全部根。它們位於單位圓上一個正七邊形的頂點。



整個解答全靠兩個事實。第一，各根之和為零：$z^7-1=(z-1)(z^6+z^5+\cdots+z+1)$，而 $\alpha\neq1$，所以

$$1+\alpha+\alpha^2+\alpha^3+\alpha^4+\alpha^5+\alpha^6=0. \qquad (*)$$

（等價的說法是：z^7-1 中 z^6 的係數為零。）第二，由於 $|\alpha|=1$，我們有 $\overline{\alpha^k}=\alpha^{-k}=\alpha^{7-k}$。所以 $\alpha^6=\overline{\alpha}$、$\alpha^5=\overline{\alpha^2}$ 及 $\alpha^3=\overline{\alpha^4}$。

求 A 和 B

設

$$p=\alpha+\alpha^2+\alpha^4,\qquad q=\alpha^3+\alpha^5+\alpha^6.$$

由上面的共軛配對，$q=\overline{p}$。這正是對另一個根最自然的猜想：一個實係數二次式若有非實數根 p，則 $\overline{p}$ 也必定是它的根。

> **想一想** 我們怎樣知道 p 不是實數？假如它是實數，「另一個根」就不一定是 $\overline{p}$ 了。

p 的虛部是 $\sin\frac{2\pi}{7}+\sin\frac{4\pi}{7}+\sin\frac{8\pi}{7}$，我們會在最後證明它是正數。所以 p 不是實數，而兩個根就是 p 和 $\overline{p}=q$。（無論如何，只要找到實數 A 和 B，使 p 和 q 為根，這個二次式便已確定。）

由 $(*)$，

$$p+q=\alpha+\alpha^2+\alpha^3+\alpha^4+\alpha^5+\alpha^6=-1.$$

至於乘積，先展開，再利用 $\alpha^7 = 1$ 把各冪次化簡：

$$\begin{aligned} pq &= (\alpha + \alpha^2 + \alpha^4)(\alpha^3 + \alpha^5 + \alpha^6) \\ &= \alpha^4 + \alpha^6 + \alpha^7 + \alpha^5 + \alpha^7 + \alpha^8 + \alpha^7 + \alpha^9 + \alpha^{10} \\ &= \alpha^4 + \alpha^6 + 1 + \alpha^5 + 1 + \alpha + 1 + \alpha^2 + \alpha^3 \\ &= 3 + (\alpha + \alpha^2 + \alpha^3 + \alpha^4 + \alpha^5 + \alpha^6) = 3 - 1 = 2. \end{aligned}$$

留意在這九項之中，除了三個 1 之外，$\alpha, \ldots, \alpha^6$ 每個都恰好出現一次。由第一部分，$A = -(p+q)$ 及 $B = pq$，所以

$$A = 1, \qquad B = 2,$$

而 p 和 q 是 $z^2 + z + 2 = 0$ 的根，即

$$z = \frac{-1 \pm \mathrm{i}\sqrt{7}}{2}.$$

三角和

取 $p = \alpha + \alpha^2 + \alpha^4$ 的實部和虛部，並利用 $\alpha^k = \cos(2k\pi/7) + \mathrm{i}\sin(2k\pi/7)$，得

$$\operatorname{Re} p = \cos\frac{2\pi}{7} + \cos\frac{4\pi}{7} + \cos\frac{8\pi}{7},$$
$$\operatorname{Im} p = \sin\frac{2\pi}{7} + \sin\frac{4\pi}{7} + \sin\frac{8\pi}{7}.$$

$z^2 + z + 2 = 0$ 的兩個根的實部都是 $-\frac{1}{2}$，所以

$$\cos\frac{2\pi}{7} + \cos\frac{4\pi}{7} + \cos\frac{8\pi}{7} = -\frac{1}{2},$$

這正是所要證明的。（另一做法：由 $p + \overline{p} = -1$ 可直接得出 $2\operatorname{Re} p = -1$。）

虛部是 $\pm\frac{\sqrt{7}}{2}$，我們必須判斷是哪一個。

> **想一想** 三個正弦值之中兩個是正數，一個是負數。我們怎樣才能肯定哪一方佔優？

把負的那一項與其中一個正項比較。由於 $\frac{8\pi}{7} = \pi + \frac{\pi}{7}$，我們有 $\sin\frac{8\pi}{7} = -\sin\frac{\pi}{7}$。此外，$\sin\frac{4\pi}{7} = \sin\left(\pi - \frac{4\pi}{7}\right) = \sin\frac{3\pi}{7}$。正弦函數在 $[0, \frac{\pi}{2}]$ 上嚴格遞增（strictly increasing），而 $0 < \frac{\pi}{7} < \frac{3\pi}{7} < \frac{\pi}{2}$，所以 $\sin\frac{3\pi}{7} > \sin\frac{\pi}{7}$。因此

$$\sin\frac{2\pi}{7} + \sin\frac{4\pi}{7} + \sin\frac{8\pi}{7} = \underbrace{\sin\frac{2\pi}{7}}_{>0} + \underbrace{\left(\sin\frac{3\pi}{7} - \sin\frac{\pi}{7}\right)}_{>0} > 0,$$

從而

$$\sin\frac{2\pi}{7} + \sin\frac{4\pi}{7} + \sin\frac{8\pi}{7} = \frac{\sqrt{7}}{2}.$$

這亦確認了 $p = \frac{-1+\mathrm{i}\sqrt{7}}{2}$ 不是實數，正如我們在上面所用到的。

驗算

用計算機（以弧度計算），$\cos\frac{2\pi}{7} \approx 0.6235$、$\cos\frac{4\pi}{7} \approx -0.2225$ 及 $\cos\frac{8\pi}{7} \approx -0.9010$，三者相加得 -0.5000。三個正弦值約為 0.7818、0.9749 和 -0.4339，相加得 1.3229。這與 $\frac{\sqrt{7}}{2} \approx 1.3229$ 吻合。你也可以直接驗證 $p^2 + p + 2 = 0$：取 $p = \frac{-1+\mathrm{i}\sqrt{7}}{2}$，則 $p^2 = \frac{-6-2\mathrm{i}\sqrt{7}}{4}$，由此即得 $p^2 + p + 2 = 0$。

常見錯誤

- 照試卷所印使用 $z^6 - 1 = 0$。這樣 α 便不是根，而「和為零」這個事實 $(*)$ 亦不成立。
- 展開 pq 時忘記用 $\alpha^7 = 1$ 化簡冪次，或數錯了等於 1 的那三項。
- 聲稱正弦之和是正數，「因為有兩項是正數」。你需要作出比較，例如 $\sin\frac{3\pi}{7} > \sin\frac{\pi}{7}$。

STEP 之外

$\{1,2,4\}$ 與 $\{3,5,6\}$ 這個分法並非隨意。1, 2, 4 這三個數正是模 7 的非零平方數，而 $2p + 1 = \mathrm{i}\sqrt{7}$ 是一個*高斯和*（Gauss sum）。高斯（Gauss）證明了：對於形如 $4k + 3$ 的質數 ℓ，α^r 對所有二次剩餘（quadratic residue）r 求和，結果是 $\frac{-1+\mathrm{i}\sqrt{\ell}}{2}$。這是兩套理論的第一步：一是關於哪些正多邊形可以用尺規作圖的理論（伽羅瓦理論，Galois theory），二是二次互反律（quadratic reciprocity）。Davenport 的 *The Higher Arithmetic* 講述二次剩餘，而 Needham 的 *Visual Complex Analysis* 則從幾何角度看單位根。

詳細解答

STEP II 1995，第 7 題

純數學 • 希羅公式

> **STEP II 1995，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/95-s2-q7

關鍵思路

內切圓（incircle）是貫穿全題的主角。由圓心到各切點的半徑都垂直於三角形的邊，正好是把 ABC 鋪滿的三個小三角形的現成的高，由此便得 (i)。第 (ii) 部分是希羅公式（Heron's formula）的經典推導：寫出 $\sin^2\alpha = 1-\cos^2\alpha$，用餘弦定理（cosine rule）消去角度，然後利用平方差（difference of two squares）分解因式，前後兩次。至於球體，則要看一個垂直截面。球體與洞口各邊的接觸點，原來正正就是內切圓與三邊相切的點。

解答

第 (i) 部分

設內切圓與 BC、CA 和 AB 分別相切於 P、Q 和 R'。通過切點的半徑垂直於切線（tangent），所以 $IP \perp BC$、$IQ \perp CA$ 及 $IR' \perp AB$，而且每一條的長度都是 r。因此三角形 BIC 的底為 a，垂直高度為 r，其餘兩個三角形亦同理：

$$[BIC] = \tfrac{1}{2}ar, \qquad [CIA] = \tfrac{1}{2}br, \qquad [AIB] = \tfrac{1}{2}cr.$$

由於 I 位於三角形之內，這三個三角形恰好拼成 ABC，互不重疊。相加得

$$\Delta = \tfrac{1}{2}r(a+b+c) = rs,$$

正是所求。

第 (ii) 部分

把 $\Delta = \frac{1}{2}bc\sin\alpha$ 平方，並利用 $\sin^2\alpha = 1 - \cos^2\alpha$，

$$\Delta^2 = \tfrac{1}{4}b^2c^2(1-\cos^2\alpha) = \tfrac{1}{16}\left[4b^2c^2 - (2bc\cos\alpha)^2\right],$$

這就是第一個結果。方括號內是一個平方差：

$$16\Delta^2 = (2bc - 2bc\cos\alpha)(2bc + 2bc\cos\alpha).$$

根據餘弦定理，$2bc\cos\alpha = b^2 + c^2 - a^2$。把它代入每個因式，

$$2bc - 2bc\cos\alpha = a^2 - (b^2 - 2bc + c^2) = a^2 - (b-c)^2,$$
$$2bc + 2bc\cos\alpha = (b^2 + 2bc + c^2) - a^2 = (b+c)^2 - a^2,$$

所以

$$\Delta^2 = \tfrac{1}{16}\left[a^2 - (b-c)^2\right]\left[(b+c)^2 - a^2\right],$$

即第二個結果。每個方括號又是一個平方差：

$$16\Delta^2 = (a-b+c)(a+b-c)(b+c-a)(b+c+a).$$

現在引入 s。由於 $a+b+c = 2s$，

$$b+c-a = 2(s-a), \qquad a-b+c = 2(s-b), \qquad a+b-c = 2(s-c),$$

所以該乘積為 $2s\cdot 2(s-a)\cdot 2(s-b)\cdot 2(s-c) = 16\,s(s-a)(s-b)(s-c)$。因此 $\Delta^2 = s(s-a)(s-b)(s-c)$。由於 $\Delta > 0$，我們取正平方根：

$$\Delta = \sqrt{s(s-a)(s-b)(s-c)}.$$

第 (iii) 部分

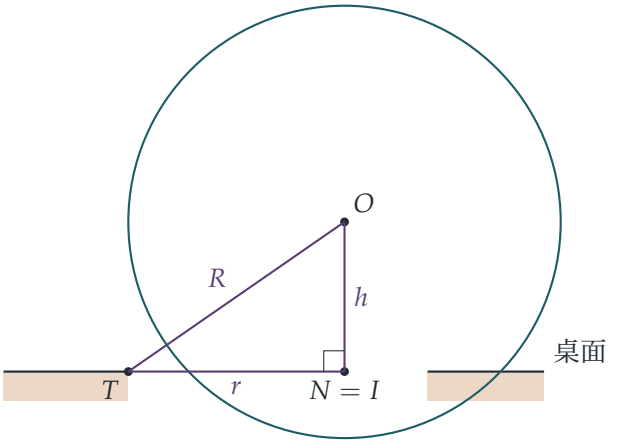
設 O 為球心，h 為球心高出桌面的高度。設 N 為由 O 到桌面所在平面的垂足（foot of the perpendicular），所以 $ON = h$。球體擱在洞口的三條邊上。假設它與邊 BC 相切於點 T。

> ? **想一想** 為甚麼 N 必定是內心（incentre）I？

在球體與直線相切的點上，半徑垂直於該直線，所以 $OT \perp BC$ 且 $OT = R$。此外，ON 垂直於整個桌面平面，所以 $ON \perp BC$。於是 BC 垂直於平面 ONT 內兩條不平行的直線，因而垂直於該平面，當然也垂直於 NT。所以 NT 就是由 N 到 BC 的垂直距離，而在直角三角形 ONT 中運用畢氏定理（Pythagoras' theorem），得

$$NT^2 = R^2 - h^2.$$

對邊 CA 和 AB 作同樣的論證，可知 N 與三條邊的距離同為 $\sqrt{R^2-h^2}$。點 N 位於三角形之內（球體坐落在洞口之上），而三角形內與三邊等距的點只有內心一個。所以 $N=I$，這個共同的距離就是 r，而各接觸點正是內切圓與三邊相切之處。這些點都在邊本身之上，而不是在邊的延線上，所以球體確實是擱在三條邊上。



上圖顯示通過 O 和 T 的垂直截面；邊 BC 穿過 T，垂直於紙面。在三角形 ONT 中運用畢氏定理，現在得到 $h^2=R^2-r^2$。由 (i) 和 (ii)，

$$r=\frac{\Delta}{s}=\frac{\sqrt{s(s-a)(s-b)(s-c)}}{s},\qquad \text{所以}\qquad r^2=\frac{(s-a)(s-b)(s-c)}{s}.$$

因此球心高出桌面的高度為

$$h=\sqrt{R^2-\frac{(s-a)(s-b)(s-c)}{s}}$$

。這需要 $R>r$：不比內切圓大的球體會直接穿過洞口掉下去。

驗算

取 3–4–5 三角形：$s=6$，希羅公式給出 $\Delta=\sqrt{6\cdot3\cdot2\cdot1}=6$，與 $\frac{1}{2}\cdot3\cdot4$ 吻合。於是 $r=\Delta/s=1$，你可以用直角三角形的公式 $r=\frac{1}{2}(3+4-5)$ 加以核實。當 $R=2$ 時，公式給出 $h=\sqrt{4-1}=\sqrt{3}\approx1.732$。當 $R\to r^+$ 時，$h\to0$：球體不斷下沉，直至其赤道與桌面齊平，並擱在內切圓上。作為更全面的檢驗，我們以數值方法測試了 18,816 個隨機三角形。在每一個情況中，置於內心上方高度 h 處的球心，與三條邊所在直線的距離都恰好是 R，而且最近的點都落在邊之內。

常見錯誤

▶ 取 $h=\sqrt{R^2-d^2}$，其中 d 是由 I 到某個頂點的距離。球體接觸的是邊，而不是角。

- 毫無理由地假設球心位於內心的正上方。這需要用上述的垂直論證，或者至少要給出清晰的對稱理由。
- 在 (ii) 中把因式的配對弄錯：$b + c - a = 2(s - a)$，而不是 $2(s - b)$ 或 $2(s - c)$。

STEP 之外

希羅公式屬於一系列關於面積和體積的對稱公式。婆羅摩笈多（Brahmagupta）公式對圓內接四邊形（cyclic quadrilateral）起着同樣的作用，而凱萊－門格行列式（Cayley–Menger determinant）則可由四面體的六條棱長求出其體積。「半徑 $\perp$ 切線，所以垂足就是內心」這一步，是三垂線定理（theorem of the three perpendiculars）的三維版本，在向量幾何中很值得掌握。想找更多這類題目，可參閱 Siklos 的 *Advanced Problems in Mathematics*。

詳細解答

STEP II 1995，第 8 題

純數學 • 邏輯斯諦微分方程

STEP II 1995，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/95-s2-q8

本題在 *Beyond STEP* 第 8.3 節收穫種羣中有完整的引導式講解：https://beyondstep.keithmcnulty.org/zh-hant/ch08/harvesting.html

關鍵思路

這條增長定律是一條邏輯斯諦方程（logistic equation），屬於可分離變數（separable）的類型：利用部分分式（partial fractions），積分便化為兩個對數。至於第二部分，把新方程右邊配方（completing the square）後可見，平移 $y = x - K + \alpha$ 會把新方程變成完全相同的邏輯斯諦方程，只是以 α 取代了 K。因此我們可以直接沿用第一部分的答案，無須再做任何積分。

解答

沒有移除時的種羣

我們取 $K > 0$（種羣由 $x = K$ 開始，故必有 $K > 0$）。方程為

$$\frac{\mathrm{d}x}{\mathrm{d}t} = (2K - x)x.$$

由 $x = K$ 出發，變化率為 $K^2 > 0$。解永遠不會到達 $x = 0$ 或 $x = 2K$，因為這兩個常數函數本身就是解，而這條方程的解不可能互相穿越。所以自始至終 $0 < x < 2K$，我們可以除以 $(2K - x)x$。寫成部分分式，

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

於是分離變數後得

$$\frac{1}{2K}\int\left(\frac{1}{x}+\frac{1}{2K-x}\right)\mathrm{d}x=\int\mathrm{d}t,$$
$$\frac{1}{2K}\Big(\ln x-\ln(2K-x)\Big)=t+C.$$

由於 x 和 $2K-x$ 均為正數，無須加上絕對值符號。在 $t=0$ 時，$x=2K-x=K$，故左邊為 0，而 $C=0$。因此

$$\ln\frac{x}{2K-x}=2Kt\quad\Longrightarrow\quad x=(2K-x)\,\mathrm{e}^{2Kt}\quad\Longrightarrow\quad x=\frac{2K\mathrm{e}^{2Kt}}{1+\mathrm{e}^{2Kt}}=\frac{2K}{1+\mathrm{e}^{-2Kt}}.$$

為了配合題目要求的形式，寫 $2=(1+\mathrm{e}^{-2Kt})+(1-\mathrm{e}^{-2Kt})$：

$$x(t)=\frac{K(1+\mathrm{e}^{-2Kt})+K(1-\mathrm{e}^{-2Kt})}{1+\mathrm{e}^{-2Kt}}=K+K\,\frac{1-\mathrm{e}^{-2Kt}}{1+\mathrm{e}^{-2Kt}},$$

這正是所需的結果。（等價地，$x=K+K\tanh Kt$。）

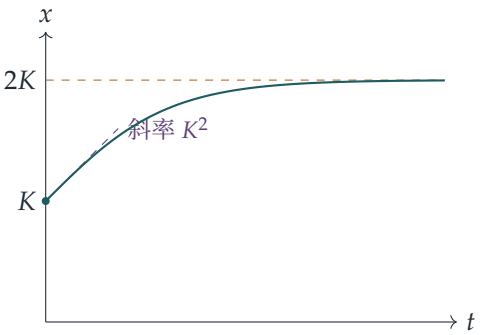
草圖。 當 $t\geq 0$ 時，$\mathrm{e}^{-2Kt}\in(0,1]$，所以 $K\leq x<2K$。於是 $\dot{x}=(2K-x)x>0$，故 x 遞增。把方程再求導一次，

$$\ddot{x}=(2K-2x)\,\dot{x}\leq 0\quad\text{因為}x\geq K,$$

所以圖像是凹向下的（concave），在 $t=0$ 處的斜率為 K^2（該處 $\ddot{x}=0$：時間的原點正是整條 S 形邏輯斯諦曲線的拐點（point of inflection））。當 $t\to\infty$ 時，$\mathrm{e}^{-2Kt}\to 0$，所以

$$x(t)\to 2K.$$

種羣數量在 $2K$ 處趨於平穩，從下方逼近它，但永遠不會到達。



以速率 L 移除

每小時移除 L 微克，即是從變化率中減去 L：

$$\frac{\mathrm{d}x}{\mathrm{d}t}=(2K-x)x-L.$$

把右邊配方：

$$(2K-x)x-L=-(x^2-2Kx)-L=K^2-L-(x-K)^2=\alpha^2-(x-K)^2,$$

其中 $\alpha=\sqrt{K^2-L}>0$，因為 $K^2>L$。

想一想 為甚麼 $y=x-K+\alpha$ 是合適的代換？

右邊 $\alpha^2-(x-K)^2$ 在 $x=K\pm\alpha$ 處為零，這就是兩個平衡（equilibrium）種羣數量。平移 $K-\alpha$ 會把較低的平衡點移到 0，而原來的邏輯斯諦方程的較低平衡點也正在那裏。由 $x-K=y-\alpha$ 及 $\dot{y}=\dot{x}$，

$$\frac{\mathrm{d}y}{\mathrm{d}t}=\alpha^2-(y-\alpha)^2=2\alpha y-y^2=(2\alpha-y)\,y.$$

這就是第一條方程，只是以 α 取代了 K。初始條件（initial condition）同樣吻合：$x(0)=K$ 給出 $y(0)=\alpha$，正好對應原來的 $x(0)=K$。由於 $\alpha>0$，第一部分的論證可以一字不改地套用，於是

$$y(t)=\frac{2\alpha}{1+\mathrm{e}^{-2\alpha t}},\qquad\text{所以}\qquad x(t)=K-\alpha+\frac{2\alpha}{1+\mathrm{e}^{-2\alpha t}}.$$

當 $t\to\infty$ 時，$\mathrm{e}^{-2\alpha t}\to0$，所以 $y\to2\alpha$，而

$$x(t)\to K-\alpha+2\alpha=K+\alpha,$$

這正是所求。當 $L>0$ 時，$\alpha<K$，所以被收割的種羣會穩定在較低的水平 $K+\alpha<2K$，一如所料。

驗算

- 在 $t=0$ 時，兩條公式都給出 $x=K$。設 $L=0$ 則 $\alpha=K$，第二條公式便化為 $2K/(1+\mathrm{e}^{-2Kt})$，即第一條公式。
- 對 $x=K+K\tanh Kt$ 求導，得 $\dot{x}=K^2\operatorname{sech}^2Kt=K^2(1-\tanh^2Kt)=(K-K\tanh Kt)(K+K\tanh Kt)=(2K-x)x$。
- 數值上，取 $K=1$，$x(1)=2/(1+\mathrm{e}^{-2})\approx1.7616$；取 $K=1$ 及 $L=0.75$（故 $\alpha=\frac{1}{2}$），$x(1)\approx1.2311$。兩者都與直接以數值方法（龍格—庫塔法，Runge–Kutta）求解微分方程所得的結果吻合，而且第二個解穩定在 $1.5=K+\alpha$。

常見錯誤

- 在部分分式中漏掉因子 $\frac{1}{2K}$，結果得出 e^{-t} 而不是 e^{-2Kt}。

- 在第二部分寫成 $+L$，或者忘記積分常數（constant of integration）必須重新求出。這裏的代換令 $y(0) = \alpha$，所以無須新的常數。

STEP 之外

第二部分讓我們初次接觸*平衡點的穩定性*（stability of equilibria）。方程 $\dot{x} = \alpha^2 - (x - K)^2$ 的平衡點在 $K \pm \alpha$：上面的一個吸引附近的解，下面的一個則排斥它們。只要看右邊的正負號便可看出這一點，完全不用解方程。這種相線（phase line）方法，以及 $L > K^2$ 時的情況（沒有平衡點，種羣在有限時間內滅絕），可參閱 Boyce 與 DiPrima 的 *Elementary Differential Equations*（微分方程入門教科書）。

