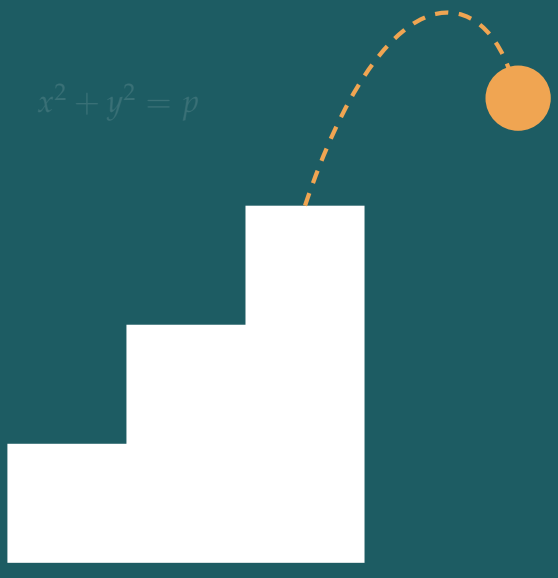
STEP 备试 · 大学数学

Beyond STEP

第 4 章

严谨地学微积分



KEITH McNULTY

Beyond STEP

以历届 STEP 试题为门
走进大学数学

第 4 章

严谨地学微积分

Keith McNulty

Beyond STEP：以历届 STEP 试题为门，走进大学数学

第 4 章：严谨地学微积分

本册是 *Beyond STEP* 全书 13 章中的第 4 章。全书、各章的单行本、每一节的可列印工作纸，以及每一道 STEP 试题的详细解答，均可在 `beyondstep.keithmcnulty.org` 免费阅读及下载。

2026 年 11 月首次于网上发表。

本 PDF 内凡属作者本人创作的内容，包括正文、解答、插图、练习及答案，均按照 Creative Commons CC0 1.0 Universal 授权条款贡献予公有领域：你可以为任何目的复制、改编及分享，毋须征求许可。详见 `creativecommons.org/publicdomain/zero/1.0`。

本书讨论的 STEP 试题，版权属 Cambridge University Press & Assessment 所有。本书并无转载这些试题：每一题均附连结，通往官方来源的题目，即 STEP Support Programme 的试题资料库或 OCR 的历届试卷。

如有错漏，欢迎到 `github.com/keithmcnulty/step_book_mcnulty` 提交 issue 指正。

章

4

严谨的微积分

积分是分析、几何与数论交汇之处，数学中一些最优美的公式亦在此诞生。

本章内容

4.1	甚么是积分?	STEP III 1989
4.2	由一个积分建立三角学	STEP III 2017
4.3	22/7 大于 π	STEP I 1997
4.4	沃利斯、Beta 与二分之一的阶乘	STEP II 1995
4.5	大二学生之梦	STEP III 2009
4.6	没有公式的积分	STEP III 2016
4.7	变分法	STEP III 2014
4.8	切球形面包	STEP I 2001
4.9	Holditch 定理	STEP III 2011
A	练习答案	第 65 页

不只是原函数

那些没有导数的函数，实在是可悲的祸害，我怀着惊恐与厌恶避之则吉。

— 埃尔米特（Charles Hermite），1893 年的一封信

问一位中学生积分（integration）是甚么，你多半会听到“微分的逆运算”。这是个了不起的事实，称为*微积分基本定理*（fundamental theorem of calculus），它令你在中学遇到的大部分积分都变成例行公事。但它是一条关于积分的定理，而不是积分的定义，而且它把积分之所以有趣的大部分原因都遮盖了。

积分比微分（differentiation）早出现约二千年。阿基米德（Archimedes）把图形切成薄片再求和，求出了抛物线弓形的面积和球的表面积，那时还远远未有人想到导数（derivative）。而且积分能到达微分无法到达的地方。每个合理的函数都可以按机械式的法则求导，但大多数函数却*无法*用公式积分出来，这就引出一个自然的问题：甚么时候可以？不可以的时候又该怎么办？积分可以定义新的函数，证明数与数之间的不等式，计算无法用其他方式描述的面积和体积，还能给出无穷和与无穷乘积的精确值。本章谈的就是这一切。

积分究竟是甚么

本章从头说起，由定义开始（Section 4.1）。积分是许多幼长矩形之和的极限（limit）；单凭这个定义、不用原函数（antiderivative）去计算 $\int_0^a x^5\,\mathrm{d}x$，已是相当费工夫的事，而这会令你懂得珍惜基本定理。有了精确的定义，随之而来的是一个精确的问题：哪些函数是可积（integrable）的？这个问题的答案，以及它所引发的种种难题，在二十世纪初催生了勒贝格（Lebesgue）的积分理论，那正是现代概率论的基础。

用来证明的积分

本章一些最出人意表的结果，是用积分去证明一些看来与面积毫无关系的事实。某个正值函数的积分恰好等于 $\frac{22}{7}-\pi$，于是一行就证明了 $\pi<\frac{22}{7}$（Section 4.3）。$\sin t$ 的幂引出沃利斯（Wallis）关于 π 的无穷乘积（infinite product），以及 $\left(\frac{1}{2}\right)!=\frac{\sqrt{\pi}}{2}$ 这个值（Section 4.4）。x^x 的积分化成了和式 $\sum(-1)^{n+1}n^{-n}$（Section 4.5）。而仅仅一个积分 $\int_0^x \frac{\mathrm{d}u}{1+u^2}$，就足以从零开始建立整套三角学（Section 4.2）。

你在大学会遇到甚么

大学一年级的分析（Analysis）课程会从积分的定义出发，严格证明微积分基本定理。其后的课程向多个方向发展。测度论（measure theory）和勒贝格积分（Lebesgue integral）令积分与极限配合得顺畅无碍，并且是概率论的根基。向量微积分（vector calculus）在曲线、曲面和立体上积分，而格林（Green）定理和斯托克斯（Stokes）定理则把它们联系起来（Section 4.9 可让你浅尝一下）。复分析（complex analysis）取道复平面来计算实积分。变分法（calculus of variations）问的是哪个函数令某个积分尽可能小（Section 4.7），并由此通向物理学中的最小作用量原理（principle of least action）。而微分代数（differential algebra）则解释了哪些积分根本有公式可言（Section 4.6）。

贯穿本章的问题

三大问题

1. **积分是甚么？哪些函数有积分？**以和的极限作出的定义，与原函数有何关连？
2. **积分能证明甚么？**一块面积怎能告诉我们有关 π、阶乘和无穷和的事？
3. **没有公式的时候，我们可以做甚么？**我们怎样知道某个积分没有公式？即使如此，我们还能算出甚么？

本章的思想

§	课题	思想
4.1	积分的定义	下和与上和（lower and upper sums）；裂项相消（telescoping）
4.2	由 $\int \frac{\mathrm{d}u}{1+u^2}$ 建立三角学	保持积分不变的代换（substitution）
4.3	$\frac{22}{7} > \pi$	把不等式写成某个正值量的积分
4.4	沃利斯与 $\left(\frac{1}{2}\right)!$	递降公式（reduction formulae）、不变量（invariant）、夹逼（squeeze）
4.5	$\int_0^1 x^x \,\mathrm{d}x$	级数展开与逐项积分（term-by-term integration）
4.6	没有公式的积分	藉多项式代数证明不可能
4.7	变分法	在积分之内配方（completing the square）
4.8	阿基米德的帽盒定理（hat-box theorem）	化简被积函数（integrand），几何意义随之显现
4.9	霍尔迪奇（Holditch）定理	由边界积分求面积；消元（elimination）

如何阅读本章

有好几节涉及繁重但例行的计算：多项式除法、分部积分法（integration by parts）、展开乘积。请先拿起纸笔亲自做一遍，然后才阅读正文。接着找出令计算得以成功的那个思想：一个不变量、一种对称、一次相消。计算是考官要检查的东西；思想才是你会记住并再次运用的东西。

4.1 甚么是积分?

题目

STEP III 1989，第 10 题

继续阅读之前，请先看题目（原题为英文）。题目载于 STEP Support Programme 的试题资料库：

`step.maths.org/questions/89-s3-q10`

引子

你做积分已经很多年了。你知道 $\int_0^a x^5\,\mathrm{d}x = \frac{1}{6}a^6$，因为“积分是微分的逆运算”。但那是一条定理，不是定义。积分究竟是甚么?

老实的答案是:积分是面积,定义为许多细长矩形面积之和的极限(limit)。这道题要你单凭这个定义计算一个积分:不用原函数(antiderivative),不用微积分基本定理(fundamental theorem of calculus)，只靠把矩形加起来再取极限。这比你想像的要难，而这正好说明，令这件事变得轻而易举的微积分基本定理，是何等了不起的成果。

动笔之前

想一想 第 (iii) 部分的两个和，都是阔度为 $\frac{a}{n}$ 的矩形的面积之和。对于递增函数（increasing function）f，哪一个和低估了曲线下的面积，哪一个高估了?

第一个和取每一条长条左端的高度，即 $x = \frac{ra}{n}$，其中 $r = 0, \ldots, n-1$。对递增函数来说，那是长条上的最低点，所以这些矩形全在曲线之下：这是下和（lower sum）。第二个和取右端，即 $r = 1, \ldots, n$，所得的矩形会突出于曲线之上：这是上和（upper sum）（图 4.1.1）。定义说明：当长条愈来愈窄时，若两个和趋向同一个极限，积分便存在。

对于 $\mathrm{f}(x) = x^5$，两个和都涉及 $\sum r^5$。第 (i) 和第 (ii) 部分分别给出这个和的上界和下界。

逐步解题

阶段 1　可以裂项相消的和

> **想一想**　设 $P(r)=r(r+1)(r+2)(r+3)(r+4)(r+5)$，即六个连续整数之积。$P(r)-P(r-1)$ 是甚么？

$P(r)$ 与 $P(r-1)=(r-1)r(r+1)(r+2)(r+3)(r+4)$ 有五个共同因子 $r(r+1)(r+2)(r+3)(r+4)$，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，右方各项裂项相消（telescoping），只余下 $P(n)-P(0)=P(n)$。所以

$$\sum_{r=1}^{n} r(r+1)(r+2)(r+3)(r+4) = \tfrac{1}{6}P(n) = \tfrac{1}{6}n(n+1)(n+2)(n+3)(n+4)(n+5).$$

■

由于对所有 $r\geq 1$ 都有 $r^5 < r(r+1)(r+2)(r+3)(r+4)$，不等式随即成立。 ■

阶段 2　下界

> **停一停，试一试**　把第 1 阶段的做法改用于五个向下数的连续整数之积：$r(r-1)(r-2)(r-3)(r-4)$。求它由 $r=0$ 到 $n-1$ 的和，并与 r^5 比较。

设 $Q(r)=(r-5)(r-4)(r-3)(r-2)(r-1)r$。则 $Q(r+1)-Q(r)=(r-4)(r-3)(r-2)(r-1)r[(r+1)-(r-5)]=6\,(r-4)(r-3)(r-2)(r-1)r$。由 $r=0$ 加到 $n-1$，裂项相消后得 $Q(n)-Q(0)=Q(n)$，所以

$$\sum_{r=0}^{n-1}(r-4)(r-3)(r-2)(r-1)r = \tfrac{1}{6}(n-5)(n-4)(n-3)(n-2)(n-1)n.$$

现在逐项比较。当 $r=0$ 时，两边都是 0。当 $1\leq r\leq 4$ 时，乘积含有一个零因子，而 $r^5>0$。当 $r\geq 5$ 时，每个因子都是正数，且不大于 r，其中有些更严格小于它，所以乘积小于 r^5。由于 $n>1$，和式包含 $r=1$ 一项，该处不等式严格成立，所以

$$\sum_{r=0}^{n-1} r^5 > \tfrac{1}{6}(n-5)(n-4)(n-3)(n-2)(n-1)n. \qquad \blacksquare$$

阶段 3　积分

想一想 把 $\mathrm{f}(x)=x^5$ 的下和与上和用 $\sum r^5$ 表示。分母中出现 n 的几次方?

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,并因为 $L_n<U_n$,

$$\frac{a^6}{6}\cdot\frac{(n-5)(n-4)\cdots n}{n^6}\;<\;L_n\;<\;U_n\;<\;\frac{a^6}{6}\cdot\frac{n(n+1)\cdots(n+5)}{n^6}.$$

两个分数各自是六个因子之积,这些因子形如 $\frac{n-3}{n}=1-\frac{3}{n}$ 或 $\frac{n+3}{n}=1+\frac{3}{n}$,每一个都趋向 1。所以外侧两个界都趋向 $\frac{a^6}{6}$,根据夹逼定理(squeeze),L_n 和 U_n 亦然。两个极限都存在而且相等,所以按定义

$$\int_0^a x^5\,\mathrm{d}x=\tfrac{1}{6}a^6.\qquad\blacksquare$$

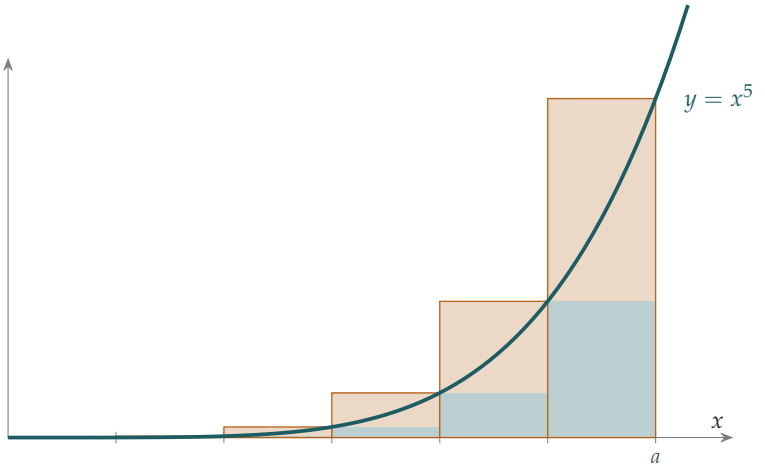


图 4.1.1 x^5 在 $[0,a]$ 上的下和与上和,分成 $n=6$ 条长条。下和是蓝绿色的面积(曲线下的矩形);上和则再加上琥珀色的部分。两者之差正好是最后、最高的那个矩形的面积 $\frac{a}{n}\,\mathrm{f}(a)$,当 n 增大时它趋向 0。

回顾

- **又是裂项相消。**连续整数之积的和一下子收拢起来,与第 2 章的和与积如出一辙。
- **先定界,再夹逼。**我们从没有求出 $\sum r^5$ 的准确值(它是 $\frac{1}{12}n^2(n+1)^2(2n^2+2n-1)$);有两边的界已经足够。
- **定义担起全部工作。**原函数由始至终没有出现过。一切都来自积分作为和的极限这个定义。

背后的数学

黎曼的定义

把细长条加起来求面积的想法，可追溯至阿基米德（Archimedes）：他在公元前 250 年左右，正是用这个论证的一个版本求出抛物线下的面积。牛顿（Newton）和莱布尼兹（Leibniz）把积分变成微分的逆运算，其后一个半世纪，大多数人都是这样理解积分的。1854 年，黎曼（Bernhard Riemann）提出了大学课程至今仍用作起点的定义：把 $[a,b]$ 分成许多小段，构作（阔度）$\times$（f 在该小段内某处的高度）之和；若小段不断缩小时，所有这样的和都趋向同一个极限，便说 f 可积（integrable）。本题的版本采用等阔长条和左右端点，是一个适用于递增函数的特殊情况；达布（Jean-Gaston Darboux）的“上和与下和”则把一般情况处理得干净利落。

并非每个函数都有积分

有了精确的定义，便可以问哪些函数可积。闭区间上的每个连续函数都可积，每个递增函数也可积。但试考虑狄利克雷（Dirichlet）函数：若 x 是有理数，$\mathrm{f}(x)=1$；若是无理数，则为 0。每一条长条无论多窄，都同时包含有理数和无理数，所以每个上和都是 1，每个下和都是 0。两者永不一致，因此 $\int_0^1 \mathrm{f}$ 在黎曼的意义下并不存在。1902 年，勒贝格（Henri Lebesgue）发明了一种更强大的积分，能够处理这个函数（它的勒贝格积分（Lebesgue integral）是 0），而且在取极限时表现好得多。勒贝格积分是现代概率论的基础，你会在大学二年级或三年级遇上它。

离散微积分

第 1 阶段的裂项相消，其实是乔装了的微积分。“上升阶乘”（rising factorial）$r(r+1)\cdots(r+k-1)$ 在求和之下的表现，与 x^k 在积分之下的表现完全一样：求和令因子的数目增加一个，并除以新的因子数目，正如 $\int x^k\,\mathrm{d}x=\frac{x^{k+1}}{k+1}$。和与积分、差分与导数之间的这种对应，称为*有限差分演算*（calculus of finite differences），是计算和式的有力工具。

