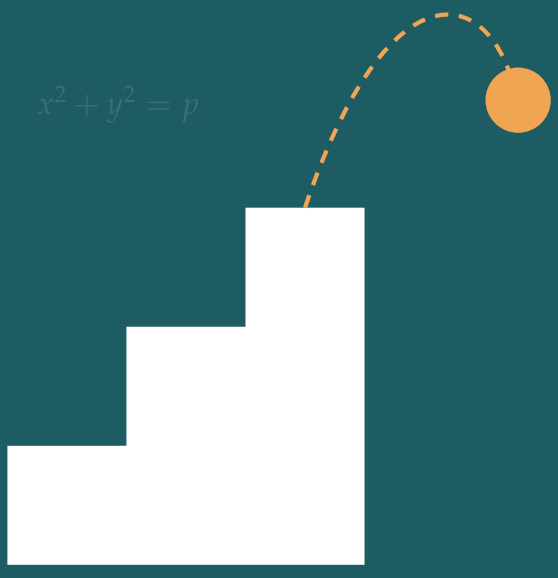
STEP 備試 · 大學數學

Beyond STEP

第 3 章

函數、不等式與嚴謹性



KEITH McNULTY

Beyond STEP

以歷屆 STEP 試題為門
走進大學數學

第 3 章

函數、不等式與嚴謹性

Keith McNulty

Beyond STEP：以歷屆 STEP 試題為門，走進大學數學

第 3 章：函數、不等式與嚴謹性

本冊是 *Beyond STEP* 全書 13 章中的第 3 章。全書、各章的單行本、每一節的可列印工作紙，以及每一道 STEP 試題的詳細解答，均可在 `beyondstep.keithmcnulty.org` 免費閱讀及下載。

2026 年 11 月首次於網上發表。

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

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

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

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 65 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，$-0.668 \lesssim g(x) \lesssim 1.083$。

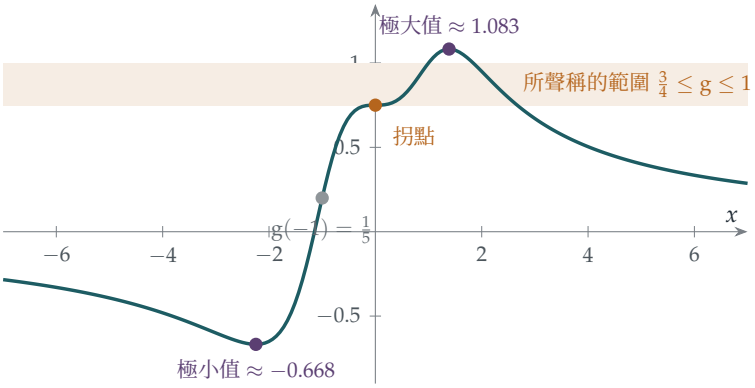


圖 3.1.1 $g(x) = \frac{2x^3+3}{x^4+4}$ 的真實圖像。該論證所聲稱的範圍（琥珀色帶）處處皆錯：$x = -1$ 不是駐點，而且 $g(-1) = \frac{1}{5}$；$x = 0$ 是拐點，不是極小點；函數更遠遠超出色帶之外，極大值高於 1，極小值低於 0。

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 $\int_{-1}^{3}(1-x)^{-3}\,\mathrm{d}x$ **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

$\left(\frac{a_1+a_2}{2}\right)^2 - a_1a_2 = \frac{a_1^2+2a_1a_2+a_2^2-4a_1a_2}{4} = \frac{(a_1-a_2)^2}{4} \geq 0.$ ■

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。 ■

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。

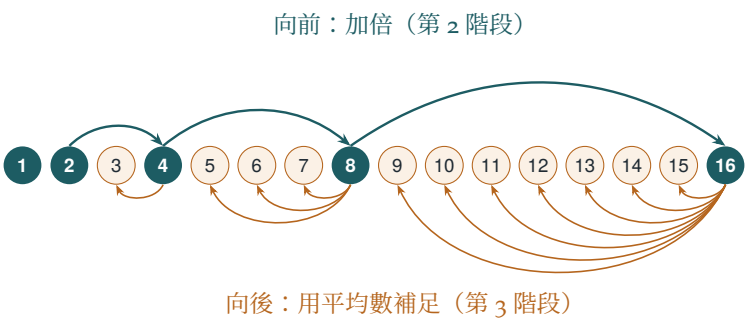


圖 3.2.1　柯西的向前—向後歸納法。藍綠色箭嘴以加倍的方式，證明 AM–GM 對 $2, 4, 8, 16, \ldots$ 個數成立；琥珀色箭嘴則把數列補足至下一個 2 的冪，從而填上其餘每一個 n。

