從祖暅原理談起

柳柏濂

一. 避開無限性 — 祖暅原理

略知一點中國數學史的人都知道, 中國古代數學的一大貢獻是祖暅原理。

祖暅(公元五到六世紀), 中國齊梁時代數學家祖沖之 (公元429-500年) 的兒子。父子兩代數學家, 父親以推算圓周率享譽世界, 而兒子, 又以“開立圓術” (由球體積求直徑的方法) 中的祖暅原理而名留青史。

祖暅原理曰:“緣冪勢既同, 則積不容異”, 意即兩同高的立體, 如果在等高處的截面積相同, 則體積亦相等, 如下圖

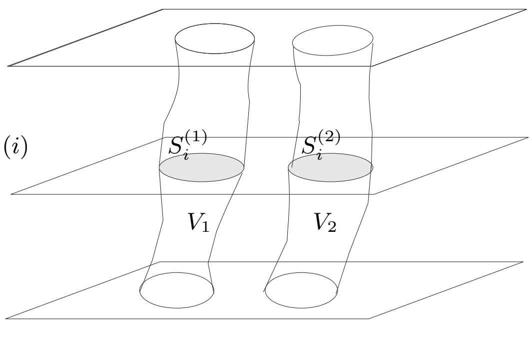

圖 1. 對任意 i, $S_i^{(1)} = S_i^{(2)} \Rightarrow V_1 = V_2$

此乃西方稱謂的卡瓦列利 (Cavalieri) 原理。 意大利人卡瓦列利 (Bonaventura Cavalieri)(1598-1647) 在1635年發現此原理, 而在此以前的一千一百多年, 祖暅已在「綴術」一書中明確提出了這一原理, 並用於準確地計算球的體積。

祖暅原理的一個直觀的表達是: 用一疊相同面值的硬幣, 拼齊豎直放在桌面上得立方體 V_1(圖2), 然後, 把這疊硬幣向一個方向挪動, 得立方體 V_2(圖3)

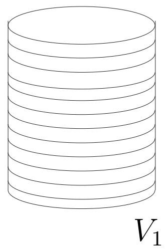

圖 2

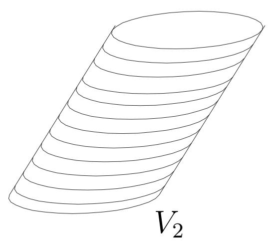

圖 3

注意到祖暅原理, 斜圓柱V_2的體積與直圓柱V_1的體積相等。

我們無從考證祖暅如何得出這一原理。「綴術」已失傳, 僅從唐朝李淳風的「九章算

術注」中得知: 祖暅原理敍述前, 有一句話“失疊棋成之積”, 意思是, 體積可看作無限小平面疊加而成, 此乃無限小分析的思想是也。

用現代的數學語言演繹, 即兩個立方體 V_1, V_2 (下面, 不妨把立方體與它的體積用同一字母表之): V_1 和 V_2 在等高處的截面積分別是 $S_i^{(1)}$ 和 $S_i^{(2)}$, $i=1,2,\ldots,k,\ldots$,

$$S_i^{(1)}=S_i^{(2)}$$

考察由面積爲 $S_i^{(1)}(S_i^{(2)})(i=1,2,\ldots,k,\ldots)$ 的面爲底, “小”高度爲 Δx_i 的柱體

則 $\sum\limits_{i=1}^{k} S_i^{(1)}\Delta x_i=\sum\limits_{i=1}^{k} S_i^{(2)}\Delta x_i$

故 $\lim\limits_{k\to\infty}\sum\limits_{i=1}^{k} S_i^{(1)}\Delta x_i=\lim\limits_{k\to\infty}\sum\limits_{i=1}^{k} S_i^{(2)}\Delta x_i$, 便得 $V_1=V_2$。