思维技巧

从定义出发

题目要求你“用定义”证明某事时，就只用定义：不用任何依赖于待证结果的定理，也不用熟悉的捷径。

这样做出奇地难，却让你确切看到那些捷径的价值。在大学，每一科的头几个星期都是这样进行的：先有定义，然后才有定理，让你不必再从定义出发。

延伸练习

1 **把极限化为积分。**试做 STEP III 1999，第 3 题（`step.maths.org/questions/99-s3-q3`）。这道题把和的极限化为积分：在每个和式中，找出曲线下的矩形。

2 **阿基米德的抛物线。**利用 $\sum_{r=1}^{n} r^2 = \frac{1}{6}n(n+1)(2n+1)$ 和第 (iii) 部分的定义，证明 $\int_0^a x^2\,\mathrm{d}x = \frac{1}{3}a^3$。

3 **上升阶乘。**证明对每个正整数 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 **准确的和。**利用恒等式 $r(r+1)(r+2)(r+3)(r+4) = r^5 + 10r^4 + 35r^3 + 50r^2 + 24r$，以及 $\sum r^k$ 在 $k \le 4$ 时的已知公式，求出 $\sum_{r=1}^{n} r^5$ 的准确值。

延伸阅读

- **David Bressoud, *A Radical Approach to Lebesgue's Theory of Integration***（Cambridge University Press）。由黎曼积分走到勒贝格积分的故事，以逼出每一步发展的问题为线索娓娓道来。
- **Michael Spivak, *Calculus***（Cambridge University Press）。第 13 章以上和与下和定义积分，与本节完全一样。
- **Graham, Knuth & Patashnik, *Concrete Mathematics***（Addison-Wesley）。第 2 章建立有限差分演算，并讨论上升阶乘和下降阶乘。

4.2 由一个积分建立三角学

题目

STEP III 2017，第 6 题

继续阅读之前，请先看题目（原题为英文）。题目载于 STEP Support Programme 的试题资料库：

`step.maths.org/questions/17-s3-q6`

引子

π 究竟是甚么？在中学，它是圆周与直径之比，而 sin 和 cos 则用三角形来定义。这对几何来说已经足够，却留下了一些棘手的问题。沿着曲线的“长度”到底是甚么？为甚么 $\sin x$ 的麦克劳林级数（Maclaurin series）偏偏是那个样子？我们又怎样知道加法公式（addition formulas）对所有角都成立，而不只限于画得出来的角？

这道题展示了一条截然不同的入门途径。忘掉三角形和圆吧。只从一个积分 $\mathrm{T}(x) = \int_0^x \frac{\mathrm{d}u}{1+u^2}$ 出发，它所需要的不过是 Section 4.1 中积分的定义。把 π 定义为 $2\mathrm{T}_\infty$。然后只用代换法（substitution），证明 $\mathrm{T}(\sqrt{3}) = \frac{\pi}{3}$ 和 $\mathrm{T}(\sqrt{2}-1) = \frac{\pi}{8}$，也就是你所熟悉的 $\arctan\sqrt{3} = 60^\circ$ 和 $\tan 22\frac{1}{2}^\circ = \sqrt{2}-1$。三角学就这样从无到有地重建起来。

动笔之前

你心里其实知道 $\mathrm{T}(x) = \arctan x$ 和 $\mathrm{T}_\infty = \frac{\pi}{2}$。题目不准你使用这些事实，而这正是用意所在：一切都必须由积分推导出来。不过，这点心照不宣的知识是很有用的向导。

? **想一想** 第 (i) 和第 (ii) 部分分别对应哪些三角恒等式（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) 部分则是乔装了的加法公式 $\tan(A+B) = \frac{\tan A+\tan B}{1-\tan A\tan B}$：设 $u = \tan A$ 和 $a = \tan B$，则 $v = \tan(A+B)$。这道题不靠任何几何，单凭积分就证明了这些恒等式。

逐步解题

阶段 1　反射 $u \mapsto 1/u$

在 $\mathrm{T}(x)$ 中代入 $u = 1/w$。于是 $\mathrm{d}u = -\mathrm{d}w/w^2$，而 $\dfrac{1}{1+u^2} = \dfrac{w^2}{w^2+1}$，所以

$$\mathrm{T}(x) = \int_{\infty}^{1/x} \frac{w^2}{w^2+1}\cdot\left(-\frac{1}{w^2}\right)\mathrm{d}w = \int_{1/x}^{\infty} \frac{\mathrm{d}w}{1+w^2} = \mathrm{T}_\infty - \mathrm{T}(x^{-1}). \qquad \blacksquare$$

换句话说，$\frac{1}{1+u^2}$ 之下由 0 到 x 的面积，等于由 $\frac{1}{x}$ 到 ∞ 的面积（图 4.2.1）。

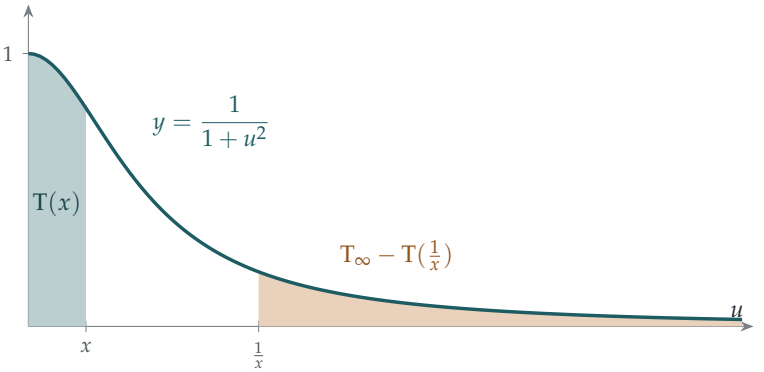


图 4.2.1　代换 $u = 1/w$ 显示 $\frac{1}{1+u^2}$ 之下的两块阴影面积相等：由 0 到 x 的面积（蓝绿色），以及由 $\frac{1}{x}$ 起无限伸延的面积（琥珀色）。图中 $x = \frac{1}{2}$。

阶段 2　加法代换

> **想一想**　用商法则（quotient rule）对 $v = \dfrac{u+a}{1-au}$ 求导。然后把 $1+v^2$ 化成单一分式，再作比较。

$$\frac{\mathrm{d}v}{\mathrm{d}u} = \frac{(1-au)+a(u+a)}{(1-au)^2} = \frac{1+a^2}{(1-au)^2},$$
$$1+v^2 = \frac{(1-au)^2+(u+a)^2}{(1-au)^2} = \frac{(1+a^2)(1+u^2)}{(1-au)^2}.$$

两式相除，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**　这表示 $\dfrac{\mathrm{d}u}{1+u^2} = \dfrac{\mathrm{d}v}{1+v^2}$。所以代换 $u \mapsto v$ 完全没有改变被积函数（integrand）。那么它对 $\mathrm{T}(x)$ 的积分上下限有甚么影响？

当 u 由 0 变到 x 时（其中 $x<\frac{1}{a}$，故 $1-au>0$ 始终成立，而 v 连续且递增），v 由 a 变到 $\frac{x+a}{1-ax}$。所以

$$\mathrm{T}(x)=\int_0^x \frac{\mathrm{d}u}{1+u^2}=\int_a^{(x+a)/(1-ax)} \frac{\mathrm{d}v}{1+v^2}=\mathrm{T}\left(\frac{x+a}{1-ax}\right)-\mathrm{T}(a). \qquad \blacksquare$$

阶段 3 再多两个恒等式

结合第 (i) 部分，并使用两次：

$$\mathrm{T}(x^{-1})=\mathrm{T}_\infty-\mathrm{T}(x)=\mathrm{T}_\infty-\mathrm{T}\left(\frac{x+a}{1-ax}\right)+\mathrm{T}(a)=2\mathrm{T}_\infty-\mathrm{T}\left(\frac{x+a}{1-ax}\right)-\mathrm{T}(a^{-1}). \qquad \blacksquare$$

现在更换变量名称：设 $y=1/x$ 和 $b=1/a$。条件 $x<1/a$ 变成 $1/y<b$，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 $\mathrm{T}(y)=2\mathrm{T}_\infty-\mathrm{T}\left(\frac{y+b}{by-1}\right)-\mathrm{T}(b)$。 $\blacksquare$

阶段 4 特殊值

> **停一停，试一试** 在最后一个恒等式中选取 y 和 b，使自变量 $\frac{y+b}{by-1}$ 刚好等于 y 本身。你会得到甚么？

取 $y=b=\sqrt{3}$：这时 $y>\frac{1}{b}$，而 $\frac{\sqrt{3}+\sqrt{3}}{3-1}=\sqrt{3}$。恒等式变成 $\mathrm{T}(\sqrt{3})=2\mathrm{T}_\infty-2\mathrm{T}(\sqrt{3})$，所以

$$\mathrm{T}(\sqrt{3})=\tfrac{2}{3}\mathrm{T}_\infty. \qquad \blacksquare$$

> **想一想** 至于 $\sqrt{2}-1$，试在加法恒等式中取 $x=a=\sqrt{2}-1$。$\frac{2x}{1-x^2}$ 是多少？$\mathrm{T}(1)$ 又是多少？

取 $x=a=\sqrt{2}-1$（由于 $x^2=3-2\sqrt{2}<1$，故 $x<\frac{1}{a}$），加法恒等式给出 $2\mathrm{T}(x)=\mathrm{T}\left(\frac{2x}{1-x^2}\right)$。而 $1-x^2=2\sqrt{2}-2=2(\sqrt{2}-1)=2x$，所以 $\frac{2x}{1-x^2}=1$，于是 $2\mathrm{T}(\sqrt{2}-1)=\mathrm{T}(1)$。最后，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

$$\mathrm{T}(\sqrt{2}-1)=\tfrac{1}{4}\mathrm{T}_\infty. \qquad \blacksquare$$

回顾

- **保持被积函数不变的代换。**两个关键步骤所用的代换，都令 $\frac{\mathrm{d}u}{1+u^2}$ 保持不变，只有积分上下限移动了。积分的这类“对称性”（symmetry）威力强大。

- **不动点（fixed point）。**$\mathrm{T}(\sqrt{3})$ 的值，来自选取适当的数，使某个表达式等于它本身，于是恒等式变成了方程。
- **定义可以来自分析。**我们从没有用过三角形，却重新得到了 $\frac{\pi}{3}$ 和 $\frac{\pi}{8}$（其中 π 定义为 $2\mathrm{T}_\infty$）。

背后的数学

建构初等函数

在严格的分析课程中，函数 exp、sin、cos 和 ln 不能随便假设存在；它们必须被*建构*出来，其性质也必须加以证明。标准的做法有三种。你可以用幂级数（power series）来定义它们（$\sin x = x - \frac{x^3}{3!} + \cdots$）。你可以把它们定义为微分方程（differential equation）的解，就如练习中那道 1986 年样本试卷的题目（$\mathrm{f}' = \mathrm{f}$、$\mathrm{f}(0) = 1$ 定义了 exp）。你也可以像这里一样，用积分来定义反函数（inverse function）：$\ln x = \int_1^x \frac{\mathrm{d}u}{u}$ 和 $\arctan x = \int_0^x \frac{\mathrm{d}u}{1+u^2}$，然后定义 $\pi = 2\int_0^\infty \frac{\mathrm{d}u}{1+u^2}$。每种做法各有所长，而证明它们全都给出相同的函数，是一道令人满足的练习。试与 Section 3.8 的函数方程（functional equation）做法比较。

圆以外：椭圆积分

积分的做法可以推广，而三角形的做法却做不到。1718 年，Giulio Fagnano 研究*双纽线*（lemniscate）这条 8 字形曲线的弧长（arc length），由此引出 $\int_0^x \frac{\mathrm{d}u}{\sqrt{1-u^4}}$，并找到一个能把弧长加倍的代换，正如我们的代换把 T 加倍一样。欧拉（Euler）在 1751 年读到 Fagnano 的著作，找出了完整的加法公式；这最终经由阿贝尔（Abel）和雅可比（Jacobi），发展成*椭圆函数*（elliptic functions）的理论：椭圆函数就是这类积分的反函数，它们之于双纽线，正如 sin 和 cos 之于圆。我们在高斯（Gauss）关于算术几何平均（AGM）的工作中（Section 2.6）遇过同一个积分。椭圆函数如今是数论的核心；Section 1.9 的椭圆曲线（elliptic curves）就是它们在几何上的投影。

计算 π

加法恒等式也是计算 π 的那些最著名公式的基础。举例说，反复运用它可以证明 $\mathrm{T}(1) = 4\mathrm{T}(\frac{1}{5}) - \mathrm{T}(\frac{1}{239})$，这就是 1706 年的梅钦公式（Machin's formula）；配合级数 $\mathrm{T}(x) = x - \frac{x^3}{3} + \frac{x^5}{5} - \cdots$，它算出了 π 的 100 位数字，并在其后两个半世纪一直是首选的方法。这是 Section 5.1 的主题。

思维技巧

积分的对称性

寻找令被积函数保持不变的代换。它们会化为该积分所定义的函数的恒等式。

这里 $u \mapsto 1/u$ 和 $u \mapsto \frac{u+a}{1-au}$ 都保持 $\frac{\mathrm{d}u}{1+u^2}$ 不变，并分别给出反射公式和加法公式。同一个想法用于其他积分，正是特殊函数（special functions）理论的基础。

延伸练习

1 **指数函数与余弦。**试做 STEP II Specimen paper，第 6 题（`step.maths.org/questions/spec-s2-q6`）。这道题只凭指数函数和余弦各自的微分方程，推导出它们的加法公式。（提示：对于余弦，找出两个理应为零的组合，并证明它们的平方和是常数。）

2 **反正弦。**试做 STEP III 2007，第 7 题（`step.maths.org/questions/07-s3-q7`）。这道题用处理本节积分的同一手法来处理反正弦（inverse sine），并以一个代换把两者联系起来。

3 **一条早期的梅钦型公式。**利用加法恒等式证明 $\mathrm{T}(\frac{1}{2}) + \mathrm{T}(\frac{1}{3}) = \mathrm{T}(1) = \frac{1}{2}\mathrm{T}_\infty$。

4 T_∞ **是有限的。**证明对所有 $x \geq 1$ 都有 $\mathrm{T}(x) \leq 1 + \int_1^x \frac{\mathrm{d}u}{u^2} < 2$，解释为甚么当 $0 < x < 1$ 时同样有 $\mathrm{T}(x) < 2$，并由单调收敛定理（monotone convergence theorem）（Section 2.2）推断 T_∞ 存在。

延伸阅读

- **Tom Körner,** ***Calculus for the Ambitious***（Cambridge University Press）。严格地建构初等函数，并讨论各种不同的做法。
- **Walter Rudin,** ***Principles of Mathematical Analysis***（McGraw-Hill）。即著名的“baby Rudin”；第 8 章的引言只用寥寥数页，便优雅地由幂级数建构出 exp、sin 和 cos。程度较深。
- **Victor Prasolov & Yuri Solovyev,** ***Elliptic Functions and Elliptic Integrals***（American Mathematical Society）。想知道双纽线通往何处，可读此书。程度较深。

4.3 22/7 大于 π

题目

STEP I 1997，第 6 题

继续阅读之前，请先看题目（原题为英文）。题目载于 STEP Support Programme 的试题资料库：

`step.maths.org/questions/97-s1-q6`

引子

一代又一代的学生都听过 π “大约是 $\frac{22}{7}$” 这个说法。这是一个非常好的近似值（approximation）：$\frac{22}{7} = 3.142857\ldots$，而 $\pi = 3.141592\ldots$。但 $\frac{22}{7}$ 究竟是偏大还是偏小？你又能否证明这一点，而不是单靠查阅 π 的各位数字？那些数字毕竟是要别人先计算出来的。

这道题给出了整个数学之中最优雅的证明之一。只需一个正函数的积分（integral），其值因而必定为正，而它竟然恰好等于 $\frac{22}{7} - \pi$。被积函数（integrand）为正，积分便为正，所以 $\frac{22}{7} > \pi$。整个证明只需寥寥数行，而且精巧得被人当作谜题印在数学 T 恤上。

动笔之前

想一想 为甚么 $\int_0^1 \frac{\text{多项式}}{1+x^2}\,\mathrm{d}x$ 会算出 π 来？这种形式的积分之中，哪一个简单的积分本身就给出 π？

$\int_0^1 \frac{\mathrm{d}x}{1+x^2} = \arctan 1 = \frac{\pi}{4}$。把一个多项式（polynomial）除以 $1+x^2$，会得到一个多项式商式（quotient）（其积分是一个有理数），再加上一个常数余式（remainder）除以 $1+x^2$（其积分是 π 的有理数倍）。题目的第一部分正是要你做这个除法。

逐步解题

阶段 1 多项式除法

先展开：$(1-x)^4 = 1 - 4x + 6x^2 - 4x^3 + x^4$，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，可以用长除法（long division），也可以用比较系数（comparing coefficients）的方法。余式是甚么?