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。 ■

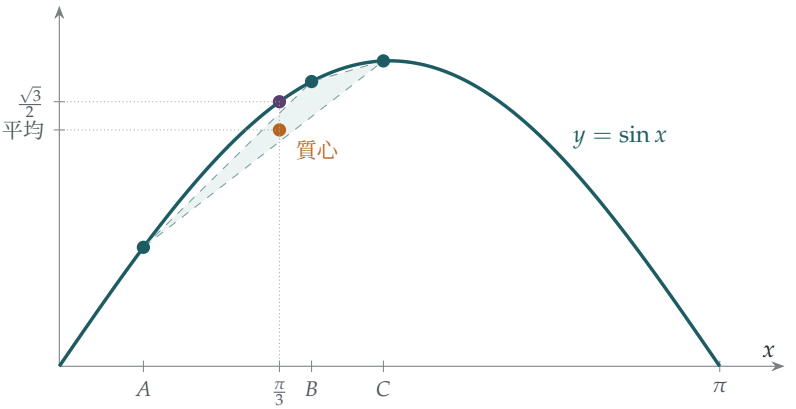


圖 3.3.1　以三角形的三個角 $A = 0.4$、$B = 1.2$、$C = \pi - 1.6$ 說明 $\sin x$ 的琴生不等式。曲線上三點的質心（琥珀色）位於陰影三角形之內，因而在曲線之下。它的高度是 $\sin A$、$\sin B$、$\sin C$ 的平均，而它正上方的曲線高度是 $\sin\frac{\pi}{3}$。

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 $(*)$。

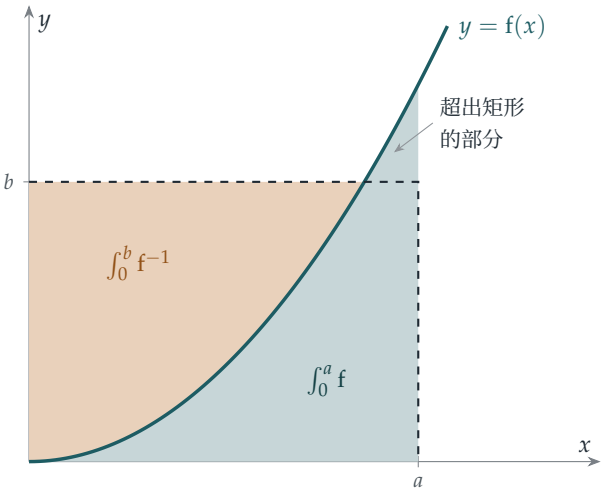


圖 3.4.1　用一幅圖看楊氏不等式。曲線下方由 0 到 a 的面積（藍綠色）與曲線左方由 0 到 b 的面積（琥珀色）合起來蓋住面積為 ab 的虛線矩形，而且還有剩餘。恰好當頂點 (a,b) 位於曲線上時，才沒有多出來的部分。

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 $a_1b_1 + a_2b_2 + a_3b_3 \le \sqrt{a_1^2 + a_2^2 + a_3^2}\sqrt{b_1^2 + b_2^2 + b_3^2}$。（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 $a, b \ge 0$，

