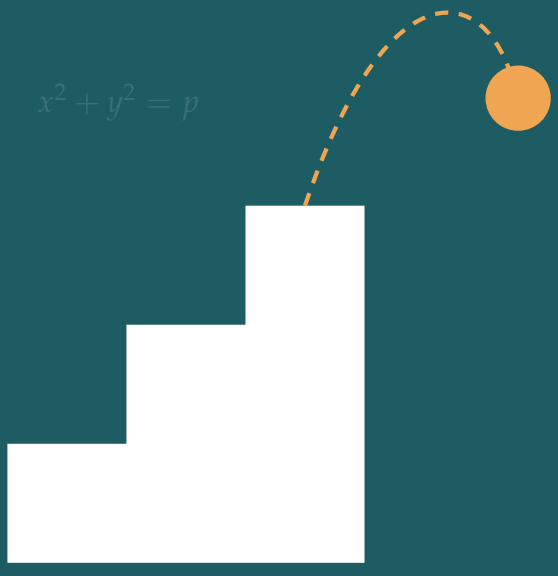
STEP 備試 · 大學數學

Beyond STEP

以 STEP 歷屆試題為門徑
通往大學數學



KEITH McNULTY

Beyond STEP

Past STEP questions as a doorway
to university mathematics

Keith McNulty

Beyond STEP: Past STEP questions as a doorway to university mathematics

First published online November 2026. This version: 2026年11月18日.

Everything in this book that is the author's own work, including the text, guided solutions, figures, exercises and answers, is dedicated to the public domain under the Creative Commons CC0 1.0 Universal licence: you may copy, adapt and share it for any purpose, without asking permission. See `creativecommons.org/publicdomain/zero/1.0`.

The STEP questions discussed in this book are the copyright of Cambridge University Press & Assessment. They are not reproduced here: each section links to its question at an official source, the STEP Support Programme's question database or OCR's past papers.

The book is also available free online, with printable worksheets for every section, at `beyondstep.keithmcnulty.org`. Corrections are welcome as issues at `github.com/keithmcnulty/step_book_mcnulty`.

本版本由機器從英文翻譯而成，尚未經以中文為母語的數學教師審閱。如有不清楚之處，概以英文原文為準。

目錄

導言 1

詞彙表 11

數學符號 23

1 數論與證明 **33**

1.1 質數工廠 39

1.2 質數分佈得有多稀疏？ 47

1.3 最小反例 53

1.4 e 是無理數，e^2 也是 59

1.5 無窮遞降法 65

1.6 兩個平方數之和 71

1.7 因數分解失效之時 77

1.8 點算有理數 83

1.9 三次曲線的切線 89

回顧 95

練習答案 99

2 數列、級數與收斂 **103**

2.1 「最終」是甚麼意思 109

2.2 作為極限的黃金比例 115

2.3 一個發散級數與它的收斂近親 121

2.4 $\pi^2/6$ 127

2.5 韋達乘積 133

2.6 算術幾何平均 139

2.7 六位還是六百位 145

2.8 平方再減二 151

2.9 逐步逼近 e 的分數 159
回顧 165
練習答案 169

3 函數、不等式與嚴謹 173

3.1 找出錯處 179
3.2 倒過來的 AM–GM 185
3.3 凸性的威力 191
3.4 以面積證明不等式 197
3.5 π^e 還是 e^π？ 203
3.6 斜率的界 209
3.7 $\pi^2 \leq 10$ 215
3.8 由加法公式決定的函數 221
3.9 創造一個新函數 227
回顧 233
練習答案 237

4 嚴謹的微積分 241

4.1 甚麼是積分？ 247
4.2 由一個積分建立三角學 253
4.3 22/7 大於 π 259
4.4 沃利斯、Beta 與二分之一的階乘 265
4.5 大二學生之夢 271
4.6 沒有公式的積分 277
4.7 變分法 283
4.8 切球形麵包 289
4.9 Holditch 定理 295
回顧 301
練習答案 305

5 複數 **309**

5.1 梅欽公式 315

5.2 單位根篩選法 321

5.3 因式分解 $z^{2n}+1$ 327

5.4 正多邊形中的距離 333

5.5 名不副實的有理角 339

5.6 費馬點 345

5.7 作為映射的複指數函數 351

5.8 圓盤中隱藏的距離 357

回顧 363

練習答案 367

6 代數與結構 **371**

6.1 以單位根為根的多項式 377

6.2 高斯的正十七邊形 383

6.3 以 $\sqrt{2}+\sqrt{3}+\sqrt{5}$ 為根的多項式 389

6.4 解四次方程 395

6.5 從公理出發的微分法 401

6.6 子羣的併集 407

6.7 四元數羣 413

6.8 由奇異矩陣組成的羣 419

回顧 425

練習答案 429

7 線性代數與幾何 **433**

7.1 對矩陣不成立的法則 439

7.2 只伸縮，不轉向 445

7.3 甚麼才算行列式 451

7.4 回到起點的矩陣 457

7.5 三維空間中的旋轉 463
7.6 雙曲旋轉與相對論 469
7.7 永恆不變的量 475
7.8 四個平方數 481
7.9 只有五種柏拉圖立體 487
回顧 493
練習答案 497

8 微分方程與動力系統 501

8.1 當解不唯一 507
8.2 不求解而估界 513
8.3 收穫種羣 519
8.4 盤旋趨向圓 525
8.5 格力犬追逐格力犬 531
8.6 量子振子 537
8.7 可以解出的混沌 543
8.8 週期三蘊含混沌 551
回顧 559
練習答案 563

9 計數與機會 567

9.1 直線分割平面 573
9.2 小型版的拉姆齊定理 579
9.3 為籬笆樁着色 585
9.4 把數列看成函數 591
9.5 湊零錢 597
9.6 奇怪的骰子 603
9.7 相等的冪和 609
9.8 沒有人拿到自己的恤衫 615

9.9 生日問題 621
9.10 總能找贖 627
9.11 集齊一套 633
9.12 隨機子集 639
9.13 做了手腳的骰子 645
回顧 651
練習答案 655

10 守恆定律、碰撞與引力 661

10.1 牛頓擺 667
10.2 疊球效應 673
10.3 火箭與槍炮 679
10.4 一個分子與一個活塞 685
10.5 有限時間內的無限次反彈 691
10.6 二體化為一體 697
10.7 球殼內的引力 703
10.8 排成三角形的三顆恆星 709
回顧 715
練習答案 719

11 振動、轉動與運動參考系 723

11.1 共振 729
11.2 簡正模式 735
11.3 真實的擺 741
11.4 等時曲線 747
11.5 加速火車中的重力 753
11.6 後旋 759
11.7 桿中的辛普森法則 765
11.8 傾斜的平衡 771

11.9 安全拋物線 777
11.10 會朝你飛回來的球 783
11.11 外野手如何接球 789
11.12 橡皮繩上的螞蟻 795
11.13 鏟雪車問題 801
11.14 是拋物線，不是懸鏈線 807
回顧 813
練習答案 817

12 概率 823

12.1 精確地取條件 829
12.2 基本比率謬誤 835
12.3 「隨機」是甚麼意思？ 839
12.4 兩兩獨立，卻非相互獨立 845
12.5 最後一位乘客的座位 851
12.6 多一枚硬幣 857
12.7 淘汰賽中誰會遇上誰 861
12.8 Penney 遊戲 865
12.9 用不公平的硬幣作公平的決定 871
12.10 把繩頭繫起來 877
12.11 郵票、杯子與 e 883
12.12 何時停止尋找 887
12.13 三人決鬥 893
12.14 混合策略 897
12.15 滅絕還是爆發 903
12.16 順向與逆向的馬可夫鏈 909
12.17 巴拿赫的火柴盒 915
12.18 無記憶性 921
12.19 平面上的隨機點 927

12.20 蒲豐投針 933
12.21 折斷一根棒 939
12.22 沒有平均值的分佈 945
12.23 由概率得出斯特靈公式 951
12.24 中心極限定理 957
12.25 眾多之中的最大值 963
回顧 969
練習答案 973

13 統計與現實世界中的隨機性 983

13.1 解讀新聞標題 987
13.2 點算田鼠 993
13.3 德國坦克問題 999
13.4 平均數還是中位數？ 1005
13.5 問尷尬的問題 1011
13.6 累積起來的誤差 1017
13.7 製造相關性 1023
回顧 1029
練習答案 1033

引言

> 我無法靠這本書教會你數學，但我相信，只要你肯下苦功，一點一點地鑿開這些題目，終會讓你內在的那位數學家顯露出來。
>
> — Stephen Siklos，*Advanced Problems in Mathematics*

這是一本關於難題的書，也是一本關於難題背後有甚麼的書。全書 136 節，每一節都由一道 STEP 歷屆試題開始。但題目只是入口。幾乎每一道好的 STEP 題目背後，都有一片你將會在大學裏正式學到的數學，而本書的目的，就是帶你穿過這道門，讓你看看門後的風景。

STEP 是甚麼，為何存在

STEP，即 Sixth Term Examination Paper，是一組在中學最後一年夏天應考的數學考試。它由劍橋大學（University of Cambridge）於 1987 年設立，作用是協助決定誰適合到那裏修讀數學：劍橋數學系發出的取錄通知，幾乎全部以 STEP 成績為條件，多年來其他大學亦有採用，當中以華威大學（Warwick）最為人熟知。起初 STEP 設有多個科目的試卷，但到了 2001 年，就只剩下數學一科。在它大部分的歷史裏，考試共有三份試卷，即 I、II 和 III；Paper I 最後一次開考是在 2019 年，現時考生應考 Paper II 和 Paper III。這項考試多年來由劍橋大學考評部（Cambridge Assessment）主辦，自 2024 年起改由 OCR 負責。

STEP 試卷與學校的考試截然不同。卷上不是十道必答的短題，而是十多道長題，考生從中選答六題，而一位實力強的考生只要在四題上取得實質進展，成績已經非常好。題目不會告訴你每一部分佔多少分，也很少要你做一些你練習過的東西。它們要你把已經懂得的數學，用在全新的地方。

對 STEP 的認識無人能及的 Stephen Siklos，曾這樣形容它的用途。它是一道門檻：事實證明，STEP 的成績能相當準確地預示學生在大學數學學位課程中的表現（據他所述，STEP 名次與劍橋一年級考試名次的相關係數達 0.63，以這類數據的標準而言非常高）。它是一種準備：STEP 所獎勵的思考方式，正是大學本科數學的思考方式。它也是對動力的考驗：準備 STEP 要下真功夫，一個不願付出這份努力的學生，大概也不會享受要求嚴格的數學學位課程。

Stephen Siklos

談 STEP 而不談 Stephen Siklos，是不可能的。他是劍橋的數學家，曾是霍金（Stephen Hawking）的研究生，研究早期宇宙和引力波的數學，在應用數學及理論物理系任講師，其後出任耶穌學院（Jesus College）的高級導師（Senior Tutor），再擔任該學院的 President。1987 年 STEP 數學科創辦時，他出任首席考官（Senior Examiner），並擔任此職三十二年，直至 2019 年 8 月離世為止。這些年間，他草擬題目，「把別人的題目修整成形」（這是他自己的說法，他說這件事長久以來都是他興趣和樂趣的來源），逐一審閱每份試卷，並督導評卷工作。他也是 STEP 復活節學校（STEP Easter School）和 STEP 支援計劃（STEP Support Programme）背後的推手，這兩個項目為學校無法提供協助的學生免費提供幫助。有一篇訃聞說，在令劍橋數學入學試向不同學校背景的學生開放這件事上，沒有人比他做得更多。

他的著作 *Advanced Problems in Mathematics: Preparing for University*（《高等數學難題：為大學作準備》）由 Open Book Publishers 在網上免費出版，至今仍是認識 STEP 的最佳入門書，我極力推薦。他在書中寫下的兩點，塑造了本書的面貌。第一點是他對甚麼才算一道好的 STEP 題目的看法：題目的第一部分向你展示一個陌生的想法，其後的部分則要你在更難的情境中再次運用它，以證明你確實明白；又或者題目的答案取決於一個要你自己去發現的參數；又或者題目由特殊情況開始，以一般結果作結。正是這類題目才會引領你到達某處，而本書的題目大多屬於這幾類。第二點是他堅持解答應該是「跳板，而不是羽毛牀」。一份解答，如果你未曾先自己與題目搏鬥過就去讀，是學不到任何東西的。

我與 STEP 的相遇

我第一次接觸 STEP 是在 1990 年代初，當時我的數學老師給了我幾份試卷，作為額外的挑戰。我做這些試卷是為了樂趣，而不是為了取錄資格，也從未真正應考，但我記得那些試卷難得出奇：題目看似與我學過的任何東西都毫無關係，一道題就寫滿一頁，我能夠起步，卻很少能做完。我也記得，在終於做通了的那些時刻，我感到自己看見了一些學校數學從未讓我看見的東西。

許多年後，我從另一邊回到 STEP：協助學生準備應考。與他們一起做過數百道題目之後，我對這些試卷有了不同的看法。並非每一道題都是瑰寶。坦白說，有些題目只是漫長的耐力測試：一整頁要小心處理的代數，一段要求很高卻屬例行公事的計算，到最後既沒有多少洞見，也沒有多少樂趣。耐力固然重要，大學數學也需要它，但令 STEP 與別不同的並不是這些題目。

令它與別不同的是，在 1987 年以來所出的 1,500 多道題目之中，散落着數百道真正優美的題目。要解釋我的意思，最好的方法就是給你看幾道。

幾顆瑰寶

以下是你稍後會在書中遇到的七道題目，我特意選來展示本書涵蓋的範圍：兩道來自組合數學（combinatorics），兩道來自動力系統（dynamical systems），兩道來自力學（mechanics），一道來自概率（probability）。每一道都由你能想像得到的東西開始，以你可能難以置信的結果作結，並通往一片在大學裏會認真研究的數學。

藏在混亂中的秩序。在紙上畫六個點，把每兩點用一條線連起來，合共十五條線，每條線隨你喜歡塗成金色或銀色。現在試試避免出現三邊同色的三角形。你每一次都會失敗。五個點是做得到的（把五邊形的邊塗成金色，五條對角線塗成銀色），但六個點就不可能，無論你的塗法多麼巧妙。1989 年 STEP I 的一道題目要你證明這件事，而且只給一個提示：看看從其中一點出發的五條線。其中三條必定同色，由此出發，證明只需寥寥數行。線的塗色方法共有 $2^{15} = 32\,768$ 種，而這個論證一次過處理了全部情況，連一個個案也不用檢查。

這個結果是拉姆齊定理（Ramsey's theorem）最小的情況。這條定理由拉姆齊（Frank Ramsey）於 1928 年證明，它說的是：徹底的無序是不可能的；任何足夠大的結構，無論拼湊得多麼雜亂無章，都必定包含一小塊完美的秩序。多大才算「足夠大」？要保證有四個點彼此全以同一顏色相連，你需要 18 個點，這是 1955 年證明的事實。至於五個點，沒有人知道：答案介乎 43 與 46 之間，而上界到 2024 年才獲證明。艾狄胥（Paul Erdős）曾開玩笑說，如果外星人威脅要毀滅地球，除非我們說出這個數，我們就應該動員所有電腦和數學家去找；但如果他們問的是六個點的答案，我們就應該設法消滅外星人。你會在大學的組合數學或圖論（graph theory）課程中遇到拉姆齊理論，而推動這個 STEP 證明的鴿巢原理（pigeonhole principle），則幾乎在你修讀的每一門課都會出現（Section 9.2）。

奇怪的骰子。擲兩顆普通骰子，總和為 7 的機會最大，總和為 2 和 12 的機會最小，各總和的擲法數目依次為大家熟悉的 $1, 2, 3, 4, 5, 6, 5, 4, 3, 2, 1$。用正整數為一對骰子編號，還有沒有別的方法，使每個總和出現的機會完全相同？看來不大可能，但這樣的方法恰好有一個：一顆骰子寫上 $1, 2, 2, 3, 3, 4$，另一顆寫上 $1, 3, 4, 5, 6, 8$。這就是 *Sicherman* 骰子，以發現它們的 George Sicherman 命名，並因馬丁·加德納（Martin Gardner）於 1978 年的介紹而廣為人知。在任何只取決於所擲總和的遊戲中，你大可用它們換掉普通骰子，而永遠不會有人察覺。

1993 年的一道題目從另一端切入這個想法。它給你一個表，列出用兩顆編號古怪的骰子擲出每個總和的方法數目，然後要你推算出骰子的編號，就像偵探根據遺下的痕跡重組案情一樣。該節接着介紹令整個課題變得簡單的想法：把每顆骰子變成一個多項式（polynomial），每一個點數為 k 的面對應一項 x^k。擲兩顆骰子相當於把它們的多項式相乘，因為 $x^i \times x^j = x^{i+j}$ 正好把兩面的數字相加。於是，尋找總和分佈相同的新骰子，就變成用另一種方式把一個多項式因式分解，情況有點像 36 可以拆成 4×9 而不是

6×6。又因為多項式與整數一樣，分解成不能再拆的因子時只有一種方式，這個因式分解同時證明了 Sicherman 骰子是唯一的另類選擇。把計數問題編碼成多項式，即*生成函數*（generating function），這個技巧是組合數學中最強大的工具之一，而背後的唯一分解（unique factorisation），則是大學第一門代數課的中心主題（Section 9.6）。

有公式的混沌。取一個介乎 0 與 1 之間的數，把它乘以 4，再乘以一減去它本身，然後不斷重複：$u_{n+1} = 4u_n(1 - u_n)$。這就是*邏輯斯諦映射*（logistic map），一個粗略的模型，描述一個不斷增長但受糧食供應限制的種羣。這條規則簡單得不能再簡單，但它產生的數卻跳來跳去，看不出任何規律，而兩個小數點後十位都相同的初始值，通常在數十步之內就會給出完全不同的數列（sequence）。生物學家 Robert May 於 1976 年在一篇著名的論文中引起大家對它的注意，論文題為「Simple mathematical models with very complicated dynamics」（具有非常複雜動態的簡單數學模型），它自此成為混沌（chaos）的標準例子，通常藉電腦來探索。

2014 年的一道題目揭示了一件出人意表的事：這個數列有一條精確的公式。如果初始值是 $\sin^2\theta$，那麼 $u_n = \sin^2(2^n\theta)$，而你只需用倍角公式（double angle formulae）和歸納法（induction）就能證明。（烏拉姆（Stanisław Ulam）和馮·諾伊曼（John von Neumann）在 1940 年代已經知道這條公式。）你也許以為有了精確的公式，混沌就會被馴服。事實上，公式反而把混沌解釋得一清二楚。每一步都把角加倍，所以角上任何微小的不確定性每一步都會加倍，而初始值的誤差平均而言也以同樣的速度增長：五十步之後，誤差已增大了約一千萬億倍。把這個角表示成 π 的分數倍，再寫成二進制（binary），就看得格外分明：映射的每一步都把二進制數字向左移一位，並丟掉第一位，所以數列是把初始值的數字逐個讀出來，而要預測一百步之後的情況，你就要知道初始值準確至一百個二進制位。同一條公式還能找出這個映射的所有循環（cycle），並顯示每一種長度的循環都存在，它們稠密地擠在 $[0, 1]$ 之內，而且每一個都不穩定。這就是動力系統這門學問，你會在大學的微分方程（differential equations）課程中遇到它，其後還會在專門講混沌的課程中再遇上（Section 8.7）。

週期三意味着混沌。取函數 $f(x) = 7 - 2|x|$，由一個數開始，不斷把答案代回去。由 1 開始，得到 $1, 5, -3, 1, 5, -3, \ldots$：一個長度為三的循環。2025 年的一道題目要你尋找這類循環，方法是描繪 f、$f(f(x))$ 和 $f(f(f(x)))$ 的圖像，找出它們與直線 $y = x$ 相交的位置。每多應用一次 f，圖像中的鋸齒數目就加倍，而該節把這個觀察化為一個證明，說明這個函數具有*每一種*長度的循環：二、三、四、一百、一百萬。

這並不是這個函數獨有的特點。1975 年，李天岩（Tien-Yien Li）和 James Yorke 證明了*任何*具有長度為三的循環的連續函數（continuous function），都必定具有其他每一種長度的循環，論文的題目令人難以抗拒：「Period three implies chaos」（週期三意味着混沌）。證明所需的不過是介值定理（intermediate value theorem），也就是每當你論證一個變號的連續函數必定穿過零時已經在用的定理，再加上一幅巧妙的圖，顯示函數把哪些區間拉伸覆蓋到哪些區間之上。後來人們才發現，烏克蘭數學家沙可夫斯基（Oleksandr

Sharkovsky）早在 1964 年已證明了一個更強的結果：他找到了正整數的一種奇特排序，以 $3, 5, 7, 9, \ldots$ 開始，使得任何一種長度的循環，都會迫使排在它後面的每一種長度的循環出現。介值定理是大學分析（analysis）課程中最先嚴格證明的結果之一；這道題目顯示了它能做到多少事情（Section 8.8）。

單車氣泵為何會發熱。替單車輪胎打氣，然後摸一摸氣泵：它是暖的。擠壓氣體會令它升溫，柴油引擎正是這樣不靠火花就點燃燃料。但為甚麼呢？2002 年的一道題目這樣回答：追蹤單獨一個空氣分子，看它在氣泵末端與正被推入的活塞之間來回彈跳。分子每次從推進中的活塞反彈，都會加快一點點，就像網球被移動中的球拍擊中一樣。題目要你準確描述一次反彈，把它化成每次碰撞時氣缸長度的遞歸關係（recurrence），然後解出這個遞歸關係。

解答揭示了一條隱藏的定律：分子的速率乘以氣缸的長度，幾乎不變。把長度減半，速率就加倍，而由於溫度量度的是分子的能量，氣體便會升溫。像這樣在系統緩慢改變時幾乎保持不變的量，稱為*絕熱不變量*（adiabatic invariants），而只要再多下一點功夫，同一個論證就能導出氣體的絕熱壓縮定律，你會在大學的熱力學中學到它。這個想法影響深遠。費米（Enrico Fermi）於 1949 年提出，宇宙射線正是以這種方式獲得巨大的能量，即在太空中移動的磁場之間反覆彈跳；而埃倫費斯特（Paul Ehrenfest）則於 1916 年指出，絕熱不變量正是量子力學必須加以量子化的量。一道關於單車氣泵的題目，原來觸及統計物理、天體物理和量子理論的開端（Section 10.4）。

會回來的球。把一個球輕輕向前拋，它會穩定地遠離你，直至落地。常識告訴我們事情總是如此：球一直向前移動，又怎會靠近你呢？但只要拋得夠陡，與水平所成的角大於約 $70.5°$，怪事就會發生。在飛行的某一段，當球下降時，它會離你*更近*，儘管它在水平方向上仍然遠離你。它的高度下降得比水平距離增長得更快。

2019 年的一道題目要你證明這件事，並求出臨界角的準確值。所用的方法你日後會一用再用：不對距離本身求導，而對距離的平方求導，這樣就避開了平方根；所得的導數（derivative）原來是時間的二次式（quadratic），於是距離會否減少的問題，就變成這個二次式有沒有實根的問題。它的判別式（discriminant）給出了答案，而臨界角原來就是正弦值恰好等於 $\frac{2\sqrt{2}}{3}$ 的那個角。這一課的意義遠遠超出拋體運動。許多「顯然」遞增的量其實並非如此，而在大學裏，透過求出導數的正負號來證明某個量是單調（monotonic）的，是證明不等式、證明數列收斂（converge），以及證明微分方程的解趨於穩定的關鍵一步（Section 11.10）。

最後一位乘客。一班列車全部滿座。第一位上車的乘客不理會自己的車票，隨機選了一個座位坐下。其後每一位乘客，如果自己的座位空着就坐自己的座位，否則就在空座位中隨機選一個。最後一位上車的乘客發現自己的座位空着的機會有多大？如果有一百個座位，各種可能性的分支令人眼花繚亂：一位被佔了座位的乘客可以佔去另一位的座位，後者又佔去另一位的，形成一條任意長度的鏈。直接計算看來毫無希望。然而，無

論座位數目是多少，答案都恰好是 $\frac{1}{2}$。

2021 年的一道題目一步一步地得出這個答案：它要你求出兩個和三個座位時的答案，再按第一位乘客所坐的位置取條件（conditioning）來建立遞歸關係，然後用歸納法證明答案永遠是 $\frac{1}{2}$。這是一個完全妥當的證明，但它令你不禁要問，答案爲甚麼如此簡單，而該節接着給出一個一句話的論證來解釋。每當有乘客要隨機選擇時，第一位乘客的座位和最後一位乘客的座位所受的對待完全相同；在其中一個被佔之前，整個過程無法分辨它們。所以最後剩下的空座位是兩者之一的機會均等，而它必定是兩者之一，因為其他每個座位，在最後一位乘客上車之前，都已被它的主人或某位被佔了座位的乘客坐了。像這樣利用對稱（symmetry）的論證，以一個觀察取代好幾頁的計算，貫穿整個大學概率論（Section 12.5）。

綜合來看，這類題目觸及數學學位一年級的每一個主要範疇：數論（number theory）、分析、微積分（calculus）、複數（complex numbers）、代數、線性代數（linear algebra）、微分方程、組合數學、力學、概率與統計。

本書的用途

這個觀察正是本書的緣起。大多數的 STEP 備試，都以通過考試為目標，這十分合理：學習技巧，操練試卷，鍛鍊速度和耐力。本書也以此為目標，而讀完它會令你更擅長應付 STEP。但它真正的目的並不在此，而在於幫助你心懷大學數學去準備 STEP：把每道題目看成一份邀請，而不是一個要清除的障礙，並且循着這份邀請一直走到盡頭。

在 STEP 中表現最好、其後在大學裏如魚得水的學生，幾乎總是那些為數學本身而熱愛數學的人。他們享受難題終於被攻破的那一刻。他們想知道一個技巧為何行得通，而不僅僅是知道它行得通。當答案簡單得令人起疑，例如 $\frac{1}{2}$ 或 $\frac{1}{e}$ 或 $\pi^2/6$，他們會懷疑背後另有更深的道理，並且動身去找。本書想培養的正是這些習慣。每一道題目之所以入選，都是因為它通往一個發現：一個比題目本身更大的想法、定理或思考方式。

思考方式與定理同樣重要。大學數學要求你用學校數學很少用到的方式思考：去證明而不是去計算，用反證法（proof by contradiction）或無窮遞降法（infinite descent）來論證，用兩種方法數算同一樣東西，找出一個永不改變的量，構造一個反例（counterexample），在不把某樣東西找出來的情況下證明它存在。STEP 題目訓練的正是這些習慣，而且往往不加明言。本書的每一節都會點明該節核心的思維技巧，好讓你再次遇上時能夠認出它。

本書的編排

本書共有十三章，按題目所通往的大學數學範疇編排，而不是按題目所出自的 STEP 試卷編排。

- **純粹數學**（第 1 至 9 章）：數論與證明；數列、級數（series）與收斂；函數、不等式與嚴謹性；正正式式的微積分；複數；代數與結構；線性代數與幾何；微分方程與動力系統；以及計數與機會，當中亦取材自概率卷的組合題。
- **力學**（第 10 和 11 章）：守恆定律、碰撞與引力；以及振動、轉動與運動參考系，最後以一組出人意表的軌跡作結。
- **概率與統計**（第 12 和 13 章）：概率，由關於取條件的清晰思考到中心極限定理（central limit theorem）；以及現實世界中的統計與隨機性。

每章以一篇短文開始：這個範疇為何重要，你在大學裏會在其中遇到甚麼，以及貫穿全章的三個大問題。每章以一篇總結收結，當中會重溫這些問題，並附上所用方法的工具箱、通往其他章節的指引，以及入讀大學前的閱讀建議。每章末尾載有練習的簡短答案。第 1 章之前有一個詞彙表，收錄書中那些在學校不大可能接觸過的術語，還有一個列出書中所用數學符號的表；如果有哪個詞或哪個符號難倒了你，就到那裏查一查。

每一節都依循同一個模式。它以*題目*開始：一條連結，通往該道 STEP 題目在官方來源的原文，與當年所出的一字不差。*引子*解釋這道題目為何比表面看來更有趣。*動筆之前*給你第一個立足點。*逐步解題*是分階段的引導式解答，中間穿插着以*問問自己*和*停下來試試*為題的方框，那些都是一位好老師會在那一刻向你提出的問題；它們是用來回答的，不是用來跳過的。*回顧*挑出關鍵的幾步。*背後的數學*是該節的核心：題目背後的大學數學，它的歷史和它的驚喜。*思維技巧*點明這道題目所鍛鍊的思考習慣。*延伸練習*提供後續的題目，當中常有相關的 STEP 題目，而*延伸閱讀*則推介一些書籍和文章，程度是實力強的中學生或大學一年級生應付得來的。

本書詳細討論了 136 道題目，但 STEP 所出的題目超過 1,500 道。網站為其中每一道都提供了詳細解答，由樣本試卷直至 2025 年。它們的寫作精神與本書一致，但篇幅短得多。每一份解答都給出解開題目的關鍵思路、逐部分的完整解答、幾項你可以對答案作出的驗算、容易犯的錯誤，以及該題所觸及的大學數學的指引。如果某道題目在本書中有完整的討論，它的解答會註明，並附上連結。

有幾節用到現行 STEP 課程綱要以外的課題，例如羣（groups）、轉動慣量（moments of inertia）和矩母函數（moment generating functions）。收錄它們，是因為那些題目好得不容割捨，而每一節都會給出它所需要的事實。

學生如何使用本書

- **先做題目。**從該節開首的連結打開題目，閱讀時把它放在手邊，因為解答會不時提到題目的各個部分和用字。每道題目至少認真地花半小時，然後才往下讀。卡住了並不代表失敗；學習正是在這裏發生的。如果你完全卡住了，先讀*動筆之前*，然後再試。

- **回答方框裏的問題。**當正文向你提問時，請停下來，在紙上作答。這些引導式解答，是寫給手裏拿着鉛筆的讀者看的。
- **不要跳過「背後的數學」各節。**它們不在考試範圍之內，卻是這本書存在的理由，也是你日後會記得的東西。
- **各節可按任何次序閱讀。**每一節都自成一體，而交叉引用會指引你到別處的相關想法。跟着你的好奇心走。
- **先做練習，再對答案。**每章末尾的答案很簡短：只給出結果和關鍵的一步，而不是完整的解答過程。
- **把詳細解答當作跳板。**用歷屆試卷練習時，先認真嘗試每道題目，然後才打開網站上的解答，而且只讀到足以讓你重新起步的地方為止。如果關鍵思路已經足夠，就關上解答，自己把題目做完。
- **善用工作紙。**網站為每一節提供可列印的工作紙，載有題目的連結、依次排列的提示、背後的數學部分和練習，方便你離開屏幕來做。
- **在本書以外多讀。**Siklos 的書、STEP 支援計劃和歷屆試卷本身，全部都可以免費取得。每一節的延伸閱讀，都是挑選入讀大學前已能讀懂的材料。

教師如何使用本書

- **作為延伸教材的來源。**每一節都是一堂自成一體的課：題目、一條貫穿題目的引導路徑，以及做完之後一個值得一去的地方。各項思維技巧很適合用作一連串課堂的主題。
- **作為討論指引。**問問自己和停下來試試的方框，是全班或小組討論的現成引子。先提出這些問題，然後才揭示下一階段。
- **配合工作紙使用。**每張工作紙都載有題目的連結、依次排列的提示、背後的數學部分和後續練習；學生可以自行決定要看多少個提示。
- **用於 STEP 備試。**書中收錄了全部三份試卷的題目，由樣本試卷直至 2025 年，其中不少在練習中附有相關題目。各章的總結和工具箱有助學生看到不同課題之間的聯繫。
- **給沒有 STEP 支援的學生。**本書的寫法，是要讓一個有決心的學生能夠獨力讀完。如果你的學校不提供 STEP 備試，本書可以承擔其中大部分的工作。

版權與勘誤

本書中凡屬我本人創作的內容，包括正文、引導式解答、插圖、練習和答案，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、改編和分享，無須徵求許可。STEP 題目本身的版權屬 Cambridge University Press & Assessment 所有，不在該授權條款的涵蓋範圍之內。因此，本書並沒有轉載這些題目。取而代之，每一節都連結到題目的官方來源：樣本試卷和 1987 至 2018 年的試卷，連結到 STEP 支援計劃的題目資料庫；2019 年起的試卷，則連結到 OCR 提供下載的各年試卷。用到其他 STEP 題目的練習，也以同樣的方式連結。

我已盡力確保書中的數學正確無誤，但這樣篇幅的一本書必定會有錯誤。如果你發現錯誤，無論多麼細微，請到本書的 GitHub 儲存庫 `github.com/keithmcnulty/step_book_mcnulty` 以 issue 的形式提出。收到勘誤後，本書的網站會隨即更正。

動筆之前的最後幾句話

Siklos 以米高安哲羅（Michelangelo）的*大衛像*的故事，為他那本書的引言作結：這座雕像是用一塊數十年來無人問津的大理石雕成的。他說，他無法教會你數學，但下苦功一點一點地鑿開題目，會讓你內在的那位數學家顯露出來。這番話我無法說得更好。本書為你提供大理石、一些工具，以及幾幅圖畫，讓你看看別人在石頭裏找到了甚麼。鑿石的功夫，就要靠你自己了。我希望你樂在其中，一如我編寫這本書時那樣樂在其中。

Keith McNulty
2026 年 11 月

詞彙表

大學數學有自己的一套詞彙，其中不少你大概從未見過。本詞彙表收錄了書中那些你在中學不大可能遇過的用語，每個詞附上簡短的解釋，並在有幫助的地方指出哪一節用到或詳細說明了這個概念。這些解釋只為幫你入門，並非精確的定義；凡有嚴格定義的詞，所指的那一節都會給出。

有幾個詞，例如*集合*、*函數*和*證明*，你早已熟悉，但大學裏使用它們時比中學更為講究，所以也一併收錄。詞條內以*斜體*標示的詞，本身另有詞條。

A

絕熱不變量（Adiabatic invariant） 當系統緩慢改變時幾乎保持不變的量。一個分子在緩慢壓縮的氣缸內來回彈跳時，它的速率乘以氣缸長度幾乎固定不變（Section 10.4）。

代數數（Algebraic number） 某個整數系數（可正、可負、可為零，但不全為零）多項式方程的根。$\sqrt{2}$、i 和 $\frac{1}{2}(1+\sqrt{5})$ 都是代數數；不是代數數的數稱為*超越數*（transcendental）（Section 6.3）。

分析（Analysis） 為微積分奠定嚴格基礎的數學分支：極限、連續性、微分、積分和無窮級數，全都精確地定義，並從基本原理出發加以證明。第 2 至 4 章是這門學問的入門。

角動量（Angular momentum） 動量在轉動中的對應量。對於繞軸自轉的物體，它等於*轉動慣量*（moment of inertia）乘以角速度；只要沒有外來的轉動效應，角動量便守恆（Section 11.6）。

算術幾何平均（Arithmetic–geometric mean） 把一對數反覆換成它們的算術平均和幾何平均，由此得到兩個數列，它們的共同極限就是算術幾何平均。其收斂速度快得驚人（Section 2.6）。

結合律（Associative） 若對所有 a、b、c 都有 $(a*b)*c = a*(b*c)$，則運算 $*$ 滿足結合律，括號因而可以省去。數的加法和乘法都滿足結合律；減法則不然。

漸近（Asymptotic） 若兩個式子之比當 $n \to \infty$ 時趨向 1，則稱它們漸近相等，記作 $f(n) \sim g(n)$。斯特靈公式（Stirling's formula）指出 $n! \sim \sqrt{2\pi n}\,(n/\mathrm{e})^n$，儘管兩邊之差會無止境地增大（Section 12.23）。

公理（Axiom） 不加證明而視為出發點的命題，其他一切都由它推導出來。大學代數有很大部分都是從一張簡短的公理清單出發，探究由此可以推出甚麼（Section 7.3、Section 6.5）。

B

基礎步驟（Base case） *歸納法*（induction）證明的第一步：驗證命題對最小的值成立，通常是 $n = 0$ 或 $n = 1$。其後的*歸納步驟*（inductive step）則證明：命題對某個值成立，便對下一個值也成立。

基本比率謬誤（Base-rate fallacy） 只憑證據判斷某事的概率（probability），卻忽略它本來有多常見，這種錯誤就是基本比率謬誤。如果某種疾病足夠罕見，即使準確的測試得出陽性結果，它是虛驚一場的可能性也可以比真正患病更大；實際是否如此，取決於疾病有多罕見和測試有多準確（Section 12.2）。

基（Basis） 一組向量，使得空間中每個向量都能以唯一的方式寫成它們的組合。在平面上，**i** 和 **j** 構成一個基；任何一對不平行的向量也是如此。一個基所含向量的數目就是空間的*維數*（dimension）（Section 6.3）。

貝葉斯定理（Bayes' theorem） 把一個條件概率倒轉過來的法則：$\mathrm{P}(A \mid B) = \mathrm{P}(B \mid A)\,\mathrm{P}(A)/\mathrm{P}(B)$。它告訴你怎樣根據證據修正原有的判斷（Section 12.2）。

偏差（Bias） 若一個*估計量*（estimator）在大量樣本上平均而言偏高或偏低，它便是有偏的；若它的期望值恰好等於所估計的量，它便是*無偏*（unbiased）的（Section 13.3）。

雙射（Bijection） 一一對應：一種把一個集合的每個成員與另一個集合的恰好一個成員配對的規則，而且兩邊都沒有剩餘。兩個有限集合大小相同，恰好就是它們之間存在雙射的時候；無限集合的大小正是這樣定義的（Section 1.8、Section 9.10）。

二元運算（Binary operation） 把兩個對象結合而得出第三個對象的規則，例如數的加法、矩陣的乘法或函數的複合。

界（Bound） 一個數集或數列的上界，是其中沒有任何數超過的數；下界則是其中沒有任何數低於的數。若數列兩者皆有，便稱為*有界*（bounded）（Section 2.2）。

分支過程（Branching process） 一種族羣模型，其中每個個體各自獨立地產生隨機數目的後代。核心問題是這個家族會否絕後，而答案來自*概率生成函數*（probability generating function）（Section 12.15）。

C

變分法（Calculus of variations） 研究這樣一類問題的學問：未知的是整個函數或整條曲線，要選取它使某個量盡可能小或盡可能大，例如最短的路徑、最快的滑道、圍出最大面積的曲線（Section 4.7）。

懸鏈線（Catenary） 懸掛的鏈條所呈的形狀，即 $y = a\cosh(x/a)$ 的圖像。它看似拋物線，其實不是（Section 11.14）。

柯西分佈（Cauchy distribution） 一種連續分佈，圖像呈鐘形，但尾部極厚，以致沒有平均值。從中抽取樣本，樣本的平均值並不比單獨一個觀測值可靠（Section 12.22）。

中心極限定理（Central limit theorem） 這個定理指出：大量獨立、同分佈而方差有限的隨機量之和，經適當平移和縮放後，近似服從正態分佈，不論那個共同的分佈是甚麼。對於分佈各不相同的隨機量也有相應的版本，條件是其中沒有少數幾個量主導整個和。它解釋了為何鐘形曲線無處不在（Section 12.24）。

混沌（Chaos） 確定性系統的一種行為：它對起點極為敏感，以致實際上無法作長期預測——兩個非常接近的起點，會導向截然不同的未來（Section 8.7、Section 8.8）。

封閉（Closed） 若把集合中兩個成員結合起來，所得的總是集合中的另一個成員，則該集合在這個運算下封閉。偶數在加法下封閉；奇數則不然。

交換律（Commutative） 若 $a * b = b * a$ 恆成立，則運算 $*$ 滿足交換律。數的乘法滿足交換律；矩陣的乘法和四元數（quaternions）的乘法則不滿足（Section 7.1、Section 6.7）。

完備性（Completeness） 實數沒有空隙這一性質：每個有上界的遞增數列都收斂於某個實數。有理數並無這個性質，因為一列分數可以逐步逼近 $\sqrt{2}$。分析有很大部分建基於此（Section 2.2）。

同餘（Congruence） $a \equiv b \pmod{n}$，讀作「a 與 b 模 n 同餘」，意思是 a 和 b 除以 n 所得的餘數相同；等價地說，n 整除 $a - b$。用同餘式進行的運算稱為模算術（modular arithmetic）（Section 1.5）。

猜想（Conjecture） 人們相信為真但尚未證明的命題。

守恆定律（Conservation law） 指出某個量（例如能量、動量或角動量）在系統演變過程中保持不變的命題。有了守恆定律，往往不必追蹤細節便能解決問題（Section 10.1）。

連續（Continuous） 若輸入的微小改變只引致輸出的微小改變，則函數是連續的；通俗地說，就是筆不離紙也能畫出它的圖像。精確地說，若當 $x \to a$ 時 $f(x) \to f(a)$，則 f 在 a 連續（Section 3.1、Section 8.8）。

反證法（Contradiction, proof by） 先假設命題不成立，再由此推出不可能的結論，這樣的證明稱為反證法。經典的例子是 $\sqrt{2}$ 為無理數的證明（Section 1.3）。

逆否命題（Contrapositive） 「若 P 則 Q」的逆否命題是「若非 Q 則非 P」。兩個命題在邏輯上等價，而證明逆否命題往往較為容易。它與*逆命題*（converse）「若 Q 則 P」並不相同，後者很可能不成立。

收斂（Converge） 若數列各項能夠任意接近某個固定的數，並一直保持接近，則數列收斂，該數就是它的*極限*（limit）。若級數的*部分和*（partial sums）收斂，則級數收斂。不收斂的數列或級數稱為*發散*（diverge）（Section 2.1、Section 2.3）。

凸（Convex） 若連接函數圖像上任意兩點的弦都位於圖像之上或與圖像重合，則函數是凸的，例如 x^2 或 e^x。許多不等式，包括 AM–GM 不等式，都是凸性的推論（Section 3.3）。

推論（Corollary） 由剛證明的*定理*（theorem）很快便能得出的結果。

相關係數（Correlation） 一個介乎 -1 與 1 之間的數，量度兩個隨機量之間的關係有多接近直線關係。它是*協方差*（covariance）除以兩個標準差所得的值（Section 13.7）。

可數（Countable） 若集合的成員可以排成第一個、第二個、第三個……這樣列下去（可能永無止境），則該集合是可數的。整數是可數的，而令人意外的是，分數也是可數的；實數則不可數（Section 1.8）。

反例（Counterexample） 顯示某個一般性命題不成立的單一例子。不論有多少例子支持某個論斷，一個反例已足以把它推翻。

協方差（Covariance） $\operatorname{Cov}(X,Y) = \operatorname{E}(XY) - \operatorname{E}(X)\operatorname{E}(Y)$，量度兩個隨機量如何一同變化。獨立的隨機量協方差為零，但協方差為零並不代表獨立（Section 12.4、Section 13.7）。

循環（Cycle） 反覆施行某個函數時依次到訪的一組點：$x_1 \mapsto x_2 \mapsto \cdots \mapsto x_n \mapsto x_1$。數 n 是它的*週期*（period），而這些點稱為*週期點*（periodic points）（Section 8.8）。

擺線（Cycloid） 滾動的輪子邊緣上一點所描出的曲線。把它上下倒轉，它既是最速降線，也是無論從哪裏滑下所需時間都相同的曲線（Section 11.4）。

分圓多項式（Cyclotomic polynomial） 以全部*本原*（primitive）n 次*單位根*（roots of unity）為根、且別無其他根的多項式 $\Phi_n(x)$。它的系數都是整數，而且*不可約*（irreducible）；$x^n - 1$ 等於 n 的所有因數 d 所對應的 $\Phi_d(x)$ 之積：例如 $x^4 - 1 = (x-1)(x+1)(x^2+1)$（Section 6.1）。

D

錯排（Derangement） 把一列東西重新排列，使得沒有任何一項留在原來的位置（Section 9.8）。

導子（Derivation） 作用於函數（或其他對象）的運算 D，滿足 $D(f+g) = Df + Dg$ 和乘積法則 $D(fg) = f\,Dg + g\,Df$。求導是標準的例子，而它的許多性質單憑這兩條法則便可推出（Section 6.5）。

行列式（Determinant） 由方陣算出的一個數，量度相應的變換把面積或體積放大多少倍。它等於零，恰好就是矩陣沒有逆矩陣的時候（Section 7.3）。

可微（Differentiable） 若函數的圖像在某點有明確的切線，即定義導數的那個極限存在，則函數在該點可微。每個可微函數都是連續的，但反之不然：$|x|$ 在 0 連續但不可微。

維數（Dimension） 一個基所含向量的數目，即向量空間（vector space）中獨立方向的數目。

發散（Diverge） 見收斂。

動力系統（Dynamical system） 描述系統狀態如何隨時間改變的規則，可以是逐步改變（由 $x_{n+1} = \mathrm{f}(x_n)$ 定義的數列），也可以是連續改變（微分方程）。第 8 章以此為主題。

E

特徵值、特徵向量（Eigenvalue, eigenvector） 若對某個數 λ 有 $\mathbf{Mv} = \lambda\mathbf{v}$，則非零向量 $\mathbf{v}$ 是矩陣 $\mathbf{M}$ 的特徵向量，而該數就是它的特徵值：這個變換只把 $\mathbf{v}$ 拉伸，而不把它轉動（Section 7.2）。

包絡線（Envelope） 與一族曲線中每一條都相切的曲線。以固定速率拋出的球，其所有軌跡的包絡線就是安全拋物線（safety parabola），即球所能到達的區域的邊界（Section 11.9）。

平衡（Equilibrium） 系統不再改變的狀態。若微小的擾動始終保持微小，平衡便是穩定（stable）的；若擾動還會逐漸消失，便是漸近穩定（asymptotically stable）的，有摩擦的擺就是這樣；沒有摩擦而垂直下垂的擺是穩定的，但輕輕一推便會令它永遠擺動下去。若有些微小的擾動會增大，平衡便是不穩定（unstable）的（Section 8.3、Section 11.8）。

估計量（Estimator） 根據數據計算出某個未知量（例如族羣的大小或分佈的平均值）的估計值的規則（Section 13.3）。

歐拉示性數（Euler characteristic） 對於有 V 個頂點、E 條棱和 F 個面的多面體，指 $V - E + F$ 這個數；沒有洞的多面體，這個數永遠是 2（Section 7.9）。

存在性證明（Existence proof） 證明某物存在，卻不一定說明怎樣把它找出來的證明。鴿巢原理（pigeonhole principle）和介值定理（intermediate value theorem）都能給出存在性證明。

期望值（Expectation） 隨機量 X 的平均值，記作 $\mathrm{E}(X)$：即多次獨立重複試驗的長期平均。期望值具有可加性，$\mathrm{E}(X+Y) = \mathrm{E}(X) + \mathrm{E}(Y)$，不論 X 與 Y 是否獨立（Section 9.11）。

極端原理（Extremal principle） 一種策略：考察某類對象中最小、最大、最先或最後的一個，並證明它具有某種性質，從而得出所求的結果或導出矛盾（Section 1.3）。

F

域（Field） 可以進行加、減、乘、除（除數不為零）而且所有慣常法則都成立的數系。有理數、實數和複數都是域；模質數的整數也是。整數不是域，因為並非總能相除（Section 6.3）。

域擴張（Field extension） 包含一個較小的域的較大的域，例如所有形如 $a + b\sqrt{2}$（其中 a、b 為有理數）的數，它們包含了有理數。域擴張的次數（degree），就是把它看成較小的域上的向量空間時的維數（Section 6.3）。

不動點（Fixed point） 滿足 $\mathrm{f}(x) = x$ 的點 x，即函數讓它留在原處的點。在圖像上，不動點就是 $y = \mathrm{f}(x)$ 與 $y = x$ 相交之處（Section 2.2、Section 8.8）。

參考系（Frame of reference） 量度位置和運動時所採取的觀點。慣性（inertial）參考系是牛頓（Newton）第一定律成立的參考系；在加速的參考系中，例如正在加速的火車上，看起來會多出一些力（Section 11.5）。

函數（Function） 一種規則，把一個集合（定義域（domain））的每個成員對應到另一個集合（上域（codomain））的恰好一個成員。在大學裏，函數並不等同於公式：許多重要的函數沒有公式，而有些公式並不定義函數。

函數方程（Functional equation） 以函數為未知量的方程，例如 $\mathrm{f}(x+y) = \mathrm{f}(x)\mathrm{f}(y)$（Section 3.8）。

代數基本定理（Fundamental theorem of algebra） 每個次數至少為 1 的複系數多項式都有一個複數根，因而可以在複數範圍內完全分解成一次因式。

G

伽瑪函數（Gamma function） 把階乘推廣至非整數的函數 Γ：對整數 n 有 $\Gamma(n+1) = n!$，而且 $\Gamma(\frac{1}{2}) = \sqrt{\pi}$，在某種意義上這是說 $(-\frac{1}{2})! = \sqrt{\pi}$（Section 4.4）。

高斯整數（Gaussian integer） 形如 $a + bi$ 的複數，其中 a 和 b 為整數。它們的表現與普通整數十分相似，有自己的質數，也有唯一分解（Section 1.6）。

生成函數（Generating function） 以某個數列各項為系數的冪級數 $a_0 + a_1 x + a_2 x^2 + \cdots$。對級數所作的運算（例如乘法）對應於對數列所作的有用運算（Section 9.4）。*概率生成函數*是對取整數值的隨機量的各個概率這樣做（Section 9.6、Section 12.15）；*矩生成函數*（moment generating function）則是它的近親（Section 12.24）。

羣（Group） 一個集合連同一個*二元運算*，這個運算滿足*結合律*，有*單位元*（identity），並且每個元素都有*逆元*（inverse）。一個圖形的所有對稱構成一個羣，非零複數在乘法下也構成一個羣。羣是描述對稱的語言（Section 6.6、Section 6.7）。

H

雙曲幾何（Hyperbolic geometry） 一種幾何，其中過已知直線外一點有無窮多條直線永不與該直線相交，而三角形的內角和小於 π。配上適當的距離量度方法的圓盤，就是它的一個模型（Section 5.8）。

I

單位元；恆等式（Identity） 運算 $*$ 的單位元是這樣的元素 e：對每個 a 都有 $e * a = a * e = a$。加法的單位元是 0，乘法的是 1，矩陣乘法的是單位矩陣 $\mathbf{I}$。另一個意思是恆等式，即對變量的每個值都成立的等式，例如 $\cos^2\theta + \sin^2\theta \equiv 1$。

當且僅當（If and only if） 「P 當且僅當 Q」，常寫作「P iff Q」或 $P \iff Q$，意思是 P 和 Q 要麼同時為真，要麼同時為假。證明它需要兩個論證，每個方向各一個。

獨立（Independent） 若知道其中一個事件已經發生並不改變另一個事件的概率，則兩個事件獨立：$\mathrm{P}(A \cap B) = \mathrm{P}(A)\mathrm{P}(B)$。三個或更多事件可以兩兩獨立，卻並非整體獨立（Section 12.4）。

歸納法（Induction） 證明某個命題對每個整數 n 都成立的方法：先證明它對第一個值成立（*基礎步驟*），再證明只要它對某個值成立，便對下一個值也成立（*歸納步驟*）。在*強歸納法*（strong induction）中，歸納步驟可以使用命題對所有較小的值都成立這一點，而不只是對前一個值（Section 1.2）。

無窮遞降法（Infinite descent） 反證法的一種形式：證明某個整數問題的任何一個

解都會引出一個更小的解，從而得到一列無窮遞減的正整數，而這是不可能的（Section 1.5）。

介值定理（Intermediate value theorem） 若連續函數取到 a 和 b 這兩個值，它便取到兩者之間的每一個值。因此，變號的連續函數必定有根（Section 8.8）。

不變量（Invariant） 在某個過程進行期間，或在某種變換之下，保持不變的量。要證明某事不可能，或者不必追蹤每一步便解決問題，尋找不變量是最有力的方法之一（Section 7.7、Section 10.4）。

逆元；反函數（Inverse） 對於具有單位元 e 的運算，元素 a 的逆元是滿足 $a * b = b * a = e$ 的元素 b。反函數則把函數的作用還原：$\mathrm{f}^{-1}(\mathrm{f}(x)) = x$。

無理數（Irrational） 不能寫成分數 $\frac{p}{q}$（其中 p 和 q 為整數，$q \neq 0$）的實數，例如 $\sqrt{2}$、π 或 e（Section 1.4）。

不可約（Irreducible） 若多項式不能寫成兩個次數較低、系數屬同一類的多項式之積，則它是不可約的。$x^2 - 2$ 在有理數範圍內不可約，但在實數範圍內可分解為 $(x - \sqrt{2})(x + \sqrt{2})$（Section 6.1）。

同構（Isomorphic） 若兩個結構（例如兩個羣）之間存在一個雙射，使它們的運算完全對應，則這兩個結構同構：它們是同一個結構，只是標籤不同（Section 6.7、Section 6.8）。

迭代（Iteration） 反覆施行同一個函數，於是 $x_1 = \mathrm{f}(x_0)$，$x_2 = \mathrm{f}(x_1)$，如此類推。數列 x_0、x_1、x_2、… 稱為 x_0 的軌道（orbit）（Section 8.7）。

L

大數定律（Law of large numbers） 這個定理指出：只要期望值存在，一個隨機量多次獨立重複所得的平均值便趨向它的期望值（Section 12.22）。

引理（Lemma） 主要為了證明某個更大的定理而先行證明的結果。

極限（Limit） 數列或函數任意接近、並一直保持接近的那個值。把這個想法精確化，是十九世紀數學的偉大成就之一，也是分析的起點（Section 2.1）。

極限環（Limit cycle） 系統運動中的一個閉合迴路，鄰近的運動會螺旋地趨近（或遠離）它，例如心跳或時鐘的穩定節奏（Section 8.4）。

線性映射（Linear map） 向量空間之間滿足 $T(\mathbf{u}+\mathbf{v}) = T(\mathbf{u}) + T(\mathbf{v})$ 和 $T(\lambda\mathbf{u}) = \lambda T(\mathbf{u})$ 的函數 T。每個線性映射都可以用矩陣表示，而線性代數就是研究線性映射的學問（Section 7.2）。

邏輯斯諦映射（Logistic map） 規則 $u_{n+1} = ru_n(1-u_n)$，是受食物供應限制的族羣的簡單模型，也是混沌的標準例子（Section 8.7）。

M

馬可夫鏈（Markov chain） 一種在各個狀態之間逐步轉移的隨機過程，其中下一步的概率只取決於當前的狀態，而與過程怎樣到達這個狀態無關（Section 12.16）。

最大似然法（Maximum likelihood） 估計未知量的一種方法：選取使觀測所得數據出現的概率最大的那個值（Section 13.2）。

無記憶性（Memoryless） 若不論你已經等了多久，餘下的等候時間的分佈都與開始時相同，則這個隨機等候時間具有無記憶性。指數分佈是唯一具有無記憶性的連續分佈（Section 12.18）。

最小多項式（Minimal polynomial） 以某個已知代數數為根、系數為有理數而首項系數為 1 的多項式之中，次數最低的一個。$\sqrt{2}$ 的最小多項式是 $x^2 - 2$（Section 6.3）。

模算術（Modular arithmetic） 見同餘。

轉動慣量（Moment of inertia） 質量在轉動中的對應量：量度要改變物體繞某軸轉動的快慢有多困難。它取決於質量離軸有多遠，而不只取決於質量有多少（Section 11.7）。

單調（Monotonic） 若數列或函數從不下降（遞增）或從不上升（遞減），則它是單調的。有界的單調數列必定收斂（Section 2.2、Section 11.10）。

N

必要及充分（Necessary and sufficient） 若沒有 P 則 Q 不可能成立（$Q \Rightarrow P$），則 P 是 Q 的必要條件；若 P 保證 Q 成立（$P \Rightarrow Q$），則它是充分條件。既必要又充分的條件與 Q 等價。

簡正模態（Normal modes） 系統的一些特殊振動形態，其中每個部分都以相同的頻率同步運動。系統的任何振動都是它的簡正模態的組合（Section 11.2）。

O

階（Order） 羣中元素 g 的階，是使 g^n 等於單位元的最小正整數 n：即必須把它施行多少次才回到起點（Section 7.4）。若不存在這樣的 n，例如整數加法羣中的數 1，則稱 g 的階為無限（infinite order）。有限羣的階是它所含元素的數目。

正交（Orthogonal） 即垂直。在大學裏，這個詞的用法推廣至向量以外：若在適當的範圍上 $\int \mathrm{f}(x)\mathrm{g}(x)\,\mathrm{d}x = 0$，則兩個函數 f 和 g 可稱為正交（Section 8.6）。

P

參數（Parameter） 在某一個問題之內固定不變、但在不同問題之間可以改變的量，例如邏輯斯諦映射中的 r。許多 STEP 題目都會問答案怎樣隨參數而變。

部分和（Partial sum） 級數首 n 項之和 $a_1 + a_2 + \cdots + a_n$。若部分和收斂，則級數收斂（Section 2.3）。

鴿巢原理（Pigeonhole principle） 若把多於 n 件物件放進 n 個盒子，則必有某個盒子裝有至少兩件。它雖然簡單，卻能證明一些出人意表的存在性結果（Section 9.2）。

泊松過程（Poisson process） 一種模型，描述在時間或空間中以恆定的平均率隨機而獨立地發生的事件：放射性衰變、電話來電、天上的星星（Section 12.19、Section 12.18）。

本原單位根（Primitive root of unity） 一個 n 次單位根，其各次冪恰好給出全部 n 個 n 次單位根；等價地說，對任何較小的 m，它都不是 m 次單位根（Section 6.1）。

概率密度函數（Probability density function） 對於連續的隨機量，指這樣的函數 f：它在某個區間上的積分，就是隨機量落在該區間內的概率。

證明（Proof） 從定義、公理和已經證明的結果出發，說明某個命題為真的完整邏輯論證。在大學裏，即使一段計算得出正確答案，除非每一步都有理據，否則不算證明。

Q

四元數（Quaternions） 哈密頓（William Rowan Hamilton）於 1843 年發現的數系，它把複數加以擴充，包含 -1 的三個平方根，記作 i、j、k，滿足 $\mathrm{i}^2 = \mathrm{j}^2 = \mathrm{k}^2 = \mathrm{ijk} = -1$。四元數的乘法不滿足交換律。它們用於描述三維空間中的旋轉（Section 6.7、Section 7.5）。

R

隨機變量（Random variable） 其值由隨機試驗的結果決定的量，例如擲十次硬幣所得正面的次數。

隨機遊走（Random walk） 由接連的隨機步伐組成的路徑，例如一枚棋子按擲硬幣的結果向左或向右移動一格（Section 9.10、Section 12.17）。

快度（Rapidity） 狹義相對論中量度快慢的一個量；與速度不同，當一個運動與另一個同方向的運動合成時，快度只需直接相加。若 c 為光速，則速度 v 對應的快度為 $\operatorname{artanh}(v/c)$（Section 7.6）。

遞推關係（Recurrence relation） 以數列中較前的項表示每一項的等式，例如斐波那契（Fibonacci）數所滿足的 $F_{n+1} = F_n + F_{n-1}$。

共振（Resonance） 系統受到週期性的推動，而推動的頻率等於或接近系統本身的固有頻率時所產生的強烈反應，就像配合鞦韆的節奏去推鞦韆上的小孩那樣（Section 11.1）。

黎曼積分（Riemann integral） 積分的精確定義：把積分定為許多細長矩形面積之和的極限（當這個極限存在時）（Section 4.1）。

嚴謹性（Rigour） 一種論證的標準：每一步都要有理據，不會因為某事看來顯然、或者在你試過的例子中都成立，便視為理所當然（第 3 章）。

單位根（Root of unity） 對某個正整數 n 滿足 $z^n = 1$ 的複數 z。全部 n 次單位根是 $e^{2\pi ik/n}$，其中 $k = 0、1、\ldots、n-1$，它們在單位圓上構成一個正 n 邊形（Section 5.2）。

S

樣本空間（Sample space） 一個隨機試驗所有可能結果組成的集合。事件是它的子集（Section 12.1）。

集合（Set） 一批對象（稱為它的*元素*（elements））合起來看成的單一整體。元素的次序無關重要，重複與否也無關重要：$\{1,2,3\}$ 和 $\{3,1,2,2\}$ 是同一個集合。

簡諧運動（Simple harmonic motion） 加速度與相對某固定點的位移成正比、並且指向該點的運動，$\ddot{x} = -\omega^2 x$。它的解是 $x = A\cos(\omega t + \varepsilon)$，而週期 $2\pi/\omega$ 與擺動的幅度無關（Section 11.4）。

穩定、不穩定（Stable, unstable） 見*平衡*。

斯特靈公式（Stirling's formula） 近似式 $n! \approx \sqrt{2\pi n}\,(n/e)^n$，其百分誤差隨 n 增大而趨向零（Section 12.23）。

子羣（Subgroup） *羣*的一個子集，它在同一運算下本身也是一個羣（Section 6.6）。

對稱性論證（Symmetry argument） 這樣的一種論證：由於問題中沒有任何東西能把兩個量區分開來，所以它們必定相等。一個好的對稱性論證可以取代好幾頁的計算（Section 12.5）。

T

定理（Theorem） 已獲證明的重要命題。較次要的結果常稱為*命題*（propositions），在通往定理的途中證明的稱為*引理*，而很快便能得出的結果則稱為*推論*。

跡（Trace） 方陣主對角線上各元素之和。與行列式一樣，把矩陣改用另一個基來表示時，它並不改變；它還等於各特徵值之和（Section 7.4、Section 7.7）。

超越數（Transcendental） 不是代數數的數：它不是任何整數系數（不全為零）多項式方程的根。e 和 π 都是超越數，不過要證明這一點並不容易（Section 1.4）。

三角不等式（Triangle inequality） $|a+b| \leq |a|+|b|$，對實數、複數或向量都成立：三角形的一邊永遠不會比另外兩邊加起來更長。

U

唯一分解（Unique factorisation） 每個大於 1 的整數都只能以一種方式寫成質數之積（不計因數的次序）這一性質。它並不如看起來那麼顯然：在有些數系中它並不成立（Section 1.7）。

唯一性定理（Uniqueness theorem） 指出某個問題只有一個解的定理。對於微分方程，當方程的性質不夠良好時，唯一性可能失效，以致同一個起點會導向幾個不同的未來（Section 8.1）。

V

向量空間（Vector space） 一批稱為向量的對象，它們可以相加，也可以乘以數，並遵從平面上箭頭所滿足的那些熟悉法則。多項式、函數以及許多微分方程的解都構成向量空間，所以線性代數的方法對它們都適用（Section 6.3）。

W

良定義（Well-defined） 若某個對象或規則的定義確實指明了單獨一個毫不含糊的東西，例如不取決於過程中所作的任意選擇，它便是良定義的。

良序原理（Well-ordering principle） 每個非空的正整數集合都有最小的成員。它與歸納法原理等價，而每一個以「取最小的反例」開始的論證，背後都是這個原理（Section 1.3）。

不失一般性（Without loss of generality） 常簡寫為 WLOG 的一句用語，意思是論證只考慮其中一種情況，因為其他情況除了要重新標記之外並無分別。「不失一般性，設 $a \leq b$」的意思是：如果 $a > b$，把兩者的名稱對調便可。

數學符號

數學寫作往往以寥寥幾個記號承載豐富的意思，而大學講師所用的許多符號，在中學課本裏甚少出現。下面各表按主題列出本書用到的符號，並說明每個符號的含義和讀法。入讀大學後不出幾個星期，其中大部分都會變得習以為常；在此之前，這裏可供你隨時查閱。

有兩項慣例值得一開始就知道。第一，本書依照 STEP 的做法，把某些具有固定含義的字母印成正體而非斜體：e 表示自然對數的底，i 表示 $\sqrt{-1}$，dy/dx 中的 d，以及函數的名稱，例如 f(x) 和 P(A)。斜體字母則表示變數。第二，同一個符號在不同場合有時會有不同的意思；$\equiv$ 和 $\sim$ 是最常見的例子，表中會在相應位置指出。

邏輯與證明

符號	含義	例子或讀法
$P \Rightarrow Q$	P 蘊含（implies）Q：若 P 成立，則 Q 亦成立	$x = 2 \Rightarrow x^2 = 4$；「$P$ 蘊含 Q」
$P \Leftarrow Q$	P 由 Q 所蘊含	「只要 Q 成立，P 就成立」
$P \iff Q$	P 當且僅當（if and only if）Q：兩者互相蘊含	$x^2 = 4 \iff x = \pm 2$；亦寫作「iff」
$\forall$	對所有、對每一個	$\forall x \in \mathbb{R},\ x^2 \geq 0$
$\exists$	存在	$\exists n \in \mathbb{Z}$ 使得 $n^2 = 9$
$\neg P$	非 P，即 P 的否定（negation）	「$\forall x,\ P(x)$」的否定是「$\exists x$ 使得 $\neg P(x)$」
$:=$	定義為	f$(x) := x^2 + 1$
■	證明完畢；亦寫作 QED	源自拉丁文 *quod erat demonstrandum*，意即「這就是所要證明的」

集合與數系

符號	含義	例子或讀法
$\{\dots\}$	由所列元素組成的集合（set）	$\{1,2,3\}$
$\{x : \dots\}$	所有滿足 … 的 x 所組成的集合；亦寫作 $\{x \mid \dots\}$	$\{x : x^2 < 2\}$；「滿足 $x^2<2$ 的 x 所組成的集合」
$x \in A$	x 是 A 的元素（element）	$\sqrt{2} \in \mathbb{R}$；「$\sqrt{2}$ 屬於 $\mathbb{R}$」
$x \notin A$	x 不是 A 的元素	$\sqrt{2} \notin \mathbb{Q}$
$A \subseteq B$	A 是 B 的子集（subset）：A 的每個元素都在 B 之內	$\mathbb{Z} \subseteq \mathbb{Q}$
$A \cup B$	併集（union）：屬於 A 或 B（或同時屬於兩者）的元素	$\{1,2\} \cup \{2,3\} = \{1,2,3\}$
$A \cap B$	交集（intersection）：同時屬於 A 和 B 的元素	$\{1,2\} \cap \{2,3\} = \{2\}$
$A \setminus B$	屬於 A 但不屬於 B 的元素	$\{1,2,3\} \setminus \{2\} = \{1,3\}$
A'	補集（complement）：所有不屬於 A 的元素	亦寫作 A^{c} 或 $\bar{A}$
$\varnothing$	空集（empty set），不含任何元素	$\{x \in \mathbb{R} : x^2 < 0\} = \varnothing$
$\lvert A \rvert$	有限集 A 的元素個數	$\lvert\{a,b,c\}\rvert = 3$
$A \times B$	所有數對 (a,b) 所組成的集合，其中 $a \in A$，$b \in B$	$\mathbb{R} \times \mathbb{R}$ 就是平面，亦寫作 $\mathbb{R}^2$
$\mathbb{N}$	自然數（natural numbers）$1,2,3,\dots$	有些作者把 0 也算在內
$\mathbb{Z}$	整數（integers）$\dots,-2,-1,0,1,2,\dots$	源自德文 *Zahlen*，意即「數」
$\mathbb{Q}$	有理數（rational numbers），即分數	源自 *quotient* 一字
$\mathbb{R}$	實數（real numbers）	
$\mathbb{C}$	複數（complex numbers）	
$\mathbb{H}$	四元數（quaternions）	以哈密頓（Hamilton）命名（Section 6.7）
$\mathbb{Z}[\mathrm{i}]$	形如 $a+b\mathrm{i}$ 的數，其中 $a,b \in \mathbb{Z}$	高斯整數（Gaussian integers）（Section 1.6）
$[a,b]$	閉區間（closed interval）：$a \le x \le b$	包含端點
(a,b)	開區間（open interval）：$a < x < b$	不包含端點；亦可表示數對或點

數、整除性與近似

符號	含義	例子或讀法
$a \mid b$	a 整除（divides）b：b 是 a 的整數倍	$3 \mid 12$；「3 整除 12」
$a \nmid b$	a 不整除 b	$5 \nmid 12$
$a \equiv b \pmod{n}$	a 與 b 除以 n 所得的餘數相同	$17 \equiv 2 \pmod{5}$；「17 與 2 模 5 同餘（congruent）」
$\equiv$	（不帶模時）恆等於：對所有值都成立	$(x+1)^2 \equiv x^2 + 2x + 1$
$\lfloor x \rfloor$	下取整（floor）：$\le x$ 的最大整數	$\lfloor 3.7 \rfloor = 3$，$\lfloor -3.7 \rfloor = -4$
$\lceil x \rceil$	上取整（ceiling）：$\ge x$ 的最小整數	$\lceil 3.2 \rceil = 4$
$n!$	n 的階乘（factorial），即 $1 \times 2 \times \cdots \times n$	按慣例 $0! = 1$
$\binom{n}{r}$	二項式係數（binomial coefficient）$\frac{n!}{r!\,(n-r)!}$	讀作「n choose r」；亦寫作 ${}^{n}\mathrm{C}_{r}$
$\approx$	約等於	$\pi \approx 3.1416$
$\sim$	漸近於（asymptotic to）：兩者之比趨向 1	$n! \sim \sqrt{2\pi n}\,(n/\mathrm{e})^n$（但另見概率部分）
$\propto$	與……成正比	$v \propto 1/d$
$\pm, \mp$	正負；負正	$\cos(A \pm B) = \cos A \cos B \mp \sin A \sin B$
$0.\overline{01}$	循環數字	$0.\overline{01} = 0.010101\ldots$

函數

符號	含義	例子或讀法
$\mathrm{f} : A \to B$	f 是由集合 A 到集合 B 的函數（function）	$\mathrm{f} : \mathbb{R} \to \mathbb{R}$
$x \mapsto \ldots$	x 被送到 ...	$x \mapsto x^2 - 2$；「x 映射到 $x^2 - 2$」
$\mathrm{f} \circ \mathrm{g}$	複合（composition）：先作 g，再作 f	$(\mathrm{f} \circ \mathrm{g})(x) = \mathrm{f}(\mathrm{g}(x))$
f^{-1}	f 的反函數（inverse function）	並非 $1/\mathrm{f}$
f^n	在本書中，通常指把 f 施行 n 次	$\mathrm{f}^3(x) = \mathrm{f}(\mathrm{f}(\mathrm{f}(x)))$；但 $\sin^2 x$ 表示 $(\sin x)^2$
$\lvert x \rvert$	模（modulus），或絕對值	$\lvert -3 \rvert = 3$
$\exp x$	指數函數（exponential function）e^x	
$\ln x$	自然對數（natural logarithm），以 e 為底	在大學裏常寫作 $\log x$
sec, cosec, cot	$1/\cos$, $1/\sin$, $1/\tan$	
arcsin, arccos, arctan	反三角函數	亦寫作 $\sin^{-1}$ 等
sinh, cosh, tanh	雙曲函數（hyperbolic functions）：$\sinh x = \frac{1}{2}(\mathrm{e}^x - \mathrm{e}^{-x})$，$\cosh x = \frac{1}{2}(\mathrm{e}^x + \mathrm{e}^{-x})$，$\tanh = \sinh / \cosh$	讀作「shine」（或「sinch」）、「cosh」和「than」（或「tanch」）
artanh	tanh 的反函數	$\operatorname{artanh} x = \frac{1}{2} \ln \frac{1+x}{1-x}$
max, min	當中最大者、最小者	$\max(3, 7) = 7$
$\Gamma(x)$	伽瑪函數（gamma function）	$\Gamma(n+1) = n!$（Section 4.4）

求和、極限與微積分

符號	含義	例子或讀法
$\sum$	求和	$\sum_{k=1}^{n} k = \frac{1}{2}n(n+1)$
$\prod$	求積	$\prod_{k=1}^{n} k = n!$
∞	無窮大：並不是一個數，而是「沒有界限」的簡寫	$\sum_{k=1}^{\infty} \frac{1}{k^2} = \frac{\pi^2}{6}$
$x \to a$	x 趨向 a	當 $n \to \infty$ 時，$\frac{1}{n} \to 0$
$\lim$	極限（limit）	$\lim_{x \to 0} \frac{\sin x}{x} = 1$
$f'(x)$, $f''(x)$	一階及二階導數（derivative）	讀作「f dash」或「f prime」
$\dot{x}$, $\ddot{x}$	對時間的導數	$\dot{x} = \frac{dx}{dt}$；讀作「x dot」、「x double dot」
$\frac{\partial f}{\partial x}$	偏導數（partial derivative）：對 x 求導，其他變數保持不變	讀作「partial f by partial x」
$\int_a^b$	由 a 到 b 的定積分（definite integral）	
$\oint$	沿閉合迴路的積分	$\oint p\,dq$（Section 10.4）
δx, Δx	x 的微小變化；x 的變化	
ε, δ	分析學中慣常用來表示細小正數的記號	「對每個 $\varepsilon > 0$，都存在 $\delta > 0$ 使得 ...」

複數

符號	含義	例子或讀法
i	-1 的一個平方根	工程師寫作 j
$\operatorname{Re} z$, $\operatorname{Im} z$	實部與虛部	若 $z = 3 + 4\mathrm{i}$，則 $\operatorname{Re} z = 3$，$\operatorname{Im} z = 4$
z^*, $\bar{z}$	共軛複數（complex conjugate）	$(3+4\mathrm{i})^* = 3 - 4\mathrm{i}$；兩種記法都有人使用
$\lvert z \rvert$	模：與 0 的距離	$\lvert 3 + 4\mathrm{i} \rvert = 5$
$\arg z$	輻角（argument）：由正實軸量起的角	$\arg \mathrm{i} = \frac{\pi}{2}$
$\mathrm{e}^{\mathrm{i}\theta}$	$\cos\theta + \mathrm{i}\sin\theta$	$\mathrm{e}^{\mathrm{i}\pi} = -1$
ω	慣常表示 1 的一個複立方根（cube root of unity）	$\omega = \mathrm{e}^{2\pi\mathrm{i}/3}$，$1 + \omega + \omega^2 = 0$

向量與矩陣

符號	含義	例子或讀法
$\mathbf{a}$, $\overrightarrow{AB}$	向量（vector）；由 A 到 B 的向量	手寫時作 $\underline{a}$
$\hat{\mathbf{n}}$	單位向量（unit vector），長度為 1	讀作「n hat」
$\lvert\mathbf{a}\rvert$	$\mathbf{a}$ 的長度（大小）	
$\mathbf{a}\cdot\mathbf{b}$	純量積（scalar product），又稱點積（dot product）	$\mathbf{a}\cdot\mathbf{b} = \lvert\mathbf{a}\rvert\lvert\mathbf{b}\rvert\cos\theta$
$\mathbf{a}\times\mathbf{b}$	向量積（vector product），又稱叉積（cross product）	同時垂直於 $\mathbf{a}$ 和 $\mathbf{b}$
$\mathbf{M}$	矩陣（matrix）	
$\mathbf{I}$	單位矩陣（identity matrix）	對每個 $\mathbf{v}$ 都有 $\mathbf{Iv} = \mathbf{v}$
$\mathbf{M}^{-1}$	逆矩陣（inverse matrix）	$\mathbf{M}\mathbf{M}^{-1} = \mathbf{I}$
$\det\mathbf{M}$	$\mathbf{M}$ 的行列式（determinant）	$\det\begin{pmatrix} a & b \\ c & d \end{pmatrix} = ad - bc$
$\operatorname{tr}\mathbf{M}$	跡（trace）：對角線上各元素之和	$\operatorname{tr}\begin{pmatrix} a & b \\ c & d \end{pmatrix} = a + d$
λ	慣常表示特徵值（eigenvalue）	$\mathbf{Mv} = \lambda\mathbf{v}$

概率與統計

符號	含義	例子或讀法
$\mathrm{P}(A)$	事件 A 的概率（probability）	
$\mathrm{P}(A \mid B)$	已知 B 發生時 A 的概率	$\mathrm{P}(A \mid B) = \mathrm{P}(A \cap B)/\mathrm{P}(B)$；讀作「$A$ given B」
$\mathrm{E}(X)$	X 的期望值（expectation），即平均值	亦寫作 μ
$\mathrm{Var}(X)$	X 的方差（variance）	$\mathrm{E}(X^2) - \mathrm{E}(X)^2$；亦寫作 σ^2
$\mathrm{Cov}(X, Y)$	協方差（covariance）	$\mathrm{E}(XY) - \mathrm{E}(X)\mathrm{E}(Y)$
$\mathrm{Corr}(X, Y)$	相關係數（correlation）	$\mathrm{Cov}(X, Y)/\sqrt{\mathrm{Var}(X)\mathrm{Var}(Y)}$；亦寫作 ρ
$X \sim \ldots$	X 的分佈為 ...	$X \sim \mathrm{B}(16, \frac{1}{2})$；「$X$ 的分佈為 ...」
$\mathrm{B}(n, p)$	二項分佈（binomial distribution）	n 次試驗中成功的次數
$\mathrm{N}(\mu, \sigma^2)$	正態分佈（normal distribution），平均值為 μ，方差為 σ^2	
$\mathrm{f}(x), \mathrm{F}(x)$	概率密度函數（probability density function）；累積分佈函數（cumulative distribution function）	$\mathrm{F}(x) = \mathrm{P}(X \leq x)$
$\mathrm{G}(t)$	概率生成函數（probability generating function）	$\mathrm{G}(t) = \mathrm{E}(t^X)$
$\mathrm{M}_X(\theta)$	矩生成函數（moment generating function）	$\mathrm{M}_X(\theta) = \mathrm{E}(\mathrm{e}^{\theta X})$
$\bar{x}$	樣本的平均值	$\bar{x} = \frac{1}{n}\sum x_i$
$\hat{\theta}$	未知量 θ 的估計值	讀作「theta hat」

希臘字母

希臘字母在數學中隨處可見，而且往往有約定俗成的含義。下表列出各字母的小寫和大寫形式（與羅馬字母外形相同的大寫字母，例如 alpha 的大寫 A，不會用作獨立的符號，故此從略）、名稱，以及一些常見用途。

字母	名稱	常見用途
α	alpha	角；方程的根
β	beta	角；方程的根；貝塔函數（beta function）
γ, Γ	gamma	歐拉（Euler）常數 $\gamma \approx 0.5772$；伽瑪函數 Γ；在相對論中，$\gamma = 1/\sqrt{1 - v^2/c^2}$
δ, Δ	delta	微小的量；Δ 表示變化，或表示三角形、判別式（discriminant）
ε, ϵ	epsilon	微小的正數量；在力學中表示相位角
ζ	zeta	黎曼（Riemann）zeta 函數，當 $s > 1$ 時 $\zeta(s) = \sum_{n=1}^{\infty} n^{-s}$（以更高深的方法可推廣至 $s = 1$ 以外的其他值）
η	eta	效率；微小的量
θ, Θ	theta	角；統計學中的未知參數
ι	iota	甚少使用
κ	kappa	曲率；常數
λ, Λ	lambda	特徵值；參數，例如泊松（Poisson）分佈的率；波長
μ	mu	平均值；摩擦係數
ν	nu	頻率
ξ, Ξ	xi	變數或隨機量
π, Π	pi	$\pi \approx 3.14159$；Π 表示求積
ρ	rho	密度；相關係數；與原點的距離
σ, Σ	sigma	標準差；Σ 表示求和
τ	tau	時間；力矩
υ, Υ	upsilon	甚少使用
ϕ, φ, Φ	phi	角；黃金比 $\phi = \frac{1}{2}(1 + \sqrt{5})$；快度（rapidity）；$\Phi$ 表示標準正態分佈函數
χ	chi	χ^2（卡方，chi-squared）分佈
ψ, Ψ	psi	量子力學中的波函數
ω, Ω	omega	角速度或角頻率；1 的立方根；Ω 表示樣本空間

力學

符號	含義	例子或讀法
g	重力加速度	在地球表面約為 $9.8\,\mathrm{m\,s^{-2}}$
G	萬有引力常數	兩個質量之間的力為 Gm_1m_2/r^2
e	恢復係數（coefficient of restitution）（以斜體印刷，切勿與 e 混淆）	沿碰撞線方向，分離速率 $= e\times$ 接近速率
ω	角速度；角頻率	簡諧運動（simple harmonic motion）中 $\ddot{x} = -\omega^2 x$
I	轉動慣量（moment of inertia）	轉動動能 $\frac{1}{2}I\omega^2$（Section 11.7）
c	光速	約為 $3 \times 10^8\,\mathrm{m\,s^{-1}}$
$\hbar$	普朗克（Planck）常數除以 2π	讀作「h bar」（Section 8.6）

章

1

數論與證明

整數是數學中最古老的對象，卻至今仍藏着數學中一些最深奧的祕密。

本章內容

1.1	質數工廠	STEP II 2002
1.2	質數分佈得有多稀疏？	STEP III 2016
1.3	最小反例	STEP III 2015
1.4	e 是無理數，e^2 也是	STEP III 2024
1.5	無窮遞降法	STEP III 2022
1.6	兩個平方數之和	STEP II 2023
1.7	因數分解失效之時	STEP I 2016
1.8	點算有理數	STEP III 2020
1.9	三次曲線的切線	STEP I 2009
A	練習答案	第 99 頁

為甚麼是整數？

> 數學是科學的皇后，而數論（*number theory*）是數學的皇后。
>
> — 高斯（Carl Friedrich Gauss）

整數是每個人最先接觸到的數學對象。你還未識字，就已經學會數數。然而，關於整數的最簡單的問題，卻可以難得出奇。每一個大於 2 的偶數，是否都是兩個質數（prime）之和？沒有人知道：哥德巴赫（Christian Goldbach）在 1742 年提出這個問題，電腦已經驗證到 4×10^{18} 以外，但至今仍然沒有人能夠證明。像 11 和 13，或 101 和 103 這樣相差 2 的質數對，是否有無窮多對？這同樣沒有人知道。不過在 2013 年，一位幾乎寂寂無名的數學家張益唐（Yitang Zhang）震驚了全世界：他證明了有無窮多對質數，彼此相差不超過 70 百萬，而這個數字後來已被降低至 246。費馬（Pierre de Fermat）在 1637 年於書頁邊的空白處隨手寫下：當 $n \geq 3$ 時，$x^n + y^n = z^n$ 沒有正整數解。人們花了超過 350 年，並動用了有史以來最精深的一些數學，才終於由懷爾斯（Andrew Wiles）在 1994 年證明他是對的。

問題連小孩子都聽得懂，答案卻可能要花上幾個世紀——正是這種組合令數論叫人欲罷不能。這也是為甚麼數論一直是證明（proof）的訓練場。整數有無窮多個，你不可能逐一檢查每個情況，於是就必須找出一個能一舉涵蓋所有情況的論證。你在往後的數學生涯中會用到的證明技巧，大部分最初都是在有關整數的問題上磨練出來的。

純粹、無用，卻支撐着互聯網

1940 年，劍橋數學家哈代（G. H. Hardy）頗為自豪地寫道：數論是所有數學中最純粹的，與日常生活相距太遠，不可能有任何實際用途，因而也不可能有任何有害的用途。他錯得不能再錯了。每當你在網上購物、發送加密訊息或為手機解鎖，你都在依靠數論。RSA 系統（1977 年）所依賴的事實是：把兩個大質數相乘很容易，但要從它們的乘積還原出這兩個質數，卻似乎難得無從入手。現代的系統所用的，正是你會在本章末遇到的那些曲線上的算術。十八、十九世紀的「無用」數學，成了二十一世紀的安全保障。

你在大學會遇到甚麼

如果你在大學修讀數學，數論會是你最先到達的地方之一，也會是你最後到達的地方之一。

- **在第一個學期**，你幾乎一定會修讀一門名為數與集合（Numbers and Sets）或數學基礎（Foundations）之類的課程。它會教你怎樣正式地書寫證明，而其中大部分素材都來自數論：歸納法（induction）、歐幾里得算法（Euclidean algorithm）、模算術（modular arithmetic）、費馬小定理（Fermat's Little Theorem）、中國餘數定理（Chinese Remainder Theorem），以及一個令人驚訝的事實——有些無窮集合比另一些無窮集合更大。
- **在羣論（group theory）中**（通常在第一年修讀），你會發現許多數論事實，原來都是同一條關於對稱的定理——拉格朗日定理（Lagrange's Theorem）——投下的影子。
- **在第二年**，一門數論課程通常會涵蓋二次互反律（quadratic reciprocity，即高斯的「黃金定理」）、連分數（continued fractions）和佩爾方程（Pell's equation）。
- **再往後**，你還可以走得更遠：代數數論（algebraic number theory），其中高斯整數（Gaussian integers）等新的數系挽救了唯一分解（unique factorisation）；解析數論（analytic number theory），它運用微積分和複分析（complex analysis）來數算質數；以及橢圓曲線（elliptic curves）的理論，它既是費馬最後定理（Fermat's Last Theorem）證明背後的支柱，也是七道懸賞一百萬美元的千禧年大獎難題（Millennium Prize Problems）之一的根源。

貫穿本章的問題

本章的九道 STEP 題目，圍繞着三個大問題打轉。做題時請把它們記在心上；我們會在本章結尾回到這些問題。

1. **質數有多少個？它們是怎樣分佈的？**人人都知道質數有無窮多個。但它們越往後變得有多稀疏？甚麼樣的論證才能告訴我們答案？
2. **怎樣證明一個數不是分數？**要說明一個數是有理數（rational），只需舉出一個例子。要說明它是無理數（irrational），卻要排除無窮多種可能。這到底是怎樣做到的？
3. **哪些方程有整數解？當我們改變數系時，又會發生甚麼事？**有時我們能證明某個方程完全沒有解。有時要了解整數，最好的方法卻是走出整數的範圍，進入複數（complex numbers），或走到一條曲線上。

關於證明的一章

在這三個問題底下，還貫穿着第四個主題：怎樣證明。每一節之所以入選，部分原因是它展示了一種不同的證明方法的實際運用；合起來，它們就是一套相當適合大學一年級

生的工具。

§	課題	思考方式
1.1	費馬數（Fermat numbers）	一個需要無窮多個質數的無窮族
1.2	質數的乘積	強歸納法（strong induction）；隱藏在二項式係數（binomial coefficients）中的信息
1.3	再談 $\sqrt{2}$	「取最小的那一個」：良序原理（well-ordering principle）
1.4	e 與 e^2	把一個整數嚴格地困在 0 與 1 之間
1.5	不可能的方程	無窮遞降法（infinite descent）；模 n 運算
1.6	兩個平方數之和	轉換場景：複數
1.7	一個奇怪的數系	質疑你視為理所當然的事
1.8	把分數逐一列出	構造雙射（bijection）；反向的遞降
1.9	立方數與曲線	用幾何解決算術問題

怎樣閱讀本章

每一節都以一道 STEP 題目的連結開始，題目與當年考試所出的完全一樣。請先自己嘗試，至少做半小時，然後才往下讀。一道你曾經苦苦思索過的題目，會比一道你只是讀過的題目令你學到多得多。當你往下讀時，問問自己方框是邀請你在答案出現之前先停下來想一想，而停一停，試一試方框則標示着關鍵思路已在你觸手可及之處的時刻。之後，每一節的背後的數學部分會把目光越過考試，望向你在大學將會遇到的數學。

各節的深度大致上逐步遞增，但它們基本上互相獨立。如果某一節未能吸引你，就先跳過，稍後再回來。

1.1 質數工廠

題目

STEP II 2002，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/02-s2-q3`

引子

你幾乎肯定早已知道質數（prime）有無窮多個。歐幾里得（Euclid）約在公元前 300 年證明了這件事，這也是每個人最早接觸到的證明之一。那麼，STEP 的題目為甚麼要你再證明一次？

因為這個證明的運作方式完全不同。歐幾里得的論證是說：給我任何一張有限的質數清單，我都能指出它並不完整。而這個證明則說：這裏有一個預先寫好、明確給出的數列（*sequence*），當中每一個數都含有一個其他各數都沒有的質數。它是一部製造新質數的機器，而它之所以行得通，全靠一條簡潔優美的恆等式（identity），題目會一步一步引領你走向它。

途中你還會遇上一些連費馬（Fermat）本人也被騙倒的數。

動筆之前

先把題目從頭到尾讀一遍，留意它的結構。它分為四個階段：

1. 計算一些例子；
2. 找出（其實是題目告訴你）一個規律，並用這些例子驗證；
3. 證明這個規律在一般情況下成立；
4. 兩次運用這個規律：先推導出關於公因數（common factor）的結論，再推導出關於質數的結論。

這種「例子 → 猜想 → 證明 → 推論」的結構，正是許多數學研究的實際做法，STEP 的

題目也經常依循這個模式。遇到這種結構時，開頭的部分通常不只是熱身：你日後需要的直覺，正是在那裏建立的。

> **想一想**　最後一行問的是無窮多個質數，中間問的是公因數。「任何兩個費馬數（Fermat number）都沒有公因數」和「質數有無窮多個」之間，究竟有甚麼關係？

請記住這個問題，我們稍後會回來討論。

逐步解題

階段 1　熟習這些數

要小心記號：2^{2^n} 是指 $2^{(2^n)}$，而不是 $(2^2)^n$。指數（exponent）本身就是 2 的冪，所以這些數增長得非常快。

$$F_0 = 2^1 + 1 = 3, \qquad F_2 = 2^4 + 1 = 17,$$
$$F_1 = 2^2 + 1 = 5, \qquad F_3 = 2^8 + 1 = 257.$$

> **想一想**　這四個數有甚麼值得留意的地方？

它們全是質數。（驗證 257：它不能被 $2, 3, 5, 7, 11$ 或 13 整除，而 $17^2 = 289 > 257$，所以可以就此停下。）下一個數 $F_4 = 2^{16} + 1 = 65537$ 也是質數。皮埃爾 · 德 · 費馬在 1640 年正正注意到這一點，並猜想（conjecture）所有這種形式的數都是質數。請把這件事記在心裏，稍後我們會看看他是否猜對了。

階段 2　驗證恆等式

現在就較小的 k 驗證 $(*)$：

- $k = 1$：左邊只有 $F_0 = 3$，而 $F_1 - 2 = 3$。✓
- $k = 2$：$F_0 F_1 = 3 \times 5 = 15$，而 $F_2 - 2 = 15$。✓
- $k = 3$：$F_0 F_1 F_2 = 15 \times 17 = 255$，而 $F_3 - 2 = 255$。✓

> **想一想**　看看由 $k = 2$ 到 $k = 3$ 的那一步。我們把 15 乘以 17，得到 255。用費馬數來表達，就是計算了 $(F_2 - 2) \times F_2$，結果得到 $F_3 - 2$。這是巧合嗎？

這個觀察就是整個證明的縮影。讓我們看看它為甚麼應該成立。

階段 3　證明恆等式

停一停，試一試 繼續讀下去之前，試試直接證明

$$(F_k - 2)\,F_k = F_{k+1} - 2.$$

提示：把所有東西都用 2 的冪來表示。$F_k - 2$ 是甚麼樣子的？

我們有 $F_k - 2 = 2^{2^k} - 1$ 和 $F_k = 2^{2^k} + 1$，所以

$$(F_k - 2)\,F_k = \left(2^{2^k} - 1\right)\left(2^{2^k} + 1\right).$$

想一想 這是一個形如 $(a-1)(a+1)$ 的乘積。它等於甚麼？

它是平方差（difference of two squares）$a^2 - 1$。取 $a = 2^{2^k}$，則

$$a^2 = \left(2^{2^k}\right)^2 = 2^{2\cdot 2^k} = 2^{2^{k+1}},$$

所以

$$(F_k - 2)\,F_k = 2^{2^{k+1}} - 1 = F_{k+1} - 2. \tag{†}$$

這就是整部機器的引擎。留意平方的作用：它把指數加倍，而 2^k 加倍後就是 2^{k+1}。費馬數的定義與平方差的因式分解（factorisation）配合得天衣無縫。

這樣一來，歸納法（induction）幾乎是水到渠成。

想一想 在歸納步驟中，我們究竟要證明甚麼？

假設 $(*)$ 對某個 $k \geq 1$ 成立，即

$$F_0 F_1 \cdots F_{k-1} = F_k - 2,$$

而我們想證明它對 $k+1$ 也成立：

$$F_0 F_1 \cdots F_{k-1} F_k = F_{k+1} - 2.$$

把假設的等式兩邊同乘以 F_k，再運用 (†)：

$$F_0 F_1 \cdots F_{k-1} F_k = (F_k - 2)\,F_k = F_{k+1} - 2.$$

我們已經驗證了 $k = 1$ 的情況，所以根據歸納法，$(*)$ 對所有 $k \geq 1$ 都成立。 ■

想一想 我們能否完全不用歸納法？

可以，而且值得看看怎樣做。由 $F_0 - 2 = 1$ 出發，反覆運用 (†)：

$$F_k - 2 = (F_{k-1} - 2)F_{k-1} = (F_{k-2} - 2)F_{k-2}F_{k-1} = \cdots = (F_0 - 2)\,F_0 F_1 \cdots F_{k-1}.$$

由於 $F_0 - 2 = 1$，證明便完成了。這種「逐層展開」的做法，其實就是把同一個歸納法完整地寫出來，不過有些人會覺得這樣更容易看明白。

階段 4　沒有公因數

我們想證明：若 $m \neq n$，則 F_m 與 F_n 沒有大於 1 的公因數。

> **想一想**　假設某個數 d 同時整除 F_m 和 F_n，其中 $m < n$。F_m 在 $(*)$ 的哪個位置出現？

在 $(*)$ 中取 $k = n$：

$$F_0 F_1 \cdots F_m \cdots F_{n-1} = F_n - 2.$$

由於 $m < n$，F_m 是左邊其中一個因數。所以若 d 整除 F_m，則 d 整除整個左邊；也就是說，d 整除 $F_n - 2$。

> **想一想**　現在 d 既整除 F_n，又整除 $F_n - 2$。它還必定整除甚麼？

它們的差：d 整除 $F_n - (F_n - 2) = 2$。所以 $d = 1$ 或 $d = 2$。

> **想一想**　d 可以是 2 嗎？

不可以：每個費馬數都是 2 的冪加 1，所以是奇數。因此 $d = 1$，即任何兩個不同的費馬數都沒有大於 1 的公因數。 ■

數學家會說費馬數*兩兩互質*（pairwise coprime）。

階段 5　無窮多個質數

現在回到我們一開始提出的問題。

> **想一想**　每個大於 1 的整數都至少有一個質因數（prime factor）。從每個費馬數中各選一個質因數。對於你選出的這些質數，你能說些甚麼？

設 p_n 為任意一個整除 F_n 的質數。假如其中兩個相等，例如 $p_m = p_n$ 而 $m \neq n$，那麼這個質數就是 F_m 與 F_n 的一個大於 1 的公因數，但我們剛剛證明了這是不可能的。所以質數

$$p_0,\ p_1,\ p_2,\ p_3,\ \ldots$$

全都*互不相同*。費馬數有無窮多個，所以不同的質數也有無窮多個。 ■

留意我們從來不需要知道每個費馬數是被*哪一個*質數整除，甚至不需要知道費馬數本身是不是質數。我們只需要知道它們永遠沒有公因數。每個費馬數都必須「用掉」至少一個其他費馬數不能用的質數。

n	$F_n = 2^{2^n} + 1$	質因數分解
0	3	質數
1	5	質數
2	17	質數
3	257	質數
4	65 537	質數
5	4 294 967 297	$641 \times 6\,700\,417$
6	18 446 744 073 709 551 617	$274\,177 \times 67\,280\,421\,310\,721$

圖 1.1.1 最初幾個費馬數。費馬猜想它們全是質數，但歐拉（Euler）在 1732 年找到了 F_5 的因數 641，而 F_4 之後至今仍未發現任何費馬數是質數。即使如此，恆等式 $(*)$ 表明不同費馬數的因數永不重疊：每一行都帶來新的質數。

回顧

值得停下來看看這道題成功的關鍵何在，因為同樣的想法會在本書中多次重現。

- **例子並不是裝飾。**看出 $15 \times 17 = 255$ 其實是「$(F_2 - 2) \times F_2 = F_3 - 2$」，正是關鍵思路。當題目給你一些簡單情況時，要問每一步在做甚麼，而不只是看數字是否吻合。
- **一條恆等式完成了所有工作。**一切都源自 $(a-1)(a+1) = a^2 - 1$。看似艱深的 STEP 題目，往往只是由一小段簡單的代數推動。
- **最後的論證着眼於結構，而非大小。**我們證明了質數有無窮多個，卻沒有找出任何一個新的質數。我們證明的是：要構成這個無窮的數族，必須用上無窮多個質數。

背後的數學

費馬猜對了嗎？

沒有。1732 年，萊昂哈德．歐拉證明了

$$F_5 = 2^{32} + 1 = 4\,294\,967\,297 = 641 \times 6\,700\,417.$$

歐拉並不是靠運氣找到 641 的。他證明了 F_n 的任何質因數都必定具有一種非常特殊的形式（當 $n = 5$ 時，它必定是 64 的倍數加一），從而大幅減少了需要考慮的可能。你可以在練習 4 中親自證明這一點。

自此以後，凡是有人成功檢驗過的費馬數，由 F_5 到 F_{32}，結果全都是合成數（composite）。就目前所知，F_0 至 F_4 可能是僅有的費馬質數，但沒有人能夠證明這一點。這是

一個貨真價實的未解難題。

妙就妙在我們的證明完全不受影響。不論是質數還是合成數，費馬數依然能產生無窮多個質數。

費馬質數與正多邊形作圖

費馬質數會在一個相當意想不到的地方出現。古希臘人懂得用直尺和圓規作出等邊三角形和正五邊形，而請留意 $3 = F_0$ 和 $5 = F_1$。其後兩千年間，沒有人能作出其他邊數為質數的正多邊形（regular polygon）。到了 1796 年，年僅 19 歲的卡爾．弗里德里希．高斯（Gauss）示範了怎樣作出正十七邊形，而 $17 = F_2$。

高斯更進一步：正 N 邊形可以用直尺和圓規作出，當且僅當 N 是 2 的冪乘以若干個互不相同的費馬質數。（Pierre Wantzel 在 1837 年補全了這個證明。）所以哪些正多邊形可以作圖，取決於費馬質數有多少個。你會在 Section 6.2（STEP II 1991，第 4 題）遇到高斯正十七邊形背後的關鍵計算，而完整的來龍去脈則屬於伽羅瓦理論（Galois theory），那是大學本科數學課程的精華之一。

一個定理，多種證明

數學家為甚麼要不厭其煩地把同一個定理證明多於一次？因為不同的證明會揭示不同的東西。這個證明一般歸功於克里斯蒂安．哥德巴赫（Goldbach）在 1730 年寫給歐拉的一封信，它表明：任何一個由大於 1 的整數組成、各項兩兩互質的無窮數列，都必然帶來無窮多個質數。費馬數只是其中一個方便的例子。（練習 3 給出另一個例子。）

它給出的結論也比歐幾里得的證明更多。質數 $p_0, p_1, \ldots, p_{n-1}$ 是 n 個不同的質數，每個都不超過 F_{n-1}。所以不大於 $2^{2^{n-1}} + 1$ 的質數至少有 n 個。這是一個關於質數如何分佈的定量陳述。它非常弱（實際情況要好得多），但它是邁向十九世紀數學其中一項偉大成果的第一步，那就是質數定理（Prime Number Theorem），它精確地描述了質數怎樣變得越來越稀疏。Section 1.2 會踏出下一步。

思維技巧

從結構性質出發建立論證

要證明 A 類對象有無窮多個，可以找出無窮多個 B 類對象，而每個 B 類對象都需要一個專屬於自己的 A 類對象。

這裏 A = 質數，B = 費馬數，而「需要一個專屬於自己的」是指「有一個不與

其他任何一個共用的質因數」。同樣的模式也出現在計數論證、線性代數（linear algebra）（無窮多個線性獨立的向量必然導致無窮維）和分析（analysis）之中。

你也練習了歸納法，尤其是養成追問「由 k 到 $k+1$ 的那一步實際上做了甚麼？」的習慣。這個問題往往比形式上的範本更有用。

延伸練習

1 **歐幾里得的版本。**寫出歐幾里得對「質數有無窮多個」的證明。它在甚麼意義上是「反證法」(by contradiction)，而費馬數的證明又在甚麼意義上是「構造性」(constructive)的？

2 **末位數字。**證明當 $n \geq 2$ 時，費馬數 F_n 的末位數字是 7。

3 **西爾維斯特（Sylvester）數列。**定義 $a_1 = 2$ 及 $a_{n+1} = a_n^2 - a_n + 1$，於是數列的開頭幾項是 $2, 3, 7, 43, 1807, \ldots$

(a) 證明 $a_{n+1} - 1 = a_1 a_2 \cdots a_n$。

(b) 由此推導出各項兩兩互質，從而給出「質數有無窮多個」的另一個證明。

(c) 較難 證明 $\dfrac{1}{a_1} + \dfrac{1}{a_2} + \cdots + \dfrac{1}{a_n} = 1 - \dfrac{1}{a_{n+1} - 1}$。

4 **歐拉的妙着。**本題需要用到費馬小定理（*Fermat's Little Theorem*）：若 p 是質數且 $p \nmid a$，則 $a^{p-1} \equiv 1 \pmod{p}$。假設質數 p 整除 F_n。

(a) 證明 $2^{2^n} \equiv -1 \pmod{p}$，從而證明 $2^{2^{n+1}} \equiv 1 \pmod{p}$。

(b) 設 d 為滿足 $2^d \equiv 1 \pmod{p}$ 的最小正整數。證明 d 整除 2^{n+1} 但不整除 2^n，並由此得出 $d = 2^{n+1}$。

(c) 利用費馬小定理，證明 d 整除 $p - 1$。由此推導出 $p \equiv 1 \pmod{2^{n+1}}$。

(d) 當 $n = 5$ 時，這表示 F_5 的任何質因數都具有 $64k + 1$ 的形式。驗證 641 確實如此。小於 641 而又具有這種形式的質數有多少個？

5 **質數的一個上界。**利用一個定理，多種證明中的論證，證明第 n 個質數不超過 $2^{2^{n-1}} + 1$。當 $n = 5$ 時，這個上界有多準確？

延伸閱讀

► **Martin Aigner & Günter M. Ziegler, *Proofs from THE BOOK*** (Springer)，第 1 章「Six proofs of the infinity of primes」（質數無窮性的六個證明）。書中除了這個證明外，

還收錄了歐幾里得的證明和另外四個證明，其中一個用上了拓撲學（topology）。這是入讀大學前最值得一讀的書之一。

- **Harold Davenport, *The Higher Arithmetic***（Cambridge University Press），第 1–2 章。一本文筆優美的入門書，介紹質數與因數分解，最適合在入讀大學前的暑假閱讀。
- **Michal Křížek, Florian Luca & Lawrence Somer, *17 Lectures on Fermat Numbers***（Springer，2001）。關於費馬數的一切，由歐拉的分解到高斯的正多邊形，應有盡有。不妨翻翻開首幾章。
- **Ian Stewart, *Why Beauty Is Truth***（Basic Books），關於高斯的一章。以一般讀者為對象，講述正十七邊形的故事。

1.2 質數分佈得有多稀疏？

題目

STEP III 2016，第 5 題

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

`step.maths.org/questions/16-s3-q5`

引子

在 Section 1.1 中，我們已看到質數（prime）有無窮多個。但「無窮多個」是個很粗略的說法。質數確實愈來愈稀疏：100 以下有 25 個質數，但 900 與 1000 之間只有 14 個。它們究竟會稀疏到甚麼程度？

這道題要證明的結論聽來相當平凡：把不超過 n 的所有質數相乘，所得的積小於 4^n。然而，這個看似平平無奇的不等式（inequality），正是回答「有多稀疏？」的第一步實質進展，也是數學史上最著名的簡短證明之一的關鍵引理（lemma）；該證明出自年輕的保羅·艾狄胥（Paul Erdős）之手，時為 1932 年。

動筆之前

記號 $P_{r,s}$ 看來有點繁瑣，但背後的意思很簡單：把某個範圍內的質數相乘。務必先弄懂題中的例子。$P_{3,7}$ 用的是大於 3 而不超過 7 的質數，即 5 和 7，所以它等於 35。

> **想一想** 第 (iv) 部分要求你對所有 $n \geq 2$ 證明一個命題。第 (iii) 部分則假設該命題對不超過 $2m$ 的每一個 k 都成立。這是哪一種歸納法（induction）？

這是強歸納法（strong induction）：要證明 n 的情況，你可以假設前面所有情況都成立，而不僅是緊接在前的那一個。請把這一點記在心裏；題目之所以這樣編排，原因正在於此。

逐步解題

階段 1　二項式係數不會太大

想一想　在 $(1+x)^{2m+1}$ 的二項展開式（binomial expansion）中代入 $x=1$，會得到甚麼？

你會得到

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

各項都是正數，所以每一項都小於 2^{2m+1}。這還不夠好；我們想要的是 2^{2m}，即它的一半。

想一想　帕斯卡三角形（Pascal's triangle）是對稱的。和式中還有哪一項等於 $\binom{2m+1}{m}$？

由對稱性，$\binom{2m+1}{m} = \binom{2m+1}{m+1}$，而這是和式中兩個不同的項（因為 $m \neq m+1$）。所以

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

這裏是嚴格不等式，因為和式中還有其他正項，例如 $\binom{2m+1}{0} = 1$。兩邊除以 2，即得 $\binom{2m+1}{m} < 2^{2m}$。 ■

階段 2　哪些質數整除這個二項式係數？

把這個二項式係數（binomial coefficient）寫出來：

$$\binom{2m+1}{m} = \frac{(2m+1)!}{m!\,(m+1)!} = \frac{(m+2)(m+3)\cdots(2m+1)}{m!}.$$

想一想　設 p 為質數，且 $m+1 < p \leq 2m+1$。p 整除分子嗎？它有可能整除分母嗎？

分子 $(m+2)(m+3)\cdots(2m+1)$ 的因子當中就有 p 本身，所以 p 整除分子。分母 $m!$ 是若干個數的積，而這些數全都小於 p；由於 p 是質數，如果它不能整除其中任何一個數，就不能整除它們的積。因此 p 整除分子而不整除分母，於是 p 整除整個二項式係數（別忘記，它是一個整數）。

停一停，試一試　我們現在知道，這個範圍內的每一個質數都整除 $\binom{2m+1}{m}$。為甚麼由此可知它們的積 $P_{m+1,2m+1}$ 也整除它？（小心：即使 4 和 6 都整除某個

數，24 也不一定整除它。）

分別在於，我們的質數是**相異質數**。如果相異質數 $p_1, p_2, \ldots$ 各自整除 N，那麼它們的積也整除 N：寫 $N = p_1 N_1$；這時 p_2 整除 $p_1 N_1$ 但不整除 p_1，所以它整除 N_1；如此類推。（這裏用到一個事實：如果一個質數整除某個積，它必整除其中一個因子。這個事實值得我們稍加敬重；Section 1.7 會展示一個世界，在那裏與它相應的命題並不成立。）

所以 $P_{m+1,2m+1}$ 整除 $\binom{2m+1}{m}$，而正數的因數不會大於該數本身，故

$$P_{m+1,\,2m+1} \le \binom{2m+1}{m} < 2^{2m}. \qquad \blacksquare$$

階段 3 把積拆開

想一想 $P_{1,2m+1}$ 是不超過 $2m+1$ 的所有質數的積，怎樣用我們已經掌握的積把它拼出來？

把不超過 $2m+1$ 的質數分成兩批：不超過 $m+1$ 的，以及大於 $m+1$ 的：

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

由第 2 階段，第二個因子小於 $2^{2m} = 4^m$。至於第一個因子，我們可以假設 $P_{1,k} < 4^k$ 對由 2 至 $2m$ 的每一個 k 都成立，而 $k = m+1$ 正在這個範圍之內（因為由 $m \ge 1$ 可得 $2 \le m+1 \le 2m$）。所以

$$P_{1,2m+1} < 4^{m+1} \times 4^m = 4^{2m+1}. \qquad \blacksquare$$

強歸納法正是在這裏大派用場：要處理 $2m+1$，我們需要的是 $m+1$ 的情況，它的大小約為一半，而不是緊接在前的那個情況。

階段 4 所有 n

想一想 第 (iii) 部分處理了奇數。對於偶數 $n = 2m+2$，情況又如何？它是質數嗎？

先驗證起始情況：$P_{1,2} = 2 < 16$ 及 $P_{1,3} = 6 < 64$。現在假設 $P_{1,k} < 4^k$ 對由 2 至 $n-1$ 的所有 k 都成立，其中 $n \ge 4$。

- 若 n 是偶數，則 n 不是質數（因為它大於 2），所以不超過 n 的質數與不超過 $n-1$ 的質數完全相同。因此 $P_{1,n} = P_{1,n-1} < 4^{n-1} < 4^n$。
- 若 n 是奇數，寫 $n = 2m+1$，其中 $m \ge 2$。這時由第 (iii) 部分直接得出 $P_{1,n} < 4^n$。

由強歸納法，$P_{1,n} < 4^n$ 對每一個 $n \ge 2$ 都成立。 $\blacksquare$

回顧

- **二項式係數蘊藏算術資訊。**$\binom{2m+1}{m}$ 是單獨一個數，它不大（小於 4^m），卻必須被許多質數整除。把許多質數擠進一個不大的數裏，就說明這些質數不可能太多。
- **當問題會把自己減半時，強歸納法是最自然的工具。**$2m+1$ 的情況取決於 $m+1$ 的情況。
- **小的情況須另行驗證。**歸納要到 $n \geq 4$ 才能開始運作，所以 $n = 2$ 和 3 必須親手計算。

背後的數學

從 4^n 到質數定理

對結果取對數：$\ln P_{1,n} < n \ln 4$，即

$$\sum_{p \leq n} \ln p < (\ln 4)\, n \approx 1.39\, n.$$

在 1850 年代，切比雪夫（Pafnuty Chebyshev）證明了正是這一類的界（bound），而且上界和下界俱備，從而說明 $\sum_{p \leq n} \ln p$ 的增長與 n 成正比。高斯（Gauss）在大約 15 歲時，已從數表中猜到不超過 n 的質數個數大約是 $n / \ln n$。這後來成為質數定理（Prime Number Theorem），最終在 1896 年由阿達馬（Jacques Hadamard）和德拉瓦萊普桑（Charles de la Vallée Poussin）運用複分析（complex analysis）證明。它的對數形式是

$$\sum_{p \leq n} \ln p \sim n,$$

所以不超過 n 的質數之積大約是 e^n。我們的界 4^n 其實相差不遠：圖 1.2.1 把三者一併畫出。

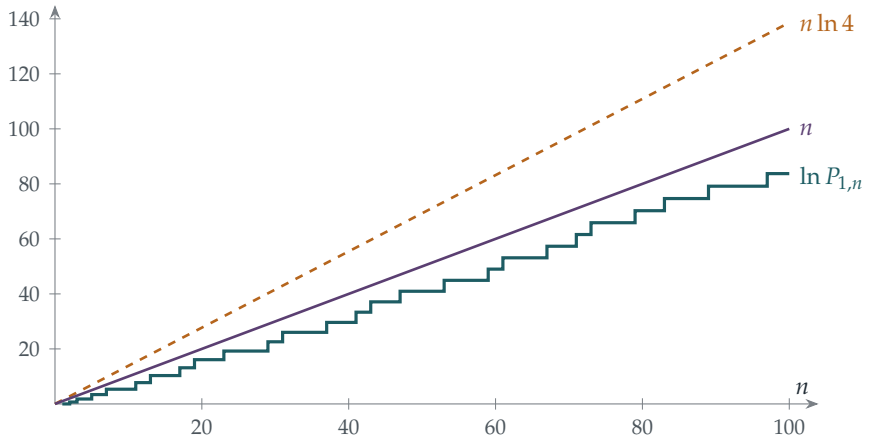


圖 1.2.1　不超過 n 的所有質數之積的對數（藍綠色），在每個質數 p 處跳升 $\ln p$。本節證明它始終低於 $n \ln 4$（琥珀色虛線）；質數定理則指出它的增長與 n 相若（紫紅色）。

伯特蘭公設

1845 年，伯特蘭（Joseph Bertrand）猜想：對每一個 $n \geq 1$，在 n 與 $2n$ 之間總有一個質數。這就是伯特蘭公設（Bertrand's postulate）。切比雪夫在 1852 年證明了它。八十年後，年僅 19 歲的艾狄胥找到了一個優美的初等證明。它的關鍵材料，就是本題的引理，再加上仔細考察哪些質數整除中間的二項式係數 $\binom{2n}{n}$。Martin Aigner 和 Günter Ziegler 把它選為 *Proofs from THE BOOK* 的第 2 章，該書匯集了數學中最優雅的證明。

接下來的方向

質數的精確分佈，由黎曼 ζ 函數（Riemann zeta function）

$$\zeta(s) = 1 + \frac{1}{2^s} + \frac{1}{3^s} + \cdots,$$

的零點（zero）所支配。你會在 Section 2.4 中遇見這個函數樸素得多的一面，那裏有 $\zeta(2) = \pi^2/6$。黎曼猜想（Riemann Hypothesis）預測了這些零點的確切位置；它若成立，便能把質數定理中的誤差限定得緊無可緊。它也許是數學中最著名的未解難題，也是千禧年大獎難題（Millennium Prize Problems）之一。

思維技巧

強歸納法，以及從巧妙的對象中提取資訊

如果 n 的情況取決於遠小於 $n-1$ 的情況，就一次過假設前面所有情況都成立。

這裏的另一個想法較為微妙，也極具數論（number theory）的特色：要證明某個範圍內所有質數的某種性質，就找出一個對象（這裏是一個二項式係數），使得這些質數每一個都必須整除它，而你又能證明它不會太大。

延伸練習

1 **感受一下數字。**對 $n = 10, 20, 30$ 計算 $P_{1,n}$，並把它與 4^n 及 e^n 比較。（這裏可以使用計算機。）

2 **中間的二項式係數。**證明 $\binom{2m}{m}$ 是 $(1+1)^{2m}$ 展開式的 $2m+1$ 項中最大的一項，並由此推出 $\binom{2m}{m} \geq \dfrac{4^m}{2m+1}$。

3 **質數不會太多。**設 $\pi(n)$ 為不超過 n 的質數個數。介乎 $\sqrt{n}$ 與 n 之間的每個質數都大於 $\sqrt{n}$，而它們的積至多為 $P_{1,n} < 4^n$。由此推出

$$\pi(n) - \pi(\sqrt{n}) < \frac{2\ln 4}{\ln n}\, n,$$

從而得出 $\pi(n) < 2.8\, n/\ln n + \sqrt{n}$。把這個結果與質數定理的 $\pi(n) \approx n/\ln n$ 比較。

4 **驗證伯特蘭公設。**對 $n = 1, 2, \ldots, 50$ 驗證伯特蘭公設，即在每個情況下找出一個介乎 n 與 $2n$ 之間的質數。你能否利用一張簡短的質數表有效率地完成，使表中每個質數都小於前一個的兩倍？

延伸閱讀

- **Aigner & Ziegler, *Proofs from THE BOOK***，第 2 章「Bertrand's postulate」（伯特蘭公設）。
- **Marcus du Sautoy, *The Music of the Primes***（Fourth Estate）。一部講述質數定理與黎曼猜想歷史的普及讀物。
- **G. J. O. Jameson, *The Prime Number Theorem***（Cambridge University Press）。嚴謹而平易近人，為大學本科生而寫；開首幾章只需要高級程度（A-level）的微積分。

1.3 最小反例

題目

STEP III 2015，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/15-s3-q5`

引子

你大概見過 $\sqrt{2}$ 是無理數(irrational)的證明：假設 $\sqrt{2} = a/b$ 為最簡分數(lowest terms)，兩邊平方，推出 a 是偶數，繼而推出 b 也是偶數，於是得出矛盾（contradiction）。這是個漂亮的證明，但它依賴一個關於偶數的技巧，而這個技巧看不出可以怎樣推廣。換成 $\sqrt[3]{2}$ 又如何？$\sqrt{7}$ 呢？

本題給出另一個證明，它的核心想法簡單得近乎荒謬：*看看最小的那一個*。在本章裏，同一個想法還會改頭換面地出現好幾次。

動筆之前

這是一道不尋常的 STEP 題目：它給你一個證明的骨架，要你補上當中欠缺的理據。大學生正是這樣學習閱讀數學的。課本裏的證明常常說「顯然」或「由此可得」，而你的部分工作，就是查證事實是否真的如此。

想一想 第 4 步說「定義 k 為 S 中最小的正整數」。我們憑甚麼可以這樣做？假如 S 是一個由正實數組成的集合，這樣做還可以嗎？

任何由正整數組成的非空集合都有最小元素。這稱為*良序原理*（well-ordering principle），它與數學歸納法（mathematical induction）原理等價。對於由正實數組成的集合，這個原理並不成立：集合 $\{x : x > 0\}$ 就沒有最小元素。所以第 4 步既依賴第 3 步（集合必須非空），也依賴 S 由整數組成這一點。

逐步解題

階段 1　第 3 步：S 非空

想一想　若 $\sqrt{2} = a/b$，其中 a 和 b 是正整數，你能寫出 S 的一個元素嗎？

若 $\sqrt{2} = a/b$，則 $b\sqrt{2} = a$，是一個整數。所以 b 屬於 S，S 非空。

階段 2　第 5 步：一個更小的元素

現在 k 是 S 的最小元素，所以 $k\sqrt{2}$ 是整數，記為 j。

停一停，試一試　要證明 $(\sqrt{2}-1)k$ 屬於 S，你必須檢查三件事。是哪三件？

我們需要 $(\sqrt{2}-1)k$ 滿足：(a) 是整數；(b) 是正數；(c) 乘以 $\sqrt{2}$ 後得到整數。

- (a) $(\sqrt{2}-1)k = k\sqrt{2} - k = j - k$，是兩個整數之差。
- (b) $\sqrt{2} > 1$，所以 $\sqrt{2} - 1 > 0$，因此 $(\sqrt{2}-1)k > 0$。
- (c) $(\sqrt{2}-1)k \cdot \sqrt{2} = 2k - k\sqrt{2} = 2k - j$，是整數。

所以 $(\sqrt{2}-1)k \in S$。

階段 3　第 6 步：矛盾

想一想　$\sqrt{2} - 1$ 有多大？

由於 $1 < \sqrt{2} < 2$，我們有 $0 < \sqrt{2} - 1 < 1$，所以 $(\sqrt{2}-1)k < k$。我們找到了 S 的一個元素，它比 S 的最小元素 k 還要小。這是不可能的，所以第 1 步的假設不成立：$\sqrt{2}$ 是無理數。 ■

留意我們從頭到尾都不需要「最簡分數」，也不需要任何關於奇偶數的事實。整個證明只建基於三個事實：$\sqrt{2} \cdot \sqrt{2}$ 是整數；$\sqrt{2}$ 嚴格地介乎兩個連續整數之間；整數具有良序性。

階段 4　立方根：把兩個數聯繫起來

設 $c = 2^{1/3}$，則 $c^2 = 2^{2/3}$ 且 $c^3 = 2$。

想一想　若 c 是有理數（rational），c^2 是否也是？反過來，若 c^2 是有理數，你能否只用乘法和除法回到 c？

若 c 是有理數，則 $c^2 = c \times c$ 也是有理數。反之，若 c^2 是有理數，則 $2/c^2 = c^3/c^2 = c$ 也是有理數。所以 $2^{1/3}$ 是有理數，當且僅當 $2^{2/3}$ 是有理數。 ■

階段 5 改造論證

? **想一想** 在第 (i) 部分，關鍵在於 $(\sqrt{2})^2 = 2$ 是整數，所以把 S 的元素乘以 $\sqrt{2}$ 兩次，就會回到整數。對於 $c = 2^{1/3}$，你需要乘三次。集合 S 應該怎樣修改？

假設 c 是有理數，則 c^2 也是。定義

$$S = \{n \text{ 為正整數} : nc \text{ 和} nc^2 \text{ 都是整數}\}.$$

- S **非空。**若 $c = a/b$ 且 $c^2 = a'/b'$，則 bb' 屬於 S。
- **設** k **為它的最小元素**，並考慮 $(c-1)k$。我們有 $1 < c < 2$（因為 $1^3 < 2 < 2^3$），所以 $0 < (c-1)k < k$。
- **它是整數：**$(c-1)k = kc - k$。
- **乘以** c**：**$(c-1)k \cdot c = kc^2 - kc$，是整數。
- **乘以** c^2**：**$(c-1)k \cdot c^2 = kc^3 - kc^2 = 2k - kc^2$，是整數。

所以 $(c-1)k \in S$，而且比 k 小：矛盾。因此 $2^{1/3}$ 是無理數，而根據第 4 階段，$2^{2/3}$ 也是無理數。 ■

回顧

- **「取最小的那一個」是一種證明技巧。**假設反例存在，取其中最小的一個，再由它造出一個更小的。這稱為*極端原理*（extremal principle），或稱*最小反例*（minimal counterexample）論證。
- **集合** S **替我們記帳。**我們不必周旋於分子和分母之間，而是把所有「好」的乘數收進同一個集合。
- **推廣，就是追問證明真正行得通的原因。**對於 $\sqrt{2}$，需要乘以 $\sqrt{2}$ 一次；對於 $\sqrt[3]{2}$，則需要乘兩次。看出這一點，我們就清楚知道該怎樣重新定義 S。

背後的數學

一個畫得出來的證明

第 (i) 部分背後有一幅優美的圖，出自 Stanley Tennenbaum。若 $\sqrt{2} = a/b$，則 $a^2 = 2b^2$：一個邊長為 a 的正方形，面積等於兩個邊長為 b 的正方形面積之和。如圖 1.3.1 所示，把兩個邊長為 b 的正方形放在邊長為 a 的正方形的兩個對角。它們重疊之處是一個小正方形，另外還有兩個小正方形沒有被蓋住。由於面積相等，重疊部分必定等於兩個未被蓋住的正方形之和，於是得到一對更小而具有同樣性質的整數。整數對不可能無止境地愈變愈小，所以這樣的正方形並不存在。這對較小的整數正是 $(2b - a,\ a - b)$，也就是我們那一步 $(\sqrt{2} - 1)k$ 的化身。

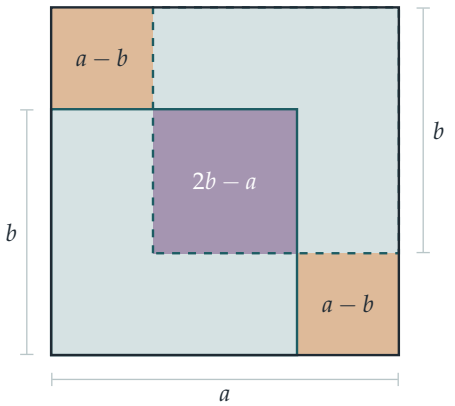


圖 1.3.1　Tennenbaum 的證明。兩個邊長為 b 的正方形（藍綠色）放在邊長為 a 的正方形的兩個對角。若 $a^2 = 2b^2$，則被計算了兩次的面積（紫紅色的重疊部分）必定等於沒有計算到的面積（兩個琥珀色的角落），所以 $(2b - a)^2 = 2(a - b)^2$：一個更小的解。

良序、歸納與遞降

良序原理與歸納法原理在邏輯上等價：兩者可以互相推導。到了大學，你會看到兩者都被用作自然數的基礎。在 Section 1.5 中，我們會以另一種形式運用同一個原理，那就是費馬（Fermat）的*無窮遞降法*（method of infinite descent）：不存在由正整數組成的無窮嚴格遞減數列。

它能走多遠？

這個論證可以證明，只要 N 不是完全平方數（perfect square），$\sqrt{N}$ 就是無理數（見練習）；更一般地，一個整數的 n 次方根要麼是整數，要麼是無理數。到了大學，這會變成一個關於代數整數（algebraic integers）的命題：若一個有理數滿足某個整數係數的首一多項式（monic polynomial）方程，例如 $x^3 - 2 = 0$，它就必定是整數。你會在 Section 6.3 再次遇到這個想法。

相傳畢達哥拉斯學派（Pythagoreans）因為希帕索斯（Hippasus）洩露了 $\sqrt{2}$ 是無理數而把他淹死，這個傳說多半並不屬實。屬實的是，這個發現迫使希臘數學家把對數的研究與對長度的研究分開；這道裂痕要等到十九世紀，待戴德金（Dedekind）構造出實數，才得以彌合。

思維技巧

極端原理

如果你想證明某件事永遠不會發生，就假設它會發生，選出最小（或最大、或最簡單）的例子，再用它造出一個更小的例子。

決定「最小」是甚麼意思，正是需要創意的一步。在這裏，它是使 $n\sqrt{2}$ 成為整數的最小的 n。

延伸練習

1. **其他平方根。**改造這個論證，證明 $\sqrt{3}$ 是無理數。哪一個數擔當了 $\sqrt{2}-1$ 的角色？
2. **所有非平方數。**設 N 為一個不是完全平方數的正整數，並設 m 為滿足 $m < \sqrt{N} < m+1$ 的整數。考慮 $(\sqrt{N} - m)k$，從而證明 $\sqrt{N}$ 是無理數。
3. **找出漏洞。**試把這個論證套用於 $\sqrt{4}$。究竟是哪一步行不通？
4. **Tennenbaum 的圖。**利用圖 1.3.1 證明：若 $a^2 = 2b^2$，則 $(2b-a)^2 = 2(a-b)^2$。再用代數方法驗證這個恆等式。

延伸閱讀

- **Daniel Velleman, *How to Prove It***（Cambridge University Press）。一本全面的證明入門書，內容包括歸納法和良序原理。

- **R. P. Burn, *A Pathway into Number Theory***（Cambridge University Press）。全書由精心編排次序的問題組成，是修讀大學一年級課程前的理想準備。
- **John Conway & Joseph Shipman, "Extreme proofs I: the irrationality of $\sqrt{2}$"**（*Mathematical Intelligencer*，2013）。同一個事實的多個證明，比較得妙趣橫生。

1.4 e 是無理數，e^2 也是

本節會用到兩道題目。第一道出自 1997 年，證明 e 是無理數（irrational number）。第二道出自 2024 年，把同一個想法用得更巧妙，證明 e^2 是無理數，而這原來確實困難得多。

熱身題

STEP III 1997，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/97-s3-q7`

主題目

STEP III 2024，第 7 題

繼續閱讀之前，請先看題目（原題為英文）。題目載於 OCR 提供下載的 2024 年 STEP 試卷（zip 檔案）內的 STEP III 試卷：

`www.ocr.org.uk/Images/725753-2024-past-papers.zip`

引子

在 Section 1.3 中，我們證明了 $\sqrt{2}$ 是無理數，用的是 $\sqrt{2}$ 滿足一條簡單方程 $x^2 = 2$ 這個事實。但就我們所知，e 並不滿足任何這樣的方程。我們手上只有它的級數（series）：

$$e = 1 + \frac{1}{1!} + \frac{1}{2!} + \frac{1}{3!} + \cdots.$$

一個級數怎能告訴你某個數不是分數？

答案是數論（number theory）中最有用的想法之一：整數（*integer*）不可能嚴格介乎 0 與 1 之間。如果你能證明某個量是整數，同時又嚴格介乎 0 與 1 之間，就得出矛盾（contradiction）。本節的一切，都是這個主題的變奏。

動筆之前

想一想　假設 $e = p/q$。把 e 乘以 q 會得到整數。那麼乘以 $q!$ 又會對 e 的級數的每一項有甚麼影響？

只要 $k \le q$，把 $\dfrac{1}{k!}$ 乘以 $q!$ 都會得到整數。因此 $q!\,e$ 可分成整數部分（首 $q+1$ 項）和「尾部」（tail，即其後的所有項）。若 $e = p/q$，則 $q!\,e$ 本身是整數，所以尾部也必定是整數。我們的計劃，就是證明尾部嚴格介乎 0 與 1 之間。

逐步解題

階段 1　熱身題：e 是無理數

級數 b_n 是等比級數（geometric series），首項和公比都是 $\frac{1}{n+1}$，所以

$$b_n = \frac{1/(n+1)}{1 - 1/(n+1)} = \frac{1}{n}.$$

想一想　逐項比較 a_n 和 b_n。哪一個較大？

兩者的首項相等。其後，$(n+1)(n+2) > (n+1)^2$，而一般來說，當 $k \ge 2$ 時 $(n+1)(n+2)\cdots(n+k) > (n+1)^k$，所以 a_n 其後的每一項都小於 b_n 的對應項。因此 $0 < a_n < b_n = 1/n$。

現在看看 a_n 究竟是甚麼。把 e 的級數乘以 $n!$：

$$n!\,e = \underbrace{\frac{n!}{0!} + \frac{n!}{1!} + \cdots + \frac{n!}{n!}}_{\text{整數}} + \underbrace{\frac{n!}{(n+1)!} + \frac{n!}{(n+2)!} + \cdots}_{= a_n}.$$

由於 $0 < a_n < 1$，$n!\,e$ 的整數部分就是第一個括號，而 a_n 就是小數部分（fractional part）：$a_n = n!\,e - [n!\,e]$。

停一停，試一試　現在完成證明。假設 $e = p/q$。你應該選取哪個 n 值？

取 $n = q$。那麼 $q!\,e = (q-1)!\,p$ 是整數，所以它的小數部分 a_q 是 0。但 $a_q > 0$。矛盾：e 是無理數。 ■

這個證明出自傅立葉（Joseph Fourier），約於 1815 年提出。

階段 2　為何 e^2 較難

想一想 試把同一個想法用於 e^2。它的級數是 $\sum 2^k/k!$。當你乘以 $n!$ 再看尾部時，哪裏出了問題？

尾部現在是 $\dfrac{2^{n+1}}{n+1}+\dfrac{2^{n+2}}{(n+1)(n+2)}+\cdots$，而 2 的冪把一切都破壞了：尾部不再細小。乘以 $n!$ 對分子中的 2^k 毫無幫助。

2024 年這道題的妙着，是完全避開 e^2。若 $e^2=p/q$，則

$$qe=\frac{p}{e}.$$

現在我們有兩個數 e 和 $1/e$，各自都有一個階乘（factorial）級數，而且沒有麻煩的 2 的冪：

$$e=\sum_{k=0}^{\infty}\frac{1}{k!},\qquad \frac{1}{e}=\sum_{k=0}^{\infty}\frac{(-1)^k}{k!}.$$

階段 3 兩個尾部

第 (i) 部分正正就是熱身題：$0<f(n)<1/n$。

至於第 (ii) 部分，$g(n)$ 是一個各項的大小遞減的交錯級數（alternating series）。

想一想 對於交錯級數 $t_1-t_2+t_3-\cdots$，其中 $t_1>t_2>t_3>\cdots>0$，為何它的和介乎 0 與 t_1 之間？

把各項兩兩配對：$(t_1-t_2)+(t_3-t_4)+\cdots$ 是若干個正括號之和，所以是正數。換另一種方式配對：$t_1-(t_2-t_3)-(t_4-t_5)-\cdots$ 是 t_1 減去若干個正括號，所以小於 t_1。這裏 $t_1=\frac{1}{n+1}$，而各項遞減，因為每一項都是前一項除以一個大於 1 的數。所以 $0<g(n)<\frac{1}{n+1}$。

階段 4 又見整數

把 e 乘以 $(2n)!$，然後分出尾部，做法與熱身題完全一樣：

$$(2n)!\,e=(\text{整數})+f(2n).$$

所以 $(2n)!\,e-f(2n)$ 是整數。現在對 $1/e$ 做同樣的事：

$$\frac{(2n)!}{e}=\sum_{k=0}^{2n}(-1)^k\frac{(2n)!}{k!}+\sum_{k=2n+1}^{\infty}(-1)^k\frac{(2n)!}{k!}.$$

想一想 尾部的首項（即 $k=2n+1$ 的一項）是甚麼符號？尾部與 $g(2n)$ 相比又如何？

尾部首項的 $k = 2n+1$ 是奇數，所以該項是負數。尾部是

$$-\frac{1}{2n+1} + \frac{1}{(2n+1)(2n+2)} - \cdots = -g(2n).$$

所以 $\dfrac{(2n)!}{e} + g(2n)$ 是整數。 ■

這就是題目用 $2n$ 而不用 n 的原因：乘數是偶數的階乘時，$1/e$ 的尾部恰好是 $-g$，符號可以預知。

階段 5　矛盾

假設 $qe = p/e$。把等式兩邊乘以 $(2n)!$，並利用第 4 步：

$$q\big(\text{整數} + f(2n)\big) = p\big(\text{整數} - g(2n)\big).$$

移項整理後，$qf(2n) + pg(2n)$ 等於若干整數的組合，所以它是整數。這就是第 (iv) 部分。

> **停一停，試一試**　現在利用第 3 步的上下界。當 n 非常大時，$qf(2n) + pg(2n)$ 會怎樣？

$f(2n)$ 和 $g(2n)$ 都是正數，所以 $qf(2n) + pg(2n) > 0$。而且

$$qf(2n) + pg(2n) < \frac{q}{2n} + \frac{p}{2n+1} < \frac{p+q}{2n}.$$

選取 $n > (p+q)/2$。那麼 $qf(2n) + pg(2n)$ 就是一個嚴格介乎 0 與 1 之間的整數。這是不可能的，所以 e^2 是無理數。 ■

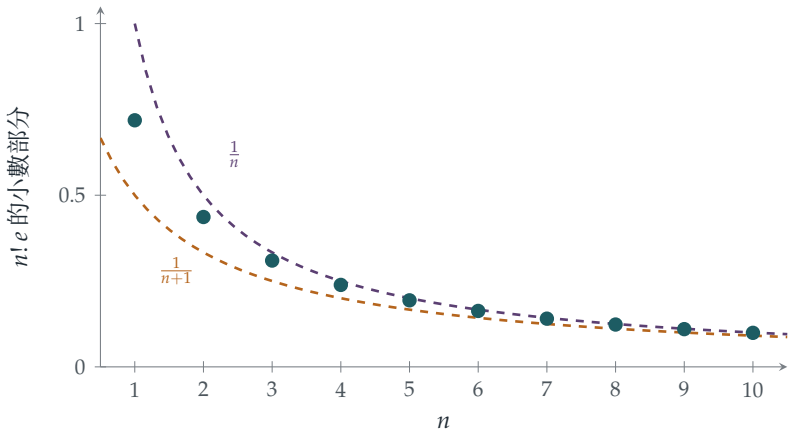


圖 1.4.1　尾部 $a_n = n!\,e - [n!\,e]$（青綠色），即 $n!\,e$ 的小數部分。它總是嚴格介乎 $\frac{1}{n+1}$ 與 $\frac{1}{n}$ 之間：永不為零，但小於 1。假如 e 是分數 $\frac{p}{q}$，那麼對每個 $n \geq q$，$n!\,e$ 都會是整數，青綠色的點就會是 0。

回顧

- **推動整個證明的是「沒有整數嚴格介乎 0 與 1 之間」。**兩個證明都構造出一個必定是整數的量，然後把它困住。
- **階乘令尾部變小。**乘以 $n!$ 把級數的開首部分變成整數，剩下一個像 $1/n$ 那樣縮小的尾部。
- **正面進攻失敗時，就重新整理。**我們沒有正面處理 $e^2 = p/q$，而是把它改寫成 $qe = p/e$，把困難分攤到兩個較易處理的數上。
- **細節很重要。**選用 $2n$ 而非 n，固定了尾部的符號。許多證明的成敗，都取決於這樣的一個小細節。

背後的數學

e 有多「無理」？

劉維爾（Joseph Liouville）在 1840 年證明了 e^2 是無理數，而且事實上它不滿足任何整數係數的二次方程。1873 年，埃爾米特（Charles Hermite）走得遠得多：他證明了 e 是*超越數*（transcendental number），意思是它不是任何整數係數多項式的根。他的證明正是上述論證的精密版本：構造一個整數，再證明它嚴格介乎 0 與 1 之間。九年後，林德曼（Ferdinand von Lindemann）改造了埃爾米特的方法，證明 π 是超越數，終於證實以尺規「化圓為方」這個古老問題是不可能做到的。

我們仍未知道的事

儘管這些方法威力強大，一些簡單得令人尷尬的問題卻仍未解決。$e + \pi$ 是無理數嗎？$e\pi$ 呢？沒有人知道，雖然兩者之中至少有一個必定是超越數（練習 4）。$\zeta(3) = 1 + \frac{1}{8} + \frac{1}{27} + \cdots$ 是無理數嗎？是的：阿培里（Roger Apéry）在 1978 年證明了這一點，震驚了整個數學界。$\zeta(5)$ 呢？沒有人知道。

逼近

這些證明背後有一個更深刻的想法：如果一個數能被分數逼近得*太好*，它就是無理數。e 的級數的部分和（partial sum）都是分數，而且很快就非常接近 e；一個分數不可能被其他分數逼近得那麼好。這就是*丟番圖逼近*（Diophantine approximation）的起點。它是數論的一個分支，你可以在 Section 2.9 進一步探索，那裏會用另一族積分得出更好的分數，步步逼近 e。

思維技巧

困住一個整數

要證明一個數是無理數，先假設它是分數，由此構造一個必定是整數的量，再證明這個量嚴格介乎 0 與 1 之間。

箇中的藝術在於選擇這個量。在這裏，乘以階乘把級數的大部分變成整數，剩下一個細小但為正數的尾部。

延伸練習

1 **只用 $1/e$。**假設 $1/e = p/q$，並只利用 $1/e$ 的交錯級數，證明 e 是無理數。

2 $\sin 1$ **是無理數。**利用 $\sin 1 = 1 - \dfrac{1}{3!} + \dfrac{1}{5!} - \cdots$ 證明 $\sin 1$ 是無理數。（乘以一個合適的奇數的階乘。）

3 **為何 π 不行？**解釋為何這個方法不能直接用於 π。然後查閱 Ivan Niven 在 1947 年發表、只有一頁的 π 是無理數的證明，當中用的是積分而不是級數。能力強的高級程度（A-level）學生應可讀懂。

4 **其中一個是超越數。**較難 證明 e 和 π 是 $x^2 - (e+\pi)x + e\pi = 0$ 的根。你可以假設：若一個多項式的係數都是代數數（algebraic number），則它的根本身也是代數數；以及 π 不是代數數（林德曼）。由此推出 $e+\pi$ 和 $e\pi$ 之中至少有一個是超越數。

延伸閱讀

- **Aigner & Ziegler,** ***Proofs from THE BOOK***，第 8 章「Some irrational numbers」（一些無理數）。正是用這些想法，證明 e、e^2、e^4 和 π^2 都是無理數。
- **Ivan Niven,** ***Irrational Numbers***（Mathematical Association of America）。一本篇幅不長的經典，內容由 $\sqrt{2}$ 一直談到超越性。
- **Julian Havil,** ***The Irrationals***（Princeton University Press）。一本淺白易讀的無理數歷史。

1.5 無窮遞降法

題目

STEP III 2022，第 2 題

繼續閱讀之前，請先看題目（原題為英文）。題目載於 OCR 提供下載的 2022 年 STEP 試卷（zip 檔案）內的 STEP III 試卷：

`www.ocr.org.uk/Images/696587-2022-past-papers.zip`

引子

怎樣才能證明一條方程沒有整數解？你不可能逐一試遍所有可能，因為可能性有無窮多個。費馬（Pierre de Fermat）在 1630 年代找到了一個方法：證明任何一個解都會產生另一個數字更小的解，然後又產生一個更小的，如此無休止地繼續下去。可是正整數不可能永無止境地愈變愈小。所以，從一開始就根本沒有解。

費馬把這個方法稱為*無窮遞降法*（method of infinite descent），並且深以為傲。它是 Section 1.3 中良序原理（well-ordering principle）的另一面，至今仍是證明方程無解的主要工具之一。

動筆之前

四個部分的格局都一樣：假設有解，把解縮小，得出矛盾（contradiction）。問題在於怎樣縮小。

想一想 在第 (i) 部分，看看 $a^3 + 2b^3 + 4c^3 = 0$。三項之中有兩項明顯是偶數。這告訴你第三項有甚麼性質？

每一次遞降都是由這類觀察開始的：某種整除性（divisibility）是你無法迴避的。

逐步解題

階段 1　第一次遞降

由於 $a^3 = -2b^3 - 4c^3$，a^3 這個數是偶數，所以 a 是偶數（奇數的立方是奇數）。設 $a = 2p$。那麼 $|p| = |a|/2 < |a|$，代入後得

$$8p^3 + 2b^3 + 4c^3 = 0 \quad \Longrightarrow \quad 4p^3 + b^3 + 2c^3 = 0.$$

想一想 新的方程是 $4p^3 + b^3 + 2c^3 = 0$。它的係數和原來一樣，只是輪換了位置。你可以再用同一招嗎？

現在 $b^3 = -4p^3 - 2c^3$ 是偶數，所以 $b = 2q$，其中 $|q| < |b|$，再除以 2 便得 $2p^3 + 4q^3 + c^3 = 0$。同樣地，c 必定是偶數：$c = 2r$，其中 $|r| < |c|$，再除以 2，

$$p^3 + 2q^3 + 4r^3 = 0.$$

這正是原來的方程，只是以 p、q、r 取代了 a、b、c。每個數的大小都減半，而且沒有一個是零，因為把非零的數減半不可能得到零。

停一停，試一試 把以上論證寫成一個證明（proof），說明非零解並不存在。你究竟應該假設甚麼是「最小」的？

在所有非零整數解之中，選取 $|a|$ 最小的一個。（只要有解存在，良序原理就容許我們這樣做。）上述論證會產生另一個非零解，並且 $|p| < |a|$，與最小性矛盾。所以沒有非零解。 ■

另一個等價的說法是：由一個解出發，我們可以構造出無窮多個解，它們滿足 $|a| > |p| > \cdots$，即一個嚴格遞減的無窮正整數數列，而這是不可能的。

階段 2　換一個質數

想一想 在 $9a^3 + 10b^3 + 6c^3 = 0$ 中，奇偶性（parity）看來幫不上忙。看看係數 9、10、6。哪一個質數（prime）整除其中兩個？

質數 3 整除 9 和 6，所以 3 整除 $10b^3$，因此 3 整除 b。設 $b = 3q$ 並除以 3：

$$3a^3 + 90q^3 + 2c^3 = 0.$$

現在 3 整除 $2c^3$，所以 $c = 3r$；再除以 3 得 $a^3 + 30q^3 + 18r^3 = 0$。於是 3 整除 a：$a = 3p$，再除以 3 一次，

$$9p^3 + 10q^3 + 6r^3 = 0.$$

這與原來的方程完全相同，只是換成了較小的非零數。根據同樣的遞降論證，方程沒有非零解。 ■

階段 3　模 3 下的平方數

想一想 每個整數都可寫成 $3n$、$3n+1$ 或 $3n-1$ 的形式。它們的平方模（modulo）3 是甚麼？

$(3n \pm 1)^2 = 9n^2 \pm 6n + 1 = 3(3n^2 \pm 2n) + 1$，比 3 的倍數多一。而 $(3n)^2$ 是 3 的倍數。所以平方數除以 3 的餘數（remainder）是 0 或 1，永遠不會是 2。

現在如果 $a^2 + b^2$ 是 3 的倍數，餘數的可能組合是 $0+0$、$0+1$、$1+0$ 和 $1+1$，分別得出 0、1、1、2。只有第一種是 3 的倍數，所以 a 和 b 都必定是 3 的倍數。

想一想 如果 $a^2 + b^2 = 3c^2$，而 $a = 3p$、$b = 3q$，你對 c 有甚麼結論？

我們得到 $9p^2 + 9q^2 = 3c^2$，所以 $c^2 = 3(p^2 + q^2)$ 是 3 的倍數，因此 $c = 3r$。於是 $9p^2 + 9q^2 = 27r^2$，即 $p^2 + q^2 = 3r^2$：同一條方程，較小的數。又是遞降。 ■

階段 4 變奏：方程變了

想一想 在 $a^2 + b^2 + c^2 = 4abc$ 中，右邊是 4 的倍數。平方數模 4 是甚麼？

偶數的平方是 4 的倍數，奇數的平方則比 4 的倍數多一（$(2n+1)^2 = 4n^2 + 4n + 1$）。所以 $a^2 + b^2 + c^2$ 除以 4 的餘數是 0、1、2 或 3，視乎 a、b、c 之中有多少個是奇數。由於右邊是 4 的倍數，它們*沒有一個*可以是奇數。設 $a = 2p$、$b = 2q$、$c = 2r$：

$$4(p^2 + q^2 + r^2) = 32\,pqr \quad \Longrightarrow \quad p^2 + q^2 + r^2 = 8\,pqr.$$

想一想 這並不是同一條方程：4 變成了 8。遞降法還行得通嗎？

竅門在這裏：*證明一個更強的命題*。考慮所有以下形式的方程

$$x^2 + y^2 + z^2 = 2^m\,xyz \qquad \text{其中} m \geq 2.$$

對任何這樣的 m，右邊都是 4 的倍數，所以完全相同的論證顯示 x、y、z 全是偶數。把它們減半，便得到含 2^{m+1} 的方程的一個解，而這條方程仍屬這種形式。因此，由 m 方程的一個非零解，我們得到 $m+1$ 方程的一個較小的非零解。在這個方程族中*任何一條*方程的所有解之中，取 $|x|$ 最小的一個，便得出矛盾。特別地，$a^2 + b^2 + c^2 = 4abc$ 沒有非零解。 ■

有時候，你想證明的命題太弱，不足以支撐它自身的歸納（induction），補救的方法是證明一個更一般、能夠自我支撐的命題。這個技巧在大學裏出奇地常見。

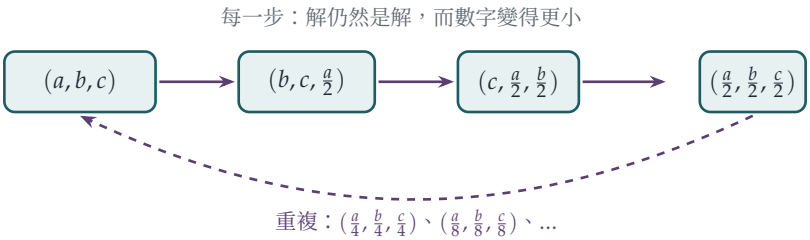


圖 1.5.1　第 (i) 部分的遞降。由 $a^3+2b^3+4c^3=0$ 的一個解出發，三步之後得到另一個解，每個分量都減半。無休止地重複下去，便會得出一個沒有盡頭的遞減正整數數列 $|a|>|\frac{a}{2}|>|\frac{a}{4}|>\cdots$，而這是不可能的。所以一開始就沒有解。

回顧

- **遞降需要整除性的來源。**每一次，我們都找到一個質數 p（這裏是 2 或 3），它必然整除所有變數，然後把它除去。
- **模 n 運算正是找出它的方法。**考察模 2、3 或 4 的餘數，把「有沒有任何可能行得通？」變成一個簡短的有限驗證。
- **如果方程變了，就推廣。**第 (iv) 部分需要一整族方程，而這一族方程在遞降步驟下是封閉的。

背後的數學

費馬、遞降法與他的最後定理

費馬用無窮遞降法證明了 $x^4+y^4=z^2$ 沒有正整數解。在他著名的「最後定理」之中，這是他唯一留下證明的情況，它解決了 $x^n+y^n=z^n$ 在 $n=4$ 時的情形。歐拉（Euler）其後用遞降法處理 $n=3$ 的情形，當中有一個漏洞，後來才被補上。遞降法至今仍是現代理論的起點：關於 Section 1.9 中那類曲線的莫德爾–韋伊定理（Mordell–Weil theorem），其證明就是一個精密的遞降。

局部障礙

本節的每一個證明，都是靠考察模某個數的餘數來完成的。如果一條方程對某個 n 而言模 n 無解，或者沒有實數解，它當然沒有整數解。這類理由稱為*局部障礙*（local obstructions）；到了大學，你會遇到*p 進數*（p-adic numbers），那是一種同時研究模 $p, p^2, p^3, \ldots$ 所有餘數的方法。

哈塞（Hasse）與閔可夫斯基（Minkowski）有一個精彩的定理：對於像 $a^2+b^2=3c^2$ 這樣的二次方程，局部障礙就是僅有的障礙——如果沒有任何「局部」的理由令它無解，它就確實有有理數解。對三次方程而言，這個結論卻徹底失效。Ernst Selmer 在 1951 年證明了

$$3x^3+4y^3+5z^3=0$$

模任何整數都有解，在實數中也有解，但沒有非零整數解。解釋「局部到整體」原理何時失效、為何失效，是現代數論的一大主題。

一條確實有解的方程

第 (iv) 部分有一個著名的近親：$x^2+y^2+z^2=3xyz$，即**馬爾可夫方程**（Markov equation）。它有無窮多個正整數解，由 $(1,1,1)$、$(1,1,2)$、$(1,2,5)$、$(1,5,13)$ 和 $(2,5,29)$ 開始，而這些解與無理數能被分數逼近到甚麼程度有關。見練習 4。

思維技巧

無窮遞降法

要證明一條方程無解，可證明任何一個解都會產生一個嚴格更小的解。這樣，最小的解便不可能存在。

把它與模 n 運算配合使用，找出推動遞降的整除性；如果命題在遞降步驟下不能保持，就要準備把命題推廣。

延伸練習

1 **STEP I 熱身題。**（*STEP I 2010 第 8 題。*）證明 $a^3+3b^3=9c^3$ 的唯一整數解是 $a=b=c=0$。然後，藉考慮末位數字，證明 $p^4+2q^4=5r^4$ 也有同樣的結論。

2 **只用模運算，毋須遞降。**只用模 4 運算，證明 $x^2+y^2=4z+3$ 完全沒有整數解。

3 **有理點。**利用第 (iii) 部分，證明圓 $x^2+y^2=3$ 上沒有坐標為有理數的點。（試與 $x^2+y^2=2$ 對比，後者包含 $(1,1)$。）

4 **馬爾可夫三元組。**證明若 (x,y,z) 滿足 $x^2+y^2+z^2=3xyz$，則 $(x,y,3xy-z)$ 亦滿足。由 $(1,1,1)$ 出發，生成上文列出的三元組。（這個「跳躍」技巧正好與遞降相反。）

延伸閱讀

- **Harold Davenport, *The Higher Arithmetic***，第 2 及第 5 章。
- **Joseph Silverman, *A Friendly Introduction to Number Theory***（Pearson）。非常易讀；包括費馬對 $x^4 + y^4 = z^2$ 的遞降證明。
- **Simon Singh, *Fermat's Last Theorem***（Fourth Estate）。為一般讀者而寫的故事，由費馬的頁邊筆記講到懷爾斯（Wiles）的證明。

1.6 兩個平方數之和

題目

STEP II 2023，第 7 題

繼續閱讀之前，請先看題目（原題為英文）。題目載於 OCR 提供下載的 2023 年 STEP 試卷（zip 檔案）內的 STEP II 試卷：

`www.ocr.org.uk/Images/707102-2023-past-papers.zip`

引子

哪些整數可以寫成兩個平方數之和？$5 = 1^2 + 2^2$ 和 $13 = 2^2 + 3^2$ 可以，但 3、7 和 21 卻不行。這裏有一個奇怪的事實：$65 = 5 \times 13$ 可以用兩種不同的方法寫成兩個平方數之和，即 $1^2 + 8^2$ 和 $4^2 + 7^2$。為甚麼兩個這樣的數相乘，會得出另一個這樣的數？又為甚麼會有兩種寫法？

答案來自一個意想不到的方向。這道題目要你放下整數，改用複數（complex number）來思考。在複數的世界裏，兩平方和變得很簡單，一些看似令人摸不着頭腦的算術也變得近乎機械化。

動筆之前

想一想 若 $z = a + \mathrm{i}b$，其中 a 和 b 為整數，那麼 $|z|^2$ 是甚麼？

它是 $a^2 + b^2$，即兩個平方數之和。所以「N 是兩個平方數之和」的意思正正就是「存在某個實部（real part）和虛部（imaginary part）皆為整數的複數 z，使得 $N = |z|^2$」。這樣的複數稱為高斯整數（Gaussian integer）。請牢記這個轉譯；整道題目都建基於此。

逐步解題

階段 1　模相乘

想一想 以 a、b、c、d 寫出 zw。關於 $|zw|^2$，你需要證明甚麼？

我們有 $zw = (ac - bd) + \mathrm{i}(ad + bc)$，所以

$$|zw|^2 = (ac - bd)^2 + (ad + bc)^2 = a^2c^2 + b^2d^2 + a^2d^2 + b^2c^2,$$

因為交叉項 $-2abcd$ 和 $+2abcd$ 互相抵消。而這個式子可以因式分解（factorise）：

$$a^2c^2 + b^2d^2 + a^2d^2 + b^2c^2 = (a^2 + b^2)(c^2 + d^2) = |z|^2|w|^2.$$

取平方根，得 $|zw| = |z||w|$。 ■

現在把它看成一個關於整數的命題：

$$(a^2 + b^2)(c^2 + d^2) = (ac - bd)^2 + (ad + bc)^2. \qquad (\star)$$

兩個兩平方和的乘積，仍是一個兩平方和，而複數更確切地告訴你那兩個平方數是甚麼。

階段 2　學以致用

第 (ii) 部分：$5 = |2 + \mathrm{i}|^2$，而 $221 = 10^2 + 11^2 = |10 + 11\mathrm{i}|^2$。於是相乘：

$$(2 + \mathrm{i})(10 + 11\mathrm{i}) = 20 + 22\mathrm{i} + 10\mathrm{i} - 11 = 9 + 32\mathrm{i},$$

得出 $5 \times 221 = 9^2 + 32^2$。（驗算：$81 + 1024 = 1105$。）

想一想 如果改為把 $10 + 11\mathrm{i}$ 乘以 $2 - \mathrm{i}$，會怎麼樣？它的模（modulus）與 $2 + \mathrm{i}$ 的相同。

$(2 - \mathrm{i})(10 + 11\mathrm{i}) = 20 + 22\mathrm{i} - 10\mathrm{i} + 11 = 31 + 12\mathrm{i}$，所以同樣有 $1105 = 31^2 + 12^2$。選用 $2 + \mathrm{i}$ 還是它的共軛（conjugate），會得出不同的表示法。這就解釋了引子中 65 的兩種寫法（圖 1.6.1）：$65 = |(2 + \mathrm{i})(3 + 2\mathrm{i})|^2 = |4 + 7\mathrm{i}|^2$ 以及 $65 = |(2 - \mathrm{i})(3 + 2\mathrm{i})|^2 = |8 + \mathrm{i}|^2$。

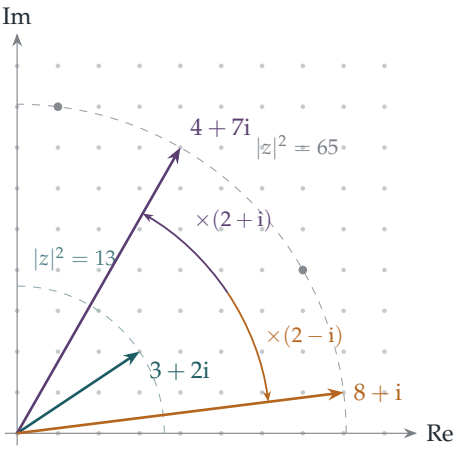


圖 1.6.1　乘以 $2 \pm \mathrm{i}$ 會把長度放大 $\sqrt{5}$ 倍，並旋轉約 26.6°：$2 + \mathrm{i}$ 為逆時針方向，$2 - \mathrm{i}$ 為順時針方向。由圓 $|z|^2 = 13$ 上的 $3 + 2\mathrm{i}$ 出發，兩種選擇會落在圓 $|z|^2 = 65$ 上不同的格點（lattice point）。灰色的點是它們對直線 $y = x$ 的反射。

階段 3　先作因式分解

? **想一想**　8045 的末位數字是 5。這提示了甚麼？

$8045 = 5 \times 1609$。我們需要把 1609 寫成兩個平方數之和；試試略小於 1609 的平方數，很快便找到 $1609 = 40^2 + 3^2$。於是

$$(2 + \mathrm{i})(40 + 3\mathrm{i}) = 77 + 46\mathrm{i}, \qquad \text{所以} \qquad 8045 = 77^2 + 46^2,$$

而共軛亦給出 $8045 = 83^2 + 34^2$。

第 (iv) 部分更快：$36 = 6^2 + 0^2$，而乘以 6 只是按比例放大：$36 \times 50805 = 612^2 + 1206^2$。

階段 4　三種寫法

? **想一想**　25 可以用多於一種方法寫成兩個平方數之和。若不計正負號和次序對調，有多少個高斯整數的模為 5？

$25 = 5^2 + 0^2 = 3^2 + 4^2$，分別對應 5、$3 + 4\mathrm{i}$ 和 $3 - 4\mathrm{i}$。接着看 1002082：由於 $1001^2 = 1002001$，我們有 $1002082 = 1001^2 + 9^2$。把 $1001 + 9\mathrm{i}$ 分別乘以這三個數，得：

$$5\,(1001 + 9\mathrm{i}) = 5005 + 45\mathrm{i},$$
$$(3 + 4\mathrm{i})(1001 + 9\mathrm{i}) = 2967 + 4031\mathrm{i},$$
$$(3 - 4\mathrm{i})(1001 + 9\mathrm{i}) = 3039 - 3977\mathrm{i}.$$

所以三對 (r, s) 是 $(45, 5005)$、$(2967, 4031)$ 和 $(3039, 3977)$。

階段 5　除法

> **停一停，試一試**　第 (vi) 部分把這部機器倒過來運轉。我們知道 $109 = 10^2 + 3^2$ 和 $1002037 = 1001^2 + 6^2$。怎樣才能不靠搜尋，就把 9193 寫成兩個平方數之和？

既然高斯整數相乘時模會相乘，那麼相除時模就應該相除。試試

$$\frac{1001 + 6\mathrm{i}}{10 + 3\mathrm{i}} = \frac{(1001 + 6\mathrm{i})(10 - 3\mathrm{i})}{109} = \frac{10028 - 2943\mathrm{i}}{109} = 92 - 27\mathrm{i}.$$

剛好整除，所以 $9193 = 92^2 + 27^2$。（驗算：$8464 + 729 = 9193$。）

> **想一想**　如果改為除以 $10 - 3\mathrm{i}$，會怎麼樣？

你會得到 $(1001 + 6\mathrm{i})(10 + 3\mathrm{i})/109 = (9992 + 3063\mathrm{i})/109$，這不是高斯整數，因為 9992/109 不是整數。在 109 的兩個「因數」之中，只有一個能整除 $1001 + 6\mathrm{i}$，正如只有某些質數（prime）能整除一個給定的整數。高斯整數有它們自己的質數，也有它們自己的因式分解。

回顧

- **轉譯、求解、再譯回來。**「兩個平方數之和」變成了「高斯整數的模的平方」。用這種語言來說，問題就變得容易。
- **先作因式分解。**每一部分都是先把一個大數拆成較小的數，而這些較小的數的表示法很容易找到。
- **共軛帶來新答案。**表示法的數目，取決於有多少種方法選擇一個因數或它的共軛。

背後的數學

費馬二平方和定理

哪些質數是兩個平方數之和？$2 = 1^2 + 1^2$、5、13、17、29、37、41 是；3、7、11、19、23、31 則不是。規律是：除 2 以外，一個質數是兩個平方數之和，當且僅當它除以 4 的餘數為 1。費馬（Fermat）於 1640 年提出這個結果；歐拉（Euler）於 1749 年給出證明。用高斯整數的語言來說：質數 $p \equiv 1 \pmod 4$ 會分裂（split）成 $p = (a + b\mathrm{i})(a - b\mathrm{i})$，而質數 $p \equiv 3 \pmod 4$ 則仍然是質數。例如 $5 = (2 + \mathrm{i})(2 - \mathrm{i})$，但 3 完全無法分解。

1990 年，Don Zagier 發表了二平方和定理的一個證明，它以全長只有一句而聞名。這

個證明連同解說收錄於 *Proofs from THE BOOK*。

一個擁有自己質數的新數系

高斯（Gauss）發現，高斯整數的表現與普通整數極為相似。它們有除法算法（division algorithm）和歐幾里得算法（Euclidean algorithm），而且每個高斯整數都可以唯一地分解成高斯質數（Gaussian prime）的乘積（不計乘以 ± 1 或 $\pm i$ 的差別）。結合 $(\star)$，你只需知道一個數的質因數分解，就能準確數出它有多少種方法寫成兩個平方數之和。

在大學裏，這是通往代數數論（algebraic number theory）的大門：研究 $\mathbb{Z}[i]$ 或 $\mathbb{Z}[\sqrt{-5}]$ 這類數系，藉此了解普通的整數。Section 1.7 說明了為甚麼需要創立這門學科：並非每個這樣的數系都像高斯整數那樣規矩。

四個平方數及其他

每個正整數都是四個平方數之和（拉格朗日（Lagrange），1770 年）。證明依賴 $(\star)$ 的四平方版本，它源自哈密頓（Hamilton）的四元數（quaternion），情形就如 $(\star)$ 源自複數一樣。你會在 Section 7.8 見到它。

思維技巧

轉換場景

一個問題在某個數系中顯得棘手，在一個更大的數系中卻可能十分自然。

兩平方和在整數中很棘手，在複數中卻不過是模而已。數學家經常這樣擴大數系：由整數到有理數，由有理數到實數，由實數到複數。每次擴大，都會令一些舊問題變得容易。

延伸練習

1. **一個障礙。**證明沒有任何形如 $4k+3$ 的數是兩個平方數之和。（平方數模 4 是甚麼？）
2. **兩種寫法。**把 65 和 85 各自用兩種不同的方法寫成兩個平方數之和，並用高斯整數解釋你的答案。
3. **四種寫法。**$1105 = 5 \times 13 \times 17$。找出四種不同的方法，把 1105 寫成兩個正平方數之和。為甚麼是四種？

4 **高斯質數。** 較難 證明若 $3 = (a + bi)(c + di)$，其中 a, b, c, d 為整數，則其中一個因數的模為 1。（取模的平方。）這說明 3 是高斯質數。

延伸閱讀

- **John Stillwell, *Elements of Number Theory***（Springer）。循序漸進講到高斯整數及更深入的內容，並附歷史背景。
- **Aigner & Ziegler, *Proofs from THE BOOK***，第 4 章，“Representing numbers as sums of two squares”（把數表示為兩個平方數之和）。
- **Joseph Silverman, *A Friendly Introduction to Number Theory***，有關平方和與高斯整數的各章。

1.7 因數分解失效之時

題目

STEP I 2016，第 7 題

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

`step.maths.org/questions/16-s1-q7`

引子

每個大於 1 的整數都可以寫成質數（prime）的乘積，而且寫法只有一種：$60 = 2 \times 2 \times 3 \times 5$，再沒有別的寫法。這個事實你大概已經用過無數次，熟悉得讓人覺得顯然成立，根本不用證明。

這道題構造了一個小小的數字世界，看上去完全合情合理，但在這個世界裏，那個「顯然」的事實卻是錯的。這是一道 STEP I 的題目，所用的數學並不艱深，但當中的教訓卻十分深刻。1847 年，一個著名的費馬最後定理（Fermat's Last Theorem）證明嘗試，正是敗在這個教訓之上，而代數學的一個全新分支亦因此誕生。

動筆之前

想一想 列出 S 和 T 的首幾個元素。按照第 (iv) 部分的定義，S 的哪些元素是S-質數？

$S = \{1, 5, 9, 13, 17, 21, 25, 29, 33, \ldots\}$，而 $T = \{3, 7, 11, 15, 19, 23, \ldots\}$。$S$ 中的一個數如果不能寫成 S 中較小元素（1 除外）的乘積，它就是 S-質數。留意 9 是 S-質數，儘管在普通整數中 $9 = 3 \times 3$，因為 3 並不在 S 之中。一個數是否「質數」，取決於你獲准使用哪些數。

逐步解題

階段 1　併集、交集與乘積

$S \cup T$ 是所有正奇數組成的集合，即兩者的併集（union），因為每個奇數除以 4 的餘數不是 1 就是 3。$S \cap T$ 是空集，即兩者的交集（intersection）沒有任何元素，因為沒有一個數會有兩個不同的餘數。

> **想一想**　S 的一般元素應該怎樣寫？兩個這樣的元素相乘，乘積是甚麼？

S 的元素是 $4a+1$，其中 $a \geq 0$，而

$$(4a+1)(4b+1) = 16ab + 4a + 4b + 1 = 4(4ab + a + b) + 1 \in S.$$

所以 S 在乘法下封閉（closed）。至於 T：

$$(4a+3)(4b+3) = 16ab + 12a + 12b + 9 = 4(4ab + 3a + 3b + 2) + 1,$$

這個數屬於 S，而不屬於 T。例如 $3 \times 7 = 21 \in S$。所以 T 在乘法下並不封閉。

階段 2　屬於 T 的質因數

> **想一想**　在 T 中取一個不是質數的數。它的質因數（prime factor）全是奇數。假如這些質因數全都屬於 S，會有甚麼後果？

這個數是奇數，所以它的質因數全是奇數，各自屬於 S 或 T。假如它們全部屬於 S，那麼由於 S 在乘法下封閉，它們的乘積也會屬於 S。但這個乘積就是我們原來的數，而它屬於 T。所以至少有一個質因數屬於 T。 ■

階段 3　奇數個因數

> **想一想**　把偶數個 T 的元素乘起來，乘積會落在哪裏？奇數個又如何？

每一對 T-數相乘都得到一個 S-數，而 S-數相乘仍然是 S-數，所以偶數個 T-因數的乘積屬於 S。奇數個的話，乘積就是（S 中的某個數）$\times$（T 的一個元素），用同類的計算可知它屬於 T。

現在取 $t \in T$。如果 t 是 T-質數，那就完成了。否則，把 t 寫成 T 的元素的乘積，其中每個元素都大於 1，因而小於 t。當中任何不是 T-質數的因數都可以再分解下去，如此類推。因數每次都變小，所以這個過程必會停止（這裏再次用到良序性質（well-ordering）），最後 t 就寫成了 T-質數的乘積。由於乘積屬於 T，因數的個數必定是奇數。 ■

階段 4　兩種分解

> **停一停，試一試** 在 S 中找一個數，它有兩種真正不同的 S-質數分解。提示：利用像 9 這樣的數，它們之所以是 S-質數，只因為它們在普通整數中的質因數並不在 S 之中。

S-質數依次為 $5, 9, 13, 17, 21, 29, 33, 37, 41, 49, \ldots$。$21 = 3 \times 7$、$9 = 3 \times 3$ 和 $49 = 7 \times 7$ 全都是 S-質數，因為 3 和 7 都不在 S 之中。而且

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

這就是 441 的兩種不同的 S-質數分解（圖 1.7.1）。唯一分解（unique factorisation）在 S 中並不成立。 ■

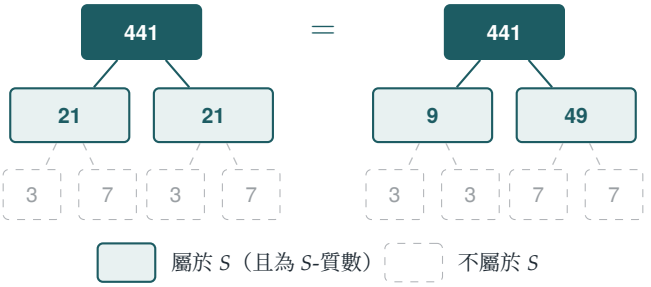


圖 1.7.1 441 的兩種 S-質數分解。在普通整數中，兩棵樹都會繼續向下分解至 $3 \times 3 \times 7 \times 7$，但 3 和 7 都不在 S 之中，所以在 S 之內，分解到第二行便停止，而兩種分解並不相同。

回顧

- **是否「質數」取決於所在的系統。**9 和 21 在整數中是合數（composite），在 S 中卻是質數，因為分解它們所需的數並不存在。
- **封閉性很重要。**S 在乘法下封閉；T 則不然。第 (iii) 和第 (iv) 部分的論證，全靠這一點分別。
- **顯然的事實也可以是一條定理。**整數的唯一分解是需要證明的。這道題說明了原因：在一個看上去非常相似的系統中，它並不成立。

背後的數學

整數有何特別之處

為甚麼分解在普通整數中確實行得通？關鍵在於歐幾里得引理（Euclid's lemma）：如果質數 p 整除乘積 ab，那麼 p 整除 a 或 p 整除 b。這條引理是用輾轉相除法（Euclidean

algorithm）證明的，而它正正是在 S 中失效的地方。在那裏，S-質數 9 整除 $21 \times 21 = 441$（因為 $441 = 9 \times 49$），但 9 在 S 中並不整除 21。到了大學一年級，你會嚴謹地證明歐幾里得引理，並看到唯一分解如何由它推導出來。

費馬最後定理的失敗證明

1847 年 3 月 1 日，拉梅（Gabriel Lamé）向巴黎科學院宣布，他已證明了費馬最後定理。他的想法是利用複數單位根（roots of unity）把 $x^n + y^n$ 分解，

$$x^n + y^n = (x+y)(x+\zeta y)(x+\zeta^2 y)\cdots(x+\zeta^{n-1}y), \qquad \zeta = e^{2\pi \mathrm{i}/n}$$

（其中 n 為奇數），然後就每個括號的質因數進行論證。劉維爾（Joseph Liouville）隨即提出質疑：由 ζ 構成的數，憑甚麼一定有唯一分解？數星期後，人們才知道庫默爾（Ernst Kummer）早已證明這類數往往沒有唯一分解。最先失效的情況是 $n = 23$。

庫默爾盡力補救，發明了「理想數」（ideal numbers），這些額外的對象能夠恢復唯一分解，他並藉此對一大類指數證明了費馬最後定理。戴德金（Richard Dedekind）把庫默爾的理想數發展成*理想*（ideal），成為近世代數的核心概念之一。

一個值得記住的較小例子

在數系 $\mathbb{Z}[\sqrt{-5}] = \{a + b\sqrt{-5} : a, b \in \mathbb{Z}\}$ 中，

$$6 = 2 \times 3 = (1+\sqrt{-5})(1-\sqrt{-5}),$$

而這四個因數在該數系中全都是「質數」（練習 3）。試把這個例子與 Section 1.6 的高斯整數（Gaussian integers）比較，後者的分解*確實*是唯一的。這類數系之中哪些具有唯一分解，是一個很深刻的問題。高斯（Gauss）猜想，在形如 $\mathbb{Z}[\sqrt{-d}]$ 的數系及與其相近的數系之中，只有九個具有唯一分解。這個猜想最終由黑格納（Heegner）、貝克（Baker）和斯塔克（Stark）在 1952 年至 1967 年間證明。

思維技巧

> **質疑你視為理所當然的事**
>
> *當一個事實看來顯然得不用證明時，問問它依賴甚麼，並找出一個令它失效的情境。*
>
> 構造反例（counterexample），也就是構造一個令「顯然」的事實不成立的世界，是弄清一個證明真正需要甚麼的最佳方法之一。大學本科代數有很大部分，就是

在探問整數那些熟悉的性質，有哪些能在更一般的系統中保留下來。

延伸練習

1 **另一個例子。**證明 693 有兩種不同的 S-質數分解。你能找出一個製造這類例子的通用方法嗎？

2 **偶數。**設 $E = \{2, 4, 6, 8, \ldots\}$。哪些數是 E-質數？證明 60 有多於一種 E-質數分解。

3 $\mathbb{Z}[\sqrt{-5}]$。定義範數（norm）$N(a + b\sqrt{-5}) = a^2 + 5b^2$。證明 $N(zw) = N(z)N(w)$。先證明沒有任何元素的範數是 2 或 3，由此推導出 2、3 和 $1 \pm \sqrt{-5}$ 在 $\mathbb{Z}[\sqrt{-5}]$ 中不能再分解。

4 **歐幾里得引理的應用。**較難 假設歐幾里得引理成立（如果一個質數整除 ab，它就整除 a 或 b），證明每個大於 1 的整數都只有一種質數分解，因數的次序不計。

延伸閱讀

- **Ian Stewart & David Tall, *Algebraic Number Theory and Fermat's Last Theorem***（CRC Press）。開首幾章講述拉梅、庫默爾與理想的故事，程度適合好學的大學一年級學生。
- **John Stillwell, *Elements of Number Theory***。把整數的唯一分解與高斯整數的唯一分解並列比較。
- **Harold Edwards, *Fermat's Last Theorem: A Genetic Introduction***（Springer）。適合有志挑戰的讀者：循歷史發展解說庫默爾的工作。

1.8 點算有理數

題目

STEP III 2020，第 8 題

繼續閱讀之前，請先看題目（原題為英文）。題目載於 OCR 提供下載的 2020 年 STEP 試卷（zip 檔案）內的 STEP III 試卷：

`www.ocr.org.uk/Images/696589-2020-past-papers.zip`

引子

你能否把所有正分數逐一列出，使每個分數都恰好出現一次？這看來是不可能的：任意兩個分數之間還有無窮多個分數，根本沒有「下一個」分數可以寫下來。然而，康托爾（Georg Cantor）在 1870 年代證明了這是做得到的。這表示，在某個精確的意義下，分數並不比整數多。

康托爾的列表頗為笨拙：表中既有 $\frac{1}{2}$，也有 $\frac{2}{4}$，重複的項要逐一刪去。本題要構造的列表卻完全沒有重複，而且來自一個規則簡單得出奇的數列（sequence）。這個列表由 Neil Calkin 和 Herbert Wilf 於 2000 年發表，不過數列本身早在 1858 年已由 Moritz Stern 研究過。

動筆之前

想一想 在證明任何東西之前，先計算最初十五項左右。你留意到比值 u_n/u_{n+1} 有甚麼特別之處？

數列的開首幾項是

$$1,\ 1,\ 2,\ 1,\ 3,\ 2,\ 3,\ 1,\ 4,\ 3,\ 5,\ 2,\ 5,\ 3,\ 4,\ 1,\ 5,\ \ldots$$

而比值 u_n/u_{n+1} 依次是

$$\frac{1}{1},\ \frac{1}{2},\ \frac{2}{1},\ \frac{1}{3},\ \frac{3}{2},\ \frac{2}{3},\ \frac{3}{1},\ \frac{1}{4},\ \frac{4}{3},\ \frac{3}{5},\ \frac{5}{2},\ \frac{2}{5},\ \frac{5}{3},\ \frac{3}{4},\ \frac{4}{1},\ \ldots$$

每個分數都以最簡分數（lowest terms）的形式出現，而且沒有一個重複。題目要你證明

這個現象會一直持續下去，並且每個正分數遲早都會出現。

逐步解題

階段 1　奇數項勝過偶數項

相鄰兩項不是 (u_{2k}, u_{2k+1})，就是 (u_{2k+1}, u_{2k+2})。

> **想一想**　試用 u_k 和 u_{k+1} 表示這兩種數對。哪一項較大？

$(u_{2k}, u_{2k+1}) = (u_k,\ u_k + u_{k+1})$，而 $(u_{2k+1}, u_{2k+2}) = (u_k + u_{k+1},\ u_{k+1})$。由於每一項都是正整數（用歸納法（induction）不難證明），$u_k + u_{k+1}$ 比 u_k 和 u_{k+1} 都大。唯一的例外是第一對 $(u_1, u_2) = (1, 1)$，它不符合上述任何一種形式，因為 u_1 是直接定義的。 ■

階段 2　相鄰兩項互質

> **想一想**　假設 $d > 1$ 同時整除 u_{2k} 和 u_{2k+1}，即同時整除 u_k 和 $u_k + u_{k+1}$。那麼 d 還整除甚麼？

d 整除 $(u_k + u_{k+1}) - u_k = u_{k+1}$，所以 d 整除相鄰數對 u_k, u_{k+1}，而這一對在數列中的位置較前。另一種情況完全相同：若 d 整除 $u_k + u_{k+1}$ 和 u_{k+1}，它便整除 u_k。

因此，相鄰數對的公因數會傳回較前的一對，然後再傳回更前的一對，如此類推。位置不可能無止境地減小（又是良序性（well-ordering)），所以 d 最終必須整除第一對 $u_1 = u_2 = 1$。當 $d > 1$ 時這是不可能的。所以每一對相鄰項都互質（coprime）。 ■

階段 3　沒有數對出現兩次

> **停一停，試一試**　階段 1 告訴你一件有用的事：看到一個相鄰數對 (a, b) 時，你可以判斷 a 位於奇數位置還是偶數位置。怎樣判斷？然後，你能否還原出階段 2 所指的、在它之前的那個數對？

若 $a < b$，則（由階段 1）b 位於奇數位置，所以對某個 k，這個數對是 $(u_{2k}, u_{2k+1}) = (u_k, u_k + u_{k+1})$，而它的「父」數對 (u_k, u_{k+1}) 就是 $(a,\ b - a)$。若 $a > b$，這個數對是 (u_{2k+1}, u_{2k+2})，其父數對是 $(a - b,\ b)$。若 $a = b$，階段 1 表明這個數對必定是最初的一對 $(1, 1)$。

所以除第一對外，每個相鄰數對都決定了它的父數對，以及父數對的位置（這個結構可以看成一棵樹，見圖 1.8.1）：當子數對位於 $2k$ 或 $2k + 1$ 時，父數對位於位置 k。現在假設 (a, b) 出現在兩個不同的位置。它們的父數對便是出現在兩個位置的同一數對，而這

些父數對的父數對亦然，如此類推。這樣一直追溯上去，我們最終會在 1 以外的位置找到 $(1,1)$，而階段 1 已排除了這個可能。 ■

階段 4 每個互質數對都會出現

> **想一想** 假設某個互質數對 (a,b) 從未以相鄰兩項的形式出現。你應該考慮哪一個這樣的數對？

在所有「缺席」的互質數對之中，取 $a+b$ 的值最小的一個：這就是極小反例（minimal counterexample）。它不可能是 $(1,1)$，因為這一對確實出現了，所以 $a \neq b$（互質而又相等，必有 $a=b=1$）。

- 若 $a > b$：數對 $(a-b,b)$ 互質（$a-b$ 和 b 的任何公因數都整除 a），而它的和 a 小於 $a+b$。由極小性，這個數對確實出現，設它位於位置 $k,k+1$。但這樣便有 $u_{2k+1} = u_k + u_{k+1} = a$ 和 $u_{2k+2} = u_{k+1} = b$，所以 (a,b) 終究還是出現了。
- 若 $a < b$：同理，$(a,b-a)$ 出現在某個 $k,k+1$，於是 $(u_{2k},u_{2k+1}) = (a,b)$。

無論哪種情況都得出矛盾。所以每一對互質的正整數都會以相鄰兩項的形式出現。 ■

（題目把第 (iv) 部分寫成遞降法（descent）：一個缺席的數對會產生一個更小的缺席數對。這其實是同一個論證，只是倒過來讀而已。）

階段 5 所有分數的列表

每個正有理數（rational number）都可以唯一地寫成最簡分數 a/b，其中 a,b 互質。由階段 4，(a,b) 會以某個相鄰數對 (u_n,u_{n+1}) 的形式出現，所以 $f(n) = a/b$：f 的值域（range）是全體正有理數。由階段 3，它只出現一次；又由階段 2，每個值 $f(n)$ 本身已是最簡分數，所以兩個不同的 n 不可能給出同一個分數。因此 f 是由正整數到正有理數的雙射（bijection），並且有逆函數（inverse）。 ■

回顧

- **遞降與它的逆向。**公因數和重複的數對都是往回追溯到數列的起點。缺席的數對則用極小反例處理，向前構造。
- **雙射要證兩半。**「每個分數都出現」（滿射，onto）和「沒有分數出現兩次」（單射，one-to-one）是兩個獨立的命題，要分開證明。
- **先動手計算。**寫出十五項之後，這些命題便顯得可信，同時也提示了推動整個證明的父子結構。

背後的數學

Calkin–Wilf 樹

把正分數排成一棵樹：把 $\frac{1}{1}$ 放在頂端，並給每個分數 $\frac{a}{b}$ 兩個子節點，左邊是 $\frac{a}{a+b}$，右邊是 $\frac{a+b}{b}$。逐行由左至右讀這棵樹，得到的正是本題的列表 $f(1), f(2), f(3), \ldots$（圖 1.8.1）。階段 3 說明每個分數都有唯一的父節點，階段 4 說明每個分數都在樹中。2003 年，Moshe Newman 發現連這棵樹也不需要：列表中每個分數都可以由前一個分數算出，公式是

$$x_{n+1} = \frac{1}{2\lfloor x_n \rfloor - x_n + 1}.$$

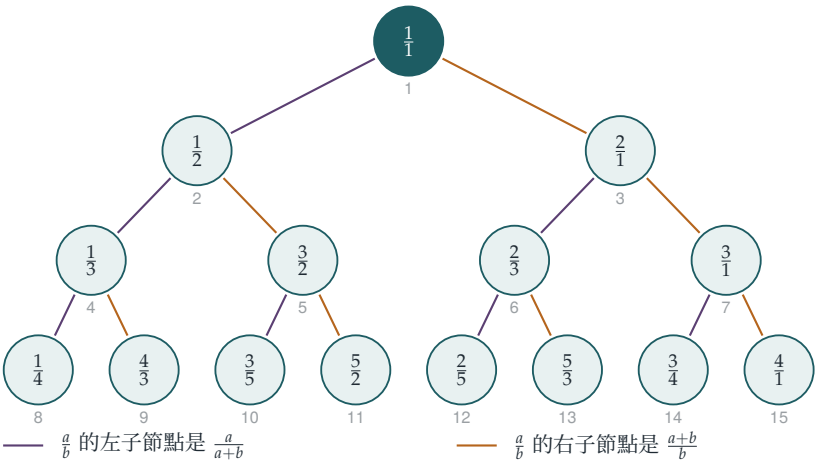


圖 1.8.1　Calkin–Wilf 樹的首四行。逐行由左至右讀，便得到 $f(1), f(2), f(3), \ldots$；灰色小數字是位置 n。每個正分數都以最簡分數的形式恰好出現一次。

不同大小的無窮

可以這樣逐一列出的集合稱為可數（countable）集合。分數是可數的，這令人意外，因為分數看來比整數多得多。在大學的第一個學期，你還會遇到康托爾的對角論證法（diagonal argument），它證明實數（real numbers）是不可數的：任何實數列表都必定有所遺漏。所以無窮確實有不同的大小，而實數所構成的無窮嚴格地大於有理數所構成的無窮。這是你在大學一年級課程中會遇到的最令人震驚的結果之一。

Stern 數列

數列 u_n 稱為 *Stern 雙原子數列*（Stern's diatomic sequence），它還有許多其他面貌。例如，u_{n+1} 等於把 n 寫成 2 的冪之和、而每個冪最多用兩次的方法數目。它與 *Stern–Brocot 樹*以及連分數（continued fractions）關係密切，而連分數能給出無理數的最佳有理逼近。

思維技巧

構造雙射

要證明兩個集合大小相同，就要把它們的元素準確地一一配對：證明目標集合的每個元素都被對應到（滿射），而且沒有元素被對應兩次（單射）。

雙射是計數的基礎，不論有限還是無窮，你在大學裏會經常用到。在本題中，兩半的證明都是把一個結構追溯回它的起點。

延伸練習

1 **找出分數。**對哪一個 n 有 $f(n) = \frac{3}{5}$？又 $f(n) = \frac{5}{3}$ 呢？利用父節點的規則倒推回 $\frac{1}{1}$。

2 **Newman 公式。**驗證由 $x_1 = 1$ 開始，Newman 公式能給出列表中的首十個分數。

3 **康托爾的對角論證法。**假設你有一個列表，以小數形式列出了 0 與 1 之間的所有實數。構造一個小數，使它與列表中第 n 個數在第 n 位數字上不同。為甚麼這就說明列表並不完整？（對於像 $0.0999\ldots = 0.1$ 這樣的小數，需要留意甚麼？）

4 **可數並集。** 較難 給出一個與正整數之間的明確雙射，從而證明由全體正整數數對 (m, n) 組成的集合是可數的。（試試 $2^{m-1}(2n-1)$。）

延伸閱讀

- **Neil Calkin & Herbert Wilf, "Recounting the rationals"**（*American Mathematical Monthly*，2000 年；〈重新數算有理數〉）。全文四頁，完全讀得懂。
- **Aigner & Ziegler, *Proofs from THE BOOK***（《數學天書中的證明》），第 17 章 "Sets, functions, and the continuum hypothesis"（集合、函數與連續統假設），當中包括 Calkin–Wilf 樹。

- **Ian Stewart & David Tall, *The Foundations of Mathematics***（《數學基礎》，Oxford University Press）。可數性、康托爾與實數，內容與大學一年級課程相若。

1.9 三次曲線的切線

題目

STEP I 2009，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/09-s1-q2`

引子

這題看來只是一道例行的坐標幾何（coordinate geometry）題：求導、找切線、解三次方程。可是最後一行卻要你找出一個立方數，它是另外三個立方數之和，而幾何竟像變魔術一樣把它變了出來。

這個竅門是過曲線上一個已知點作直線，從而找出新的點，它可追溯至公元三世紀亞歷山大港的丟番圖（Diophantus）。到了現代數學家手上，它發展成橢圓曲線（elliptic curves）的理論。這套理論位於當今數論的核心，是費馬最後定理（Fermat's Last Theorem）證明的基礎，也保障着全球一大部分互聯網通訊的安全。

動筆之前

? **想一想** 點 $(-a, b)$ 在曲線上。為甚麼？這跟立方和又有甚麼關係？

代入：$(-a)^3 + a^3 + b^3 = b^3$，所以確實在曲線上。在 $y^3 = x^3 + a^3 + b^3$ 上坐標為有理數（rational）的點正是我們想要的：若 $x = q/s'$ 而 $y = p/s'$，去分母後便得到一條立方數之間的關係式。我們的計劃是從一個容易找到的點出發，利用幾何找出一個沒那麼明顯的點。

逐步解題

階段 1　切線

對 $y^3 = x^3 + a^3 + b^3$ 作隱函數求導（implicit differentiation）：$3y^2\dfrac{dy}{dx} = 3x^2$，所以 $\dfrac{dy}{dx} = \dfrac{x^2}{y^2}$。在 $(-a, b)$ 處斜率（gradient）為 a^2/b^2，所以切線（tangent）是

$$y - b = \frac{a^2}{b^2}(x + a) \quad\Longrightarrow\quad b^2y - b^3 = a^2x + a^3 \quad\Longrightarrow\quad b^2y - a^2x = a^3 + b^3. \qquad \blacksquare$$

階段 2　切線與曲線再次相交之處

取 $a = 1$、$b = 2$，曲線是 $y^3 = x^3 + 9$，而在 $(-1, 2)$ 的切線是 $4y - x = 9$，即 $y = \dfrac{x + 9}{4}$。代入曲線方程：

$$\frac{(x+9)^3}{64} = x^3 + 9 \quad\Longrightarrow\quad x^3 + 27x^2 + 243x + 729 = 64x^3 + 576,$$

整理後得 $63x^3 - 27x^2 - 243x - 153 = 0$，再除以 9 便得 $7x^3 - 3x^2 - 27x - 17 = 0$。 $\blacksquare$

階段 3　解三次方程

> **想一想**　不用做任何代數運算，你已經知道這個三次方程的一個根。是哪一個？它為甚麼應該出現兩次？

切線在 $x = -1$ 處與曲線相切（圖 1.9.1），而「相切」意味着重根（repeated root），正如拋物線的切線與拋物線相交於某個二次方程的二重根。所以 $(x+1)^2$ 應該是一個因式。驗算：$7(-1)^3 - 3(1) - 27(-1) - 17 = -7 - 3 + 27 - 17 = 0$，而導數 $21x^2 - 6x - 27$ 在 $x = -1$ 處也等於零。相除後得

$$7x^3 - 3x^2 - 27x - 17 = (x+1)^2(7x - 17).$$

> **停一停，試一試**　所以切線與曲線再次相交於 $x = \frac{17}{7}$ 之處。求出 y，然後把這個點化為一條立方數之間的關係式。

$y = \dfrac{17/7 + 9}{4} = \dfrac{80/7}{4} = \dfrac{20}{7}$。點 $\left(\frac{17}{7}, \frac{20}{7}\right)$ 在 $y^3 = x^3 + 9$ 上，所以

$$\frac{20^3}{7^3} = \frac{17^3}{7^3} + 9 \quad\Longrightarrow\quad 20^3 = 17^3 + 9 \times 7^3.$$

> **想一想**　我們希望 9×7^3 是兩個立方數之和。回想一下那個 9 是怎樣來的。

$9 = a^3 + b^3 = 1^3 + 2^3$，所以 $9 \times 7^3 = 7^3 + 14^3$。因此

$$20^3 = 17^3 + 14^3 + 7^3.$$

驗算：$8000 = 4913 + 2744 + 343$。 ■

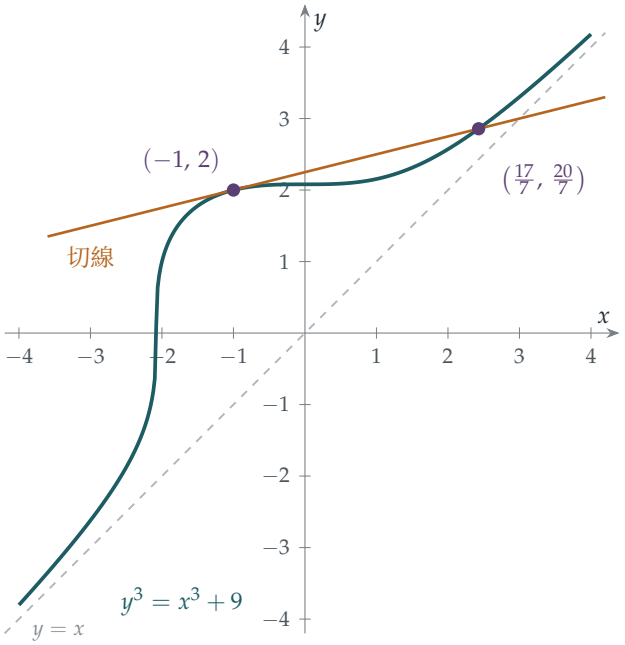


圖 1.9.1　曲線 $y^3 = x^3 + 9$（藍綠色）在遠離原點處趨近直線 $y = x$。在有理點 $(-1, 2)$ 的切線（琥珀色）在該處與曲線相切，這算作二重交點，並在另一個有理點 $\left(\frac{17}{7}, \frac{20}{7}\right)$ 再與曲線相交一次。

回顧

- **相切即二重根。**這令一個沒有明顯根的三次方程，變成可以即時因式分解的方程。
- **新的根自動是有理數。**三個根之中一旦有兩個是有理數（這裏兩個都是 -1），第三個也必定是有理數，因為各根之和 $\frac{3}{7}$ 是有理數。這正是整個方法的精髓。
- **幾何帶來了算術。**我們從沒有刻意搜尋立方數；是一條切線把它們找了出來。

背後的數學

弦切法

取一條三次曲線（cubic curve），例如 $y^3 = x^3 + 9$，或者更標準的形式 $y^2 = x^3 + Ax + B$。假如你知道曲線上兩個坐標為有理數的點，那麼過這兩點的直線恰好與曲線再相交於一點，而該點的坐標也是有理數（理由見回顧中的論證）。假如你只知道一個點，就

改用該點的切線。無論哪種情況，舊的有理點都能產生新的有理點。

值得注意的是，這個「弦切」（chord–tangent）作圖法可以變成一種把曲線上的點相加的方法，於是有理點便構成一個羣（group），這是你在大學一年級會學到的一種代數結構。莫德爾（Louis Mordell）在 1922 年證明了這個羣由有限個點生成：所有有理點，不論有多少個，都可以由一個有限集合出發，利用弦和切線構造出來。他的證明是 Section 1.5 中費馬（Fermat）下降法的一個版本。

今日的橢圓曲線

這類曲線稱為橢圓曲線（並非因為它們是橢圓，而是因為歷史上它們與橢圓的弧長有關）。它們在現代數學中無處不在。

- **費馬最後定理。**在 1980 年代，Gerhard Frey 留意到 $x^n + y^n = z^n$ 的一個解會產生一條非常古怪的橢圓曲線。懷爾斯（Andrew Wiles）在 1994 年證明了這樣古怪的曲線不可能存在，從而證明了費馬最後定理。
- **貝赫與斯維訥通－戴爾猜想（Birch and Swinnerton-Dyer conjecture）。**有理點所構成的羣需要多少個生成元（generators）？在 1960 年代，有人憑藉劍橋一些數學史上最早的電腦實驗提出了一個猜想，它用一個函數來預測答案，而這個函數是由點算曲線在模質數之下的點數構造出來的。它是七條千禧年大獎難題（Millennium Prize Problems）之一。
- **密碼學（cryptography）。**在有限數系上的橢圓曲線上，把一個點 P 自己相加 n 次是很快的，即使 n 有數百位數字也一樣（把 P 反覆加倍，就像你用反覆平方來計算高次冪那樣）。但已知 P 和結果 nP，要找出 n 卻似乎極之困難：這就是橢圓曲線離散對數問題（elliptic-curve discrete logarithm problem）。這種不對稱性保護着今日大部分的網上通訊。

同次冪之和

能寫成三個正立方數之和的最小立方數是 $6^3 = 3^3 + 4^3 + 5^3$；切線法找到的是 $20^3 = 7^3 + 14^3 + 17^3$，它並非前者的倍數，而重複使用這個方法會不斷產生新的恆等式（雖然有些會涉及負數）。要多少個 k 次冪才能湊成另一個 k 次冪？歐拉（Euler）在 1769 年猜想，至少需要 k 個 k 次冪才能湊成另一個 k 次冪。對於四次冪和五次冪，他錯了：1966 年 Leon Lander 和 Thomas Parkin 以電腦搜尋找到

$$27^5 + 84^5 + 110^5 + 133^5 = 144^5$$

而在 1988 年，Noam Elkies 利用橢圓曲線找到了四次冪的反例（counterexamples）。

思維技巧

以幾何解決算術

要找出一個方程的整數解或有理數解，不妨把它看成一條曲線，然後尋找過已知點的直線。

背後的代數很簡單：若一個有理係數多項式（polynomial）除了一個根之外其餘的根都是有理數，那麼最後一個根也是有理數。這樣在圖像與方程之間來回轉換，是數學家所能擁有的最強大的習慣之一。

延伸練習

1 **圓上的有理點。**過 $(-1,0)$ 而斜率為 t 的直線與圓 $x^2+y^2=1$ 再次相交。證明另一個交點是 $\left(\frac{1-t^2}{1+t^2}, \frac{2t}{1+t^2}\right)$，並由此得出無窮多組畢氏三元數（Pythagorean triples）。

2 **一條落空的弦。**點 $(1,2)$ 和 $(2,1)$ 在 $x^3+y^3=9$ 上。證明過這兩點的直線與曲線再沒有其他交點。那個「第三點」到哪裏去了？（查一查無窮遠點（point at infinity）。）

3 **故技重施。**求 $y^3=x^3+9$ 在 $\left(\frac{17}{7}, \frac{20}{7}\right)$ 的切線，以及它與曲線相交的第三點。（使用電腦代數系統會有幫助；留意數字增長得有多快。）

4 **的士數（taxicab numbers）。**$1729 = 1^3+12^3 = 9^3+10^3$。求過 $(1,12)$ 和 $(9,10)$ 的直線與曲線 $x^3+y^3=1729$ 相交的第三點，並由此推導出一條形如 $p^3=q^3+r^3+s^3$ 的新恆等式。

延伸閱讀

- **Joseph Silverman & John Tate, *Rational Points on Elliptic Curves***（Springer）。為大學本科生而寫；第 1 章所需的知識不超出英國高考（A-level）程度。
- **Avner Ash & Robert Gross, *Elliptic Tales***（Princeton University Press）。一本普及讀物，一路講到貝赫與斯維訥通－戴爾猜想。
- **Simon Singh, *Fermat's Last Theorem***。講述橢圓曲線如何解決一個有 350 年歷史的難題。

我們學到了甚麼

在本章開首，我們就整數提出了三個問題，並預告還有第四個主題貫穿其中：怎樣證明（prove）一件事。現在是時候看看，九道 STEP 題目把我們帶到了多遠。

1. 質數有多少個？它們是怎樣分佈的？

在 Section 1.1 中，我們證明了質數（prime）有無窮多個，卻連一個新的質數也沒有找出來。我們構造了一族無窮多個的數，即費馬數（Fermat numbers），它們兩兩之間沒有公因數，所以每一個都需要一個屬於自己的質數。這個證明談的是*結構*，而不是大小。

在 Section 1.2 中，我們更進一步，證明了不超過 n 的所有質數之積小於 4^n。這是我們第一個定量的結論：它限制了質數可以排得多密。取對數後，它說的是 $\sum_{p \le n} \ln p < 1.39\,n$，而質數定理（Prime Number Theorem）指出真實的值大約是 n。我們這個 STEP 程度的論證，已經相當接近了。

我們*沒有*做到的，是確切找出質數怎樣分佈。這由黎曼 ζ 函數（Riemann zeta function）所支配，而把這件事說得精確，就是黎曼猜想（Riemann Hypothesis），也就是數學中最著名的未解決問題。你現在所知的，已足以明白它在說甚麼。

2. 怎樣證明一個數不是分數？

我們找到了三種截然不同的方法。

- **取最小的那一個**（Section 1.3）。假如 $\sqrt{2}$ 是分數，就會有一個最小的正整數 k 使 $k\sqrt{2}$ 為整數；但這樣一來，$(\sqrt{2}-1)k$ 就是一個更小的。同樣的論證也適用於立方根。
- **把一個整數困在 0 與 1 之間**（Section 1.4）。假如 e 或 e^2 是分數，乘以一個階乘（factorial）後，就會得出一個嚴格介乎 0 與 1 之間的整數。埃爾米特（Hermite）和林德曼（Lindemann）證明 e 和 π 是超越數（transcendental），背後正是這個想法。
- **找出一個不可能成立的整除關係**（Section 1.5，習題 3）。證明 $x^2 + y^2 = 3$ 上沒有有理點（rational point），等同於證明當 x 是有理數時，$\sqrt{3 - x^2}$ 永遠不是有理數。

我們也認識到還有多少事情仍然未知：至今沒有人能證明 $e + \pi$ 是無理數（irrational）。

3. 哪些方程有整數解？

有時答案是*沒有*，而 Section 1.5 展示了費馬（Fermat）證明這一點的方法：任何一個解都會產生一個更小的解，如此無休無止。以一個精心挑選的質數為模（modulo）來運算，就能告訴我們該利用哪一個整除關係。

有時答案是*有*，*而且這裏有一部製造解的機器*。在 Section 1.6 中，複數（complex numbers）把二平方和變成了一條乘法法則。在 Section 1.9 中，三次曲線（cubic curve）的一條切線給出了 $20^3 = 17^3 + 14^3 + 7^3$。

還有些時候，正確的一步是*換一個數系*。高斯整數（Gaussian integers）令二平方和變得容易。但 Section 1.7 提出了一個警告：在一個看來完全合理的數系中，數可以用兩種不同的方式分解成「質數」。這個警告在 1847 年被人忽視，結果催生了理想（ideals）和近世代數。

Section 1.8 還另添了一重轉折：分數看似遠比整數多，卻可以逐一列舉出來。無窮比乍看之下有更豐富的結構。

4. 證明的工具箱

以下是你在本章用過的方法。每一種都會在大學數學課程中一再出現。

方法	你在哪裏用過
歸納法（induction）與強歸納法（strong induction）	費馬數（1.1）；質數之積（1.2）
良序性（well-ordering）、最小反例（minimal counterexample）	$\sqrt{2}$ 與 $\sqrt[3]{2}$（1.3）；列舉分數（1.8）
無窮遞降法（infinite descent）	不可能成立的方程（1.5）；互質（coprime）數對（1.8）
反證法（proof by contradiction）	貫穿全章，尤其是 1.3 和 1.4
模 n 運算	1.5；1.7
估界與夾逼（bounding and squeezing）	1.2；1.4
構造雙射（bijection）	1.8
構造反例（counterexample）	非唯一分解（1.7）
轉換場景	複數（1.6）；幾何（1.9）

回頭看看這些題目，你會留意到 STEP 的擬題人往往替你走了第一步：「考慮二項展開式（binomial expansion）…」、「考慮複數 $2 + \mathrm{i}$ 和 $10 + 11\mathrm{i}$…」。到了大學，這種幫助會逐漸消失，而找出正確的第一步就成了真正的本領。最好的準備正是你一直在做的事：每做完一道題，問問自己那個提示*爲甚麼*奏效，下次你就能自己想到。

下一步

本章有好幾條線索會在書中其他地方延續下去。

- 以分數逼近 e 的課題會在 Section 2.9 再次出現，而黎曼 ζ 函數則見於 Section 2.4，那裏有 $\zeta(2) = \pi^2/6$。
- 單位根（roots of unity）曾在拉梅（Lamé）那個失敗的證明中出現，它是 Chapter 5 的主題；而在 Section 1.1 中預告過的高斯（Gauss）正十七邊形作圖法，就是 Section 6.2。
- Section 1.6 中那條恆等式的四平方版本，會透過四元數（quaternions）在 Section 7.8 中出現。

如果你想在上大學前走得更遠

- **Harold Davenport, *The Higher Arithmetic***：初等數論入門書中最好的一本。
- **Aigner & Ziegler, *Proofs from THE BOOK***：第 1–4、8 和 17 章與本章直接相關。
- **Joseph Silverman, *A Friendly Introduction to Number Theory***：幾乎涵蓋本章的全部內容，而且還有更多，節奏平緩。
- **Marcus du Sautoy, *The Music of the Primes*** 和 **Simon Singh, *Fermat's Last Theorem***：數學背後的人物故事。

數論總有辦法把人吸引進去。它的問題簡單得第一天就可以提出，卻又深奧得足以耗上一生。如果其中某些章節令你很想知道接下來會怎樣，那正是本書的用意所在。

延伸練習答案

以下只是簡短的答案和提示，並非完整解答：只列出結果和關鍵步驟，方便你核對自己的做法。請先認真做過每一道練習，然後才看答案。

1.1 質數工廠

1. 假設 $p_1, \dots, p_k$ 就是全部的質數（prime）；那麼 $p_1 \cdots p_k + 1$ 有一個不在此列之中的質因數（prime factor），矛盾（contradiction）。用費馬數（Fermat number）的證明則是構造性（constructive）的：它具體給出無窮多個兩兩互質（pairwise coprime）的數 $F_0, F_1, \dots$，每一個都提供一個新的質數。

2. 當 $n \geq 2$ 時，2^n 是 4 的倍數，所以 $2^{2^n} = 16^{2^{n-2}}$，而 16 的任何次冪的末位數字都是 6。因此 F_n 的末位數字是 7。

3. (a) 用歸納法（induction）：$a_{n+1} - 1 = a_n(a_n - 1) = a_n \cdot a_1 \cdots a_{n-1}$。(b) a_m 與 a_n（$m < n$）的公因數必整除 $a_1 \cdots a_{n-1} = a_n - 1$ 和 a_n，所以只能是 1；每一項都有一個不與其他項共有的質因數。(c) $\frac{1}{a_n} = \frac{1}{a_n - 1} - \frac{1}{a_{n+1} - 1}$；求和時各項裂項相消（telescope），得 $1 - \frac{1}{a_{n+1} - 1}$。

4. (a) 由 $p \mid F_n$ 得 $2^{2^n} \equiv -1$；兩邊平方得 $2^{2^{n+1}} \equiv 1$。(b) 凡滿足 $2^e \equiv 1$ 的 e 都能被 d 整除，所以 $d \mid 2^{n+1}$；又因為 $2^{2^n} \equiv -1 \not\equiv 1$（因 p 是奇數），所以 $d \nmid 2^n$。故 $d = 2^{n+1}$。(c) $2^{p-1} \equiv 1$，所以 $2^{n+1} \mid p - 1$。(d) $641 = 64 \times 10 + 1$。小於 641 而形如 $64k + 1$ 的質數有 193、257、449 和 577：只需試四個候選數，歐拉（Euler）正是這樣找到 641 的。

5. 第 n 個質數不會超過 $F_0, \dots, F_{n-1}$ 之中任何一個數的最大質因數，所以 $p_n \leq F_{n-1} = 2^{2^{n-1}} + 1$。當 $n = 5$ 時，這個上界是 65 537，而 $p_5 = 11$：結論正確，但極為粗疏。

1.2 質數分佈得有多稀疏？

1. $P_{1,10} = 210$，$P_{1,20} = 9\,699\,690$，$P_{1,30} = 6\,469\,693\,230$。與 $4^{30} \approx 1.15 \times 10^{18}$ 和 $e^{30} \approx 1.07 \times 10^{13}$ 比較：這個乘積遠小於 4^n，增長速度大致與 e^n 相若（事實上 $\ln P_{1,n} \sim n$，這是質數定理（Prime Number Theorem）的一種形式）。

2. $\binom{2m}{k+1} / \binom{2m}{k} = \frac{2m-k}{k+1}$，此值大於 1 當且僅當 $k < m$，所以各係數先遞增至中間一項，

然後遞減。$2m+1$ 項之和為 4^m，其中最大的一項至少等於它們的平均值。

3. 區間 $(\sqrt{n}, n]$ 內的質數每個都大於 $\sqrt{n}$，而它們的乘積小於 4^n，所以 $(\sqrt{n})^{\pi(n)-\pi(\sqrt{n})} < 4^n$；兩邊取對數（logarithm）。由於 $2\ln 4 \approx 2.77$ 且 $\pi(\sqrt{n}) \le \sqrt{n}$，即得所求的上界。它具有正確的形式 $n/\ln n$，只是常數為 2.8，而非真正的 1。

4. 取質數 $2, 3, 5, 7, 13, 23, 43, 83$，其中每一個都小於前一個的兩倍。對任意 $n \le 50$，此列中第一個大於 n 的質數小於前一個質數的兩倍，而前一個質數不超過 n，所以它介乎 n 與 $2n$ 之間。

1.3 最小反例

1. 設 k 是使 $k\sqrt{3}$ 為整數的最小正整數，則 $k' = (\sqrt{3}-1)k$ 是一個更小的正整數，而且 $k'\sqrt{3} = 3k - k\sqrt{3}$ 是整數。這裏 $\sqrt{3}-1$ 所起的作用，與 $\sqrt{2}-1$ 相同。

2. 由於 $0 < \sqrt{N} - m < 1$，$k' = (\sqrt{N}-m)k$ 是一個小於 k 的正整數，而 $k'\sqrt{N} = kN - mk\sqrt{N}$ 是整數：矛盾。

3. 當 $N = 4$ 時，$m = 2 = \sqrt{4}$，所以 $\sqrt{N} - m = 0$，而 $k' = 0$ 並非正整數。「$0 < \sqrt{N} - m$」這一步不成立。

4. 在圖中，邊長為 a 的正方形內有兩個邊長為 b 的正方形互相重疊，它們把中央一個邊長為 $2b-a$ 的正方形覆蓋了兩次，並留下兩個邊長為 $a-b$ 的正方形未被覆蓋。用代數表示，就是 $(2b-a)^2 - 2(a-b)^2 = 2b^2 - a^2 = 0$。

1.4 e 是無理數，e^2 也是

1. 若 $\frac{1}{e} = \frac{p}{q}$，把 $\frac{1}{e} = \sum_{k\ge 0} \frac{(-1)^k}{k!}$ 乘以 $q!$。$k \le q$ 的各項都是整數，而尾部 $\sum_{k>q}(-1)^k q!/k!$ 是各項遞減的交錯級數（alternating series），所以其絕對值嚴格介乎 0 與 $\frac{1}{q+1}$ 之間。因此 $q!/e$ 不是整數：矛盾。

2. 若 $\sin 1 = \frac{p}{q}$，取 $2n+1 \ge q$，兩邊乘以 $(2n+1)!$。左邊是整數，級數中直至 $\frac{1}{(2n+1)!}$ 的各項也是整數；尾部是交錯級數，首項的大小為 $\frac{1}{(2n+2)(2n+3)} < 1$，所以它不是整數。

3. 這個方法需要一個各項均為有理數（rational）的級數，而且其尾部縮小的速度要比分母增長的速度快，就像 $\frac{1}{k!}$ 那樣；對於 π，目前並不知道有這樣的級數。Niven 的證明用到 $\int_0^\pi \frac{x^n(\pi-x)^n}{n!}\sin x\,\mathrm{d}x$：假如 π 是有理數，當 n 足夠大時，這個積分就會是一個小於 1 的正整數。

4. $x^2 - (e+\pi)x + e\pi$ 的根是 e 和 π。假如 $e+\pi$ 和 $e\pi$ 都是代數數(algebraic number)，π 就是一個代數係數多項式的根，因而也是代數數，這與林德曼（Lindemann）定理矛盾。

1.5 無窮遞降法

1. $3 \mid a$；設 $a = 3a_1$，得 $9a_1^3 + b^3 = 3c^3$，所以 $3 \mid b$，於是 $3a_1^3 + 9b_1^3 = c^3$，所以 $3 \mid c$，而 (a_1, b_1, c_1) 是一個更小的解：這就是遞降（descent）。至於 $p^4 + 2q^4 = 5r^4$：四次方數模（modulo）5 只能是 0 或 1（它們的末位數字是 0、1、5 或 6），所以 $p^4 + 2q^4 \equiv 0$ 迫使 $5 \mid p$ 且 $5 \mid q$；於是 $5^4 \mid 5r^4$，所以 $5 \mid r$，同樣可以遞降。
2. 平方數模 4 只能是 0 或 1，所以 $x^2 + y^2$ 只能是 0、1 或 2，絕不會是 3。
3. 若有有理點 $(\frac{a}{c}, \frac{b}{c})$，便會得到 $a^2 + b^2 = 3c^2$，其中 $c \neq 0$，而第 (iii) 部分已排除這個可能。
4. 固定 x、y 時，方程是二次方程 $z^2 - 3xyz + (x^2 + y^2) = 0$，其兩根之和為 $3xy$。由 $(1,1,1)$ 出發：得 $(1,1,2)$；把 $(1,2)$ 所對的那個 1 換掉，得 $3 \cdot 1 \cdot 2 - 1 = 5$，即 $(1,2,5)$；然後是 $(1,5,13)$ 和 $(2,5,29)$。

1.6 兩個平方數之和

1. 平方數模 4 只能是 0 或 1，所以兩個平方數之和模 4 絕不會是 3。
2. $65 = 1^2 + 8^2 = 4^2 + 7^2$ 及 $85 = 2^2 + 9^2 = 6^2 + 7^2$。利用 $65 = (2+\mathrm{i})(2-\mathrm{i})(3+2\mathrm{i})(3-2\mathrm{i})$：$(2+\mathrm{i})(3+2\mathrm{i}) = 4 + 7\mathrm{i}$ 及 $(2+\mathrm{i})(3-2\mathrm{i}) = 8 - \mathrm{i}$。同理，對於 85，$(2+\mathrm{i})(4 \pm \mathrm{i})$ 給出 $7 + 6\mathrm{i}$ 和 $9 - 2\mathrm{i}$。
3. $1105 = 4^2 + 33^2 = 9^2 + 32^2 = 12^2 + 31^2 = 23^2 + 24^2$。5、13、17 各自都可分解為一個高斯整數（Gaussian integer）與其共軛（conjugate）之積；從每一對中各選一個因數，共得 8 個乘積，它們兩兩互為共軛（給出相同的平方和），所以剩下 4 個。
4. 取模的平方，得 $9 = (a^2 + b^2)(c^2 + d^2)$。由於 3 不是兩個平方數之和，其中一個因數必為 1。

1.7 因數分解失效之時

1. $693 = 9 \times 77 = 21 \times 33$，四個因數全是 S-質數。方法：取四個形如 $4k+3$ 的質數 p, q, r, s；那麼 $(pq)(rs) = (pr)(qs)$ 就給出兩種分解為 S-質數的方式。
2. E-質數就是形如 $2 \pmod 4$ 的數（它們不可能是兩個偶數之積）。$60 = 2 \times 30 = 6 \times 10$。
3. 展開 $(ac - 5bd)^2 + 5(ad + bc)^2$ 即得 $N(zw) = N(z)N(w)$。$a^2 + 5b^2$ 絕不會等於 2 或 3。由於 $N(2) = 4$、$N(3) = 9$ 及 $N(1 \pm \sqrt{-5}) = 6$，若要分解為非單位（non-unit）之積，就需要一個範數（norm）為 2 或 3 的因數。
4. 設 n 是具有兩種不同分解 $p_1 \cdots p_r = q_1 \cdots q_s$ 的最小的數。根據歐幾里得（Euclid）

引理，p_1 整除某個 q_j，所以 $p_1 = q_j$；約去後得到一個更小而又有兩種不同分解的數，矛盾。

1.8 點算有理數

1. $\frac{3}{5}$：逐級追溯其父節點，$(3,5) \leftarrow (3,2) \leftarrow (1,2) \leftarrow (1,1)$，對應的位置為 $1 \to 2 \to 5 \to 10$，所以 $f(10) = \frac{3}{5}$。$\frac{5}{3}$：$(5,3) \leftarrow (2,3) \leftarrow (2,1) \leftarrow (1,1)$，位置為 $1 \to 3 \to 6 \to 13$，所以 $f(13) = \frac{5}{3}$。

2. Newman 的法則 $x_{n+1} = \frac{1}{2\lfloor x_n \rfloor - x_n + 1}$ 給出 1、$\frac{1}{2}$、2、$\frac{1}{3}$、$\frac{3}{2}$、$\frac{2}{3}$、3、$\frac{1}{4}$、$\frac{4}{3}$、$\frac{3}{5}$，與 $f(1), \ldots, f(10)$ 吻合。

3. 第 n 位數字一律取 5，除非第 n 個數的第 n 位數字是 5，這時則取 4。新的小數與列表中每一個數都不相同，而且由於避開了 0 和 9，它只有一種小數展開式，所以確實是一個不同的數。

4. 每個正整數都可唯一地寫成 2 的冪乘以一個奇數，即 $2^{m-1}(2n-1)$，所以 $(m,n) \mapsto 2^{m-1}(2n-1)$ 是一個雙射（bijection）。

1.9 三次曲線的切線

1. 把 $y = t(x+1)$ 代入 $x^2 + y^2 = 1$，得 $(1+t^2)x^2 + 2t^2 x + t^2 - 1 = 0$，其根為 -1 和 $\frac{1-t^2}{1+t^2}$。取 $t = \frac{q}{p}$，該點為 $\left(\frac{p^2-q^2}{p^2+q^2}, \frac{2pq}{p^2+q^2}\right)$，由此得到三元數組 $(p^2-q^2, 2pq, p^2+q^2)$。

2. 該直線為 $x + y = 3$，而 $x^3 + (3-x)^3 = 9$ 可化簡為 $9x^2 - 27x + 18 = 0$：三次項互相抵消，只剩下 $x = 1, 2$。這條直線與漸近線（asymptote）$y = -x$ 平行，第三個交點在無窮遠處。

3. 切線(tangent)的斜率(gradient)為 $\frac{x^2}{y^2} = \frac{289}{400}$，它與曲線再相交於 $\left(-\frac{188479}{90391}, -\frac{36520}{90391}\right)$。去分母後，得 $188479^3 = 36520^3 + 90391^3 + 180782^3$。

4. 該直線與 $x^3 + y^3 = 1729$ 再相交於 $\left(-\frac{37}{3}, \frac{46}{3}\right)$，所以 $46^3 = 37^3 + 27 \times 1729 = 37^3 + 3^3 + 36^3$（這裏用到 $27 = 3^3$ 和 $27 \times 1728 = 36^3$）。

章

2

數列、級數與收斂

無窮是數學變得有趣的地方，也是最容易出錯的地方。

本章內容

2.1	「最終」是甚麼意思	STEP III 2015
2.2	作為極限的黃金比例	STEP II 2013
2.3	一個發散級數與它的收斂近親	STEP III 2014
2.4	$\pi^2/6$	STEP III 2018
2.5	韋達乘積	STEP III 2022
2.6	算術幾何平均	STEP II 2024
2.7	六位還是六百位	STEP II 2023
2.8	平方再減二	STEP III 2021
2.9	逐步逼近 e 的分數	STEP II 2021
A	練習答案	第 169 頁

馴服無窮

> 無窮！從來沒有別的問題如此深刻地觸動人類的心靈。
>
> — 希爾伯特（David Hilbert）

兩千五百年前，埃利亞的芝諾（Zeno of Elea）論證說運動是不可能的。要走過一個房間，你必須先走一半路程，然後走餘下路程的一半，再走那餘下的一半，如此類推，永無止境：這是無窮多項任務，肯定無法在有限時間內完成。人人都知道芝諾錯了，因為人們確實能走過房間，但在很長的一段時間裏，沒有人能確切說出他爲甚麼錯。答案是：無窮和 $\frac{1}{2}+\frac{1}{4}+\frac{1}{8}+\cdots$ 有一個有限的值，就是 1。可是，無窮多個數加起來等於某個數，這句話究竟是甚麼意思？

這個問題佔據數學的核心達兩千年之久。十七世紀時，牛頓（Newton）和萊布尼茲（Leibniz）在它之上建立了微積分（calculus），卻始終沒有真正回答它；他們的後繼者既得出了許多精彩的結果，也製造了一些荒謬絕倫的謬論。歐拉（Euler）是有史以來最偉大的計算家，他毫不猶豫地寫下 $1-1+1-1+\cdots=\frac{1}{2}$。另一些人則就同一個無窮級數（infinite series）「證明」出互相矛盾的結論。直到十九世紀，經過柯西（Cauchy）、魏爾斯特拉斯（Weierstrass）等人的努力，數學家才有了足夠精確的定義，可以分辨哪些結果是真、哪些是假。這些定義就是本章的主題。

為甚麼重要

你一生中用來計算的數，幾乎每一個都是某樣東西的極限（limit）。你的計算機並不知道 $\sqrt{2}$、π 或 e 是甚麼：它只知道一些能夠愈來愈接近這些數的計算步驟，並在足夠接近時停下來。明白一套步驟逼近得有多快，正是一毫秒內完成的演算法與永遠算不完的演算法之間的分別。用來預測天氣、設計飛機和為金融產品定價的電腦模擬，背後也是同樣的想法。而在最純粹的一面，對極限的研究引出了數學中一些最優美的公式：在本章中，你將會證明

$$1+\frac{1}{4}+\frac{1}{9}+\frac{1}{16}+\cdots=\frac{\pi^2}{6},$$

這個結果令歐拉一舉成名，至今看來仍然近乎神奇。

你在大學會遇到甚麼

幾乎在每一所大學，數學分析（Analysis）都是你最先修讀的課程之一，而對許多學生來說，它帶來的衝擊也最大。在中學數學裏，你學的是計算極限；在數學分析裏，你學的

是從精確的定義出發，證明關於極限的命題。你會定義數列（sequence）收斂（converge）是甚麼意思，用的正是 Section 2.1 中出現的「對於每一個 $\varepsilon > 0$」和「存在 N」這些說法。你會證明每一個有界遞增數列都收斂，並發現這有賴於實數的一項性質——*完備性*（completeness），而有理數並不具備這項性質。你會學習判別級數是否收斂的各種方法，然後運用同樣的想法，為連續性、微分和積分賦予嚴格的意義。

往後的課程會在這個基礎上向多個方向發展。*數值分析*（Numerical analysis）探討近似值收斂得有多快，以及怎樣令它們收斂得更快（Section 2.7、Section 2.6）。*動力系統*（Dynamical systems）研究反覆套用同一條規則時會發生甚麼事，有時會趨向一個極限，有時則產生混沌（chaos）（Section 2.2、Section 2.8）。*複分析*（Complex analysis）解釋了為甚麼 π 會在那麼多無窮和與無窮乘積中出現（Section 2.4、Section 2.5）。而*解析數論*（analytic number theory）則利用極限來回答關於整數的問題，例如 Section 2.9。

貫穿本章的問題

> **三大問題**
>
> 1. **一個無窮的過程有答案，是甚麼意思？**我們怎樣才能精確地說一個數列「趨近」某個極限，或者一個無窮和「加起來等於」某個數？
> 2. **怎樣在不知道極限是甚麼的情況下，證明某樣東西收斂？**而一旦知道極限存在，又怎樣把它求出來？
> 3. **有多快？**兩個過程可以有相同的極限，逼近的速度卻有天壤之別。為甚麼會這樣？我們又怎樣量度這種速度？

第一節（Section 2.1）談的是語言：數學家怎樣精確地表達「最終」這個概念。接着的五節談的是著名的極限：以連分數（continued fraction）表示的黃金比（golden ratio）、一個發散（divergent）級數和它那個收斂的近親、$\pi^2/6$、韋達（Viète）關於 $2/\pi$ 的乘積，以及高斯（Gauss）的算術幾何平均（arithmetic–geometric mean）。最後三節談的是速度與結構：線性收斂與二次收斂的對比、一個要麼爆炸式增長要麼混沌地遊走的迭代（iteration），以及一族逐步逼近 e 的積分。

處理收斂的工具

有少數幾種工具，你會一再遇到，每次的場合都不一樣。

§	主題	工具
2.1	「最終較小」	量詞（quantifiers）：找出 m、否定命題
2.2	黃金比	單調有界子數列（monotone bounded subsequences）
2.3	調和級數（harmonic series）與 Kempner 級數	分段比較
2.4	$\pi^2/6$	多項式的根，然後用夾逼法（squeeze）
2.5	韋達乘積	裂項相消（telescoping），然後求一個簡單的極限
2.6	算術幾何平均（AGM）	兩個數列互相夾逼；一個不變量（invariant）
2.7	逼近 $\sqrt{2}$	誤差的遞推關係（recurrences）：線性與二次
2.8	$x \mapsto x^2 - 2$	一個令映射變得簡單的代換（substitution）
2.9	逼近 e 的分數	很小但不為零的積分

怎樣閱讀本章

與第 1 章一樣，每一節都以一道 STEP 原題開始。請先自己嘗試，然後才往下讀。本章的題目往往會告訴你該用哪一種工具（「考慮……」、「你可以使用以下定理……」）；在解題的過程中，請留意同樣的幾種工具出現得有多頻密，因為到了大學，你便要自己作出選擇。

有好幾道題目要求你假設極限存在，或者假設無窮和與無窮乘積收斂。遇到這種情況時，該節會解釋要證明這一點還需要些甚麼。意識到「這顯然行得通」與「我能證明這行得通」之間的差距，是學會像分析學家那樣思考的第一步。

2.1 「最終」是甚麼意思

題目

STEP III 2015，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/15-s3-q2`

引子

大學第一門分析（analysis）課程裏，幾乎每一個命題都是同一個模樣：*存在*某個東西，使得*對所有*某些東西，某個條件都成立。「數列（sequence）趨向 L」的意思是：對每一個容許誤差 $\varepsilon > 0$，都*存在*一個位置 N，使得*對所有* $n \geq N$，都有 $|a_n - L| < \varepsilon$。能夠純熟運用這種句式的學生，會覺得分析引人入勝；不熟悉的學生，則會覺得摸不着頭腦。

這道題目設計精妙，正好是這種說話方式的一個平易入門。它定義了一個新概念「最終小於」，然後就這個概念問了四條是非題。四條之中有一條是假的，而它為假的原因，正正是大多數大學一年級學生初次嘗試*否定*（negate）含有量詞（quantifier）的命題時會掉進的陷阱。

動筆之前

慢慢細讀定義。$(s_n) \leq (t_n)$ 並不是說對每個 n 都有 $s_n \leq t_n$，而是說*由某一項開始*，這個不等式一直成立。最初的幾項完全無關痛癢：即使你把其中一個數列開首的一百萬項改掉，$(s_n) \leq (t_n)$ 是否成立也不受影響。

> **想一想** 定義中有「存在正整數 m」這句話。要證明 $(s_n) \leq (t_n)$，你究竟要拿出甚麼來？要證明它是*假*的，你又要說明甚麼？

要證明它，你必須*拿出*一個合適的 m，然後說明*每一個* $n \geq m$ 都符合要求。要否證它，你必須說明*沒有任何* m 行得通：無論別人提出哪一個 m，總有某個 $n \geq m$ 使得 $s_n > t_n$。請把後面這個說法記在心裏，第 (ii) 部分會用到。

逐步解題

階段 1　線性函數對二次函數

第 (i) 部分問的是 $1000n \le n^2$ 是否最終成立。由於 $n > 0$，這等同於 $1000 \le n$。所以取 $m = 1000$ 便可：對每個 $n \ge 1000$，都有 $n^2 = n \cdot n \ge 1000n$。此命題為**真**。 ■

在最初的 999 項，$1000n$ 遠大於 n^2，但這並不要緊。「最終」二字不理會任何有限多項。

階段 2　「最終小於」跟普通的「小於」相似嗎？

對普通的數而言，如果 $a \le b$ 不成立，那麼 $b < a$，因此當然有 $b \le a$。第 (ii) 部分問的是，我們這個新關係是否也有同樣的性質。

> **想一想**　「$(s_n) \le (t_n)$ 不成立」究竟是甚麼意思？把它完整地寫出來，而且句首不要用「不」字。

「存在 m，使得對所有 $n \ge m$，都有 $s_n \le t_n$」這個命題為假，恰好就是

$$\text{對每個 } m\text{，都有某個 } n \ge m \text{ 使得 } s_n > t_n\text{。}$$

換句話說，$s_n > t_n$ 出現無限多次。這比「由某一項開始一直有 $s_n > t_n$」弱得多。兩個數列大可以不斷互換位置。

> **停一停，試一試**　試找出兩個正數數列，它們永遠不斷互換位置，以致任何一個都不是最終小於另一個。

取 $s_n = 2 + (-1)^n$，即 $s_n = 1, 3, 1, 3, \ldots$，並對所有 n 取 $t_n = 2$。對每個 m，都有一個偶數 $n \ge m$ 使得 $s_n = 3 > 2 = t_n$，所以 $(s_n) \le (t_n)$ 不成立。但同樣地，對每個 m，都有一個奇數 $n \ge m$ 使得 $t_n = 2 > 1 = s_n$，所以 $(t_n) \le (s_n)$ 也不成立。此命題為**假**。 ■

階段 3　串連起來

> **想一想**　在第 (iii) 部分，你有兩個 m 的值，每個假設各給一個，兩者可能不同。要得出結論，應該用哪一個？

假設對所有 $n \ge m_1$ 都有 $s_n \le t_n$，並且對所有 $n \ge m_2$ 都有 $t_n \le u_n$。若 $n \ge \max(m_1, m_2)$，兩個不等式同時成立，所以 $s_n \le t_n \le u_n$。因此取 $m = \max(m_1, m_2)$ 便可，此命題為**真**。 ■

在兩個門檻之中取較大者，是你在分析裏會用上數百次的一着。舉例說，要證明兩個收斂（convergent）數列之和也收斂，用的就是這一着。

階段 4 二次函數對指數函數

先試幾個數值：

n	1	2	3	4	5	6	7	8
n^2	1	4	9	16	25	36	49	64
2^n	2	4	8	16	32	64	128	256

可見 $n^2 \le 2^n$ 在 $n = 1, 2$ 時成立，在 $n = 3$ 時不成立，然後由 $n = 4$ 起似乎一直成立（圖 2.1.1）。很自然會猜 $m = 4$。

? **想一想** 要用歸納法（induction）證明對所有 $n \ge 4$ 都有 $n^2 \le 2^n$，你需要由 $n^2 \le 2^n$ 推到 $(n+1)^2 \le 2^{n+1} = 2 \cdot 2^n$。$(n+1)^2$ 與 n^2 之間要有怎樣的不等式，這一步才行得通？

只要知道 $(n+1)^2 \le 2n^2$ 就足夠，因為這樣便有 $(n+1)^2 \le 2n^2 \le 2 \cdot 2^n$。而 $2n^2 - (n+1)^2 = n^2 - 2n - 1 = (n-1)^2 - 2$，當 $n \ge 3$ 時為正數。於是：

- 基礎步驟：$4^2 = 16 = 2^4$。
- 歸納步驟：若 $n \ge 4$ 且 $n^2 \le 2^n$，則 $(n+1)^2 \le 2n^2 \le 2^{n+1}$。

因此對所有 $n \ge 4$ 都有 $n^2 \le 2^n$，此命題為**真**，可取 $m = 4$。 ■

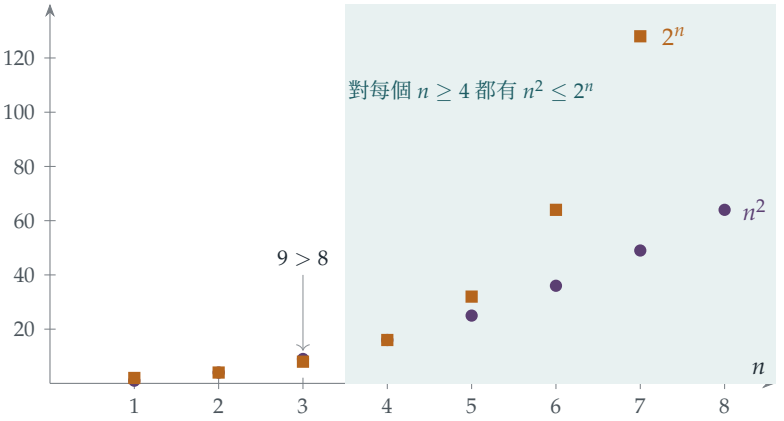


圖 2.1.1 n^2（紫色圓點）對 2^n（琥珀色方塊）。兩個數列在最初幾項交叉了兩次，但由 $n = 4$ 起，2^n 永遠領先。這正是 $(n^2) \le (2^n)$ 的意思，其中 $m = 4$。

回顧

- **證明「存在」，就是要拿出來。**在第 (i)、(iii) 和 (iv) 部分，證明的內容就是指明 m 是甚麼，然後加以驗證。

- **否定會把量詞對調。**「並非（存在 m，對所有 $n \geq m$，P）」變成「對所有 m，存在 $n \geq m$，非 P」。第 (ii) 部分考的就是能否把這一點做對。
- **新的關係不一定跟舊的關係性質相同。**「最終小於」跟 $\leq$ 一樣具有傳遞性（transitive），但跟 $\leq$ 不同的是，它並不能比較每一對數列。

背後的數學

分析的語言

在牛頓（Newton）和萊布尼茲（Leibniz）之後約 150 年間，數學家運用極限（limit）和無窮小（infinitesimal）出神入化，卻說不清楚它們確切是甚麼意思。貝克萊主教（Bishop Berkeley）在 1734 年嘲諷無窮小是「已逝之量的鬼魂」，他說的不無道理。補救要等到十九世紀：柯西（Augustin-Louis Cauchy）以及其後的魏爾斯特拉斯（Karl Weierstrass），以由「對每個」和「存在」構成的命題，取代了「變得無限接近」這類含糊的說法。現代的定義

$$a_n \to L \iff \text{對每個} \varepsilon > 0\text{，都存在} N\text{，使得} |a_n - L| < \varepsilon \text{ 適用於所有} n \geq N$$

跟本題的定義模樣完全相同，只是前面多了一個量詞。這是你在大學分析課程裏最先遇到的東西。

增長的階

第 (iv) 部分是一道著名階梯（增長的階，orders of growth）的其中一小級。對任何固定的 $k > 0$ 和 $c > 1$，

$$(\ln n) \leq (n^k), \qquad (n^k) \leq (c^n), \qquad (c^n) \leq (n!), \qquad (n!) \leq (n^n),$$

而事實上，每個數列不單最終小於下一個，更是最終小於下一個的*任何常數倍*。哈代（G. H. Hardy）為這類比較寫了一整本書：*Orders of Infinity*（1910）。今天它已是計算機科學的日常語言，「這個算法需要 $O(n^2)$ 步」的意思，就是步數最終不超過 n^2 的某個常數倍。n^2 與 2^n 之別，就是一個會完成的程式，與一個到太陽燃盡也完成不了的程式之別。

並非全序的序

第 (ii) 部分說明「最終小於」不是*全序*（total order）：有些數列對根本無法比較。這類具有傳遞性但不能比較所有東西的關係，稱為*預序*（preorder）或*偏序*（partial order），而且隨處可見：整數的整除性、集合的包含關係，以及「是……的子羣（subgroup）」都

是例子。辨別一個關係具有哪些性質、不具有哪些性質，是數學家訓練的一部分。

思維技巧

處理量詞

要證明「存在 m，使得對所有 $n \geq m$...」，先指明 m，再檢查每一個 n。要否證它，則說明對每個 m，都有某個 $n \geq m$ 不符合要求。

當一個命題含有多個量詞時，先把它完整寫出來，才嘗試證明或否證；否定時由左至右逐個處理：每個「對所有」變成「存在」，每個「存在」變成「對所有」，而末尾的條件則取其否定。

延伸練習

1 **樣本試卷熱身題。**試做 STEP III Specimen paper，第 2 題（`step.maths.org/questions/spec-s3-q2`），此題出自 1987 年首屆 STEP 舉行之前發出的試卷。（最後一部分的提示：利用二項展開式（binomial expansion）的下一項。）

2 **兩邊都成立。**證明若 $(s_n) \leq (t_n)$ 且 $(t_n) \leq (s_n)$，則對所有足夠大的 n 都有 $s_n = t_n$。

3 **競逐無窮。**試找出兩個遞增的正數數列，兩者皆趨向無窮大，而任何一個都不是最終小於另一個。

4 **拾級而上。**證明 $(2^n) \leq (n!)$，並給出一個明確的 m。然後證明 $(n^{10}) \leq (2^n)$，並求出可行的最小 m。（第二部分用試算表會有幫助；答案大得出人意表。）

延伸閱讀

- **Lara Alcock, *How to Study for a Mathematics Degree***（牛津大學出版社 Oxford University Press）。第 3 章講的正是這種邏輯語言，以及大學新生為何覺得它困難。
- **Lara Alcock, *How to Think About Analysis***（牛津大學出版社 Oxford University Press）。一本親切易讀的指南，講解大學一年級分析的各個定義。
- **Daniel Velleman, *How to Prove It***（劍橋大學出版社 Cambridge University Press）。有系統地介紹邏輯、量詞與證明。

2.2 作為極限的黃金比例

題目

STEP II 2013，第 6 題

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

`step.maths.org/questions/13-s2-q6`

引子

寫下這個式子

$$1+\cfrac{1}{1+\cfrac{1}{1+\cfrac{1}{1+\cdots}}}$$

然後問它等於甚麼。一個快捷的答案是：設它為 x；第一條分數線下面的部分又是 x，所以 $x = 1 + 1/x$，由此得 $x^2 = x + 1$，於是 $x = \frac{1+\sqrt{5}}{2}$，即黃金比（golden ratio）φ。

但這個論證成立嗎？它假設了這個無窮式子*真的有意義*，即逐層構造下去的過程最終會穩定於一個數。把同樣的手法用在 $1-1+1-1+\cdots$ 上（「設它為 S；那麼 $S = 1 - S$，所以 $S = \frac{1}{2}$」），得出的卻是荒謬的結果。本題示範怎樣把這件事做得嚴謹，所用的是 STEP 考生在事先不知道極限（limit）的情況下證明收斂（convergence）的唯一工具：*單調數列*（*monotone sequences*）。

動筆之前

先計算幾項。取 $u_1 = 1$：

$$1,\quad 2,\quad \frac{3}{2},\quad \frac{5}{3},\quad \frac{8}{5},\quad \frac{13}{8},\quad \frac{21}{13},\quad \frac{34}{21},\quad \ldots$$

想一想 你認得這些分子和分母嗎？再看看小數值：$1, 2, 1.5, 1.667, 1.6, 1.625, 1.615, \ldots$。這個數列是遞增的嗎？還是遞減的？

它們是相鄰斐波那契數（Fibonacci numbers）之比。這個數列既不遞增也不遞減：它呈鋸齒狀擺動，時而過高，時而過低。因此，題目給出的定理是關於遞增數列的，不能用於整個數列。不過，奇數項 $1, 1.5, 1.6, 1.615, \ldots$ 看來在上升，而偶數項 $2, 1.667, 1.625, \ldots$ 看來在下降。這正是本題的佈局。

逐步解題

階段 1　比較相隔兩項的項

第 (i) 部分比較相隔兩步的 u_{n+2} 與 u_n。最自然的第一步，是利用 $(*)$ 把它們各自用前一項表示。

> **想一想**　運用 $(*)$ 兩次，把 $u_{n+2} - u_n$ 用 $u_{n+1} - u_{n-1}$ 表示，再把 $u_{n+1} - u_{n-1}$ 用 $u_n - u_{n-2}$ 表示。

由 $(*)$，

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

同理 $u_{n+1} - u_{n-1} = -\dfrac{u_n - u_{n-2}}{u_n u_{n-2}}$。把兩式合起來，兩個負號互相抵消：

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

最後，由 $(*)$ 可知 $u_{n+1}u_n = u_n + 1$ 及 $u_{n-1}u_{n-2} = u_{n-2} + 1$，於是得到所求的公式。（我們需要 $n \geq 3$，u_{n-2} 才存在。）∎

階段 2　一個陷阱

> **停一停，試一試**　證明對所有 n 均有 $1 \leq u_n \leq 2$。

對 $u_1 = 1$ 這是成立的。若 $1 \leq u_n \leq 2$，則 $\frac{1}{2} \leq \frac{1}{u_n} \leq 1$，所以 $\frac{3}{2} \leq u_{n+1} \leq 2$，這當然介乎 1 與 2 之間。由歸納法（induction），對所有 n 均有 $1 \leq u_n \leq 2$。∎

特別地，每個 u_n 都是正數，所以第 (i) 部分中的分母 $(1 + u_n)(1 + u_{n-2})$ 是正數。

階段 3　兩個單調子數列

> **想一想**　第 (i) 部分的公式說明，$u_{n+2} - u_n$ 是 $u_n - u_{n-2}$ 的正數倍。由此你能看出 $u_3 - u_1$、$u_5 - u_3$、$u_7 - u_5$、...的正負號有甚麼關係嗎？

它們的正負號全部相同！由於 $u_3 - u_1 = \frac{3}{2} - 1 > 0$，由第 (i) 部分和歸納法可得 $u_5 >$

u_3、$u_7 > u_5$，如此類推：奇數項是遞增的。它們有上界 2，所以根據定理，它們趨向一個極限；設它為 L。

同樣地，$u_4 - u_2 = \frac{5}{3} - 2 < 0$，所以偶數項是遞減的。它們有下界 1。定理是就遞增數列而言的，但 $-u_2, -u_4, -u_6, \ldots$ 是遞增的，且有上界 -1，所以它收斂，因而偶數項也收斂；設它們的極限為 M。

階段 4 求出極限

想一想 要求出 L，你需要一條把 u_{n+2} 與 u_n 直接聯繫起來的方程。你能從 $(*)$ 得出這樣的方程嗎？

運用 $(*)$ 兩次，

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

令 $n \to \infty$ 而只取奇數值，則 u_n 和 u_{n+2} 都趨向 L，所以

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

方程的根是 $\frac{1 \pm \sqrt{5}}{2}$，而由於 $L \geq 1$，必有 $L = \frac{1+\sqrt{5}}{2} = \varphi$。對偶數值作完全相同的論證，可得 $M = \varphi$。

想一想 奇數項趨向 φ，偶數項也趨向 φ。為甚麼這就表示整個數列趨向 φ？

直觀地說，每一項不是奇數項就是偶數項，而兩類項都在逼近 φ。嚴格地說：給定任意容許誤差 $\varepsilon > 0$，奇數項由某個位置 N_1 起與 φ 的距離都小於 ε，偶數項則由某個位置 N_2 起如此。所以由 $\max(N_1, N_2)$ 起，*所有*項與 φ 的距離都小於 ε，而這正是收斂的意思。（這就是 Section 2.1 中「取較大的那個門檻」的做法。）

階段 5 從別處出發

取 $u_1 = 3$ 時，數列是

$$3, \quad \frac{4}{3}, \quad \frac{7}{4}, \quad \frac{11}{7}, \quad \frac{18}{11}, \quad \frac{29}{18}, \quad \cdots$$

現在 $u_1 = 3$ 在 $[1, 2]$ 之外，所以第 (ii) 部分的歸納法無法起步。但 $u_2 = \frac{4}{3}$ 在區間之內，歸納法由此開始便行得通：對所有 $n \geq 2$ 均有 $1 \leq u_n \leq 2$。第 (i) 部分只用到由 u_{n-2} 起的各項，而且只需要它們是正數，而它們確實全是正數。這一次 $u_3 - u_1 = \frac{7}{4} - 3 < 0$ 而 $u_4 - u_2 = \frac{11}{7} - \frac{4}{3} > 0$，所以奇數項遞減，偶數項遞增：角色互換了，但每個子數列（subsequence）仍然是單調且有界（bounded）的。兩個極限滿足同一條方程，而且都是正數，所以都是 φ。結論**沒有**改變。 ■

回顧

- **數列不單調時，找出單調的部分。**把數列分成奇數項和偶數項，鋸齒狀的數列就變成了兩個規矩的數列。
- **先證明收斂，再求極限。**方程 $L = (2L+1)/(L+1)$ 只有在我們知道 L 存在之後才站得住腳。引子中那個快捷的論證所略過的，正是這一步。
- **檢查該取哪個根。**極限方程有兩個解；第 (ii) 部分的上下界告訴我們哪一個才對。

背後的數學

蛛網圖與不動點

每個由 $u_{n+1} = f(u_n)$ 定義的數列，都可以畫成一幅蛛網圖（*cobweb diagram*）（圖 2.2.1）。畫出 $y = f(x)$ 和 $y = x$。由 x 軸上的 u_1 出發，垂直走到曲線（高度為 $u_2 = f(u_1)$），再水平走到直線 $y = x$（這樣便移到 $x = u_2$），然後再垂直走到曲線，如此類推。這裏 $f(x) = 1 + 1/x$ 是遞減的，所以路徑會向內盤旋，交替地過高和過低，趨向曲線與直線相交的那一點：不動點（*fixed point*）φ，該處 $f(\varphi) = \varphi$。