祖暅原理不涉及面積的個數問題, 因而避開了無限性的干擾。當我們承認了連續公理及用一些微積分的知識, 祖暅原理可以被證明爲定理。

若把高爲 h 的立方體 V_1, V_2 的截面積分別表達爲變量 (高) z 的函數 $S_1(z), S_2(z)$, 祖暅原理的現代化表述爲: 對任何 z, $0\le z\le h$ 若 $S_1(z)=S_2(z)$ 則

$$V_1=\int_0^h S_1(z)dz=\int_0^h S_2(z)dz=V_2$$

二. 減少一維 — 祖暅面積原理

把祖暅原理從空間體積轉向平面的面積, 我們考察平面上的祖暅原理。

做爲祖暅原理的直接類比, 可以有:

(2.1) 夾在兩平行直線間的兩個平面被平行這兩條直線的任意一條直線所截, 如果截得的兩條線段的長度總相等, 那麼, 這兩個平面圖形的面積也相等。

然而, 我們可以把這一原理再放寬一些。容易知道, 下面的原理也是正確的。並且是上述原理 (2.1) 的推廣。

(2.2) 夾在兩平行直線間的兩個平面圖形被平行這兩條直線的任意一條直線所截, 如果截得的兩條線段的長度之比恒爲一個常數, 那麼, 這兩個平面圖形的面積之比等於這個常數。

作爲一個應用, 可以用祖暅面積原理推導橢圓面積公式。

考察在直角座標系中, 橢圓 $\frac{x^2}{a^2}+\frac{y^2}{b^2}=1$ $(a>b>0)$ 和圓 $x^2+y^2=a^2$ 。記此橢圓和圓面積分別是 $S_{橢}, S_{圓}$。

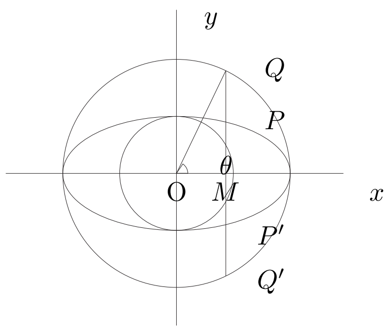

圖 4

做任一垂直於 X 軸的直線 $x=a\cos\theta$ 交 X 軸於 M, 交橢圓於 P, P', 交大圓於 Q, Q', 則 P, P', Q, Q' 的座標分別是

$(a\cos\theta, b\sin\theta), (a\cos\theta, -b\sin\theta)$

$(a\cos\theta, a\sin\theta), (a\cos\theta, -b\sin\theta)$。

故

$$\frac{|PM|}{|QM|}=\frac{|PP'|}{|QQ'|}=\frac{b}{a}\quad。$$

由原理(2.2)

$$\frac{S_{橢}}{S_{圓}}=\frac{b}{a}$$

便得 $S_{橢}=\frac{b}{a}\cdot\pi a^2=\pi ab$

三. 往下再走一步 — 天下大亂?

由體積到面積, 從三維到二維, 步步類比, 引來學子躍躍欲試, 在一次講座之餘, 一個聽者問道: “我試做一維祖暅原理, 可爲何結果令人迷惘?”

他順手畫出兩條平行直線 l_1, l_2 及夾在 l_1, l_2 間的兩條不相等長的線段 a, b ($a>b$, 這裡也把線段記號和長度用同一個字母表示), 用任一條與 l_1 (也與l_2) 平行的直線 l 去截 a, b(如圖5), 顯然, 其截面的長 (或面積) 相等 (均爲0)。

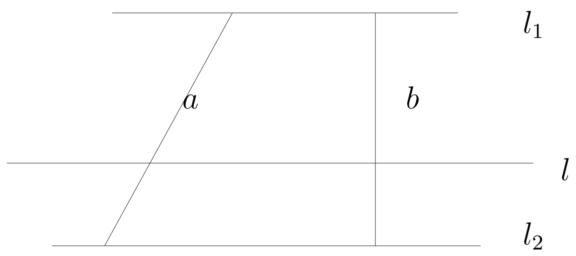

圖 5

按一維祖暅原理, 便得 $a=b$, 矛盾!