詳細解答

STEP II 1995，第 9 題

力學 • 摩擦力與平衡

STEP II 1995，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：

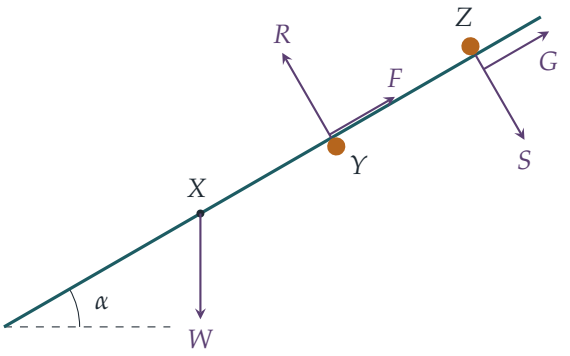
step.maths.org/questions/95-s2-q9

關鍵思路

桿沿着一條最大斜率線（line of greatest slope）放置，所以很自然會沿桿的方向和垂直於桿的方向分解（resolve）各力。桿擱在其上的那根橫桿把它向上推，使它離開平面；桿從其下穿過的那根橫桿則只能把它向下推。對 Y 取力矩（moment），便可知道桿的中心必須在哪裏；然後令兩個摩擦力（friction）都沿斜面向上並達到極限值（limiting value），就得到 μ 的最小值。

解答

我們把桿看成一根重量為 W 的均勻剛性桿（uniform rigid rod），把兩根橫桿看成固定的細桿，各自對桿施加一個垂直於桿的法向反作用力（normal reaction）和一個沿桿方向的摩擦力。由於桿與兩根水平橫桿都垂直，它便沿着兩根橫桿所在平面的一條最大斜率線放置，因此它與水平面的夾角為 α。設 R 為在 Y 處的反作用力（把桿推離平面，因為桿擱在這根橫桿之上），S 為在 Z 處的反作用力（把桿推向平面，因為桿在這根橫桿之下），其中 $R, S \geq 0$。設 F 和 G 分別為在 Y 和 Z 處的摩擦力，以沿斜面向上為正。



中心必須在哪裏

沿垂直於桿的方向分解，並對 Y 取力矩（距離沿桿量度）：

$$R = S + W\cos\alpha,$$
$$S\,d = W\cos\alpha \cdot XY.$$

S 對 Y 的力矩，必須由重量分量 $W\cos\alpha$ 的力矩抵消，而這個分量與 S 一樣指向平面。兩個方向相同的平行力，只有分別作用在 Y 的兩側時，它們對 Y 的力矩才可以互相抵消。由於 Z 在 Y 的斜面上方，中心 X 必須在 Y 處或在 Y 的下方，於是

$$XZ = XY + YZ, \qquad \text{所以} \qquad l = XY + d \geq d.$$

（假如 X 在 Y 與 Z 之間，或在 Z 之外，力矩方程便會迫使 $S < 0$：在 Z 處的橫桿就要把桿拉住，但它做不到，桿便會轉動而離開這根橫桿。）等式 $l = d$ 是臨界情況，此時 X 位於 Y 而 $S = 0$。

想一想 試卷上寫的是 $l \leq d$。為甚麼我們得到相反的不等式？

試卷上的不等式（inequality）是印刷錯誤：題目本身接着取 $l = 2d$，而這只有在 $l \geq d$ 時才可能。上面的論證說明，平衡（equilibrium）所要求的正是 $l \geq d$，我們以下就用這個條件。

μ 的最小值

以 $XY = l - d$ 代入，兩條方程給出

$$S = \frac{W\cos\alpha\,(l-d)}{d}, \qquad R = \frac{W\cos\alpha\,l}{d}, \qquad \text{所以} \qquad R + S = \frac{W\cos\alpha\,(2l-d)}{d}.$$

沿桿的方向分解，摩擦力必須承托重量沿斜面向下的分量：

$$F + G = W\sin\alpha.$$

摩擦定律要求 $|F| \le \mu R$ 及 $|G| \le \mu S$，所以 $F+G \le \mu(R+S)$。因此平衡需要

$$W\sin\alpha \le \mu\,\frac{W\cos\alpha\,(2l-d)}{d}, \qquad \text{即} \qquad \mu \ge \frac{d\tan\alpha}{2l-d}.$$

反過來，若 μ 等於這個值，則 $F=\mu R$ 和 $G=\mu S$（兩者都沿斜面向上，兩者都達到極限值）滿足每一條方程，所以桿可以處於平衡。因此最小的可能值是

$$\mu_{\min} = \frac{d\tan\alpha}{2l-d}.$$

當 $l=2d$ 時，上式變成 $\mu_{\min} = \dfrac{d\tan\alpha}{3d} = \frac{1}{3}\tan\alpha$，所以 $\mu \ge \frac{1}{3}\tan\alpha$，正是所要證明的結果。

驗算

- 當 $l=d$ 時，中心位於較低的橫桿之上，$S=0$，我們得到 $\mu \ge \tan\alpha$：這正是單一質點（particle）靜止在粗糙斜面上的熟悉條件。
- 當 $l \to \infty$ 時，反作用力無限增大，所以只需要極少的摩擦，而的確有 $\mu_{\min} \to 0$。
- 數值上，取 $\alpha = 0.7$ 及 $l = 2d$，公式給出 $0.28076\ldots$，即 $\frac{1}{3}\tan 0.7$。我們亦用電腦核對了對隨機選取的點所列的力方程和力矩方程，並驗證了平衡在略高於 $\mu_{\min}$ 時成立，在略低於它時則不成立。

常見錯誤

- 把兩個反作用力取成同一方向。桿從其下穿過的那根橫桿，只能把桿推向平面。
- 假設只有一根橫桿處的摩擦達到極限。要令 μ 盡可能小，兩根橫桿都必須提供最大的摩擦力，方向沿斜面向上。

STEP 之外

這是一道帶有不等式約束的靜力學（statics）小問題：接觸力必須滿足一組線性方程，同時滿足摩擦不等式，而 μ 的最小值就是可行集（feasible set）剛好變成非空的那一點。同樣的結構，即線性方程加上線性不等式，正是線性規劃（linear programming）以及工程學中摩擦接觸分析的基礎。想進一步了解剛體靜力學和摩擦，可參閱 Kleppner and Kolenkow, *An Introduction to Mechanics*（一本大學力學入門教科書）。

詳細解答

STEP II 1995，第 10 題

力學 • 碰撞與恢復係數

STEP II 1995，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/95-s2-q10

關鍵思路

每一次碰撞都是一個由兩條方程組成的問題：動量守恆（conservation of momentum）和牛頓實驗定律（Newton's experimental law）。依次處理兩次碰撞，令第二個球的最終速度等於 u，然後解出 m_1；不等式則來自一個簡單的事實：質量必須是正數。至於最後一部分，留意碰撞只取決於相對速度，所以只要在一個以速率 v 運動的參考系（frame）中觀察這個運動，問題便化為第一部分。

解答

我們把各球視為質點（particle），在光滑桌面上沿同一直線運動，因此水平方向的力就只有球與球之間的衝量（impulse）。取第一個球的運動方向為正方向。



第一次碰撞

設第一次碰撞剛結束時，第一個球和第二個球的速度分別為 v_1 和 w。由動量守恆和牛頓定律（分離速率等於接近速率的 e 倍）可得

$$m_1 v_1 + m_2 w = m_1 u, \qquad w - v_1 = eu.$$

把 $v_1 = w - eu$ 代入第一條方程，得 $(m_1 + m_2)w = m_1(1+e)u$，所以

$$w = \frac{m_1(1+e)u}{m_1 + m_2}, \qquad v_1 = \frac{(m_1 - em_2)u}{m_1 + m_2}.$$

第二次碰撞

現在第二個球以速度 w 撞向靜止的第三個球。這與剛才的情況完全相同，只是 m_1, m_2, u 換成了 m_2, m_3, w，所以第二個球碰撞後的速度是

$$\frac{(m_2 - em_3)w}{m_2 + m_3}.$$

題目告訴我們這個速度等於 u。代入 w 的表達式，並約去 $u \neq 0$，得

$$\frac{(m_2 - em_3)\, m_1(1+e)}{(m_2+m_3)(m_1+m_2)} = 1, \qquad \text{所以} \qquad m_1(1+e)(m_2 - em_3) = (m_1+m_2)(m_2+m_3).$$

把含 m_1 的項合併，

$$m_1\big[(1+e)m_2 - e(1+e)m_3 - m_2 - m_3\big] = m_2(m_2+m_3),$$

而方括號內的式子可化簡為 $em_2 - m_3(1+e+e^2)$。因此

$$\boxed{m_1 = \frac{m_2(m_2+m_3)}{em_2 - m_3(1+e+e^2)}}.$$

想一想 我們手上只有一條等式，不等式從何而來？

答案來自物理：質量是正數。分子 $m_2(m_2+m_3)$ 是正數，而我們用來相除的那個方括號不可能是零（否則前一條等式就會變成 $0 = m_2(m_2+m_3)$）。所以 $m_1 > 0$ 迫使分母為正數：

$$em_2 - m_3(1+e+e^2) > 0, \qquad \text{即} \qquad m_2 e > m_3(1+e+e^2),$$

這正是所要證明的結果。

由於最後一部分要數算碰撞的次數，值得確認一下之後再沒有其他事情發生。第二次碰撞後，第二個球以速度 u 運動，而第一個球的速度是 $v_1 = u - \frac{m_2(1+e)u}{m_1+m_2} < u$，所以第一個球永遠不會再追上第二個球。第三個球則以相對速率 $ew > 0$ 遠離第二個球。因此碰撞恰好發生兩次。

最後一部分

現在三個球都有共同的速度 v，只有第一個球例外，它以 $u > v$ 的速度運動。我們在一個沿該直線以恆定速度 v 運動的參考系中觀察這個運動。在這個參考系中：

- 第一個球以 $u - v > 0$ 的速度朝第二個球運動，而第二個球和第三個球都是靜止的；
- 每次碰撞中動量仍然守恆，因為把每個速度都減去 v，動量方程兩邊的改變量相同，都是 $(m_a + m_b)v$；

- 牛頓定律保持不變，因為它只涉及速度之差。

所以在這個運動的參考系中，情況與第一部分完全一樣，只是 u 換成了 $u-v$，而各質量仍滿足同一關係式（該關係式根本不含 u）。根據第一部分，兩次碰撞之後，第二個球相對於這個參考系的速度是 $u-v$。回到桌面的參考系，我們把 v 加回去：第二個球在兩次碰撞之後的速度是

$$\boxed{u}.$$

條件 $u>v$ 保證了第一次碰撞確實會發生；至於 $v>0$ 這個假設，論證中其實用不着。

驗算

取 $e=\frac{1}{2}$、$m_3=1$ 和 $m_2=4$；則 $em_2=2>\frac{7}{4}=m_3(1+e+e^2)$，而公式給出 $m_1=\frac{4\cdot 5}{1/4}=80$。取 $u=7$，第一次碰撞後第二個球的速度是 $w=\frac{80\cdot\frac{3}{2}\cdot 7}{84}=10$，第二次碰撞後它的速度則是 $\frac{(4-\frac{1}{2})\cdot 10}{5}=7=u$，與預期相符。我們亦用電腦以分數作精確運算，對超過 10^4 組隨機選取、符合條件的質量和 e 值模擬了碰撞過程，每一次都確認碰撞恰好發生兩次、第二個球最終的速度是 u，而且把共同速度 $v<u$ 加到第二個球和第三個球上之後，最終速度同樣是 u。

常見錯誤

- 把牛頓定律的方向寫反，例如寫成 $v_1-w=eu$；在前面的球，最終的速度必定不小於後面的球。
- 在最後一部分從頭再計算各次碰撞，而沒有利用轉換參考系的方法，題目其實已明確提示可以這樣做。
- 忘記在最後把 v 加回去，以 $u-v$ 作為答案。

STEP 之外

最後一部分是伽利略不變性(Galilean invariance)的一個例子：在每一個以恆定速度運動的參考系中，碰撞定律的形式都相同，所以你可以選擇令問題最容易處理的參考系。在大學的力學課程中，最常用的選擇是質心參考系（centre-of-mass frame），在這個參考系中總動量為零，而一次碰撞只不過是把各速度反向，再乘以 e。Kleppner 與 Kolenkow 合著的 *An Introduction to Mechanics*（力學入門教科書）在處理碰撞時大量運用了這個想法。

詳細解答

STEP II 1995，第 11 題

力學 • 線性空氣阻力

STEP II 1995，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：

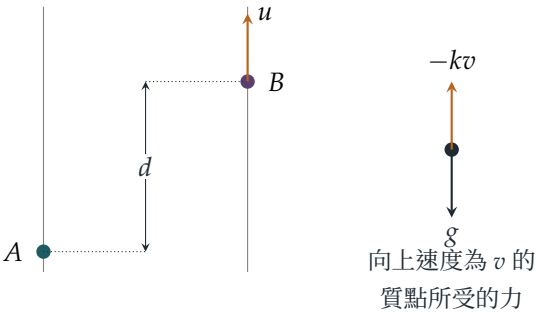
step.maths.org/questions/95-s2-q11

關鍵思路

由於阻力與速度成線性關係，運動方程（equation of motion）是線性的，而且無論質點正在上升還是下降，同一條方程都成立。因此把兩條運動方程相減：重力項互相抵消，相對速度（relative velocity）純粹按指數衰減。把它積分一次便直接得到兩者的距離，完全不必知道任何一個質點身在何處。

解答

把每個物體視為在鉛垂線上運動的單位質量質點（particle），受到重量 g 以及大小為 $k|v|$、方向與運動相反的阻力（resistance）作用。取向上為正方向，以 A 為原點，並設質點 1（由 A 出發）和質點 2（由 B 出發）在時間 t 的高度為 x_1、x_2，速度為 v_1、v_2。



運動方程

對於向上速度為 v 的質點，阻力為 $-kv$：若 $v>0$（上升），阻力向下；若 $v<0$（下降），阻力向上，兩種情況都正好是「與運動方向相反」。因此，對單位質量應用牛頓第二定律（Newton's second law），每個質點都有

$$\frac{\mathrm{d}v_1}{\mathrm{d}t} = -g - kv_1, \qquad \frac{\mathrm{d}v_2}{\mathrm{d}t} = -g - kv_2,$$

對所有 $t \geq 0$ 均成立。

> **想一想** 質點 2 先上升，停下來，然後下降。難道不需要為每個階段各寫一條方程嗎？

假如阻力與速率的平方成正比，那就需要：這時的力是 $-kv|v|$，當 v 變號時，它的形式會隨之改變。但這裏的力是 $-kv$，這單一的表達式在 v 反向時已自動反向，所以一條方程便涵蓋整個運動。

相對運動

設 $w = v_2 - v_1$ 為質點 2 相對於質點 1 的速度。把兩條方程相減，重量項互相抵消：

$$\frac{\mathrm{d}w}{\mathrm{d}t} = -kw, \qquad w(0) = u - 0 = u,$$

所以 $w = u\,\mathrm{e}^{-kt}$。現在設 $s = x_2 - x_1$，由於 B 在 A 上方高度 d 處，故 $s(0) = d$。於是 $\mathrm{d}s/\mathrm{d}t = w$，而

$$s = d + \int_0^t u\,\mathrm{e}^{-k\tau}\,\mathrm{d}\tau = d + \frac{u}{k}\bigl(1 - \mathrm{e}^{-kt}\bigr).$$

由於 $u \geq 0$ 及 $1 - \mathrm{e}^{-kt} \geq 0$，我們有 $s \geq d > 0$：質點 2 永遠在質點 1 之上，兩個質點永不相遇。因此它們在時間 t 的距離是

$$\boxed{d + \frac{u}{k}\bigl(1 - \mathrm{e}^{-kt}\bigr)}$$

當 $t \to \infty$ 時，因為 $k > 0$，所以 $\mathrm{e}^{-kt} \to 0$，距離趨向 $d + \dfrac{u}{k}$，正如題目所求。

個別質點的運動（如果你寧願走遠路）

分別解每條方程（它是一階線性方程（first-order linear），也是可分離變量（separable）的方程），得 $v = -\frac{g}{k} + \bigl(v(0) + \frac{g}{k}\bigr)\mathrm{e}^{-kt}$，再積分一次，

$$x_1 = -\frac{gt}{k} + \frac{g}{k^2}\bigl(1 - \mathrm{e}^{-kt}\bigr),$$

$$x_2 = d - \frac{gt}{k} + \frac{1}{k}\Bigl(u + \frac{g}{k}\Bigr)\bigl(1 - \mathrm{e}^{-kt}\bigr).$$

兩式相減，gt/k 和 g/k^2 的項互相抵消，得回同一個 s。兩個質點的速度都趨近終端速度（terminal velocity）$-g/k$，這就是它們的距離最終穩定下來的原因。

驗算

- 在 $t = 0$ 時距離為 d；而當 t 很小時，距離約為 $d + ut$，與沒有阻力時的情況一樣（那時兩個質點在重力下以相同方式下墜）。
- 若 $u = 0$，兩個質點同時由靜止釋放，運動完全相同，一直相距 d，與公式吻合。
- 單位：k 的單位是 s^{-1}（每單位質量、每單位速度的力），所以 u/k 是長度。
- 數值上，以龍格－庫塔法（Runge–Kutta method）對隨機選取的 k、u、d、t 逐步求解兩條運動方程，結果與公式吻合至約 10^{-12}。取 $k = 0.5$、$u = 10$、$d = 3$，距離在 $t = 1$ 時為 10.87，在 $t = 20$ 時為 22.999，趨近 $d + u/k = 23$。

常見錯誤

- 把質點 2 的運動分成「上升」和「下降」兩個階段，並為阻力配上不同的正負號：這裏並無必要，而且經常引致正負號出錯。
- 當 v 已經是（負值的）向上速度時，仍把下降質點的阻力寫成 $+kv$，這樣便把正負號重複計算了。

STEP 之外

把方程相減這一招之所以奏效，是因為運動方程是線性的：兩個解之差滿足齊次方程（homogeneous equation）$\dot{w} = -kw$。這就是疊加原理（superposition principle），它貫穿整套線性微分方程（linear differential equations）理論，其中每個解都是「特解加餘函數」（particular solution plus complementary function）。若阻力與速率的平方成正比（quadratic drag），這個原理便不再成立，兩者的距離會有截然不同的表現。Boyce 與 DiPrima 的 *Elementary Differential Equations*（微分方程入門教材）闡述線性理論；Kleppner 與 Kolenkow 的 *An Introduction to Mechanics*（力學入門教材）則討論兩種阻力下的運動。

詳細解答

STEP II 1995，第 12 題

概率與統計 • 連續決鬥

> **STEP II 1995，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/95-s2-q12

本題在 *Beyond STEP* 第 12.13 節三人決鬥中有完整的引導式講解：https://beyondstep.keithmcnulty.org/zh-hant/ch12/truel.html

關鍵思路

在一場持續到有人被擊中為止的決鬥中，每當雙方都擲不中，遊戲便「重新開始」。因此所求的概率（probability）滿足一條以它本身表示的簡單方程，根本不必求無窮級數（infinite series）之和（當然你也可以這樣做）。至於三人決鬥，Algernon 第一擲的每個選擇都很快變成二人決鬥，而第一個結果正好用來處理這些情況。出人意表的是，Algernon 的最佳一着竟是誰也不擊中。

解答

二人決鬥

設 P 為 Percy 先擲時他勝出的概率。考慮第一對投擲：

- Percy 第一擲便擊中（概率為 p），他立即勝出；
- Percy 擲不中，然後 Rodney 擊中（概率為 $(1-p)r$），Percy 落敗；
- 兩人都擲不中（概率為 $(1-p)(1-r)$），於是回到起點，仍由 Percy 先擲。各次投擲互相獨立（independent），所以由此開始 Percy 勝出的概率仍然是 P。

因此

$$P = p + (1-p)(1-r)P \quad\Longrightarrow\quad P\big(1-(1-p)(1-r)\big) = p \quad\Longrightarrow\quad P = \frac{p}{p+r-pr},$$

這是因為 $1-(1-p)(1-r)=p+r-pr$，而由於 $p>0$，此式為正數。這就是所要求的結果。

想一想 遊戲是否必定會結束，使得 Percy 落敗確實等同於 Rodney 勝出？

是的。首 n 回合全部以雙方擲不中告終的概率是 $\big((1-p)(1-r)\big)^n$，由於 $p>0$，它趨向 0。換一個做法，求等比級數（geometric series）$\sum_{k\geq 0} p\big((1-p)(1-r)\big)^k$ 之和，也會得到相同的答案。

在三人決鬥中，A、B 和 C 擊中的概率分別為 $\frac{2}{5}$、$\frac{3}{5}$ 和 1。由 A 先擲的 A、B 二人決鬥會出現兩次，所以我們現在先用上面的結果，取 $p=\frac{2}{5}$、$r=\frac{3}{5}$ 把它算出來：

$$\frac{\frac{2}{5}}{\frac{2}{5}+\frac{3}{5}-\frac{6}{25}}=\frac{\frac{2}{5}}{\frac{19}{25}}=\frac{10}{19}.$$

第 (i) 部分

若 Algernon 擊中 Cuthbert，便只剩下 Algernon 和 Bertie，接着輪到 Bertie 擲，目標是 Algernon。只有當 Bertie 擲不中（概率為 $\frac{2}{5}$）時，Algernon 才能留在場上。之後便是由 A 先擲的 A、B 二人決鬥，Algernon 勝出的概率為 $\frac{10}{19}$。所以

$$\mathrm{P}(\text{A 勝出})=\frac{2}{5}\cdot\frac{10}{19}=\frac{4}{19}.$$

第 (ii) 部分

若 Algernon 擊中 Bertie，接着輪到 Cuthbert 擲，而他只剩 Algernon 一個目標。Cuthbert 從不失手，所以 Algernon 勝出的機會是 0。

第 (iii) 部分

若 Algernon 擲不中，三人仍然全部在場。接着輪到 Bertie 擲，他瞄準 Cuthbert。

- Bertie 有 $\frac{3}{5}$ 的概率擊中 Cuthbert。這時剩下 Algernon 和 Bertie，輪到 Algernon 擲，Algernon 勝出的概率為 $\frac{10}{19}$。
- Bertie 有 $\frac{2}{5}$ 的概率擲不中。Cuthbert 隨即瞄準 Bertie，並且必定擊中他。這時剩下 Algernon 和 Cuthbert，輪到 Algernon 擲。Algernon 這一擲必須擊中，否則 Cuthbert 便會擊中他。所以他勝出的概率為 $\frac{2}{5}$。

因此

$$\mathrm{P}(\text{A 勝出})=\frac{3}{5}\cdot\frac{10}{19}+\frac{2}{5}\cdot\frac{2}{5}=\frac{6}{19}+\frac{4}{25}=\frac{150+76}{475}=\frac{226}{475}.$$

