

基础数学讲义之一

《基础代数学》

项武义

香港科技大学数学系

目录

代序	v
绪论	xvii
0.1 数系运算律	xviii
0.2 代数问题和代数方法简介	xxii
0.2.1 解代数方程式的基本原理和未知数符号之引入	xxii
0.2.2 韩信点兵法，善用分配律的启蒙者	xxv
0.2.3 向量代数与空间本质的线性化	xxviii
0.2.4 线性代数和轨尺作图	xxxii
0.2.5 研讨代数学的几个基本方法	xxxiv
0.3 例题、习题与思考题	xxxvii
0.3.1 数系运算律与数系扩张	xxxvii
0.3.2 辗转相除法与算术基本定理	xxxviii
0.3.3 多项式基本公式举例	xl
一 多项式的基础理论	1
1.1 多项式运算	1
1.2 多项式函数	4
1.3 韩信点兵法和插值公式	9
1.4 求和公式 (Summation Formula)	14
1.5 插值法与因式分解	22
二 二项定理与泰勒公式	27
2.1 二项定理 (The Binomial Theorem)	27
2.2 泰勒公式与多项式的局部展开式	30

2.3 泰勒公式与局部分析	35
三 多项式函数的微积分	43
3.1 变率与微分	43
3.2 总和与积分	48
四 线性方程组与行列式的基础理论	53
4.1 代入法和消元法	54
4.2 二阶和三阶行列式	60
4.3 四阶行列式	65
五 行列式的基本性质与应用范例	73
5.1 n -阶行列式的归纳定义	73
5.2 斜对称多线性函数与行列式的界定定理	78
5.3 行列式的常用基本性质	82
5.4 矩阵的乘法公式和行列式的乘法公式	84
5.4.1 矩阵运算	86
5.4.2 行列式乘法公式	89
5.4.3 平行体体积与行列式	89
5.5 行列式的几个应用范例	94
5.5.1 几何图形的坐标方程	94
5.5.2 行列式与插值公式	96
5.5.3 圆锥截线交点与求解四次方程	99
附录：域上的线性空间与域的代数扩张	107

代序

<节录：第四届苏步青数学教育奖颁奖会的演讲讲义>

精简实用、平实近人、引人入胜 ——论基础数学教育之本质与精要

在即将来临的廿一世纪，科技与智力在总体国力中将居于主导成份。因此，科教乃是兴国必经之途，而基础教育的优良素质则是立国的根本。再者，在中、小学的素质教育中，基础数学教育乃是启发脑力和培训解析思维的主要学科。自古以来，它也一直是教导学生善于认识问题、善于解决问题的最佳园地。相传在古希腊雅典柏拉图学院的门前，曾书有：「不懂几何学者不得入」的字句。这其实就是当年的教育家们对于数学教育的基本重要性的一种体认，数学是学习和研究科学与技术不可或缺的基本工具。但是数学在科技和工程上的应用还只能说是「小乘应用数学」，而基础数学在教育上的用场才是「大乘应用数学」。在佛学中：「小乘独善其身，大乘普渡众生。」基础数学教育乃是普渡众生，成为善于认识问题、善于解决问题者的不二法门。

显然，上述普渡众生的任务也给基础数学教育提出了高素质、高效力的要求，而这也就是我们从事数学教育的播种者与耕耘者，所要集思广益、钻研实践，才能克成其功者。今天我想和各位志同道合的新、老朋友们，从基础数学教育的本质和精要所在，来谈一谈如何把它教得精简实用、平实近人和引人入胜。我觉得唯有如此，基础数学教育才能真正做到普渡众生，并且大幅度地提升国民的思维素质和创造能力。

现在，让我们从下列几点来讨论上述话题：

I. 基础数学的源起与本质

概括地说，人类文明对于大自然的认知和理解的进化过程是由定性层面向定量层面深化。例如先定性地认识到我们所在的大地乃是一个地球，然后再进而估计和测量地球的大小。基础数学的起源就是上述认知定量化的自然产物，而基础数学本身的进程则可以大体简述如下：

- (i) 数系的构造与逐步扩充：例如自然数系、整数系和分数系；这乃是算术的范畴。
- (ii) 由算术进步到代数的关键在于数系运算律的系统运用，亦即以通性求通解。
- (iii) 几何学乃是人类对于其所所在的空间本质的认知的逐步深化，其演进过程大体如下：
实验几何→定性平面几何→定量平面几何→立体几何→坐标解析几何→向量几何。
- (iv) 解析几何乃是代数与几何的自然结合。由此再产生研讨变量问题的基础理论——微分与积分——则是水到渠成、顺理成章的更上一层楼。

现在让我们再来看一看上述基础数学的本质是什麼。归根究底，代数的根本在于数的运算和运算律；几何的根本在于空间的基本结构和基本性质，例如连结于两点之间的直线段乃是两点之间的唯一最短通路，这就是空间的基本结构，而空间对于任给平面的反射对称性则是空间的基本性质；微分和积分运算则是函数的「变率」和「求和」的解析表述和有效计算，它们也就是分析变量问题的基础理论。总之它们的本质都是精简朴实的，它们的根源都是自然而且富有直观内含的。其实，上述简朴、平实、近人的本质也就是为什麼基础数学教育，自古以来一直是锻链脑力、培养思维能力的「益智游乐园」。因此，在基础数学的教育中必须随时随地充分体现它精简实用、平实近人的本质，才能发挥其启发脑力、培训思维的功效，真正成为普渡众生的慈航。

II. 精中求简、以简御繁、至精至简

基础数学的范畴，自然也随著科学文明的蓬勃进展而逐步扩张。例如在古希腊到了高级学府柏拉图学院才要求的几何学，在现代已经成为中学数学中，人人必学的基本学科。而且年轻的学生们还要同时学习各种各样的其他学科。学生负担过重的确是现代教育中一个亟待宽解的重大问题，而要减轻学生的负担则必须要简化教学题材！但是如何把基础数学的教材简化呢？却又是众说纷芸、莫衷一是。在这方面基本上可以归纳为下述两种不同的方针和想法：其一是简略的简化法，其二则是「精中求简」。前者的思路是探讨如何对于基础数学的现有题材作适当的加或减来达成简化课程；而後者的思路则是除了简略所有枝节性的题材之外，还要对于基础数学的核心部分，在总体结构上探讨精中求简的出路。我觉得唯有做好精中求简的研究才能真正提高教学质量与效果，也唯有这样，才能使得基础数学易学、好懂、能懂、会用，从而达成减轻学生的负担。

今天限于时间，且以定性平面几何和定量平面几何为例，简约地讨论精中求简的一些具体途径。

(1) 定性平面几何：定性平面几何所要研讨的主题是「全等形」和「平行性」。在本质上，前者乃是平面对于任给直线的反射对称性的具体反映，而後者则是三角形的内角和恒等于一个平角所表达的「平直性」。我们可以由 S.A.S. 叠合公理和上述「内角和」这样两个基本性质为起点，引导学生去研讨等腰三角形和平行四边形的各种各样性质。然後让他们集思广益，共同研讨那些性质已经构成这两种基本图形的特征性质。亦即让他们自主地、自动地去发现下述两个定性平面几何论证的基本工具：

(A) 等腰三角形的特征性质：（如 [图-1] 所示）

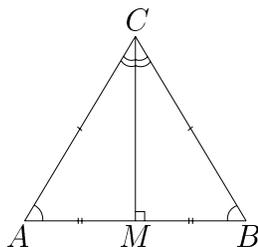
(i) $\overline{CA} = \overline{CB}$ (定义)；

(ii) $\angle A = \angle B$ ；

(iii) $\angle C$ 的分角线 \overline{CM} 垂直底边 \overline{AB} ；

(iv) 中线 \overline{CM} 垂直底边；

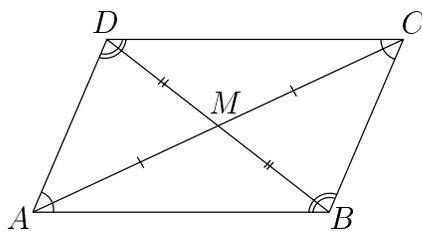
(v) 垂线 \overline{CM} 平分顶角。



[图-1]

(B) 平行四边形的特征性质：(如 [图-2] 所示)

- (i) $\overline{AB} \parallel \overline{DC}$ 而且 $\overline{AD} \parallel \overline{BC}$ ；
- (ii) $\overline{AB} = \overline{DC}$ 而且 $\overline{AD} = \overline{BC}$ ；
- (iii) $\angle A = \angle C$ ，而且 $\angle B = \angle D$ ；
- (iv) \overline{AC} 和 \overline{BD} 互相平分；
- (v) $\overline{AB} \parallel \overline{DC}$ 而且 $\overline{AB} = \overline{DC}$ (或 $\overline{AD} \parallel \overline{BC}$ 而且 $\overline{AD} = \overline{BC}$)。



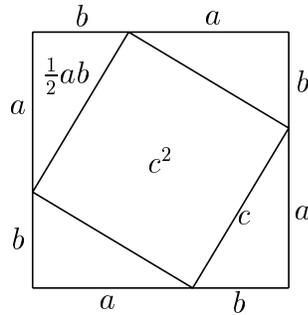
[图-2]

接著引导学生逐步运用上述两个基本工具，一以贯之地论证、解答所有其他平面几何中的定理、例题和习题。其实，等腰三角形的特征性质之间的转换是充分体现了对称性的具体用法，而平行四边形的特征性质之间的转换则是充分体现了平行性的作用和表现。这样的教学途径，不但可以达成定性平面几何的求精中求简，而且也使得学生逐渐学会抓本质，和逐步体认在认知上的「以简御繁」。

(2) 定量平面几何：定量平面几何的基本定理和精要所在乃是三角形面积公式、勾股定理和相似三角形定理。所以定量平面几何教学首务之

要在于简明扼要地推导上述三者，然后引导学生把它们用来解答或论证各种各样定量平面几何的问题。在这里，中国古算中善用面积公式的创见，提供了既简朴又直截了当的途径，其具体做法是以长方形面积等于长乘宽为起点，用熟知的割补推导出三角形面积等于二分之一底乘高。然后用 [图-3] 和 [图-4]，以面积计算推导勾股定理和直角三角形的相似比。

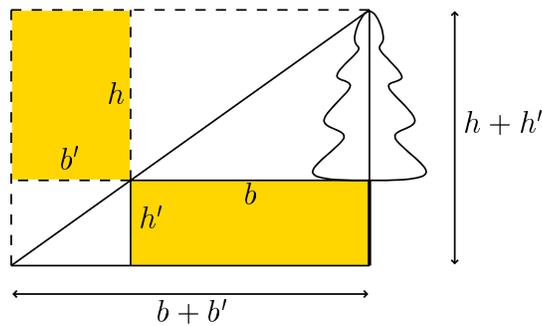
勾股定理：



$$(a+b)^2 = c^2 + 4 \cdot \frac{1}{2}ab \Rightarrow a^2 + b^2 = c^2$$

[图-3]

出入相补：



$$b \cdot h' = h \cdot b' \Rightarrow \frac{b}{b'} = \frac{h}{h'} \Rightarrow \frac{b+b'}{b'} = \frac{h+h'}{h'}$$

[图-4]

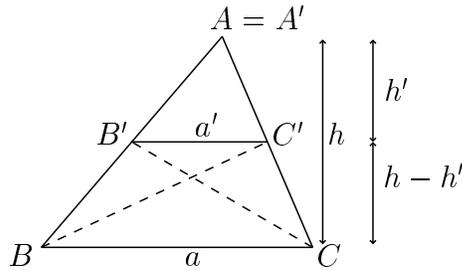
相似三角形定理的面积证法：

[证一]：出入相补所得之比例式

$$\frac{b+b'}{b'} = \frac{h+h'}{h'}$$

其实就是直角三角形的相似形定理。而一般的三角形均可用其垂线分割成两个直角三角形。由此可见，容易结合上述比例式和勾股定理直接推导一般的相似三角形定理。

[证二]：用两种分割法计算 [图-5] 中的梯形面积：



[图-5]

$$\square BCC'B' = \frac{1}{2}ah - \frac{1}{2}a'h', \quad \square BCC'B' = \frac{1}{2}(h-h')(a+a')$$

$$\Rightarrow \frac{h}{h'} = \frac{a}{a'}$$

$$\Rightarrow \frac{\Delta}{\Delta'} = \frac{\frac{1}{2}ah}{\frac{1}{2}a'h'} = \left(\frac{a}{a'}\right) \left(\frac{h}{h'}\right) = \left(\frac{a}{a'}\right)^2$$

$$\Rightarrow \left(\frac{a}{a'}\right)^2 = \left(\frac{b}{b'}\right)^2 = \left(\frac{c}{c'}\right)^2 = \frac{\Delta}{\Delta'}$$

$$\Rightarrow \frac{a}{a'} = \frac{b}{b'} = \frac{c}{c'}$$

[证三]：如 [图-5] 所示，

$$\triangle BB'C' = \triangle CB'C' \quad (\text{同底同高})$$

$$\Rightarrow \triangle ABC' = \triangle AB'C$$

$$\Rightarrow \frac{\triangle ABC'}{\triangle ABC} = \frac{\triangle AB'C}{\triangle ABC} \Rightarrow \frac{b'}{b} = \frac{c'}{c}$$

以上只是基础数学中精中求简的二个实例。其实，基础数学教育在各方面精中求简乃是一个大有可为，值得全面探讨的课题。而这方面的探

讨成果，将是一方面大幅提升教学效益而另一方面还可以真正减轻学生负担的改革途径。当然，这还得要由教材的妥为编写，教学方法的全面改进和教育制度能彻底从「应试教育」中解脱出来等等的配合。

总之，我们要集思广益，共同努力把基础数学的各个环节，各种方面都下功夫。精益求精，务必使得基础数学中的精要所在都以至简形式展示在学生面前。这种至精以至简形式展现的「精简合一」，当然就是基础数学能够用来以简御繁的基本道理。其实，这也就是基础数学最能真正引人入胜的地方。

长话短说，基础数学的「至精至简」可以大体上用下述四套运算律来加以总结。即：

1. 集合运算的运算律乃是逻辑思维的至精至简（亦即布尔代数）。
2. 数系运算的运算律乃是代数学的至精至简。
3. 向量运算的运算律乃是几何学的至精至简。
4. 微分积分的运算律乃是分析学（亦即变量数学）的至精至简。

学习基础数学就是要学会有效地、有系统地运用上述四套运算律去解决和认知大自然中的各种各样问题，这也就是以简御繁的具体实践！它们是人类理性文明中的「大巧若拙」，是人人能懂、到处有用的大道理。若能把基础数学教得精到简朴、能懂好用和平实近人，则引人入胜的基础数学是可望可及的。

III. 返璞归真、平实近人、谆谆善诱、引人入胜

基础数学的本质和基本思想都是平实朴素的，而它们的逐步深入和有系统的运用，则又可以用来探索自然、以简御繁、妙用无穷。但是要把基础数学的教学真正做到平实近人、引人入胜，当然还是很有考究的；而其中好多精微细致之处，则著实是耐人寻味、值得推敲。我觉得教者和教材必须要对于基础数学的本质和基本思想下一番深切的返璞归真的功夫，才能把它教得平实近人。再者，在题材编组和教学方式上，要尽可能善导善诱，让同学逐步渐进地体会精简的妙用；要把基础数学不但教得能懂，而且要懂了又能广泛有用，唯有如此才能引人入胜。

让我们且以代数学上的几个实例，来对上述想法作简要的解说：

(1) 解代数方程原理和代数的起源：

从表面的形式来看，初中代数和小学算术的差别在于代数中引进不定元和多项式运算，然後把它们用来解代数方程式、求公式等等。假若我们对于由算术到代数的进化，下一番返璞归真的功夫去细致推敲，就会发现上述进化历程中真正的突破点在于下述简朴的基本思想：

数系的加、乘和指数运算满足一系列普遍成立的运算律。

例如加和乘的交换律、结合律；乘对于加的分配律和指数定则。骤看起来，像

$$\text{分配律： } m \cdot a + n \cdot a = (m + n) \cdot a$$

其所表达的是： m 个 a 之和再加以 n 个 a 之和即为 $(m+n)$ 个 a 之和。像这种本质十分明显的事实，虽然是普遍成立，但是它们会有什麼用场呢？其实，整个代数学所发展的就是有系统、有效力地运用这一系列简朴、普遍成立的数系运算律，去解决各种各样代数问题。此事的具体做法的首次成功就是把运算律用来解代数方程式。而这个用法的基本思想就是下述解代数方程原理：

在各种各样的代数问题之中，最为简朴的类型是某些待解的「未知量」和某些已经给定的「已知量」之间具有某些特定的代数关系。把这些特定的代数关系简明扼要地列述出来，就是一组含有「未知数」符号的代数方程式。所以代数学中最为简朴基本的问题就是如何由一组代数方程式，去确定其中所含的「未知数」所应取之值，这也就是求解代数方程。解代数方程的基本原理究竟是什麼呢？这也就是每一位开始学习代数的学生理当掌握的。它的基本想法是：那些运算律是对于任何数都普遍成立，所以它们对于「未知数」也当然成立、可用。归根究底，解代数方程的基本原理就是有系统地运用运算律把所给的代数方程简化，从而确定其中所含的「未知数」所应取之值，亦即有意识、有系统地达成化未知为已知的目的。解代数方程的原理就是上述极为简朴的思想，而它在解决代数问题上又是那麽简明有力、妙用无穷。当然，在教学上，我们还要谆谆善诱地由简入深，顺理成章

地由二、三元一次方程组，一元二次方程等入门，然后再逐步引导出多项式运算和多项式函数的基本性质和定理。

例如鸡兔同笼问题是一个同学原来已学过的问题，我们可以把它用来作一次温故知新，解说上述原理的用法。鸡数和兔数是待解的未知数，可以用未知数符号 x, y 分别表示之；而头数和足数则是已知量，设其分别是 36 和 108。则它们之间的代数关系就可以用下述代数式简明扼要地表达之，即

$$\begin{aligned}x + y &= 36 \\2x + 4y &= 108\end{aligned}$$

用第一式可得 $y = (36 - x)$ ，代入第二式，即得

$$(*) \quad 2x + 4 \cdot (36 - x) = 108$$

若以算术中一成不改按部就班地先算小括号再算中括号的办法就无法再简化 $(*)$ -式了，因为小括号中含有未知数！但是只要用上述解代数方程原理，就可以把 $(*)$ -式用分配律逐步简化如下，即：

$$\begin{aligned}\Rightarrow & \quad 2x + 4 \times 36 - 4x = 108 \\ \Rightarrow & \quad 4 \times 36 - 2x = 108 \\ \Rightarrow & \quad 4 \times 36 - 108 = 2x \\ \Rightarrow & \quad x = (4 \times 36 - 108) \div 2 = 18\end{aligned}$$

由此可见，分配律的运用就足以解决算术中括号内含有未知量的困扰，在代数中我们用分配律有系统地简化含有未知量的括式。我们还可以用下述形象的说法来描述这个看来简朴无华而实质上影响深远的进步，即：

在算术中基本上不用分配律，乃是数学的石器时代；及至代数，则有系统地运用分配律去简化各种各样代数式和代数关系，这就进步到数学的铜器时代了！

(2) 归纳法和代数公式、代数定理的发现和证明

代数学和几何学有一个本质上的差异。几何学所研讨的空间是我们生活的所在，我们对于它的种种形象和基本性质，与生俱来地有丰富的

感性认识和相当可靠的直观，所以好多几何学上的基本定理如矩形面积公式，相似三角形定理等等是可以直观地看到、想象到它们应该是对的。但是代数学所研讨的是数系的结构和各种各样公式，它们在本质上是逐步归纳、复化所构造而得者，它们的直观性比之于几何就相去甚远了。长话短说，代数学中的公式和定理绝大部分都是用归纳法由低次到高次，由一元，二元到多元逐步归纳而发现，然后再用归纳论证去确立其正确性。因此，归纳乃是整个代数学的基本大法和基本功。但是这里要特别强调，归纳法的内含是归纳地去探索、发现，然后再归纳地定义（例如 n 阶行列式用 $(n-1)$ -阶行列式加以定义），然后再归纳地论证。唯有这样才是完整的归纳法教学，才真正能做好代数学中，公式和定理的返璞归真，也唯有这样归根究底，才能把代数学教得平实近人。像目下把所要证的公式定理看成书本的宣示，然后再加以归纳论证，乃是只教后半段的不完整的归纳法教学，是亟待补正的！