乍看起來, 這的確是一個“危險”的推理, 按此辦理, 豈不是可以證明: 世間任兩條線段長均相等嗎? 天下必然大亂了!

且慢, 稍爲氣定神閑, 必可看出荒謬所在。

事實上, 圖5的“所作所爲”亦是祖暅面積原理的特款。易見, 線段 a 的面積是0, 線段 b 的面積也是0, 兩者面積確是相等, 又何謬之有呢?

如果要考察長度, 就必然要觸及到另一個問題: 設 l_1, l_2 的距離爲 h, 用與 l_2 距離爲 i 的直線 l_i $(0\le i\le h)$ 截a,b所得的“截面積”分別記爲$S_i^{(a)}, S_i^{(b)}$ (實質上, 應是 a,b 上的“小”長度), 我們面臨要回答的問題是: 當 $S_i^{(a)}=S_i^{(b)}=0$, $0\le i\le h$ (i 取遍 $[0,h]$ 中的實數),

$$\sum_{0\le i\le h} S_i^{(a)}=\sum_{1\le i\le h} S_i^{(b)} \qquad (3.1)$$

成立嗎?

圖5的例子告訴我們:(3.1) 式是不一定能成立的。

這位學子雖被無限性引入岐路, 他卻向我們提出一個有趣的問題: 無限多個長度爲0的集, 合起來有多長?

四. 無限多段長為零的線合起來有多長?

“0”這個東西, 在數學上是再簡單不過了。可是, 一旦把它放到無限數學中去, 有時連大數學家也受不了。十七世紀, 牛頓創立微積分的初期, 由於沒有嚴謹的分析理論基礎, 無窮小被英國大主教貝克萊譏爲“逝去的鬼魂” — 一會非零一會又是零, 牛頓也無言以對。

要回答本節的問題, 必須弄清長度這個最基本的概念。正因爲它“基本”, 因而要用公理的語言來描述。爲更直觀起見, 在本文中, 我們僅以數直線 (數軸) 上的集來討論問題。不應把“無限個長度爲零的線段” 僅僅理解爲首尾相接在一起的線段, 我們可以把每一線段用區間 [a,b]表示。

把直線 l 上有限區間的集記爲 P, 在 P 上定義一個非負集合函數, $\mu(E), E\epsilon P$ 它滿足下面三條公理:

(4.1) 正則性: 對 $E=[a,b]\epsilon P,(a\leq b$ 則 $\mu(E)=b-a$。

(4.2) 運動不變性: 若 $E_1,E_2\epsilon P$ 且 $E_1=E_2$, 則 $\mu(E_1)=\mu(E_2)$。

(4.3) 有限可加性: 若 $E_1,E_2\ldots,\ E_n$ 是 P 中元素, 兩兩不相重疊, 則 $\mu(\cup_{i=1}^n E_i)=\sum_{i=1}^n\mu(E_i)$。

$\mu(E)$ 稱爲 E 的長度, 也稱爲 Jordan 測度, 這裡簡稱爲 J 測度。

上面三條公理“譯”成通俗的語言, 便是 (1) 線段有長;(2) 線段的長度在運動中保持不變; (3) 有限條線段的總長度是它們各長度的和。這顯然是我們日常中所接觸到的長度概念是一致的。當然, 這也並非不言而喻的。愛因斯坦的相對論告訴我們: 在高速運動中, 長度並沒有不變性, 因此, 嚴格來說, 公理 (4.1)–(4.3) 乃是一種數學上的約定。按照上述長度 (J 測度) 的定義, 對於一個點 $\{a\}$, 可以看作[a,a], 由 (4.1), 它的長度是

$$\mu(\{a\})=a-a=0\ 。$$

如果我們把一個區間 $[a,b]$ 去掉一個點 (不妨端點), 則由有限可加性 ((4,3))

$$\begin{aligned}\mu([a,b]) &= \mu(\{a\}\cup(a,b])\\ &= \mu(\{a\})+\mu((a,b])\\ &= \mu((a,b])\ 。\end{aligned}$$

