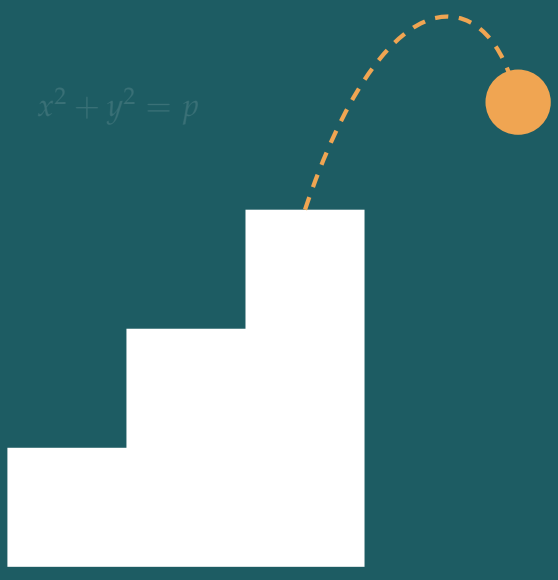
STEP 備試 · 大學數學

Beyond STEP

STEP I, II 及 III

1994 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

1994 年詳細解答

Keith McNulty

Beyond STEP：以歷屆 STEP 試題為門，走進大學數學

1994 年詳細解答

1994 年 STEP 試卷（STEP I, II 及 III）全部 38 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 2 題 對數微分法 純數學 . . . 3

第 4 題 反正切積分 純數學 . . . 7

第 5 題 拋物線上的軌跡 純數學 . . . 10

第 6 題 複函數的複合 純數學 . . . 14

第 8 題 代換積分法 純數學 . . . 17

第 9 題 拋體運動 力學 . . . 20

第 10 題 圓周運動 力學 . . . 24

第 11 題 鏈面上的連接質點 力學 . . . 28

第 13 題 指示變量 概率與統計 . . . 32

第 14 題 離數分佈的順序統計量 概率與統計 . . . 35

STEP II

第 1 題 整除性與十進制數字 純數學 . . . 38

第 2 題 勒讓德多項式 純數學 . . . 41

第 3 題 函數方程 純數學 . . . 44

第 4 題 橢圓的面積 純數學 . . . 48

第 5 題 實根的計數 純數學 . . . 53

第 6 題 三角函數裂項相消級數 純數學 . . . 57

第 7 題 阿波羅尼奧斯圓 純數學 . . . 60

第 8 題 二次模型的最優化 純數學 . . . 64

第 9 題 剛體的平衡 力學 . . . 68

第 10 題 彈性繩與簡諧運動 力學 . . . 72

第 11 題 拋體與爆炸 力學 . . . 75

第 12 題 條件概率 概率與統計 . . . 79

第 13 題 優惠券收集問題 概率與統計 82
第 14 題 幾何概率 概率與統計 86

STEP III

第 1 題 調和級數的發散 純數學 90
第 2 題 回文四次方程 純數學 94
第 3 題 球面、平面與扣連 純數學 98
第 4 題 常微分方程的奇解 純數學 102
第 5 題 麥克勞林級數 純數學 106
第 6 題 阿根圖上的旋轉 純數學 110
第 7 題 群同構 純數學 114
第 8 題 四平方和恆等式 純數學 117
第 9 題 等時曲線與碰撞 力學 120
第 10 題 相對速度 力學 124
第 11 題 剛體的平衡 力學 128
第 12 題 蹲步分析 概率與統計 132
第 13 題 撞數分佈的壽命 概率與統計 136
第 14 題 幾何概率 概率與統計 139

詳細解答

STEP I 1994，第 2 題

純數學 • 對數微分法

> **STEP I 1994，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/94-s1-q2

關鍵思路

當變數在底數而指數固定時，可以用冪法則（power rule）；當底數固定而變數在指數時，可以用指數函數的求導法則（exponential rule）。一旦 x 同時出現在這兩個位置，兩條法則都不能直接套用。解決辦法是把所有式子寫成 e 的冪，或者等價地，先取對數（logarithm），再作隱函數微分（implicit differentiation）：對數能把難纏的冪化為乘積，而乘積正是我們懂得處理的。

全題假設 $x > 0$（否則 x^x 和 $\ln x$ 都沒有意義），而在 (ii) 中底數 a 必須為正數。

解答

我們會反覆使用的唯一工具：對 $u > 0$，

$$u^v = \mathrm{e}^{v \ln u},$$

所以若 $y = u^v$，則 $\ln y = v \ln u$，再用鏈式法則（chain rule）求導，得 $\dfrac{1}{y}\dfrac{\mathrm{d}y}{\mathrm{d}x} = \dfrac{\mathrm{d}}{\mathrm{d}x}(v \ln u)$。

第 (i) 部分

這裏指數是常數，所以用的就是普通的冪法則：

$$\frac{\mathrm{d}}{\mathrm{d}x}\, x^a = a\, x^{a-1}.$$

如果你想知道為何它對任意實數 a（而不只是整數）都成立，可以寫 $x^a = \mathrm{e}^{a \ln x}$；由鏈式法則得 $\mathrm{e}^{a \ln x} \cdot \frac{a}{x} = a\, x^a / x = a\, x^{a-1}$。

第 (ii) 部分

現在底數是常數。寫成 $a^x = \mathrm{e}^{x\ln a}$，其中 $\ln a$ 只是一個常數，

$$\frac{\mathrm{d}}{\mathrm{d}x}\,a^x = \ln a\,\mathrm{e}^{x\ln a} = a^x \ln a.$$

請留意這與 (i) 有多大分別：答案並不是 $x\,a^{x-1}$。

第 (iii) 部分

設 $y = x^x$。取對數，得 $\ln y = x\ln x$。兩邊求導，右邊用積法則（product rule），

$$\frac{1}{y}\frac{\mathrm{d}y}{\mathrm{d}x} = \ln x + x\cdot\frac{1}{x} = 1 + \ln x,$$

所以

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

想一想 為甚麼不能直接用冪法則，得出 $x\cdot x^{x-1} = x^x$？

冪法則假設指數是常數，而指數函數的求導法則假設底數是常數。這裏兩者都在變。事實上，答案恰好是兩個「錯誤」答案之和：把指數當作固定，得 $x\cdot x^{x-1} = x^x$；把底數當作固定，得 $x^x\ln x$；而真正的導數是 $x^x + x^x\ln x$。這其實是多元鏈式法則（multivariable chain rule）的變相：當 $y = u^v$ 時，每個變數各自貢獻它的變化率。

第 (iv) 部分

設 $y = x^{(x^x)}$，這是一座冪塔，其指數本身就是 x^x。取對數，

$$\ln y = x^x \ln x.$$

用積法則對右邊求導，其中 x^x 的導數取自 (iii)：

$$\begin{aligned}\frac{1}{y}\frac{\mathrm{d}y}{\mathrm{d}x} &= x^x(1+\ln x)\cdot\ln x + x^x\cdot\frac{1}{x}\\ &= x^x\left(\frac{1}{x} + \ln x + (\ln x)^2\right).\end{aligned}$$

乘以 $y = x^{(x^x)}$，並把 x 的冪合併，

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

等價地，把 $\frac{1}{x}$ 收進括號內，得 $x^{x^x+x-1}(1 + x\ln x + x(\ln x)^2)$。

第 (v) 部分

? **想一想** $(x^x)^x$ 不就是 (iv) 中的式子嗎？

不是。根據指數定律（law of indices）$(u^m)^n = u^{mn}$，我們有 $(x^x)^x = x^{x\cdot x} = x^{x^2}$，而 (iv) 中的指數是 x^x，不是 x^2。兩者碰巧在 $x = 1$ 和 $x = 2$ 時相等（在 $x = 2$ 時都等於 16），但在 $x = 3$ 時，一個是 3^{27}，另一個只是 3^9。

因此設 $y = x^{x^2}$。於是 $\ln y = x^2 \ln x$，由積法則得

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

因此

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

（你也可以直接對 $(x^x)^x$ 求導：$\ln y = x\ln(x^x) = x^2\ln x$，結果一樣。）

驗算

- 在 $x = 1$ 時，(iii)、(iv) 和 (v) 的導數全都等於 1：例如在 (iv) 中，$1^{1+1}(1+0+0) = 1$。用中心差分（central difference）$\big(f(1.000001) - f(0.999999)\big)/0.000002$ 計算，結果吻合。
- x^x 在 $x = \mathrm{e}^{-1}$ 處有極小值，該處 $1+\ln x = 0$；而 $(x^x)^x = x^{x^2}$ 在 $x = \mathrm{e}^{-1/2}$ 處有極小值，該處 $1+2\ln x = 0$。兩者都可以在圖像計算機上看到。
- 我們在 $0.2 \le x \le 2.5$（以及 $0.2 \le a \le 4$）的範圍內隨機取了 2000 個點，把每條公式與數值導數比較；相對誤差始終小於 10^{-9}。

常見錯誤

- 寫成 $\frac{\mathrm{d}}{\mathrm{d}x}x^x = x\cdot x^{x-1}$ 或 $\frac{\mathrm{d}}{\mathrm{d}x}a^x = x\,a^{x-1}$：冪法則要求指數是常數。
- 把 (iv) 和 (v) 當作同一個函數。冪運算不滿足結合律(associative)：一般而言 $x^{(x^x)} \neq (x^x)^x$。
- 對 $\ln y$ 求導之後，忘記乘回 y。

STEP 之外

當 a 是無理數時，例如 $x^{\sqrt{2}}$，x^a 究竟是甚麼意思？大學裏的答案是：先用冪級數（power series）定義 exp，再把 ln 定義為它的反函數，然後對 $u > 0$ 定義 $u^v = \mathrm{e}^{v \ln u}$；這樣一來，所有指數定律和上面的導數公式都成為定理。Abbott 的 *Understanding Analysis*，或 Bartle 與 Sherbert 的 *Introduction to Real Analysis*，都對此有嚴謹的處理。(iii) 中「兩個錯誤答案之和」，正是對偏導數（partial derivative）和多元鏈式法則的初步一瞥。

詳細解答

STEP I 1994，第 4 題

純數學 • 反正切積分

> **STEP I 1994，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/94-s1-q4

關鍵思路

第 (i) 部分是一條半角恆等式（half-angle identity），第 (ii) 部分則是標準的「先配方（complete the square），再用 $\tan^{-1}$」積分。最後的積分就是第 (ii) 部分取 $k = \cos\alpha$ 的情況，這時 $\sqrt{1-k^2} = \sin\alpha$，剛好與分子的因子 $\sin\alpha$ 約去。唯一真正要小心的地方在最後：對某個角的正切取 $\tan^{-1}$，只有當該角落在 $\left(-\frac{\pi}{2}, \frac{\pi}{2}\right)$ 之內時，才會得回原來的角。

解答

第 (i) 部分

利用二倍角公式(double-angle formulae)$1 - \cos\alpha = 2\sin^2\frac{1}{2}\alpha$ 及 $\sin\alpha = 2\sin\frac{1}{2}\alpha\cos\frac{1}{2}\alpha$，

$$\frac{1-\cos\alpha}{\sin\alpha} = \frac{2\sin^2\frac{1}{2}\alpha}{2\sin\frac{1}{2}\alpha\cos\frac{1}{2}\alpha} = \frac{\sin\frac{1}{2}\alpha}{\cos\frac{1}{2}\alpha} = \tan\tfrac{1}{2}\alpha,$$

只要 $\sin\alpha \neq 0$（使等式兩邊均有定義），上式便成立。這正是所要證明的結果。

第 (ii) 部分

將分母配方：

$$1 - 2kx + x^2 = (x-k)^2 + (1-k^2).$$

由於 $|k| < 1$，常數 $1-k^2$ 為正數，故可寫成 $1-k^2 = a^2$，其中 $a = \sqrt{1-k^2} > 0$。（條件 $|k| < 1$ 正是用在這裏：否則分母可以因式分解，得到的便會是對數而不是反正切。）作

代換（substitution）$x-k=a\tan\theta$，則 $\mathrm{d}x=a\sec^2\theta\,\mathrm{d}\theta$，

$$\begin{aligned}\int\frac{\mathrm{d}x}{(x-k)^2+a^2}&=\int\frac{a\sec^2\theta}{a^2\tan^2\theta+a^2}\,\mathrm{d}\theta=\int\frac{1}{a}\,\mathrm{d}\theta=\frac{\theta}{a}+C\\&=\frac{1}{\sqrt{1-k^2}}\tan^{-1}\left(\frac{x-k}{\sqrt{1-k^2}}\right)+C,\end{aligned}$$

這就是所求的結果。　（另一做法是對右邊求導：由鏈式法則（chain rule）可得 $\frac{1}{a}\cdot\frac{1/a}{1+(x-k)^2/a^2}=\frac{1}{(x-k)^2+a^2}$。）

最後的積分

取 $k=\cos\alpha$。由於 $0<\alpha<\pi$，我們有 $|k|<1$，所以第 (ii) 部分適用；又因為 $\sin\alpha>0$，可得

$$\sqrt{1-k^2}=\sqrt{\sin^2\alpha}=\sin\alpha$$

（而不是 $-\sin\alpha$，因為在這個區間上 $\sin\alpha>0$）。因此分子的因子 $\sin\alpha$ 與第 (ii) 部分的因子 $1/\sqrt{1-k^2}$ 互相約去：

$$\begin{aligned}\int_0^1\frac{\sin\alpha}{1-2x\cos\alpha+x^2}\,\mathrm{d}x&=\left[\tan^{-1}\left(\frac{x-\cos\alpha}{\sin\alpha}\right)\right]_0^1\\&=\tan^{-1}\left(\frac{1-\cos\alpha}{\sin\alpha}\right)-\tan^{-1}\left(\frac{-\cos\alpha}{\sin\alpha}\right)\\&=\tan^{-1}\left(\tan\tfrac{1}{2}\alpha\right)+\tan^{-1}(\cot\alpha),\end{aligned}$$

其中第一項用了第 (i) 部分，第二項則用了 $\tan^{-1}$ 是奇函數（odd function）這一事實。

> **想一想** $\tan^{-1}(\tan\theta)$ 是否總是等於 θ？

不是：$\tan^{-1}$ 的取值範圍是 $\left(-\frac{\pi}{2},\frac{\pi}{2}\right)$，所以只有當 θ 在這個區間內時，才有 $\tan^{-1}(\tan\theta)=\theta$。我們逐項檢查。

- 由於 $0<\alpha<\pi$，我們有 $0<\frac{1}{2}\alpha<\frac{\pi}{2}$，所以 $\tan^{-1}\left(\tan\frac{1}{2}\alpha\right)=\frac{1}{2}\alpha$。
- 寫成 $\cot\alpha=\tan\left(\frac{\pi}{2}-\alpha\right)$。由於 $0<\alpha<\pi$，角 $\frac{\pi}{2}-\alpha$ 落在 $\left(-\frac{\pi}{2},\frac{\pi}{2}\right)$ 之內，所以 $\tan^{-1}(\cot\alpha)=\frac{\pi}{2}-\alpha$。

兩項相加，

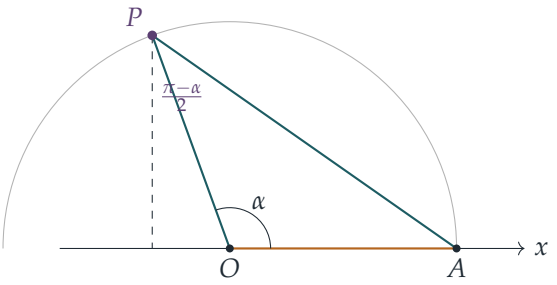
$$\int_0^1\frac{\sin\alpha}{1-2x\cos\alpha+x^2}\,\mathrm{d}x=\frac{\alpha}{2}+\frac{\pi}{2}-\alpha=\frac{\pi-\alpha}{2},$$

這就是所求的結果。

另一做法：一幅圖

這個答案背後有一個很漂亮的幾何原因。設 $P=(\cos\alpha,\sin\alpha)$ 為單位圓上的一點，O 為原點，$A=(1,0)$。當點 $X=(x,0)$ 沿 x 軸移動時，PX 與過 P 向下的鉛垂線之間的夾

角 θ 滿足 $\tan\theta = (x-\cos\alpha)/\sin\alpha$，而被積函數（integrand）正好就是 $\mathrm{d}\theta/\mathrm{d}x$。所以這個積分就是 X 由 O 移到 A 時所掃過的角，即角 OPA。三角形 OPA 是等腰三角形，$OP = OA = 1$，在 O 處的頂角為 α，所以兩個底角各為 $\frac{1}{2}(\pi-\alpha)$。



驗算

- 當 $\alpha = \frac{\pi}{2}$ 時，積分為 $\int_0^1 \frac{\mathrm{d}x}{1+x^2} = \frac{\pi}{4}$，而公式給出 $\frac{1}{2}\left(\pi - \frac{\pi}{2}\right) = \frac{\pi}{4}$。
- 數值上，取 $\alpha = 1$，以辛普森法則（Simpson's rule）計算得 $1.0707963\ldots$，與 $\frac{1}{2}(\pi - 1) = 1.0707963\ldots$ 吻合。
- 當 $\alpha \to \pi^-$ 時，對 $[0,1]$ 內的每個 x，被積函數都趨向 0，而 $\frac{1}{2}(\pi-\alpha)$ 亦然。

常見錯誤

- 寫下 $\tan^{-1}(\cot\alpha) = \frac{\pi}{2} - \alpha$，卻沒有檢查 $\frac{\pi}{2} - \alpha$ 是否落在 $\tan^{-1}$ 的值域之內；當 α 在 $(0,\pi)$ 以外時，同樣的做法便會出錯。
- 把 $\sqrt{1-\cos^2\alpha} = |\sin\alpha|$ 當作 $\sin\alpha$，卻沒有說明原因；這裏的原因是 $0 < \alpha < \pi$。

STEP 之外

當 $\alpha \to 0^+$ 時，答案趨向 $\frac{\pi}{2}$，儘管在 $\alpha = 0$ 時被積函數恆等於零：被積函數在 $x = 1$ 附近變成一個又高又窄的尖峯，而其面積並不縮小。這就是半平面的泊松核（Poisson kernel），而隨意交換極限與積分的次序便會得出錯誤的答案。這種交換在甚麼情況下才容許，是大學一年級分析學（analysis）的核心問題；參見 Abbott 的 *Understanding Analysis*。至於「所對的角」這個幾何觀點，Needham 的 *Visual Complex Analysis* 有十分精彩的闡述。

詳細解答

STEP I 1994，第 5 題

純數學 • 拋物線上的軌跡

> **STEP I 1994，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/94-s1-q5

關鍵思路

先把幾何條件化為代數條件。原點處的角是直角，即位置向量（position vector）$\overrightarrow{OP}$ 與 $\overrightarrow{OQ}$ 互相垂直；對於 $y = x^2$ 上的點，這個條件可簡化為 p 與 q 之間一個簡潔的關係式。這樣一來，R 和 T 兩點的坐標都是關於 p 和 q 對稱的，所以我們把一切都用 $p+q$ 和 pq 表示，然後消元。

解答

直角的條件

P 和 Q 都不可能是原點，否則該角便沒有定義，所以 $p \neq 0$ 且 $q \neq 0$。角 POQ 是直角，當且僅當 $\overrightarrow{OP} = (p, p^2)$ 與 $\overrightarrow{OQ} = (q, q^2)$ 的純量積（scalar product）為零：

$$pq + p^2q^2 = 0 \quad \Longleftrightarrow \quad pq(1+pq) = 0 \quad \Longleftrightarrow \quad pq = -1,$$

其中最後一步除以 $pq \neq 0$。（等價的說法是：OP 與 OQ 的斜率（gradient）分別為 p 和 q，兩者之積為 -1。）

因此我們要找的數對，正是 $q = -1/p$，其中 $p \neq 0$。當 p 取遍所有非零實數時，兩數之和

$$s = p + q = p - \frac{1}{p}$$

可取得每一個實數值：任意給定 s，二次方程 $p^2 - sp - 1 = 0$ 的判別式（discriminant）為 $s^2 + 4 > 0$，所以它有實根，而且兩根之積為 -1，故兩根皆不為零。我們稍後需要這個結論，才能說明下面的軌跡（locus）是完整的曲線，而不只是曲線的一部分。

第 (i) 部分

PQ 的中點（midpoint）是

$$R = \left(\frac{p+q}{2}, \frac{p^2+q^2}{2} \right).$$

設 $R = (x, y)$，則 $p + q = 2x$。利用 $pq = -1$，

$$y = \frac{(p+q)^2 - 2pq}{2} = \frac{4x^2+2}{2} = 2x^2 + 1.$$

由於 $p + q$ 可取得每一個實數值，x 亦然，所以 R 的軌跡是整條拋物線（parabola）

$$y = 2x^2 + 1,$$

其頂點（vertex）為 $(0, 1)$，對稱軸為 y 軸。它比 $y = x^2$ 窄，而且完全位於後者的上方，因為 $2x^2 + 1 - x^2 = x^2 + 1 > 0$；這與 R 是凸（convex）曲線上一條弦（chord）的中點這一事實吻合。

第 (ii) 部分

對 $y = x^2$ 求導，可知在 P 的切線（tangent）斜率為 $2p$，所以切線方程為

$$y - p^2 = 2p(x - p), \qquad \text{即} \qquad y = 2px - p^2,$$

同理，在 Q 的切線為 $y = 2qx - q^2$。這兩條切線並不平行，因為 $p \neq q$（兩者之積為 -1，所以不可能相等）。把兩條方程相減，

$$2(p - q)x = p^2 - q^2 \quad \Longrightarrow \quad x = \frac{p+q}{2},$$

其中除以了 $p - q \neq 0$。於是

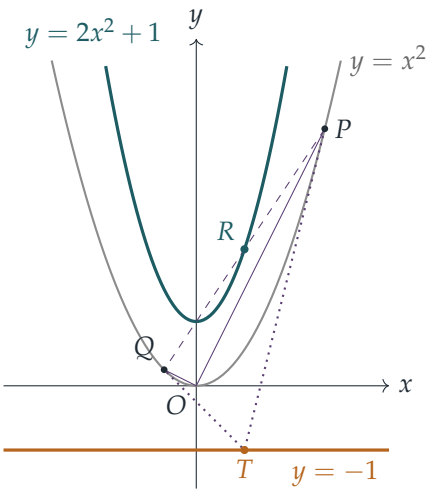
$$y = 2p \cdot \frac{p+q}{2} - p^2 = pq = -1.$$

想一想 T 的 y 坐標是常數。那麼軌跡是單獨一點、一條線段，還是整條直線？

T 的 x 坐標是 $\frac{1}{2}(p + q)$，而我們已證明它可取得每一個實數值。所以 T 可以位於該直線上的任何位置，T 的軌跡就是整條直線

$$y = -1.$$

還可留意，T 總是在 R 的正下方：兩者的 x 坐標都是 $\frac{1}{2}(p + q)$。



圖中所示是其中一個位置，$p = 2$，$q = -\frac{1}{2}$：此時 $R = (\frac{3}{4}, \frac{17}{8})$ 在藍綠色的拋物線上，而 $T = (\frac{3}{4}, -1)$ 在琥珀色的直線上。

驗算

取 $p = 1$，$q = -1$：則 $P = (1,1)$，$Q = (-1,1)$，而 OP、OQ 就是直線 $y = \pm x$，顯然互相垂直。中點是 $(0,1)$，即 $y = 2x^2 + 1$ 的頂點；兩條切線 $y = 2x - 1$ 與 $y = -2x - 1$ 相交於 $(0,-1)$，在 $y = -1$ 上。再取圖中的 $p = 2$，$q = -\frac{1}{2}$，則 $2 \cdot (\frac{3}{4})^2 + 1 = \frac{17}{8}$，與預期相符。另外以 10^5 個隨機的 p 值作數值檢驗，結果確認 R 滿足 $y = 2x^2 + 1$，誤差約在 10^{-10} 以內，而 T 的坐標總是滿足 $y = -1$。

常見錯誤

- 寫到 $pq(1 + pq) = 0$ 便停下，沒有解釋為何可排除 $pq = 0$：這種情況會令 P 或 Q 落在原點。
- 只寫出軌跡的方程，卻沒有檢查是否每一個 x 值都能取得，以確定所畫的草圖應是整條曲線而不是其中一部分。
- 以為 T 位於準線（directrix）上。$y = x^2$ 的準線是 $y = -\frac{1}{4}$；當 PQ 通過焦點（focus）時，兩條切線才會在準線上相交，這與 $\angle POQ = 90°$ 是不同的條件。

STEP 之外

這裏的弦 PQ 總是通過定點 $(0,1)$：它的方程是 $y=(p+q)x-pq=(p+q)x+1$。通過一個定點的直線族，以及與之對應、由切線相交而成的「極點」（pole）（這裏它們位於直線 $y=-1$ 上，即 $(0,1)$ 的極線（polar)），正是射影幾何（projective geometry）的起點；射影幾何會有系統地研究圓錐曲線（conics）及其極點與極線的關係。想做更多同類題目，可參考 Siklos 的 *Advanced Problems in Mathematics*。

詳細解答

STEP I 1994，第 6 題

純數學 • 複函數的複合

> **STEP I 1994，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/94-s1-q6

關鍵思路

要求一個商的實部（real part）和虛部（imaginary part），可把分子和分母同乘以分母的共軛（conjugate）；由於 x 是實數，這個共軛很容易寫出來。至於重複施行函數的部分，與其把 (i) 中繁複的實部和虛部再代入公式一次，不如對一般的 z 以代數方法把 f 與自身複合（compose），這樣簡潔得多。一旦把 $\mathrm{f}(\mathrm{f}(z))$ 化成整齊的形式，最後一部分只需一行代入便可完成。

解答

由於 x 是實數，$\mathrm{i}x+1$ 的實部為 1，所以它永不為零，因此 $\mathrm{f}(x)$ 對每個實數 x 都有定義。

第 (i) 部分

因為 x 是實數，$\mathrm{i}x+1$ 的共軛是 $1-\mathrm{i}x$。把分子和分母同乘以它，

$$\begin{aligned}\mathrm{f}(x) = \frac{(\mathrm{i}x-1)(1-\mathrm{i}x)}{(1+\mathrm{i}x)(1-\mathrm{i}x)} &= \frac{\mathrm{i}x+x^2-1+\mathrm{i}x}{1+x^2} \\ &= \frac{x^2-1}{x^2+1} + \mathrm{i}\,\frac{2x}{x^2+1}.\end{aligned}$$

兩個分數都是實數，所以它們正是所求的實部和虛部。

第 (ii) 部分

對任何複數 w，$ww^* = (\operatorname{Re} w)^2 + (\operatorname{Im} w)^2$。利用第 (i) 部分，

$$\mathrm{f}(x)\mathrm{f}(x)^* = \frac{(x^2-1)^2+4x^2}{(x^2+1)^2} = \frac{x^4+2x^2+1}{(x^2+1)^2} = 1.$$

所以 $|\mathrm{f}(x)| = 1$：f 把每個實數映射到單位圓（unit circle）上的一點。

第 (iii) 部分

我們不把 (i) 的答案代入 f，而是對一般的 z 進行複合。設 $w = \mathrm{f}(z) = \dfrac{\mathrm{i}z-1}{\mathrm{i}z+1}$，並把 $\mathrm{f}(w) = \dfrac{\mathrm{i}w-1}{\mathrm{i}w+1}$ 的分子和分母同乘以 $\mathrm{i}z+1$：

$$\begin{aligned}\mathrm{f}(\mathrm{f}(z)) &= \frac{\mathrm{i}(\mathrm{i}z-1)-(\mathrm{i}z+1)}{\mathrm{i}(\mathrm{i}z-1)+(\mathrm{i}z+1)} = \frac{-z-\mathrm{i}-\mathrm{i}z-1}{-z-\mathrm{i}+\mathrm{i}z+1} \\ &= \frac{-(1+\mathrm{i})(z+1)}{(1-\mathrm{i})(1-z)} = \frac{1+\mathrm{i}}{1-\mathrm{i}}\cdot\frac{z+1}{z-1} = \mathrm{i}\,\frac{z+1}{z-1},\end{aligned}$$

這是因為 $\dfrac{1+\mathrm{i}}{1-\mathrm{i}} = \dfrac{(1+\mathrm{i})^2}{2} = \mathrm{i}$。

> **想一想** $\mathrm{f}(\mathrm{f}(x))$ 是否對每個實數 x 都有定義？

不是。我們需要 $\mathrm{i}w+1 \neq 0$，即 $\mathrm{f}(x) \neq \mathrm{i}$。由第 (i) 部分，$\mathrm{f}(x) = \mathrm{i}$ 需要 $x^2-1=0$ 及 $2x = x^2+1$，而這只在 $x=1$ 時成立。這與上式分母中的因式 $x-1$ 吻合。所以對實數 $x \neq 1$，

$$\mathrm{f}(\mathrm{f}(x)) = \mathrm{i}\,\frac{x+1}{x-1}, \qquad \text{所以} \qquad \operatorname{Re}\mathrm{f}(\mathrm{f}(x)) = 0, \quad \operatorname{Im}\mathrm{f}(\mathrm{f}(x)) = \frac{x+1}{x-1}.$$

第一次施行 f 把實數線映射到單位圓，第二次則把單位圓映射到虛軸。

第 (iv) 部分

寫 $\mathrm{f}(\mathrm{f}(x)) = \mathrm{i}y$，其中 $y = \dfrac{x+1}{x-1}$ 是實數。於是 $\mathrm{i}(\mathrm{i}y) = -y$，所以

$$\mathrm{f}(\mathrm{i}y) = \frac{-y-1}{-y+1} = \frac{y+1}{y-1}.$$

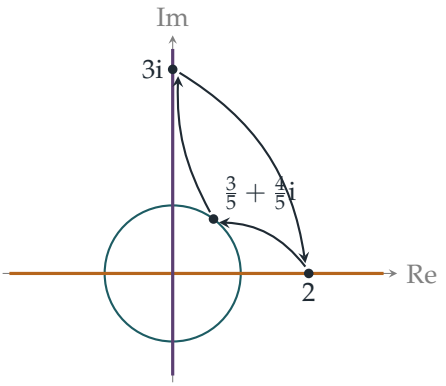
這個式子有定義，因為 $y=1$ 需要 $x+1 = x-1$，而這是不可能的。現在

$$y+1 = \frac{2x}{x-1}, \qquad y-1 = \frac{2}{x-1},$$

所以

$$\mathrm{f}(\mathrm{f}(\mathrm{f}(x))) = x \qquad \text{對每個實數} x \neq 1.$$

施行 f 三次便回到起點：第三次施行把虛軸送回實數線。下圖顯示 $x=2$ 時的循環，即 $2 \mapsto \frac{3}{5}+\frac{4}{5}\mathrm{i} \mapsto 3\mathrm{i} \mapsto 2$。



驗算

- $x = 0$：$\mathrm{f}(0) = -1$，它在單位圓上，與 (i) 一致。然後 $\mathrm{f}(-1) = \dfrac{-\mathrm{i}-1}{-\mathrm{i}+1} = -\mathrm{i}$，與 $\mathrm{i}\frac{0+1}{0-1}$ 吻合，而 $\mathrm{f}(-\mathrm{i}) = \frac{1-1}{1+1} = 0$。
- $x = 2$：第 (i) 部分給出 $\frac{3}{5}+\frac{4}{5}\mathrm{i}$，第 (iii) 部分給出 3i，第 (iv) 部分則回到 2。以 10^4 個隨機實數 x（以及隨機複數 z）作數值檢驗，結果證實以上每條公式都準確至捨入誤差之內。

常見錯誤

- 把 (i) 的實部和虛部直接代入 f。這樣做是可行的，但代數運算繁重，容易出錯。以 z 來進行複合則簡潔得多。
- 忘記在 (iii) 和 (iv) 中必須排除 $x = 1$，因為 $\mathrm{f}(1) = \mathrm{i}$，而 $\mathrm{f}(\mathrm{i})$ 的分母為零。

STEP 之外

形如 $z \mapsto \dfrac{az+b}{cz+d}$ 的函數稱為莫比烏斯變換（*Möbius transformation*）。這類映射把圓和直線映射到圓和直線，正如我們在這裏所見：實軸到單位圓，再到虛軸，然後返回實軸。把兩個這樣的變換複合，相當於把 2×2 矩陣 $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ 相乘。事實上，對每個複數 z 都有 $\mathrm{f}(\mathrm{f}(\mathrm{f}(z))) = z$，所以 f 在莫比烏斯變換所組成的羣（group）中的階（order）為 3。此外，在 (i) 中代入 $x = \tan\frac{\theta}{2}$，可得 $\mathrm{f}(x) = -\cos\theta + \mathrm{i}\sin\theta$，這就是大家熟悉的「$t$ 代換」圓參數化（parametrisation）。Needham 的 *Visual Complex Analysis* 中有一章精彩地講述莫比烏斯變換及其幾何。

詳細解答

STEP I 1994，第 8 題

純數學 • 代換積分法

STEP I 1994，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/94-s1-q8

關鍵思路

代換（substitution）$\theta = \frac{1}{4}\pi - \phi$ 把區間 $[0, \frac{1}{4}\pi]$ 反射到自身，而複角公式（addition formula）則把 $1 + \tan(\frac{1}{4}\pi - \phi)$ 化為 $2/(1 + \tan\phi)$。於是原來的積分會帶着負號在等式右邊重現，我們只須把它解出來，而不必計算任何困難的積分。後面兩個積分其實是同一個積分的變相：一個經代換 $x = \tan\theta$ 得出，另一個則經半角公式（half-angle formulae）得出。

解答

第一個積分

把這個積分記為 I。設 $\theta = \frac{1}{4}\pi - \phi$，則 $\mathrm{d}\theta = -\mathrm{d}\phi$；當 $\theta = 0$ 時，$\phi = \frac{1}{4}\pi$，而當 $\theta = \frac{1}{4}\pi$ 時，$\phi = 0$。因此

$$I = \int_{\frac{1}{4}\pi}^{0} \ln\bigl(1 + \tan(\tfrac{1}{4}\pi - \phi)\bigr)\,(-\mathrm{d}\phi) = \int_{0}^{\frac{1}{4}\pi} \ln\bigl(1 + \tan(\tfrac{1}{4}\pi - \phi)\bigr)\,\mathrm{d}\phi.$$

利用複角公式及 $\tan\frac{1}{4}\pi = 1$，

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

（在 $[0, \frac{1}{4}\pi]$ 上有 $\tan\phi \geq 0$，所以沒有任何式子等於零，而每個對數都是取正數的對數。）故此

$$I = \int_{0}^{\frac{1}{4}\pi} \bigl(\ln 2 - \ln(1 + \tan\phi)\bigr)\,\mathrm{d}\phi = \tfrac{1}{4}\pi \ln 2 - I,$$

因為最後一個積分正是 I，只是啞變數（dummy variable）換了名稱。

想一想 我們由始至終沒有求出原函數（antiderivative）。為甚麼可以這樣做？

我們根本不需要原函數。I 只是一個數，而代換給出了這個數必須滿足的方程 $I = \frac{1}{4}\pi\ln 2 - I$。解之，得

$$2I = \tfrac{1}{4}\pi\ln 2 \quad\Longrightarrow\quad \int_0^{\frac{1}{4}\pi} \ln(1+\tan\theta)\,\mathrm{d}\theta = \tfrac{1}{8}\pi\ln 2,$$

正是所求。

含 $1+x^2$ 的積分

因子 $\dfrac{\mathrm{d}x}{1+x^2}$ 提示我們設 $x = \tan\theta$，於是 $\mathrm{d}x = \sec^2\theta\,\mathrm{d}\theta = (1+x^2)\,\mathrm{d}\theta$。積分上下限 $x=0$ 和 $x=1$ 變成 $\theta = 0$ 和 $\theta = \frac{1}{4}\pi$。因此

$$\int_0^1 \frac{\ln(1+x)}{1+x^2}\,\mathrm{d}x = \int_0^{\frac{1}{4}\pi} \ln(1+\tan\theta)\,\mathrm{d}\theta = \tfrac{1}{8}\pi\ln 2.$$

含 $\sin x$ 和 $\cos x$ 的積分

把它記為 J。最快捷的方法是利用對稱性（symmetry）：代換 $x = \frac{1}{2}\pi - u$ 把 sin 和 cos 互換，並把 $[0, \frac{1}{2}\pi]$ 映射到自身，所以

$$J = \int_0^{\frac{1}{2}\pi} \ln\left(\frac{1+\cos u}{1+\sin u}\right)\mathrm{d}u = -J,$$

因此 $J = 0$。（換個說法，$\int_0^{\frac{1}{2}\pi}\ln(1+\sin x)\,\mathrm{d}x$ 與 $\int_0^{\frac{1}{2}\pi}\ln(1+\cos x)\,\mathrm{d}x$ 相等，所以兩者之差為零。）

想一想 第一個結果在哪裏派上用場？

J 與 I 之間的關係值得一看，這大概也是題目把它放在這裏的原因。設 $t = \frac{1}{2}x$，由半角公式得

$$1+\cos x = 2\cos^2 t, \qquad 1+\sin x = (\cos t + \sin t)^2,$$

所以

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

當 x 取遍 $[0, \frac{1}{2}\pi]$ 時，t 取遍 $[0, \frac{1}{4}\pi]$，且 $\mathrm{d}x = 2\,\mathrm{d}t$。因此

$$J = \int_0^{\frac{1}{4}\pi} \big(2\ln(1+\tan t) - \ln 2\big)\,2\,\mathrm{d}t = 4I - \tfrac{1}{2}\pi\ln 2 = \tfrac{1}{2}\pi\ln 2 - \tfrac{1}{2}\pi\ln 2 = 0,$$