(3) 插值法、待定系数法

插值法 (interpolation) 和待定系数法 (method of undetermined coefficients) 是初等代数学的要点和局部制高点，而因式分解则是可以大幅度地加以简略。从初等代数的教学题材的总体分析，我觉得在传统教学中占用大量篇幅的因式分解是可以而且应该大幅度的简略。其实，除了极为简朴的公式如

$$(x^n - a^n) = (x - a)(x^{n-1} + x^{n-2}a + \dots + xa^{n-2} + a^{n-1})$$

和二项定理之外，绝大部分其他因式分解的公式并无真正的用场。而且大量的因式分解的例、习题也并不能让学生增加代数上的体会和修养。而另一方面，插值法和待定系数法则是充分体现代数学基本精神，而且是非常有用、有力的基本功。但是一般初等代数，甚至于高中代数的教学中，却往往忽略未教，或者只是轻轻带过、未加深入。这种轻重倒置的格局是亟待改正的；而把它改正过来，则是使得初等代数的教学题材在总体上能够做到精简实用、引人入胜的基本途径。

再者，插值法的源头其实就是韩信点兵法（西方称之为中国剩馀定理），而插值法乃是代数学中求公式的根本方法。所以其教学应该从

韩信点兵教起，然後把它用来去求得各种有用的公式如求和公式：

$$\begin{aligned}0 + 1 + 2 + \dots + (n-1) &= S_1(n), \\0^2 + 1^2 + 2^2 + \dots + (n-1)^2 &= S_2(n), \\0^3 + 1^3 + 2^3 + \dots + (n-1)^3 &= S_3(n), \\0^4 + 1^4 + 2^4 + \dots + (n-1)^4 &= S_4(n).\end{aligned}$$

插值法的基本原理是：一个待求的 k 次多项式可以由它在 $k+1$ 点所取之值去求解，而待定系数法则是把一个待求的公式中的待定系数转化成一组代数方程式的未知数来求解而得。总之，这两种基本方法是初等代数中普遍有用的想法和算法，理当在初等代数的教学中多用举例说明，逐步深化的手法，引导学生学会其用法，体认其用场。

最後，请允许我向各届获奖的得奖人再一次衷心祝恭，并以下述几句话作为我们珍重再见的祝愿：

愿祖国大地，山永青，水常绿，苗木遍野；
愿我们以播种者，耕耘者，共勉互励。

1999 年 10 月 24 日于上海复旦大学

绪论——大巧若拙的运算律； 代数问题与代数方法简介

同学们在初中都已学习过「代数」，例如多项式运算，解代数方程等等。现在让我们来谈一谈代数学的基本思想和基本方法究竟是个什麼？要明确地说明这样一个代数学的本质性问题，当然就得从代数学的起源、代数问题的范例以及代数方法的精要作一次探本究源的工作。这也就是我们今天开始学习「基础代数学」所要研讨的首要课题。

开宗明义，代数学的根源在于代数运算，亦即加、减、乘、除、乘方、开方等等。所有能够用代数运算加以表达的问题都统称之谓代数问题。而代数学 (Algebra) 这门学问所要研讨者就是如何有效、有系统地去解决各种各样的代数问题。代数学中用来解决多种多样的代数问题的基本思想和基本方法究竟是个什麼呢？这就得由数系的代数运算的本质和通性说起了。

代数运算具有一系列普遍成立 (universally valid) 的运算律 (laws of operations)，亦即

加、乘的交换律： $a + b = b + a, a \cdot b = b \cdot a$

加、乘的结合律： $a + (b + c) = (a + b) + c, a \cdot (b \cdot c) = (a \cdot b) \cdot c$

分配律： $(b + c) \cdot a = (b \cdot a) + (c \cdot a)$

指数定则： $a^m \cdot a^n = a^{m+n}, (a^m)^n = a^{m \cdot n}, (a \cdot b)^n = a^n \cdot b^n$

代数学的基本思想就是有效、有系统地运用这些普遍成立的运算律去解答多种多样的代数问题！由此可见，今天要和同学们一起来做的「代数学寻根之旅」理当由「数系运算律」的寻根做起。

0.1 数系运算律

话说当年，同学们在小学、初中的数学课程中，逐步渐进地学习了逐级扩充的数系运算：由最为原始的自然数系 \mathbb{N} （亦即正整数系）起始，到正负整数系 \mathbb{Z} ；再到正负分数系 \mathbb{Q} （亦称为有理数系）；然後在高中再到实数系 \mathbb{R} 和复数系 \mathbb{C} 。同学们现在应该来回想、温习一下：在上述这一系列逐级扩充的过程之中，每一次扩充究竟添加了那一类「新数」？添加了它们之後的扩张数系比之于原先者究竟又多有了那些好处？再者，所引进的新数之间的运算究竟是如何归结到原有者之间的运算来加以定义的？例如分数运算就是用下列算式归于整数运算加以定义的，即

$$\frac{m}{n} + \frac{p}{q} = \frac{m \cdot q + n \cdot p}{n \cdot q}, \quad \frac{m}{n} \cdot \frac{p}{q} = \frac{m \cdot p}{n \cdot q}$$

而复数之间的运算则是用下列算式归于实数运算加以定义的，即

$$\begin{aligned}(a + b\sqrt{-1}) + (c + d\sqrt{-1}) &= (a + c) + (b + d)\sqrt{-1} \\ (a + b\sqrt{-1}) \cdot (c + d\sqrt{-1}) &= (ac - bd) + (ad + bc)\sqrt{-1}\end{aligned}$$

对于上述数系的逐级扩充，即

$$\mathbb{N} \subset \mathbb{Z} \subset \mathbb{Q} \subset \mathbb{R} \subset \mathbb{C}$$

作一次总复习乃是每一位同学都应该自动自发地去回想、深究的大好习题。而下这样一番功夫肯定会让你体会到温故知新的得益和乐趣的！

现在让我们先来讨论下述两个问题：

- (一) 这些看起来简朴得几乎是笨拙的运算律究竟能有多大用场呢？
- (二) 运算律普遍成立的原由何在？

因为数系运算律是普遍成立而且又是极为简朴易用的算式，所以它们是广泛能用而且简单有力的代数学基本工具。其实，运算律就是整个代数学的基础！但是运算律广大无穷的用场和它们多彩多姿，既简朴又精到的种种用法，则有待在往後的章节以及更进一阶的代数学中逐步、逐样地去开拓、展现它。所以对于第一个问题的回答乃是：「且

听下回和下下回分解。」我们现在要和同学们作比较详细的初步分析的是第二个问题。

在数系逐阶扩张中，我们可以清楚地看到下述两点：第一点、在 $\mathbb{N} \subset \mathbb{Z} \subset \mathbb{Q} \subset \mathbb{R} \subset \mathbb{C}$ 这一串的逐阶扩张中，「後者的运算律之普遍成立」乃是由「前者运算律之普遍成立」推导而得者。第二点、使得运算律在逐步扩张的过程中得以保持其运算律的普遍成立，其实就是每一步扩张的做法的一个指导思想。总之，运算律普遍成立的原由，归根究底就得要仔细考察一下自然数系的运算律为什麼普遍成立的呢？在这里，也许有些同学会问：2+3 和 3+2 都等于 5；2×3 和 3×2 都等于 6；这岂不是熟知而且极为明显的事实！？还有什麼好讨论的？叫人觉得这简直是庸人自扰！？但是请大家注意，这种举例说明，或者逐一验证是无法说明运算律是普遍成立的！试问有那一位同学曾经对于三个「百位数」的乘法，真的去验算过结合律是否成立的？显然，这种验算一来极为繁琐、费时而且极易出错；二来于事无补，因为即使费了千牛万虎之力，还是只能对于无穷多个情形之中的小小一角做了验算吧了，是不？由此可见，追根究底地问自然数系运算律为什麼普遍成立——此事绝非庸人自扰，乃是学习代数不可缺的奠基与起步！

这里，我们得先说明一下「证明」这件事的本质。证明者也，乃是用某些事物的正确性去说明其他一些事物的正确性，例如以整数系运算律为基点，可以进而证明分数系运算律的成立。总之，任何论证都必须有所本，证明是不可能「无中生有」的！由此可见，要证明自然数系的运算律，也必须要有本；而这样一个最、最原始的根本究竟是什麼呢？且让我们来剖析一下自然数系的原始本质吧。

自然数系是我们用来数「个数」的数学工具 (mathematical tool for counting)，它的本质就是一个顺序排列的体系，其起始者为 1，往後顺序地「後者」表示比「前者」多加 1 个。亦即

$$2 = 1 + 1, \quad 3 = 2 + 1, \quad 4 = 3 + 1, \quad 5 = 4 + 1, \quad \dots\dots$$

如此逐个加 1 以至于「无穷」。由此可见，自然数系最、最原始根本的结构就是「+1」运算。任何自然数都可以由 1 起始，逐步「+1」而达到它。上述对于自然数系本质的朴素的描述也就是通常叫做数学归纳法原理 (Principle of Mathematical Induction) 者也！这也就是我们即将论证自然

数系运算律的所本。其实，认清了上述基本出发点，就不难看到下面关于加法，乘法和乘方的一系列返朴归真的剖析了。

(一) 加法是「+1」的复合，例如

$$a + 2 = (a + 1) + 1, \quad a + 3 = (a + 2) + 1 = ((a + 1) + 1) + 1$$

亦即「+ n 」乃是把「+1」连续做 n 次的缩写。所以 $+(n + 1)$ 乃是「+ n 」之後再多做一次「+1」者也。改用算式表达，亦即

$$(0.1) \quad a + (n + 1) = (a + n) + 1$$

其实，上述算式也就是自然数系的加法运算的归纳定义 (inductive definition)。

(二) 乘法是自相加的缩写，例如

$$1 \cdot a = a, \quad 2 \cdot a = a + a, \quad 3 \cdot a = 2 \cdot a + a = (a + a) + a$$

亦即

$$(0.2) \quad (n + 1) \cdot a = n \cdot a + a$$

其实，上述算式和 $1 \cdot a = a$ 也就归纳地定义了自然数系中的乘法运算。

(三) 乘方是自相乘的缩写，例如

$$a^1 = a, \quad a^2 = a \cdot a, \quad a^3 = a^2 \cdot a = (a \cdot a) \cdot a$$

亦即 a^n 乃是 n 个 a 自相乘的缩写，所以

$$(0.3) \quad a^{n+1} = a^n \cdot a, \quad (a^1 = a)$$

也就是乘方运算的归纳定义。

由上述逐步归纳定义，我们剖析了自然数系中的加、乘和乘方运算是如何由最、最原始的「+1」逐步归纳组合而成的。由此也就不难想到它们的运算律是可以由它们的归纳定义为起点，逐步地加以归纳论证的。下面只是举三个关键性的归纳论证为例，其他的逐一论证则正好是同学们开始练习归纳论证的自然起点。

(i) 加法结合律的归纳证明：

$$(0.4) \quad (a + b) + c = a + (b + c)$$

对于 c 作归纳证明：在起始的 $c = 1$ 时，上式就是加法的归纳定义式，即 (0.1)-式。再者，我们归纳地设 (0.4)-式对于 c 成立，从而去论证 (0.4)-式对于 $(c + 1)$ 也因而成立。其证明如下：

$$\begin{aligned} (a + b) + (c + 1) &= ((a + b) + c) + 1 && \text{[加法定义式]} \\ &= (a + (b + c)) + 1 && \text{[归纳假设]} \\ &= a + ((b + c) + 1) && \text{[加法定义式]} \\ &= a + (b + (c + 1)) && \text{[加法定义式]} \end{aligned}$$

(ii) 分配律的归纳证明：

$$(0.5) \quad m \cdot a + n \cdot a = (m + n) \cdot a$$

对于 n 作归纳论证：在起始的 $n = 1$ 时，(0.5)-式就是定义式 (0.2)。而 (0.5)-式的归纳论证如下：

$$\begin{aligned} m \cdot a + (n + 1) \cdot a &= m \cdot a + (n \cdot a + a) && \text{[乘法定义式]} \\ &= (m \cdot a + n \cdot a) + a && \text{[加法结合律]} \\ &= (m + n) \cdot a + a && \text{[归纳假设]} \\ &= [(m + n) + 1] \cdot a && \text{[乘法定义式]} \\ &= [m + (n + 1)] \cdot a && \text{[加法定义式]} \end{aligned}$$

(iii) 指数定则 $a^m \cdot a^n = a^{(m+n)}$ 的归纳论证：

在起始的 $n = 1$ 时， $a^m \cdot a = a^{m+1}$ 就是定义式 (0.3)。对于 n 的归纳论证如下：

$$\begin{aligned} a^m \cdot a^{n+1} &= a^m \cdot (a^n \cdot a) = (a^m \cdot a^n) \cdot a \\ &= a^{m+n} \cdot a = a^{(m+n)+1} = a^{m+(n+1)} \end{aligned}$$

总之，自然数系的加、乘和乘方运算都是由最、最原始的「+1」运算，逐步复合而得者，而 (0.1), (0.2) 和 (0.3)-式分别是它们的归

纳定义式。关于这三种基本代数运算的一系列运算律，则可以逐步地以上述三个归纳定义式为起点，加以归纳证明之！这样就建立起运算律普遍成立这个代数学的基础。这也就是我们学习基础代数学的首务之要：代数学的探源与奠基工作。

接著让我们来谈一谈这些简朴的运算律的用场与用法。概括地说，运算律的用场在于它们是解答各种各样代数问题的基本工具，而其多姿多采的用法也就是我们将要逐步介绍其梗概的种种代数方法。其实，学习代数学，就是要学会善用运算律去有效、有系统地解决多种多样的代数问题。再者，多姿多采的代数方法则当然是随著代数问题的拓展与深化而亦步亦趋地发展起来的；新的代数问题往往激发起新的代数方法，而新的代数方法又可以把更多样的代数问题得以系统地理解并且把代数学应用的领域得以拓展。整个代数学就是这样地蓬勃开展起来的。长话短说，下面就让我们先对代数问题和代数方法作一个简短的初步介绍。

0.2 代数问题和代数方法简介

代数问题的范畴既广且深，这里只举几种初等的类型为例。它们在代数学的启蒙与奠基阶段，则扮演著重要的角色。

0.2.1 解代数方程式的基本原理和未知数符号之引入

同学们在小学时学习算术，在初中时进而学习代数。在数学的发展史中，代数乃是由算术进步而得者。现在让我们来剖析一下，从算术到代数的进步，其要点何在？突破点何在？回顾当年，在小学算术中所学的四则应用题，如年龄问题、鸡兔问题等等，大体上来说，当时是通过对于每一类型的问题详加分析，各别地求得其所针对的公式来加以解答的。但是到了初中代数，这些多样的四则题统统都可以用简单的线性方程组加以解答。其中究竟使了什麼「巧妙」呢？骤看起来，其妙处好像在于引进了「未知数」符号 (unknown)。例如在鸡兔问题中，可以用 x, y 分别表示「兔数」与「鸡数」；则所给问题中的数量

关系可以列成两个方程式，即

$$\begin{cases} x + y = \text{头数} \\ 4x + 2y = \text{足数} \end{cases}$$

在算术中虽然不用 x, y 这种符号，却也可以同样地写下

$$\begin{cases} \text{兔数} + \text{鸡数} = \text{头数} \\ 4 \times \text{兔数} + 2 \times \text{鸡数} = \text{足数} \end{cases}$$

由此可见，到此为止，代数和算术根本是毫无本质上的差别。其实，代数解法的真正妙处并不在于引进未知数符号和列方程式，而是在于能够由代数方程去解得未知数所应有的值！现在让我们以头数、足数分别是 15 和 38 为例，来分析一下代数解法和算术解法的差别何在？

(i) 代数解法：用第一式解得 $y = 15 - x$ ，代入第二式即得

$$4x + 2 \cdot (15 - x) = 38$$

用分配律将上式简化，即得

$$\begin{aligned} 4x + 30 - 2x &= 2x + 30 = 38 \\ \Rightarrow 2x &= 38 - 30 = 8 \Rightarrow x = 4 \end{aligned}$$

(ii) 算术解法：若笼中全部是鸡（即 15 只），则足数应为 $2 \times 15 = 30$ 。今逐一以兔换鸡，则头数保持不变而足数则每次增加 2。由此可见，要使得足数由 30 增加到 38，所需做的更换次数是 $(38 - 30) \div 2 = 4$ 。这样就求得兔数 = 4。

【比较分析】：在代数解法中，我们用运算律和「移项」即可解得 x, y 所应有之值（化未知为已知！）。而在算术解法中，老师其实早已「心中有式」，即

$$\text{兔数} = (\text{足数} - 2 \times \text{头数}) \div 2$$

然後再用上述这样一套「说词」来解释其合理性。这里也许同学们会问：「那麽老师『心中有式』的上述公式究竟是如何想到的呢？」此事大致有两种途径，其一是用代数方程求解而得，即

$$\begin{cases} x + y = a \\ 4x + 2y = b \end{cases} \Rightarrow 2x + 2(a - x) = b$$

$$\Rightarrow 2x = b - 2a \Rightarrow x = (b - 2a) \div 2$$

其二是由实验归纳而得，即

兔数	鸡数	足数
0	a	$2a$
1	$a - 1$	$2a + 2$
2	$a - 2$	$2a + 4$
3	$a - 3$	$2a + 6$
\vdots	\vdots	\vdots

上面所讨论的虽然只是一个简单的实例，但是我们已经可以看到两种解法的真正区别是：在算术中未能有效利用运算律；而代数则有系统地用运算律去简化给定的数量关系，从而化未知为已知。要把这个区别认识得更加清楚一些，请同学们再回想一下，在小学时学习算术四则运算，是不是有一个口诀：「先算小括号，再算中括号，然後算大括号。」总之，在算术四则中，括号是由里及外，逐步去括号来加以计算的，是不？这种算法在括号中含有未知量时，例如 $4x + 2 \cdot (15 - x) = 38$ ，就行不通了；而在代数中，我们可以用分配律把 $2 \cdot (15 - x)$ 改写成 $30 - 2x$ 。这叫做「穷则变，变则通！」

上述「变则通」的原理如下：在算式中的未知数在尚未解出其应有之值之前，虽然不知其值为何，但肯定是一个数，所以那种对于任何数皆普遍成立的运算律是肯定可用的。其中最为有力者是分配律，它使得算式中的括号可去、可加，不会因为某一括号中含有未知量而「受阻」。这也就是解代数方程式的基本原理！这才是算术与代数真正的分野！

可以这麼说，在算术年代，是没学会用分配律的「石器时代」；到了代数，则进步到大用运算律（特别是分配律）的「铜器时代」。在这里，我们还要再谈一谈「引进未知数符号」这个措举的真正好处和意义。

由上述分析可见，解代数方程式的基本原理在于系统地用运算律，把含有未知量的算式加以简化、分析。引进符号如 x, y, z 等来表达未知量的做法就是要把上述原理具体化、形象化；亦即未知数就是那种在运算上满足运算律的符号。这样做，就使得我们在对于「未知量」使用运算律时「有的放矢」（未知数符号是「的」，运算律是「矢」），运用起来自然更能得心应手，是不？再者「代数」这个名词

，顾名思义，乃是因为这种以符号代表未知数等等这种做法在代数学中广泛运用而得其名者也。其实，这种做法真正好处在于使得对于各种各样具有数系通性的事物如未知数、变数、待定系数等等，能够有的放矢地运用运算律，得以正规化、系统化。在中学代数中所学的多项式运算就是这个做法的具体形式。

在中学代数课程中，所学习的单元（或多元）多项式运算，其实就是含有单个（或多个）字母符号的算式之间的运算。那些叫做「元」的符号的本质就是它们在运算上满足运算律（亦即对于任何数皆普遍成立的通性）。例如两个多项式的乘积是用分配律把它归于单项式的乘积之和式来计算，而单项式的乘积则是用乘法的交换律、结合律和指数定则来计算的。所以多项式本质上就是具有运算律的符号所组合而成的代数式。它们之间的代数运算，则就是对于这些符号运用运算律所能做的形式运算。例如

$$\begin{aligned}(x+1)(x-1) &= x^2 - 1, \\(x^2+x+1)(x-1) &= x^3 - 1, \\(x^3+x^2+x+1)(x-1) &= x^4 - 1, \\(x+y)^2 &= x^2 + 2xy + y^2, \\(x+y)^3 &= x^3 + 3x^2y + 3xy^2 + y^3, \\(x+y)^4 &= x^4 + 4x^3y + 6x^2y^2 + 4xy^3 + y^4, \text{ 等等}\end{aligned}$$

从下一章开始，我们将会有系统地研讨单元多项式的基础理论，而它的一个重要应用则是它为微分学提供了基础；微分学的本质就是对于一个可微函数描述其多项式函数的局部逼近。

0.2.2 韩信点兵法，善用分配律的启蒙者

在数系的运算律之中，分配律是唯一连结加、乘这两种基本运算者。归本究源，分配律起源于「乘乃是自相加的复合」这个「乘的本质」而且它本身又简明扼要地总结了加、乘之间的关联。在代数学中，我们对于所有运算律当然都不分彼此地一起运用，但是其中以分配律最为重要、有力，而且它的用场和用法上也特别有讲究。在很多基本代数问题的研讨上，是否能够善用分配律往往就是成败的关键。在这一

点上，中国古算中有一个独到的创见，这也就是我们接著要详加剖析的「韩信点兵法」，它也就是在数论中具有基本重要性的「中国剩餘定理」(Chinese Remainder Theorem)。