無論這次擲不中是怎樣造成的，是故意還是瞄準後失手，結果都一樣，因為 Bertie 和

Cuthbert 其後的打法並無分別。

Algernon 的最佳策略

> **想一想** Algernon 不能選擇擊中某人，只能選擇瞄準某人。每個選擇實際上給他帶來甚麼？

若他瞄準某人，他擊中的概率為 $\frac{2}{5}$，擲不中的概率為 $\frac{3}{5}$，而擲不中便歸入第 (iii) 部分的情況。所以：

$$\begin{aligned}
\text{瞄準 Cuthbert：}\quad & \frac{2}{5}\cdot\frac{4}{19}+\frac{3}{5}\cdot\frac{226}{475}=\frac{200+678}{2375}=\frac{878}{2375}\approx 0.370,\\
\text{瞄準 Bertie：}\quad & \frac{2}{5}\cdot 0+\frac{3}{5}\cdot\frac{226}{475}=\frac{678}{2375}\approx 0.285,\\
\text{故意擲不中：}\quad & \frac{226}{475}=\frac{1130}{2375}\approx 0.476.
\end{aligned}$$

即使必定擊中 Cuthbert，也只能得到 $\frac{4}{19}=\frac{100}{475}$，比 $\frac{226}{475}$ 還要小。擊中任何一個對手都只會降低他的勝算。因此 Algernon 應該**第一擲故意擲不中**。這樣 Bertie 和 Cuthbert 便會互相攻擊，而 Algernon 可以對倖存的一方先擲。採用這個策略，他勝出的概率為 $\frac{226}{475}$，正如題目所求。

驗算

- 若 $r=0$，決鬥公式給出 1；若 $p=1$，它也給出 1，兩者都合乎預期。
- 三個選擇的排序為 $0.476>0.370>0.285$。以蒙地卡羅模擬（Monte Carlo simulation）模擬整場三人決鬥（每種策略 3×10^5 局），得出 0.476、0.369 和 0.284，與精確值吻合。
- 以第 (iii) 部分作合理性檢查：$\frac{6}{19}\approx 0.316$ 與 $\frac{4}{25}=0.16$ 相加得 0.476。

常見錯誤

- 弄錯第一擲之後的投擲次序。在第 (i) 部分，接着擲的是 Bertie，所以要待 Bertie 擲不中之後，決鬥公式才適用。
- 把「瞄準 Cuthbert」當作「擊中 Cuthbert」。第 (i)–(iii) 部分都是以結果為條件的，制訂策略時必須分別以 $\frac{2}{5}$ 和 $\frac{3}{5}$ 為它們加權。

STEP 之外

把一個概率寫成走了一步之後以它本身表示的式子，這種方法稱為首步分析（first-step analysis）。它是研究馬可夫鏈（Markov chain）的基本工具，可把到達概率（hitting probability）和期望到達時間（expected hitting time）表示為線性方程組的解。「三人決鬥」（truel）是博弈論（game theory）中一道著名的難題，當中最弱的參與者往往置身事外才最有利。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*（概率論導引）。

詳細解答

STEP II 1995，第 13 題

概率與統計 • 泊松分佈的最優化

> **STEP II 1995，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/95-s2-q13

關鍵思路

儘管題庫把它歸入另一個類別，這其實是一道披着商業外衣的概率題：缺席乘客的人數是一個二項（binomial）計數，試驗次數很多而成功概率極小，這正是泊松分佈（Poisson distribution）登場的時候。之後，要對整數 k 求最優值，竅門不在微積分，而在一階差分（first difference）$u_{k+1}-u_k$：只要多賣一張機票能令期望利潤（expected profit）增加，就繼續賣下去。一旦證明了這個差分是遞減的，最佳的 k 就是它首次不再為正的地方。

解答

泊松模型

假設 $N+k$ 名持票人各自獨立地以概率 ϵ 缺席。那麼缺席人數 R 服從二項分佈 $\mathrm{B}(N+k,\epsilon)$，其平均值為 $(N+k)\epsilon$。由於 k 與 N 相比非常小，這個平均值近似等於 $N\epsilon=\lambda$。試驗次數很多而成功概率極小的二項分佈，可以用平均值相同的泊松分佈很好地近似：對固定的 r，當 $n\to\infty$ 而 $n\epsilon=\lambda$ 保持不變時，

$$\binom{n}{r}\epsilon^r(1-\epsilon)^{n-r}=\frac{n(n-1)\cdots(n-r+1)}{n^r}\cdot\frac{\lambda^r}{r!}\Big(1-\frac{\lambda}{n}\Big)^{n-r}\;\to\;\frac{\lambda^r}{r!}\,\mathrm{e}^{-\lambda},$$

因為第一個因子趨向 1，而 $(1-\lambda/n)^n\to\mathrm{e}^{-\lambda}$。所以近似地有 $R\sim\mathrm{Po}(\lambda)$。從現在起，我們把這個分佈當作精確成立，並記

$$p_r=\mathrm{P}(R=r)=\frac{\lambda^r}{r!}\,\mathrm{e}^{-\lambda}.$$

利潤

航空公司售出 $N+k$ 張機票，每張利潤為 £A，所以無論發生甚麼事，這部分都是 $A(N+k)$。若有 r 名乘客缺席，則有 $N+k-r$ 人到場，爭奪 N 個座位。

- 若 $r \geq k$，到場的人最多只有 N 個，人人都可以登機：無人需要賠償，而 $\max(0, k-r) = 0$。
- 若 $r < k$，則有 $k-r$ 人無法登機，每人令公司損失 £B，而 $\max(0, k-r) = k-r$。

兩種情況下，以英鎊計算的利潤都是 $A(N+k) - B\max(0, k-r)$。

期望利潤 $u_0, \ldots, u_3$

取期望值時，只有 $r = 0, 1, \ldots, k-1$ 對罰款有貢獻，所以

$$u_k = A(N+k) - B\sum_{r=0}^{k-1}(k-r)\,p_r.$$

代入 $p_0 = \mathrm{e}^{-\lambda}$、$p_1 = \lambda\mathrm{e}^{-\lambda}$ 和 $p_2 = \frac{1}{2}\lambda^2\mathrm{e}^{-\lambda}$：

$$\begin{aligned}
u_0 &= AN,\\
u_1 &= A(N+1) - B\mathrm{e}^{-\lambda},\\
u_2 &= A(N+2) - B\mathrm{e}^{-\lambda}(2+\lambda),\\
u_3 &= A(N+3) - B\mathrm{e}^{-\lambda}\left(3 + 2\lambda + \tfrac{1}{2}\lambda^2\right).
\end{aligned}$$

差分 v_k

把一般公式相減，

$$\begin{aligned}
v_k = u_{k+1} - u_k &= A - B\left[\sum_{r=0}^{k}(k+1-r)\,p_r - \sum_{r=0}^{k-1}(k-r)\,p_r\right]\\
&= A - B\left[p_k + \sum_{r=0}^{k-1}\big((k+1-r) - (k-r)\big)p_r\right]\\
&= A - B\sum_{r=0}^{k} p_r,
\end{aligned}$$

其中第一個和式裏 $r = k$ 的一項 $(k+1-k)p_k = p_k$，在第二個和式中沒有對應的項。用文字來說，就是 $v_k = A - B\,\mathrm{P}(R \leq k)$。這是合理的：第 $(k+1)$ 張額外機票必定賺得 £A，而它恰好在該乘客最終被拒登機時才多花 £B，這種情況在 $R \leq k$ 時發生。

想一想 為甚麼方括號內的式子會化簡成單純的概率之和？

對每個 $r \le k-1$，當 k 增加一時，p_r 的係數恰好增加一（在該結果下多了一名被拒登機的乘客），而結果 $r = k$ 則開始要付一筆賠償。所以由 0 到 k 的每個 r 都各貢獻一次 p_r。

於是

$$v_k - v_{k+1} = B\,p_{k+1} = B\,\frac{\lambda^{k+1}}{(k+1)!}\,\mathrm{e}^{-\lambda} > 0,$$

因為 $B > 0$ 且 $\lambda > 0$。所以對每個 k 都有 $v_k > v_{k+1}$：差分嚴格遞減。

最後，$\sum_{r=0}^{k} p_r$ 是 $\mathrm{e}^{-\lambda}\sum_{r\ge 0}\lambda^r/r! = \mathrm{e}^{-\lambda}\mathrm{e}^{\lambda} = 1$ 的部分和（partial sum），所以當 $k \to \infty$ 時它趨向 1，而

$$v_k \to A - B \qquad \text{當} k \to \infty,$$

正如題目所求。由於 $B > A$，這個極限是負數。

給 Fly By Night 的建議

記 $F(k) = \mathrm{P}(R \le k) = \sum_{r=0}^{k} p_r$，則 $v_k = A - BF(k)$，而

$$u_{k+1} > u_k \iff v_k > 0 \iff F(k) < \frac{A}{B}.$$

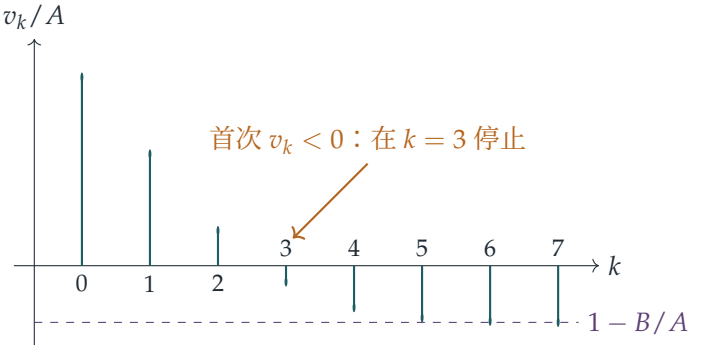
因為 v_k 嚴格遞減而極限為負，數列 $v_0, v_1, v_2, \ldots$ 在某一點之前為正，其後為負（或有一次為零）。所以 u_k 先升後降，其最大值出現在第一個令下一張機票不再有利可圖的 k：

$$\boxed{\text{選取最小的} k \text{ 使得} F(k) = \sum_{r=0}^{k} \frac{\lambda^r}{r!}\mathrm{e}^{-\lambda} \;\ge\; \frac{A}{B}.}$$

特別地，若 $\mathrm{e}^{-\lambda} \ge A/B$（乘客非常可靠，或罰款非常重），則 $v_0 \le 0$，航空公司完全不應超賣：$k = 0$。若恰好有 $F(k) = A/B$，則 $v_k = 0$，而 k 與 $k+1$ 給出相同的期望利潤。由於 $A/B < 1$ 且 $F(k) \to 1$，這樣的 k 必定存在。

所以建議是：只要最多有 k 人缺席的概率仍小於機票利潤與賠償額之比 A/B，就繼續多賣機票；到了第一個令這個概率達到 A/B 的 k 就停止。

例如，取 $\lambda = 2$ 和 $B = 1.25A$（即 $A/B = 0.8$），可得 $F(0) \approx 0.135$、$F(1) \approx 0.406$、$F(2) \approx 0.677$ 和 $F(3) \approx 0.857$，所以最佳選擇是 $k = 3$。下圖顯示這個情況下的 v_k/A。



驗算

- 由一般公式，$u_1 - u_0 = A - B\mathrm{e}^{-\lambda} = A - Bp_0$ 及 $u_2 - u_1 = A - B\mathrm{e}^{-\lambda}(1+\lambda) = A - B(p_0 + p_1)$，與 v_0 和 v_1 吻合。
- 極限情況：若 $B \to A$，則 $A/B \to 1$，k 無止境地增大（拒載幾乎沒有代價）；若 $B \to \infty$，則 $A/B \to 0$，$k = 0$。
- 數值上，我們對數百組隨機選取的 N, A, B, λ，直接由利潤公式計算 u_k，並確認了 $u_0, \ldots, u_3$ 和 v_k 的公式、$v_k - v_{k+1} = Bp_{k+1}$，以及令 u_k 最大的 k 正是滿足 $F(k) \geq A/B$ 的最小 k。當 $N = 400$、$k = 3$、$\epsilon = 0.01$ 時，有 $r = 3$ 人缺席的二項概率和泊松概率分別是 0.194 和 0.195。

常見錯誤

- 把 $r \geq k$ 的項也計入罰款的和式：這些項毫無貢獻，因為此時每個到場的人都有座位。
- 把 u_k 對 k 求導。這個變量是整數，差分 v_k 才是合適的工具。
- 把答案寫成「$v_k = 0$」。一般而言沒有任何 k 能令 v_k 恰好為零；答案是第一個令 $v_k \leq 0$ 的 k。

STEP 之外

「差分遞減，所以在它變號之處停止」這個論證，是凹函數（concave function）在導數為零處取得最大值的離散版本。「當 $\mathrm{P}(R \leq k)$ 達到 A/B 時停止」這條法則，就是運籌學（operations research）中經典的報童（newsvendor，又稱臨界分位數 critical fractile）法則，而航空公司超賣機票正是它的一個標準應用。上文嚴格證明了的二項分佈的泊松極限，是大學概率課程中最早遇到的極限定理（limit theorem）之一。Grimmett 與 Welsh 的 *Probability: An Introduction*（概率論入門教材）既涵蓋泊松近似，也討論像 $\max(0, k - R)$ 這類函數的期望值。

詳細解答

STEP II 1995，第 14 題

概率與統計 • 假設檢驗

> **STEP II 1995，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/95-s2-q14

關鍵思路

把密度寫成 $x^2\phi(x)$，其中 ϕ 是標準正態密度（standard normal density）。這樣一來，我們需要的每一個積分都變成標準正態變量 Z 的矩（moment），題目中關於 $\mathrm{E}(Z^4)$ 的提示正是用在這裏。要判斷數據是否「動搖我們的信心」，我們要問：假如我們的假設（hypothesis）（$\mu = 0$）成立，出現至少這樣極端的結果的概率有多大？如果這個概率極小，而把 μ 移動一下就能令數據變得平平無奇，那麼數據就對我們的假設不利。

解答

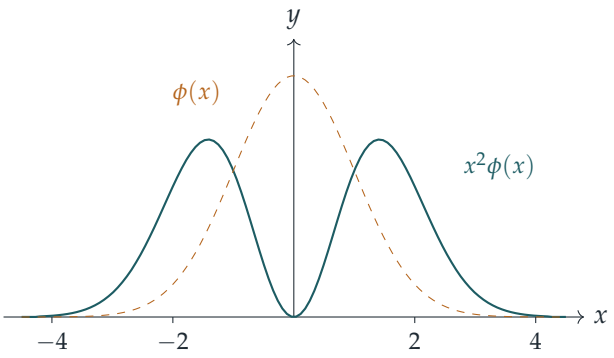
求 A

用分部積分法（integration by parts），取 $u = x$ 及 $\mathrm{d}v = x\mathrm{e}^{-x^2/2}\,\mathrm{d}x$，於是 $v = -\mathrm{e}^{-x^2/2}$：

$$\int_{-\infty}^{\infty} x^2\mathrm{e}^{-x^2/2}\,\mathrm{d}x = \Big[-x\mathrm{e}^{-x^2/2}\Big]_{-\infty}^{\infty} + \int_{-\infty}^{\infty} \mathrm{e}^{-x^2/2}\,\mathrm{d}x = 0 + \sqrt{2\pi},$$

這裏用了標準正態積分。總概率必須等於 1，所以

$$A = \frac{1}{\sqrt{2\pi}}, \qquad \text{即} \qquad \mathrm{f}(x) = x^2\phi(x), \quad \phi(x) = \tfrac{1}{\sqrt{2\pi}}\mathrm{e}^{-x^2/2}.$$



這個密度是對稱的，在 $x = \pm\sqrt{2}$ 處有兩個峯，而在 $x = 0$ 處為零。由對稱性可知 $\mathrm{E}(X) = 0$，並且

$$\mathrm{E}(X^2) = \int_{-\infty}^{\infty} x^4\phi(x)\,\mathrm{d}x = \mathrm{E}(Z^4) = 3, \qquad \text{所以} \qquad \mathrm{Var}(X) = 3.$$

第 (i) 部分

在我們的假設之下，出現像 $a = 87.3$ 這樣偏遠的結果，可能性有多大？像之前一樣用分部積分法（這次 $v = -\phi(x)$，因為 $\phi'(x) = -x\phi(x)$），

$$\mathrm{P}(X \geq a) = \int_a^{\infty} x \cdot x\phi(x)\,\mathrm{d}x = a\phi(a) + \mathrm{P}(Z \geq a).$$

當 $a > 0$ 時，我們還有 $\mathrm{P}(Z \geq a) = \int_a^{\infty} \phi(x)\,\mathrm{d}x \leq \int_a^{\infty} \frac{x}{a}\phi(x)\,\mathrm{d}x = \frac{\phi(a)}{a}$，因為在積分範圍內 $x/a \geq 1$。因此，由對稱性，

$$\mathrm{P}(|X| \geq 87.3) \leq 2\left(a + \frac{1}{a}\right)\phi(a) \approx 70\,\mathrm{e}^{-3810.6} \approx 10^{-1653}.$$

這個數值小得荒謬：假如我們的假設成立，我們基本上永遠不會見到這樣的數值。另一方面，如果 μ 大約是 85 或 89（使得 87.3 正好落在 $\mathrm{f}(x - \mu)$ 的其中一個峯上），那麼 87.3 這個結果就完全平常。所以，單單一個 87.3 的結果，就應該徹底動搖我們的信心。唯一的出路是懷疑結果本身（例如記錄出錯），而明智的科學家在放棄假設之前，會先查證這一點。

? **想一想** 每一個個別數值的概率都是零，只憑一次觀測就下判斷，豈不是不公平？

重點不在於恰好取得 87.3 這個數值的概率，而在於取得一個離我們的假設所集中的位置至少這麼遠的數值的概率。同樣地，在我們的假設之下，87.3 處的密度大約是某個合適的備擇假設之下該處密度的 10^{-1650} 倍。無論從哪個角度看，觀測結果極端到這個程度，一次就綽綽有餘。

第 (ii) 部分

設 $X_1, \ldots, X_{1000}$ 為各個獨立的結果，$\bar{X}$ 為它們的平均值。在我們的假設之下，每個 X_i 的平均值為 0，方差（variance）為 3，所以

$$\mathrm{E}(\bar{X}) = 0, \qquad \mathrm{Var}(\bar{X}) = \frac{3}{1000}, \qquad \mathrm{sd}(\bar{X}) = \sqrt{0.003} \approx 0.0548.$$

方差是有限的，而 1000 又夠大，所以根據中心極限定理（central limit theorem），$\bar{X}$ 近似服從 $\mathrm{N}(0, 0.003)$，儘管 f 本身看起來一點也不像正態密度。觀測所得的平均值與 0 相距

$$\frac{0.23}{0.0548} \approx 4.2$$

個標準差（standard deviation），而

$$\mathrm{P}(|\bar{X}| \geq 0.23) \approx \mathrm{P}(|Z| \geq 4.2) \approx 3 \times 10^{-5}.$$

所以，假如我們的假設成立，把整個實驗重複十萬次，樣本平均值離 0 這麼遠的情況大約只會出現三次。而在 $\mu \approx 0.23$ 的備擇假設之下，同樣的數據就完全典型。這同樣應該嚴重動搖我們的信心，只是不及 (i) 那樣具壓倒性。在任何常用的顯著性水平（significance level）下，例如 1% 甚至 0.1%，我們都會否定 $\mu = 0$。

> **想一想** 我們可否改用切比雪夫不等式（Chebyshev's inequality），從而避開中心極限定理？

切比雪夫不等式給出 $\mathrm{P}(|\bar{X}| \geq 0.23) \leq 0.003/0.23^2 \approx 0.057$，當中不涉及任何近似。這個結果有參考價值，但不足以下定論，因為這樣只能在大約 6% 的水平上排除如此極端的結果。中心極限定理給出的數字要精確得多；而且有 1000 個觀測值，分佈的尾部又像 $\mathrm{e}^{-x^2/2}$ 那樣衰減，這個近似是可以信賴的。

驗算

- 辛普森法則（Simpson's rule）給出 $\int x^2 \mathrm{e}^{-x^2/2}\,\mathrm{d}x = 2.506628\ldots = \sqrt{2\pi}$，以及 $\int x^2 \mathrm{f}(x)\,\mathrm{d}x = 3.000$。
- 尾部公式 $\mathrm{P}(|X| \geq a) = 2(a\phi(a) + \mathrm{P}(Z \geq a))$ 在 $a = 3$ 時與數值積分的結果吻合（兩者皆為 0.029291）。
- $|X|$ 的密度，與由三個獨立 $\mathrm{N}(0,1)$ 變量組成的向量的長度的密度相同。用這個方法模擬 X，每次取 1000 個數值求平均，重複 4000 次，結果約有 4.9% 的平均值落在 2 sd 之外，接近正態分佈所預測的 4.6%，而且沒有一個超出 0.23。

常見錯誤

- 因為指數部分看似標準正態分佈，就取 $\mathrm{Var}(X) = 1$。額外的 x^2 把概率質量向外推，令方差變成 3。
- 在 (ii) 中，對單一個 X 而不是對 $\bar{X}$ 作正態檢驗，或者忘記把方差除以 1000。
- 在 (i) 中，辯稱「一次觀測證明不了甚麼」。結果極端到這個程度，一次觀測就足以說明問題。

STEP 之外

這是從基本原理出發看假設檢驗（hypothesis testing）。大學一年級的統計學課程會用 p 值（p-value）、似然比（likelihood ratio）（奈曼–皮爾遜引理，Neyman–Pearson lemma）和中心極限定理把它形式化。順帶一提，$|X|$ 服從自由度（degrees of freedom）為 3 的卡分佈（chi distribution），亦即氣體中分子速率的麥克斯韋–玻爾茲曼分佈（Maxwell–Boltzmann distribution）。Grimmett 與 Welsh 合著的 *Probability: An Introduction*（概率論入門教材）涵蓋了這裏用到的極限定理。

詳細解答

STEP III 1995，第 1 題

純數學 • 聯立線性方程

STEP III 1995，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/95-s3-q1

關鍵思路

第二條和第三條方程的 y 項和 z 項完全相同，所以兩式相減便可一次過消去兩個未知數，只剩下一條關於 x 的方程。這道題的重點，大部分在於找出 a 和 b 取哪些值時，這一步或其後的某次除法會失效。係數的行列式（determinant）能預先告訴我們這些值在哪裏，所以我們先把它算出來，然後逐一處理每個例外情況。

解答

把方程編號為

$$a^2x + ay + z = a^2 \tag{1}$$

$$ax + y + bz = 1 \tag{2}$$

$$a^2bx + y + bz = b. \tag{3}$$

哪裏會出問題？

把係數矩陣（coefficient matrix）的行列式沿第一行展開，

$$\begin{vmatrix} a^2 & a & 1 \\ a & 1 & b \\ a^2b & 1 & b \end{vmatrix} = a^2(b-b) - a(ab - a^2b^2) + (a - a^2b)$$

$$= a^3b^2 - 2a^2b + a = a(ab-1)^2.$$

所以除非 $a = 0$ 或 $ab = 1$，否則方程組恰有一個解。在這兩種情況下，方程組要麼無解，要麼有無限多個解，我們必須回到方程本身才能判斷是哪一種。

一般情況：$a \neq 0$ 且 $ab \neq 1$

由 (3) 減去 (2)，便可消去 y 和 z：

$$(a^2b - a)x = b - 1, \qquad \text{即} \qquad a(ab-1)\,x = b - 1.$$

由於 $a \neq 0$ 且 $ab \neq 1$，我們可以相除，得

$$x = \frac{1-b}{a(1-ab)}.$$

接着，由 (1) 得 $z = a^2 - a^2x - ay$。把它代入 (2)：

$$ax + y + a^2b - a^2bx - aby = 1 \quad \Longrightarrow \quad (1-ab)\,y = (1 - a^2b) - ax(1-ab).$$