$$ab \le e^a - 1 - a + (1+b)\ln(1+b) - b.$$

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。

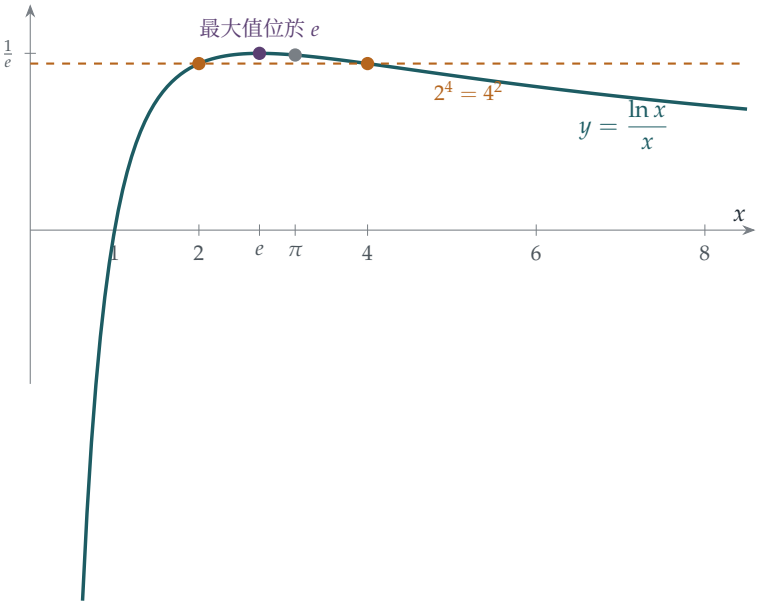


圖 3.5.1　$h(x) = \frac{\ln x}{x}$ 的圖像。它上升至 $x = e$ 處的最大值 $\frac{1}{e}$，然後下降。因為 $h(e) > h(\pi)$，所以 $e^\pi > \pi^e$。高度為 $\frac{\ln 2}{2}$ 的水平線（琥珀色）與圖像相交於 $x = 2$ 和 $x = 4$：這是唯一一對滿足 $n^m = m^n$ 的相異整數。

$x^y = y^x$ 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。

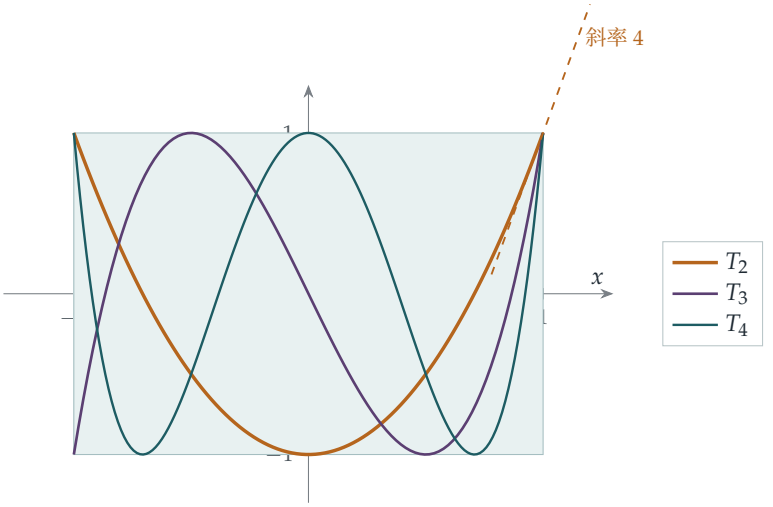


圖 3.6.1　切比雪夫多項式 $T_2 = 2x^2 - 1$、$T_3 = 4x^3 - 3x$ 和 $T_4 = 8x^4 - 8x^2 + 1$ 都留在方框 $|x| \le 1$、$|y| \le 1$ 之內，並且盡可能多次觸及方框的頂邊和底邊。在同次數的多項式中，它們在兩端最陡：$T_n'(1) = n^2$。虛線是 T_2 在 $x = 1$ 處的切線，斜率為 4。

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）。

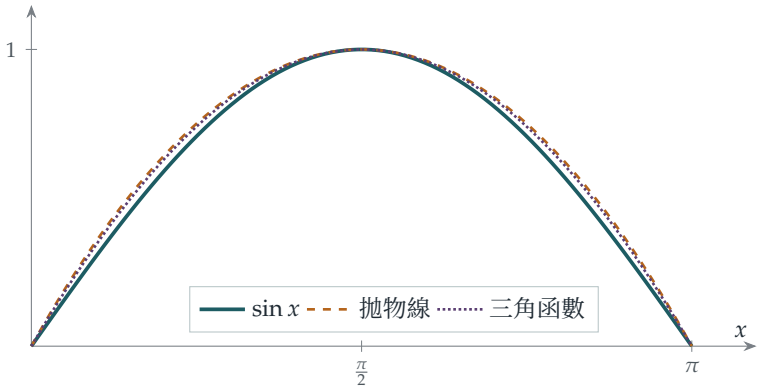


圖 3.7.1　本節的三個函數，各自縮放至最大值為 1：極端情況 $\sin x$（實線），它令 $(*)$ 的等號成立；拋物線 $x(\pi - x)$（虛線），它給出 $\pi^2 \le 10$；以及 $\sin\frac{x}{2} + \cos\frac{x}{2} - 1$（點線），它給出 $\pi \le \frac{22}{7}$。兩個測試函數都非常接近 $\sin x$，所以得出的界才這樣好。

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 $\frac{1.8}{1.81}c \approx 0.9945c$。