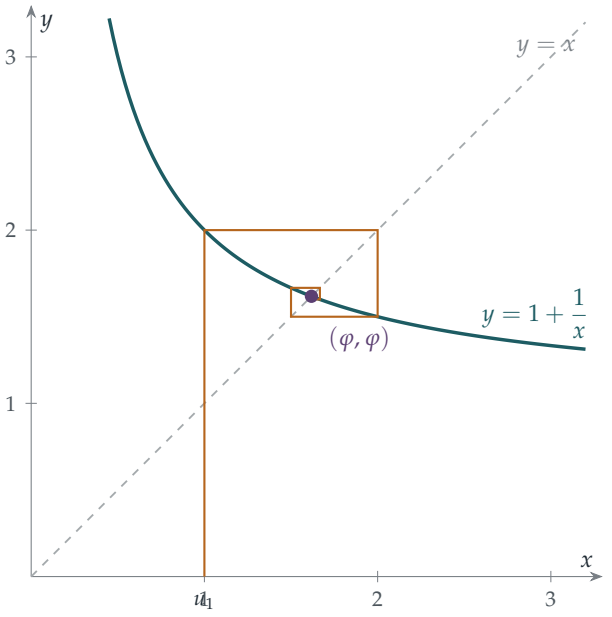


圖 2.2.1 $u_{n+1} = 1 + 1/u_n$ 由 $u_1 = 1$ 出發的蛛網圖。路徑向上走到曲線，再橫向走到直線 $y = x$，如此不斷重複。由於曲線向下傾斜，路徑會向內盤旋，趨向不動點 $\varphi \approx 1.618$，並且交替落在兩側：這就是奇數項與偶數項的鋸齒狀擺動。

有多快，為甚麼

第 (i) 部分中的因子 $(1+u_n)(1+u_{n-2})$ 從何而來？在不動點附近，$u_{n+1} - \varphi = f(u_n) - f(\varphi) \approx f'(\varphi)(u_n - \varphi)$，而 $f'(\varphi) = -1/\varphi^2 \approx -0.38$。所以每走一步，誤差大約乘以 -0.38：正負號反轉（鋸齒狀擺動），大小縮減（收斂）。走兩步則乘以 $1/\varphi^4 \approx 0.146$，這正是 $1/\big((1+u_n)(1+u_{n-2})\big)$ 所趨向的值。到了大學，這會發展成*壓縮映射定理*（*contraction mapping theorem*）：若在不動點附近有 $|f'| \le k < 1$，則迭代會收斂至該不動點，而誤差至少以幾何級數的速度縮小。它是數學中最有用的定理之一，也是用來證明微分方程有解的方法（第 8 章）。在 Section 2.8 中，我們會遇到一個不動點並不吸引的映射，結果截然不同。

最無理的數

各項 $u_n = F_{n+1}/F_n$ 是同等大小的分數之中對 φ 最好的逼近，但所謂「最好」其實並不太好：φ 比任何其他數都更難用分數逼近。原因可從它的連分數（continued fraction）$[1;1,1,1,\ldots]$ 看出：連分數中出現大的項，就表示有非常好的逼近（試想 $\pi = [3;7,15,1,292,\ldots]$ 和極佳的 $\frac{355}{113}$），而 φ 的每一項自始至終都是最小的。赫維茲（Adolf Hurwitz）於 1891 年證明，對每個無理數（irrational）x，都有無窮多個分數滿足

$|x - p/q| < 1/(\sqrt{5}\,q^2)$，而且 $\sqrt{5}$ 不能再改進，原因正正在於 φ。這種「最大無理性」，就是黃金角出現在向日葵種子螺旋圖案中的原因。

思維技巧

單調有界必收斂

要在不知道極限的情況下證明數列收斂，可找出一個單調有界的數列，或者把數列分成若干單調有界的部分。

只有在知道極限存在之後，才可以在遞推關係（recurrence）中「令 $n \to \infty$」來求出極限。在大學裏，這個定理是實數完備性（*completeness*）的推論，完備性正是 $\mathbb{R}$ 有別於 $\mathbb{Q}$ 的性質，你會看到它由公理出發的證明。

延伸練習

1 **藏在其中的斐波那契數。**證明 $u_n = F_{n+1}/F_n$，其中 $F_1 = F_2 = 1$ 及 $F_{n+1} = F_n + F_{n-1}$。

2 **誤差縮小。**證明 $u_{n+1} - \varphi = -\dfrac{u_n - \varphi}{\varphi\, u_n}$，並由此推出對所有 n 均有 $|u_n - \varphi| \le \varphi^{-(n-1)}$。（利用 $\varphi^2 = \varphi + 1$ 及 $u_n \ge 1$。）

3 **令數列中斷。**證明若 $u_1 = \frac{1-\sqrt{5}}{2}$，則數列是常數數列。求出所有使得某個 $u_n = 0$、以致數列無法繼續下去的初始值 u_1。（提示：由 $u_n = 0$ 倒推回去。）

4 **$\sqrt{2}$ 的近親。**設 $v_1 = 1$ 及 $v_{n+1} = 1 + \dfrac{1}{1 + v_n}$。仿照本題的方法，證明 $v_n \to \sqrt{2}$，並把 $\sqrt{2}$ 寫成連分數。

延伸閱讀

- **Mario Livio, *The Golden Ratio***（Broadway）。講述 φ 的歷史，把真正的數學與眾多迷思區分開來。
- **C. D. Olds, *Continued Fractions***（Mathematical Association of America）。一本簡短而非常易讀的入門書，對象是能力較強的中學生。
- **J. C. Burkill, *A First Course in Mathematical Analysis***（Cambridge University Press）。嚴謹地處理單調數列與完備性，並附有大量習題。

2.3 一個發散級數與它的收斂近親

題目

STEP III 2014，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/14-s3-q8`

引子

調和級數（harmonic series）$1+\frac{1}{2}+\frac{1}{3}+\frac{1}{4}+\cdots$ 是最著名的例子：它的各項縮小至零，但總和卻無止境地增長。它增長得慢得離譜：要超過 10，需要多於 12,000 項；要超過 100，則需要大約 1.5×10^{43} 項。但它終究會超過任何一個數。

現在把分母含有數字 2 的項全部丟掉：$\frac{1}{2}$、$\frac{1}{12}$、$\frac{1}{20}$ 至 $\frac{1}{29}$、$\frac{1}{32}$ 等等統統刪去。你也許會以為，對一個本來就是無窮大的和來說，刪去十分之一左右的項不會有多大分別。事實卻是，餘下各項加起來小於 80。這個出人意表的結果由 A. J. Kempner 於 1914 年證明，而本題只用一個簡單而漂亮的想法，就把你一步步帶到這個結論：把和式切成一塊一塊。

動筆之前

想一想 第一個不等式談的是由 $r=k^n$ 至 $r=k^{n+1}-1$ 的那一塊。取 $k=2$ 和 $n=3$，這一塊包含哪些項？共有多少項？

它們是 $f(8), f(9), \ldots, f(15)$，共八項。一般而言，這一塊由 k^n 排到 $k^{n+1}-1$，所以共有 $k^{n+1}-k^n = k^n(k-1)$ 項。不等式所說的是：這一塊的和介乎（項數）×（各項可能取的最小值）與（項數）×（最大的一項）之間。

逐步解題

階段 1　估計一塊的上下界

由於 f 是遞減（decreasing）的，這一塊內的每一項都滿足

$$f(k^{n+1}) < f(k^{n+1}-1) \le f(r) \le f(k^n) \qquad (k^n \le r \le k^{n+1}-1).$$

把這 $k^n(k-1)$ 項加起來，便得到該不等式。 ■

階段 2　調和級數

> **想一想**　取 $k=2$ 和 $f(r)=1/r$，由 2^n 至 $2^{n+1}-1$ 的那一塊，兩個界分別是甚麼？

下界（lower bound）是 $2^n \cdot \frac{1}{2^{n+1}} = \frac{1}{2}$，上界（upper bound）是 $2^n \cdot \frac{1}{2^n} = 1$。所以每一塊，不論排得多後，其貢獻都介乎 $\frac{1}{2}$ 與 1 之間（圖 2.3.1）。$n = 0, 1, \dots, N$ 的各塊恰好拼合成由 $r=1$ 至 $r = 2^{N+1}-1$ 的和，所以把 $N+1$ 塊加起來，便得

$$\frac{N+1}{2} \le \sum_{r=1}^{2^{N+1}-1} \frac{1}{r} \le N+1.$$

當 $N \to \infty$ 時，左邊無止境地增長，所以部分和（partial sum）無界，級數發散（diverge）。 ■

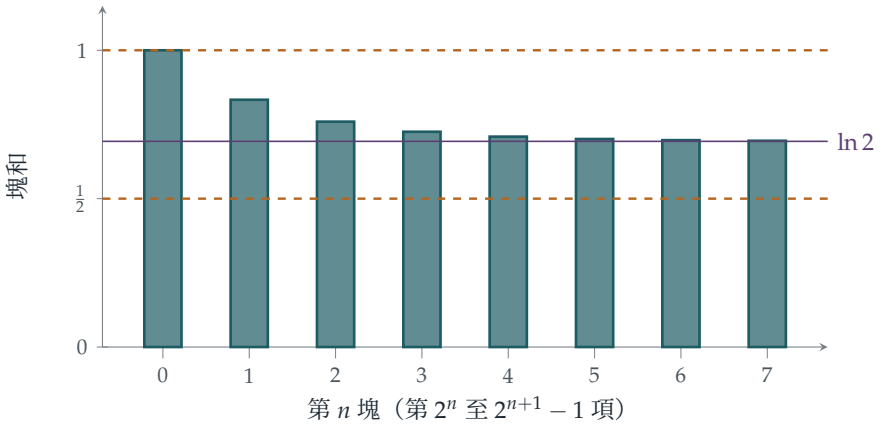


圖 2.3.1　把調和級數切成分別有 $1, 2, 4, 8, \dots$ 項的各塊。每一塊的和都介乎 $\frac{1}{2}$ 與 1 之間（琥珀色虛線），而事實上各塊的和會逐漸趨近 $\ln 2 \approx 0.693$（紫紅色線）。無窮多塊，每塊至少值 $\frac{1}{2}$，總和不可能有限。

階段 3　立方倒數之和

現在取 $f(r) = 1/r^3$，仍然用 $k=2$。第 n 塊的上界是

$$2^n \cdot \frac{1}{(2^n)^3} = \frac{1}{4^n}.$$

$n = 0, 1, 2, \ldots$ 的各塊把每個正整數恰好覆蓋一次，所以

$$\sum_{r=1}^{\infty} \frac{1}{r^3} \le \sum_{n=0}^{\infty} \frac{1}{4^n} = \frac{1}{1 - \frac{1}{4}} = \frac{4}{3} = 1\tfrac{1}{3}. \qquad \blacksquare$$

（嚴格來說，這只證明了每個部分和都不超過 $\frac{4}{3}$；一個遞增而有上界的數列必定收斂（converge），所以這個無窮和存在，而且不超過 $\frac{4}{3}$。它的真正數值是 1.2020 . . .）

階段 4 數一數不含 2 的數

想一想 如果容許在前面補零（例如把 7 寫成 007），由 0 至 999 的整數之中，有多少個可以只用數字 $0, 1, 3, 4, 5, 6, 7, 8, 9$ 寫出來？

三個數位每個都有 9 種選擇，所以這樣的數字串共有 9^3 個。它們與由 0 至 999 之中不含 2 的數一一對應，再剔除 0（它不是正數），$S(1000)$ 中便剩下 $9^3 - 1 = 728$ 個數。 $\blacksquare$

階段 5 Kempner 級數

分塊的想法現在同樣適用，只是改為按位數來分塊，這等同於取 $k = 10$。

停一停，試一試 有多少個 d 位數不含 2？而一個 d 位數最小可以是多少？

首位數字可以是 $1, 3, 4, \ldots, 9$ 中任何一個（八種選擇：不能是 0，也不能是 2），其餘 $d - 1$ 個數位各有九種選擇。所以這樣的數共有 $8 \times 9^{d-1}$ 個，而每一個都至少是 10^{d-1}。因此，$S(n)$ 中的 d 位數對 $\sigma(n)$ 的貢獻至多是

$$8 \times 9^{d-1} \times \frac{1}{10^{d-1}} = 8\left(\frac{9}{10}\right)^{d-1}$$

。對所有位數 $d = 1, 2, 3, \ldots$ 求和，

$$\sigma(n) \le \sum_{d=1}^{\infty} 8\left(\frac{9}{10}\right)^{d-1} = \frac{8}{1 - \frac{9}{10}} = 80.$$

至於嚴格不等式，只須留意大部分的界都遠遠未達到相等（例如 $\frac{1}{3} < 1$，而我們卻用 1 作為每個一位數項的上界），所以對所有 n 都有 $\sigma(n) < 80$。 $\blacksquare$

回顧

- **把各項分組。**兩個和逐項處理都不容易。把它們分成大小合適的塊之後，每個和都可以與一個等比級數（geometric series）比較。
- **怎樣分塊視乎問題而定。**2 的冪適合 $1/r$ 和 $1/r^3$；10 的冪則適合一道關於十進制數字的題目。

- **有界而遞增。**兩個收斂證明都建基於與 Section 2.2 相同的原理：正項的部分和是遞增的，所以只需要找到一個界便足夠。

背後的數學

凝聚判別法

題目中的第一個不等式正是柯西凝聚判別法（Cauchy's condensation test）的核心：若 f 為正且遞減，則

$$\sum_{r=1}^{\infty} f(r) \text{ 收斂} \iff \sum_{n=0}^{\infty} 2^n f(2^n) \text{ 收斂。}$$

它把困難的級數變成容易的級數。應用於 $f(r) = 1/r^p$，便得到 $\sum 2^n \cdot 2^{-np} = \sum (2^{1-p})^n$，這是一個等比級數，恰好在 $p > 1$ 時收斂。應用於 $1/(r \ln r)$，可知連這個級數也發散；應用於 $1/(r(\ln r)^2)$，則可知這個級數收斂。這個論證可追溯至奧雷姆（Nicole Oresme），他在大約 1350 年正是用圖 2.3.1 中的分塊方法，證明了調和級數發散。

調和級數發散得有多慢

圖 2.3.1 顯示每一塊的貢獻大約是 $\ln 2$，所以首 N 項之和應該大約是 $\log_2 N \times \ln 2 = \ln N$。事實上

$$1 + \frac{1}{2} + \cdots + \frac{1}{N} = \ln N + \gamma + (\text{某個趨向 } 0 \text{ 的量}),$$

其中 $\gamma = 0.5772\ldots$ 是歐拉－馬斯刻若尼常數（Euler–Mascheroni constant）。至今沒有人知道 γ 是否有理數（rational）。你會在大學一年級的分析課中用積分證明這條公式，而調和數（harmonic numbers）會在 Section 9.11 和 Section 11.12 再次出現。

大多數的數都含有 2

Kempner 級數為甚麼收斂？因為在大數之中，「不含 2 的數」非常罕有：在小於 10^d 的數之中，不含 2 的所佔比例至多是 $(0.9)^d$，而這個比例趨向 0。在 100 位數之中，只有不足 0.003% 能避開數字 2。所以我們刪去的並不是十分之一的項，而是幾乎所有的項。我們求得的界 80 相當寬鬆：1979 年，Robert Baillie 用一種加速方法把這個和計算到小數點後許多位，其值約為 19.257。

冪的倒數之和

第 (ii) 部分的和是 $\zeta(3)$，即我們在 Section 1.2 遇過的黎曼（Riemann）zeta 函數的一個值。在 Section 2.4 我們會求出 $\zeta(2) = \pi^2/6$ 的準確值，而歐拉（Euler）亦找到了

$\zeta(4), \zeta(6), \ldots$ 的公式。但至今沒有人找到 $\zeta(3)$ 的公式。1978 年，Roger Apéry 證明了 $\zeta(3)$ 是無理數（irrational），震驚了整個數學界；他所用的積分與 Section 2.9 中的頗為相似。至於 $\zeta(5)$ 是否無理數，至今仍然未知。

思維技巧

分塊比較

要判斷一個正項級數是否收斂，可以把各項分成若干塊，使每塊內的項大致不變，然後把每一塊與一個簡單的級數比較。

比較判別法（comparison test）是大學一年級分析課的家常便飯。技巧在於選擇與甚麼比較：在這裏，就是選大小為 2^n 或 10^d 的塊，塊內各項的變化不超過一個固定的倍數。

延伸練習

1 **感受一下有多慢。**利用 $\sum_{r \le N} 1/r \approx \ln N + 0.5772$，估計調和級數需要多少項才超過 10，以及需要多少項才超過 100。

2 **判別法的實際運用。**用第 (i) 部分的方法，取 $f(r) = 1/r^2$，證明 $\sum 1/r^2 \le 2$。然後把第一項分開處理，從而改進這個界。

3 **在收斂與發散之間。**利用 $k = 2$ 的分塊不等式，證明 $\sum_{r \ge 2} \dfrac{1}{r \ln r}$ 發散，而 $\sum_{r \ge 2} \dfrac{1}{r (\ln r)^2}$ 收斂。

4 **排除更多數字。**較難證明：十進制表示中既不含 2 亦不含 7 的所有正整數，其倒數（reciprocal）之和小於 35。（你看得出 35 是怎樣來的嗎？）

延伸閱讀

- **Julian Havil, *Gamma: Exploring Euler's Constant***（Princeton University Press）。講述調和級數、γ 以及更多相關課題，為程度較高的中學生而寫。
- **Mary Hart, *Guide to Analysis***（Palgrave）。一本簡明的大學一年級教科書，其中論級數與比較判別法的一章寫得十分清晰。
- **Robert Baillie, "Sums of reciprocals of integers missing a given digit"**, *American Mathe-*

matical Monthly 86 (1979)。一篇短論文（論缺少某個指定數字的整數的倒數之和），讀完本節後你便有能力讀懂。

2.4 $\pi^2/6$

題目

STEP III 2018，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/18-s3-q7`

引子

1650 年，Pietro Mengoli 提出了一個簡單的問題：

$$1+\frac{1}{4}+\frac{1}{9}+\frac{1}{16}+\frac{1}{25}+\cdots ?$$

的值是甚麼？各項縮小得夠快，足以令這個和收斂（converge）（Section 2.3 已證明它不超過 2），但沒有人能求出它的值。伯努利（Bernoulli）兄弟試過，卻無功而還；雅各布·伯努利（Jacob Bernoulli）更公開懇求他人相助。這個問題後來以伯努利兄弟的家鄉命名，稱為巴塞爾問題（*Basel problem*）。

1734 年，年輕的歐拉（Leonhard Euler）解決了這個問題，答案令他一舉成名：這個和等於 $\pi^2/6$。π 明明是一個關於圓的數，究竟為甚麼會出現在平方倒數之和裏？本題給出一個完整而嚴格的證明，所用的工具全部不超出高級程度（A-level）的範圍：棣美弗定理（De Moivre's theorem）、多項式（polynomial）的根之和，以及夾逼（squeeze）。

動筆之前

在處理細節之前，值得先看看整體的計劃。

- 第 (i) 部分構造一個多項式，它的根就是 $\cot^2\left(\frac{m\pi}{2n+1}\right)$ 這些數。
- 第 (ii) 部分利用係數（coefficient），在不知道各個根的情況下把這些根加起來。
- 第 (iii) 部分利用「當 θ 很小時 $\cot\theta \approx 1/\theta$」這個事實，把 $\cot^2$ 值之和變成 $1/m^2$ 值之和，然後夾逼。

> **想一想**　若多項式 $a_n x^n + a_{n-1}x^{n-1} + \cdots + a_0$ 的根為 $r_1, \ldots, r_n$，那麼 $r_1 + \cdots + r_n$ 以係數表示是甚麼？

答案是 $-a_{n-1}/a_n$。這個事實正是第 (ii) 部分的動力所在。

逐步解題

階段 1　棣美弗定理

> **想一想**　把 $\cot\theta + \mathrm{i}$ 寫成單一分數。你留意到分子有甚麼特別？

$\cot\theta + \mathrm{i} = \dfrac{\cos\theta + \mathrm{i}\sin\theta}{\sin\theta}$，同樣地 $\cot\theta - \mathrm{i} = \dfrac{\cos\theta - \mathrm{i}\sin\theta}{\sin\theta}$。根據棣美弗定理，

$$(\cot\theta \pm \mathrm{i})^{2n+1} = \frac{\cos(2n+1)\theta \pm \mathrm{i}\sin(2n+1)\theta}{\sin^{2n+1}\theta}.$$

兩式相減，餘弦項互相抵消，正弦項則變成兩倍，再除以 $2\mathrm{i}$ 便得

$$\frac{\sin(2n+1)\theta}{\sin^{2n+1}\theta}. \qquad \blacksquare$$

階段 2　多項式

現在改用二項式定理（binomial theorem）展開左邊。設 $c = \cot\theta$ 及 $N = 2n+1$。

> **想一想**　在 $(c+\mathrm{i})^N - (c-\mathrm{i})^N$ 中，兩個二項展開式有哪些項互相抵消，又有哪些項變成兩倍？

含 i^j 的項在第一個展開式中帶 $+$ 號，在第二個展開式中則帶 $(-1)^j$。因此 j 為偶數的項互相抵消，j 為奇數的項變成兩倍：

$$(c+\mathrm{i})^N - (c-\mathrm{i})^N = 2\sum_{j\ \text{為奇數}} \binom{N}{j} c^{N-j}\,\mathrm{i}^j.$$

取 $j = 2k+1$，則 $\mathrm{i}^j = \mathrm{i}\,(-1)^k$ 及 $c^{N-j} = c^{2n-2k} = (c^2)^{n-k}$。除以 $2\mathrm{i}$，得

$$\frac{\sin(2n+1)\theta}{\sin^{2n+1}\theta} = \binom{2n+1}{1}x^n - \binom{2n+1}{3}x^{n-1} + \cdots + (-1)^n,$$

其中 $x = \cot^2\theta$。

> **停一停，試一試** 左邊在哪些 θ 的值時等於零？利用這一點找出多項式的 n 個不同的根。

當 $\sin(2n+1)\theta = 0$ 而 $\sin\theta \neq 0$ 時，左邊等於零，例如當 $\theta = \frac{m\pi}{2n+1}$，其中 $m = 1, 2, \ldots, n$。這些角全部嚴格介乎 0 與 $\frac{\pi}{2}$ 之間，而 cot 在這個範圍內為正且嚴格遞減，所以 $\cot^2\left(\frac{m\pi}{2n+1}\right)$ 這 n 個值各不相同。n 次多項式最多有 n 個根，所以這些就是它的全部根。 ■

階段 3 把根加起來

根之和等於 $-$(以下一項的係數：x^{n-1})/(以下一項的係數：x^n)：

$$\sum_{m=1}^{n} \cot^2\left(\frac{m\pi}{2n+1}\right) = \frac{\binom{2n+1}{3}}{\binom{2n+1}{1}} = \frac{(2n+1)(2n)(2n-1)}{6\,(2n+1)} = \frac{n(2n-1)}{3}. \qquad ■$$

以 $n = 1$ 驗算：$\cot^2(\pi/3) = \frac{1}{3}$，而 $\frac{1\times 1}{3} = \frac{1}{3}$。

階段 4 夾逼

由 $0 < \sin\theta < \theta < \tan\theta$ 出發，取倒數（不等號方向因而倒轉）再平方：

$$\cot^2\theta < \frac{1}{\theta^2} < \frac{1}{\sin^2\theta} = 1 + \cot^2\theta.$$

> **想一想** 代入 $\theta = \frac{m\pi}{2n+1}$，並對 $m = 1, \ldots, n$ 求和。外側的兩個和變成甚麼？中間的和與 $\sum 1/m^2$ 有甚麼關係？

中間的和是 $\displaystyle\sum_{m=1}^{n} \frac{(2n+1)^2}{m^2\pi^2}$。根據第 (ii) 部分，左邊的和是 $\frac{n(2n-1)}{3}$，而右邊的和是 $n + \frac{n(2n-1)}{3} = \frac{2n(n+1)}{3}$。各項同乘以 $\frac{\pi^2}{(2n+1)^2}$：

$$\frac{\pi^2}{3}\cdot\frac{n(2n-1)}{(2n+1)^2} < \sum_{m=1}^{n}\frac{1}{m^2} < \frac{\pi^2}{3}\cdot\frac{2n(n+1)}{(2n+1)^2}.$$

當 $n \to \infty$ 時，$\frac{n(2n-1)}{(2n+1)^2}$ 和 $\frac{2n(n+1)}{(2n+1)^2}$ 都趨向 $\frac{2}{4} = \frac{1}{2}$（把分子和分母同除以 n^2）。所以上下兩個界都趨向 $\frac{\pi^2}{6}$，而部分和（partial sum）被夾在兩者之間，亦必定趨向 $\frac{\pi^2}{6}$（圖 2.4.1）。 ■

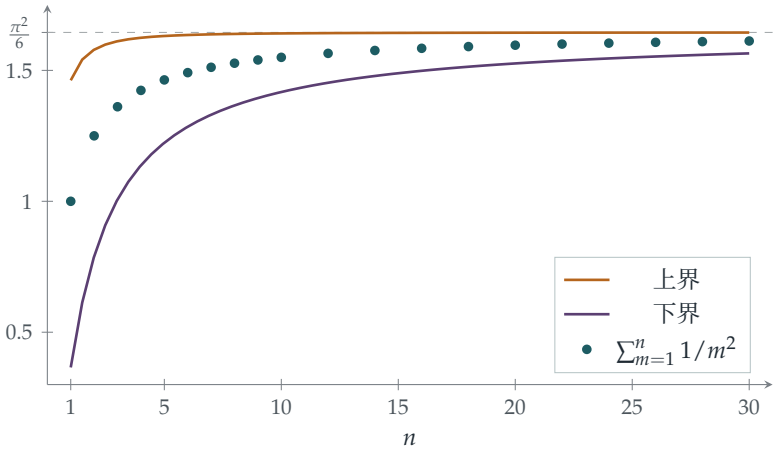


圖 2.4.1　第 (iii) 部分的夾逼。$\sum 1/m^2$ 的部分和（藍綠色點）被夾在下界（紫紅色）與上界（琥珀色）之間，而兩個界都趨向 $\pi^2/6 \approx 1.645$。

回顧

- **從係數取得資訊。**我們始終沒有把 $\cot^2(m\pi/(2n+1))$ 這些根化成任何更簡單的形式，卻準確地求出了它們的和。
- **先近似，再夾逼。**$\cot\theta$ 只是近似等於 $1/\theta$，但誤差在兩個方向上都受到精確的不等式控制。
- **由有限到無限。**除了最後一步，每一步處理的都是有限和。無窮和只在我們於不等式中令 $n \to \infty$ 時才出現，這正是處理無限的穩妥方法。

背後的數學

歐拉的大膽論證

以現代標準來看，歐拉本人在 1734 年的論證並不嚴格，卻令人目眩神迷。一個多項式若以 $\pm a, \pm b, \pm c, \ldots$ 為根，且在 0 處取值 1，便可分解為 $(1 - x^2/a^2)(1 - x^2/b^2)\cdots$。函數 $\frac{\sin x}{x}$ 在 0 處取值 1，而它的根是 $\pm\pi, \pm 2\pi, \pm 3\pi, \ldots$，於是歐拉把它當作一個「無限多項式」，寫下

$$\frac{\sin x}{x} = \left(1 - \frac{x^2}{\pi^2}\right)\left(1 - \frac{x^2}{4\pi^2}\right)\left(1 - \frac{x^2}{9\pi^2}\right)\cdots.$$

右邊 x^2 的係數是 $-\frac{1}{\pi^2}\left(1 + \frac{1}{4} + \frac{1}{9} + \cdots\right)$，而由麥克勞林級數 (Maclaurin series) $\frac{\sin x}{x} = 1 - \frac{x^2}{6} + \cdots$ 可知它是 $-\frac{1}{6}$。兩相比較便得 $\pi^2/6$。我們這個 STEP 證明正是同一想法的

有限而嚴格的版本：一個多項式，它的根與 π 的倍數息息相關，再加上它的第二個係數。

平方倒數以外

歐拉其後還求出 $\sum 1/m^4 = \pi^4/90$、$\sum 1/m^6 = \pi^6/945$，而一般而言 $\sum 1/m^{2k}$ 是 π^{2k} 的有理數倍，當中涉及伯努利數（*Bernoulli numbers*）。奇數次冪則是另一回事：正如我們在 Section 2.3 所見，沒有人知道 $\sum 1/m^3$ 的公式。這些和就是黎曼 ζ 函數（Riemann zeta function）的值 $\zeta(2), \zeta(4), \ldots$，而這個函數在複數自變量上的表現，支配着質數的分佈（Section 1.2）。

日後在哪裏再遇上

在大學裏，你會至少再遇到 $\zeta(2) = \pi^2/6$ 的三個證明：一個來自傅立葉級數（*Fourier series*）（把函數寫成正弦和餘弦之和），一個來自二重積分（double integral），還有一個來自複分析（complex analysis）。這個數也在概率中出現：隨機選出的兩個整數沒有公因數的機會是 $6/\pi^2 \approx 0.61$。（質數 p 同時整除兩者的概率是 $1/p^2$，所以沒有任何質數同時整除兩者的機會是 $\prod_p (1 - 1/p^2)$，而歐拉證明了這個乘積等於 $1/\zeta(2)$。）π 在概率極限中的另一次現身，是 Section 12.23 的主題。

思維技巧

夾逼，以及根的對稱函數（symmetric functions）

要求一個極限，可把該量夾在兩個極限相同而較簡單的式子之間。

而當你需要求一些難纏的數之和（或積）時，不妨試找一個以它們為根的多項式：係數會把這些數的對稱組合白白送給你。

延伸練習

1 **1992 年的版本。**試做 STEP III 1992，第 8 題（`step.maths.org/questions/92-s3-q8`）。它利用一個與本節相似的多項式，求出正切平方與餘割平方的準確和。（提示：若你知道一個多項式的根是 r_i，那麼哪個多項式的根是 $1/r_i$？）

2 **奇數的平方。**由 $\sum 1/m^2 = \pi^2/6$ 推導出 $1 + \frac{1}{9} + \frac{1}{25} + \frac{1}{49} + \cdots = \frac{\pi^2}{8}$。

3 **四次方。** 較難 利用 $\sum r_i^2 = (\sum r_i)^2 - 2\sum_{i<j} r_i r_j$ 和多項式的係數，求 $\sum_{m=1}^{n} \cot^4\!\left(\frac{m\pi}{2n+1}\right)$，並改動夾逼的論證以證明 $\sum 1/m^4 = \pi^4/90$。

4 **有多快？** 利用第 (iii) 部分的下界，證明 $\frac{\pi^2}{6} - \sum_{m=1}^{n} \frac{1}{m^2} < \frac{\pi^2}{4n}$。按這個結果，$n$ 要多大才能保證部分和與 $\pi^2/6$ 相差不超過 0.001？試與真正的誤差比較，後者接近 $1/n$。

延伸閱讀

- **William Dunham, *Euler: The Master of Us All***（Mathematical Association of America）。第 3 章講述巴塞爾問題的故事，並完整交代當中的數學。
- **Aigner & Ziegler, *Proofs from THE BOOK***，第 9 章「Three times $\pi^2/6$」。當中包括本節的證明。
- **Paul Nahin, *Dr. Euler's Fabulous Formula***（Princeton University Press）。內容涵蓋複數、傅立葉級數，以及類似本節的求和。

2.5 韋達乘積

題目

STEP III 2022，第 4 題

繼續閱讀之前，請先看題目（原題為英文）。題目載於 OCR 提供下載的 2022 年 STEP 試卷（zip 檔案）內的 STEP III 試卷：

`www.ocr.org.uk/Images/696587-2022-past-papers.zip`

引子

這道題的最後一行，是數學中最優美的公式之一，也是同類公式中最古老的一條。它由法國律師兼數學家韋達（François Viète）於 1593 年發表，是歷史上第一次有人把 π 準確地寫成一個無窮過程的結果。在韋達之前，π 只能求近似值；在他之後，它可以用一條公式來定義，只要一直算下去，就能得出每一個數位。

這條公式看來很神祕：一層套一層的 2 的平方根，即嵌套平方根（nested square roots），究竟從何而來？這道題揭示，它其實來自反覆運用倍角公式（double-angle formula），而且同一套方法還順道得出另一條名氣較小的公式，是關於 $\ln 2$ 的。

動筆之前

想一想 sinh 的倍角公式是甚麼？（如果不肯定，可把 $\sinh 2y$ 和 $\sinh y \cosh y$ 用指數函數寫出來，再作比較。）

$\sinh 2y = 2\sinh y \cosh y$，與 $\sin 2y = 2\sin y \cos y$ 完全一樣。整道題就是靠反覆運用這一條恆等式（identity）來推動的。題目一直使用雙曲函數（hyperbolic functions），到最後一部分我們才轉用普通的三角函數。

逐步解題

階段 1　反覆減半

$n = 1$ 的情況就是取 $y = x/2$ 的倍角公式：

$$\sinh x = 2\cosh\left(\frac{x}{2}\right)\sinh\left(\frac{x}{2}\right).$$

> **想一想**　假設公式對 n 成立。最後一個因子是 $\sinh(x/2^n)$。怎樣改寫它，才能得到 $n+1$ 時的公式？

對它運用倍角公式：$\sinh(x/2^n) = 2\cosh(x/2^{n+1})\sinh(x/2^{n+1})$。代入後得

$$\sinh x = 2^{n+1}\cosh\left(\frac{x}{2}\right)\cdots\cosh\left(\frac{x}{2^n}\right)\cosh\left(\frac{x}{2^{n+1}}\right)\sinh\left(\frac{x}{2^{n+1}}\right),$$

這正是 $n+1$ 時的公式。根據歸納法（induction），公式對所有 n 成立。 ■

至於隨後的推論，先把兩邊除以 x，再把兩邊乘以 $\dfrac{x/2^n}{\sinh(x/2^n)}$。在右邊，$2^n\sinh(x/2^n)\cdot\dfrac{x/2^n}{x\sinh(x/2^n)} = 1$，只剩下各個 cosh 項的乘積。 ■

階段 2　極限

由級數（series）可知，當 $y \neq 0$ 時，

$$\frac{\sinh y}{y} = 1 + \frac{y^2}{3!} + \frac{y^4}{5!} + \cdots,$$

當 $y \to 0$ 時它趨向 1，所以它的倒數 $y/\sinh y$ 也趨向 1。（如果想嚴謹一點：尾部 $\frac{y^2}{3!} + \frac{y^4}{5!} + \cdots$ 中的每個階乘都至少是 6，所以當 $0 < |y| < 1$ 時，尾部至多是 $\frac{1}{6}(y^2 + y^4 + \cdots) = \frac{y^2}{6(1-y^2)}$，而這個量趨向 0。）

現在固定 x，在第 (i) 部分中令 $n \to \infty$。這時 $y = x/2^n \to 0$，所以左邊趨向 $\dfrac{\sinh x}{x} \times 1$。右邊是部分積（partial product），因此它也趨向 $\dfrac{\sinh x}{x}$。這正是無窮乘積（infinite product）的意思。 ■

階段 3　$\ln 2$ 的公式

取 $x = \ln 2$，則 $e^{x/2} = \sqrt{2}$，所以

$$\cosh\left(\frac{x}{2}\right) = \frac{2^{1/2} + 2^{-1/2}}{2} = \frac{2+1}{2\sqrt{2}} = \frac{3}{2\sqrt{2}}.$$

同樣地，$e^{x/4} = 2^{1/4}$，所以 $\cosh(x/4) = \dfrac{2^{1/4} + 2^{-1/4}}{2}$，把分子和分母同乘以 $2^{1/4}$，便得 $\dfrac{2^{1/2}+1}{2\times 2^{1/4}}$。 ■

想一想 $\cosh(x/2^k)$ 的一般規律是甚麼？把所有分母 $2\times 2^{1/2^k}$ 乘在一起，會有甚麼結果？

同樣的計算顯示，對每個 $k\geq 1$，

$$\cosh\left(\frac{x}{2^k}\right)=\frac{1+2^{1/2^{k-1}}}{2}\times 2^{-1/2^k}.$$

（當 $k=1$ 時，這就是 $\frac{1+2}{2}\times 2^{-1/2}=\frac{3}{2\sqrt{2}}$，與前面一致。）對 $k=1,2,3,\ldots$ 連乘，2 的各個冪合併成 $2^{-(1/2+1/4+1/8+\cdots)}=2^{-1}$，而第一個分數是 $\frac{1+2}{2}=\frac{3}{2}$。所以這個無窮乘積是

$$\frac{3}{2}\times\frac{1}{2}\times\frac{1+2^{1/2}}{2}\times\frac{1+2^{1/4}}{2}\times\frac{1+2^{1/8}}{2}\cdots.$$

另一邊，$\sinh(\ln 2)=\frac{2-\frac{1}{2}}{2}=\frac{3}{4}$，所以 $\dfrac{\sinh x}{x}=\dfrac{3}{4\ln 2}$。兩邊的因子 $\frac{3}{4}$ 互相約去，剩下

$$\frac{1}{\ln 2}=\frac{1+2^{1/2}}{2}\times\frac{1+2^{1/4}}{2}\times\frac{1+2^{1/8}}{2}\cdots.$$ ■

階段 4 韋達公式

想一想 第 (iv) 部分含有 π，所以它必定來自普通的三角函數，而不是雙曲函數。如果把 sinh 和 cosh 換成 sin 和 cos，第 (i) 和第 (ii) 部分的一切是否仍然成立？

仍然成立。歸納法只用到倍角公式，而 $\sin 2y=2\sin y\cos y$。極限只用到當 $y\to 0$ 時 $\frac{\sin y}{y}\to 1$，這可以用同樣的方法由 $\sin y=y-\frac{y^3}{3!}+\cdots$ 推出。所以當 $x\neq 0$（並且 x 不是 π 的倍數，以確保過程中沒有因子為零）時，

$$\frac{\sin x}{x}=\cos\left(\frac{x}{2}\right)\cos\left(\frac{x}{4}\right)\cos\left(\frac{x}{8}\right)\cdots.$$

停一停，試一試 選取 x 使左邊等於 $2/\pi$，然後算出最初幾個因子。半角公式（half-angle formula）$\cos\frac{\theta}{2}=\sqrt{\frac{1+\cos\theta}{2}}$（適用於 $0\leq\theta\leq\pi$）會派上用場。

取 $x=\frac{\pi}{2}$。這時 $\frac{\sin x}{x}=\frac{1}{\pi/2}=\frac{2}{\pi}$，而各因子是 $\cos\frac{\pi}{4},\cos\frac{\pi}{8},\cos\frac{\pi}{16},\ldots$。第一個是 $\cos\frac{\pi}{4}=\frac{\sqrt{2}}{2}$。如果某個因子是 $\cos\theta=\frac{c}{2}$，由半角公式可知下一個因子是

$$\cos\frac{\theta}{2}=\sqrt{\frac{1+c/2}{2}}=\frac{\sqrt{2+c}}{2}.$$

所以分子依次是 $\sqrt{2}$、$\sqrt{2+\sqrt{2}}$、$\sqrt{2+\sqrt{2+\sqrt{2}}}$、...，每一個都是把前一個加上 2 再開平方而得。這就是韋達公式。 ■

回顧

- **乘積的相消（telescoping）。**第 (i) 部分的恆等式是一個會收攏的乘積，正如某些和式可以裂項相消一樣。
- **先有有限的恆等式，再取極限。**與 Section 2.4 一樣，每一步對有限的 n 都是準確的，唯一需要的極限就是非常簡單的 $\frac{\sin y}{y} \to 1$。
- **雙曲函數與圓函數是孿生兄弟。**同一個證明對兩者都適用，因為兩者滿足相同的倍角公式和相同的小角極限。這並非巧合：$\cosh(ix) = \cos x$，你會在 Section 5.7 看到。

背後的數學

阿基米德（Archimedes）的多邊形

韋達公式有一個幾何意義，可追溯至阿基米德。內接（inscribed）於半徑為 1 的圓的正 2^n 邊形，面積是 $A_n = 2^{n-1}\sin\frac{2\pi}{2^n}$。把邊數加倍，面積就乘以

$$\frac{A_{n+1}}{A_n} = \frac{2^n \sin(\pi/2^n)}{2^{n-1}\cdot 2\sin(\pi/2^n)\cos(\pi/2^n)} = \frac{1}{\cos(\pi/2^n)}.$$

由正方形（$A_2 = 2$）開始，不斷把邊數加倍，面積便趨向圓的面積 π。所以 $\pi = 2 \div \left(\cos\frac{\pi}{4}\cos\frac{\pi}{8}\cos\frac{\pi}{16}\cdots\right)$，這正是韋達公式。乘積中的每個因子，都是圖 2.5.1 中相鄰兩個多邊形的面積之比。

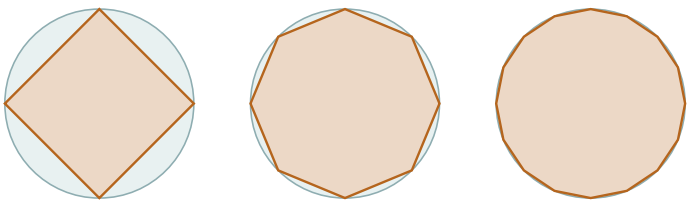


圖 2.5.1　內接於半徑為 1 的圓的正多邊形，邊數分別為 4、8 和 16。邊數每加倍一次，面積就多除以韋達乘積中的一個因子，而面積不斷增大並趨向 π。

無窮乘積

韋達乘積是數學史上第一個無窮乘積。其後陸續有其他例子：沃利斯（John Wallis）在 1655 年發現 $\frac{\pi}{2} = \frac{2}{1}\cdot\frac{2}{3}\cdot\frac{4}{3}\cdot\frac{4}{5}\cdot\frac{6}{5}\cdot\frac{6}{7}\cdots$，而歐拉（Euler）關於 $\frac{\sin x}{x}$ 的乘積已在 Section 2.4 出現過。到了大學，你會學到無窮乘積在甚麼情況下收斂（converge）。粗略地說，當 $\prod(1+a_n)$ 中的 a_n 是細小的正數時，它的表現與 $\sum a_n$ 相似，因為 $\ln(1+a_n)\approx a_n$。這裏 $1-\cos(\pi/2^n)\approx\frac{1}{2}(\pi/2^n)^2$，每次縮小 4 倍，所以這個乘積收斂得很快。

有多快？

乘了 n 個因子之後，部分積恰好等於 $\dfrac{1}{2^n\sin(\pi/2^{n+1})}$，其誤差大約是 $\frac{\pi^2}{6\cdot 4^{n+1}}$ 乘以 $\frac{2}{\pi}$。每多一個因子，答案的準確度大約提高 4 倍，即多得大約 0.6 個小數位。十個因子可使 $2/\pi$ 準確至六個小數位。這已經不錯，但遠遠不及 Section 2.6 的方法那麼快，那些方法每一步都令準確的位數加倍。

思維技巧

反覆運用恆等式

一條把 $f(x)$ 與 $f(x/2)$ 聯繫起來的恆等式，可以一次又一次地運用；取極限之後，結果往往是一個無窮和或無窮乘積。

找出 n 步之後的「餘項」（這裏是 $2^n\sinh(x/2^n)$），然後另行求出它的極限。這個模式在分析學中隨處可見，由無窮乘積以至分形（fractals）的構造都有它的蹤影。

延伸練習

1 **1990 年的版本。**試做 STEP I 1990，第 8 題（`step.maths.org/questions/90-s1-q8`）。它循一條與本節不同的途徑得出韋達乘積，所用的不外乎 sin 的倍角公式。你覺得哪一種方法較為自然？

2 **另一個角。**在 $\frac{\sin x}{x}$ 的乘積中取 $x=\frac{\pi}{3}$，把 $\frac{3\sqrt{3}}{2\pi}$ 寫成無窮乘積，並化簡最初兩個因子。

3 **要多少個因子？**用計算機找出韋達乘積需要多少個因子，才能使 π 準確至 4 個小數位。解釋為甚麼每多一個因子大約多得 0.6 個數位。

4 $\ln 3$ **的公式。**取 $x = \ln 3$ 重做第 (iii) 部分，把 $\frac{1}{\ln 3}$ 寫成無窮乘積。需要多少個因子才能使 $\frac{1}{\ln 3}$ 準確至兩個小數位？為甚麼這個乘積收斂得比韋達乘積慢？

延伸閱讀

- **Petr Beckmann, *A History of* π**（St Martin's Press）。阿基米德、韋達、沃利斯等人的故事，寫得興味盎然。
- **Jonathan Borwein & Peter Borwein, *Pi and the AGM***（Wiley）。關於計算 π 的嚴謹數學，包括 Section 2.6 的方法。讀完本章之後，第 1 章應該讀得懂。
- **Eli Maor, *Trigonometric Delights***（Princeton University Press）。書中有一章專門講述韋達和他的乘積。

2.6 算術幾何平均

題目

STEP II 2024，第 8 題

繼續閱讀之前，請先看題目（原題為英文）。題目載於 OCR 提供下載的 2024 年 STEP 試卷（zip 檔案）內的 STEP II 試卷：

`www.ocr.org.uk/Images/725753-2024-past-papers.zip`

引子

取兩個正數，例如 $\sqrt{2}$ 和 1。把它們換成兩者的算術平均（arithmetic mean）和幾何平均（geometric mean）。再做一次，然後再做一次。結果如下：

n	x_n（算術）	y_n（幾何）
0	1.4142135623730950488	1
1	1.2071067811865475244	1.1892071150027210667
2	1.1981569480946342956	1.1981235214931201226
3	1.1981402347938772091	1.1981402346773072058
4	1.1981402347355922074	1.1981402347355922074

只需四步，兩個數便有二十位小數完全相同。它們的共同極限（limit）就是算術幾何平均（arithmetic–geometric mean，簡稱 AGM）$M(\sqrt{2},1) = 1.19814\ldots$

1799 年 5 月 30 日，年僅 21 歲的高斯（Carl Friedrich Gauss）在日記中寫道，他發現 $1/M(\sqrt{2},1)$ 與一個常數有十一位小數相同，而這個常數出自一個完全不同的問題：一條稱為雙紐線（lemniscate）的曲線的弧長。他寫道：「這必定會開闢一個全新的分析領域。」事實果然如此。本題要證明的，正是高斯這項發現背後的關鍵事實：AGM 能求出一個用任何常規方法都計算不了的積分。

動筆之前

> **想一想**　為甚麼兩個不相等的正數的算術平均，總是大於它們的幾何平均？你能用一行證明嗎？

這就是算術幾何平均不等式（AM–GM inequality），而第 (i) 部分的提示正好給出這個一行證明：當 $x \neq y$ 時，$\left(\sqrt{x}-\sqrt{y}\right)^2 > 0$，展開後便是 $x+y > 2\sqrt{xy}$。第 (i) 部分的一切，都由這個不等式，以及「每個平均都介乎原來兩個數之間」這個事實推出。

逐步解題

階段 1　兩個數列互相夾緊

若 $y_n < x_n$，則 $\sqrt{x_n} \neq \sqrt{y_n}$，所以 $x_n - 2\sqrt{x_n y_n} + y_n > 0$，即 $x_{n+1} > y_{n+1}$。由於 $y_0 < x_0$，用歸納法（induction）可知對所有 n 都有 $y_n < x_n$。

> **想一想**　為甚麼 x_n 與 y_n 的算術平均小於 x_n，而幾何平均大於 y_n？

因為 $y_n < x_n$：$x_{n+1} = \frac{x_n+y_n}{2} < \frac{x_n+x_n}{2} = x_n$，而 $y_{n+1} = \sqrt{x_n y_n} > \sqrt{y_n y_n} = y_n$。所以

$$y_0 < y_1 < y_2 < \cdots < x_2 < x_1 < x_0.$$

y_n 遞增，並以 x_0 為上界（bounded above），所以根據題目所給的定理，它趨向一個極限 M。

至於兩者的差距，

$$x_{n+1} - y_{n+1} < x_{n+1} - y_n = \frac{x_n+y_n}{2} - y_n = \frac{1}{2}(x_n - y_n),$$

而由第一部分可知它是正數。所以差距每次至少減半：$0 < x_n - y_n < (x_0 - y_0)/2^n$，而後者趨向 0。

最後，$x_n = y_n + (x_n - y_n)$ 是一個趨向 M 的數列與一個趨向 0 的數列之和，所以 $x_n \to M$ 亦成立。 ■

階段 2　代換

這個代換（substitution）看來無緣無故，但它其實經過精心設計，令根號下的兩個括號都變成完全平方（或接近完全平方）。

想一想 設 $t = \frac{1}{2}\left(x - \frac{pq}{x}\right)$，計算 $t^2 + pq$。然後計算 $t^2 + \frac{1}{4}(p+q)^2$。兩者之中，有沒有哪一個可以漂亮地因式分解？

首先，

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

其次，

$$\begin{aligned} t^2 + \tfrac{1}{4}(p+q)^2 &= \frac{1}{4}\left(x^2 - 2pq + \frac{p^2q^2}{x^2} + p^2 + 2pq + q^2\right) \\ &= \frac{x^4 + (p^2+q^2)x^2 + p^2q^2}{4x^2} = \frac{(x^2+p^2)(x^2+q^2)}{4x^2}. \end{aligned}$$

另外 $\dfrac{\mathrm{d}t}{\mathrm{d}x} = \dfrac{1}{2}\left(1 + \dfrac{pq}{x^2}\right) = \dfrac{x^2+pq}{2x^2}$。

停一停，試一試 把以上結果合起來。當 x 由 0 走到 ∞ 時，t 怎樣變化？關於 t 的積分會變成甚麼？

當 x 由 0 增加至 ∞ 時，t 由 $-\infty$ 遞增至 ∞（因為它的導數是正數）。對 $x > 0$，

$$\sqrt{\left(t^2 + \tfrac{1}{4}(p+q)^2\right)\left(t^2+pq\right)} = \frac{\sqrt{(x^2+p^2)(x^2+q^2)}}{2x} \cdot \frac{x^2+pq}{2x},$$

所以

$$\frac{\mathrm{d}t}{\sqrt{\cdots}} = \frac{x^2+pq}{2x^2} \cdot \frac{4x^2}{(x^2+pq)\sqrt{(x^2+p^2)(x^2+q^2)}}\,\mathrm{d}x = \frac{2\,\mathrm{d}x}{\sqrt{(x^2+p^2)(x^2+q^2)}}.$$

因此關於 t 的積分等於 $2\,\mathrm{I}(p,q)$。

另一方面，關於 t 的積分與 I 的形式完全一樣，只是 p 和 q 分別換成了 $\frac{1}{2}(p+q) = \mathrm{a}(p,q)$ 和 $\sqrt{pq} = \mathrm{g}(p,q)$，而且積分範圍是整條實數線。被積函數（integrand）是 t 的偶函數（even function），所以關於 t 的積分等於 $2\,\mathrm{I}\big(\mathrm{a}(p,q), \mathrm{g}(p,q)\big)$。兩者比較，得

$$\mathrm{I}(p,q) = \mathrm{I}\big(\mathrm{a}(p,q), \mathrm{g}(p,q)\big). \qquad \blacksquare$$

階段 3 求積分的值

想一想 由 (x_0, y_0) 開始，反覆應用這個恆等式。經過 n 步之後，$\mathrm{I}(x_0,y_0)$ 等於甚麼？當 $n \to \infty$ 時又會怎樣？

對每個 n，$\mathrm{I}(x_0,y_0) = \mathrm{I}(x_1,y_1) = \mathrm{I}(x_2,y_2) = \cdots = \mathrm{I}(x_n,y_n)$。當 $n \to \infty$ 時，兩個變數都趨向 M，而

$$\mathrm{I}(M,M) = \int_0^\infty \frac{\mathrm{d}x}{M^2+x^2} = \left[\frac{1}{M}\arctan\frac{x}{M}\right]_0^\infty = \frac{\pi}{2M}.$$

所以

$$\mathrm{I}(x_0, y_0) = \frac{\pi}{2M}. \qquad \blacksquare$$

題目容許我們直接接受「$\mathrm{I}(x_n, y_n) \to \mathrm{I}(M, M)$」這一步，但其實可以簡潔地加以證明。把 p 或 q 增大，被積函數便會變小，而 $y_n < x_n$，所以

$$\frac{\pi}{2x_n} = \mathrm{I}(x_n, x_n) < \mathrm{I}(x_n, y_n) < \mathrm{I}(y_n, y_n) = \frac{\pi}{2y_n}.$$

上下兩個界都趨向 $\frac{\pi}{2M}$，而對每個 n，$\mathrm{I}(x_n, y_n)$ 都是同一個數 $\mathrm{I}(x_0, y_0)$，所以它必定等於 $\frac{\pi}{2M}$。這又是一次夾逼（squeeze），與 Section 2.4 中的做法一樣。

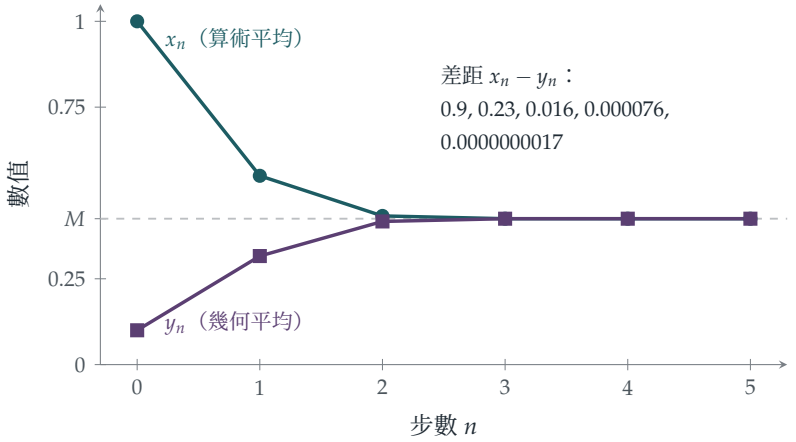


圖 2.6.1　$x_0 = 1$ 與 $y_0 = 0.1$ 的算術幾何平均。算術平均 x_n（藍綠色）遞減，幾何平均 y_n（紫紅色）遞增，兩者以驚人的速度逼近共同極限 $M = 0.425040709\ldots$：一旦兩者接近，正確位數每一步大約增加一倍。

回顧

- **兩個單調數列互相夾緊。**一個上升，一個下降，而兩者的差距縮小至零。要證明兩個數列有共同極限，這是最乾淨利落的方法。
- **不變量（invariant）。**$\mathrm{I}(x_n, y_n)$ 這個量在每一步都保持不變。在其他一切都在變動時，找出某樣固定不變的東西，是數學中最有力的想法之一。
- **巧妙的代換是設計出來的，不是猜出來的。**選用這個代換，正是為了令根號下的式子可以因式分解。

背後的數學

比題目所說的還要快

第 (i) 部分證明了差距每一步至少減半，這相當於每一步大約多得一個二進制位。引子中的表格所顯示的卻驚人得多：正確位數每一步大約*增加一倍*（$2, 4, 10, 21, 43, \ldots$）。事實上

$$x_{n+1} - y_{n+1} = \frac{1}{2}\bigl(\sqrt{x_n} - \sqrt{y_n}\bigr)^2 = \frac{(x_n - y_n)^2}{2\bigl(\sqrt{x_n} + \sqrt{y_n}\bigr)^2},$$

所以每個差距大約是一個常數乘以前一個差距的*平方*。這稱為*二次收斂*（quadratic convergence），我們會在 Section 2.7 中仔細探討。

橢圓積分

我們的積分 $\mathrm{I}(p, q)$ 其實是一個喬裝了的*完全橢圓積分*（complete elliptic integral）。這類積分在計算橢圓的弧長時出現，而且不能用你所認識的函數表示。它最著名的應用是單擺。高級程度（A-level）課程中的週期公式 $T = 2\pi\sqrt{L/g}$ 只是小幅度擺動時的近似。由角度 θ_0 釋放的單擺，其準確週期是

$$T = \frac{2\pi\sqrt{L/g}}{M\bigl(1, \cos\frac{\theta_0}{2}\bigr)},$$

這是一個可用 AGM 求值的橢圓積分。當 $\theta_0 = 90^\circ$ 時，真正的週期比簡單公式所預測的長約 18%。高斯那則日記所談的是雙紐線，一條 8 字形的曲線，它的總弧長涉及 $\int_0^1 \frac{\mathrm{d}t}{\sqrt{1-t^4}}$；這個積分其實就是喬裝了的 $\mathrm{I}(\sqrt{2}, 1)$，所以可以用 AGM 求值。

計算 π

1976 年，Richard Brent 與 Eugene Salamin 各自獨立地把高斯的想法變成一個計算 π 的算法（algorithm），每一步都令正確位數增加一倍。二十五步便可得出數千萬位。在 1980 及 1990 年代，這類算法曾用於多次刷新紀錄的 π 計算；時至今日，凡是需要高精度的地方，類似的二次收斂算法已成為標準工具。試比較 Section 2.5 中的韋達（Viète）乘積，它每一步只多得約 0.6 位。

思維技巧

> **尋找不變量**
>
> 當一個過程同時改變許多東西時，找出一個不變的量。然後在過程的終點求它的值，因爲那裏一切都很簡單。
>
> 在這裏，過程就是 AGM 的每一步，而不變量就是積分 I。在過程的終點，兩個變數相等，積分變成初等積分。不變量在數學中隨處可見，由力學中的能量守恆（第 10 章），以至證明某些謎題無解。

延伸練習

1. **二次收斂。**證明恆等式 $x_{n+1}-y_{n+1}=\dfrac{(x_n-y_n)^2}{2(\sqrt{x_n}+\sqrt{y_n})^2}$，並由此推出 $x_{n+1}-y_{n+1}\le \dfrac{(x_n-y_n)^2}{8y_0}$。

2. **親手計算。**用計算機求 $M(2,1)$，準確至六位小數。你需要多少步？

3. **縮放。**證明對 $\lambda>0$ 有 $M(\lambda x,\lambda y)=\lambda M(x,y)$，以及 $\mathrm{I}(\lambda p,\lambda q)=\lambda^{-1}\mathrm{I}(p,q)$。驗證這兩個事實與 $\mathrm{I}(x_0,y_0)=\pi/(2M)$ 互相吻合。

4. **單擺。**利用上面的公式，求由 60° 釋放的單擺的週期比小角度近似的預測值長多少，答案準確至三位有效數字。

延伸閱讀

- **David Cox, “The arithmetic–geometric mean of Gauss”**, *L'Enseignement Mathématique* 30 (1984)。講述這段歷史和有關數學的經典文章；開首幾節只需要微積分知識。
- **Jonathan Borwein & Peter Borwein, *Pi and the AGM*** (Wiley)。關於 AGM 與 π 的快速計算的權威著作。
- **Jonathan Borwein & Keith Devlin, *The Computer as Crucible*** (A. K. Peters)。講述實驗數學，包括像高斯那樣憑着留意數值上的巧合而得出的發現。

2.7 六位還是六百位

題目

STEP II 2023，第 5 題

繼續閱讀之前，請先看題目（原題為英文）。題目載於 OCR 提供下載的 2023 年 STEP 試卷（zip 檔案）內的 STEP II 試卷：

`www.ocr.org.uk/Images/707102-2023-past-papers.zip`

引子

有兩個求 $\sqrt{2}$ 近似值的簡單方法，同樣由 1 開始，同樣只用加法、乘法和除法。十步之後，第一個方法準確至小數點後約 8 位，第二個方法則準確至超過 700 位。

兩者的分別，並不在於第二個方法對付 $\sqrt{2}$ 的手法較為高明，而在於誤差的形態不同：第一個方法每走一步，誤差乘以一個固定的分數；第二個方法每走一步，誤差就被平方一次。這道題目把這一點說得精確，同時引入數值分析（numerical analysis）的一個核心概念；每一部計算機、每一次天氣預報、每一個工程模擬，背後都是這門學科。

動筆之前

先以分數形式計算兩個數列（sequence）的首幾項。

$$x_n:\quad 1,\ \frac{3}{2},\ \frac{7}{5},\ \frac{17}{12},\ \frac{41}{29},\ \frac{99}{70},\ \cdots \qquad y_n:\quad 1,\ \frac{3}{2},\ \frac{17}{12},\ \frac{577}{408},\ \frac{665857}{470832},\ \cdots$$

想一想 你留意到這兩列數有甚麼特別嗎？（看看哪些分數在兩列中都有出現。）

每一個 y_n 都在 x 的那一列中出現，只是位置愈來愈後：$y_2 = x_3$，$y_3 = x_7$，一般而言 $y_n = x_{2^n-1}$。第二個方法其實是在同一列近似值之中，以指數速度向前跳。請記住這一點，我們會在本節末解釋。

逐步解題

階段 1　第一個方法的各項不小於 1

我們可以寫成 $x_{n+1} = 1 + \dfrac{1}{x_n + 1}$。若 $x_n \geq 1$，則 $x_n + 1 > 0$，所以 $x_{n+1} > 1$。由於 $x_0 = 1$，由歸納法（induction）可知對所有 n 都有 $x_n \geq 1$。 ■

階段 2　第一個方法的誤差

> **想一想**　題目以 $x_n^2 - 2$ 而不是 $x_n - \sqrt{2}$ 來量度誤差。試把 $x_{n+1}^2 - 2$ 寫成以 x_n 表示的單一分數。分子有甚麼特別之處？

$$x_{n+1}^2 - 2 = \frac{(x_n + 2)^2 - 2(x_n + 1)^2}{(x_n + 1)^2} = \frac{x_n^2 + 4x_n + 4 - 2x_n^2 - 4x_n - 2}{(x_n + 1)^2} = \frac{2 - x_n^2}{(x_n + 1)^2}.$$

分子正好是 $-(x_n^2 - 2)$。所以 $x_{n+1}^2 - 2$ 與 $x_n^2 - 2$ 異號（近似值交替落在 $\sqrt{2}$ 的兩側），而且由 $x_n \geq 1$ 可得 $(x_n + 1)^2 \geq 4$，因此

$$|x_{n+1}^2 - 2| = \frac{|x_n^2 - 2|}{(x_n + 1)^2} \leq \frac{1}{4}|x_n^2 - 2|. \qquad \blacksquare$$

階段 3　十步之後

把上式連續應用十次，得 $|x_{10}^2 - 2| \leq 4^{-10}|x_0^2 - 2| = 4^{-10} = \frac{1}{1048576} < 10^{-6}$。至於正負號：$x_0^2 - 2 = -1 < 0$，而正負號逐項交替，所以對每個偶數 n 都有 $x_n^2 - 2 < 0$。於是 $x_{10}^2 < 2$，兩者合起來便得 $2 - 10^{-6} \leq x_{10}^2 \leq 2$。 ■

x_n^2 的誤差每一步最多乘以 $\frac{1}{4}$。這稱為**線性收斂**（linear convergence）：每一步增加固定數目的準確位數（這裏最少是 $\log_{10} 4 \approx 0.6$；實際上較接近 0.77）。

階段 4　第二個方法的誤差

> **想一想**　試把 $y_{n+1} - \sqrt{2}$ 寫成單一分數。分子應該會令你想起完全平方。

$$y_{n+1} - \sqrt{2} = \frac{y_n^2 + 2 - 2\sqrt{2}\,y_n}{2y_n} = \frac{\left(y_n - \sqrt{2}\right)^2}{2y_n}. \qquad \blacksquare$$

若 $y_n > 0$，右方為非負數，所以 $y_{n+1} \geq \sqrt{2} > 1$。由於 $y_0 = 1$，由歸納法可知對所有 n 都有 $y_n \geq 1$。 ■（事實上，對所有 $n \geq 1$ 都有 $y_n \geq \sqrt{2}$：走完第一步之後，第二個方法永遠從上方逼近 $\sqrt{2}$。）

階段 5　把誤差平方

設 $e_n = y_n - \sqrt{2}$，則當 $n \geq 1$ 時 $e_n \geq 0$，而且因為 $y_n \geq 1$，有 $e_{n+1} = \dfrac{e_n^2}{2y_n} \leq \dfrac{e_n^2}{2}$。

> **停一停，試一試** 用歸納法證明：當 $n \geq 1$ 時 $e_n \leq 2r^{2^n}$，其中 $r = \frac{\sqrt{2}-1}{2}$。請仔細檢查基礎情況。

基礎情況：$e_1 = \dfrac{(1-\sqrt{2})^2}{2 \cdot 1} = \dfrac{(\sqrt{2}-1)^2}{2}$，而 $2r^2 = 2 \cdot \dfrac{(\sqrt{2}-1)^2}{4} = \dfrac{(\sqrt{2}-1)^2}{2}$。兩者相等，所以基礎情況成立（而且是等式）。

歸納步驟：若 $e_n \leq 2r^{2^n}$，則

$$e_{n+1} \leq \frac{e_n^2}{2} \leq \frac{4r^{2^{n+1}}}{2} = 2r^{2^{n+1}}. \qquad \blacksquare$$

階段 6 六百位小數

由於 $\sqrt{2} - 1 < \frac{1}{2}$，我們有 $r < \frac{1}{4}$，所以

$$0 \leq y_{10} - \sqrt{2} \leq 2\left(\frac{1}{4}\right)^{2^{10}} = 2 \cdot 2^{-2048} = 2^{-2047}.$$

而 $2^{10} = 1024 > 10^3$，所以 $2^{-2047} < 2^{-2040} = (2^{10})^{-204} < 10^{-612}$，這當然小於 10^{-600}。 $\blacksquare$

回顧

- **找出誤差的遞推關係。**兩部分的做法，都是把新誤差寫成舊誤差的一個精確函數，其餘一切便隨之而來。
- **線性與二次之別。**「新誤差 $\leq$ 常數 $\times$ 舊誤差」表示每一步增加固定數目的準確位數。「新誤差 $\leq$ 常數 $\times$ (舊誤差)^2」則表示每一步令準確位數加倍（圖 2.7.1）。
- **平方量決定正負號。**在第 (ii) 部分，走一步之後的誤差是一個平方數除以一個正數，所以永遠不會是負數：迭代從上方逼近。

背後的數學

牛頓法

第二個方法其實是把牛頓法（Newton's method）應用於 $f(y) = y^2 - 2$。牛頓法求 f 的根的做法，是反覆把猜測值 y 換成 y 處的切線與橫軸的交點：

$$y_{\text{新}} = y - \frac{f(y)}{f'(y)} = y - \frac{y^2 - 2}{2y} = \frac{y^2 + 2}{2y}.$$

在單根（simple root）α 附近，由泰勒展開式（Taylor expansion）可知新誤差約為 $\frac{f''(\alpha)}{2f'(\alpha)} \times (\text{舊誤差})^2$，這就是牛頓法二次收斂（quadratic convergence）的原因。這個特殊情況，即「取 y 與 $2/y$ 的平均值」，古巴比倫人早已知道；一塊約公元前 1700 年的泥板（現稱為 YBC 7289）以 60 進制給出 $\sqrt{2}$ 的值，誤差小於百萬分之一。

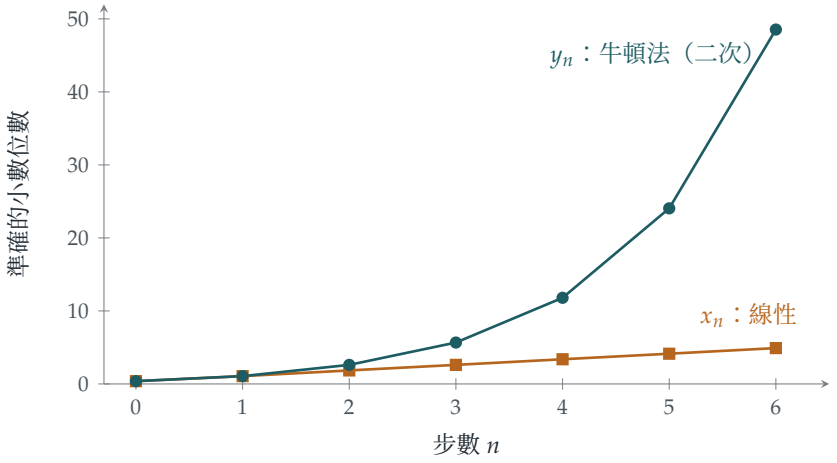


圖 2.7.1　兩個方法各走 n 步之後，$\sqrt{2}$ 的準確小數位數。線性的方法每一步增加約 0.77 位；牛頓法則每一步令準確位數大約加倍，到 $n = 10$ 時達到約 780 位。

不動點及其收斂速度

兩個方法都具有 $u_{n+1} = g(u_n)$ 的形式，而且 $g(\sqrt{2}) = \sqrt{2}$，所以 $\sqrt{2}$ 是一個不動點（fixed point），正如 Section 2.2 的蛛網圖（cobweb diagram）所示。在不動點 α 附近，誤差每一步大約乘以 $g'(\alpha)$。對第一個方法而言，$g(x) = \frac{x+2}{x+1}$，而 $g'(\sqrt{2}) = -\frac{1}{(\sqrt{2}+1)^2} = -(3 - 2\sqrt{2}) \approx -0.17$：負號解釋了正負號為何交替，0.17 則解釋了每一步為何增加 0.77 位（$\log_{10}(1/0.17) \approx 0.77$）。對牛頓法而言，$g'(\alpha) = 0$，所以線性項消失，誤差由下一

項（即平方項）決定。設計出滿足 $g'(\alpha) = 0$ 的迭代（iteration），正是快速數值方法的精髓所在。

為何出現相同的分數

分數 $x_n = p_n/q_n$ 滿足 $p_{n+1} = p_n + 2q_n$ 和 $q_{n+1} = p_n + q_n$，並且 $p_n^2 - 2q_n^2 = \pm 1$，正負交替。它們是佩爾方程（Pell's equation）$p^2 - 2q^2 = \pm 1$ 的解，也是連分數（continued fraction）$\sqrt{2} = [1; 2, 2, 2, \ldots]$ 的漸近分數（convergents）。它們來自 $1 + \sqrt{2}$ 的冪：$p_n + q_n\sqrt{2} = (1 + \sqrt{2})^{n+1}$。牛頓法的每一步把 $p + q\sqrt{2}$ 平方（不妨一試：$(p + q\sqrt{2})^2 = p^2 + 2q^2 + 2pq\sqrt{2}$，而 $\frac{p^2+2q^2}{2pq}$ 正是把牛頓法的公式應用於 p/q 的結果）。平方令指數加倍，這就是 $y_n = x_{2^n-1}$ 的原因。同樣的「平方」模式，也是 Section 2.8 中迭代的動力來源。

思維技巧

分析誤差，而不是答案

要了解一個近似方法的表現，就寫下某一步的誤差如何取決於上一步的誤差。

這是數值分析的起點。兩個方法的答案本身（$\sqrt{2}$）是一樣的；不同之處，亦是實際應用中最要緊之處，在於誤差如何演變。

延伸練習

1 **佩爾方程。**證明：若 $y_n = p/q$ 為最簡分數，且 $p^2 - 2q^2 = 1$，則 $y_{n+1} = \dfrac{p^2 + 2q^2}{2pq}$ 亦滿足這個方程。驗證 $577^2 - 2 \times 408^2 = 1$。

2 **立方根。**寫出應用於 $f(y) = y^3 - 2$ 的牛頓法，並證明新誤差為 $\dfrac{(y_n - \alpha)^2(2y_n + \alpha)}{3y_n^2}$，其中 $\alpha = \sqrt[3]{2}$。由此推出：若 $y_n > \alpha$，則新誤差最多是舊誤差平方的 $\frac{1}{\alpha}$ 倍。（提示：利用 $\alpha^3 = 2$。）

3 **不用除法的除法。**證明應用於 $f(y) = \frac{1}{y} - a$ 的牛頓法給出 $y_{n+1} = y_n(2 - ay_n)$，當中無需除法。若 $e_n = 1 - ay_n$，證明 $e_{n+1} = e_n^2$。（電腦處理器正是用這類迭代來做除法的。）

4 **牛頓法失效的時候。**把牛頓法應用於 $f(y) = y^{1/3}$，它唯一的根是 0。證明無論從哪裏開始，每一步都令與根的距離加倍。誤差公式的哪一個假設不成立？

延伸閱讀

- **Endre Süli & David Mayers, *An Introduction to Numerical Analysis***（Cambridge University Press）。標準的大學本科數值分析教科書；第 1 章講述不動點迭代和牛頓法。
- **Nick Higham (ed.), *The Princeton Companion to Applied Mathematics***（Princeton University Press）。收錄有關牛頓法和數值分析的短文，由專家執筆，面向廣大讀者。
- **David Fowler & Eleanor Robson, "Square root approximations in Old Babylonian mathematics: YBC 7289 in context"**（古巴比倫數學中的平方根近似值），*Historia Mathematica* 25 (1998)。

2.8 平方再減二

題目

STEP III 2021，第 8 題

繼續閱讀之前，請先看題目（原題為英文）。題目載於 OCR 提供下載的 2021 年 STEP 試卷（zip 檔案）內的 STEP III 試卷：

`www.ocr.org.uk/Images/696588-2021-past-papers.zip`

引子

「平方後減二」幾乎是可以用來生成數列（sequence）的最簡單規則。然而它把實數線分成兩個世界。從 $[-2,2]$ 之外出發，各項會以比指數增長還要快的速度衝向無窮大。從區間之內出發，各項則永遠被困在其中，四處遊走，而且在一個精確的意義下完全無法預測：這是混沌（chaos）最簡單的例子之一。

同一條規則，用在非常大的整數上，就是盧卡斯－萊默檢驗法（Lucas–Lehmer test）的核心。這是一個高效的檢驗法，用來判斷梅森數（Mersenne number）2^p-1 是否質數（prime）；現時的最大質數紀錄，以及其他已知最大的質數中的大部分，都是靠它找到的。本題探討的是故事中「逃逸」的一面，並找出一條隱藏的公式，令這個數列完全可解。

動筆之前

先試幾個初始值。取 $a=3$：$3, 7, 47, 2207, \ldots$ 取 $a=1$：$1, -1, -1, -1, \ldots$ 取 $a=0$：$0, -2, 2, 2, 2, \ldots$ 取 $a=0.3$：$0.3, -1.91, 1.648, 0.716, -1.487, 0.211, \ldots$

想一想 為甚麼 $[-2,2]$ 內的數永遠無法逃出去？當 $-2 \le x \le 2$ 時，x^2-2 最大和最小可以是多少？

若 $-2 \le x \le 2$，則 $0 \le x^2 \le 4$，所以 $-2 \le x^2-2 \le 2$。區間 $[-2,2]$ 被映射到它自身之內。這就是第 (ii) 部分的一半。

逐步解題

階段 1　逃逸

想一想　第 (i) 部分出現了 4^{n-1}，暗示與 2 的距離每一步至少變成四倍。設 $x_n = 2 + d$。以 d 表示，x_{n+1} 是甚麼？

$x_{n+1} = (2+d)^2 - 2 = 2 + 4d + d^2$。所以若 $d \geq 0$，新一項高出 2 的距離是 $4d + d^2 \geq 4d$。現在用歸納法 (induction)。當 $n = 1$ 時，$x_1 = a = 2 + 4^0(a-2)$。若 $x_n \geq 2 + 4^{n-1}(a-2)$，則 $d = x_n - 2 \geq 4^{n-1}(a-2) \geq 0$，所以

$$x_{n+1} - 2 \geq 4d \geq 4^n(a-2). \qquad \blacksquare$$

階段 2　逃逸當且僅當 $|a| > 2$

想一想　「當且僅當」(if and only if) 表示要分別作兩個證明。第 (i) 部分差不多已經給出哪一個方向？還欠甚麼？

若 $|a| > 2$：當 $a > 2$ 時，由第 (i) 部分得 $x_n \geq 2 + 4^{n-1}(a-2) \to \infty$。當 $a < -2$ 時，第一步已把正負號糾正過來：$x_2 = a^2 - 2 > 4 - 2 = 2$，而由 x_2 開始的數列，就是以 $x_2 > 2$ 為初始值的同一個數列，所以它趨向無窮大。

僅當：若 $|a| \leq 2$，則 $a \in [-2, 2]$，由上面的觀察可知每一項都留在 $[-2, 2]$ 之內。有界 (bounded) 的數列不可能趨向無窮大。 $\blacksquare$

階段 3　一條隱藏的公式

第 (iii) 部分看來令人卻步。在動手之前，值得先找出 x_n 的公式。

想一想　若 $a > 2$，則存在一個數 $t > 1$ 使得 $a = t + \dfrac{1}{t}$。(為甚麼？) 以 t 表示，x_2 和 x_3 是甚麼？

當 t 由 1 增至 ∞ 時，函數 $t + 1/t$ 由 2 遞增至 ∞，所以每個 $a > 2$ 都可以寫成這個形式。於是

$$x_2 = \left(t + \frac{1}{t}\right)^2 - 2 = t^2 + \frac{1}{t^2}, \qquad x_3 = \left(t^2 + \frac{1}{t^2}\right)^2 - 2 = t^4 + \frac{1}{t^4},$$

再由歸納法得

$$x_n = t^{2^{n-1}} + t^{-2^{n-1}}.$$

平方後減 2，只不過是把指數加倍。這樣題目就變得容易處理了。

階段 4 乘積裂項相消

想一想 $\left(t-\frac{1}{t}\right)\left(t+\frac{1}{t}\right)$ 是甚麼？若把結果再乘以 $t^2+\frac{1}{t^2}$，會怎樣？

$\left(t-\frac{1}{t}\right)\left(t+\frac{1}{t}\right)=t^2-\frac{1}{t^2}$，然後 $\left(t^2-\frac{1}{t^2}\right)\left(t^2+\frac{1}{t^2}\right)=t^4-\frac{1}{t^4}$，如此類推：每個因子 x_k 都把兩個冪之差中的指數加倍。所以

$$\left(t-\frac{1}{t}\right)x_1x_2\cdots x_n=t^{2^n}-t^{-2^n}.$$

停一停，試一試 以 t 表示並計算 x_{n+1}^2-4，然後與上面的乘積比較。

$$x_{n+1}^2-4=\left(t^{2^n}+t^{-2^n}\right)^2-4=\left(t^{2^n}-t^{-2^n}\right)^2.$$

由於 $t>1$，$t^{2^n}-t^{-2^n}>0$，所以 $\sqrt{x_{n+1}^2-4}=t^{2^n}-t^{-2^n}=\left(t-\frac{1}{t}\right)x_1x_2\cdots x_n$。因此

$$\frac{\sqrt{x_{n+1}^2-4}}{x_{n+1}}=\frac{\left(t-\frac{1}{t}\right)x_1\cdots x_n}{x_{n+1}},$$

而它對每個 n 都等於 y_n，恰好就是 $A=t-\frac{1}{t}$ 的時候。最後，$\left(t-\frac{1}{t}\right)^2=\left(t+\frac{1}{t}\right)^2-4=a^2-4$，所以條件是 $A=\sqrt{a^2-4}$，即

$$a=\sqrt{A^2+4}.$$

（因為 $A>0$，這個值大於 2，符合要求。） ■

階段 5 y_n 是否收斂？

取這個 a 值時，

$$y_n=\frac{\sqrt{x_{n+1}^2-4}}{x_{n+1}}=\sqrt{1-\frac{4}{x_{n+1}^2}}.$$

由第 (ii) 部分知 $x_{n+1}\to\infty$，所以 $\frac{4}{x_{n+1}^2}\to 0$，而 $y_n\to 1$。數列收斂（converge），極限為 1。 ■

所以無窮乘積（infinite product）$x_1x_2x_3\cdots$ 的增長速度與 x_{n+1}/A 完全相同：我們找到了一個由急速增長的項所組成的乘積的精確增長率。

回顧

- **不變區域（invariant region）。**$[-2,2]$ 被映射到它自身之內，單憑這一點就解決了第 (ii) 部分的一半。

- **把問題線性化的代換（substitution）。**寫成 $a = t + 1/t$ 之後，「平方後減 2」就變成「把指數加倍」。雜亂的非線性行為變得簡單。
- **又是裂項相消（telescoping）。**與 Section 2.5 中的乘積一樣，由許多因子組成的乘積之所以能摺疊起來，是因為每個因子都與下一個互相配合。

背後的數學

秩序與混沌

在 $[-2, 2]$ 之內會發生甚麼事？在那裏 $a = t + 1/t$ 行不通（它需要 $|a| \geq 2$），但三角函數的版本卻可行：設 $a = 2\cos\theta$。則 $x_2 = 4\cos^2\theta - 2 = 2\cos 2\theta$，而一般地

$$x_n = 2\cos\bigl(2^{n-1}\theta\bigr).$$

所以在區間之內，這個映射只不過是*把角度加倍*。聽起來無傷大雅，但把角度加倍其實是混沌的。把 θ/π 寫成二進制（binary）：加倍會把二進制小數點向右移一位，並丟掉最前面的一位數字。經過 50 步之後，x_n 取決於 θ/π 的第 51 位二進制數字，所以初始值若有 10^{-15} 的誤差（一般電腦的精確度），50 步之後的答案就完全不可靠。這種對初始條件的敏感依賴（sensitive dependence on initial conditions）正是混沌的標誌，見圖 2.8.1。映射 $x \mapsto x^2 - 2$ 經變數代換後，等價於著名的邏輯斯諦映射（logistic map）$x \mapsto 4x(1-x)$，我們會在 Section 8.7 再遇上它。

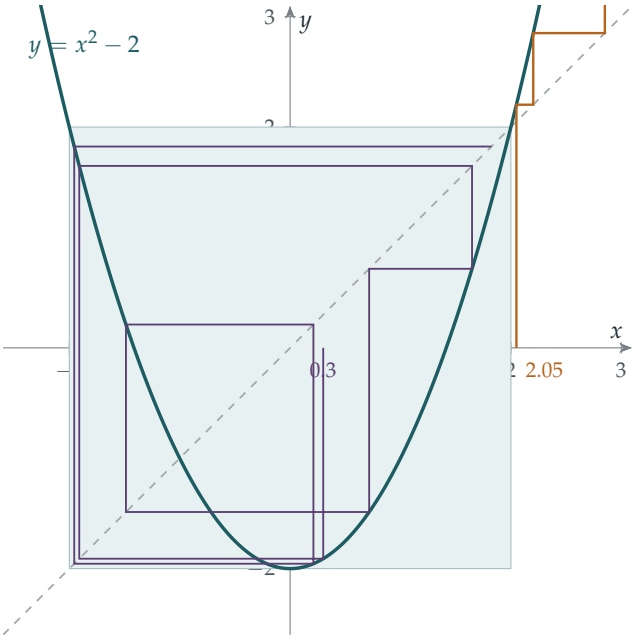


圖 2.8.1 $x_{n+1} = x_n^2 - 2$ 的蛛網圖（cobweb diagram）。由 0.3 出發（紫紅色），軌道（orbit）永遠留在陰影正方形 $[-2,2]^2$ 之內，混沌地遊走。由區間外不遠處的 2.05 出發（琥珀色），它就會逃逸：$2.05, 2.20, 2.85, 6.13, \ldots$

尋找最大的質數

梅森數是指 $M_p = 2^p - 1$。盧卡斯（Édouard Lucas）在 1876 年發現了一個出色的檢驗法，用來判斷 M_p 是否質數（其中 p 為奇質數），其後萊默（Derrick Lehmer）在 1930 年把它完善：由 $s_0 = 4$ 開始，反覆平方後減 2，每次都取模（modulo）M_p 化簡。則

$$M_p \text{ 是質數} \iff s_{p-2} \equiv 0 \pmod{M_p}.$$

以 $M_7 = 127$ 為例：$s = 4, 14, 194 \equiv 67, 67^2 - 2 \equiv 42, 42^2 - 2 \equiv 111, 111^2 - 2 \equiv 0$，所以 127 是質數。它與本題的聯繫就在於那個代換：$4 = t + 1/t$，其中 $t = 2 + \sqrt{3}$，所以 $s_k = (2+\sqrt{3})^{2^k} + (2-\sqrt{3})^{2^k}$，而檢驗法的證明用到形如 $a + b\sqrt{3}$ 的數的算術，與 Section 1.6 的高斯整數（Gaussian integers）頗為相似。由於每一步只需要一次平方，這個檢驗法極為快速，這就是為甚麼已知最大的質數中的大部分，包括現時的紀錄保持者，都是梅森質數（並非全部都是：其他特殊形式，例如 $a^{2^n} + 1$，也有各自的快速檢驗法）。在本書撰寫之時，紀錄是 $2^{136279841} - 1$，這個數有多於 41 百萬位數字，於 2024 年發現。

曼德博集合

把 -2 換成任意常數 c，並容許複數：$z \mapsto z^2 + c$。令 0 的軌道保持有界的所有 c 所組成的集合，就是曼德博集合（Mandelbrot set）。我們的映射對應 $c = -2$，即曼德博集合最左邊的一點，而軌道保持有界的初始值所組成的集合（它的茹利亞集合（Julia set））正好是我們在第 (ii) 部分找到的區間 $[-2, 2]$。對於幾乎所有其他的 c，這個集合都是分形（fractal）。

思維技巧

轉換變數，令映射變得簡單

若一個迭代（iteration）看來很複雜，就找一個代換，把它變成簡單的東西，例如加倍或加上一個常數。

這裏 $x = t + 1/t$ 和 $x = 2\cos\theta$ 把 $x \mapsto x^2 - 2$ 變成了加倍。在動力系統（dynamical systems）理論中，這稱為*共軛*（conjugacy），而找到共軛往往是徹底理解一個系統的關鍵。

延伸練習

1 **區間之內。**證明若 $a = 2\cos\theta$，則 $x_n = 2\cos(2^{n-1}\theta)$。求所有使得 $x_3 = a$ 但 $x_2 \neq a$ 的 a 值。（*提示*：你需要 $4\theta \equiv \pm\theta$，模 2π。）

2 **手算盧卡斯－萊默檢驗法。**用這個檢驗法證明 $M_5 = 31$ 是質數，而 $M_{11} = 2047$ 不是。（對 M_{11} 你需要九步；用計算機會有幫助。把 2047 作因數分解，以驗證你的結論。）

3 **一個具體例子。**取 $A = \frac{3}{2}$。求 a 和 t，並以數值驗證公式 $y_1 = \sqrt{x_2^2 - 4}/x_2$。

4 **裂項相消的和。**較難 設 $a = t + 1/t$ 且 $t > 1$，證明

$$\frac{1}{x_1 x_2 \cdots x_n} = \frac{t - t^{-1}}{t^{2^n} - t^{-2^n}}.$$

利用恆等式 $\dfrac{1}{z-1} - \dfrac{1}{z^2-1} = \dfrac{z}{z^2-1}$，其中 $z = t^{2^n}$，由此推出

$$\frac{1}{x_1} + \frac{1}{x_1 x_2} + \frac{1}{x_1 x_2 x_3} + \cdots = \frac{1}{t}.$$

當 $a = 3$ 時，以數值驗證這個結果。

延伸閱讀

- **Ian Stewart, *Does God Play Dice?***（Penguin）。講述混沌的經典科普著作，當中包括邏輯斯諦映射。
- **Steven Strogatz, *Nonlinear Dynamics and Chaos***（CRC Press）。標準的大學本科入門教材；第 10 章講述像本題這樣的一維映射。
- **Marcus du Sautoy, *The Music of the Primes***（Fourth Estate）。書中包括搜尋梅森質數的故事。

2.9 逐步逼近 e 的分數

題目

STEP II 2021，第 8 題

繼續閱讀之前，請先看題目（原題為英文）。題目載於 OCR 提供下載的 2021 年 STEP 試卷（zip 檔案）內的 STEP II 試卷：

`www.ocr.org.uk/Images/696588-2021-past-papers.zip`

引子

怎樣才能找到逼近 e 的好分數？你可以把級數（series）$1+1+\frac{1}{2}+\frac{1}{6}+\cdots$ 截斷，但以分母的大小來衡量，這樣得到的分數並不算特別好。這道題的做法高明得多：它構造出一族積分，每個積分都是形如「整數 + 整數 $\times e$」的一個很小的數。如果 $a+be$ 極小，那麼 e 就非常接近 $-a/b$，於是這些積分便為我們送上一連串極佳的分數：$\frac{19}{7}$、$\frac{193}{71}$、$\frac{2721}{1001}$、…

同樣的技巧，即找出既很小、又形如「你所關心的那個數的整數組合」的積分，正是數學中一些最深刻的無理性證明背後的想法，包括埃爾米特（Hermite）對 e 是超越數（transcendental）的證明，以及阿佩里（Apéry）對 $\zeta(3)$ 是無理數（irrational）的證明。

動筆之前

看看 $[0,1]$ 上的被積函數（integrand）$\frac{t^n(1-t)^n}{n!}e^t$。

> **想一想** $t(1-t)$ 在 $[0,1]$ 上的最大值是多少？由此可知，當 n 增大時 T_n 的大小會怎樣變化？

$t(1-t)$ 的最大值是 $\frac{1}{4}$，在 $t=\frac{1}{2}$ 處取得。所以被積函數至多為 $\frac{e}{4^n n!}$，它縮小的速度快得驚人（圖 2.9.1）。這個「小」只是故事的一半。另一半，即 T_n 具有 a_n+b_ne 這種特殊形式，則來自分部積分法（integration by parts）。

逐步解題

階段 1　求導兩次

設 $p(t) = t^n(1-t)^n$。由積法則（product rule），

$$p'(t) = nt^{n-1}(1-t)^n - nt^n(1-t)^{n-1} = nt^{n-1}(1-t)^{n-1}(1-2t).$$

> **想一想**　再求導一次。你會得到兩項；把公因式 $nt^{n-2}(1-t)^{n-2}$ 抽出來。要配合題目要求的形式，你需要把 $(1-2t)^2$ 改寫成以 $t(1-t)$ 表示的式子。怎樣做？

用積法則對 $n\,[t(1-t)]^{n-1}(1-2t)$ 求導（留意 $\frac{\mathrm{d}}{\mathrm{d}t}[t(1-t)] = 1-2t$）：

$$\begin{aligned} p''(t) &= n(n-1)[t(1-t)]^{n-2}(1-2t)^2 - 2n[t(1-t)]^{n-1} \\ &= nt^{n-2}(1-t)^{n-2}\Big[(n-1)(1-2t)^2 - 2t(1-t)\Big]. \end{aligned}$$

現在 $(1-2t)^2 = 1-4t+4t^2 = 1-4t(1-t)$，所以方括號內的式子是 $(n-1) - 4(n-1)t(1-t) - 2t(1-t) = (n-1) - 2(2n-1)t(1-t)$。 ■

階段 2　分部積分兩次

> **想一想**　$p(t)$ 在 $t=0$ 和 $t=1$ 處都等於零。當 $n \geq 2$ 時，$p'(t)$ 又如何？在對 $p(t)e^t$ 作分部積分時，這一點為甚麼重要？

當 $n \geq 2$ 時，$p'(t)$ 含有因式 $t^{n-1}(1-t)^{n-1}$，所以它在兩個端點也等於零。分部積分兩次，每一個邊界項都是零：

$$\int_0^1 p\,e^t\,\mathrm{d}t = \Big[p\,e^t\Big]_0^1 - \int_0^1 p'\,e^t\,\mathrm{d}t = -\Big[p'\,e^t\Big]_0^1 + \int_0^1 p''\,e^t\,\mathrm{d}t = \int_0^1 p''\,e^t\,\mathrm{d}t.$$

由第 1 階段，$p''(t) = n(n-1)\,t^{n-2}(1-t)^{n-2} - 2n(2n-1)\,t^{n-1}(1-t)^{n-1}$。除以 $n!$：

$$\frac{p''(t)}{n!} = \frac{t^{n-2}(1-t)^{n-2}}{(n-2)!} - 2(2n-1)\frac{t^{n-1}(1-t)^{n-1}}{(n-1)!}.$$

乘以 e^t 再積分，便得 $T_n = T_{n-2} - 2(2n-1)T_{n-1}$。 ■

階段 3　整數結構

$T_0 = \int_0^1 e^t\,\mathrm{d}t = e-1$。至於 $T_1 = \int_0^1 (t-t^2)e^t\,\mathrm{d}t$，由分部積分得 $\int_0^1 te^t\,\mathrm{d}t = 1$ 及 $\int_0^1 t^2e^t\,\mathrm{d}t = e-2$，所以 $T_1 = 1-(e-2) = 3-e$。

想一想 $T_0 = -1 + 1\cdot e$ 而 $T_1 = 3 + (-1)\cdot e$。為甚麼每個 T_n 都必定具有 $a_n + b_n e$ 的形式，其中 a_n、b_n 是整數？

用強歸納法（strong induction）：若 $T_{n-2} = a_{n-2} + b_{n-2}e$ 及 $T_{n-1} = a_{n-1} + b_{n-1}e$，其中係數都是整數，則

$$T_n = (a_{n-2} - 2(2n-1)a_{n-1}) + (b_{n-2} - 2(2n-1)b_{n-1})e,$$

而新的係數也是整數。 ■

階段 4 很小，因而帶來資訊

當 $0 < t < 1$ 時，被積函數是正數，並且 $t^n(1-t)^n \le 1$ 及 $e^t < e$。所以

$$0 < T_n < \int_0^1 \frac{e}{n!}\,\mathrm{d}t = \frac{e}{n!}.$$

停一停，試一試 把 $0 < a_n + b_n e < \dfrac{e}{n!}$ 除以 $|b_n|$。這說明 e 與 $-a_n/b_n$ 之間的距離如何？

除以 $|b_n|$（它是正整數，所以至少是 1），

$$\left| e - \left(\frac{-a_n}{b_n}\right)\right| = \frac{|a_n + b_n e|}{|b_n|} < \frac{e}{n!\,|b_n|} \le \frac{e}{n!}.$$

右邊趨向 0，所以 $-a_n/b_n \to e$。 ■

以下是由遞推關係（recurrence）算出的頭幾個分數：

n	2	3	4	5	6
$-a_n/b_n$	$\frac{19}{7}$	$\frac{193}{71}$	$\frac{2721}{1001}$	$\frac{49171}{18089}$	$\frac{1084483}{398959}$
誤差	4×10^{-3}	3×10^{-5}	1×10^{-7}	3×10^{-10}	5×10^{-13}

分數 $\frac{2721}{1001}$ 準確至小數點後 6 位，而分母只有四位數。

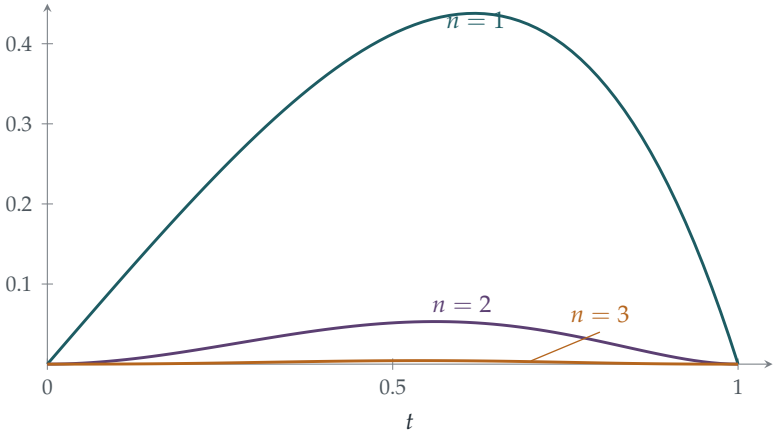


圖 2.9.1　$n = 1, 2, 3$ 時的被積函數 $\frac{t^n(1-t)^n}{n!}e^t$。每一個在 $(0, 1)$ 上都是正數，所以每個 T_n 都是正數；而且它們縮小得極快，到 $n = 3$ 時曲線已幾乎看不見。是正數但極小，正是這個論證所需要的。

回顧

- **分部積分是製造遞推關係的機器。**由於多項式及其導數在兩個端點都等於零，分部積分給出了一個沒有邊界項的簡潔遞推關係。
- **結構與大小並用。**遞推關係說明 T_n 具有 $a_n + b_n e$ 的形式；上界說明它很小。單憑其中任何一項，都無法說明怎樣逼近 e。
- **不需要的就不要算。**題目叫我們不必求出 a_n 和 b_n 的值，而論證也從來用不着它們；所需要的只是它們是整數。

背後的數學

e 是無理數的另一個證明

這道題離一個著名定理只差一步。假設 $e = p/q$，其中 p, q 是正整數。那麼

$$q\,T_n = q a_n + b_n p$$

是一個整數。但是 $0 < q\,T_n < \dfrac{qe}{n!}$，而當 n 很大時，右邊小於 1。於是 qT_n 就會是一個嚴格介乎 0 與 1 之間的整數，這是不可能的。因此 e 是無理數。這正是 Section 1.4 中「把一個整數困在 0 與 1 之間」的想法，只是這次由積分而不是級數來實現。正是這一類積分，配合 $t^n(1-t)^n$ 這樣的多項式，在 1873 年被夏爾·埃爾米特（Charles Hermite）用來證明 e 是*超越數*：它不是任何整係數多項式的根。

連分數

表中的分數不僅是好；它們更屬於*最佳之列*。e 的連分數（continued fraction）由歐拉（Euler）發現，有一個優美的規律：

$$e = [2; 1, 2, 1, 1, 4, 1, 1, 6, 1, 1, 8, \ldots],$$

而它的漸近分數（convergents）是 $2, 3, \frac{8}{3}, \frac{11}{4}, \frac{19}{7}, \frac{87}{32}, \frac{106}{39}, \frac{193}{71}, \ldots, \frac{2721}{1001}, \ldots$。我們的 $-a_n/b_n$ 恰好是每隔兩個取一個的漸近分數。沒有任何分母更小的分數能更接近 e。歐拉發現的這個規律本身就是 e 是無理數的第一個證明，因為只有有理數才有有限的連分數。

阿佩里的奇蹟

1978 年，羅傑·阿佩里（Roger Apéry）宣佈證明了 $\zeta(3) = 1 + \frac{1}{8} + \frac{1}{27} + \cdots$ 是無理數，這是自歐拉以來一直懸而未決的問題（見 Section 2.3）。他的證明實在太出人意表，以致在場聽眾起初並不相信。一年後，Frits Beukers 找到一個簡短的版本，所用的積分形如

$$\int_0^1 \int_0^1 \frac{-\ln(xy)}{1-xy}\, P_n(x) P_n(y)\, \mathrm{d}x\, \mathrm{d}y,$$

其中 P_n 由 $\frac{\mathrm{d}^n}{\mathrm{d}x^n}\left(x^n(1-x)^n\right)$ 構造而成。這些積分很小，是正數，並且形如「整數 + 整數 $\times \zeta(3)$」再除以一個受控的分母：這正是本題的策略。至於 $\zeta(5)$ 或 $e + \pi$ 是否無理數，至今仍然未知。

思維技巧

很小但不是零

如果已知某個量是若干個數的整數組合，又已知它很小但不是零，它就帶有資訊：關於逼近的資訊，有時還有關於無理性的資訊。

這個模式就是「結構來自代數，大小來自分析」。它在數論中反覆出現，也是分析與代數交匯得最有成果的地方之一。

延伸練習

1. **運行遞推關係。**由遞推關係計算 $n = 2, 3, 4$ 時的 a_n 和 b_n，並核對表中的分數。證明 b_n 的正負號交替出現。

2. **更精確的上界。**證明 $0 < T_n < \dfrac{e}{4^n\, n!}$。這把 $-a_n/b_n$ 的誤差估計改進了多少？

3. **無理性的證明。**把上面方框中的論證仔細寫出來。你究竟在哪一處用到 $T_n > 0$？

4. **換一個權函數。**較難把 e^t 換成 e^{2t}，重做這道題（即 $U_n = \int_0^1 \frac{t^n(1-t)^n}{n!} e^{2t}\,\mathrm{d}t$）。求 U_n 的遞推關係，證明 U_n 是 1 和 e^2 的有理係數組合，並探究這個論證是否仍能證明 e^2 是無理數。（與 Section 1.4 比較。）

延伸閱讀

- **Ivan Niven, *Irrational Numbers***（Mathematical Association of America）。關於無理性證明的經典小書，當中包括埃爾米特的方法。
- **Alfred van der Poorten, 「A proof that Euler missed」**, *The Mathematical Intelligencer* 1 (1979)。生動地記述阿佩里的證明及其所得到的迴響。
- **Julian Havil, *The Irrationals***（Princeton University Press）。一部無理數的歷史，為一般數學讀者而寫。

我們學到了甚麼

我們由芝諾（Zeno）說起，並提出了關於無窮過程的三個問題：說無窮過程有答案，究竟是甚麼意思？怎樣證明它們確實有答案？它們又以多快的速度到達答案？做過九道 STEP 題目之後，讓我們看看自己走了多遠。

1. 說無窮過程有答案，是甚麼意思？

在 Section 2.1 中，我們學會了這套語言。凡是描述數列（sequence）長遠表現的命題，都具有這樣的形式：「存在某一點 m，使得對所有 $n \geq m$...」。要證明它，就要把 m 找出來；要否定它，就要證明任何人提出的 m 都不管用。極限（limit）的定義只是在最前面多加一個量詞（quantifier）：「對任意容許誤差 $\varepsilon > 0$」。有了這個定義，芝諾悖論便迎刃而解：無論你要求的容許誤差多小，$\frac{1}{2} + \frac{1}{4} + \cdots + \frac{1}{2^n}$ 由某一點起，與 1 的距離都在這個誤差之內，而「和等於 1」的全部意思，不過如此。

這套語言也能防止我們說出毫無意義的話。在 Section 2.2 中，「設 x 等於那個無窮表達式」這種取巧的論證，為黃金比（golden ratio）給出了正確答案，但那只是因為我們能夠另行證明極限存在。把同一個論證用在 $1 - 1 + 1 - \cdots$ 上，得出的答案卻毫無意義。

2. 不知道極限是甚麼，怎樣證明收斂？

本章大部分小節所用的，都不外乎以下幾種工具。

- **單調有界。**遞增而有上界的數列必定收斂（converge）（Section 2.2、Section 2.6）。各項皆為正數的部分和（partial sum）亦然（Section 2.3）。到了大學，你會發現這其實是關於實數本身的一個命題：它就是實數的*完備性*（completeness）。
- **比較。**一個正項級數（series）如果比某個收斂的級數小，它就收斂；如果比某個發散（diverge）的級數大，它就發散。把各項分組成段之後，比較起來便很容易，由此可見調和級數（harmonic series）發散，而 Kempner 那個刪減了部分項的版本卻收斂（Section 2.3）。
- **夾逼。**把要研究的量夾在兩個較簡單而極限相同的量之間。我們就是這樣證明 $\sum 1/m^2 = \pi^2/6$（Section 2.4），並嚴格地計算出算術幾何平均（AGM）積分的值（Section 2.6）。

- **有限階段的精確公式。**在 Section 2.5 和 Section 2.8 中，我們利用裂項相消（telescoping）求出 n 步之後的精確表達式，到了最後一刻才讓 $n \to \infty$。

一旦知道極限存在，往往只需在遞推關係（recurrence）中「讓 $n \to \infty$」便可求出極限，例如黃金比的 $L = (2L+1)/(L+1)$。但次序十分重要：先證存在，後求數值。

3. 有多快？

在這方面，本章有些出人意表的發現。韋達（Viète）乘積（Section 2.5）每多乘一個因子，大約多得 0.6 個小數位。Section 2.7 中計算 $\sqrt{2}$ 的簡單方法，每一步大約多得 0.8 個數位。牛頓法（Newton's method）和 AGM（Section 2.6）則每一步都令正確數位的數目*倍增*：牛頓法只需十步，便可得出 $\sqrt{2}$ 的 700 多個數位。差別在於誤差遞推關係的形式：一種是「新誤差 $\approx$ 常數 $\times$ 舊誤差」（線性），另一種是「新誤差 $\approx$ 常數 $\times$（舊誤差）^{2}」（二次）。而在 Section 2.8 中，我們見到了另一個極端：由僅僅大於 2 的數出發，映射（map）$x \mapsto x^2 - 2$ 會以比指數增長還要快的速度把數推開；而在 $[-2, 2]$ 之內，它會把微小的誤差不斷放大，直至完全無法預測。

速度對理論同樣重要。Section 2.9 中的積分所產生的分數，趨近 e 的速度極快，足以證明 e 是無理數（irrational）：任何分數 p/q 都不可能被其他分數以這樣快的速度逼近。

4. 分析學的工具箱

方法	你在哪裏用過
找出量詞所要求的對象與否定量詞	「最終」（2.1）
單調有界數列	黃金比（2.2）；AGM（2.6）
比較與凝聚（condensation）	調和級數與 Kempner 級數（2.3）
夾逼	$\pi^2/6$（2.4）；AGM 積分（2.6）
根的對稱函數（symmetric function）	$\cot^2$ 值之和（2.4）
裂項相消的和與積	韋達（2.5）；$x^2 - 2$（2.8）
不變量（invariant）	AGM 積分（2.6）
誤差遞推關係	線性收斂與二次收斂（2.7）
以變數代換簡化映射	$a = t + 1/t$、$a = 2\cos\theta$（2.8）
細小但非零的整數組合	逼近 e（2.9）

留意這些工具當中，有多少附帶着一條「你可以直接使用的定理」。STEP 經常直接把單調收斂定理（monotone convergence theorem）交給你，或者叫你假設極限存在。到了大學，你要由實數的公理（axiom）出發證明這些定理，而且當一個論證不動聲色地假設了極限存在，你也要自己察覺得到。閱讀解答的時候，不妨養成一個很好的習慣，問問自己：*究竟在哪一步取了極限？這一步有沒有根據？*

下一步

本章的想法貫穿全書其餘部分。

- 不動點（fixed point）和蛛網圖（cobweb diagram）會在 Section 8.7 再次出現，在那裏，一個混沌（chaotic）映射原來有精確解。
- 調和數（harmonic number）$1 + \frac{1}{2} + \cdots + \frac{1}{n}$ 會在 Section 9.11 的概率和 Section 11.12 的力學中出現。
- π 另一次在極限中現身，是在 Section 12.23，這次來自概率。
- 對極限的審慎態度是第 3 章的主題，該章以 Section 3.1 開始：一些看似正確、其實不然的論證。

如果你想在上大學之前再進一步

- **Lara Alcock, *How to Think About Analysis***：大學一年級的分析學是怎樣的，又該如何應付。
- **David Bressoud, *A Radical Approach to Real Analysis***：循歷史脈絡講授分析學，由那些迫使數學家不得不嚴謹起來的難題說起。
- **Tom Körner, *Calculus for the Ambitious***（Cambridge University Press）：由一位劍橋的分析學家執筆，正是為本書這樣的讀者而寫。
- **Julian Havil, *Gamma*** 和 **William Dunham, *Euler: The Master of Us All***：本章許多級數背後的故事。

芝諾認為無窮令運動變得不可能。數學家花了兩千年才回應他，而答案原來是人類發現過最強而有力的想法之一：一種能夠精確地描述永無止境的過程的方法。這套語言最初的幾個詞語，你現在已經認識了。到了分析學，你便會學懂怎樣流利地說這套語言。

延伸練習答案

以下只是簡短的答案和提示，並非完整解答：這裏只給出結果和關鍵步驟，方便你核對自己的做法。請先認真做過每一道練習，然後才翻看。

2.1 「最終」是甚麼意思

1. $(1+x)^n = 1 + nx + \binom{n}{2}x^2 + \cdots$ 的各項均為非負，所以 $(1+x)^n \geq 1 + nx > nx$。設 $y = 1 + x$，則只要 $n > K/(y-1)$，便有 $y^n > n(y-1) > K$。利用下一項，$y^n > \binom{n}{2}(y-1)^2$，所以只要 $n > 1 + 2K/(y-1)^2$，便有 $y^n/n > \frac{1}{2}(n-1)(y-1)^2 > K$。

2. 若當 $n \geq m_1$ 時 $s_n \leq t_n$，而當 $n \geq m_2$ 時 $t_n \leq s_n$，則當 $n \geq \max(m_1, m_2)$ 時 $s_n = t_n$。

3. 例如：當 n 為偶數時 $s_n = n + 0.6$，當 n 為奇數時 $s_n = n$；而當 n 為奇數時 $t_n = n + 0.6$，當 n 為偶數時 $t_n = n$。兩個數列（sequence）都是遞增的，而且各自都有無限多次大於對方。

4. 當 $n \geq 4$ 時 $2^n \leq n!$（用歸納法（induction）：$2^{n+1} = 2 \cdot 2^n \leq (n+1)\,n!$）。至於 $n^{10} \leq 2^n$，這個不等式對由 2 至 58 的每一個 n 都不成立，而對所有 $n \geq 59$ 都成立，所以最小的 m 是 59。

2.2 作為極限的黃金比例

1. 用歸納法：$u_1 = 1 = F_2/F_1$，而若 $u_n = F_{n+1}/F_n$，則 $u_{n+1} = 1 + F_n/F_{n+1} = F_{n+2}/F_{n+1}$。

2. 由於 $\varphi - 1 = 1/\varphi$，$u_{n+1} - \varphi = \frac{1}{u_n} - \frac{1}{\varphi} = -\frac{u_n - \varphi}{\varphi u_n}$。因為 $u_n \geq 1$，所以 $|u_{n+1} - \varphi| \leq |u_n - \varphi|/\varphi$，又 $|u_1 - \varphi| < 1$，故 $|u_n - \varphi| \leq \varphi^{-(n-1)}$。

3. $\psi = \frac{1-\sqrt{5}}{2}$ 滿足 $\psi^2 = \psi + 1$，所以 $1 + 1/\psi = \psi$。倒推回去，$u_{n-1} = \frac{1}{u_n - 1}$：由 0 出發，得到 $-1, -\frac{1}{2}, -\frac{2}{3}, -\frac{3}{5}, -\frac{5}{8}, \ldots$，即對某個 $k \geq 1$ 有 $u_1 = -F_{k-1}/F_k$。這些數值趨近 ψ。

4. 不動點（fixed point）滿足 $v = 1 + \frac{1}{1+v}$，即 $v^2 = 2$。誤差每次乘以 $-\frac{1}{(1+v_n)(1+\sqrt{2})}$，

其大小不超過 $\frac{1}{2(1+\sqrt{2})}$，所以 $v_n \to \sqrt{2}$。並且 $\sqrt{2} = 1 + \cfrac{1}{2 + \cfrac{1}{2 + \cdots}}$。

2.3 一個發散級數與它的收斂近親

1. 由 $\ln N + 0.5772 = 10$ 得 $N \approx 12\,400$；要超過 100，大約需要 $e^{99.4} \approx 1.5 \times 10^{43}$ 項。

2. 取 $k = 2$，第 n 組之和不超過 $2^n \cdot 4^{-n} = 2^{-n}$，所以總和不超過 $\sum_{n\ge 0} 2^{-n} = 2$。把 $1 + \frac14 + \frac19$ 準確加起來，再以 $\frac14 + \frac18 + \cdots = \frac12$ 作為由 $r = 4$ 起各組的上界，便得到約 1.86。（真確值是 $\frac{\pi^2}{6} \approx 1.645$。）

3. 對於 $1/(r \ln r)$，由 2^n 至 $2^{n+1} - 1$ 的一組之和至少是 $\frac{2^n}{2^{n+1}(n+1)\ln 2} = \frac{1}{2(n+1)\ln 2}$，而這些數之和的表現與調和級數（harmonic series）相同：級數發散（divergence）。對於 $1/(r(\ln r)^2)$，該組之和不超過 $\frac{1}{n^2(\ln 2)^2}$，而 $\sum \frac{1}{n^2}$ 收斂（converges）。

4. 可用的數字：首位數字有 7 個選擇，其餘每一位各有 8 個選擇，所以符合條件的 d 位數最多有 $7 \cdot 8^{d-1}$ 個，而每個都至少是 10^{d-1}。總和不超過 $\sum_{d\ge 1} 7(0.8)^{d-1} = \frac{7}{1-0.8} = 35$。

2.4 $\pi^2/6$

1. 各個 $\tan^2 = 1/\cot^2$ 的值是倒序多項式 $\sum_j (-1)^j \binom{2n+1}{2j+1} y^j$ 的根（roots），其根之和為 $\binom{2n+1}{2n-1} = n(2n+1)$。然後由 $\operatorname{cosec}^2 = 1 + \cot^2$ 得 $n + \frac{n(2n-1)}{3} = \frac{2n(n+1)}{3}$。

2. 偶數項之和為 $\frac14 \sum \frac{1}{m^2} = \frac{\pi^2}{24}$，所以奇數項之和為 $\frac{\pi^2}{6} - \frac{\pi^2}{24} = \frac{\pi^2}{8}$。

3. 這些根滿足 $e_1 = \frac{n(2n-1)}{3}$ 和 $e_2 = \binom{2n+1}{5}\big/\binom{2n+1}{1}$，由此得 $\sum \cot^4 = e_1^2 - 2e_2 = \frac{n(2n-1)(4n^2+10n-9)}{45}$。利用 $\cot^4\theta < \theta^{-4} < (1+\cot^2\theta)^2$ 作夾逼（squeezing），再令 $n \to \infty$，便得到 $\frac{\pi^4}{90}$。

4. 由下界得 $\sum_{m\le n} \frac{1}{m^2} > \frac{\pi^2 n(2n-1)}{3(2n+1)^2}$，所以誤差小於 $\frac{\pi^2(6n+1)}{6(2n+1)^2} < \frac{\pi^2}{4n}$。這保證了當 $n \ge 2468$ 時準確度達到 0.001；事實上 $n \approx 1000$ 已經足夠。

2.5 韋達乘積

1. 反覆使用 $\sin\theta = 2\sin\frac{\theta}{2}\cos\frac{\theta}{2}$：乘積會逐項相消（telescopes）。令 $n \to \infty$（於是 $2^n \sin(\alpha/2^n) \to \alpha$），並取 $\alpha = \frac{\pi}{2}$，再利用 $\cos\frac{\pi}{4} = \sqrt{\frac12}$ 和 $\cos\frac{\theta}{2} = \sqrt{\frac12 + \frac12\cos\theta}$。

2. $\frac{\sin(\pi/3)}{\pi/3} = \frac{3\sqrt3}{2\pi} = \cos\frac{\pi}{6}\cos\frac{\pi}{12}\cos\frac{\pi}{24}\cdots$，其中 $\cos\frac{\pi}{6} = \frac{\sqrt3}{2}$，而 $\cos\frac{\pi}{12} = \frac{\sqrt{2+\sqrt3}}{2} = \frac{\sqrt6+\sqrt2}{4}$。

3. 取 8 個因子得 3.14157，四捨五入後是 3.1416。取 n 個因子後，誤差約為 $\frac{\pi^3}{6\cdot 4^{n+1}}$，因為被略去的因子是 $\cos\frac{\pi}{2^{j+1}} \approx 1 - \frac{\pi^2}{2\cdot 4^{j+1}}$：每多一個因子，誤差約縮小至原來的 4 分之一，而 $\log_{10} 4 \approx 0.6$。

4. 設 $x = \ln a$，同樣的論證給出

$$\frac{a-1}{\ln a} = \prod_{j\geq 1} \frac{1+a^{1/2^j}}{2},$$

所以 $\frac{1}{\ln 3} = \frac{1}{2}\prod_{j\geq 1}\frac{1+3^{1/2^j}}{2}$。取七個因子得 0.906，四捨五入後是 0.91（真確值是 0.9102）。這裏的因子趨近 1 的方式如同 $1 + \frac{\ln 3}{2^{j+1}}$，所以每多一個因子只能把誤差減半（約 0.3 位數字）；而韋達（Viète）的因子趨近 1 的方式如同 $1 - \frac{\pi^2}{2\cdot 4^{j+1}}$，每次把誤差縮小至四分之一。

2.6 算術幾何平均

1. $x_{n+1} - y_{n+1} = \frac{1}{2}(\sqrt{x_n} - \sqrt{y_n})^2 = \frac{(x_n - y_n)^2}{2(\sqrt{x_n}+\sqrt{y_n})^2}$，而 $(\sqrt{x_n} + \sqrt{y_n})^2 \geq 4y_n \geq 4y_0$。

2. $M(2,1) = 1.456791$；三步之後，x_3 與 y_3 已經有七位小數相同。

3. 迭代（iteration）的每一步都按 λ 的比例縮放，所以 $M(\lambda x, \lambda y) = \lambda M(x,y)$。在積分中代入 $x = \lambda u$，得 $\mathrm{I}(\lambda p, \lambda q) = \lambda^{-1}\mathrm{I}(p,q)$。$\mathrm{I} = \frac{\pi}{2M}$ 的兩邊都按 λ^{-1} 的比例縮放。

4. $M(1, \cos 30^\circ) \approx 0.9318$，所以週期變長至原來的 1.073 倍：即長了約 7.32%。

2.7 六位還是六百位

1. $(p^2+2q^2)^2 - 2(2pq)^2 = (p^2-2q^2)^2 = 1$；並且 $577^2 - 2\times 408^2 = 332\,929 - 332\,928 = 1$。

2. $y_{n+1} = \frac{2y_n^3+2}{3y_n^2}$，而 $y_{n+1} - \alpha = \frac{2y^3 - 3\alpha y^2 + \alpha^3}{3y^2} = \frac{(y-\alpha)^2(2y+\alpha)}{3y^2}$。若 $y > \alpha$，則 $\frac{2y+\alpha}{3y^2} < \frac{3y}{3y^2} = \frac{1}{y} < \frac{1}{\alpha}$。

3. $y - \frac{1/y - a}{-1/y^2} = y + y - ay^2 = y(2-ay)$，而 $1 - ay(2-ay) = (1-ay)^2$。

4. $y - \frac{y^{1/3}}{\frac{1}{3}y^{-2/3}} = y - 3y = -2y$。誤差公式要求在根的附近 f′ 不等於零，而且 f″ 有界（bounded）；但這裏 f′ 在 0 處是無限大。

2.8 平方再減二

1. $(2\cos u)^2 - 2 = 2\cos 2u$，所以由歸納法得 $x_n = 2\cos(2^{n-1}\theta)$。$x_3 = a$ 需要 $4\theta \equiv \pm\theta$：即 3θ 或 $5\theta \equiv 0 \pmod{2\pi}$。排除 $x_2 = a$ 的情況（不動點 $a = 2$ 和 $a = -1$）

後，剩下 $\theta = \frac{2\pi}{5}, \frac{4\pi}{5}$：即 $a = \frac{\sqrt{5}-1}{2}$ 或 $a = -\frac{\sqrt{5}+1}{2}$，也就是一個 2-循環（cycle）的兩個點。

2. 對於 M_5，數列是 4，14，$194 \equiv 8$，$62 \equiv 0$：是質數（prime）。對於 M_{11}，數列是 4，14，194，788，701，119，1877，240，282，$1736 \neq 0$：是合成數（composite），而事實上 $2047 = 23 \times 89$。

3. 由 $A = t - \frac{1}{t} = \frac{3}{2}$ 得 $t = 2$ 和 $a = \frac{5}{2}$。於是 $x_2 = \frac{17}{4}$，$y_1 = \frac{Ax_1}{x_2} = \frac{15}{17}$，而 $\frac{\sqrt{x_2^2-4}}{x_2} = \frac{15/4}{17/4} = \frac{15}{17}$。

4. $x_k = t^{2^{k-1}} + t^{-2^{k-1}}$，而把 $x_1 \cdots x_n$ 乘以 $t - t^{-1}$ 後，逐項相消得 $t^{2^n} - t^{-2^n}$。所求之和變成 $(t - t^{-1}) \sum_{n\geq 1} \left(\frac{1}{t^{2^n}-1} - \frac{1}{t^{2^{n+1}}-1} \right) = \frac{t-t^{-1}}{t^2-1} = \frac{1}{t}$。當 $a = 3$ 時，$t = \frac{3+\sqrt{5}}{2}$ 而 $\frac{1}{t} \approx 0.3820$；級數 $\frac{1}{3} + \frac{1}{21} + \frac{1}{987} + \cdots$ 與此吻合。

2.9 逐步逼近 e 的分數

1. $T_0 = e - 1$，$T_1 = 3 - e$，$T_2 = 7e - 19$，$T_3 = 193 - 71e$，$T_4 = 1001e - 2721$，由此得到 $\frac{19}{7}$，$\frac{193}{71}$，$\frac{2721}{1001}$。由於 $b_n = b_{n-2} - 2(2n-1)b_{n-1}$，而（按歸納法）$b_{n-1}$ 與 b_{n-2} 異號，所以 b_n 與 b_{n-2} 同號。

2. $t(1-t) \leq \frac{1}{4}$，所以 $T_n < \frac{e}{4^n n!}$；$-a_n/b_n$ 的誤差是 $T_n/|b_n|$，現在改進了 4^n 倍。

3. 若 $e = \frac{p}{q}$，則 $qT_n = qa_n + pb_n$ 是整數。由 $T_n > 0$ 可知它是正整數，所以至少是 1；但當 n 足夠大時，$qT_n < \frac{qe}{n!} < 1$。

4. 分部積分（integration by parts）兩次（邊界項為零），得 $U_n = \frac{1}{4}\big(U_{n-2} - 2(2n-1)U_{n-1}\big)$，其中 $U_0 = \frac{e^2-1}{2}$ 和 $U_1 = \frac{1}{2}$。所以 $2^{2n+1}U_n = A_n + B_n e^2$，其中 A_n，B_n 為整數，而 $0 < U_n < \frac{e^2}{4^n n!}$。若 $e^2 = \frac{p}{q}$，則 $q2^{2n+1}U_n$ 是一個小於 $\frac{2qe^2}{n!}$ 的正整數，當 n 足夠大時這是不可能的。所以這個論證確實證明了 e^2 是無理數（irrational）。

章

3

函數、不等式與嚴謹

知道一件事是對的，不等於知道它爲甚麼對；數學就活在這兩者的差別之中。

本章內容

3.1	找出錯處	STEP I 1991
3.2	倒過來的 AM–GM	STEP II 1993
3.3	凸性的威力	STEP II 2007
3.4	以面積證明不等式	STEP III 2011
3.5	π^e 還是 e^π？	STEP II 2013
3.6	斜率的界	STEP I 1988
3.7	$\pi^2 \leq 10$	STEP II 2014
3.8	由加法公式決定的函數	STEP III 2019
3.9	創造一個新函數	STEP II 2021
A	練習答案	第 237 頁

把事情做對

在數學裏，你並不是理解了甚麼，你只是習慣了它們。

— 馮・諾伊曼（John von Neumann）

以下是 $1 = 2$ 的一個「證明」（proof）。設 $a = b$。那麼 $a^2 = ab$，所以 $a^2 - b^2 = ab - b^2$，因式分解後得 $(a + b)(a - b) = b(a - b)$。約去 $a - b$ 得 $a + b = b$，而由於 $a = b$，這就是說 $2b = b$，所以 $2 = 1$。

你大概已經看出問題所在：由於 $a = b$，「約去 $a - b$」其實就是除以零。但請留意，每一步看起來都多麼合理，而錯誤又多麼容易藏在一個例行的步驟之中。真正的數學裏，真正的錯誤很少這樣明顯，有史以來最偉大的一些數學家也曾被它們騙倒。本章要談的，正是能把這些錯誤揪出來的思考習慣：精確的定義、小心運用定理，以及對看來太過容易的論證抱持適度的懷疑。

為甚麼嚴謹很重要

在中學，數學主要是求出答案。到了大學，數學主要是知道自己的答案是對的，並且能夠說服別人。這個轉變是中學數學與大學數學之間最大的一個變化，也可能令人不安。一件「顯然」成立的事，為甚麼要花一整頁去證明？

答案是：「顯然」並不可靠。分析（analysis）的歷史上，滿是看似顯然、實則錯誤的結果：連續（continuous）函數除了可能在少數幾個孤立點之外必定可導（differentiable）（安培（André-Marie Ampère）在 1806 年發表過一個「證明」；到了 1872 年，魏爾斯特拉斯（Weierstrass）找到一個處處不可導的連續函數）；連續函數的收斂（convergent）和仍然連續（柯西（Cauchy）相信這一點；阿貝爾（Abel）找到了反例（counterexample））；收斂級數（series）的各項可以重新排列而不改變其和（一般而言，不可以）。嚴謹（rigour）並不是吹毛求疵，而是分辨真確與似是而非的唯一可靠方法。

不等式：分析的基本工具

本章的第二個主題是不等式（inequality）。在中學數學裏，大部分結果都是等式：這個等於那個。在分析裏，大部分結果都是不等式：這個至多是那個。要證明一個數列收斂，你要證明誤差小於任何容許的限度；要證明一個積分存在，你要證明某個和有界；要估計一個近似值有多好，你要給誤差定出一個界。專業的分析學家花在證明不等式上的時間，遠多於證明等式，而你將在這裏遇到的經典不等式（AM–GM、詹森（Jensen）、

楊（Young）、赫爾德（Hölder）、柯西—施瓦茲（Cauchy–Schwarz））正是這一行的看家工具。

這些不等式本身也十分精彩。一條關於函數的不等式，配上一個精心挑選的例子，就能證明 $\pi^2 \leq 10$（Section 3.7）。關於對數圖像形狀的一個事實，就能判定 π^e 與 e^π 哪個較大（Section 3.5）。一位化學家關於量度誤差的提問，引出了數學中最重要的一些多項式（Section 3.6）。

你在大學會遇到甚麼

嚴謹是大學一年級分析（Analysis）課程的主題，也是另一門常稱為數與集合（Numbers and Sets）或數學基礎（Foundations）的課程的主題，你會在那裏學習怎樣精確地書寫證明。你會遇到連續性和可導性的正式定義，以及帶有精確假設（hypothesis）的定理（介值定理（intermediate value theorem）、中值定理（mean value theorem）、極值定理（extreme value theorem）），而這些假設確實缺一不可。往後的課程會把本章的不等式發展成一門門完整的學科：泛函分析（functional analysis）研究函數空間，並以 Section 3.4 和 Section 3.7 中那類積分來量度其中的函數；凸分析（convex analysis）和最優化（optimisation）由 Section 3.3 發展而來；逼近論（approximation theory）由 Section 3.6 發展而來；而像 Section 3.9 中那樣的特殊函數（special functions），則在應用數學和物理學中隨處可見。

貫穿本章的問題

三個大問題

1. **怎樣知道一個論證是正確的？**看似合理的論證通常會在哪些地方出錯？又應該怎樣檢查？
2. **怎樣證明一條不等式？**經典不等式從何而來？又怎樣選擇用哪一條？
3. **函數究竟是甚麼？**一個函數可否由它具有的某種性質來定義，而不是由公式來定義？當舊有的函數不夠用時，我們可否創造新的函數？

本章以找錯開始（Section 3.1），然後逐一探討各條經典不等式，每條都從不同的角度切入：一種不尋常的歸納法（induction）、凸性（convexity）、面積、圖像的形狀、多項式插值（polynomial interpolation）和試驗函數（test function）。最後兩節討論由性質定義的函數：由加法定律（addition law）得出的 e^x 和 $\tanh x$，以及一個以反函數（inverse）形式定義的全新函數。

本章的數學思想

§	課題	思想
3.1	似是而非的錯誤論證	檢查計算、邏輯和隱藏的假設
3.2	AM–GM	前推—後推歸納法
3.3	詹森不等式	凹（concave）曲線位於其弦之上
3.4	楊氏不等式	在圖中比較面積
3.5	π^e 還是 e^π？	化為單一函數並研究其圖像
3.6	二次函數的斜率	以已知量表示未知量
3.7	$\pi^2 \leq 10$	把試驗函數代入一般的不等式
3.8	e^x 與 $\tanh x$	由函數方程（functional equation）定義的函數
3.9	W 函數	以反函數的形式創造一個函數

怎樣閱讀本章

和之前一樣，請先自己嘗試每道題目，然後才閱讀解答。在這一章，請再多做一件事：讀完每個解答之後，回頭找出粗心的學生會做錯的一兩個步驟。隱藏的假設會在哪裏悄悄混進來？哪條定理的哪個假設需要驗證？這樣帶著批判眼光、主動地閱讀，是你在上大學之前所能養成的最有用的一個習慣。

3.1 找出錯處

題目

STEP I 1991，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/91-s1-q6`

引子

大多數考試題目都要求你寫出一個論證。這一題卻給你兩個論證，要你把它們逐一拆解。這是一種截然不同的技巧，也是你在大學裏最常用到的技巧：檢查自己的作業、閱讀書中的證明（proof）並判斷是否相信它、批改朋友的解答、審閱論文。

這裏的論證錯得很有啟發性。每一步看起來都像你見過上百次的東西。有些步驟含有計算上的失誤，但更嚴重的是邏輯上的錯誤：即使計算無誤，結論也推不出來。這類錯誤會一直殘留到專業數學之中，而本章要談的正是如何把它們找出來。

動筆之前

? **想一想** 當題目要求你「評論每一步」時，你在每一步應該檢查甚麼？動筆之前先列一張清單。

一張好的清單至少有三個問題。計算正確嗎？這一步能從前面的步驟推出來嗎？有沒有隱藏的假設（例如函數是連續的、極限存在、駐點必定是極大點或極小點）？題目還問到結論會否無論如何都是對的。錯誤的論證也可以有正確的結論，所以這一點需要另行檢查。

逐步解題

階段 1　導數對嗎？

想一想 用商法則（quotient rule）對 g 求導。$g'(x) = 0$ 在哪裏成立？直接檢查 $x = 1$ 和 $x = -1$。

由商法則，

$$g'(x) = \frac{6x^2(x^4+4) - (2x^3+3)\cdot 4x^3}{(x^4+4)^2} = \frac{-2x^6 - 12x^3 + 24x^2}{(x^4+4)^2} = \frac{-2x^2(x^4+6x-12)}{(x^4+4)^2}.$$

所以當 $x = 0$ 或 $x^4 + 6x - 12 = 0$ 時，$g'(x) = 0$。在 $x = 1$，$x^4 + 6x - 12 = -5$；在 $x = -1$，其值為 -17。兩者都不是駐點（stationary point）。第一個斷言是**錯的**。

四次多項式 $x^4 + 6x - 12$ 在 $x = 0$ 的值是 -12，而且在兩個方向上都趨向 $+\infty$，所以它有（至少）一個正根和一個負根；它恰好有兩個實根，因為它的二階導數 $12x^2$ 非負，所以它是凸（convex）的，最多只能穿過零兩次。數值上，兩個根是 $x \approx 1.386$ 和 $x \approx -2.247$。

階段 2　函數值對嗎？

$g(0) = \frac{3}{4}$ 是正確的，而 $g(1) = \frac{5}{5} = 1$。但 $g(-1) = \frac{-2+3}{1+4} = \frac{1}{5}$，不是 1。即使按它自己的思路，這個論證也有計算錯誤：它似乎假設了 g 是偶函數。

階段 3　比較函數值能判斷哪個是極大點嗎？

想一想 假設一個函數恰好有兩個駐點，駐點值分別是 $\frac{3}{4}$ 和 1。能否由此推出前者是極小點、後者是極大點？

不能。駐點的性質是局部（local）的性質：它取決於函數在該點附近的表現，而不是該點的函數值與別處的函數值孰大孰小。一個函數的局部極大值可以低於局部極小值（想想 $x + 2\sin x$），駐點也可以兩者皆不是。在這裏，在 $x = 0$ 附近，因式 $x^4 + 6x - 12$ 接近 -12，所以

$$g'(x) = \frac{-2x^2(x^4+6x-12)}{(x^4+4)^2} > 0 \qquad \text{對足夠小的} x \neq 0.$$

導數在 0 的兩側都是正的：$x = 0$ 是一個駐拐點（stationary point of inflection），不是極小點。

階段 4　駐點值能界定函數的範圍嗎？

停一停，試一試 即使駐點都求對了，為甚麼仍不能斷定 g 的每一個值都介乎最小與最大的駐點值之間？找出一個使 $g(x) < \frac{3}{4}$ 的 x 值。

駐點只描述局部的表現。全局（global）的表現還取決於定義域兩端的情況，在這裏即 $x \to \pm\infty$ 時的情況。由於當 $x \to \pm\infty$ 時 $g(x) \to 0$，而且每當 $2x^3 + 3 < 0$，即 $x <$

$-\sqrt[3]{3/2} \approx -1.14$ 時，$g(x) < 0$，所以這個函數會取負值。例如 $g(-2) = \frac{-13}{20}$。因此結論 $\frac{3}{4} \le g(x) \le 1$ 是**錯的**（圖 3.1.1）。

正確的說法是：g 遞增至 $x \approx 1.386$ 處約為 1.083 的全局極大值，然後向 0 遞減；在左邊，它由 0 遞減至 $x \approx -2.247$ 處約為 -0.668 的全局極小值，然後遞增，經過 $x = 0$（在該處有一個水平拐點），一直升到極大值。所以對所有 x，$-0.668 \lesssim g(x) \lesssim 1.083$。

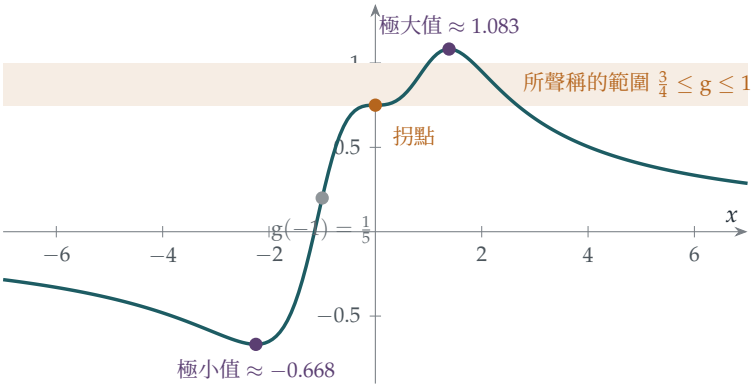


圖 3.1.1 $g(x) = \frac{2x^3+3}{x^4+4}$ 的真實圖像。該論證所聲稱的範圍（琥珀色帶）處處皆錯：$x = -1$ 不是駐點，而且 $g(-1) = \frac{1}{5}$；$x = 0$ 是拐點，不是極小點；函數更遠遠超出色帶之外，極大值高於 1，極小值低於 0。

階段 5 積分

想一想 對 $-3(1-x)^{-4}$ 求導。你會得到 $(1-x)^{-3}$ 嗎？

$\frac{\mathrm{d}}{\mathrm{d}x}\left[-3(1-x)^{-4}\right] = -3 \cdot (-4)(1-x)^{-5} \cdot (-1) = -12(1-x)^{-5}$。冪和係數都錯了。正確的原函數（antiderivative）是

$$\int (1-x)^{-3}\,\mathrm{d}x = \tfrac{1}{2}(1-x)^{-2} + C,$$

你可以求導來驗證。

想一想 用正確的原函數，$\left[\frac{1}{2}(1-x)^{-2}\right]_{-1}^{3} = \frac{1}{2}\cdot\frac{1}{4} - \frac{1}{2}\cdot\frac{1}{4} = 0$。那麼答案原來還是對的。真的嗎？

並不是。微積分基本定理（fundamental theorem of calculus）$\int_a^b F'(x)\,\mathrm{d}x = F(b) - F(a)$ 要求被積函數在整個區間 $[a,b]$ 上連續（continuous）。這裏 $(1-x)^{-3}$ 在 $x = 1$ 處沒有定義，而這一點位於 $[-1,3]$ 之內，而且在 $x = 1$ 附近函數值極其巨大。這個積分必須當作瑕積分（improper integral）來處理，在出問題的那一點拆開：

$$\int_{-1}^{1} (1-x)^{-3}\,\mathrm{d}x = \lim_{\varepsilon\to 0^+}\left[\tfrac{1}{2}(1-x)^{-2}\right]_{-1}^{1-\varepsilon} = \lim_{\varepsilon\to 0^+}\left(\frac{1}{2\varepsilon^2} - \frac{1}{8}\right) = +\infty.$$

由 -1 到 1 的積分是無窮大（而由 1 到 3 的積分是 $-\infty$），所以 $\int_{-1}^{3}(1-x)^{-3}\,dx$ **不存在**。結論「$= 0$」是錯的，原因不在於真正的值是另一個數，而在於根本沒有值。 ■

回顧

- **三類錯誤。**計算錯誤（導數、$g(-1)$、原函數）、邏輯錯誤（靠比較函數值來判別駐點的性質；把駐點值當作全局的界）以及隱藏的假設（微積分基本定理中的連續性）。
- **答案對，理由錯。**在第 (ii) 部分改正原函數之後，仍然得出答案 0，但論證依然無效。答案看似正確，並不能證明推理站得住腳。
- **局部與全局。**導數告訴你的是函數在一點附近的表現。要得出全局的結論，還需要更多資料：定義域兩端的情況，或者函數表現異常之處的情況。

背後的數學

大數學家也會出錯

分析學的奠基者也犯過一模一樣的錯誤。柯西（Augustin-Louis Cauchy）在 1821 年的 *Cours d'analyse* 中「證明」了連續函數的收斂（convergent）和必定連續。1826 年，阿貝爾（Niels Henrik Abel）指出了一個反例（counterexample）：傅立葉級數（Fourier series）

$$\sin x - \tfrac{1}{2}\sin 2x + \tfrac{1}{3}\sin 3x - \cdots$$

對每一個 x 都收斂，而且每一項都連續，但它的和在 $x = \pm\pi$ 處出現跳躍（當 $-\pi < x < \pi$ 時，它等於 $x/2$）。要弄清楚柯西的論證究竟錯在哪裏，花了數十年時間，並由此產生了一致收斂（uniform convergence）的概念，如今每個修讀分析的學生都要學。隱藏的假設，尤其是關於交換兩個極限次序的假設，是整個分析學中最常見的錯誤來源。

帶有前提的定理

這一題中有兩個錯誤，都源於忽視了定理的前提（hypotheses）。

- 極值定理（extreme value theorem）指出，在有界閉區間 $[a,b]$ 上的連續函數必定取得最大值和最小值。要找出它們，須檢查端點、駐點，以及函數不可導的點：$|x|$ 在 $[-1,1]$ 上的最小值在 0 處取得，而這一點既非端點，亦非駐點。如果 f 在區間內部處處可導，則位於內部的最大點或最小點必定是駐點。在整條實數線上，沒有端點可供檢查，函數甚至可能根本沒有最大值；你必須改為考察 $x \to \pm\infty$ 時的表現。

- *微積分基本定理要求被積函數在 $[a,b]$ 上連續。如果不連續，積分可能仍然存在（試試 $\int_0^1 x^{-1/2}\,\mathrm{d}x$），也可能不存在（如本題），你必須用極限來檢查。*

在大學裏，你遇到的每一條定理都附有一張精確的前提清單。訓練的一部分，就是每次使用定理之前，都要把這些前提檢查一遍。

發散積分的一個值

第 (ii) 部分還有一個耐人尋味的註腳。區間 $[-1,3]$ 關於出問題的點 $x=1$ 對稱，而 $(1-x)^{-3}$ 關於該點是「奇」的。如果你挖去一個對稱的缺口 $(1-\varepsilon,1+\varepsilon)$，再令 $\varepsilon\to 0$，兩塊無窮大的部分便會互相抵消，你確實會得到 0。這稱為*柯西主值*（Cauchy principal value），在物理學和複分析中確實很有用。但它是一個刻意選取的*不同定義*，一個誤打誤撞得出這個值的論證，甚麼也沒有證明。

思維技巧

以批判的眼光閱讀論證

每一步都要檢查三件事：計算對不對，這一步在邏輯上是否成立，以及有甚麼是沒有明說卻已假設了的？

然後另行檢查結論：不正確的論證可以有正確的結論，而「正確」的答案也可以出自無效的論證。對自己的作業也要同樣嚴格審視；你最容易相信的論證，正是你自己的論證。

延伸練習

1. **會說謊的代換。**試做 STEP I 1987，第 8 題（`step.maths.org/questions/87-s1-q8`）。一個代換（substitution）似乎顯示某個正的積分等於一個負的積分。準確找出論證在哪裏出錯，然後正確地計算兩個積分。

2. **是對還是錯？**試做 STEP I 1998，第 3 題（`step.maths.org/questions/98-s1-q3`）：有四個斷言要你證明或否證。對每一個斷言，動筆之前先判斷你預期是哪一種。

3. **所有馬都是同一種顏色。**找出以下用歸納法（induction）寫成的「證明」的漏洞：任何一匹馬都與自己顏色相同；給定 $n+1$ 匹馬，由歸納假設，前 n 匹顏色相同，後 n 匹也是；兩組有重疊，所以全部 $n+1$ 匹顏色相同。

4 **一個確實存在的積分。**證明 $\int_{-1}^{1} |x|^{-1/2}\,\mathrm{d}x$ 存在，儘管被積函數在 0 處沒有定義，並求出它的值。

延伸閱讀

- **E. A. Maxwell, *Fallacies in Mathematics***（Cambridge University Press）。一本經典的謬誤論證集，收錄各種似是而非的錯誤論證，按課題編排，為高中第六學級（sixth form）學生而寫。
- **Bernard Gelbaum & John Olmsted, *Counterexamples in Analysis***（Dover）。數以百計的例子，說明定理的前提為何重要。開始修讀分析課程之後，不妨隨手翻閱。
- **Imre Lakatos, *Proofs and Refutations***（Cambridge University Press）。一篇精彩的對話，探討數學家如何發現並修正證明中的錯誤。

3.2 倒過來的 AM–GM

題目

STEP II 1993，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/93-s2-q8`

引子

算術平均—幾何平均不等式（arithmetic mean–geometric mean inequality，簡稱 AM–GM）也許是數學中最有用的一條不等式。它說的是：對於正數，普通的平均數總是不小於「乘法意義下的平均數」，而且只有當所有數都相等時，等號（equality）才成立。它不用微積分就能解決最優化（optimisation）問題，可以用來證明其他不等式，並且在分析學（analysis）的每一個分支中都會出現。

這道題用一種奇特而優美的歸納法（induction）來證明它，這個方法由柯西（Augustin-Louis Cauchy）於 1821 年提出。它不是由 n 走到 $n+1$，而是由 2 躍到 4，再到 8、16，只涵蓋 2 的冪，然後掉頭往下走，把中間的數補上。它提醒我們，「歸納法」這個想法遠比課本上的版本靈活。

動筆之前

網上試題庫所載的這份試卷有一處誤植：它寫的是「$b_1 = a_2$」。那當然應該是 $b_1 = a_1$；意思是 b 這一列數就是 a 那一列數，再用它們的平均數的若干個副本補足。

? **想一想** 為甚麼由 2^m 個數走到 2^{m+1} 個數，可能比由 n 走到 $n+1$ 容易？

因為 2^{m+1} 個數可以自然地分成兩組，每組 2^m 個；你可以對每一半應用歸納假設（inductive hypothesis），然後用兩個數的情形把兩半合起來。由 n 走到 $n+1$，就沒有這樣對稱的結構可供利用。

逐步解題

階段 1　兩個數

$$\left(\frac{a_1+a_2}{2}\right)^2 - a_1a_2 = \frac{a_1^2+2a_1a_2+a_2^2-4a_1a_2}{4} = \frac{(a_1-a_2)^2}{4} \geq 0. \qquad \blacksquare$$

階段 2　加倍

$m = 1$ 的情形就是第 1 階段。假設 $(*)$ 對 2^m 個數成立，現取 2^{m+1} 個正數。

> **想一想**　把這 2^{m+1} 個數分成前半和後半，設 P 和 Q 分別為兩半的算術平均。歸納假設對每一半說了甚麼？

對每一半應用 $(*)$，

$$a_1 \cdots a_{2^m} \leq P^{2^m} \qquad \text{及} \qquad a_{2^m+1} \cdots a_{2^{m+1}} \leq Q^{2^m}.$$

兩式相乘（所有數都是正數），全部 2^{m+1} 個數的積不大於 $(PQ)^{2^m}$。現在由兩個數的情形得 $PQ \leq \left(\frac{P+Q}{2}\right)^2$，而 $\frac{P+Q}{2}$ 正好是全部 2^{m+1} 個數的平均數。所以

$$a_1 \cdots a_{2^{m+1}} \leq (PQ)^{2^m} \leq \left(\frac{P+Q}{2}\right)^{2^{m+1}} = \left(\frac{a_1 + \cdots + a_{2^{m+1}}}{2^{m+1}}\right)^{2^{m+1}},$$

這就是 $m+1$ 時的 $(*)$。由歸納法可知，只要項數是 2 的冪，AM–GM 就成立。 $\blacksquare$

階段 3　填補空隙

現在取任意的 n，並選取 m 使得 $2^m > n$。我們在數列 $a_1, \ldots, a_n$ 之後補上 $2^m - n$ 個它自己的平均數 A。

> **想一想**　補足後的數列 $b_1, \ldots, b_{2^m}$ 的平均數是多少？

其總和為 $(a_1 + \cdots + a_n) + (2^m - n)A = nA + (2^m - n)A = 2^m A$，所以平均數就是 A 本身。加入平均數的副本並不會改變平均數。因此，對 b 這一列數應用 $(*)$，得

$$a_1 \cdots a_n \cdot A^{2^m - n} \leq A^{2^m}.$$

> **停一停，試一試**　完成這個證明。

兩邊同除以 A^{2^m-n}（這是正數）：$a_1 \cdots a_n \leq A^n$。取 n 次方根，

$$(a_1 \cdots a_n)^{1/n} \leq A = \frac{a_1 + \cdots + a_n}{n}. \qquad \blacksquare$$

回顧

- **沿另一條路徑作歸納。**第 2 階段是對 m 作歸納，涵蓋 $n = 2, 4, 8, \ldots$；第 3 階段則是由任何一個 2 的冪往下走到任何較小的 n 的方法（圖 3.2.1）。
- **用平均數補足。**把數列加長而不改變其平均數，這個技巧值得記住；它能把關於任意 n 的問題，化為關於一個方便的 n 的問題。
- **等號何時成立？**沿着證明往回追溯，等號成立需要每一次使用兩個數的情形時等號都成立，而這只有在所有數都相等時才會發生。

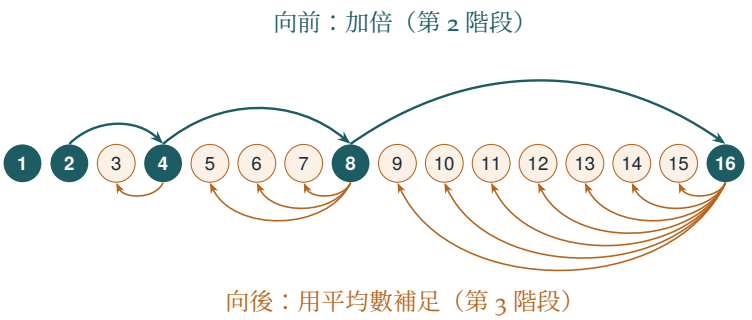


圖 3.2.1　柯西的向前—向後歸納法。藍綠色箭嘴以加倍的方式，證明 AM–GM 對 $2, 4, 8, 16, \ldots$ 個數成立；琥珀色箭嘴則把數列補足至下一個 2 的冪，從而填上其餘每一個 n。

背後的數學

柯西的歸納法

柯西在他的《*Cours d'analyse*》（1821）中發表了這個證明；我們在第 2 章談過，正是這本書開創了分析學的嚴格處理。這種技巧現在稱為**向前—向後歸納法**（forward–backward induction）：先對無窮多個情形組成的一個序列（這裏是 2 的冪）證明命題成立，然後證明 n 的情形蘊含 $n-1$ 的情形。兩者合起來，就能到達每一個 n。我們的第 3 階段是由 2^m 直接跳到 n，但你也可以一步一步來：在 $n-1$ 個數之後補上一個它們的平均數，就能證明 n 的情形蘊含 $n-1$ 的情形。

一條不等式，多種證明

AM–GM 有數十種證明，把它們互相比較，是認識不同數學風格的好方法。

- **利用凸性（convexity）。**對數（logarithm）函數是凹（concave）函數，所以對數的平均不大於平均的對數。取指數之後，這就是 AM–GM；這也是 Section 3.3 的主題。
- **波利亞（George Pólya）的證明。**利用 $e^{x-1} \geq x$（e^{x-1} 在 $x=1$ 處的切線位於曲線下方）。設 A 為平均數，把這條不等式應用於每一個 a_i/A，然後相乘：$\prod \frac{a_i}{A} \leq e^{\sum(a_i/A-1)} = e^0 = 1$。波利亞說這個證明是他在夢中想到的。
- **利用磨光法（smoothing）。**如果其中兩個數不相等，就把它們都換成兩者的平均數。算術平均不變，而由兩個數的情形可知，乘積會增大。重複這個步驟，所有數就會被推向平均數。

不用微積分的最優化

AM–GM 能把許多最優化問題變成一行就能解決的問題。在表面積固定的所有長方體盒子當中，哪一個體積最大？在周界固定的所有長方形當中，哪一個面積最大？（是正方形，由兩個數的情形可知：$xy \leq \left(\frac{x+y}{2}\right)^2$，等號當 $x=y$ 時成立。）等號成立的條件告訴你最優值在哪裏出現，而不等式則告訴你那確實是最優值，而不僅僅是一個駐點（stationary point）。可與 Section 3.1 中的告誡比較。

思維技巧

靈活的歸納法

歸納法不一定要按 $1, 2, 3, \ldots$ 的次序進行。只要一條路徑能到達每一個情形，你就可以沿着它作歸納：可以加倍，可以用強歸納法（*strong induction*），也可以先在一個子數列（*subsequence*）上向前走，再向後走以填補空隙。

當由 n 到 $n+1$ 這顯而易見的一步難以走通時，不妨問問哪一步會容易走，以及能否靠這一步到達每一個情形。

延伸練習

1. **後退一步。**補上一個平均數的副本，直接證明：若 AM–GM 對 n 個數成立，則它對 $n-1$ 個數也成立。
2. **一個遞增數列。**把 AM–GM 應用於 $1+\frac{1}{n}, \ldots, 1+\frac{1}{n}, 1$ 這 $n+1$ 個數（其中 $1+\frac{1}{n}$ 有 n 個），證明 $\left(1+\frac{1}{n}\right)^n < \left(1+\frac{1}{n+1}\right)^{n+1}$。

3 **最好的盒子。**一個密封的長方體盒子的表面積為 S。把 AM–GM 用於三個不同的面的面積，證明其體積不大於 $(S/6)^{3/2}$，等號在盒子為正方體時成立。

4 **調和平均（harmonic mean）。**把 AM–GM 應用於 $1/a_i$ 這些數，證明調和平均 $\dfrac{n}{1/a_1 + \cdots + 1/a_n}$ 不大於幾何平均。

延伸閱讀

- **J. Michael Steele, *The Cauchy–Schwarz Master Class***（Cambridge University Press）。一本以問題帶動、講不等式的出色著作，正好由這個程度起步。第 2 章講的就是 AM–GM。
- **Hardy, Littlewood & Pólya, *Inequalities***（Cambridge University Press）。經典參考書，由二十世紀三位偉大的分析學家撰寫。
- **Arthur Engel, *Problem-Solving Strategies***（Springer）。第 7 章收錄了大量來自數學奧林匹克的不等式題目。

3.3 凸性的威力

題目

STEP II 2007，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/07-s2-q7`

引子

在 Section 3.2 中，我們用一個巧妙的歸納法（induction）證明了 AM–GM 不等式。在這一節，它只需一行便水到渠成，因為它不過是某個更宏大的結果的特殊情形。這個更宏大的東西就是**凹性**（concavity）：向下彎的曲線所具有的形狀，就像山頂一樣。關於這類曲線，有一個簡單的幾何事實：曲線位於它的弦（chord）之上。單憑這一點，竟然可以證明範圍廣得驚人的各種不等式（inequality），從關於三角形的一個結果，到 AM–GM 不等式，以至概率（probability）、信息論（information theory）和經濟學中的不等式。

到了大學，這個想法會發展成**凸性**（convexity），它是現代數學中最能貫通各個領域的重要主題之一。有些學科，例如凸優化（convex optimisation），整個就是建基於它之上。

動筆之前

想一想 畫一條向下彎的曲線，例如 $(0, \pi)$ 上的 $y = \sin x$。在曲線上標出兩點，並用直線把它們連起來。這條弦在曲線之上還是之下？

在曲線之下。對於凹曲線，每一條弦在兩個端點之間都位於曲線之下。兩個函數值的平均，就是弦的中點的高度；而函數在平均值處的取值，就是該中點正上方的曲線高度。因此，「弦在曲線之下」正正就是 $n = 2$ 時的琴生不等式（Jensen's inequality）：

$$\frac{f(x_1) + f(x_2)}{2} \leq f\left(\frac{x_1 + x_2}{2}\right).$$

若有 n 個點，則各點 $(x_k, f(x_k))$ 的平均就是它們的質心（centre of mass），它位於曲線下方的區域之內，所以同樣在曲線之下（圖 3.3.1）。

逐步解題

階段 1　驗證凹性

當 $0 < x < \pi$ 時，$\frac{\mathrm{d}^2}{\mathrm{d}x^2}\sin x = -\sin x < 0$；當 $x > 0$ 時，$\frac{\mathrm{d}^2}{\mathrm{d}x^2}\ln x = -\frac{1}{x^2} < 0$。 ■

階段 2　三角形

想一想　應該用哪個函數、取哪些點？它們的平均是甚麼？三角形的各個角是否都落在函數呈凹性的區間之內？

取 $\mathrm{f}(x) = \sin x$ 和三個點 A, B, C。每個角都嚴格介乎 0 與 π 之間，而 sin 在這個區間上是凹的，它們的平均是 $\frac{A+B+C}{3} = \frac{\pi}{3}$。由琴生不等式得

$$\frac{\sin A + \sin B + \sin C}{3} \le \sin\frac{\pi}{3} = \frac{\sqrt{3}}{2},$$

所以 $\sin A + \sin B + \sin C \le \frac{3\sqrt{3}}{2}$，而等號恰好在等邊三角形（equilateral triangle）時成立。 ■

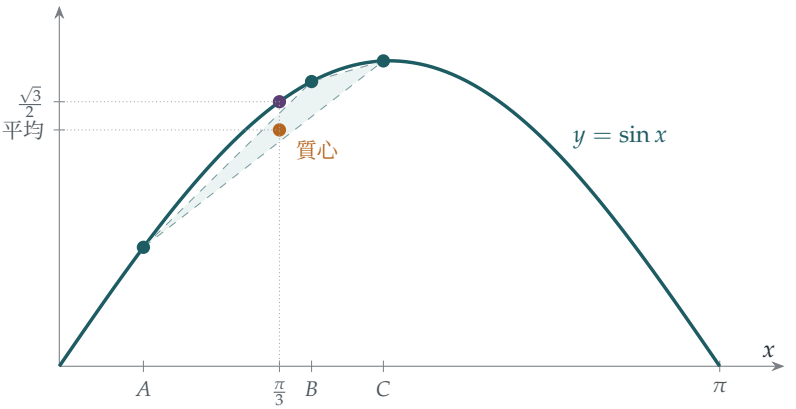


圖 3.3.1　以三角形的三個角 $A = 0.4$、$B = 1.2$、$C = \pi - 1.6$ 說明 $\sin x$ 的琴生不等式。曲線上三點的質心（琥珀色）位於陰影三角形之內，因而在曲線之下。它的高度是 $\sin A$、$\sin B$、$\sin C$ 的平均，而它正上方的曲線高度是 $\sin\frac{\pi}{3}$。

階段 3　一行證明 AM–GM

想一想　AM–GM 不等式比較的是一個積與一個和。哪個凹函數能把積變成和？

對數（logarithm）。把琴生不等式應用於 $\mathrm{f} = \ln$ 和正數 $t_1, \ldots, t_n$：

$$\frac{\ln t_1 + \cdots + \ln t_n}{n} \le \ln\left(\frac{t_1 + \cdots + t_n}{n}\right).$$

左邊是 $\ln\left(\sqrt[n]{t_1 \cdots t_n}\right)$。由於 ln 是遞增函數，我們可以把兩邊的對數同時去掉，這就是 AM–GM 不等式。等號恰好在所有 t_k 都相等時成立。 ■

階段 4 運用 AM–GM

想一想 在 (a) 中，左邊是四項之和，右邊含有三個變量的積。你應該對多少項應用 AM–GM 不等式？那個 16 又為何有用？

對四個正數 x^4、y^4、z^4、16 應用 AM–GM 不等式：

$$\frac{x^4 + y^4 + z^4 + 16}{4} \ge \sqrt[4]{16\,x^4y^4z^4} = 2xyz.$$

兩邊乘以 4，得 $x^4 + y^4 + z^4 + 16 \ge 8xyz$。 ■

停一停，試一試 現在求 $x^5 + y^5 + z^5 - 5xyz$ 的最小值。（提示：你想要的是 $x^5y^5z^5$ 的五次方根，所以需要五個數。另外兩個數應該是甚麼？）

對 $x^5, y^5, z^5, 1, 1$ 應用 AM–GM 不等式：

$$\frac{x^5 + y^5 + z^5 + 1 + 1}{5} \ge \sqrt[5]{x^5y^5z^5} = xyz,$$

所以 $x^5 + y^5 + z^5 - 5xyz \ge -2$。等號在 $x^5 = y^5 = z^5 = 1$ 時成立，即 $x = y = z = 1$，這時該式的值為 $3 - 5 = -2$。因此最小值是 -2。 ■

留意等號成立的條件在這裏身兼兩職：它說明這個界確實可以取到，所以 -2 是最小值，而不僅僅是一個下界（lower bound）。

回顧

- **一個想法，多個不等式。**關於凹曲線的一個幾何事實，既給出了一個關於三角形的結果，又給出了 AM–GM 不等式，而所需的計算不過是驗證一個二階導數（second derivative）。
- **選擇函數和點。**運用琴生不等式，就是要決定把它應用於哪個函數、取哪些點。對數能把積變成和，是常用之選。
- **補項以配合題目。**在 (b) 中，補上兩個 1 令 AM–GM 不等式切合題目，正如在 Section 3.2 中補上平均值一樣。

背後的數學

凸函數

在大學裏，凸函數（convex function，即把凹函數上下倒轉）的標準定義完全不涉及導數：若

$$\mathrm{f}\big(tx_1 + (1-t)x_2\big) \le t\,\mathrm{f}(x_1) + (1-t)\,\mathrm{f}(x_2) \qquad \text{對所有} x_1, x_2 \text{及} 0 \le t \le 1,$$

則稱 f 為凸函數，這正正是說弦位於曲線之上。這個定義對不可微（not differentiable）的函數（例如 $|x|$）以及多元函數同樣適用。丹麥工程師兼數學家琴生（Johan Jensen）在 1906 年提出了以他命名的不等式，使凸函數成為一門獨立的學問。

無處不在的凸性

凸性在數學及其應用中隨處可見。

- **概率。**琴生不等式指出，對於凹函數 f 和任意隨機變量（random variable）X，有 $\mathrm{E}[\mathrm{f}(X)] \le \mathrm{f}(\mathrm{E}[X])$。在經濟學中，若以 f 為一個凹的「效用」（utility）函數，這就解釋了為何人們寧願穩拿 £100，也不願參與一場各有一半機會得到 £0 或 £200 的賭局，亦解釋了為何會有保險。
- **信息論。**均勻分佈（uniform distribution）使熵（entropy）達到最大，這個結果是數據壓縮的基礎，而它的證明就是 ln 的琴生不等式。
- **優化。**凸函數除了全局最小值之外，再沒有其他局部極小值（local minima），所以「一直往下坡走」總能找到最佳答案。這就是凸優化問題能夠有效率地求解的原因，也是人們花那麼多工夫把問題改寫成凸形式的原因。

在 Section 3.4 中，我們會看到凸性的另一面：函數與其反函數（inverse）之間的一種幾何對偶（duality）。

思維技巧

> **找出一般原理**
>
> 當幾個結果的證明相似時，要找出它們背後的同一個原理。然後把這個原理證明一次，再多次應用。

AM–GM 不等式、關於正弦的三角形不等式，以及許多其他結果，全都是「凹曲線位於它的弦之上」的實例。認出這個原理，就能把一堆零散的技巧變成一套方法。

延伸練習

1 **定義。**STEP II 2018，第 2 題（`step.maths.org/questions/18-s2-q2`）的第 (i) 及 (ii) 部分由上面方框中以弦刻劃凹性的定義出發，逐步建立三個點的琴生不等式。

2 **正弦之積。**試做 STEP II 2018，第 2 題（`step.maths.org/questions/18-s2-q2`）的第 (iii) 部分。

3 **銳角三角形。**證明 $\tan x$ 在 $(0, \frac{\pi}{2})$ 上是凸的，並由此推出在銳角三角形(acute-angled triangle）中，$\tan A + \tan B + \tan C \geq 3\sqrt{3}$。

4 **冪平均。**證明 x^2 是凸的，並由此推出 $\left(\frac{a_1+\cdots+a_n}{n}\right)^2 \leq \frac{a_1^2+\cdots+a_n^2}{n}$。試把這個結果詮釋為關於一組數的方差（variance）的命題。

延伸閱讀

- **J. Michael Steele, *The Cauchy–Schwarz Master Class***（Cambridge University Press）。第 6 章講述凸性和琴生不等式。
- **Claudi Alsina & Roger Nelsen, *When Less Is More: Visualizing Basic Inequalities***（Mathematical Association of America）。以圖形證明不等式，當中不少用到凸性。
- **Stephen Boyd & Lieven Vandenberghe, *Convex Optimization***（Cambridge University Press）。一部現代經典，可在網上免費閱讀；開首幾章說明凸性在實際應用中為何重要。

3.4 以面積證明不等式

題目

STEP III 2011，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/11-s3-q4`

引子

這裏有一個不等式（inequality），證明它的圖簡單得可以畫在一枚郵票上。畫一條通過原點的遞增曲線。曲線下方的面積，加上曲線左方的面積，總能蓋住一個矩形。單憑這一個觀察，就得出楊氏不等式（Young's inequality）$ab \le \frac{a^p}{p} + \frac{b^q}{q}$，而它又是通往赫爾德不等式（Hölder's inequality）和柯西–施瓦茲不等式（Cauchy–Schwarz inequality）的鑰匙，這兩個不等式是分析學中最重要的工具之一。

這幅圖還蘊藏着一個更深刻的想法。它以完全對稱的方式把一個函數與它的反函數（inverse function）配成一對；把這種配對加以推廣，就是勒讓德變換（Legendre transform），它把古典力學的兩大表述方式聯繫起來，也把熱力學的兩種主要描述聯繫起來。

動筆之前

想一想 $\int_0^a \mathrm{f}(x)\,\mathrm{d}x$ 是曲線與 x 軸之間的面積。在同一幅圖上，$\int_0^b \mathrm{f}^{-1}(y)\,\mathrm{d}y$ 代表哪一塊面積？

$x = \mathrm{f}^{-1}(y)$ 的圖像與 $y = \mathrm{f}(x)$ 是同一條曲線，只是橫着看而已。所以 $\int_0^b \mathrm{f}^{-1}(y)\,\mathrm{d}y$ 是曲線與 y 軸之間、由高度 0 到高度 b 的面積。明白了這一點，第 (i) 部分只需要畫圖。

逐步解題

階段 1　圖形

想一想　畫出以 $(0,0)$ 和 (a,b) 為頂點的矩形。用一種顏色塗上曲線下方由 0 到 a 的區域，再用另一種顏色塗上曲線左方由 0 到 b 的區域。這兩個區域能蓋住整個矩形嗎？甚麼時候既沒有重疊，也沒有多出來的部分？

這兩個區域合起來蓋住整個矩形 $[0,a]\times[0,b]$，通常還蓋住矩形以外的一些面積（圖 3.4.1）。矩形內的每一點 (x,y)，不是在曲線下方（因為 $x \le a$，所以屬於第一個區域），就是在曲線左方（因為 $y \le b$，所以屬於第二個區域）。因此塗色部分的總面積至少等於矩形的面積 ab。這就是 $(*)$。

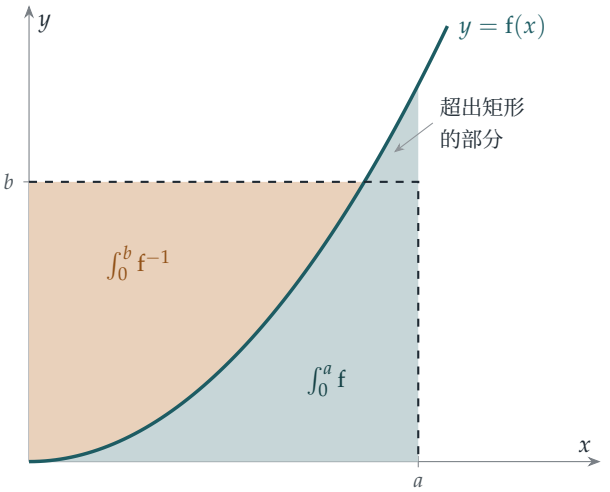


圖 3.4.1　用一幅圖看楊氏不等式。曲線下方由 0 到 a 的面積（藍綠色）與曲線左方由 0 到 b 的面積（琥珀色）合起來蓋住面積為 ab 的虛線矩形，而且還有剩餘。恰好當頂點 (a,b) 位於曲線上時，才沒有多出來的部分。

多出來的部分，就是曲線與矩形頂點之間那一小塊曲邊區域。它恰好在曲線通過頂點 (a,b) 時消失，也就是當

$$b = f(a).$$

階段 2　楊氏不等式

想一想　取 $f(x) = x^{p-1}$，$f^{-1}(y)$ 是甚麼？計算兩個積分，並用 $\frac{1}{p}+\frac{1}{q}=1$ 化簡第二個。

由於 $p>1$，$\mathrm{f}(x)=x^{p-1}$ 連續而且遞增，並且 $\mathrm{f}(0)=0$，而 $\mathrm{f}^{-1}(y)=y^{1/(p-1)}$。於是

$$\int_0^a x^{p-1}\,\mathrm{d}x=\frac{a^p}{p},\qquad \int_0^b y^{1/(p-1)}\,\mathrm{d}y=\frac{b^{1+1/(p-1)}}{1+1/(p-1)}.$$

由 $\frac{1}{p}+\frac{1}{q}=1$ 得 $q=\frac{p}{p-1}$，所以 $1+\frac{1}{p-1}=\frac{p}{p-1}=q$，第二個積分就是 $\frac{b^q}{q}$。因此 $(*)$ 變成

$$ab\le\frac{a^p}{p}+\frac{b^q}{q}.\qquad\blacksquare$$

等號應在 $b=\mathrm{f}(a)=a^{p-1}$ 時成立。驗證：這時 $ab=a^p$，而 $b^q=a^{(p-1)q}=a^p$，所以右邊是 $a^p\left(\frac{1}{p}+\frac{1}{q}\right)=a^p$。$\blacksquare$

取 $p=q=2$，這就是 $ab\le\frac{a^2+b^2}{2}$，其實是 Section 3.2 中兩個數的算術–幾何平均不等式（AM–GM inequality）換了個樣子。

階段 3　反正弦

現在取 $[0,\frac{\pi}{2}]$ 上的 $\mathrm{f}(x)=\sin x$，它連續而且遞增，並且 $\mathrm{f}(0)=0$，而在 $[0,1]$ 上 $\mathrm{f}^{-1}=\arcsin$。

想一想　$\int_0^a \sin x\,\mathrm{d}x$ 很容易。你會怎樣求 $\arcsin y$ 的積分？

$\int_0^a \sin x\,\mathrm{d}x=1-\cos a$。至於反正弦，把它看成 $\arcsin y\times 1$，用分部積分法（integration by parts）：

$$\int_0^b \arcsin y\,\mathrm{d}y=\Big[y\arcsin y\Big]_0^b-\int_0^b\frac{y}{\sqrt{1-y^2}}\,\mathrm{d}y=b\arcsin b+\sqrt{1-b^2}-1.$$

兩式相加，由 $(*)$ 得 $ab\le b\arcsin b+\sqrt{1-b^2}-\cos a$。$\blacksquare$

停一停，試一試　現在推導關於 $\arcsin(t^{-1})$ 的不等式。應該選哪一個 a 的值？（試試最簡單的那一個。）

取 $a=0$：這時 $0\le b\arcsin b+\sqrt{1-b^2}-1$，即 $b\arcsin b\ge 1-\sqrt{1-b^2}$。（這不過是說 $\int_0^b\arcsin y\,\mathrm{d}y\ge 0$。）設 $b=1/t$，其中 $t\ge 1$，於是 $0<b\le 1$，再除以 b：

$$\arcsin\left(\frac{1}{t}\right)\ge t\left(1-\sqrt{1-\frac{1}{t^2}}\right)=t-\sqrt{t^2-1}.\qquad\blacksquare$$

回顧

- **以圖為證。**只要看出合適的圖，第 (i) 部分就是一個完整的證明。把幾何事實轉化為積分之間的不等式，是一種威力強大的技巧。

- **把一般結果特殊化。**同一個一般的不等式，選取不同的 f，就得出楊氏不等式和一個關於 arcsin 的不等式。
- **最簡單的選擇可能正是對的選擇。**最後的推導用了 $a = 0$，這時不等式幾乎是顯然的。簡單的情況一定要試試。

背後的數學

從楊氏到赫爾德

楊（William Henry Young）在 1912 年發表他的不等式。它之所以重要，在於它所引出的結果。假設 $\sum a_i^p = 1$ 及 $\sum b_i^q = 1$。對每一對數應用楊氏不等式，然後相加，

$$\sum a_i b_i \le \frac{\sum a_i^p}{p} + \frac{\sum b_i^q}{q} = \frac{1}{p} + \frac{1}{q} = 1.$$

把各個 a 和各個 b 按比例縮放，使兩個和都等於 1，就得到**赫爾德不等式**：對正數，

$$\sum a_i b_i \le \Big(\sum a_i^p\Big)^{1/p} \Big(\sum b_i^q\Big)^{1/q}.$$

取 $p = q = 2$，這就是**柯西—施瓦茲不等式** $\sum a_i b_i \le \sqrt{\sum a_i^2}\sqrt{\sum b_i^2}$，它說的是兩個向量夾角的餘弦最多為 1。赫爾德不等式是 L^p 空間理論的基礎；在這些函數空間裏，「距離」是用 $\big(\int |f|^p\big)^{1/p}$ 來量度的，你會在泛函分析（functional analysis）和概率的課程中遇到它們。

勒讓德變換

圖 3.4.1 對 f 和 f^{-1} 的處理完全對稱。記 $F(a) = \int_0^a \mathrm{f}$ 及 $G(b) = \int_0^b \mathrm{f}^{-1}$。不等式說的是 $F(a) + G(b) \ge ab$，在曲線上等號成立，所以

$$G(b) = \max_a \big(ab - F(a)\big).$$

這個把凸函數（convex function）F 變成另一個凸函數 G 的運算，就是**勒讓德變換**，而連續做兩次就會回到 F。在力學中，它把**拉格朗日量**（Lagrangian，位置與速度的函數）轉換成**哈密頓量**（Hamiltonian，位置與動量的函數）：這是牛頓（Newton）定律的兩種等價表述，你會在大學二年級的課程中學到。在熱力學中，它在能量、焓（enthalpy）和自由能（free energy）之間互相轉換。對 $F(a) = a^p/p$，變換的結果是 $G(b) = b^q/q$，這就是楊氏不等式中兩個指數由 $\frac{1}{p} + \frac{1}{q} = 1$ 聯繫起來的原因。

思維技巧

從圖形得出不等式

積分就是面積。在一幅精心挑選的圖中比較面積，可以證明一個從代數上難以看出的不等式。

留意這樣的圖：其中一個區域明顯包含另一個區域，或者幾個區域明顯蓋住第三個區域。然後把每塊面積轉化為積分。

延伸練習

1 **1987 年的版本。**試做 STEP I 1987，第 6 題（`step.maths.org/questions/87-s1-q6`）。它用同一類圖來求反函數的積分。用求導來驗算你的答案。

2 **柯西–施瓦茲。**用 $p = q = 2$ 的楊氏不等式，證明對正數有 $a_1b_1 + a_2b_2 + a_3b_3 \leq \sqrt{a_1^2 + a_2^2 + a_3^2}\sqrt{b_1^2 + b_2^2 + b_3^2}$。（提示：先處理兩個平方根都等於 1 的情況。）

3 **指數版本。**在 $(*)$ 中取 $\mathrm{f}(x) = e^x - 1$，證明對 $a, b \geq 0$，

$$ab \leq e^a - 1 - a + (1 + b)\ln(1 + b) - b.$$

等號何時成立？

4 **這個界有多好？**較難用計算機比較 $t = 1, 2, 10$ 時的 $\arcsin(1/t)$ 與 $t - \sqrt{t^2 - 1}$。證明當 t 很大時，$t - \sqrt{t^2 - 1} \approx \frac{1}{2t}$，而 $\arcsin(1/t) \approx \frac{1}{t}$。為甚麼這個界相差了一個因子 2？（看看 $a = 0$ 時的圖：多出來的部分有多少被浪費了？）

延伸閱讀

- **J. Michael Steele, *The Cauchy–Schwarz Master Class***（Cambridge University Press）。第 1 章和第 9 章講述柯西–施瓦茲不等式和赫爾德不等式，途中也談到楊氏不等式。
- **Roger Nelsen, *Proofs Without Words***（Mathematical Association of America）。收錄了楊氏不等式和許多其他的圖形證明。
- **V. I. Arnold, *Mathematical Methods of Classical Mechanics***（Springer）。第 14 節正是用這幅圖來解釋勒讓德變換。內容較深，但相關的一節很短。

3.5 π^e 還是 e^π？

題目

STEP II 2013，第 1 題

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

`step.maths.org/questions/13-s2-q1`

引子

π^e 和 e^π，哪一個較大？計算機一秒鐘就能告訴你：$e^\pi \approx 23.14$ 而 $\pi^e \approx 22.46$。但計算機只告訴你結果是這樣，卻沒有告訴你爲甚麼，更無法一次過解答無窮多條同類的問題。3^π 比 π^3 大嗎？除了顯而易見的 $n = m$ 之外，還有沒有整數滿足 $n^m = m^n$？

原來這些問題全都關乎同一條曲線 $y = \ln x$，以及與它相交的直線。本題說明，只要運用得當，一幅草圖就能解決一些單靠代數看似無從入手的大小比較，也說明 e 這個數怎樣自然而然地成為分界點。

動筆之前

想一想 取對數（logarithm）後，$a^b = b^a$ 等同於 $b \ln a = a \ln b$。試把這條式重新整理，使 a 只出現在一邊，b 只出現在另一邊。

$b \ln a = a \ln b$ 等同於

$$\frac{\ln a}{a} = \frac{\ln b}{b}.$$

因此，每次比較 a^b 與 b^a，其實都是比較同一個函數 $h(x) = \frac{\ln x}{x}$ 在 a 和 b 兩處的值。而 $\frac{\ln a}{a}$ 正是由原點（origin）連到曲線 $y = \ln x$ 上的點 $(a, \ln a)$ 的直線的斜率（gradient）。這就是本題與直線有關的原因。

逐步解題

階段 1　通過原點的切線

? **想一想**　若 $y = mx$ 與 $y = \ln x$ 在 $x = x_0$ 處相切，哪兩條方程必須成立？

直線與曲線必須相交，即 $mx_0 = \ln x_0$，而且斜率相同，即 $m = 1/x_0$。把第二式代入第一式，得 $1 = \ln x_0$，所以 $x_0 = e$ 及 $m = 1/e$。由原點引出的切線（tangent）與曲線相切於 $(e, 1)$。

若 $y = mx$ 改為與曲線相交於 $x = a$ 和 $x = b$，則 $\ln a = ma$ 及 $\ln b = mb$，所以 $\frac{\ln a}{a} = m = \frac{\ln b}{b}$，由此得 $b \ln a = a \ln b$，因而 $a^b = b^a$。 ■

至於草圖：一條通過原點、比切線稍為平坦的直線會與曲線相交兩次，切點兩側各一次，所以 $a < e < b$。斜率大於 $1/e$ 的直線則完全碰不到曲線，因為 $\ln x$ 是凹（concave）函數，整條曲線都在其切線 $y = x/e$ 之下。

階段 2　與 y 軸相交於原點上方的直線

? **想一想**　現在直線是 $y = mx + c$，其中 $c > 0$。試以 m 和 c 表示 $\frac{\ln p}{p}$ 和 $\frac{\ln q}{q}$。

由 $\ln p = mp + c$ 及 $\ln q = mq + c$，得

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

由於 $c > 0$ 且 $p < q$，可知 $\frac{c}{p} > \frac{c}{q}$，所以 $\frac{\ln p}{p} > \frac{\ln q}{q}$。兩邊乘以 $pq > 0$，得 $q \ln p > p \ln q$，因此 $p^q > q^p$。 ■

從幾何角度看：若通過 $(p, \ln p)$ 和 $(q, \ln q)$ 兩點的直線與 y 軸相交於原點上方，那麼從原點望去，前一點的仰角比後一點的更陡。

階段 3　e 右方的弦

? **想一想**　設 $e \le p < q$。由 $(p, \ln p)$ 連到 $(q, \ln q)$ 的弦（chord），其斜率小於曲線在 p 處的斜率。為甚麼？又為甚麼這樣會令它的 y 截距（intercept）為正數？

因為 $\ln x$ 是凹函數，它的斜率 $1/x$ 遞減，所以由 p 到 q 的弦比 p 處的切線平坦：弦的斜率 s 滿足 $s < \frac{1}{p}$。弦的 y 截距是

$$c = \ln p - sp > \ln p - \frac{1}{p} \cdot p = \ln p - 1 \ge 0,$$

當中用到 $p \geq e$。所以這條弦與 y 軸相交於原點上方。 ■

> **停一停，試一試** 利用這個結果和第 2 階段，判斷 π^e 與 e^π 哪一個較大。

取 $p = e$ 及 $q = \pi$。連接這兩點的弦的截距為正數，所以根據第 2 階段，$p^q > q^p$，即

$$e^\pi > \pi^e.$$

階段 4 與切線平行的弦

現在假設由 p 到 q 的弦的斜率剛好是 $\frac{1}{e}$，與切線 $y = \frac{x}{e}$ 的斜率相同。

> **想一想** 這條弦與 $(e, 1)$ 處的切線平行。由於 $\ln x$ 是凹函數，曲線位於該切線之下。這條弦與 y 軸相交於原點的上方還是下方？

這條弦是直線 $y = \frac{x}{e} + c$，其中 c 為某個常數。它的兩個端點都在曲線上，而曲線位於切線 $y = \frac{x}{e}$ 上或其下方，並且只在 $x = e$ 處與切線相接。兩個端點不可能同時位於 $x = e$，所以最少有一個端點嚴格位於切線下方，於是整條平行的弦都在切線之下：$c < 0$。現在以 $c < 0$ 重做第 2 階段：不等式的方向反轉，$\frac{\ln p}{p} < \frac{\ln q}{q}$，因此 $q^p > p^q$。 ■

回顧

- **化為單一函數。**比較 a^b 與 b^a，變成了比較 $h(a)$ 與 $h(b)$，其中 $h(x) = \frac{\ln x}{x}$。
- **有理有據的草圖。**每一處「藉草圖說明」的背後，都有確切的事實支持：ln 的凹性（弦在曲線之下，切線在曲線之上），以及截距的正負。
- e **的特殊角色。**由原點引出的切線在 $x = e$ 處與曲線相切，而這一點正是以上所有比較的分界點。

背後的數學

$\frac{\ln x}{x}$ 的圖像

本題的一切，都可以在 $h(x) = \frac{\ln x}{x}$ 的圖像中看到（圖 3.5.1）。它的導數（derivative）是 $h'(x) = \frac{1 - \ln x}{x^2}$，所以 h 在 $(0, e)$ 上遞增，在 $x = e$ 處達到最大值（maximum）$\frac{1}{e}$，然後遞減並趨向 0。一條高度為 m 的水平線，若高度介乎 0 與 $\frac{1}{e}$ 之間，便會與圖像相交兩次，交點為 $a < e < b$，且 $h(a) = h(b)$，這正正就是 $a^b = b^a$。在 e 的右方，h 遞減，所以當 $e \leq p < q$ 時，我們有 $h(p) > h(q)$，即 $p^q > q^p$。特別地，$e^\pi > \pi^e$；更一般而

言，對每個 $x \neq e$ 都有 $e^x > x^e$：在這場較量中，沒有其他數能勝過 e。

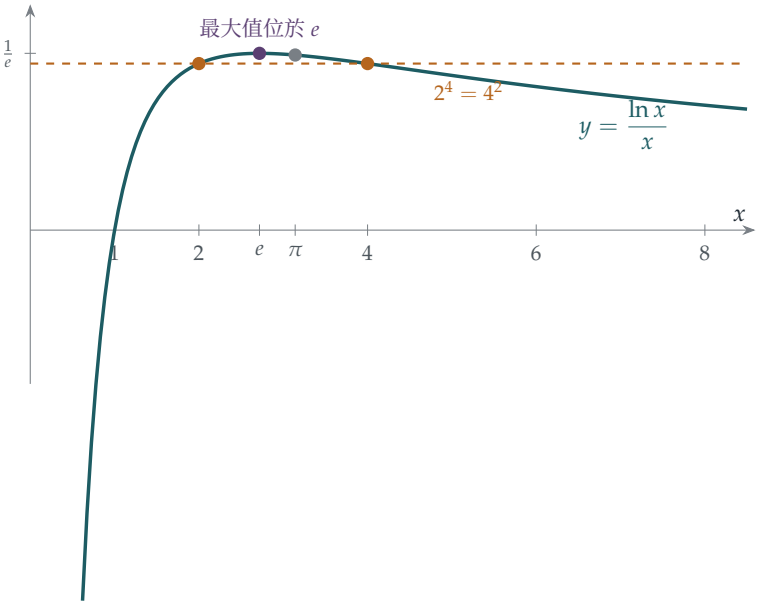


圖 3.5.1　$h(x) = \frac{\ln x}{x}$ 的圖像。它上升至 $x = e$ 處的最大值 $\frac{1}{e}$，然後下降。因為 $h(e) > h(\pi)$，所以 $e^\pi > \pi^e$。高度為 $\frac{\ln 2}{2}$ 的水平線（琥珀色）與圖像相交於 $x = 2$ 和 $x = 4$：這是唯一一對滿足 $n^m = m^n$ 的相異整數。

$x^y = y^x$ 的所有解

哪些相異的正實數對滿足 $x^y = y^x$？有一個巧妙的方法可以把它們全部找出來。設 $y = tx$，其中 $t > 1$；那麼 $x^{tx} = (tx)^x$ 給出 $x^t = tx$，所以 $x^{t-1} = t$，並且

$$x = t^{1/(t-1)}, \qquad y = t^{t/(t-1)}.$$

寫成 $t = 1 + \frac{1}{n}$，便得 $x = \left(1 + \frac{1}{n}\right)^n$ 及 $y = \left(1 + \frac{1}{n}\right)^{n+1}$。取 $n = 1$ 時，這就是 $(2, 4)$；取 $n = 2$ 時則是 $\left(\frac{9}{4}, \frac{27}{8}\right)$，而的確有 $\left(\frac{9}{4}\right)^{27/8} = \left(\frac{27}{8}\right)^{9/4}$。當 $n \to \infty$ 時，兩者都趨向 e，也就是兩個分支會合之處。1728 年，丹尼爾·伯努利（Daniel Bernoulli）與哥德巴赫（Christian Goldbach）曾在書信中討論這個問題，其後歐拉（Euler）亦有研究。

最經濟的底

函數 $x^{1/x} = e^{h(x)}$ 同樣在 $x = e$ 處取得最大值。這解答了一個有趣的問題：假如一個記數系統的「成本」是它的底（base）乘以所需的數位數目，那麼最經濟的底就是 e；若只考慮整數，則是 3（僅僅勝過 2）。有少數實驗性質的電腦採用過以 3 為底的記數法，例

如 1958 年蘇聯的 Setun。

思維技巧

重新整理以分離變數，然後研究單一函數

一條不等式或方程，若當中兩個量糾纏在一起，往往可以化爲關於同一個函數在兩點取值的命題。

一旦找到這個函數，它的圖像（在哪裏遞增、在哪裏遞減、在哪裏達到頂點）就能一次過回答所有問題。在這裏，a^b 與 b^a 的比較變成了 $h(a)$ 與 $h(b)$ 的比較。

延伸練習

1 **整數。**試做 STEP I 1988，第 1 題（`step.maths.org/questions/88-s1-q1`）。

2 **更多比較。**試做 STEP II 2025，第 3 題（載於 OCR 提供下載的 2025 年試卷，`www.ocr.org.uk/Images/750486-2025-past-papers.zip`）。該題利用本節的圖像，解決更多 a^b 與 b^a 之間的比較，其中一個要靠 e 的數值界限才能完成。

3 **一行證明。**利用當 $x \neq 0$ 時 $e^x > 1 + x$，並取 $x = \frac{\pi}{e} - 1$，直接證明 $e^\pi > \pi^e$。

延伸閱讀

- **Eli Maor, *e: The Story of a Number***（Princeton University Press）。講述 e 的歷史，包括方程 $x^y = y^x$。
- **Paul Nahin, *An Imaginary Tale***（Princeton University Press）。不只談 $\sqrt{-1}$：書中滿載正是這一類的有趣計算。
- **Michael Spivak, *Calculus***（Cambridge University Press）。第 11 章以完全嚴謹的方式，講解如何運用導數了解圖像的形狀。

3.6 斜率的界

題目

STEP I 1988，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/88-s1-q7`

引子

假如你知道一個函數在某個區間上的值總是介乎 -1 與 1 之間，你能否說出它有多陡？對一般的函數來說，答案是不能：$\sin(1000x)$ 的值介乎 -1 與 1 之間，但它的斜率可以大至 1000。不過二次函數（quadratic）不可能這樣擺動。這道題目要證明的是：一個困在方框 $|x| \leq 1$、$|y| \leq 1$ 之內的二次函數，在方框內的斜率永遠不會大於 4，而且 4 已是最佳的界。

這個問題的來歷出人意表。1887 年，以元素週期表聞名的化學家門捷列夫（Dmitri Mendeleev）正在研究溶液的密度如何隨濃度變化，並用二次函數擬合他的數據。他需要知道數據中的誤差對曲線斜率的影響最壞可以有多大，於是證明了與本題完全相同的結果。兩年後，年輕的馬可夫（Andrei Markov）解決了任意次數多項式（polynomial）的一般問題，而答案直接引出了數學中最重要的一些多項式。

動筆之前

想一想 $f'(x)$ 的公式只用到 f 的三個值：在 -1、0 和 1 的值。為甚麼三個值就足以確定一個二次函數的導數（derivative）？

二次函數有三個係數，所以只要知道它在任意三點的值，就能把它確定下來。它的導數同樣是由這些係數決定的函數，因此必定可以用這三個值表示出來。題中的公式正好告訴你怎樣表示。而這恰恰是你所需要的，因為假設 $|f(x)| \leq 1$ 告訴你的是 f 的值，而結論談的卻是它的導數。

逐步解題

階段 1　公式

$\mathrm{f}(1) = a+b+c$，$\mathrm{f}(-1) = a-b+c$，而 $\mathrm{f}(0) = c$。所以右方是

$$(a+b+c)\left(x+\tfrac{1}{2}\right) + (a-b+c)\left(x-\tfrac{1}{2}\right) - 2cx.$$

x 的係數是 $(a+b+c)+(a-b+c)-2c = 2a$，常數項是 $\tfrac{1}{2}(a+b+c) - \tfrac{1}{2}(a-b+c) = b$。所以右方等於 $2ax+b = \mathrm{f}'(x)$。 ■

階段 2　上界

> **想一想**　對公式取絕對值，然後運用三角不等式（triangle inequality），再配合 $|\mathrm{f}(1)|, |\mathrm{f}(-1)|, |\mathrm{f}(0)| \le 1$。你會得到 $|\mathrm{f}'(x)|$ 的甚麼上界（寫成 x 的函數）？

根據三角不等式 $|u+v+w| \le |u|+|v|+|w|$，

$$|\mathrm{f}'(x)| \le |\mathrm{f}(1)|\,|x+\tfrac{1}{2}| + |\mathrm{f}(-1)|\,|x-\tfrac{1}{2}| + 2|\mathrm{f}(0)|\,|x| \le |x+\tfrac{1}{2}| + |x-\tfrac{1}{2}| + 2|x|.$$

> **停一停，試一試**　證明當 $|x| \le 1$ 時，$|x+\tfrac{1}{2}| + |x-\tfrac{1}{2}| + 2|x| \le 4$。（按正負號分情況討論。）

若 $|x| \le \tfrac{1}{2}$，則 $x+\tfrac{1}{2} \ge 0$ 而 $x-\tfrac{1}{2} \le 0$，所以 $|x+\tfrac{1}{2}| + |x-\tfrac{1}{2}| = (x+\tfrac{1}{2}) - (x-\tfrac{1}{2}) = 1$，總和是 $1+2|x| \le 2$。

若 $\tfrac{1}{2} \le |x| \le 1$，兩個括號的正負號都與 x 相同，所以 $|x+\tfrac{1}{2}| + |x-\tfrac{1}{2}| = 2|x|$，總和是 $4|x| \le 4$。

無論哪種情況，都有 $|\mathrm{f}'(x)| \le 4$。 ■

這個上界只會在 $x = \pm 1$ 處達到；在區間中間的一半，斜率最多只是 2。

階段 3　一個極端例子

> **想一想**　要在 $x = 1$ 處達到上界，第 2 階段中的每個不等式都必須取等號。$\mathrm{f}(1)$、$\mathrm{f}(-1)$ 和 $\mathrm{f}(0)$ 必須取甚麼值？正負號又是甚麼？

在 $x=1$ 處，公式是 $\mathrm{f}'(1) = \tfrac{3}{2}\mathrm{f}(1) + \tfrac{1}{2}\mathrm{f}(-1) - 2\mathrm{f}(0)$。要令它等於 4，我們需要 $\mathrm{f}(1)=1$、$\mathrm{f}(-1)=1$ 和 $\mathrm{f}(0)=-1$。通過 $(-1,1)$、$(0,-1)$ 和 $(1,1)$ 的二次函數是

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

所以 $a = 2$，$b = 0$，$c = -1$。在 $|x| \le 1$ 上，$x^2 \in [0,1]$，所以 $2x^2 - 1 \in [-1,1]$，而 $\mathrm{f}'(1) = 4$。 ■

回顧

- **用已知的量表示想求的量。**公式把導數改寫成 f 的值，而假設提供的資料正正在這些值上。
- **三角不等式。**用各項絕對值之和來估計一個和的上界，是分析學（analysis）中最常見的手法之一。
- **從等號成立的條件找出極值情況。**追問每個不等式何時取等號，就直接得出例子 $2x^2 - 1$。

背後的數學

馬可夫不等式與切比雪夫多項式

馬可夫在 1889 年證明：若 p 是 n 次多項式，且在 $[-1,1]$ 上 $|p(x)| \le 1$，則

$$|p'(x)| \le n^2 \qquad \text{當} |x| \le 1.$$

當 $n = 2$ 時，這就是我們的上界 4。極端情況是*切比雪夫多項式*（Chebyshev polynomials）T_n，其定義為

$$T_n(\cos\theta) = \cos n\theta.$$

例如 $\cos 2\theta = 2\cos^2\theta - 1$ 給出 $T_2(x) = 2x^2 - 1$，即我們的極端二次函數；而 $\cos 3\theta = 4\cos^3\theta - 3\cos\theta$ 給出 $T_3(x) = 4x^3 - 3x$。由於 $T_n(x) = \cos n\theta$，這些多項式在 -1 與 1 之間來回振盪的次數，已是 n 次多項式所能做到的最多，而且它們在兩端也是最陡的（圖 3.6.1）。馬可夫的弟弟 Vladimir 在 1892 年把結果推廣到高階導數。

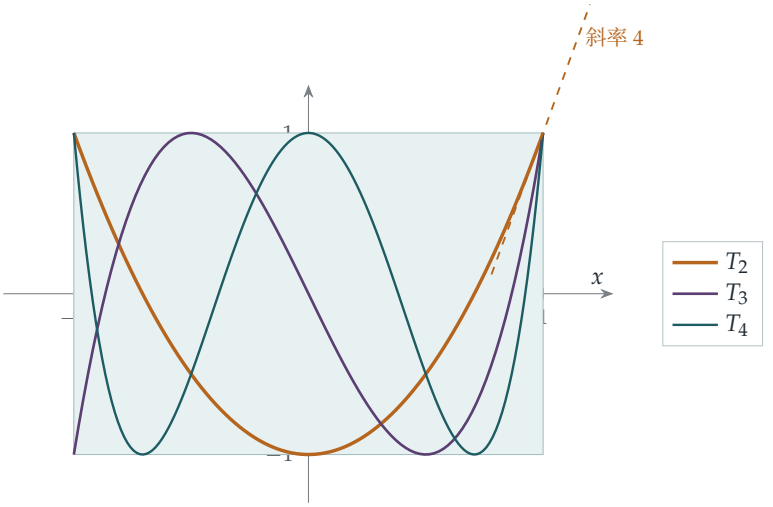


圖 3.6.1　切比雪夫多項式 $T_2 = 2x^2 - 1$、$T_3 = 4x^3 - 3x$ 和 $T_4 = 8x^4 - 8x^2 + 1$ 都留在方框 $|x| \le 1$、$|y| \le 1$ 之內，並且盡可能多次觸及方框的頂邊和底邊。在同次數的多項式中，它們在兩端最陡：$T_n'(1) = n^2$。虛線是 T_2 在 $x = 1$ 處的切線，斜率為 4。

插值

題目中的公式是*拉格朗日插值*（Lagrange interpolation）的一個例子：已知函數在若干點的值，寫出通過這些點而次數最低的唯一多項式，再用它來逼近該函數或其導數。這裏的三個點是 $-1, 0, 1$。這類公式在數值分析（numerical analysis）中無處不在，例如用來從量度所得的數據估計導數。馬可夫不等式就是隨之而來的警告：函數值的小誤差，在導數中最多可以被放大 n^2 倍，這正是門捷列夫當年擔心的問題。

切比雪夫多項式的實際應用

切比雪夫多項式是逼近理論（approximation theory）的主力工具。在 T_n 的根處對函數插值，而不是在等距的點上插值，可以把誤差控制住；若用等距的點，高次插值在區間兩端附近可以錯得非常離譜（龍格現象，Runge's phenomenon）。那些把函數計算至機器精度極限、或以極高準確度求解微分方程（differential equation）的軟件，內部往往使用切比雪夫展開式。

思維技巧

用已知的量表示未知的量

如果假設告訴你的是某一類量，而結論談的是另一類量，就去找一個把兩者連繫起來的恆等式，然後逐項估計上界。

得到一個界之後，還要問它是否最佳。沿着等號成立的情況倒推回去，通常會把你帶到極端例子，而極端例子本身往往就很值得研究。

延伸練習

1 **一般的想法。**證明若 f 是二次函數，則 $\mathrm{f}(x) = \mathrm{f}(-1)\dfrac{x(x-1)}{2} + \mathrm{f}(0)(1-x^2) + \mathrm{f}(1)\dfrac{x(x+1)}{2}$，然後求導，重新得出題目中的公式。

2 **唯一性。**較難 證明在 $[-1,1]$ 上滿足 $|\mathrm{f}(x)| \le 1$、且在 $[-1,1]$ 內某處有 $|\mathrm{f}'(x)| = 4$ 的二次函數，只有 $\pm(2x^2-1)$。

3 **切比雪夫多項式。**利用 $\cos 3\theta = 4\cos^3\theta - 3\cos\theta$ 驗證 $T_3(x) = 4x^3 - 3x$，並驗證 $T_3'(1) = 9$。證明 $T_{n+1}(x) = 2xT_n(x) - T_{n-1}(x)$，並求 T_4。

4 **T_n 的斜率。**把 $T_n(\cos\theta) = \cos n\theta$ 對 θ 求導，證明 $T_n'(\cos\theta) = \dfrac{n\sin n\theta}{\sin\theta}$，再令 $\theta \to 0$，推導出 $T_n'(1) = n^2$。

延伸閱讀

- **Lloyd N. Trefethen, *Approximation Theory and Approximation Practice***（SIAM）。一本出色的現代著作，講述切比雪夫多項式與插值，能力強的中學生也讀得懂，全書穿插電腦實驗。
- **Peter Borwein & Tamás Erdélyi, *Polynomials and Polynomial Inequalities***（Springer）。想了解馬可夫不等式及其相關結果的來龍去脈，可讀此書。程度較深。
- **George M. Phillips, *Interpolation and Approximation by Polynomials***（Springer）。一本大學本科教科書，涵蓋拉格朗日插值、切比雪夫多項式和馬可夫型不等式。

3.7 $\pi^2 \leq 10$

題目

STEP II 2014，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/14-s2-q2`

引子

每個中學生都知道 π 比 3 大一點。但你會怎樣證明它的一個上界（upper bound）呢？阿基米德（Archimedes）用的是多邊形。這道題的做法卻完全出人意表：取一條關於函數的一般不等式，代入一條簡單的拋物線（parabola），便直接讀出 $\pi^2 \leq 10$。換另一個函數，更可得出著名的近似值 $\pi \leq \frac{22}{7}$。

不等式 $(*)$ 稱為 *Wirtinger 不等式*（Wirtinger's inequality）。它說的是：一個兩端都固定為零的函數，不可能只「大」而不「陡」。它的背後是數學物理中最核心的思想之一：弦的振動、量子粒子的能級，以及特徵值（eigenvalue）理論。

動筆之前

想一想 一個在 $[0, \pi]$ 兩端都為零的函數，如果在中間取值很大，為甚麼就必須很陡？條件 $f(0) = f(\pi) = 0$ 為甚麼重要？

要在長度為 π 的距離之內，由 0 升到某個高度 H 再降回 0，函數上升的平均速率至少要有 $H/(\pi/2)$。因此它的導數（derivative）不可能處處都小。若沒有兩端的條件，就沒有這種限制：常數函數可以任意大，而導數為零。不等式 $(*)$ 把這個想法精確化，以平方的積分（integral）來量度「大小」和「陡峭程度」。

逐步解題

階段 1　正弦波

想一想　$\int_0^\pi \sin^2 nx\,\mathrm{d}x$ 等於甚麼？$\int_0^\pi \cos^2 nx\,\mathrm{d}x$ 呢？（利用 $\sin^2\theta = \frac{1}{2}(1-\cos 2\theta)$。）

兩者都等於 $\frac{\pi}{2}$，因為 $\int_0^\pi \cos 2nx\,\mathrm{d}x = 0$。取 $\mathrm{f}(x) = \sin nx$，則 $\mathrm{f}'(x) = n\cos nx$，所以

$$\int_0^\pi \mathrm{f}^2 = \frac{\pi}{2} \le \frac{n^2\pi}{2} = \int_0^\pi (\mathrm{f}')^2,$$

因為 $n \ge 1$。　■留意 $n = 1$ 時等號成立，所以這條不等式已經緊得不能再緊。

階段 2　端點條件不可或缺

想一想　找一個簡單的函數，滿足 $\mathrm{f}(0) \ne 0$，而 $(*)$ 對它不成立。然後再找一個滿足 $\mathrm{f}(0) = 0$ 但 $\mathrm{f}(\pi) \ne 0$ 的。

常數函數 $\mathrm{f}(x) = 1$ 有 $\mathrm{f}(0) = 1 \ne 0$，而且 $\int_0^\pi 1\,\mathrm{d}x = \pi > 0 = \int_0^\pi 0\,\mathrm{d}x$。

至於第二個，$\mathrm{f}(x) = x$ 有 $\mathrm{f}(0) = 0$ 但 $\mathrm{f}(\pi) = \pi$，而且

$$\int_0^\pi x^2\,\mathrm{d}x = \frac{\pi^3}{3} \approx 10.3 > \pi = \int_0^\pi 1^2\,\mathrm{d}x.$$

所以 $(*)$ 在兩種情況下都不成立。　■

階段 3　拋物線

想一想　哪個二次函數在 $x = 0$ 和 $x = \pi$ 處都等於零？（若不計常數倍，就只有一個。為甚麼這個倍數無關緊要？）

這個二次函數必定是 $x(\pi - x)$ 的倍數，而把 f 乘以常數 k，$(*)$ 的兩邊都會乘以 k^2，所以我們可以取 $\mathrm{f}(x) = x(\pi - x)$，即 $a = -1$、$b = \pi$、$c = 0$。於是 $\mathrm{f}'(x) = \pi - 2x$，再用代換（substitution）$u = x - \frac{\pi}{2}$（這樣兩個被積函數都變成偶函數），

$$\int_0^\pi x^2(\pi - x)^2\,\mathrm{d}x = \int_{-\pi/2}^{\pi/2} \left(\frac{\pi^2}{4} - u^2\right)^2 \mathrm{d}u = \frac{\pi^5}{30},$$

$$\int_0^\pi (\pi - 2x)^2\,\mathrm{d}x = \int_{-\pi/2}^{\pi/2} 4u^2\,\mathrm{d}u = \frac{\pi^3}{3}.$$

（第一個積分：$\int_{-h}^{h}(h^2 - u^2)^2\,\mathrm{d}u = 2(h^5 - \frac{2}{3}h^5 + \frac{1}{5}h^5) = \frac{16}{15}h^5$，其中 $h = \frac{\pi}{2}$。）所以 $(*)$ 給出 $\frac{\pi^5}{30} \le \frac{\pi^3}{3}$，即

$$\pi^2 \le 10. \qquad ■$$

階段 4 三角函數的選擇

想一想 對於 $f(x) = p\sin\frac{1}{2}x + q\cos\frac{1}{2}x + r$，$f(0) = 0$ 和 $f(\pi) = 0$ 對 p、q、r 施加了甚麼條件？

$f(0) = q + r = 0$ 及 $f(\pi) = p + r = 0$。所以 $p = q = -r$，經縮放後，我們取 $f(x) = \sin\frac{x}{2} + \cos\frac{x}{2} - 1$。

停一停，試一試 對這個函數計算 $\int_0^\pi f^2$ 和 $\int_0^\pi (f')^2$。由此得出關於 π 的甚麼不等式？

記 $s = \sin\frac{x}{2}$ 及 $c = \cos\frac{x}{2}$。則 $(s+c)^2 = 1 + 2sc = 1 + \sin x$，所以

$$f^2 = (s+c)^2 - 2(s+c) + 1 = 2 + \sin x - 2\sin\tfrac{x}{2} - 2\cos\tfrac{x}{2}.$$

利用 $\int_0^\pi \sin x = 2$、$\int_0^\pi \sin\frac{x}{2} = 2$ 及 $\int_0^\pi \cos\frac{x}{2} = 2$，

$$\int_0^\pi f^2\,dx = 2\pi + 2 - 4 - 4 = 2\pi - 6.$$

另外 $f' = \frac{1}{2}(c - s)$，所以 $(f')^2 = \frac{1}{4}(1 - \sin x)$，而 $\int_0^\pi (f')^2\,dx = \frac{1}{4}(\pi - 2)$。於是 $(*)$ 給出

$$2\pi - 6 \le \frac{\pi - 2}{4} \quad\Longrightarrow\quad 8\pi - 24 \le \pi - 2 \quad\Longrightarrow\quad \pi \le \frac{22}{7}. \qquad \blacksquare$$

階段 5 哪個較好？

由 $\pi \le \frac{22}{7}$ 可得 $\pi^2 \le \frac{484}{49} \approx 9.878$，比 $\pi^2 \le 10$ 更好。（真實值是 $\pi^2 \approx 9.870$。）留意這裏證明的是 $\frac{22}{7}$ 為 π 的一個上界；而 Section 4.3 中一個截然不同的積分，則證明它並不等於 π。

回顧

- **一般的不等式，特定的函數。**選定一個測試函數（test function），就把一個關於所有函數的命題，變成一個關於某個數的命題。
- **縮放。**$(*)$ 的兩邊按同一個因子縮放，所以在每一族函數中，我們都可以把那個自由常數固定下來。
- **測試函數愈好，界愈好。**測試函數愈接近極端情況 $\sin x$，不等式就愈接近等式，估計也就愈準確（圖 3.7.1）。

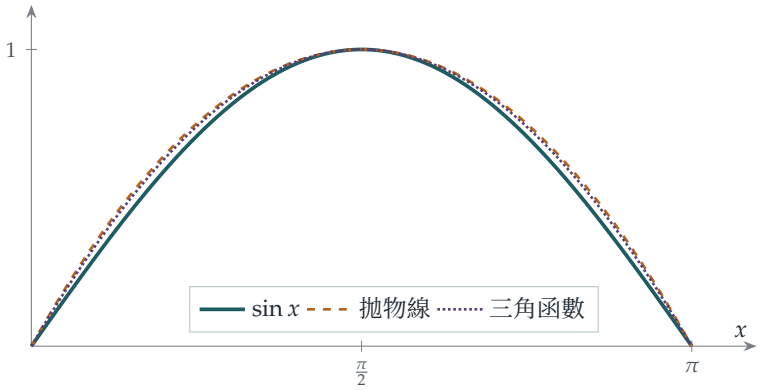


圖 3.7.1　本節的三個函數，各自縮放至最大值為 1：極端情況 $\sin x$（實線），它令 $(*)$ 的等號成立；拋物線 $x(\pi - x)$（虛線），它給出 $\pi^2 \le 10$；以及 $\sin\frac{x}{2} + \cos\frac{x}{2} - 1$（點線），它給出 $\pi \le \frac{22}{7}$。兩個測試函數都非常接近 $\sin x$，所以得出的界才這樣好。

背後的數學

用傅立葉級數的證明

為甚麼 $(*)$ 成立？到了大學你會學到，$[0, \pi]$ 上任何在兩端都為零的合理函數，都可以寫成傅立葉正弦級數（Fourier sine series）：

$$\mathrm{f}(x) = b_1 \sin x + b_2 \sin 2x + b_3 \sin 3x + \cdots .$$

由於當 $m \neq n$ 時 $\int_0^\pi \sin mx \sin nx \, \mathrm{d}x = 0$，平方後再積分，所有交叉項都會消失，得出

$$\int_0^\pi \mathrm{f}^2 = \frac{\pi}{2}\left(b_1^2 + b_2^2 + b_3^2 + \cdots\right), \qquad \int_0^\pi (\mathrm{f}')^2 = \frac{\pi}{2}\left(1^2 b_1^2 + 2^2 b_2^2 + 3^2 b_3^2 + \cdots\right).$$

右式的每一項都不小於左式的對應項，這就證明了 $(*)$，同時說明等號只在 f 是 $\sin x$ 的倍數時成立。題目的第 (i) 部分就是只有一項的情況。

振動的弦與量子粒子

比值 $\int (\mathrm{f}')^2 / \int \mathrm{f}^2$ 稱為瑞利商（Rayleigh quotient），而 $(*)$ 說的是：對於固定在 0 和 π 的函數，它的最小可能值是 1。這個數有物理意義。對一條兩端固定的弦來說，$\int (\mathrm{f}')^2$ 量度拉伸的能量，$\int \mathrm{f}^2$ 量度位移的大小，而這個比值的最小值決定了弦所能奏出的最低音。在量子力學（quantum mechanics）中，同一個最小值就是困在盒子裏的粒子的最低能量：它不可能是零，因為兩端固定的波必須彎曲。瑞利勳爵（Lord Rayleigh）正是

用這一招——猜一個測試函數，再計算比值——來估計振動系統的頻率；由於極端情況給出最小值，任何猜測都會給出真正最低頻率的一個上界，而好的猜測更會給出非常準確的上界。

變分法

找出令某個積分之比取得最小值的函數，是*變分法*（calculus of variations）的問題，當中的未知量是整個函數，而不是一個數。這裏取得最小值的函數 $\sin x$ 滿足微分方程（differential equation）$f'' + f = 0$，而這並非巧合：求積分的最小值，總會導出一條由取得最小值的函數所滿足的微分方程。我們會在 Section 4.7 再談這個想法。

思維技巧

測試函數

一條對某類函數中每一個都成立的不等式，可以套用在一個巧妙選取的函數上，從而得到關於某個特定數的資訊。

箇中藝術在於選擇：接近極端情況的測試函數，會給出近乎最佳的結果。這個想法貫穿分析學、物理學和工程學，由估計特徵值到有限元法（finite element method）皆然。

延伸練習

1 **四次函數。**在 $(*)$ 中取 $f(x) = x^2(\pi - x)^2$。證明它只能給出 $\pi^2 \le 12$。為甚麼這個測試函數比拋物線差？（把它的形狀與 $\sin x$ 比較。）

2 **兩個正弦波。**對 $f(x) = \sin x + \sin 2x$ 直接驗證 $(*)$，把兩邊都計算出來。

3 **另一個區間。**用代換證明：若 $g(0) = g(L) = 0$，則 $\displaystyle\int_0^L g^2 \le \frac{L^2}{\pi^2}\int_0^L (g')^2$。這對不同長度的弦說明了甚麼？

4 **瑞利法。**較難 取 $f(x) = x(\pi - x) + \lambda x^2(\pi - x)^2$，並選取 λ 使 π^2 的界盡可能好。（代數運算相當繁重，用電腦代數系統會有幫助。）

延伸閱讀

- **T. W. Körner, *Fourier Analysis***（Cambridge University Press）。一本精彩而旁徵博引的

書，講述傅立葉級數及其應用，作者是劍橋的分析學家。

- **Hardy, Littlewood & Pólya, *Inequalities***（Cambridge University Press）。第 7 章討論 Wirtinger 不等式及相關的不等式。
- **Richard Feynman, *The Feynman Lectures on Physics***，第 II 卷第 19 章（「The principle of least action」，最小作用量原理）。費曼（Feynman）關於求積分最小值的一講，通篇貫穿物理直覺。可於網上免費閱讀。

3.8 由加法公式決定的函數

題目

STEP III 2019，第 2 題

繼續閱讀之前，請先看題目（原題為英文）。題目載於 OCR 提供下載的 2019 年 STEP 試卷（zip 檔案）內的 STEP III 試卷：

`www.ocr.org.uk/Images/696590-2019-past-papers.zip`

引子

你怎樣知道 e^x 是甚麼？在學校裏，它是透過複利息引入的，或者說成是導數（derivative）等於自身的函數，又或者寫成 $\lim(1+x/n)^n$。這道題給出另一個答案：e^{kx} 是唯一能把加法變成乘法的可微（differentiable）函數，即 $\mathrm{f}(x+y)=\mathrm{f}(x)\mathrm{f}(y)$。單憑這條規則，再加上一個數 $\mathrm{f}'(0)$，就足以把這個函數完全確定下來。

第二部分對一條較陌生的規則做同樣的事：$\mathrm{g}(x+y)=\frac{\mathrm{g}(x)+\mathrm{g}(y)}{1+\mathrm{g}(x)\mathrm{g}(y)}$。如果你覺得這條公式似曾相識，那就對了：它正是狹義相對論（special relativity）中愛因斯坦（Einstein）的速度相加法則，只是把光速設為 1。這條規則迫使函數必然是 tanh，而這正好解釋了為何任何物體都不能被加速至超越光速。

動筆之前

像 $\mathrm{f}(x+y)=\mathrm{f}(x)\mathrm{f}(y)$ 這樣以函數為未知量的方程，稱為*函數方程*（functional equation）。

? **想一想** 這條方程對所有 x 和 y 都成立。最值得首先嘗試的是甚麼？

特殊值。設 $y=0$（或 $x=y$，或 $y=-x$），往往不費甚麼工夫就能得到大量資料。兩個部分都是這樣起步的。

逐步解題

階段 1　在零處的值

代入 $x = y = 0$：$\mathrm{f}(0) = \mathrm{f}(0)^2$，所以 $\mathrm{f}(0) = 0$ 或 1。

> **想一想**　如果 $\mathrm{f}(0) = 0$，會出甚麼問題？

若 $\mathrm{f}(0) = 0$，則代入 $y = 0$ 可得對每個 x 都有 $\mathrm{f}(x) = \mathrm{f}(x)\mathrm{f}(0) = 0$。於是 f 恆等於零，且 $\mathrm{f}'(0) = 0$，與 $\mathrm{f}'(0) = k \neq 0$ 矛盾。所以 $\mathrm{f}(0) = 1$。 ■

階段 2　由零處的導數得出處處的導數

> **想一想**　在 $(*)$ 中，寫成 $\mathrm{f}(x+h) = \mathrm{f}(x)\mathrm{f}(h)$。有甚麼可以抽到極限之外？

$$\begin{aligned}\mathrm{f}'(x) &= \lim_{h\to 0}\frac{\mathrm{f}(x)\mathrm{f}(h) - \mathrm{f}(x)}{h} = \mathrm{f}(x)\lim_{h\to 0}\frac{\mathrm{f}(h)-1}{h}\\ &= \mathrm{f}(x)\lim_{h\to 0}\frac{\mathrm{f}(0+h)-\mathrm{f}(0)}{h} = \mathrm{f}(x)\,\mathrm{f}'(0) = k\,\mathrm{f}(x).\end{aligned}$$ ■

函數方程讓我們把 0 附近的性態搬到其他每一點去。

階段 3　求解 $\mathrm{f}' = k\mathrm{f}$

考慮 $\mathrm{u}(x) = \mathrm{f}(x)e^{-kx}$。則 $\mathrm{u}'(x) = \big(\mathrm{f}'(x) - k\mathrm{f}(x)\big)e^{-kx} = 0$，所以 u 是常數，而 $\mathrm{u}(0) = \mathrm{f}(0) = 1$。因此

$$\mathrm{f}(x) = e^{kx}.$$ ■

（分離變數（separating the variables），$\int \frac{\mathrm{df}}{\mathrm{f}} = \int k\,\mathrm{d}x$，會得到相同的答案，但需要 $\mathrm{f}(x) \neq 0$；用 u 這個技巧就免卻了這重顧慮。）

階段 4　第二條規則

> **想一想**　在 g 的規則中代入 $y = 0$。這告訴你 $\mathrm{g}(0)$ 是甚麼？$|\mathrm{g}(x)| < 1$ 在哪裏派上用場？

取 $y = 0$：$\mathrm{g}(x)\big(1 + \mathrm{g}(x)\mathrm{g}(0)\big) = \mathrm{g}(x) + \mathrm{g}(0)$，化簡後得 $\mathrm{g}(0)\big(\mathrm{g}(x)^2 - 1\big) = 0$。由於 $|\mathrm{g}(x)| < 1$，括號內的式子永不為零，所以 $\mathrm{g}(0) = 0$。

> **停一停，試一試**　現在利用這條規則計算 $\mathrm{g}(x+h) - \mathrm{g}(x)$，化簡，除以 h，再令 $h \to 0$。

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

除以 h。當 $h\to 0$ 時，$\frac{\mathrm{g}(h)}{h}=\frac{\mathrm{g}(h)-\mathrm{g}(0)}{h}\to \mathrm{g}'(0)=k$，而 $\mathrm{g}(h)\to \mathrm{g}(0)=0$（可微函數必定連續（continuous）），所以分母趨向 1。因此

$$\mathrm{g}'(x)=k\bigl(1-\mathrm{g}(x)^2\bigr). \qquad \blacksquare$$

階段 5 求解

> ? **想一想** 分離變數。當 $|g|<1$ 時，哪個函數的導數是 $\frac{1}{1-g^2}$？

由於 $|\mathrm{g}|<1$，我們可以相除：$\dfrac{\mathrm{g}'}{1-\mathrm{g}^2}=k$。函數 $\operatorname{artanh} g=\frac{1}{2}\ln\frac{1+g}{1-g}$ 的導數是 $\frac{1}{1-g^2}$，所以 $\frac{\mathrm{d}}{\mathrm{d}x}\operatorname{artanh}\mathrm{g}(x)=k$，而 $\operatorname{artanh}\mathrm{g}(x)=kx+C$。由 $\mathrm{g}(0)=0$ 得 $C=0$，所以

$$\mathrm{g}(x)=\tanh kx. \qquad \blacksquare$$

驗算：加法公式 $\tanh(A+B)=\dfrac{\tanh A+\tanh B}{1+\tanh A\tanh B}$ 正正就是題目給出的規則。

回顧

- **先試特殊值。**代入 $y=0$ 便求得 $\mathrm{f}(0)$ 和 $\mathrm{g}(0)$。
- **由一點到處處。**加法定律（addition law）把 0 處的導數搬到每個 x，把函數方程化為微分方程（differential equation）。
- **唯一性（uniqueness）。**這個論證不單是驗證 e^{kx} 和 $\tanh kx$ 符合要求，更證明了它們是僅有的可能。函數方程之所以能成為定義，原因正在於此。

背後的數學

相對論中的速度

在日常生活中，如果一列火車以速率 u 行駛，而你在車廂內以速率 v 向前走，你相對於地面的速率就是 $u+v$。愛因斯坦在 1905 年證明這只是一個近似。正確的法則（其中 c 為光速）是

$$u\oplus v=\frac{u+v}{1+uv/c^2}.$$

若以 c 的分數來量度速率，這正是 g 的規則。我們的結果說明，如果把每個速度寫成

$v = \tanh\phi$，那麼速度的合成就是把各個 ϕ 相加。$\phi = \operatorname{artanh} v$ 這個量稱為快度（rapidity）。火箭引擎每次推進都會增加固定份量的快度，但由於對每個 ϕ 都有 $\tanh\phi < 1$，速度只會愈來愈接近光速，卻永遠達不到（圖 3.8.1）。把 $0.9c$ 與 $0.9c$ 合成，得到的不是 $1.8c$，而是 $\frac{1.8}{1.81}c \approx 0.9945c$。

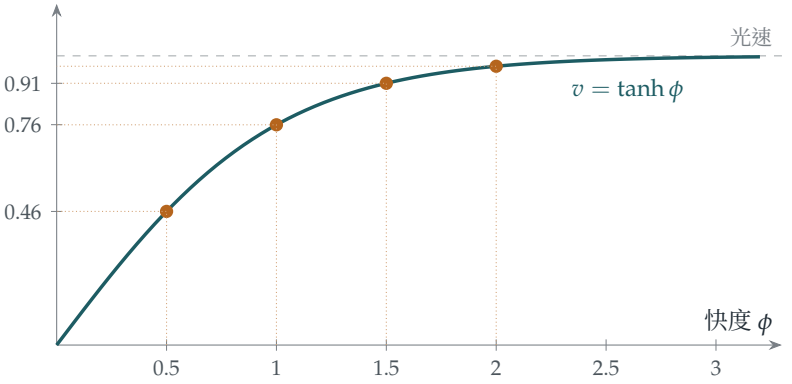


圖 3.8.1　速度（以光速的分數表示）對快度的圖像。快度每次增加相同的份量（引擎每次相同的推進所提供的正是這樣），速度的增幅卻愈來愈小：0.46、0.76、0.91、0.96、…，永遠達不到 1。

柯西函數方程

最簡單的函數方程是 $f(x+y) = f(x) + f(y)$，柯西（Cauchy）在 1821 年研究過它。如果 f 可微，用本題的方法可證明 $f(x) = kx$。事實上只要連續就夠了。但若完全不作任何正則性（regularity）假設，就會有其他極其怪異的解，它們的圖像稠密地散佈在整個平面上。Georg Hamel 在 1905 年構造出這些解；這個構造需要用到*選擇公理*（axiom of choice），而且不可能把其中任何一個明確地寫出來。這就是題目堅持 f 可微的原因之一：這個假設並非例行公事。

同態

$f(x+y) = f(x)f(y)$ 的意思是，f 把實數的加法變成正實數的乘法，並且保持兩者的結構。用大學一年級代數的語言來說，f 是由 $(\mathbb{R}, +)$ 到 $(\mathbb{R}_{>0}, \times)$ 的*羣同態*（group homomorphism），而對數就是它的反函數。數學中許多最重要的函數，都是這樣由它們所保持的結構來刻劃的。在 Section 7.6 中，我們會看到 tanh 的幾何意義：它量度的是「雙曲角」（hyperbolic angle），而它的加法定律來自雙曲旋轉的複合。

思維技巧

用性質來定義函數

不用公式，而是用函數所遵守的規則來描述它，並證明只有一個函數遵守這條規則。

數學家往往就是這樣看待 exp、sin、cos 和許多其他函數的：不把它們看成公式，而是看成函數方程或微分方程的唯一解。這解釋了它們爲何重要，而不只是怎樣計算它們。

延伸練習

1 **正切的規則。**設 g 可微，$g'(0) = k \neq 0$，且只要等式兩邊都有定義，就有 $g(x+y) = \dfrac{g(x)+g(y)}{1-g(x)g(y)}$。證明 $g'(x) = k(1+g(x)^2)$，並由此推出在 $x=0$ 附近 $g(x) = \tan kx$。

2 **柯西方程。**若 f 可微，且對所有 x、y 都有 $f(x+y) = f(x)+f(y)$，證明 $f(x) = kx$。

3 **不用導數。**較難只假設 $f(x+y) = f(x)f(y)$ 及 $f(1) = a > 0$。證明對正整數 n 有 $f(n) = a^n$，然後證明 $f(1/n) = a^{1/n}$，再證明對每個正有理數 q 都有 $f(q) = a^q$。（為甚麼要再進一步就需要連續性？）

4 **火箭科學。**一枚火箭每次點火，都能相對於它當時所處的參考系把速率增加 $0.5c$。由靜止開始，經過 2、3 和 4 次點火後，它（相對於起點）的速率是多少？請使用快度。

延伸閱讀

- **Christopher Small, *Functional Equations and How to Solve Them***（Springer）。一本淺白易讀的函數方程入門書，以喜歡解題的讀者為對象。
- **Edwin Taylor & John Wheeler, *Spacetime Physics***（W. H. Freeman）。狹義相對論的經典入門書，全書都使用快度。
- **Tom Körner, *Calculus for the Ambitious***（Cambridge University Press）。由指數函數和三角函數的性質出發來定義它們，與本節的精神完全一致。

3.9 創造一個新函數

題目

STEP II 2021，第 4 題

繼續閱讀之前，請先看題目（原題為英文）。題目載於 OCR 提供下載的 2021 年 STEP 試卷（zip 檔案）內的 STEP II 試卷：

`www.ocr.org.uk/Images/696588-2021-past-papers.zip`

引子

你能解 $e^{-x} = 5x$ 嗎？計算機上沒有任何一個函數幫得上忙。移項整理也無濟於事：x 同時困在指數函數的裏面和外面。你可以用試誤法（trial and improvement）求出數值答案，但並沒有甚麼「公式」。

真的沒有嗎？這道題踏出了大膽的一步：如果某一族方程無法用已知的函數解出來，那就發明一個新函數來解它們。數學家當年正是這樣得到 $\sqrt{x}$（x^2 的反函數）、$\ln x$（e^x 的反函數）和 $\arcsin x$ 的。這裏的新函數是 xe^x 的反函數（inverse function），現稱為蘭伯特 W 函數（Lambert W function）。有了它，從 $x^x = 10$ 到一條物理定律，範圍廣得出人意料的各種方程都可以精確地解出來。

動筆之前

想一想 「解」一條方程，究竟是甚麼意思？$x = \sqrt{2}$ 是 $x^2 = 2$ 的解，還是只不過是那個解的名字？

兩者皆是。$\sqrt{2}$ 是「平方等於 2 的那個正數」的名字；我們接受它作為答案，是因為函數 $\sqrt{\ }$ 已為人熟知、容易計算，而且性質清楚。一個答案算不算「公式」，取決於我們約定了為哪些函數命名。這道題為這份名單再添一個。

逐步解題

階段 1　曲線 $y = xe^{x}$

$\frac{dy}{dx} = e^{x} + xe^{x} = (1+x)e^{x}$，它只在 $x = -1$ 處為零，此時 $y = -e^{-1}$。導數（derivative）在 $x < -1$ 時為負，在 $x > -1$ 時為正，所以這是一個極小值（minimum）。

> **想一想**　當 $x \to +\infty$ 和 $x \to -\infty$ 時，xe^{x} 會怎樣？曲線在哪裏與坐標軸相交？

當 $x \to \infty$ 時，$xe^{x} \to \infty$。當 $x \to -\infty$ 時，$xe^{x} \to 0$（從下方趨近），因為指數衰減勝過線性增長。曲線通過原點。所以它從負 x 軸的稍下方進來，下降至極小點 $(-1, -\frac{1}{e})$，然後上升穿過 $(0,0)$，再急速向上（圖 3.9.1）。

階段 2　反函數

要有反函數，f 必須是一對一（one-to-one）的，所以必須把它限制在一個嚴格遞增（strictly increasing）或嚴格遞減的區域上。包含右邊部分的最大區間是 $[-1,\infty)$，所以 $a = -1$。在 $[-1,\infty)$ 上，f 由 $-\frac{1}{e}$ 遞增至 ∞，所以反函數 g 定義在 $[-\frac{1}{e},\infty)$ 上，由 $g(-\frac{1}{e}) = -1$ 開始遞增，經過 $g(0) = 0$，並且增長緩慢（當 x 很大時，比 $\ln x$ 還要慢）。它的圖像是 $y = xe^{x}$ 右邊部分關於直線 $y = x$ 的反射，在 $(-\frac{1}{e}, -1)$ 處有一條垂直切線。

g 的定義性質是

$$g(z)\, e^{g(z)} = z, \qquad g(z) \geq -1.$$

所以 (iii) 中每一部分的策略都一樣：把方程整理成 (某個式子) $e^{(\text{同一個式子})}$ = (一個數) 的形式。

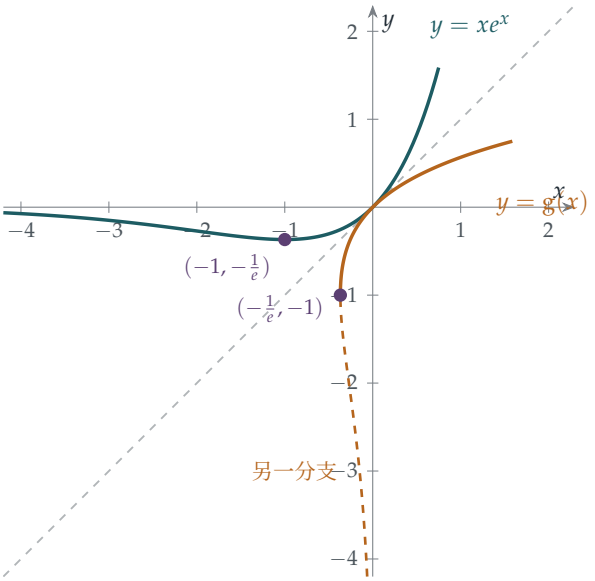


圖 3.9.1 曲線 $y = xe^x$（藍綠色）在 $(-1, -\frac{1}{e})$ 處有極小值。限制在 $x \geq -1$ 上，再關於 $y = x$ 反射，便得到反函數 g（琥珀色實線），即蘭伯特 W 函數。琥珀色虛線是 $y = xe^x$ 左邊部分的反射：這是第二個分支，第 (iii) 部分的第二個根要用到它。

階段 3 四條方程

(a) $e^{-x} = 5x$。乘以 e^x 再除以 5：$xe^x = \frac{1}{5}$。所以 $x = g(\frac{1}{5}) \approx 0.169$。由於 xe^x 在 $x < 0$ 時為負，任何根都是正數，而 xe^x 在 $x > 0$ 時遞增，所以這是唯一的根。

> **想一想** 在 (b) 和 (c) 中，未知數以 $x \ln x$ 的形式出現。用甚麼代換（substitution）可以把 $x \ln x$ 變成 ue^u 的形式？

設 $x = e^u$，則 $\ln x = u$，而 $x \ln x = ue^u$。

(b) $2x \ln x + 1 = 0$ 變成 $ue^u = -\frac{1}{2}$。但對所有實數 u 都有 $ue^u \geq -\frac{1}{e} \approx -0.368$，而 $-\frac{1}{2} < -\frac{1}{e}$。所以**沒有實根**。

(c) $3x \ln x + 1 = 0$ 變成 $ue^u = -\frac{1}{3}$，而 $-\frac{1}{e} < -\frac{1}{3} < 0$。高度為 $-\frac{1}{3}$ 的水平線與 $y = ue^u$ 相交兩次，在極小點 $u = -1$ 的兩側各一次。右邊的根是 $u = g(-\frac{1}{3})$，由此得

$$x = e^{g(-1/3)} \approx 0.538.$$

另一個根滿足 $u < -1 < g(-\frac{1}{3})$，而 $x = e^u$ 隨 u 遞增，所以我們找到的根是兩者之中**較大**的一個。（另一個約為 0.220。）

停一停，試一試 現在試做 (d)，$x = 3\ln x$。（提示：兩邊取指數，設法湊成 (某個式子) $e^{\text{某個式子}}$ 的形式，其中「某個式子」是 x 的倍數。）

(d) $x = 3\ln x$ 即 $e^{x/3} = x$，所以 $xe^{-x/3} = 1$，再乘以 $-\frac{1}{3}$，

$$\left(-\frac{x}{3}\right)e^{-x/3} = -\frac{1}{3}.$$

所以 $-\frac{x}{3}$ 是在 (c) 中找到的 $ue^u = -\frac{1}{3}$ 的兩個解之一。取由 g 給出的那一個，$x = -3\,\mathrm{g}(-\frac{1}{3}) \approx 1.857$。另一個根來自另一個解 $u < \mathrm{g}(-\frac{1}{3})$，它給出較大的 $x = -3u$。所以這一次我們找到的根是兩者之中**較小**的一個。（另一個約為 4.536。） ■

階段 4 $x^x = 10$

取對數，$x\ln x = \ln 10$。設 $x = e^u$，這就是 $ue^u = \ln 10$，所以 $u = \mathrm{g}(\ln 10)$，而

$$x = e^{\mathrm{g}(\ln 10)} \approx 2.506.$$

由於 $\mathrm{g}(z)e^{\mathrm{g}(z)} = z$，我們也可以把它寫成 $x = \dfrac{\ln 10}{\mathrm{g}(\ln 10)}$。 ■

回顧

- **為缺少的函數命名。**當一條方程無法用已知函數解出時，把一個新函數定義為反函數，是正當的做法，而且往往富有啟發性。
- **整理成標準形式。** (iii) 和 (iv) 的每一部分最終都歸結為 $ue^u =$ 常數，用的是 $x = e^u$ 或 $u = -\frac{x}{3}$ 之類的代換。
- **用圖像數根的個數。** $ue^u = c$ 有零個、一個還是兩個解，只取決於 c 相對於極小值 $-\frac{1}{e}$ 的位置。

背後的數學

歷史悠久、名字卻很新的函數

蘭伯特（Johann Heinrich Lambert）在 1758 年研究過這類方程，歐拉（Euler）在 1783 年作了進一步的研究。但這個函數一直沒有標準名稱，直到 1990 年代，電腦代數系統（computer algebra system）Maple 的開發者需要為它起一個名字。1996 年，Robert Corless、Gaston Gonnet、David Hare、David Jeffrey 和 Donald Knuth 發表了論文「On the Lambert W function」，匯集了它的性質和應用。這篇論文已被引用數千次，

而 W 現已內置於大多數數學軟件之中。題目中的函數 g 是主分支（principal branch）W_0；圖 3.9.1 中的虛線，即給出 (c) 和 (d) 中第二個根的那條曲線，是另一個實分支 W_{-1}。

W 出現在哪裏

- **維恩位移定律（Wien's displacement law）。**熱的物體輻射最強的波長與其溫度成反比。當中的常數來自解方程 $5(1-e^{-x})=x$，其正數解為 $x=5+W\left(-5e^{-5}\right)\approx 4.965$。這就是為甚麼約 5800 K 的太陽看起來是黃白色，而溫度較低的紅矮星看起來是紅色。
- **冪塔（power tower）。**無窮冪塔 $x^{x^{x^{\cdots}}}$ 在 $e^{-e}\le x\le e^{1/e}$ 時收斂（converge），其值為 $\dfrac{W(-\ln x)}{-\ln x}$。
- **人口與傳染病模型。**在最簡單的模型中，一場傳染病的最終規模，以及某些增長模型中達到指定人口所需的時間，都由 W 給出。

初等函數

W 真的是新函數嗎？還是可以用 exp、ln、根式和三角函數寫出來？它確實是新的：一條可追溯至 1830 年代劉維爾（Joseph Liouville）的定理表明，W 不是*初等函數*（elementary function）。同一套理論還表明，一些看似平平無奇的積分，例如 $\int e^{-x^2}\,dx$，無法用初等函數表示，這就是為甚麼正態分佈（normal distribution）需要自己專用的特殊函數。

思維技巧

以反函數作為解

方程 $f(x)=c$ 的解是 $x=f^{-1}(c)$，前提是把 f 限制成一對一的函數。這個限制決定了你得到的是哪一個解。

許多看似不同的方程，其實是同一條方程的不同偽裝；找出能揭示共同形式的代換，就是關鍵的一步。

延伸練習

1 **二、四，還有一個。**方程 $x^2 = 2^x$ 有顯而易見的解 $x = 2$ 和 $x = 4$（比較 Section 3.5）。證明它還有一個負數解，而這個解是

$$x = -\frac{2}{\ln 2}\,\mathrm{g}\left(\frac{\ln 2}{2}\right) \approx -0.767.$$

2 **一個導數。**對 $\mathrm{g}(x)e^{\mathrm{g}(x)} = x$ 作隱函數微分（implicit differentiation），證明當 $x \neq 0$ 時，

$$\mathrm{g}'(x) = \frac{\mathrm{g}(x)}{x\big(1 + \mathrm{g}(x)\big)}.$$

在 $x = -\frac{1}{e}$ 處會怎樣？

3 **一個積分。**利用代換 $x = we^w$ 證明 $\displaystyle\int_0^e \mathrm{g}(x)\,\mathrm{d}x = e - 1$。

4 **冪塔。**假設冪塔 $T = \sqrt{2}^{\sqrt{2}^{\sqrt{2}^{\cdots}}}$ 收斂，證明 $T = \sqrt{2}^{T}$，並由此推出 $T = 2$ 或 $T = 4$。究竟是哪一個？（計算數列 $t_1 = \sqrt{2}$、$t_{n+1} = \sqrt{2}^{t_n}$ 的頭幾項，並想想 Section 2.2 中的蛛網圖（cobweb diagram）。）

延伸閱讀

- **Robert Corless et al., "On the Lambert W function"**，*Advances in Computational Mathematics* 5 (1996)。為這個函數命名的論文。讀完本節之後，論文的頭幾節應可讀懂。
- **Brian Hayes, "Why W?"**，*American Scientist* 93 (2005)。一篇生動的文章，講述這個函數的歷史和用途。可在網上免費閱讀。
- **Brian Conrad, "Impossibility theorems for elementary integration"**。網上可找到的講義，解釋劉維爾的理論。程度較深。

我們學到了甚麼

本章以一個 $1 = 2$ 的「證明」(proof) 和三個問題開始：怎樣知道一個論證是正確的？怎樣證明不等式 (inequality)？函數 (function) 究竟是甚麼？以下是本章九道題目教曉我們的東西。

1. 怎樣知道一個論證是正確的？

Section 3.1 給了我們一份檢查清單。每走一步，都要問三件事：計算對嗎？這一步在邏輯上成立嗎？有甚麼假設是沒有明說的？那一節出現的錯誤都很典型。比較兩個駐點 (stationary point) 上的函數值，並不能說明哪一個是極大值。在無限的定義域上，駐點的函數值不能作為函數的界。微積分基本定理 (fundamental theorem of calculus) 要求被積函數連續 (continuous)。此外，一個論證也可能只是碰巧得出「看似正確」的答案。

同樣的警覺貫穿本章其餘部分。Section 3.7 用反例 (counterexample) 說明，去掉一個假設 ($f(0) = f(\pi) = 0$)，不等式便不再成立。在 Section 3.8 中，函數可微 (differentiable) 這個假設必不可少：沒有它，柯西 (Cauchy) 函數方程 (functional equation) 便有一些誰也寫不出來的怪異解。而在 Section 3.6 和 Section 3.3 中，等號成立的條件告訴我們某個界是否已是最佳，這正是嚴謹的數學家總會問的那類問題。

2. 怎樣證明不等式？

我們從幾個不同方向接觸了各個經典不等式，也看到它們之間的關係。

- **平方非負。**兩個數的算術－幾何平均不等式 (AM–GM) 其實就是 $(a-b)^2 \geq 0$。柯西的前推－後推歸納法 (forward–backward induction) 由此建立起一般情形 (Section 3.2)。
- **凸性 (convexity)。**凹 (concave) 曲線位於它的弦之上。單憑這一點便得出琴生不等式 (Jensen's inequality)；配合對數，一行便得出 AM–GM；配合 $\sin x$，則得出一個關於三角形的精確不等式 (Section 3.3)。
- **面積。**曲線之下的面積加上曲線旁邊的面積，足以蓋過一個長方形。這給出楊氏不等式 (Young's inequality)，再由此推出赫爾德 (Hölder) 不等式和柯西－施瓦茲 (Cauchy–Schwarz) 不等式 (Section 3.4)。

- **圖像的形狀。**把 a^b 與 b^a 的比較改寫成 $\frac{\ln x}{x}$ 在兩點上的比較，便解決了 π^e 與 e^π 孰大孰小的問題（Section 3.5）。
- **插值（interpolation）與三角不等式（triangle inequality）。**把導數用函數值表示，然後逐項估計其界，便得出二次多項式的馬爾可夫不等式（Markov's inequality）（Section 3.6）。
- **試驗函數（test function）。**把一個一般的不等式應用於巧妙選取的函數，便得出 $\pi^2 \leq 10$ 和 $\pi \leq \frac{22}{7}$（Section 3.7）。

在這些方法之中作出選擇，是一種要靠練習才能養成的技巧。粗略的指引是：見到積與和，可考慮 AM–GM；見到函數值的平均，可考慮琴生不等式；見到一個函數和它的反函數（inverse function），可考慮楊氏不等式；要比較兩個形式相同的式子，則可嘗試找出它們背後的那一個函數。

3. 函數究竟是甚麼？

最後兩節從兩方面回答了這個問題。在 Section 3.8 中，e^{kx} 和 $\tanh kx$ 根本不是由公式給出的；它們是滿足某些加法定律（addition law）的僅有可微函數。這是比公式更深一層的理解：它解釋了為何這些函數不斷出現，既見於複利息，也見於愛因斯坦（Einstein）的速度相加公式。在 Section 3.9 中，我們更進一步，創造了一個函數，即 xe^x 的反函數，因為手上已有的函數不足以解一族很自然的方程。你計算機上的每一個函數，當初都是這樣誕生的新事物。

4. 嚴謹論證與不等式的工具箱

方法	用於何處
檢查每一步和每一個假設	似是而非的錯誤論證（3.1）
局部性質與整體性質	駐點與無窮遠處的極限（3.1）
前推－後推歸納法	AM–GM（3.2）
補上額外的項	AM–GM（3.2、3.3）
琴生不等式	三角形；AM–GM（3.3）
比較面積	楊氏不等式（3.4）
分離變量，歸結為單一函數	π^e 與 e^π 的比較（3.5）
三角不等式、等號成立的情況	馬爾可夫不等式（3.6）
試驗函數	$\pi^2 \leq 10$（3.7）
函數方程	e^x 和 $\tanh x$（3.8）
反函數與代換	W 函數（3.9）

下一步

- 在 Section 3.7 中略窺一二的變分法（calculus of variations），是 Section 4.7 的主題。
- tanh 及其加法定律背後的幾何，見於 Section 7.6。
- 第 4 章會以本章這種審慎的態度看待積分（integration）本身，由積分究竟是甚麼談起。

如果你想在入讀大學前再進一步

- **J. Michael Steele, *The Cauchy–Schwarz Master Class***：不等式的最佳入門書，全書由題目構成。
- **Daniel Velleman, *How to Prove It***：有系統地教你寫出正確證明的指南。
- **Michael Spivak, *Calculus***：由根基建立起來的嚴謹微積分；不少劍橋學生在入學前的那個暑假讀這本書。
- **E. A. Maxwell, *Fallacies in Mathematics***：用來練習找出哪裏出了錯。

馮紐曼（Von Neumann）曾打趣說，在數學裏你不是理解事物，只是習慣了它們。這話不無道理：本章那些精確的定義和嚴謹的論證，起初或會令人覺得陌生。但正是習慣了它們，一個懂得計算的學生才會變成一個能夠確信無誤的數學家。$1 = 2$ 的證明再也騙不到你；數學學位的目標，就是讓每一個更隱蔽的錯誤都同樣一眼可見。

「延伸練習」答案

以下只是簡短的答案和提示，並非完整解答：這裏只列出結果和關鍵步驟，方便你核對自己的做法。請先認真做過每道練習，然後才看答案。

3.1 找出錯處

1. 當 x 由 -1 變到 1 時會經過 0，而 $t = 1/x$ 在該處沒有定義：這個代換（substitution）在區間上並不連續（continuous）。正確的結果是 $\int_{-1}^{1} \frac{\mathrm{d}x}{(1+x^2)^2} = \frac{1}{2} + \frac{\pi}{4}$（用 $x = \tan\theta$），而 $\int_{-1}^{1} \frac{-t^2}{(1+t^2)^2}\,\mathrm{d}t = \frac{1}{2} - \frac{\pi}{4}$。

2. (a) 正確：兩邊的對數（logarithm）都是 $\ln a \ln b$。(b) 錯誤：當 $\theta = 0$ 時，$\cos 0 = 1 \neq \sin 1$。（事實上兩者永不相等，因為 $|\sin\theta \pm \cos\theta| \leq \sqrt{2} < \frac{\pi}{2}$。）(c) 錯誤：非常數的多項式（polynomial）是無界的，而一個常數不可能同時與 1 和 -1 相差不超過 10^{-6}。(d) 正確：由算術—幾何平均不等式（AM–GM），$x^4 + x^{-4} \geq 2$，等號在 $x = 1$ 時成立。

3. 由 $n = 1$ 推到 $n = 2$ 的一步並不成立：兩匹馬當中的「前 1 匹」與「後 1 匹」並不重疊。

4. $\int_{\varepsilon}^{1} x^{-1/2}\,\mathrm{d}x = 2 - 2\sqrt{\varepsilon} \to 2$，所以由對稱性，該積分（作為極限）存在，其值為 4。