同理可知

$$\begin{aligned}\mu((a,b]) &= \mu([a,b))=\mu((a,b))\\ &= \mu([a,b])=b-a\end{aligned}$$

這表明, 一個區間的端點的去留, 並不改變區間的長度。

容易看出, P 對有限個元素的“並”運算和兩個元素的“差”運算是封閉的, P 中的元素, 總可以表示爲有限個點和有限多個非零區間之和。(非零區間指長度非零的區間)

上面多次提到有限, 讓我們試圖觸摸一下“無限”。

考察一個非零區間 $[a,b]$ 所表示的線段, 顯然, 它是由點所構成, 用數學語言表達爲 $[a,b]=\cup_{a\leq x\leq b}\{x\}$, 既然, 每點的長度 $\mu(\{x\})=0$, 於是

$$\mu(\cup_{a\leq x\leq b}\{x\})=\sum_{a\leq x\leq b}\mu(\{x\}) \qquad (4.4)$$

便有 $\mu([a,b]=0$, 但由正則性 $\mu([a,b])=b-a\neq 0$, 矛盾!

這一矛盾與上面第三節的謬誤如出一轍, 它的根源在於在 (4.4) 式中, 用“無限可加性” 偷換了公理 (4.3) 的“有限可加性”。

因此, 一般來說, J 測度的無限可加性是不成立的。無限多個長度爲0的集 $\{x\}$, 可以並成一個長度爲定值的集, 這個定值可以是0或非零。

坦率一句, 用 J 測度這把"尺子", 我們不能量度在數直線上無限個點並起來的集的長度。

五. 換一把尺子—突破有限的束縛

J 測度的有限可加性公理, 來自於"總長度等於各分長度之和"這個直觀意念。那麼, 如果我們把這個分長度的個數擴充成可列個—即能與自然數的集一一對應的一列無限的長度 。也就是把公理 (4.3) 的有限可加性擴充為下列的可列可加性。

$$\mu(\cup_{i=1}^{\infty}E_i)=\sum_{i=1}^{\infty}\mu(E_i) \tag{5.1}$$

(請區分 (5.1) 式與 (4.4) 的不同)

擴充後的可加概念, 仍然會得到合理的解釋嗎? 回答是肯定的。請看線段 (0,1] 由長度的正則性 ((4.1))

$$\mu((0,1])=1-0=1 \tag{5.2}$$

又 $(0,1]=\cup_{i=1}^{\infty}(\frac{1}{i+1},\frac{1}{i}]$

用可列可加性,

$$\mu(\cup_{i=1}^{\infty}(\frac{1}{i+1},\frac{1}{i}])=\sum_{i=1}^{\infty}(\frac{1}{i}-\frac{1}{i+1})=1 \tag{5.3}$$

(5.2) 和 (5.3) 的結果是一致的。

我們又考察一列點 $1,\frac{1}{2},\frac{1}{2^2},\ldots,\frac{1}{2^n},\ldots$ 並起來的集合 $\cup_{n=1}^{\infty}\{\frac{1}{2^n}\}$ 按可列可加性, 可以得到它的長度

$$\mu(\cup_{n=1}^{\infty}\{\frac{1}{2^n}\})=\sum_{n=0}^{\infty}0=0 \tag{5.4}$$

可以從另一個角度驗證 (5.4) 式的合理性, 因

$$[0,1]=(\bigcup_{n=1}^{\infty}\{\frac{1}{2^n}\})\bigcup(\bigcup_{n=1}^{\infty}(\frac{1}{2^n},\frac{1}{2^{n-1}}))\text{ 即}$$

$$\begin{aligned}\mu[0,1]&=\mu(\bigcup_{n=1}^{\infty}\{\frac{1}{2^n}\})+\mu(\bigcup_{n=1}^{\infty}(\frac{1}{2^n},\frac{1}{2^{n-1}}))\\&=\mu(\bigcup_{n=1}^{\infty}\{\frac{1}{2^n}\})+\sum_{n=1}^{\infty}\frac{1}{2^n}\end{aligned}$$