在 $(a_6x^6 + \cdots + a_0)(x^2+1) + b$ 中由最高次项开始比较系数：$a_6 = 1$；$a_5 = -4$；$a_4 + a_6 = 6$，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 $a_0 = 4$；最后 $a_1 = 0$（x 的系数）及 $a_0 + b = 0$，所以 $b = -4$。即是说，

$$x^4(1-x)^4 = (x^6 - 4x^5 + 5x^4 - 4x^2 + 4)(x^2+1) - 4.$$

阶段 2 计算积分

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

$$\int_0^1 \frac{x^4(1-x)^4}{1+x^2}\,\mathrm{d}x = \frac{1}{7} - \frac{4}{6} + \frac{5}{5} - \frac{4}{3} + 4 - 4\cdot\frac{\pi}{4} = \frac{1}{7} - \frac{2}{3} + 1 - \frac{4}{3} + 4 - \pi.$$

由于 $-\frac{2}{3} - \frac{4}{3} = -2$，有理数部分是 $\frac{1}{7} + 3 = \frac{22}{7}$，因此积分等于 $\frac{22}{7} - \pi$。 ■

阶段 3 估计积分的界

想一想 被积函数在 $(0,1)$ 上为正，所以立即得到 $\frac{22}{7} > \pi$。要得到 π 的下界（lower bound），我们需要这个积分的上界（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

$$\tfrac{1}{2}x^4(1-x)^4 \le \frac{x^4(1-x)^4}{1+x^2} \le x^4(1-x)^4,$$

而且除了在端点之外，不等式（inequality）都是严格成立的。

停一停，试一试 计算 $\int_0^1 x^4(1-x)^4\,\mathrm{d}x$，然后完成证明。

把展开式逐项积分，

$$\int_0^1 x^4(1-x)^4\,\mathrm{d}x = \frac{1}{9} - \frac{4}{8} + \frac{6}{7} - \frac{4}{6} + \frac{1}{5} = \frac{1}{630}.$$

（以 2520 为公分母：$280 - 1260 + 2160 - 1680 + 504 = 4$，而 $\frac{4}{2520} = \frac{1}{630}$。）所以

$$0 < \frac{22}{7} - \pi < \frac{1}{630},$$

亦即 $\frac{22}{7} > \pi > \frac{22}{7} - \frac{1}{630}$。 ■

以数值计算，$\frac{22}{7} - \frac{1}{630} = 3.14127\ldots$，所以我们已证明 $3.1412 < \pi < 3.1429$。若同时用上左边的不等式，便有 $\frac{1}{1260} < \frac{22}{7} - \pi$，这把上界改进为 $\pi < 3.14207$。

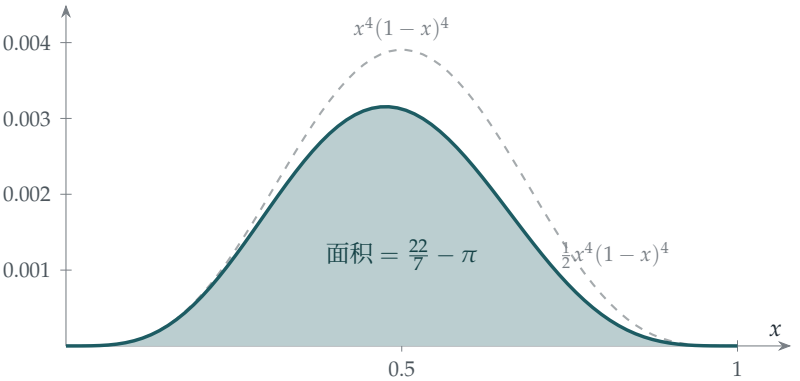


图 4.3.1 被积函数 $\frac{x^4(1-x)^4}{1+x^2}$（蓝绿色）为正，但数值极小：它的最大值约为 0.003。阴影部分的面积恰好是 $\frac{22}{7} - \pi \approx 0.00126$。被积函数介乎 $\frac{1}{2}x^4(1-x)^4$（点线）与 $x^4(1-x)^4$（虚线）之间，这两条曲线下的面积分别是 $\frac{1}{1260}$ 和 $\frac{1}{630}$。

回顾

- **以正值证明不等式。**正函数的积分为正。一旦认出积分等于 $\frac{22}{7} - \pi$，不等式便立即成立。
- **被积函数愈小，界愈精确。**因子 $x^4(1-x)^4$ 在 $[0,1]$ 上极小（最多只有 $\frac{1}{256}$），这使积分很小，所得的界也就很精确。
- **有理数加上 π 的倍数。**正如 Section 2.9 所见，一个形如“有理数 + 某常数的有理数倍”的积分，若其值很小，便给出该常数的一个良好有理近似值。

背后的数学

一个有来历的证明

阿基米德（Archimedes）约于公元前 250 年证明了 $\frac{223}{71} < \pi < \frac{22}{7}$，方法是计算圆内接和圆外切正 96 边形的周界（可与 Section 2.5 中的多边形比较）。积分证明则晚近得多。它由 D. P. Dalzell 于 1944 年发表，其后在 1968 年被选为美国大学本科生普特南数学竞赛（Putnam competition）的题目，因而广为人知。它的迷人之处在于完全不需要任何数值计算：不等式 $\frac{22}{7} > \pi$ 直接来自“正函数的积分为正”这个事实。

用同一思路得出更佳的近似值

这个方法可以推广。以 $x^8(1-x)^8$ 取代 $x^4(1-x)^4$，同样的除法给出

$$\int_0^1 \frac{x^8(1-x)^8}{1+x^2}\,\mathrm{d}x = 4\pi - \frac{188684}{15015},$$

由此证明 $\pi > \frac{47171}{15015} = 3.1415917\ldots$，准确至六位有效数字（significant figures）。2009 年，Stephen Lucas 找到了多族这类被积函数非负的积分，它们能证明关于其他著名近似值的不等式，其中包括准确至七位有效数字的 $\frac{355}{113}$。

近似值与无理性

为甚么 π 会有 $\frac{22}{7}$ 和 $\frac{355}{113}$ 这样好的分数近似值？每个无理数（irrational number）都有无穷多个分数 $\frac{p}{q}$ 与它相距不超过 $\frac{1}{q^2}$，而 π 碰巧在很早的阶段就有几个特别好的，因为它的连分数（continued fraction）$[3;7,15,1,292,\ldots]$ 含有 292 这个很大的项（见 Section 2.2）。像本题这样的积分，若令其愈来愈小，便是另一类证明的起点：这类证明指出某些数不可能被分数逼近得太好，因而是无理数。Ivan Niven 那个著名的、只有一页的 π 是无理数的证明（1947 年），用的就是 $\frac{x^n(1-x)^n}{n!}\sin(\pi x)$ 的积分，它们是 Section 2.9 中那些积分的三角函数近亲。

思维技巧

把不等式写成某个正值函数的积分

要证明 $A > B$，可找一个积分等于 $A - B$ 的正函数。

这样便把一个不等式化为一个恒等式（identity）再加上一个显然的事实。这是分析学（analysis）中最有用的技巧之一；同一思路，以平方和取代积分，就在

Section 3.2 中证明了算术平均—几何平均不等式（AM–GM）。

延伸练习

1 **核对数字。**计算 $\frac{22}{7} - \pi$ 至四位有效数字，并确认它介乎 $\frac{1}{1260}$ 与 $\frac{1}{630}$ 之间。

2 **一个 Beta 积分。**利用公式 $\int_0^1 x^n(1-x)^n\,\mathrm{d}x = \frac{(n!)^2}{(2n+1)!}$（在 Section 4.4 中证明）验证 $\int_0^1 x^4(1-x)^4\,\mathrm{d}x = \frac{1}{630}$。

3 **一个偏小的近似值。**证明

$$\int_0^1 \frac{x(1-x)^2}{1+x^2}\,\mathrm{d}x = \frac{\pi}{2} - \frac{3}{2} \quad \text{及} \quad \int_0^1 \frac{x^2(1-x)^4}{1+x^2}\,\mathrm{d}x = \pi - \frac{47}{15},$$

并由此推出 $\pi > 3$ 及 $\pi > \frac{47}{15}$。为甚么这些近似值不及 $\frac{22}{7}$ 准确?

4 **更进一步。**较难验证上文给出的 $\int_0^1 \frac{x^8(1-x)^8}{1+x^2}\,\mathrm{d}x$ 的恒等式。（除法可以交给电脑代数系统去做；重点是看出余式是 16，从而得出 4π。）

延伸阅读

- **Petr Beckmann, *A History of* π**（St Martin's Press）。详细讲述阿基米德的多边形方法，另有许多其他内容。
- **Stephen Lucas, "Approximations to π derived from integrals with nonnegative integrands"**, *American Mathematical Monthly* 116 (2009)。篇幅短小，容易阅读。
- **Ivan Niven, *Irrational Numbers***（Mathematical Association of America）。收录了 Niven 对 π 是无理数的证明。

4.4 沃利斯、Beta 与二分之一的阶乘

题目

STEP II 1995，第 4 题

继续阅读之前，请先看题目（原题为英文）。题目载于 STEP Support Programme 的试题资料库：

`step.maths.org/questions/95-s2-q4`

引子

$\left(\frac{1}{2}\right)!$ 是甚么？这个问题听来毫无道理：$n!$ 的意思是 $1 \times 2 \times \cdots \times n$，而你不可能把“半个数那么多”的项相乘。然而它确有一个自然的答案，就是

$$\left(\frac{1}{2}\right)! = \frac{\sqrt{\pi}}{2}.$$

π 从何而来？来自 $\sin t$ 的幂的积分（integral），这正是本题研究的对象。过程中，本题证明了关于 π 的其中一条最古老、最奇特的公式，即沃利斯（John Wallis）在 1655 年提出的无穷乘积（infinite product）

$$\frac{\pi}{2} = \frac{2}{1}\cdot\frac{2}{3}\cdot\frac{4}{3}\cdot\frac{4}{5}\cdot\frac{6}{5}\cdot\frac{6}{7}\cdots,$$

并且找出 $n!$ 的斯特灵公式（Stirling's formula）中隐藏的常数，这条公式是每一位概率学家和统计学家都会用到的。

动笔之前

想一想 u_0 和 u_1 是甚么？当 n 很大时，$\sin^n t$ 在 $[0, \frac{\pi}{2}]$ 上的图像大致是甚么样子？

$u_0 = \int_0^{\pi/2} 1\,\mathrm{d}t = \frac{\pi}{2}$，而 $u_1 = \int_0^{\pi/2} \sin t\,\mathrm{d}t = 1$。当 n 很大时，$\sin^n t$ 极小，只有在 $t = \frac{\pi}{2}$ 附近例外，因为那里 $\sin t$ 接近 1：图像变成紧贴右边界的一个狭窄尖峰（图 4.4.1）。所以 u_n 随 n 增大而缩小。本题准确地找出它缩小得有多快：与 $\sqrt{\pi/(2n)}$ 相若。

逐步解题

阶段 1　一条递推公式

想一想　用分部积分法（integration by parts）计算 $\int \sin t \cdot \sin^{n-1} t \, \mathrm{d}t$：对 $\sin t$ 积分，对 $\sin^{n-1} t$ 求导。新的被积函数（integrand）含有 $\cos^2 t$；怎样才能变回 $\sin t$ 的幂?

取 $\mathrm{d}v = \sin t \, \mathrm{d}t$（故 $v = -\cos t$）及 $w = \sin^{n-1} t$（故 $w' = (n-1)\sin^{n-2} t \cos t$），

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

当 $n \geq 2$ 时，边界项为零。写成 $\cos^2 t = 1 - \sin^2 t$，该积分便是 $u_{n-2} - u_n$，所以 $u_n = (n-1)(u_{n-2} - u_n)$，即

$$n u_n = (n-1) u_{n-2}. \qquad \blacksquare$$

两边同乘以 u_{n-1}，得 $n u_n u_{n-1} = (n-1) u_{n-1} u_{n-2}$。 ■

阶段 2　一个不变量

想一想　这个关系式说明，当 n 减一时，$n u_n u_{n-1}$ 这个量保持不变。它在 $n = 1$ 时等于甚么?

反复运用，$n u_n u_{n-1} = (n-1) u_{n-1} u_{n-2} = \cdots = 1 \cdot u_1 u_0 = 1 \cdot 1 \cdot \frac{\pi}{2}$。所以对所有 $n \geq 1$，$n u_n u_{n-1} = \frac{1}{2}\pi$。 ■

阶段 3　夹逼

对于 $0 < t < \frac{\pi}{2}$，我们有 $0 < \sin t < 1$，所以 $\sin^n t < \sin^{n-1} t$：除端点外，$\sin^n t$ 的图像位于 $\sin^{n-1} t$ 的图像之下。积分后得 $0 < u_n < u_{n-1}$。

停一停，试一试　利用 $u_n < u_{n-1}$ 和 $n u_n u_{n-1} = \frac{1}{2}\pi$ 把 $n u_n^2$ 夹住。

把 u_{n-1} 换成较小的 u_n，得 $n u_n^2 < n u_n u_{n-1} = \frac{1}{2}\pi$。把 u_n 换成较大的 u_{n-1}，得 $\frac{1}{2}\pi = n u_n u_{n-1} < n u_{n-1}^2$。所以 $n u_n^2 < \frac{1}{2}\pi < n u_{n-1}^2$。 ■

在右边的不等式中以 $n+1$ 代替 n：$\frac{1}{2}\pi < (n+1) u_n^2$，所以 $n u_n^2 > \frac{n}{n+1} \cdot \frac{1}{2}\pi$。合起来，

$$\left(\frac{n}{n+1}\right)\frac{1}{2}\pi < n u_n^2 < \frac{1}{2}\pi. \qquad \blacksquare$$

由于 $\frac{n}{n+1} \to 1$，由夹逼可得 $n u_n^2 \to \frac{1}{2}\pi$。 ■

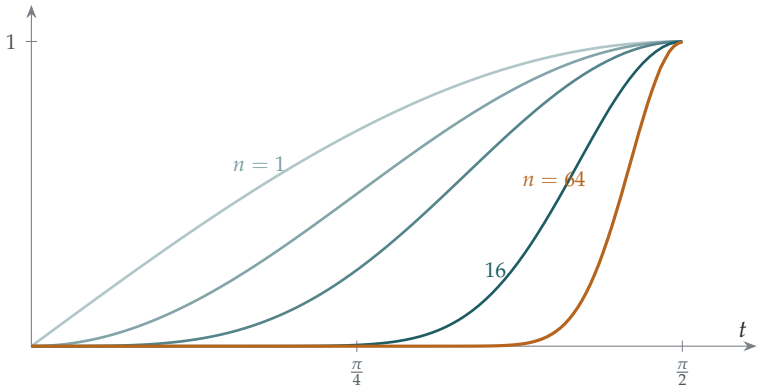


图 4.4.1 $n = 1, 2, 4, 16, 64$ 时 $\sin^n t$ 的图像。随着 n 增大，曲线下的面积 u_n 不断缩小，并集中在 $t = \frac{\pi}{2}$ 附近。本题证明 $u_n \approx \sqrt{\pi/(2n)}$；当 $n = 64$ 时，这个值约为 0.157。

回顾

- **递推公式（reduction formulae）。**分部积分法把 u_n 化为 u_{n-2}，所以所有 u_n 都由 u_0 和 u_1 决定。
- **寻找不变量（invariant）。**乘积 nu_nu_{n-1} 保持不变，正如 Section 2.6 中的积分在 AGM 步骤下保持不变一样。
- **单调性（monotonicity）加上不变量，便得出夹逼（squeeze）。**知道 $u_n < u_{n-1}$ 以及 nu_nu_{n-1} 的准确值，便把 nu_n^2 夹在两个极限相同的量之间。

背后的数学

沃利斯乘积

把递推公式一直用到 u_0 或 u_1，便得到明确的值：

$$u_{2m} = \frac{1\cdot 3\cdot 5\cdots(2m-1)}{2\cdot 4\cdot 6\cdots 2m}\cdot\frac{\pi}{2}, \qquad u_{2m+1} = \frac{2\cdot 4\cdot 6\cdots 2m}{3\cdot 5\cdot 7\cdots(2m+1)}.$$

由夹逼可知 $u_{2m}/u_{2m+1} \to 1$。把这个比写出来再整理，便得到沃利斯乘积

$$\frac{\pi}{2} = \lim_{m\to\infty}\frac{2\cdot 2}{1\cdot 3}\cdot\frac{4\cdot 4}{3\cdot 5}\cdots\frac{2m\cdot 2m}{(2m-1)(2m+1)}.$$