我们先谈一谈一个引人入胜的传说：话说当年，在楚汉相争的年代，韩信是刘邦的大将。有一天，刘邦亲临韩信的军营，而在兵场上约有二千名士兵在操练。韩信命令士兵们先後以七人一组、十一人一组及十三人一组结集成小组，并把每次餘下不能组成七（或十一、十三）人小组的人数报上，然後他便可以快速地算出士兵的确切数目。例如，在组七人小组时餘下 3 人，组十一人小组时餘下 4 人，组十三人小组时餘下 8 人，韩信很快便算出其确切数目是 1984 人；其後用一个跟著一个点的费时数法确认了人数正是 1984，而且当中更包括一些由刘邦暗中吩咐混入士兵们中的数名近卫军，以防韩信事前已知道士兵数目的可能性。故事我们就只说到这里，接著让我们来讨论上述「韩信点兵法」背後的数学内涵。

韩信点兵法在数学方面的分析：

- (i) 设 x 为士兵数目，则当 x 除以 7、11、13 时的餘数便分别是 3、4、8（简称餘数集为 $(3, 4, 8)$ ）。首先要注意，上述条件并不能唯一地确定 x 的值，但若两数 x 与 x' 具有相同的餘数集，则两者的差别乃是一个 $7 \times 11 \times 13 = 1001$ 的倍数。因此，韩信很容易就可以选定其中最接近粗估其人数约为 2000 的那个 x 为答案。
- (ii) 设 x_1 、 x_2 、 x_3 分别是以 $(1, 0, 0)$ 、 $(0, 1, 0)$ 、 $(0, 0, 1)$ 为餘数集而且小于 1001 的解。根据假设 x_1 是 $11 \times 13 = 143$ 的倍数，即 $x_1 = 143 \cdot y_1$ ，不难求出 $y_1 = 5$ ，即得 $x_1 = 715$ 。同理， $x_2 = 7 \times 13 \cdot y_2 = 91 \cdot y_2$ ， $y_2 = 4$ ，即得 $x_2 = 364$ ； $x_3 = 7 \times 11 \cdot y_3 = 77 \cdot y_3$ ， $y_3 = 12$ ，即得 $x_3 = 924$ 。这些 x_1 、 x_2 、 x_3 便是韩信预先早已算妥记得的「胸有成竹」。
- (iii) 跟著由分配律可直接得出 $r_1 \cdot x_1 + r_2 \cdot x_2 + r_3 \cdot x_3$ 的餘数集便是 (r_1, r_2, r_3) ！其中 $0 \leq r_1 < 7$ ， $0 \leq r_2 < 11$ ， $0 \leq r_3 < 13$ 。在上述例子中， $(r_1, r_2, r_3) = (3, 4, 8)$ ，所以韩信很快便能算出 x 的值是 1984。

韩信点兵法的基本想法和启示：

现在让我们来剖析一下「韩信点兵法」的基本想法和这样一个成功范例所展现的启示：

基本想法：在求解剩餘问题时，当餘数之中只有一个是 1，其他皆为 0 的特殊情形，不但容易解答，而且可以用这一系列特殊解，把一般情形的解答简洁地用下述公式表达之。设 x_1, x_2, \dots, x_k 是对于给定一组公因数为 1（亦即互质）的除数 $\{a_1, a_2, \dots, a_k\}$ ，其餘数组分别是 $(1, 0, \dots, 0), (0, 1, 0, \dots, 0), \dots, (0, \dots, 0, 1)$ 的特殊解，则

$$(0.6) \quad x = r_1 x_1 + r_2 x_2 + \dots + r_k x_k$$

乃是一个餘数组是 (r_1, r_2, \dots, r_k) 的解。而且任何以 (r_1, r_2, \dots, r_k) 为餘数组的解和上述 x 相差一个 $a_1 a_2 \dots a_k$ 的整数倍。上述公式 (0.6) 之所以普遍成立的理由就是分配律。由分配律可见 $r_i x_i$ 被 a_i 除的餘数是 r_i 而且 $x - r_i x_i = \sum_{j \neq i} r_j x_j$ 被 a_i 除的餘数是 0（因为每一个 $r_j x_j, j \neq i$ ，都含有因子 a_i ）。由此可见，上述算法的基本思想就是善用分配律，把剩餘问题的一般情形，直截了当地归于易解好算的特殊情形来系统解答之。

有鉴于韩信本人乃是楚汉相争时代的大将军，在此不妨借用军事术语来比喻上述基本思想：假如我们把剩餘问题看做「战场的全局」的话，则那些特殊餘数组的解就是能够简明扼要地控制全局的「战略要点」。所以上述分解组合由特殊解去全局控制一般解的想法，的确展现了韩信在认识问题、解决问题上的「大将之风范」。他给後学後进的一般性启示大致如下：

「当我们在研讨某一种问题，不宜拘泥也不必局限于原给的问题，而是要把它放到恰到好处的广度和范畴之中去处理。唯有这样才有可能把同类的问题妥加组织（例如在剩餘问题中由分配律的运用所得的 (0.6)-式），然後再用这种组织去实现分解组合系统解答的战略思想。」

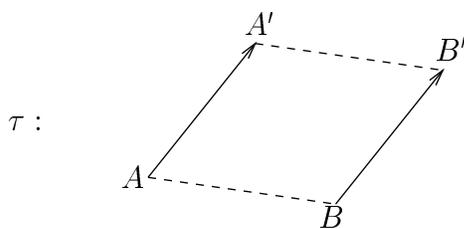
当然，如何对于某一种问题把其相应的广度和范畴选得恰到好处和如何把所选定的全局妥加组织乃是密切相关的。此事一来随著待解的问题而定，二来也有赖于解题者对于该问题的认识深度而定，所以无

法有定规，也难以有一概之论。但是在基础数学中却自然而然地有些基本常用的问题，也都可以同样地用分配律加以组织，使得一般解和特殊解之间的关系也和 (0.6)-式同样地是一种倍数组组合（或称为线性组合）。这一类问题通称之谓线性问题。研究有效解答线性问题，并且拓展其应用的课题则叫做线性代数 (Linear Algebra)，这其实就是我们在大学基础数学课程中所要研讨的一个中心课题，而且它也是在整个数学中一个最广泛有用的精简所在。

总结上面这一段讨论，可以说线性代数的基本想法就是系统地运用分配律去解决各种各样线性问题。而这种善用分配律去解决线性问题的基本方法和思路，业已在韩信点兵法中充分体现、简明扼要地表达得「淋漓尽致」。将来同学们在学习线性代数的起步时刻，当知线性代数的基本方法，实乃起源于中国古算的「中华祖法」。

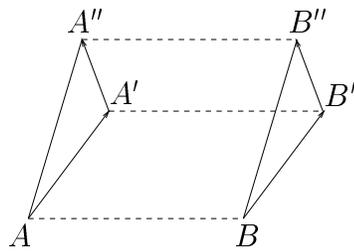
0.2.3 向量代数与空间本质的线性化

上面两小节的讨论，初步说明了分配律在代数学上的重要性。大体上来说，分配律在代数学中的基本重要性是很自然也相当明显的，但是分配律在几何学中也同样具有基本重要性，则是并非显而易见的了。长话短说，在空间中把每个点都作一同向等距的移动（如 [图 0-1] 所示）叫做一个平移（平行移动的简称）。用「平行四边形定理」（亦即一个四边形若有一对对边平行且等长则其另一对对边也平行且等长）即可证明两个平移的组合必定还是一个平移（见 [图 0-2]）。



[图 0-1]

同理也可证两个平移的组合是可交换的，即 $\tau_2 \circ \tau_1$ 和 $\tau_1 \circ \tau_2$ 恒为同一平移。以上事实为基础，就自然可以把平移的组合看成平移之间的「加法」。而且把它想成一种带有方向成份的量，称之为「位移向量」，往後将用小粗体拉丁字母如 **a, b, c** 等表示位移向量。



[图 0-2]

接著我们还可以引进一个向量的倍积 (Scalar Multiplication), 即 $k \cdot \mathbf{a}$ 在 k 是正 (或负) 实数时乃是那个和 \mathbf{a} 同向 (或反向) 而且其长度则是 \mathbf{a} 的 $|k|$ -倍的位移向量。例如正整数倍 $n \cdot \mathbf{a}$ 其实就是 n 个 \mathbf{a} 自相加的总和, 分数倍 $\frac{m}{n} \cdot \mathbf{a}$ 其实就是那个唯一满足 $n \cdot \mathbf{x} = m \cdot \mathbf{a}$ 的向量。由上述倍积的定义, 不难看到它满足下列运算律, 即

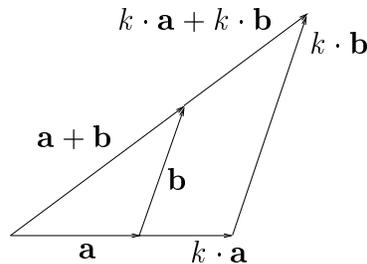
$$k_1 \cdot (k_2 \cdot \mathbf{a}) = (k_1 \cdot k_2) \cdot \mathbf{a}$$

$$(k_1 + k_2) \cdot \mathbf{a} = k_1 \cdot \mathbf{a} + k_2 \cdot \mathbf{a}$$

但是另外一个分配律, 即

$$k \cdot (\mathbf{a} + \mathbf{b}) = k \cdot \mathbf{a} + k \cdot \mathbf{b}$$

是否也成立呢? 如 [图 0-3] 所示, 其实和上述分配律相应的几何事实就是平面几何中极为重要有用的相似三角形定理! 所以不但上述倍积分配律是普遍成立的, 它本身根本就是相似三角形定理的代数化!



[图 0-3]

一个位移向量 \mathbf{a} 同时具有长度和方向这样两种内含。我们将用 $|\mathbf{a}|$ 表示 \mathbf{a} 的长度, 用 $\angle(\mathbf{a}, \mathbf{b})$ 表示 \mathbf{a}, \mathbf{b} 这一对向量的方向差, 亦即两者之间的夹角。在平面几何学的研讨中, 三角形是既精且简的基本图形

。用向量来表达三角形，则它的三个有向边就可以分别表达成 \mathbf{a} , \mathbf{b} 和 $\mathbf{a} + \mathbf{b}$ 。由平面几何中所熟知的 S.S.S. 叠合条件可见夹角 $\angle(\mathbf{a}, \mathbf{b})$ 业已被其三边边长 $|\mathbf{a}|$, $|\mathbf{b}|$, $|\mathbf{a} + \mathbf{b}|$ 所唯一确定（亦即余弦定律）。再者，中国古算中的勾股定理可以改写成

$$\text{若 } \angle(\mathbf{a}, \mathbf{b}) = \text{直角}, \text{ 则有 } |\mathbf{a} + \mathbf{b}|^2 = |\mathbf{a}|^2 + |\mathbf{b}|^2$$

在一般 \mathbf{a} , \mathbf{b} 并非互相垂直的情形则 $|\mathbf{a} + \mathbf{b}|^2 - |\mathbf{a}|^2 - |\mathbf{b}|^2 \neq 0$ 。例如当 $\mathbf{a} = \mathbf{b}$ 的特殊情形，则有

$$|\mathbf{a} + \mathbf{b}|^2 - |\mathbf{a}|^2 - |\mathbf{b}|^2 = 4|\mathbf{a}|^2 - |\mathbf{a}|^2 - |\mathbf{a}|^2 = 2|\mathbf{a}|^2$$

总之，对于任给两个位移向量 \mathbf{a} , \mathbf{b} ，下述函数

$$(0.7) \quad f(\mathbf{a}, \mathbf{b}) = \frac{1}{2} \{ |\mathbf{a} + \mathbf{b}|^2 - |\mathbf{a}|^2 - |\mathbf{b}|^2 \}$$

是一个值得研讨的几何量。例如 $f(\mathbf{a}, \mathbf{b}) = 0$ 乃是 \mathbf{a} , \mathbf{b} 互相垂直的充要条件，而且 $f(\mathbf{a}, \mathbf{a}) = |\mathbf{a}|^2$ 。所以它显然是一个和 \mathbf{a} , \mathbf{b} 的长度、夹角都密切相关的几何量。但是归根究底 (0.7)-式所定义的几何量是否真正有用、好用，还得要看它是否具有简洁好用的优良性质。它显然具有对称性，即 $f(\mathbf{a}, \mathbf{b}) = f(\mathbf{b}, \mathbf{a})$ ，而详加研讨的结果会发现它其实还具有下述简洁易算的性质，即

$$(0.8) \quad f(\mathbf{a}, \mathbf{b} + \mathbf{c}) = f(\mathbf{a}, \mathbf{b}) + f(\mathbf{a}, \mathbf{c})$$

上述几何量将是用向量去研讨几何广泛有用的有力工具。在向量代数中，我们索兴把它想成是一种由两个向量求得一个数值的一种乘积，叫做向量的内积 (Inner Product) 而且改用符号 $\mathbf{a} \cdot \mathbf{b}$ 表达之，亦即以

$$(0.7') \quad \mathbf{a} \cdot \mathbf{b} = \frac{1}{2} \{ |\mathbf{a} + \mathbf{b}|^2 - |\mathbf{a}|^2 - |\mathbf{b}|^2 \}$$

为向量内积的定义式。这样做的基本原由就是使得性质 (0.8)-式可以直截了当地改写成

$$(0.8') \quad \mathbf{a} \cdot (\mathbf{b} + \mathbf{c}) = \mathbf{a} \cdot \mathbf{b} + \mathbf{a} \cdot \mathbf{c}$$

这种分配律的形式，使得它运用起来能够更加得心应手。

总结这一小节的讨论，我们可以用位移向量的加法、倍积和内积简洁自然地把空间的结构全面代数化，而且把定量几何学中的基本定理如相似三角形定理和勾股定理分别转化为倍积和内积的分配律。这也就是为什么分配律的代数又成为研讨几何学的利器的道理。

0.2.4 线性代数和轨尺作图

在古希腊几何学中，相传有三个待解的轨尺作图难题，一直找不到其解法：

1. 其一是：如何作一正立方体，其体积为一个给定正立方体的两倍？
2. 其二是：如何将一个给定角三等分？
3. 其三是：如何作一正方形和单位圆具有同等面积？

此事世代相传，一直无法解答，一直到近世才真相大白：原来三者都是无法用有限次轨尺作图来达成的！换句话说，古希腊几何学家所寻求的作图法，其实是根本不存在的！而前两个的不存在性的证明乃是运用线性代数的架构，把原本的几何问题转化成代数问题而得以解决的。此事乃是大学线性代数课程的讨论范围，在此则仅作一简明扼要的介绍。

[注]：详细证明请参看附录的讨论。

从定量几何的观点来说，第一个作图题所要达成者是由给定单位长去求作一个长度等于 $\sqrt[3]{2}$ 的直线段，而第二个作图题所要寻求者，则是如何由一个给定角 θ 的正弦 $\sin\theta$ （或其余弦 $\cos\theta$ ）去作出长度是 $\sin\frac{\theta}{3}$ （或 $\cos\frac{\theta}{3}$ ）的直线段。开宗明义，问题本身并非长度分别是 $\sqrt[3]{2}$ 和 $\sin\frac{\theta}{3}$ 的直线段是否存在，而是它们是否能用有限次的轨尺作图构造而得？由此可见，要证明上述两个作法之「不存在性」，当然就得对于那些可以用有限次轨尺作图构造而得的线段长有一种明确的总体认识！

【分析】：归根究底，在轨尺作图中，逐步所做者不外乎一次次去求 { 线、线 }，{ 线、圆 } 和 { 圆、圆 } 的交截。以近代解析几何的观点来表达，那就是求解它们的联立方程式的解点坐标。而其中每次所求解的方程组的系数，则是原给线段或前面逐次作图所得者的长度。稍加分析，就可以看到 { 线、线 } 交点的坐标总是可以用系数的四则运算

加以表达；而 { 线、圆 } 和 { 圆、圆 } 的交点坐标则可以用系数的四则运算加上平方根的表式来表达之。由此可见，那种可以用有限次轨尺作图求得的长度乃是可以由原给长度，以有限次四则运算和平方根的组合加以表达者。反之，任何能够用给定长度再用上述五种运算的有限次组合加以表达的长度，也都可以用有限次轨尺作图去逐步构造之！

【基本想法与具体做法】：

(一) 改用代数的描述。 $\sqrt[3]{2}$ 和 $\sin \frac{\theta}{3}$ 分别是下述三次方程式的根，即

$$\begin{aligned}x^3 - 2 &= 0 \\4x^3 - 3x + \sin \theta &= 0\end{aligned}$$

若将 θ 取为 $\frac{\pi}{6}$ ($= 30^\circ$)，则有

$$\begin{aligned}4x^3 - 3x + \frac{1}{2} &= 0 \\ \text{亦即} \quad \tilde{x}^3 - 3\tilde{x} + 1 &= 0, \quad \tilde{x} = 2x\end{aligned}$$

显然 $x^3 - 2$ 和 $\tilde{x}^3 - 3\tilde{x} + 1$ 都是有理系数不可约的三次多项式（因为它们显然没有有理根）。

概括地来说，我们所要论证者，就是像上述不可约三次多项式的根是无法改用有限次二次方程的求根和四则组合来表达的！换句话说，单个三次方程的求解就是无法改用任何有限次的各种各样二次方程求根和四则组合去替代之的。这种不可替代性强烈地显示了三次方程的根和二次方程的根肯定具有本质性的差别。唯有精确地认清两者之间的相异本质，才能严格论证这种不可替代性。

(二) 在上面所讨论的代数问题中，所涉及的三次多项式的不可约性是至关重要的。而多项式的不可约性则是要在系数域明确界定之下才能决定者也。设 F 是一个给定的系数域， α 是一个 F 上不可约二次多项式的根，则有

$$F(\alpha) = \{c_0 + c_1\alpha; c_0, c_1 \in F\}$$

构成一个新的域，而且它也是 F 上的一个 2-维线性空间。

再者，设 V 是一个以 $F(\alpha)$ 为其系数域的 m 维线性空间。设 $\{\mathbf{b}_i, 1 \leq i \leq m\}$ 是 V 的一组基底，则 V 中任给元素 \mathbf{x} 皆可唯一地表达为 $\{\mathbf{b}_i, 1 \leq i \leq m\}$ 的 $F(\alpha)$ -线性组合，即

$$\mathbf{x} = \sum_{i=1}^m \xi_i \mathbf{b}_i, \quad \xi_i = a_i + b_i \cdot \alpha, \quad a_i, b_i \in F$$

亦即

$$\mathbf{x} = \sum_{i=1}^m a_i \mathbf{b}_i + \sum_{i=1}^m b_i (\alpha \mathbf{b}_i)$$

所以，当我们把 V 想成是一个 F 上的线性空间时，则 V 中的任给元素皆可唯一地表达成 $\{\mathbf{b}_i, \alpha \mathbf{b}_i, 1 \leq i \leq m\}$ 的 F -线性组合。由此可见 V 是 F 上的 $2m$ -维线性空间，亦即有

$$\dim_F V = 2 \cdot \dim_{F(\alpha)} V$$

(三) 设 V 是由 F 起始，逐步添加二次方程的根扩张所得的域，即有

$$\begin{aligned} V &= F(\alpha_1)(\alpha_2) \dots (\alpha_k) \supseteq F(\alpha_1)(\alpha_2) \dots (\alpha_{k-1}) \\ &\supseteq \dots \supseteq F(\alpha_1)(\alpha_2) \supseteq F(\alpha_1) \supseteq F \end{aligned}$$

其中 α_j 是一个以 $F(\alpha_1) \dots (\alpha_{j-1})$ 为系数域的二次不可约多项式的根。直接用上述维数公式逐步推进就有

$$\dim_F V = 2^k$$

(四) 设 β 是一个以 F 为系数域的不可约三次多项式 $f(x)$ 的根。由上述所设可见 $\{1, \beta, \beta^2\}$ 线性无关，而 $\{1, \beta, \beta^2, \beta^3\}$ 则线性相关。再者，任何一个次数至多为 2 的多项式 $g(x)$ 当然是和 $f(x)$ 互素的，亦即 $f(x)$ 和 $g(x)$ 的最高公因式等于 1。由辗转相除求最高公因式可得关系式

$$1 = A(x)f(x) + B(x)g(x)$$

将 β 代入上式即得

$$1 = A(\beta)f(\beta) + B(\beta)g(\beta) = B(\beta)g(\beta)$$

亦即 $g(\beta)$ 的倒数也可以表达为 $1, \beta, \beta^2$ 的线性组合。所以

$$F(\beta) = \{c_0 + c_1\beta + c_2\beta^2, c_0, c_1, c_2 \in F\}$$

业已构成一个域，它显然是 F 上的一个 3-维线性空间。

(五) 最後我们要证明上述 β 不可能被包含在像前述 V 这种由 F 逐步添加有限个二次方程的根所扩张而得的域之中。假如不然，则 $V \supseteq F(\beta)$ 。所以 V 也是 $F(\beta)$ 上的一个线性空间，设其维数为 n ， $\{\mathbf{a}_i, 1 \leq i \leq n\}$ 是 V 的一组 $F(\beta)$ -基底。则 $\{\mathbf{a}_i, \beta\mathbf{a}_i, \beta^2\mathbf{a}_i, 1 \leq i \leq n\}$ 就构成了 V 的一组 F -基底，亦即

$$2^k = \dim_F V = 3 \cdot \dim_{F(\beta)} V = 3 \cdot n$$

这显然是一个矛盾，因为 2^k 不可能被 3 整除！

当 $\beta = \sqrt[3]{2}$ 时，上述论证说明了[问题一]是无解的；而当 $\beta = \sin 10^\circ$ 时，上述论证说明了[问题二]在 $\theta = 30^\circ$ 时已经是无解的了。