應有$\mu(\cup_{n=1}^{\infty}\{\frac{1}{2^n}\})=0$

在這裡, 如果把上一節 J 測度的有限可加性 (公理 (4.3)) 換為可列可加性 (5.1), 則我們得到一個十分重要的新測度 — 勒貝格 (Lebesque) 測度。

設 l 是數直線, S 是 l 上子集所構成的集合類。S 中的一列元素之並集仍在 S 內, 任兩個元素的差仍在 S 內, 這時, 若有一個定義在 S 上的非負函數 $m(E), E\epsilon S$ 滿足:

(1) 正則性: 任意閉區間 $[a,b]\epsilon S$, $m[a,b]=b-a$

(2) 運動不變性: 若 $E_1,E_2\epsilon S,E_1$ 和 E_2 全等, 則 $m(E_1)=m(E_2)$

(3) 可列可加性: 設 $E_1,E_2,\ldots,E_n,\ldots$ 是 S 中一列集合, 兩兩不相重疊, 則

$$m(\cup_{n=1}^{\infty}E_n)=\sum_{n=1}^{\infty}m(E_n)\text{ 。}$$

這時, 稱集合 $m(E)$ 為勒貝格測度, 簡稱為 L 測度。

細心對照下, J 測度和 L 測度, 除了前者的有限可加拓廣為後者的可列可加外 — 我們說它"拓廣", 還因為由可列可加性可以推導出有限可加性, 前者的定義域 P 僅是有限個點和有限個區間之和, 而後者的定義域 S 中, 一列元素可以相加, 一列點, 一列區間都可以屬於 S, 也就是說 $m(E)$ 的定義域 S

包含了 $\mu(E)$ 的定義域 P, 並有明顯的擴大化和複雜化。

L 測度比起 J 測度是一個很大的進步, 一些用 J 測度無從"量度"長度的點集, 可以換一把新 的尺子 — L 測度把它的長度量出來。

試看區間 [0,1]中有理數全體構成的集合 Q_1, [0,1]中無理數全體構成的集合 M_1。

Q_1, M_1 不屬於 P, 它們沒有 J 測度, 但 Q_1, M_1 在 S 中, 我們可以把它的 L 測度算出來。

先考察 $m(Q_1)$。

我們把 Q_1 排成一列, 以證明 Q_1 是一個可列集: 因爲 [0,1]中的有理數都是既約眞分數 $\frac{p}{q}$, p, q 是正整數且 $p \le q, (p,q) = 1$。我們排列 $\frac{p}{q}$, 規則如下: 分母小的在前, 分母大的在後, 如果分母相同, 則分子小的在前, 大的在後, 於是, Q_1 的有理數可排成 $\frac{1}{1}, \frac{1}{2}, \frac{1}{3}, \frac{2}{3}, \frac{1}{4}, \frac{3}{4}, \frac{1}{5}, \frac{2}{5}, \frac{3}{5}, \frac{1}{6}, \frac{5}{6}, \frac{1}{7}, \ldots$, 每一個 $\frac{p}{q}$ 總要出現在這一列中且位置完全確定, 於是,$Q_1 = \{r_n\}$ 即 Q_1 是可列集。

由可列可加性, 便得

$$\begin{aligned} m(Q_1) &= m(\cup_{n=1}^{\infty} r_n) = \sum_{n=1}^{\infty} m(r_n) \\ &= \sum_{n=1}^{\infty} 0 = 0 \end{aligned}$$

於是

$$m(M_1) = m([0,1]) - m(Q_1) = 1 - 0 = 1。$$

這從另一個側面說明:[0,1]中的無理數雖然也是無限集, 但它不是可列的。$m(M_1) = m([0,1])$ 表明了這樣一個有趣的事實: 一個集合去掉了可列個點 (被挖掉了無限多個"洞"!) 它的 L 測度還是沒有改變。