除以 $1 - ab \neq 0$，並利用 $ax = \dfrac{1-b}{1-ab}$，

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

最後，

$$\begin{aligned} z = a^2 - a(ax) - ay &= a^2 - \frac{a(1-b)}{1-ab} - \frac{ab(1-a^2)}{1-ab} \\ &= \frac{a^2 - a^3b - a + ab - ab + a^3b}{1-ab} = \frac{a(a-1)}{1-ab}. \end{aligned}$$

所以當 $a \neq 0$ 且 $ab \neq 1$ 時，唯一解為

$$\boxed{x = \frac{1-b}{a(1-ab)}, \qquad y = \frac{b(1-a^2)}{1-ab}, \qquad z = \frac{a(a-1)}{1-ab}.}$$

$a = 0$ 的情況

方程變成

$$z = 0, \qquad y + bz = 1, \qquad y + bz = b,$$

所以 $z = 0$，$y = 1$，而最後一條方程則表示 $1 = b$。因此：

- 若 $a = 0$ 且 $b \neq 1$，方程組**無解**；
- 若 $a = 0$ 且 $b = 1$，解為 $x = \lambda$，$y = 1$，$z = 0$，其中 λ 為任意實數：這些解構成一條直線，因為 x 已不再出現在任何一條方程之中。

$ab = 1$ 的情況（故 $a \neq 0$）

想一想 當 $ab = 1$ 時，方程 $a(ab-1)x = b-1$ 的左方等於 0。這告訴我們甚麼？

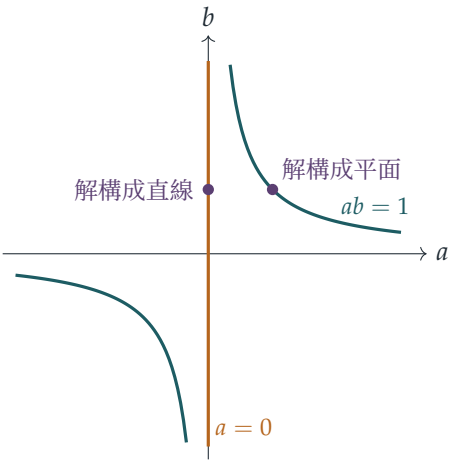
由 (3) 減去 (2) 這一步仍然成立；它現在變成 $0 = b-1$。所以若 $b \neq 1$，方程組便不相容（inconsistent）。在 $ab = 1$ 之下，條件 $b = 1$ 必然令 $a = 1$。因此：

- 若 $ab = 1$ 且 $a \neq 1$，方程組**無解**；
- 若 $a = b = 1$，三條方程全都變成 $x + y + z = 1$，所以解為 $x = \lambda$，$y = \mu$，$z = 1 - \lambda - \mu$，其中 λ, μ 為任意實數：這些解構成一整個平面。

這與行列式中的平方因子 $(ab-1)^2$ 吻合。在 $a = b = 1$ 時，矩陣的秩（rank）一直降至 1，所以有兩個自由參數；而在 $a = 0$、$b = 1$ 時，矩陣的秩為 2，只有一個自由參數。

總結

下圖所示為 (a, b) 平面。在直線 $a = 0$ 和雙曲線（hyperbola）$ab = 1$ 以外，方程組有唯一解。在這兩條線上，除了兩個標示出來的點之外，方程組無解。



- $a \neq 0$，$ab \neq 1$：上面方框內的唯一解。
- $a = 0$，$b = 1$：$(x, y, z) = (\lambda, 1, 0)$，$\lambda \in \mathbb{R}$。
- $a = b = 1$：$(x, y, z) = (\lambda, \mu, 1 - \lambda - \mu)$，$\lambda, \mu \in \mathbb{R}$。
- $a = 0$ 而 $b \neq 1$，或 $ab = 1$ 而 $a \neq 1$：無解。

驗算

取 $a=2$，$b=3$。公式給出 $x=\frac{1}{5}$，$y=\frac{9}{5}$，$z=-\frac{2}{5}$，而確實有 $4\cdot\frac{1}{5}+2\cdot\frac{9}{5}-\frac{2}{5}=4=a^2$，$\frac{2}{5}+\frac{9}{5}-\frac{6}{5}=1$ 以及 $\frac{12}{5}+\frac{9}{5}-\frac{6}{5}=3=b$。另一個快捷的測試是取 $b=1$，$a\neq 0,1$：公式給出 $x=0$，$y=1+a$，$z=-a$，而這些值滿足全部三條方程。我們亦以精確的有理數運算，對 2000 組隨機選取的 (a,b) 檢驗過行列式和一般公式。

常見錯誤

- 未有先把 $a=0$ 和 $ab=1$ 分開處理，便除以 $a(ab-1)$。題目說「對所有可能的值」，意思就是特殊情況正是這道題的重點所在。
- 聲稱「行列式為零，所以無解」。行列式為零只排除了唯一解的可能；在這道題中，有兩組特殊的參數值會給出無限多個解。
- 忽略了 $ab=1$ 與 $b=1$ 同時成立必然令 $a=1$，所以雙曲線上只有一個特殊點。

STEP 之外

這是一個小例子，說明矩陣的秩如何支配一個線性方程組（linear system）。方程組相容，當且僅當係數矩陣與增廣矩陣（augmented matrix）的秩相同；這時所有解構成一個維數（dimension）為 $3-$ 秩 的集合。這就是為甚麼我們得到的是單獨一點、一條直線或一個平面。高斯消去法（Gaussian elimination）令這一切變得有系統，而 Strang 的 *Introduction to Linear Algebra*（線性代數入門）則以大量例題詳細闡述這個課題。

詳細解答

STEP III 1995，第 2 題

純數學 • 歸約公式

STEP III 1995，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/95-s3-q2

關鍵思路

因子 $\sqrt{x}$ 和 $\sqrt{a-x}$ 都提示我們使用代換（substitution）$x = a\sin^2\theta$，這樣兩者會同時化為 $\sqrt{a}\sin\theta$ 和 $\sqrt{a}\cos\theta$。（求 I_0 還有一條巧妙的幾何捷徑。）至於歸約公式（reduction formula），可用分部積分法（integration by parts），使 x 的冪次降一。這時 $a-x$ 的冪次會升一，我們再利用 $a-x$ 把多出的一個因子拆開。這樣便會重新得到 I_n 本身，於是可以合併同類項。

全題設 $a > 0$。

解答

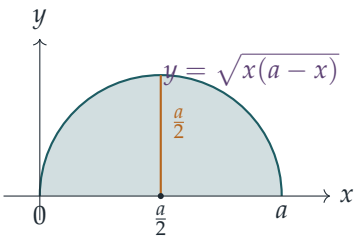
求 I_0

我們要求 $\mathrm{I}_0 = \int_0^a x^{1/2}(a-x)^{1/2}\,\mathrm{d}x$。設 $x = a\sin^2\theta$，則 $\mathrm{d}x = 2a\sin\theta\cos\theta\,\mathrm{d}\theta$。當 x 由 0 變到 a 時，θ 由 0 變到 $\frac{\pi}{2}$，而在這個範圍內 $\sin\theta, \cos\theta \geq 0$。所以 $\sqrt{x} = \sqrt{a}\sin\theta$ 及 $\sqrt{a-x} = \sqrt{a}\cos\theta$，不會有正負號的問題。於是

$$\begin{aligned}\mathrm{I}_0 &= \int_0^{\pi/2} (\sqrt{a}\sin\theta)(\sqrt{a}\cos\theta)\,2a\sin\theta\cos\theta\,\mathrm{d}\theta = 2a^2\int_0^{\pi/2}\sin^2\theta\cos^2\theta\,\mathrm{d}\theta \\ &= \frac{a^2}{2}\int_0^{\pi/2}\sin^2 2\theta\,\mathrm{d}\theta = \frac{a^2}{4}\int_0^{\pi/2}(1-\cos 4\theta)\,\mathrm{d}\theta = \frac{a^2}{4}\cdot\frac{\pi}{2} = \frac{\pi a^2}{8},\end{aligned}$$

這正是所求的結果。這裏用到了 $\sin\theta\cos\theta = \frac{1}{2}\sin 2\theta$ 和 $\sin^2 2\theta = \frac{1}{2}(1-\cos 4\theta)$，而 $\cos 4\theta$ 在 $[0, \frac{\pi}{2}]$ 上的積分為零。

一幅圖。若 $y = \sqrt{x(a-x)}$，則 $y^2 = ax - x^2$，整理後得 $\left(x - \frac{a}{2}\right)^2 + y^2 = \left(\frac{a}{2}\right)^2$，其中 $y \geq 0$。所以 I_0 就是半徑為 $\frac{a}{2}$ 的半圓的面積，即 $\frac{1}{2}\pi\left(\frac{a}{2}\right)^2 = \frac{\pi a^2}{8}$。



歸約公式

取 $n \geq 1$。以 $u = x^{n+\frac{1}{2}}$ 和 $\mathrm{d}v = (a-x)^{1/2}\,\mathrm{d}x$ 作分部積分。則 $\mathrm{d}u = \left(n+\frac{1}{2}\right)x^{n-\frac{1}{2}}\,\mathrm{d}x$ 及 $v = -\frac{2}{3}(a-x)^{3/2}$。由此得

$$\mathrm{I}_n = \left[-\tfrac{2}{3}x^{n+\frac{1}{2}}(a-x)^{3/2}\right]_0^a + \frac{2}{3}\left(n+\frac{1}{2}\right)\int_0^a x^{n-\frac{1}{2}}(a-x)^{3/2}\,\mathrm{d}x.$$

邊界項在兩端都等於零：在 $x = a$ 處是因為有因子 $(a-x)^{3/2}$，在 $x = 0$ 處則是因為有 $x^{n+\frac{1}{2}}$。留意 $\frac{2}{3}\left(n+\frac{1}{2}\right) = \frac{2n+1}{3}$。

想一想 新的積分含有 $(a-x)^{3/2}$，並不是 I_{n-1} 的形式。我們該怎樣處理它？

把它寫成 $(a-x)^{3/2} = (a-x)\,(a-x)^{1/2}$，然後展開第一個因子：

$$\int_0^a x^{n-\frac{1}{2}}(a-x)^{3/2}\,\mathrm{d}x = a\int_0^a x^{n-\frac{1}{2}}(a-x)^{1/2}\,\mathrm{d}x - \int_0^a x^{n+\frac{1}{2}}(a-x)^{1/2}\,\mathrm{d}x = a\,\mathrm{I}_{n-1} - \mathrm{I}_n.$$

所以

$$\mathrm{I}_n = \frac{2n+1}{3}\left(a\,\mathrm{I}_{n-1} - \mathrm{I}_n\right) \quad\Longrightarrow\quad (2n+4)\,\mathrm{I}_n = (2n+1)\,a\,\mathrm{I}_{n-1},$$

這是兩邊乘以 3 再合併 I_n 各項後所得的（$3+2n+1 = 2n+4$）。這就是所求的結果。

計算 I_n

依次對 $k = n, n-1, \ldots, 1$ 應用 $\mathrm{I}_k = \dfrac{(2k+1)a}{2k+4}\,\mathrm{I}_{k-1}$，最後到達 I_0，得

$$\mathrm{I}_n = a^n\,\frac{\prod_{k=1}^n (2k+1)}{\prod_{k=1}^n (2k+4)}\cdot\frac{\pi a^2}{8}.$$

現在整理這兩個乘積。分子是奇數的乘積 $3\cdot 5\cdots(2n+1)$。我們補上缺少的偶數 $2\cdot 4\cdots(2n) = 2^n n!$，得

$$\prod_{k=1}^{n}(2k+1) = \frac{(2n+1)!}{2^n\,n!}.$$

分母是 $\prod_{k=1}^n 2(k+2) = 2^n\cdot 3\cdot 4\cdots(n+2) = 2^n\,\dfrac{(n+2)!}{2}$。因此

$$\mathrm{I}_n = \frac{\pi a^{n+2}}{8}\cdot\frac{(2n+1)!}{2^n\,n!}\cdot\frac{2}{2^n\,(n+2)!} = \boxed{\frac{\pi\,a^{n+2}\,(2n+1)!}{4^{n+1}\,n!\,(n+2)!}}$$

當 $n = 0$ 時，這個公式同樣給出正確的值 $\frac{\pi a^2}{8}$。

驗算

- $n = 1$：遞推關係（recurrence）給出 $6\mathrm{I}_1 = 3a\mathrm{I}_0$，所以 $\mathrm{I}_1 = \frac{\pi a^3}{16}$。公式則給出 $\pi a^3 \cdot 3!/(16 \cdot 1! \cdot 3!) = \frac{\pi a^3}{16}$，兩者吻合。
- 對稱性：把 x 換成 $a - x$，可見 $\mathrm{I}_1 = \int_0^a x^{1/2}(a-x)^{3/2}\,\mathrm{d}x$ 亦成立。把兩個表達式相加，得 $2\mathrm{I}_1 = a\mathrm{I}_0$，這就獨立地驗證了 $\mathrm{I}_1 = \frac{\pi a^3}{16}$。
- 取 $a = 1$ 作數值計算，用森遜法則（Simpson's rule）或中點法則作數值積分，得 $\mathrm{I}_0 \approx 0.392699 = \pi/8$ 及 $\mathrm{I}_3 \approx 0.085903$。公式給出 $\pi \cdot 7!/(4^4 \cdot 3! \cdot 5!) = 35\pi/1280 \approx 0.085903$。

常見錯誤

- 弄錯 v 的正負號：$\int (a-x)^{1/2}\,\mathrm{d}x = -\frac{2}{3}(a-x)^{3/2}$，其中的負號來自鏈式法則（chain rule）。
- 一心想直接得出 I_{n-1}，卻沒有留意到 I_n 在右邊再次出現，必須把它合併。
- 在 $3 \cdot 4 \cdots (n+2) = (n+2)!/2$ 中漏掉因子 2。

STEP 之外

這些積分都是貝塔函數（Beta function）$B(p,q) = \int_0^1 t^{p-1}(1-t)^{q-1}\,\mathrm{d}t$ 的特殊情況。代入 $x = at$ 可得 $\mathrm{I}_n = a^{n+2}B\left(n+\frac{3}{2},\frac{3}{2}\right)$。我們證明的遞推關係，正是恆等式 $B(p+1,q) = \frac{p}{p+q}B(p,q)$ 的一個特例。在大學一年級的課程中，你會遇到公式 $B(p,q) = \Gamma(p)\Gamma(q)/\Gamma(p+q)$，它解釋了 π 的來歷：$\Gamma\left(\frac{1}{2}\right) = \sqrt{\pi}$。Siklos 的 *Advanced Problems in Mathematics* 一書中，還有更多這類風格的歸約公式題目。

詳細解答

STEP III 1995，第 3 題

純數學 • 阻尼諧運動

STEP III 1995，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/95-s3-q3

關鍵思路

試代入 $x = \mathrm{e}^{\lambda t}$：微分方程（differential equation）便化為一條關於 λ 的二次方程，而題目的三種情況，正好對應其判別式（discriminant）的三種可能（兩個相異實根、一個重根、複數根）。在有阻尼而仍然振盪的情況下，駐點（stationary point）每隔半個周期出現一次，時間間隔相等；每經過這樣的一段時間，振盪因子只是變號，而指數因子則按相同的比例縮小。這就是相繼極值之比為常數的原因。

解答

代入 $x = \mathrm{e}^{\lambda t}$，得輔助方程（auxiliary equation）

$$\lambda^2 + 2k\lambda + 1 = 0, \qquad \lambda = -k \pm \sqrt{k^2 - 1}.$$

以下 A 和 B 均為任意常數。

第 (i) 部分：$k > 1$

兩根為相異實數，故

$$x = A\,\mathrm{e}^{(-k+\sqrt{k^2-1})\,t} + B\,\mathrm{e}^{(-k-\sqrt{k^2-1})\,t}.$$

（由於 $\sqrt{k^2-1} < k$，兩個指數均為負數：這是過阻尼（overdamping），沒有振盪。）

第 (ii) 部分：$k = 1$

輔助方程為 $(\lambda+1)^2 = 0$，有重根（repeated root）$\lambda = -1$。第二個獨立解是 $t\,\mathrm{e}^{-t}$，故

$$x = (A + Bt)\,\mathrm{e}^{-t}.$$

（驗證：若 $x = te^{-t}$，則 $x' = (1-t)e^{-t}$ 及 $x'' = (t-2)e^{-t}$，而 $(t-2)+2(1-t)+t = 0$。）

第 (iii) 部分： $0 < k < 1$

此時 $k^2 - 1 < 0$，故兩根為 $\lambda = -k \pm i\omega$，其中

$$\omega = \sqrt{1-k^2} > 0.$$

取 $e^{(-k\pm i\omega)t}$ 的實組合，得

$$x = e^{-kt}\big(A\cos\omega t + B\sin\omega t\big).$$

相繼的極值

由 $x(0) = 0$ 得 $A = 0$，故 $x = Be^{-kt}\sin\omega t$，且 $B \neq 0$（否則 $x \equiv 0$，沒有極大值或極小值可言）。求導得

$$\frac{dx}{dt} = Be^{-kt}\big(\omega\cos\omega t - k\sin\omega t\big).$$

由於 $k^2 + \omega^2 = 1$，可設 $k = \sin\varphi$、$\omega = \cos\varphi$，其中 $0 < \varphi < \frac{\pi}{2}$，於是

$$\omega\cos\omega t - k\sin\omega t = \cos(\omega t + \varphi).$$

因此 $\dfrac{dx}{dt} = Be^{-kt}\cos(\omega t + \varphi)$，它等於零當且僅當 $\omega t + \varphi$ 是 $\frac{\pi}{2}$ 的奇數倍，並且在每個這樣的時刻都會變號。所以（在 $t > 0$ 時）駐點位於

$$t_n = t_1 + \frac{(n-1)\pi}{\omega}, \qquad n = 1, 2, 3, \ldots,$$

彼此相隔 π/ω，間距相等；又因為 dx/dt 在每個駐點都變號，它們確實是極大值與極小值交替出現。

想一想 不必明確求出 t_1，怎樣比較 $x(t_{n+1})$ 與 $x(t_n)$？

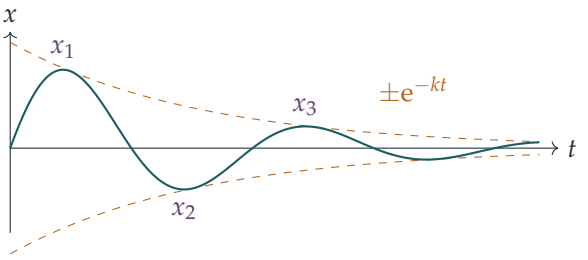
時間推前 π/ω，ωt 便增加 π，而 $\sin(\theta + \pi) = -\sin\theta$。所以

$$x_{n+1} = Be^{-k(t_n + \pi/\omega)}\sin(\omega t_n + \pi) = -e^{-k\pi/\omega}\,Be^{-kt_n}\sin\omega t_n = -e^{-k\pi/\omega}\,x_n.$$

此外 $x_n \neq 0$，因為若 $\sin\omega t_n = 0$，則必有 $\cos\omega t_n = \pm 1$，於是 $dx/dt = \pm B\omega e^{-kt_n} \neq 0$。因此

$$\left|\frac{x_{n+1}}{x_n}\right| = \alpha, \qquad \text{其中} \alpha = e^{-k\pi/\omega} = \exp\left(-\frac{k\pi}{\sqrt{1-k^2}}\right),$$

此值與 n 無關，正是所要證明的結果。（式中的負號印證了極大值與極小值交替出現。）



k^2 的公式

取對數，得 $\ln\alpha = -\dfrac{k\pi}{\sqrt{1-k^2}}$，故

$$(\ln\alpha)^2 = \frac{k^2\pi^2}{1-k^2} \implies (\ln\alpha)^2 - k^2(\ln\alpha)^2 = k^2\pi^2 \implies k^2 = \frac{(\ln\alpha)^2}{\pi^2 + (\ln\alpha)^2},$$

這就是所求的結果。

驗算

- 當 $k \to 0$（沒有阻尼）時，$\alpha \to 1$ 而 $\ln\alpha \to 0$，公式給出 $k^2 \to 0$，與預期相符：各個極值的大小全都相同。
- 當 $k \to 1^-$ 時，$\alpha \to 0$ 而 $(\ln\alpha)^2 \to \infty$，故 $k^2 \to 1$：在臨界阻尼（critical damping）下，運動不再往回擺動。
- 數值上，取 $k = 0.5$ 及 $x(0) = 0$，把方程逐步向前數值求解，得出相繼極值之比為 0.163034，與 $e^{-0.5\pi/\sqrt{0.75}} \approx 0.163034$ 吻合；而且 $(\ln\alpha)^2/(\pi^2 + (\ln\alpha)^2) = 0.25 = k^2$。

常見錯誤

- 把一個極大值與下一個極小值之間的時間當作一整個周期 $2\pi/\omega$，而不是半個周期 π/ω；這樣會得出 α^2 而不是 α。
- 以為極值出現在 $\sin\omega t = \pm 1$ 之處。其實不然：衰減因子令極值提早出現，移到 $\tan\omega t = \omega/k$ 之處。上面的論證自始至終不需要知道極值的確切位置，只需要知道它們的間距。

STEP 之外

$-\ln\alpha$ 這個量稱為對數衰減率（logarithmic decrement）：工程師從記錄下來的振盪中量度它，再用本題的公式把它換算成阻尼 k。Boyce 與 DiPrima 所著的 *Elementary Differential Equations*（初等微分方程教科書）詳細討論了過阻尼、臨界阻尼和欠阻尼三種情況。

詳細解答

STEP III 1995，第 4 題

純數學 • 三角級數求和

STEP III 1995，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/95-s3-q4

關鍵思路

餘弦之和是複指數（complex exponential）之和的實部（real part），而 $\sum e^{ik\theta}$ 是一個等比級數（geometric series）。先把它求和，再從分子和分母各抽出一個「半角」（half-angle）因子 $e^{i\theta/2}$，就能把難以處理的 $e^{im\theta}-1$ 項化為正弦，分母中的 $\sin(\frac{1}{2}\theta)$ 正是由此而來。至於不等式，訣竅是改寫 $C_n(\theta)$，使 $\frac{1}{2}$ 明顯地分離出來。

解答

和式 C_n 與 S_n

設 $z=e^{i\theta}$。由於 $0<\theta<2\pi$，我們有 $z\neq 1$，故等比級數求和公式適用：

$$C_n(\theta)+iS_n(\theta)=\sum_{k=0}^{n}e^{ik\theta}=\frac{e^{i(n+1)\theta}-1}{e^{i\theta}-1}.$$