與對稱性的論證吻合。所以

$$\int_0^{\frac{1}{2}\pi} \ln\left(\frac{1+\sin x}{1+\cos x}\right)\mathrm{d}x = 0.$$

驗算

用辛普森法則(Simpson's rule)計算，tan 的積分和含 $1+x^2$ 的積分都得出 0.2721982613，與 $\frac{1}{8}\pi\ln 2 = 0.2721982613\ldots$ 相符；第三個積分在機器精度之內等於 0。也可以用手粗略驗算：被積函數（integrand）$\ln(1+\tan\theta)$ 由 0 上升至 $\ln 2$，所以 I 應約為 $\frac{1}{4}\pi\ln 2$ 的一半。這恰好就是答案。

常見錯誤

- 忘記了 $\mathrm{d}\theta = -\mathrm{d}\phi$，同時也忘記了上下限要對調。這兩次變號會互相抵消；若只保留其中一次，便會得出 $I = -\frac{1}{4}\pi\ln 2 - I$（以及一個正積分竟得負數答案）。
- 寫成 $\tan(\frac{1}{4}\pi - \phi) = 1 - \tan\phi$，漏掉了複角公式中的分母。
- 在第三個積分上花了很長時間，卻沒有留意到 $x \mapsto \frac{1}{2}\pi - x$ 這個對稱性。

STEP 之外

以反射作代換，再把結果與原式相加，從而得到 $\int_a^b f(x)\,\mathrm{d}x = \int_a^b f(a+b-x)\,\mathrm{d}x$，這個技巧在某些書中稱為 King 性質（King property）。它同樣能給出經典結果 $\int_0^{\frac{1}{2}\pi}\ln\sin x\,\mathrm{d}x = -\frac{1}{2}\pi\ln 2$。在大學裏，這類結果屬於含參數定積分（definite integrals with parameters）和圍道積分（contour integration）的理論，當中一整族的積分都可以不求原函數而算出來。Needham 的 *Visual Complex Analysis*（《複分析：可視化方法》）展示了複變方法如何得出這類結果，而 Siklos 的 *Advanced Problems in Mathematics*（進階數學難題集）則有更多運用這類對稱技巧的 STEP 式練習。

詳細解答

STEP I 1994，第 9 題

力學 • 拋體運動

> **STEP I 1994，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/94-s1-q9

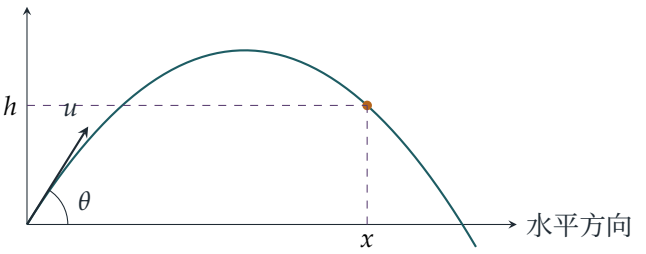
關鍵思路

發射角 θ 是未知的量，所以要把它消去：水平方向和垂直方向的方程分別含有 $u\cos\theta$ 或 $u\sin\theta$，把兩式平方後相加，就可以完全消去 θ。這樣得出的方程是 t^2 的二次方程（quadratic），而二次方程只有在判別式（discriminant）非負時才有實根（real roots）。這個條件正好就是射程的上界。

解答

關於 t 的方程

把炮彈看成只受重力作用的質點（particle），從原點以仰角 θ 發射。經過時間 t 後，它的水平距離為 x，高度為 h，所以它與大炮的距離是 $\sqrt{x^2+h^2}$，與題目所述相符。



利用向下的恆定加速度（constant acceleration）g，以及水平方向不受力，

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

把含 θ 的項分離出來：

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

兩式平方後相加，並利用 $\cos^2\theta + \sin^2\theta = 1$，

$$\begin{aligned} u^2t^2 &= x^2 + \left(h + \tfrac{1}{2}gt^2\right)^2 \\ &= x^2 + h^2 + ght^2 + \tfrac{1}{4}g^2t^4, \end{aligned}$$

整理後得

$$\tfrac{1}{4}g^2t^4 - (u^2 - gh)t^2 + h^2 + x^2 = 0,$$

這正是所求。

射程的上界

現在固定 u 和 h，考慮在高度 h 處可以達到哪些水平距離 x。如果炮彈在某時刻 t 位於 (x, h)，那麼 $T = t^2$ 就是以下方程的一個實根：

$$\tfrac{1}{4}g^2T^2 - (u^2 - gh)T + (h^2 + x^2) = 0.$$

> **想一想** 這條方程涉及 x、h、u 和 t，但我們想要的是一個只關於 x 的結論。二次方程若有實根，必須滿足甚麼條件？

有實根就需要判別式非負：

$$\begin{aligned} (u^2 - gh)^2 - 4\cdot\tfrac{1}{4}g^2(h^2 + x^2) &\geq 0 \\ u^4 - 2u^2gh + g^2h^2 - g^2h^2 - g^2x^2 &\geq 0 \\ g^2x^2 &\leq u^2(u^2 - 2gh). \end{aligned}$$

由於 $u^2 > 2gh$，右邊為正數，兩邊取非負平方根得

$$x \leq \frac{u\sqrt{u^2 - 2gh}}{g}.$$

所以在高度 h 處的水平射程永遠不會超過這個值，這正是我們要證明的。（如果 $u^2 < 2gh$，則不等式對任何 x 都不成立：高度 h 超過了炮彈所能達到的最大高度 $u^2/2g$。）

上界可以達到

等號成立表示判別式為零，所以方程有重根（double root）

$$t^2 = T = \frac{u^2 - gh}{2\cdot\frac{1}{4}g^2} = \frac{2(u^2 - gh)}{g^2}.$$

這個值是正數（若 $h \geq 0$，因為 $u^2 > 2gh \geq gh$；若 $h < 0$，則顯然成立），所以它確實給出一個時刻 $t > 0$。我們仍須檢查，是否有某個角 θ 能在 $x = u\sqrt{u^2 - 2gh}/g$ 的情況下給出這個時刻。由兩條分量方程，我們需要

$$\sin\theta = \frac{h + \frac{1}{2}gt^2}{ut}, \qquad \cos\theta = \frac{x}{ut}.$$

代入上面 t^2 的值，$h + \frac{1}{2}gt^2 = h + (u^2 - gh)/g = u^2/g$，所以

$$\sin\theta = \frac{u}{gt}, \qquad \cos\theta = \frac{\sqrt{u^2 - 2gh}}{gt}.$$

想一想 我們分別為 $\sin\theta$ 和 $\cos\theta$ 選定了數值。怎樣知道有同一個角同時取得這兩個值？

兩個值都是正數，而且根據 t^2 的值，

$$\sin^2\theta + \cos^2\theta = \frac{u^2 + u^2 - 2gh}{g^2t^2} = \frac{2(u^2 - gh)}{g^2t^2} = 1$$

。所以在 0 與 $\frac{1}{2}\pi$ 之間恰好有一個角 θ 符合要求，即

$$\tan\theta = \frac{u}{\sqrt{u^2 - 2gh}}.$$

以這個角度發射，炮彈在時刻 t 位於高度 h，水平距離恰好是 $u\sqrt{u^2 - 2gh}/g$，所以當 $u^2 > 2gh$ 時，上界總是可以達到的。

留意當 $h > 0$ 時 $\tan\theta > 1$：要在高於大炮的某個指定高度上射得最遠，發射角要比 45° 更陡。此外，該時刻的垂直速度為 $u\sin\theta - gt = (u^2 - g^2t^2)/(gt) = (2gh - u^2)/(gt) < 0$，所以炮彈是在下降途中到達高度 h 處的最遠點。

驗算

- 當 $h = 0$ 時，上界為 u^2/g，而 $\tan\theta = 1$：這就是大家熟悉的平地最大射程，發射角為 45°。
- 當 $u = 20\,\mathrm{m\,s^{-1}}$、$g = 9.8\,\mathrm{m\,s^{-2}}$ 而 $h = 10\,\mathrm{m}$ 時，上界約為 29.15 m，此時 $\theta \approx 54.5°$。
- 我們以數值方法確認了該四次方程（quartic）在隨機選取的軌跡上均成立，而且掃描數千個發射角，結果從未超過上界，但可以接近至相差不足 10^{-4}。上面求得的角則恰好達到上界。

常見錯誤

- 嘗試把 x 對 θ 求導來找出最佳的 θ。這樣做是可以的，但過程繁複。判別式已經替你完成了求最優值的工作。
- 忘記最後一步：判別式為零只給出 t 的一個值，但你仍須驗證 $\sin^2\theta + \cos^2\theta = 1$，以證明確實有一個實數角能給出這個值。

STEP 之外

炮彈以速率 u 發射時所能到達的點構成一個區域，其邊界是所有軌跡的 包絡線（envelope），即「安全拋物線」（parabola of safety）$y = \frac{u^2}{2g} - \frac{gx^2}{2u^2}$。把本題的上界重新整理，可以看出 (x, h) 可以到達，當且僅當它位於這條拋物線上或其下方。曲線族的包絡線同樣源自「重根」這個想法，並且會在變分法（calculus of variations）和光學（焦散線，caustics）中再次出現。想進一步了解拋體（projectiles）及其背後的力學，可參閱 Kleppner 與 Kolenkow 的 *An Introduction to Mechanics*（力學入門）。

詳細解答

STEP I 1994，第 10 題

力學 • 圓周運動

> **STEP I 1994，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/94-s1-q10

關鍵思路

這是一個用可伸長的繩繫着的圓錐擺（conical pendulum）。質點沒有垂直方向的加速度，所以張力（tension）的垂直分量與重量平衡，而水平分量則提供向心力（centripetal force）。新的地方在於圓的半徑取決於伸長量（extension），而伸長量又取決於張力，於是虎克定律（Hooke's law）把一切連繫起來。至於不等式，則來自這個問題：$\cos\theta$ 取哪些值才真正描述一個圓？

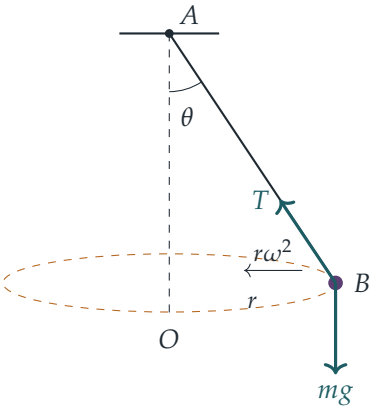
解答

求 $\cos\theta$

把 B 視為質點（particle），把 AB 視為符合虎克定律的輕彈性繩（light elastic string）。設伸長量為 x，張力為 T，則

$$T = \frac{\lambda x}{l}, \qquad AB = l + x.$$

質點繞着 A 正下方的點 O，在半徑為 $r = (l + x)\sin\theta$ 的水平圓上運動。作用於質點的力只有它的重量和沿 BA 方向的張力。



沿垂直方向分解（沒有垂直加速度）：

$$T\cos\theta = mg. \tag{1}$$

沿水平方向分解，指向圓心 O（加速度為 $r\omega^2$）：

$$T\sin\theta = m\omega^2(l+x)\sin\theta.$$

要是真正的圓，必須 $\sin\theta \neq 0$，所以可以把它約去：

$$T = m\omega^2(l+x). \tag{2}$$

現在由虎克定律得 $\lambda x/l = m\omega^2(l+x)$，即

$$x\left(\lambda - m\omega^2 l\right) = m\omega^2 l^2.$$

由於 $x > 0$（由 (1) 得 $T > 0$，所以繩處於拉伸狀態）而且 $m\omega^2 l^2 > 0$，可知 $\lambda - m\omega^2 l > 0$，並且

$$x = \frac{m\omega^2 l^2}{\lambda - m\omega^2 l}, \qquad l + x = \frac{l\lambda}{\lambda - m\omega^2 l}.$$

於是由 (2) 得

$$T = \frac{m\omega^2 l\lambda}{\lambda - m\omega^2 l},$$

再由 (1) 得

$$\boxed{\cos\theta = \frac{mg}{T} = \frac{g\left(\lambda - m\omega^2 l\right)}{l\lambda\,\omega^2}}$$

運動何時可能發生

想一想 對 $\cos\theta$ 的每一個限制，在物理上各代表甚麼？

要質點在 A 下方的水平圓上運動，而且繩保持拉緊，我們需要 $0 < \theta < \frac{1}{2}\pi$，即 $0 < \cos\theta < 1$。反過來說，如果 ω 代入公式後得出 $0 < \cos\theta < 1$，那麼把上述步驟逆推回去（其中 $\lambda - m\omega^2 l > 0$），便會得到滿足 (1) 和 (2) 的正伸長量和正張力，所以這個運動確實存在。

上界。由於其餘因子均為正數，$\cos\theta > 0$ 成立當且僅當 $\lambda - m\omega^2 l > 0$。這就是

$$\omega^2 < \frac{\lambda}{ml}.$$

物理意義：如果 ω 太大，繩無論伸長多少（只要是有限的）都無法提供所需的向心力，因為多出來的長度所需要的額外的力，增長得比繩所能提供的還要快。

下界。$\cos\theta < 1$ 即是

$$g\lambda - gm\omega^2 l < l\lambda\omega^2 \iff g\lambda < \omega^2 l(\lambda + mg) \iff \omega^2 > \frac{g\lambda}{l(\lambda + mg)},$$

其中我們先除以 $l\lambda\omega^2 > 0$，最後除以 $l(\lambda + mg) > 0$，所以不等號的方向保持不變。物理意義：如果 ω 太小，公式便會要求 $\cos\theta \geq 1$；唯一的可能就是質點靜止地懸掛在軸上（$\theta = 0$），而這並不是圓周運動（circular motion）。

把兩者合起來，這個運動只有在以下條件下才可能發生：

$$\frac{g\lambda}{l(\lambda + mg)} < \omega^2 < \frac{\lambda}{ml},$$

這正是所要證明的結果。

驗算

- 在下界處，$\cos\theta = 1$：這時繩是垂直的，張力為 mg，長度為 $l(\lambda + mg)/\lambda$，恰好就是質點懸掛時的平衡長度。留意 $\omega^2 = g/(\text{該長度})$，即熟悉的圓錐擺臨界條件 $\omega^2 > g/L$。
- 當 $\lambda \to \infty$（不可伸長的繩）時，結果變成 $\cos\theta = g/(l\omega^2)$，即標準的圓錐擺，而範圍則變成 $\omega^2 > g/l$。
- 數值上，取 $g = 9.8$ 並隨機選取 m、l、λ 和 ω，直接用二分法（bisection）解虎克定律和水平方向的方程，所得的 $\cos\theta$ 與公式相同，而且兩條分解所得的方程都成立；在 20 000 個隨機個案中，運動存在當且僅當不等式成立。

常見錯誤

- 把半徑取為 $l\sin\theta$ 而不是 $(l+x)\sin\theta$：繩是被拉長了的。
- 忘記說明為何可以除以 $\sin\theta$，因而看不出為甚麼要排除 $\theta = 0$（從而排除 $\cos\theta = 1$）。
- 只證明了 $\cos\theta < 1$，卻忘記 $\cos\theta > 0$（等價於伸長量為正）才給出上界。

STEP 之外

這個上界是失控（runaway）或不穩定性（instability）的一個簡單例子：當 $\omega^2 \to \lambda/(ml)$ 時，所需的伸長量會趨向無窮大。同樣的效應也出現在旋轉的彈性桿以及自旋物體的穩定性之中；而下界則是一個分歧（bifurcation），懸掛的平衡狀態在那裏讓位予圓周運動。想進一步了解旋轉參考系中的力學，可參閱 Kleppner 與 Kolenkow 的 *An Introduction to Mechanics*（力學入門）。

詳細解答

STEP I 1994，第 11 題

力學 • 斜面上的連接質點

STEP I 1994，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：

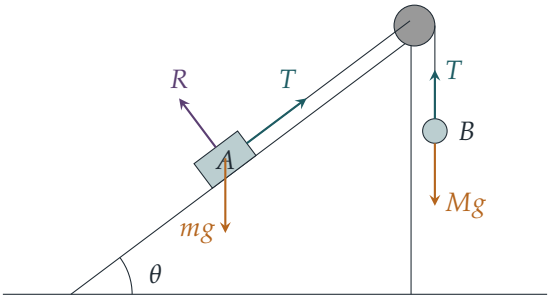
step.maths.org/questions/94-s1-q11

關鍵思路

先從幾何入手。繩子的長度恰好等於軌道的長度，所以當貨車位於底部時，小球正好懸在滑輪處。因此小球下落的距離就是斜面的高度 $d\sin\theta$，而貨車亦沿斜面向上移動相同的距離。繩子拉緊期間，整個系統以同一個恆定加速度運動。小球一着地，繩子便會鬆弛，貨車只是沿着光滑斜面向上滑行。所以第 (ii) 部分要問的是：貨車餘下的速率能否把它帶到滑輪處。

解答

我們把貨車和小球視為質點（particle），軌道視為光滑（smooth），繩子視為輕而不可伸長（light and inextensible），並跨過一個輕而無摩擦的滑輪（pulley）。因此兩邊的張力（tension）T 相同，而且在繩子拉緊期間，貨車和小球的速率相同。



第 (i) 部分

設 a 為小球向下的加速度，這也是貨車沿斜面向上的加速度。對小球（沿鉛垂方向）和對貨車（沿斜面方向）分別應用牛頓第二定律（Newton's second law），得

$$Mg - T = Ma,$$
$$T - mg\sin\theta = ma.$$

法向反作用力（normal reaction）R 與斜面垂直，並不出現在方程之中。把兩式相加便可消去 T：

$$a = \frac{(M - m\sin\theta)\,g}{M + m}.$$

系統由靜止開始，所以只有當 $a > 0$ 時小球才會移動。（若 $a \leq 0$，則甚麼也不會發生：當 $a = 0$ 時系統保持平衡；而繩子不能推物，所以 $a < 0$ 是不可能的。）因此，小球落到地面，當且僅當

$$M > m\sin\theta.$$

當貨車位於軌道底部時，整條長度為 d 的繩子都貼着斜面，所以小球起初與滑輪同高，離地面的高度為 $d\sin\theta$。它由靜止開始以恆定加速度 a 下落這段距離，所以由 $v^2 = u^2 + 2as$ 得

$$v^2 = 2\cdot\frac{(M - m\sin\theta)g}{M + m}\cdot d\sin\theta = \frac{2g(M - m\sin\theta)\,d\sin\theta}{M + m},$$

這正是所要求的結果。

第 (ii) 部分

小球着地時，貨車已沿軌道向上移動了 $d\sin\theta$，並正以速率 v 運動。此時它與滑輪之間還餘下距離

$$d - d\sin\theta = d(1 - \sin\theta)$$

。

想一想 小球着地之後，繩子會怎樣？

小球停了下來，但貨車仍然沿斜面向上、朝着滑輪運動。所以繩子變得鬆弛，$T = 0$。這時沿斜面方向唯一的力就是重力的分量，所以貨車以 $g\sin\theta$ 減速。它再走一段距離 s 後便會靜止，這段距離由 $0 = v^2 - 2g\sin\theta\, s$ 給出：

$$s = \frac{v^2}{2g\sin\theta} = \frac{(M - m\sin\theta)\,d}{M + m}.$$

（其後貨車會滑回斜面下方，滑輪附近不會再有任何事情發生。）

如果貨車在到達滑輪之前停下，即 $s < d(1-\sin\theta)$，它便不會與滑輪相撞。由於 $d > 0$ 及 $M + m > 0$，這個條件等價於

$$\begin{aligned} M - m\sin\theta &< (M+m)(1-\sin\theta) \\ &= M + m - M\sin\theta - m\sin\theta, \end{aligned}$$

化簡後得 $M\sin\theta < m$。

把這個結果與第 (i) 部分結合起來，小球到達地面而貨車又在滑輪之前停下，恰好是在以下情況：

$$m\sin\theta < M < \frac{m}{\sin\theta}.$$

若 $M\sin\theta = m$，貨車剛好以零速率到達滑輪。那只是輕輕碰到，不算相撞。

驗算

- 貨車沿軌道向上所能到達的最遠距離是 $d\sin\theta + s = \dfrac{Md(1+\sin\theta)}{M+m}$。這個距離小於 d，恰好是在 $M\sin\theta < m$ 的時候，與上面的條件吻合。當 $m \to 0$ 時，它趨向 $d(1+\sin\theta) > d$：一輛沒有重量的貨車必定會撞上滑輪，正如你所預料的一樣。
- 當 $\theta = 90°$ 時，條件 $M < m$ 與第 (i) 部分的 $M > m\sin\theta = m$ 互相矛盾。這是對的：在鉛垂的軌道上，貨車從滑輪正下方的底部出發，小球從頂部出發，而貨車上升的距離恰好等於小球下落的距離，也就是一直升到滑輪處。
- 以數值計算，取 $m = 2$、$M = 1.5$、$\theta = 30°$、$d = 3$ 及 $g = 9.8$，公式給出 $v^2 = 4.2$，而最遠點為 $\frac{27}{14} \approx 1.93 < 3$。以細小的時間步長逐步模擬這個運動，得到相同的數值。隨機抽取的例子亦證實，相撞恰好在 $M\sin\theta > m$ 時發生。

常見錯誤

- 把小球下落的距離當作 d 而不是 $d\sin\theta$。小球從滑輪處出發，而滑輪的高度是 $d\sin\theta$。
- 假設小球着地後繩子仍然拉緊，或者完全忘記了滑行階段，只把 $d\sin\theta$ 與 d 作比較。
- 繩子鬆弛之後，計算貨車的減速度時用了 g 而不是 $g\sin\theta$。

STEP 之外

這道題把運動分成約束條件不同的幾個階段：先是繩子拉緊，然後是繩子鬆弛。這是單側約束（*unilateral constraint*）最簡單的例子：這種約束只能拉，不能推。這類約束在力

學中隨處可見，從滑落桌面的鏈條，到機械人學中的接觸問題都有。這裏用能量方法可以得到一個快捷的另解：在滑行階段，貨車的動能（kinetic energy）$\frac{1}{2}mv^2$ 轉化為勢能（potential energy）$mgs\sin\theta$。Kleppner 與 Kolenkow 所著的 *An Introduction to Mechanics*（力學入門教科書）對這兩種觀點都有細緻的闡述。

詳細解答

STEP I 1994，第 13 題

概率與統計 • 指示變量

> **STEP I 1994，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/94-s1-q13

關鍵思路

把一個計數寫成若干指示變量（indicator variable，每個取值 1 或 0）之和，可以把難算的期望值（expectation）變成易算的，因為無論各部分是否獨立（independent），期望值都是線性的。這裏各次抽取肯定不獨立，但對稱性告訴我們每個 X_i 的期望值都相同。然後我們用第二種方法計算 $\mathrm{E}(X)$，即直接由 X 的分佈出發；令兩個答案相等，幾乎不費工夫就得到一條二項式係數（binomial coefficient）恆等式。

解答

全文設 $1 \le n \le M$，並約定當 $b > a$ 或 $b < 0$ 時 $\binom{a}{b} = 0$。

第 (i) 部分

若第 i 個取出的代幣是紅色，X_i 為 1，否則為 0。因此在和 $X_1 + X_2 + \cdots + X_n$ 之中，每個取出的紅色代幣恰好貢獻 1，每個非紅色代幣貢獻 0。所以這個和數出的正是取出的 n 個代幣中紅色代幣的數目，即 X。

第 (ii) 部分

由於 X_i 只取 0 和 1 兩個值，

$$\mathrm{E}(X_i) = 0 \cdot \mathrm{P}(X_i = 0) + 1 \cdot \mathrm{P}(X_i = 1) = \mathrm{P}(\text{第 } i \text{ 個取出的代幣是紅色}).$$

? **想一想** 第 i 個代幣是紅色的機會，難道不取決於之前抽到甚麼嗎？

就條件概率而言，是的；就無條件概率而言，不是。試想像把代幣編號，然後逐一全部抽出：M 個代幣的每一種排列次序都同樣可能出現。因此，位置 i 上的代幣同樣可能是 M 個代幣中的任何一個，正如位置 1 上的代幣一樣。抽了 n 次便停下來，並不會改變位置 $i \le n$ 上發生的事。由於 M 個代幣中有 m 個是紅色，

$$\mathrm{E}(X_i) = \mathrm{P}(X_i = 1) = \frac{m}{M} \qquad \text{對每個} i = 1, \ldots, n.$$

（對於 $i = 2$，你可以用較長的方法驗證：$\frac{m}{M} \cdot \frac{m-1}{M-1} + \frac{M-m}{M} \cdot \frac{m}{M-1} = \frac{m(M-1)}{M(M-1)} = \frac{m}{M}$。）

第 (iii) 部分

由第 (i) 部分及期望值的線性性質（無需獨立性），

$$\mathrm{E}(X) = \mathrm{E}(X_1) + \cdots + \mathrm{E}(X_n) = n \cdot \frac{m}{M} = \frac{mn}{M},$$

正是所求。

第 (iv) 部分

取出的 n 個代幣所成的集合，同樣可能是 $\binom{M}{n}$ 個大小為 n 的子集中的任何一個。要恰好有 k 個紅色代幣，我們從 m 個紅色代幣中選 k 個，再從其餘 $M - m$ 個中選 $n - k$ 個，所以

$$\mathrm{P}(X = k) = \frac{\binom{m}{k}\binom{M-m}{n-k}}{\binom{M}{n}}, \qquad k = 0, 1, \ldots, n.$$

按照關於二項式係數的約定，當 $k > m$（紅色代幣不足）或 $n - k > M - m$（其他代幣不足）時，上式正確地等於 0。這就是超幾何分佈（hypergeometric distribution）。

第 (v) 部分

直接由第 (iv) 部分計算 $\mathrm{E}(X)$；$k = 0$ 的一項為零，所以

$$\mathrm{E}(X) = \sum_{k=1}^{n} k\,\mathrm{P}(X = k) = \frac{1}{\binom{M}{n}} \sum_{k=1}^{n} k \binom{m}{k}\binom{M-m}{n-k}.$$

令此式與第 (iii) 部分的結果相等，

$$\sum_{k=1}^{n} k \binom{m}{k}\binom{M-m}{n-k} = \frac{mn}{M}\binom{M}{n}.$$

最後，

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

所以右邊等於 $m\binom{M-1}{n-1}$，這就是所求的恆等式。

驗算

- 取 $M = 5$，$m = 2$，$n = 3$。第 (iv) 部分給出 $\mathrm{P}(X = 0, 1, 2) = \frac{1}{10}, \frac{6}{10}, \frac{3}{10}$，三者之和為 1，而 $\mathrm{E}(X) = \frac{6+6}{10} = \frac{6}{5} = \frac{mn}{M}$。恆等式則為 $1 \cdot 2 \cdot 3 + 2 \cdot 1 \cdot 3 = 12 = 2\binom{4}{2}$。
- 若 $n = M$（全部取出），則必有 $X = m$，而確實 $\mathrm{E}(X) = mM/M = m$。
- 我們對所有 $M \leq 6$ 的袋子，列出代幣的每一種排列次序，驗證了 $\mathrm{E}(X_i)$、$\mathrm{E}(X)$ 和 $\mathrm{P}(X = k)$；並對每個 $M < 30$ 驗證了 (v) 中的恆等式。

常見錯誤

- 聲稱 $\mathrm{E}(X) = \sum \mathrm{E}(X_i)$ 需要獨立性。其實不需要：期望值的線性性質永遠成立。
- 把各次抽取當作獨立，在 (iv) 中寫出二項分佈（binomial distribution）$\mathrm{B}(n, \frac{m}{M})$。這樣得出的平均值是對的，但概率卻是錯的，因為抽取是不放回的。

STEP 之外

指示變量法是概率論中最有用的工具之一：它可以求出隨機置換（random permutation）的不動點（fixed point）數目的期望值、擲硬幣時連串（run）數目的期望值、隨機圖（random graph）中孤立頂點數目的期望值，而且還可藉 $\mathrm{Var}(\sum X_i) = \sum \mathrm{Var}(X_i) + \sum_{i \neq j} \mathrm{Cov}(X_i, X_j)$ 推廣至方差（variance）。以兩種方法計算同一個量來證明組合恆等式，這種手法同樣普遍適用。Blitzstein 與 Hwang 所著的 *Introduction to Probability*（概率論入門）全書貫徹使用指示變量，是觀摩這兩種想法實際運用的絕佳去處。

詳細解答

STEP I 1994，第 14 題

概率與統計 • 指數分佈的順序統計量

STEP I 1994，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/94-s1-q14

關鍵思路

每個 T_i 都服從指數分佈（exponential distribution），所以第 (i) 部分只是兩個標準的分部積分（integration by parts）。至於最後一篇和第一篇文章，竅門在於把關於最大值或最小值的陳述，轉化為關於每一個 T_i 的陳述：最後一篇文章在時刻 u 或之前交到，當且僅當所有文章都在那時或之前交到；第一篇文章在時刻 t 仍未交到，當且僅當所有文章在那時都仍未交到。獨立性（independence）隨即把「所有」變成乘積，再對分佈函數（distribution function）求導，便得到密度函數（density）。

解答

第 (i) 部分

以 X 表示任意一個 T_i，其密度函數為 $f(t) = \lambda \mathrm{e}^{-\lambda t}$（$t \geq 0$）。用分部積分，取 $u = t$ 及 $\mathrm{d}v = \lambda \mathrm{e}^{-\lambda t}\,\mathrm{d}t$（故 $v = -\mathrm{e}^{-\lambda t}$），

$$\mathrm{E}(X) = \int_0^\infty t\,\lambda \mathrm{e}^{-\lambda t}\,\mathrm{d}t = \Big[-t\mathrm{e}^{-\lambda t}\Big]_0^\infty + \int_0^\infty \mathrm{e}^{-\lambda t}\,\mathrm{d}t = 0 + \frac{1}{\lambda} = \frac{1}{\lambda},$$

因為當 $t \to \infty$ 時 $t\mathrm{e}^{-\lambda t} \to 0$（指數函數的增長勝過冪函數）。同樣地，

$$\mathrm{E}(X^2) = \int_0^\infty t^2 \lambda \mathrm{e}^{-\lambda t}\,\mathrm{d}t = \Big[-t^2\mathrm{e}^{-\lambda t}\Big]_0^\infty + \int_0^\infty 2t\,\mathrm{e}^{-\lambda t}\,\mathrm{d}t = \frac{2}{\lambda}\,\mathrm{E}(X) = \frac{2}{\lambda^2},$$

其中我們認出 $\int_0^\infty t\,\mathrm{e}^{-\lambda t}\,\mathrm{d}t$ 就是 $\frac{1}{\lambda}\mathrm{E}(X)$。因此

$$\mathrm{E}(T_i) = \frac{1}{\lambda}, \qquad \mathrm{Var}(T_i) = \frac{2}{\lambda^2} - \frac{1}{\lambda^2} = \frac{1}{\lambda^2}.$$

第 (ii) 部分

首先我們需要單獨一個 T_i 的分佈函數：對於 $u \geq 0$，

$$\mathrm{P}(T_i \leq u) = \int_0^u \lambda \mathrm{e}^{-\lambda t}\,\mathrm{d}t = 1 - \mathrm{e}^{-\lambda u}.$$

現在 $U = \max(T_1, \ldots, T_n)$，而各個 T_i 中最大的一個不超過 u，當且僅當每一個都不超過 u：

$$\{U \leq u\} = \{T_1 \leq u\} \cap \{T_2 \leq u\} \cap \cdots \cap \{T_n \leq u\}.$$

? **想一想** 為甚麼這裏可以把各個概率相乘？

事件 $\{T_i \leq u\}$ 涉及不同的 T_i，而按假設它們互相獨立，所以交集的概率等於各概率的乘積。因此，對於 $u \geq 0$，

$$\mathrm{P}(U \leq u) = \prod_{i=1}^{n} \mathrm{P}(T_i \leq u) = \left(1 - \mathrm{e}^{-\lambda u}\right)^n,$$

正是題目所求。U 的密度函數是這個分佈函數的導數；由鏈式法則（chain rule），

$$f_U(u) = n\left(1 - \mathrm{e}^{-\lambda u}\right)^{n-1} \cdot \lambda \mathrm{e}^{-\lambda u} = n\lambda\, \mathrm{e}^{-\lambda u}\left(1 - \mathrm{e}^{-\lambda u}\right)^{n-1} \qquad (u \geq 0),$$

而當 $u < 0$ 時 $f_U(u) = 0$。

第 (iii) 部分

現在 $T = \min(T_1, \ldots, T_n)$。這裏有用的是補事件（complement）：第一篇文章在時刻 t 之後才交到，當且僅當到那時沒有一篇交到，即每一個 $T_i > t$。利用 $\mathrm{P}(T_i > t) = \mathrm{e}^{-\lambda t}$，並再次運用獨立性，

$$\mathrm{P}(T > t) = \prod_{i=1}^{n} \mathrm{P}(T_i > t) = \left(\mathrm{e}^{-\lambda t}\right)^n = \mathrm{e}^{-n\lambda t} \qquad (t \geq 0).$$

所以 $\mathrm{P}(T \leq t) = 1 - \mathrm{e}^{-n\lambda t}$，求導得

$$f_T(t) = n\lambda\, \mathrm{e}^{-n\lambda t} \qquad (t \geq 0),$$

而當 $t < 0$ 時 $f_T(t) = 0$。

第 (iv) 部分

第 (iii) 部分的密度函數，正是第 (i) 部分的密度函數把 λ 換成 $n\lambda$：T 服從參數為 $n\lambda$ 的指數分佈。所以我們可以由第 (i) 部分直接寫出

$$\mathrm{E}(T) = \frac{1}{n\lambda}, \qquad \mathrm{Var}(T) = \frac{1}{n^2\lambda^2}.$$

這合乎情理：有 n 名學生時，第一篇文章交到的時間，應該比任何指定一篇快大約 n 倍。

驗算

- 當 $n = 1$ 時只有一篇文章，所以 $T = U = T_1$；兩個密度函數的確都簡化為 $\lambda e^{-\lambda t}$。
- $P(U \le u)$ 由 $u = 0$ 時的 0 上升至 $u \to \infty$ 時的 1，符合分佈函數應有的性質。
- 我們以 $n = 5$、$\lambda = 0.7$ 模擬了 2×10^5 個班別。例如，$U \le 2$ 所佔的比例為 0.2425，對比 $(1 - e^{-1.4})^5 \approx 0.2427$；$T$ 的平均值為 0.2860，對比 $1/(5 \times 0.7) \approx 0.2857$。用辛普森法則（Simpson's rule）亦驗證了 f_U 的積分為 1。

常見錯誤

- 處理最小值時直接計算 $P(T \le t)$：「至少有一個 $T_i \le t$」並不是交集，所以不能分解成乘積。應該經由 $P(T > t)$ 來計算。
- 忘記說明乘積之所以成立，是因為獨立性。
- 寫成 $\mathrm{Var}(T_i) = E(T_i^2) = 2/\lambda^2$，忘記減去 $(E\,T_i)^2$。

STEP 之外

T 和 U 分別是樣本中最小和最大的順序統計量（order statistics）。獨立指數隨機變量的最小值仍服從指數分佈，而其率參數是各率參數之和，這是泊松過程（Poisson process）和排隊論（queueing theory）的基礎。由無記憶性（memoryless property）還可得出一個漂亮的結果：相繼兩篇文章之間的時間間隔是互相獨立的指數隨機變量，率參數依次為 $n\lambda, (n-1)\lambda, \ldots, \lambda$，所以 $E(U) = \frac{1}{\lambda}\left(1 + \frac{1}{2} + \cdots + \frac{1}{n}\right)$。這兩個結果在 Grimmett 與 Welsh 的 *Probability: An Introduction*（概率入門），以及 Blitzstein 與 Hwang 的 *Introduction to Probability*（概率導論）中都有詳細討論。

詳細解答

STEP II 1994，第 1 題

純數學 • 整除性與十進制數字

> **STEP II 1994，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/94-s2-q1

關鍵思路

乘積的最後幾位數字，只取決於各因數的最後幾位數字，所以本題的一切運算都是模（modulo）10 的某個冪的算術。末位為 1、3、7 或 9 的數與 10 沒有公因數，因此它在模 10 之下是可逆（invertible）的，事實上在模 10 的任何冪之下也是可逆的。中間部分把 2 和 5 這些因數從 n 中抽出來，再把它們變成零。最後部分則利用模 10^{10} 之下的可逆性，迫使乘積的最後十位數字成為我們想要的任何一組數字。

解答

第 1 部分：末位為 1 的倍數

$n \times m$ 的末位數字，就是（n 的末位數字）×（m 的末位數字）的末位數字，因為其餘各位數字只會貢獻 10 的倍數。由於 n 的末位既不是 5 也不是偶數，所以它的末位是 1、3、7 或 9，而

$$1 \times 1 = 1, \qquad 3 \times 7 = 21, \qquad 7 \times 3 = 21, \qquad 9 \times 9 = 81.$$