3.2 倒過來的 AM–GM

1. 設 $a_1, \ldots, a_{n-1}$ 的平均值為 A，幾何平均（geometric mean）為 G，再加入 A 作為第 n 個數。平均值仍然是 A，所以 $(G^{n-1}A)^{1/n} \leq A$，由此得 $G \leq A$。

2. 這 $n+1$ 個數的平均值為 $1 + \frac{1}{n+1}$，幾何平均為 $(1 + \frac{1}{n})^{n/(n+1)}$，而且它們並非全部相等，所以不等式是嚴格的；把兩邊同時取 $n+1$ 次方即可。

3. 各個面的面積為 ab、bc、ca，而 $S = 2(ab + bc + ca)$。由 AM–GM 得 $\frac{S}{6} \geq (abc)^{2/3} = V^{2/3}$，等號在 $ab = bc = ca$ 時成立，即正方體。

4. 對 $1/a_i$ 應用 AM–GM：$\frac{1}{n}\sum \frac{1}{a_i} \geq \frac{1}{G}$。再取倒數。

3.3 凸性的威力

1. 若 $\mathrm{f}'' < 0$，斜率遞減，所以曲線位於每條弦（chord）的上方。取 $t = \frac{1}{3}$：$\mathrm{f}\big(\frac{u+v+w}{3}\big) \geq \frac{1}{3}\mathrm{f}(u) + \frac{2}{3}\mathrm{f}\big(\frac{v+w}{2}\big) \geq \frac{1}{3}\big(\mathrm{f}(u) + \mathrm{f}(v) + \mathrm{f}(w)\big)$。

2. $(\ln\sin x)'' = -\operatorname{cosec}^2 x < 0$，所以 $\ln\sin$ 在 $(0,\pi)$ 上是凹（concave）函數。由延森不等式（Jensen's inequality）得 $\frac{1}{3}\ln(\sin A \sin B \sin C) \leq \ln\sin\frac{\pi}{3}$，所以該乘積最大為 $\big(\frac{\sqrt{3}}{2}\big)^3$。

3. 在 $(0,\frac{\pi}{2})$ 上 $(\tan x)'' = 2\sec^2 x \tan x > 0$；由延森不等式得 $\tan A + \tan B + \tan C \geq 3\tan\frac{\pi}{3} = 3\sqrt{3}$。

4. $(x^2)'' = 2 > 0$；由延森不等式即得所求的不等式，它說明 $\frac{1}{n}\sum a_i^2 - \bar{a}^2$（即方差（variance））是非負的。

3.4 以面積證明不等式

1. 取 $\mathrm{f} = \sin$，定義在 $[0,\frac{\pi}{2}]$ 上，等號在 $b = \mathrm{f}(a)$ 時成立；取 $a = \arcsin t$、$b = t$：$t\arcsin t = (1 - \cos(\arcsin t)) + \int_0^t \arcsin y\,\mathrm{d}y$。把所得結果的右邊求導，得 $\arcsin t + \frac{t}{\sqrt{1-t^2}} - \frac{t}{\sqrt{1-t^2}} = \arcsin t$。

2. 在楊氏不等式（Young's inequality）中取 $p = q = 2$，得 $ab \leq \frac{1}{2}(a^2 + b^2)$。若 $\sum a_i^2 = \sum b_i^2 = 1$，逐項相加得 $\sum a_i b_i \leq 1$。一般情況下，把這個結果應用於 $a_i/\sqrt{\sum a_j^2}$ 和 $b_i/\sqrt{\sum b_j^2}$。

3. $\int_0^a (e^x - 1)\,\mathrm{d}x = e^a - 1 - a$，而 $\int_0^b \ln(1+y)\,\mathrm{d}y = (1+b)\ln(1+b) - b$。等號在 $b = e^a - 1$ 時成立。

4. $t = 1$：1.571 對 1；$t = 2$：0.524 對 0.268；$t = 10$：0.1002 對 0.0501。當 t 很大時，$t - \sqrt{t^2-1} = \frac{1}{t+\sqrt{t^2-1}} \approx \frac{1}{2t}$。這個界來自 $a = 0$ 的情況，此時 $(*)$ 只是說 $0 \leq \int_0^b \arcsin y\,\mathrm{d}y$；當 b 很小時，這塊面積約為 $\frac{b^2}{2}$，是 $b\arcsin b \approx b^2$ 的一半，把它捨去便損失了因子 2。

3.5 π^e 還是 e^π？

1. $n^m = m^n$ 即 $h(n) = h(m)$。h 在 $(1,e)$ 上遞增，其後遞減，所以兩數之中有一個小於 e：它必定是 2（因為 $h(1) = 0$ 這個值不會再次取到），於是另一個是 4。唯一的一對是 $\{2,4\}$。

2. $e < 3 < \pi$，而 h 在該處遞減，所以 $h(3) > h(\pi)$，即 $3^\pi > \pi^3$。$\frac{9}{4}$ 和 $\sqrt{5}$ 都小於 e，而 $\frac{9}{4} > \sqrt{5}$，所以 $h(\frac{9}{4}) > h(\sqrt{5})$，由此得第二個結果。對於 $1 < x < 2$：由於 $3 < x+2 < 4$，有 $h(x) < h(2) = h(4) < h(x+2)$，所以 $x^{x+2} < (x+2)^x$。最後，

$9^{\sqrt{2}} = 3^{2\sqrt{2}}$，而 $\sqrt{2}^{9} = 16\sqrt{2} = (2\sqrt{2})^3$。因為 $e^2 < 8$，所以 $e < 2\sqrt{2} < 3$，而 h 在該處遞減，故 $h(2\sqrt{2}) > h(3)$，由此得 $(2\sqrt{2})^3 > 3^{2\sqrt{2}}$：$\sqrt{2}^{9}$ 較大。

3. $e^{\pi/e-1} > \frac{\pi}{e}$，所以 $e^{\pi/e} > \pi$；再把兩邊取 e 次方。

3.6 斜率的界

1. 兩邊都是二次多項式，並且在 $x = -1, 0, 1$ 處取值相同，所以兩者相等（拉格朗日插值（Lagrange interpolation））。求導得 $\mathrm{f}(-1)(x - \frac{1}{2}) - 2\mathrm{f}(0)x + \mathrm{f}(1)(x + \frac{1}{2})$。

2. 在 $[-1, 1]$ 上 $|\mathrm{f}'(x)| \le |x - \frac{1}{2}| + 2|x| + |x + \frac{1}{2}| \le 4$，等號只在 $x = \pm 1$ 時成立，而且必須 $\mathrm{f}(1)$、$\mathrm{f}(-1)$、$-\mathrm{f}(0)$ 全部等於 1 或全部等於 -1。二次多項式由這三個值唯一確定，於是得 $\pm(2x^2 - 1)$。

3. 由 $T_3(\cos\theta) = \cos 3\theta$ 得 $4x^3 - 3x$，而 $T_3'(1) = 12 - 3 = 9$。由 $\cos(n+1)\theta + \cos(n-1)\theta = 2\cos\theta\cos n\theta$，得 $T_{n+1} = 2xT_n - T_{n-1}$，以及 $T_4 = 8x^4 - 8x^2 + 1$。

4. $-\sin\theta\, T_n'(\cos\theta) = -n\sin n\theta$。當 $\theta \to 0$ 時，$\frac{n\sin n\theta}{\sin\theta} \to n^2$。

3.7 $\pi^2 \le 10$

1. $\int_0^\pi \mathrm{f}^2 = \frac{\pi^9}{630}$，而 $\int_0^\pi (\mathrm{f}')^2 = \frac{2\pi^7}{105}$，所以由 $(*)$ 得 $\pi^2 \le 12$。與 $\sin x$ 相比，這個四次多項式在兩端附近過於平坦，在中間又過於尖聳，所以它的斜率與大小之比離最優值較遠。

2. $\int_0^\pi (\sin x + \sin 2x)^2 = \frac{\pi}{2} + \frac{\pi}{2} = \pi$（交叉項的積分為 0），而 $\int_0^\pi (\cos x + 2\cos 2x)^2 = \frac{\pi}{2} + 2\pi = \frac{5\pi}{2}$。

3. 設 $\mathrm{g}(x) = \mathrm{f}(\pi x/L)$。左邊按 $\frac{L}{\pi}$ 的比例縮放，而 $\int(\mathrm{g}')^2$ 則按 $\frac{\pi}{L}$ 的比例縮放。對弦線而言，最低頻率與 $\frac{1}{L}$ 成正比：長度減半，音高便加倍。

4. 界 $\pi^2 \le \int(\mathrm{f}')^2 / \int \mathrm{f}^2$（採用適當的單位）在 $\lambda \approx 1.133/\pi^2$ 時取得最小值，由此得 $\pi^2 \le 9.86975$，而真確值為 9.86960。

3.8 由加法公式決定的函數

1. 代入 $x = y = 0$ 得 $\mathrm{g}(0) = 0$。對 y 求導並取 $y = 0$，得 $\mathrm{g}'(x) = k(1 + \mathrm{g}(x)^2)$，所以 $\frac{\mathrm{d}}{\mathrm{d}x}\arctan\mathrm{g}(x) = k$，於是 $\mathrm{g}(x) = \tan kx$。

2. 對 y 求導得 $\mathrm{f}'(x+y) = \mathrm{f}'(y)$，所以 f' 是常數 k；又 $\mathrm{f}(0) = 0$（代入 $x = y = 0$）。

3. $\mathrm{f}(x) = \mathrm{f}(x/2)^2 \ge 0$，而且 f 永不為 0（否則它處處為 0）。用歸納法（induction）得 $\mathrm{f}(n) = a^n$；由 $\mathrm{f}(1/n)^n = \mathrm{f}(1) = a$ 及 $\mathrm{f}(1/n) > 0$ 得 $a^{1/n}$；於是 $\mathrm{f}(m/n) = \mathrm{f}(1/n)^m$。若沒有連續性，函數在無理數處的值便不受限制：存在一些利用「哈默

爾基（Hamel basis）」構造出來的奇怪解，它們並不是 a^x。

4. 每次點火都使快度（rapidity）增加 $\phi = \operatorname{artanh}\frac{1}{2}$，所以點火 k 次之後，速率為 $\tanh k\phi = \frac{3^k-1}{3^k+1}$：$0.8c$、$\frac{13}{14}c \approx 0.929c$ 和 $\frac{40}{41}c \approx 0.976c$。

3.9 創造一個新函數

1. 當 $x = -1$ 時，$x^2 - 2^x = \frac{1}{2} > 0$；當 $x = 0$ 時其值為 -1：所以有一個負根。設 $x = -2u/\ln 2$，則 $x^2 = 2^x$（其中 $x < 0$，取 $x = -2^{x/2}$）化為 $ue^u = \frac{\ln 2}{2}$，所以 $u = \mathrm{g}\left(\frac{\ln 2}{2}\right)$，而 $x \approx -0.7667$。

2. $\mathrm{g}'e^{\mathrm{g}}(1+\mathrm{g}) = 1$，而 $e^{\mathrm{g}} = x/\mathrm{g}$。當 $x = -\frac{1}{e}$ 時，$\mathrm{g} = -1$，斜率為無限大（圖像在該處是垂直的）。

3. 設 $x = we^w$，則 $\mathrm{d}x = (1+w)e^w\,\mathrm{d}w$，而 w 由 0 變到 1：$\int_0^1 w(1+w)e^w\,\mathrm{d}w = \left[(w^2-w+1)e^w\right]_0^1 = e-1$。

4. $T = \sqrt{2}^{\,T}$ 即 $\frac{\ln T}{T} = \frac{\ln 2}{2}$，所以 $T = 2$ 或 4。各項 $1.414, 1.633, 1.761, \ldots$ 遞增，並且一直小於 2（若 $t_n < 2$，則 $t_{n+1} < \sqrt{2}^{\,2} = 2$），所以 $T = 2$；不動點（fixed point）4 是排斥（repelling）的。

章

4

嚴謹的微積分

積分是分析、幾何與數論交匯之處，數學中一些最優美的公式亦在此誕生。

本章內容

4.1	甚麼是積分？	STEP III 1989
4.2	由一個積分建立三角學	STEP III 2017
4.3	22/7 大於 π	STEP I 1997
4.4	沃利斯、Beta 與二分之一的階乘	STEP II 1995
4.5	大二學生之夢	STEP III 2009
4.6	沒有公式的積分	STEP III 2016
4.7	變分法	STEP III 2014
4.8	切球形麵包	STEP I 2001
4.9	Holditch 定理	STEP III 2011
A	練習答案	第 305 頁

不只是原函數

> 那些沒有導數的函數，實在是可悲的禍害，我懷着驚恐與厭惡避之則吉。
>
> — 埃爾米特（Charles Hermite），1893 年的一封信

問一位中學生積分（integration）是甚麼，你多半會聽到「微分的逆運算」。這是個了不起的事實，稱為*微積分基本定理*（fundamental theorem of calculus），它令你在中學遇到的大部分積分都變成例行公事。但它是一條關於積分的定理，而不是積分的定義，而且它把積分之所以有趣的大部分原因都遮蓋了。

積分比微分（differentiation）早出現約二千年。阿基米德（Archimedes）把圖形切成薄片再求和，求出了拋物線弓形的面積和球的表面積，那時還遠遠未有人想到導數（derivative）。而且積分能到達微分無法到達的地方。每個合理的函數都可以按機械式的法則求導，但大多數函數卻*無法*用公式積分出來，這就引出一個自然的問題：甚麼時候可以？不可以的時候又該怎麼辦？積分可以定義新的函數，證明數與數之間的不等式，計算無法用其他方式描述的面積和體積，還能給出無窮和與無窮乘積的精確值。本章談的就是這一切。

積分究竟是甚麼

本章從頭說起，由定義開始（Section 4.1）。積分是許多幼長矩形之和的極限（limit）；單憑這個定義、不用原函數（antiderivative）去計算 $\int_0^a x^5\,\mathrm{d}x$，已是相當費工夫的事，而這會令你懂得珍惜基本定理。有了精確的定義，隨之而來的是一個精確的問題：哪些函數是可積（integrable）的？這個問題的答案，以及它所引發的種種難題，在二十世紀初催生了勒貝格（Lebesgue）的積分理論，那正是現代概率論的基礎。

用來證明的積分

本章一些最出人意表的結果，是用積分去證明一些看來與面積毫無關係的事實。某個正值函數的積分恰好等於 $\frac{22}{7}-\pi$，於是一行就證明了 $\pi<\frac{22}{7}$（Section 4.3）。$\sin t$ 的冪引出沃利斯（Wallis）關於 π 的無窮乘積（infinite product），以及 $\left(\frac{1}{2}\right)!=\frac{\sqrt{\pi}}{2}$ 這個值（Section 4.4）。x^x 的積分化成了和式 $\sum(-1)^{n+1}n^{-n}$（Section 4.5）。而僅僅一個積分 $\int_0^x \frac{\mathrm{d}u}{1+u^2}$，就足以從零開始建立整套三角學（Section 4.2）。

你在大學會遇到甚麼

大學一年級的*分析*（Analysis）課程會從積分的定義出發，嚴格證明微積分基本定理。其後的課程向多個方向發展。*測度論*（measure theory）和*勒貝格積分*（Lebesgue integral）令積分與極限配合得順暢無礙，並且是概率論的根基。*向量微積分*（vector calculus）在曲線、曲面和立體上積分，而格林(Green)定理和斯托克斯(Stokes)定理則把它們聯繫起來(Section 4.9 可讓你淺嘗一下)。*複分析*(complex analysis)取道複平面來計算實積分。*變分法*(calculus of variations)問的是哪個函數令某個積分盡可能小(Section 4.7)，並由此通向物理學中的最小作用量原理（principle of least action）。而*微分代數*（differential algebra）則解釋了哪些積分根本有公式可言（Section 4.6）。

貫穿本章的問題

三大問題

1. **積分是甚麼？哪些函數有積分？**以和的極限作出的定義，與原函數有何關連？
2. **積分能證明甚麼？**一塊面積怎能告訴我們有關 π、階乘和無窮和的事？
3. **沒有公式的時候，我們可以做甚麼？**我們怎樣知道某個積分沒有公式？即使如此，我們還能算出甚麼？

本章的思想

§	課題	思想
4.1	積分的定義	下和與上和（lower and upper sums）；裂項相消（telescoping）
4.2	由 $\int \frac{\mathrm{d}u}{1+u^2}$ 建立三角學	保持積分不變的代換（substitution）
4.3	$\frac{22}{7} > \pi$	把不等式寫成某個正值量的積分
4.4	沃利斯與 $\left(\frac{1}{2}\right)!$	遞降公式（reduction formulae）、不變量（invariant）、夾逼（squeeze）
4.5	$\int_0^1 x^x \,\mathrm{d}x$	級數展開與逐項積分（term-by-term integration）
4.6	沒有公式的積分	藉多項式代數證明不可能
4.7	變分法	在積分之內配方（completing the square）
4.8	阿基米德的帽盒定理（hat-box theorem）	化簡被積函數（integrand），幾何意義隨之顯現
4.9	霍爾迪奇（Holditch）定理	由邊界積分求面積；消元（elimination）

如何閱讀本章

有好幾節涉及繁重但例行的計算：多項式除法、分部積分法（integration by parts）、展開乘積。請先拿起紙筆親自做一遍，然後才閱讀正文。接着找出令計算得以成功的那個思想：一個不變量、一種對稱、一次相消。計算是考官要檢查的東西；思想才是你會記住並再次運用的東西。

4.1 甚麼是積分？

題目

STEP III 1989，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/89-s3-q10`

引子

你做積分已經很多年了。你知道 $\int_0^a x^5\,\mathrm{d}x = \frac{1}{6}a^6$，因為「積分是微分的逆運算」。但那是一條定理，不是定義。積分究竟是甚麼？

老實的答案是：積分是面積，定義為許多細長矩形面積之和的極限（limit）。這道題要你單憑這個定義計算一個積分：不用原函數（antiderivative），不用微積分基本定理（fundamental theorem of calculus），只靠把矩形加起來再取極限。這比你想像的要難，而這正好說明，令這件事變得輕而易舉的微積分基本定理，是何等了不起的成果。

動筆之前

想一想 第 (iii) 部分的兩個和，都是闊度為 $\frac{a}{n}$ 的矩形的面積之和。對於遞增函數（increasing function）f，哪一個和低估了曲線下的面積，哪一個高估了？

第一個和取每一條長條左端的高度，即 $x = \frac{ra}{n}$，其中 $r = 0, \ldots, n-1$。對遞增函數來說，那是長條上的最低點，所以這些矩形全在曲線之下：這是下和（lower sum）。第二個和取右端，即 $r = 1, \ldots, n$，所得的矩形會突出於曲線之上：這是上和（upper sum）（圖 4.1.1）。定義說明：當長條愈來愈窄時，若兩個和趨向同一個極限，積分便存在。

對於 $\mathrm{f}(x) = x^5$，兩個和都涉及 $\sum r^5$。第 (i) 和第 (ii) 部分分別給出這個和的上界和下界。

逐步解題

階段 1　可以裂項相消的和

想一想　設 $P(r) = r(r+1)(r+2)(r+3)(r+4)(r+5)$，即六個連續整數之積。$P(r) - P(r-1)$ 是甚麼？

$P(r)$ 與 $P(r-1) = (r-1)r(r+1)(r+2)(r+3)(r+4)$ 有五個共同因子 $r(r+1)(r+2)(r+3)(r+4)$，所以

$$\begin{aligned} P(r) - P(r-1) &= r(r+1)(r+2)(r+3)(r+4)[(r+5) - (r-1)] \\ &= 6\,r(r+1)(r+2)(r+3)(r+4). \end{aligned}$$

由 $r = 1$ 加到 n，右方各項裂項相消（telescoping），只餘下 $P(n) - P(0) = P(n)$。所以

$$\sum_{r=1}^{n} r(r+1)(r+2)(r+3)(r+4) = \tfrac{1}{6}P(n) = \tfrac{1}{6}n(n+1)(n+2)(n+3)(n+4)(n+5).$$

■

由於對所有 $r \geq 1$ 都有 $r^5 < r(r+1)(r+2)(r+3)(r+4)$，不等式隨即成立。 ■

階段 2　下界

停一停，試一試　把第 1 階段的做法改用於五個向下數的連續整數之積：$r(r-1)(r-2)(r-3)(r-4)$。求它由 $r = 0$ 到 $n-1$ 的和，並與 r^5 比較。

設 $Q(r) = (r-5)(r-4)(r-3)(r-2)(r-1)r$。則 $Q(r+1) - Q(r) = (r-4)(r-3)(r-2)(r-1)r[(r+1) - (r-5)] = 6\,(r-4)(r-3)(r-2)(r-1)r$。由 $r = 0$ 加到 $n-1$，裂項相消後得 $Q(n) - Q(0) = Q(n)$，所以

$$\sum_{r=0}^{n-1} (r-4)(r-3)(r-2)(r-1)r = \tfrac{1}{6}(n-5)(n-4)(n-3)(n-2)(n-1)n.$$

現在逐項比較。當 $r = 0$ 時，兩邊都是 0。當 $1 \leq r \leq 4$ 時，乘積含有一個零因子，而 $r^5 > 0$。當 $r \geq 5$ 時，每個因子都是正數，且不大於 r，其中有些更嚴格小於它，所以乘積小於 r^5。由於 $n > 1$，和式包含 $r = 1$ 一項，該處不等式嚴格成立，所以

$$\sum_{r=0}^{n-1} r^5 > \tfrac{1}{6}(n-5)(n-4)(n-3)(n-2)(n-1)n. \qquad ■$$

階段 3　積分

想一想 把 $\mathrm{f}(x)=x^5$ 的下和與上和用 $\sum r^5$ 表示。分母中出現 n 的幾次方？

取 $\mathrm{f}(x)=x^5$，

$$L_n=\sum_{r=0}^{n-1}\frac{a}{n}\left(\frac{ra}{n}\right)^5=\frac{a^6}{n^6}\sum_{r=0}^{n-1}r^5,\qquad U_n=\frac{a^6}{n^6}\sum_{r=1}^{n}r^5.$$

由第 (i) 和第 (ii) 部分，並因為 $L_n<U_n$，

$$\frac{a^6}{6}\cdot\frac{(n-5)(n-4)\cdots n}{n^6}\;<\;L_n\;<\;U_n\;<\;\frac{a^6}{6}\cdot\frac{n(n+1)\cdots(n+5)}{n^6}.$$

兩個分數各自是六個因子之積，這些因子形如 $\frac{n-3}{n}=1-\frac{3}{n}$ 或 $\frac{n+3}{n}=1+\frac{3}{n}$，每一個都趨向 1。所以外側兩個界都趨向 $\frac{a^6}{6}$，根據夾逼定理（squeeze），L_n 和 U_n 亦然。兩個極限都存在而且相等，所以按定義

$$\int_0^a x^5\,\mathrm{d}x=\tfrac{1}{6}a^6.\qquad\blacksquare$$

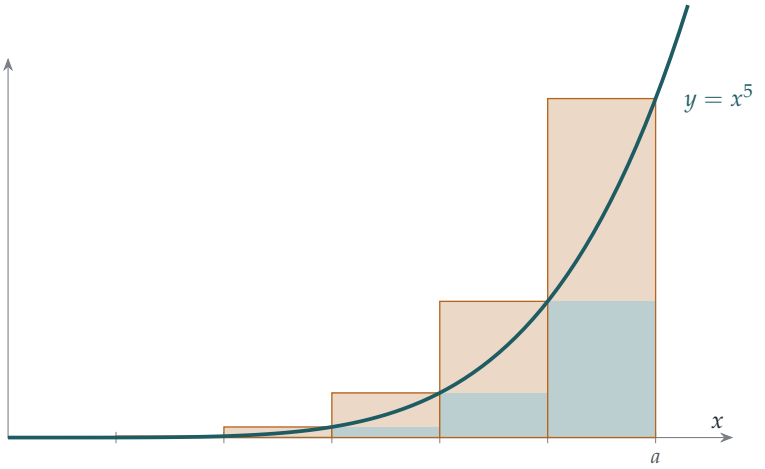


圖 4.1.1 x^5 在 $[0,a]$ 上的下和與上和，分成 $n=6$ 條長條。下和是藍綠色的面積（曲線下的矩形）；上和則再加上琥珀色的部分。兩者之差正好是最後、最高的那個矩形的面積 $\frac{a}{n}\,\mathrm{f}(a)$，當 n 增大時它趨向 0。

回顧

- **又是裂項相消。**連續整數之積的和一下子收攏起來，與第 2 章的和與積如出一轍。
- **先定界，再夾逼。**我們從沒有求出 $\sum r^5$ 的準確值（它是 $\frac{1}{12}n^2(n+1)^2(2n^2+2n-1)$）；有兩邊的界已經足夠。
- **定義擔起全部工作。**原函數由始至終沒有出現過。一切都來自積分作為和的極限這個定義。

背後的數學

黎曼的定義

把細長條加起來求面積的想法，可追溯至阿基米德（Archimedes）：他在公元前 250 年左右，正是用這個論證的一個版本求出拋物線下的面積。牛頓（Newton）和萊布尼茲（Leibniz）把積分變成微分的逆運算，其後一個半世紀，大多數人都是這樣理解積分的。1854 年，黎曼（Bernhard Riemann）提出了大學課程至今仍用作起點的定義：把 $[a, b]$ 分成許多小段，構作（闊度）$\times$（f 在該小段內某處的高度）之和；若小段不斷縮小時，所有這樣的和都趨向同一個極限，便說 f 可積（integrable）。本題的版本採用等闊長條和左右端點，是一個適用於遞增函數的特殊情況；達布（Jean-Gaston Darboux）的「上和與下和」則把一般情況處理得乾淨利落。

並非每個函數都有積分

有了精確的定義，便可以問哪些函數可積。閉區間上的每個連續函數都可積，每個遞增函數也可積。但試考慮狄利克雷（Dirichlet）函數：若 x 是有理數，$\mathrm{f}(x) = 1$；若是無理數，則為 0。每一條長條無論多窄，都同時包含有理數和無理數，所以每個上和都是 1，每個下和都是 0。兩者永不一致，因此 $\int_0^1 \mathrm{f}$ 在黎曼的意義下並不存在。1902 年，勒貝格（Henri Lebesgue）發明了一種更強大的積分，能夠處理這個函數（它的勒貝格積分（Lebesgue integral）是 0），而且在取極限時表現好得多。勒貝格積分是現代概率論的基礎，你會在大學二年級或三年級遇上它。

離散微積分

第 1 階段的裂項相消，其實是喬裝了的微積分。「上升階乘」（rising factorial）$r(r+1)\cdots(r+k-1)$ 在求和之下的表現，與 x^k 在積分之下的表現完全一樣：求和令因子的數目增加一個，並除以新的因子數目，正如 $\int x^k \,\mathrm{d}x = \frac{x^{k+1}}{k+1}$。和與積分、差分與導數之間的這種對應，稱為*有限差分演算*（calculus of finite differences），是計算和式的有力工具。

思維技巧

從定義出發

題目要求你「用定義」證明某事時，就只用定義：不用任何依賴於待證結果的定理，也不用熟悉的捷徑。

這樣做出奇地難，卻讓你確切看到那些捷徑的價值。在大學，每一科的頭幾個星期都是這樣進行的：先有定義，然後才有定理，讓你不必再從定義出發。

延伸練習

1 **把極限化為積分。**試做 STEP III 1999，第 3 題（`step.maths.org/questions/99-s3-q3`）。這道題把和的極限化為積分：在每個和式中，找出曲線下的矩形。

2 **阿基米德的拋物線。**利用 $\sum_{r=1}^{n} r^2 = \frac{1}{6}n(n+1)(2n+1)$ 和第 (iii) 部分的定義，證明 $\int_0^a x^2\,\mathrm{d}x = \frac{1}{3}a^3$。

3 **上升階乘。**證明對每個正整數 k，都有 $\displaystyle\sum_{r=1}^{n} r(r+1)\cdots(r+k-1) = \frac{n(n+1)\cdots(n+k)}{k+1}$。

4 **準確的和。**利用恆等式 $r(r+1)(r+2)(r+3)(r+4) = r^5+10r^4+35r^3+50r^2+24r$，以及 $\sum r^k$ 在 $k \le 4$ 時的已知公式，求出 $\sum_{r=1}^{n} r^5$ 的準確值。

延伸閱讀

- **David Bressoud, *A Radical Approach to Lebesgue's Theory of Integration***（Cambridge University Press）。由黎曼積分走到勒貝格積分的故事，以逼出每一步發展的問題為線索娓娓道來。
- **Michael Spivak, *Calculus***（Cambridge University Press）。第 13 章以上和與下和定義積分，與本節完全一樣。
- **Graham, Knuth & Patashnik, *Concrete Mathematics***（Addison-Wesley）。第 2 章建立有限差分演算，並討論上升階乘和下降階乘。

4.2 由一個積分建立三角學

題目

STEP III 2017，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/17-s3-q6`

引子

π 究竟是甚麼？在中學，它是圓周與直徑之比，而 sin 和 cos 則用三角形來定義。這對幾何來說已經足夠，卻留下了一些棘手的問題。沿着曲線的「長度」到底是甚麼？為甚麼 $\sin x$ 的麥克勞林級數（Maclaurin series）偏偏是那個樣子？我們又怎樣知道加法公式（addition formulas）對所有角都成立，而不只限於畫得出來的角？

這道題展示了一條截然不同的入門途徑。忘掉三角形和圓吧。只從一個積分 $\mathrm{T}(x) = \int_0^x \frac{\mathrm{d}u}{1+u^2}$ 出發，它所需要的不過是 Section 4.1 中積分的定義。把 π 定義為 $2\mathrm{T}_\infty$。然後只用代換法（substitution），證明 $\mathrm{T}(\sqrt{3}) = \frac{\pi}{3}$ 和 $\mathrm{T}(\sqrt{2}-1) = \frac{\pi}{8}$，也就是你所熟悉的 $\arctan\sqrt{3} = 60^\circ$ 和 $\tan 22\frac{1}{2}^\circ = \sqrt{2}-1$。三角學就這樣從無到有地重建起來。

動筆之前

你心裏其實知道 $\mathrm{T}(x) = \arctan x$ 和 $\mathrm{T}_\infty = \frac{\pi}{2}$。題目不准你使用這些事實，而這正是用意所在：一切都必須由積分推導出來。不過，這點心照不宣的知識是很有用的嚮導。

想一想 第 (i) 和第 (ii) 部分分別對應哪些三角恆等式（trigonometric identities）？（想想 $\arctan x + \arctan\frac{1}{x}$，以及 $\tan(A+B)$ 的公式。）

第 (i) 部分就是 $\arctan x + \arctan\frac{1}{x} = \frac{\pi}{2}$。第 (ii) 部分則是喬裝了的加法公式 $\tan(A+B) = \frac{\tan A+\tan B}{1-\tan A\tan B}$：設 $u = \tan A$ 和 $a = \tan B$，則 $v = \tan(A+B)$。這道題不靠任何幾何，單憑積分就證明了這些恆等式。

逐步解題

階段 1　反射 $u \mapsto 1/u$

在 $\mathrm{T}(x)$ 中代入 $u = 1/w$。於是 $\mathrm{d}u = -\mathrm{d}w/w^2$，而 $\dfrac{1}{1+u^2} = \dfrac{w^2}{w^2+1}$，所以

$$\mathrm{T}(x) = \int_{\infty}^{1/x} \frac{w^2}{w^2+1} \cdot \left(-\frac{1}{w^2}\right) \mathrm{d}w = \int_{1/x}^{\infty} \frac{\mathrm{d}w}{1+w^2} = \mathrm{T}_\infty - \mathrm{T}(x^{-1}). \qquad \blacksquare$$

換句話說，$\frac{1}{1+u^2}$ 之下由 0 到 x 的面積，等於由 $\frac{1}{x}$ 到 ∞ 的面積（圖 4.2.1）。

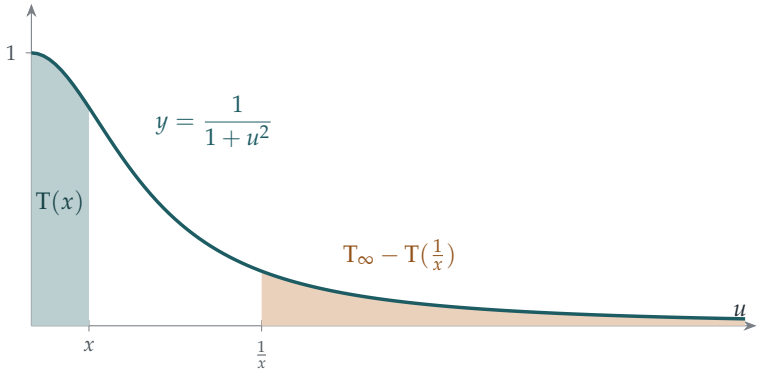


圖 4.2.1　代換 $u = 1/w$ 顯示 $\frac{1}{1+u^2}$ 之下的兩塊陰影面積相等：由 0 到 x 的面積（藍綠色），以及由 $\frac{1}{x}$ 起無限伸延的面積（琥珀色）。圖中 $x = \frac{1}{2}$。

階段 2　加法代換

想一想　用商法則（quotient rule）對 $v = \dfrac{u+a}{1-au}$ 求導。然後把 $1+v^2$ 化成單一分式，再作比較。

$$\frac{\mathrm{d}v}{\mathrm{d}u} = \frac{(1-au)+a(u+a)}{(1-au)^2} = \frac{1+a^2}{(1-au)^2},$$
$$1+v^2 = \frac{(1-au)^2+(u+a)^2}{(1-au)^2} = \frac{(1+a^2)(1+u^2)}{(1-au)^2}.$$

兩式相除，得 $\dfrac{1+v^2}{1+u^2} = \dfrac{1+a^2}{(1-au)^2} = \dfrac{\mathrm{d}v}{\mathrm{d}u}$。 $\blacksquare$

想一想　這表示 $\dfrac{\mathrm{d}u}{1+u^2} = \dfrac{\mathrm{d}v}{1+v^2}$。所以代換 $u \mapsto v$ 完全沒有改變被積函數（integrand）。那麼它對 $\mathrm{T}(x)$ 的積分上下限有甚麼影響？

當 u 由 0 變到 x 時（其中 $x < \frac{1}{a}$，故 $1-au>0$ 始終成立，而 v 連續且遞增），v 由 a 變到 $\frac{x+a}{1-ax}$。所以

$$\mathrm{T}(x)=\int_0^x \frac{\mathrm{d}u}{1+u^2}=\int_a^{(x+a)/(1-ax)} \frac{\mathrm{d}v}{1+v^2}=\mathrm{T}\!\left(\frac{x+a}{1-ax}\right)-\mathrm{T}(a). \qquad \blacksquare$$

階段 3 再多兩個恆等式

結合第 (i) 部分，並使用兩次：

$$\mathrm{T}(x^{-1})=\mathrm{T}_\infty-\mathrm{T}(x)=\mathrm{T}_\infty-\mathrm{T}\!\left(\frac{x+a}{1-ax}\right)+\mathrm{T}(a)=2\mathrm{T}_\infty-\mathrm{T}\!\left(\frac{x+a}{1-ax}\right)-\mathrm{T}(a^{-1}). \qquad \blacksquare$$

現在更換變數名稱：設 $y=1/x$ 和 $b=1/a$。條件 $x<1/a$ 變成 $1/y<b$，即 $y>1/b$，而

$$\frac{x+a}{1-ax}=\frac{1/y+1/b}{1-1/(by)}=\frac{b+y}{by-1}.$$

所以 $\mathrm{T}(y)=2\mathrm{T}_\infty-\mathrm{T}\!\left(\frac{y+b}{by-1}\right)-\mathrm{T}(b)$。 $\blacksquare$

階段 4 特殊值

> **停一停，試一試** 在最後一個恆等式中選取 y 和 b，使自變量 $\frac{y+b}{by-1}$ 剛好等於 y 本身。你會得到甚麼？

取 $y=b=\sqrt{3}$：這時 $y>\frac{1}{b}$，而 $\dfrac{\sqrt{3}+\sqrt{3}}{3-1}=\sqrt{3}$。恆等式變成 $\mathrm{T}(\sqrt{3})=2\mathrm{T}_\infty-2\mathrm{T}(\sqrt{3})$，所以

$$\mathrm{T}(\sqrt{3})=\tfrac{2}{3}\mathrm{T}_\infty. \qquad \blacksquare$$

> **想一想** 至於 $\sqrt{2}-1$，試在加法恆等式中取 $x=a=\sqrt{2}-1$。$\frac{2x}{1-x^2}$ 是多少？$\mathrm{T}(1)$ 又是多少？

取 $x=a=\sqrt{2}-1$（由於 $x^2=3-2\sqrt{2}<1$，故 $x<\frac{1}{a}$），加法恆等式給出 $2\mathrm{T}(x)=\mathrm{T}\!\left(\frac{2x}{1-x^2}\right)$。而 $1-x^2=2\sqrt{2}-2=2(\sqrt{2}-1)=2x$，所以 $\frac{2x}{1-x^2}=1$，於是 $2\mathrm{T}(\sqrt{2}-1)=\mathrm{T}(1)$。最後，在第 (i) 部分取 $x=1$，得 $\mathrm{T}(1)=\mathrm{T}_\infty-\mathrm{T}(1)$，所以 $\mathrm{T}(1)=\frac{1}{2}\mathrm{T}_\infty$。因此

$$\mathrm{T}(\sqrt{2}-1)=\tfrac{1}{4}\mathrm{T}_\infty. \qquad \blacksquare$$

回顧

- **保持被積函數不變的代換。**兩個關鍵步驟所用的代換，都令 $\frac{\mathrm{d}u}{1+u^2}$ 保持不變，只有積分上下限移動了。積分的這類「對稱性」（symmetry）威力強大。

- **不動點（fixed point）。**$\mathrm{T}(\sqrt{3})$ 的值，來自選取適當的數，使某個表達式等於它本身，於是恆等式變成了方程。
- **定義可以來自分析。**我們從沒有用過三角形，卻重新得到了 $\frac{\pi}{3}$ 和 $\frac{\pi}{8}$（其中 π 定義為 $2\mathrm{T}_\infty$）。

背後的數學

建構初等函數

在嚴格的分析課程中，函數 exp、sin、cos 和 ln 不能隨便假設存在；它們必須被建構出來，其性質也必須加以證明。標準的做法有三種。你可以用冪級數（power series）來定義它們（$\sin x = x - \frac{x^3}{3!} + \cdots$）。你可以把它們定義為微分方程（differential equation）的解，就如練習中那道 1986 年樣本試卷的題目（$\mathrm{f}' = \mathrm{f}$、$\mathrm{f}(0) = 1$ 定義了 exp）。你也可以像這裏一樣，用積分來定義反函數（inverse function）：$\ln x = \int_1^x \frac{\mathrm{d}u}{u}$ 和 $\arctan x = \int_0^x \frac{\mathrm{d}u}{1+u^2}$，然後定義 $\pi = 2\int_0^\infty \frac{\mathrm{d}u}{1+u^2}$。每種做法各有所長，而證明它們全都給出相同的函數，是一道令人滿足的練習。試與 Section 3.8 的函數方程（functional equation）做法比較。

圓以外：橢圓積分

積分的做法可以推廣，而三角形的做法卻做不到。1718 年，Giulio Fagnano 研究雙紐線（lemniscate）這條 8 字形曲線的弧長（arc length），由此引出 $\int_0^x \frac{\mathrm{d}u}{\sqrt{1-u^4}}$，並找到一個能把弧長加倍的代換，正如我們的代換把 T 加倍一樣。歐拉（Euler）在 1751 年讀到 Fagnano 的著作，找出了完整的加法公式；這最終經由阿貝爾（Abel）和雅可比（Jacobi），發展成*橢圓函數*（elliptic functions）的理論：橢圓函數就是這類積分的反函數，它們之於雙紐線，正如 sin 和 cos 之於圓。我們在高斯（Gauss）關於算術幾何平均（AGM）的工作中（Section 2.6）遇過同一個積分。橢圓函數如今是數論的核心；Section 1.9 的橢圓曲線（elliptic curves）就是它們在幾何上的投影。

計算 π

加法恆等式也是計算 π 的那些最著名公式的基礎。舉例說，反覆運用它可以證明 $\mathrm{T}(1) = 4\mathrm{T}(\frac{1}{5}) - \mathrm{T}(\frac{1}{239})$，這就是 1706 年的梅欽公式（Machin's formula）；配合級數 $\mathrm{T}(x) = x - \frac{x^3}{3} + \frac{x^5}{5} - \cdots$，它算出了 π 的 100 位數字，並在其後兩個半世紀一直是首選的方法。這是 Section 5.1 的主題。

思維技巧

積分的對稱性

尋找令被積函數保持不變的代換。它們會化爲該積分所定義的函數的恆等式。

這裏 $u \mapsto 1/u$ 和 $u \mapsto \frac{u+a}{1-au}$ 都保持 $\frac{\mathrm{d}u}{1+u^2}$ 不變，並分別給出反射公式和加法公式。同一個想法用於其他積分，正是特殊函數（special functions）理論的基礎。

延伸練習

1 **指數函數與餘弦。**試做 STEP II Specimen paper，第 6 題（`step.maths.org/questions/spec-s2-q6`）。這道題只憑指數函數和餘弦各自的微分方程，推導出它們的加法公式。（提示：對於餘弦，找出兩個理應為零的組合，並證明它們的平方和是常數。）

2 **反正弦。**試做 STEP III 2007，第 7 題（`step.maths.org/questions/07-s3-q7`）。這道題用處理本節積分的同一手法來處理反正弦（inverse sine），並以一個代換把兩者聯繫起來。

3 **一條早期的梅欽型公式。**利用加法恆等式證明 $\mathrm{T}(\frac{1}{2}) + \mathrm{T}(\frac{1}{3}) = \mathrm{T}(1) = \frac{1}{2}\mathrm{T}_\infty$。

4 T_∞ **是有限的。**證明對所有 $x \geq 1$ 都有 $\mathrm{T}(x) \leq 1 + \int_1^x \frac{\mathrm{d}u}{u^2} < 2$，解釋為甚麼當 $0 < x < 1$ 時同樣有 $\mathrm{T}(x) < 2$，並由單調收斂定理（monotone convergence theorem）（Section 2.2）推斷 T_∞ 存在。

延伸閱讀

- **Tom Körner,** ***Calculus for the Ambitious***（Cambridge University Press）。嚴格地建構初等函數，並討論各種不同的做法。
- **Walter Rudin,** ***Principles of Mathematical Analysis***（McGraw-Hill）。即著名的「baby Rudin」；第 8 章的引言只用寥寥數頁，便優雅地由冪級數建構出 exp、sin 和 cos。程度較深。
- **Victor Prasolov & Yuri Solovyev,** ***Elliptic Functions and Elliptic Integrals***（American Mathematical Society）。想知道雙紐線通往何處，可讀此書。程度較深。

4.3 22/7 大於 π

題目

STEP I 1997，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/97-s1-q6`

引子

一代又一代的學生都聽過 π「大約是 $\frac{22}{7}$」這個說法。這是一個非常好的近似值（approximation）：$\frac{22}{7} = 3.142857\ldots$，而 $\pi = 3.141592\ldots$。但 $\frac{22}{7}$ 究竟是偏大還是偏小？你又能否證明這一點，而不是單靠查閱 π 的各位數字？那些數字畢竟是要別人先計算出來的。

這道題給出了整個數學之中最優雅的證明之一。只需一個正函數的積分（integral），其值因而必定為正，而它竟然恰好等於 $\frac{22}{7} - \pi$。被積函數（integrand）為正，積分便為正，所以 $\frac{22}{7} > \pi$。整個證明只需寥寥數行，而且精巧得被人當作謎題印在數學 T 恤上。

動筆之前

想一想 為甚麼 $\int_0^1 \frac{\text{多項式}}{1+x^2}\,\mathrm{d}x$ 會算出 π 來？這種形式的積分之中，哪一個簡單的積分本身就給出 π？

$\int_0^1 \frac{\mathrm{d}x}{1+x^2} = \arctan 1 = \frac{\pi}{4}$。把一個多項式（polynomial）除以 $1+x^2$，會得到一個多項式商式（quotient）（其積分是一個有理數），再加上一個常數餘式（remainder）除以 $1+x^2$（其積分是 π 的有理數倍）。題目的第一部分正是要你做這個除法。

逐步解題

階段 1　多項式除法

先展開：$(1-x)^4 = 1 - 4x + 6x^2 - 4x^3 + x^4$，所以

$$x^4(1-x)^4 = x^8 - 4x^7 + 6x^6 - 4x^5 + x^4.$$

想一想 把它除以 x^2+1，可以用長除法（long division），也可以用比較係數（comparing coefficients）的方法。餘式是甚麼？

在 $(a_6x^6 + \cdots + a_0)(x^2+1) + b$ 中由最高次項開始比較係數：$a_6 = 1$；$a_5 = -4$；$a_4 + a_6 = 6$，所以 $a_4 = 5$；$a_3 + a_5 = -4$，所以 $a_3 = 0$；$a_2 + a_4 = 1$，所以 $a_2 = -4$；$a_1 + a_3 = 0$，所以 $a_1 = 0$；$a_0 + a_2 = 0$，所以 $a_0 = 4$；最後 $a_1 = 0$（x 的係數）及 $a_0 + b = 0$，所以 $b = -4$。即是說，

$$x^4(1-x)^4 = (x^6 - 4x^5 + 5x^4 - 4x^2 + 4)(x^2+1) - 4.$$

階段 2　計算積分

除以 $1+x^2$，得

$$\frac{x^4(1-x)^4}{1+x^2} = x^6 - 4x^5 + 5x^4 - 4x^2 + 4 - \frac{4}{1+x^2},$$

所以

$$\int_0^1 \frac{x^4(1-x)^4}{1+x^2}\,\mathrm{d}x = \frac{1}{7} - \frac{4}{6} + \frac{5}{5} - \frac{4}{3} + 4 - 4\cdot\frac{\pi}{4} = \frac{1}{7} - \frac{2}{3} + 1 - \frac{4}{3} + 4 - \pi.$$

由於 $-\frac{2}{3} - \frac{4}{3} = -2$，有理數部分是 $\frac{1}{7} + 3 = \frac{22}{7}$，因此積分等於 $\frac{22}{7} - \pi$。 ■

階段 3　估計積分的界

想一想 被積函數在 $(0,1)$ 上為正，所以立即得到 $\frac{22}{7} > \pi$。要得到 π 的下界（lower bound），我們需要這個積分的上界（upper bound）。$\frac{1}{1+x^2}$ 在 $[0,1]$ 上最大可以有多大？

在 $[0,1]$ 上，$1 \le 1+x^2 \le 2$，所以

$$\tfrac{1}{2}x^4(1-x)^4 \le \frac{x^4(1-x)^4}{1+x^2} \le x^4(1-x)^4,$$

而且除了在端點之外，不等式（inequality）都是嚴格成立的。

停一停，試一試 計算 $\int_0^1 x^4(1-x)^4\,\mathrm{d}x$，然後完成證明。

把展開式逐項積分，

$$\int_0^1 x^4(1-x)^4\,\mathrm{d}x = \frac{1}{9} - \frac{4}{8} + \frac{6}{7} - \frac{4}{6} + \frac{1}{5} = \frac{1}{630}.$$

（以 2520 為公分母：$280 - 1260 + 2160 - 1680 + 504 = 4$，而 $\frac{4}{2520} = \frac{1}{630}$。）所以

$$0 < \frac{22}{7} - \pi < \frac{1}{630},$$

亦即 $\frac{22}{7} > \pi > \frac{22}{7} - \frac{1}{630}$。 ■

以數值計算，$\frac{22}{7} - \frac{1}{630} = 3.14127\ldots$，所以我們已證明 $3.1412 < \pi < 3.1429$。若同時用上左邊的不等式，便有 $\frac{1}{1260} < \frac{22}{7} - \pi$，這把上界改進為 $\pi < 3.14207$。

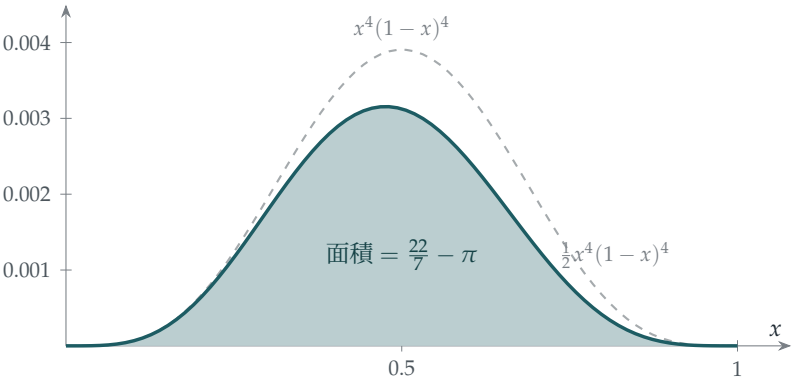


圖 4.3.1　被積函數 $\frac{x^4(1-x)^4}{1+x^2}$（藍綠色）為正，但數值極小：它的最大值約為 0.003。陰影部分的面積恰好是 $\frac{22}{7} - \pi \approx 0.00126$。被積函數介乎 $\frac{1}{2}x^4(1-x)^4$（點線）與 $x^4(1-x)^4$（虛線）之間，這兩條曲線下的面積分別是 $\frac{1}{1260}$ 和 $\frac{1}{630}$。

回顧

- **以正值證明不等式。**正函數的積分為正。一旦認出積分等於 $\frac{22}{7} - \pi$，不等式便立即成立。
- **被積函數愈小，界愈精確。**因子 $x^4(1-x)^4$ 在 $[0,1]$ 上極小（最多只有 $\frac{1}{256}$），這使積分很小，所得的界也就很精確。
- **有理數加上 π 的倍數。**正如 Section 2.9 所見，一個形如「有理數 + 某常數的有理數倍」的積分，若其值很小，便給出該常數的一個良好有理近似值。

背後的數學

一個有來歷的證明

阿基米德（Archimedes）約於公元前 250 年證明了 $\frac{223}{71} < \pi < \frac{22}{7}$，方法是計算圓內接和圓外切正 96 邊形的周界（可與 Section 2.5 中的多邊形比較）。積分證明則晚近得多。它由 D. P. Dalzell 於 1944 年發表，其後在 1968 年被選為美國大學本科生普特南數學競賽（Putnam competition）的題目，因而廣為人知。它的迷人之處在於完全不需要任何數值計算：不等式 $\frac{22}{7} > \pi$ 直接來自「正函數的積分為正」這個事實。

用同一思路得出更佳的近似值

這個方法可以推廣。以 $x^8(1-x)^8$ 取代 $x^4(1-x)^4$，同樣的除法給出

$$\int_0^1 \frac{x^8(1-x)^8}{1+x^2}\,\mathrm{d}x = 4\pi - \frac{188684}{15015},$$

由此證明 $\pi > \frac{47171}{15015} = 3.1415917\ldots$，準確至六位有效數字（significant figures）。2009 年，Stephen Lucas 找到了多族這類被積函數非負的積分，它們能證明關於其他著名近似值的不等式，其中包括準確至七位有效數字的 $\frac{355}{113}$。

近似值與無理性

為甚麼 π 會有 $\frac{22}{7}$ 和 $\frac{355}{113}$ 這樣好的分數近似值？每個無理數（irrational number）都有無窮多個分數 $\frac{p}{q}$ 與它相距不超過 $\frac{1}{q^2}$，而 π 碰巧在很早的階段就有幾個特別好的，因為它的連分數（continued fraction）$[3;7,15,1,292,\ldots]$ 含有 292 這個很大的項（見 Section 2.2）。像本題這樣的積分，若令其愈來愈小，便是另一類證明的起點：這類證明指出某些數不可能被分數逼近得太好，因而是無理數。Ivan Niven 那個著名的、只有一頁的 π 是無理數的證明（1947 年），用的就是 $\frac{x^n(1-x)^n}{n!}\sin(\pi x)$ 的積分，它們是 Section 2.9 中那些積分的三角函數近親。

思維技巧

把不等式寫成某個正值函數的積分

要證明 $A > B$，可找一個積分等於 $A - B$ 的正函數。

這樣便把一個不等式化為一個恆等式（identity）再加上一個顯然的事實。這是分析學（analysis）中最有用的技巧之一；同一思路，以平方和取代積分，就在

Section 3.2 中證明了算術平均—幾何平均不等式（AM–GM）。

延伸練習

1 **核對數字。**計算 $\frac{22}{7} - \pi$ 至四位有效數字，並確認它介乎 $\frac{1}{1260}$ 與 $\frac{1}{630}$ 之間。

2 **一個 Beta 積分。**利用公式 $\int_0^1 x^n(1-x)^n\,\mathrm{d}x = \frac{(n!)^2}{(2n+1)!}$（在 Section 4.4 中證明）驗證 $\int_0^1 x^4(1-x)^4\,\mathrm{d}x = \frac{1}{630}$。

3 **一個偏小的近似值。**證明

$$\int_0^1 \frac{x(1-x)^2}{1+x^2}\,\mathrm{d}x = \frac{\pi}{2} - \frac{3}{2} \quad \text{及} \quad \int_0^1 \frac{x^2(1-x)^4}{1+x^2}\,\mathrm{d}x = \pi - \frac{47}{15},$$

並由此推出 $\pi > 3$ 及 $\pi > \frac{47}{15}$。為甚麼這些近似值不及 $\frac{22}{7}$ 準確？

4 **更進一步。**較難驗證上文給出的 $\int_0^1 \frac{x^8(1-x)^8}{1+x^2}\,\mathrm{d}x$ 的恆等式。（除法可以交給電腦代數系統去做；重點是看出餘式是 16，從而得出 4π。）

延伸閱讀

- **Petr Beckmann, *A History of* π**（St Martin's Press）。詳細講述阿基米德的多邊形方法，另有許多其他內容。
- **Stephen Lucas, "Approximations to π derived from integrals with nonnegative integrands"**, *American Mathematical Monthly* 116 (2009)。篇幅短小，容易閱讀。
- **Ivan Niven, *Irrational Numbers***（Mathematical Association of America）。收錄了 Niven 對 π 是無理數的證明。

4.4 沃利斯、Beta 與二分之一的階乘

題目

STEP II 1995，第 4 題

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

`step.maths.org/questions/95-s2-q4`

引子

$\left(\frac{1}{2}\right)!$ 是甚麼？這個問題聽來毫無道理：$n!$ 的意思是 $1 \times 2 \times \cdots \times n$，而你不可能把「半個數那麼多」的項相乘。然而它確有一個自然的答案，就是

$$\left(\tfrac{1}{2}\right)! = \frac{\sqrt{\pi}}{2}.$$

π 從何而來？來自 $\sin t$ 的冪的積分（integral），這正是本題研究的對象。過程中，本題證明了關於 π 的其中一條最古老、最奇特的公式，即沃利斯（John Wallis）在 1655 年提出的無窮乘積（infinite product）

$$\frac{\pi}{2} = \frac{2}{1}\cdot\frac{2}{3}\cdot\frac{4}{3}\cdot\frac{4}{5}\cdot\frac{6}{5}\cdot\frac{6}{7}\cdots,$$

並且找出 $n!$ 的斯特靈公式（Stirling's formula）中隱藏的常數，這條公式是每一位概率學家和統計學家都會用到的。

動筆之前

想一想 u_0 和 u_1 是甚麼？當 n 很大時，$\sin^n t$ 在 $[0, \frac{\pi}{2}]$ 上的圖像大致是甚麼樣子？

$u_0 = \int_0^{\pi/2} 1\,\mathrm{d}t = \frac{\pi}{2}$，而 $u_1 = \int_0^{\pi/2} \sin t\,\mathrm{d}t = 1$。當 n 很大時，$\sin^n t$ 極小，只有在 $t = \frac{\pi}{2}$ 附近例外，因為那裏 $\sin t$ 接近 1：圖像變成緊貼右邊界的一個狹窄尖峯（圖 4.4.1）。所以 u_n 隨 n 增大而縮小。本題準確地找出它縮小得有多快：與 $\sqrt{\pi/(2n)}$ 相若。

逐步解題

階段 1　一條遞推公式

想一想　用分部積分法（integration by parts）計算 $\int \sin t \cdot \sin^{n-1} t \, \mathrm{d}t$：對 $\sin t$ 積分，對 $\sin^{n-1} t$ 求導。新的被積函數（integrand）含有 $\cos^2 t$；怎樣才能變回 $\sin t$ 的冪？

取 $\mathrm{d}v = \sin t \, \mathrm{d}t$（故 $v = -\cos t$）及 $w = \sin^{n-1} t$（故 $w' = (n-1)\sin^{n-2} t \cos t$），

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

當 $n \geq 2$ 時，邊界項為零。寫成 $\cos^2 t = 1 - \sin^2 t$，該積分便是 $u_{n-2} - u_n$，所以 $u_n = (n-1)(u_{n-2} - u_n)$，即

$$n u_n = (n-1) u_{n-2}. \qquad \blacksquare$$

兩邊同乘以 u_{n-1}，得 $n u_n u_{n-1} = (n-1) u_{n-1} u_{n-2}$。 $\blacksquare$

階段 2　一個不變量

想一想　這個關係式說明，當 n 減一時，$n u_n u_{n-1}$ 這個量保持不變。它在 $n = 1$ 時等於甚麼？

反覆運用，$n u_n u_{n-1} = (n-1) u_{n-1} u_{n-2} = \cdots = 1 \cdot u_1 u_0 = 1 \cdot 1 \cdot \frac{\pi}{2}$。所以對所有 $n \geq 1$，$n u_n u_{n-1} = \frac{1}{2}\pi$。 $\blacksquare$

階段 3　夾逼

對於 $0 < t < \frac{\pi}{2}$，我們有 $0 < \sin t < 1$，所以 $\sin^n t < \sin^{n-1} t$：除端點外，$\sin^n t$ 的圖像位於 $\sin^{n-1} t$ 的圖像之下。積分後得 $0 < u_n < u_{n-1}$。

停一停，試一試　利用 $u_n < u_{n-1}$ 和 $n u_n u_{n-1} = \frac{1}{2}\pi$ 把 $n u_n^2$ 夾住。

把 u_{n-1} 換成較小的 u_n，得 $n u_n^2 < n u_n u_{n-1} = \frac{1}{2}\pi$。把 u_n 換成較大的 u_{n-1}，得 $\frac{1}{2}\pi = n u_n u_{n-1} < n u_{n-1}^2$。所以 $n u_n^2 < \frac{1}{2}\pi < n u_{n-1}^2$。 $\blacksquare$

在右邊的不等式中以 $n+1$ 代替 n：$\frac{1}{2}\pi < (n+1) u_n^2$，所以 $n u_n^2 > \frac{n}{n+1} \cdot \frac{1}{2}\pi$。合起來，

$$\left(\frac{n}{n+1}\right)\frac{1}{2}\pi < n u_n^2 < \frac{1}{2}\pi. \qquad \blacksquare$$

由於 $\frac{n}{n+1} \to 1$，由夾逼可得 $n u_n^2 \to \frac{1}{2}\pi$。 $\blacksquare$

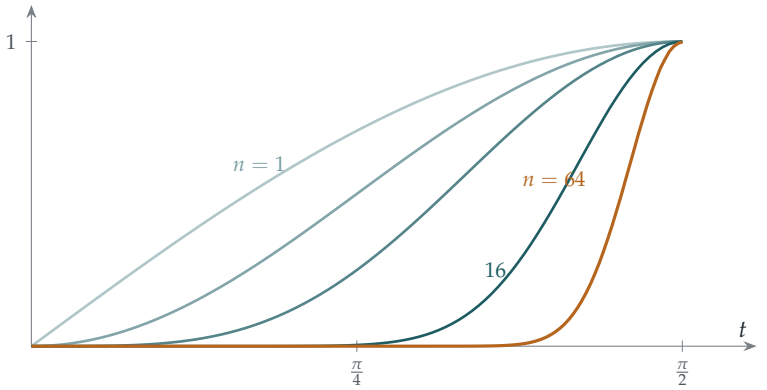


圖 4.4.1　$n = 1, 2, 4, 16, 64$ 時 $\sin^n t$ 的圖像。隨着 n 增大，曲線下的面積 u_n 不斷縮小，並集中在 $t = \frac{\pi}{2}$ 附近。本題證明 $u_n \approx \sqrt{\pi/(2n)}$；當 $n = 64$ 時，這個值約為 0.157。

回顧

- **遞推公式（reduction formulae）。**分部積分法把 u_n 化為 u_{n-2}，所以所有 u_n 都由 u_0 和 u_1 決定。
- **尋找不變量（invariant）。**乘積 nu_nu_{n-1} 保持不變，正如 Section 2.6 中的積分在 AGM 步驟下保持不變一樣。
- **單調性（monotonicity）加上不變量，便得出夾逼（squeeze）。**知道 $u_n < u_{n-1}$ 以及 nu_nu_{n-1} 的準確值，便把 nu_n^2 夾在兩個極限相同的量之間。

背後的數學

沃利斯乘積

把遞推公式一直用到 u_0 或 u_1，便得到明確的值：

$$u_{2m} = \frac{1\cdot 3\cdot 5\cdots(2m-1)}{2\cdot 4\cdot 6\cdots 2m}\cdot\frac{\pi}{2}, \qquad u_{2m+1} = \frac{2\cdot 4\cdot 6\cdots 2m}{3\cdot 5\cdot 7\cdots(2m+1)}.$$

由夾逼可知 $u_{2m}/u_{2m+1} \to 1$。把這個比寫出來再整理，便得到沃利斯乘積

$$\frac{\pi}{2} = \lim_{m\to\infty}\frac{2\cdot 2}{1\cdot 3}\cdot\frac{4\cdot 4}{3\cdot 5}\cdots\frac{2m\cdot 2m}{(2m-1)(2m+1)}.$$

沃利斯在 1655 年以一種大膽的插值（interpolation）方法發現了它：他把 $\int_0^1(1-x^2)^{1/2}\,dx$ 放進一個指數為整數的積分表中，從而猜出它「應該」是甚麼。這是數學史

上最早的無窮乘積之一，僅次於韋達（Viète）的乘積（Section 2.5），也是第一個只用有理數的無窮乘積。

二分之一的階乘

沃利斯的插值想法，在 1729 年由歐拉（Euler）賦予了嚴格的意義。他找到一個函數，現稱為*伽瑪函數*（Gamma function），

$$\Gamma(s) = \int_0^\infty t^{s-1} e^{-t}\,\mathrm{d}t,$$

它滿足 $\Gamma(s+1) = s\,\Gamma(s)$（用分部積分法），而且對整數 n 有 $\Gamma(n+1) = n!$。它把階乘（factorial）平滑地延拓至所有正實數。與它密切相關的*貝塔函數*（Beta function）

$$\mathrm{B}(p,q) = \int_0^1 x^{p-1}(1-x)^{q-1}\,\mathrm{d}x = \frac{\Gamma(p)\Gamma(q)}{\Gamma(p+q)}$$

包含了 Section 4.3 和 Section 2.9 中的積分，而且經代換 $x = \sin^2 t$ 之後，也包含我們的 u_n。特別地，$\int_0^1 x^{1/2}(1-x)^{1/2}\,\mathrm{d}x = \frac{\pi}{8}$（練習 2），由此得 $\Gamma(\frac{3}{2})^2/\Gamma(3) = \frac{\pi}{8}$，即 $\left(\frac{1}{2}\right)! = \Gamma(\frac{3}{2}) = \frac{\sqrt{\pi}}{2}$。等價地，$\Gamma(\frac{1}{2}) = \sqrt{\pi}$，這與著名的高斯積分（Gaussian integral）$\int_{-\infty}^{\infty} e^{-x^2}\,\mathrm{d}x = \sqrt{\pi}$ 是同一個事實，也正是 π 出現在正態分佈（normal distribution）公式中的原因。

斯特靈公式

斯特靈公式說 $n! \approx \sqrt{2\pi n}\,(n/e)^n$。其中 $(n/e)^n$ 的部分相對容易得到，但常數 $\sqrt{2\pi}$ 則較難，而求出它的標準方法正是本題：u_{2m} 的明確值涉及 $(2m)!/(m!)^2$，而極限 $nu_n^2 \to \frac{\pi}{2}$ 則把這個常數確定下來。另有一條截然不同的途徑，通過概率得到同一個常數，見 Section 12.23。

思維技巧

由積分得出遞推關係

分部積分法往往把帶有參數 n 的積分，與參數較小的同一類積分聯繫起來。這個遞推關係（recurrence）隨後可以解出，或用來分析 $n \to \infty$ 時的表現。

留意遞推關係下保持不變的量，以及對 n 的單調性；兩者結合，往往不需要任何明確公式便能求出極限。

延伸練習

1 **沃利斯乘積。**利用 u_{2m} 和 u_{2m+1} 的明確值，以及 $u_{2m+1} < u_{2m} < u_{2m-1}$，證明沃利斯乘積。

2 **一個貝塔積分。**試做 STEP II 2013，第 2 題（`step.maths.org/questions/13-s2-q2`）。這題用遞推公式計算 $\int_0^1 x^n(1-x)^n\,\mathrm{d}x$，然後問當 $n=\frac{1}{2}$ 時答案有何啟示：這是對二分之一的階乘的初次一瞥。

3 **貝塔函數。**試做 STEP III 2025，第 1 題（載於 OCR 提供下載的 2025 年試卷，`www.ocr.org.uk/Images/750486-2025-past-papers.zip`）。這題建立一般的貝塔積分 $\int_0^1 x^{p-1}(1-x)^{q-1}\,\mathrm{d}x$、它的對稱性和一個遞推關係，並用它們計算一個在 $[0,\infty)$ 上的積分。

4 **中心二項式係數（central binomial coefficients）。**較難證明 $u_{2m}=\binom{2m}{m}\dfrac{\pi}{2^{2m+1}}$，並利用 $nu_n^2\to\frac{\pi}{2}$ 推導出當 m 很大時 $\binom{2m}{m}\approx\dfrac{4^m}{\sqrt{\pi m}}$。就 $m=10$ 驗證這個結果。

延伸閱讀

- **Julian Havil, *Gamma***（Princeton University Press）。內容包括伽瑪函數及其歷史。
- **Philip J. Davis, “Leonhard Euler’s integral: a historical profile of the Gamma function”,** *American Mathematical Monthly* 66 (1959)。一篇獲獎的經典文章，讀過本節之後便能看懂。
- **Emil Artin, *The Gamma Function***（Dover）。一本簡短而優美的專著，適合學過一些分析之後閱讀。

4.5 大二學生之夢

題目

STEP III 2009，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/09-s3-q8`

引子

每個學生遲早都會學到 $(a+b)^2 \neq a^2+b^2$。那條錯誤的等式有個戲稱，叫做「大一新生之夢」（freshman's dream）。本題要證明的結果，看起來同樣像初學者會犯的錯：

$$\int_0^1 x^x \,\mathrm{d}x = \sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n^n} = 1 - \frac{1}{2^2} + \frac{1}{3^3} - \frac{1}{4^4} + \cdots,$$

以及與它配對的 $\int_0^1 x^{-x}\,\mathrm{d}x = \sum_{n=1}^{\infty} \frac{1}{n^n}$。$x^x$ 的積分（integral）竟然變成 n^{-n} 的和：這似乎好得令人難以置信，所以它被稱為**大二學生之夢**（sophomore's dream）。但它確實成立，由約翰·伯努利（Johann Bernoulli）於 1697 年發現。

證明用到三個你在大學會不斷遇到的想法：比較指數函數與冪的極限（limit）、歸約公式（reduction formula），以及把函數展開成冪級數（power series）再逐項積分（integrating term by term）。

動筆之前

? **想一想** 函數 x^x 在 $x=0$ 處無法用你所知的任何公式定義（0^0 的意義並不明確）。當 x 從上方趨近 0 時，它的表現如何？試用計算機計算 $x=0.1$、0.01、0.001 的情況。

$0.1^{0.1} \approx 0.794$，$0.01^{0.01} \approx 0.955$，$0.001^{0.001} \approx 0.993$。看來它趨近 1。第 (i) 部分正是證明這一點，由此可知被積函數（integrand）在 0 附近表現良好，這個積分是有意義的。x^x 在 $(0,1]$ 上的圖像由接近 1 的地方出發，下降至極小值，再在 $x=1$ 處回到 1（圖 4.5.1）。

逐步解題

階段 1　冪勝過對數

想一想　當 $x \to 0^+$ 時，$\ln x \to -\infty$。哪一個代換（substitution）可以把 $x \to 0^+$ 變成 $t \to \infty$，並把 $x^m (\ln x)^n$ 變成類似 $e^{-mt} t^n$ 的式子？

設 $x = e^{-t}$。當 $x \to 0^+$ 時，$t \to \infty$，而 $\ln x = -t$，所以

$$x^m (\ln x)^n = e^{-mt} (-t)^n = (-1)^n e^{-mt} t^n \to 0.$$

用文字來說：當 $x \to 0$ 時，x 的任何正數次冪趨向零的速度，都快過 $\ln x$ 的任何次冪趨向無窮的速度。 ■

取 $m = n = 1$，得 $x \ln x \to 0$；由於指數函數是連續（continuous）的，所以 $x^x = e^{x \ln x} \to e^0 = 1$。 ■

階段 2　歸約公式

想一想　用分部積分法（integration by parts）計算 $I_{n+1} = \int_0^1 x^m (\ln x)^{n+1} \, dx$。應該對哪一個因子積分？邊界項（boundary term）在 $x = 0$ 處會怎樣？

對 x^m 積分，對 $(\ln x)^{n+1}$ 求導：

$$I_{n+1} = \left[\frac{x^{m+1}}{m+1} (\ln x)^{n+1} \right]_0^1 - \int_0^1 \frac{x^{m+1}}{m+1} \cdot \frac{(n+1)(\ln x)^n}{x} \, dx.$$

在 $x = 1$ 處，因為 $\ln 1 = 0$，邊界項等於 0；當 $x \to 0$ 時，根據第 (i) 部分（取 $m+1$ 和 $n+1$），它趨向 0。餘下的積分是 $\frac{n+1}{m+1} I_n$，所以

$$I_{n+1} = -\frac{n+1}{m+1} I_n. \qquad \blacksquare$$

由 $I_0 = \int_0^1 x^m \, dx = \frac{1}{m+1}$ 出發，應用這條公式 n 次，得

$$I_n = (-1)^n \frac{n!}{(m+1)^n} \cdot \frac{1}{m+1} = \frac{(-1)^n \, n!}{(m+1)^{n+1}}. \qquad \blacksquare$$

階段 3　級數

想一想　利用 e^y 的級數展開 $x^x = e^{x \ln x}$。每一項是甚麼樣子？m 取甚麼值，才會出現第 (ii) 部分的積分？

由於對任何 y 都有 $e^y = \sum_{n=0}^{\infty} \frac{y^n}{n!}$，

$$x^x = e^{x \ln x} = \sum_{n=0}^{\infty} \frac{x^n (\ln x)^n}{n!}.$$

停一停，試一試 由 0 至 1 逐項積分，並使用第 (ii) 部分中 $m = n$ 的情況（$n = 0$ 另行處理）。你得到甚麼？

$n = 0$ 的一項是 1，積分後得 1。當 $n \geq 1$ 時，在第 (ii) 部分取 $m = n$，得 $\int_0^1 x^n (\ln x)^n \, dx = \frac{(-1)^n \, n!}{(n+1)^{n+1}}$，所以第 n 項的貢獻是 $\frac{(-1)^n}{(n+1)^{n+1}}$。把各項相加，

$$\int_0^1 x^x \, dx = 1 - \frac{1}{2^2} + \frac{1}{3^3} - \frac{1}{4^4} + \cdots . \qquad \blacksquare$$

（$n = 0$ 的一項也符合同一規律，即 $\frac{1}{1^1}$。）各項縮小得極快：取七項，便得到準確至小數點後六位的數值 0.783430 . . .。

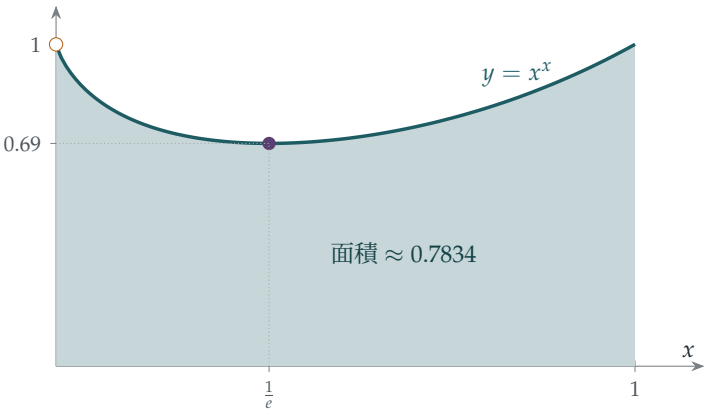


圖 4.5.1 x^x 在 $(0, 1]$ 上的圖像。當 $x \to 0$ 時它趨向 1（空心圓點），在 $x = \frac{1}{e}$ 處下降至極小值 $e^{-1/e} \approx 0.692$，再在 $x = 1$ 處回到 1。陰影部分的面積是 $\sum (-1)^{n+1} n^{-n} \approx 0.7834$。

回顧

- **用代換比較增長速度。**設 $x = e^{-t}$，便把在 0 處的極限變成在 ∞ 處一個熟悉的極限。
- **邊界項必須檢查。**歸約公式有賴邊界項在 $x = 0$ 處消失，這正是用到第 (i) 部分的地方。
- **先展開成級數，再積分。**把被積函數展開成冪級數並逐項積分，便把一個沒有初等原函數（elementary antiderivative）的積分，變成一個收斂得很快的和。

背後的數學

交換求和與積分

第 3 階段不動聲色地把無窮和與積分交換了次序：$\int \sum = \sum \int$。這樣做並非總是容許的。在大學裏，你會遇到交換後不成立的例子，也會學到一些定理，說明甚麼時候可以放心交換。最簡單的一個定理指出，當級數一致收斂(converges uniformly)時，交換是安全的：即級數的尾部可以同時在每一個 x 處都變得很小。在這裏，$|x \ln x| \le \frac{1}{e}$ 在 $(0,1]$ 上成立，所以無論 x 是甚麼，第 n 項的大小最多是 $\frac{e^{-n}}{n!}$。這些上界之和收斂，因此尾部一致地小，交換也就有了根據。這個論證稱為魏爾斯特拉斯 M 判別法(Weierstrass M-test)，是分析學（analysis）中最常用的工具之一。

為何這個「夢」會成真

這條恆等式很特別。它有賴 $\frac{(x\ln x)^n}{n!}$ 中的冪與積分 $\int_0^1 x^n(\ln x)^n\,\mathrm{d}x = \frac{(-1)^n n!}{(n+1)^{n+1}}$ 完全吻合，當中的 $n!$ 恰好全部約去。追根究柢，這個積分其實是喬裝了的伽瑪函數（Gamma function）：代換 $x = e^{-t/(n+1)}$ 把它變成 $\frac{(-1)^n}{(n+1)^{n+1}}\int_0^\infty t^n e^{-t}\,\mathrm{d}t = \frac{(-1)^n\,\Gamma(n+1)}{(n+1)^{n+1}}$，與 Section 4.4 連繫起來。

名字的由來

「大二學生之夢」這個名字，是 Jonathan Borwein 與合著者在 2004 年出版的 *Experimentation in Mathematics* 一書中所創，用來與「大一新生之夢」$(a+b)^n = a^n + b^n$ 配對（有趣的是，後者在某些數系中是成立的：在模質數 p 的算術中，$(a+b)^p = a^p + b^p$，因為中間所有的二項式係數（binomial coefficient）都可被 p 整除）。Borwein 是實驗數學（experimental mathematics）的先驅，這門學問利用電腦去發現像本題這樣的恆等式，然後再以人手證明。

思維技巧

展開後逐項積分

如果被積函數有冪級數展開，而你又能對其中各項積分，便對每一項積分，再把結果加起來。

對於沒有初等原函數的積分，這是一個很有效的求值方法。緊記要問：交換求和與積分是否有根據？通常是有的，但你應該知道原因。

延伸練習

1 **配對的等式。**證明 $\displaystyle\int_0^1 x^{-x}\,\mathrm{d}x = \sum_{n=1}^{\infty} \frac{1}{n^n} \approx 1.2913$。

2 **極小值。**證明 x^x 在 $(0,\infty)$ 上的極小值出現於 $x = \frac{1}{e}$，其值為 $e^{-1/e}$。

3 **需要多少項？**對於各項遞減的交錯級數（alternating series），截斷後的誤差小於第一個被略去的項。要準確至小數點後六位，需要取 $\int_0^1 x^x\,\mathrm{d}x$ 的級數的多少項？

4 **一個近親。**用同樣的方法證明 $\displaystyle\int_0^1 x^2 \ln x\,\mathrm{d}x = -\frac{1}{9}$，以及 $\displaystyle\int_0^1 \frac{\ln x}{1-x}\,\mathrm{d}x = -\sum_{n=1}^{\infty} \frac{1}{n^2}$。（根據 Section 2.4，這等於 $-\frac{\pi^2}{6}$。）

延伸閱讀

- **Jonathan Borwein, David Bailey & Roland Girgensohn, *Experimentation in Mathematics***（A. K. Peters）。這個名字的出處，書中還有許多由電腦發現的恆等式。
- **William Dunham, *The Calculus Gallery***（Princeton University Press）。由牛頓（Newton）到勒貝格（Lebesgue）的分析學傑作，包括伯努利的工作。
- **Mary Hart, *Guide to Analysis***（Palgrave）。清楚講解一致收斂，以及何時可以交換求和與積分。

4.6 沒有公式的積分

題目

STEP III 2016，第 3 題

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

`step.maths.org/questions/16-s3-q3`

引子

求導（differentiation）是機械式的工作：有一套法則可循，任何由常見部件砌成的函數，只要套用這些法則就能求導。積分（integration）卻不然。每個學生都遇過各種技巧都攻不破的積分，並懷疑是否只是自己不夠聰明。有時答案是：誰也不夠聰明，因為根本沒有公式。

本題就在一個特殊情況下證明這樣的事。積分 $\int \frac{e^x}{x+1}\,\mathrm{d}x$ 作為面積當然存在，但它不能寫成（有理函數）$\times e^x$ 的形式，而這正是用分部積分法（integration by parts）時預期會得到的形式。證明某件事不可能做到，而不是把某樣東西找出來，是另一種數學，也極具數學的特色。

動筆之前

想一想 若 $\int \mathrm{f}(x)e^x\,\mathrm{d}x = \frac{\mathrm{P}}{\mathrm{Q}}e^x + C$，P 和 Q 必須滿足甚麼方程？（對右邊求導。）

求導得 $\dfrac{\mathrm{d}}{\mathrm{d}x}\left(\dfrac{\mathrm{P}}{\mathrm{Q}}e^x\right) = \dfrac{\mathrm{P}'\mathrm{Q} - \mathrm{P}\mathrm{Q}' + \mathrm{P}\mathrm{Q}}{\mathrm{Q}^2}e^x$，而這必須等於 $\mathrm{f}(x)e^x$。所以整道題目其實圍繞着以下的多項式恆等式（polynomial identity）：

$$\mathrm{P}'\mathrm{Q} - \mathrm{P}\mathrm{Q}' + \mathrm{P}\mathrm{Q} = \mathrm{f}\,\mathrm{Q}^2.$$

關於積分的問題，就此變成關於多項式的問題：它們的因式和次數。

逐步解題

階段 1　因式 $x+1$

取 $\mathrm{f}(x)=\dfrac{x^3-2}{(x+1)^2}$，把恆等式兩邊同乘 $(x+1)^2$：

$$(x+1)^2(\mathrm{P}'\mathrm{Q}-\mathrm{P}\mathrm{Q}'+\mathrm{P}\mathrm{Q})=(x^3-2)\,\mathrm{Q}^2. \tag{†}$$

> **想一想**　在 (†) 中代入 $x=-1$。必然會出現甚麼情況？

當 $x=-1$ 時，左邊等於 0，右邊則是 $(-1-2)\,\mathrm{Q}(-1)^2=-3\,\mathrm{Q}(-1)^2$。所以 $\mathrm{Q}(-1)=0$，而根據因式定理（factor theorem），$x+1$ 是 Q 的因式。 ■

階段 2　次數

設 P 的次數（degree）為 p，Q 的次數為 q。

> **想一想**　在 $\mathrm{P}'\mathrm{Q}-\mathrm{P}\mathrm{Q}'+\mathrm{P}\mathrm{Q}$ 中，哪一項的次數最高？它會被其餘各項抵消嗎？

PQ 的次數是 $p+q$，而 $\mathrm{P}'\mathrm{Q}$ 和 $\mathrm{P}\mathrm{Q}'$ 的次數最多是 $p+q-1$。所以 (†) 左邊的次數恰好是 $2+p+q$，右邊的次數是 $3+2q$。令兩者相等，得 $p=q+1$。 ■

階段 3　求 P

取 $\mathrm{Q}=x+1$，則 $\mathrm{Q}'=1$，而 P 的次數為 2。把 (†) 除以 $(x+1)^2$：

$$\mathrm{P}'(x+1)-\mathrm{P}+\mathrm{P}(x+1)=x^3-2, \qquad \text{即} \qquad (x+1)\mathrm{P}'+x\mathrm{P}=x^3-2.$$

試設 $\mathrm{P}=\alpha x^2+\beta x+\gamma$。左邊是

$$(x+1)(2\alpha x+\beta)+\alpha x^3+\beta x^2+\gamma x=\alpha x^3+(2\alpha+\beta)x^2+(2\alpha+\beta+\gamma)x+\beta.$$

與 x^3-2 比較係數：$\alpha=1$；$2\alpha+\beta=0$，故 $\beta=-2$；於是 $\gamma=0$；常數項 $\beta=-2$ 亦吻合。所以

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

（有興趣的話可以求導驗算，這是練習商法則（quotient rule）的好題目。）

階段 4　一件不可能的事

現在假設 $\int \frac{e^x}{x+1}\,\mathrm{d}x=\frac{\mathrm{P}}{\mathrm{Q}}e^x+C$，其中 P 和 Q 沒有公因式（common factor）。恆等式變成

$$(x+1)\big(\mathrm{P}'\mathrm{Q}-\mathrm{P}\mathrm{Q}'+\mathrm{P}\mathrm{Q}\big)=\mathrm{Q}^2.$$

一如第 1 階段，代入 $x = -1$ 可知 $Q(-1) = 0$。寫成 $Q = (x+1)^k R$，其中 $k \geq 1$ 且 $R(-1) \neq 0$（即盡量抽出 $x+1$ 這個因式）。

> **停一停，試一試** 代入 $Q = (x+1)^k R$，盡量約去 $(x+1)$ 的因式，然後代入 $x = -1$。你能得知 $P(-1)$ 的甚麼性質？

這時 $Q' = k(x+1)^{k-1}R + (x+1)^k R'$，而

$$P'Q - PQ' + PQ = (x+1)^{k-1}\Big[(x+1)\big(P'R - PR' + PR\big) - kPR\Big].$$

所以恆等式變成 $(x+1)^k\big[(x+1)(P'R - PR' + PR) - kPR\big] = (x+1)^{2k}R^2$。除以 $(x+1)^k$：

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

當 $x = -1$ 時，由於 $k \geq 1$，上式給出 $-kP(-1)R(-1) = 0$。因為 $k \neq 0$ 且 $R(-1) \neq 0$，必有 $P(-1) = 0$。但這樣 $x+1$ 就同時整除 P 和 Q，與兩者沒有公因式的假設矛盾。所以這樣的 P 和 Q 並不存在。 ■

回顧

- **轉化為代數。**「這個積分能否寫成這種形式？」變成了「這個多項式恆等式有沒有解？」
- **在特殊點取值。**代入 $x = -1$ 就得到關於因式的資訊，前後共用了三次。
- **利用最簡分數的反證法（proof by contradiction）。**正如 Section 1.3 中處理 $\sqrt{2}$ 的做法，先假設分數已約至最簡，再找出一個公因式，便得到矛盾。

背後的數學

極點與對數

第 (ii) 部分背後有一幅簡單的圖像。被積函數（integrand）$\frac{e^x}{x+1}$ 在 $x = -1$ 處有一個單極點（simple pole）：在該點附近，它的表現像 $\frac{e^{-1}}{x+1}$。對 $\frac{1}{x+1}$ 積分得 $\ln|x+1|$，所以原函數（antiderivative）在 -1 附近的表現應該像 $e^{-1}\ln|x+1|$，以對數的速度趨向 $-\infty$（圖 4.6.1）。但有理函數乘以 e^x 只能以冪 $(x+1)^{-k}$ 的方式趨向無窮，其導數因而有一個階數為 $k+1 \geq 2$ 的極點，絕不會是單極點。有理函數無法產生對數式的表現；第 4 階段的代數運算，正是這個觀察的嚴格版本。相反，在第 (i) 部分，被積函數有一個二階極點，而這正是帶有單極點的函數求導後的模樣，所以答案才有可能是有理的。

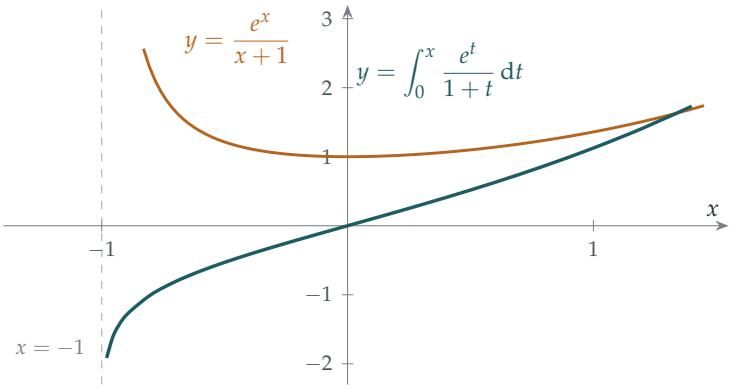


圖 4.6.1　被積函數 $\frac{e^x}{x+1}$（琥珀色）在 $x = -1$ 處有一個單極點。它的原函數（藍綠色，以數值方法計算）完全存在，但當 $x \to -1^+$ 時，它以對數的速度趨向 $-\infty$，表現像 $e^{-1}\ln(x+1)$。沒有任何有理函數乘以 e^x 能有這樣的表現。

劉維爾與 Risch 算法

在 1830 年代，劉維爾（Joseph Liouville）建立了一套理論，研究哪些積分可以用*初等函數*（elementary functions）表示（即由多項式、指數函數、對數函數、根式和三角函數經通常的運算構成的函數）。他的其中一個結果指出：若 f 和 g 是有理函數而 g 不是常數，則 $\int \mathrm{f}(x)e^{\mathrm{g}(x)}\,dx$ 只有在等於 $\mathrm{R}(x)e^{\mathrm{g}(x)}$（另加一個常數）時才是初等的，其中 R 是某個有理函數。（對 g 的條件不可或缺：當 $\mathrm{g} = 0$ 而 $\mathrm{f}(x) = \frac{1}{x}$ 時，積分是 $\ln|x|$，它是初等函數，卻不是有理函數。）所以本題排除了有理形式，其實就證明了 $\int \frac{e^x}{x+1}\,dx$ 根本不是初等函數。同一定理的其他版本可證明 $\int e^{-x^2}\,dx$ 和 $\int \frac{\sin x}{x}\,dx$ 都不是初等的。1969 年，Robert Risch 把劉維爾的想法發展成一套算法（algorithm），可就一大類函數判定初等積分是否存在，並在存在時把它求出來。電腦代數系統（computer algebra system）內部所用的，就是這套算法的不同版本。

特殊函數

當一個積分沒有初等形式時，數學家會給它起個名字，把它當作新的函數來研究，正如我們在 Section 3.9 中處理朗伯 W 函數（Lambert W function）那樣。這裏的積分本質上就是*指數積分*（exponential integral）$\mathrm{Ei}(x)$，它出現於熱傳導問題，也出現於*對數積分*（logarithmic integral）$\mathrm{li}(x) = \mathrm{Ei}(\ln x)$ 之中，後者是不超過 x 的質數（prime）個數的最佳簡單近似（Section 1.2）。當 $x < 0$ 時，它是 $\int_{-\infty}^{x} \frac{e^t}{t}\,dt$。當 $x > 0$ 時，這個積分會穿過 $t = 0$ 處的無窮尖峯，所以 $\mathrm{Ei}(x)$ 的定義是先挖去尖峯周圍的一個小區間 $(-\varepsilon, \varepsilon)$，再令 $\varepsilon \to 0$：這就是*柯西主值*（Cauchy principal value）。

思維技巧

證明不可能

要證明某件事做不到，先假設它做得到，把這個假設轉化爲精確的代數式，然後推出矛盾。

困難在於轉化：要為「做得到」的意思找出一個精確的陳述。在這裏，它的意思是「存在沒有公因式的多項式 P、Q 滿足這個恆等式」。數學中許多最深刻的結果，都是這類不可能性證明。

延伸練習

1 **兩個可行的例子。**證明 $\int \frac{xe^x}{(x+1)^2}\,\mathrm{d}x = \frac{e^x}{x+1} + C$ 及 $\int \frac{(x^2+1)e^x}{(x+1)^2}\,\mathrm{d}x = \frac{x-1}{x+1}e^x + C$。

2 **高斯函數。**證明不存在多項式 P 使得 $\int e^{-x^2}\,\mathrm{d}x = \mathrm{P}(x)e^{-x^2} + C$。（提示：求導後比較次數。）

3 **另一個極點。**仿照第 (ii) 部分的論證，證明 $\int \frac{e^x}{x}\,\mathrm{d}x$ 不能寫成 $\frac{\mathrm{P}(x)}{\mathrm{Q}(x)}e^x + C$ 的形式。

4 **何時可行？**假設 f 是多項式，且對某個多項式 P 有 $\int \frac{\mathrm{f}(x)}{(x+1)^2}e^x\,\mathrm{d}x = \frac{\mathrm{P}(x)}{x+1}e^x + C$。證明 $(x+1)\mathrm{P}' + x\mathrm{P} = \mathrm{f}$，並透過計算此式及其導數在 $x = -1$ 的值，證明 $\mathrm{f}(-1) + \mathrm{f}'(-1) = 0$。驗證這個條件對 $\mathrm{f}(x) = x^3 - 2$ 以及練習 1 的例子都成立。較難你能證明這個條件也是充分的嗎？

延伸閱讀

- **Brian Conrad, “Impossibility theorems for elementary integration”**（初等積分的不可能性定理）。可免費取得的講義，配合例子講解劉維爾定理。
- **Elena Anne Marchisotto & Gholam-Ali Zakeri, “An invitation to integration in finite terms”**（有限形式積分導引），*The College Mathematics Journal* 25 (1994)。劉維爾理論的淺白入門。
- **Manuel Bronstein, *Symbolic Integration I***（Springer）。講述電腦代數系統實際上如何求積分。程度較深。

4.7 變分法

題目

STEP III 2014，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/14-s3-q4`

引子

中學裏的最優化（optimisation）問題，大多是求最佳的數：例如求令面積最大的 x 值。*變分法*（calculus of variations）要求的卻是最佳的*函數*。連接兩點的曲線之中，珠子沿哪一條滑下最快？懸掛的鏈條會呈甚麼形狀？光線穿過透鏡時走哪一條路徑？在這些例子中，要優化的量都是一個積分，而未知的是一整個函數。

本題研究的是最簡單而又有趣的例子。積分 $J = \int_0^1 ((y')^2 - a^2 y^2)\,\mathrm{d}x$ 在兩個互相競爭的項之間取得平衡：函數愈陡，積分愈大；函數值愈大，積分愈小。當 a 較小時，陡峭程度總是佔上風，所以對每個函數都有 $J \geq 0$。但 a 有一個臨界值，平衡會在那裏逆轉，而本題正好把它準確求出：$a = \frac{\pi}{2}$。π 從何而來，就是本節要講的故事。

動筆之前

想一想 為甚麼 $I \geq 0$ 並非顯而易見？試用 $y = 1 - x$，它滿足 $y(1) = 0$。

I 是兩個正數量之差，所以它可正可負。對於 $y = 1 - x$：$\int_0^1 1\,\mathrm{d}x = 1$ 而 $\int_0^1 (1-x)^2\,\mathrm{d}x = \frac{1}{3}$，所以 $I = \frac{2}{3} > 0$。但若函數選得更巧妙，或許可以令 $\int y^2$ 相對於 $\int (y')^2$ 變得更大。題目的竅門，是把 I 寫成一個*平方*的積分，而這顯然是非負的。這就是積分版本的「配方法」（completing the square）。

逐步解題

階段 1　配方

把 I_1 中的平方展開：

$$I_1 = \int_0^1 ((y')^2 + 2yy' \tan x + y^2 \tan^2 x)\,\mathrm{d}x,$$

所以

$$I - I_1 = -\int_0^1 (y^2 + y^2 \tan^2 x + 2yy' \tan x)\,\mathrm{d}x = -\int_0^1 (y^2 \sec^2 x + 2yy' \tan x)\,\mathrm{d}x.$$

> **想一想**　被積函數 $y^2 \sec^2 x + 2yy' \tan x$ 恰好是某個函數的導數（exact derivative）。是哪個函數？

根據乘積法則（product rule），$\dfrac{\mathrm{d}}{\mathrm{d}x}(y^2 \tan x) = 2yy' \tan x + y^2 \sec^2 x$。所以

$$I - I_1 = -\Big[y^2 \tan x\Big]_0^1 = -(y(1)^2 \tan 1 - y(0)^2 \tan 0) = 0,$$

因為 $y(1) = 0$ 而 $\tan 0 = 0$。所以 $I = I_1 \geq 0$，因為它是一個平方的積分。 ■

階段 2　何時 $I = 0$？

> **想一想**　如果一個連續的非負函數的積分是零，你對這個函數能下甚麼結論？那麼若 $I = 0$，y 必須滿足甚麼方程？

積分為零的連續非負函數必定恆等於零（假如它在某一點取正值，它在該點附近的一個小區間上也會取正值，從而貢獻一個正的量）。所以 $I = 0$ 迫使 $y' + y \tan x = 0$ 在 $[0, 1]$ 上成立。乘以 $\sec x$ 後，這就是說 $\frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{y}{\cos x}\right) = \frac{y' \cos x + y \sin x}{\cos^2 x} = 0$，所以 $y = C \cos x$。於是 $y(1) = C \cos 1 = 0$，而由於 $\cos 1 \neq 0$，得 $C = 0$。所以只有當 $y \equiv 0$ 時才有 $I = 0$。 ■

階段 3　一般情況

> **想一想**　用 $\int_0^1 (y' + ay \tan bx)^2\,\mathrm{d}x$ 重複第 1 步。b 必須是甚麼，餘下的項才能恰好組成一個導數？

把這個積分記作 K。那麼

$$\begin{aligned} J - K &= -\int_0^1 (a^2y^2 + a^2y^2\tan^2 bx + 2ayy'\tan bx)\,\mathrm{d}x \\ &= -\int_0^1 (a^2y^2\sec^2 bx + 2ayy'\tan bx)\,\mathrm{d}x. \end{aligned}$$

另一方面，$\frac{\mathrm{d}}{\mathrm{d}x}(ay^2\tan bx) = 2ayy'\tan bx + ab\,y^2\sec^2 bx$。兩者恰好在 $ab = a^2$ 時吻合，即 $b = a$。於是

$$J - K = -\Big[ay^2\tan ax\Big]_0^1 = -a\,y(1)^2\tan a = 0,$$

所以 $J = K \geq 0$。

> **停一停，試一試** 這個論證在哪裏可能失效？（留意 $\tan ax$ 在區間 $0 \leq x \leq 1$ 上的情況。）

這個論證需要 $\tan ax$ 對 $[0,1]$ 內所有 x 都有定義而且連續；否則 K 未必存在，微積分基本定理（fundamental theorem of calculus）亦不能應用。當 x 取遍 $[0,1]$ 時，ax 取遍 $[0,a]$，而 tan 在該處連續的充要條件正是 $a < \frac{\pi}{2}$。所以這個證明在以下情況成立：

$$a < k = \frac{\pi}{2}. \qquad \blacksquare$$

階段 4 臨界情況

> **想一想** 當 $a = \frac{\pi}{2}$ 時，證明失效。在第 2 步，令 $I = 0$ 的函數須滿足 $y' + y\tan x = 0$。現在對應的方程是甚麼？它的解又是甚麼？

對應的方程是 $y' + ay\tan ax = 0$，它的解是 $y = C\cos ax$。當 $a = \frac{\pi}{2}$ 時，$y = \cos\frac{\pi x}{2}$ 滿足 $y(1) = \cos\frac{\pi}{2} = 0$。直接驗證：

$$\begin{aligned} J &= \frac{\pi^2}{4}\int_0^1 \left(\sin^2\frac{\pi x}{2} - \cos^2\frac{\pi x}{2}\right)\mathrm{d}x = -\frac{\pi^2}{4}\int_0^1 \cos\pi x\,\mathrm{d}x \\ &= -\frac{\pi^2}{4}\left[\frac{\sin\pi x}{\pi}\right]_0^1 = 0. \qquad \blacksquare \end{aligned}$$

所以在臨界值處，有一個非零函數令 J 恰好等於零。

回顧

- **在積分中配方。**加上再減去一個恰好是導數的式子，便把 J 化成一個平方的積分，再加上一個等於零的邊界項（boundary term）。

- **查看工具在哪裏失效。**這個方法只在 $\tan ax$ 在區間上保持有限時才行得通。它失效之處 $a = \frac{\pi}{2}$，正正就是臨界值。
- **極端情況是某個微分方程的解。**令 $J = 0$ 的函數來自一階微分方程（differential equation）$y' + ay\tan ax = 0$，而它同時滿足 $y'' + a^2 y = 0$。

背後的數學

最速降線與一門學科的誕生

1696 年，約翰·伯努利（Johann Bernoulli）向歐洲的數學家發出挑戰：求一條曲線，使珠子沿着它從一點滑到另一點所需的時間最短。這就是最速降線（brachistochrone）問題。提交解答的有牛頓（Newton）（他匿名投稿，一夜之間解出；伯努利說他「從爪子認出了獅子」）、萊布尼茲（Leibniz）、洛必達（l'Hôpital）和雅各布·伯努利（Jakob Bernoulli）。答案是一條擺線（cycloid）。到了 1750 年代，歐拉（Euler）和年輕的拉格朗日（Joseph-Louis Lagrange）把這些個別的技巧變成一套通用的方法：如果 y 令 $\int F(x, y, y')\,\mathrm{d}x$ 取平穩值，那麼 y 滿足*歐拉－拉格朗日方程*（Euler–Lagrange equation）

$$\frac{\mathrm{d}}{\mathrm{d}x}\frac{\partial F}{\partial y'} = \frac{\partial F}{\partial y}.$$

對於我們的 J，$F = (y')^2 - a^2y^2$，而歐拉－拉格朗日方程就是 $y'' + a^2y = 0$：簡諧運動（simple harmonic motion）。

為何是 $\pi/2$？

當 $a < \frac{\pi}{2}$ 時，J 的最小值是 0，而且只有 $y \equiv 0$ 才能取得。當 $a = \frac{\pi}{2}$ 時，最小值仍然是 0，但多了一個函數能取得它。當 $a > \frac{\pi}{2}$ 時，同一個函數 $y = \cos\frac{\pi x}{2}$ 給出

$$J = \frac{\pi^2}{8} - \frac{a^2}{2} < 0,$$

而只要把 y 按比例放大，J 就可以負得要多大有多大：根本沒有最小值（圖 4.7.1）。臨界值就是第一個令 $y'' + a^2y = 0$ 有非零解滿足邊界條件的 a（邊界條件是 $y(1) = 0$，以及在自由端的「自然」條件 $y'(0) = 0$）。這是一個*特徵值問題*（eigenvalue problem），而 $a^2 = \frac{\pi^2}{4}$ 是它最小的特徵值。試與 Section 3.7 比較：那裏把函數在長度為 π 的區間的兩端都固定，得出臨界值 1；這裏區間長度為 1，而且只固定一端，相當於長度為 2 的區間的一半，所以臨界的 a 是 $\frac{\pi}{2}$。

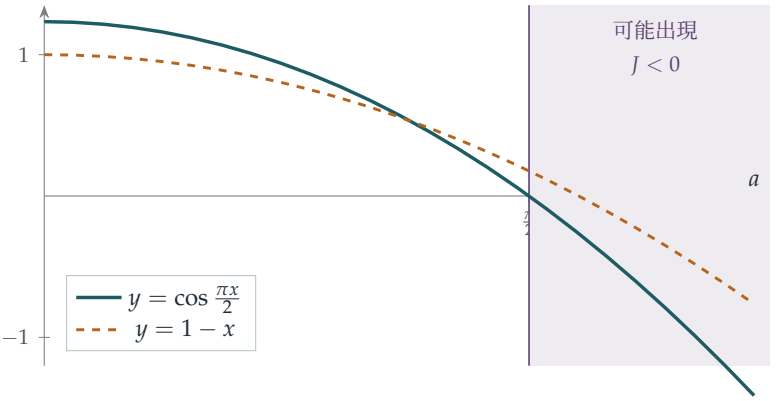


圖 4.7.1 對於兩個滿足 $y(1) = 0$ 的試驗函數，J 的值隨 a 的變化。對於 $y = 1 - x$，$J = 1 - \frac{a^2}{3}$ 要到 $a > \sqrt{3}$ 才變成負數。對於 $y = \cos\frac{\pi x}{2}$，$J = \frac{\pi^2}{8} - \frac{a^2}{2}$ 在 $a > \frac{\pi}{2}$ 時隨即變成負數，而且沒有函數能做得更好：當 $a < \frac{\pi}{2}$ 時，對每個 y 都有 $J \geq 0$。

物理學與最小作用量原理

後來發現，變分法描述的竟是大自然本身。在十八和十九世紀，人們逐漸明白運動定律可以這樣表述：力學系統的運動方式，會令一個稱為作用量（action）的積分取平穩值，這就是最小作用量原理（principle of least action）。同樣，光沿着需時最短的路徑前進（費馬原理，Fermat's principle），這解釋了折射現象。到了二十世紀，同一個想法通過費曼（Richard Feynman）的「路徑求和」（sum over paths），成為量子力學的基礎。本題中的二次積分是最簡單的例子，而它們引出的特徵值問題，會在 Section 8.6 以量子振子能級的面貌再次出現。

思維技巧

先配方，再留意邊界

要證明一個二次式非負，可把它寫成一個平方加上一些你能控制的東西。對於積分，這些「東西」往往恰好是某個函數的導數，積分後便成爲邊界項。

此外，永遠要問這個方法在哪裏不再奏效。本題中方法失效之處並非偶然；它正是題目中最重要的那個數。

延伸練習

1 **越過臨界值。**證明若 $a > \frac{\pi}{2}$，則 $y = \cos\frac{\pi x}{2}$ 給出 $J < 0$。

2 **歐拉－拉格朗日方程。**較難設 y 在滿足 $y(1) = 0$ 的函數之中令 J 取最小值。對任何滿足 $\eta(1) = 0$ 的函數 η 和任何數 ε，都有 $J(y + \varepsilon\eta) \geq J(y)$。把 $J(y + \varepsilon\eta)$ 按 ε 的冪展開，並由此推出 $\int_0^1 (y'\eta' - a^2 y\eta)\,\mathrm{d}x = 0$。用分部積分法（integration by parts）證明 $y'' + a^2 y = 0$ 及 $y'(0) = 0$（假設 y 足夠光滑）。

3 **較長的區間。**用代換法（substitution）證明：若 $y(L) = 0$，則 $\int_0^L \left((y')^2 - a^2 y^2\right)\mathrm{d}x \geq 0$ 對所有 y 成立，當且僅當 $a \leq \frac{\pi}{2L}$；並證明只有當 $a = \frac{\pi}{2L}$ 時，才有非零的 y 令等號成立。

4 **擺線。**查閱最速降線問題，並找出擺線的參數方程（parametric equations）。證明在一條倒置的擺線上，珠子無論從哪一點由靜止釋放，到達最低點所需的時間都相同（這是等時（tautochrone）性質，由惠更斯（Huygens）於 1659 年發現）。

延伸閱讀

- **Paul Nahin, *When Least Is Best***（Princeton University Press）。講述最優化的歷史與數學，包括最速降線，為學生而寫。
- **Richard Feynman, *The Feynman Lectures on Physics***，第二卷第 19 章（「The principle of least action」，最小作用量原理）。可在網上免費閱讀。
- **Bruce van Brunt, *The Calculus of Variations***（Springer）。一本大學本科教科書；頭幾章只需要高級程度（A-level）的微積分知識。

4.8 切球形麵包

題目

STEP I 2001，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/01-s1-q6`

引子

把一個球形麵包切成厚度相等的薄片。靠近中間的幾片是又大又厚的圓片，只有邊緣一圈窄窄的麵包皮；靠近兩端的幾片則是小小的球冠，看起來幾乎全是麵包皮。兩端的薄片，麵包皮一定比較多吧？

其實不然。不論切在靠近中間還是靠近兩端的位置，每一片的麵包皮面積都完全相同。中間那圈窄帶周長很大，但幾乎是垂直的；兩端的球冠周長很小，但傾斜得很厲害，所以麵包皮鋪開的闊度反而較大。這兩種效應恰好互相抵消。這是阿基米德（Archimedes）的一條定理，他對這條定理珍愛有加，甚至要求在自己的墓碑上刻上一個內切於圓柱（cylinder）的球（sphere）。題目的第二部分是一則較為輕鬆的「規例」，但它所依據的正是同一個計算。

動筆之前

想一想 半徑為 a 的球，是由哪一條曲線繞 x-軸旋轉而成的？該曲線的 $\frac{\mathrm{d}y}{\mathrm{d}x}$ 是甚麼？

把半圓（semicircle）$y = \sqrt{a^2 - x^2}$, $-a \leq x \leq a$ 旋轉，便得到這個球。於是 $\frac{\mathrm{d}y}{\mathrm{d}x} = \frac{-x}{\sqrt{a^2-x^2}} = -\frac{x}{y}$。請記住這一點：斜率（slope）在兩端附近（$x \to \pm a$）很陡，在中間則很平緩，這正是引子中所描述的效應。

逐步解題

階段 1　麵包皮公式的化簡

? **想一想**　對這個半圓計算 $y\sqrt{1+(y')^2}$。奇妙的事情出現了。

由於 $y' = -x/y$，

$$1 + (y')^2 = 1 + \frac{x^2}{y^2} = \frac{y^2 + x^2}{y^2} = \frac{a^2}{y^2},$$

所以 $y\sqrt{1+(y')^2} = y \cdot \frac{a}{y} = a$，竟然是一個常數！因此，任意兩個平面 $x = x_1$ 與 $x = x_2$ 之間的麵包皮面積為

$$2\pi \int_{x_1}^{x_2} a \, \mathrm{d}x = 2\pi a (x_2 - x_1).$$

它只取決於兩個平面之間的距離，而與平面的位置無關。

階段 2　麵包皮與片數成正比

如果把麵包切成厚度為 h 的薄片，每一片的麵包皮面積都是 $2\pi a h$，片片相同。不論是從一端吃、從兩端吃，還是從任何地方吃掉若干片，剩下的麵包皮面積都是 $2\pi a h$ 乘以剩下的片數。 ■

階段 3　「脆皮」的條件

題目的提示所描述的，是從一端吃掉若干片之後剩下的麵包，我們就依照提示去做：假設從 $x = a$ 那一端開始吃，於是剩下的部分由 $x = -a$ 延伸至 $x = -a + t$。一項關於「任意片數」的規定，應該不論麵包切得多薄都成立，所以我們把 t 看成可以取 0 與 $2a$ 之間的任何值。（第 4 階段會探討從兩端吃的情況。）

? **想一想**　利用題目給出的公式計算 V。使用代換（substitution）$s = x + a$ 或許能令代數運算更簡潔。

設 $s = x + a$，則 $y^2 = a^2 - x^2 = a^2 - (s-a)^2 = 2as - s^2$，所以

$$V = \pi \int_0^t (2as - s^2)\, \mathrm{d}s = \pi \left(at^2 - \frac{t^3}{3} \right), \qquad A = 2\pi a t,$$

以及

$$\frac{V}{A} = \frac{at^2 - t^3/3}{2at} = \frac{t}{2} - \frac{t^2}{6a}.$$

停一停，試一試 規定要求吃掉任意片數之後都有 $V/A < 1$，即是對所有滿足 $0 < t \le 2a$ 的 t 都成立。在這個範圍內，$\frac{t}{2} - \frac{t^2}{6a}$ 的最大值是多少？

對 t 求導：$\frac{1}{2} - \frac{t}{3a} = 0$ 於 $t = \frac{3a}{2}$ 時成立，這個值位於範圍 $(0, 2a]$ 之內，而且是極大值（maximum）（該式是 t 的一條開口向下的拋物線）。最大值為

$$\frac{3a}{4} - \frac{9a^2/4}{6a} = \frac{3a}{4} - \frac{3a}{8} = \frac{3a}{8}.$$

脆皮麵包需要對所有 t 都有 $V/A < 1$，特別是在 $t = \frac{3a}{2}$ 時也要成立，所以 $\frac{3a}{8} < 1$，即 $a < \frac{8}{3} = 2\frac{2}{3}$ 米。 ■

階段 4 題目之外：$2\frac{2}{3}$ 足夠嗎？

想一想 題目只要求找出半徑必須滿足的條件。$a < \frac{8}{3}$ 是否同時也足夠？如果從兩端吃又如何？

並不足夠。從兩端吃會在中間剩下一條環帶，而以同樣多的麵包皮計算，中間的環帶比末端的球冠包含更多麵包。對於環帶 $-\frac{w}{2} \le x \le \frac{w}{2}$，

$$V = \pi \int_{-w/2}^{w/2} (a^2 - x^2)\,\mathrm{d}x = \pi\left(a^2 w - \frac{w^3}{12}\right), \qquad A = 2\pi a w, \qquad \frac{V}{A} = \frac{a}{2} - \frac{w^2}{24a},$$

當環帶很窄時，這個值接近 $\frac{a}{2}$。例如，一個半徑為 2.4 米的麵包，明顯小於 $2\frac{2}{3}$，切成四片後吃掉兩端的兩片，剩下的環帶有 $w = a$ 及 $\frac{V}{A} = \frac{11a}{24} = 1.1$。事實上，對於任何一組薄片都有 $\frac{V}{A} < \frac{a}{2}$，所以如果可以從任何地方吃掉薄片，麵包是脆皮的當且僅當 $a \le 2$ 米。題目中的 $2\frac{2}{3}$ 是一個正確的必要條件（necessary condition），但這項規定其實比題目所透露的更為嚴格。

回顧

- **先化簡，後積分。**麵包皮的被積函數（integrand）看似複雜，結果卻是常數。動手積分之前，務必先把被積函數化簡。
- **仔細閱讀「任意」二字。**條件必須在吃掉任意片數之後都成立，所以我們需要的是 V/A 的最大值，而不是整個麵包的值（即 $\frac{a}{3}$）。
- **簡單的計算，出人意表的答案。**「麵包皮相等」這個性質聽起來難以置信，但計算只有兩行。

背後的數學

阿基米德的帽盒定理

阿基米德在《論球與圓柱》（*On the Sphere and Cylinder*）中證明了：球面上夾在兩個平行平面之間的環帶，其面積等於這兩個平面在包裹球的圓柱上所截出的環帶的面積（圖 4.8.1）。我們的計算正是這條帽盒定理（hat-box theorem）：半徑為 a 的圓柱每單位長度的麵包皮面積為 $2\pi a$，球也是如此。考慮整個球，它的表面積（surface area）等於圓柱的曲面面積，即 $2\pi a \times 2a = 4\pi a^2$。阿基米德認為，這個結果連同「球的體積是圓柱體積的三分之二」這一事實，是他畢生最出色的成就。西塞羅（Cicero）於公元前 75 年到訪敍拉古（Syracuse）時，正是憑着尋找刻在墓上的球與圓柱，找到了那座雜草叢生的墳墓。

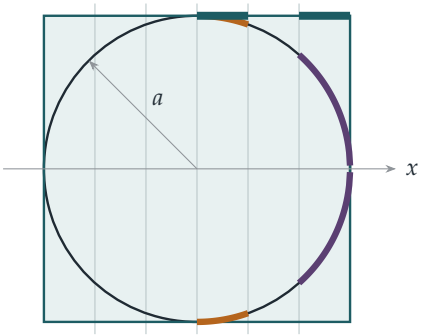


圖 4.8.1　半徑為 a 的球形麵包置於包裹它的圓柱之內的橫截面，麵包被切成厚度相等的薄片。中間的薄片（琥珀色）的麵包皮是一段短而接近水平的弧，離軸較遠；末端的薄片（紫紅色）的麵包皮是一段長而陡峭的弧，離軸較近。繞 x-軸旋轉後，兩者的麵包皮面積都恰好等於圓柱上對應環帶（藍綠色）的面積：$2\pi a$ 乘以厚度。

地球的地圖

帽盒定理提供了一種繪製保持面積不變的地球地圖的方法。把地球儀上的每一點沿水平方向向外投影到包裹它的圓柱上，然後把圓柱攤開。球面上的環帶會映射到圓柱上面積相同的環帶，而任何區域也是如此。這就是蘭伯特（*Lambert*）圓柱等積投影（Lambert's cylindrical equal-area projection，1772 年）。它在兩極附近會令形狀嚴重失真，但面積是如實的，不像人們熟悉的麥卡托（Mercator）投影那樣，令格陵蘭看起來與非洲一樣大（其實格陵蘭大約只有非洲的十四分之一）。高斯（Gauss）有一條著名的定理，稱為「絕妙定理」（*Theorema Egregium*，1827 年），它說明沒有任何球面地圖可以同時保持面

積和形狀：每一幅平面的地球地圖都必定有所失真。

球面上的隨機點

以下是一個實際的應用。要在球面上均勻（uniformly）隨機地選取一點，你或許會嘗試均勻地選取緯度（latitude）和經度（longitude），但這樣會令點在兩極附近過於密集。帽盒定理給出了正確的做法：在 $-a$ 與 a 之間均勻地選取高度 x，並在 0 與 2π 之間均勻地選取經度。由於面積與高度成正比，這樣選出的點恰好是均勻分佈的。物理學和電腦圖像的電腦模擬都採用這個方法。（值得一提的是，這個方法只在三維空間中有效：對於四維或更高維的球面，相應的命題並不成立。）

思維技巧

尋找隱藏的簡潔

當計算得出出乎意料地簡單的結果（例如一個常數）時，請停下來問一問爲甚麼。背後通常有幾何上的原因，而找出這個原因，才是眞正理解之所在。

在這裏，被積函數 $y\sqrt{1+(y')^2}$ 是常數，因為在球面上，陡峭程度與離軸的距離恰好互相補償：這就是阿基米德的定理。

延伸練習

1. **整個麵包。**證明對於整個未切開的麵包，$V/A = \frac{a}{3}$。為甚麼題目中的條件比 $\frac{a}{3} < 1$ 更難滿足？

2. **熱帶地區。**地球表面有多大比例位於南緯 23.4° 與北緯 23.4° 之間？（把地球視為一個球。由於面積與高度成正比，你只需要用到 $\sin 23.4°$。）

3. **其他形狀。**證明對於拋物面（paraboloid）形狀的麵包（把 $y = \sqrt{x}$ 繞 x-軸旋轉），每片的麵包皮並非常數。對於哪些曲線 $y(x)$，$y\sqrt{1+(y')^2}$ 是常數？（解出相應的微分方程（differential equation）。）

4. **體積。**利用體積公式，證明球的體積是包裹它的圓柱體積的三分之二，正如阿基米德所發現的那樣。

延伸閱讀

- **Sherman Stein, *Archimedes: What Did He Do Besides Cry Eureka?***（Mathematical Association of America）。為現代讀者講解阿基米德的數學。
- **Timothy Gowers, *Mathematics: A Very Short Introduction***（Oxford University Press）。書中關於幾何的一章介紹了曲面的幾何。
- **John Snyder, *Flattening the Earth***（University of Chicago Press）。兩千年來的地圖投影史。

4.9 Holditch 定理

題目

STEP III 2011，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/11-s3-q5`

引子

隨便取一個卵形，甚麼形狀都可以，只要它是凸（convex）的，然後讓一根長度固定的桿在裏面滑動一圈，兩端始終留在曲線上。在桿上塗一個點，把桿分成長度為 a 和 b 的兩段。桿繞行一周時，這個點會描出一個較小的卵形。兩條曲線之間的環形區域，面積是多少？

答案令人驚訝：*無論原來的曲線是甚麼形狀，面積都是* πab。又長又窄的卵形、蛋形、圓角三角形：環形的面積永遠相同，只取決於那個點在桿上的位置。這個結果由劍橋大學岡維爾與凱斯學院（Gonville and Caius College, Cambridge）的院士 Hamnet Holditch 牧師於 1858 年發現，是初等幾何中最出人意表的定理之一。本題只用一條面積公式就把它證明出來。

動筆之前

想一想 先在最簡單的情況下檢驗這個定理。若 $\mathcal{C}$ 是半徑為 R 的圓，P 描出的是甚麼曲線？環形的面積是多少？

由對稱性可知，P 描出一個同心的圓，設其半徑為 r。桿 AB 是大圓上一條通過 P 的弦，而*相交弦定理*（intersecting chords theorem，又稱「圓冪」（power of a point））指出，對任何通過 P 的弦，都有 $PA \times PB = R^2 - r^2$。所以 $ab = R^2 - r^2$，環形的面積是 $\pi R^2 - \pi r^2 = \pi ab$。定理說的是：這個結論對每一條凸曲線都成立，不限於圓。

逐步解題

階段 1　由坐標求面積

想一想　設 $x = r\cos\theta$ 和 $y = r\sin\theta$，全部都是 t 的函數，計算 $x\dot{y} - y\dot{x}$，其中上方的點表示 $\frac{\mathrm{d}}{\mathrm{d}t}$。

$\dot{x} = \dot{r}\cos\theta - r\sin\theta\,\dot{\theta}$，而 $\dot{y} = \dot{r}\sin\theta + r\cos\theta\,\dot{\theta}$。於是

$$x\dot{y} - y\dot{x} = r\cos\theta\,(\dot{r}\sin\theta + r\cos\theta\,\dot{\theta}) - r\sin\theta\,(\dot{r}\cos\theta - r\sin\theta\,\dot{\theta}) = r^2\dot{\theta}.$$

所以 $\frac{1}{2}\int r^2\,\mathrm{d}\theta = \frac{1}{2}\int r^2\dot{\theta}\,\mathrm{d}t = \frac{1}{2}\int (x\dot{y} - y\dot{x})\,\mathrm{d}t$。 ■

這條公式不需要極坐標（polar coordinates），只需要把笛卡兒坐標（cartesian coordinates）下的位置寫成某個參數（parameter）的函數。當一點沿着一條包圍 O 的閉曲線逆時針繞行一周時，公式給出的就是曲線內的面積。

階段 2　桿的兩端

桿由 A 指向 B 的方向是 $(\cos t, \sin t)$，而 P 與 A 的距離是 a，與 B 的距離是 b。所以

$$A = (x - a\cos t,\ y - a\sin t), \qquad B = (x + b\cos t,\ y + b\sin t).$$

階段 3　OA 掃過的面積

想一想　把 $(*)$ 應用於 A。展開 $x_A\dot{y}_A - y_A\dot{x}_A$，並按 a 的冪次把各項整理好。

由 $x_A = x - a\cos t$、$y_A = y - a\sin t$，可得 $\dot{x}_A = \dot{x} + a\sin t$ 和 $\dot{y}_A = \dot{y} - a\cos t$。展開得

$$\begin{aligned} x_A\dot{y}_A - y_A\dot{x}_A &= (x - a\cos t)(\dot{y} - a\cos t) - (y - a\sin t)(\dot{x} + a\sin t) \\ &= (x\dot{y} - y\dot{x}) + a^2(\cos^2 t + \sin^2 t) - a[(x + \dot{y})\cos t + (y - \dot{x})\sin t]. \end{aligned}$$

由 0 積分至 2π 再除以二：第一項給出 $[P]$，第二項給出 $\frac{1}{2}a^2 \cdot 2\pi = \pi a^2$，第三項給出 $-af$。所以

$$[A] = [P] + \pi a^2 - af. \qquad ■$$

停一停，試一試　現在求 $[B]$。（提示：把 A 的公式中的 a 換成甚麼，就得到 B？）

$B = (x + b\cos t, y + b\sin t)$ 就是把 A 的公式中的 a 換成 $-b$。所以

$$[B] = [P] + \pi b^2 + bf.$$

階段 4 消去 f

想一想 A 和 B 都沿同一條曲線 $\mathcal{C}$ 繞行一周。由此可知 $[A]$ 和 $[B]$ 有甚麼關係？

當 t 由 0 變到 2π 時，A 和 B 各自沿 $\mathcal{C}$ 逆時針繞行一周，所以 $[A] = [B] =$ 面積$(\mathcal{C})$。令兩個表達式相等，

$$\pi a^2 - af = \pi b^2 + bf \quad\Longrightarrow\quad f(a+b) = \pi(a^2 - b^2) \quad\Longrightarrow\quad f = \pi(a-b).$$

由於 P 沿 $\mathcal{D}$ 繞行一周，$[P] =$ 面積$(\mathcal{D})$，於是

$$\text{面積}(\mathcal{C}) - \text{面積}(\mathcal{D}) = [A] - [P] = \pi a^2 - a\pi(a-b) = \pi ab. \qquad \blacksquare$$

那個取決於曲線形狀的神祕量 f，自始至終沒有計算過：它被消去了。

回顧

- **切合問題的公式。**面積公式 $(*)$ 只需要把坐標寫成參數的函數，而桿的運動正是這樣描述的。
- **兩個表達式，一個未知量。**我們無法計算 f，但手上有兩條含有它的方程，於是把它消去，就像解聯立方程（simultaneous equations）一樣。
- **先檢驗特殊情況。**動筆之前，圓的情況已經給出答案 πab，讓我們知道目標是甚麼。

背後的數學

格林定理與求積儀

公式 $(*)$ 是格林定理（Green's theorem）的特殊情況；格林定理把沿閉曲線的積分與曲線內區域上的積分聯繫起來。它的面積形式指出，曲線所包圍的面積是 $\frac{1}{2}\oint(x\,\mathrm{d}y - y\,\mathrm{d}x)$，這個積分只需沿邊界走一圈就能算出。工程師利用這一點造出了求積儀（planimeter），這是十九世紀發明的一種機械儀器（其中以 Jakob Amsler 在 1854 年的設計最為成功）：用指針沿着地圖或圖表上某個圖形的邊界描一圈，一個小輪就會記錄所包圍的面積。求積儀曾用來量度土地面積、根據示功圖（indicator diagram）計算蒸汽機所作的功，還有許多其他用途，直至被數碼方法取代為止。Holditch 定理與此密切相關：那根桿就像求積儀的臂。

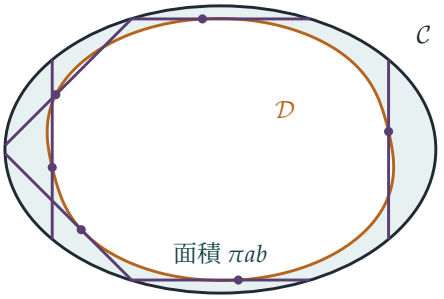


圖 4.9.1　Holditch 定理的例子：橢圓的半軸（semi-axes）為 3 和 2，桿長 2.5，其中 $a = 1$、$b = 1.5$。圖中畫出了桿的若干位置（紫色），並標出點 P。P 所描出的曲線 $\mathcal{D}$（琥珀色）是用數值方法算出的，而兩條曲線之間的陰影環形面積為 $4.7124\ldots = 1.5\pi$，與定理的預測完全吻合。

細節條款

Holditch 陳述這個定理時，對條件不大講究，但定理其實需要一些條件。桿必須能夠在兩端都留在曲線上的情況下繞足一整圈，而描出的曲線 $\mathcal{D}$ 必須是只描一次的簡單閉曲線（simple closed curve）；如果桿相對於 $\mathcal{C}$ 的闊度太長，$\mathcal{D}$ 可能會自相交，這時公式給出的「面積」會把某些區域計算兩次，或以負號計算。題目假設 $\mathcal{D}$ 是一條包含 O 的凸曲線，從而避開了這些困難。找出定理究竟在甚麼條件下成立，正好說明一個優美的想法必須經過嚴格化才站得住腳，這也是第 3 章的主題。

思維技巧

算不出來的，就把它消去

如果某個量太複雜而無法計算，就找兩條含有它的不同方程，把它消去。

這裏 f 取決於曲線的每一個細節，但它同時出現在 $[A]$ 和 $[B]$ 的表達式中，而 $[A] = [B]$ 這個事實把它消除了。答案 πab 之所以能夠與曲線無關，全因取決於曲線的部分互相抵消了。

延伸練習

1 **橢圓的面積。**利用 $(*)$，取 $x = p\cos t$、$y = q\sin t$，證明半軸為 p 和 q 的橢圓面積為 πpq。

2 **圓的情況。**證明相交弦定理：若圓的兩條弦相交於 P，則每條弦被分成的兩段，其長度之積相等。由此推導出圓的 Holditch 定理。

3 **中點。**證明由 $f = \pi(a - b)$ 可知，當 P 是桿的中點時，$f = 0$。這時環形的面積是多少？答案以桿長 L 表示。

4 **三角形。**較難利用 $(*)$，沿邊界繞行一周，求頂點為 $(0,0)$、(x_1, y_1)、(x_2, y_2) 的三角形的面積，並證明它等於 $\frac{1}{2}|x_1 y_2 - x_2 y_1|$。這就是求任意多邊形面積的「鞋帶公式」（shoelace formula）的起點。

延伸閱讀

- **Mark Levi, *The Mathematical Mechanic***（Princeton University Press）。以桿、輪和彈簧等物理推理來證明數學結果，與本節的精神一脈相承。
- **H. M. Schey, *Div, Grad, Curl, and All That***（Norton）。向量微積分（vector calculus）的非正式入門，內容包括格林定理。
- **Hamnet Holditch, "Geometrical theorem"**，*Quarterly Journal of Pure and Applied Mathematics* 2 (1858)。原始的短文，只有短短一段，完全沒有提及任何細節條款。

我們學到了甚麼

本章開始時，我們提出了三個問題：積分（integral）究竟是甚麼？積分可以證明甚麼？當一個積分沒有公式時，又該怎麼辦？以下是本章九道題目給出的答案。

1. 積分是甚麼？哪些函數有積分？

在 Section 4.1 中，我們把積分定義為由窄長矩形組成的下和（lower sum）與上和（upper sum）的共同極限，並單憑這個定義計算出 $\int_0^a x^5\,\mathrm{d}x$。這要花一番工夫：先以裂項相消求和（telescoping sum）處理連續整數乘積之和，再從兩邊估計上下界，最後用夾逼（squeeze）得出結果。花了多少力氣，正好說明微積分基本定理（fundamental theorem of calculus）讓你不勞而獲的有多少。

這個定義也告訴我們，並非每個函數都可積（integrable）。狄利克雷（Dirichlet）函數在有理數上取值 1，在無理數上取值 0，它的上和全部等於 1，下和全部等於 0。正是這類問題，促使勒貝格（Lebesgue）建立了一套更強大的積分理論，你日後在大學便會學到。

2. 積分可以證明甚麼？

遠遠不止面積。

- **數與數之間的不等式。**一個正函數的積分等於 $\frac{22}{7}-\pi$，所以 $\frac{22}{7} > \pi$（Section 4.3）。「把 $A-B$ 寫成某個正函數的積分」這個想法既簡單，又非常有用。
- **無窮乘積與特殊值。**利用 $\int \sin^n t\,\mathrm{d}t$ 的歸約公式（reduction formula）、一個不變量（invariant）和一次夾逼，便得到沃利斯（Wallis）乘積、斯特靈（Stirling）公式中的常數，以及 $\left(\frac{1}{2}\right)! = \frac{\sqrt{\pi}}{2}$（Section 4.4）。
- **無窮和。**把 x^x 展開成冪級數（power series）再逐項積分，便得到「二年級生之夢」（sophomore's dream）（Section 4.5）。
- **三角學的基礎。**定義 $\mathrm{T}(x) = \int_0^x \frac{\mathrm{d}u}{1+u^2}$ 及 $\pi = 2\mathrm{T}(\infty)$ 之後，單靠代換（substitution）便推出了加法公式，以及 $\mathrm{T}(\sqrt{3}) = \frac{\pi}{3}$ 等特殊值（Section 4.2）。
- **幾何。**一個化簡後變成常數的被積函數（integrand），揭示了阿基米德（Archimedes）的帽盒定理（hat-box theorem）（Section 4.8）；而把面積公式寫成沿邊界的積分，則

證明了令人驚歎的 Holditch 定理（Section 4.9）。

- **最優函數。**在積分內配方（completing the square），便準確指出 $\int\left((y')^2 - a^2y^2\right)$ 在甚麼情況下必定非負，而臨界值 $\frac{\pi}{2}$ 原來是一個特徵值（eigenvalue）（Section 4.7）。

3. 沒有公式時，我們可以怎麼辦？

在 Section 4.6 中，我們證明了 $\int \frac{e^x}{x+1}\,\mathrm{d}x$ 不能寫成有理函數（rational function）乘以 e^x 的形式，而劉維爾（Liouville）的理論更表明它根本沒有初等（elementary）公式。從極點（pole）的角度看，原因是其原函數（antiderivative）在 $x = -1$ 附近的表現必須像對數函數，而有理函數做不到這一點。然而這個積分確確實實存在，我們也可以把它算出來：可以用數值方法，可以用級數（如 Section 4.5 中處理 $\int x^x$ 那樣），也可以給它起一個名字，把它當作一個新函數來研究，就像我們在 Section 3.9 中對待 W 函數那樣。

4. 積分工具箱

方法	你在哪裏用過
從定義出發	由矩形求 $\int_0^a x^5$（4.1）
裂項相消求和	連續整數乘積之和（4.1）
保持積分不變的代換	由 $\int \frac{\mathrm{d}u}{1+u^2}$ 建立三角學（4.2）
多項式除法	$\frac{22}{7} - \pi$（4.3）
正性與被積函數的估界	$\pi < \frac{22}{7}$（4.3）
歸約公式	$\int \sin^n t$（4.4）；$\int x^m(\ln x)^n$（4.5）
不變量與夾逼	$nu_nu_{n-1} = \frac{\pi}{2}$（4.4）
逐項積分	二年級生之夢（4.5）
因式與次數論證	不存在有理原函數（4.6）
配方法	變分法（calculus of variations）（4.7）
先化簡，後積分	球形麵包（4.8）
以邊界積分表示面積	Holditch 定理（4.9）

下一站

- Section 4.4 中的伽瑪函數（Gamma function）和貝塔函數（Beta function），以及斯特靈公式中的常數 $\sqrt{2\pi}$，會在 Section 12.23 的概率問題中再度出現。
- Section 4.7 的特徵值問題，到了 Section 8.6 會變成量子諧振子（quantum harmonic oscillator）。
- 梅欽（Machin）公式由 Section 4.2 的加法恆等式建立而成，是 Section 5.1 的主題。
- 第 5 章轉向複數（complex numbers），它解釋了為何本章有這麼多積分都與 π 有關。

如果你想在上大學前走得更遠

- **Paul Nahin, *Inside Interesting Integrals***（Springer）：一場巧妙積分的盛宴，並道出背後的種種竅門。
- **William Dunham, *The Calculus Gallery***：從牛頓（Newton）到勒貝格的分析學傑作。
- **Michael Spivak, *Calculus***：嚴謹地講述積分的定義和微積分基本定理。
- **Sherman Stein, *Archimedes***：追溯積分的古代源頭。

埃爾米特（Hermite）對沒有導數的函數避之則吉，但積分從來沒有這個問題：它十分寬容，連微分無法處理的函數也照單全收，而且不斷帶來驚喜。如果說有一個主題貫穿本章，那就是：積分不只是一種求面積的技巧。它是一種語言，用來表達和證明關於數、函數和圖形的事實，而在你往後的數學生涯中，你都會一直說這種語言。

延伸練習答案

以下只是簡短的答案和提示，並非完整解答：只列出結果和關鍵步驟，方便你核對自己的做法。請先認真做過每道練習，然後才看答案。

4.1 甚麼是積分？

1. 這個和是 $y=\mathrm{f}(x)$ 之下、在 $[1,2]$ 上 n 個闊度為 $\frac{1}{n}$ 的長方形的總面積。取 $\mathrm{f}(x)=\frac{1}{x}$，得 $\frac{1}{n}\sum\frac{1}{1+m/n}=\sum\frac{1}{n+m}\to\ln 2$。第二個和是 $\frac{1}{n}\sum\frac{1}{1+(m/n)^2}\to\int_0^1\frac{\mathrm{d}x}{1+x^2}=\frac{\pi}{4}$。

2. 分成 n 條時，上和（upper sum）是 $\frac{a}{n}\sum_{r=1}^{n}\left(\frac{ra}{n}\right)^2=\frac{a^3(n+1)(2n+1)}{6n^2}$，下和（lower sum）則只須把第一個因子中的 $n+1$ 換成 $n-1$；兩者都趨向 $\frac{a^3}{3}$。

3. 各項可以裂項相消（telescope）：$r(r+1)\cdots(r+k)-(r-1)r\cdots(r+k-1)=(k+1)\,r(r+1)\cdots(r+k-1)$。

4. $\sum_{r=1}^{n}r^5=\frac{n^2(n+1)^2(2n^2+2n-1)}{12}$。

4.2 由一個積分建立三角學

1. 導數（derivative）是 $-\mathrm{f}(a-x)\mathrm{f}(x)+\mathrm{f}(a-x)\mathrm{f}(x)=0$，所以 $\mathrm{f}(a-x)\mathrm{f}(x)=\mathrm{f}(a)$；代入 $a=x+y$ 即可。至於餘弦，固定 y，設 $p(x)=\mathrm{c}(x+y)-\mathrm{c}(x)\mathrm{c}(y)+\mathrm{s}(x)\mathrm{s}(y)$ 及 $q(x)=\mathrm{s}(x+y)-\mathrm{s}(x)\mathrm{c}(y)-\mathrm{c}(x)\mathrm{s}(y)$。於是 $p'=-q$ 而 $q'=p$，所以 p^2+q^2 是常數，而它在 $x=0$ 時等於 0。

2. $\frac{\mathrm{d}y}{\mathrm{d}u}=(1+u^2)^{-3/2}$ 而 $\sqrt{1-y^2}=(1+u^2)^{-1/2}$，所以 $\frac{\mathrm{d}y}{\sqrt{1-y^2}}=\frac{\mathrm{d}u}{1+u^2}$。於是 $\mathrm{S}(\frac{1}{\sqrt{2}})=\mathrm{T}(1)$，而在第 (i) 部分取 $x=1$ 便得 $\mathrm{T}(1)=\frac{1}{2}\mathrm{T}_\infty$。

3. 取 $a=\frac{1}{2}$ 及 $x=\frac{1}{3}$，則 $\frac{x+a}{1-ax}=1$，所以 $\mathrm{T}(\frac{1}{3})=\mathrm{T}(1)-\mathrm{T}(\frac{1}{2})$。

4. 在 $[0,1]$ 上 $\frac{1}{1+u^2}\le 1$，在 $[1,x]$ 上則 $\le\frac{1}{u^2}$，所以當 $x\ge 1$ 時，$\mathrm{T}(x)\le 1+(1-\frac{1}{x})<2$；當 $0<x<1$ 時，$\mathrm{T}(x)<\mathrm{T}(1)\le 1$。（上界 $1+\int_1^x u^{-2}\,\mathrm{d}u$ 在 $x<1$ 時並不成立：在 $x=\frac{1}{2}$ 時它等於 0。）T 遞增而且有界（bounded），所以它有極限（limit）。

4.3 22/7 大於 π

1. $\frac{22}{7}-\pi\approx 0.001264$，介乎 $\frac{1}{1260}\approx 0.000794$ 與 $\frac{1}{630}\approx 0.001587$ 之間。

2. $\frac{(4!)^2}{9!} = \frac{576}{362\,880} = \frac{1}{630}$。

3. $\frac{x(1-x)^2}{1+x^2} = x - 2 + \frac{2}{1+x^2}$，積分（integrate）後得 $\frac{1}{2} - 2 + \frac{\pi}{2}$。同樣地，$\frac{x^2(1-x)^4}{1+x^2} = x^4 - 4x^3 + 5x^2 - 4 + \frac{4}{1+x^2}$，積分後得 $\frac{1}{5} - 1 + \frac{5}{3} - 4 + \pi = \pi - \frac{47}{15}$。兩個積分都是正數。這兩個被積函數（integrand）遠大於 $\frac{x^4(1-x)^4}{1+x^2}$（其最大值約為 0.003），所以誤差 $\frac{\pi}{2} - \frac{3}{2} \approx 0.07$ 和 $\pi - \frac{47}{15} \approx 0.008$ 也較大。

4. 多項式除法（polynomial division）餘下 $\frac{16}{1+x^2}$，積分後得 4π；整個積分是 $4\pi - \frac{188684}{15015} \approx 3.6 \times 10^{-6}$，所以 $\pi \approx 3.14159174$。

4.4 沃利斯、Beta 與二分之一的階乘

1. $u_{2m} = \frac{1\cdot3\cdots(2m-1)}{2\cdot4\cdots2m} \cdot \frac{\pi}{2}$ 而 $u_{2m+1} = \frac{2\cdot4\cdots2m}{3\cdot5\cdots(2m+1)}$。於是 $1 < \frac{u_{2m}}{u_{2m+1}} < \frac{u_{2m-1}}{u_{2m+1}} = \frac{2m+1}{2m}$，所以這個比趨向 1，整理後即得 $\frac{\pi}{2} = \prod_{k\geq1} \frac{(2k)^2}{(2k-1)(2k+1)}$。

2. 分部積分法（integration by parts）給出 $I_n = \frac{n}{n+1}\int_0^1 x^{n+1}(1-x)^{n-1}\,\mathrm{d}x$，再寫成 $x^{n+1}(1-x)^{n-1} = x^n(1-x)^{n-1} - x^n(1-x)^n$ 並利用對稱性，便得 $I_n = \frac{n}{2(2n+1)} I_{n-1}$；用歸納法（induction）可得 $\frac{(n!)^2}{(2n+1)!}$。取 $x = \sin^2\theta$，得 $I_{1/2} = \int_0^{\pi/2} 2\sin^2\theta\cos^2\theta\,\mathrm{d}\theta = \frac{\pi}{8}$。公式提示 $\left((\frac{1}{2})!\right)^2/2! = \frac{\pi}{8}$，所以 $(\frac{1}{2})! = \frac{\sqrt{\pi}}{2}$（即 $\Gamma(\frac{3}{2})$）。

3. 對稱性來自 $x \mapsto 1-x$；遞推關係（recurrence）來自 $x^p = x^{p-1} - x^{p-1}(1-x)$；而取 $x = \frac{t}{1+t}$ 時，$1 - x = \frac{1}{1+t}$ 且 $\mathrm{d}x = \frac{\mathrm{d}t}{(1+t)^2}$。該積分是 $\mathrm{b}(\frac{5}{2}, \frac{7}{2})$；利用遞推關係和對稱性化簡至 $\mathrm{b}(\frac{1}{2}, \frac{1}{2}) = \pi$，便得 $\frac{3\pi}{256}$。

4. 由第 (i) 部分的顯式數值，$u_{2m} = \frac{(2m)!}{(2^m m!)^2} \cdot \frac{\pi}{2} = \binom{2m}{m}\frac{\pi}{2^{2m+1}}$。於是由 $2mu_{2m}^2 \to \frac{\pi}{2}$ 得 $\binom{2m}{m}^2 \approx \frac{4^{2m}}{\pi m}$。當 $m = 10$ 時：$\binom{20}{10} = 184\,756$，而 $\frac{4^{10}}{\sqrt{10\pi}} \approx 187\,079$（高出約 1.3%）。

4.5 大二學生之夢

1. $x^{-x} = e^{-x\ln x} = \sum_{n\geq0} \frac{(-x\ln x)^n}{n!}$，而 $\int_0^1 \frac{x^n(-\ln x)^n}{n!}\,\mathrm{d}x = \frac{1}{(n+1)^{n+1}}$（代入 $x = e^{-u}$）。所以該積分是 $\sum_{n\geq1} n^{-n} \approx 1.2913$。

2. $\frac{\mathrm{d}}{\mathrm{d}x}x^x = x^x(\ln x + 1)$，在 $x = \frac{1}{e}$ 時為零；其值為 $(e^{-1})^{e^{-1}} = e^{-1/e} \approx 0.692$。

3. 各項是 $\frac{1}{n^n}$。取 7 項後停止，誤差小於 $\frac{1}{8^8} \approx 6 \times 10^{-8}$，已經足夠；只取 6 項的話，上界 $\frac{1}{7^7} \approx 1.2 \times 10^{-6}$ 並不足夠。所以要取 7 項。

4. 用分部積分法：$\int_0^1 x^2 \ln x\,\mathrm{d}x = \left[\frac{x^3\ln x}{3}\right]_0^1 - \int_0^1 \frac{x^2}{3}\,\mathrm{d}x = -\frac{1}{9}$。一般而言 $\int_0^1 x^n \ln x\,\mathrm{d}x = -\frac{1}{(n+1)^2}$，而 $\frac{\ln x}{1-x} = \sum_{n\geq0} x^n \ln x$。

4.6 沒有公式的積分

1. 求導：$\frac{\mathrm{d}}{\mathrm{d}x}\frac{e^x}{x+1} = \frac{xe^x}{(x+1)^2}$，而 $\frac{\mathrm{d}}{\mathrm{d}x}\frac{(x-1)e^x}{x+1} = e^x\left(\frac{x-1}{x+1} + \frac{2}{(x+1)^2}\right) = \frac{(x^2+1)e^x}{(x+1)^2}$。

2. $\frac{\mathrm{d}}{\mathrm{d}x}\left(\mathrm{P}e^{-x^2}\right) = (\mathrm{P}' - 2x\mathrm{P})e^{-x^2}$，而 $\mathrm{P}' - 2x\mathrm{P}$ 的次數是 $\deg \mathrm{P} + 1 \geq 1$（當 $\mathrm{P} \neq 0$），所以它不可能等於 1。

3. 假如 $\mathrm{R} = \mathrm{P}/\mathrm{Q}$（已約至最簡）在 0 處有 $k \geq 1$ 階的極點（pole），則 $\mathrm{R}' + \mathrm{R}$ 在該處會有 $k+1 \geq 2$ 階的極點，但 $\frac{1}{x}$ 的極點只是 1 階。假如 R 在 0 處沒有極點，則 $\mathrm{R}' + \mathrm{R}$ 也沒有。無論哪種情況，都有 $(\mathrm{R}e^x)' \neq \frac{e^x}{x}$。

4. 對 $\frac{\mathrm{P}e^x}{x+1}$ 求導，得 $e^x\frac{(x+1)\mathrm{P}' + x\mathrm{P}}{(x+1)^2}$。在 $x = -1$ 處：$\mathrm{f}(-1) = -\mathrm{P}(-1)$，而由 $\mathrm{f}' = (x+1)\mathrm{P}'' + (x+1)\mathrm{P}' + \mathrm{P}$ 得 $\mathrm{f}'(-1) = \mathrm{P}(-1)$。對於 $x^3 - 2$：$-3+3=0$；對於 x：$-1+1=0$；對於 x^2+1：$2-2=0$。這個條件也是充分的：映射 $\mathrm{P} \mapsto (x+1)\mathrm{P}' + x\mathrm{P}$ 把次數小於 d 的多項式映到次數不超過 d 的多項式，而且是單射（injective），所以它的像（image）的維數（dimension）是 d，並且包含在由滿足 $\mathrm{f}(-1) + \mathrm{f}'(-1) = 0$ 的 f 所組成的 d 維空間之內；因此像就是整個空間。

4.7 變分法

1. 取 $y = \cos\frac{\pi x}{2}$，得 $J = \frac{\pi^2}{4}\int_0^1 \sin^2 - a^2\int_0^1 \cos^2 = \frac{1}{2}\left(\frac{\pi^2}{4} - a^2\right) < 0$。

2. $J(y+\varepsilon\eta) = J(y) + 2\varepsilon\int_0^1 (y'\eta' - a^2 y\eta) + \varepsilon^2 J(\eta)$；要在 $\varepsilon = 0$ 處取得最小值，線性項必須為零。分部積分後，對每個滿足 $\eta(1) = 0$ 的 η 都有 $-y'(0)\eta(0) - \int_0^1 (y'' + a^2 y)\eta = 0$，由此必有 $y'' + a^2 y = 0$ 及 $y'(0) = 0$。

3. 取 $x = Lu$，得 $\int_0^L \left((y')^2 - a^2y^2\right)\mathrm{d}x = \frac{1}{L}\int_0^1 \left((Y')^2 - (aL)^2Y^2\right)\mathrm{d}u$，所以條件是 $aL \leq \frac{\pi}{2}$。（當 $aL = \frac{\pi}{2}$ 時，正如正文所述，積分仍然不會是負數，但 $y = \cos\frac{\pi x}{2L}$ 會令它等於零；當 $aL > \frac{\pi}{2}$ 時，同一個函數會令它變成負數。）

4. 擺線（cycloid）是 $x = R(\theta - \sin\theta), y = R(1-\cos\theta)$。關於等時曲線（tautochrone），見 Section 11.4：以弧長（arc length）表示時，該運動恰好是簡諧運動（simple harmonic motion）。

4.8 切球形麵包

1. $V/A = \frac{\frac{4}{3}\pi a^3}{4\pi a^2} = \frac{a}{3}$。但餘下一塊高度為 t 的部分有 $A = 2\pi a t$ 及 $V = \frac{1}{3}\pi t^2(3a-t)$，所以 $V/A = \frac{t(3a-t)}{6a}$，它在 $t = \frac{3}{2}a$ 時取得最大值 $\frac{3a}{8}$：比 $\frac{a}{3}$ 還要大。

2. 這條帶的高度是 $2R\sin 23.4°$，而球的高度是 $2R$，所以所佔比例是 $\sin 23.4° \approx 0.40$。

3. 對於 $y = \sqrt{x}$，$y\sqrt{1+(y')^2} = \sqrt{x + \frac{1}{4}}$，並非常數。由 $y\sqrt{1+(y')^2} = c$ 得 $(y')^2 =$

$\frac{c^2}{y^2} - 1$，其解為 $y = c$（圓柱）以及圓 $(x - x_0)^2 + y^2 = c^2$（球）。

4. 球的體積是 $\frac{4}{3}\pi a^3$，圓柱的體積是 $\pi a^2 \cdot 2a = 2\pi a^3$；兩者之比是 $\frac{2}{3}$。

4.9 Holditch 定理

1. $\frac{1}{2}\int_0^{2\pi}(p\cos t \cdot q\cos t + q\sin t \cdot p\sin t)\,\mathrm{d}t = \frac{1}{2}\cdot 2\pi pq = \pi pq$。

2. 把兩條弦（chord）的端點連起來所得的兩個三角形相似（同弧所對的角相等），所以 $PA \cdot PB = PC \cdot PD$。對於半徑為 R 的圓，若 P 與圓心的距離為 r，則通過 P 的直徑給出 $(R - r)(R + r)$，所以 $ab = R^2 - r^2$，而圓環的面積是 $\pi R^2 - \pi r^2 = \pi ab$。

3. 若 $a = b$，則 $f = \pi(a - b) = 0$，而圓環的面積是 $\pi a^2 = \frac{\pi L^2}{4}$：即以該桿為直徑的圓的面積。

4. 沿着通過 O 的兩條邊，$x\,\mathrm{d}y - y\,\mathrm{d}x = 0$。第三條邊以 $(x_1, y_1) + s(x_2 - x_1, y_2 - y_1)$（其中 $0 \le s \le 1$）作參數表示，在這條邊上它等於 $(x_1 y_2 - x_2 y_1)\,\mathrm{d}s$。所以面積是 $\frac{1}{2}|x_1 y_2 - x_2 y_1|$。

章

5

複數

實數域中兩個眞理之間的最短路徑，要取道複數域。

本章內容

5.1	梅欽公式	STEP III 1987
5.2	單位根篩選法	STEP I 1990
5.3	因式分解 $z^{2n}+1$	STEP III 2013
5.4	正多邊形中的距離	STEP III 2016
5.5	名不副實的有理角	STEP III 2022
5.6	費馬點	STEP III 2000
5.7	作為映射的複指數函數	STEP II 1997
5.8	圓盤中隱藏的距離	STEP III 2013
A	練習答案	第 367 頁

虛數，實在的答案

實數域中兩個眞理之間的最短路徑，要取道複數域。

— 阿達馬（Jacques Hadamard）

1545 年，卡爾達諾（Gerolamo Cardano）發表了一條解三次方程（cubic equation）的公式，卻發現了一件令人不安的事。有些三次方程明明有三個不折不扣的實根，公式在計算途中卻要求取負數的平方根。這些平方根到最後會互相抵消，留下實數答案，但在計算的中途，卻出現了一些按當時任何標準都不存在的量。卡爾達諾形容這種運算「既精妙，又無用」。其後兩個世紀，數學家用起這些「虛」數（imaginary number）時總是惴惴不安，只當它是一種碰巧行得通的竅門。

要到十九世紀初，Caspar Wessel、Jean-Robert Argand 和高斯（Gauss）把複數（complex number）看成平面上的點，複數才得到認可。自此之後，人們發現複數並非竅門，而是一種啟示。複數解釋了實數世界的現象為何如此表現，而且不斷為有關實數的結果提供最短的路徑，正如阿達馬所言。

複數為何重要

每一條多項式方程在複數中都有解：這就是代數基本定理（fundamental theorem of algebra）。因此，在一個確切的意義下，代數到了複數才算完備。複數也是描述旋轉和振盪的天然語言。乘以一個模（modulus）為 1 的複數，就是把平面旋轉，所以凡是涉及角度、波動或週期現象的問題，用複數處理都會較為簡單。電機工程師用複數分析交流電；物理學家沒有複數就寫不出量子力學；由手提電話到磁力共振（MRI）掃描器，訊號處理全靠複數傅立葉變換（Fourier transform）運作。而在純數學中，複變函數原來有着驚人的剛性和美感：只要知道一個複函數在一小片區域上的取值，它在各處的取值就全部確定了。

你在大學會遇到甚麼

你從第一個星期起就會用到複數，但真正的一大步是複分析（complex analysis），通常在二年級修讀。它的核心對象是全純（holomorphic）函數，即可以求導的複函數，而這套理論處處是驚喜。每個全純函數都自動地無限次可導，並且等於它的泰勒級數（Taylor series）。沿閉曲線的積分，只需察看函數趨向無窮的那些點便可求出（柯西留數定理，Cauchy's residue theorem），讓你能夠計算其他方法都束手無策的實積分。這套理論再

往前走，就通往保角映射（conformal map）（Section 5.7）、黎曼曲面（Riemann surface），以及支配質數的黎曼 zeta 函數（Riemann zeta function）（Section 1.2）。與此同時，單位根（roots of unity）的代數（Section 5.2、Section 5.3）通往伽羅瓦理論（Galois theory），而複幾何則通往 Section 5.8 的非歐幾何（non-Euclidean geometry）。

貫穿本章的問題

三個大問題

1. **複數怎樣幫助解決實數問題？**為甚麼繞道複平面（complex plane），竟能得出關於 π、二項式係數（binomial coefficient）或餘弦的事實？
2. **單位根有甚麼特別之處？**為甚麼 $z^n = 1$ 的解不斷在求和、乘積、多邊形和角度之中出現？
3. **複函數是甚麼模樣的？**輸入和輸出都是平面上的點的函數，我們怎樣把它畫出來？它們又揭示了怎樣的幾何？

本章由輻角（argument）與乘法出發（梅欽公式），經過單位根的代數（篩選、因式分解、多邊形的距離和無理角），走到用複數處理的幾何（費馬點），最後到達作為映射的複函數（指數函數，以及單位圓盤中一個隱藏的距離）。

本章的主要想法

§	課題	想法
5.1	梅欽（Machin）公式	複數相乘時，輻角相加
5.2	每隔三項的二項式係數	對單位根取平均，可按餘數（residue）篩選
5.3	餘弦的乘積	按各根分解因式，然後代入
5.4	正多邊形的弦	距離就是模；$z^n - 1 = \prod(z - \omega^k)$
5.5	$\cos^{-1}\frac{1}{3}$ 是無理的	棣美弗（De Moivre）定理、奇偶性（parity）與最小反例（minimal counterexample）
5.6	費馬點（Fermat point）	旋轉就是乘法；選取好的坐標
5.7	把 e^w 看成映射	把複函數畫出來；多對一映射（many-to-one map）
5.8	一個隱藏的距離	把不等式看成雙曲幾何（hyperbolic geometry）中的三角不等式（triangle inequality）

怎樣閱讀本章

動手畫圖。本章幾乎每一個結果都有幾何意義，隨手畫一幅複平面的草圖（單位根在哪裏、某個乘法把平面轉向哪一邊、網格線落在何處），往往在動用任何代數之前，就已

經告訴你該怎樣做。然後要小心地做代數運算，因為在正負號和 2π 的倍數上，圖形可能會誤導你。

5.1 梅欽公式

題目

STEP III 1987，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/87-s3-q3`

引子

1706 年，倫敦數學家梅欽（John Machin）徒手把 π 計算到小數點後 100 位，在此之前從未有人接近過這個紀錄。他的祕訣就是第 (ii) 部分的公式：這條公式把 $\frac{\pi}{4}$ 用幾個角表示出來，而這些角的正切都是很小的分數。分數愈小，級數（series）$\tan^{-1} x = x - \frac{x^3}{3} + \frac{x^5}{5} - \cdots$ 收斂（converge）得愈快。梅欽公式及其各種變體，在其後兩個半世紀一直是計算 π 的標準方法，直到電腦和 Section 2.6 的方法出現才被取代。

這條公式看起來像是三角恆等式（identity）變出來的奇蹟。這道題目卻說明，它其實是關於複數（complex number）乘法的一個事實：複數相乘時，它們的輻角（argument）相加。只要把幾個實部、虛部都是整數的複數乘一次，整條公式就得證。

動筆之前

想一想 $5 - \mathrm{i}$ 的輻角是甚麼？$1 + \mathrm{i}$ 的呢？複數相乘時，輻角會怎樣變化？

$\arg(5 - \mathrm{i}) = -\tan^{-1}\frac{1}{5}$，是一個很小的負角，而 $\arg(1 + \mathrm{i}) = \frac{\pi}{4}$。複數相乘時，模（modulus）相乘，輻角相加（可相差 2π 的倍數）。因此

$$\arg((5 - \mathrm{i})^4(1 + \mathrm{i})) = \frac{\pi}{4} - 4\tan^{-1}\frac{1}{5} \quad \text{（可能再加上若干個}2\pi\text{）}.$$

如果我們能算出這個乘積，再用另一個方法求出它的輻角，就得到一條恆等式。

逐步解題

階段 1　模的上下界

想一想　求上界（upper bound）時，把兩邊平方。求下界（lower bound）時，把 $|x|\cos\alpha + |y|\sin\alpha$ 看成兩個向量的點積（dot product）。

上界：$(|x|+|y|)^2 = x^2+y^2+2|x||y| \geq x^2+y^2 = |z|^2$，而且兩邊都非負，所以 $|z| \leq |x|+|y|$。

下界：$|x|\cos\alpha + |y|\sin\alpha$ 是長度為 $|z|$ 的向量 $(|x|,|y|)$ 與單位向量（unit vector）$(\cos\alpha, \sin\alpha)$ 的點積。點積不會超過兩個向量長度之積，所以它不超過 $|z|$。（代數證法：$|z|^2 - (|x|\cos\alpha + |y|\sin\alpha)^2 = (|x|\sin\alpha - |y|\cos\alpha)^2 \geq 0$。） ■

階段 2　關鍵的乘法

想一想　先計算 $(5-\mathrm{i})^2$，再把結果平方，然後乘以 $1+\mathrm{i}$。

$(5-\mathrm{i})^2 = 25 - 10\mathrm{i} - 1 = 24 - 10\mathrm{i}$。於是 $(24-10\mathrm{i})^2 = 576 - 480\mathrm{i} - 100 = 476 - 480\mathrm{i}$。最後

$$(476-480\mathrm{i})(1+\mathrm{i}) = 476 + 476\mathrm{i} - 480\mathrm{i} + 480 = 956 - 4\mathrm{i}.$$

停一停，試一試　現在比較輻角。$956-4\mathrm{i}$ 的輻角是甚麼？你憑甚麼肯定沒有多出 2π 的倍數？

$956 - 4\mathrm{i} = 4(239-\mathrm{i})$，其輻角為 $-\tan^{-1}\frac{1}{239}$。因此

$$\frac{\pi}{4} - 4\tan^{-1}\frac{1}{5} = -\tan^{-1}\frac{1}{239} + 2k\pi$$

其中 k 是某個整數。但 $0 < \tan^{-1}\frac{1}{5} < \frac{1}{5}$，所以左邊介乎 $\frac{\pi}{4} - \frac{4}{5}$ 與 $\frac{\pi}{4}$ 之間，即大約介乎 -0.02 與 0.79 之間；而右邊除去 $2k\pi$ 後介乎 -0.005 與 0 之間。由於 2π 遠大於兩者的差距，故 $k=0$。移項後得

$$\frac{\pi}{4} = 4\tan^{-1}\frac{1}{5} - \tan^{-1}\frac{1}{239}.$$ ■

圖 5.1.1 顯示了整個過程：由 $45°$ 出發，每乘一次 $5-\mathrm{i}$ 就順時針旋轉約 $11.3°$，乘了四次之後，剛好越過實軸，只過了 $0.24°$。

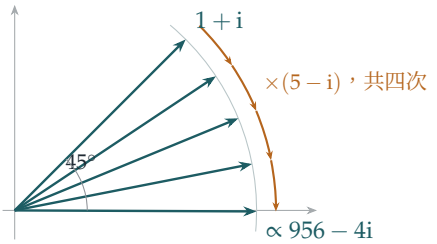


圖 5.1.1 把 $1+\mathrm{i}$ 乘以 $5-\mathrm{i}$ 四次。每乘一次就順時針旋轉 $\tan^{-1}\frac{1}{5}\approx 11.31°$（同時也會拉長，這裏略去不計）。四次旋轉把 $45°$ 帶到 $-0.24°$，即 $956-4\mathrm{i}$ 的輻角，也就是 $-\tan^{-1}\frac{1}{239}$。

階段 3 第二條公式

想一想 依樣畫葫蘆，這次應該計算哪一個乘積？

考慮 $(4+\mathrm{i})^3(20+\mathrm{i})(1985+\mathrm{i})$，它的輻角是 $3\tan^{-1}\frac{1}{4}+\tan^{-1}\frac{1}{20}+\tan^{-1}\frac{1}{1985}$。現在 $(4+\mathrm{i})^2=15+8\mathrm{i}$，所以 $(4+\mathrm{i})^3=(15+8\mathrm{i})(4+\mathrm{i})=52+47\mathrm{i}$。於是 $(52+47\mathrm{i})(20+\mathrm{i})=1040+52\mathrm{i}+940\mathrm{i}-47=993+992\mathrm{i}$，而

$$(993+992\mathrm{i})(1985+\mathrm{i})=(1971105-992)+(993+1969120)\mathrm{i}=1970113+1970113\mathrm{i}.$$

實部（real part）等於虛部（imaginary part），所以輻角是 $\frac{\pi}{4}$。三個角都是很小的正角（它們的和小於 $3\cdot\frac{1}{4}+\frac{1}{20}+\frac{1}{1985}<1<2\pi$），所以毋須擔心 2π 的倍數，公式由此得證。■

回顧

- **相乘時輻角相加。**單憑這一個事實，高斯整數（Gaussian integers）的乘積就變成了角與角之間的恆等式。
- **留意 2π 的倍數。**輻角只能確定到相差 2π 的倍數，所以每條恆等式都要靠估計大小來確定那個倍數。
- **整數包辦一切。**所有運算都只涉及整數；關於 π 的恆等式是精確地得出的，沒有用到任何近似。

背後的數學

徒手計算 π

為甚麼梅欽公式這樣適合用來計算？級數 $\tan^{-1} x = x - \frac{x^3}{3} + \frac{x^5}{5} - \cdots$ 在 $|x| \le 1$ 時收斂，但在 $x = 1$ 時慢得令人難受：要用 $1 - \frac{1}{3} + \frac{1}{5} - \cdots$ 求得 $\frac{\pi}{4}$ 的三位小數，大約需要一千項。取 $x = \frac{1}{5}$ 時，每一項約為前一項的 25 分之一，即每項大約帶來 1.4 位小數；取 $x = \frac{1}{239}$ 時，每項約 4.8 位。梅欽的 100 位小數，只需要第一個級數約 70 項和第二個級數約 20 項。1873 年，William Shanks 用一條梅欽型公式算出 707 位，單是運算就花了多年時間；到了 1944 年，人們發現他在第 528 位出了錯，其後的數字全部不對。

再談高斯整數

形如 $a + b\mathrm{i}$ 而 a、b 為整數的數，就是 Section 1.6 的高斯整數。$5 - \mathrm{i}$ 之所以這樣好用，原因在於一個因數分解（factorisation）：$239^2 + 1 = 57122 = 2 \times 13^4$，在高斯整數中寫成 $(239 + \mathrm{i})(239 - \mathrm{i}) = (1 + \mathrm{i})(1 - \mathrm{i})\,(3 + 2\mathrm{i})^4(3 - 2\mathrm{i})^4$，而 $5 - \mathrm{i} = (1 - \mathrm{i})(3 + 2\mathrm{i})$ 正是由同一批高斯質數（Gaussian primes）組成的。尋找梅欽型公式，其實就是把形如 $n^2 + 1$ 的數作因數分解的問題。1899 年，Carl Størmer 正是用這個想法證明了：形如 $\frac{\pi}{4} = a\tan^{-1}\frac{1}{m} + b\tan^{-1}\frac{1}{n}$ 而 a、b、m、n 為整數的公式只有四條，即梅欽的一條，以及歐拉（Euler）、Hermann 和 Hutton 分別發現的三條。

旋轉

乘以一個模為 1 的複數，就是把平面旋轉（rotation）。正因如此，凡是涉及旋轉和角的事情，複數都是最自然的語言：由本章稍後的幾何，到電機工程中的交流電，以至量子力學中的波函數，莫不如此。在 Section 4.2 中，$\tan^{-1}$ 的加法公式（addition formula）來自一個巧妙的代換（substitution）；在這裏，它來自複數的乘法。兩者其實是同一個事實，只是從不同的角度去看。

思維技巧

> **把角轉化為複數乘法**
>
> 角的和之間的恆等式，就是乘積的輻角之間的恆等式。把乘積精確地算出來，再讀出它的輻角。

複數的運算機械而精確，而處理反三角函數（inverse trigonometric functions）則容易出錯。把後者轉化為前者，往往是最快捷的證明途徑。

延伸練習

1 **一個近親。**試做 STEP II 2000，第 4 題（`step.maths.org/questions/00-s2-q4`）。這道題用與本節相同的方法，證明另一條關於 $\frac{\pi}{4}$ 的反正切公式：把複數相乘，然後讀出輻角。

2 **歐拉公式。**考慮 $(2+\mathrm{i})(3+\mathrm{i})$，證明 $\tan^{-1}\frac{1}{2}+\tan^{-1}\frac{1}{3}=\frac{\pi}{4}$。

3 **Hutton 公式。**證明 $\frac{\pi}{4}=2\tan^{-1}\frac{1}{3}+\tan^{-1}\frac{1}{7}$。

4 **要多少項？**估計一下，要把 π 計算到小數點後 20 位，梅欽公式中的兩個級數各需要多少項。然後用計算機實際算出最初幾位。

延伸閱讀

- **Jonathan Borwein & Peter Borwein, *Pi and the AGM***（Wiley）。書中載有梅欽型公式的歷史。
- **Jörg Arndt & Christoph Haenel, *Pi Unleashed***（Springer）。生動地介紹古今各種計算 π 的方法。
- **Tristan Needham, *Visual Complex Analysis***（Oxford University Press）。要從幾何角度理解複數，這是最好的一本書。第 1 章講的就是乘法即旋轉加縮放。

5.2 單位根篩選法

題目

STEP I 1990，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/90-s1-q2`

引子

帕斯卡三角形（Pascal's triangle）同一行的二項式係數（binomial coefficients）加起來等於 2^n：只要在 $(1+x)^n$ 中代入 $x=1$ 即可。若改為代入 $x=-1$，便得到交錯和，其值為 0；把兩式相加，可知位於偶數位置的係數之和為 2^{n-1}，剛好是一半。所以偶數位置和奇數位置的項把總和平分。

那麼每隔兩項取一個、即每第三個係數又如何？沒有任何實數 x 能夠挑出位置 $0,3,6,\ldots$。但有一個複數可以。三次單位根（cube roots of unity）就像一個濾波器：對 1、ω 和 ω^2 取平均，冪 ω^k 在 k 是 3 的倍數時給出 1，否則給出 0。這個想法稱為單位根濾波器（roots-of-unity filter），是離散傅立葉變換（discrete Fourier transform）的小型版本，而後者是有史以來最重要的算法之一。

動筆之前

? **想一想** ω 是 $z^3-1=0$ 的一個根。試把 z^3-1 因式分解。為甚麼由此可得 $1+\omega+\omega^2=0$？

$z^3-1=(z-1)(z^2+z+1)$。由於 $\omega^3=1$ 但 $\omega\neq 1$，ω 必定是第二個因式的根：$\omega^2+\omega+1=0$。從幾何上看，1、ω 和 ω^2 是一個以原點為中心的等邊三角形的頂點，它們的和就是該三角形的質心，即 0。

逐步解題

階段 1　模與輻角

由 $1+\omega+\omega^2=0$，得 $1+\omega^2=-\omega$。因為 $\omega=e^{2\pi\mathrm{i}/3}$，所以 $-\omega=e^{\mathrm{i}\pi}e^{2\pi\mathrm{i}/3}=e^{5\pi\mathrm{i}/3}=e^{-\pi\mathrm{i}/3}$。故 $1+\omega^2$ 的模（modulus）為 1，輻角（argument）為 $-\frac{\pi}{3}$。同理，$1+\omega=-\omega^2=e^{\pi\mathrm{i}/3}$，其模為 1，輻角為 $\frac{\pi}{3}$。

階段 2　濾波器

> **想一想**　當 k 是 3 的倍數時，$1+\omega^k+\omega^{2k}$ 等於甚麼？不是倍數時又如何？

若 $3\mid k$，則 $\omega^k=1$，和為 3。若否，ω^k 是 ω 或 ω^2，而無論哪一種情況，都有 $1+\omega^k+\omega^{2k}=1+\omega+\omega^2=0$。因此，表達式 $\frac{1}{3}(1+\omega^k+\omega^{2k})$ 在 $3\mid k$ 時等於 1，否則等於 0：它是一個偵測 3 的倍數的探測器。

> **停一停，試一試**　把 $(1+\omega^r)^n$ 對 $r=1,2,3$ 相加，每一項都用二項式定理（binomial theorem）展開。你得到甚麼？

由於 $\omega^3=1$，

$$(1+\omega)^n+(1+\omega^2)^n+(1+1)^n=\sum_{k=0}^{n}\binom{n}{k}(\omega^k+\omega^{2k}+1)=3\sum_{3\mid k}\binom{n}{k}.$$

另一方面，由階段 1，

$$(1+\omega)^n+(1+\omega^2)^n+2^n=e^{n\pi\mathrm{i}/3}+e^{-n\pi\mathrm{i}/3}+2^n=2\cos\frac{n\pi}{3}+2^n.$$

兩邊除以 3 便得到所求公式。 ■

階段 3　兩個計算

對於 $n=25$：$\frac{25\pi}{3}=8\pi+\frac{\pi}{3}$，所以 $\cos\frac{25\pi}{3}=\frac{1}{2}$，於是

$$\binom{25}{0}+\binom{25}{3}+\cdots+\binom{25}{24}=\frac{33554432+1}{3}=11184811.$$

> **想一想**　第二個和需要的是滿足 $k\equiv 2\pmod 3$ 的位置。怎樣修改濾波器，使它偵測的是 $k\equiv 2$ 而不是 $k\equiv 0$？

把探測器平移：$\frac{1}{3}(1+\omega^{k-2}+\omega^{2(k-2)})$ 恰好在 $k\equiv 2\pmod 3$ 時等於 1。由於 $\omega^{-2}=\omega$

及 $\omega^{-4}=\omega^2$，由此可得

$$\sum_{k\equiv 2}\binom{n}{k}=\frac{1}{3}\Big(2^n+\omega(1+\omega)^n+\omega^2(1+\omega^2)^n\Big).$$

取 $n=24$，則 $(1+\omega)^{24}=e^{8\pi i}=1$，同樣 $(1+\omega^2)^{24}=1$，所以括號內的值為 $2^{24}+\omega+\omega^2=2^{24}-1$。由 $2^{24}=16777216$，

$$\binom{24}{2}+\binom{24}{5}+\cdots+\binom{24}{23}=\frac{16777215}{3}=5592405.\qquad\blacksquare$$

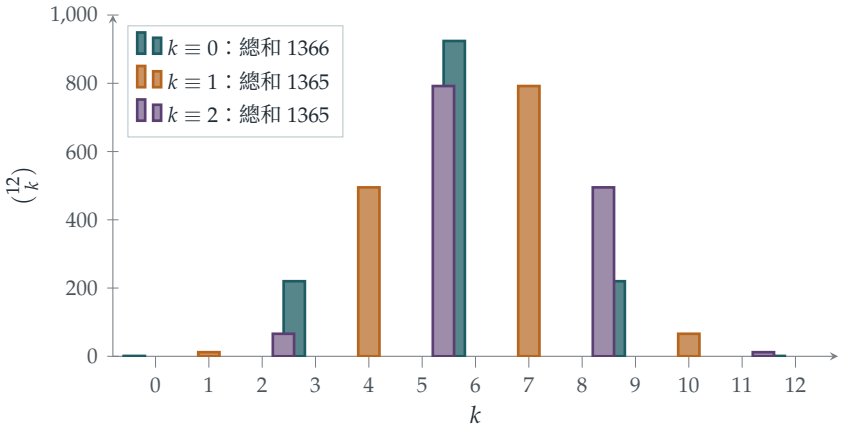


圖 5.2.1　帕斯卡三角形的第 12 行，按 k 除以 3 的餘數着色。三個類別的總和分別是 1366、1365 和 1365：各自幾乎剛好是 $2^{12}=4096$ 的三分之一。單位根濾波器準確地預測了它們與 $\frac{4096}{3}$ 之間的微小差異：$+\frac{2}{3}$、$-\frac{1}{3}$ 和 $-\frac{1}{3}$。

回顧

- **對單位根取平均，可按餘數篩選。**$\frac{1}{3}\sum_r \omega^{rk}$ 等於 1 還是 0，視乎 3 是否整除 k。把多項式在各單位根處的值相加，便能挑出每第三個係數。
- **平移濾波器。**取平均之前先乘以 ω^{-rj}，便能挑出滿足 $k\equiv j$ 的係數。
- **用複數求實數答案。**答案是整數，題目問的也是整數，但解題的路徑卻穿過了複平面。

背後的數學

離散傅立葉變換

把 3 換成任意的 N，並把 ω 換成 $e^{2\pi i/N}$。對於一列數 $a_0, a_1, \ldots, a_{N-1}$，以下各數

$$\hat{a}_r = \sum_{k=0}^{N-1} a_k \, \omega^{rk}$$

構成它的離散傅立葉變換，而正是濾波恆等式 $\sum_r \omega^{rk} = N$ 或 0，讓你可以從 $\hat{a}_r$ 還原出 a_k。這個變換把一個訊號分解成它的各個頻率成分。1965 年，James Cooley 和 John Tukey 發表了一種方法，只需大約 $N \log_2 N$ 步而非 N^2 步便能把它算出，這就是快速傅立葉變換（fast Fourier transform，簡稱 FFT）；高斯（Gauss）早在 1805 年已發現本質上相同的想法，但從未發表。FFT 應用於音訊和圖像壓縮（MP3 和 JPEG）、流動電話訊號、醫學掃描儀，以及極大數字的乘法，並且經常獲列為二十世紀最重要的算法之一。

生成函數

這個濾波器只是一個更宏大想法的其中一例：數列 $a_0, a_1, a_2, \ldots$ 可以包裝成一個多項式或冪級數（power series）$A(x) = \sum a_k x^k$，稱為它的生成函數（generating function），而對數列進行的運算就變成對函數進行的運算。在 1 處求值，就是把數列求和；在單位根處求值，就是挑出各個剩餘類（residue classes）；求導則是把各項以 k 加權。這樣一來，許多計數問題都化為代數問題，而這正是第 9 章的核心技巧。同一個濾波器，配合五次單位根應用於 $(1+x)^n$，便會與正五邊形和黃金比（golden ratio）扯上關係；我們會在 Section 6.1 繼續探討單位根的代數。

思維技巧

利用對稱性來篩選

要從一個和之中挑出滿足某個條件的項，可找一個函數，使它在你想要的項上取值 1、在其餘的項上取值 0，而這個函數是由對稱平均構造出來的。

對單位根取平均就是經典的例子。同一個原理——「對一個對稱羣取平均，以分離出你想要的部分」——在代數中隨處可見，由單位根濾波器以至羣表示論（representation theory of groups）皆然。

延伸練習

1 **偶數位置。**利用二次單位根（1 和 -1）的濾波器，證明 $\binom{n}{0}+\binom{n}{2}+\binom{n}{4}+\cdots=2^{n-1}$。

2 **另一個餘數。**求 $\binom{n}{1}+\binom{n}{4}+\binom{n}{7}+\cdots$ 的公式，並驗證三條公式加起來等於 2^n。

3 **每第四個。**利用 i，證明 $\binom{n}{0}+\binom{n}{4}+\binom{n}{8}+\cdots=\frac{1}{4}\left(2^n+2^{n/2+1}\cos\frac{n\pi}{4}\right)$。

4 **整除性。**$\{1,2,\ldots,12\}$ 有多少個子集的元素個數可被 3 整除？有多少個子集的元素之和可被 3 整除？（第二問的提示：考慮 $\prod_{j=1}^{12}(1+x^j)$ 在三次單位根處的值。）

延伸閱讀

- **Herbert Wilf, *generatingfunctionology***（A. K. Peters；可於網上免費閱覽）。一本引人入勝的生成函數入門書；單位根濾波器在書中很早便出現。
- **Arthur Benjamin & Jennifer Quinn, *Proofs That Really Count***（Mathematical Association of America）。以計數方法證明二項式恆等式。
- **T. W. Körner, *Fourier Analysis***（Cambridge University Press）。內容包括離散傅立葉變換和 FFT。

5.3 因式分解 $z^{2n}+1$

題目

STEP III 2013，第 4 題

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

`step.maths.org/questions/13-s3-q4`

引子

$\cos\frac{\pi}{8}\cos\frac{3\pi}{8}\cos\frac{5\pi}{8}\cos\frac{7\pi}{8}$ 等於多少？每個因子都是涉及多重根號的難看數字，但它們的乘積卻恰好是 $\frac{1}{8}$。等間隔角度的餘弦之積，總是這樣化簡成簡潔的數，而這道題正好解釋了箇中原因。竅門在於看出這些角度是某個多項式（polynomial）的根（roots）的輻角（argument），把該多項式因式分解（factorise），然後代入（substitute）一個巧妙選取的值。

在複數（complex numbers）範圍內，每個多項式都可以完全分解成線性因式（linear factors）。在實數範圍內，非實根以共軛對（conjugate pairs）的形式出現，每一對合併成一個實二次因式（real quadratic factor）。這兩項事實在本題中都展現出來，而 $z^{2n}+1$ 明明完全沒有實根，卻可以寫成實二次式之積，靠的正是第二項事實。

動筆之前

想一想 $z^{2n}=-1$ 的解在複數平面（complex plane）上的甚麼位置？試就 $n=2$ 和 $n=4$ 把它們畫出來。

它們的模（modulus）全都是 1，而輻角 θ 滿足 $2n\theta=\pi+2m\pi$，所以 θ 是 $\frac{\pi}{2n}$ 的奇數倍。它們在單位圓（unit circle）上構成一個正 $2n$ 邊形，並且旋轉到沒有任何頂點落在實軸上（圖 5.3.1）。它們兩兩配成共軛對 $e^{\pm i\theta}$，即關於實軸互為鏡像。

逐步解題

階段 1　共軛對給出實二次式

$$(z-e^{\mathrm{i}\theta})(z-e^{-\mathrm{i}\theta}) = z^2-(e^{\mathrm{i}\theta}+e^{-\mathrm{i}\theta})z+1 = z^2-2z\cos\theta+1. \qquad \blacksquare$$

階段 2　因式分解

滿足 $-\pi<\theta\le\pi$ 的 -1 的 $2n$ 次方根是 $e^{\pm\mathrm{i}\theta_k}$，其中

$$\theta_k=\frac{(2k-1)\pi}{2n},\qquad k=1,2,\dots,n.$$

這是 $2n$ 個互不相同的數，全都是 $2n$ 次多項式 $z^{2n}+1$ 的根，所以它們就是該多項式的全部根，於是

$$z^{2n}+1=\prod_{k=1}^{n}(z-e^{\mathrm{i}\theta_k})(z-e^{-\mathrm{i}\theta_k})=\prod_{k=1}^{n}(z^2-2z\cos\theta_k+1). \qquad \blacksquare$$

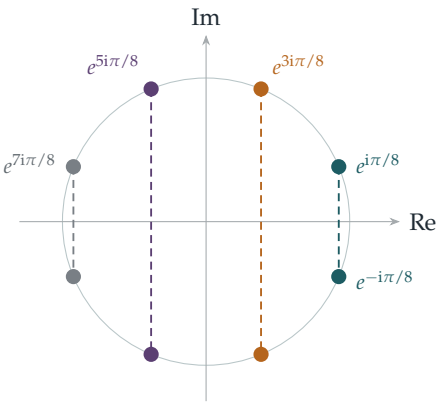


圖 5.3.1　$z^8=-1$ 的八個解（即 $n=4$ 的情況）。它們配成四個共軛對，以虛線相連；每一對給出 z^8+1 的一個實二次因式 $z^2-2z\cos\theta_k+1$。

階段 3　代入 $z=\mathrm{i}$

想一想　$\mathrm{i}^2-2\mathrm{i}\cos\theta+1$ 等於甚麼？當 n 是偶數時，$\mathrm{i}^{2n}+1$ 又等於甚麼？

每個因式變成

$$\mathrm{i}^2-2\mathrm{i}\cos\theta_k+1=-2\mathrm{i}\cos\theta_k,$$

所以右邊是 $(-2\mathrm{i})^n\prod_k\cos\theta_k$。左邊是 $\mathrm{i}^{2n}+1=(-1)^n+1$，當 n 是偶數時等於 2。而對於偶數 n，

$$(-2\mathrm{i})^n=2^n(-\mathrm{i})^n=2^n\big((-\mathrm{i})^2\big)^{n/2}=2^n(-1)^{n/2}.$$

因此

$$\prod_{k=1}^{n}\cos\theta_k=\frac{2}{2^n(-1)^{n/2}}=(-1)^{\frac12 n}2^{1-n}.\qquad\blacksquare$$

當 $n=4$ 時，這正是引子中的乘積：$(-1)^2 2^{-3}=\frac18$。

階段 4　當 n 是奇數時

> **想一想**　對於奇數 n，其中一個角 θ_k 等於 $\frac{\pi}{2}$。是哪一個？為甚麼這會令代入 $z=\mathrm{i}$ 這一步在現有形式下毫無用處？

當 $k=\frac{n+1}{2}$ 時，$\theta_k=\frac{n\pi}{2n}=\frac{\pi}{2}$，所以該因式是 z^2+1，它在 $z=\mathrm{i}$ 時等於零；兩邊都是 0，我們甚麼也得不到。解決辦法是先把這個因式除去，而題目給出的恆等式正是這樣做的。

> **停一停，試一試**　把因式分解式的兩邊除以 $1+z^2$，代入 $z=\mathrm{i}$，然後化簡。你需要把餘下的餘弦兩兩配對。

除以因式 z^2+1（即 $k=\frac{n+1}{2}$ 的那一項），得

$$1-z^2+z^4-\cdots+z^{2n-2}=\prod_{k\neq\frac{n+1}{2}}\left(z^2-2z\cos\theta_k+1\right).$$

當 $z=\mathrm{i}$ 時，左邊每一項都是 $(-z^2)^j=1^j=1$，而共有 n 項，所以左邊等於 n。右邊是 $(-2\mathrm{i})^{n-1}\prod_{k\neq\frac{n+1}{2}}\cos\theta_k$。由於 $n-1$ 是偶數，$(-2\mathrm{i})^{n-1}=2^{n-1}(-1)^{(n-1)/2}$。

現在把各角配對：$\theta_{n+1-k}=\pi-\theta_k$，所以 $\cos\theta_{n+1-k}=-\cos\theta_k$。餘下的 $n-1$ 個因子組成 $\frac{n-1}{2}$ 個這樣的對，其中 $k=1,\dots,\frac{n-1}{2}$，而每一對貢獻 $-\cos^2\theta_k$。因此

$$n=2^{n-1}(-1)^{(n-1)/2}\cdot(-1)^{(n-1)/2}\prod_{k=1}^{(n-1)/2}\cos^2\theta_k=2^{n-1}\prod_{k=1}^{(n-1)/2}\cos^2\theta_k,$$

從而得到 $\prod\cos^2\theta_k=n2^{1-n}$。最後一個因子，即 $k=\frac{n-1}{2}$ 的那個，是 $\cos^2\frac{(n-2)\pi}{2n}$，與題目所寫的一致。$\blacksquare$

以 $n=3$ 驗算：$\cos^2\frac{\pi}{6}=\frac34=3\cdot2^{-2}$。

回顧

- **知道根，就知道多項式。**一個多項式由它的根決定（只差一個常數倍），所以列出 $z^{2n}+1$ 的根，就得到它的因式分解。
- **代入一個巧妙的值。**取 $z=\mathrm{i}$，每個二次因式都變成餘弦的倍數。

- **當代入後得到 $0 = 0$，就先做除法。**對於奇數 n，有一個因式在 $z = \mathrm{i}$ 時等於零，所以我們在代入之前先把它除去。

背後的數學

代數基本定理

$z^{2n} + 1$ 可以在 $\mathbb{C}$ 上完全分解成線性因式，這是代數基本定理（fundamental theorem of algebra）的一個特殊情況：每個複係數的非常數多項式都有一個複根，因而可以完全分解成線性因式。高斯（Gauss）在 1799 年的博士論文中給出了第一個獲廣泛接受的證明，其後又多次重新探討這條定理，提出新的證明。由此可推出，每個實多項式都可以分解成實線性因式和實二次因式，因為非實根以共軛對的形式出現，正如第 1 步所見。這條定理名為「代數」，實質卻屬於分析：每個證明都要用到實數的某種性質，例如連續性（continuity），而你會在複分析（complex analysis）的入門課程中見到其中一個最巧妙的證明。

正弦與餘弦之積

這個方法可以給出一整族恆等式。以 $z^n - 1$ 取代 $z^{2n} + 1$，就得到

$$\prod_{k=1}^{n-1} \sin\frac{k\pi}{n} = \frac{n}{2^{n-1}},$$

它的幾何形式在 Section 5.4 中有證明。這類恆等式會在振動系統的分析中出現（以彈簧相連的 n 個質點的頻率涉及 $\sin\frac{k\pi}{2(n+1)}$），也會在計數問題中出現（某些圖的生成樹（spanning trees）數目，以及長方形的多米諾骨牌鋪砌（domino tilings）數目，都可以用餘弦之積表示）。Kasteleyn 在 1961 年提出的 $m \times n$ 長方形多米諾骨牌鋪砌公式，是一個雙重乘積，其中各項是含 $\cos^2\frac{j\pi}{m+1}$ 和 $\cos^2\frac{k\pi}{n+1}$ 的式子。

思維技巧

從根到乘積

要計算等間隔角度的三角函數值之積，先看出它們來自某個簡單多項式的根，再作因式分解，然後代入一個方便的值。

選取甚麼值，是一門藝術：本題中 $z = \mathrm{i}$ 把每個二次因式變成一個餘弦。其他選擇（例如 $z = 1$ 或 $z = -1$）會給出其他恆等式。

延伸練習

1 **一個七邊形的乘積。**利用第 (ii) 部分，取 $n=7$，證明 $\cos\frac{\pi}{14}\cos\frac{3\pi}{14}\cos\frac{5\pi}{14}=\frac{\sqrt{7}}{8}$。然後證明 $\cos\frac{\pi}{7}\cos\frac{2\pi}{7}\cos\frac{3\pi}{7}=\frac{1}{8}$。（提示：利用 $\cos\theta=\sin(\frac{\pi}{2}-\theta)$ 把各角聯繫起來，或者反覆使用恆等式 $\sin 2x=2\sin x\cos x$。）

2 **另一個代入。**在因式分解式中代入 $z=1$，證明

$$\prod_{k=1}^{n}\sin\frac{(2k-1)\pi}{4n}=2^{\frac{1}{2}-n}.$$

3 **分解 z^6+1。**把 z^6+1 寫成三個實二次式之積，並把它們乘開以驗算答案。

4 **正弦。**以類似方法分解 $z^{2n}-1$，再除以 z^2-1，證明

$$\prod_{k=1}^{n-1}\sin\frac{k\pi}{2n}=\frac{\sqrt{n}}{2^{n-1}}.$$

延伸閱讀

- **Tristan Needham, *Visual Complex Analysis***（Oxford University Press）。第 2 章講述單位根（roots of unity）與多項式，配有精美的圖畫。
- **Benjamin Fine & Gerhard Rosenberger, *The Fundamental Theorem of Algebra***（Springer）。收錄這條定理的多個證明，分別來自分析、代數和拓撲，為大學本科生而寫。
- **Paul Nahin, *An Imaginary Tale***（Princeton University Press）。一部關於 $\sqrt{-1}$ 的歷史，包括單位根及其應用。

5.4 正多邊形中的距離

題目

STEP III 2016，第 7 題

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

`step.maths.org/questions/16-s3-q7`

引子

在半徑為 1 的圓內畫一個正七邊形（regular heptagon）。從其中一個頂點出發，畫出連往其餘各頂點的六條弦（chord），量度它們的長度，再把這六個長度乘起來。答案恰好是 7。若換成正 100 邊形，從一個頂點出發的 99 條弦，長度之積恰好是 100。

這麼簡潔的結果，為甚麼會成立？各條弦的長度本身並不整齊（$2\sin\frac{\pi}{7}$、$2\sin\frac{2\pi}{7}$ 等等）。這道題目說明，這個結果其實是一個簡單代數恆等式在幾何上的投影，這個恆等式就是 $z^n - 1$ 按單位根（roots of unity）所作的因式分解（factorisation）；題目同時說明，平面上的距離就是複數（complex number）的模（modulus）。

動筆之前

想一想 把多邊形的頂點放在 n 次單位根上，即 $X_k = \omega^k$。若 P 是點 z，$|PX_k|$ 用 z 表示是甚麼？那麼所有距離之積又是甚麼？

$|PX_k| = |z - \omega^k|$，而所有距離之積是 $|(z-1)(z-\omega)\cdots(z-\omega^{n-1})|$，按題目第一部分，這等於 $|z^n - 1|$。n 個距離之積變成了單獨一個模。餘下的工夫，全在於巧妙地選取 z。

逐步解題

階段 1　因式分解

每個 ω^k 都滿足 $(\omega^k)^n = (\omega^n)^k = 1$，所以 $1, \omega, \ldots, \omega^{n-1}$ 都是 $z^n - 1$ 的根。它們互不相同（它們的輻角（argument）$\frac{2\pi k}{n}$ 各不相同，且都在 0 與 2π 之間），所以它們就是這個首項係數（leading coefficient）為 1 的 n 次多項式的全部 n 個根，於是 $z^n - 1 = \prod_{k=0}^{n-1}(z - \omega^k)$。 ■

階段 2　垂直平分線上的一點

把圖形旋轉，使得 $X_k = \omega^k$。與 $X_0 = 1$ 和 $X_1 = \omega$ 等距的點，組成 X_0X_1 的垂直平分線（perpendicular bisector），即通過 O 而傾角為 $\frac{\pi}{n}$ 的直線。

想一想　所以 P 就是 $z = te^{\mathrm{i}\pi/n}$，其中 t 是某個實數（可正可負），而 $|OP| = |t|$。z^n 是甚麼？

$z^n = t^n e^{\mathrm{i}\pi} = -t^n$。所以

$$|PX_0| \times \cdots \times |PX_{n-1}| = |z^n - 1| = |-t^n - 1| = |t^n + 1|.$$

若 n 是偶數，$t^n = |t|^n \geq 0$，所以上式等於 $|OP|^n + 1$。 ■

停一停，試一試　n 是奇數時又如何？分開考慮 $t > 0$ 和 $t < 0$ 兩種情況，並說明每種情況在幾何上代表甚麼。

若 n 是奇數，t^n 與 t 同號。

- 若 $t \geq 0$，即 P 與 X_0X_1 的中點位於 O 的同一側，則所求之積是 $|OP|^n + 1$。
- 若 $t < 0$，即 P 位於 O 的另一側（朝向與邊 X_0X_1 相對的頂點），則 $t^n = -|OP|^n$，所求之積是 $\left|1 - |OP|^n\right|$。

當 n 是奇數時，通過 O 而垂直於 X_0X_1 的直線會穿過另一側的一個頂點；在第二種情況下，當 P 就是該頂點（$|OP| = 1$）時，所求之積等於零，這是理所當然的，因為此時其中一個距離是零。

階段 3　從一個頂點出發的弦

想一想　現在 $P = X_0$，其中一個距離 $|X_0X_0|$ 是零，所以 $|z^n - 1| = 0$ 派不上用場。哪一個多項式以 $\omega, \omega^2, \ldots, \omega^{n-1}$ 為根，卻不以 1 為根？

把 $z^n - 1$ 除以 $z - 1$，

$$(z - \omega)(z - \omega^2)\cdots(z - \omega^{n-1}) = \frac{z^n - 1}{z - 1} = 1 + z + z^2 + \cdots + z^{n-1}.$$

右邊是一個多項式，所以這個等式對所有 z 都成立，包括 $z = 1$，此時它的值是 n。在

$z = 1$ 處取模，

$$|X_0X_1| \times |X_0X_2| \times \cdots \times |X_0X_{n-1}| = |1-\omega|\,|1-\omega^2|\cdots|1-\omega^{n-1}| = n. \qquad \blacksquare$$

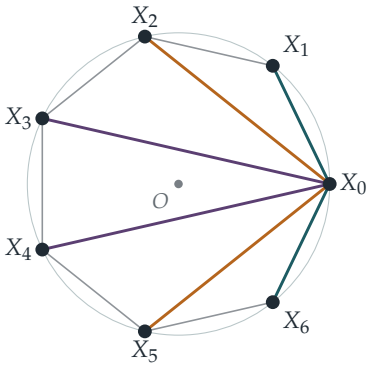


圖 5.4.1　內接於半徑為 1 的圓的正七邊形。從 X_0 出發的六條弦，長度為 $2\sin\frac{k\pi}{7}$，約為 0.868（藍綠色）、1.564（琥珀色）和 1.950（紫紅色），每個長度各出現兩次。它們的積恰好是 7。

回顧

- **距離就是模。**$|PX_k| = |z-\omega^k|$，所以距離之積就是積的模，而這個積有閉合形式（closed form）。
- **利用對稱性選取坐標。**把圖形旋轉，使頂點成為單位根，再利用通過 O 的對稱軸，z^n 便容易計算。
- **先消去零因子，再代入。**一如 Section 5.3，先除去在所關注的點上等於零的因子，代入才行得通。

背後的數學

科茨定理

第 (i) 部分是科茨（Roger Cotes）的一條定理。科茨是劍橋一位才華橫溢的年輕數學家，曾編訂牛頓（Newton）《自然哲學的數學原理》（*Principia*）的第二版，1716 年去世，年僅三十三歲（牛頓曾說：「假如他還在世，我們或許已經學到一點東西了。」）。科茨發現，若 P 位於由 O 出發、通過正 n 邊形一個頂點的射線上，則 P 到各頂點的距離之積是 $\big||OP|^n - 1\big|$；若 P 位於通過某條邊中點的射線上，則這個積是 $|OP|^n + 1$。他沒有用到複數，就憑這個結果把 $x^n \pm 1$ 分解成實二次因式，而這正是他計算有理函數

（rational function）積分時所需要的。這條定理在他去世後，於 1722 年才發表。

正弦之積

由於 $|1-\omega^k| = |e^{-\mathrm{i}\pi k/n} - e^{\mathrm{i}\pi k/n}| = 2\sin\frac{k\pi}{n}$，第 (ii) 部分說的就是

$$\prod_{k=1}^{n-1} \sin\frac{k\pi}{n} = \frac{n}{2^{n-1}}.$$

這與 Section 5.3 中的餘弦之積互相配對。另一個相關的積給出 z^n-1 的判別式（discriminant）：所有成對的相異 n 次單位根之間距離的平方，乘起來等於 n^n。判別式量度一個多項式的根彼此相隔多遠，在多項式和數域（number field）的代數理論中佔有核心地位，我們會在第 6 章討論。

靜電學與等距分佈

這個結果還有一個物理上的解讀。把 n 個相等的電荷放在一個圓上；令勢能（potential energy）最小的排列（採用二維靜電學的對數勢）就是正多邊形，而能量所涉及的正是 $\sum \ln|X_jX_k|$，即距離之積的對數。「分佈得盡量疏散」的點，在數值分析（numerical analysis）中十分重要（見 Section 3.6，當中最佳的插值點是由正多邊形經 $x=\cos\theta$ 得來的），在隨機矩陣（random matrix）理論中亦然。

思維技巧

用複數把幾何化為代數

平面上的點是複數，距離是模，正多邊形是單位根。距離之積就變成多項式的模。

經過這樣的轉換，像 z^n-1 的因式分解這類代數恆等式，便可以用來證明一些很難直接看出的幾何事實。

延伸練習

1. **檢查小的情況。**就正方形（$n=4$）和等邊三角形（$n=3$）直接驗證第 (ii) 部分。
2. **通過頂點。**證明：若 P 位於由 O 出發、通過 X_0 的射線上，則 $|PX_0|\times\cdots\times|PX_{n-1}| = \big||OP|^n - 1\big|$。

3 **平均距離。**利用第 (ii) 部分和 Section 3.2 的算術平均—幾何平均不等式（AM–GM inequality），證明從 X_0 出發的 $n-1$ 條弦，平均長度至少是 $n^{1/(n-1)}$。就 $n=6$ 求出實際的平均值 $\frac{1}{n-1}\sum_{k} 2\sin\frac{k\pi}{n}$，並加以比較。

4 **所有的弦。**較難證明：內接於單位圓的正 n 邊形，全部 $\binom{n}{2}$ 條弦的長度之積是 $n^{n/2}$。

延伸閱讀

- **Tristan Needham, *Visual Complex Analysis***（Oxford University Press）。用複數處理幾何，插圖精美。
- **Ross Honsberger, *Mathematical Gems***（Mathematical Association of America）。收錄了好幾道關於正多邊形的精巧問題。
- **Ronald Gowing, *Roger Cotes: Natural Philosopher***（Cambridge University Press）。科茨的傳記，並介紹他的數學工作。

5.5 名不副實的有理角

題目

STEP III 2022，第 8 題

繼續閱讀之前，請先看題目（原題為英文）。題目載於 OCR 提供下載的 2022 年 STEP 試卷（zip 檔案）內的 STEP III 試卷：

`www.ocr.org.uk/Images/696587-2022-past-papers.zip`

引子

中學幾何裏遇到的角，度數通常是整數：30°、45°、60°。它們的餘弦是 $\frac{\sqrt{3}}{2}$、$\frac{1}{2}$ 這類數。反過來又如何？如果餘弦是個簡單的分數，例如 $\frac{1}{3}$，那麼這個角的度數會不會也是個漂亮的數？

$\cos^{-1}\frac{1}{3} \approx 70.528779\ldots^\circ$ 這個角在科學中隨處可見。它是正四面體（regular tetrahedron）兩個面之間的夾角，而它的補角 109.47° 則是甲烷分子中兩條化學鍵之間的夾角。本題要證明這個角的度數不是有理數（rational number）：不是 $70\frac{1}{2}$，不是 $\frac{1411}{20}$，也不是任何分數。它的小數展開永不循環。這個看似冷僻的事實，解決了大衛·希爾伯特（David Hilbert）在 1900 年提出的著名問題之一。

動筆之前

想一想 假設 θ 是有理數，設 $\theta = \frac{p}{q}$ 度。試找出一個正整數 k，使得 $\sin(k\theta) = 0$。

取 $k = 180q$：這時 $k\theta = 180p$ 度，是半圈的整數倍，所以 $\sin(k\theta) = 0$。因此，度數為有理數的角，總有某個倍數會落在 180° 的倍數上。要證明 θ 是無理數（irrational number），只需證明*沒有*任何倍數 $k\theta$ 的正弦為零。第 (ii) 部分的做法是：假如這樣的 k 存在，就取其中最小的一個，然後推出矛盾（contradiction）。這又是 Section 1.3 的良序原理（well-ordering principle）。

逐步解題

階段 1　倍角公式

想一想　寫出 $\cos k\theta + \mathrm{i}\sin k\theta = (\cos\theta + \mathrm{i}\sin\theta)^k$，並提出因子 $\cos^k\theta$。括號內剩下甚麼？

$(\cos\theta + \mathrm{i}\sin\theta)^k = \cos^k\theta\,(1 + \mathrm{i}\tan\theta)^k = \cos^k\theta \sum_{j=0}^{k} \binom{k}{j} \mathrm{i}^j \tan^j\theta$。虛部（imaginary part）來自奇數 $j = 2m+1$ 的各項，這時 $\mathrm{i}^j = \mathrm{i}(-1)^m$：

$$\begin{aligned}\sin k\theta &= \cos^k\theta \sum_{m} (-1)^m \binom{k}{2m+1} \tan^{2m+1}\theta \\ &= \sin\theta\cos^{k-1}\theta \sum_{m} (-1)^m \binom{k}{2m+1} \tan^{2m}\theta.\end{aligned}$$

由於 $\tan^2\theta = \sec^2\theta - 1$，這就是所求的公式。■實部（real part）來自偶數 $j = 2m$ 的各項，由此得

$$\cos k\theta = \cos^k\theta \left(1 - \binom{k}{2}(\sec^2\theta - 1) + \binom{k}{4}(\sec^2\theta - 1)^2 - \cdots\right).$$

階段 2　奇偶性

現在 $\sec\theta = a$ 是奇整數，所以 $\sec^2\theta - 1 = a^2 - 1$ 是偶整數。而且 θ 嚴格介乎 0° 與 90° 之間，所以 $\sin\theta \neq 0$ 且 $\cos\theta \neq 0$。

想一想　如果 $\sin(k\theta) = 0$，括號 $k - \binom{k}{3}(a^2-1) + \cdots$ 必須等於零。第一項之後的每一項都是 $a^2 - 1$ 的倍數。由此可知 k 有甚麼性質？

括號等於 k 加上若干個偶數 $a^2 - 1$ 的倍數之和。假如 k 是奇數，括號就是奇數，而奇整數不可能等於 0。所以 k 必定是偶數。

於是 $\sin(k\theta) = 2\sin(\tfrac{1}{2}k\theta)\cos(\tfrac{1}{2}k\theta) = 0$。但 $\tfrac{1}{2}k$ 是比 k 小的正整數，根據 k 的最小性，$\sin(\tfrac{1}{2}k\theta) \neq 0$。因此 $\cos(\tfrac{1}{2}k\theta) = 0$。 ■

階段 3　矛盾

停一停，試一試　取 $m = \tfrac{1}{2}k$，套用餘弦公式。為甚麼 $\cos(m\theta)$ 永遠不可能是零？

根據第 1 步，$\cos(m\theta) = a^{-m}\left(1 - \binom{m}{2}(a^2-1) + \binom{m}{4}(a^2-1)^2 - \cdots\right)$。括號等於 1 加上若干個偶數 $a^2 - 1$ 的倍數，所以它是奇數，自然不等於零。因此 $\cos(m\theta) \neq 0$，與第 2

步矛盾。

所以不存在使 $\sin(k\theta) = 0$ 的正整數 k。（假如有這樣的正整數，其中必有最小的一個，而我們剛剛證明了這個最小的不可能存在。）根據動筆之前的觀察，θ 的度數不可能是有理數。 ■

階段 4 餘切版本

? **想一想** 如果 $\cot\phi = \frac{1}{b}$，$\sec^2\phi - 1$ 是多少？它是偶數嗎？

$\tan\phi = b$，所以 $\sec^2\phi - 1 = \tan^2\phi = b^2$；由於 b 是偶數，這個數是偶數（其實更是 4 的倍數）。此外 $0° < \phi < 90°$，所以 $\sin\phi$ 和 $\cos\phi$ 都不等於零。完全相同的論證現在同樣適用：如果 k 是使 $\sin(k\phi) = 0$ 的最小正整數，括號 $k - \binom{k}{3}b^2 + \cdots$ 必須等於零，於是 k 必為偶數；然後 $\cos(\frac{1}{2}k\phi) = 0$；但餘弦的括號 $1 - \binom{m}{2}b^2 + \cdots$ 是奇數。這就產生矛盾，所以 ϕ 是無理數。 ■

例如，$\cot^{-1}\frac{1}{2} = \tan^{-1}2 \approx 63.43°$ 是無理數。

回顧

- **有理角的倍數具週期性。**這樣重新表述之後，無理性就變成一個關於所有 k 的 $\sin(k\theta)$ 的命題。
- **奇偶性論證。**以模 2 計算，可知某些整數是奇數，因而不等於零。方法簡單，卻一錘定音。
- **最小反例。**假設有最小的 k，再證明必定有更小的一個（這裏是 $\frac{1}{2}k$），正是第 1 章的良序論證。

背後的數學

無理旋轉

如果 θ 的度數是有理數，不斷重複旋轉 θ，最終會回到起點。如果 θ 是無理數，就永遠不會回到起點，而且會出現奇妙的現象：各點 $k\theta$（以模 360° 計算）會均勻地散佈在整個圓周上，與每一個角都可以任意接近。這就是等分佈定理（equidistribution theorem），由赫爾曼．外爾（Hermann Weyl）、瓦茨瓦夫．謝爾賓斯基（Wacław Sierpiński）和 Piers Bohl 在 1910 年前後各自獨立證明。圖 5.5.1 顯示 $\cos^{-1}\frac{1}{3}$ 的首 40 個倍數：它們已經逐漸佈滿圓周，而且永遠不會重複。

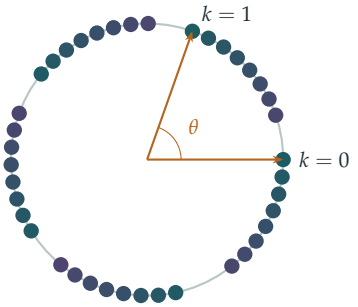


圖 5.5.1　位於角 $k\theta$ 的各點，其中 $k = 0, 1, \ldots, 39$，而 $\theta = \cos^{-1}\frac{1}{3} \approx 70.53°$（顏色隨 k 增加由藍綠色漸變為紫紅色）。由於 θ 是無理數，沒有任何一點會重複出現，而各點均勻地散佈在整個圓周上。

希爾伯特第三問題

1900 年，大衛．希爾伯特為新世紀列出了二十三個問題。第三個問題是：如果兩個多面體（polyhedron）體積相同，是否總可以把其中一個切成有限多塊多面體，再重新拼合成另一個？（對平面上的多邊形而言，答案是肯定的：這就是 Wallace–Bolyai–Gerwien 定理。）不出一年，希爾伯特的學生馬克斯．德恩（Max Dehn）證明了答案是否定的：正四面體不能切開再重新拼合成體積相同的立方體。德恩的證明為每個多面體賦予一個由其棱長和二面角（dihedral angle）構成的量，現稱為德恩不變量（Dehn invariant），它在切割和重新拼合之下保持不變；立方體的德恩不變量是零，而正四面體的德恩不變量不等於零，原因正是二面角 $\cos^{-1}\frac{1}{3}$ 不是 $180°$ 的有理數倍。這恰恰就是本題所證明的事實。

尼雲定理

這個結論可以推到多遠？1956 年，伊凡．尼雲（Ivan Niven）證明：如果 θ 的度數是有理數，而且 $\cos\theta$ 是有理數，那麼 $\cos\theta$ 只能是 0、$\pm\frac{1}{2}$、± 1 其中之一。所以餘弦為有理數的「漂亮」角，就只有 $0°$、$60°$、$90°$、$120°$、$180°$ 以及與它們相關的角。證明用到這個事實：$2\cos\theta$ 是某個整數係數首一多項式（monic polynomial）的根，即代數整數（algebraic integer），也就是 Section 1.3 末尾提到的概念。

思維技巧

以奇偶性證明不等於零

要證明一個整數算式永不爲零，可證明它恆爲奇數：把它以模 2 化簡。

這個簡單的想法與最小反例（minimal counterexample）結合，就證明了一個深刻的事實。以一個小的數為模作化簡，從而排除某個方程成立的可能，是數論中最常用的手法之一，Section 1.5 便是一例。

延伸練習

1 **正四面體。**證明正四面體兩個面之間的夾角是 $\cos^{-1}\frac{1}{3}$，並證明甲烷分子中兩條化學鍵（由正四面體的中心連到兩個頂點）之間的夾角是 $\cos^{-1}\left(-\frac{1}{3}\right)$。

2 **3–4–5 三角形。**證明 3–4–5 三角形最小的角的度數是無理數。（提示：這個角是 $\cot^{-1}\frac{4}{3}$；可改寫上述論證，或考慮高斯整數（Gaussian integer）$4+3\mathrm{i}$ 的各次冪，並證明它們全都不是實數。）

3 **為甚麼 a 要是奇數？**當 $a=2$ 時，論證在哪裏出了問題？$\cos^{-1}\frac{1}{2}$ 是有理數嗎？

4 **尼雲定理的簡單情況。**找出介乎 0° 與 90° 之間、度數為整數而 $\cos\theta$ 為有理數的所有角 θ，並驗證你所列的結果與尼雲定理吻合。

延伸閱讀

- **Ivan Niven, *Irrational Numbers***（Mathematical Association of America）。書中載有關於三角函數有理數值的尼雲定理。
- **Martin Aigner & Günter Ziegler, *Proofs from THE BOOK***，論希爾伯特第三問題的一章。對德恩的解法有精彩的闡述。
- **Benjamin Yandell, *The Honors Class: Hilbert's Problems and Their Solvers***（A. K. Peters）。講述全部二十三個問題背後的故事。

5.6 費馬點

題目

STEP III 2000，第 3 題

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

`step.maths.org/questions/00-s3-q3`

引子

三條村莊想興建一個把它們全部連接起來的道路網絡，而且路程愈短愈好。交匯處應該設在哪裏？大約在 1640 年，費馬（Pierre de Fermat）提出了這樣一個問題：在三角形內找出使 $PA + PB + PC$ 最小的點 P。托里拆利（Evangelista Torricelli）解決了這個問題：從這一點望向三條邊，所見的角都是 120°。

本題利用複數（complex numbers）來證明這一點的一個漂亮作圖法。在三角形的每條邊上向外作一個等邊三角形（equilateral triangle），再把每個新頂點與原三角形中相對的頂點連起來。這三條直線相交於同一點，即費馬點（Fermat point），而且三條線段的長度都相同，這個長度正是道路總長的最小值。有了複數，再加上一條關於 $e^{\mathrm{i}\pi/3}$ 的恆等式，整件事只需幾行代數運算。

動筆之前

想一想 A、B 和 C 在哪裏？它們的輻角（argument）是多少？OA、OB 和 OC 彼此之間的夾角是多少？

$A = p$ 位於正實軸上（輻角為 0°）。$B = q\alpha^2 = qe^{2\pi\mathrm{i}/3}$ 的輻角是 120°。$C = -r\alpha = re^{\mathrm{i}\pi}e^{\mathrm{i}\pi/3}$ 的輻角是 240°。因此原點是三角形內的一點，從這一點望向三個頂點，方向兩兩相隔 120°，距離分別是 p、q 和 r。按照這個設定，原點就是費馬點；本題要證明的，是等邊三角形作圖法確實能找到它。

逐步解題

階段 1　關鍵恆等式

$\alpha^3 = e^{i\pi} = -1$，所以 α 是 $z^3+1 = (z+1)(z^2-z+1)$ 的一個根。由於 $\alpha \neq -1$，故 $\alpha^2 - \alpha + 1 = 0$，即 $1 + \alpha^2 = \alpha$。 ■

有兩個推論稍後會用得着：$1-\alpha = -\alpha^2$，以及 $\alpha - \alpha^2 = 1$。此外，$1-\alpha = \frac{1}{2} - \frac{\sqrt{3}}{2}\mathrm{i} = e^{-i\pi/3}$，即 α 的共軛複數（complex conjugate）。

階段 2　把等邊三角形看成旋轉

> **想一想**　三角形 CAN 是等邊三角形，頂點按順時針方向標記。假如你站在 C，怎樣從 A 走到 N？試以 $A - C$ 表示 $N - C$。

按順時針方向依次經過 C、A、N，可知點 N 是把 A 繞 C 順時針旋轉（rotation）60° 而得。乘以 $e^{-i\pi/3} = 1-\alpha$ 相當於順時針旋轉 60°，所以

$$\begin{aligned} N &= C + (1-\alpha)(A-C) = -r\alpha + (1-\alpha)(p + r\alpha) \\ &= (1-\alpha)p + r(\alpha - \alpha^2 - \alpha) = (1-\alpha)p - \alpha^2 r. \end{aligned}$$ ■

（順時針而非逆時針旋轉，會使 N 落在 CA 的另一側，與原點分處兩邊，即在三角形之外。）

> **停一停，試一試**　用同樣的方法求出 L 和 M，並利用第 1 階段的恆等式化簡。

同理，$L = A + (1-\alpha)(B-A)$ 及 $M = B + (1-\alpha)(C-B)$：

$$\begin{aligned} L &= p + (1-\alpha)(q\alpha^2 - p) = \alpha p + (\alpha^2 - \alpha^3)q = \alpha p + (\alpha^2+1)q = \alpha(p+q), \\ M &= q\alpha^2 + (1-\alpha)(-r\alpha - q\alpha^2) = q\alpha^3 + r(\alpha^2 - \alpha) = -q - r. \end{aligned}$$

再利用 $1-\alpha = -\alpha^2$，可得 $N = -\alpha^2 p - \alpha^2 r = -\alpha^2(p+r)$。

階段 3　三條通過原點的直線

> **想一想**　比較 $L = \alpha(p+q)$ 與 $C = -r\alpha$。你留意到甚麼？對 M 與 A，以及 N 與 B，也作同樣的比較。

L 是 α 的正數倍，而 C 是 α 的負數倍。所以 L 和 C 分別位於 O 的兩側，並且在同一條通過 O 的直線上；線段 LC 因而通過原點，長度為 $|L-C| = |\alpha|(p+q+r) = p+q+r$。

同樣地，$M = -(q+r)$ 和 $A = p$ 都在實軸上，分別位於 O 的兩側，所以 MA 通過 O，

長度為 $p+q+r$。而 $N=-\alpha^2(p+r)$ 和 $B=q\alpha^2$ 都在通過 O、方向為 α^2 的直線上，並且分處兩側，所以 NB 通過 O，長度為 $p+q+r$。 ■

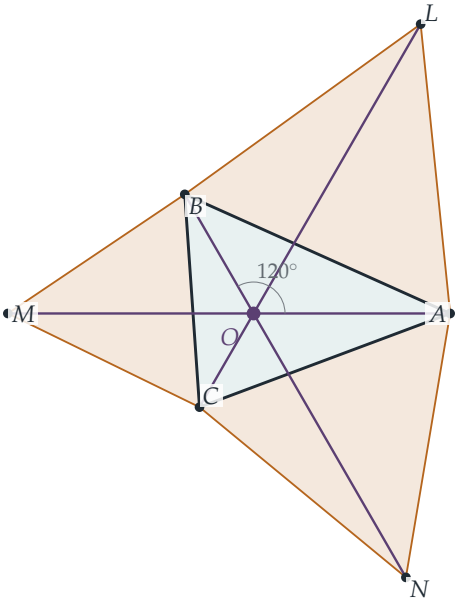


圖 5.6.1　費馬點，其中 $p=2$，$q=1.4$，$r=1.1$。在 ABC 的各邊上向外作等邊三角形（琥珀色）。直線 LC、MA 和 NB（紫紅色）全部通過原點 O，而 ABC 的各邊在該點所對的角均為 120°；每條線段的長度都是 $p+q+r=4.5$。

回顧

- **旋轉即乘法。**作等邊三角形，就是把一個頂點繞另一個頂點旋轉 60°，也就是乘以 $e^{\pm i\pi/3}$。
- **巧選坐標。**把費馬點放在原點，並讓各頂點沿 1、α^2、$-\alpha$ 這三個方向排列，每條算式都隨即大幅簡化。
- **一條恆等式，反覆使用。**每一步化簡都來自 $\alpha^2-\alpha+1=0$ 和 $\alpha^3=-1$。

背後的數學

為甚麼費馬點使距離最小

為甚麼 $PA + PB + PC$ 會在各邊所對的角均為 120° 的那一點取得最小值？以下是一個精彩的論證。把整幅圖繞 B 旋轉 60°，把 A 送到 L'（即 AB 上等邊三角形的頂點），把 P 送到 P'。三角形 BPP' 是等邊三角形，所以 $PB = PP'$，而且 $PA = P'L'$。於是 $PA + PB + PC = L'P' + P'P + PC$，這是一條由 L' 到 C 的路徑的長度。它至少等於 $|L'C|$，而等號恰好在路徑成一直線時成立，這種情況出現於 P 處的各個角均為 120° 之時。所以總長度的最小值是 $|LC| = p + q + r$，與題目所得的結果一致。（這個論證成立的前提是三角形的每個角都小於 120°；否則最小值會在鈍角的頂點取得。）

肥皂膜與斯坦納樹

把一個裝有三根小釘（代表 A、B、C）的框架浸入肥皂水中：肥皂膜會令其表面張力能量減至最小，形成三面薄壁，在費馬點以 120° 相交。若有更多的點，最短的連接網絡是一棵斯坦納樹（Steiner tree），其中有額外的交匯點，每個交匯點都有三條邊以 120° 相交。一般而言，求斯坦納樹在計算上十分困難（這是一個 NP 困難（NP-hard）問題），但它們在實際應用中相當重要，例如設計道路網絡、管道網絡以及電腦晶片上的線路。

拿破崙定理

同一個圖形中還藏着另一件瑰寶：在任意三角形的三邊上所作的三個等邊三角形，它們的中心構成一個等邊三角形。傳統上這個結果歸功於拿破崙（Napoleon Bonaparte）（他是一位熱衷數學的業餘愛好者），不過並無證據顯示這是他發現的。這個定理同樣可以用複數簡短地證明，用到的正是 $1 + \alpha^2 = \alpha$。

思維技巧

選擇切合問題的坐標

把最重要的點放在原點，並按照問題的對稱性來安排各個方向。選得好，代數運算幾乎可以變得輕而易舉。

在這裏，原點就是費馬點，而三個頂點分別位於 -1 的三個立方根的方向上（或相差一個正負號）。若改用其他坐標，同樣的計算會繁複得多。

延伸練習

1 **檢查字母次序。**證明若各等邊三角形的頂點按逆時針方向標記（即向內作），則三條直線一般不會通過 O。

2 **拿破崙定理。**證明等邊三角形 ABL 的中心是 $\frac{1}{3}(A+B+L)$。利用上文的表達式，證明三個中心構成一個等邊三角形。

3 **最小值。**取 $p=2$，$q=1.4$，$r=1.1$，分別就 P 位於每個頂點以及位於形心（centroid）$\frac{1}{3}(A+B+C)$ 的情況計算 $PA+PB+PC$，並驗證所得的值全部大於 4.5。

4 **等邊三角形的判別法。**證明 z_1、z_2、z_3 構成等邊三角形（不論方向），當且僅當 $z_1^2+z_2^2+z_3^2=z_1z_2+z_2z_3+z_3z_1$。

延伸閱讀

- **Richard Courant & Herbert Robbins, *What Is Mathematics?***（Oxford University Press）。第 7 章討論費馬點、斯坦納樹和肥皂膜。
- **Liang-shin Hahn, *Complex Numbers and Geometry***（Mathematical Association of America）。以複數證明多個幾何定理，程度正好與本節相若。
- **Tristan Needham, *Visual Complex Analysis***（Oxford University Press）。第 1 章包括拿破崙定理及其他幾何結果。

5.7 作為映射的複指數函數

題目

STEP II 1997，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/97-s2-q5`

引子

你所認識的指數函數（exponential function）是一條曲線 $y = e^x$。但複變數的函數卻無法畫成曲線：它的輸入是平面上的一點，輸出也是平面上的一點。要想像這樣的函數，應把它看成一個映射（map），也就是一個變換（transformation）：它把一個平面上的每一點搬到另一個平面上的某一點，沿途把圖形拉伸、彎曲。

題目中的公式是 $z = e^u(\cos v + \mathrm{i}\sin v) = e^{u+\mathrm{i}v} = e^w$。作為映射來看，複指數函數（complex exponential）做了一件奇妙的事：它把平面上的一條水平帶狀區域捲成整個平面（除去原點），把鉛垂線繞成圓，把水平線鋪成射線。又因為它是繞圈捲起來的，所以它不是一對一（one-to-one）的：每一點都有無窮多個點映到它那裏。這正是複對數（complex logarithm）「多值」（multi-valued）的原因，也是複分析（complex analysis）最早帶來的驚奇之一。

動筆之前

想一想 $z = e^u(\cos v + \mathrm{i}\sin v)$ 的模（modulus）和輻角（argument）用 u 和 v 表示是甚麼？

$|z| = e^u$，而 $\arg z = v$（可相差 2π 的倍數）。所以 w 的實部 u 以指數方式控制與原點的距離，而虛部 v 則控制方向。單憑這一點觀察，題目幾乎每一部分都能解答。

逐步解題

階段 1 **解** $e^w = \mathrm{i}e$

> **想一想** $\mathrm{i}e$ 的模和輻角是甚麼？

$|\mathrm{i}e| = e$，而 $\arg(\mathrm{i}e) = \frac{\pi}{2}$。所以我們需要 $e^u = e$，得 $u = 1$，並且 $v = \frac{\pi}{2} + 2k\pi$，其中 k 為任意整數。方程的解為

$$w = 1 + \mathrm{i}\left(\tfrac{\pi}{2} + 2k\pi\right), \qquad k \in \mathbb{Z}.$$

解有無窮多個，沿鉛垂方向每隔 2π 有一個。

階段 2 **格線的像**

在鉛垂線 $u = c$ 上，$|z| = e^c$ 固定不變，而 v 則在變動：它的像（image）是以原點為圓心、半徑為 e^c 的圓，當 v 取遍所有實數時，這個圓被繞了無窮多圈。

在水平線 $v = d$ 上，$\arg z = d$ 固定不變，而 $|z| = e^u$ 取遍所有正值：它的像是由原點出發、方向角為 d 的開射線（半直線），不包括原點本身。直線 $v = d$ 與 $v = d + 2\pi$ 給出同一條射線。

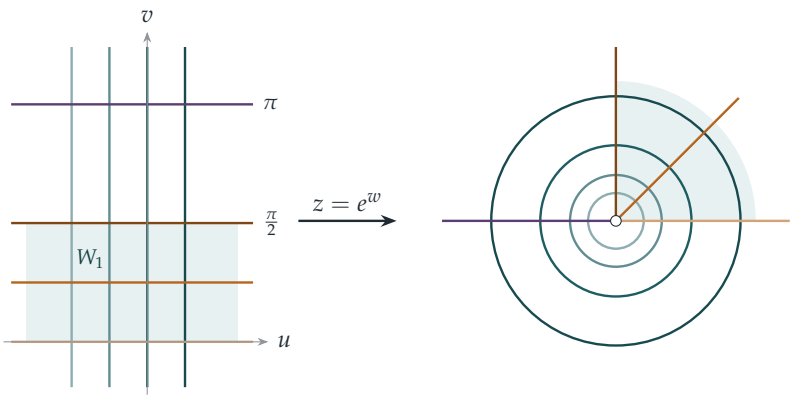


圖 5.7.1　把複指數函數看成映射。w 平面上的鉛垂線 $u = c$（藍綠色）映到半徑為 e^c 的圓；水平線 $v = d$（琥珀色、紫紅色）映到方向角為 d 的射線。陰影帶狀區域 $W_1 = \{0 < v < \frac{\pi}{2}\}$ 恰好映滿第一象限。留意兩幅圖中的格線都相交成直角：這個映射保持角度不變。原點（空心圓圈）永遠不會被映到。

階段 3 **一對一的部分**

想一想 哪些點 w 會映入第一象限 $x>0$、$y>0$？為甚麼這個集合太大，不能與該象限構成一一對應（one-to-one correspondence）？

z 位於開的第一象限，當且僅當 $\arg z$ 嚴格介乎 0 與 $\frac{\pi}{2}$ 之間（模 2π），而模可取任意值。所以 w 必須滿足 $0<v<\frac{\pi}{2}$ 或 $2\pi<v<\frac{5\pi}{2}$ 或 ...：共有無窮多條水平帶狀區域，每一條都映滿整個象限。任取其中一條，它都與該象限一一對應，因為在一條帶狀區域之內，u 和 v 由 $|z|$ 和 $\arg z$ 唯一確定。所以我們可以取

$$W_1=\{u+\mathrm{i}v:0<v<\tfrac{\pi}{2}\},\qquad W_2=\{u+\mathrm{i}v:2\pi<v<\tfrac{5\pi}{2}\}.$$

停一停，試一試 現在試在 w 平面上找出兩個集合，使它們各自與去心圓盤（punctured disc）$0<|z|<1$ 一一對應。

$|z|<1$ 即 $e^u<1$，亦即 $u<0$；而每一個輻角都必須恰好被覆蓋一次，所以 v 的取值範圍應是一個長度恰為 2π 的區間，包括其中一個端點而不包括另一個。例如

$$W_3=\{u+\mathrm{i}v:u<0,\ 0\le v<2\pi\},\qquad W_4=\{u+\mathrm{i}v:u<0,\ -\pi<v\le\pi\}.$$

這些都是半帶狀區域，一條邊包括在內，另一條邊不包括在內。（若兩條邊都包括在內，正實軸上的點便會被映到兩次；若兩條邊都不包括，這些點便會被遺漏。）■

回顧

- **模和輻角把變數分開。**$|e^w|=e^{\operatorname{Re}w}$，而 $\arg e^w=\operatorname{Im}w$，所以 w 的實部和虛部分別獨立地控制距離和方向。
- **把複函數想像成映射。**追蹤格線被映到哪裏，便能清楚看見一個無法畫成圖像的函數的面貌。
- **周期性破壞單射性（injectivity）。**$e^{w+2\pi\mathrm{i}}=e^w$，所以這個映射是多對一的，必須限制在一條高度為 2π 的帶狀區域上，才有反函數。

背後的數學

複對數

在實數線上，ln 是 exp 的反函數。在複平面上，對每一個 $z\neq 0$，方程 $e^w=z$ 都有無窮多個解 w，彼此相差 $2\pi\mathrm{i}$ 的倍數，所以「$\log z$」並不是單獨一個數。要得到一個真正的函數，就必須選定一個分支（branch），也就是選定一條像 W_3 或 W_4 那樣的帶狀

區域：*主分支*（principal branch）取 $-\pi < \arg z \le \pi$。每一個分支都必然在某處有跳躍，跳躍發生在由原點出發的一條射線上，這條射線稱為*分支割線*（branch cut）。黎曼（Bernhard Riemann）在 1851 年的博士論文中提出了解決辦法：不再堅持對數函數要定義在一個平面上。它其實定義在一道無窮的螺旋樓梯上，即一個*黎曼曲面*（Riemann surface）；在這個曲面上，對數函數是單值而且光滑的。時至今日，黎曼曲面在幾何、數論和弦論中都佔有核心地位。

角度保持不變

再看一次圖 5.7.1。鉛垂和水平的格線相交成直角，它們變成的圓和射線也是如此。這並非巧合：凡是複函數的導數不為零之處，該函數都會保持曲線之間的夾角不變。這樣的映射稱為*保角*（conformal）映射。在局部範圍內，保角映射不過是一次旋轉加上一次縮放，而這正是乘以複數 $f'(z)$ 所起的作用。保角映射可用來解決流體流動、靜電學和熱傳導的問題，方法是把一個複雜的區域變換成一個答案已知的簡單區域。

麥卡托地圖

這裏有一個意想不到的發現，把本節與 Section 4.8 連繫起來。麥卡托（Gerardus Mercator）在 1569 年繪製的地圖，設計的目的是令羅盤方位恆定的航線成為一條直線，這就要求地圖保持角度不變。如果你先把地球以球極投影（stereographic projection）的方式（從北極出發）投影到一個平面上，然後取複對數，所得的結果恰好就是麥卡托投影：經線變成鉛垂線，緯線變成水平線，正如射線和圓在 log 之下變成水平線和鉛垂線一樣。把本節的指數映射倒過來運行，就是航海圖背後的數學。

思維技巧

把函數想像成變換

當一個函數無法畫成圖像時，就把它所做的事畫出來：取一組由簡單曲線組成的格網，看看它們被映到哪裏。

對於複函數，水平線和鉛垂線的像，或者圓和射線的像，往往已能說明一切。這種着眼於映射如何改變圖形的思考方式，貫穿整個幾何與分析。

延伸練習

1 **解方程。**求所有滿足 $e^w = -1$ 的複數 w，以及所有滿足 $e^w = 1 + \mathrm{i}$ 的 w。

2 **正方形的像。**求頂點為 0、1、$1+\mathrm{i}$ 和 i 的正方形在 $z = e^w$ 之下的像，並畫出草圖。

3 **直角。**直接證明圓 $|z| = e^c$ 與射線 $\arg z = d$ 相交成直角，並解釋為甚麼這與直線 $u = c$ 和 $v = d$ 互相垂直是一致的。

4 **平方。**設 c 和 d 為非零實數。求直線 $x = c$ 和 $y = d$ 在映射 $z \mapsto z^2$ 之下的像。證明它們都是拋物線，並且凡相交之處都相交成直角。（若 $c = 0$ 或 $d = 0$，情況會怎樣？）

延伸閱讀

- **Tristan Needham, *Visual Complex Analysis***（Oxford University Press）。第 2 章和第 4 章以精彩的插圖講解指數函數、對數函數和保角映射。
- **Hilary Priestley, *Introduction to Complex Analysis***（Oxford University Press）。一本清晰易懂的大學本科入門教材。
- **Ian Stewart & David Tall, *Complex Analysis***（Cambridge University Press）。另一本出色的本科入門書，着重幾何觀點。

5.8 圓盤中隱藏的距離

題目

STEP III 2013，第 6 題

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

`step.maths.org/questions/13-s3-q6`

引子

乍看之下，這只是一道運算模（modulus）和共軛（conjugate）的練習，最後得出一條模樣古怪的不等式。為甚麼會有人關心 $\frac{|z-w|}{|1-zw^*|}$ 呢？

因為它是一個距離。它不是 z 與 w 之間的普通距離，而是另一種幾何中的距離：在這種幾何裏，單位圓盤（unit disc）本身就是一個無限大的完整宇宙，直線是圓弧，三角形的內角和小於 180°。這就是*雙曲幾何*（*hyperbolic geometry*）。兩千年來，人們一直試圖證明歐幾里得（Euclid）的平行公設（parallel postulate），到了十九世紀才發現這種幾何；而本題的不等式，原來正是它的三角不等式（triangle inequality），只是喬裝了一番。更令人意外的是，不等式的右邊正是狹義相對論（special relativity）中的速度相加公式，我們在 Section 3.8 已經見過。

動筆之前

? **想一想** 怎樣從圖形看出 $|z-w| \le |z|+|w|$？（把 z 和 $-w$ 想成向量。）

$z-w = z+(-w)$。先把 z 畫成由原點出發的箭嘴，再把 $-w$ 畫成由 z 的尖端出發的箭嘴。由原點指向終點的箭嘴就是 $z-w$，三支箭嘴組成一個三角形，邊長分別為 $|z|$、$|w|$ 和 $|z-w|$。三角形任何一邊的長度都不超過另外兩邊之和。

另外要記得，對任何複數 a 都有 $|a|^2 = aa^*$。這樣就可以把題目中每一個模都化為代數運算。

逐步解題

階段 1　展開各個模

想一想　展開 $|z-w|^2=(z-w)(z^*-w^*)$，再與 $(|z|+|w|)^2$ 比較。

$|z-w|^2=zz^*-zw^*-z^*w+ww^*=|z|^2+|w|^2-zw^*-z^*w$。另一方面，$(|z|+|w|)^2-E=|z|^2+2|z||w|+|w|^2-zw^*-z^*w-2|zw|$，而由於 $|z||w|=|zw|$，兩者相等。 ■

$E=(zw^*)+(zw^*)^*+2|zw|=2\,\mathrm{Re}(zw^*)+2|zw|$，這是實數。而且根據三角不等式，$E=(|z|+|w|)^2-|z-w|^2\geq 0$。 ■

階段 2　第二條恆等式

$|1-zw^*|^2=(1-zw^*)(1-z^*w)=1-zw^*-z^*w+|z|^2|w|^2$，而 $(1+|zw|)^2-E=1+2|zw|+|zw|^2-zw^*-z^*w-2|zw|$，兩者相等。 ■

階段 3　不等式

想一想　設 $A=(|z|+|w|)^2$ 和 $B=(1+|zw|)^2$。把不等式兩邊平方後，它就是 $\frac{A-E}{B-E}\leq\frac{A}{B}$。這在甚麼時候成立？

所求不等式的兩邊都是非負的，所以它等價於平方後的不等式 $\frac{A-E}{B-E}\leq\frac{A}{B}$。假設 $B-E=|1-zw^*|^2>0$，兩邊乘開：

$$(A-E)B\leq A(B-E)\iff -EB\leq -EA\iff E(B-A)\geq 0.$$

由於 $E\geq 0$，只要 $B\geq A$，即 $1+|zw|\geq|z|+|w|$，上式便成立。

停一停，試一試　證明 $1+|z||w|-|z|-|w|$ 可以因式分解。它在甚麼時候非負？

$1+|z||w|-|z|-|w|=(1-|z|)(1-|w|)$。若 $|z|>1$ 且 $|w|>1$，兩個因式都是負數，乘積為正。所以 $B\geq A$，不等式成立。 ■

若改為 $|z|<1$ 且 $|w|<1$，兩個因式都是正數，乘積同樣為正，同一個論證依然適用：**是的**，不等式在這個情況下也成立。（在這兩個情況下都有 $|zw^*|\neq 1$，所以 $1-zw^*\neq 0$，左邊有定義。）不等式可能不成立的，是混合的情況：一點在單位圓內，另一點在單位圓外。

回顧

- $|a|^2 = aa^*$。每一條恆等式都是由展開共軛的乘積而來。這是用代數方法處理模的標準做法。
- **歸結為一個量的正負。**平方並交叉相乘之後，整條不等式只取決於 $(1-|z|)(1-|w|)$ 的正負。
- **檢查其他情況。**最後一問考慮 $|z|, |w| < 1$；同一個因式分解立即給出答案。

背後的數學

非歐幾里得幾何

兩千年來，數學家一直試圖由歐幾里得的其他公理，證明他的第五公設（過直線外一點，恰好有一條直線與該直線平行）。在 1820 和 1830 年代，羅巴切夫斯基（Nikolai Lobachevsky）、鮑耶（János Bolyai）和（私下研究的）高斯（Gauss）意識到這是辦不到的：存在一種完全相容的幾何，其中有無限多條平行線。1882 年，龐加萊（Henri Poincaré）在單位圓盤內找到了這種幾何的一個優美模型。點就是開圓盤內的點；「直線」是直徑，以及與邊界圓成直角相交的圓弧；而 z 與 w 之間的距離是

$$d(z,w) = 2\tanh^{-1}\rho(z,w), \qquad \text{其中} \qquad \rho(z,w) = \frac{|z-w|}{|1-zw^*|}.$$

當一點趨近邊界圓時，它與圓心的距離趨向無限大：這個圓盤是一整個無限的平面，只是透過一塊令影像變形的透鏡來觀看。

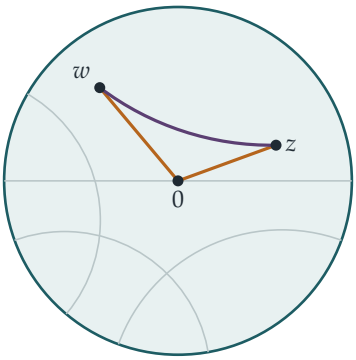


圖 5.8.1 龐加萊圓盤（Poincaré disc）。雙曲「直線」是直徑，或與邊界成直角相交的圓弧（灰色）。以 0、z 和 w 為頂點的雙曲三角形，有兩條邊沿着直徑（琥珀色），一條邊沿着圓弧（紫紅色）。題目中的不等式說明，以雙曲距離計算，紫紅色的邊不會比另外兩條邊加起來更長。

喬裝了的三角不等式

對於 $|z|, |w| < 1$，題目所證明的其實是以下結果。設 $|z| = \tanh a$ 和 $|w| = \tanh b$，則由 0 到 z 和 w 的雙曲距離分別是 $2a$ 和 $2b$。根據 Section 3.8 中 tanh 的加法公式，

$$\frac{|z| + |w|}{1 + |zw|} = \frac{\tanh a + \tanh b}{1 + \tanh a \tanh b} = \tanh(a + b).$$

所以該不等式就是 $\tanh\big(\tfrac{1}{2} d(z, w)\big) \le \tanh(a + b)$，即

$$d(z, w) \le 2a + 2b = d(z, 0) + d(0, w).$$

這就是雙曲距離的三角不等式（圖 5.8.1）。相對論的速度相加公式與雙曲幾何，骨子裏是同一回事：狹義相對論中所有可能速度組成的集合，配上適當的距離概念，就是一個雙曲空間。我們會在 Section 7.6 再談這一點。

圓盤的對稱

為甚麼 ρ 是合適的量？對於 $|a| < 1$，映射 $z \mapsto \dfrac{z - a}{1 - a^* z}$ 把單位圓盤映到自身，保持雙曲距離不變，並把 a 移到圓心。它們是平移在雙曲幾何中的對應物，而 $\rho(z, w)$ 正是把 w 移到圓心時的 z 映射後的 $|$。這些莫比烏斯變換（*Möbius transformation*）是複分析（complex analysis）的核心課題。艾雪（M. C. Escher）的木刻版畫《圓極限》（*Circle Limit*），靈感來自數學家考克斯特（H. S. M. Coxeter）的一幅圖；畫中以全等的魚和天使鑲嵌（tiling）龐加萊圓盤，它們愈近邊緣便愈小，因為雙曲距離在那裏會無限增大。

思維技巧

> **問一問算式有甚麼意義**
>
> 當某個代數式不斷出現時，問一問它是否有某種意義：一個距離、一個角、一個概率、一種對稱。
>
> 在這裏，一個古怪的模之比原來是另一種幾何中的距離，而一條古怪的不等式原來是它的三角不等式。認出當中的意義，就能解釋結果為何成立，以及它通往何處。

延伸練習

1 **數值驗算。**取 $z = \frac{1}{2}$ 和 $w = -\frac{1}{2}$。計算不等式的兩邊。然後計算雙曲距離 $d(z, 0)$、$d(0, w)$ 和 $d(z, w)$，並驗證三角不等式在這裏取等號。為甚麼？

2 **混合的情況。**求滿足 $|z| < 1 < |w|$ 的 z、w，使不等式不成立。

3 **等距映射（isometry）。**設 $f(z) = \dfrac{z - a}{1 - a^{*}z}$，其中 $|a| < 1$。證明當 $|z| = 1$ 時必有 $|f(z)| = 1$，並且 $f(a) = 0$。證明 $\rho(f(z), f(w)) = \rho(z, w)$。

4 **無限近在咫尺。**證明由 0 到 r（其中 $0 < r < 1$）的雙曲距離是 $\ln \frac{1+r}{1-r}$，並且當 $r \to 1$ 時它趨向無限大。一點要多接近邊界，才會與圓心相距雙曲距離 10？

延伸閱讀

- **Caroline Series, “Hyperbolic geometry” lecture notes**（華威大學 University of Warwick，可免費取得）。一份清晰的入門課程講義，由龐加萊圓盤講起。
- **Jeremy Gray, *Ideas of Space***（牛津大學出版社 Oxford University Press）。由歐幾里得到愛因斯坦（Einstein）的非歐幾里得幾何史。
- **Tristan Needham, *Visual Complex Analysis***（牛津大學出版社 Oxford University Press）。第 6 章講述非歐幾里得幾何和莫比烏斯變換，並配有精美的插圖。

我們學到了甚麼

本章開首，我們談到卡爾達諾（Cardano）那個令人不安的發現：要得到實數答案，中途可能非用虛數不可。我們也提出了三個問題：複數（complex number）怎樣幫助我們解決實數問題？單位根（roots of unity）有何特別之處？複函數（complex function）是甚麼模樣？以下是本章八道題目所揭示的答案。

1. 複數怎樣幫助我們解決實數問題？

一次又一次，關於實數的事實，取道複平面（complex plane）便有簡短的證明。

- **角變成乘積。**梅欽（Machin）的 π 公式，說的其實是 $(5-\mathrm{i})^4(1+\mathrm{i})$ 是 $239-\mathrm{i}$ 的實數倍，只需整數運算便可驗證（Section 5.1）。
- **計數變成取平均。**每隔三項取一個二項式係數（binomial coefficient）所得的和，來自在三次單位根處計算 $(1+x)^n$ 的值（Section 5.2）。
- **三角函數的乘積變成多項式的值。**等間隔角的餘弦之積，來自把 $z^{2n}+1$ 因式分解，再代入 $z=\mathrm{i}$（Section 5.3）。
- **幾何變成代數。**距離變成模（modulus），所以由正 n 邊形的一個頂點引出的各條弦，其長度之積是 $|1+1+\cdots+1|=n$（Section 5.4）；旋轉變成乘法，所以費馬點（Fermat point）的作圖化簡為恆等式 $1+\alpha^2=\alpha$（Section 5.6）。
- **由棣美弗定理推出無理性。**展開 $(\cos\theta+\mathrm{i}\sin\theta)^k$，再從奇偶性（parity）着手論證，便知 $\cos^{-1}\frac{1}{3}$ 的度數不是有理數；這正是德恩（Dehn）解決希爾伯特（Hilbert）第三問題的關鍵事實（Section 5.5）。

2. 單位根有何特別之處？

n 次單位根是正 n 邊形的頂點，是最簡單的多項式 z^n-1 的根，並且在乘法之下構成一個羣（group）。這三個事實合起來，解釋了它們為何無處不在。由於它們均勻地分佈在一個圓上，所以總和為零；Section 5.2 的篩選法由此而來，推而廣之就是離散傅立葉變換（discrete Fourier transform）。由於它們是 z^n-1 的根，所以遍取它們的乘積，就是簡單多項式的值（Section 5.3、Section 5.4）。又由於它們對乘法封閉，單位根的冪會循環往復，而像 $\cos^{-1}\frac{1}{3}$ 這樣的角，欠缺的正是這一點（Section 5.5）。到了第 6 章，單位根會成為代數的核心，並通往伽羅瓦理論（Galois theory）。

3. 複函數是甚麼模樣？

在 Section 5.7 中，我們把 e^w 想像成一個把帶狀區域捲成整個平面的映射：它把鉛垂線變成圓，把水平線變成射線，並且無窮無盡地一圈圈繞下去，所以對數是多值的。我們看到它保持角度不變，是共形（conformal）映射的第一個例子，也是麥卡托（Mercator）海圖背後的原理。在 Section 5.8 中，一個古怪的不等式原來就是雙曲幾何（hyperbolic geometry）的三角不等式（triangle inequality）；在雙曲幾何裏，單位圓盤是一個無限的平面，直線是圓弧，而距離則按相對論的速度相加公式合成。

4. 複數工具箱

方法	用於何處
相乘時輻角（argument）相加	梅欽公式（5.1）
檢查 2π 的倍數	梅欽公式（5.1）
對單位根取平均	每隔三項的係數（5.2）
按根因式分解，然後代入	餘弦乘積（5.3）；弦（5.4）
約去一個等於零的因式	5.3；5.4
棣美弗定理（De Moivre's theorem）	倍角公式（5.5）
奇偶性與最小反例（minimal counterexample）	無理角（5.5）
以乘法表示旋轉	等邊三角形（5.6）
映射網格	指數映射（5.7）
$\lvert a\rvert^2 = aa^*$	龐加萊（Poincaré）圓盤不等式（5.8）

下一步

- 單位根是第 6 章的主角，當中包括 Section 6.1 的分圓多項式（cyclotomic polynomial），以及 Section 6.2 中高斯（Gauss）的正 17 邊形作圖。
- 雙曲幾何、tanh 與相對論之間的聯繫，在 Section 7.6 中有進一步的探討。
- 在梅欽公式中出現過的高斯整數（Gaussian integer），是 Section 1.6 的主題。