現在把 $e^{i\phi}-1=e^{i\phi/2}\left(e^{i\phi/2}-e^{-i\phi/2}\right)=2ie^{i\phi/2}\sin(\frac{1}{2}\phi)$ 用於分子（取 $\phi=(n+1)\theta$）和分母（取 $\phi=\theta$）：

$$C_n(\theta)+iS_n(\theta)=\frac{2ie^{i(n+1)\theta/2}\sin\left(\frac{1}{2}(n+1)\theta\right)}{2ie^{i\theta/2}\sin(\frac{1}{2}\theta)}=e^{in\theta/2}\,\frac{\sin\left(\frac{1}{2}(n+1)\theta\right)}{\sin(\frac{1}{2}\theta)}.$$

因為 $0<\frac{1}{2}\theta<\pi$，分母不為零。這個分數是實數，所以取實部和虛部（利用 $e^{in\theta/2}=\cos\frac{1}{2}n\theta+i\sin\frac{1}{2}n\theta$）便得

$$C_n(\theta)=\frac{\cos(\frac{1}{2}n\theta)\sin\left(\frac{1}{2}(n+1)\theta\right)}{\sin(\frac{1}{2}\theta)},\qquad S_n(\theta)=\frac{\sin(\frac{1}{2}n\theta)\sin\left(\frac{1}{2}(n+1)\theta\right)}{\sin(\frac{1}{2}\theta)}.$$

第一式就是題目要求的結果，第二式則是 S_n 的相應表達式。

想一想 有沒有不用複數也能看出這個結果的方法？

有：把每一項乘以 $2\sin(\frac{1}{2}\theta)$，再用積化和差公式（product-to-sum formula）

$$2\sin(\tfrac{1}{2}\theta)\cos k\theta = \sin\big((k+\tfrac{1}{2})\theta\big) - \sin\big((k-\tfrac{1}{2})\theta\big).$$

由 $k=0$ 加到 n，各項裂項相消（telescope），只剩下最後一項和最初一項：

$$2\sin(\tfrac{1}{2}\theta)\,\mathrm{C}_n(\theta) = \sin\big((n+\tfrac{1}{2})\theta\big) - \sin(-\tfrac{1}{2}\theta) = \sin\big((n+\tfrac{1}{2})\theta\big) + \sin(\tfrac{1}{2}\theta). \qquad (*)$$

根據和差化積公式（sum-to-product formula）$\sin A + \sin B = 2\sin\frac{1}{2}(A+B)\cos\frac{1}{2}(A-B)$，取 $A=(n+\frac{1}{2})\theta$，$B=\frac{1}{2}\theta$，右邊等於 $2\sin\big(\frac{1}{2}(n+1)\theta\big)\cos(\frac{1}{2}n\theta)$，與上面求得的結果一致。

不等式

式 $(*)$ 正是我們想要的形式。兩邊除以 $2\sin(\frac{1}{2}\theta)\neq 0$，得

$$\mathrm{C}_n(\theta) = \frac{1}{2} + \frac{\sin\big((n+\frac{1}{2})\theta\big)}{2\sin(\frac{1}{2}\theta)}.$$

（若改由題目給出的公式出發，可用 $2\cos(\frac{1}{2}n\theta)\sin\big(\frac{1}{2}(n+1)\theta\big) = \sin\big((n+\frac{1}{2})\theta\big) + \sin(\frac{1}{2}\theta)$ 得到同一行。）因此

$$\left|\mathrm{C}_n(\theta) - \frac{1}{2}\right| = \frac{\left|\sin\big((n+\frac{1}{2})\theta\big)\right|}{2\sin(\frac{1}{2}\theta)} \le \frac{1}{2\sin(\frac{1}{2}\theta)},$$

其中我們用到了當 $0<\theta<2\pi$ 時 $\sin(\frac{1}{2}\theta)>0$（所以分母不需要加上模（modulus）號），以及 $|\sin x|\le 1$。這就是所求的不等式，而等號成立當且僅當 $\sin\big((n+\frac{1}{2})\theta\big)=\pm 1$。

想一想 為甚麼 $\sin(\frac{1}{2}\theta)$ 為正這一點很重要？

假如容許 θ 取 $(0,2\pi)$ 以外的值，$\sin(\frac{1}{2}\theta)$ 便可能是負數，不等式的右邊也會變成負數，而模不可能小於負數。在題目給定的範圍內，$\frac{1}{2}\theta\in(0,\pi)$，所以這個上界確實是一個正數。

驗算

- $n=1$：$\mathrm{C}_1 = 1+\cos\theta = 2\cos^2(\frac{1}{2}\theta)$，而公式給出 $\cos(\frac{1}{2}\theta)\sin\theta/\sin(\frac{1}{2}\theta) = 2\cos^2(\frac{1}{2}\theta)$。同樣，$\mathrm{S}_1=\sin\theta$ 與 $\sin(\frac{1}{2}\theta)\sin\theta/\sin(\frac{1}{2}\theta)$ 相符。
- 當 $\theta\to 0$ 時，C_n 的公式趨向 $(n+1)$，正是 $n+1$ 個一之和；上界 $1/(2\sin\frac{1}{2}\theta)$ 則趨向無窮大，這是必然的。

- 數值驗算，取 $n=3$，$\theta=1$：直接求和得 $\mathrm{C}_3(1)\approx 0.1342$，與公式相符，而且 $|\mathrm{C}_3-\frac{1}{2}|\approx 0.366\le 1/(2\sin\frac{1}{2})\approx 1.043$。等號成立的例子有 $n=2$，$\theta=\pi/5$，此時兩邊都等於 $1.618\ldots$。

常見錯誤

- 忘記和式由 $k=0$ 開始，所以共有 $n+1$ 項：等比級數的分子是 $e^{\mathrm{i}(n+1)\theta}$，而不是 $e^{\mathrm{i}n\theta}$。
- 沒有先把分母化為實數，就對 $\left(e^{\mathrm{i}(n+1)\theta}-1\right)/\left(e^{\mathrm{i}\theta}-1\right)$ 取實部；半角因式分解能乾淨利落地做到這一點。
- 寫下 $|\sin(\frac{1}{2}\theta)|$，卻沒有說明它為何等於 $\sin(\frac{1}{2}\theta)$。

STEP 之外

函數 $\mathrm{C}_n(\theta)-\frac{1}{2}=\sin\left((n+\frac{1}{2})\theta\right)/\left(2\sin\frac{1}{2}\theta\right)$ 除了相差一個常數因子外，就是狄利克雷核（Dirichlet kernel），它是傅立葉級數（Fourier series）理論的核心對象：一個函數的第 n 個傅立葉部分和，就是該函數對這個核取的平均。這裏證明的上界，加上分子的振盪，是證明傅立葉級數何時收斂的起點（也是用於 $\sum\cos(k\theta)/k$ 這類級數的狄利克雷判別法（Dirichlet's test）的起點）。Abbott 的 *Understanding Analysis* 有一章講傅立葉級數，寫得淺白易讀，用的正是這些和式。

詳細解答

STEP III 1995，第 5 題

純數學 • 麥克勞林級數

> **STEP III 1995，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/95-s3-q5

關鍵思路

設 $\theta = \sin^{-1} x$，於是 $\mathrm{d}\theta/\mathrm{d}x = 1/\sqrt{1-x^2}$。把 $y^{(1)}$ 乘以 $\sqrt{1-x^2}$，便可消去這個礙事的因子；再求導一次，就得到不含平方根的二階方程。接着用萊布尼茲法則（Leibniz's rule）把方程化為遞推關係（recurrence），而在 $x = 0$ 處，因子 $1 - x^2$ 變成 1，含 x 的項亦隨之消失。這樣，遞推關係便把 0 處的各階導數（derivative）全部交到我們手上。

解答

微分方程

設 $\theta = \sin^{-1} x$，則 $\sin\theta = x$，且當 $|x| < 1$ 時，$\cos\theta = \sqrt{1-x^2} > 0$，並有 $\dfrac{\mathrm{d}\theta}{\mathrm{d}x} = \dfrac{1}{\sqrt{1-x^2}}$。於是 $y = \sin^2 m\theta$，由鏈式法則（chain rule）得

$$y^{(1)} = 2m \sin m\theta \cos m\theta \cdot \frac{1}{\sqrt{1-x^2}} = \frac{m \sin 2m\theta}{\sqrt{1-x^2}}, \qquad \text{所以} \qquad \sqrt{1-x^2}\, y^{(1)} = m \sin 2m\theta.$$

把上式對 x 求導，左邊用積法則（product rule）：

$$\sqrt{1-x^2}\, y^{(2)} - \frac{x}{\sqrt{1-x^2}}\, y^{(1)} = 2m^2 \cos 2m\theta \cdot \frac{1}{\sqrt{1-x^2}}.$$

全式乘以 $\sqrt{1-x^2}$，並利用 $\cos 2m\theta = 1 - 2\sin^2 m\theta = 1 - 2y$：

$$(1-x^2)\, y^{(2)} = x\, y^{(1)} + 2m^2(1-2y),$$

這就是所求的方程。

想一想 為甚麼不先把 $\sqrt{1-x^2}\,y^{(1)} = m\sin 2m\theta$ 兩邊平方？

平方後得 $(1-x^2)\big(y^{(1)}\big)^2 = 4m^2y(1-y)$，再求導的話，每一項都會帶着因子 $2y^{(1)}$。這樣我們就要除以 $y^{(1)}$，而它在 $x=0$ 處等於零，偏偏那正是我們稍後最關心的一點。先求導而不先平方，就完全不必做這個除法。

遞推關係

利用萊布尼茲法則，把方程求導 n 次，其中 $n \geq 1$。由於 $1-x^2$ 只有兩個非零的導數（$-2x$ 和 -2），而 x 只有一個，

$$\begin{aligned}
\frac{\mathrm{d}^n}{\mathrm{d}x^n}\Big[(1-x^2)y^{(2)}\Big] &= (1-x^2)y^{(n+2)} - 2nx\,y^{(n+1)} - n(n-1)\,y^{(n)},\\
\frac{\mathrm{d}^n}{\mathrm{d}x^n}\Big[x\,y^{(1)}\Big] &= x\,y^{(n+1)} + n\,y^{(n)},\\
\frac{\mathrm{d}^n}{\mathrm{d}x^n}\Big[2m^2(1-2y)\Big] &= -4m^2\,y^{(n)}.
\end{aligned}$$

最後一行需要 $n \geq 1$，因為只有這樣，常數 $2m^2$ 求導後才會變成零。令兩邊相等，

$$(1-x^2)y^{(n+2)} - 2nx\,y^{(n+1)} - n(n-1)y^{(n)} = x\,y^{(n+1)} + n\,y^{(n)} - 4m^2y^{(n)},$$

再合併同類項（共有 $2n+1$ 個 $xy^{(n+1)}$，以及 $n(n-1)+n = n^2$），得

$$(1-x^2)\,y^{(n+2)} = (2n+1)\,x\,y^{(n+1)} + (n^2-4m^2)\,y^{(n)} \qquad (n \geq 1),$$

正如所求。

麥克勞林級數

在 $x=0$ 處有 $\theta = 0$，所以

$$y(0) = \sin^2 0 = 0, \qquad y^{(1)}(0) = m\sin 0 = 0, \qquad y^{(2)}(0) = 0 + 2m^2(1-0) = 2m^2,$$

其中最後一個由微分方程得出。在遞推關係中代入 $x=0$，得

$$y^{(n+2)}(0) = (n^2-4m^2)\,y^{(n)}(0) \qquad (n \geq 1).$$

奇數階導數。由於 $y^{(1)}(0) = 0$，遞推關係給出 $y^{(3)}(0) = 0$，繼而 $y^{(5)}(0) = 0$，如此類推：所有奇數階導數在 0 處都等於零。這是合理的，因為 y 是 x 的偶函數（even function）。

偶數階導數。取 $n = 2j$：

$$y^{(2j+2)}(0) = (4j^2-4m^2)\,y^{(2j)}(0) = -4(m^2-j^2)\,y^{(2j)}(0) \qquad (j \geq 1).$$

由 $y^{(2)}(0) = 2m^2$ 出發，依次對 $j = 1, 2, \ldots, k-1$ 應用上式，

$$y^{(2k)}(0) = 2m^2 \prod_{j=1}^{k-1} \bigl(-4(m^2 - j^2)\bigr) = (-1)^{k-1}\, 2^{2k-1} \prod_{j=0}^{k-1} (m^2 - j^2),$$

這裏我們把最前面的 m^2 併入乘積，作為 $j = 0$ 的因子。（當 $k = 1$ 時，公式給出 $2m^2$，與應有的結果相符。）

因此麥克勞林級數（Maclaurin series）為

$$y = \sin^2\bigl(m \sin^{-1} x\bigr) = \sum_{k=1}^{\infty} \frac{(-1)^{k-1}\, 2^{2k-1}}{(2k)!}\, m^2 (m^2 - 1^2)(m^2 - 2^2) \cdots \bigl(m^2 - (k-1)^2\bigr)\, x^{2k},$$

其首數項為

$$y = m^2 x^2 - \frac{m^2(m^2 - 1)}{3}\, x^4 + \frac{2m^2(m^2 - 1)(m^2 - 4)}{45}\, x^6 - \cdots.$$

若 m 是整數，則凡 $k > |m|$ 的項都含有因子 $m^2 - m^2 = 0$，所以級數到此為止，y 是一個 $2|m|$ 次多項式。

驗算

- $m = 1$：$y = \sin^2(\sin^{-1} x) = x^2$，而級數給出 x^2，其後各項全為零。$m = 2$：級數給出 $4x^2 - 4x^4$，而確實有 $\sin^2 2\theta = 4 \sin^2\theta (1 - \sin^2\theta) = 4x^2 - 4x^4$。
- $m = \frac{1}{2}$，$x = \frac{1}{2}$：$y = \sin^2(\pi/12) = \frac{1}{2}\bigl(1 - \frac{\sqrt{3}}{2}\bigr) \approx 0.066987$。級數為 $\frac{1}{4}x^2 + \frac{1}{16}x^4 + \frac{1}{32}x^6 + \cdots$，其首三項之和為 0.066895，多加幾項，差距便會逐漸縮小。
- 我們以數值方法驗證過：微分方程在隨機選取的 x 和 m 上成立；級數的係數精確地滿足該方程；以及當 $|x| \le 0.8$ 時，級數（取至 300 項）與 y 相符，誤差約為 10^{-15}。

常見錯誤

- 在 $n = 0$ 時使用遞推關係。它只對 $n \ge 1$ 成立，因為 $n = 0$ 時常數 $2m^2$ 不會消失。$n = 0$ 的情況就是原來的方程，我們正是由它得出 $y^{(2)}(0) = 2m^2$。
- 弄錯萊布尼茲法則的係數：$-n(n-1)y^{(n)}$ 這一項來自 $\binom{n}{2} \cdot (-2)$，而不是 $\binom{n}{2} \cdot (-2x)$。
- 寫通項時漏掉正負交替的符號。每一步都乘以 $4(j^2 - m^2)$，而當 $m^2 > j^2$ 時這是負數。

STEP 之外

由導數砌出來的麥克勞林級數，未必等於原來的函數，但在這裏，當 $|x| < 1$ 時兩者確實相等。這個級數的收斂半徑（radius of convergence）為 1（當 m 不是整數時，可用比值審斂法（ratio test）得出），而 $\sin^{-1} x$ 到了 $x = \pm 1$ 就不再性質良好。m 為整數時所得的多項式，是切比雪夫多項式（Chebyshev polynomials）的近親，因為 $\cos 2m\theta$ 是 $\sin\theta$ 的多項式。想知道這類級數為何真的收斂到原來的函數，可閱讀 Abbott 的 *Understanding Analysis* 中關於帶餘項的泰勒定理（Taylor's theorem with remainder）的部分。Needham 的 *Visual Complex Analysis* 則解釋了為何收斂半徑取決於 ± 1 處的毛病。

詳細解答

STEP III 1995，第 6 題

純數學 • 軌跡與反演

> **STEP III 1995，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/95-s3-q6

關鍵思路

圓 $|z-\mathrm{i}|=1$ 通過原點，所以它的極式（polar form）方程特別簡單：模（modulus）是輻角（argument）的單一個三角函數。一旦把 z 寫成模—輻角形式，$1/z$ 便立即可得（取模的倒數，並把輻角變號），再把它寫成直角坐標形式（Cartesian form），便會發現這個圓變成了一條直線。最後一部分是同一現象的逆向版本，直接設 $w=x+\mathrm{i}y$ 來做最快捷。

解答

z 的模

設 $z=r(\cos\theta+\mathrm{i}\sin\theta)$，其中 $r>0$。則 $z-\mathrm{i}=r\cos\theta+\mathrm{i}(r\sin\theta-1)$，所以

$$\begin{aligned}|z-\mathrm{i}|^2=1 &\iff r^2\cos^2\theta+r^2\sin^2\theta-2r\sin\theta+1=1\\ &\iff r^2-2r\sin\theta=0\\ &\iff r=2\sin\theta,\end{aligned}$$

其中最後一步除以 r，這是容許的，因為 $z\neq 0$。所以

$$|z|=2\sin\theta.$$

由於 $r>0$，必有 $\sin\theta>0$，即 $0<\theta<\pi$：這個圓（圓心 i，半徑 1）位於上半平面，只在原點與實軸相切，而原點已被排除。反過來，每一個 $\theta\in(0,\pi)$ 都給出圓上的一點 $z=2\sin\theta\,(\cos\theta+\mathrm{i}\sin\theta)$，所以當 θ 走遍 $(0,\pi)$ 時，z 描出除原點外的整個圓。

想一想 為甚麼從圖形來看，$|z| = 2\sin\theta$ 是「顯而易見」的？

由 0 至 2i 的線段是圓的一條直徑。根據半圓上的圓周角定理（angle in a semicircle），以 0、z、2i 為頂點的三角形在 z 處是直角，而在 0 處的角是 $\frac{\pi}{2} - \theta$。所以 $|z| = 2\cos(\frac{\pi}{2} - \theta) = 2\sin\theta$。

$1/z$ 的模和輻角

由於 $z = 2\sin\theta\,(\cos\theta + \mathrm{i}\sin\theta)$，

$$\frac{1}{z} = \frac{1}{2\sin\theta}\big(\cos(-\theta) + \mathrm{i}\sin(-\theta)\big),$$

所以

$$\left|\frac{1}{z}\right| = \frac{1}{2\sin\theta}, \qquad \arg\frac{1}{z} = -\theta.$$

因為 $0 < \theta < \pi$，角 $-\theta$ 位於 $(-\pi, 0)$ 之內，所以它就是輻角主值（principal argument）。

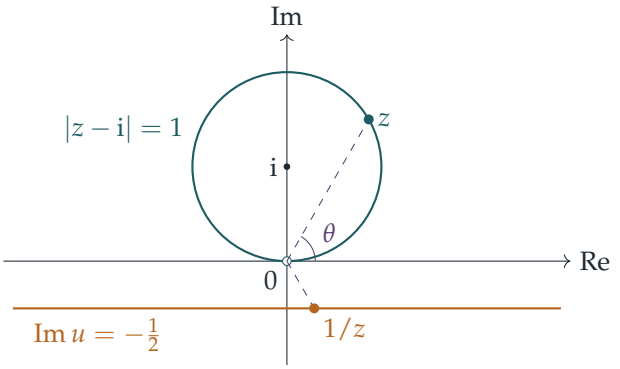
軌跡

要看出 $1/z$ 描出甚麼圖形，可把它轉為直角坐標形式：

$$\frac{1}{z} = \frac{\cos\theta - \mathrm{i}\sin\theta}{2\sin\theta} = \tfrac{1}{2}\cot\theta - \tfrac{1}{2}\mathrm{i}.$$

虛部恆為 $-\frac{1}{2}$，而當 θ 走遍 $(0, \pi)$ 時，實部 $\frac{1}{2}\cot\theta$ 取遍每一個實數值恰好一次（它由 $+\infty$ 遞減至 $-\infty$）。所以 $1/z$ 描出整條直線 $\operatorname{Im} u = -\frac{1}{2}$。

因此 z 位於圓心為 i、半徑為 1 的圓上（原點除外），而 $1/z$ 位於通過 $-\frac{1}{2}\mathrm{i}$ 的水平線上。下圖畫出兩者，並標出 $\theta = \frac{\pi}{3}$ 時的一對點：這時 $z = \frac{\sqrt{3}}{2} + \frac{3}{2}\mathrm{i}$，而 $1/z = \frac{1}{2\sqrt{3}} - \frac{1}{2}\mathrm{i}$。這兩點分別位於角度為 θ 和 $-\theta$ 的射線上，而它們與 0 的距離相乘得 1。當 z 趨近原點時，$1/z$ 沿着直線走向無窮遠；圓的最高點 $z = 2\mathrm{i}$ 則對應點 $-\frac{1}{2}\mathrm{i}$，即直線上最接近原點的一點。



軌跡 C

設 $w = x + \mathrm{i}y \neq 0$（要 $1/w$ 存在，必須有 $w \neq 0$）。則

$$\frac{1}{w} = \frac{\bar{w}}{|w|^2} = \frac{x - \mathrm{i}y}{x^2 + y^2}, \qquad \text{所以} \qquad \operatorname{Re}\frac{1}{w} = \frac{x}{x^2 + y^2}.$$

因此，對於 $w \neq 0$，

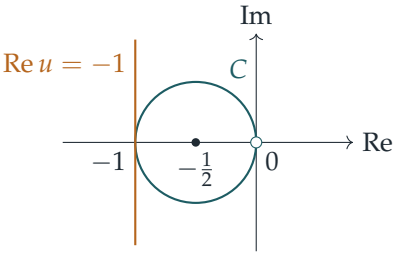
$$\begin{aligned} \operatorname{Re}\frac{1}{w} = -1 &\iff x = -(x^2 + y^2) \\ &\iff x^2 + x + y^2 = 0 \\ &\iff \left(x + \tfrac{1}{2}\right)^2 + y^2 = \tfrac{1}{4}. \end{aligned}$$

每一步都是可逆的（我們乘以了 $x^2 + y^2$，而它不等於零），所以這是名副其實的「當且僅當」。因此 C 是圓

$$\left|w + \tfrac{1}{2}\right| = \tfrac{1}{2},$$

圓心為 $-\frac{1}{2}$，半徑為 $\frac{1}{2}$，但要除去原點。原點確實在這個圓上，但 $1/w$ 在該處沒有定義，所以它不屬於 C。

這與第一部分互相呼應：在那裏，一個通過 0 的圓被 $z \mapsto 1/z$ 映射到一條不通過 0 的直線；在這裏，直線 $\operatorname{Re} u = -1$ 被同一個映射（它是自己的逆映射）映射到一個通過 0 的圓。