0.2.5 研讨代数学的几个基本方法

基于代数问题的本质和代数学的基本思想，就自然产生了下列几种常用、有用的基本方法，它们乃是在代数学的研讨和各种各样代数问题的解法中贯穿全局，经常有用者。

1. 归纳发现、归纳论证与唯一性

几何学与代数学都是研究大自然的有力工具，但是两者的研讨方法却有著本质上的差异。在几何学的研讨中，我们有著相当可靠的几何直观 (geometric intuition) 的引导，往往在讨论初期便已得知结论的梗概，而且几何直观往往又能带领著我们逐步迈向所需的结论。相比之下，在代数学的研讨中，虽然我们具备各种各样有效能算的运算律，但是缺乏相应的「代数直观」的引导，我们所能「见」者并不甚远。这便是几何学和代数学两者在研讨方法上的一种重要差别。所以，在代数学的研讨中，寻找目标以作「有的放矢」乃是十分重要的一步，而这个「寻找『的』用以放矢」的工作，便是同学们曾接触过的归纳法，也是我们在第一节「数系运算律」的讨论中所用过的方法。

在中学的课程中，归纳法的应用往往是在于「归纳论证」的部分。其实由上面一小段的分析，大家可看到归纳法的要点其实在于「归纳发现」这一部分，而「归纳论证」只是用来确立「归纳发现」之所得者的理据。所以，在一个完整的归纳法中，理应可以分为下述三个部分：

- 一、 归纳地发现具有某种有用特性的事物 (inductive discovery)；
- 二、 归纳地定义该事物 (inductive definition)；
- 三、 归纳地证明上述归纳定义者的确具有所应有之特性 (inductive proof)。

所以若只著重「归纳证明」而忽略了「归纳发现」的讨论可以说是有点舍本逐末，断章取义的教学。在往後研讨线性方程的基础理论中，我们将以行列式的归纳发现、归纳定义和归纳论证为例来体现一个完整的归纳法的具体做法和用场。

在此，同学们可能会问：「如果采用另一种途径来归纳发现、定义上述行列式，会否得到另一种的『行列式』？」这个便是所谓的「唯一性」问题。在代数学的研讨中，我们会用五花八门的方法来寻找不同类型的公式，当然有可能会出现上述不同方法得出不同公式的情况。所以，在发现公式及论证其正确性之後，我们应当问一问这种「唯一性」的问题，亦即所求的公式是否业已由用来寻找它的特性所「唯一」地确定。大家要注意在这里的「唯一」并不只限于「一个」，而是包括「一类」、「一系列」等等较为广泛的意思，例如某组合所得者必为某类常数等等。我们将会证明行列式会由三个条件所唯一地确定，这个结果我们称之为「行列式的界定定理」。行列式的界定定理的用途广泛多样，例如就可以用于证明有向面积或有向体积其实就是行列式这个结果。

其实，代数学中很多重要的公式和定理，都是由低次到高次、由少元到多元逐步归纳发现，然後再行归纳论证其普遍性，确立其唯一性。

2. 线性结构和线性化

韩信点兵法启示我们如何善用分解组合，把所待解的线性问题系统地归于某一组易解的特殊解的线性组合而获得其通解。由此可见，

若能把多种多样的代数问题（或几何、分析问题）妥加组织，适当地把它们转化成线性问题，则韩信点兵法就可以大行其道了。在 §0.2.3 和 §0.2.4 中所讨论的乃是其中几个重要的实例。

3. 待定系数法

好些代数问题，可以适当地引进待定系数，把它归于解方程的理论来研讨之。这是代数学研讨中常用、好用的一种手法。下面所举者乃是一些简单的实例，暂且作为其初步简介之用。

在平面上，相异两点定一直线，不共线三点定一圆；这些都是很基本的几何事实。在坐标几何的讨论中，我们经常要找出这些线和圆的坐标方程，以便作后续的讨论和计算。但如果当中某些点含有未知的坐标时，求解上述的坐标方程便很麻烦了，例如要把某个系数是否为零的情况分开来讨论。但无论如何，直线方程必定是 $Ax + By + C = 0$ 这种样子，圆的方程必定是 $A(x^2 + y^2) + Dx + Ey + F = 0$ 这种样子，是不？

待定系数法的精神就是假设所求的公式就是上述样子，然后用已给的资料列出相应的方程组，最后把原来问题转化成相应的方程组有否某种特定解而取代之。例如，设直线 $Ax + By + C = 0$ 通过给定（相异）两点 (x_1, y_1) 和 (x_2, y_2) ，则有

$$\begin{cases} Ax_1 + By_1 + C = 0 \\ Ax_2 + By_2 + C = 0 \end{cases}$$

若 (x, y) 是在线上的另给一点，则还有 $Ax + By + C = 0$ ，即

$$\begin{cases} Ax_1 + By_1 + C = 0 \\ Ax_2 + By_2 + C = 0 \\ Ax + By + C = 0 \end{cases}$$

由几何结论我们知道应可求出 A, B, C 的一个非平凡解，即 A, B, C 不全为 0 的一组解。所以，若把上述方程组想成以 A, B, C 为变元的线性方程组：

$$\begin{cases} x_1 \cdot A + y_1 \cdot B + 1 \cdot C = 0 \\ x_2 \cdot A + y_2 \cdot B + 1 \cdot C = 0 \\ x \cdot A + y \cdot B + 1 \cdot C = 0 \end{cases}$$

则 A, B, C 不可能只有唯一解 $A = B = C = 0$ ，亦即其系数的行列式必须为 0（详见第四章关于行列式的讨论），所以得出

$$\begin{vmatrix} x_1 & y_1 & 1 \\ x_2 & y_2 & 1 \\ x & y & 1 \end{vmatrix} = 0$$

这便是过 (x_1, y_1) 和 (x_2, y_2) 两点的直线方程。

同理，过不共线三点 $(x_1, y_1), (x_2, y_2)$ 和 (x_3, y_3) 的圆方程 $A(x^2 + y^2) + Dx + Ey + F = 0$ 就是

$$\begin{vmatrix} x_1^2 + y_1^2 & x_1 & y_1 & 1 \\ x_2^2 + y_2^2 & x_2 & y_2 & 1 \\ x_3^2 + y_3^2 & x_3 & y_3 & 1 \\ x^2 + y^2 & x & y & 1 \end{vmatrix} = 0$$

总结本章对于代数学的绪论，可见代数问题的范畴其实是既广且深的。但是代数学的基本精神则是一贯的，亦即有系统、有效地运用运算律。当然，随着代数问题的深化和多样化，这种善用运算律的代数方法也就自然而然地深化和多样化。再者，为了把代数方法用来解决更多彩多姿的数学问题，也自然地产生了更多样、更精细的代数结构。其中最为简朴常用的就是线性结构，和它们的研讨所展现的线性代数学。这也是大学基础代数的一个中心课题。

0.3 例题、习题与思考题

0.3.1 数系运算律与数系扩张

(1) 自然数系运算律的归纳论证：加、乘、乘方的归纳定义乃是论证之所本，即

$$a + (n + 1) = (a + n) + 1$$

$$(n + 1) \cdot a = n \cdot a + a$$

$$a^{(n+1)} = a^n \cdot a$$

基于上述归纳定义式，即可逐步归纳证明自然数系的运算律，即加、乘的交换律与结合律，分配律与指数定则。在 §0.1 中业已列述了加法结合律，分配律和 $a^m \cdot a^n = a^{m+n}$ 的归纳证明。同学们应该以此为例进而逐步归纳论证其他各个运算律。请注意，这个奠基性的逐步论证运算律，在顺序上是很有考究的。

(2) 由自然数系到整数系的扩张及整数系运算律的验证：整数系比自然数系添加了 0 和负整数 $\{-n, n \in \mathbb{N}\}$ ，其特征性质分别是

$$a + 0 = a \quad \text{和} \quad (-n) + n = 0$$

试分析整数系的加、乘运算的定义及其运算律之普遍成立，并从而认清两者之间的密切相关。

(3) 由整数系到分数系的扩张及分数系运算律的验证：分数 $\frac{m}{n}$ ，其中 m, n 是整数而且 $n \neq 0$ ，它的特征性质就是

$$n \cdot \frac{m}{n} = m, \quad \text{而且} \quad \frac{m}{n} = \frac{p}{q} \Leftrightarrow m \cdot q = n \cdot p$$

试分析分数系的加、乘运算的定义及其运算律的普遍成立，并从而认清两者之间的密切相关。

0.3.2 辗转相除法与算术基本定理

(1) 带余除法和辗转相除法：给定一个正整数 m ，我们暂时还没有一个有效能算的方法来寻找其 1 和 m 以外的因数。因为这个分解过程没有有效计算方法，现今常用的密码就是用到这个计算上的盲点来编制的。但是若给定一对正整数 $n < m$ ，我们就有熟知的带余除法，亦即存在（唯一的）整数 q_1 和 $0 \leq r_1 < n$ ，使得

$$m = q_1 \cdot n + r_1$$

若 $r_1 = 0$ 则 n 是 m 的一个因数，否则即不是。在此还可以看出一个好用的事实，即 (m, n) 的最大公因数等同于 (n, r_1) 的最大公因数！同学们可尝试验证之。

若 $r_1 > 0$ ，我们可以再用带余除法于 (n, r_1) ，则有

$$\begin{aligned} n &= q_2 \cdot r_1 + r_2 \\ r_1 &= q_3 \cdot r_2 + r_3 \\ &\vdots \\ r_{k-3} &= q_{k-1} \cdot r_{k-2} + r_{k-1} \\ r_{k-2} &= q_k \cdot r_{k-1} + r_k \\ r_{k-1} &= q_{k+1} \cdot r_k \end{aligned}$$

直至 $r_{k+1} = 0$ 为止。由此可见 (m, n) 的最大公因数就是 r_k ！上述运算我们称之为「辗转相除法」，而在西方则称之为“Euclidean Algorithm”，因为它写在 Euclid 的“Elements”中。

我们还可以运用上述各式把 r_k 重写成之前的 r_i 的倍数组组合，即

$$\begin{aligned} r_k &= r_{k-2} - q_k r_{k-1} = r_{k-2} - q_k(r_{k-3} - q_{k-1} r_{k-2}) \\ &= -q_k r_{k-3} + (1 - q_k q_{k-1}) r_{k-2} \end{aligned}$$

如此逐步反代，最後就可以把 r_k 重写成 m 和 n 的倍数组组合，即

$$r_k = A \cdot m + B \cdot n$$

其中 A, B 是适当的整数。由此可见 (m, n) 的最大公因数 r_k 其实可以写成 (m, n) 的倍数组组合，易见它就是在其中的最小正整数（试证之）。

(2) 正整数的因数分解与算术基本定理：当我们从某些途径找到整数 n 的一个因数 $1 < k < n$ 时，我们当然可以把 n 分解成 $n = k \cdot \ell$ 。如此分解下去，我们最後就可以把 n 写成

$$n = k_1 \cdot k_2 \cdots k_m$$

其中每个 $k_i > 1$ 都只有 1 和 k_i 才是自己的因数。我们称这些 k_i 为 n 的质因数。一个自然的问题就是上述分解是否唯一呢（除了那些质因数的次序外）？稍加分析後，我们发现其证明的重点引理就是：

【引理】：若一个质数 p 可以整除整数乘积 ab ，则它至少可以整除 a 或 b 其中之一。

同学们不妨验证由此即得上述唯一分解性的结论。但是如何证明这个引理呢？原来就是用到 (1) 的结论。如今反观，基本上别无他法。

引理证明：设 p 不能整除 a ，则需要证明 p 必须整除 b 。因为 (p, a) 的最大公因数为 1，由 (1) 的结论知道，存在整数 A, B 使得 $1 = Ap + Ba$ 。在等式两旁乘上 b ，则有

$$b = Apb + Bab$$

因为上式右方可以被 p 所整除，所以 p 必须整除 b 。□

【算术基本定理】 (Fundamental Theorem of Arithmetics)：给定一个整数 $n > 1$ ，存在有唯一的质数列 $p_1 < p_2 < \dots < p_r$ 和指数列 $(\alpha_1, \alpha_2, \dots, \alpha_r)$ ，其中 $\alpha_i \geq 1, 1 \leq i \leq r$ ，使得

$$n = p_1^{\alpha_1} \cdot p_2^{\alpha_2} \cdots p_r^{\alpha_r}$$

0.3.3 多项式基本公式举例

(1) $(x-1)(x+1) = x^2 - 1, (x-1)(x^2+x+1) = x^3 - 1$ 的推广：

- (i) 试求 $q(x)$ 使得 $(x-1)q(x) = x^n - 1$ 。
- (ii) 试求 $g(x)$ 使得 $(x-a)g(x) = x^n - a^n$ 。
- (iii) 试证 $f(x) - f(a)$ 含有因式 $(x-a)$ 。
- (iv) 若 $f(x)$ 在相异的三个数 a_1, a_2, a_3 的值为零，试证 $f(x)$ 含有因式 $(x-a_1)(x-a_2)(x-a_3)$ 。
- (v) 上述事实是否可以推而广之？其叙述为何？
- (vi) 试问一个 n 次多项式是否能够有多于 n 个相异的「根」（亦即使得 $f(x)$ 在其上取值为零者）？试证明你的论断。

(2) 二项展开公式的归纳构造

- (i) 试运用分配律，把 $(x+y)^n$ 归纳地展开成单项式之和，其中 $n = 2, 3, 4, \dots, 10$ 。

(ii) 把上题展开式所得的系数重新排列成如下的三角阵式：

$$\begin{array}{rcccc} n = 0 & & & & 1 \\ n = 1 & & & 1 & 1 \\ n = 2 & & & 1 & 2 & 1 \\ n = 3 & & & 1 & 3 & 3 & 1 \\ & & & \vdots & & & \\ n = 10 & & & 1 & 10 & \cdots & 10 & 1 \end{array}$$

试找出相邻两层系数之间的关系。

第一章

多项式的基础理论

在各种各样的代数问题中，我们总是运用各种代数运算（如加法、乘法等等）来分析量与量之间的代数关联。在运算过程中，运算律的普遍成立乃是常用好用的要点。这些运算律的普遍性让我们可以有系统、有效地分析所给代数问题中未知量和已知量之间的关联，从而「化未知为已知」。由于我们常用的数系运算律（如分配律、指数定则等等）对于所有数字皆普遍成立，所以其做法都可以广泛地应用到任何一个只需用到那些数系运算律的代数系统（即可以假设所处理的符号满足数系通性）。同学们在初中时所学习的多项式代数，就是上述做法的一个典型例子。

1.1 多项式运算

在这里我们首先列出一些关于多项式的基本事实，作为同学们对于多项式代数的简要复习。

1. 一个满足所有数系运算律的符号我们称之为「不定元」(indeterminant)，通常以 x, y, z 等符号表示之（亦即 x, y, z 等符号的确实意义还有待确定；但无论如何，它们所代表的量必定会满足所有数系运算律的）。将有限个数字（称之为「系数」(coefficients)）和不定元以加法和乘法结合起来的代数表达式，就是所谓的「多项式」(polynomial)。例如：

- (i) 单项式 (monomial) 如 $x^5, \sqrt{2}x^3, -5x^2y^3, (\sqrt{2} + \sqrt{3})xyz^2, ax^\ell y^m z^n$ 等都是多项式的特例。
- (ii) $(x-a), x^3+ax^2+a^2x+a^3, (x-a)\cdot(x^3+ax^2+a^2x+a^3), x+\sqrt{2}y-\sqrt{3}z, (x+y+z)(x^2+y^2+z^2-xy-yz-zx), x^3+y^3+z^3-3xyz, (x+y)^5, (x+y)^7$ 等等都是多项式。
- (iii) 运用运算律 (尤其是分配律), 我们可以有系统地把任何多项式展开重组, 把它重新表达成单项式之和。例如:

$$\begin{aligned} (x-a)(x+a) &= x^2 - a^2, \\ (x-a)(x^2+ax+a^2) &= x^3 - a^3, \\ (x-a)(x^3+ax^2+a^2x+a^3) &= x^4 - a^4, \\ (x+y+z)(x^2+y^2+z^2-xy-yz-zx) &= x^3+y^3+z^3-3xyz, \\ (x+y)^4 &= x^4+4x^3y+6x^2y^2+4xy^3+y^4, \\ (x+y)^5 &= x^5+5x^4y+10x^3y^2+10x^2y^3+5xy^4+y^5 \text{ 等}; \\ (a_0+a_1x+a_2x^2)(b_0+b_1x+b_2x^2) &= a_0b_0+(a_0b_1+a_1b_0)x+ \\ &\quad +(a_0b_2+a_1b_1+a_2b_0)x^2+(a_1b_2+a_2b_1)x^3+a_2b_2x^4 \end{aligned}$$

2. 在只有单个不定元的情况之下, 任何多项式都能以该不定元的幂数 (powers 或 exponents) 由低幂至高幂, 或由高幂至低幂顺序排列来表达之, 即:

$$a_0 + a_1x + a_2x^2 + \dots + a_nx^n \quad \text{或} \quad a_nx^n + a_{n-1}x^{n-1} + \dots + a_1x + a_0$$

前者称为升幂表达式 (ascending exponent normal form), 而後者则称为降幂表达式 (decending exponent normal form)。在上述两种表达式中, 某些系数 a_i 可以是 0 的, 而在习惯上我们总是会假设最高幂数的系数不等于 0 (即上式中的 $a_n \neq 0$)。

3. 一个单项式的「次数」(degree) 定义为其 (所有) 不定元的指数之总和。例如: $\frac{1}{10}x^9, 5^{10}x^2, -x^3y^2$ 和 $ax^\ell y^m z^n$ 的次数分别就是 9, 2, 5 和 $(\ell+m+n)$ 。而一个多项式 f 的次数则定义为实质出现在其表达式中的单项式之最高次数, 记为 $\deg f$, 如:

$$\begin{aligned} f(x) &= a_0 + a_1x + \dots + a_mx^m, \quad a_m \neq 0, \quad \deg f(x) = m; \\ g(x) &= b_0 + b_1x + \dots + b_nx^n, \quad b_n \neq 0, \quad \deg g(x) = n \end{aligned}$$

易见 $\deg(f(x) \cdot g(x)) = m + n$ ，因为 $f(x) \cdot g(x)$ 包含单项式 $(a_m \cdot b_n)x^{m+n}$ ，而这个单项式的次数显然是在 $f(x) \cdot g(x)$ 之中最高的。因此，下述公式是显而易见的：

$$(1.1) \quad \deg(f(x) \cdot g(x)) = \deg f(x) + \deg g(x)$$

[注意]：一个非零常数多项式的次数我们定义为 0，而「零多项式」的次数则定义为「 $-\infty$ 」。采用这个定义，则 (1.1)-式在下述特殊情况下亦可言之成理，即

$$(1.2) \quad \deg(0) = \deg(0 \cdot g(x)) = \deg(0) + \deg g(x)$$

仍然成立，因为在一般的理解之下「 $-\infty = -\infty + \text{任何有限整数}$ 」。

然而对于多项式的加法，我们并没有特定的公式从 $\deg f(x)$ 和 $\deg g(x)$ 来计算 $\deg(f(x) + g(x))$ 。在一般的情形我们至多只可以说

$$(1.3) \quad \deg(f(x) + g(x)) \leq \max\{\deg f(x), \deg g(x)\}$$

而且等式在 $\deg f(x) \neq \deg g(x)$ 时总是成立。

[注]：当 $f(x)$ 与 $g(x)$ 两者的最高次数的单项式互相抵消时，不等式“ $<$ ”就会成立。

4. 令

$$\begin{aligned} f(x) &= a_m x^m + a_{m-1} x^{m-1} + \dots + a_0 \\ g(x) &= b_n x^n + b_{n-1} x^{n-1} + \dots + b_0 \end{aligned}$$

并假设 $\deg f(x) = m > n = \deg g(x)$ ，则易见

$$(1.4) \quad \deg \left(\left(f(x) - \frac{a_m}{b_n} x^{m-n} \cdot g(x) \right) \right) < m = \deg f(x)$$

因为 $f(x)$ 与 $\frac{a_m}{b_n} x^{m-n} g(x)$ 两者的最高次数单项式刚好互相抵消。其实上述运算就是我们惯用常做的多项式长除法的步骤！只需重复地做上述运算，我们最终就可得出一个（唯一的） $g(x)$ 的多项式倍式，且以 $q(x) \cdot g(x)$ 记之，使得

$$(1.5) \quad \begin{aligned} &\deg(f(x) - q(x) \cdot g(x)) < \deg g(x), \quad \text{或} \\ f(x) &= q(x) \cdot g(x) + r(x), \quad \deg r(x) < \deg g(x) \end{aligned}$$

我们通常把 (1.5)-式称之为「多项式的带余除法」(division algorithm for polynomials)。

[注意]: 上述带余除法只适用于单元多项式的范畴! 多元多项式是没有这种带余除法的!