我們看到, 一列點的並所成的集合的 L 測度爲 0, 但是, 它的逆不眞, 請看下面構造一個非可列的無限個集的並。

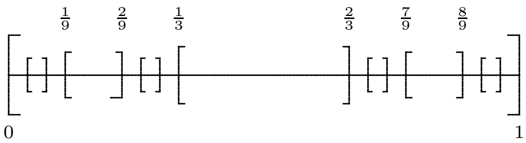

圖 6. Cantor 集

將閉區間 $[0,1]$ 三等分, 去掉中間一個開區間 $I_1^{(1)} = (\frac{1}{3}, \frac{2}{3})$, 把剩下的兩個閉區間 $[0, \frac{1}{3}][\frac{2}{3}, 1]$ 分別再三等分, 再各去掉中間的開區間 $I_1^{(2)} = (\frac{1}{9}, \frac{2}{9})$, $I_2^{(2)} = (\frac{7}{9}, \frac{8}{9})$, 餘下四個閉區間, $[0, \frac{1}{9}]$ $[\frac{2}{9}, \frac{3}{9}]$ $[\frac{6}{9}, \frac{7}{9}]$ $[\frac{8}{9}, 1]$ 又分別把這些閉區間三等分, 再各去掉其中間的開區間 $I_1^{(3)} = (\frac{1}{27}, \frac{2}{27})$, $I_2^{(3)} = (\frac{7}{27}, \frac{8}{27})$, $I_3^{(3)} = (\frac{19}{27}, \frac{20}{27})$, $I_4^{(3)} = (\frac{25}{27}, \frac{26}{27})$, 如此繼續, 在第 n 次三等分時去掉的開區間是 $I_1^{(n)} = (\frac{1}{3^n}, \frac{2}{3^n}), I_2^{(n)} = (\frac{7}{3^n}, \frac{8}{3^n}), \ldots, I_{2^{n-1}}^{(n)} = (\frac{3^n-2}{3^n}, \frac{3^n-1}{3^n})$。令 $O_c = \cup_{n,k} I_k^{(n)}$,$K = [0,1] - O_c$, K 稱爲康托 (Cantor) 集。K 是一個無限集族。可以證明,K 不是一個可列集, 而 K 的 L 測度 $m(K) = 0$。

六. 柳暗花明又一村—從測度到積分

從祖暅原理開始, 我們用無限多個面積來刻劃具有有限體積的立方體, 祖暅的思想正是通過截面積的測度, 避開無限性的麻

煩。然而, 當我們的問題步步深入時, 無限性是不可避免的, 從有限到無限, 這是數學研究從初等到高等的重要轉折。一百多年前, 康托把“無限”引入了數學, 帶來了整個數學面貌的更新。從 J 測度到 L 測度, 正體現了這種質的飛躍。

L 測度使我們在處理極限運算時暢通無阻, 由此, 把建立在 J 測度上的黎曼積分發展爲以 L 測度爲基礎的勒貝格積分。由於函數值的急劇變化而帶來的黎曼積分的不可積問題, 用 L 測度理論可以得到解決。例如,Dirichlet 函數 (在有理數處函數值爲1, 在無理數處函數值爲0的函數), 在任何區間 $[a,b](a<b)$ 上的振幅都等於1, 因而, 在黎曼意義下不可積, 但它卻是勒貝格可積的, 且積分值爲0(請聯繫 $m(Q_1)=0, m(M_1)=1$ 想一想)。

測度理論的發展, 勒貝格積分的出現是二十世紀數學的重大成就之一, 它在“疑無路”的不可積問題中找到了“又一村”, 從而不僅發展了泛函分析的新學科, 而且在概率論, 數理統計, 理論物理方面有很多重要的應用。

附記: 本文寫作於作者應邀訪問香港中文大學數學系期間, 作者對岑嘉評博士, 區國強博士的盛情接待表示感謝。

—本文作者任教於中國華南師範大學數學系—