驗算

- 取 $\theta = \frac{\pi}{6}$：則 $|z| = 1$ 且 $z = \frac{\sqrt{3}}{2} + \frac{1}{2}\mathrm{i}$，而確實有 $|z - \mathrm{i}| = \left|\frac{\sqrt{3}}{2} - \frac{1}{2}\mathrm{i}\right| = 1$。它的倒數就是它的共軛複數（conjugate），$\frac{\sqrt{3}}{2} - \frac{1}{2}\mathrm{i} \approx 0.866 - 0.5\mathrm{i}$，其虛部為 $-\frac{1}{2}$，與預期相符。
- $\theta = \frac{\pi}{2}$ 給出 $z = 2\mathrm{i}$，即圓的最高點，而 $1/z = -\frac{1}{2}\mathrm{i}$。
- 對於 C：$w = -1$ 在圓上，而 $1/w = -1$；$w = -\frac{1}{2} + \frac{1}{2}\mathrm{i}$ 在圓上，而 $1/w = -1 - \mathrm{i}$，實部為 -1。

在每條軌跡上各取 10 000 個隨機點作數值測試，三個結果全部得到確認。

常見錯誤

- 忘記範圍 $0 < \theta < \pi$：公式 $|z| = 2\sin\theta$ 只在 $\sin\theta > 0$ 時才有意義。
- 把 $1/z$ 的軌跡說成半直線或線段。由於 $\cot\theta$ 在 $(0, \pi)$ 上取遍所有實數值，軌跡是整條直線。
- 把原點算入 C。原點滿足圓的方程，但 $1/w$ 在該處並不存在。

STEP 之外

映射 $z \mapsto 1/z$ 是莫比烏斯變換（Möbius transformation）$z \mapsto (az + b)/(cz + d)$ 的一個例子。這類映射的一個核心性質，是它們把「廣義圓」（圓和直線）映射到廣義圓，而通過原點的圓正正就是被 $1/z$ 映射到直線的那些圓。這是黎曼球面（Riemann sphere）幾何的起點；在黎曼球面上，直線只不過是通過無窮遠點的圓。Needham 的 *Visual Complex Analysis*（《可視化複分析》）配以圖像，把這一切闡述得十分精彩。

詳細解答

STEP III 1995，第 7 題

純數學 • 群與同構

STEP III 1995，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/95-s3-q7

關鍵思路

結合律（associativity）已經給定，所以對每個集合只需檢查三件事：封閉性（closure）、單位元（identity），以及每個元素都有逆元（inverse）。只要其中一項不成立，該集合就不是羣（group），而且舉出一個明確的反例便已足夠。至於同構（isomorphism），則要找出任何同構都必須保持的結構特徵：元素的個數，以及各元素的階（order）。

解答

以下所說的「乘法」，視乎情況而定，分別指複數的普通乘法、模（modulo）n 乘法，或矩陣乘法。

第 (i) 部分

設 $S = \{z \in \mathbb{C} : |z| = 1\}$。

- 封閉性：若 $|z| = |w| = 1$，則 $|zw| = |z||w| = 1$。
- 單位元：$1 \in S$，且 $1z = z1 = z$。
- 逆元：若 $|z| = 1$，則 $z \neq 0$，而 $z^{-1} = \bar{z}$（因為 $z\bar{z} = |z|^2 = 1$），其模（modulus）同樣是 1。

所以 S 是一個羣（而且是無限羣）。

第 (ii) 部分

在模 4 乘法之下，集合 $\{0,1,2,3\}$ 是封閉的，並有單位元 1，但 0 沒有逆元，因為對任何 b 都有 $0\times b\equiv 0\not\equiv 1$。所以它不是羣。（即使把 0 拿走也無補於事：2 仍然沒有逆元，因為 $2b$ 必為偶數；事實上 $2\times 2\equiv 0$，所以 $\{1,2,3\}$ 連封閉性也不具備。）

第 (iii) 部分

利用複角公式（compound-angle formulae），

$$\mathrm{M}(\alpha)\mathrm{M}(\beta)=\begin{pmatrix} c_\alpha c_\beta - s_\alpha s_\beta & -(s_\alpha c_\beta + c_\alpha s_\beta) \\ s_\alpha c_\beta + c_\alpha s_\beta & c_\alpha c_\beta - s_\alpha s_\beta \end{pmatrix} = \mathrm{M}(\alpha+\beta),$$

其中 $c_\alpha=\cos\alpha$，$s_\alpha=\sin\alpha$，如此類推。由於 cos 和 sin 的週期為 2π，故 $\mathrm{M}(\alpha+\beta)=\mathrm{M}(\gamma)$，其中 γ 是把 $\alpha+\beta$ 約化至 $[0,2\pi)$ 內所得的角。因此：

- *封閉性*：集合中兩個矩陣的積仍在集合之內。
- *單位元*：$\mathrm{M}(0)=I$。
- *逆元*：當 $0<\theta<2\pi$ 時，$\mathrm{M}(\theta)\mathrm{M}(2\pi-\theta)=\mathrm{M}(2\pi)=I$；而 $\mathrm{M}(0)$ 的逆元就是它自己。

所以這個集合是一個羣。

第 (iv) 部分

模 8 之下，

$$3^2=9\equiv 1,\qquad 5^2=25\equiv 1,\qquad 7^2=49\equiv 1,$$
$$3\times 5=15\equiv 7,\qquad 3\times 7=21\equiv 5,\qquad 5\times 7=35\equiv 3.$$

連同涉及 1 的平凡乘積，這就給出了整個乘法表（乘法滿足交換律（commutative）），而每個乘積都在 $\{1,3,5,7\}$ 之內，所以集合是封閉的。單位元是 1，而每個元素的逆元都是它自己。所以這是一個羣，階為 4，其中除 1 以外每個元素的階都是 2。

第 (v) 部分

單位矩陣 I 的元素都是整數，而且集合是封閉的（整數的積與和仍是整數），但逆元這一項不成立。零矩陣根本沒有逆矩陣。即使是可逆矩陣也可能出問題：$A=\begin{pmatrix}2&0\\0&1\end{pmatrix}$ 在實矩陣中的逆矩陣是 $\begin{pmatrix}\frac{1}{2}&0\\0&1\end{pmatrix}$，而逆矩陣是唯一的，所以 A 沒有元素全為整數的逆矩陣。所以這個集合不是羣。

想一想 單單指出零矩陣真的足夠嗎？還是需要就所有矩陣說明一下？

只要有一個元素沒有逆元，就足以證明該集合不是羣。第二個例子則說明，問題並不止於奇異矩陣（singular matrix）。一個整數矩陣有整數逆矩陣，當且僅當它的行列式（determinant）是 ± 1，因為 $\det A \det A^{-1} = 1$，而兩個行列式都是整數。若只考慮這些矩陣，便確實得到一個羣。

第 (vi) 部分

模 5 之下，2 的各次冪是

$$2^1 = 2, \quad 2^2 = 4, \quad 2^3 = 8 \equiv 3, \quad 2^4 = 16 \equiv 1.$$

所以 $\{1,2,3,4\} = \{2^0, 2^1, 2^2, 2^3\}$，而 $2^a 2^b = 2^{a+b}$，其中指數按模 4 計算。因此集合是封閉的，1 是單位元，而 2^a 的逆元是 2^{4-a}。它是一個羣：階為 4 的循環羣（cyclic group），由 2（或由 3）生成。

同構

是羣的有 (i)、(iii)、(iv) 和 (vi)，共可組成六對。

(i) 與 (iii) 同構。每個單位複數都可以唯一地寫成 $\mathrm{e}^{\mathrm{i}\theta}$，其中 $0 \le \theta < 2\pi$。定義 $\phi(\mathrm{e}^{\mathrm{i}\theta}) = \mathrm{M}(\theta)$。這是一個雙射（bijection），因為 $[0, 2\pi)$ 內不同的 θ 給出不同的矩陣（第一列 $(\cos\theta, \sin\theta)$ 已決定了 θ）。它也保持乘法：$\mathrm{e}^{\mathrm{i}\alpha}\mathrm{e}^{\mathrm{i}\beta} = \mathrm{e}^{\mathrm{i}(\alpha+\beta)}$ 及 $\mathrm{M}(\alpha)\mathrm{M}(\beta) = \mathrm{M}(\alpha+\beta)$，而在兩個羣中，角都是按模 2π 約化的。從幾何角度看，兩個羣都是平面繞原點的旋轉。

(i) 或 (iii) 與 (iv) 或 (vi)：不同構。同構是雙射，所以兩個羣的元素個數必須相同。羣 (i) 和 (iii) 是無限的，而 (iv) 和 (vi) 各有四個元素。這樣就排除了四對。

想一想 (iv) 和 (vi) 都有四個元素，而且都滿足交換律。有甚麼可以把它們區分開來？

(iv) 與 (vi)：不同構。同構 ϕ 保持每個元素的階：$\phi(g)^n = \phi(g^n)$，而它等於單位元當且僅當 g^n 是單位元，因為 ϕ 是單射（injective），並把單位元映到單位元。在 (vi) 中，元素 2 的階是 4。在 (iv) 中，每個元素 g 都滿足 $g^2 = 1$，所以沒有元素的階是 4。因此不存在同構。

總括而言，唯一同構的一對是 (i) 和 (iii)。

驗算

我們用電腦對 (iv) 和 (vi) 的乘法表逐一窮舉，確認了封閉性、單位元和逆元三項性質。電腦亦發現，模 4 之下 0 和 2 沒有逆元，而元素的階在 (iv) 中是 $1,2,2,2$，在 (vi) 中是 $1,4,4,2$（依次對應 $1,2,3,4$）。它還在 10^4 個隨機角上驗證了 $\mathrm{M}(\alpha)\mathrm{M}(\beta)=\mathrm{M}(\alpha+\beta)$ 以及與 $\mathrm{e}^{i\theta}$ 的對應關係。(iv) 可以很快用手算驗證：任何奇數的平方在模 8 之下都是 1，因為 $(2k+1)^2=4k(k+1)+1$，而 $k(k+1)$ 是偶數。

常見錯誤

- 說 (iv) 和 (vi) 同構，「因為兩者都有四個元素」。元素個數相等是必要條件，但並非充分條件。
- 在 (v) 中只檢查 $\det A\neq 0$。逆矩陣的元素也必須全為整數。
- 把 (ii) 誤讀成模 4 加法。在加法之下，它會是一個階為 4 的循環羣，與 (vi) 同構，但題目指明的是乘法。

STEP 之外

在同構意義下，階為 4 的羣恰好只有兩個：循環羣 C_4，即 (vi)；以及克萊因四元羣（Klein four-group）$C_2\times C_2$，即 (iv)。按元素的階把小階數的羣分類，是大學羣論課程最先處理的課題之一。(i) 和 (iii) 的圓羣（circle group）是李羣（Lie group）最簡單的例子；李羣是指同時也是光滑曲線或曲面的羣。(v) 中行列式為 ±1 的整數矩陣構成羣 $\mathrm{GL}_2(\mathbb{Z})$，它在數論中十分重要。想進一步了解，Allenby 的 *Rings, Fields and Groups* 是一本很好的讀物。

詳細解答

STEP III 1995，第 8 題

純數學 • 平面與球面

> **STEP III 1995，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/95-s3-q8

關鍵思路

平面截球面時，所得的圓的圓心就是由球心向該平面所作垂線的垂足（foot of the perpendicular）。因此，球心必定位於通過每一個圓的圓心的法線（normal）上，而畢氏定理（Pythagoras）則把球心沿每條法線的距離與球的半徑聯繫起來。題目餘下的部分，不過是把這一條向量方程與精心挑選的向量作點積（dot product）而已。

解答

法線與距離

通過 X 而垂直於 π 的法線，就是通過 $\mathbf{x}$ 而方向為 $\mathbf{n}$ 的直線：

$$\mathbf{r} = \mathbf{x} + t\,\mathbf{n}, \qquad t \in \mathbb{R}.$$

它與 π 相交於 $(\mathbf{x}+t\mathbf{n})\cdot\mathbf{n} = p$ 之處，即（因為 $\mathbf{n}\cdot\mathbf{n} = 1$）$t = p - \mathbf{x}\cdot\mathbf{n}$ 之處。所以垂足是 $\mathbf{x} + (p - \mathbf{x}\cdot\mathbf{n})\mathbf{n}$，而由於 $\mathbf{n}$ 是單位向量（unit vector），X 到 π 的垂直距離為

$$|\mathbf{x}\cdot\mathbf{n} - p|.$$

同時通過兩個圓的球面

假設有一個球心為 $\mathbf{c}$、半徑為 R 的球面包含這兩個圓。固定 i。C_i 上的每一點 $\mathbf{y}$ 都滿足 $|\mathbf{y} - \mathbf{r}_i| = 1$ 及 $(\mathbf{y} - \mathbf{r}_i)\cdot\mathbf{n}_i = 0$；當 $\mathbf{y}$ 繞 C_i 走一圈時，向量 $\mathbf{u} = \mathbf{y} - \mathbf{r}_i$ 會取遍與 $\mathbf{n}_i$ 垂直的每一個單位向量。

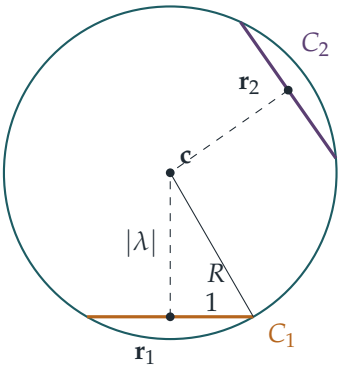
作分解 $\mathbf{c}-\mathbf{r}_i=\mathbf{a}+\mu_i\mathbf{n}_i$，其中 $\mathbf{a}\perp\mathbf{n}_i$。於是

$$\begin{aligned}R^2=|\mathbf{y}-\mathbf{c}|^2&=|\mathbf{u}-\mathbf{a}-\mu_i\mathbf{n}_i|^2\\&=1-2\,\mathbf{u}\cdot\mathbf{a}+|\mathbf{a}|^2+\mu_i^2,\end{aligned}$$

這裏用到 $\mathbf{u}\cdot\mathbf{n}_i=\mathbf{a}\cdot\mathbf{n}_i=0$。左邊對 C_i 上的每一點都相同，所以 $\mathbf{u}\cdot\mathbf{a}$ 對每一個容許的 $\mathbf{u}$ 都取相同的值。但只要 $\mathbf{u}$ 是容許的，$-\mathbf{u}$ 也是容許的，所以對平面上每一個單位向量 $\mathbf{u}$ 都有 $\mathbf{u}\cdot\mathbf{a}=-\mathbf{u}\cdot\mathbf{a}=0$，而取 $\mathbf{u}$ 與 $\mathbf{a}$ 平行（假如 $\mathbf{a}\neq\mathbf{0}$）便會得出矛盾。因此 $\mathbf{a}=\mathbf{0}$，並且

$$\mathbf{c}=\mathbf{r}_i+\mu_i\mathbf{n}_i,\qquad R^2=1+\mu_i^2.$$

用文字來說：球心位於通過 C_i 的圓心而垂直於 π_i 的法線上，這正是第一部分所求得的直線。



把以上結果應用於 $i=1$ 和 $i=2$，得 $1+\mu_1^2=R^2=1+\mu_2^2$，所以 $\mu_2=\pm\mu_1$。記 $\lambda=\mu_1$，則

$$\mathbf{r}_1+\lambda\mathbf{n}_1=\mathbf{c}=\mathbf{r}_2\pm\lambda\mathbf{n}_2,$$

所以只要存在合適的球面，這樣的 λ 就必定存在，這正是所要證明的。

兩個必要條件

移項後得

$$\mathbf{r}_1-\mathbf{r}_2=-\lambda\mathbf{n}_1\pm\lambda\mathbf{n}_2.$$

兩邊與 $\mathbf{n}_1\times\mathbf{n}_2$ 作點積，而這個向量同時垂直於 $\mathbf{n}_1$ 和 $\mathbf{n}_2$：

$$(\mathbf{r}_1-\mathbf{r}_2)\cdot(\mathbf{n}_1\times\mathbf{n}_2)=0.$$

想一想 p_1 從何而來？它並沒有在向量方程中出現。

圓心 $\mathbf{r}_1$ 位於 π_1 上，所以 $\mathbf{n}_1 \cdot \mathbf{r}_1 = p_1$；同理 $\mathbf{n}_2 \cdot \mathbf{r}_2 = p_2$。記 $k = \mathbf{n}_1 \cdot \mathbf{n}_2$。把 $\mathbf{r}_1 + \lambda \mathbf{n}_1 = \mathbf{r}_2 \pm \lambda \mathbf{n}_2$ 先與 $\mathbf{n}_1$ 作點積，再與 $\mathbf{n}_2$ 作點積，得

$$p_1 + \lambda = \mathbf{n}_1 \cdot \mathbf{r}_2 \pm \lambda k,$$
$$\mathbf{n}_2 \cdot \mathbf{r}_1 + \lambda k = p_2 \pm \lambda.$$

所以

$$p_1 - \mathbf{n}_1 \cdot \mathbf{r}_2 = \lambda(\pm k - 1), \qquad p_2 - \mathbf{n}_2 \cdot \mathbf{r}_1 = \lambda(k \mp 1).$$

取上方的符號時，兩式右邊都等於 $\lambda(k-1)$；取下方的符號時，它們分別是 $-\lambda(k+1)$ 和 $\lambda(k+1)$。無論哪一種情況，兩式左邊不是相等就是互為相反數，所以

$$(p_1 - \mathbf{n}_1 \cdot \mathbf{r}_2)^2 = (p_2 - \mathbf{n}_2 \cdot \mathbf{r}_1)^2.$$

幾何意義

第一個條件。它說明 $\mathbf{r}_1 - \mathbf{r}_2$、$\mathbf{n}_1$ 和 $\mathbf{n}_2$ 共面（coplanar），即純量三重積（scalar triple product）等於零。等價地說，通過 C_1 的圓心而垂直於 π_1 的法線，與通過 C_2 的圓心而垂直於 π_2 的法線，位於同一個平面上：它們不是異面直線（skew lines），所以它們或者相交（交於球心），或者平行。

第二個條件。根據第一部分，$|p_1 - \mathbf{n}_1 \cdot \mathbf{r}_2|$ 是 C_2 的圓心到 π_1 的距離，而 $|p_2 - \mathbf{n}_2 \cdot \mathbf{r}_1|$ 是 C_1 的圓心到 π_2 的距離。所以這個條件說明這兩個距離相等：每個圓的圓心到另一個圓所在平面的距離，等於另一個圓心到這個圓所在平面的距離。這反映了圖形的對稱性：兩個圓與 $\mathbf{c}$ 的距離同為 $|\lambda|$，所以對某個通過 $\mathbf{c}$ 的合適平面作反射（reflection），便會把兩個圓互換，同時把這兩個距離互換。

驗算

- 取 $\pi_1 : z = 0$ 和 $\pi_2 : z = 2$，兩個單位圓的圓心都在 z 軸上。它們位於球心為 $(0, 0, 1)$、半徑為 $\sqrt{2}$ 的球面上；這時 $\lambda = 1$，取下方的符號，$\mathbf{n}_1 \times \mathbf{n}_2 = \mathbf{0}$，而第二個條件中的兩個距離都等於 2。
- 在數值方面，我們隨機構造了 5000 個半徑 $R > 1$ 的球面，每個球面配上兩個隨機平面，各截出一個單位圓，並確認了圓上每一點都在球面上，結果發現兩個條件在每一個情況下都成立，誤差不超過 10^{-9}，而且兩種符號都有出現。至於隨機選取、互不相關的兩個圓，則每一次都不滿足第一個條件。

常見錯誤

- 未加論證便假設球心位於通過 $\mathbf{r}_i$ 的法線上；這正是整個證明的核心。

- 遺漏了 $\pm$：球心可以位於兩個平面各自的法向量所指的同一側，也可以位於相反的兩側。
- 忘記了由於 C_1 的圓心位於 π_1 上，所以 $\mathbf{n}_1 \cdot \mathbf{r}_1 = p_1$。

STEP 之外

這裏的條件只是必要條件（necessary conditions）：要得到完整的「當且僅當」（if and only if），還需要兩條法線確實相交於一個與兩個圓等距的點。這類題目是踏入二次曲面（quadrics）解析幾何（analytic geometry）的第一步，在那裏，平面、球面以及它們的交集都用線性代數（linear algebra）來處理；純量三重積會以量度體積的 3×3 行列式（determinant）的形式再次出現。Strang 的 *Introduction to Linear Algebra*（線性代數入門）對這些概念有細緻的闡述。

詳細解答

STEP III 1995，第 9 題

力學 • 複擺

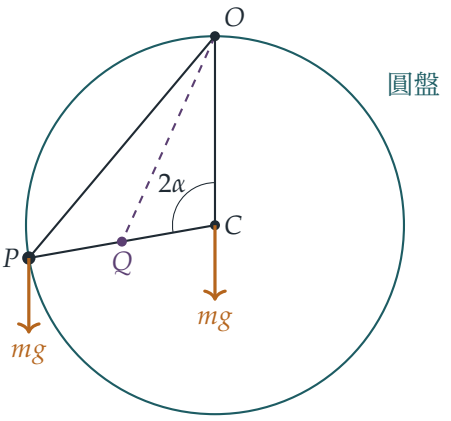
> **STEP III 1995，第 9 題**
>
> 繼續閱讀之前，請先看題目（原題為英文）。題目載於 STEP Support Programme 的試題資料庫：
>
> step.maths.org/questions/95-s3-q9

關鍵思路

圓盤與質點的質量相同，所以整個系統的質心（centre of mass）就是 CP 的中點 Q。差不多所有結論都由此而來：在穩定平衡（stable equilibrium）時，Q 位於 O 的正下方；而第 (iii) 部分的能量方程只需要 Q 的高度。餘下的工作，就是等腰三角形 OCP 的幾何，以及一個標準的轉動慣量（moment of inertia）積分。

解答

我們把圓盤視為剛性薄片（rigid lamina），把 P 視為固定在其邊緣上的質點（particle）；轉軸是光滑的，所以機械能守恆。全題中 $OC = CP = r$，且 $\angle OCP = 2\alpha$。



第 (i) 部分

由於圓盤（質心為 C）與質點（位於 P）的質量同為 m，系統的質心就是 CP 的中點 Q。作用於系統的力只有重力和轉軸上的反作用力，所以系統處於平衡，當且僅當 Q 與 O 在同一鉛垂線上；而當 Q 在 O 的下方時，平衡是穩定的（此時任何位移都會令 Q 升高，使勢能增加）。因此在穩定平衡時 OQ 是鉛垂的，而 β 就是角 $\angle COQ$。

在圓盤所在的平面上取坐標軸，以 C 為原點，並令 O 位於 $(r,0)$；於是 $P = r(\cos 2\alpha, \sin 2\alpha)$，$Q = \frac{r}{2}(\cos 2\alpha, \sin 2\alpha)$。因此

$$\overrightarrow{OC} = (-r,\, 0), \qquad \overrightarrow{OQ} = \tfrac{r}{2}\big(-(2-\cos 2\alpha),\, \sin 2\alpha\big).$$

由於 $2 - \cos 2\alpha > 0$，向量 $\overrightarrow{OQ}$ 沿 $\overrightarrow{OC}$ 方向的分量為正數 $\frac{r}{2}(2-\cos 2\alpha)$，而與之垂直的分量為 $\frac{r}{2}\sin 2\alpha$，所以 β 是銳角，且

$$\tan\beta = \frac{\frac{r}{2}\sin 2\alpha}{\frac{r}{2}(2-\cos 2\alpha)} = \frac{\sin 2\alpha}{2-\cos 2\alpha},$$