1.2 多项式函数

在一个系统中, 当一个量的值唯一地确定另一个量的值时, 我们称后者 (因变量, dependent variable) 为前者 (自变量, independent variable) 的「函数」(function)。例如, 圆盘面积和球体体积都是其半径的函数, 而在各自的公式中, 即:

$$(1.6) \quad A = \pi r^2 \quad \text{或} \quad V = \frac{4}{3}\pi r^3$$

它们之间的函数关联是可以由一些明确的运算来确定。一般地说, 若 y 的值是由 x_1, \dots, x_n 这 n 个量之值所唯一确定, 则称 y 为 x_1, \dots, x_n 的函数。例如: 一个三角形的面积是其三边边长的函数, 而一个矩形的面积则是其长和宽的函数。它们的明确关系式则分别是:

$$A(\triangle) = \frac{1}{4}\sqrt{(a+b+c)(a+b-c)(a-b+c)(-a+b+c)},$$

$$A(\square) = \ell \cdot w$$

在往后的讨论中我们将会集中讨论一系列比较特殊的函数, 其中每一个函数的明确表达式是可以由某一个特定的多项式来表达者, 即:

$$(1.7) \quad y = c_0 + c_1x + \dots + c_nx^n$$

这类函数我们称之为「多项式函数」(polynomial function), 往后我们将以 $y = f(x)$ 来表达由多项式 $f(x)$ 所决定的多项式函数 y , 并以 $f(a)$ 代表当 $x = a$ 时 y 所对应的值, 一般称作「 y (或 $f(x)$) 在 (位置) $x = a$ 的值」。

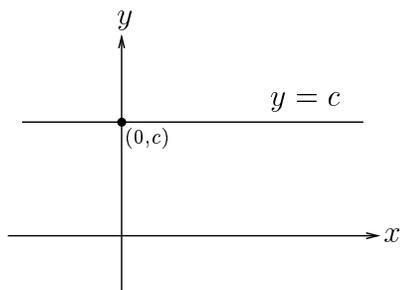
若我们让 x 在实数线上随意变动, 则在坐标平面上所有以 $(x, f(x))$ 为坐标的点就组成一条曲线。这条曲线就是以几何方式来表达函数 $y = f(x)$, 我们称之为 $y = f(x)$ 的「图象」, 即:

$$(1.8) \quad f(x) \text{ 的图象} = \{(x, f(x)); x \in \mathbb{R}\}$$

下面是一些基本的多项式函数的图象的例子。

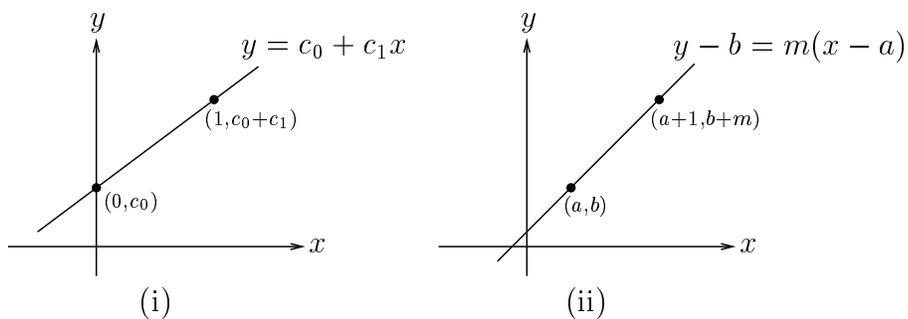
【例子】：

(i) $\deg f(x) = 0 : y = c$ (c 为常数)



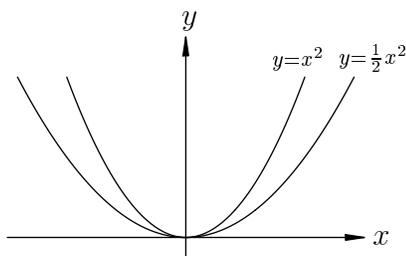
[图 1-1]

(ii) $\deg f(x) = 1 : y = c_0 + c_1x$ 或 $y - b = m(x - a)$



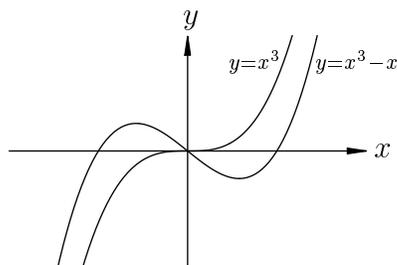
[图 1-2]

(iii) $\deg f(x) = 2 : y = kx^2, k = 1, \frac{1}{2}$



[图 1-3]

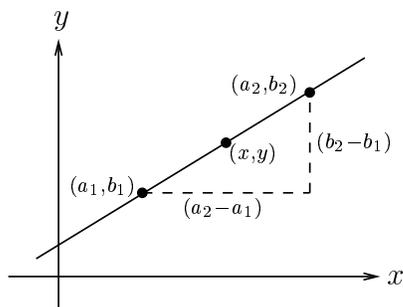
(iv) $\deg f(x) = 3 : y = x^3, y = x^3 - x$



[图 1-4]

根、值和插值问题：

显然一个 1 次（或 0 次）的多项式函数是可被其在两个（或一个）位置的值所唯一地确定。因为 1 次多项式函数的图象是一条直线，所以上述现象就是对应于「相异两点定一直线」这个众所周知的几何事实。



[图 1-5]

例如，若 1 次多项式函数 $y = f(x)$ 满足 $b_1 = f(a_1)$ 和 $b_2 = f(a_2)$ ，则其表达式为：

$$(1.9) \quad (y - b_1) = \frac{b_2 - b_1}{a_2 - a_1}(x - a_1)$$

上述简单事实可以自然而然地推广为下述唯一性问题，即：

【问题一】：一个 n 次多项式函数是否由其在 $(n+1)$ 个相异位置的值所唯一决定？换句话说，若 $f(x)$ 和 $g(x)$ 为两个次数至多为 n 的多项式，它们在 $(n+1)$ 个相异位置 $\{a_i; 0 \leq i \leq n\}$ 的值相同，即 $f(a_i) = g(a_i)$, $0 \leq i \leq n$ ，问 $f(x)$ 和 $g(x)$ 是否必然相等？

令 $h(x) = f(x) - g(x)$ ，则易见 $h(x)$ 的次数不高于 n ，而且 $h(a_i) = f(a_i) - g(a_i) = 0, 0 \leq i \leq n$ 。所以，若我们可以证明任何一个次数不高于 n 的非零多项式最多只能有 n 个根（roots，即 h 取 0 为值的位置），则由此可见 $h(x)$ 必为零多项式，并给出了上述[问题一]的一个正面的解答。

【定理 1.1】：一个 n 次非零多项式最多只有 n 个根。再者，一个次数不高于 n 的多项式函数是可以由其在 $(n+1)$ 个相异位置的值所唯一确定。

证明：容易验证 $c_k(x^k - a^k) = (x - a) \cdot c_k(x^{k-1} + ax^{k-2} + \dots + a^{k-1})$ ，所以

$$\begin{aligned} f(x) - f(a) &= [c_0 + c_1x + \dots + c_kx^k + \dots + c_nx^n] \\ &\quad - [c_0 + c_1a + \dots + c_ka^k + \dots + c_na^n] \\ &= c_1(x - a) + c_2(x^2 - a^2) + \dots + c_k(x^k - a^k) + \dots + c_n(x^n - a^n) \\ &= (x - a) \cdot q(x) \end{aligned}$$

其中

$$\begin{aligned} q(x) &= c_1 + c_2(x + a) + \dots + c_k(x^{k-1} + ax^{k-2} + \dots + a^{k-1}) \\ &\quad + \dots + c_n(x^{n-1} + ax^{n-2} + \dots + a^{n-1}) \end{aligned}$$

所以， a 是 $f(x)$ 的一个根的充要条件为 $f(x) = (x - a) \cdot q(x)$ ，亦即 $(x - a)$ 是 $f(x)$ 的一个因式。如此逐步地做下去，即得当 $\{a_1, \dots, a_k\}$ 是 $f(x)$ 的 k 个相异根时，则 $(x - a_1) \dots (x - a_k)$ 乃是 $f(x)$ 的一个因式，亦即

$$(1.10) \quad f(x) = (x - a_1) \dots (x - a_k) \cdot g(x)$$

因此，若一个 n 次多项式 $f(x)$ 已具有 n 个相异根 $\{a_i, 1 \leq i \leq n\}$ ，则可将 $f(x)$ 写成

$$(1.11) \quad f(x) = (x - a_1) \dots (x - a_n) \cdot g_n(x)$$

其中 $\deg g_n(x) = 0$ ，亦即 $g_n(x)$ 其实是一个非零常数（即 $g_n(x) = c_n$ ）。所以，若 a 为 $\{a_i, 1 \leq i \leq n\}$ 之外的任何一点，则有

$$(1.12) \quad f(a) = (a - a_1) \cdot (a - a_2) \dots (a - a_n) \cdot c_n, \quad (a - a_i) \neq 0, 1 \leq i \leq n$$

亦即 $f(a)$ 等于 $(n+1)$ 个非零数值之乘积，当然不会为零；所以一个 n 次非零多项式最多只有 n 个根。由此亦可推得一个次数不高于 n 的多项式函数是可以由其在 $(n+1)$ 个相异位置的值所唯一确定。 \square

上述「唯一性」定理很自然地带出下述「存在性」问题，即：

【问题二】：设 $\{a_i; 0 \leq i \leq n\}$ 为任意给定的 $(n+1)$ 个相异位置。令 $\{b_i; 0 \leq i \leq n\}$ 为任意给定的 $(n+1)$ 个值（不必相异）。问是否存在一个次数不高于 n 的多项式函数 $y = f(x)$ 使得 $f(a_i) = b_i, 0 \leq i \leq n$ ？再者，若这个多项式函数是存在的（由[定理 1.1]知道此函数必定是唯一的），如何去求得这个函数呢？

（上述乃是一个非常重要的插值问题，而其完满解答则是韩信点兵法的直接推广，详见下一节的讨论。）

【习题】：

- (1) 对于下列每一对的 $\{f(x), g(x)\}$ ，运用除法算式找出相应的 $q(x)$ 和 $r(x)$ 使得：

$$f(x) = q(x) \cdot g(x) + r(x), \quad \deg r(x) < \deg g(x)$$

(i) $f(x) = 5x^5 - 4x^4 + 3x^3 - 2x^2 + x - 1, g(x) = x - 1$

(ii) $f(x)$ 同上， $g(x) = x^2 - x + 1$

(iii) $f(x) = x^7 + x^4 - 10x^2 - 1, g(x) = x^2 + x + 1$

(iv) $f(x) = x^6 - 2x^4 + x^2 - 1, g(x) = x^2 + 1$

- (2) 把下列的 2 次多项式函数重写成 $(y - b) = k(x - a)^2$ 的模样，并描绘出它们的图象：

(i) $y = 3x^2 - 6x + 8$

(ii) $y = -x^2 + \sqrt{7}x - 13$

(iii) $y = \sqrt{7}x^2 - \sqrt{21}x + 10$

- (3) 在上题中所列出的三个 2 次多项式函数的极大值（或极小值）分别是多少？分别在那里出现？

(4) 把下列的 3 次多项式函数重写成 $(y - b) = k(x - a)^3 + m(x - a)$ 的模样：

$$(i) y = x^3 - 6x^2 + x - 7$$

$$(ii) y = x^3 + 6x^2 - 10x + 1$$

$$(iii) y = 5x^3 - 15x^2 + 10x + 7$$

(5) 试求公式计算下述等差级数之总和，即：

$$a + (a + d) + (a + 2d) + \dots + (a + (n - 1)d) = ?$$

(6) 试求公式计算下述等比级数之总和，即：

$$a + ar + ar^2 + \dots + ar^{n-1} = ?$$

1.3 韩信点兵法和插值公式

在绪论中所提到的「韩信点兵法」，乃是有系统地把馀数问题归结到一组简朴的特殊解而加以组合之，是一种善用分配律的解题方法。现在让我们先重温一下「韩信点兵法」的基本想法：

在求解剩馀问题时，当馀数之中只有一个是 1，其他皆为 0 的特殊情形，不但容易解答，而且可以用这一系列特殊解，把一般情形的解答简洁地用下述公式表达之。设 x_1, x_2, \dots, x_k 是对于给定一组公因数为 1（亦即互质）的除数 $\{a_1, a_2, \dots, a_k\}$ ，其馀数组分别是 $(1, 0, \dots, 0)$, $(0, 1, 0, \dots, 0)$, \dots , $(0, \dots, 0, 1)$ 的特殊解，则

$$(1.13) \quad x = r_1x_1 + r_2x_2 + \dots + r_kx_k$$

乃是一个馀数组是 (r_1, r_2, \dots, r_k) 的解。而且任何以 (r_1, r_2, \dots, r_k) 为馀数组的解和上述 x 相差一个 $a_1a_2 \dots a_k$ 的整数倍。上述公式 (1.13) 之所以普遍成立的理由就是分配律。由分配律可见 r_ix_i 被 a_i 除的馀数是 r_i 而且 $x - r_ix_i = \sum_{j \neq i} r_jx_j$ 被 a_i 除的馀数是 0（因为每一个 r_jx_j , $j \neq i$, 都含有因子 a_i ）。由此可见，上述算法的基本思想就是善用分配律，把剩馀问题的一般情形，直截了当地归于易解好算的特殊情形来系统解答之。

在数论里，这是一个很重要的基本定理，称作「中国剩馀定理」。现在让我们再次运用这种大巧若拙的算法用来研究多项式函数的插值问题。

插值公式 (interpolation formula) :

为简化讨论起见，我们考虑 $n = 3$ 作为典型情况讨论之；由于解法是可以作明显的推广（就像馀数问题一样），我们可以沿用同样方法求得在一般情况下的结果。

设 $\{a_0, a_1, a_2, a_3\}$ 为四个相异位置， $\{b_0, b_1, b_2, b_3\}$ 为四个任意给定数值。则在上一节所提出的[问题二]即为：是否存在一个次数不高于 3 的多项式函数 $y = f(x)$ ，它在 $\{a_0, a_1, a_2, a_3\}$ 上的值恰好就是 $\{b_0, b_1, b_2, b_3\}$ ？现在我们就用「韩信点兵法」的想法来写下上述所求的多项式 $f(x)$ ，直截了当地说明其存在性。

注意当 $f(x)$ 被除以 $(x - a_i)$ 时，其馀数就是 $f(a_i) = b_i, 0 \leq i \leq 3$ （详见[定理 1.1]的证明）。因此上述插值问题只是「馀数问题」的直接推广，即由整数除法推广至多项式除法，而现在的除式就是 $\{(x - a_i); 0 \leq i \leq 3\}$ 。所以，我们的祖先业已教导我们应先讨论一些特殊的情况，即 $\{b_0, b_1, b_2, b_3\}$ 分别为 $(1, 0, 0, 0)$ 、 $(0, 1, 0, 0)$ 、 $(0, 0, 1, 0)$ 、 $(0, 0, 0, 1)$ 这四种情况。令 $f_i(x), 0 \leq i \leq 3$ 分别为具有上述馀数集的 3 次多项式。这样， $f_0(x)$ 乃是 $(x - a_1)(x - a_2)(x - a_3)$ 的倍式，亦即：

$$(1.14) \quad f_0(x) = c(x - a_1)(x - a_2)(x - a_3)$$

由假设 $f_0(a_0) = 1$ 得：

$$(1.15) \quad 1 = f_0(a_0) = c(a_0 - a_1)(a_0 - a_2)(a_0 - a_3)$$

因此 $c = [(a_0 - a_1)(a_0 - a_2)(a_0 - a_3)]^{-1}$ ，亦即：

$$(1.16) \quad f_0(x) = \frac{(x - a_1)(x - a_2)(x - a_3)}{(a_0 - a_1)(a_0 - a_2)(a_0 - a_3)}$$

同理可得：

$$\begin{aligned} f_1(x) &= \frac{(x-a_0)(x-a_2)(x-a_3)}{(a_1-a_0)(a_1-a_2)(a_1-a_3)}; \\ f_2(x) &= \frac{(x-a_0)(x-a_1)(x-a_3)}{(a_2-a_0)(a_2-a_1)(a_2-a_3)}; \\ f_3(x) &= \frac{(x-a_0)(x-a_1)(x-a_2)}{(a_3-a_0)(a_3-a_1)(a_3-a_2)} \end{aligned}$$

到此阶段，前述的插值问题的解答已是「呼之欲出」，即：

$$(1.17) \quad f(x) = b_0 f_0(x) + b_1 f_1(x) + b_2 f_2(x) + b_3 f_3(x)$$

就是我们所要的多项式。上述想法显然是可以推广至任意 $(n+1)$ 个相异位置 $\{a_i, 0 \leq i \leq n\}$ 的情况，且让我们将其结论写成下述定理：

【定理 1.2】（拉格朗日插值公式，Lagrange interpolation formula）：设 $\{a_i; 0 \leq i \leq n\}$ 为任意给定的相异位置， $\{b_i; 0 \leq i \leq n\}$ 为任意给定的数值（不必相异），则存在唯一的多项式 $f(x)$ ，其次数不高于 n ，使得 $f(a_i) = b_i, 0 \leq i \leq n$ 。再者， $f(x)$ 的明确表达式就是：

$$(1.18) \quad f(x) = b_0 f_0(x) + b_1 f_1(x) + \dots + b_n f_n(x) = \sum_{i=0}^n b_i f_i(x)$$

其中

$$f_0(x) = \frac{(x-a_1)(x-a_2)\dots(x-a_n)}{(a_0-a_1)(a_0-a_2)\dots(a_0-a_n)} = \prod_{i=1}^n (x-a_i) / \prod_{i=1}^n (a_0-a_i)$$

.....

$$f_j(x) = \prod_{i \neq j} (x-a_i) / \prod_{i \neq j} (a_j-a_i)$$

.....

$$f_n(x) = \prod_{i=0}^{n-1} (x-a_i) / \prod_{i=0}^{n-1} (a_n-a_i)$$

[注] :

(i) 推导上述公式的方法和我们原先在讨论 $n = 3$ 时的方法其实是完全一样的；而唯一有别的地方就是要改用和、积记号如 $\sum_{i=0}^n$ 、 $\prod_{i \neq j}$ 。

(ii) $\{a_i; 0 \leq i \leq n\}$ 为相异的条件保证了每个出现在公式分母中的乘积 $\prod_{i \neq j} (a_j - a_i)$ 皆不为 0。

(iii) $\{a_i\}$, $\{b_i\}$, $\{f_i(x)\}$ 和 $f(x)$ 之间的关系可以简洁地用下表表示之：

x	a_0	a_1	\cdots	a_i	\cdots	a_n
$f_0(x)$	1	0	\cdots	0	\cdots	0
$f_1(x)$	0	1	\cdots	0	\cdots	0
\vdots	\vdots	\vdots	\ddots	\vdots	\ddots	\vdots
$f_i(x)$	0	0	\cdots	1	\cdots	0
\vdots	\vdots	\vdots	\ddots	\vdots	\ddots	\vdots
$f_n(x)$	0	0	\cdots	0	\cdots	1
$f(x)$	b_0	b_1	\cdots	b_i	\cdots	b_n

而结论 $f(x) = \sum_{i=0}^n b_i f_i(x)$ 则可以由下述公式

$$f(a_j) = \sum_{i=0}^n b_i f_i(a_j) = b_j$$

所得出，因为 $f_j(a_j) = 1$ ，而对于其他所有 $i \neq j$ 而言， $f_i(a_j) = 0$ 。

(iv) 上述的插值公式让我们可以把一个「尚待确定」的多项式的样子，用它在不同地方的值直截了当地写下来。这个尚待确定的多项式往往就是一些未知的理论公式或实验公式，所以插值公式不管在理论上或实际应用上都是很重要的。在下一节的讨论中，我们将会应用插值法来研究「求和公式」的性质和其明确表达式。

【习题】：

(1) 令 $g_k(x)$ 为那个唯一确定的 k 次多项式使得 $g_k(i) = 0, 0 \leq i \leq k-1$ 和 $g_k(k) = 1$ 。试找出这系列的多项式 $\{g_k(x), k = 1, 2, 3, 4, \dots\}$ 。