沃利斯在 1655 年以一种大胆的插值(interpolation)方法发现了它：他把 $\int_0^1(1-x^2)^{1/2}\,dx$ 放进一个指数为整数的积分表中，从而猜出它“应该”是甚么。这是数学史上最早的无

穷乘积之一，仅次于韦达（Viète）的乘积（Section 2.5），也是第一个只用有理数的无穷乘积。

二分之一的阶乘

沃利斯的插值想法，在 1729 年由欧拉（Euler）赋予了严格的意义。他找到一个函数，现称为*伽玛函数*（Gamma function），

$$\Gamma(s) = \int_0^\infty t^{s-1} e^{-t}\,\mathrm{d}t,$$

它满足 $\Gamma(s+1) = s\,\Gamma(s)$（用分部积分法），而且对整数 n 有 $\Gamma(n+1) = n!$。它把阶乘（factorial）平滑地延拓至所有正实数。与它密切相关的*贝塔函数*（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 中的积分，而且经代换 $x = \sin^2 t$ 之后，也包含我们的 u_n。特别地，$\int_0^1 x^{1/2}(1-x)^{1/2}\,\mathrm{d}x = \frac{\pi}{8}$（练习 2），由此得 $\Gamma(\frac{3}{2})^2/\Gamma(3) = \frac{\pi}{8}$，即 $\left(\frac{1}{2}\right)! = \Gamma(\frac{3}{2}) = \frac{\sqrt{\pi}}{2}$。等价地，$\Gamma(\frac{1}{2}) = \sqrt{\pi}$，这与著名的高斯积分（Gaussian integral）$\int_{-\infty}^{\infty} e^{-x^2}\,\mathrm{d}x = \sqrt{\pi}$ 是同一个事实，也正是 π 出现在正态分布（normal distribution）公式中的原因。

斯特灵公式

斯特灵公式说 $n! \approx \sqrt{2\pi n}\,(n/e)^n$。其中 $(n/e)^n$ 的部分相对容易得到，但常数 $\sqrt{2\pi}$ 则较难，而求出它的标准方法正是本题：u_{2m} 的明确值涉及 $(2m)!/(m!)^2$，而极限 $nu_n^2 \to \frac{\pi}{2}$ 则把这个常数确定下来。另有一条截然不同的途径，通过概率得到同一个常数，见 Section 12.23。

思维技巧

由积分得出递推关系

分部积分法往往把带有参数 n 的积分，与参数较小的同一类积分联系起来。这个递推关系（recurrence）随后可以解出，或用来分析 $n \to \infty$ 时的表现。

留意递推关系下保持不变的量，以及对 n 的单调性；两者结合，往往不需要任何明确公式便能求出极限。

延伸练习

1 **沃利斯乘积。**利用 u_{2m} 和 u_{2m+1} 的明确值，以及 $u_{2m+1} < u_{2m} < u_{2m-1}$，证明沃利斯乘积。

2 **一个贝塔积分。**试做 STEP II 2013, 第 2 题（`step.maths.org/questions/13-s2-q2`）。这题用递推公式计算 $\int_0^1 x^n(1-x)^n\,\mathrm{d}x$，然后问当 $n=\frac{1}{2}$ 时答案有何启示：这是对二分之一的阶乘的初次一瞥。

3 **贝塔函数。**试做 STEP III 2025，第 1 题（载于 OCR 提供下载的 2025 年试卷，`www.ocr.org.uk/Images/750486-2025-past-papers.zip`）。这题建立一般的贝塔积分 $\int_0^1 x^{p-1}(1-x)^{q-1}\,\mathrm{d}x$、它的对称性和一个递推关系，并用它们计算一个在 $[0,\infty)$ 上的积分。

4 **中心二项式系数（central binomial coefficients）。**较难证明 $u_{2m}=\binom{2m}{m}\dfrac{\pi}{2^{2m+1}}$，并利用 $nu_n^2\to\frac{\pi}{2}$ 推导出当 m 很大时 $\binom{2m}{m}\approx\dfrac{4^m}{\sqrt{\pi m}}$。就 $m=10$ 验证这个结果。

延伸阅读

- **Julian Havil, *Gamma***（Princeton University Press）。内容包括伽玛函数及其历史。
- **Philip J. Davis, “Leonhard Euler’s integral: a historical profile of the Gamma function”,** *American Mathematical Monthly* 66 (1959)。一篇获奖的经典文章，读过本节之后便能看懂。
- **Emil Artin, *The Gamma Function***（Dover）。一本简短而优美的专著，适合学过一些分析之后阅读。

4.5 大二学生之梦

题目

STEP III 2009，第 8 题

继续阅读之前，请先看题目（原题为英文）。题目载于 STEP Support Programme 的试题资料库：

`step.maths.org/questions/09-s3-q8`

引子

每个学生迟早都会学到 $(a+b)^2 \neq a^2+b^2$。那条错误的等式有个戏称，叫做“大一新生之梦”（freshman's dream）。本题要证明的结果，看起来同样像初学者会犯的错：

$$\int_0^1 x^x \,\mathrm{d}x = \sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n^n} = 1 - \frac{1}{2^2} + \frac{1}{3^3} - \frac{1}{4^4} + \cdots,$$

以及与它配对的 $\int_0^1 x^{-x}\,\mathrm{d}x = \sum_{n=1}^{\infty} \frac{1}{n^n}$。$x^x$ 的积分（integral）竟然变成 n^{-n} 的和：这似乎好得令人难以置信，所以它被称为大二学生之梦（sophomore's dream）。但它确实成立，由约翰·伯努利（Johann Bernoulli）于 1697 年发现。

证明用到三个你在大学会不断遇到的想法：比较指数函数与幂的极限（limit）、归约公式（reduction formula），以及把函数展开成幂级数（power series）再逐项积分（integrating term by term）。

动笔之前

想一想 函数 x^x 在 $x=0$ 处无法用你所知的任何公式定义（0^0 的意义并不明确）。当 x 从上方趋近 0 时，它的表现如何？试用计算机计算 $x = 0.1$、0.01、0.001 的情况。

$0.1^{0.1} \approx 0.794$，$0.01^{0.01} \approx 0.955$，$0.001^{0.001} \approx 0.993$。看来它趋近 1。第 (i) 部分正是证明这一点，由此可知被积函数（integrand）在 0 附近表现良好，这个积分是有意义的。x^x 在 $(0,1]$ 上的图像由接近 1 的地方出发，下降至极小值，再在 $x=1$ 处回到 1（图 4.5.1）。

逐步解题

阶段 1　幂胜过对数

想一想　当 $x \to 0^+$ 时，$\ln x \to -\infty$。哪一个代换（substitution）可以把 $x \to 0^+$ 变成 $t \to \infty$，并把 $x^m (\ln x)^n$ 变成类似 $e^{-mt} t^n$ 的式子？

设 $x = e^{-t}$。当 $x \to 0^+$ 时，$t \to \infty$，而 $\ln x = -t$，所以

$$x^m (\ln x)^n = e^{-mt} (-t)^n = (-1)^n e^{-mt} t^n \to 0.$$

用文字来说：当 $x \to 0$ 时，x 的任何正数次幂趋向零的速度，都快过 $\ln x$ 的任何次幂趋向无穷的速度。 ■

取 $m = n = 1$，得 $x \ln x \to 0$；由于指数函数是连续（continuous）的，所以 $x^x = e^{x \ln x} \to e^0 = 1$。 ■

阶段 2　归约公式

想一想　用分部积分法（integration by parts）计算 $I_{n+1} = \int_0^1 x^m (\ln x)^{n+1}\,\mathrm{d}x$。应该对哪一个因子积分？边界项（boundary term）在 $x = 0$ 处会怎样？

对 x^m 积分，对 $(\ln x)^{n+1}$ 求导：

$$I_{n+1} = \left[\frac{x^{m+1}}{m+1}(\ln x)^{n+1}\right]_0^1 - \int_0^1 \frac{x^{m+1}}{m+1} \cdot \frac{(n+1)(\ln x)^n}{x}\,\mathrm{d}x.$$

在 $x = 1$ 处，因为 $\ln 1 = 0$，边界项等于 0；当 $x \to 0$ 时，根据第 (i) 部分（取 $m+1$ 和 $n+1$），它趋向 0。余下的积分是 $\frac{n+1}{m+1} I_n$，所以

$$I_{n+1} = -\frac{n+1}{m+1} I_n. \qquad ■$$

由 $I_0 = \int_0^1 x^m\,\mathrm{d}x = \frac{1}{m+1}$ 出发，应用这条公式 n 次，得

$$I_n = (-1)^n \frac{n!}{(m+1)^n} \cdot \frac{1}{m+1} = \frac{(-1)^n\, n!}{(m+1)^{n+1}}. \qquad ■$$

阶段 3　级数

想一想　利用 e^y 的级数展开 $x^x = e^{x \ln x}$。每一项是甚么样子？m 取甚么值，才会出现第 (ii) 部分的积分？

由于对任何 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，试一试 由 0 至 1 逐项积分，并使用第 (ii) 部分中 $m = n$ 的情况（$n = 0$ 另行处理）。你得到甚么？

$n = 0$ 的一项是 1，积分后得 1。当 $n \geq 1$ 时，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 项的贡献是 $\frac{(-1)^n}{(n+1)^{n+1}}$。把各项相加，

$$\int_0^1 x^x \, dx = 1 - \frac{1}{2^2} + \frac{1}{3^3} - \frac{1}{4^4} + \cdots . \qquad \blacksquare$$

（$n = 0$ 的一项也符合同一规律，即 $\frac{1}{1^1}$。）各项缩小得极快：取七项，便得到准确至小数点后六位的数值 0.783430…。

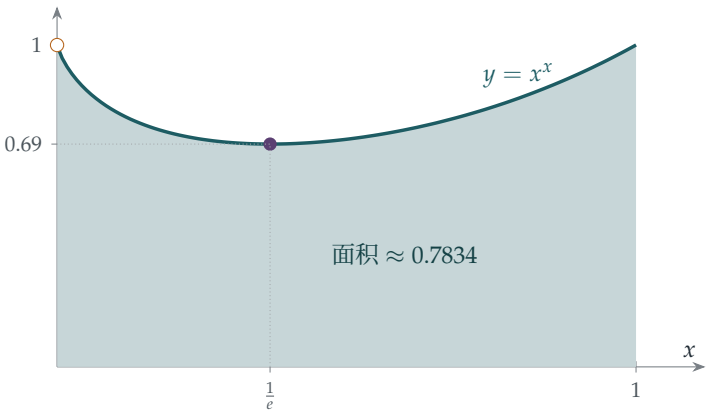


图 4.5.1 x^x 在 $(0, 1]$ 上的图像。当 $x \to 0$ 时它趋向 1（空心圆点），在 $x = \frac{1}{e}$ 处下降至极小值 $e^{-1/e} \approx 0.692$，再在 $x = 1$ 处回到 1。阴影部分的面积是 $\sum (-1)^{n+1} n^{-n} \approx 0.7834$。

回顾

- **用代换比较增长速度。**设 $x = e^{-t}$，便把在 0 处的极限变成在 ∞ 处一个熟悉的极限。
- **边界项必须检查。**归约公式有赖边界项在 $x = 0$ 处消失，这正是用到第 (i) 部分的地方。
- **先展开成级数，再积分。**把被积函数展开成幂级数并逐项积分，便把一个没有初等原函数（elementary antiderivative）的积分，变成一个收敛得很快的和。

背后的数学

交换求和与积分

第 3 阶段不动声色地把无穷和与积分交换了次序：$\int \sum = \sum \int$。这样做并非总是容许的。在大学里，你会遇到交换后不成立的例子，也会学到一些定理，说明甚么时候可以放心交换。最简单的一个定理指出，当级数一致收敛（converges uniformly）时，交换是安全的：即级数的尾部可以同时在每一个 x 处都变得很小。在这里，$|x \ln x| \le \frac{1}{e}$ 在 $(0,1]$ 上成立，所以无论 x 是甚么，第 n 项的大小最多是 $\frac{e^{-n}}{n!}$。这些上界之和收敛，因此尾部一致地小，交换也就有了根据。这个论证称为*魏尔斯特拉斯 M 判别法*（Weierstrass M-test），是分析学（analysis）中最常用的工具之一。

为何这个“梦”会成真

这条恒等式很特别。它有赖 $\frac{(x\ln x)^n}{n!}$ 中的幂与积分 $\int_0^1 x^n(\ln x)^n\,dx = \frac{(-1)^n n!}{(n+1)^{n+1}}$ 完全吻合，当中的 $n!$ 恰好全部约去。追根究柢，这个积分其实是乔装了的伽玛函数（Gamma function）：代换 $x = e^{-t/(n+1)}$ 把它变成 $\frac{(-1)^n}{(n+1)^{n+1}}\int_0^\infty t^n e^{-t}\,dt = \frac{(-1)^n\,\Gamma(n+1)}{(n+1)^{n+1}}$，与 Section 4.4 连系起来。

名字的由来

“大二学生之梦”这个名字，是 Jonathan Borwein 与合著者在 2004 年出版的 *Experimentation in Mathematics* 一书中所创，用来与“大一新生之梦”$(a+b)^n = a^n + b^n$ 配对（有趣的是，后者在某些数系中是成立的：在模质数 p 的算术中，$(a+b)^p = a^p + b^p$，因为中间所有的二项式系数（binomial coefficient）都可被 p 整除）。Borwein 是*实验数学*（experimental mathematics）的先驱，这门学问利用电脑去发现像本题这样的恒等式，然后再以人手证明。

思维技巧

展开后逐项积分

如果被积函数有幂级数展开，而你又能对其中各项积分，便对每一项积分，再把结果加起来。

对于没有初等原函数的积分，这是一个很有效的求值方法。紧记要问：交换求和与积分是否有根据？通常是有的，但你应该知道原因。

延伸练习

1 **配对的等式。**证明 $\displaystyle\int_0^1 x^{-x}\,\mathrm{d}x = \sum_{n=1}^{\infty} \frac{1}{n^n} \approx 1.2913$。

2 **极小值。**证明 x^x 在 $(0,\infty)$ 上的极小值出现于 $x = \frac{1}{e}$，其值为 $e^{-1/e}$。

3 **需要多少项？**对于各项递减的交错级数（alternating series），截断后的误差小于第一个被略去的项。要准确至小数点后六位，需要取 $\int_0^1 x^x\,\mathrm{d}x$ 的级数的多少项？

4 **一个近亲。**用同样的方法证明 $\displaystyle\int_0^1 x^2 \ln x\,\mathrm{d}x = -\frac{1}{9}$，以及 $\displaystyle\int_0^1 \frac{\ln x}{1-x}\,\mathrm{d}x = -\sum_{n=1}^{\infty} \frac{1}{n^2}$。（根据 Section 2.4，这等于 $-\frac{\pi^2}{6}$。）

延伸阅读

- **Jonathan Borwein, David Bailey & Roland Girgensohn, *Experimentation in Mathematics***（A. K. Peters）。这个名字的出处，书中还有许多由电脑发现的恒等式。
- **William Dunham, *The Calculus Gallery***（Princeton University Press）。由牛顿（Newton）到勒贝格（Lebesgue）的分析学杰作，包括伯努利的工作。
- **Mary Hart, *Guide to Analysis***（Palgrave）。清楚讲解一致收敛，以及何时可以交换求和与积分。

4.6 没有公式的积分

题目

STEP III 2016，第 3 题

继续阅读之前，请先看题目（原题为英文）。题目载于 STEP Support Programme 的试题资料库：

`step.maths.org/questions/16-s3-q3`

引子

求导（differentiation）是机械式的工作：有一套法则可循，任何由常见部件砌成的函数，只要套用这些法则就能求导。积分（integration）却不然。每个学生都遇过各种技巧都攻不破的积分，并怀疑是否只是自己不够聪明。有时答案是：谁也不够聪明，因为根本没有公式。

本题就在一个特殊情况下证明这样的事。积分 $\int \frac{e^x}{x+1}\,\mathrm{d}x$ 作为面积当然存在，但它不能写成（有理函数）$\times e^x$ 的形式，而这正是用分部积分法（integration by parts）时预期会得到的形式。证明某件事不可能做到，而不是把某样东西找出来，是另一种数学，也极具数学的特色。

动笔之前

想一想 若 $\int \mathrm{f}(x)e^x\,\mathrm{d}x = \frac{\mathrm{P}}{\mathrm{Q}}e^x + C$，$\mathrm{P}$ 和 Q 必须满足甚么方程？（对右边求导。）

求导得 $\dfrac{\mathrm{d}}{\mathrm{d}x}\left(\dfrac{\mathrm{P}}{\mathrm{Q}}e^x\right) = \dfrac{\mathrm{P}'\mathrm{Q} - \mathrm{P}\mathrm{Q}' + \mathrm{P}\mathrm{Q}}{\mathrm{Q}^2}e^x$，而这必须等于 $\mathrm{f}(x)e^x$。所以整道题目其实围绕着以下的多项式恒等式（polynomial identity）：

$$\mathrm{P}'\mathrm{Q} - \mathrm{P}\mathrm{Q}' + \mathrm{P}\mathrm{Q} = \mathrm{f}\mathrm{Q}^2.$$

关于积分的问题，就此变成关于多项式的问题：它们的因式和次数。

逐步解题

阶段 1　因式 $x+1$

取 $\mathrm{f}(x)=\dfrac{x^3-2}{(x+1)^2}$，把恒等式两边同乘 $(x+1)^2$：

$$(x+1)^2(\mathrm{P}'\mathrm{Q}-\mathrm{P}\mathrm{Q}'+\mathrm{P}\mathrm{Q})=(x^3-2)\,\mathrm{Q}^2. \tag{†}$$

> ? **想一想**　在 (†) 中代入 $x=-1$。必然会出现甚么情况?