如果你想在入讀大學之前再進一步

- **Tristan Needham, *Visual Complex Analysis***（《複分析：可視化方法》）：有史以來最優美的複數著作，圖畫豐富。由第 1 章讀起。
- **Paul Nahin, *An Imaginary Tale***（《虛數的故事》）：$\sqrt{-1}$ 的歷史，由卡爾達諾講到現代物理。

- **Liang-shin Hahn, *Complex Numbers and Geometry***（《複數與幾何》）：用複數處理歐幾里得幾何，程度與本章完全相同。
- **Hilary Priestley, *Introduction to Complex Analysis***（《複分析導論》）：讓你初步認識大學的複分析課程。

卡爾達諾認為他的虛數量既精微又無用。五個世紀之後，在數學或物理之中，已經很難找到一個沒有它們蹤影的角落。本章的啟示，正是阿達馬（Hadamard）的那句話：當一個關於實數的問題很難時，不妨試試走出實數線。複平面更加廣闊，許多看似複雜的事情，在那裏原來都很簡單。

延伸練習答案

以下只是簡短的答案和提示，並非完整解答：只列出結果和關鍵步驟，方便你核對自己的做法。請先認真做過每道練習，然後才看答案。

5.1 梅欽公式

1. $(5-\mathrm{i})^2(1+\mathrm{i}) = (24-10\mathrm{i})(1+\mathrm{i}) = 34+14\mathrm{i}$，其輻角（argument）為 $\tan^{-1}\frac{7}{17}$。它的輻角同時也是 $-2\tan^{-1}\frac{1}{5}+\frac{\pi}{4}$。

2. $(2+\mathrm{i})(3+\mathrm{i}) = 5+5\mathrm{i}$，輻角為 $\frac{\pi}{4}$。

3. $(3+\mathrm{i})^2(7+\mathrm{i}) = (8+6\mathrm{i})(7+\mathrm{i}) = 50+50\mathrm{i}$。

4. 這些級數（series）都是交錯級數，所以誤差小於第一個被略去的項。對於 $16\tan^{-1}\frac{1}{5}$，需要被略去的項 $\frac{16}{(2N+1)5^{2N+1}}$ 小於 5×10^{-21}，即要取 $N=14$ 項；對於 $4\tan^{-1}\frac{1}{239}$，取 4 項便足夠。用計算機可得 $16\bigl(\frac{1}{5}-\frac{1}{3\cdot5^3}+\frac{1}{5\cdot5^5}\bigr)-\frac{4}{239}\approx3.14162$。

5.2 單位根篩選法

1. $(1+1)^n+(1-1)^n = 2\bigl(\binom{n}{0}+\binom{n}{2}+\cdots\bigr)$，所以當 $n\ge1$ 時，該和為 2^{n-1}。

2. 先分別乘以 ω^{-1} 和 ω^{-2}，然後取平均：$\frac{1}{3}\bigl(2^n+\omega^{-1}(1+\omega)^n+\omega^{-2}(1+\omega^2)^n\bigr)$。由於 $1+\omega=\mathrm{e}^{\mathrm{i}\pi/3}$，上式等於 $\frac{1}{3}\bigl(2^n+2\cos\frac{(n-2)\pi}{3}\bigr)$。第三個剩餘類（residue）給出 $\frac{1}{3}\bigl(2^n+2\cos\frac{(n+2)\pi}{3}\bigr)$，而三個餘弦 $\cos x+\cos(x-\frac{2\pi}{3})+\cos(x+\frac{2\pi}{3})$ 之和為 0。

3. 對 $j=0,1,2,3$ 取 $(1+\mathrm{i}^j)^n$ 的平均：該和為 $\frac{1}{4}\bigl(2^n+(1+\mathrm{i})^n+0+(1-\mathrm{i})^n\bigr)$，而 $(1\pm\mathrm{i})^n = 2^{n/2}\mathrm{e}^{\pm\mathrm{i}n\pi/4}$。

4. 按元素個數計：$\frac{1}{3}(2^{12}+2\cos4\pi) = 1366$。按元素之和計：$\prod(1+\omega^j)$ 中，2、$1+\omega$ 和 $1+\omega^2$ 這三種因子各出現四次，而 $(1+\omega)(1+\omega^2)=1$，所以該乘積等於 16（在 ω^2 處的乘積亦然）。所求的數目為 $\frac{1}{3}(4096+16+16) = 1376$。

5.3 因式分解 $z^{2n}+1$

1. 在第 (ii) 部分取 $n=7$，得 $\cos^2\frac{\pi}{14}\cos^2\frac{3\pi}{14}\cos^2\frac{5\pi}{14} = \frac{7}{64}$，而三個餘弦均為正數。至於第二個結果，設 $c=\cos\frac{\pi}{7}\cos\frac{2\pi}{7}\cos\frac{4\pi}{7}$。則 $8c\sin\frac{\pi}{7} = \sin\frac{8\pi}{7} = -\sin\frac{\pi}{7}$，所以 $c=-\frac{1}{8}$；另外 $\cos\frac{4\pi}{7} = -\cos\frac{3\pi}{7}$。

2. 在 $z=1$ 處，$2=\prod\bigl(2-2\cos\frac{(2k-1)\pi}{2n}\bigr)=\prod 4\sin^2\frac{(2k-1)\pi}{4n}$，所以各正弦（全為正數）的乘積為 $\sqrt{2/4^n}=2^{\frac12-n}$。

3. $z^6+1=(z^2-\sqrt3 z+1)(z^2+1)(z^2+\sqrt3 z+1)$；首尾兩個因式相乘得 z^4-z^2+1。

4. $z^{2n}=1$ 的根是 ±1 以及共軛（conjugate）對 $\mathrm{e}^{\pm \mathrm{i}k\pi/n}$，其中 $1\le k\le n-1$，所以 $\frac{z^{2n}-1}{z^2-1}=\prod_{k=1}^{n-1}\bigl(z^2-2z\cos\frac{k\pi}{n}+1\bigr)$。令 $z\to1$：左邊趨向 n，右邊則為 $\prod 4\sin^2\frac{k\pi}{2n}$。

5.4 正多邊形中的距離

1. 正方形：$\sqrt2\times2\times\sqrt2=4$。三角形：$\sqrt3\times\sqrt3=3$。

2. 設 $X_k=\omega^k$，並把 $P=r$ 放在正實軸上。則 $\prod|r-\omega^k|=|r^n-1|$。

3. 由算術幾何平均不等式（AM–GM），平均值不小於幾何平均（geometric mean）$n^{1/(n-1)}$。當 $n=6$ 時，各弦（chord）的長度為 1、$\sqrt3$、2、$\sqrt3$、1，平均值為 $\frac{4+2\sqrt3}{5}\approx1.493$，而 $6^{1/5}\approx1.431$。

4. 由第 (ii) 部分，從每個頂點出發的 $n-1$ 條弦的乘積為 n。對全部 n 個頂點相乘得 n^n，其中每條弦都計算了兩次，所以 $\binom{n}{2}$ 條弦的乘積為 $n^{n/2}$。

5.5 名不副實的有理角

1. 取頂點 $(1,1,1)$、$(1,-1,-1)$、$(-1,1,-1)$、$(-1,-1,1)$，中心為 O。兩個鍵向量的點積（dot product）為 -1，長度均為 $\sqrt3$，所以鍵角（bond angle）為 $\cos^{-1}(-\frac13)$。一個面的向外法向量（outward normal），是指向對面頂點的向量的負向量，所以兩個向外法向量之間的夾角同樣是 $\cos^{-1}(-\frac13)$，而兩個面之間的夾角是它的補角（supplement），即 $\cos^{-1}\frac13$。

2. 設 $(4+3\mathrm{i})^k=a_k+\mathrm{i}b_k$。由於 $4+3\mathrm{i}$ 及其共軛是 $t^2-8t+25$ 的根，故 $b_{k+1}=8b_k-25b_{k-1}\equiv3b_k\pmod5$，而 $b_1=3$，所以 $b_k\equiv3^k\not\equiv0\pmod5$。因此沒有任何一個冪是實數。假如該角的度數是有理數（rational number），它的某個倍數便會是 180° 的倍數，相應的冪就會是實數。

3. 當 $a=2$ 時，$a^2-1=3$ 是奇數，所以奇偶性（parity）論證不成立。事實上 $\cos^{-1}\frac12=60^\circ$ 的確是有理數；取 $k=3$ 時，正弦的括號部分為 $3-\binom33\cdot3=0$。

4. 0°、60° 和 90°，餘弦值分別為 1、$\frac12$ 和 0。尼雲定理（Niven's theorem）指出：在 180° 的有理數倍處，cos 的有理數值只有 0、$\pm\frac12$ 和 ±1。

5.6 費馬點

1. 向內作時，各頂點為 $A+(1-\bar\alpha)(B-A)$，餘此類推。在 $p=2$、$q=1.4$、$r=1.1$ 的例子中，通過 C 和新的 L 的直線並不經過 O（檢驗條件 $\operatorname{Im}(\bar CL')=0$ 不成立），

另外兩條直線也是如此。

2. 等邊三角形（equilateral triangle）的中心就是它的形心（centroid）。設 $L = \alpha(p+q)$、$M = -(q+r)$ 和 $N = -\alpha^2(p+r)$，則 ABL、BCM、CAN 的中心 G_1、G_2、G_3 滿足 $G_3 - G_1 = \alpha(G_2 - G_1)$（利用 $\alpha^2 = \alpha - 1$ 展開），這說明 $G_1G_2G_3$ 是等邊三角形。在該例子中，它的邊長為 $\frac{p+q+r}{\sqrt{3}} \approx 2.598$。

3. $PA + PB + PC$ 在 A 處約為 5.68，在 B 處約為 5.13，在 C 處約為 4.89，在形心處約為 4.54；全部都大於 4.5。

4. 等邊（兩種定向（orientation）之一）即是 $z_1 + \omega z_2 + \omega^2 z_3 = 0$ 或 $z_1 + \omega^2 z_2 + \omega z_3 = 0$，其中 $\omega = \mathrm{e}^{2\pi\mathrm{i}/3}$。利用 $\omega + \omega^2 = -1$，兩式左邊的乘積為 $z_1^2 + z_2^2 + z_3^2 - z_1z_2 - z_2z_3 - z_3z_1$。

5.7 作為映射的複指數函數

1. 當 $w = (2k+1)\pi\mathrm{i}$ 時 $\mathrm{e}^w = -1$；當 $w = \frac{1}{2}\ln 2 + \mathrm{i}\left(\frac{\pi}{4} + 2k\pi\right)$ 時 $\mathrm{e}^w = 1 + \mathrm{i}$，其中 k 為整數。

2. 區域 $1 \le |z| \le \mathrm{e}$、$0 \le \arg z \le 1$（一弧度）：圓環（annulus）的一部分。底邊映到實軸上的 $[1, \mathrm{e}]$，右邊映到半徑為 e 的一段圓弧，頂邊映到角度為 1 的一段徑向線段，左邊映到單位圓的一段圓弧。

3. 射線沿半徑方向；圓的切線則與半徑垂直。這與直線 $u = c$ 和 $v = d$ 的情況吻合，因為 e^w 保持角度不變（它的導數永不為 0）。

4. $(x + \mathrm{i}y)^2 = x^2 - y^2 + 2\mathrm{i}xy$。直線 $x = c$ 映到 $u = c^2 - \frac{v^2}{4c^2}$，而 $y = d$ 映到 $u = \frac{v^2}{4d^2} - d^2$：分別是開口向左和向右的拋物線（parabola），兩者的焦點（focus）都在 0。令兩式相等，得 $v^2 = 4c^2d^2$，所以它們相交兩次，交點為 $u = c^2 - d^2$、$v = \pm 2cd$。在 $v = 2cd$ 處（即兩直線交點 $c + \mathrm{i}d$ 的像），斜率 $\frac{\mathrm{d}u}{\mathrm{d}v}$ 分別為 $-\frac{d}{c}$ 和 $\frac{c}{d}$；在 $v = -2cd$ 處則為 $\frac{d}{c}$ 和 $-\frac{c}{d}$。兩種情況下乘積都是 -1。（第二個交點既是第一條直線上 $c - \mathrm{i}d$ 的像，也是第二條直線上 $-c + \mathrm{i}d$ 的像：平方把 z 和 $-z$ 映到同一點。）若 $c = 0$，$x = 0$ 的像是射線 $u \le 0$、$v = 0$；若 $d = 0$，$y = 0$ 的像是射線 $u \ge 0$、$v = 0$：這些都不是拋物線。只要兩直線的交點不在 0，射線仍會與另一個像成直角相交；但若 $c = d = 0$，兩個像都落在實軸上：在 0 處這個映射把角度加倍（它的導數 $2z$ 在該處為零），所以兩軸之間的直角變成了平角。

5.8 圓盤中隱藏的距離

1. 兩邊都等於 $\frac{1}{5/4} = \frac{4}{5}$。於是 $d(z,0) = d(0,w) = 2\tanh^{-1}\frac{1}{2} = \ln 3$，而 $d(z,w) = 2\tanh^{-1}\frac{4}{5} = \ln 9$，所以等號成立：0 位於由 z 到 w 的雙曲線段（hyperbolic line segment，這裏是一條直徑）上。

2. 由第 (i) 和第 (ii) 部分，當 $E > 0$ 時，該不等式等價於 $|z| + |w| \le 1 + |zw|$，即 $(1-|z|)(1-|w|) \ge 0$。在混合的情況下這並不成立，例如取 $z = \tfrac{1}{2}$、$w = 3$：左邊是 $\tfrac{5/2}{1/2} = 5$，右邊是 $\tfrac{7/2}{5/2} = \tfrac{7}{5}$。

3. 若 $|z| = 1$，則 $|1 - a^{*}z| = |\bar{z} - a^{*}| = |z - a|$。直接計算，$f(z) - f(w) = \tfrac{(1-|a|^{2})(z-w)}{(1-a^{*}z)(1-a^{*}w)}$ 及 $1 - f(z)f(w)^{*} = \tfrac{(1-|a|^{2})(1-zw^{*})}{(1-a^{*}z)(1-aw^{*})}$；兩式相除再取模（modulus），得 $\rho(f(z), f(w)) = \rho(z, w)$。

4. $d(0, r) = 2\tanh^{-1} r = \ln \tfrac{1+r}{1-r}$，當 $r \to 1$ 時趨向 ∞。距離為 10 需要 $r = \tanh 5 \approx 0.99991$：離邊界只有約 9×10^{-5}。

章

6

代數與結構

代數中每一個巧妙的技巧背後都有一個結構，而理解這個結構，比掌握多少技巧都更有價值。

本章內容

6.1	以單位根為根的多項式	STEP III 2010
6.2	高斯的正十七邊形	STEP II 1991
6.3	以 $\sqrt{2}+\sqrt{3}+\sqrt{5}$ 為根的多項式	STEP II 2023
6.4	解四次方程	STEP III 2017
6.5	從公理出發的微分法	STEP II 1992
6.6	子羣的併集	STEP II 1988
6.7	四元數羣	STEP III 1993
6.8	由奇異矩陣組成的羣	STEP III 1996
A	練習答案	第 429 頁

從方程到結構

數學是給不同事物起相同名字的藝術。

— 龐加萊（Henri Poincaré）

在歷史上的大部分時間裏，代數就是解方程。四千年前，巴比倫人已懂得解二次方程；十六世紀的意大利數學家解決了三次和四次方程；其後近三百年間，五次方程卻令一切嘗試無功而還。到了 1820 至 1830 年代，問題終於有了答案，而這個答案徹底改變了代數。阿貝爾（Niels Henrik Abel）證明了五次方程不存在一般的求根公式，伽羅瓦（Évariste Galois）則確切說明了哪些方程可以求解，他研究的不是方程本身，而是方程各根的對稱性。為此，他必須創造一種全新的數學對象：羣（group）。

從計算答案轉向研究結構，這個轉變正是本章的主題。現代代數問的不再是「解是甚麼？」，而是「這是哪一類對象？它容許哪些運算？」這些問題的答案，既解釋了舊有技巧為何奏效，也解決了單靠技巧永遠無法解決的問題。

結構為何重要

以下三個古老的問題，兩千年來令一切嘗試都無功而還：只用直尺和圓規，作出體積為已知立方體兩倍的立方體；三等分任意角；作出正 7 邊形。三者都不可能做到，而在十九世紀找到的證明，靠的並不是巧妙的作圖，而是關於結構的論證：所討論的是尺規所能到達的數所組成的域(field)的次數(degree)(Section 6.3)。同一套理論也解釋了為何正 17 邊形可以作出 (Section 6.2)，以及為何四次方程可解而五次方程不可解 (Section 6.4)。

結構還可以搬到別處去。晶體的對稱、多項式各根的置換（permutation）、三維空間的旋轉、扭計骰的各種轉法，以及橢圓曲線（elliptic curve）上的點，全都滿足羣的公理（axiom）。只要從公理出發把一條定理證明一次，它便在以上所有情況中成立。這種簡約，即龐加萊所說的給「不同事物起相同名字」，正是抽象代數（abstract algebra）如此強大的原因。

你在大學會遇到甚麼

在一年級，你會遇到羣（帶有一個運算並滿足三條公理的集合），通常還有環（ring）和域（帶有加法和乘法的集合）。你會證明拉格朗日定理（Lagrange's theorem），即子羣（subgroup）的大小整除羣的大小，並認識置換、對稱羣（symmetry group）和矩陣羣（matrix group）。到了二、三年級，這些內容會引向伽羅瓦理論(Galois theory)，它把域

和羣聯繫起來，回答有關方程和作圖的經典問題；引向表示論（representation theory），它透過矩陣來研究羣；還會引向代數數論（algebraic number theory），Section 1.7 中的想法在那裏會得到嚴格的基礎。本章最後三節講的是羣。羣已不在 STEP 的課程範圍之內（已於 2019 年刪去），但這些題目本身自成一體，每一節都會給出所需的定義。

貫穿本章的問題

三個大問題

1. **甚麼可以由甚麼構造出來？**從有理數出發，透過開平方根或解方程，可以到達哪些數？又怎樣證明某個數無法到達？
2. **方程的公式為何奏效？**二次、三次和四次方程都有求根公式，背後隱藏的原因是甚麼？到了五次，又有甚麼改變了？
3. **抽象化帶來甚麼？**如果我們只憑對象所遵守的規則來描述它們，會有甚麼結果？

本章的想法

§	課題	想法
6.1	分圓多項式（cyclotomic polynomials）	按階（order）把單位根（roots of unity）分類
6.2	正 17 邊形	高斯週期（Gauss periods）：和與積給出二次方程
6.3	最小多項式（minimal polynomials）	消去根式；次數相乘
6.4	解四次方程	對稱函數（symmetric functions）；三次預解式（resolvent cubic）
6.5	從公理出發的微分法	由規則決定的運算
6.6	子羣的併集（union）	從羣公理出發作論證
6.7	四元數羣（quaternion group）	用矩陣表示運算
6.8	由奇異矩陣（singular matrices）組成的羣	按字面理解定義

如何閱讀本章

這些題目當中，有不少可以靠蠻力完成：把所有式子展開、比較係數、逐一檢查每種情況。如有需要，不妨先這樣做。然後回頭想一想，是甚麼令蠻力行得通，背後有沒有結構上的原因。結構上的原因通常較為簡短，總是更具啟發性，而當蠻力再也行不通時，

你需要的正是它。

6.1 以單位根為根的多項式

題目

STEP III 2010，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/10-s3-q3`

引子

n 次單位根（roots of unity），即 $z^n = 1$ 的解，位於正 n 邊形的各個頂點上。但它們並非全都一樣。例如在 12 次單位根之中，1 本身已是一次單位根，-1 是二次單位根，i 是四次單位根，而十二個之中只有四個，即本原（primitive）的那些，才需要走足十二步方能回到 1。

按所需的步數把單位根分類，再由每一類構造一個多項式，便得到分圓多項式（cyclotomic polynomials，此名源自希臘文，意即「分割圓周」）。它們是 $z^n - 1$ 的基本構件，正如質數是整數的基本構件一樣；高斯（Gauss）的正 17 邊形作圖法（Section 6.2）、單位根篩選法（Section 5.2）以及現代數論中的不少內容，背後都有它們的身影。

動筆之前

想一想 每個 n 次單位根都滿足 $z^n = 1$。但使它等於 1 的最小正整數次冪，可能小於 n。對於四個 4 次單位根 1、i、-1、$-\mathrm{i}$，這個最小的冪次分別是多少？

$1^1 = 1$，$(-1)^2 = 1$，而 $\mathrm{i}^4 = (-\mathrm{i})^4 = 1$，且沒有更小的冪次可行。所以最小的冪次依次是 1、4、2、4。這個最小的冪次稱為該單位根的階（order），而本原 n 次單位根正是階為 n 的那些單位根。整道題的關鍵事實是：每個單位根都恰好有一個階。

逐步解題

階段 1　最初的幾個

本原 4 次單位根是 i 和 $-\mathrm{i}$，所以 $\mathrm{C}_4(x) = (x-\mathrm{i})(x+\mathrm{i}) = x^2+1$。 ■

> **想一想**　哪些單位根的階是 1、2、3、5 和 6？（畫出 6 次單位根，並按階把它們分類。）

階為 1：只有 1，所以 $\mathrm{C}_1(x) = x-1$。階為 2：只有 -1，所以 $\mathrm{C}_2(x) = x+1$。

階為 3：除 1 以外的三次單位根，即 $\frac{x^3-1}{x-1}$ 的根。所以 $\mathrm{C}_3(x) = x^2+x+1$。

階為 5：除 1 以外的五次單位根，它們是 $\frac{x^5-1}{x-1}$ 的根，所以 $\mathrm{C}_5(x) = x^4+x^3+x^2+x+1$。

階為 6：在六次單位根 $e^{k\pi\mathrm{i}/3}$，$k = 0,\dots,5$ 之中，$k = 0$、3 的階分別是 1、2，而 $k = 2$、4 的階是 3。餘下的是 $e^{\pm\pi\mathrm{i}/3}$，所以

$$\mathrm{C}_6(x) = (x-e^{\mathrm{i}\pi/3})(x-e^{-\mathrm{i}\pi/3}) = x^2 - 2x\cos\tfrac{\pi}{3} + 1 = x^2 - x + 1. \qquad ■$$

階段 2　哪個 n 給出 x^4+1？

$x^4+1=0$ 的根滿足 $x^4=-1$，所以 $x^8=1$ 但 $x^4\neq 1$。它們的階整除 8 但不整除 4，所以階是 8。反過來，階為 8 的單位根滿足 $x^8=1$，$x^4\neq 1$，所以 $x^4=-1$。本原 8 次單位根有四個，x^4+1 的根也有四個，所以 $n=8$。 ■

階段 3　質數

若 p 是質數，則 p 次單位根的階整除 p，所以階是 1 或 p。因此除 1 以外的每個 p 次單位根都是本原的，而

$$\mathrm{C}_p(x) = \frac{x^p-1}{x-1} = x^{p-1}+x^{p-2}+\cdots+x+1. \qquad ■$$

階段 4　沒有分圓多項式是另外兩個分圓多項式之積

> **停一停，試一試**　假設 $\mathrm{C}_q = \mathrm{C}_r\mathrm{C}_s$。取 C_r 的一個根。它的階是多少？這條恆等式對它的階有何說法？由此必然得出甚麼？

C_r 的每個根（至少有一個，因為 $e^{2\pi\mathrm{i}/r}$ 是本原的）都是 $\mathrm{C}_r\mathrm{C}_s = \mathrm{C}_q$ 的根，所以它的階是 q。但它的階同時也是 r，而一個單位根只有一個階，所以 $r=q$。於是 $\mathrm{C}_q = \mathrm{C}_q\,\mathrm{C}_s$，又因為 C_q 不是零多項式，所以 $\mathrm{C}_s = 1$。這是不可能的，因為 C_s 至少有一個根 $e^{2\pi\mathrm{i}/s}$。所以這樣的 q、r、s 並不存在。 ■

回顧

- **按階分類。**每個單位根都有唯一的階，按階分組便把 $z^n - 1$ 的根分成互不相交的類。
- **用根來定義多項式。**每個 C_n 都是藉着確切指明哪些數是它的根而確定下來的。
- **唯一性帶來不可能性。**第 (iv) 部分只需要一個想法：一個根不可能有兩個不同的階。

背後的數學

分解 $x^n - 1$

每個 n 次單位根都有某個整除 n 的階 d；而當 $d \mid n$ 時，每個階為 d 的 d 次單位根都是 n 次單位根。所以 $x^n - 1$ 的根按階分成若干類，由此得到

$$x^n - 1 = \prod_{d \mid n} C_d(x).$$

例如 $x^6 - 1 = C_1C_2C_3C_6 = (x-1)(x+1)(x^2+x+1)(x^2-x+1)$。圖 6.1.1 展示了按這種方式分類的十二個 12 次單位根。本原 n 次單位根的個數，亦即 C_n 的次數，等於 1 至 n 之間與 n 沒有公因數的 k 的個數，也就是歐拉函數（Euler's totient function）$\varphi(n)$。比較兩邊的次數，便得到恆等式 $\sum_{d \mid n} \varphi(d) = n$。

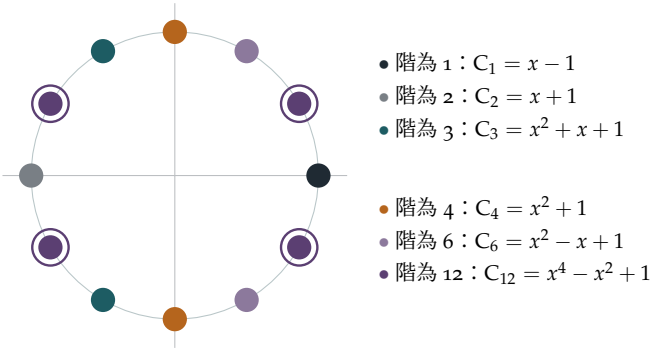


圖 6.1.1 十二個 12 次單位根，按階着色。四個本原 12 次單位根（以圓圈標示）是 C_{12} 的根。各類合起來給出 $x^{12} - 1 = C_1C_2C_3C_4C_6C_{12}$，次數為 $1+1+2+2+2+4 = 12$。

整數係數，以及一個意外

單看定義並不明顯，但每個分圓多項式的係數都是整數。（利用 $x^n - 1$ 的因式分解和歸納法（induction）：C_n 等於 $x^n - 1$ 除以若干個較早的分圓多項式之積，而以首一（monic）整係數多項式對整係數多項式作長除法，所得係數仍是整數。）在很長一段時間裏，所有已知例子的係數都只有 0、1 和 -1。第一個例外是 C_{105}，它含有 $-2x^7$ 這一項，由 Migotti 於 1883 年發現；$105 = 3 \times 5 \times 7$ 是具有三個相異奇質因數的最小的數。事實上，分圓多項式的係數可以要多大有多大。

不可約性

第 (iv) 部分證明了沒有分圓多項式是另外兩個分圓多項式之積。其實更強的結論也成立：C_n 根本不能分解成次數較低的有理係數多項式之積。高斯對 n 為質數的情況證明了這一點（它可由艾森斯坦判別法（Eisenstein's criterion）推出，你會在大學的第一門代數課中遇到），而克羅內克（Kronecker）和戴德金（Dedekind）則證明了一般 n 的情況。正是這種不可約性（irreducibility），使分圓多項式成為與單位根有關的多項式之中的「質數」，它也是單位根的伽羅瓦理論（Galois theory）的起點。

思維技巧

先分類，後構造

用一個定義明確的不變量（*invariant*），把問題中的對象分成若干類，然後逐類建立結構。

這裏的不變量是單位根的階。由於每個根恰好有一個階，各類互不重疊，於是 $x^n - 1$ 可分解成遍及各類的乘積。同樣的模式（由唯一的不變量得出不相交的分解）也是質因數分解和有限阿貝爾羣（finite abelian groups）分類的基礎。

延伸練習

1. **檢查因式分解。**驗證 $(x-1)(x+1)(x^2+x+1)(x^2-x+1) = x^6 - 1$，並驗證 $x^{12} - 1$ 的因式分解如圖 6.1.1 所示。
2. **質數的兩倍。**證明若 p 是奇質數，則 $C_{2p}(x) = C_p(-x)$。就 $p = 3$ 驗證這個結果。
3. **歐拉函數。**證明對於質數 p，有 $\varphi(p^k) = p^k - p^{k-1}$，並且 $C_{p^k}(x) = C_p\left(x^{p^{k-1}}\right)$。利用這個結果求 C_9。

4 較長的一題。 較難 利用 $x^{15}-1=\mathrm{C}_1\mathrm{C}_3\mathrm{C}_5\mathrm{C}_{15}$，證明 $\mathrm{C}_{15}(x)=x^8-x^7+x^5-x^4+x^3-x+1$。

延伸閱讀

- **Ian Stewart, *Galois Theory***（CRC Press）。大學本科的標準入門書；分圓多項式和正多邊形在書中很早便出現。
- **Harold Davenport, *The Higher Arithmetic***（Cambridge University Press）。內容包括單位根及其算術性質。
- **Marcus du Sautoy, *Finding Moonshine***（Fourth Estate）。一本講述對稱與羣論的科普讀物，當中包括伽羅瓦的故事。

6.2 高斯的正十七邊形

題目

STEP II 1991，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/91-s2-q4`

引子

1796 年 3 月 30 日，哥廷根（Göttingen）一名十八歲的學生在他的數學日記寫下了第一則記錄，這本日記他此後寫了一輩子。他發現了怎樣用尺規作圖（ruler and compasses construction）作出正 17 邊形。古希臘人早已懂得作出邊數為 3、4、5 和 15 的正多邊形（regular polygon），以及邊數為這些數目反覆加倍的正多邊形，但其後兩千年間，再沒有人添上任何一個新的。高斯（Carl Friedrich Gauss）後來說，正是這項發現令他決心成為數學家，而不是語文學家；相傳他還曾要求在自己的墓碑上刻一個正 17 邊形。

這個作圖法的核心，就是本題的後半部分：四個特定餘弦之和等於 $\frac{\sqrt{17}-1}{4}$，這個數只涉及一個平方根。由平方根構成的長度都可以用尺規作出，而高斯再用三個同類的步驟，就能作出 $\cos\frac{2\pi}{17}$ 本身。

動筆之前

? **想一想** 為甚麼一道關於正 17 邊形的題目，第一部分卻是關於 $\cos\frac{2\pi}{5}$ 的？

那是熱身，把同一個想法用在較小的多邊形上。對正五邊形（pentagon）而言，和 $\cos\phi + \cos 2\phi$ 滿足一條有理係數的二次方程（quadratic），所以解一條二次方程就能求出它。對正 17 邊形而言，四個餘弦之和 z 原來也滿足一條二次方程。兩種情況的竅門都是把餘弦相乘，再用 $\cos A\cos B = \frac{1}{2}(\cos(A+B)+\cos(A-B))$ 把乘積變回同一批角的餘弦之和。

逐步解題

階段 1　正五邊形

想一想　展開 $y^2=(\cos\phi+\cos 2\phi)^2$，把每個乘積化為和，再利用 $5\phi=2\pi$ 化簡 3ϕ 和 4ϕ 這類角。

$$y^2=\cos^2\phi+2\cos\phi\cos 2\phi+\cos^2 2\phi=\tfrac12(1+\cos 2\phi)+(\cos 3\phi+\cos\phi)+\tfrac12(1+\cos 4\phi).$$

由於 $5\phi=2\pi$，有 $\cos 3\phi=\cos(2\pi-2\phi)=\cos 2\phi$ 和 $\cos 4\phi=\cos\phi$。所以

$$y^2=1+\tfrac12(\cos\phi+\cos 2\phi)+(\cos\phi+\cos 2\phi)=1+\tfrac32 y,$$

即 $2y^2=3y+2$。 ■方程 $2y^2-3y-2=(2y+1)(y-2)=0$ 的根是 $y=2$ 和 $y=-\tfrac12$。但 $y=2$ 需要兩個餘弦都等於 1，而事實並非如此，所以 $y=-\tfrac12$。 ■

階段 2　十七次單位根之和為零

$\sum_{k=0}^{16}\cos k\theta$ 是 $\sum_{k=0}^{16}\omega^k$ 的實部（real part），其中 $\omega=e^{\mathrm{i}\theta}$ 是一個 17 次本原單位根（primitive root of unity）。這個等比級數（geometric series）等於 $\frac{\omega^{17}-1}{\omega-1}=0$。 ■

所以 $\sum_{k=1}^{16}\cos k\theta=-1$。由於 $\cos(17-k)\theta=\cos k\theta$，各項可以兩兩配對，於是

$$\sum_{k=1}^{8}\cos k\theta=-\tfrac12.$$

階段 3　把餘弦分成兩組

想一想　設 $z=\cos\theta+\cos 2\theta+\cos 4\theta+\cos 8\theta$，並設 z' 為其餘四個餘弦之和 $\cos 3\theta+\cos 5\theta+\cos 6\theta+\cos 7\theta$。$z+z'$ 是多少？

由第 2 階段，$z+z'=-\tfrac12$。如果還能求出 zz'，那麼 z 和 z' 就是某條二次方程的兩個根。

停一停，試一試　計算 zz'。把它展開成 16 個乘積 $\cos a\theta\cos b\theta$，把每個乘積化為兩個餘弦之和，然後數一數 $\cos\theta,\cos 2\theta,\dots,\cos 8\theta$ 各出現多少次。（利用 $\cos k\theta=\cos(17-k)\theta$ 和 $\cos(-k\theta)=\cos k\theta$ 化簡各角。）

每個乘積給出 $\cos a\theta\cos b\theta=\tfrac12(\cos(a+b)\theta+\cos(a-b)\theta)$，所以 zz' 是 32 個餘弦之和的 $\tfrac12$ 倍。逐一考慮 $a\in\{1,2,4,8\}$ 和 $b\in\{3,5,6,7\}$ 的 16 對組合：例如 $a=1$、$b=3$ 給出 $\cos 4\theta$ 和 $\cos 2\theta$；$a=8$、$b=7$ 給出 $\cos 15\theta=\cos 2\theta$ 和 $\cos\theta$。對全部十六對都這樣

做，$\cos\theta,\ldots,\cos 8\theta$ 每一個都恰好出現四次。因此

$$zz' = \tfrac{1}{2}\times 4\times\sum_{k=1}^{8}\cos k\theta = 2\times\left(-\tfrac{1}{2}\right) = -1.$$

階段 4 二次方程

所以 z 和 z' 是 $t^2+\frac{1}{2}t-1=0$ 的兩個根，即 $4t^2+2t-4=0$，其根為 $t=\frac{-1\pm\sqrt{17}}{4}$。哪一個是 z？按數值計算，$\cos\theta\approx 0.932$、$\cos 2\theta\approx 0.739$、$\cos 4\theta\approx 0.092$ 和 $\cos 8\theta\approx -0.983$，所以 $z\approx 0.78>0$。因此

$$z=\frac{-1+\sqrt{17}}{4}=-\frac{1-\sqrt{17}}{4}. \qquad \blacksquare$$

回顧

- **積化和。**餘弦的積化和公式把 y^2 和 zz' 變成同一批餘弦之和，於是代數運算得以自成一體。
- **和與積給出二次方程。**知道 $z+z'$ 和 zz'，就等於知道以 z 和 z' 為根的二次方程。
- **分組要分得對。**把八個餘弦分成 $\{1,2,4,8\}$ 和 $\{3,5,6,7\}$，令每個指標在 zz' 中出現的次數相同。隨意分組是行不通的。

背後的數學

高斯週期

為甚麼是 $\{1,2,4,8\}$？以 17 為模（modulo），3 的冪 $3^0,3^1,3^2,\ldots$ 會走遍所有非零剩餘（residue）：

$$1,3,9,10,13,5,15,11,16,14,8,7,4,12,2,6.$$

每隔一個取一個，即 $3^0,3^2,3^4,\ldots$，得出 1、9、13、15、16、8、4、2，以 17 為模就是 ±1、±2、±4、±8：正好是 z 中的角。另一半給出 $\pm3,\pm5,\pm6,\pm7$，即 z' 中的角。高斯把 ω^k 在這類剩餘類上的和稱為週期（periods）。由於 $17-1=16=2^4$，這些剩餘可以對分四次，而在每一階段，週期都滿足一些二次方程，其係數涉及上一階段的週期。四條二次方程，也就是四個平方根，就從有理數一路通往 $\cos\frac{2\pi}{17}$：

$$\cos\tfrac{2\pi}{17}=\tfrac{1}{16}\Big(-1+\sqrt{17}+\sqrt{34-2\sqrt{17}}\\ +2\sqrt{17+3\sqrt{17}-\sqrt{34-2\sqrt{17}}-2\sqrt{34+2\sqrt{17}}}\Big).$$

每一步都只涉及平方根，而平方根可以用尺規作出（利用相似三角形或半圓），所以正 17 邊形是可作圖（constructible）的。

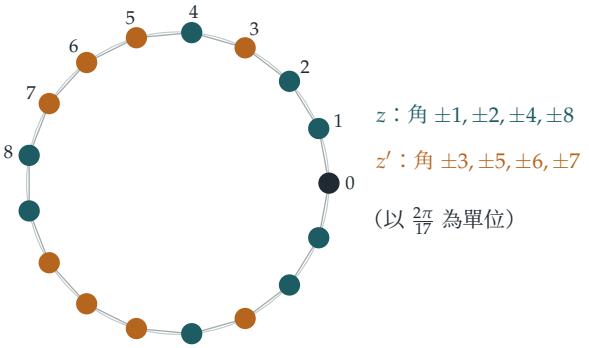


圖 6.2.1　正 17 邊形的頂點，按高斯的兩個週期分成兩組。藍綠色頂點位於角 $\pm 1, \pm 2, \pm 4, \pm 8$（以 $\frac{2\pi}{17}$ 為單位），它們的 x 坐標之和是 $2z = \frac{-1+\sqrt{17}}{2}$；琥珀色頂點則給出 $2z' = \frac{-1-\sqrt{17}}{2}$。

哪些正多邊形可以作圖？

高斯斷言，而 Pierre Wantzel 在 1837 年證明：正 n 邊形可以用尺規作出，當且僅當 n 是 2 的冪乘以若干個互不相同的費馬質數（Fermat primes）之積，費馬質數即形如 $2^{2^m}+1$ 的質數。已知的費馬質數有 3、5、17、257 和 65537，它們是 Section 1.1 中的費馬數。（費馬（Fermat）猜想所有這類數都是質數，但歐拉（Euler）證明了 $2^{32}+1$ 可被 641 整除，而此後再沒有人找到其他費馬質數。）所以正 257 邊形和正 65537 邊形都可以作圖；十九世紀時，Johann Hermes 花了大約十年時間推算出正 65537 邊形的作圖法，他的手稿現存於哥廷根。正 7 邊形和正 11 邊形不可作圖，因為 7 和 11 不是費馬質數（$7-1=6$ 和 $11-1=10$ 都不是 2 的冪）；正 9 邊形也不可作圖，因為 $9=3^2$ 把費馬質數 3 重複了。Wantzel 的證明用到 Section 6.3 中關於域次數（field degrees）的理論。

思維技巧

計算和與積來求出二次方程

要求出一個量，可以爲它配上一個自然的夥伴，計算兩者的和與積，再解由此得出的二次方程。

箇中的藝術在於選對夥伴，使和與積都算得出來。這裏的夥伴來自以 17 為模的剩餘的對稱性，而把這個想法重複四次，就得出高斯的作圖法。

延伸練習

1 **正七邊形（heptagon）。**試做 STEP II 1995，第 6 題（`step.maths.org/questions/95-s2-q6`）。這道題把高斯將單位根分組的想法用於七次單位根，並求出餘弦之和與正弦之和的準確值。判斷正弦之和的正負號是當中最微妙的一步。

2 **正五邊形。**證明 $\cos\frac{2\pi}{5} = \frac{\sqrt{5}-1}{4}$，並由此描述一個正五邊形的尺規作圖法。

3 **另一個週期。**用數值驗證 $z' = \cos 3\theta + \cos 5\theta + \cos 6\theta + \cos 7\theta = \frac{-1-\sqrt{17}}{4}$。

4 **生成元（generators）。**證明以 17 為模時，3 的冪給出所有非零剩餘，但 2 的冪卻不然。2 的冪給出哪些剩餘？

延伸閱讀

- **Ian Stewart, *Galois Theory***（CRC Press）。當中有一章講述尺規作圖和正 17 邊形。
- **Richard Courant & Herbert Robbins, *What Is Mathematics?***（Oxford University Press）。第 3 章以淺白優美的筆觸介紹作圖問題，包括正 17 邊形。
- **Daniel Kehlmann, *Measuring the World***（Quercus）。一本關於高斯和洪堡（Alexander von Humboldt）的詼諧小說，可作輕鬆讀物。

6.3 以 $\sqrt{2}+\sqrt{3}+\sqrt{5}$ 為根的多項式

題目

STEP II 2023，第 4 題

繼續閱讀之前，請先看題目（原題為英文）。題目載於 OCR 提供下載的 2023 年 STEP 試卷（zip 檔案）內的 STEP II 試卷：

`www.ocr.org.uk/Images/707102-2023-past-papers.zip`

引子

$\sqrt{2}$ 是 x^2-2 的根，而這是一個整數係數多項式（polynomial）。同樣，$\sqrt{3}$ 是 x^2-3 的根。那麼 $\sqrt{2}+\sqrt{3}$ 是否也是某個整數係數多項式的根？$\sqrt{2}+\sqrt{3}+\sqrt{5}$ 又如何？$\sqrt[3]{2}+\sqrt[3]{3}$ 呢？

凡是整數係數多項式的根，都稱為*代數數*（algebraic number）。這道題透過親手消元（elimination），說明代數數之和仍是代數數。計算所得的次數 4、8、6、9 並非偶然：它們量度每個數有多「複雜」，而這種量度，即*域擴張的次數*（degree of a field extension），正是最終解決古希臘倍立方與三等分角問題的關鍵。

動筆之前

想一想 要找一個以 $\sqrt{2}$ 為根的多項式，你會把 $x=\sqrt{2}$ 平方，從而消去平方根。若有兩個不同的平方根，怎樣才能逐一消去？

先把其中一個平方根單獨移到一邊再平方，它便會消失；然後把另一個平方根單獨移到一邊，再平方一次。每次平方都可能令次數加倍。整道題的策略就是如此：逐一消去根式（radical）。

逐步解題

階段 1　兩個平方根

若 $(x-\sqrt{2})^2=3$，則 $x^2-2\sqrt{2}\,x+2=3$，所以 $x^2-1=2\sqrt{2}\,x$。平方便可消去餘下的平方根：$(x^2-1)^2=8x^2$，即

$$x^4-10x^2+1=0. \qquad \blacksquare$$

而 $x=\sqrt{2}+\sqrt{3}$ 滿足 $(x-\sqrt{2})^2=(\sqrt{3})^2=3$，所以 $\mathrm{f}(\sqrt{2}+\sqrt{3})=0$。 $\blacksquare$

階段 2　三個平方根

想一想　設 $x=\sqrt{2}+\sqrt{3}+\sqrt{5}$，則 $x-\sqrt{5}=\sqrt{2}+\sqrt{3}$ 是 f 的根。把 $\mathrm{f}(x-\sqrt{5})=0$ 寫成 $P(x)+\sqrt{5}\,Q(x)=0$ 的形式。之後怎樣消去 $\sqrt{5}$？

利用 $(x-\sqrt{5})^2=x^2+5-2\sqrt{5}x$ 和 $(x-\sqrt{5})^4=\bigl(x^2+5-2\sqrt{5}x\bigr)^2=x^4+30x^2+25-4\sqrt{5}\,x(x^2+5)$ 展開：

$$\mathrm{f}(x-\sqrt{5})=\bigl(x^4+30x^2+25\bigr)-4\sqrt{5}\,x(x^2+5)-10\bigl(x^2+5\bigr)+20\sqrt{5}\,x+1.$$

合併同類項，得 $P(x)=x^4+20x^2-24$ 和 $Q(x)=-4x^3-20x+20x=-4x^3$。所以 $x^4+20x^2-24=4\sqrt{5}\,x^3$。平方得

$$\bigl(x^4+20x^2-24\bigr)^2=80x^6,$$

展開後得

$$\mathrm{g}(x)=x^8-40x^6+352x^4-960x^2+576. \qquad \blacksquare$$

（驗算：$(x^4+20x^2-24)^2=x^8+40x^6+352x^4-960x^2+576$，其中用到 $400-48=352$；再減去 $80x^6$ 便得 g。）

階段 3　一個三次方程與一個平方根

想一想　若 $x=a+\sqrt{2}$，則 $x-\sqrt{2}$ 是 t^3-3t+1 的根。代入後，把含有 $\sqrt{2}$ 的項與不含它的項分開，然後平方。

$(x-\sqrt{2})^3=x^3-3\sqrt{2}\,x^2+6x-2\sqrt{2}$，所以

$$(x-\sqrt{2})^3-3(x-\sqrt{2})+1=\bigl(x^3+3x+1\bigr)-\sqrt{2}\bigl(3x^2-1\bigr)=0.$$

因此 $(x^3+3x+1)^2=2(3x^2-1)^2$。展開得 $x^6+6x^4+2x^3+9x^2+6x+1=18x^4-12x^2+2$，所以

$$\mathrm{h}(x)=x^6-12x^4+2x^3+21x^2+6x-1. \qquad \blacksquare$$

同樣的論證對 $b+\sqrt{2}$ 和 $c+\sqrt{2}$ 也成立，因為我們只用到 a 是 t^3-3t+1 的根這一事實。（h 的另外三個根是 $a-\sqrt{2}$、$b-\sqrt{2}$ 和 $c-\sqrt{2}$，它們是在平方時引入的。）

階段 4　兩個立方根

停一停，試一試 設 $x=\sqrt[3]{2}+\sqrt[3]{3}$。把兩邊立方。你能否把結果用 x 和單獨一個立方根表示？

利用 $(u+v)^3=u^3+v^3+3uv(u+v)$，取 $u=\sqrt[3]{2}$，$v=\sqrt[3]{3}$：

$$x^3=2+3+3\sqrt[3]{6}\,x, \qquad \text{所以} \qquad x^3-5=3\sqrt[3]{6}\,x.$$

再立方一次，便消去最後一個立方根：$(x^3-5)^3=27\cdot 6\,x^3=162x^3$。展開 $(x^3-5)^3=x^9-15x^6+75x^3-125$，得

$$\mathrm{k}(x)=x^9-15x^6-87x^3-125. \qquad \blacksquare$$

回顧

- **逐一消去根式。**把一個根式單獨移到一邊，取適當的次方，然後重複這個步驟。
- **把有理部分與無理部分分開。**把式子寫成 $P+\sqrt{5}\,Q$，再利用 $P^2=5Q^2$，是關鍵的一步。
- **平方會帶來額外的根。**每次平方都會引入「共軛」（conjugate）根，例如除了 $a+\sqrt{2}$ 之外還有 $a-\sqrt{2}$。最後得到的多項式以它們全部為根。

背後的數學

域及其次數

所有形如 $p+q\sqrt{2}$（其中 p、q 為有理數）的數所組成的集合 $\mathbb{Q}(\sqrt{2})$，對加、減、乘、除（除數不為零）都封閉：它是一個域（field）。它是 $\mathbb{Q}$ 上的二維「向量空間」（vector space），以 $1,\sqrt{2}$ 為基底（basis），所以我們說它在 $\mathbb{Q}$ 上的次數（degree）是 2。添加 $\sqrt{3}$ 便得到 $\mathbb{Q}(\sqrt{2},\sqrt{3})$，其基底為 $1,\sqrt{2},\sqrt{3},\sqrt{6}$，次數為 4；再添加 $\sqrt{5}$，次數便是 8（圖 6.3.1）。*塔定律*（tower law）指出，沿着一座域塔往上，次數會相乘。由一個數生成的域，其次數等於該數的*最小多項式*（minimal polynomial）的次數，即該數所滿足的有理係數多項式之中次數最低的一個。可以證明，本題的多項式 f 和 g 都是最小多項式，所以 $\sqrt{2}+\sqrt{3}$ 的次數是 4，而 $\sqrt{2}+\sqrt{3}+\sqrt{5}$ 的次數是 8：事實上，單憑 $\sqrt{2}+\sqrt{3}+\sqrt{5}$ 這一個數，就足以生成整個 8 次域。

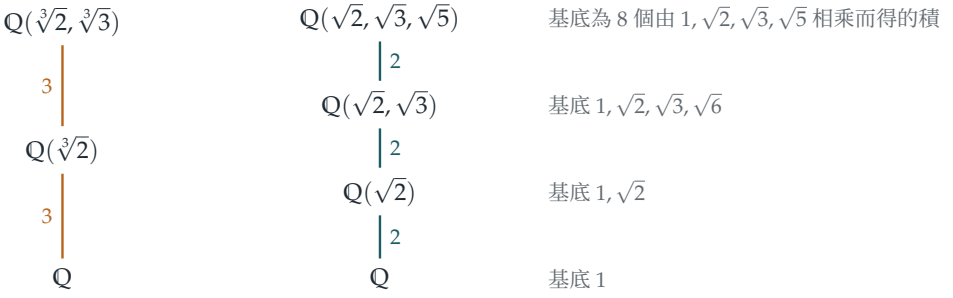


圖 6.3.1　域塔。右：每添加一個平方根，次數便加倍，得出 $2\times 2\times 2=8$，即 g 的次數。左：每添加一個立方根，次數便變為三倍，得出 $3\times 3=9$，即 k 的次數。沿着域塔往上，次數會相乘；這條塔定律正是證明某些尺規作圖不可能完成的關鍵。

倍立方

尺規作圖（ruler-and-compass construction）所能作出的長度，只限於由有理數出發、反覆解一次和二次方程而得到的數。每一步都給出一個次數為 1 或 2 的域擴張，所以根據塔定律，每個可作圖（constructible）的長度都位於一個次數為 2 的冪的域之內，而它本身的次數也是 2 的冪。但 $\sqrt[3]{2}$ 的最小多項式是 x^3-2，次數為 3。由於 3 不是 2 的冪，$\sqrt[3]{2}$ 不可作圖，因此約於公元前 430 年提出的提洛島（Delos）問題，即用尺規把立方體的體積加倍，是不可能做到的。旺策爾（Pierre Wantzel）在 1837 年證明了這一點，同時證明了一般的角不可能三等分，並完成了可作圖正多邊形的分類（Section 6.2）。兩千年來的種種嘗試，就此被一個關於次數的論證畫上句號。

代數數構成一個域

本題的方法透過例子說明，代數數之和仍是代數數。一般情況則要用線性代數（linear algebra）來證明：若 α 和 β 的次數分別為 m 和 n，則所有 $\alpha^i\beta^j$（其中 $i<m$，$j<n$）這些數所張成的空間，維數至多為 mn，於是 $mn+1$ 個冪 $1,\alpha+\beta,(\alpha+\beta)^2,\ldots$ 必定線性相關（linearly dependent），由此得出一個次數至多為 mn 的多項式。我們得到的次數 $4=2\times 2$、$6=3\times 2$ 和 $9=3\times 3$ 與這個規律完全吻合。不是代數數的數稱為超越數（transcendental number）；e 和 π 都是例子（分別由埃爾米特（Hermite）於 1873 年及林德曼（Lindemann）於 1882 年證明），而林德曼關於 π 是超越數的定理，也證明了化圓為方同樣不可能。

思維技巧

先消去，再量度

要找出一個含根式的式子所滿足的多項式，可把根式逐個單獨移到一邊，再取適當的次方，逐一消去。

然後問一問：次數說明了甚麼？最小多項式的次數是該數的一個不變量（invariant），而比較次數是證明某事不可能的最有力方法之一。

延伸練習

1 **本原元（primitive element）。**設 $y=\sqrt{2}+\sqrt{3}$。證明 $y^3=11\sqrt{2}+9\sqrt{3}$，並由此推出 $\sqrt{2}=\frac{1}{2}(y^3-9y)$。因此每個形如 $p+q\sqrt{2}+r\sqrt{3}+s\sqrt{6}$ 的數都可以寫成 y 的多項式。

2 **混合根式。**求一個次數為 6 的整數係數多項式，使它以 $\sqrt{2}+\sqrt[3]{3}$ 為根。

3 **不可約性（irreducibility）。**證明 x^4-10x^2+1 沒有有理根，而且不能寫成兩個整數係數二次式的積。（提示：試用 $(x^2+ax+b)(x^2-ax+c)$。）由此推出 $\sqrt{2}+\sqrt{3}$ 的次數是 4。

4 **三等分角。**證明 $\cos 3\theta=4\cos^3\theta-3\cos\theta$，並由此推出 $\cos 20^\circ$ 是 $8x^3-6x-1=0$ 的根。假設這個三次方程沒有有理根（請自行驗證），解釋為何不能用尺規把 60° 的角三等分。

延伸閱讀

- **Ian Stewart, *Galois Theory***（CRC Press）。講述域擴張、塔定律及古典作圖問題，為大學本科生而寫。
- **Richard Courant & Herbert Robbins, *What Is Mathematics?***（Oxford University Press）。第 3 章證明古希臘幾大問題不可能解決，毋須任何預備知識。
- **Robin Hartshorne, *Geometry: Euclid and Beyond***（Springer）。以近世代數處理尺規作圖。

6.4 解四次方程

題目

STEP III 2017，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/17-s3-q3`

引子

每個學生都學過二次方程（quadratic equation）的求根公式。三次方程（cubic）和四次方程（quartic）其實也有公式；這些公式在十六世紀的意大利被發現，其間充滿祕密比試，以及為爭奪發現權而起的激烈爭執。卡爾達諾（Gerolamo Cardano）在 1545 年出版的《大術》（*Ars Magna*）中把兩者一併發表：三次方程的解法來自塔爾塔利亞（Niccolò Tartaglia），而塔爾塔利亞曾要他發誓保密；四次方程的解法則來自他自己的學生費拉里（Lodovico Ferrari）。此後近三百年，數學家一直苦苦尋找五次方程（quintic）的公式，卻始終徒勞無功。

這道題目展示的，正是解開四次方程的那個竅門，而這個竅門背後大有深意。從四個根出發，按三種可能的方式把根兩兩配對，得出三個新的量。這三個量滿足一個三次方程，而三次方程是你解得到的；有了它們，原來的四個根便隨之而來。爲甚麼這個方法行得通？為甚麼四個根可以化約為三個量，五個根卻不能如法炮製？正是這個問題，引領伽羅瓦（Galois）創立了羣論（group theory）。

動筆之前

想一想 把四件物件 $\alpha, \beta, \gamma, \delta$ 分成兩對，共有多少種分法？

三種：$\{\alpha\beta, \gamma\delta\}$、$\{\alpha\gamma, \beta\delta\}$ 和 $\{\alpha\delta, \beta\gamma\}$（圖 6.4.1）。題目中的三個量 $\alpha\beta + \gamma\delta$ 等等，正好對應這三種配對。把四個根任意重新排列，只會令這三種配對彼此對調，所以這三個量的任何對稱式（symmetric expression），都是四個根的對稱式，因而可以用 p、q、r、s 表示。這就是為甚麼它們滿足一個係數只涉及 p、q、r、s 的三次方程。

逐步解題

階段 1　欠缺的係數

想一想　在三次方程 $y^3 + Ay^2 + By + C = 0$ 中，根之和用 A 表示是甚麼？而 $(\alpha\beta + \gamma\delta) + (\alpha\gamma + \beta\delta) + (\alpha\delta + \beta\gamma)$ 用四次方程的係數表示又是甚麼？

三次方程的根之和是 $-A$。那三個量之和，就是 $\alpha, \beta, \gamma, \delta$ 中兩兩相乘所得全部六個積之和；對四次多項式 $x^4 + px^3 + qx^2 + rx + s$ 而言，這個和是 q。所以 $A = -q$。

階段 2　具體的三次方程

代入 $p = 0$、$q = 3$、$r = -6$、$s = 10$，三次方程是

$$y^3 - 3y^2 + (0 - 40)y + (120 - 0 - 36) = y^3 - 3y^2 - 40y + 84 = 0.$$

想一想　試試 84 的細小因數。

$y = 2$ 給出 $8 - 12 - 80 + 84 = 0$。作除法，得 $y^3 - 3y^2 - 40y + 84 = (y - 2)(y^2 - y - 42) = (y - 2)(y - 7)(y + 6)$。最大的根是 7，所以 $\alpha\beta + \gamma\delta = 7$。 ■

階段 3　根的配對

$(\alpha + \beta)(\gamma + \delta) = \alpha\gamma + \alpha\delta + \beta\gamma + \beta\delta$，這是全部六個兩兩之積的和減去 $\alpha\beta + \gamma\delta$，即 $q - 7 = 3 - 7 = -4$。

此外 $\alpha\beta \cdot \gamma\delta = \alpha\beta\gamma\delta = s = 10$。所以 $\alpha\beta$ 和 $\gamma\delta$ 之和為 7，積為 10：它們是 $t^2 - 7t + 10 = (t - 2)(t - 5)$ 的根。由於 $\alpha\beta > \gamma\delta$，得 $\alpha\beta = 5$ 及 $\gamma\delta = 2$。 ■

階段 4　解四次方程

停一停，試一試　求 $\alpha + \beta$ 和 $\gamma + \delta$。你已知道它們的和（由 p 得出）和它們的積（由第 3 步得出）；要判斷哪個是哪個，就用 r。

$\alpha + \beta + \gamma + \delta = -p = 0$ 而 $(\alpha + \beta)(\gamma + \delta) = -4$，所以 $\{\alpha + \beta, \gamma + \delta\} = \{2, -2\}$。係數 r 給出四個三根之積的和：$\alpha\beta\gamma + \alpha\beta\delta + \alpha\gamma\delta + \beta\gamma\delta = -r = 6$。這可以寫成 $\alpha\beta(\gamma + \delta) + \gamma\delta(\alpha + \beta) = 5(\gamma + \delta) + 2(\alpha + \beta)$。若 $\alpha + \beta = 2$，此式等於 $-10 + 4 = -6$；若 $\alpha + \beta = -2$，則等於 $10 - 4 = 6$。所以 $\alpha + \beta = -2$ 及 $\gamma + \delta = 2$。

於是 α、β 是 $t^2 + 2t + 5 = 0$ 的根，即 $-1 \pm 2\mathrm{i}$；而 γ、δ 是 $t^2 - 2t + 2 = 0$ 的根，即

$1 \pm \mathrm{i}$。四個根是

$$x = -1 \pm 2\mathrm{i}, \qquad x = 1 \pm \mathrm{i}. \qquad \blacksquare$$

驗算：$(x^2 + 2x + 5)(x^2 - 2x + 2) = x^4 + 3x^2 - 6x + 10$。

回顧

- **對稱函數（symmetric functions）。**每一步都用上了根與係數的關係：根之和與根之積。
- **降低次數。**一個四次方程的問題變成了一個三次方程（即*預解三次方程*（resolvent cubic)），然後再變成兩個二次方程。
- **用盡每一項資料。**p 和 q 給出了各個和與積；s 把 $\alpha\beta$ 和 $\gamma\delta$ 分開；r 則決定了哪一對根配哪一個和。

背後的數學

這個竅門為何行得通：根的置換

1770 年，拉格朗日（Joseph-Louis Lagrange）追問三次和四次方程的解法為何行得通，並在根的*置換*（permutation）中找到了答案。四個根共有 24 種重新排列的方式。在所有這些排列之下，三個量 $y_1 = \alpha\beta + \gamma\delta$、$y_2 = \alpha\gamma + \beta\delta$、$y_3 = \alpha\delta + \beta\gamma$ 只不過是彼此對調：四個根的每一種重新排列，都引致三種配對的一種重新排列（圖 6.4.1）。所以它們滿足一個 3 次多項式，其係數是各根的對稱式。用現代的語言來說，從對稱羣（symmetric group）S_4（全部 24 個置換）到 S_3（三種配對的 6 個置換）有一個滿射的羣同態（group homomorphism），這就是四次方程可以化約為三次方程的原因。

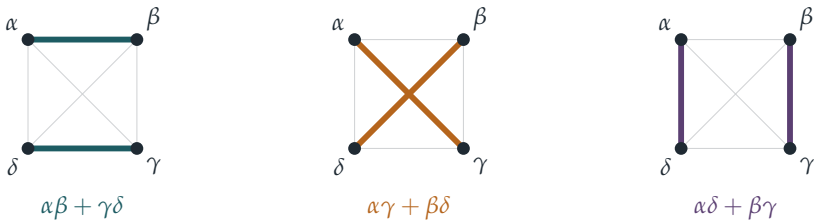


圖 6.4.1　把四個根分成兩對的三種方式，以及預解三次方程中與之對應的根。$\alpha, \beta, \gamma, \delta$ 的任何重新排列，都只會令這三幅圖彼此置換，這就是為甚麼那三個量滿足一個三次方程，而其係數可以用四次方程的係數表示。

五次方程

對於五個根，拉格朗日嘗試尋找一個類似的量，使它在根的置換之下只取寥寥幾個值，卻找不到有用的。1799 年魯菲尼（Paolo Ruffini）提出證明，1824 年阿貝爾（Niels Henrik Abel）把證明補全，兩人證明了：只用四則運算和根式（radicals），並不存在五次方程的一般求根公式。數年之後，伽羅瓦（Évariste Galois）確切地說明了哪些多項式可以用根式求解：就是那些根的對稱羣能以某種特定方式分解成較簡單部分的多項式。對稱羣 S_4 可以這樣分解（經由 S_3 和更小的羣），但 S_5 不可以，因為它包含一個大小為 60 的「單」（simple）部分，無法再進一步分解。1832 年，伽羅瓦在決鬥中喪生，年僅二十歲；決鬥前一晚，他才匆匆把自己的理論寫下來。這套理論如今是數學學位課程中的瑰寶之一，通常在三年級學到。

解三次方程

三次方程本身也可以用類似的竅門求解：把一個缺二次項的三次式（depressed cubic）$x^3 - 3px + q$ 寫成兩個完全立方的組合，從而化約為一個二次方程。這就是練習中那道 2011 年題目所用的方法，它與卡爾達諾公式（Cardano's formula）等價。當三次方程有三個實根時，這條公式在運算途中仍然需要用到複數（complex numbers），正是這個現象迫使數學家認真看待複數（見第 5 章的引言）。

思維技巧

> **對稱性化繁為簡**
>
> 如果由根構造出來的一些量，在根被置換時只是彼此置換，那麼它們滿足一個次數較低的多項式，而其係數是你可以計算出來的。
>
> 這是伽羅瓦理論（Galois theory）的種子：透過理解方程各根的對稱性來理解方程本身。

延伸練習

1. **驗證三次方程。**驗證 $y_1y_2 + y_2y_3 + y_3y_1 = pr - 4s$，其中 $y_1 = \alpha\beta + \gamma\delta$，餘類推。（展開並利用 $\sum\alpha = -p$、$\sum\alpha\beta\gamma = -r$、$\alpha\beta\gamma\delta = s$。）

2. **把三次式寫成立方之和。**試做 STEP III 2011，第 3 題（`step.maths.org/questions/11-s3-q3`）。這道題把三次式寫成兩個立方的組合來求解，與本節的路線截然不同。請用數值方法核對它的答案。

3 **另一個四次方程。**用本節的方法解 $x^4 - 4x^3 + 7x^2 - 6x + 2 = 0$。（預解三次方程有一個根是細小的整數。）

4 **數算置換。**在 $\alpha, \beta, \gamma, \delta$ 的 24 個置換中，找出所有把配對 $\{\alpha\beta, \gamma\delta\}$ 變回自身的置換。證明這樣的置換共有 8 個。（它們構成 S_4 的一個子羣（subgroup）。）

延伸閱讀

- **Mario Livio, *The Equation That Couldn't Be Solved***（Simon & Schuster）。五次方程的故事與羣論的誕生，適合一般讀者。
- **Ian Stewart, *Why Beauty Is Truth: A History of Symmetry***（Basic Books）。從巴比倫人的方程講到伽羅瓦，再到其後的發展。
- **Ian Stewart, *Galois Theory***（CRC Press）。大學本科教科書，其歷史導論所涵蓋的正是本節的內容。

6.5 從公理出發的微分法

題目

STEP II 1992，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/92-s2-q4`

引子

甚麼是微分法（differentiation）？在中學，它是用極限（limit）$\lim_{h\to 0}\frac{f(x+h)-f(x)}{h}$ 來定義的，然後你再學各種法則：和的導數（derivative）等於導數的和、積法則（product rule），諸如此類。這道題把整件事倒轉過來。請把極限完全拋開。假設你對某個運算所知的，就只有它符合和法則、積法則和常數的法則，並且把 x 變成 1。題目要證明的是：這樣的運算必定就是微分，至少對多項式（polynomial）而言是如此。

這是代數學家看微積分的角度：一個運算的特徵在於它所遵從的法則，而不在於某條公式。遵從積法則的運算稱為導子（derivation），它們在代數、幾何和物理中隨處可見。出人意料的是，訓練現代機器學習系統的軟件，內裏用的也是同一個想法。

動筆之前

想一想 法則 (i) 就是積法則。積法則怎可能告訴你關於常數的任何事情？（試把常數看成一個乘積。）

數 1 等於 1×1。把積法則應用於 1×1，便得出 $\Diamond 1$ 與它自身之間的一個關係，從而迫使它等於零。一般而言，處理這類公理（axiom）的竅門，是把它們應用於精心挑選的簡單輸入。

逐步解題

階段 1　常數

想一想　取 $\mathrm{f} = \mathrm{g} = 1$，應用法則 (i)。然後用法則 (iv) 處理其他任何常數。

$\Diamond 1 = \Diamond(1 \cdot 1) = 1 \cdot \Diamond 1 + 1 \cdot \Diamond 1 = 2\Diamond 1$，所以 $\Diamond 1 = 0$。對任意常數 λ，由法則 (iv) 得 $\Diamond \lambda = \Diamond(\lambda \cdot 1) = \lambda \Diamond 1 = 0$。 ■

階段 2　冪

利用法則 (iii)，$\Diamond x^2 = \Diamond(x \cdot x) = x\Diamond x + x\Diamond x = 2x$。於是

$$\Diamond x^3 = \Diamond(x^2 \cdot x) = x\Diamond x^2 + x^2 \Diamond x = x \cdot 2x + x^2 = 3x^2.$$

停一停，試一試　用歸納法（induction）證明：對每個正整數 n，均有 $\Diamond x^n = nx^{n-1}$。

$n = 1$ 的情況就是法則 (iii)。若 $\Diamond x^n = nx^{n-1}$，則

$$\Diamond x^{n+1} = \Diamond(x^n \cdot x) = x\Diamond x^n + x^n \Diamond x = x \cdot nx^{n-1} + x^n = (n+1)x^n. \qquad \blacksquare$$

階段 3　所有多項式

一般的多項式是 $\mathrm{h}(x) = a_0 + a_1 x + \cdots + a_m x^m$。由法則 (ii)（重複應用）和法則 (iv)，

$$\begin{aligned} \Diamond \mathrm{h}(x) &= \Diamond a_0 + a_1 \Diamond x + a_2 \Diamond x^2 + \cdots + a_m \Diamond x^m \\ &= a_1 + 2a_2 x + \cdots + m a_m x^{m-1} = \mathrm{h}'(x). \end{aligned} \qquad \blacksquare$$

回顧

- **公理決定運算。**四條法則，完全沒有提及極限，就把 $\Diamond$ 在多項式上的作用完全確定下來。
- **生成元（generator）。**每個多項式都是由常數和 x 經加法和乘法構造出來的。只要知道 $\Diamond$ 對這些基本構件的作用，以及它與這些構造運算如何互動，就能確定它在所有多項式上的作用。
- **沿着構造過程作歸納。**證明的步驟緊隨多項式的構造方式：先以重複相乘得出冪，然後再求和。

背後的數學

導子

在一個代數系統上，若運算 D 具有可加性（additive），並且滿足積法則 $D(fg) = fD(g) + gD(f)$，便稱為導子。這道題說明，在多項式上，一個把常數變成零的導子完全由 $D(x)$ 決定：事實上 $D(h) = h'(x)\,D(x)$，所以唯一的自由度就是多項式 $D(x)$ 的選擇。法則 (iii) 規定了 $D(x) = 1$。導子遍佈數學各個領域：在幾何中，曲面上的向量場（vector field）可以定義為該曲面上的函數的一個導子（它沿着向量場的方向對函數求導）；在代數中，一個結構的導子量度它的「無窮小對稱」（infinitesimal symmetries）；在物理中，量子力學的交換子（commutator）$[A, B] = AB - BA$ 對每個變元而言都是一個導子。

沒有極限的微積分

在極限完全沒有意義的場合，代數式的導數依然行得通。在有限域（finite field）上，例如模質數 p 的算術，並沒有極限可言，但多項式的形式導數（formal derivative）仍由同樣的法則定義，而且仍然能夠偵測重根（repeated root）：多項式 f 有重根，當且僅當 f 與 f' 有公因式。舉例說，我們就是這樣證明當 $p \nmid n$ 時 $x^n - 1$ 沒有重根的；要在有限域上定義 Section 6.1 的分圓多項式（cyclotomic polynomial），就需要這個事實。編碼理論和密碼學也用到同樣的想法。

自動微分

以下是一個現代應用。創造一個數 ε，使得 $\varepsilon^2 = 0$ 但 $\varepsilon \neq 0$，然後用對偶數（dual number）$a + b\varepsilon$ 來做算術。這時 $(a + b\varepsilon)(c + d\varepsilon) = ac + (ad + bc)\varepsilon$：$\varepsilon$ 的部分恰好遵從積法則。因此對任何多項式，

$$f(x + \varepsilon) = f(x) + f'(x)\,\varepsilon,$$

而在對偶數 $x + \varepsilon$ 處計算 f，就能同時求出函數值和導數，結果完全準確，既不用極限，也沒有因微小差值而產生的捨入誤差（圖 6.5.1）。這稱為*前向模式自動微分*（forward-mode automatic differentiation）。它的近親「反向模式」（reverse mode），就是用來訓練擁有數十億個參數的神經網絡（neural network）的算法（*反向傳播*，backpropagation）：這類系統每學習一次，都是在把積法則和鏈式法則（chain rule）當作導子，一步一步地應用。

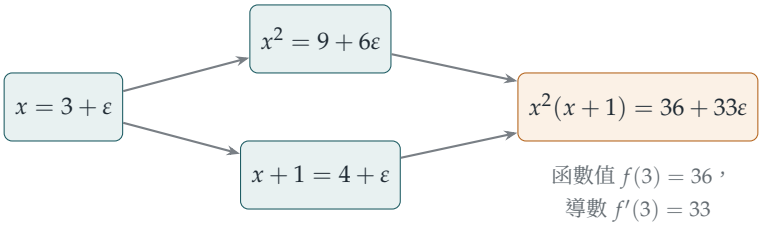


圖 6.5.1　用對偶數（$\varepsilon^2 = 0$）對 $f(x) = x^2(x+1)$ 在 $x = 3$ 處作自動微分。每一步都遵從題目中的法則：$(3+\varepsilon)^2 = 9 + 6\varepsilon$，以及 $(9+6\varepsilon)(4+\varepsilon) = 36 + (9+24)\varepsilon$。$\varepsilon$ 的係數就是 $f'(3) = 3\cdot 9 + 2\cdot 3 = 33$，計算結果完全準確。

思維技巧

以性質刻劃對象

用一個運算所遵從的法則來描述它，然後證明這些法則唯一地確定了它。

這就是公理化方法（axiomatic method），是大學代數的核心。它的威力在於：凡是由公理證明出來的結果，都適用於每一個滿足這些公理的運算，不論那是多項式的微分、曲面上函數的微分，還是某種尚未發明的東西。

延伸練習

1. **另一個導子。**假設 $\Diamond$ 滿足法則 (i)、(ii) 和 (iv)，但 $\Diamond x = x$ 而不是 1。證明 $\Diamond \mathrm{h}(x) = x\,\mathrm{h}'(x)$。$\Diamond$ 會把 x^n 變成甚麼？

2. **商法則（quotient rule）。**把 $\Diamond$ 推廣到有理函數（rational function）f/g，並假設法則 (i) 仍然成立。把法則 (i) 應用於 $\mathrm{g}\cdot\frac{\mathrm{f}}{\mathrm{g}} = \mathrm{f}$，從而推導出商法則。

3. **重根。**證明若 $(x-a)^2$ 整除多項式 f，則 $x-a$ 整除 f′。利用這個結果證明 $x^n - 1$ 沒有重根。

4. **對偶數。**設 $f(x) = (x^2+1)^3$，用對偶數計算 $f(2)$ 和 $f'(2)$，並用鏈式法則驗算。

延伸閱讀

- **Peter Cameron, *Introduction to Algebra***（Oxford University Press）。大學一年級程度的環、多項式和域，並包括形式導數。

- **Andreas Griewank & Andrea Walther, *Evaluating Derivatives***（SIAM）。自動微分的標準參考書。程度較深。
- **Michael Nielsen, *Neural Networks and Deep Learning***（網上免費閱讀）。第 2 章清楚講解反向傳播。

6.6 子羣的併集

題目

STEP II 1988，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/88-s2-q9`

引子

這是全書最短的題目之一，也是最具大學數學特色的題目之一。它談的是羣（group），即描述對稱的數學語言，問的問題聽起來像個謎題：一個羣能否被它裏面兩個較小的羣完全覆蓋？答案是不能，而證明只有三行。不過三個較小的羣卻可以做到，這倒是出人意表。

羣論（group theory）在 2019 年以前一直屬於 STEP 的課程範圍，如今則是大學一年級的必修內容。這道題只需要用到定義，下文會一一列出；它正好讓你初嘗羣論學者的思考方式：一個簡短而巧妙的論證，除了公理（axiom）之外甚麼都不用。

動筆之前

羣是一個集合 G，連同一個運算（寫成乘法 ab），滿足：

- 運算符合結合律（associative）：$(ab)c = a(bc)$；
- 存在單位元（identity）e，對每個 a 都有 $ea = ae = a$；
- 每個 a 都有逆元（inverse）a^{-1}，使得 $aa^{-1} = a^{-1}a = e$。

例子：整數（正數、負數和零）在加法下構成羣；非零實數在乘法下構成羣；正方形的旋轉在複合（composition）下構成羣。子羣（subgroup）是 G 的一個子集 K，它在同一運算下本身就是一個羣：它包含 e，而且若 $a, b \in K$，則 $ab \in K$ 及 $a^{-1} \in K$。

? **想一想** 若 K 是子羣，$a \in K$ 但 $b \notin K$，那麼 ab 可以屬於 K 嗎？

不可以。若 $ab \in K$，則 $b = a^{-1}(ab)$ 是 K 中兩個元素的乘積，因而屬於 K。這個小小的觀察就是整個證明。

逐步解題

階段 1　兩個子羣無法覆蓋

用反證法（proof by contradiction）：假設 $G = G_1 \cup G_2$，但 G_1 和 G_2 都不是整個 G。

> **想一想**　由於 $G_1 \neq G$，必有一個元素不在 G_1 之中。它必定在哪裏？G_2 的情況也一樣。現在把這兩個元素相乘。

取 $a \in G$ 使得 $a \notin G_1$；由於每個元素都屬於 G_1 或 G_2，故 $a \in G_2$。取 $b \in G$ 使得 $b \notin G_2$；於是 $b \in G_1$。考慮乘積 ab。它屬於 G_1 或 G_2。

- 若 $ab \in G_1$：由於 $b \in G_1$，$a = (ab)b^{-1} \in G_1$。矛盾。
- 若 $ab \in G_2$：由於 $a \in G_2$，$b = a^{-1}(ab) \in G_2$。矛盾。

無論哪種情況都得出矛盾，所以 $G_1 = G$ 或 $G_2 = G$。 ■

留意我們從未用過 G 是有限羣這個條件。這個結果對所有羣都成立。

階段 2　三個子羣卻可以

> **停一停，試一試**　找出一個羣，它是三個子羣的併集（union），而這三個子羣都不是整個羣。（提示：試試一個每個元素都是自己逆元的小羣。）

取克萊因四元羣（Klein four-group）$V = \{e, a, b, c\}$，其運算規則為 $a^2 = b^2 = c^2 = e$ 以及 $ab = ba = c$、$bc = cb = a$、$ca = ac = b$。（具體來說：一個非正方形的長方形的四個對稱，即保持不動、旋轉 $180°$，以及沿兩條對稱軸之一作反射。又或者是由位元 0、1 組成的數對 (x, y)，在模 2 加法下構成的羣。）它的三個子羣

$$H_1 = \{e, a\}, \qquad H_2 = \{e, b\}, \qquad H_3 = \{e, c\}$$

各有兩個元素，沒有一個是整個羣，而它們合起來包含了 V 的每一個元素（圖 6.6.1）。 ■

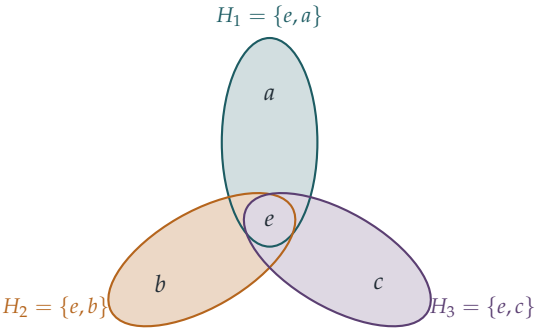


圖 6.6.1 克萊因四元羣 $\{e, a, b, c\}$ 是它三個 2 階子羣的併集，這三個子羣只在單位元 e 處重疊。根據題目的第一部分，沒有任何羣可以只是兩個真子羣（proper subgroup）的併集。

回顧

- **選出見證元素，再把它們結合。**證明中在每個子羣之外各選一個元素，然後把它們相乘。這種「各取一個再結合」的手法在代數中極為常見。
- **封閉性威力強大。**證明中只用到子羣的一個性質：它們在乘法和取逆元之下封閉（closed）。
- **檢查你真正需要的假設。**題目給了有限這個條件，卻沒有用上。察覺到某個假設並非必要，是理解一個證明的一部分。

背後的數學

對稱的語言

羣在十九世紀同時從幾個方向出現：伽羅瓦（Galois）對多項式的根的置換（permutation）的研究（Section 6.4）、柯西（Cauchy）關於置換的工作，以及幾何學和晶體學中對對稱的研究。上文那個抽象定義，即凡滿足三條公理的東西都是羣，到該世紀末已經確立。它的威力在於一般性：由公理證明的定理，例如本題中的那一個，同時適用於由數、矩陣、對稱和置換構成的各種羣。今天羣無處不在：在粒子物理學中（標準模型建基於對稱羣）、在化學中（分子的對稱性決定光譜）、在密碼學中（橢圓曲線羣，Section 1.9），以及在扭計骰（Rubik's cube）的解法之中。

要用多少個子羣？

這道題引出一個自然的問題：最少要多少個真子羣，它們的併集才是整個羣？根據本題，這個數目永遠不會是 2。1926 年，Gaetano Scorza 證明了一個羣是三個真子羣的併集，當且僅當它以克萊因四元羣為「商羣」（quotient）。1994 年，J. H. E. Cohn 研究了一般的問題，並提出了若干猜想；1997 年，M. J. Tomkinson 證明了沒有任何羣的這個最小數目是七，儘管由 3 到 6 的每個數都會出現。這類問題敘述簡單，卻出奇地微妙，正是羣論的典型。

向量空間

類似的論證可以證明，實數上的向量空間（vector space）（例如平面或三維空間）永遠不可能是有限多個真子空間（proper subspace）的併集：平面不是有限多條通過原點的直線的併集。但在有限域（finite field）上卻可以：兩個元素的域上的平面，其實又是克萊因四元羣，它正是其三條直線的併集。

思維技巧

從公理出發論證

在抽象代數中，你對所處理的對象往往一無所知，只知道它們滿足哪些公理。證明必須只用這些公理。

這迫使論證簡短而乾淨利落，就像本題的論證一樣，也使它們放諸四海皆準。這需要一些時間去適應，而它正是大學一年級代數課中最獨特的體驗之一。

延伸練習

1. **驗證這個羣。**寫出克萊因四元羣的乘法表，並驗證各條羣公理。證明它與乘法下的羣 $\{1, \mathrm{i}, -1, -\mathrm{i}\}$ 並不相同。（提示：哪些元素是自己的逆元？）

2. **循環羣。**若一個羣的每個元素都是某一個元素 g 的冪，則稱它為循環羣（cyclic group）。證明循環羣永遠不是真子羣的併集，不論用多少個。（提示：哪個子羣包含 g？）

3. **整數。**直接證明 $(\mathbb{Z}, +)$ 不是子羣 $2\mathbb{Z}$（偶數）和 $3\mathbb{Z}$（3 的倍數）的併集。它是否某兩個真子羣的併集？

4 **平面上的直線。**較難證明平面 $\mathbb{R}^2$ 不是有限多條通過原點的直線的併集。（提示：每條直線除原點外，最多只包含拋物線 $y = x^2$ 上的一個點。）

延伸閱讀

- **M. A. Armstrong, *Groups and Symmetry***（Springer）。一本插圖精美的入門教材，以幾何例子帶動全書。
- **Nathan Carter, *Visual Group Theory***（Mathematical Association of America）。用圖像講解羣，最適合在入讀大學前閱讀。
- **Peter Neumann, Gabrielle Stoy & Edward Thompson, *Groups and Geometry***（Oxford University Press）。牛津大學一年級的經典教材。

6.7 四元數羣

題目

STEP III 1993，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/93-s3-q5`

引子

1843 年 10 月 16 日，哈密頓（William Rowan Hamilton）與妻子沿着都柏林的皇家運河散步時，苦思多年的問題忽然靈光一閃，有了答案。他一直想仿照複數把一對對的數相乘那樣，把三個一組的數相乘，卻屢試屢敗。他突然看出，答案需要的是四個數，而且乘法的次序不可調換：$ij \neq ji$。他把這些法則刻在布魯穆橋（Brougham Bridge）的石頭上：

$$\mathrm{i}^2 = \mathrm{j}^2 = \mathrm{k}^2 = \mathrm{ijk} = -1.$$

這就是四元數（quaternions）。本題用複數對把四元數構造出來，並在其中找到一個只有八個元素、乘法帶點扭曲的小羣。時至今日，每一款電腦遊戲、每一部飛行模擬器、每一套太空船導航系統，都用四元數來描述三維空間中的旋轉。

動筆之前

羣的各項定義見 Section 6.6。這裏再補充兩個用語：羣中元素 g 的階（order），是指使 $g^n = e$ 成立的最小正整數 n；若兩個羣之間存在一個保持運算的一一對應，則稱這兩個羣同構（isomorphic），也就是說，它們是「換了標籤的同一個羣」。

想一想 把數對 (z_1, z_2) 看成代表一個「數」$z_1 + z_2\mathrm{j}$，其中 j 是一個新符號，滿足 $\mathrm{j}^2 = -1$，並且對任意複數 w 有 $\mathrm{j}w = w^*\mathrm{j}$。試用這些法則計算 $(z_1 + z_2\mathrm{j})(w_1 + w_2\mathrm{j})$。得出的是否就是 $\circ$ 的公式？

$(z_1 + z_2\mathrm{j})(w_1 + w_2\mathrm{j}) = z_1w_1 + z_1w_2\mathrm{j} + z_2\mathrm{j}w_1 + z_2\mathrm{j}w_2\mathrm{j} = z_1w_1 + z_1w_2\mathrm{j} + z_2w_1^*\mathrm{j} + z_2w_2^*\mathrm{j}^2$，即 $(z_1w_1 - z_2w_2^*) + (z_1w_2 + z_2w_1^*)\mathrm{j}$。答案是肯定的：運算 $\circ$ 其實就是喬裝了的四元數乘

法。

逐步解題

階段 1　借助矩陣證明結合律

直接驗證 $(a \circ b) \circ c = a \circ (b \circ c)$，計算十分冗長。其實有更巧妙的方法。

> **想一想**　把 (z_1, z_2) 對應到複矩陣 $M = \begin{pmatrix} z_1 & z_2 \\ -z_2^* & z_1^* \end{pmatrix}$。矩陣乘積 $M(z_1,z_2)\,M(w_1,w_2)$ 的第一行是甚麼？

乘積的第一行是 $(z_1 w_1 - z_2 w_2^*,\ z_1 w_2 + z_2 w_1^*)$，正好是 $(z_1, z_2) \circ (w_1, w_2)$；第二行是 $(-(z_1 w_2 + z_2 w_1^*)^*,\ (z_1 w_1 - z_2 w_2^*)^*)$，所以乘積仍具有同樣的特殊形式。因此 $M(a \circ b) = M(a) M(b)$。矩陣乘法滿足結合律（associativity），而 M 是一對一的（從第一行便可讀出 z_1、z_2），所以 $\circ$ 滿足結合律。 ■

階段 2　不滿足交換律

$(0,1) \circ (\mathrm{i}, 0) = (0 \cdot \mathrm{i} - 1 \cdot 0,\ 0 \cdot 0 + 1 \cdot \mathrm{i}^*) = (0, -\mathrm{i})$，但 $(\mathrm{i}, 0) \circ (0, 1) = (0, \mathrm{i})$。所以 $\circ$ 不滿足交換律（commutativity）。

四個特殊乘積可由定義直接得出：

$$(z,0) \circ (w,0) = (zw, 0), \qquad (z,0) \circ (0,w) = (0, zw),$$
$$(0,z) \circ (w,0) = (0, zw^*), \qquad (0,z) \circ (0,w) = (-zw^*, 0).$$

階段 3　八個元素的羣

用四元數的記法，$A = 1$，$C = \mathrm{i}$，$B = \mathrm{j}$，$D = \mathrm{ij} = \mathrm{k}$，而 E, G, F, H 則分別是它們的相反數。S_1 的每個元素都具有 $(u, 0)$ 或 $(0, u)$ 的形式，其中 $u \in \{1, \mathrm{i}, -1, -\mathrm{i}\}$。

> **想一想**　利用四個特殊乘積，證明 S_1 是封閉的。單位元是甚麼？

由四個特殊乘積可知，兩個形如 $(u, 0)$ 或 $(0, u)$ 的元素相乘，結果仍是這兩種形式之一，其第一個或第二個分量為 uv、uv^* 或 $-uv^*$。由於 $\{1, \mathrm{i}, -1, -\mathrm{i}\}$ 在乘法、取共軛及取相反數之下都封閉（closed），乘積仍在 S_1 之內。單位元（identity）是 $A = (1, 0)$：$(1,0) \circ (w_1, w_2) = (w_1, w_2)$，而且 $(w_1, w_2) \circ (1, 0) = (w_1, w_2)$。

> **停一停，試一試**　求出每個元素的階和逆元。

A 的階為 1。$E \circ E = (-1,0) \circ (-1,0) = (1,0)$，所以 E 的階為 2。對於 $C = (\mathrm{i},0)$：$C \circ C = (-1,0) = E$，所以 C 的階為 4，而 $C^{-1} = C^3 = E \circ C = G$。對於 $B = (0,1)$：$B \circ B = (-1 \cdot 1^*, 0) = (-1,0) = E$，所以 B 的階為 4，而 $B^{-1} = F$。同理，$D \circ D = (-\mathrm{i} \cdot \mathrm{i}^*, 0) = (-\mathrm{i} \cdot (-\mathrm{i}), 0) = (-1,0) = E$，所以 D 的階為 4，逆元（inverse）為 H。至於 G、F、H，它們分別是 C、B、D 的逆元，階同樣為 4。總結如下：

元素	A	B	C	D	E	F	G	H
逆元	A	F	G	H	E	B	C	D
階	1	4	4	4	2	4	4	4

既有封閉性和結合律（第 1 步），又有單位元和逆元，S_1 是一個羣。 ■這個羣稱為四元數羣（quaternion group）Q_8。

階段 4 一個子羣

$\{A, B, E, F\} = \{1, \mathrm{j}, -1, -\mathrm{j}\}$ 由 B 的各次冪組成：$B^1 = B$，$B^2 = E$，$B^3 = F$，$B^4 = A$。所以它是封閉的，並且包含各元素的逆元：這是一個子羣（subgroup），而且是階為 4 的循環羣（cyclic group）。矩陣 $J = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$ 滿足 $J^2 = -I$，$J^3 = -J$，$J^4 = I$，所以它同樣生成一個階為 4 的循環羣。對應 $B^k \leftrightarrow J^k$ 是一個同構。 ■（事實上，J 正是 $M(0,1)$，即 B 所對應的矩陣。）

回顧

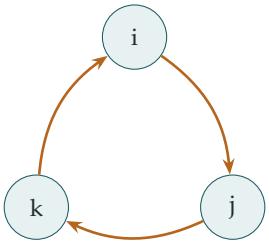
- **用矩陣來表示。**結合律直接驗證起來十分費勁，這裏卻可從矩陣乘法繼承過來。
- **找出公式背後的結構。**古怪的運算。原來就是四元數乘法，看穿了這一點，一切都變得容易。
- **階和逆元刻劃了一個羣。**單看各元素的階，已可見 Q_8 與別不同：一個階為 2 的元素，六個階為 4 的元素。

背後的數學

複數之外

實數構成一個一維的數系，在其中可以做加法、乘法和除法；複數則構成一個二維的數系。哈密頓的四元數是四維的，除法依然可行，但乘法不再滿足交換律（圖 6.7.1）。1878 年，弗羅貝尼烏斯（Ferdinand Frobenius）證明了 $\mathbb{R}$、$\mathbb{C}$ 和 $\mathbb{H}$（取自 Hamilton 的首字母）這三者，是實數域上僅有的、乘法滿足結合律的有限維「可除代數」（division

algebras）。如果放棄結合律，但保留一個可乘的大小，即 $|xy| = |x|\,|y|$，一如實數、複數和四元數那樣，那就恰好還多一個：八維的八元數（octonions），由 John Graves 和凱萊（Arthur Cayley）在 1840 年代發現。赫維茲（Adolf Hurwitz）於 1898 年證明，這類數系就只有這四個。（若不設關於大小的條件，還有其他更古怪的非結合可除代數，但 1958 年一條由博特（Raoul Bott）、米爾諾（John Milnor）和 Michel Kervaire 得出的深刻定理指出，任何有限維實可除代數的維數必為 1、2、4 或 8。）



順着箭頭：$\mathrm{ij} = \mathrm{k}$，$\mathrm{jk} = \mathrm{i}$，$\mathrm{ki} = \mathrm{j}$

逆着箭頭：$\mathrm{ji} = -\mathrm{k}$，$\mathrm{kj} = -\mathrm{i}$，$\mathrm{ik} = -\mathrm{j}$

以及 $\mathrm{i}^2 = \mathrm{j}^2 = \mathrm{k}^2 = -1$

圖 6.7.1　哈密頓的乘法法則。把 i、j、k 其中兩個順着箭頭方向相乘，得到第三個；逆着箭頭方向相乘，則得到第三個的相反數。$\pm 1, \pm\mathrm{i}, \pm\mathrm{j}, \pm\mathrm{k}$ 這八個元素構成本題的四元數羣 Q_8。

三維空間中的旋轉

模（modulus）為 1 的四元數描述三維空間的旋轉：繞單位軸 $\mathbf{u}$ 轉過角度 θ 的旋轉，對應於 $q = \cos\frac{\theta}{2} + \sin\frac{\theta}{2}(u_1\mathrm{i} + u_2\mathrm{j} + u_3\mathrm{k})$，它作用於向量 $\mathbf{v}$（寫成 $v_1\mathrm{i} + v_2\mathrm{j} + v_3\mathrm{k}$）的方式是 $\mathbf{v} \mapsto q\mathbf{v}q^{-1}$。旋轉的複合就是四元數的相乘。這比使用矩陣或角度更有效率，數值上也更穩定，而且避開了「萬向節鎖」（gimbal lock）：那是失去一個自由度的現象，當年阿波羅（Apollo）太空人使用以陀螺儀為本的導航系統時，就不得不設法繞過這個問題。半角 $\frac{\theta}{2}$ 也透露出一點深刻的端倪：旋轉 360° 得到的是 $q = -1$，而不是 1，物理學家在電子的行為（自旋，spin）中正好看到這個事實。

四平方和

四元數的模是可乘的，$|pq| = |p||q|$，與複數的模一樣。把它展開寫出，便得到歐拉（Euler）的*四平方恆等式*（four-square identity）：兩個四平方和之積，本身也是一個四平方和。這正是拉格朗日（Lagrange）定理的關鍵一步，該定理指出每個正整數都是四個平方數之和；我們會在 Section 7.8 證明這條定理，它是 Section 1.6 中二平方和結果的四維近親。

思維技巧

把抽象的對象具體地表示出來

如果一個抽象的運算難以推理，就找一個具體的模型（例如矩陣），讓它在其中變成熟悉的運算。

這樣，結合律之類的性質便不費吹灰之力。這個想法，即把羣表示為矩陣來研究，後來發展成一門龐大而優美的學科：*表示論*（representation theory）。

延伸練習

1 **哈密頓的法則。**利用數對，驗證 $C \circ B = D$ 及 $B \circ C = H$，即 $\mathrm{ij} = \mathrm{k}$ 及 $\mathrm{ji} = -\mathrm{k}$。

2 **只有一個階為 2 的元素。**證明 $E = -1$ 是 Q_8 中唯一階為 2 的元素。由此推斷 Q_8 與正方形的對稱羣（它有五個階為 2 的元素）並不同構。

3 **四平方和。**證明 $\det M(z_1, z_2) = |z_1|^2 + |z_2|^2$。利用 $\det(M_1 M_2) = \det M_1 \det M_2$，推斷兩個四平方和之積是一個四平方和。

4 **子羣。**找出 Q_8 的所有子羣。證明除了 $\{A\}$ 之外，每個子羣都包含 E，並證明 Q_8 是它三個階為 4 的循環子羣的併集（試與 Section 6.6 比較）。

延伸閱讀

- **John Conway & Derek Smith, *On Quaternions and Octonions***（A. K. Peters）。講述四元數和八元數的幾何與算術。程度較深，但引人入勝。
- **John Baez,「The octonions」**，*Bulletin of the American Mathematical Society* 39 (2002)；網上亦可免費閱覽。一篇精彩的綜述；引言部分現在就讀得懂。
- **Andrew Hanson, *Visualizing Quaternions***（Morgan Kaufmann）。講述以四元數處理旋轉，並介紹其在電腦圖像方面的應用。

6.8 由奇異矩陣組成的羣

題目

STEP III 1996，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/96-s3-q6`

引子

每個學生都學過：行列式（determinant）為零的矩陣沒有逆矩陣。那麼，一個由矩陣組成的羣（group），既然其中每個元素都有逆元（inverse），就肯定不可能包含任何奇異矩陣（singular matrix）吧？

其實可以。關鍵在於「逆」這個字。在羣裏，$\mathbf{A}$ 的逆元是指與 $\mathbf{A}$ 結合後得出該羣的單位元（identity element）的那個元素，而沒有任何規定說羣的單位元必須是單位矩陣（identity matrix）$\mathbf{I}$。本題構造出一個羣，其中每個矩陣都是奇異的，而單位元本身就是一個行列式為零的矩陣。這一課教我們要按字面去讀定義，也教我們分清抽象結構與我們最初認識它時所見的那個熟悉例子。

動筆之前

羣的定義見 Section 6.6。

想一想 在第 (i) 部分，單位矩陣 $\mathbf{I}$ 並不是 $\begin{pmatrix} a & a \\ a & a \end{pmatrix}$ 這種形式。那麼 S 的單位元會是甚麼？

它必須是 S 中的某個矩陣 $\mathbf{E}$，使得對 S 中每個 $\mathbf{X}$ 都有 $\mathbf{EX} = \mathbf{XE} = \mathbf{X}$。這個條件只涉及 $\mathbf{E}$ 怎樣作用於這種特殊形式的矩陣，而不是所有矩陣。解方程就可以把它找出來。

逐步解題

階段 1　封閉性（closure）

以 $\mathbf{M}_a$ 表示四個項全都等於 a 的矩陣。那麼

$$\mathbf{M}_a\mathbf{M}_b = \begin{pmatrix} a & a \\ a & a \end{pmatrix}\begin{pmatrix} b & b \\ b & b \end{pmatrix} = \begin{pmatrix} 2ab & 2ab \\ 2ab & 2ab \end{pmatrix} = \mathbf{M}_{2ab},$$

而當 $a, b \neq 0$ 時，$2ab \neq 0$。所以 S 是封閉的。 ■

階段 2　羣公理（group axioms）

> **想一想**　矩陣乘法滿足結合律。至於單位元，解 $\mathbf{M}_e\mathbf{M}_a = \mathbf{M}_a$。至於逆元，解 $\mathbf{M}_a\mathbf{M}_b = \mathbf{M}_e$。

結合律（*associativity*）由矩陣乘法繼承而來。單位元：$\mathbf{M}_e\mathbf{M}_a = \mathbf{M}_{2ea} = \mathbf{M}_a$ 對所有 a 成立，當且僅當 $e = \tfrac{1}{2}$；由於公式是對稱的，$\mathbf{M}_a\mathbf{M}_{1/2} = \mathbf{M}_a$ 同樣成立。所以

$$\mathbf{E} = \begin{pmatrix} \tfrac{1}{2} & \tfrac{1}{2} \\ \tfrac{1}{2} & \tfrac{1}{2} \end{pmatrix}$$

是 S 的單位元。逆元：當 $b = \tfrac{1}{4a}$ 時，$\mathbf{M}_a\mathbf{M}_b = \mathbf{M}_{2ab} = \mathbf{M}_{1/2}$，而這個值不為零，所以 $\mathbf{M}_{1/(4a)}$ 是 $\mathbf{M}_a$ 在 S 中的逆元。因此 S 是一個羣。■每個元素的行列式都是 $a^2 - a^2 = 0$。

階段 3　奇異性會蔓延

> **停一停，試一試**　假設 $\mathbf{A} \in G$ 是奇異的，並設 $\mathbf{B}$ 為 G 的任意元素。把 $\mathbf{B}$ 寫成一個含有 $\mathbf{A}$ 的乘積，然後取行列式。

以 $\mathbf{A}^{-1}$ 表示 $\mathbf{A}$ 在羣 G 中的逆元（不一定是逆矩陣）。那麼 $\mathbf{C} = \mathbf{A}^{-1}\mathbf{B}$ 屬於 G，而且

$$\mathbf{AC} = \mathbf{A}\mathbf{A}^{-1}\mathbf{B} = \mathbf{EB} = \mathbf{B}.$$

取行列式，得 $\det\mathbf{B} = \det\mathbf{A}\,\det\mathbf{C} = 0$。所以 G 的每個元素都是奇異的。 ■特別地，如果一個矩陣羣中有一個元素作為矩陣是可逆（invertible）的，那麼所有元素都可逆；一個矩陣羣要麼全部可逆，要麼全部奇異。

階段 4　一個 3×3 的例子

想一想 你能把第 (i) 部分推廣到 3×3 矩陣嗎？或者找一個更簡單的例子，用一個「無視」三個坐標中其中兩個的矩陣？

取

$$G = \left\{ \begin{pmatrix} a & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} : a \neq 0 \right\}.$$

角上的項分別為 a 和 b 的兩個矩陣，乘積就是角上的項為 ab 的矩陣，所以 G 是封閉的；乘法滿足結合律；單位元是角上的項為 1 的矩陣；角上的項為 a 的矩陣，其逆元是角上的項為 $\frac{1}{a}$ 的矩陣。所以 G 是一個羣，而且每個元素的行列式都是 0。 ■（另一個例子：九個項全都等於 $a \neq 0$ 的 3×3 矩陣，其單位元的所有項都是 $\frac{1}{3}$。）

回顧

- **按字面去讀定義。**羣的單位元就是*在該羣之內*起單位元作用的那個元素；它不一定是我們熟悉的那個單位矩陣。
- **解方程求出結構。**S 的單位元和逆元都是靠解方程找出來的，而不是靠猜。
- **行列式具有可乘性。**單憑這一個性質，就說明了奇異性會由一個元素蔓延到所有元素。

背後的數學

冪等元與隱藏的可逆性

任何矩陣羣的單位元 $\mathbf{E}$ 都滿足 $\mathbf{E}^2 = \mathbf{E}$：它是冪等（*idempotent*）的。從幾何上看，冪等矩陣就是一個投影（*projection*）：它把每個向量映射到某個子空間（subspace）上，而原本已在該子空間內的向量則保持不變。在第 (i) 部分，$\mathbf{E}$ 把平面投影到直線 $y = x$ 上，而每個矩陣 $\mathbf{M}_a$ 都是先投影到這條直線上，然後沿着它以因子 $2a$ 拉伸（圖 6.8.1）。在這條直線上，這個羣的作用是可逆的拉伸；它只是無視了與直線垂直的方向。所以 S 其實就是非零實數的乘法羣（透過 $\mathbf{M}_a \mapsto 2a$），只不過它棲身於平面內的一條直線上。這個結論普遍成立：每個由矩陣組成的羣，經過換基（change of basis）之後，都是一個由可逆矩陣組成的羣，作用於其單位元的像上，其餘位置則以零填充。

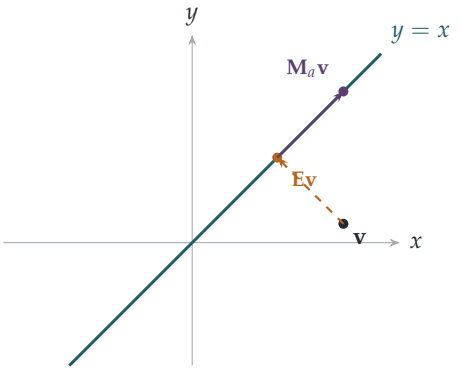


圖 6.8.1　第 (i) 部分的矩陣所起的作用。單位元 $\mathbf{E}$ 把每個向量 $\mathbf{v}$ 投影到直線 $y = x$ 上（琥珀色虛線）。每個 $\mathbf{M}_a$ 都先作同樣的投影，然後沿着直線以因子 $2a$ 拉伸（紫紅色；圖中 $2a = \frac{16}{9}$）。在這條直線本身之上，這個羣的作用就是普通的可逆縮放。

抽象與具體

人們最早研究的羣都是具體的：置換 (permutation)、對稱、可逆矩陣。抽象的定義 (一個集合連同一個滿足三條公理的運算) 是從這些例子中提煉出來的。本題顯示了抽象觀點的價值：S 與非零實數的乘法羣是同構（*isomorphic*）的，即同一個羣換了一身衣服，儘管一個由奇異矩陣組成，另一個由數組成。辨認出兩個外表不同的結構何時「相同」，是代數最核心的技巧之一，而數學家正是藉同構這個概念把這種「相同」說得精確。

思維技巧

質疑對熟悉對象的隱藏假設

當一個熟悉的對象出現在抽象定義中時，要查清楚定義實際上要求它具備哪些我們熟悉的性質。

「矩陣羣的單位元是 $\mathbf{I}$」聽起來理所當然，卻不是定義的一部分；放下這個假設，就得出了一個真正嶄新的例子。許多新的數學結構，正是這樣靠質疑假設而發現的。

延伸練習

1 **一個同構。**證明映射 $\mathbf{M}_a \mapsto 2a$ 是由 S 到非零實數的一一對應（one-to-one correspondence），並且把矩陣乘法變成數的乘法。

2 **不是羣。**證明所有形如 $\begin{pmatrix} a & b \\ 0 & 0 \end{pmatrix}$ 且 $a \neq 0$ 的矩陣所組成的集合在乘法下封閉，但不是一個羣。（提示：試找單位元。有沒有左單位元？有沒有右單位元？）

3 **冪等元。**求所有滿足 $\mathbf{E}^2 = \mathbf{E}$ 的 2×2 實矩陣 $\mathbf{E}$，並從幾何角度描述它們。

4 **所有項相等。**證明所有項都等於 $a \neq 0$ 的 $n \times n$ 矩陣在乘法下構成一個羣，並求它的單位元。

延伸閱讀

- **M. A. Armstrong, *Groups and Symmetry***（Springer）。講述矩陣羣與同構，並附有大量幾何例子。
- **Gilbert Strang, *Introduction to Linear Algebra***（Wellesley-Cambridge）。對投影與冪等矩陣有清晰的講解。
- **Joseph Gallian, *Contemporary Abstract Algebra***（CRC Press）。一本平易近人的入門教材，載有許多不尋常的羣的例子。

我們學到了甚麼

本章開首，我們談到代數的重心由解方程轉移到研究結構，並提出了三個問題：甚麼可以由甚麼構造出來？方程的公式為何行得通？抽象化帶來了甚麼好處？以下是本章八道題目所揭示的答案。

1. 甚麼可以由甚麼構造出來？

在 Section 6.3 中，我們逐一消去根式（radical），找出 $\sqrt{2}+\sqrt{3}$、$\sqrt{2}+\sqrt{3}+\sqrt{5}$ 和 $\sqrt[3]{2}+\sqrt[3]{3}$ 所滿足的多項式，次數分別為 4、8 和 9。這些次數是不變量（invariant）：它們就是該數所生成的域（field）的次數。由於次數沿着域塔（tower of fields）逐層相乘，凡是可以用平方根得到的數，其次數必為 2 的冪。單憑這一點，便足以證明倍立方不可能做到（$\sqrt[3]{2}$ 的次數是 3），也證明 $60°$ 角不可能三等分。

反過來，Section 6.2 展示了某些東西可以怎樣構造出來。把 17 次單位根（roots of unity）分拆成高斯（Gauss）週期（period）之後，和 $\cos\frac{2\pi}{17}+\cos\frac{4\pi}{17}+\cos\frac{8\pi}{17}+\cos\frac{16\pi}{17}$ 滿足一個二次方程，其根為 $\frac{-1+\sqrt{17}}{4}$；再解三個二次方程便可得到 $\cos\frac{2\pi}{17}$，從而作出正 17 邊形。這兩節的背後，都是 Section 6.1 的分圓多項式（cyclotomic polynomial）：它們是 x^n-1 的基本構件，把單位根按階（order）分類便可得到。

2. 方程的公式為何行得通？

Section 6.4 揭示了四次方程（quartic）的祕密：由四個根組成三個量 $\alpha\beta+\gamma\delta$、$\alpha\gamma+\beta\delta$、$\alpha\delta+\beta\gamma$，每個量對應把四個根兩兩配對的一種方法。無論怎樣重新排列這些根，都只會把這三個量互相調換，所以它們滿足一個係數已知的三次方程。先解這個三次方程，再解兩個二次方程，便得出四個根。拉格朗日（Lagrange）看出，這個方法之所以行得通，在於四個物件的置換（permutation）如何作用於那三種配對；伽羅瓦（Galois）則看出，這個方法對五個根失效，是因為五個物件的置換羣（group）之中，有一部分無法再分解。這些公式，其實是羣結構的影子。

3. 抽象化帶來了甚麼好處？

最後四節都以物件所遵守的規則來描述物件。

- 在 Section 6.5 中，四條規則（可加性、乘積法則、常數、$\Diamond x=1$）迫使一個運算必然是求導（differentiation），完全不涉及極限。同樣的規則，正是現代軟件中自動微分

（automatic differentiation）的基礎。

- 在 Section 6.6 中，由羣公理（axiom）出發的一個三行論證，一舉對所有羣證明了：沒有任何羣是兩個真子羣（proper subgroup）的併集。
- 在 Section 6.7 中，一個定義在複數對上的古怪運算，原來就是哈密頓（Hamilton）的四元數（quaternion）乘法；用矩陣來表示它，結合律（associativity）便自動成立。
- 在 Section 6.8 中，按字面理解羣的定義，便得出一個由奇異矩陣（singular matrix）組成的羣，它的單位元（identity）並不是單位矩陣，而它「其實」是非零實數羣的化身。

在每一個例子中，抽象的觀點都解釋了一些單靠計算會被掩蓋的東西。

4. 代數工具箱

方法	在哪裏用過
按唯一的不變量分類	按階把單位根分類（6.1）
由和與積得出二次方程	高斯週期（6.2）；四次方程（6.4）
積化和差公式	正 17 邊形（6.2）
消去根式	最小多項式（minimal polynomial）（6.3）
次數與塔定律（tower law）	作圖的不可能性（6.3）
根的對稱函數（symmetric function）	三次預解式（resolvent cubic）（6.4）
由公理確定一個運算	導子（derivation）（6.5）
由羣公理出發論證	子羣的併集（6.6）
矩陣表示（matrix representation）	四元數（6.7）
同構（isomorphism）	奇異矩陣羣（6.8）

下一步

- 四元數通往 Section 7.8 的拉格朗日四平方和定理（four-square theorem）。
- 把矩陣視為自有一套規則的代數物件，是第 7 章的主題。
- 像高斯整數（Gaussian integers）這類數系的算術，當中因數分解（factorisation）可能失效，是 Section 1.7 的主題。

如果你想在上大學之前再走遠一點

- **Mario Livio, *The Equation That Couldn't Be Solved***：五次方程、伽羅瓦與羣論的誕生。

- **Nathan Carter, *Visual Group Theory***：用圖像認識羣，是初次接觸的理想讀物。
- **M. A. Armstrong, *Groups and Symmetry***：建基於幾何的大學入門課程。
- **Ian Stewart, *Galois Theory***：留待你準備好了解整個故事的時候。

伽羅瓦在一場決鬥之後去世，年僅二十歲，只留下數十頁手稿，數學界花了幾十年才讀懂。他看出，一個方程最深層的祕密並不在它的係數之中，而在它的解的對稱性之中。這個洞見——給不同的東西起同一個名字，並研究它們共有的結構——如今貫穿整個數學。你已經在多項式、多邊形、四元數和矩陣中見過它發揮作用；到了大學，你幾乎處處都會遇見它。

延伸練習答案

以下只是簡短的答案和提示，並非完整解答：這裏只列出結果和關鍵步驟，方便你核對自己的做法。請先認真做過每一道練習，然後才看答案。

6.1 以單位根為根的多項式

1. $(x-1)(x+1) = x^2-1$ 而 $(x^2+x+1)(x^2-x+1) = x^4+x^2+1$，兩者之積為 x^6-1。同理，$x^{12}-1 = (x^6-1)(x^6+1)$，而 $x^6+1 = (x^2+1)(x^4-x^2+1) = \mathrm{C}_4\mathrm{C}_{12}$。

2. $2p$ 的因數（divisor）為 $1、2、p、2p$，所以 $\mathrm{C}_{2p} = \frac{x^{2p}-1}{(x-1)(x+1)\mathrm{C}_p(x)} = \frac{x^p+1}{x+1}$。另一方面，$\mathrm{C}_p(-x) = \frac{(-x)^p-1}{-x-1} = \frac{x^p+1}{x+1}$，因為 p 是奇數。當 $p=3$ 時：$\mathrm{C}_6 = x^2-x+1 = \mathrm{C}_3(-x)$。

3. 不超過 p^k 而又與 p^k 有公因數的數，就是那 p^{k-1} 個 p 的倍數。本原（primitive）p^k 次單位根是 $x^{p^k}-1$ 的根之中不屬於 $x^{p^{k-1}}-1$ 的根的那些，所以 $\mathrm{C}_{p^k} = \frac{x^{p^k}-1}{x^{p^{k-1}}-1} = \mathrm{C}_p\left(x^{p^{k-1}}\right)$。因此 $\mathrm{C}_9 = x^6+x^3+1$。

4. $\mathrm{C}_{15} = \frac{(x^{15}-1)(x-1)}{(x^3-1)(x^5-1)}$。把 $(x-1)(x^2+x+1)(x^4+x^3+x^2+x+1)$ 與題中給出的八次式相乘並展開，驗證結果為 $x^{15}-1$。

6.2 高斯的正十七邊形

1. $\eta+\eta'$ 是全部六個非實七次單位根之和，即 -1。展開 $\eta\eta'$ 得到九項 α^{a+b}，其指數為 $4,6,0,5,0,1,0,2,3$（模 7）：即三個 1，以及每個非實根各出現一次，所以 $\eta\eta' = 3-1 = 2$。因此 $\eta、\eta'$ 就是 $\frac{-1\pm\mathrm{i}\sqrt{7}}{2}$。$\eta$ 的實部給出餘弦之和。它的虛部是 $\sin\frac{2\pi}{7} + \sin\frac{4\pi}{7} - \sin\frac{\pi}{7}$，由於 $\sin\frac{4\pi}{7} = \sin\frac{3\pi}{7} > \sin\frac{\pi}{7}$，這個值為正；所以它等於 $+\frac{\sqrt{7}}{2}$。

2. 設 $\zeta = \mathrm{e}^{2\pi\mathrm{i}/5}$，則兩個週期（period）$\zeta+\zeta^4 = 2\cos\frac{2\pi}{5}$ 和 $\zeta^2+\zeta^3$ 之和為 -1，之積為 -1，所以它們是 $t^2+t-1=0$ 的根，而其中的正根給出 $\cos\frac{2\pi}{5} = \frac{\sqrt{5}-1}{4}$。作圖方法：在單位圓內，兩條直角邊為 1 和 $\frac{1}{2}$ 的直角三角形，斜邊長為 $\frac{\sqrt{5}}{2}$；減去 $\frac{1}{2}$ 再取一半，便得 $\cos\frac{2\pi}{5}$。在軸上這個距離處作垂線，它與圓相交於一個與 $(1,0)$ 相鄰的頂點；再用這條弦沿圓周逐步截取。

3. 兩者準確至小數點後五位都是 -1.28078。

4. 3 的各次冪依次為 $3,9,10,13,5,15,11,16,14,8,7,4,12,2,6,1$：十六個數全部出現。2 的各次冪在八步之後重複（$2^8 = 256 \equiv 1$），給出 $\{1,2,4,8,9,13,15,16\}$，即模 17 的非零平方數。

6.3 以 $\sqrt{2}+\sqrt{3}+\sqrt{5}$ 為根的多項式

1. $y^3 = 2\sqrt{2}+6\sqrt{3}+9\sqrt{2}+3\sqrt{3} = 11\sqrt{2}+9\sqrt{3}$，而 $9y = 9\sqrt{2}+9\sqrt{3}$，所以 $y^3 - 9y = 2\sqrt{2}$。於是 $\sqrt{3} = y - \sqrt{2}$ 和 $\sqrt{6} = \sqrt{2}\sqrt{3}$ 也都是 y 的多項式。

2. 設 $x - \sqrt{2} = \sqrt[3]{3}$ 並取立方：$x^3 + 6x - 3 = \sqrt{2}(3x^2+2)$。兩邊平方得 $x^6 - 6x^4 - 6x^3 + 12x^2 - 36x + 1 = 0$。

3. 可能的有理根只有 ± 1，而兩者都不是根。由於沒有 x^3 項，分解成兩個首一（monic）二次式時必定形如 $(x^2+ax+b)(x^2-ax+c)$，其中 $a(c-b) = 0$、$b+c-a^2 = -10$ 及 $bc = 1$。若 $a = 0$，則 $b+c = -10$ 且 $bc = 1$，在整數範圍內不可能；若 $b = c = \pm 1$，則 $a^2 = 12$ 或 8，同樣不可能。所以這個四次式不可約（irreducible），而 $\sqrt{2}+\sqrt{3}$ 的次數（degree）為 4。

4. 由棣美弗（De Moivre）定理，$\cos 3\theta = 4\cos^3\theta - 3\cos\theta$；取 $\theta = 20^\circ$，得 $\frac{1}{2} = 4c^3 - 3c$。候選的有理根 $\pm 1, \pm\frac{1}{2}, \pm\frac{1}{4}, \pm\frac{1}{8}$ 全都不成立。所以 $\cos 20^\circ$ 的次數為 3，但每個可作圖（constructible）長度的次數都是 2 的冪，因此無法由 60° 作出 20°。

6.4 解四次方程

1. $y_1y_2 + y_2y_3 + y_3y_1$ 是十二項 $\alpha^2\beta\gamma$（一個根的平方乘以另外兩個根）之和。而 $\left(\sum\alpha\right)\left(\sum\alpha\beta\gamma\right)$ 則是這十二項再加上 $4\alpha\beta\gamma\delta$。所以所求之和為 $(-p)(-r) - 4s$。

2. 比較係數得 $a+b = 1$、$a\alpha + b\beta = 0$ 及 $a\alpha^2 + b\beta^2 = -p$，當 $\alpha \neq \beta$ 時有解，而這正是那兩個條件所保證的：$p \neq 0$ 使關於 t 的方程是真正的二次方程（若 $p = 0$，它會退化成 $-qt = 0$，只有單一個根 0），而 $q^2 \neq 4p^3$ 則使它的判別式（discriminant）$q^2 - 4p^3$ 不為零。對於 $x^3 - 24x + 48$：$p = 8$，$q = 48$，而 $8t^2 - 48t + 64 = 0$ 給出 $\alpha = 2$，$\beta = 4$。於是 $\left(\frac{x-2}{x-4}\right)^3 = -\frac{b}{a} = \frac{\alpha}{\beta} = \frac{1}{2}$，所以 $x - 2 = 2^{-1/3}(x-4)$，整理後即得題中給出的根，$x \approx -5.6946$。

3. 預解三次式（resolvent cubic）為 $y^3 - 7y^2 + 16y - 12 = (y-3)(y-2)^2$。由根 $y = 3$ 得出 $(x^2-2x+1)(x^2-2x+2)$，所以各根為 1（二重根）和 $1 \pm \mathrm{i}$。

4. 在每一對之內可以對換或不對換（4 種選擇），兩對之間也可以互換或不互換（2 種選擇）：恆等置換、$(\alpha\beta)$、$(\gamma\delta)$、$(\alpha\beta)(\gamma\delta)$、$(\alpha\gamma)(\beta\delta)$、$(\alpha\delta)(\beta\gamma)$、$(\alpha\gamma\beta\delta)$ 和 $(\alpha\delta\beta\gamma)$。這與正方形的對稱羣（symmetry group）結構完全相同。

6.5 從公理出發的微分法

1. 常數仍然給出 0，而由規則 (i) 及歸納法（induction）可得 $\lozenge x^n = x^{n-1}\lozenge x + x\lozenge x^{n-1} = nx^n$。再由線性（linearity）得 $\lozenge \mathrm{h} = x\mathrm{h}'$。
2. $\lozenge \mathrm{f} = \frac{\mathrm{f}}{\mathrm{g}}\lozenge \mathrm{g} + \mathrm{g}\lozenge\frac{\mathrm{f}}{\mathrm{g}}$，所以 $\lozenge\frac{\mathrm{f}}{\mathrm{g}} = \frac{\mathrm{g}\lozenge \mathrm{f} - \mathrm{f}\lozenge \mathrm{g}}{\mathrm{g}^2}$。
3. 若 $\mathrm{f} = (x-a)^2\mathrm{g}$，則 $\mathrm{f}' = (x-a)\big(2\mathrm{g} + (x-a)\mathrm{g}'\big)$。$x^n - 1$ 的重根（repeated root）必定也是 nx^{n-1} 的根，所以只能是 0；但 0 並不是 $x^n - 1$ 的根。
4. $(2+\varepsilon)^2 + 1 = 5 + 4\varepsilon$，而 $(5+4\varepsilon)^3 = 125 + 300\varepsilon$。所以 $f(2) = 125$ 及 $f'(2) = 300$；用鏈式法則（chain rule）則得 $3(x^2+1)^2 \cdot 2x = 3 \cdot 25 \cdot 4 = 300$。

6.6 子羣的併集

1. 設 $V = \{e, a, b, c\}$：$a^2 = b^2 = c^2 = e$，而 a、b、c 之中任意兩個之積都是第三個。每個元素都是自己的逆元（inverse）。在 $\{1, \mathrm{i}, -1, -\mathrm{i}\}$ 中，i 的逆元是 $-\mathrm{i}$，所以這兩個羣並不相同。
2. 假如 g 屬於某個真子羣（proper subgroup）H，那麼 g 的每個冪也都屬於它，於是 H 就是整個羣。
3. 1 既不是偶數，也不是 3 的倍數。更一般地，$\mathbb{Z}$ 是循環羣（cyclic group，由 1 生成），所以根據上一題，它不是任意多個真子羣的並集（union）。
4. 直線 $y = mx$ 與 $y = x^2$ 只相交於 $x = 0$ 和 $x = m$ 兩處，而直線 $x = 0$ 與它只相交於原點。所以 k 條直線最多只能包含 k 個形如 (t, t^2) 而 $t \neq 0$ 的點，但這樣的點有無窮多個。

6.7 四元數羣

1. $C \circ B = (\mathrm{i}, 0) \circ (0, 1) = (0, \mathrm{i}) = D$，而 $B \circ C = (0, 1) \circ (\mathrm{i}, 0) = (0, \mathrm{i}^*) = (0, -\mathrm{i}) = H$。
2. A 的階（order）為 1，E 的階為 2；其餘六個元素的平方都是 E，所以階為 4。正方形的對稱羣有五個階為 2 的元素（四個反射和一個半周旋轉），所以這兩個羣並不同構（isomorphic）。
3. $\det M = z_1 z_1^* + z_2 z_2^* = |z_1|^2 + |z_2|^2$，這是四個實數平方之和。由於 $M(z)M(w)$ 是 $z \circ w$ 的矩陣，把行列式（determinant）相乘，便可把 $(a_1^2 + \cdots + a_4^2)(b_1^2 + \cdots + b_4^2)$ 寫成 $|u_1|^2 + |u_2|^2$，同樣是四個平方之和。
4. 各子羣為 $\{A\}$、$\{A, E\}$、$\{A, E, C, G\}$、$\{A, E, B, F\}$、$\{A, E, D, H\}$ 和 Q_8。除 A 以外，每個元素都有某個冪等於 E，所以除 $\{A\}$ 以外的任何子羣都包含 E。三個階為 4 的子羣合起來包含了全部八個元素。

6.8 由奇異矩陣組成的羣

1. $\mathbf{M}_a\mathbf{M}_b = \mathbf{M}_{2ab}$，而 $2(2ab) = (2a)(2b)$。這個映射是一對一（one-to-one）的，並且映成（onto）全體非零實數。
2. $\begin{pmatrix} a & b \\ 0 & 0 \end{pmatrix}\begin{pmatrix} c & d \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} ac & ad \\ 0 & 0 \end{pmatrix}$，其中 $ac \neq 0$，所以這個集合是封閉（closed）的。每個 $\begin{pmatrix} 1 & e \\ 0 & 0 \end{pmatrix}$ 都是左單位元（left identity）。但若要 $\mathbf{XE} = \mathbf{X}$，其中 $\mathbf{E} = \begin{pmatrix} e & f \\ 0 & 0 \end{pmatrix}$，便需要 $e = 1$ 及 $f = b/a$，而後者取決於 $\mathbf{X}$：所以沒有右單位元，因而沒有單位元，也就不成羣。
3. $\mathbf{0}$、$\mathbf{I}$，以及滿足 $bc = a(1-a)$ 的矩陣 $\begin{pmatrix} a & b \\ c & 1-a \end{pmatrix}$（跡（trace）為 1，行列式為 0）。這些就是把平面沿一條直線的方向投影（projection）到另一條直線上的變換。
4. 所有元素均為 a 的矩陣 $\mathbf{J}_a$ 滿足 $\mathbf{J}_a\mathbf{J}_b = \mathbf{J}_{nab}$。單位元是 $\mathbf{J}_{1/n}$，而 $\mathbf{J}_a$ 的逆元是 $\mathbf{J}_{1/(n^2a)}$。

章

7

線性代數與幾何

矩陣不是一張數字表。它是空間的一種變換，而且自有其生命。

本章內容

7.1	對矩陣不成立的法則	STEP II 1993
7.2	只伸縮，不轉向	STEP III 1990
7.3	甚麼才算行列式	STEP II 2019
7.4	回到起點的矩陣	STEP II 2020
7.5	三維空間中的旋轉	STEP III 2022
7.6	雙曲旋轉與相對論	STEP III 1991
7.7	永恆不變的量	STEP III 2024
7.8	四個平方數	STEP III 1994
7.9	只有五種柏拉圖立體	STEP III 1999
A	練習答案	第 497 頁

空間的變換

我們思考時不依賴基底，寫作時也不依賴基底，但到了緊要關頭，我們就關上辦公室的門，拼命地用矩陣計算。

— Irving Kaplansky，談他與 Paul Halmos 處理線性代數的方式

在中學，矩陣（matrix）通常以解聯立方程的工具出現，或者是變換點的一套步驟：把矩陣乘以一個列向量（column vector），然後讀出新的坐標。兩者都有用，但它們掩蓋了矩陣之所以成為數學中最重要概念之一的原因。矩陣是一個線性變換（linear transformation）：一種把空間拉伸、旋轉、反射和錯切的方式，而且矩陣可以相加、相乘、取冪和求導，就像數一樣。只不過它們又不像數。矩陣的乘法取決於次序，兩個矩陣可以相乘得零，而且它們有隱藏的不變量（invariant）——跡（trace）、行列式（determinant）、特徵值（eigenvalue）——無論矩陣寫成甚麼樣子，這些量都能刻劃它的本質。

線性代數為何重要

線性代數（linear algebra）大概是應用最廣的數學分支。每一個電腦圖像引擎，由電子遊戲到電影動畫，都是一部做矩陣乘法的機器：角色的每一次旋轉、每一個鏡頭角度、每一個透視投影，都是一個矩陣。互聯網上的搜尋引擎和推薦系統建基於特徵向量（eigenvector）。工程師利用特徵值求出橋樑和建築物的振動頻率；統計學家利用它們找出數據中的主要規律；量子力學完全以線性代數的語言寫成。而現代的機器學習，核心就是數量極其龐大的矩陣乘法。

線性代數無處不在，是有原因的。線性問題是我們能夠徹底解決的問題，而幾乎每一個非線性問題，只要放大到足夠近，看起來都是線性的。這正是微積分的想法，也意味着幾乎所有事物的局部性態，都是用線性代數這種語言來描述的。

你在大學會遇到甚麼

線性代數與分析（analysis）並列為大學一年級的兩大支柱。你會遇到向量空間（vector space，即可以做加法和純量乘法的集合，平面和三維空間只是最初的例子）以及它們之間的線性映射（linear map），並且學到矩陣其實就是用坐標寫出來的線性映射。你會學習行列式、特徵值和特徵向量，以及譜定理（spectral theorem），它指出對稱矩陣（symmetric matrix）可以藉一個旋轉來對角化（diagonalise）。往後的課程處處用到這些內容：在微分方程（第 8 章）中，在幾何中（旋轉和反射在那裏構成羣），在描述連續對

稱的李羣（Lie group）理論中，以及在泛函分析（functional analysis）中，也就是無限多維的線性代數。

貫穿本章的問題

三大問題

1. **矩陣在哪些方面像數，又在哪些方面不同？**哪些算術法則仍然成立，哪些不再成立？
2. **矩陣究竟做了甚麼？**我們怎樣才能看出一個變換的幾何，又有哪些數能夠不依賴坐標而刻劃它？
3. **甚麼保持不變？**哪些量在變換之下，或在運動之下保持不變，它們又告訴我們甚麼？

本章由矩陣代數的法則開始（Section 7.1），然後轉到變換的幾何：特徵向量、行列式、有限階矩陣、三維空間中的旋轉，以及狹義相對論中的雙曲旋轉（hyperbolic rotation）。接着是隨時間演化的矩陣的守恆量（conserved quantity）、由行列式得出的四平方恆等式（four-square identity），最後以一個計數論證作結，即歐拉（Euler）公式，它把柏拉圖立體（Platonic solids）分類。

本章的概念

§	主題	概念
7.1	不成立的法則	非交換性（non-commutativity）；零因子（zero divisor）
7.2	特徵向量	變換不會轉動的方向
7.3	行列式	由可乘性刻劃；有向面積
7.4	有限階矩陣	跡、行列式與凱萊–哈密頓（Cayley–Hamilton）定理
7.5	三維旋轉	由叉積（cross product）得出羅德里格斯（Rodrigues）公式
7.6	洛倫茲推進（Lorentz boosts）	雙曲角可以相加
7.7	守恆量	交換子（commutator）的跡為零
7.8	四平方	行列式的可乘性
7.9	柏拉圖立體	算兩次（double counting）與歐拉公式

如何閱讀本章

對本章的每一個矩陣，都把它所做的事畫出來。取單位正方形，或單位圓，或幾個向量，然後畫出它們的像的草圖。花兩分鐘畫一幅草圖，往往在你動手計算之前就讓你看到答案（一個特徵向量、一個反例、一條不變直線），事後也有助你檢查計算。

7.1 對矩陣不成立的法則

題目

STEP II 1993，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/93-s2-q6`

引子

矩陣（matrix）可以相加，也可以相乘，所以我們很容易把它們當作數來看待。這道題正是要考驗這種想法：它列出四條對數成立的「法則」，問哪些對矩陣仍然成立。結果大部分都不成立，而它們不成立的方式很有啟發性。矩陣乘法不滿足交換律（not commutative）；有些非零矩陣的乘積是零；而一個矩陣的二次方程可以有無窮多個解。

這些「失效」並不是缺陷，反而正是矩陣有用的原因：矩陣描述的是空間的變換（transformation），而以不同的次序進行兩個變換，確實會得到不同的結果。1925 年，海森堡（Werner Heisenberg）發現量子力學中的物理量的表現像矩陣而不像數，而 $\mathbf{AB} = \mathbf{BA}$ 不成立，正是不確定性原理（uncertainty principle）的數學形式。

動筆之前

想一想 對實數而言，為甚麼 $ab = 0$ 可推出 $a = 0$ 或 $b = 0$？這個論證用到了數的哪一項性質？這項性質對矩陣成立嗎？

若 $ab = 0$ 而 $a \neq 0$，兩邊除以 a 便得 $b = 0$。這個論證需要每個非零的數都有逆（inverse）。非零矩陣卻不一定有逆（行列式（determinant）為零的矩陣就沒有），所以這個論證行不通。同時也要記住幾何上的圖像：矩陣是平面的一個變換，而乘積 $\mathbf{AB}$ 的意思是「先做 $\mathbf{B}$，再做 $\mathbf{A}$」。

逐步解題

階段 1　能否有 $\mathbf{AB} = \mathbf{0}$ 但 $\mathbf{BA} \neq \mathbf{0}$？

想一想　從幾何角度去想。你希望 $\mathbf{B}$ 把所有東西送到一條直線上，然後 $\mathbf{A}$ 把這條直線壓成零。但次序調轉時，$\mathbf{A}$ 不應把所有東西都壓到會被 $\mathbf{B}$ 消去的地方。

設 $\mathbf{A} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$，它是到 x 軸上的投影（projection），把 y 軸消去；又設 $\mathbf{B} = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$，它把 (x, y) 送到 $(0, x)$，即送到 y 軸上。於是

$$\mathbf{AB} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}, \qquad \mathbf{BA} = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \neq \mathbf{0}.$$

這個斷言是**錯**的（圖 7.1.1）。 ■

階段 2　平方差

小心展開，保持每個乘積的次序：

$$(\mathbf{A} - \mathbf{B})(\mathbf{A} + \mathbf{B}) = \mathbf{A}^2 + \mathbf{AB} - \mathbf{BA} - \mathbf{B}^2.$$

這個式子等於 $\mathbf{A}^2 - \mathbf{B}^2$，當且僅當 $\mathbf{AB} = \mathbf{BA}$。取第 1 階段的矩陣，$\mathbf{AB} - \mathbf{BA} = -\mathbf{BA} \neq \mathbf{0}$，所以這個斷言是**錯**的。 ■

階段 3　$\mathbf{AX} = \mathbf{0}$ 的非零解

想一想　若 $\det \mathbf{A} \neq 0$，你可以對 $\mathbf{AX} = \mathbf{0}$ 做甚麼？若 $\det \mathbf{A} = 0$，你能否找到一個非零的列向量（column vector）$\mathbf{v}$，使得 $\mathbf{Av} = \mathbf{0}$？

若 $\det \mathbf{A} \neq 0$，則 $\mathbf{A}^{-1}$ 存在，而 $\mathbf{X} = \mathbf{A}^{-1}\mathbf{AX} = \mathbf{0}$：沒有非零解。

若 $\det \mathbf{A} = 0$，寫 $\mathbf{A} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$，其中 $ad = bc$。向量 $\begin{pmatrix} b \\ -a \end{pmatrix}$ 和 $\begin{pmatrix} d \\ -c \end{pmatrix}$ 滿足 $\mathbf{A}\begin{pmatrix} b \\ -a \end{pmatrix} = \begin{pmatrix} ab - ba \\ cb - da \end{pmatrix} = \mathbf{0}$，另一個向量亦同理。由於 $\mathbf{A} \neq \mathbf{0}$，兩者之中至少有一個非零；把它記作 $\mathbf{v}$。那麼，兩列都是 $\mathbf{v}$ 的矩陣 $\mathbf{X}$ 是非零的，並且 $\mathbf{AX} = \mathbf{0}$。所以這個斷言是**對**的。 ■

階段 4　有多少個平方根？

停一停，試一試 找出非零的 $\mathbf{A}$ 和 $\mathbf{B}$，使得 $\mathbf{X}^2 + \mathbf{AX} + \mathbf{B} = \mathbf{0}$ 有無窮多個解。（提示：把左邊變成完全平方 $(\mathbf{X} - \mathbf{I})^2$，再想想平方為零的矩陣。）

取 $\mathbf{A} = -2\mathbf{I}$ 和 $\mathbf{B} = \mathbf{I}$，兩者都非零。方程變成 $(\mathbf{X} - \mathbf{I})^2 = \mathbf{0}$。任何滿足 $\mathbf{N}^2 = \mathbf{0}$ 的矩陣 $\mathbf{N}$ 都給出一個解 $\mathbf{X} = \mathbf{I} + \mathbf{N}$，而這樣的 $\mathbf{N}$ 有無窮多個：

$$\mathbf{N} = \begin{pmatrix} 0 & t \\ 0 & 0 \end{pmatrix}, \qquad \mathbf{N}^2 = \mathbf{0} \text{ 對任意實數} t.$$

所以對每一個 t，$\mathbf{X} = \begin{pmatrix} 1 & t \\ 0 & 1 \end{pmatrix}$ 都是解，這個斷言是**錯**的。 ■

回顧

- **次序很重要。**每次展開都必須保持乘積的次序；$\mathbf{AB}$ 和 $\mathbf{BA}$ 通常並不相同。
- **零因子（zero divisors）。**非零矩陣相乘可以得零，這就破壞了「二次方程最多有兩個根」背後的論證。
- **從變換的角度去想。**只要想像矩陣對平面做了甚麼，每個反例都很容易找到。

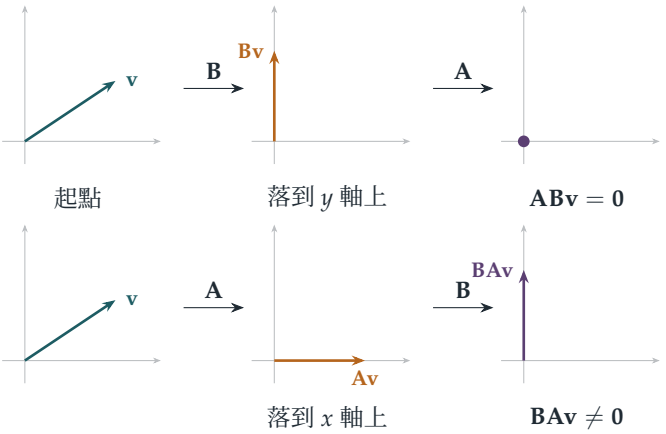


圖 7.1.1 為甚麼 $\mathbf{AB} = \mathbf{0}$ 但 $\mathbf{BA} \neq \mathbf{0}$。上：$\mathbf{B}$ 把每個向量送到 y 軸上，然後 $\mathbf{A}$（到 x 軸上的投影）把它壓成零。下：次序調轉時，$\mathbf{A}$ 保留 x 分量，$\mathbf{B}$ 再把它移到 y 軸上，它在那裏得以保存下來。

背後的數學

矩陣作為一種代數

凱萊（Arthur Cayley）在 1858 年引入矩陣代數，把矩陣看成可以相加和相乘的單一對象，而不僅僅是係數表。他立刻注意到乘法不滿足交換律，並證明了凱萊—哈密頓定理（*Cayley–Hamilton theorem*）：每個方陣都滿足它自己的特徵方程（characteristic equation）。對於跡（trace）為 t、行列式為 d 的 2×2 矩陣，這就是說 $\mathbf{M}^2 - t\mathbf{M} + d\mathbf{I} = \mathbf{0}$，本章會反覆用到這個事實（Section 7.4）。2×2 實矩陣構成一個環（*ring*）：可以加、減、乘，但乘法不滿足交換律，而且存在零因子，即乘積為零的非零元素。

無窮多個平方根

第 4 階段的結論還可以變得更加驚人。方程 $\mathbf{X}^2 = \mathbf{I}$ 有明顯的解 $\pm\mathbf{I}$，但平面上對任何一條通過原點的直線的反射（reflection）$\mathbf{X} = \begin{pmatrix} \cos\theta & \sin\theta \\ \sin\theta & -\cos\theta \end{pmatrix}$ 也都是解，因為反射兩次等於甚麼也沒做。這就是 $\mathbf{I}$ 的一整個圓那麼多的平方根。係數為數的 n 次多項式最多有 n 個根，是因為數沒有零因子；矩陣卻有，於是這個根的數目上限便不再成立。

物理學中的非交換性

在經典力學中，位置 x 和動量 p 都是數，並且 $xp = px$。在量子力學中，它們變成算子（operator，即無窮矩陣），滿足 $xp - px = \mathrm{i}\hbar$，其中 $\hbar$ 是普朗克（Planck）常數除以 2π。單憑這一條非交換關係，就可推出海森堡的不確定性原理：位置和動量不可能同時被準確得知。海森堡在 1925 年發現量子力學的法則時，並不知道矩陣是甚麼；是玻恩（Max Born）認出海森堡所發明的那條奇怪的乘法法則就是矩陣乘法。

思維技巧

在新的情境中檢驗熟悉的法則

遇到一種新的對象時，要檢查哪些熟悉的法則仍然成立，並弄清楚每條法則究竟依賴哪一項性質。

法則「$ab = 0$ 可推出 $a = 0$ 或 $b = 0$」依賴除法；「$(a-b)(a+b) = a^2 - b^2$」依賴交換律。知道每條法則建基於甚麼，你就知道甚麼時候仍然可以使用它。

延伸練習

1 **甚麼時候可交換？**求所有與 $\begin{pmatrix}1 & 1\\ 0 & 1\end{pmatrix}$ 可交換的 2×2 矩陣。

2 **零的平方根。**求所有滿足 $\mathbf{N}^2=\mathbf{0}$ 的 2×2 實矩陣 $\mathbf{N}$。（提示：利用凱萊–哈密頓定理證明 $\operatorname{tr}\mathbf{N}=\det\mathbf{N}=0$。）

3 **反射。**證明每個反射矩陣 $\begin{pmatrix}\cos\theta & \sin\theta\\ \sin\theta & -\cos\theta\end{pmatrix}$ 的平方都是 $\mathbf{I}$，並求出它所對應的反射軸。

4 **凱萊–哈密頓定理。**以直接計算證明每個 2×2 矩陣都滿足 $\mathbf{M}^2-(\operatorname{tr}\mathbf{M})\mathbf{M}+(\det\mathbf{M})\mathbf{I}=\mathbf{0}$。

延伸閱讀

- **Gilbert Strang, *Introduction to Linear Algebra***（Wellesley-Cambridge）。最受歡迎的線性代數入門教科書，着重理解。
- **Sheldon Axler, *Linear Algebra Done Right***（Springer；可在網上免費閱讀）。取向較為抽象，以線性映射而非矩陣為中心。
- **Jim Baggott, *The Quantum Story***（Oxford University Press）。量子力學的歷史，包括海森堡的矩陣力學。

7.2 只伸縮，不轉向

題目

STEP III 1990，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/90-s3-q6`

引子

矩陣變換（matrix transformation）通常會把每個向量轉到新的方向。不過，許多矩陣都有幾個特別的方向，它完全不會轉動這些方向，只會沿着它們拉伸或壓縮。這些方向上的向量就是矩陣的*特徵向量*（eigenvector，源自德文 *eigen*，意思是「自身的」或「特有的」），而拉伸的倍數就是它的*特徵值*（eigenvalue）。當特徵向量的數目足夠時（像本題這類對稱矩陣就必定足夠），這個變換便一目了然：在合適的坐標下，它只不過是沿兩條軸的拉伸。

本題要找出某個矩陣的特徵向量，再用它們來理解這個矩陣對一個圓有甚麼作用。特徵向量是整個應用數學中最有用的概念之一。它們描述橋樑的固有振動、旋轉物體的主軸、原子的能級、一團數據中最重要的方向，以及 Google 最初的 PageRank 演算法怎樣為網頁排名。

動筆之前

想一想 T 使一個向量「方向不變，只乘以一個純量」，這是甚麼意思？試把它寫成一條等式。

意思是存在某個數 λ 使得 $T\mathbf{v} = \lambda\mathbf{v}$：$\mathbf{v}$ 是特徵值為 λ 的特徵向量。留意題目中的兩個向量 $(1, 2)$ 和 $(2, -1)$ 互相垂直。這並非巧合：這個矩陣是*對稱*（symmetric）的，而對稱矩陣總有互相垂直的特徵向量。

逐步解題

階段 1　特徵向量

$$T\begin{pmatrix}1\\2\end{pmatrix}=\frac{2}{5}\begin{pmatrix}9-4\\-2+12\end{pmatrix}=\frac{2}{5}\begin{pmatrix}5\\10\end{pmatrix}=2\begin{pmatrix}1\\2\end{pmatrix},\qquad T\begin{pmatrix}2\\-1\end{pmatrix}=\frac{2}{5}\begin{pmatrix}20\\-10\end{pmatrix}=4\begin{pmatrix}2\\-1\end{pmatrix}.$$

所以 $(1,2)$ 被拉伸至 2 倍，而 $(2,-1)$ 被拉伸至 4 倍。 ■

階段 2　圓的像

> **想一想**　要求出 E 的方程，可先用 X 和 Y 表示 x 和 y（即求矩陣的逆），再代入 $x^2+y^2=1$。

矩陣 $\frac{2}{5}\begin{pmatrix}9&-2\\-2&6\end{pmatrix}$ 的行列式（determinant）是 $\frac{4}{25}(54-4)=8$，它的逆矩陣（inverse）是 $\frac{1}{20}\begin{pmatrix}6&2\\2&9\end{pmatrix}$。所以 $x=\frac{6X+2Y}{20}$ 及 $y=\frac{2X+9Y}{20}$，而 $x^2+y^2=1$ 變成

$$(6X+2Y)^2+(2X+9Y)^2=400\quad\Longrightarrow\quad 40X^2+60XY+85Y^2=400,$$

兩邊除以 5，便得 $8X^2+12XY+17Y^2=80$。 ■

> **想一想**　E 是甚麼形狀？它的面積是多少？試利用特徵向量作答。

單位圓在每個方向上的半徑都是 1。在 $(2,-1)$ 的方向上，它被拉伸至 4 倍；在與之垂直的 $(1,2)$ 方向上，則被拉伸至 2 倍。所以 E 是一個橢圓（ellipse），半軸長為 4 和 2，分別沿着兩個特徵向量的方向（圖 7.2.1），其面積為 $\pi\times4\times2=8\pi$。（換個說法：任何線性映射（linear map）都會把面積乘以 $|\det T|=8$，而圓的面積是 π。）

階段 3　X 的最大值

> **想一想**　在 C 上，$X=\frac{2}{5}(9x-2y)$。當 $x^2+y^2=1$ 時，$9x-2y$ 最大可以是多少？

$9x-2y$ 是 $(9,-2)$ 與單位向量 (x,y) 的點積（dot product），所以它的最大值是 $|(9,-2)|=\sqrt{85}$，在 (x,y) 指向 $(9,-2)$ 的方向時取得。因此 X 的最大值是 $\frac{2}{5}\sqrt{85}=2\sqrt{\frac{85}{25}}=2\sqrt{\frac{17}{5}}$。 ■

階段 4　切線

停一停，試一試 C 上的點 $\frac{1}{5}(3,4)$ 的切線（tangent）是 $3x+4y=5$。線性映射會把切線變成切線。試利用這一點求出 E 的切線。

把 $x=\frac{6X+2Y}{20}$、$y=\frac{2X+9Y}{20}$ 代入 $3x+4y=5$：

$$\frac{3(6X+2Y)+4(2X+9Y)}{20}=5 \quad\Longrightarrow\quad 26X+42Y=100,$$

即 $13X+21Y=50$。 ■

（驗算：該點本身的像是 $T(\frac{3}{5},\frac{4}{5})=(\frac{38}{25},\frac{36}{25})$，而 $13\cdot\frac{38}{25}+21\cdot\frac{36}{25}=\frac{494+756}{25}=50$。）

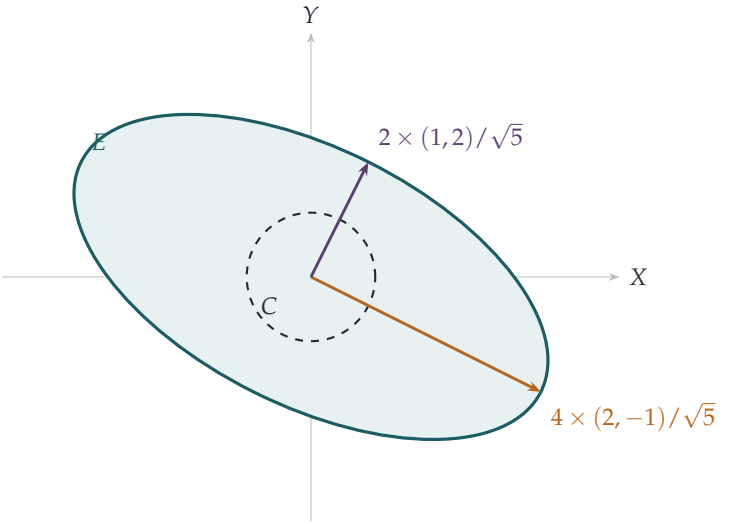


圖 7.2.1 單位圓 C（虛線）及它在 T 之下的像 E。沿着特徵向量 $(2,-1)$，圓被拉伸至 4 倍；沿着與之垂直的特徵向量 $(1,2)$，則被拉伸至 2 倍。所以 E 是一個橢圓，半軸長為 4 和 2，分別沿着這兩個方向，面積為 8π。

回顧

- **特徵向量是天然的坐標軸。**在特徵向量所給出的坐標下，T 只是分別拉伸至 4 倍和 2 倍，它對圓的作用顯而易見。
- **行列式量度面積。**E 的面積是 C 的面積的 $|\det T|$ 倍。
- **線性映射保持相切關係。**把圓的切線加以變換，幾乎不用計算便得到橢圓的切線。

背後的數學

譜定理

每個實對稱矩陣都有一整組互相垂直的特徵向量，而且特徵值都是實數。這就是譜定理（spectral theorem），是大學第一門線性代數（linear algebra）課程的核心結果之一。它告訴我們，每個對稱的線性映射，在合適的垂直坐標下，都只是沿各條軸的拉伸。換個說法，每個像 $8X^2 + 12XY + 17Y^2$ 這樣的二次型（quadratic form），都可以經旋轉化成 $\lambda_1 u^2 + \lambda_2 v^2$ 的形式，所以每條有心圓錐曲線（central conic）都是軸互相垂直的橢圓或雙曲線（主軸定理，principal axis theorem）。這個想法可追溯至 1820 年代的柯西（Cauchy），他用它來求二次曲面的主軸，以及固體內應力的主軸。

這一切之所以成立，全靠對稱性。不對稱的矩陣就未必這樣合作：旋轉 90° 會轉動每一個方向，所以它完全沒有實特徵向量；而錯切（shear）$\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ 只有一個特徵向量方向，所以它不是沿任何一對軸的拉伸。

無處不在的特徵向量

以下是特徵向量出現的其中幾個地方。

- **振動。**由質點和彈簧組成的系統會以簡正模態（normal mode）振動，即各部分以同一頻率同步運動的模式。這些模態是某個矩陣的特徵向量，而頻率則來自特徵值（Section 11.2）。
- **數據。**一團數據點分佈得最散的方向，就是其協方差矩陣（covariance matrix）的特徵向量。找出這些方向的方法稱為主成分分析（principal component analysis），是統計學和機器學習的基本工具之一（Section 13.7）。
- **量子力學。**量度一個物理量可能得到的結果，是某個矩陣（或算子）的特徵值，而原子的能級也是特徵值。
- **網上搜尋。**Google 的 PageRank 是這個搜尋引擎背後的原始構想：有一個描述網頁之間連結的巨型矩陣，PageRank 就按它那個特徵值為 1 的特徵向量的各個分量來為網頁排名。

思維技巧

選擇配合變換的坐標

要理解一個線性變換，就找出它不會轉動的方向。在這些坐標下，它會變得簡單。

這與 Section 5.6 中選費馬點（Fermat point）作原點，或 Section 2.8 中以變數代換來簡化映射，是同一個想法：選對了坐標，結構便一目了然。

延伸練習

1 **求特徵值。**證明 $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ 的特徵值是 $\lambda^2 - (a+d)\lambda + (ad-bc) = 0$ 的根，並驗證題目中的矩陣給出 $\lambda = 2$ 和 4。

2 **互相垂直的特徵向量。**證明若 $\mathbf{M}$ 是對稱矩陣，且 $\mathbf{Mu} = \lambda\mathbf{u}$、$\mathbf{Mv} = \mu\mathbf{v}$，其中 $\lambda \neq \mu$，則 $\mathbf{u}\cdot\mathbf{v} = 0$。（提示：用兩種方法計算 $\mathbf{v}\cdot\mathbf{Mu}$。）

3 **旋轉。**證明除非 θ 是 π 的倍數，否則旋轉矩陣 $\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$ 沒有實特徵向量。它的複特徵值是甚麼？

4 **冪。**利用特徵向量，求 $T^n \begin{pmatrix} 1 \\ 0 \end{pmatrix}$。（提示：把 $(1,0)$ 寫成 $(1,2)$ 和 $(2,-1)$ 的組合。）

延伸閱讀

- **Gilbert Strang, *Introduction to Linear Algebra***（Wellesley-Cambridge）。第 6 章是特徵值及其應用的出色入門。
- **Kurt Bryan & Tanya Leise, "The $25,000,000,000 eigenvector: the linear algebra behind Google"**，*SIAM Review* 48 (2006)。淺白易讀，可免費取閱。
- **3Blue1Brown, *Essence of Linear Algebra***（影片系列，可在網上免費觀看）。以精美的動畫呈現的，正是本節的幾何概念。

7.3 甚麼才算行列式

題目

STEP II 2019，第 8 題

繼續閱讀之前，請先看題目（原題為英文）。題目載於 OCR 提供下載的 2019 年 STEP 試卷（zip 檔案）內的 STEP II 試卷：

`www.ocr.org.uk/Images/696590-2019-past-papers.zip`

引子

2×2 矩陣（matrix）的行列式（determinant）是 $ad - bc$。這條公式從何而來？它為甚麼如此重要？一個答案是：它是矩陣把面積放大或縮小的倍數。另一個答案，也就是本題要探討的，是行列式基本上是矩陣上唯一能把矩陣乘法變成普通乘法的函數：$\det(\mathbf{MN}) = \det \mathbf{M} \det \mathbf{N}$。

題目給你一個未知函數 f，它具有這一個性質（另加兩個用來排除平凡情況的寬鬆條件），然後要求你一步一步推導出它必然有怎樣的表現。與 Section 6.5 的導子一樣，這是一道單憑性質來確定一個對象的練習。到最後，f 被迫在兩行對調時變號、在奇異矩陣（singular matrix）上取值為零、在錯切（shear）上等於 1：這正是行列式的表現。

動筆之前

想一想 題目中每個特殊矩陣各有甚麼作用？對一般的 $\mathbf{M} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$，計算 **JM**、**MJ**、**KM** 和 **PM**。

JM 把 **M** 的兩行對調；**MJ** 把它的兩列對調。**KM** 把第二行乘以 k。**PM** 把第二行加到第一行上。這些就是高斯消元法（Gaussian elimination）中的初等行運算（elementary row operations）。由於 f 把乘積變成乘積，只要知道 f 在這些特殊矩陣上的值，就知道 f 對每種行運算有何反應。

逐步解題

階段 1　單位矩陣

$\mathrm{f}(\mathbf{M}) = \mathrm{f}(\mathbf{MI}) = \mathrm{f}(\mathbf{M})\,\mathrm{f}(\mathbf{I})$。由於 $\mathrm{f}(\mathbf{M}) \neq 0$，兩邊相除即得 $\mathrm{f}(\mathbf{I}) = 1$。 ■

階段 2　對調

$\mathbf{J}^2 = \mathbf{I}$，所以 $\mathrm{f}(\mathbf{J})^2 = \mathrm{f}(\mathbf{I}) = 1$，由此得 $\mathrm{f}(\mathbf{J}) = \pm 1$。由於 $\mathrm{f}(\mathbf{J}) \neq 1$，故 $\mathrm{f}(\mathbf{J}) = -1$。

現在 $\begin{pmatrix} c & d \\ a & b \end{pmatrix} = \mathbf{JM}$，所以它的值是 $\mathrm{f}(\mathbf{J})\mathrm{f}(\mathbf{M}) = -\mathrm{f}(\mathbf{M})$。而 $\begin{pmatrix} d & c \\ b & a \end{pmatrix} = \mathbf{JMJ}$（先對調兩行，再對調兩列），所以它的值是 $\mathrm{f}(\mathbf{J})^2\mathrm{f}(\mathbf{M}) = \mathrm{f}(\mathbf{M})$。 ■

階段 3　線性相關的兩行

想一想　先考慮一個兩行相等的矩陣 $\mathbf{B}$。$\mathbf{JB}$ 是甚麼？這對 $\mathrm{f}(\mathbf{B})$ 說明了甚麼？

若 $\mathbf{B}$ 的兩行都是 (a, b)，把它們對調不會帶來任何改變：$\mathbf{JB} = \mathbf{B}$。所以 $\mathrm{f}(\mathbf{B}) = \mathrm{f}(\mathbf{J})\mathrm{f}(\mathbf{B}) = -\mathrm{f}(\mathbf{B})$，這就迫使 $\mathrm{f}(\mathbf{B}) = 0$。

現在若 $\mathbf{A}$ 的兩行是 (a, b) 和 $k(a, b)$，則 $\mathbf{A} = \mathbf{KB}$，其中 $\mathbf{B}$ 的兩行都等於 (a, b)。所以 $\mathrm{f}(\mathbf{A}) = \mathrm{f}(\mathbf{K})\mathrm{f}(\mathbf{B}) = 0$。 ■（這包括 $k = 0$ 的情況，此時第二行為零。）

階段 4　錯切

停一停，試一試　計算 $\mathbf{P}^2$ 和 $\mathbf{K}^{-1}\mathbf{PK}$。選取 k 使兩者相等。由此可得 $\mathrm{f}(\mathbf{P})$ 是甚麼？

$\mathbf{P}^2 = \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix}$，而當 $k \neq 0$ 時，

$$\mathbf{K}^{-1}\mathbf{PK} = \begin{pmatrix} 1 & 0 \\ 0 & \frac{1}{k} \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & k \end{pmatrix} = \begin{pmatrix} 1 & k \\ 0 & 1 \end{pmatrix}.$$

取 $k = 2$，則 $\mathbf{K}^{-1}\mathbf{PK} = \mathbf{P}^2$。對兩邊取 f，並利用 $\mathrm{f}(\mathbf{K}^{-1})\mathrm{f}(\mathbf{K}) = \mathrm{f}(\mathbf{I}) = 1$，得

$$\mathrm{f}(\mathbf{P}) = \mathrm{f}(\mathbf{P})^2,$$

所以 $\mathrm{f}(\mathbf{P}) = 0$ 或 1。但 $\mathbf{P}$ 是可逆的（invertible），而 $\mathrm{f}(\mathbf{P})\mathrm{f}(\mathbf{P}^{-1}) = \mathrm{f}(\mathbf{I}) = 1$，所以 $\mathrm{f}(\mathbf{P}) \neq 0$。因此 $\mathrm{f}(\mathbf{P}) = 1$。 ■

回顧

- **特殊元素揭示函數的面貌。**把乘法性質應用於 $\mathbf{I}$、$\mathbf{J}$、$\mathbf{K}$ 和 $\mathbf{P}$，便得出 f 在每個矩陣上的值，以及它對每種行運算的影響。
- **與自身的變換相等的元素。**兩行相等的矩陣經行對調後保持不變，而行對調會改變 f 的符號；所以 f 在該處必為零。
- **共軛（conjugation）保持 f 不變。**$\mathrm{f}(\mathbf{K}^{-1}\mathbf{P}\mathbf{K}) = \mathrm{f}(\mathbf{P})$，這個事實在代數中一再用到。

背後的數學

每個乘性函數都是行列式的函數

題目在 $\mathrm{f}(\mathbf{P}) = 1$ 處便停下，但這個論證是可以完成的。每個可逆矩陣都是初等矩陣（elementary matrices，即行對調、行縮放和錯切）的乘積，而題目已說明 f 在每一種上的表現。由此可知 $\mathrm{f}(\mathbf{M}) = \phi(\det \mathbf{M})$，其中 ϕ 是某個單一實數變量的函數，並滿足 $\phi(xy) = \phi(x)\phi(y)$。行列式本身就是 $\phi(x) = x$ 的情況，但 $\phi(x) = x^3$ 或 $\phi(x) = |x|^{1/2}\operatorname{sign}(x)$ 同樣可行。換言之，*行列式是矩陣上唯一眞正嶄新的乘性函數*。用羣論（group theory）的語言來說，可逆矩陣所成的羣之中，凡是交換羣（commutative group）所能看見的一切，行列式都已盡數捕捉。

行列式作為面積

從幾何上看，2×2 矩陣把單位正方形變成一個平行四邊形，而 $\det \mathbf{M}$ 就是它的*有向面積*（signed area）：若定向（orientation）保持不變則為正，若定向反轉則為負（圖 7.3.1）。這樣，乘法性質的意義便一目了然：相繼進行兩個變換，面積的比例因子便會相乘。行對調會反轉定向，所以 $\det \mathbf{J} = -1$；錯切把正方形推斜成面積相同的平行四邊形，所以 $\det \mathbf{P} = 1$；而兩行線性相關的矩陣會把正方形壓扁，所以它的行列式為零。題目的每一步都有對應的圖像。



圖 7.3.1　行列式作為有向面積。兩個基向量（basis vectors）的像（琥珀色、梅紫色）張成一個平行四邊形。錯切保持面積不變；對調基向量會反轉定向；線性相關的兩行（或兩列）把正方形壓縮到一條直線上。相繼進行兩個變換，有向面積便會相乘，這正是性質 $\mathrm{f}(\mathbf{M}\mathbf{N}) = \mathrm{f}(\mathbf{M})\mathrm{f}(\mathbf{N})$。

高維的行列式

對於 $n \times n$ 矩陣，行列式是有向體積的比例因子，並由三個性質所刻劃：它對每一行都是線性的，兩行對調時它會變號，而且它在單位矩陣上等於 1。單憑這三點，你就能推導出熟悉的展開公式，而高斯消元法則提供了一個高效的計算方法。行列式出現在多重積分的換元公式（雅可比行列式，Jacobian）之中，出現在逆矩陣的公式之中，並且如我們在 Section 5.4 所見，也出現在多項式的判別式（discriminant）之中。

思維技巧

> **用特殊輸入探測未知函數**
>
> 當一個函數只能透過某個性質來認識時，把這個性質應用於你能想到的最簡單的特殊輸入，看看它會迫使甚麼結果成立。
>
> 在這裏，單位矩陣、一個對調、一個縮放和一個錯切，已足以確定幾乎一切。選取那些等於自身的平方、逆或共軛的輸入，尤其有效。

延伸練習

1. **其他例子。**證明 $\mathrm{f}(\mathbf{M}) = (\det \mathbf{M})^3$ 滿足題目的所有條件，並證明 $\mathrm{f}(\mathbf{M}) = |\det \mathbf{M}|$ 滿足乘法性質，但不滿足 $\mathrm{f}(\mathbf{J}) \neq 1$。

2. **列運算。**證明對於任何第二列是第一列的倍數的矩陣，f 的值同樣為零。

3. **縮放。**設 $\mathbf{K}_k = \begin{pmatrix} 1 & 0 \\ 0 & k \end{pmatrix}$。證明 $\mathrm{f}\begin{pmatrix} k & 0 \\ 0 & 1 \end{pmatrix} = \mathrm{f}(\mathbf{K}_k)$，並證明 $\mathrm{f}(\mathbf{K}_k)\mathrm{f}(\mathbf{K}_l) = \mathrm{f}(\mathbf{K}_{kl})$。

4. **歸結為行列式。**較難 證明每個可逆的 2×2 矩陣都可以寫成形如 $\mathbf{J}$、$\mathbf{K}_k$ 和 $\begin{pmatrix} 1 & t \\ 0 & 1 \end{pmatrix}$ 的矩陣的乘積。由此推導出 $\mathrm{f}(\mathbf{M})$ 只取決於 $\det \mathbf{M}$。

延伸閱讀

- **Sheldon Axler, *Linear Algebra Done Right***（Springer；網上免費閱覽）。從行列式的性質出發來定義行列式，與本題的精神一脈相承。
- **Gilbert Strang, *Introduction to Linear Algebra***（Wellesley-Cambridge）。第 5 章從三條規則出發建立行列式的理論。

- **3Blue1Brown,「The determinant」**（影片，網上免費觀看）。以精美的動畫呈現行列式的面積詮釋。

7.4 回到起點的矩陣

題目

STEP II 2020，第 6 題

繼續閱讀之前，請先看題目（原題為英文）。題目載於 OCR 提供下載的 2020 年 STEP 試卷（zip 檔案）內的 STEP II 試卷：

`www.ocr.org.uk/Images/696589-2020-past-papers.zip`

引子

有些變換（transformation）只要重複足夠多次，就會令一切回到原位。旋轉 90° 要四步，反射則只要兩步。哪些 2×2 矩陣（matrix）具有這種性質，即對某個 n 有 $\mathbf{M}^n = \mathbf{I}$？又怎樣判斷？

這道題令人意外之處，在於只需兩個數就足以判定：跡（trace）和行列式（determinant）。你不必知道那四個元素，只需知道 $a + d$ 和 $ad - bc$。而且答案對現實世界有一個驚人的推論：它解釋了為甚麼晶體可以有 2 重、3 重、4 重和 6 重對稱，卻永遠沒有 5 重對稱。這條晶體學「定律」屹立了兩個世紀，直至 1982 年的一項發現才被打破，而該發現更贏得了諾貝爾獎。

動筆之前

想一想 每個 2×2 矩陣都滿足 $\mathbf{M}^2 - (\operatorname{tr}\mathbf{M})\mathbf{M} + (\det\mathbf{M})\mathbf{I} = \mathbf{0}$（這是凱萊—哈密頓定理（Cayley–Hamilton theorem）；把它乘開來驗證一下）。這對處理有關 $\mathbf{M}^2$ 的問題有甚麼幫助？

它把 $\mathbf{M}^2$ 用 $\mathbf{M}$ 和 $\mathbf{I}$ 表示出來：$\mathbf{M}^2 = t\mathbf{M} - d\mathbf{I}$，其中 $t = \operatorname{tr}\mathbf{M}$，$d = \det\mathbf{M}$。於是，關於 $\mathbf{M}^2$ 的條件就變成一個把 $\mathbf{M}$、t 和 d 聯繫起來的條件。這就是整道題的關鍵。

逐步解題

階段 1　平方的跡

$\mathbf{M}^2 = \begin{pmatrix} a^2+bc & ab+bd \\ ac+cd & bc+d^2 \end{pmatrix}$，所以 $\operatorname{tr}(\mathbf{M}^2) = a^2 + 2bc + d^2 = (a+d)^2 - 2(ad-bc)$。 ■

階段 2　$\pm\mathbf{I}$ 的平方根

> **想一想**　假設 $\mathbf{M}^2 = \mathbf{I}$。把它代入 $\mathbf{M}^2 = t\mathbf{M} - d\mathbf{I}$。若 $t \neq 0$，這說明 $\mathbf{M}$ 是怎樣的矩陣？

若 $\mathbf{M}^2 = \mathbf{I}$，則 $\mathbf{I} = t\mathbf{M} - d\mathbf{I}$，所以 $t\mathbf{M} = (1+d)\mathbf{I}$。若 $t \neq 0$，則 $\mathbf{M}$ 是 $\mathbf{I}$ 的倍數，設為 $c\mathbf{I}$，而 $c^2 = 1$，所以 $\mathbf{M} = \pm\mathbf{I}$，但這已被排除。因此 $t = 0$，於是 $(1+d)\mathbf{I} = \mathbf{0}$，所以 $d = -1$。

反過來，若 $t = 0$ 且 $d = -1$，則 $\mathbf{M}^2 = t\mathbf{M} - d\mathbf{I} = \mathbf{I}$，而且 $\mathbf{M} \neq \pm\mathbf{I}$，因為 $\operatorname{tr}(\pm\mathbf{I}) = \pm 2 \neq 0$。 ■

至於 $\mathbf{M}^2 = -\mathbf{I}$：代入後得 $t\mathbf{M} = (d-1)\mathbf{I}$。若 $t \neq 0$，則 $\mathbf{M} = c\mathbf{I}$，其中 $c^2 = -1$，這對實數 c 是不可能的。所以 $t = 0$ 且 $d = 1$。反過來，由 $t = 0$ 和 $d = 1$ 可得 $\mathbf{M}^2 = -\mathbf{I}$。 ■

階段 3　四次方根

> **停一停，試一試**　設 $\mathbf{N} = \mathbf{M}^2$，則 $\mathbf{M}^4 = \mathbf{I}$ 即是 $\mathbf{N}^2 = \mathbf{I}$。根據第 (ii) 部分，要麼 $\mathbf{N} = \pm\mathbf{I}$，要麼 $\operatorname{tr}\mathbf{N} = 0$ 且 $\det\mathbf{N} = -1$。為甚麼第二種情況不可能發生？

$\det\mathbf{N} = \det(\mathbf{M}^2) = (\det\mathbf{M})^2 \geq 0$，所以 $\det\mathbf{N} \neq -1$。因此 $\mathbf{N} = \mathbf{M}^2 = \pm\mathbf{I}$。逆命題顯然成立。 ■

至於 $\mathbf{M}^4 = -\mathbf{I}$：把第 (ii) 部分應用於 $\mathbf{N} = \mathbf{M}^2$，可知它成立當且僅當 $\operatorname{tr}(\mathbf{M}^2) = 0$ 且 $\det(\mathbf{M}^2) = 1$，即 $t^2 - 2d = 0$ 且 $d^2 = 1$。由於 $t^2 = 2d \geq 0$，必須有 $d = 1$，於是 $t^2 = 2$。所以

$$\mathbf{M}^4 = -\mathbf{I} \iff \det\mathbf{M} = 1 \text{ 且 } \operatorname{tr}\mathbf{M} = \pm\sqrt{2}. \qquad \blacksquare$$

階段 4　一個不是旋轉的 I 的八次方根

任何滿足 $\mathbf{M}^4 = -\mathbf{I}$ 的矩陣都有 $\mathbf{M}^8 = \mathbf{I}$。旋轉 45° 的矩陣的跡是 $\sqrt{2}$，行列式是 1；我們想找另一個跡和行列式都相同、但不是旋轉的矩陣。取

$$\mathbf{M} = \begin{pmatrix} \sqrt{2} & -1 \\ 1 & 0 \end{pmatrix}, \qquad \operatorname{tr}\mathbf{M} = \sqrt{2}, \quad \det\mathbf{M} = 1.$$

根據第 3 步，$\mathbf{M}^4 = -\mathbf{I}$，所以 $\mathbf{M}^8 = \mathbf{I}$。但 $\mathbf{M}$ 既不是旋轉也不是反射，因為旋轉和反射都保持長度不變，而 $\mathbf{M}\begin{pmatrix}1\\0\end{pmatrix} = \begin{pmatrix}\sqrt{2}\\1\end{pmatrix}$ 的長度是 $\sqrt{3} \neq 1$。 ■

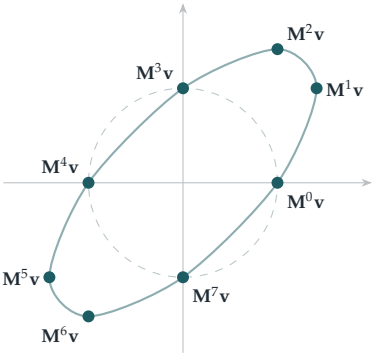


圖 7.4.1 $\mathbf{v} = (1,0)$ 在 $\mathbf{M} = \begin{pmatrix}\sqrt{2} & -1\\1 & 0\end{pmatrix}$ 作用下的軌道（orbit）。八步之後它回到起點。這些點並不在同一個圓（虛線）上，所以 $\mathbf{M}$ 不是旋轉，但它們確實位於一個橢圓上：$\mathbf{M}$ 是在斜坐標下看到的 45° 旋轉。

回顧

- **凱萊－哈密頓定理把冪化為線性組合（linear combination）。**2×2 矩陣的每一個冪都是 $\mathbf{M}$ 和 $\mathbf{I}$ 的組合，其係數只取決於跡和行列式。
- **平方的行列式是非負的。**這個簡單的事實在第 (iii) 部分排除了一整種情況。
- **跡和行列式是不變量（invariant）。**兩個跡和行列式都相同的矩陣（且都不是 $\mathbf{I}$ 的倍數），不論外表是否相似，取冪時的表現都一樣。

背後的數學

相似矩陣

為甚麼說第 4 步的 $\mathbf{M}$ 是「偽裝了的旋轉」？若存在可逆（invertible）矩陣 $\mathbf{S}$ 使得 $\mathbf{M} = \mathbf{SRS}^{-1}$，則稱兩個矩陣 $\mathbf{M}$ 和 $\mathbf{R}$ *相似*（similar）：它們在不同的坐標下表示同一個變換。相似矩陣有相同的跡和行列式，而且（對於不是 $\mathbf{I}$ 的倍數的 2×2 矩陣）逆命題也成立。所以任何跡為 $\sqrt{2}$、行列式為 1 的實矩陣都與 45° 的旋轉相似，而在 $\mathbf{S}$ 所給出的斜坐標下，它就只是在旋轉（圖 7.4.1）。一般而言，滿足 $\mathbf{M}^n = \mathbf{I}$ 的 2×2 實矩陣必定與某個旋轉或反射相似。

為甚麼晶體沒有五重對稱

晶體是原子有規律、重複的排列。晶體的旋轉對稱把原子所成的格（lattice）映射到它自身，所以在以這個格為基礎的坐標下，它由一個元素全是整數的矩陣給出。因此它的跡是整數。但是旋轉角為 θ 的旋轉，在任何坐標下的跡都是 $2\cos\theta$，而 $2\cos\theta$ 只有在等於 $-2, -1, 0, 1$ 或 2 時才是整數，也就是對應階（order）為 2、3、4 和 6（以及 1）的旋轉。五重對稱需要 $2\cos 72^\circ = \frac{\sqrt{5}-1}{2}$，而這不是整數。這就是*晶體學限制定理*（crystallographic restriction theorem）。1982 年，Dan Shechtman 觀察到一種看似具有五重對稱的合金，起初還遭人嘲笑；後來才知道那是一種*準晶體*（quasicrystal），有序但沒有周期性，就像彭羅斯鋪磚（Penrose tilings）一樣，因此不在該定理的涵蓋範圍之內。他於 2011 年獲頒諾貝爾化學獎。

思維技巧

> **尋找能刻劃本質行為的不變量**
>
> *當題目涉及一個有很多參數的對象時，問問自己：是否只需這些參數的幾個組合，即那些轉換觀點後仍然不變的組合，就已經足夠？*
>
> 在這裏，四個元素化約為兩個不變量，即跡和行列式，因為這個問題（「$\mathbf{M}^n = \mathbf{I}$ 是否成立？」）並不取決於坐標。找到合適的不變量，往往是解決分類問題的關鍵。

延伸練習

1. **三階的旋轉。**試做 STEP II 2021，第 7 題（載於 OCR 提供下載的 2021 年試卷，`www.ocr.org.uk/Images/696588-2021-past-papers.zip`）。它從一個具有不尋常性質的旋轉矩陣出發，最後得出一個結論，適用於所有立方為單位矩陣（identity matrix）的 2×2 實矩陣。

2. **六階。**求跡和行列式須滿足的條件，使得 $\mathbf{M}^6 = \mathbf{I}$，但 $\mathbf{M}^2 \neq \mathbf{I}$ 且 $\mathbf{M}^3 \neq \mathbf{I}$。

3. **晶體學限制。**證明：若 $\mathbf{M}$ 的元素全是整數，$\det \mathbf{M} = 1$，且對某個 n 有 $\mathbf{M}^n = \mathbf{I}$，則 $\operatorname{tr} \mathbf{M} \in \{-2, -1, 0, 1, 2\}$。找出一個階為 6 的整數矩陣。

4. **以凱萊－哈密頓定理求冪。**利用 $\mathbf{M}^2 = t\mathbf{M} - d\mathbf{I}$，把 $\mathbf{M}^3$ 和 $\mathbf{M}^4$ 表示成 $p\mathbf{M} + q\mathbf{I}$ 的形式，並藉此重新推導 $\mathbf{M}^4 = -\mathbf{I}$ 的條件。

延伸閱讀

- **M. A. Armstrong, *Groups and Symmetry***（Springer）。內容包括晶體學限制和牆紙圖案的分類。
- **Marjorie Senechal, *Quasicrystals and Geometry***（Cambridge University Press）。準晶體與非周期有序結構的數學。
- **Gilbert Strang, *Introduction to Linear Algebra***（Wellesley-Cambridge）。相似矩陣與凱萊－哈密頓定理。

7.5 三維空間中的旋轉

題目

STEP III 2022，第 7 題

繼續閱讀之前，請先看題目（原題為英文）。題目載於 OCR 提供下載的 2022 年 STEP 試卷（zip 檔案）內的 STEP III 試卷：

`www.ocr.org.uk/Images/696587-2022-past-papers.zip`

引子

在二維空間中，旋轉（rotation）一個點很容易：乘以 $e^{i\theta}$ 便可。到了三維空間，事情就困難得多。沒有一個複數可以讓我們直接相乘，繞不同軸的旋轉不可交換（commute），而要寫出「繞通過 $(1,2,2)$ 的軸旋轉 $37°$」的矩陣，看來更是令人卻步。

本題利用叉積（cross product），為任意的軸和任意的角構造出這個矩陣。所得的結果就是*羅德里格旋轉公式*（Rodrigues' rotation formula），由羅德里格（Olinde Rodrigues）於 1840 年發表。機械人技術、電腦圖像和太空船導航每天都用到它，而它亦揭示了一個優美的事實：與 $\mathbf{n}$ 作叉積，就是「繞 $\mathbf{n}$ 旋轉 $90°$」，而每一個繞 $\mathbf{n}$ 的旋轉都由它構成，正如 $e^{i\theta} = \cos\theta + i\sin\theta$ 是由乘以 i 構成一樣。

動筆之前

想一想 當 $\mathbf{r}$ 垂直於 $\mathbf{n}$ 時，$\mathbf{n}\times\mathbf{r}$ 在幾何上是甚麼樣子？當 $\mathbf{r}$ 平行於 $\mathbf{n}$ 時又如何？

若 $\mathbf{r}\parallel\mathbf{n}$，則 $\mathbf{n}\times\mathbf{r}=\mathbf{0}$。若 $\mathbf{r}\perp\mathbf{n}$，則 $\mathbf{n}\times\mathbf{r}$ 的長度與 $\mathbf{r}$ 相同（因為 $|\mathbf{n}|=1$，而夾角是 $90°$），與兩者皆垂直，而且根據右手定則（right-hand rule），它就是 $\mathbf{r}$ 繞 $\mathbf{n}$ 逆時針旋轉 $90°$ 的結果（沿 $\mathbf{n}$ 向下看）。因此 f 消去 $\mathbf{r}$ 沿 $\mathbf{n}$ 方向的分量（component），並把垂直的分量旋轉 $90°$。這幅圖像足以解釋整條題目。

逐步解題

階段 1　作兩次叉積

$\mathrm{f}(\mathbf{r}) = \mathbf{n}\times\mathbf{r} = (bz-cy,\ cx-az,\ ay-bx)$。再與 $\mathbf{n}=(a,b,c)$ 作一次叉積，$\mathrm{f}(\mathrm{f}(\mathbf{r}))$ 的 x 分量是

$$b(ay-bx)-c(cx-az) = -x(b^2+c^2)+aby+acz. \qquad \blacksquare$$

由於 $a^2+b^2+c^2=1$，上式等於 $-x(1-a^2)+aby+acz = a(ax+by+cz)-x$，這正是 $(\mathbf{n}.\mathbf{r})\mathbf{n}-\mathbf{r}$ 的 x 分量。其餘分量可由對稱性得出，所以 $\mathrm{f}(\mathrm{f}(\mathbf{r})) = (\mathbf{n}.\mathbf{r})\mathbf{n}-\mathbf{r}$。 $\blacksquare$

> **想一想**　設 $\mathbf{r}=\mathbf{r}_{\parallel}+\mathbf{r}_{\perp}$，其中 $\mathbf{r}_{\parallel}=(\mathbf{n}.\mathbf{r})\mathbf{n}$ 是沿 $\mathbf{n}$ 方向的分量。用這兩部分表示，$\mathrm{f}(\mathrm{f}(\mathbf{r}))$ 是甚麼？

$\mathrm{f}(\mathrm{f}(\mathbf{r})) = \mathbf{r}_{\parallel}-\mathbf{r} = -\mathbf{r}_{\perp}$。施行 f 兩次，會消去平行的部分，並把垂直的部分繞 $\mathbf{n}$ 旋轉 $180°$，亦即把它反向（圖 7.5.1）。這就是 $\mathrm{i}^2=-1$ 的類比。

階段 2　旋轉

> **想一想**　計算 $\mathrm{g}(\mathbf{n})$。然後對於 $\mathbf{r}\perp\mathbf{n}$ 的情況計算 $\mathrm{g}(\mathbf{r})$，並利用 $\mathrm{f}(\mathrm{f}(\mathbf{r}))=-\mathbf{r}$ 對這樣的 $\mathbf{r}$ 成立這一點。

$\mathrm{f}(\mathbf{n}) = \mathbf{n}\times\mathbf{n} = \mathbf{0}$，所以 $\mathrm{g}(\mathbf{n})=\mathbf{n}$：軸保持不動。若 $\mathbf{r}\perp\mathbf{n}$，則 $\mathrm{f}(\mathrm{f}(\mathbf{r}))=-\mathbf{r}$，於是

$$\mathrm{g}(\mathbf{r}) = \mathbf{r}+\sin\theta\,(\mathbf{n}\times\mathbf{r})-(1-\cos\theta)\mathbf{r} = \cos\theta\,\mathbf{r}+\sin\theta\,(\mathbf{n}\times\mathbf{r}).$$

由於 $\mathbf{n}\times\mathbf{r}$ 是 $\mathbf{r}$ 繞 $\mathbf{n}$ 轉過 $90°$ 的結果，長度不變，所以上式就是 $\mathbf{r}$ 繞 $\mathbf{n}$ 旋轉 θ 的結果，與平面上的 $(\cos\theta+\mathrm{i}\sin\theta)z$ 如出一轍。由於 g 是線性（linear）的，而每個向量都可寫成 $\mathbf{n}$ 的倍數與一個垂直於 $\mathbf{n}$ 的向量之和，所以 g 就是**繞通過原點、方向為 $\mathbf{n}$ 的軸旋轉角 θ 的旋轉**（從 $\mathbf{n}$ 的尖端望去為逆時針方向）。 $\blacksquare$

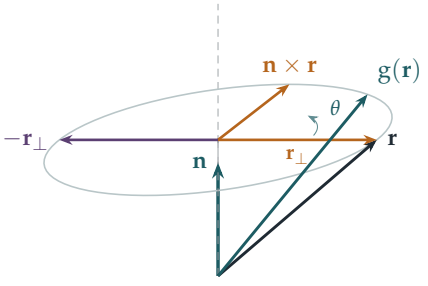


圖 7.5.1　羅德里格公式。向量 $\mathbf{r}$ 分解為沿軸 $\mathbf{n}$ 的部分和與軸垂直的部分 $\mathbf{r}_{\perp}$，後者掃出一個圓。與 $\mathbf{n}$ 作叉積，會把 $\mathbf{r}_{\perp}$ 轉過 $90°$（琥珀色）；作兩次叉積則得 $-\mathbf{r}_{\perp}$（紫紅色）。旋轉後的向量 $\mathrm{g}(\mathbf{r})$ 是 $\mathbf{r}_{\parallel}+\cos\theta\,\mathbf{r}_{\perp}+\sin\theta\,(\mathbf{n}\times\mathbf{r})$。

階段 3 反射

> **停一停，試一試** 利用第 (i) 部分化簡 $\mathrm{h}(\mathbf{s}) = -\mathbf{s} - 2\mathrm{f}(\mathrm{f}(\mathbf{s}))$。它對沿 $\mathbf{n}$ 方向的向量有甚麼作用？對垂直於 $\mathbf{n}$ 的向量又有甚麼作用？

$\mathrm{h}(\mathbf{s}) = -\mathbf{s} - 2\big((\mathbf{n}.\mathbf{s})\mathbf{n} - \mathbf{s}\big) = \mathbf{s} - 2(\mathbf{n}.\mathbf{s})\mathbf{n}$。垂直於 $\mathbf{n}$ 的向量保持不變，而 $\mathrm{h}(\mathbf{n}) = -\mathbf{n}$。所以 h 是**對通過原點且垂直於 n 的平面的反射（reflection）**。 ■

回顧

- **分解為平行和垂直的部分。**一旦把 $\mathbf{r}$ 分解為沿 $\mathbf{n}$ 方向和與之垂直的分量，每一步都變得清晰。
- **線性減省工夫。**要辨認一個線性映射（linear map），只需知道它對幾個精心挑選的向量有甚麼作用便足夠。
- **複數的類比。**在垂直於 $\mathbf{n}$ 的向量上，f 的表現就像乘以 i：$\mathrm{f}^2 = -1$，而 $\mathrm{g} = \cos\theta + \sin\theta\,\mathrm{f}$。

背後的數學

歐拉旋轉定理

三維空間中每一個固定原點的旋轉，都是繞某條通過原點的軸的旋轉。歐拉（Euler）於 1775 年證明了這一點。用矩陣的語言來說，每個 3×3 旋轉矩陣都有一個特徵值（eigenvalue）為 1 的特徵向量（eigenvector），也就是旋轉軸。羅德里格公式給出其逆命題：對每一條軸 $\mathbf{n}$ 和每一個角 θ，旋轉矩陣是

$$\mathbf{R} = \mathbf{I} + \sin\theta\,\mathbf{K} + (1-\cos\theta)\,\mathbf{K}^2, \qquad \mathbf{K} = \begin{pmatrix} 0 & -c & b \\ c & 0 & -a \\ -b & a & 0 \end{pmatrix},$$

其中 $\mathbf{K}$ 是 $\mathbf{r} \mapsto \mathbf{n}\times\mathbf{r}$ 的矩陣。由於 $\mathbf{K}^3 = -\mathbf{K}$，指數級數（exponential series）$e^{\theta\mathbf{K}} = \mathbf{I} + \theta\mathbf{K} + \frac{\theta^2}{2!}\mathbf{K}^2 + \cdots$ 恰好化簡成這條公式，它就是 $e^{\mathrm{i}\theta} = \cos\theta + \mathrm{i}\sin\theta$ 的三維版本。旋轉與「無窮小旋轉」$\mathbf{K}$ 的指數之間的這種聯繫，正是李羣（Lie groups）理論的起點。

旋轉不可交換

把一本書平放，先繞垂直軸旋轉 90°，再繞水平軸旋轉 90°；然後從頭再來，以相反的次序做這兩個旋轉。書最終的位置並不相同。三維空間中旋轉的複合（composition）不符合交換律，與二維的情況不同。這正是 Section 6.7 的四元數（quaternion）有用的原因之一：繞 $\mathbf{n}$ 旋轉 θ 對應於四元數 $\cos\frac{\theta}{2} + \sin\frac{\theta}{2}(a\mathrm{i} + b\mathrm{j} + c\mathrm{k})$，而複合旋轉就是把四元數相乘，不可交換性已包含在其中。羅德里格本人在 1840 年的論文中，已找到旋轉的複合公式，比哈密頓（Hamilton）發明四元數還早三年。

思維技巧

> **分解為不變的部分**
>
> 要理解一個變換，可把空間分成變換處理得簡單的幾塊：在這裏，就是軸（保持不動）和與軸垂直的平面（被旋轉）。
>
> 這與 Section 7.2 中特徵向量的想法相同：找出變換所保持的子空間（subspace），然後在每一塊上分別研究它。

延伸練習

1 **平面上的旋轉。**試做 STEP III 2017，第 2 題（`step.maths.org/questions/17-s3-q2`）。該題利用複數，證明平面上的兩個旋轉複合起來仍是一個旋轉，只有一種例外。找出這個例外，以及此時複合的結果是甚麼。

2 **一個具體的旋轉。**求繞方向為 $(1,1,1)$ 的軸旋轉 120° 的矩陣。證明它把各坐標軸循環置換。

3 **K 的冪。**證明 $\mathbf{K}^2\mathbf{r} = (\mathbf{n}.\mathbf{r})\mathbf{n} - \mathbf{r}$ 及 $\mathbf{K}^3 = -\mathbf{K}$，並由指數級數推導出羅德里格公式。

4 **兩次反射構成一個旋轉。**證明先後對兩個通過原點、互成 α 角的平面作反射，結果是繞兩平面的交線旋轉 2α。

延伸閱讀

- **Simon Altmann, *Rotations, Quaternions, and Double Groups***（Dover）。關於三維旋轉的權威論著，當中包括羅德里格的貢獻。
- **Simon Altmann, "Hamilton, Rodrigues, and the quaternion scandal"**, *Mathematics Maga-*

zine 62 (1989)。一篇簡短而生動的歷史敘述。

- **John Stillwell, *Naive Lie Theory***（Springer）。為大學本科生而寫的旋轉、四元數與李羣入門。

7.6 雙曲旋轉與相對論

題目

STEP III 1991，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/91-s3-q6`

引子

旋轉矩陣（rotation matrix）$\begin{pmatrix} \cos u & -\sin u \\ \sin u & \cos u \end{pmatrix}$ 把點沿着圓 $x^2 + y^2 = 1$ 移動，而先旋轉 u 再旋轉 w，就等於旋轉 $u + w$。把 cos 和 sin 換成 cosh 和 sinh，再調整一個正負號，便得到本題的矩陣。它改為把點沿着雙曲線（hyperbola）$x^2 - y^2 = 1$ 移動，而且它的「角」同樣可以相加。這就是*雙曲旋轉*（hyperbolic rotation）。

這並非只是一件趣聞。若以 x 代表時間、y 代表距離（採用光速為 1 的單位），這個矩陣就是*勞侖茲推進*（Lorentz boost）：它是愛因斯坦（Einstein）狹義相對論（special relativity）中發現的變換，聯繫着兩個以恆定相對速度運動的觀察者各自的量度結果。「雙曲角」u 稱為*快度*（rapidity），而快度可以相加這一事實，正好解釋了我們在 Section 3.8 遇過的那條奇怪的速度相加定律。

動筆之前

想一想 $\cosh u$ 和 $\sinh u$ 滿足哪一條恆等式，與 $\cos^2 u + \sin^2 u = 1$ 相對應？$\cosh(u+v)$ 和 $\sinh(u+v)$ 的加法公式又是甚麼？

$\cosh^2 u - \sinh^2 u = 1$，以及

$$\cosh(u+v) = \cosh u \cosh v + \sinh u \sinh v,$$
$$\sinh(u+v) = \sinh u \cosh v + \cosh u \sinh v.$$

你可以由定義 $\cosh u = \frac{1}{2}(e^u + e^{-u})$ 和 $\sinh u = \frac{1}{2}(e^u - e^{-u})$ 驗證這些公式。點

$(\cosh v, \sinh v)$ 位於雙曲線 $x^2 - y^2 = 1$ 上，正如 $(\cos v, \sin v)$ 位於圓上一樣。

逐步解題

階段 1　雙曲線保持不變

$$\begin{aligned} x'^2 - y'^2 &= (x\cosh u + y\sinh u)^2 - (x\sinh u + y\cosh u)^2 \\ &= (x^2 - y^2)(\cosh^2 u - \sinh^2 u) = x^2 - y^2, \end{aligned}$$

因為交叉項 $2xy\cosh u \sinh u$ 互相抵消。所以 T 保持 $x^2 - y^2$ 不變，特別是把雙曲線 $x^2 - y^2 = 1$ 映射到它自身。 ■

> **想一想**　哪些通過原點的直線會被映射到自身？（試找特徵向量（eigenvector）。）

$T\begin{pmatrix}1\\1\end{pmatrix} = (\cosh u + \sinh u)\begin{pmatrix}1\\1\end{pmatrix} = e^u\begin{pmatrix}1\\1\end{pmatrix}$，而 $T\begin{pmatrix}1\\-1\end{pmatrix} = e^{-u}\begin{pmatrix}1\\-1\end{pmatrix}$。所以直線 $y = x$ 和 $y = -x$ 各自被映射到自身：一條被拉長 e^u 倍，另一條被縮短至 e^{-u} 倍。這兩條直線正是雙曲線的漸近線（asymptote）。 ■

階段 2　雙曲角相加

斜率（gradient）為 $\tanh v$ 的直線，其方向向量為 $(\cosh v, \sinh v)$。它的像的方向為

$$T\begin{pmatrix}\cosh v\\ \sinh v\end{pmatrix} = \begin{pmatrix}\cosh u\cosh v + \sinh u\sinh v\\ \sinh u\cosh v + \cosh u\sinh v\end{pmatrix} = \begin{pmatrix}\cosh(u+v)\\ \sinh(u+v)\end{pmatrix},$$

其斜率為 $\tanh(u+v)$。線性映射（linear map）把平行線變成平行線，所以無論直線是否通過原點，這個結論都成立。因此 $v' = v + u$。 ■

階段 3　夾角何時保持不變？

> **想一想**　斜率分別為 m_1、m_2 的兩條直線，其夾角 ψ 滿足 $\tan\psi = \left|\frac{m_1 - m_2}{1 + m_1 m_2}\right|$。當 $m_i = \tanh v_i$ 時，化簡這個式子。（提示：把所有項通分，以 $\cosh v_1 \cosh v_2$ 為分母。）

利用 $\tanh v_1 - \tanh v_2 = \dfrac{\sinh(v_1 - v_2)}{\cosh v_1 \cosh v_2}$ 和 $1 + \tanh v_1 \tanh v_2 = \dfrac{\cosh(v_1 + v_2)}{\cosh v_1 \cosh v_2}$，得

$$\tan\psi = \left|\frac{\sinh(v_1 - v_2)}{\cosh(v_1 + v_2)}\right|.$$

變換之後，$v_1 - v_2$ 不變，而 $v_1 + v_2$ 則變成 $v_1 + v_2 + 2u$。

> **停一停，試一試** 甚麼時候 $\cosh(v_1 + v_2) = \cosh(v_1 + v_2 + 2u)$？

假設兩條直線不平行（$v_1 \neq v_2$），則兩個夾角相等，當且僅當 $\cosh(v_1 + v_2) = \cosh(v_1 + v_2 + 2u)$。由於 cosh 是偶函數，並且在 $[0, \infty)$ 上嚴格遞增，這表示 $v_1 + v_2 = \pm(v_1 + v_2 + 2u)$。取正號得 $u = 0$，而這已被排除，所以

$$v_1 + v_2 = -u. \qquad \blacksquare$$

階段 4 圖像

當 $v_1 + v_2 = -u$ 時，變換後的角為 $v_1' = v_1 + u = -v_2$ 和 $v_2' = v_2 + u = -v_1$。所以 ℓ_1 的像的斜率為 $-\tanh v_2$：它是 ℓ_2 對 x 軸的反射，而 ℓ_2 的像則是 ℓ_1 的反射。對於通過原點的直線，這個變換把兩條直線對調並加以反射，這就解釋了為何它們之間的夾角不變（圖 7.6.1）。

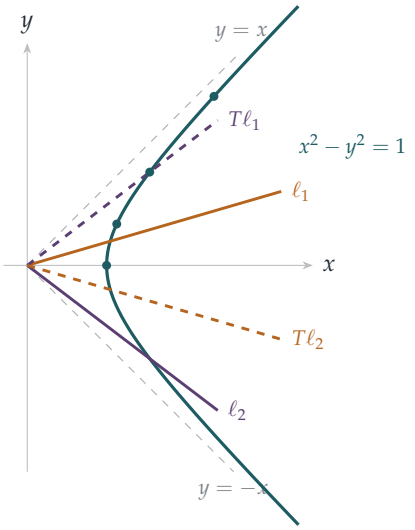


圖 7.6.1 把 $u = 0.7$ 的雙曲旋轉施於通過原點的直線 ℓ_1、ℓ_2，其中 $v_1 = 0.3$，$v_2 = -1.0$，因此 $v_1 + v_2 = -u$。ℓ_1 的像（紫紅色虛線）是 ℓ_2 的鏡像，而 ℓ_2 的像（琥珀色虛線）是 ℓ_1 的鏡像，所以兩線之間的夾角保持不變。雙曲線上的藍綠色點按雙曲角等距分佈（$v = 0, 0.5, 1, 1.5$）；T 把每一點沿曲線再移前 0.7，到達新的位置。

回顧

- **不變的二次型（quadratic form）。** T 保持 $x^2 - y^2$ 不變，正如旋轉保持 $x^2 + y^2$ 不變。

- **特徵向量就是漸近線。**保持不變的直線是 $y = \pm x$，分別按 $e^{\pm u}$ 的倍數拉長和縮短。
- **雙曲角可以相加。**以 $\tanh v$ 把方向參數化之後，T 的作用便化為簡單的加法 $v \mapsto v + u$。

背後的數學

狹義相對論

在愛因斯坦的狹義相對論（1905 年）中，以恆定相對速度運動的不同觀察者，對時間和距離的量度並不一致，但他們對 $t^2 - x^2$ 這個量（取光速 $c = 1$）卻意見一致，這個量稱為時空間隔（spacetime interval）。保持它不變的變換就是勞侖茲變換（Lorentz transformation），而在一個空間維度的情況下，它們恰好就是本題的矩陣，其中 x 為時間、y 為距離。一個以速度 $v = \tanh u$ 運動的觀察者，所用的坐標與我們的坐標就是由這個矩陣聯繫起來，其中 $\cosh u = \frac{1}{\sqrt{1-v^2}}$，即時間膨脹（time dilation）中著名的因子 γ。閔可夫斯基（Hermann Minkowski）在 1908 年看出，這使時空成為一種四維幾何，其「距離」之中帶有一個負號：「從今以後，空間本身和時間本身，都注定要消退成為影子。」

題目的物理意義

- **不變直線** $y = \pm x$ 是光線的路徑。它們對每個觀察者都相同：光速對所有人都一樣，這正是相對論的奠基原理。
- **雙曲線** $x^2 - y^2 = 1$ 是與原點相距一個單位「原時」（proper time）的所有事件所成的集合，而原時是由以各種可能的恆定速度運動的觀察者所量度的。
- **快度相加。**通過原點而斜率為 $\tanh v$ 的直線，是一個以速度 $\tanh v$ 運動的物體的路徑。快度為 u 的推進把它的速度變為 $\tanh(v+u) = \frac{\tanh v + \tanh u}{1 + \tanh v \tanh u}$，這就是 Section 3.8 中愛因斯坦的速度相加公式。速度不能直接相加，快度卻可以。

而在 Section 5.8 中，我們見過同一條加法公式，它是龐加萊圓盤（Poincaré disc）中的距離法則：相對論中所有可能的速度所成的集合，以快度來量度，就是一個雙曲空間（hyperbolic space）。

思維技巧

找出合適的參數化

選用一個參數來描述你的對象，使變換在這個參數之下變得簡單。

若以斜率描述方向，T 的作用是一條繁複的公式；若以 v 描述方向，其中斜率為 $\tanh v$，它就只是加上 u。同樣的想法，曾令旋轉在使用角度時變得簡單，也曾在 Section 2.8 中令映射 $x \mapsto x^2 - 2$ 變得簡單。

延伸練習

1 **推進的複合。**證明 $T(u)T(w) = T(u+w)$，其中 $T(u)$ 是題目中的矩陣。$T(u)^{-1}$ 是甚麼？

2 **時間膨脹。**一艘太空船以光速的 0.8 倍航行。求它的快度 u 和 $\cosh u$。若地球上過了 10 年，船上過了多少時間？（利用 $t_{\text{船}} = t_{\text{地球}} / \cosh u$。）

3 **雙曲面積。**證明由雙曲線 $x^2 - y^2 = 1$、x 軸，以及由原點至 $(\cosh v, \sinh v)$ 的線段所圍成的面積為 $\frac{v}{2}$，正如圓的相應面積為 $\frac{\theta}{2}$。為甚麼 T 必定保持這類面積不變？

4 **擠壓。**把坐標轉換為 $p = x + y$，$q = x - y$。證明 T 變成 $p \mapsto e^u p$，$q \mapsto e^{-u} q$，而雙曲線則變成 $pq = 1$。試在這組坐標下，以幾何方式描述這個變換。

延伸閱讀

- **Edwin Taylor & John Wheeler, *Spacetime Physics***（W. H. Freeman）。以快度和時空圖講解狹義相對論，適合具備高級程度（A-level）數學基礎的學生。
- **Tevian Dray, *The Geometry of Special Relativity***（CRC Press）。把相對論看成雙曲三角學，與本節的精神完全一致。
- **Albert Einstein, *Relativity: The Special and the General Theory***（多種版本）。愛因斯坦親自撰寫的通俗讀物，至今仍然十分易讀。

7.7 永恆不變的量

題目

STEP III 2024，第 5 題

繼續閱讀之前，請先看題目（原題為英文）。題目載於 OCR 提供下載的 2024 年 STEP 試卷（zip 檔案）內的 STEP III 試卷：

`www.ocr.org.uk/Images/725753-2024-past-papers.zip`

引子

物理學中一些最深刻的事實是*守恆定律*（conservation laws）：無論運動多麼複雜，能量、動量和角動量始終保持不變。當一個系統有足夠多的守恆量時，它的運動可以精確地解出；守恆量太少時，運動則可能是混沌的。

這道題展示了一種製造守恆量的絕妙方法。若一個矩陣按照 $\frac{\mathrm{d}\mathbf{M}}{\mathrm{d}t} = \mathbf{MN} - \mathbf{NM}$ 隨時間變化，那麼無論它的元素變化得多麼劇烈，它的跡（trace）和行列式（determinant），以至它的特徵值（eigenvalues），都永遠不變。這樣的一對 $(\mathbf{M}, \mathbf{N})$ 稱為*拉克斯對*（Lax pair），以彼得·拉克斯（Peter Lax）命名。他在 1968 年提出這個想法，用來解釋*孤立子*（solitons）的驚人表現：孤立子是一種孤立波，彼此穿過之後絲毫不變。

動筆之前

想一想 表達式 $\mathbf{MN} - \mathbf{NM}$ 稱為 $\mathbf{M}$ 與 $\mathbf{N}$ 的*交換子*（commutator）。它在甚麼時候等於零？根據第 (i) 部分，它的跡是多少？

它等於零，當且僅當 $\mathbf{M}$ 與 $\mathbf{N}$ 可交換。而根據第 (i) 部分，它的跡是 $\mathrm{tr}(\mathbf{MN}) - \mathrm{tr}(\mathbf{NM}) = 0$，恆常如此。這一個事實，就是題目中每一條守恆定律的起點。

逐步解題

階段 1　跡

設 $\mathbf{M} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ 及 $\mathbf{N} = \begin{pmatrix} p & q \\ r & s \end{pmatrix}$，則 $\mathbf{MN}$ 的對角線元素是 $ap + br$ 和 $cq + ds$，所以 $\operatorname{tr}(\mathbf{MN}) = ap + br + cq + ds$。把 $\mathbf{M}$ 與 $\mathbf{N}$ 的角色對調，得 $pa + qc + rb + sd$，結果相同。顯然 $\operatorname{tr}(\mathbf{M} + \mathbf{N}) = \operatorname{tr}\mathbf{M} + \operatorname{tr}\mathbf{N}$。 ■

階段 2　行列式的導數

想一想　用積法則求 $\det\mathbf{M} = ad - bc$ 的導數。然後利用 $\mathbf{M}^{-1} = \frac{1}{\det\mathbf{M}}\begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$，計算 $\mathbf{M}^{-1}\frac{\mathrm{d}\mathbf{M}}{\mathrm{d}t}$ 及其跡。

以字母上的一點表示導數，則 $\frac{\mathrm{d}}{\mathrm{d}t}\det\mathbf{M} = \dot{a}d + a\dot{d} - \dot{b}c - b\dot{c}$。而

$$\mathbf{M}^{-1}\dot{\mathbf{M}} = \frac{1}{\det\mathbf{M}}\begin{pmatrix} d & -b \\ -c & a \end{pmatrix}\begin{pmatrix} \dot{a} & \dot{b} \\ \dot{c} & \dot{d} \end{pmatrix}$$

的對角線元素是 $\frac{d\dot{a} - b\dot{c}}{\det\mathbf{M}}$ 和 $\frac{-c\dot{b} + a\dot{d}}{\det\mathbf{M}}$。兩者之和是 $\frac{1}{\det\mathbf{M}}\frac{\mathrm{d}}{\mathrm{d}t}\det\mathbf{M}$。 ■

階段 3　三個守恆量

想一想　對 $\dot{\mathbf{M}} = \mathbf{MN} - \mathbf{NM}$ 取跡。至於行列式，則用第 (ii) 部分：$\mathbf{M}^{-1}(\mathbf{MN} - \mathbf{NM})$ 是甚麼？

跡：$\frac{\mathrm{d}}{\mathrm{d}t}\operatorname{tr}\mathbf{M} = \operatorname{tr}\dot{\mathbf{M}} = \operatorname{tr}(\mathbf{MN}) - \operatorname{tr}(\mathbf{NM}) = 0$。

行列式：根據第 (ii) 部分，

$$\frac{1}{\det\mathbf{M}}\frac{\mathrm{d}}{\mathrm{d}t}\det\mathbf{M} = \operatorname{tr}(\mathbf{M}^{-1}(\mathbf{MN} - \mathbf{NM})) = \operatorname{tr}\mathbf{N} - \operatorname{tr}(\mathbf{M}^{-1}\mathbf{NM}) = \operatorname{tr}\mathbf{N} - \operatorname{tr}\mathbf{N} = 0,$$

當中用到第 (i) 部分的以下形式：$\operatorname{tr}(\mathbf{M}^{-1}(\mathbf{NM})) = \operatorname{tr}((\mathbf{NM})\mathbf{M}^{-1}) = \operatorname{tr}\mathbf{N}$。

平方的跡：$\operatorname{tr}(\mathbf{M}^2) = (\operatorname{tr}\mathbf{M})^2 - 2\det\mathbf{M}$（見 Section 7.4），由於兩項都是常數，所以它也是常數。 ■

階段 4　一個顯式解

停一停，試一試　設 $\mathbf{N} = t\mathbf{I} + t\mathbf{E}$，其中 $\mathbf{E} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$，計算 $\mathbf{MN} - \mathbf{NM}$，並寫出 a、b、c、d 所滿足的微分方程。按以下次序求解：先解 c，然後 a 和 d，最後 b。

$t\mathbf{I}$ 這部分與任何矩陣都可交換，所以 $\mathbf{MN}-\mathbf{NM}=t(\mathbf{ME}-\mathbf{EM})$。現在 $\mathbf{ME}=\begin{pmatrix}0 & a\\ 0 & c\end{pmatrix}$ 而 $\mathbf{EM}=\begin{pmatrix}c & d\\ 0 & 0\end{pmatrix}$，所以

$$\dot a=-tc,\qquad \dot b=t(a-d),\qquad \dot c=0,\qquad \dot d=tc.$$

因此 $c=C$ 是常數；於是 $a=A-\frac12Ct^2$ 及 $d=D+\frac12Ct^2$；繼而 $\dot b=t(A-D-Ct^2)$，所以 $b=B+\frac12(A-D)t^2-\frac14Ct^4$。故此

$$\mathbf{M}=\begin{pmatrix}A-\frac12Ct^2 & B+\frac12(A-D)t^2-\frac14Ct^4\\ C & D+\frac12Ct^2\end{pmatrix}.\qquad\blacksquare$$

驗算：跡是 $A+D$，而行列式是 $(A-\frac12Ct^2)(D+\frac12Ct^2)-C\big(B+\frac12(A-D)t^2-\frac14Ct^4\big)=AD-BC$，一如所料是常數（圖 7.7.1）。

階段 5 另一條方程

想一想 若 $\dot{\mathbf{M}}=\mathbf{MN}$，則 $\mathbf{N}=\mathbf{M}^{-1}\dot{\mathbf{M}}$，而根據第 (ii) 部分，$\operatorname{tr}\mathbf{N}$ 就是 $\det\mathbf{M}$ 的對數導數（logarithmic derivative）。$\mathbf{M}$ 的跡能否保持不變，而它的行列式卻在改變？

當然可以。取 $\mathbf{M}=\begin{pmatrix}1 & t\\ t & 1\end{pmatrix}$，其中 $|t|<1$（此時 $\det\mathbf{M}=1-t^2\neq0$）。它的跡恆等於 2。定義 $\mathbf{N}=\mathbf{M}^{-1}\dot{\mathbf{M}}$，使得 $\dot{\mathbf{M}}=\mathbf{MN}$：

$$\mathbf{N}=\frac{1}{1-t^2}\begin{pmatrix}1 & -t\\ -t & 1\end{pmatrix}\begin{pmatrix}0 & 1\\ 1 & 0\end{pmatrix}=\frac{1}{1-t^2}\begin{pmatrix}-t & 1\\ 1 & -t\end{pmatrix},$$

這是一個非奇異（non-singular）矩陣，而且當 $t\neq0$ 時，$\operatorname{tr}\mathbf{N}=\frac{-2t}{1-t^2}\neq0$。所以答案是**否定的**，$\operatorname{tr}\mathbf{N}$ 不一定等於零。$\blacksquare$

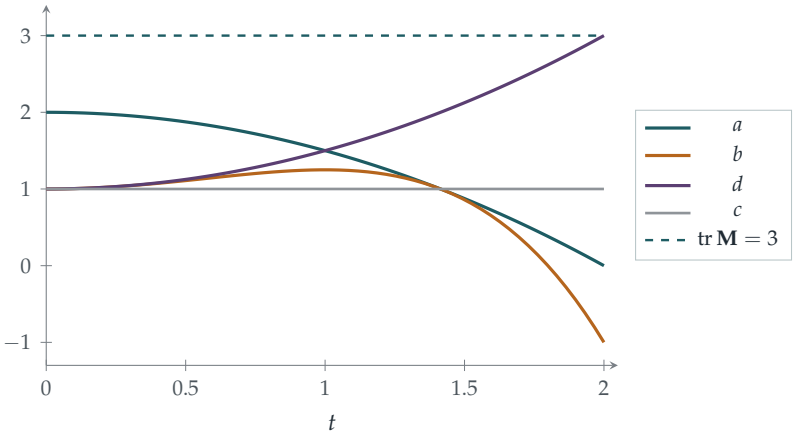


圖 7.7.1　第 4 階段中 $\mathbf{M}(t)$ 的各個元素，其中 $A = 2$、$B = 1$、$C = 1$、$D = 1$。四個元素之中除了 c 以外全都在變，但跡始終是 3，行列式始終是 1。所以特徵值 $\frac{3 \pm \sqrt{5}}{2}$ 永遠不變：這個流是等譜（isospectral）的。

回顧

- **交換子的跡等於零。**由 $\operatorname{tr}(\mathbf{MN}) = \operatorname{tr}(\mathbf{NM})$ 出發，每一條守恆定律都隨之而來。
- **對數導數。**第 (ii) 部分把行列式的導數化為一個跡，而跡是線性的，容易處理。
- **在顯式解中檢查守恆。**我們確認了 $\mathbf{M}(t)$ 的跡和行列式都是常數，藉此驗算它的公式。

背後的數學

孤立子與拉克斯對

1834 年，蘇格蘭工程師 John Scott Russell 策馬追蹤一個沿愛丁堡附近的聯合運河（Union Canal）前進的孤立波，追了超過一英里，波形始終沒有改變。這類波由科特韋赫－德弗里斯（*Korteweg–de Vries*，簡稱 *KdV*）方程描述。到了 1960 年代，Norman Zabusky 與 Martin Kruskal 的電腦實驗顯示，兩個這樣的「孤立子」可以相撞，然後原封不動地離開，彷彿它們是粒子一樣。要解釋這個現象，需要無窮多個守恆量。彼得．拉克斯在 1968 年證明，KdV 方程可以寫成 $\dot{L} = LP - PL$，其中 L 和 P 是一對微分算子（differential operators）；於是，與這道題完全一樣，L 的特徵值是守恆的，而它們正好提供了那無窮多條守恆定律。擁有拉克斯對的系統稱為可積（integrable）系統，它們屬於少數可以精確求解的非線性系統。

等譜流

由於 $\mathbf{M}$ 的跡和行列式決定了它的特徵多項式（characteristic polynomial）$\lambda^2 - (\operatorname{tr}\mathbf{M})\lambda + \det\mathbf{M}$，$\mathbf{M}$ 的特徵值不隨時間改變。流 $\dot{\mathbf{M}} = [\mathbf{M}, \mathbf{N}]$ 使 $\mathbf{M}$ 在特徵值固定的矩陣所組成的集合之內移動。事實上，解總是具有 $\mathbf{M}(t) = \mathbf{S}(t)^{-1}\mathbf{M}(0)\mathbf{S}(t)$ 的形式，其中 $\mathbf{S}(t)$ 是某個可逆矩陣：這個矩陣始終與它的初始值相似（similar）。同一個想法貫穿整個物理學：在量子力學中，可觀測量（observable）$\mathbf{A}$ 按 $\frac{\mathrm{d}\mathbf{A}}{\mathrm{d}t} = \frac{\mathrm{i}}{\hbar}[\mathbf{H}, \mathbf{A}]$ 演化，其中 $\mathbf{H}$ 是能量，所以凡是與能量可交換的量都是守恆的。

雅可比公式

第 (ii) 部分對任何大小的可逆矩陣都成立，稱為雅可比公式（Jacobi's formula）。最穩妥的寫法是不用對數，寫成 $\frac{\mathrm{d}}{\mathrm{d}t}\det\mathbf{M} = \det\mathbf{M}\,\operatorname{tr}\!\left(\mathbf{M}^{-1}\dot{\mathbf{M}}\right)$，因為 $\det\mathbf{M}$ 可能是負數；對數形式 $\frac{\mathrm{d}}{\mathrm{d}t}\ln\det\mathbf{M} = \operatorname{tr}\!\left(\mathbf{M}^{-1}\dot{\mathbf{M}}\right)$ 需要 $\det\mathbf{M} > 0$（一般情況下則用 $\ln|\det\mathbf{M}|$）。由此可得一個推論：若 $\dot{\mathbf{M}} = \mathbf{A}\mathbf{M}$，其中 $\mathbf{A}$ 是常數矩陣，則 $\det\mathbf{M}(t) = e^{t\operatorname{tr}\mathbf{A}}\det\mathbf{M}(0)$。在微分方程的研究中，這個結果告訴你相空間（phase space）中的體積如何增長或縮小；對於經典力學的方程，這個跡等於零，體積保持不變，這個結果稱為劉維爾定理（Liouville's theorem）。

思維技巧

> **找出不變的東西**
>
> 當某樣東西在演化時，試找出由它構成而又保持不變的量。這些量限制了運動，並往往揭示出它的結構。
>
> 在這裏，跡和行列式之所以守恆，是因為變化率是一個交換子。在力學中，守恆量來自對稱性（諾特定理，Noether's theorem），而找出守恆量往往是解題的第一步。

延伸練習

1. **驗算解。**透過求導，驗證第 4 階段的矩陣 $\mathbf{M}(t)$ 滿足 $\dot{\mathbf{M}} = \mathbf{M}\mathbf{N} - \mathbf{N}\mathbf{M}$。

2. **相似性。**證明：若 $\mathbf{M}(t) = \mathbf{S}^{-1}\mathbf{M}_0\mathbf{S}$，其中 $\mathbf{M}_0$ 是常數矩陣，而 $\mathbf{S}(t)$ 滿足 $\dot{\mathbf{S}} = \mathbf{S}\mathbf{N}$，則 $\dot{\mathbf{M}} = \mathbf{M}\mathbf{N} - \mathbf{N}\mathbf{M}$。（提示：$\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{S}^{-1} = -\mathbf{S}^{-1}\dot{\mathbf{S}}\mathbf{S}^{-1}$。）

3 **體積增長。**設 $\mathbf{M}(t) = \begin{pmatrix} e^{2t} & 0 \\ 0 & e^{-t} \end{pmatrix}$。求滿足 $\dot{\mathbf{M}} = \mathbf{M}\mathbf{N}$ 的 $\mathbf{N}$，並驗證 $\frac{\mathrm{d}}{\mathrm{d}t} \ln \det \mathbf{M} = \operatorname{tr} \mathbf{N}$。

4 **更高次冪。**直接證明：當 $\dot{\mathbf{M}} = \mathbf{M}\mathbf{N} - \mathbf{N}\mathbf{M}$ 時，對每個正整數 k 都有 $\frac{\mathrm{d}}{\mathrm{d}t} \operatorname{tr}(\mathbf{M}^k) = 0$。

延伸閱讀

- **P. G. Drazin & R. S. Johnson, *Solitons: An Introduction***（Cambridge University Press）。大學本科程度的標準入門書，內容包括拉克斯對。
- **Thierry Dauxois & Michel Peyrard, *Physics of Solitons***（Cambridge University Press）。書中講述了 Scott Russell 那個孤立波的故事。
- **Dwight Neuenschwander, *Emmy Noether's Wonderful Theorem***（Johns Hopkins University Press）。講解對稱性如何帶來守恆定律。

7.8 四個平方數

題目

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`

引子

7 不是兩個平方數之和，也不是三個平方數之和，卻是四個平方數之和：$7 = 4 + 1 + 1 + 1$。1770 年，拉格朗日（Joseph-Louis Lagrange）證明了每一個正整數都是四個平方數之和（容許使用 0^2）。這條定理早在 1621 年已由 Bachet 提出猜想，費馬（Fermat）亦聲稱自己證得，但一如以往沒有留下證明；歐拉（Euler）為此鑽研了四十年，始終差一步未能完成。

關鍵的工具由歐拉於 1748 年發現，正正就在這道題目之中：兩個四平方和的積，本身也是一個四平方和。這表示只需對質數（prime number）證明拉格朗日定理，其餘情況靠相乘便可得出。題目利用複矩陣（complex matrix）的行列式（determinant），寥寥數行便推導出這條恆等式（identity），而這些矩陣其實就是 Section 6.7 中的四元數（quaternion），只是換了個面貌。

動筆之前

想一想 回想兩平方的版本：$(a^2 + b^2)(p^2 + q^2)$ 是兩個平方數之和，因為對複數 $z = a + b\mathrm{i}$、$w = p + q\mathrm{i}$ 有 $|z|^2|w|^2 = |zw|^2$。矩陣可以怎樣把這個結果推廣到四個平方？

想法是找出一種由四個實數構成的「數」，它的「大小的平方」是 $a^2 + b^2 + c^2 + d^2$，而且具有可乘性（multiplicative）。矩陣的行列式正有可乘性：$\det(\mathbf{AB}) = \det \mathbf{A} \det \mathbf{B}$。因此我們要找一個由 a, b, c, d 構成的矩陣，其行列式為 $a^2 + b^2 + c^2 + d^2$，並且這類矩陣相乘後仍然是同一形式。

逐步解題

階段 1　行列式是四個平方數之和

想一想　$\begin{pmatrix} z_1 & z_2 \\ -z_2^* & z_1^* \end{pmatrix}$ 的行列式是甚麼？若 $z_1 = a + b\mathrm{i}$ 及 $z_2 = c + d\mathrm{i}$，以 a, b, c, d 表示又是甚麼？

$\det = z_1 z_1^* + z_2 z_2^* = |z_1|^2 + |z_2|^2 = a^2 + b^2 + c^2 + d^2$。同樣地，取 $z_3 = p + q\mathrm{i}$ 及 $z_4 = r + s\mathrm{i}$，第二個矩陣的行列式為 $p^2 + q^2 + r^2 + s^2$。

階段 2　乘積具有相同形式

相乘得

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

設 $w_1 = z_1 z_3 - z_2 z_4^*$ 及 $w_2 = z_1 z_4 + z_2 z_3^*$，則下面一行是 $(-w_2^*, w_1^*)$，所以乘積是 $\begin{pmatrix} w_1 & w_2 \\ -w_2^* & w_1^* \end{pmatrix}$，形式相同。它的行列式是 $|w_1|^2 + |w_2|^2$，而由於行列式可以相乘，

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

階段 3　讀出四個平方

停一停，試一試　把 $w_1 = (a + b\mathrm{i})(p + q\mathrm{i}) - (c + d\mathrm{i})(r - s\mathrm{i})$ 及 $w_2 = (a + b\mathrm{i})(r + s\mathrm{i}) + (c + d\mathrm{i})(p - q\mathrm{i})$ 展開，寫成實部（real part）和虛部（imaginary part）。

$$w_1 = (ap - bq - cr - ds) + \mathrm{i}(aq + bp + cs - dr),$$
$$w_2 = (ar - bs + cp + dq) + \mathrm{i}(as + br - cq + dp).$$

所以

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

每一個對 a, b, c, d 及對 p, q, r, s 都是線性（linear）的，並且 $(a^2 + b^2 + c^2 + d^2)(p^2 + q^2 + r^2 + s^2) = L_1^2 + L_2^2 + L_3^2 + L_4^2$。 ■

例如，$7 = 2^2 + 1^2 + 1^2 + 1^2$ 而 $3 = 1^2 + 1^2 + 1^2 + 0^2$。取 $(a, b, c, d) = (2, 1, 1, 1)$ 及 $(p, q, r, s) = (1, 1, 1, 0)$，由公式得 $(L_1, L_2, L_3, L_4) = (0, 2, 4, 1)$，而果然 $0^2 + 2^2 + 4^2 + 1^2 = 21 = 7 \times 3$。

回顧

- **可乘性是關鍵。**這條恆等式來自 $\det(\mathbf{AB}) = \det\mathbf{A}\det\mathbf{B}$，正如兩平方恆等式來自 $|zw| = |z||w|$ 一樣。
- **對乘法封閉（closure）。**$\begin{pmatrix} z_1 & z_2 \\ -z_2^* & z_1^* \end{pmatrix}$ 這個特殊形式在相乘之下保持不變，結果才會是另一個四平方和。
- **細心記賬十分重要。**最後一步雖屬例行運算，卻容易出錯；正負號和共軛（conjugate）都必須小心處理。

背後的數學

拉格朗日四平方和定理

拉格朗日的證明大致如下。根據歐拉恆等式，只需證明每個質數 p 都是四個平方數之和。先用一個計數論證，證明存在某個倍數 mp（其中 $m < p$）是四個平方數之和。然後用一個巧妙的「遞降法」(descent) 證明：若 mp 是四個平方數之和而 $1 < m < p$，則對某個更小的 m'，$m'p$ 也是四個平方數之和；如此重複下去，便會到達 $m = 1$。(可與 Section 1.5 中費馬的遞降法比較。) 1834 年，雅可比（Jacobi）更進一步，求出把 n 寫成四個平方數之和的方法的確切數目：n 的所有非 4 的倍數的因數之和的 8 倍。例如，1 有 8 種寫法（± 1 可放在四個位置中的任何一個）。

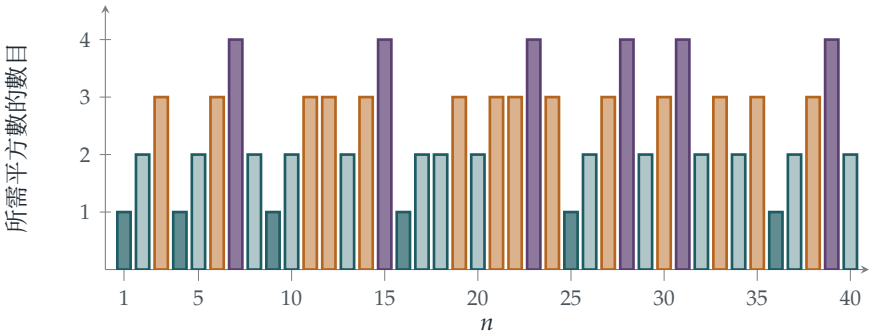


圖 7.8.1　由 1 至 40 的每個 n 最少需要多少個平方數才能組成。根據拉格朗日定理，每個數最多只需四個。需要四個的數（紫紅色）是 $7, 15, 23, 28, 31, 39, \ldots$：根據勒讓德（Legendre）和高斯（Gauss）的一條定理，它們正好是形如 $4^k(8m+7)$ 的數。

二、三、四

不同維數之間的規律相當微妙。兩平方和來自複數，而一個數是兩個平方數之和，當且僅當它的質因數中形如 $4k+3$ 的都以偶數次冪出現（Section 1.6）。四平方和來自四元數，而每個數都是四平方和。但三平方和卻完全沒有乘法恆等式（並不存在一個「大小」具可乘性的三維數系；這正是哈密頓（Hamilton）研究三元組時失敗的原因，見 Section 6.7），相應的定理也較難：勒讓德和高斯證明了，除非 $n = 4^k(8m+7)$，否則 n 是三個平方數之和（圖 7.8.1）。四以上，八維的八元數（octonion）給出一條八平方恆等式，由 Degen 和 Graves 發現；而胡爾維茲（Hurwitz）於 1898 年證明，這類恆等式只在 1、2、4 和 8 個平方的情況下存在。

華林問題

拉格朗日定理是華林（Edward Waring）於 1770 年提出的一個問題的第一個情況：對每個 k，是否存在一個數 $g(k)$，使得每個正整數都是最多 $g(k)$ 個 k 次冪之和？希爾伯特（David Hilbert）於 1909 年證明了確實存在。現在我們知道 $g(2) = 4$、$g(3) = 9$（例如 $23 = 2\cdot 2^3 + 7\cdot 1^3$ 需要九個立方數）以及 $g(4) = 19$。

思維技巧

尋找具可乘性的結構

要證明某個性質在乘法之下得以保持，可尋找一個代數結構（複數、四元數、矩陣），使該性質就是其中某個對象的「大小」，而大小具有可乘性。

一旦找到這樣的結構，乘積恆等式便唾手可得，而一個關於所有數的問題，也就化為一個關於質數的問題。

延伸練習

1. **一個實例。**利用 $L_1, \dots, L_4$ 的公式，取 $(a,b,c,d) = (p,q,r,s) = (2,1,1,1)$，把 $49 = 7\times 7$ 寫成四個平方數之和，並驗算你的答案。它與顯而易見的 $49 = 7^2+0^2+0^2+0^2$ 是否相同？

2. **四平方中的兩平方。**證明若 $c = d = r = s = 0$，這條恆等式便簡化為 Section 1.6 的兩平方恆等式。

3 **需要四個平方的數。**考慮平方數模 8 的餘數，證明形如 $8m+7$ 的數都不是三個平方數之和。

4 **雅可比公式。**數一數把 2 表示成四個平方數之和的方法有多少種（次序和正負號不同均視為不同），並驗證答案是 $8\times(1+2)=24$。

延伸閱讀

- **Harold Davenport, *The Higher Arithmetic***（Cambridge University Press）。第 V 章證明了拉格朗日四平方和定理。
- **John Conway & Derek Smith, *On Quaternions and Octonions***（A. K. Peters）。講述四元數的算術，包括利用四元數整數證明四平方和定理。
- **Aigner & Ziegler, *Proofs from THE BOOK***，「Representing numbers as sums of two squares」（把數表示成兩個平方數之和）一章。與本節相配的兩平方版本。

7.9 只有五種柏拉圖立體

題目

STEP III 1999，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/99-s3-q4`

引子

正多邊形（regular polygon）有無窮多種：三角形、正方形、五邊形，如此類推，任何邊數都有。你或會以為正多面體（regular polyhedron）也有無窮多種。其實恰好只有五種：正四面體（tetrahedron）、正方體（cube）、正八面體（octahedron）、正十二面體（dodecahedron）和正二十面體（icosahedron）。柏拉圖（Plato）把其中四種與四大元素（火、土、氣、水）對應，第五種則對應宇宙；歐幾里得（Euclid）的《幾何原本》（*Elements*）是有史以來影響最深遠的數學著作，全書最後就是證明再沒有其他正多面體；開普勒（Kepler）更用它們建構了一個太陽系模型。

本題只用寥寥數行就證明了這個分類，所依據的只是歐拉（Euler）在 1750 年發現的一條非凡公式：對任何凸多面體（convex polyhedron），$V - E + F = 2$。一條關於點算的公式，化成了一個關於形狀的定理，而 2 這個數原來正是球面的「指紋」。

動筆之前

想一想 用正方體驗證歐拉公式：它有多少個頂點（vertex）、多少條棱（edge）、多少個面（face）？以正方形為底的棱錐體（pyramid）又如何？

正方體有 $V = 8$，$E = 12$，$F = 6$，而 $8 - 12 + 6 = 2$。以正方形為底的棱錐體有 5 個頂點、8 條棱和 5 個面，而 $5 - 8 + 5 = 2$。任何沒有洞的多面體，不論多麼不規則，這條公式都成立。

想一想　在正多面體中，每個面有 m 條棱，而每條棱是兩個面的公共邊。由此可得 E 與 F 之間的甚麼方程？同樣地，E 與 V 之間又有甚麼方程？

用兩種方法點算棱與面的關聯：F 個面各有 m 條棱，而每條棱被數了兩次（它所屬的兩個面各數一次），所以 $mF = 2E$。同樣地，每個頂點連着 n 條棱，而每條棱有兩個端點，所以 $nV = 2E$。這種「用兩種方法點算同一樣東西」的技巧稱為**雙重計數**（double counting）。

逐步解題

階段 1　F 的公式

由 $mF = 2E$ 和 $nV = 2E$，可得 $E = \frac{mF}{2}$ 和 $V = \frac{mF}{n}$。代入歐拉公式，

$$\frac{mF}{n} - \frac{mF}{2} + F = 2 \quad\Longrightarrow\quad F\,\frac{2m - mn + 2n}{2n} = 2 \quad\Longrightarrow\quad F = \frac{4n}{2m + 2n - mn}.$$

而 $4 - (n-2)(m-2) = 4 - (mn - 2m - 2n + 4) = 2m + 2n - mn = h$。所以 $F = \frac{4n}{h}$。 ■

階段 2　m 和 n 可以取哪些值？

想一想　m（每個面的邊數）和 n（每個頂點處的面數）最小可以是多少？又既然 $F > 0$，h 必須是正還是負？

多邊形最少要有 3 條邊，所以 $m \geq 3$；立體的每個頂點處最少要有 3 個面相交，所以 $n \geq 3$。由於 $F > 0$，我們需要 $h > 0$，即 $(n-2)(m-2) < 4$。兩個因子都至少是 1，所以 $(m-2, n-2)$ 只可能是 $(1,1)$、$(1,2)$、$(2,1)$、$(1,3)$ 和 $(3,1)$。

停一停，試一試　就這五種可能逐一計算 h，然後求 F，再求 $E = \frac{mF}{2}$ 和 $V = \frac{mF}{n}$。指出每種情況是哪一個立體。

m	n	h	F	E	V	立體
3	3	3	4	6	4	正四面體
3	4	2	8	12	6	正八面體
4	3	2	6	12	8	正方體
3	5	1	20	30	12	正二十面體
5	3	1	12	30	20	正十二面體

可能的情況只有這些，所以正多面體最多只有五種；而這五種全部存在（歐幾里得說明

了如何逐一構作）。 ■

回顧

- **雙重計數。**用兩種方法點算棱與面、棱與頂點的關聯，得出 $mF = 2E$ 和 $nV = 2E$。
- **由正值得出不等式。**$F > 0$ 這個簡單的要求，把無窮多個候選者篩剩五個。
- **存在性要另行處理。**這個論證只說明最多有五種；每一種是否真的存在，還須實際構作出來才能確定。

背後的數學

歐拉公式與拓撲學

歐拉在 1750 年提出 $V - E + F = 2$，但他的證明有漏洞；這條公式其後經歷了一段引人入勝的歷史，當中有一次又一次的嘗試證明、反例（counterexample）和修補。一個「畫框」形的多面體，中間穿了一個正方形的洞，有 $V - E + F = 0$；由兩個正方體在一個頂點處相連而成的多面體，則有 $V - E + F = 3$。拉卡托斯（Imre Lakatos）把這段歷史寫成一本名著 *Proofs and Refutations*（《證明與反駁》），探討數學概念如何在反例的磨礪下變得精確。最終的答案是：$V - E + F$ 只取決於曲面在拓撲學（topology）那種寬鬆意義下的形狀：凡是可以變形成球面的曲面，它等於 2；環面（torus，即冬甩形）是 0；有 g 個洞的曲面則是 $2 - 2g$。這個數稱為**歐拉示性數**（Euler characteristic），是拓撲學最早的概念之一；拓撲學研究的，就是經拉伸和彎曲後仍然不變的性質。

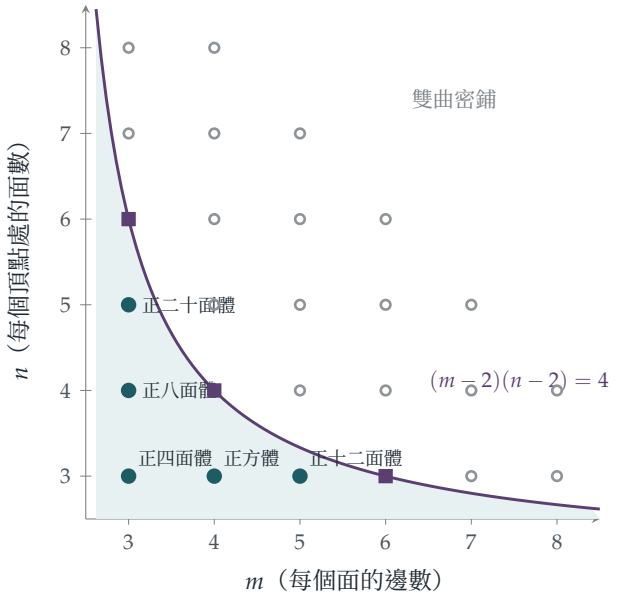


圖 7.9.1　滿足 $m, n \geq 3$ 的所有可能數對 (m, n)。在曲線 $(m-2)(n-2)=4$ 下方（陰影部分）的是五種柏拉圖立體（藍綠色）。在曲線上的是平面的三種密鋪（紫紅色方點）：分別用三角形（$m=3$，$n=6$）、正方形（4, 4）和六邊形（6, 3）。曲線上方的每一點（空心圓圈）都對應雙曲平面的一種正則密鋪。

球面以外

當 $h = 0$ 或 $h < 0$ 時又會怎樣？當 $(m-2)(n-2) = 4$ 時，有三組解：$(3,6)$、$(4,4)$ 和 $(6,3)$，即六個三角形、四個正方形或三個六邊形在一點相接，各角之和恰好是 360°。這就是平面的三種正則密鋪（regular tiling），也就是浴室地磚和蜂巢的圖案。而當 $(m-2)(n-2) > 4$ 時，一個頂點處各角之和大於 360°，這在平面上不可能出現，但在 Section 5.8 的雙曲平面（hyperbolic plane）上卻完全可行。雙曲平面有無窮多種正則密鋪，圖 7.9.1 中曲線上方的每一點各對應一種；艾雪（Escher）的木刻版畫 *Circle Limit*（《圓極限》）系列，畫的就是其中幾種。所以，柏拉圖立體的分類其實是一個三分法（trichotomy）的第一種情況，這個三分法（球面、平面、雙曲平面）貫穿整個幾何學。

對稱羣

柏拉圖立體的對稱構成羣（group）：正四面體有 12 個旋轉，正方體和正八面體有 24 個（兩者互為對偶（dual）：把面與頂點互換，$m \leftrightarrow n$），正十二面體和正二十面體則有 60 個。這個階（order）為 60 的正二十面體羣是一個「單」羣（simple group）；正如我們在 Section 6.4 提過，正是它令五次方程（quintic）不能用根式（radicals）求解。克萊因（Felix Klein）寫了一整本書 *Lectures on the Icosahedron*（《正二十面體講義》，1884），把

正二十面體與五次方程的解法聯繫起來。

思維技巧

雙重計數

用兩種不同的方法點算同一組東西，然後令兩個結果相等。

這裏所點算的，是所有（棱，面）配對所成的集合，其中棱在該面上。雙重計數是組合數學（combinatorics）中最強大的工具之一，在第 9 章會一再出現。

延伸練習

1 **對偶性。**證明把 m 和 n 互換，會令 V 和 F 互換，而 E 保持不變。哪幾對柏拉圖立體互為對偶？哪一個是自對偶的？

2 **足球。**傳統的足球由 12 個五邊形和 20 個六邊形組成，每個頂點處有三個面相接。利用雙重計數和歐拉公式，驗證 $V - E + F = 2$。再證明更一般的結果：任何這樣由五邊形和六邊形組成、每個頂點處有三個面相接的多面體，都恰好有 12 個五邊形。

3 **環面。**造一個「畫框」形的多面體（一個正方形方塊，中間有一個正方形的洞），數出 V、E 和 F。驗證 $V - E + F = 0$。

4 **頂點處的角。**正 m 邊形的內角（interior angle）是 $180°(1 - \frac{2}{m})$。證明 n 個這樣的角能圍着凸立體的一個頂點拼合（總和小於 360°），當且僅當 $(m-2)(n-2) < 4$，從而得出這個分類的第二個證明。

延伸閱讀

- **David Richeson, *Euler's Gem***（Princeton University Press）。講述 $V - E + F = 2$ 的歷史，以及它如何引出拓撲學。非常適合高中第六學級（sixth form）的學生閱讀。
- **Imre Lakatos, *Proofs and Refutations***（Cambridge University Press）。一場關於歐拉公式和證明本質的對話。
- **H. S. M. Coxeter, *Regular Polytopes***（Dover）。論述正多面體及其高維同類的經典著作。

我們學到了甚麼

本章開首我們說過，矩陣（matrix）是空間的一種變換（transformation），自有其生命；我們還提出了三個問題：矩陣與數有何相似、有何不同，矩陣究竟做了甚麼，以及甚麼保持不變。以下是本章九道題目所揭示的答案。

1. 矩陣與數有何相似，又有何不同？

矩陣可以相加和相乘，但 Section 7.1 指出了哪些熟悉的法則不再成立。乘法不符合交換律（commutative law），所以 $(\mathbf{A}-\mathbf{B})(\mathbf{A}+\mathbf{B})$ 未必等於 $\mathbf{A}^2-\mathbf{B}^2$；矩陣中存在零因子（zero divisor），所以 $\mathbf{AB}=\mathbf{0}$ 並不保證 $\mathbf{BA}=\mathbf{0}$；而一條二次矩陣方程可以有無限多個根。每一項法則的失效都有幾何解釋，可以從這些矩陣對平面的作用來理解。至關重要的是，有一項性質保留了下來：行列式（determinant）可乘，即 $\det(\mathbf{AB})=\det\mathbf{A}\det\mathbf{B}$。Section 7.3 說明這項性質基本上刻劃了行列式，而 Section 7.8 則用它證明了歐拉（Euler）的四平方恆等式（four-square identity），這正是拉格朗日（Lagrange）定理的關鍵一步；該定理指出每個正整數都是四個平方數之和。

2. 矩陣究竟做了甚麼？

有幾節透過找出合適的觀點來回答這個問題。

- **特徵向量（eigenvector）**（Section 7.2）是矩陣不會轉動的方向。在以特徵向量為軸的坐標下，對稱矩陣（symmetric matrix）只不過是一個伸縮，而圓的像是一個橢圓，其軸和面積都可以直接讀出。
- **行列式**（Section 7.3）是帶正負號的面積比例因子：錯切（shear）保持面積不變，對換會把定向（orientation）反轉，奇異矩陣（singular matrix）則把平面壓扁。
- **跡（trace）與行列式**（Section 7.4）透過凱萊–哈密頓定理（Cayley–Hamilton theorem），決定一個矩陣會否在若干步之後回到單位矩陣（identity matrix）。滿足 $\mathbf{M}^8=\mathbf{I}$ 的矩陣未必是旋轉；它是在歪斜坐標下看到的旋轉。又由於整數矩陣的跡是整數，晶體不可能有五重對稱。
- **三維空間中的旋轉**（Section 7.5）可由向量積（cross product）透過羅德里格斯公式（Rodrigues′ formula）構造出來，這是 $e^{i\theta}=\cos\theta+i\sin\theta$ 在三維的對應版本。

- **雙曲旋轉（hyperbolic rotation）**（Section 7.6）使點沿雙曲線移動，並把它們的雙曲角相加。它們就是狹義相對論中的洛倫茲變換（Lorentz boost），而它們的不變直線（invariant line）就是光的路徑。

3. 甚麼保持不變？

不變量（invariant）無處不在。旋轉保持 $x^2 + y^2$ 不變，洛倫茲變換保持 $x^2 - y^2$ 不變，而相似矩陣（similar matrices）有相同的跡和行列式。在 Section 7.7 中，一個按 $\dot{\mathbf{M}} = \mathbf{MN} - \mathbf{NM}$ 演化的矩陣，不論其元素如何變化，它的跡和行列式都保持不變，因而它的特徵值（eigenvalue）也保持不變：這就是孤立子（soliton）各條守恆定律背後的機制。而在 Section 7.9 中，不變量是一個數，$V - E + F = 2$，它對每個凸多面體（convex polyhedron）都成立，並把正多面體限定為僅有五種。這個數稱為歐拉示性數（Euler characteristic），是拓撲學（topology）中最早出現的不變量之一。

4. 線性代數工具箱

方法	你在哪裏用過
幾何反例	失效的法則（7.1）
以特徵向量為自然坐標軸	橢圓（7.2）；洛倫茲變換（7.6）
用特殊矩陣作試探	行列式（7.3）
凱萊–哈密頓定理	有限階（7.4）
分解為平行和垂直部分	羅德里格斯公式（7.5）
選擇合適的參數	快度（rapidity）（7.6）
$\mathrm{tr}(\mathbf{MN}) = \mathrm{tr}(\mathbf{NM})$	守恆量（7.7）
行列式的可乘性	四平方和（7.8）
重複計數（double counting）	柏拉圖立體（Platonic solids）（7.9）

下一步

- 在 Section 11.2 中，特徵值和特徵向量支配着彈簧與質點系統的振動。
- 在 Section 13.7 中，協方差矩陣（covariance matrix）的特徵向量會在統計學中出現。
- 微分方程（differential equation）和動力系統（dynamical system）是第 8 章的主題。

如果你想在入讀大學之前再走遠一點

- **3Blue1Brown, *Essence of Linear Algebra***（線性代數的本質）：一套免費的影片系列，展示矩陣究竟做了甚麼。請先看這個。

- **Gilbert Strang, *Introduction to Linear Algebra***（線性代數導論）：經典的入門課程，十分着重理解和應用。
- **Sheldon Axler, *Linear Algebra Done Right***（線性代數應該這樣學）：較為抽象的處理手法，以線性映射（linear map）為中心；可在網上免費閱讀。
- **David Richeson, *Euler's Gem***（歐拉的瑰寶）：歐拉公式與拓撲學的誕生。

Kaplansky 曾打趣說，他和 Halmos 思考時不用坐標，但關起門來卻「拼命地」用矩陣計算。這句話正好道出本章想說明的事：矩陣既具體又可以計算，但每一項計算背後都有幾何（一個伸縮、一個旋轉、一個投影、一個洛倫茲變換），還有一個不變量，它不取決於你怎樣把它寫下來。學會在這兩種觀點之間來回轉換，是數學家所能擁有的最寶貴技能之一。

延伸練習答案

以下只是簡短的答案和提示，並非完整解答：這裏給出結果和關鍵步驟，讓你核對自己的做法。請先認真做過每道練習，然後才看答案。

7.1 對矩陣不成立的法則

1. 設 $\mathbf{X} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$，比較 $\mathbf{XP}$ 與 $\mathbf{PX}$，必有 $c = 0$ 及 $a = d$：所求的矩陣（matrix）為 $\begin{pmatrix} a & b \\ 0 & a \end{pmatrix}$。

2. 由凱萊－哈密頓定理（Cayley–Hamilton theorem）得 $\mathbf{N}^2 = t\mathbf{N} - d\mathbf{I}$，所以 $t\mathbf{N} = d\mathbf{I}$。若 $t \neq 0$，則 $\mathbf{N}$ 是 $\mathbf{I}$ 的倍數，而 $\mathbf{N}^2 = \mathbf{0}$ 迫使 $\mathbf{N} = \mathbf{0}$，與 $t \neq 0$ 矛盾。所以 $t = 0$，從而 $d = 0$。答案是：$\begin{pmatrix} a & b \\ c & -a \end{pmatrix}$，其中 $a^2 + bc = 0$。

3. 平方後的對角線元素為 $\cos^2\theta + \sin^2\theta = 1$，非對角線元素為 0。保持不變的向量（vector）是 $(\cos\frac{\theta}{2}, \sin\frac{\theta}{2})$，所以它是對通過原點、傾角為 $\frac{\theta}{2}$ 的直線的反射（reflection）。