這正是所要證明的結果。

第 (ii) 部分

我們把 $\rho x/r$ 理解為每單位面積的質量。半徑介乎 x 與 $x + \mathrm{d}x$ 之間的圓環，面積為 $2\pi x\,\mathrm{d}x$，所以

$$m = \int_0^r \frac{\rho x}{r}\, 2\pi x\,\mathrm{d}x = \frac{2\pi\rho r^2}{3}.$$

該圓環上每一點與通過 C 且垂直於圓盤的軸的距離都是 x，所以對該軸的轉動慣量為

$$I_C = \int_0^r \frac{\rho x}{r}\, 2\pi x \cdot x^2\,\mathrm{d}x = \frac{2\pi\rho r^4}{5} = \frac{3}{5}mr^2,$$

其中用了 $2\pi\rho r^2 = 3m$。通過 O 的軸與這條軸平行（兩者都垂直於圓盤所在的鉛垂平面），相距 $OC = r$，所以由平行軸定理（parallel axis theorem）得

$$I_O = \frac{3}{5}mr^2 + mr^2 = \frac{8}{5}mr^2.$$

第 (iii) 部分

系統的轉動慣量。三角形 OCP 是等腰三角形，頂角為 2α，所以弦長 $OP = 2r\sin\alpha$。質點的貢獻是 $m(2r\sin\alpha)^2$，於是

$$I = \frac{8}{5}mr^2 + 4mr^2\sin^2\alpha = \frac{4}{5}mr^2\,(2 + 5\sin^2\alpha).$$

質心下降的高度。釋放時 OQ 是水平的，所以 Q 與 O 同高。在最終位置，P 在 O 的正下方，所以 OP 是鉛垂的，而任何一點在 O 以下的深度，就是它沿 OP 方向與 O 的距離。

- P 的深度為 $OP = 2r\sin\alpha$。
- 三角形 OCP 的底角為 $\frac{1}{2}(\pi - 2\alpha) = \frac{\pi}{2} - \alpha$，所以 $\angle COP = \frac{\pi}{2} - \alpha$，而 C 的深度為 $r\cos(\frac{\pi}{2} - \alpha) = r\sin\alpha$。

Q 是 CP 的中點，所以它的深度是兩者的平均值，即 $\frac{3}{2}r\sin\alpha$。因此系統（總質量為 $2m$）損失的勢能為 $2mg \cdot \frac{3}{2}r\sin\alpha = 3mgr\sin\alpha$。

> ? **想一想** 系統真的會到達 P 在 O 下方的那個位置嗎？

系統在 Q 與 O 同高時由靜止釋放，其後便像複擺（compound pendulum）一樣，在 OQ 水平的兩個位置之間擺動，並經過 Q 低於 O 所在水平的每一個位置。在最終位置，Q 的深度為 $\frac{3}{2}r\sin\alpha > 0$，所以系統確實會到達該位置（而「$C$ 起初比 P 低」這個條件，只是決定了擺動的方向）。

能量。由靜止開始，設角速率（angular speed）為 ω，則

$$\tfrac{1}{2}I\omega^2 = 3mgr\sin\alpha \quad\Longrightarrow\quad \omega^2 = \frac{6gr\sin\alpha}{\frac{4}{5}r^2(2+5\sin^2\alpha)} = \frac{15g\sin\alpha}{2r(2+5\sin^2\alpha)}.$$

C 沿着以 O 為圓心、半徑為 $OC = r$ 的圓運動，所以 $v = r\omega$，且

$$v^2 = \frac{15gr\sin\alpha}{2(2+5\sin^2\alpha)},$$

這正是所要證明的結果。

極大值。設 $s = \sin\alpha$；由於 $0 < \alpha \le \pi/2$，故 $s \in (0,1]$。令 $f(s) = \dfrac{s}{2+5s^2}$，則

$$f'(s) = \frac{(2+5s^2) - 10s^2}{(2+5s^2)^2} = \frac{2-5s^2}{(2+5s^2)^2},$$

當 $s < \sqrt{2/5}$ 時為正，當 $s > \sqrt{2/5}$ 時為負。由於 $\sqrt{2/5} < 1$ 在容許的範圍之內，f 在該處取得極大值，$f = \dfrac{\sqrt{2/5}}{4} = \dfrac{\sqrt{10}}{20}$。因此

$$v^2_{\max} = \frac{15gr}{2}\cdot\frac{\sqrt{10}}{20} = \frac{3\sqrt{10}}{8}gr \approx 1.19\,gr, \qquad \text{當 } \sin\alpha = \sqrt{2/5}.$$

驗算

- 若 $\alpha \to 0$，P 就落在 O 上，只有圓盤懸掛着，OC 是鉛垂的：果然 $\tan\beta \to 0$。若 $\alpha = \pi/2$，P 與 O 位於直徑的兩端，Q 在 OC 上，同樣有 $\beta = 0$。
- 取 $\alpha = \pi/2$：$I = \frac{28}{5}mr^2$，質量 $2m$ 下降了 $\frac{3}{2}r$，所以 $v^2 = r^2 \cdot 3mgr \cdot \frac{10}{28mr^2} = \frac{15}{14}gr$，與公式吻合。

- 我們就幾個不同的 α 值，由釋放位置開始對運動方程作數值積分（RK4）；例如當 $\alpha = 0.7$、$r = 1.3$、$g = 9.81$ 時，模擬結果為 $v^2 = 15.1206$，與公式吻合至六位有效數字。直接搜尋亦確認極大值為 $3\sqrt{10}/8 \approx 1.1859$（以 gr 為單位）。

常見錯誤

- I_C 用了均勻圓盤的數值 $\frac{1}{2}mr^2$；這裏的密度向外遞增，令 I_C 增大至 $\frac{3}{5}mr^2$。
- 在第 (iii) 部分忘記了質點對轉動慣量的貢獻 $m \cdot OP^2$。
- 答了 ω^2 而不是 v^2：題目問的是 C 的速率，即 $r\omega$。

STEP 之外

繞水平軸轉動的剛體（rigid body）就是複擺：它的表現就像一個擺錘位於質心的單擺（simple pendulum），只是慣量由 Md^2 換成 I。在第 (i) 部分的平衡位置附近作小幅振動時，週期為 $2\pi\sqrt{I/(Mgd)}$，其中 $d = OQ$；而「等效單擺」的長度 $I/(Md)$ 則引出了 Kater 的可倒擺（reversible pendulum），這種擺曾用來精確量度 g。Kleppner 與 Kolenkow 合著的 *An Introduction to Mechanics*（力學入門）正是以這種思路處理剛體轉動和平行軸定理。

詳細解答

STEP III 1995，第 10 題

力學 • 斜面上的拋體

STEP III 1995，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/95-s3-q10

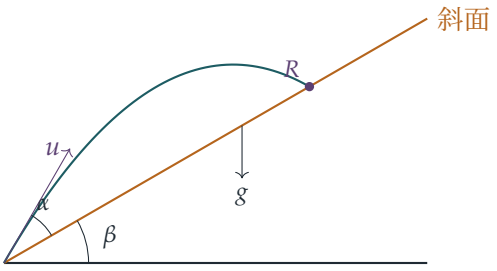
關鍵思路

在斜面（inclined plane）上，最自然的坐標軸不是水平和垂直方向，而是沿着斜面和垂直於斜面的方向。在這組坐標軸下，重力有兩個恆定的分量，飛行時間（time of flight）單憑垂直於斜面的運動便可求得，而射程（range）則是三角函數的乘積。利用積化和差公式（factor formula）把這個乘積化為兩個正弦之差，最大值便一目了然，完全不需要微積分。

解答

建立模型

把炮彈視為只受重力作用而自由運動的質點（particle），不計空氣阻力，從斜面底部沿最大斜率線（line of greatest slope）向上發射。設發射方向與斜面的夾角為 α，則它相對於水平面的仰角（angle of elevation）為 $\theta = \alpha + \beta$；我們需要 $0 < \alpha < \frac{\pi}{2} - \beta$，炮彈才會真正沿斜面向上飛行。



取 x-軸沿最大斜率線向上，y-軸垂直於斜面並指向離開斜面的一方。把加速度 g（鉛垂

向下）沿這兩個方向分解，得

$$\ddot{x} = -g\sin\beta, \qquad \ddot{y} = -g\cos\beta,$$

而初速度沿斜面的分量為 $u\cos\alpha$，垂直於斜面的分量為 $u\sin\alpha$。因此在時間 t 時

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

飛行時間與射程

當 $t > 0$ 而再次有 $y = 0$ 時，炮彈着地：

$$T = \frac{2u\sin\alpha}{g\cos\beta}.$$

於是炮彈沿斜面向上所走的距離為

$$\begin{aligned} R = x(T) &= \frac{2u^2\sin\alpha\cos\alpha}{g\cos\beta} - \tfrac{1}{2}g\sin\beta\cdot\frac{4u^2\sin^2\alpha}{g^2\cos^2\beta} \\ &= \frac{2u^2\sin\alpha}{g\cos^2\beta}\big(\cos\alpha\cos\beta - \sin\alpha\sin\beta\big) \\ &= \frac{2u^2\sin\alpha\cos(\alpha+\beta)}{g\cos^2\beta}. \end{aligned}$$

想一想 怎樣才能不用求導而求出 $\sin\alpha\cos(\alpha+\beta)$ 的最大值？

在積化和差公式 $2\sin A\cos B = \sin(A+B) + \sin(A-B)$ 中取 $A = \alpha$ 和 $B = \alpha+\beta$，得

$$2\sin\alpha\cos(\alpha+\beta) = \sin(2\alpha+\beta) - \sin\beta,$$

所以

$$R = \frac{u^2\big(\sin(2\alpha+\beta) - \sin\beta\big)}{g\cos^2\beta}.$$

只有 $\sin(2\alpha+\beta)$ 這一項與 α 有關，而它最大只能是 1。當 α 取遍 $(0, \frac{\pi}{2}-\beta)$ 時，角 $2\alpha+\beta$ 取遍 $(\beta, \pi-\beta)$，而這個區間包含 $\frac{\pi}{2}$；所以最大值 1 確實可以取到，而且恰好在 $2\alpha+\beta = \frac{\pi}{2}$ 時取到，即

$$\alpha = \frac{\pi}{4} - \frac{\beta}{2}, \qquad \text{等價地} \qquad \theta = \alpha+\beta = \frac{\pi}{4} + \frac{\beta}{2}.$$

因此大炮應以高於水平面 $\frac{\pi}{4} + \frac{\beta}{2}$ 的角度發射：發射方向平分斜面與鉛垂線之間的夾角。

最大射程

取 $\sin(2\alpha+\beta) = 1$，並利用 $\cos^2\beta = 1-\sin^2\beta = (1-\sin\beta)(1+\sin\beta)$，

$$R_{\max} = \frac{u^2(1-\sin\beta)}{g(1-\sin\beta)(1+\sin\beta)} = \frac{u^2}{g(1+\sin\beta)},$$

其中由於 $\beta < \frac{\pi}{2}$，約去 $1-\sin\beta$ 是合理的。這就是所求的距離。

驗算

- 平地，即 $\beta = 0$：最佳角度是 $45°$，射程為 u^2/g，正是大家熟悉的結果。
- 當 $\beta \to \frac{\pi}{2}$（一個鉛垂的「斜面」）時，最佳方向趨向垂直向上，而 $R_{\max} \to u^2/(2g)$，即垂直向上發射的炮彈所能達到的最大高度。
- 數值上，取 $u = 20\,\mathrm{m\,s^{-1}}$、$g = 9.81\,\mathrm{m\,s^{-2}}$ 和 $\beta = 30°$，公式給出在仰角 $60°$ 時 $R_{\max} \approx 27.18\,\mathrm{m}$。在水平和垂直坐標下逐步計算炮彈的運動，直至它穿過斜面，對隨機選取的 u, β, α 都能重現 $R(\alpha)$；再對 α 作精細搜索，最大值確實出現在 $\frac{\pi}{4} - \frac{\beta}{2}$。

常見錯誤

- 混淆了與水平面的夾角和與斜面的夾角；答案 $\frac{\pi}{4} + \frac{\beta}{2}$ 是從水平面量起的。
- 遺漏了沿斜面方向的分量 $-g\sin\beta$，結果得到的只是比例出錯的平地射程公式。

STEP 之外

這個問題有一個優雅的幾何看法：從同一點以速率 u 發射的所有軌跡，都位於一條「安全拋物線」（parabola of safety，即這些軌跡的包絡線（envelope））之內，其焦點（focus）就在發射點，而任何斜面上的最大射程，就在斜面與這條拋物線相交之處。利用焦點–準線（directrix）性質，一行便可得出 $R_{\max} = u^2/(g(1+\sin\beta))$，同時解釋了為何最佳方向平分斜面與鉛垂線之間的夾角。Kleppner 與 Kolenkow 所著的 *An Introduction to Mechanics*（力學入門）以向量（vector）方法處理拋體（projectile）問題，是進一步學習的好選擇。

詳細解答

STEP III 1995，第 11 題

力學 • 相對速度

> **STEP III 1995，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/95-s3-q11

關鍵思路

改用船的參考系（frame of reference）來看。相對於船，飛機必須沿西北方向的直線筆直飛出，再沿同一直線筆直飛回。空氣帶着飛機一起移動，所以在這個參考系中，飛機的速度等於它的空速（airspeed）向量再加上一股「水流」：風的速度減去船的速度。一旦知道這股水流，每一程都是一道渡河問題。我們選取航向（heading），使合速度（resultant velocity）指向這條直線的方向，然後由來回所需的時間定出距離。

解答

全題取 x 軸向東、y 軸向北，速率以節（knot，即每小時海里數）為單位，於是距離便以海里（nautical mile）為單位。船的速度是 $(-V, 0)$。若空氣以速度 $\mathbf{w}$ 移動，而飛機相對於空氣的速度是 $\mathbf{a}$，其中 $|\mathbf{a}| = kV$，則飛機相對於船的速度（相對速度，relative velocity）是

$$\mathbf{a} + \mathbf{c}, \qquad \mathbf{c} = \mathbf{w} - (-V, 0) = \mathbf{w} + (V, 0).$$

飛機一直保持在船的西北方，所以 $\mathbf{a} + \mathbf{c}$ 在去程必須指向 $\mathbf{u} = \frac{1}{\sqrt{2}}(-1, 1)$ 的方向，在回程則必須指向 $-\mathbf{u}$ 的方向。

有風時：去程

風是由北向南吹的，所以 $\mathbf{w} = (0, -V)$，而

$$\mathbf{c} = (V, -V) = -\sqrt{2}\, V\, \mathbf{u}.$$

這股水流的速率是 $\sqrt{2}\,V$，方向為正東南，恰好沿着巡邏線。因此機師只需把機頭對準這條直線。去程（向西北）時她正面逆流而飛；回程時則順流而飛：

$$\text{去程速率} = (k-\sqrt{2})V, \qquad \text{回程速率} = (k+\sqrt{2})V.$$

這裏需要 $k > \sqrt{2}$：否則飛機永遠無法沿西北方向的直線離開船。

若飛機飛到距離船 d_w 之處，而每次來回需時 $\frac{1}{2}$ 小時，則

$$\frac{d_w}{(k-\sqrt{2})V} + \frac{d_w}{(k+\sqrt{2})V} = \frac{1}{2}$$
$$\frac{d_w}{V}\cdot\frac{2k}{k^2-2} = \frac{1}{2},$$

所以

$$d_w = \frac{(k^2-2)V}{4k}.$$

因此去程需時

$$\frac{d_w}{(k-\sqrt{2})V} = \frac{k^2-2}{4k(k-\sqrt{2})} = \frac{k+\sqrt{2}}{4k}\ \text{小時} = 15\left(1+\frac{\sqrt{2}}{k}\right)\ \text{分鐘}.$$

這個時間必定多於 15 分鐘，正合預期：去程是逆流而行的。

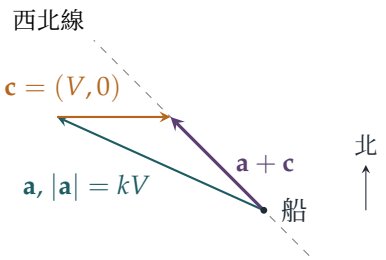
想一想 在無風的情況下，飛機為甚麼不能同樣只把機頭對準西北？

無風時，水流不再沿着巡邏線。假如機師把機頭對準西北，水流中向東的部分便會把飛機帶離這條直線。她必須把機頭稍為偏向迎着水流的一方，而這會耗去她一部分空速。

無風時

現在 $\mathbf{w} = \mathbf{0}$，所以 $\mathbf{c} = (V, 0)$：相對於船，空氣以速率 V 向東流動。把 $\mathbf{c}$ 分解為沿直線的分量（component）和垂直於直線的分量，前者用 $\mathbf{u}$，後者用 $\mathbf{v} = \frac{1}{\sqrt{2}}(1,1)$：

$$\mathbf{c} = -\frac{V}{\sqrt{2}}\,\mathbf{u} + \frac{V}{\sqrt{2}}\,\mathbf{v}.$$



合速度不能有 $\mathbf{v}$ 分量，所以 $\mathbf{a}$ 的 $\mathbf{v}$ 分量必須是 $-V/\sqrt{2}$。由於 $|\mathbf{a}| = kV$，它沿行進方向的分量便是

$$\sqrt{k^2V^2 - \tfrac{1}{2}V^2} = V\sqrt{k^2 - \tfrac{1}{2}}.$$

沿着這條直線，水流在去程的貢獻是 $-V/\sqrt{2}$，在回程則是 $+V/\sqrt{2}$（因為 $\mathbf{c}\cdot(-\mathbf{u}) = V/\sqrt{2}$）。設 $A = V\sqrt{k^2 - \frac{1}{2}}$ 及 $B = V/\sqrt{2}$，則

$$\text{去程速率} = A - B, \qquad \text{回程速率} = A + B.$$

去程速率是正數，因為 $A^2 - B^2 = V^2(k^2 - 1) > 0$。

來回同樣需時半小時，於是

$$\frac{d_c}{A-B} + \frac{d_c}{A+B} = \frac{2A\,d_c}{A^2-B^2} = \frac{1}{2} \quad\Longrightarrow\quad d_c = \frac{A^2-B^2}{4A} = \frac{V(k^2-1)}{4\sqrt{k^2-\frac{1}{2}}}.$$

比值

兩式相除，

$$\frac{d_w}{d_c} = \frac{(k^2-2)V}{4k}\cdot\frac{4\sqrt{k^2-\frac{1}{2}}}{V(k^2-1)} = \frac{(k^2-2)\sqrt{k^2-\frac{1}{2}}}{k(k^2-1)}.$$

由於 $\sqrt{k^2-\frac{1}{2}} = \frac{1}{2}\sqrt{4k^2-2}$，

$$\frac{d_w}{d_c} = \frac{k^2-2}{2k(k^2-1)}\sqrt{4k^2-2},$$

這正是所求的結果。

驗算

- 當 $k \to \infty$ 時，風和船幾乎無關重要。去程時間趨向 15 分鐘，而 $d_w/d_c \to 1$（比值的分子和分母都以 $2k^3$ 的速度增長）。
- 當 $k \to \sqrt{2}^{+}$ 時，有風時的去程速率降至零，所以 $d_w \to 0$，比值 $\to 0$，與公式相符。
- 當 $k = 2$ 時，去程需時 $15(1 + 1/\sqrt{2}) \approx 25.6$ 分鐘，而 $d_w/d_c = \frac{2}{12}\sqrt{14} \approx 0.624$。我們直接解出航向，並逐步模擬整個運動，以數值方法確認了這個結果。

常見錯誤

- 把風向弄反：「由北向南」的風是朝着南方吹的。

- 把船的速度加上去，而不是減去。相對於一艘向西航行的船，靜止的空氣是向東漂移的。
- 在無風時把機頭對準西北，並以 $kV \pm V/\sqrt{2}$ 作為速率。必須先抵銷橫向的水流。

STEP 之外

這是一次伽利略參考系變換（Galilean change of reference frame）：在彼此作勻速相對運動的參考系中量度所得的速度，只相差一個常向量。邁克生—莫雷（Michelson–Morley）的「河中小船」論證，背後也是同一套計算。在那個論證中，橫渡來回與順逆流來回所需的時間並不相同，情況與本題如出一轍，而這正是後來被狹義相對論（special relativity）推翻的經典計算。Kleppner 與 Kolenkow 所著的 *An Introduction to Mechanics*（力學入門教材）正是以這種思路先講相對運動，再講相對論。

詳細解答

STEP III 1995，第 12 題

概率與統計 • 徑向對稱密度函數

STEP III 1995，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/95-s3-q12

本題在 *Beyond STEP* 第 12.22 節沒有平均值的分佈中有完整的引導式講解：https://beyondstep.keithmcnulty.org/zh-hant/ch12/cauchy.html

關鍵思路

徑向對稱（radially symmetrical）的密度只透過與原點的距離 $r = \sqrt{u^2+v^2}$ 而取決於 (u,v)，所以繞原點作任何旋轉之後，它看起來都一樣。特別地，點落在以原點為頂點的扇形內的概率，只取決於該扇形的角度。事件 $U/V < k$ 是由直線 $u = kv$ 與 u 軸割出的一對相對的扇形，所以它的概率不過是一個角度除以 2π。

解答

聯合密度與徑向對稱

X 和 Y 各自的密度為 $\phi(t) = \frac{1}{\sqrt{2\pi}}\mathrm{e}^{-t^2/2}$。由於 X 與 Y 獨立（independent），它們的聯合密度（joint density）是邊際密度（marginal density）的乘積：

$$\mathrm{f}(x,y) = \phi(x)\,\phi(y) = \frac{1}{\sqrt{2\pi}}\mathrm{e}^{-x^2/2}\cdot\frac{1}{\sqrt{2\pi}}\mathrm{e}^{-y^2/2} = \frac{1}{2\pi}\mathrm{e}^{-\frac{1}{2}(x^2+y^2)},$$

對所有實數 x 和 y 成立，這正是所求。

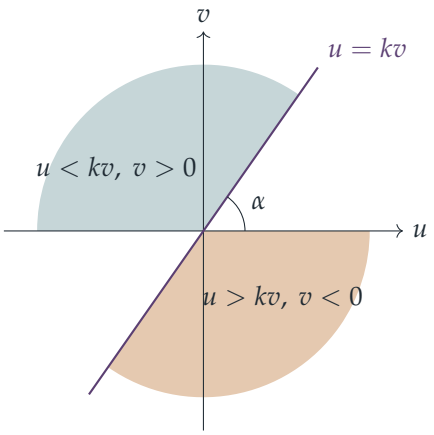
說這個密度是*徑向對稱*的，意思是它在點 (x,y) 的值只取決於該點與原點的距離 $r = \sqrt{x^2+y^2}$，而與該點的方向無關。這裏 $\mathrm{f}(x,y) = \frac{1}{2\pi}\mathrm{e}^{-r^2/2}$。換一個等價的說法：把平面繞原點作任何旋轉，密度都不變，而它的等高線是以原點為圓心的圓。