当 $x=-1$ 时，左边等于 0，右边则是 $(-1-2)\,\mathrm{Q}(-1)^2=-3\,\mathrm{Q}(-1)^2$。所以 $\mathrm{Q}(-1)=0$，而根据因式定理（factor theorem），$x+1$ 是 Q 的因式。 ■

阶段 2　次数

设 P 的次数（degree）为 p，Q 的次数为 q。

> ? **想一想**　在 $\mathrm{P}'\mathrm{Q}-\mathrm{P}\mathrm{Q}'+\mathrm{P}\mathrm{Q}$ 中，哪一项的次数最高?它会被其余各项抵消吗?

$\mathrm{P}\mathrm{Q}$ 的次数是 $p+q$，而 $\mathrm{P}'\mathrm{Q}$ 和 $\mathrm{P}\mathrm{Q}'$ 的次数最多是 $p+q-1$。所以 (†) 左边的次数恰好是 $2+p+q$，右边的次数是 $3+2q$。令两者相等，得 $p=q+1$。 ■

阶段 3　求 P

取 $\mathrm{Q}=x+1$，则 $\mathrm{Q}'=1$，而 P 的次数为 2。把 (†) 除以 $(x+1)^2$：

$$\mathrm{P}'(x+1)-\mathrm{P}+\mathrm{P}(x+1)=x^3-2, \qquad \text{即} \qquad (x+1)\mathrm{P}'+x\mathrm{P}=x^3-2.$$

试设 $\mathrm{P}=\alpha x^2+\beta x+\gamma$。左边是

$$(x+1)(2\alpha x+\beta)+\alpha x^3+\beta x^2+\gamma x=\alpha x^3+(2\alpha+\beta)x^2+(2\alpha+\beta+\gamma)x+\beta.$$

与 x^3-2 比较系数：$\alpha=1$；$2\alpha+\beta=0$，故 $\beta=-2$；于是 $\gamma=0$；常数项 $\beta=-2$ 亦吻合。所以

$$\mathrm{P}(x)=x^2-2x, \qquad \int \frac{x^3-2}{(x+1)^2}\,e^x\,\mathrm{d}x=\frac{x^2-2x}{x+1}\,e^x+C. \qquad ■$$

（有兴趣的话可以求导验算，这是练习商法则（quotient rule）的好题目。）

阶段 4　一件不可能的事

现在假设 $\int \frac{e^x}{x+1}\,\mathrm{d}x=\frac{\mathrm{P}}{\mathrm{Q}}e^x+C$，其中 P 和 Q 没有公因式（common factor）。恒等式变成

$$(x+1)\left(\mathrm{P}'\mathrm{Q}-\mathrm{P}\mathrm{Q}'+\mathrm{P}\mathrm{Q}\right)=\mathrm{Q}^2.$$

一如第 1 阶段，代入 $x=-1$ 可知 $Q(-1)=0$。写成 $Q=(x+1)^k R$，其中 $k\geq 1$ 且 $R(-1)\neq 0$（即尽量抽出 $x+1$ 这个因式）。

> **停一停，试一试** 代入 $Q=(x+1)^k R$，尽量约去 $(x+1)$ 的因式，然后代入 $x=-1$。你能得知 $P(-1)$ 的甚么性质？

这时 $Q'=k(x+1)^{k-1}R+(x+1)^k R'$，而

$$P'Q-PQ'+PQ=(x+1)^{k-1}\Big[(x+1)\big(P'R-PR'+PR\big)-kPR\Big].$$

所以恒等式变成 $(x+1)^k\big[(x+1)(P'R-PR'+PR)-kPR\big]=(x+1)^{2k}R^2$。除以 $(x+1)^k$：

$$(x+1)\big(P'R-PR'+PR\big)-kPR=(x+1)^k R^2.$$

当 $x=-1$ 时，由于 $k\geq 1$，上式给出 $-kP(-1)R(-1)=0$。因为 $k\neq 0$ 且 $R(-1)\neq 0$，必有 $P(-1)=0$。但这样 $x+1$ 就同时整除 P 和 Q，与两者没有公因式的假设矛盾。所以这样的 P 和 Q 并不存在。 ■

回顾

- **转化为代数。**“这个积分能否写成这种形式?”变成了“这个多项式恒等式有没有解?”
- **在特殊点取值。**代入 $x=-1$ 就得到关于因式的信息，前后共用了三次。
- **利用最简分数的反证法（proof by contradiction）。**正如 Section 1.3 中处理 $\sqrt{2}$ 的做法，先假设分数已约至最简，再找出一个公因式，便得到矛盾。

背后的数学

极点与对数

第 (ii) 部分背后有一幅简单的图像。被积函数（integrand）$\frac{e^x}{x+1}$ 在 $x=-1$ 处有一个**单极点**（simple pole）：在该点附近，它的表现像 $\frac{e^{-1}}{x+1}$。对 $\frac{1}{x+1}$ 积分得 $\ln|x+1|$，所以原函数（antiderivative）在 -1 附近的表现应该像 $e^{-1}\ln|x+1|$，以对数的速度趋向 $-\infty$（图 4.6.1）。但有理函数乘以 e^x 只能以幂 $(x+1)^{-k}$ 的方式趋向无穷，其导数因而有一个阶数为 $k+1\geq 2$ 的极点，绝不会是单极点。有理函数无法产生对数式的表现；第 4 阶段的代数运算，正是这个观察的严格版本。相反，在第 (i) 部分，被积函数有一个二阶极点，而这正是带有单极点的函数求导后的模样，所以答案才有可能是有理的。

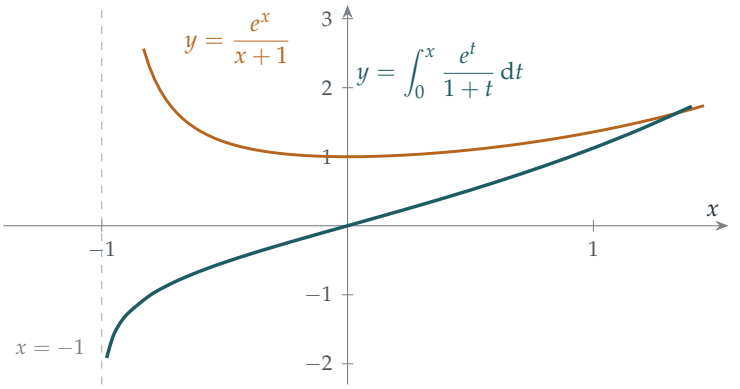


图 4.6.1　被积函数 $\frac{e^x}{x+1}$（琥珀色）在 $x = -1$ 处有一个单极点。它的原函数（蓝绿色，以数值方法计算）完全存在，但当 $x \to -1^+$ 时，它以对数的速度趋向 $-\infty$，表现像 $e^{-1}\ln(x+1)$。没有任何有理函数乘以 e^x 能有这样的表现。

刘维尔与 Risch 算法

在 1830 年代，刘维尔（Joseph Liouville）建立了一套理论，研究哪些积分可以用*初等函数*（elementary functions）表示（即由多项式、指数函数、对数函数、根式和三角函数经通常的运算构成的函数）。他的其中一个结果指出：若 f 和 g 是有理函数而 g 不是常数，则 $\int \mathrm{f}(x)e^{\mathrm{g}(x)}\,\mathrm{d}x$ 只有在等于 $\mathrm{R}(x)e^{\mathrm{g}(x)}$（另加一个常数）时才是初等的，其中 R 是某个有理函数。（对 g 的条件不可或缺：当 $\mathrm{g} = 0$ 而 $\mathrm{f}(x) = \frac{1}{x}$ 时，积分是 $\ln|x|$，它是初等函数，却不是有理函数。）所以本题排除了有理形式，其实就证明了 $\int \frac{e^x}{x+1}\,\mathrm{d}x$ 根本不是初等函数。同一定理的其他版本可证明 $\int e^{-x^2}\,\mathrm{d}x$ 和 $\int \frac{\sin x}{x}\,\mathrm{d}x$ 都不是初等的。1969 年，Robert Risch 把刘维尔的想法发展成一套算法（algorithm），可就一大类函数判定初等积分是否存在，并在存在时把它求出来。电脑代数系统（computer algebra system）内部所用的，就是这套算法的不同版本。

特殊函数

当一个积分没有初等形式时，数学家会给它起个名字，把它当作新的函数来研究，正如我们在 Section 3.9 中处理朗伯 W 函数（Lambert W function）那样。这里的积分本质上就是*指数积分*（exponential integral）$\mathrm{Ei}(x)$，它出现于热传导问题，也出现于*对数积分*（logarithmic integral）$\mathrm{li}(x) = \mathrm{Ei}(\ln x)$ 之中，后者是不超过 x 的质数（prime）个数的最佳简单近似（Section 1.2）。当 $x < 0$ 时，它是 $\int_{-\infty}^{x} \frac{e^t}{t}\,\mathrm{d}t$。当 $x > 0$ 时，这个积分会穿过 $t = 0$ 处的无穷尖峰，所以 $\mathrm{Ei}(x)$ 的定义是先挖去尖峰周围的一个小区间 $(-\varepsilon, \varepsilon)$，再令 $\varepsilon \to 0$：这就是*柯西主值*（Cauchy principal value）。

思维技巧

证明不可能

要证明某件事做不到，先假设它做得到，把这个假设转化为精确的代数式，然后推出矛盾。

困难在于转化：要为“做得到”的意思找出一个精确的陈述。在这里，它的意思是“存在没有公因式的多项式 P、Q 满足这个恒等式”。数学中许多最深刻的结果，都是这类不可能性证明。

延伸练习

1 **两个可行的例子。**证明 $\int \frac{xe^x}{(x+1)^2}\,dx = \frac{e^x}{x+1} + C$ 及 $\int \frac{(x^2+1)e^x}{(x+1)^2}\,dx = \frac{x-1}{x+1}e^x + C$。

2 **高斯函数。**证明不存在多项式 P 使得 $\int e^{-x^2}\,dx = P(x)e^{-x^2} + C$。（提示：求导后比较次数。）

3 **另一个极点。**仿照第 (ii) 部分的论证，证明 $\int \frac{e^x}{x}\,dx$ 不能写成 $\frac{P(x)}{Q(x)}e^x + C$ 的形式。

4 **何时可行？**假设 f 是多项式，且对某个多项式 P 有 $\int \frac{f(x)}{(x+1)^2}e^x\,dx = \frac{P(x)}{x+1}e^x + C$。证明 $(x+1)P' + xP = f$，并通过计算此式及其导数在 $x = -1$ 的值，证明 $f(-1) + f'(-1) = 0$。验证这个条件对 $f(x) = x^3 - 2$ 以及练习 1 的例子都成立。较难你能证明这个条件也是充分的吗？

延伸阅读

- **Brian Conrad, “Impossibility theorems for elementary integration”**（初等积分的不可能性定理）。可免费取得的讲义，配合例子讲解刘维尔定理。
- **Elena Anne Marchisotto & Gholam-Ali Zakeri, “An invitation to integration in finite terms”**（有限形式积分导引），*The College Mathematics Journal* 25 (1994)。刘维尔理论的浅白入门。
- **Manuel Bronstein, *Symbolic Integration I***（Springer）。讲述电脑代数系统实际上如何求积分。程度较深。

4.7 变分法

题目

STEP III 2014，第 4 题

继续阅读之前，请先看题目（原题为英文）。题目载于 STEP Support Programme 的试题资料库：

`step.maths.org/questions/14-s3-q4`

引子

中学里的最优化（optimisation）问题，大多是求最佳的*数*：例如求令面积最大的 x 值。*变分法*（calculus of variations）要求的却是最佳的*函数*。连接两点的曲线之中，珠子沿哪一条滑下最快？悬挂的链条会呈甚么形状？光线穿过透镜时走哪一条路径？在这些例子中，要优化的量都是一个积分，而未知的是一整个函数。

本题研究的是最简单而又有趣的例子。积分 $J = \int_0^1 ((y')^2 - a^2 y^2)\,\mathrm{d}x$ 在两个互相竞争的项之间取得平衡：函数愈陡，积分愈大；函数值愈大，积分愈小。当 a 较小时，陡峭程度总是占上风，所以对每个函数都有 $J \geq 0$。但 a 有一个临界值，平衡会在那里逆转，而本题正好把它准确求出：$a = \frac{\pi}{2}$。π 从何而来，就是本节要讲的故事。

动笔之前

想一想 为甚么 $I \geq 0$ 并非显而易见？试用 $y = 1 - x$，它满足 $y(1) = 0$。

I 是两个正数量之差，所以它可正可负。对于 $y = 1 - x$：$\int_0^1 1\,\mathrm{d}x = 1$ 而 $\int_0^1 (1-x)^2\,\mathrm{d}x = \frac{1}{3}$，所以 $I = \frac{2}{3} > 0$。但若函数选得更巧妙，或许可以令 $\int y^2$ 相对于 $\int (y')^2$ 变得更大。题目的窍门，是把 I 写成一个*平方*的积分，而这显然是非负的。这就是积分版本的“配方法”（completing the square）。

逐步解题

阶段 1　配方

把 I_1 中的平方展开：

$$I_1 = \int_0^1 \left((y')^2 + 2yy' \tan x + y^2 \tan^2 x\right) \mathrm{d}x,$$

所以

$$I - I_1 = -\int_0^1 \left(y^2 + y^2 \tan^2 x + 2yy' \tan x\right) \mathrm{d}x = -\int_0^1 \left(y^2 \sec^2 x + 2yy' \tan x\right) \mathrm{d}x.$$

> **想一想**　被积函数 $y^2 \sec^2 x + 2yy' \tan x$ 恰好是某个函数的导数（exact derivative）。是哪个函数？

根据乘积法则（product rule），$\dfrac{\mathrm{d}}{\mathrm{d}x}(y^2 \tan x) = 2yy' \tan x + y^2 \sec^2 x$。所以

$$I - I_1 = -\Big[y^2 \tan x\Big]_0^1 = -(y(1)^2 \tan 1 - y(0)^2 \tan 0) = 0,$$

因为 $y(1) = 0$ 而 $\tan 0 = 0$。所以 $I = I_1 \geq 0$，因为它是一个平方的积分。 ■

阶段 2　何时 $I = 0$？

> **想一想**　如果一个连续的非负函数的积分是零，你对这个函数能下甚么结论？那么若 $I = 0$，y 必须满足甚么方程？

积分为零的连续非负函数必定恒等于零（假如它在某一点取正值，它在该点附近的一个小区间上也会取正值，从而贡献一个正的量）。所以 $I = 0$ 迫使 $y' + y \tan x = 0$ 在 $[0, 1]$ 上成立。乘以 $\sec x$ 后，这就是说 $\frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{y}{\cos x}\right) = \frac{y' \cos x + y \sin x}{\cos^2 x} = 0$，所以 $y = C \cos x$。于是 $y(1) = C \cos 1 = 0$，而由于 $\cos 1 \neq 0$，得 $C = 0$。所以只有当 $y \equiv 0$ 时才有 $I = 0$。 ■

阶段 3　一般情况

> **想一想**　用 $\int_0^1 (y' + ay \tan bx)^2 \,\mathrm{d}x$ 重复第 1 步。b 必须是甚么，余下的项才能恰好组成一个导数？

把这个积分记作 K。那么

$$\begin{aligned} J - K &= -\int_0^1 \left(a^2 y^2 + a^2 y^2 \tan^2 bx + 2ayy' \tan bx\right) \mathrm{d}x \\ &= -\int_0^1 \left(a^2 y^2 \sec^2 bx + 2ayy' \tan bx\right) \mathrm{d}x. \end{aligned}$$