4. 設 $\mathbf{M} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$，則 $\mathbf{M}^2 = \begin{pmatrix} a^2 + bc & b(a+d) \\ c(a+d) & d^2 + bc \end{pmatrix}$；減去 $(a+d)\mathbf{M}$ 再加上 $(ad - bc)\mathbf{I}$，便得 $\mathbf{0}$。

7.2 只伸縮，不轉向

1. $\mathbf{Mu} = \lambda\mathbf{u}$ 且 $\mathbf{u} \neq \mathbf{0}$，必須有 $\det(\mathbf{M} - \lambda\mathbf{I}) = 0$，這就是該二次方程。這裏跡（trace）為 6，行列式（determinant）為 $\frac{4}{25}(54 - 4) = 8$，所以 $\lambda^2 - 6\lambda + 8 = 0$：$\lambda = 2$（對應 $(1, 2)$）和 4（對應 $(2, -1)$）。

2. $\mathbf{v} \cdot \mathbf{Mu} = \lambda\,\mathbf{v} \cdot \mathbf{u}$，又因為 $\mathbf{M}$ 是對稱矩陣（symmetric matrix），$\mathbf{v} \cdot \mathbf{Mu} = \mathbf{Mv} \cdot \mathbf{u} = \mu\,\mathbf{v} \cdot \mathbf{u}$。所以 $(\lambda - \mu)\,\mathbf{u} \cdot \mathbf{v} = 0$。

3. 特徵方程（characteristic equation）$\lambda^2 - 2\lambda\cos\theta + 1 = 0$ 的判別式（discriminant）為 $-4\sin^2\theta$，除非 θ 是 π 的倍數，否則必為負數。特徵值（eigenvalue）是 $\mathrm{e}^{\pm i\theta}$。

4. $(1, 0) = \frac{1}{5}(1, 2) + \frac{2}{5}(2, -1)$，所以 $T^n(1, 0) = \frac{2^n}{5}(1, 2) + \frac{2 \cdot 4^n}{5}(2, -1) = \frac{1}{5}(2^n +$

$4^{n+1},\ 2^{n+1} - 2\cdot 4^n)$。當 $n = 1$ 時，結果是 $(\frac{18}{5}, -\frac{4}{5})$，正如所料。

7.3 甚麼才算行列式

1. $(\det \mathbf{MN})^3 = (\det \mathbf{M})^3 (\det \mathbf{N})^3$，而 $(\det \mathbf{J})^3 = -1 \neq 1$。函數 $|\det \mathbf{M}|$ 具有可乘性（multiplicative），但 $|\det \mathbf{J}| = 1$。

2. $\begin{pmatrix} a & \lambda a \\ c & \lambda c \end{pmatrix} = \begin{pmatrix} a & 0 \\ c & 0 \end{pmatrix} \begin{pmatrix} 1 & \lambda \\ 0 & 0 \end{pmatrix}$，而第二個因子的第二行是第一行的 0 倍，所以它的 f 值為 0。

3. $\begin{pmatrix} k & 0 \\ 0 & 1 \end{pmatrix} = \mathbf{J}\mathbf{K}_k\mathbf{J}$，所以它的值是 $\mathrm{f}(\mathbf{J})^2 \mathrm{f}(\mathbf{K}_k) = \mathrm{f}(\mathbf{K}_k)$。另外 $\mathbf{K}_k\mathbf{K}_l = \mathbf{K}_{kl}$。

4. 行簡化（row reduction）可把任何可逆矩陣（invertible matrix）寫成初等矩陣（elementary matrix）的乘積：行對換（$\mathbf{J}$）、行倍乘（$\mathbf{K}_k$ 或 $\mathbf{J}\mathbf{K}_k\mathbf{J}$），以及錯切（shear）（$\begin{pmatrix} 1 & t \\ 0 & 1 \end{pmatrix}$ 或它經 $\mathbf{J}$ 共軛（conjugate）後所得的矩陣）。錯切的 $\mathrm{f} = 1$（做法如第 (iv) 部分，利用 $\mathbf{K}^{-1}\mathbf{P}\mathbf{K}$）。此外，$\mathbf{J}$ 是若干錯切與 $\mathbf{K}_{-1}$ 的乘積，所以 $\mathrm{f}(\mathbf{K}_{-1}) = \mathrm{f}(\mathbf{J}) = -1$。因此 $\mathrm{f}(\mathbf{M}) = \mathrm{g}(\det \mathbf{M})$，其中 $\mathrm{g}(k) = \mathrm{f}(\mathbf{K}_k)$；若 $\mathbf{M}$ 是奇異矩陣（singular matrix），則由第 (iii) 部分得 $\mathrm{f}(\mathbf{M}) = 0$。

7.4 回到起點的矩陣

1. 由 $|1 + \mathrm{e}^{\mathrm{i}\phi}|^2 = 2 + 2\cos\phi = 1$ 得 $\phi = \pm\frac{2\pi}{3}$，所以 $\mathbf{R}^3 = \mathbf{I}$。至於 $\mathbf{S}$：它的特徵值是 1 的立方根。假如兩個特徵值都是 1，則由於 $x^3 - 1$ 的根各不相同，$\mathbf{S}$ 可對角化（diagonalisable），因而等於 $\mathbf{I}$。實矩陣的非實數特徵值必以共軛對的形式出現，所以特徵值是 $\mathrm{e}^{\pm 2\pi\mathrm{i}/3}$：$\det \mathbf{S} = 1$ 及 $\operatorname{tr} \mathbf{S} = -1$。

2. 特徵值必定是六次本原單位根（primitive sixth roots of unity）$\mathrm{e}^{\pm\mathrm{i}\pi/3}$（若選取其他六次單位根並與其共軛配對，都會得出 $\mathbf{M}^2 = \mathbf{I}$ 或 $\mathbf{M}^3 = \mathbf{I}$）。所以 $\operatorname{tr} \mathbf{M} = 1$ 及 $\det \mathbf{M} = 1$。

3. 特徵值是乘積為 1 的單位根，所以它們是 $\mathrm{e}^{\pm\mathrm{i}\theta}$（或同為 ± 1），而跡 $2\cos\theta$ 是 $[-2, 2]$ 內的整數。階（order）為 6 的例子：$\begin{pmatrix} 1 & -1 \\ 1 & 0 \end{pmatrix}$。

4. $\mathbf{M}^3 = (t^2 - d)\mathbf{M} - td\mathbf{I}$ 及 $\mathbf{M}^4 = (t^3 - 2td)\mathbf{M} - d(t^2 - d)\mathbf{I}$。要使 $\mathbf{M}^4 = -\mathbf{I}$（$\mathbf{M}$ 不可能是 $\mathbf{I}$ 的倍數），須有 $t(t^2 - 2d) = 0$ 及 $d(t^2 - d) = 1$。$t = 0$ 給出 $-d^2 = 1$，不可能成立；所以 $t^2 = 2d$，$d^2 = 1$，於是 $d = 1$，$t^2 = 2$。

7.5 三維空間中的旋轉

1. $z \mapsto a + \mathrm{e}^{\mathrm{i}\theta}(z-a)$。複合變換是 $z \mapsto \mathrm{e}^{\mathrm{i}(\theta+\phi)}z + k$，其中 $k = \mathrm{e}^{\mathrm{i}\phi}a(1-\mathrm{e}^{\mathrm{i}\theta}) + b(1-\mathrm{e}^{\mathrm{i}\phi})$；若 $\mathrm{e}^{\mathrm{i}(\theta+\phi)} \neq 1$，它有不動點（fixed point）$c = \frac{k}{1-\mathrm{e}^{\mathrm{i}(\theta+\phi)}}$。若 $\theta+\phi = 2\pi$，它就是位移為 $(b-a)(1-\mathrm{e}^{-\mathrm{i}\theta})$ 的平移（translation）。
2. $\begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$：它保持 $(1,1,1)$ 不變，跡為 $0 = 1 + 2\cos 120°$，並把 $\mathbf{i} \to \mathbf{j} \to \mathbf{k} \to \mathbf{i}$。
3. $\mathbf{K}^2\mathbf{r} = \mathbf{n} \times (\mathbf{n} \times \mathbf{r}) = (\mathbf{n}.\mathbf{r})\mathbf{n} - \mathbf{r}$，於是 $\mathbf{K}^3\mathbf{r} = \mathbf{n} \times \big((\mathbf{n}.\mathbf{r})\mathbf{n} - \mathbf{r}\big) = -\mathbf{K}\mathbf{r}$。在級數（series）中，奇數次冪給出 $(\theta - \frac{\theta^3}{3!} + \cdots)\mathbf{K} = \sin\theta\,\mathbf{K}$，偶數次冪（由 $\mathbf{K}^2$ 起）給出 $(\frac{\theta^2}{2!} - \frac{\theta^4}{4!} + \cdots)\mathbf{K}^2 = (1-\cos\theta)\mathbf{K}^2$。
4. 複合變換保持交線上的每一點不變，而且行列式為 $(-1)^2 = 1$。在每個垂直於該交線的平面上，它是對兩條夾角為 α 的直線先後作反射的乘積，即轉角為 2α 的旋轉（rotation）。

7.6 雙曲旋轉與相對論

1. $T(u)T(w)$ 的元素是 $\cosh u\cosh w + \sinh u\sinh w = \cosh(u+w)$ 和 $\sinh u\cosh w + \cosh u\sinh w = \sinh(u+w)$。所以 $T(u)^{-1} = T(-u)$。
2. 由 $\tanh u = 0.8$ 得 $u = \frac{1}{2}\ln\frac{1.8}{0.2} = \ln 3 \approx 1.099$ 及 $\cosh u = \frac{1}{2}(3 + \frac{1}{3}) = \frac{5}{3}$。在飛船上，過去的時間是 $10 \times \frac{3}{5} = 6$ 年。
3. 面積是 $\frac{1}{2}\oint(x\,\mathrm{d}y - y\,\mathrm{d}x)$。兩條直邊沒有貢獻；沿着雙曲線（hyperbola），取 $x = \cosh s$，$y = \sinh s$，被積函數（integrand）是 $(\cosh^2 s - \sinh^2 s)\,\mathrm{d}s = \mathrm{d}s$，得出 $\frac{v}{2}$。T 的行列式為 $\cosh^2 u - \sinh^2 u = 1$，所以它保持面積不變；它把雙曲線映射到自身，並把參數為 v 的點移到參數為 $v+u$ 的點。
4. $x' + y' = (\cosh u + \sinh u)(x+y) = \mathrm{e}^u p$ 及 $x' - y' = \mathrm{e}^{-u}q$；另外 $x^2 - y^2 = pq$。在這組坐標下，T 是一個擠壓映射（squeeze）：沿漸近線（asymptote）$y = x$ 拉伸 e^u 倍，沿 $y = -x$ 壓縮至 e^{-u} 倍，使各點沿着每條雙曲線 $pq =$ 常數滑動。

7.7 永恆不變的量

1. $\dot{a} = -Ct = -tc$，$\dot{d} = Ct = tc$，$\dot{c} = 0$，以及 $\dot{b} = (A-D)t - Ct^3 = t(a-d)$，與 $\mathbf{M}\mathbf{N} - \mathbf{N}\mathbf{M} = t\begin{pmatrix} -c & a-d \\ 0 & c \end{pmatrix}$ 吻合。
2. $\dot{\mathbf{M}} = -\mathbf{S}^{-1}\dot{\mathbf{S}}\mathbf{S}^{-1}\mathbf{M}_0\mathbf{S} + \mathbf{S}^{-1}\mathbf{M}_0\dot{\mathbf{S}} = -\mathbf{N}\mathbf{M} + \mathbf{M}\mathbf{N}$，當中用到 $\dot{\mathbf{S}} = \mathbf{S}\mathbf{N}$。

3. $\mathbf{N} = \mathbf{M}^{-1}\dot{\mathbf{M}} = \begin{pmatrix} 2 & 0 \\ 0 & -1 \end{pmatrix}$，而 $\ln\det\mathbf{M} = 2t - t = t$，其導數（derivative）為 $1 = \operatorname{tr}\mathbf{N}$。

4. 利用跡的循環性質（cyclic property），$\frac{\mathrm{d}}{\mathrm{d}t}\operatorname{tr}(\mathbf{M}^k) = \sum_j \operatorname{tr}(\mathbf{M}^j\dot{\mathbf{M}}\mathbf{M}^{k-1-j}) = k\operatorname{tr}(\mathbf{M}^{k-1}\dot{\mathbf{M}})$；再用一次循環性質，$\operatorname{tr}(\mathbf{M}^{k-1}(\mathbf{MN} - \mathbf{NM})) = \operatorname{tr}(\mathbf{M}^k\mathbf{N}) - \operatorname{tr}(\mathbf{M}^{k-1}\mathbf{NM}) = 0$。

7.8 四個平方數

1. $(L_1, L_2, L_3, L_4) = (1, 4, 4, 4)$，而 $1 + 16 + 16 + 16 = 49$。這個表示法與 $7^2 + 0^2 + 0^2 + 0^2$ 不同。

2. 取 $c = d = r = s = 0$：$L_1 = ap - bq$，$L_2 = aq + bp$，$L_3 = L_4 = 0$，即 $(a^2 + b^2)(p^2 + q^2) = (ap - bq)^2 + (aq + bp)^2$。

3. 平方數模（modulo）8 只可能是 0、1 或 4。三個這樣的數之和模 8 可得 $0, 1, 2, 3, 4, 5, 6$，但永遠不會是 7。

4. $2 = (\pm1)^2 + (\pm1)^2 + 0^2 + 0^2$：選出兩個非零的位置（6 種方法）及它們的正負號（4 種方法），共 24 種。由於 2 不能被 4 整除，雅可比（Jacobi）公式給出 $8\sigma(2) = 8 \times 3 = 24$。

7.9 只有五種柏拉圖立體

1. 設 $h = 2m + 2n - mn$，則 $V = \frac{4m}{h}$，$E = \frac{2mn}{h}$ 及 $F = \frac{4n}{h}$；把 m 和 n 對調，V 和 F 也隨之對調。正方體與正八面體（octahedron）互為對偶（dual），正十二面體（dodecahedron）與正二十面體（icosahedron）亦然；正四面體（tetrahedron）則是自對偶（self-dual）的。

2. $F = 32$，$E = \frac{1}{2}(12 \cdot 5 + 20 \cdot 6) = 90$ 及 $V = \frac{1}{3}(180) = 60$，而 $60 - 90 + 32 = 2$。一般而言，若有 P 個五邊形和 H 個六邊形，則 $E = \frac{1}{2}(5P + 6H)$ 及 $V = \frac{1}{3}(5P + 6H)$，所以 $V - E + F = -\frac{1}{6}(5P + 6H) + P + H = \frac{P}{6}$。這個值等於 2，所以 $P = 12$。

3. 把頂面和底面（兩者都有洞）各分成四個梯形（trapezium）。於是 $V = 16$，$F = 4 + 4 + 4 + 4 = 16$ 及 $E = 12 + 12 + 8 = 32$，所以 $V - E + F = 0$。

4. $180^\circ n(1 - \frac{2}{m}) < 360^\circ$ 即 $n(m - 2) < 2m$，也就是 $mn - 2m - 2n < 0$，或 $(m - 2)(n - 2) < 4$。當 $m, n \geq 3$ 時，解為 $(3, 3)$、$(3, 4)$、$(4, 3)$、$(3, 5)$ 和 $(5, 3)$。

章

8

微分方程與動力系統

絕大多數微分方程都解不出來。學問就在於：解不出來，也要讀懂它。

本章內容

8.1	當解不唯一	STEP III 1994
8.2	不求解而估界	STEP III 2013
8.3	收穫種羣	STEP II 1995
8.4	盤旋趨向圓	STEP III 2015
8.5	格力犬追逐格力犬	STEP III 1988
8.6	量子振子	STEP III 1988
8.7	可以解出的混沌	STEP I 2014
8.8	週期三蘊含混沌	STEP III 2025
A	練習答案	第 563 頁

變化的法則

大自然對積分的種種困難一笑置之。

— 相傳為拉普拉斯（Pierre-Simon Laplace）所言

在中學，微分方程（differential equation）是拿來解的。你把變量分離，或者找一個積分因子（integrating factor），又或者猜一個形如 $e^{\lambda x}$ 的解，最後得到一條公式。課本裏的方程都是特意挑選的，所以這些方法才行得通。可惜的是，來自現實世界的微分方程（三顆行星的運動、天氣、捕食者與獵物的種羣數量、大幅度擺動的單擺），幾乎全都沒有可以用我們熟悉的函數寫出來的解。

本章要談的，就是這時候該怎麼辦。我們不再問「公式是甚麼？」，而是問一些不靠公式也能回答的問題。解到底存在嗎？是否只有一個？它會保持有界（bounded）嗎？它會趨於一個穩定狀態，還是趨於一個周而復始的循環，抑或永遠游移不定？長遠來說它會怎樣？這種由「求解」轉為「理解」的觀點，始於 1880 年代的龐加萊（Henri Poincaré），並由此開創了今天稱為*動力系統*（dynamical systems）的學科。

微分方程為何重要

微分方程是科學的語言。牛頓（Newton）第二定律是一條微分方程；麥克斯韋（Maxwell）的電磁學方程組、量子力學的薛定諤（Schrödinger）方程，以及用來預測天氣和設計飛機的流體流動方程，也都是微分方程。流行病學家用它們為疾病的傳播建立模型，生態學家用來描述魚類資源的數量，經濟學家用來描述經濟增長，工程師則用來描述各種控制系統的響應，由恆溫器到自動駕駛儀無一例外。凡是一個量的變化率取決於它當前的狀態，就有一條微分方程，而且通常是一條沒有人能精確解出的方程。

與這些連續模型並列的還有離散模型，其中時間是一步一步推進的：今年的種羣數量決定明年的種羣數量。這些*迭代映射*（iterated maps）寫起來比較簡單，但最後兩節會說明，它們的行為可以同樣狂野。1970 年代的重大發現之一，就是非常簡單的法則也能產生*混沌*（chaos）：這種行為完全是確定的，實際上卻無法預測。

你在大學會遇到甚麼

大學一年級的微分方程課程會講授標準的求解方法，但也會引入本章的這些問題。你會遇到存在性與唯一性定理（existence and uniqueness theorems），它們告訴你解在甚麼時候存在，又在甚麼時候由初始值決定。你會學習繪畫*相圖*（phase portraits），即把所

有解一次過畫出來的圖，並學習透過線性化（linearising）把平衡點（equilibria）分類為穩定或不穩定。往後的課程會走得遠得多：*非線性動力學*（nonlinear dynamics）研究極限環（limit cycles）、分歧（bifurcations）和混沌；*數學生物學*（mathematical biology）把這些概念應用於種羣和流行病；而偏微分方程（partial differential equations）、傅立葉級數（Fourier series）和正交函數（orthogonal functions）的理論（在 Section 8.6 可以略窺一二）則是量子力學和大部分應用數學的基礎。

貫穿本章的問題

三大問題

1. **現在能決定將來嗎？**一條微分方程配上一個初始值，在甚麼時候恰好只有一個解？
2. **不求解，我們能知道甚麼？**在沒有公式的情況下，怎樣為解定出上下界、找出它最終停留在哪裏，或者描述它的形狀？
3. **長遠來說會怎樣？**解會趨於穩定狀態、趨近一個循環，還是呈現混沌的行為？我們又怎樣分辨？

本章首先討論唯一性，以及唯一性不成立時會發生甚麼事（Section 8.1），然後發展一些不必求解也能理解方程的工具：不會增加的類能量的量、平衡點及其穩定性，以及配合問題對稱性的坐標。有兩節把這些想法應用於物理問題：在正方形內互相追逐的幾隻狗，以及量子諧振子（quantum harmonic oscillator）。本章最後討論離散映射：在那裏，一條公式揭示了混沌，而單單一個長度為三的循環，就迫使任何長度的循環都必然存在。

本章的概念

§	課題	概念
8.1	唯一性	包絡線（envelopes）與奇解（singular solutions）；黏合
8.2	能量方法	一個永不增加的量
8.3	種羣模型	平衡點、穩定性與臨界點（tipping point）
8.4	極限環	極坐標與伯努利代換（Bernoulli substitution）
8.5	追逐曲線（pursuit curves）	對稱性把四個物體化約為一個
8.6	正交函數	朗斯基行列式（Wronskian）與正交性
8.7	混沌	變量代換：邏輯斯諦映射（logistic map）其實是喬裝了的倍角映射
8.8	週期軌道（periodic orbits）	區間覆蓋；沙可夫斯基定理（Sharkovskii's theorem）

如何閱讀本章

在動手解任何東西之前，先畫草圖。對於微分方程 $y' = F(y)$，標出 F 在哪裏為正、為負和為零，並在 y 軸上畫上箭頭；對於映射 $x \mapsto f(x)$，把它的圖像連同直線 $y = x$ 一起畫出來，再用蛛網圖法（cobweb method）追蹤幾個數列。這些圖只需一分鐘就能畫好，而且往往在代數運算開始之前，就已經告訴你答案應該是甚麼模樣。

8.1 當解不唯一

題目

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`

引子

微分方程（differential equation）是一條告訴你某個量如何變化的規則：牛頓（Newton）定律告訴行星，在知道自己目前位置的情況下該如何運動。古典物理學的圖像是，這樣就已足夠：只要知道世界此刻的狀態，定律便決定了它的整個未來。拉普拉斯（Pierre-Simon Laplace）曾設想有這樣一個智者，只要知道每一個位置和速度，便能預測一切。這幅圖像建基於一個數學假設：一條微分方程加上一個初始條件（starting condition），恰好只有一個解。

這道題說明這個假設可以不成立。看似平平無奇的方程 $(y')^2 = 4y$，有些解會沿着 x 軸走，然後在自己選定的一刻離軸而去。經過某些點的解有無限多個。知道現在，並不能決定未來。找出唯一性（uniqueness）究竟在甚麼時候成立，是大學微分方程課程中最先遇到的定理之一。

動筆之前

想一想 把兩邊開平方：$y' = \pm 2\sqrt{y}$。試試 $y = (x-c)^2$。它滿足這條方程嗎？有沒有更簡單的解？

若 $y = (x-c)^2$，則 $y' = 2(x-c)$，而 $(y')^2 = 4(x-c)^2 = 4y$。所以對每一個 x，每一條拋物線（parabola）$y = (x-c)^2$ 都是解。而 $y = 0$ 也是解：$(0)^2 = 4 \times 0$。請記住最後這一個解；它是整道題的關鍵。

逐步解題

階段 1　經過一點的解

想一想　哪些拋物線 $y=(x-c)^2$ 經過 (a,b^2)？

我們需要 $(a-c)^2=b^2$，所以 $a-c=\pm b$，即 $c=a-b$ 或 $c=a+b$。由於 $b\neq 0$，這兩個值並不相同，於是得出兩個解

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

（我們還應該檢查有沒有遺漏：在 $y>0$ 的區域上，$y'=\pm 2\sqrt{y}$ 而正負號固定，並且 $\frac{\mathrm{d}}{\mathrm{d}x}\sqrt{y}=\frac{y'}{2\sqrt{y}}=\pm 1$，所以 $\sqrt{y}=\pm(x-c)$，得出的正是這些拋物線。因此在給定的點附近，只要 y 保持為正，就只有這兩個解。不過，解一旦到達 x 軸，便可以與零解黏合起來，我們將在第 3 階段看到這一點，而這樣會多出無限多個對所有 x 都有定義的解。題目所說的「兩個解」，就是這兩條拋物線。）

階段 2　兩個點與一個共同的解

經過原點的解必定是 $y=x^2$（即 $c=0$ 的那條拋物線）。它在 $x=1$ 和 $x=-1$ 處達到高度 1。所以取 $(a_1,1)=(1,1)$ 和 $(a_2,1)=(-1,1)$。

經過 $(1,1)$（即 $a=1$，$b=\pm 1$）的解是 $y=x^2$ 和 $y=(x-2)^2$。經過 $(-1,1)$ 的解是 $y=(x+2)^2$ 和 $y=x^2$。兩種情況下，經過原點的解都是同一條曲線 $y=x^2$。 $\blacksquare$

階段 3　整個曲線族與一條公切線

停一停，試一試　就不同的 c 值，畫出幾條拋物線 $y=(x-c)^2$ 的草圖。哪一條直線與它們每一條都相切？這條直線本身是不是解？

每一條拋物線 $y=(x-c)^2$ 都在其頂點 $(c,0)$ 與 x 軸相切，所以 x 軸是整個曲線族的公切線（common tangent），亦即它們的包絡線（envelope）（圖 8.1.1）。而 x 軸，即 $y=0$，本身也是這條微分方程的解。它不是那些拋物線的其中一條；它是另外一個獨立的解，稱為奇解（singular solution）。

階段 4　為何重要

想一想　從原點出發。哪些解經過原點？你能否構造一個解，先沿着 x 軸走一段，然後離開？

經過原點的有 $y=x^2$ 和 $y=0$。但事實遠不止於此。由於每條拋物線都與 x 軸相切（兩

者在該處的斜率都是 0），你可以把各段平滑地黏合起來：對任意 $c_1 \le 0 \le c_2$，

$$y = \begin{cases} (x-c_1)^2 & x < c_1, \\ 0 & c_1 \le x \le c_2, \\ (x-c_2)^2 & x > c_2 \end{cases}$$

是一個可微函數（differentiable function），處處滿足 $(y')^2 = 4y$，並且經過原點。經過 x 軸上每一點的解都有無限多個：一個解可以隨意在 $y = 0$ 上停留多久，然後才離開。這條公切線正正就是唯一性失效的地方。 ■

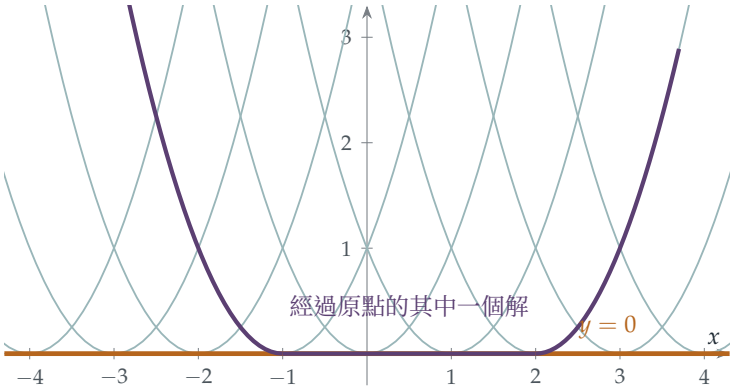


圖 8.1.1 $(y')^2 = 4y$ 的解 $y = (x-c)^2$（藍綠色）全都與 x 軸相切；這條軸是它們的包絡線，本身也是解（琥珀色）。把各段黏合起來可得出更多的解，例如紫紅色的那一個：它沿着一條拋物線落到軸上，沿軸而行，然後在自己選定的時刻再次離開。經過 x 軸上任何一點的解都有無限多個。

回顧

- **先找簡單的解。**試一試 $y = (x-c)^2$ 和 $y = 0$，便看清了整個局面。
- **包絡線給出奇解。**一族解的公切線，本身可以是一個不屬於該族的解。
- **黏合。**兩個解相切之處，可以把它們接合成一個新的可微解。非唯一性就是這樣蔓延開去的。

背後的數學

解在甚麼時候是唯一的？

皮卡—林德洛夫定理（Picard–Lindelöf theorem，由皮卡（Émile Picard）和林德洛夫（Ernst Lindelöf）約於 1890–1894 年證明）指出，方程 $y' = F(x, y)$ 連同 $y(x_0) = y_0$ 在 x_0 附近恰好有一個解，條件是 F 對 y 滿足利普希茨（Lipschitz）條件：存在常數 L 使得 $|F(x, y_1) - F(x, y_2)| \le L|y_1 - y_2|$。粗略地說，當 y 變化時，F 不可以變化得無限快。我們的方程是 $y' = \pm 2\sqrt{y}$，而 $\sqrt{y}$ 在 $y = 0$ 處的斜率是無限大：它在該處不滿足利普希茨條件，而那裏正是唯一性失效的地方。（皮亞諾（Giuseppe Peano）在 1886 年證明了單憑 F 的連續性（continuity）已能保證解*存在*，但不能保證解唯一。）皮卡–林德洛夫定理的證明用到一個迭代（iteration），它憑藉 Section 2.2 的壓縮映射原理（contraction mapping principle）而收斂。

漏水的水桶

以下是一個物理例子。一個底部有孔的水桶按照托里拆利定律（Torricelli's law）流乾：水的高度 h 滿足 $h' = -k\sqrt{h}$。讓時間向前走，一切都沒有問題：水桶流乾，然後一直是空的。但讓時間倒流：假如你發現一個空桶，這條方程無法告訴你它是甚麼時候流乾的。在所有時間都有 $h = 0$ 的解，以及那些直到某一刻仍然盛滿、然後才流乾的解，全都與「現在是空的」相符。這種非唯一性，正是「關於過去的資訊已經失去」這個事實在數學上的投影。

Norton 圓頂

2003 年，哲學家 John Norton 描述了一個無摩擦的圓頂，其形狀使得一個恰好放在頂端的球所遵從的牛頓定律，形式為 $r'' = \sqrt{r}$。其中一個解是球永遠停在頂端。但還有其他的解：球在頂端停留任意長的時間，然後在沒有任何原因的情況下，自發地朝任意方向滾下。這個例子至今仍有爭論：古典力學到底是否決定論（deterministic）的？從數學上看，這正是本題所見的那種非唯一性。

思維技巧

不要假設唯一性

找到一個解，並不表示它是唯一的解。要問一問是否還有其他解，尤其是當方程涉及不光滑的函數時。

在大學裏，存在性與唯一性定理（existence and uniqueness theorems）會確切地告訴你需要甚麼條件。在這裏，一個在零處斜率無限大的平方根，已足以破壞唯一性。

延伸練習

1 **兩種增長率。**試做 STEP II 2018，第 8 題（`step.maths.org/questions/18-s2-q8`）。該題用代換法（substitution）解兩條增長方程。然後想一想，為甚麼 $y = 0$ 也是具有相同初始值的解，而這對唯一性又說明了甚麼。

2 **漏水的水桶。**解 $h' = -k\sqrt{h}$，其中 $h(0) = h_0 > 0$。水桶在甚麼時候變空？寫出兩個不同的解，使得對某個 T 有 $h(T) = 0$，並且對所有 $t \geq T$ 有 $h(t) = 0$。

3 **恢復唯一性。**證明 $y' = 2y$ 連同 $y(0) = 0$ 只有 $y = 0$ 這一個解。（提示：考慮 $\frac{\mathrm{d}}{\mathrm{d}x}(ye^{-2x})$。）$2y$ 具有甚麼性質，是 $2\sqrt{y}$ 所沒有的？

4 **立方。**求 $y' = 3y^{2/3}$ 經過原點的所有解。

延伸閱讀

- **Morris Hirsch, Stephen Smale & Robert Devaney, *Differential Equations, Dynamical Systems, and an Introduction to Chaos***（Academic Press）。講述存在性與唯一性，以及本章其餘的大部分內容。
- **George Simmons, *Differential Equations with Applications and Historical Notes***（CRC Press）。一本經典著作，富於歷史掌故；書中討論了奇解與包絡線。
- **John Norton, 「The dome: an unexpectedly simple failure of determinism」**，*Philosophy of Science* 75 (2008)。篇幅短小，容易閱讀。

8.2 不求解而估界

題目

STEP III 2013，第 7 題

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

`step.maths.org/questions/13-s3-q7`

引子

大多數微分方程（differential equation）都是解不出來的。$y'' + y^3 = 0$ 的解無法用我們熟悉的函數寫成公式，第 (iii) 部分那頭怪物就更不用說了。然而，我們仍然可以對這些解說出很多東西：它們保持有界（bounded）、它們會振盪、它們會漸趨平穩。訣竅在於找出一個由解及其導數（derivative）構成、而且永不增加的量。這樣一來，解就被困住了。

在物理學中，這個量通常就是能量（energy）。在碗中滾動的球不可能爬到比起始高度更高的地方，因為能量守恆，而在有摩擦時能量更會減少。本題把同樣的推理用於三條方程，其中最後一條是特意挑選的，要你自己發明那個「能量」。這個方法由李雅普諾夫（Aleksandr Lyapunov）於 1892 年提出，是研究微分方程最重要的工具之一。

動筆之前

? **想一想** 在第 (i) 部分，把 y 看成質點的位置，把 x 看成時間。那麼 $y'' = -y^3$ 就是牛頓定律，其中力為 $-y^3$。動能（kinetic energy）和勢能（potential energy）分別是甚麼？

動能是 $\frac{1}{2}(y')^2$（取質量為 1），而力 $-y^3$ 來自勢函數 $V(y) = \frac{1}{4}y^4$，因為力等於 $-V'(y)$。總能量是 $\frac{1}{2}(y')^2 + \frac{1}{4}y^4$，即 E 的一半。能量守恆（conservation of energy）所說的，正正就是 E 為常數。

逐步解題

階段 1　守恆

由鏈式法則（chain rule），

$$\frac{\mathrm{dE}}{\mathrm{d}x} = 2y'y'' + 2y^3y' = 2y'(y'' + y^3) = 0.$$

所以 E 是常數，等於它在 $x = 0$ 時的值：$\mathrm{E} = 0 + \tfrac{1}{2} = \tfrac{1}{2}$。由於 $(y')^2 \geq 0$，我們得到 $\tfrac{1}{2}y^4 \leq \tfrac{1}{2}$，所以 $y^4 \leq 1$，即 $|y| \leq 1$。 ■

這個解在 -1 與 1 之間振盪，在 (y, y') 平面上描出閉曲線 $(y')^2 + \tfrac{1}{2}y^4 = \tfrac{1}{2}$（圖 8.2.1）。

階段 2　耗散

想一想　對 $\mathrm{E} = (v')^2 + 2\cosh v$ 求導。利用微分方程把 v'' 換掉。結果的正負號是甚麼？

$$\frac{\mathrm{dE}}{\mathrm{d}x} = 2v'v'' + 2\sinh v\, v' = 2v'(v'' + \sinh v) = 2v'(-xv') = -2x(v')^2 \leq 0 \quad \text{當} x \geq 0.$$

xv' 這一項的作用有如摩擦，會帶走能量。所以當 $x \geq 0$ 時，$\mathrm{E}(x) \leq \mathrm{E}(0)$。在 $x = 0$ 處，$\cosh(\ln 3) = \frac{3+\frac{1}{3}}{2} = \frac{5}{3}$，所以 $\mathrm{E}(0) = \frac{10}{3}$。因此

$$2\cosh v \leq (v')^2 + 2\cosh v = \mathrm{E}(x) \leq \tfrac{10}{3}, \qquad \text{所以} \qquad \cosh v(x) \leq \tfrac{5}{3}.$$ ■

階段 3　發明能量

第 (iii) 部分沒有給任何提示。我們需要一個不會增加的函數 $\mathrm{E} = (w')^2 + F(w)$。

想一想　若 $\mathrm{E} = (w')^2 + F(w)$，則 $\mathrm{E}' = 2w'w'' + F'(w)w'$。$F'(w)$ 應該是甚麼，才可以利用方程把 w'' 換掉，最後只剩下 $(w')^2$ 的倍數？

我們希望 $\mathrm{E}' = 2w'(w'' + w\cosh w + 2\sinh w)$，所以需要 $F'(w) = 2(w\cosh w + 2\sinh w)$。由於 $\frac{\mathrm{d}}{\mathrm{d}w}(w\sinh w) = \sinh w + w\cosh w$，我們可以取

$$F(w) = 2w\sinh w + 2\cosh w, \qquad \mathrm{E} = (w')^2 + 2w\sinh w + 2\cosh w.$$

於是，利用方程可得

$$\frac{\mathrm{dE}}{\mathrm{d}x} = -2(5\cosh x - 4\sinh x - 3)(w')^2.$$

停一停，試一試 證明對所有 x 都有 $5\cosh x - 4\sinh x - 3 \geq 0$。（把它用 e^x 表示，然後找出一個完全平方。）然後完成整個論證。

$5\cosh x - 4\sinh x - 3 = \frac{5}{2}(e^x + e^{-x}) - 2(e^x - e^{-x}) - 3 = \frac{1}{2}e^x - 3 + \frac{9}{2}e^{-x} = \frac{1}{2}\big(e^{x/2} - 3e^{-x/2}\big)^2 \geq 0$。所以 E 永不增加。在 $x = 0$ 處，$\mathrm{E}(0) = \frac{1}{2} + 0 + 2 = \frac{5}{2}$。由於對任何 w 都有 $w\sinh w \geq 0$（兩個因子同號），

$$2\cosh w \leq (w')^2 + 2w\sinh w + 2\cosh w = \mathrm{E}(x) \leq \tfrac{5}{2}, \qquad \text{所以} \qquad \cosh w(x) \leq \tfrac{5}{4}. \quad \blacksquare$$

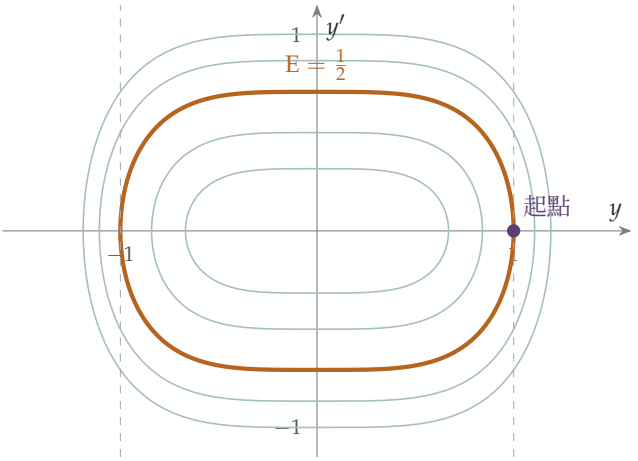


圖 8.2.1 能量 $\mathrm{E} = (y')^2 + \frac{1}{2}y^4$ 在 (y, y') 平面上的等高線。由於 E 是常數，第 (i) 部分的解從 $y = 1$、$y' = 0$ 出發，永遠留在曲線 $\mathrm{E} = \frac{1}{2}$（琥珀色）上，因而永不離開帶狀區域 $|y| \leq 1$（虛線）。我們根本不必找出 y 的公式，就已經知道這一點。

回顧

- **乘以 y'。**在每一部分，E′ 都是 $2y'$ 乘以方程的左方，再加上餘下的項。正是這個餘下的項決定了 E 是守恆還是遞減。
- **能量限制位置。**由於動能部分 $(y')^2$ 非負，E 的上界就是勢能部分的上界，從而也給出 y 的界。
- **留意隱藏的平方。**第 (iii) 部分的阻尼係數看來可能會變號，但它其實是一個經過偽裝的完全平方。

背後的數學

李雅普諾夫函數

李雅普諾夫在 1892 年的博士論文中研究運動的穩定性（stability）：如果一個系統受到輕微擾動而偏離平衡，它會回到平衡，還是愈走愈遠？他的「第二方法」是找出一個函數，它在平衡點處取最小值，而且沿着解永不增加；這種函數現在稱為李雅普諾夫函數（Lyapunov function）。這樣，從平衡點附近出發的解就會一直留在平衡點附近。這個方法的美妙之處，在於它不需要解的公式，只需要微分方程本身，而這正是本題所展示的。現今，從飛機控制系統、電網以至生態系統，各種各樣事物的穩定性都是用李雅普諾夫函數來證明的。

物理與工程中的能量

第 (i) 部分是一個非線性振子（nonlinear oscillator）：它像彈簧，但回復力與位移的立方成正比。與簡諧振子不同，它的週期取決於振幅（振盪愈大，速度愈快），但能量守恆仍然把它困住。第 (ii) 部分加入了隨時間增強的阻尼（damping），能量因而逐漸流失。在工程學中，人們用能量方法來證明結構不會把自己搖散，以及控制系統不會振盪至失控。在偏微分方程（partial differential equation）的理論中，這類「能量估計」是證明解存在而且性質良好的主要工具；有關流體流動方程的問題，包括千禧年大獎難題（Millennium Prize Problems）之一的納維－斯托克斯（Navier–Stokes）問題，歸根究底就是能否找到這樣的界。

思維技巧

找出一個不能增加的量

如果你解不出一條方程，就找一個由解構成、而且導數有確定正負號的量。它會給出解的界。

這種尋找一方面由物理引導（「能量是甚麼？」），另一方面由代數引導（「我要加上甚麼，導數才會化簡？」）。第 (iii) 部分示範了在沒有物理直覺可以依靠時，如何走代數這條路。

延伸練習

1 **簡諧運動。**對於 $y'' + y = 0$，其中 $y(0) = A$、$y'(0) = 0$，求守恆的能量，並由此推出 $|y| \le A$。用精確解核對。

2 **單擺。**單擺的角度 θ 滿足 $\theta'' + \sin\theta = 0$（採用適當的單位）。證明 $(\theta')^2 - 2\cos\theta$ 是常數。若單擺在角度 θ_0 處由靜止開始，其中 $|\theta_0| \le \pi$，證明 $|\theta|$ 永不超過 $|\theta_0|$。若 $|\theta_0| > \pi$，哪裏會出問題？

3 **阻尼。**對於 $y'' + 2ky' + y = 0$，其中 $k > 0$，證明 $(y')^2 + y^2$ 遞減。它是否必定趨向零？

4 **逃逸。**對於 $y'' - y^3 = 0$，其中 $y(0) = 1$、$y'(0) = 0$，求一個守恆量，並證明它不能給出 y 的界。哪裏出了問題？

延伸閱讀

- **Steven Strogatz, *Nonlinear Dynamics and Chaos***（CRC Press）。第 6 章和第 7 章講述相平面、保守系統和李雅普諾夫函數。
- **D. W. Jordan & P. Smith, *Nonlinear Ordinary Differential Equations***（Oxford University Press）。一本標準的大學本科教科書，載有大量能量方法的例子。
- **Keith Devlin, *The Millennium Problems***（Basic Books）。書中有一篇關於納維－斯托克斯問題的淺白介紹。

8.3 收穫種羣

題目

STEP II 1995，第 8 題

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

`step.maths.org/questions/95-s2-q8`

引子

食物供應無限的種羣會以指數增長（exponential growth）的方式擴張。現實中的種羣卻總會用盡空間：當種羣數量接近環境所能維持的上限，即其*環境承載力*（carrying capacity）時，增長便會放緩。描述這個現象的最簡單模型是*邏輯斯諦方程*（logistic equation），它出現在本題的前半部分，由培養皿中的細菌到新科技的普及，都可以用它來描述。

後半部分所問的，是一個在現實中關係重大的問題：假如我們以穩定的速率收穫（harvest）這個種羣，會發生甚麼事？在某個閾值（threshold）以下，種羣會調整到一個新的、較低的平衡點（equilibrium）。但只要收穫率越過這個閾值，哪怕只是超出少許，平衡點便會消失，種羣數量隨即崩潰至零。這就是漁業崩潰背後的數學，也是設立捕魚配額的原因。

動筆之前

想一想 在動手求解之前，先把增長率 $(2K - x)x$ 當作 x 的函數，畫出它的略圖。它在哪裏為零？在哪裏為正？這告訴你 x 會怎樣變化？

增長率在 $x = 0$ 和 $x = 2K$ 處為零，在兩者之間為正，超過 $2K$ 則為負。所以，由 0 與 $2K$ 之間出發的種羣會向 $2K$ 增長，而由 $2K$ 以上出發的種羣則會向它縮減。環境承載力是 $2K$，而增長在 $x = K$ 時最快，那時的增長率是 K^2。不必解方程，你也能預測長遠的變化。

逐步解題

階段 1　解邏輯斯諦方程

想一想　分離變量（separate the variables）：$\displaystyle\int \frac{\mathrm{d}x}{(2K-x)x} = \int \mathrm{d}t$。把左邊拆成部分分式（partial fractions）。

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

$$\frac{1}{2K}\ln\frac{x}{2K-x} = t + C.$$

當 $t=0$ 時，$x=K$，所以 $\ln 1 = 0$，而 $C=0$。於是 $\frac{x}{2K-x} = e^{2Kt}$，由此得

$$x = \frac{2Ke^{2Kt}}{1+e^{2Kt}} = \frac{2K}{1+e^{-2Kt}} = K + K\frac{1-e^{-2Kt}}{1+e^{-2Kt}}. \qquad \blacksquare$$

這個分式就是 $\tanh(Kt)$，所以 $x = K + K\tanh Kt$：這是一條由 K 上升並趨向 $2K$ 的 S 形曲線。當 $t\to\infty$ 時，$e^{-2Kt}\to 0$，而 $x\to 2K$。

階段 2　收穫

以速率 L 移走細菌，便得

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

想一想　對 $(2K-x)x - L$ 進行配方（complete the square）。它的零點在哪裏？

$(2K-x)x - L = K^2 - L - (x-K)^2 = \alpha^2 - (x-K)^2 = \big(\alpha-(x-K)\big)\big(\alpha+(x-K)\big)$。平衡點是 $x = K\pm\alpha$。

停一停，試一試　現在作代換（substitution）$y = x - K + \alpha$。y 滿足甚麼方程？把它與原來的邏輯斯諦方程比較一下。

設 $y = x-K+\alpha$，則 $x-K = y-\alpha$，所以 $\alpha+(x-K) = y$，而 $\alpha-(x-K) = 2\alpha - y$。因此

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

這又是邏輯斯諦方程，只是 K 換成了 α。初始時 $y(0) = K - K + \alpha = \alpha$，正好對應第一部分的 $x(0)=K$。所以根據第 1 階段的結果（以 α 代替 K），$y(t) = \alpha + \alpha\tanh\alpha t \to 2\alpha$，因此

$$x(t) = K + \alpha\tanh\alpha t \to K+\alpha. \qquad \blacksquare$$

回顧

- **先作定性分析。**還未動手求解，增長率的正負號已經預測了長遠的變化。
- **配方揭示平衡點。**有收穫的方程的平衡點在 $K \pm \alpha$。
- **代換把新問題化為舊問題。**把變量平移之後，有收穫的方程便變成了早已解出的邏輯斯諦方程。

背後的數學

臨界點

有收穫的方程有兩個平衡點，$K + \alpha$ 和 $K - \alpha$。上面的一個是穩定（stable）的：種羣數量稍高於或稍低於它，都會回到它那裏。下面的一個是不穩定（unstable）的：種羣數量只要稍低於 $K - \alpha$，便會一直下跌，直至滅絕。隨着收穫率 L 增加，$\alpha = \sqrt{K^2 - L}$ 會縮小，兩個平衡點互相靠近；到 $L = K^2$ 時兩者合而為一，而當 $L > K^2$ 時則完全沒有平衡點，任何種羣都會崩潰（圖 8.3.1）。這是一個鞍結分岔（saddle-node bifurcation），是最簡單的一種臨界點（tipping point）。它的可怕之處，在於事前沒有任何明顯的預兆：當 L 接近 K^2 時，平衡種羣數量只是緩緩下降，然後平衡點突然不復存在。事後才減少收穫未必有用，因為種羣數量可能已經跌到不穩定的閾值以下。

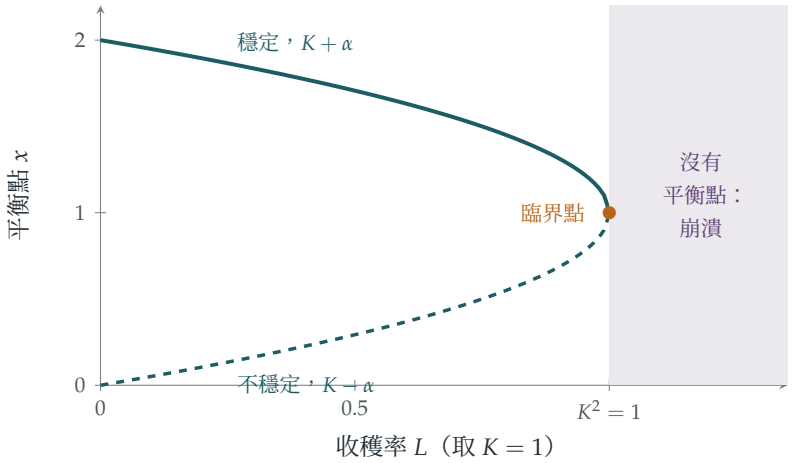


圖 8.3.1 平衡種羣數量與收穫率 L 的關係，其中 $K = 1$。隨着 L 增加，穩定平衡點（實線）與不穩定閾值（虛線）互相靠近，並在 $L = K^2$ 處相遇。越過這一點之後，平衡點不再存在，無論種羣有多大都會崩潰。

最大可持續產量

種羣所能承受的最大收穫量是多少？答案是 $L = K^2$，即增長率 $(2K - x)x$ 的峯值，在種羣數量維持於 K（環境承載力的一半）時達到。二十世紀中葉，漁業科學採納了這個「最大可持續產量」（maximum sustainable yield）作為管理目標。可是，恰好按最大值收穫，就等於把系統恰好放在臨界點上，只要稍為高估了 K，或者遇上一個壞年頭，便會引致崩潰。紐芬蘭對開的大淺灘（Grand Banks）鱈魚漁場，五百年來一直是全世界最豐饒的漁場之一，卻在 1990 年代初崩潰；1992 年，加拿大政府頒布禁捕令，令數以萬計的人失去工作。魚類資源至今仍未完全恢復。

其他地方的邏輯斯諦增長

邏輯斯諦方程由 Pierre-François Verhulst 於 1838 年提出，用來為人類人口建立模型。同一條 S 形曲線，也描述了疫症初期的傳播、新科技的普及、自催化化學反應以及腫瘤的生長。它的離散時間版本，即邏輯斯諦映射（logistic map），表現卻截然不同：它可以是混沌（chaotic）的，我們會在 Section 8.7 看到。

思維技巧

平衡點及其穩定性

對於方程 $x' = f(x)$，平衡點就是 f 的零點，而 f 在各零點之間的正負號，告訴你解會向哪個方向移動。

若 f 在某個平衡點由正變負，該平衡點是穩定的；若由負變正，則是不穩定的。畫出 f 的圖像，往往不必求解任何東西，便能回答最重要的問題。

延伸練習

1. **由低處出發。**在有收穫且 $K^2 > L$ 的情況下，證明若 $x(0) < K - \alpha$，則種羣數量會在有限時間內降至零。（提示：當 $y(0) < 0$ 時，y 會怎樣變化？）

2. **過度收穫。**若 $L > K^2$，設 $\beta = \sqrt{L - K^2}$。在 $x(0) = 2K$ 的條件下解 $x' = -(x - K)^2 - \beta^2$，並求種羣數量何時降至零。

3. **按比例收穫。**假設改為移走種羣中某個固定的比例：$x' = (2K - x)x - hx$。求平衡點，並證明對所有 $h < 2K$，種羣都能存活。為甚麼這個方案較為安全？

4. **半程時間。**在沒有收穫的情況下，求種羣數量達到 $\frac{3}{2}K$ 的時間。

延伸閱讀

- **Steven Strogatz, *Nonlinear Dynamics and Chaos***（CRC Press）。第 3 章討論分岔，並以這個收穫模型作為習題。
- **Nicholas Britton, *Essential Mathematical Biology***（Springer）。為大學本科生而寫的種羣模型。
- **Mark Kurlansky, *Cod: A Biography of the Fish That Changed the World***（Penguin）。鱈魚漁業及其崩潰的歷史。

8.4 盤旋趨向圓

題目

STEP III 2015，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/15-s3-q8`

引子

自然界中有許多系統，無論起始狀態如何，最終都會進入同一種規律的節奏：心臟的跳動、蟋蟀的鳴叫、穩定拉奏的小提琴弦、捕食者與獵物的種羣數量週期。稍加擾動，它們又會回到原來的節奏。用微分方程（differential equation）的語言來說，它們的解會趨近一條閉合曲線，稱為極限環（limit cycle）。

這道題目正是要構造一個極限環。方程以 x 和 y 表示時看似無從入手，但換成極坐標（polar coordinates）後，就化簡成可以求解的形式，而且各個解原來都會從內側和外側盤旋着貼近單位圓。當中的關鍵一步，是選取與問題的對稱性（symmetry）相配的坐標，這也是本書一直沿用的手法。

動筆之前

想一想 兩條方程其實都暗藏 $x\,\mathrm{d}x + y\,\mathrm{d}y$ 和 $x\,\mathrm{d}y - y\,\mathrm{d}x$ 這兩個組合。它們在極坐標中是甚麼？

由於 $r^2 = x^2 + y^2$，求導後得 $x\,\mathrm{d}x + y\,\mathrm{d}y = r\,\mathrm{d}r$。另外 $x\,\mathrm{d}y - y\,\mathrm{d}x = r^2\,\mathrm{d}\theta$（這就是 Section 4.9 中的面積公式）。有了這兩條恆等式（identity），變數代換（change of variables）幾乎毫不費力：只要把方程改寫成只含這兩個組合的形式即可。

逐步解題

階段 1　以極坐標處理第 (i) 部分

把方程寫成 $(y+x)\,\mathrm{d}y-(y-x)\,\mathrm{d}x=0$，再重新整理：

$$(x\,\mathrm{d}x+y\,\mathrm{d}y)+(x\,\mathrm{d}y-y\,\mathrm{d}x)=0.$$

想一想　展開算式，驗證以上的整理無誤。然後運用上面兩條恆等式。

$(y+x)\,\mathrm{d}y-(y-x)\,\mathrm{d}x = y\,\mathrm{d}y+x\,\mathrm{d}y-y\,\mathrm{d}x+x\,\mathrm{d}x$，正是上面的算式。化成極坐標形式就是 $r\,\mathrm{d}r+r^2\,\mathrm{d}\theta=0$，兩邊除以 $r^2\,\mathrm{d}\theta$（其中 $r>0$），得

$$\frac{\mathrm{d}r}{\mathrm{d}\theta}+r=0. \qquad \blacksquare$$

方程的解是 $r=Ce^{-\theta}$：這些是對數螺線（logarithmic spiral），隨着 θ 增加而向原點捲入，隨着 θ 減少而向無窮遠處捲出。

階段 2　第 (ii) 部分

停一停，試一試　新方程是在舊方程之上，加入含有 $x^2+y^2=r^2$ 的額外項。用同樣的方法改寫，並求出 $\frac{\mathrm{d}r}{\mathrm{d}\theta}$。

方程是 $\big(y+x-xr^2\big)\,\mathrm{d}y-\big(y-x-yr^2\big)\,\mathrm{d}x=0$，即

$$\big[(y+x)\,\mathrm{d}y-(y-x)\,\mathrm{d}x\big]-r^2\big(x\,\mathrm{d}y-y\,\mathrm{d}x\big)=0.$$

第一個括號和之前一樣，等於 $r\,\mathrm{d}r+r^2\,\mathrm{d}\theta$，第二個括號則是 $r^2\,\mathrm{d}\theta$。因此 $r\,\mathrm{d}r+r^2(1-r^2)\,\mathrm{d}\theta=0$，由此得

$$\frac{\mathrm{d}r}{\mathrm{d}\theta}=r^3-r.$$

想一想　這條方程對 r 而言並非線性（linear），但換成 $u=r^{-2}$ 後就變成線性。u 滿足甚麼方程？

設 $u=r^{-2}$，則 $\frac{\mathrm{d}u}{\mathrm{d}\theta}=-2r^{-3}\frac{\mathrm{d}r}{\mathrm{d}\theta}=-2r^{-3}(r^3-r)=-2+2u$。所以 $\frac{\mathrm{d}u}{\mathrm{d}\theta}-2u=-2$，其解為 $u=1+Ae^{2\theta}$。因此

$$r^2=\frac{1}{1+Ae^{2\theta}}. \qquad \blacksquare$$

階段 3　三類解

- $A = 0$：$r = 1$。單位圓是一個解。
- $A > 0$：恆有 $r < 1$。當 $\theta \to -\infty$ 時，$r \to 1$；當 $\theta \to +\infty$ 時，$r \to 0$。曲線在圓與原點之間盤旋。
- $A < 0$：$r > 1$，而且公式只在 $1 + Ae^{2\theta} > 0$ 時才有意義，即 $\theta < \frac{1}{2}\ln\left(-\frac{1}{A}\right)$。當 $\theta \to -\infty$ 時，$r \to 1$；當 θ 趨近該上界時，$r \to \infty$。曲線在圓與無窮遠之間盤旋。

所以當 $\theta \to -\infty$ 時，每條解曲線都會從內側或外側趨近單位圓（圖 8.4.1）。

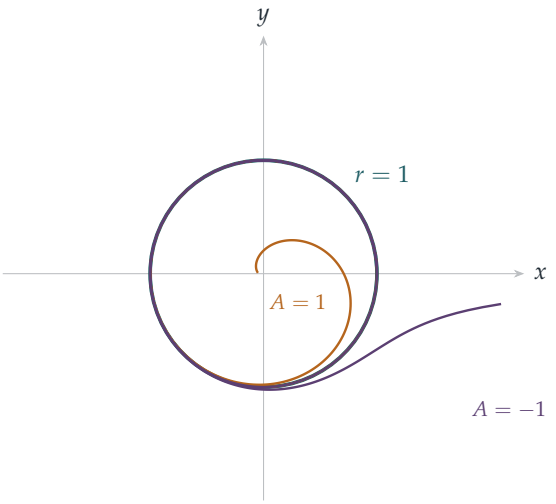


圖 8.4.1 第 (ii) 部分的解。單位圓（藍綠色）是一個解。$A = 1$ 的解（琥珀色）位於圓內，從圓出發（當 $\theta \to -\infty$ 時）盤旋着走向原點；$A = -1$ 的解（紫紅色）位於圓外，從圓出發盤旋着走向無窮遠。反方向來看（θ 遞減），每個解都會纏繞到圓上：這就是極限環。

回顧

- **善用對稱性。**這些方程對所有方向一視同仁，而極坐標令這一點顯露出來。
- **認出標準組合。**$x\,\mathrm{d}x + y\,\mathrm{d}y = r\,\mathrm{d}r$ 和 $x\,\mathrm{d}y - y\,\mathrm{d}x = r^2\,\mathrm{d}\theta$ 完成了大部分工作。
- **代換可使方程線性化。**非線性方程 $r' = r^3 - r$ 換成 $u = r^{-2}$ 後變成線性（這是一條*伯努利方程*（Bernoulli equation））。

背後的數學

極限環

龐加萊（Henri Poincaré）在 1880 年代研究平面上微分方程的定性行為時，引入了極限環。不妨把題目中的方程理解為描述一個點在平面上運動的路徑。單位圓是一條孤立的閉軌道（closed orbit），附近的每條軌道都會纏繞着向它靠近（或遠離它，視乎時間的方向而定）。1926 年，荷蘭工程師 Balthasar van der Pol 在一個裝有真空管的電路中發現了極限環，並提出類似的方程可以描述心臟的跳動。凡是具有自持而穩定節奏的系統，都可以用極限環作為模型：生理時鐘、神經脈衝、捕食者與獵物種羣數量的週期。

平面上沒有混沌

龐加萊—本迪克松定理（Poincaré–Bendixson theorem）指出，對於平面上的光滑系統，如果一條軌道一直留在某個有界區域內，又不趨近任何平衡點（equilibrium），它就必定趨近一條閉軌道。所以在二維的情況下，長期行為只限於平衡點和環：混沌（chaos）沒有容身之處，因為軌道不能相交。混沌至少需要三維，正如洛倫茲（Edward Lorenz）在 1963 年以他那著名的三變數對流模型（「蝴蝶」）所發現的那樣。至於離散時間的情況，我們會在 Section 8.7 看到，即使只有一維，混沌也可能出現。

一個未解決的問題

1900 年，希爾伯特（David Hilbert）在他的第十六問題中問道：當 P 和 Q 是 n 次多項式時，系統 $x' = P(x, y)$，$y' = Q(x, y)$ 最多可以有多少個極限環？即使是 $n = 2$ 的情況，答案至今仍然未知：已知有些系統具有四個極限環，但是否可能有更多，則無人知曉。在希爾伯特的問題之中，這是極少數至今仍然完全懸而未決的一個。

思維技巧

轉換坐標以配合對稱性

如果問題具有旋轉對稱性，就用極坐標；如果具有其他對稱性，就找出與之相配的坐標。

合適的坐標把一條難以處理的方程變成可分離變數（separable）的方程，並令解的幾何形狀（圓和螺線）一目了然。

延伸練習

1 **螺線。**對於第 (i) 部分的解 $r = e^{-\theta}$，證明曲線與每條由原點出發的射線都以相同的角度 45° 相交。（這就是對數螺線又稱為等角螺線（equiangular spiral）的原因。）

2 **極坐標恆等式。**證明 $x\,\mathrm{d}x + y\,\mathrm{d}y = r\,\mathrm{d}r$ 和 $x\,\mathrm{d}y - y\,\mathrm{d}x = r^2\,\mathrm{d}\theta$。

3 **隨時間變化的系統。**考慮系統 $\dot{r} = r(1 - r^2)$，$\dot{\theta} = 1$。證明當 $t \to \infty$ 時，每個滿足 $r(0) > 0$ 的解都趨近單位圓。把這個系統以 x 和 y 表示。

4 **兩個環。**找出一個形如 $\dot{r} = f(r)$，$\dot{\theta} = 1$ 的系統，使它有兩個極限環，分別位於 $r = 1$ 和 $r = 2$，其中一個是吸引的（attracting），另一個是排斥的（repelling）。

延伸閱讀

- **Steven Strogatz, *Nonlinear Dynamics and Chaos***（CRC Press）。第 7 章講述極限環和龐加萊–本迪克松定理。
- **Steven Strogatz, *Sync***（Hyperion）。一本講述自然界中的節奏與同步現象的科普讀物。
- **James Gleick, *Chaos: Making a New Science***（Penguin）。講述洛倫茲與發現混沌的故事。

8.5 格力犬追逐格力犬

題目

STEP III 1988，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/88-s3-q10`

引子

四隻狗分別站在一塊正方形田地的四個角上，每隻狗都追趕牠左邊的那一隻。牠們會在哪裏相遇？要花多少時間？所走的路徑又是怎樣的？這些路徑是優美的螺線，繞着中心轉無窮多圈，然而每隻狗只跑了有限的距離，恰好等於正方形的邊長。

這是一個追逐問題（pursuit problem），這類問題最早由 Pierre Bouguer 於 1732 年研究，當時考慮的是一艘船追趕海盜船的情形。令問題得以解決的訣竅是對稱性（symmetry）：發生在一隻狗身上的事，同樣發生在其餘各隻身上，所以四隻狗始終構成一個正方形。一旦看出這一點，微分方程（differential equation）幾乎就自然而然地寫出來了。

動筆之前

? **想一想** 根據對稱性，四隻狗在每一刻構成甚麼形狀？它的中心在哪裏？

把整個情形繞正方形的中心旋轉四分之一圈（即 $A \to B \to C \to D \to A$），情形保持不變，所以四隻狗始終構成一個正方形，而且中心始終是同一點 O。這個正方形一邊旋轉一邊縮小，但始終是正方形。特別地，A 和 B 與 O 的距離始終同為 r，而 OB 始終比 OA 超前四分之一圈。

逐步解題

階段 1　運動的方向

想一想　A 朝着 B 跑。在三角形 OAB 中，AB 與 AO（指向中心的方向）所成的角是多少？

三角形 OAB 中 $OA = OB$，且在 O 處是直角，所以其餘兩個角都是 $45°$。因此 A 的速度（velocity）大小為 v，方向與向內的徑向（radial）方向成 $45°$，並且偏向 B 的一方，即 θ 增加的方向（假設我們標記角度時，令 B 位於 A 的逆時針方向）。

階段 2　微分方程

想一想　把 A 的速度分解為徑向和橫向（transverse）兩部分。在極坐標（polar coordinates）中，徑向速度是 $\dot{r}$，橫向速度是 $r\dot{\theta}$。

徑向分量（向內）是 $v\cos 45°$，橫向分量是 $v\sin 45°$，所以

$$\dot{r} = -\frac{v}{\sqrt{2}}, \qquad r\dot{\theta} = \frac{v}{\sqrt{2}}.$$

兩式相除，得 $\dfrac{\mathrm{d}r}{\mathrm{d}\theta} = \dfrac{\dot{r}}{\dot{\theta}} = -r$，所以 $r = \lambda e^{-\theta}$，其中若最初 $\theta = 0$，則 λ 是最初與中心的距離。 ■

階段 3　n 隻狗

停一停，試一試　對於正 n 邊形，三角形 OAB 中在 O 處的角是多少？AB 與 AO 所成的角是多少？重複第 2 階段的做法。

在 O 處的角是 $\frac{2\pi}{n}$，所以兩個相等的底角是 $\frac{1}{2}(\pi - \frac{2\pi}{n}) = \frac{\pi}{2} - \frac{\pi}{n}$。速度與向內的半徑所成的角就是這個角，所以

$$\dot{r} = -v\cos\left(\frac{\pi}{2} - \frac{\pi}{n}\right) = -v\sin\frac{\pi}{n}, \qquad r\dot{\theta} = v\sin\left(\frac{\pi}{2} - \frac{\pi}{n}\right) = v\cos\frac{\pi}{n}.$$

因此 $\dfrac{\mathrm{d}r}{\mathrm{d}\theta} = -r\tan\dfrac{\pi}{n}$，於是

$$r = \lambda e^{-\theta\tan(\pi/n)}. \qquad ■$$

當 $n = 4$ 時，$\tan\frac{\pi}{4} = 1$，這就回到正方形的情形。

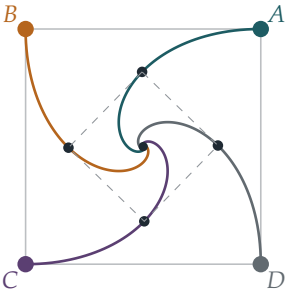


圖 8.5.1 四隻狗從正方形的四個角出發互相追逐。每隻狗都沿着一條對數螺線 $r = \lambda e^{-\theta}$（相對於牠自己的起始方向）前進，繞着中心轉無窮多圈。在每一刻，四隻狗都構成一個正方形（虛線所示為其中一個中間時刻），這個正方形一邊旋轉一邊縮小，直至牠們在中心相遇。

回顧

- **對稱性把多個物體化為一個。**四隻互相影響的狗，變成了一個遵從簡單法則的質點（particle），因為整個佈局始終是正方形。
- **繞中心的運動宜用極坐標。**速度的徑向和橫向分量 $\dot{r}$ 和 $r\dot{\theta}$ 直接給出了方程。
- **相除以消去時間。**比值 $\dot{r}/\dot{\theta}$ 給出了路徑的形狀，而毋須知道何時到達每一點。

背後的數學

無窮多圈，有限距離

徑向速率 $\dot{r} = -v\sin\frac{\pi}{n}$ 是常數，所以各隻狗在時間 $T = \frac{r_0}{v\sin(\pi/n)}$ 之後到達中心，每隻狗跑過的距離是 $vT = \frac{r_0}{\sin(\pi/n)}$。對於邊長為 s 的正 n 邊形，$r_0 = \frac{s}{2\sin(\pi/n)}$，所以每隻狗跑過的距離是

$$\frac{s}{2\sin^2(\pi/n)}.$$

對於四隻狗，這恰好是 s：每隻狗跑過的距離正好是一條邊的長度。（這裏有一個巧妙的看法：B 的速度始終與 AB 垂直，所以它不會改變距離 AB；只有 A 以速率 v 沿着 AB 筆直前進的運動才會縮短兩者的距離，而這個距離最初是 s。）對於三隻狗，距離是 $\frac{2}{3}s$；對於六隻狗，距離是 $2s$。然而在每一種情形下，螺線都繞着中心轉無窮多圈，因為當 $r \to 0$ 時 $\theta \to \infty$。一條長度有限卻轉了無窮多圈的曲線並不是悖論：那些圈縮小得足夠快。

對數螺線

曲線 $r = \lambda e^{-k\theta}$ 稱為對數螺線（logarithmic spiral）或等角螺線（equiangular spiral）：它與每一條從中心出發的射線相交的角度都相同。它出現在大自然之中，例如鸚鵡螺的殼和某些星系的旋臂，因為保持形狀不變的生長就會產生這種曲線。雅各布·伯努利（Jacob Bernoulli）對它極為着迷，甚至要求把它刻在自己的墓碑上，並附上格言 *Eadem mutata resurgo*（「縱然改變，我仍將依舊復起」），指的是它在任何尺度下看起來都一樣。結果石匠刻的卻是一條阿基米德螺線（Archimedean spiral）。

實際中的追逐

「純追逐」（pure pursuit），即始終筆直朝着目標前進，是導彈或捕食者最簡單的導引法則。但它並不是最好的：朝着目標將會到達的位置前進（「比例導引」，proportional navigation）更有效率，而現代導彈以及某些捕食者（例如蜻蜓和隼）實際上正是這樣做的。追逐與逃避的數學在二十世紀自成一門學科，名為微分博弈（differential games）。

思維技巧

利用對稱性簡化多體問題

如果一個系統具有對稱性，而其法則又不破壞這種對稱性，那麼對稱性就會一直保持下去，你只需追蹤其中一部分。

這個想法能解決力學和動力學中的許多問題；它是對稱與守恆（conservation）之間深刻聯繫的一個簡單例子。

延伸練習

1. **巧妙的論證。**仔細解釋為何在四隻狗的問題中，距離 AB 恰好以速率 v 減少，並由此推出四隻狗在時間 s/v 之後相遇。

2. **三隻狗。**三隻狗從邊長為 s 的等邊三角形的三個角出發。求每條路徑的方程、牠們相遇所需的時間，以及每隻狗跑過的距離。

3. **轉了多少圈？**設四隻狗出發時距離中心 100 m，牠們轉了完整一圈之後距離中心多近？轉了多少個完整的圈之後，牠們與中心的距離在 1 mm 以內？

4. **兩隻狗。**當 $n = 2$（兩隻狗面對面）時會怎樣？公式是否仍然有意義？

延伸閱讀

- **Paul Nahin, *Chases and Escapes: The Mathematics of Pursuit and Evasion***（Princeton University Press）。關於追逐曲線的一切，從 Bouguer 到導彈。
- **Eli Maor, *e: The Story of a Number***（Princeton University Press）。內容包括對數螺線和伯努利的墓碑。
- **D'Arcy Thompson, *On Growth and Form***（Cambridge University Press）。講述大自然中的數學的經典著作，包括螺線式的生長。

8.6 量子振子

題目

STEP III 1988，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/88-s3-q7`

引子

繫在彈簧上的質點，振動的能量可以隨你喜歡：把它拉開少許也好，拉開很遠也好。但在 1900 年，普朗克（Max Planck）發現，要解釋高溫物體所發出的光的顏色，唯一的辦法是假定振動中的原子只能具有某些特定的能量，每級相差一個固定的大小。在原子的尺度上，能量是一份一份的，稱為量子（quanta）。二十五年後，量子力學（quantum mechanics）解釋了箇中原因。

這道題目正包含了這個解釋。題中的微分方程（differential equation）經過單位轉換後，就是一個量子粒子處於彈簧式勢能中的薛定諤方程（Schrödinger equation），亦即量子諧振子（quantum harmonic oscillator）。只有當能量取某些特殊的值時，它才有性質良好的解，即在無窮遠處衰減至零的解；這些能量值以 $n = 0, 1, 2, \ldots$ 標記，彼此等距排列。而且不同的 n 所對應的解是正交（orthogonal）的：它們乘積的積分為零，就像傅立葉級數（Fourier series）中的 $\sin mx$ 與 $\sin nx$ 一樣。量子諧振子是研究分子、固體和光的物理學的起點。

動筆之前

想一想 為甚麼像 $e^{-\lambda x^2}$ 這樣的函數會是一個合理的猜測？當 $x \to \pm\infty$ 時，$e^{-\lambda x^2}$ 及其導數會怎樣？

當 $|x|$ 很大時，方程主要由 $y'' \approx \omega^2 x^2 y$ 支配，而高斯函數（Gaussian）$e^{-\lambda x^2}$ 在 x 很大時滿足 $y'' \approx 4\lambda^2 x^2 y$，只要 $4\lambda^2 = \omega^2$，兩者便吻合。若 $\lambda > 0$，高斯函數及其所有導數衰減至 0 的速度比 x 的任何次冪都要快。因此邊界條件（boundary conditions）會自動

成立；真正要下工夫的，是令函數符合方程。

逐步解題

階段 1　基態

想一想　把 $y_0 = e^{-\lambda x^2}$ 求導兩次，代入 $n = 0$ 時的方程，然後比較 x^2 的係數和常數項。

$y_0' = -2\lambda x e^{-\lambda x^2}$，而 $y_0'' = (4\lambda^2 x^2 - 2\lambda)e^{-\lambda x^2}$。$n = 0$ 時的方程要求

$$4\lambda^2 x^2 - 2\lambda - \omega^2 x^2 = -\omega$$

對所有 x 成立。比較係數（match coefficients）：$4\lambda^2 = \omega^2$ 及 $2\lambda = \omega$。取 $\lambda = \tfrac{\omega}{2}$ 時兩式同時成立（取 $\lambda > 0$，使得 $y_0 \to 0$）。而且當 $x \to \pm\infty$ 時，y_0、$y_0' \to 0$。 ■

階段 2　第一激發態

$y_1 = x e^{-\lambda x^2}$ 滿足 $y_1' = (1 - 2\lambda x^2)e^{-\lambda x^2}$ 及 $y_1'' = (4\lambda^2 x^3 - 6\lambda x)e^{-\lambda x^2}$。$n = 1$ 時的方程要求

$$4\lambda^2 x^3 - 6\lambda x - \omega^2 x^3 = -3\omega x,$$

而同樣取 $\lambda = \tfrac{\omega}{2}$ 時此式成立。y_1 與 y_1' 同樣在 $\pm\infty$ 處衰減至零。 ■

階段 3　一個恆等式

想一想　用積法則（product rule）對 $y_m y_n' - y_n y_m'$ 求導。四項之中有兩項互相抵銷。然後利用微分方程代換 y_n'' 和 y_m''。

$\dfrac{\mathrm{d}}{\mathrm{d}x}(y_m y_n' - y_n y_m') = y_m' y_n' + y_m y_n'' - y_n' y_m' - y_n y_m'' = y_m y_n'' - y_n y_m''$。由方程可得 $y_n'' = (\omega^2 x^2 - (2n+1)\omega) y_n$，對 m 亦有類似的式子。所以

$$y_m y_n'' - y_n y_m'' = y_m y_n[\omega^2 x^2 - (2n+1)\omega - \omega^2 x^2 + (2m+1)\omega] = 2(m-n)\omega y_m y_n.$$ ■

階段 4　正交性

停一停，試一試　把恆等式由 $-\infty$ 積分至 ∞。左方是甚麼？

左方積分後得 $\big[y_m y_n' - y_n y_m'\big]_{-\infty}^{\infty} = 0 - 0 = 0$，因為每個 y 和 y' 在兩端都趨向 0。所以

$2(m-n)\omega\int_{-\infty}^{\infty} y_m y_n\,\mathrm{d}x = 0$，而由於 $m \neq n$ 及 $\omega > 0$，

$$\int_{-\infty}^{\infty} y_m(x)y_n(x)\,\mathrm{d}x = 0. \qquad \blacksquare$$

對於 y_0 和 y_1，這一點很容易直接看出：$y_0 y_1 = xe^{-\omega x^2}$ 是奇函數（odd function）。但這個恆等式毋須知道函數本身是甚麼，便一舉證明了每一對解都正交。

回顧

- **比較係數。**代入猜測的形式，並要求方程對所有 x 成立，便確定了 λ。
- **朗斯基行列式（Wronskian）的技巧。**$y_m y_n' - y_n y_m'$ 這個組合求導後，其中的二階導數可以用方程代換，餘下的式子十分簡單。
- **邊界條件消去邊界項。**解在無窮遠處衰減，令左方積分後為零，從而得出正交性。

背後的數學

薛定諤方程

1926 年，薛定諤（Erwin Schrödinger）提出，量子粒子可用波函數（wave function）$\psi(x)$ 來描述；若粒子的能量為 E，並處於勢能 $V(x)$ 之中，則波函數滿足

$$-\frac{\hbar^2}{2m}\psi'' + V(x)\psi = E\psi.$$

對於彈簧，$V(x) = \tfrac{1}{2}m\Omega^2 x^2$，把 x 重新標度後，這正是題目中的方程，其中 $E = \left(n + \tfrac{1}{2}\right)\hbar\Omega$。物理上要求 ψ 在無窮遠處衰減至零（這樣粒子才確實身處某處），而只有這些能量才能滿足這個要求：這就是量子化（quantisation），是由數學自然推導出來的結果。留意最低的能量並不是零，而是 $\tfrac{1}{2}\hbar\Omega$：量子諧振子永遠不可能完全靜止。這種零點能（zero-point energy）真實存在，而且可以量度；正因如此，液態氦在常壓下即使到了絕對零度也不會凝固。

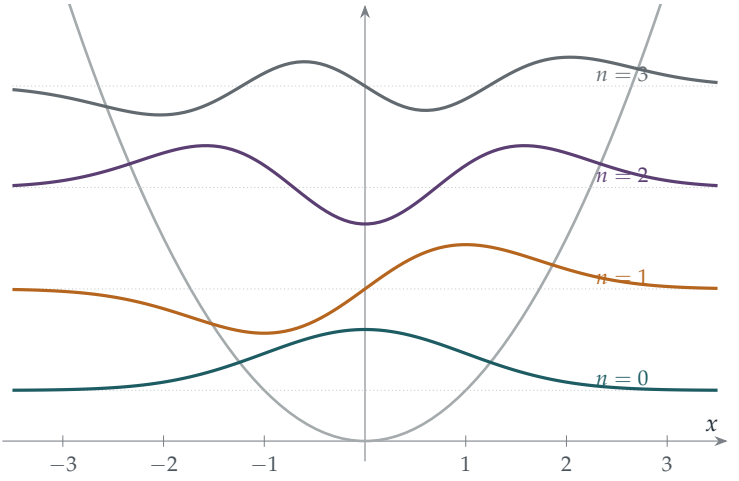


圖 8.6.1　量子諧振子（取 $\omega = 1$）。拋物線是勢能。每個波函數 y_n 都畫在其能級 $n + \frac{1}{2}$ 的高度上；各能級等距排列，而最低的能級高於零。函數 y_n 穿過零 n 次，而且其中任意兩個都正交。

埃爾米特函數與施圖姆－劉維爾理論

通解為 $y_n = H_n(\sqrt{\omega}x)e^{-\omega x^2/2}$，其中 H_n 是一個 n 次多項式，稱為埃爾米特多項式（Hermite polynomial）（圖 8.6.1）。正交性並非這條方程獨有。只要一條方程可以寫成 $(py')' + qy = -\lambda wy$ 並配上適當的邊界條件，屬於不同 λ 值的解便互相正交，所用的論證與第 4 階段完全相同。這就是施圖姆－劉維爾理論（Sturm–Liouville theory，1830 年代），它涵蓋了傅立葉級數中的正弦函數（Section 3.7）、鼓面的振動、靜電學中的勒讓德多項式（Legendre polynomials），以及氫原子的波函數。Section 4.7 的特徵值問題（eigenvalue problem）是另一個例子：在那裏求得的臨界值就是特徵值 λ，而那些正交函數就是系統固有的「模態」（modes）。

思維技巧

由恆等式得出正交性

要證明相關方程的兩個解正交，可找出一個恰當導數（*exact derivative*），使它的積分因邊界條件而爲零，而它本身又等於兩個解的乘積的某個倍數。

這比直接計算積分有力得多：它一次過適用於每一對解，即使你無法把解寫出來也行。

延伸練習

1 **下一個態。**證明 $y_2 = (4\lambda x^2 - 1)e^{-\lambda x^2}$（其中 $\lambda = \frac{\omega}{2}$）滿足 $n = 2$ 時的方程。利用 $\int_{-\infty}^{\infty} e^{-\omega x^2}\,\mathrm{d}x = \sqrt{\pi/\omega}$ 及 $\int_{-\infty}^{\infty} x^2 e^{-\omega x^2}\,\mathrm{d}x = \frac{1}{2\omega}\sqrt{\pi/\omega}$，直接驗證 $\int_{-\infty}^{\infty} y_0 y_2\,\mathrm{d}x = 0$。

2 **沒有其他能量。**把 n 換成一般的數 ν，並嘗試 $y = e^{-\lambda x^2}$。證明只有當 $\nu = 0$ 時它才是解。（事實上，只有當 ν 是非負整數時，才存在衰減的解。）

3 **正弦波。**由 $y'' = -n^2 y$ 及 $y(0) = y(\pi) = 0$ 出發，用同樣的方法證明當 $m \neq n$ 時 $\int_0^{\pi} \sin mx \sin nx\,\mathrm{d}x = 0$。

4 **階梯。**證明若 y_n 是方程的解，則 $y_{n+1} = \omega x y_n - y_n'$ 是把 n 換成 $n+1$ 後的方程的解。利用這個結果，由 y_0 生成 y_1 和 y_2。（這就是量子力學中的「升算符」（raising operator）。）

延伸閱讀

- **David Griffiths, *Introduction to Quantum Mechanics***（Cambridge University Press）。大學本科物理的標準教科書；第 2 章討論諧振子。
- **費曼（Richard Feynman）, *QED: The Strange Theory of Light and Matter***（Princeton University Press）。由量子力學的一位大師撰寫的科普讀物。
- **Mary Boas, *Mathematical Methods in the Physical Sciences***（Wiley）。為數學和物理學生講解埃爾米特多項式和施圖姆－劉維爾理論。

8.7 可以解出的混沌

題目

STEP I 2014，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/14-s1-q6`

引子

規則 $u_{n+1} = ku_n(1-u_n)$ 稱為邏輯斯諦映射（logistic map），是描述資源有限的種羣的最簡單模型：u_n 是第 n 年的種羣數量佔環境所能承載的最大數量的比例，k 是繁殖率，而因子 $1-u_n$ 則代表過度擁擠的影響。1976 年，生物學家羅伯特·梅（Robert May）在《自然》（*Nature*）上發表了一篇短文，題為「Simple mathematical models with very complicated dynamics」（動力學行為非常複雜的簡單數學模型）。他指出，當 k 逐漸增大至 4 時，這條只有一行的規則會由趨於穩定，變為來回振盪，再變為看似完全隨機。它成為了混沌（chaos）理論的奠基例子之一。

混沌系統通常無法求解：u_n 沒有公式，要求出 u_{100}，唯一的方法就是把它之前的各項逐一算出來。本題找到的正是那個例外。在 $k = 4$ 這個最混沌的取值上，公式確實存在，而且它恰好解釋了這個數列為何無法預測。

動筆之前

想一想 你以前在哪裏見過 $4s(1-s)$ 這個式子，其中 $s = \sin^2\theta$？哪一條恆等式把 $\sin^2\theta$、$\cos^2\theta$ 和 $\sin 2\theta$ 聯繫起來？

若 $s = \sin^2\theta$，則 $1-s = \cos^2\theta$，所以 $4s(1-s) = 4\sin^2\theta\cos^2\theta = (2\sin\theta\cos\theta)^2 = \sin^2 2\theta$。映射 $s \mapsto 4s(1-s)$ 其實就是喬裝了的倍角公式（double-angle formula）。

逐步解題

階段 1　猜出規律

設 $u_0 = \sin^2\theta$，由上面的恆等式可得

$$u_1 = 4\sin^2\theta\cos^2\theta = \sin^2 2\theta, \qquad u_2 = 4\sin^2 2\theta\cos^2 2\theta = \sin^2 4\theta.$$

想一想　每一步之後，角度有甚麼變化？你對 u_n 有甚麼猜想？

角度每次都加倍，所以我們猜想（conjecture）$u_n = \sin^2(2^n\theta)$。

階段 2　證明猜想

想一想　一個數列由「從一項到下一項」的規則定義，而我們要證明的命題涉及數列的每一項。哪一種證明方法最合適？

歸納法（induction）。命題在 $n = 0$ 時成立，因為 $u_0 = \sin^2\theta = \sin^2(2^0\theta)$。假設對某個 n 有 $u_n = \sin^2(2^n\theta)$，則

$$u_{n+1} = 4\sin^2(2^n\theta)(1 - \sin^2(2^n\theta)) = (2\sin(2^n\theta)\cos(2^n\theta))^2 = \sin^2(2^{n+1}\theta),$$

這正是 $n+1$ 時的命題。所以對所有 $n \geq 0$，都有 $u_n = \sin^2(2^n\theta)$。 ■

留意 $[0,1]$ 內的每一個初始值 u 都可以寫成 $\sin^2\theta$（取 $\theta = \arcsin\sqrt{u}$），所以對於 $[0,1]$ 內的任何初始值，這個公式都解出了 $k = 4$ 的邏輯斯諦映射。

階段 3　變數代換

想一想　把 $v_n = \alpha u_n + \beta$ 代入 v_{n+1} 的規則。你需要結果是 $u_{n+1} = 4u_n - 4u_n^2$。這會給出多少個條件？你又有多少個未知數？

代入後得

$$\alpha u_{n+1} + \beta = -p(\alpha u_n + \beta)^2 + q(\alpha u_n + \beta) + r,$$

所以

$$u_{n+1} = -p\alpha\, u_n^2 + (q - 2p\beta)\, u_n + \frac{-p\beta^2 + q\beta + r - \beta}{\alpha}.$$

要使它等於 $4u_n - 4u_n^2$，需要滿足三個條件：

$$p\alpha = 4, \qquad q - 2p\beta = 4, \qquad -p\beta^2 + (q-1)\beta + r = 0.$$

頭兩個條件確定了 $\alpha = 4/p$ 和 $\beta = (q-4)/(2p)$。這時已經沒有未知數剩下，所以第三條方程是對 p、q 和 r 的一個真正的條件。代入 β 並乘以 $4p$，得

$$-(q-4)^2 + 2(q-1)(q-4) + 4pr = 0,$$

展開後得 $-(q^2-8q+16)+(2q^2-10q+8)+4pr=0$，即 $q^2-2q-8+4pr=0$，亦即

$$4pr = 8 + 2q - q^2. \qquad \blacksquare$$

> **停一停，試一試** 三條方程，卻只有兩個未知數：這就是出現條件的原因。你能否在未計算之前，就看出為甚麼一般的二次映射不應該總是與 $(*)$ 等價？（想想線性的變數代換能改變甚麼，又不能改變甚麼。）

階段 4 例子

對於 $v_{n+1} = -v_n^2 + 2v_n + 2$，我們有 $p=1$、$q=2$、$r=2$，而 $4pr = 8 = 8+4-4$，所以條件成立。於是 $\alpha = 4$ 及 $\beta = -1$，所以 $v_n = 4u_n - 1$，其中 u_n 滿足 $(*)$。由於 $v_0 = 1$，故 $u_0 = \frac{1}{2} = \sin^2\frac{\pi}{4}$，所以

$$u_n = \sin^2\!\left(\frac{2^n\pi}{4}\right), \qquad v_n = 4\sin^2\!\left(\frac{2^n\pi}{4}\right) - 1 = 1 - 2\cos\!\left(\frac{2^n\pi}{2}\right).$$

由此得 $v_0 = 1$、$v_1 = 3$，而對所有 $n \geq 2$ 都有 $v_n = -1$。（直接驗算：$-1+2+2=3$，然後 $-9+6+2=-1$，再然後 $-1-2+2=-1$。） $\blacksquare$

這個初始值是特意選定的，好讓角度落在 π 的倍數上，數列從此不再變化。這樣一來，有趣的部分就被掩蓋了，而那正是接下來要談的。

回顧

- **識破喬裝。**一旦寫成 $s = \sin^2\theta$，映射 $s \mapsto 4s(1-s)$ 就變成了角度加倍。
- **先猜想，再用歸納法證明。**頭兩項提示了規律，歸納法則證明了它。
- **數一數方程和未知數。**線性代換有兩個自由參數，二次式有三個係數，所以必定剩下一個條件。

背後的數學

數列為何無法預測

公式 $u_n = \sin^2(2^n\theta)$ 看來已是故事的結局：任何一項都可以直接算出。但請細看它說了甚麼。由於 $\sin^2$ 的週期（period）是 π，所以只有 $2^n\theta$ 模（modulo）π 的值才有影響。

設 $\theta = \pi x$，其中 $0 \le x < 1$，並把 x 寫成二進制（binary）：

$$x = 0.b_1 b_2 b_3 \ldots \quad (\text{底為} 2).$$

乘以 2 會把二進制小數點向右移一位，而把結果取模 1，就是丟掉越過小數點的那個數位。所以 u_n 只取決於 x 由 b_{n+1} 起的二進制數位。邏輯斯諦映射每走一步，就多讀取初始值的一個數位，同時丟掉最前面的一個。

這就是*對初始條件的敏感依賴*（sensitive dependence on initial conditions），也是混沌的決定性特徵。兩個初始值若在首 20 個二進制位上相同（約六個小數位），它們產生的數列一般會在大約 20 步之內吻合，其後便毫無關係（圖 8.7.1）。要預測 u_{100}，你需要知道 u_0 準確至 100 個二進制位。初始角度的任何誤差，無論多小，每一步都會加倍。（種羣數量 u 本身的誤差並非每次都以完全相同的倍數增長，因為這個映射對 $[0,1]$ 某些部分的拉伸比其他部分大，但平均而言，它的增長率是一樣的。）公式幫不上忙：它只是令這個機制顯現出來。

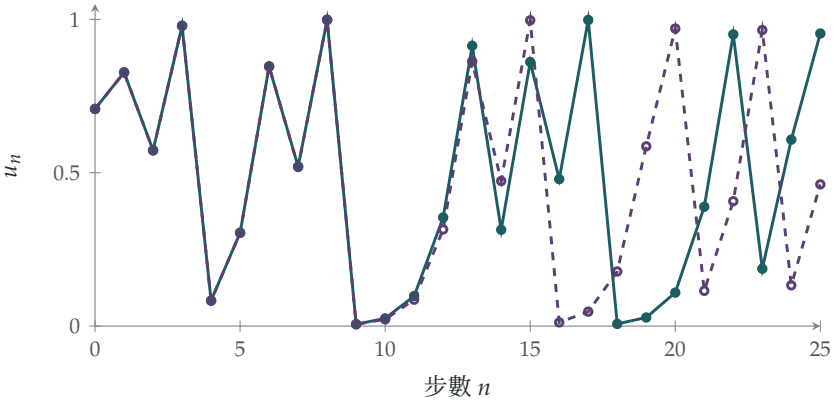


圖 8.7.1　$k = 4$ 的邏輯斯諦映射的兩個數列，分別由 $\sin^2(1)$（藍綠色）和 $\sin^2(1.00001)$（紫紅色虛線）開始。在大約頭十步內兩者無法分辨，到第 12 步已有明顯差別，其後便毫無關係：到第 12 步，角度之差 10^{-5} 已經加倍了十二次，變成約 $2^{12} \times 10^{-5} \approx 0.04$。

週期點無處不在

這個公式還能找出映射的每一個循環（cycle）。數列在 N 步之後回到初始值，當且僅當 $\sin^2(2^N\theta) = \sin^2\theta$，即 $2^N\theta \equiv \pm\theta \pmod{\pi}$。所以 $\theta = m\pi/(2^N - 1)$ 或 $\theta = m\pi/(2^N + 1)$ 給出一個週期整除 N 的點。例如 $\theta = \pi/5$ 滿足 $4\theta \equiv -\theta$，所以 $\sin^2\frac{\pi}{5} \to \sin^2\frac{2\pi}{5} \to \sin^2\frac{4\pi}{5} = \sin^2\frac{\pi}{5}$ 是一個 2-循環。每一個週期都有對應的週期點（periodic point），而且它們在 $[0,1]$ 中稠密（dense）：在每一個區間內，無論多小，都有一個初始值使數列重複出現。然而每一個週期點都是不穩定的，因為任何誤差都會加倍。稠密的週期點，卻把附近的一切都推開，這是混沌的第二個標誌。

通往混沌之路

對於較小的 k，並沒有公式可用，而數列的行為會隨 k 劇烈變化（圖 8.7.2）。當 $1 < k < 3$ 時，每一個由 $(0,1)$ 內出發的數列都會趨於不動點（fixed point）$1 - 1/k$。在 $k = 3$ 時，不動點變得不穩定，由一個 2-循環取而代之（這是 STEP II 2002 Q5 的主題；見延伸練習）。在 $k = 1 + \sqrt{6} \approx 3.449$ 時，2-循環讓位予 4-循環，然後是 8-循環，週期加倍來得愈來愈快，直至 $k \approx 3.5699$ 時，循環的長度變成無限，混沌由此開始。在混沌區域之內還有一些有序的窗口：最明顯的一個在 $k = 3.83$ 附近，其中有一個 3-循環，那是 Section 8.8 的主題。

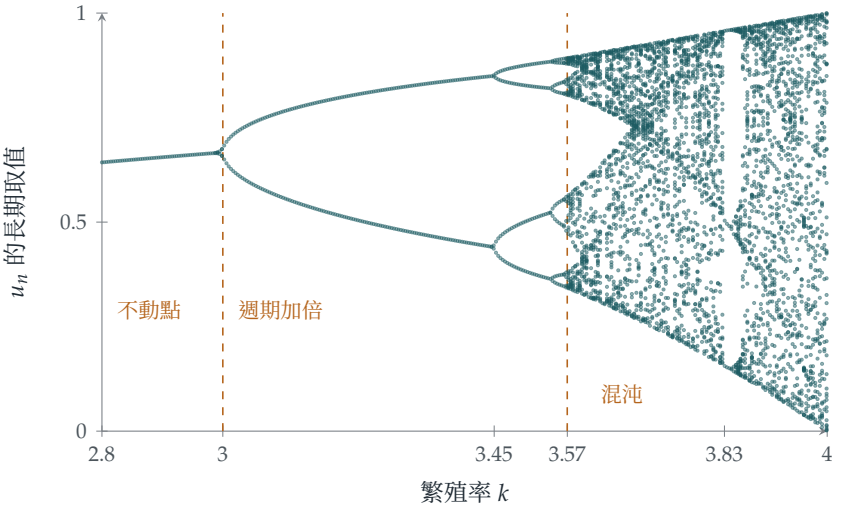


圖 8.7.2 邏輯斯諦映射的分岔圖（bifurcation diagram）。對每一個 k，由 $u_0 = 0.31416$ 開始把映射迭代 300 次，然後畫出其後的 40 個值。單一個不動點在 $k = 3$ 時分裂成一個 2-循環，然後是 4、8、...，直至在 $k = 3.57$ 附近進入混沌。$k = 3.83$ 附近那道淺色的垂直空隙，就是含有穩定 3-循環的窗口。在 $k = 4$ 時，這些值填滿了整個 $[0,1]$。

費根鮑姆常數

1975 年，費根鮑姆（Mitchell Feigenbaum）在洛斯阿拉莫斯（Los Alamos）工作時，用一部可編程的袖珍計算機，量度了週期加倍時的各個 k 值：$k_1 = 3$、$k_2 \approx 3.4495$、$k_3 \approx 3.5441$、$k_4 \approx 3.5644$、...。他留意到這些間距以一個固定的倍數縮小，

$$\frac{k_n - k_{n-1}}{k_{n+1} - k_n} \to 4.6692\ldots,$$

然後更令人驚訝地發現，同一個數 4.6692... 也出現在 $x \mapsto k\sin(\pi x)$ 之中，以及其他所有只有單一個峯、而峯頂像拋物線般圓滑（即二次極大值，quadratic maximum）的

光滑映射之中。這個常數對這一整類映射是普適（universal）的；至於峯頂較平坦的映射，例如形狀像 $-x^4$ 的那種，則各有自己的普適常數。其後，人們在滴水的水龍頭、受熱的流體和電子電路的實驗中都量度到這個常數。要解釋箇中原因，需要借用相變（phase transition）物理學的想法，而嚴格的證明要到 1982 年才由 Oscar Lanford 藉助電腦完成。

同一招，用三次

第 (ii) 部分的代換是共軛（conjugacy）的一個例子：若一個變數代換能把一個映射變成另一個，這兩個映射便稱為共軛，而它們的動力學行為完全相同。本題證明了每一個滿足該條件的二次映射，都與 $k = 4$ 的邏輯斯諦映射共軛，而後者又透過 $u = \sin^2\theta$，成為喬裝了的角度加倍。（嚴格來說，最後這一環是半共軛（semiconjugacy）：θ 和 $\pi - \theta$ 給出相同的 u，所以這個變數代換並非一對一（one-to-one）。邏輯斯諦映射真正共軛的對象是帳篷映射（tent map），即當 $t \le \frac{1}{2}$ 時 $t \mapsto 2t$、當 $t \ge \frac{1}{2}$ 時 $t \mapsto 2 - 2t$，所用的代換是 $u = \sin^2\frac{\pi t}{2}$，它在 $[0,1]$ 上是一對一的。）在 Section 2.8 中，你遇過 $x \mapsto x^2 - 2$，它由 $x = 2\cos\theta$ 解出，角度每一步加倍；它同樣滿足該條件（不妨試試）。所以這些全都是同一個映射的不同喬裝。在 Section 8.8 中，這個家族的第三位成員將會登場：一個由直線構成的映射，即帳篷映射。

思維技巧

變數代換，看清結構

一個好的代換可以把複雜的規則變成簡單的規則。設法找出一個變數代換，把問題變成你已經明白的問題。

在這裏，$u = \sin^2\theta$ 把一個二次映射變成了角度加倍，而一個線性代換則把一整族二次式歸結為一個。在大學裏，共軛是一件核心工具：在動力系統(dynamical systems)中，在線性代數中(把矩陣對角化，就是選取一組令它變得簡單的坐標)，以及在羣論中，都是如此。

延伸練習

1 **邏輯斯諦映射的循環。**試做 STEP II 2002，第 5 題（`step.maths.org/questions/02-s2-q5`）。然後證明只有當 $k > 3$ 時才存在 2-循環。

2 **平方減二。**證明 $v_{n+1} = v_n^2 - 2$ 滿足第 (ii) 部分的條件（把它寫成 $-pv_n^2 + qv_n + r$，其中 $p = -1$）。求 α 和 β，並驗證你的答案與 $v_n = 2\cos(2^n\phi)$ 一致。

3 **一個 3-循環。**利用公式 $u_n = \sin^2(2^n\theta)$，求 $u_{n+1} = 4u_n(1-u_n)$ 的一個 3-循環。（提示：試試 $\theta = \pi/7$，或 $\theta = \pi/9$。）週期恰好為 3 的點有多少個？

4 **丟失數位。**電腦儲存數字時，大約準確至 53 個二進制位。如果你在電腦上迭代 $u_{n+1} = 4u_n(1-u_n)$，大約多少步之後，計算出來的值便與真正的數列毫無關係？到那時，計算出來的數列是否就沒有用處？

延伸閱讀

- **Robert May，「Simple mathematical models with very complicated dynamics」**，*Nature* 261 (1976)。經典之作，至今仍然十分易讀。
- **James Gleick，*Chaos: Making a New Science***（Penguin）。講述梅、費根鮑姆以及混沌被發現的故事。
- **Steven Strogatz，*Nonlinear Dynamics and Chaos***（CRC Press），第 10 章。一維映射、邏輯斯諦映射和費根鮑姆常數，程度相當於大學一年級。
- **Ian Stewart，*Does God Play Dice?***（Penguin）。一本介紹混沌的科普讀物，其中有一章對邏輯斯諦映射講解得很好。

8.8 週期三蘊含混沌

題目

STEP III 2025，第 2 題

繼續閱讀之前，請先看題目（原題為英文）。題目載於 OCR 提供下載的 2025 年 STEP 試卷（zip 檔案）內的 STEP III 試卷：

`www.ocr.org.uk/Images/750486-2025-past-papers.zip`

引子

1975 年，李天岩（Tien-Yien Li）與約克（James Yorke）在《美國數學月刊》（*American Mathematical Monthly*）發表了一篇論文，題目是數學史上最著名的標題之一：「Period three implies chaos」（週期三蘊涵混沌）。他們證明了：如果一個由區間映到自身的連續函數（continuous function）有一個週期（period）為 3 的點（即一個恰好經過三步後返回原處的起始值），那麼它就有任何週期的點：1、2、4、5、17、1000，無一例外。這篇論文亦把「混沌」（chaos）一詞帶進了數學。

這道 2025 年的題目，正是讓你親手走一遍這個定理。函數 $\mathrm{f}(x) = 7 - 2|x|$ 由兩條直線組成，所以一切都可以精確計算。你會找出它的 2-循環（cycle）和 3-循環，並在過程中看到為甚麼其中一種循環存在，就迫使其餘的循環也必然存在。

動筆之前

想一想 $y = 7 - 2|x|$ 的圖像是甚麼樣子的？它在哪裏與直線 $y = x$ 相交？它把區間 $[-7, 7]$ 送到哪裏？

圖像是一個倒轉的 V 形，頂點在 $(0, 7)$，斜率（gradient）為 ± 2。它與 $y = x$ 相交於 $x = 7 - 2x$（其中 $x \geq 0$）之處，得 $x = \frac{7}{3}$；以及 $x = 7 + 2x$（其中 $x < 0$）之處，得 $x = -7$。這些就是*不動點*（fixed point）：從這裏出發，數列永遠不會移動。此外 $\mathrm{f}(-7) = \mathrm{f}(7) = -7$ 而 $\mathrm{f}(0) = 7$，所以 f 把 $[-7, 7]$ 映滿自身：先把它拉長至兩倍，再對摺起來。

逐步解題

階段 1　$\mathrm{f}\circ\mathrm{f}$ 的圖像

想一想　當 x 由 -7 走到 7 時，$\mathrm{f}(x)$ 由 -7 升至 7，再降回 -7。那麼 $\mathrm{f}(\mathrm{f}(x))$ 會怎樣變化？

當 x 由 -7 走到 0 時，$\mathrm{f}(x)$ 走遍整個 $[-7,7]$，所以 $\mathrm{f}(\mathrm{f}(x))$ 描出一個完整的 V 形：由 -7 升至 7，然後再降下來。當 x 由 0 走到 7 時，情況相同。因此在 $[-7,7]$ 上，$\mathrm{f}\circ\mathrm{f}$ 的圖像是一個倒轉的 W 形：四段斜率為 ± 4 的直線段，每段都在高度 -7 與 7 之間來回，頂點在 $x=\pm\frac{7}{2}$（該處 $\mathrm{f}(x)=0$），並在 $x=0$ 處下降至 -7（圖 8.8.1）。在 $[-7,7]$ 之外，兩個圖像都急速下降。

階段 2　解 $\mathrm{f}(\mathrm{f}(x))=x$

想一想　四段直線段每一段都與對角線 $y=x$ 恰好相交一次。其中兩個交點你早已知道，為甚麼？怎樣才能不必展開多重絕對值而找出另外兩個？

不動點 $\frac{7}{3}$ 和 -7 當然滿足 $\mathrm{f}(\mathrm{f}(x))=x$。任何其他的解 p 都屬於一個 2-循環 $p\to q\to p$，其中 $q=\mathrm{f}(p)\neq p$。試試各種正負號的組合。如果 p 和 q 都是正數，則 $q=7-2p$ 及 $p=7-2q$，這迫使 $p=q=\frac{7}{3}$；如果兩者都是負數，同樣的論證給出 -7。所以必定一正一負，設 $p>0>q$：

$$q=7-2p,\qquad p=7+2q.$$

於是 $p=7+14-4p$，所以 $p=\frac{21}{5}$ 及 $q=-\frac{7}{5}$，而這確實滿足 $q<0$。

在 $[-7,7]$ 之外並沒有解：當 $x>7$ 時，有 $\mathrm{f}(\mathrm{f}(x))=21-4x<x$；當 $x<-7$ 時，有 $\mathrm{f}(\mathrm{f}(x))=21+4x<x$。所以四個解是

$$x=-7,\quad -\tfrac{7}{5},\quad \tfrac{7}{3},\quad \tfrac{21}{5}.\qquad\blacksquare$$

數列的週期為 2，當且僅當 $a=\frac{21}{5}$ 或 $a=-\frac{7}{5}$。（在不動點處，數列是常數數列，週期為 1。）

階段 3　遲了才進入循環

取 $a=\frac{28}{5}$：$u_1=7-\frac{56}{5}=-\frac{21}{5}$，$u_2=7-\frac{42}{5}=-\frac{7}{5}$，$u_3=7-\frac{14}{5}=\frac{21}{5}$，$u_4=-\frac{7}{5}$，如此類推。由 u_2 開始，數列就是那個 2-循環，但 $u_0=\frac{28}{5}$ 和 $u_1=-\frac{21}{5}$ 並不在循環之內。$\blacksquare$

想一想　你自己可以怎樣找到 $\frac{28}{5}$？試從後倒推：f 把哪些數送到 $-\frac{7}{5}$？

$f(x) = -\frac{7}{5}$ 即是 $|x| = \frac{21}{5}$，所以 $x = \pm\frac{21}{5}$。其中一個在循環之內；另一個，即 $-\frac{21}{5}$，就是一個入口。然後 $f(x) = -\frac{21}{5}$ 即是 $|x| = \frac{28}{5}$。由於 f 是二對一的，$(-7,7)$ 內每一點都有兩個原像（preimage），所以每個循環都有一整棵由起始值組成的樹，最終匯入這個循環。

階段 4 三循環

停一停，試一試 計算由 $a = 1$ 和 $a = -\frac{7}{9}$ 出發的數列。

由 $a = 1$ 出發：$1 \to 5 \to -3 \to 1$。由 $a = -\frac{7}{9}$ 出發：$-\frac{7}{9} \to \frac{49}{9} \to -\frac{35}{9} \to -\frac{7}{9}$。兩者都是 3-循環。

想一想 $f(f(f(x))) = x$ 可以有多少個解？像第 1 階段那樣，想想它的圖像的形狀。

重複第 1 階段的論證，每多施行一次 f，直線段的數目就增加一倍。所以在 $[-7,7]$ 上，$f \circ f \circ f$ 的圖像是由八段組成的鋸齒形，斜率為 ± 8，每段都走遍由 -7 到 7 的整個高度（圖 8.8.2）。每一段都與對角線恰好相交一次，所以在 $[-7,7]$ 內恰好有八個解，而（一如之前）區間之外則沒有解。我們已經找到八個：兩個不動點，以及兩個 3-循環的六個點。所以解是

$$x = -7,\ \tfrac{7}{3},\ 1,\ 5,\ -3,\ -\tfrac{7}{9},\ \tfrac{49}{9},\ -\tfrac{35}{9}. \qquad \blacksquare$$

階段 5 遲了才進入 3-循環

由循環 $1 \to 5 \to -3$ 從後倒推。我們想要 u_3 在循環之內，但 u_2 不在。1 的原像滿足 $|x| = 3$，所以 $x = \pm 3$；-3 在循環之內，故取 $u_2 = 3$。3 的原像滿足 $|x| = 2$；取 $u_1 = 2$。2 的原像滿足 $|x| = \frac{5}{2}$；取 $u_0 = \frac{5}{2}$。於是

$$\tfrac{5}{2} \to 2 \to 3 \to 1 \to 5 \to -3 \to 1 \to \cdots,$$

而 $\frac{5}{2}$、2、3 全都不在循環之內。所以 $a = \frac{5}{2}$ 符合要求。 $\blacksquare$

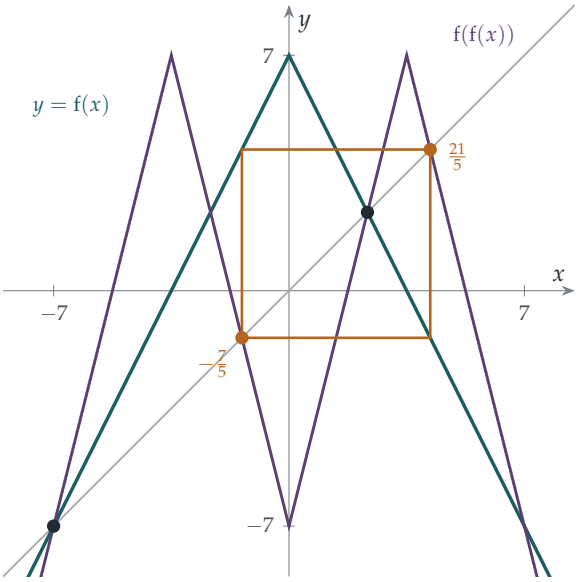


圖 8.8.1　$y = \mathrm{f}(x)$（藍綠色）、$y = \mathrm{f}(\mathrm{f}(x))$（紫紅色）和 $y = x$（灰色）的圖像。紫紅色的圖像與對角線相交四次：在不動點 -7 和 $\frac{7}{3}$（黑色），以及在 2-循環 $\frac{21}{5} \leftrightarrow -\frac{7}{5}$（琥珀色）。琥珀色的正方形是這個 2-循環的蛛網圖（cobweb）：向上走到 f 的圖像，橫向走到對角線，然後重複。

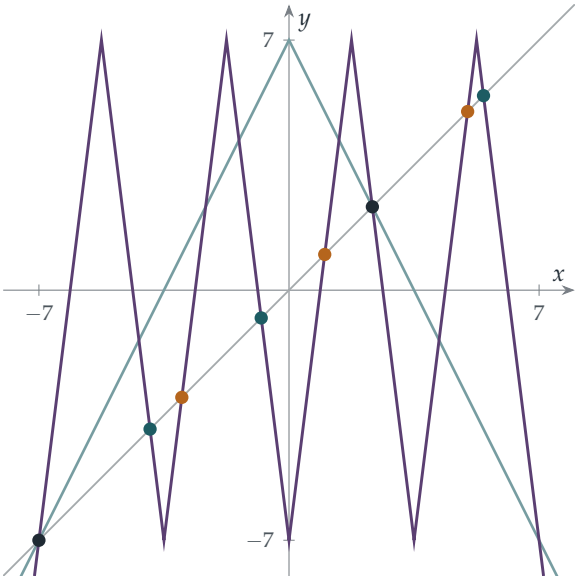


圖 8.8.2　$y = \mathrm{f}(\mathrm{f}(\mathrm{f}(x)))$ 的圖像（紫紅色）在 $[-7, 7]$ 上是由八段組成的鋸齒形，斜率為 ± 8，每段與對角線相交一次。八個交點分別是兩個不動點（黑色）、3-循環 $1 \to 5 \to -3$（琥珀色）和 3-循環 $-\frac{7}{9} \to \frac{49}{9} \to -\frac{35}{9}$（藍綠色）。

回顧

- **憑形狀點算。**我們沒有展開多重絕對值，而是點算圖像中直線段的數目。這告訴我們要找多少個解，也就知道甚麼時候已經全部找齊。
- **利用循環的結構。**$f(f(x)) = x$ 的解若非不動點，就是成對出現；$f(f(f(x))) = x$ 的解若非不動點，就是三個一組出現。
- **從後倒推。**第 (i)(d) 和 (ii)(c) 部分中那些「遲來者」，都是靠原像構造出來的。

背後的數學

任何長度的循環

第 4 階段的點算方法適用於任何步數。把 f^n 記作施行 f 共 n 次。在 $[-7,7]$ 上，f^n 的圖像有 2^n 段直線段，每段都走遍整個高度，所以 $f^n(x) = x$ 恰好有 2^n 個解。其中有些解的週期較短，是某個整除 n 的 d，但對每個 d 而言，這樣的解最多只有 2^d 個，而且

$$2^n > 2^{n/2} + 2^{n/3} + \cdots + 2^1 \quad \text{（對所有滿足 } d < n \text{ 的 } n \text{ 的因數求和）}$$

因為所有小於 n 的因數都不超過 $n/2$，而 $2^1 + 2^2 + \cdots + 2^{\lfloor n/2 \rfloor} < 2^{\lfloor n/2 \rfloor + 1} \leq 2^n$。所以對每個 n，都存在週期恰好為 n 的點。這個函數有任何長度的循環。

為甚麼週期三迫使一切出現

李天岩與約克的定理指出，這並非我們這個函數獨有的性質：任何由區間映到自身的連續函數 g，只要有一個 3-循環，就有任何長度的循環。證明所用的不外乎介值定理（intermediate value theorem），而我們的 3-循環正好展示了證明的思路。

> **想一想** 設 $I_1 = [-3,1]$ 及 $I_2 = [1,5]$，即循環 $1 \to 5 \to -3$ 各點之間的兩個區間。f 把它們分別送到哪裏？

在 I_2 上，f 由 $f(1) = 5$ 遞減至 $f(5) = -3$，所以 $f(I_2) = [-3,5] = I_1 \cup I_2$。在 I_1 上，f 由 $f(-3) = 1$ 升至 $f(0) = 7$，再降至 $f(1) = 5$，所以 $f(I_1) = [1,7]$，它包含 I_2。留意我們只用到循環的三個點：任何具有 3-循環 $a \to b \to c$、$a < b < c$ 的連續函數，都以同樣的方式覆蓋它的區間。我們可以用一幅圖來概括哪個區間覆蓋哪個區間（圖 8.8.3）：

$$I_1 \to I_2, \qquad I_2 \to I_1, \qquad I_2 \to I_2.$$

定理於是由以下兩個事實推出，兩者都是介值定理的推論。

- 如果 $g(I) \supseteq I$，則 g 在 I 內有一個不動點。（對 $g(x) - x$ 應用介值定理。）

▶ 如果 $g(I) \supseteq J$，則存在一個較小的閉區間 $I' \subseteq I$，使得 $g(I') = J$ 恰好成立。

要找一個週期為 n 的點，可沿着圖中的迴路 $I_2 \to I_2 \to \cdots \to I_2 \to I_1 \to I_2$ 走：$n-1$ 步留在 I_2 之內，另有一步走出去到 I_1 再回來。反覆運用第二個事實，我們可以把 I_2 收窄成一個由起始點組成的區間，這些起始點的軌道（orbit）都依循這條路線，而 g^n 把這個區間映滿整個 I_2。第一個事實給出一點 x，滿足 $g^n(x) = x$，並且只在第 $n-1$ 步到訪 I_1。它的週期不可能小於 n，否則對 I_1 的到訪便會在更早的時候重複出現。（這裏有一個微妙之處，因為 I_1 和 I_2 共有一點 1；驗證這不會造成問題，留作習題。）

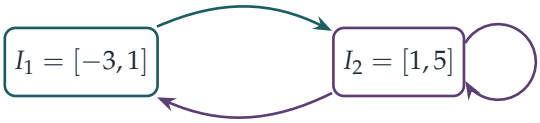


圖 8.8.3　3-循環的覆蓋圖（covering diagram）。由一個區間指向另一個區間的箭嘴，表示前者的像包含後者。圖中每一條迴路都給出函數的一個循環：迴路 $I_2 \to I_2$ 給出一個不動點，$I_1 \to I_2 \to I_1$ 給出一個 2-循環，而先繞 I_2 處的迴路 $n-1$ 次再到訪 I_1，則給出一個長度為 n 的循環。

沙可夫斯基定理

李天岩與約克並不是最先發現的人。早在 1964 年，烏克蘭數學家沙可夫斯基（Oleksandr Sharkovskii）已經證明了一個完備得多的結果，以俄文發表於《烏克蘭數學雜誌》（*Ukrainian Mathematical Journal*），其後十年間西方對此一無所知。把正整數排列成

$$3,\ 5,\ 7,\ 9,\ \ldots,\ \ 2\cdot 3,\ 2\cdot 5,\ \ldots,\ \ 2^2\cdot 3,\ 2^2\cdot 5,\ \ldots,\ \ \ldots,\ \ 2^3,\ 2^2,\ 2,\ 1:$$

先是由 3 開始的奇數，然後是奇數的兩倍，再然後是奇數的四倍，如此類推，最後以 2 的冪按遞減次序作結。沙可夫斯基定理指出：如果區間上的一個連續函數有一個長度為 m 的循環，那麼凡是在這個列表中排在 m 之後的長度，它都有相應的循環。由於 3 排在最前，週期 3 便蘊涵所有週期。而且這個定理是最佳的：對每個 m，都存在一個函數，它有長度為 m 的循環，卻沒有列表中排在更前的長度的循環。列表末端 2 的冪，正正就是 Section 8.7 中邏輯斯蒂映射（logistic map）的倍週期級聯（period-doubling cascade），而 3-循環在那裏出現於 $k = 1 + \sqrt{8}$。

帳篷映射與邏輯斯蒂映射

作代換 $x = 14y - 7$，它把 $[0,1]$ 映滿 $[-7,7]$。簡單計算可見，f 變成了帳篷映射（tent map）

$$T(y) = \begin{cases} 2y & 0 \le y \le \frac{1}{2}, \\ 2 - 2y & \frac{1}{2} \le y \le 1, \end{cases}$$

而再作代換 $u = \sin^2(\pi y/2)$，就把帳篷映射變成邏輯斯蒂映射 $u \mapsto 4u(1-u)$。所以這道題目中的函數，其實就是 Section 8.7 中那個混沌的邏輯斯蒂映射，只是被拉直了。它的 $f^3(x) = x$ 的八個解，對應於 $\sin^2(8\theta) = \sin^2\theta$ 在 $[0, \frac{\pi}{2}]$ 內的八個解。

思維技巧

以介值定理證明存在性

要證明解存在，不一定要把它找出來。只需證明一個連續函數變號，或者一個區間被映射後蓋過自身。

在這道題目中，每一個解都可以算出來。但李天岩與約克的定理適用於我們根本無從計算的函數，而且它只用介值定理和一幅顯示哪個區間覆蓋哪個區間的圖，就證明了無窮多個循環的存在。在大學裏，這種論證方式，即不作構造而證明存在，會出現在分析（analysis）（介值定理本身）、拓撲學（topology）（不動點定理）以及動力系統（dynamical systems）的各個角落。

延伸練習

1 **一個 4-循環。**利用覆蓋圖預測 f 的一個 4-循環應該在哪裏（三步在 I_2，一步在 I_1）。然後把它精確地求出來。（提示：這個循環有三個正的點和一個負的點；寫下四條線性方程。）

2 **微妙之處。**在「週期 3 蘊涵週期 n」的證明中，解釋為甚麼當 $n \geq 4$ 時，所找到的點 x 不可能是點 1（或原來那個 3-循環的任何一點）。

3 **有週期 5 而沒有週期 3。**草繪一個在 $[1,5]$ 上的連續函數的圖像，它滿足 $g(1) = 3$、$g(3) = 4$、$g(4) = 2$、$g(2) = 5$、$g(5) = 1$，並以直線連接這些點。就區間 $[1,2]$、$[2,3]$、$[3,4]$、$[4,5]$ 畫出它的覆蓋圖。你能找到一條長度為 3 的迴路嗎？（沙可夫斯基定理說它沒有 3-循環。）

4 **帳篷映射。**驗證 $x = 14y - 7$ 把 f 變成帳篷映射。然後證明帳篷映射把以二進制寫成的 y 加倍，並且每當首位數字是 1 時就把所有數字反轉；再利用這一點找出它的 3-循環。

延伸閱讀

► **Tien-Yien Li & James Yorke, "Period three implies chaos"**, *American Mathematical*

Monthly 82 (1975)。篇幅短小，懂得介值定理便可讀懂。

- **Robert Devaney, *A First Course in Chaotic Dynamical Systems*** (CRC Press)。為大學本科生而寫；涵蓋帳篷映射、邏輯斯蒂映射和沙可夫斯基定理，並附完整證明。
- **Keith Burns & Boris Hasselblatt, "The Sharkovsky theorem: a natural direct proof"**, *American Mathematical Monthly* 118 (2011)。一個運用覆蓋圖的現代證明。

我們學到了甚麼

本章開首我們說過，大多數微分方程（differential equation）都無法求解，而真正的功夫在於即使解不出來，也能理解它們。我們提出了三個問題：現在是否決定未來？不求解的話，我們能知道些甚麼？長遠而言會發生甚麼事？以下是本章八道題目所揭示的答案。

1. 現在是否決定未來？

通常是，但並非必然。在 Section 8.1 中，方程 $(y')^2 = 4y$ 經過 x 軸上每一點都有無窮多個解：一個解可以隨意在 $y = 0$ 停留任意長的時間，然後才離開。罪魁禍首是平方根：它在零處的斜率為無窮大，破壞了皮卡–林德勒夫定理（Picard–Lindelöf theorem）中的利普希茨條件（Lipschitz condition）。同一機制也解釋了為何一個漏水的水桶無法告訴你它是何時漏空的，而它亦是諾頓圓頂（Norton's dome）背後的原理——那是古典力學中一個關於決定論（determinism）的難題。

即使未來已被決定，它也未必可以預測。Section 8.7 的邏輯斯諦映射（logistic map）有一條精確的公式 $u_n = \sin^2(2^n\theta)$，然而每走一步，起始角 θ 的任何誤差都會加倍，而起始值的誤差平均而言也增長得同樣快。原則上的決定論，並不等於實際上的可預測性。

2. 不求解的話，我們能知道些甚麼？

能知道的非常多。

- **能量。**在 Section 8.2 中，把方程乘以 y' 便得出一個不可能增加的量，從而為一條無人能解的方程的解定出了界限。這類量稱為*李雅普諾夫函數*（Lyapunov function），是證明穩定性（stability）的主要工具。
- **平衡點與穩定性。**在 Section 8.3 中，增長率的正負號預示了魚羣的命運，並顯示捕撈量一旦超過某個臨界值，魚羣便會突然崩潰。
- **合適的坐標。**在 Section 8.4 中，極坐標（polar coordinates）把一個耦合方程組化為只含 r 的單一方程，顯示每個解都盤旋着趨向一個圓。在 Section 8.5 中，對稱性把四隻互相追逐的狗簡化為一隻，再以相除消去時間，便得出牠們路徑的形狀。
- **由恆等式看出結構。**在 Section 8.6 中，朗斯基行列式（Wronskian）證明了某方程的解互相正交（orthogonal），而無須對其中任何一個解進行積分。

- **點算與覆蓋。**在 Section 8.8 中，點算圖像的段數便找出了所有循環（cycle），而介值定理（intermediate value theorem）則證明了任意長度的循環都存在。

3. 長遠而言會發生甚麼事？

本章遇到了各種各樣的長期行為。解或會趨於穩定的平衡點（Section 8.3）、趨近一個不斷重複的循環（Section 8.4）、在有限時間內盤旋收於一點（Section 8.5），又或者表現出混沌（chaos）行為（Section 8.7）。就邏輯斯諦映射而言，單單一個參數便可令系統經歷以上各種情況：先是一個不動點（fixed point），然後是一連串由費根鮑姆（Feigenbaum）普適常數支配的倍週期（period doubling），最後是混沌，其間夾雜着一些有序的窗口，例如 3-循環。而 Section 8.8 則證明了，一旦出現 3-循環，任意長度的循環都必定存在——這是沙可夫斯基（Sharkovskii）整數排序的第一步。

4. 微分方程與動力學的工具箱

方法	你在哪裏用過
尋找奇解（singular solution）	非唯一性（8.1）
一個不可能增加的量	定界（8.2）
增長率的正負號	捕撈（8.3）
配合對稱性的坐標	極限環（limit cycle）（8.4）；追逐（8.5）
以代換法線性化	伯努利方程（Bernoulli equation）（8.4）
朗斯基行列式	正交性（8.6）
共軛（conjugacy）	邏輯斯諦映射（8.7）
歸納法（induction）	$u_n = \sin^2(2^n\theta)$（8.7）
點算圖像的段數	f^n 的循環（8.8）
介值定理	週期三（8.8）

下一步

- 能量與守恆定律會在力學中再次出現，由 Section 10.1 開始。
- 含多個未知函數的線性微分方程可用特徵向量（eigenvector）求解，它們支配着 Section 11.2 中的振動系統；至於無法以初等函數求解的非線性單擺，則見於 Section 11.3。
- 在 Section 12.15 中，反覆迭代一個映射以找出其不動點，可判定一個姓氏會否消亡；而隨機數列的長期行為則是 Section 12.16 的主題。

如果你想在上大學之前再進一步

- **Steven Strogatz, *Nonlinear Dynamics and Chaos***：動力系統（dynamical systems）最好的入門書，應用和圖解十分豐富。Strogatz 根據此書講授的課堂錄影可在網上免費觀看。
- **James Gleick, *Chaos: Making a New Science***：講述混沌是如何被發現的故事。
- **George Simmons, *Differential Equations with Applications and Historical Notes***：經典的入門課本，歷史材料豐富。
- **Robert Devaney, *A First Course in Chaotic Dynamical Systems***：迭代映射，附有證明，程度相當於大學一年級。

拉普拉斯（Laplace）曾經設想，只要準確知道現在，便可以算出整個未來。本章顯示了這個夢想在何處止步。唯一性可以失效；而即使唯一性成立，混沌也可以把最微小的不確定性不斷放大，直至無法預測。但本章同時顯示，沒有公式也能理解很多東西：界限、平衡點、循環及其穩定性，以至混沌本身的結構。由「解方程」轉為「理解方程的解」，是大學數學課程要求你作出的最重要的思維轉變之一。

延伸練習答案

以下只是簡短的答案和提示，並非完整解答：只列出結果和關鍵步驟，方便你核對自己的做法。請先認真做過每道練習，然後才翻看。

8.1 當解不唯一

1. 設 $y = v^2$：$2vv' = \alpha v - \beta v^2$，所以 $v' = \frac{1}{2}(\alpha - \beta v)$，而 $y_1 = \frac{\alpha^2}{\beta^2}\left(1 - \mathrm{e}^{-\beta t/2}\right)^2$。設 $y = v^3$：$v' = \frac{1}{3}(\alpha - \beta v)$，而 $y_1 = \frac{\alpha^3}{\beta^3}\left(1 - \mathrm{e}^{-\beta t/3}\right)^3$。常數函數 $y = 0$ 令等式兩邊都等於零；推導時除以 v，就把這個解丟掉了。

2. $2\sqrt{h} = 2\sqrt{h_0} - kt$，所以 $h = \left(\sqrt{h_0} - \frac{1}{2}kt\right)^2$，水桶在 $T = \frac{2\sqrt{h_0}}{k}$ 時流乾。$h \equiv 0$，以及當 $t \le T$ 時 $h = \frac{k^2}{4}(T - t)^2$、當 $t \ge T$ 時 $h = 0$ 的函數，兩者都滿足方程，而且由 T 起都等於零：看着一個空桶，你無法知道它是何時流乾的。

3. $\frac{\mathrm{d}}{\mathrm{d}x}(y\mathrm{e}^{-2x}) = (y' - 2y)\mathrm{e}^{-2x} = 0$，所以 $y = C\mathrm{e}^{2x}$，而 $y(0) = 0$ 迫使 $C = 0$。函數 $2y$ 滿足利普希茨（Lipschitz）條件（它的變化至多為 $2|y_1 - y_2|$），而 $2\sqrt{y}$ 在 0 處的斜率為無窮大。

4. 對任意 $b \le 0 \le a$（容許 $b = -\infty$ 或 $a = \infty$）：當 $x < b$ 時 $y = (x - b)^3$，當 $b \le x \le a$ 時 $y = 0$，當 $x > a$ 時 $y = (x - a)^3$。$a = b = 0$ 的情況就是 $y = x^3$。

8.2 不求解而估界

1. $\frac{\mathrm{d}}{\mathrm{d}x}\left((y')^2 + y^2\right) = 2y'(y'' + y) = 0$，所以 $(y')^2 + y^2 = A^2$，因而 $|y| \le A$。精確解是 $A\cos x$。

2. $\frac{\mathrm{d}}{\mathrm{d}t}\left((\theta')^2 - 2\cos\theta\right) = 2\theta'(\theta'' + \sin\theta) = 0$。由靜止開始，$(\theta')^2 = 2(\cos\theta - \cos\theta_0) \ge 0$，所以 $\cos\theta \ge \cos\theta_0$；當 $|\theta_0| < \pi$ 時，連續性（continuity）令 θ 留在 $[-|\theta_0|, |\theta_0|]$ 之內，而當 $|\theta_0| = \pi$ 時，擺倒立平衡，並一直停在那裏。若 $|\theta_0| > \pi$，例如 $\theta_0 = \frac{3\pi}{2}$，則 $\theta''(0) = -\sin\frac{3\pi}{2} = 1 > 0$，所以 θ 起初遞增：擺其實位於角度 $\theta_0 - 2\pi = -\frac{\pi}{2}$，並在由該處量起的 $\pm\frac{\pi}{2}$ 之間擺動。

3. $\frac{\mathrm{d}}{\mathrm{d}t}\left((y')^2 + y^2\right) = 2y'(y'' + y) = -4k(y')^2 \le 0$。它確實趨向零，但單憑單調性（monotonicity）並不足以證明這一點：你還需要其他工具，例如顯式解，其指數 $-k \pm \sqrt{k^2 - 1}$ 的實部為負。

4. $(y')^2 - \frac{1}{2}y^4$ 是常數，等於 $-\frac{1}{2}$。由此得 $(y')^2 = \frac{1}{2}(y^4 - 1)$，但這給不出任何界：「位

能」－$\frac{1}{2}y^4$ 沒有可以困住 y 的位阱，而事實上 y 會在有限時間內趨向無窮（blow up）。

8.3 收穫種羣

1. 方程是 $x' = \alpha^2 - (x-K)^2$，可化為 $y' = y(2\alpha - y)$。設 $w = \frac{1}{y}$，則 $w' = 1 - 2\alpha w$，所以 $w = \frac{1}{2\alpha} + Ce^{-2\alpha t}$。若 $y(0) < 0$，則 $w(0) < 0 < \frac{1}{2\alpha}$，所以 w 在有限時間內經過 0，而 $y \to -\infty$；x 在此之前已到達 0。

2. 設 $z = x - K$，則 $z' = -(z^2 + \beta^2)$，所以 $\tan^{-1}\frac{z}{\beta} = \tan^{-1}\frac{K}{\beta} - \beta t$。種羣數量在 $t = \frac{2}{\beta}\tan^{-1}\frac{K}{\beta}$ 時降至零（$z = -K$）。

3. $x' = x(2K - h - x)$ 的平衡點（equilibria）是 0 和 $2K - h$。當 $h < 2K$ 時，這是一條環境承載量（carrying capacity）為正的邏輯斯諦方程（logistic equation），每個正數解都趨向 $2K - h$。捕獲量隨種羣數量下降而減少，所以永遠不會把種羣數量逼至零。

4. $x = K + K\tanh Kt$，而 $\tanh Kt = \frac{1}{2}$ 給出 $t = \frac{1}{K}\tanh^{-1}\frac{1}{2} = \frac{\ln 3}{2K}$。

8.4 盤旋趨向圓

1. 曲線與半徑之間的夾角 ψ 滿足 $\tan\psi = \frac{r}{\mathrm{d}r/\mathrm{d}\theta} = -1$，在每一點都相同：曲線與每條射線都以 45° 相交。

2. $x\,\mathrm{d}x + y\,\mathrm{d}y = \frac{1}{2}\mathrm{d}(x^2+y^2) = r\,\mathrm{d}r$。另外 $x\,\mathrm{d}y - y\,\mathrm{d}x = r\cos\theta(\sin\theta\,\mathrm{d}r + r\cos\theta\,\mathrm{d}\theta) - r\sin\theta(\cos\theta\,\mathrm{d}r - r\sin\theta\,\mathrm{d}\theta) = r^2\,\mathrm{d}\theta$。

3. 設 $s = r^2$，則 $\dot{s} = 2s(1-s)$，是一條邏輯斯諦方程，所以對任意 $s(0) > 0$ 都有 $s \to 1$。寫成笛卡兒（Cartesian）形式：$\dot{x} = x(1 - x^2 - y^2) - y$，$\dot{y} = y(1 - x^2 - y^2) + x$。

4. $\dot{r} = r(r-1)(r-2)$，$\dot{\theta} = 1$。在 0 與 1 之間，$\dot{r} > 0$；在 1 與 2 之間，$\dot{r} < 0$；超過 2 之後，$\dot{r} > 0$。所以 $r = 1$ 是吸引的，而 $r = 2$ 是排斥的。

8.5 格力犬追逐格力犬

1. 由對稱性可知，幾隻狗始終構成一個正方形。A 沿 AB 以速率 v 奔跑，令距離以速率 v 縮短；B 沿 BC 奔跑，而這個方向垂直於 AB，所以在一階近似下不會改變距離。因此 AB 恰好以速率 v 縮短，由 s 減至 0，需時 $\frac{s}{v}$。

2. $r = \lambda e^{-\sqrt{3}\theta}$。距離以速率 $v - v\cos 120° = \frac{3}{2}v$ 縮短，所以牠們在時間 $\frac{2s}{3v}$ 後相遇，每隻各跑了 $\frac{2s}{3}$（這與 $\frac{s}{2\sin^2(\pi/n)}$ 在 $n = 3$ 時的值吻合）。

3. 每轉一圈，r 就乘以 $e^{-2\pi} \approx \frac{1}{535}$：轉一圈後約為 19 cm。轉兩圈後約為 0.35 mm，所以轉足兩圈便已足夠。

4. 兩隻狗筆直地向對方奔跑，在時間 $\frac{s}{2v}$ 後於中點相遇。公式在此失效：$\tan\frac{\pi}{2}$ 為無窮大，正好對應完全沒有轉彎這個事實。

8.6 量子振子

1. 這只需直接求導（雖然頗為繁瑣）；另一個做法是由第 4 題推出。取 $2\lambda=\omega$，則 $\int y_0y_2=\int(2\omega x^2-1)\mathrm{e}^{-\omega x^2}\,\mathrm{d}x=2\omega\cdot\frac{1}{2\omega}\sqrt{\frac{\pi}{\omega}}-\sqrt{\frac{\pi}{\omega}}=0$。

2. $y''=(4\lambda^2x^2-2\lambda)y$，所以方程要求 $4\lambda^2=\omega^2$ 及 $2\lambda=(2\nu+1)\omega$。取 $\lambda=\frac{\omega}{2}>0$（使解衰減），得 $\nu=0$。

3. $\frac{\mathrm{d}}{\mathrm{d}x}(y_my_n'-y_ny_m')=y_my_n''-y_ny_m''=(m^2-n^2)y_my_n$。由 0 積分至 π：因為 y_m 和 y_n 在兩端都等於零，左邊得 0，而 $m^2\neq n^2$。

4. 記 $c=(2n+1)\omega$，於是 $y_n''=(\omega^2x^2-c)y_n$。把 $z=\omega xy_n-y_n'$ 求導兩次，並運用上式（及其導數），得 $z''-\omega^2x^2z=-(c+2\omega)z$，而 $c+2\omega=(2n+3)\omega$。由 $y_0=\mathrm{e}^{-\omega x^2/2}$ 出發：先得 $2\omega x\mathrm{e}^{-\omega x^2/2}$，再得 $2\omega(2\omega x^2-1)\mathrm{e}^{-\omega x^2/2}$，它們分別是 y_1 和 y_2 的倍數。

8.7 可以解出的混沌

1. (i) 用歸納法（induction）：若 $0<x<1$，則 $0<x(1-x)\le\frac{1}{4}$，所以 $0<kx(1-x)<1$。(ii) $a=ka(1-a)$ 給出 $k=\frac{1}{1-a}$。(iii) 由於 $x_1=ka(1-a)=ab$，條件 $x_2=a$ 即 $kab(1-ab)=a$；代入 $k=\frac{b}{1-a}$ 後，化為 $ab^3-b^2+1-a=0$。根 $b=1$ 對應不動點（fixed point）（$x_1=a$），所以把它提出來：$ab^2+(a-1)b+(a-1)=0$；解出 b，便得 $k=\frac{b}{1-a}$。等價的說法是：2-循環（cycle）的兩點是 $k^2x^2-k(k+1)x+(k+1)=0$ 的兩個根，其判別式（discriminant）$k^2(k+1)(k-3)$ 只在 $k>3$ 時為正。

2. $p=-1$，$q=0$，$r=-2$：$4pr=8=8+2q-q^2$。比較係數得 $\alpha=\frac{4}{p}=-4$ 及 $\beta=\frac{q-4}{2p}=2$，所以 $v_n=2-4\sin^2(2^n\theta)=2\cos(2^{n+1}\theta)$，即 $2\cos(2^n\phi)$，其中 $\phi=2\theta$。

3. $u_3=u_0$ 要求 $\sin^2 8\theta=\sin^2\theta$，所以 $8\theta=\pm\theta+k\pi$。取 $\theta=\frac{\pi}{7}$：得循環 $\sin^2\frac{\pi}{7}\to\sin^2\frac{2\pi}{7}\to\sin^2\frac{4\pi}{7}\to\sin^2\frac{8\pi}{7}=\sin^2\frac{\pi}{7}$。取 $\theta=\frac{\pi}{9}$：$\sin^2\frac{\pi}{9}$，$\sin^2\frac{2\pi}{9}$，$\sin^2\frac{4\pi}{9}$。映射（map）$u\mapsto u_3$ 有 $2^3=8$ 個不動點；其中兩個（0 和 $\frac{3}{4}$）是原映射本身的不動點，餘下 6 個點的週期恰好是 3：即兩個 3-循環。

4. 每一步都把 θ 加倍，因而損失一個二進制數位的準確度，所以大約 50 步之後，計算所得的數值與真實數值已毫無關係。但這些數值並非沒有用處：計算所得的數列始終貼近一條由稍為不同的起點出發的真實軌道（orbit）（這就是影隨（shadowing）性質），所以它的統計特性仍然正確。

8.8 週期三蘊含混沌

1. 設 $a、b、c > 0$ 而 $d < 0$：$b = 7 - 2a$，$c = 7 - 2b$，$d = 7 - 2c$，$a = 7 + 2d$。解得 $17a = 49$，所以循環是 $\frac{49}{17} \to \frac{21}{17} \to \frac{77}{17} \to -\frac{35}{17} \to \frac{49}{17}$，其中三點在 $I_2 = [1,5]$ 內，一點在 $I_1 = [-3,1]$ 內。

2. 所找到的點的行程（itinerary）是 $I_2, I_2, \ldots, I_2, I_1$（其中有 $n - 1 \geq 3$ 步在 I_2 內）。循環 $1 \to 5 \to -3$ 不可能依循這個行程：-3 不在 I_2 內，所以 1（它的第二個像是 -3）、5（它的第一個像是 -3）以及 -3 本身，都不可能是該點。

3. 覆蓋圖（covering diagram）是 $[1,2] \to [3,4],[4,5]$；$[2,3] \to [4,5]$；$[3,4] \to [2,3],[3,4]$；$[4,5] \to [1,2]$。長度為 3 的迴路只有 $[3,4] \to [3,4] \to [3,4]$，它給出的是不動點，而不是 3-循環。

4. 設 $x = 14y - 7$，則 $\frac{f(x)+7}{14} = 1 - |2y - 1|$，即帳篷映射（tent map）。加倍會把二進制數位向左移一位；若移出的數位是 1，結果便是 $2 - 2y$，即把餘下的數位逐一反轉。3-循環有 $\frac{2}{7} \to \frac{4}{7} \to \frac{6}{7}$ 和 $\frac{2}{9} \to \frac{4}{9} \to \frac{8}{9}$（即 $x = -3, 1, 5$ 和 $x = -\frac{35}{9}, -\frac{7}{9}, \frac{49}{9}$）；以二進制表示，$\frac{2}{7} = 0.\overline{010}$ 而 $\frac{2}{9} = 0.\overline{001110}$。

章

9

計數與機會

計數是最古老的數學，但要數得巧妙，至今仍是最困難的事情之一。

本章內容

9.1	直線分割平面	STEP II 1990
9.2	小型版的拉姆齊定理	STEP I 1989
9.3	為籬笆樁着色	STEP II 1995
9.4	把數列看成函數	STEP I 2016
9.5	湊零錢	STEP I 1997
9.6	奇怪的骰子	STEP I 1993
9.7	相等的冪和	STEP I 1992
9.8	沒有人拿到自己的恤衫	STEP II 2016
9.9	生日問題	STEP II 2007
9.10	總能找贖	STEP I 2011
9.11	集齊一套	STEP II 1994
9.12	隨機子集	STEP III 2019
9.13	做了手腳的骰子	STEP I 1999
A	練習答案	第 655 頁

計數的藝術

組合數學是拓撲學的貧民窟。

— J. H. C. Whitehead，說於 1950 年代。今天已經沒有人抱持這種看法。

人人都會數數。但若要問：用硬幣湊出 £1 有多少種方法？安排十個人入座，使得沒有人坐在自己的位子上，有多少種方法？用四種顏色為一幅地圖着色，又有多少種方法？單純的點算很快就行不通了。情況多得無法逐一列出，而且各種情況還會互相重疊。組合數學（combinatorics）是關於計數與排列的數學，它是一門藝術：把這類問題整理得井井有條，使我們不必逐一列舉也能解決。

正如 Whitehead 那句話所反映的，在很長一段時間裏，組合數學只被視為一堆巧妙技巧的集合，而不是一門嚴肅的學科。這種情況在二十世紀下半葉改變了，推動力一方面來自計算機科學（computer science），因為點算一個算法（algorithm）所需的步驟數目，本身就是一個組合問題；另一方面則來自人們發現它與代數、幾何和概率（probability）之間有深刻的聯繫。今天，它是數學中最活躍的領域之一。2022 年，許埈珥（June Huh）獲頒數學界的最高榮譽菲爾茲獎（Fields Medal），部分原因是他運用代數幾何（algebraic geometry）的思想，證明了一個 1968 年提出的猜想，這個猜想所涉及的，正是 Section 9.3 中的色多項式（chromatic polynomial）。

計數為何重要

概率最簡單的形式就是計數：當所有結果的可能性均等時，一個事件的概率等於有利結果的數目除以結果的總數。因此，每一條關於紙牌、骰子和彩票的問題，都是計數問題；本章後半部分取材自概率試卷，正是這個緣故。但計數的重要性遠遠不止於遊戲。計算機科學家要點算一個算法所需的運算次數，以及它可能出錯的方式。化學家要點算分子可能具有的結構。生物學家要點算一條 DNA 鏈可以摺疊的方式。密碼學家要點算攻擊者必須嘗試的密鑰數目，以及（如 Section 9.9 所示）要有多少個，才會出現兩個相同的情況。

你在大學會遇到甚麼

大多數大學都會在一年級或二年級開設組合數學或離散數學（discrete mathematics）的課程，其主題正是本章的主題。你會學到基本的計數原理（乘法原理（product rule）、雙射（bijection）、容斥原理（inclusion–exclusion）、鴿巢原理（pigeonhole principle）），以及怎

樣建立和求解遞推關係（recurrence relation）。你會遇到生成函數（generating function），它把數列變成函數，把計數變成代數。你會研究圖（graph），從着色（colouring）到拉姆齊理論（Ramsey theory），還有偏序集（partially ordered set），例如子集所構成的格（lattice）。概率課程也會運用同樣的工具，再加上期望值的線性性質（linearity of expectation），來分析隨機結構。再往後，組合數學會與代數（對稱函數、表示論）、數論（分拆、加性組合）以及計算機科學（算法分析、複雜度理論）連繫起來。

貫穿本章的問題

三大問題

1. **怎樣不逐一列舉而完成計數？**有哪些結構（遞推關係、雙射、生成函數）能讓你準確地點算龐大的集合？
2. **甚麼時候某件事必然發生？**甚麼時候單憑數量之大，就足以迫使某種規律、巧合或配對出現？
3. **計數與概率怎樣互相幫助？**甚麼時候最好用計數來求概率，甚麼時候又最好用概率來求數目？

本章分為兩半。前七節取材自純數學試卷，先講遞推關係（直線分割平面、為籬笆樁着色）和鴿巢原理（拉姆齊定理），然後建立生成函數，並用它來點算硬幣、設計奇特的骰子，以及把數字分成冪和相等的隊伍。後六節是概率試卷中的經典計數問題：錯排（derangement）與 $1/e$ 的出現、生日問題、選票問題及其反射原理（reflection principle）、優惠券收集問題（coupon collector）、隨機子集，以及為何不可能把骰子灌鉛，使得每個總和出現的可能性均等。

本章的思想

§	課題	思想
9.1	平面的區域	多加一個時，有甚麼改變
9.2	拉姆齊理論	鴿巢原理
9.3	圖着色	耦合遞推關係；色多項式
9.4	生成函數	數列變成函數
9.5	分拆（partition）	透過係數來計數
9.6	Sicherman 骰子	多項式的因式分解
9.7	Thue–Morse	重根帶來平衡的和
9.8	錯排	容斥原理；$1/e$
9.9	巧合	補事件；平方根定律
9.10	選票問題	反射原理
9.11	優惠券收集問題	期望值的線性性質；$n \ln n$
9.12	隨機子集	每次只看一個元素
9.13	灌鉛骰子	以算術—幾何平均不等式（AM–GM）證明不可能

怎樣閱讀本章

永遠從小的情況入手。畫三條直線，列出三張牌的六種次序，數一數湊出 5p 的方法。小的情況能給你數據去猜想，最後也可以用來檢驗你的公式。但在找到理由之前，不要輕信任何規律：Section 9.1 中有一個著名的數列 $1, 2, 4, 8, 16, \ldots$，它的下一項是 31。

9.1 直線分割平面

題目

STEP II 1990，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/90-s2-q4`

引子

用 n 刀直線把一個薄餅切開，各塊不必大小相同，最多可以切成多少塊？一刀得 2 塊，兩刀得 4 塊，三刀得 7 塊（不是 8 塊：不妨試試），四刀得 11 塊。數列 $1, 2, 4, 7, 11, 16, \ldots$ 稱為*懶惰宴會承辦人數列*（lazy caterer's sequence），而本題要求的正是它的公式。

這道題是一種貫穿整個組合數學（combinatorics）的技巧的第一課：不要試圖一次過把所有東西數完，而應該問：多加一個對象時，有甚麼改變？而答案原來是歐拉公式（Euler's formula）$V - E + F = 2$ 的一個特殊情況，你在 Section 7.9 討論柏拉圖立體時已經見過這條公式。

動筆之前

> **停一停，試一試** 仔細畫出 $n = 1, 2, 3, 4$ 的情況，並保持各直線處於一般位置（general position）（沒有兩條平行，沒有三條共點）。數一數每種情況中線段和區域的數目。你能猜出當中的規律嗎？

你應該會得到：

n	0	1	2	3	4
線段	0	1	4	9	16
區域	1	2	4	7	11

線段的數目看來是 n^2。區域的數目則依次增加 $1, 2, 3, 4$：相鄰兩項之差正是自然數。

逐步解題

階段 1　數線段

> **想一想**　集中看其中一條直線。其餘的直線中有多少條與它相交？它們把這條直線切成多少段？

每條直線都與其餘 $n-1$ 條直線逐一相交（沒有一條與它平行），而且交於 $n-1$ 個不同的點（沒有三條直線交於同一點）。這 $n-1$ 個點把該直線切成 n 段：兩端各有一段無限長的，中間有 $n-2$ 段有限長的。共有 n 條直線，所以共有

$$n \times n = n^2 \text{ 條線段.} \qquad \blacksquare$$

這是直接點算，無需使用歸納法（induction）：我們是逐條直線去數線段的。

階段 2　加一條直線時會怎樣？

> **想一想**　假設已畫好 $n-1$ 條直線，現在加上第 n 條。它被其餘直線切成 n 段。每一段對它所穿過的區域有甚麼影響？

畫出新直線時，它的 n 段各自穿過一個原有的區域，並把該區域一分為二（圖 9.1.1）。不同的線段位於不同的區域內，因為新直線要由一個區域走到另一個區域，就必須越過一條原有的直線。因此，加上第 n 條直線後，區域的數目恰好增加 n：

$$R(n) = R(n-1) + n.$$

階段 3　解遞歸關係

由 $R(0) = 1$（沒有直線，只有一個區域）開始，

$$R(n) = 1 + (1 + 2 + \cdots + n) = 1 + \frac{n(n+1)}{2} = \frac{n^2 + n + 2}{2}.$$

若要寫成歸納法證明：公式對 $n = 1$ 成立，因為 $\frac{1}{2}(1+1+2) = 2$；而若 $R(n-1) = \frac{1}{2}\big((n-1)^2 + (n-1) + 2\big) = \frac{1}{2}(n^2 - n + 2)$，則 $R(n) = \frac{1}{2}(n^2 - n + 2) + n = \frac{1}{2}(n^2 + n + 2)$。 $\blacksquare$

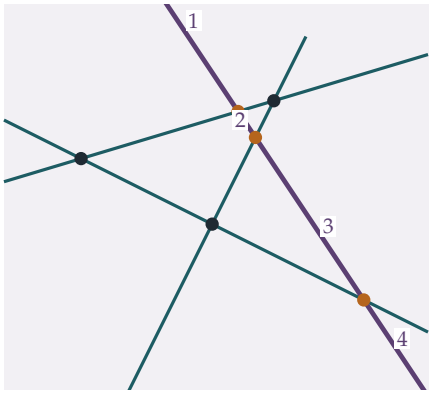


圖 9.1.1　三條直線（藍綠色）把平面分成 7 個區域。第四條直線（紫紅色）與其餘三條相交於琥珀色的點，這些點把它切成四段，編號為 1 至 4。每一段把一個區域一分為二，所以第四條直線增加了 4 個區域，合共 11 個。

回顧

- **每次只數一部分。**線段是逐條直線去數的；區域則是按每條新直線所增加的數目去數的。
- **把幾何上的一步化為遞歸關係。**「每條新線段把一個區域一分為二」這項觀察，化成了 $R(n) = R(n-1) + n$。
- **一般位置十分重要。**若有平行的直線，或有三條直線共點，區域的數目就會較少。試找出每項假設在哪裏用過。

背後的數學

喬裝的歐拉公式

以下是第二個證明，它不僅推導出答案，更解釋了答案從何而來。把這些直線組成的圖形看成一幅地圖：交點是頂點（vertices），線段是邊（edges），區域是面（faces）。圖中有 $V = \binom{n}{2}$ 個頂點（每一對直線對應一個）和 $E = n^2$ 條邊。無限長的線段只有一個端點，所以我們額外加上一個「在無窮遠處」的頂點，讓它們全部在那裏相遇；這樣，整幅圖就成為畫在球面上的地圖（想像把平面包在一個球上，無窮遠點位於球頂）。由 Section 7.9 的歐拉公式 $V - E + F = 2$ 可得

$$\left(\binom{n}{2} + 1\right) - n^2 + F = 2, \qquad \text{所以} \qquad F = n^2 - \frac{n(n-1)}{2} + 1 = \frac{n^2 + n + 2}{2}.$$

同一個數，卻來自完全不同的方向。這些區域還可以整齊地分成兩類：其中 $2n$ 個是無界的（每兩條相鄰的無限射線之間有一個），其餘 $\frac{1}{2}(n^2 + n + 2) - 2n = \binom{n-1}{2}$ 個是有界

的。

二項式係數與高維空間

這條公式可以寫成

$$R(n) = \binom{n}{0} + \binom{n}{1} + \binom{n}{2},$$

而這並非巧合。把圖形轉動，使得沒有一條直線是水平的。一個下方有界的區域必有一個最低點，而這個點必定是交點；每個交點又恰好是一個區域（即它正上方的那個）的最低點：這樣就得到 $\binom{n}{2}$ 個區域，每一對直線對應一個。其餘每個區域都向下無限延伸，因而會與一條畫在所有交點下方很遠處的水平線相遇；那 n 條直線把這條水平線切成 $n+1$ 個區間，每個這樣的區域內各有一個。合共 $\binom{n}{2} + n + 1$ 個。在三維空間中，n 個處於一般位置的平面把空間切成

$$\binom{n}{0} + \binom{n}{1} + \binom{n}{2} + \binom{n}{3} = \frac{n^3 + 5n + 6}{6}$$

塊（即蛋糕數（cake numbers）$1, 2, 4, 8, 15, 26, \ldots$），而這個規律在任何維數都成立。證明用的是同一個歸納法：新加的平面被其餘平面切成若干區域，其排列方式就如平面上的直線所分成的區域，而每個這樣的區域都把一塊空間一分為二。

施泰納與超平面排列

直線和平面所分割出的區域數目的公式，是施泰納（Jakob Steiner）在 1826 年發現的。今天，它是*超平面排列*（hyperplane arrangements）理論的起點；這套理論點算任意一組直線、平面及其高維類比所形成的區域數目，包括並非處於一般位置的情況。Thomas Zaslavsky 在 1975 年證明了一條非凡的定理：利用一個記錄了哪些超平面相交的多項式，就可以數出任何排列的區域數目。同樣的想法也出現在機器學習之中：由「ReLU」單元構成的神經網絡，會用超平面把輸入空間切成多個區域，而區域的數目正好衡量它所能表示的函數有多複雜。

思維技巧

> **問一問：多加一個時，有甚麼改變**
>
> *要點算由 n 個對象組成的構形，可把它與 $n-1$ 個對象的構形比較。點算其中的改變，往往比點算整體容易得多。*
>
> 這是把歸納法用作發現的工具，而不僅是證明的工具。它是組合數學中最常見的手法之一，能把幾何問題化為可以求解的遞歸關係（recurrence relation）。

延伸練習

1 **圓。**證明：若 n 個圓中每兩個都交於兩點，且沒有三個圓共點，則它們把平面分成 $n^2 - n + 2$ 個區域。

2 **平行線。**若 n 條直線中有 k 條互相平行（其餘方面均處於一般位置），共有多少個區域？

3 **蛋糕數。**用歸納法證明：n 個處於一般位置的平面把空間分成 $\frac{1}{6}(n^3 + 5n + 6)$ 個區域。

4 **會失效的規律。**在圓周上取 n 個點，並把每一對點用弦連起來，其中沒有三條弦在圓內交於同一點。當 $n = 1, 2, 3, 4, 5$ 時，分別得到 $1, 2, 4, 8, 16$ 個區域。數一數 $n = 6$ 時的區域數目。（提示：使用歐拉公式；每四個點給出一個交點。）證明區域的數目是 $1 + \binom{n}{2} + \binom{n}{4}$。

延伸閱讀

- **Ronald Graham, Donald Knuth & Oren Patashnik, *Concrete Mathematics***（Addison-Wesley），第 1 章。開首討論的正是這個問題，還有「薄餅」和「之字形」的變化題。
- **Miklós Bóna, *A Walk Through Combinatorics***（World Scientific）。一本平易近人的入門教材，設有歸納法、圖和歐拉公式等章節。
- **Richard Guy, "The strong law of small numbers"**，*American Mathematical Monthly* 95 (1988)。收錄了許多看似明顯、其實不成立的規律，包括練習中的圓問題。

9.2 小型版的拉姆齊定理

題目

STEP I 1989，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/89-s1-q4`

引子

在任何一個六人派對上，要麼有三個人彼此都認識，要麼有三個人彼此都不認識。無論這六個人是誰，這一點都必定成立。換成五個人，就沒有這個保證了。這個小小的事實是拉姆齊定理（*Ramsey's theorem*）的第一個情形，它道出了一條貫穿現代組合數學（combinatorics）的原則：完全的無序是不可能的。任何足夠大的結構，無論怎樣着色或排列，都必定包含一個高度有序的部分。

證明只需四行。但順理成章的下一個問題（要多少人才能保證有四個互相認識的人或四個互不認識的人？五個呢？）卻通向數學中一些最困難的未解難題。

動筆之前

想一想 那六個點之所以重要，只在於哪些點對以哪種顏色相連：幾何只會令人分心。一共有多少條線段？每個點又連出多少條？

一共有 $\binom{6}{2} = 15$ 條線段，每個點連出 5 條。（一般位置（general position）這個條件只是確保任意三點都構成一個真正的三角形。這道題其實是關於六個頂點上的完全圖（*complete graph*）K_6，它的邊以兩種顏色着色。）

逐步解題

階段 1　五條線段，兩種顏色

想一想 從 A 連出五條線段，每條不是金色就是銀色。關於它們的顏色，你可以肯定甚麼？

其中至少有三條顏色相同。假如金色的至多兩條，銀色的也至多兩條，那麼線段至多只有四條。這就是鴿巢原理（*pigeonhole principle*）：把五件物件放進兩個盒子，必有一個盒子裏至少有三件。不妨設（有需要時把兩種顏色的名稱對調）AB、AC 和 AD 全是金色。

階段 2　看看另一端

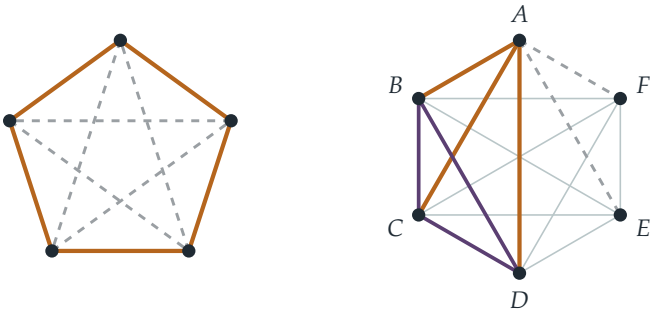
想一想 現在看看三角形 BCD。如果它有任何一條邊是金色，會怎樣？如果一條都不是呢？

如果三角形 BCD 有任何一條邊是金色，比如 BC，那麼 ABC 的三條邊 AB、AC 和 BC 全是金色。如果 BCD 沒有一條邊是金色，那麼 BC、CD 和 DB 全是銀色，BCD 就是一個銀色三角形。無論哪種情況，都有一個單色三角形（圖 9.2.1，右）。 ■

階段 3　五個點並不足夠

想一想 若只有五個點，每個點只連出四條線段。怎樣着色，才能使每個點連出的線段恰好每種顏色各兩條？

把五個點放成一個正五邊形（如果想要它們處於一般位置，可以把它們稍微提離平面）。把五邊形的五條邊塗成金色，五條對角線塗成銀色（圖 9.2.1，左）。金色的邊構成一個五圈，當中沒有三角形；銀色的邊構成一個五角星，它同樣是一個五圈，當中也沒有三角形。所以並沒有單色三角形。 ■



五個點：沒有單色三角形　　六個點：看看 BCD

圖 9.2.1　左：五個點，五邊形為金色（實線），五角星為銀色（虛線），沒有單色三角形。右：六個點時，從 A 連出的五條線段中有三條顏色相同，比如連到 B、C、D 的三條都是金色。這時三角形 BCD（紫紅色）決定了結果：只要有一條邊是金色，就與 A 構成一個金色三角形；沒有金色的邊，BCD 就是一個銀色三角形。

回顧

- **撇除無關的東西。**空間中的點變成了一個着色的圖；重要的只有邊的顏色。
- **鴿巢原理。**五條邊、兩種顏色，迫使其中三條同色。
- **對合適的對象分情況討論。**一旦找到三條金色的邊，它們另一端的三角形就決定了一切。
- **緊確性（sharpness）。**五邊形的例子說明「六」這個數不能再降低。一條好的定理，總會附帶一個例子，說明它已是最佳的結果。

背後的數學

拉姆齊數

拉姆齊數（*Ramsey number*）$R(s,t)$ 是滿足以下條件的最小的 N：把 K_N 的邊以金、銀兩色着色，無論怎樣着色，都必定包含一個金色的 K_s（即 s 個點，它們之間的所有邊都是金色）或一個銀色的 K_t。這道題已證明了 $R(3,3)=6$：六個點必定可行，而五邊形說明五個點並不可行。

弗蘭克·拉姆齊（Frank Ramsey）在 1928 年證明了這些數必定存在：對任意的 s 和 t，總有某個 N 符合要求。證明用的是歸納法（induction），把上面的論證加以推廣，並得出

$R(s,t) \le R(s-1,t) + R(s,t-1)$。但要找出確切的數值卻異常困難：

$$R(3,3) = 6, \qquad R(4,4) = 18, \qquad 43 \le R(5,5) \le 46.$$

$R(4,4)$ 的值是在 1955 年求得的。儘管已動用電腦作大規模搜尋，$R(5,5)$ 至今仍然未知；上界 46 到 2024 年才獲證明。電腦無法直接檢查 K_{43} 的每一種着色：一共有 2^{903} 種，遠遠多於宇宙中原子的數目。

艾狄胥與外星人

保羅·艾狄胥（Paul Erdős）對拉姆齊理論（Ramsey theory）發展的貢獻無人能及，他很喜歡講這個故事。假設外星人入侵地球，並威脅說除非我們告訴他們 $R(5,5)$ 的值，否則就毀滅地球。我們應當調動所有電腦和數學家，這樣或許能把它找出來。但假如他們問的是 $R(6,6)$，我們就應該設法消滅外星人。艾狄胥還在 1947 年證明了 $R(k,k) > 2^{k/2}$，所用的方法令人驚歎：他證明了，在少於 $2^{k/2}$ 個點上隨機着色的圖，有正的概率不包含任何單色的 K_k，所以好的着色必定存在，儘管沒有人知道怎樣把它構造出來。這種*概率方法*（*probabilistic method*）如今是組合數學中最強大的工具之一。

亂中有序

拉姆齊定理有很多近親，它們表達的都是同一個想法：足夠大的結構無法避開秩序。*范德瓦爾登定理*（*Van der Waerden's theorem*）說，如果把所有整數以有限種顏色着色，就存在任意長的單色等差數列（arithmetic progression）。*艾狄胥—塞凱賴什定理*（*Erdős–Szekeres theorem*）說，任何由 n^2+1 個相異的數組成的數列，都包含一個長度為 $n+1$ 的遞增或遞減子數列（subsequence）。而拉姆齊本人感興趣的是一個邏輯問題：他是在一篇關於判定邏輯命題真假的論文中，把這條定理作為引理（lemma）來證明的。

思維技巧

鴿巢原理

如果把多於 kn 件物件放進 n 個盒子，則必有一個盒子裏有多於 k 件物件。

這條原理顯而易見，但要用得好卻是一門藝術：功夫在於選定甚麼是物件、甚麼是盒子。在這道題中，物件是某一點上的五條邊，盒子是兩種顏色。在大學裏，鴿巢原理的論證會出現在數論（以分數逼近無理數）、分析（波爾查諾–魏爾斯特拉斯定理（Bolzano–Weierstrass theorem））之中，更遍及整個組合數學。

延伸練習

1 **至少兩個。**證明在 K_6 的任何雙色着色中，其實至少有兩個單色三角形。（提示：數一數「雙色角」的數目，即在同一個頂點上的一對不同顏色的邊。若一個點有 g 條金色的邊和 $5-g$ 條銀色的邊，它就是 $g(5-g) \leq 6$ 個這樣的角的頂點。每個非單色的三角形恰好包含兩個這樣的角。）

2 **三種顏色。**如果把 K_{17} 的邊以三種顏色着色，證明其中有一個單色三角形。（提示：有一個點至少有六條同色的邊。）

3 **三或四。**證明 $R(3,4) \leq 10$：在 K_{10} 的任何金、銀着色中，都有一個金色三角形或一個銀色的 K_4。（提示：一個點要麼至少有 4 條金色的邊，要麼至少有 6 條銀色的邊。）事實上 $R(3,4) = 9$。

4 **單調子數列。**就 $n = 2$ 的情形證明艾狄胥–塞凱賴什定理：任何由 5 個相異的數組成的數列，都包含一個長度為 3 的遞增或遞減子數列。找出一個由 4 個數組成、兩者皆無的數列。

延伸閱讀

- **Miklós Bóna, *A Walk Through Combinatorics***（World Scientific），組合數學入門書，可讀其中關於鴿巢原理和拉姆齊理論的章節。
- **Ronald Graham & Joel Spencer,「Ramsey theory」**，*Scientific American*（1990 年 7 月）。由兩位頂尖研究者撰寫的科普介紹。
- **Noga Alon & Joel Spencer, *The Probabilistic Method***（Wiley）。講述艾狄胥這套方法的標準著作；第一章不難讀懂。
- **Paul Hoffman, *The Man Who Loved Only Numbers***（Fourth Estate）。艾狄胥的傳記。

9.3 為籬笆樁着色

題目

STEP II 1995，第 2 題

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

`step.maths.org/questions/95-s2-q2`

引子

為排成一列的柱子髹色很容易：第一根有 3 種選擇，其後每根各有 2 種（只要不與前一根同色即可），共有 $3 \times 2^{n-1}$ 種髹法。把這一列彎成一個圓圈，問題卻變得出奇地棘手，因為最後一根柱子既要與相鄰的前一根不同色，又要與第一根不同色，而這究竟排除了一種顏色還是沒有排除任何顏色，取決於前面所有柱子的髹法。

題目用一個經典技巧繞過這個難處：按最後一根柱子的顏色把計數分成幾種情況，再找出各種情況之間如何互相轉化。答案 $2^n + 2(-1)^n$ 是*色多項式*（chromatic polynomial）的第一個例子，而色多項式正是四色定理（four colour theorem）背後的數學的一部分。

動筆之前

停一停，試一試 數一數為圍成一圈的 3 根柱子髹色有多少種方法，4 根又有多少種。$3 \times 2^{n-1}$ 有沒有多算？多算了多少？

圍成一圈的三根柱子兩兩相鄰，所以需要三種不同的顏色：共 $3! = 6$ 種方法。圍成一圈的四根柱子：18 種方法（你可以在最後驗證這一點）。與排成一列時的 $3 \times 2^2 = 12$ 和 $3 \times 2^3 = 24$ 比較：圓圈少了的，正是最後一根柱子與第一根同色的那些着色。

逐步解題

階段 1　兩類着色

想一想 考慮 $n+1$ 根柱子的一種着色（colouring），其中第一根和最後一根都是紅色。把最後一根柱子拿走。餘下的首 n 根柱子的着色有甚麼特點？

第 n 根柱子與第 $n+1$ 根相鄰，而後者是紅色，所以第 n 根不是紅色。拿走第 $n+1$ 根後，剩下的是 n 根柱子的一種着色，其中第一根是紅色而最後一根不是紅色：這正是 s_n 種着色之一。反過來，s_n 種着色中的任何一種都恰好有一種方法延伸，就是把第 $n+1$ 根柱子髹成紅色。所以兩個集合完全對應，$r_{n+1} = s_n$。

想一想 $r_n + s_n$ 是甚麼？

每一種第一根柱子為紅色的着色，其最後一根柱子不是紅色就是非紅色，所以 $r_n + s_n$ 數的是 n 根柱子中第一根為紅色的全部着色。第一根之後，每根柱子有 2 種選擇，所以 $r_n + s_n = 2^{n-1}$。因此

$$r_{n+1} + r_n = s_n + r_n = 2^{n-1}. \qquad \blacksquare$$

階段 2 用數學歸納法（induction）證明公式

當 $n = 1$ 時只有一根柱子，髹成紅色，所以 $r_1 = 1$，而公式給出 $\frac{1}{3}(1+2) = 1$。假設公式對 r_n 成立。那麼

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

這正是 r_{n+1} 的公式。 $\blacksquare$

停一停，試一試 這條公式從何而來？遞歸關係（recurrence relation）$r_{n+1} + r_n = 2^{n-1}$ 就像一條微分方程（differential equation）：解是一個「特解」（particular solution）$c \cdot 2^n$ 加上一個滿足 $r_{n+1} + r_n = 0$ 的「餘函數」（complementary function）$A(-1)^n$。求 c 和 A。

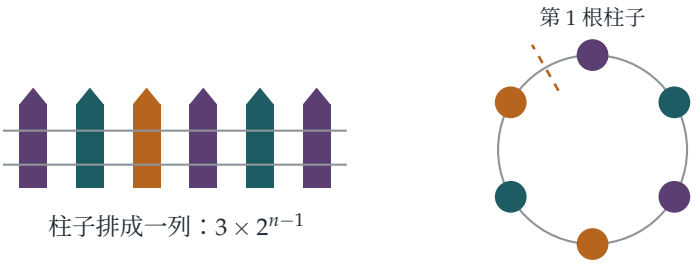
階段 3 圓圈

想一想 在 n 根柱子圍成的圓圈中，固定第 1 根柱子的顏色，例如紅色。現在在第 n 根與第 1 根柱子之間把圓圈剪開，拉直成一列。哪些直列着色對應於合乎要求的圓圈着色？

拉直之後，圓圈變成由紅色開始的一列 n 根柱子，其中最後一根還必須與第 1 根不同色：最後一根不是紅色。這些正好就是那 s_n 種着色。第 1 根柱子可以是 3 種顏色中的任何一種，所以圓圈着色的數目是

$$3s_n = 3(2^{n-1} - r_n) = 3 \cdot 2^{n-1} - 2^{n-1} - 2(-1)^{n-1} = 2^n + 2(-1)^n. \qquad \blacksquare$$

驗算：$n = 3$ 給出 $8 - 2 = 6$，而 $n = 4$ 給出 $16 + 2 = 18$。



柱子排成一列：$3 \times 2^{n-1}$

圍成一圈：$2^n + 2(-1)^n$

圖 9.3.1 左：排成一列時，第一根之後的每根柱子只須與前一根不同色。右：圍成一圈時，最後一根柱子還必須與第 1 根柱子不同色。沿虛線把圓圈剪開，圓圈着色就變成首尾兩根柱子不同色的直列着色：即題目中所數的 s_n 種着色之一。

回顧

- **分成互相轉化的情況。**記下最後一根柱子是否紅色，便得到兩個數列 r_n 和 s_n，它們由簡單的遞歸關係聯繫起來。
- **雙射（bijection）。**$r_{n+1} = s_n$ 是靠把着色一一配對來證明的，而不是靠計算。
- **剪開圓圈。**一個循環的問題變成了一個對兩端附加條件的直線問題。

背後的數學

色多項式

柱子和它們之間的相鄰關係構成一個圖（graph）：排成一列時是一條路徑（path），圍成一圈時是一個圈（cycle）。*正常着色*（proper colouring）給相鄰的頂點（vertex）不同的顏色。用 k 種顏色的正常着色數目是 k 的一個多項式，稱為該圖的*色多項式*，由伯克霍夫（George Birkhoff）於 1912 年引入。用 k 種顏色重複題目中的論證，可得

$$\text{路徑：} k(k-1)^{n-1}, \qquad \text{圈：} (k-1)^n + (-1)^n(k-1).$$

取 $k = 3$，圈的公式就是 $2^n + 2(-1)^n$。伯克霍夫希望透過證明每幅地圖所對應的圖的色多項式在 $k = 4$ 時為正數，從而證明四色定理（每幅地圖都可以用四種顏色着色，使相鄰的國家顏色不同）。這個方法始終未能成功；定理最終在 1976 年由阿佩爾（Kenneth Appel）和哈肯（Wolfgang Haken）借助大量電腦運算證明。色多項式至今仍是一個活躍的研究領域，並與統計物理（statistical physics）有出人意表的聯繫：在那裏它們以描

述磁性的波茨模型（Potts model）的形式出現。

用矩陣數出同一個答案

還有另一種方法可以看出答案，用的是第 7 章的線性代數（linear algebra）。設 $\mathbf{M}$ 為對角線上是 0、其餘位置是 1 的 3×3 矩陣：當顏色 j 可以跟在顏色 i 之後時，第 (i, j) 項為 1。那麼 $\mathbf{M}^{n-1}$ 的第 (i, j) 項數的是由顏色 i 開始、以顏色 j 結束的直列着色，所以 r_n 是 $\mathbf{M}^{n-1}$ 的一個對角線項。圓圈着色是經過 n 步後回到起點的一個序列，所以圓圈着色的數目就是 $\mathbf{M}^n$ 的跡（trace）。由於 $\mathbf{M} = \mathbf{J} - \mathbf{I}$，其中 $\mathbf{J}$ 的每一項都是 1，$\mathbf{M}$ 的特徵值（eigenvalue）是 $3 - 1 = 2$（特徵向量（eigenvector）為 $(1, 1, 1)$）和 -1（兩重），而 $\mathbf{M}^n$ 的跡是各特徵值的 n 次方之和：

$$\operatorname{tr} \mathbf{M}^n = 2^n + (-1)^n + (-1)^n = 2^n + 2(-1)^n.$$

這種轉移矩陣（transfer matrix）方法在組合數學（combinatorics）和統計物理中隨處可見，用來數算逐步構造出來的排列。

思維技巧

建立互相耦合的遞歸關係

當一個數列單靠自己無法滿足簡單的遞歸關係時，不妨多記錄一點資料（這裏是最後一根柱子的顏色），再找出聯繫各種情況的遞歸關係。

這就是狀態（state）的概念：只記下繼續下去所必須知道的資料，不多也不少。它是電腦科學中動態規劃（dynamic programming）的基礎，也是概率中馬可夫鏈（Markov chains）（Section 12.16）和轉移矩陣方法的基礎。

延伸練習

1 **k 種顏色。**用 k 種顏色重做這道題。證明 $r_{n+1} = s_n$ 仍然成立，並且 $r_{n+1} + r_n = (k-1)^{n-1}$，求 r_n，並由此推導出圈的色多項式 $(k-1)^n + (-1)^n(k-1)$。

2 **沒有兩根紅柱相鄰。**把排成一列的 n 根柱子髹成紅色或白色，而且沒有兩根相鄰的柱子同為紅色，共有多少種方法？（你應該會得到斐波那契數（Fibonacci numbers）。）

3 **輪子。**加一根位於中心的柱子，與圓圈上的每根柱子都相鄰（一個「輪子」）。用 4 種顏色為它着色有多少種方法？（提示：中心用去一種顏色，圓圈必須使用其餘三種。）

4 **項鏈。**如果圓圈的兩種着色在其中一種可由另一種旋轉而得時視為相同，那麼當 $n = 4$ 時有多少種不同的着色？$n = 5$ 時呢？（這是伯恩賽德引理（Burnside's lemma）的開端。）

延伸閱讀

- **Robin Wilson, *Introduction to Graph Theory***（Pearson）。標準的圖論入門教材；涵蓋着色和四色問題。
- **Robin Wilson, *Four Colours Suffice***（Penguin）。四色定理的故事，為一般讀者而寫。
- **Herbert Wilf, *generatingfunctionology***（可在網上免費閱讀）。講述遞歸關係及其解法，包括轉移矩陣方法。

9.4 把數列看成函數

題目

STEP I 2016，第 8 題

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

`step.maths.org/questions/16-s1-q8`

引子

用計算機算一算 1/89：0.011235955…。開首的數字是 $0, 1, 1, 2, 3, 5$，正是斐波那契數（Fibonacci numbers）。這是巧合嗎？不是，而這道題目正好解釋了原因。

這個想法是組合數學（combinatorics）中最強大的想法之一。取一個無窮數列，把各項用作冪級數（power series）的係數，從而把整個數列包裝成單一個函數。這樣，關於數列的事實就變成關於函數的事實，你便可以用代數和微積分來處理。Herbert Wilf 有一句令人難忘的話：「生成函數（generating function）是一條晾衣繩，我們把一列數字掛在上面展示。」

動筆之前

想一想 哪個數列的生成函數是 $\dfrac{1}{1-x}$？$\dfrac{1}{1-2x}$ 又如何？如果把一個數列的生成函數乘以 x，數列會有甚麼變化？

等比級數（geometric series）$\frac{1}{1-x} = 1 + x + x^2 + \cdots$ 是 $1, 1, 1, \ldots$ 的生成函數，而 $\frac{1}{1-2x} = 1 + 2x + 4x^2 + \cdots$ 則是 2 的各次冪的生成函數。乘以 x 會把每一項向後移一位：$x \cdot \frac{1}{1-x} = x + x^2 + \cdots$ 對應的數列是 $0, 1, 1, 1, \ldots$。

逐步解題

階段 1　自然數與三角形數

想一想　用二項級數（binomial series）展開 $(1-x)^{-2}$。x^n 的係數是甚麼？

由二項級數可得

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

所以 $x(1-x)^{-2} = x + 2x^2 + 3x^3 + \cdots = \sum nx^n$：這就是 $u_n = n$ 的生成函數。

同樣地，$(1-x)^{-3}$ 中 x^n 的係數是 $\frac{(-3)(-4)\cdots(-3-n+1)}{n!}(-1)^n = \frac{(n+1)(n+2)}{2}$，所以

$$x(1-x)^{-3} = \sum_{n\geq 1} \frac{n(n+1)}{2}\, x^n = x + 3x^2 + 6x^3 + 10x^4 + \cdots,$$

這就是三角形數（triangle numbers）$u_n = \frac{1}{2}n(n+1)$ 的生成函數。

想一想　怎樣用 n 和 $\frac{1}{2}n(n+1)$ 來表示 n^2？

由於 $n^2 = 2\cdot\frac{1}{2}n(n+1) - n$，而生成函數可以像數列一樣相加，

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

階段 2　遞推關係變成方程

想一想　把 $u_n = ku_{n-1}$ 乘以 x^n，然後對所有 $n \geq 1$ 求和。兩邊用 f 表示，分別是甚麼？

左邊是 $\sum_{n\geq 1} u_n x^n = \mathrm{f}(x) - u_0 = \mathrm{f}(x) - a$。右邊是 $kx\sum_{n\geq 1} u_{n-1}x^{n-1} = kx\,\mathrm{f}(x)$。所以 $\mathrm{f}(x) = a + kx\,\mathrm{f}(x)$，於是

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

當然，$u_n = ak^n$；把 $\frac{a}{1-kx}$ 展開成等比級數，便可確認這一點。

階段 3　斐波那契數

對 $u_n = u_{n-1} + u_{n-2}$（其中 $n \geq 2$）作同樣的處理：

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

即

$$\mathrm{f}(x) - u_0 - u_1 x = x\big(\mathrm{f}(x) - u_0\big) + x^2\,\mathrm{f}(x).$$

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

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

階段 4 **解釋** $1/89$

> **停一停，試一試** 暫且不理收斂（convergence）的問題（這裏沒有問題：級數在 $|x| < 0.618$ 時收斂），把 $x = \frac{1}{10}$ 代入 f。你得到甚麼？

$$\frac{1}{10}F_1 + \frac{1}{100}F_2 + \frac{1}{1000}F_3 + \cdots = \mathrm{f}\Big(\frac{1}{10}\Big) = \frac{1/10}{1 - 1/10 - 1/100} = \frac{10}{89}.$$

除以 10，得 $\frac{1}{89} = \sum F_n/10^{n+1}$。斐波那契數依次放在相繼的小數位上，當它們大於 9 時便會互相重疊（圖 9.4.1）。$1/89$ 開首的數字顯示出斐波那契數；其後的數字則因進位而變得模糊。

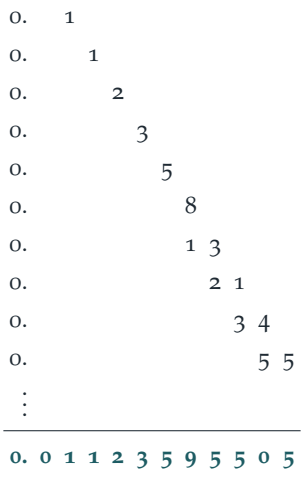


圖 9.4.1 把 $F_n/10^{n+1}$ 對 $n = 1, 2, 3, \ldots$ 相加。每個斐波那契數的末位都比前一個向右多移一個小數位。一旦它們有兩位數字，便會互相重疊，而進位會改變第七位起的數字：$\frac{1}{89} = 0.01123595505\ldots$。

回顧

- **把數列變成函數。**整個無窮數列變成了一個有理函數（rational function）。

- **遞推關係變成方程。**乘以 x^n 再求和，便把規則 $u_n = u_{n-1} + u_{n-2}$ 變成一條關於 f 的方程，可以用代數方法解出。
- **從舊的造出新的。**把生成函數相加、平移和求導，便能得出新的生成函數，無須從頭做起。

背後的數學

從生成函數到公式

生成函數不只是把數列編碼：它還可以把數列解出來。分母可以因式分解為 $1 - x - x^2 = (1 - \varphi x)(1 - \psi x)$，其中 $\varphi = \frac{1+\sqrt{5}}{2}$ 和 $\psi = \frac{1-\sqrt{5}}{2}$ 是 $t^2 = t + 1$ 的根。由部分分式（partial fractions）可得

$$\frac{x}{1 - x - x^2} = \frac{1}{\sqrt{5}}\left(\frac{1}{1 - \varphi x} - \frac{1}{1 - \psi x}\right),$$

再把每個分式展開成等比級數，

$$F_n = \frac{\varphi^n - \psi^n}{\sqrt{5}}.$$

這就是*比內公式*（*Binet's formula*）。由於 $|\psi| < 1$，第二項非常小，所以 F_n 是最接近 $\varphi^n/\sqrt{5}$ 的整數，而相鄰兩個斐波那契數之比趨向黃金比（golden ratio）（這是 Section 2.2 的主題）。同一方法可以解任何常係數線性遞推關係（linear recurrence with constant coefficients）；練習中有一個來自 STEP I 2004 的例子。

生成函數為何如此有用

真正的威力在相乘時才顯現出來。乘積 $f(x)g(x)$ 中 x^n 的係數是 $\sum_k u_k v_{n-k}$，即對所有把 n 拆成兩部分的方法求和。要數算由兩個部分組成的東西，這正是我們所需要的。接下來的兩節會用它來數算找換零錢的方法（Section 9.5），以及找出點數古怪、但總和的分佈與普通骰子相同的骰子（Section 9.6）。在概率中，同一想法給出*概率生成函數*（*probability generating function*），它把獨立隨機變量之和變成乘積。

棣莫弗、歐拉與拉普拉斯

棣莫弗（Abraham de Moivre）在 1720 年代用生成函數來解線性遞推關係，所用的方法基本上就是上面的方法。歐拉（Leonhard Euler）在 1748 年用它們來研究整數的分拆（partitions），而拉普拉斯（Pierre-Simon Laplace）則以它們作為其著作 *Théorie analytique des probabilités*（1812）的基礎。到了二十世紀，生成函數成為組合數學的核心，而解析方法（把生成函數看成複變數的函數來研究）如今能為複雜得無法精確求解的計數問題給出精確的估計。

思維技巧

轉換表示方式

一個關於數列的問題，在把數列變成函數、矩陣或圖形之後，可能會變得容易。

生成函數把遞推關係翻譯成代數，把計數翻譯成乘法。在大學裏你會遇到許多這類翻譯：傅立葉級數（Fourier series）把函數變成數列，拉普拉斯變換（Laplace transforms）把微分方程變成代數，而矩陣則把幾何變成算術。

延伸練習

1 **每次兩位數字。**試做 STEP II 2022，第 3 題（載於 OCR 提供下載的 2022 年試卷，`www.ocr.org.uk/Images/696587-2022-past-papers.zip`）。它解釋了為何 $\frac{1}{89}$ 的小數展開式以斐波那契數開首，並要求找出一個分數，其小數每次兩位數字地顯示這個規律。

2 **解遞推關係。**試做 STEP I 2004，第 7 題（`step.maths.org/questions/04-s1-q7`）。它用生成函數和部分分式解一個線性遞推關係，然後把這個方法倒過來做。

3 **逐一數數。**證明 $\frac{1}{9801} = 0.00\,01\,02\,03\,04\ldots$，其中每兩位數字為一組，逐一向上數。這個規律最先在哪裏出錯？為甚麼？

4 **平方和。**解釋為何 $\frac{1}{1-x}f(x)$ 是部分和 $u_0 + u_1 + \cdots + u_n$ 的生成函數。利用這一點和第 (i) 部分的答案，求 $1^2 + 2^2 + \cdots + n^2$ 的公式。

延伸閱讀

- **Herbert Wilf, *generatingfunctionology***（A K Peters；網上免費閱讀）。經典的入門書。讀完本節後，最適合接着讀第 1 章。
- **Ronald Graham, Donald Knuth & Oren Patashnik, *Concrete Mathematics***（Addison-Wesley），第 7 章。
- **Philippe Flajolet & Robert Sedgewick, *Analytic Combinatorics***（Cambridge University Press；網上免費閱讀）。遠遠超出大學一年級的程度，但開首數頁展示了這門學問的去向。

9.5 湊零錢

題目

STEP I 1997，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/97-s1-q1`

引子

這看來是全書最容易的一題：把所有可能逐一列出便是。可是列舉很快就會失控。用英國人口袋裏的硬幣，湊出 20p 有 41 種方法，湊出 50p 有 451 種，湊出 £1 更有 4563 種。這道題真正要求的是一套*系統*，也就是一種整理計數的方法，使得既無遺漏，亦無重複。

這套系統背後是一段優美的數學。每種硬幣對應一個等比級數（geometric series），湊出 n 便士的方法數就是這些級數之積的一個係數（coefficient）；把同一個想法一併用於所有正整數，便得到歐拉（Euler）的分拆（partition）理論，那是數論中最深刻的課題之一。

動筆之前

想一想 只用 2p 和 1p 硬幣，湊出 n 便士有多少種方法？為甚麼答案只取決於你用了多少枚 2p 硬幣？

一旦選定 2p 硬幣的數目，其餘的就必須全是 1p 硬幣，所以 2p 硬幣的每個可能數目 $0, 1, \ldots, \lfloor n/2 \rfloor$ 都恰好對應一種方法。這樣共有 $\lfloor n/2 \rfloor + 1$ 種方法，記作 $a(n)$。例如 $a(10) = 6$，$a(5) = 3$。

逐步解題

階段 1　十便士

想一想　按面值最大的硬幣來整理計數。你可以用多少枚 10p 硬幣？定下來之後，又可以用多少枚 5p 硬幣？

- 一枚 10p 硬幣：1 種方法。
- 不用 10p，用兩枚 5p：1 種方法。
- 不用 10p，用一枚 5p：餘下的 5p 由 2p 和 1p 湊成，有 $a(5) = 3$ 種方法。
- 不用 10p 也不用 5p：10p 全由 2p 和 1p 湊成，有 $a(10) = 6$ 種方法。

總數是 $1 + 1 + 3 + 6 = 11$。 ■

階段 2　逐層建立

想一想　設 $b(n)$ 為用 5p、2p 和 1p 硬幣湊出 n 便士的方法數。怎樣用函數 a 表示 $b(n)$？

選定 5p 硬幣的數目 k；餘下的 $n - 5k$ 便士由 2p 和 1p 硬幣湊成。所以

$$b(n) = a(n) + a(n-5) + a(n-10) + \cdots .$$

由此得

n	0	5	10	15	20
$a(n)$	1	3	6	8	11
$b(n) = a(n) + b(n-5)$	1	4	10	18	29

同樣地，設 $c(n)$ 為連 10p 硬幣也可使用時的方法數；則 $c(n) = b(n) + b(n-10) + \cdots$，所以 $c(10) = 10 + 1 = 11$（與第 1 階段的結果一致），而 $c(20) = b(20) + b(10) + b(0) = 29 + 10 + 1 = 40$。

階段 3　二十便士

最後輪到 20p 硬幣：我們要麼用它（一種方法），要麼不用（$c(20)$ 種方法）。所以共有

$$40 + 1 = 41 \text{ 種方法.} \qquad \blacksquare$$

停一停，試一試 留意當中的結構：每加入一種新硬幣就多一層，而每一層都是把上一層按新硬幣的面值為間隔取值再相加。你能看出電腦會怎樣為 £1 甚至 £100 進行這項計算嗎？

回顧

- **按最大的部分來整理。**先固定面值最大的硬幣的數目，計數便有條不紊。
- **建立表格。**每加入一種新硬幣，就多出一行，由上一行計算出來。這就是動態規劃（dynamic programming），是電腦計算這類數目的標準方法。
- **與已知的情況核對。**表格重現了 10p 的 11 種方法。

背後的數學

把硬幣看成等比級數

以下這個想法把問題化為代數。所用 5p 硬幣的數目可以是 $0, 1, 2, \ldots$，分別貢獻 $0, 5, 10, \ldots$ 便士。把這一點記錄成級數

$$1 + x^5 + x^{10} + x^{15} + \cdots = \frac{1}{1 - x^5}.$$

對每種硬幣都這樣做，然後相乘：

$$\frac{1}{(1-x)(1-x^2)(1-x^5)(1-x^{10})(1-x^{20})} = \sum_n w(n) x^n.$$

把各括號乘開時，每一項 x^n 都來自在每個括號中各選一項，也就是為每種硬幣各選一個數目，而總值為 n。所以 x^n 的係數 $w(n)$ 正好就是湊出 n 便士的方法數。第 2 階段的表格，其實就是把這個乘積逐個括號乘開。這正是 Section 9.4 中母函數（generating function）的想法，而乘法正是它大顯威力之處：母函數之積所計算的，就是各項選擇的組合數目。

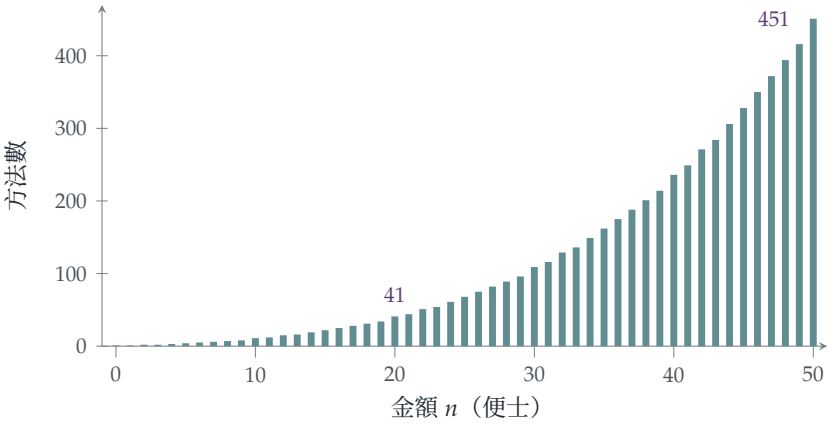


圖 9.5.1　用 1p、2p、5p、10p、20p 和 50p 硬幣湊出 n 便士的方法數。在 5 和 10 的倍數處出現的跳升，顯示新硬幣首次可用的位置。有六種硬幣時，方法數的增長大致與 n^5 相若：一般而言，增長有如一個多項式，其次數比硬幣種數少一。

分拆

假如每個正整數都可以作為硬幣的面值，又會怎樣？這時，不計次序而把 n 寫成正整數之和的方法數，就是**分拆數**（partition number）$p(n)$。例如 $p(4) = 5$，因為 $4 = 3+1 = 2+2 = 2+1+1 = 1+1+1+1$。歐拉在 1748 年看出

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

並用它發現了一些非凡的恆等式（identity）。其中之一是：把 n 寫成*互不相同*的部分之和的方法數，等於把它寫成*奇數*部分之和的方法數。（用母函數來證明，只需一行代數：請在練習中一試。）

哈代、拉馬努金與 $p(200)$

分拆數增長得很快：$p(10) = 42$，$p(100) = 190\,569\,292$。1918 年，哈代（G. H. Hardy）與拉馬努金（Srinivasa Ramanujan）為 $p(n)$ 找到一條令人驚歎的公式，那是一個無窮級數（infinite series），其首項為

$$p(n) \approx \frac{1}{4n\sqrt{3}}\, e^{\pi\sqrt{2n/3}}.$$

為了檢驗這條公式，Percy MacMahon 利用歐拉的恆等式，徒手算出 $p(200) = 3\,972\,999\,029\,388$。哈代與拉馬努金的級數只取首幾項，所得結果與正確答案相差不足一。拉馬努金還發現 $p(5n+4)$ 必定可被 5 整除，$p(7n+5)$ 必定可被 7 整除，而 $p(11n+6)$ 必定可被 11 整除。至於為何再沒有其他這類規律，則要到二十一世紀才得

到解釋。

思維技巧

藉係數來計數

當一個對象由若干獨立的選擇構成時，每項選擇就是一個因式，而總大小爲 n 的對象的數目，就是乘積中 x^n 的係數。

這把計數化為代數。在大學裏，它是計數組合學（enumerative combinatorics）的基礎，而概率母函數（probability generating function）也正是這樣處理獨立隨機變量（independent random variables）之和的。

延伸練習

1 **兩種六。**試做 STEP II 2012，第 1 題（`step.maths.org/questions/12-s2-q1`）。該題利用等比級數之積來計數。做完之後，請描述一個以這些係數為答案的硬幣計數問題。

2 **互異與奇數。**證明 $(1+x)(1+x^2)(1+x^3)\cdots$ 是分拆成互不相同部分的母函數。利用 $1+x^k = \frac{1-x^{2k}}{1-x^k}$ 證明它等於 $\frac{1}{(1-x)(1-x^3)(1-x^5)\cdots}$，並由此推出歐拉定理。就 $n=6$ 驗證這個結果。

3 **一條公式。**證明只用 1p、2p 和 5p 硬幣湊出 n 便士的方法數，是最接近 $\frac{(n+4)^2}{20}$ 的整數。（先用較小的 n 驗證。）

4 **次序有別。**把 1p 和 2p 硬幣逐枚投入投幣口來湊出 n 便士，若次序有別，共有多少種方法？（這個數列（sequence）你以前見過。）

延伸閱讀

- **George Andrews & Kimmo Eriksson, *Integer Partitions***（Cambridge University Press）。一本為大學本科生而寫的入門書，篇幅短小，淺白易讀。
- **波利亞（George Pólya），「On picture-writing」**，*American Mathematical Monthly* 63 (1956)。波利亞以兌換一美元的問題（共有 292 種方法）來講解母函數的經典文章。
- **Robert Kanigel, *The Man Who Knew Infinity***（Abacus）。拉馬努金的生平，以及他與哈

代的合作。

9.6 奇怪的骰子

題目

STEP I 1993，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/93-s1-q1`

引子

兩顆六面骰子，各面標上正整數，除了平常的標法之外，還有沒有別的標法，可以令由 2 到 12 的每個點數和出現的次數，都與普通骰子完全相同？1978 年，Martin Gardner 在《科學人》（*Scientific American*）雜誌上宣布，這樣的標法恰好只有一種：一顆骰子標上 $1,2,2,3,3,4$，另一顆標上 $1,3,4,5,6,8$。這對骰子由 George Sicherman 上校發現，現在市面上有售，稱為 *Sicherman* 骰子。用它們來玩一局大富翁（Monopoly），與用普通骰子玩根本無從分辨。

本題是一道熱身題：從點數和的列表，推斷兩顆骰子各面標了甚麼數字。要有系統地解決這個問題，以至證明 Sicherman 骰子是唯一的一對奇異骰子，方法都是把每顆骰子化為一個多項式（polynomial）。

動筆之前

想一想 假設第一顆骰子有 p_i 個面顯示 i，第二顆有 q_j 個面顯示 j，那麼擲出點數和 t 共有多少種方法？

方法的數目是 $\sum_i p_i q_{t-i}$：先選定第一顆骰子的點數 i，第二顆就必須顯示 $t-i$。就題目中的第一對骰子而言，點數和 6 可以來自 $1+5$（有 1×1 種方法）、$2+4$（2×1 種）、$3+3$（2×2 種），此外再無其他可能，因為第一顆骰子沒有 4 或 5，而第二顆沒有 0。所以共有 $1+2+4=7$ 種方法。 ■

逐步解題

階段 1　兩端的情況

把兩顆骰子稱為 A 和 B，並以 a_i、b_i 分別表示兩顆骰子上顯示 i 的面數。

> **想一想**　點數和不可能是 2，也不可能是 12。這對 1 和 6 的分佈有甚麼啟示？擲出 3 卻確實有兩種方法。哪一顆骰子上有 1？

擲不出點數和 2，即 $a_1 b_1 = 0$：最多只有一顆骰子上有 1。點數和 3 只可能是 $1+2$ 或 $2+1$，而且有兩種方法，所以其中一顆骰子確實有 1；把它稱為 A，於是 $b_1 = 0$。那麼

$$\text{點數和3：}\quad a_1 b_2 = 2, \qquad \text{所以} (a_1, b_2) = (2,1) \text{ 或} (1,2).$$

同理，擲不出點數和 12，即 $a_6 b_6 = 0$。

階段 2　逐步向上推

> **想一想**　寫出點數和為 4 的方程，記住 $b_1 = 0$。在兩種情況下，這個方程分別說明了甚麼？

點數和 4 是 $1+3$ 或 $2+2$（由於 $b_1 = 0$，$3+1$ 不可能出現），所以 $a_1 b_3 + a_2 b_2 = 1$。

- **情況 (a)**，$a_1 = 2$，$b_2 = 1$：$2b_3 + a_2 = 1$，所以 $b_3 = 0$ 及 $a_2 = 1$。
- **情況 (b)**，$a_1 = 1$，$b_2 = 2$：$b_3 + 2a_2 = 1$，所以 $a_2 = 0$ 及 $b_3 = 1$。

點數和 5：$a_1 b_4 + a_2 b_3 + a_3 b_2 = 1$。

- 情況 (a)：$2b_4 + 0 + a_3 = 1$，所以 $b_4 = 0$，$a_3 = 1$。至此 A 有 $1,1,2,3$，B 有 2；B 其餘五個面都是 5 或 6。
- 情況 (b)：$b_4 + 0 + 2a_3 = 1$，所以 $a_3 = 0$，$b_4 = 1$。至此 A 有 1，B 有 $2,2,3,4$；A 其餘五個面取自 $4,5,6$，B 其餘兩個面取自 $5,6$。

階段 3　情況 (a) 不成立

> **想一想**　在情況 (a) 中，寫出點數和為 7 的方程，並利用 $b_5 + b_6 = 5$。

點數和 7：$a_1 b_6 + a_2 b_5 + a_3 b_4 + a_4 b_3 + a_5 b_2 = 2b_6 + b_5 + 0 + 0 + a_5 = 3$。代入 $b_6 = 5 - b_5$，得 $10 - b_5 + a_5 = 3$，所以 $b_5 = 7 + a_5 \geq 7$。但 B 只有六個面。情況 (a) 不可能成立。

階段 4 情況 (b) 成立

想一想 在情況 (b) 中，利用點數和 6 和 11，以及 $a_6b_6 = 0$ 這個事實。

點數和 6：$a_1b_5 + a_4b_2 = b_5 + 2a_4 = 4$（其餘各個乘積都含有 $a_2 = a_3 = 0$ 或 $b_1 = 0$）。由於 $b_5 \le 2$，而且 $b_5 = 4 - 2a_4$ 是偶數，所以要麼 $b_5 = 0, a_4 = 2$，要麼 $b_5 = 2, a_4 = 1$。

- 若 $b_5 = 0$，則 B 是 $2,2,3,4,6,6$，於是 $a_6 = 0$，而 A 是 $1,4,4,5,5,5$。但這樣點數和 7 就有 $a_1b_6 + a_4b_3 + a_5b_2 = 2 + 2 + 6 = 10$ 種方法，而不是 3 種。
- 若 $b_5 = 2$，則 B 是 $2,2,3,4,5,5$，其中 $b_6 = 0$，而 A 是 $1,4$ 再加上四個取自 5 和 6 的面。點數和 11：$a_6b_5 = 2a_6 = 6$，所以 $a_6 = 3$ 及 $a_5 = 1$。

唯一的可能是

$$A:\ 1,4,5,6,6,6, \qquad B:\ 2,2,3,4,5,5.$$

最後，我們必須驗證這對骰子確實給出整個列表，因為我們只用了其中一部分點數和。把乘積展開（或利用下文的多項式），即可確認每一項都吻合。 ■

回顧

- **從兩端入手。**最小和最大的點數和只牽涉一兩個乘積，所以提供的資訊最有力。
- **分情況討論，逐一排除。**每個點數和都把可能性收窄一些，直到只剩下一種情況。
- **檢查充分性（sufficiency）。**推論說明了骰子*必須*是甚麼樣子；最後的驗證則說明這對骰子確實可行。

背後的數學

把骰子看成多項式

把每顆骰子寫成一個多項式：每個顯示 i 的面對應一項 x^i。普通骰子是 $x + x^2 + x^3 + x^4 + x^5 + x^6$，而題目中的兩顆骰子是

$$A(x) = x + x^4 + x^5 + 3x^6, \qquad B(x) = 2x^2 + x^3 + x^4 + 2x^5.$$

相乘時，每一對面貢獻 $x^i x^j = x^{i+j}$，所以 $A(x)B(x)$ 中 x^t 的係數（coefficient），就是擲出點數和 t 的方法數目。這又是 Section 9.4 中母函數（generating function）的想法：點數和的列表其實就是乘積 $A(x)B(x)$，而我們逐個情況的推論，正是把它因式分解（factorise）的一種方法。

Sicherman 骰子

兩顆普通骰子的乘積是 $(x+x^2+\cdots+x^6)^2$。任何一對點數和分佈相同的骰子，其多項式 $P(x)$、$Q(x)$ 必須滿足 $P(x)Q(x)=(x+\cdots+x^6)^2$。現在作因式分解：

$$x+x^2+\cdots+x^6=x(1+x)(1+x+x^2)(1-x+x^2),$$

而這些因式不能再進一步分解（成為整數係數的多項式）。所以 P 和 Q 都必須由因式 x、$1+x$、$1+x+x^2$、$1-x+x^2$ 組成，而每個因式合共出現兩次。以下三個條件把選擇收窄：

- 每顆骰子有 6 個面，所以 $P(1)=Q(1)=6$；各因式在 $x=1$ 時的值分別是 1、2、3、1，所以每顆骰子各得一個 $1+x$ 和一個 $1+x+x^2$；
- 沒有任何面顯示 0，所以 $P(0)=Q(0)=0$：每顆骰子各得一個因式 x；
- 餘下唯一的自由，是兩個 $1-x+x^2$ 如何分配。每顆各得一個，就是普通骰子。兩個都歸同一顆骰子，則得

$$x(1+x)(1+x+x^2)=x+2x^2+2x^3+x^4,$$
$$x(1+x)(1+x+x^2)(1-x+x^2)^2=x+x^3+x^4+x^5+x^6+x^8.$$

這就是 Sicherman 骰子，1,2,2,3,3,4 和 1,3,4,5,6,8（圖 9.6.1），而以上論證亦證明了它們是唯一的另一種可能。關鍵在於：多項式與整數一樣，可以唯一地分解為不可約（irreducible）因式的乘積——數論與組合數學（combinatorics）在此相遇。

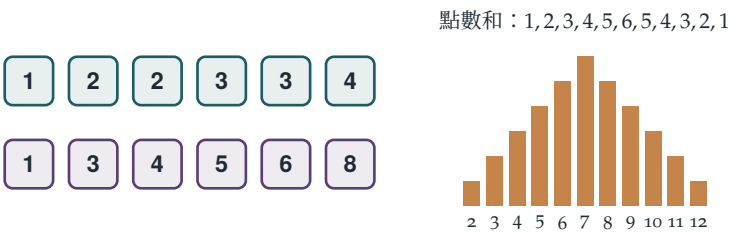


圖 9.6.1　Sicherman 骰子。兩顆一起擲，由 2 到 12 的每個點數和出現的頻數，都與兩顆普通骰子完全相同（右）：擲出 2 有一種方法，擲出 3 有兩種方法，如此類推，直至擲出 7 有六種方法。

概率母函數

在概率論中，同一個想法以*概率母函數*（probability generating function）的形式出現：取值 $0,1,2,\ldots$ 的隨機變量（random variable）X 有 $G(t)=E(t^X)=\sum P(X=k)t^k$，而對於獨立（independent）的 X 和 Y，$E(t^{X+Y})=E(t^X)E(t^Y)$。獨立隨機變量之和，對應於母函數之積。STEP III 2003 Q14（見練習）就是用這個方法找出 Sicherman 骰子，

而 Section 9.13 則用一個相關的論證，說明無論怎樣在兩顆骰子上做手腳，都不能令所有點數和出現的機會均等。

思維技巧

把計數問題編碼成多項式

當物件以數值相加的方式組合時，把每個物件表示成 x 的多項式。組合變成相乘，而有關組合的問題就變成有關因式分解的問題。

在這裏，一個關於骰子的問題化成了多項式的唯一分解（unique factorisation）。在大學裏，這類編碼隨處可見：在組合數學中，在概率論中，也在代數中——在代數裏，多項式本身就是研究對象，它們構成一個環（ring），其算術與整數的算術相似。

延伸練習

1 **四面體骰子。**試做 STEP III 2003，第 14 題（`step.maths.org/questions/03-s3-q14`），該題就四面骰子提出 Sicherman 的問題。（提示：$x + x^2 + x^3 + x^4 = x(1+x)(1+x^2)$。）

2 **驗證列表。**展開 $(x + x^4 + x^5 + 3x^6)(2x^2 + x^3 + x^4 + 2x^5)$，並與題目中的列表對照。

3 **三顆骰子。**是否存在三顆六面骰子，各面標上正整數，不全是標準骰子，而其點數和的分佈與三顆標準骰子相同？（提示：$(x + \cdots + x^6)^3$ 的因式現在包括三個 $1 - x + x^2$。）

4 **一枚硬幣與一個八面體。**一顆標上 $1, \ldots, 8$ 的八面骰子，和一顆標上 $1, 2$ 的兩面「骰子」（即一枚硬幣）：能否用另一種方式重新標號，而點數和的分佈保持不變？

延伸閱讀

- **Martin Gardner, *Penrose Tiles to Trapdoor Ciphers*** (Mathematical Association of America)。收錄了關於 Sicherman 骰子的那篇專欄文章。
- **Joseph Gallian & David Rusin, "Cyclotomic polynomials and nonstandard dice"**, *Discrete Mathematics* 27 (1979)。一般理論，適用於任意面數的骰子。

► **Herbert Wilf, *generatingfunctionology***（可於網上免費閱覽），第 1 章。

9.7 相等的冪和

題目

STEP I 1992，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/92-s1-q5`

引子

把 1 至 16 這些數分成兩隊，每隊八個：

$$\{1,4,6,7,10,11,13,16\} \quad 和 \quad \{2,3,5,8,9,12,14,15\}.$$

兩隊的和都是 68。兩隊的平方和都是 748。兩隊的立方和都是 9248。很難相信這不是純粹的巧合，但它其實是一個無窮無盡的規律的第三步：只要用正確的方法把 1 至 32 分開，連四次方之和也會相等。

這個分法由 Eugène Prouhet 於 1851 年發現，而它背後的規則，即 *Thue–Morse* 數列（Thue–Morse sequence），出現的地方多得驚人：數論、計算理論、國際象棋，以至最公平的輪流方式。本題透過一個多項式（polynomial）及其導數（derivative）把它發掘出來。

動筆之前

想一想 若多項式 $\mathrm{p}(x) = \sum c_k x^k$ 滿足 $\mathrm{p}(1) = 0$，這對各係數（coefficient）有甚麼含意？若 $\mathrm{p}'(1) = 0$ 又如何？

$\mathrm{p}(1) = \sum c_k$，所以 $\mathrm{p}(1) = 0$ 表示各係數之和為零。而 $\mathrm{p}'(x) = \sum k c_k x^{k-1}$，所以 $\mathrm{p}'(1) = \sum k c_k$。若所有係數都是 ± 1，這兩個條件就是說：$+1$ 與 -1 所在的位置數目相同，而且位置之和也相等。

逐步解題

階段 1　一個三重根

? **想一想**　把 $1-x^2$ 和 $1-x^4$ 因式分解。當中共出現多少個因式 $1-x$？

由於 $1-x^2=(1-x)(1+x)$ 及 $1-x^4=(1-x)(1+x)(1+x^2)$，

$$\mathrm{p}_0(x)=(1-x)^3(1+x)^2(1+x^2),$$

所以 $(1-x)^3$ 是一個因式（factor）。 ■

階段 2　在 1 處的導數

? **想一想**　若 $(1-x)^3$ 整除 p_0，那麼 $(1-x)$ 的多少次方整除 p_0'？接着 $\mathrm{p}_1=x\mathrm{p}_0'$ 和 p_1' 又如何？

寫成 $\mathrm{p}_0(x)=(1-x)^3\mathrm{q}(x)$。根據積法則（product rule），$\mathrm{p}_0'(x)=-3(1-x)^2\mathrm{q}(x)+(1-x)^3\mathrm{q}'(x)$，它有因式 $(1-x)^2$。所以 $\mathrm{p}_0'(1)=0$。於是 $\mathrm{p}_1(x)=x\mathrm{p}_0'(x)$ 同樣有因式 $(1-x)^2$，而用相同的論證可知 p_1' 有因式 $(1-x)$，所以 $\mathrm{p}_1'(1)=0$。 ■

階段 3　讀出各係數

展開可得 $(1-x)(1-x^2)=1-x-x^2+x^3$，然後

$$\mathrm{p}_0(x)=1-x-x^2+x^3-x^4+x^5+x^6-x^7.$$

? **想一想**　以係數 c_k 表示，$\mathrm{p}_0'(1)$ 和 $\mathrm{p}_1'(1)$ 分別是甚麼？

由於 $\mathrm{p}_1(x)=x\mathrm{p}_0'(x)=\sum kc_kx^k$，我們有 $\mathrm{p}_1'(1)=\sum k^2c_k$。係數在 $k=0,3,5,6$ 時為 $+1$，在 $k=1,2,4,7$ 時為 -1。所以 $\mathrm{p}_0'(1)=0$ 和 $\mathrm{p}_1'(1)=0$ 即是說

$$0+3+5+6=1+2+4+7\quad(=14),$$
$$0^2+3^2+5^2+6^2=1^2+2^2+4^2+7^2\quad(=70),$$

略去 0 之後，這正是所需的恆等式。 ■

階段 4　再多一個因式

> **停一停，試一試** 要令立方和也平衡，我們需要第四個因式 $(1-x)$。應該用哪一個多項式取代 p_0？

取 $\mathrm{P}(x)=(1-x)(1-x^2)(1-x^4)(1-x^8)$，它有因式 $(1-x)^4$。它的係數來自 $\mathrm{p}_0(x)-x^8\mathrm{p}_0(x)$：先是 0 至 7 的正負號規律，然後在 8 至 15 重複同一規律，但正負號全部相反：

k	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
c_k	+	−	−	+	−	+	+	−	−	+	+	−	+	−	−	+

把第 2 階段的論證再用一次，可知對於 $r=0,1,2,3$ 都有 $\sum c_k k^r=0$。取 $+$ 號的位置 $\{0,3,5,6,9,10,12,15\}$ 和 $-$ 號的位置 $\{1,2,4,7,8,11,13,14\}$。當 $r=1,2,3$ 時，來自 0 的一項為零，所以可以略去。因此以下的次序

$$a_1,\ldots,a_8=1,2,4,7,8,11,13,14,\qquad a_9,\ldots,a_{15}=3,5,6,9,10,12,15$$

符合要求：兩半的和都是 60，平方和都是 620，立方和都是 7200。 ■

回顧

- **導數能偵測重根（repeated root）。**因式 $(1-x)^m$ 會迫使首 $m-1$ 階導數在 $x=1$ 處全部為零。
- **先乘以 x 再求導**把 $\sum kc_k$ 變成 $\sum k^2c_k$，如此類推便得到更高次方。
- **倍增。**每多一個因式 $(1-x^{2^m})$，長度便增加一倍，同時多一個次方達到平衡。

背後的數學

Thue–Morse 數列

$(1-x)(1-x^2)(1-x^4)(1-x^8)\cdots$ 各係數的正負號構成 *Thue–Morse* 數列。以 0 代表 $+$、1 代表 $-$，寫出來就是：

$$0110100110010110\ \ldots$$

它有三種等價的描述。

- **倍增：**由 0 開始，不斷把手上已有的部分的補（complement，即把 0 與 1 互換）接在後面：$0\to01\to0110\to01101001\to\cdots$。
- **二進制數字：**把 k 寫成二進制（binary）時，若當中 1 的個數是偶數，第 k 項便是 0；若是奇數，便是 1。（這就是 x^k 的正負號為 $(-1)^{1\text{ 的個數，取自}k}$ 的原因：要從乘積中得出 x^k，就要選取那些在二進制中組成 k 的因子 x^{2^m}，而每一個都帶來一個負號。）

► **多項式：**$\prod_m (1 - x^{2^m})$ 的各係數，一如題目所示。

按 Thue–Morse 數列把 $0, 1, \ldots, 2^N - 1$ 分開，所得的兩個集合對於每個 $r < N$ 都有相等的 r 次方之和（圖 9.7.1）。這就是 *Prouhet–Tarry–Escott 問題*（Prouhet–Tarry–Escott problem）：找出兩個不同的整數集合，使它們的 r 次方之和在 $r = 1, \ldots, k$ 時全部相等。Prouhet 的解用了 2^{k+1} 個數；究竟最少需要多少個數，是一個著名的未解決問題。

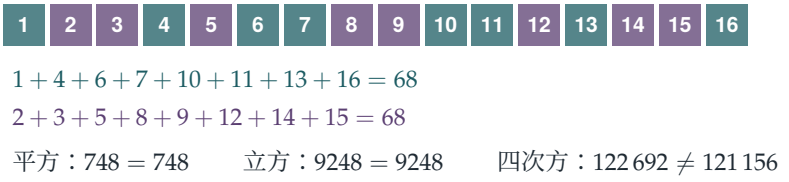


圖 9.7.1　按 Thue–Morse 數列分開的 1 至 16。兩個集合的和相等、平方和相等、立方和也相等，但四次方之和並不相等；要做到這一點，便需要 1 至 32。由左至右讀出顏色，得到的規律是藍綠–紫–紫–藍綠、紫–藍綠–藍綠–紫、...。

最公平的輪流方式

兩個小孩輪流從一堆糖果中挑選，最好的先被選走。若他們以 $ABABAB\ldots$ 的次序交替，那麼每一對糖果中較好的一顆總是歸 A，這是一種系統性的優勢。$ABBA$ 的次序較為公平，而 $ABBABAAB$ 更加公平：這正是 Thue–Morse 數列，而次方和平衡恰恰就是箇中原因。若糖果的價值由其位置的一個光滑函數給出，把該函數展開成多項式，便可看到 Thue–Morse 次序會把它的首幾項抵銷。同樣的想法曾被提議用於足球的互射十二碼（足球規則制定者自 2017 年起試行的「$ABBA$」次序），而網球決勝局的發球次序亦採用這個方式。

Thue、Morse 與 Euwe

這個數列曾多次被人發現。Prouhet 在 1851 年用過它。Axel Thue 在 1912 年研究沒有重複規律的數列時重新發現了它：它永遠不會連續三次出現同一段（即它是「無立方」(cube-free) 的）。Marston Morse 在 1921 年研究曲面上的測地線 (geodesics) 時再次發現它。而國際象棋世界冠軍尤偉（Max Euwe）在 1929 年用它證明了，當時生效的重複局面和棋規則並不能防止對局無限延續下去。時至今日，它是*自動數列*（automatic sequences）理論中的核心例子；自動數列是指由簡單機器讀取 n 的二進制數字而計算出來的數列。

思維技巧

把組合性質編碼成多項式

次方和是否平衡，關鍵在於 $+$ 號和 $-$ 號落在哪些位置。以這些正負號作爲係數的多項式，可把這個性質轉化爲一個關於根的命題。

本題從頭到尾都沒有要求你去搜尋：$x = 1$ 處的一個重根已經包辦了所有工作。利用代數來構造組合（combinatorial）對象，而不是靠反覆試驗去找，是大學組合數學中一再出現的想法。

延伸練習

1 **平移。**試做 STEP II 2025，第 8 題（載於 OCR 提供下載的 2025 年試卷，`www.ocr.org.uk/Images/750486-2025-past-papers.zip`）。該題說明怎樣把一個令和與平方和平衡的分法，透過平移其中一列再合併，變成一個更大而且連立方和也平衡的分法。

2 **平方的平方。**（同樣出自 STEP II, 2025。）集合 $A = \{1,3,4,5,9,11\}$ 和 $B = \{2,6,7,8,10\}$ 的和相等，平方和也相等。對於每個正整數 n，把 $\{n^2,(n+1)^2,\ldots,(n+11)^2\}$ 分成兩個和相等的集合。（提示：把 0 加入第二個集合。）

3 **平移不變性(translation invariance)。**證明：若兩個數集的 r 次方之和在 $r = 0,1,\ldots,k$ 時全部相等，那麼把每個元素都加上同一個數 c 所得的兩個集合亦然。為甚麼這表示題目中的集合 $\{1,\ldots,15\}$ 同樣可以換成任意 16 個連續整數？

4 **公平分糖果。**糖果的價值為 $1, \frac{1}{2}, \frac{1}{3}, \ldots, \frac{1}{8}$。比較 A 在 $ABABABAB$ 和 $ABBABAAB$ 兩種次序下所得的總值。

延伸閱讀

- **Jean-Paul Allouche & Jeffrey Shallit, “The ubiquitous Prouhet–Thue–Morse sequence”**，載於 *Sequences and Their Applications*（Springer, 1999）。綜述這個數列在各處的出現；可在網上免費閱讀。
- **Matt Parker, *Things to Make and Do in the Fourth Dimension***（Penguin）。內容包括等次方和及其他數字趣聞。
- **Joshua Cooper & Aaron Dutle, “Greedy Galois games”**，*American Mathematical Monthly*

120 (2013)。公平輪流與 Thue–Morse 數列。

9.8 沒有人拿到自己的恤衫

題目

STEP II 2016，第 12 題

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

`step.maths.org/questions/16-s2-q12`

引子

一支十一人的足球隊匆忙更衣，各人隨手抓起一件球衣就穿。沒有任何人穿上自己球衣的機會有多大？再設想那是一支五十人的大軍，甚至是一百萬人的人羣。人多了，出錯的人也多了，機會是否隨之上升？還是因為有更多人可能碰巧拿對，機會反而下降？出人意料的是，這個機會幾乎完全不變：由五個人開始，它與 $1/e \approx 0.368$ 的差距不超過半個百分點。

這是組合概率（combinatorial probability）中最古老的問題，由 Pierre de Montmort 於 1708 年提出，稱為 *problème des rencontres*（相遇問題）。本題以容斥原理（inclusion–exclusion principle）解決這個問題；容斥原理是點算互相重疊的事物的工具，而所得的答案，是 e 在數學中最令人意想不到的現身之一。

動筆之前

> **停一停，試一試** 先試 $n = 3$：列出卡 $1, 2, 3$ 的全部六種次序，數一數其中有多少種是沒有任何一張卡在自己位置上的。如果你有耐性，再試 $n = 4$（共有 24 種次序）。

當 $n = 3$ 時，只有 231 和 312 沒有任何卡在自己的位置上，所以概率是 $\frac{2}{6} = \frac{1}{3}$。當 $n = 4$ 時，24 種次序中有 9 種符合條件，概率為 $\frac{3}{8} = 0.375$。沒有任何一張卡留在原位的排列，稱為錯排（derangement）。

逐步解題

階段 1　三個事件的容斥原理

想一想　怎樣把 $A \cup B \cup C$ 寫成兩個事件的併集（union），使兩個事件的公式可以派上用場？

把 $A \cup B$ 看成單一事件，對它和 C 應用公式：

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

現在 $(A \cup B) \cap C = (A \cap C) \cup (B \cap C)$，再應用一次公式，得 $\mathrm{P}((A \cap C) \cup (B \cap C)) = \mathrm{P}(A \cap C) + \mathrm{P}(B \cap C) - \mathrm{P}(A \cap B \cap C)$。把 $\mathrm{P}(A \cup B)$ 也展開，再合併各項，便得出結果。 ■

四個事件時，規律依舊：把四個單一事件的概率相加，減去六個兩兩交集（intersection）的概率，加上四個三重交集的概率，再減去 $\mathrm{P}(A \cap B \cap C \cap D)$。

階段 2　各個概率

想一想　如果卡 i 在位置 i，其餘的卡有多少種排法？如果卡 i 和卡 j 都固定了呢？

共有 $n!$ 種機會均等的次序。如果卡 i 固定在位置 i，其餘 $n-1$ 張卡可以任意排列：共 $(n-1)!$ 種排法。所以

$$\mathrm{P}(E_i) = \frac{(n-1)!}{n!} = \frac{1}{n}, \qquad \mathrm{P}(E_i \cap E_j) = \frac{(n-2)!}{n!} = \frac{1}{n(n-1)},$$
$$\mathrm{P}(E_i \cap E_j \cap E_k) = \frac{(n-3)!}{n!} = \frac{1}{n(n-1)(n-2)}.$$

一般而言，指定的 k 張卡全部在自己位置上的概率是 $\frac{(n-k)!}{n!}$。

階段 3　加起來

想一想　在 $\mathrm{P}(E_1 \cup \cdots \cup E_n)$ 的容斥公式中，涉及 k 個事件交集的項有多少個？它們的總和是多少？

從這些事件中選出 k 個，有 $\binom{n}{k}$ 種方法，而每個交集的概率都是 $\frac{(n-k)!}{n!}$。所以各個 k 重項的總和是

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

除了 $k!$ 之外，全部約去。配上正負相間的符號，

$$\mathrm{P}(\text{最少有一張卡在自己的位置上}) = 1 - \frac{1}{2!} + \frac{1}{3!} - \cdots + (-1)^{n+1}\frac{1}{n!}. \qquad \blacksquare$$

所以出現錯排的概率是

$$d_n = 1 - 1 + \frac{1}{2!} - \frac{1}{3!} + \cdots + (-1)^n \frac{1}{n!} = \sum_{k=0}^{n} \frac{(-1)^k}{k!}.$$

階段 4 恰好一張

想一想 要恰好有一張卡在原位，是哪一張卡？其餘的卡又必須怎樣？

先選定哪一張卡在自己的位置上（n 種選法）。卡 i 在原位的概率是 $\frac{1}{n}$；在這個條件下，其餘 $n-1$ 張卡必須在它們之間形成一個錯排，其概率為 d_{n-1}。所以

$$\mathrm{P}(\text{恰好一張}) = n \cdot \frac{1}{n} \cdot d_{n-1} = \sum_{k=0}^{n-1} \frac{(-1)^k}{k!}. \qquad \blacksquare$$

回顧

- **先多算，再修正。**把各個 $\mathrm{P}(E_i)$ 相加，會把有兩張卡固定的排列算了兩次；減去兩兩交集的項可以修正，但對三張卡固定的情況又減得太多；如此類推。
- **對稱性令各項相等。**每個 k 重交集的概率都相同，所以公式的每一層都是一個數乘以一個二項式係數（binomial coefficient）。
- **化為較小的情況。**「恰好一張」變成了 $n-1$ 張卡的錯排。

背後的數學

e 從何而來

級數 $\sum(-1)^k/k!$ 就是指數級數（exponential series）$e^x = \sum x^k/k!$ 在 $x = -1$ 時的值。所以當 $n \to \infty$ 時，

$$d_n \to e^{-1} \approx 0.3679,$$

而恰好有一張卡對上的概率也趨向同一個極限。收斂（convergence）的速度快得驚人（圖 9.8.1）：在 $1/n!$ 一項停下來所產生的誤差小於 $1/(n+1)!$，所以在 $n = 7$ 時，答案已經與 $1/e$ 吻合至小數點後四位。事實上，n 件物件的錯排數目 $n!\,d_n$ 恰好是最接近 $n!/e$ 的整數。

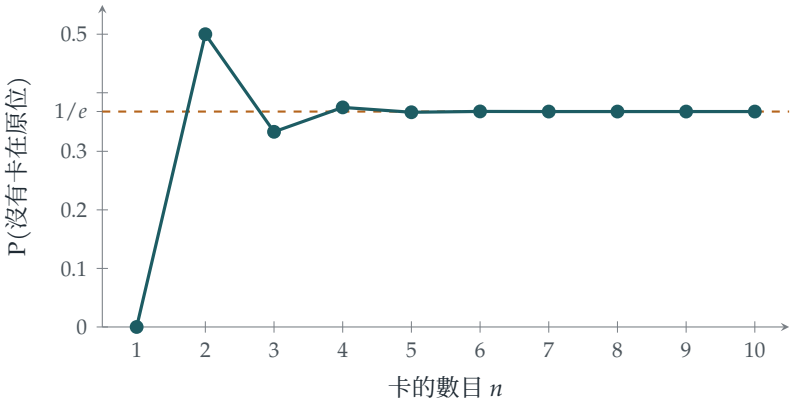


圖 9.8.1　把 n 張卡隨機洗勻後，沒有任何一張卡留在自己位置上的概率。它在 $1/e$ 的上下交替出現，而由 $n = 6$ 開始，在這個比例尺下已經與 $1/e$ 無法分辨。

對上的數目服從泊松分佈

結論還不止於此。設 N 為在自己位置上的卡的數目。那麼對每個 k，都有 $\mathrm{P}(N = k) \to \frac{e^{-1}}{k!}$：在極限之下，對上的數目服從平均值為 1 的泊松分佈（Poisson distribution）。平均值為甚麼是 1，不用任何計算也看得出來：N 是 $E_1, \ldots, E_n$ 的指示變量（indicator）之和，每個的期望值（expectation）都是 $\frac{1}{n}$，而即使這些事件並非獨立，期望值仍然可以相加。STEP III 2004 Q12（見練習）證明方差（variance）同樣是 1。每當有許多機會很微、又近乎獨立的事件時，泊松分佈就會出現；錯排是一個經典例子，也是第 12 章 Section 12.11 的主題。

Montmort 的 *treize* 遊戲

Pierre Rémond de Montmort 是一位法國貴族，他在著作 *Essay d'analyse sur les jeux de hazard*（1708）中研究各種博彩遊戲。在 *treize*（十三）這個遊戲中，玩家派出十三張牌，同時口中數着「一、二、三……」，只要有一次派出的牌與所數的數字相同，便算勝出。Montmort 求出了勝出的概率，並在與尼古拉·伯努利（Nicolaus Bernoulli）的通信中，於 1713 年或之前確立了一般公式。歐拉（Euler）在 1753 年重新發現了這條公式。容斥原理其後由棣莫弗（Abraham de Moivre）加以形式化，而它的一般形式則由西爾維斯特（James Joseph Sylvester）在十九世紀建立。

思維技巧

容斥原理

要點算互相重疊的集合的併集，先把各集合的大小相加，減去兩兩重疊的部分，加上三重重疊的部分，如此類推。

當對稱性令同一大小的重疊部分全部相等時，這個方法最見威力，本題正是如此。在大學裏，它會在組合數學（combinatorics）、概率論和數論（number theory）中出現；在數論中，它用來點算不能被一系列質數中任何一個整除的整數，並由此引出歐拉函數（Euler's totient function）和默比烏斯函數（Möbius function）。

延伸練習

1 **球衣。**試做 STEP III 2004，第 12 題（`step.maths.org/questions/04-s3-q12`）。這道題用指示變量點算穿上自己球衣的球員人數，求出平均值和方差，並把結果與泊松分佈比較。

2 **遞推關係。**設 D_n 為 n 件物件的錯排數目。考慮卡 1 去了哪個位置，以及原本在該位置的卡是否去了位置 1，證明 $D_n = (n-1)(D_{n-1} + D_{n-2})$。用 $D_3 = 2$、$D_4 = 9$ 驗算。

3 **恰好 k 張。**證明恰好有 k 張卡在自己位置上的概率是 $\frac{1}{k!}d_{n-k}$，並且它趨向 $\frac{e^{-1}}{k!}$。

4 **互質的整數。**用容斥原理點算由 1 至 30 的整數中，不能被 2、3 和 5 任何一個整除的有多少個。你的答案與 $30\left(1-\frac{1}{2}\right)\left(1-\frac{1}{3}\right)\left(1-\frac{1}{5}\right)$ 有甚麼關係？

延伸閱讀

- **Geoffrey Grimmett & Dominic Welsh, *Probability: An Introduction***（Oxford University Press）。大學一年級教科書，涵蓋容斥原理和配對問題。
- **William Feller, *An Introduction to Probability Theory and Its Applications*, Vol. 1**（Wiley），第 IV 章。經典之作，載有配對問題的許多變化。
- **Miklós Bóna, *A Walk Through Combinatorics***（World Scientific），關於容斥原理的部分。

9.9 生日問題

題目

STEP II 2007，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/07-s2-q13`

引子

一個房間裏要有多少人，才會令其中兩人同一天生日的機會大於一半？大多數人會猜 180 左右，即 365 的一半。答案卻是 23。在一班三十人的課堂裏，有人同一天生日的概率（probability）是 70%；在五十人的派對上，則是 97%。

這個結果令人意外，是因為我們把兩個不同的問題混為一談：「有沒有人與我同一天生日？」（這確實需要大約 253 人）和「有沒有任何兩個人同一天生日？」。23 個人可以組成 $\binom{23}{2} = 253$ 對，而每一對都是一次出現相同生日的機會。這道題把這個直覺變得精確，而同樣的數學亦決定了保護你銀行戶口的密碼技術有多安全。

動筆之前

想一想 為甚麼計算沒有兩位賓客同一天生日的概率會比較容易？請賓客排成一行：第二位避開第一位生日的機會是多少？第三位同時避開前兩位的機會又是多少？

「至少有兩人相同」包含許多互相重疊的情況（兩人相同、三人相同、兩對不同的人各自相同……）。它的補事件（complement）「全部不同」卻是單一而簡潔的事件。第二位賓客避開第一位生日的概率是 $\frac{364}{365}$；在此條件下，第三位同時避開前兩位的概率是 $\frac{363}{365}$；如此類推。

逐步解題

階段 1　一個有用的近似

想一想　$y = e^{-x}$ 在 $x = 0$ 處的切線（tangent line）是甚麼？

由於 $e^{-x} = 1 - x + \frac{x^2}{2} - \cdots$，當 x 很小時，我們有 $e^{-x} \approx 1 - x$。取 $x = r/n$，得

$$\frac{n-r}{n} = 1 - \frac{r}{n} \approx e^{-r/n}. \qquad \blacksquare$$

誤差約為 $\frac{r^2}{2n^2}$，當 r 遠小於 n 時，這個誤差很小。

階段 2　有人同一天生日的概率

k 個生日全部不同的概率是

$$\frac{365}{365} \cdot \frac{364}{365} \cdot \frac{363}{365} \cdots \frac{365-(k-1)}{365} = \prod_{r=0}^{k-1} \frac{365-r}{365}.$$

對每個因子使用上述近似（approximation），再利用指數函數相乘時指數相加這個性質，

$$\prod_{r=0}^{k-1} e^{-r/365} = \exp\Big(-\frac{0+1+\cdots+(k-1)}{365}\Big) = \exp\Big(-\frac{k(k-1)}{730}\Big).$$

因此，至少有一對人同一天生日的概率約為 $1 - e^{-k(k-1)/730}$。 $\blacksquare$

想一想　看看指數：$\frac{k(k-1)}{730} = \binom{k}{2}\Big/365$。用文字表達，這說明了甚麼？

它是賓客的配對數目，乘以某一對人同一天生日的概率 $\frac{1}{365}$：也就是生日相同的配對數目的期望值（expected number）。這個近似告訴我們，*沒有任何配對相同的機會是* $e^{-(\text{相同配對的期望數目})}$，這又一次透露出泊松分佈（Poisson distribution）的影子，正如 Section 9.8 所見。

階段 3　二十三

我們需要 $1 - e^{-k(k-1)/730} \geq \frac{1}{2}$，即 $e^{-k(k-1)/730} \leq \frac{1}{2}$，亦即

$$\frac{k(k-1)}{730} \geq \ln 2 \approx \frac{253}{365}, \qquad \text{即} \qquad k(k-1) \geq 506.$$

由於 $22 \times 21 = 462$ 而 $23 \times 22 = 506$，最小的值是 $k = 23$。 $\blacksquare$

（23 位賓客時的精確概率是 0.507，22 位賓客時是 0.476，所以這個近似給出了正確答案。）

階段 4 與主人同一天生日

想一想 對每位賓客來說，不與主人同一天生日的概率是多少？在這方面，各位賓客之間是否互相獨立（independent）？

每位賓客各自獨立地以概率 $\frac{364}{365} \approx e^{-1/365}$ 避開主人的生日。因此 k 位賓客中無人與主人同一天生日的概率是 $\left(\frac{364}{365}\right)^k \approx e^{-k/365}$，而我們需要

$$e^{-k/365} \le \tfrac{1}{2}, \qquad \text{即} \qquad k \ge 365 \ln 2 \approx 253.$$

所以必須有 253 位賓客。 ■

這個對比正是重點所在：要與某一個指定的人生日相同，需要 253 位賓客；但要有某一對人生日相同，只需 23 人，因為 23 個人可以組成 253 對。

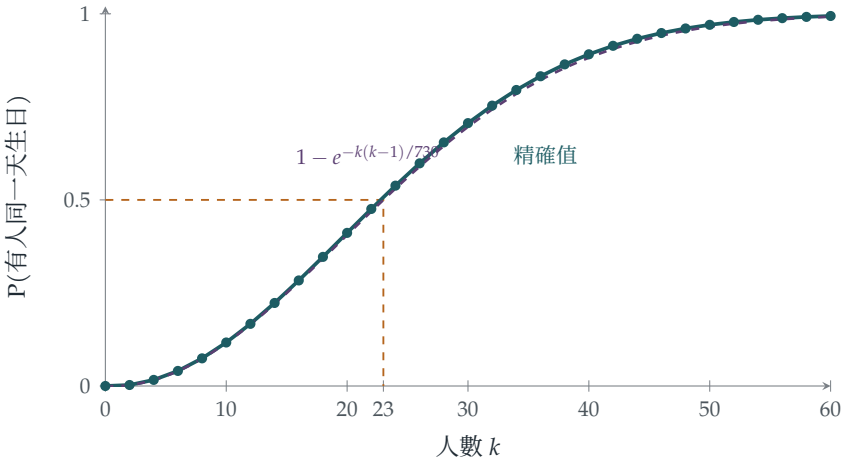


圖 9.9.1 k 個人中至少有兩人同一天生日的概率：精確值（藍綠色）與近似值 $1 - e^{-k(k-1)/730}$（紫色虛線）。兩者非常接近，並且都在 $k = 22$ 與 $k = 23$ 之間越過 $\frac{1}{2}$。

回顧

- **利用補事件。**「全部不同」只是一個乘積；「至少有一對相同」卻是一團亂麻。
- **先線性化，再取指數。**把每個因子 $1 - x$ 換成 e^{-x}，乘積就變成了和。
- **數配對。**指數正是配對的數目乘以一對人相同的機會。$k(k-1)$ 中的平方，就是這個悖論的全部解釋。

背後的數學

平方根定律

若把 365 天換成 n 個等可能的日子，同樣的論證給出

$$\mathrm{P}(\text{有人相同，人數為}k) \approx 1 - e^{-k^2/(2n)},$$

所以只要 k 達到 $\sqrt{n}$ 左右，而不是 n，就很可能出現相同的情況。更準確地說，概率在 $k \approx 1.18\sqrt{n}$ 時達到 $\frac{1}{2}$，而在首次出現重複之前你需要遇到的人數，期望值約為 $\sqrt{\pi n/2}$（當 $n = 365$ 時，約為 24.6）。這條*平方根定律*（square-root law）正是巧合如此常見的原因：在任何一大批事件之中，總很可能有某一對在某方面吻合，儘管任何一個指定的吻合都不大可能發生。

生日攻擊

電腦會計算檔案的*雜湊值*（hash），以檢查檔案有否被竄改：那是一個指紋，可能的取值有（比方說）$n = 2^{128}$ 個。偽造者若想得到兩份雜湊值相同的文件（一份無害的合約和一份欺詐的合約），並不需要試 2^{128} 份文件。根據生日問題的論證，為每份文件各自產生大約 $\sqrt{2^{128}} = 2^{64}$ 個變體，就很可能得到一對相同的雜湊值。這種*生日攻擊*（birthday attack）正是現代雜湊函數採用 256 位元或以上的原因：有一半的位元因平方根定律而白白損失。類似的計算亦決定了隨機識別碼必須有多長，才能令兩部電腦永遠不會碰巧選中同一個。

誰最先提出？

生日問題通常歸功於馮·米塞斯（Richard von Mises），他在 1939 年討論過這個問題，不過它在此之前已經流傳。據說 Harold Davenport 亦更早知道這個問題，只是覺得它太簡單而沒有發表。按 Warren Weaver 的說法，他在第二次世界大戰期間喜歡在軍官的晚宴上示範這個問題：他打賭在場二十二人之中有兩人同一天生日，結果輸了，直至一位女侍應透露她與其中一位軍官同一天生日。

思維技巧

計算補事件；小心地近似

當「至少一個」很複雜時，就去求「一個也沒有」的概率。當乘積難以處理時，把每個因子近似爲指數函數，然後把指數相加。

這兩個做法在大學裏都是標準手法。第二個，即 $1 - x \approx e^{-x}$，是泊松近似理論的起點，也是漸近方法（asymptotic methods）的起點；後者用來估計那些複雜得無法精確計算的量。

延伸練習

1 **首次重複。**試做 STEP II 2014，第 13 題（`step.maths.org/questions/14-s2-q13`）。這道題研究隨機數產生器（random number generator）首次出現重複的值，其實是換了個面貌的生日問題，並把有關概率化成兩個出人意表的恆等式（identities）。

2 **三人相同。**大約需要多少人，才會令某三人同一天生日的機會大於一半？（提示：同一天生日的三人組數目的期望值是 $\binom{k}{3}/365^2$。令它約等於 $\ln 2$。）真正的答案是 88。

3 **差之毫釐。**需要多少人，才會令兩人的生日在同一天或相鄰兩天的機會達到一半？

4 **不均勻的生日。**現實中的生日並非均勻（uniform）分佈。證明若各日的概率為 $p_1, \ldots, p_{365}$，則兩個人同一天生日的機會是 $\sum p_i^2$，並且當所有 p_i 相等時這個值最小。（所以不均勻只會令相同的情況更容易出現。）

延伸閱讀

- **William Feller, *An Introduction to Probability Theory and Its Applications*, Vol. 1**（Wiley），第 II 章。生日問題及其各種近似。
- **Persi Diaconis & Frederick Mosteller,「Methods for studying coincidences」**，*Journal of the American Statistical Association* 84 (1989)。解釋巧合為何常見；非常易讀。
- **David Hand, *The Improbability Principle***（Bantam）。平方根定律，以及不大可能的事情為何時刻都在發生的其他原因。

9.10 總能找贖

題目

STEP I 2011，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/11-s1-q12`

引子

這個規律實在叫人難以抗拒：$\frac{m}{m+1}$、$\frac{m-1}{m+1}$、$\frac{m-2}{m+1}$。若有 n 個人拿着 £2 硬幣，答案想必就是 $\frac{m-n+1}{m+1}$。事實正是如此，而最巧妙的證明（proof）用上了組合數學（combinatorics）中數一數二優雅的想法：*反射原理*（reflection principle）。它把圖形的一部分沿一條直線反射，從而將一個困難的計數問題（從不缺零錢找贖的隊伍有多少種）變成一個容易的計數問題。

同一個問題有許多不同的面貌：點票過程中勝出者從未落後的選舉（1887 年的伯特蘭投票問題（Bertrand's ballot problem））、永不破產的賭徒、正確配對的括號。當 $m = n$ 時，數出來的就是*卡塔蘭數*（Catalan numbers），這大概是組合數學中除二項式係數（binomial coefficients）以外最常出現的數列（sequence）。

動筆之前

想一想 每位拿 £1 的顧客會給你一枚硬幣；每位拿 £2 的顧客會取走一枚（作為找贖）。你甚麼時候會無法繼續？

留意你手上 £1 硬幣的數目。每來一位 £1 顧客，數目加 1；每來一位 £2 顧客，數目減 1。你無法繼續的情況，恰好就是一位 £2 顧客來到而你手上沒有硬幣：數目會變成 -1。所以你能成功，當且僅當（if and only if）在隊伍中的每一刻，到目前為止的 £1 顧客人數都不少於到目前為止的 £2 顧客人數。

逐步解題

階段 1　一枚 £2 硬幣

當 $n = 1$ 時，只有 £2 顧客排第一才會失敗。£2 顧客排在 $m+1$ 個位置中任何一個的可能性均等，所以

$$\mathrm{P}(\text{成功}) = 1 - \frac{1}{m+1} = \frac{m}{m+1}. \qquad \blacksquare$$

階段 2　兩枚 £2 硬幣

> **想一想**　第一個人必須拿着 £1 硬幣。第二個人可以出現甚麼情況？在每種情況下，你餘下面對的是甚麼局面？

第一個人拿着 £1 硬幣的概率（probability）是 $\frac{m}{m+2}$；否則你立即失敗。這時你手上有一枚硬幣，而餘下的隊伍有 $m-1$ 個 £1 的人和兩個 £2 的人。

- 若第二個人拿着 £1 硬幣（概率為 $\frac{m-1}{m+1}$），你手上便有兩枚硬幣，無論兩枚 £2 硬幣何時出現都足夠找贖。成功是必然的。
- 若第二個人拿着 £2 硬幣（概率為 $\frac{2}{m+1}$），你手上又變回沒有硬幣，餘下 $m-1$ 個 £1 的人和一個 £2 的人。根據第 (i) 部分，以 $m-1$ 代替 m，你成功的概率是 $\frac{m-1}{m}$。

所以

$$\mathrm{P}(\text{成功}) = \frac{m}{m+2}\left(\frac{m-1}{m+1} + \frac{2}{m+1}\cdot\frac{m-1}{m}\right) = \frac{m}{m+2}\cdot\frac{m-1}{m+1}\cdot\frac{m+2}{m} = \frac{m-1}{m+1}. \qquad \blacksquare$$

> **停一停，試一試**　試用同樣的方法處理 $n = 3$。這是行得通的，但情況會愈分愈多。有沒有方法可以一次過數出所有成功的隊伍？

階段 3　把隊伍看成路徑

> **想一想**　以你手上的硬幣數目對已服務的人數作圖。一條隊伍看起來是甚麼樣子？一條成功的隊伍又是甚麼樣子？

一條隊伍就是一條路徑，由 m 步上升（£1）和 n 步下降（£2）組成，從高度 0 出發，終於高度 $m-n$。這樣的路徑共有 $\binom{m+n}{n}$ 條，每條出現的可能性均等。一條隊伍成功，恰好就是它的路徑從不低於 0。所以我們需要數出壞路徑，即那些觸及高度 -1 的路徑。

階段 4　反射

想一想 取一條壞路徑。找出它第一次觸及 -1 的時刻，把路徑在這一刻之前的部分沿直線 $y=-1$ 反射。新的路徑從哪裏開始，又在哪裏結束？

反射後的路徑從 -2 出發（即 0 對直線 $y=-1$ 的鏡像），終點與原來的路徑相同，即 $m-n$（圖 9.10.1）。反過來說，每條由 -2 到 $m-n$ 的路徑都必定在某處穿過 $y=-1$（因為 $m\geq n$，它起點在這條線之下，終點在這條線之上）；把它第一次到達這條線之前的部分反射，便得回一條由 0 出發的壞路徑。所以由 0 出發的壞路徑，與由 -2 到 $m-n$ 的*所有*路徑恰好一一對應。