關於 $\mathrm{P}(U/V < k)$ 的恆等式

設 U 和 V 具有徑向對稱的聯合密度 $h(u,v) = \rho\left(\sqrt{u^2+v^2}\right)$。直線 $v=0$ 的概率為零（它的面積為零），所以我們可以不理會 $V=0$ 的情況。把不等式 $U/V < k$ 乘以 V，並在 $V<0$ 時把不等號反轉，得

$$\left\{\frac{U}{V} < k\right\} = \{U < kV,\ V > 0\} \cup \{U > kV,\ V < 0\},$$

而右邊的兩個事件是互斥（disjoint）的。



如圖所示（圖中取 $k>0$），通過原點的直線 $u=kv$ 與 u 軸把平面分成四個扇形，其中兩個塗上陰影的扇形合起來就是事件 $U/V<k$。映射 $(u,v)\mapsto(-u,-v)$ 是繞原點旋轉 π，它把區域 $\{u<kv,\ v>0\}$ 恰好映到 $\{u>kv,\ v<0\}$ 上（把兩個不等式都乘以 -1 即可）。徑向對稱的密度滿足 $h(-u,-v)=h(u,v)$，所以這兩個區域的概率相同。因此

$$\mathrm{P}\left(\frac{U}{V} < k\right) = 2\,\mathrm{P}(U < kV,\ V > 0),$$

這就是我們要證明的結果。

U/V 的密度

採用極坐標（polar coordinates）$u = r\cos\theta$，$v = r\sin\theta$。區域 $\{u<kv,\ v>0\}$ 是扇形 $\alpha<\theta<\pi$，其中 $\alpha\in(0,\pi)$ 是射線 $u=kv$，$v>0$ 與正 u 軸所成的角，所以 $\cot\alpha = k$。（事實上，當 $0<\theta<\pi$ 時有 $\sin\theta>0$，於是 $u<kv$ 變成 $\cot\theta<k$；由於 cot 在 $(0,\pi)$ 上遞減，這就是 $\theta>\alpha$。）

> **想一想** 在不知道函數 ρ 的情況下，(U,V) 落在以原點為頂點、角度為 β 的扇形內的概率是多少？

在極坐標中，面積元素是 $r\,\mathrm{d}r\,\mathrm{d}\theta$，所以扇形 $\theta_1 < \theta < \theta_1 + \beta$ 的概率為

$$\int_{\theta_1}^{\theta_1+\beta}\int_0^\infty \rho(r)\,r\,\mathrm{d}r\,\mathrm{d}\theta = \beta\int_0^\infty \rho(r)\,r\,\mathrm{d}r = \frac{\beta}{2\pi},$$

因為取 $\beta = 2\pi$ 便得到整個平面，所以 $2\pi\int_0^\infty \rho(r)\,r\,\mathrm{d}r = 1$。換句話說，無論 ρ 是甚麼，(U,V) 的方向都是均勻分佈（uniformly distributed）的。

我們的扇形的角度是 $\pi - \alpha$，所以

$$\mathrm{G}(k) = \mathrm{P}\Big(\frac{U}{V} < k\Big) = 2\cdot\frac{\pi-\alpha}{2\pi} = 1 - \frac{\alpha}{\pi}.$$

由於 $\cot\alpha = k$ 且 $\alpha \in (0,\pi)$，我們有 $\alpha = \frac{\pi}{2} - \arctan k$（兩邊都在 $(0,\pi)$ 內，而且餘切都等於 k，因為 $\cot(\frac{\pi}{2} - \psi) = \tan\psi$）。因此

$$\mathrm{G}(k) = \frac{1}{2} + \frac{1}{\pi}\arctan k, \qquad -\infty < k < \infty.$$

它由 0 上升至 1，正如分佈函數（distribution function）所必須的那樣。求導便得到 U/V 的密度：

$$\mathrm{g}(k) = \mathrm{G}'(k) = \frac{1}{\pi(1+k^2)}, \qquad -\infty < k < \infty,$$

這正是所求。特別地，這個結果適用於第一部分中兩個獨立標準正態變量（standard normal）之比 X/Y。

驗算

- $\mathrm{G}(0) = \frac{1}{2}$：由對稱性，$U/V$ 取正值與取負值的可能性相同。
- $\int_{-\infty}^{\infty}\frac{\mathrm{d}k}{\pi(1+k^2)} = \frac{1}{\pi}\big[\arctan k\big]_{-\infty}^{\infty} = 1$。
- 用三個不同的徑向對稱分佈（兩個獨立的標準正態變量、單位圓盤上的均勻分佈，以及與 e^{-r} 成正比的密度）各取 2×10^5 個點進行模擬，每次都得到 $\mathrm{P}(U/V < 0.7) \approx 0.694$，與 $\frac{1}{2} + \frac{1}{\pi}\arctan 0.7 \approx 0.6944$ 吻合；對於若干個 k 值，恆等式兩邊的差距都在抽樣誤差之內。

常見錯誤

- 寫成 $\{U/V < k\} = \{U < kV\}$：當 $V < 0$ 時，乘以 V 會令不等號反轉。
- 不加細想便取 $\alpha = \arctan(1/k)$：當 $k < 0$ 時它落在 $(-\frac{\pi}{2}, 0)$ 內，而我們需要的角卻在 $(0,\pi)$ 內。寫成 $\alpha = \frac{\pi}{2} - \arctan k$ 則對每個 k 都適用。

STEP 之外

密度 $\frac{1}{\pi(1+k^2)}$ 就是標準柯西分佈（*Cauchy distribution*）。它以沒有平均值而聞名（$|k|\,\mathrm{g}(k)$ 的積分發散），所以大數定律（law of large numbers）對它並不成立：n 個獨立柯西變量的平均值，與其中任何一個變量具有相同的柯西分佈。這裏的極坐標論證，讓你初嘗以雅可比行列式（Jacobian）變換聯合密度的滋味，這是大學一年級的標準技巧。Grimmett 與 Welsh 的 *Probability: An Introduction*（概率論入門）對兩者都有論述。

詳細解答

STEP III 1995，第 13 題

概率與統計 • 有噪聲的二元信道

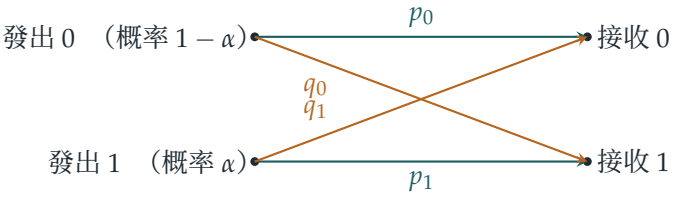
> **STEP III 1995，第 13 題**
>
> 繼續閱讀之前，請先看題目（原題為英文）。題目載於 STEP Support Programme 的試題資料庫：
>
> step.maths.org/questions/95-s3-q13

關鍵思路

一切都由逐個數字考慮而來。每個數字都經過一條「二元信道」（binary channel）：它以 0 或 1 發出，然後要麼原樣保留，要麼被翻轉。一旦弄清楚單一數字的情況，整段訊息不過是許多獨立的複本，所以概率相乘即可。至於增強站，竅門在於留意到：一個數字經過兩段電纜後正確到達，當且僅當它被翻轉了偶數次。

解答

全篇以 $N = 10^k$ 表示訊息的長度。我們按照題目的原意，假設原訊息的各個數字互相獨立（independent）（每個數字是 1 的概率為 α），並假設電纜對每個數字的處理與其他數字無關。單一數字的情況如下圖所示。



整段訊息正確接收

設 C 為某個指定數字被正確接收的事件。這個數字要麼是保留下來的 0，要麼是保留下來的 1，所以

$$\mathrm{P}(C) = (1-\alpha)p_0 + \alpha p_1.$$

由於 N 個數字互相獨立，

$$\mathrm{P}(\text{整段訊息正確}) = \left[(1-\alpha)p_0 + \alpha p_1\right]^{10^k}.$$

概率 β

一個數字被接收為 1，要麼它原本是 1 而且保留下來，要麼它原本是 0 而被翻轉：

$$\beta = \alpha p_1 + (1-\alpha)q_0.$$

現在

$$\begin{aligned}\beta = \alpha &\iff \alpha p_1 + (1-\alpha)q_0 = \alpha \iff (1-\alpha)q_0 = \alpha(1-p_1) = \alpha q_1\\ &\iff q_0 = \alpha(q_0+q_1) \iff \alpha = \frac{q_0}{q_0+q_1},\end{aligned}$$

其中最後一步除以 q_0+q_1，而除非電纜完美無誤（$q_0 = q_1 = 0$，此時 $\beta = \alpha$ 顯然成立），否則這個數必為正數。每一步都可以逆推，所以這就是所要求的「當且僅當」(if and only if)。用文字來說：電纜不改變 1 所佔的比例，恰好就是 $0 \to 1$ 的流量（即 $(1-\alpha)q_0$）與 $1 \to 0$ 的流量（即 αq_1）互相抵銷之時。

全是零的訊息

假設上述條件成立，而且接收到的每個數字都是 0。由於各個位置上的（發出數字，接收數字）這些數對互相獨立，整段訊息正確的條件概率（conditional probability）就是各個單一數字的概率 $\mathrm{P}(\text{發出0} \mid \text{接收0})$ 的乘積。

> ? **想一想** 怎樣把「接收到 0」轉化為一句關於發出了甚麼的陳述？

用貝葉斯定理（Bayes' theorem）。一個數字被接收為 0 的概率是 $1-\beta$，而條件說明 $\beta = \alpha$，所以 $1-\beta = 1-\alpha$。因此

$$\mathrm{P}(\text{發出0} \mid \text{接收0}) = \frac{(1-\alpha)p_0}{1-\beta} = \frac{(1-\alpha)p_0}{1-\alpha} = p_0,$$

於是

$$\mathrm{P}(\text{訊息正確} \mid \text{接收到的全是零}) = p_0^{10^k}.$$

一半一半的機會

當 $q_0 = q_1 = q$ 且 $\alpha = \frac{1}{2}$ 時，每個數字正確的概率是 $\frac{1}{2}(1-q) + \frac{1}{2}(1-q) = 1-q$。對於一段有 10^6 個數字的訊息，我們需要

$$(1-q)^{10^6} = \tfrac{1}{2} \quad\Longrightarrow\quad 10^6 \ln(1-q) = -\ln 2.$$

由於 q 極小，$\ln(1-q) \approx -q$，得出

$$q \approx \frac{\ln 2}{10^6} \approx 6.93 \times 10^{-7}.$$

加設增強站

長度減半會使 q 除以 $2^2 = 4$，因為 q 與長度的平方成正比。所以每半段電纜翻轉一個數字的概率是 $q/4$，而且與另一半段無關（增強站只是把收到的東西原樣再發出）。

> **想一想** 「兩半段都正確傳送」是數字正確到達的唯一途徑嗎？

不是：一個在兩半段都被翻轉的數字會被翻回原狀。所以

$$\mathrm{P}(\text{數字正確}) = \left(1-\frac{q}{4}\right)^2 + \left(\frac{q}{4}\right)^2 = 1 - \frac{q}{2} + \frac{q^2}{8} \approx 1 - \frac{q}{2},$$

這裏略去了 q^2 階的項。（在這裏，雙重翻轉正是這類項之一，所以它對答案沒有影響，但明白箇中原因仍然值得。）整段訊息正確的概率現在是

$$\left(1-\frac{q}{2}\right)^{10^6}.$$

在同一階的近似下，$(1-\frac{q}{2})^2 \approx 1-q$，所以

$$\left(1-\frac{q}{2}\right)^{10^6} \approx (1-q)^{10^6/2} = \left[(1-q)^{10^6}\right]^{1/2} = \left(\tfrac{1}{2}\right)^{1/2} = \frac{1}{\sqrt{2}} \approx 0.707.$$

所以增強站把訊息完全正確的機會由 $\frac{1}{2}$ 提高到約 0.71。留意我們在這裏根本用不着 q 的數值：原來的條件 $(1-q)^{10^6} = \frac{1}{2}$ 已經包辦了一切。同樣的想法配合 $e^{-10^6 q} = \frac{1}{2}$，可得 $e^{-10^6 q/2} = 1/\sqrt{2}$。

驗算

- 把長度為 4 的所有發出訊息和接收訊息逐一準確列舉（取 $p_0 = 0.9$，$p_1 = 0.8$），可以重現 $[(1-\alpha)p_0 + \alpha p_1]^4$，以及在 α 滿足該條件時的條件概率 $p_0^4 = 0.6561$。
- 準確解出 $(1-q)^{10^6} = \frac{1}{2}$ 可得 $q = 6.9315 \times 10^{-7}$，與 $\ln 2/10^6$ 吻合至五位有效數字。
- 取這個 q 值，增強站情況下的準確概率 $\left[(1-\frac{q}{4})^2 + (\frac{q}{4})^2\right]^{10^6}$ 是 0.707107，與 $1/\sqrt{2}$ 相符。
- 合理性檢查：若 $q_0 = q_1$，條件給出 $\alpha = \frac{1}{2}$，正如對稱性所要求。

常見錯誤

- 把每半段的出錯概率當作 $q/2$。出錯概率與長度的平方成正比，所以應是 $q/4$。

► 在「全是零」的部分只答 $(1-\alpha)^N$ 或單單一個 p_0：我們需要的是每個數字的條件概率，再取 10^k 次方。

STEP 之外

這就是信息論（information theory）中的二元信道。當 $q_0 = q_1$ 時，它就是二元對稱信道（binary symmetric channel），是香農（Shannon）理論的起點——這套理論研究有噪聲的信道能夠傳送多少信息；它也是糾錯碼（error-correcting codes）的起點，糾錯碼加入冗餘，使偶爾出現的翻轉可以被偵測和修復，而不僅僅是變得較為罕見。增強站的計算也是一條兩步的馬可夫鏈（Markov chain）：翻轉概率的複合，就是把轉移矩陣（transition matrices）相乘。Grimmett 與 Welsh 的 *Probability: An Introduction*（概率論入門）以恰到好處的程度講解條件概率和馬可夫鏈。

詳細解答

STEP III 1995，第 14 題

概率與統計 • 矩生成函數

> **STEP III 1995，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/95-s3-q14

關鍵思路

矩母函數（moment generating function，簡稱 MGF）能把獨立隨機變量（independent random variables）之和化為乘積：若 T_1 與 T_2 獨立，則 $\mathrm{E}\left[\mathrm{e}^{\theta(T_1+T_2)}\right] = \mathrm{E}\left[\mathrm{e}^{\theta T_1}\right]\mathrm{E}\left[\mathrm{e}^{\theta T_2}\right]$。因此我們先求出一個 MGF，把它自乘若干次，再認出所得結果正是同一族中某個密度函數（density）的 MGF。由於 MGF 決定其分佈，認出它便已足夠。

解答

T 的 MGF

對實數 θ，

$$M_T(\theta) = \mathrm{E}\left[\mathrm{e}^{\theta T}\right] = \int_0^\infty t\,\mathrm{e}^{-(1-\theta)t}\,\mathrm{d}t.$$

這個積分只在 $\theta < 1$ 時收斂（否則被積函數不趨向 0），所以下文一律假設 $\theta < 1$。我們會反覆用到以下積分：

$$\int_0^\infty t^k \mathrm{e}^{-\lambda t}\,\mathrm{d}t = \frac{k!}{\lambda^{k+1}} \qquad (\lambda > 0,\ k = 0, 1, 2, \ldots),$$

這可由分部積分法（integration by parts）連做 k 次得出（或代入 $s = \lambda t$，再利用 $\int_0^\infty s^k \mathrm{e}^{-s}\,\mathrm{d}s = k!$）。取 $k = 1$ 及 $\lambda = 1 - \theta > 0$，得

$$\boxed{M_T(\theta) = \frac{1}{(1-\theta)^2}}, \qquad \theta < 1.$$

簡單核對一下：$M_T(0) = 1$，這是每個 MGF 都必須滿足的。

U 的 MGF

設 T_1 和 T_2 為做這兩題所用的時間，則 $U = T_1 + T_2$，其中 T_1 和 T_2 的密度函數均為 f，而且互相獨立（題目已說明）。於是 $\mathrm{e}^{\theta T_1}$ 與 $\mathrm{e}^{\theta T_2}$ 亦互相獨立，所以它們乘積的期望值（expectation）等於各自期望值的乘積：

$$M_U(\theta) = \mathrm{E}\left[\mathrm{e}^{\theta T_1}\mathrm{e}^{\theta T_2}\right] = M_T(\theta)^2 = \frac{1}{(1-\theta)^4}, \qquad \theta < 1.$$

U 的密度函數

現在計算題目所給的密度函數 $\mathrm{g}(u) = \frac{1}{6}u^3\mathrm{e}^{-u}$ 的 MGF。在上述積分中取 $k = 3$ 及 $\lambda = 1-\theta$，

$$\int_0^\infty \mathrm{e}^{\theta u}\,\tfrac{1}{6}u^3\mathrm{e}^{-u}\,\mathrm{d}u = \frac{1}{6}\cdot\frac{3!}{(1-\theta)^4} = \frac{1}{(1-\theta)^4}.$$

（取 $\theta = 0$ 可確認 g 的積分為 1，所以它確實是一個密度函數。）

> **想一想** 我們已證明 g 與 U 有相同的 MGF。為甚麼這就表示 U 的密度函數是 g？

這裏用到 MGF 的唯一性定理（uniqueness theorem）：若兩個隨機變量的 MGF 在包含 $\theta = 0$ 的某個開區間上存在而且相等，則兩者有相同的分佈。這裏兩個 MGF 對所有 $\theta < 1$ 都相等，當中包括區間 $(-1, 1)$。所以 U 的密度函數為

$$\mathrm{g}(u) = \frac{1}{6}u^3\mathrm{e}^{-u}, \qquad u \geq 0,$$

正如題目所求。

如果你不想使用唯一性定理，卷積（convolution）公式可直接給出同一結果。由於 $T_1, T_2 \geq 0$，對 $u \geq 0$，

$$\begin{aligned}\mathrm{g}(u) &= \int_0^u \mathrm{f}(s)\,\mathrm{f}(u-s)\,\mathrm{d}s = \mathrm{e}^{-u}\int_0^u s(u-s)\,\mathrm{d}s \\ &= \mathrm{e}^{-u}\left[\tfrac{1}{2}us^2 - \tfrac{1}{3}s^3\right]_0^u = \frac{1}{6}u^3\mathrm{e}^{-u}.\end{aligned}$$

做 n 題所用的總時間

設 $U_n = T_1 + \cdots + T_n$，其中各 T_i 互相獨立，密度函數均為 f。同樣的乘積論證給出

$$M_{U_n}(\theta) = M_T(\theta)^n = \frac{1}{(1-\theta)^{2n}}, \qquad \theta < 1.$$

由 $n = 2$ 的規律可以猜想密度函數為 $\mathrm{h}_n(u) = c\,u^{2n-1}\mathrm{e}^{-u}$。利用上述積分並取 $k = 2n-1$，它的 MGF 是 $c\,(2n-1)!/(1-\theta)^{2n}$。這恰好在 $c = 1/(2n-1)!$ 時等於 M_{U_n}。再次根

據唯一性，做 n 題所用的總時間的密度函數為

$$\boxed{\mathrm{h}_n(u) = \frac{u^{2n-1}\mathrm{e}^{-u}}{(2n-1)!}, \qquad u \geq 0.}$$

當 $n = 1$ 時這就是 f，當 $n = 2$ 時這就是 g，與預期一致。

驗算

- $\mathrm{E}[T] = M_T'(0) = 2$，與 $\int_0^\infty t^2\mathrm{e}^{-t}\,\mathrm{d}t = 2$ 吻合。所以 $\mathrm{E}[U_n] = 2n$。由 h_n 亦得出同一數值：$\int_0^\infty u^{2n}\mathrm{e}^{-u}\,\mathrm{d}u/(2n-1)! = (2n)!/(2n-1)! = 2n$。
- 數值上，用森遜法則（Simpson's rule）可得 $\int_0^\infty \mathrm{e}^{0.5t}\,t\mathrm{e}^{-t}\,\mathrm{d}t = 4.000$，與 $1/(1-0.5)^2$ 吻合。模擬 2×10^5 次，每次從 f 抽取三個時間並求和，得出 $\mathrm{P}(U_3 \leq 5) \approx 0.386$，而由 h_3 算得 $\int_0^5 u^5\mathrm{e}^{-u}/120\,\mathrm{d}u \approx 0.384$。

常見錯誤

- 忘記 $\theta < 1$ 這個限制：當 $\theta \geq 1$ 時 MGF 並不存在。
- 把 MGF 相加而不是相乘。U 的 MGF 是 $M_T(\theta)^2$，而不是 $2M_T(\theta)$。（$2T$ 的 MGF 則是 $M_T(2\theta)$，又是另一回事。）
- 弄錯歸一化常數（normalising constant）中的階乘。u 的冪是 $2n-1$，所以常數是 $1/(2n-1)!$，而不是 $1/(2n)!$。

STEP 之外

密度函數 $t\mathrm{e}^{-t}$ 是形狀參數（shape）為 2、速率參數（rate）為 1 的伽瑪分佈（gamma distribution），而 h_n 則是形狀參數為 $2n$ 的伽瑪分佈。這道題說明：速率相同的獨立伽瑪變量相加時，形狀參數亦相加。特別地，每個 T 的表現就如兩個獨立指數分佈（exponential）時間之和，而伽瑪分佈正是這樣作為泊松過程（Poisson process）中第 k 個事件的等候時間而出現的。在大學裏，你會遇到特徵函數（characteristic function）$\mathrm{E}[\mathrm{e}^{\mathrm{i}\theta X}]$，它必定存在，並可用來證明中心極限定理（central limit theorem）。Grimmett 與 Welsh 的 *Probability: An Introduction*（概率論入門）對母函數（generating functions）及這些概念有清晰的闡述。

Beyond STEP

STEP 是大多數中學生一生中應考過最艱深的數學考試，也是最出色的考試之一：題目簡短、陌生而深刻，當中最精彩的，其實是喬裝了的大學數學小品。

本書精選最有趣的 STEP 試題，由 1987 年首次開考之前的樣本試卷（specimen papers），到 2025 年的試卷，並以每一道題目作為一扇門。每道題目都會仔細解答，沿途提出優秀學生會問的問題，然後帶你穿過這扇門：這個想法從何而來，到了大學會演變成甚麼，以及接下來該讀些甚麼。

本書為正在準備 STEP、有天分又有志投身數學的學生而寫。他們想要的不止是模擬試卷：他們想開始像數學系本科生那樣思考，並一探前方有甚麼正等着自己。

- 共 13 章，涵蓋純數學（pure mathematics）、力學（mechanics）、概率（probability）與統計（statistics）
- 136 道 STEP 歷屆試題，每題均附引導式解答
- 每道題目背後的大學數學
- 附答案的延伸練習，以及延伸閱讀

1995 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。