因此分別取 $m = 1$、7、3 或 9 即可。這樣 $n \times m$ 的末位便是 1，正如所求。

第 2 部分：任意的 n

寫成

$$n = 2^a 5^b k,$$

其中 $a, b \geq 0$，而 k 既不能被 2 整除，也不能被 5 整除。於是 k 的末位既不是 5 也不是偶數，所以由第 1 部分可知，存在 m' 使得 km' 的末位是 1。現在設

$$m = 2^b 5^a m'.$$

那麼

$$n \times m = 2^{a+b} 5^{a+b} km' = 10^{a+b} (km').$$

乘以 10^{a+b} 只不過是在 km' 之後添上 $a + b$ 個零。所以，若 $a + b = 0$，則 $n \times m$ 的末位是 1；否則它的結尾是 1 後面跟着 $a + b$ 個零。這正是我們所需要的。

例如，$n = 40 = 2^3 \cdot 5$ 對應 $k = 1$，所以 $m = 2 \cdot 5^3 = 250$，而 $n \times m = 10\,000$。

第 3 部分：齊集十個數字

寫成 $n = 10^r q$，其中 $r \geq 0$，而 q 的末位是 1。在末尾添零不會令任何數字消失，所以只需找出 m，使 $q \times m$ 包含每一個數字；這樣 $n \times m = 10^r (qm)$ 也會包含它們。

設 $N = 10^{10}$ 及 $X = 1234567890$，後者包含每一個數字。我們的目標是

$$qm \equiv X \pmod{N},$$

也就是說，qm 的最後十位數字是 1234567890。

> **想一想** 為甚麼 q 必定有某個倍數，除以 10^{10} 時餘數（remainder）是 X？

考慮 N 個數 $q \times 0$, $q \times 1$, …, $q \times (N-1)$，以及它們模 N 的餘數。假設其中兩個餘數相等，譬如 $qs \equiv qt \pmod{N}$，其中 $0 \leq t < s \leq N - 1$。那麼 10^{10} 整除 $q(s-t)$。但 q 的末位是 1，所以它是奇數，也不是 5 的倍數，因而與 $10^{10} = 2^{10}5^{10}$ 沒有公共的質因數。由此可知 $2^{10}5^{10}$ 整除 $s - t$，但這是不可能的，因為 $0 < s - t < N$。所以這 N 個餘數各不相同。可能出現的餘數只有 N 個，即 $0, 1, \ldots, N-1$，所以每一個都會出現。特別地，存在某個滿足 $0 \leq m \leq N - 1$ 的 m，使得 $qm \equiv X \pmod{N}$。

這個 m 不是 0，因為 $X \neq 0$。所以 qm 是一個正的倍數，且 $qm \geq X \geq 10^9$。因此它的最後十位數字正好是 $1, 2, 3, 4, 5, 6, 7, 8, 9, 0$，而 X 開首的 1 也確實存在。於是 $q \times m$ 包含 $0, 1, \ldots, 9$ 所有數字，$n \times m$ 亦然。

舉例來說，當 $n = 21$ 時，我們可以取 $m = 7\,201\,646\,090$，這時 $21 \times m = 151\,234\,567\,890$。

> **想一想** 關於 n 的假設在哪裏用過？為甚麼需要它？

我們用它只是為了知道 q 沒有因數 2 或 5。沒有這一點，餘數 $qm \bmod 10^{10}$ 就不能取遍所有的值。例如，2 的每個倍數都是偶數，所以結尾不可能是 $\ldots 1$。正是這個假設讓我們可以指定乘積的結尾。（第 2 部分證明了任意的 n 都有一個這種形式的倍數，所以與第 3 部分結合起來，每一個 n 都有一個包含全部十個數字的倍數。）

驗算

- 用手驗算第 1 部分：$13 \times 7 = 91$ 及 $29 \times 9 = 261$。
- 電腦驗算證實，第 1 部分的方法對每一個末位既不是 5 也不是偶數的 $n \le 20\,000$ 都成立。它亦證實第 2 部分的構造對每一個 $n \le 20\,000$，以及 20 000 個不超過 10^{15} 的隨機 n 都成立。它還為超過 10 000 個符合所需形式的 n 值計算了第 3 部分的乘數，當中包括末尾有好幾個零的數，而在每一個情況下 $n \times m$ 都包含全部十個數字。例子：$21 \times 7\,201\,646\,090 = 151\,234\,567\,890$ 及 $11 \times 4\,657\,687\,990 = 51\,234\,567\,890$。

常見錯誤

- 在第 3 部分，忘記檢查 1234567890 這組數字是否真的完整出現。開首的 1 需要 $qm \ge 10^9$，而這是成立的，因為 $qm \equiv X$ 且 $qm > 0$。
- 在第 2 部分，乘以 10^{a+b} 而不是 $2^b 5^a$。這樣做雖然添了零，但 2 和 5 這些原本已在 n 之中的因數卻原封不動。

STEP 之外

第 3 部分的鴿巢原理（pigeonhole principle）論證，其實證明了一個一般的事實：若 $\gcd(q, N) = 1$，則 q 在模 N 之下有乘法逆元（multiplicative inverse）。可逆的剩餘（residue）構成一個羣（group），記作 $(\mathbb{Z}/N\mathbb{Z})^\times$，這是初等數論和 RSA 密碼學的起點。另一個技巧可以證明，最後部分對任何 n 都成立：只要 $10^k \ge n$，就有 n 的某個倍數介乎 $X \cdot 10^k$ 與 $X \cdot 10^k + 10^k$ 之間，而這樣的倍數是以 1234567890 開首的。Davenport 的 *The Higher Arithmetic* 是一本介紹同餘（congruence）與剩餘的優秀入門書。

詳細解答

STEP II 1994，第 2 題

純數學 • 勒讓德多項式

STEP II 1994，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/94-s2-q2

關鍵思路

每次求導，因式 $(x-a)^m$ 都會保留下來，只是次數減一，所以 m 重的重根（repeated root）同時也是首 $m-1$ 階導數（derivative）的根。多項式 $(x^2-1)^n=(x-1)^n(x+1)^n$ 在 $[-1,1]$ 的兩個端點都有 n 重根，因此對 $x^m\mathrm{P}_n(x)$ 作分部積分（integration by parts）時，每一個邊界項都會消失。每作一次分部積分，就有一階導數從 $(x^2-1)^n$ 轉移到 x^m 上，而 x^m 會比 $(x^2-1)^n$ 更早把導數用盡。

解答

重根與導數

由積法則（product rule），

$$\begin{aligned}\mathrm{P}'(x) &= m(x-a)^{m-1}\mathrm{Q}(x)+(x-a)^m\mathrm{Q}'(x)\\ &= (x-a)^{m-1}\big[m\mathrm{Q}(x)+(x-a)\mathrm{Q}'(x)\big]=(x-a)^{m-1}\mathrm{R}(x),\end{aligned}$$

其中 $\mathrm{R}(x)=m\mathrm{Q}(x)+(x-a)\mathrm{Q}'(x)$ 是多項式，因為多項式的和與積仍是多項式。這正是所要求的形式。

這個結果說明：若多項式含有因式 $(x-a)^k$（$k\geq 1$），則它的導數含有因式 $(x-a)^{k-1}$。由 $k=m$ 開始反覆應用這個結果，可得

$$\mathrm{P}^{(r)}(x)=(x-a)^{m-r}\mathrm{R}_r(x)\qquad(0\leq r\leq m)$$

其中 R_r 為某些多項式。每一步都是容許的，因為我們只在 $r\leq m-1$ 時才對 $\mathrm{P}^{(r)}$ 求導以得出 $\mathrm{P}^{(r+1)}$，而此時指數 $m-r$ 至少為 1。若 $1\leq r\leq m-1$，則 $m-r\geq 1$，所以

$(x-a)$ 仍是 $\mathrm{P}^{(r)}(x)$ 的因式，因此 $\mathrm{P}^{(r)}(a)=0$。（這對 $r=0$ 同樣成立，但對 $r=m$ 一般不成立。）

P_n 是 n 次多項式

設 $\mathrm{F}(x)=(x^2-1)^n$。由二項式定理（binomial theorem），$\mathrm{F}(x)$ 等於 x^{2n} 加上次數不超過 $2n-2$ 的各項。把 x^k 求導 n 次，結果或是 0（若 $k<n$），或是 x^{k-n} 的倍數，所以 $\mathrm{P}_n=\mathrm{F}^{(n)}$ 是次數不超過 n 的多項式。它的 x^n 項只來自 x^{2n}：

$$\frac{\mathrm{d}^n}{\mathrm{d}x^n}x^{2n}=2n(2n-1)\cdots(n+1)\,x^n=\frac{(2n)!}{n!}\,x^n.$$

這個係數不等於零，所以 P_n 的次數恰好是 n。

等於零的積分

由於 $\mathrm{F}(x)=(x-1)^n(x+1)^n$，在第一部分取 $a=1$、$m=n$ 及 $\mathrm{Q}(x)=(x+1)^n$，可知當 $1\le r\le n-1$ 時 $\mathrm{F}^{(r)}(1)=0$。取 $a=-1$ 作同樣的論證，可得 $\mathrm{F}^{(r)}(-1)=0$。另外，直接可見 $\mathrm{F}(\pm1)=0$。所以

$$\mathrm{F}^{(r)}(\pm1)=0\qquad\text{對於}0\le r\le n-1.\tag{$*$}$$

現在取 $j\ge1$ 及 $1\le k\le n$。作分部積分，對 x^j 求導、對 $\mathrm{F}^{(k)}$ 積分，可得

$$\begin{aligned}\int_{-1}^{1}x^j\mathrm{F}^{(k)}(x)\,\mathrm{d}x&=\Big[x^j\mathrm{F}^{(k-1)}(x)\Big]_{-1}^{1}-j\int_{-1}^{1}x^{j-1}\mathrm{F}^{(k-1)}(x)\,\mathrm{d}x\\&=-j\int_{-1}^{1}x^{j-1}\mathrm{F}^{(k-1)}(x)\,\mathrm{d}x,\end{aligned}$$

因為 $k-1\le n-1$，所以由 $(*)$ 可知邊界項等於零。

想一想 這個步驟可以用多少次，才會出問題？

每用一次，x 的冪次減一，導數的階數也減一。我們由 x^m 和 $\mathrm{F}^{(n)}$ 開始，所以可以應用 m 次。這個步驟需要 $k\ge1$，而我們用到的最小的 k 是 $n-m+1\ge1$。經過 m 步之後，

$$\int_{-1}^{1}x^m\mathrm{P}_n(x)\,\mathrm{d}x=(-1)^m m!\int_{-1}^{1}\mathrm{F}^{(n-m)}(x)\,\mathrm{d}x.$$

（當 $m=0$ 時，這就是原來的積分。）

若 $0\le m\le n-1$，則 $n-m\ge1$，所以最後的被積函數（integrand）本身是一個導數，於是

$$\int_{-1}^{1}x^m\mathrm{P}_n(x)\,\mathrm{d}x=(-1)^m m!\Big[\mathrm{F}^{(n-m-1)}(x)\Big]_{-1}^{1}=0,$$

這是由 $(*)$ 得出的，因為 $0\le n-m-1\le n-1$。這就是所要求的結果。

$m = n$ 的情況

取 $m = n$ 時，同樣的化簡過程仍然有效（最後一步用到 $k = 1$），並給出

$$\int_{-1}^{1} x^n \mathrm{P}_n(x)\,\mathrm{d}x = (-1)^n n! \int_{-1}^{1} (x^2-1)^n\,\mathrm{d}x = n! \int_{-1}^{1} (1-x^2)^n\,\mathrm{d}x,$$

因為 $(x^2-1)^n = (-1)^n(1-x^2)^n$。被積函數是偶函數(even function)，而代換 $x = \sin t$、$\mathrm{d}x = \cos t\,\mathrm{d}t$ 把 $[0,1]$ 變為 $[0, \frac{\pi}{2}]$，且 $1 - x^2 = \cos^2 t$。所以，利用提示中的公式，

$$\int_{-1}^{1} (1-x^2)^n\,\mathrm{d}x = 2\int_{0}^{\pi/2} \cos^{2n+1} t\,\mathrm{d}t = \frac{2^{2n+1}(n!)^2}{(2n+1)!}.$$

因此

$$\boxed{\int_{-1}^{1} x^n \mathrm{P}_n(x)\,\mathrm{d}x = \frac{2^{2n+1}(n!)^3}{(2n+1)!}.}$$

驗算

當 $n = 1$ 時，$\mathrm{P}_1(x) = 2x$，所以 $\int_{-1}^{1} x \cdot 2x\,\mathrm{d}x = \frac{4}{3}$，而公式給出 $\frac{8\cdot 1}{6} = \frac{4}{3}$。當 $n = 2$ 時，$\mathrm{P}_2(x) = 12x^2 - 4$。於是 $\int_{-1}^{1}(12x^2-4)\,\mathrm{d}x = 8 - 8 = 0$，且 $\int_{-1}^{1} x(12x^2-4)\,\mathrm{d}x = 0$（被積函數是奇函數），與預期相符。另外 $\int_{-1}^{1} x^2(12x^2-4)\,\mathrm{d}x = \frac{24}{5} - \frac{8}{3} = \frac{32}{15}$，與 $\frac{2^5\cdot 8}{120} = \frac{32}{15}$ 吻合。我們已用電腦對 $n \le 10$ 的所有情況作了精確驗證。

常見錯誤

- 忘記邊界項除了需要 $1 \le r \le n-1$ 時 $\mathrm{F}^{(r)}(\pm 1) = 0$ 之外，還需要 $r = 0$ 時也成立。$m = n - 1$ 的最後一步正是用到 $r = 0$。
- 弄錯正負號：$(x^2-1)^n = (-1)^n(1-x^2)^n$ 抵銷了分部積分所產生的 $(-1)^n$，所以答案是正數。這是合理的，因為 $x^n\mathrm{P}_n(x)$ 的首項係數是正數。

STEP 之外

除了相差一個常數倍之外，P_n 就是 n 次勒讓德多項式（Legendre polynomial），而它的這個公式稱為羅德里格斯公式（Rodrigues' formula）。第一個結果說明，就內積（inner product）$\int_{-1}^{1} fg\,\mathrm{d}x$ 而言，P_n 與每一個次數較低的多項式都正交（orthogonal）。勒讓德級數（Legendre series）和高斯求積法（Gaussian quadrature）背後所依靠的正是這個性質。在球坐標（spherical coordinates）下求解拉普拉斯方程（Laplace's equation）時，勒讓德多項式也會出現。Boyce 與 DiPrima 的 *Elementary Differential Equations*（微分方程入門教材）透過勒讓德微分方程（Legendre's differential equation）引入這些多項式。

詳細解答

STEP II 1994，第 3 題

純數學 • 函數方程

> **STEP II 1994，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/94-s2-q3

關鍵思路

關於 f，我們只知道一條對*所有* x 和 y 都成立的恆等式（identity），以及一個函數值 $\mathrm{f}(0) = 1$，除此之外一無所知。因此唯一的工具就是代入（substitution）：選取適當的 x 和 y，使大部分自變數變成 0 或 a，因為函數在這些點的值我們已經知道（或即將知道）。這條恆等式應該令你聯想到 $\cos(x-y) = \cos x \cos y + \sin x \sin y$，其中 $\mathrm{f}(a-x)$ 扮演 $\pm \sin x$ 的角色；這個類比既告訴你可以預期甚麼結果，也告訴你例子會是甚麼。

解答

以下一律把題目給出的恆等式寫成

$$\mathrm{f}(x-y) = \mathrm{f}(x)\,\mathrm{f}(y) - \mathrm{f}(a-x)\,\mathrm{f}(a+y) \qquad \text{對所有} x, y. \tag{$*$}$$

下面每一步都只是代入 $(*)$，所以我們沒有假設連續性（continuity）、可微性（differentiability）或任何其他性質。

證明 $\mathrm{f}(a) = 0$

在 $(*)$ 中取 $x = y = 0$：

$$\mathrm{f}(0) = \mathrm{f}(0)^2 - \mathrm{f}(a)\,\mathrm{f}(a), \qquad \text{所以} \qquad 1 = 1 - \mathrm{f}(a)^2.$$

因此 $\mathrm{f}(a)^2 = 0$，而 $\mathrm{f}(a)$ 是實數，故 $\mathrm{f}(a) = 0$。

第 (i) 部分

在 $(*)$ 中取 $x = 0$ 和 $y = t$：

$$\mathrm{f}(-t) = \mathrm{f}(0)\,\mathrm{f}(t) - \mathrm{f}(a)\,\mathrm{f}(a+t) = 1 \cdot \mathrm{f}(t) - 0 = \mathrm{f}(t),$$

這裏用到 $\mathrm{f}(0) = 1$ 和 $\mathrm{f}(a) = 0$。所以 f 是偶函數（even function）。

第 (ii) 部分

在 $(*)$ 中取 $x = y = a$：

$$\mathrm{f}(0) = \mathrm{f}(a)^2 - \mathrm{f}(0)\,\mathrm{f}(2a), \qquad \text{所以} \qquad 1 = 0 - \mathrm{f}(2a),$$

由此得 $\mathrm{f}(2a) = -1$。

第 (iii) 部分

想一想 哪一個代入可以令右邊第一個乘積消失，同時讓 t 保持自由？

取 $x = a$ 可以消去 $\mathrm{f}(x)\mathrm{f}(y)$ 這一項（因為 $\mathrm{f}(a) = 0$），並把 $\mathrm{f}(a-x)$ 變成 $\mathrm{f}(0) = 1$，而 y 仍然可以任意選取。當 $x = a$ 時，$(*)$ 變成

$$\mathrm{f}(a-y) = 0 \cdot \mathrm{f}(y) - \mathrm{f}(0)\,\mathrm{f}(a+y) = -\mathrm{f}(a+y) \qquad \text{對所有} y.$$

所以 f「關於 a 是奇的」。現在取 $y = a - t$：這時 $a - y = t$ 而 $a + y = 2a - t$，所以

$$\mathrm{f}(t) = -\mathrm{f}(2a - t),$$

即對所有 t 都有 $\mathrm{f}(2a - t) = -\mathrm{f}(t)$，這正是所要證明的結果。（取 $t = 0$ 便重新得到第 (ii) 部分的結果，這是一個有用的一致性檢查。）

第 (iv) 部分

把 (iii) 用兩次，中間用一次 (i)。首先，在 (iii) 中把 t 換成 $-t$，得

$$\mathrm{f}(2a + t) = -\mathrm{f}(-t) = -\mathrm{f}(t), \tag{1}$$

這裏用到 (i)。所以把自變數平移 $2a$ 會令函數值變號。再平移 $2a$ 一次，即在 (1) 中把 t 換成 $2a + t$，

$$\mathrm{f}(4a + t) = -\mathrm{f}(2a + t) = -\big(-\mathrm{f}(t)\big) = \mathrm{f}(t).$$

所以 f 是週期函數（periodic function），週期為 $4a$（當 $a \neq 0$ 時），正如餘弦函數的週期是 $4 \cdot \frac{\pi}{2} = 2\pi$ 一樣。

例子

當 $a = \pi/2$ 時。取 $\mathrm{f}(x) = \cos x$。這時 $\mathrm{f}(0) = 1$，而且由於 $\cos(\frac{\pi}{2} - x) = \sin x$ 和 $\cos(\frac{\pi}{2} + y) = -\sin y$，

$$\mathrm{f}(x)\mathrm{f}(y) - \mathrm{f}(\tfrac{\pi}{2} - x)\mathrm{f}(\tfrac{\pi}{2} + y) = \cos x \cos y + \sin x \sin y = \cos(x - y) = \mathrm{f}(x - y).$$

當 $a = -2$ 時。試用 $\mathrm{f}(x) = \cos kx$。由上面的結果可知，我們需要 $\mathrm{f}(a) = 0$，即 $\cos(2k) = 0$，並且週期為 $4|a| = 8$；所以取 $k = \frac{\pi}{4}$，即 $\mathrm{f}(x) = \cos\frac{\pi x}{4}$。驗證如下：$\mathrm{f}(0) = 1$，並且設 $u = \frac{\pi x}{4}$、$v = \frac{\pi y}{4}$，則

$$\mathrm{f}(-2-x) = \cos\big(-\tfrac{\pi}{2} - u\big) = \cos\big(\tfrac{\pi}{2} + u\big) = -\sin u, \qquad \mathrm{f}(-2+y) = \cos\big(v - \tfrac{\pi}{2}\big) = \sin v,$$

所以

$$\begin{aligned}\mathrm{f}(x)\mathrm{f}(y) - \mathrm{f}(-2-x)\mathrm{f}(-2+y) &= \cos u \cos v + \sin u \sin v \\ &= \cos(u - v) = \cos\tfrac{\pi(x-y)}{4} = \mathrm{f}(x - y).\end{aligned}$$

這兩個函數顯然都不是常數函數。（更一般地說，只要 $\cos ka = 0$，$\cos kx$ 對該 a 值便適用，所以當 $a = -2$ 時 $\cos\frac{3\pi x}{4}$ 同樣可以。）

驗算

在第 (iii) 部分取 $t = 0$，得 $\mathrm{f}(2a) = -\mathrm{f}(0) = -1$，與 (ii) 吻合；在 (iii) 取 $t = a$，得 $\mathrm{f}(a) = -\mathrm{f}(a)$，與 $\mathrm{f}(a) = 0$ 吻合。再以數值檢驗：對 $\mathrm{f}(x) = \cos\frac{\pi x}{4}$ 和 $a = -2$，取 $x = 1$、$y = 3$，左邊是 $\mathrm{f}(-2) = 0$，右邊是 $\mathrm{f}(1)\mathrm{f}(3) - \mathrm{f}(-3)\mathrm{f}(1) = \cos\frac{\pi}{4}\cos\frac{3\pi}{4} - \cos\frac{3\pi}{4}\cos\frac{\pi}{4} = 0$。用電腦對兩個例子各以 20 000 組隨機數對 (x, y) 測試，結果恆等式在 10^{-13} 的誤差範圍內成立。

常見錯誤

- 弄錯平移半個週期時的正負號：由 (iii) 和 (i) 得 $\mathrm{f}(t + 2a) = -\mathrm{f}(t)$，所以週期是 $4a$，而不是 $2a$。
- 在證明 (i)–(iv) 時假設 f 可微（或假設它就是餘弦函數）。題目不容許這樣做；只可以使用代入。
- 對於 $a = -2$，只寫出 $\cos\frac{\pi x}{4}$ 而沒有驗證它滿足恆等式：符合 $\mathrm{f}(a) = 0$ 和週期的要求只是很好的指引，並不是證明。

STEP 之外

由函數所滿足的恆等式來確定該函數，正是*函數方程*（functional equations）這門學問的課題。在本題中，單憑一條恆等式，便足以推出偶函數性質、「半週期」變號以及週期性。在分析學（analysis）課程中會遇到一些經典結果：柯西（Cauchy）方程 $\mathrm{f}(x+y)=\mathrm{f}(x)+\mathrm{f}(y)$ 在加上連續性的條件後，必然得出 $\mathrm{f}(x)=cx$；若沒有這類條件，便會出現性質極為古怪的解。同樣地，餘弦的減法公式在連續函數之中刻劃了 $\cos kx$。至於連續性如何在這類論證中發揮作用，可參閱 Abbott 的 *Understanding Analysis*（一本講解嚴謹的分析學入門書）。

詳細解答

STEP II 1994，第 4 題

純數學 • 橢圓的面積

STEP II 1994，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/94-s2-q4

關鍵思路

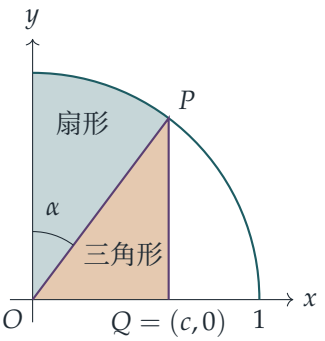
第一個積分是四分之一圓（quarter circle）下方的面積，而這塊面積可以乾淨俐落地分成一個直角三角形和一個扇形（sector），所以毋須作任何代換（substitution）。橢圓（ellipse）是把圓沿兩個方向以不同倍數拉伸而成的，而拉伸會把面積乘以一個常數，於是第二個結果立即得出。至於最後一部分，對稱性（symmetry）幾乎包辦了所有工作：兩個橢圓互為對直線 $y = x$ 的反射（reflection），所以我們只需求交集（intersection）的八分之一，而這一小塊正是第二個結果減去一個三角形。

解答

$(1 - x^2)^{1/2}$ 的積分

所討論的區域由四分之一圓 $y = (1 - x^2)^{1/2}$、兩條坐標軸和直線 $x = c$ 圍成。（這條曲線須理解為邊界：區域是滿足 $0 \le y \le (1 - x^2)^{1/2}$ 和 $0 \le x \le c$ 的點所組成的集合。）它的面積正是我們要求的積分。

設 $P = \left(c, (1 - c^2)^{1/2}\right)$ 為圓上位於 $x = c$ 上方的點，$Q = (c, 0)$，並設 $\alpha = \sin^{-1} c$，於是 $0 < \alpha \le \frac{\pi}{2}$。連接 O 與 P，便把區域分成兩塊。



- 三角形 OQP 在 Q 處是直角，底為 c，高為 $(1-c^2)^{1/2}$，所以它的面積是 $\frac{1}{2}c(1-c^2)^{1/2}$。
- 餘下的部分是單位圓在 OP 與 y 軸之間的扇形。OP 與 y 軸的夾角是 α，因為 P 與 y 軸的水平距離是 $c = 1\cdot\sin\alpha$。半徑為 1、角為 α 的扇形，面積是 $\frac{1}{2}\alpha = \frac{1}{2}\sin^{-1}c$。

把兩塊相加，

$$\int_0^c (1-x^2)^{\frac{1}{2}}\,\mathrm{d}x = \tfrac{1}{2}\left[c(1-c^2)^{\frac{1}{2}} + \sin^{-1}c\right], \qquad 0 < c \le 1,$$

正是所求。（當 $c = 1$ 時，三角形退化成一條線段，我們得到四分之一圓的面積 $\frac{\pi}{4}$。）

橢圓下方的面積

同樣地，區域位於曲線上半部分 $y = b\left(1-x^2/a^2\right)^{1/2}$ 之下，其中 $0 \le x \le c$，所以它的面積是

$$\int_0^c b\left(1-\frac{x^2}{a^2}\right)^{\frac{1}{2}}\mathrm{d}x.$$

代入 $x = au$，則 $\mathrm{d}x = a\,\mathrm{d}u$，積分上下限變成 0 和 c/a（由於 $0 < c \le a$，故 $0 < c/a \le 1$），上式變成

$$ab\int_0^{c/a}(1-u^2)^{\frac{1}{2}}\,\mathrm{d}u = \frac{ab}{2}\left[\frac{c}{a}\left(1-\frac{c^2}{a^2}\right)^{\frac{1}{2}} + \sin^{-1}\left(\frac{c}{a}\right)\right],$$

這裏用了第一個結果，並把 c 換成 c/a。這就是所求的面積。

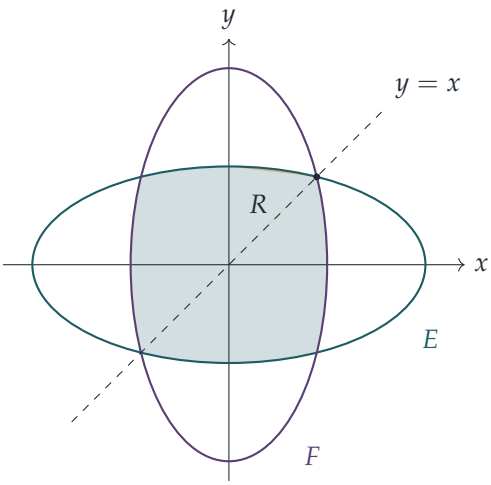
從幾何角度看：橢圓是把單位圓沿水平方向拉伸 a 倍、沿垂直方向拉伸 b 倍而成的，而這個拉伸把每一塊面積都乘以 ab。

兩個橢圓的交集

E 和 F 在 $x \mapsto -x$ 和 $y \mapsto -y$ 之下都保持不變，所以 $E \cap F$ 也是如此，它的面積等於它在象限（quadrant）$x, y \ge 0$ 內那部分面積的四倍。此外，把 x 和 y 對調會把 E 變成 F、

把 F 變成 E，所以 $E \cap F$ 對直線 $y = x$ 對稱。因此，第一象限內的那四分之一由互為鏡像的兩半組成，而

$$\text{面積}(E \cap F) = 8 \times \text{區域面積} R, \qquad R = \{(x, y) \in E \cap F : \ 0 \le x \le y\}.$$



想一想 在 R 之中，兩個橢圓裏究竟是哪一個構成邊界？

對任意一點，

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

當 $0 \le x \le y$ 時，$x^2 - y^2 \le 0$；又因為 $0 < b \le a$，所以 $\frac{1}{b^2} - \frac{1}{a^2} \ge 0$。因此這個差 ≤ 0：在 $y = x$ 的那一側，E 內的任何一點都自動在 F 內。於是

$$R = \{(x, y) \in E : \ 0 \le x \le y\},$$

只有 E 起作用。（這裏用到了 $b \le a$；若 $b > a$，兩個橢圓的角色便會互換。）

E 的邊界在第一象限內與 $y = x$ 相交之處滿足 $x^2\left(\frac{1}{a^2} + \frac{1}{b^2}\right) = 1$，即交於 $x = y = x_0$，其中

$$x_0 = \frac{ab}{\sqrt{a^2 + b^2}}.$$

當 $0 \le x \le x_0$ 時，E 的上半部分位於 $y = x$ 之上或恰在其上（它從高度 $b > 0$ 出發，只在 x_0 處與該直線相交），而當 $x > x_0$ 時則位於該直線之下。所以 R 是 $0 \le x \le x_0$ 時介乎 $y = x$ 與曲線 $y = b(1 - x^2/a^2)^{1/2}$ 之間的區域：

$$\text{面積}(R) = \underbrace{\int_0^{x_0} b\left(1 - \frac{x^2}{a^2}\right)^{\frac{1}{2}} \mathrm{d}x}_{\text{上一個結果，取} c = x_0} - \underbrace{\tfrac{1}{2}x_0^2}_{\text{三角形，位於此線之下：} y = x}.$$

取 $c = x_0$（並且符合所需的條件 $0 < x_0 < a$），

$$\frac{c}{a} = \frac{b}{\sqrt{a^2+b^2}}, \qquad \left(1-\frac{c^2}{a^2}\right)^{\frac{1}{2}} = \frac{a}{\sqrt{a^2+b^2}}, \qquad \frac{c}{a}\left(1-\frac{c^2}{a^2}\right)^{\frac{1}{2}} = \frac{ab}{a^2+b^2}.$$

所以

$$\begin{aligned} \text{面積}(E\cap F) &= 8\,\text{面積}(R) \\ &= 4ab\left[\frac{ab}{a^2+b^2} + \sin^{-1}\left(\frac{b}{\sqrt{a^2+b^2}}\right)\right] - \frac{4a^2b^2}{a^2+b^2} \\ &= 4ab\sin^{-1}\left(\frac{b}{\sqrt{a^2+b^2}}\right), \end{aligned}$$

因為兩個代數項恰好互相抵消。這就是所求的面積。

? 想一想 為甚麼代數項會互相抵消？

$\frac{ab}{2}\cdot\frac{c}{a}(1-c^2/a^2)^{1/2} = \frac{1}{2}x_0y_0$ 這一項，是圓的圖中那個三角形經拉伸後得到的三角形 O、$(x_0,0)$、(x_0,y_0)。這裏 $y_0 = x_0$，所以它恰好就是我們減去的、位於 $y = x$ 之下的那個三角形。剩下的是一個經拉伸的扇形，這就是答案只含一個反正弦（inverse sine）的原因。

驗算

- 若 $a = b$，兩個橢圓是同一個圓，公式給出 $4a^2\sin^{-1}\frac{1}{\sqrt{2}} = 4a^2\cdot\frac{\pi}{4} = \pi a^2$，與應有的結果相符。
- 當 a 固定而 $b \to 0$ 時，面積的表現近似於 $4ab\cdot\frac{b}{a} = 4b^2$：交集幾乎就是邊長為 $2b$ 的正方形，即 E 那條窄長的水平帶與 F 那條窄長的垂直帶交叉之處。
- 當 $a = 2$、$b = 1$ 時，公式給出 $8\sin^{-1}\frac{1}{\sqrt{5}} \approx 3.70918$。對兩條邊界中較低的一條使用辛普森法則（Simpson's rule），得 3.70918；以 2×10^5 個點作蒙地卡羅（Monte Carlo）估算，得 3.707。頭兩個結果亦已就多組隨機的 a、b、c 以數值求積（quadrature）核對過。

常見錯誤

- 弄錯扇形的角：扇形位於 OP 與 y 軸之間，角是 $\sin^{-1}c$，而不是 $\cos^{-1}c$。
- 沒有檢查是哪一個橢圓圍住該區域。在 $0 \le x \le y$ 這一半之中，邊界是 E，全因 $b \le a$；若選了另一個橢圓，答案中反正弦裏面的 b 便會變成 a。

STEP 之外

線性映射（linear map）$(x, y) \mapsto (ax, by)$ 把每一塊面積都乘以 ab；一般而言，線性映射會把面積乘以其行列式（determinant）的絕對值。這正是重積分（multiple integrals）變數代換公式（change-of-variables formula）中雅可比（Jacobian）因子的二維情形，你會在大學一年級的微積分課程中遇到。Strang 的 *Introduction to Linear Algebra*（線性代數入門）把行列式解釋為面積（以及體積）的縮放因子。

詳細解答

STEP II 1994，第 5 題

純數學 • 實根的計數

> **STEP II 1994，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/94-s2-q5

關鍵思路

要數 $f(x) = c$ 有多少個根（root），即是問水平線 $y = c$ 與 f 的圖像相交多少次，所以關鍵在於 f 的形狀：它在哪裏取最小值，在哪裏遞增、在哪裏遞減。本題的每個式子都由兩個關於某個中心點對稱的括號組成。把變量平移，使中心點變成原點，則奇次項互相抵消（四次方的情況）或偶次項互相抵消（三次方的情況），圖像的形狀便一目了然。

解答

第 (i) 部分

用二項式定理（binomial theorem）展開。在 $(x+1)^4$ 和 $(x-1)^4$ 中，x 的奇次項符號相反，所以相加時互相抵消：

$$f(x) = (x-1)^4 + (x+1)^4 = 2\big(x^4 + 6x^2 + 1\big) = 2x^4 + 12x^2 + 2.$$

因此 f 是偶函數（even function），而且 $f(x) = 2 + 2x^4 + 12x^2 \geq 2$，等號只在 $x = 0$ 時成立。對於 $x \geq 0$，當 $x > 0$ 時 $f'(x) = 8x^3 + 24x > 0$，所以 f 在 $[0, \infty)$ 上嚴格遞增（strictly increasing），而且當 $x \to \infty$ 時 $f(x) \to \infty$。由於 f 是連續（continuous）的，它在 $[0, \infty)$ 上取 $[2, \infty)$ 中的每個值恰好一次；根據對稱性，它在 $(-\infty, 0]$ 上亦然。

- 若 $c > 2$：恰有一個根 $x = a > 0$，於是 $x = -a$ 是第二個根；這兩個根互不相同，此外再無其他根。所以**恰好有兩個**實根。
- 若 $c = 2$：由 $f(x) = 2$ 必得 $2x^4 + 12x^2 = 0$，所以 $x = 0$ 是**唯**一的根。
- 若 $c < 2$：對所有 x 都有 $f(x) \geq 2 > c$，所以**沒有**根。

這正是題目所要求的結論。

（另一做法：$2x^4 + 12x^2 + 2 - c = 0$ 是關於 $u = x^2$ 的二次方程，其根為 $u = -3 \pm \sqrt{8 + c/2}$；只有取 $+$ 號的根才可能非負，而它是正數、零還是負數，分別對應 $c > 2$、$c = 2$ 和 $c < 2$。）

第 (ii) 部分

括號 $x - 3$ 和 $x - 1$ 以 $x = 2$ 為中心，所以設 $y = x - 2$。於是

$$(x-3)^4 + (x-1)^4 = (y-1)^4 + (y+1)^4,$$

這正是第 (i) 部分的式子，只是變量換成了 y。由於 $x \mapsto y = x - 2$ 是實數之間的一一對應（one-to-one correspondence），這個方程的實根數目與第 (i) 部分相同：

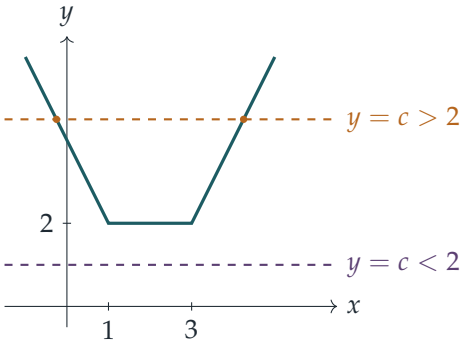
$$\text{有兩個根，若} c > 2, \qquad \text{有一個根} (x = 2)\text{，若} c = 2, \qquad \text{沒有根，若} c < 2.$$