一條以 $m+n$ 步由 -2 走到 $m-n$ 的路徑共上升了 $m-n+2$，所以它有 $m+1$ 步上升和 $n-1$ 步下降：這樣的路徑有 $\binom{m+n}{n-1}$ 條。因此

$$\mathrm{P}(\text{失敗})=\frac{\binom{m+n}{n-1}}{\binom{m+n}{n}}=\frac{n!\,m!}{(n-1)!\,(m+1)!}=\frac{n}{m+1},$$

而

$$\mathrm{P}(\text{成功})=1-\frac{n}{m+1}=\frac{m-n+1}{m+1}.$$

當 $n=3$ 時，這就是 $\frac{m-2}{m+1}$。 ■

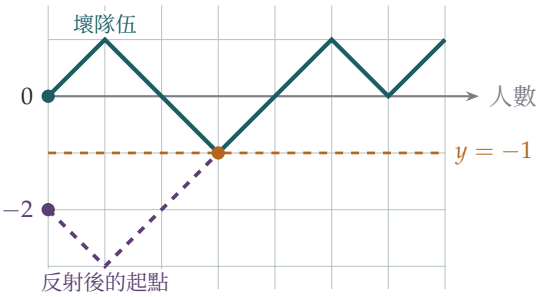


圖 9.10.1 一條有 $m=4$ 枚 £1 硬幣和 $n=3$ 枚 £2 硬幣的隊伍，畫成一條路徑（藍綠色）。它在第三個人時首次到達 -1（琥珀色），這時你無法找贖。把路徑在這一點之前的部分沿直線 $y=-1$ 反射，便得到一條由 -2 出發（紫色虛線）而終點相同的路徑。壞隊伍與由 -2 出發的路徑之間的這種配對，就是反射原理。

回顧

- **找出合適的狀態。**唯一要緊的是手上硬幣的數目，這把隊伍變成了一次遊走。
- **以最初幾步為條件。**第 (i) 和第 (ii) 部分的做法，是考慮最初幾個人的情況，從而化約為較小的情形。

- **用雙射（bijection）數出補集。**壞路徑難以直接數出，但反射把它們與一個容易數算的集合配對起來。

背後的數學

卡塔蘭數

當 $m = n$（£1 硬幣剛好足夠）時，成功隊伍的數目是

$$\binom{2n}{n} - \binom{2n}{n-1} = \frac{1}{n+1}\binom{2n}{n},$$

即第 n 個卡塔蘭數 C_n：$1, 1, 2, 5, 14, 42, 132, \ldots$。Richard Stanley 的著作 *Catalan Numbers* 列出了超過二百種由卡塔蘭數計算數目的事物。其中包括：把 n 對括號正確排列的方法數目（「(())()」是允許的，「())(()」則不允許，理由正是上述的硬幣論證，只要把「(」看成 £1、「)」看成 £2）、用互不相交的對角線把一個 $n+2$ 邊形分割成三角形的方法數目，以及為 $n+1$ 個因子的乘積加上括號的方法數目。

隨機遊走

隊伍就是隨機遊走（random walk），而反射原理是研究隨機遊走的基本工具之一。舉例說，它可以證明由 0 出發的公平隨機遊走必定終會返回 0，也可以證明它在最初 $2n$ 步一直保持嚴格為正的概率是 $\frac{1}{2}\binom{2n}{n}2^{-2n} \approx \frac{1}{2\sqrt{\pi n}}$。在連續的形式下，把同樣的反射論證應用於布朗運動（Brownian motion），便可得出其最大值的分佈（distribution）。這是某些金融期權定價的基礎，這類期權的回報取決於股價有否越過某個界限。

伯特蘭投票問題

1887 年，約瑟夫·伯特蘭（Joseph Bertrand）提出：在一場選舉中，候選人 A 得 a 票，B 得 $b < a$ 票，那麼 A 在整個點票過程中一直嚴格領先的概率是多少？答案是 $\frac{a-b}{a+b}$，同年由 Désiré André 證明，他的論證通常被視為反射原理（不過歷史學家指出，André 實際的證明用的是一個稍有不同的雙射）。我們的抽獎券問題是一個容許平手的變體：成功的意思是 £1 的數目從不低於 £2 的數目。其中一道練習會由我們的答案推導出伯特蘭的答案。

思維技巧

雙射計數

要數算一個集合，就找出它與一個你懂得數算的集合之間的一一對應（*one-to-one correspondence*）。

反射把壞隊伍與由另一個起點出發、不受限制的路徑配對起來。雙射是組合數學中最令人滿足的證明，因為它們解釋了兩個數爲何相等，而不只是證明它們相等。在大學裏，你會在組合數學中處處遇到雙射證明，也會在研究隨機遊走和布朗運動時再次遇上反射原理。

延伸練習

1 **分情況處理三枚。**仿照第 (ii) 部分，以隊伍中最初幾個人為條件來證明第 (iii) 部分，看看反射原理省卻了多少工夫。

2 **伯特蘭投票問題。**在一場選舉中，A 得 a 票，B 得 $b < a$ 票。證明 A 在整個點票過程中一直嚴格領先的概率是 $\frac{a-b}{a+b}$。（提示：第一票必須投給 A；此後，A 落後於 B 的票數絕不可多於…。）

3 **括號。**列出三對括號的 5 種正確排列，以及把五邊形分割成三角形的 5 種方法。你能找出兩者之間的一個雙射嗎？

4 **一張 £5 紙幣。**一條隊伍中有 m 個人拿着 £1 硬幣，n 個人拿着 £5 紙幣，而一張票的價錢是 £1。解釋為何反射原理不再給出簡單的答案。（這個推廣困難得多。）

延伸閱讀

- **William Feller, *An Introduction to Probability Theory and Its Applications*, Vol. 1**（Wiley），第 III 章。有關隨機遊走和反射原理的經典論述。
- **Richard Stanley, *Catalan Numbers***（Cambridge University Press）。超過二百種詮釋，並附練習。
- **Marc Renault, "Lost (and found) in translation: André's actual method and its application to the generalized ballot problem"**，*American Mathematical Monthly* 115 (2008)。反射原理的歷史。

9.11 集齊一套

題目

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`

引子

一本足球貼紙簿要集齊 600 張貼紙，貼紙每次隨機發售一張。要貼滿整本貼紙簿，你預計要買多少張貼紙？最初幾百張很快到手，因為幾乎每一張都是新的。但貼紙簿愈貼愈滿，新貼紙就愈來愈罕有，單是最後一張，平均就要買 600 次。總數大約是 4200 張，是貼紙簿大小的七倍。

這就是*集券問題*（coupon collector's problem）。它的答案結合了兩個想法：等候一次成功所需的時間，以及調和級數（harmonic series）$1 + \frac{1}{2} + \frac{1}{3} + \cdots$，後者的增長與 $\ln n$ 相若。你在 Section 2.3 見過調和級數發散（diverge）；在這裏，它那緩慢的增長正正就是集齊一套收藏品所要付出的代價。

動筆之前

想一想 如果每包各自獨立地以概率 p 開出一張新卡，那麼直至開出第一張新卡為止（連這一包計算在內），你開包數目的期望值（expectation）是多少？

開包數目 X 服從*幾何分佈*（geometric distribution）：$\mathrm{P}(X = k) = (1-p)^{k-1}p$。它的期望值是 $\frac{1}{p}$。（有一個快捷的看法：$\mathrm{E}(X) = 1 + (1-p)\mathrm{E}(X)$，因為你先開一包，然後有 $1-p$ 的概率回到起點。）如果有三分之一的機會開出新卡，你預計要等三包。

逐步解題

階段 1　等候一張新卡

手上已有 r 張不同的卡時，一包開出新卡的概率是 $\frac{n-r}{n}$，與之前各包獨立。因此，直至開出下一張新卡為止，所需包數的期望值是

$$\frac{1}{(n-r)/n} = \frac{n}{n-r}. \qquad \blacksquare$$

階段 2　比較和與積分

> **想一想**　在區間 $r \le x \le r+1$ 上，$\frac{1}{x}$ 介乎哪兩個值之間？這對曲線下的面積有甚麼啟示？

由於 $\frac{1}{x}$ 是遞減的，在 $[r, r+1]$ 上有 $\frac{1}{r+1} \le \frac{1}{x} \le \frac{1}{r}$。在長度為 1 的區間上積分，得

$$\frac{1}{r+1} \le \int_r^{r+1} \frac{\mathrm{d}x}{x} \le \frac{1}{r}. \qquad \blacksquare$$

用圖來看（圖 9.11.1）：曲線下的面積夾在高度為 $\frac{1}{r+1}$ 的長方形與高度為 $\frac{1}{r}$ 的長方形之間。把這些不等式由 $r = 1$ 加到 $r = n-1$，各個積分便連接成 $\int_1^n \frac{\mathrm{d}x}{x} = \ln n$，於是

$$\frac{1}{2} + \frac{1}{3} + \cdots + \frac{1}{n} \le \ln n \le 1 + \frac{1}{2} + \cdots + \frac{1}{n-1}. \qquad \blacksquare$$

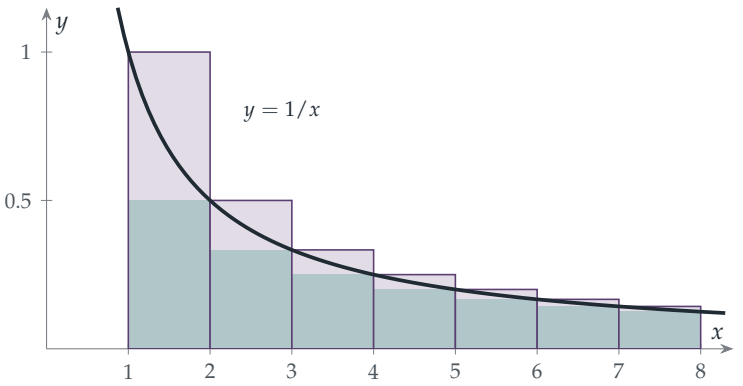


圖 9.11.1　在每個區間 $[r, r+1]$ 上，$y = 1/x$ 之下的面積介乎藍綠色長方形（高度 $\frac{1}{r+1}$）與較大的紫紅色長方形（高度 $\frac{1}{r}$）之間。全部加起來，$\ln n$ 便介乎 $\frac{1}{2} + \cdots + \frac{1}{n}$ 與 $1 + \frac{1}{2} + \cdots + \frac{1}{n-1}$ 之間。曲線與紫紅色長方形之間的空隙，加起來是一個有限的總和：歐拉常數（Euler's constant）$\gamma \approx 0.577$。

階段 3　集齊全套

想一想 總包數是等候第一張卡、第二張新卡……各段等候時間之和。它的期望值是多少？

設 X_r 為由 r 張不同的卡增至 $r+1$ 張所需的包數。總數是 $X_0 + X_1 + \cdots + X_{n-1}$，而期望值可以相加，所以

$$\mathrm{E}(\text{總數}) = \sum_{r=0}^{n-1} \frac{n}{n-r} = n\Big(\frac{1}{n} + \frac{1}{n-1} + \cdots + 1\Big) = nH_n,$$

其中 $H_n = 1 + \frac{1}{2} + \cdots + \frac{1}{n}$ 是第 n 個調和數（harmonic number）。由第 2 步，$\ln n + \frac{1}{n} \le H_n \le \ln n + 1$，所以當 n 很大時 H_n 約為 $\ln n$，而我預計大約要吃掉

$$n \ln n \text{ 包.} \qquad \blacksquare$$

回顧

- **把漫長的等候分成多個階段。**總時間拆成 n 段等候時間，每一段都屬於簡單的幾何分佈。
- **期望值可以相加。**我們從頭到尾都不需要總數的分佈，只需要各部分的期望值。
- **比較和與積分。**那幅圖把一個我們無法求值的和化成一個對數，誤差至多為 1。

背後的數學

這個估計有多準確？

這個比較還可以再精確一些。圖 9.11.1 中曲線與上方長方形之間的陰影空隙，總面積是有限的，所以 $H_n - \ln n$ 趨向一個極限，

$$\gamma = \lim_{n\to\infty} (H_n - \ln n) = 0.5772156649\ldots,$$

即歐拉常數。（沒有人知道 γ 是否有理數。）因此包數的期望值是 $n\ln n + \gamma n + \frac{1}{2} + \cdots$。以 600 張貼紙的貼紙簿為例，$600 \times (\ln 600 + 0.577) \approx 4185$，這就是引子中所引述的數字。

時間都花到哪裏去了？各個階段極不平均。集齊全套的前一半只需大約 $n\ln 2 \approx 0.69n$ 包，而單是最後一張卡，平均就要 n 包。大部分工夫都花在最後幾張卡上，這正是收藏者要互相交換重複卡的原因。

集中現象

期望值只是個開始：總數的變動有多大？總數的方差（variance）是 $\sum nr/(n-r)^2$，約為 $\frac{\pi^2}{6}n^2$（Section 2.4 中的和 $\sum 1/k^2 = \pi^2/6$ 在這裏出現了）。所以標準差（standard deviation）約為 $1.28n$，當 n 很大時遠小於平均值 $n\ln n$：總數集中在其平均值附近。艾狄胥（Paul Erdős）和雷尼（Alfréd Rényi）在 1961 年證明了一個更精確的結果：總數減去 $n\ln n$ 再除以 n，會有一個極限分佈（limiting distribution），即耿貝爾分佈（Gumbel distribution），這個分佈也描述了許多隨機量之中最大的一個（Section 12.25）。

無處不在的集券問題

這個問題可追溯至棣莫弗（Abraham de Moivre，1712 年）和拉普拉斯（Laplace）。如今它以許多不同的面貌出現。要檢查一個系統的每個組件，需要做多少次隨機測試？在一個網絡上進行隨機漫步（random walk），要多久才會走遍每個節點？電腦的雜湊表（hash table）要探查多少次？每一種情況的答案都涉及 $n\ln n$。在生物學中，同樣的計算可用來估計要有多少個定序「讀段」（reads）才能覆蓋整個基因組。

思維技巧

期望值的線性性質（linearity of expectation）

和的期望值等於期望值之和，不論各部分是否獨立。把一個複雜的隨機量拆成簡單的部分。

這是概率論中最有用的事實之一。在這裏，它把集券問題簡化為 n 段幾何等候；在 Section 9.8 中，它不用任何計算就給出了配對數目的期望值。到了大學，同一個想法推動着演算法分析，以及組合數學中的概率方法（probabilistic method）。

延伸練習

1 **半套。**證明集齊 $\frac{n}{2}$ 張不同的卡（n 為偶數）所需包數的期望值是 $n(H_n - H_{n/2})$，並且這個值約為 $n\ln 2$。

2 **不同的卡。**證明開了 k 包之後，你手上不同的卡的數目的期望值是 $n\big(1-(1-\frac{1}{n})^k\big)$。（提示：對每一張卡，求出你擁有它的概率，然後相加。）大約要開多少包，才能得到全套的 95%？

3 **兩位收藏者。**兩個小孩收集同一套卡，並共用所開的各包，因此對其中一人來說是重複的卡，對另一人來說可能是新的。證明集齊兩套所需包數的期望值約為 $n(\ln n + \ln\ln n)$。（這題較難；先說服自己這個數目遠小於 $2n\ln n$。）

4 **更精確的界。**改用梯形法則（trapezium rule）的圖像而不用長方形，證明 $H_n - \ln n$ 是遞減的，而 $H_n - \ln n - \frac{1}{2n}$ 是遞增的。由較小的 n 值推導出 $0.5 < \gamma < 0.7$。

延伸閱讀

- **Geoffrey Grimmett & David Stirzaker, *One Thousand Exercises in Probability***（Oxford University Press）。集券問題的許多變化。
- **Julian Havil, *Gamma: Exploring Euler's Constant***（Princeton University Press）。關於 γ 和調和級數的一切，寫得深入淺出。
- **Michael Mitzenmacher & Eli Upfal, *Probability and Computing***（Cambridge University Press）。集券問題及其在電腦科學中的應用。

9.12 隨機子集

題目

STEP III 2019，第 12 題

繼續閱讀之前，請先看題目（原題為英文）。題目載於 OCR 提供下載的 2019 年 STEP 試卷（zip 檔案）內的 STEP III 試卷：

`www.ocr.org.uk/Images/696590-2019-past-papers.zip`

引子

完全隨機地選出 $\{1,2,\ldots,100\}$ 的兩個子集（subset）。它們沒有任何共同元素的可能性有多大？直覺也許會告訴你：「相當大：大多數成對的集合都不相同」。事實上，這個概率是 $\left(\frac{3}{4}\right)^{100}$，約為 3×10^{-13}。隨機集合很大（大約佔全部元素的一半），所以它們幾乎總會重疊。

這道題很短，當中唯一的想法卻很有威力：不要把集合當作整體來想，而要每次只想一個元素。一個隨機子集不過是 n 次獨立（independent）的擲硬幣，每個元素擲一次；於是，凡是關於隨機子集的問題，都變成關於這些擲硬幣結果的問題。

動筆之前

? **想一想** 怎樣用 n 條是非題的一串答案，來描述 $\{1,\ldots,n\}$ 的一個子集？

對每個元素 i，回答「i 在集合裏嗎？」這條問題。子集決定了答案，答案也決定了子集。每個元素有 2 個選擇，所以一共有 2^n 個子集。均勻隨機（uniformly at random）地選一個子集，就等於獨立地擲一枚公平硬幣來回答每條問題。

逐步解題

階段 1　一個元素

按照擲硬幣的描述，$\mathrm{P}(1 \in A_1) = \frac{1}{2}$。事實上，$2^n$ 個子集中恰好有一半包含 1：把每個不包含 1 的子集，與加入 1 之後所得的子集配成一對即可。

階段 2　不相交的集合

想一想　要使 $A_1 \cap A_2 = \varnothing$，每個元素 i 必須滿足甚麼條件？就單一個元素而言，這件事的概率是多少？為甚麼可以把各個元素的概率相乘？

A_1 與 A_2 不相交（disjoint），當且僅當沒有元素同時屬於兩者。對單一個元素 i 來說，有四種等可能的情況：兩個集合都不在、只在 A_1、只在 A_2、兩個都在（圖 9.12.1 左）。只有最後一種是禁止的，所以 i 不同時屬於兩者的概率是 $\frac{3}{4}$。不同的元素由不同而且獨立的擲硬幣結果決定，所以

$$\mathrm{P}(A_1 \cap A_2 = \varnothing) = \left(\frac{3}{4}\right)^n. \qquad \blacksquare$$

對三個集合而言，元素 i 同時屬於三者的概率是 $\frac{1}{8}$，所以 $\mathrm{P}(A_1 \cap A_2 \cap A_3 = \varnothing) = \left(\frac{7}{8}\right)^n$。一般而言，

$$\mathrm{P}(A_1 \cap \cdots \cap A_m = \varnothing) = \left(1 - \frac{1}{2^m}\right)^n. \qquad \blacksquare$$

階段 3　子集

想一想　要使 $A_1 \subseteq A_2$，一個元素的四種情況之中，哪一種是禁止的？

$A_1 \subseteq A_2$ 不成立，當且僅當有某個元素在 A_1 裏而不在 A_2 裏。同樣是四種模式中有一種被禁止，所以

$$\mathrm{P}(A_1 \subseteq A_2) = \left(\frac{3}{4}\right)^n,$$

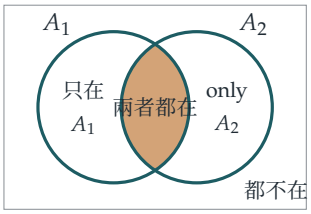
與不相交的情況相同。這並非巧合：$A_1 \subseteq A_2$ 當且僅當 A_1 與 A_2 的補集（complement）不相交，而隨機集合的補集也是一個隨機集合。

想一想　要使 $A_1 \subseteq A_2 \subseteq A_3$，一個元素的 8 種歸屬模式（在或不在每個集合裏）之中，哪些是容許的？

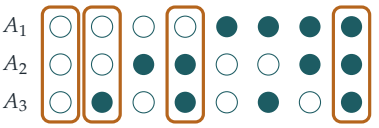
沿着 A_1, A_2, A_3 這條鏈（chain）走，一個元素一旦進入，就必須一直留在裏面。所以容許的模式是「不在、不在、不在」、「不在、不在、在」、「不在、在、在」和「在、在、在」：8 種之中的 4 種（圖 9.12.1 右）。因此 $\mathrm{P}(A_1 \subseteq A_2 \subseteq A_3) = \left(\frac{4}{8}\right)^n = \left(\frac{1}{2}\right)^n$。

對於由 m 個集合組成的鏈，一個容許的模式取決於該元素最先進入哪一個集合（$A_1, \ldots, A_m$ 其中之一，或者從不進入），所以 2^m 種模式之中有 $m+1$ 種是容許的：

$$\mathrm{P}(A_1 \subseteq A_2 \subseteq \cdots \subseteq A_m) = \left(\frac{m+1}{2^m}\right)^n. \qquad \blacksquare$$



每個元素：4 個等可能的位置



符合 $A_1 \subseteq A_2 \subseteq A_3$ 的模式：4 種（共 8 種）

圖 9.12.1 左：S 的一個元素相對於 A_1 和 A_2 可以處於四個位置之一，每個位置的概率都是 $\frac{1}{4}$。不相交這個條件禁止「兩者都在」（琥珀色）；$A_1 \subseteq A_2$ 則禁止「只在 A_1」。右：元素對 A_1、A_2、A_3 的八種歸屬模式（實心圓表示「在」）。鏈 $A_1 \subseteq A_2 \subseteq A_3$ 只容許加了方框的四種模式，即元素一旦進入便一直留在裏面。

回顧

- **按元素分解。**每個集合層面的條件（不相交、子集、鏈），都是分別加在每個元素上的條件。
- **元素之間互相獨立。**由於各個元素由獨立的擲硬幣結果決定，概率可以相乘。
- **數算模式。**每個答案都是（容許的模式數 ÷ 全部模式數）^{n}。

背後的數學

不用概率的計數

把每個概率乘以選擇的總數，就得到一個計數。滿足 $A_1 \subseteq A_2$ 的有序對 (A_1, A_2) 的數目是 $4^n \times \left(\frac{3}{4}\right)^n = 3^n$。你也可以直接數算這些有序對：先選定 A_2 的大小 k，然後以 $\binom{n}{k}$ 種方法選出 A_2 本身，再從它的 2^k 個子集中任選一個作為 A_1。所以

$$\sum_{k=0}^{n} \binom{n}{k} 2^k = 3^n,$$

這正是 $(1+2)^n$ 的二項式定理（binomial theorem）。用兩種方法數算同一樣東西，一次按元素、一次按大小，便證明了一條恆等式：這就是算兩次（double counting）的方法，你在 Section 7.9 討論柏拉圖立體時已經見過。同樣地，鏈 $A_1 \subseteq \cdots \subseteq A_m$ 的數目是 $(m+1)^n$。

子集的格

$\{1, \ldots, n\}$ 的所有子集按包含關係排序，形成一個稱為布爾格（Boolean lattice）的結構，可以把它想像成一個 n 維立方體：每個子集是一個頂點，以它那一串是非題的答案作標

記，而沿着一條棱移動，就是加入或移走一個元素。鏈就是向上走的路徑。有些深刻的結果描述了在給定條件下，子集族可以有多大。施佩納定理（Sperner's theorem，1928 年）指出，在沒有一個集合包含另一個集合的子集族之中，最大的一族有 $\binom{n}{\lfloor n/2 \rfloor}$ 個集合；艾狄胥—柯—拉多定理（Erdős–Ko–Rado theorem，1961 年）則說明，最多可以選出多少個 k 元集合，使得任意兩個都有共同元素。兩者都是組合數學（combinatorics）入門課程的標準內容，而把子集看成擲硬幣的觀點，給出了其中一些最漂亮的證明。

布爾的代數

布爾（George Boole）的 *The Laws of Thought*（《思維規律》，1854 年）把邏輯命題當作集合，把「與」、「或」、「非」當作交集、併集和補集。把一個集合編碼成一串是非題的答案，正是電腦表示集合的方法：$\{1, \ldots, 64\}$ 的一個子集就是單獨一個 64 位元的數，而交集就是「逐位元 AND」（bitwise AND）指令。本題的方法，即每次只看一個元素（一個位元），正是分析這類運算的方法。

思維技巧

每次只想一個元素

當對象是由每個元素各自的獨立選擇構成時，一個分別加在每個元素上的條件，其概率等於一個乘積。

這個方法把一個關於 2^n 個子集的問題，變成一個關於 4 種或 8 種模式的問題。到了大學，它會在組合數學（集合與多重集的生成函數（generating functions））、概率論（獨立的坐標）和計算機科學（逐位元的推理）中再次出現。

延伸練習

1 **覆蓋。**求 $\mathrm{P}(A_1 \cup A_2 = S)$ 和 $\mathrm{P}(A_1 \cup A_2 \cup \cdots \cup A_m = S)$。

2 **大小相同。**證明 $\mathrm{P}(|A_1| = |A_2|) = \binom{2n}{n}/4^n$。（提示：$|A_1| = |A_2|$ 當且僅當 $|A_1| + |S \setminus A_2| = n$。）證明當 n 很大時，這個概率約為 $1/\sqrt{\pi n}$。

3 **分成三份。**證明 S 的不相交子集所組成的有序對 (A, B) 的數目是 3^n，並求由不相交子集所組成的無序對 $\{A, B\}$ 的數目。

4 **小 n 的施佩納定理。**對 $n = 3$ 和 $n = 4$，求 $\{1, \ldots, n\}$ 的子集族中，沒有一個集合包含另一個集合的最大一族。

延伸閱讀

- **Miklós Bóna, *A Walk Through Combinatorics***（World Scientific），其中關於子集和偏序集（partially ordered sets）的章節。
- **Béla Bollobás, *Combinatorics: Set Systems, Hypergraphs, Families of Vectors and Combinatorial Probability***（Cambridge University Press）。一本簡短而優雅的書，講述子集族，由施佩納定理說起。
- **Peter Cameron, *Combinatorics: Topics, Techniques, Algorithms***（Cambridge University Press）。一門入門課程，特別着重用兩種方法計數。

9.13 做了手腳的骰子

題目

STEP I 1999，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/99-s1-q12`

引子

擲兩顆公平骰子時，點數總和為 7 的可能性是總和為 2 的六倍。桌上遊戲的設計者有時會希望情況並非如此：能否在骰子上做手腳（加上配重，使某些面較常出現），令 2 至 12 的每個總和都有相同的機會，即 $\frac{1}{11}$？那十二個概率（probability）可以任你選擇，看來迴旋的餘地相當充裕。

答案是不能，而證明堪稱一絕：它只用到十一個總和中的三個，以及不等式（inequality）$x + \frac{1}{x} \geq 2$。這是證明某件事不可能的典範；與證明某件事辦得到相比，這往往更難，也更有趣。

動筆之前

想一想 對於 $x > 0$，你會怎樣證明 $x + \frac{1}{x} \geq 2$？等號何時成立？

兩邊同乘 $x > 0$：不等式變成 $x^2 + 1 \geq 2x$，即 $(x-1)^2 \geq 0$，這顯然成立。等號只在 $x = 1$ 時成立。（另一做法是 $x + \frac{1}{x} - 2 = (\sqrt{x} - \frac{1}{\sqrt{x}})^2$。）這是算術平均—幾何平均不等式（AM–GM inequality）最簡單的形式：x 與 $\frac{1}{x}$ 的平均值不小於它們的幾何平均（geometric mean）1。

逐步解題

階段 1 兩端的總和

想一想　擲出總和 2 有多少種方法？總和 12 呢？試以 q_i 和 r_j 表示它們的概率。

總和為 2 需要兩顆骰子都擲出 1，總和為 12 則需要兩顆都擲出 6，所以 $\mathrm{P}(2) = q_1 r_1$ 及 $\mathrm{P}(12) = q_6 r_6$。

為了引出矛盾，假設全部十一個總和的可能性相等，每個的概率都是 $\frac{1}{11}$。那麼

$$q_1 r_1 = \frac{1}{11}, \qquad q_6 r_6 = \frac{1}{11}.$$

特別地，q_1、r_1、q_6、r_6 四者全為正數。

階段 2　中間的總和

想一想　總和 7 可以由六種方法得出。其中哪兩種只涉及 1 和 6 這兩個面？試利用上面的等式寫出它們的概率。

總和 7 包括結果 $(1, 6)$ 和 $(6, 1)$，所以

$$\mathrm{P}(7) \geq q_1 r_6 + q_6 r_1.$$

現在利用第 1 步的等式引入一個比。由於 $q_1 = \frac{1}{11 r_1}$ 及 $q_6 = \frac{1}{11 r_6}$，

$$q_1 r_6 + q_6 r_1 = \frac{1}{11}\left(\frac{r_6}{r_1} + \frac{r_1}{r_6}\right) \geq \frac{1}{11} \times 2 = \frac{2}{11},$$

這裏用了 $x + \frac{1}{x} \geq 2$，其中 $x = r_6 / r_1 > 0$。因此 $\mathrm{P}(7) \geq \frac{2}{11}$，與 $\mathrm{P}(7) = \frac{1}{11}$ 矛盾。各個總和不可能全部等可能。 ■

停一停，試一試　看看這個證明為何行得通。要令 2 和 12 與 7 同樣可能出現，1 和 6 就必須有相當高的概率，於是單憑 $(1, 6)$ 和 $(6, 1)$ 這兩個「混合」結果，已足以令 7 的概率過高。圖 9.13.1 顯示了所涉及的方格。



圖 9.13.1 兩顆骰子的 36 個結果，各以其總和標示。證明只用到四個角落的方格。兩個琥珀色方格的概率之積，等於兩個深藍綠色方格的概率之積，即 $q_1r_1q_6r_6 = q_1r_6 \cdot q_6r_1$；而根據算術平均—幾何平均不等式，藍綠色那一對之和至少是琥珀色那一對的幾何平均的兩倍。

階段 3 同一顆骰子擲兩次

想一想 如果把一顆骰子擲兩次，兩次點數相同的概率是多少？試以 q_j 表示。對公平骰子而言，這個概率是多少？

兩次擲出相同點數的概率是 $q_1^2 + q_2^2 + \cdots + q_6^2$。對公平骰子而言，這是 $6 \times \frac{1}{36} = \frac{1}{6}$，已經遠大於 $\frac{1}{36}$。

想一想 在骰子上做手腳，能否令 $\sum q_j^2$ 小於 $\frac{1}{6}$？你知道 $\sum q_j = 1$。

寫成 $q_j = \frac{1}{6} + e_j$，其中各偏差 e_j 之和為零。那麼

$$\sum q_j^2 = \sum\left(\frac{1}{36} + \frac{e_j}{3} + e_j^2\right) = \frac{1}{6} + \frac{1}{3}\sum e_j + \sum e_j^2 = \frac{1}{6} + \sum e_j^2 \geq \frac{1}{6}.$$

所以擲出重複點數的概率必定不小於 $\frac{1}{6}$，而且只有公平骰子才令等號成立。它當然不可能小於 $\frac{1}{36}$。 ■

回顧

- **反證法（proof by contradiction）。**我們假設所有總和等可能，然後推出不可能成立的結論。
- **只用所需的條件。**十一條等式之中，三條已經足夠。選對等式正是關鍵的一步。
- **以不等式證明不可能。**第 (i) 部分的算術平均—幾何平均不等式，以及第 (ii) 部分的「平方和非負」，都一舉排除了一整類的可能性。

背後的數學

用多項式來證明

以下是一個完全不同的證明，用的是 Section 9.6 的生成函數（generating function）。設 $Q(x) = q_1x + \cdots + q_6x^6$ 及 $R(x) = r_1x + \cdots + r_6x^6$。假如所有總和的概率都是 $\frac{1}{11}$，則

$$Q(x)R(x) = \frac{x^2 + x^3 + \cdots + x^{12}}{11} = \frac{x^2}{11}\bigl(1 + x + \cdots + x^{10}\bigr).$$

現在 $Q(x)/x$ 是一個 5 次實係數多項式（polynomial），而每個奇數次的實多項式都有一個實根（它的圖像由 $-\infty$ 走到 $+\infty$，或者反過來）。所以左邊有一個非零的實根。但 $1 + x + \cdots + x^{10} = \frac{x^{11}-1}{x-1}$，它的根是 1 以外的 11 次單位根（roots of unity），當中沒有一個是實數。矛盾。同樣的論證表明，只要 $n - 1$ 是奇數，即 n 是偶數，無論怎樣在兩顆 n 面骰子上做手腳，都不能令各個總和等可能。（當 n 是奇數時，多項式的論證需要多花些工夫，但上面的算術平均—幾何平均證明對每個 $n \geq 2$ 都適用。）

碰撞概率

第 (ii) 部分中的量 $\sum q_j^2$，即兩次獨立抽取結果相同的機會，稱為一個分佈的*碰撞概率*（collision probability）。第 (ii) 部分表明，均勻分佈（uniform distribution）的碰撞概率最小。它量度一個分佈有多集中：在生態學中，$\sum q_j^2$ 是辛普森多樣性指數（Simpson's index of diversity），其中 q_j 是物種 j 在某個生境中所佔的比例；在經濟學中，它是量度市場集中度的赫芬達爾指數（Herfindahl index）；在密碼學中，它量度一個密碼分佈有多容易被猜中。它也解釋了 Section 9.9 結尾的那句話：各個生日的概率不相等，只會令生日相同的情況更有可能出現。

不可能性的證明

數學中一些最著名的結果，說的正是某件事辦不到：沒有分數的平方等於 2；沒有尺規作圖法能三等分任意角；沒有根式公式能解所有五次方程；沒有算法能判定每一個數學命題。每一個這樣的結果都改變了所屬的學科，因為要證明*沒有*任何方法行得通，就必須了解所有方法構成的整個空間。灌鉛骰子問題是一個縮影：要排除每一種做手腳的方式，我們需要一個每種方式都必須遵守的不變性質（算術平均—幾何平均不等式）。

思維技巧

以反證法和不等式證明不可能

要證明沒有任何參數的選取行得通，先假設有一組行得通，找出互相抵觸的那幾個條件，再用不等式證明它們不可能同時成立。

在大學裏，反證法配合不等式是證明某些對象不存在的標準方法，見於分析、組合數學和最優化；在這些領域中，算術平均—幾何平均不等式和柯西—施瓦茲不等式（Cauchy–Schwarz inequality）這類不等式為可達到的結果設下了界限。

延伸練習

1 **柯西—施瓦茲不等式。**利用柯西—施瓦茲不等式 $\left(\sum a_j b_j\right)^2 \le \sum a_j^2 \sum b_j^2$，取 $a_j = q_j$ 及 $b_j = 1$，為第 (ii) 部分給出一行的證明。

2 **兩顆不同的骰子。**如果把兩顆不同的灌鉛骰子各擲一次，它們點數相同的概率能否小於 $\frac{1}{6}$？能否等於 0？

3 **面上數字不同時的均勻總和。**證明如果兩顆骰子各面的數字另作安排，它們可以給出等可能的總和：試取一顆骰子的數字為 $0, 1, 2, 3, 4, 5$，另一顆為 $0, 6, 12, 18, 24, 30$。會出現哪些總和？為甚麼這與題目並不矛盾？

4 **三顆骰子。**三顆各面為 1 至 6 的六面灌鉛骰子，能否令 3 至 18 的所有總和等可能？（提示：利用總和 3、8、13 和 18，以及六個數的算術平均—幾何平均不等式。為甚麼多項式證明這次行不通？）

延伸閱讀

- **J. Michael Steele, *The Cauchy–Schwarz Master Class***（Cambridge University Press）。一本出色的不等式入門書，由算術平均—幾何平均不等式講起。
- **Geoffrey Grimmett & Dominic Welsh, *Probability: An Introduction***（Oxford University Press）。講述生成函數，以及它們在上述這類證明中的應用。
- **Ian Stewart, *Galois Theory***（CRC Press），開首數章。最著名的不可能性證明：五次方程。

我們學到了甚麼

本章開首說過，計數是最古老的數學，而要數得巧妙，至今仍是最難的事情之一。我們提出了三個問題：怎樣不逐一列舉而數出答案？甚麼時候某件事必然發生？計數與概率（probability）怎樣互相幫助？以下是本章十三道題目所揭示的答案。

1. 怎樣不逐一列舉而數出答案？

本章逐步建立了一套工具。

- **遞歸關係（recurrences）。**在平面上多加一條直線（Section 9.1），或在圍欄上多加一根柱子（Section 9.3），數目都只會以簡單的方式改變，由此得出可以求解的遞歸關係。只要多記錄一點資料，即最後一根柱子的顏色，這些遞歸關係便能自成一體。
- **生成函數（generating functions）。**把數列（sequence）包裝成冪級數（power series），遞歸關係便化為代數運算（Section 9.4），既解釋了為何 $1/89 = 0.01123\ldots$，也把數算組合變成多項式相乘：Section 9.5 的硬幣、Section 9.6 的骰子，以及 Section 9.7 的等冪和，其中 $x = 1$ 處的重根（repeated root）包辦了一切。
- **雙射（bijections）。**反射原理（reflection principle）（Section 9.10）把不合格的隊列與不受限制的路徑一一配對，從而數出前者的數目；而取隨機集合的補集（complement），則把子集變成互不相交的集合對（Section 9.12）。
- **容斥原理（inclusion–exclusion）。**先多算、後修正，便求得錯排（derangements）的數目（Section 9.8）。
- **重複計數（double counting）。**以歐拉公式（Euler's formula）數算區域（Section 9.1），以及按大小數算子集對（Section 9.12），都是循不同途徑得出相同答案，並順道證明了一些恆等式。

2. 甚麼時候某件事必然發生？

有好幾節顯示，單憑規模已足以迫出結構。根據鴿巢原理（pigeonhole principle），六個人當中必有三人互相認識，或有三人互不相識（Section 9.2），但五個人則不足夠。二十三人當中有人同一天生日的機會已過半（Section 9.9），因為配對的數目隨人數的平方增長。一個大集合的隨機子集幾乎必定互相重疊（Section 9.12）。還有一些事情，無論怎樣嘗試都不可能發生：不論怎樣為兩顆骰子灌鉛，都無法令每個點數總和出現的機會均

等（Section 9.13）；而除了 Sicherman 的編號方式外，再沒有其他骰子編號能重現普通骰子的點數總和分佈（Section 9.6）。

3. 計數與概率怎樣互相幫助？

在本章的後半部分，概率是靠計數求得的：錯排用容斥原理，生日問題用連乘積，選票問題（ballot problem）用反射。但反過來，概率也幫了計數的忙。期望值的線性性質（linearity of expectation）給出了集換券收集問題（coupon collector）的 nH_n（Section 9.11），也給出了洗牌後配對數目的期望值，完全毋須計數。把隨機子集看成擲 n 次硬幣（Section 9.12），3^n 和 $(m+1)^n$ 這類數目便成了一行就能完成的計算。而艾狄胥（Erdős）的概率方法（probabilistic method）（Section 9.2）則透過證明隨機着色有正概率符合要求，證明了理想的着色確實存在。有兩個常數憑空出現：e 出現在錯排概率的極限中，而歐拉常數 γ 則出現在集齊一整套所需的代價中。

4. 計數的工具箱

方法	用於何處
多加一個時有何改變	直線（9.1）；圍欄柱子（9.3）
鴿巢原理	拉姆齊（Ramsey）（9.2）
耦合遞歸關係	圍欄柱子（9.3）
生成函數	9.4、9.5、9.6、9.7、9.13
從極端情況入手分類討論	奇異骰子（9.6）；灌鉛骰子（9.13）
容斥原理	錯排（9.8）
補集與指數近似	生日（9.9）
反射原理	選票問題（9.10）
期望值的線性性質	集換券（9.11）；配對（9.8）
每次處理一個元素	隨機子集（9.12）
以算術—幾何平均不等式（AM–GM）導出矛盾	灌鉛骰子（9.13）

下一步

- 錯排的泊松極限（Poisson limit）會在 Section 12.11 再次出現，而隨機變量（random variables）的生成函數則推動 Section 12.15 對家族樹的研究。
- 耦合遞歸關係在 Section 12.16 會變成馬可夫鏈（Markov chains）。
- 集換券收集問題中的岡貝爾分佈（Gumbel distribution），在 Section 12.25 用來描述眾多隨機量中的最大者。

如果你想在入讀大學前再進一步

- **Miklós Bóna, *A Walk Through Combinatorics***：一本平易近人的入門教材，幾乎涵蓋本章的每一個概念。
- **Herbert Wilf, *generatingfunctionology***：生成函數的經典著作，可於網上免費閱讀。
- **Ronald Graham, Donald Knuth & Oren Patashnik, *Concrete Mathematics***：求和、遞歸關係與生成函數，筆調風趣，習題豐富。
- **Geoffrey Grimmett & Dominic Welsh, *Probability: An Introduction***：大學一年級概率論中與計數有關的部分。

懷特黑德（Whitehead）曾稱組合數學（combinatorics）為「拓撲學的貧民窟」。本章嘗試說明，為何今天再沒有人會這樣說。每一個巧妙的計數背後，都有一個影響遠超計數本身的想法：一個生成函數、一個雙射、一個遞歸關係、一個不等式。這些想法把計數與代數、分析和概率連繫起來；學會尋找這些想法，而不是再多找一個竅門，正是令計數成為數學的關鍵。

延伸練習答案

以下只是簡短的答案和提示，並非完整解答：這裏只給出結果和關鍵步驟，方便你核對自己的做法。請先認真做過每道練習，然後才翻看。

9.1 直線分割平面

1. 一個圓把平面分成 2 個區域。第 n 個圓與其餘各圓相交於 $2(n-1)$ 點，這些點把它分成 $2(n-1)$ 段弧，每段弧把一個區域一分為二。所以 $R_n = R_{n-1} + 2(n-1)$，由此得 $n^2 - n + 2$。

2. 每條直線所增加的區域數目，是 1 加上該直線上的交點數目，所以總數是 $1 + n +$ (交點數目)。那 k 條平行線少了 $\binom{k}{2}$ 個交點：$1 + n + \binom{n}{2} - \binom{k}{2}$。

3. 第 n 個平面被其餘平面截出 $n-1$ 條處於一般位置（general position）的直線，這些直線把它分成 $\frac{1}{2}(n^2 - n + 2)$ 塊，每一塊把空間中的一個區域一分為二。把這個數加到 $\frac{1}{6}\big((n-1)^3 + 5(n-1) + 6\big)$ 上，便得 $\frac{1}{6}(n^3 + 5n + 6)$。

4. 當 $n = 6$ 時有 31 個區域，而不是 32 個。在圓內，這個圖有 $V = n + \binom{n}{4}$ 個頂點（vertex）（每個交點由四個點決定）和 $E = n + \binom{n}{2} + 2\binom{n}{4}$ 條邊（edge）（圓上的各段弧，加上各條弦，而每個交點再多添兩段）。歐拉（Euler）公式 $V - E + F = 2$ 把圓外的區域也算在內，所以區域的數目是 $F - 1 = 1 + \binom{n}{2} + \binom{n}{4}$。

9.2 小型版的拉姆齊定理

1. 雙色角（bichromatic angle）最多有 $6 \times 6 = 36$ 個。每個非單色的三角形恰好含有兩個雙色角，所以這類三角形最多有 18 個。由於 $\binom{6}{3} = 20$，最少有 2 個三角形是單色的。

2. 每個點有 16 條邊，所以其中最少有 6 條同色，設為紅色，並設它們連到 $P_1, \ldots, P_6$ 各點。若 P_i 之間有任何一條邊是紅色，便有一個紅色三角形。若沒有，則各 P_i 構成一個以兩種顏色著色的 K_6，其中必有單色三角形。

3. 點 v 有 9 條邊，所以要麼最少有 4 條是金色，要麼最少有 6 條是銀色。若有 4 個金色鄰點：它們之間只要有一條金色邊，便與 v 構成金色三角形，否則它們構成一個銀色的 K_4。若有 6 個銀色鄰點：它們之中必有一個單色三角形；若是金色，證明完成，若是銀色，它便與 v 構成一個銀色的 K_4。

4. 給每一項標上 (i, d)：分別是以該項結尾的最長遞增子序列（subsequence）和最長遞減子序列的長度。對任意兩項，較後的一項不是 i 較大，就是 d 較大，所以各個標記互不相同。假如沒有長度為 3 的子序列，那麼 5 項就只有 4 個可能的標記。兩者皆無的例子：$2, 1, 4, 3$。

9.3 為籬笆樁着色

1. 與之前一樣，$r_{n+1} = s_n$（第 n 根柱不是紅色，所以第 $n+1$ 根柱可以是紅色），而 $r_n + s_n = (k-1)^{n-1}$，因為第一根之後的每根柱都有 $k-1$ 種選擇。所以 $r_{n+1} + r_n = (k-1)^{n-1}$，而 $r_n = \frac{(k-1)^{n-1}+(k-1)(-1)^{n-1}}{k}$。至於圍成一圈的情況：第 1 根柱的顏色有 k 種選擇，然後是一排 $n+1$ 根柱，最後一根與第一根同色，由此得 $kr_{n+1} = (k-1)^n + (-1)^n(k-1)$。

2. $2, 3, 5, 8, \ldots$：設共有 a_n 種方法，以白色結尾的一排可接在任何一排 $n-1$ 根柱之後，而以紅色結尾的一排則接在以白色結尾的一排 $n-2$ 根柱之後，所以 $a_n = a_{n-1} + a_{n-2}$。答案是 F_{n+2}。

3. 中心有 4 種選擇，然後是 n-圈的 3-著色：$4\big(2^n + 2(-1)^n\big)$。

4. 當 $n = 4$ 時：有 18 種著色。轉四分之一圈不會保持任何一種著色不變，轉半圈則保持其中 6 種（即相對的柱同色的那些）不變，所以有 $\frac{1}{4}(18 + 0 + 6 + 0) = 6$ 條項鏈。當 $n = 5$ 時：有 30 種著色，沒有任何旋轉能保持其中一種不變，所以是 6。

9.4 把數列看成函數

1. 餘下各項之和最多是 $F_7 \sum_{r\geq 7} 2^{r-7}10^{-r} = \frac{13\times 10^{-7}}{0.8} < 2 \times 10^{-6}$。由於 $\sum F_r 10^{-r} = \frac{10}{89}$，而首六項給出 0.112358，所以 $\frac{10}{89}$ 介乎 0.112358 與 0.112360 之間，於是 $\frac{1}{89} = 0.011235\ldots$。若改以兩位數為一組，則 $\sum F_r 100^{-r-1} = \frac{1}{9899}$。

2. 把遞推關係（recurrence）求和，得 $\mathrm{f}(x)(1 - 7x + 10x^2) = 2 - 7x$，而 $\frac{2-7x}{(1-2x)(1-5x)}$ 可如題目所述分拆。至於 b_n：由 $10p + 5q = 40$ 和 $40p + 10q = 100$ 得 $p = 1$，$q = 6$；$t^2 = t + 6$ 的根是 3 和 -2，而 $b_n = 4 \cdot 3^n + (-2)^n$。

3. $\sum nx^n = \frac{x}{(1-x)^2}$；在 $x = \frac{1}{100}$ 時它等於 $\frac{100}{9801}$，所以 $\frac{1}{9801} = \sum n \cdot 100^{-n-1}$。這個規律在 98 這一組出錯：100 那一組把 1 進位到 99 那一組，後者又再進位到 98 那一組，於是展開式是 $\ldots 96\,97\,99\,00\,01\ldots$。

4. 乘以 $\frac{1}{1-x} = 1 + x + x^2 + \cdots$ 之後，x^n 的係數變成 $u_0 + \cdots + u_n$。平方數的生成函數（generating function）是 $\frac{x(1+x)}{(1-x)^3}$，所以其部分和的生成函數是 $\frac{x+x^2}{(1-x)^4}$，而 x^n 的係數是 $\binom{n+2}{3} + \binom{n+1}{3} = \frac{n(n+1)(2n+1)}{6}$。

9.5 湊零錢

1. x^{24} 的係數就是 $6a+6b+3c=24$ 的解的數目：對由 0 到 4 的每個 $s=a+b$，各有 $s+1$ 種選擇，合共 15 種。一般而言，x^n 的係數是 0，除非 $n=3m$，這時它等於 $\frac{1}{2}(k+1)(k+2)$，其中 $k=\lfloor m/2\rfloor$。加上額外的因式後，係數變成部分和：x^{24}、x^{25} 和 x^{66} 的係數分別是 55、55 和 650。它們所數的，是用 1p 和 3p 硬幣以及兩種不同的 6p 硬幣付款的方法數目。

2. 每個部分 k 要麼用一次，要麼完全不用。這個乘積可以逐項相消：$\prod \frac{1-x^{2k}}{1-x^k}$ 把分母中指數為偶數的因式全部約去，剩下 $\prod \frac{1}{1-x^{\text{奇數}}}$，也就是分拆（partition）成奇數部分的生成函數。當 $n=6$ 時：6、$5+1$、$4+2$、$3+2+1$，以及 $5+1$、$3+3$、$3+1+1+1$、$1+1+1+1+1+1$：兩類各有四個。

3. 當 $n=0,1,2,\ldots$ 時，數值依次是 $1,1,2,2,3,4,5,6,7,8,10,\ldots$，而 $\frac{(n+4)^2}{20}$ 與它們的差距總不超過 $\frac{1}{2}$（已驗算至 $n=60$；證明則要用到 $\frac{1}{(1-x)(1-x^2)(1-x^5)}$ 的部分分式（partial fractions））。

4. 最後一枚硬幣是 1p 或 2p，所以 $c_n=c_{n-1}+c_{n-2}$：這就是斐波那契（Fibonacci）數，$c_n=F_{n+1}$。

9.6 奇怪的骰子

1. 每顆骰子都是 x 乘以 $1+x$、$1+x$、$1+x^2$、$1+x^2$ 這些因式中若干個的乘積，而且在 $x=1$ 時的值是 4，所以它要取其中兩個因式。除了標準的一對之外，唯一的選擇是 $x(1+x)^2$ 和 $x(1+x^2)^2$：各面是 $1,2,2,3$ 和 $1,3,3,5$。

2. 乘積是 $2x^3+x^4+x^5+4x^6+3x^7+8x^8+6x^9+5x^{10}+6x^{11}$，與表中數字吻合。

3. 可以：取 Sicherman 骰子對 $1,2,2,3,3,4$ 和 $1,3,4,5,6,8$，再加一顆標準骰子。

4. $(x+x^2)(x+\cdots+x^8)=x^2(1+x)^2(1+x^2)(1+x^4)$。硬幣需要一個值為 2 的因式：除了標準的選擇 $x(1+x)$ 之外，它可以是 $x(1+x^2)$（兩面是 $1,3$，配上八面體（octahedron）骰子 $1,2,2,3,5,6,6,7$），也可以是 $x(1+x^4)$（兩面是 $1,5$，配上 $1,2,2,3,3,4,4,5$）。

9.7 相等的冪和

1. $\sum(c+b_k)^3-\sum(c+a_k)^3 = 3c^2\left(\sum b_k-\sum a_k\right)+3c\left(\sum b_k^2-\sum a_k^2\right)+\sum b_k^3-\sum a_k^3$，而首兩個括號都等於零。由 $\{1,4,6,7\}$、$\{2,3,5,8\}$ 出發，取 $c=8$：得 $\{1,4,6,7,10,11,13,16\}$ 和 $\{2,3,5,8,9,12,14,15\}$。

2. A 和 $B\cup\{0\}$ 都有 6 個元素，和為 33，平方和為 253。於是 $\sum_{a\in A}(n+a)^2=\sum_{b\in B\cup\{0\}}(n+b)^2$，這樣便把 $\{n^2,\ldots,(n+11)^2\}$ 分成兩組。

3. 用二項式定理（binomial theorem）展開 $(x+c)^r$：式中只涉及不高於 r 次的冪，而這些冪的和都相等。題目中的兩個集合，在第二個加入 0 之後，合起來就是 $\{0,\dots,15\}$，再作平移便得任意 16 個連續整數。

4. 按 $ABABABAB$ 的次序，A 約得 1.676，B 約得 1.042。按 $ABBABAAB$ 的次序，A 約得 1.560，B 約得 1.158：仍不相等，但已接近得多。

9.8 沒有人拿到自己的恤衫

1. $\mathrm{E}(C_1)=\frac{1}{m}$，$\mathrm{Var}(C_1)=\frac{1}{m}-\frac{1}{m^2}$，而 $\mathrm{Cov}(C_1,C_2)=\frac{1}{m(m-1)}-\frac{1}{m^2}=\frac{1}{m^2(m-1)}$。於是 $\mathrm{E}(N)=1$，而 $\mathrm{Var}(N)=m\bigl(\frac{1}{m}-\frac{1}{m^2}\bigr)+m(m-1)\frac{1}{m^2(m-1)}=1$。當 $m=4$ 時：$\mathrm{P}(N=0)=\frac{9}{24}=0.375$，相比之下 $\mathrm{e}^{-1}\approx 0.368$。

2. 第 1 張牌落在某個位置 i（有 $n-1$ 種選擇）。若第 i 張牌落在位置 1，餘下的就是 $n-2$ 張牌的錯排（derangement）；若不是，則把位置 1 當作第 i 張牌的禁位，得到 $n-1$ 張牌的錯排。驗算：$D_3=2(1+0)=2$，$D_4=3(2+1)=9$。

3. 先選出 k 張留在原位的牌，再把其餘的牌錯排：$\binom{n}{k}\frac{D_{n-k}}{n!}=\frac{1}{k!}\cdot\frac{D_{n-k}}{(n-k)!}=\frac{d_{n-k}}{k!}$，它趨向 $\frac{\mathrm{e}^{-1}}{k!}$。

4. $30-(15+10+6)+(5+3+2)-1=8$，它等於 $30\cdot\frac{1}{2}\cdot\frac{2}{3}\cdot\frac{4}{5}$：容斥原理（inclusion–exclusion）的和式可以因式分解（這就是歐拉的 $\varphi(30)$）。

9.9 生日問題

1. $\mathrm{P}(X=r)=\bigl(1-\frac{1}{n}\bigr)\cdots\bigl(1-\frac{r-2}{n}\bigr)\frac{r-1}{n}$，而這些概率之和為 1。至於期望值（expectation），$\mathrm{P}(Y=r)$ 在 $\mathrm{P}(Y\ge 1),\dots,\mathrm{P}(Y\ge r)$ 中各被計算一次。記 $Q_k=\mathrm{P}(X>k)=\prod_{j=1}^{k-1}\bigl(1-\frac{j}{n}\bigr)$。於是 $\mathrm{E}(X)=\sum(k+1)\frac{k}{n}Q_k$，同時 $\mathrm{E}(X)=1+\sum Q_k$；利用 $\sum\frac{k}{n}Q_k=1$ 並將兩式相減，得 $\sum Q_k\bigl(1-\frac{k^2}{n}\bigr)=0$。

2. 由 $\binom{k}{3}\approx 365^2\ln 2$ 得 $k\approx 82$；這個近似忽略了各個三人組之間的相依關係，真正的答案是 88。

3. 14 人。（每一對人「接近」的概率是 $\frac{3}{365}$，而由 $\binom{k}{2}\frac{3}{365}\approx\ln 2$ 得 $k\approx 14$；精確計算的結果是：14 人時概率約為 0.537，13 人時約為 0.483。）

4. 兩人同在第 i 天生日的概率是 p_i^2；把它們加起來。根據柯西–施瓦茲不等式（Cauchy–Schwarz），$1=\bigl(\sum p_i\bigr)^2\le 365\sum p_i^2$，等號僅在所有 p_i 都相等時成立。

9.10 總能找贖

1. 第一個人必須持有一枚 £1 硬幣。若按其後幾個人的情況取條件，便要分成許多情形（視乎每位 £2 顧客之後你手上有哪些硬幣），每種情形最後都歸結到第 (i) 和第

(ii) 部分的答案；把它們合起來，得 $\frac{m-2}{m+1}$。用反射原理（reflection principle）則一行便可完成。

2. 第一票必須投給 A。在以 A 開始的點票過程之中，壞的那些（即曾經出現平手的）可藉着反射第一次平手之前的部分，與以 B 開始的點票過程一一對應。所以好的點票過程的數目是 $\binom{a+b-1}{a-1} - \binom{a+b-1}{a}$，除以 $\binom{a+b}{a}$ 便得 $\frac{a-b}{a+b}$。

3. $((()))$、$(()())$、$(())()$、$()(())$、$()()()$。一個雙射（bijection）：第一對括號把序列分成括號之內的部分和括號之後的部分；相應地，以某條固定的邊為底的三角形把多邊形分成兩個較小的多邊形。對兩部分分別遞歸處理。

4. 每位 £5 顧客需要找回四枚 £1 硬幣，所以路徑每步上升 1 或下降 4。路徑可以越過界線而不碰到它，於是反射再不能把壞的路徑與一個簡單的集合配對起來。

9.11 集齊一套

1. 各段等待時間的期望值之和是 $\frac{n}{n} + \frac{n}{n-1} + \cdots + \frac{n}{n/2+1} = n(H_n - H_{n/2}) \approx n(\ln n - \ln \frac{n}{2}) = n \ln 2$。

2. 每張卡缺少的概率是 $(1 - \frac{1}{n})^k$；對 n 張卡求和。要達到 95%，需 $(1 - \frac{1}{n})^k \approx \mathrm{e}^{-k/n} = 0.05$，所以 $k \approx n \ln 20 \approx 3n$。

3. 買了 t 包之後，某一張指定的卡出現少於兩次的概率約為 $\mathrm{e}^{-t/n}(1 + \frac{t}{n})$。當這類卡的期望數目 $n\mathrm{e}^{-t/n}(1 + \frac{t}{n})$ 降至約 1 時，兩個孩子便大致都集齊了；由此得 $\frac{t}{n} \approx \ln n + \ln \ln n$，遠小於 $2 \ln n$。

4. $\frac{1}{n+1} < \ln \frac{n+1}{n}$（$\frac{1}{x}$ 在 $[n, n+1]$ 上的曲線下面積），所以 $H_n - \ln n$ 遞減。由凸性（convexity）可知梯形（trapezium）高估了面積：$\ln \frac{n+1}{n} < \frac{1}{2}\left(\frac{1}{n} + \frac{1}{n+1}\right)$，所以 $H_n - \ln n - \frac{1}{2n}$ 遞增。取 $n = 1$，得 $\gamma > \frac{1}{2}$；取 $n = 4$，得 $\gamma < H_4 - \ln 4 \approx 0.697$。

9.12 隨機子集

1. 每個元素必須屬於最少一個集合：$\left(\frac{3}{4}\right)^n$，一般情況則是 $\left(1 - 2^{-m}\right)^n$。

2. $|S \setminus A_2|$ 與 $|A_2|$ 有相同的分佈（distribution），並且與 A_1 獨立（independent），所以所求概率就是 $|A_1| + |S \setminus A_2| = n$ 的概率。$(A_1, S \setminus A_2)$ 這一對集合對應於一個有 $2n$ 個元素的集合的隨機子集，其大小為 n 的概率是 $\binom{2n}{n}/4^n$。由斯特靈（Stirling）公式（或沃利斯（Wallis）乘積，見 Section 4.4）得 $\approx \frac{1}{\sqrt{\pi n}}$。

3. 每個元素放進 A、放進 B，或兩者都不放：3^n。滿足 $A = B$ 的有序對只有 $(\varnothing, \varnothing)$，所以共有 $\frac{3^n-1}{2} + 1 = \frac{3^n+1}{2}$ 個無序對。

4. 當 $n = 3$ 時，是 3 個集合（全部單元素集，或全部二元素集）。當 $n = 4$ 時，是 6 個集合（全部二元素集）。一般情況是 $\binom{n}{\lfloor n/2 \rfloor}$（Sperner 定理，Sperner's theorem）。

9.13 做了手腳的骰子

1. $1 = \left(\sum q_j\right)^2 \le 6\sum q_j^2$，所以 $\mathrm{P}(\text{相同}) = \sum q_j^2 \ge \frac{1}{6} > \frac{1}{36}$。

2. 兩問的答案都是「可以」：所求概率是 $\sum q_j r_j$，而比方說，若一顆骰子永遠擲出 1，另一顆永遠擲出 2，它便等於 0。

3. 點數總和是 $0, 1, \ldots, 35$，每個恰好出現一次：是均勻分佈。題目要求兩顆骰子都標上 1 至 6，而這裏的骰子並非如此。

4. 不可以。記 $\mathrm{P}(3) = q_1 r_1 s_1$ 和 $\mathrm{P}(18) = q_6 r_6 s_6$，兩者都是 $\frac{1}{16}$。$\mathrm{P}(8) + \mathrm{P}(13)$ 包含 $q_1 r_1 s_6, \ldots, q_1 r_6 s_6$ 這六項（兩個 1 和一個 6，或兩個 6 和一個 1），它們的乘積是 $(\mathrm{P}(3)\mathrm{P}(18))^3$。由算術–幾何平均不等式（AM–GM）得 $\mathrm{P}(8) + \mathrm{P}(13) \ge 6 \cdot \frac{1}{16} > \frac{2}{16}$。原來那個多項式證明照搬過來並不成立，因為 $1 + x + \cdots + x^{15}$ *確實*有一個實根 -1。（不過它可以補救：三個實係數五次多項式需要三個實根，而 -1 只是單根。）

章

10

守恆定律、碰撞與引力

有些東西永遠不變。知道是哪些，物理學就懂了一半。

本章內容

10.1	牛頓擺	STEP II 1990
10.2	疊球效應	STEP II 2010
10.3	火箭與槍炮	STEP III 2017
10.4	一個分子與一個活塞	STEP I 2002
10.5	有限時間內的無限次反彈	STEP III 2002
10.6	二體化為一體	STEP III 2004
10.7	球殼內的引力	STEP III 1999
10.8	排成三角形的三顆恆星	STEP I 1999
A	練習答案	第 719 頁

甚麼保持不變

> 有一個事實，或者你願意的話，可稱之爲一條定律，支配着迄今所知的一切自然現象。這條定律至今未見任何例外。
>
> — 費曼（Richard Feynman）論能量守恆，*The Feynman Lectures on Physics*（1963）

學校裏的力學（mechanics）通常離不開力：畫一幅圖，分解各力，再應用 $F = ma$。這套方法對付斜面上的木塊綽綽有餘，卻對付不了歷時僅千分之一秒的碰撞（collision），因為其間的力既巨大又未知；也對付不了一邊飛行一邊改變質量的火箭，或者三顆同時互相牽引的恆星。在這些情況下，出路是不再追問力，而改為追問甚麼不能改變。

只要沒有外力作用，動量（momentum）就守恆。在彈性碰撞（elastic collision）中能量守恆，而在其他碰撞中，能量的損失可以量度出來。質量、角動量（angular momentum）和質心（centre of mass）的運動也遵從類似的規律。這些定律不會告訴你一切，但它們永遠成立，適用於複雜得根本無法仔細分析的系統，而且往往在動手計算之前，就已排除了大部分的可能性。

守恆定律為何重要

守恆定律（conservation law）是物理學中最可靠的工具。在稱為 β 衰變（beta decay）的放射過程中，射出的粒子的能量加起來對不上數；泡利（Wolfgang Pauli）沒有放棄能量守恆，反而在 1930 年預言有一種看不見的粒子帶走了失去的能量，而中微子（neutrino）在二十六年後果然被探測到。工程師靠動量和能量的收支帳來設計防撞欄、安全氣袋和火箭。天文學家靠觀察質心來秤量恆星、尋找行星。到了 1918 年，諾特（Emmy Noether）解釋了這些定律為何成立：每一條守恆定律都是大自然某種對稱性（symmetry）的投影。動量守恆，是因為物理定律在任何地方都相同；能量守恆，是因為物理定律在任何時刻都相同。

引力（gravity）是本章的第二個主題。牛頓（Newton）正是用這種力把天上與地上統一起來，而它與守恆定律是天生的拍檔：軌道、引力彈弓和塌縮的星雲，全都要透過能量和動量，再加上一個關於球體的奇妙事實，才能理解。

你在大學會遇到甚麼

大學一年級的動力學（dynamics）課程就從這裏開始。你會遇到質點系統、質心與約化質量（reduced mass）、不同參考系（frame）中的碰撞、火箭之類的變質量系統，以及

物體在引力下的運動：開普勒（Kepler）定律、軌道、逃逸速度（escape velocity）。往後的課程會用拉格朗日（Lagrange）和哈密頓（Hamilton）的語言重新表述力學，屆時諾特定理（Noether's theorem）會成為一個精確的命題，而守恆定律可以從單獨一個函數的對稱性中直接讀出。再往前走，便是天體力學與三體問題（three-body problem）、分子運動論（kinetic theory）與統計力學（statistical mechanics）（Section 10.4 的活塞在那裏會變成氣體理論），以及愛因斯坦（Einstein）的引力理論。

貫穿本章的問題

三大問題

1. **守恆定律單憑自身能決定甚麼？**動量和能量何時能決定結果，何時只能排除某些可能？
2. **小步驟如何累積？**一個簡單的事件，例如一次碰撞或一次發射，重複多次甚至無限多次，會發生甚麼事？
3. **如何馴服多個物體？**哪些坐標的選擇、哪些對稱性，能把涉及幾個相互作用物體的問題，化為我們解得到的問題？

本章由碰撞開始：牛頓擺（Newton's cradle）做得到甚麼、做不到甚麼（Section 10.1），以及一串球如何令速率倍增（Section 10.2）。接着探討重複的過程：從貨車上開炮與火箭方程（rocket equation）、分子撞向活塞而愈來愈快，以及在有限時間內彈跳無限多次的冰球。最後三節轉向互相吸引或排斥的物體：質心與約化質量、球殼的引力，以及拉格朗日那個由恆星組成的旋轉三角形。

本章的概念

§	課題	概念
10.1	牛頓擺	守恆定律作為約束條件
10.2	疊球	迭代（iteration）同一種碰撞；引力彈弓
10.3	火箭	變質量；火箭方程
10.4	分子運動論	絕熱不變量（adiabatic invariant）
10.5	非彈性碰撞（inelastic collision）	有限時間內無限多次彈跳
10.6	二體問題（two-body problem）	質心與約化質量
10.7	牛頓引力	球殼定理（shell theorem）；自由落體式塌縮
10.8	三體問題	拉格朗日的等邊三角形解

如何閱讀本章

計算之前，先列出甚麼是守恆的，以及在哪個參考系中守恆。然後問一問：換個位置看，問題會否簡單一些？例如從質心看、從移動中的活塞看、從彈跳中的球看。本章不少題目在合適的參考系中幾乎變得不費吹灰之力，而力學的大部分功夫，正在於找出這個參考系。

10.1 牛頓擺

題目

STEP II 1990，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/90-s2-q14`

引子

把牛頓擺（Newton's cradle）的一個球拉開，然後放手：另一端會有一個球飛出。拉開兩個，就有兩個飛出。為甚麼？為甚麼不是一個球以兩倍速率飛出，或者三個球以三分之二的速率飛出？這兩種情況的動量同樣正確。

答案要從力學的兩大守恆定律（conservation laws）說起。動量（momentum）與能量（energy）合起來所施加的限制，遠比其中任何一個單獨施加的嚴格得多，你想像得到的結果大多會被它們排除。本題把這兩條定律用作約束條件（constraints）：不是用來計算會發生甚麼，而是用來判斷甚麼可以發生。把守恆定律看成每一個可能結果都必須遵守的規則，這種想法貫穿整個物理學。

動筆之前

想一想 一粒質量為 m 的彈珠以速率 u 運動，撞上一粒質量為 M 的靜止彈珠，而動能（kinetic energy）守恆。碰撞後兩者的速度是多少？若 $M = m$ 又如何？

設碰撞後的速度為 p 和 q。由動量得 $mp + Mq = mu$，由能量得 $mp^2 + Mq^2 = mu^2$。把這兩式寫成 $m(u-p) = Mq$ 和 $m(u-p)(u+p) = Mq^2$，然後相除，得 $u + p = q$。於是

$$p = \frac{m-M}{m+M}\,u, \qquad q = \frac{2m}{m+M}\,u.$$

當 $M = m$ 時，第一粒彈珠停下，第二粒以速率 u 離開：兩者只是交換了速度。這就是

質量相等的彈珠能把運動沿着隊列傳遞下去的原因。

逐步解題

階段 1　只有最後兩粒運動

想一想　假設碰撞後只有 A_n（速度 p）和 A_{n+1}（速度 q）運動。寫出守恆定律。哪一粒彈珠的運動限制了 p 的正負號？

由動量和能量得

$$mp + \lambda m q = mu, \qquad \tfrac{1}{2}mp^2 + \tfrac{1}{2}\lambda m q^2 = \tfrac{1}{2}mu^2.$$

這正是質量為 m 的彈珠撞上質量為 λm 的彈珠時的方程，所以（除了 $p = u$、$q = 0$ 這個解之外，而該解需要 A_n 穿過 A_{n+1}）

$$p = \frac{1-\lambda}{1+\lambda}\,u, \qquad q = \frac{2}{1+\lambda}\,u.$$

若 $\lambda < 1$，兩者皆為正，且 $p < q$：A_n 跟在 A_{n+1} 後面，但走得較慢，所以兩者分開，一切並無矛盾。若 $\lambda > 1$，則 $p < 0$：A_n 向後運動。但 A_{n-1} 靜止在它的正後方，並與它接觸，而 A_n 不能穿過它。所以這種運動不可能發生。 ■

階段 2　三粒彈珠

想一想　現在假設只有 A_{n-1}、A_n 和 A_{n+1} 運動，速度分別為 a、b、c。已知哪些彈珠靜止、哪一粒在哪一粒後面，a、b、c 必須滿足甚麼不等式（inequalities）？

彈珠 A_{n-2} 靜止在 A_{n-1} 後面（它確實存在，因為 $n \geq 3$），所以 $a \geq 0$。而且沒有彈珠可以越過前面的那一粒，所以 $a \leq b \leq c$。合起來就是 $0 \leq a \leq b \leq c$。守恆定律為

$$a + b + \lambda c = u, \qquad a^2 + b^2 + \lambda c^2 = u^2.$$

想一想　把動量方程平方，再與能量方程比較。剩下甚麼？

平方後，$u^2 = a^2 + b^2 + \lambda^2 c^2 + 2ab + 2\lambda ac + 2\lambda bc$。減去能量方程，

$$0 = \lambda(\lambda - 1)c^2 + 2ab + 2\lambda c(a + b).$$

由於 $\lambda > 1$ 且 $a, b, c \geq 0$，右邊每一項都非負，所以每一項都等於零。特別地，$c = 0$，於是 $0 \leq a \leq b \leq c$ 迫使 $a = b = 0$。但這樣總動量就是 0，而不是 mu。矛盾：這種運動不可能發生。 ■

停一停，試一試 留意這個論證的結構。把動量方程平方會產生 $2ab$ 之類的交叉項（cross terms），而能量方程卻沒有這些項。兩者之差量度了運動有多「分散」。當 $\lambda > 1$ 時，同一個論證可以排除只有 $A_k, \ldots, A_{n+1}$ 以非負速度運動的任何結果，你看得出原因嗎？

階段 3 兩粒可以運動的彈珠

想一想 向前運動的彈珠必須全部位於遠端，而這種情況既然行不通，也許有一粒彈珠是向後運動的。哪一粒彈珠可以自由地向後運動？

只有 A_0 的後面沒有任何東西。不妨試試：A_0 以速度 p 運動（可能是負數），A_{n+1} 以速度 q 運動，其餘全部靜止。守恆定律再次與質量為 m 和 λm 的兩個物體碰撞時相同，所以

$$p = -\frac{\lambda - 1}{\lambda + 1}\,u, \qquad q = \frac{2}{\lambda + 1}\,u.$$

A_0 反彈，A_{n+1} 向前運動，中間的彈珠保持靜止，一切並無矛盾。這就完全像是 A_0 直接撞上了 A_{n+1}，而中間那一列相等的彈珠充當了完美的信差。 ■

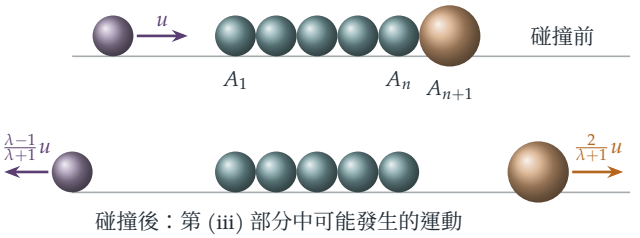


圖 10.1.1 一列相等的彈珠（藍綠色），末端是一粒較重的彈珠（琥珀色，$\lambda > 1$）。碰撞後，A_0 反彈，重彈珠離開，中間每一粒彈珠都保持靜止，完全像是 A_0 直接撞上了 A_{n+1}。

回顧

- **以守恆定律作為約束條件。**我們從來沒有仔細追蹤每一次碰撞，只是問：動量、能量和幾何條件容許哪些結果？
- **物理上的不等式十分重要。**「彈珠不能互相穿過」給出了不等式 $0 \le a \le b \le c$，而排除第 (i) 和第 (ii) 部分的，正是這些不等式，而不是守恆定律。
- **平方後相減。**把動量的平方與能量比較，會得出一個由交叉項組成的和，而這個和必須等於零。

背後的數學

為甚麼只有一個球飛出

在由相等的球組成的牛頓擺中，假設有 k 個球以速率 w 一起離開。由動量得 $kw = u$，由能量得 $kw^2 = u^2$。兩式相除，得 $w = u$ 及 $k = 1$。所以在離開的球一起運動的各種結果之中，只有「一個球以原來的速率離開」同時符合兩條定律。兩個球以一半速率離開，動量正確，但能量只有一半；一個球以兩倍速率離開，則只有當它是半個球時動量才正確。

但問題並未就此解決。若有 n 個球，就有 n 個未知的速度，而守恆定律只有兩條，所以同時符合兩者的結果有無窮多個（試找出一個三個球以不同速率運動的結果）。實際出現哪一個，取決於各球在接觸期間如何變形。1980 年代初，Frank Herrmann 和他的同事證明，我們熟悉的現象有賴於衝擊沿着隊列傳遞時不會擴散開去，而堅硬的鋼球之間留有極小空隙時正是如此。若用軟球，或者把球緊緊壓在一起，牛頓擺的表現就沒有那麼乾淨利落；仔細的量度更顯示，其他的球總會稍為移動。

守恆定律與對稱

守恆定律從何而來？1918 年，諾特（Emmy Noether）證明了一條定理，它是現代物理學的基石之一：物理系統的每一種連續對稱（continuous symmetry），都對應一個守恆量（conserved quantity）。如果物理定律在這裏與在左邊一米處相同（空間平移對稱），動量就守恆。如果物理定律今天與明天相同（時間平移對稱），能量就守恆。如果物理定律在每個方向都相同（旋轉對稱），角動量（angular momentum）就守恆。你會在大學二年級的拉格朗日力學（Lagrangian mechanics）課程中學到諾特定理（Noether's theorem）；本章自始至終都在運用它的推論。

惠更斯、雷恩與沃利斯

1668 年，英國皇家學會（Royal Society）徵求物體碰撞問題的解答。沃利斯（John Wallis）處理了非彈性碰撞（inelastic collisions），而雷恩（Christopher Wren）和惠更斯（Christiaan Huygens）則各自獨立地解決了彈性（elastic）的情況。惠更斯的關鍵洞見是轉換參考系（frame）：在一艘隨質心（centre of mass）運動的船上觀察碰撞，這時兩個相等的球只是以相反方向、原來的速率彈回，然後再變換回岸上的觀點。他還注意到某一個量，即各質量乘以速率平方之和，在碰撞前後相同；不過要再過一個世紀，動能本身才被確認為一個守恆量。

思維技巧

利用守恆量排除不可能的情況

當一個過程複雜得無法追蹤時，列出不會改變的量，看看它們容許哪些結果。

物理學家正是這樣推斷粒子加速器中的碰撞：當中相互作用的細節並不清楚，但動量、能量和電荷必須平衡。再配合一個不等式（這裏是彈珠不能互相穿過），就可以證明某個結果不可能發生。

延伸練習

1 **同一個問題再現。**試做 STEP I 2006，第 11 題（`step.maths.org/questions/06-s1-q11`）：槽中有一列粒子，末端是較重的一粒。哪些最終狀態是可能的？

2 **追蹤各次碰撞。**把第 (iii) 部分的撞擊看成一連串獨立的兩粒彈珠碰撞，由 A_0 撞上 A_1 開始。逐步追蹤這一連串碰撞，並證明最後得到的正是第 (iii) 部分所求得的運動。

3 **兩個球撞入。**在一個由五個相等的球組成的牛頓擺中，兩個球同時釋放，以速率 u 撞上其餘三個。證明「兩個球以速率 u 離開」與兩條守恆定律並無矛盾，並找出另一個同樣符合的結果。

4 **末端較輕。**若 $\lambda < 1$，追蹤一連串的兩粒彈珠碰撞。除了 A_n 和 A_{n+1} 之外，最後還有其他彈珠在運動嗎？

延伸閱讀

- **David Morin, *Introduction to Classical Mechanics***（Cambridge University Press），第 5 章。動量守恆與能量守恆，附有大量關於碰撞的習題。
- **Frank Herrmann & Peter Schmälzle, "Simple explanation of a well-known collision experiment"**，*American Journal of Physics* 49 (1981)。解釋牛頓擺為何行得通。
- **Dwight Neuenschwander, *Emmy Noether's Wonderful Theorem***（Johns Hopkins University Press）。深入淺出地講述對稱與守恆之間的聯繫。

10.2 疊球效應

題目

STEP II 2010，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/10-s2-q10`

引子

把一個網球放在籃球上面，然後讓兩個球一起落下。籃球從地面反彈，網球隨即向上飛射，升到原來下落高度的好幾倍，有時甚至高得撞上天花板。這個實驗請到戶外去做。

題目的第 (i) 部分說明，單獨一次碰撞有一個速率上限：無論撞過來的球有多重，一個靜止的球被撞之後，速率都不可能超過來球速率的兩倍。第 (ii) 部分則說明怎樣打破這個上限：讓運動沿着一串愈來愈輕的球傳遞下去，每傳一步，速率都會再乘上一個倍數。把太空探測器送往外行星，靠的正是同一個想法；也有人提出，超新星（supernova）爆發的機制當中有一部分也是如此。

動筆之前

想一想 根據 Section 10.1：若質量為 M、以速率 u 運動的物體與質量為 m 的靜止物體發生彈性碰撞（elastic collision），碰撞後 m 的速率是多少？

那一節的結果（把兩個質量的角色對調）給出被撞粒子的速率為 $\frac{2M}{M+m}u$。還有另一個記法：在彈性碰撞中，相對速度（relative velocity）會反向。如果重球幾乎沒有減慢，輕球離開時的速率就必定比重球的速率大約多出 u，也就是大約 $2u$。

逐步解題

階段 1　速率上限

取 $M = bm$，碰撞後 A 的速率為

$$v_A = \frac{2bm}{bm + m}\,u = \frac{2b}{b+1}\,u.$$

想一想 $\frac{2b}{b+1}$ 與 2 相比，哪個較大？當 $b \to \infty$ 時會怎樣？

由於對任何 b 都有 $\frac{2b}{b+1} = 2 - \frac{2}{b+1} < 2$，所以 A 的速率總是小於 $2u$，無論 B 有多重。當 $b \to \infty$ 時，這個速率趨近 $2u$：一個無限重的球會把 A 彈開，就像 A 撞上一堵移動中的牆一樣。 ■

階段 2　沿着一串粒子傳遞

想一想 在第 (ii) 部分，每個粒子的質量都是下一個粒子的 λ 倍。對於這串粒子中的每一次碰撞，第 (i) 部分告訴了我們甚麼？

B_1（質量 $\lambda^n m$）撞上 B_2（質量 $\lambda^{n-1} m$），後者的質量是前者的 λ 分之一。在第 (i) 部分取 $b = \lambda$，可知 B_2 以速率 $\frac{2\lambda}{\lambda+1}u$ 離開。接着 B_2 撞上 B_3，後者的質量同樣是前者的 λ 分之一，於是 B_3 以速率 $\left(\frac{2\lambda}{\lambda+1}\right)^2 u$ 離開。經過 n 次這樣的碰撞之後（最後一次是 B_n 撞上 A），

$$v_A = \left(\frac{2\lambda}{\lambda+1}\right)^n u.$$

（各個 B 之間其後再發生的碰撞，只會把 A 推得更快，不會使它減慢，因為從來沒有東西從前方撞向 A。）

想一想 $\frac{2\lambda}{\lambda+1}$ 比 1 大還是比它小？

由於 $\lambda > 1$，故 $2\lambda > \lambda + 1$，因此比值 $r = \frac{2\lambda}{\lambda+1}$ 大於 1，而且當 $n \to \infty$ 時 $r^n \to \infty$。所以只要 n 取得夠大，A 的速率要多快就有多快。 ■

階段 3　快二十倍

取 $\lambda = 4$，則 $r = \frac{8}{5} = 1.6$，我們需要 $1.6^n > 20$。取以 10 為底的對數（logarithm），

$$\log_{10} 1.6 = \log_{10} 16 - 1 = 4\log_{10} 2 - 1 \approx 0.20412, \qquad \log_{10} 20 = 1 + \log_{10} 2 \approx 1.30103,$$

所以我們需要 $n > \frac{1.30103}{0.20412} \approx 6.37$。最小的值是 $n = 7$。（事實上 $1.6^6 \approx 16.8$，而 $1.6^7 \approx 26.8$。） ■

回顧

- **把單一結果反覆套用。**第 (i) 部分的結果在第 (ii) 部分用了 n 次，把有界的增益變成無界的增益。

- **永遠達不到的極限。**速率可以趨近上界 $2u$，卻永遠達不到，所以正確的寫法是嚴格不等式（strict inequality）。
- **對數把乘方化為乘積。**「需要多少步」這個問題，變成了一道除法。

背後的數學

籃球與網球

設兩個球從高度 H 一起落下，到達地面時的速率為 $U = \sqrt{2gH}$。假設一切碰撞都是完全彈性的，而且籃球比網球重得多。籃球從地面反彈，以速率 U 向上運動，隨即遇上仍以速率 U 向下運動的網球。在籃球的參考系（frame）中，網球以 $2U$ 的速率迎面而來，撞上一個非常重的物體後反彈，再以 $2U$ 的速率離開。回到地面參考系，那就是以 $2U + U = 3U$ 的速率向上運動（圖 10.2.1）。由於上升高度與拋射速率的平方成正比，網球會升到 $9H$ 的高度。

現實中的球並不是完全彈性的。若恢復系數（coefficient of restitution）為 e，質量比為 M/m，網球離開時的速率就是 $\frac{Me(2+e)-m}{M+m}U$。練習中的 STEP I 2014 Q10，就是由量度所得的高度反推質量比。

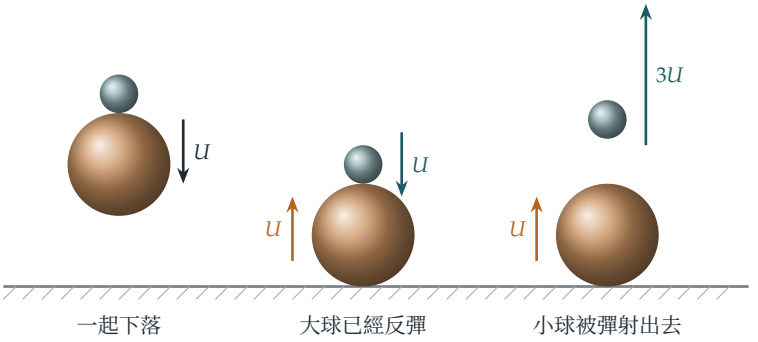


圖 10.2.1 疊球效應：兩個球都是完全彈性的，而且下面的球重得多。大球以速率 U 向上反彈，遇上以速率 U 落下的小球。相對於大球，小球以 $2U$ 的速率到達，又以 $2U$ 的速率離開，所以在地面參考系中它以 $3U$ 的速率離開，升到原來高度的九倍。

引力彈弓

太空探測器飛掠行星時並沒有碰到行星，但引力的作用就像與一個非常重的物體發生一次非常柔和的彈性碰撞。在行星的參考系中，探測器繞過行星，離開時的速率與到達時相同。如果行星正以速率 V 朝探測器運動，那麼在太陽的參考系中，探測器的速率最多可以增加 $2V$，與網球的情況完全一樣。1977 年發射的兩艘航行者號（*Voyager*）探測器，

利用了外行星一次罕見的排列，從木星飛往土星（航行者 2 號（*Voyager 2*）更繼續飛往天王星和海王星），每次與行星相遇都獲得加速。假如沒有這些引力助推（gravity assist），前往海王星的旅程就要花上大約三十年，而不是十二年。能量來自行星，而行星因此減慢的幅度小得無法量度。

彈力球與超新星

1968 年，W. R. Mellen 在 *American Journal of Physics* 上描述了一種「彈力球反彈拋射體」（superball rebound projectile）：把幾個由大至小的球疊起來，最頂的球可以用撞擊速率的許多倍彈射出去。天體物理學家指出，這與核心塌縮超新星（core-collapse supernova）有一點粗略的相似之處：恆星向內墜落的外層物質撞上堅硬而緻密的核心後反彈，衝擊波（shock wave）把動量（momentum）向外帶進密度愈來愈低的物質，把這些物質加速到極高的速率。疊球效應的示範常用來向學生解釋這個想法，不過真實的超新星所涉及的物理比彈跳的球多得多，中微子（neutrino）就是其中之一。

思維技巧

把單一結果反覆套用

如果一個過程只是把一個簡單步驟重複多次，就先把這一步解決一次，然後把它複合起來。

每一步的增益因子都小於 2，但經過很多步之後，增益卻變成無界，因為這些因子是相乘的。到了大學，這就是轉移矩陣（transfer matrix）、多級放大器和多級火箭（Section 10.3）背後的想法：一條鏈的增益，等於鏈上各個環節的增益的乘積。

延伸練習

1 **真實的情況。**試做 STEP I 2014，第 10 題（`step.maths.org/questions/14-s1-q10`）：這是帶有真實數據、反彈並不完美的疊球問題。

2 **非彈性的鏈。**試做 STEP I 2017，第 10 題（`step.maths.org/questions/17-s1-q10`）。這一題讓碰撞沿着一串質量按幾何級數增長的粒子傳遞下去。在甚麼情況下，每個粒子都恰好參與兩次碰撞？

3 **三個球。**把三個完全彈性的球疊起來，每個球都比它下面的球輕得多。證明最頂的球以 $7U$ 的速率離開，並升到 $49H$ 的高度。若有 k 個球，證明最頂的球以 $(2^k - 1)U$

的速率離開。

4 **能量從何而來？**在第 (ii) 部分，證明即使 A 的速率無界，總動能（kinetic energy）仍然守恆。那些重的粒子會怎樣？

延伸閱讀

- **David Morin, *Introduction to Classical Mechanics***（Cambridge University Press），第 5 章，當中包括疊球問題。
- **W. R. Mellen,「Superball rebound projectiles」**, *American Journal of Physics* 36 (1968)。篇幅短小，讀來有趣。
- **Gary Flandro,「Fast reconnaissance missions to the outer solar system utilizing energy derived from the gravitational field of Jupiter」**, *Astronautica Acta* 12 (1966)。這篇論文提出了航行者號的「行星大巡遊」（grand tour）。

10.3 火箭與槍炮

題目

STEP III 2017，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/17-s3-q11`

引子

太空中的火箭沒有任何東西可以借力。它靠把質量向後拋出而前進，其餘的就交給動量守恆（conservation of momentum）。一輛從車尾開炮的貨車，就是一枚拆成一塊塊的火箭：每枚拋射體（projectile）都是一團噴出的廢氣。

那麼，這些炮應該同時發射，還是逐一發射？這道題給出了明確的答案，但要看清楚它比較的是甚麼。同時發射會令貨車得到較大的速率，但也消耗較多能量，所以在兩種情況下，這些炮所做的並不是同一件事。只要改變保持不變的量，答案就會逆轉。把這一點弄清楚，便會引出太空飛行中最重要的方程——齊奧爾科夫斯基（Tsiolkovsky）的*火箭方程*（rocket equation），也會明白為甚麼真正的火箭要分成多級。

動筆之前

想一想 某一門炮發射之前，貨車以 V 行駛。發射之後，拋射體的速度是 $V - v$（即以速率 $v - V$ 向後運動）。拋射體*相對於發射前的貨車*的速度是多少？

答案是 $(V - v) - V = -v$。所以每門炮都以速率 v 把拋射體向後拋出，這個速率是相對於貨車在發射前一刻的速度而言的。這就是本題的固定規則；請牢記它，因為答案正是由它決定的。

逐步解題

階段 1　同時發射

系統最初靜止，動量為零。發射之後，貨車（質量為 M）以 u 向前運動，而 n 枚拋射體則以速率 v 向後運動（因為發射前 $V=0$）。所以

$$Mu - nmv = 0, \qquad u = \frac{nmv}{M}.$$

所獲得的動能（kinetic energy）為

$$K = \tfrac{1}{2}Mu^2 + \tfrac{1}{2}nmv^2 = \tfrac{1}{2}\frac{n^2m^2v^2}{M} + \tfrac{1}{2}nmv^2 = \tfrac{1}{2}nmv^2\Big(1 + \frac{nm}{M}\Big). \qquad \blacksquare$$

階段 2　逐一發射

想一想　在第 r 門炮發射前的一刻，貨車連同餘下的拋射體的質量是多少？速率又是多少？寫出第 r 次發射的動量守恆式。

第 r 次發射之前，貨車載有 $n-r+1$ 枚拋射體，所以它的質量是 $M+(n-r+1)m$，速率是 u_{r-1}。發射之後，載有 $n-r$ 枚拋射體的貨車速率為 u_r，而射出的拋射體速度為 $u_{r-1}-v$。動量守恆：

$$\big(M+(n-r+1)m\big)u_{r-1} = \big(M+(n-r)m\big)u_r + m(u_{r-1}-v).$$

把含 mu_{r-1} 的項整理後，得 $\big(M+(n-r)m\big)(u_r-u_{r-1}) = mv$，這就是 $(*)$。 $\blacksquare$

把各個增量相加，

$$u_n = \sum_{r=1}^{n}\frac{mv}{M+(n-r)m} = \frac{mv}{M} + \frac{mv}{M+m} + \cdots + \frac{mv}{M+(n-1)m}.$$

第一項之後的每一項都小於 $\frac{mv}{M}$，而且 $n>1$，所以 $u_n < \frac{nmv}{M} = u$。 $\blacksquare$

停一停，試一試　為甚麼在這裏逐一發射的效果較差？第一門炮發射時，它要加速的不只是貨車，還有那 $n-1$ 枚未發射的拋射體，所以它帶來的速率增幅，比起作用在空車上時要小。

階段 3　能量的帳單

想一想　寫出第 r 次發射時動能的變化。$(*)$ 怎樣幫助你化簡含 $u_r^2 - u_{r-1}^2$ 的那一項？

動能的變化為

$$K_r - K_{r-1} = \tfrac{1}{2}\big(M + (n-r)m\big)u_r^2 + \tfrac{1}{2}m(u_{r-1} - v)^2 - \tfrac{1}{2}\big(M + (n-r+1)m\big)u_{r-1}^2.$$

最後兩項合併成 $\frac{1}{2}m(v^2 - 2vu_{r-1}) - \frac{1}{2}\big(M + (n-r)m\big)u_{r-1}^2$。然後，把平方差因式分解，並利用 $(*)$，

$$\tfrac{1}{2}\big(M + (n-r)m\big)(u_r^2 - u_{r-1}^2) = \tfrac{1}{2}mv(u_r + u_{r-1}).$$

所以

$$K_r - K_{r-1} = \tfrac{1}{2}mv(u_r + u_{r-1}) + \tfrac{1}{2}mv^2 - mvu_{r-1} = \tfrac{1}{2}mv^2 + \tfrac{1}{2}mv(u_r - u_{r-1}). \qquad \blacksquare$$

由 $r = 1$ 加到 n，增量 $u_r - u_{r-1}$ 逐項相消（telescope），餘下 u_n：

$$K_n = \tfrac{1}{2}nmv^2 + \tfrac{1}{2}mvu_n. \qquad \blacksquare$$

最後，$K = \frac{1}{2}nmv^2 + \frac{1}{2}nmv^2 \cdot \frac{nm}{M} = \frac{1}{2}nmv^2 + \frac{1}{2}mvu$，而 $u_n < u$，所以 $K_n < K$。 $\blacksquare$

回顧

- **每一步都用動量。**每次發射都是一個小型的二體問題：一邊是貨車連同所載的東西，另一邊是一枚拋射體。
- **裂項求和。**$(*)$ 的增量和能量的增量都能乾淨利落地加起來。
- **仔細閱讀結論。**同時發射得到較大的速率，但要付出較多能量。題目比較的是兩個做功多少並不相同的過程。

背後的數學

應該讓甚麼保持不變？

真正的炮或火箭引擎，是以相對於*自身*的固定速率拋出拋射體的，而這個速率是在拋射體剛離開之後量度的。本題的參照對象卻是貨車在發射*之前*的速率。對一枚細小的拋射體來說，兩者相差極微；但當 n 枚拋射體同時發射時，差別就很大：在第 (i) 部分，一大批沉重的拋射體以速率 v 被向後拋出，彷彿貨車完全沒有後座（recoil）一樣，這需要大量能量。

假如改為每枚拋射體都以相對於發射*之後*的貨車的速率 v 離開，結論就會逆轉。同時發射得 $u = \frac{nmv}{M+nm}$；逐一發射得 $u_n = \sum_{k=1}^{n} \frac{mv}{M+km}$，後者較大，因為每一項都至少是 $\frac{mv}{M+nm}$。這時逐一發射勝出：每枚拋射體都是從一輛已在運動的貨車上蹬出去的，而且越後發射的炮，要加速的質量越少。（請在練習中驗證這一點。）

火箭方程

現在讓拋射體變得極小而數目極多，並保持總質量 $\mu = nm$ 不變：這就是一枚連續燃燒燃料的火箭。u_n 的和變成一個積分（兩種約定給出相同的極限）：

$$u_n = \sum_{k=0}^{n-1} \frac{mv}{M+km} \;\to\; \int_0^{\mu} \frac{v\,\mathrm{d}s}{M+s} = v \ln \frac{M+\mu}{M}.$$

這就是*齊奧爾科夫斯基火箭方程*：若火箭噴出的廢氣以相對於火箭的速率 v 離開，火箭所增加的速率為

$$\Delta u = v \ln \frac{m_{初}}{m_{末}}.$$

圖 10.3.1 比較了這三個答案。麻煩就出在對數（logarithm）上：若排氣速率約為 4.5 km/s（優良的化學燃料），要達到約 7.8 km/s 的軌道速率，質量比必須是 $e^{7.8/4.5} \approx 5.7$；一旦把空氣阻力和重力計算在內，這個比值更接近 10 或以上。火箭幾乎全是燃料。

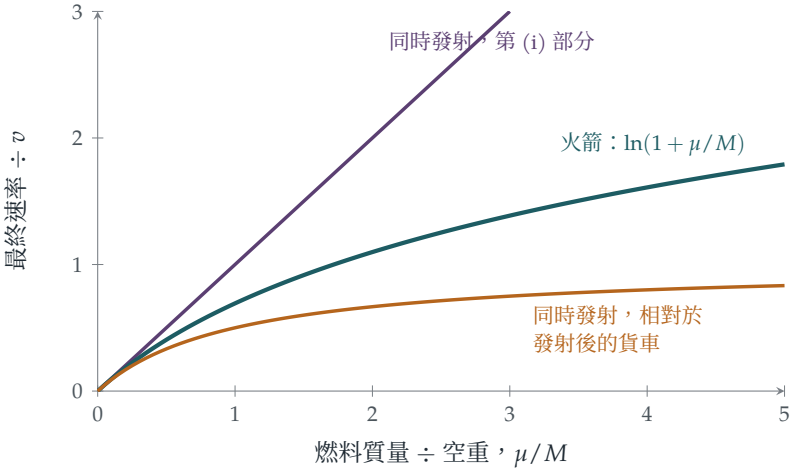


圖 10.3.1　最終速率與燃料質量對空重之比的關係。按題目的約定，全部同時發射得 μ/M（紫紅色）；但若排氣速率是相對於後座中的貨車量度，則只得 $\frac{\mu/M}{1+\mu/M}$（琥珀色）。連續發射的火箭介乎兩者之間，為 $\ln(1+\mu/M)$（藍綠色）。這幅圖所表達的就是不等式 $\frac{x}{1+x} \le \ln(1+x) \le x$。

為甚麼火箭要分級

對數也解釋了分級（staging）的道理。火箭若在升空途中拋掉用完的燃料箱，就不必再加速這些無用的質量。各級所得的速率增量是*相加*的（就像那些炮一樣），而每一級的增量則取決於該級質量的*比*，所以把火箭分成多級，就能達到總質量相同的單級火箭無法達到的速率。把太空人送上月球的土星五號（Saturn V）有三級；它的第一級在不足三分鐘內燒掉約 2000 公噸燃料。

齊奧爾科夫斯基

康斯坦丁·齊奧爾科夫斯基（Konstantin Tsiolkovsky）是俄羅斯外省小城卡盧加（Kaluga）一位失聰的教師，他在 1903 年一篇題為「利用反作用裝置探索外太空」的論文中發表了火箭方程。多級火箭、液態氫和液態氧燃料，以及太空站，他都一一推算過，比這些東西真正造得出來早了數十年。這條方程其實更早就在一些較不起眼的地方出現過（William Moore 在 1810 年已推導出其中一個版本），但齊奧爾科夫斯基是第一個看出它對太空飛行意味着甚麼的人。他寫道：「地球是人類的搖籃，但人不能永遠活在搖籃裏。」

思維技巧

從離散過程到極限

連續的過程，往往可以理解爲許多細小離散步驟的極限（*limit*）。先把離散的問題精確地解出來，再讓步長縮小：和就變成積分。

在這裏，n 門炮變成了一枚火箭，一個和變成了一個對數。到了大學，每逢推導物理定律，你都會見到同一手法：先分析一個微小的單元，再把結果加起來，成為一個積分或一條微分方程（differential equation）。

延伸練習

1 **另一種約定。**假設每枚拋射體都以相對於發射之後的貨車的速率 v 離開，重做第 (i) 和第 (ii) 部分。證明這時逐一發射所得的最終速率較大。

2 **塵埃雲中的彗星。**試做 STEP III 2009，第 11 題（`step.maths.org/questions/09-s3-q11`）。它和火箭一樣是變質量（variable mass）問題，但彗星的質量是增加而不是減少。留意它的速率趨向甚麼值。

3 **兩級火箭。**一枚火箭的排氣速率為 v，空重為 M，燃料質量為 $9M$。求它的最終速率。現在把它分成兩級，每級的空重為 $\frac{1}{2}M$，燃料為 $4.5M$，第一級在燃料用盡後拋掉。比較兩個最終速率。

4 **不等式。**在 $[0, x]$ 上把 $\frac{1}{1+s}$ 分別與 1 和 $\frac{1}{1+x}$ 比較，從而證明當 $x \geq 0$ 時，$\frac{x}{1+x} \leq \ln(1+x) \leq x$。

延伸閱讀

- **Daniel Kleppner & Robert Kolenkow, *An Introduction to Mechanics***（Cambridge University Press），第 4 章。動量、變質量和火箭方程。
- **George Sutton & Oscar Biblarz, *Rocket Propulsion Elements***（Wiley）。工程學上的標準教科書；讀完本節後，開首幾章已可讀懂。
- **Hermann Oberth, *Ways to Spaceflight***（NASA 譯本，可在網上免費閱讀）。一位先驅的親身記述。

10.4 一個分子與一個活塞

題目

STEP I 2002，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/02-s1-q10`

引子

替單車輪胎打氣時摸一摸氣泵：它會發熱。壓縮氣體會令氣體變暖，柴油引擎（單靠壓縮就能點燃燃料）、雪櫃以至恆星的形成，背後都是這個原理。但為甚麼壓縮會令氣體升溫？溫度就是分子的平均動能（kinetic energy），所以問題其實是：為甚麼分子會加速？

這道題目只用一個分子就回答了這個問題。分子每次從推進中的活塞（piston）反彈，都會多得一點速率，就像網球碰上揮動中的球拍一樣。把每次反彈準確地追蹤下去，便會發現一條隱藏的定律：分子的速率乘以氣缸的長度，幾乎保持不變。這就是一個絕熱不變量（adiabatic invariant），也是你在物理課遇到的氣體定律在微觀層面的來源。

動筆之前

想一想 一個以速率 w 運動的球，撞上一堵以速率 u 迎面移來的牆，並作彈性（elastic）反彈。反彈後球的速率是多少？（試在牆的參考系中思考。）

在牆的參考系（frame）中，球以 $w+u$ 接近；由於它是從一個極重的物體上彈性反彈，所以亦以 $w+u$ 離開。回到地面參考系，牆正以 u 朝着與反彈後的球相同的方向運動，所以球的速率是 $w+2u$。每次從推進中的活塞反彈，速率都增加 $2u$。（從氣缸靜止的一端反彈，則甚麼也不會改變。）

逐步解題

階段 1　反彈 n 次後的速率

質點（particle）起初的速率是 v。每次與活塞碰撞（collision），速率增加 $2u$，而與末端碰撞則不會改變速率。所以與活塞碰撞 n 次之後，

$$v_n = v + 2nu. \qquad \blacksquare$$

階段 2　長度的遞推關係

> **想一想**　在第 n 次與第 $(n+1)$ 次碰撞之間，質點由活塞走到末端再折返。它走了多遠？活塞又移動了多遠？

在時間 t_n 內，活塞移動了 ut_n，所以 $d_n - d_{n+1} = ut_n$。與此同時，質點以速率 v_n 由活塞走到末端（距離 d_n），再返回活塞（距離 d_{n+1}），所以 $d_n + d_{n+1} = v_n t_n$。把兩條方程相除，便可消去 t_n：

$$\frac{d_n + d_{n+1}}{d_n - d_{n+1}} = \frac{v_n}{u}, \qquad \text{所以} \qquad d_{n+1}(v_n + u) = d_n(v_n - u).$$

代入 $v_n = v + 2nu$，得

$$d_{n+1} = \frac{v + (2n-1)u}{v + (2n+1)u}\, d_n. \qquad \blacksquare$$

這就是所求的遞推關係（recurrence）。

階段 3　裂項相消

> **想一想**　把 d_n 寫成 d_1 乘以 $n-1$ 個這樣的分數之積。哪些項會互相抵消？

$$d_n = d_1 \cdot \frac{v+u}{v+3u} \cdot \frac{v+3u}{v+5u} \cdots \frac{v+(2n-3)u}{v+(2n-1)u} = \frac{v+u}{v+(2n-1)u}\, d_1. \qquad \blacksquare$$

每個分子都與前一個分母抵消，只剩下第一個分子和最後一個分母。

階段 4　$v = u$ 的情況

這時 $d_n = \frac{2u}{2nu} d_1 = \frac{d_1}{n}$，而

$$ut_n = d_n - d_{n+1} = d_1\Big(\frac{1}{n} - \frac{1}{n+1}\Big) = \frac{d_1}{n(n+1)}. \qquad \blacksquare$$

把這些加起來，$u(t_1 + t_2 + \cdots) = d_1\big(1 - \tfrac{1}{2} + \tfrac{1}{2} - \tfrac{1}{3} + \cdots\big) = d_1$：活塞在第一次碰撞後經過時間 d_1/u 到達氣缸末端，這正是理所當然的結果。在活塞到達之前，質點反彈了無限多次（圖 10.4.1）。

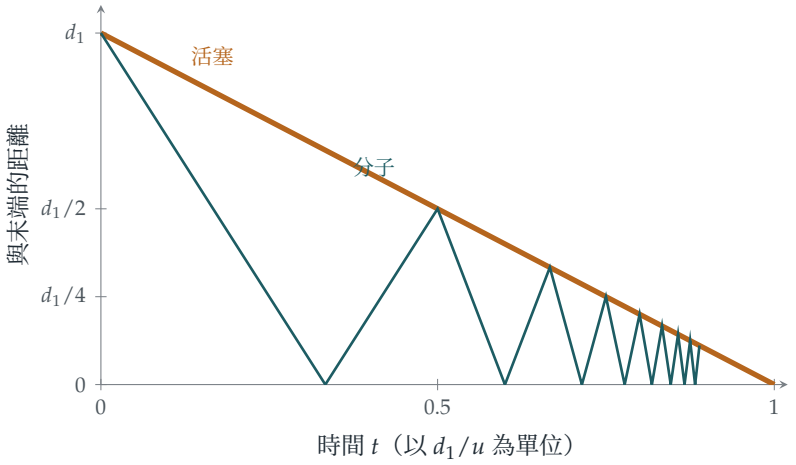


圖 10.4.1 $v = u$ 情況下的時空圖：與氣缸末端的距離對時間。活塞（琥珀色）穩步推進；分子（藍綠色）在活塞與末端之間來回反彈。它的鋸齒形軌跡愈來愈陡（愈來愈快），也愈來愈短。第 n 次與活塞的碰撞發生在距離 d_1/n 處，所以分子的速率大致與 1/(長度) 成正比。

回顧

- **處理碰撞時轉換參考系。**在活塞的參考系中，反彈再簡單不過。
- **兩條方程，消去未知量。**質點與活塞在同一段時間內所走的距離，給出了一個不含 t_n 的遞推關係。
- **裂項相消（telescoping）。**那個乘積一下子化簡了，正如 $\frac{1}{n(n+1)}$ 之和一樣。

背後的數學

一個絕熱不變量

把速率乘以長度：

$$v_n d_n = (v + 2nu) \cdot \frac{v+u}{v+(2n-1)u}\, d_1 \approx (v+u)d_1$$

這在 n 很大時成立，在 u 遠小於 v 時亦然。所以

$$\text{速率} \times \text{長度} \approx \text{常數}.$$

如果活塞相對於分子移動得很慢，那麼無論活塞移動多遠，乘積 vd 都幾乎不變。像這樣在系統緩慢改變時幾乎保持不變的量，稱為*絕熱不變量*。一般而言，往復運動的絕熱不變量是 $\oint p\,\mathrm{d}q$，即動量（momentum）沿一個週期的積分；在這裏它就是 $mv \times 2d$。

從一個分子到一團氣體

由於 $v \propto 1/d$，動能 $\frac{1}{2}mv^2$ 與 $1/d^2$ 成正比。對於由許多沿同一方向運動的分子組成的氣體，溫度與平均動能成正比，所以 $T \propto 1/d^2$：長度減半，溫度便變成四倍。在三維空間中，能量由三個運動方向分擔，而同樣的論證（費曼（Feynman）在他的 *Lectures* 中給出了這個論證）可推出單原子氣體（monatomic gas）的絕熱壓縮定律，

$$TV^{2/3} = \text{常數}, \qquad \text{或等價地} \qquad pV^{5/3} = \text{常數}.$$

指數 $\frac{5}{3}$ 就是單原子氣體著名的比熱比（ratio of specific heats）γ。把單車氣泵內的空氣壓縮至原來體積的五分之一，如果做得夠快而且沒有熱量散失，其絕對溫度會升高至原來的 $5^{0.4} \approx 1.9$ 倍（空氣分子有兩個原子，因而 $\gamma = 1.4$）：由 20°C 升至約 280°C。氣泵之所以沒有那麼熱，只是因為壓縮得慢，熱量又會散走。

費米的宇宙射線

1949 年，恩里科·費米（Enrico Fermi）提出，宇宙射線（cosmic rays）這些從太空而來、能量極高的粒子，其加速方式與本題的分子如出一轍：它們在星際氣體雲中運動的磁場上反覆反彈，每次獲得一點能量。現時認為，「費米加速」（Fermi acceleration）在超新星殘骸的激波處發生，粒子在那裏來回穿越激波，就像分子在活塞與牆壁之間往返一樣。絕熱不變量在量子力學的誕生中也扮演過一個角色：保羅·埃倫費斯特（Paul Ehrenfest）在 1916 年指出，被量子化的量必定是絕熱不變量，因為緩慢的改變不可能令系統在量子態之間躍遷。

思維技巧

從物理過程建立遞推關係

當一個過程不斷重複時，把其中一個週期準確地描述出來，並寫下狀態怎樣由一個週期變到下一個週期。

這裏的狀態是每次碰撞時的一對量（速率，長度）。這個遞推關係可以精確地解出，而它的解揭示了一個近似守恆的量。在大學裏，同樣的方法（先考慮一個週期，然後得出遞推關係或「回歸映射」（return map））會用於軌道、振子，以及 Section 8.7 的迭代映射。

延伸練習

1 **不變量的精確形式。**證明 $v_n d_n - (v+u)d_1 = \frac{u(v+u)d_1}{v+(2n-1)u}$，並由此推出當 $n \to \infty$ 時 $v_n d_n \to (v+u)d_1$。

2 **把活塞拉出。**設活塞改為以速率 $u < v$ 向外抽出，把題目重做一次。證明質點會減速，並求 $v_n d_n$。（氣體膨脹時會冷卻。）

3 **能量。**在 $v = u$ 的情況下，隨着活塞逼近末端，分子的速率會無止境地增長。求第 n 個週期內活塞所受的平均力（動量的改變除以 t_n），並證明它按 n^3 的速度增長。這說明了把活塞一直推到盡頭所需的功有甚麼特點？

4 **彈跳的球。**一個球在重力作用下於地板上作彈性彈跳，而地板被緩慢地升高。試猜出這個運動的一個絕熱不變量。（提示：與 vd 相對應的量是平均速率乘以彈跳的高度，或等價地，能量乘以兩次彈跳之間的時間。）

延伸閱讀

- **Richard Feynman, Robert Leighton & Matthew Sands, *The Feynman Lectures on Physics*, Vol. I**（《費曼物理學講義》第一卷，可在網上免費閱讀），第 39 章「The kinetic theory of gases」（氣體動力論）。同一個活塞論證，並推廣至三維。
- **Stephen Blundell & Katherine Blundell, *Concepts in Thermal Physics***（Oxford University Press）。大學一年級程度的絕熱過程與氣體動力論。
- **Lev Landau & Evgeny Lifshitz, *Mechanics***（Butterworth-Heinemann），§49。對絕熱不變量的精簡論述，留待日後閱讀。

10.5 有限時間內的無限次反彈

題目

STEP III 2002，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/02-s3-q11`

引子

埃利亞的芝諾（Zeno of Elea）曾經論證，你永遠走不到一堵牆：你必須先走完一半的距離，然後再走完餘下距離的一半，如此類推，永無止境。化解這個悖論的關鍵，在於無窮多步所需的總時間可以是有限的，也就是一個收斂（converge）的無窮級數（infinite series）。這道題目正是芝諾悖論（Zeno's paradox）的真實版本：一個冰球滑進楔形的角落，在兩面牆之間反彈無窮多次，卻在有限時間內到達角落。

大自然也會這樣做。一個球在地板上彈跳，若恢復系數（coefficient of restitution）小於 1，它會在有限時間內彈跳無窮多次，然後靜止下來。在顆粒物質（沙、穀物、雪）的電腦模擬中，這種「非彈性坍縮」（inelastic collapse）是個實實在在的難題：電腦要在有限時間內處理無窮多次碰撞。

動筆之前

想一想 當一個質點（particle）以恢復系數 e 撞向一面光滑的牆時，速度沿牆方向的分量會怎樣？垂直於牆的分量又會怎樣？

光滑的牆不會沿着牆面施力，所以速度平行於牆的分量保持不變。垂直分量則會反向，並乘以 e。因此，若質點以與牆成 α 的角度到達，並以 β 的角度離開，則

$$\tan\beta = e\tan\alpha,$$

而由於平行分量不變，（入射速率）$\cos\alpha$ =（離開速率）$\cos\beta$。

逐步解題

階段 1　離開的角度

想一想　質點到達 M_1 時與牆 OA 成角 θ（在遠離 O 的一側量度）。它離開時與 OA 所成的角 φ 是多少？利用三角形 OM_1N_1，以 θ 表示 φ。

質點離開 M_1 時朝着 O 的一側前進，與牆成角 φ，所以角 OM_1N_1 等於 φ。在三角形 OM_1N_1 中，位於 O 的角是 30°。在 N_1，角 M_1N_1B 等於 θ，而 O、N_1、B 在同一直線上，所以角 ON_1M_1 等於 $180^\circ - \theta$。三角形的內角和為 180°：

$$30^\circ + \varphi + (180^\circ - \theta) = 180^\circ, \qquad \text{所以} \qquad \varphi = \theta - 30^\circ.$$

（特別地，$\theta > 30^\circ$。）於是由恢復定律 $\tan\varphi = e\tan\theta$ 可得

$$e = \frac{\tan(\theta - 30^\circ)}{\tan\theta},$$

而由平行分量守恆 $u_1\cos\theta = v_1\cos\varphi$ 可得

$$\frac{v_1}{u_1} = \frac{\cos\theta}{\cos(\theta - 30^\circ)}. \qquad \blacksquare$$

階段 2　規律重複出現

想一想　在 N_1，質點到達時與牆 OB 成角 θ（在遠離 O 的一側量度）。把這個情況與它到達 M_1 時的情況比較一下。接下來會發生甚麼事？

N_1 的情況是 M_1 的情況的鏡像：質點以角 θ 撞上一面牆，而這個角位於遠離 O 的一側。所以它以角 $\theta - 30^\circ$ 離開，再以角 θ 撞上 OA，如此類推，永無止境。每一次反彈都是第一次反彈按比例縮放後的複製品。三角形 OM_kN_k、ON_kM_{k+1}、... 每一個都與前一個相似（圖 10.5.1），而每次反彈時速率都乘以同一個因子。

階段 3　兩個比

想一想　由一次反彈到下一次反彈，與 O 的距離按甚麼因子 ρ 縮小？速率按甚麼因子 k 縮小？那麼每一段路程所需的時間又按甚麼因子縮小？

在三角形 OM_1N_1 中應用正弦定理（sine rule），可得

$$\rho = \frac{ON_1}{OM_1} = \frac{\sin\varphi}{\sin(180^\circ - \theta)} = \frac{\sin(\theta - 30^\circ)}{\sin\theta}.$$

下一個三角形的所有長度都是這個三角形的 ρ 倍，所以路徑的每一段都是前一段的 ρ 倍長。每次反彈時速率都乘以 $k = \frac{\cos\theta}{\cos(\theta-30^\circ)}$。時間等於長度除以速率，所以每一段所需的時間是前一段的

$$\frac{\rho}{k} = \frac{\sin(\theta-30^\circ)\cos(\theta-30^\circ)}{\sin\theta\cos\theta} = \frac{\sin(2\theta-60^\circ)}{\sin 2\theta}$$

倍。

階段 4 一個等比級數

總時間是 $T_1(1 + r + r^2 + \cdots)$，其中 T_1 是第一段所需的時間，而 $r = \frac{\sin(2\theta-60^\circ)}{\sin 2\theta}$。這是一個等比級數（geometric series），若 $0 < r < 1$，總時間便是有限的。

想一想 $\sin(2\theta - 60^\circ) < \sin 2\theta$ 在甚麼時候成立？利用恆等式 $\sin X - \sin Y = 2\cos\frac{1}{2}(X+Y)\sin\frac{1}{2}(X-Y)$。

$$\sin 2\theta - \sin(2\theta-60^\circ) = 2\cos(2\theta-30^\circ)\sin 30^\circ = \cos(2\theta-30^\circ),$$

當 $2\theta - 30^\circ < 90^\circ$，即 $\theta < 60^\circ$ 時，上式為正。另外，由於 $\theta > 30^\circ$，故 $r > 0$。所以當 $30^\circ < \theta < 60^\circ$ 時，級數收斂：質點反彈無窮多次，它與 O 的距離每次按因子 $\rho < 1$ 縮小，並在有限時間 $\frac{T_1}{1-r}$ 之後到達 O。 ■

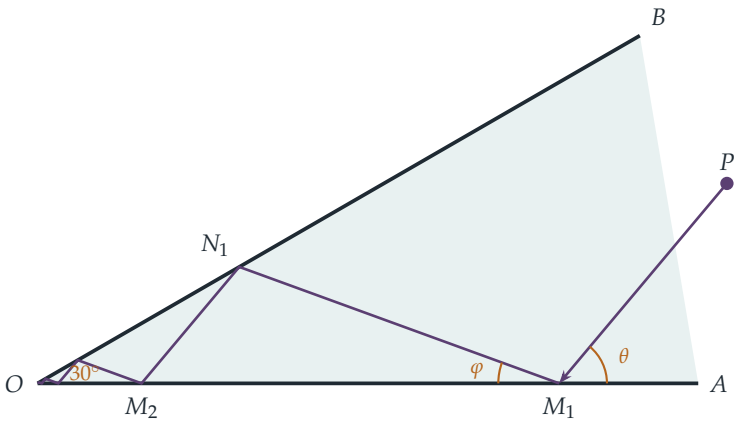


圖 10.5.1 $\theta = 50^\circ$ 時質點在楔形內的路徑，此時 $\varphi = 20^\circ$，$e = \tan 20^\circ / \tan 50^\circ \approx 0.31$。每一次反彈都是前一次的複製品，按 $\rho \approx 0.45$ 的比例縮小，而每一段所需的時間約為前一段的 0.65 倍。質點經過無窮多次反彈後，在有限時間內到達 O。

回顧

- **自相似性（self-similarity）。**一旦弄清楚一次反彈，其後每一次反彈都只是按比例縮放的複製品，所以整個無窮過程就歸結為兩個比。
- **正確地組合各個比。**距離按 ρ 縮小，速率按 k 縮小，所以時間按 ρ/k 縮小。
- **無窮多個事件，有限的總和。**公比小於 1 的等比級數收斂。

背後的數學

當 $\theta \geq 60°$ 時

當 $\theta = 60°$ 時，這個比恰好等於 1：每一段所需的時間都相同，質點雖然愈來愈接近 O，卻永遠到不了。（速率和每段的長度都按同一個因子縮小。）當 $\theta > 60°$ 時，每一段所需的時間愈來愈長。在所有情況下，質點都會趨近 O，並反彈無窮多次；只有當 $\theta < 60°$ 時，它才會在有限時間內到達。$e = \tan 30° / \tan 60° = \frac{1}{3}$ 這個值就是分界線。

彈跳的球與非彈性坍縮

同一現象最簡單的例子：一個球掉落到地板上，反彈至高度 e^2H，然後是 e^4H，如此類推。兩次彈跳之間的時間與每次彈跳時的速率成正比，而速率按因子 e 縮小，所以總時間是一個等比級數：設 $T_0 = \sqrt{2H/g}$ 為第一次下落的時間，則球在

$$T_0 + 2eT_0 + 2e^2T_0 + \cdots = T_0\,\frac{1+e}{1-e}.$$

之後停止彈跳。無窮多次彈跳，然後靜止。1992 年，Sean McNamara 和 William Young 證明了，在模擬排成一直線的多個非彈性質點時，同樣的事情也會發生：其中一組質點可以在有限時間內發生無窮多次碰撞，這就是「非彈性坍縮」。沙和顆粒流的模擬必須加入特別的規則來避免這種情況，而這個現象也有助解釋為何顆粒氣體會自發地聚集成團。

芝諾與收斂

芝諾的悖論約於公元前 450 年提出，困擾了哲學家兩千年。數學上的解答，即無窮和可以有有限的值，要到十九世紀才變得嚴謹，當時柯西（Cauchy）、魏爾斯特拉斯（Weierstrass）等人以極限（limit）來定義收斂（這是第 2 章的主題）。這道題目令人驚訝之處，在於這個過程在物理上是無窮的：確確實實有無窮多次碰撞，發生在有限時間之內。研究多邊形內的桌球運動和碰撞質點運動的數學家，把這種行為稱為芝諾現象（Zeno phenomenon）；要證明由互相碰撞的硬球組成的系統在所有時刻都有明確定

義的運動，這正是其中一個微妙之處。

思維技巧

為無窮過程求和

當一個過程以固定的縮放因子不斷重複時，任何可加的量（時間、距離、能量）的總和都是一個等比級數。根據公比判斷它是否收斂。

這裏的關鍵是找出正確的比：不是距離之比，也不是速率之比，而是時間之比。在大學裏，同樣的推理支配着迭代法(iterative methods)的收斂性、分形(fractal)理論，以及對反覆把問題減半的演算法（algorithm）的分析。

延伸練習

1 **總距離。**證明對於 30° 與 90° 之間的每一個 θ，由 M_1 到 O 所走的總距離都是有限的，並以 OM_1 和 θ 表示這個總距離。

2 **其他楔形。**把 30° 換成角 α 的楔形，重做這道題目。證明 $\varphi = \theta - \alpha$，並求出質點在有限時間內到達 O 時 θ 須滿足的條件。

3 **彈性牆。**若 $e = 1$，質點有可能到達 O 嗎？(提示：把楔形沿它的牆反覆反射，使路徑變成一條直線。圍繞 O 可以放下多少個 30° 的楔形？)

4 **彈跳的球。**一個球從 1 m 的高度掉下，$e = 0.8$。它會彈跳多久？它總共走了多遠？(取 $g = 9.8\,\mathrm{m\,s^{-2}}$。)

延伸閱讀

- **David Morin, *Introduction to Classical Mechanics***（Cambridge University Press），第 5 章。書中有多道關於彈跳的球和無窮多次連續碰撞的題目。
- **Sean McNamara & William Young, "Inelastic collapse and clumping in a one-dimensional granular medium"**（一維顆粒介質中的非彈性坍縮與聚團），*Physics of Fluids A* 4 (1992)。
- **Serge Tabachnikov, *Geometry and Billiards***（American Mathematical Society；可在網上免費閱讀)。講述桌球的數學，包括練習中用到的展開技巧。

10.6 二體化為一體

題目

STEP III 2004，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/04-s3-q11`

引子

兩個物體互相吸引或排斥：地球與月球、兩顆恆星、一個質子與一個電子。每個物體的運動都取決於另一個物體的位置，因此兩條方程是耦合（coupled）的，看來非要同時求出兩個運動不可。牛頓（Newton）的重大簡化在於更換變數：不用兩個位置，而改用質心（centre of mass）和間距（separation）。質心的運動彷彿完全不受力，而間距的運動則像一個質量經過修正的單一質點。一個難題就此化為兩個易題。

本題就兩個在直線上運動的質點，實際做一次這樣的分離。過程中還有一個小小的奇蹟：間距原來是 t 的二次式的平方根，而箇中原因與題目本身的設定毫無關係，卻完全源於平面上的運動。

動筆之前

想一想 寫出每個質點的牛頓第二定律（Newton's second law）。每個力朝哪個方向作用？

這個力是排斥力，所以它把（左邊的）P 推向左，把（右邊的）Q 推向右：

$$2\ddot{x}_1 = -\frac{2}{z^3}, \qquad 1\cdot\ddot{x}_2 = +\frac{2}{z^3}.$$

這兩條方程是耦合的，因為每一條都含有 $z = x_2 - x_1$。

逐步解題

階段 1　間距

想一想　兩條方程的哪一個組合只含有 z？

相減：$\ddot{z} = \ddot{x}_2 - \ddot{x}_1 = \frac{2}{z^3} + \frac{1}{z^3} = \frac{3}{z^3}$。 ■

間距的表現就像一個質量為 $\frac{2}{3}$ 的單一質點，受力 $\frac{2}{z^3}$ 推動。$\frac{2}{3} = \frac{2\times 1}{2+1}$ 這個數就是約化質量（reduced mass）$\frac{m_1 m_2}{m_1+m_2}$。

階段 2　間距的能量

想一想　利用 $\ddot{z} = v\frac{\mathrm{d}v}{\mathrm{d}z}$，分離變數然後積分。$z$ 和 $v = \dot{z}$ 的初始值是多少？

由 $v\frac{\mathrm{d}v}{\mathrm{d}z} = \frac{3}{z^3}$，積分得 $\frac{1}{2}v^2 = -\frac{3}{2z^2} + C$。起初 $z = 1$ 而 $\dot{z} = \dot{x}_2 - \dot{x}_1 = 0 - 1 = -1$，所以 $\frac{1}{2} = -\frac{3}{2} + C$，得 $C = 2$：

$$v^2 = 4 - \frac{3}{z^2}.$$

（這就是間距的能量守恆（conservation of energy）：以每單位約化質量計，動能加上勢能 $\frac{3}{2z^2}$ 是常數。）

階段 3　出奇簡單的解

想一想　要證明的答案是 $z^2 = 4t^2 - 2t + 1$。所以不妨考慮 $w = z^2$，而不是 z。以 w 表示，$\dot{w}^2$ 是甚麼？

設 $w = z^2$，則 $\dot{w} = 2z\dot{z}$，所以

$$\dot{w}^2 = 4z^2\dot{z}^2 = 4z^2\left(4 - \frac{3}{z^2}\right) = 16w - 12.$$

求導得 $2\dot{w}\ddot{w} = 16\dot{w}$，所以（只要 $\dot{w} \neq 0$）$\ddot{w} = 8$。因此 w 是 t 的二次式：$w = 4t^2 + \alpha t + \beta$。起初 $w = 1$ 而 $\dot{w} = 2z\dot{z} = -2$，所以 $\beta = 1$ 及 $\alpha = -2$：

$$z = \sqrt{4t^2 - 2t + 1}. \qquad ■$$

（你可以直接驗證：$\dot{w}^2 = (8t-2)^2 = 64t^2 - 32t + 4 = 16(4t^2 - 2t + 1) - 12$。）兩個質點在 $t = \frac{1}{4}$ 時最接近，此時 $z = \frac{\sqrt{3}}{2}$，其後便永遠愈離愈遠。

階段 4　質心

想一想　$2x_1 + x_2$ 滿足甚麼方程？為甚麼它是最自然要考慮的組合？

把兩條運動方程相加，$2\ddot{x}_1 + \ddot{x}_2 = -\frac{2}{z^3} + \frac{2}{z^3} = 0$。兩個力大小相等、方向相反（牛頓第三定律），所以總動量（momentum）$2\dot{x}_1 + \dot{x}_2$ 是常數。它的初始值是 $2 \times 1 + 0 = 2$，而 $2x_1 + x_2$ 的初始值是 1，所以

$$2x_1 + x_2 = 1 + 2t.$$

質心位於 $\frac{2x_1+x_2}{3}$，以速率 $\frac{2}{3}$ 勻速移動。再結合 $x_2 - x_1 = z$，

$$x_1 = \frac{1 + 2t - \sqrt{4t^2 - 2t + 1}}{3}, \qquad x_2 = \frac{1 + 2t + 2\sqrt{4t^2 - 2t + 1}}{3}. \qquad \blacksquare$$

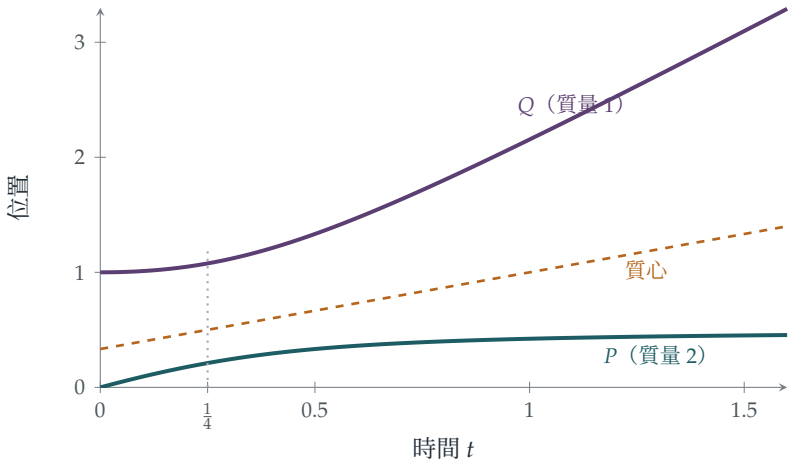


圖 10.6.1 P（藍綠色）和 Q（紫紅色）的位置隨時間的變化。P 向 Q 移動，受排斥力減慢並折返，逐漸趨近一個固定位置，而 Q 則被推開。兩者在 $t = \frac{1}{4}$ 時最接近。質心（琥珀色虛線）自始至終以恆定速率移動。

回顧

- **更換變數以解除耦合。**兩條方程之差只含有 z；加權之和只涉及質心。
- **能量積分。**寫成 $\ddot{z} = v\frac{\mathrm{d}v}{\mathrm{d}z}$，便把二階方程化為一階方程：這就是能量守恆。
- **更換未知量。**z 的方程不好處理；z^2 的方程卻簡單不過。

背後的數學

為甚麼 z^2 是二次式

對任何形如 $\ddot{z} = k/z^3$ 的力律，同樣的計算都給出 $\frac{\mathrm{d}^2}{\mathrm{d}t^2}(z^2) = 2\dot{z}^2 + 2z\ddot{z} = 2\dot{z}^2 + \frac{2k}{z^2} =$ 常數，因為 $\frac{1}{2}\dot{z}^2 + \frac{k}{2z^2}$ 正是守恆的能量。所以 z^2 必定是 t 的二次式。為甚麼？

隱藏的原因是這樣的。設想一個質點在平面上自由運動，即沿直線以恆定速率前進。它與原點的距離 r 滿足 $r^2 = $（一個二次式，變數為 t），因為 $r^2 = (x_0 + at)^2 + (y_0 + bt)^2$。而在極坐標（polar coordinates）中，它的徑向運動方程是 $\ddot{r} - r\dot{\theta}^2 = 0$；由於角動量（angular momentum）$h = r^2\dot{\theta}$ 是常數，方程變成

$$\ddot{r} = \frac{h^2}{r^3}.$$

立方反比的排斥力，正正就是隱藏的角向運動所產生的離心（centrifugal）效應。在我們的問題中，$z^2 = 4t^2 - 2t + 1 = (2t - \tfrac{1}{2})^2 + \tfrac{3}{4}$ 是某一點與原點距離的平方，該點以速率 2 沿一條直線運動，而這條直線與原點的距離為 $\tfrac{\sqrt{3}}{2}$。一維上互相排斥的兩個質點，就是二維自由運動的影子。

二體問題

只要兩個物體之間的力僅取決於它們的間距，並沿兩者的連線作用，本題的方法便適用，一維如是，三維亦如是。質心作勻速運動，而間距向量 $\mathbf{r}$ 服從 $\mu\ddot{\mathbf{r}} = \mathbf{F}(\mathbf{r})$，其中 $\mu = \frac{m_1 m_2}{m_1 + m_2}$。對引力而言，這就是開普勒問題（Kepler problem），由牛頓在《自然哲學的數學原理》（*Principia*）中解決：間距描出的軌跡是橢圓、拋物線或雙曲線。地球和太陽都繞着它們的共同質心運行，而這個質心位於太陽內部；地球和月球則繞着距離地心約 4700 km 的一點運行，這一點在地球內部，但離地心頗遠，所以地球每月都會擺動一次。天文學家正是靠偵測恆星的這種擺動，來尋找環繞其他恆星的行星。

可積系統

1969 年，Francesco Calogero 研究直線上多個質點以平方反比勢（即立方反比力）互相排斥的情況，發現無論質點數目多少，這個系統竟然都可以精確求解。Jürgen Moser 在 1975 年解釋了箇中原因，他用上一個矩陣，其特徵值（eigenvalues）在矩陣隨時間演化時保持不變：這就是 Section 7.7 的 *Lax* 對（Lax pairs）。*Calogero–Moser* 系統如今是可積系統（integrable system）的核心例子，所謂可積系統，就是守恆量多得足以完全求解的系統；它把力學與隨機矩陣、量子場論及孤立子理論聯繫起來。

思維技巧

選擇能解除耦合的坐標

當兩個量互相耦合時，找出不耦合的組合。和與差，配上適當的常數作加權，往往能把問題拆成互相獨立的部分。

質心坐標與相對坐標是第一個例子。到了大學，同一個想法會以矩陣對角化（簡正模態（normal modes），Section 11.2）、偏微分方程的分離變數法等形式出現，也見於量子力學。

延伸練習

1 **繩上的兩個質量。**試做 STEP III 2023，第 9 題（載於 OCR 提供下載的 2023 年試卷，`www.ocr.org.uk/Images/707102-2023-past-papers.zip`）。題目考察以繩相連的兩個質點，並利用質心把運動分離。留意其中一個質點何時向後移動。

2 **吸引力。**把力改為吸引力 $\frac{2}{z^3}$，重做本題。證明 z^2 仍然是 t 的二次式，並求兩個質點何時相撞。

3 **隱藏的平面。**就題目中的運動，求出平面上與原點距離為 $z(t)$ 的那條直線運動軌跡，以及掃過的角 $\theta(t)$。驗證 $z^2\dot{\theta}$ 是常數，且 $(z^2\dot{\theta})^2 = 3$。

4 **開普勒第三定律。**兩顆質量分別為 m_1 和 m_2 的恆星相距 a，在引力作用下繞它們的質心作圓周運動。證明週期 T 滿足 $T^2 = \frac{4\pi^2 a^3}{G(m_1+m_2)}$。

延伸閱讀

- **John Taylor, *Classical Mechanics***（University Science Books），第 8 章。二體問題、約化質量與軌道。
- **Daniel Kleppner & Robert Kolenkow, *An Introduction to Mechanics***（Cambridge University Press），第 10 章，論有心力。
- **Francesco Calogero, *Classical Many-Body Problems Amenable to Exact Treatments***（Springer）。給有志挑戰的讀者：精確可解系統的世界。

10.7 球殼內的引力

題目

STEP III 1999，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/99-s3-q9`

引子

試想像一顆空心的行星：一層薄薄的岩石球殼（spherical shell），裏面空無一物。假如你站在球殼內靠近殼壁的地方，也許會以為自己會被附近的岩石吸過去。事實上你會浮在原處：球殼內任何一處，球殼的引力都恰好是零。球殼較近的一側距離較短，但較遠的一側面積較大，兩種效應剛好完全抵銷。

這就是*牛頓球殼定理*（*Newton's shell theorem*）。牛頓（Newton）花了多年才把它證明出來，而它對牛頓的引力理論至為重要：連同與它配對的另一個結果（在球體之外，引力就如全部質量都集中在球心一樣），它說明了為何可以把地球和太陽當作點來處理。在練習和「背後的數學」一節中，我們會用這個定理計算一團氣體雲塌縮成恆星需要多少時間。

動筆之前

想一想 考慮圓環上兩小段位於同一直徑兩端的部分。畫出它們對質點（particle）所施的力。哪些分量（component）互相抵銷，哪些相加？

這兩小段與質點的距離相同。它們的引力沿軸方向的分量相等，因而相加；垂直於軸的分量大小相等、方向相反，因而抵銷。把整個圓環這樣逐對配起來，可知合力沿着軸，指向圓環的中心。

逐步解題

階段 1　圓環

圓環上每一小段與質點的距離都是 $\sqrt{x^2+r^2}$，而連接兩者的直線與軸成角 ψ，其中 $\cos\psi = \frac{x}{\sqrt{x^2+r^2}}$。質量為 δm 的一小段所施的引力為 $\frac{G\,\delta m}{x^2+r^2}$，其中沿軸方向的分量為 $\frac{G\,\delta m}{x^2+r^2}\cdot\frac{x}{\sqrt{x^2+r^2}}$。每一小段的距離和角度都相同，所以這些分量直接相加，得

$$\frac{Gmx}{(x^2+r^2)^{3/2}}. \qquad \blacksquare$$

階段 2　把球殼切成圓環

> **想一想**　設球殼的中心為 O，質點位於軸上，與 O 相距 b。把球殼切成多個垂直於軸的薄圓環，並以圓環上的點在 O 處與軸所成的角 ϕ 來標記各圓環。圓環的半徑、位置和質量分別是甚麼？

角度為 ϕ 的圓環半徑為 $a\sin\phi$，其中心沿軸與 O 相距 $a\cos\phi$，因此與質點相距 $a\cos\phi - b$。圓環的寬度為 $a\,\delta\phi$，所以面積為 $2\pi a\sin\phi\cdot a\,\delta\phi$，佔球殼面積 $4\pi a^2$ 的 $\frac{1}{2}\sin\phi\,\delta\phi$。它的質量是 $\frac{1}{2}M\sin\phi\,\delta\phi$。根據第 1 階段的結果，它沿着軸把質點拉向 $\phi = 0$ 的一端，力的大小為

$$\frac{G\cdot\frac{1}{2}M\sin\phi\,\delta\phi\,(a\cos\phi - b)}{\left((a\cos\phi-b)^2 + a^2\sin^2\phi\right)^{3/2}} = \frac{GM}{2}\cdot\frac{(a\cos\phi-b)\sin\phi\,\delta\phi}{(a^2-2ab\cos\phi+b^2)^{3/2}}.$$

（正負號會自動處理妥當：位於質點前方、滿足 $a\cos\phi > b$ 的圓環把質點拉向一方，位於質點後方的圓環則把它拉向另一方。）

階段 3　合適的變量

> **想一想**　分母含有 $s^2 = a^2 - 2ab\cos\phi + b^2$。$s$ 在幾何上代表甚麼？$\mathrm{d}s$ 以 $\mathrm{d}\phi$ 表示是甚麼？

根據餘弦定理（cosine rule），s 是質點到圓環的距離。求導可得 $2s\,\mathrm{d}s = 2ab\sin\phi\,\mathrm{d}\phi$，而由 $s^2 = a^2 - 2ab\cos\phi + b^2$ 解出 $a\cos\phi$，可得 $a\cos\phi - b = \frac{a^2-b^2-s^2}{2b}$。所以合力為

$$\frac{GM}{2}\int\frac{a^2-b^2-s^2}{2b\,s^3}\cdot\frac{s\,\mathrm{d}s}{ab} = \frac{GM}{4ab^2}\int_{a-b}^{a+b}\left(\frac{a^2-b^2}{s^2}-1\right)\mathrm{d}s,$$

其中 s 由 $a-b$（最近的點，$\phi=0$）變到 $a+b$（最遠的點，$\phi=\pi$）。

階段 4　結果為零

$$\int_{a-b}^{a+b}\Big(\frac{a^2-b^2}{s^2}-1\Big)\,\mathrm{d}s=(a^2-b^2)\Big(\frac{1}{a-b}-\frac{1}{a+b}\Big)-2b=(a^2-b^2)\cdot\frac{2b}{a^2-b^2}-2b=0.$$

球殼對質點不施任何力。 ■

> **停一停，試一試** 如果質點在球殼之外，即 $b>a$，情況會有甚麼不同？s 的上下限變成 $b-a$ 和 $b+a$。重做最後一行，並證明所受的力為 $\frac{GM}{b^2}$，指向球心，就如全部質量都集中在球心一樣。

回顧

- **先看對稱性（symmetry）。**把相對的小段配對，未經任何計算便已看出力的方向。
- **由較簡單的形狀逐步建構。**點組成圓環；圓環組成球殼。
- **選擇切合物理情境的變量。**把變量由角 ϕ 換成距離 s，一個棘手的積分便變得輕而易舉。

背後的數學

牛頓的圓錐論證

牛頓本人的證明（《自然哲學的數學原理》（*Principia*）第一卷命題 70）完全沒有用到積分。以質點為頂點，畫一個狹窄的對頂圓錐，它在球殼上截出兩小塊曲面（圖 10.7.1）。這兩小塊的面積與它們到質點的距離的平方 r_1^2 和 r_2^2 成正比（兩小塊與圓錐所成的角相等，因為一條弦在兩端與圓所成的角相等）。所以它們的質量與 r_1^2 和 r_2^2 成正比，而它們的引力與質量除以距離平方成正比，因此大小相等、方向相反。每一個通過質點的對頂圓錐都互相抵銷，所以合力為零。這種抵銷之所以成立，正是因為平方反比定律（inverse-square law）：若換成距離的任何其他次方，球殼便會對質點產生推力或拉力。

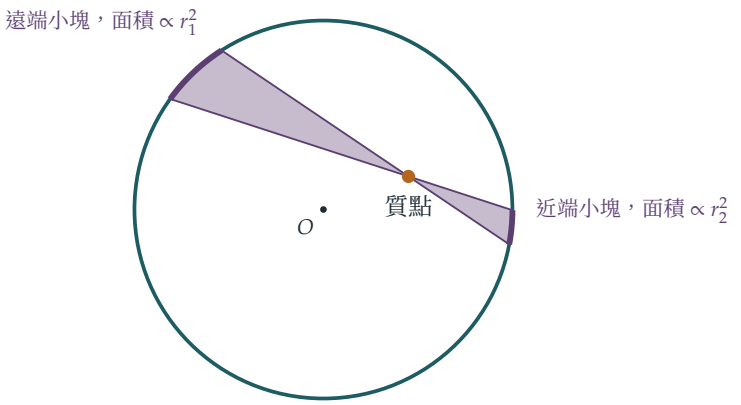


圖 10.7.1　牛頓的論證。由質點出發的一個細窄對頂圓錐在球殼上截出兩小塊。遠端的一塊較大，面積之比等於距離平方之比，但它的引力也恰好按同一比例減弱。兩個引力互相抵銷，其餘每一對亦然：在球殼之內，引力消失。

塌縮的雲團

STEP III 1987 Q14 把這個定理反過來用。在均勻球體之外，位於球面的質點所受的引力為 GMm/r^2，就如全部質量都集中在球心一樣。現在取一團靜止的均勻球狀塵埃雲，半徑為 R，質量為 M，讓它在自身引力下塌縮。球面上的質點只感受到它以內的質量，而在所有物質向內墜落的過程中，這質量一直是 M，所以它向內的速率 $v = -\dot{r}$ 滿足

$$v\frac{\mathrm{d}v}{\mathrm{d}r} = -\frac{GM}{r^2}, \qquad \text{所以} \qquad v^2 = 2GM\Big(\frac{1}{r} - \frac{1}{R}\Big).$$

代入 $r = R\cos^2\alpha$（其中一道練習），可得到達半徑 r 所需的時間為 $\sqrt{\frac{R^3}{2GM}}\big(\alpha + \frac{1}{2}\sin 2\alpha\big)$。當 $\alpha = \frac{\pi}{2}$ 時，整團雲塌縮成一點，所需時間就是**自由落體時間**（*free-fall time*）

$$t_{\mathrm{ff}} = \frac{\pi}{2}\sqrt{\frac{R^3}{2GM}} = \sqrt{\frac{3\pi}{32G\rho}},$$

其中 ρ 是密度。值得注意的是，雲團的每一層球殼都在同一刻到達中心，因為這個時間只取決於密度（圖 10.7.2）。對於一團典型的星際雲，每立方米約有 10^{-17} kg，t_{ff} 約為 700 000 年：這決定了恆星形成的時間尺度。

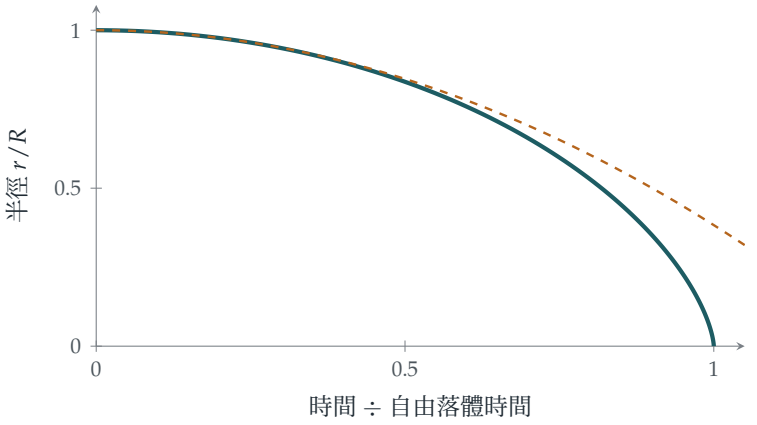


圖 10.7.2 塌縮中的均勻塵埃雲的半徑隨時間的變化（青綠色曲線）。它起初很慢，就像一塊剛放手的石頭（琥珀色虛線拋物線是恆定加速度的近似 $r = R - \frac{1}{2}\frac{GM}{R^2}t^2$），其後急劇加速，並在 $t = t_{\text{ff}}$ 時降至零。這條曲線是半條擺線（cycloid）。

牛頓的延遲與高斯定律

牛頓早在 1660 年代已掌握平方反比定律，但到 1687 年才發表他的引力理論。據一些歷史學家所說，其中一個原因是他在證明球殼定理之前，無法說明為何可以把地球當作點質量（point mass）；一旦證明了，《自然哲學的數學原理》的其餘部分便水到渠成。今天，這個定理只是高斯定律（*Gauss's law*）的一個簡短推論。高斯定律指出，穿過任何閉合曲面向外流出的引力場總量，與曲面內的質量成正比。在球殼內取一個同心的球面。它沒有包圍任何質量，所以穿過它的總流量為零。單憑這一點並不足以令場為零（均勻場穿過任何閉合曲面的總流量都是零，從一邊流入，從另一邊流出），但這裏對稱性幫了忙：場必定筆直指向球心或背離球心，而且在球面上各處強度相同，所以總流量為零就意味着場為零。同一條定律也支配電場，這就是為甚麼帶電金屬殼內沒有電場，也是為甚麼汽車能保護你免受雷擊。

思維技巧

積分即求和

要求出一個連續物體的效應，可把它切成效應已知的小塊，再把它們加起來。利用對稱性來選擇怎樣切。

點組成圓環，圓環組成球殼，球殼組成實心球。在大學裏，每一種場（引力場、電場、磁場）都是這樣由場源計算出來的，由此更引出高斯（Gauss）定理和斯托克斯（Stokes）定理的向量微積分（vector calculus）。

延伸練習

1 **球殼之外。**完成「暫停思考」方框中的問題：證明在球殼之外，所受的力為 $\frac{GM}{b^2}$。由此推斷，質量為 M 的均勻實心球對球外質點的引力，就如其質量集中在球心一樣。

2 **塌縮時間。**試做 STEP III 1987，第 14 題（`step.maths.org/questions/87-s3-q14`）。這道題求一團塵埃雲在自身引力下塌縮需要多少時間；用三角代換（trigonometric substitution）便可算出積分。

3 **穿越地球的隧道。**證明在均勻實心球之內，與球心相距 b 處所受的力為 $\frac{GMb}{a^3}$，與 b 成正比。由此推斷，把一個球投進一條穿過地球中心的隧道，它會來回振盪，週期（period）為 $2\pi\sqrt{a^3/(GM)}$，約 84 分鐘，與一顆貼着地面飛行的衛星的週期相同。

4 **另一種力的定律。**假如引力遵從立方反比定律，球殼內所受的力是否仍然為零？（重做第 3 階段，以 $\frac{1}{s^3}$ 取代 $\frac{1}{s^2}$。）

延伸閱讀

- **Daniel Kleppner & Robert Kolenkow, *An Introduction to Mechanics***（Cambridge University Press），關於萬有引力的章節，當中包括球殼定理。
- **Subrahmanyan Chandrasekhar, *Newton's Principia for the Common Reader***（Oxford University Press）。由一位諾貝爾獎得主講解牛頓的證明，包括命題 70 和 71。
- **Andrew Pontzen, *The Universe in a Box***（Riverhead）。講述引力、塌縮和模擬如何解釋宇宙的結構。

10.8 排成三角形的三顆恆星

題目

STEP I 1999，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/99-s1-q11`

引子

兩個以引力互相吸引的天體，其運動可以完全解出，這是牛頓（Newton）證明了的，Section 10.6 亦已解釋過：兩者各自沿一條圓錐曲線（conic）繞共同的質心（centre of mass）運動，若兩者互相束縛，軌道是橢圓；若能量足以互相逃逸，則是拋物線或雙曲線。再加入第三個天體，一切就不同了。三體問題（three-body problem）沒有能用常見函數表示的通解，而在 1890 年，龐加萊（Henri Poincaré）更發現它的解可以是混沌的，這是混沌（chaos）在數學中首次出現。

然而，有些特殊解卻十分簡單。1772 年，拉格朗日（Joseph-Louis Lagrange）發現，位於等邊三角形（equilateral triangle）三個頂點的三個天體可以作剛性旋轉，整個三角形像車輪一樣轉動。本題要找的，就是質量相等時的這個解。拉格朗日的三角形並非只是趣聞：太陽、木星和數以千計的「特洛伊」（Trojan）小行星正好構成這種位形，今天的太空望遠鏡亦停泊在類似的點上。

動筆之前

想一想 一顆星要以角速度（angular velocity）ω 沿半徑為 R 的圓運動，它所受的總力必須是甚麼？

總力的大小必須是 $m\omega^2 R$，方向指向圓心。因此我們要檢查每顆星所受的總引力是否指向三角形的中心，並求出使其大小恰好相符的 ω。

逐步解題

階段 1　幾何

想一想　邊長為 a 的等邊三角形，每個頂點與中心相距多遠？

中心位於每條中線（median）上，與頂點的距離是中線長度的三分之二。中線的長度為 $a\sin 60^\circ = \frac{\sqrt{3}}{2}a$，所以每個頂點到中心的距離是

$$R = \frac{2}{3}\cdot\frac{\sqrt{3}}{2}a = \frac{a}{\sqrt{3}}.$$

由對稱性（symmetry）可知，三角形的中心就是三顆相同的星的質心。

階段 2　一顆星所受的力

想一想　另外兩顆星各以大小為 $\frac{\gamma m^2}{a^2}$ 的力拉它。這兩個拉力之間的夾角是多少？它們的合力（resultant）是多少，又指向哪個方向？

兩個拉力分別沿着三角形在該星處相交的兩條邊，所以它們之間的夾角是 60°。由對稱性可知，合力沿着這個角的角平分線（bisector），而角平分線通過中心。每個拉力與角平分線成 30°，所以合力的大小是

$$2\cdot\frac{\gamma m^2}{a^2}\cos 30^\circ = \frac{\sqrt{3}\,\gamma m^2}{a^2},$$

方向指向中心（圖 10.8.1）。

階段 3　角速度

令合力等於 $m\omega^2 R$：

$$m\omega^2\frac{a}{\sqrt{3}} = \frac{\sqrt{3}\,\gamma m^2}{a^2}, \qquad \text{所以} \qquad \omega^2 = \frac{3\gamma m}{a^3}.$$

取這個角速度時，同樣的計算（由對稱性）適用於每一顆星，所以三顆星都繞中心作圓周運動，三角形保持剛性不變。 ■

階段 4　中心的第四顆星

想一想　一顆質量為 λm 的星位於中心，與每個頂點的距離是 $\frac{a}{\sqrt{3}}$。它對每顆頂點上的星額外施加多大的力？質心會移動嗎？中心的星有沒有受到淨力？

中心的星不會改變質心的位置，而由對稱性可知，它所受的三個拉力互相抵消，所以它保持靜止。它把每顆頂點上的星拉向中心，力的大小是 $\frac{\gamma\lambda m^2}{(a/\sqrt{3})^2} = \frac{3\gamma\lambda m^2}{a^2}$。現在頂點上的

星所受指向中心的總力是 $(\sqrt{3}+3\lambda)\frac{\gamma m^2}{a^2}$，所以

$$m\omega^2\frac{a}{\sqrt{3}}=(\sqrt{3}+3\lambda)\frac{\gamma m^2}{a^2},\qquad \omega^2=\frac{(3+3\sqrt{3}\,\lambda)\gamma m}{a^3}=\frac{3\gamma m}{a^3}(1+\sqrt{3}\,\lambda).\qquad\blacksquare$$

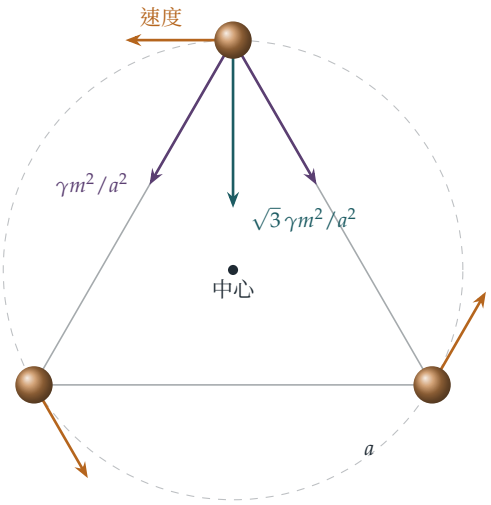


圖 10.8.1 三顆相同的星位於等邊三角形的三個頂點。頂部那顆星所受的兩個拉力（紫色）合成一個指向中心的力（藍綠色）。若每顆星以適當的速率（琥珀色）沿虛線圓運動，整個三角形便作剛性旋轉。這就是拉格朗日對三體問題的等邊三角形解。

回顧

- **猜一個對稱的解。**我們並沒有解出三體問題；我們只是猜了一種高度對稱的運動，再檢查牛頓定律是否容許它。
- **對稱性決定方向。**合力必定沿着角平分線，所以只需計算它的大小。
- **檢查每一個天體。**由對稱性可知，驗證了一顆星就等於驗證了全部三顆；中心的星則需要另作（容易的）檢查。

背後的數學

喬裝的開普勒定律

把答案寫成 $\omega^2a^3=\gamma(3m)$。對於相距 a、沿圓形軌道運行的兩個天體，開普勒（Kepler）第三定律（Section 10.6 的一道練習）指出 $\omega^2a^3=\gamma(m_1+m_2)$，右邊出現的是總質量。

三顆星遵從同一條定律，總質量為 $3m$。事實上，拉格朗日證明了等邊三角形解對任意三個質量都存在，而不限於相等的質量：三角形繞質心旋轉，並且 $\omega^2 a^3 = \gamma(m_1+m_2+m_3)$。歐拉（Euler）在幾年前已找到第二族解，其中三個天體位於一條旋轉的直線上。這五種位形（三種共線、兩種三角形，不計標號的差別）是三個天體能保持形狀不變的僅有情況。

拉格朗日點

取 m_1 為太陽，m_2 為一顆行星，m_3 為一顆微小的小行星。等邊三角形解告訴我們，小行星可以停留在與太陽和行星構成等邊三角形的第三個頂點上，在行星軌道上位於行星的前方或後方。這就是**拉格朗日點**（Lagrange points）L_4 和 L_5；歐拉的共線解則給出 L_1、L_2 和 L_3。木星有超過一萬顆已知的「特洛伊」小行星，聚集在它的 L_4 和 L_5 點附近。詹姆斯·韋伯太空望遠鏡（James Webb Space Telescope）在日地系統的 L_2 點附近運行，位於地球以外 1.5 百萬公里處，在那裏它可以把太陽、地球和月球全部擋在遮陽罩之後。

穩定性

造星者的三角形能維持下去嗎？對於相等的質量，答案是不能：最輕微的擾動也會不斷增長，幾顆星最終不是四散飛離，就是互相碰撞。Edward Routh 在 1875 年證明，拉格朗日的三角形只有在其中一個質量佔絕對優勢時才是穩定（stable）的，確切的條件是

$$27(m_1m_2 + m_2m_3 + m_3m_1) < (m_1+m_2+m_3)^2.$$

對於太陽和木星，這個條件綽綽有餘地成立，這就是特洛伊小行星得以長存的原因。對於三個相等的質量，$27 \times 3m^2 > 9m^2$，所以這個位形並不穩定。

龐加萊的錯誤與 8 字形軌道

1887 年，瑞典國王奧斯卡二世（Oscar II）懸賞徵求多體問題的解。龐加萊在 1889 年憑一篇關於三體問題的長篇論文獲獎，但論文付印期間，一位編輯發現了其中一處漏洞。在修正的過程中，龐加萊發現有些解的表現「複雜得驚人，我甚至無法嘗試把它畫出來」：這是人類對混沌的第一瞥。已印好的論文全數收回銷毀，費用由龐加萊承擔。新的特殊解至今仍不斷被發現：1993 年 Cristopher Moore 以數值方法發現，而 2000 年 Alain Chenciner 與 Richard Montgomery 證明，三個相等的質量可以沿一條 8 字形曲線互相追逐。此後，人們借助電腦又找到了數以千計的周期軌道。

思維技巧

猜一個對稱的解，然後驗證

當一個問題難以在一般情況下解出時，不妨尋找對稱性盡可能高的解。對稱的猜想能把許多未知量減少到一兩個，而驗證它往往很容易。

這是數學和物理中一個強而有力的策略：對稱的平衡態、穩態、行波和剛性旋轉的位形，往往是最先找到的解，也是研究其餘各種解的起點。接下來要問的，永遠是這個對稱解是否穩定。

延伸練習

1 **正方形上的四顆星。**四顆質量為 m 的星位於邊長為 a 的正方形的四個頂點。證明它們可以繞中心作剛性旋轉，其中 $\omega^2 = \frac{(4+\sqrt{2})\gamma m}{2a^3}$。

2 **兩顆相同的星和一顆行星。**兩顆質量為 M 的星相距 a，繞它們的質心運行。一顆質量可忽略的行星位於中心。它是否處於平衡？對於沿兩星連線方向的微小位移，它是否穩定？垂直於連線的方向又如何？

3 **不相等的質量。**質量分別為 m_1、m_2、m_3 的三顆星位於邊長為 a 的等邊三角形的三個頂點，並繞它們的質心 G 旋轉。證明 m_1 所受的引力為 $\frac{\gamma m_1(m_1+m_2+m_3)}{a^3}\overrightarrow{P_1G}$，其中 P_1 是它的位置。（提示：把各力寫成向量，並利用質心的定義。）由此推導出拉格朗日的結果。

4 n **顆星。**n 顆相同的星位於一個正 n 邊形的各個頂點，該正多邊形內接於半徑為 R 的圓。求剛性旋轉的 ω^2，並把它表示成 $\frac{1}{\sin(k\pi/n)}$ 對 k 求和的形式。

延伸閱讀

- **June Barrow-Green, *Poincaré and the Three Body Problem***（American Mathematical Society）。講述那次懸賞以及龐加萊發現混沌的故事。
- **Richard Montgomery, "A new solution to the three-body problem"**, *Notices of the American Mathematical Society* 48 (2001)。由 8 字形軌道的其中一位發現者親自講解；可於網上免費閱讀。
- **Florin Diacu & Philip Holmes, *Celestial Encounters***（Princeton University Press）。一部關於三體問題與混沌的通俗歷史。

我們學到了甚麼

本章開首說過，有些東西永遠不變，而知道哪些東西不變，就已經掌握了半門物理學。我們提出了三個問題：守恆定律（conservation laws）單憑自身能決定甚麼、細小的步驟如何累積起來，以及怎樣馴服多個物體組成的系統。以下是本章八道題目給出的答案。

1. 守恆定律單憑自身能決定甚麼？

兩個物體沿直線作彈性碰撞（elastic collision）時，動量（momentum）和能量（energy）完全決定了碰撞的結果，而這一個結論在本章一再派上用場。物體再多一些，情況就不同了。在 Section 10.1 中，一排彈珠有許多種結果都同時符合這兩條定律，要再加上一項物理事實（彈珠不能互相穿過），才能排除其中一些。把動量的平方減去能量，得出一個由交叉項組成的和，而這個和必須等於零；這個不等式（inequality）論證證明了某些運動根本不可能發生。至於牛頓擺（Newton's cradle）那種人所熟知的表現，原來取決於守恆定律以外的細節：衝擊如何在球與球之間傳遞。

守恆定律也揭露了隱藏的假設。在 Section 10.3 中，所有炮同時發射所得的速率，比逐一發射為高，但這只是因為同時發射用了更多能量；一旦把該固定的量固定下來，結論就反過來了。

2. 細小的步驟如何累積起來？

有好幾節都是把一個簡單的步驟重複又重複。在 Section 10.2 中，一次碰撞最多只能令速率加倍，但經過一串愈來愈輕的球，增幅就變得沒有上限：這正是籃球疊網球背後的原理，也是引力彈弓（gravitational slingshot）背後的原理。在 Section 10.3 中，許多次細小的發射化成了一個積分（integral），並得出齊奧爾科夫斯基（Tsiolkovsky）的對數（logarithm），由此解釋了為何火箭的大部分都是燃料，以及為何火箭要分級。在 Section 10.4 中，分子每次從移動中的活塞反彈，速率都增加 $2u$，而一個精確的遞推關係（recurrence）顯示，速率與長度的乘積幾乎保持不變：這是一個絕熱不變量（adiabatic invariant），也是氣體受壓縮時升溫的原因。而在 Section 10.5 中，無窮多次反彈只花了有限的時間，因為各段時間構成一個收斂（convergent）的等比級數（geometric series）。

3. 怎樣馴服多個物體組成的系統？

方法是選取合適的坐標，並善用對稱性（symmetry）。在 Section 10.6 中，質心（centre of mass）作勻速運動，而兩物體的間距則表現得像一個帶有約化質量（reduced mass）的單一質點；一個隱藏的平面甚至解釋了為何間距的平方是時間的二次式。在 Section 10.7 中，對稱性加上積分，證明了球殼（spherical shell）對內部不施加引力，對外部則有如一個質點；於是一團正在坍縮的塵埃雲可以用單獨一條方程來處理，而當中的自由下落時間（free-fall time）決定了恆星形成的步伐。而在 Section 10.8 中，一個對稱的猜想解決了無法求解的三體問題（three-body problem）的一個特殊情況：拉格朗日（Lagrange）的旋轉三角形，它在特洛伊小行星（Trojan asteroids）身上得到了體現。

4. 力學工具箱

方法	你在哪裏用過
以守恆定律作為約束條件	牛頓擺（10.1）
轉換參考系（frame）	碰撞（10.2）；活塞（10.4）
反覆運用同一結果	疊球（10.2）
質量可變時的動量	火箭（10.3）
由求和到積分	火箭方程（10.3）；球殼（10.7）
遞推關係與裂項相消（telescoping）	活塞（10.4）
等比級數	楔（10.5）
質心與約化質量	二體（10.6）
能量積分 $v\,\mathrm{d}v/\mathrm{d}z$	二體（10.6）；坍縮（10.7）
對稱的猜想	三星（10.8）

下一步

- 第 11 章由守恆轉到振盪（oscillation）與轉動（rotation）：共振（resonance）、簡正模式（normal modes）、真實的擺，以及在加速參考系和旋轉參考系中的運動。
- Section 8.2 的能量方法和 Section 8.4 的極限環（limit cycles），是同一批思想在微分方程（differential equation）方面的體現。
- 龐加萊（Poincaré）在三體問題中發現的混沌（chaos），是 Section 8.7 的主題。

如果你想在入讀大學之前再走遠一點

- **Daniel Kleppner & Robert Kolenkow, *An Introduction to Mechanics***（力學導論）：經典的大學一年級教科書；要求很高，但收穫甚豐。

- **David Morin, *Introduction to Classical Mechanics***（古典力學導論）：收錄數以百計的題目，當中不少與本章的題目一樣出人意表。
- **Richard Feynman, *The Character of Physical Law***（物理定律的本性）：為一般聽眾而設的講座，內容涵蓋守恆定律、對稱性和引力。
- **Dwight Neuenschwander, *Emmy Noether's Wonderful Theorem***（諾特的奇妙定理）：為何有對稱性就有守恆。

費曼（Feynman）說過，能量守恆至今未見任何例外。本章嘗試展示守恆定律的威力有多大：排除不可能的結果、揭示隱藏的不變量，並把複雜得令人絕望的系統化為我們能夠理解的系統。但本章也展示了它們的局限。守恆定律能夠約束，卻不一定能夠決定；在它們止步之處，就要由力、對稱性和參考系的細節接手。

延伸練習答案

以下只是簡短的答案和提示，並非完整解答：只列出結果和關鍵步驟，方便你核對自己的做法。每道練習請先認真做過，然後才看答案。

10.1 牛頓擺

1. 所有碰撞（collision）結束後，運動中的質點後面不可能跟着一個較慢的質點。因此，若恰好只有一個質點在動，它必定是 A_n，而 $\lambda m v = mu$、$\lambda m v^2 = mu^2$ 在 $\lambda \neq 1$ 時互相矛盾。若只有 A_{n-1} 和 A_n 在動，則方程與質量為 m 和 λm 的兩個質點相碰時的方程相同，得出 A_{n-1} 的速度為 $\frac{1-\lambda}{1+\lambda}u < 0$；但這樣它便會撞上 A_{n-2}。有三個質點在動的情況，可完全仿照第 (ii) 部分排除。在動的兩個質點是 A_0（以 $\frac{\lambda-1}{\lambda+1}u$ 向後運動）和 A_n（以 $\frac{2u}{\lambda+1}$ 向前運動）。

2. 質量相等的彈珠相碰時會交換速度，所以速率 u 一直傳到 A_n。之後 A_n 撞上 A_{n+1}：A_{n+1} 以 $\frac{2u}{1+\lambda}$ 離開，而 A_n 則以 $\frac{\lambda-1}{\lambda+1}u$ 反彈。這個向後的速度沿着整列彈珠傳回 A_0，使它以該速率向後離開。

3. 「兩球以 u 彈出」所帶的動量（momentum）是 $2mu$，能量是 mu^2，與碰撞前相同。另一個解（之後不再有碰撞）是：最後三個球分別以 $\frac{u}{3}$、$\frac{u}{3}$ 和 $\frac{4u}{3}$ 離開（驗算：總和為 $2u$，平方和為 $2u^2$）。真實的牛頓擺選擇第一個解，是因為各球之間並非完全貼緊，所以碰撞是一對一對地先後發生的。

4. 速率一直傳到 A_n，它再撞上較輕的 A_{n+1}：A_{n+1} 以 $\frac{2u}{1+\lambda}$ 離開，而 A_n 則以較慢的 $\frac{1-\lambda}{1+\lambda}u$ 跟在後面。其餘質點全都不動。

10.2 疊球效應

1. 球以 $U = \sqrt{2gH}$ 到達地面，再以 eU 反彈，上升 e^2H。把小球放在上面時：大球以 eU 向上運動，遇上以 U 向下運動的小球。由動量守恆和恢復係數（restitution）的關係，可得小球向上的速率為 $w = U\frac{Me(2+e)-m}{M+m}$。小球球心由 $2R + r = 0.45$ 出發，升至 4.5，所以 $\frac{w^2}{2g} = 4.05$，而 $\frac{w}{U} = \sqrt{\frac{4.05}{1.8}} = 1.5$。代入 $e = \frac{2}{3}$，得 $\frac{16}{9}M - m = 1.5(M + m)$，於是 $M = 9m$。

2. 由動量守恆和恢復係數的關係，P_1 的速率為 $\frac{1+e}{1+\lambda}u$；每次碰撞的質量比都相同，情況重複出現，所以 P_n 以 $\left(\frac{1+e}{1+\lambda}\right)^n u$ 運動。速率為 s 的質點撞上下一個質點後，餘下

的速率為 $\frac{1-\lambda e}{1+\lambda}s$，而下一個質點經過自己的下一次碰撞後，速率減至 $\frac{(1+e)(1-\lambda e)}{(1+\lambda)^2}s$。當 $\lambda e<1$ 時，前者只有在 $\frac{1-\lambda e}{1+\lambda}>\frac{(1+e)(1-\lambda e)}{(1+\lambda)^2}$ 時才會再次追上後者，即 $\lambda>e$；所以若 $e>\lambda$，每個質點只會被撞一次，也只會撞別的質點一次。若 $\lambda e=1$，每個質點撞上下一個之後便立即停下。

3. 最底的球以 U 反彈，遇上中間的球（正以 U 向下運動）：相對速率為 $2U$，所以中間的球以 $3U$ 離開。最上面的球與它相遇時相對速率為 $4U$，於是以 $7U$ 離開，上升 $49H$。一般而言 $u_{j+1}=2u_j+U$，所以 $u_j+U=2^jU$。

4. 彈性碰撞（elastic collision）中動能（kinetic energy）守恆。A 的能量是 $\frac12 m\big(\frac{2\lambda}{1+\lambda}\big)^{2n}u^2$，而最初的能量是 $\frac12\lambda^n mu^2$，又因 $\frac{4\lambda}{(1+\lambda)^2}<1$，所以 A 得到的能量永遠不會多於原有的總能量。較重的質點繼續向前運動，各自保留原來速率的 $\frac{\lambda-1}{\lambda+1}$ 倍，餘下的能量便留在它們身上。

10.3 火箭與槍炮

1. 一次過全部發射：$Mu=nm(v-u)$，所以 $u=\frac{nmv}{M+nm}$。逐個發射：第 r 次發射給出 $\big(M+(n-r+1)m\big)(u_r-u_{r-1})=mv$，所以 $u_n=\sum_{j=1}^{n}\frac{mv}{M+jm}$，而除最後一項外，每一項都大於 $\frac{mv}{M+nm}$。因此 $u_n>u$。

2. 動量：$M(1+bx)v=MV$。於是 $(1+bx)\frac{\mathrm{d}x}{\mathrm{d}t}=V$，所以 $x+\frac12 bx^2=Vt$，而 $x=\frac{\sqrt{1+2bVt}-1}{b}$。有外力時，$\frac{\mathrm{d}}{\mathrm{d}t}\big(M(1+bx)v\big)=Mf$ 給出 $(1+bx)v=ft+V$。若速率為常數 c，則須有 $x=ct$ 及 $(1+bct)c=ft+V$，所以 $c=V$ 且 $bV^2=f$。否則 $x+\frac12 bx^2=\frac12 ft^2+Vt$，所以 $x\sim t\sqrt{f/b}$，而 $v\to\sqrt{f/b}$。

3. 單級：$v\ln 10\approx 2.30v$。兩級：$v\ln\frac{10}{5.5}+v\ln\frac{5}{0.5}\approx 0.60v+2.30v=2.90v$。

4. 當 $0\le s\le x$ 時，$\frac{1}{1+x}\le\frac{1}{1+s}\le 1$；在 $[0,x]$ 上積分即可。

10.4 一個分子與一個活塞

1. 這個遞推關係（recurrence）逐項相消後得 $d_n=\frac{(v+u)d_1}{v+(2n-1)u}$，所以 $v_nd_n=\frac{(v+2nu)(v+u)d_1}{v+(2n-1)u}$，它比 $(v+u)d_1$ 大 $\frac{u(v+u)d_1}{v+(2n-1)u}\to 0$。

2. 把 u 換成 $-u$：每次從後退中的活塞反彈，速率都減少 $2u$，所以 $v_n=v-2nu$（只要質點仍能追上活塞），$d_n=\frac{(v-u)d_1}{v-(2n-1)u}$，而 $v_nd_n=\frac{(v-2nu)(v-u)d_1}{v-(2n-1)u}$，同樣接近 $(v-u)d_1$。

3. 當 $v=u$ 時，$v_n=(2n+1)u$。第 n 次碰撞時動量的改變量為 $m(v_{n-1}+v_n)=4nmu$，而 $t_n=\frac{d_1}{un(n+1)}$，所以平均力為 $\frac{4mu^2n^2(n+1)}{d_1}$，量級為 n^3。每個循環所作的功約為力 $\times\, ut_n$，大約是 $4nmu^2$，這些功加起來是無窮大：要把活塞完全推到底，需要無窮多的功，正如分子的能量 $\frac12 m(2n+1)^2u^2$ 會無止境地增長一樣。

4. 不變量（invariant）是 $\oint p\,\mathrm{d}q = 2\int_0^h m\sqrt{2g(h-z)}\,\mathrm{d}z \propto h^{3/2}$，其中 h 是彈跳離地板的高度。所以當地板緩慢上升時，$h^{3/2}$（等價地說，即 v^3，或能量乘以週期）幾乎保持不變。

10.5 有限時間內的無限次反彈

1. 由正弦定理（sine rule），$M_1N_1 = \frac{OM_1\sin 30^\circ}{\sin\theta}$，而相繼各段路程按比率 $\rho = \frac{\sin(\theta-30^\circ)}{\sin\theta} < 1$ 縮短。總和為 $\frac{M_1N_1}{1-\rho} = \frac{OM_1}{2\left(\sin\theta - \sin(\theta-30^\circ)\right)} = \frac{OM_1}{4\sin 15^\circ\cos(\theta-15^\circ)}$。

2. 由三角形 OM_1N_1 的各角可得 $\varphi = \theta - \alpha$。時間之比為 $\frac{\rho}{k} = \frac{\sin(\theta-\alpha)\cos(\theta-\alpha)}{\sin\theta\cos\theta} = \frac{\sin(2\theta-2\alpha)}{\sin 2\theta}$，它小於 1 當且僅當 $\alpha < \theta < 45^\circ + \frac{1}{2}\alpha$。

3. 把楔形沿它的壁反射，可將路徑展開成一條直線。圍繞 O 剛好可放十二個楔形，而一條不通過 O 的直線最多穿過其中六個。所以質點只會反彈有限次，然後逃離，除非它正對着 O 射去。

4. 它首先下落 $t_0 = \sqrt{2/9.8} \approx 0.452\,\mathrm{s}$。每次彈跳歷時 $2e^k t_0$，所以總時間為 $t_0\left(1 + \frac{2e}{1-e}\right) = 9t_0 \approx 4.07\,\mathrm{s}$。總路程為 $1 + 2\sum_{k\geq 1} e^{2k} = 1 + \frac{2(0.64)}{0.36} \approx 4.56\,\mathrm{m}$。

10.6 二體化為一體

1. 繩子保持拉緊，相對速度與繩子垂直，大小恆為 u，所以 $\theta = \frac{ut}{r}$。由動量可得 $my + MY = mut$。利用 $A = G - \frac{M}{m+M}\overrightarrow{AB}$ 和 $\overrightarrow{AB} = r(\cos\theta, -\sin\theta)$，得 $y = \frac{1}{m+M}\left(mut + Mr\sin\frac{ut}{r}\right)$。於是 $\dot{y} = \frac{u}{m+M}\left(m + M\cos\frac{ut}{r}\right)$，它在 $\cos\frac{ut}{r} < -\frac{m}{M}$ 時為負，而這在 $M > m$ 時才有可能。

2. 現在 $\ddot{z} = -\frac{3}{z^3}$，所以 $\dot{z}^2 = \frac{3}{z^2} - 2$（利用 $z = 1$、$\dot{z} = -1$）。設 $w = z^2$，則 $\ddot{w} = 2\dot{z}^2 + 2z\ddot{z} = -4$，所以 $z^2 = 1 - 2t - 2t^2$。兩者在 $t = \frac{\sqrt{3}-1}{2} \approx 0.366$ 時相撞。

3. $z^2 = 4t^2 - 2t + 1 = |\mathbf{p} + \mathbf{q}t|^2$，其中 $\mathbf{p} = (1,0)$，$\mathbf{q} = (-1,\sqrt{3})$：這是通過 $(1,0)$、方向為 $(-1,\sqrt{3})$ 的直線，而 $\theta = \tan^{-1}\frac{\sqrt{3}t}{1-t}$。於是 $z^2\dot{\theta} = |\mathbf{p}\times\mathbf{q}| = \sqrt{3}$。這個斥力正好就是平面上自由運動的「離心」項 $\frac{h^2}{z^3}$，其角動量（angular momentum）為 $h = \sqrt{3}$。

4. 兩顆星與質心（centre of mass）的距離分別為 $\frac{m_2 a}{m_1+m_2}$ 和 $\frac{m_1 a}{m_1+m_2}$。對第一顆星，$\frac{Gm_1m_2}{a^2} = m_1\omega^2\frac{m_2 a}{m_1+m_2}$，所以 $\omega^2 = \frac{G(m_1+m_2)}{a^3}$，而 $T = \frac{2\pi}{\omega}$。

10.7 球殼內的引力

1. 在球殼外（$b > a$），與第 3 階段相同的積分改為對 s 由 $b-a$ 積至 $b+a$，結果是 $\frac{GM}{b^2}$。實心球是多個同心球殼的併集，每個球殼的作用都如同其質量集中於球心。

2. 設 $r = R\cos^2\alpha$，則 $\frac{1}{r} - \frac{1}{R} = \frac{\tan^2\alpha}{R}$，所以速率為 $\sqrt{2GM/R}\tan\alpha$。於是 $\mathrm{d}t = \frac{2R\sin\alpha\cos\alpha\,\mathrm{d}\alpha}{\sqrt{2GM/R}\tan\alpha} = \sqrt{\frac{R^3}{2GM}}\,2\cos^2\alpha\,\mathrm{d}\alpha$；再積分即可。

3. 根據球殼定理（shell theorem），只有半徑 b 以內的質量（即 $M\frac{b^3}{a^3}$）產生引力，結果是 $\frac{GMb}{a^3}$。這是簡諧運動（simple harmonic motion），其中 $\omega^2 = \frac{GM}{a^3}$，週期為 $2\pi\sqrt{a^3/GM} \approx 84$ 分鐘，這裏取 $a = 6.37\times10^6\,\mathrm{m}$ 及 $GM = 3.99\times10^{14}\,\mathrm{m^3\,s^{-2}}$。在半徑 a 處運行的衛星滿足 $\omega^2 a = \frac{GM}{a^2}$：$\omega$ 相同。

4. 不會。換成 $\frac{1}{s^3}$ 後，被積函數不再化簡為常數，球殼較近的一側佔了上風：質點會被拉向球殼最近的部分。只有平方反比定律（inverse-square law）才會完全抵消（球殼上一小塊的面積按 s^2 增長，剛好抵消引力的減弱）。

10.8 排成三角形的三顆恆星

1. 每顆星與中心的距離為 $\frac{a}{\sqrt{2}}$。相鄰兩顆星的引力指向中心的分量合共為 $\frac{2\gamma m^2}{a^2}\cos45^\circ = \frac{\sqrt{2}\gamma m^2}{a^2}$，而對面那顆星的引力為 $\frac{\gamma m^2}{2a^2}$。令 $\frac{\gamma m^2}{a^2}\left(\sqrt{2}+\frac{1}{2}\right) = m\omega^2\frac{a}{\sqrt{2}}$，得 $\omega^2 = \frac{(4+\sqrt{2})\gamma m}{2a^3}$。

2. 在中心處各引力互相抵消，所以行星處於平衡（equilibrium）。沿着兩顆星的連線方向，平衡是不穩定的：向其中一顆星移動距離 x，便會受到約為 $\frac{32\gamma Mx}{a^3}$ 的向外淨力。在垂直於軌道平面的方向，平衡是穩定的，回復力約為 $\frac{16\gamma Mz}{a^3}$。（在軌道平面內垂直於連線的方向，引力有回復作用，但旋轉會把這個方向的運動與不穩定的方向耦合起來。）

3. 作用於 m_1 的力為 $\frac{\gamma m_1}{a^3}\big(m_2(P_2-P_1)+m_3(P_3-P_1)\big)$。由於 $m_1P_1+m_2P_2+m_3P_3 = (m_1+m_2+m_3)G$，這個力等於 $\frac{\gamma m_1(m_1+m_2+m_3)}{a^3}(G-P_1)$。每顆星都被拉向 G，引力與它跟 G 的距離成正比，而且比例常數相同，所以三顆星可以像剛體一樣一起旋轉，其中 $\omega^2 = \frac{\gamma(m_1+m_2+m_3)}{a^3}$。

4. 相隔 k 個位置的那顆星，距離為 $2R\sin\frac{k\pi}{n}$，其引力指向中心的分量佔 $\sin\frac{k\pi}{n}$。所以 $m\omega^2R = \frac{\gamma m^2}{4R^2}\sum_{k=1}^{n-1}\frac{1}{\sin(k\pi/n)}$，得 $\omega^2 = \frac{\gamma m}{4R^3}\sum_{k=1}^{n-1}\frac{1}{\sin(k\pi/n)}$。當 $n=3$、$R=\frac{a}{\sqrt{3}}$ 時，結果是 $\frac{3\gamma m}{a^3}$。

章

11

振動、轉動與運動參考系

會擺動的東西、會旋轉的東西，以及從移動中的位置所看到的東西。

本章內容

11.1	共振	STEP II 1991
11.2	簡正模式	STEP III 2014
11.3	真實的擺	STEP III 1998
11.4	等時曲線	STEP III 1994
11.5	加速火車中的重力	STEP II 2021
11.6	後旋	STEP III 1997
11.7	桿中的辛普森法則	STEP II 1989
11.8	傾斜的平衡	STEP III 2025
11.9	安全拋物線	STEP II 2014
11.10	會朝你飛回來的球	STEP II 2019
11.11	外野手如何接球	STEP II 1998
11.12	橡皮繩上的螞蟻	STEP II 2018
11.13	鏟雪車問題	STEP III Specimen paper
11.14	是拋物線，不是懸鏈線	STEP I 1991
A	練習答案	第 817 頁

看清運動

年輕理論物理學家的事業，就是在愈來愈抽象的層次上處理諧振子。

— 相傳出自 Sidney Coleman

上一章利用守恆定律（conservation laws）來理解碰撞和引力。本章轉向動力學（dynamics）入門課程的另外幾個重要主題：會振動的東西、會轉動的東西，以及從最有用的觀點描述運動的藝術。

振動（oscillation）無處不在。單擺和彈簧是教科書上的例子，但同樣的數學也描述了分子的振動、建築物的搖晃、收音機裏的電路，以及在量子力學中光本身的行為。原因在於，幾乎任何處於穩定平衡（stable equilibrium）附近的東西，表現都像彈簧：把它拉開一點點，回復力（restoring force）便與位移成正比。因此，簡諧振子（simple harmonic oscillator）是幾乎所有物理系統的一階近似，而物理學有很大一部分，就是要弄清楚當這個近似不太準確時會發生甚麼事。

為甚麼這很重要

工程師必須知道所建造的每一座結構的固有頻率（natural frequency），使它們遠離風、交通和地震的頻率。化學家憑分子的振動來辨認分子。計時這門技藝，從惠更斯（Huygens）的擺鐘到石英錶和原子鐘，就是製造完美振子的藝術。轉動支配着一切，從桌球的旋轉到人造衛星的穩定性皆然。至於哪些參考系（frame of reference）才算「自然」這個問題，則經由引力與加速度的等效性，引出了愛因斯坦（Einstein）的廣義相對論。

本章最後有六道題目，歸為一組，稱為出人意表的軌跡：拋體、螞蟻、外野手和鏟雪車，它們的表現都與直覺所預料的不同。這些題目短小而優雅，每一道的核心都有一個真正的數學概念：包絡線（envelope）、單調性（monotonicity）、反饋（feedback）、發散（divergence）、線性微分方程（linear differential equation），以及拋物線（parabola）與懸鏈線（catenary）之間的分別。

你在大學會遇到甚麼

大學一年級的動力學課程涵蓋簡諧運動、阻尼（damped）諧運動和受迫（forced）諧運動、共振（resonance），以及耦合振子（coupled oscillators）及其簡正模態（normal modes）（這是線性代數與力學交匯之處）。課程也會處理剛體（rigid body）、轉動慣量

（moment of inertia）和滾動，以及在加速參考系和旋轉參考系中的運動，當中會出現離心力（centrifugal force）和科里奧利力（Coriolis force）。往後的課程會把這一切發展成拉格朗日力學（Lagrangian mechanics）和哈密頓力學（Hamiltonian mechanics），坐標的選擇在其中成為核心主題；又會發展成波動理論，而波就是在空間中散佈開來的振動。非線性振動，例如大振幅的單擺，則通向橢圓函數（elliptic functions）和第 8 章的動力系統（dynamical systems）。

貫穿本章的問題

三大問題

1. **運動在甚麼時候是簡單的？**一個系統在甚麼時候會以單一而固定的頻率振動？若不是這樣，又會發生甚麼事？
2. **甚麼才是合適的觀點？**哪些坐標、哪些變量、哪個參考系能令問題變得容易？
3. **直覺可信嗎？**日常對運動的直覺在哪裏出錯？數學又如何糾正它？

本章由振動開始：共振(Section 11.1)、簡正模態、真實的單擺和等時曲線(tautochrone)。接着轉向參考系和剛體：一列加速中的火車、一顆旋轉的彈珠、一根表現得像三個質點的桿，以及一個可以按三種不同傾斜角度靜止的立體。最後六節就是那些出人意表的軌跡。

本章的概念

§	課題	概念
11.1	受迫振動	共振與拍（beats）
11.2	耦合振子	簡正模態就是特徵向量（eigenvectors）
11.3	真實的單擺	橢圓積分（elliptic integrals）；算術幾何平均（AGM）
11.4	等時曲線	以弧長（arc length）表示的簡諧運動
11.5	運動參考系	假想力（fictitious forces）；等效原理（equivalence principle）
11.6	滾動與滑動	一個守恆的角動量（angular momentum）
11.7	轉動慣量	動力等效（dynamical equivalence）；辛普森法則（Simpson's rule）
11.8	穩定性	通過質心（centre of mass）的法線
11.9	拋體	軌跡的包絡線
11.10	拋體	距離不一定遞增
11.11	接球	一條反饋法則
11.12	伸長的繩	分數與調和級數（harmonic series）
11.13	鏟雪車	把時間看成距離的函數
11.14	懸索橋	是拋物線，不是懸鏈線

Section 11.6 和 Section 11.7 這兩節用到轉動動力學（rotational dynamics）和轉動慣量，這些內容已於 2019 年從 STEP 考試範圍中刪去。收錄它們，是因為這兩道都是極好的題目；每一節都會給出所需的知識。

如何閱讀本章

動手計算之前，先問問題目想要的是哪個變量。等時曲線用弧長，耦合彈簧用 $x \pm y$，螞蟻用繩子的分數，鏟雪車則把時間看成距離的函數：在每一種情況下，選對了變量，難題就會變成熟悉的問題。

11.1 共振

題目

STEP II 1991，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/91-s2-q12`

引子

推鞦韆上的小孩時，只要每一下都推在恰當的時刻，輕輕的推力也會累積成大幅的擺動；節奏不對的話，就幾乎毫無效果。每個振動系統都有其**固有頻率**（natural frequency），以這個頻率去驅動它，所得的反應會大得與所花的力氣完全不成比例。這就是**共振**（resonance）。歌手能夠震碎酒杯，士兵過橋時要打亂步伐，收音機能夠從數以千計的電台中選出其中一個，都是這個緣故。

這道題目搖動一條彈性繩的頂端，問繩子在甚麼情況下會保持拉緊。答案含有因子 $\frac{1}{n-p}$，當驅動頻率 p 趨近固有頻率 n 時，這個因子會趨於無窮大：在接近共振時，輕輕一搖已足以令繩子鬆弛。

動筆之前

想一想 質點懸掛在平衡位置，繩的伸長量（extension）為 c。若繩的彈性模量（modulus）為 λ，則 λ、a、c、m 與 g 之間有甚麼關係？當頂端向下移動了 $b\sin pt$，而質點向下移動了 x 時，伸長量是多少？

在平衡狀態下，張力（tension）等於重量：$\frac{\lambda c}{a} = mg$。當頂端向下位移 $b\sin pt$，而質點向下位移 x 時，繩的長度增加了 $x - b\sin pt$，所以伸長量是 $c + x - b\sin pt$。

逐步解題

階段 1　運動方程

取向下為正，作用於質點的力有它的重量 mg，以及向上的張力 $\frac{\lambda}{a}(c + x - b\sin pt)$：

$$m\ddot{x} = mg - \frac{\lambda}{a}(c + x - b\sin pt) = -\frac{\lambda}{a}(x - b\sin pt),$$

這裏用了 $\frac{\lambda c}{a} = mg$。由於 $\frac{\lambda}{ma} = \frac{g}{c} = n^2$，上式即 $\ddot{x} = -n^2(x - b\sin pt)$。 ■

質點受到一個彈簧力，把它拉向移動中的點 $b\sin pt$。假如沒有驅動，它會以角頻率（angular frequency）n 振動，這就是它的固有頻率。

階段 2　求解

> **想一想**　方程是 $\ddot{x} + n^2x = n^2b\sin pt$。試用 $x = C\sin pt$ 這種形式的特解（particular solution），然後加上 $\ddot{x} + n^2x = 0$ 的通解（general solution）。質點滿足哪些初始條件？

代入 $x = C\sin pt$ 得 $(-p^2 + n^2)C = n^2b$，所以 $C = \frac{n^2b}{n^2 - p^2}$（這裏需要 $p \neq n$）。餘函數（complementary function）是 $D\cos nt + E\sin nt$。質點由平衡位置從靜止開始運動：$x(0) = 0$ 令 $D = 0$，而 $\dot{x}(0) = 0$ 給出 $Cp + En = 0$，所以 $E = -\frac{nbp}{n^2 - p^2}$。綜合起來，

$$x(t) = \frac{bn}{n^2 - p^2}(n\sin pt - p\sin nt). \qquad ■$$

階段 3　保持拉緊

> **想一想**　繩子拉緊，當且僅當它的伸長量為正。試把伸長量 $c + x - b\sin pt$ 以 t 表示。

從 $x(t)$ 減去 $b\sin pt$，

$$\begin{aligned} x - b\sin pt &= \frac{b}{n^2 - p^2}\left(n^2\sin pt - np\sin nt - (n^2 - p^2)\sin pt\right) \\ &= \frac{bp}{n^2 - p^2}(p\sin pt - n\sin nt). \end{aligned}$$

所以繩子自始至終保持拉緊，當且僅當

$$c + \frac{bp}{n^2 - p^2}(p\sin pt - n\sin nt) \geq 0 \qquad \text{對所有} t \geq 0.$$

想一想 $p\sin pt - n\sin nt$ 最小可以負到甚麼程度？

由於 $|\sin| \le 1$，我們有 $p\sin pt - n\sin nt \ge -(p+n)$。所以伸長量至少是

$$c - \frac{bp(n+p)}{n^2-p^2} = c - \frac{bp}{n-p},$$

當 $pb < (n-p)c$ 時，這個值為正。 ■

停一停，試一試 當 p 趨近 n 時，這個條件會變成怎樣？當 $p = 0.9n$ 時，搖動的振幅（amplitude）b 要有多小？當 $p = 0.99n$ 時又如何？

回顧

- **由平衡位置量起。**利用平衡時的伸長量 c，重力便從方程中消去了。
- **特解加餘函數。**頻率為 p 的受迫響應，加上頻率為 n 的自由振動，其中的常數由初始條件決定。
- **估界，不必求解。**我們沒有準確求出伸長量的最小值；對正弦作一個粗略的估界，已足以得出一個充分條件（sufficient condition）。

背後的數學

共振極限

在解中令 $p \to n$。分子和分母都趨於零，計算極限(或使用洛必達法則(l'Hôpital's rule)，對 p 求導）可得

$$x(t) \to \frac{b}{2}(\sin nt - nt\cos nt).$$

$nt\cos nt$ 這一項會無止境地增大：恰好以固有頻率驅動時，振幅隨時間線性增長，每一下推動都與運動同步，為它再添一分。在略低於共振的頻率下，兩個頻率時而同步、時而錯開，產生拍（beats）：振幅以緩慢的頻率 $n-p$ 時漲時落（圖 11.1.1）。兩條結他弦的音高幾乎調準時，聽到的也是同樣的拍。

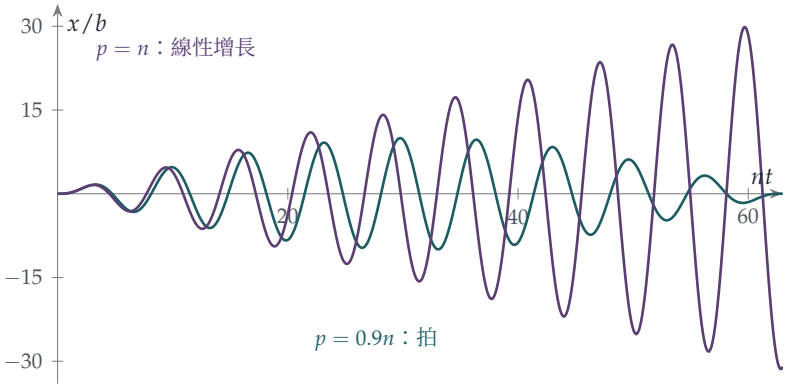


圖 11.1.1　質點的響應（不考慮繩子鬆弛），以搖動振幅 b 為單位。當 $p = 0.9n$（藍綠色）時，響應增大至約為 b 的十倍，然後再度消退，拍的周期為 $\frac{2\pi}{n-p}$。在恰好共振，即 $p = n$（紫紅色）時，振幅無止境地增長。

阻尼與共振曲線

現實中的系統會因摩擦或空氣阻力而損失能量，這就限制了振幅的增長。有阻尼（damping）時，系統對頻率為 p 的驅動力的穩態響應，其振幅在固有頻率附近達到峯值；峯愈尖銳，阻尼就愈小。工程師用品質因數（quality factor）Q 來描述這一點：粗略地說，它是系統在失去大部分能量之前所完成的振動次數。酒杯的 Q 數以千計，所以一個恰好等於其固有頻率的持續音可以把它震碎。汽車的懸掛系統則刻意設計成 Q 值很低，以免路面顛簸令車身彈跳不止。

會動的橋

1831 年，一隊士兵以整齊步伐操過曼徹斯特附近的布勞頓吊橋（Broughton Suspension Bridge），令橋身振動，最終倒塌；自此以後，軍隊過橋時都奉命打亂步伐。倫敦千禧橋（London Millennium Bridge）於 2000 年 6 月啟用，人羣一踏上橋，橋身便開始左右搖擺：行人不自覺地調整步伐去配合搖擺，同步地為它注入能量，結果該橋關閉了兩年，以便加裝阻尼器。1940 年塔科馬海峽吊橋（Tacoma Narrows Bridge）那次著名的倒塌，常被說成是共振，但其實是另一種現象，即氣動彈性顫振（aeroelastic flutter），當中風所施的力取決於橋本身的運動。要分析它，需要用到自激振動（self-excited oscillation）和極限環（limit cycle）的概念（Section 8.4）。

思維技巧

把受迫響應與自由響應疊加

對於受周期力驅動的線性方程，其解是驅動頻率下的響應，加上固有頻率下的自由振動，後者按初始條件選定。

這個方法之所以可行，全靠線性（linearity）：解可以相加。在大學裏，同一原理應用於多個驅動頻率之和，便引出傅立葉分析（Fourier analysis）、工程學中的傳遞函數（transfer function），以及物理學中的格林函數（Green's function）。

延伸練習

1 **下墜的升降機。**試做 STEP II 1996，第 11 題（`step.maths.org/questions/96-s2-q11`）：升降機內有一個質量繫於彈性繩上，升降機先開始加速，其後停止加速。留意最終的振幅如何取決於加速持續的時間。

2 **共振之時。**試用特解 $x = Ct\cos nt$，直接解 $\ddot{x} + n^2 x = n^2 b \sin nt$，其中 $x(0) = \dot{x}(0) = 0$。驗證所得結果與上文的極限一致。

3 **拍。**證明 $n\sin pt - p\sin nt$ 可寫成兩項之和，而當 p 接近 n 時，其中一項近似於 $2n\cos\frac{(n+p)t}{2}\sin\frac{(p-n)t}{2}$。由此解釋拍的周期為何是 $\frac{2\pi}{n-p}$。

4 **有阻尼的共振。**對於 $\ddot{x} + 2k\dot{x} + n^2 x = F\cos pt$，其中 $k > 0$ 且很小，求穩態振幅，並證明當 $p^2 = n^2 - 2k^2$ 時振幅最大。

延伸閱讀

- **A. P. French, *Vibrations and Waves***（MIT Introductory Physics Series）。第 4 章講述受迫振動與共振，是這個課題的經典論述。
- **Steven Strogatz, *Sync***（Penguin）。講述事物如何趨於同步，從螢火蟲談到千禧橋。
- **Henry Petroski, *To Engineer Is Human***（Vintage）。講述工程上的失敗個案，包括歷史專欄中提到的幾座橋。

11.2 簡正模式

題目

STEP III 2014，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/14-s3-q10`

引子

一個掛在彈簧上的質點，會以簡單而規律的方式上下彈動。把兩個質點用彈簧連起來，運動看起來就複雜得多：每個質點都不規則地忽動忽停，能量在兩者之間來回流轉。然而在這複雜的表象背後，藏着兩種極其簡單的運動，稱為**簡正模態**（normal modes）：兩個質點或一起同向運動，或恰好反向運動，而每個模態各以自己的頻率（frequency）振盪。系統的任何運動，都是這兩個模態的組合。

本題要找出夾在三條彈簧之間的兩個質點的簡正模態，然後證明一個令人驚訝的結論：若系統由靜止開始，而其中一個質點有初始位移，則系統永遠不會回到起始位置。原因在於兩個頻率之比是 $\sqrt{3}$，而 $\sqrt{3}$ 是無理數（irrational number）。

動筆之前

想一想 設 X 與 P 的距離為 $a+x$，Y 與 Q 的距離為 $a+y$，那麼中間的彈簧 XY 有多長？當 x 和 y 皆為正數時，哪些彈簧被拉長，哪些被壓縮？

三段長度之和為 $3a$，所以 $XY = 3a-(a+x)-(a+y) = a-x-y$。當 $x, y > 0$ 時，外側兩條彈簧被拉長（伸長量分別為 x 和 y），中間一條則被壓縮（壓縮量為 $x+y$）。

逐步解題

階段 1　運動方程

沿着由 P 指向 Q 的方向量度 x。彈簧 PX 的伸長量（extension）為 x，所以它的張力（tension）$\frac{\lambda x}{a}$ 把 X 拉回 P 的方向。彈簧 XY 的伸長量為 $-(x+y)$，所以它的張力 $-\frac{\lambda(x+y)}{a}$ 把 X 拉向 Y（負的張力即是推力）。因此

$$m\ddot{x} = -\frac{\lambda x}{a} - \frac{\lambda(x+y)}{a} = -\frac{\lambda}{a}(2x+y). \qquad \blacksquare$$

根據對稱性（把 X 與 Y、P 與 Q 的角色互換），

$$m\ddot{y} = -\frac{\lambda}{a}(2y+x).$$

階段 2　解耦

想一想　每條方程都同時涉及 x 和 y。把兩條方程怎樣組合，才會只涉及 x 和 y 的某一個組合？

把兩式相減及相加，得

$$m(\ddot{x}-\ddot{y}) = -\frac{\lambda}{a}(x-y), \qquad m(\ddot{x}+\ddot{y}) = -\frac{3\lambda}{a}(x+y).$$

兩者都是簡諧運動（simple harmonic motion）的方程：$x-y$ 以 $\omega^2 = \frac{\lambda}{ma}$ 振盪，而 $x+y$ 則以 $3\omega^2$ 振盪。

階段 3　初始條件

起初 $XP = a$ 及 $YQ = \frac{1}{2}a$，所以 $x = 0$ 及 $y = -\frac{1}{2}a$，而且兩個質點都處於靜止。因此在 $t = 0$ 時，$x - y = \frac{1}{2}a$ 及 $x + y = -\frac{1}{2}a$，而兩者的變化率皆為零，由此得

$$x - y = \tfrac{1}{2}a\cos\omega t, \qquad x + y = -\tfrac{1}{2}a\cos\sqrt{3}\,\omega t.$$

所以 $A = \frac{1}{2}a$ 及 $B = 0$。 $\blacksquare$

階段 4　永不返回

想一想　把 y 寫成 t 的函數。若 Y 要回到起始位置，便須 $y = -\frac{1}{2}a$。這時兩個餘弦必須各等於甚麼？

兩式相減，得 $y = \frac{1}{2}\big((x+y)-(x-y)\big) = -\frac{1}{4}a\big(\cos\sqrt{3}\,\omega t + \cos\omega t\big)$。要 $y = -\frac{1}{2}a$，便需要 $\cos\sqrt{3}\,\omega t + \cos\omega t = 2$；由於每個餘弦最大只能是 1，兩者必須都等於 1：

$$\omega t = 2\pi k, \qquad \sqrt{3}\,\omega t = 2\pi j$$

其中 k 和 j 為整數。當 $t > 0$ 時，$k \geq 1$，兩式相除得 $\sqrt{3} = \frac{j}{k}$，但 $\sqrt{3}$ 是無理數，所以這是不可能的。因此 Y 永遠不會回到起始位置。 $\blacksquare$

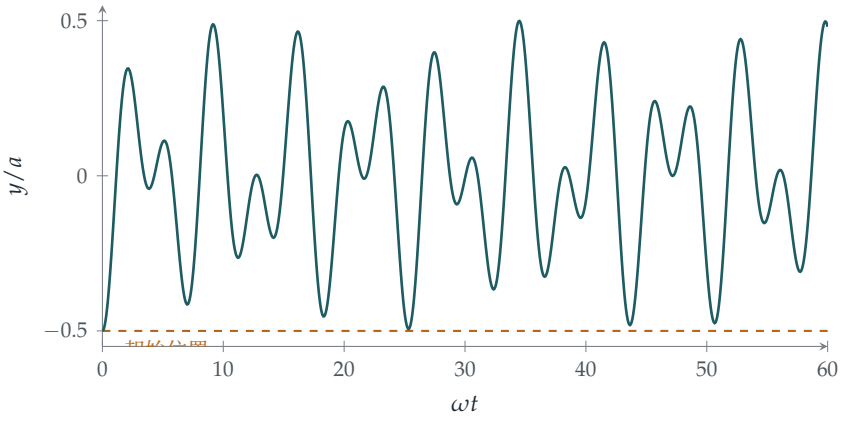


圖 11.2.1 質點 Y 的位移 y。它是兩個振盪之和，而兩者頻率之比為 $\sqrt{3}:1$，所以圖形永不重複。Y 一次又一次地接近起始位置（虛線），例如在 $\omega t \approx 25$ 附近（該處 $\omega t \approx 8\pi$ 及 $\sqrt{3}\,\omega t \approx 14\pi$），但永遠不會剛好到達。

回顧

- **找出能夠解耦的組合。**$x-y$ 和 $x+y$ 各自滿足一條屬於自己的方程。
- **對稱性會提示這些組合。**把左右兩邊互換，系統保持不變，所以對稱和反對稱的組合是很自然的選擇。
- **無理性會帶來物理後果。**運動之所以不是週期性的，是因為頻率之比是無理數。

背後的數學

簡正模態就是特徵向量

把兩條方程寫成一條矩陣方程：

$$m\begin{pmatrix}\ddot{x}\\ \ddot{y}\end{pmatrix} = -\frac{\lambda}{a}\begin{pmatrix}2 & 1\\ 1 & 2\end{pmatrix}\begin{pmatrix}x\\ y\end{pmatrix}.$$

能夠解耦的組合，正是這個矩陣的特徵向量（eigenvectors）：$(1,-1)$，其特徵值（eigenvalue）為 1；以及 $(1,1)$，其特徵值為 3（Section 7.2）。在第一個模態中，$x=-y$：兩個質點沿直線同向運動（記住 x 和 y 是從相反的兩端量度的），中間的彈簧長度保持不變。在第二個模態中，$x=y$：兩個質點反向運動，把中間的彈簧一壓一拉，頻率因而較高。頻率的平方就是特徵值乘以 $\frac{\lambda}{ma}$。

無論有多少個質點和彈簧，這個方法都適用。一個由 N 個耦合振子（coupled oscillators）組成的系統有 N 個簡正模態，它們由一個對稱矩陣的特徵向量給出，而譜定理（spectral

theorem）保證這些模態必定存在，並且互相垂直。分子以簡正模態振動，化學家正是憑分子所吸收的紅外光來辨認它們。建築物和橋樑也有簡正模態，工程師必須令這些模態的頻率遠離風和地震的頻率。當質點的數目趨向無限時，簡正模態就變成振動弦線的諧波（harmonics），這正是傅立葉分析（Fourier analysis）的起點。

接近而不到達

雖然 Y 永遠不會剛好返回，但它可以接近到任意程度。原因是克羅內克（Kronecker）的一條定理：若 α 是無理數，則各倍數 $\alpha, 2\alpha, 3\alpha, \ldots$ 取模 1 之後，可以任意接近 $[0,1]$ 內的每一個數。因此我們可以找到一些時刻，使 ωt 是 2π 的倍數，而 $\sqrt{3}\,\omega t$ 與另一個這樣的倍數之差小於任何指定的容許誤差。這種運動是*準週期*（quasi-periodic）的：它在一個環面（torus）上不斷纏繞（每個模態對應一個角），卻永遠不會閉合。有理與無理頻率比之間的這種區別，同樣是關於太陽系穩定性的 KAM 定理的核心。

思維技巧

以變數代換把方程解耦

當方程互相耦合時，設法找出各自獨立演變的變數組合。在線性問題中，這些組合就是特徵向量。

這與 Section 10.6 中質心和相對坐標的想法如出一轍。到了大學，它就成為對角化（diagonalisation）：選取一組令矩陣成為對角矩陣的基（basis），一組耦合的方程便會分拆成各自獨立的方程。

延伸練習

1. **能量。**證明總能量 $\tfrac{1}{2}m(\dot{x}^2+\dot{y}^2)+\tfrac{\lambda}{2a}\bigl(x^2+y^2+(x+y)^2\bigr)$ 是常數，並證明它可以分拆成兩項之和，每個簡正模態各佔一項，而每一項本身都是常數。

2. **純模態。**應該怎樣讓兩個質點開始運動，才能使運動是週期性的，且週期為 $\frac{2\pi}{\omega}$？若週期為 $\frac{2\pi}{\sqrt{3}\,\omega}$ 又如何？

3. **較硬的中間彈簧。**若中間彈簧的彈性模量（modulus of elasticity）是 μ 而不是 λ，求兩個簡正模態的頻率。當 μ/λ 取哪些值時，所有運動都是週期性的？

4. **三個質點。**三個質量相等的質點排成一線，由四條相同的彈簧連接在兩個固定點之間。求對應的矩陣及其特徵值 $(2-\sqrt{2},\,2,\,2+\sqrt{2})$，並描述三個簡正模態。

延伸閱讀

- **A. P. French, *Vibrations and Waves***（MIT Introductory Physics Series），第 5 章，講述耦合振子與簡正模態。
- **Gilbert Strang, *Introduction to Linear Algebra***（Wellesley-Cambridge），關於特徵值的一章，當中以彈簧與質點作為主要例子。
- **Steven Strogatz, *Nonlinear Dynamics and Chaos***（CRC Press），第 8 章，講述環面上的準週期運動。

11.3 真實的擺

題目

STEP III 1998，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/98-s3-q11`

引子

每個修讀物理的學生都學過：單擺（pendulum）的週期（period）是 $2\pi\sqrt{l/g}$，與擺動的幅度無關。伽利略（Galileo）對此深信不疑，相傳他曾在比薩大教堂以自己的脈搏為一盞搖晃的吊燈計時；擺鐘之所以能夠報時，靠的也正是這一點。然而這並不完全正確。這條公式源自近似式 $\sin\theta \approx \theta$，而擺動弧度大的單擺，所需的時間會較長。在 90° 時，週期長了 18%；若單擺從幾乎倒立的位置釋放，所需的時間則要多長有多長。

本題先把精確的週期寫成一個積分，再求出中等擺幅下的第一個修正項。這個積分無法用初等函數求出：它是一個*橢圓積分*（elliptic integral），而人們如何馴服這類積分的故事，由高斯（Gauss）的算術幾何平均一直延伸到橢圓曲線的理論。

動筆之前

想一想 擺錘沿半徑為 l 的圓周運動。以 θ 和 $\dot{\theta}$ 表示，它的速率和高度（相對於最低點）分別是甚麼？

速率是 $l\dot{\theta}$，而距最低點的高度是 $l(1-\cos\theta)$。由於繩子不作功（其張力恆與運動方向垂直），能量守恆（conservation of energy）。

逐步解題

階段 1 能量

在最大擺角 γ 處，擺錘瞬時靜止。由能量守恆得

$$\tfrac{1}{2}ml^2\dot\theta^2 + mgl(1-\cos\theta) = mgl(1-\cos\gamma),$$

兩邊除以 ml，得 $\tfrac{1}{2}l\dot\theta^2 = g(\cos\theta - \cos\gamma)$。 ■

階段 2　把週期寫成積分

想一想　怎樣用半角表示 $\cos\theta - \cos\gamma$？另外，由 $\theta = 0$ 擺到 $\theta = \gamma$ 佔了週期的多少？

利用 $\cos\theta = 1 - 2\sin^2\tfrac{\theta}{2}$，可得 $\cos\theta - \cos\gamma = 2(\sin^2\tfrac{\gamma}{2} - \sin^2\tfrac{\theta}{2})$，所以

$$\dot\theta = 2\sqrt{\frac{g}{l}}\left(\sin^2\frac{\gamma}{2} - \sin^2\frac{\theta}{2}\right)^{1/2}.$$

根據對稱性，由 0 擺到 γ 需時四分之一個週期，所以

$$P = 4\int_0^{\gamma}\frac{\mathrm{d}\theta}{\dot\theta} = 2\sqrt{\frac{l}{g}}\int_0^{\gamma}\left(\sin^2\frac{\gamma}{2} - \sin^2\frac{\theta}{2}\right)^{-1/2}\mathrm{d}\theta. \qquad ■$$

階段 3　消除奇點

想一想　被積函數（integrand）在 $\theta = \gamma$ 處變成無窮大。設 $k = \sin\tfrac{\gamma}{2}$，並作代換（substitution）$\sin\tfrac{\theta}{2} = k\sin\phi$。$0 \le \theta \le \gamma$ 對應於 ϕ 的甚麼範圍？被積函數會變成怎樣？

當 θ 由 0 變到 γ 時，ϕ 由 0 變到 $\tfrac{\pi}{2}$。求導得 $\tfrac{1}{2}\cos\tfrac{\theta}{2}\,\mathrm{d}\theta = k\cos\phi\,\mathrm{d}\phi$，所以

$$\mathrm{d}\theta = \frac{2k\cos\phi\,\mathrm{d}\phi}{\cos\frac{\theta}{2}} = \frac{2k\cos\phi\,\mathrm{d}\phi}{\sqrt{1-k^2\sin^2\phi}}.$$

另一方面，$\sin^2\tfrac{\gamma}{2} - \sin^2\tfrac{\theta}{2} = k^2(1-\sin^2\phi) = k^2\cos^2\phi$。那些棘手的因子互相抵銷：

$$P = 2\sqrt{\frac{l}{g}}\int_0^{\pi/2}\frac{1}{k\cos\phi}\cdot\frac{2k\cos\phi}{\sqrt{1-k^2\sin^2\phi}}\,\mathrm{d}\phi = 4\sqrt{\frac{l}{g}}\int_0^{\pi/2}\frac{\mathrm{d}\phi}{\sqrt{1-k^2\sin^2\phi}}.$$

由於 $k < 1$，新的被積函數性質十分良好。

階段 4　第一個修正項

想一想 當 k 很小時，用二項式定理(binomial theorem)展開 $(1-k^2\sin^2\phi)^{-1/2}$。$\int_0^{\pi/2}\sin^2\phi\,d\phi$ 等於甚麼？

準確至一階，$(1-k^2\sin^2\phi)^{-1/2}\approx 1+\frac{1}{2}k^2\sin^2\phi$，而 $\int_0^{\pi/2}\sin^2\phi\,d\phi=\frac{\pi}{4}$。所以

$$P\approx 4\sqrt{\frac{l}{g}}\left(\frac{\pi}{2}+\frac{k^2}{2}\cdot\frac{\pi}{4}\right)=2\pi\sqrt{\frac{l}{g}}\left(1+\frac{k^2}{4}\right).$$

當 γ 很小時，$k=\sin\frac{\gamma}{2}\approx\frac{\gamma}{2}$，故 $k^2\approx\frac{\gamma^2}{4}$，於是

$$P\approx 2\pi\sqrt{\frac{l}{g}}\left(1+\frac{\gamma^2}{16}\right). \qquad \blacksquare$$

擺幅為 20°（$\gamma\approx 0.35$）時，修正量不足 1%；在 60° 時則約為 7%（圖 11.3.1）。

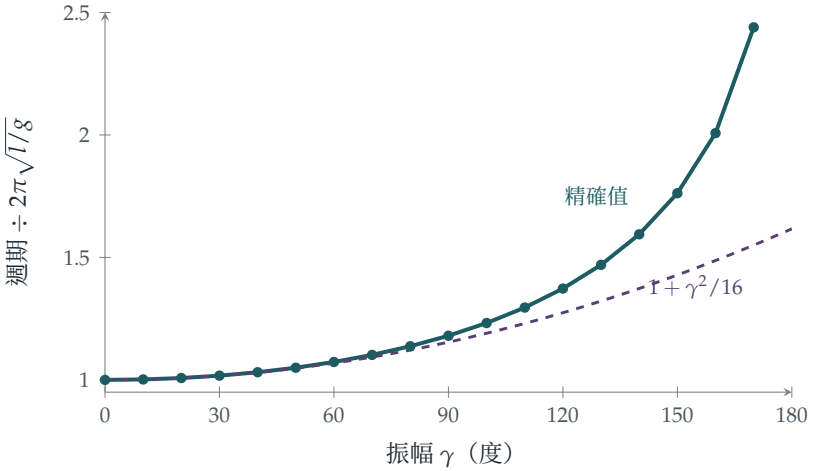


圖 11.3.1 單擺的精確週期（藍綠色）相對於小角度值的比，對振幅（amplitude）作圖，並與近似式 $1+\gamma^2/16$（紫紅色虛線）比較。在大約 60° 以內，這個近似極佳。當 $\gamma\to 180°$ 時，真正的週期趨向無窮大：一個倒立平衡的單擺要花無限長的時間才會倒下。

回顧

- **能量給出首次積分（first integral）。**二階方程變成了一階方程，而週期則變成了一個積分。
- **代換馴服奇點（singularity）。**被積函數在轉向點處趨向無窮大；代換 $\sin\frac{\theta}{2}=k\sin\phi$ 消除了這個問題。

- **先求精確，再作近似。**我們先求出精確的公式，然後將它展開，這樣同時可以知道近似的效果有多好。

背後的數學

橢圓積分

積分

$$K(k) = \int_0^{\pi/2} \frac{\mathrm{d}\phi}{\sqrt{1 - k^2 \sin^2 \phi}}$$

稱為第一類完全橢圓積分（complete elliptic integral of the first kind），這個名稱的由來，是因為與它相近的另一個積分給出橢圓的周界。精確的週期是 $P = 4\sqrt{l/g}\,K\bigl(\sin\tfrac{\gamma}{2}\bigr)$。把被積函數完全展開，再利用 Section 4.4 中的沃利斯積分（Wallis integrals）$\int_0^{\pi/2} \sin^{2n}\phi\,\mathrm{d}\phi$ 逐項積分，便得

$$P = 2\pi\sqrt{\frac{l}{g}}\Bigl(1 + \bigl(\tfrac{1}{2}\bigr)^2 k^2 + \bigl(\tfrac{1\cdot 3}{2\cdot 4}\bigr)^2 k^4 + \bigl(\tfrac{1\cdot 3\cdot 5}{2\cdot 4\cdot 6}\bigr)^2 k^6 + \cdots\Bigr).$$

$K(k)$ 無法以有限個熟悉的函數組合而成；它是一個名副其實的新函數。

高斯的非凡公式

計算 K 有一個好得多的方法。你在 Section 2.6 已見過算術幾何平均（arithmetic–geometric mean）：由 a 和 b 開始，反覆以它們的算術平均和幾何平均取代它們。高斯在 1799 年發現

$$K(k) = \frac{\pi}{2\,\mathrm{AGM}\bigl(1, \sqrt{1 - k^2}\bigr)},$$

所以振幅為 γ 的單擺，其週期恰好是

$$P = \frac{2\pi\sqrt{l/g}}{\mathrm{AGM}\bigl(1, \cos\tfrac{\gamma}{2}\bigr)}.$$

AGM 收斂得極快：擺幅在大約 160° 以內時，只需三四步便可求得準確至十位有效數字的週期，圖 11.3.1 中的數值正是這樣算出來的。接近 180° 時，兩個初始值 1 和 $\cos\frac{\gamma}{2}$ 相距甚遠，需要多走一兩步，但即使在 179°，五步也已足夠。高斯在日記中寫道，這項發現「開闢了一個全新的分析領域」，而他說得一點不錯：把橢圓積分反演，引領阿貝爾（Abel）和雅可比（Jacobi）發現了橢圓函數（elliptic functions），而橢圓函數又引出了橢圓曲線（elliptic curves）。橢圓曲線如今是數論的核心（它們正是證明費馬最後定理（Fermat's Last Theorem）的關鍵），也是保障互聯網安全的密碼學的核心。

伽利略的吊燈與惠更斯的時鐘

伽利略大約在 1602 年提出，單擺的週期與振幅無關；對於他所研究的小幅擺動，這是個很好的近似。惠更斯（Christiaan Huygens）意識到這並不精確，並在 1659 年找到了一條曲線，沿着它擺動時週期確實與振幅無關：那就是擺線（cycloid），即 Section 11.4 的主題。他由 1656 年起製造的擺鐘，把計時的誤差由每天約一刻鐘減至約十五秒。航海天文鐘（marine chronometer）解決了在海上測定經度的難題，它後來完全捨棄了單擺，但直至 1920 年代，擺鐘仍然是世上最準確的計時器。

思維技巧

由精確積分到級數近似

當一個量由無法求值的積分給出時，先把它變換成性質良好的形式，然後展開被積函數並逐項積分。

物理學家正是這樣為理想化的結果求出修正項的，這也是微擾理論（perturbation theory）的起點。到了大學，你會學到逐項積分在甚麼情況下才成立（一致收斂（uniform convergence）），以及怎樣估計截斷級數所造成的誤差。

延伸練習

1 **繫着的質點。**試做 STEP III 2015，第 9 題（`step.maths.org/questions/15-s3-q9`）：這個振子即使在小幅振動時，週期也取決於振幅。為甚麼它會這樣，而單擺卻不會？

2 **下一項。**把二項式展開多取一項，證明 $P \approx 2\pi\sqrt{\frac{l}{g}}\left(1+\frac{\gamma^2}{16}+\frac{11\gamma^4}{3072}\right)$。

3 **徒手計算 AGM。**用三步 AGM 估算振幅為 90° 的單擺的週期，並與圖中的數值 1.1803 比較。

4 **時鐘。**一座擺鐘經調校後，在振幅為 4° 時走時完全準確。若它的振幅增至 5°，它每天會慢多少秒？

延伸閱讀

- **David Morin, *Introduction to Classical Mechanics***（Cambridge University Press），第 4

章，當中討論了大振幅的單擺。

- **Jonathan Borwein & Peter Borwein, *Pi and the AGM***（Wiley）。高斯的發現及其應用，包括計算 π 的最快方法。
- **Dava Sobel, *Longitude***（Fourth Estate）。海上計時的故事，講述單擺以外的發展。

11.4 等時曲線

題目

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`

引子

在一個碗內同時放開兩粒彈珠，一粒靠近碗邊，另一粒只比碗底高一點。哪一粒先到碗底？在普通的碗裏，通常是起點較低的那粒勝出。不過有一種碗，無論彈珠從哪裏出發，總是同時到達。這就是**等時曲線**（tautochrone，希臘文意為「相同時間」），它的橫切面是一條**擺線**（cycloid），即滾動的輪子邊緣上一點所描出的曲線。

惠更斯（Huygens）在 1659 年發現了這個性質，並用它設計出一種週期與擺幅無關的擺鐘，修正了 Section 11.3 中發現的缺陷。本題只用寥寥數行便證明了等時性質，關鍵在於選對變數：用沿曲線量度的距離，而不是水平位置。

動筆之前

想一想 碗面是光滑的。有哪些力作用在質點上？其中哪些力有作功？

重力和法向反作用力（normal reaction）。法向反作用力垂直於碗面，因而垂直於運動方向，所以不作功。只有重力作功，故能量守恆。

逐步解題

階段 1　沿曲線的能量

由於 s 是弧長（arc length），質點沿碗面的速率是 $\left|\frac{\mathrm{d}s}{\mathrm{d}t}\right|$。由能量守恆（conservation of

energy）得 $\frac{1}{2}m\dot{s}^2 + mgy =$ 常數，所以 $\dot{s}^2 + 2gy$ 是常數。 ■

階段 2　以 s 表示的簡諧運動

想一想　利用 $s = 2\sqrt{ky}$，把 y 換成 s 的函數。所得的方程描述的是哪一種運動？

由於 $y = \frac{s^2}{4k}$，能量方程變成

$$\dot{s}^2 + \frac{g}{2k}s^2 = \text{常數} = \frac{g}{2k}s_0^2,$$

其中 $s_0 = 2\sqrt{kh}$ 是起始弧長（質點由靜止出發）。求導後得 $2\dot{s}\ddot{s} + \frac{g}{k}s\dot{s} = 0$，所以

$$\ddot{s} = -\frac{g}{2k}\,s.$$

這是以 s 為變數的簡諧運動（simple harmonic motion），其中 $\omega^2 = \frac{g}{2k}$，與振幅（amplitude）無關：$s = s_0\cos\omega t$。

階段 3　到達底部所需的時間

質點經過四分之一個週期後到達 V（該處 $s = 0$）：

$$\frac{\pi}{2\omega} = \frac{\pi}{2}\sqrt{\frac{2k}{g}} = \pi\sqrt{\frac{k}{2g}},$$

這與 h 無關。 ■

停一停，試一試　為甚麼這樣行得通？沿曲線的「回復力」是 g 乘以斜率，即 $g\frac{\mathrm{d}y}{\mathrm{d}s} = \frac{g}{2k}s$：恰好與沿曲線量度、離底部的距離成正比。這正是簡諧運動的定義性質。起點較高的質點要走的路較長，但所受的推動也按比例加大。

階段 4　兩個質點

想一想　兩個質點何時到達 V？到達時速率是多少？誰的動量較大？

兩者從相反的兩邊，在同一時刻 $T = \pi\sqrt{k/(2g)}$ 到達 V。在 V 的速率是 ω 乘以 s 的振幅，與 $\sqrt{\text{高度}}$ 成正比。設較重的質點從高度 h 放開，在 V 的速率為 U；較輕的質點從 $\alpha^2 h$ 放開，速率為 αU。它們的動量（momentum）分別是 $\alpha m \cdot U$ 和 $m \cdot \alpha U$，方向相反：大小相等，方向相反。碰撞時的總動量為零。

想一想　總動量為零，而恢復係數（coefficient of restitution）為 α，碰撞剛結束時兩者的速度是多少？

動量保持為零，所以碰撞後兩個質點仍然朝相反方向運動，動量大小相等，而它們的相對速率是接近速率 $\alpha U + U$ 的 α 倍。若較輕的質點（質量 αm）此時以速率 w 運動，則較重的質點（質量 m）以速率 αw 運動，且 $w + \alpha w = \alpha(1+\alpha)U$，所以 $w = \alpha U$。每個質點反彈時，速率都乘以 α：質量為 αm 的質點以 αU 離開，質量為 m 的質點以 $\alpha^2 U$ 離開。

階段 5 **整個運動**

每個質點現在沿自己的一邊爬升，所達高度縮減為原來的 α^2 倍（因為高度與速率的平方成正比）：質量為 αm 的質點升至 $\alpha^2 h$，質量為 m 的質點升至 $\alpha^4 h$。由等時性質，兩者在 $2T$ 之後恰好同一時刻回到 V，動量仍然大小相等、方向相反。同樣的碰撞再次發生，只是所有速率都按 α 的比例縮小。因此：

- 兩個質點總是在 V 碰撞，時刻為 $T, 3T, 5T, \ldots$；
- 第 n 次碰撞後，它們分別升至高度 $\alpha^{2n} h$ 和 $\alpha^{2n+2} h$；
- 運動按幾何級數式衰減，但永不停止：碰撞有無限多次，並且永遠持續下去。 ■

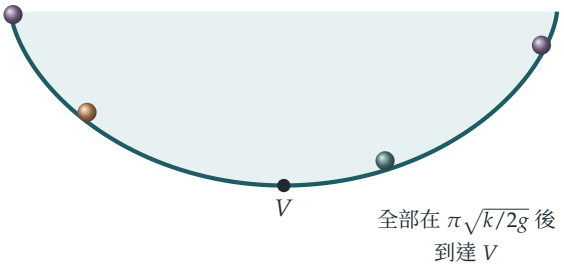


圖 11.4.1 等時曲線：一條倒轉的擺線，由沿直線滾動的輪子邊緣上一點描出。彈珠從曲線上任何位置由靜止放開，都在同一時刻到達最低點 V。對於由半徑為 R 的輪子生成的擺線，弧長與高度滿足 $s^2 = 8Ry$，所以 $k = 2R$。

回顧

- **選擇自然的變數。**以弧長 s 表示時，運動恰好是簡諧運動。以 x 或 y 表示，則一團糟。
- **能量給出方程。**對能量方程求導，便直接得出運動方程。
- **時間上的對稱。**由於每一趟到達底部的行程所需時間都相同，兩個質點永遠保持同步，碰撞也就原樣重複。

背後的數學

為甚麼是擺線？

半徑為 R 的輪子在一條直線下方滾動，描出擺線 $x = R(t + \sin t)$，$y = R(1 - \cos t)$。於是 $\frac{dx}{dt} = R(1 + \cos t)$ 及 $\frac{dy}{dt} = R\sin t$，所以沿曲線的速率是 $R\sqrt{2 + 2\cos t} = 2R\cos\frac{t}{2}$。積分後，由底部起計的弧長是 $s = 4R\sin\frac{t}{2}$，而 $y = 2R\sin^2\frac{t}{2}$。因此 $s^2 = 8Ry$：擺線正是滿足 $s = 2\sqrt{ky}$ 的曲線，其中 $k = 2R$。反過來說，要求回復力與 s 成正比，就迫使曲線必定是擺線。

最速降線

1696 年，約翰·伯努利（Johann Bernoulli）向歐洲的數學家提出挑戰：一粒珠子從一點滑到較低的另一點（不在其正下方），沿哪一條曲線所需的時間最短？答案是最速降線（brachistochrone），它同樣是一條擺線。據說牛頓（Newton）收到題目後，一個晚上就解決了；他匿名發表了自己的解法，但伯努利認出了作者，說是「從爪認出獅子」。這個問題是*變分法*（calculus of variations）的開山例子，你在 Section 4.7 已經見過：找出使某個由積分定義的量取最小值的曲線。

惠更斯的擺線鐘

惠更斯知道單擺計時並不完美，因為它的週期取決於振幅。他在《*Horologium Oscillatorium*》（1673 年）中說明了怎樣令擺錘沿擺線運動：把擺線掛在兩塊彎成擺線形狀的「夾板」之間，擺線便會貼着夾板捲繞，使擺錘本身描出一條擺線（擺線的漸伸線（involute）是另一條擺線）。實際上，夾板因摩擦而引入的誤差比它消除的還要多，鐘錶匠於是改為保持小幅擺動。然而惠更斯的分析開創了漸屈線（evolute）與漸伸線的理論，他這部著作也是十七世紀力學的偉大作品之一。

思維技巧

找出令運動變得簡單的變數

沿曲線的運動，往往以沿曲線的距離來表示最爲簡單。要找出令方程變成線性的那個變數。

到了大學，這會發展成拉格朗日力學（Lagrangian mechanics）中廣義坐標（generalised coordinates）的概念：任何能描述系統位形的變數都可以採用，選得好，就能令問題變得可解。

延伸練習

1 **不同的高度。**一個質點在等時曲線碗內從高度 h 放開。證明它在 V 的速率是 $\sqrt{2gh}$（由能量守恆，必然如此），並驗證這與 ωs_0 一致。

2 **普通的碗。**在半徑為 R 的半球形碗內，比較從很小的高度到達底部所需的時間與從碗邊出發所需的時間。哪一個較長？（利用 Section 11.3 的結果。）

3 **質量相等。**把題目的第二部分重做一次，這次兩個質點質量同為 m，分別從高度 h_1 和 h_2 放開，恢復係數為 e。它們是否仍然總是在 V 碰撞？

4 **擺線的弧長。**證明擺線的一拱長度為 $8R$，而拱下的面積為 $3\pi R^2$，即滾動圓面積的三倍。

延伸閱讀

- **George Simmons, *Differential Equations with Applications and Historical Notes***（CRC Press）。書中包括等時曲線和最速降線，並附有相關歷史。
- **E. A. Whitman, “Some historical notes on the cycloid”**，*American Mathematical Monthly* 50 (1943)。講述擺線為何被稱為「幾何學家的海倫」。
- **Paul Nahin, *When Least Is Best***（Princeton University Press）。以歷史故事的方式講述最速降線及其他最優化問題。

11.5 加速火車中的重力

題目

STEP II 2021，第 10 題

繼續閱讀之前，請先看題目（原題為英文）。題目載於 OCR 提供下載的 2021 年 STEP 試卷（zip 檔案）內的 STEP II 試卷：

`www.ocr.org.uk/Images/696588-2021-past-papers.zip`

引子

火車駛離車站時，你會感到自己被壓向椅背；桌上那杯咖啡的液面傾斜了；地板上的球滾向車尾。從車廂內看，一切就像重力傾斜了一樣：除了向下，還稍微指向後方。愛因斯坦（Einstein）領悟到這不僅僅是比喻，並把這個領悟稱為「我一生中最快樂的想法」：在局部範圍內，重力與加速度是無法區分的。

這道題把一顆珠子穿在彎曲的金屬線上，再放進一列正在加速的火車裏。從火車上看，珠子的運動就像身處一個傾斜了的重力場（gravitational field）之中，傾斜的角度的正切為 a/g。金屬線的底部不再是最低點，珠子會圍繞一個新的平衡位置（equilibrium）振動。

動筆之前

想一想 火車以加速度 a 向西（即沿負 x 方向）加速。對車上的觀察者來說，質量為 m 的物體看似受到怎樣的「假想」力作用？

在火車的參考系(frame)中，只要加上一個假想力(fictitious force)$-m\times$(參考系的加速度)，牛頓第二定律（Newton's second law）便仍然成立。參考系在 x 方向的加速度為 $-a$，所以假想力是 $+ma$，指向東。連同重力一併考慮，珠子每單位質量所感受到的有效重力（effective gravity）為 $(a,-g)$：向東傾斜。

逐步解題

階段 1　不能靜止於原點

在原點處，金屬線 $ky = x^2$ 是水平的。該處的法向反作用力（normal reaction）是鉛垂的，因此無法抵消向東的假想力 ma。沿金屬線方向的合力為 $ma \neq 0$，所以珠子不可能相對於火車保持靜止。 ■

階段 2　能量方程

> **想一想**　在火車的參考系中寫出牛頓定律，當中要包括法向反作用力 $\mathbf{N}$。然後與速度 $(\dot{x}, \dot{y})$ 取點積。為甚麼 $\mathbf{N}$ 會消失？

在火車的參考系中，

$$m(\ddot{x}, \ddot{y}) = \mathbf{N} + (ma, -mg).$$

金屬線是光滑的，所以 $\mathbf{N}$ 垂直於金屬線，因而也垂直於沿金屬線方向的速度 $(\dot{x}, \dot{y})$。與速度取點積（dot product），$\mathbf{N}$ 便會消去：

$$m\dot{x}\ddot{x} + m\dot{y}\ddot{y} = ma\dot{x} - mg\dot{y}, \qquad \text{所以} \qquad \dot{x}(\ddot{x} - a) + \dot{y}(\ddot{y} + g) = 0. \quad ■$$

左邊是 $\frac{1}{2}(\dot{x}^2 + \dot{y}^2) - ax + gy$ 的導數，因此這個量是常數。 ■

這就是火車參考系中的能量守恆（conservation of energy），其中每單位質量的勢能（potential energy）$-ax + gy$ 已把假想力計算在內。

階段 3　能升到多高？

> **想一想**　由於珠子從原點由靜止出發，該常數為 0。在最高點，珠子瞬時靜止。由此可得出甚麼方程？

靜止時，$\frac{1}{2}(\dot{x}^2 + \dot{y}^2) = 0$，所以 $ax = gy$。代入 $y = \frac{x^2}{k}$，得 $ax = \frac{gx^2}{k}$，所以（除了起點 $x = 0$ 之外）$x = \frac{ak}{g}$，而

$$b = \frac{a^2 k}{g^2}. \quad ■$$

珠子沿金屬線向東擺上至點 $\left(\frac{ak}{g}, \frac{a^2 k}{g^2}\right)$，然後返回原點，如此往復不息。

階段 4　最快的一點

> **想一想**　把速率的平方寫成只含 x 的函數。它在哪裏最大？

速率的平方是 $2(ax - gy) = 2\left(ax - \frac{gx^2}{k}\right)$，這是 x 的二次式，在 $x = \frac{ak}{2g}$ 處取得最大值，剛好在兩個轉向點（turning point）的正中間。該處速率的平方為 $2\left(\frac{a^2k}{2g} - \frac{a^2k}{4g}\right) = \frac{a^2k}{2g}$，所以最大速率是

$$a\sqrt{\frac{k}{2g}}. \qquad \blacksquare$$

想一想 求金屬線在 $x = \frac{ak}{2g}$ 處的斜率。它與有效重力 $(a, -g)$ 的方向有甚麼關係？

斜率是 $\frac{2x}{k} = \frac{a}{g}$。所以在這一點，金屬線垂直於有效重力：它是金屬線在傾斜重力場中的「最低」點，而珠子就圍繞它振動（圖 11.5.1）。

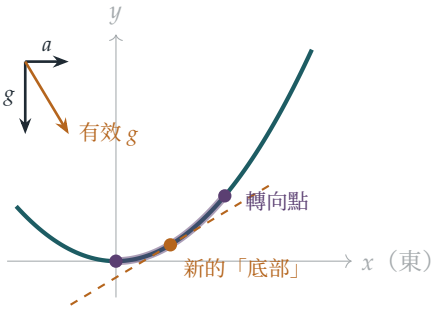


圖 11.5.1 火車參考系中的金屬線 $ky = x^2$（藍綠色，取 $k = 1$ 及 $a/g = 0.6$）。有效重力（琥珀色箭頭）向東傾斜，傾斜角的正切為 a/g。珠子在原點與 $x = ak/g$ 之間振動（紫紅色），圍繞着金屬線與有效重力垂直的那一點，該點就是新的最低點。虛線是傾斜重力場中的一條「水平」線。

回顧

- **在方便的參考系中解題，並加上假想力。**在火車的參考系中，金屬線是固定的，唯一的代價是多了一個恆力。
- **與速度取點積。**這樣可以消去約束力（constraint force），並得出能量方程。
- **詮釋答案。**轉向點和最快的一點都有幾何意義：有效重力傾斜了。

背後的數學

假想力

牛頓定律在慣性參考系(inertial frame)中成立，即沒有加速的參考系。在加速參考系中，只要加上假想力，牛頓定律便可繼續使用：對於加速度為 $\mathbf{A}$ 的參考系，要加上 $-m\mathbf{A}$；而在旋轉參考系中，則要加上離心力（centrifugal force）和科里奧利力（Coriolis force）。它們之所以稱為假想力，是因為沒有任何物體施加這些力；它們純粹是參考系造成的假象。但對身處其中的人來說，它們卻是千真萬確的。離心力使地球在兩極變得扁平；科里奧利力決定了颶風旋轉的方向（在北半球為逆時針），也令遠程炮彈偏離原來的方向。

等效原理

恆定的假想力 $-m\mathbf{A}$ 與質量成正比，這一點與重力完全相同。在一個密封的箱子內，沒有任何實驗能夠分辨箱子是靜止在行星上，還是在虛空中加速：這就是*等效原理*（equivalence principle）。愛因斯坦以它作為廣義相對論（general relativity）的出發點。既然光橫越一個加速中的箱子時看來會彎曲，光在重力場中也必定會彎曲；既然加速火箭的前端和後端的時鐘快慢不同，重力場中不同高度的時鐘也必定如此。這兩個預測都已得到證實，而每一顆 GPS 衛星都必須就第二個效應作出修正：單是這個效應，就令衛星上的時鐘每天比地面的時鐘快約 45 微秒。

愛因斯坦最快樂的想法

1907 年，愛因斯坦在伯爾尼的專利局工作，當時他正在撰寫一篇關於狹義相對論（special relativity）的綜述文章。據他後來憶述，就在這時，「我一生中最快樂的想法」出現了：一個從屋頂自由墮下的人，不會感覺到自己的重量。只要選取一個下墮中的參考系，重力就可以被「變換掉」。他再花了八年時間，並借助大量新的數學（由高斯（Gauss）和黎曼（Riemann）發展的彎曲空間微分幾何），才把這個想法變成 1915 年的廣義相對論。

思維技巧

在運動的參考系中尋找守恆量

在加速度恆定的參考系中，有一個恆定的假想力在起作用。它可以併入勢能之中，而能量在該參考系中守恆。

把運動方程與速度取點積以消去約束力，這個技巧在任何參考系中都適用。到了大學，它會成為一個普遍的原理：若一個系統的方程並不顯含時間，該系統便有一個守恆的能量。

延伸練習

1 **火車上的大炮。**試做 STEP I 2000，第 9 題（step.maths.org/questions/00-s1-q9）：在加速中的車廂內發射拋體（projectile），其中加速度的作用就如重力傾斜了一樣。

2 **逆風。**試做 STEP II 2006，第 11 題（step.maths.org/questions/06-s2-q11）：拋體迎着恆定的風飛行，風可以在它落地之前把它吹回你這邊。

3 **火車上的咖啡。**證明在加速中的火車上，一杯咖啡的液面是一個與水平面成 $\arctan(a/g)$ 角的傾斜平面。

4 **微小振動。**求珠子圍繞 $x = \frac{ak}{2g}$ 作微小振動的週期。（提示：在新的底部附近，金屬線近似於一個圓，其半徑等於金屬線在該處的曲率半徑（radius of curvature）。）

延伸閱讀

- **Daniel Kleppner & Robert Kolenkow, *An Introduction to Mechanics***（Cambridge University Press），當中講述非慣性參考系與假想力的一章。
- **Albert Einstein, *Relativity: The Special and the General Theory***（有多個版本）。愛因斯坦親自撰寫的普及讀物，當中包括加速箱子的討論。
- **Pedro Ferreira, *The Perfect Theory***（Little, Brown）。一部廣義相對論的傳記。

11.6 後旋

題目

STEP III 1997，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/97-s3-q10`

引子

用手指按住一粒波子的後半部向下一壓，波子便會一邊向後旋轉，一邊向前彈出。它從你面前滑開，慢下來，停住，然後彷彿改變了主意似的滾回來。桌球（snooker）球手打「縮桿」（screw 或 draw）時用的正是同一效應：擊打母球的下部，使它向後旋轉，碰到目標球時停住，再退回來。

本題要求的是：後旋（backspin）恰好要多大，波子才會完全停定；又要多大，波子才會以初速的七分之一返回。計算時要追蹤兩個量，即線速度（linear velocity）和角速度（angular velocity），兩者都由同一個摩擦力（frictional force）改變。這背後有一條漂亮的捷徑：兩者的某個組合完全不受摩擦力影響。

關於課程大綱的說明

轉動動力學（rotational dynamics）已於 2019 年從 STEP 考試範圍中刪去。這裏你只需要知道：質量為 m、半徑為 a 的實心球，繞直徑的轉動慣量（moment of inertia）為 $I = \frac{2}{5}ma^2$；而力矩（torque）τ 按 $I\dot{\omega} = \tau$ 改變其角速度，這就是 $m\dot{v} = F$ 的轉動版本。

動筆之前

想一想　波子以 v 向前運動，並以 ω 向後旋轉。波子的最低點相對桌面移動得多快？摩擦力朝哪個方向作用？

最低點因波子本身的運動而有速度 v，再加上因後旋而來的 $a\omega$（後旋會把最低點帶向前）。因此它以 $v + a\omega$ 向前滑動，而摩擦力向後作用，阻礙滑動。只有當 $v + a\omega = 0$ 時，波子才會純滾動（rolling without slipping），即只滾不滑。

逐步解題

階段 1　摩擦力的作用

想一想　設摩擦力為 F，方向向後。它對 v 有甚麼影響？它對球心的力矩為 Fa：這會增加還是減少後旋？

摩擦力使波子減速：$m\dot{v} = -F$，所以 $v = v_0 - \frac{F}{m}t$。由於它在波子底部向後作用，它同時產生一個阻礙後旋的力矩 Fa：$I\dot{\omega} = -Fa$，於是代入 $I = \frac{2}{5}ma^2$ 得

$$\omega = \omega_0 - \frac{5F}{2ma}t.$$

波子滑動期間，v 和 ω 都穩定地減少。

階段 2　完全停定

想一想　波子要完全停下來，v 和 ω 必須在同一刻變成零。兩者各在何時變成零？

$v = 0$ 於 $t = \frac{mv_0}{F}$，而 $\omega = 0$ 於 $t = \frac{2ma\omega_0}{5F}$。當 $v_0 = \frac{2}{5}a\omega_0$ 時兩者相等，即

$$\frac{v_0}{a\omega_0} = \frac{2}{5}. \qquad \blacksquare$$

在那一刻，波子既無線速度亦無角速度，而摩擦力（只在有滑動時才作用）隨即消失：波子就此停在原地。

階段 3　滾回來

想一想　若後旋較大，v 變成零時 ω 仍為正。最低點仍在向前滑動，所以摩擦力仍然向後作用。接下來會發生甚麼事？滑動何時停止？

摩擦力繼續把波子向後推，所以 v 變為負值，而 ω 繼續減少。當 $v + a\omega = 0$ 時滑動停止：

$$v_0 + a\omega_0 - \Big(\frac{F}{m} + \frac{5F}{2m}\Big)t = 0, \qquad t^* = \frac{2m(v_0 + a\omega_0)}{7F}.$$