另一方面，$\frac{\mathrm{d}}{\mathrm{d}x}(ay^2\tan bx) = 2ayy'\tan bx + ab\,y^2\sec^2 bx$。两者恰好在 $ab = a^2$ 时吻合，即 $b = a$。于是

$$J - K = -\Big[ay^2\tan ax\Big]_0^1 = -a\,y(1)^2\tan a = 0,$$

所以 $J = K \geq 0$。

> **停一停，试一试** 这个论证在哪里可能失效？（留意 $\tan ax$ 在区间 $0 \leq x \leq 1$ 上的情况。）

这个论证需要 $\tan ax$ 对 $[0,1]$ 内所有 x 都有定义而且连续；否则 K 未必存在，微积分基本定理（fundamental theorem of calculus）亦不能应用。当 x 取遍 $[0,1]$ 时，ax 取遍 $[0,a]$，而 tan 在该处连续的充要条件正是 $a < \frac{\pi}{2}$。所以这个证明在以下情况成立：

$$a < k = \frac{\pi}{2}. \qquad \blacksquare$$

阶段 4 临界情况

> **想一想** 当 $a = \frac{\pi}{2}$ 时，证明失效。在第 2 步，令 $I = 0$ 的函数须满足 $y' + y\tan x = 0$。现在对应的方程是甚么？它的解又是甚么？

对应的方程是 $y' + ay\tan ax = 0$，它的解是 $y = C\cos ax$。当 $a = \frac{\pi}{2}$ 时，$y = \cos\frac{\pi x}{2}$ 满足 $y(1) = \cos\frac{\pi}{2} = 0$。直接验证：

$$\begin{aligned} J &= \frac{\pi^2}{4}\int_0^1 \left(\sin^2\frac{\pi x}{2} - \cos^2\frac{\pi x}{2}\right)\mathrm{d}x = -\frac{\pi^2}{4}\int_0^1 \cos\pi x\,\mathrm{d}x \\ &= -\frac{\pi^2}{4}\left[\frac{\sin\pi x}{\pi}\right]_0^1 = 0. \qquad \blacksquare \end{aligned}$$

所以在临界值处，有一个非零函数令 J 恰好等于零。

回顾

- **在积分中配方。**加上再减去一个恰好是导数的式子，便把 J 化成一个平方的积分，再加上一个等于零的边界项（boundary term）。
- **查看工具在哪里失效。**这个方法只在 $\tan ax$ 在区间上保持有限时才行得通。它失效之处 $a = \frac{\pi}{2}$，正正就是临界值。
- **极端情况是某个微分方程的解。**令 $J = 0$ 的函数来自一阶微分方程（differential equation）$y' + ay\tan ax = 0$，而它同时满足 $y'' + a^2y = 0$。

背后的数学

最速降线与一门学科的诞生

1696 年，约翰·伯努利（Johann Bernoulli）向欧洲的数学家发出挑战：求一条曲线，使珠子沿着它从一点滑到另一点所需的时间最短。这就是最速降线（brachistochrone）问题。提交解答的有牛顿（Newton）（他匿名投稿，一夜之间解出；伯努利说他“从爪子认出了狮子”）、莱布尼兹（Leibniz）、洛必达（l'Hôpital）和雅各布·伯努利（Jakob Bernoulli）。答案是一条摆线（cycloid）。到了 1750 年代，欧拉（Euler）和年轻的拉格朗日（Joseph-Louis Lagrange）把这些个别的技巧变成一套通用的方法：如果 y 令 $\int F(x, y, y')\,\mathrm{d}x$ 取平稳值，那么 y 满足*欧拉－拉格朗日方程*（Euler–Lagrange equation）

$$\frac{\mathrm{d}}{\mathrm{d}x}\frac{\partial F}{\partial y'} = \frac{\partial F}{\partial y}.$$

对于我们的 J，$F = (y')^2 - a^2y^2$，而欧拉－拉格朗日方程就是 $y'' + a^2y = 0$：简谐运动（simple harmonic motion）。

为何是 $\pi/2$?

当 $a < \frac{\pi}{2}$ 时，J 的最小值是 0，而且只有 $y \equiv 0$ 才能取得。当 $a = \frac{\pi}{2}$ 时，最小值仍然是 0，但多了一个函数能取得它。当 $a > \frac{\pi}{2}$ 时，同一个函数 $y = \cos\frac{\pi x}{2}$ 给出

$$J = \frac{\pi^2}{8} - \frac{a^2}{2} < 0,$$

而只要把 y 按比例放大，J 就可以负得要多大有多大：根本没有最小值（图 4.7.1）。临界值就是第一个令 $y'' + a^2y = 0$ 有非零解满足边界条件的 a（边界条件是 $y(1) = 0$，以及在自由端的“自然”条件 $y'(0) = 0$）。这是一个*特征值问题*（eigenvalue problem），而 $a^2 = \frac{\pi^2}{4}$ 是它最小的特征值。试与 Section 3.7 比较：那里把函数在长度为 π 的区间的两端都固定，得出临界值 1；这里区间长度为 1，而且只固定一端，相当于长度为 2 的区间的一半，所以临界的 a 是 $\frac{\pi}{2}$。

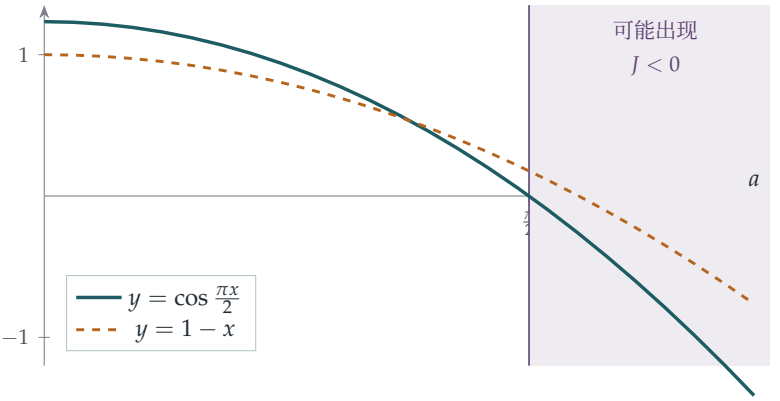


图 4.7.1 对于两个满足 $y(1) = 0$ 的试验函数，J 的值随 a 的变化。对于 $y = 1 - x$，$J = 1 - \frac{a^2}{3}$ 要到 $a > \sqrt{3}$ 才变成负数。对于 $y = \cos \frac{\pi x}{2}$，$J = \frac{\pi^2}{8} - \frac{a^2}{2}$ 在 $a > \frac{\pi}{2}$ 时随即变成负数，而且没有函数能做得更好：当 $a < \frac{\pi}{2}$ 时，对每个 y 都有 $J \geq 0$。

物理学与最小作用量原理

后来发现，变分法描述的竟是大自然本身。在十八和十九世纪，人们逐渐明白运动定律可以这样表述：力学系统的运动方式，会令一个称为作用量（action）的积分取平稳值，这就是最小作用量原理（principle of least action）。同样，光沿着需时最短的路径前进（费马原理，Fermat's principle），这解释了折射现象。到了二十世纪，同一个想法通过费曼（Richard Feynman）的“路径求和”（sum over paths），成为量子力学的基础。本题中的二次积分是最简单的例子，而它们引出的特征值问题，会在 Section 8.6 以量子振子能级的面貌再次出现。

思维技巧

先配方，再留意边界

要证明一个二次式非负，可把它写成一个平方加上一些你能控制的东西。对于积分，这些“东西”往往恰好是某个函数的导数，积分后便成为边界项。

此外，永远要问这个方法在哪里不再奏效。本题中方法失效之处并非偶然；它正是题目中最重要的那个数。

延伸练习

1 **越过临界值。**证明若 $a > \frac{\pi}{2}$，则 $y = \cos\frac{\pi x}{2}$ 给出 $J < 0$。

2 **欧拉－拉格朗日方程。**较难设 y 在满足 $y(1) = 0$ 的函数之中令 J 取最小值。对任何满足 $\eta(1) = 0$ 的函数 η 和任何数 ε，都有 $J(y + \varepsilon\eta) \geq J(y)$。把 $J(y + \varepsilon\eta)$ 按 ε 的幂展开，并由此推出 $\int_0^1 (y'\eta' - a^2 y\eta)\,\mathrm{d}x = 0$。用分部积分法（integration by parts）证明 $y'' + a^2 y = 0$ 及 $y'(0) = 0$（假设 y 足够光滑）。

3 **较长的区间。**用代换法（substitution）证明：若 $y(L) = 0$，则 $\int_0^L \left((y')^2 - a^2 y^2\right)\mathrm{d}x \geq 0$ 对所有 y 成立，当且仅当 $a \leq \frac{\pi}{2L}$；并证明只有当 $a = \frac{\pi}{2L}$ 时，才有非零的 y 令等号成立。

4 **摆线。**查阅最速降线问题，并找出摆线的参数方程（parametric equations）。证明在一条倒置的摆线上，珠子无论从哪一点由静止释放，到达最低点所需的时间都相同（这是等时（tautochrone）性质，由惠更斯（Huygens）于 1659 年发现）。

延伸阅读

- **Paul Nahin, *When Least Is Best***（Princeton University Press）。讲述最优化的历史与数学，包括最速降线，为学生而写。
- **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）。可在网上免费阅读。
- **Bruce van Brunt, *The Calculus of Variations***（Springer）。一本大学本科教科书；头几章只需要高级程度（A-level）的微积分知识。

4.8 切球形面包

题目

STEP I 2001，第 6 题

继续阅读之前，请先看题目（原题为英文）。题目载于 STEP Support Programme 的试题资料库：

`step.maths.org/questions/01-s1-q6`

引子

把一个球形面包切成厚度相等的薄片。靠近中间的几片是又大又厚的圆片，只有边缘一圈窄窄的面包皮；靠近两端的几片则是小小的球冠，看起来几乎全是面包皮。两端的薄片，面包皮一定比较多吧?

其实不然。不论切在靠近中间还是靠近两端的位置，每一片的面包皮面积都完全相同。中间那圈窄带周长很大，但几乎是垂直的；两端的球冠周长很小，但倾斜得很厉害，所以面包皮铺开的阔度反而较大。这两种效应恰好互相抵消。这是阿基米德（Archimedes）的一条定理，他对这条定理珍爱有加，甚至要求在自己的墓碑上刻上一个内切于圆柱（cylinder）的球（sphere）。题目的第二部分是一则较为轻松的“规例”，但它所依据的正是同一个计算。

动笔之前

想一想 半径为 a 的球，是由哪一条曲线绕 x-轴旋转而成的？该曲线的 $\frac{\mathrm{d}y}{\mathrm{d}x}$ 是甚么？

把半圆（semicircle）$y = \sqrt{a^2 - x^2}$, $-a \leq x \leq a$ 旋转，便得到这个球。于是 $\frac{\mathrm{d}y}{\mathrm{d}x} = \frac{-x}{\sqrt{a^2-x^2}} = -\frac{x}{y}$。请记住这一点：斜率（slope）在两端附近（$x \to \pm a$）很陡，在中间则很平缓，这正是引子中所描述的效应。

逐步解题

阶段 1　面包皮公式的化简

> **想一想**　对这个半圆计算 $y\sqrt{1+(y')^2}$。奇妙的事情出现了。

由于 $y' = -x/y$，

$$1 + (y')^2 = 1 + \frac{x^2}{y^2} = \frac{y^2 + x^2}{y^2} = \frac{a^2}{y^2},$$

所以 $y\sqrt{1+(y')^2} = y \cdot \frac{a}{y} = a$，竟然是一个常数！因此，任意两个平面 $x = x_1$ 与 $x = x_2$ 之间的面包皮面积为

$$2\pi \int_{x_1}^{x_2} a \, \mathrm{d}x = 2\pi a (x_2 - x_1).$$

它只取决于两个平面之间的距离，而与平面的位置无关。

阶段 2　面包皮与片数成正比

如果把面包切成厚度为 h 的薄片，每一片的面包皮面积都是 $2\pi a h$，片片相同。不论是从一端吃、从两端吃，还是从任何地方吃掉若干片，剩下的面包皮面积都是 $2\pi a h$ 乘以剩下的片数。 ■

阶段 3　“脆皮”的条件

题目的提示所描述的，是从一端吃掉若干片之后剩下的面包，我们就依照提示去做：假设从 $x = a$ 那一端开始吃，于是剩下的部分由 $x = -a$ 延伸至 $x = -a + t$。一项关于“任意片数”的规定，应该不论面包切得多薄都成立，所以我们把 t 看成可以取 0 与 $2a$ 之间的任何值。（第 4 阶段会探讨从两端吃的情况。）

> **想一想**　利用题目给出的公式计算 V。使用代换（substitution）$s = x + a$ 或许能令代数运算更简洁。

设 $s = x + a$，则 $y^2 = a^2 - x^2 = a^2 - (s-a)^2 = 2as - s^2$，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，试一试 规定要求吃掉任意片数之后都有 $V/A < 1$，即是对所有满足 $0 < t \le 2a$ 的 t 都成立。在这个范围内，$\frac{t}{2} - \frac{t^2}{6a}$ 的最大值是多少?

对 t 求导：$\frac{1}{2} - \frac{t}{3a} = 0$ 于 $t = \frac{3a}{2}$ 时成立，这个值位于范围 $(0, 2a]$ 之内，而且是极大值（maximum）（该式是 t 的一条开口向下的抛物线）。最大值为

$$\frac{3a}{4} - \frac{9a^2/4}{6a} = \frac{3a}{4} - \frac{3a}{8} = \frac{3a}{8}.$$

脆皮面包需要对所有 t 都有 $V/A < 1$，特别是在 $t = \frac{3a}{2}$ 时也要成立，所以 $\frac{3a}{8} < 1$，即 $a < \frac{8}{3} = 2\frac{2}{3}$ 米。 ■

阶段 4 题目之外：$2\frac{2}{3}$ 足够吗?

想一想 题目只要求找出半径必须满足的条件。$a < \frac{8}{3}$ 是否同时也足够? 如果从两端吃又如何?

并不足够。从两端吃会在中间剩下一条环带，而以同样多的面包皮计算，中间的环带比末端的球冠包含更多面包。对于环带 $-\frac{w}{2} \le x \le \frac{w}{2}$，

$$V = \pi \int_{-w/2}^{w/2} (a^2 - x^2)\,\mathrm{d}x = \pi\left(a^2 w - \frac{w^3}{12}\right), \qquad A = 2\pi a w, \qquad \frac{V}{A} = \frac{a}{2} - \frac{w^2}{24a},$$

当环带很窄时，这个值接近 $\frac{a}{2}$。例如，一个半径为 2.4 米的面包，明显小于 $2\frac{2}{3}$，切成四片后吃掉两端的两片，剩下的环带有 $w = a$ 及 $\frac{V}{A} = \frac{11a}{24} = 1.1$。事实上，对于任何一组薄片都有 $\frac{V}{A} < \frac{a}{2}$，所以如果可以从任何地方吃掉薄片，面包是脆皮的当且仅当 $a \le 2$ 米。题目中的 $2\frac{2}{3}$ 是一个正确的必要条件（necessary condition），但这项规定其实比题目所透露的更为严格。

回顾

- **先化简，后积分。**面包皮的被积函数（integrand）看似复杂，结果却是常数。动手积分之前，务必先把被积函数化简。
- **仔细阅读“任意”二字。**条件必须在吃掉任意片数之后都成立，所以我们需要的是 V/A 的最大值，而不是整个面包的值（即 $\frac{a}{3}$）。
- **简单的计算，出人意表的答案。**“面包皮相等”这个性质听起来难以置信，但计算只有两行。

背后的数学

阿基米德的帽盒定理

阿基米德在《论球与圆柱》（*On the Sphere and Cylinder*）中证明了：球面上夹在两个平行平面之间的环带，其面积等于这两个平面在包裹球的圆柱上所截出的环带的面积（图 4.8.1）。我们的计算正是这条帽盒定理（hat-box theorem）：半径为 a 的圆柱每单位长度的面包皮面积为 $2\pi a$，球也是如此。考虑整个球，它的表面积（surface area）等于圆柱的曲面面积，即 $2\pi a \times 2a = 4\pi a^2$。阿基米德认为，这个结果连同“球的体积是圆柱体积的三分之二”这一事实，是他毕生最出色的成就。西塞罗（Cicero）于公元前 75 年到访叙拉古（Syracuse）时，正是凭着寻找刻在墓上的球与圆柱，找到了那座杂草丛生的坟墓。

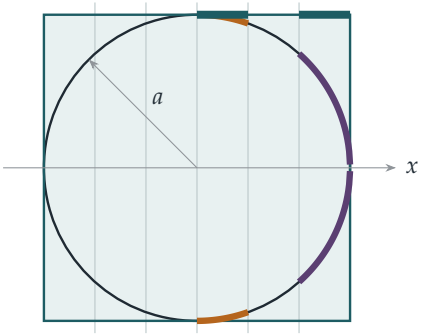


图 4.8.1　半径为 a 的球形面包置于包裹它的圆柱之内的横截面，面包被切成厚度相等的薄片。中间的薄片（琥珀色）的面包皮是一段短而接近水平的弧，离轴较远；末端的薄片（紫红色）的面包皮是一段长而陡峭的弧，离轴较近。绕 x-轴旋转后，两者的面包皮面积都恰好等于圆柱上对应环带（蓝绿色）的面积：$2\pi a$ 乘以厚度。