(2) 设除数为 $\{7, 11, 13\}$ 。试找出小于 1001 的正整数使得其余数集分别为 (i) $(0, 3, 7)$ (ii) $(4, 2, 2)$ (iii) $(4, 5, 9)$

(3) 试求公式来计算下述的总和

$$0^2 + 1^2 + 2^2 + \dots + (n-1)^2 = f(n)$$

亦即上述定义的 $f(n)$ 应该是什麼？

[提示：先假设所求的 $f(n)$ 是一个 n 的 3 次多项式。]

(4) 试求公式来计算下述的总和

$$\frac{1}{2} \cdot 0(0-1) + \frac{1}{2} \cdot 1(1-1) + \dots + \frac{1}{2}(n-1)(n-2) = F(n)$$

亦即上述定义的 $F(n)$ 应该是什麼？

(5) 试求公式来计算下述的总和

$$0^3 + 1^3 + 2^3 + \dots + (n-1)^3 = g(n)$$

亦即上述定义的 $g(n)$ 应该是什麼？

[提示：先假设所求的 $g(n)$ 是一个 n 的 4 次多项式。]

(6) 试求公式来计算下述的总和

$$\frac{1}{6} \cdot 0(0-1)(0-2) + \frac{1}{6} \cdot 1(1-1)(1-2) + \dots + \frac{1}{6}(n-1)(n-2)(n-3) = G(n)$$

亦即上述定义的 $G(n)$ 应该是什麼？

1.4 求和公式 (Summation Formula)

现在让我们先看一看几个简单的求和公式的例子：

$$(i) S_n = 0 + 1 + 2 + \dots + (n-1) = \frac{1}{2}n(n-1)$$

在历史上来说，上述级数的求和公式可以用下述巧妙的方法来发现，即

$$\begin{array}{r} S_n = 0 + 1 + 2 + \dots + (n-1) \\ +) S_n = (n-1) + (n-2) + (n-3) + \dots + 0 \\ \hline 2S_n = (n-1) + (n-1) + (n-1) + \dots + (n-1) = n \cdot (n-1) \end{array}$$

因此

$$S_n = \frac{1}{2}n(n-1)$$

(ii) 运用同样方法也可以找出下述等差级数的求和公式：

$$\tilde{S}_n = a + (a+d) + (a+2d) + \dots + [a + (n-1)d] = \frac{1}{2}n[2a + (n-1)d]$$

即以互逆方向相加 $\tilde{S}_n + \tilde{S}_n$ ，得出

$$\begin{aligned} 2\tilde{S}_n &= [2a + (n-1)d] + [2a + (n-1)d] + \dots + [2a + (n-1)d] \\ &= n \cdot [2a + (n-1)d] \end{aligned}$$

其实，我们还可以应用已知 S_n 的公式来间接地求得上述公式：

$$\begin{aligned} \sum_{i=0}^{(n-1)} (a+id) &\stackrel{(\text{分配律})}{=} n \cdot a + d \cdot \sum_{i=0}^{(n-1)} i = n \cdot a + d \cdot \frac{1}{2}n(n-1) \\ &= \frac{1}{2}n[2a + (n-1)d] \end{aligned}$$

$$(iii) \sum_{i=0}^{(n-1)} i^2 = \frac{1}{6}n(n-1)(2n-1)$$

据历史记载，古代中国和古希腊的数学家都知道上述级数的求和公式，但是他们究竟是如何求得此公式，现已无法考证了。不过

，只要他们能够从某些途径或想法得出一个正确的「猜想公式」，再去验证这个猜想公式的正确性其实并不是一件很困难的事情。以现今的说法，验证工作只需要证明下述恒等式：

$$\frac{1}{6}n(n-1)(2n-1) + n^2 \stackrel{?}{=} \frac{1}{6}(n+1) \cdot n \cdot (2n+1)$$

用上述恒等式所提供者作为「归纳步骤」，就可以用数学归纳法来证明上述求和公式的正确性。

(iv) 其实古代中国和古希腊的数学家亦知道另一个类似的求和公式：

$$\sum_{i=0}^{(n-1)} \frac{1}{2}i(i-1) = \frac{1}{6}n(n-1)(n-2)$$

同理，以现今的方法我们只需验证：

$$\begin{aligned} \frac{1}{6}n(n-1)(n-2) + \frac{1}{2}n(n-1) &= \frac{1}{6}n(n-1)[n-2+3] \\ &= \frac{1}{6}(n+1)n(n-1) \end{aligned}$$

即可用归纳法证得上述求和公式的正确性。

由此可见，假若我们能够从某些途径找出一个正确的猜想公式，则馀下所需的验证工作其实是相当简单的「归纳证明」。因此，问题的要点在于怎样去发现和写出这个「尚待确定」的求和公式。现在，让我们应用插值法来找出一个「有系统的」求和公式探讨，作为插值法应用的一个好例子。

首先，如「韩信点兵法」所示，我们不妨先把问题放在一个适当广度的范畴来讨论：

【问题三】：给定任意的多项式 $f(x)$ ，如 $1, x, x^2, x^3$ ，或 $\frac{1}{2}x(x-1)$ ， $\frac{1}{6}x(x-1)(x-2)$ 等等，怎样找出求和公式 $S_f(n)$ 来计算下述级数之总和

$$(1.19) \quad \sum_{i=0}^{(n-1)} f(i) = f(0) + f(1) + f(2) + \dots + f(n-1) = ?$$

亦即要找出适当公式 $S_f(n)$ 使得

$$(1.20) \quad \sum_{i=0}^{(n-1)} f(i) = S_f(n)$$

对于所有 $n = 1, 2, 3, \dots$ 皆成立。

分析：

- (i) 不妨先假设上述求和公式 $S_f(n)$ 是 n 的多项式。则 $S_f(n)$ 满足下述的特征性质：

$$\begin{aligned} S_f(n+1) - S_f(n) &= f(n), \\ S_f(0) &= 0 \quad [\text{零项的总和当然应该是 } 0] \end{aligned}$$

- (ii) 反之，假若有一个多项式 $S_f(x)$ 满足条件 $S_f(0) = 0$ 和

$$S_f(x+1) - S_f(x) = f(x)$$

其中的 $S_f(x+1)$ 表示把 $S_f(x)$ 内所有的 x 都换成 $x+1$ 者，则

$$\begin{aligned} S_f(i+1) - S_f(i) &= f(i), \quad i \geq 0 \\ \sum_{i=0}^{(n-1)} f(i) &= \sum_{i=0}^{(n-1)} (S_f(i+1) - S_f(i)) \\ &= S_f(n) - S_f(0) = S_f(n) \end{aligned}$$

【定理 1.3】：给出任何一个 k 次多项式 $f(x)$ ，存在一个唯一的 $k+1$ 次多项式 $S_f(x)$ ，它满足条件 $S_f(0) = 0$ 和

$$(1.21) \quad S_f(x+1) - S_f(x) = f(x)$$

亦即

$$(1.22) \quad \sum_{i=0}^{(n-1)} f(i) = S_f(n)$$

对于所有 $n = 1, 2, 3, \dots$ 皆成立。

证明：让我们先对 k 用归纳法来证明 $S_f(x)$ 的存在性。当 $k = 0$ 时， $f(x)$ 实质上是一个常数，即 $f(x) \equiv c$ 。易知这时 $S_f(x) = cx$ ，亦即

$$(1.23) \quad S_f(x+1) - S_f(x) = c(x+1) - cx = c = f(x)$$

归纳假设对所有次数不高于 k 的多项式 $f(x)$ ，求和公式 $S_f(x)$ 皆存在，而且其次数不高于 $k+1$ ；现在进而证明对任意给出的 $(k+1)$ 次多项式 $f(x)$ ，其所相应的 $S_f(x)$ 的存在性如下。令

$$(1.24) \quad \begin{aligned} f(x) &= cx^{k+1} + g(x), \quad \deg g(x) \leq k \\ \frac{c}{k+2} [(x+1)^{k+2} - x^{k+2}] &= cx^{k+1} + h(x) \end{aligned}$$

易见

$$(1.25) \quad (x+1)^{k+2} = x^{k+2} + (k+2)x^{k+1} + \text{不高于 } k \text{ 次的多项式}$$

所以有 $\deg h(x) \leq k$ 和 $\deg(g(x) - h(x)) \leq k$ 。由归纳假设得知存在一个次数不高于 $(k+1)$ 的多项式 $G(x)$ 使得：

$$(1.26) \quad G(0) = 0 \quad \text{和} \quad G(x+1) - G(x) = g(x) - h(x)$$

令

$$(1.27) \quad S_f(x) = \frac{c}{k+2} x^{k+2} + G(x)$$

则 $S_f(0) = 0$ 和

$$(1.28) \quad \begin{aligned} S_f(x+1) - S_f(x) &= \frac{c}{k+2} [(x+1)^{k+2} - x^{k+2}] + [G(x+1) - G(x)] \\ &= cx^{k+1} + h(x) + [g(x) - h(x)] \\ &= cx^{k+1} + g(x) = f(x) \end{aligned}$$

这样便归纳地证明了 $S_f(x)$ 的存在性。

至于 $S_f(x)$ 的唯一性则可以由[定理 1.1]的结论和 $S_f(x)$ 的特征性质：

$$(1.29) \quad \sum_{i=0}^{(n-1)} f(i) = S_f(n), \quad n = 1, 2, 3, \dots$$

直接推论而得。 □

[注]: 当求和公式的存在性和唯一性业已验证妥当後, 插值法便提供了一种非常有效的方法来求出 $S_f(x)$ 的明确公式。

一个 k 次多项式 $f(x)$ 是可以由其 $(k+1)$ 个值 $\{f(i); 0 \leq i \leq k\}$ 所唯一确定。而相应的求和公式 $S_f(x)$, 由于它是 $(k+1)$ 次多项式, 则需要由其 $(k+2)$ 个值所决定, 即

$$\begin{aligned} S_f(0) &= 0, S_f(1) = f(0), S_f(2) = f(0) + f(1), \dots \\ S_f(k+1) &= f(0) + f(1) + \dots + f(k) \end{aligned}$$

因此, 我们很自然地便想到可以用插值法来从上述 $(k+2)$ 个接连的和值来求出 $S_f(x)$ 的明确公式。

一系列特殊的多项式: 令

$$\begin{aligned} g_0(x) &= 1, g_1(x) = x, g_2(x) = \frac{1}{2}x(x-1), \dots, \\ g_k(x) &= \frac{1}{k!}x(x-1)\dots(x-k+1), \dots \end{aligned}$$

上述 $g_k(x)$ 的选取是使得 $g_k(i) = 0, 0 \leq i \leq (k-1)$ 和 $g_k(k) = 1$, 所以 $g_k(n)$ 的首 k 个总和皆为 0 (即 $S_{g_k}(i) = 0, 1 \leq i \leq k$), 而第 $(k+1)$ 个总和是 1 (即 $S_{g_k}(k+1) = 1$)。这刚好表示 $g_k(x)$ 的求和公式其实就是此系列多项式的下一位成员, 即

$$S_{g_k}(x) = g_{k+1}(x)$$

上述极为简洁的关系式自然而然地成为我们用来寻找求和公式的有效工具 (就像在余数问题中的 x_1, \dots, x_k 一样)。给定任意一个 ℓ 次的多项式 $f(x)$, 其求和公式 $S_f(x)$ 可以用下述方法来有系统地确定之:

(i) 首先要注意 $f(x)$ 是可以唯一地写成 $\{g_k(x); 0 \leq k \leq \ell\}$ 的常数倍之组合, 亦即存在系数集 $\{c_k; 0 \leq k \leq \ell\}$ 使得

$$\begin{aligned} f(x) &= c_0 g_0(x) + c_1 g_1(x) + \dots + c_k g_k(x) + \dots + c_\ell g_\ell(x) \\ &= \sum_{k=0}^{\ell} c_k g_k(x) \end{aligned}$$

上述组合一般称之为 $\{g_k(x), 0 \leq k \leq \ell\}$ 的「线性组合」(linear combination)。

(ii) 然後，由分配律可得：

$$\begin{aligned} \sum_{i=0}^{(n-1)} f(i) &= \sum_{i=0}^{(n-1)} (c_0 g_0(i) + c_1 g_1(i) + \dots + c_k g_k(i) + \dots + c_\ell g_\ell(i)) \\ &= c_0 \sum_{i=0}^{(n-1)} g_0(i) + \dots + c_k \sum_{i=0}^{(n-1)} g_k(i) + \dots + c_\ell \sum_{i=0}^{(n-1)} g_\ell(i) \\ &= c_0 g_1(n) + c_1 g_2(n) + \dots + c_k g_{k+1}(n) + \dots + c_\ell g_{\ell+1}(n) \end{aligned}$$

即

$$S_f(x) = \sum_{k=0}^{\ell} c_k g_{k+1}(x)$$

由此可见若要写下这个待定的 $S_f(x)$ 的明确表达式，就只需去求得系数集 $\{c_k; 0 \leq k \leq \ell\}$ 。所以接著的工作就是如何有效地从 $\{f(i); 0 \leq i \leq \ell\}$ 这 $(\ell + 1)$ 个值来确定系数集 $\{c_k; 0 \leq k \leq \ell\}$ 。

(iii) 首个系数 c_0 显然等于 $f(0)$ ，因为 $g_k(0) = 0, k > 0$ 。为了方便描述下面一套有系统的方法来从 $\{f(i); 0 \leq i \leq \ell\}$ 这 $(\ell + 1)$ 值求得对应系数集 $\{c_k; 0 \leq k \leq \ell\}$ ，我们引进下述运算符号（称之为差分算子，difference operator）：

$$\Delta f(x) = f(x+1) - f(x)$$

不难直接验证 Δ 算子具有下述易算好用的性质：

$$\Delta(f(x) + h(x)) = \Delta f(x) + \Delta h(x), \quad \Delta(cf(x)) = c \Delta(f(x))$$

另一方面，因为 $g_k(x)$ 是 $g_{k-1}(x)$ 的求和公式，易见

$$\Delta g_k(x) = g_{k-1}(x)$$

由此可得

$$\begin{aligned} \Delta f(x) &= \Delta(c_0 + c_1 g_1(x) + \dots + c_\ell g_\ell(x)) \\ &= 0 + c_1 g_0(x) + c_2 g_1(x) + \dots + c_\ell g_{\ell-1}(x) \end{aligned}$$

即

$$c_1 = \Delta f(0) = f(1) - f(0)$$

再次应用 Δ 算子，我们便可以如下算出下一个系数 c_2 ：

$$\Delta^2 f(x) = c_2 g_0(x) + c_3 g_1(x) + \dots + c_\ell g_{\ell-2}(x)$$

$$\begin{aligned} c_2 &= \Delta^2 f(0) = \Delta f(1) - \Delta f(0) \\ &= (f(2) - f(1)) - (f(1) - f(0)) \\ &= f(2) - 2f(1) + f(0) \end{aligned}$$

如此类推，我们不断重复应用 Δ 算子，然后再代入 $x = 0$ ，便可逐步从 $\{f(i); 0 \leq i \leq \ell\}$ 求得 $\{c_k; 0 \leq k \leq \ell\}$ ：

$$\begin{aligned} (c_0 =) & f(0), f(1), f(2), f(3), \dots, f(\ell-1), f(\ell) \\ (c_1 =) & \Delta f(0), \Delta f(1), \Delta f(2), \dots, \Delta f(\ell-1) \\ (c_2 =) & \Delta^2 f(0), \Delta^2 f(1), \dots, \Delta^2 f(\ell-2) \\ (c_3 =) & \Delta^3 f(0), \dots, \Delta^3 f(\ell-3) \\ & \quad \vdots \quad \quad \quad \vdots \\ (c_\ell =) & \Delta^\ell f(0) \end{aligned}$$

[下一层的数值是由上一层相邻的两个数值之差所得出的。]

【例子】：

(i) $f(x) = x^2$

$$\begin{array}{r} 0 \quad 1 \quad 4 \\ 1 \quad 3 \\ 2 \end{array}$$

$$f(x) = g_1(x) + 2g_2(x)$$

$$S_f(x) = g_2(x) + 2g_3(x) = \frac{1}{6}x(x-1)(2x-1)$$

(ii) $\underline{f(x) = x^3}$

$$\begin{array}{cccc} 0 & 1 & 8 & 27 \\ & 1 & 7 & 19 \\ & & 6 & 12 \\ & & & 6 \end{array}$$

$$f(x) = g_1(x) + 6g_2(x) + 6g_3(x)$$

$$S_f(x) = g_2(x) + 6g_3(x) + 6g_4(x) = \frac{1}{4}x^2(x-1)^2$$

(iii) $\underline{f(x) = x^4}$

$$\begin{array}{cccccc} 0 & 1^4 & 2^4 & 3^4 & 4^4 \\ & 1 & 15 & 65 & 175 \\ & & 14 & 50 & 110 \\ & & & 36 & 60 \\ & & & & 24 \end{array}$$

$$f(x) = g_1(x) + 14g_2(x) + 36g_3(x) + 24g_4(x)$$

$$S_f(x) = g_2(x) + 14g_3(x) + 36g_4(x) + 24g_5(x)$$

【习题】：

(1) 试找出下列 $f(x)$ 的求和公式 $S_f(x)$ ：

(i) $f(x) = x^2(x-1)$

(ii) $f(x) = x^2(x-1)(x-2)$

(iii) $f(x) = x^4 + x^3 - x^2$

(iv) $f(x) = x^3(x-1)(x-2)$

(v) $f(x) = x^2(x+1)(x-1)(x-2)$

(vi) $f(x) = x(x+1)(x+2)(x+3)(x+4)$

(2) 试证明 $\{g_k(x), k=0, 1, 2, \dots\}$ 在所有整数点 (即包括正、负和零) 上恒取整数值。

- (3) 若一个多项式 $f(x)$ 在所有的整数点上取整数值，问 $f(x)$ 与 $\{g_k(x), k = 0, 1, 2, \dots\}$ 的关系是什麼？
- (4) 设 $f(x)$ 为一个 n 次多项式，它在 $0, 1, 2, \dots, n$ 这 $(n+1)$ 点上皆取整数值。问 $f(x)$ 在其他的整数点上是否也取整数值？
- (5) 试求一个 4 次多项式 $f(x)$ 在 (i) $0, 1, 2, 3, 5$; (ii) $0, 1, 2, 3, 6$ 皆取整数值的充要条件。
- (6) 若一个 k 次多项式 $f(x)$ 在 $0, 2, 4, \dots, 2k$ 上皆取整数值，问 $f(x)$ 在其他的偶数点上是否也取整数值？
- (7) 设一个 k 次多项式在 $(k+1)$ 个连接的整数点上皆取偶数值（或奇数值），试问它是否在所有整数点上皆取偶数值（或奇数值）？试证明你的论断。

1.5 插值法与因式分解

因式分解在初中的代数课程里是常常遇到的题目。同学们试回想一下当时究竟是用那一种方法来找一个给定多项式的因式分解？相信大多数是采用「盲目撞试」的方法。及至引进了徐式定理後，才有一个比较有系统的方法来撞试其一次因式的可能性。但假若给定的多项式并没有整数系数的一次因式，例如

$$2x^4 - 4x^3 + x^2 + 3x + 3$$

相信大家对于寻找这个多项式的因式分解也就无从入手了。现在我们有了插值法和插值公式，就可以对这些因式分解问题给与一个有效能算的解法。下面的讨论且以一个 4 次多项式的因式分解作为特例说明，但易见其方法是可以作明显推广的。首先我们引进一些术语：

【定义】：一个整数系数的多项式 $f(x)$ ，若可以写成两个同样具有整数系数的非常数因式之乘积，则称 $f(x)$ 在整数系数内是可约的 (reducible over integer coefficients)。

【例 1】： $x^2 - 1 = (x+1)(x-1)$ 是可约的。

【例 2】： $x^2 - 2$ 在整数系数内是不可约的。虽然 $x^2 - 2 = (x + \sqrt{2})(x - \sqrt{2})$ ，但是这个因式分解中用到非整数系数： $\pm\sqrt{2}$ 。

在下面的讨论中，我们都是假设所用到的系数皆为整数。让我们先来温习一下寻找一次因式的方法：

令 $f(x) = ax^4 + bx^3 + cx^2 + dx + e$ 为一个 4 次多项式。我们不妨假设 $e \neq 0$ ，否则即可提出因式 x ，并把情况简化到 3 次多项式的情形来讨论。若 $f(x)$ 具有 1 次因式 $(px + q)$ ，则易见 p, q 分别可整除 a, e ：

$$f(x) = (px + q)(c_3x^3 + c_2x^2 + c_1x + c_0) \Rightarrow p \cdot c_3 = a, q \cdot c_0 = e$$