在那一刻，

$$v = v_0 - \frac{2(v_0 + a\omega_0)}{7} = \frac{5v_0 - 2a\omega_0}{7}.$$

此後波子只滾不滑，摩擦力不再作用，v 保持不變。要它以速率 $\frac{v_0}{7}$ 滾回來，我們需要 $5v_0 - 2a\omega_0 = -v_0$，所以

$$\frac{v_0}{a\omega_0} = \frac{1}{3}. \qquad \blacksquare$$

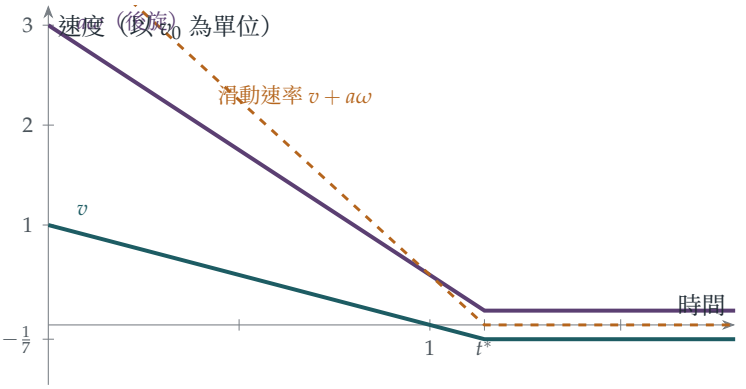


圖 11.6.1 $v_0/(a\omega_0) = \frac{1}{3}$ 的情況，時間以 mv_0/F 為單位。摩擦力以恆定的速率減少向前的速度 v（藍綠色）和後旋 $a\omega$（紫紅色），兩個速率之比為 2 : 5。波子在 $t = 1$ 時停止向前運動，但仍在滑動，所以被向後推，直至滑動速率（琥珀色）在 $t^* = \frac{8}{7}$ 時變成零。其後它以 $\frac{v_0}{7}$ 滾回來。

回顧

- **追蹤兩個量。**摩擦力同時改變 v 和 ω，兩者的變率成固定比例。
- **知道力何時消失。**滑動摩擦力只在有滑動時才作用；波子一旦滾動，它便停止作用。
- **滑動速率是關鍵變量。**它線性地減少，而它變成零的一刻，正是由滑動轉為滾動的一刻。

背後的數學

一個守恆量

要求最終速率，還有一個更快的方法。考慮波子對它與桌面接觸點的角動量（angular momentum）。摩擦力正正作用於該點，所以對該點沒有力矩，而法向反作用力（normal reaction）和重力都沿着通過該點的鉛垂線作用。因此波子滑動期間，這個角動量守恆（conserved）。它等於 mva（來自球心的運動）減去 $I\omega$（後旋以負值計算）：

$$L = mav - \tfrac{2}{5}ma^2\omega.$$

當波子以速率 V 滾動時，$a\omega = -V$，角動量為 $maV + \tfrac{2}{5}maV = \tfrac{7}{5}maV$。令它等於初始值，

$$\tfrac{7}{5}maV = mav_0 - \tfrac{2}{5}ma^2\omega_0, \qquad V = \frac{5v_0 - 2a\omega_0}{7},$$

答案相同，卻完全不必提及摩擦力的大小，甚至不必理會它是否恆定。最終速率單憑守恆便已確定；摩擦力只決定要花多少時間才達到這個速率。

桌球、保齡球與網球

同樣的分析可以解釋許多運動中的現象。桌球被擊中正中央時，起初滑動，然後開始滾動，最終速率為初速的 $\frac{5}{7}$（在公式中取 $\omega_0 = 0$）。十瓶保齡球帶側旋擲出時，會沿球道滑行，到開始滾動時便「拐彎」。帶上旋（topspin）的網球會急速下墜，不過這裏的主要效應來自空氣動力：旋轉的球會使周圍的空氣偏轉，這就是馬格努斯效應（Magnus effect）。

科里奧利與桌球

科里奧利（Gaspard-Gustave Coriolis）如今因旋轉參考系中的虛擬力（fictitious force）而為人所知（Section 11.5），他在 1835 年寫了一整本關於桌球數學的書：*Théorie mathématique des effets du jeu de billard*。他證明了不少結果，其中包括：帶旋轉擊出的球，在滑動期間沿拋物線運動，到滾動時才走直線；他亦精確地分析了本題的縮桿打法。研究這項運動的物理學家至今仍會引用此書。

思維技巧

追蹤由同一個力改變的兩個量

當一個力同時影響幾個量時，寫下它怎樣改變每一個量，並找出一個它完全不會改變的組合。

這裏摩擦力改變了 v 和 ω，卻沒有改變對接觸點的角動量。在大學裏，尋找這類組合是一個反覆出現的主題：選擇對哪一點取矩，使未知的力自動消去，正是選擇參考系的轉動版本。

延伸練習

1 **上旋。**若波子起初帶的是向前的旋轉 ω_0 而非後旋，求它最終的滾動速率。$v_0/(a\omega_0)$ 取甚麼值時它會加速？

2 **桌球。**一個球被擊中正中央，以 v_0 開始滑動而不帶旋轉。證明它最終以 $\frac{5}{7}v_0$ 滾動，並求動能（kinetic energy）損失的比例。（滾動時的動能為 $\frac{1}{2}mV^2 + \frac{1}{2}I\omega^2$。）

3 **圓環。**對一個在自身所在平面內滾動的薄圓環重做本題，其 $I = ma^2$。現在 $v_0/(a\omega_0)$ 為何值時，它會以 $\frac{v_0}{7}$ 返回？

4 **縮桿打法。**桌球球桿必須擊中球的甚麼位置（在球心以上多高，以 a 表示），球才會從一開始便只滾不滑？（提示：考慮衝量（impulse）作用期間對接觸點的角動量。）

延伸閱讀

- **David Morin, *Introduction to Classical Mechanics***（Cambridge University Press），第 8 章，講述角動量及帶滑動的滾動。
- **Daniel Kleppner & Robert Kolenkow, *An Introduction to Mechanics***（Cambridge University Press），關於剛體運動的各章。
- **David Alciatore, *The Illustrated Principles of Pool and Billiards***（Sterling）。這項運動的物理學，附有大量圖解。

11.7 桿中的辛普森法則

題目

STEP II 1989，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/89-s2-q13`

引子

一根均勻棒（uniform rod）受到推動和旋轉時，會有甚麼反應？原來就力學上的一切用途而言，一根棒的表現與三個質點（point mass）完全相同：兩端各佔其質量的六分之一，中點佔三分之二。無論你對這根棒做甚麼（推它，或讓它繞任何軸旋轉），都無法把它與這三個質點區分開來。

$\frac{1}{6}$、$\frac{2}{3}$、$\frac{1}{6}$ 這幾個數也許似曾相識。它們正是*辛普森法則*（Simpson's rule）$\int_{-a}^{a} f \approx 2a\left(\frac{1}{6}f(-a) + \frac{2}{3}f(0) + \frac{1}{6}f(a)\right)$ 中的權（weight），這是你可能在高級程度（A level）課程中學過的數值積分（numerical integration）方法。這並非巧合：動力等效（dynamically equivalent）系統與數值積分法則，解決的其實是同一個問題。

關於課程範圍的說明

轉動慣量（moment of inertia）已於 2019 年從 STEP 考試範圍中刪去。一個物體對某軸的轉動慣量是 $\sum m_i d_i^2$，即每份質量乘以它與該軸距離的平方之總和（對連續物體而言則是一個積分）。質量為 m、長度為 $2a$ 的均勻棒，對通過其中心的垂直軸的轉動慣量為 $\frac{1}{3}ma^2$。*平行軸定理*（parallel axis theorem）指出：對任何軸的轉動慣量，等於對通過質心（centre of mass）的平行軸的轉動慣量加上 md^2，其中 d 是兩軸之間的距離。

動筆之前

想一想　兩個系統的質量相同、質心相同，而且對通過質心的每一條軸的轉動慣量也相同。為甚麼平行軸定理隨即保證，它們對其他任何軸的轉動慣量也相同？

對任何軸而言，兩個系統的轉動慣量都等於（對通過 O 的平行軸的轉動慣量）$+md^2$。根據假設，第一項相同，而兩個系統的 m 和 d 也相同。 ■

逐步解題

階段 1　棒

想一想　在兩端各放質量 p，在中點放質量 q。由對稱性可知質心位置正確。總質量和對通過中心的垂直軸的轉動慣量，分別要求甚麼條件？

我們需要 $2p+q=m$，而對通過中點的垂直軸的轉動慣量必須是 $\frac{1}{3}ma^2$：$2pa^2=\frac{1}{3}ma^2$。所以 $p=\frac{1}{6}m$，$q=\frac{2}{3}m$。

想一想　這是否對通過中心的每一條軸都成立，而不只是垂直的那一條？

考慮一條通過中心、與棒成 ψ 角的軸：棒上與中心相距 r 的點，與該軸的距離為 $r\sin\psi$。因此每一個轉動慣量（不論是棒的還是質點的）都是垂直軸情況下數值的 $\sin^2\psi$ 倍，所以兩者對所有 ψ 都相等。（不在包含棒的平面上的軸也是同樣道理：只有各點與軸的距離才有關係，而對於一條通過 O 的直線上的點，這個距離與 r 成正比。）所以這根棒與質量 $\frac{1}{6}m$、$\frac{2}{3}m$、$\frac{1}{6}m$ 動力等效。 ■

階段 2　長方形

想一想　把薄片（lamina）切成與其中一邊平行的幼條，再把每一幼條換成它的三個等效質點。你會得到甚麼？接下來又可以怎樣做？

每一幼條都是一根棒，等效於把它質量的 $\frac{1}{6}$、$\frac{2}{3}$、$\frac{1}{6}$ 分別放在左邊、中線和右邊。把所有幼條加起來，薄片便等效於三根均勻棒（分別沿左邊、中間和右邊），質量為 $\frac{1}{6}M$、$\frac{2}{3}M$、$\frac{1}{6}M$。現在再把這三根棒各自換成比例相同的三個質點。每一點的質量是兩個分數之積（圖 11.7.1）：

$$\text{四角：}\frac{1}{6}\cdot\frac{1}{6}=\frac{1}{36},\qquad \text{各邊中點：}\frac{1}{6}\cdot\frac{2}{3}=\frac{1}{9},\qquad \text{中心：}\frac{2}{3}\cdot\frac{2}{3}=\frac{4}{9},$$

全部再乘以 M。 ■

（嚴格來說，每一次替換都必須保持對通過整塊薄片中心的每一條軸的轉動慣量不變，而事實確是如此：每一次替換都保持了該部分的質量、質心，以及對通過該部分質心的各軸的轉動慣量，其餘的就交給平行軸定理。）

階段 3 正方形對其對角線

對於邊長為 $2a$ 的正方形，對角線通過兩個角和中心，這些點沒有任何貢獻。另外兩個角與對角線的距離為 $\sqrt{2}a$，而四個邊的中點與它的距離為 $\frac{a}{\sqrt{2}}$。所以

$$I = 2\cdot\tfrac{1}{36}M\cdot 2a^2 + 4\cdot\tfrac{1}{9}M\cdot\tfrac{a^2}{2} = \tfrac{1}{9}Ma^2 + \tfrac{2}{9}Ma^2 = \tfrac{1}{3}Ma^2. \qquad ■$$

這與對通過中心、平行於一邊的軸的轉動慣量相同。（這並非偶然：對正方形而言，在其平面上通過中心的所有軸都給出相同的答案，這可由練習中的一個對稱性論證得出。）

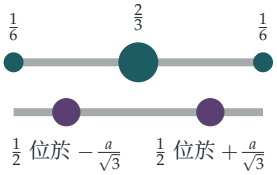
階段 4 非均勻棒

? **想一想** 設棒的長度為 L，在距離其中一端 x 處的密度為 $\frac{2mx}{L^2}$。在 $x = 0$、$\frac{L}{2}$、L 處分別放質量 p_0、p_1、p_2。令總質量、質心和對該端的轉動慣量一一吻合，便得到三條方程。它們是甚麼？

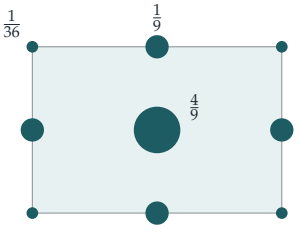
這根棒的總質量為 m，對較輕一端的一階矩（first moment）為 $\int_0^L x\cdot\frac{2mx}{L^2}\,\mathrm{d}x = \frac{2}{3}mL$，二階矩（second moment）為 $\int_0^L x^2\cdot\frac{2mx}{L^2}\,\mathrm{d}x = \frac{1}{2}mL^2$。令這些量吻合：

$$p_0 + p_1 + p_2 = m, \qquad \tfrac{1}{2}Lp_1 + Lp_2 = \tfrac{2}{3}mL, \qquad \tfrac{1}{4}L^2p_1 + L^2p_2 = \tfrac{1}{2}mL^2.$$

把第二式（除以 L）減去第三式（除以 L^2），得 $\frac{1}{4}p_1 = \frac{1}{6}m$，所以 $p_1 = \frac{2}{3}m$，繼而 $p_2 = \frac{1}{3}m$，$p_0 = 0$。只需要兩個質點：$\frac{2}{3}m$ 位於中點，$\frac{1}{3}m$ 位於較重的一端。 ■



兩根等效的棒



薄片：$\frac{1}{6}, \frac{2}{3}, \frac{1}{6}$ 的乘積

圖 11.7.1 左上：一根均勻棒與位於兩端和中心的質量 $\frac{1}{6}$、$\frac{2}{3}$、$\frac{1}{6}$（以佔其質量的分數表示）動力等效，這些正是辛普森法則的權。左下：它也等效於兩份各佔一半的質量，位於與中心相距 $\frac{a}{\sqrt{3}}$ 之處，即兩點高斯積分的節點。右：一塊長方形薄片等效於九個質量，圖中圓點的面積與質量成正比。

回顧

- **令各階矩吻合。**動力等效的意思，就是令質量分佈的零階矩、一階矩和二階矩一一吻合。
- **逐個維度建構。**長方形是由棒組成的棒，所以它的權是乘積。
- **數一數方程和未知數的數目。**要吻合的矩有三個，未知的質量也有三個；而有時解中恰好會出現一個零。

背後的數學

辛普森法則

令質量分佈的各階矩吻合，與設計一條數值積分法則是同一個問題。一條法則 $\int_{-a}^{a} f(x)\,\mathrm{d}x \;\approx\; w_0 f(-a) + w_1 f(0) + w_2 f(a)$ 若對 $f = 1, x, x^2$ 精確成立，就必須滿足

$$w_0 + w_1 + w_2 = 2a, \qquad -a w_0 + a w_2 = 0, \qquad a^2 w_0 + a^2 w_2 = \tfrac{2}{3}a^3,$$

這正是棒所要滿足的條件（其中 $f = 1, x, x^2$ 分別扮演質量、一階矩和轉動慣量的角色）。其解為 $w = 2a(\tfrac{1}{6}, \tfrac{2}{3}, \tfrac{1}{6})$：這就是辛普森法則。對稱性還帶來額外的收穫：這條法則對 $f = x^3$ 也精確成立，因為兩邊都等於零，所以辛普森法則能精確地積分三次多項式，而它對光滑函數的誤差與四階導數成正比。因此，一根棒與它的三個質點不僅在質量、質心和轉動慣量上吻合，連三階矩也吻合。

高斯求積法與慣性張量

如果位置和質量都可以自由選擇，兩個質點便已足夠：在 $\pm\frac{a}{\sqrt{3}}$ 兩點各放一半質量。驗算：$2 \cdot \frac{1}{2}m \cdot \frac{a^2}{3} = \frac{1}{3}ma^2$。這就是兩點高斯－勒讓德法則（Gauss–Legendre rule）$\int_{-a}^{a} f \approx a\big(f(-\frac{a}{\sqrt{3}}) + f(\frac{a}{\sqrt{3}})\big)$，它同樣對三次多項式精確成立。高斯（Carl Friedrich Gauss）在 1814 年證明，只要巧妙地選取 n 個點，便能精確地積分次數為 $2n-1$ 的多項式；高斯求積法（Gaussian quadrature）如今已是數值分析（numerical analysis）的標準方法。

在力學中，一個物體對通過其中心的所有軸的轉動慣量，都記錄在一個 3×3 對稱矩陣之中，稱為慣性張量（inertia tensor）。兩個物體若質量、質心和慣性張量都相同，對所有力的反應便完全一樣，這就是為甚麼工程師在模擬中可以用幾個質點取代一個複雜的部件。慣性張量的特徵向量（eigenvector）就是主軸（principal axes），即物體可以繞之穩定旋轉的軸。

思維技巧

把系統換成等效的系統

如果兩個系統在問題所涉及的每一個量上都一致，就把複雜的那個換成簡單的那個。

這裏只有三個矩有關係，所以一根連續的棒可以換成三個質點。在大學裏，這個想法會以多種面貌出現：高斯求積法、電磁學和引力理論中的多極展開（multipole expansion）（在遠處，電荷或質量分佈只有頭幾個矩有影響），以及統計學中的矩估計法（method of moments）。

延伸練習

1 **辛普森法則。**用只有一個區間的辛普森法則估算 $\int_0^1 \frac{\mathrm{d}x}{1+x^2}$，並與 $\frac{\pi}{4}$ 比較。驗證這條法則對 $\int_{-1}^1 x^3\,\mathrm{d}x$ 和 $\int_{-1}^1 x^2\,\mathrm{d}x$ 精確成立，但對 $\int_{-1}^1 x^4\,\mathrm{d}x$ 則不然。

2 **所有軸都一樣。**證明正方形薄片對在其平面上通過中心的任何軸的轉動慣量都是 $\frac{1}{3}Ma^2$。（提示：九個等效質點在旋轉 90° 之下是對稱的；這對一個關於 $\cos\psi$ 和 $\sin\psi$ 的二次式說明了甚麼？）

3 **三角形。**證明一塊質量為 M 的均勻三角形薄片，與位於其三邊中點、質量各為 $\frac{1}{3}M$ 的三個質點動力等效。

4 **實心立方體。**為一個均勻實心立方體找出一個由 27 個質點組成的動力等效系統，並用它求出邊長為 $2a$ 的立方體對其一條棱的轉動慣量。

延伸閱讀

- **Daniel Kleppner & Robert Kolenkow, *An Introduction to Mechanics***（Cambridge University Press），關於剛體和轉動慣量的一章。
- **Endre Süli & David Mayers, *An Introduction to Numerical Analysis***（Cambridge University Press），關於數值積分的各章，包括辛普森法則和高斯求積法。
- **Philip Davis & Philip Rabinowitz, *Methods of Numerical Integration***（Dover）。經典參考書。

11.8 傾斜的平衡

題目

STEP III 2025，第 9 題

繼續閱讀之前，請先看題目（原題為英文）。題目載於 OCR 提供下載的 2025 年 STEP 試卷（zip 檔案）內的 STEP III 試卷：

`www.ocr.org.uk/Images/750486-2025-past-papers.zip`

引子

把一個欖球放在桌上，它會側躺着停下來；把它豎起來，只要十分小心，它也可以這樣平衡，但輕輕一碰就會倒下。每一個放在桌上的剛體（rigid body）都有一組平衡位置（equilibrium positions），其中有些是穩定（stable）的，有些是不穩定（unstable）的，而哪些位置存在，只取決於它的形狀和質心（centre of mass）。一個物體可以有多少個平衡位置？最少又可以有多少個？

這道題構造了一個立體，形狀像陀螺，上半部是圓頂，下半部是尖底，並證明它除了可以正立或倒立之外，還可以在三個不同的傾斜角度上靜止。關鍵思路是一個簡單的幾何條件：物體能以其表面上某一點着地而靜止，當且僅當表面在該點的法線（normal）通過質心。這背後有一項近年的發現：一個稱為 Gömböc 的凸（convex）形體，它只有一個穩定位置和一個不穩定位置。

動筆之前

想一想 一個立體放在桌上，只以其曲面上的一點 X 與桌面接觸。作用於它的力是它的重量（作用於質心 G）和在 X 的反作用力（reaction）。這兩個力何時平衡？

桌面在 X 與曲面相切，而反作用力作用於 X，並垂直於桌面，即沿着曲面的法線。重量則沿鉛垂方向作用，並通過 G。兩個力只有在沿同一直線作用時才能平衡，所以通過 X 的鉛垂線，也就是在 X 的法線，必須通過 G。因此：*物體能以 X 着地而靜止，當且僅當在 X 的法線通過 G。*

逐步解題

階段 1　質心

想一想　把立體切成薄薄的圓柱殼（cylindrical shells），半徑為 x，厚度為 δx，高度為 y。一個圓柱殼的體積是多少？它的質心在哪裏？

由於 $y = \frac{r^n - x^n}{r^{n-1}}$，一個圓柱殼的體積是 $2\pi x y\,\delta x$，而它的質心在高度 $\frac{y}{2}$ 處。所以

$$V = \int_0^r 2\pi x y\,\mathrm{d}x = \frac{2\pi}{r^{n-1}}\Big(\frac{r^{n+2}}{2} - \frac{r^{n+2}}{n+2}\Big) = \frac{n\pi r^3}{n+2},$$

而對底面的矩（moment）是

$$\begin{aligned}\int_0^r 2\pi x \cdot \frac{y^2}{2}\,\mathrm{d}x &= \frac{\pi}{r^{2n-2}}\int_0^r x(r^n - x^n)^2\,\mathrm{d}x \\ &= \pi r^4\Big(\frac{1}{2} - \frac{2}{n+2} + \frac{1}{2n+2}\Big) = \frac{n^2\pi r^4}{2(n+1)(n+2)}.\end{aligned}$$

（括號可以化簡，是因為 $(n+1)(n+2) - 4(n+1) + (n+2) = n^2$。）兩者相除，得

$$\bar{y} = \frac{nr}{2(n+1)}.$$ ■

當 $n = 2$（拋物面，paraboloid）時，這是 $\frac{r}{3}$，正是大家熟悉的結果；當 $n \to \infty$ 時，立體變成圓柱，而 $\bar{y} \to \frac{r}{2}$。

階段 2　法線與軸相交之處

曲線在 $x = rp$ 的斜率（gradient）是 $-\frac{nx^{n-1}}{r^{n-1}} = -np^{n-1}$，所以法線的斜率是 $\frac{1}{np^{n-1}}$。沿法線由 $\big(rp, r(1-p^n)\big)$ 走到 y-軸，x 減少了 rp，而 y 減少了 $\frac{rp}{np^{n-1}} = \frac{r}{np^{n-2}}$。所以

$$Y = r\Big(1 - p^n - \frac{1}{np^{n-2}}\Big).$$ ■

想一想　當 $n = 4$ 時，$Y = r\big(1 - p^4 - \frac{1}{4p^2}\big)$。它在哪裏取最大值？當 $p \to 0$ 時，以及在 $p = 1$ 時，情況如何？

$\frac{\mathrm{d}Y}{\mathrm{d}p} = r\big(-4p^3 + \frac{1}{2p^3}\big) = 0$ 成立於 $p^6 = \frac{1}{8}$，即 $p^2 = \frac{1}{2}$。這時 $Y = r\big(1 - \frac{1}{4} - \frac{1}{2}\big) = \frac{1}{4}r$。 ■

當 $p \to 0$ 時，$Y \to -\infty$（圓頂在頂部附近非常平坦，它的法線幾乎是鉛垂的，要在很低的地方才與軸相交）；在 $p = 1$ 時，$Y = -\frac{1}{4}r$。所以當 p 由 0 增至 1 時，Y 由 $-\infty$ 上升至 $\frac{1}{4}r$，再回落至 $-\frac{1}{4}r$。

階段 3　陀螺

想一想 第 (iii) 部分的立體，是一個圓頂 ($n = 4$) 疊在一個倒轉的拋物面 ($n = 2$) 之上。它的質心在哪裏？

由步驟 1，圓頂的體積是 $\frac{2\pi r^3}{3}$，質心在 $\frac{2r}{5}$；拋物面的體積是 $\frac{\pi r^3}{2}$，而由於它是倒轉的，質心在 $-\frac{r}{3}$。所以

$$\bar{y} = \frac{\frac{2}{3}\cdot\frac{2}{5} - \frac{1}{2}\cdot\frac{1}{3}}{\frac{2}{3}+\frac{1}{2}}\, r = \frac{\frac{4}{15}-\frac{1}{6}}{\frac{7}{6}}\, r = \frac{3r}{35}.$$

G 位於 $\left(0, \frac{3r}{35}\right)$，比最闊之處略高。

階段 4 數算平衡位置

想一想 傾斜的平衡位置需要有一點（不是 A 或 B），其法線通過 G。在曲面的每一部分上，有多少個點的法線與軸相交於高度 $\frac{3r}{35}$？

圓頂。它的法線與軸相交的高度 Y 由 $-\infty$ 上升至 $\frac{1}{4}r$，再下降至 $-\frac{1}{4}r$。由於 $-\frac{1}{4}r < \frac{3r}{35} < \frac{1}{4}r$，根據介值定理（intermediate value theorem），$\frac{3r}{35}$ 這個值在上升的一段取得一次，在下降的一段也取得一次：共兩點。

尖底。由第 (ii) 部分取 $n = 2$，再對 x-軸作反射，可知在參數為 p 的點上，法線與軸相交於 $-r\left(1 - p^2 - \frac{1}{2}\right) = r\left(p^2 - \frac{1}{2}\right)$，這個值由 $-\frac{1}{2}r$ 遞增至 $\frac{1}{2}r$。它恰好有一次等於 $\frac{3r}{35}$，就是在 $p^2 = \frac{41}{70}$ 時。

這三點各給出一個平衡位置，而三條法線與軸所成的角各不相同，所以 $\overrightarrow{AB}$ 與鉛垂線成三個不同的非零角（圖 11.8.1）。 ■

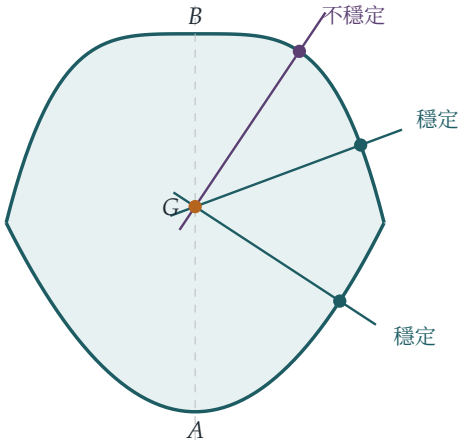


圖 11.8.1 第 (iii) 部分的立體的截面，以及它的質心 G（琥珀色）。在三個標記的點上，曲面的法線都通過 G，所以立體可以用該點着地而靜止。其中兩個平衡位置是穩定的（曲面在該處局部地離 G 最近），一個是不穩定的（局部地最遠）。

回顧

- **把物理條件化為幾何。**平衡變成了「法線通過 G」。
- **不用求解也能證明存在。**我們從沒有求出那些角。介值定理，加上函數 $Y(p)$ 的形狀，已說明了它們有多少個。
- **重用一般結果。**第 (ii) 部分對一般的 n 所得的結果，一併處理了第 (iii) 部分中曲面的兩個部分。

背後的數學

平衡位置即臨界點

當立體以點 X 着地而靜止時，G 離桌面的高度就是由 G 到在 X 的切平面（tangent plane）的距離；對凸體來說，當 GX 沿着法線時，這個距離就是 $|GX|$。一般而言，平衡位置就是由 G 到曲面的距離的*臨界點*（*critical points*）：使 $|GX|$ 取駐值的點 X，正好就是法線通過 G 的點。穩定的平衡位置是這個距離的局部極小值（G 盡可能低），而不穩定的則是局部極大值或鞍點（saddle points）。

沿着截面由 A 走到 B，與 G 的距離依次是：在 A 較大（極大值：以尖端平衡，不穩定），在尖底上的傾斜點較小（極小值，穩定），在邊緣較大，在圓頂上較低的傾斜點較小（穩定），在較高的傾斜點較大（不穩定），而在 B 較小（穩定：以圓頂的頂部着地）。正如圖中所示，極小值和極大值必定交替出現。（事實上還有一個題目沒有計算在內的平衡位置：立體可以不穩定地平衡在它尖銳的邊緣上，曲面在那裏沒有唯一的法線。）

Gömböc

一個凸體最少可以有多少個平衡位置？任何物體都至少有一個穩定位置和一個不穩定位置：就是曲面上離 G 最近和最遠的點。密度均勻的二維凸形狀總是至少有四個（距離的兩個極小值和兩個極大值，交替出現）。1995 年，阿諾爾德（Vladimir Arnold）提出一個問題：密度均勻的三維凸體，能否只有兩個平衡位置，即一個穩定、一個不穩定？2006 年，Gábor Domokos 和 Péter Várkonyi 構造出這樣的一個形體，就是 *Gömböc*：無論以甚麼姿勢把它放下，它都會滾動搖晃，直至到達它唯一的穩定位置為止。Domokos 後來留意到，有些陸龜的殼很接近這個形狀，這有助牠們在被翻轉後自行翻身。

阿基米德與浮體

對物體穩定性的研究可以追溯至阿基米德（Archimedes）。他在《論浮體》（*On Floating Bodies*）中，根據形狀和密度，分析了一個浮着的拋物面（船身的一個模型）何時能穩

定地直立，何時會翻側。他只用幾何便得出答案，當中涉及的正是本節所用的那種質心和法線的計算。造船工程師至今仍在運用他的想法，形式就是穩心（*metacentre*）：如果通過浮心（centre of buoyancy）的鉛垂線與船的軸線的交點位於質心之上，船便是穩定的。

思維技巧

勢能地形

平衡位置是勢能（*potential energy*）的駐點（*stationary points*），而穩定的平衡位置是它的極小值。與其把駐點逐一解出來，不如數算它們的數目。

這裏的勢能與 G 的高度成正比，也就是由 G 到曲面的距離。在大學裏，這個觀點貫穿力學、化學（分子會落入能量的極小值）和拓撲學；在拓撲學中，莫爾斯理論（*Morse theory*）說明，曲面上任何函數的極小值、極大值和鞍點的數目，都受曲面的形狀所限制。

延伸練習

1 **不倒翁玩具。**試做 STEP II 1996，第 9 題（`step.maths.org/questions/96-s2-q9`）：一個置於半球上的圓錐，何時會自行回正？（提示：相對於半球的球心，G 必須在哪裏？）

2 **球體。**對於均勻實心球體，每一條法線都通過球心。這說明它的平衡位置有甚麼特點？如果球體的質心稍為偏離球心，又會怎樣？

3 **橢圓。**證明一塊橢圓形圓盤在其自身所在的平面內沿一條直線滾動時，恰好有四個平衡位置，並指出哪些是穩定的。

4 **兩個拋物面。**把第 (iii) 部分的圓頂換成拋物面 $ry = r^2 - x^2$，使立體對 x-軸對稱。證明 G 位於原點，求出傾斜的平衡位置，並解釋為何它們成對出現，且每對的傾斜角相等。

延伸閱讀

► **Péter Várkonyi & Gábor Domokos,「Static equilibria of rigid bodies: dice, pebbles, and the Poincaré–Hopf theorem」**, *Journal of Nonlinear Science* 16 (2006)。提出 Gömböc 的

論文。

- **Daniel Kleppner & Robert Kolenkow, *An Introduction to Mechanics***（Cambridge University Press），論平衡與穩定性。
- **Chris Rorres,「Completing Book II of Archimedes's *On Floating Bodies*」**, *Mathematical Intelligencer* 26 (2004)。阿基米德的穩定性分析，配以現代圖解。

11.9 安全拋物線

題目

STEP II 2014，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/14-s2-q10`

引子

一座噴泉以相同的速率向四方八面噴水。水柱形成一個閃閃生輝的圓頂，外緣輪廓分明：在裏面，每一點都有水經過；在外面，水一滴也到不了。這條邊界是一條拋物線（parabola），稱為*安全拋物線*（parabola of safety），因為炮口速率固定的大炮，無論以甚麼角度發射，都打不中它外面的目標。

令人意外的是，這條邊界本身並不是其中一條軌跡。它是整族軌跡的*包絡線*（envelope），即一條與每條軌跡都相切一次的曲線，就像 Section 8.1 中直線 $y = 0$ 與那族拋物線相切一樣。這道題目只需用到二次式的判別式（discriminant），便能把它求出來。

動筆之前

? **想一想** 用 t 表示 x 和 y，然後消去 t。怎樣用 $\lambda = \tan\theta$ 表示 $\sec^2\theta$？

由 $x = ut\cos\theta$ 和 $y = ut\sin\theta - \frac{1}{2}gt^2$，代入 $t = \frac{x}{u\cos\theta}$：

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

其中用到 $\sec^2\theta = 1 + \tan^2\theta$。

逐步解題

階段 1　固定 x，改變角度

想一想　對於固定的 $x > 0$，把 y 看成 λ 的函數。它是哪一類函數？當 λ 取遍所有實數時，它會取哪些值？

對於固定的 x，y 是 λ 的二次函數，首項係數 $-\frac{gx^2}{2u^2}$ 為負：它是以 λ 為變量、開口向下的拋物線。它有極大值，而且當 $\lambda \to \pm\infty$ 時，$y \to -\infty$。由於它是連續的，所以它會取遍不超過極大值的每一個值。極大值出現在 $\frac{\mathrm{d}y}{\mathrm{d}\lambda} = x - \frac{gx^2}{u^2}\lambda = 0$ 之處，即 $\lambda = \frac{u^2}{gx}$，此時

$$Y = \frac{u^2}{gx}\,x - \frac{gx^2}{2u^2}\left(1 + \frac{u^4}{g^2x^2}\right) = \frac{u^2}{2g} - \frac{gx^2}{2u^2}. \qquad \blacksquare$$

（等價的說法是：點 (x, y) 可以到達，當且僅當二次方程 $\frac{gx^2}{2u^2}\lambda^2 - x\lambda + \frac{gx^2}{2u^2} + y = 0$ 有實根 λ，而判別式的條件給出 $y \le Y$。）

階段 2　區域

曲線 $Y = \frac{u^2}{2g} - \frac{gx^2}{2u^2}$ 是一條開口向下的拋物線，頂點位於 $\left(0, \frac{u^2}{2g}\right)$，即垂直向上發射所達到的高度；它與地面相交於 $x = \pm\frac{u^2}{g}$，即最大射程。在它上面或下方（並在地面以上）的每一點都可以到達，而在它上方的點則全都到不了（圖 11.9.1）。

想一想　經過嚴格位於曲線下方的一點，有多少條軌跡？經過曲線上的一點呢？

曲線下方的一點對應一個有兩個實根的 λ 二次方程：兩個拋射角，一高一低。曲線上的一點對應重根：恰好一條軌跡，並且在該點與曲線相切。

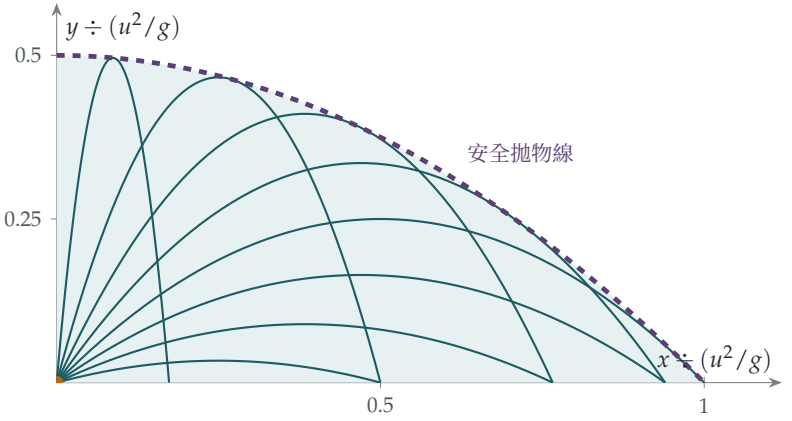


圖 11.9.1　以固定速率 u 及角度 $15°, 25°, \ldots, 85°$ 發射的軌跡（藍綠色）。每條軌跡都與虛線所示的安全拋物線 $y = \frac{u^2}{2g} - \frac{gx^2}{2u^2}$ 恰好相切一次，而它下方的區域（陰影部分）正好是所有可到達的點所組成的集合。原點 O（琥珀色）是安全拋物線的焦點（focus）。

階段 3 最遠的點

? **想一想** 距離 O 最遠的可到達點（在地面上或地面以上）位於邊界上。把距離的平方 x^2+Y^2 寫成 x^2 的函數。它是遞增還是遞減？

設 $X=x^2$。則

$$x^2+Y^2=X+\left(\frac{u^2}{2g}-\frac{gX}{2u^2}\right)^2,\qquad \frac{\mathrm{d}}{\mathrm{d}X}(x^2+Y^2)=1-\frac{g}{u^2}\left(\frac{u^2}{2g}-\frac{gX}{2u^2}\right)=\frac{1}{2}+\frac{g^2X}{2u^4}>0.$$

因此，沿着邊界向外移動時，距離不斷增加，而最遠的點就是邊界與地面相交之處，即 $x=\frac{u^2}{g}$。最遠距離為

$$\frac{u^2}{g},$$

由平常以 45° 發射的最大射程一擊達到。 ■

回顧

- **把問題反過來問。**不問某條給定的軌跡會去到哪裏，而問哪些軌跡會經過某個給定的點。
- **以參數為未知數的二次方程。**對於每一點，$\lambda=\tan\theta$ 所滿足的方程都是二次方程，而它的判別式決定該點能否到達。
- **兩個、一個或沒有。**區域內的點有兩個角度可以到達，邊界上的點有一個，區域外的點則一個也沒有。

背後的數學

包絡線

單參數曲線族（one-parameter family of curves）$\mathrm{F}(x,y,\lambda)=0$ 往往有一條*包絡線*：一條與族中每個成員都相切的曲線。求法是聯立解 $\mathrm{F}=0$ 和 $\frac{\partial \mathrm{F}}{\partial\lambda}=0$，這正是我們對 λ 求 y 的極大值時所做的事。包絡線在科學中隨處可見。一杯咖啡表面上那條明亮的光曲線，就是反射光線的包絡線，稱為*焦散線*（caustic）。超音速飛機的音爆，由它所發出聲波的包絡面傳送，那就是*馬赫錐*（Mach cone）。而在 Section 8.1 中，一個微分方程（differential equation）的一族解的包絡線，本身也是一個解，即所謂的「奇異」解。

焦點

安全拋物線有一個奇妙的性質：它的焦點就是發射點 O。（拋物線 $y=c-\frac{x^2}{4c}$ 的焦點位於原點；這裏 $c=\frac{u^2}{2g}$。）每條軌跡也都是拋物線，而所有軌跡的焦點都落在以 O 為圓心、

半徑為 $\frac{u^2}{2g}$ 的圓上。這可以用能量來解釋：拋射體在高度 y 時的速率是 $\sqrt{u^2 - 2gy}$，在準線（directrix）的高度 $\frac{u^2}{2g}$ 處為零；而以同一條直線為準線並且經過 O 的拋物線，其焦點都在那個圓上。

托里拆利的拋物線

伽利略（Galileo）在 1609 年前後證明了拋射體沿拋物線運動，並且在水平地面上，最大射程出現在 45°。他的學生托里拆利（Evangelista Torricelli）以發明氣壓計而較為人熟悉，他在 1644 年的著作 *De motu gravium* 中發現了安全拋物線。對炮手來說，它回答了一個實際問題：究竟哪些目標在射程之內？托里拆利還證明了，經過給定一點的軌跡成對出現，兩者的角度關於該點的方向與鉛垂線之間的角平分線對稱，這就是本題中的高射和低射。

思維技巧

曲線族的包絡線

要找出一族曲線所覆蓋區域的邊界，可固定一點，然後問：參數取哪些值時，會有曲線經過這一點？邊界就是解的數目發生變化的地方。

同一個想法，即隨着點的移動來數算解的數目，是光學中焦散線研究的基礎，也是動力系統（dynamical systems）中分歧（bifurcation）理論的基礎，更是突變理論（catastrophe theory）的基礎；突變理論把方程的解的數目可以怎樣變化加以分類。

延伸練習

1 **爆開的煙花。**試做 STEP II 2023，第 10 題（載於 OCR 提供下載的 2023 年試卷，`www.ocr.org.uk/Images/707102-2023-past-papers.zip`）。這道題把爆炸產生的粒子看成一個球面，其球心作自由下落，並找出一顆沿着安全拋物面（paraboloid of safety）運動的粒子。

2 **向斜坡上發射。**求速率為 u 的拋射體沿傾角為 β 的斜面向上發射時的最大射程。（提示：直線 $y = x\tan\beta$ 與安全拋物線在哪裏相交？）證明最佳角度平分斜坡與鉛垂線之間的夾角。

3 **兩個角度。**證明能到達點 P 的兩個角度，關於 OP 與鉛垂線夾角的平分線對稱。

4 **焦點。**證明拋射角為 θ 的軌跡，其焦點位於 $\left(\frac{u^2}{2g}\sin 2\theta, -\frac{u^2}{2g}\cos 2\theta\right)$，並由此證明這些焦點都落在以 O 為圓心的一個圓上。

延伸閱讀

- **David Morin, *Introduction to Classical Mechanics***（Cambridge University Press），第 3 章，當中討論了軌跡的包絡線。
- **J. W. Bruce, P. J. Giblin & C. G. Gibson, "On caustics of plane curves"**, *American Mathematical Monthly* 88 (1981)。幾何學中的包絡線與焦散線。
- **E. H. Lockwood, *A Book of Curves***（Cambridge University Press）。包絡線、焦散線及許多經典曲線，並附作圖方法。

11.10 會朝你飛回來的球

題目

STEP II 2019，第 9 題

繼續閱讀之前，請先看題目（原題為英文）。題目載於 OCR 提供下載的 2019 年 STEP 試卷（zip 檔案）內的 STEP II 試卷：

`www.ocr.org.uk/Images/696590-2019-past-papers.zip`

引子

把球向前拋出，它在落地之前會一直離你而去。這是理所當然的。把球垂直向上拋，它卻會回到你手上。在這兩種情況之間，必定有些事情起了變化：必定存在某些角度，令球先遠離你，然後向你*靠近*，到落地之前又再遠離。本題要找出的正是那個確切的分界線：$\sin\alpha = \frac{2\sqrt{2}}{3}$，即大約 70.5°。

這個結果是數學中一種常見情況的好例子：人人都以為某個量是單調（monotonic）的，其實不然。它同時展示了一個在大學裏經常用到的標準技巧：要判斷一個距離是否正在增加，就對它的*平方*求導（differentiate），這樣便可避開平方根。

動筆之前

想一想 為甚麼處理 OP^2 比處理 OP 容易？「OP 遞增」與「OP^2 遞增」是同一回事嗎？

$OP^2 = x^2 + y^2$ 是 t 的多項式（polynomial），而 OP 則涉及平方根。由於 $OP \geq 0$，而平方運算在非負數上是遞增的，所以 OP 遞增的時候，正正就是 OP^2 遞增的時候。

逐步解題

階段 1　距離的平方

由 $x = ut\cos\alpha$ 及 $y = ut\sin\alpha - \frac{1}{2}gt^2$，可得

$$OP^2 = u^2t^2\cos^2\alpha + u^2t^2\sin^2\alpha - ugt^3\sin\alpha + \tfrac{1}{4}g^2t^4 = u^2t^2 - ugt^3\sin\alpha + \tfrac{1}{4}g^2t^4.$$

想一想 對 t 求導。抽出因子 t 之後，剩下的是甚麼？它在甚麼情況下對所有 t 皆為正？

$$\frac{\mathrm{d}}{\mathrm{d}t}OP^2 = 2u^2t - 3ugt^2\sin\alpha + g^2t^3 = t(g^2t^2 - 3ug\sin\alpha\, t + 2u^2).$$

當 $t > 0$ 時，其正負號與二次式（quadratic）$g^2t^2 - 3ug\sin\alpha\, t + 2u^2$ 的正負號相同。這個二次式的首項係數為正，所以只要它的判別式（discriminant）為負，它對每一個 t 都是正的：

$$9u^2g^2\sin^2\alpha - 8u^2g^2 < 0 \quad\Longleftrightarrow\quad \sin^2\alpha < \frac{8}{9} \quad\Longleftrightarrow\quad \sin\alpha < \frac{2\sqrt{2}}{3}.$$

在這情況下，OP 在整個飛行過程中都在遞增。 ■

階段 2　較陡的拋擲

想一想 若 $\sin\alpha > \frac{2\sqrt{2}}{3}$，這個二次式有兩個實根。它們是正數嗎？較小的那個根是否出現在質點落地之前？

兩根之和為 $\frac{3u\sin\alpha}{g} > 0$，兩根之積為 $\frac{2u^2}{g^2} > 0$，所以兩根皆為正數；而在兩根之間，這個二次式為負，因而 $\frac{\mathrm{d}}{\mathrm{d}t}OP^2$ 也為負。質點在 $T = \frac{2u\sin\alpha}{g}$ 時落地。較小的根是

$$t_1 = \frac{u}{2g}(3\sin\alpha - \sqrt{9\sin^2\alpha - 8}) < \frac{u}{2g}\cdot 3\sin\alpha < \frac{2u\sin\alpha}{g} = T.$$

所以緊接 t_1 之後（而且在落地之前），OP 正在遞減。 ■

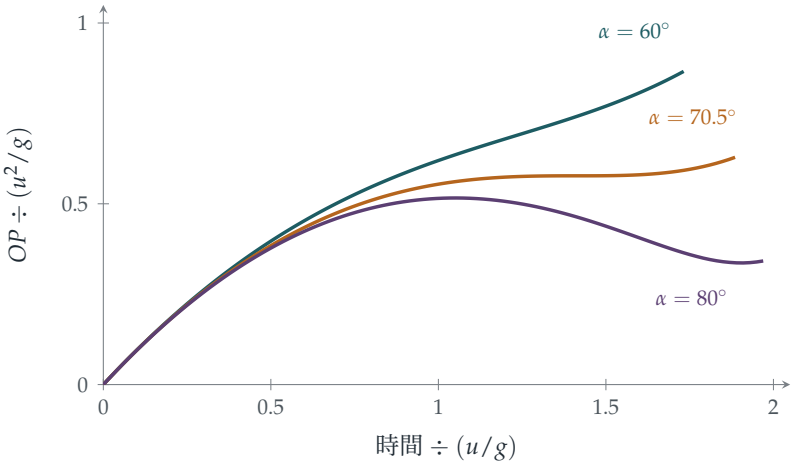


圖 11.10.1 三個不同角度下，球與拋球者的距離隨時間的變化，直至落地為止。在 60°（藍綠色）時，距離一直增加。在臨界角（critical angle）$\arcsin\frac{2\sqrt{2}}{3} \approx 70.5°$（琥珀色）時，距離有一瞬間停頓。在 80°（紫紅色）時，球先遠離，然後有一段時間向拋球者靠近，到落地前一刻又再遠離。

階段 3 移動的觀察者

想一想 Q 沿地面朝相反方向以速率 v 離開 O。從 Q 量度，水平距離是 $x + vt$。這對計算有甚麼影響？

間距 QP 的水平分量為 $(u\cos\alpha + v)t$，垂直分量為 $ut\sin\alpha - \frac{1}{2}gt^2$。所以

$$QP^2 = \big((u\cos\alpha + v)^2 + u^2\sin^2\alpha\big)t^2 - ugt^3\sin\alpha + \tfrac{1}{4}g^2t^4,$$

與之前的式子相同，只是第一項中的 u^2 換成了 $w^2 = (u\cos\alpha + v)^2 + u^2\sin^2\alpha$。其導數（derivative）為 $t\big(g^2t^2 - 3ug\sin\alpha\, t + 2w^2\big)$，而它對所有 $t > 0$ 皆為正的條件是

$$9u^2\sin^2\alpha < 8w^2 \quad\Longleftrightarrow\quad u^2\sin^2\alpha < 8(u\cos\alpha + v)^2,$$

這是在兩邊消去 $8u^2\sin^2\alpha$ 之後所得的結果。兩邊取平方根（$u\sin\alpha$ 與 $u\cos\alpha + v$ 皆為正數），即得 $u\sin\alpha < 2\sqrt{2}(u\cos\alpha + v)$，整理後就是

$$2\sqrt{2}\,v > (\sin\alpha - 2\sqrt{2}\cos\alpha)u. \qquad \blacksquare$$

（這個條件足以令 QP 在整個飛行過程中遞增，因為此時導數對所有 $t > 0$ 皆為正。）

回顧

- **對平方求導。**距離的平方是多項式，而它的導數與距離本身的導數正負號相同。

- **恆為正的二次式。**「在所有時間都遞增」這個條件，化成了一個關於判別式的條件。
- **把新問題化為舊問題。**移動的觀察者只改變了一個係數，所以之前的分析可以直接套用。

背後的數學

幾何觀點

OP 遞減的時候，正正就是 P 的速度指向「內側」的時候，即速度與向量 $\overrightarrow{OP}$ 成鈍角（obtuse angle），也就是 $\overrightarrow{OP} \cdot \mathbf{v} < 0$ 的時候。對於陡峭的拋擲，質點在接近飛行最高點時，位於拋球者的高處而略為偏前，幾乎是水平地向外移動，但已開始下墜；相對於 $\overrightarrow{OP}$ 的方向而言，一旦速度中向下的部分蓋過向外的部分，質點便會靠近。對於較平的拋擲，質點總是在前方足夠遠的位置，以致它的速度從不指回 O。臨界情況就是速度恰好只在一個瞬間與 $\overrightarrow{OP}$ 垂直，亦即圖 11.10.1 中距離停頓的那一刻。

單調性並非必然

許多「顯然」會增加的量，其實並不如此。月球與地球的距離會來回振盪；在彎曲的路上，即使其中一輛車一直領先，兩車之間的距離仍可以減少；獵物大量繁殖的時候，捕食者的數量卻可以下降。要證明一個量是單調的，通常就是要證明某個導數的正負號固定不變，而這裏所用的技巧（化為多項式，再檢查它的根或判別式）正是標準的做法。在分析學（analysis）中，單調性往往是證明收斂（convergence）的關鍵：有界的單調數列必定收斂，正如你在第 2 章所見。

思維技巧

藉導數檢查單調性

要判斷一個量是否一直遞增，可對它（或對它的某個方便處理的函數，例如它的平方）求導，然後問：導數會否變號？

在大學裏，這類關於正負號的論證會用來證明不等式、找出極大值和極小值，以及研究微分方程（differential equation）解的穩定性；在後者中，一個一直遞減的量（如 Section 8.2 所示）便證明了系統會趨於平穩。

延伸練習

1 **有多近？**對於 $\alpha = 80°$，求 OP 取得局部極大值（local maximum）和局部極小值（local minimum）的時間，以及該兩刻的距離，以 u^2/g 為單位。

2 **在斜面上。**把球沿一個與水平成 β 角的斜面向上拋出，並落在斜面上，就這個情況重做第 (i) 部分。臨界角（從斜面量起）會改變嗎？

3 **考慮空氣阻力。**假設球還受到一個與其速度成正比的阻力（resistance）。不作計算，你預期臨界角會大於還是小於 70.5°？為甚麼？

4 **從高處拋出。**若球從懸崖頂拋出，並落在遠低於崖頂的地方，證明無論以甚麼角度拋出，OP 最終都會遞增，並判斷是否存在某些角度，令它在某個階段遞減。

延伸閱讀

- **James Walker, *The Flying Circus of Physics***（Wiley）。數以百計出人意表的日常物理謎題。
- **David Morin, *Introduction to Classical Mechanics***（Cambridge University Press），第 3 章，講述拋體運動，並附有數道有違直覺的題目。
- **J. M. Steele, *The Cauchy–Schwarz Master Class***（Cambridge University Press）。證明一樣東西恆大於另一樣東西的藝術。

11.11 外野手如何接球

題目

STEP II 1998，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/98-s2-q11`

引子

球被高高擊向外野。出色的外野手會立刻起步，並且剛好在球落下的一刻趕到。這是怎樣做到的？沒有人能在不足一秒之內心算出拋體運動（projectile motion）的方程，何況外野手根本不知道球的速率和角度。然而外野手，以至接飛碟的狗，都能輕易做到。

1968 年，物理學家 Seville Chapman 提出了一個答案，而本題正是要證明它。假如外野手留意着球的仰角（angle of elevation）θ，並且移動自己的位置，使 $\tan\theta$ 以恆定的速率增加，她就會剛好到達球的落點。她毋須懂得任何物理，只需能察覺球看起來是升得比穩定上升快，還是比穩定上升慢。一個困難的預測問題，就這樣換成了一條簡單的反饋（feedback）法則。

動筆之前

想一想 設擊球手位於原點，外野手在距離 D 處。若球位於 (x, y)，$\tan\theta$ 是甚麼？

球在外野手眼睛上方高度 y 處，並在她前方水平距離 $D - x$ 處（球正朝她飛來）。所以 $\tan\theta = \frac{y}{D-x}$。

逐步解題

階段 1　位置恰到好處的外野手

想一想 設球的水平速率為 U，初始垂直速率為 W，並在時間 T 落下（回到眼睛的高度）。試以 U、W、T、g 和 t 表示 x、y 和 D。你能把 y 因式分解嗎？

我們有 $x = Ut$ 和 $y = Wt - \frac{1}{2}gt^2$，而球在 $T = \frac{2W}{g}$ 時回到眼睛的高度，外野手正在該處：$D = UT$。由於 $W = \frac{1}{2}gT$，

$$y = \tfrac{1}{2}gt(T - t), \qquad D - x = U(T - t).$$

因子 $T - t$ 可以約去：

$$\tan\theta = \frac{\frac{1}{2}gt(T-t)}{U(T-t)} = \frac{g}{2U}\,t.$$

所以 $\frac{\mathrm{d}}{\mathrm{d}t}\tan\theta = \frac{g}{2U}$，是一個常數。 ■

階段 2　奔跑中的外野手

想一想 現在外野手從距離 D_0 處起步，以速率 v 跑向擊球手，所以她的位置是 $D_0 - vt$。她在時間 T 接到球。D_0 是甚麼？把剛才的計算重做一次。

要在時間 T 接到球，她必須身處球的落點：$D_0 - vT = UT$，所以 $D_0 = (U + v)T$。她與球的水平距離是 $D_0 - vt - Ut = (U + v)(T - t)$，而該因子同樣可以約去：

$$\tan\theta = \frac{\frac{1}{2}gt(T-t)}{(U+v)(T-t)} = \frac{g}{2(U+v)}\,t.$$

所以只要她以適當的速率奔跑，$\tan\theta$ 就會以恆定的速率增加。

想一想 題目問的是逆命題（converse）：假如她奔跑時令 $\frac{\mathrm{d}}{\mathrm{d}t}\tan\theta$ 保持不變，她會接到球嗎？

假設她以速率 v 奔跑，而 $\tan\theta = \frac{\frac{1}{2}gt(T-t)}{D_0-(U+v)t}$ 以恆定的速率 c 增加，所以 $\tan\theta = ct$（在 $t = 0$ 時它等於 0）。於是

$$\tfrac{1}{2}g(T - t) = c(D_0 - (U + v)t) \quad \text{對所有} t.$$

比較 t 的係數和常數項：$\frac{1}{2}g = c(U + v)$ 及 $\frac{1}{2}gT = cD_0$。兩式相除，得 $D_0 = (U + v)T$：這正是在時間 T 身處落點的條件。所以她應選擇 v 使 $\tan\theta$ 保持穩定上升，這樣她就能接到球。 ■

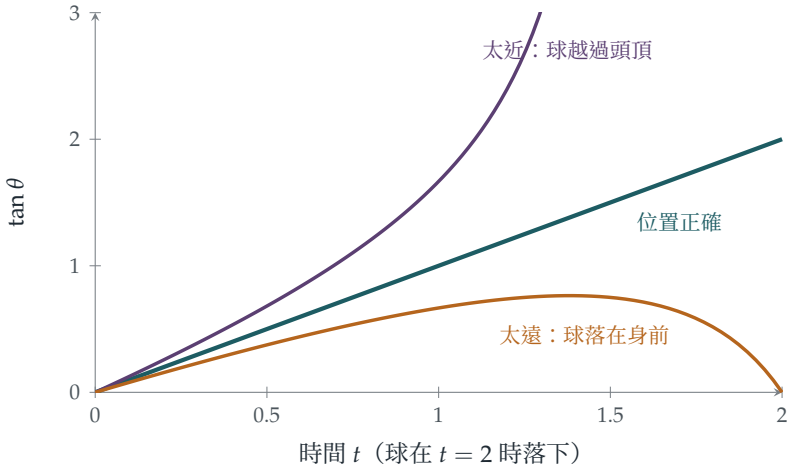


圖 11.11.1 站着不動的外野手所見的情況，球的落點在距離 2 處（取 $g = 2$，水平速率為 1）。站在正確的位置時，$\tan\theta$ 沿直線上升（藍綠色）。站得太近，$\tan\theta$ 會加速上升，球越過她的頭頂（紫紅色）；站得太後，它升得愈來愈慢，然後下降（琥珀色）。法則是：移動自己的位置，使球的影像穩定上升。

回顧

- **找出公因子。**高度和水平距離都含有 $T - t$，把它約去後，$\tan\theta$ 就變成線性。
- **證明逆命題。**要說明這條法則行得通（即它確實令人接到球），需要反方向的論證，也就是比較係數。
- **容易察覺的不變量（invariant）。**$\tan\theta$ 是否穩定上升，憑肉眼就能判斷；落點則不能。

背後的數學

光學加速度抵消

Chapman 的法則現稱為*光學加速度抵消*（optical acceleration cancellation）：外野手移動自己的位置，使「光學加速度」$\frac{\mathrm{d}^2}{\mathrm{d}t^2}\tan\theta$ 保持為零。假如球的影像加速上升，球就會越過她的頭頂，所以她要後退；假如影像減速，球就會落在她身前，所以她要向前跑（圖 11.11.1）。這個策略沒有告訴她球會落在哪裏，卻能把她帶到那裏。這是一條反饋控制律（feedback control law），與導引導彈所用的同出一轍：導彈採用*比例導引*（proportional navigation），使指向目標的視線角保持不變，那正是同一個想法的橫向版本。

外野手實際怎樣做

真實的外野手是否採用 Chapman 的法則？1990 年代有人把攝影機裝在棒球外野手的頭上做實驗，結果並不一致。Michael McBeath 及其同事於 1995 年在 *Science* 上提出，對於打向一側的球，外野手會沿彎曲的路線奔跑，使球的影像在視野中沿直線移動（即「線性光學軌跡」，linear optical trajectory），這同樣能把他們帶到接球的位置。其後的研究顯示，這兩種策略以及兩者的混合都有人使用，而且外野手往往並非以恆定速率奔跑。2004 年的一項研究發現，接飛碟的狗所走的路線，與同樣的光學法則吻合。無論細節如何，這個令人驚嘆的結論依然成立：接球並不需要預測，只需要正確的反饋。

Chapman 的論文

Seville Chapman 的「Catching a baseball」於 1968 年刊登在 *American Journal of Physics*，全文只有三頁，內容基本上就是本題的計算。他指出這個策略只在沒有空氣阻力時才成立，而真實棒球的軌跡會因阻力和旋轉而變形，令情況變得複雜。儘管如此，大腦利用視野中簡單的不變量來解決複雜的控制問題，這個想法在心理學和機械人學中已變得十分重要。

思維技巧

找出容易觀察的不變量

當某個量難以預測時，不妨找一個與它相關、可以直接觀察的量，而這個量恰好在一切順利時表現得簡單。

在這裏，落點難以預測，但 $\tan\theta$ 的變化率卻容易監察。在大學裏，這就是反饋控制的想法，也是選擇一個可觀察量、讓它包含你所需信息的想法。

延伸練習

1. **向後跑。**假如球會落在外野手起步位置的後方，證明她應該背離擊球手奔跑，而同一條法則依然適用。
2. **不是 θ 本身。**證明對於位置恰到好處的外野手，$\frac{d\theta}{dt}$ 並非常數。為甚麼 $\tan\theta$ 才是應該留意的量？
3. **眼睛的高度。**假如球是從外野手眼睛高度以下 h 的位置擊出，對於位置恰當的外野手，$\tan\theta$ 是否仍然是 t 的線性函數？（假設她在眼睛的高度接球。）

4 **空氣阻力。**在線性空氣阻力之下，水平速度會按指數衰減。對於位置恰當的外野手，$\tan\theta$ 是否仍然是 t 的線性函數？

延伸閱讀

- **Seville Chapman，「Catching a baseball」**，*American Journal of Physics* 36 (1968)。
- **Michael McBeath、Dennis Shaffer 及 Mary Kaiser，「How baseball outfielders determine where to run to catch fly balls」**，*Science* 268 (1995)。
- **Dennis Shaffer 等，「How dogs navigate to catch Frisbees」**，*Psychological Science* 15 (2004)。

11.12 橡皮繩上的螞蟻

題目

STEP II 2018，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/18-s2-q10`

引子

一隻螞蟻沿着一條長一公里的橡皮繩出發，每秒走一厘米。每過一秒，繩子都會均勻地再拉長一公里。繩的遠端以每秒一千米的速度飛馳而去，比螞蟻快十萬倍。螞蟻最終能走到繩的盡頭嗎？

令人驚訝的是，答案是能。繩子拉伸時會帶着螞蟻一起前進；雖然螞蟻每次走過的路程佔繩長的比例愈來愈小，但這些比例加起來就像調和級數（harmonic series），而調和級數是發散（diverge）的（Section 2.3）。螞蟻終會到達，只是所需的時間遠遠超過宇宙的年齡。本題精確地解出這個問題的連續版本，並發現所需時間隨兩個速率之比呈指數增長。

動筆之前

想一想 繩子是均勻拉伸的：每一小段都按相同的倍數伸長。在時間 t，繩長為 $a+ut$。若距離釘子 x 的一點位於全長的 $\frac{x}{a+ut}$ 處，這一點移動得有多快？

繩上每一點所佔全長的比例 f 保持不變，所以它與釘子的距離是 $f(a+ut)$，速率是 $fu = \frac{ux}{a+ut}$。靠近釘子的點幾乎不動；遠端則以 u 的速率移動。

逐步解題

階段 1　螞蟻的方程

螞蟻相對於繩子的速率是 v，還要加上牠腳下那段繩子的速率：

$$\frac{\mathrm{d}x}{\mathrm{d}t} = v + \frac{ux}{a+ut}.$$

想一想 用商法則（quotient rule）對 $\frac{x}{a+ut}$ 求導，並代入 $\dot{x}$ 的方程。

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{x}{a+ut}\right) = \frac{\dot{x}}{a+ut} - \frac{ux}{(a+ut)^2} = \frac{1}{a+ut}\left(v + \frac{ux}{a+ut}\right) - \frac{ux}{(a+ut)^2} = \frac{v}{a+ut}. \qquad \blacksquare$$

螞蟻已走過的繩長比例以 $\frac{v}{a+ut}$ 的速率增加：即牠自身的速率除以當時的繩長。拉伸的影響完全消失了，因為拉伸不會改變任何一點所佔的比例。

階段 2　到達盡頭

想一想 由 0 積分至 t。比例何時達到 1？

螞蟻從釘子出發，所以

$$\frac{x}{a+ut} = \int_0^t \frac{v\,\mathrm{d}s}{a+us} = \frac{v}{u}\ln\frac{a+ut}{a}.$$

當比例等於 1 時，螞蟻到達盡頭：$\ln\frac{a+uT}{a} = \frac{u}{v} = k$，所以 $a + uT = ae^k$，即

$$uT = a(e^k - 1). \qquad \blacksquare$$

對數可以無限增大，所以無論繩子拉得多快，螞蟻總會到達。

階段 3　回程

想一想 往回走時，螞蟻相對於繩子的速度是 $-v$。現在牠所佔比例的變化率是多少？牠在時間 T 由比例 1 出發。

以 $-v$ 重複同樣的計算，得 $\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{x}{a+ut}\right) = -\frac{v}{a+ut}$，所以

$$\frac{x}{a+ut} = 1 - \frac{v}{u}\ln\frac{a+ut}{a+uT}.$$

當 $a + ut = (a+uT)e^k$ 時，比例降至 0。由於 $a + uT = ae^k$，回程所需時間為

$$t - T = \frac{(a+uT)(e^k-1)}{u} = \frac{ae^k \cdot uT/a}{u} = Te^k. \qquad \blacksquare$$

回程所需的時間是去程的 e^k 倍：螞蟻掉頭的時候，繩子已是開始時的 e^k 倍長。

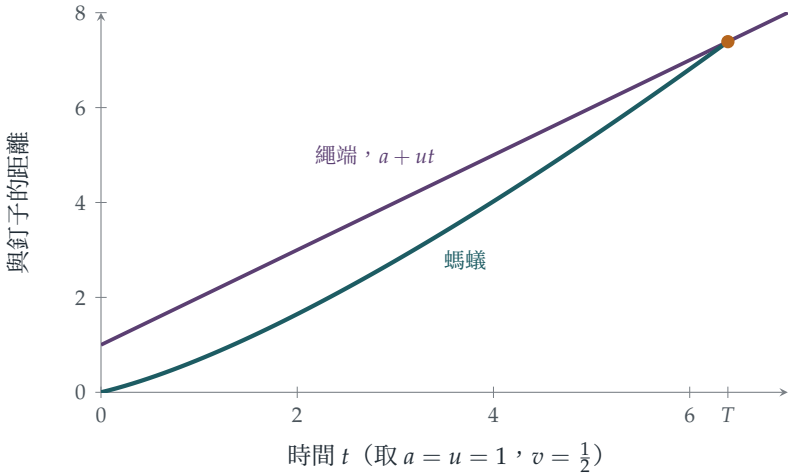


圖 11.12.1 螞蟻（藍綠色）沿着繩子行走，繩端（紫紅色）移動的速率是螞蟻步行速率的兩倍（$k=2$）。螞蟻被拉伸帶着前進，牠與釘子的距離增長得愈來愈快。牠在 $T=e^2-1\approx 6.39$ 時到達盡頭。

回顧

- **選對變量。**已走過的繩長比例是最自然的量：拉伸不會改變它。
- **喬裝了的積分因子（integrating factor）。**乘以 $\frac{1}{a+ut}$ 之後，一條線性微分方程（linear differential equation）就變成了一個恰好的導數。
- **發散勝過速度。**比例像對數一樣增長，而對數是無界的，所以螞蟻總會到達。

背後的數學

調和級數的版本

這個謎題通常是以離散的拉伸來敍述的。一條繩長 1 公里；螞蟻在第一秒走 1 厘米；然後繩子拉長 1 公里；然後螞蟻再走一厘米，如此類推。在第 n 秒，螞蟻在長 n 公里的繩上走了 1 厘米：所佔比例為 $\frac{1}{100\,000\,n}$。拉伸不會改變牠所佔的比例，所以 N 秒之後，牠已走過繩子的

$$\frac{1}{100\,000}\left(1+\frac{1}{2}+\frac{1}{3}+\cdots+\frac{1}{N}\right)\approx\frac{\ln N+0.577}{100\,000}$$

。當 $H_N\approx 100\,000$ 時牠便到達，那大約是 $e^{100\,000}$ 秒之後，這個數有超過 43 000 位數字。（作為比較，宇宙的年齡約為 4×10^{17} 秒。）我們的連續版答案與此吻合：$k=\frac{u}{v}=100\,000$，而 $T\approx\frac{a}{u}e^k$。

膨脹宇宙中的光

繩上的螞蟻，正是光穿越膨脹宇宙的一個模型。遙遠的星系被空間本身的膨脹帶着遠離我們，而它們發出的光則以速率 c 在膨脹的空間之中向我們傳播，就像螞蟻沿着拉伸中的繩子行走一樣。遙遠星系的光最終能否到達我們這裏，取決於膨脹率如何隨時間變化。在一個以穩定速率膨脹的宇宙中，就像那條繩子一樣，光最終總會到達。而在一個加速膨脹的宇宙中（我們的宇宙看來正是如此），有些星系的光永遠不會到達我們這裏：這就是*宇宙事件視界*（cosmic event horizon）。對螞蟻來說，相應的情況是一條按指數拉伸的繩子，繩端的速率與繩長成正比地增長（可在練習中試試）。

加德納的螞蟻

這個謎題似乎是法國數學家 Denys Wilquin 在 1972 年發明的，當時說的是橡皮筋上的一條蟲；其後馬丁·加德納（Martin Gardner）在他的 *Scientific American* 專欄和著作 *Time Travel and Other Mathematical Bewilderments* 中介紹，令它廣為人知。它是最著名的例子之一：結果違反直覺，卻又容易證明；它也是人們說明調和級數發散時最愛用的例子。

思維技巧

找出不受複雜因素影響的量

當問題中某個部分改變了一切（這裏是拉伸），就去找一個它不會改變的量。

拉伸會改變每一段距離，卻不改變任何比例。改用比例來處理，拉伸就從問題中消失了。在大學裏，這個想法會以宇宙學和流體力學中的*共動坐標*（comoving coordinates）的形式出現，更一般地說，就是尋找在某個變換下的不變量（invariant）。

延伸練習

1. **中途。**螞蟻何時走到繩子的一半？那是在時間 $\frac{T}{2}$ 之前還是之後？
2. **指數拉伸。**假設繩長改為 $ae^{\lambda t}$。證明螞蟻僅當 $v > \lambda a$ 時才能到達盡頭。（這就是宇宙事件視界。）
3. **加速的繩子。**若繩長為 $a + \frac{1}{2}bt^2$，螞蟻是否總能到達？

4 **離散的謎題。**在離散版本中，螞蟻需要多少秒才能走過繩子的 1%？請使用 $H_N \approx \ln N + 0.577$。

延伸閱讀

- **Martin Gardner, *Time Travel and Other Mathematical Bewilderments***（W. H. Freeman）。書中收錄了橡皮繩上的螞蟻。
- **Julian Havil, *Gamma: Exploring Euler's Constant***（Princeton University Press）。調和級數，以及這隻螞蟻。
- **Barbara Ryden, *Introduction to Cosmology***（Cambridge University Press）。膨脹的空間、視界與共動坐標，屬大學本科程度。

11.13 鏟雪車問題

題目

STEP III Specimen paper，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/spec-s3-q11`

引子

雪不停地均勻落下。一輛剷雪車（snowplough）沿着一條路出發，前方的積雪愈來愈深，車速也愈來愈慢。稍後，第二輛剷雪車沿着已經清理過的路面跟上。它遇到的只是第一輛車駛過之後才落下的雪，積雪較淺，所以走得較快，最終會追上前車。那是甚麼時候？

這是一道著名問題的變體，原題由 Ralph Agnew 於 1942 年發表，後來成為微分方程（differential equations）中最廣為人知的趣題之一。它的妙處在於，第二輛車的速率取決於第一輛車何時駛過每一點，所以第 (i) 部分的解會代入第 (ii) 部分的方程之中。而這條方程「倒過來」寫，把時間看成距離的函數，才最為巧妙。

動筆之前

想一想 第一輛剷雪車在時間 t 遇到的積雪深度為 $h + kt$。它的速率是多少？求 x 作為 t 的函數較容易，還是求 t 作為 x 的函數較容易？

它的速率是 $\frac{\mathrm{d}x}{\mathrm{d}t} = \frac{k}{\alpha(h+kt)}$。這條方程可以乾淨利落地分離變數（separation of variables）。把它倒過來，$\frac{\mathrm{d}t}{\mathrm{d}x} = \frac{\alpha(h+kt)}{k}$，這是一條以 t 為 x 的函數的簡單線性（linear）方程。對第一輛車來說，兩種寫法都行得通；對第二輛車來說，倒過來的形式卻必不可少。

逐步解題

階段 1　第一輛剷雪車

由 $\frac{\mathrm{d}t}{\mathrm{d}x} = \alpha(\frac{h}{k} + t)$ 可知，$\frac{h}{k} + t$ 這個量隨 x 按指數增長（exponential growth）：$\frac{h}{k} + t = Ce^{\alpha x}$。當 $x = 0$ 時 $t = 0$，所以 $C = \frac{h}{k}$，於是

$$t_1(x) = \frac{h}{k}(e^{\alpha x} - 1). \qquad \blacksquare$$

第一輛車按指數規律減慢：走過相等的距離，所需的時間按指數遞增。

階段 2　第二輛剷雪車的方程

> **想一想**　第二輛剷雪車在時間 t 到達位置 y。那裏的積雪有多深？（第一輛車剛駛過之後，深度是多少？）

第一輛車在時間 $t_1(y)$ 清除了位置 y 上的所有積雪。自此之後，雪以速率 k 累積，所以在時間 t 的深度為 $z = k\big(t - t_1(y)\big)$。因此第二輛車的速率是

$$\frac{\mathrm{d}y}{\mathrm{d}t} = \frac{k}{\alpha k\big(t - t_1(y)\big)} = \frac{1}{\alpha\big(t - t_1(y)\big)},$$

倒過來便得

$$\frac{1}{\alpha}\frac{\mathrm{d}t}{\mathrm{d}y} = t - \frac{h}{k}\big(e^{\alpha y} - 1\big). \qquad \blacksquare$$

階段 3　求解

> **想一想**　這條方程對 t 是線性的：$\frac{\mathrm{d}t}{\mathrm{d}y} - \alpha t = -\frac{\alpha h}{k}(e^{\alpha y} - 1)$。應該用哪個積分因子（integrating factor）？初始條件是甚麼？

乘以 $e^{-\alpha y}$，

$$\frac{\mathrm{d}}{\mathrm{d}y}\big(te^{-\alpha y}\big) = -\frac{\alpha h}{k}\big(1 - e^{-\alpha y}\big), \qquad te^{-\alpha y} = -\frac{\alpha h}{k}y - \frac{h}{k}e^{-\alpha y} + C.$$

第二輛車在 $t = T$ 時由 $y = 0$ 出發，所以 $T = -\frac{h}{k} + C$，得 $C = T + \frac{h}{k}$。再乘回 $e^{\alpha y}$ 並整理，

$$t_2(y) = \frac{h}{k}\big(e^{\alpha y} - 1\big) + \Big(T - \frac{\alpha h y}{k}\Big)e^{\alpha y}. \qquad \blacksquare$$

（題目要求驗證，方法是：求導後代入；各項合併後正好是 $\alpha\big(t - t_1(y)\big)$。）

階段 4　相撞

想一想 兩輛剷雪車在同一時間到達同一點時便會相撞。試比較 $t_2(y)$ 與 $t_1(y)$。

兩者之差為 $t_2(y) - t_1(y) = \left(T - \frac{\alpha h y}{k}\right)e^{\alpha y}$，即兩輛車在位置 y 的時間差距。這個差距縮減至零時，

$$y = \frac{kT}{\alpha h}. \qquad \blacksquare$$

在這一點，第二輛車追上第一輛車：它前方的積雪深度為零，而在這個模型中，它的速率變成無限大（圖 11.13.1）。

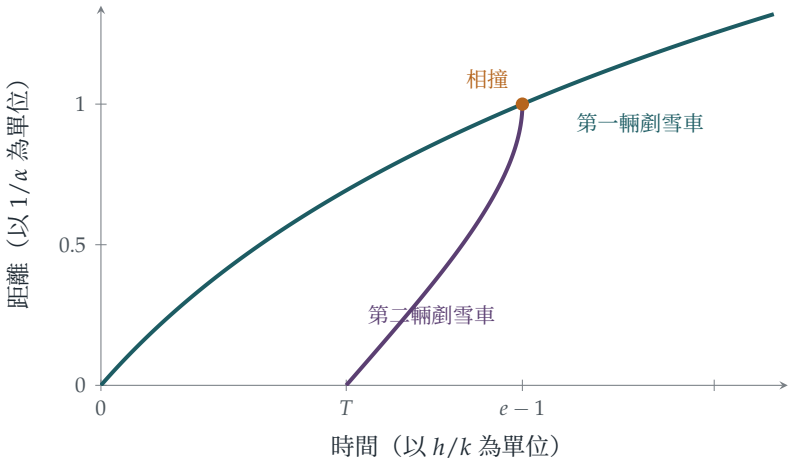


圖 11.13.1 兩輛剷雪車的距離對時間圖，其中 $T = \frac{h}{k}$（即第二輛車在積雪深度倍增之時出發）。第一輛車（藍綠色）隨積雪加深而減慢。第二輛車（紫紅色）較遲出發，路面已清理了一部分，它一邊拉近距離一邊加速，在距離 $\frac{kT}{\alpha h} = \frac{1}{\alpha}$、時間 $(e-1)\frac{h}{k}$ 處追上第一輛車。

回顧

- **把方程倒過來。**把 t 看成位置的函數，兩條方程都是線性的。
- **前一個解代入後一條方程。**第一輛車的時間表 $t_1(y)$ 出現在第二輛車的方程之中。
- **作比較，不必解兩次。**相撞的條件來自差 $t_2 - t_1$，而這個差有一個簡單的因子。

背後的數學

Agnew 的剷雪車

Ralph Palmer Agnew 的版本出自他 1942 年的教科書 *Differential Equations*：某天早上開始下起大雪，雪勢均勻不變；一輛剷雪車在正午出發，第一個小時走了 2 英里，第二個

小時走了 1 英里；雪是甚麼時候開始下的？剷雪車的速率與積雪深度成反比，而深度則與開始下雪以來所經過的時間成正比。答案（見下面的一道練習）涉及黃金比（golden ratio）：雪大約在正午前 37 分鐘開始下。其後才陸續出現像本題這樣有兩輛剷雪車的版本。

會傳遞的資訊

第二輛車的運動取決於路面上記錄下來的歷史：每一點最後一次被清理的時間。這是具有記憶（memory）的系統的一個簡單例子，在這類系統中，某處的狀態取決於該處先前發生過的事。交通流、冰川的運動和山火的蔓延都有這種特性，而它們背後的數學，即一階偏微分方程（first-order partial differential equations）及其特徵線（characteristics）的理論，是大學二年級的標準課題。把時間寫成位置的函數（「前沿何時到達這裏？」），正是處理這類問題時所用的到達時間（arrival times）方法。

思維技巧

選擇自變量（independent variable）

把變量的角色對調，微分方程可能會簡單得多。如果變化率取決於某處發生某事的時間，就把時間寫成位置的函數。

本題的兩條方程寫成 $t(y)$ 的方程後都變成線性。在大學裏，自由選擇哪個變量作自變量（並在方便時更換）是經常用到的手法，由流體力學中的速度圖法（hodograph method），到幾何中以弧長（arc length）作參數，比比皆是。

延伸練習

1 **Agnew 的問題。**雪在正午前某個未知的時間開始以穩定的速率落下。一輛剷雪車的速率與積雪深度成反比，它在正午出發，第一個小時走了 2 英里，第二個小時走了 1 英里。證明若雪在正午前 t_0 小時開始下，則 $\left(1+\frac{2}{t_0}\right)^2=\left(1+\frac{1}{t_0}\right)^3$，由此推出 $\frac{1}{t_0}$ 是黃金比，並求雪何時開始下。

2 **驗證。**用求導驗證 $t_2(y)$ 滿足第 (ii) 部分的微分方程。

3 **時間差距。**證明若 $kT>h$，時間差距 $t_2(y)-t_1(y)$ 先增加，然後減少至零，並求它在何處最大。若 $kT\le h$，又會怎樣？試解釋其意義。

4 **三輛剷雪車。**第三輛剷雪車在時間 $2T$ 出發。寫出它的運動方程。它追上第二輛車，是在第二輛車追上第一輛車之前還是之後？

延伸閱讀

- **Ralph Palmer Agnew, *Differential Equations***（McGraw-Hill, 1942）。剷雪車問題的原始出處。
- **George Simmons, *Differential Equations with Applications and Historical Notes***（CRC Press），書中討論了剷雪車問題和許多其他問題。
- **Richard Haberman, *Mathematical Models***（SIAM）。交通流與特徵線，程度淺易。

11.14 是拋物線，不是懸鏈線

題目

STEP I 1991，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/91-s1-q10`

引子

用手拿着一條鏈的兩端，它便會垂成一條稱為懸鏈線（catenary）的曲線（此名源自拉丁文的「鏈」字）。伽利略（Galileo）以為它是拋物線（parabola），其實不然，少年時代的惠更斯（Huygens）在 1646 年已證明了這一點。不過，若像懸索橋（suspension bridge）那樣，用等距排列的吊桿把沉重的路面掛在鏈上，鏈確實會形成拋物線。分別在於重量如何分佈：懸鏈線的重量沿彎曲的鏈身分佈，而橋的重量則沿水平方向均勻分佈。

本題只需在每個節點分解力（resolving forces），便能證明這條拋物線。關鍵在於張力（tension）的水平分量在整條鏈上處處相同，所以每一段的斜率都與它所承托的重量成正比。結果是一個優美的規律：相鄰節點之間的落差按奇數 $1, 3, 5, 7, \ldots$ 遞增，而這些奇數的部分和（partial sums）正是完全平方數。

動筆之前

想一想 在每個節點 A_i 上有三個力作用：相鄰兩段鏈的張力，以及吊桿上所掛的重量 w。重量無法影響張力的哪一個分量？

重量是鉛垂的，所以在每個節點上，兩個張力的水平分量必須互相平衡。因此，張力的水平分量（設為 H）在鏈的每一段都相同。

逐步解題

階段 1　每個節點上的方程

設線段 $A_{i-1}A_i$ 與水平線所成的俯角為 ϕ_i（由 A_{i-1} 向中間方向量度），則其張力 T_i 的水平分量為 $T_i\cos\phi_i$，鉛垂分量為 $T_i\sin\phi_i$。對 $1\le i\le n$，在 A_i 分解力：

$$T_i\cos\phi_i = T_{i+1}\cos\phi_{i+1} = H, \qquad T_i\sin\phi_i - T_{i+1}\sin\phi_{i+1} = w.$$

（外側的一段向上向外拉，鉛垂分量較大；內側的一段向內並稍微向下拉；兩者之差承托着吊桿。）

> **想一想**　在中間的節點 A_{n+1} 會怎樣？試利用橋的對稱性。

由對稱性，在 A_{n+1} 相接的兩段互為鏡像，所以各自承托該吊桿的一半重量：$T_{n+1}\sin\phi_{n+1} = \tfrac{1}{2}w$。

階段 2　由重量求斜率

由中間向外逐段計算，鉛垂分量為

$$T_{n+1}\sin\phi_{n+1} = \tfrac{1}{2}w, \quad T_n\sin\phi_n = \tfrac{3}{2}w, \quad \ldots, \quad T_1\sin\phi_1 = \tfrac{2n+1}{2}w.$$

把各式除以 H，得斜率為 $\tan\phi_i = \dfrac{(2(n+1-i)+1)w}{2H}$。線段 $A_{i-1}A_i$ 的斜率同時等於 $\dfrac{y_{i-1}-y_i}{h}$，所以

$$\frac{y_n - y_{n+1}}{h} = \frac{w}{2H}, \qquad \frac{y_{n-1}-y_n}{h} = \frac{3w}{2H}, \qquad \ldots, \qquad \frac{y_0 - y_1}{h} = \frac{(2n+1)w}{2H}.$$

取倒數後再乘以相應的奇數，可知 $\dfrac{h}{y_n-y_{n+1}}$、$\dfrac{3h}{y_{n-1}-y_n}$、…、$\dfrac{(2n+1)h}{y_0-y_1}$ 各項均等於 $\dfrac{2H}{w}$。 ■

階段 3　拋物線

> **想一想**　設 j 為由中間向外數的節點序數，使得 A_{n+1-j} 與 A_{n+1} 的水平距離為 jh。它比 A_{n+1} 高出多少？

把各段落差相加，並設 $c = \dfrac{wh}{2H}$：

$$y_{n+1-j} - y_{n+1} = c(1+3+5+\cdots+(2j-1)) = cj^2.$$

若 x 由中間量起，則 $x = jh$，所以該點高出最低節點 $\dfrac{c}{h^2}x^2 = \dfrac{w}{2Hh}x^2$。由對稱性，另一邊亦然，而兩個支承點（$j = n+1$）也符合同一公式。因此所有點都位於拋物線

$$y = y_{n+1} + \frac{w}{2Hh}(x - x_{n+1})^2.$$ ■

上。

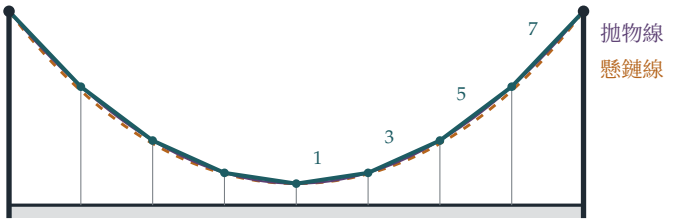


圖 11.14.1 一條在相等水平間距上承受相等負載的鏈（藍綠色）。它的節點位於一條拋物線上（紫紅色），而由中間向外，相鄰節點之間的落差之比為 $1:3:5:7$。一條只受自身重量、兩端和最低點相同的自由鏈，則會垂成懸鏈線（琥珀色虛線），其底部較平坦，近支承點處較陡。

回顧

- **找出不變的量。**只有鉛垂負載時，水平張力處處相同。
- **由中間開始累加。**每一段都承托它與中間之間的全部重量，所以鉛垂分量每次增加 w。
- **奇數之和是平方數。**$1+3+\cdots+(2j-1)=j^2$ 把一列斜率變成了一條拋物線。

背後的數學

拋物線還是懸鏈線？

把上述論證改為連續的版本，便能同時解釋這兩條曲線。對於一條水平張力恆為 H 的鏈 $y(x)$，張力的鉛垂分量為 Hy'，而在一小段上，它的變化等於該段所承受的負載：$H\,\mathrm{d}y' = \mathrm{d}W$。

- 若負載沿水平方向均勻分佈（沉重的路面，每單位路面長度的重量為 ρ），則 $Hy'' = \rho$，所以 y 是二次函數：一條拋物線。
- 若負載是鏈本身的重量，沿鏈身均勻分佈（每單位鏈長的重量為 ρ），則 $Hy'' = \rho\sqrt{1+y'^2}$，其解為懸鏈線 $y = c\cosh\frac{x}{c}$，其中 $c = H/\rho$。

在底部附近，鏈幾乎是水平的，兩者差不多一樣；它們在支承點附近才有分別，因為那裏的鏈較陡，同一段水平距離內包含較多的鏈。真實的懸索橋，橋面遠比纜索重，所以形狀十分接近拋物線。

倒轉的鏈與拱

鏈處於純張力狀態；把它的形狀上下倒轉，便得到一個處於純壓力（compression）狀態的拱，完全沒有彎曲的傾向。胡克（Robert Hooke）在 1675 年以一個字謎（anagram）的形式發表了這個發現，解開後的意思是：「柔軟的繩怎樣懸垂，倒轉過來，堅硬的拱便怎樣屹立。」雷恩（Christopher Wren）與胡克把這個想法用於聖保羅大教堂（St Paul's Cathedral）的圓頂，而高迪（Antoni Gaudí）設計巴塞隆拿的聖家堂（Sagrada Família）時，則從天花板垂下掛着重物的繩子，再把照片倒轉來看。聖路易斯的拱門（Gateway Arch）就是一條（加權的）倒置懸鏈線。

懸鏈線的挑戰

1690 年，雅各布·伯努利（Jakob Bernoulli）向數學家提出挑戰：求出懸垂的鏈的形狀。不出一年，他的弟弟約翰（Johann）、萊布尼茲（Leibniz）和惠更斯都發表了解答。約翰·伯努利後來得意地憶述，出題的哥哥自己反而解不出來。懸鏈線是新興的微積分（calculus）最早的勝利之一，並引出了雙曲函數（hyperbolic functions）cosh 和 sinh；到了十八世紀，Vincenzo Riccati 和蘭伯特（Johann Heinrich Lambert）對這些函數作了有系統的研究。

思維技巧

在每個節點分解力

對於由節點和直桿件組成的結構，在每個節點上令各力平衡。不變的分量和逐步累積的負載，往往會揭示結構的形狀。

到了大學，這會發展成工程學中桁架（trusses）和框架（frameworks）的分析，而在連續極限下，則成為纜索、樑和薄膜的微分方程（differential equations）。由離散的鏈過渡到連續的鏈，與 Section 10.3 中由求和過渡到積分是同一步。

延伸練習

1 **偶數條吊桿。**把題目改為 $2n$ 條吊桿再做一次，此時鏈的最低處是中間的一段水平線段。各節點是否仍然位於一條拋物線上？

2 **張力。**以 H、w 和 n 表示鏈中的最大張力。它出現在哪裏？

3 **懸鏈線。**驗證 $y = c\cosh\frac{x}{c}$ 滿足 $y'' = \frac{1}{c}\sqrt{1+y'^2}$。證明懸鏈線由 $x = 0$ 到 $x = X$ 的長度為 $c\sinh\frac{X}{c}$。

4 **伽利略的錯誤。**一條長度為 $2L$ 的鏈懸掛在相距 $2d$ 的兩個支承點之間。當 L 只比 d 稍大時，證明懸鏈線的垂度（sag）與拋物線的垂度近似相同。為甚麼伽利略的實驗沒有顯示出兩者的分別？

延伸閱讀

- **E. H. Lockwood, *A Book of Curves***（Cambridge University Press）。懸鏈線、拋物線及其相關曲線。
- **Jacques Heyman, *The Stone Skeleton***（Cambridge University Press）。由胡克的鏈說起，講述磚石拱和圓頂如何屹立不倒。
- **David Billington, *The Tower and the Bridge***（Princeton University Press）。橋樑的藝術與工程。

我們學到了甚麼

這一章由會擺動的物體、會旋轉的物體，以及從運動中的位置所看到的景象說起。我們提出了三個問題：運動在甚麼時候才算簡單？怎樣才是合適的觀點？直覺是否可靠？以下是本章十四道題目所揭示的答案。

1. 運動在甚麼時候才算簡單？

簡諧運動（simple harmonic motion）只有單一頻率（frequency），而且頻率與振幅（amplitude）無關，是最理想的情況。以固有頻率（natural frequency）驅動它，便會產生共振（resonance）（Section 11.1）：稍為偏離共振時出現拍（beats），恰好處於共振時則無止境地增長。兩個耦合振子（coupled oscillators）原來只是兩個各自獨立的簡單振子喬裝而成（Section 11.2），但兩者頻率之比是無理數 $\sqrt{3}$，所以合成的運動永不重複。真實的擺一點也不簡單（Section 11.3）：它的週期（period）隨振幅增大，當中涉及一個橢圓積分（elliptic integral），而高斯（Gauss）的算術幾何平均（arithmetic–geometric mean）能以驚人的速度把它算出來。至於等時曲線（tautochrone）（Section 11.4），則說明了只要把軌道彎成擺線（cycloid），運動便可以再次變得完全簡單。

2. 怎樣才是合適的觀點？

幾乎每一道題目，都是靠選對變量或選對參考系（frame）而解決的。

- **坐標。**弧長（arc length）令等時曲線變得簡單；組合 $x \pm y$ 把兩個彈簧解耦（decouple）；改用佔繩長的比例，便消去了螞蟻問題中繩子的伸長（Section 11.12）；把時間寫成距離的函數，令鏟雪車的方程變成線性（Section 11.13）。
- **參考系。**在加速的火車中（Section 11.5），一個虛擬力（fictitious force）使重力傾斜，珠子於是圍繞新的最低點振動。愛因斯坦（Einstein）把這個觀察發展成廣義相對論（general relativity）。
- **取矩的參考點。**對於旋轉的彈珠（Section 11.6），由於摩擦力正正作用於接觸點，所以對接觸點的角動量（angular momentum）守恆，毋須知道摩擦力的大小也可求得最終速率。
- **等效系統。**一根桿的表現與三個質點完全相同（Section 11.7），各質點的權重正是辛普森法則（Simpson's rule）中的權重；而一個立體的平衡位置，就是其法線（normal）

穿過質心（centre of mass）之處（Section 11.8）。

3. 直覺是否可靠？

往往並不可靠。大炮射程所及區域的邊界並不是一條彈道，而是所有彈道的包絡線（envelope）(Section 11.9)。把球拋得夠陡，它在落地之前會朝你飛回來(Section 11.10)。外野手毋須預測球的落點，只要令球的影像保持穩定上升，便可以把球接住（Section 11.11）。一隻螞蟻沿着繩子慢慢爬行，即使繩子伸長的速度比牠快十萬倍，牠仍然能到達盡頭(Section 11.12)。懸垂的鏈並不是拋物線，但承托着橋面的鏈卻是（Section 11.14）。在每一個例子中，一段簡短的計算都以精確的陳述取代了自信卻錯誤的直覺，而且通常還附帶一個解釋。

4. 處理振動與參考系的工具箱

方法	在哪裏用過
特解加餘函數（particular plus complementary solution）	共振（11.1）
以特徵向量（eigenvector）解耦	簡正模式（normal modes）（11.2）
能量積分與代換	擺（11.3）
自然變量	等時曲線（11.4）；螞蟻（11.12）
虛擬力	火車（11.5）
對某一點的角動量	回旋（11.6）
配對各階矩（matching moments）	辛普森桿（11.7）
數算駐點（stationary point）	傾斜的平衡（11.8）
判別式（discriminant）與包絡線	安全拋物線（11.9）
導數的正負號	飛回來（11.10）
比較係數	外野手（11.11）
積分因子（integrating factor）	螞蟻（11.12）；鏟雪車（11.13）
在接點處分解力	橋（11.14）

下一步

- 接下來的兩章轉向概率與統計，由 Section 12.1 開始。
- 簡正模式就是特徵向量（Section 7.2），而 Section 8.6 的量子諧振子（quantum harmonic oscillator），就是本章的振子在量子力學中的版本。
- 極限環(limit cycle)(Section 8.4)是能夠自我維持的振動，而混沌(chaos)(Section 8.7)則是非線性振子受驅動時可能出現的現象。

如果你想在入讀大學之前再進一步

- **A. P. French, *Vibrations and Waves***：講解振動、共振與簡正模式的經典入門書。
- **David Morin, *Introduction to Classical Mechanics***：轉動、參考系，以及許多出人意表的問題。
- **Steven Strogatz, *Sync***：由螢火蟲到橋樑，振子如何變得步伐一致。
- **George Simmons, *Differential Equations with Applications and Historical Notes***：等時曲線、最速降線（brachistochrone）與鏟雪車問題，連同它們的歷史。

Coleman 曾打趣說，物理學家的一生，就是以愈來愈抽象的方式處理諧振子。本章嘗試說明箇中原因：振子是最簡單的運動，而只要選對觀點，幾乎任何事物看起來都像一個振子。學會尋找這個觀點，不論是一個變量、一個參考系，還是一個組合，正是力學所教的最有用的技巧。

延伸練習答案

以下只是簡短的答案和提示，並非完整解答：只列出結果和關鍵步驟，讓你核對自己的做法。請先認真做完每道練習，然後才翻看。

11.1 共振

1. 在升降機的參考系（frame）中，重力實際上相當於 $g - f$，而弦線的勁度（stiffness）是 $\frac{mg}{a}$（在 mg 的作用下它伸長 a）。因此 $\ddot{y} + \frac{g}{a}y = g - f$，開始時 $y = a$、$\dot{y} = 0$。新的平衡位置（equilibrium）是 $a - \frac{af}{g}$，所以 $y = a - \frac{af}{g}(1 - \cos\omega t)$：振幅（amplitude）為 $\frac{af}{g}$，而最小伸長量 $a(1 - \frac{2f}{g})$ 是正數。升降機停止加速時，平衡位置回到 a，此時質點的位移是 $-\frac{af}{g}(1 - \cos\omega T)$，速度是 $-\frac{af}{g}\omega\sin\omega T$。新的振幅是 $\frac{af}{g}\sqrt{(1 - \cos\omega T)^2 + \sin^2\omega T} = \frac{2af}{g}\left|\sin\frac{1}{2}\omega T\right|$。

2. 由 $x = Ct\cos nt$ 得 $\ddot{x} + n^2x = -2Cn\sin nt$，所以 $C = -\frac{nb}{2}$。再加上 $\frac{b}{2}\sin nt$ 以符合初始條件：$x = \frac{b}{2}(\sin nt - nt\cos nt)$。對題目中的公式就 p 使用洛必達法則（L'Hôpital's rule），當 $p \to n$ 時會得到相同的結果。

3. $n\sin pt - p\sin nt = n(\sin pt - \sin nt) + (n - p)\sin nt = 2n\cos\frac{(n+p)t}{2}\sin\frac{(p-n)t}{2} + (n - p)\sin nt$。第二項很小。第一項是一個快速振盪，被包在緩慢變化的包絡（envelope）$\left|\sin\frac{(n-p)t}{2}\right|$ 之內，而這個包絡每隔 $\frac{2\pi}{n-p}$ 重複一次。

4. 振幅是 $\frac{F}{\sqrt{(n^2-p^2)^2+4k^2p^2}}$。把分母對 p^2 求極小：$-2(n^2 - p^2) + 4k^2 = 0$，所以 $p^2 = n^2 - 2k^2$。

11.2 簡正模式

1. 求導後，能量的變化率是 $\dot{x}\big(m\ddot{x} + \frac{\lambda}{a}(2x + y)\big) + \dot{y}\big(m\ddot{y} + \frac{\lambda}{a}(x + 2y)\big) = 0$。設 $u = x - y$ 及 $v = x + y$，能量便是 $\big(\frac{m}{4}\dot{u}^2 + \frac{\lambda}{4a}u^2\big) + \big(\frac{m}{4}\dot{v}^2 + \frac{3\lambda}{4a}v^2\big)$，每個括號都是一個簡諧振子（simple harmonic oscillator）的能量，所以各自都是常數。

2. 週期 $\frac{2\pi}{\omega}$：以 $x = -y$ 及 $\dot{x} = -\dot{y}$ 開始（使得 $v \equiv 0$）。週期 $\frac{2\pi}{\sqrt{3}\omega}$：以 $x = y$ 及 $\dot{x} = \dot{y}$ 開始。

3. 作用於 X 的力是 $-\frac{\lambda}{a}x - \frac{\mu}{a}(x + y)$。模式（mode）$x - y$ 仍然有 $\omega^2 = \frac{\lambda}{ma}$，而 $x + y$ 現在則是 $\frac{\lambda+2\mu}{ma}$。當且僅當頻率之比 $\sqrt{1 + 2\mu/\lambda}$ 是有理數時，所有運動都是週期

性的。

4. 矩陣是 $\frac{k}{m}\begin{pmatrix}2&-1&0\\-1&2&-1\\0&-1&2\end{pmatrix}$，特徵向量（eigenvector）為 $(1,\sqrt{2},1)$、$(1,0,-1)$ 和 $(1,-\sqrt{2},1)$，分別對應 $2-\sqrt{2}$、2 和 $2+\sqrt{2}$。三個模式是：三者一同擺動，中間的擺幅最大；外側兩個反向擺動，中間的靜止不動；以及中間的與外側兩個反向擺動。

11.3 真實的擺

1. 能量：$\frac{1}{2}m\dot{x}^2+\frac{\lambda}{2a}\left(\sqrt{x^2+a^2}-a\right)^2=\frac{1}{2}mv^2$。位移最大時 $\sqrt{x^2+a^2}=a+\frac{v}{k}$，所以 $x_{\max}=\sqrt{\frac{2av}{k}+\frac{v^2}{k^2}}$。當 v 很小時，伸長量約為 $\frac{x^2}{2a}$，所以 $\dot{x}^2\approx v^2-\frac{k^2x^4}{4a^2}$。設 $x=Xu$ 及 $X=\sqrt{2av/k}$，週期 $4\int_0^X\frac{\mathrm{d}x}{\dot{x}}$ 便化為 $\frac{4X}{v}\int_0^1\frac{\mathrm{d}u}{\sqrt{1-u^4}}$。回復力（restoring force）與 x^3 成正比，而不是與 x 成正比（在 $x=0$ 時弦線剛好處於自然長度），所以不存在簡諧極限。

2. $(1-m\sin^2\phi)^{-1/2}\approx 1+\frac{m}{2}\sin^2\phi+\frac{3m^2}{8}\sin^4\phi$，其中 $m=\sin^2\frac{\gamma}{2}\approx\frac{\gamma^2}{4}-\frac{\gamma^4}{48}$。對 ϕ 取平均得 $1+\frac{m}{4}+\frac{9m^2}{64}$，而 γ^4 項是 $-\frac{1}{192}+\frac{9}{1024}=\frac{11}{3072}$。

3. 由 1 和 $\cos 45^\circ=0.70711$ 開始：$(0.85355,0.84090)$，然後 $(0.84722,0.84720)$，然後 0.847213。週期因子是 $\frac{1}{0.847213}=1.18034$。

4. 週期因子的變化是 $\frac{1}{16}(\gamma_5^2-\gamma_4^2)\approx 1.71\times10^{-4}$（以弧度計），所以時鐘每天大約慢 15 秒。

11.4 等時曲線

1. 由於 $y=\frac{s^2}{4k}$，$\ddot{s}=-g\frac{\mathrm{d}y}{\mathrm{d}s}=-\frac{g}{2k}s$：這是 $\omega^2=\frac{g}{2k}$ 的簡諧運動（simple harmonic motion）。由 $s_0=2\sqrt{kh}$ 出發，在 V 處的速率是 $\omega s_0=\sqrt{2gh}$。

2. 這個碗相當於一個長度為 R 的擺。從很小的高度出發，到達底部所需的時間是 $\frac{\pi}{2}\sqrt{R/g}$。從碗邊（$\gamma=90^\circ$）出發，所需時間是這個值的 1.180 倍，約為 $1.85\sqrt{R/g}$。

3. 會。兩者都在四分之一週期後到達 V，並在該處碰撞。無論它們碰撞後以甚麼速度離開，各自都以 s 作簡諧運動，而且週期相同，所以各自恰好在半個週期後回到 V，並再次在 V 碰撞。（碰撞後靜止於 V 的質點則會一直留在該處。）

4. 設 $x=R(\theta-\sin\theta)$ 及 $y=R(1-\cos\theta)$，則 $\mathrm{d}s=2R\sin\frac{\theta}{2}\,\mathrm{d}\theta$，所以一拱的長度是 $\int_0^{2\pi}2R\sin\frac{\theta}{2}\,\mathrm{d}\theta=8R$。面積是 $\int y\,\mathrm{d}x=R^2\int_0^{2\pi}(1-\cos\theta)^2\,\mathrm{d}\theta=3\pi R^2$。

11.5 加速火車中的重力

1. 在車廂的參考系中，每單位質量受到一個向後（向南）的力 a，方向與發射方向相同。飛行時間為 $\frac{2V\sin\theta}{g}$，所以射程是 $\frac{V^2}{g}\big(\sin 2\theta + \frac{a}{g}(1-\cos 2\theta)\big)$。當 $\tan 2\theta = -\frac{g}{a}$ 時它的導數為零，所以 $2\theta \approx \frac{\pi}{2} + \frac{a}{g}$。極大值是 $\frac{V^2}{g}\big(\sqrt{1+a^2/g^2} + \frac{a}{g}\big) \approx \frac{V^2}{g} + \frac{V^2 a}{g^2}$。

2. $x = ut\cos\theta - \frac{1}{2}ft^2$，在 $t = \frac{2u\sin\theta}{g}$ 時着地，着地點為 $x = \frac{2u^2\sin\theta}{g^2}(g\cos\theta - f\sin\theta)$。若 $\tan\theta > \frac{g}{2f}$，水平速度 $u\cos\theta - ft$ 會在着地前變為負數，所以 $\alpha = \tan^{-1}\frac{g}{2f}$。當 $f = g$ 及 $\theta = 45°$ 時，加速度剛好與初速度方向相反：拋體（projectile）沿直線 $y = x$ 筆直向上，停下，然後沿原路直接返回 O。

3. 有效重力（effective gravity）是向下的 g 加上向後的 a。自由液面與它垂直，所以傾斜的角度是 $\tan^{-1}\frac{a}{g}$。

4. 有效重力的大小是 $\sqrt{g^2+a^2}$，在 $x = \frac{ak}{2g}$ 處斜率是 $\frac{a}{g}$，曲率半徑（radius of curvature）是 $\rho = \frac{k}{2}\big(1+\frac{a^2}{g^2}\big)^{3/2}$。週期是 $2\pi\sqrt{\rho/\sqrt{g^2+a^2}} = 2\pi\sqrt{\frac{k(g^2+a^2)}{2g^3}}$，當 $a = 0$ 時簡化為 $2\pi\sqrt{\frac{k}{2g}}$。

11.6 後旋

1. 摩擦力作用於接觸點，所以對該點的角動量（angular momentum）守恆：$maV + \frac{2}{5}ma^2\frac{V}{a} = mav_0 + \frac{2}{5}ma^2\omega_0$。因此 $V = \frac{5v_0+2a\omega_0}{7}$，當 $\frac{v_0}{a\omega_0} < 1$ 時它大於 v_0。

2. $V = \frac{5}{7}v_0$。最終能量是 $\frac{1}{2}m\big(\frac{5}{7}v_0\big)^2\big(1+\frac{2}{5}\big) = \frac{5}{7}\cdot\frac{1}{2}mv_0^2$，所以損失了 $\frac{2}{7}$ 的能量。

3. 以 $I = ma^2$ 重複同一論證，得 $V = \frac{1}{2}(v_0 - a\omega_0)$。令 $V = -\frac{v_0}{7}$，得 $\frac{v_0}{a\omega_0} = \frac{7}{9}$。

4. 在球心上方高度 h 處施加衝量（impulse）J，得 $mv = J$ 及 $\frac{2}{5}ma^2\omega = Jh$。滾動需要 $v = a\omega$，所以 $h = \frac{2}{5}a$。

11.7 桿中的辛普森法則

1. $\frac{1}{6}\big(1 + 4\cdot\frac{4}{5} + \frac{1}{2}\big) = \frac{47}{60} \approx 0.7833$，而 $\frac{\pi}{4} \approx 0.7854$。對於 $\int_{-1}^{1}$，法則 $\frac{1}{3}\big(\mathrm{f}(-1) + 4\mathrm{f}(0) + \mathrm{f}(1)\big)$ 對 x^3 給出 0，對 x^2 給出 $\frac{2}{3}$（兩者都是準確值），但對 x^4 給出 $\frac{2}{3}$，而它的積分其實是 $\frac{2}{5}$。

2. 對於與基準方向成角 ψ 的軸，轉動慣量（moment of inertia）是一個二次型（quadratic form）$A\cos^2\psi + B\sin^2\psi + 2C\sin\psi\cos\psi$。在 $\psi \mapsto \psi + 90°$ 之下不變，迫使 $A = B$ 及 $C = 0$，所以對每一條軸都相同：$\frac{1}{3}Ma^2$，即對平行於一邊的軸的值。

3. 在每邊中點各放置 $\frac{1}{3}$ 的面積，便能準確地在三角形上積分任何二次式（可在直角三

角形上驗證 1、x、x^2、xy）。所以這三個質點與該薄片（lamina）有相同的質量、質心（centre of mass）和二階矩（second moments）。

4. 在每個方向上，取桿在 $-a$、0、a 處的權重 $\frac{1}{6}$、$\frac{2}{3}$、$\frac{1}{6}$ 的乘積：頂點處為 $\frac{M}{216}$，棱的中點處為 $\frac{M}{54}$，面心處為 $\frac{2M}{27}$，中心處為 $\frac{8M}{27}$。對通過中心且平行於一條棱的軸，$I = M\left(\frac{a^2}{3} + \frac{a^2}{3}\right) = \frac{2}{3}Ma^2$。根據平行軸定理（parallel axis theorem），對一條棱的轉動慣量是 $\frac{2}{3}Ma^2 + M(\sqrt{2}a)^2 = \frac{8}{3}Ma^2$。

11.8 傾斜的平衡

1. 由公共面起向圓錐方向量度，G 位於 $\frac{-\frac{2}{3}\cdot\frac{3a}{8} + \frac{\lambda}{3}\cdot\frac{\lambda a}{4}}{\frac{2}{3}+\frac{\lambda}{3}} = \frac{a(\lambda^2-3)}{4(2+\lambda)}$。在半球上，接觸點處的法線必定通過該面的中心，所以當 G 低於該中心，即 $\lambda < \sqrt{3}$ 時，玩具會自行回正。側臥時，玩具沿圓錐的一條母線（generator）及在邊緣處與桌面接觸；若通過 G 的鉛垂線落在頂點與邊緣之間，它便會停留在那裏。這需要 G 的高度至少為 $-\frac{a}{\lambda}$，化簡後即 $\lambda^3 + \lambda + 8 > 0$：恆成立。

2. 均勻球體上的每一點都是平衡位置（隨遇平衡）。若質心稍微偏離球心，便只有兩個：G 在最低處（穩定）和 G 在最高處（不穩定）。

3. 平衡出現在法線通過中心之處：即兩軸的四個端點。靠在短軸的一端是穩定的（G 盡可能低）；靠在長軸的一端則不穩定。

4. 由對稱性，G 在原點。當 $n = 2$ 時，$Y = r\left(\frac{1}{2} - p^2\right)$，而傾斜平衡需要法線通過 G，所以 $Y = 0$ 及 $p = \frac{1}{\sqrt{2}}$。該處的斜率是 $-\sqrt{2}$，所以 $\overrightarrow{AB}$ 傾斜的角度是 $\tan^{-1}\sqrt{2} \approx 54.7°$，若靠在另一個拋物面上則是 $125.3°$。兩者互為關於 x 軸的反射，這就是傾斜角成對出現的原因。

11.9 安全拋物線

1. 每個質點位於 $ut\,\hat{\mathbf{n}} - \frac{1}{2}gt^2\mathbf{j}$，即在以 $-\frac{1}{2}gt^2\mathbf{j}$ 為球心、半徑為 ut 的球面上。從 $(0, \frac{u^2}{2g})$ 水平射出的質點沿 $y = \frac{u^2}{2g} - \frac{gx^2}{2u^2}$ 運動：這正是安全拋物線（parabola of safety）。以角 α 射出的質點在時間 $\frac{u}{g\sin\alpha}$ 於 $x = \frac{u^2}{g\tan\alpha}$ 處與它相切；水平射出的質點需時 $\frac{x}{u}$ 到達該處，所以它必須在時間 $\frac{u(1-\cos\alpha)}{g\sin\alpha} = \frac{u}{g}\tan\frac{\alpha}{2}$ 發射（這需要 $\alpha \geq 45°$，使切點在地面之上）。

2. 該直線與安全拋物線相交於 $R\sin\beta = \frac{u^2}{2g} - \frac{gR^2\cos^2\beta}{2u^2}$ 之處，得 $R = \frac{u^2}{g(1+\sin\beta)}$。在極大值處，練習 3 的兩個角重合，所以最佳角度是角平分線的方向，即 $45° + \frac{1}{2}\beta$。

3. 設 $\lambda = \tan\theta$，兩個值滿足 $\lambda_1 + \lambda_2 = \frac{2u^2}{gx}$ 及 $\lambda_1\lambda_2 = 1 + \frac{2u^2y}{gx^2}$。所以 $\tan(\theta_1 + \theta_2) = -\frac{x}{y}$ 及 $\theta_1 + \theta_2 = 90° + \varphi$，其中 φ 是 OP 的仰角。OP 與鉛垂線的角平分線位於 $\frac{1}{2}(90° + \varphi)$。

4. 頂點位於 $\left(\frac{u^2}{2g}\sin 2\theta, \frac{u^2}{2g}\sin^2\theta\right)$，焦距（focal length）是 $\frac{u^2\cos^2\theta}{2g}$，所以焦點（focus）的高度是 $\frac{u^2}{2g}(\sin^2\theta - \cos^2\theta)$。無論 θ 是多少，它與 O 的距離都是 $\frac{u^2}{2g}$。

11.10 會朝你飛回來的球

1. 取單位使 $u = g = 1$，則 $\frac{\mathrm{d}}{\mathrm{d}t}OP^2 = t(t^2 - 3t\sin\alpha + 2)$，其根為 $t \approx 1.050$（局部極大，$OP \approx 0.516\frac{u^2}{g}$）和 $t \approx 1.904$（局部極小，$OP \approx 0.336\frac{u^2}{g}$）。球在 $t \approx 1.970$ 着地。

2. 沿斜面及垂直於斜面方向分解，$OP^2 = u^2t^2 - ugt^3\sin(\alpha' + \beta) + \frac{1}{4}g^2t^4$，其中 α' 由斜面起量度。這就是平地的公式，只是 $\alpha = \alpha' + \beta$。所以條件 $\sin(\alpha' + \beta) > \frac{2\sqrt{2}}{3}$ 所涉及的是與水平所成的角；若由斜面起量度，臨界角便降至 $70.5^\circ - \beta$（前提是球尚未着地）。

3. 較小。阻力（drag）抵消水平運動，並使下降變得更陡，所以球落回時更接近鉛垂方向，令 OP 更容易縮短。

4. OP^2 的公式不涉及着地點，所以與之前一樣，當且僅當 $\sin\alpha > \frac{2\sqrt{2}}{3}$ 時，OP 會有一段時間遞減（而飛行時間較長，意味着遞減階段必定在着地之前出現）。最終下落佔主導，OP 便會遞增。

11.11 外野手如何接球

1. 若她站在 D_0 處，並以速度 v 背離擊球手（batsman）跑動，球便在她前方 $D_0 + vt - Ut$ 的距離處。要接到球，她需要 $D_0 + vT = UT$，這時距離是 $(U - v)(T - t)$，所以 $\tan\theta = \frac{gt}{2(U-v)}$，同樣是線性的。若球落在她身後，則 $UT > D_0$ 及 $v > 0$：她向後跑。

2. $\theta = \tan^{-1}\frac{gt}{2U}$，所以 $\dot\theta = \frac{g/2U}{1+(gt/2U)^2}$，它是遞減的。$\tan\theta$ 之所以是線性的，是因為高度 $\frac{1}{2}gt(T - t)$ 和距離 $U(T - t)$ 有公因式 $T - t$。

3. 會。相對於她眼睛的高度是一個在接球時刻為零的二次式，所以可因式分解為 $(T - t)(c + dt)$；除以 $U(T - t)$ 後剩下一個線性函數（它起初是負數：她最初要向下看）。

4. 不會。有阻力時，水平距離是 $\frac{U}{k}(\mathrm{e}^{-kt} - \mathrm{e}^{-kT})$，而高度並不等於它乘以 t 的一個線性函數，所以 $\tan\theta$ 不再是線性的。

11.12 橡皮繩上的螞蟻

1. $\frac{x}{a+ut} = \frac{1}{k}\ln\frac{a+ut}{a}$，它在 $t = \frac{a}{u}(\mathrm{e}^{k/2} - 1)$ 時等於 $\frac{1}{2}$。由於 $\frac{T}{2} = \frac{a}{2u}(\mathrm{e}^{k/2} - 1)(\mathrm{e}^{k/2} + 1)$ 而 $\frac{1}{2}(\mathrm{e}^{k/2} + 1) > 1$，這個時刻在 $\frac{T}{2}$ 之前。

2. $\frac{\mathrm{d}}{\mathrm{d}t}\frac{x}{L} = \frac{v}{L} = \frac{v}{a}\mathrm{e}^{-\lambda t}$，所以 $\frac{x}{L} = \frac{v}{a\lambda}(1-\mathrm{e}^{-\lambda t})$，只有當 $v > \lambda a$ 時才會達到 1。
3. $\frac{x}{L} = v\int_0^t \frac{\mathrm{d}s}{a+\frac{1}{2}bs^2}$，它趨向 $\frac{v\pi}{\sqrt{2ab}}$。所以只有當 $v > \frac{\sqrt{2ab}}{\pi}$ 時，螞蟻才會到達。
4. 這需要 $\frac{H_N}{100\,000} = \frac{1}{100}$，所以 $H_N = 1000$，而 $N \approx \mathrm{e}^{999.4} \approx 10^{434}$ 秒。

11.13 鏟雪車問題

1. 正午後時間 t 的積雪深度與 $t+t_0$ 成正比，所以鏟雪車走過的距離是 $c\ln\frac{t+t_0}{t_0}$。條件 $x(1) = 2$ 及 $x(2) = 3$ 給出 $\left(1+\frac{1}{t_0}\right)^3 = \left(1+\frac{2}{t_0}\right)^2$。設 $u = \frac{1}{t_0}$，這便化簡為 $u^2 - u - 1 = 0$，所以 $u = \frac{1+\sqrt{5}}{2}$ 及 $t_0 \approx 0.618$ 小時：大約在 11:23 開始下雪。
2. $\frac{\mathrm{d}t}{\mathrm{d}y} = \frac{\alpha h}{k}\mathrm{e}^{\alpha y} - \frac{\alpha h}{k}\mathrm{e}^{\alpha y} + \alpha\left(T - \frac{\alpha h y}{k}\right)\mathrm{e}^{\alpha y}$，而 $\alpha\left(t - (\mathrm{e}^{\alpha y}-1)\frac{h}{k}\right)$ 與它相同。
3. 時間差距是 $\left(T - \frac{\alpha h y}{k}\right)\mathrm{e}^{\alpha y}$，它的導數在 $y = \frac{kT-h}{\alpha h}$ 處為零。若 $kT > h$，差距會一直增大到該處，然後縮小，在 $y = \frac{kT}{\alpha h}$ 處減至 0。若 $kT \le h$，導數 $\alpha\mathrm{e}^{\alpha y}\left(T - \frac{h}{k} - \frac{\alpha h y}{k}\right)$ 對所有 $y > 0$ 都是負數，所以差距從一開始便縮小。第二輛鏟雪車面對的積雪深度是 $k\times$（差距）；只要這比第一輛鏟雪車在同一地點遇到的積雪深，它便會落後得更遠。
4. $\frac{1}{\alpha}\frac{\mathrm{d}t_3}{\mathrm{d}y} = t_3 - t_2(y)$。寫成 $t_3 = t_2 + g$，得 $g = \mathrm{e}^{\alpha y}\left(T - \alpha T y + \frac{\alpha^2 h y^2}{2k}\right)$。若 $kT \ge 2h$，它在 $y = \frac{k}{\alpha h}\left(T - \sqrt{T^2 - 2hT/k}\right)$ 處為零，早於第二輛鏟雪車追上第一輛之處；若 $kT < 2h$，則第二輛會先追上第一輛。

11.14 是拋物線，不是懸鏈線

1. 是。水平張力 H 仍然是常數，中間一段的斜率為 0，其後斜率每次遞增 $\frac{w}{H}$。以水平一段的中點為起點量度 x，第 k 個接點位於 $x = (k-\frac{1}{2})h$，比最低的接點高出 $\frac{wh}{2H}k(k-1)$，即 $\frac{w}{2Hh}\left(x^2 - \frac{h^2}{4}\right)$：是一條拋物線。
2. 張力是 $\sqrt{H^2+V^2}$，在垂直分量 V 最大之處最大：即在兩端的一段，該處 V 承托一半的負載，即 $\frac{(2n+1)w}{2}$。所以 $T_{\max} = \sqrt{H^2 + \frac{1}{4}(2n+1)^2w^2}$。
3. $y' = \sinh\frac{x}{c}$ 及 $y'' = \frac{1}{c}\cosh\frac{x}{c} = \frac{1}{c}\sqrt{1+\sinh^2\frac{x}{c}}$。長度是 $\int_0^X \cosh\frac{x}{c}\,\mathrm{d}x = c\sinh\frac{X}{c}$。
4. 對於垂度(sag)為 s 的平緩曲線，懸鏈線(catenary)和拋物線的長度都約為 $2d + \frac{4s^2}{3d}$，所以兩者都給出 $s \approx \sqrt{\frac{3}{2}d(L-d)}$；它們只在 s^4 階才有分別。對於一條幾乎拉緊的鏈，這個差別遠小於伽利略（Galileo）所能量度的任何數值。

章

12

概率

不確定的事情，可以精確地推理。直覺呢，原來不行。

本章內容

12.1	精確地取條件	STEP I 2007
12.2	基本比率謬誤	STEP II 1998
12.3	「隨機」是甚麼意思？	STEP III 2022
12.4	兩兩獨立，卻非相互獨立	STEP III 2021
12.5	最後一位乘客的座位	STEP II 2021
12.6	多一枚硬幣	STEP I 2016
12.7	淘汰賽中誰會遇上誰	STEP II 樣本試卷
12.8	Penney 遊戲	STEP III 2001
12.9	用不公平的硬幣作公平的決定	STEP I 2019
12.10	把繩頭繫起來	STEP III 2008
12.11	郵票、杯子與 e	STEP I 1996
12.12	何時停止尋找	STEP I 2001
12.13	三人決鬥	STEP II 1995
12.14	混合策略	STEP I 1989
12.15	滅絕還是爆發	STEP II 2003
12.16	順向與逆向的馬可夫鏈	STEP III 1996
12.17	巴拿赫的火柴盒	STEP II 1992
12.18	無記憶性	STEP III 2021
12.19	平面上的隨機點	STEP II 2012
12.20	蒲豐投針	STEP I 1993
12.21	折斷一根棒	STEP III 1990
12.22	沒有平均值的分佈	STEP III 1995
12.23	由概率得出斯特靈公式	STEP II 2016
12.24	中心極限定理	STEP III 2008
12.25	眾多之中的最大值	STEP III 2025
A	練習答案	第 973 頁

機會的推理

概率論不過是化爲計算的常識而已。

— 拉普拉斯（Pierre-Simon Laplace），*Théorie analytique des probabilités*（1812）

拉普拉斯這句名言只說對了一半。概率（probability）確實是精確化了的常識；但正如本章一再顯示的，任由常識自行其是，就會出錯。人們以為，罕見疾病的檢測結果呈陽性，就表示你很可能患了病；以為若第一位登機的乘客坐錯了座位，最後一位乘客就很難坐到自己的座位；以為一串擲硬幣的結果，不可能勝過另一串同樣長度的結果；以為多次量度的平均值總會穩定下來。這些全都是錯的，而每一次只需簡短的計算，便能揭示真相，更可貴的是，還能揭示當中的原因。

概率始於十七世紀有關賭博的問題，見於帕斯卡（Blaise Pascal）與費馬（Pierre de Fermat）之間的書信。直到 1933 年，柯爾莫哥洛夫（Andrey Kolmogorov）訂立了它的公理（axiom），概率才成為嚴格數學的一個分支。在這期間以及其後，它已發展成數學中應用最廣的部分之一。

概率為何重要

凡是要處理不確定性的領域，都用到概率：醫學（診斷檢測與臨牀試驗）、法律（證據的分量）、金融（風險與保險）、工程（可靠性與洪水）、電腦科學（隨機化演算法、密碼學、機器學習）、物理（量子力學與統計力學）以及生物學（遺傳學與流行病）。現代人工智能在很大程度上就是應用概率：語言模型所做的，是預測下一個字詞出現的概率。而能夠正確地就風險、機會和證據進行推理，是任何受過教育的人所能擁有的最寶貴技能之一。

你在大學會遇到甚麼

大學一年級的概率課程涵蓋公理、條件概率（conditional probability）與貝葉斯定理（Bayes' theorem）、獨立性（independence）、離散與連續隨機變量（random variable）、期望值（expectation）與方差（variance）、母函數（generating function），以及兩大極限定理（limit theorem）：大數定律（law of large numbers）和中心極限定理（central limit theorem）。其後的課程會研究隨機過程（random process），包括馬可夫鏈（Markov chain）、泊松過程（Poisson process）、隨機漫步（random walk）和分支過程（branching process），也會研究測度論概率（measure-theoretic probability，它為極限和連續分佈

奠定穩固的基礎）和鞅（martingale），並把概率與統計學、分析、組合數學和物理聯繫起來。

貫穿本章的問題

三個大問題

1. **問題究竟在問甚麼？**隨機試驗（random experiment）是甚麼？以甚麼為條件？「隨機」又是甚麼意思？
2. **巧妙的切入點在哪裏？**甚麼時候可以用對稱性（symmetry）、指示變量（indicator）、取條件（conditioning）或精心選定的狀態，來取代冗長的計算？
3. **長遠而言會怎樣？**隨機過程隨時間如何演變？許多隨機量的和、平均值和最大值又是甚麼模樣？

本章分為五部分。*清晰思考*（Section 12.1 至 Section 12.4）處理各種陷阱：取條件、基礎比率（base rate）、「隨機」的含義，以及獨立的含義。*對稱性與巧妙論證*（Section 12.5 至 Section 12.9）收集了一些解法既短且美的問題。*期望值、博弈與停止*（Section 12.10 至 Section 12.14）轉向平均值與決策：何時停止、向誰開槍、怎樣隨機化。*隨機過程*（Section 12.15 至 Section 12.19）追蹤隨機性在時間和空間中的表現。*而極限與連續概率*（Section 12.20 至 Section 12.25）則以幾何概率（geometric probability）和幾個重要的極限定理作結。

本章的思想

§	課題	思想
12.1	取條件	「指定的一個」與「至少一個」之別
12.2	貝葉斯定理	基礎比率與自然頻率（natural frequency）
12.3	貝特朗悖論（Bertrand's paradox）	「隨機」需要定義
12.4	獨立性	兩兩獨立（pairwise independent）不等於相互獨立（mutually independent）
12.5	飛機座位	兩個座位之間的對稱性
12.6	多一枚硬幣	互補而對稱的事件
12.7	錦標賽	數算比賽場數與配對
12.8	彭尼博弈（Penney's game）	首步分析（first-step analysis）；非傳遞性（non-transitivity）
12.9	公平硬幣	從偏差中提取公平
12.10	隨機環圈	無記憶的指示變量之和
12.11	杯子與 e	階乘（factorial）與指數級數（exponential series）
12.12	最優停止（optimal stopping）	$\frac{1}{e}$ 法則
12.13	三人決鬥（truel）	策略上的弱勢
12.14	博弈論（game theory）	隨機化策略；極小化極大（minimax）
12.15	分支過程	滅絕作為不動點（fixed point）
12.16	馬可夫鏈	平穩分佈（stationary distribution）；逆轉
12.17	巴拿赫（Banach）的火柴盒	負二項分佈（negative binomial distribution）
12.18	指數分佈（exponential distribution）	無記憶性（memorylessness）與獨立性
12.19	空間隨機性	最近鄰；瑞利（Rayleigh）分佈
12.20	蒲豐投針（Buffon's needle）	幾何概率與 π
12.21	隨機三角形	以面積表示概率
12.22	柯西分佈（Cauchy distribution）	沒有平均值的分佈
12.23	斯特靈公式（Stirling's formula）	由正態曲線（normal curve）得出階乘
12.24	中心極限定理	矩母函數（moment generating function）
12.25	極值（extreme value）	岡貝爾分佈（Gumbel distribution）

其中一節，即 Section 12.24，用到矩母函數，而矩母函數已於 2019 年從 STEP 考試範圍中刪去；該節會解釋所需的內容。

怎樣閱讀本章

當一道概率題令人摸不着頭腦時，不妨就一個細小的情況寫出樣本空間（sample space），然後逐一點算。當答案簡單得令人起疑時，就去找出解釋它的對稱性。而當你的直覺與

計算結果不一致時，請相信計算，然後弄清楚直覺為何出錯：真正的學習正在於此。

12.1 精確地取條件

題目

STEP I 2007，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/07-s1-q13`

引子

「已知」（given that）是概率中最危險的字眼。這道題最後四個部分的說法聽起來幾乎一模一樣：編號為 3 的圓片最先被取出；編號為 3 的圓片被取出；最先取出的是一塊有編號的圓片；取出的圓片中有一塊有編號。其中三個的答案都是 $\frac{1}{2}$，第四個卻不是。

這個分別，正正就是著名的「男孩還是女孩」謎題背後的分別（一個家庭有兩個孩子，其中至少一個是男孩；兩個都是男孩的機會有多大？）。這道謎題難倒了一代又一代的學生，也難倒了不少專業數學家。當中的教訓是：條件（condition）是一個精確的陳述，說明哪些結果仍然可能出現；而資訊是怎樣得來的，與資訊本身同樣重要。

動筆之前

? **想一想** 從十一塊圓片中取出四塊。對於「四塊都有編號」這個事件（event）而言，抽取的次序有沒有影響？四塊圓片組成的等可能（equally likely）組合共有多少個？

每一個由四塊圓片組成的組合被抽中的可能性都相同，所以我們可以不理會次序：共有 $\binom{11}{4} = 330$ 個等可能的組合。

逐步解題

階段 1 沒有條件

由四塊有編號圓片組成的組合有 $\binom{5}{4} = 5$ 個，所以

$$\mathrm{P}(\text{四塊都有編號}) = \frac{5}{330} = \frac{1}{66}.$$

階段 2 3 號最先取出

想一想 編號為 3 的圓片最先被抽出之後，餘下的三次抽取是怎樣的？

其餘三塊圓片是從餘下十塊（四塊有編號、六塊空白）中隨機抽出的三塊，全部 $\binom{10}{3} = 120$ 個組合都是等可能的。所以

$$\text{(ii)}\quad \frac{\binom{4}{3}}{120} = \frac{4}{120} = \frac{1}{30}, \qquad \text{(iii)}\quad \frac{\binom{4}{1}\binom{6}{2}}{120} = \frac{4 \times 15}{120} = \frac{1}{2},$$

因為在 (iii) 中，其餘三塊之中必須恰好有一塊有編號。

階段 3 3 號被取出（在某一次）

想一想 編號為 3 的圓片是在哪一次抽出的，有沒有關係？已知 3 號在四塊之中，對其餘三塊你能說些甚麼？

根據對稱性（symmetry），已知這四塊圓片中包含 3 號，其餘三塊就是從餘下十塊中隨機抽出的三塊，與第 2 階段完全一樣。所以 (iv) 的答案也是 $\frac{1}{2}$。

階段 4 最先取出的是有編號的圓片

已知第一塊圓片有編號（不論是哪個號碼），其餘三塊同樣是從餘下十塊圓片中隨機抽出的三塊，當中四塊有編號、六塊空白。所以 (v) 的答案也是 $\frac{1}{2}$。

階段 5 取出的圓片中有一塊有編號

想一想 現在的條件是「四塊之中至少有一塊有編號」。這與之前任何一個條件相同嗎？330 個組合之中有多少個符合這個條件？

這個條件有所不同：它沒有指明任何一塊特定的圓片。只有 $\binom{6}{4} = 15$ 個組合完全沒有編號圓片，所以有 315 個符合條件。其中 $\binom{5}{2}\binom{6}{2} = 10 \times 15 = 150$ 個恰好包含兩塊有編號的圓片。所以

$$\text{(vi)}\quad \frac{150}{315} = \frac{10}{21}. \qquad \blacksquare$$

停一停，試一試 為甚麼 (vi) 與 (v) 不同？在 (v) 中，我們知道哪一塊圓片有編號（第一塊），而其餘三塊不受限制。在 (vi) 中，我們只知道有某一塊圓片有編號。含有三塊或四塊編號圓片的組合，可以用更多方式滿足「某一塊圓片有編號」，所以它們不會像在 (v) 中那樣比重偏低。

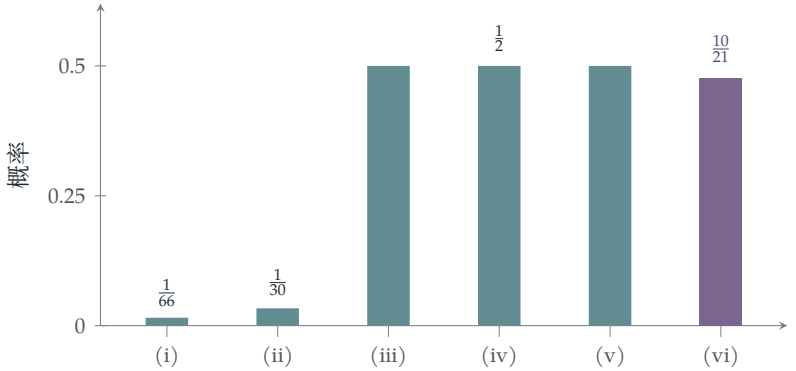


圖 12.1.1 六個答案。第 (iii)、(iv) 和 (v) 部分的條件都是某一塊特定的圓片有編號，答案是 $\frac{1}{2}$。第 (vi) 部分的條件只是有某一塊圓片有編號，答案是 $\frac{10}{21}$。

回顧

- **條件會收窄樣本空間（sample space）。** $\mathrm{P}(A \mid B)$ 是 B 的結果之中同時屬於 A 的比例。
- **對稱性可以省功夫。** 3 號是最先抽出，還是在某一次抽出，並無分別，因為所有位置的地位都相同。
- **「特定的一個」不等於「至少一個」。** 聽起來相似的條件，選出來的結果集合可以大不相同。

背後的數學

男孩還是女孩的謎題

馬丁·加德納（Martin Gardner）在 1959 年提出這個問題：史密斯先生有兩個孩子，其中至少一個是男孩；兩個都是男孩的概率是多少？列出等可能的家庭組合 BB、BG、GB、GG，剔除 GG 後剩下三個，其中一個是 BB：答案是 $\frac{1}{3}$。但如果我們知道年長的孩子是男孩，答案就是 $\frac{1}{2}$。這正是我們這道題第 (vi) 和第 (v) 部分的縮影。加德納後來承認，他最初提出的問題含糊不清：答案取決於我們是怎樣得知有一個男孩的。如果我們隨機遇上其中一個孩子，而他是男孩，答案是 $\frac{1}{2}$；如果我們問「你有沒有至少一個兒子？」而

對方回答有，答案就是 $\frac{1}{3}$。概率取決於產生資訊的過程，而不僅是資訊本身。

蒙提霍爾問題（Monty Hall）

概率中爭議最多的謎題，背後也是同一個教訓。在一個遊戲節目中，三道門的其中一道後面有一輛汽車。你選了 1 號門。主持人知道汽車在哪裏，他必定會打開另外兩道門的其中一道，讓你看到門後是一隻山羊，並且必定會讓你選擇是否換門；當汽車就在你所選的門後，以致他有選擇餘地時，他會在另外兩道門之間隨機挑選。他打開了 3 號門，並讓你改選 2 號門。你應該換嗎？應該：換門後勝出的概率是 $\frac{2}{3}$。主持人選哪道門並非隨機；這取決於汽車的位置，因而帶有資訊。確切的規則十分重要。如果主持人只要可以就一定打開 3 號門，那麼看到 3 號門打開，1 號門和 2 號門就變得同樣可能，換門後勝出的概率只有 $\frac{1}{2}$；然而，一個事先決定必定換門的參賽者，仍然有 $\frac{2}{3}$ 的機會勝出，因為換門勝出，恰好就是第一次選錯的時候。Marilyn vos Savant 在 1990 年於雜誌專欄中給出這個答案時，收到數以千計的來信指她錯了，其中有些更是出自數學家之手。用電腦模擬一次，或者仔細列出各個等可能的情況，就可以解決爭議。

貝葉斯（Bayes）與拉普拉斯（Laplace）

條件概率（conditional probability）的系統化基礎，是由托馬斯·貝葉斯（Thomas Bayes）和皮埃爾–西蒙·拉普拉斯（Pierre-Simon Laplace）奠定的：貝葉斯的論文在他死後於 1763 年發表，而拉普拉斯則在 1774 年以一般形式寫出了現稱為貝葉斯定理（Bayes' theorem）的法則。安德烈·柯爾莫哥洛夫（Andrey Kolmogorov）在 1933 年提出的公理（axioms），是現今每一個大學課程的起點，當中定義 $\mathrm{P}(A \mid B) = \mathrm{P}(A \cap B)/\mathrm{P}(B)$，而概率論的其餘部分都建基於這個定義之上。下一節 Section 12.2 會展示貝葉斯定理在實際應用中的作用。

思維技巧

準確把握條件

計算條件概率之前，先確切寫下哪些結果符合條件。問一問資訊是怎樣得來的。

在大學裏，條件概率和條件期望（conditional expectation）是概率與統計中最重要的工具之一，也是最常出錯的地方。把樣本空間寫出來，至少就一個小規模的情況寫出來，是最好的防範方法。

延伸練習

1 **推廣。**設有 b 塊空白圓片和 n 塊有編號的圓片，從中抽出 k 塊，證明 (iv) 和 (v) 的答案永遠相等，並求出 (vi) 在甚麼情況下與它們相等。

2 **星期二出生的男孩。**一個家庭有兩個孩子，其中至少一個是在星期二出生的男孩。證明兩個都是男孩的概率是 $\frac{13}{27}$。(列出 14×14 個等可能的情況，或者用巧妙的方法計數。)

3 **蒙提霍爾問題。**列出各個結果，驗證 $\frac{2}{3}$ 這個答案。如果主持人不知道汽車在哪裏，隨機打開另外兩道門的其中一道，而門後剛好是一隻山羊，那又如何？

4 **兩張 A。**一手 13 張的橋牌中至少有一張 A。這手牌中至少有兩張 A 的概率是多少？試與已知這手牌中有葵扇 A 時的概率比較。

延伸閱讀

- **Jason Rosenhouse, *The Monty Hall Problem***（Oxford University Press）。講述這道謎題、它的歷史，以及它的眾多變化。
- **Geoffrey Grimmett & Dominic Welsh, *Probability: An Introduction***（Oxford University Press），有關條件概率的各章。
- **Joseph Blitzstein & Jessica Hwang, *Introduction to Probability***（CRC Press；可於網上免費閱讀）。對條件概率的處理十分出色，並載有許多悖論。

12.2 基本比率謬誤

題目

STEP II 1998，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/98-s2-q12`

引子

某種疾病的檢測，對患者的準確度達 90%，而且甚少誤報。Frankie 的檢測結果呈陽性。Frankie 應該有多擔心？大多數人都會答「大約 90%」，而在許多研究中，大多數醫生也是這樣回答。在這道題目中，正確答案是 $\frac{1}{11}$，約為 9%，足足小了十倍。

錯處在於忽略了*基礎比率*（base rate）：這種疾病很罕見，而健康的人為數眾多，所以即使假陽性（false positive）率很低，誤報的數目仍會遠遠多於真正的病例。這種*基礎比率謬誤*（base-rate fallacy）在醫學、法律（在法律上稱為檢控官謬誤（prosecutor's fallacy））、保安檢查和機器學習中都舉足輕重。貝葉斯定理（Bayes' theorem）正是對症良藥。

動筆之前

想一想 試想像有 1000 人。當中有多少人患病？這些患者之中有多少人檢測呈陽性？健康的人之中又有多少人檢測呈陽性？

在 1000 人當中，有 10 人患病，其中 9 人檢測呈陽性。在 990 個健康的人當中，有 $\frac{1}{11}$ 檢測呈陽性，即 90 個假陽性。因此共有 99 人檢測呈陽性，其中只有 9 人真正患病。

逐步解題

階段 1 貝葉斯定理

想一想　以 D 表示「Frankie 患有此病」，以 $+$ 表示「AL 檢測呈陽性」。你已知 $\mathrm{P}(+ \mid D)$，想求 $\mathrm{P}(D \mid +)$。兩者有甚麼關係？

根據條件概率（conditional probability）的定義，$\mathrm{P}(D \mid +) = \frac{\mathrm{P}(D \cap +)}{\mathrm{P}(+)}$，而 $\mathrm{P}(+)$ 可按 Frankie 是否患病拆成兩部分：

$$\begin{aligned}\mathrm{P}(D \mid +) &= \frac{\mathrm{P}(+ \mid D)\mathrm{P}(D)}{\mathrm{P}(+ \mid D)\mathrm{P}(D) + \mathrm{P}(+ \mid \bar{D})\mathrm{P}(\bar{D})} \\ &= \frac{0.9 \times 0.01}{0.9 \times 0.01 + \frac{1}{11} \times 0.99} = \frac{0.009}{0.009 + 0.09} = \frac{1}{11}.\end{aligned}$$

■

這正是那一千人當中的 $\frac{9}{99}$。

階段 2　第二次檢測

想一想　只有 AL 檢測呈陽性的人才會接受第二次檢測。在這羣人之中，患者佔多大比例？把這個比例當作新的「基礎比率」。

在 AL 檢測呈陽性的人當中，有 $\frac{1}{11}$ 是患者，$\frac{10}{11}$ 不是。以這兩個數作為先驗概率（prior probability），再次應用貝葉斯定理，得

$$\mathrm{P}(D \mid \text{兩次均呈陽性}) = \frac{0.9 \times \frac{1}{11}}{0.9 \times \frac{1}{11} + p \times \frac{10}{11}} = \frac{0.9}{0.9 + 10p}.$$

令上式等於 $\frac{1}{2}$，得 $10p = 0.9$，所以 $p = 0.09$。 ■

（這裏假設在已知 Frankie 是否患病的條件下，兩項檢測的誤差互相獨立（independent），題目亦隱含這個假設。）

停一停，試一試　當 $p = 0.09$ 時，STEP 的假陽性率與 AL 的相若，但第二次陽性結果卻令概率由 $\frac{1}{11}$ 升至 $\frac{1}{2}$。為甚麼質素相同的檢測，第二次的影響會大得多？

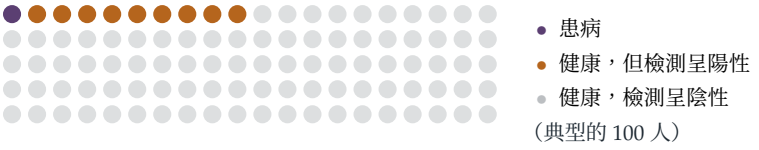


圖 12.2.1　典型的一百人。其中一人患病（紫色），十次檢測中有九次會呈陽性。在 99 個健康的人當中，約有 9 人會被誤判為陽性（琥珀色）。在檢測呈陽性的人當中，健康的人與患者的數目之比約為十比一。

回顧

- **自然頻率（natural frequencies）。**想像 1000 人，答案便一目了然，也容易驗算。
- **先驗乘以似然（likelihood）。**後驗概率（posterior probability）既取決於檢測的準確度，也取決於基礎比率。
- **分階段更新。**昨天的後驗就是今天的先驗：第二次檢測是由 $\frac{1}{11}$ 開始，而不是 $\frac{1}{100}$。

背後的數學

勝算與證據

以勝算（odds）來表達，貝葉斯定理最為清晰。D 的勝算是 $\mathrm{P}(D) : \mathrm{P}(\bar{D})$。一項證據 E 會把勝算乘以*似然比*（likelihood ratio）$\frac{\mathrm{P}(E|D)}{\mathrm{P}(E|\bar{D})}$：

$$\text{後驗勝算} = \text{先驗勝算} \times \text{似然比}.$$

這裏的先驗勝算是 $1:99$，而 AL 的似然比是 $0.9 \div \frac{1}{11} = 9.9$，於是後驗勝算為 $9.9:99 = 1:10$，即概率為 $\frac{1}{11}$。第二項獨立檢測的似然比若為 10（$p = 0.09$ 時的 STEP 正是如此），便會把勝算帶到 $1:1$。取對數之後，證據就可以*相加*：每項獨立檢測所貢獻的，是其似然比的對數。圖靈（Alan Turing）在布萊切利園（Bletchley Park）破解恩尼格瑪（Enigma）密碼電文時，用的正是這個方法，並以「ban」和「deciban」作為量度證據的單位。

檢控官謬誤

1999 年，事務律師 Sally Clark 被裁定謀殺兩名年幼兒子罪名成立，定罪的部分依據是一項證據，指同一家庭發生兩宗嬰兒猝死的機會是「7300 萬分之一」。姑且不論這個數字本身有錯（它假設兩宗死亡互相獨立，忽略了遺傳和環境因素），這個推論還把 P(證據 | 無辜) 與 P(無辜 | 證據) 混為一談。母親連殺兩子同樣極為罕見，真正應該比較的是這兩種解釋。皇家統計學會（Royal Statistical Society）發表公開信指出這個錯誤，她的定罪於 2003 年獲撤銷。當中的數學與本題完全相同。

Gigerenzer 與自然頻率

在 1990 年代，心理學家 Gerd Gigerenzer 指出，醫生若獲得以條件概率表達的檢測統計數字，遇到本題這類問題時大多錯得離譜；但同樣的資料若以自然頻率表達（「在 1000 名女性當中，有 10 人患乳癌；其中 8 人檢測呈陽性…」），大多數醫生都能答對。他的研究改變了許多國家向病人傳達醫療風險的方式。

思維技巧

用貝葉斯定理把條件概率倒轉過來

$\mathrm{P}(A \mid B)$ 與 $\mathrm{P}(B \mid A)$ 是兩個不同的數。要由其中一個求出另一個，你需要基礎比率 $\mathrm{P}(A)$。

貝葉斯推理，即根據證據更新信念，是現代統計學和機器學習的基礎，由垃圾郵件過濾器以至醫學診斷都用得上。到了大學，你會遇到它的進階形式——貝葉斯推斷（Bayesian inference），屆時處理的不再是單一的概率，而是先驗分佈和後驗分佈。

延伸練習

1 **雪櫃裏的屍體。**試做 STEP II 1987，第 16 題（`step.maths.org/questions/87-s2-q16`）：貝葉斯定理應用於兩名可靠程度不同的證人，是 STEP 中較令人難忘的題目情境之一。

2 **要多罕見？**某項檢測的靈敏度（sensitivity）為 0.9，假陽性率為 $\frac{1}{11}$。疾病要有多普遍，陽性結果才表示患病的機會高於一半？

3 **重複檢測。**假如重複進行 AL 檢測，而在已知受測者真實狀況的條件下各次結果互相獨立，那麼兩次 AL 檢測均呈陽性後，患病的概率是多少？三次之後呢？

4 **篩查。**某項篩查計劃為一百萬人檢測某種病症，其患病率（prevalence）為 10 000 分之 1，所用檢測的靈敏度為 99%，假陽性率為 1%。有多少人會被告知可能患有此病症？當中又有多少人真的患病？

延伸閱讀

- **Gerd Gigerenzer, *Reckoning with Risk***（Penguin）。自然頻率、醫學檢測與法庭。
- **David Spiegelhalter, *The Art of Statistics***（Pelican）。內容包括醫學篩查和 Sally Clark 案。
- **Sharon Bertsch McGrayne, *The Theory That Would Not Die***（Yale University Press）。貝葉斯定理的歷史，包括圖靈如何運用這個定理。

12.3 「隨機」是甚麼意思？

題目

STEP III 2022，第 12 題

繼續閱讀之前，請先看題目（原題為英文）。題目載於 OCR 提供下載的 2022 年 STEP 試卷（zip 檔案）內的 STEP III 試卷：

`www.ocr.org.uk/Images/696587-2022-past-papers.zip`

引子

「在圓上隨機選一條弦（chord），它的期望長度（expected length）是多少？」這條問題聽起來好像有一個答案。其實沒有，除非你先說清楚「隨機」是甚麼意思。在圓周上隨機取兩點再連起來，期望長度是 $\frac{4a}{\pi} \approx 1.27a$。在圓內隨機取一點，再隨機取一個方向作過該點的弦，答案卻是 $\frac{16a}{3\pi} \approx 1.70a$。兩種做法都是選取「一條隨機弦」的合理方法。

這就是貝特朗悖論（Bertrand's paradox），由約瑟夫·貝特朗（Joseph Bertrand）於 1889 年發表，用意是提出警告。只有在指明隨機試驗（random experiment）之後，概率（probability）才有明確的定義。這個悖論促使數學家在二十世紀為概率論建立嚴格的基礎：在這套基礎之中，每一個概率都定義在一個精確界定的結果空間之上。

動筆之前

想一想 若角 AOB 為 θ，弦 AB 的長度是多少？（由 O 作垂線。）

由 O 所作的垂線平分該弦，也平分該角，所以 $AB = 2a\sin\frac{\theta}{2}$。

逐步解題

階段 1　圓周上的兩點

$$\mathrm{E}(AB) = \frac{1}{2\pi}\int_0^{2\pi} 2a\sin\frac{\theta}{2}\,\mathrm{d}\theta = \frac{a}{\pi}\Big[-2\cos\frac{\theta}{2}\Big]_0^{2\pi} = \frac{a}{\pi}\cdot 4 = \frac{4a}{\pi}. \qquad \blacksquare$$

階段 2　圓內的一個隨機點

想一想　對於 $0 \le r \le a$，$\mathrm{P}(R \le r)$ 是多少？對它求導，便得到密度。

$\mathrm{P}(R \le r)$ 是半徑為 r 的圓盤面積除以整個圓盤的面積，即 $\frac{r^2}{a^2}$。所以 R 在 $[0,a]$ 上的密度（density）是 $\frac{2r}{a^2}$。（點較有可能離圓心較遠，因為那裏的面積較大。）

想一想　一條過 C 的弦與 OC 成角 t。這條弦與 O 相距多遠？

由 O 到該直線的距離是 $R\sin t$，所以弦長為 $2\sqrt{a^2 - R^2\sin^2 t}$。

階段 3　對 C 取平均

設 $s = \sin t$，

$$\begin{aligned} L(t) &= \int_0^a 2\sqrt{a^2 - r^2 s^2}\cdot\frac{2r}{a^2}\,\mathrm{d}r = \frac{4}{a^2}\Big[-\frac{(a^2 - r^2s^2)^{3/2}}{3s^2}\Big]_0^a \\ &= \frac{4}{a^2}\cdot\frac{a^3 - a^3(1-s^2)^{3/2}}{3s^2} = \frac{4a(1-\cos^3 t)}{3\sin^2 t}. \qquad \blacksquare \end{aligned}$$

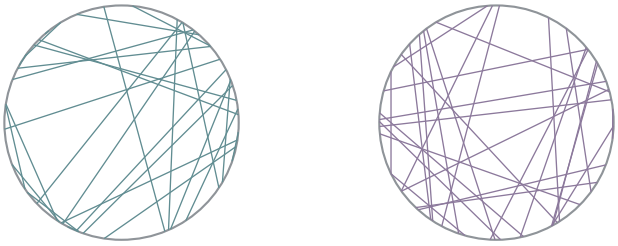
想一想　把 $1 - \cos^3 t$ 和 $\sin^2 t = 1 - \cos^2 t$ 因式分解。有甚麼可以約去？

設 $c = \cos t$，則 $\frac{1-c^3}{1-c^2} = \frac{(1-c)(1+c+c^2)}{(1-c)(1+c)} = \frac{1+c+c^2}{1+c} = c + \frac{1}{1+c}$，而 $1 + \cos t = 2\cos^2\frac{t}{2}$。所以

$$L(t) = \frac{4a}{3}\Big(\cos t + \tfrac{1}{2}\sec^2\tfrac{t}{2}\Big). \qquad \blacksquare$$

階段 4　方向也隨機

$$\begin{aligned} \mathrm{E}\big(L(T)\big) &= \frac{2}{\pi}\int_0^{\pi/2}\frac{4a}{3}\Big(\cos t + \tfrac{1}{2}\sec^2\tfrac{t}{2}\Big)\,\mathrm{d}t \\ &= \frac{8a}{3\pi}\Big[\sin t + \tan\tfrac{t}{2}\Big]_0^{\pi/2} = \frac{8a}{3\pi}(1+1) = \frac{16a}{3\pi}. \qquad \blacksquare \end{aligned}$$



兩個隨機端點
平均長度 $\frac{4a}{\pi} \approx 1.27a$

圓盤內的隨機中點
平均長度 $\frac{4a}{3} \approx 1.33a$

圖 12.3.1 以兩種不同方法選出的二十八條「隨機弦」。左：圓周上兩個獨立的均勻分佈點。右：以圓盤內一個均勻隨機的點作為弦的中點。右邊的方法在邊緣附近產生的短弦較少。不同的方法給出不同的分佈，也給出不同的答案。

回顧

- **指明試驗。**「一條隨機弦」在每一部分都化成了一個精確的步驟，而每個步驟給出不同的答案。
- **由面積求密度。**$\mathrm{P}(R \leq r)$ 直接來自幾何；求導便得到密度。
- **分階段取平均。**先對 C 的位置取平均（t 固定），再對 t 取平均。

背後的數學

貝特朗的三個答案

貝特朗原本的問題是：一條隨機弦比內接等邊三角形的邊長 $\sqrt{3}a$ 更長的概率是多少？他給出了三種方法和三個答案。

- 隨機端點：固定一個端點；若另一個端點落在圓周中間的三分之一之內，弦便是長的。概率為 $\frac{1}{3}$。
- 隨機取一條半徑，再在其上隨機取一點作為中點：若弦與 O 的距離小於 $\frac{a}{2}$，弦便是長的。概率為 $\frac{1}{2}$。
- 在圓盤內任意位置隨機取中點：若中點與 O 的距離在 $\frac{a}{2}$ 之內，弦便是長的。概率為 $\frac{1}{4}$。

每一種方法都是弦的集合上一個不同的概率分佈（probability distribution），當中並無矛盾：它們回答的是不同的問題。這裏的教訓是，「均勻隨機」只有相對於某個選定的參數化（parametrisation）才有意義，而對於連續的集合，通常並沒有唯一自然的參數化。

有沒有正確答案？

1973 年，物理學家 Edwin Jaynes 主張第二種方法才是「正確」的，理由是把圓稍為移動或縮放，答案都不應改變：一條隨機弦應該就像一根拋到地上所畫圓圈上的飲管。只有與圓心距離均勻分佈的那個分佈才具有這種不變性（invariance），而拋飲管的實驗也證實了 $\frac{1}{2}$。同一個想法，即正確的分佈就是尊重情境中各種對稱性的那一個，正是積分幾何（integral geometry）的基礎；積分幾何是研究隨機直線的數學，它也解釋了 Section 12.20 中的布豐投針（Buffon's needle）。

從悖論到公理

貝特朗在 *Calcul des probabilités*（1889）中發表了他的弦悖論，部分原因是要嘲諷他眼中對「幾何概率」的草率運用。龐加萊（Henri Poincaré）也討論過它。1900 年，希爾伯特（David Hilbert）把概率的公理化（axiomatisation）列入他為新世紀提出的著名問題之中，而柯爾莫哥洛夫（Andrey Kolmogorov）在 1933 年解決了這個問題：概率是結果空間上的一個測度（measure），任何關於「隨機」對象的問題，都必須先說明是哪一個空間、哪一個測度。貝特朗悖論正是說明為何必須如此的標準例子。

思維技巧

界定隨機試驗

計算概率之前，先準確描述隨機對象是怎樣產生的。不同的描述可以給出不同的答案。

在大學裏，每一個概率計算都由一個樣本空間（sample space）和一個概率測度（probability measure）出發。在統計學和物理學中，選擇合適的隨機模型（往往藉助對稱性）是工作的一大部分。

延伸練習

1 **貝特朗的概率。**驗證上文的三個概率 $\frac{1}{3}$、$\frac{1}{2}$、$\frac{1}{4}$。

2 **另一個平均值。**證明若弦的中點在圓盤內均勻選取（右邊的圖），期望長度是 $\frac{4a}{3}$。（提示：這就是 $L\left(\frac{\pi}{2}\right)$。為甚麼？）

3 **隨機距離。**若弦與 O 的距離在 $[0, a]$ 上均勻分佈，證明期望長度是 $\frac{\pi a}{2}$。

4 **三種方法，三個分佈。**對三種方法中的每一種，求弦長 ℓ 在 $[0, 2a]$ 上的概率密度（probability density）。

延伸閱讀

- **Edwin Jaynes, "The well-posed problem"**, *Foundations of Physics* 3 (1973)。不變性論證，易讀而又發人深省。
- **Gábor Székely, *Paradoxes in Probability Theory and Mathematical Statistics*** (Reidel)。貝特朗悖論及其他許多悖論。
- **Daniel Klain & Gian-Carlo Rota, *Introduction to Geometric Probability*** (Cambridge University Press)。積分幾何，留待日後研讀。

12.4 兩兩獨立，卻非相互獨立

題目

STEP III 2021，第 12 題

繼續閱讀之前，請先看題目（原題為英文）。題目載於 OCR 提供下載的 2021 年 STEP 試卷（zip 檔案）內的 STEP III 試卷：

`www.ocr.org.uk/Images/696588-2021-past-papers.zip`

引子

玩家 1 和玩家 2 擲出相同的點數。這能告訴你玩家 2 和玩家 3 是否擲出相同點數嗎？出人意料的是，不能：這兩個事件是獨立（independent）的。玩家 1 和玩家 3 的情況也一樣。但如果 1 與 2 相同，而 2 又與 3 相同，那麼 1 必定與 3 相同。三個事件，每兩個之間都獨立，合起來卻不獨立。

兩兩獨立（pairwise independence）比完全獨立弱，而這個差別至關重要。就方差（variance）而言，兩兩獨立已經足夠，本題正是利用了這一點。但在許多其他用途上，這並不足夠。電腦科學家會刻意使用兩兩獨立的隨機數，因為產生它們的成本遠低於產生完全獨立的隨機數，而對許多演算法來說已經夠用。

動筆之前

? **想一想** $\mathrm{P}(X_{12}=1)$ 是多少？$\mathrm{P}(X_{12}=1 \text{ 及 } X_{23}=1)$ 又是多少？

無論玩家 1 擲出甚麼，玩家 2 與之相同的概率都是 $\frac{1}{6}$。$X_{12}=1$ 和 $X_{23}=1$ 同時成立，即三人擲出相同的點數：概率為 $\frac{6}{216}=\frac{1}{36}$。

逐步解題

階段 1 **兩兩獨立**

由於 $\mathrm{P}(X_{12}=1, X_{23}=1)=\frac{1}{36}=\frac{1}{6}\cdot\frac{1}{6}=\mathrm{P}(X_{12}=1)\mathrm{P}(X_{23}=1)$，而每個變量只取兩個值，所以 X_{12} 與 X_{23} 獨立。（對於只取兩個值的變量，聯合分佈表中的一格便決定了其餘各格；練習 1 會把這一點說清楚。）沒有共同玩家的兩對，例如 X_{12} 與 X_{34}，顯然是獨立的。因此所有 X_{ij} 都兩兩獨立。

階段 2 得分的平均值與方差

得分為 $S=\sum_{i<j}X_{ij}$，是 $\binom{n}{2}$ 個變量之和，每個變量的平均值為 $\frac{1}{6}$，方差為 $\frac{1}{6}\cdot\frac{5}{6}=\frac{5}{36}$。期望值（expectation）總是可以相加的，所以 $\mathrm{E}(S)=\frac{1}{6}\binom{n}{2}$。

> **想一想** 和的方差在甚麼時候等於方差之和？

和的方差等於方差之和，再加上協方差（covariance）之和的兩倍，而獨立的一對變量協方差為零。兩兩獨立已經足夠：

$$\mathrm{E}(S)=\frac{n(n-1)}{12},\qquad \mathrm{Var}(S)=\frac{5}{36}\binom{n}{2}=\frac{5n(n-1)}{72}.$$

階段 3 方差公式

若各 Y_i 的平均值為零，則 $\mathrm{Var}(\sum Y_i)=\mathrm{E}((\sum Y_i)^2)$，把平方展開後，每個 Y_i^2 出現一次，每個乘積 Y_iY_j（其中 $i<j$）出現兩次。取期望值便得到公式。 ■

階段 4 奇偶遊戲

> **想一想** $\mathrm{E}(Z_{12})$ 是多少？$Z_{12}Z_{23}$ 可以取哪些值？它在甚麼時候不等於零？

$Z_{12}=1$ 的概率為 $\frac{3}{36}=\frac{1}{12}$（兩人都擲出 2、都擲出 4 或都擲出 6），而取值 -1 的概率為 $\frac{1}{12}$；所以 $\mathrm{E}(Z_{12})=0$ 而 $\mathrm{E}(Z_{12}^2)=\frac{1}{6}$。

只有當三位玩家都擲出相同點數時，才有 $Z_{12}Z_{23}\neq 0$，而此時 $Z_{12}=Z_{23}$（同為 $+1$ 或同為 -1），所以乘積為 1。因此 $\mathrm{E}(Z_{12}Z_{23})=\frac{1}{36}$，這並不等於 $\mathrm{E}(Z_{12})\mathrm{E}(Z_{23})=0$：它們並不獨立。 ■

（更直接的看法是，$\mathrm{P}(Z_{12}=1, Z_{23}=-1)=0$，而每個事件單獨發生的概率都是 $\frac{1}{12}$。）

階段 5 方差

總得分 $T=\sum Z_{ij}$ 的平均值為 0。根據第 (ii) 部分，$\mathrm{Var}(T)$ 等於 $\binom{n}{2}\cdot\frac{1}{6}$，再加上 $\mathrm{E}(Z_{ij}Z_{kl})$ 對所有「兩對玩家」的組合求和後的兩倍。沒有共同玩家的兩對，貢獻為 0（它們獨立，而且各自的平均值為零）。恰好有一位共同玩家的兩對，每組貢獻 $\frac{1}{36}$，而這樣的組合共

有 $n\binom{n-1}{2}$ 個（先選共同的玩家，再選另外兩人）。所以

$$\begin{aligned}\operatorname{Var}(T) &= \frac{n(n-1)}{12} + 2\cdot\frac{1}{36}\cdot\frac{n(n-1)(n-2)}{2} \\ &= \frac{n(n-1)}{36}\big(3+(n-2)\big) = \frac{n(n^2-1)}{36}. \qquad \blacksquare\end{aligned}$$

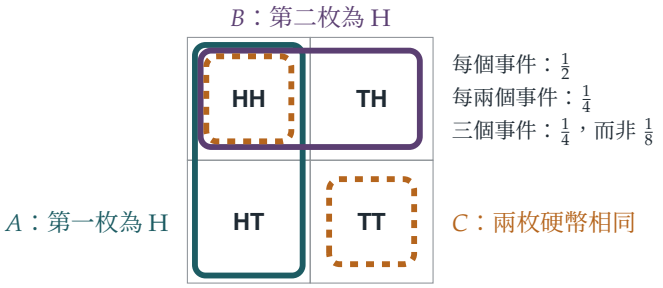


圖 12.4.1 伯恩斯坦（Bernstein）的例子：兩枚公平硬幣。事件「第一枚硬幣是正面」（藍綠色）、「第二枚硬幣是正面」（紫紅色）和「兩枚硬幣結果相同」（琥珀色虛線）的概率各為 $\frac{1}{2}$，而每兩個事件都相交於概率為 $\frac{1}{4}$ 的單一方格，所以它們兩兩獨立。但三個事件全部相交於 HH，其概率為 $\frac{1}{4}$，而非 $\frac{1}{8}$：其中任何兩個事件都決定了第三個。

回顧

- **指示變量（indicator variables）。**把得分寫成指示變量之和，一切便化為對一對對玩家的考慮。
- **方差只需要看成對的變量。**和的方差只取決於兩兩之間的協方差，所以兩兩獨立已足以令方差相加。
- **獨立性要驗證，不要假設。**兩個看似相同的遊戲，卻有不同的相依結構。

背後的數學

多個事件的獨立性

如果對每一個子族，交集的概率都等於各概率之積，則稱事件 $A_1,\dots,A_n$ *獨立*。對三個事件來說，這意味着四個條件：三個是關於每一對的，另一個是 $\mathrm{P}(A\cap B\cap C) = \mathrm{P}(A)\mathrm{P}(B)\mathrm{P}(C)$。伯恩斯坦（Sergei Bernstein）在 1920 年代提出圖 12.4.1 中的例子，說明兩兩之間的條件並不能推出三者的條件。（反過來，單憑三者的條件也不能推出兩兩之間的條件。）題目中「點數相同」的事件 X_{12}、X_{23}、X_{13} 與伯恩斯坦的硬幣如出一轍。

不相關不等於獨立

協方差為零，即 $\mathrm{E}(XY) = \mathrm{E}(X)\mathrm{E}(Y)$，是比獨立更弱的條件，稱為不相關（uncorrelated）。STEP III 2000 Q14（見練習）說明，對於只取兩個值的變量，這兩個概念是一致的，但取三個值時則不然。一個簡單的例子：設 X 在 $\{-1, 0, 1\}$ 上均勻分佈，並設 $Y = X^2$。那麼 $\mathrm{E}(XY) = \mathrm{E}(X^3) = 0 = \mathrm{E}(X)\mathrm{E}(Y)$，然而 Y 完全由 X 決定。協方差只量度線性關聯。

廉價的隨機性

對電腦來說，產生隨機數的成本很高。1979 年，Larry Carter 和 Mark Wegman 證明，對許多演算法而言，例如把數據儲存在偽隨機位置的雜湊（hashing），兩兩獨立的隨機選擇與完全獨立的同樣有效。兩兩獨立的數值可以用極低的成本產生：例如，取隨機的 a 和 b，則 $a + bk \pmod{p}$ 對 $k = 0, 1, \ldots, p-1$ 給出 p 個數值，其中任何兩個都是獨立的。本題的方差計算正是它行得通的原因：切比雪夫（Chebyshev）不等式只需要方差，而方差只需要看成對的變量。

思維技巧

分解為指示變量

要求一個計數的平均值和方差，可把它寫成指示變量之和，每一個可能被數到的對象對應一個指示變量。求平均值只需要考慮單個指示變量；求方差則需要考慮成對的指示變量。

這是概率論中最有用的技巧之一，早在 Section 9.8 已經出現過。在大學裏，隨機圖、演算法和統計檢驗的分析都以它為基礎。

延伸練習

1 **兩個值。**試做 STEP III 2000，第 14 題（`step.maths.org/questions/00-s3-q14`）。該題證明，對於各自只取兩個值的隨機變量，$\mathrm{E}(XY) = \mathrm{E}(X)\mathrm{E}(Y)$ 必然推出獨立，並要求給出一個取三個值的反例。

2 **人數更多的隊伍。**在第 (i) 部分中取 $n = 7$，求得分的平均值和標準差，並把例子中的得分 7 與它們比較。

3 **再談生日。**利用指示變量，求 k 個人當中生日相同的人數對的平均值和方差（Section 9.9）。這些指示變量是否兩兩獨立？

4 **三者成立但兩兩不獨立。**找出三個事件，每個的概率都是 $\frac{1}{2}$，滿足 $\mathrm{P}(A \cap B \cap C) = \frac{1}{8}$，但並非兩兩獨立。

延伸閱讀

- **Geoffrey Grimmett & David Stirzaker, *Probability and Random Processes***（Oxford University Press），開首幾章討論獨立性。
- **Michael Mitzenmacher & Eli Upfal, *Probability and Computing***（Cambridge University Press）。兩兩獨立及其在演算法中的應用。
- **Joseph Blitzstein & Jessica Hwang, *Introduction to Probability***（CRC Press；可於網上免費閱讀），關於指示變量與協方差。

12.5 最後一位乘客的座位

題目

STEP II 2021，第 11 題

繼續閱讀之前，請先看題目（原題為英文）。題目載於 OCR 提供下載的 2021 年 STEP 試卷（zip 檔案）內的 STEP II 試卷：

`www.ocr.org.uk/Images/696588-2021-past-papers.zip`

引子

一百名乘客登上一班滿座的飛機。第一位乘客遺失了登機證，於是隨便找個座位坐下。其後每位乘客，若自己的座位仍然空着就坐自己的座位，否則就在餘下的空位中隨機選一個。最後一位乘客坐到自己座位的概率（probability）是多少？混亂或會蔓延至整架飛機，所以你可能預期這個概率很小；又或者大部分乘客根本不受影響，概率應該接近 1。

無論乘客人數是多少，答案都恰好是 $\frac{1}{2}$。本題以遞推關係（recurrence）和歸納法（induction）證明這個結果。不過，還有一個利用對稱性（symmetry）的漂亮論證，一句話就能說完；它解釋了為甚麼答案是 $\frac{1}{2}$，更可讓你立即寫出任何一位乘客的答案。

動筆之前

停一停，試一試 列出所有可能發生的情況，求出 P_2 和 P_3。

若有兩個座位，T_1 會坐 S_1（這時 T_2 坐到 S_2）或 S_2，兩者的概率各為 $\frac{1}{2}$：$P_2 = \frac{1}{2}$。若有三個座位：如果 T_1 坐 S_1（概率為 $\frac{1}{3}$），人人都沒有問題；如果 T_1 坐 S_3，T_3 就坐不到自己的座位；如果 T_1 坐 S_2，則 T_2 要在 S_1 和 S_3 之間隨機選擇，而 T_3 只有在 T_2 選中 S_1 時才沒有問題。所以 $P_3 = \frac{1}{3} + \frac{1}{3} \cdot \frac{1}{2} = \frac{1}{2}$。

逐步解題

階段 1 遞推關係

想一想 假設 T_1 坐了 S_k，其中 $2 \le k \le n-1$。乘客 $T_2, \dots, T_{k-1}$ 會怎樣？當 T_k 到達時，哪些座位是空的？這個局面與一列較小的列車相比又如何？

乘客 $T_2, \dots, T_{k-1}$ 發現自己的座位空着，便坐了下來。當 T_k 到達時，空着的座位是 S_1 和 $S_{k+1}, \dots, S_n$：共有 $n-k+1$ 個。T_k 從中隨機選一個，此後每位乘客若自己的座位空着就坐自己的座位，否則就隨機選一個。這正是有 $n-k+1$ 個座位的原來問題，T_k 扮演第一位乘客的角色，而 S_1 則扮演 T_k「自己的」座位的角色（即一旦被選中，一切便回復正常的那個座位）。因此，T_n 坐到 S_n 的條件概率（conditional probability）是 P_{n-k+1}。

現在按 T_1 所坐的座位分情況考慮，每個座位的概率都是 $\frac{1}{n}$：坐 S_1 必定成功，坐 S_n 必定失敗，而坐 S_k 則給出 P_{n-k+1}。所以

$$P_n = \frac{1}{n}\Big(1 + 0 + \sum_{k=2}^{n-1} P_{n-k+1}\Big) = \frac{1}{n}\Big(1 + \sum_{r=2}^{n-1} P_r\Big). \qquad \blacksquare$$

階段 2 答案

想一想 $P_2 = P_3 = \frac{1}{2}$。猜出 P_n，並用強歸納法（strong induction）證明。

斷言：對所有 $n \ge 2$，$P_n = \frac{1}{2}$。當 $n = 2$ 時斷言成立。若 $P_r = \frac{1}{2}$ 對 $2 \le r \le n-1$ 成立，則

$$P_n = \frac{1}{n}\Big(1 + \frac{n-2}{2}\Big) = \frac{1}{n} \cdot \frac{n}{2} = \frac{1}{2}. \qquad \blacksquare$$

階段 3 倒數第二位乘客

想一想 試找一個更巧妙的論證。到最後，留給 T_n 的是哪個座位？有可能是 S_2 嗎？

最後留給 T_n 的座位不是 S_1 就是 S_n。它不可能是 S_k（其中 $2 \le k \le n-1$），因為 T_k 早就會把它坐了。而在此之前的每一次隨機選擇，對 S_1 和 S_n 都是對稱的：每當有被佔了座位的乘客要選位，S_1 和 S_n 被選中的機會均等。兩者之中先被選中的那個決定了結局，所以兩者同樣可能成為最後剩下的座位：$P_n = \frac{1}{2}$。

同一個論證適用於任何一位乘客 T_k，$k \ge 2$。只有當某位較早的乘客佔了 T_k 的座位，他們才會失去自己的座位。座位被佔的連鎖反應會一直持續，直至有人選中 S_1 為止；在 $S_1, S_k, S_{k+1}, \dots, S_n$ 這些座位之中，T_k 到達之前的選位者對它們一視同仁，所以最先被

選中的座位同樣可能是這 $n-k+2$ 個座位中的任何一個，而 T_k 只有在最先被選中的是 S_k 時才會失去座位。所以

$$\mathrm{P}(T_k \text{ 坐在} S_k) = \frac{n-k+1}{n-k+2}.$$

當 $k = n-1$（其中 $n \geq 3$）時，這給出 $Q_n = \frac{2}{3}$；而 $Q_2 = \frac{1}{2}$，因為這時 $T_{n-1} = T_1$，他們坐在 S_1 的概率是 $\frac{1}{2}$。 ■

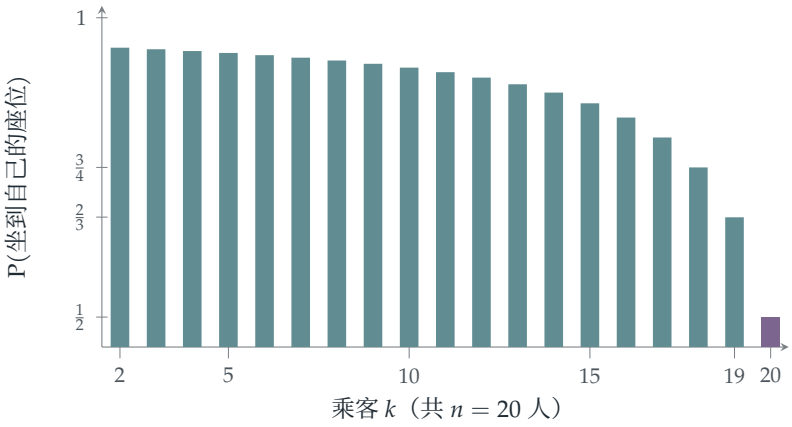


圖 12.5.1 乘客 k 坐到自己座位的概率，列車共有 $n = 20$ 個座位。當 $2 \leq k \leq n$ 時，概率是 $\frac{n-k+1}{n-k+2}$：由 T_2 的 $\frac{19}{20}$ 一直下降至 T_{19} 的 $\frac{2}{3}$，以及最後一位乘客的 $\frac{1}{2}$（紫紅色）。

回顧

- **化為較小的同類問題。**在 T_1 坐了 S_k 之後，餘下的問題就是座位較少的同一個問題。
- **強歸納法。**遞推關係牽涉所有較小的情況，所以歸納時要假設結果對所有這些情況都成立。
- **找出對稱性。**座位 S_1 和 S_n 自始至終可以互換，這樣不用任何代數運算就解釋了 $\frac{1}{2}$ 的由來。

背後的數學

對稱論證

第二個證明體現了概率論中一個威力強大的想法：若兩個結果可由隨機過程的某種對稱性互相對應，它們的概率就相同。這裏的過程在 S_1 和 S_n 其中一個被選中之前，並不區分這兩個座位，所以兩者成為最後剩下的座位的可能性必定相等。對稱論證往往只憑一

個觀察，就能取代冗長的計算。接下來的幾節有很多這樣的例子：Section 12.6 中多擲一次硬幣，Section 12.7 中的淘汰賽抽籤。

有多少位乘客被迫換位？

被迫換位的乘客（即坐不到自己座位的乘客）的數目同樣有很好的規律。T_1 除非選中 S_1，否則就算是換了位；而 T_k（其中 $k \geq 2$）被迫換位的概率是 $\frac{1}{n-k+2}$。把這些概率相加，被迫換位的乘客數目的期望值（expected number）是

$$\left(1-\frac{1}{n}\right)+\left(\frac{1}{n}+\frac{1}{n-1}+\cdots+\frac{1}{2}\right)=H_n-\frac{1}{n},$$

大約是 $\ln n$：即使飛機上有 300 個座位，這場混亂平均也只影響約六位乘客。調和數（harmonic numbers）又再出現了（Section 9.11）。

思維技巧

尋找對稱性

如果隨機過程無法分辨兩個結果，它們的可能性就相等。找出它們最初變得可以分辨的那一刻。

在大學，對稱論證在概率論中隨處可見，由反射原理（reflection principle）（Section 9.10）到可交換性（exchangeability）皆是；可交換性是指一列隨機變量（random variables）無論以任何次序排列都有相同的分佈，這個性質是不少現代統計學的基礎。

延伸練習

1 **兩位健忘的乘客。**如果最先的兩位乘客都隨機選位，最後一位乘客坐到自己座位的概率是多少？

2 **被迫換位的乘客。**列出所有可能情況，驗證 $n = 3$ 時被迫換位的乘客數目的期望值。

3 **迂迴的路線。**利用遞推關係，證明 $\sum_{r=2}^{n-1} P_r = \frac{n-2}{2}$ 與對稱論證互相吻合。

4 **最後一個座位。**證明在最後一位乘客坐不到自己座位的條件下，最後一位乘客坐到 S_1 的概率是 1。為甚麼？

延伸閱讀

- **Peter Winkler, *Mathematical Puzzles: A Connoisseur's Collection***（A K Peters）。收錄了飛機座位謎題，以及許多其他可用對稱性解決的謎題。
- **Frederick Mosteller, *Fifty Challenging Problems in Probability***（Dover）。一本經典的題集。
- **Joseph Blitzstein & Jessica Hwang, *Introduction to Probability***（CRC Press；可於網上免費閱覽）。

12.6 多一枚硬幣

題目

STEP I 2016，第 12 題

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

`step.maths.org/questions/16-s1-q12`

引子

Bob 比 Alice 多擲一枚硬幣。他擲得的正面（heads）比 Alice 多的概率（probability）是多少？他佔了優勢，但多擲一次不保證多一個正面，而且平手算他輸。用硬算的方法做出頭兩個情況，兩次的答案都是 $\frac{1}{2}$。這是巧合嗎？

不是巧合，而且有一個只需一句話的證明：只要留意到 Bob 擲得的正面或反面（tails）必有其一比 Alice 多，但不會兩者皆多。這道題目是一個典範，示範了計算如何提示出規律，而巧妙的對稱性（symmetry）又如何解釋這個規律。

動筆之前

? **想一想** Alice 擲兩次：她擲得正面的次數的分佈（distribution）是甚麼？擲三次的 Bob 又如何？

Alice 擲得 $0, 1, 2$ 個正面的概率分別是 $\frac{1}{4}, \frac{2}{4}, \frac{1}{4}$；Bob 擲得 $0, 1, 2, 3$ 個正面的概率分別是 $\frac{1}{8}, \frac{3}{8}, \frac{3}{8}, \frac{1}{8}$。

逐步解題

階段 1 硬算

第 (i) 部分，按 Alice 擲得的正面數目逐項相加：

$$\begin{aligned}\mathrm{P}(B > A) &= \tfrac{1}{4}\mathrm{P}(B \geq 1) + \tfrac{2}{4}\mathrm{P}(B \geq 2) + \tfrac{1}{4}\mathrm{P}(B = 3) \\ &= \tfrac{1}{4}\cdot\tfrac{7}{8} + \tfrac{1}{2}\cdot\tfrac{4}{8} + \tfrac{1}{4}\cdot\tfrac{1}{8} = \frac{7+8+1}{32} = \frac{1}{2}.\end{aligned}$$

第 (ii) 部分是同樣的計算，只是項數較多；數一數 $8 \times 16 = 128$ 個等可能的結果，其中 64 個是 Bob 領先（圖 12.6.1），所以答案同樣是 $\frac{1}{2}$。

階段 2 拆開 Bob 的投擲

想一想 把 Bob 的 $n+1$ 次投擲看成先擲 n 次，再額外擲一次。擲完頭 n 次之後，他可能領先 Alice、與她平手，或者落後。每種情況下會發生甚麼事？

二人各擲 n 次之後，Bob 領先的概率是 p_2，平手的概率是 p_1，落後的概率是 p_2（由 Alice 與 Bob 之間的對稱性可知）。如果他領先，他會保持領先。如果平手，他額外的一擲有 $\frac{1}{2}$ 的概率令他領先。如果他落後，多一個正面最多只能追成平手。所以

$$\mathrm{P}(\text{Bob 的正面較多}) = p_2 + \tfrac{1}{2}p_1.$$

由於這三種情況涵蓋了所有可能，$2p_2 + p_1 = 1$，因此 $p_2 + \frac{1}{2}p_1 = \frac{1}{2}$。多了一枚硬幣，Bob 擲得較多正面的概率對每個 n 都剛好是 $\frac{1}{2}$。 ■

停一停，試一試 以下就是那個一句話的證明。Bob 擲 $n+1$ 次，所以他的正面和反面數目加起來是 $n+1$，比 Alice 的多一。試證明 Bob 的正面比 Alice 多，當且僅當他的反面不比 Alice 多。由於正面與反面是對稱的，這兩個事件的概率相等，所以各為 $\frac{1}{2}$。

事實上，記 $H_A + T_A = n$ 及 $H_B + T_B = n+1$：若 $H_B > H_A$，則 $H_B \geq H_A + 1$，所以 $T_B = n + 1 - H_B \leq n - H_A = T_A$；反過來，若 $T_B \leq T_A$，則 $H_B \geq H_A + 1$。所以「正面較多」與「反面較多」是互補（complementary）的，而把正面與反面對調（由於硬幣是公平的，這不會改變概率）就會把兩者互換。

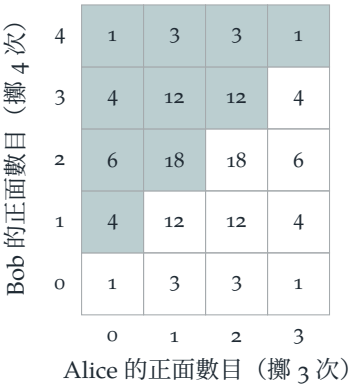


圖 12.6.1 第 (ii) 部分：在 128 個等可能的結果之中，每種正面數目組合所佔的個數。Bob 在有底色的格子中領先，這些格子剛好包含一半的結果。

回顧

- **計算小的情況。**兩次計算都提示答案是 $\frac{1}{2}$。
- **按自然的劃分來取條件。**把 Bob 額外的一擲分開，問題便化為一場勢均力敵的比試再加一枚硬幣。
- **找出互補的一對。**「正面較多」與「反面較多」互為補事件，並且在一個對稱之下互換。

背後的數學

對稱與補事件

那個漂亮的證明結合了兩個事實：在一個對稱之下互換的兩個事件概率相等，而兩個互補事件的概率之和為 1。若兩者同時成立，則每個事件的概率都是 $\frac{1}{2}$。同樣的推理在概率論中隨處可見。在航班座位問題 (Section 12.5) 中，S_1 與 S_n 在一個對稱之下互換，而最後剩下的恰好是其中一個。在六貫棋 (Hex) 中，由棋盤的對稱性，加上恰有一方必勝這個事實，可知先手一方佔優，不過在大棋盤上無人知道必勝策略是甚麼。

平手的可能性有多大？

當二人投擲次數相同時，平手的概率 p_1 就很重要。Alice 與 Bob 平手，即 Alice 的正面數目等於 Bob 的正面數目；利用對稱性，把 Bob 的正面換成他的反面，這個概率就等於 Alice 的正面數目加上 Bob 的正面數目等於 n 的概率，亦即在 $2n$ 次投擲中共有 n 個正

面。所以

$$p_1 = \binom{2n}{n}\frac{1}{4^n} \approx \frac{1}{\sqrt{\pi n}},$$

這裏用了斯特靈近似（Stirling's approximation）（Section 12.23）。平手會變得罕見，但變得很慢：各擲 100 次，平手的機會仍然約有 5.6%。

思維技巧

先計算，再猜想，然後解釋

當計算得出一個簡單得可疑的答案時，就要找出原因。簡單的答案通常有簡單的解釋。

在大學裏，這個習慣正是理解與計算的分野。這裏的第一個證明確認了答案；第二個證明則解釋了答案，而且可以立即推廣（試做練習 2）。

延伸練習

1 **多擲兩枚硬幣。**若 Bob 擲 $n+2$ 枚硬幣而 Alice 擲 n 枚，證明 $\mathrm{P}(\text{Bob 的正面較多}) = \frac{1}{2} + \frac{1}{2}\mathrm{P}(H' = H_A)$，其中 H' 是 Bob 頭 $n+1$ 次投擲中的正面數目。

2 **不公平的硬幣。**若二人所用的硬幣擲出正面的概率都是 p，多擲一次時答案是否仍然是 $\frac{1}{2}$？那個漂亮的證明中哪一步不再成立？

3 **平手。**用正文中的技巧證明 $p_1 = \binom{2n}{n}/4^n$，並求它在 $n = 1, 2, 3$ 時的值。

4 **骰子。**Alice 擲 n 粒骰子，Bob 擲 $n+1$ 粒。Bob 擲得較多六點的概率是否等於 $\frac{1}{2}$？

延伸閱讀

- **Peter Winkler, *Mathematical Mind-Benders***（A K Peters）。收錄大量可用對稱性解決的概率謎題。
- **Gabor Székely, *Paradoxes in Probability Theory and Mathematical Statistics***（Reidel）。
- **Frederick Mosteller, *Fifty Challenging Problems in Probability***（Dover）。

12.7 淘汰賽中誰會遇上誰

題目

STEP II Specimen paper，第 16 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/spec-s2-q16`

引子

溫布頓網球錦標賽（Wimbledon）有 128 名選手參加抽籤。隨機選出其中兩人：他們在賽事中某個階段相遇的機會有多大？你也許會嘗試把每一輪相遇的概率（probability）加起來，而每一項都是「雙方晉級」與「被抽在一起」這些機會的乘積。這樣做是可行的，但有一個好得多的方法。數一數比賽的場數：每場比賽都由一對選手對賽，而比賽恰好有 127 場。根據對稱性（symmetry），每一對選手成為其中任何一場的對賽雙方的可能性都相同。因此所求的機會就是 127 除以選手配對的總數。

這道題是「數對了東西，概率問題便迎刃而解」的絕佳例子。答案 $\frac{1}{2^{n-1}}$ 只需一行便可得出。

動筆之前

想一想 有 2^n 名選手的淘汰賽（knockout tournament）共要進行多少場比賽？（每場比賽淘汰一名選手。）

每場比賽恰好淘汰一名選手，而除了一人之外所有選手都被淘汰，所以共有 $2^n - 1$ 場比賽。

逐步解題

階段 1 第一輪

把選出的兩名選手稱為 X 和 Y。在第一輪，X 的對手是其餘 $2^n - 1$ 名選手中任何一人的可能性都相同，所以

$$\mathrm{P}(\text{相遇於輪次1}) = \frac{1}{2^n - 1}.$$

階段 2 關鍵的對稱性

想一想 每一輪的抽籤都是隨機的，而每場比賽都相當於擲一次公平的硬幣。這個過程對待某一對選手的方式，會否與對待另一對有所不同？

不會。由於抽籤是隨機的，而且每場比賽的勝負機會都是 50–50，把選手重新標號並不會改變任何概率。所以任何一場指定的比賽（例如決賽）由 $\binom{2^n}{2}$ 對選手中任何一對對賽的可能性都相同。

階段 3 決賽

決賽是一場特定的比賽，所以它由 X 和 Y 對賽的概率是

$$\frac{1}{\binom{2^n}{2}} = \frac{2}{2^n(2^n - 1)}.$$

（驗算：兩人都必須打進決賽，這需要他們分處不同的半區，並且各自勝出 $n - 1$ 場比賽。）

階段 4 在任何階段

想一想 X 和 Y 最多只能相遇一次。這一點怎樣讓你可以把各場比賽的概率相加？

設 I_m 在第 m 場比賽由 X 和 Y 對賽時取值 1。在 $2^n - 1$ 場比賽中，每一場都有 $\mathrm{P}(I_m = 1) = \binom{2^n}{2}^{-1}$，而這些事件是互斥（mutually exclusive）的（他們不可能相遇兩次），所以

$$\mathrm{P}(\text{相遇}) = \frac{2^n - 1}{\binom{2^n}{2}} = \frac{2(2^n - 1)}{2^n(2^n - 1)} = \frac{1}{2^{n-1}}. \qquad \blacksquare$$

以溫布頓為例，$n = 7$，兩名指定選手相遇的概率是 $\frac{1}{64}$。作為驗算，在 $2^n - 1$ 場比賽中，第一輪佔了 2^{n-1} 場，由此得出 $\frac{2^{n-1}}{\binom{2^n}{2}} = \frac{1}{2^n - 1}$，與第 1 階段的結果一致。

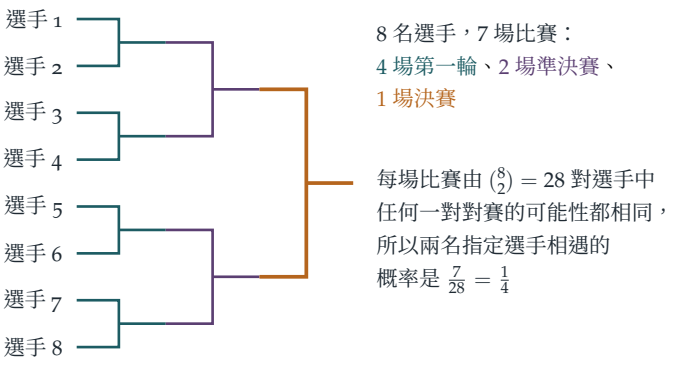


圖 12.7.1 $2^3 = 8$ 名選手的淘汰賽（圖中畫成固定的賽程表，但題目中每一輪都會重新隨機抽籤）。比賽共有 7 場。根據對稱性，每一場由 28 對選手中任何一對對賽的可能性都相同，所以選出的兩名選手相遇的概率是 $\frac{7}{28} = \frac{1}{4} = \frac{1}{2^{n-1}}$。

回顧

- **數比賽的場數。**比賽的場數 $2^n - 1$ 是關鍵的事實。
- **各對選手之間的對稱性。**每一對選手成為任何一場指定比賽的對賽雙方的可能性都相同。
- **對各場比賽的線性性質。**由於兩名選手最多相遇一次，相遇的概率就是相遇次數的期望值，也就是對各場比賽求和。

背後的數學

期望次數

第 4 階段的論證其實是一個期望值（expectation）的計算：X 和 Y 相遇的次數是若干個指示變量（indicator）之和，每場比賽對應一個；而它的期望值（由於它只取 0 或 1，所以等於相遇的概率）就是各個指示變量的期望值之和。這就是期望值的線性性質（linearity of expectation），在 Section 9.11 和 Section 12.4 中都曾用到。它最大的好處是完全不必理會各輪之間錯綜複雜的相依關係。

設有種子的賽事

現實中的賽事設有種子（seeded）：最強的選手會被編排在適當的位置，使他們要到後段輪次才可能相遇。而且現實中的比賽也不是 50–50 的。這兩點都破壞了對稱性，於是計算兩名指定選手相遇的機會便成了一項貨真價實的運算。為體育運動建立模型的統計

學家正是運用這類計算來預測賽事，當中每場比賽的概率都是根據選手的評分估計出來的。對稱的情況則是一條基線，用來衡量各種現實因素的影響。

思維技巧

數對的東西

只要選對了對象，概率往往可以用（有利對象的數目）÷（等可能對象的數目）求得。在這裏，對象就是比賽和選手配對。

選擇數甚麼，正是組合概率（combinatorial probability）的藝術。在大學裏，同一個想法會以算兩次（double counting）的形式出現（Section 7.9），也會出現在概率方法（probabilistic method）之中：在那裏，某種結構的期望數目可以用來證明它存在。

延伸練習

1 **逐輪計算。**直接證明 X 和 Y 在第 k 輪相遇的概率是 $\frac{2^{n-k}}{\binom{2^n}{2}}$，並驗證這些概率加起來等於 $\frac{1}{2^{n-1}}$。

2 **固定的賽程表。**假如賽程表改為預先固定（如圖所示），但所有選手都被隨機編入其中，答案會否改變？

3 **一名強手。**假設選手 X 每場比賽都勝出。現在 X 與 Y 相遇的概率是多少？

4 **聯賽。**在一個有 $2n$ 支球隊的聯賽中，每兩隊都對賽一次，共有多少場比賽？相應的「相遇概率」是甚麼？

延伸閱讀

- **Frederick Mosteller, *Fifty Challenging Problems in Probability***（Dover），書中收錄了有關淘汰賽的問題。
- **David Spiegelhalter, *The Art of Statistics***（Pelican），講述建立模型與預測。
- **Joseph Blitzstein & Jessica Hwang, *Introduction to Probability***（CRC Press；可於網上免費閱讀），講述指示變量與對稱性。

12.8 Penney 遊戲

題目

STEP III 2001，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/01-s3-q13`

引子

你和朋友各自選定一個由三次擲硬幣結果組成的序列，然後不斷擲硬幣，直至其中一個序列出現為止。在任何一個時刻，每個三次投擲的序列出現的機會都相同，那麼無論選哪一個，理應同樣好吧？完全不是。無論朋友選哪個序列，你都可以選出一個勝過它的序列，勝算（odds）有時更高達 7 比 1。

這就是 *Penney 遊戲*（Penney's game），由 Walter Penney 於 1969 年發明。本題要證明 A 勝 B，B 勝 C，C 勝 D，而 D 又勝 A：這是一個循環，就像包剪揼一樣。「勝過」這種關係並不具傳遞性（transitive）。長遠而言，後選的一方總能勝出。

動筆之前

想一想 假設至今的投擲結果以 TT 結尾，而且仍未有人勝出。THH 有沒有可能比 TTH 先出現？

從 TT 出發，之後每擲出一個 T，最後兩次投擲仍然是 TT，而第一個 H 一出現就湊成 TTH。要 THH 出現，必須先有 TH，而 TH 只能由一個 T 後接一個 H 而來，但這個 H 已經湊成了 TTH。所以 TT 一旦出現，A 必勝。

逐步解題

階段 1 狀態

想一想 兩個序列都不可能在兩次投擲內完成。擲了兩次之後，從每個狀態（state）出發，下一次投擲會帶來甚麼結果？

由於遊戲只取決於最後兩次投擲，我們可以追蹤狀態（即最後兩次投擲的結果）：

- 從 TT 出發：A 勝（上面已證），所以 $p_{\mathrm{TT}} = 1$。
- 從 HT 出發：下一次投擲會得出 TH 或 TT，所以 $p_{\mathrm{HT}} = \frac{1}{2}p_{\mathrm{TH}} + \frac{1}{2}p_{\mathrm{TT}}$。
- 從 TH 出發：擲出 H 便湊成 THH（B 勝）；擲出 T 則得出 HT。所以 $p_{\mathrm{TH}} = \frac{1}{2}\cdot 0 + \frac{1}{2}p_{\mathrm{HT}}$。
- 從 HH 出發：擲出 H 仍是 HH；擲出 T 則得出 HT。所以 $p_{\mathrm{HH}} = \frac{1}{2}p_{\mathrm{HH}} + \frac{1}{2}p_{\mathrm{HT}}$，由此得 $p_{\mathrm{HH}} = p_{\mathrm{HT}}$。

階段 2 求解

把 $p_{\mathrm{TH}} = \frac{1}{2}p_{\mathrm{HT}}$ 代入 HT 的方程：$p_{\mathrm{HT}} = \frac{1}{4}p_{\mathrm{HT}} + \frac{1}{2}$，所以 $p_{\mathrm{HT}} = \frac{2}{3}$、$p_{\mathrm{HH}} = \frac{2}{3}$ 及 $p_{\mathrm{TH}} = \frac{1}{3}$。四個起始狀態的概率（probability）各為 $\frac{1}{4}$，所以

$$\mathrm{P}(A \text{ 勝}) = \tfrac{1}{4}\left(\tfrac{2}{3} + \tfrac{2}{3} + \tfrac{1}{3} + 1\right) = \tfrac{2}{3}. \qquad \blacksquare$$

階段 3 B 對 C

想一想 C 想要 HHT，B 想要 THH。C 在甚麼情況下勝出？（留意第一個 HH 之前緊接着的必定是甚麼。）

如果頭兩次投擲是 HH，那麼其後第一個 T 便湊成 HHT，THH 來不及出現：C 勝。否則，在 HH 首次出現之前的某處必有一個 T，因此第一個 HH 的前一次投擲必定是 T：THH 就在那一刻出現，早於 HHT（HHT 還需要在 HH 之後再擲出一個 T）。所以只有當遊戲以 HH 開始時 C 才會勝出：

$$\mathrm{P}(B \text{ 勝}) = \tfrac{3}{4}. \qquad \blacksquare$$

階段 4 循環

用同樣的方法（或「背後的數學」一節的捷徑）可以證明 C（HHT）勝 D（HTT）的概率是 $\frac{2}{3}$，而 D（HTT）勝 A（TTH）的概率是 $\frac{3}{4}$；後者正是第 3 階段把正面與反面對調後的鏡像。所以

$$A \text{ 勝} B, \qquad B \text{ 勝} C, \qquad C \text{ 勝} D, \qquad D \text{ 勝} A. \qquad \blacksquare$$

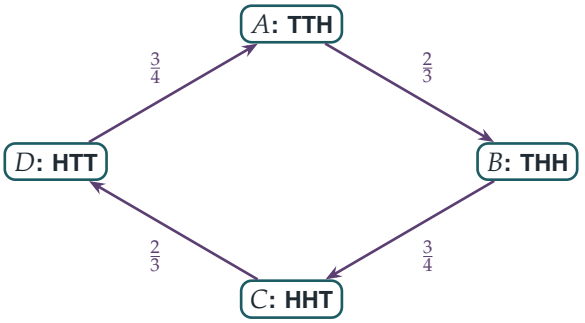


圖 12.8.1 Penney 循環。由 X 指向 Y 的箭頭表示 X 的序列勝過 Y 的序列，箭頭旁為相應的概率。每個序列都被另一個序列擊敗，所以後選的一方總能挑出勝算較高的序列。

回顧

- **選對狀態。**只有最後兩次投擲才有影響，所以四個狀態、四條方程便概括了整個遊戲。
- **首步分析（first-step analysis）。**每條方程都是以下一次投擲為條件：從某狀態出發的概率，就是下一次投擲所到達的各個狀態的概率的平均值。
- **找出決定性的事件。**在第 3 階段，「第一個 HH 之前是 T」這一點無須列方程便決定了勝負。

背後的數學

為甚麼有些序列較好

在任何一個位置，所有三次投擲的序列出現的機會都相同，而每個序列的期望等待時間（expected waiting time）也不難計算。但各序列的等待時間並不相同：TTH 平均要等 8 次投擲，TTT 卻要等 14 次，因為與自身重疊的序列（例如 TTT，只差一步的 TTH 會令已累積的兩次投擲前功盡廢）出現得較慢。而在兩個序列的競賽中，關鍵在於每個序列與對方如何重疊。THH 勝 HHT，是因為 HHT 每次出現（除了在最開頭之外），前面都有一個 T，這表示 THH 剛剛已經出現。後選一方的竅門總是一樣：取對手序列的頭兩個符號，再在前面加上與對手中間符號相反的那個符號。

康威公式（Conway's formula）

約翰·康威發現了一條計算勝算的絕妙公式。對於長度為 n 的兩個序列 X 和 Y，定義相關數（correlation）XY 為一個 n 位的二進制數：由左至右，若 X 中由第 i 個符號開始的

部分與 Y 開頭相同長度的部分相同，則第 i 位為 1。那麼 Y 比 X 先出現的勝算是

$$(XX - XY) : (YY - YX).$$

取 X = HHT 及 Y = THH：$XX = 100_2 = 4$（只有整個 HHT 與自身吻合）；$XY = 001_2 = 1$（只有 HHT 末尾的 T 與 THH 的開頭吻合）；$YY = 100_2 = 4$；而 $YX = 011_2 = 3$（THH 末尾的 HH 和 H 都與 HHT 的開頭吻合）。所以 THH 先出現的勝算是 $(4-1):(4-3) = 3:1$，即第 3 階段的 $\frac{3}{4}$。康威公式適用於任何長度、任何字母表的序列，其證明用到鞅（martingale）的理論：想像有一間賭場，賭客就序列的出現下注，再利用公平遊戲的期望利潤為零這一事實。

Penney 與加德納（Gardner）

Walter Penney 於 1969 年在 *Journal of Recreational Mathematics* 的一篇短文中介紹了這個遊戲。馬丁·加德納於 1974 年在他的 *Scientific American* 專欄中介紹了它，使它廣為人知，文中還包括康威公式，這條公式是康威發現的，但他一直沒有發表。這類非傳遞關係（埃弗龍骰子（Efron's dice）是另一個例子：骰子 A 勝 B，B 勝 C，C 勝 D，而 D 勝 A）一再令人驚訝，在投票理論中也很重要，因為多數人的偏好可以形成循環：這就是孔多塞悖論（Condorcet's paradox）。

思維技巧

首步分析

若一個過程在各狀態之間轉移，要求出與它有關的概率或期望值（expectation），可以第一步爲條件。這樣每個狀態都給出一條線性方程。

這是處理馬可夫鏈（Markov chain）的基本技巧，你會在 Section 12.16 以及大學概率課程中一再遇到。同樣的方法可以求出先到達某一狀態而非另一狀態的概率、被吸收前的期望時間，以及賭博問題中的破產概率。

延伸練習

1 **四人同場。**試做 STEP II 2015，第 12 題（`step.maths.org/questions/15-s2-q12`）：四名玩家同時參與的 Penney 遊戲。

2 **等待時間。**證明直至 HH 首次出現的期望投擲次數是 6，而直至 HT 首次出現的則是 4。（使用首步分析。）

3 **最佳對策。**對於八個長度為三的序列，逐一找出以最大概率勝過它的序列。哪個序列最容易被擊敗？

4 **康威公式。**用康威公式驗證 HHT 勝 HTT 的概率是 $\frac{2}{3}$。

延伸閱讀

- **Martin Gardner, *Time Travel and Other Mathematical Bewilderments***（W. H. Freeman）。書中包括 Penney 遊戲和康威公式。
- **Ronald Graham, Donald Knuth & Oren Patashnik, *Concrete Mathematics***（Addison-Wesley），第 8.4 節。以生成函數（generating function）處理 Penney 遊戲。
- **David Williams, *Probability with Martingales***（Cambridge University Press）。以鞅給出的證明，留待日後再讀。

12.9 用不公平的硬幣作公平的決定

題目

STEP I 2019，第 11 題

繼續閱讀之前，請先看題目（原題為英文）。題目載於 OCR 提供下載的 2019 年 STEP 試卷（zip 檔案）內的 STEP I 試卷：

`www.ocr.org.uk/Images/696590-2019-past-papers.zip`

引子

你要作出一個公平（fair）的決定，但手上唯一的硬幣是彎的，而且你不知道它偏差（bias）有多大。這樣還能公平地作出選擇嗎？馮紐曼（John von Neumann）在 1951 年給出的答案是：擲兩次。若結果是 HT，就選酒店；若是 TH，就選茶室；若是 HH 或 TT，就再擲兩次。由於 HT 和 TH 的概率（probability）都是 pq，所以不論 p 是多少，這個決定都完全公平。

本題的程序剛好相反：它在出現 HH 和 TT 時作決定，這樣做很快，卻不公平（選酒店的概率是 $\frac{p^2}{p^2+q^2}$）。本題要證明的是，不論硬幣如何，這個程序都不會拖得很久。比較這兩個程序，可以看到產生隨機性時的一個基本取捨（trade-off）：速度與公平之間的取捨。

動筆之前

想一想 在一輪之中，作出決定的概率是多少？沒有決定的概率又是多少？$2pq$ 最大可以取甚麼值？

若兩枚硬幣結果相同，便作出決定，概率是 p^2+q^2；否則他們繼續下去，概率是 $2pq$。由於 $pq = p(1-p) \le \frac{1}{4}$（當 $p = \frac{1}{2}$ 時等號成立），所以 $2pq \le \frac{1}{2}$。

逐步解題

階段 1 在第 n 輪作出決定

各輪之間互相獨立（independent）。要在第 n 輪作出決定，須先有 $n-1$ 輪結果不同，然後一輪結果相同：

$$\mathrm{P}(\text{作出決定的輪次為}n) = (2pq)^{n-1}(p^2+q^2). \qquad \blacksquare$$

階段 2 在第 n 輪或之前作出決定

想一想 在首 n 輪中沒有作出決定的概率是多少？

n 輪之內沒有決定的概率是 $(2pq)^n \le \left(\frac{1}{2}\right)^n$，所以不論 p 是多少，在第 n 輪或之前作出決定的概率至少是 $1-\frac{1}{2^n}$。 $\blacksquare$

階段 3 三個人

想一想 第一輪出現兩個正面和一個反面之後，第三個人把反面翻成正面。第二輪要發生甚麼事才會有決定？如果第一輪是兩個反面和一個正面呢？

第一輪作出決定的概率是 p^3+q^3。否則，有兩枚硬幣結果相同，另一枚不同。若多數是正面（概率為 $3p^2q$），那枚不同的硬幣會被翻成正面，而第二輪要作出決定，就需要另外兩人都擲出正面：概率是 p^2（此時已不可能全是反面）。同理，若多數是反面（概率為 $3pq^2$），要作出決定就需要兩個反面：q^2。所以

$$\mathrm{P}(\text{作出決定的輪次不遲於}2) = p^3+q^3+3p^2q\cdot p^2+3pq^2\cdot q^2.$$

想一想 把所有東西用 $s=pq$ 表示。（利用 $p+q=1$。）

我們有 $p^3+q^3 = (p+q)^3-3pq(p+q) = 1-3s$，以及 $3p^4q+3pq^4 = 3s(p^3+q^3) = 3s(1-3s)$。所以

$$\mathrm{P} = 1-3s+3s-9s^2 = 1-9s^2 \ge 1-\frac{9}{16} = \frac{7}{16},$$

因為 $s \le \frac{1}{4}$。 $\blacksquare$

回顧

- **補事件（complement）。**「在第 n 輪或之前至少發生一次」最容易的算法是 $1-$（「n 輪之內從未發生」）。
- **為最壞情況定界。**未知的 p 是靠不等式（inequality）$pq \le \frac{1}{4}$ 來處理的。
- **對稱函數（symmetric functions）。**對 p 和 q 對稱的式子，用 $p+q=1$ 和 pq 來表示時便會簡化。

背後的數學

馮紐曼的妙計

題目中的程序很快，但有偏差。馮紐曼的程序則是公平的：擲兩次，出現 HT 或 TH 就作決定，出現 HH 或 TT 就重複。每次嘗試成功的概率是 $2pq$，所以平均需要 $\frac{1}{2pq}$ 次嘗試；當硬幣偏差很大時，這會很慢（圖 12.9.1）。題目的程序每輪作出決定的概率是 $p^2+q^2 \geq \frac{1}{2}$，但當 $p > \frac{1}{2}$ 時，它的結果偏向酒店。

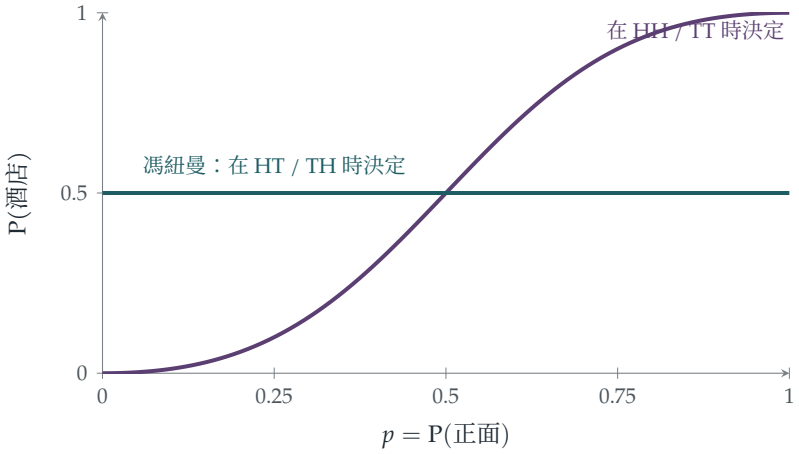


圖 12.9.1 選中酒店的概率，對應硬幣的偏差。在兩枚硬幣結果相同時作決定（紫紅色）很快，但偏向較可能出現的一面。馮紐曼的程序（藍綠色）對每一個 p 都完全公平，代價是當硬幣偏差很大時需時較長。

提取隨機性

馮紐曼的妙計是第一個隨機性提取器（randomness extractor）：一種把不完美的隨機性變成完美隨機性的程序。它浪費很多（每兩次投擲最多只得到一個公平的位元，而且往往一個也得不到）；更好的方法會把被棄掉的 HH 和 TT 重新利用，從而提取硬幣所含的幾乎全部隨機性，而這隨機性的多少是用硬幣的熵（entropy）來量度的，即每次投擲 $-p\log_2 p - q\log_2 q$ 個位元。提取器用於電腦的硬件隨機數產生器：這些產生器收集不可預測但有偏差的物理雜訊，並必須把它變成無偏差的位元，供密碼學（cryptography）使用。

馮紐曼與隨機數

馮紐曼在 1951 年的論文「Various techniques used in connection with random digits」中描述了這個妙計。當時他和同事正在為設計核武器而進行一些最早期的電腦模擬（即

「蒙地卡羅方法」（Monte Carlo method）。在同一篇論文中，他就用算術方法產生隨機數說了一句名言：「任何人若考慮用算術方法來產生隨機數字，當然就是身處罪惡之中。」

思維技巧

利用對稱性消去未知量

當某個參數未知時，找出那些不論參數取何值、概率都相等的事件，並以它們作爲決定的依據。

HT 和 TH 有相同的概率 pq，所以比較這兩者就把 p 完全消去了。在大學裏，同樣的想法在統計學中以樞軸量（pivotal quantities）的形式出現，也出現於實驗設計之中：在那裏，隨機化（randomisation）和對稱性會把未知的效應抵消掉。

延伸練習

1 **輪數的期望值。**求第 (i) 部分的程序所需輪數的期望值（expected number），以及馮紐曼的程序所需輪數的期望值。對哪些 p 而言，馮紐曼的程序較慢？

2 **公平的三選一。**利用一枚有偏差的硬幣，設計一個程序，從三個選項中選出一個，使每個選項的概率都恰好是 $\frac{1}{3}$。

3 **用公平硬幣造出有偏差的骰子。**利用一枚公平的硬幣，設計一個程序，模擬一枚擲出正面的概率恰好是 $\frac{1}{3}$ 的硬幣。所需投擲次數的期望值是多少？

4 **再看三個人。**在第 (ii) 部分中，他們去酒店的概率是多少？請以 p 的函數表示。當 $p = \frac{1}{2}$ 時，這個程序公平嗎？

延伸閱讀

- **John von Neumann，「Various techniques used in connection with random digits」**，*National Bureau of Standards Applied Mathematics Series* 12 (1951)。篇幅短小而著名。
- **Donald Knuth，*The Art of Computer Programming*, Vol. 2**（Addison-Wesley），第 3 章，論隨機數。
- **Michael Mitzenmacher & Eli Upfal，*Probability and Computing***（Cambridge University

Press），論提取隨機性與熵。

12.10 把繩頭繫起來

題目

STEP III 2008，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/08-s3-q13`

引子

把一百條繩子放進袋中，只讓繩頭露在外面。隨機選兩個繩頭綁在一起，如此重複，直至再沒有鬆開的繩頭為止。最後你會得到一堆糾纏的繩圈。究竟有多少個？一百條繩子原則上可以結成一百個小環，也可以結成一個巨大的圈。出人意表的是，繩環數目的期望值（expected number）只有約 3.3；即使有四萬條繩子，也不過約 6。

只要看出一個關鍵之處，問題便幾乎不費吹灰之力：每打一個結之後，不論剛才發生了甚麼事，箱中都只是少了一條繩子。這個過程會「忘記」自己的歷史，於是繩環的數目就變成一連串獨立（independent）擲硬幣結果之和，只是每次的概率（probability）各有不同。

動筆之前

想一想 把兩個繩頭綁在一起之後，還剩下多少條繩子（結成環的不再算作一「條」）？這是否取決於有沒有結成環？

如果兩個繩頭屬於同一條繩子，這條繩子便成為一個環，餘下 $n-1$ 條。如果兩個繩頭屬於不同的繩子，這兩條繩子便接成較長的一條，同樣餘下 $n-1$ 條。無論是哪一種情況，我們都剩下 $n-1$ 條繩子，每條各有兩個繩頭。

逐步解題

階段 1　第一個結

繩頭共有 $2n$ 個。先選定一個；與它相綁的繩頭等可能地是其餘 $2n-1$ 個繩頭中的任何一個，而當中恰好有一個屬於同一條繩子。因此第一個結結成環的概率是 $\frac{1}{2n-1}$，而第一步所產生的繩環數目的期望值就是 $\frac{1}{2n-1}$。

階段 2　所有的結

想一想　根據上面的觀察，打第二個結之前的情況是怎樣的？第三個呢？設 I_k 為指示變量，表示在剩下 k 條繩子時所打的結是否結成環。各個 $\mathrm{E}(I_k)$ 是多少？

每次打結之前，箱中有 k 條繩子，其中 $k = n, n-1, \ldots, 1$，而這個結結成環的概率是 $\frac{1}{2k-1}$。（最後一個結在只剩一條繩子時打，必定結成環。）繩環的總數是 $R = I_n + I_{n-1} + \cdots + I_1$，所以

$$\mathrm{E}(R) = \sum_{k=1}^{n} \frac{1}{2k-1} = 1 + \frac{1}{3} + \frac{1}{5} + \cdots + \frac{1}{2n-1}. \qquad \blacksquare$$

階段 3　方差

想一想　各個指示變量（indicator）I_k 是否獨立？（某一步結成環的概率，是否取決於之前發生過的事？）

在每一步，不論之前發生過甚麼事，結成環的概率都是 $\frac{1}{2k-1}$，因為當時的格局永遠是「k 條繩子」。所以各個 I_k 互相獨立，方差（variance）可以相加：

$$\mathrm{Var}(R) = \sum_{k=1}^{n} \frac{1}{2k-1}\left(1 - \frac{1}{2k-1}\right) = \sum_{k=1}^{n} \frac{2k-2}{(2k-1)^2}. \qquad \blacksquare$$

階段 4　四萬條繩子

想一想　試用調和數（harmonic numbers）表示直至 $2n-1$ 為止所有奇數的倒數之和。

把直至 $2n$ 為止所有正整數的倒數之和，減去其中偶數的倒數之和，便得

$$\sum_{k=1}^{n} \frac{1}{2k-1} = H_{2n} - \tfrac{1}{2}H_n \approx \ln(2n) - \tfrac{1}{2}\ln n = \tfrac{1}{2}\ln(4n).$$

當 $n = 40\,000$ 時，$4n = 160\,000 = 20^4$，所以 $\mathrm{E}(R) \approx \frac{1}{2} \cdot 4 \ln 20 \approx 6$。 ■（準確值是 6.28；兩者之差來自歐拉常數（Euler's constant），因為 $H_n \approx \ln n + 0.577$。）

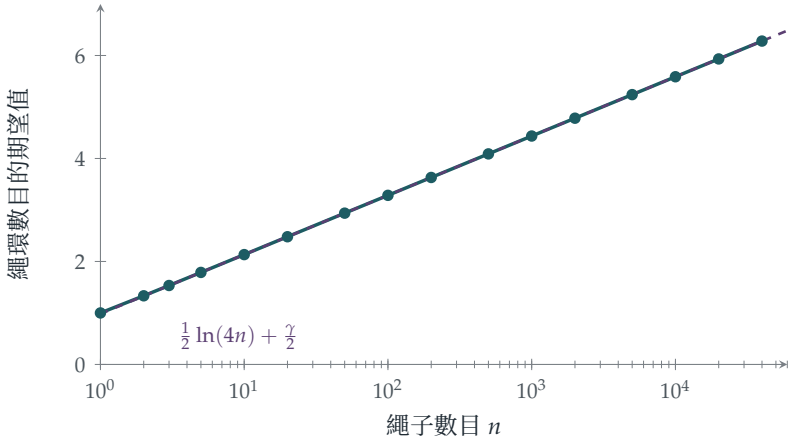


圖 12.10.1 繩環數目的期望值與繩子數目的關係，橫軸採用對數尺度。準確值（藍綠色）落在直線 $\frac{1}{2}\ln(4n) + \frac{\gamma}{2}$（紫紅色虛線）上，其中 $\gamma \approx 0.577$ 是歐拉常數。把繩子數目乘以 10，繩環只會多出約 1.15 個。

回顧

- **找出過程忘記了甚麼。**每打一個結之後，不論過去的經過如何，箱中都是「k 條繩子」。
- **又是指示變量。**繩環的數目是指示變量之和，每個結對應一個。
- **有獨立性便有方差。**由於每一步的概率都不取決於過去，各個指示變量互相獨立。

背後的數學

對數增長

繩環數目的期望值以 $\frac{1}{2}\ln n$ 的速度增長，而方差與平均值相當接近（兩者都是一些很小的概率之和）。根據中心極限定理（central limit theorem）（Section 12.24），當 n 很大時，繩環數目近似服從正態分佈（normal distribution），平均值和方差都約為 $\frac{1}{2}\ln(4n)$。若有 40 000 條繩子，你通常會得到 2 至 11 個環。

隨機置換的循環

同一個對數也出現在 n 件物件的隨機置換（random permutation）的循環（cycle）數目之中：平均有 $H_n \approx \ln n$ 個（繩子問題是它的近親，對應的是 $\frac{1}{2}\ln n$）。隨機置換的循環很少：通常有一個巨大的循環包含了大部分元素，而最大循環所佔元素的比例，平均約為 0.62，即 Golomb–Dickman 常數。隨機地把繩子綁在一起，情況也是一樣：大部分繩子最終都落在一兩個長圈之中。這些事實屬於隨機組合結構（random combinatorial structures）的研究範圍，而這門學問又與數論（number theory）有關（一個隨機整數的質因數，其表現就像一個隨機置換的循環）。

隨機紐結與聚合物

綁繩子問題是某些物理系統的簡化模型。溶液中的長鏈聚合物分子會首尾相接，形成環和鏈，而成環的統計規律也遵從類似的法則。在紐結理論（knot theory）中，數學家問的是一個更難的問題：一個隨機的閉合圈有多大機會打了結？對於很長的隨機圈，答案趨向 1，這個結果由 De Witt Sumners 和 Stuart Whittington 於 1988 年證明。

思維技巧

找出無記憶的結構

如果每一步之後，系統所處的狀態只取決於已經走了多少步，那麼各步的表現便互相獨立，期望值和方差都可以直接相加。

這正是優惠券收集問題（coupon collector）（Section 9.11）背後的想法，當中每個階段都是一次全新的幾何等待；它也在馬可夫鏈（Markov chains）的理論中再次出現，在那裏只有當前的狀態才起作用。

延伸練習

1. **小的情況。**對於 $n = 2$，列出所有可能情況，並驗證繩環數目的期望值是 $\frac{4}{3}$。

2. **一個大圈。**證明整個過程最終得到單一個包含所有繩子的圈的概率是 $\prod_{k=2}^{n}\left(1 - \frac{1}{2k-1}\right)$，並證明當 n 很大時，這個概率約為 $\sqrt{\frac{\pi}{4n}}$。（提示：找出這個乘積與 $\binom{2n}{n}$ 的關係。）

3. **置換的循環。**證明 n 件物件的隨機置換的循環數目的期望值是 H_n，方法是逐個元素地構造這個置換（即「中國餐館」構造法）。

4 方差。證明 $\mathrm{Var}(R) = \mathrm{E}(R) - \sum_{k=1}^{n} \frac{1}{(2k-1)^2}$，並證明被減去的那個和趨向 $\frac{\pi^2}{8}$。

延伸閱讀

- **Peter Winkler, *Mathematical Puzzles: A Connoisseur's Collection***（A K Peters），書中收錄了這道綁繩子謎題。
- **Richard Arratia, Andrew Barbour & Simon Tavaré, *Logarithmic Combinatorial Structures***（European Mathematical Society）。講述循環、質數與隨機結構，可留待日後閱讀。
- **Colin Adams, *The Knot Book***（American Mathematical Society）。紐結理論入門，內容包括隨機紐結。

12.11 郵票、杯子與 e

題目

STEP I 1996，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/96-s1-q13`

引子

有這樣一種古怪的賣郵票方法：買家不斷猜郵票藏在哪一個杯子下面，每一輪多加一個杯子；到他第一次猜錯時，桌上有多少個杯子，他就要付多少英鎊。平均而言，他要付 £e，約為 £2.72。

e 這個數在概率（probability）中出現之頻繁，簡直不可思議：沒有人拿回自己襯衫的機會（Section 9.8）、祕書問題（secretary problem，Section 12.12）、泊松分佈（Poisson distribution）。本題也許是它眾多現身之中最簡單的一次，而且原來它與另一道著名謎題有着完全相同的分佈（distribution）：要把多少個介乎 0 與 1 之間的隨機數相加，總和才會超過 1？平均而言，是 e 個。

動筆之前

? **想一想** 若 $N = 4$，必須發生甚麼事？概率是多少？

Jim 必須在有 2 個杯子時猜中（概率為 $\frac{1}{2}$），在有 3 個杯子時也猜中（$\frac{1}{3}$），然後在有 4 個杯子時猜錯（$\frac{3}{4}$）。所以 $\mathrm{P}(N = 4) = \frac{1}{2} \cdot \frac{1}{3} \cdot \frac{3}{4} = \frac{3}{4!} = \frac{1}{8}$。

逐步解題

階段 1 **分佈**

若 $N = k$（其中 $k \geq 2$），Jim 在有 $2, 3, \ldots, k-1$ 個杯子時都猜中，而在有 k 個杯子時猜錯：

$$\mathrm{P}(N = k) = \frac{1}{2} \cdot \frac{1}{3} \cdots \frac{1}{k-1} \cdot \frac{k-1}{k} = \frac{k-1}{k!}.$$

（驗算：$\frac{1}{2!} + \frac{2}{3!} + \frac{3}{4!} + \cdots = \sum_{k \geq 2} \left(\frac{1}{(k-1)!} - \frac{1}{k!} \right) = 1$，這是一個裂項和（telescoping sum）。）

階段 2　期望值

想一想　化簡 $k \cdot \frac{k-1}{k!}$。你會得到甚麼級數（series）？

$$\mathrm{E}(N) = \sum_{k \geq 2} \frac{k(k-1)}{k!} = \sum_{k \geq 2} \frac{1}{(k-2)!} = \frac{1}{0!} + \frac{1}{1!} + \cdots = e. \qquad \blacksquare$$

階段 3　方差

想一想　求 $\mathrm{E}(N^2)$：可寫成 $k^2 = k(k-1) + k$，或直接化簡 $\frac{k^2(k-1)}{k!}$。

$$\mathrm{E}(N^2) = \sum_{k \geq 2} \frac{k^2(k-1)}{k!} = \sum_{k \geq 2} \frac{k}{(k-2)!} = \sum_{j \geq 0} \frac{j+2}{j!} = \sum_{j \geq 1} \frac{1}{(j-1)!} + 2 \sum_{j \geq 0} \frac{1}{j!} = e + 2e = 3e.$$

所以 $\mathrm{Var}(N) = 3e - e^2 = e(3-e) \approx 0.766$。 $\blacksquare$

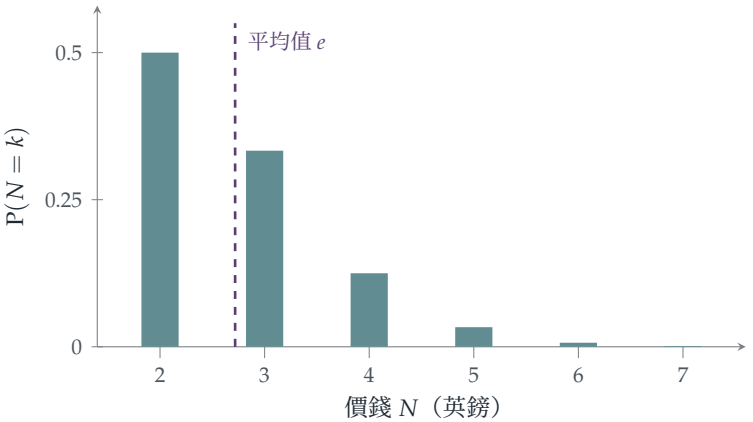


圖 12.11.1　價錢的分佈，$\mathrm{P}(N = k) = \frac{k-1}{k!}$。有一半機會 Jim 只需付 £2；那條又長又細的尾巴把平均值拉高至 £$e \approx$ £2.72。

回顧

▶ **概率的乘積。**一連串獨立（independent）的成功之後接着一次失敗，所得的乘積可

化簡為 $\frac{k-1}{k!}$。

- **約去階乘。**乘以 k 或 k^2 再約簡，那些和式就變成了指數級數（exponential series）。
- **以平移求 $\mathrm{E}(N^2)$。**以 $j = k-2$ 重新編排下標，和式便分拆成兩份 e。

背後的數學

換了裝扮的同一個分佈

從 $[0,1]$ 中均勻（uniformly）地隨機選取數 $U_1, U_2, \ldots$，並設 M 為令它們的總和超過 1 所需的個數。首 k 個數之和小於 1 的概率，就是區域 $\{u_1 + \cdots + u_k < 1, u_i \geq 0\}$ 的體積，即 $\frac{1}{k!}$（面積為 $\frac{1}{2}$ 的三角形、體積為 $\frac{1}{6}$ 的四面體，如此類推）。所以

$$\mathrm{P}(M > k) = \frac{1}{k!}, \qquad \mathrm{P}(M = k) = \frac{1}{(k-1)!} - \frac{1}{k!} = \frac{k-1}{k!},$$

正正就是 N 的分佈。在兩個問題中都有 $\mathrm{P}(N > k) = \frac{1}{k!}$：就杯子而言，這是連續猜中 $k-1$ 次的概率。利用 $\mathrm{E}(N) = \sum_{k \geq 0} \mathrm{P}(N > k)$（Section 9.9 中的一條公式），一行便得出 $\mathrm{E}(N) = \sum \frac{1}{k!} = e$。

為何 e 不斷出現

每當概率中含有 $\frac{1}{k!}$，不論它來自 k 件物件按某個特定次序排列、k 個概率依次為 $\frac{1}{2}, \frac{1}{3}, \ldots$ 的相繼事件，還是單純形（simplex）的體積，對 k 求和都會產生指數級數。錯排（derangements，Section 9.8）給出 $\sum \frac{(-1)^k}{k!} = e^{-1}$；這裏我們則有 $\sum \frac{1}{k!} = e$。至於形如 $\left(1 - \frac{1}{n}\right)^n \to e^{-1}$ 的極限（limit），它們源自大量獨立而不大可能發生的事件，正好解釋了祕書問題和泊松分佈中的 e。

蒙地卡羅 e

「總和超過 1」這個問題提供了一個以實驗估計 e 的方法：產生隨機數，數一數需要多少個才超過 1，重複多次，然後取平均值。它是蒲豐投針（Buffon's needle，Section 12.20）的近親，後者用來估計 π。這類蒙地卡羅（Monte Carlo）方法以賭場命名，利用隨機抽樣來估計難以直接計算的量；如今它們廣泛應用於科學的各個範疇，由粒子物理學到金融，主要用於計算高維積分。

思維技巧

認出級數

當期望值是一個含有階乘的和式時，把各項化簡，直至出現一個標準級數（指數級數、幾何級數、二項級數）。

在大學，生成函數(generating function)把這種做法系統化：概率生成函數(probability generating function) $\mathrm{E}(t^N)$ 把整個分佈包裝起來，而它在 $t=1$ 處的導數則給出平均值和方差。

延伸練習

1 **生成函數。**證明 $\mathrm{E}(t^N) = 1 + (t-1)e^t$，並用它求 $\mathrm{E}(N)$ 和 $\mathrm{E}(N(N-1))$。

2 **尾部公式。**對於取非負整數值的隨機變量（random variable），證明 $\mathrm{E}(N) = \sum_{k\geq 0} \mathrm{P}(N > k)$，並用它一行求出 $\mathrm{E}(N)$。

3 **由一個杯子開始。**假如遊戲由一個杯子開始 (Jim 不可能猜錯)，分佈會有甚麼改變？價錢的期望值是多少？

4 **單純形。**證明 $\{u_1 + u_2 + u_3 < 1, u_i \geq 0\}$ 的體積為 $\frac{1}{6}$，並用歸納法（induction）證明 k 維的情況為 $\frac{1}{k!}$。

延伸閱讀

- **Eli Maor, *e: The Story of a Number***（Princeton University Press）。e 的歷史，包括它在概率中的各次現身。
- **Geoffrey Grimmett & Dominic Welsh, *Probability: An Introduction***（Oxford University Press），關於生成函數。
- **Paul Nahin, *Digital Dice***（Princeton University Press）。以蒙地卡羅模擬解決的概率謎題。

12.12 何時停止尋找

題目

STEP I 2001，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/01-s1-q14`

引子

你正在逐一面試應徵者，而且必須即場決定：取錄眼前這一位，然後停止；還是永遠拒絕這一位，再見下一位。你想要的是最好的那一位。如果太早取錄，你很可能錯過更好的人；如果等得太久，最好的那一位很可能已經走了。你應該怎樣做？

答案是概率（probability）論中最著名的結果之一：先拒絕最初 37% 的應徵者，然後取錄第一位比這些人都好的應徵者。這個做法大約有 37% 的機會成功，而 37% 正是 $\frac{1}{e}$。本題計算其中的基本概率，而與它配對的 STEP II 2006 題目（見練習）則分析這個策略（strategy）最簡單的版本：只拒絕第一位應徵者。

動筆之前

想一想 在隨機次序之下，全體之中最好的應徵者落在最先面試的 n 位之中的概率是多少？

最好的應徵者落在 N 個位置中任何一個的可能性均等，所以概率是 $\frac{n}{N}$。

逐步解題

階段 1 最先 n 位中的最佳者就是全體最佳者

最先 n 位中的最佳者是全體最佳者，當且僅當全體最佳者在最先 n 位之中：概率為 $\frac{n}{N}$。

■

階段 2　最佳或次佳

想一想　最先 n 位中的最佳者是全體次佳者，當且僅當次佳者在最先 n 位之中而最佳者不在其中。這個概率是多少？

次佳者在最先 n 位之中的概率是 $\frac{n}{N}$；在此條件下，最佳者在其餘 $N-1$ 個位置的其中一個，當中有 $N-n$ 個位置排在最先 n 位之後。所以

$$\mathrm{P}(\text{最佳或次佳}) = \frac{n}{N} + \frac{n}{N}\cdot\frac{N-n}{N-1}. \qquad \blacksquare$$

當 $N=4$，$n=2$ 時：$\frac{1}{2}+\frac{1}{2}\cdot\frac{2}{3}=\frac{5}{6}$。把最好的兩位應徵者在四個位置中的排法逐一列出（全部 12 種有序排法的可能性均等），除了兩位頂尖應徵者同時落在第 3 和第 4 位的 2 種排法之外，最先兩位中的最佳者都是頂尖兩位之一：$\frac{10}{12}=\frac{5}{6}$。

階段 3　最優策略

想一想　策略：先拒絕最初 r 位應徵者，然後取錄第一位比他們全部都好的應徵者。如果最佳應徵者在位置 $k>r$，這個策略在甚麼情況下會選中此人？

設最佳應徵者在位置 k，這個策略選中此人，當且僅當位置 $r+1,\ldots,k-1$ 上沒有任何應徵者勝過最初 r 位全部的人，即是說，最先 $k-1$ 位應徵者中的最佳者在最初 r 位之中。根據第 1 階段（把 N 換成 $k-1$），這件事的概率是 $\frac{r}{k-1}$。最佳應徵者落在每個位置 k 的概率都是 $\frac{1}{N}$，所以

$$\mathrm{P}(\text{成功}) = \frac{1}{N}\sum_{k=r+1}^{N}\frac{r}{k-1} = \frac{r}{N}\left(\frac{1}{r}+\frac{1}{r+1}+\cdots+\frac{1}{N-1}\right).$$

想一想　當 N 很大時，記 $x=\frac{r}{N}$。這個和大約等於 $\ln\frac{N}{r}$。x 取甚麼值時 $x\ln\frac{1}{x}$ 達到最大？

成功的概率大約是 $x\ln\frac{1}{x}$。求導可得，$\ln\frac{1}{x}-1=0$ 於 $x=\frac{1}{e}$ 時成立，而最大值是 $\frac{1}{e}\approx 0.368$（圖 12.12.1）。

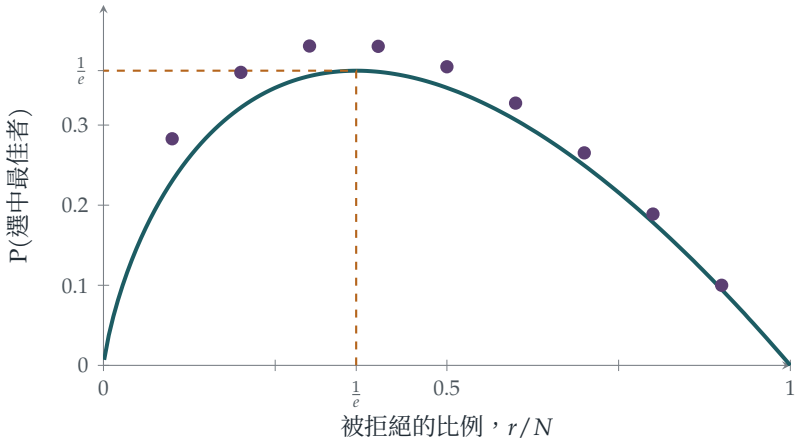


圖 12.12.1 若先拒絕佔比例 r/N 的最初一批應徵者，選中最佳應徵者的概率：極限曲線 $x \ln \frac{1}{x}$（藍綠色），以及 $N = 10$ 時的精確值（紫紅色）。最佳策略是拒絕大約 $\frac{1}{e} \approx 37\%$ 的應徵者，成功的概率大約是 $\frac{1}{e}$，不論 N 有多大。

回顧

- **位置的對稱性。**在隨機次序之下，每位應徵者落在任何位置的可能性均等，由此立即得出 $\frac{n}{N}$。
- **以最佳者的位置為條件。**計算策略的成功概率時，我們以最佳應徵者所在的位置為條件分情況處理。
- **求和變成積分。**當 N 很大時，調和和（harmonic sum）變成了對數，再經一次求最大值便得出 $\frac{1}{e}$。

背後的數學

最優停止

祕書問題（secretary problem）是最優停止（optimal stopping）的奠基例子：根據隨時間陸續到來的資訊，決定何時採取行動，以使期望回報最大。可以證明，在所有策略之中（而不只是「先拒絕 r 位，然後取錄下一位至今最佳者」這一類），這種閾值規則（threshold rule）是最優的，論證的方法是由最後一位應徵者開始向前逆推。同樣的數學也決定了何時賣出股票、何時行使金融期權，以及應該花多久去找泊車位。一個值得注意的特點是，在極限之下答案與 N 無關：不論應徵者有一百位還是一百萬位，你成功的概率都是 $\frac{1}{e}$。

各種變化

如果你只要選中頂尖兩位應徵者的其中一位便滿意，最佳策略大約有 57% 的機會成功。如果你想做到的不是使選中最佳者的機會最大，而是使所選應徵者的期望*名次*最優，最優策略便複雜得多，而期望名次趨向大約 3.87（Chow、Moriguti、Robbins 和 Samuels，1964）。如果你有一定的概率可以回頭召回已被拒絕的應徵者，又或者你知道應徵者分數的分佈（即「完全資訊」問題），答案又會不同。練習中的雪糕題是最簡單的情形：只拒絕一個，然後取下一個比它好的。

是誰解決的？

這個問題在 1950 年代流傳於數學家之間，馬丁·加德納（Martin Gardner）於 1960 年以「googol 遊戲」之名把它普及開來（在紙條上寫下數字；你的對手逐張翻開，並且必須在最大的那個數出現時停下）。Thomas Ferguson 的文章「Who solved the secretary problem?」(1989)追溯了它錯綜複雜的歷史；有一種說法指，天文學家開普勒(Johannes Kepler）在 1611 年挑選第二任妻子時，會見了十一位人選，最後向第五位求婚。

思維技巧

把策略參數化，然後求最優

當最佳策略未知時，考慮一族帶有參數的自然策略，把成功概率寫成該參數的函數，然後求最優。

在大學裏，這種方法用於決策論、博弈論和統計學。要證明這一族之中的最佳策略就是全體之中的最佳策略，是更困難也更有趣的一步，通常要用*逆向歸納法*（backward induction）來完成。

延伸練習

1. **雪糕。**試做 STEP II 2006，第 13 題（`step.maths.org/questions/06-s2-q13`）：這是祕書問題的小型版本，規則很簡單。拒絕第一個，然後取第一個勝過它的。

2. **較小的 N。**當 $N = 4$ 時，求 r 的最佳值和成功的概率。用列舉法驗算。

3. **最佳或次佳。**利用第 (ii) 部分，論證「最先 n 位中的最佳者是頂尖兩位之一」的可能性遠高於「最先 n 位中的最佳者就是最佳者」。這對於選中頂尖兩位之一的策略有甚麼啟示？

4 **精確的最大值。**當 $N = 10$ 時，從圖中驗證 $r = 3$ 是最優的。當 $N = 100$ 時，證明 $r = 37$ 是最優的（準確至四位小數，$r = 36$、37、38 的成功概率分別是 0.3710、0.3710、0.3708）。

延伸閱讀

- **Thomas Ferguson，「Who solved the secretary problem?」**，*Statistical Science* 4 (1989)。歷史與各種變化；可在網上免費閱讀。
- **Brian Christian & Tom Griffiths，*Algorithms to Live By***（William Collins）。把最優停止及其他算法應用於日常決策。
- **Martin Gardner，*New Mathematical Diversions***（Mathematical Association of America）。googol 遊戲。

12.13 三人決鬥

題目

STEP II 1995，第 12 題

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

`step.maths.org/questions/95-s2-q12`

引子

三個人用餐包決鬥。Cuthbert 百發百中，Bertie 每五次擲中三次，而 Algernon 準頭最差，每五次只擲中兩次。Algernon 先擲。他應該瞄準誰？答案出人意表：誰也不瞄準。他應該故意把餐包擲偏。

原因在於策略。只要三人都仍在場，Bertie 和 Cuthbert 這兩個危險的對手就會互相瞄準。假如 Algernon 淘汰了其中一人，餘下的那個便會轉而對付他。他故意擲失，就能讓那兩人先行分出勝負，然後由他向倖存者先擲一次。事實上，只要採用最佳方案，準頭最差的人反而最有可能勝出。這場「三人決鬥」（truel）是博弈論（game theory）中常被引用的例子，說明弱勢有時反而是優勢。

動筆之前

? **想一想** 在一場二人決鬥中，Percy 先擲。他要麼一擲即中，要麼兩人都擲失，於是局面完全重演。試寫出一條關於 Percy 勝出概率（probability）的方程。

設 P 為 Percy 勝出的概率。他一擲即中的概率是 p；否則（概率為 $1-p$）輪到 Rodney 擲，而 Rodney 擲失的概率是 $1-r$，此時局面回到起點。所以 $P = p + (1-p)(1-r)P$。

逐步解題

階段 1 二人決鬥

解方程，得 $P(1-(1-p)(1-r))=p$，而 $1-(1-p)(1-r)=p+r-pr$，所以

$$P=\frac{p}{p+r-pr}. \qquad \blacksquare$$

若 Algernon ($p=\frac{2}{5}$) 對 Bertie ($r=\frac{3}{5}$)，由 Algernon 先擲，這個概率是 $\frac{2/5}{2/5+3/5-6/25}=\frac{10}{19}$；若由 Bertie 先擲，Algernon 勝出的概率是 $\left(1-\frac{3}{5}\right)\cdot\frac{10}{19}=\frac{4}{19}$。

階段 2 Algernon 的三個選擇

想一想 在每種情況下，還剩下誰？下一個輪到誰擲？

(i) *Algernon* 擲中 *Cuthbert*。接着由 Bertie 擲，目標是 Algernon。Algernon 要在 Bertie 先擲的二人決鬥中勝出：概率為 $\frac{4}{19}$。

(ii) *Algernon* 擲中 *Bertie*。接着由 Cuthbert 擲，目標是 Algernon，而他從不失手。Algernon 的勝出機會是 0。

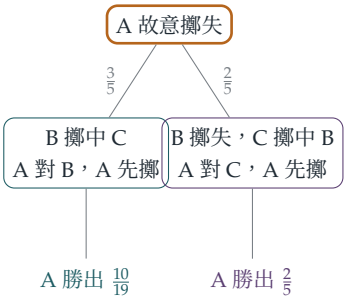
(iii) *Algernon* 擲失。Bertie 向 Cuthbert 擲。他有 $\frac{3}{5}$ 的概率擲中，剩下 Algernon 和 Bertie，並由 Algernon 先擲：Algernon 勝出的概率是 $\frac{10}{19}$。Bertie 有 $\frac{2}{5}$ 的概率擲失，然後 Cuthbert 擲中 Bertie，Algernon 便要面對 Cuthbert，而且只有一次出手機會：他必須擲中才能勝出，概率為 $\frac{2}{5}$。合計為

$$\frac{3}{5}\cdot\frac{10}{19}+\frac{2}{5}\cdot\frac{2}{5}=\frac{6}{19}+\frac{4}{25}=\frac{150+76}{475}=\frac{226}{475}\approx 0.476.$$

階段 3 建議

想一想 如果 Algernon 瞄準 Cuthbert，他可能擲中（概率為 $\frac{2}{5}$），也可能擲失。他整體的勝出機會是多少？試與故意擲失比較。

瞄準 Cuthbert 的勝出概率是 $\frac{2}{5}\cdot\frac{4}{19}+\frac{3}{5}\cdot\frac{226}{475}=\frac{200+678}{2375}=\frac{878}{2375}\approx 0.370$。瞄準 Bertie 則更差。故意擲失的勝出概率是 $\frac{226}{475}\approx 0.476$，是三者中最高的。所以 Algernon 應該把餐包擲偏，這樣他勝出的概率是 $\frac{226}{475}$。 $\blacksquare$



採用這個方案：
A 勝出 $\frac{226}{475} \approx 0.48$
B 勝出 $\frac{135}{475} \approx 0.28$
C 勝出 $\frac{114}{475} \approx 0.24$

圖 12.13.1 Algernon 的最佳方案。他故意擲失，讓兩個較強的對手先行交鋒，並確保自己能向倖存者先擲一次。結果準頭最差的人最有可能勝出，而百發百中的人反而最不可能勝出。

回顧

- **重複出現的局面可以列出方程。**兩人都擲失後，決鬥「重新開始」，由此得出一條關於勝出概率的線性方程（linear equation）。
- **考慮對手的反應。**Algernon 的每個選擇，都會改變 Bertie 和 Cuthbert 接下來的行動。
- **比較所有選擇，包括甚麼都不做。**最好的一着，正是看來最差的那一着。

背後的數學

策略與博弈論

三人決鬥展示了博弈論的核心思想：最佳行動取決於其他參與者會如何回應。擲中最危險的對手 Cuthbert 看似理所當然，但這樣做會改變局勢，反而對 Algernon 不利。在本題中，Bertie 和 Cuthbert 的策略（strategy）是固定的（他們互相瞄準）；在完整的博弈論分析中，每個參與者都會自行選擇策略，並且知道其他人也在這樣做。這類分析始於馮·諾伊曼（John von Neumann）與摩根斯坦（Oskar Morgenstern）在 1944 年的工作，以及納殊（John Nash）在 1950 年的工作，如今已成為經濟學、演化生物學和政治學的基礎。