地球的地图

帽盒定理提供了一种绘制保持面积不变的地球地图的方法。把地球仪上的每一点沿水平方向向外投影到包裹它的圆柱上，然后把圆柱摊开。球面上的环带会映射到圆柱上面积相同的环带，而任何区域也是如此。这就是兰伯特（*Lambert*）圆柱等积投影（Lambert's cylindrical equal-area projection，1772 年）。它在两极附近会令形状严重失真，但面积是如实的，不像人们熟悉的麦卡托（Mercator）投影那样，令格陵兰看起来与非洲一样大（其实格陵兰大约只有非洲的十四分之一）。高斯（Gauss）有一条著名的定理，称为“绝妙定理”（*Theorema Egregium*，1827 年），它说明没有任何球面地图可以同时保持面积和

形状：每一幅平面的地球地图都必定有所失真。

球面上的随机点

以下是一个实际的应用。要在球面上均匀（uniformly）随机地选取一点，你或许会尝试均匀地选取纬度（latitude）和经度（longitude），但这样会令点在两极附近过于密集。帽盒定理给出了正确的做法：在 $-a$ 与 a 之间均匀地选取高度 x，并在 0 与 2π 之间均匀地选取经度。由于面积与高度成正比，这样选出的点恰好是均匀分布的。物理学和电脑图像的电脑模拟都采用这个方法。（值得一提的是，这个方法只在三维空间中有效：对于四维或更高维的球面，相应的命题并不成立。）

思维技巧

寻找隐藏的简洁

当计算得出出乎意料地简单的结果（例如一个常数）时，请停下来问一问为甚么。背后通常有几何上的原因，而找出这个原因，才是真正理解之所在。

在这里，被积函数 $y\sqrt{1+(y')^2}$ 是常数，因为在球面上，陡峭程度与离轴的距离恰好互相补偿：这就是阿基米德的定理。

延伸练习

1 **整个面包。**证明对于整个未切开的面包，$V/A = \frac{a}{3}$。为甚么题目中的条件比 $\frac{a}{3} < 1$ 更难满足？

2 **热带地区。**地球表面有多大比例位于南纬 23.4° 与北纬 23.4° 之间？（把地球视为一个球。由于面积与高度成正比，你只需要用到 $\sin 23.4°$。）

3 **其他形状。**证明对于抛物面（paraboloid）形状的面包（把 $y = \sqrt{x}$ 绕 x-轴旋转），每片的面包皮并非常数。对于哪些曲线 $y(x)$，$y\sqrt{1+(y')^2}$ 是常数？（解出相应的微分方程（differential equation）。）

4 **体积。**利用体积公式，证明球的体积是包裹它的圆柱体积的三分之二，正如阿基米德所发现的那样。

延伸阅读

- **Sherman Stein, *Archimedes: What Did He Do Besides Cry Eureka?***（Mathematical Association of America）。为现代读者讲解阿基米德的数学。
- **Timothy Gowers, *Mathematics: A Very Short Introduction***（Oxford University Press）。书中关于几何的一章介绍了曲面的几何。
- **John Snyder, *Flattening the Earth***（University of Chicago Press）。两千年来的地图投影史。

4.9 Holditch 定理

题目

STEP III 2011，第 5 题

继续阅读之前，请先看题目（原题为英文）。题目载于 STEP Support Programme 的试题资料库：

`step.maths.org/questions/11-s3-q5`

引子

随便取一个卵形，甚么形状都可以，只要它是凸（convex）的，然后让一根长度固定的杆在里面滑动一圈，两端始终留在曲线上。在杆上涂一个点，把杆分成长度为 a 和 b 的两段。杆绕行一周时，这个点会描出一个较小的卵形。两条曲线之间的环形区域，面积是多少？

答案令人惊讶：*无论原来的曲线是甚么形状，面积都是* πab。又长又窄的卵形、蛋形、圆角三角形：环形的面积永远相同，只取决于那个点在杆上的位置。这个结果由剑桥大学冈维尔与凯斯学院（Gonville and Caius College, Cambridge）的院士 Hamnet Holditch 牧师于 1858 年发现，是初等几何中最出人意表的定理之一。本题只用一条面积公式就把它证明出来。

动笔之前

想一想 先在最简单的情况下检验这个定理。若 $\mathcal{C}$ 是半径为 R 的圆，P 描出的是甚么曲线？环形的面积是多少？

由对称性可知，P 描出一个同心的圆，设其半径为 r。杆 AB 是大圆上一条通过 P 的弦，而*相交弦定理*（intersecting chords theorem，又称“圆幂”（power of a point））指出，对任何通过 P 的弦，都有 $PA \times PB = R^2 - r^2$。所以 $ab = R^2 - r^2$，环形的面积是 $\pi R^2 - \pi r^2 = \pi ab$。定理说的是：这个结论对每一条凸曲线都成立，不限于圆。

逐步解题

阶段 1　由坐标求面积

想一想　设 $x = r\cos\theta$ 和 $y = r\sin\theta$，全部都是 t 的函数，计算 $x\dot{y} - y\dot{x}$，其中上方的点表示 $\frac{\mathrm{d}}{\mathrm{d}t}$。

$\dot{x} = \dot{r}\cos\theta - r\sin\theta\,\dot{\theta}$，而 $\dot{y} = \dot{r}\sin\theta + r\cos\theta\,\dot{\theta}$。于是

$$x\dot{y} - y\dot{x} = r\cos\theta\,(\dot{r}\sin\theta + r\cos\theta\,\dot{\theta}) - r\sin\theta\,(\dot{r}\cos\theta - r\sin\theta\,\dot{\theta}) = r^2\dot{\theta}.$$

所以 $\frac{1}{2}\int r^2\,\mathrm{d}\theta = \frac{1}{2}\int r^2\dot{\theta}\,\mathrm{d}t = \frac{1}{2}\int (x\dot{y} - y\dot{x})\,\mathrm{d}t$。 ■

这条公式不需要极坐标（polar coordinates），只需要把笛卡儿坐标（cartesian coordinates）下的位置写成某个参数（parameter）的函数。当一点沿着一条包围 O 的闭曲线逆时针绕行一周时，公式给出的就是曲线内的面积。

阶段 2　杆的两端

杆由 A 指向 B 的方向是 $(\cos t, \sin t)$，而 P 与 A 的距离是 a，与 B 的距离是 b。所以

$$A = (x - a\cos t,\ y - a\sin t), \qquad B = (x + b\cos t,\ y + b\sin t).$$

阶段 3　OA 扫过的面积

想一想　把 $(*)$ 应用于 A。展开 $x_A\dot{y}_A - y_A\dot{x}_A$，并按 a 的幂次把各项整理好。

由 $x_A = x - a\cos t$、$y_A = y - a\sin t$，可得 $\dot{x}_A = \dot{x} + a\sin t$ 和 $\dot{y}_A = \dot{y} - a\cos t$。展开得

$$\begin{aligned} x_A\dot{y}_A - y_A\dot{x}_A &= (x - a\cos t)(\dot{y} - a\cos t) - (y - a\sin t)(\dot{x} + a\sin t) \\ &= (x\dot{y} - y\dot{x}) + a^2(\cos^2 t + \sin^2 t) - a[(x + \dot{y})\cos t + (y - \dot{x})\sin t]. \end{aligned}$$

由 0 积分至 2π 再除以二：第一项给出 $[P]$，第二项给出 $\frac{1}{2}a^2 \cdot 2\pi = \pi a^2$，第三项给出 $-af$。所以

$$[A] = [P] + \pi a^2 - af. \qquad ■$$

停一停，试一试　现在求 $[B]$。（提示：把 A 的公式中的 a 换成甚么，就得到 B?）

$B = (x + b\cos t, y + b\sin t)$ 就是把 A 的公式中的 a 换成 $-b$。所以

$$[B] = [P] + \pi b^2 + bf.$$

阶段 4 消去 f

想一想 A 和 B 都沿同一条曲线 $\mathcal{C}$ 绕行一周。由此可知 $[A]$ 和 $[B]$ 有甚么关系?

当 t 由 0 变到 2π 时，A 和 B 各自沿 $\mathcal{C}$ 逆时针绕行一周，所以 $[A] = [B] =$ 面积$(\mathcal{C})$。令两个表达式相等，

$$\pi a^2 - af = \pi b^2 + bf \quad\Longrightarrow\quad f(a+b) = \pi(a^2 - b^2) \quad\Longrightarrow\quad f = \pi(a-b).$$

由于 P 沿 $\mathcal{D}$ 绕行一周，$[P] =$ 面积$(\mathcal{D})$，于是

$$\text{面积}(\mathcal{C}) - \text{面积}(\mathcal{D}) = [A] - [P] = \pi a^2 - a\pi(a-b) = \pi ab. \qquad \blacksquare$$

那个取决于曲线形状的神秘量 f，自始至终没有计算过：它被消去了。

回顾

- **切合问题的公式。**面积公式 $(*)$ 只需要把坐标写成参数的函数，而杆的运动正是这样描述的。
- **两个表达式，一个未知量。**我们无法计算 f，但手上有两条含有它的方程，于是把它消去，就像解联立方程（simultaneous equations）一样。
- **先检验特殊情况。**动笔之前，圆的情况已经给出答案 πab，让我们知道目标是甚么。

背后的数学

格林定理与求积仪

公式 $(*)$ 是*格林定理*（Green's theorem）的特殊情况；格林定理把沿闭曲线的积分与曲线内区域上的积分联系起来。它的面积形式指出，曲线所包围的面积是 $\frac{1}{2}\oint(x\,\mathrm{d}y - y\,\mathrm{d}x)$，这个积分只需沿边界走一圈就能算出。工程师利用这一点造出了*求积仪*（planimeter），这是十九世纪发明的一种机械仪器（其中以 Jakob Amsler 在 1854 年的设计最为成功）：用指针沿着地图或图表上某个图形的边界描一圈，一个小轮就会记录所包围的面积。求积仪曾用来量度土地面积、根据示功图（indicator diagram）计算蒸汽机所作的功，还有许多其他用途，直至被数码方法取代为止。Holditch 定理与此密切相关：那根杆就像求积仪的臂。

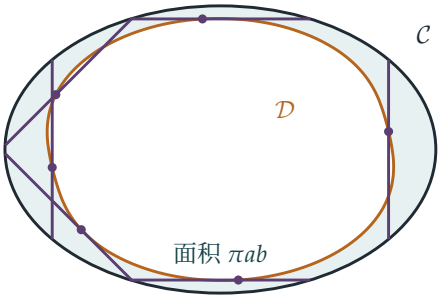


图 4.9.1　Holditch 定理的例子：椭圆的半轴（semi-axes）为 3 和 2，杆长 2.5，其中 $a = 1$、$b = 1.5$。图中画出了杆的若干位置（紫色），并标出点 P。P 所描出的曲线 $\mathcal{D}$（琥珀色）是用数值方法算出的，而两条曲线之间的阴影环形面积为 $4.7124\ldots = 1.5\pi$，与定理的预测完全吻合。

细节条款

Holditch 陈述这个定理时，对条件不大讲究，但定理其实需要一些条件。杆必须能够在两端都留在曲线上的情况下绕足一整圈，而描出的曲线 $\mathcal{D}$ 必须是只描一次的简单闭曲线（simple closed curve）；如果杆相对于 $\mathcal{C}$ 的阔度太长，$\mathcal{D}$ 可能会自相交，这时公式给出的“面积”会把某些区域计算两次，或以负号计算。题目假设 $\mathcal{D}$ 是一条包含 O 的凸曲线，从而避开了这些困难。找出定理究竟在甚么条件下成立，正好说明一个优美的想法必须经过严格化才站得住脚，这也是第 3 章的主题。

思维技巧

算不出来的，就把它消去

如果某个量太复杂而无法计算，就找两条含有它的不同方程，把它消去。

这里 f 取决于曲线的每一个细节，但它同时出现在 $[A]$ 和 $[B]$ 的表达式中，而 $[A] = [B]$ 这个事实把它消除了。答案 πab 之所以能够与曲线无关，全因取决于曲线的部分互相抵消了。

延伸练习

1 **椭圆的面积。**利用 $(*)$，取 $x = p\cos t$、$y = q\sin t$，证明半轴为 p 和 q 的椭圆面积为 πpq。

2 **圆的情况。**证明相交弦定理：若圆的两条弦相交于 P，则每条弦被分成的两段，其长度之积相等。由此推导出圆的 Holditch 定理。

3 **中点。**证明由 $f = \pi(a - b)$ 可知，当 P 是杆的中点时，$f = 0$。这时环形的面积是多少？答案以杆长 L 表示。

4 **三角形。**较难利用 $(*)$，沿边界绕行一周，求顶点为 $(0,0)$、(x_1, y_1)、(x_2, y_2) 的三角形的面积，并证明它等于 $\frac{1}{2}|x_1 y_2 - x_2 y_1|$。这就是求任意多边形面积的“鞋带公式”（shoelace formula）的起点。

延伸阅读

- **Mark Levi, *The Mathematical Mechanic***（Princeton University Press）。以杆、轮和弹簧等物理推理来证明数学结果，与本节的精神一脉相承。
- **H. M. Schey, *Div, Grad, Curl, and All That***（Norton）。向量微积分（vector calculus）的非正式入门，内容包括格林定理。
- **Hamnet Holditch, “Geometrical theorem”**，*Quarterly Journal of Pure and Applied Mathematics* 2 (1858)。原始的短文，只有短短一段，完全没有提及任何细节条款。

我们学到了甚么

本章开始时，我们提出了三个问题：积分（integral）究竟是甚么？积分可以证明甚么？当一个积分没有公式时，又该怎么办？以下是本章九道题目给出的答案。

1. 积分是甚么？哪些函数有积分？

在 Section 4.1 中，我们把积分定义为由窄长矩形组成的下和（lower sum）与上和（upper sum）的共同极限，并单凭这个定义计算出 $\int_0^a x^5\,\mathrm{d}x$。这要花一番工夫：先以裂项相消求和（telescoping sum）处理连续整数乘积之和，再从两边估计上下界，最后用夹逼（squeeze）得出结果。花了多少力气，正好说明微积分基本定理（fundamental theorem of calculus）让你不劳而获的有多少。

这个定义也告诉我们，并非每个函数都可积（integrable）。狄利克雷（Dirichlet）函数在有理数上取值 1，在无理数上取值 0，它的上和全部等于 1，下和全部等于 0。正是这类问题，促使勒贝格（Lebesgue）建立了一套更强大的积分理论，你日后在大学便会学到。

2. 积分可以证明甚么？

远远不止面积。

- **数与数之间的不等式。**一个正函数的积分等于 $\frac{22}{7}-\pi$，所以 $\frac{22}{7}>\pi$（Section 4.3）。“把 $A-B$ 写成某个正函数的积分”这个想法既简单，又非常有用。
- **无穷乘积与特殊值。**利用 $\int \sin^n t\,\mathrm{d}t$ 的归约公式（reduction formula）、一个不变量（invariant）和一次夹逼，便得到沃利斯（Wallis）乘积、斯特灵（Stirling）公式中的常数，以及 $\left(\frac{1}{2}\right)!=\frac{\sqrt{\pi}}{2}$（Section 4.4）。
- **无穷和。**把 x^x 展开成幂级数（power series）再逐项积分，便得到“二年级生之梦”（sophomore's dream）（Section 4.5）。
- **三角学的基础。**定义 $\mathrm{T}(x)=\int_0^x \frac{\mathrm{d}u}{1+u^2}$ 及 $\pi=2\mathrm{T}(\infty)$ 之后，单靠代换（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）。
- **几何。**一个化简后变成常数的被积函数（integrand），揭示了阿基米德（Archimedes）的帽盒定理（hat-box theorem）（Section 4.8）；而把面积公式写成沿边界的积分，则证明了令人惊叹的 Holditch 定理（Section 4.9）。

- **最优函数。**在积分内配方（completing the square），便准确指出 $\int\left((y')^2 - a^2y^2\right)$ 在甚么情况下必定非负，而临界值 $\frac{\pi}{2}$ 原来是一个特征值（eigenvalue）（Section 4.7）。

3. 没有公式时，我们可以怎么办？

在 Section 4.6 中，我们证明了 $\int \frac{e^x}{x+1}\,\mathrm{d}x$ 不能写成有理函数（rational function）乘以 e^x 的形式，而刘维尔（Liouville）的理论更表明它根本没有初等（elementary）公式。从极点（pole）的角度看，原因是其原函数（antiderivative）在 $x = -1$ 附近的表现必须像对数函数，而有理函数做不到这一点。然而这个积分确确实实存在，我们也可以把它算出来：可以用数值方法，可以用级数（如 Section 4.5 中处理 $\int x^x$ 那样），也可以给它起一个名字，把它当作一个新函数来研究，就像我们在 Section 3.9 中对待 W 函数那样。