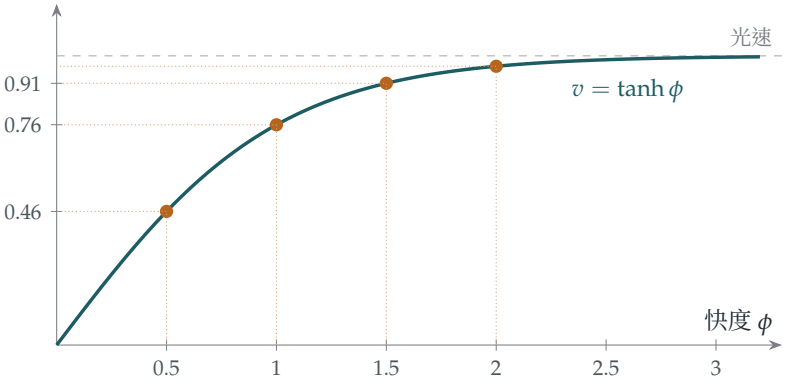


圖 3.8.1　速度（以光速的分數表示）對快度的圖像。快度每次增加相同的份量（引擎每次相同的推進所提供的正是這樣），速度的增幅卻愈來愈小：0.46、0.76、0.91、0.96、…，永遠達不到 1。

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，$\mathrm{g}'(0) = k \neq 0$，且只要等式兩邊都有定義，就有 $\mathrm{g}(x+y) = \dfrac{\mathrm{g}(x)+\mathrm{g}(y)}{1-\mathrm{g}(x)\mathrm{g}(y)}$。證明 $\mathrm{g}'(x) = k\big(1+\mathrm{g}(x)^2\big)$，並由此推出在 $x=0$ 附近 $\mathrm{g}(x) = \tan kx$。

2 **柯西方程。**若 f 可微，且對所有 x、y 都有 $\mathrm{f}(x+y) = \mathrm{f}(x)+\mathrm{f}(y)$，證明 $\mathrm{f}(x) = kx$。

3 **不用導數。**較難只假設 $\mathrm{f}(x+y) = \mathrm{f}(x)\mathrm{f}(y)$ 及 $\mathrm{f}(1) = a > 0$。證明對正整數 n 有 $\mathrm{f}(n) = a^n$，然後證明 $\mathrm{f}(1/n) = a^{1/n}$，再證明對每個正有理數 q 都有 $\mathrm{f}(q) = a^q$。（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 $y = xe^{x}$

$\frac{\mathrm{d}y}{\mathrm{d}x} = e^{x} + xe^{x} = (1+x)e^{x}$，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 $\mathrm{g}(-\frac{1}{e}) = -1$ 開始遞增，經過 $\mathrm{g}(0) = 0$，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

$$\mathrm{g}(z)\, e^{\mathrm{g}(z)} = z, \qquad \mathrm{g}(z) \geq -1.$$

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。

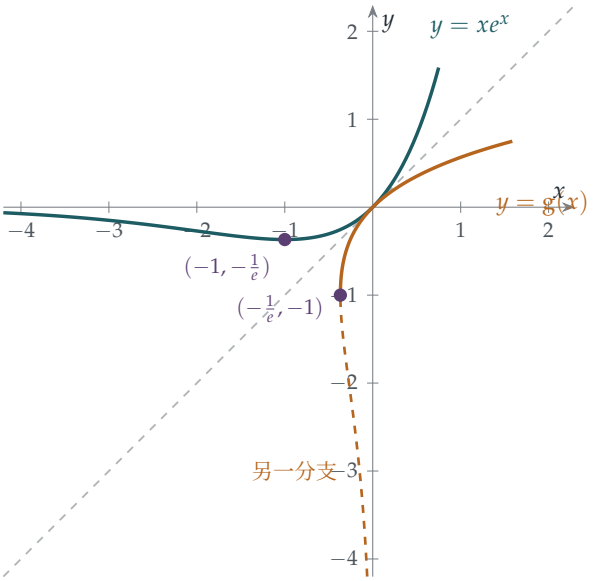


圖 3.9.1 曲線 $y = xe^x$（藍綠色）在 $(-1, -\frac{1}{e})$ 處有極小值。限制在 $x \geq -1$ 上，再關於 $y = x$ 反射，便得到反函數 g（琥珀色實線），即蘭伯特 W 函數。琥珀色虛線是 $y = xe^x$ 左邊部分的反射：這是第二個分支，第 (iii) 部分的第二個根要用到它。

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。

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

全書共 13 章，本冊為第 3 章。 全書各章及所有詳細解答均可免費閱讀、下載及分享，網址為

beyondstep.keithmcnulty.org

文字內容已貢獻予公有領域（CC0 1.0）。STEP 試題 © Cambridge University Press & Assessment。