由于 a, e 的因数（包括正负）只有有限个选择，所以只需直接验算每一个可能情况，最後便可得知 $f(x)$ 是否具有 1 次因式。

【例 3】：若 $(px + q)$ 可整除 $f(x) = 2x^4 + 3x^2 + x - 1$ ，则 p, q 必须分别整除 2, 1，由此可见 (p, q) 的可能性为

$$(1, 1), (1, -1), (2, 1), (2, -1)$$

[注： $(-1, 1)$ 对应 $(1, -1)$ ，此等情况可以不用再考虑。]运用余式定理，我们只需分别验算

$$f(-1) = 3 \neq 0, f(1) = 5 \neq 0, f(-\frac{1}{2}) = -\frac{5}{8} \neq 0, f(\frac{1}{2}) = \frac{3}{8} \neq 0$$

即知 $f(x)$ 没有一次因式。

现在让我们考虑 $f(x)$ 的 2 次因式。若 $g(x)$ 为 $f(x)$ 的一个 2 次因式，即 $f(x) = g(x) \cdot h(x)$ ，则 $f(-1) = g(-1) \cdot h(-1)$ ， $f(0) = g(0) \cdot h(0)$ 和 $f(1) = g(1) \cdot h(1)$ ，亦即 $g(-1), g(0), g(1)$ 分别可整除 $f(-1), f(0), f(1)$ 。因为 2 次多项式可以由其在三个相异位置的值所唯一决定，所以我们只需列出 $g(-1), g(0)$ 和 $g(1)$ 的所有可能性，然後以插值公式写下相应的 $g(x)$ ，再验算此等 $g(x)$ 是否有为 $f(x)$ 的因式者，最後便可得知 $f(x)$ 是否具有 2 次因式了。

【例 4】：若 $g(x)$ 可整除 $f(x) = 2x^4 + 3x^2 + x - 1$ ，则 $g(-1), g(0)$ 和 $g(1)$ 分别可以整除 $f(-1) = 3, f(0) = -1$ 和 $f(1) = 5$ ，即

x	-1	0	1
$f(x)$	3	-1	5
$g(x)$	$\pm 1, \pm 3$	± 1	$\pm 1, \pm 5$

不妨选定 $g(0) = 1$ ，则总共有 16 种 $g(-1)$ 和 $g(1)$ 的组合：

	$g(-1)$	$g(0)$	$g(1)$
1.	1	1	1
2.	1	1	-1
3.	1	1	5
4.	1	1	-5
5.	-1	1	1
6.	-1	1	-1

	$g(-1)$	$g(0)$	$g(1)$
7.	-1	1	5
8.	-1	1	-5
9.	3	1	1
10.	3	1	-1
11.	3	1	5
12.	3	1	-5

	$g(-1)$	$g(0)$	$g(1)$
13.	-3	1	1
14.	-3	1	-1
15.	-3	1	5
16.	-3	1	-5

运用插值公式，把每一种组合所相应的 $g(x)$ 如下列出来：

$$\begin{array}{lll}
 1. & 1 & 7. \quad x^2 + 3x + 1 & 13. \quad -2x^2 + 2x + 1 \\
 2. & -x^2 - x + 1 & 8. \quad -4x^2 - 2x + 1 & 14. \quad -3x^2 + x + 1 \\
 3. & 2x^2 + 2x + 1 & 9. \quad x^2 - x + 1 & 15. \quad 4x + 1 \\
 4. & -3x^2 - 3x + 1 & 10. \quad -2x + 1 & 16. \quad -5x^2 - x + 1 \\
 5. & -x^2 + x + 1 & 11. \quad 3x^2 + x + 1 & \\
 6. & -2x^2 + 1 & 12. \quad -2x^2 - 4x + 1 &
 \end{array}$$

耐心地逐个验算，我们便会发现上述所得的 13 个 2 次多项式皆不能整除 $f(x)$ 。由此可见， $f(x)$ 也没有 2 次因式。再者，因为 $f(x)$ 有 3 次因式等同于 $f(x)$ 有 1 次因式，所以我们其实已经证明了 $f(x) = 2x^4 + 3x^2 + x - 1$ 是不可约的。

当然，选用 $x = -1, 0, 1$ 为插值点只是为了方便计算，它们并不是唯一的选择，亦即可以选取任意三个整数为插值点。另一方面，若要找出一个多项式有没有 3 次因式，则只需用到四个插值点的插值法。

其实，上述特例也列出了在 $x = -1, 0, 1$ 分别取值 3, -1, 5 的多项式所能具有的 2 次因式的所有可能性，亦即我们可以反过来写下有那些 4 次可约多项式能够在 $x = -1, 0, 1$ 取值 3, -1, 5 的所有可能性。当一个 4 次多项式不是其中之一时，即知它为不可约者（在整系数多项式中）。由此可见，插值法的用途广泛多样，重点在于其「插值想法」，并不只是在于其表达公式。

【思考题】：

- (1) 在[例 3]中的寻找 1 次因式的方法能否一起合并入[例 4]的插值法之内？

-
- (2) 在[例 4]的讨论中， $f(x)$ 于 $x = -1, 0, 1$ 取值 $3, -1, 5$ 或 $3, 1, 5$ 这两种情况的讨论有没有分别？
- (3) 试用插值方法寻找 $f(x) = 2x^4 - 4x^3 + x^2 + 3x + 3$ 的因式分解。
- (4) 设 $\ell(x)$ 为一个 4 次多项式，它在 $x = -1, 0, 1$ 上的值分别是 $-1, 1, 5$ 。问 $\ell(x)$ 可以写成什麼模样？
- (5) 给定任意的整数 a, b ，问 $f(x) = (ax + b)x(x + 1)(x - 1) + 1$ 是否可约？

第二章

二项定理与泰勒公式

在多项式的运算中，把 $(x+y)^n$ 展开成单项式之和（即 $\sum_{k=0}^n c_k x^{n-k} y^k$ ）的公式是有著根本的重要性的。例如在上一章讨论求和公式的时候，我们需要同时讨论 $S_f(x)$ 在 $x = k$ 和 $x = k+1$ 时的表达式，所以寻找 $(x+1)^n$ 的展开式就是一个关键的步骤。二项定理就是把上述 $(x+y)^n = \sum_{k=0}^n c_k x^{n-k} y^k$ 中的每个系数 c_k ($0 \leq k \leq n$) 明确地表达出来。

泰勒公式在本质上其实就是二项定理的推广，它也是引进多项式的微分学的一个重要起点。

2.1 二项定理 (The Binomial Theorem)

当 n 为 2, 3, 4, 5 等等这些较低的指数时，不难直接运用分配律归纳地写下 $(x+y)^n$ 的展开式：

$$n = 0 : (x+y)^0 = 1$$

$$n = 1 : (x+y)^1 = x+y$$

$$n = 2 : (x+y)^2 = x^2 + 2xy + y^2$$

$$n = 3 : (x+y)^3 = x^3 + 3x^2y + 3xy^2 + y^3$$

$$n = 4 : (x+y)^4 = x^4 + 4x^3y + 6x^2y^2 + 4xy^3 + y^4$$

$$n = 5 : (x+y)^5 = x^5 + 5x^4y + 10x^3y^2 + 10x^2y^3 + 5xy^4 + y^5$$

假如把上述每个展开式中的 $x^{n-k}y^k$ 的系数分离出来，我们发现它们之间可以自然而然地排列成下述三角阵式，并且满足一种简单的构造规

律。这个三角阵式我们称之为「贾宪三角」或「杨辉三角」，而在西方则称之为「帕斯卡三角」(Pascal triangle)。

$$(2.1) \quad \begin{array}{ccccccccc} & & & & 1 & & & & \\ & & & & 1 & & 1 & & \\ & & & & 1 & & 2 & & 1 & \\ & & & & 1 & & 3 & & 3 & & 1 & \\ & & & & 1 & & 4 & & 6 & & 4 & & 1 & \\ & & & & 1 & & 5 & & 10 & & 10 & & 5 & & 1 & \\ & & & & 1 & & 6 & & 15 & & 20 & & 15 & & 6 & & 1 & \\ & & & & \vdots & & & & \vdots & & & & & & & & & \vdots & \end{array}$$

例如在 $n = 3$ 时，其相应的系数集就是 $\{1, 3, 3, 1\}$ ，亦即

$$(2.2) \quad (x + y)^3 = 1x^3 + 3x^2y + 3xy^2 + 1y^3$$

若将 (2.2)-式乘以 $(x + y)$ ，则用分配律可得

$$(2.3) \quad \begin{aligned} (x + y)^4 &= x \cdot (1x^3 + 3x^2y + 3xy^2 + 1y^3) \\ &\quad + y \cdot (1x^3 + 3x^2y + 3xy^2 + 1y^3) \end{aligned}$$

亦即

$$\begin{aligned} (x + y)^4 &= 1x^4 + 3x^3y + 3x^2y^2 + 1xy^3 + (0y^4) \\ &\quad + (0x^4) + 1x^3y + 3x^2y^2 + 3xy^3 + 1y^4 \\ \hline &= 1x^4 + 4x^3y + 6x^2y^2 + 4xy^3 + 1y^4 \end{aligned}$$

由此可见，第五行 ($n = 4$) 中的每个系数其实就是上一行 ($n = 3$) 中的相邻系数之和（在首、尾补上两个零）。而在一般的情况，令

$$(2.4) \quad (x + y)^n = \sum_{k=0}^n C_k^n x^{n-k} y^k$$

则类似地可得

$$(2.5) \quad \begin{aligned} (x + y)^{n+1} &= (x + y)(x + y)^n = x \cdot (x + y)^n + y \cdot (x + y)^n \\ &= \sum_{k=0}^n C_k^n x^{n+1-k} y^k + \sum_{k=0}^n C_k^n x^{n-k} y^{k+1} \\ &= \sum_{k=0}^{n+1} (C_k^n + C_{k-1}^n) x^{n+1-k} y^k \quad (\text{补上 } C_{n+1}^n = 0 = C_{-1}^n) \end{aligned}$$

另一方面，由 C_k^n 的定义式 (2.4) 得

$$(2.6) \quad (x+y)^{n+1} = \sum_{k=0}^{n+1} C_k^{n+1} x^{n+1-k} y^k$$

由此可得

$$C_k^{n+1} = C_k^n + C_{k-1}^n \quad (\Leftrightarrow C_k^{n+1} - C_k^n = C_{k-1}^n)$$

这个归纳递推公式 (inductive recurrence relation) 来表达那些系数与系数之间的关系，亦即可以用 $(x+y)^n$ 展开后的系数来表达 $(x+y)^{n+1}$ 展开后的系数。这也就是描述上述三角阵式的构造规律。

现在让我们从另一角度来看上述递推公式的意义。先把某一个 $k \geq 0$ 固定不动，然后将 C_k^n 看成一个函数 $f_k(x)$ 在 $x = n$ 时的取值。易见

$$(2.7) \quad f_0(x) \equiv 1, \quad f_1(x) = x \quad (\text{因为 } C_1^n = n)$$

则前述的递推公式可以重新写成

$$(2.8) \quad f_k(n+1) - f_k(n) = f_{k-1}(n) \quad (\text{即 } C_k^{n+1} - C_k^n = C_{k-1}^n)$$

显然上述的讨论说明了

$$(2.9) \quad \Delta f_k(x) = f_{k-1}(x) \Leftrightarrow f_k(x) = S_{f_{k-1}}(x)$$

亦即上述所定义的一系列函数 $\{f_k(x), k = 0, 1, 2, \dots\}$ 刚好就是在前一章寻找求和公式时引入的特殊函数集 $\{g_k(x), k = 0, 1, 2, \dots\}$ 。

【定理 2.1】二项定理 (The Binomial Theorem) :

$$(2.10) \quad (x+y)^n = \sum_{k=0}^n C_k^n x^{n-k} y^k$$

而其中出现的二项式系数 (binomial coefficients) C_k^n 的明确表达式就是

$$(2.11) \quad C_k^n = g_k(n) = \frac{1}{k!} n(n-1)\dots(n-k+1) = \frac{n!}{k!(n-k)!} \quad (0! = 1)$$

【推论一】 : $C_k^n = C_{n-k}^n$ 。

【推论二】 : $\sum_{k=0}^n C_k^n = (1+1)^n = 2^n$, $\sum_{k=0}^n (-1)^k C_k^n = (1-1)^n = 0$ 。

2.2 泰勒公式与多项式的局部展开式

设 $f(x)$ 为一个给定的多项式, a 为一个给定点。若我们把 x 局限于 a 的 (足够小) 邻近时, $f(x)$ 在这个范围的性质便称之为 $f(x)$ 在 $x=a$ 的邻域的「局部性质」(local properties)。当我们研究 $f(x)$ 的局部性质时, 宜以 $x = a+t$ 代入然后以二项定理把 $f(x) = f(a+t)$ 展开为 t 的升幂表达式, 其中 $|t|$ 是足够小者, 因而展开后的各项的绝对值乃是随着 t 的次数的升高而大幅缩小, 因此它们的局部影响力显然是高次项要远小于低次项, 可以说乃是「阶段分明, 一目了然」者也。总之, 这是一种常用、好用的办法, 例如:

(i) $f(x) = x^3$:

$$f(a+t) = a^3 + 3a^2t + 3at^2 + t^3$$

(ii) $f(x) = x^3 - 2x^2 + 5x + 1$:

$$f(a+t) = (a^3 - 2a^2 + 5a + 1) + (3a^2 - 4a + 5)t + (3a - 2)t^2 + t^3$$

(iii) $f(x) = cx^n$:

$$(2.12) \quad f(a+t) = ca^n + nca^{n-1}t + \frac{n(n-1)}{2}ca^{n-2}t^2 + \dots + ct^n$$

由上述「局部展开」(local expansion) 的例子可见, 引进下述对于多项式的形式运算¹ (formal operation) 将会有很大的用途:

【定义】: 设 $f(x)$ 为任意给定的多项式

$$f(x) = a_0 + a_1x + \dots + a_nx^n = \sum_{i=0}^n a_i x^i$$

定义算子 D 为把 $f(x)$ 变成下述多项式 $Df(x)$ 的形式运算:

$$(2.13) \quad Df(x) = a_1 + 2a_2x + \dots + na_nx^{n-1} = \sum_{i=1}^n ia_ix^{i-1}$$

¹形式运算: 单纯在符号层面上的运算, 其实质意义还有待确定。

再者，定义 D^k 为上述形式运算 D 的重复 k 次者。例如：

$$\begin{aligned} D^2 f(x) &= 2a_2 + 6a_3x + \dots + n(n-1)a_nx^{n-2} \\ &= \sum_{i=2}^n i(i-1)a_ix^{i-2} \\ D^k(x^n) &= n(n-1)\dots(n-k+1)x^{n-k} = k!C_k^n x^{n-k} \end{aligned}$$

易证上述所定义的算子 D 满足下述简单好用的性质，即

$$(2.14) \quad D(c_1f_1(x) + c_2f_2(x)) = c_1Df_1(x) + c_2Df_2(x)$$

(事实上， D 本身就可以用上述性质和基本定义 $D(x^n) = nx^{n-1}$ 所唯一确定者。)

运用算子符号 D ，我们可以把 (2.12)-式中的 $f(x) = cx^n$ 局部展开式重新写成下述非常简洁的形式，即

$$\begin{aligned} f(a+t) &= f(a) + Df(a)t + \frac{D^2f(a)}{2!}t^2 + \dots + \frac{D^k f(a)}{k!}t^k + \dots + \frac{D^n f(a)}{n!}t^n \\ &= \sum_{k=0}^n \frac{D^k f(a)}{k!}t^k \end{aligned}$$

在 $f(x) = x^n$ 的特殊情形，上述公式其实就是二项定理的另一写法，也许大家会问这岂非仅是旧酒装新瓶的做法？但其实不然！上述公式乃是带领著我们迈向一个崭新的领域，它开启了进入微分学范畴的大门（详见第三章）。

【定理 2.2】多项式的泰勒公式 (Taylor's formula for polynomials)：
设 $f(x)$ 为一个 n 次多项式，则

$$\begin{aligned} f(a+t) &= f(a) + Df(a)t + \frac{D^2f(a)}{2!}t^2 + \dots + \frac{D^k f(a)}{k!}t^k + \dots + \frac{D^n f(a)}{n!}t^n \\ &= \sum_{k=0}^n \frac{D^k f(a)}{k!}t^k \end{aligned}$$

[注]：上述局部展开式在 $k = n$ 时会自然终止，因为对于所有 $k > n = \deg f(x)$ 来说， $D^k f(x) \equiv 0$ 。

证明：我们采用归纳法按多项式次数 n 作归纳证明。 $\deg f(x) = 0$ 时的情况是显而易见的，现在归纳假设泰勒公式已对于所有次数不高于 $(n-1)$ 的多项式普遍成立。令 $f(x)$ 为任意一个 n 次的多项式，把 $f(x)$ 重新写成：

$$(2.15) \quad f(x) = cx^n + f_2(x) = f_1(x) + f_2(x), \quad \deg f_2(x) \leq n-1$$

则由归纳假设可得

$$(2.16) \quad f_2(a+t) = \sum_{k=0}^n \frac{1}{k!} D^k f_2(a) t^k \quad (D^n f_2(a) = 0)$$

由前述 cx^n 的展开式例子（即 (2.12)-式）：

$$(2.17) \quad f_1(a+t) = \sum_{k=0}^n \frac{1}{k!} D^k f_1(a) t^k$$

由此可得

$$\begin{aligned} f(a+t) &= f_1(a+t) + f_2(a+t) \\ &= \sum_{k=0}^n \frac{1}{k!} D^k f_1(a) t^k + \sum_{k=0}^n \frac{1}{k!} D^k f_2(a) t^k \\ &= \sum_{k=0}^n \frac{1}{k!} (D^k f_1(a) + D^k f_2(a)) t^k \\ &= \sum_{k=0}^n \frac{1}{k!} D^k f(a) t^k \quad (\text{即 (2.14)-式}) \quad \square \end{aligned}$$

【推论一】：当 $f(x)$ 被除以 $(x-a)^k$ 时，其商式为：

$$f(a) + Df(a)(x-a) + \dots + \frac{1}{(k-1)!} D^{k-1} f(a)(x-a)^{k-1}$$

证明：

$$\begin{aligned} f(x) &= f(a+(x-a)) = \sum_{j=0}^n \frac{1}{j!} D^j f(a)(x-a)^j \\ &= \left\{ \sum_{j=k}^n \frac{1}{j!} D^j f(a)(x-a)^{(j-k)} \right\} \cdot (x-a)^k + \left\{ \sum_{j=0}^{k-1} \frac{1}{j!} D^j f(a)(x-a)^j \right\} \quad \square \end{aligned}$$

【推论二】： $f(x)$ 可被 $(x-a)^2$ 整除的充要条件为 $f(a) = 0$ 和 $Df(a) = 0$ ，即 $x = a$ 是 $f(x)$ 和 $Df(x)$ 的公共根，亦即 $(x-a)$ 是 $f(x)$ 和 $Df(x)$ 的一个公因式。

【推论三】： $D(f(x) \cdot g(x)) = (Df(x)) \cdot g(x) + f(x) \cdot Dg(x)$ 。

证明：只需证明对任意 a 恒有

$$(2.18) \quad D(f \cdot g)(a) = Df(a) \cdot g(a) + f(a) \cdot Dg(a)$$

由[推论一]知

$$\begin{aligned} f(x) &= q_1(x) \cdot (x-a)^2 + [f(a) + Df(a)(x-a)] \\ g(x) &= q_2(x) \cdot (x-a)^2 + [g(a) + Dg(a)(x-a)] \\ f(x) \cdot g(x) &= q_3(x) \cdot (x-a)^2 + [f(a) \cdot g(a) + D(f \cdot g)(a)(x-a)] \end{aligned}$$

把上述的第一式和第二式相乘後可得

$$\begin{aligned} & f(x) \cdot g(x) \\ &= \{q_1(x) \cdot g(x) + q_2(x) \cdot [f(a) + Df(a)(x-a)] + Df(a) \cdot Dg(a)\} \cdot (x-a)^2 \\ & \quad + \{f(a) \cdot g(a) + [Df(a) \cdot g(a) + f(a) \cdot Dg(a)](x-a)\} \end{aligned}$$

现在，把上述两种 $f(x) \cdot g(x)$ 的展开式中的 $(x-a)$ 项作一比较，即得所求证的等式 $D(f \cdot g)(a) = Df(a) \cdot g(a) + f(a) \cdot Dg(a)$ ，亦即

$$D(f \cdot g) = Df \cdot g + f \cdot Dg \quad \square$$

【习题】：

(1) 试证明

$$D(f \cdot g \cdot h) = (Df) \cdot g \cdot h + f \cdot (Dg) \cdot h + f \cdot g \cdot (Dh)$$

(2) 试以归纳法证明

$$D[f(x)^n] = n[f(x)]^{n-1} \cdot Df(x)$$

(3) $D[(x^2 + x + 1)^5] = ?$

(4) 试证 $D^2(f(x) \cdot g(x)) = D^2(f(x)) + 2D(f(x))D(g(x)) + D^2(g(x))$ 。问 $D^3(f(x) \cdot g(x)) = ?$ 能否给出 $D^n(f(x) \cdot g(x))$ 的公式?