第 (iii) 部分

設 $g(x) = |x - 3| + |x - 1|$，並在 $x = 1$ 和 $x = 3$ 所分出的三個區間上分別去掉絕對值（modulus）符號：

$$g(x) = \begin{cases} (3-x) + (1-x) = 4 - 2x, & x < 1, \\ (3-x) + (x-1) = 2, & 1 \le x \le 3, \\ (x-3) + (x-1) = 2x - 4, & x > 3. \end{cases}$$

所以圖像是一個平底的 V 形：在 1 與 3 之間恆等於 2，在左邊嚴格遞減，在右邊嚴格遞增，而且在 $[1, 3]$ 之外 $g > 2$。



- 若 $c < 2$：處處都有 $g(x) \ge 2 > c$，所以**沒有**根。
- 若 $c = 2$：凡滿足 $1 \le x \le 3$ 的 x 都是根，此範圍以外則沒有根，所以有**無窮多個**根。

- 若 $c > 2$：在 $[1,3]$ 內沒有根；在 $x < 1$ 上，由 $4 - 2x = c$ 得 $x = 2 - \frac{c}{2}$，由於 $c > 2$，它確實小於 1；在 $x > 3$ 上，由 $2x - 4 = c$ 得 $x = 2 + \frac{c}{2} > 3$。所以**恰好有兩個**根，即 $x = 2 \pm \frac{c}{2}$。

> **想一想** 為甚麼 $c = 2$ 時的答案與第 (ii) 部分相差這麼遠？

對於四次方，由於兩項都是嚴格凸（strictly convex）的，它們的和在中心點有唯一而尖銳的最小值。對於絕對值，當 x 在 1 與 3 之間移動時，其中一個距離增加多少，另一個距離便恰好減少多少，所以它們的和在整個區間上都保持為 2（即 1 與 3 之間的距離）。

第 (iv) 部分

同樣設 $y = x - 2$。對於三次方，互相抵消的是 y 的偶次項：

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

設 $h(y) = 2y^3 + 6y$。則對任意實數 y 都有

$$h'(y) = 6y^2 + 6 \geq 6 > 0$$

，所以 h 在整條實數線上嚴格遞增，因而每個值最多取一次。此外，當 $y \to \infty$ 時 $h(y) \to \infty$，當 $y \to -\infty$ 時 $h(y) \to -\infty$，而 h 是連續的，所以根據介值定理（intermediate value theorem），它取每個實數值至少一次。

因此，對**每一個**實數 c，方程都**恰好有一個**實根。

驗算

- 在中心點 $x = 2$，兩個四次方的式子都等於 $1 + 1 = 2$，與 $c = 2$ 的情況吻合，而且 $(2-3)^3 + (2-1)^3 = 0$。
- 在第 (iv) 部分，$c = 8$ 應給出 $x = 3$：的確 $0^3 + 2^3 = 8$，而取 $y = 1$ 時 $h(1) = 2 + 6 = 8$。
- 在第 (iii) 部分取 $c = 4$，根 $x = 0$ 和 $x = 4$ 分別給出 $3 + 1 = 4$ 和 $1 + 3 = 4$。

我們以數值方法確認了各展開式在隨機選取的點上成立，並對一系列 c 值（例如在 (i) 和 (ii) 中取 $c = 1.9$、2.5、100，在 (iv) 中取 c 由 -1000 至 5000），在細密的網格上數算每個函數減去 c 後的變號次數：所得次數分別是 0、2、2，以及恆為 1，與上述結論一致。

常見錯誤

- 在第 (iii) 部分 $c = 2$ 時答「一個根」或「兩個根」：整個區間 $1 \leq x \leq 3$ 上的點全都是根。
- 在第 (iv) 部分，只說三次方程「有一個或三個實根」，卻沒有判斷是哪一種情況。$h'(y) > 0$ 處處成立這一事實便足以定奪。
- 忘記檢查 $x = 2 - \frac{c}{2}$ 是否真的落在使用公式 $4 - 2x$ 的區域 $x < 1$ 之內。

STEP 之外

這裏的論證建基於兩個定理，大學第一門分析（analysis）課程會仔細證明它們：其一是介值定理（區間上的連續函數會取遍兩個函數值之間的每一個值），其二是導數為正的函數必定嚴格遞增，這可由中值定理（mean value theorem）推出。第 (iii) 部分亦讓你初嘗凸性（convexity）和 ℓ^1 距離的滋味：到兩點的距離之和，在連接這兩點的整條線段上都取得最小值。Abbott 的 *Understanding Analysis* 一書嚴謹地闡述了這些概念。

詳細解答

STEP II 1994，第 6 題

純數學 • 三角函數裂項相消級數

> **STEP II 1994，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/94-s2-q6

關鍵思路

一切都建基於一條小小的恆等式（identity）：$\tan x = \cot x - 2\cot 2x$。它把和式中的每一項化為形如 $\frac{1}{2^r}\cot\frac{\theta}{2^r}$ 的相鄰兩項之差，於是和式可以裂項相消（telescope），而歸納步驟之所以行得通，靠的也正是這條恆等式。至於無窮和，只需用到極限（limit）$\frac{\sin\varphi}{\varphi}\to 1$。

解答

一條有用的恆等式

對任何使 $\sin x$、$\cos x$ 和 $\sin 2x$ 皆不為零的 x，

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

所以

$$\tan x = \cot x - 2\cot 2x. \qquad (*)$$

我們只會在 $x = \theta/2^r$（其中 $r \geq 1$）的情況下使用 $(*)$。由於 $0 < \theta < \pi$，這樣的 x 必在 $(0, \frac{\pi}{2})$ 之內，因此 $\sin x$、$\cos x$ 和 $\sin 2x$ 全為正數，我們寫下的每一個 tan 和 cot 都有定義。

以歸納法求有限和

以 S_n 表示左方 $\sum_{r=1}^{n}\frac{1}{2^r}\tan\frac{\theta}{2^r}$，並設 $P(n)$ 為命題

$$S_n = \frac{1}{2^n}\cot\frac{\theta}{2^n} - \cot\theta.$$

基礎情況 $n = 1$。在 $(*)$ 中代入 $x = \theta/2$，再除以二：

$$S_1 = \tfrac{1}{2}\tan\tfrac{\theta}{2} = \tfrac{1}{2}\cot\tfrac{\theta}{2} - \cot\theta,$$

這就是 $P(1)$。

歸納步驟。假設 $P(k)$ 對某個 $k \geq 1$ 成立。在 $(*)$ 中代入 $x = \theta/2^{k+1}$，於是 $2x = \theta/2^k$，再除以 2^{k+1}：

$$\frac{1}{2^{k+1}}\tan\frac{\theta}{2^{k+1}} = \frac{1}{2^{k+1}}\cot\frac{\theta}{2^{k+1}} - \frac{1}{2^k}\cot\frac{\theta}{2^k}.$$

把此式與 $P(k)$ 相加，

$$\begin{aligned} S_{k+1} &= S_k + \frac{1}{2^{k+1}}\tan\frac{\theta}{2^{k+1}} \\ &= \left(\frac{1}{2^k}\cot\frac{\theta}{2^k} - \cot\theta\right) + \left(\frac{1}{2^{k+1}}\cot\frac{\theta}{2^{k+1}} - \frac{1}{2^k}\cot\frac{\theta}{2^k}\right) \\ &= \frac{1}{2^{k+1}}\cot\frac{\theta}{2^{k+1}} - \cot\theta, \end{aligned}$$

這就是 $P(k+1)$。根據數學歸納法（induction），$P(n)$ 對所有 $n \geq 1$ 和所有 $\theta \in (0, \pi)$ 成立，命題得證。

想一想 假如你從未見過恆等式 $(*)$，怎樣才能把它找出來？

從歸納法倒推回去。由 $n = k$ 到 $n = k+1$，右方的變化是 $\frac{1}{2^{k+1}}\cot\frac{\theta}{2^{k+1}} - \frac{1}{2^k}\cot\frac{\theta}{2^k}$，左方的變化是 $\frac{1}{2^{k+1}}\tan\frac{\theta}{2^{k+1}}$。乘以 2^{k+1} 並記 $x = \theta/2^{k+1}$，可見兩邊的變化相等，當且僅當 $\tan x = \cot x - 2\cot 2x$。換言之，題目本身已告訴你要證明哪一條恆等式；之後只需用倍角公式（double-angle formulae）驗證便可。

無窮和

各項全為正數，而且我們已有部分和（partial sums）的精確公式，所以只需求出 S_n 在 $n \to \infty$ 時的極限。設 $\varphi_n = \theta/2^n$，則 $\varphi_n \to 0^+$ 且 $\frac{1}{2^n} = \frac{\varphi_n}{\theta}$。於是

$$\frac{1}{2^n}\cot\frac{\theta}{2^n} = \frac{\varphi_n}{\theta}\cdot\frac{\cos\varphi_n}{\sin\varphi_n} = \frac{1}{\theta}\cdot\frac{\varphi_n}{\sin\varphi_n}\cdot\cos\varphi_n \longrightarrow \frac{1}{\theta}\cdot 1\cdot 1 = \frac{1}{\theta},$$

這裏用到了標準極限：當 $\varphi \to 0$ 時，$\frac{\sin\varphi}{\varphi} \to 1$ 及 $\cos\varphi \to 1$。由於 $\cot\theta$ 與 n 無關，

$$\sum_{r=1}^{\infty}\frac{1}{2^r}\tan\frac{\theta}{2^r} = \lim_{n\to\infty} S_n = \frac{1}{\theta} - \cot\theta,$$

這正是所求的結果。特別地，此級數（series）對所有 $\theta \in (0, \pi)$ 均收斂（converge）。

想一想 為甚麼可以說「無窮和等於 S_n 的極限」？

這正是無窮級數之和的定義：$\sum_{r=1}^{\infty} a_r$ 的意思就是 $\lim_{n\to\infty}\sum_{r=1}^{n} a_r$，前提是該極限存在。我們已把極限算了出來，從而證明它存在，所以毋須再做其他工夫。

驗算

- $\theta = \frac{\pi}{2}$：右方為 $\frac{2}{\pi} - 0 \approx 0.63662$。以數值方法把級數的首 60 項相加，同樣得到 0.6366198。
- $\theta = 1$：級數與 $1 - \cot 1$ 均得 0.3579074。
- θ 很小時：$\tan\frac{\theta}{2^r} \approx \frac{\theta}{2^r}$，所以級數約為 $\theta\sum 4^{-r} = \frac{\theta}{3}$。這與展開式 $\frac{1}{\theta} - \cot\theta = \frac{\theta}{3} + O(\theta^3)$ 吻合。
- 當 $\theta \to \pi^-$ 時，$\cot\theta \to -\infty$，而首項 $\frac{1}{2}\tan\frac{\theta}{2} \to +\infty$，所以兩邊同時趨向無窮大，正如所料。

常見錯誤

- 把 $(*)$ 中的因子 2 弄錯，例如寫成 $\tan x = \cot x - \cot 2x$。可用 $x = \frac{\pi}{4}$ 檢驗：$1 = 1 - 2\cdot 0$。
- 因為 $\frac{1}{2^n} \to 0$ 便聲稱 $\frac{1}{2^n}\cot\frac{\theta}{2^n} \to 0$。餘切的增長速度與 $2^n/\theta$ 相若，所以乘積趨向 $\frac{1}{\theta}$，而不是 0。
- 忘記說明為何各個正切和餘切都有定義：正是條件 $0 < \theta < \pi$ 確保每一個角 $\theta/2^r$ 都在 $(0, \frac{\pi}{2})$ 之內。

STEP 之外

函數 $\frac{1}{\theta} - \cot\theta$ 屬於一個著名的家族：對部分分式展開（partial-fraction expansion）$\cot\theta = \frac{1}{\theta} + \sum_{k\ge 1}\frac{2\theta}{\theta^2 - k^2\pi^2}$ 求導，可得巴塞爾和（Basel sum）$\sum 1/k^2 = \pi^2/6$，而歐拉（Euler）正是從 $\sin\theta$ 的無窮乘積（infinite product）得出這類結果。對韋達（Viète）乘積 $\frac{\sin\theta}{\theta} = \prod_{r\ge 1}\cos\frac{\theta}{2^r}$ 取對數再求導，恰好得到本題的級數，這是理解此結果的另一途徑。如何令這類對無窮和與無窮乘積的運算變得嚴謹，是大學一年級分析（analysis）課程的課題；可參閱 Abbott, *Understanding Analysis*（一本數學分析入門教科書）。

詳細解答

STEP II 1994，第 7 題

純數學 • 阿波羅尼奧斯圓

> **STEP II 1994，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/94-s2-q7

關鍵思路

從距離 r 處觀看高度為 h 的旗杆頂，仰角（angle of elevation）的正切是 h/r，所以「仰角相等」其實只是說：士兵與兩個杆腳的距離保持固定比 $h_1 : h_2$。把這個條件兩邊平方，得到的方程正好是第一部分的形式，因此軌跡（locus）是一個圓（即阿波羅尼奧斯圓（circle of Apollonius））。至於最後一部分，把旗杆放在一般位置，並把每個圓心寫成兩個杆腳的加權組合；恆等式便自然得出，而且其係數之和為零，正是這一點保證了三點共線（collinear）。

解答

圓

由於 $a > 0$，我們可以除以 a，然後配方（complete the square）：

$$x^2 + y^2 + \frac{2g}{a}x + \frac{2f}{a}y + \frac{c}{a} = 0 \iff \left(x + \frac{g}{a}\right)^2 + \left(y + \frac{f}{a}\right)^2 = \frac{g^2 + f^2 - ac}{a^2}.$$

因為 $f^2 + g^2 > ac$，右邊是正數，所以這是一個圓，其

$$\text{圓心為} \left(-\frac{g}{a}, -\frac{f}{a}\right), \qquad \text{半徑為} \frac{\sqrt{g^2 + f^2 - ac}}{a}.$$

（若沒有這個條件，這個「圓」便會是單獨一點，或者是空集。）

兩支旗杆

把高度為 h_1 的旗杆的杆腳放在原點，另一支的杆腳放在 $(d, 0)$。若士兵位於 (x, y)，與兩個杆腳的距離分別是 r_1 和 r_2，則仰角 θ_1, θ_2 滿足 $\tan\theta_i = h_i / r_i$。兩個角都在 $(0, \frac{\pi}{2})$ 之

內，而 tan 在這個區間上是一對一（one-to-one）的，所以

$$\theta_1 = \theta_2 \iff \frac{h_1}{r_1} = \frac{h_2}{r_2} \iff h_2^2 r_1^2 = h_1^2 r_2^2,$$

最後一步可以逆推，因為所有量都是正數。（士兵不可能站在杆腳上：$r_1 = 0$ 會迫使 $r_2 = 0$ 同時成立。）所以路徑是

$$h_2^2(x^2+y^2) = h_1^2\big((x-d)^2+y^2\big) \iff (h_2^2-h_1^2)(x^2+y^2) + 2h_1^2 d\,x - h_1^2 d^2 = 0.$$

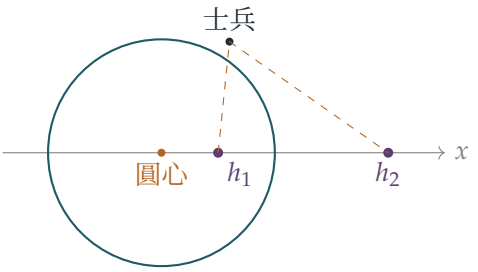
為旗杆編號，使得 $h_2 > h_1$（若兩者高度不同，總有一支較高）。這樣，上式就是第一部分的方程，其中

$$a = h_2^2 - h_1^2 > 0, \quad g = h_1^2 d, \quad f = 0, \quad c = -h_1^2 d^2,$$

而且

$$g^2 + f^2 - ac = h_1^4 d^2 + (h_2^2 - h_1^2)h_1^2 d^2 = h_1^2 h_2^2 d^2 > 0.$$

所以士兵沿着一個圓步操，圓心為 $\left(-\dfrac{h_1^2 d}{h_2^2 - h_1^2}, 0\right)$，半徑為 $\dfrac{h_1 h_2 d}{h_2^2 - h_1^2}$。圓心位於通過兩個杆腳的直線上，在較矮旗杆的外側。這個圓圍繞着較矮的旗杆，這很合理：要令矮杆看起來與高杆一樣高，就必須站得離它近一些。



（圖中取 $h_1 = 1$，$h_2 = 2$，$d = 3$：圓心是 $(-1, 0)$，半徑是 2。）

若 $h_1 = h_2$，$x^2 + y^2$ 的項互相抵消，方程變成 $2h_1^2 d\,x = h_1^2 d^2$，即 $x = \frac{1}{2}d$。士兵沿着連接兩個杆腳的線段的垂直平分線（perpendicular bisector）行走，那是一條直線而不是圓。這正是上述各圓的極限情況：當 $h_2 \to h_1$ 時，半徑趨向無限大。

三支旗杆

現在我們無法把三個杆腳全部放在同一條坐標軸上，所以設高度為 h_i 的旗杆立於 $\mathbf{p}_i = (p_i, q_i)$。以 $r_i^2 = (x-p_i)^2 + (y-q_i)^2$ 重複上面的論證，負責旗杆 i 和 j 的士兵所走的路徑是

$$(h_j^2 - h_i^2)(x^2+y^2) - 2(h_j^2 p_i - h_i^2 p_j)x - 2(h_j^2 q_i - h_i^2 q_j)y + \text{常數} = 0.$$

這是一個圓（移動原點不會改變 g^2+f^2-ac 的值，所以它仍然是 $h_i^2h_j^2|\mathbf{p}_i-\mathbf{p}_j|^2>0$；若 $h_j<h_i$，則先把全式乘以 -1）。利用第一部分，其圓心 $\mathbf{c}_{ij}=(x_{ij},y_{ij})$ 滿足

$$(h_j^2-h_i^2)\,\mathbf{c}_{ij}=h_j^2\mathbf{p}_i-h_i^2\mathbf{p}_j.$$

把 i 和 j 對調，這條式子不變（兩邊同時變號），所以 $\mathbf{c}_{31}$ 的下標次序怎樣寫都沒有關係。

> **想一想** 怎樣不必解出任何未知量，就把三條這樣的公式變成題目給出的恆等式？

恆等式中每一項在 x_{ij} 旁邊都已經有因子 $\pm(h_j^2-h_i^2)$，所以我們可以直接代入上面的公式，並使用向量，把 x 坐標和 y 坐標一併處理：

$$\begin{aligned}
h_3^2(h_1^2-h_2^2)\,\mathbf{c}_{12}&=-h_3^2(h_2^2\mathbf{p}_1-h_1^2\mathbf{p}_2)=-h_2^2h_3^2\,\mathbf{p}_1+h_1^2h_3^2\,\mathbf{p}_2,\\
h_1^2(h_2^2-h_3^2)\,\mathbf{c}_{23}&=-h_1^2(h_3^2\mathbf{p}_2-h_2^2\mathbf{p}_3)=-h_1^2h_3^2\,\mathbf{p}_2+h_1^2h_2^2\,\mathbf{p}_3,\\
h_2^2(h_3^2-h_1^2)\,\mathbf{c}_{31}&=\;\;h_2^2(h_3^2\mathbf{p}_1-h_1^2\mathbf{p}_3)=\;\;h_2^2h_3^2\,\mathbf{p}_1-h_1^2h_2^2\,\mathbf{p}_3.
\end{aligned}$$

三式相加，各項兩兩抵消，所以

$$h_3^2(h_1^2-h_2^2)\,\mathbf{c}_{12}+h_1^2(h_2^2-h_3^2)\,\mathbf{c}_{23}+h_2^2(h_3^2-h_1^2)\,\mathbf{c}_{31}=\mathbf{0}.$$

其 x 分量就是所求的 x_{ij} 之間的關係式，而 y 分量則是 y_{ij} 之間的同一關係式。

> **想一想** 為甚麼三點之間有一個線性關係，就能令它們共線？

把三個係數記為 $A=h_3^2(h_1^2-h_2^2)$，$B=h_1^2(h_2^2-h_3^2)$，$C=h_2^2(h_3^2-h_1^2)$。各高度都是正數而且互不相同，所以 A,B,C 全都不是零。關鍵的額外事實是它們的和為零：

$$A+B+C=h_1^2h_3^2-h_2^2h_3^2+h_1^2h_2^2-h_1^2h_3^2+h_2^2h_3^2-h_1^2h_2^2=0.$$

所以 $C=-A-B$，關係式變成

$$A(\mathbf{c}_{12}-\mathbf{c}_{31})+B(\mathbf{c}_{23}-\mathbf{c}_{31})=\mathbf{0},\qquad\text{所以}\qquad\mathbf{c}_{12}-\mathbf{c}_{31}=-\frac{B}{A}\,(\mathbf{c}_{23}-\mathbf{c}_{31}).$$

由 $\mathbf{c}_{31}$ 指向另外兩個圓心的向量互相平行（或者其中一個是零向量，若有兩個圓心重合），所以三個圓心位於同一條直線上。假如沒有 $A+B+C=0$，一個線性關係根本證明不了甚麼：任意三點都滿足某個這類的關係。

驗算

- 取 $h_1=1$，$h_2=2$，$d=3$，我們的公式給出圓心 $(-1,0)$ 和半徑 2。點 $(1,0)$ 與兩個杆腳的距離是 1 和 2，而 $(-3,0)$ 的距離是 3 和 6：兩者之比都是 $1:2$，正如所料。

- 我們做了一次數值檢驗：在每個圓上隨機取點，確認兩個仰角相等。此外，對 2000 組位置隨機的三支旗杆，也確認了恆等式成立，以及三個圓心共線。
- 預期共線的另一個理由：若某點同時在其中兩個圓上，則 $r_1 : r_2 = h_1 : h_2$ 且 $r_2 : r_3 = h_2 : h_3$，所以 $r_1 : r_3 = h_1 : h_3$，它也在第三個圓上。當這些圓相交時，它們全都通過相同的兩點，所以它們的圓心全都在這兩點的垂直平分線上。

常見錯誤

- 直接寫 $h_1/r_1 = h_2/r_2$，卻沒有說明為何仰角相等就得出正切相等（兩個角都是銳角）。
- 忘記檢查 $g^2 + f^2 > ac$：第一部分只在這個條件下才給出一個圓。
- 單憑恆等式就推斷共線，而沒有用到 $A + B + C = 0$ 和 $A, B, C \neq 0$。

STEP 之外

與兩個定點的距離成固定比 $k \neq 1$ 的點所組成的集合，是一個阿波羅尼奧斯圓。取遍 k 的所有值，便得到一個共軸圓系（*coaxal system*），即其中任意兩個圓都有相同根軸（radical axis）的一族圓。這類圓族正是反演幾何（inversive geometry）的起點；在複平面上，它們就是圓 $|z - a| = k|z - b|$，而莫比烏斯變換（Möbius transformation）會把這些圓互相映射。Coxeter 與 Greitzer 的 *Geometry Revisited*（重溫幾何）涵蓋共軸圓和反演；Needham 的 *Visual Complex Analysis*（視覺化的複分析）則展示複數方面的觀點。

詳細解答

STEP II 1994，第 8 題

純數學 • 二次模型的最優化

> **STEP II 1994，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/94-s2-q8

關鍵思路

追蹤一個送貨週期內的金錢流向。存貨以穩定的速率 r 減少，直至售罄為止，所以 $S(t)$ 只是一條簡單的斜線，而倉租就是 b 乘以這條斜線下的面積。於是每個週期的利潤是 X 的二次函數（quadratic）。在容許範圍 $0 < X \leq rT$ 內求它的最大值，便得出所需條件，而假設 $T \geq p/b$ 的作用，正是令拋物線（parabola）的頂點（vertex）落在這個範圍之內。

解答

為何 $X \leq rT$

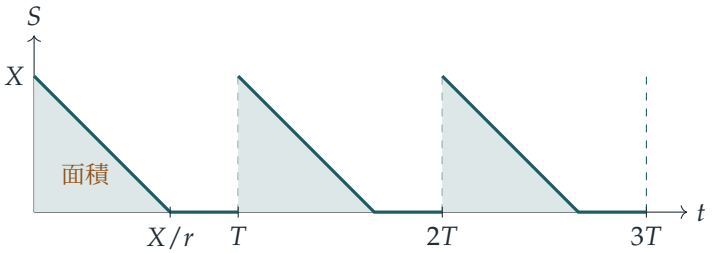
在 T 小時內，她最多只能賣出 rT 件貨。若 $X > rT$，下一輛貨車到達時，貨架上至少仍有 $X - rT$ 件貨。此後每個週期都會令這批賣不出的剩貨至少再增加 $X - rT$ 件，所以經過 n 個週期後，剩貨至少有 $n(X - rT)$ 件。這個數量會無止境地增長，按每件每小時 b 計算的倉租亦然。她要為這些貨品付款（還要付倉租），但它們不會帶來額外的銷售，因為每個週期的銷量無論如何都以 rT 為上限。因此每次訂貨多於 rT 只會增加成本，她應取 $X \leq rT$。

存貨量

當 $X \le rT$ 時，每個週期開始時恰好有 X 件貨。在連續模型（continuous model）中，只要有存貨，她便以速率 r 售貨，所以

$$S(t) = \begin{cases} X - rt, & 0 \le t \le X/r, \\ 0, & X/r \le t \le T. \end{cases}$$

圖像是一條直線，由 X 下降至 $t = X/r$ 時的零，然後保持在零，直到 $t = T$ 下一次送貨為止（屆時跳回 X）。經過數個週期，便形成鋸齒形的圖像。



每個週期的利潤

在長 T 小時的一個週期內：

- 銷售。全部 X 件貨都會售出（因為 $X \le rT$），每件售 £$(p+q)$，收入為 £$(p+q)X$。
- 送貨。£$(a+qX)$。
- 倉租。在一段短時間 δt 內存放 $S(t)$ 件貨的費用為 $bS(t)\,\delta t$，所以總數是

  $$b\int_0^T S(t)\,\mathrm{d}t = b\int_0^{X/r} (X - rt)\,\mathrm{d}t = b \cdot \tfrac{1}{2} \cdot X \cdot \frac{X}{r} = \frac{bX^2}{2r},$$

  即陰影三角形的面積乘以 b。
- 營運開支。£cT。

含 qX 的項互相抵銷，所以每個週期的利潤是

$$P(X) = pX - \frac{bX^2}{2r} - a - cT.$$

可見 p 才是每件貨的真正利潤：q 只是轉嫁給顧客而已。

生意何時可以有利可圖？

T 固定時，她可以在 $0 < X \le rT$ 的範圍內選擇 X。配方（completing the square）後，

$$P(X) = \frac{p^2 r}{2b} - a - cT - \frac{b}{2r}\left(X - \frac{pr}{b}\right)^2.$$

想一想 最佳的 X 真的在容許範圍之內嗎？

拋物線在 $X = pr/b$ 處達到最高點，而我們需要 $pr/b \le rT$，即 $T \ge p/b$。這正是題目的假設。所以她可以選 $X = pr/b$，此時

$$P = \frac{p^2 r}{2b} - (a + cT).$$

這個值為正，當且僅當 $p^2 r > 2b(a + cT)$，即

$$p^2 > 2\,\frac{(a + cT)b}{r}.$$

因此，若這個不等式（inequality）成立，訂貨量 $X = pr/b$ 便能令生意有利可圖，這正是所要證明的。（事實上 pr/b 是最佳的選擇，所以當 $T \ge p/b$ 時，這個條件也是必要的。）

反之，若 $T < p/b$，頂點便落在 rT 之外。這時 P 在整個容許範圍內都是遞增的，她最多只能取 $X = rT$。這就是為何要得出這個條件，必須有關於 T 的假設。

驗算

- 取 $a = 10$、$b = 1$、$r = 10$、$p = 4$、$c = 1$、$T = 5$。則 $T \ge p/b = 4$，且 $p^2 = 16 > 2(15)(1)/10 = 3$。最佳訂貨量為 $X = 40 \le rT = 50$，利潤為 $160 - 10 - 80 - 5 = 65 = \frac{16 \cdot 10}{2} - 15$。
- 電腦就 30 組隨機參數，以數值方法逐步計算一個週期內的 $S(t)$，結果與 $P(X)$ 相差不超過 10^{-7}。另就 2000 個滿足 $T \ge p/b$ 的隨機個案，對 X 作網格搜尋，證實最大值為 $\frac{p^2 r}{2b} - a - cT$。
- 單位：$p^2 r/b$ 的單位是 $(£/件)^2 \times (件/小時)/(£/件/小時) = £$，與 a 和 cT 相同。

常見錯誤

- 把全部 X 件貨在整整 T 小時內的倉租都計算在內（即 bXT），而不是按實際存貨量計算的 $b \int S \, \mathrm{d}t$。
- 忘記檢查 $X = pr/b$ 是否滿足 $X \le rT$。整道題只有這一步用到條件 $T \ge p/b$。

STEP 之外

這是運籌學（operations research）中經典的經濟訂貨量（economic order quantity）模型。在每次訂貨的固定成本與每件貨每單位時間的存貨成本之間取得平衡，會得出一條

平方根法則。在本題中，若她還可以選擇 T，則要令利潤率 P/T 最大，便須在 a 與倉租之間作出取捨。這類約束優化（constrained optimisation）問題（這裏是頂點與邊界 $X = rT$ 之間的比較）進一步發展下去，便是線性規劃（linear programming）、非線性規劃（nonlinear programming）以及 Karush–Kuhn–Tucker 條件。若想養成建立和檢驗這類模型的習慣，Siklos 的 *Advanced Problems in Mathematics* 是很好的參考書。

詳細解答

STEP II 1994，第 9 題

力學 • 剛體的平衡

> **STEP II 1994，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/94-s2-q9

關鍵思路

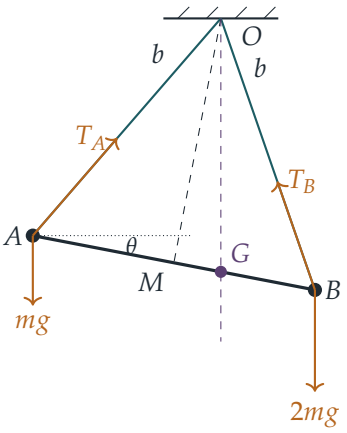
把桿、兩條繩和兩個質點視為一個剛體（rigid body）。作用於它的外力只有兩個重量和 O 點的支承力，因此要達致平衡（equilibrium），作用於質心（centre of mass）G 的合重量，其作用線必須通過 O：即 G 懸於 O 的正下方。這樣一來，求角度就變成純粹的等腰三角形幾何問題。至於張力（tension），則對 B 取力矩（moments），這樣可消去作用於 B 的兩個未知力。

解答

模型。桿和繩都是輕的，繩不可伸長，而 A、B 是質點（particles）。由於 $OA = OB = b$ 且 $AB = 2a$，三角形 OAB 是等腰三角形。設 M 為 AB 的中點；則 $OM \perp AB$，並由畢氏定理（Pythagoras）得

$$OM = h = \sqrt{b^2 - a^2}.$$

較重的質點 B 懸得較低，桿與水平線成角 θ。



角度

考慮整個系統：桿、繩和質點合在一起。作用於它的外力是 A 處的重量 mg、B 處的重量 $2mg$，以及 O 點支承所施的力。對 O 取力矩，支承力沒有力矩，所以兩個重量對 O 的總力矩必須為零。兩個重量互相平行，故可合成為一個通過兩質點質心 G 的單一力 $3mg$。它對 O 的力矩為零，當且僅當 G 位於通過 O 的鉛垂線上。所以 G 在 O 的正下方。

現在求 G 的位置。它在 AB 上，且 $AG : GB = 2m : m = 2 : 1$，所以 $AG = \frac{2}{3} \cdot 2a = \frac{4a}{3}$，因此

$$MG = AG - AM = \tfrac{4a}{3} - a = \tfrac{a}{3},$$

而 G 在 M 靠近 B 的一側。

? **想一想** 為甚麼角 MOG 等於 θ？

OG 是鉛垂的，而 OM 垂直於桿。把鉛垂線旋轉到 OM 上的旋轉，正是把水平線轉到桿上的那個旋轉，所以 OG 與 OM 之間的夾角等於桿與水平線之間的夾角，即 θ。

三角形 OMG 在 M 處是直角，所以

$$\tan\theta = \frac{MG}{OM} = \frac{a/3}{\sqrt{b^2 - a^2}} = \frac{a}{3\sqrt{b^2 - a^2}},$$

正如所求。

AO 的張力

再次把桿和質點合起來考慮。作用於這個系統的力是沿 AO 的 T_A、沿 BO 的 T_B，以及兩個重量。（桿內的力是內力。）對 B 取力矩，便可消去 T_B 和重量 $2mg$。

由 B 到直線 OA 的垂直距離，等於三角形 OAB 面積的兩倍除以 OA。該面積為 $\frac{1}{2} \cdot 2a \cdot$

$h = ah$，所以這距離是 $2ah/b$。A 處重量 mg 的力臂（moment arm）等於 A 與 B 之間的水平距離，即 $2a\cos\theta$。這兩個力矩方向相反，令它們互相平衡，得

$$T_A \cdot \frac{2ah}{b} = mg \cdot 2a\cos\theta \quad \Longrightarrow \quad T_A = \frac{mgb\cos\theta}{h}.$$

由 $\tan\theta = \dfrac{a}{3h}$，且 θ 為銳角，可得 $\cos\theta = \dfrac{3h}{\sqrt{9h^2 + a^2}}$，而 $9h^2 + a^2 = 9b^2 - 8a^2$。因此

$$T_A = \frac{mgb}{h} \cdot \frac{3h}{\sqrt{9b^2 - 8a^2}} = \frac{3mgb}{\sqrt{9b^2 - 8a^2}}.$$

（以同樣的論證對 A 取力矩，可得 $T_B = 2T_A$：每條繩的張力與它所支承的質量成正比。）

驗算

- 若 $a \to 0$，兩條繩變成鉛垂，而 $T_A \to 3mgb/(3b) = mg$，正好是質點 A 所需的張力。
- 對整個系統沿鉛垂方向分解力（resolving），結果應為 $3mg$。取 $a = 1$、$b = 2$、$m = 1$、$g = 9.8$，得 $T_A \approx 11.11\,\text{N}$ 和 $T_B \approx 22.22\,\text{N}$。把它們沿繩的方向分解，鉛垂分量的總和為 $29.4\,\text{N} = 3mg$。
- 以數值方法驗算：隨機選取 2000 組 (a, b, m)，把桿置於所預測的角度，然後分別對每個質點求解力的平衡。兩個質點所得的桿內力一致，而且是推力（thrust），張力則與公式相符，誤差約為 10^{-15}。

常見錯誤

- 把 G 放在距離 M 為 $\frac{2a}{3}$ 之處，或放在錯誤的一側。質心較接近較重的質點，與中點相距 $\frac{a}{3}$。
- 把 T_A 對 B 的力臂當作 $2a$ 或 b。所需的是由 B 到直線 OA 的垂直距離。
- 答案仍以 θ 表示，但題目要求以 m、g、a 和 b 表示。

STEP 之外

從一點自由懸掛的物體，靜止時質心必在該點的正下方；這個原理其實是一個更深刻事實的最簡單情況：穩定平衡使勢能（potential energy）取最小值，而這裏的勢能不過是 $3mg$ 乘以 G 的高度。到了大學，這會發展成虛功原理（virtual work principle）和拉格朗日力學（Lagrangian mechanics），當中求平衡位置的方法是對能量求導，而不是分解

力。Kleppner 與 Kolenkow 所著的 *An Introduction to Mechanics*（力學入門教科書）對剛體靜力學和能量方法有細緻的論述。

詳細解答

STEP II 1994，第 10 題

力學 • 彈性繩與簡諧運動

STEP II 1994，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/94-s2-q10

關鍵思路

着眼於繩子的伸長量（extension），而不是拖卡的位置。由於貨車的速度恆定，伸長量所滿足的方程與拖卡加速度的方程完全相同，只差一個正負號，而這個方程正是簡諧運動（simple harmonic motion）。因此繩子會在半個振動週期內保持拉緊，然後放開拖卡，而此時拖卡已比貨車走得更快。之後便只是一道普通的匀速追及問題。

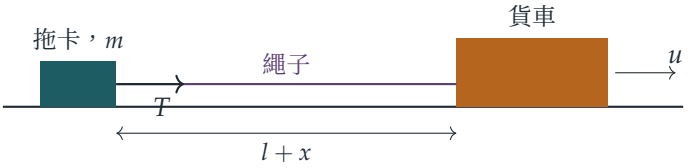
解答

建立模型

把拖卡視為質點（particle），把繩子視為輕繩，拉緊時服從胡克定律（Hooke's law），鬆弛時不施任何力，並忽略摩擦。由繩子首次拉緊的一刻起計時間 t，距離則沿運動方向量度。在 $t = 0$ 時，貨車以速率 u 行駛（並一直保持這個速率），拖卡靜止，而貨車在拖卡前方距離 l 處。