強大反成弱點

三人決鬥中誰會勝出？若 Algernon 採用最佳方案，Algernon 勝出的概率是 $\frac{226}{475}$，Bertie 是 $\frac{135}{475}$，Cuthbert 是 $\frac{114}{475}$。準頭最好的人最不可能倖存，因為凡是可以選擇目標的人都會瞄準他。同樣的規律也見於有多名候選人的選舉、企業之間的競爭，以及自然界：最強

的競爭者可能成為首個被攻擊的目標。Martin Shubik 在 1950 年代把三人決鬥引入博弈論，並藉此探問：最適者是否必然生存？

思維技巧

多想一步

選擇行動之前，先弄清楚它會怎樣改變其他各人面對的局面，以及他們會如何回應。

到了大學，這會發展成博弈論中的逆向歸納法（backward induction）：先分析博弈的終局，再往回推，從而決定最佳的第一步。

延伸練習

1 **Bertie 與 Cuthbert。**驗證：若 Algernon 故意擲失，Bertie 勝出的概率是 $\frac{135}{475}$，Cuthbert 勝出的概率是 $\frac{114}{475}$。

2 **準頭較好的投手。**假如 Algernon 擲中的概率是 a 而不是 $\frac{2}{5}$，當 a 取哪些值時，故意擲失仍然是他最佳的第一步？

3 **Bertie 應該瞄準 Cuthbert 嗎？**已知 Algernon 故意擲失，Bertie 瞄準 Cuthbert 而不瞄準 Algernon，是否對他自己有利？（假設 Cuthbert 的行為不變。）

4 **投擲次數有限的決鬥。**在一場每人只可擲兩次的二人決鬥中，求 Percy 勝出的概率，並與不限次數的決鬥比較。

延伸閱讀

- **D. Marc Kilgour & Steven Brams，「The truel」**，*Mathematics Magazine* 70 (1997)。一篇淺白易讀的三人決鬥綜述。
- **Avinash Dixit & Barry Nalebuff，*Thinking Strategically***（Norton）。寫給一般讀者的博弈論讀物，當中也談到三人決鬥。
- **Ken Binmore，*Game Theory: A Very Short Introduction***（Oxford University Press）。

12.14 混合策略

題目

STEP I 1989，第 16 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/89-s1-q16`

引子

玩「包剪揼」時，任何固定的出法都會被對手利用：每次都出揼（石頭），對手很快便學會出包（布）。唯一穩妥的策略就是隨機出手。本題是一個更精細的版本：各種結果的得失並不相等，怎樣隨機化才對也並不明顯。結果發現，只要以概率 $\frac{1}{4}$、$\frac{1}{2}$、$\frac{1}{4}$ 分別猜 1、2、3，不論對手怎樣做，玩家都可以保證長遠而言不會輸錢。

這讓我們初次窺見馮紐曼（John von Neumann）在 1928 年提出的**極小極大定理**（minimax theorem），它是博弈論（game theory）的奠基結果：在任何一方所得即另一方所失的二人博弈中，雙方各有一個隨機化策略（randomised strategy），能保證在對手發揮得完美無瑕時，仍取得可能範圍內的最佳結果。

動筆之前

想一想 寫出一個表，列出每一對猜測之下 B 贏得的金額（B 的猜測列於左側，A 的猜測列於頂端）。

按規則逐一推算（例如，若 B 說 1 而 A 說 3，兩數相差 2，說出較小數的 B 要付 £2）：

B \ A	1	2	3
1	0	+1	−2
2	−1	0	+1
3	+2	−1	0

這個表是*反對稱*（antisymmetric）的（把兩位玩家對調，正負號便會改變），所以這個博

弈是公平的，意思是雙方都沒有與生俱來的優勢。

逐步解題

階段 1　對付固定的猜測

若 A 每次都說 1，B 應該說 3，每局贏 £2。若 A 每次都說 2，B 說 1，贏 £1。若 A 每次都說 3，B 說 2，贏 £1。任何固定的猜測都會被對手利用。

階段 2　對付隨機的猜測

> **想一想**　若 A 以均勻的概率隨機猜測，當 B 說 j 時，B 的期望收益（expected gain）是多少？（求表中第 j 行的平均值。）

各行的平均值分別是：$j=1$ 時為 $\frac{1}{3}(0+1-2)=-\frac{1}{3}$，$j=2$ 時為 0，$j=3$ 時為 $\frac{1}{3}(2-1+0)=\frac{1}{3}$。所以 B 每局贏得金額的期望值是

$$-\tfrac{1}{3}b_1+0\cdot b_2+\tfrac{1}{3}b_3,$$

當 $b_3=1$ 時取得最大值：每次都說 3，平均每局贏 $\frac{1}{3}$ 鎊。均勻的隨機化並不能保護 A。

階段 3　一般情況

> **想一想**　現在設 A 以概率 a_j 說 j。把 B 贏得金額的期望值按 A 的猜測歸項：a_1 的係數（即第 1 列以 b_j 為權的加權平均）是甚麼？

沿着表中各列由上而下讀，

$$\text{B 贏得金額的期望值}=(-b_2+2b_3)a_1+(b_1-b_3)a_2+(-2b_1+b_2)a_3,$$

所以 $X=-b_2+2b_3$，$Y=b_1-b_3$，$Z=-2b_1+b_2$。

> **想一想**　你能令三者同時為零嗎？（共有三條方程，但它們並非獨立；為甚麼？）

它們並非獨立：$X+2Y+Z=0$ 恆成立（請自行驗證）。所以由 $Y=0$ 得 $b_1=b_3$，由 $X=0$ 得 $b_2=2b_3$，兩者合起來便迫使 $Z=0$ 亦成立。再配合 $b_1+b_2+b_3=1$：

$$b_1=\tfrac{1}{4},\qquad b_2=\tfrac{1}{2},\qquad b_3=\tfrac{1}{4},$$

於是 $X=Y=Z=0$。這樣，不論 a_j 取甚麼值，B 贏得金額的期望值都是 0，B 長遠而言不會有任何損失。 ■

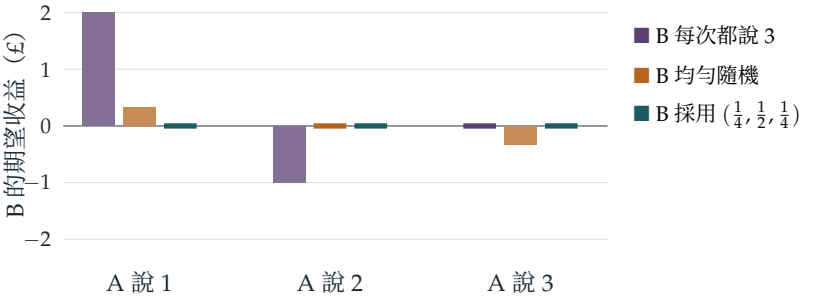


圖 12.14.1 B 採用三種不同策略時，面對 A 每一種選擇的期望收益。每次都說 3，遇上 A 說 1 會贏，遇上 A 說 2 卻會輸；均勻隨機猜測，遇上 A 說 3 會輸。策略 $(\frac{1}{4}, \frac{1}{2}, \frac{1}{4})$ 則不論 A 怎樣做，收益都恰好是 0（圖中以幼條表示），所以永遠不會被對手利用。

回顧

- **寫出得失表。**所有推理都圍繞一個 3×3 矩陣的行和列。
- **純策略（pure strategy）會被對手利用。**面對任何固定的猜測，甚至面對均勻隨機的猜測，對手都有一個有利可圖的應對方法。
- **拉平。**穩妥的策略令對手選甚麼都沒有分別：A 的每一種選擇都帶來相同的期望結果。

背後的數學

極小極大定理

在一個二人零和（zero-sum）博弈中，設得失矩陣（payoff matrix）為 M，B 選取概率 $\mathbf{b}$，A 選取概率 $\mathbf{a}$，則 B 贏得金額的期望值是 $\mathbf{b}^{\mathsf{T}} M \mathbf{a}$。馮紐曼的極小極大定理指出

$$\max_{\mathbf{b}} \min_{\mathbf{a}} \mathbf{b}^{\mathsf{T}} M \mathbf{a} = \min_{\mathbf{a}} \max_{\mathbf{b}} \mathbf{b}^{\mathsf{T}} M \mathbf{a}.$$

用文字說：B 預先公開自己的隨機化策略而仍能保證取得的金額，等於 A 預先公開自己的策略而能把 B 壓制到的金額。這個共同的數值稱為博弈的值（value）。在本題中，由反對稱性可知博弈的值是 0，而策略 $(\frac{1}{4}, \frac{1}{2}, \frac{1}{4})$ 對雙方都是最優的。這個定理可用線性代數（linear algebra）和凸性（convexity）證明（它等價於線性規劃（linear programming）的對偶定理），而它推廣到多人博弈的版本，即納殊均衡（Nash equilibrium），為納殊（John Nash）贏得諾貝爾經濟學獎。

為甚麼要隨機化？

這裏的隨機性無關運氣，而是關乎資訊。如果 B 的選擇可以被預測，A 便能加以利用；B 以恰當的概率隨機選擇，既令自己的選擇無從預測，又確保任何預測都幫不了對手。現實中的例子包括足球的十二碼罰球——對職業球員的研究發現，主射球員和守門員在左右兩邊之間隨機選擇，比例與博弈論的最優值十分接近——以及網球的發球。

馮紐曼 1928 年的論文

馮紐曼在 1928 年證明了極小極大定理，論文題為「On the theory of games of strategy」（論策略博弈的理論），研究動機來自撲克，因為撲克中的「偷雞」（bluffing，即虛張聲勢）正是一種隨機化。1944 年，他與經濟學家 Oskar Morgenstern 出版了 *Theory of Games and Economic Behavior*（博弈論與經濟行為），博弈論自此成為一門學科。馮紐曼後來憶述，他一直視極小極大定理為關鍵結果，沒有它，就沒有值得發表的博弈理論。

思維技巧

以隨機化消除對手的優勢

當任何固定的選擇都會被對手利用時，就隨機選擇，而所用的概率要令對手無論選哪一項都沒有分別。

在大學裏，博弈論、線性規劃和演算法理論都會用到這個想法。例如，隨機演算法（randomised algorithm）會作出任何對手都無法預測的隨機選擇，藉此防範最壞情況的輸入。

延伸練習

1 **A 的最佳應對。**若 B 採用 $\left(\frac{1}{4}, \frac{1}{2}, \frac{1}{4}\right)$，證明 A 的每一種策略所得的期望值都是 0。若 B 改用 $\left(\frac{1}{3}, \frac{1}{3}, \frac{1}{3}\right)$，求 A 的最佳應對。

2 **包剪揼。**寫出「包剪揼」的得失矩陣，並求最優策略。為甚麼這裏的最優策略是均勻的，而本題的卻不是？

3 **唯一性。**證明 $\left(\frac{1}{4}, \frac{1}{2}, \frac{1}{4}\right)$ 是 B 能保證不輸的唯一策略。（提示：它必須令 X、Y、Z 全部非負；利用 $X + 2Y + Z = 0$。）

4 **改動後的博弈。**若相差 2 時的罰款由 £2 改為 £3，求最優策略。

延伸閱讀

- **Ken Binmore, *Playing for Real***（Oxford University Press）。一本詳盡而平易近人的博弈論入門書。
- **Avinash Dixit & Barry Nalebuff, *The Art of Strategy***（Norton）。以體育和商業的例子講解博弈論。
- **William Poundstone, *Prisoner's Dilemma***（Anchor）。馮紐曼、博弈論與冷戰。

12.15 滅絕還是爆發

題目

STEP II 2003，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/03-s2-q12`

引子

1873 年，高爾頓（Francis Galton）在《*Educational Times*》上提出一個問題：為甚麼貴族的姓氏那麼容易失傳？每個男人所生的兒子數目是隨機的，而每個兒子都會把姓氏傳下去。即使每個男人平均有多於一個兒子，這個姓氏就能永遠流傳下去嗎？這個問題開創了分支過程（branching process）的理論，如今這套理論用來模擬核連鎖反應、傳染病的擴散、細胞羣體的增長，以及突變能否存活下去。

本題建立的是最簡單的模型：每個粒子分裂成 0、1 或 2 個粒子。粒子總數的期望值（expectation）按幾何級數增長或縮減。但期望值並不是故事的全部。即使期望的粒子總數急劇膨脹，整個族系完全滅絕的概率仍然是正數，而一個漂亮的不動點（fixed point）論證就能把這個概率求出來。

動筆之前

? **想一想** 如果某一階段有兩個粒子，下一階段粒子數目的分佈是甚麼？

兩個粒子各自獨立地產生後代，後代數目服從 $\mathrm{B}(2, p)$，所以總數服從 $\mathrm{B}(4, p)$：共有四次獨立的「機會」，每次成功的概率為 p。

逐步解題

階段 1　先是兩個，然後又是兩個

$\mathrm{P}(X_1 = 2) = p^2$，然後 $\mathrm{P}(X_2 = 2 \mid X_1 = 2) = \binom{4}{2} p^2 q^2 = 6p^2q^2$。所以

$$\mathrm{P}(X_1 = 2, X_2 = 2) = 6p^4q^2. \qquad \blacksquare$$

階段 2　逆向推理

> **想一想**　$X_2 = 2$ 可以有多少種方式發生？（X_1 可以是甚麼？）

要使 $X_2 = 2$，可以是 $X_1 = 1$（概率為 $2pq$），其後產生 2 個後代（p^2）；也可以是 $X_1 = 2$，其後由 $\mathrm{B}(4, p)$ 得出 2 個。（$X_1 = 0$ 則甚麼也不剩。）所以 $\mathrm{P}(X_2 = 2) = 2p^3q + 6p^4q^2$，而

$$\mathrm{P}(X_1 = 2 \mid X_2 = 2) = \frac{6p^4q^2}{2p^3q + 6p^4q^2} = \frac{3pq}{1 + 3pq}.$$

令此式等於 $\frac{9}{25}$：$75pq = 9 + 27pq$，所以 $pq = \frac{3}{16}$，即 $p(1-p) = \frac{3}{16}$，由此得 $p = \frac{1}{4}$ 或 $p = \frac{3}{4}$。 $\blacksquare$

階段 3　粒子總數的期望值

> **想一想**　如果在第 $n-1$ 階段有 k 個粒子，第 n 階段粒子數目的期望值是多少？然後對 k 取平均。

每個粒子平均有 $\mathrm{E}(X) = 2p$ 個後代，所以 k 個粒子平均有 $2pk$ 個後代：$\mathrm{E}(X_n \mid X_{n-1} = k) = 2pk$。對 k 取平均，得 $\mathrm{E}(X_n) = 2p\,\mathrm{E}(X_{n-1})$。由於 $\mathrm{E}(X_0) = 1$，

$$\mathrm{E}(X_n) = (2p)^n.$$

當 $p = \frac{1}{4}$ 時，$(2p)^n = 2^{-n} \to 0$；當 $p = \frac{3}{4}$ 時，$(2p)^n = \left(\frac{3}{2}\right)^n \to \infty$。 $\blacksquare$

回顧

- **獨立二項變量之和。**k 個粒子各自給出 $\mathrm{B}(2, p)$，合起來就是 $\mathrm{B}(2k, p)$。
- **時間上的貝葉斯定理（Bayes' theorem）。**$\mathrm{P}(X_1 = 2 \mid X_2 = 2)$ 是由較後的觀察結果逆向推斷較早的狀態。
- **條件期望值（conditional expectation）。**以上一代的情況為條件，期望值就化成一個等比數列（geometric sequence）。

背後的數學

滅絕的概率

當 $p = \frac{3}{4}$ 時，粒子總數的期望值無止境地增長。這是否表示羣體一定能存活下去？不一定：如果第一個粒子沒有後代（概率為 $q^2 = \frac{1}{16}$），這個族系便立即終結。設 e 為最終滅絕的概率。第一個粒子有 k 個後代，而每個後代的族系各自獨立地以概率 e 滅絕，所以

$$e = q^2 + 2pq\,e + p^2e^2 = (q + pe)^2 = \mathrm{G}(e),$$

其中 $\mathrm{G}(s) = (q + ps)^2$ 是後代數目的*概率生成函數*（probability generating function）。滅絕概率是 G 的一個不動點：事實上是最小的非負不動點。當 $p = \frac{3}{4}$ 時，由 $(1+3e)^2 = 16e$ 得 $9e^2 - 10e + 1 = 0$，所以 $e = \frac{1}{9}$ 或 1，而滅絕概率是 $\frac{1}{9}$。當 $p = \frac{1}{4}$ 時，在 $[0,1]$ 內唯一的不動點是 $e = 1$：滅絕是必然的（圖 12.15.1）。

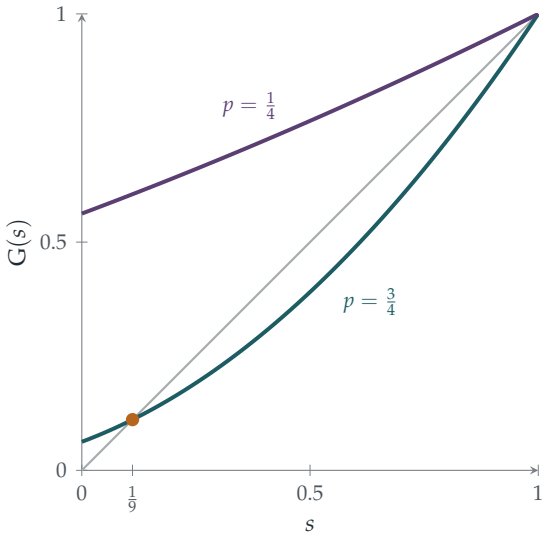


圖 12.15.1　後代數目的生成函數 $\mathrm{G}(s) = (q + ps)^2$ 與直線 $y = s$。滅絕概率就是兩者相交的最小一點。當 $p = \frac{1}{4}$（紫紅色，平均值 $\frac{1}{2}$）時，該點是 $s = 1$：滅絕是必然的。當 $p = \frac{3}{4}$（藍綠色，平均值 $\frac{3}{2}$）時，該點是 $s = \frac{1}{9}$（琥珀色）：羣體以概率 $\frac{8}{9}$ 永遠存活下去。

臨界門檻

一般而言，如果後代數目的平均值小於 1，分支過程必然滅絕；平均值等於 1 時也是如此，唯一的例外是每個個體都恰好有一個後代的平凡情況（這時族系只是一個接一個地永遠延續下去）；如果平均值大於 1，則分支過程以正概率存活下去。後代數目的平均值就是流行病學中的*繁殖數*（reproduction number）R：$R > 1$ 的疫症有可能大規模爆發，但即使如此，早期的爆發仍可能純粹因為偶然而無疾而終，其概率正好由上面的不動點方程給出。在核物理學中，如果每次裂變平均恰好引發另一次裂變，連鎖反應就處於臨

界（critical）狀態，而操作反應堆的學問，就在於讓它一直保持在這個狀態。

高爾頓、Watson 與 Bienaymé

高爾頓的問題在 1874 年由 Henry William Watson 牧師作答，但答案有部分是錯的：他找到了不動點方程，卻漏掉了第二個根，因而斷定每個姓氏最終都會失傳。其實正確的分析早在 1845 年已由法國統計學家 Irénée-Jules Bienaymé 提出，只是他的研究一直被人遺忘，直到 1970 年代才重見天日。這個模型現在稱為*高爾頓—Watson 過程*（Galton–Watson process，有時稱為 Bienaymé–Galton–Watson 過程），是每一門隨機過程（random processes）課程中的核心例子。

思維技巧

以第一代為條件

一個會自我更新的過程（每個後代都開始一個獨立的、與整個過程相同的複本），必定滿足一條以第一步爲條件而得出的方程。

方程 $e = \mathrm{G}(e)$ 是一個不動點問題，與 Section 2.2 和 Section 8.7 中遇到的屬同一類。在大學裏，生成函數和不動點是研究分支過程、隨機樹和隨機網絡的主要工具。

延伸練習

1 **一般情況下的滅絕。**設後代數目的分佈為 $\mathrm{B}(2, p)$，證明當 $p \le \frac{1}{2}$ 時滅絕概率為 1，而當 $p > \frac{1}{2}$ 時滅絕概率為 $\left(\frac{q}{p}\right)^2$。

2 **方差（variance）。**證明 $\mathrm{Var}(X_n)$ 滿足 $\mathrm{Var}(X_n) = (2p)^2\mathrm{Var}(X_{n-1}) + 2pq\,\mathrm{E}(X_{n-1})$，並求 $p = \frac{1}{2}$ 時的方差。

3 **在第 n 代或之前滅絕。**設 $e_n = \mathrm{P}(X_n = 0)$。證明 $e_n = \mathrm{G}(e_{n-1})$，並就 $p = \frac{3}{4}$ 計算 e_1、e_2、e_3。它們是否趨近 $\frac{1}{9}$？

4 **姓氏。**如果每個男人有 0、1、2 或 3 個兒子的概率各為 $\frac{1}{4}$，他的姓氏最終失傳的概率是多少？

延伸閱讀

- **Geoffrey Grimmett & David Stirzaker, *Probability and Random Processes***（Oxford University Press），論分支過程的一節。
- **Adam Kucharski, *The Rules of Contagion***（Profile）。講述繁殖數與疫症爆發，適合一般讀者。
- **David Kendall, "Branching processes since 1873"**，*Journal of the London Mathematical Society* 41 (1966)。講述這段歷史，包括 Bienaymé 的貢獻。

12.16 順向與逆向的馬可夫鏈

題目

STEP III 1996，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/96-s3-q12`

引子

Ecks 教授的定理有一個規律：她下一個證明甚麼，只取決於她上一個證明了甚麼。一個過程的未來若只取決於現在而與過去無關，就稱為**馬可夫鏈**（*Markov chain*）。馬可夫鏈無處不在：天氣模型、蛇梯棋這種圖版遊戲、DNA 模型、Google 最初的 PageRank 演算法，以至統計學家用來模擬複雜概率分佈的方法，全都用得上它。

本題要求出各類定理在長遠而言所佔的比例，即**平穩分佈**（*stationary distribution*）：長遠來說，她的定理每六個之中有一個是新而正確的，而有一半是錯的。然後題目把這條鏈在時間上倒過來運行，並問：已知今天的定理，昨天的定理是哪一類？

動筆之前

? **想一想** 「每一行（row）之和為 1」這個條件是甚麼意思？各列（column）之和是否也是 1？

第 i 行列出的是：已知現時的類型為 i，下一個定理屬於各類型的概率；三者必有其一發生，所以每一行之和為 1。各列則不一定：第二列之和恰好是 1，但第一列之和是 0.6，第三列之和是 1.4。

逐步解題

階段 1 **一步**

根據全概率定律（law of total probability），以現時定理的類型 i 為條件，

$$b_j = \sum_i \mathrm{P}(\text{現時}i)\,\mathrm{P}(\text{下一個}j \mid \text{現時}i) = a_1p_{1j} + a_2p_{2j} + a_3p_{3j}. \qquad \blacksquare$$

用矩陣的語言來說，行向量（row vector）$\mathbf{b} = \mathbf{a}P$。

階段 2 平穩分佈

想一想 我們要求 $\mathbf{a} = \mathbf{a}P$，且 $a_1 + a_2 + a_3 = 1$。把各方程寫出來。哪些最簡單？

各方程為

$$\begin{aligned} a_1 &= 0.3a_1 + 0.2a_2 + 0.1a_3, \\ a_2 &= 0.3a_1 + 0.4a_2 + 0.3a_3, \\ a_3 &= 0.4a_1 + 0.4a_2 + 0.6a_3. \end{aligned}$$

（它們並非獨立：把三式相加，兩邊都得 $a_1 + a_2 + a_3$。）由第二式得 $0.6a_2 = 0.3(a_1 + a_3) = 0.3(1 - a_2)$，所以 $a_2 = \frac{1}{3}$。由第一式得 $7a_1 = 2a_2 + a_3 = \frac{2}{3} + a_3$，而 $a_1 + a_3 = \frac{2}{3}$，所以 $8a_1 = \frac{4}{3}$，$a_1 = \frac{1}{6}$，$a_3 = \frac{1}{2}$：

$$(a_1, a_2, a_3) = \left(\tfrac{1}{6}, \tfrac{1}{3}, \tfrac{1}{2}\right). \qquad \blacksquare$$

階段 3 把鏈倒過來運行

想一想 $q_{ij} = \mathrm{P}(\text{上一個}i \mid \text{現時}j)$。運用貝葉斯定理（Bayes' theorem），上一個定理和現時的定理都採用平穩分佈。

根據貝葉斯定理，

$$q_{ij} = \frac{\mathrm{P}(\text{上一個}i)\,\mathrm{P}(\text{現時}j \mid \text{上一個}i)}{\mathrm{P}(\text{現時}j)} = \frac{a_ip_{ij}}{a_j}.$$

例如 $q_{11} = \frac{\frac{1}{6}\times 0.3}{\frac{1}{6}} = 0.3$ 及 $q_{21} = \frac{\frac{1}{3}\times 0.2}{\frac{1}{6}} = 0.4$。全部結果如下（以 i 標示行，j 標示列）：

$$(q_{ij}) = \begin{pmatrix} 0.3 & 0.15 & \frac{2}{15} \\ 0.4 & 0.4 & \frac{4}{15} \\ 0.3 & 0.45 & 0.6 \end{pmatrix}. \qquad \blacksquare$$

現在是每一列之和為 1，因為第 j 列列出的是上一個定理各種可能的類型。

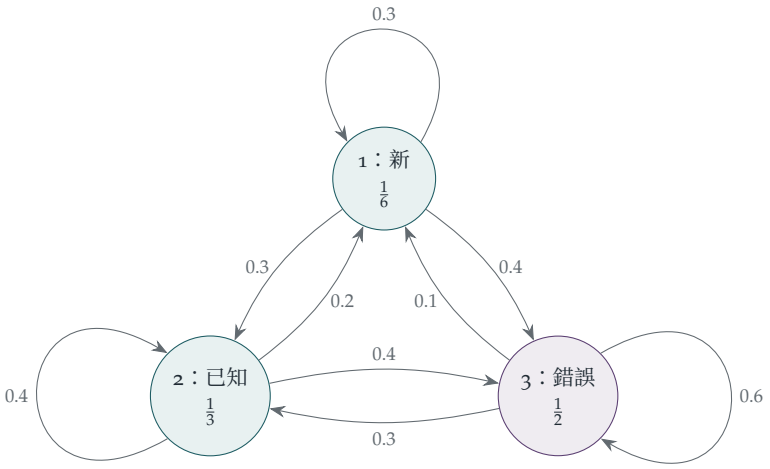


圖 12.16.1　把 Ecks 教授的定理看成一條馬可夫鏈。箭頭顯示下一個定理屬於各類型的概率。圓圈內的分數是平穩分佈：長遠來說，每六個定理有一個是新而正確的，每三個有一個是已知的，而有一半是錯的。

回顧

- **以矩陣表達全概率定律。**鏈每走一步，就是把一個行向量乘以矩陣 P。
- **平穩即特徵向量（eigenvector）。**$\mathbf{a} = \mathbf{a}P$ 說明 $\mathbf{a}$ 是 P 的左特徵向量，對應的特徵值（eigenvalue）為 1。
- **用貝葉斯定理倒推。**逆向概率構成另一個隨機矩陣（stochastic matrix），由正向的隨機矩陣和平穩分佈構造而成。

背後的數學

收斂至平衡

讓 Ecks 教授從任何一個定理開始；再證明 n 個之後，類型的分佈是 $\mathbf{a}_0 P^n$。對於像這樣的一條鏈，即每種類型都可以緊接在任何一種類型之後，不論起點為何，P^n 都會收斂至一個每一行均等於平穩分佈 $(\frac{1}{6}, \frac{1}{3}, \frac{1}{2})$ 的矩陣。收斂的速度由 P 的其他特徵值（這裏是 0.2 和 0.1）決定，它們以幾何方式縮小：只需幾個定理之後，她的過去就被遺忘了。這個關於馬可夫鏈的*遍歷定理*（*ergodic theorem*）是大學二年級概率課程的核心結果，馬可夫鏈蒙地卡羅（Markov chain Monte Carlo）方法之所以行得通，靠的正是它：要從一個複雜的分佈中抽樣，就設計一條以該分佈為平穩分佈的馬可夫鏈，然後讓它運行一段長時

間。

可逆性

逆向鏈 (q_{ij}) 與正向鏈 (p_{ij}) 並不相同：從她的定理的一份長紀錄中，你可以分辨出時間的方向。若一條鏈的逆向版本與正向版本一致，即對所有 i、j 都有 $a_i p_{ij} = a_j p_{ji}$，就稱為*可逆*（*reversible*）；網絡上的隨機遊走（random walk）以及處於熱平衡的物理系統的模型都屬此類。可逆性是一個威力強大的性質，因為它令平穩分佈很容易求出。

馬可夫與普希金

馬可夫（Andrey Markov）在 1906 年引入他的鏈，部分原因是要證明大數定律（law of large numbers）並不需要獨立性；當時他正與數學家 Pavel Nekrasov 爭論，後者聲稱獨立性是必需的。1913 年，馬可夫把他的想法應用於普希金（Pushkin）的詩體小說《葉甫蓋尼·奧涅金》（*Eugene Onegin*）的首 20 000 個字母，把每個字母分為元音或輔音，並證明每個字母的類型取決於前一個字母的類型，而這種關係可用一條兩個狀態的鏈很好地描述。這是史上第一次對文本作統計分析，也是所有預測輸入系統的始祖。

思維技巧

概率作向量，轉移作矩陣

當一個系統以固定的概率在各狀態之間轉移時，把分佈寫成向量，把轉移寫成矩陣。這樣，長遠行爲就變成一個關於特徵向量的問題。

在大學裏，這樣做會把第 7 章的線性代數帶進概率論。佩龍－弗羅賓尼斯定理（Perron–Frobenius theorem）保證平穩分佈存在，而其他特徵值則量度到達平穩分佈有多快。

延伸練習

1 **兔子窩。**試做 STEP III 2003，第 13 題（`step.maths.org/questions/03-s3-q13`）：在一個正方形的四個頂點上的隨機遊走，每個頂點各有一個出口，以首步方程求解。

2 **兩步。**若 Ecks 教授上一個定理是新而正確的，求再下一個定理屬於三種類型的概率。

3 **特徵值。**證明 P 的特徵值為 1、0.2 和 0.1。

4 **可逆嗎？**檢查 Ecks 教授的鏈是否滿足 $a_i p_{ij} = a_j p_{ji}$。答案告訴你甚麼？

延伸閱讀

- **James Norris, *Markov Chains***（Cambridge University Press）。標準的大學本科教科書。
- **Olle Häggström, *Finite Markov Chains and Algorithmic Applications***（Cambridge University Press）。篇幅短小，講馬可夫鏈蒙地卡羅尤其出色。
- **Brian Hayes,「First links in the Markov chain」**，*American Scientist* 101 (2013)。重述馬可夫對普希金作品的分析。

12.17 巴拿赫的火柴盒

題目

STEP II 1992，第 16 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/92-s2-q16`

引子

相傳數學家巴拿赫（Stefan Banach）煙癮很大，兩邊口袋各放一盒火柴。每次要用火柴，他就隨機選一個口袋。有一天，他發現其中一盒空了。另一盒還剩多少根火柴？這就是巴拿赫火柴盒問題（Banach's matchbox problem），由斯坦豪斯（Hugo Steinhaus）為向巴拿赫致意而設計；本題則是的士司機的版本，把火柴換成了糖果。

微妙之處在於「發現」一詞。司機取走一包糖的最後一顆時，並不會察覺那包已經空了，要到下一次伸手去拿那包糖時才知道。所以這個事件是：她伸手去拿太妃糖，而全部 n 顆太妃糖早已吃完。計算這件事在第 N 次伸手時發生的方法數目，是二項分佈（binomial distribution）的一道巧妙練習；而巴拿赫問題的答案有一個出人意表的特點：另一盒最可能剩下的火柴數目是 0（與 1 並列）。

動筆之前

想一想 若司機在第 N 次伸手時發現太妃糖已經吃完，第 N 次伸手必須是甚麼？之前的 $N-1$ 次伸手又必須包含甚麼？

第 N 次伸手必須是去拿太妃糖，而之前的 $N-1$ 次伸手必須剛好取走了 n 顆太妃糖（即全部）和 $N-1-n$ 顆薄荷糖。此外，她在此之前不可以已經發現薄荷糖空了。

逐步解題

階段 1　發現太妃糖吃完了

> **想一想**　頭 $N-1$ 次伸手可以按甚麼次序出現？$N-1-n$ 要符合甚麼條件，才能確保她從未發現那包薄荷糖是空的？

那 n 次拿太妃糖的伸手，可以是頭 $N-1$ 次伸手中的任何 n 次。只要她最多取了 m 顆薄荷糖，即 $N-1-n \le m$，她就從未在薄荷糖吃完之後再伸手去拿薄荷糖。因此，對於 $n+1 \le N \le n+m+1$，

$$\mathrm{P}(T_N) = \binom{N-1}{n} p^{n+1} q^{N-1-n},$$

其中 T_N 是事件「她在第 N 次伸手時發現太妃糖吃完了」（因子 p^{n+1} 已包括第 N 次伸手本身）。

階段 2　發現薄荷糖吃完了

把兩者的角色對調，用同樣的論證可得，對於 $m+1 \le N \le n+m+1$，

$$\mathrm{P}(M_N) = \binom{N-1}{m} q^{m+1} p^{N-1-m}.$$

階段 3　條件概率

> **想一想**　已知在第 N 次伸手時有所發現。那次拿的是太妃糖的概率是多少？

事件 T_N 與 M_N 互斥（exclusive），所以

$$\mathrm{P}(\text{太妃糖} \mid \text{有所發現的伸手次數為} N) = \frac{\binom{N-1}{n} p^{n+1} q^{N-1-n}}{\binom{N-1}{n} p^{n+1} q^{N-1-n} + \binom{N-1}{m} q^{m+1} p^{N-1-m}},$$

其中二項式係數（binomial coefficient）在上述範圍以外一律視為 0。（若 $N \le m$，則只有太妃糖可能已經吃完，概率為 1；同樣，若 $N \le n$ 但 $N > m$，概率為 0。）∎

> **停一停，試一試**　當 $n = m$ 且 $p = q = \frac{1}{2}$ 時，驗算一下這條公式。根據對稱性，對每一個可能的 N，答案都應該是 $\frac{1}{2}$。是否如此？

回顧

- **準確理解事件。**「發現」是指伸手去拿一包已空的糖，而不是取走最後一顆糖。

- **固定最後一步，再數其餘。**第 N 次伸手已經固定；頭 $N-1$ 次可以按任何次序出現，於是得出一個二項式係數。
- **檢查邊界。**限制條件 $N-1-n \le m$ 和 $N-1-m \le n$ 排除了不可能出現的過程。

背後的數學

巴拿赫的火柴盒

在巴拿赫的版本中，每盒起初有 n 根火柴，選中每個口袋的機會均等。當他首次發現其中一盒空了，設 K 為另一盒的火柴數目。若 $K=k$，他必定已從空的那盒取了 n 根，從另一盒取了 $n-k$ 根，次序不拘，然後再一次伸手去拿空的那盒；而兩盒之中任何一盒都可以是空的那盒：

$$\mathrm{P}(K=k) = 2\binom{2n-k}{n}\left(\frac{1}{2}\right)^{2n-k+1} = \binom{2n-k}{n}\left(\frac{1}{2}\right)^{2n-k}.$$

這些概率隨 k 增大而遞減，唯一例外是 $\mathrm{P}(K=0)=\mathrm{P}(K=1)$（圖 12.17.1）。剩餘火柴數目的期望值（expected value）約為 $2\sqrt{n/\pi}-1$：若每盒有 50 根火柴，約為 7。這個 $\sqrt{n}$ 是經過 $2n$ 次隨機選擇之後，兩盒之間差距的典型大小，與生日問題（Section 9.9）和隨機漫步（random walk）中的平方根定律如出一轍。

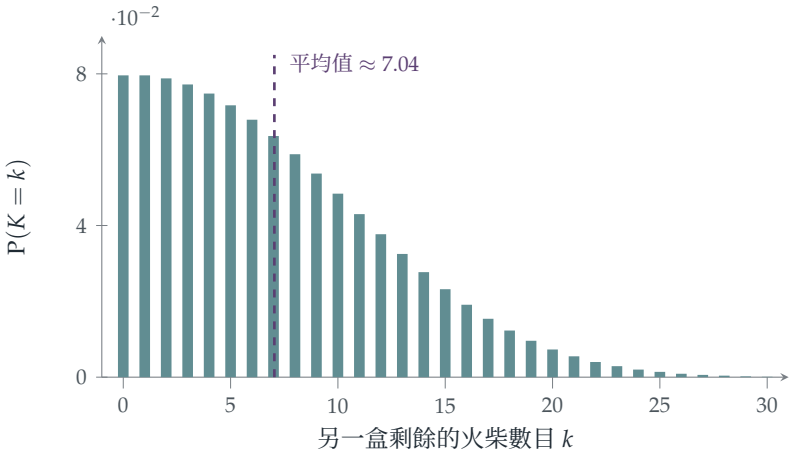


圖 12.17.1 每盒有 50 根火柴的巴拿赫火柴盒問題：首次發現其中一盒空了的時候，另一盒剩餘火柴數目的分佈。最可能的數值是 0 和 1，兩者概率相同，但平均值約為 7。

負二項分佈

在一連串獨立試驗（independent trials）中，每次試驗成功的概率為 p，則取得第 r 次成功所需的試驗次數服從負二項分佈（negative binomial distribution）：$\mathrm{P}(N=k)=$

$\binom{k-1}{r-1}p^r q^{k-r}$。的士司機的問題，是兩個負二項過程（太妃糖和薄荷糖）之間的競賽，每個過程所需的「成功」次數都比該包糖的數目多一次。負二項分佈見於保險（索償宗數）、生態學（動物聚集分佈的數目）和遺傳學，往往用來代替泊松分佈（Poisson distribution），因為它更有彈性。

巴拿赫與蘇格蘭咖啡館

巴拿赫是泛函分析（functional analysis）的奠基人之一，他的數學研究有不少是在 Lwów（今稱 Lviv）的蘇格蘭咖啡館（Scottish Café）裏完成的。他和同事把問題寫在一本由侍應保管的筆記簿上，這就是《蘇格蘭書》（Scottish Book），解題獎品從一杯啤酒到一隻活鵝不等。斯坦豪斯是巴拿赫的恩師，他為了一篇向巴拿赫的吸煙習慣致意的演說，設計了這道火柴盒問題。最先在書刊中分析這個問題的，是費勒（William Feller）1950 年出版的著名概率論教科書。

思維技巧

固定最終事件，數出通往它的路徑

要求一個過程在某一時刻以某種方式結束的概率，可先固定最後一步發生的事，再數出通往這一步的所有過程。

在大學裏，這是負二項分佈、隨機漫步的首達時間（first-passage time），以及 Section 9.10 中反射原理（reflection principle）的基礎。

延伸練習

1 **小規模情況。**設 $n = m = 1$ 及 $p = \frac{1}{2}$，列出所有可能的序列，並求 N 的分佈。

2 **巴拿赫公式。**對 $n = 1$ 和 $n = 2$，驗證 $\sum_{k=0}^{n} \binom{2n-k}{n} 2^{-(2n-k)} = 1$。

3 **並列最高。**證明在巴拿赫問題中，對每一個 n 都有 $\mathrm{P}(K = 0) = \mathrm{P}(K = 1)$。

4 **哪一包先吃完？**就的士司機的問題，當 $n = m = 2$ 及 $p = \frac{2}{3}$ 時，求太妃糖是首先被發現空了的那一包（不論何時發現）的概率。

延伸閱讀

► **William Feller, *An Introduction to Probability Theory and Its Applications*, Vol. 1** (Wiley)，

第 VI 章，正是這本書把巴拿赫問題介紹給全世界。

- **R. Daniel Mauldin (ed.), *The Scottish Book***（Birkhäuser）。收錄蘇格蘭咖啡館的問題，並附評註。
- **Geoffrey Grimmett & Dominic Welsh, *Probability: An Introduction***（Oxford University Press），內容涉及負二項分佈。

12.18 無記憶性

題目

STEP III 2021，第 11 題

繼續閱讀之前，請先看題目（原題為英文）。題目載於 OCR 提供下載的 2021 年 STEP 試卷（zip 檔案）內的 STEP III 試卷：

`www.ocr.org.uk/Images/696588-2021-past-papers.zip`

引子

一顆放射性原子在實驗室裏待了一百萬年，一直沒有衰變。它是否「快要」衰變了？不是：它在下一秒內衰變的機會，與一顆剛剛產生的原子完全相同。放射性原子不會老化。直至衰變為止的等候時間服從指數分佈（*exponential distribution*），而這個分佈的決定性特徵，就是它沒有記憶。

這道題以一種令人印象深刻的形式展示了這種無記憶性（memorylessness）。把一個指數等候時間拆成整數個單位和小數部分。原來這兩部分是獨立（*independent*）的：知道你等了多少個完整的小時，對於分鐘數毫無提示。這道題還會在練習中解釋，為甚麼你等巴士的時間，可以顯得比兩班巴士之間的間隔還要長。

動筆之前

想一想 對於 $x \geq 0$，$\mathrm{P}(X > x)$ 是甚麼？$\mathrm{P}(X > s+t \mid X > s)$ 又是甚麼？

$\mathrm{P}(X > x) = \int_x^\infty \lambda e^{-\lambda u}\,\mathrm{d}u = e^{-\lambda x}$。於是 $\mathrm{P}(X > s+t \mid X > s) = \frac{e^{-\lambda(s+t)}}{e^{-\lambda s}} = e^{-\lambda t} = \mathrm{P}(X > t)$：已經等了 s 之後，餘下的等候時間與從頭開始等候的分佈相同。這就是無記憶性質（*memoryless property*）。

逐步解題

階段 1　整數部分

$Y = n$ 當且僅當 $n \le X < n+1$：

$$\mathrm{P}(Y = n) = e^{-\lambda n} - e^{-\lambda(n+1)} = (1 - e^{-\lambda})e^{-n\lambda}. \qquad \blacksquare$$

所以 Y 服從幾何分佈（geometric distribution）：每一個完整單位都以概率 $e^{-\lambda}$ 被「捱過」。

階段 2 小數部分

想一想 $Z < z$ 表示 X 落在其中一個區間 $[n, n+z)$ 內。把這些區間的概率加起來。

$$\mathrm{P}(Z < z) = \sum_{n \ge 0} \left(e^{-\lambda n} - e^{-\lambda(n+z)}\right) = (1 - e^{-\lambda z}) \sum_{n \ge 0} e^{-\lambda n} = \frac{1 - e^{-\lambda z}}{1 - e^{-\lambda}}. \qquad \blacksquare$$

階段 3 Z 的平均值

想一想 對於取值於 $[0,1]$ 的隨機變量（random variable），$\mathrm{E}(Z) = \int_0^1 \mathrm{P}(Z > z)\,\mathrm{d}z$。運用這條公式。

設 $c = 1 - e^{-\lambda}$，則 $\mathrm{P}(Z > z) = 1 - \frac{1-e^{-\lambda z}}{c}$，所以

$$\mathrm{E}(Z) = 1 - \frac{1}{c}\left(1 - \frac{1 - e^{-\lambda}}{\lambda}\right) = 1 - \frac{1}{c} + \frac{1}{\lambda} = \frac{1}{\lambda} - \frac{e^{-\lambda}}{1 - e^{-\lambda}} = \frac{1}{\lambda} - \frac{1}{e^{\lambda} - 1}. \qquad \blacksquare$$

（當 λ 很小，即等候時間很長時，這個值趨向 $\frac{1}{2}$：小數部分幾乎是均勻分佈的。當 λ 很大時，它接近 $\frac{1}{\lambda}$，因為 X 幾乎總是小於 1。）

階段 4 獨立性

$$\mathrm{P}(Y = n, z_1 < Z < z_2) = \mathrm{P}(n + z_1 < X < n + z_2) = e^{-\lambda n}\left(e^{-\lambda z_1} - e^{-\lambda z_2}\right).$$

想一想 利用第 (i) 及第 (ii) 部分，與 $\mathrm{P}(Y = n)\,\mathrm{P}(z_1 < Z < z_2)$ 作比較。

由 (i) 及 (ii)，$\mathrm{P}(Y = n)\mathrm{P}(z_1 < Z < z_2) = (1 - e^{-\lambda})e^{-n\lambda} \cdot \frac{e^{-\lambda z_1} - e^{-\lambda z_2}}{1 - e^{-\lambda}}$，兩者相同。所以 Y 與 Z 獨立。 $\blacksquare$

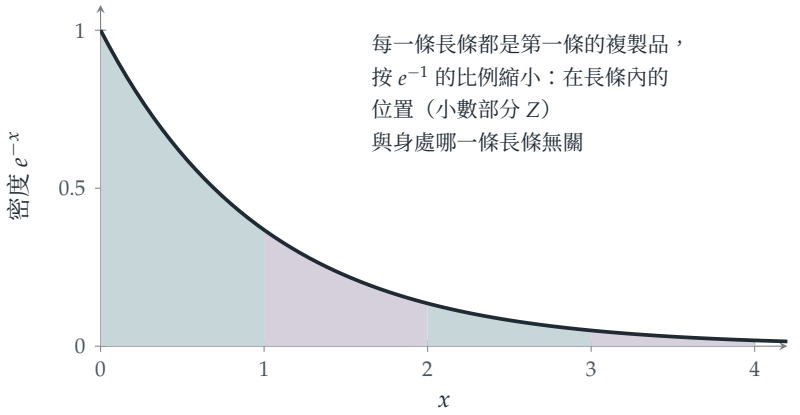


圖 12.18.1 $\lambda = 1$ 的指數密度，分割成單位寬度的長條。由於 $e^{-(n+z)} = e^{-n}e^{-z}$，每一條長條的形狀都相同。$X$ 落在哪一條長條（整數部分 Y），與它落在長條內的甚麼位置（小數部分 Z），兩者是獨立的。

回顧

- **逐段求和。**Z 的分佈和聯合概率（joint probability），都是把各區間 $[n, n+1)$ 上的概率相加而得的。
- **尾部公式。**$\mathrm{E}(Z) = \int \mathrm{P}(Z > z)\,\mathrm{d}z$ 讓我們不必求出密度。
- **能夠分解即是獨立。**聯合概率分拆成一個 n 的函數乘以一個 z_1、z_2 的函數。

背後的數學

指數分佈為何特別

Y 與 Z 的獨立性直接來自 $e^{-\lambda(n+z)} = e^{-\lambda n}e^{-\lambda z}$，而這正是無記憶性質的另一副面貌。事實上，指數分佈是唯一沒有記憶的連續分佈：若對所有 $s, t \geq 0$ 都有 $\mathrm{P}(X > s+t) = \mathrm{P}(X > s)\mathrm{P}(X > t)$，則 $g(t) = \ln \mathrm{P}(X > t)$ 滿足 $g(s+t) = g(s) + g(t)$，而（對於單調（monotonic）的 g）這就迫使 $g(t) = -\lambda t$。正因如此，指數分佈可用來描述那些「隨機」發生、沒有任何累積過程的事件的等候時間：放射性衰變、顧客的到達、不會損耗的電子元件的故障。

最小值與等候時間悖論

若 X_1 與 X_2 是獨立的指數隨機變量，率分別為 λ_1 和 λ_2，則 $\mathrm{P}(\min > x) = e^{-\lambda_1 x}e^{-\lambda_2 x}$，所以最小值服從率為 $\lambda_1 + \lambda_2$ 的指數分佈。等候兩個獨立隨機事件中最先發生的一個，

就如同以合併後的率等候單一事件（STEP III 2002 Q13，見練習）。無記憶性也解釋了等候時間悖論（*waiting-time paradox*）：若巴士隨機到站，平均間隔為 10 分鐘，則一位在隨機時刻到達的乘客平均要等 10 分鐘，而不是 5 分鐘，因為餘下的等候時間是無記憶的；而乘客到達時所處的那個間隔，平均長達 20 分鐘，因為較長的間隔更容易被碰上。

盧瑟福（Rutherford）與泊松過程（Poisson process）

1910 年，歐內斯特·盧瑟福與漢斯·蓋革（Hans Geiger）以 7.5 秒為一個時段，數算一個放射源所放出的阿爾法粒子（alpha particles），發現這些計數服從泊松分佈（Poisson distribution）；數學家 Harry Bateman 證明了，只要各次放射以恆定的率獨立發生，結果就理應如此。這樣一來，相鄰兩次放射之間的時間便服從指數分佈。泊松過程，即間隔服從指數分佈的一連串事件，如今已成為描述時間上隨機性的基本模型，應用於電話通話、保險索償以及神經元的放電；下一節 Section 12.19 會把它推廣到空間。

思維技巧

以因式分解檢驗獨立性

若兩個隨機變量的聯合概率可分解爲其中一個變量的函數乘以另一個變量的函數，則這兩個隨機變量獨立。

在大學裏，聯合密度（joint density）能否分解是檢驗獨立性的標準方法，你亦會藉此在多變量問題中辨認出獨立的分量，例如 Section 12.22 中服從正態分佈（normal distribution）的點的半徑與角度。

延伸練習

1 **兩條巴士路線。**試做 STEP III 2002，第 13 題（`step.maths.org/questions/02-s3-q13`）。這道題利用兩個指數隨機變量的最小值，求出等候下一班巴士的時間。

2 **離散的無記憶性。**證明幾何分佈是無記憶的：$\mathrm{P}(Y \geq m+n \mid Y \geq m) = \mathrm{P}(Y \geq n)$。

3 **等候時間悖論。**巴士的到站是一個率為 λ 的泊松過程。證明包含某固定時刻 t（t 很大）的那個間隔，其期望長度為 $\frac{2}{\lambda}$，即平均間隔的兩倍。

4 Z **的密度。**求 Z 的概率密度（probability density），並就 $\lambda = 1$ 和 $\lambda = 5$ 繪出其草圖。

延伸閱讀

- **Geoffrey Grimmett & Dominic Welsh, *Probability: An Introduction***（Oxford University Press），論及指數分佈與泊松過程。
- **Sheldon Ross, *Introduction to Probability Models***（Academic Press）。指數分佈、泊松過程與排隊。
- **David Spiegelhalter, *The Art of Uncertainty***（Pelican）。日常生活中的隨機性與等候。

12.19 平面上的隨機點

題目

STEP II 2012，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/12-s2-q13`

引子

假如超級市場、樹木、恆星或流動電話發射站「隨機」散佈在某個區域內，你要走多遠才到達最近的一個？答案只取決於密度 k（每單位面積的平均數目），而且有一個獨特的形狀，就是瑞利分佈（*Rayleigh distribution*）。平均而言，最近的一個位於距離 $\frac{1}{2\sqrt{k}}$ 處：若每平方公里有一間超級市場，這距離就是 500 米。

這個模型稱為空間泊松過程（*spatial Poisson process*），即時間上隨機到達（Section 12.18）的二維版本。計算過程簡短而優美，但它真正的價值在於充當基準：生態學家把實際的最近鄰距離與這條公式比較，以判斷樹木或鳥巢究竟是聚集成羣、隨機散佈，還是刻意彼此隔開。

動筆之前

想一想 一個平均值為 μ 的泊松隨機變量（Poisson random variable），取值 0 的概率是多少？

$\mathrm{P}(0) = e^{-\mu}$。半徑為 y 的圓面積是 πy^2，所以圓內超級市場數目的平均值為 $k\pi y^2$。

逐步解題

階段 1　附近沒有超級市場

在距離 y 以內沒有超級市場的概率是 $e^{-\pi k y^2}$。

階段 2 最近距離的分佈

? **想一想** $Y \geq y$ 表示在距離 y 以內沒有超級市場。那麼 $\mathrm{P}(Y < y)$ 是甚麼？

$\mathrm{P}(Y < y) = 1 - e^{-\pi k y^2}$，求導後便得到密度函數（density）

$$\mathrm{f}(y) = 2\pi k y e^{-\pi k y^2}, \qquad y \geq 0. \qquad \blacksquare$$

階段 3 平均值

? **想一想** 可以用分部積分法（integration by parts）計算 $y\,\mathrm{f}(y)$ 的積分，也可以利用 $\mathrm{E}(Y) = \int_0^\infty \mathrm{P}(Y > y)\,\mathrm{d}y$。然後作代換，以便使用題目給出的高斯積分（Gaussian integral）。

$$\mathrm{E}(Y) = \int_0^\infty e^{-\pi k y^2}\,\mathrm{d}y.$$

代入 $x = y\sqrt{2\pi k}$，使得 $\pi k y^2 = \frac{1}{2}x^2$，便得 $\mathrm{E}(Y) = \frac{1}{\sqrt{2\pi k}}\sqrt{\frac{\pi}{2}} = \frac{1}{2\sqrt{k}}$。 $\blacksquare$

階段 4 方差

$$\mathrm{E}(Y^2) = \int_0^\infty y^2 \cdot 2\pi k y e^{-\pi k y^2}\,\mathrm{d}y.$$

代入 $u = \pi k y^2$（因此 $\mathrm{d}u = 2\pi k y\,\mathrm{d}y$），上式變成 $\int_0^\infty \frac{u}{\pi k} e^{-u}\,\mathrm{d}u = \frac{1}{\pi k}$。因此

$$\mathrm{Var}(Y) = \frac{1}{\pi k} - \frac{1}{4k} = \frac{4-\pi}{4\pi k}. \qquad \blacksquare$$

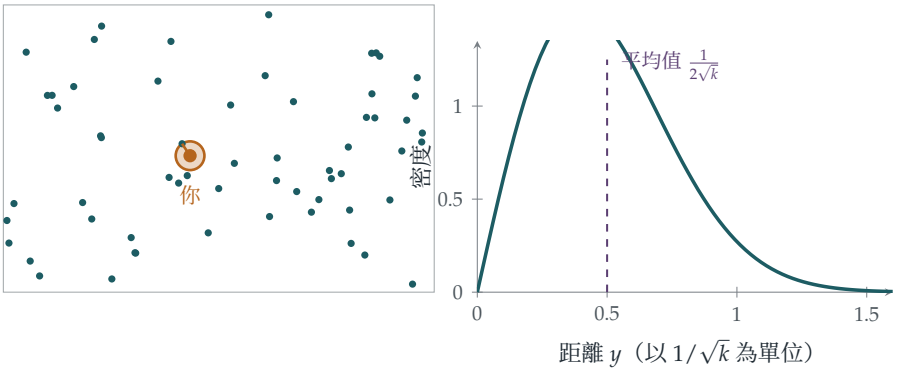


圖 12.19.1　左：六十個點隨機散佈在面積為 24 的區域內（所以 $k = 2.5$），以及由選定的一點到其中最近一點的距離（琥珀色）。右：該距離的瑞利密度函數 $2\pi k y e^{-\pi k y^2}$，距離以 $1/\sqrt{k}$ 為單位。其平均值為 $\frac{1}{2\sqrt{k}}$。

回顧

- **把距離問題化為計數問題。**「最近的一個比 y 更遠」就是「在 y 以內一個也沒有」。
- **先求累積分佈，再求密度。**$\mathrm{P}(Y < y)$ 可以直接寫出；求導後便得到密度函數。
- **以代換化為已知積分。**每個矩（moment）都化成了高斯積分或指數積分。

背後的數學

平面上的泊松過程

強度（intensity）為 k 的空間泊松過程，是一種隨機散佈的點集，滿足：互不相交的區域內的點數彼此獨立，而面積為 A 的區域內的點數服從平均值為 kA 的泊松分佈。把極大量的點獨立而均勻地投放到一片極大的區域上，並保持適當的平均密度，得到的就是這種過程：這正是「稀有事件定律」（law of rare events），它同樣解釋了錯排（derangement）的現象（Section 9.8）。它是完全空間隨機性（complete spatial randomness）的自然模型，實際的分佈形態都以它為比較標準。*Clark–Evans 檢驗*（*Clark–Evans test*）由兩位生態學家於 1954 年提出，把觀察所得的平均最近鄰距離與 $\frac{1}{2\sqrt{k}}$ 比較：小得多表示有聚集現象（種子落在母樹附近），大得多則表示間隔規則（有領域行為的鳥類，或植物之間爭奪水分）。

三維的情況

在三維空間中作同樣的計算，可得到最近的恆星、分子或星系的距離滿足 $\mathrm{P}(Y > y) = e^{-\frac{4}{3}\pi k y^3}$。錢德拉塞卡（Subrahmanyan Chandrasekhar）在 1943 年用它來研究一顆恆星受隨機分佈的鄰近恆星所施的引力，物理學家則用它來計算氣體分子的平均自由程（mean free path）。

瑞利

密度函數 $\frac{y}{\sigma^2} e^{-y^2/(2\sigma^2)}$ 以瑞利勳爵（Lord Rayleigh）命名，他在 1880 年研究聲學理論時遇上它：它描述了許多相位隨機的波疊加後的振幅。若一點的兩個坐標是獨立的正態隨機變量（normal random variable），它也給出該點與原點距離的分佈，這是理解它為何在這裏出現的另一個角度（練習 3）。今天，它用來模擬流動電話訊號的衰落，因為訊號是沿着許多相位隨機的路徑到達接收器的。

思維技巧

先求分佈函數

要求最小值、最大值或最近距離的分佈時，$\mathrm{P}(Y > y)$ 往往很容易直接寫出。求導便得到密度函數。

這個方法會在 Section 12.25 再次出現，用來處理許多隨機壽命中最長的一個。在大學裏，這是處理任何「順序統計量」（order statistic）的標準第一步。

延伸練習

1 **中位數。**求到最近超級市場的距離的中位數（median），並與平均值比較。

2 **第二近的。**求到第二近的超級市場的距離的分佈。（提示：$\mathrm{P}(Y_2 > y)$ 是在 y 以內最多只有一間超級市場的概率。）

3 **正態坐標。**若 U 和 V 是獨立的 $\mathrm{N}(0, \sigma^2)$ 隨機變量，證明 $\sqrt{U^2 + V^2}$ 的密度函數為 $\frac{r}{\sigma^2} e^{-r^2/(2\sigma^2)}$。$\sigma$ 取何值時會得到超級市場的分佈？

4 **Clark–Evans。**在一項調查中，一片面積 1 km^2 的樹林內有 400 棵樹，而一棵樹到其最近鄰的平均距離為 35 m。這與完全空間隨機性相符嗎？還是顯示有聚集或間隔的現象？

延伸閱讀

- **Peter Diggle, *Statistical Analysis of Spatial Point Patterns***（Arnold）。這方面的標準參考書，載有 Clark–Evans 檢驗及其後續發展的方法。
- **John Kingman, *Poisson Processes***（Oxford University Press）。一本優美的短篇著作，討論任意維數的泊松過程。
- **Geoffrey Grimmett & David Stirzaker, *Probability and Random Processes***（Oxford University Press），論泊松過程的一章。

12.20 蒲豐投針

題目

STEP I 1993，第 16 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/93-s1-q16`

引子

把一根針拋到鋪有平行地板條的地板上。針橫跨板縫的機會竟然與 π 有關，儘管眼前根本沒有任何圓形。拋很多次並數一數橫跨板縫的次數，就可以用實驗來估算 π。這就是*布豐投針問題*（*Buffon's needle*），由布豐伯爵喬治-路易·勒克萊爾（Georges-Louis Leclerc, Comte de Buffon）於 1733 年提出，並於 1777 年由他本人解決：它是*幾何概率*（*geometric probability*）的第一個問題，也是所有蒙地卡羅方法（Monte Carlo method）的始祖。

這個 π 來自角度：針的方向在半圈之內均勻分佈（uniform），把針橫跨平行線方向的「影子」長度對所有方向取平均，便得出因子 $\frac{2}{\pi}$。本題則以 $\cos^{-1}$ 的積分來完成這個計算。

動筆之前

想一想 針的中心與最近一條直線的距離為 x，而針與這些直線的垂直線所成的角為 y。針朝着該直線能伸展多遠？

針的每一半長度為 a，它在垂直於直線方向上的分量是 $a\cos y$。因此，針與最近的直線相交，當且僅當 $a\cos y \geq x$。（由於 $a < b$，針不可能觸及其他直線。）

逐步解題

階段 1 積分

設 $t = \cos y$，當 t 由 0 變到 1 時，$y = \cos^{-1} t$ 由 $\frac{\pi}{2}$ 變到 0，而 $\mathrm{d}t = -\sin y\,\mathrm{d}y$。所以

$$\int_0^1 \cos^{-1} t\,\mathrm{d}t = \int_0^{\pi/2} y \sin y\,\mathrm{d}y = [-y\cos y]_0^{\pi/2} + \int_0^{\pi/2} \cos y\,\mathrm{d}y = 0 + 1 = 1. \qquad \blacksquare$$

（或者：這個積分是 $\cos^{-1}$ 的圖像與兩條坐標軸之間的面積，它等於 $y = \cos x$ 由 0 到 $\frac{\pi}{2}$ 的曲線下面積，兩者只是沿 $y = x$ 反射而已。）

階段 2　給定位置

想一想　對於固定的 x，哪些角度 y 會使針與直線相交？這些角度的概率是多少？

若 $x \geq a$，針無法觸及直線。若 $x < a$，相交的條件是 $\cos y \geq \frac{x}{a}$，即 $y \leq \cos^{-1}\frac{x}{a}$。由於 Y 在 $[0, \frac{\pi}{2}]$ 上均勻分佈，

$$\mathrm{P}(\text{相交} \mid X = x) = \frac{2}{\pi}\cos^{-1}\frac{x}{a} \qquad (0 \leq x < a). \qquad \blacksquare$$

階段 3　對位置取平均

X 在 $[0, b]$ 上均勻分佈，其密度（density）為 $\frac{1}{b}$，所以

$$\mathrm{P}(\text{相交}) = \frac{1}{b}\int_0^a \frac{2}{\pi}\cos^{-1}\frac{x}{a}\,\mathrm{d}x = \frac{2a}{\pi b}\int_0^1 \cos^{-1} t\,\mathrm{d}t = \frac{2a}{\pi b}, \qquad \blacksquare$$

其中用了代換 $x = at$。

停一停，試一試　若針的長度恰好等於半間距（$a = \frac{b}{2}$，即針的長度 $2a$ 等於 b），相交的概率是 $\frac{1}{\pi}$。要把 π 估算至小數點後兩位，你需要拋多少次？

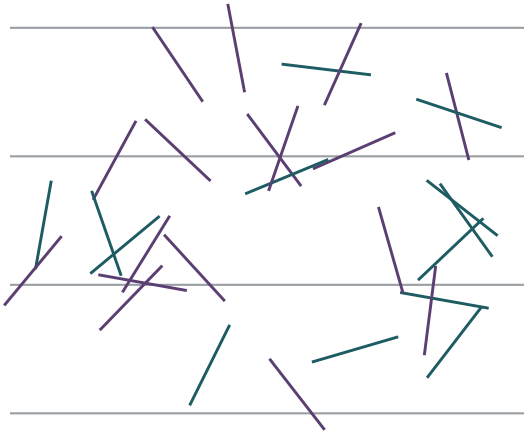


圖 12.20.1　三十根長度為 $2a = 1.4$ 的針，拋在相距 $2b = 2$ 的直線上。與直線相交的針以紫紅色表示。這裏有 17 根相交，而期望值是 13.4 根：小樣本只能給出粗略的估算，而布豐投針正是出名緩慢的 π 計算方法。

回顧

- **兩個獨立的坐標。**針的位置和角度互相獨立（independent）且均勻分佈，所以所求概率是一個二重積分（double integral），可以逐個變數計算。
- **先取條件，再求平均。**先固定 x，對角度取平均；然後再對 x 取平均。
- **代換顯露出常數。**以 $x = at$ 作縮放，答案便化約為單一個通用的積分。

背後的數學

布豐的麵條

以下的方法幾乎不用計算就能看出答案。把針屈曲成任意形狀，成為一根「麵條」，並設 C 為它與直線相交的次數。根據期望值的線性（linearity of expectation），一條曲線相交次數的期望值，等於它各小段的期望值之和，所以與長度成正比：$\mathrm{E}(C) = \kappa L$，其中 κ 是某個常數。要求出 κ，可取麵條為一個直徑等於直線間距 $2b$ 的圓：它總是恰好相交兩次，而其長度為 $2\pi b$。所以 $\kappa \cdot 2\pi b = 2$，得出 $\kappa = \frac{1}{\pi b}$。對於長度為 $2a < 2b$ 的直針，它最多相交一次，故 $\mathrm{P}(\text{相交}) = \mathrm{E}(C) = \frac{2a}{\pi b}$。這個論證出自 Joseph-Émile Barbier（1860 年），是概率推理中的瑰寶：那個 π 來自用來校準常數的圓。

估算 π，以及錯誤示範

若拋出 n 根針，其中 c 根相交，則 $\frac{c}{n} \approx \frac{2a}{\pi b}$，而 $\pi \approx \frac{2an}{bc}$。誤差只會按 $\frac{1}{\sqrt{n}}$ 的速度縮小，所以要把 π 算至小數點後三位，大約需要拋一百萬次。1901 年，意大利數學家 Mario Lazzarini 聲稱拋了 3408 次便得出 $\pi \approx \frac{355}{113} = 3.1415929\ldots$（準確至小數點後六位）。統計學家指出，這個結果好得令人難以置信：拋 3408 次，再配合一根精心挑選的針，恰好 1808 次相交便會得出 $\frac{355}{113}$，而只要在適當的時刻停止實驗，就能製造出這個結果。這是關於選擇性報告的經典警世故事。

布豐

布豐是一位博物學家，曾任巴黎皇家植物園園長，並著有一套 36 卷的自然史。他於 1733 年向法國科學院（Académie des Sciences）提出投針問題（以及一個使用方形地磚的版本，即「franc-carreau」遊戲），並於 1777 年發表解答。他的目的是要說明，概率中不單可以運用算術，也可以運用幾何。積分幾何（integral geometry）——研究隨機直線及其與曲線和曲面的交點的理論——正是由此發展而來，如今應用於體視學（stereology），即從二維切片估算三維結構（例如組織中的細胞）的大小和形狀的科學。

思維技巧

把概率看成區域上的積分

當結果取決於數個獨立的連續隨機變量時，一個事件的概率就是聯合密度（joint density）在該事件發生的區域上的積分。

在大學裏，這成為連續概率的基礎：聯合密度、邊際分佈（marginal distribution）和變數代換（change of variable），這些全都會在 Section 12.21 和 Section 12.22 再次出現。

延伸練習

1 **長針。**若 $a > b$，針可能與多於一條直線相交。證明相交次數的期望值仍然是 $\frac{2a}{\pi b}$。

2 **麵條。**把一根長度為 $2a$ 的針屈成直角（兩臂各長 a）。相交次數的期望值是多少？

3 **方形地磚。**把一枚半徑為 r 的硬幣拋到鋪有邊長為 $s > 2r$ 的方形地磚的地板上。求硬幣完全落在一塊地磚之內的概率。

4 **Lazzarini 的針。**設 $a = \frac{5}{6}b$（使得 $\frac{2a}{\pi b} = \frac{5}{3\pi}$），證明拋 3408 次而有 1808 次相交，會給出估算值 $\pi \approx \frac{355}{113}$。

延伸閱讀

- **Daniel Klain & Gian-Carlo Rota, *Introduction to Geometric Probability***（Cambridge University Press），此書由布豐投針問題講起。
- **Lee Badger, "Lazzarini's lucky approximation of π"**，*Mathematics Magazine* 67 (1994)。一個可疑實驗背後的統計學。
- **Paul Nahin, *Digital Dice***（Princeton University Press）。以蒙地卡羅方法處理概率問題。

12.21 折斷一根棒

題目

STEP III 1990，第 16 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/90-s3-q16`

引子

在一根木棒上隨機選兩點，把它折斷。三段木棒能否組成三角形？著名的答案是：概率（probability）為 $\frac{1}{4}$。如果能組成三角形，它是銳角（acute）三角形還是鈍角（obtuse）三角形？通常是鈍角三角形：所組成的三角形中，只有大約三分之一是銳角三角形。這類關於隨機圖形的問題，自十九世紀起便令概率學家樂此不疲，也是練習處理連續概率（continuous probability）的絕佳場地：區域內的一個隨機點、一個界定子區域的條件，以及一個計算其面積的積分。

本題採用兩種不同的折棒方法，並在每種情況下問出現鈍角的可能性有多大。答案 $3-2\sqrt{2}$ 和 $3-4\ln 2$ 都是出人意表的數，卻源自簡單的幾何。

動筆之前

想一想 三個長度能組成三角形，當且僅當每一個都小於其餘兩個之和。如果三個長度之和為 L，這對最長的一個有何含意？邊 c 所對的角何時是鈍角？

$c < a+b$ 連同 $a+b+c = L$，即 $c < \frac{L}{2}$；所以三段能組成三角形，正正是當每一段都短於總長度的一半。根據餘弦公式（cosine rule），c 所對的角是鈍角，正正是當 $c^2 > a^2+b^2$。

逐步解題

階段 1　先折一次，再對半分

三段的長度是 X、$\frac{1-X}{2}$、$\frac{1-X}{2}$。兩段相等的木棒必定短於 $\frac{1}{2}$，所以能組成三角形當且僅當 $X < \frac{1}{2}$。這個三角形是等腰（isosceles）三角形；兩個相等的角都是銳角，所以只有當 X 所對的角是鈍角時，它才是鈍角三角形：

$$X^2 > 2\Big(\frac{1-X}{2}\Big)^2 = \frac{(1-X)^2}{2}, \qquad \text{即} \qquad \sqrt{2}X > 1-X, \quad X > \frac{1}{1+\sqrt{2}} = \sqrt{2}-1.$$

已知 $X < \frac{1}{2}$ 時，X 在 $[0, \frac{1}{2}]$ 上均勻分佈（uniform），所以

$$\mathrm{P}(\text{鈍角} \mid \text{三角形}) = \frac{\frac{1}{2} - (\sqrt{2}-1)}{\frac{1}{2}} = 3 - 2\sqrt{2} \approx 0.172. \qquad \blacksquare$$

階段 2　第二種構造方法

> **想一想**　設 (x, y) 位於以 $(1,0)$、$(1,1)$、$(0,1)$ 為頂點的三角形內，對邊長 x、y 和 $2-x-y$ 檢查三條三角不等式（triangle inequality）。

該區域是 $\{x \le 1, y \le 1, x+y \ge 1\}$。三條不等式分別是：$x+y \ge 2-x-y$，即 $x+y \ge 1$；$x+(2-x-y) \ge y$，即 $y \le 1$；以及 $y+(2-x-y) \ge x$，即 $x \le 1$。它們在區域內全部成立（在區域的邊上等號成立，此時三角形是退化（degenerate）的）。 $\blacksquare$

> **想一想**　B 處的角所對的邊是 $CA = y$。把條件 $y^2 > x^2 + (2-x-y)^2$ 寫成關於 y 的不等式。

展開得 $y^2 > x^2 + 4 + x^2 + y^2 - 4x - 4y + 2xy$，所以 $4y - 2xy > 2x^2 - 4x + 4$，即 $2y(2-x) > 2(x^2 - 2x + 2)$。由於 $2 - x > 0$，

$$\angle ABC \text{ 為鈍角} \iff y > \frac{x^2 - 2x + 2}{2 - x}. \qquad \blacksquare$$

階段 3　面積

> **想一想**　以 $2-x$ 除 $x^2 - 2x + 2$。然後求區域中位於曲線上方那部分的面積，再除以區域的面積 $\frac{1}{2}$。

由於 $x^2 - 2x + 2 = -x(2-x) + 2$，曲線是 $y = -x + \frac{2}{2-x}$。它由 $(0,1)$ 伸延至 $(1,1)$，而當 $0 < x < 1$ 時，它位於直線 $y = 1 - x$ 之上、$y = 1$ 之下（圖 12.21.1）。鈍角區域位於

曲線與 $y=1$ 之間，面積為

$$\int_0^1 \left(1+x-\frac{2}{2-x}\right)\mathrm{d}x = 1+\frac{1}{2}+\left[2\ln(2-x)\right]_0^1 = \frac{3}{2}-2\ln 2.$$

除以 $\frac{1}{2}$：$\mathrm{P}(\angle ABC \text{ 為鈍角}) = 3-4\ln 2 \approx 0.227$。 ■

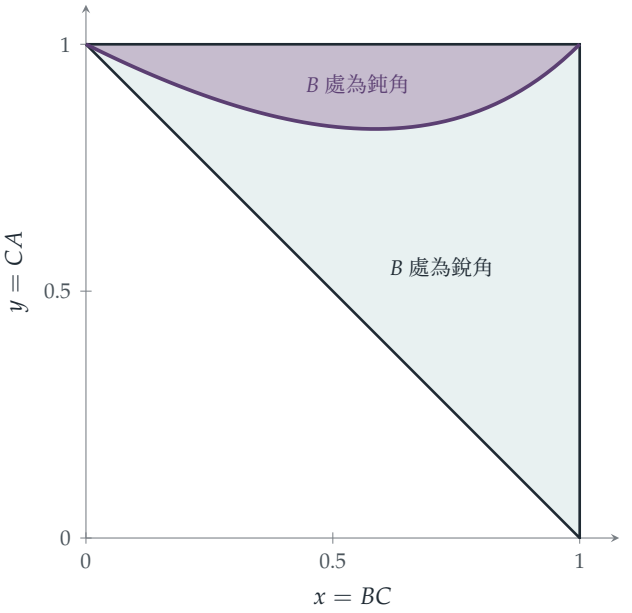


圖 12.21.1 第 (ii) 部分中 (x,y) 所有可能取值構成的三角形。在曲線 $y=\frac{x^2-2x+2}{2-x}$ 上方（紫色區域），B 處的角是鈍角；這個區域佔三角形的比例為 $3-4\ln 2 \approx 0.23$。

回顧

- **把幾何轉化為不等式。**「能組成三角形」和「是鈍角三角形」都化成了各長度之間的不等式。
- **連續變量的條件概率（conditional probability）。**已知 $X<\frac{1}{2}$ 時，X 在較小的區間上均勻分佈。
- **概率即面積之比。**對於區域內均勻分佈的點，概率等於面積除以總面積。

背後的數學

經典的折棒問題

在兩個獨立而均勻分佈的點 U 和 V 處把木棒折斷。數對 (U, V) 在單位正方形上均勻分佈，而三段能組成三角形，正正是當每一段都短於 $\frac{1}{2}$：這個條件切去四個小三角形，總面積為 $\frac{3}{4}$，餘下的概率是 $\frac{1}{4}$。已知能組成三角形時，它是銳角三角形的概率是 $\frac{1}{4}$ 當中的 $3\ln 2 - 2 \approx 0.079$，即大約 32%。這三段木棒是均勻隨機分割（uniform random partition）的一個例子，它們的聯合分佈（joint distribution）是狄利克雷分佈（Dirichlet distribution），這種分佈出現在貝葉斯統計學（Bayesian statistics）中，也出現在描述資源如何隨機分配的模型中。

怎樣折斷木棒，大有關係

請留意，題目中的兩種方法，對聽起來相似的問題給出了不同的答案 $3 - 2\sqrt{2}$ 和 $3 - 4\ln 2$，而兩者又都與經典方法的答案不同。正如貝特朗（Bertrand）的弦（Section 12.3）一樣，在指明具體程序之前，「隨機」一詞毫無意義。

隨機三角形

關於隨機三角形的問題，是維多利亞時代流行的趣題。路易斯·卡羅（Lewis Carroll）的 *Pillow Problems*（1893 年）問及「隨機」選取的三角形是鈍角三角形的概率，而他的答案，正如本節的答案一樣，完全取決於他所選擇的隨機性模型。關於隨機三角形的問題一直帶來驚喜：例如，在正方形內均勻隨機選取三點，它們組成鈍角三角形的概率約為 0.725，這個數值到二十世紀才有人準確算出。

思維技巧

樣本空間（sample space）的幾何

把隨機結果表示爲區域內的一點，把事件表示爲子區域，並以面積之比（或密度（*density*）的積分）計算概率。

在大學裏，這幅圖像是多元概率（multivariate probability）的基礎，也是變數代換公式（change-of-variables formula）的基礎；這條公式說明坐標改變時，密度如何隨之變換。

延伸練習

1 **經典問題。**證明：若在兩個獨立而均勻分佈的點處把木棒折斷，則各段能組成三角形的概率是 $\frac{1}{4}$。

2 **折斷較長的一段。**在一個均勻分佈的點處把木棒折斷，然後在較長的一段上取一個均勻分佈的點，再把它折斷。證明三段能組成三角形的概率是 $2\ln 2 - 1 \approx 0.386$。

3 **銳角還是鈍角。**在第 (ii) 部分中，求三角形 ABC 在 C 處（對邊為 $AB = 2 - x - y$）有鈍角的概率。由此求它有任何一個角是鈍角的概率。

4 **最長一段的期望長度。**對於經典的折棒問題，證明最長一段的期望（expected）長度是 $\frac{11}{18}$。

延伸閱讀

- **Martin Gardner, *The Colossal Book of Mathematics***（Norton），書中收錄了折棒問題。
- **Frederick Mosteller, *Fifty Challenging Problems in Probability***（Dover）。
- **Lewis Carroll, *Pillow Problems***（Dover，與 *A Tangled Tale* 合訂）。維多利亞時代的趣題，包括隨機三角形。

12.22 沒有平均值的分佈

題目

STEP III 1995，第 12 題

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

`step.maths.org/questions/95-s3-q12`

引子

把許多次量度的結果取平均，隨機誤差便會互相抵消：平均值會穩定下來，趨向真實的數值。這條*大數定律*（*law of large numbers*）是整個統計學的基石。但它暗藏一個假設：這些量度必須有有限的平均值(mean)。本題構造了一個再自然不過的隨機變量(random variable)，即兩個獨立（independent）標準正態（standard normal）變量之比，而它根本沒有平均值。它的平均永遠不會穩定下來；事實上，一百萬個這樣的變量的平均，其分佈與單獨一個的分佈完全相同。

這就是*柯西分佈*（*Cauchy distribution*）。它的密度（density）看來像一個人畜無害的鐘形，但它的尾部實在太厚，偶然出現的極大數值足以左右任何平均。本題用一個漂亮的幾何論證把它找出來：兩個坐標之比就是一條直線的斜率，而徑向對稱令這條直線的角度呈均勻分佈。

動筆之前

想一想 若 (U, V) 是一個隨機點，其密度只取決於它與原點的距離，那麼由原點到該點的直線與坐標軸所成的角 Θ，你能說些甚麼？

把平面旋轉並不會改變密度，所以每個方向的可能性都相同：角 Θ 在 $[0, 2\pi)$ 上呈均勻（uniform）分佈。而 $\frac{U}{V}$ 只取決於方向，與距離無關。

逐步解題

階段 1　聯合密度（joint density）

由於 X 與 Y 獨立，它們的聯合密度就是各自密度的乘積：

$$\mathrm{f}(x,y) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2} \cdot \frac{1}{\sqrt{2\pi}} e^{-y^2/2} = \frac{1}{2\pi} e^{-\frac{1}{2}(x^2+y^2)}.$$

它只透過 x^2+y^2（即與原點距離的平方）取決於 x 和 y：在以原點為圓心的圓上，密度是常數。這就是「徑向對稱」（radially symmetrical）的意思。

階段 2　運用對稱性

> **想一想**　$\frac{U}{V} < k$ 可以在兩種情況下發生，視乎 V 的正負號而定。這兩種情況與旋轉 180° 的變換 $(u,v) \mapsto (-u,-v)$ 有甚麼關係？

$\frac{U}{V} < k$ 的意思是：要麼 $V > 0$ 且 $U < kV$，要麼 $V < 0$ 且 $U > kV$。旋轉 $(u,v) \mapsto (-u,-v)$ 把第一個區域映射到第二個區域上，並且保持密度不變。所以兩者的概率相同，於是

$$\mathrm{P}\Big(\frac{U}{V} < k\Big) = 2\mathrm{P}(U < kV, V > 0). \qquad \blacksquare$$

階段 3　一個角

> **想一想**　在以 u 為橫軸、v 為縱軸的平面上，畫出區域 $\{V > 0, U < kV\}$。它是一個頂點在原點的楔形。它的角度是多少？

區域 $V > 0$ 是上半平面。通過原點的直線 $u = kv$ 與正 v 軸所成的角為 $\arctan k$（當 $k > 0$ 時偏向正 u 的一方）。上半平面中位於這條直線左方的部分，是一個角度為 $\frac{\pi}{2} + \arctan k$ 的楔形。根據徑向對稱，楔形的概率等於它的角度除以 2π，所以

$$\mathrm{P}\Big(\frac{U}{V} < k\Big) = 2 \cdot \frac{\frac{\pi}{2} + \arctan k}{2\pi} = \frac{1}{2} + \frac{1}{\pi} \arctan k.$$

對 k 求導，便得到密度

$$\mathrm{g}(k) = \frac{1}{\pi(1+k^2)}. \qquad \blacksquare$$

回顧

- **獨立的正態變量具有徑向對稱性。**兩個正態密度的乘積只取決於距離。
- **比值即方向。**$\frac{U}{V}$ 只取決於該點的角度，而這個角度呈均勻分佈。

► **楔形的概率。**在徑向對稱之下，楔形的概率與它的角度成正比。

背後的數學

沒有平均值

柯西分佈的平均值本應是 $\int_{-\infty}^{\infty} \frac{k}{\pi(1+k^2)}\,\mathrm{d}k$，但 $\int_0^K \frac{k}{1+k^2}\,\mathrm{d}k = \frac{1}{2}\ln(1+K^2) \to \infty$：這個積分並不收斂（converge），所以平均值不存在。（對稱性或會令人以為答案是 0，但正負兩部分都是無窮大，而 $\infty - \infty$ 並沒有值。）方差（variance）同樣是無窮大。

更糟的是，n 個獨立柯西變量的平均仍然是柯西變量，而且分佈完全相同。（其中一個看出這點的方法，是利用特徵函數（characteristic function）$\mathrm{E}(e^{itX}) = e^{-|t|}$。）所以取平均完全無補於事：一百萬次量度的平均，並不比單獨一次可靠。圖 12.22.1 顯示了實際情況：正態樣本的累計平均會穩定下來，而柯西樣本的累計平均卻永遠東歪西倒，不時被偶然出現的極大數值撞得偏離原位。

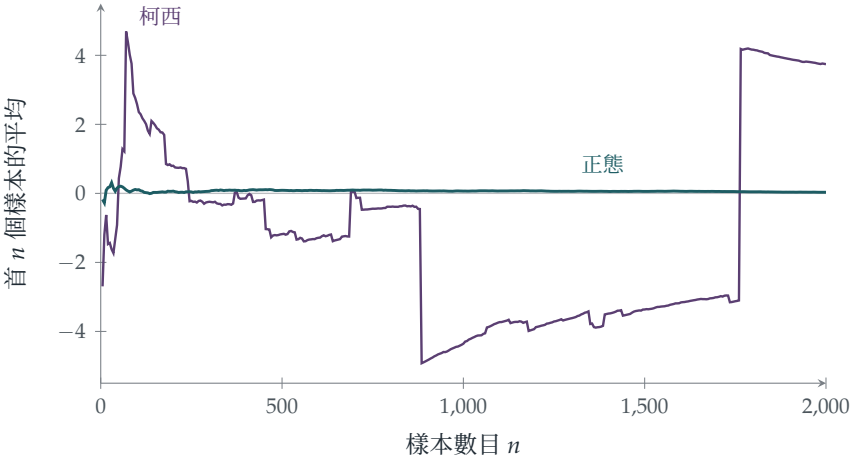


圖 12.22.1 2000 個獨立樣本的累計平均。對於標準正態樣本（藍綠色），平均值一如大數定律所預測，穩定下來並趨向 0。對於柯西樣本（紫紅色），它永遠不會穩定：每隔一段時間，單單一個極大的數值就會把它拋到另一個位置。

柯西變量從何而來

拿一支電筒，朝一個均勻隨機的方向照向一面牆：光點在牆上的位置便服從柯西分佈。在物理學中，同一個密度稱為勞侖茲分佈（*Lorentzian*），用來描述譜線和共振峯的形狀（比較 Section 11.1）。在金融學中，像柯西分佈這類厚尾（heavy-tailed）分佈，會用來為偶爾出現的股災建立模型，而正態模型往往嚴重低估這類事件。這對統計學是一個嚴肅的教訓：建基於平均值和方差的方法，遇上厚尾數據便會失效，這時就需要穩健的方

法，例如使用中位數（median）（柯西分佈的中位數完全沒有問題）。

泊松、柯西與一場論戰

這個分佈在 1824 年由泊松（Siméon Denis Poisson）研究過，他注意到它違反大數定律。柯西（Augustin-Louis Cauchy）則在 1853 年與 Irénée-Jules Bienaymé 就最小二乘法（method of least squares）公開論戰時遇上它：柯西用它來說明，當誤差屬於厚尾時，最小二乘法可以是一個很差的方法。這個分佈最終冠上了柯西的名字，正是 Stephen Stigler 所稱的「史蒂格勒命名定律」（Stigler's law of eponymy）的一例：沒有任何科學發現是以其最初發現者命名的。

思維技巧

檢查期望值是否存在

使用平均值或方差之前，先確定定義它的積分是收斂的。對稱性並不能令一個發散的積分變成零。

在大學裏，可積性條件是概率論中每一條定理的一部分：大數定律需要有限的平均值，中心極限定理（central limit theorem）（Section 12.24）需要有限的方差。當這些條件不成立時，柯西分佈就是標準的反例。

延伸練習

1 **電筒。**一支電筒與一面長而直的牆相距 1，所指的方向與垂直線成角 Θ，其中 Θ 在 $(-\frac{\pi}{2}, \frac{\pi}{2})$ 上呈均勻分佈。證明光點的位置 $\tan\Theta$ 服從柯西分佈。

2 **倒數。**證明若 C 服從柯西分佈，則 $\frac{1}{C}$ 亦然。（提示：$\frac{U}{V}$ 與 $\frac{V}{U}$。）

3 **中位數行得通。**證明柯西分佈的中位數是 0，而它的四分位數（quartiles）是 ± 1。

4 **半徑與角度。**設 X、Y 為獨立的標準正態變量，透過轉換成極坐標（polar coordinates），證明 $R^2 = X^2 + Y^2$ 服從平均值為 2 的指數分佈（exponential distribution），並且與角度獨立。（這是用來產生正態隨機數的 Box–Muller 方法的基礎。）

延伸閱讀

► **Geoffrey Grimmett & Dominic Welsh, *Probability: An Introduction*** (Oxford University

Press），論隨機變量的變換。

- **Nassim Nicholas Taleb, *The Black Swan***（Penguin）。講述厚尾現象及其重要性，適合一般讀者。
- **Stephen Stigler, *The History of Statistics***（Harvard University Press）。書中記述了柯西與 Bienaymé 的論戰。

12.23 由概率得出斯特靈公式

題目

STEP II 2016，第 13 題

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

`step.maths.org/questions/16-s2-q13`

引子

100! 有多大？要逐項乘出來得花點工夫：它有 158 位數字。斯特靈公式（*Stirling's formula*）$n! \approx \sqrt{2\pi n}\,(n/e)^n$ 只用一行，就能給出誤差在 0.1% 以內的答案。它是數學中最有用的近似式之一，在概率（probability）、統計物理和演算法分析中經常用到。

它有很多種證明（proof），可以用積分（integral）（Section 4.4），也可以用歐拉–麥克勞林公式（Euler–Maclaurin formula）。本題給出的是最出人意表的一種：用概率來證明。平均值（mean）為 n 的泊松分佈（Poisson distribution）近似於正態分佈（normal distribution），在峯頂比較兩者，$n!$ 的值就被確定下來。斯特靈公式中的 $\sqrt{2\pi}$，正正就是正態密度（normal density）中的 $\sqrt{2\pi}$。

動筆之前

? **想一想** 設離散隨機變量（discrete random variable）X 的平均值為 μ、方差（variance）為 σ^2，並且近似於正態分佈，那麼 $\mathrm{P}(X = \mu)$ 大約是多少？（試想一個寬度為 1、以 μ 為中心的區間的概率。）

按正態近似（normal approximation），$\mathrm{P}(X = \mu) \approx \mathrm{P}(\mu - \frac{1}{2} < X < \mu + \frac{1}{2})$，這大約等於寬度 1 乘以峯頂處的密度：$\frac{1}{\sigma\sqrt{2\pi}}$。

逐步解題

階段 1　十六枚硬幣

$X \sim \mathrm{B}(16, \tfrac{1}{2})$ 的平均值為 8，方差為 $16 \cdot \tfrac{1}{2} \cdot \tfrac{1}{2} = 4$，所以 $\sigma = 2$。由正態近似，

$$\mathrm{P}(X = 8) \approx \frac{1}{2\sqrt{2\pi}} \approx 0.1995.$$

準確值是 $\binom{16}{8}/2^{16} = \tfrac{12870}{65536} \approx 0.1964$，兩者相差不到 2%。

階段 2　中央二項式係數

> **想一想**　$\mathrm{B}(2n, \tfrac{1}{2})$ 的平均值和標準差（standard deviation）是多少？分別寫出 $\mathrm{P}(X = n)$ 的準確值和近似值。

平均值是 n，方差是 $\tfrac{n}{2}$，所以 $\sigma = \sqrt{n/2}$。準確值是 $\mathrm{P}(X = n) = \binom{2n}{n}\tfrac{1}{2^{2n}}$；近似值是 $\tfrac{1}{\sqrt{n/2}\sqrt{2\pi}} = \tfrac{1}{\sqrt{\pi n}}$。令兩者相等，

$$\frac{(2n)!}{(n!)^2\, 2^{2n}} \approx \frac{1}{\sqrt{\pi n}}, \qquad (2n)! \approx \frac{2^{2n}(n!)^2}{\sqrt{n\pi}}. \qquad \blacksquare$$

階段 3　斯特靈公式

> **想一想**　參數為 n 的泊松分佈，其平均值和方差都等於 n。分別寫出 $\mathrm{P}(X = n)$ 的準確值和近似值。

準確值是 $\mathrm{P}(X = n) = e^{-n}\tfrac{n^n}{n!}$。當 n 很大時，泊松分佈近似於正態分佈（它是 n 個獨立的 Poisson(1) 變量之和），其中 $\sigma = \sqrt{n}$，所以 $\mathrm{P}(X = n) \approx \tfrac{1}{\sqrt{2\pi n}}$。令兩者相等，

$$n! \approx \sqrt{2\pi n}\, e^{-n} n^n. \qquad \blacksquare$$

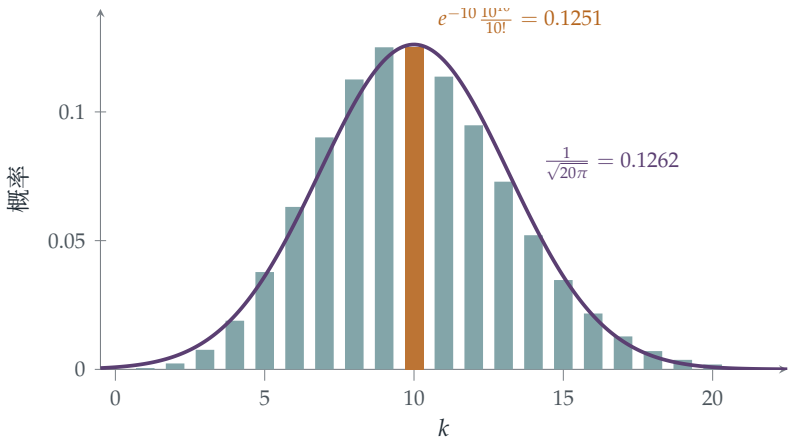


圖 12.23.1 平均值為 10 的泊松分佈（長條），以及平均值和方差與之相同的正態密度（紫色曲線）。在峯頂，泊松概率 $e^{-10}10^{10}/10!$（琥珀色）與正態密度 $\frac{1}{\sqrt{2\pi\cdot 10}}$ 十分接近。把這個近似等式移項整理，就是斯特靈公式。

回顧

- **一點上的概率 ≈ 密度乘以寬度。**對於格點分佈（lattice distribution），某一點上的正態密度近似於該點的概率。
- **準確公式遇上近似式。**二項分佈（binomial distribution）和泊松分佈的概率是準確的；令它們等於各自的正態近似，便得到階乘（factorial）的估計。
- **按公式的形狀選擇分佈。**$\binom{2n}{n}$ 令人想到 $\mathrm{B}(2n, \frac{1}{2})$；$\frac{n^n}{n!}$ 令人想到 $\mathrm{Poisson}(n)$。

背後的數學

這個近似有多準確？

這個論證只是啟發式（heuristic）的：它在單獨一點上使用正態近似，而一般的中心極限定理（central limit theorem）並不足以支持這樣做。它有一個嚴格的版本，稱為*局部中心極限定理*（*local central limit theorem*），所做的正是這件事，並使這個論證成為真正的證明。斯特靈公式其實出奇地準確：比值 $\frac{n!}{\sqrt{2\pi n}(n/e)^n}$ 大約是 $1 + \frac{1}{12n}$，所以即使 $n = 1$，它給出的也是 0.92 而不是 1；而當 $n = 10$ 時，它只偏小了 0.8%。完整的級數（series）$1 + \frac{1}{12n} + \frac{1}{288n^2} - \cdots$ 是*漸近級數*（*asymptotic series*）的一個例子：對於大的 n，它給出的近似值越來越好；但對於固定的 n，它最終會發散（diverge）。

階乘與正態分佈為何相連

斯特靈公式和正態密度中的 $\sqrt{2\pi}$ 是同一個常數，因為 $\int_{-\infty}^{\infty} e^{-x^2/2}\,\mathrm{d}x = \sqrt{2\pi}$。棣莫弗（Abraham de Moivre）在 1733 年正是循這條路找到二項分佈的正態近似的：他需要在 n 很大時近似 $\binom{n}{k}$；他找到了 $n! \approx Cn^{n+1/2}e^{-n}$ 這個形式，卻沒有找到那個常數。斯特靈（James Stirling）證明了 $C = \sqrt{2\pi}$。正態分佈是從階乘的漸近性質中誕生的。

棣莫弗與斯特靈

棣莫弗是一位逃亡到倫敦的法國新教徒，靠當私人教師和為賭徒出謀獻策維生。他在 1718 年出版了 *The Doctrine of Chances*（《機會論》），並在 1733 年的一本小冊子中找到二項分佈的正態近似，這是正態曲線首次面世。蘇格蘭數學家斯特靈在他的 *Methodus Differentialis*（1730）中發表了該常數的值，棣莫弗亦大方地把功勞歸於他。階乘的近似至今仍是活躍的課題：組合數學（combinatorics）和密碼學（cryptography）都需要對極大的 n 計算 $n!$。

思維技巧

用概率去證明與概率無關的事情

把一個準確的概率與一個近似值作比較，可以得知公式中那些數的資訊。

這種「概率方法（probabilistic method）」是雙向的。這裏概率給出了階乘的估計；艾狄胥（Erdős）則用它來證明某些圖（graph）的存在（Section 9.2）。到了大學，你會見到概率被用來證明分析、數論和組合數學中的定理。

延伸練習

1. **檢驗公式。**用斯特靈公式估計 10! 和 20!，並與真確值 3 628 800 和 2 432 902 008 176 640 000 比較。

2. **由 (iii) 到 (ii)。**證明第 (iii) 部分的斯特靈公式可推出第 (ii) 部分的近似式。

3. **100! 的位數。**利用 $\log_{10} n! \approx n\log_{10}\frac{n}{e} + \frac{1}{2}\log_{10}(2\pi n)$ 證明 100! 有 158 位數字。

4. **擲硬幣打成平手。**利用第 (ii) 部分，證明擲一枚公平硬幣 1000 次而恰好得到 500 次正面的概率約為 0.025。

延伸閱讀

- **William Feller, *An Introduction to Probability Theory and Its Applications*, Vol. 1**（Wiley），第 II 章，當中證明了斯特靈公式。
- **Ronald Graham, Donald Knuth & Oren Patashnik, *Concrete Mathematics***（Addison-Wesley），第 9 章，講述漸近分析。
- **Stephen Stigler, *The History of Statistics***（Harvard University Press），講述棣莫弗與正態曲線的誕生。

12.24 中心極限定理

題目

STEP III 2008，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/08-s3-q12`

引子

把許多微小而獨立的隨機效應加起來，不論每一個效應本身服從甚麼分佈，總和都會服從鐘形的正態曲線。這就是*中心極限定理*（central limit theorem），它也許是概率論中最重要的定理。它解釋了為何身高、量度誤差和考試分數都大致服從正態分佈（normal distribution），也解釋了為何統計學家即使對背後的分佈一無所知，仍能說出一個平均值有多可靠，只要該分佈的方差（variance）有限便可（試與 Section 12.22 比較）。

本題在一個特殊情況下證明這個定理。一隻青蛙隨機跳躍，每次跳躍服從雙邊指數分佈（two-sided exponential distribution），其密度函數（density）有一個尖峯，完全不像鐘形。跳了許多次之後，牠的總位移卻服從正態分佈。所用的工具是*矩母函數*（moment generating function），它把獨立變量之和化為乘積，再取對數，又把乘積化為和。

關於課程大綱的說明

矩母函數已於 2019 年從 STEP 考試大綱中刪除。隨機變量（random variable）X 的矩母函數是 $\mathrm{M}_X(\theta) = \mathrm{E}(e^{\theta X})$。它在 0 處的各階導數給出各階矩（moments）：$\mathrm{M}'(0) = \mathrm{E}(X)$，$\mathrm{M}''(0) = \mathrm{E}(X^2)$。若 X 與 Y 獨立（independent），則 $\mathrm{M}_{X+Y} = \mathrm{M}_X\mathrm{M}_Y$，而且 $\mathrm{M}_{cX}(\theta) = \mathrm{M}_X(c\theta)$。若矩母函數收斂，則（在適當的意義下）分佈也收斂。

動筆之前

想一想　密度函數 $\frac{1}{2}e^{-|x|}$ 是對稱的。它的圖像是甚麼樣子？它的平均值（mean）是多少？

它是一條帳篷形的曲線，在 0 處有一個高度為 $\frac{1}{2}$ 的尖峯，並向兩邊按指數衰減（圖 12.24.1）。由對稱性可知平均值為 0。

逐步解題

階段 1　矩母函數

當 $|\theta| < 1$ 時，把積分在 0 處分開，

$$\mathrm{M}_X(\theta) = \frac{1}{2}\int_{-\infty}^{0} e^{(1+\theta)x}\,\mathrm{d}x + \frac{1}{2}\int_{0}^{\infty} e^{-(1-\theta)x}\,\mathrm{d}x = \frac{1}{2}\Big(\frac{1}{1+\theta} + \frac{1}{1-\theta}\Big) = \frac{1}{1-\theta^2}. \qquad \blacksquare$$

展開得 $\mathrm{M}_X(\theta) = 1 + \theta^2 + \theta^4 + \cdots$，所以 $\mathrm{E}(X) = 0$，$\mathrm{E}(X^2) = 2$（即 $\frac{\theta^2}{2!}$ 的係數）。方差為 2。

階段 2　青蛙的位移

想一想　$Y = X_1 + \cdots + X_n$ 的矩母函數是甚麼？$Y/\sqrt{2n}$ 的又是甚麼？

由獨立性得 $\mathrm{M}_Y(\theta) = (1-\theta^2)^{-n}$，再作縮放，

$$\mathrm{M}_{Y/\sqrt{2n}}(\theta) = \mathrm{M}_Y\Big(\frac{\theta}{\sqrt{2n}}\Big) = \Big(1 - \frac{\theta^2}{2n}\Big)^{-n}.$$

（除以 Y 的標準差（standard deviation）$\sqrt{2n}$ 作縮放，令方差變為 1。）

階段 3　極限

想一想　取對數，並利用當 u 很小時 $\ln(1-u) \approx -u$。

$\ln \mathrm{M}_{Y/\sqrt{2n}}(\theta) = -n\ln\big(1 - \frac{\theta^2}{2n}\big) = -n\big(-\frac{\theta^2}{2n} - \frac{\theta^4}{8n^2} - \cdots\big) \to \frac{\theta^2}{2}$，所以

$$\mathrm{M}_{Y/\sqrt{2n}}(\theta) \to e^{\theta^2/2}, \qquad \blacksquare$$

這正是標準正態分佈的矩母函數。

階段 4　青蛙的機會

想一想 當 n 很大時，Y 近似服從 $\mathrm{N}(0, 2n)$。何時有 $\mathrm{P}(|Y| \geq 25) = 0.05$？

對於標準正態變量 Z，$\mathrm{P}(|Z| \geq 1.96) = 0.05$。因此我們需要 $\frac{25}{\sqrt{2n}} = 1.96$，由此得 $2n = \left(\frac{25}{1.96}\right)^2 \approx 162.7$，即 $n \approx 81.3$。由於落在遠處的機會隨 n 增大而增加，要有至少 5% 的機會，最少的跳躍次數約為 82。 ■

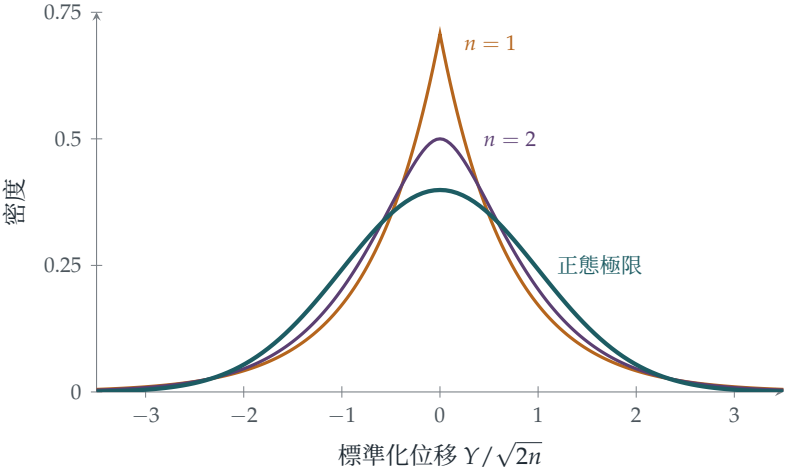


圖 12.24.1 青蛙跳了 $n = 1$ 次（琥珀色：一個拉普拉斯（Laplace）尖峯）和 $n = 2$ 次（紫紅色）之後，其標準化位移的密度函數，與標準正態密度函數（藍綠色）的比較。只跳了兩次，尖峯已經變得圓滑；中心極限定理指出，當 $n \to \infty$ 時，這些曲線收斂至正態曲線。

回顧

- **變換把和化為乘積。**獨立變量之和的矩母函數，等於各變量矩母函數的乘積。
- **對數把乘積化為和。**$\ln\left(1 - \frac{\theta^2}{2n}\right)^{-n}$ 變成了一個和，其極限很容易求得。
- **標準化（standardise）。**除以標準差 $\sqrt{2n}$ 令方差保持為 1，所以極限既不是 0，也不是無窮大。

背後的數學

一般的定理

同樣的論證適用於任何平均值為 0、方差 σ^2 有限的分佈（還須有矩母函數，不過這個條件可以避開）。這時 $\mathrm{M}(\theta) = 1 + \frac{1}{2}\sigma^2\theta^2 + o(\theta^2)$，而且

$$\left(\mathrm{M}\left(\frac{\theta}{\sigma\sqrt{n}}\right)\right)^n = \left(1 + \frac{\theta^2}{2n} + o\left(\frac{1}{n}\right)\right)^n \to e^{\theta^2/2}.$$

取極限之後只有方差留下來：原來分佈的形狀已被遺忘。這就是正態分佈具有普遍性的原因。在大學裏，這個定理是用*特徵函數*（characteristic function）$\mathrm{E}(e^{i\theta X})$ 來證明的，特徵函數必定存在，而分佈的收斂則由萊維連續性定理（Lévy's continuity theorem）賦予嚴格的意義。

有多快？

中心極限定理說明極限情況下會發生甚麼，卻沒有說明收斂得有多快。*Berry–Esseen* 定理（Berry–Esseen theorem）給出誤差的上界：一個常數乘以 $\frac{\mathrm{E}|X|^3}{\sigma^3\sqrt{n}}$。對於像青蛙跳躍那樣的對稱分佈，跳了幾十次之後近似已相當準確。對於偏態（skewed）分佈，或者在極端的尾部，即使 n 很大，近似也可能很差：正態近似會低估極罕見事件發生的機會，這個事實對保險和金融都有影響。

從棣美弗到 Lindeberg

棣美弗（De Moivre）在 1733 年發現了二項分佈（binomial distribution）的正態近似（Section 12.23），拉普拉斯在 1810 年把它推廣至更一般的變量之和，並首先提出量度誤差是許多微小效應之和這個想法。嚴格的證明要到很久之後才出現：李亞普諾夫（Aleksandr Lyapunov）在 1901 年給出證明，Jarl Lindeberg 則在 1922 年給出證明，並找到了基本上是令定理成立的最弱條件。「中心極限定理」這個名稱是波利亞（George Pólya）在 1920 年起的，意思是這個定理在概率論中處於中心地位。

思維技巧

變換、化簡、再變換回來

一個困難的運算（把獨立隨機變量相加，即對它們的密度函數作卷積（*convolution*）），經過變換之後變成一個容易的運算（乘法）。在那裏完成工作，然後詮釋所得的結果。

同樣的策略也是生成函數（generating functions）（Section 9.4）、分析學和工程學中的傅立葉（Fourier）變換和拉普拉斯變換，以及把乘法化為加法的對數背後

的原理。

延伸練習

1 **兩次跳躍的精確結果。**證明 $Y = X_1 + X_2$ 的密度函數為 $\frac{1}{4}(1 + |y|)e^{-|y|}$，並驗證其方差為 4。

2 **從二項到正態。**設 $X \sim \mathrm{B}(1, \frac{1}{2})$，並把它平移使平均值為 0（即取值 $\pm\frac{1}{2}$），求出 $\mathrm{M}_X(\theta) = \cosh\frac{\theta}{2}$，並證明 n 個這樣的變量的標準化和的矩母函數趨向 $e^{\theta^2/2}$。

3 **青蛙的精確結果。**利用電腦（或以 $n = 2$ 時的精確密度函數作檢驗），看看當 $n = 82$ 時正態近似有多準確。

4 **為何柯西不行？**柯西分佈（Cauchy distribution）（Section 12.22）沒有矩母函數。論證中哪一步失效？n 個柯西變量的平均值又會怎樣？

延伸閱讀

- **Geoffrey Grimmett & Dominic Welsh, *Probability: An Introduction***（Oxford University Press），關於生成函數和極限定理的一章。
- **Hans Fischer, *A History of the Central Limit Theorem***（Springer）。從拉普拉斯到 Lindeberg。
- **David Spiegelhalter, *The Art of Statistics***（Pelican），講述為何正態分佈無處不在。

12.25 眾多之中的最大值

題目

STEP III 2025，第 11 題

繼續閱讀之前，請先看題目（原題為英文）。題目載於 OCR 提供下載的 2025 年 STEP 試卷（zip 檔案）內的 STEP III 試卷：

`www.ocr.org.uk/Images/750486-2025-past-papers.zip`

引子

下一個世紀最高的洪水會有多高？一千個燈泡之中，壽命最長的那個能用多久？在 n 名參賽者之中，跑得最快的馬拉松選手會有多快？這類關於極值（extremes）的問題，即許多隨機數值之中最大或最小的一個，在工程、保險和氣候科學中都極為重要，而極值的表現與平均值截然不同。

本題要找出 n 個指數分佈（exponential distribution）壽命中最大者的分佈。它的典型大小按 $\ln n$ 增長，雖然緩慢，卻沒有上限。然而它的離散程度卻完全不會增長：一千個壽命中的最大值，其可預測的程度與十個壽命中的最大值大致相同。在極限情況下，它的分佈形狀就是耿貝爾分佈（Gumbel distribution），也就是極值世界中與正態曲線相對應的角色。

動筆之前

想一想 $X_1, \ldots, X_n$ 中的最大者不超過 y，恰好等同於發生了甚麼事？

$Y \le y$ 恰好在每個 $X_i \le y$ 時成立。利用獨立性（independence），這個概率便化為一個乘積。

逐步解題

階段 1　最大值的分佈

每個 X_i 在 $y \ge 0$ 時都有 $\mathrm{F}(y) = 1 - e^{-\lambda y}$，由獨立性可得

$$\mathrm{G}(y) = \mathrm{P}(X_1 \le y, \dots, X_n \le y) = (1 - e^{-\lambda y})^n. \qquad \blacksquare$$

階段 2　分位數

> **想一想**　由 $\mathrm{G}(L) = \alpha$ 解出 L。然後解 $\mathrm{G}(U) = 1 - \alpha$。

由 $\left(1 - e^{-\lambda L}\right)^n = \alpha$ 得 $e^{-\lambda L} = 1 - \alpha^{1/n}$，所以

$$L(\alpha) = -\frac{1}{\lambda}\ln\!\left(1 - \alpha^{1/n}\right), \qquad U(\alpha) = -\frac{1}{\lambda}\ln\!\left(1 - (1-\alpha)^{1/n}\right). \qquad \blacksquare$$

階段 3　n 很大的情況

> **想一想**　把 $\alpha^{1/n} = e^{(\ln \alpha)/n}$ 寫出來。當 n 很大時，指數很小。那麼 $1 - \alpha^{1/n}$ 近似於甚麼？

$\alpha^{1/n} = e^{(\ln\alpha)/n} \approx 1 + \frac{\ln \alpha}{n}$，所以 $1 - \alpha^{1/n} \approx \frac{-\ln\alpha}{n} = \frac{\ln(1/\alpha)}{n}$。因此

$$\lambda L(\alpha) \approx -\ln\frac{\ln(1/\alpha)}{n} = \ln n - \ln\!\left(\ln\frac{1}{\alpha}\right). \qquad \blacksquare$$

同理，$\lambda U(\alpha) \approx \ln n - \ln\!\left(\ln \frac{1}{1-\alpha}\right)$。

階段 4　增長與離散程度

中位數（median）是 $L(\frac{1}{2})$，而 $\lambda L(\frac{1}{2}) \approx \ln n - \ln\ln 2 \to \infty$。但是

$$\lambda\big(U(\alpha) - L(\alpha)\big) \approx \ln\!\left(\ln\frac{1}{\alpha}\right) - \ln\!\left(\ln\frac{1}{1-\alpha}\right),$$

當中的 $\ln n$ 項已互相抵消：區間的闊度與 n 無關。 $\blacksquare$

階段 5　一個 90% 區間

> **想一想**　取 $\alpha = 0.05$。$\ln\frac{1}{\alpha} = \ln 20$ 和 $\ln\frac{1}{0.95}$ 大約是多少？

當 $\alpha = 0.05$ 時，$\ln 20 \approx 3$，所以 $\ln\ln\frac{1}{\alpha} \approx \ln 3$。另外 $\ln\frac{1}{0.95} = -\ln(1-0.05) \approx 0.05 = \frac{1}{20}$，所以 $\ln\ln\frac{1}{1-\alpha} \approx \ln\frac{1}{20} \approx -3$。因此闊度約為 $\frac{1}{\lambda}(\ln 3 + 3) \approx \frac{4.1}{\lambda}$，而 Y 以概率 0.9 落在 $[L(0.05), U(0.05)]$ 之內，這個區間的闊度約為 $4\lambda^{-1}$。 $\blacksquare$

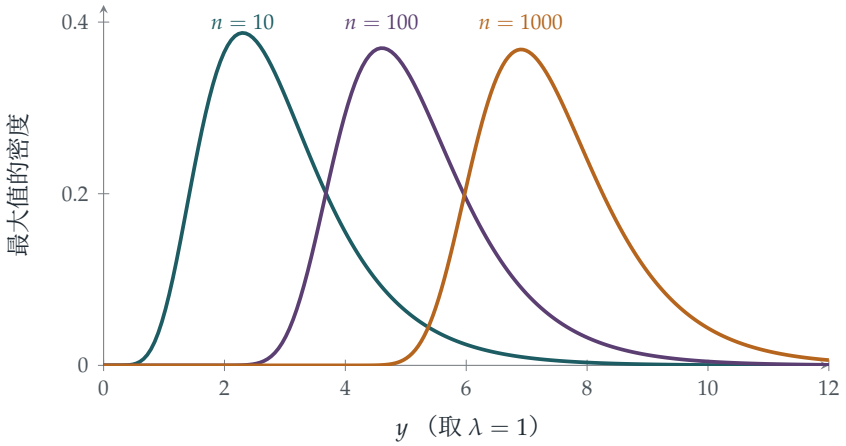


圖 12.25.1 n 個指數分佈壽命（平均值為 1）中最大者的密度，分別對應 $n = 10$、100 和 1000。n 每增至十倍，曲線便向右平移 $\ln 10 \approx 2.3$，但其形狀和闊度基本上保持不變：這就是平移了 $\ln n$ 的耿貝爾分佈。

回顧

- **所有值都不超過 y 時，最大值亦然。**獨立性使最大值的分佈函數成為一個冪。
- **以反函數求分位數。**解 $G(y) = \alpha$ 便得出區間兩端的精確公式。
- **先近似，再比較。**L 和 U 的近似式都含有 $\ln n$，在計算闊度時互相抵消。

背後的數學

耿貝爾分佈

由於 $\mathrm{P}(\lambda Y - \ln n \le x) = \left(1 - \frac{e^{-x}}{n}\right)^n \to \exp(-e^{-x})$，平移後的最大值 $\lambda Y - \ln n$ 在極限下服從耿貝爾分佈，其分佈函數（distribution function）為 $\exp(-e^{-x})$。它的平均值是歐拉常數（Euler's constant）$\gamma \approx 0.577$。你其實早已遇過它：集齊全部 n 款優惠券所需的時間（Section 9.11）就是 n 個（接近獨立的）等待時間中的最大值，這正是它約為 $n \ln n + \gamma n$ 的原因，而其波動屬耿貝爾型，數量級為 n。

極值理論

費雪（Fisher）、Tippett 和格涅堅科（Gnedenko）有一個非凡的定理（1928–1943）：許多獨立同分佈（independent identically distributed）的變量，其最大值經適當平移和縮放後，極限形狀只可能有三種：耿貝爾型（適用於輕尾分佈，例如指數分佈或正態分佈）、弗雷歇（Fréchet）型（適用於重尾分佈，例如柯西（Cauchy）分佈）和韋布爾（Weibull）型

（適用於有硬性上限的變量）。這是中心極限定理（central limit theorem）（Section 12.24）在極值方面的對應結果，所屬的學問稱為極值理論（extreme value theory）。工程師用它來設計防洪設施和海上平台，以抵禦「百年一遇的巨浪」；保險公司用它為巨災保障定價；氣候科學家則用它來估計熱浪正在怎樣變化。

耿貝爾與洪水

Emil Julius Gumbel（耿貝爾）是德國統計學家，直言不諱地反對納粹，1932 年被迫離開大學教席，最終逃往美國。他在 1930 和 1940 年代把極值理論應用於河流氾濫的研究。他的著作 *Statistics of Extremes*（1958）使這門學問成為實用的工具。荷蘭的三角洲工程（Delta Works）是在 1953 年北海洪災之後為保護荷蘭而興建的，設計時就運用了極值統計學，以抵禦預計約 10 000 年一遇的風暴潮。

思維技巧

最大值與最小值的分佈

對於獨立變量的最大值，$\mathrm{P}(\max \le y) = \prod \mathrm{P}(X_i \le y)$；對於最小值，$\mathrm{P}(\min > y) = \prod \mathrm{P}(X_i > y)$。

這一個想法就能處理最近的超級市場（Section 12.19）、第一班巴士（Section 12.18）和壽命最長的燈泡。在大學裏，它是次序統計量（order statistics）和極值理論的起點。

延伸練習

1 **最大值的期望值。**證明 $\mathrm{E}(Y) = \frac{1}{\lambda}\left(1 + \frac{1}{2} + \cdots + \frac{1}{n}\right)$。（提示：根據無記憶性（memorylessness），n 個壽命中的第一個結束之後，餘下的 $n-1$ 個會重新開始。）

2 **最小值。**證明 $X_1, \ldots, X_n$ 中的最小者服從參數為 $n\lambda$ 的指數分佈。

3 **中位數的精確值。**求 Y 的精確中位數，並在 $n = 10$ 時與近似值 $\frac{1}{\lambda}(\ln n - \ln \ln 2)$ 比較。

4 **均勻分佈的最大值。**若 $X_1, \ldots, X_n$ 在 $[0,1]$ 上均勻分佈（uniform），求它們的最大值 M 的分佈，並證明當 n 很大時，$n(1-M)$ 近似服從指數分佈。這屬於三種極值形狀中的哪一種？

延伸閱讀

- **Stuart Coles, *An Introduction to Statistical Modeling of Extreme Values***（Springer）。標準的入門書，附有洪水和海平面的真實數據。
- **Emil Julius Gumbel, *Statistics of Extremes***（Dover）。經典之作。
- **David Spiegelhalter, *The Art of Uncertainty***（Pelican），談罕見事件以及應該怎樣思考它們。

我們學到了甚麼

本章開首我們說過：不確定性是可以精確地推理的，但單憑直覺卻不行。我們提出了三個問題：題目問的究竟是甚麼？巧妙的切入點在哪裏？長遠而言會發生甚麼事？以下是本章二十五道題目所揭示的答案。

1. 題目問的究竟是甚麼？

首四節談的都是如何把問題弄清楚。「已知取出的是一塊有編號的圓片」與「已知取出的是編號為 3 的圓片」，兩者給出不同的答案（Section 12.1）。一種罕見疾病的檢測即使準確度達 90%，檢測呈陽性的病人真正患病的可能性也只有 9%，因為基本比率（base rate）舉足輕重（Section 12.2）。「隨機的弦」在我們說明它是如何選取之前，根本沒有期望長度可言（Section 12.3），而折斷的木棒（Section 12.21）說明的也是同一道理。此外，「獨立」（independent）原來有好幾種意思：事件可以兩兩獨立而並非全體獨立（Section 12.4），而不相關（uncorrelated）的變量也不一定獨立。在每一個例子中，解決之道都是確切地寫下有哪些可能的結果，以及它們是如何產生的。

2. 巧妙的切入點在哪裏？

許多題目都既有冗長的計算，也有簡短的洞見。

- **對稱性（symmetry）。**過程本身無法分辨的兩個座位（Section 12.5）；「正面較多」與「反面較多」是互補的鏡像（Section 12.6）；每一對選手在任何一場比賽中相遇的可能性均等（Section 12.7）；對任何硬幣而言，HT 與 TH 出現的可能性相同（Section 12.9）；徑向對稱的點，其方向服從均勻分佈（Section 12.22）。
- **指示變量（indicator）與線性（linearity）。**把配對（Section 12.4）、繩結（Section 12.10）和比賽場數（Section 12.7）寫成指示變量之和來點算，毋須求出完整的分佈，便可得到平均值和方差。
- **狀態與首步分析（first-step analysis）。**Penney 遊戲（Section 12.8）、兔子洞穴與 Ecks 教授（Section 12.16），以及姓氏的消亡（Section 12.15），全都是以下一步為條件來求解的。
- **策略。**三人決鬥（truel）中槍法最差的人應該故意射失（Section 12.13）；隨機化的策略無法被對手利用（Section 12.14）；先拒絕最初 $\frac{1}{e}$ 的應徵者是最優的做法（Sec-

tion 12.12）。

3. 長遠而言會發生甚麼事？

本章最後的部分追蹤隨機性如何隨時間演變，以及在大量試驗中的表現。分支過程（branching process）要麼消亡，要麼爆發式增長，而即使是爆發式增長的過程，也可能因偶然而消亡（Section 12.15）。馬可夫鏈（Markov chain）會趨於一個平穩分佈（stationary distribution），並忘卻自己的過去（Section 12.16）。服從指數分佈的等候時間則完全沒有記憶（Section 12.18）。平均值通常會穩定下來，但柯西（Cauchy）分佈卻不然，因為它沒有平均值（Section 12.22）。根據中心極限定理（central limit theorem），大量獨立效應之和會趨於正態分佈（Section 12.24），而這種正態性之中更隱藏着 $n!$ 的斯特靈（Stirling）公式（Section 12.23）。最大值的表現又截然不同：它們按 $\ln n$ 的速度增長，離散程度卻完全不會增長（Section 12.25）。而常數 e、π 和 γ 則不斷出現，在杯子與郵票之中，也在針與繩圈之中。

4. 概率的工具箱

方法	你在哪裏用過
寫下樣本空間（sample space）	條件概率（12.1）；貝特朗（Bertrand）（12.3）
貝葉斯定理（Bayes' theorem）與賠率（odds）	基本比率（12.2）；馬可夫逆推（12.16）
對稱性	12.5、12.6、12.7、12.22
指示變量與線性	12.4、12.7、12.10
首步分析	Penney（12.8）；兔子洞穴（12.16）；分支過程（12.15）
補事件（complement）與界	公平的決定（12.9）
以級數求期望值	郵票與 e（12.11）
參數化並求最優	祕書問題（12.12）
博弈論（game theory）式推理	三人決鬥（12.13）；混合策略（12.14）
點算通往終點的路徑	火柴盒（12.17）
先求分佈函數（distribution function）	12.18、12.19、12.25
把概率看成面積或積分	蒲豐（Buffon）（12.20）；折斷的木棒（12.21）
正態近似	斯特靈（12.23）；中心極限定理（12.24）

下一步

- 第 13 章由概率轉向統計：從數據估計未知的量，由 Section 13.1 開始。
- Section 9.8、Section 9.9 和 Section 9.11 的計數問題，是本章內容在組合數學方面的一面。

- 馬可夫鏈用到第 7 章的線性代數，而分支過程的不動點則用到 Section 8.7 的動力系統理論。

如果你想在入讀大學之前再進一步

- **Joseph Blitzstein & Jessica Hwang, *Introduction to Probability***：一流的入門課程，可在網上免費閱讀，並附有講課影片。
- **Geoffrey Grimmett & Dominic Welsh, *Probability: An Introduction***：簡潔而嚴謹，屬大學一年級程度。
- **Frederick Mosteller, *Fifty Challenging Problems in Probability***：經典謎題，附完整解答。
- **David Spiegelhalter, *The Art of Statistics* 及 *The Art of Uncertainty***：現實世界中的概率與風險。

拉普拉斯（Laplace）說過，概率就是化為計算的常識。本章說明了，之所以往往需要計算，正正是因為常識會失靈；而最好的計算不止於得出數字，它們還能解釋箇中道理。一種對稱性、一個指示變量、一個選得恰當的狀態，或一條極限定理，都能把一道謎題化為一個洞見，而大學的概率課程正是建基於這些洞見之上。

延伸練習答案

以下只是簡短的答案和提示，並非完整解答：它們給出結果和關鍵步驟，方便你核對自己的做法。請先認真做過每道練習，然後才看答案。

12.1 精確地取條件

1. 在 (iv) 和 (v) 中，其餘 $k-1$ 塊圓片都是從餘下的 $b+n-1$ 塊中隨機抽出的，其中 $n-1$ 塊有編號，所以兩者都等於 $\frac{(n-1)\binom{b}{k-2}}{\binom{b+n-1}{k-1}}$。(vi) 的答案是 $\frac{\binom{n}{2}\binom{b}{k-2}}{\binom{b+n}{k}-\binom{b}{k}}$。設 $N=b+n$，兩者只在 $(2N-nk)\binom{N-1}{k-1}=2k\binom{b}{k}$ 時才相等。這只會在個別特殊情況下發生，例如 $(b,n,k)=(2,4,3)$、$(3,3,4)$ 或 $(6,11,3)$，又或者是 $n=2$ 而所有圓片都被抽出這種平凡情況。對於題目所給的數值，等式並不成立。

2. 兩個孩子的（性別，星期幾）組合共有 $14^2=196$ 種，各種組合可能性相等，其中 $196-13^2=27$ 種包含至少一個星期二出生的男孩。在這些組合之中，$7^2-6^2=13$ 種是兩個男孩。所以概率是 $\frac{13}{27}$。

3. 你選了 1 號門。若汽車在 1 號門後（概率 $\frac{1}{3}$），換門會輸；若汽車在 2 號或 3 號門後，主持人必須打開另一扇後面是山羊的門，換門便會贏：概率為 $\frac{2}{3}$。若主持人隨機打開一扇門，而門後碰巧是山羊，那麼根據你所見，「汽車在你選的門後」和「汽車在另一扇關着的門後」這兩個結果的可能性相等：概率為 $\frac{1}{2}$。

4. $\mathrm{P}(\geq 2 \mid \geq 1)=\frac{1-\mathrm{P}(0)-\mathrm{P}(1)}{1-\mathrm{P}(0)}\approx 0.370$，但 $\mathrm{P}(\geq 2 \mid \text{葵扇 A})=1-\binom{48}{12}/\binom{51}{12}\approx 0.561$。知道是哪一張 A，會令手上有第二張 A 的可能性增加。

12.2 基本比率謬誤

1. 若屍體在那裏，兩者都說「有屍體」的概率是 pq；若不在，則是 $(1-p)(1-q)$。先驗勝算（prior odds）為 $1:1$，所以後驗概率（posterior probability）是 $\frac{pq}{pq+(1-p)(1-q)}=\frac{pq}{1-p-q+2pq}$。至於第二部分，在三種可能性相等的情況中，有兩種情況下「A 在那裏」為真，所以 $\mathrm{P}(\text{兩者都在} \mid \text{兩者都說}A)=\frac{pq}{2pq+(1-p)(1-q)}$。

2. 我們需要 $0.9\pi > \frac{1}{11}(1-\pi)$，即 $\pi > \frac{10}{109}\approx 9.2\%$。

3. 每一次 AL 測試呈陽性，都會把勝算乘以 $\frac{0.9}{1/11}=9.9$。由先驗勝算 $\frac{1}{99}$ 開始：一次陽性之後，勝算為 0.1（概率 $\frac{1}{11}$）；兩次之後，勝算為 0.99（概率約 0.497）；三次之後，勝算為 9.8（概率約 0.907）。

4. 在 100 名患者之中，99 人測試呈陽性；在 999 900 名沒有患病的人之中，約有 9999 人呈陽性。所以約有 10 100 人被告知可能患病，而當中只有 99 人（約 1%）真正患病。

12.3 「隨機」是甚麼意思？

1. 弦（chord）比內接等邊三角形的邊長 $\sqrt{3}a$ 長的條件分別是：它的遠端落在與近端相對的那三分之一圓周上（$\frac{1}{3}$）；它與 O 的距離小於 $\frac{a}{2}$（$\frac{1}{2}$）；它的中點落在半徑為 $\frac{a}{2}$ 的圓內（$\frac{1}{4}$）。

2. 以 C 為中點的弦垂直於 OC，所以它就是通過 C 而角度為 $t = \frac{\pi}{2}$ 的弦，而 $L(\frac{\pi}{2}) = \frac{4a}{3}\left(0 + \frac{1}{2}\sec^2\frac{\pi}{4}\right) = \frac{4a}{3}$。

3. 弦長為 $2\sqrt{a^2 - d^2}$，平均值為 $\frac{2}{a}\int_0^a \sqrt{a^2 - d^2}\,\mathrm{d}d = \frac{2}{a}\cdot\frac{\pi a^2}{4} = \frac{\pi a}{2}$。

4. 隨機端點：$\frac{2}{\pi\sqrt{4a^2 - \ell^2}}$。隨機距離：$\frac{\ell}{2a\sqrt{4a^2 - \ell^2}}$。隨機中點：$\frac{\ell}{2a^2}$。

12.4 兩兩獨立，卻非相互獨立

1. 展開後得 $\mathrm{E}(XY) - \mathrm{E}(X)\mathrm{E}(Y) = (a - pq)(x_1 - x_2)(y_1 - y_2)$，所以 $a = pq$；其餘三格隨之確定，於是兩者獨立。例子：在 $(0,0)$ 放概率 $\frac{1}{3}$，在 $(\pm1, \pm1)$ 各點放概率 $\frac{1}{6}$。兩個邊際分佈（marginal distribution）都是均勻的，而且 $\mathrm{E}(AB) = 0$，但 $\mathrm{P}(A = 0, B = 1) = 0 \neq \frac{1}{9}$。

2. 共有 21 對，每對相同的概率是 $\frac{1}{6}$，而各指示變量（indicator）兩兩獨立（pairwise independent）。所以平均值是 3.5，方差（variance）是 $21 \cdot \frac{1}{6} \cdot \frac{5}{6} \approx 2.92$，標準差約為 1.71。得分 7 大約比平均值高出兩個標準差。

3. 平均值是 $\binom{k}{2}\frac{1}{365}$。各指示變量兩兩獨立，與骰子的情況完全一樣，所以方差是 $\binom{k}{2}\frac{1}{365}\cdot\frac{364}{365}$。（它們並非相互獨立（mutually independent）：若 1,2 相同，2,3 也相同，則 1,3 亦相同。）

4. 取八個可能性相等的結果，$A = B = \{1,2,3,4\}$，$C = \{1,5,6,7\}$。那麼 $\mathrm{P}(A \cap B \cap C) = \frac{1}{8}$，但 $\mathrm{P}(A \cap B) = \frac{1}{2} \neq \frac{1}{4}$。

12.5 最後一位乘客的座位

1. 對所有 $n \geq 3$，答案都是 $\frac{1}{3}$。最後一個座位現在是 S_1、S_2 和 S_n 其中之一，而每一次隨機選擇對這三個座位都是對稱的。

2. 若 T_1 坐了 S_1，沒有人要讓位。若 T_1 坐了 S_3，則 T_1 和 T_3 都坐錯了位。若 T_1 坐了 S_2，便由 T_2 選擇：選 S_1 有 2 人坐錯位，選 S_3 則有 3 人。期望值（expectation）是 $\frac{1}{3}\left(0 + 2 + \frac{5}{2}\right) = \frac{3}{2}$，與公式給出的 $\frac{2}{3} + \frac{1}{3} + \frac{1}{2}$ 一致。

3. 當 $r \geq 2$ 時 $P_r = \frac{1}{2}$，於是 $\frac{1}{n}\left(1 + \frac{n-2}{2}\right) = \frac{1}{2}$，所以遞推關係（recurrence）與 $P_n = \frac{1}{2}$ 相容。

4. 最後一個空位必定是 S_1 或 S_n（任何其他 S_k 最遲在 T_k 登機時已被佔去）。所以若它不是 S_n，就是 S_1。

12.6 多一枚硬幣

1. 設 H' 為 Bob 首 $n+1$ 次擲幣的結果，X 為他最後一次擲幣的結果。若 $H' > H_A$，或者 $H' = H_A$ 而 X 是正面，Bob 便勝出。所以概率是 $\mathrm{P}(H' > H_A) + \frac{1}{2}\mathrm{P}(H' = H_A)$，而根據「多一枚硬幣」的結果，$\mathrm{P}(H' > H_A) = \frac{1}{2}$。

2. 不成立。當 $n = 0$ 時，Bob 勝出的概率是 p。那個巧妙的證明把正面和反面對調，而這只有對公平硬幣才是一種對稱。

3. $H_A = H_B$ 恰好在 $H_A + (\text{Bob 的反面數目}) = n$ 時成立，即一個 $\mathrm{B}(2n, \frac{1}{2})$ 變量等於 n。數值：$\frac{1}{2}$、$\frac{3}{8}$、$\frac{5}{16}$。

4. 不成立：擲出六點就像一枚 $p = \frac{1}{6}$ 的硬幣。當 $n = 0$ 時概率是 $\frac{1}{6}$，而一般而言概率低於 $\frac{1}{2}$。

12.7 淘汰賽中誰會遇上誰

1. 第 k 輪有 2^{n-k} 場比賽，而根據對稱性，每場比賽由 $\binom{2^n}{2}$ 對球員中任何一對對賽的可能性都相等。把整個淘汰賽的 $2^n - 1$ 場比賽加起來，得 $\frac{2^n-1}{\binom{2^n}{2}} = \frac{2}{2^n}$。

2. 不會：在固定的對陣表中隨機編排位置，對稱性依然保持，同樣的計數方法仍然適用。

3. X 在第 k 輪的對手來自一組 2^{k-1} 名球員，每人從該組勝出的可能性相等。所以 Y 是該對手的概率為 $\frac{2^{k-1}}{2^n-1} \cdot \frac{1}{2^{k-1}} = \frac{1}{2^n-1}$，而 n 輪的總和是 $\frac{n}{2^n-1}$。

4. 共有 $\binom{2n}{2} = n(2n-1)$ 場比賽；每一對都會相遇，所以概率是 1。

12.8 Penney 遊戲

1. HHT 只有在頭兩次擲出 HH 時才會勝出；否則在湊成 HH 之前已出現一個 T，THH 必定先出現。所以 HHT 勝出的概率是 $\frac{1}{4}$。四人一起玩時：HH 之後，HHT 必勝；TT 之後，TTH 必勝；由 HT 和 TH 開始，HTT 和 THH 的勝出機會分別是 $\frac{2}{3}$ 和 $\frac{1}{3}$（或倒過來），於是每名玩家的勝出概率是 $\frac{1}{4} \cdot 1$ 或 $\frac{1}{4}\left(\frac{2}{3} + \frac{1}{3}\right)$，即 $\frac{1}{4}$。TTH 擊敗 THH 的概率是 $\frac{2}{3}$（第 (i) 部分）。

2. 對於 HH：$e_0 = 1 + \frac{1}{2}e_{\mathrm{H}} + \frac{1}{2}e_0$ 及 $e_{\mathrm{H}} = 1 + \frac{1}{2}e_0$，得 $e_0 = 6$。對於 HT：一旦有了一個 H，便等待一個 T（平均擲兩次），而第一個 H 平均要擲兩次：共 4 次。

3. HHH 和 TTT 分別被 THH 和 HTT 擊敗，概率為 $\frac{7}{8}$；HHT 和 TTH 分別被 THH 和 HTT 擊敗，概率為 $\frac{3}{4}$；其餘四個（HTH、HTT、THH、THT）分別被 HHT、HHT、TTH、TTH 擊敗，概率為 $\frac{2}{3}$。HHH 和 TTT 最容易被擊敗。

4. 設 $A = \mathrm{HTT}$，$B = \mathrm{HHT}$，各相關數（correlation）為 $AA = 4$、$AB = 0$、$BB = 4$、$BA = 2$。B 勝出的勝算是 $(AA - AB) : (BB - BA) = 4 : 2$，即概率為 $\frac{2}{3}$。

12.9 用不公平的硬幣作公平的決定

1. 所需輪數服從幾何分佈（geometric distribution）：第 (i) 部分的平均值是 $\frac{1}{p^2+q^2}$，而馮·諾伊曼（von Neumann）的方法（HT 或 TH）的平均值是 $\frac{1}{2pq}$。由於 $p^2 + q^2 - 2pq = (p-q)^2$，對所有 $p \neq \frac{1}{2}$，馮·諾伊曼的方法都較慢。不過，公平的卻是它。

2. 用馮·諾伊曼的技巧製造公平的位元（bit）。每次取兩個：00、01 和 10 分別代表三個選項，11 則表示重新開始。

3. 把擲幣結果逐次與二進制展開式 $\frac{1}{3} = 0.010101\ldots$ 比較（把 H 讀作 1，T 讀作 0）。在第一次出現不同的那一擲，若擲幣結果小於該位數字（T 對着 1），便宣佈「正面」，否則宣佈「反面」。這樣得出正面的概率恰好是 $\frac{1}{3}$，平均用 2 次擲幣。

4. 在未能決定的一輪之後，其中一面有兩個，另一面有一個，而整個過程是這兩個狀態上的馬可夫鏈（Markov chain）。解之得 $\mathrm{P}(\text{酒店}) = \frac{p^3}{p^3+q^3}$，當 $p = \frac{1}{2}$ 時等於 $\frac{1}{2}$：這時是公平的，其他情況則不然。

12.10 把繩頭繫起來

1. 第一次打結形成一個環的概率是 $\frac{1}{3}$（其餘三個繩端之中，有一個屬於同一條繩），而最後一次打結則必定形成一個環。環的數目：2 個的概率是 $\frac{1}{3}$，1 個的概率是 $\frac{2}{3}$；平均值是 $\frac{4}{3}$。

2. 餘下 k 條繩時，一次打結不形成環的概率是 $1 - \frac{1}{2k-1}$。乘積 $\prod_{k=2}^{n} \frac{2k-2}{2k-1}$ 等於 $\frac{4^n}{2n\binom{2n}{n}}$，約為 $\frac{\sqrt{\pi n}}{2n} = \sqrt{\frac{\pi}{4n}}$。

3. 依次插入元素 $1, 2, \ldots, n$：元素 k 要麼開始一個新的循環（cycle）（概率 $\frac{1}{k}$），要麼放在之前 $k-1$ 個元素其中一個之後。把各指示變量加起來便得 H_n。

4. 各次打結的成環指示變量互相獨立，概率為 $\frac{1}{2k-1}$，所以方差是 $\sum \frac{1}{2k-1}\left(1 - \frac{1}{2k-1}\right)$。另外 $\sum_{k\geq 1} \frac{1}{(2k-1)^2} = \frac{\pi^2}{6} - \frac{1}{4} \cdot \frac{\pi^2}{6} = \frac{\pi^2}{8}$。

12.11 郵票、杯子與 e

1. $\sum_{k\geq 2} \frac{k-1}{k!} t^k = te^t - (e^t - 1) = 1 + (t-1)e^t$。求導後得 $\mathrm{E}(N) = G'(1) = e$ 及 $\mathrm{E}(N(N-1)) = G''(1) = 2e$，所以 $\mathrm{Var}(N) = 3e - e^2$。

2. $\mathrm{P}(N=j)$ 在 $k=0,\dots,j-1$ 的 $\mathrm{P}(N>k)$ 中出現，即共 j 次。這裏 $\mathrm{P}(N>k)=\frac{1}{k!}$，所以 $\mathrm{E}(N)=\sum\frac{1}{k!}=\mathrm{e}$。

3. 完全沒有改變：只有一個杯時，Jim 必定「猜」中。期望價格仍然是 e。

4. $\int_0^1\frac{(1-u)^2}{2}\,\mathrm{d}u=\frac{1}{6}$。一般而言，在 $u_k=u$ 處的截面是一個按 $1-u$ 縮放的 $(k-1)$ 維單純形（simplex），所以 $V_k=V_{k-1}\int_0^1(1-u)^{k-1}\,\mathrm{d}u=\frac{V_{k-1}}{k}$。

12.12 何時停止尋找

1. 若最大的在第 $k\geq 2$ 位，你恰好在第一個雪糕是首 $k-1$ 個之中最大時才會得到它，概率為 $\frac{1}{k-1}$。對 k 取平均得 $\frac{1}{n}\sum_{k=2}^{n}\frac{1}{k-1}$，當 $n=4$ 時是 $\frac{11}{24}$。列出全部 24 種次序，可知得到第二、第三和第四大的概率分別是 $\frac{7}{24}$、$\frac{1}{6}$ 和 $\frac{1}{12}$。

2. 拒絕 $r=0,1,2,3$ 人分別給出 $\frac{1}{4}$、$\frac{11}{24}$、$\frac{5}{12}$、$\frac{1}{4}$。所以應拒絕一名候選人，成功概率為 $\frac{11}{24}$。

3. 首 n 人之中最好的一個位列頭兩名的概率是 $\frac{n}{N}+\frac{n}{N}\cdot\frac{N-n}{N-1}$，當 n 很小時接近 $\frac{n}{N}$ 的兩倍。若頭兩名任何一個都可以接受，你便可以少看一些候選人才作決定。（以此為目標的最佳策略，成功概率約為 0.57。）

4. 設 $\mathrm{P}(r)=\frac{r}{N}\sum_{k=r}^{N-1}\frac{1}{k}$：當 $N=10$ 時，$r=2,3,4$ 給出 0.366、0.399、0.398；當 $N=100$ 時，$r=36,37,38$ 給出 0.3710、0.3710、0.3708，其中 $r=37$ 僅僅領先。

12.13 三人決鬥

1. Algernon 擲失後，Bertie 擲向 Cuthbert。若他擲中 $(\frac{3}{5})$，便是 Algernon 對 Bertie，由 Algernon 先擲，Bertie 勝出的概率是 $1-\frac{2/5}{2/5+3/5-6/25}=\frac{9}{19}$。若他擲失，Cuthbert 會擲中 Bertie，然後除非 Algernon 先擲中 $(\frac{3}{5})$，否則 Cuthbert 勝出。所以 Bertie 勝出的概率是 $\frac{3}{5}\cdot\frac{9}{19}=\frac{135}{475}$，Cuthbert 則是 $\frac{2}{5}\cdot\frac{3}{5}=\frac{114}{475}$。

2. 對所有 $a<1$ 都成立。擲失給出 $\frac{3}{5}\cdot\frac{a}{a+\frac{3}{5}-\frac{3}{5}a}+\frac{2}{5}a$；擲中 Cuthbert 給出 $1-\frac{3/5}{3/5+a-\frac{3}{5}a}$（這時 Bertie 先向他擲），這個值總是較小；擲中 Bertie 則給出 0。

3. 不應該。若 Bertie 擲向 Algernon 而擲中，Cuthbert 隨即擲中 Bertie；若他擲失，Cuthbert 照樣擲中 Bertie。無論如何 Bertie 都一無所得，而擲向 Cuthbert 則有 $\frac{27}{95}$。

4. $p+(1-p)(1-r)p$，它小於 $\frac{p}{p+r-pr}$，即整個等比級數（geometric series）之和。

12.14 混合策略

1. 取 $b=(\frac{1}{4},\frac{1}{2},\frac{1}{4})$，則 $X=Y=Z=0$，所以無論對方怎樣做，B 都贏得 0。對着 $(\frac{1}{3},\frac{1}{3},\frac{1}{3})$，A 猜 1、2 和 3 的期望贏款分別是 $-\frac{1}{3}$、0 和 $\frac{1}{3}$，所以 A 應該每次都猜 3。

2. 收益矩陣（payoff matrix）的對角線上是 0，其餘位置是 ± 1，把三個策略循環置換後矩陣不變，所以最優策略是 $(\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$。在原題中，相差 1 和相差 2 的罰則不同，所以各策略不能互換。

3. 若 $X、Y、Z \geq 0$ 且 $X + 2Y + Z = 0$，則三者都是 0。由此必有 $b_1 = b_3$ 及 $b_2 = 2b_3$。

4. 現在 $X = -b_2 + 3b_3、Y = b_1 - b_3$ 及 $Z = -3b_1 + b_2$，而 $X + 3Y + Z = 0$。令三者都等於零，得 $(\frac{1}{5}, \frac{3}{5}, \frac{1}{5})$。

12.15 滅絕還是爆發

1. 滅絕概率（extinction probability）是 $s = (q + ps)^2$ 在 $[0, 1]$ 內的最小根。兩個根是 1 和 $\left(\frac{q}{p}\right)^2$，而第二個根恰好在 $p > \frac{1}{2}$ 時小於 1。

2. 給定 X_{n-1}，X_n 的平均值是 $2pX_{n-1}$，方差是 $2pqX_{n-1}$，所以 $\mathrm{Var}(X_n) = \mathrm{E}(2pqX_{n-1}) + (2p)^2\mathrm{Var}(X_{n-1})$。當 $p = \frac{1}{2}$ 時：$\mathrm{Var}(X_n) = \mathrm{Var}(X_{n-1}) + \frac{1}{2}$，所以 $\mathrm{Var}(X_n) = \frac{n}{2}$。

3. 在第 n 代或之前滅絕，表示第一代每一顆粒子的後代都在 $n - 1$ 代之內消亡。取 $p = \frac{3}{4}$：$e_1 = 0.0625$，$e_2 \approx 0.0881$，$e_3 \approx 0.0999$，遞增並趨向 $\frac{1}{9} \approx 0.111$。

4. $s = \frac{1}{4}(1 + s + s^2 + s^3)$ 可因式分解為 $(s - 1)(s^2 + 2s - 1) = 0$，所以滅絕概率是 $\sqrt{2} - 1 \approx 0.414$。

12.16 順向與逆向的馬可夫鏈

1. 根據對稱性 $p_B = p_D$；另外 $p_A = \frac{1}{3} + \frac{2}{3}p_B$、$p_B = \frac{1}{3}p_A + \frac{1}{3}p_C$ 及 $p_C = \frac{2}{3}p_B$。所以 $p_A = \frac{7}{15}$。由 A 出發，不經過 C 而在 A 離開：$x_A = \frac{1}{3} + \frac{2}{3}x_B$ 及 $x_B = \frac{1}{3}x_A$，所以 $x_A = \frac{3}{7}$。條件概率是 $\frac{3/7}{7/15} = \frac{45}{49}$。

2. 把第一行乘以 P：$(0.19, 0.33, 0.48)$。

3. 跡（trace）是 1.3，行列式（determinant）是 0.02，而 1 是一個特徵值（eigenvalue）（各行之和為 1）。另外兩個特徵值之和為 0.3，積為 0.02：它們是 0.2 和 0.1。

4. $a_1 p_{12} = \frac{1}{20}$ 但 $a_2 p_{21} = \frac{1}{15}$，所以這條鏈不是可逆的（reversible）：倒過來運行時，它的轉移概率並不相同（第 (iii) 部分的 q_{ij} 與 p_{ij} 不同）。

12.17 巴拿赫的火柴盒

1. 若第二次伸手要拿的與第一次是同一種，她當下便發現那包已空（$N = 2$）；否則她拿走另一顆糖果，並在第三次伸手時發現有一包已空。所以 $N = 2$ 或 3，概率各為 $\frac{1}{2}$。

2. $n=1$：$\frac{1}{2}+\frac{1}{2}=1$。$n=2$：$\frac{6}{16}+\frac{6}{16}+\frac{4}{16}=1$。
3. $\binom{2n-1}{n}2^{-(2n-1)}=\frac{1}{2}\binom{2n}{n}\cdot 2\cdot 4^{-n}=\binom{2n}{n}4^{-n}$。
4. 就餘下的拖肥糖和薄荷糖數目作首步遞推（first-step recursion），得 $\frac{64}{81}\approx 0.79$。

12.18 無記憶性

1. $\mathrm{P}(\min\ >\ t)\ =\ \mathrm{e}^{-\lambda_1 t}\mathrm{e}^{-\lambda_2 t}$。等候下一班巴士的時間服從指數分佈（exponential distribution），率為每分鐘 $\frac{1}{15}+\frac{1}{30}=\frac{1}{10}$，所以其中位數（median）是 $10\ln 2\approx 6.9$ 分鐘。
2. $\mathrm{P}(Y\geq k)=(1-p)^k$，所以比值是 $(1-p)^n$。
3. 包含 t 的那段間隔，由回溯至上一班巴士的時間和前望至下一班巴士的時間組成，兩者都服從平均值為 $\frac{1}{\lambda}$ 的指數分佈（當 t 很大時），所以間隔的期望長度是 $\frac{2}{\lambda}$。較長的間隔較有可能包含某個指定時刻。
4. 在 $[0,1]$ 上為 $\frac{\lambda \mathrm{e}^{-\lambda z}}{1-\mathrm{e}^{-\lambda}}$。當 $\lambda=1$ 時，它由約 1.58 緩緩下降至 0.58；當 $\lambda=5$ 時，它集中在 0 附近，由約 5.03 下降至 0.03。

12.19 平面上的隨機點

1. 由 $1-\mathrm{e}^{-\pi k y^2}=\frac{1}{2}$ 得中位數為 $\sqrt{\frac{\ln 2}{\pi k}}\approx\frac{0.470}{\sqrt{k}}$，比平均值 $\frac{1}{2\sqrt{k}}$ 略小。
2. $\mathrm{P}(Y_2>y)=\mathrm{e}^{-\pi k y^2}(1+\pi k y^2)$，所以概率密度（density）是 $2\pi^2 k^2 y^3 \mathrm{e}^{-\pi k y^2}$。
3. 在極坐標下，R 的密度是 $\frac{1}{2\pi\sigma^2}\mathrm{e}^{-r^2/(2\sigma^2)}\cdot 2\pi r$。當 $\sigma^2=\frac{1}{2\pi k}$ 時，它與超級市場的分佈吻合。
4. 取 $k=4\times 10^{-4}$ 每 m^2，完全隨機的假設預測平均值為 $\frac{1}{2\sqrt{k}}=25\,\mathrm{m}$。觀察所得的 35 m 較大，這顯示樹木是互相疏開的（競爭），而不是聚集成叢。

12.20 蒲豐投針

1. 把針切成很多小段。相交次數的期望值具有可加性，而每一小段的貢獻與其長度成正比，所以無論 a 是多少，總和都是 $\frac{2a}{\pi b}$。
2. 根據同樣的可加性（蒲豐投麵條，Buffon's noodle），仍然是 $\frac{2a}{\pi b}$。
3. 圓心與每條邊的距離都必須至少為 r：$\frac{(s-2r)^2}{s^2}$。
4. 由 $\frac{1808}{3408}\approx\frac{5}{3\pi}$ 得 $\pi\approx\frac{5\times 3408}{3\times 1808}=\frac{17\,040}{5424}=\frac{355}{113}$：完美得令人起疑。

12.21 折斷一根棒

1. 要組成三角形，每一段都必須短於 $\frac{1}{2}$。在由折斷點組成的單位正方形中，這個條件劃出一個面積為 $\frac{1}{4}$ 的區域。

2. 假設第一個折斷點在 $x < \frac{1}{2}$（另一情況是對稱的），而較長的一段 $1-x$ 再被均勻地折斷。新的兩段都必須短於 $\frac{1}{2}$，其概率為 $\frac{x}{1-x}$。所以答案是 $2\int_0^{1/2} \frac{x}{1-x}\,\mathrm{d}x = 2\ln 2 - 1$。

3. 當 $(2-x-y)^2 > x^2+y^2$，即 $y < \frac{2-2x}{2-x}$ 時，C 處的角是鈍角。這條曲線與 $x+y=1$ 之間的面積是 $\frac{3}{2} - 2\ln 2$，所以概率是 $3-4\ln 2$，與 B 處的相同。這並非巧合：三條邊各自最多為 1，總和為 2，並且是均勻分佈的，所以它們可以互換。因此 $\mathrm{P}(\text{鈍角三角形}) = 9 - 12\ln 2 \approx 0.68$。

4. 根據容斥原理（inclusion–exclusion），最長的一段 M 滿足 $\mathrm{P}(M > m) = 3(1-m)^2 - 3(1-2m)_+^2 + (1-3m)_+^2$，其中 $x_+ = \max(x,0)$。在 $[0,1]$ 上積分得 $\mathrm{E}(M) = 1 - \frac{1}{2} + \frac{1}{9} = \frac{11}{18}$。

12.22 沒有平均值的分佈

1. $\mathrm{P}(\tan\Theta \le k) = \frac{1}{\pi}\left(\tan^{-1}k + \frac{\pi}{2}\right)$，求導後得 $\frac{1}{\pi(1+k^2)}$。

2. 把 U 和 V 對調，徑向對稱的密度保持不變，所以 $\frac{V}{U}$ 與 $\frac{U}{V}$ 有相同的分佈。

3. 密度是對稱的，所以中位數是 0；而 $\mathrm{P}(C \le 1) = \frac{1}{2} + \frac{1}{\pi}\tan^{-1}1 = \frac{3}{4}$。

4. $\frac{1}{2\pi}\mathrm{e}^{-r^2/2}\,r\,\mathrm{d}r\,\mathrm{d}\theta$，而令 $s = r^2$ 後這就是 $\frac{1}{2\pi}\,\mathrm{d}\theta \cdot \frac{1}{2}\mathrm{e}^{-s/2}\,\mathrm{d}s$：角度是均勻分佈的，$R^2$ 服從平均值為 2 的指數分佈，而且兩者獨立。

12.23 由概率得出斯特靈公式

1. $\sqrt{20\pi}(10/\mathrm{e})^{10} \approx 3\,598\,696$（約低 0.8%）及 $\sqrt{40\pi}(20/\mathrm{e})^{20} \approx 2.4228 \times 10^{18}$（約低 0.4%）。

2. $\frac{(2n)!}{(n!)^2} \approx \frac{\sqrt{4\pi n}(2n/\mathrm{e})^{2n}}{2\pi n(n/\mathrm{e})^{2n}} = \frac{4^n}{\sqrt{\pi n}}$。

3. $100\log_{10}\frac{100}{\mathrm{e}} + \frac{1}{2}\log_{10}(200\pi) \approx 157.97$，所以 100! 有 158 位數字。

4. $\binom{1000}{500}2^{-1000} \approx \frac{1}{\sqrt{500\pi}} \approx 0.0252$。

12.24 中心極限定理

1. 作卷積（convolution），得 $\frac{1}{4}\int \mathrm{e}^{-|x|-|y-x|}\,\mathrm{d}x = \frac{1}{4}(1+|y|)\mathrm{e}^{-|y|}$。方差是 $2+2=4$。

2. $\mathrm{E}(\mathrm{e}^{\theta X}) = \frac{1}{2}(\mathrm{e}^{\theta/2} + \mathrm{e}^{-\theta/2}) = \cosh\frac{\theta}{2}$。標準化後的和 $\frac{2}{\sqrt{n}}\sum X_i$ 的母函數（generating function）是 $\cosh^n \frac{\theta}{\sqrt{n}}$，其對數為 $n\ln\bigl(1 + \frac{\theta^2}{2n} + \cdots\bigr) \to \frac{\theta^2}{2}$。

3. 把 Y 寫成兩個獨立的 $\Gamma(82,1)$ 變量之差，再作數值積分，得 $\mathrm{P}(|Y| \geq 25) \approx 0.051$：正態近似（normal approximation）非常準確。

4. 它沒有矩母函數（moment generating function），甚至連平均值也沒有。事實上，n 個獨立柯西（Cauchy）變量的平均值本身也服從柯西分佈：取平均完全不會令它變得集中。

12.25 眾多之中的最大值

1. n 個壽命之中，最先完結的一個平均需時 $\frac{1}{n\lambda}$；之後餘下 $n-1$ 個，一切重新開始，如此類推。所以 $\mathrm{E}(Y) = \frac{1}{\lambda}\bigl(\frac{1}{n} + \frac{1}{n-1} + \cdots + 1\bigr)$。

2. $\mathrm{P}(\min > t) = (\mathrm{e}^{-\lambda t})^n = \mathrm{e}^{-n\lambda t}$。

3. 由 $(1 - \mathrm{e}^{-\lambda y})^n = \frac{1}{2}$ 得 $\lambda y = -\ln(1 - 2^{-1/n})$。當 $n = 10$ 時這是 2.704，而近似值為 $\ln 10 - \ln\ln 2 \approx 2.669$。

4. $\mathrm{P}(M \leq m) = m^n$，所以 $\mathrm{P}\bigl(n(1-M) > s\bigr) = \bigl(1 - \frac{s}{n}\bigr)^n \to \mathrm{e}^{-s}$。這屬於韋布爾（Weibull）型：各變量有上界，而最大值會擠向這個上界。

章

13

統計與現實世界中的隨機性

概率由模型推向數據，統計則由數據返回現實世界。

本章內容

13.1	解讀新聞標題	STEP I 1992
13.2	點算田鼠	STEP II 2000
13.3	德國坦克問題	STEP III 1998
13.4	平均數還是中位數？	STEP II 2000
13.5	問尷尬的問題	STEP II 2023
13.6	累積起來的誤差	STEP II 1993
13.7	製造相關性	STEP III 2010
A	練習答案	第 1033 頁

從數據中學習

所有模型都是錯的，但有些是有用的。

— 統計學家 George Box

在概率（probability）中，我們知道骰子怎樣被動了手腳，然後問會發生甚麼事。在統計（statistics）中，我們看到已發生的事，然後問骰子必定是怎樣被動了手腳。概率由因推果，統計則由果溯因。對總體的每一個估計、每一次民意調查、每一項臨牀試驗、每一次科學測量，都是一個統計推斷（statistical inference）：根據有限而帶有雜訊的數據，對現實世界作出結論。

與數學的大部分分支相比，統計是一門年輕的學科。它的現代面貌是在二十世紀初由皮爾遜（Karl Pearson）、費雪（Ronald Fisher）、奈曼（Jerzy Neyman）等人建立的，背後的推動力往往是實際問題：釀造啤酒（t 檢驗就是在健力士（Guinness）啤酒廠發明的）、種植農作物、測試藥物。時至今日，隨着數據量爆炸式增長，統計及其近親機器學習（machine learning）已成為數學最重要的應用之一。

統計為何重要

有關藥物、公共衞生、經濟政策、氣候和司法的決定，全都建基於統計證據。誤讀這些證據會帶來實實在在的後果：病人被假陽性（false positive）結果嚇壞（Section 12.2），政策建基於一次偶然的波動(Section 13.1)，電腦系統中的微小誤差不斷累積(Section 13.6)。明白數據能告訴你甚麼、不能告訴你甚麼，以及任何結論有多不確定，是作為一個知情公民的一部分，對科學家、醫生、工程師和經濟學家而言更是專業上的必需。

你在大學會遇到甚麼

統計的入門課程涵蓋估計（estimation，即為未知量找出好的估計值，並說明其精確程度）、假設檢驗（hypothesis testing，即判斷數據是否與某個假設相符）、置信區間（confidence interval）、迴歸（regression，即把一個變量與其他變量聯繫起來），以及實驗和調查的設計。核心概念是*抽樣分佈*（sampling distribution，即估計值本身是一個隨機變量（random variable））、*偏差*（bias）與*方差*（variance），以及*似然*（likelihood）。其後的課程會加入貝葉斯推斷（Bayesian inference）、多元分析（multivariate analysis）、時間序列（time series）和機器學習的方法，而機器學習其實就是規模極其龐大的統計。

貫穿本章的問題

三大問題

1. **是真有其事，還是只是雜訊？**單憑偶然，我們應預期有多大的變動？一項變化要到甚麼程度才值得認真看待？
2. **我們應該怎樣估計？**有了數據，一個未知量的最佳估計是甚麼？我們又怎樣比較不同的方法？
3. **隨機性怎樣進入數據？**數據的收集、保護或處理方式，會對我們的結論造成甚麼影響？

本章先從以批判的眼光閱讀報章上的統計數字開始（Section 13.1），然後轉向估計：用捉放法（capture–recapture）點算田鼠的數目、由序號估計總數，以及比較平均值（mean）與中位數（median）。最後是三道關於隨機性怎樣進入數據的題目：保護私隱的隨機化調查、會累積或互相抵消的捨入誤差（rounding error），以及構造具有指定相關性的變量。

本章的概念

§	課題	概念
13.1	解讀統計數字	比率、自然變異、趨勢
13.2	捉放法	最大似然（maximum likelihood）
13.3	序號	偏差、方差與充分性（sufficiency）
13.4	估計中心位置	效率（efficiency）與穩健性（robustness）
13.5	隨機化回答（randomised response）	混合分佈（mixture）與總方差（total variance）
13.6	捨入誤差	系統誤差（systematic error）與隨機誤差（random error）
13.7	相關（correlation）	楚列斯基（Cholesky）構造法

Section 13.7 會用到協方差（covariance）和相關，這兩個課題已於 2019 年從 STEP 考試範圍中刪去；該節會解釋所需的知識。

怎樣閱讀本章

請記住，每一個估計值都是隨機變量：假如數據有所不同，估計值也會隨之不同。對於本章的每一種方法，都要問一問：重複很多次之後，它的分佈會是怎樣的？平均而言它是否正確？它的變動又有多大？這一個問題，正是統計思維的核心。

13.1 解讀新聞標題

題目

STEP I 1992，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/92-s1-q14`

引子

統計學（statistics）是從數據中學習的科學，而它的第一課就是：數字本身不會說話。「安全六倍」、「升幅驚人」、「死傷日益慘重」：新聞標題每天都把數字變成論斷，而這些論斷往往是錯的；錯不在數字本身有假，而在於比較的對象不對，或是把尋常的隨機波動（random variation）誤當成趨勢。

這道題完全沒有公式，在 STEP 中甚為罕見。它要求你像統計學家那樣思考：要證明每個標題成立，我們需要知道甚麼？單憑偶然，我們又應預期有多大的波動？答案要用到泊松分佈（Poisson distribution）（Section 12.18），而當中的推理正是本章其餘部分的基礎。

動筆之前

想一想 假如死亡事件以穩定的比率獨立而隨機地發生，每年的總數應有多大的變動？若平均每年為 1080，每年與平均值的典型偏差是多少？

獨立罕見事件的計數近似服從泊松分佈，而平均值為 μ 的泊松計數，其方差（variance）為 μ，標準差（standard deviation）為 $\sqrt{\mu}$。取 $\mu = 1080$，標準差約為 33：每年總數落在約 1015 至 1145 之間，都不足為奇。

逐步解題

階段 1　「坐飛機比步行安全六倍」

> **想一想**　每年有 1080 名行人死亡，180 人死於空難。六倍，究竟是甚麼的六倍？

這個標題比較的是死亡總數，但安全與否關乎的是風險：每單位風險承擔量（exposure）的死亡人數。幾乎人人每天都步行，坐飛機的人卻相對較少，而且只是偶一為之。公平的比較應採用每公里行程、每小時行程或每次行程的死亡人數，而這幾種算法可以得出截然不同的答案（按每公里計，航空旅行極為安全；按每次行程計則沒有那麼出色，因為大多數空難死亡都發生在起飛和降落之時）。題目所給的數字根本不能「證明」兩者孰更安全。此外，「在道路意外中喪生的行人」是個狹窄的類別；整體的道路死亡人數要大得多。

階段 2　「我們的天空更安全：1990 年僅 125 人喪生」

> **想一想**　空難死亡像獨立的罕見事件嗎？你預期每年總數的變動有多大？

空難死亡是成羣出現的：一次墜機就可能奪去數百人的性命。因此每年的總數主要取決於墜機的次數，而這個次數很小，於是總數每年的變動極大。某一年有 125 人死亡，相對於平均的 180 人，與情況毫無改變完全吻合：只要多一次墜機，這個數字就會徹底改觀。一年的數據不能顯示趨勢。

階段 3　「道路死傷日益慘重：星期二 7 人喪生」

> **想一想**　平均每天有多少名行人喪生？若每天的死亡人數服從泊松分佈，一天之內出現 7 人或以上的情況有多常見？

平均每天有 $\frac{1080}{365} \approx 2.96$ 名行人死亡。若每天的計數服從這個平均值的泊松分佈，則 $P(\geq 7) \approx 0.031$：純粹出於偶然，每年就約有 11 天會有 7 名或以上的行人死亡。（而且「道路死傷」想必還包括司機和乘客，那麼一天之內有 7 人死於道路意外就更不值得大驚小怪了。）單單一天的數字，對趨勢毫無說明作用。這個標題從一個變動極大的數列中挑出一個極端值，是典型的謬誤。

階段 4　「升幅驚人：1990 年 1350 名行人喪生」

想一想 把 1350 與泊松標準差比較。這純屬偶然嗎？若不是，還有甚麼可以解釋它？這又足以稱為「升幅驚人」嗎？

這是全題最有分量的部分。在穩定的泊松模型之下，1350 比平均值 1080 高出 270，約為 8 個標準差：這極不可能是偶然所致。所以很可能確實有些事情改變了。但這個標題仍然言過其實，原因有以下幾方面：

- *被平均值掩蓋的趨勢*。1080 這個數字是 16 年的平均值。假如死亡人數由 1974 年的約 900 穩步上升至 1989 年的約 1260，平均值仍然是 1080，而 1990 年的 1350 只不過是延續既有的趨勢，並非突如其來的「驚人」升幅（圖 13.1.1）。我們需要逐年的數字。
- *風險承擔量*。人口增長、步行的人增多、交通更繁忙或人口老化，都會令死亡人數上升，而道路按每次行程計卻沒有變得更危險。按人均計算，或按每公里步行路程計算的比率，才是恰當的量度。
- *定義*。死亡個案的記錄或分類方式一旦改變（例如意外發生後多久之內死亡才計算在內），就可以造成數字的跳升。
- *過度離散*（*over-dispersion*）。現實中每年計數的變動比泊松模型所顯示的要大，因為各種條件（天氣、經濟活動）都在變。對「8 個標準差」這個說法應有所保留。

升幅很可能是真實的，但它是否驚人、成因為何，則需要更多數據才能判斷。

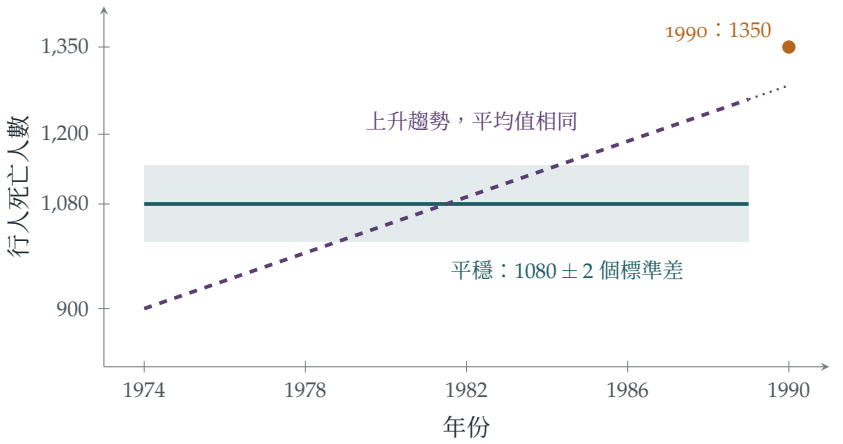


圖 13.1.1 兩段截然不同的歷史，16 年的平均值同樣是 1080。假如死亡人數一直平穩（藍綠色，連同泊松分佈典型波動的範圍帶），1990 年的 1350 就是急劇的跳升。假如死亡人數一直穩步上升（紫紅色），1350 就只是略高於趨勢。單憑平均值，無法分辨這兩種情況。

回顧

- **同類才可相比。**風險需要有分母：每人、每公里或每次行程的死亡人數。
- **認識自然波動。**泊松分佈的變異性，以及成羣出現的事件那大得多的變異性，讓我們知道甚麼時候的變化才值得留意。
- **平均值會掩蓋趨勢。**單憑一個概括的數字，無法描述一個量如何隨時間變化。

背後的數學

訊號與雜訊

每一個統計問題，歸根究底都是要把訊號（signal，即真實的效應或趨勢）與雜訊（noise，即隨機波動）分開。泊松模型為獨立事件計數中的雜訊提供了一把尺：標準差為 $\sqrt{\mu}$。罕見事件的相對波動很大：計數為 9 時，波動約為 ±3，即其大小的三分之一；而計數為 10 000 時，波動只有約 1%。正因如此，小數目（某一天的死亡人數、某一年的墜機次數）才會屢屢被人過度解讀，而小鎮上的癌症「羣組」最終也多半只是偶然。

均值回歸

還有另一個陷阱。道路安全攝影機往往安裝在剛剛經歷了異常惡劣一年的地點。翌年，意外減少，功勞就歸於攝影機。但異常惡劣的一年過後，本來就很可能迎來較為平常的一年：這就是均值回歸（regression to the mean），最早由高爾頓（Francis Galton）在 1880 年代提出。要把攝影機的效果與均值回歸分開，就必須與沒有安裝攝影機的類似地點作比較，這正是對照實驗（controlled experiment）背後的理念。

南丁格爾的統計學

南丁格爾（Florence Nightingale）以克里米亞戰爭中的護士身分聞名於世，同時也是統計學的先驅。她在 1858 年繪製的「玫瑰圖」顯示，死於可預防疾病的士兵遠多於死於創傷的士兵，並說服了政府改革軍隊的衞生狀況。她堅持使用比率（每年每千名士兵的死亡人數）而非原始總數，並堅持同類才可相比，這些正是本題所考核的要點。她是首位獲選為皇家統計學會（Royal Statistical Society）會士的女性。

思維技巧

問一句：「跟甚麼相比？」

在接受一個統計論斷之前，先問清楚：比較的對象是甚麼，分母是甚麼，以及單憑偶然會預期有多大的波動。

到了大學，統計學課程會把這些問題化為精確的方法：假設檢驗（hypothesis test）探究一項觀察結果是否與偶然相符；置信區間（confidence interval）說明一個量測量得有多精確；而實驗設計（design of experiments）則確保比較公平。

延伸練習

1 **泊松分佈的尾部。**設 X 服從平均值為 2.96 的泊松分佈，驗證 $\mathrm{P}(X \geq 7) \approx 0.031$。你預期每年有多少天會有 9 名或以上的行人死亡？

2 **成羣出現的死亡。**假設墜機事故按泊松過程（Poisson process）發生，比率為每年 1.2 宗，而每宗奪去 150 人的性命。求每年空難死亡人數的平均值和標準差。一年之內有 125 人或以下死亡，是否不尋常？有沒有可能某一年恰好有 125 人死亡？

3 **顯著的升幅？**假如 1974–1989 年的數字一直平穩地維持在 1080 ± 33，試（利用正態近似（normal approximation））求單一年份出現 1350 或以上的概率。

4 **擬一個更好的標題。**為四個標題逐一改寫，各擬一個數據真正支持的標題。

延伸閱讀

- **David Spiegelhalter, *The Art of Statistics***（Pelican）。講述如何從數據中學習，並附有大量取自新聞的例子。
- **Tim Harford, *How to Make the World Add Up***（Bridge Street）。思考傳媒中的數字的十條法則。
- **Darrell Huff, *How to Lie with Statistics***（Penguin）。1954 年出版的短篇經典，至今仍不過時。

13.2 點算田鼠

題目

STEP II 2000，第 13 題

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

`step.maths.org/questions/00-s2-q13`

引子

一個島上住了多少隻田鼠？逐隻去數是不可能的：牠們會躲藏，會走動，而且長得一模一樣。不過，你可以先捉一些，做上記號，放走，然後再捉一批。假如有 200 隻做了記號，而第二次捉到的 200 隻當中有 11 隻帶有記號，那麼整個種羣中大約有 $\frac{11}{200}$ 帶有記號，所以種羣的大小約為 $200 \times \frac{200}{11} \approx 3636$。

這種*標記重捕法*（capture–recapture）用於估計魚類、鳥類和鯨魚的數量，也用於估計吸毒者、無家者和戰爭傷亡者的人數，即任何無法完全點算的總體。這道題用*最大似然*（maximum likelihood）原理為這個直觀的估計打下穩固的基礎。最大似然是統計估計中最重要的思想：選取使觀察到的數據出現概率最大的那個未知量的值。

動筆之前

想一想 種羣共有 N 隻田鼠，其中 200 隻帶有記號。現隨機抽取 200 隻作為樣本。恰好有 11 隻帶有記號的概率是多少？

從 200 隻有記號的田鼠中選出 11 隻，再從 $N - 200$ 隻沒有記號的田鼠中選出 189 隻，而等可能的樣本共有 $\binom{N}{200}$ 個：

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

這就是*超幾何分佈*（hypergeometric distribution）。

逐步解題

階段 1　似然

把含有 N 的二項式係數（binomial coefficient）展開寫出，

$$\binom{N-200}{189} = \frac{(N-200)!}{189!\,(N-389)!}, \qquad \binom{N}{200} = \frac{N!}{200!\,(N-200)!},$$

所以

$$p_N = \binom{200}{11}\frac{200!}{189!}\cdot\frac{((N-200)!)^2}{N!\,(N-389)!} = k\,\frac{((N-200)!)^2}{N!\,(N-389)!}. \qquad \blacksquare$$

把數據固定，視之為未知量 N 的函數時，p_N 稱為*似然*（likelihood）。

階段 2　對整數求最大值

> **想一想**　N 是整數，所以用微積分並不方便。不如比較 p_{N+1} 與 p_N。兩者之比何時大於 1？

大部分階乘都會互相約去：

$$\frac{p_{N+1}}{p_N} = \frac{(N-199)^2}{(N+1)(N-388)}.$$

當 $(N-199)^2 > (N+1)(N-388)$ 時，這個比大於 1，即 $N^2 - 398N + 39601 > N^2 - 387N - 388$，亦即 $11N < 39989$，$N < 3635.4$。所以 p_N 一直遞增至 $N = 3636$，其後遞減：最大似然估計值為

$$\hat{N} = 3636. \qquad \blacksquare$$

這正是 $\frac{200\times 200}{11} = 3636.4$ 的整數部分，也就是那個直觀的估計。

階段 3　第三個樣本

抽取第二個樣本之後，其中 189 隻沒有記號的田鼠也被做上記號，所以現在共有 389 隻田鼠帶有記號。取 $N = 3636$，則有 $3636 - 389 = 3247$ 隻沒有記號。第三個含 200 隻田鼠的樣本中恰好有 j 隻帶有記號的概率為

$$\frac{\binom{389}{j}\binom{3247}{200-j}}{\binom{3636}{200}}.$$

這些概率（對 $j = 0, 1, \ldots, 200$）之和必定等於 1，由此即得該恆等式。 $\blacksquare$

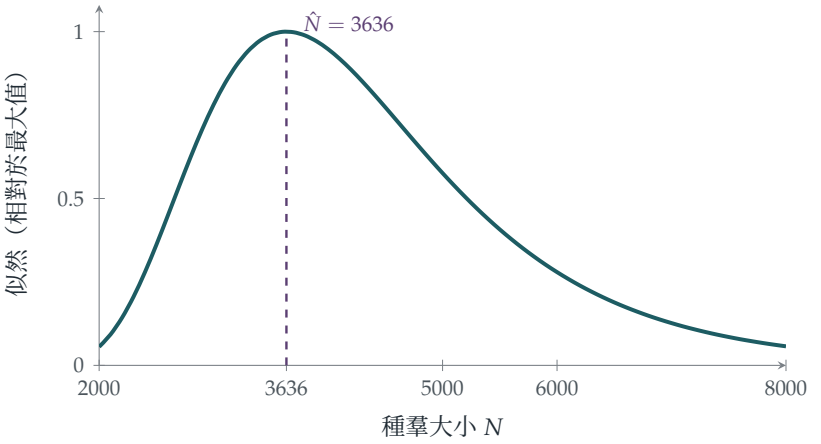


圖 13.2.1 在第二個樣本中發現 11 隻有記號田鼠的似然 p_N，表示為種羣大小 N 的函數，並按比例縮放使其最大值為 1。它在 $N = 3636$ 處達到峯值，但曲線寬闊而且偏斜：由大約 2500 至 6500 的種羣大小，都與數據相當吻合。

回顧

- **把數據的概率看成未知量的函數。**似然把慣常的觀點倒轉過來：數據固定不變，參數（parameter）才是變量。
- **用比值求離散最大值。**比較相鄰兩項，便找出了一個整數變量函數的最大值所在。
- **概率之和為 1。**把一個分佈對所有結果求和，便證明了一個組合恆等式（范德蒙恆等式，Vandermonde's identity）。

背後的數學

最大似然

最大似然原理由費雪（Ronald Fisher）在 1910 至 1920 年代發展出來，內容是：以使觀察到的數據出現概率最大的那個值，作為未知參數的估計。它是至今應用最廣的估計方法，由以曲線擬合數據，到訓練機器學習中的神經網絡，無一不用；神經網絡的「損失函數」（loss function）通常就是負對數似然。費雪證明了，對於大樣本，最大似然估計值近似服從正態分佈（normal distribution），而且其精確程度不遜於任何合理的估計量（estimator），前提是模型具備某些良好性質，而這些性質對大多數日常模型都成立。不過也有重要的例外。在下一節（Section 13.3）中，未知量是可能出現的最大值，即分佈的一個端點，在那裏最大似然估計值的表現大不相同：它的誤差並不近似服從正態分佈，而且誤差是以 $\frac{1}{n}$ 而非 $\frac{1}{\sqrt{n}}$ 的速度縮小。

這個估計有多精確？

圖 13.2.1 顯示似然曲線很寬闊：11 次重捕所提供的信息並不多。粗略計算一下：第二個樣本中有記號田鼠的數目近似服從泊松分佈（Poisson distribution），平均值為 $\frac{200\times 200}{N}$，所以其標準差（standard deviation）約為 $\sqrt{11}\approx 3.3$，相對誤差約為 30%，而這個誤差會傳遞到 N 的估計上。若要把 N 的估計誤差控制在 10% 以內，生物學家需要大約 100 次重捕。統計學家還會修正這個簡單估計的輕微偏差（bias），採用*查普曼估計量*（Chapman estimator）$\frac{(M+1)(C+1)}{R+1}-1$，其中 M 和 C 是兩次捕捉的數量，R 是重捕到的數目。

拉普拉斯點算法國人口

拉普拉斯（Laplace）在 1802 年用了同樣的想法，在沒有人口普查的情況下估計法國的人口。他知道全國一年內登記的出生人數，並且就一批抽樣的市鎮而言，同時知道其人口和出生人數。他假設人口與出生人數之比在各地都相同，從而估計法國人口約為 28 百萬，甚至給出了不確定程度的度量。後來這個方法在動物種羣研究中被重新發現：C. G. J. Petersen 在 1890 年代於丹麥的峽灣為鰈魚做記號，Frederick Lincoln 則在 1930 年代使用鳥環；這個方法至今仍常被稱為*林肯—彼得森*（Lincoln–Petersen）估計量。

思維技巧

選取使數據最有可能出現的參數

要估計一個未知量，先把你所觀察到的結果的概率寫成該未知量的函數，然後求其最大值。

在大學裏，最大似然是統計推斷（statistical inference）的骨幹，並有一套理論說明估計值有多精確（費雪信息量，Fisher information），以及如何比較互相競爭的模型（似然比檢驗，likelihood ratio test）。

延伸練習

1 **一般公式。**若有 M 隻動物做了記號，而第二個含 C 隻動物的樣本中有 R 隻帶有記號，證明 N 的最大似然估計值是 $\frac{MC}{R}$ 的整數部分。

2 **范德蒙恆等式。**考慮從 a 名女士和 b 名男士中選出 c 人組成委員會的方法數目，以此證明 $\sum_j \binom{a}{j}\binom{b}{c-j}=\binom{a+b}{c}$。

3 **擲硬幣。**一枚硬幣擲出正面的概率 p 未知，擲 n 次後觀察到 h 次正面。證明 p 的最大似然估計值是 $\frac{h}{n}$。

4 **沒有重捕。**如果第二個樣本中沒有任何帶記號的田鼠，最大似然估計值會變成怎樣？這說明了這個方法的甚麼特點？

延伸閱讀

- **David Spiegelhalter, *The Art of Statistics***（Pelican），談估計與不確定性。
- **George Seber, *The Estimation of Animal Abundance***（Blackburn）。標記重捕法的標準參考書。
- **John Rice, *Mathematical Statistics and Data Analysis***（Duxbury）。數理統計的入門課本，內容包括最大似然。

13.3 德國坦克問題

題目

STEP III 1998，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/98-s3-q14`

引子

第二次世界大戰期間，盟軍情報部門想知道德國每月生產多少輛坦克。間諜的報告顯示每月約 1400 輛。統計學家則另闢蹊徑：繳獲的坦克在變速箱和其他零件上都刻有序號，而這些序號實際上就是從 $1, 2, \ldots, N$ 中抽出的一個隨機樣本（random sample）。單憑這些數字，他們估計產量約為每月 250 輛。按照一般引用的數字，戰後查閱的德國紀錄顯示為 245 輛。統計學家的估計遠比間諜準確。

本題比較兩種利用序號估計 N 的方法。兩者平均而言都是正確的，但其中一種運用數據的效率遠高於另一種；弄清楚是哪一種，就會引出估計理論（estimation theory）的核心思想：在平均而言正確的估計量（estimator）當中，應選取波動最小的一個。

動筆之前

想一想 設 M 是 $[0, N]$ 上 n 個獨立均勻分佈（uniform distribution）取值中的最大者，那麼 $\mathrm{P}(M \le m)$ 是多少？

$M \le m$ 當且僅當每個 $X_i \le m$，所以當 $0 \le m \le N$ 時，$\mathrm{P}(M \le m) = \left(\frac{m}{N}\right)^n$（可與 Section 12.25 比較）。求導後可知，M 的概率密度（density）為 $\frac{nm^{n-1}}{N^n}$。

逐步解題

階段 1　兩者皆無偏

對於 M：

$$\mathrm{E}(M) = \int_0^N m \cdot \frac{n m^{n-1}}{N^n}\,\mathrm{d}m = \frac{n}{n+1}N, \qquad \text{所以} \qquad \mathrm{E}(Z_1) = \frac{n+1}{n} \cdot \frac{n}{n+1}N = N.$$

對於平均值：每個 X_i 的平均值為 $\frac{N}{2}$，所以 $\mathrm{E}(Z_2) = \frac{2}{n} \cdot n \cdot \frac{N}{2} = N$。 ■

兩個估計量都是無偏（unbiased）的：平均而言，它們會給出正確的答案。因子 $\frac{n+1}{n}$ 用來修正這個事實：觀察到的最大序號總是比 N 略小一點。

階段 2　兩者的方差

> **想一想**　求 $\mathrm{E}(M^2)$，從而求 $\mathrm{Var}(M)$。至於 Z_2，單一個 $[0, N]$ 上的均勻變量的方差（variance）是多少？

$\mathrm{E}(M^2) = \int_0^N m^2 \cdot \frac{n m^{n-1}}{N^n}\,\mathrm{d}m = \frac{n}{n+2}N^2$，所以

$$\mathrm{Var}(M) = \frac{n}{n+2}N^2 - \frac{n^2}{(n+1)^2}N^2 = \frac{nN^2}{(n+1)^2(n+2)},$$
$$\mathrm{Var}(Z_1) = \frac{(n+1)^2}{n^2}\mathrm{Var}(M) = \frac{N^2}{n(n+2)}.$$

$[0, N]$ 上的均勻變量的方差為 $\frac{N^2}{12}$，而各個 X_i 互相獨立，所以

$$\mathrm{Var}(Z_2) = \frac{4}{n^2} \cdot n \cdot \frac{N^2}{12} = \frac{N^2}{3n}. \qquad ■$$

階段 3　哪一個較好？

> **想一想**　比較 $\frac{1}{n(n+2)}$ 與 $\frac{1}{3n}$。當 n 增大時，兩者如何變化？

兩者之比為 $\frac{\mathrm{Var}(Z_1)}{\mathrm{Var}(Z_2)} = \frac{3}{n+2}$，當 $n \geq 2$ 時小於 1（當 $n = 1$ 時則相等，此時兩個估計量完全相同）。所以 Z_1 較好，而且 n 愈大就好得愈多：它的標準差（standard deviation）以 $\frac{1}{n}$ 的速度縮小，而 Z_2 的標準差只以 $\frac{1}{\sqrt{n}}$ 的速度縮小（圖 13.3.1）。若有 $n = 10$ 個序號，Z_1 的標準差約為 N 的 9%，而 Z_2 的約為 18%。

選用 Z_1 還有另一個理由：Z_2 有時會得出荒謬的結果。若序號為 10、20 和 300，則 $Z_2 = \frac{2}{3} \cdot 330 = 220$，比我們實際見過的一個序號還要小。而 $Z_1 = \frac{4}{3} \cdot 300 = 400$ 永遠不會小於最大的觀察值。 ■

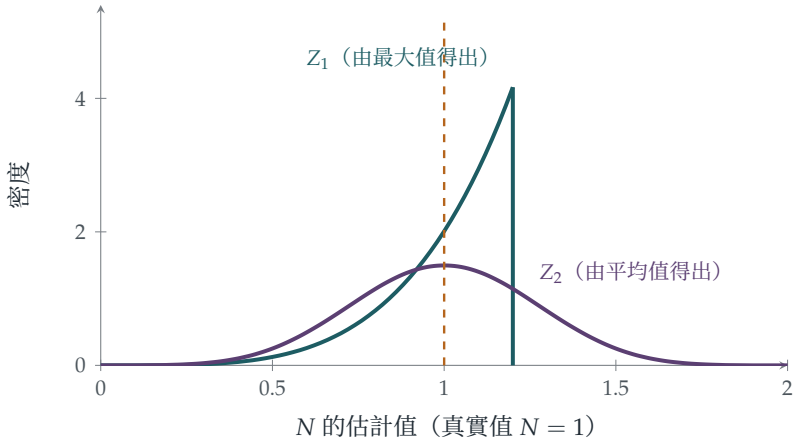


圖 13.3.1 兩個估計量的抽樣分佈（sampling distribution），其中序號數目 $n = 5$，真實值 $N = 1$。兩者的平均值都恰好是 1，但 Z_1（藍綠色）比 Z_2（紫紅色）更緊密地集中在真實值附近，而且它永遠不會低於觀察到的最大序號。

回顧

- **無偏性。**每個估計量都經過縮放，使其期望值（expected value）恰好等於 N。
- **比較方差。**在無偏估計量當中，方差較小的一個較為可靠。
- **選用信息量最大的統計量。**最大值包含了樣本中關於 N 的全部信息；平均值則浪費了其中一大部分。

背後的數學

充分性

為甚麼最大值好這麼多？一旦知道最大值 M，其餘各值不論 N 是多少，都只是在 $[0, M]$ 上均勻分佈：它們再沒有帶來關於 N 的任何信息。最大值是一個*充分統計量*（sufficient statistic）。C. R. Rao 與 David Blackwell 的一條定理（1945–47）指出，任何無偏估計量，只要換成它在給定充分統計量下的期望值，都可以得到改進，或至少不會變差；在這裏，這個做法正好把 Z_2 變成 Z_1。*Lehmann–Scheffé* 定理進一步證明，Z_1 是方差最小的無偏估計量。還要留意，Z_1 的方差以 $\frac{1}{n^2}$ 的速度下降，比大多數估計量的 $\frac{1}{n}$ 更快：這是因為 N 是分佈的端點，數據會擠向這個端點。

離散版本

真實的序號是整數 $1, 2, \ldots, N$，並且是以不放回（without replacement）方式抽樣的。這時最佳的無偏估計是 $\hat{N} = m + \frac{m}{k} - 1$，其中 m 是 k 個觀察到的序號中的最大者：即最大值加上各觀察值之間的平均間距。盟軍統計學家運用這類估計，主要根據變速箱和車輪的序號，估計德國的產量約為每月 246 輛坦克；相比之下，其後在德國紀錄中找到的數字是 245 輛，而情報報告所顯示的是 1400 輛。

歷史上的序號

此後這個方法曾多次派上用場，例如分析員曾根據最先售出的幾件產品的序號，估計某款新消費產品的生產數量。製造商汲取了教訓：現在不少製造商都會把序號編碼，使人無法把它們當作簡單的順序編號來解讀。Richard Ruggles 與 Henry Brodie 於 1947 年在 *Journal of the American Statistical Association* 上記述了這項戰時工作。

思維技巧

以抽樣分佈評價估計量

估計量是一個隨機變量（*random variable*）。要問它平均而言是否正確（偏差），以及它的波動有多大（方差），然後選取誤差較小的一個。

在大學，這會引出有效估計（efficient estimation）的理論：Cramér–Rao 下界（Cramér–Rao bound）指出無偏估計量的方差最小可以小到甚麼程度，而充分性則說明如何達到這個下界。

延伸練習

1. **均方誤差（mean squared error）。**有偏估計量 M 滿足 $\mathrm{E}(M-N)^2 = \mathrm{Var}(M) + \big(\mathrm{E}(M) - N\big)^2$。計算此值，並與 $\mathrm{Var}(Z_1)$ 比較。求使 $\mathrm{E}(cM - N)^2$ 取最小值的常數 c。

2. **最小值也可以。**證明 $(n+1)\min\{X_i\}$ 也是 N 的無偏估計量，並求它的方差。這個估計量好用嗎？

3. **離散公式。**對於序號 $1, \ldots, N$，以不放回方式抽取 k 個，其中最大者為 m。證明 $\mathrm{P}(\max \le m) = \binom{m}{k} / \binom{N}{k}$，並驗證當 $k = 1$ 時 $m + \frac{m}{k} - 1$ 是無偏的。

4. **坦克。**若 5 輛繳獲的坦克的序號為 19、40、42、60、98，試估計坦克的生產數量。

延伸閱讀

- **Richard Ruggles & Henry Brodie, "An empirical approach to economic intelligence in World War II"**，*Journal of the American Statistical Association* 42 (1947)。最初的記述。
- **John Rice, *Mathematical Statistics and Data Analysis***（Duxbury），論及偏差、方差與充分性。
- **David Spiegelhalter, *The Art of Statistics***（Pelican）。

13.4 平均數還是中位數？

題目

STEP II 2000，第 14 題

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

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

引子

要估計一組量度數據的中心，你可以取平均值（mean），也可以取中位數（median）。哪一個較好？答案視乎數據而定。假如量度結果偶爾會出現離譜的誤差，中位數就可靠得多：一個荒謬的數值足以把平均值拉到任何地方，卻幾乎動搖不了中位數。至於規規矩矩的數據，通常是平均值勝出。本題要量度的，正是在數據服從均勻分佈（uniform distribution）時，平均值究竟勝出多少：中位數的方差（variance）是平均值的三倍，所以要令中位數與平均值同樣精確，就需要三倍數量的觀測值。

這個比較引入了*效率*（efficiency）的概念，它是選擇統計方法時的主要考慮之一；同時也用到 Beta 積分（beta integral），這個積分在 Section 4.4 也出現過。

動筆之前

想一想 為甚麼中位數的密度（density）與 $y^n(1-y)^n$ 成正比？（要 Y 接近 y，$2n+1$ 個值之中須有多少個在其下，又有多少個大於 y？）

中位數要落在 y，就必須有一個值在 y，n 個值在其下（各自的概率為 y），另有 n 個值在其上（各自的概率為 $1-y$）。這就得出因子 $y^n(1-y)^n$；常數則是數算哪些值扮演哪個角色的選法數目。

逐步解題

階段 1　常數

密度的積分必須等於 1：

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

階段 2　平均值與方差

密度關於 $\frac{1}{2}$ 對稱（把 y 換成 $1-y$ 即可看出），所以 $\mathrm{E}(Y)=\frac{1}{2}$。

> **想一想**　利用 Beta 積分計算 $\mathrm{E}(Y^2) = k\int_0^1 y^{n+2}(1-y)^n\,\mathrm{d}y$，並化簡當中的階乘。

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

因此

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

階段 3　與平均值比較

> **想一想**　$\bar{X}$ 是 $2n+1$ 個獨立均勻變量的平均，它的平均值和方差是多少？

每個 X_i 的平均值為 $\frac{1}{2}$，方差為 $\frac{1}{12}$，所以 $\mathrm{E}(\bar{X})=\frac{1}{2}$，而 $\mathrm{Var}(\bar{X}) = \frac{1}{12(2n+1)}$。當 n 很大時，

$$\mathrm{Var}(Y)\approx\frac{1}{8n}, \qquad \mathrm{Var}(\bar{X})\approx\frac{1}{24n}.$$

> **想一想**　假設 Y 和 $\bar{X}$ 都服從正態分佈（normal distribution），平均值為 $\frac{1}{2}$，標準差（standard deviation）分別為 σ_Y 和 $\sigma_{\bar{X}}$，試用標準正態分佈函數 Φ 表示這兩個概率。

$\mathrm{P}\big(|Y-\frac{1}{2}|<\frac{d}{\sqrt{n}}\big) = 2\Phi\big(\frac{d}{\sqrt{n}\,\sigma_Y}\big)-1$，而 $\sqrt{n}\,\sigma_Y\to\frac{1}{\sqrt{8}}$；$\bar{X}$ 的情況相似，其中 $\sqrt{n}\,\sigma_{\bar{X}}\to\frac{1}{\sqrt{24}}$。所以兩個概率分別趨向 $2\Phi(d\sqrt{8})-1$ 和 $2\Phi(d\sqrt{24})-1$。由於 Φ 是遞增函數，而且 $\sqrt{24}>\sqrt{8}$，後者較大，因此對所有足夠大的 n，不等式都成立。 ■平均值較有可能接近真正的中心。

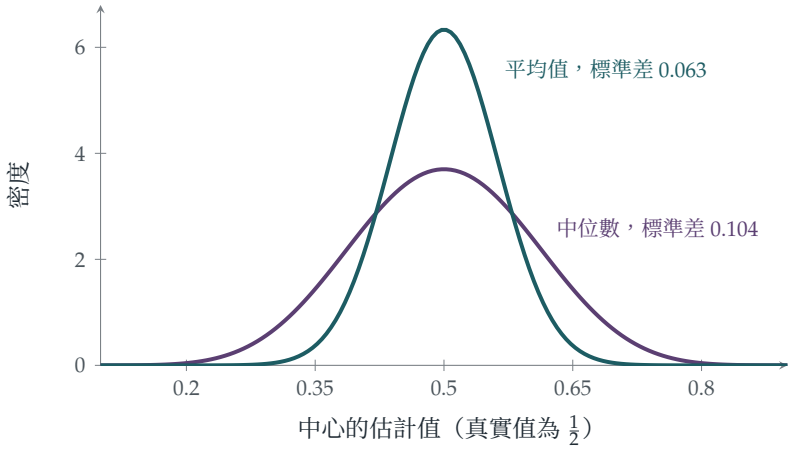


圖 13.4.1 21 個獨立均勻觀測值的中位數（紫紅色）和平均值（藍綠色）的抽樣分佈（sampling distribution）。兩者都以真實值 $\frac{1}{2}$ 為中心，但平均值集中得多：它的方差大約是中位數方差的三分之一。

回顧

- **Beta 積分。** $\int_0^1 y^r(1-y)^s\,\mathrm{d}y$ 這一條公式，同時給出了歸一化常數（normalising constant）和各階矩（moments）。
- **對稱性給出平均值。** 求 $\mathrm{E}(Y)$ 根本不用積分。
- **在合適的尺度上比較離散程度。** 把標準差乘以 $\sqrt{n}$，比較的結果便與 n 無關。

背後的數學

效率

兩個無偏估計量（unbiased estimator）的*相對效率*（relative efficiency）是它們方差之比；它說明其中一個需要多少倍的觀測值，才能與另一個看齊。對於均勻分佈的數據，中位數相對於平均值的效率是 $\frac{1}{3}$。對於正態分佈的數據，這個數字是 $\frac{2}{\pi} \approx 0.64$：中位數大約需要多 57% 的觀測值。但遇上厚尾（heavy-tailed）數據，形勢就逆轉了。對於柯西（Cauchy）分佈（Section 12.22），平均值毫無用處，因為 n 個觀測值的平均並不比單獨一個觀測值好，而中位數卻表現良好。對於拉普拉斯（Laplace）分佈（Section 12.24），中位數的效率是平均值的兩倍。

穩健統計學

中位數是*穩健*（robust）的：即使把接近一半的數據改成任意數值，它也不會移動太遠。平均值卻沒有這種保障：一個錯誤的數值就足以把它毀掉。在 1960 及 1970 年代，John Tukey、Peter Huber 等人發展出*穩健統計學*（robust statistics），即一些估計量，對乾淨的數據而言效率幾乎與平均值一樣高，卻又能抵禦離羣值（outlier），例如*截尾平均值*（trimmed mean）（捨去最大和最小的 10%，再取餘下數據的平均），跳水、體操等運動項目的評分就採用這種方法。

拉普拉斯與中位數

拉普拉斯在 1774 年證明，假如誤差服從雙指數（拉普拉斯）分佈，則真實值的最佳估計是中位數；高斯（Gauss）在 1809 年證明，假如誤差服從正態分佈，則最佳估計是平均值，他並由此得出最小二乘法（method of least squares）。誰是誰非，取決於誤差的表現如何；每遇上一種新的數據，統計學家至今仍會為這個問題爭論一番。

思維技巧

以變異程度比較不同方法

兩種方法可能平均而言都正確；要比較的是它們的答案變動有多大，以及數據增加時這種變動收縮得有多快。

在大學裏，漸近效率（asymptotic efficiency），即比較大樣本下的方差，是在不同統計方法之間作取捨的標準工具；與之並用的還有穩健性，以應付模型可能出錯的情況。

延伸練習

1 **中位數的密度。**用數算的方法推導密度 $\frac{(2n+1)!}{(n!)^2}y^n(1-y)^n$：一個值落在 $[y, y+\delta y]$ 內，n 個在其下，n 個在其上。

2 **三個觀測值。**當 $n=1$（三個均勻觀測值）時，求 $\mathrm{Var}(Y)$ 和 $\mathrm{Var}(\bar{X})$ 的準確值。哪一個較小？

3 **中列數。**中心的另一個估計量是中列數（midrange）$\frac{1}{2}(\max+\min)$。證明對於均勻分佈的數據，若有 $m=2n+1$ 個觀測值，它的方差是 $\frac{1}{2(m+1)(m+2)}$。它與平均值相比如何？（試與 Section 13.3 比較。）

4 **一個離羣值。**十一個量度結果為 9.8、9.9、10.0、10.0、10.1、10.1、10.2、10.2、10.3、10.4 和 104（打字錯誤）。比較平均值和中位數。

延伸閱讀

- **John Rice, *Mathematical Statistics and Data Analysis***（Duxbury），論及順序統計量（order statistics）和效率。
- **Rand Wilcox, *Fundamentals of Modern Statistical Methods***（Springer）。為非專家講解穩健方法為何重要。
- **Stephen Stigler, *The Seven Pillars of Statistical Wisdom***（Harvard University Press）。從歷史角度看數據的匯總、平均值和中位數。

13.5 問尷尬的問題

題目

STEP II 2023，第 11 題

繼續閱讀之前，請先看題目（原題為英文）。題目載於 OCR 提供下載的 2023 年 STEP 試卷（zip 檔案）內的 STEP II 試卷：

`www.ocr.org.uk/Images/707102-2023-past-papers.zip`

引子

怎樣才能知道有多少比例的學生曾在考試中作弊，或者有多少比例的僱員曾偷取公司財物？直接問的話，很多人都會說謊。1965 年，統計學家 Stanley Warner 想出了一個巧妙的解決辦法：隨機化回答（randomised response）。每個人私下擲一枚硬幣或一顆骰子，由結果決定他是如實回答那條敏感問題，還是回答與之相反的問題。訪問員永遠不知道是哪一種情況，所以任何一個人的答案都不足以成為對他不利的證據；然而，由於隨機化的概率（probability）是已知的，整體的比例仍然可以估計出來。

本題探討的正是這背後的數學。讓每個人獨立地（independently）隨機化（第 (ii) 部分），所得的總數性質良好，足以確定該比例。用同一次擲硬幣為所有人隨機化（第 (iii) 部分），所得的總數雖然平均值（mean）相同，分佈（distribution）卻截然不同，呈雙峯形狀，幾乎不能告訴我們任何事。隨機性放在哪裏，至關重要。

動筆之前

想一想 Y 以概率 p 等於 X_1，以概率 q 等於 X_2。$\mathrm{E}(Y)$ 是甚麼？$\mathrm{E}(Y^2)$ 又是甚麼？

以選中哪一個變量為條件，可得 $\mathrm{E}(Y) = p\mu_1 + q\mu_2$，以及 $\mathrm{E}(Y^2) = p\,\mathrm{E}(X_1^2) + q\,\mathrm{E}(X_2^2) = p(\sigma_1^2 + \mu_1^2) + q(\sigma_2^2 + \mu_2^2)$。

逐步解題

階段 1　混合分佈的方差

$$\operatorname{Var}(Y) = p(\sigma_1^2 + \mu_1^2) + q(\sigma_2^2 + \mu_2^2) - (p\mu_1 + q\mu_2)^2.$$

含 μ 的各項為 $p\mu_1^2 + q\mu_2^2 - p^2\mu_1^2 - 2pq\mu_1\mu_2 - q^2\mu_2^2 = pq\mu_1^2 + pq\mu_2^2 - 2pq\mu_1\mu_2$，當中用到 $p - p^2 = pq$ 和 $q - q^2 = pq$。因此

$$\operatorname{Var}(Y) = p\sigma_1^2 + q\sigma_2^2 + pq(\mu_1 - \mu_2)^2. \qquad \blacksquare$$

混合分佈（mixture）的方差（variance），等於各成分內部方差的平均，再加上各成分平均值之間的方差。

階段 2　獨立隨機化

> **想一想**　$\mathrm{P}(B = 1)$ 是多少？B 的 n 個值是否獨立？

$\mathrm{P}(B = 1) = \frac{1}{2}\cdot\frac{1}{6} + \frac{1}{2}\cdot\frac{5}{6} = \frac{1}{2}$。這 n 個值互相獨立，每個的成功概率都是 $\frac{1}{2}$，所以 $Z_1 \sim \mathrm{B}(n, \frac{1}{2})$，其平均值為 $\frac{n}{2}$，標準差（standard deviation）為 $\frac{\sqrt{n}}{2}$。

與平均值相差不超過 10%，即是 $|Z_1 - \frac{n}{2}| < \frac{n}{20}$。以標準單位表示，這就是 $|Z| < \frac{n/20}{\sqrt{n}/2} = \frac{\sqrt{n}}{10}$，而此值趨向無窮大，所以該概率趨向 1。 $\blacksquare$

階段 3　所有人共用一枚硬幣

> **想一想**　現在只擲一次硬幣，便決定全部 n 顆骰子的情況。已知擲得正面，Z_2 的分佈是甚麼？已知擲得反面呢？試用第 (i) 部分的結果。

已知擲得正面，$Z_2 \sim \mathrm{B}(n, \frac{1}{6})$：平均值為 $\frac{n}{6}$，方差為 $\frac{5n}{36}$。已知擲得反面，$Z_2 \sim \mathrm{B}(n, \frac{5}{6})$：平均值為 $\frac{5n}{6}$，方差為 $\frac{5n}{36}$。在第 (i) 部分取 $p = q = \frac{1}{2}$，得

$$\mathrm{E}(Z_2) = \frac{n}{2}, \qquad \operatorname{Var}(Z_2) = \frac{5n}{36} + \frac{1}{4}\left(\frac{4n}{6}\right)^2 = \frac{5n}{36} + \frac{n^2}{9}.$$

> **想一想**　大部分的概率集中在哪裏？在平均值 $\frac{n}{2}$ 附近的概率多嗎？

Z_2 有兩個峯，分別在 $\frac{n}{6}$ 和 $\frac{5n}{6}$ 附近，每個峯的闊度約為 $\sqrt{n}$（圖 13.5.1），而兩峯之間幾乎甚麼都沒有。具有相同平均值和方差的正態分佈（normal distribution）則只有一個峯，中心位於 $\frac{n}{2}$，正正是 Z_2 幾乎從不出現的位置。要 Z_2 與 $\frac{n}{2}$ 相差不超過 10%，一個服從 $\mathrm{B}(n, \frac{1}{6})$ 或 $\mathrm{B}(n, \frac{5}{6})$ 的變量必須落在 $[0.45n, 0.55n]$ 之內。對 $\mathrm{B}(n, \frac{1}{6})$ 而言，這比其平均值

高出至少 $0.28n$，約為 $\frac{0.28n}{0.37\sqrt{n}} \approx 0.75\sqrt{n}$ 個標準差，所以該概率趨向 0；$\mathrm{B}(n, \frac{5}{6})$ 的情況亦類似。 ■

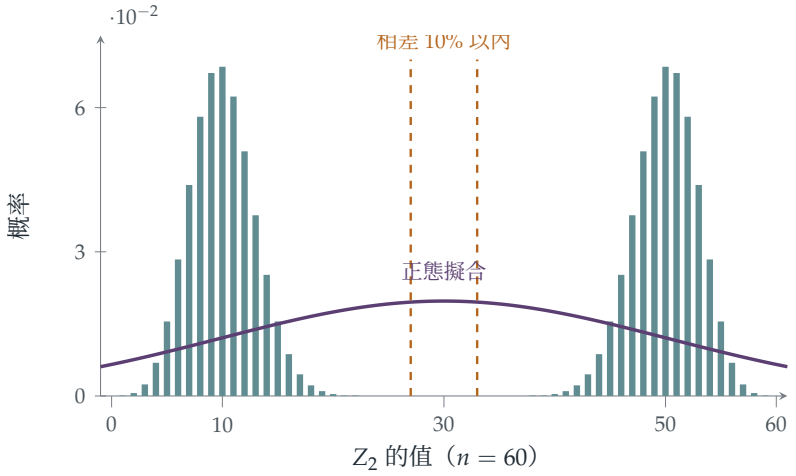


圖 13.5.1 $n = 60$ 顆骰子時 Z_2 的分佈（長條）：兩個峯分別位於 10 和 50，來自那一次擲硬幣的兩種結果。具有相同平均值（30）和方差的正態分佈（紫紅色）擬合得很差，而且幾乎沒有概率落在與平均值相差 10% 以內的範圍（兩條琥珀色線之間）。

回顧

- **以隱藏的選擇為條件。**混合分佈的平均值和方差，都是以選中哪一個成分為條件而求得的。
- **內部方差加上組間方差。**額外的一項 $pq(\mu_1 - \mu_2)^2$ 量度了兩個成分之間的離散程度。
- **平均值和方差並非全部。**Z_2 與 Z_1 的平均值相同，但形狀完全不同。

背後的數學

隨機化回答

在 Warner 的方案中，每位受訪者私下作出選擇，以概率 p 回答「你是否屬於 A 組？」，以概率 $1 - p$ 回答「你是否不屬於 A 組？」。若屬於 A 的人所佔比例為 π，則答「是」的概率為

$$\lambda = p\pi + (1 - p)(1 - \pi),$$

因此由觀察所得答「是」的比例 $\hat{\lambda}$，我們可以估計 $\hat{\pi} = \frac{\hat{\lambda} - (1-p)}{2p-1}$（前提是 $p \neq \frac{1}{2}$）。本題第 (ii) 部分正是這種情況，由每個人自己的骰子來決定。私隱的代價是精確度：除以

$2p - 1$ 會令方差增大，所以隨機化調查所需的受訪者比直接調查更多。不過每個人都是獨立地隨機化的，所以大數定律（law of large numbers）能夠發揮作用。第 (iii) 部分顯示了共用隨機化時會出甚麼問題：那一枚硬幣所帶來的不確定性，永遠不會被平均抵銷。

差分私隱

隨機化回答是*差分私隱*（differential privacy）的前身。差分私隱是 Cynthia Dwork 及其同事於 2006 年提出的一個關於私隱的數學定義：如果加入或移除任何一個人的數據，只會令任何輸出的概率改變至多一個很小的倍數，該程序便算是保障私隱的。科技公司用它來收集使用統計數據，美國人口普查局亦在 2020 年人口普查中採用了它，在公佈的數字上加入經仔細校準的隨機雜訊，令任何個人的答案都無法被還原。

全方差定律

第 (i) 部分的恆等式是*全方差定律*（law of total variance）的一個特例：$\operatorname{Var}(Y) = \operatorname{E}\big(\operatorname{Var}(Y \mid C)\big) + \operatorname{Var}\big(\operatorname{E}(Y \mid C)\big)$，其中 C 是隱藏的選擇。它是*方差分析*（analysis of variance，ANOVA）的基礎。方差分析由費雪（Ronald Fisher）於 1920 年代在 Rothamsted 為農業實驗而引入，它把農作物產量的變異分拆為兩部分：可由肥料或品種解釋的部分，以及未能解釋的部分。

思維技巧

要看整個分佈，而不只是它的概括數值

平均值和方差能很好地描述單峯分佈，但對混合分佈卻可能造成嚴重誤導。要問一問：概率實際上落在哪裏。

在大學裏，混合模型（mixture model）用來描述由幾個羣組組成的總體，而全方差定律則是統計學的標準工具，從調查設計到階層模型（hierarchical model）都會用到。

延伸練習

1 **雙胞胎。**試做 STEP II 2005，第 12 題（`step.maths.org/questions/05-s2-q12`）。這其實是一道喬裝了的隨機化回答謎題：在甚麼情況下，一個答案完全不能透露硬幣的結果？

2 **估計 π。**在一項 $p = \frac{3}{4}$ 的 Warner 調查中，有 400 人作答，其中 160 人答「是」。試估計 A 組所佔的比例 π，並求出你的近似標準誤差（standard error）。

3 **全方差。**設 C 為取多個值的離散隱藏變量，證明全方差定律。

4 **共用硬幣。**在第 (iii) 部分中，$Z_2 > \frac{n}{2}$ 的概率是多少？試解釋為何當 n 很大時，這個概率接近 $\frac{1}{2}$，儘管 Z_2 幾乎從不接近 $\frac{n}{2}$。

延伸閱讀

- **Stanley Warner, "Randomized response: a survey technique for eliminating evasive answer bias"**，*Journal of the American Statistical Association* 60 (1965)。
- **Cynthia Dwork & Aaron Roth, *The Algorithmic Foundations of Differential Privacy***（可於網上免費閱讀）。現代理論；開首幾章不難讀懂。
- **David Spiegelhalter, *The Art of Statistics***（Pelican），談調查及其陷阱。

13.6 累積起來的誤差

題目

STEP II 1993，第 15 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/93-s2-q15`

引子

1982 年，溫哥華證券交易所推出了一個新指數，起始值為 1000。每當有股價變動，指數便會重新計算，並截斷（truncate）至小數點後三位。每次截斷平均損失 0.0005，看來微不足道。可是每天大約要重新計算 3000 次，二十二個月後，儘管大市上升，指數卻只剩下約 520。經正確重算後，指數一夜之間躍升至約 1100。

這道題目解釋了箇中原因。截斷誤差（truncation error）全部朝同一方向，因此會穩步累積，與運算次數成正比。捨入誤差（rounding error）則有正有負，大部分互相抵銷：它們的增長速度只相當於運算次數的平方根，這正是中心極限定理（central limit theorem）（Section 12.24）所預測的隨機漫步（random walk）行為。

動筆之前

? **想一想** 若某數的小數部分 U 在 $[0,1)$ 上均勻分佈（uniform），把該數截斷時誤差是多少？捨入時又是多少？兩者的平均值（mean）和方差（variance）各是多少？

截斷是以 $\lfloor a \rfloor$ 取代 a，誤差（真確值減近似值）為 U，在 $[0,1)$ 上均勻分佈：平均值為 $\frac{1}{2}$，方差為 $\frac{1}{12}$。捨入所得的誤差在 $(-\frac{1}{2}, \frac{1}{2}]$ 上均勻分佈：平均值為 0，方差為 $\frac{1}{12}$。

逐步解題

階段 1　總誤差

加了 1500 個數之後，總誤差是 1500 個獨立（independent）誤差之和，根據中心極限定理，它近似於正態分佈（normal distribution）：

- LEP（捨入）：平均值為 0，方差為 $\frac{1500}{12} = 125$，標準差（standard deviation）為 $\sqrt{125} \approx 11.18$。
- VOZ（截斷）：平均值為 $\frac{1500}{2} = 750$，標準差同樣是 11.18。

階段 2　出現大誤差的概率

想一想　對 LEP 而言，若 $E \sim \mathrm{N}(0, 125)$，$\mathrm{P}(|E| > 15)$ 是多少？對 VOZ 而言，誤差幾乎必然落在哪裏？

對 LEP 而言，$\frac{15}{11.18} \approx 1.342$，所以

$$\mathrm{P}(|E| > 15) \approx 2\big(1 - \Phi(1.342)\big) \approx 2 \times 0.090 = 0.18.$$

對 VOZ 而言，誤差約為 750 ± 11；其絕對值小於 15 的概率基本上是零，所以 $\mathrm{P}(|E| > 15) \approx 1$。 ■

階段 3　可以做多少次加法？

想一想　做了 n 次加法後，LEP 的誤差的標準差為 $\sqrt{n/12}$。何時有 $\mathrm{P}(|E| < 10) \geq 0.9$？

我們需要 $\frac{10}{\sqrt{n/12}} \geq 1.645$（因為 $\mathrm{P}(|Z| < 1.645) = 0.9$），即 $\frac{n}{12} \leq \left(\frac{10}{1.645}\right)^2 \approx 36.95$，所以 $n \leq 443$。LEP 可以做大約 443 次加法。 ■

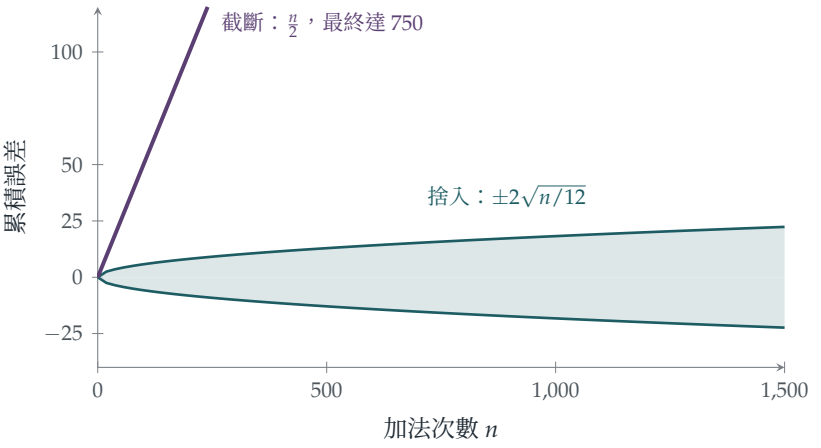


圖 13.6.1 誤差如何累積。截斷誤差（紫紅色）全部同號，與 n 成正比增長：做了 1500 次加法後，總誤差約為 750。捨入誤差（藍綠色帶狀區域，即零的兩側各兩個標準差）平均而言互相抵銷，增長速度只相當於 $\sqrt{n}$：做了 1500 次加法後，誤差通常在約 ±22 之內。

回顧

- **為每個誤差建立模型。**小數部分均勻分佈，令每個誤差都有簡單的分佈。
- **把平均值和方差相加。**平均值可以相加（得出偏差 $\frac{n}{2}$），而對於獨立的誤差，方差同樣可以相加。
- **正態近似。**中心極限定理把 1500 個均勻誤差之和化為正態分佈。

背後的數學

偏差與方差

截斷是*有偏*（biased）的：其誤差的平均值不是零，而偏差（bias）會線性累積。捨入是*無偏*（unbiased）的，其誤差的表現有如隨機漫步，增長速度相當於 $\sqrt{n}$。經過一百萬次運算後，兩者的分別就是數量級為 500 000 的誤差與數量級為 300 的誤差之別。這與實驗中系統誤差（systematic error，例如儀器校準失準）和隨機測量誤差（random error）之間的區別如出一轍：把多次測量結果取平均，可以減少隨機誤差，對系統誤差卻毫無作用。

浮點數

所有電腦都只能以有限的精確度儲存實數，每次算術運算都要把結果捨入。現代電腦遵從 IEEE 754 標準，把結果捨入至最接近的可表示數；若剛好位於兩數正中，則取末位為偶數的一個，目的正是令捨入平均而言無偏。即使如此，把大量數字相加時若不小心處理，仍會損失準確度；William Kahan 的*補償求和*（compensated summation）算法（1965 年）會記錄失去的低位數字，再把它們加回去。1991 年波斯灣戰爭期間，一個愛國者導彈連未能攔截一枚來襲導彈，原因是它的時鐘以十分之一秒為單位計時，而這個數無法以二進制準確儲存；連續運作 100 小時後，截斷誤差已增至約三分之一秒，足以令目標位置的估算偏差超過半公里。

溫哥華指數

溫哥華證券交易所指數於 1982 年 1 月設立，起始值為 1000.000。到 1983 年 11 月，儘管大市普遍上升，指數卻已跌至約 524。顧問發現，指數每天重新計算數千次，而每次都截斷（而非捨入）至小數點後三位。經過一個周末的正確重算，指數重開時約為 1098。這件事已成為數值計算課程中的經典警示。

思維技巧

分清系統誤差與隨機誤差

正負號一致的誤差會線性累積；平均值爲零的獨立誤差則按平方根的速度累積。先認清你面對的是哪一種。

在大學，數值分析（numerical analysis）研究誤差如何在計算過程中傳播，統計學則區分偏差與方差；「偏差–方差權衡」（bias–variance trade-off）是現代機器學習的核心概念。

延伸練習

1. **修正 VOZ。**若 VOZ 在最後把 $\frac{n}{2}$ 加到答案上，其誤差的分佈是甚麼？它現在是否與 LEP 一樣好？

2. **捨入至小數點後一位。**若把各數捨入至小數點後一位，而被捨去的數字均勻分佈，求把 10 000 個數相加時總誤差的標準差。

3 **溫哥華指數。**若指數每天重新計算 3000 次，每次截斷平均損失 0.0005，那麼在一年 250 個交易日內，指數會損失多少？

4 **精確分佈。**求兩個捨入誤差之和的精確密度函數（density，呈三角形）。若是三個誤差之和，試繪出其草圖。它多快便開始貌似正態分佈？

延伸閱讀

- **Nicholas Higham, *Accuracy and Stability of Numerical Algorithms***（SIAM）。關於捨入誤差的權威著作；第一章淺白易讀。
- **David Goldberg, "What every computer scientist should know about floating-point arithmetic"**, *ACM Computing Surveys* 23 (1991)。經典之作，可在網上免費閱讀。
- **Tim Harford, *How to Make the World Add Up***（Bridge Street），探討數據中的系統誤差。

13.7 製造相關性

題目

STEP III 2010，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/10-s3-q13`

引子

一間銀行想知道，手上的股票投資組合(portfolio)有多大機會虧蝕。股價往往一起變動：一隻股票下跌，其他股票也傾向跟着下跌。要模擬數以千計的可能未來，銀行必須產生一批隨機數，而這些隨機數之間的相關程度要恰到好處。如果手上只有互相獨立（independent）的隨機變量（random variable），怎樣才能製造出具有指定相關係數（correlation）的隨機變量呢？

這道題就三個變量的情形，一步一步給出了配方。每個新變量都是先前各個獨立材料的組合，再加上一份全新的獨立材料，用來補足方差（variance）。這個構造其實是線性代數(linear algebra)中一個著名結果的化身，即科列斯基分解(Cholesky decomposition)，它同時揭示了一個隱藏的限制：並非每一組相關係數都有可能出現。

關於課程範圍的說明

協方差（covariance）和相關係數已於 2019 年從 STEP 的考試範圍中刪去。協方差是 $\mathrm{Cov}(U,V) = \mathrm{E}(UV) - \mathrm{E}(U)\mathrm{E}(V)$；獨立變量的協方差為零，而且協方差對每個變元都是線性的：$\mathrm{Cov}(aU+bV,W) = a\,\mathrm{Cov}(U,W) + b\,\mathrm{Cov}(V,W)$。此外 $\mathrm{Var}(U) = \mathrm{Cov}(U,U)$。相關係數必定介乎 -1 與 1 之間。

動筆之前

想一想　Z_1 和 Z_2 互相獨立。它們的協方差是多少？相關係數又是多少？

獨立變量滿足 $\mathrm{E}(Z_1Z_2) = \mathrm{E}(Z_1)\,\mathrm{E}(Z_2)$。所以協方差是 0，而 $\mathrm{Corr}(Z_1, Z_2) = 0$。

逐步解題

階段 1　兩個變量

$\mathrm{E}(Y_2) = 0$。由於 Z_1 和 Z_2 互相獨立，方差均為 1，

$$\mathrm{Var}(Y_2) = \rho_{12}^2 + (1-\rho_{12}^2) = 1, \qquad \mathrm{Cov}(Y_1, Y_2) = \mathrm{Cov}(Z_1, \rho_{12}Z_1 + \sqrt{1-\rho_{12}^2}\,Z_2) = \rho_{12}.$$

兩個方差都是 1，所以 $\mathrm{Corr}(Y_1, Y_2) = \rho_{12}$。 ■

想一想　把 Z_1 和 Z_2 想像成互相垂直的單位向量。Y_2 在哪裏？Y_1 與 Y_2 之間夾角的餘弦是多少？

$Y_2 = \rho_{12}Z_1 + \sqrt{1-\rho_{12}^2}\,Z_2$ 是一個單位向量，與 $Z_1 = Y_1$ 的夾角為 θ，其中 $\cos\theta = \rho_{12}$。把變量看成向量，相關係數就是它們之間夾角的餘弦（圖 13.7.1）。

階段 2　第三個變量

想一想　試以 a、b、c 寫出 $\mathrm{Var}(Y_3) = 1$、$\mathrm{Cov}(Y_1, Y_3) = \rho_{13}$、$\mathrm{Cov}(Y_2, Y_3) = \rho_{23}$ 這三個條件。

利用線性和獨立性：

$$a^2 + b^2 + c^2 = 1, \qquad a = \rho_{13}, \qquad \rho_{12}a + \sqrt{1-\rho_{12}^2}\,b = \rho_{23}.$$

所以

$$a = \rho_{13}, \qquad b = \frac{\rho_{23} - \rho_{12}\rho_{13}}{\sqrt{1-\rho_{12}^2}}, \qquad c = \sqrt{1 - \rho_{13}^2 - \frac{(\rho_{23} - \rho_{12}\rho_{13})^2}{1-\rho_{12}^2}}.$$ ■

（由於 $c \geq 0$，這裏取非負平方根。）

階段 3　任意的均值和方差

縮放和平移都不會改變相關係數，所以

$$X_i = \mu_i + \sigma_i Y_i \qquad (i = 1, 2, 3)$$

的均值為 μ_i，方差為 σ_i^2，相關係數為 ρ_{ij}。 ■

> **停一停，試一試** c 的公式要求平方根內的量非負。試取 $\rho_{12} = \rho_{13} = \rho_{23} = -0.6$。哪裏出了問題？三個變量不可能兩兩之間都有很強的負相關，為甚麼這是合理的？

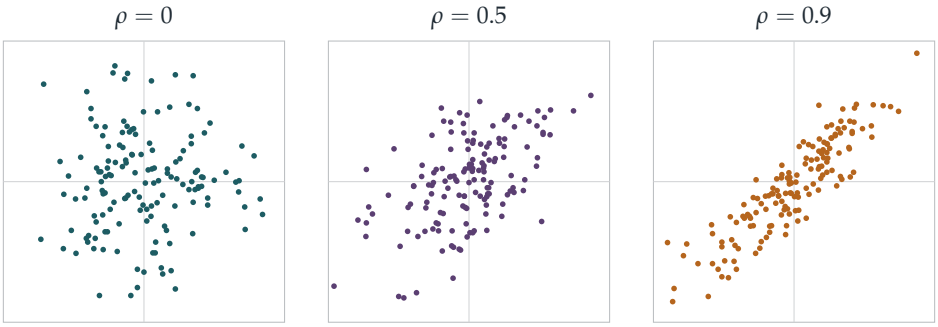


圖 13.7.1 (Y_1, Y_2) 的樣本，由獨立的標準正態（standard normal）變量 Z_1, Z_2 按配方 $Y_1 = Z_1, Y_2 = \rho Z_1 + \sqrt{1-\rho^2}\, Z_2$ 製成，分別對應 $\rho = 0$、0.5 和 0.9。同樣的獨立材料，可以造出我們任意選定的相關係數。

回顧

- **每次只構造一個變量。**每個新變量都從先前的材料中取用適當的份量，再加入一份全新的獨立材料。
- **協方差是線性的。**所有條件都化成了關於係數的簡單方程。
- **相關係數是餘弦。**把標準化後的變量看成單位向量，相關係數就變成了夾角。

背後的數學

科列斯基分解

用矩陣形式來寫，這個構造就是 $\mathbf{Y} = L\mathbf{Z}$，其中的矩陣是下三角（lower triangular）矩陣

$$L = \begin{pmatrix} 1 & 0 & 0 \\ \rho_{12} & \sqrt{1-\rho_{12}^2} & 0 \\ a & b & c \end{pmatrix},$$

而 $\mathbf{Y}$ 的相關矩陣（correlation matrix）是 $R = LL^{\mathsf{T}}$。把一個對稱矩陣寫成 LL^{T} 的形式，就是科列斯基分解，由法國軍事工程師科列斯基（André-Louis Cholesky）約於 1910 年

發現，用來求解大地測量中的方程組（他在第一次世界大戰中陣亡，這項成果在他身後才發表）。只要 R 是正定（positive definite）的，這個分解就行得通，而計算過程正是我們剛才用手做的：逐行求解。在金融、工程和氣候模型中，這是模擬相關變量的標準方法。

哪些相關係數是可能的？

c 的平方根內的數必須非負。展開後，條件是

$$1 - \rho_{12}^2 - \rho_{13}^2 - \rho_{23}^2 + 2\rho_{12}\rho_{13}\rho_{23} \geq 0,$$

這正是相關矩陣的行列式（determinant）非負的條件。從幾何上看：三個單位向量兩兩之間的夾角為 θ_{ij}（其中 $\cos\theta_{ij} = \rho_{ij}$），它們必須真的能在空間中拼合起來。如果 $\rho_{12} = \rho_{13} = \rho_{23} = -0.6$，各個夾角都約為 127°，而三個兩兩相隔 127° 的向量，夾角之和就要超過 360°：這是不可能的。一般而言，一個矩陣要成為相關矩陣，必須是半正定（positive semi-definite）的，即所有特徵值（eigenvalue）均非負（Section 7.2）。如果各個相關係數是根據不完整的數據分別估計出來的，這個條件就可能不成立，統計學家於是要找出最接近的有效相關矩陣，這個問題由 Nicholas Higham 在 2002 年解決。

高爾頓、皮爾遜與相關係數

高爾頓（Francis Galton）在 1888 年研究父母與子女身高的關係時引入了相關的概念，而皮爾遜（Karl Pearson）在 1896 年給出了本題所用的積矩（product moment）公式。高爾頓還注意到均值回歸（regression to the mean）的現象（Section 13.1）。皮爾遜於 1911 年在倫敦大學學院（University College London）創立了全世界第一個大學統計學系。「相關不等於因果」這句警告，幾乎與相關係數本身一樣古老。

思維技巧

逐步構造，每次加入全新的獨立材料

要構造具有指定關係的對象，可以逐個加入，每一個都是先前各個對象的組合，再加上一些全新而獨立的東西。

這是線性代數中的格拉姆–施密特（Gram–Schmidt）正交化過程倒過來做，也就是科列斯基分解。在大學裏，它會出現在最小二乘法（least squares）、多元正態分佈（multivariate normal distribution）的理論，以及隨機過程的模擬之中。

延伸練習

1 **不可能的相關係數。**證明：除非 $\rho \geq -\frac{1}{2}$，否則三個隨機變量不可能兩兩之間的相關係數都是 ρ。對於 n 個變量，相應的結論是甚麼？

2 **檢查行列式。**證明 $1 - \rho_{13}^2 - b^2 \geq 0$ 等價於 $\det R \geq 0$，其中 R 是 3×3 相關矩陣。

3 **投資組合的風險。**兩隻股票回報的標準差（standard deviation）分別為 $\sigma_1 = 0.2$ 和 $\sigma_2 = 0.3$，相關係數為 ρ。若投資組合中兩隻股票各佔一半，求 $\rho = -1$、0 和 1 時該組合的標準差。分散投資達到了甚麼效果？

4 **還原材料。**給定題目中的 Y_1、Y_2，試以 Y_1 和 Y_2 表示 Z_2。（這就是格拉姆–施密特過程中的一步。）

延伸閱讀

- **Gilbert Strang,** ***Introduction to Linear Algebra***（Wellesley-Cambridge），講述正定矩陣和科列斯基分解。
- **Paul Glasserman,** ***Monte Carlo Methods in Financial Engineering***（Springer）。講述實際應用中如何模擬相關變量；內容較深，但開頭幾章不難讀。
- **Theodore Porter,** ***Karl Pearson: The Scientific Life in a Statistical Age***（Princeton University Press）。

我們學到了甚麼

本章開首說過，概率（probability）是由模型走向數據，而統計（statistics）則是由數據回到現實世界。我們提出了三個問題：這是真有其事，還是純屬雜訊？應該怎樣估計？隨機性是怎樣進入數據的？以下是本章七道題目所揭示的答案。

1. 是真有其事，還是純屬雜訊？

在 Section 13.1 中，四則新聞標題把數字變成了論斷。比較總數而不理會分母、單憑一年或一日的數據就看出趨勢、把趨勢隱藏在平均數之中，這些全都是錯誤；而泊松模型（Poisson model）則提供了一把尺，用來量度計數數據的自然變異。上升了 8 個標準差（standard deviation），幾乎肯定是真有其事；某個星期二有七人死亡，則幾乎肯定不是。再加上均值回歸（regression to the mean）和公平比較的必要，整幅圖畫便完整了。

2. 應該怎樣估計？

有三道題目發展了估計的理論。

- **最大似然（maximum likelihood）**（Section 13.2）：選取令觀察所得數據出現的概率最大的未知數值。在捉放法（capture–recapture）中，它給出了自然的估計值 $\frac{200\times 200}{11} \approx 3636$，而寬闊的似然函數則顯示出這個估計有多不確定。
- **偏差（bias）與方差（variance）**（Section 13.3）：潛艇數目有兩個無偏估計量（unbiased estimator），一個以最大值為基礎，另一個以平均值為基礎。最大值是充分統計量（sufficient statistic），精確得多，其方差按 $\frac{1}{n^2}$ 而非 $\frac{1}{n}$ 的速度縮小。同樣的推理，在戰時估計德軍坦克數量一事上，勝過了情報部門。
- **效率（efficiency）與穩健性（robustness）**（Section 13.4）：對於均勻分佈的數據，平均值的效率是中位數（median）的三倍；但對於重尾（heavy-tailed）數據，則是中位數勝出，而且它不受極端誤差影響，十分穩健。

3. 隨機性是怎樣進入數據的？

隨機性可以是刻意加入的：隨機化回答（randomised response）（Section 13.5）既保障每位受訪者的私隱，又仍然可以對總體作出估計，條件是每個人各自獨立地進行隨機化；假如所有人共用同一枚硬幣，便會得出一個雙峯的混合分佈，幾乎不能告訴我們任何東

西。隨機性也會在處理數據的過程中進入：四捨五入的誤差會互相抵銷，只按 $\sqrt{n}$ 的速度增長，而截尾誤差（truncation error）則會線性地累積，溫哥華證券交易所（Vancouver Stock Exchange）便親身領教過（Section 13.6）。隨機性還可以經人手設計出來：把獨立的成分以 Cholesky 構造法（Cholesky construction）組合起來，便可產生具有任何一組有效相關係數（correlation）的變量（Section 13.7），但並非每一組相關係數都是有效的。

4. 統計學的工具箱

方法	在哪裏用過
泊松變異	新聞標題（13.1）
似然及其最大值	田鼠（13.2）
無偏性與方差	序號（13.3）；中位數（13.4）
充分統計量	序號（13.3）
正態近似（normal approximation）	13.4、13.5、13.6
全方差定律（law of total variance）	隨機化回答（13.5）
偏差與隨機誤差	四捨五入（13.6）
協方差（covariance）的線性性質	相關係數（13.7）

下一步往哪裏去

- 貝葉斯定理（Bayes' theorem）（Section 12.2）是貝葉斯統計學的基礎，而貝葉斯統計學正是本章各種方法以外最主要的另一套方法。
- 中心極限定理（central limit theorem）（Section 12.24）為本章處處用到的正態近似提供了理據；柯西分佈（Cauchy distribution）（Section 12.22）則顯示這些近似在甚麼時候失效。
- 相關矩陣（correlation matrix）是半正定矩陣（positive semi-definite matrix），可以用 Section 7.2 的特徵向量（eigenvector）來研究。

如果你想在入讀大學之前再走遠一點

- **David Spiegelhalter, *The Art of Statistics***：為一般讀者而寫的統計思維入門書之中最好的一本。
- **Tim Harford, *How to Make the World Add Up***：教你怎樣解讀新聞中的數字。
- **John Rice, *Mathematical Statistics and Data Analysis***：大學數理統計的第一門課程。
- **Stephen Stigler, *The Seven Pillars of Statistical Wisdom***：統計學的重大思想及

其來源。

George Box 說過，所有模型都是錯的，但有些是有用的。本章嘗試說明怎樣令模型變得有用：知道應該預期多大的變異，以方法在多次重複之下的表現來評價方法，並且思考數據是怎樣得來的。正是這些習慣，把數字變成知識。

最後的話

全書到此結束。在這十三章裏，我們以 STEP 題目為門徑，走進大學一年級課程的數學：數論與分析、微積分與複數、代數與線性代數、微分方程與動力系統、組合數學、力學、概率與統計。每一道題目都是一份邀請，請你追問它背後藏着甚麼，而答案幾乎每一次都是一整門學科。如果你曾經手執鉛筆，認真做過其中哪怕只是幾節，思考頁邊提出的問題，並嘗試每節末的習題，你就已經開始像數學家那樣思考了。餘下的旅程，就由你自己去走。

延伸練習答案

以下只是簡短的答案和提示，並非完整解答：這裏只給出結果和關鍵步驟，方便你核對自己的做法。每道練習都應先認真做過，然後才看答案。

13.1 解讀新聞標題

1. $1 - \sum_{k=0}^{6} \mathrm{e}^{-2.96}\frac{2.96^k}{k!} \approx 0.0315$，所以一年之中大約有 11 天會出現七宗或以上的死亡個案。$\mathrm{P}(X \geq 9) \approx 0.0035$：大約一年一天。

2. 死亡人數為 $150K$，其中 K 服從平均值（mean）為 1.2 的泊松分佈（Poisson distribution）：平均值為 180，方差（variance）為 $150^2 \times 1.2 = 27\,000$，標準差（standard deviation）約為 164。死亡人數為 125 或以下的年份，就是沒有發生空難的年份，其概率為 $e^{-1.2} \approx 0.30$：完全是尋常事。但在這個模型中，死亡人數不可能恰好是 125，因為死亡人數只能是 150 的倍數。平均值和標準差都顯示不出這一點：個案扎堆出現，令分佈變得一團一團，全部概率都集中在 0、150、300 等數值上。

3. $z = \frac{1349.5-1080}{33} \approx 8.2$：概率的數量級為 10^{-16}。穩定不變的發生率無法解釋 1350 這個數字；必定是有些事情真的改變了（又或者是統計死亡人數的方法改變了）。

4. 例如：(i)「單憑總數，無法判斷乘飛機與步行哪樣較安全」；(ii)「1990 年空難死亡人數屬正常範圍」；(iii)「星期二七名行人死亡：不幸的一天，但這樣的日子一年約有十一次」；(iv)「1990 年行人死亡人數急升：必須追查原因」。

13.2 點算田鼠

1. $\frac{p_N}{p_{N-1}} = \frac{(N-M)(N-C)}{N(N-M-C+R)}$，而這個比值不小於 1，當且僅當 $NR \leq MC$。所以 p_N 一直遞增至 $N = \frac{MC}{R}$，然後遞減，而最大值就在該數的整數部分取得。

2. 從 a 名女性和 b 名男性中選出 c 人組成委員會，按其中女性的人數 j 來數：對每個 j，共有 $\binom{a}{j}\binom{b}{c-j}$ 種。

3. $\frac{\mathrm{d}}{\mathrm{d}p}\big(h \ln p + (n-h)\ln(1-p)\big) = \frac{h}{p} - \frac{n-h}{1-p} = 0$ 給出 $p = \frac{h}{n}$。

4. 當 $R = 0$ 時，p_N 永遠遞增，估計值是無限大。這個方法只有在第二個樣本中有足夠多已標記的動物時才行得通；重複捕獲的數目少的話，估計值會非常不穩定。

13.3 德國坦克問題

1. 設 M 為最大值，則 $\mathrm{E}(M) = \frac{nN}{n+1}$ 及 $\mathrm{Var}(M) = \frac{nN^2}{(n+1)^2(n+2)}$，所以 $\mathrm{E}(M-N)^2 = \frac{2N^2}{(n+1)(n+2)}$。這是 $\mathrm{Var}(Z_1) = \frac{N^2}{n(n+2)}$ 的 $\frac{2n}{n+1}$ 倍：較差。最佳的倍數是 $c = \frac{\mathrm{NE}(M)}{\mathrm{E}(M^2)} = \frac{n+2}{n+1}$，其均方誤差（mean squared error）為 $\frac{N^2}{(n+1)^2}$，比 Z_1 稍佳。

2. $\mathrm{E}(\min) = \frac{N}{n+1}$，而 min 的方差與 M 的相同，所以 $\mathrm{Var}\big((n+1)\min\big) = \frac{nN^2}{n+2}$。這個方差不會隨 n 增大而縮小：是個差劣的估計量（estimator）。

3. 最大值不超過 m，當且僅當全部 k 個數都來自 $\{1,\dots,m\}$。當 $k=1$ 時，估計量是 $2X-1$，其平均值為 $2\cdot\frac{N+1}{2}-1 = N$。

4. $98 + \frac{98}{5} - 1 = 116.6$：大約 117 輛坦克。

13.4 平均數還是中位數？

1. 先選定哪一個值是中位數（median）（有 $2n+1$ 種選法），再選定其餘各值中哪 n 個在它之下（有 $\binom{2n}{n}$ 種選法）。所求概率為 $(2n+1)\binom{2n}{n}y^n(1-y)^n\,\delta y = \frac{(2n+1)!}{(n!)^2}y^n(1-y)^n\,\delta y$。

2. $\mathrm{Var}(Y) = \frac{1}{4(2n+3)} = \frac{1}{20}$ 及 $\mathrm{Var}(\bar{X}) = \frac{1}{12(2n+1)} = \frac{1}{36}$。平均值較佳。

3. $\mathrm{Var}(\max) = \mathrm{Var}(\min) = \frac{m}{(m+1)^2(m+2)}$ 及 $\mathrm{Cov}(\max,\min) = \frac{1}{(m+1)^2(m+2)}$，由此得出 $\frac{1}{2(m+1)(m+2)}$。它縮小的速度與 $\frac{1}{m^2}$ 同階，遠快於平均值的 $\frac{1}{12m}$：對於均勻分佈（uniform distribution）的數據，資訊都在極端值之中，與坦克問題的情況一樣。

4. 平均值約為 18.6，被一個打字錯誤完全破壞；中位數是 10.1，不受影響。

13.5 問尷尬的問題

1. 設 $t = ap + bq$，即答案屬實的概率。回答「是」的概率為 $\frac{1}{2}t + \frac{1}{2}(1-t) = \frac{1}{2}$，所以 $\mathrm{P}(\text{正面} \mid \text{是}) = t$，而它等於 $\frac{1}{2}$，當且僅當 $2t = 1$。若由同一名孿生兒回答兩次，則不論硬幣擲出的是正面還是反面，先答「是」後答「否」的概率都是 $pa(1-a)+qb(1-b)$，所以後驗概率（posterior probability）是 $\frac{1}{2}$。兩次都答「是」的概率分別為 pa^2+qb^2（正面）和 $p(1-a)^2+q(1-b)^2$（反面），兩者相等當且僅當 $2(ap+bq)=1$。

2. $\mathrm{P}(\text{是}) = \frac{3}{4}\pi + \frac{1}{4}(1-\pi)$，所以 $\hat{\pi} = \frac{0.4-0.25}{0.5} = 0.3$。標準誤（standard error）為 $\frac{\sqrt{0.4\times0.6/400}}{0.5} \approx 0.049$。

3. $\mathrm{E}(Y^2) = \sum_c \mathrm{P}(C=c)\big(\mathrm{Var}(Y\mid c) + \mathrm{E}(Y\mid c)^2\big)$；減去 $\mathrm{E}(Y)^2 = \big(\sum_c \mathrm{P}(C=c)\mathrm{E}(Y\mid c)\big)^2$ 便得 $\mathrm{E}(\mathrm{Var}(Y\mid C)) + \mathrm{Var}(\mathrm{E}(Y\mid C))$。

4. Z_2 接近 $\frac{n}{6}$（正面）或 $\frac{5n}{6}$（反面），兩種情況的概率各為 $\frac{1}{2}$。所以 $\mathrm{P}(Z_2 > \frac{n}{2}) \to \frac{1}{2}$，儘管 Z_2 幾乎從不接近 $\frac{n}{2}$。

13.6 累積起來的誤差

1. 每個截尾誤差（truncation error）都在 $(-1, 0]$ 上均勻分佈；各加上 $\frac{1}{2}$ 後，便在 $(-\frac{1}{2}, \frac{1}{2}]$ 上均勻分佈，與四捨五入(rounding)的情況完全一樣。所以修正後的 VOZ 與 LEP 一樣好。

2. 每個誤差都在 $(-0.05, 0.05)$ 上均勻分佈，方差為 $\frac{0.1^2}{12}$，所以總和的標準差為 $\sqrt{10\,000 \times 0.01/12} \approx 2.9$。

3. $3000 \times 250 \times 0.0005 = 375$ 點。

4. 兩個誤差時，是 $[-1, 1]$ 上的 $1 - |x|$。三個誤差時：當 $|x| \leq \frac{1}{2}$ 時為 $\frac{3}{4} - x^2$，當 $\frac{1}{2} \leq |x| \leq \frac{3}{2}$ 時為 $\frac{1}{2}\left(\frac{3}{2} - |x|\right)^2$，已經很接近正態曲線（normal curve）。

13.7 製造相關性

1. 對於標準化變量（standardised variables），$\mathrm{Var}(X_1 + X_2 + X_3) = 3 + 6\rho \geq 0$，所以 $\rho \geq -\frac{1}{2}$。對於 n 個變量，$n + n(n-1)\rho \geq 0$ 給出 $\rho \geq -\frac{1}{n-1}$。

2. 取 $b = \frac{\rho_{23} - \rho_{12}\rho_{13}}{\sqrt{1-\rho_{12}^2}}$，則 $(1 - \rho_{12}^2)(1 - \rho_{13}^2 - b^2) = 1 - \rho_{12}^2 - \rho_{13}^2 - \rho_{23}^2 + 2\rho_{12}\rho_{13}\rho_{23} = \det R$。由於 $1 - \rho_{12}^2 > 0$，這兩個條件等價。

3. $\mathrm{Var} = \frac{1}{4}(0.04 + 0.09 + 0.12\rho)$：當 $\rho = -1$、0、1 時，標準差分別為 0.05、0.18 和 0.25。除非兩隻股票完全相關（perfectly correlated），否則這個組合的風險低於兩者風險的平均值（0.25）。

4. $Z_2 = \dfrac{Y_2 - \rho_{12}Y_1}{\sqrt{1-\rho_{12}^2}}$。

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

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。