4. 积分工具箱

方法	你在哪里用过
从定义出发	由矩形求 $\int_0^a x^5$（4.1）
裂项相消求和	连续整数乘积之和（4.1）
保持积分不变的代换	由 $\int \frac{\mathrm{d}u}{1+u^2}$ 建立三角学（4.2）
多项式除法	$\frac{22}{7} - \pi$（4.3）
正性与被积函数的估界	$\pi < \frac{22}{7}$（4.3）
归约公式	$\int \sin^n t$（4.4）；$\int x^m(\ln x)^n$（4.5）
不变量与夹逼	$nu_nu_{n-1} = \frac{\pi}{2}$（4.4）
逐项积分	二年级生之梦（4.5）
因式与次数论证	不存在有理原函数（4.6）
配方法	变分法（calculus of variations）（4.7）
先化简，后积分	球形面包（4.8）
以边界积分表示面积	Holditch 定理（4.9）

下一站

- Section 4.4 中的伽玛函数（Gamma function）和贝塔函数（Beta function），以及斯特灵公式中的常数 $\sqrt{2\pi}$，会在 Section 12.23 的概率问题中再度出现。
- Section 4.7 的特征值问题，到了 Section 8.6 会变成量子谐振子（quantum harmonic oscillator）。
- 梅钦（Machin）公式由 Section 4.2 的加法恒等式建立而成，是 Section 5.1 的主题。
- 第 5 章转向复数（complex numbers），它解释了为何本章有这么多积分都与 π 有关。

如果你想在上大学前走得更远

- **Paul Nahin, *Inside Interesting Integrals***（Springer）：一场巧妙积分的盛宴，并道出背后的种种窍门。
- **William Dunham, *The Calculus Gallery***：从牛顿（Newton）到勒贝格的分析学杰作。
- **Michael Spivak, *Calculus***：严谨地讲述积分的定义和微积分基本定理。
- **Sherman Stein, *Archimedes***：追溯积分的古代源头。

埃尔米特（Hermite）对没有导数的函数避之则吉，但积分从来没有这个问题：它十分宽容，连微分无法处理的函数也照单全收，而且不断带来惊喜。如果说有一个主题贯穿本章，那就是：积分不只是一种求面积的技巧。它是一种语言，用来表达和证明关于数、函数和图形的事实，而在你往后的数学生涯中，你都会一直说这种语言。

延伸练习答案

以下只是简短的答案和提示，并非完整解答：只列出结果和关键步骤，方便你核对自己的做法。请先认真做过每道练习，然后才看答案。

4.1 甚么是积分？

1. 这个和是 $y=\mathrm{f}(x)$ 之下、在 $[1,2]$ 上 n 个阔度为 $\frac{1}{n}$ 的长方形的总面积。取 $\mathrm{f}(x)=\frac{1}{x}$，得 $\frac{1}{n}\sum\frac{1}{1+m/n}=\sum\frac{1}{n+m}\to\ln 2$。第二个和是 $\frac{1}{n}\sum\frac{1}{1+(m/n)^2}\to\int_0^1\frac{\mathrm{d}x}{1+x^2}=\frac{\pi}{4}$。

2. 分成 n 条时，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）则只须把第一个因子中的 $n+1$ 换成 $n-1$；两者都趋向 $\frac{a^3}{3}$。

3. 各项可以裂项相消（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 由一个积分建立三角学

1. 导数（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。至于余弦，固定 y，设 $p(x)=\mathrm{c}(x+y)-\mathrm{c}(x)\mathrm{c}(y)+\mathrm{s}(x)\mathrm{s}(y)$ 及 $q(x)=\mathrm{s}(x+y)-\mathrm{s}(x)\mathrm{c}(y)-\mathrm{c}(x)\mathrm{s}(y)$。于是 $p'=-q$ 而 $q'=p$，所以 p^2+q^2 是常数，而它在 $x=0$ 时等于 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 $\frac{\mathrm{d}y}{\sqrt{1-y^2}}=\frac{\mathrm{d}u}{1+u^2}$。于是 $\mathrm{S}(\frac{1}{\sqrt{2}})=\mathrm{T}(1)$，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 $x=\frac{1}{3}$，则 $\frac{x+a}{1-ax}=1$，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 $[1,x]$ 上则 $\le\frac{1}{u^2}$，所以当 $x\ge 1$ 时，$\mathrm{T}(x)\le 1+(1-\frac{1}{x})<2$；当 $0<x<1$ 时，$\mathrm{T}(x)<\mathrm{T}(1)\le 1$。（上界 $1+\int_1^x u^{-2}\,\mathrm{d}u$ 在 $x<1$ 时并不成立：在 $x=\frac{1}{2}$ 时它等于 0。）T 递增而且有界（bounded），所以它有极限（limit）。

4.3 $22/7$ 大于 π

1. $\frac{22}{7}-\pi\approx 0.001264$，介乎 $\frac{1}{1260}\approx 0.000794$ 与 $\frac{1}{630}\approx 0.001587$ 之间。

2. $\frac{(4!)^2}{9!} = \frac{576}{362\,880} = \frac{1}{630}$。

3. $\frac{x(1-x)^2}{1+x^2} = x - 2 + \frac{2}{1+x^2}$，积分（integrate）后得 $\frac{1}{2} - 2 + \frac{\pi}{2}$。同样地，$\frac{x^2(1-x)^4}{1+x^2} = x^4 - 4x^3 + 5x^2 - 4 + \frac{4}{1+x^2}$，积分后得 $\frac{1}{5} - 1 + \frac{5}{3} - 4 + \pi = \pi - \frac{47}{15}$。两个积分都是正数。这两个被积函数（integrand）远大于 $\frac{x^4(1-x)^4}{1+x^2}$（其最大值约为 0.003），所以误差 $\frac{\pi}{2} - \frac{3}{2} \approx 0.07$ 和 $\pi - \frac{47}{15} \approx 0.008$ 也较大。

4. 多项式除法（polynomial division）余下 $\frac{16}{1+x^2}$，积分后得 4π；整个积分是 $4\pi - \frac{188684}{15015} \approx 3.6 \times 10^{-6}$，所以 $\pi \approx 3.14159174$。

4.4 沃利斯、Beta 与二分之一的阶乘

1. $u_{2m} = \frac{1\cdot3\cdots(2m-1)}{2\cdot4\cdots2m} \cdot \frac{\pi}{2}$ 而 $u_{2m+1} = \frac{2\cdot4\cdots2m}{3\cdot5\cdots(2m+1)}$。于是 $1 < \frac{u_{2m}}{u_{2m+1}} < \frac{u_{2m-1}}{u_{2m+1}} = \frac{2m+1}{2m}$，所以这个比趋向 1，整理后即得 $\frac{\pi}{2} = \prod_{k\geq1} \frac{(2k)^2}{(2k-1)(2k+1)}$。

2. 分部积分法（integration by parts）给出 $I_n = \frac{n}{n+1} \int_0^1 x^{n+1}(1-x)^{n-1}\,\mathrm{d}x$，再写成 $x^{n+1}(1-x)^{n-1} = x^n(1-x)^{n-1} - x^n(1-x)^n$ 并利用对称性，便得 $I_n = \frac{n}{2(2n+1)} I_{n-1}$；用归纳法（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 $\left(\left(\frac{1}{2}\right)!\right)^2/2! = \frac{\pi}{8}$，所以 $\left(\frac{1}{2}\right)! = \frac{\sqrt{\pi}}{2}$（即 $\Gamma\left(\frac{3}{2}\right)$）。

3. 对称性来自 $x \mapsto 1-x$；递推关系（recurrence）来自 $x^p = x^{p-1} - x^{p-1}(1-x)$；而取 $x = \frac{t}{1+t}$ 时，$1-x = \frac{1}{1+t}$ 且 $\mathrm{d}x = \frac{\mathrm{d}t}{(1+t)^2}$。该积分是 $\mathrm{b}\left(\frac{5}{2}, \frac{7}{2}\right)$；利用递推关系和对称性化简至 $\mathrm{b}\left(\frac{1}{2}, \frac{1}{2}\right) = \pi$，便得 $\frac{3\pi}{256}$。

4. 由第 (i) 部分的显式数值，$u_{2m} = \frac{(2m)!}{(2^m m!)^2} \cdot \frac{\pi}{2} = \binom{2m}{m} \frac{\pi}{2^{2m+1}}$。于是由 $2mu_{2m}^2 \to \frac{\pi}{2}$ 得 $\binom{2m}{m}^2 \approx \frac{4^{2m}}{\pi m}$。当 $m = 10$ 时：$\binom{20}{10} = 184\,756$，而 $\frac{4^{10}}{\sqrt{10\pi}} \approx 187\,079$（高出约 1.3%）。

4.5 大二学生之梦

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 $x = e^{-u}$）。所以该积分是 $\sum_{n\geq1} n^{-n} \approx 1.2913$。

2. $\frac{\mathrm{d}}{\mathrm{d}x}x^x = x^x(\ln x + 1)$，在 $x = \frac{1}{e}$ 时为零；其值为 $(e^{-1})^{e^{-1}} = e^{-1/e} \approx 0.692$。

3. 各项是 $\frac{1}{n^n}$。取 7 项后停止，误差小于 $\frac{1}{8^8} \approx 6 \times 10^{-8}$，已经足够；只取 6 项的话，上界 $\frac{1}{7^7} \approx 1.2 \times 10^{-6}$ 并不足够。所以要取 7 项。

4. 用分部积分法：$\int_0^1 x^2 \ln x\,\mathrm{d}x = \left[\frac{x^3\ln x}{3}\right]_0^1 - \int_0^1 \frac{x^2}{3}\,\mathrm{d}x = -\frac{1}{9}$。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 没有公式的积分

1. 求导：$\frac{\mathrm{d}}{\mathrm{d}x}\frac{e^x}{x+1}=\frac{xe^x}{(x+1)^2}$，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 $\mathrm{P}'-2x\mathrm{P}$ 的次数是 $\deg\mathrm{P}+1\geq 1$（当 $\mathrm{P}\neq 0$），所以它不可能等于 1。

3. 假如 $\mathrm{R}=\mathrm{P}/\mathrm{Q}$（已约至最简）在 0 处有 $k\geq 1$ 阶的极点（pole），则 $\mathrm{R}'+\mathrm{R}$ 在该处会有 $k+1\geq 2$ 阶的极点，但 $\frac{1}{x}$ 的极点只是 1 阶。假如 R 在 0 处没有极点，则 $\mathrm{R}'+\mathrm{R}$ 也没有。无论哪种情况，都有 $(\mathrm{R}e^x)'\neq\frac{e^x}{x}$。

4. 对 $\frac{\mathrm{P}e^x}{x+1}$ 求导，得 $e^x\frac{(x+1)\mathrm{P}'+x\mathrm{P}}{(x+1)^2}$。在 $x=-1$ 处：$\mathrm{f}(-1)=-\mathrm{P}(-1)$，而由 $\mathrm{f}'=(x+1)\mathrm{P}''+(x+1)\mathrm{P}'+\mathrm{P}$ 得 $\mathrm{f}'(-1)=\mathrm{P}(-1)$。对于 x^3-2：$-3+3=0$；对于 x：$-1+1=0$；对于 x^2+1：$2-2=0$。这个条件也是充分的：映射 $\mathrm{P}\mapsto(x+1)\mathrm{P}'+x\mathrm{P}$ 把次数小于 d 的多项式映到次数不超过 d 的多项式，而且是单射（injective），所以它的像（image）的维数（dimension）是 d，并且包含在由满足 $\mathrm{f}(-1)+\mathrm{f}'(-1)=0$ 的 f 所组成的 d 维空间之内；因此像就是整个空间。

4.7 变分法

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. $J(y+\varepsilon\eta)=J(y)+2\varepsilon\int_0^1(y'\eta'-a^2y\eta)+\varepsilon^2J(\eta)$；要在 $\varepsilon=0$ 处取得最小值，线性项必须为零。分部积分后，对每个满足 $\eta(1)=0$ 的 η 都有 $-y'(0)\eta(0)-\int_0^1(y''+a^2y)\eta=0$，由此必有 $y''+a^2y=0$ 及 $y'(0)=0$。

3. 取 $x=Lu$，得 $\int_0^L\left((y')^2-a^2y^2\right)\mathrm{d}x=\frac{1}{L}\int_0^1\left((Y')^2-(aL)^2Y^2\right)\mathrm{d}u$，所以条件是 $aL\leq\frac{\pi}{2}$。（当 $aL=\frac{\pi}{2}$ 时，正如正文所述，积分仍然不会是负数，但 $y=\cos\frac{\pi x}{2L}$ 会令它等于零；当 $aL>\frac{\pi}{2}$ 时，同一个函数会令它变成负数。）

4. 摆线（cycloid）是 $x=R(\theta-\sin\theta),y=R(1-\cos\theta)$。关于等时曲线（tautochrone），见 Section 11.4：以弧长（arc length）表示时，该运动恰好是简谐运动（simple harmonic motion）。

4.8 切球形面包

1. $V/A=\frac{\frac{4}{3}\pi a^3}{4\pi a^2}=\frac{a}{3}$。但余下一块高度为 t 的部分有 $A=2\pi at$ 及 $V=\frac{1}{3}\pi t^2(3a-t)$，所以 $V/A=\frac{t(3a-t)}{6a}$，它在 $t=\frac{3}{2}a$ 时取得最大值 $\frac{3a}{8}$：比 $\frac{a}{3}$ 还要大。

2. 这条带的高度是 $2R\sin 23.4^\circ$，而球的高度是 $2R$，所以所占比例是 $\sin 23.4^\circ\approx 0.40$。

3. 对于 $y=\sqrt{x}$，$y\sqrt{1+(y')^2}=\sqrt{x+\frac{1}{4}}$，并非常数。由 $y\sqrt{1+(y')^2}=c$ 得 $(y')^2=\frac{c^2}{y^2}-1$，其解为 $y=c$（圆柱）以及圆 $(x-x_0)^2+y^2=c^2$（球）。

4. 球的体积是 $\frac{4}{3}\pi a^3$，圆柱的体积是 $\pi a^2 \cdot 2a = 2\pi a^3$；两者之比是 $\frac{2}{3}$。

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. 把两条弦（chord）的端点连起来所得的两个三角形相似（同弧所对的角相等），所以 $PA \cdot PB = PC \cdot PD$。对于半径为 R 的圆，若 P 与圆心的距离为 r，则通过 P 的直径给出 $(R-r)(R+r)$，所以 $ab = R^2 - r^2$，而圆环的面积是 $\pi R^2 - \pi r^2 = \pi ab$。
3. 若 $a = b$，则 $f = \pi(a-b) = 0$，而圆环的面积是 $\pi a^2 = \frac{\pi L^2}{4}$：即以该杆为直径的圆的面积。
4. 沿着通过 O 的两条边，$x\,\mathrm{d}y - y\,\mathrm{d}x = 0$。第三条边以 $(x_1, y_1) + s(x_2 - x_1, y_2 - y_1)$（其中 $0 \le s \le 1$）作参数表示，在这条边上它等于 $(x_1 y_2 - x_2 y_1)\,\mathrm{d}s$。所以面积是 $\frac{1}{2}|x_1 y_2 - x_2 y_1|$。

Beyond STEP

STEP 是大多数中学生一生中应考过最艰深的数学考试，也是最出色的考试之一：题目简短、陌生而深刻，当中最精彩的，其实是乔装了的大学数学小品。

本书精选最有趣的 STEP 试题，由 1987 年首次开考之前的样本试卷（specimen papers），到 2025 年的试卷，并以每一道题目作为一扇门。每道题目都会仔细解答，沿途提出优秀学生会问的问题，然后带你穿过这扇门：这个想法从何而来，到了大学会演变成甚么，以及接下来该读些甚么。

本书为正在准备 STEP、有天分又有志投身数学的学生而写。他们想要的不止是模拟试卷：他们想开始像数学系本科生那样思考，并一探前方有甚么正等着自己。

- 共 13 章，涵盖纯数学（pure mathematics）、力学（mechanics）、概率（probability）与统计（statistics）
- 136 道 STEP 历届试题，每题均附引导式解答
- 每道题目背后的大学数学
- 附答案的延伸练习，以及延伸阅读

全书共 13 章，本册为第 4 章。 全书各章及所有详细解答均可免费阅读、下载及分享，网址为

beyondstep.keithmcnulty.org

文字内容已贡献予公有领域（CC0 1.0）。STEP 试题 © Cambridge University Press & Assessment。