設拖卡位於 $y(t)$，其中 $y(0) = 0$，則貨車位於 $l + ut$。繩子的伸長量為

$$x = (l + ut) - y - l = ut - y.$$



繩子拉緊期間

當 $x > 0$ 時，張力（tension）為 $T = \lambda x/l$，它是拖卡所受的唯一水平力，把拖卡向前拉。所以

$$m\ddot{y} = \frac{\lambda x}{l}.$$

由於 $x = ut - y$ 而 u 為常數，$\ddot{x} = -\ddot{y}$，因此

$$\ddot{x} = -\omega^2 x, \qquad \text{其中} \omega^2 = \frac{\lambda}{ml}.$$

這就是簡諧運動。初始條件（initial conditions）為 $x(0) = 0$（繩子剛剛拉緊）和 $\dot{x}(0) = u - \dot{y}(0) = u$，所以

$$x = \frac{u}{\omega}\sin\omega t, \qquad \dot{x} = u\cos\omega t.$$

拖卡的速度為 $\dot{y} = u - \dot{x} = u(1 - \cos\omega t)$。

> **想一想** 按照簡諧運動的公式，過了一段時間後 x 會變成負數。那時公式還成立嗎？

不成立。方程 $\ddot{x} = -\omega^2 x$ 用到張力 $\lambda x/l$，而這只在 $x > 0$ 時適用；繩子不能推東西。公式在 $0 \le t \le \pi/\omega$ 內成立，此時 $\sin\omega t \ge 0$，而繩子再次鬆弛的時刻為

$$t_1 = \frac{\pi}{\omega} = \pi\left(\frac{lm}{\lambda}\right)^{1/2}.$$

在那一刻伸長量為零，所以拖卡又一次落後貨車距離 l，而它的速度為

$$\dot{y}(t_1) = u(1 - \cos\pi) = 2u.$$

繩子鬆弛之後

當 $t > t_1$ 時，拖卡以 $2u$ 行駛，貨車以 u 行駛，所以兩者之間的距離（在 $t = t_1$ 時為 l）以穩定的速率 $2u - u = u$ 縮短。此後這段距離小於 l，所以繩子保持鬆弛，拖卡仍然不受任何水平力，它的速率確實維持在 $2u$。再過時間 l/u，這段距離便縮短至零。

因此，拖卡與貨車相撞的時刻是繩子首次拉緊後的

$$t_1 + \frac{l}{u} = \pi\left(\frac{lm}{\lambda}\right)^{1/2} + \frac{l}{u}$$

正如題目所求。

驗算

- 單位：λ 是一個力，所以 lm/λ 的單位是 $\mathrm{m}\cdot\mathrm{kg}/\mathrm{N} = \mathrm{s}^2$，兩項都是時間。

- 轉換參考系（frame）：繩子的作用就像一條彈簧，令拖卡與一輛質量實際上無限大的貨車作彈性碰撞（elastic collision）。在貨車的參考系中，拖卡以 u 接近，再以 u 沿相反方向離開，所以它最後相對地面的速率是 $2u$，與上面的結果一致。
- 對拖卡的運動方程作數值積分（numerical integration）（繩子鬆弛時把力關掉），隨機選取 m、l、λ、u 的值，結果與公式吻合；例如選取 m, l, λ, u 使得 $\pi/\omega \approx 2.9944$ 及 $l/u \approx 2.3017$ 時，得出的碰撞時間為 5.2962，而公式的預測值為 5.2961。

常見錯誤

- 把簡諧運動的解用於所有時間，忘記了伸長量回到零時繩子會鬆弛。
- 直接就拖卡的位置 y 寫出運動方程，得到 $m\ddot{y} = \lambda(ut - y)/l$，這是一個受迫方程（forced equation）。它是可以解的，但代換 $x = ut - y$ 能立即把它化成非受迫方程。
- 漏掉最後的 l/u：繩子鬆弛時，拖卡仍然落後貨車整整一個自然長度（natural length）。

STEP 之外

轉到貨車的參考系後，整個過程就變成拖卡與一條固定在不動牆上的彈簧之間的一次彈性「碰撞」：拖卡以速率 u 進來，以速率 u 離開，而不論速率是多少，接觸的時間都恰好是半個週期。這是說明伽利略不變性（Galilean invariance）和能量守恆（energy conservation）如何約束碰撞的最簡單例子；這個想法貫穿大學一年級的力學，並引向散射理論（scattering theory）。Kleppner 與 Kolenkow 的 *An Introduction to Mechanics*（力學入門教科書）正是以這種精神處理振動和碰撞。

詳細解答

STEP II 1994，第 11 題

力學 • 拋體與爆炸

STEP II 1994，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/94-s2-q11

關鍵思路

爆炸發生在一個已知的高度，即垂直飛行的最高點，而那一刻炮彈是靜止的。因此兩塊碎片由同一點出發，速度大小相等、方向相反。若用一個角 θ 來描述方向，每塊碎片的落地時間都可由一條二次方程（quadratic equation）求得，而兩個落地點之間的距離就成為 θ 的單一函數。之所以出現兩種情況，是因為該函數的最大值要麼位於容許範圍的端點，要麼位於範圍內部的轉向點（turning point）。

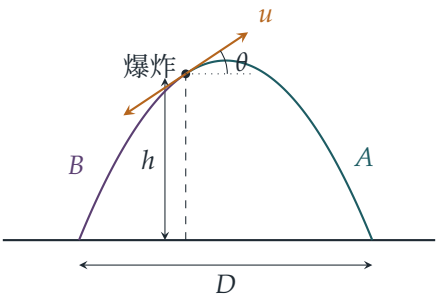
解答

建立模型

我們把炮彈和碎片都視為只受重力作用的質點（particle）。炮彈以初速 v 上升，所以它在以下高度靜止：

$$h = \frac{v^2}{2g}, \qquad \text{即} \qquad 2gh = v^2.$$

把剛爆炸後的兩塊碎片稱為 A 和 B。它們的速度大小相等、方向相反，所以兩者都在通過爆炸點的同一個垂直平面內。在該平面內取 x 軸為水平方向，y 軸垂直向上，爆炸點位於 $(0,h)$。我們可以適當地命名碎片，使 A 不向下運動。於是 A 的速度為 $(u\cos\theta,\ u\sin\theta)$，$B$ 的速度為 $(-u\cos\theta,\ -u\sin\theta)$，其中 $0 \le \theta \le \frac{\pi}{2}$。記 $s = \sin\theta$ 及 $c = \cos\theta$，故 $c \ge 0$。



落地時間

爆炸後的時間 t，碎片 A 的高度為 $h + ust - \frac{1}{2}gt^2$。它落地時

$$\tfrac{1}{2}gt^2 - us\,t - h = 0.$$

兩根之積為 $-2h/g < 0$，所以恰好有一個正根：

$$t_A = \frac{us + \sqrt{u^2s^2 + 2gh}}{g} = \frac{us + \sqrt{u^2s^2 + v^2}}{g}.$$

對於 B，垂直速度為 $-us$，所以我們把 s 換成 $-s$：

$$t_B = \frac{-us + \sqrt{u^2s^2 + v^2}}{g}.$$

這個值是正數，因為 $\sqrt{u^2s^2 + v^2} > us$。

落地點之間的距離

在水平方向上，A 沿 x 軸正方向移動了 $uc\,t_A$，B 沿負方向移動了 $uc\,t_B$，所以它們落在發射點的兩側。因此兩個落地點之間的距離為

$$D = uc\,(t_A + t_B) = \frac{2uc\sqrt{u^2s^2 + v^2}}{g}.$$

由於 $D \geq 0$，求 D 的最大值等同於求 D^2 的最大值。設 $x = s^2 \in [0,1]$，則 $c^2 = 1 - x$。於是

$$\frac{g^2D^2}{4u^2} = (1-x)(v^2 + u^2x) = f(x), \qquad \text{記作此式,}$$

並且

$$f(x) = v^2 + (u^2 - v^2)x - u^2x^2.$$

這是 x 的二次函數，首項係數為負，其頂點（vertex）位於

$$x^* = \frac{u^2 - v^2}{2u^2}.$$

想一想 f 的頂點是最大值點，那麼為甚麼答案不總是 $f(x^*)$？

只有 $x \in [0,1]$ 才是容許的，因為 $x = \sin^2\theta$。留意 $x^* < \frac{1}{2}$ 恆成立，所以頂點永遠不會超出上端。頂點是否在容許範圍內，只取決於它的正負。

情況 $u \le v$。此時 $x^* \le 0$。在 $[0,1]$ 上，我們位於一條開口向下的拋物線（parabola）的頂點右方，所以 f 在該區間上遞減，最大值在 $x = 0$ 處取得，而 $f(0) = v^2$。這對應 $\theta = 0$，即水平方向的爆炸。因此

$$D_{\max} = \frac{2u}{g}\sqrt{v^2} = \frac{2uv}{g}.$$

情況 $v \le u$。此時 $0 \le x^* < \frac{1}{2}$，所以頂點在容許範圍內，並給出最大值：

$$\begin{aligned} f(x^*) &= v^2 + \frac{(u^2-v^2)^2}{2u^2} - \frac{(u^2-v^2)^2}{4u^2} = v^2 + \frac{(u^2-v^2)^2}{4u^2} \\ &= \frac{4u^2v^2 + u^4 - 2u^2v^2 + v^4}{4u^2} = \frac{(u^2+v^2)^2}{4u^2}. \end{aligned}$$

所以

$$D_{\max} = \frac{2u}{g}\cdot\frac{u^2+v^2}{2u} = \frac{u^2+v^2}{g}.$$

這在 $\sin^2\theta = \dfrac{u^2-v^2}{2u^2}$ 時發生，所以碎片是與水平方向成一個角度飛出的。

當 $u = v$ 時，兩條公式都給出 $2u^2/g$，所以兩個答案在兩種情況的分界處一致，這是必然的。證明至此完成。

驗算

- 單位：uv/g 和 $(u^2+v^2)/g$ 都是（速率）^{2}/（加速度），即長度。
- 取 $g = 9.8$、$u = 30$ 及 $v = 10$，最佳角度為 $\theta \approx 41.8°$，而由 D 的一般公式和 $(u^2+v^2)/g$ 都得出 $D = 1000/9.8 \approx 102.04\,\text{m}$。水平方向的爆炸只給出 $2\cdot 30\cdot 10/9.8 \approx 61.2\,\text{m}$。
- 以數值方法逐步計算兩條軌跡，並在細密的網格上掃描 θ，對數百組隨機的 (u,v) 都能重現這兩個最大值。
- 當 $v = 0$ 時，炮彈在地面爆炸，答案變成 u^2/g。這正是以速率 u、仰角 $45°$ 發射的拋體（projectile）通常的最大射程，理應如此：B 立即撞入地面，而 $x^* = \frac{1}{2}$ 給出 $\theta = 45°$。

常見錯誤

- 假設最佳的爆炸方向是水平的。這只在 $u \leq v$ 時才成立。
- 對兩塊碎片使用相同的飛行時間。向上拋出的那一塊在空中停留得較久。
- 求最大值時忘記了約束條件 $0 \leq \sin^2\theta \leq 1$，因而漏掉了分情況討論。

STEP 之外

這個簡潔的答案有一個漂亮的解釋。質心（centre of mass）只是筆直下墜，而相對於質心，兩塊碎片以相對速率 $2u$ 沿一條直線互相遠離。但它們在不同的時間着地，所以問題就變成在約束條件下求最優值，而最大值會隨着參數的變化由邊界轉移到內部。這種轉換會在大學一年級最優化（optimisation）課程的拉格朗日乘數（Lagrange multiplier）和 Karush–Kuhn–Tucker 條件中再次出現。關於動量守恆（conservation of momentum）、質心參考系和更深入的拋體問題，可參閱 Kleppner 與 Kolenkow 的 *An Introduction to Mechanics*（力學入門教材）。

詳細解答

STEP II 1994，第 12 題

概率與統計 • 條件概率

STEP II 1994，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/94-s2-q12

關鍵思路

第一次擲骰一旦定出「目標點數」(point) s，其後每一次擲骰都是同一個獨立 (independent) 試驗，只有三種結果：擲出 s（勝）、擲出 7（負），或擲出其他點數（繼續）。因此，在給定目標點數的條件下，遊戲結束於第幾次擲骰服從幾何分佈 (geometric distribution)，而整道題目就是按目標點數分情況討論，再用全概率公式（law of total probability）加起來。最後一部分是同樣的計算，只是換了另一個作為條件的事件。

解答

設 S 為兩顆公平骰子的點數和。數算 36 個等可能的結果，可得

$$\mathrm{P}(S=s)=\frac{6-|s-7|}{36},\qquad s=2,3,\ldots,12,$$

所以點數和 $2,3,4,5,6,7,8,9,10,11,12$ 的概率為 $\frac{1}{36}\times(1,2,3,4,5,6,5,4,3,2,1)$。

第一次擲骰即獲勝

若第一次的點數和是 7 或 11，她立即獲勝：

$$\mathrm{P}(\text{獲勝擲次為1})=\frac{6}{36}+\frac{2}{36}=\frac{2}{9}.$$

已知擲骰多於一次，在第 n 次擲骰獲勝

她擲骰多於一次，當且僅當第一次的點數和是目標點數 $4,5,6,8,9,10$ 其中之一，其概率為

$$\mathrm{P}(M)=\frac{3+4+5+5+4+3}{36}=\frac{2}{3},$$

其中 M 是事件「她擲骰多於一次」。

固定一個目標點數 s，並記 $p_s = \mathrm{P}(S = s)$。其後每一次擲骰都與之前的擲骰獨立，而在該次擲骰中，

$$\mathrm{P}(\text{點數和為}s) = p_s, \qquad \mathrm{P}(\text{點數和為}7) = \tfrac{1}{6}, \qquad \mathrm{P}(\text{兩者皆非}) = r_s = 1 - p_s - \tfrac{1}{6}.$$

（由於 $s \neq 7$，這兩個事件互斥（disjoint）。）她在第 $n \geq 2$ 次擲骰獲勝，當且僅當第 $2, \ldots, n-1$ 次擲骰全是「兩者皆非」，而第 n 次擲出 s。所以，由獨立性，

$$\mathrm{P}(\text{獲勝擲次為}n \mid \text{目標點數為}s) = r_s^{n-2}\, p_s.$$

目標點數可以配成概率相等的三對（4 與 10、5 與 9、6 與 8）。記 $p_s = k/36$，其中 $k = 3, 4, 5$；則 $r_s = (30-k)/36$，分別為 $\frac{3}{4}$、$\frac{13}{18}$、$\frac{25}{36}$。由全概率公式，對於 $n \geq 2$，

$$\mathrm{P}(\text{獲勝擲次為}n) = \sum_s p_s \cdot r_s^{n-2} p_s = 2\sum_{k=3}^{5} \left(\frac{k}{36}\right)^2 \left(\frac{30-k}{36}\right)^{n-2}.$$

在第 $n \geq 2$ 次擲骰獲勝必然意味着 M 發生，所以除以 $\mathrm{P}(M) = \frac{2}{3}$ 就相當於把每個係數乘以 $\frac{3}{2}$：

$$2 \cdot \frac{k^2}{1296} \cdot \frac{3}{2} = \frac{k^2}{432} \quad \text{得出} \quad \frac{9}{432} = \frac{1}{48}, \quad \frac{16}{432} = \frac{1}{27}, \quad \frac{25}{432}.$$

因此

$$\mathrm{P}(\text{獲勝擲次為}n \mid M) = \frac{1}{48}\left(\frac{3}{4}\right)^{n-2} + \frac{1}{27}\left(\frac{13}{18}\right)^{n-2} + \frac{25}{432}\left(\frac{25}{36}\right)^{n-2},$$

這正是所要求的結果。

想一想 為甚麼不可以只用一個「平均」目標點數和單一個幾何分佈來計算？

因為繼續下去的概率 r_s 取決於目標點數。三個幾何分佈的混合（mixture）本身並不是幾何分佈，所以答案中有三個不同的冪，而不是只有一個。

已知擲骰多於 m 次

設 M_m 為事件「她擲骰多於 m 次」，其中 $m > 1$。這個事件發生，當且僅當第一次擲骰定出目標點數 s，而第 $2, \ldots, m$ 次擲骰全是「兩者皆非」，所以

$$\mathrm{P}(M_m) = \sum_s p_s\, r_s^{m-1} = \frac{2}{36}\sum_{k=3}^{5} k\left(\frac{30-k}{36}\right)^{m-1}.$$

若 $n \leq m$，在第 n 次擲骰獲勝與 M_m 不能同時發生，所以條件概率（conditional probability）為 0。若 $n > m$，在第 n 次擲骰獲勝必然意味着 M_m 發生，所以

$$\mathrm{P}(\text{獲勝擲次為}n \mid M_m) = \frac{\mathrm{P}(\text{獲勝擲次為}n)}{\mathrm{P}(M_m)} = \frac{9\left(\frac{3}{4}\right)^{n-2} + 16\left(\frac{13}{18}\right)^{n-2} + 25\left(\frac{25}{36}\right)^{n-2}}{36\left[3\left(\frac{3}{4}\right)^{m-1} + 4\left(\frac{13}{18}\right)^{m-1} + 5\left(\frac{25}{36}\right)^{m-1}\right]},$$

其中因子 $\frac{2}{1296}$ 和 $\frac{2}{36}$ 已經約去。等價地，把分子和分母同乘以 $\frac{1}{432}$，得

$$\frac{\frac{1}{48}\left(\frac{3}{4}\right)^{n-2}+\frac{1}{27}\left(\frac{13}{18}\right)^{n-2}+\frac{25}{432}\left(\frac{25}{36}\right)^{n-2}}{\frac{1}{4}\left(\frac{3}{4}\right)^{m-1}+\frac{1}{3}\left(\frac{13}{18}\right)^{m-1}+\frac{5}{12}\left(\frac{25}{36}\right)^{m-1}}, \qquad n>m.$$

留意分母正是 $\mathrm{P}(M_m \mid M)$，這是理所當然的：分子是先前以 M 為條件的答案，而 $M_m \subseteq M$。

驗算

- 取 $m=1$ 時，分母變成 $\frac{1}{4}+\frac{1}{3}+\frac{5}{12}=1$，還原為上一部分的結果。
- 當 $n=2$ 時，條件概率公式給出 $\frac{1}{48}+\frac{1}{27}+\frac{25}{432}=\frac{50}{432}=\frac{25}{216}\approx 0.1157$；直接計算，$\frac{3}{2}\cdot\frac{2(9+16+25)}{1296}$ 給出相同的值。
- 當 $m=2$、$n=3$ 時，答案是 $\frac{107}{930}\approx 0.1151$。我們用分數把遊戲逐步精確推演，核對了所有公式；另外模擬了 4×10^5 局遊戲，結果在抽樣誤差範圍內吻合。

常見錯誤

- 忘記除以 $\mathrm{P}(M)=\frac{2}{3}$，結果得到係數 $\frac{1}{72}$、$\frac{2}{81}$、$\frac{25}{648}$。
- 以單一個公比寫成 $\mathrm{P}(M_m)=\left(\text{某數}\right)^{m-1}$：作為條件的事件同樣是對各個目標點數的混合。
- 把「兩者皆非」的擲骰次數算成 $n-1$ 而不是 $n-2$：第 1 次擲骰用來定出目標點數，第 n 次擲骰則是獲勝的一擲。

STEP 之外

這個遊戲是一條小型的馬可夫鏈（Markov chain），有「勝」和「負」兩個吸收態（absorbing states），而這裏的吸收時間是幾何分佈的混合。以擲過 m 次仍未結束為條件，會令混合的權重移向維持得最久的目標點數（這裏是 4 和 10，即最難再次擲出的點數）；這是一個簡單的例子，說明無記憶（memoryless）分佈的混合本身並不具有無記憶性。Grimmett and Welsh, *Probability: An Introduction*（概率論入門）以這個程度講解條件概率和馬可夫鏈。

詳細解答

STEP II 1994，第 13 題

概率與統計 • 贈券收集問題

STEP II 1994，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/94-s2-q13

本題在 *Beyond STEP* 第 9.11 節集齊一套中有完整的引導式講解：https://beyondstep.keithmcnulty.org/zh-hant/ch09/coupon-collector.html

關鍵思路

這是經典的「優惠券收集問題」（coupon collector problem），雖然資料庫把它歸入別的類別，它其實是純粹的概率（probability）題。把集齊全套所需的漫長等待拆成 n 段較短的等待：先等第一張新卡，再等第二張，如此類推。每段短等待都服從幾何分佈（geometric distribution），所以平均值很容易求得，再用期望值的線性性質（linearity of expectation）把它們加起來。總和是 n 乘以一個調和和（harmonic sum），而題中的積分不等式正是用來告訴我們這個和有多大。

解答

等待一張新卡

我們假設每包內有 n 張卡的其中一張，每張出現的概率均為 $\frac{1}{n}$，而且各包之間互相獨立（independent）。

假設我已經擁有 r 張不同的卡，其中 $0 \le r \le n-1$。打開一包得到新卡的概率是

$$p = \frac{n-r}{n} > 0,$$

得到舊卡的概率則是 $q = 1 - p = \frac{r}{n}$。設 X 為我直至第一次得到新卡為止（包括該包在內）所打開的包數。由於各包互相獨立，$X = k$ 表示先有 $k-1$ 張舊卡，然後是一張新

卡，所以

$$\mathrm{P}(X=k)=q^{k-1}p, \qquad k=1,2,3,\ldots$$

因此

$$\mathrm{E}(X)=\sum_{k=1}^{\infty} kpq^{k-1}=p\,(1-q)^{-2}=\frac{p}{p^2}=\frac{1}{p},$$

其中用到當 $|q|<1$ 時 $\sum_{k\geq 1} kq^{k-1}=(1-q)^{-2}$（把幾何級數（geometric series）求導，或者用二項式定理展開 $(1-q)^{-2}$）。這裏 $0\leq q<1$，因為 $r\leq n-1$。所以

$$\mathrm{E}(X)=\frac{n}{n-r},$$

這正是所要證明的結果。

（一個更簡潔的做法：第一包要麼給出新卡，我們就此停下；要麼以概率 q 令我們回到起點，所以 $\mathrm{E}(X)=1+q\,\mathrm{E}(X)$，由此得 $\mathrm{E}(X)=1/(1-q)=1/p$。這個做法假設了 $\mathrm{E}(X)$ 是有限的，而上面的級數已確認了這一點。）

積分不等式

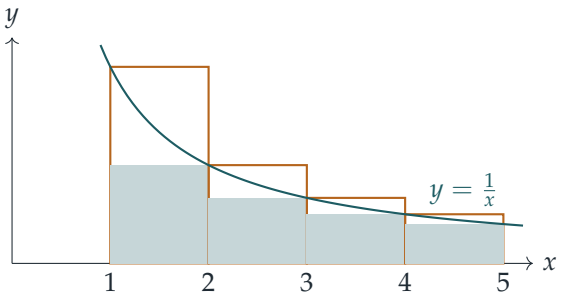
設 $r\geq 1$。當 $r\leq x\leq r+1$ 時，我們有

$$\frac{1}{r+1}\leq\frac{1}{x}\leq\frac{1}{r},$$

因為 $1/x$ 在 $x>0$ 時是遞減的。在長度為 1 的區間上積分，這些不等式仍然成立，所以

$$\frac{1}{r+1}\leq\int_r^{r+1}\frac{1}{x}\,\mathrm{d}x\leq\frac{1}{r}.$$

在圖中，曲線 $y=1/x$ 在 $x=r$ 與 $x=r+1$ 之間的面積，介乎高度為 $\frac{1}{r+1}$ 的矮長方形與高度為 $\frac{1}{r}$ 的高長方形之間。



（為清楚起見，縱軸的比例已經拉長。）塗上陰影的長方形位於曲線之下；只畫出外框的長方形則高於曲線。

現在把 $r = 1, 2, \ldots, n-1$ 的不等式加起來，其中 $n \geq 2$。各個積分會連接成一個：

$$\sum_{r=1}^{n-1} \frac{1}{r+1} \leq \int_1^n \frac{1}{x}\,\mathrm{d}x \leq \sum_{r=1}^{n-1} \frac{1}{r}.$$

中間的一項是 $\ln n - \ln 1 = \ln n$，而在左邊的和式中，我們把 $r+1$ 改記為 r，它由 2 取到 n。所以

$$\sum_{r=2}^{n} \frac{1}{r} \leq \ln n \leq \sum_{r=1}^{n-1} \frac{1}{r},$$

這正是所要證明的結果。

集齊全套

設 X_r 為我恰好擁有 r 張不同的卡期間所打開的包數，以給出第 $r+1$ 張卡的那一包作結。集齊全套所需的總包數是

$$T = X_0 + X_1 + \cdots + X_{n-1}.$$

> **想一想** 各個 X_r 取決於之前發生過的事。我們仍然可以直接把它們的期望值相加嗎？

可以。期望值的線性性質 $\mathrm{E}(A+B) = \mathrm{E}(A) + \mathrm{E}(B)$ 完全不需要獨立性。（事實上，無論之前的經過如何，X_r 都服從 $p = \frac{n-r}{n}$ 的幾何分佈，所以第一部分的結果適用於每一個 X_r。）因此

$$\mathrm{E}(T) = \sum_{r=0}^{n-1} \frac{n}{n-r} = n\left(\frac{1}{n} + \frac{1}{n-1} + \cdots + \frac{1}{1}\right) = n\sum_{k=1}^{n} \frac{1}{k},$$

其中記 $k = n - r$。

記 $H_n = \sum_{k=1}^{n} \frac{1}{k}$。不等式的兩邊分別給出

$$H_n = 1 + \sum_{k=2}^{n} \frac{1}{k} \leq 1 + \ln n, \qquad H_n = \sum_{k=1}^{n-1} \frac{1}{k} + \frac{1}{n} \geq \ln n + \frac{1}{n}.$$

所以

$$n\ln n + 1 \leq \mathrm{E}(T) \leq n\ln n + n.$$

上下界之間的差距小於 n，當 n 很大時，這與 $n\ln n$ 相比微不足道（兩者之比 $1/\ln n \to 0$）。所以當 n 很大時，我預計大約要吃

$$n\ln n \text{ 包。}$$

例如，100 位數學家就意味着大約 460 包 Cruncho。

驗算

- $r = 0$：第一包必定是新卡，而 $n/(n-0) = 1$。$r = n-1$：只欠某一張指定的卡，它出現的概率是 $\frac{1}{n}$，所以我們預計需要 n 包，與 $n/1$ 吻合。
- $n = 2$：$\mathrm{E}(T) = 2(1+\frac{1}{2}) = 3$，而上下界分別給出 $2\ln 2 + 1 \approx 2.39$ 和 $2\ln 2 + 2 \approx 3.39$。
- 以 $n = 20$ 模擬 10^5 次完整的收集過程，平均需要 71.9 包，而 $20H_{20} \approx 71.95$；上下界是 60.9 和 79.9。

常見錯誤

- 由 $r = 0$ 開始把積分不等式相加：界 $\frac{1}{r}$ 在該處不成立，而且 $\int_0^1 \frac{1}{x}\,\mathrm{d}x$ 發散（diverge）。求和必須由 $r = 1$ 開始。
- 以為各個 X_r 必須互相獨立，才可以把期望值相加。線性性質並不需要獨立性。
- 答 $\ln n$ 或 n：包數的期望值的增長速度與 $n\ln n$ 相若，比 n 快。

STEP 之外

差 $H_n - \ln n$ 遞減至一個極限，即歐拉－馬斯刻若尼常數（Euler–Mascheroni constant）$\gamma \approx 0.5772$，所以更精確的答案是 $\mathrm{E}(T) \approx n(\ln n + \gamma)$；當 $n = 100$ 時，這給出大約 519 包，而不是 460 包。像這樣把和與積分作比較，就是級數的積分審斂法（integral test），在大學第一門分析課程的初段就會遇到。優惠券收集問題本身亦會在演算法分析中再次出現，也會在研究要隨機洗牌多少次才能把一副牌洗勻時出現。Grimmett 與 Welsh 合著的 *Probability: An Introduction*（概率入門教科書）是一個好去處，可以看到幾何等待時間與期望值的線性性質如何一併運用。

詳細解答

STEP II 1994，第 14 題

概率與統計 • 幾何概率

> **STEP II 1994，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/94-s2-q14

關鍵思路

「在圓盤上均勻分佈（uniform）」的意思是概率與面積成正比。因此，落點與圓心的距離不超過 r 的概率，就是較小圓盤的面積除以鏢靶的面積，由此立刻得到 R 的分佈函數（distribution function）。第 (ii) 部分於是化為三件事：每次投擲的單一概率 p、一個幾何分佈式的「至少成功一次」計算，以及求一個整數變量函數的最小值。最後一項我們不用微積分，而是考察相鄰兩項之差。

解答

第 (i) 部分

把每枝飛鏢的落點看成在半徑為 a 的圓盤上均勻分佈的一點，於是飛鏢落在鏢靶上任何區域的概率，就是該區域的面積除以 πa^2。

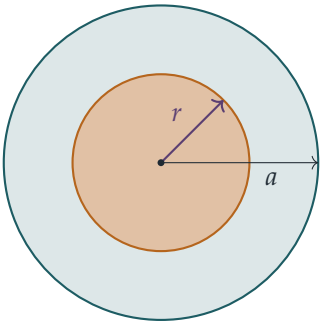
對於 $0 \le r \le a$，事件 $R \le r$ 表示飛鏢落在半徑為 r 的同心圓盤內，所以

$$\mathrm{P}(R \le r) = \frac{\pi r^2}{\pi a^2} = \frac{r^2}{a^2}.$$

求導可得 R 的概率密度函數（probability density function）為

$$f(r) = \frac{2r}{a^2} \qquad (0 \le r \le a),$$

而在其他情況下 $f(r) = 0$。



注意 R 在 $[0,a]$ 上並非均勻分佈：半徑越大，面積越多（薄圓環越長），這就是密度隨 r 線性增長的原因。

現在

$$\mathrm{E}(R) = \int_0^a r \cdot \frac{2r}{a^2}\,\mathrm{d}r = \frac{2}{a^2}\cdot\frac{a^3}{3} = \frac{2a}{3},$$
$$\mathrm{E}(R^2) = \int_0^a r^2 \cdot \frac{2r}{a^2}\,\mathrm{d}r = \frac{2}{a^2}\cdot\frac{a^4}{4} = \frac{a^2}{2},$$

所以

$$\mathrm{Var}(R) = \mathrm{E}(R^2) - \big(\mathrm{E}(R)\big)^2 = \frac{a^2}{2} - \frac{4a^2}{9} = \frac{a^2}{18},$$

這正是所要證明的方差（variance）。

第 (ii) 部分

（我們把題目中的「with $a/\sqrt{10}$」理解為「within $a/\sqrt{10}$」。）由第 (i) 部分，單次投擲落在距圓心 $a/\sqrt{10}$ 以內的概率為

$$p = \frac{(a/\sqrt{10})^2}{a^2} = \frac{1}{10}.$$

我們假設各次投擲互相獨立（independent）（每次都是鏢靶上一個全新的均勻隨機點）。那麼 m 次投擲中沒有一次落在 $a/\sqrt{10}$ 以內的概率為 $(0.9)^m$，所以

$$\mathrm{P}(\text{他贏回£12}) = 1 - (0.9)^m.$$

他的損失是所付的 $£(12+m)$ 減去贏回的金額，所以他的期望損失（expected loss，以英鎊計）為

$$L(m) = 12 + m - 12\big(1-(0.9)^m\big) = m + 12(0.9)^m.$$

每多投一次，肯定要多付 £1，但贏得 £12 的機會也隨之增加；我們需要找出兩者在何處取得平衡。

想一想 一個只在整數上有定義的函數，怎樣求它的最小值？

比較相鄰的值。對於 $m \geq 1$，

$$L(m+1) - L(m) = 1 + 12(0.9)^m(0.9-1) = 1 - 1.2\,(0.9)^m.$$

由於 $(0.9)^m$ 隨 m 增大而減小，這個差隨 m 遞增，所以當差為負時 L 遞減，一旦差變為正便遞增。差為負當且僅當 $(0.9)^m > \frac{1}{1.2} = \frac{5}{6}$。現在

$$(0.9)^1 = 0.9 > \tfrac{5}{6}, \qquad (0.9)^2 = 0.81 < \tfrac{5}{6},$$

所以 $L(2) - L(1) < 0$，而對每個 $m \geq 2$ 都有 $L(m+1) - L(m) > 0$。因此

$$L(1) > L(2) < L(3) < L(4) < \cdots,$$

而事實上 $L(1) = 11.80$，$L(2) = 11.72$，$L(3) = 11.748$。（他決意要玩，所以不容許 $m = 0$；無論如何 $L(0) = 12$ 也較大。）

他應選擇投 $m = 2$ 次，期望損失為 £11.72。由於這個最小值是嚴格的，$m = 2$ 是唯一符合要求的值。

驗算

- 在單位圓盤內模擬 4×10^5 個均勻分佈的點，得出平均值 0.667、平方的平均值 0.500 及方差 $0.0556 \approx \frac{1}{18}$，與 $\frac{2}{3}$、$\frac{1}{2}$ 及 $\frac{1}{18}$ 吻合；約有 10% 的點落在距圓心 $1/\sqrt{10}$ 以內。
- 準確值為 $L(1) = \frac{59}{5}$，$L(2) = \frac{293}{25}$，$L(3) = \frac{2937}{250}$，而在 $1 \leq m \leq 99$ 範圍內逐一搜尋，亦確認 $m = 2$ 為最小值所在。
- 合理性檢查：方差 $a^2/18$ 遠小於 $[0, a]$ 上均勻分佈變量的方差 $a^2/12$，因為 R 集中在靠近邊緣的位置。

常見錯誤

- 把 R 當作在 $[0, a]$ 上均勻分佈，這樣會得出方差 $a^2/12$。
- 忘記固定的 £12 只是成本的一部分：損失是 $12 + m$ 減去期望贏得的金額，而不只是 m。
- 令 $m + 12(0.9)^m$ 的導數等於零會得出 $m \approx 2.3$，這並不是答案；你必須比較它兩側的整數值。

STEP 之外

由平面上一個均勻分佈的點求 R 的分佈，是把聯合密度（joint density）變換到極坐標（polar coordinates）的最簡單例子，其中 $\mathrm{d}A = r\,\mathrm{d}r\,\mathrm{d}\theta$ 裏的雅可比（Jacobian）因子 r，正是令 R 的密度與 r 成正比的原因。聯合密度的變量代換，以及在離散選擇上求最優，都是大學一年級概率課程的標準課題。Grimmett 與 Welsh 所著的 *Probability: An Introduction*（概率入門教材）涵蓋了這兩個課題。

詳細解答

STEP III 1994，第 1 題

純數學 • 歸約公式

STEP III 1994，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/94-s3-q1

關鍵思路

對於第一個積分，把 $\operatorname{sech} t$ 寫成 $\dfrac{\cosh t}{\cosh^2 t} = \dfrac{\cosh t}{1+\sinh^2 t}$：分子正是 $\sinh t$ 的導數（derivative），所以這其實是一個喬裝了的 arctan。至於歸約公式（reduction formula），最自然的做法是分出一個因子 $\operatorname{sech}^2 t$，即 $\tanh t$ 的導數，然後用分部積分法（integration by parts）。恆等式（identity）$\tanh^2 t = 1 - \operatorname{sech}^2 t$ 隨即令 I_n 在右邊重新出現，我們再把它合併起來。

解答

$\operatorname{sech} t$ 的積分

把分子和分母同乘以 $\cosh t$，並利用 $\cosh^2 t = 1 + \sinh^2 t$，

$$\int_0^x \operatorname{sech} t \,\mathrm{d}t = \int_0^x \frac{\cosh t}{1+\sinh^2 t}\,\mathrm{d}t.$$

設 $s = \sinh t$，則 $\mathrm{d}s = \cosh t \,\mathrm{d}t$，積分上下限變成 0 和 $\sinh x$，所以

$$\int_0^x \operatorname{sech} t \,\mathrm{d}t = \int_0^{\sinh x} \frac{\mathrm{d}s}{1+s^2} = \arctan(\sinh x).$$

另一個同樣好的做法是取 $\operatorname{sech} t = \dfrac{2\mathrm{e}^t}{\mathrm{e}^{2t}+1}$ 並設 $u = \mathrm{e}^t$，所得答案的形式是 $2\arctan(\mathrm{e}^x) - \frac{\pi}{2}$。兩者是一致的：它們在 $x = 0$ 時都等於零，而且導數都是 $\operatorname{sech} x$。（這個以 x 為變量的函數稱為古德曼函數（Gudermannian）。）

歸約公式

取 $n \geq 2$，並寫 $\operatorname{sech}^n t = \operatorname{sech}^{n-2} t \cdot \operatorname{sech}^2 t$。用分部積分法，取

$$u = \operatorname{sech}^{n-2} t, \qquad \frac{\mathrm{d}v}{\mathrm{d}t} = \operatorname{sech}^2 t,$$

於是 $v = \tanh t$，而根據鏈式法則（chain rule）（利用 $\frac{\mathrm{d}}{\mathrm{d}t}\operatorname{sech} t = -\operatorname{sech} t \tanh t$），

$$\frac{\mathrm{d}u}{\mathrm{d}t} = (n-2)\operatorname{sech}^{n-3} t \cdot (-\operatorname{sech} t \tanh t) = -(n-2)\operatorname{sech}^{n-2} t \tanh t.$$

於是

$$\begin{aligned} I_n &= \left[\operatorname{sech}^{n-2} t \tanh t\right]_0^x + (n-2)\int_0^x \operatorname{sech}^{n-2} t \tanh^2 t \,\mathrm{d}t \\ &= \operatorname{sech}^{n-2} x \tanh x + (n-2)\int_0^x \operatorname{sech}^{n-2} t \left(1 - \operatorname{sech}^2 t\right) \mathrm{d}t \\ &= \operatorname{sech}^{n-2} x \tanh x + (n-2) I_{n-2} - (n-2) I_n, \end{aligned}$$

其中在 0 處的邊界項等於零，因為 $\tanh 0 = 0$。

想一想 我們似乎兜了一個圈：I_n 又在右邊出現了。這是個問題嗎？

完全不是，這正是我們想要的。右邊的 I_n 的係數是 $-(n-2)$，而不是 $+1$，所以把它移到左邊便得到 $(n-1)I_n$，而我們可以除以 $n - 1 \neq 0$。歸約公式就是

$$\boxed{(n-1) I_n = \operatorname{sech}^{n-2} x \tanh x + (n-2)\, I_{n-2} \qquad (n \geq 2).}$$

求 I_5

奇數的一串由 $I_1 = \arctan(\sinh x)$ 開始。代入 $n = 3$：

$$2I_3 = \operatorname{sech} x \tanh x + I_1, \qquad \text{所以} \qquad I_3 = \tfrac{1}{2}\operatorname{sech} x \tanh x + \tfrac{1}{2}\arctan(\sinh x).$$

代入 $n = 5$：

$$4I_5 = \operatorname{sech}^3 x \tanh x + 3I_3,$$

所以

$$I_5 = \tfrac{1}{4}\operatorname{sech}^3 x \tanh x + \tfrac{3}{8}\operatorname{sech} x \tanh x + \tfrac{3}{8}\arctan(\sinh x).$$

求 I_6

偶數的一串由 $I_0 = \int_0^x 1 \,\mathrm{d}t = x$ 開始，或者更直接地由 $I_2 = \int_0^x \operatorname{sech}^2 t \,\mathrm{d}t = \tanh x$ 開始（在公式中代入 $n = 2$ 可以確認這一點：$I_2 = \tanh x + 0$）。代入 $n = 4$：

$$3I_4 = \operatorname{sech}^2 x \tanh x + 2I_2, \qquad \text{所以} \qquad I_4 = \tfrac{1}{3}\operatorname{sech}^2 x \tanh x + \tfrac{2}{3}\tanh x.$$

代入 $n=6$：

$$5I_6 = \operatorname{sech}^4 x \tanh x + 4I_4,$$

所以

$$I_6 = \tfrac{1}{5}\operatorname{sech}^4 x \tanh x + \tfrac{4}{15}\operatorname{sech}^2 x \tanh x + \tfrac{8}{15}\tanh x.$$

偶數次冪的「其他方法」

偶數次冪有一個更簡潔的直接做法。設 $u = \tanh t$，則 $\mathrm{d}u = \operatorname{sech}^2 t\,\mathrm{d}t$，而 $\operatorname{sech}^2 t = 1-u^2$，所以

$$I_6 = \int_0^x (\operatorname{sech}^2 t)^2 \operatorname{sech}^2 t\,\mathrm{d}t = \int_0^{\tanh x} (1-u^2)^2\,\mathrm{d}u = \tanh x - \tfrac{2}{3}\tanh^3 x + \tfrac{1}{5}\tanh^5 x.$$

在先前的答案中把 $\operatorname{sech}^2 x$ 換成 $1-\tanh^2 x$ 再展開，所得的正是這個式子，所以兩種形式是一致的。

驗算

- 當 $x \to \infty$ 時，$\tanh x \to 1$、$\operatorname{sech} x \to 0$ 且 $\arctan(\sinh x) \to \frac{\pi}{2}$，所以 $I_5 \to \frac{3\pi}{16} \approx 0.58905$ 而 $I_6 \to \frac{8}{15}$。在 $[0,30]$ 上用森遜法則（Simpson's rule）得出 0.589049 和 0.533333。
- 在 $x=1$ 處，森遜法則給出 $I_5 \approx 0.561566$ 和 $I_6 \approx 0.518343$，與公式吻合；我們亦在 200 個隨機的 x 值上，對 $2 \le n \le 8$ 檢驗過歸約公式。
- I_5 和 I_6 都是 x 的奇函數（odd function），這是必然的，因為 sech 是偶函數（even function）。

常見錯誤

- 弄錯 $\frac{\mathrm{d}}{\mathrm{d}t}\operatorname{sech} t$ 的正負號：它是 $-\operatorname{sech} t \tanh t$，而正是這個負號令邊界項成為正數。
- 把兩個 I_n 合併之後，除以 n 而不是 $n-1$。
- 把答案留作不定積分（indefinite integral）：積分上下限是 0 和 x，所以沒有任意常數，而且每一項在 $x=0$ 時都應該等於零。

STEP 之外

像這樣的歸約公式，讓你可以用遞歸（recursion）的方法求出一整族積分的值；同一個想法也是沃利斯（Wallis）公式的基礎，該公式處理 $\int_0^{\pi/2} \sin^n t\,\mathrm{d}t$，並由此引出一個關

於 π 的著名無窮乘積（infinite product）。這裏求得的極限 $\int_0^\infty \operatorname{sech}^n t \, \mathrm{d}t$ 是貝塔函數（Beta function）的值，而貝塔函數是大學一年級分析和概率課程中的核心對象。Siklos 的 *Advanced Problems in Mathematics* 有更多歸約公式的練習，而 Abbott 的 *Understanding Analysis* 則為反常積分（improper integral）提供嚴格的處理。

詳細解答

STEP III 1994，第 2 題

純數學 • 回文四次方程

> **STEP III 1994，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/94-s3-q2

關鍵思路

(i) 中的四次方程是迴文（palindromic）的：它的係數由前往後讀和由後往前讀都一樣。除以 x^2 之後，x^2 與 x^{-2} 配成一對，x 與 x^{-1} 配成一對，而這兩對都可以用 $y = x + x^{-1}$ 表示，於是四次方程（quartic）便簡化成 y 的二次方程（quadratic）。(ii) 沒有提示，所以我們要找出同一種對稱性（symmetry）的變奏：只要容許加入一個縮放因子，係數 1, 1, −10, −4, 16 同樣可以配對，由此想到代換（substitution）$y = x - 4/x$。

解答

第 (i) 部分

常數項為 $1 \neq 0$，所以 $x = 0$ 不是根（root），我們可以除以 x^2：

$$x^2 + 10x + 26 + \frac{10}{x} + \frac{1}{x^2} = 0, \qquad \text{即} \qquad \left(x^2 + \frac{1}{x^2}\right) + 10\left(x + \frac{1}{x}\right) + 26 = 0.$$

設 $y = x + x^{-1}$，則 $y^2 = x^2 + 2 + x^{-2}$，所以 $x^2 + x^{-2} = y^2 - 2$。方程變成

$$y^2 - 2 + 10y + 26 = 0 \iff y^2 + 10y + 24 = 0 \iff (y+4)(y+6) = 0,$$

所以 $y = -4$ 或 $y = -6$。現在回到 x，把 $x + x^{-1} = y$ 兩邊乘以 x（因為 $x \neq 0$，這一步同樣沒有問題）：

$$x + \frac{1}{x} = -4 \iff x^2 + 4x + 1 = 0 \iff x = -2 \pm \sqrt{3},$$

$$x + \frac{1}{x} = -6 \iff x^2 + 6x + 1 = 0 \iff x = -3 \pm 2\sqrt{2}.$$

這是四個互不相同的數，而四次方程最多只有四個根，所以方程的解是

$$x = -2 \pm \sqrt{3}, \qquad x = -3 \pm 2\sqrt{2}.$$

留意這些根是以互為倒數（reciprocal）的形式成對出現的：$(-2+\sqrt{3})(-2-\sqrt{3}) = 1$ 和 $(-3+2\sqrt{2})(-3-2\sqrt{2}) = 1$。迴文多項式（polynomial）必然如此，因為把 x 換成 $1/x$ 再乘以 x^4，便會得回同一條方程。

第 (ii) 部分

想一想 係數 1, 1, −10, −4, 16 並不對稱。有沒有類似 (i) 的代換仍然行得通？

比較外側的兩對係數。常數項是首項係數（leading coefficient）的 $16 = 4^2$ 倍，而 x 的係數是 x^3 的係數的 -4 倍。因此各項可以這樣配對：

$$x^4 + 16 = x^2\left(x^2 + \frac{16}{x^2}\right), \qquad x^3 - 4x = x^2\left(x - \frac{4}{x}\right).$$

這就是把 $1/x$ 換成 $-4/x$ 之後的迴文規律。由於 $x = 0$ 不是根（常數項是 16），除以 x^2 可得

$$\left(x^2 + \frac{16}{x^2}\right) + \left(x - \frac{4}{x}\right) - 10 = 0.$$

設 $y = x - 4/x$。則 $y^2 = x^2 - 8 + 16/x^2$，所以 $x^2 + 16/x^2 = y^2 + 8$，而

$$y^2 + 8 + y - 10 = 0 \iff y^2 + y - 2 = 0 \iff (y-1)(y+2) = 0,$$

得 $y = 1$ 或 $y = -2$。把 $x - 4/x = y$ 乘以 x：

$$x - \frac{4}{x} = 1 \iff x^2 - x - 4 = 0 \iff x = \frac{1 \pm \sqrt{17}}{2},$$
$$x - \frac{4}{x} = -2 \iff x^2 + 2x - 4 = 0 \iff x = -1 \pm \sqrt{5}.$$

這同樣是四個互不相同的根，所以全部的根就是這些：

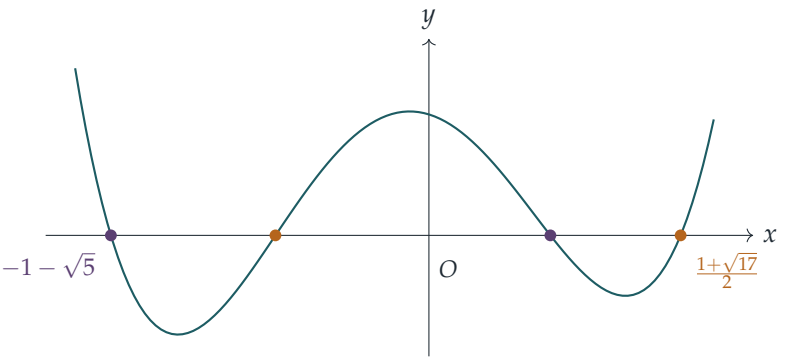
$$x = \frac{1 \pm \sqrt{17}}{2}, \qquad x = -1 \pm \sqrt{5}.$$

這次各對根的積是 -4 而不是 1：對稱性現在是 $x \mapsto -4/x$。

作為驗算，以上的運算相當於因式分解（factorisation）

$$x^4 + x^3 - 10x^2 - 4x + 16 = (x^2 - x - 4)(x^2 + 2x - 4),$$

你可以展開來確認：x^3 項給出 $2 - 1 = 1$，x^2 項給出 $-4 - 2 - 4 = -10$，而 x 項給出 $4 - 8 = -4$。



(ii) 中四次多項式的圖像與橫軸相交四次：琥珀色的根來自 $x^2 - x - 4 = 0$，紫紅色的根來自 $x^2 + 2x - 4 = 0$。

驗算

- 根的積：在 (i) 中，四個根相乘得 $1 \cdot 1 = 1$，即常數項；在 (ii) 中，它們相乘得 $(-4)(-4) = 16$，同樣是常數項。
- 根的和：在 (i) 中，$(-4) + (-6) = -10$，即 x^3 係數的相反數；在 (ii) 中，$1 + (-2) = -1$，情況相同。
- 數值上，$x = -2 + \sqrt{3} \approx -0.268$ 和 $x = -3 - 2\sqrt{2} \approx -5.828$ 都令 (i) 中的四次多項式的值為零，誤差在 10^{-9} 之內；而 $x = \frac{1}{2}(1 + \sqrt{17}) \approx 2.562$ 和 $x = -1 - \sqrt{5} \approx -3.236$ 在 (ii) 中亦然。

常見錯誤

- 忘記交叉項(cross term)：$x^2 + x^{-2} = y^2 - 2$，而不是 y^2；在 (ii) 中則是 $x^2 + 16x^{-2} = y^2 + 8$，這裏是加號，因為 $(x - 4/x)^2$ 的交叉項是 -8。
- 在 (ii) 中嘗試 $y = x + 4/x$。這樣可得 $x^2 + 16/x^2 = y^2 - 8$，卻無法得出 $x^3 - 4x$，因為它的正負號要求用 $x - 4/x$。
- 求出 y 的值便停下來。題目要求的是 x，而每個 y 的值對應兩個 x 的值。

STEP 之外

這個技巧之所以行得通，是因為這些多項式在某個代換之下保持不變(invariant)：(i) 中是 $x \mapsto 1/x$，(ii) 中是 $x \mapsto -4/x$。y 這個量在該代換之下不變，而任何在該代換下不變

的多項式表達式，都可以用 y 來表示。這是對伽羅瓦理論（Galois theory）的初步一瞥：在這套理論中，多項式各根之間的對稱性決定了對應的方程可以怎樣求解。Allenby 的 *Rings, Fields and Groups*（環、域與羣）是了解這些想法如何發展下去的好去處。

詳細解答

STEP III 1994，第 3 題

純數學 • 球面、平面與扣連

STEP III 1994，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/94-s3-q3

關鍵思路

每個圓各自位於一個平面上，而這兩個平面相交於直線 L。因此 C_1 能夠穿過 C_2 所在平面的地方，就只有它在 L 上的兩個點，反過來也是一樣。於是，兩個圓是否互相扣連，歸根究底是一個一維的問題：L 上的四個點沿着這條直線是怎樣排列的。

解答

平面與球面

設一個球面（sphere）的球心為 Q、半徑為 r，並設 d 為 Q 到平面 Π 的距離。設 F 為由 Q 到 Π 的垂足（foot of the perpendicular）。對於 Π 上的一點 X，由畢氏定理（Pythagoras）得 $QX^2 = d^2 + FX^2$，所以 X 在球面上當且僅當 $FX^2 = r^2 - d^2$。因此：

- 若 $d > r$，平面與球面不相交；
- 若 $d = r$，平面與球面相切於唯一的一點 F（它是一個切平面（tangent plane））；
- 若 $d < r$，交集是 Π 上的一個圓，圓心為 F，半徑為 $\sqrt{r^2 - d^2}$。當 $d = 0$ 時，這是一個大圓（great circle），半徑為 r。

已知資料

配方（completing the square）後，S_2 為

$$x^2 + (y-3)^2 + (z-2)^2 = 3,$$

球心為 $(0,3,2)$，半徑為 $\sqrt{3}$。球面 S_1 的球心為 O，半徑為 $\sqrt{7}$。

O 到 P_1 的距離為 $\frac{|-1|}{\sqrt{10}} = \frac{1}{\sqrt{10}} < \sqrt{7}$，所以 C_1 確實是一個圓，半徑為 $\sqrt{7 - \frac{1}{10}}$。$(0,3,2)$ 到 P_2 的距離為 $\frac{|0-3+1|}{\sqrt{2}} = \sqrt{2} < \sqrt{3}$，所以 C_2 是一個半徑為 $\sqrt{3-2} = 1$ 的圓。它的圓心就是垂足 $(0,3,2) + (1,-1,0) = (1,2,2)$；驗證一下：$1 - 2 + 1 = 0$。

直線 L

把兩個平面方程相減，得 $2x - 2 = 0$，所以 $x = 1$，繼而 $y = x + 1 = 2$，而 z 可取任意值。所以

$$L = \{(1,2,t) : t \in \mathbb{R}\},$$

即通過 $(1,2,0)$ 的鉛垂線。

在 S_1 上：$1 + 4 + t^2 = 7$，所以 $t = \pm\sqrt{2}$。因此 L 與 S_1 相交於

$$(1,2,\sqrt{2}) \quad \text{和} \quad (1,2,-\sqrt{2}).$$

在 S_2 上：$1 + 1 + (t-2)^2 = 3$，所以 $t - 2 = \pm 1$。因此 L 與 S_2 相交於

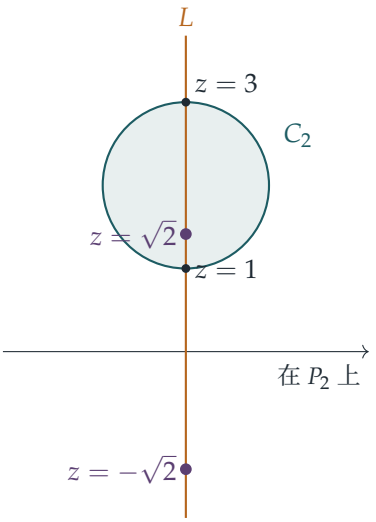
$$(1,2,1) \quad \text{和} \quad (1,2,3).$$

由於 L 位於 P_1 上，L 在 S_1 上的點正好就是 L 在 C_1 上的點；同樣地，L 在 S_2 上的點就是 L 在 C_2 上的點。特別地，$(1,2,1)$ 和 $(1,2,3)$ 是 C_2 一條直徑的兩個端點，這與它的圓心 $(1,2,2)$ 和半徑 1 吻合。

兩個圓是否互相扣連？

首先，兩個圓並不相交。任何公共點都必定位於 $P_1 \cap P_2 = L$ 上，而 C_1 在 L 上的點（$z = \pm\sqrt{2}$）與 C_2 在 L 上的點（$z = 1,3$）並不相同。

設 D_2 為 P_2 上以 C_2 為邊界的平面圓盤（disc）。它與 L 的交集是由 $(1,2,1)$ 到 $(1,2,3)$ 的直徑，即滿足 $1 \le t \le 3$ 的點 $(1,2,t)$。



圖中所示是平面 P_2，當中 C_2 及其圓盤以陰影表示。紫色的點是 C_1 穿過 P_2 的位置。

想一想 為甚麼 C_1 在這兩個點真的是穿過 P_2，而不只是與它相切？

直線 L 把平面 P_1 分成兩個半平面（half-plane），而由於 P_1 沿着 L 橫越 P_2，這兩個半平面分別位於 P_2 的兩側。圓 C_1 與 L 相交於兩個不同的點，所以 L 把 C_1 切成兩段弧，每個半平面各有一段。因此 C_1 的一段弧完全位於 P_2 的一側，另一段弧則完全位於另一側，而 C_1 在 $(1, 2, \sqrt{2})$ 和 $(1, 2, -\sqrt{2})$ 穿過 P_2，此外再無其他穿越點。

現在 $1 < \sqrt{2} < 3$，所以在 $(1, 2, \sqrt{2})$ 的穿越點位於圓盤 D_2 之內。但 $-\sqrt{2} < 1$，所以在 $(1, 2, -\sqrt{2})$ 的穿越點位於圓盤之外。沿 C_1 繞一圈，我們恰好穿過 C_2 所張成的圓盤一次，然後在圓盤之外越過 P_2 返回。換言之，C_1 穿過 C_2 一次，就像鏈條上相扣的兩個環。

沿着 L，這四個點的次序是

$$-\sqrt{2} < 1 < \sqrt{2} < 3,$$

所以來自 C_1 的一對點與來自 C_2 的一對點互相交錯（interleave）。正正是這一點令 C_1 穿過 D_2 一次。（假如兩對點都沒有把對方分隔開，C_1 就會穿過 D_2 兩次或完全不穿過，兩個圓便不會互相扣連。）因此**圓 C_1 與 C_2 互相扣連**（linked）。

我們無法把它們拉開：任何把它們變成兩個相距甚遠的圓的連續移動，都必須把 C_1 穿過 C_2 的圓盤的次數由奇數變為偶數。穿越點總是成對地出現和消失，除非其中一個圓穿過另一個圓，而在不剪斷的情況下把它們分開，是不容許這樣做的。

驗算

把各點逐一代入：例如 $(1,2,1)$ 給出 $3-2-1=0$、$1-2+1=0$ 和 $1+4+1-12-4+10=0$。數值方面，我們把兩個圓參數化（parametrise），並計算高斯（Gauss）的環繞積分（linking integral）$\frac{1}{4\pi}\oint\oint \frac{(\mathbf{r}_1-\mathbf{r}_2)\cdot(\mathrm{d}\mathbf{r}_1\times\mathrm{d}\mathbf{r}_2)}{|\mathbf{r}_1-\mathbf{r}_2|^3}$，得到 1.002（不扣連的一對圓會給出 0）。我們亦求得兩個圓之間的最短距離約為 0.09，所以它們十分接近，但並不相碰。

常見錯誤

- 只檢查兩個圓是否相交。它們並不相交，但單憑這一點完全無法說明它們是否扣連。
- 忘記了 C_1 與 P_2 的穿越點必定位於 L 上，反而在別處尋找。
- 弄錯 S_2 的球心：$-6y$ 和 $-4z$ 兩項給出的球心是 $(0,3,2)$，而不是 $(0,-3,-2)$。

STEP 之外

兩條互不相交的閉曲線（closed curve）是否扣連，是由一個整數不變量（invariant）來量度的，稱為**環繞數**（linking number）。計算它的方法，可以是數算其中一條曲線穿過另一條曲線所張成的曲面的帶符號穿越次數（這正是我們所做的），也可以是使用「驗算」一節中高斯的二重積分（double integral）。環繞數是踏入紐結理論（knot theory）和代數拓撲（algebraic topology）的第一步。Colin Adams 的 *The Knot Book*（一本紐結理論入門書）是十分易讀的入門讀物。

詳細解答

STEP III 1994，第 4 題

純數學 • 常微分方程的奇解

STEP III 1994，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/94-s3-q4

本題在 *Beyond STEP* 第 8.1 節當解不唯一中有完整的引導式講解：https://beyondstep.keithmcnulty.org/zh-hant/ch08/non-uniqueness.html

關鍵思路

兩邊開平方：方程即 $\mathrm{d}y/\mathrm{d}x = \pm 2\sqrt{y}$，馬上可以分離變數（separation of variables）。更值得留意的是，每一條拋物線（parabola）$y = (x-c)^2$ 都滿足這個方程，所以方程的解不過是把 $y = x^2$ 沿 x 軸滑動所得的各個副本。一旦看出這幅圖像，其後各部分要問的，就是哪些副本通過哪些點，以及所有副本有甚麼共通之處。

解答

通解（general solution）

由於 $(\mathrm{d}y/\mathrm{d}x)^2 \geq 0$，任何解都有 $y \geq 0$。在 $y > 0$ 之處，可以寫成

$$\frac{\mathrm{d}y}{\mathrm{d}x} = \pm 2\sqrt{y} \quad \Longrightarrow \quad \frac{1}{2\sqrt{y}}\frac{\mathrm{d}y}{\mathrm{d}x} = \pm 1 \quad \Longrightarrow \quad \sqrt{y} = \pm(x-c)$$

其中 c 為常數，而無論取哪一個符號，平方後都得到

$$y = (x-c)^2.$$

值得直接驗證整條拋物線都是解，包括左右兩半和頂點（vertex）：$y' = 2(x-c)$，所以對每一個 x 都有 $(y')^2 = 4(x-c)^2 = 4y$。

? **想一想** 兩個符號得出同一條公式。那麼「±」到哪裏去了？

它藏在每條拋物線的兩半之中。在 $y=(x-c)^2$ 上，斜率（gradient）$2(x-c)$ 在頂點左方等於 $-2\sqrt{y}$，在右方等於 $+2\sqrt{y}$，所以單單一條拋物線已經用上兩個符號。到我們要令曲線通過某一點時，符號的選擇便以 c 的選擇這個形式重新出現。

通過 (a,b^2) 的解

我們需要 $(a-c)^2=b^2$，所以 $a-c=\pm b$，即 $c=a-b$ 或 $c=a+b$。兩個解是

$$y=(x-a+b)^2 \qquad 及 \qquad y=(x-a-b)^2.$$

兩者並不相同，因為 $b\neq 0$ 令 $a-b\neq a+b$。（假如 $b=0$，兩者便會合併成一條頂點在 $(a,0)$ 的拋物線；這時該點亦會位於解 $y=0$ 之上，這個解我們在最後會再遇到。）把 b 換成 $-b$ 只會把兩個解對調，這是理所當然的，因為點 (a,b^2) 並無改變。

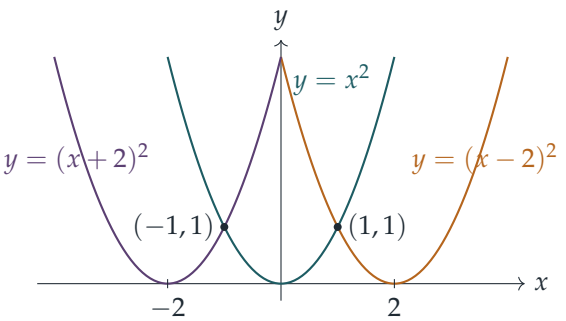
點 $(a_1,1)$ 和 $(a_2,1)$

取 $b=1$：通過 $(a,1)$ 的解是 $y=(x-a+1)^2$ 和 $y=(x-a-1)^2$。拋物線 $y=(x-c)^2$ 通過原點，當且僅當 $c=0$，所以我們需要 $a-1=0$ 或 $a+1=0$。因此

$$a_1=1, \qquad a_2=-1.$$

通過 $(1,1)$ 的解是 $y=x^2$（來自 $c=a_1-1=0$）和 $y=(x-2)^2$。通過 $(-1,1)$ 的解是 $y=(x+2)^2$ 和 $y=x^2$（來自 $c=a_2+1=0$）。

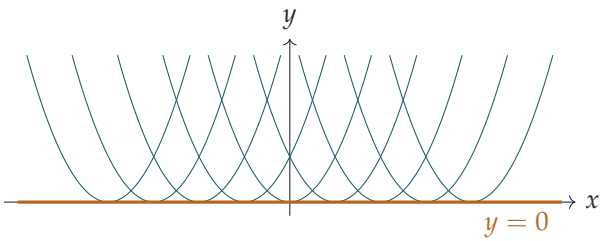
對這兩點來說，通過原點的解都是同一條曲線 $y=x^2$，所以它們的圖像重合。這並不令人意外：$(1,1)$ 和 $(-1,1)$ 正是 $y=x^2$ 上高度為 1 的兩點，分別位於頂點兩側。另外兩個解是 $y=(x-2)^2$ 和 $y=(x+2)^2$，即把 $y=x^2$ 分別向右和向左平移 2。



在 $(1,1)$，曲線 $y=x^2$ 的斜率是 $+2$，而 $y=(x-2)^2$ 的斜率是 -2：通過同一點的兩個解總是在該點相交，斜率為 $\pm 2b$，每個平方根各佔其一。

曲線族及其公切線（common tangent）

改變 a 和 b，只不過是令 $c = a \pm b$ 取遍所有實數，所以這個曲線族就是 $y = x^2$ 的所有水平平移（translate）。



每一條拋物線都在其頂點 $(c, 0)$ 與 x 軸相切，該處的斜率為 0。所以直線 $y = 0$ 是整個曲線族的公切線（它是這族曲線的*包絡線*（*envelope*））。

它之所以重要，是因為 $y = 0$ 本身也是一個解：這時 $y' = 0$，而 $(y')^2 = 0 = 4y$。但無論 c 取甚麼值，它都不是其中一條拋物線：它是一個*奇異解*（*singular solution*），分離變數時因為除以 $\sqrt{y}$ 而遺漏了它。在每一點 $(c, 0)$，解 $y = 0$ 與拋物線 $y = (x - c)^2$ 有相同的值和相同的斜率，所以通過 x 軸上各點的解並不唯一。事實上我們可以把它們拼接起來：對於 $p < q$，

$$y = \begin{cases} (x - p)^2, & x < p, \\ 0, & p \le x \le q, \\ (x - q)^2, & x > q \end{cases}$$

有連續的導數，並且處處滿足方程。所以嚴格來說，通過 (a, b^2) 的「兩個解」是指在該點附近與那兩條拋物線吻合的兩個解；離開該點之後，解一旦到達 x 軸，便可以沿着該軸走一段才離開。

驗算

- 代回原式：對於 $y = (x - a \pm b)^2$，在 $x = a$ 時得 $y = b^2$，而 $(y')^2 = 4y$ 恆成立。
- 在 (a, b^2) 的兩個斜率是 $2(a - c) = \pm 2b$，它們的平方都是 $4b^2 = 4y$，符合要求。
- 數值上，以中心差分（central difference）計算導數，確認了隨機選取的拋物線以及上述取 $p = -1$、$q = 1$ 的拼接曲線都滿足 $(y')^2 = 4y$，亦確認了通過 $(1, 1)$ 和 $(-1, 1)$ 的解分別有 $c \in \{0, 2\}$ 和 $c \in \{-2, 0\}$，兩者只共有 $c = 0$。

常見錯誤

- 只寫 $y' = 2\sqrt{y}$，結果通過每一點只得到半條拋物線。
- 忘記了 $y = 0$：它是一個真正的解，而且正正是最後一部分所要找的公切線。
- 以為通過 $(a_1, 1)$ 和 $(a_2, 1)$ 而又經過原點的兩個解是不同的曲線；其實兩者都是 $y = x^2$。

STEP 之外

這個方程是解的唯一性（uniqueness）失效的標準例子。皮卡–林德勒夫定理（Picard–Lindelöf theorem）保證，當 f 對 y 滿足利普希茨（Lipschitz）條件時，$y' = f(x, y)$ 通過一點的解恰好只有一個；但 $\pm 2\sqrt{y}$ 在 $y = 0$ 附近並不滿足利普希茨條件，而那裏正是拋物線與奇異解相遇之處。單參數曲線族的包絡線，以及它們為克萊羅（Clairaut）方程一類方程所帶來的奇異解，是常微分方程（ODE）入門課程中的經典課題。Boyce and DiPrima, *Elementary Differential Equations*（常微分方程入門教科書），以類似這個的例子討論了解的存在性和唯一性。

詳細解答

STEP III 1994，第 5 題

純數學 • 麥克勞林級數

STEP III 1994，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/94-s3-q5

關鍵思路

把 $\sin^{-1} x$ 求導一次會得到一個平方根，再求導一次，便得到 f'' 與 f' 之間一個不含平方根的關係式。然後利用萊布尼茲定理（Leibniz's theorem），就可以把這一條方程變成所有高階導數的遞推關係（recurrence）。在 $x = 0$ 處，大部分的項都會消失，於是得到麥克勞林係數（Maclaurin coefficients）的一個簡單遞推關係。至於 g，利用對數定律便可把一切化為我們熟悉的 $\ln(1 \pm x)$ 的級數。

解答

微分方程

對於 $-1 < x < 1$，

$$\mathrm{f}'(x) = \frac{1}{\sqrt{1-x^2}} = (1-x^2)^{-1/2}, \qquad \mathrm{f}''(x) = -\tfrac{1}{2}(1-x^2)^{-3/2} \cdot (-2x) = \frac{x}{(1-x^2)^{3/2}}.$$

把第二式乘以 $1 - x^2$，得 $(1-x^2)\mathrm{f}''(x) = x(1-x^2)^{-1/2} = x\mathrm{f}'(x)$，所以

$$(1-x^2)\mathrm{f}''(x) - x\mathrm{f}'(x) = 0,$$

這正是所需的結果。

遞推關係

利用萊布尼茲定理 $(uv)^{(n)} = \sum_r \binom{n}{r} u^{(r)} v^{(n-r)}$，把這個恆等式求導 n 次。在第一個乘積中，取 $u = 1 - x^2$。它的各階導數依次為 $-2x$、-2，之後全為 0，所以只有三項留下

來：

$$\frac{\mathrm{d}^n}{\mathrm{d}x^n}\left[(1-x^2)\mathrm{f}''\right] = (1-x^2)\mathrm{f}^{(n+2)} + n(-2x)\mathrm{f}^{(n+1)} + \binom{n}{2}(-2)\mathrm{f}^{(n)}.$$

在第二個乘積中 $u = x$，只有兩項留下來：

$$\frac{\mathrm{d}^n}{\mathrm{d}x^n}\left[x\mathrm{f}'\right] = x\mathrm{f}^{(n+1)} + n\mathrm{f}^{(n)}.$$

兩式相減，並利用 $2\binom{n}{2} = n(n-1)$，

$$\begin{aligned} 0 &= (1-x^2)\mathrm{f}^{(n+2)} - (2n+1)x\mathrm{f}^{(n+1)} - \big(n(n-1)+n\big)\mathrm{f}^{(n)} \\ &= (1-x^2)\mathrm{f}^{(n+2)} - (2n+1)x\mathrm{f}^{(n+1)} - n^2\mathrm{f}^{(n)}, \end{aligned}$$

這就是所需的結果。它對每個 $n \geq 1$ 都成立，而當 $n = 0$ 時，它就是最初的那條方程。（如果你喜歡，也可以改用歸納法（induction）來證明：把 n 的情形求導一次，再整理各項，便得到 $n+1$ 的情形。這其實是同一個計算，只是逐步進行而已。）

f 的麥克勞林級數

想一想 x 取哪個值，會令這個遞推關係容易使用？

代入 $x = 0$。中間的一項消失，而 $1 - x^2 = 1$，所以

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

我們由 $\mathrm{f}(0) = \sin^{-1} 0 = 0$ 和 $\mathrm{f}'(0) = 1$ 出發。

偶數階導數。由於 $\mathrm{f}(0) = 0$，根據遞推關係，每個偶數階導數在 0 處都等於零。這與我們的預期相符，因為 $\sin^{-1}$ 是奇函數（odd function）。

奇數階導數。取 $n = 1, 3, 5, \ldots$，

$$\mathrm{f}'''(0) = 1^2, \quad \mathrm{f}^{(5)}(0) = 3^2 \cdot 1^2, \quad \mathrm{f}^{(7)}(0) = 5^2 \cdot 3^2 \cdot 1^2,$$

而一般地

$$\mathrm{f}^{(2k+1)}(0) = \big(1 \cdot 3 \cdot 5 \cdots (2k-1)\big)^2 \qquad (k \geq 1).$$

因此

$$\sin^{-1} x = x + \frac{1^2}{3!}x^3 + \frac{1^2 \cdot 3^2}{5!}x^5 + \frac{1^2 \cdot 3^2 \cdot 5^2}{7!}x^7 + \cdots = x + \frac{x^3}{6} + \frac{3x^5}{40} + \frac{5x^7}{112} + \cdots.$$

x^{2k+1} 的係數可以化簡得很整齊。寫出 $(2k+1)! = \big(1 \cdot 3 \cdots (2k-1)\big)\big(2 \cdot 4 \cdots 2k\big)(2k+1)$，再約去一個奇數乘積：

$$a_{2k+1} = \frac{\big(1 \cdot 3 \cdots (2k-1)\big)^2}{(2k+1)!} = \frac{1 \cdot 3 \cdot 5 \cdots (2k-1)}{2 \cdot 4 \cdot 6 \cdots (2k)} \cdot \frac{1}{2k+1}.$$

由比值審斂法（ratio test）可知，這個級數在 $|x| < 1$ 時收斂，因為 $a_{2k+3}/a_{2k+1} \to 1$。至於它在該範圍內確實等於 $\sin^{-1} x$，則可以把 $(1-x^2)^{-1/2}$ 的二項級數（binomial series）逐項積分來確認，所得的係數完全相同。

函數 g

根據對數定律，$\mathrm{g}(x) = \frac{1}{2}\big(\ln(1+x) - \ln(1-x)\big)$。對於 $|x| < 1$，

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

兩者相減時，偶數次冪互相抵消，奇數次冪則變成兩倍。所以

$$\mathrm{g}(x) = x + \frac{x^3}{3} + \frac{x^5}{5} + \cdots = \sum_{k=0}^{\infty} \frac{x^{2k+1}}{2k+1}.$$

（你也許會認出 g 就是 $\tanh^{-1} x$。）

比較係數

> **想一想** 兩個係數都有因子 $\frac{1}{2k+1}$。那麼還剩下甚麼要比較？

對於 $n \geq 1$，x^{2n+1} 的係數在 g 中是 $\dfrac{1}{2n+1}$，而在 f 中是

$$\frac{1}{2} \cdot \frac{3}{4} \cdot \frac{5}{6} \cdots \frac{2n-1}{2n} \cdot \frac{1}{2n+1}$$

。這 n 個因子 $\frac{2r-1}{2r}$ 每一個都嚴格介乎 0 與 1 之間，而且因為 $n \geq 1$，這樣的因子至少有一個。所以它們的乘積嚴格小於 1，從而

$$\frac{1 \cdot 3 \cdots (2n-1)}{2 \cdot 4 \cdots (2n)} \cdot \frac{1}{2n+1} < \frac{1}{2n+1}.$$

因此 g 中的係數較大，這正是所需的結果。當 $n = 0$ 時，兩個係數都等於 1，這就是題目把這個情形排除在外的原因。

驗算

- 在 $x = -\frac{1}{2}$ 處，把級數求和至 x^{59} 一項，得 -0.5235987756，與 $\sin^{-1}(-\frac{1}{2}) = -\pi/6$ 吻合。
- 係數 $1, \frac{1}{6}, \frac{3}{40}, \frac{5}{112}$ 與直接由 $\mathrm{f}^{(n+2)}(0) = n^2 \mathrm{f}^{(n)}(0)$ 生成的係數一致。$\frac{1}{6} < \frac{1}{3}$、$\frac{3}{40} < \frac{1}{5}$、$\frac{5}{112} < \frac{1}{7}$ 這幾個比較亦如所述成立。
- 由於 f 的每個係數都不超過 g 中對應的係數，對於 $0 < x < 1$ 應有 $\sin^{-1} x < \tanh^{-1} x$。在 $x = 0.5$ 處：$0.5236 < 0.5493$。

常見錯誤

- 使用萊布尼茲定理時漏掉 $\binom{n}{2}$ 的一項，以致最後的係數變成 n 而不是 n^2。

- 寫成 $f^{(2k+1)}(0) = 1 \cdot 3 \cdots (2k-1)$ 而漏掉平方。遞推關係的每一步都是乘以 n^2，而不是 n。
- 計算麥克勞林係數時，只除以 $(2k)!$ 或 $(2k+1)$，而不是除以 $(2k+1)!$。

STEP 之外

比值 $\frac{1\cdot3\cdots(2n-1)}{2\cdot4\cdots(2n)}$ 等於 $\binom{2n}{n}/4^n$，而當 n 很大時，它的表現近似於 $1/\sqrt{\pi n}$。這是沃利斯乘積（Wallis's product）和斯特靈公式（Stirling's formula）的推論。$\sin^{-1} x$ 的級數在 $x = 1$ 處同樣收斂，其和為 $\pi/2$，但要證明這一點，需要用到阿貝爾定理（Abel's theorem），該定理處理冪級數（power series）在收斂區間（interval of convergence）端點的情況。當一個函數的各階導數滿足某條微分方程時，便可推出其係數的遞推關係；這就是常微分方程（ODE）的級數解法，Boyce 與 DiPrima 的 *Elementary Differential Equations*（微分方程入門教材）對此有所講解。至於泰勒級數（Taylor series）在甚麼情況下真正等於原來的函數，嚴謹的證明可參閱 Abbott 的 *Understanding Analysis*（數學分析入門教材）。

詳細解答

STEP III 1994，第 6 題

純數學 • 阿根圖上的旋轉

> **STEP III 1994，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/94-s3-q6

關鍵思路

以點 p 為中心、轉角為 θ 的旋轉（rotation）就是映射 $z \mapsto p + \mathrm{e}^{\mathrm{i}\theta}(z-p)$：先把 p 移到原點，旋轉，再移回去。把若干個這樣的映射複合（compose）起來，結果必定形如 $z \mapsto \lambda z + \mu$，而「$P_{\text{最後}} = P_1$ 對每一個 P_1 都成立」就表示這個映射是恆等映射（identity），所以 $\lambda = 1$ 且 $\mu = 0$。餘下的工作，就是從幾何角度解讀這兩條等式。

解答

求 z_5

正方向的四分之一圈旋轉相當於乘以 i，所以把 P_1 繞 A 旋轉便得

$$z_2 = a + \mathrm{i}(z_1 - a) = (1-\mathrm{i})a + \mathrm{i}z_1.$$

同理，$z_3 = (1-\mathrm{i})b + \mathrm{i}z_2$，$z_4 = (1-\mathrm{i})c + \mathrm{i}z_3$，以及 $z_5 = (1-\mathrm{i})d + \mathrm{i}z_4$。逐步代回去，

$$\begin{aligned} z_5 &= (1-\mathrm{i})\big(d + \mathrm{i}c + \mathrm{i}^2 b + \mathrm{i}^3 a\big) + \mathrm{i}^4 z_1 \\ &= (1-\mathrm{i})\big(d - b + \mathrm{i}(c-a)\big) + z_1. \end{aligned}$$

P_5 何時回到 P_1？

四個四分之一圈合起來是一整圈，所以 z_1 的係數恰好是 1，而

$$z_5 - z_1 = (1-\mathrm{i})\big(d - b + \mathrm{i}(c-a)\big),$$

這完全與 z_1 無關。複合映射是一個純粹的平移（translation）。由於 $1-\mathrm{i}\neq 0$，

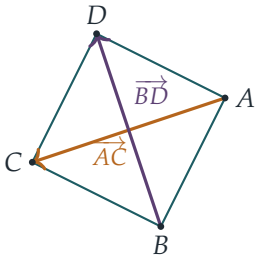
$$\begin{aligned} z_5 = z_1 \text{ 對所有} z_1 &\iff d-b+\mathrm{i}(c-a)=0 \\ &\iff \mathrm{i}(c-a)=b-d \\ &\iff a-c=\mathrm{i}(b-d), \end{aligned}$$

其中最後一步是把兩邊同乘以 $-\mathrm{i}$（而 $-\mathrm{i}\cdot\mathrm{i}=1$）。每一步都可逆，所以這就是所要求的「若且僅若」（if and only if）。（如果條件不成立，平移量便不是零，於是無論怎樣選取 P_1，都有 $P_5\neq P_1$。）

想一想 $a-c=\mathrm{i}(b-d)$ 說明了圖形 $ABCD$ 有甚麼特點？

乘以 -1：條件變成 $c-a=\mathrm{i}(d-b)$。這裏 $c-a$ 代表向量 $\overrightarrow{AC}$，而 $d-b$ 代表向量 $\overrightarrow{BD}$。乘以 i 會把向量旋轉 $\pi/2$ 而不改變其長度。所以：

$\overrightarrow{AC}$ 是 $\overrightarrow{BD}$ 沿正方向轉過一個直角所得；也就是說，四邊形 $ABCD$ 的對角線（diagonal）AC 與 BD 長度相等且互相垂直，而 AC 是由 BD 逆時針轉四分之一圈而得。



在圖中的例子裏，$\overrightarrow{BD}=-1+3\mathrm{i}$，而 $\overrightarrow{AC}=-3-\mathrm{i}=\mathrm{i}(-1+3\mathrm{i})$。定向（orientation）是有關係的：頂點按順時針次序標記的正方形滿足這個條件，但按逆時針次序標記的則不滿足（那時 $\overrightarrow{AC}=-\mathrm{i}\,\overrightarrow{BD}$）。

三次轉角為 θ 的旋轉

設 $w=\mathrm{e}^{\mathrm{i}\theta}$。與之前完全一樣，每一步都是 $z\mapsto(1-w)p+wz$，所以

$$\begin{aligned} z_4 &= (1-w)c+w\big((1-w)b+w\big((1-w)a+wz_1\big)\big) \\ &= (1-w)\big(c+wb+w^2a\big)+w^3z_1. \end{aligned}$$

這就是 $\lambda z_1+\mu$，其中 $\lambda=w^3$，$\mu=(1-w)(c+wb+w^2a)$。它對所有 z_1 都等於 z_1，若且僅若 $\lambda=1$ 且 $\mu=0$（取 $z_1=0$ 得 $\mu=0$，再取 $z_1=1$ 得 $\lambda=1$）。

轉角。$w^3=1$ 表示 3θ 是 2π 的倍數。我們把 $\theta\neq 0$ 理解為「不是整數圈」，所以 $w\neq 1$；因此 w 是一個非實數的三次單位根（cube root of unity），而

$$\theta=\pm\tfrac{2\pi}{3}\pmod{2\pi}.$$

各點。由於 $w \neq 1$，$\mu = 0$ 變成 $c + wb + w^2 a = 0$。乘以 w 並利用 $w^3 = 1$，得 $a + w^2 b + wc = 0$；又因為 $1 + w + w^2 = 0$（因 $w^3 = 1$，$w \neq 1$），我們可以改寫 $a = -(w + w^2)a$，所以

$$w(c-a) + w^2(b-a) = 0 \quad \Longrightarrow \quad c - a = -w\,(b-a) = \mathrm{e}^{\mathrm{i}(\theta+\pi)}(b-a) = \mathrm{e}^{\mathrm{i}(\theta-\pi)}(b-a).$$

? **想一想** 當 $\theta = \pm 2\pi/3$ 時，$\mathrm{e}^{\mathrm{i}(\theta-\pi)}$ 是甚麼？

當 $\theta = 2\pi/3$ 時，它是 $\mathrm{e}^{-\mathrm{i}\pi/3}$；當 $\theta = -2\pi/3$ 時，它是 $\mathrm{e}^{\mathrm{i}\pi/3}$（因為 $\mathrm{e}^{-5\pi\mathrm{i}/3} = \mathrm{e}^{\mathrm{i}\pi/3}$）。所以 $\overrightarrow{AC}$ 是 $\overrightarrow{AB}$ 轉過 $\mp\pi/3$ 所得，長度不變：ABC 是**等邊**（equilateral）三角形，當 $\theta = 2\pi/3$ 時 A, B, C 按**順時針**次序排列，當 $\theta = -2\pi/3$ 時則按**逆時針**次序排列。反過來，以上每一步都可逆，所以這些條件也能推出 $\mu = 0$。

由於 A, B, C 位於單位圓上，這個圓就是三角形的外接圓（circumcircle），而等邊三角形的外心（circumcentre）就是它的中心，所以三個頂點平均分佈在 $|z| = 1$ 上。兩種情況可以簡潔地合併成一句：

$$\theta = \pm\tfrac{2\pi}{3} \pmod{2\pi}, \qquad b = a\,\mathrm{e}^{-\mathrm{i}\theta}, \quad c = b\,\mathrm{e}^{-\mathrm{i}\theta}.$$

用文字來說，三次旋轉各轉三分之一圈，而 A, B, C 平均分佈在單位圓上，其排列方向與旋轉方向相反。

驗算

- 用電腦隨機選取 A, B, C, D，再對隨機選取的點施行四次四分之一圈旋轉，結果與 z_5 的公式吻合，誤差在 10^{-9} 之內；取 $c = a - \mathrm{i}(b-d)$ 時，每個 P_1 都回到原位，而把 c 稍微移動 0.1 之後便不再成立。
- 在單位圓上隨機選取 a，並取 $\theta = \pm 2\pi/3$，則 $b = a\mathrm{e}^{-\mathrm{i}\theta}$、$c = b\mathrm{e}^{-\mathrm{i}\theta}$ 對隨機的 z_1 都給出 $z_4 = z_1$；把標記次序倒轉或改用其他 θ 則不成立。
- 手算：把 $z_1 = 0$ 依次繞順時針標記的單位正方形 $a = 0$、$b = \mathrm{i}$、$c = 1 + \mathrm{i}$、$d = 1$ 旋轉。你會得到 $0 \to 0 \to 1 + \mathrm{i} \to 1 + \mathrm{i} \to 0$。回到原位了。

常見錯誤

- 把繞 A 的旋轉寫成 $z \mapsto \mathrm{i}z$，忘記了要先把 A 移到原點。
- 忽略定向：對角線之間的旋轉方向，以及 $\theta = \pm 2\pi/3$ 兩種情況下三角形各自的排列方向，都是條件的一部分。

STEP 之外

滿足 $|\lambda| = 1$ 的映射 $z \mapsto \lambda z + \mu$ 就是平面上保持定向的等距變換（isometry），它們構成一個羣（group）。轉角可以相加，所以若干個旋轉的複合是一個轉角為各角總和的旋轉；若總和是整數圈，則是一個平移：這就是 $z_5 - z_1$ 結果是常數的原因。Needham 的 *Visual Complex Analysis* 把複數的這種幾何觀點闡述得十分精彩。

詳細解答

STEP III 1994，第 7 題

純數學 • 群同構

> **STEP III 1994，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/94-s3-q7

關鍵思路

兩個羣（group）要同構（isomorphic），就必須共有同構映射（isomorphism）所保持的每一項結構性質。其中最有用的兩項，是各元素的階（order）的清單，以及該羣是否可交換（commutative，即阿貝爾羣，abelian）。只要列出各元素的階，已足以把 S_3 和 Z_6 區分開來。要證明兩個羣確實同構，我們就寫出一個明確的雙射（bijection），並驗證它與運算相容。最後一部分歸結為交換性：凡是由複數組成的羣都是阿貝爾羣，而 S_3 不是。

解答

S_3 和 Z_6 的元素

把三件物件標記為 $1,2,3$，並採用循環記法（cycle notation），由右至左複合，即 $\sigma\tau$ 表示「先做 τ，再做 σ」。六個置換（permutation）列於下表。

S_3 的元素	e	(12)	(13)	(23)	(123)	(132)
階	1	2	2	2	3	3

對換（transposition）把兩件物件互換，做兩次便一切還原，所以它的階是 2。像 (123) 這樣的 3-循環把 $1 \to 2 \to 3 \to 1$，要施行三次才能使每件物件回到原位，所以它的階是 3。

在 $Z_6 = \{0,1,2,3,4,5\}$ 中，單位元（identity）是 0，而 a 的階是滿足 $na \equiv 0 \pmod 6$ 的最小 $n \geq 1$，即 $6/\gcd(a,6)$。

Z_6 的元素	0	1	2	3	4	5
階	1	6	3	2	3	6

例如，$4+4=8\equiv 2$ 而 $4+4+4=12\equiv 0$，所以 4 的階是 3。

S_3 和 Z_6 同構嗎？

不同構。同構映射 θ 保持元素的階：若 g 的階是 n，則 $\theta(g)^k=\theta(g^k)$，而由於 θ 是單射（injective），這個元素是單位元當且僅當 g^k 是單位元。因此，同構的羣中，每一種階的元素數目都相同。但 Z_6 有兩個階為 6 的元素，S_3 卻一個也沒有，所以兩者並不同構。（兩者在階為 2 的元素上也有分別：S_3 有三個，Z_6 只有一個。）

還有第二個理由，我們在最後會再用到。Z_6 是阿貝爾羣，因為加法可交換。S_3 則不是：

$$(12)(13)=(132),\qquad (13)(12)=(123).$$

就第一個乘積而言，(13) 把 $1\to 3$，然後 (12) 保持 3 不變，所以 $1\to 3$。同理 $3\to 2$ 及 $2\to 1$。同構映射把可交換的一對元素映到可交換的一對元素，所以非阿貝爾羣不可能與阿貝爾羣同構。

C_6 與 Z_6 同構

定義 $\phi: Z_6\to C_6$ 為 $\phi(k)=\alpha^k$，其中 $k=0,1,\dots,5$。

- 雙射。當 $k=0,\dots,5$ 時，各數 $\alpha^k=\mathrm{e}^{\mathrm{i}k\pi/3}$ 的輻角（argument）為 $0,\frac{\pi}{3},\dots,\frac{5\pi}{3}$。這些是 $[0,2\pi)$ 內互不相同的角，所以這六個值是單位圓上互不相同的點。因此 ϕ 是單射，從而是兩個大小同為 6 的集合之間的雙射。
- 運算。設 $a,b\in Z_6$，並寫 $a+b=6q+r$，其中 $r=(a+b)\bmod 6$。由於 $\alpha^6=\mathrm{e}^{2\pi\mathrm{i}}=1$，

$$\phi(a)\,\phi(b)=\alpha^a\alpha^b=\alpha^{a+b}=(\alpha^6)^q\alpha^r=\alpha^r=\phi\big((a+b)\bmod 6\big).$$

所以 ϕ 把模 6 加法變成複數乘法，它是一個同構映射。因此 $C_6\cong Z_6$。作為核對，α 和 α^5（即 1 和 5 的像）的階是 6，α^2 和 α^4 的階是 3，而 $\alpha^3=-1$ 的階是 2，與 Z_6 的表吻合。

$\mathbb{C}$ 有與 S_3 同構的子羣嗎？

> **想一想** 所有由複數組成的羣都共有甚麼性質？

它們全都是阿貝爾羣。$(\mathbb{C},+)$ 的子羣（subgroup）H 是阿貝爾羣，因為對所有複數 z,w 都有 $z+w=w+z$。非零複數在乘法下所成的羣的子羣 H 是阿貝爾羣，因為 $zw=wz$。子羣沿用大羣的同一運算，所以繼承了交換性。

現在假設某個這樣的 H 有同構映射 $\theta : S_3 \to H$。把它應用於上面那對不可交換的元素。由於 H 是阿貝爾羣，

$$\theta\big((12)(13)\big) = \theta(12)\,\theta(13) = \theta(13)\,\theta(12) = \theta\big((13)(12)\big),$$

這裏以乘法形式書寫（若是加法的情況，把乘積換成和即可）。因為 θ 是單射，由此會得出 $(12)(13) = (13)(12)$，而這是錯的。所以複數的子羣，不論是加法的還是乘法的，都**沒有**一個與 S_3 同構。

驗算

- 在每個羣中，元素的階都整除 6（拉格朗日定理，Lagrange's theorem）：S_3 中是 $1, 2, 3$，Z_6 中是 $1, 2, 3, 6$。
- 我們用電腦重新計算了 S_3、Z_6 和 C_6 中每個元素的階，確認了 $(12)(13) \neq (13)(12)$，並對全部 36 對 (a, b) 以數值方法檢驗了 $\alpha^{(a+b) \bmod 6} = \alpha^a \alpha^b$，誤差在 10^{-12} 以內。

常見錯誤

- 說「兩者都有六個元素，所以同構」。大小相同是必要條件，但遠遠不是充分條件。
- 說 S_3 中階為 3 的元素是 (12) 和 (13)。對換的階是 2；只有 3-循環的階才是 3。
- 在最後一部分，只排除了有限的乘法子羣（單位根，roots of unity）。交換性的論證一次過涵蓋所有子羣，不論有限還是無限、加法還是乘法。

STEP 之外

這道題講的是不變量（invariant）：同構的羣必須共有的性質，例如元素的階和交換性。事實上，每個階為 6 的羣都與 Z_6 或 S_3 其中之一同構。更一般地，非零複數在乘法下所成的羣的任何有限子羣都是循環羣（cyclic group），所以它是一個由單位根組成的羣。這兩個結果，你在第一門羣論課程的初段就會遇到。Allenby 的 *Rings, Fields and Groups*（環、域與羣）寫得平易近人，很適合用來閱讀這些內容。

詳細解答

STEP III 1994，第 8 題

純數學 • 四平方和恆等式

STEP III 1994，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/94-s3-q8

本題在 *Beyond STEP* 第 7.8 節四個平方數中有完整的引導式講解：https://beyondstep.keithmcnulty.org/zh-hant/ch07/four-squares.html

關鍵思路

形如 $\begin{pmatrix} z & w \\ -w^* & z^* \end{pmatrix}$ 的矩陣(matrix)，其行列式(determinant)為 $zz^* + ww^* = |z|^2 + |w|^2$，是四個實數平方之和。若把 a, b, c, d 裝進一個這樣的矩陣，再把 p, q, r, s 裝進另一個，則由 $\det(AB) = \det A \det B$ 便得到左邊。要做的工夫是證明乘積 AB 具有同樣的特殊形式，因為這樣它的行列式也是四個平方之和，而這些平方正是 L_i^2。

解答

準備工作

對複數（complex number）z 和 w，記

$$M(z,w) = \begin{pmatrix} z & w \\ -w^* & z^* \end{pmatrix}, \qquad \det M(z,w) = zz^* - w(-w^*) = |z|^2 + |w|^2.$$

取

$$z_1 = a + b\mathrm{i}, \quad z_2 = c + d\mathrm{i}, \quad z_3 = p + q\mathrm{i}, \quad z_4 = r + s\mathrm{i}.$$

則

$$\det M(z_1, z_2) = a^2 + b^2 + c^2 + d^2, \qquad \det M(z_3, z_4) = p^2 + q^2 + r^2 + s^2.$$

乘積具有相同的形式

乘開得

$$M(z_1, z_2)\, M(z_3, z_4) = \begin{pmatrix} z_1 z_3 - z_2 z_4^* & z_1 z_4 + z_2 z_3^* \\ -z_2^* z_3 - z_1^* z_4^* & -z_2^* z_4 + z_1^* z_3^* \end{pmatrix}.$$

設 $w_1 = z_1 z_3 - z_2 z_4^*$ 及 $w_2 = z_1 z_4 + z_2 z_3^*$，即第一行的兩個元素。由於乘積的共軛（conjugate）等於共軛的乘積，而且 $(z^*)^* = z$，

$$w_1^* = z_1^* z_3^* - z_2^* z_4, \qquad -w_2^* = -z_1^* z_4^* - z_2^* z_3,$$

而這兩個正好就是右下角和左下角的元素。所以

$$M(z_1, z_2)\, M(z_3, z_4) = M(w_1, w_2).$$

想一想 乘積具有這個特殊形狀，為甚麼這樣重要？

對一般的 2×2 複矩陣而言，行列式根本不是平方和。正因為乘積仍然是 $M(w_1, w_2)$ 的形式，它的行列式才是 $|w_1|^2 + |w_2|^2$，從而可拆成四個實數平方。

取行列式

利用 $\det(AB) = \det A \, \det B$，

$$(a^2 + b^2 + c^2 + d^2)(p^2 + q^2 + r^2 + s^2) = \det M(w_1, w_2) = |w_1|^2 + |w_2|^2.$$

餘下只須求出 w_1 和 w_2 的實部（real part）和虛部（imaginary part）：

$$\begin{aligned} w_1 &= (a + b\mathrm{i})(p + q\mathrm{i}) - (c + d\mathrm{i})(r - s\mathrm{i}) \\ &= \big[(ap - bq) + (aq + bp)\mathrm{i}\big] - \big[(cr + ds) + (dr - cs)\mathrm{i}\big] \\ &= (ap - bq - cr - ds) + (aq + bp + cs - dr)\mathrm{i}, \end{aligned}$$

$$\begin{aligned} w_2 &= (a + b\mathrm{i})(r + s\mathrm{i}) + (c + d\mathrm{i})(p - q\mathrm{i}) \\ &= \big[(ar - bs) + (as + br)\mathrm{i}\big] + \big[(cp + dq) + (dp - cq)\mathrm{i}\big] \\ &= (ar - bs + cp + dq) + (as + br - cq + dp)\mathrm{i}. \end{aligned}$$

因此可取

$$L_1 = ap - bq - cr - ds, \qquad L_2 = aq + bp + cs - dr,$$
$$L_3 = ar - bs + cp + dq, \qquad L_4 = as + br - cq + dp,$$

於是 $|w_1|^2 + |w_2|^2 = L_1^2 + L_2^2 + L_3^2 + L_4^2$，得出

$$(a^2 + b^2 + c^2 + d^2)(p^2 + q^2 + r^2 + s^2) = L_1^2 + L_2^2 + L_3^2 + L_4^2.$$

每個 L_i 都是若干項之和，每項的形式為（a, b, c, d 中的一個字母）$\times$（p, q, r, s 中的一個字母），所以它對兩組變數分別都是線性（linear）的，符合題目要求。（答案並不唯一：改變任何 L_i 的正負號，或以不同方式選取 z_j，都會得出其他合用的組合。）

驗算

- 取 $(a, b, c, d) = (1, 2, 3, 4)$ 及 $(p, q, r, s) = (5, 6, 7, 8)$。左邊為 $30 \times 174 = 5220$。由公式得 $(L_1, L_2, L_3, L_4) = (-60, 12, 30, 24)$，而 $3600 + 144 + 900 + 576 = 5220$。
- 設 $c = d = r = s = 0$。則 $L_3 = L_4 = 0$，恆等式(identity)簡化為 $(a^2 + b^2)(p^2 + q^2) = (ap - bq)^2 + (aq + bp)^2$，即由 $|z_1 z_3| = |z_1||z_3|$ 得出的二平方和恆等式。
- 電腦驗算已就八個變數的 20,000 組隨機有理數取值，確認恆等式精確成立，並以數值方式確認了 $M(z_1, z_2)M(z_3, z_4) = M(w_1, w_2)$。

常見錯誤

- 相乘時忘記第二行的共軛，以致破壞了乘積的特殊形式。
- $(c + d\mathrm{i})(r - s\mathrm{i})$ 中的正負號出錯：$-ds\,\mathrm{i}^2$ 一項會變成 $+ds$。

STEP 之外

這就是歐拉(Euler)四平方和恆等式。矩陣 $M(z, w)$ 是哈密頓(Hamilton)四元數(quaternion)的一個具體模型：它們在乘法下封閉（closed），而行列式就是「長度」的平方，具有可乘性(multiplicative)。這個恆等式是拉格朗日(Lagrange)定理(每個正整數都是四個平方之和)證明中的關鍵一步，因為它把問題歸結為質數(prime)的情況。Davenport 的 *The Higher Arithmetic*（一本數論入門經典）涵蓋平方和的課題，而 Needham 的 *Visual Complex Analysis*（以幾何觀點講解複分析）則為二平方和的情況提供幾何直觀。

詳細解答

STEP III 1994，第 9 題

力學 • 等時曲線與碰撞

> **STEP III 1994，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/94-s3-q9

本題在 *Beyond STEP* 第 11.4 節等時曲線中有完整的引導式講解：https://beyondstep.keithmcnulty.org/zh-hant/ch11/tautochrone.html

關鍵思路

沿曲面的能量守恆（energy conservation）給出一個首次積分（first integral），而碗的形狀 $y = s^2/(4k)$ 把它化為以弧長（arc-length）s 表示的簡諧運動（simple harmonic motion）能量方程。簡諧運動的週期與振幅無關，所以每個由靜止釋放的質點，無論從哪裏出發，都在相同的時間後到達 V。正因如此，最後一部分的兩個質點必定在 V 相遇，而且一次又一次地相遇。

解答

常數

我們把質點看作光滑曲面上的一個質點模型（point mass）。作用在它身上的力有兩個：重量和法向反作用力（normal reaction）。質點由靜止出發，所以它的速度一開始並沒有垂直於「通過起點和軸的鉛垂平面」的水平分量。兩個力都沒有垂直於該平面的分量（軸對稱曲面的反作用力位於通過軸的平面之內），所以質點一直留在這個鉛垂截面內。它沿曲線 $s = 2\sqrt{ky}$ 運動，速率為 $|\mathrm{d}s/\mathrm{d}t|$。

曲面是光滑的，所以法向反作用力與速度垂直，不作功。只有重力作功，因此機械能守恆：

$$\tfrac{1}{2}m\left(\frac{\mathrm{d}s}{\mathrm{d}t}\right)^2 + mgy = \text{常數}.$$

乘以 $2/m$ 可見 $\left(\frac{\mathrm{d}s}{\mathrm{d}t}\right)^2 + 2gy$ 是常數，正如所求。質點在高度 h 由靜止釋放，故該常數為 $2gh$。

到達 V 所需的時間

在曲線上，$y = s^2/(4k)$。代入後得

$$\left(\frac{\mathrm{d}s}{\mathrm{d}t}\right)^2 + \frac{g}{2k}s^2 = 2gh.$$

對 t 求導，再除以 $2\,\mathrm{d}s/\mathrm{d}t$（質點在運動時此步成立），得

$$\frac{\mathrm{d}^2 s}{\mathrm{d}t^2} = -\frac{g}{2k}s,$$

這是關於 s 的簡諧運動，其中 $\omega^2 = \dfrac{g}{2k}$，中心在 $s = 0$，即點 V。

想一想 為甚麼由釋放到抵達 V 的時間是四分之一個週期？

質點由靜止從端點位置 $s_0 = 2\sqrt{kh}$ 出發，所以 $s = s_0 \cos \omega t$。當 $\omega t = \frac{\pi}{2}$ 時它首次到達 $s = 0$，這是週期 $2\pi/\omega$ 的四分之一。所以到達 V 所需的時間是

$$T = \frac{\pi}{2\omega} = \frac{\pi}{2}\sqrt{\frac{2k}{g}} = \pi\sqrt{\frac{k}{2g}},$$

正如所求。這個結果與 h 無關：這個碗是一條等時曲線（tautochrone）。

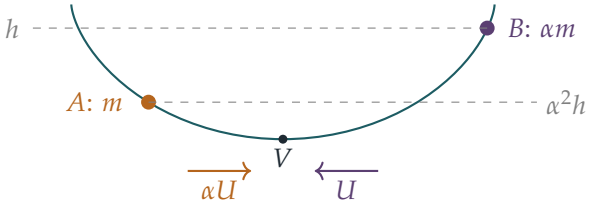
為方便稍後使用，質點在 V 的速率是 ωs_0；或者直接在能量方程中取 $y = 0$，得 $\sqrt{2gh}$。

兩個質點

把兩個質點看作在 V 正面碰撞的質點。該處的切線是水平的，所以碰撞沿一條水平線發生。把質量為 m 的質點稱為質點 A（在高度 $\alpha^2 h$ 釋放），把質量為 αm 的質點稱為質點 B（在另一側的高度 h 釋放）。

它們在 V 相遇。根據上一部分，每個質點不論起始高度如何，都在時間 T 後到達 V。它們同時釋放，所以同時到達 V，並在時間 T 於該處碰撞。

碰撞前一刻的速率。記 $U = \sqrt{2gh}$。由能量方程，A 到達時的速率是 $\sqrt{2g\alpha^2 h} = \alpha U$，$B$ 到達時的速率是 U，兩者運動方向相反。取 A 的運動方向為正。



碰撞前的動量是

$$m(\alpha U) + \alpha m(-U) = 0.$$

這就是題目如此選取質量和高度的原因。

第一次碰撞之後。設碰撞後的速度為 v_A 和 v_B。由動量守恆（conservation of momentum）和牛頓實驗定律（Newton's experimental law）得

$$mv_A + \alpha m\, v_B = 0, \qquad v_B - v_A = \alpha\big(\alpha U - (-U)\big) = \alpha(1+\alpha)U.$$

由第一條方程得 $v_B = -v_A/\alpha$。代入第二條方程得 $-v_A(1+\alpha)/\alpha = \alpha(1+\alpha)U$，所以

$$v_A = -\alpha^2 U, \qquad v_B = \alpha U.$$

兩個速度都反向，並乘以 α。（總動量為零時，相對速度和每個質點各自的速度全都乘以 $-\alpha$。）所以每個質點都沿碗中自己那一側往回上升。由能量守恆，A 升至高度 $\alpha^4 h$，B 升至高度 $\alpha^2 h$。

規律重複出現。兩個質點在同一時刻，即碰撞後時間 T，各自在新的最高點靜止下來。然後各自再用時間 T 返回 V。所以它們在時間 $3T$ 再次於 V 碰撞，此時 A 來自高度 $\alpha^4 h$，B 來自高度 $\alpha^2 h$。這正是把 h 換成 $\alpha^2 h$ 之後的初始狀況。總動量再次為零，兩個速度也再次反向並乘以 α。

由歸納法（induction），第 n 次碰撞在時間 $(2n-1)T$ 於 V 發生，而在這次碰撞之後

$$A \text{ 升至高度} \alpha^{2n+2}h, \qquad B \text{ 升至高度} \alpha^{2n}h,$$

各自留在碗中它出發的那一側。兩個質點在各自的一側振盪，並在等間隔的時刻 $T, 3T, 5T, \ldots$ 於 V 碰撞。每次碰撞時速率都按因子 α 減小，所以高度按因子 α^2 減小，動能損失的比例為 $1-\alpha^2$。碰撞永不停止，因為相鄰兩次碰撞的間隔總是 $2T$。振幅按等比方式遞減，當 $t \to \infty$ 時兩個質點都趨向靜止於 V，但不會在有限時間內真正靜止在那裏。

驗算

- 量綱（dimensions）：k 是長度（因為 $s^2 = 4ky$），所以 $\sqrt{k/g}$ 是時間。
- 數值方面，我們對隨機選取的 k 和 h 以數值求積法(quadrature)計算了 $\int_0^{s_0} \mathrm{d}s/\sqrt{2g(h-y)}$。取 $k=1$、$g=9.81$ 時，結果是 $0.7092\ldots = \pi\sqrt{1/19.62}$，與 h 無關。
- 我們以數值方法逐步計算兩個質點的運動（取 $\alpha = 0.6$ 和 $\alpha = 0.3$），並在每次相遇時應用碰撞定律。碰撞發生在 $t/T = 1, 3, 5, 7$，每次都在 V，而且總動量為零。取 $\alpha = 0.6$ 時，第一次碰撞後的高度是 $0.1296h = \alpha^4 h$ 和 $0.36h = \alpha^2 h$。
- 極限情況 $\alpha \to 1$：質量相等、高度相等，而且是完全彈性碰撞（perfectly elastic），所以每個質點都永遠反彈回到起始高度，與預期相符。

常見錯誤

- 以為 V 距離起點半個週期：質點從振盪的端點出發，所以到 V 只需四分之一個週期。
- 在牛頓定律中忘記方向反轉：要使用帶正負號的速度，而不只是速率。
- 聲稱質點在有限次碰撞後靜止下來。碰撞之間的時間間隔是固定的，並不會縮短（這與球在地板上彈跳不同），所以有無限多次碰撞，分佈在無限長的時間之內。

STEP 之外

內稟方程（intrinsic equation）$s^2 = 4ky$ 描述的是一條擺線（cycloid）。這道題重新發現了惠更斯（Huygens）的等時曲線：在倒置的擺線上，滑到最低點所需的時間與起點無關。惠更斯利用這一點設計出週期與擺幅無關的擺鐘。同一條曲線也是最速降線（brachistochrone），即下降最快的曲線，它可用變分法（calculus of variations）求得，而變分法是大學二年級力學的標準課題。想進一步了解振盪和碰撞，可參閱 Kleppner 與 Kolenkow 合著的 *An Introduction to Mechanics*（力學入門教科書）。

詳細解答

STEP III 1994，第 10 題

力學 • 相對速度

STEP III 1994，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：

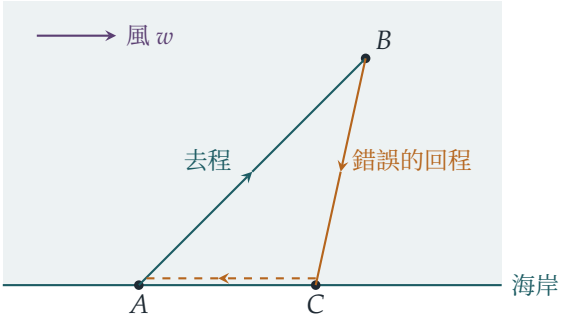
step.maths.org/questions/94-s3-q10

關鍵思路

一切都是向量加法（vector addition）：飛機相對地面的速度，等於它相對空氣的速度加上風的速度。把海岸放在 x 軸上，以 Alphaport 為原點，並把去程的空速（air velocity）寫成分量（components）。把航向（heading）反過來，空速會反向，但風不會反向，因此回程時，運動的向北部分和向東部分表現截然不同，而這正是第一部分的全部要點。至於第二部分，兩個正確的地速（ground speed）來自同一個二次方程（quadratic），於是韋達定理（Vieta's formulas）就能代勞。

解答

我們把飛機看成一個質點（particle），它相對空氣以大小為 v 的恆定速度運動，而空氣則以恆定速度 w 向正東移動（來自西面的風是吹向東面的）。取 Alphaport A 為原點，x 軸向東，y 軸向北，於是海岸就是直線 $y = 0$，而島嶼位於 $y > 0$ 的一側。Betaville B 位於 $\left(\frac{d}{\sqrt{2}}, \frac{d}{\sqrt{2}}\right)$。速率以 $\mathrm{km\,h^{-1}}$ 為單位，時間以小時為單位。（試卷上把這兩個地方稱為 Alphabet/Alphaport 和 Betaville/Betaport；它們指的是同樣的地方。）



去程

設去程的空速為 (a, b)，則 $a^2 + b^2 = v^2$。相對地面的速度是 $(a + w,\ b)$，它必須指向東北，所以兩個分量相等而且為正：

$$a + w = b > 0.$$

她以 y 方向的速率 b 飛過 y 方向的距離 $\frac{d}{\sqrt{2}}$，而去程所需時間為 t，所以

$$bt = \frac{d}{\sqrt{2}} = (a + w)t. \tag{1}$$

錯誤的回程

採用相反的航向時，她的空速是 $(-a, -b)$，但風仍然加上 $(w, 0)$，所以她相對地面的速度是

$$(w - a,\ -b).$$

? **想一想** 她會在哪裏到達海岸？在 Alphaport 的哪一邊？

向北方向的運動不受風影響：她的 y 坐標由 $y = \frac{d}{\sqrt{2}}$ 起以速率 b 遞減，所以由 (1) 可知，她恰好在時間 t 之後到達 $y = 0$，與去程所需的時間相同。在這段時間內，她的 x 坐標改變了 $(w - a)t$，所以她越過海岸的位置是

$$x = \frac{d}{\sqrt{2}} + (w - a)t = (a + w)t + (w - a)t = 2wt,$$

這裏再次用到 (1)。因此她降落在 C，位於 Alphaport 以東 $2wt$ 處：去程時，風把她推向東面，而她選定的航向正是為了抵銷這個影響；但回程時，航向所作的修正反了過來，風的推力卻沒有反過來，所以偏移加倍。由於 $2wt > 0$，沿海岸轉向西面的確是正確的方向。

她以空速 v 逆風向正西飛行，地速為 $v - w > 0$（因為 $v > w$）。沿海岸的一段需時 $\dfrac{2wt}{v - w}$，所以整個回程需時

$$t + \frac{2wt}{v - w} = \frac{(v - w) + 2w}{v - w}\, t = \left(\frac{v + w}{v - w}\right) t \text{ 小時,}$$

正如所求。

正確的回程

把沿直線 AB 的相對地面速度寫成 $\pm(p, p)$，其中 $\sqrt{2}\, p$ 是地速。空速等於相對地面的速度減去風速，而且大小為 v。

去程時，相對地面的速度為 (p_1, p_1)，空速為 $(p_1 - w,\ p_1)$：

$$(p_1 - w)^2 + p_1^2 = v^2 \iff 2p_1^2 - 2wp_1 + (w^2 - v^2) = 0.$$

回程時，相對地面的速度為 $(-p_2, -p_2)$，其中 $p_2 > 0$，空速為 $(-p_2 - w,\ -p_2)$：

$$(p_2 + w)^2 + p_2^2 = v^2 \iff 2(-p_2)^2 - 2w(-p_2) + (w^2 - v^2) = 0.$$

所以 p_1 和 $-p_2$ 都是同一個二次方程 $2p^2 - 2wp + (w^2 - v^2) = 0$ 的根。它的兩根之積為 $\frac{w^2 - v^2}{2} < 0$，所以一個根為正，一個根為負；p_1 是正的那個，$-p_2$ 是負的那個，因此它們就是兩個相異的根。根據韋達定理，

$$p_1 + (-p_2) = w, \qquad p_1(-p_2) = \frac{w^2 - v^2}{2},$$

即 $p_1 - p_2 = w$ 及 $p_1 p_2 = \frac{1}{2}(v^2 - w^2)$。

來回的距離都是 d，所以去程所需時間為 $t = \dfrac{d}{\sqrt{2}\,p_1}$，而正確回程所需時間為 $\dfrac{d}{\sqrt{2}\,p_2}$。因此

$$\begin{aligned}\frac{d}{\sqrt{2}\,p_2} &= t + \frac{d}{\sqrt{2}}\left(\frac{1}{p_2} - \frac{1}{p_1}\right) = t + \frac{d}{\sqrt{2}} \cdot \frac{p_1 - p_2}{p_1 p_2} \\ &= t + \frac{d}{\sqrt{2}} \cdot \frac{2w}{v^2 - w^2} = t + \frac{\sqrt{2}\,wd}{v^2 - w^2} \text{ 小時,}\end{aligned}$$

這就是所求的結果。

驗算

- 若 $w = 0$，兩個回程時間都化為 t，在無風的情況下理應如此。
- 取 $v = 300$、$w = 60$、$d = 400$：$t \approx 1.1785\,\text{h}$，錯誤的回程（以數值方法逐步模擬運動而求得）需時 $1.7678\,\text{h} = \frac{360}{240}t$，而正確的回程需時 $1.5713\,\text{h}$，與 $t + \frac{\sqrt{2}\cdot 60\cdot 400}{86400}$ 吻合。對許多組 $v > w$ 和 d 作隨機試算，結果在捨入誤差範圍內一致。
- 錯誤的回程總是比正確的回程慢：這就是不等式 $\frac{2wt}{v-w} > \frac{\sqrt{2}\,wd}{v^2-w^2}$，即 $\sqrt{2}(v+w)t > d$；它之所以成立，是因為去程的地速 $\frac{d}{t} = \sqrt{2}p_1 = \sqrt{2}(p_2 + w)$ 小於 $\sqrt{2}(v+w)$，原因是 $p_2 < v$。

常見錯誤

- 把相對地面的速度反向，而不是把空速反向：飛行員掉頭時，風並不會跟着掉頭。
- 以為風是吹向西面的。「來自西面」的意思是風把飛機帶向東面。
- 用求根公式逐一解兩個二次方程，再把繁複根式（surds）的倒數相減；只要看出 p_1 和 $-p_2$ 是同一個二次方程的根，就可以完全避開這些計算。

STEP 之外

這是相對速度（relative velocity）最簡單的形式：飛機在地面參考系（frame）中的速度，等於它在空氣參考系中的速度加上空氣的速度，也就是一次伽利略（Galilean）參考系變換。同一套記帳方法，只要做得仔細，就是導航、大學一年級力學中的渡河問題，以及（當它在高速下失效時）狹義相對論（special relativity）的動機的基礎。Kleppner 與 Kolenkow 的 *An Introduction to Mechanics*（力學入門）對運動參考系有詳盡的論述。

詳細解答

STEP III 1994，第 11 題

力學 • 剛體的平衡

STEP III 1994，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/94-s3-q11

關鍵思路

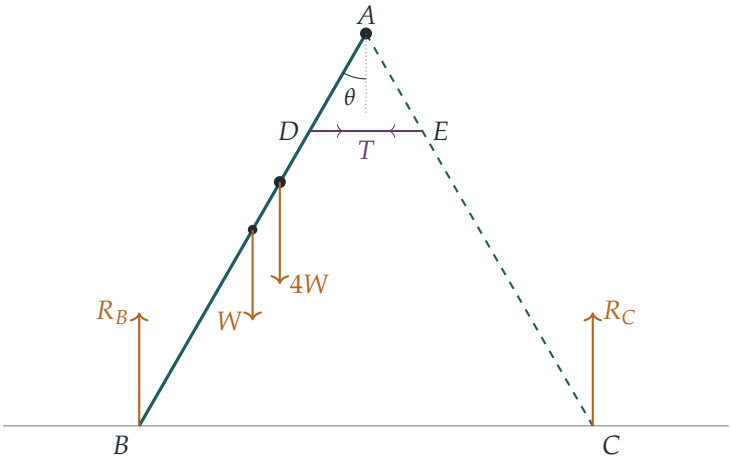
由於地面光滑（smooth），地面施加的力只能沿鉛垂方向，所以把整座梯子視為一個整體，便可立即求出 C 處的反作用力（reaction）。接下來，巧妙之處全在輕質的梯腳 AC：作用於它的力只有三個（地面的反作用力、繩的張力和鉸鏈的力），因此對鉸鏈 A 取矩（taking moments），便可消去未知的鉸鏈力，並把繩的張力（tension）與 C 處的反作用力聯繫起來。虎克定律（Hooke's law）把張力與梯子的角度連在一起，由此可把 β 表示成該角度的函數，然後求其最大值。

解答

建立模型

我們把 AB 視為均勻（uniform）桿，把 AC 視為輕質（light）桿，兩者長度均為 $4a$，在 A 處以光滑鉸鏈（hinge）相連，並把繩視為輕質彈性繩。那人則視為質點（particle）。由對稱性（$AD = AE$ 及 $AB = AC$），繩 DE 是水平的。設每條梯腳與鉛垂線的夾角為 θ，其中 $0 < \theta < \frac{\pi}{2}$。

由於地面光滑，B 和 C 處的反作用力 R_B 和 R_C 都沿鉛垂方向。AB 的重量 W 作用於其中點，與 B 的水平距離為 $2a\sin\theta$；那人的重量 $4W$ 的作用點與 B 的水平距離為 $\beta a\sin\theta$。張力 T 把 D 和 E 互相拉近。



C 處的反作用力

對整座摺梯以 B 為支點取矩（鉸鏈力和張力都是內力，所以不會出現），其中 $BC = 8a\sin\theta$：

$$R_C \cdot 8a\sin\theta = W \cdot 2a\sin\theta + 4W \cdot \beta a\sin\theta \quad \Longrightarrow \quad R_C = \tfrac{1}{4}W(1+2\beta).$$

沿鉛垂方向分解力，得 $R_B = 5W - R_C = \frac{1}{4}W(19-2\beta)$；由於 $\beta \le 4$（梯級在梯子上），此值為正。所以梯腳永遠不會離開地面。

繩的張力

繩的長度為 $DE = 2a\sin\theta$，自然長度(natural length)為 $\frac{a}{4}$。假如繩是鬆弛的，則 $T = 0$，而下文對 AC 以 A 為支點取矩便會迫使 $R_C = 0$，這是不可能的。所以繩處於拉伸狀態，$\sin\theta > \frac{1}{8}$，而由虎克定律得

$$T = \frac{W\left(2a\sin\theta - \frac{a}{4}\right)}{a/4} = W(8\sin\theta - 1).$$

輕質梯腳 AC 的平衡

想一想 為甚麼要單獨考慮 AC，而不是 AB？

梯腳 AC 是輕質的，所以作用於它的力只有三個：C 處的 R_C、E 處的張力 T，以及 A 處的鉸鏈力。對 A 取矩便可消去未知的鉸鏈力。張力沿水平方向，作用點在 A 之下、鉛垂深度為 $a\cos\theta$ 之處；反作用力 R_C 沿鉛垂方向，作用點與 A 的水平距離為 $4a\sin\theta$。所以

$$R_C \cdot 4a\sin\theta = T \cdot a\cos\theta.$$

代入 R_C 和 T：

$$W(1+2\beta)\sin\theta = W(8\sin\theta-1)\cos\theta,$$
$$\text{所以}\quad 1+2\beta = 8\cos\theta-\cot\theta,$$
$$\text{而}\quad \beta = \tfrac{1}{2}\left(8\cos\theta-\cot\theta-1\right).$$

對於那人的每個位置 β，平衡（equilibrium）可以成立，當且僅當這條方程有解 θ（且 $\sin\theta > \frac{1}{8}$）。所以 β 的最大可能值，由 $f(\theta) = 8\cos\theta-\cot\theta$ 的最大值決定。

求最大值

求導數，得

$$f'(\theta) = -8\sin\theta+\operatorname{cosec}^2\theta = \frac{1-8\sin^3\theta}{\sin^2\theta}.$$

由於 $\sin\theta$ 在 $(0,\frac{\pi}{2})$ 上遞增，當 $\sin\theta<\frac{1}{2}$ 時上式為正，當 $\sin\theta>\frac{1}{2}$ 時上式為負。所以 f 在 $\sin\theta=\frac{1}{2}$（即 $\theta=\frac{\pi}{6}$）時取得最大值，此時

$$f\left(\tfrac{\pi}{6}\right) = 8\cdot\tfrac{\sqrt{3}}{2}-\sqrt{3} = 3\sqrt{3},\qquad \beta_{\max} = \tfrac{1}{2}\left(3\sqrt{3}-1\right)\approx 2.10.$$

此值小於 4，所以梯級確實在梯子上；而且 $\sin\theta=\frac{1}{2}>\frac{1}{8}$，所以繩是拉緊的，與我們的假設相符。

梯級的高度

梯級位於由 B 起沿 BA 方向距離 βa 之處，而 BA 與鉛垂線的夾角為 θ，所以它的高度為

$$h = \beta a\cos\theta = \tfrac{1}{2}\left(3\sqrt{3}-1\right)a\cdot\tfrac{\sqrt{3}}{2} = \frac{(9-\sqrt{3})\,a}{4}\approx 1.82a.$$

驗算

- 當 $\theta=\frac{\pi}{6}$ 並取 $W=a=1$ 時：$T=8\cdot\frac{1}{2}-1=3$，$R_C=\frac{1}{4}(1+3\sqrt{3}-1)=\frac{3\sqrt{3}}{4}\approx 1.299$，而確實有 $R_C\cdot 4\sin\theta = 2.598 = T\cos\theta$。
- 以數值方法驗算：分別對兩條梯腳解出力和力矩的方程，並讓 θ 掃遍 $(0,\frac{\pi}{2})$，結果發現 β 的最大值為 2.09808，出現於 $\theta\approx 30°$，對應的高度為 $1.81699a$。
- 極限情況也合理：當 $\sin\theta\to\frac{1}{8}$ 時，繩變得鬆弛，β 便須為負數；當 $\theta\to\frac{\pi}{2}$ 時，繩的力臂 $a\cos\theta$ 趨於零，所以 β 同樣必須下降。最大值出現在兩者之間，正如所料。

常見錯誤

- 以為 A 處的鉸鏈力是水平的。一般而言並非如此，因為兩條梯腳所受的負荷不同；對 A 取矩，便完全毋須用到這個力。
- 把整段長度 $DE = 2a\sin\theta$ 當作伸長量（extension）。伸長量應為 $2a\sin\theta - \frac{a}{4}$。
- 答案寫成 $\beta_{\max}$（沿梯子量度的距離），而不是題目所問的高度；求高度還須多乘一個因子 $a\cos\theta$。

STEP 之外

這是一道靜力學（statics）題目，其中的位形（configuration，即 θ）並非已知，而是由彈性繩決定，所以各個平衡位置構成一個單參數族，而問題便變成在這個族上求最優化（optimisation）。在大學裏，這正是虛功（virtual work）和勢能（potential energy）方法背後的想法：平衡位置就是勢能的駐點（stationary point），而追問哪些平衡是穩定的，便會引出分歧（bifurcation）理論和突變理論（catastrophe theory）。Kleppner 與 Kolenkow 合著的 *An Introduction to Mechanics*（力學入門）是深入學習力學的好起點。

詳細解答

STEP III 1994，第 12 題

概率與統計 • 首步分析

> **STEP III 1994，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/94-s3-q12

關鍵思路

把各局兩局兩局地成對來看。領先 2 局的情況只可能在打完偶數局之後出現，而每打完一對局，兩名球員各發球一次，於是發球權又回到原來的一方。這樣，整場比賽就變成一個簡單的重複試驗：每一對局不是由 A 全勝，就是由 B 全勝，又或者各勝一局，而各勝一局的話，雙方便回到平手的局面。標準網球其實是同一回事，只是頭兩對局要特別處理。

解答

全篇記

$$a = p_A p_B, \qquad b = q_A q_B, \qquad s = p_A q_B + q_A p_B.$$

在連續的一對局中（兩名球員各發球一局），a 是 A 兩局全勝的概率（probability），b 是 B 兩局全勝的概率，s 是兩人各勝一局的概率（或是 A 保住發球局而 B 也保住發球局，或是雙方都破發）。我們按題目的原意，假設不同局的結果互相獨立（independent）。結果只有這三種，所以

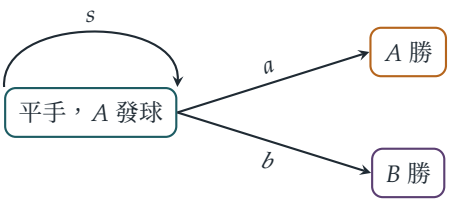
$$a + b + s = (p_A + q_A)(p_B + q_B) = 1.$$

簡化網球

每打一局，比分之差就改變 ± 1，所以打完偶數局後分差是偶數，打完奇數局後分差是奇數。因此，恰好領先 2 局的情況，最早只可能在第 $2, 4, 6, \ldots$ 局之後出現；而如果打完偶數局後比賽仍未結束，雙方必定是平手。

於是把比賽分成一對一對（第 1–2 局、第 3–4 局、...）。由平手、輪到 A 發球的局面開

始，下一對局有概率 a 令比賽以 A 勝出告終，有概率 b 令比賽以 B 勝出告終，另有概率 s 令我們回到平手、又輪到 A 發球的局面。這時的情況與開始時完全一樣，所以



$$P_{\text{short}} = a + sP_{\text{short}} \quad \Longrightarrow \quad P_{\text{short}} = \frac{a}{1-s} = \frac{a}{a+b} = \frac{p_A p_B}{p_A p_B + q_A q_B},$$

其中用到 $1 - s = a + b$。（除以 $1 - s$ 並無問題，除非 $s = 1$；那是每一對局都各勝一局、比賽永遠不會結束的退化情況。）

如果由 B 先發球，每一對局就是先打一局 B 的發球局，再打一局 A 的發球局。A 兩局全勝、B 兩局全勝、兩人各勝一局的概率分別是 $p_B p_A$、$q_B q_A$ 和 $q_B p_A + p_B q_A$：仍然是同樣的 a、b、s，因為一對局之中兩局的先後次序並無影響。同樣的論證給出同樣的答案，所以 P_{short} 與誰先發球無關。

4 局內分出勝負

未打完 4 局之前，比賽不可能分出勝負。打完 4 局後（即兩對局，由 A 先發球），分差是偶數，而比賽分出勝負，當且僅當比分是 4–0、3–1、1–3 或 0–4。用兩對局來表示：

- A 以 4–0 勝出：A 兩對局全勝，概率為 $a^2 = p_A^2 p_B^2$；
- 0–4：B 兩對局全勝，概率為 $b^2 = q_A^2 q_B^2$；
- A 以 3–1 勝出：A 全勝其中一對局，另一對局各勝一局，兩種次序都可以，概率為 $2as$；
- 1–3：同理，概率為 $2bs$。

相加得

$$\Pr(4 \text{ 局內分出勝負}) = p_A^2 p_B^2 + q_A^2 q_B^2 + 2(p_A p_B + q_A q_B)(p_A q_B + q_A p_B),$$

正是所求。

概率相等時的標準網球

如果比賽在 4 局內未分勝負，比分就是 2–2。這種情況的概率是 $s^2 + 2ab$（兩對局都各勝一局，或兩名球員各全勝一對局，兩種次序都可以）。由 2–2 開始，「領先 2 局」的規則

正好就是簡化網球由平手開始的規則，而第 5 局由 A 發球；所以由這裏開始，A 勝出的概率是 P_{short}。因此，一般而言，

$$P_{\text{stan}} = a^2 + 2as + (s^2 + 2ab)P_{\text{short}}.$$

現在取 $p_A = p_B = p$ 和 $q_A = q_B = q$，於是 $a = p^2$，$b = q^2$，$s = 2pq$，而 $P_{\text{short}} = \dfrac{p^2}{p^2+q^2}$。那麼

$$P_{\text{stan}} = p^4 + 4p^3q + \frac{6p^4q^2}{p^2+q^2}.$$

想一想 有沒有辦法比較 P_{stan} 與 P_{short}，而無須把所有式子按 p 和 q 的冪展開？

有：把簡化網球同樣在頭四局上展開。把 $P_{\text{short}} = a + sP_{\text{short}}$ 連用兩次，

$$P_{\text{short}} = a + as + s^2P_{\text{short}}.$$

用文字來說：A 全勝第一對局；或者第一對局各勝一局而 A 全勝第二對局；或者兩對局都各勝一局，一切重新開始。從 P_{stan} 的公式減去此式，s^2P_{short} 兩項互相抵消：

$$\begin{aligned} P_{\text{stan}} - P_{\text{short}} &= a^2 + as - a + 2abP_{\text{short}} \\ &= a(a+s-1) + 2abP_{\text{short}} \\ &= -ab + \frac{2a^2b}{a+b} \qquad (\text{因為 } a+s-1 = -b) \\ &= \frac{ab(a-b)}{a+b}. \end{aligned}$$

這個結果對任何 p_A、p_B 都成立。代入 $a = p^2$ 和 $b = q^2$，

$$P_{\text{stan}} - P_{\text{short}} = \frac{p^2q^2(p^2-q^2)}{p^2+q^2} = \frac{p^2q^2(p-q)(p+q)}{p^2+q^2} = \frac{p^2q^2(p-q)}{p^2+q^2},$$

因為 $p+q=1$。這就是所求的結果。

箇中原因從運算過程中已可看出：兩種賽制只在 A 全勝第一對局而 B 全勝第二對局（或次序相反）時才有分別。簡化網球把勝利判給全勝第一對局的一方；標準網球則由 2–2 繼續打下去。兩種次序的概率都是 ab，所以兩者之差是 $ab(2P_{\text{short}} - 1)$，對較強的球員有利。

驗算

- 若 $p = q = \frac{1}{2}$，兩者之差為 0，按對稱性理應如此：兩個概率都是 $\frac{1}{2}$。
- 取 $p = 0.6$：$P_{\text{short}} = 0.36/0.52 \approx 0.6923$，$P_{\text{stan}} = 0.1296 + 0.3456 + 0.124416/0.52 \approx 0.7145$，兩者之差 0.02215 與 $0.0576 \times 0.2/0.52$ 吻合。

- 4 局內分出勝負的概率 $a^2+b^2+2(a+b)s$，與打成 2–2 的概率 s^2+2ab，相加得 $(a+b+s)^2=1$，理應如此。
- 按實際規則用電腦模擬，取 $p_A=0.7$ 和 $p_B=0.45$，無論誰先發球，A 在簡化網球中勝出的次數都約佔 65.6%，與 $0.315/0.48=0.65625$ 吻合。

常見錯誤

- 忘記 3–1 可以由兩種次序得出（先各勝一局再全勝，或先全勝再各勝一局），因而漏掉因數 2。
- 以為每一局的勝出概率都相同：其實它取決於誰發球，這正是成對分組如此有用的原因。
- 寫出一個無窮等比級數（infinite geometric series），卻沒有留意到它其實就是方程 $P=a+sP$。

STEP 之外

這種逐對分析的論證，是帶有吸收態（absorbing states）的馬可夫鏈（Markov chain）的入門例子，而方程 $P=a+sP$ 就是首步分析（first-step analysis），是求吸收概率（absorption probabilities）的標準方法。同樣的想法可以求出贏得一整盤網球或一局壁球的概率，並進一步引向賭徒破產問題（gambler's ruin）和隨機漫步（random walks）。Grimmett 與 Welsh 合著的 *Probability: An Introduction*（概率論入門）正好以這個程度講解隨機漫步和馬可夫鏈。

詳細解答

STEP III 1994，第 13 題

概率與統計 • 指數分佈的壽命

> **STEP III 1994，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/94-s3-q13

關鍵思路

一切都源自指數分佈（exponential distribution）壽命的生存函數（survival function）$\mathrm{P}(T > t) = \mathrm{e}^{-\lambda t}$。以「已存活至 t_0」為條件，只不過是把兩個這樣的指數相除，著名的*無記憶性*（memoryless property）正是由此而來。直至兩條繩都失效為止的時間，是兩個獨立（independent）壽命的*最大值*，而最大值不超過 t 恰好等價於每個壽命都不超過它。因此我們先求出它的分佈函數（distribution function），然後求導。至於最後一部分，「兩條繩在同一場表演中失效」是指兩條繩在同一個區間 $((n-1)h,\, nh]$ 內失效，再對 n 求和。這個和是一個等比級數（geometric series）。

解答

已在使用中的繩

設 T 為該繩的總壽命，則對 $t \geq 0$ 有

$$\mathrm{P}(T > t) = \int_t^{\infty} \lambda \mathrm{e}^{-\lambda u}\, \mathrm{d}u = \mathrm{e}^{-\lambda t}.$$

已知該繩已經撐過了 t_0 小時，即 $T > t_0$。設 $S = T - t_0$ 為它還能再用的時間。對 $s \geq 0$，事件 $\{T > t_0 + s\}$ 包含於 $\{T > t_0\}$ 之內，所以

$$\mathrm{P}(S > s \mid T > t_0) = \frac{\mathrm{P}(T > t_0 + s)}{\mathrm{P}(T > t_0)} = \frac{\mathrm{e}^{-\lambda(t_0+s)}}{\mathrm{e}^{-\lambda t_0}} = \mathrm{e}^{-\lambda s}.$$

因此，剩餘時間 S 的概率密度（density）在 $s \geq 0$ 時為 $\lambda \mathrm{e}^{-\lambda s}$（在 $s < 0$ 時為 0）。無論 t_0 是多少，它都服從參數*同樣*為 λ 的指數分佈。

詮釋。用過的繩與新繩完全一樣好。繩並不「記得」自己過去用了多久，無論繩齡多大，

它在接下來一小段時間 δt 內失效的機會都約為 $\lambda\,\delta t$（即失效率（failure rate）為常數）。這種模型適用於由突發的隨機意外所造成的失效。但對真實的繩來說並不切合實際：繩會磨損、起毛，所以舊繩應該更容易在短期內失效。採用失效率隨繩齡遞增的模型會更好。表演者堅持使用新繩，在現實中是明智的，儘管在這個模型之下並無分別。

直至兩條新繩都失效的時間

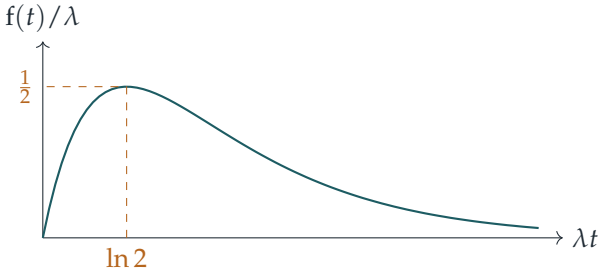
設 T_1, T_2 為兩條新繩的壽命。兩者互相獨立，各自的分佈函數為 $\mathrm{P}(T_i \le t) = 1 - \mathrm{e}^{-\lambda t}$ $(t \ge 0)$。兩條繩到時刻 t 都已失效，恰好等價於較遲的那次失效已經發生，所以我們要求的時間是 $M = \max(T_1, T_2)$。現在 $M \le t$ 當且僅當 $T_1 \le t$ 並且 $T_2 \le t$，所以由獨立性得

$$\mathrm{P}(M \le t) = \mathrm{P}(T_1 \le t)\,\mathrm{P}(T_2 \le t) = \left(1 - \mathrm{e}^{-\lambda t}\right)^2 \qquad (t \ge 0),$$

而對 $t < 0$ 則有 $\mathrm{P}(M \le t) = 0$。利用鏈式法則（chain rule）求導，

$$\mathrm{f}(t) = 2\left(1 - \mathrm{e}^{-\lambda t}\right) \cdot \lambda \mathrm{e}^{-\lambda t} = 2\lambda \mathrm{e}^{-\lambda t}\left(1 - \mathrm{e}^{-\lambda t}\right) \quad (t \ge 0),$$

而對 $t < 0$ 則有 $\mathrm{f}(t) = 0$，正如題目所求。



與單一指數壽命不同，M 在 $t = 0$ 處的密度為 0，因為兩次失效不可能一開始就同時發生。密度在 $\mathrm{e}^{-\lambda t} = \frac{1}{2}$ 處達到最高點。

兩條繩都在第 n 場表演中失效

第 n 場表演對應表演時間的區間 $\big((n-1)h,\, nh\big]$。記 $q = \mathrm{e}^{-\lambda h}$。對其中一條繩，

$$\mathrm{P}\big((n-1)h < T_i \le nh\big) = \mathrm{e}^{-\lambda(n-1)h} - \mathrm{e}^{-\lambda nh} = q^{n-1}(1-q).$$

想一想 我們應否把 $\mathrm{f}(t)$ 在第 n 場表演的時段上積分？

不應該。$\int_{(n-1)h}^{nh} \mathrm{f}(t)\,\mathrm{d}t$ 是第二次失效發生在第 n 場表演中的概率。第一條繩可能早在之前的某場表演中已經失效。我們需要的是每一條繩都在第 n 場表演中失效，而兩條繩互相獨立，所以

$$p_n = \left(q^{n-1}(1-q)\right)^2 = \mathrm{e}^{-2\lambda(n-1)h}\left(1 - \mathrm{e}^{-\lambda h}\right)^2.$$

兩條繩在同一場表演中失效

對 $n = 1, 2, 3, \ldots$，各個「兩條繩都在第 n 場表演中失效」的事件互斥（mutually exclusive），所以我們把它們的概率相加。這是一個公比為 $q^2 < 1$ 的等比級數（因為 $\lambda h > 0$）：

$$\sum_{n=1}^{\infty} p_n = (1-q)^2 \sum_{n=1}^{\infty} q^{2(n-1)} = \frac{(1-q)^2}{1-q^2} = \frac{1-q}{1+q}.$$

現在把分子和分母同乘以 $\mathrm{e}^{\lambda h/2}$：

$$\frac{1-\mathrm{e}^{-\lambda h}}{1+\mathrm{e}^{-\lambda h}} = \frac{\mathrm{e}^{\lambda h/2} - \mathrm{e}^{-\lambda h/2}}{\mathrm{e}^{\lambda h/2} + \mathrm{e}^{-\lambda h/2}} = \tanh\left(\frac{\lambda h}{2}\right),$$

這就是所要求的結果。

驗算

- $\int_0^{\infty} \mathrm{f}(t)\,\mathrm{d}t = \left[(1-\mathrm{e}^{-\lambda t})^2\right]_0^{\infty} = 1$，所以 f 確實是一個概率密度。
- 極限情況：當 $h \to \infty$（一場永不完結的表演）時，答案 $\tanh(\lambda h/2) \to 1$。當 $h \to 0$ 時，它趨向 0，因為兩個連續型壽命幾乎不可能落在同一個極短的區間內。當 h 很小時，它約為 $\lambda h/2$。
- 數值上，取 $\lambda = 0.7$ 和 $h = 0.8$，級數與 $\tanh(0.28)$ 都給出 0.27291。模擬 2×10^5 對繩所得的結果為 0.2746，而 $p_1 \approx 0.1839$ 亦與模擬所得的 0.1848 吻合。

常見錯誤

- 用了最小值而非最大值。最小值是直至第一條繩失效的時間，它服從參數為 2λ 的指數分佈。
- 把 f 在第 n 場表演的時段上積分來求 p_n。這些概率加起來是 1，而不是 $\tanh(\lambda h/2)$。

STEP 之外

無記憶性刻劃了指數分佈：它是唯一具有這個性質的連續型分佈。在可靠性理論（reliability theory）中，「失效率」（風險函數，hazard function）$\mathrm{f}(t)/\mathrm{P}(T > t)$ 只有在指數分佈下才是常數。磨損則以遞增的風險函數來建模，例如採用韋布爾分佈（Weibull distribution）。獨立變量的最大值和最小值是次序統計量（order statistics）的兩個極端情況，你會在大學的第一門概率課程中遇到它們。Grimmett 與 Welsh 所著的 *Probability: An Introduction*（概率論入門教材）涵蓋了以上所有內容。

詳細解答

STEP III 1994，第 14 題

概率與統計 • 幾何概率

STEP III 1994，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/94-s3-q14

關鍵思路

圓周角（inscribed angle）等於所對的弧在圓心所張的角的一半。因此，三角形的三個角正好是 π 乘以 P、Q、R 把圓分成的三段弧，而弧長以佔圓周的比例來量度。問題於是變成：一個圓怎樣被隨機分成三段？而這其實是一個關於三角形內面積的問題。

解答

建立模型

以佔圓周的比例來量度圓上的位置，這樣圓的「長度」就是 1。由旋轉對稱性（rotational symmetry），P 的位置無關重要，所以把 P 放在 0。於是 Q 和 R 的位置是獨立的，並在 $[0,1)$ 上均勻分佈（uniformly distributed）。

設 $U < V$ 為 Q 和 R 的位置由小至大的排列。數對 (U,V) 在區域 $0 < u < v < 1$ 上均勻分佈，而這個區域的面積是 $\frac{1}{2}$，所以它的聯合密度（joint density）在該區域上是 2。（兩點重合的概率是 0，下文所有邊界情況亦然。）三段弧是

$$X = U, \qquad Y = V - U, \qquad Z = 1 - V, \qquad X + Y + Z = 1.$$

變數代換 $(u,v) \mapsto (x,y) = (u, v-u)$ 的雅可比行列式（Jacobian）是 1，所以 (X,Y) 在三角形

$$T = \{(x,y) : x > 0,\ y > 0,\ x + y < 1\},$$

上的密度是 2，而這個三角形的面積是 $\frac{1}{2}$。因此，下文每一個概率都是 T 內某塊面積的兩倍。

想一想 為甚麼三角形的每個角都是 π 乘以一段弧？

三角形在 R 的角是一個圓周角，它所對的是不包含 R 的那段弧 PQ。根據圓周角定理（inscribed angle theorem），它等於該弧在圓心所張的角的一半。一段佔圓周比例為 x 的弧，在圓心所張的角是 $2\pi x$，所以在 R 的角是 πx。每個頂點的情況都一樣，所以 PQR 的三個角是 πX、πY 和 πZ（它們的和是 π，正如所料）。因此

$$\text{某個角大於}k\pi \iff \text{對應的弧大於}k.$$

設 A_X、A_Y 和 A_Z 分別為事件 $X > k$、$Y > k$ 和 $Z > k$。我們要求 $\mathrm{P}(A_X \cup A_Y \cup A_Z)$。

一段弧大於 k

對於 $0 \le k \le 1$：

- A_X 是 T 中 $x > k$ 的部分：一個兩條直角邊均為 $1-k$ 的三角形，面積是 $\frac{1}{2}(1-k)^2$。
- A_Y 是 $y > k$ 的部分：同樣的計算給出面積 $\frac{1}{2}(1-k)^2$。
- A_Z 是 $x + y < 1 - k$ 的部分：位於角落、兩條直角邊均為 $1-k$ 的三角形，面積是 $\frac{1}{2}(1-k)^2$。

乘以二，得 $\mathrm{P}(A_X) = \mathrm{P}(A_Y) = \mathrm{P}(A_Z) = (1-k)^2$。

$\frac{1}{2} \le k \le 1$ 的情況

最多只有一段弧可以大於 k：假如有兩段，它們的總和就會大於 $2k \ge 1$，但三段弧加起來才是 1。所以事件 A_X、A_Y、A_Z 互斥（mutually exclusive），而

$$\mathrm{P}(\text{某個角大於}k\pi) = 3(1-k)^2,$$

正是所求。

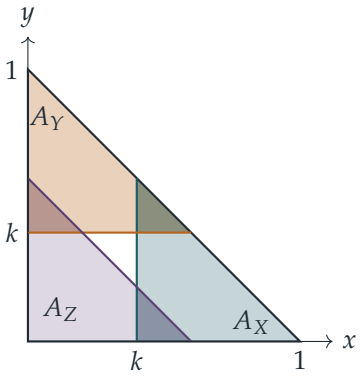
$\frac{1}{3} \le k \le \frac{1}{2}$ 的情況

現在可以有兩段弧同時大於 k，所以我們使用容斥原理（inclusion–exclusion）。

兩段弧。 $A_X \cap A_Y$ 是 T 中 $x > k$、$y > k$、$x + y < 1$ 的部分。代入 $x = k + s$、$y = k + t$，條件變成 $s, t > 0$、$s + t < 1 - 2k$，這是一個兩條直角邊均為 $1 - 2k \ge 0$ 的三角形，面積是 $\frac{1}{2}(1-2k)^2$。所以 $\mathrm{P}(A_X \cap A_Y) = (1-2k)^2$。另外兩對的結果相同：例如 $A_X \cap A_Z$ 是 $x > k$、$y > 0$、$x + y < 1 - k$，同樣是一個兩條直角邊均為 $1 - 2k$ 的三角形。

三段弧。 三段弧全部大於 k 就需要 $X + Y + Z > 3k \ge 1$，這是不可能的。所以三個事件的交集的概率是 0。

下圖顯示 $k = 0.4$ 時的 T：三個位於角落的三角形是 A_X、A_Y、A_Z，顏色較深的重疊部分是兩兩的交集。



由容斥原理，

$$\begin{aligned}\mathrm{P}(A_X \cup A_Y \cup A_Z) &= 3(1-k)^2 - 3(1-2k)^2 + 0 \\ &= 3\big[(1-k)-(1-2k)\big]\big[(1-k)+(1-2k)\big] \\ &= 3k(2-3k).\end{aligned}$$

所以當 $\frac{1}{3} \le k \le \frac{1}{2}$ 時，至少有一個角大於 $k\pi$ 的概率是

$$3k(2-3k) = 6k - 9k^2.$$

驗算

- 在 $k = \frac{1}{2}$ 時，兩條公式都給出 $\frac{3}{4}$，所以它們連續地銜接起來。這就是那個經典結果：隨機的圓內接三角形是鈍角三角形的概率為 $\frac{3}{4}$。
- 在 $k = \frac{1}{3}$ 時，我們得到 $3 \cdot \frac{1}{3} \cdot 1 = 1$。這是對的：最大的角永遠不小於 $\frac{\pi}{3}$，因為三個角加起來是 π。
- 在 $k = 1$ 時，我們得到 0，因為沒有角可以大於 π。
- 模擬 4×10^5 個隨機三角形，直接由坐標計算各個角，在 $k = 0.4$ 時得到 0.9602（公式給出 0.96），在 $k = 0.6$ 時得到 0.4783（公式給出 0.48）。

常見錯誤

- 用了整段弧 PQ，而不是不包含第三個頂點的那段弧。兩段弧 PQ 之中，只有一段是 R 所對的弧。

- 把 X、Y、Z 當作獨立的均勻隨機變數。它們受 $X+Y+Z=1$ 所約束，而老老實實處理這個約束的方法，就是面積圖。
- 在 $k=\frac{1}{2}$ 以下仍然套用 $3(1-k)^2$。當 $k<\frac{1}{2}$ 時，它把有兩個大角的三角形重複計算了，而在 $k=\frac{1}{3}$ 時它會給出 $\frac{4}{3}>1$。

STEP 之外

弧 (X,Y,Z) 在單純形（simplex）$x+y+z=1$ 上均勻分佈。這就是狄利克雷分佈（Dirichlet distribution），它描述隨機點把一個區間或一個圓切開後留下的「間距」。有 n 個點時，同樣的容斥原理給出 $\mathrm{P}(\text{最大間距}>k)=\sum_{j}(-1)^{j+1}\binom{n}{j}(1-jk)_{+}^{n-1}$，這條公式可追溯至 Whitworth。Grimmett 與 Welsh 的 *Probability: An Introduction*（概率論入門）對幾何概率和聯合密度有很好的論述。

Beyond STEP

STEP 是大多數中學生一生中應考過最艱深的數學考試，也是最出色的考試之一：題目簡短、陌生而深刻，當中最精彩的，其實是喬裝了的大學數學小品。

本書精選最有趣的 STEP 試題，由 1987 年首次開考之前的樣本試卷（specimen papers），到 2025 年的試卷，並以每一道題目作為一扇門。每道題目都會仔細解答，沿途提出優秀學生會問的問題，然後帶你穿過這扇門：這個想法從何而來，到了大學會演變成甚麼，以及接下來該讀些甚麼。

本書為正在準備 STEP、有天分又有志投身數學的學生而寫。他們想要的不止是模擬試卷：他們想開始像數學系本科生那樣思考，並一探前方有甚麼正等着自己。

- 共 13 章，涵蓋純數學（pure mathematics）、力學（mechanics）、概率（probability）與統計（statistics）
- 136 道 STEP 歷屆試題，每題均附引導式解答
- 每道題目背後的大學數學
- 附答案的延伸練習，以及延伸閱讀

1994 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。