【定义】：若 $f(x)$ 含有因式 $(x-a)^2$ ，则称 a 为 $f(x)$ 的一个重根。若 $f(x)$ 含有因式 $(x-a)^k$, $k \geq 2$ ，而不含有因式 $(x-a)^{k+1}$ ，则称 a 为 $f(x)$ 的 k -重根。

[注]：若 $k=1$ ，则称 a 为 $f(x)$ 的一个单根。有时候为了方便讨论起见，我们会把单根的情况当作 1-重根来讨论。

(5) 设 $d(x)$ 是 $f(x)$ 和 $Df(x)$ 的最高公因式，试证 a 为 $f(x)$ 的重根的充要条件就是 a 为 $d(x)$ 的根。

[注]：所以，若 $f(x)$ 和 $Df(x)$ 互素，亦即 $d(x) \equiv 1$ ，则 $f(x)$ 不可能有任何重根。

(6) 试证 a 是 $f(x)$ 的 k -重根 ($k \geq 2$) 的充要条件乃是 a 是 $f(x)$ 的根和 $Df(x)$ 的 $(k-1)$ -重根。

【定义】：给定某一点 $x=a$ ，若存在有一个足够小的 $\delta > 0$ 使得对于所有满足 $|x-a| < \delta$ 的 x 恒有

$$f(x) \geq f(a) \quad (\text{或 } f(x) \leq f(a))$$

则称 $x=a$ 为 $f(x)$ 的局部极小点 (或局部极大点)。

(7) 假设 $Df(a) = 0$ 及 $D^2f(a) > 0$ (或 $D^2f(a) < 0$)。那么 $x=a$ 是否 $f(x)$ 的一个局部极小点 (或局部极大点)? 为什麼?

(8) 设 $Df(a) = D^2f(a) = \dots = D^{k-1}f(a) = 0$ 但 $D^k f(a) \neq 0$ 。 $f(x)$ 在 $x=a$ 的邻域的局部性质应是怎样?

[提示：分开处理 k 为奇、偶数的情况。注意 $f(x)$ 与 $y = \frac{1}{k!} D^k f(a) \cdot (x-a)^k$ 在 $x=a$ 的邻域的局部性质极为相似。]

2.3 泰勒公式与局部分析：局部性质和局部逼近

从代数学方面来说，上一节所导出的多项式泰勒公式已经是二项定理的一个深远推广。其实，泰勒公式不单只是二项定理的一个推广公式，它也是研究一般函数局部性质的重要工具。在这一节里我们暂且集中讨论泰勒公式在多项式函数的应用，其研讨方法也自然而然地带领著我们迈向微分学的范畴。

局部性质和局部逼近 (local properties and local approximations)：

首先让我们清楚界定「局部性质」的明确意义。若一个给定函数 $y = f(x)$ 的某些性质只取决于该函数在 $x = a$ 的足够小邻近内的值（即只需把 $f(x)$ 局限于 $|x - a| < \delta$ 的范围内，其中 δ 为一个足够小的正实数），则称该性质为 $f(x)$ 在 $x = a$ 时的局部性质。例如：

- (i) 我们称 $y = f(x)$ 在 $x = a$ 时有「局部极小值」（或局部极大值） $f(a)$ ，其意义为

$$f(x) \geq f(a) \quad (\text{或 } f(x) \leq f(a))$$

对于所有 x 在 $x = a$ 的足够小邻近内恒成立。

- (ii) 我们称 $y = f(x)$ 在 $x = a$ 的足够小邻近内是「递增」的（或递减的），其意义为

$$f(x_1) \leq f(x_2) \quad (\text{或 } f(x_1) \geq f(x_2))$$

能对于所有满足 $(a - \delta) < x_1 < x_2 < (a + \delta)$ 的 x_1, x_2 恒成立。

接著请同学们留意下述一个十分简单的事实。在应用泰勒公式于研究函数的局部性质的时候，它将会扮演著一个很常用的角色。

【基本事实一】：一个微小量的高次幂的绝对值要远比其低次幂的绝对值为小。例如，设

$$|\Delta x| \leq \frac{1}{10^6} = 10^{-6}$$

则

$$|\Delta x^{k+\ell}| = |\Delta x|^{k+\ell} \leq 10^{-6\ell} \cdot |\Delta x^k|$$

亦即 $|\Delta x^{k+1}|$ 的值会小于 $|\Delta x^k|$ 的值的百万分之一。

在研究函数的局部性质时，我们惯用 Δx 来代表 $(x-a)$ ，并且会假设 $|\Delta x|$ 为足够小的值。所以，在研究多项式函数 $f(x)$ 于 $x=a$ 的局部性质时中，由泰勒公式所得的展开式

$$\begin{aligned} f(x) &= f(a) + Df(a) \cdot \Delta x + \frac{1}{2}D^2f(a)\Delta x^2 + \dots \\ &= f(a) + \sum_{k=1}^n \frac{1}{k!}D^k f(a) \cdot \Delta x^k \end{aligned}$$

我们容易看见下述另一个基本事实：

【基本事实二】：设 $D^k f(a)$ 为 $D^i f(a)$ ($1 \leq i \leq n$) 列中的第一个非零值，则

$$(2.19) \quad g(x) = f(a) + \frac{1}{k!}D^k f(a)\Delta x^k \quad (\Delta x = (x-a))$$

为 $f(x)$ 于 $x=a$ 的一个「 k -阶」的局部逼近函数，亦即

$$(2.20) \quad f(x) - g(x) = \sum_{\ell=k+1}^n \frac{1}{\ell!}D^\ell f(a)\Delta x^\ell$$

的绝对值要远比 $\frac{1}{k!}D^k f(a)\Delta x^k$ 的绝对值小。因此，有很多 $f(x)$ 的局部性质都会和逼近函数 $g(x)$ 的局部性质相同，所以我们不妨改用较为简单的 $g(x)$ (一个 k 次的多项式函数) 来研究 $f(x)$ (一个次数至少为 k 的多项式函数) 的那些局部性质。上述看似单纯的想法其实业已蕴含著很有用的局部逼近方法！

【引理一】：若 $Df(a) > 0$ (或 $Df(a) < 0$)，则 $y = f(x)$ 在 $x = a$ 的邻近范围是递增的 (或递减的)。

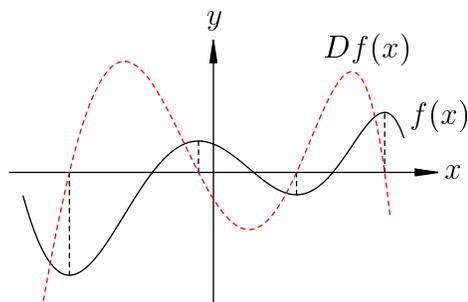
证明想法： $y = f(x)$ 可用下述函数来局部逼近：

$$f(a) + Df(a)\Delta x = f(a) + Df(a) \cdot (x-a)$$

当 $Df(a) > 0$ (或 $Df(a) < 0$) 时，上述函数显然会在 $x = a$ 邻近范围内递增 (或递减)。 $[Df(a)$ 实际上就是上述线性函数的斜率。] \square

单调区间和极大、极小点：

给出一个 n 次多项式 $f(x)$ ，则 $Df(x)$ 为一个 $(n-1)$ 次多项式。若把整条实数轴以 $Df(x)$ 的（部分）零点分割为一些开区间线段，使得在每一个区间中 $Df(x)$ 维持其正负号不变（并假设由一个区间走向相邻区间时 $Df(x)$ 会改变其正负号），则由[引理一]易知 $y = f(x)$ 在每一个区间中恒为单调递增或单调递减（由 $Df(x)$ 在其上的正负号决定之）。我们把这些区间称为函数 $f(x)$ 的「单调区间」。再者，那些单调区间的分点 $\{a_j\}$ 乃是函数 $f(x)$ 改变其单调性的转向点，易见若 $f(x)$ 由 \nearrow 改变为 \searrow ，则转向点为其局部极大点；若 $f(x)$ 由 \searrow 改变为 \nearrow ，则转向点为其局部极小点。



[图 2-1]

【例一】： $y = f(x) = Ax^2 + Bx + C, A > 0$

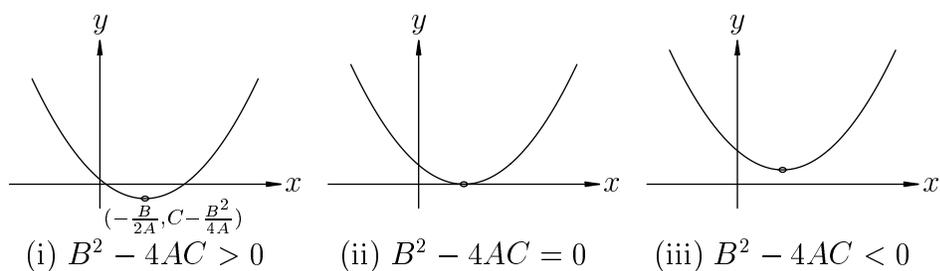
$$Df(x) = 2Ax + B, \text{ 而在 } a_1 = -\frac{B}{2A} \text{ 时 } Df(a) = 0$$

$$a \begin{cases} < \\ > \end{cases} -\frac{B}{2A}, \quad Df(a) \begin{cases} < \\ > \end{cases} 0$$

$$\Rightarrow a = -\frac{B}{2A} \text{ 乃是其极小点，而其极小值为 } f\left(-\frac{B}{2A}\right) = C - \frac{B^2}{4A}$$

其实，二次函数的极值问题亦可以用配方法得之，即

$$f(x) = A\left(x + \frac{B}{2A}\right)^2 + \left(C - \frac{B^2}{4A}\right) \Rightarrow f(x) \geq \left(C - \frac{B^2}{4A}\right)$$



[图 2-2]

【例二】： $f(x) = x^3 + Bx^2 + Cx + D$

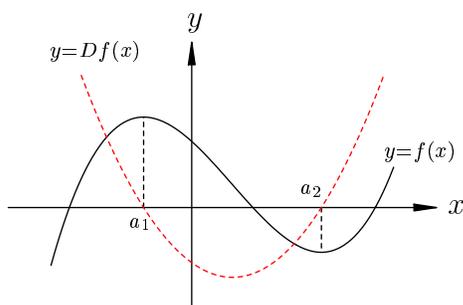
由[例一]的讨论，可见 $Df(a)$ 的正负区间有下述三种情形，即

$$(i) B^2 - 3C > 0, \quad (ii) B^2 - 3C = 0, \quad (iii) B^2 - 3C < 0$$

我们现在分别讨论上述每一个情况：

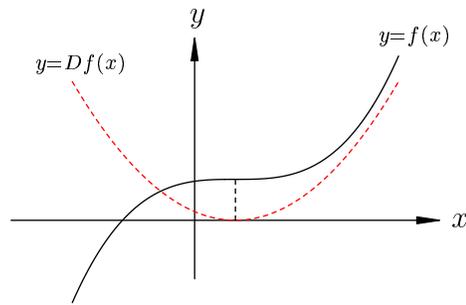
(i) $B^2 - 3C > 0$ ：在此情况， $Df(x) = 0$ 有两个相异的根 $\{a_1, a_2\}$ 。两者把实数轴分为三个 $f(x)$ 的单调区间，即

x	a_1		a_2		
$Df(x)$	+	0	-	0	+
$f(x)$	↗	→	↘	→	↗



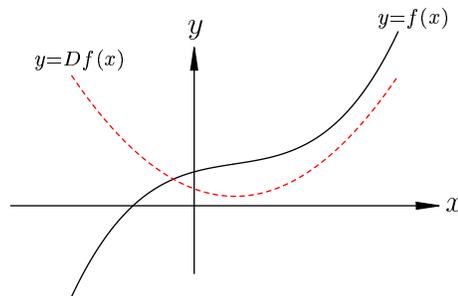
[图 2-3]

(ii) $B^2 - 3C = 0$ ：在此情况，我们恒有 $Df(x) \geq 0$ ，因此 $f(x)$ 在整条实数轴上是单调递增的。



[图 2-4]

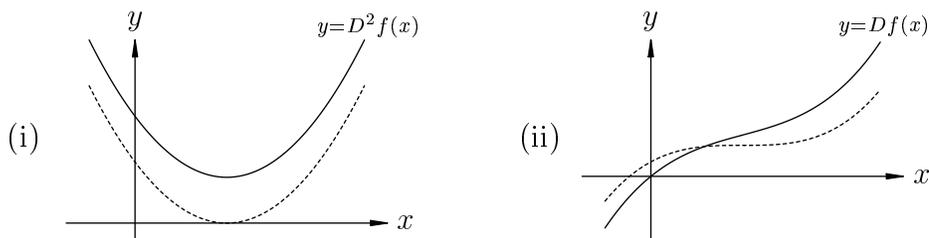
(iii) $B^2 - 3C < 0$: 在此情况, 恒有 $Df(x) > 0$, 因此 $f(x)$ 在整条实数轴上是严格单调递增的。



[图 2-5]

【例三】: $f(x) = x^4 + Bx^3 + Cx^2 + Dx + E$

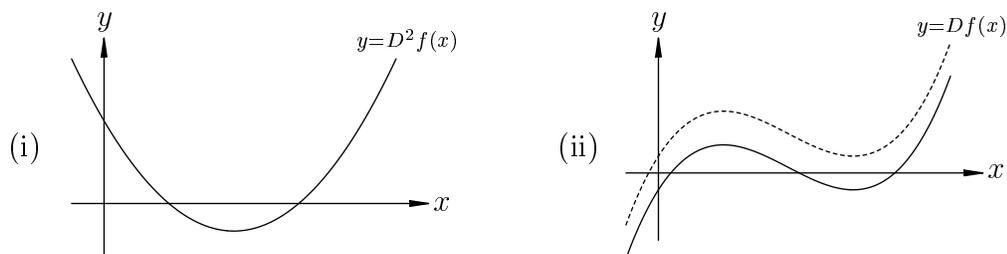
易见 $Df(x) = 4x^3 + 3Bx^2 + 2Cx + D$ 和 $D^2f(x) = 12x^2 + 6Bx + 2C$, 我们依然尝试用二次的 $D^2f(x)$ 来把 $f(x)$ 分类。假设 $D^2f(x)$ 的样子是属于[例一]的 (ii) 或 (iii) 型, 即如 [图 2-6(i)] 所示的虚线或实线图形:



[图 2-6]

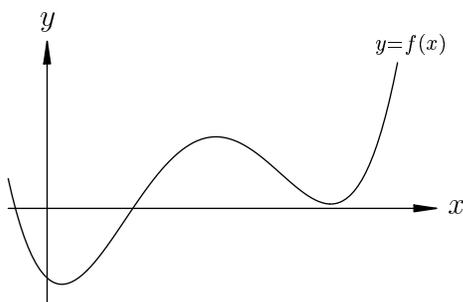
则 $Df(x)$ 的样子就会是如 [图 2-6(ii)] 所示者, 亦即 $f(x)$ 在这两种情况

之下都只有两个单调区间。另一方面，假若 $D^2f(x)$ 的样子是[例一]的(i)型，即如[图 2-7(i)]所示者：



[图 2-7]

则 $Df(x)$ 的样子会如[图 2-7(ii)]所示，但有两种不同位置。当 $Df(x)$ 是[图 2-7(ii)]中虚线所示的位置， $f(x)$ 只有两个单调区间；而当 $Df(x)$ 是[图 2-7(ii)]中实线所示的位置， $f(x)$ 就会有四个单调区间，即如下图所示：



[图 2-8]

函数的 k -阶局部逼近：

在局部分析中，我们还可引入下述关于「微小量的阶」和「逼近的阶」的概念，即

【定义】：一个微小量 Δx 的函数称为「 k -阶微小量」若它可写成 Δx^k 的常数倍再加上一些更高阶的微小量。

【定义】：一个函数 $g(x)$ 称为函数 $f(x)$ 在 $x = a$ 的「 k -阶局部逼近」若 $|f(x) - g(x)|$ 为足够小的 Δx 的 $(k+1)$ -阶微小量， $\Delta x = (x - a)$ 。

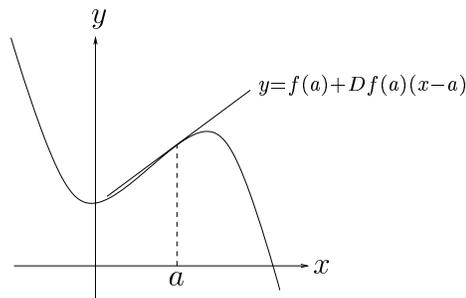
【例子】：

(i) 下述线性（即次数为 1）多项式函数

$$y = f(a) + Df(a)(x - a)$$

为 $y = f(x)$ 于 $x = a$ 的 1-阶局部逼近。

从几何方面来看，线性多项式的图象为一条直线。所以上述的线性局部逼近函数的图象就是在 $y = f(x)$ 的图象中于 $(a, f(a))$ 点的切线。



[图 2-9]

(ii) 下述函数

$$y = f(a) + Df(a)\Delta x + \dots + \frac{1}{k!}D^k f(a)\Delta x^k = \sum_{j=0}^k D^j f(a)\Delta x^j$$

为 $y = f(x)$ 于 $x = a$ 的 k -阶局部逼近。

【习题】：

(1) 试找出下列多项式函数的单调区间：

(i) $y = x^2 + 2x - 3$

(ii) $y = -5x^2 + 20x - 14$

(iii) $y = x^3 - 3x^2 - 45x + 17$

(iv) $y = x^3 - 3x^2 + 13x - 10$

$$(v) y = 9x^4 - 34x^3 + 20x^2 + 50x - 6$$

- (2) 对于每一个在题 (1) 中所列出的多项式函数, 若它们的极大点或极小点存在, 试找出它们的所在位置。
- (3) 令 $D^k f(a)$ 为 $D^i f(a)$ ($1 \leq i \leq n$) 列中的第一个非零值。
- (i) 设 k 为偶数。试证明 $x = a$ 其实是 $f(x)$ 的一个局部极值点。试讨论在什麼情况下它是一个局部极大点或局部极小点。
- (ii) 设 k 为奇数。试证明当 $D^k f(a) > 0$ 时 $f(x)$ 是在 a 点邻近局部递增的; 而当 $D^k f(a) < 0$ 时 $f(x)$ 是在 a 点邻近局部递减的。
- (4) 设 $D^2 f(a) > 0$ (或 $D^2 f(a) < 0$)。试证明 $y = f(x)$ 在 $x = a$ 邻近的图象是位于其在 $x = a$ 的切线之上 (或之下)。试描绘出以上两种情况的局部图象之相对位置。
- (5) 若 $D^2 f(a) = 0$ 但 $D^3 f(a) \neq 0$, 则 $f(x)$ 在 $x = a$ 的局部图象与 $f(x)$ 在 $x = a$ 的切线之相对位置如何?
- (6) 在题 (5) 中所描述之点称为拐点。试求出题 (1) 的 (iii) 及 (iv), 以及第一章第一节习题内的 3 次多项式函数的拐点。

第三章

多项式函数的微积分

概括地说，一个函数 $y = f(x)$ 所描述者，乃是一个变数的值如何随著另一个变数的改变而变化的方式。由此可见，函数的「变化速度」乃是有著根本的重要性。另一方面，从宏观的角度来说，函数的整体效应亦有著根本的重要性，它纪录了这个函数在某一个范围内所作的影响的「总和」。

在这一章中，我们将会对上述两种具有根本重要性的函数性质——「变率」与「总和」——做一次探本究源的工作。在多项式函数的范畴中，两者则分别对应于上两章所讨论的泰勒公式和求和公式。

3.1 变率与微分

首先让我们看看线性函数的情形作为简单例子。设 $y = Ax + B$ ，当 x 由 x_1 改变至 x_2 时，则相应会有 y 由 $y_1 = Ax_1 + B$ 改变至 $y_2 = Ax_2 + B$ 。因此 y 的改变值与 x 的改变值之间的比率为：

$$(3.1) \quad \frac{y_2 - y_1}{x_2 - x_1} = \frac{A(x_2 - x_1)}{x_2 - x_1} = A$$

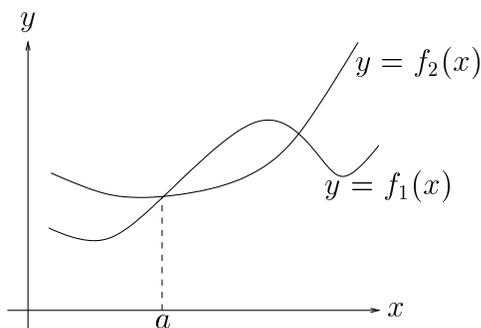
所以 $y = Ax + B$ 的「变率」就自然而然地定义为上述常数 A 。反之，若某一个函数 $y = f(x)$ 在每一点的「变率」恒等于常数 A ，则易见 $y = Ax + f(0)$ 。

但对于其他的多项式函数而言（即使像 $y = x^2$ 这样简单的函数），上面所计算的「 x, y 改变值之比」是会随著 (x_1, x_2) 的不同位置而不

同者，所以这类函数的「变率」其实还是一个「尚待定义」的概念！无论如何，现在让我们先来重新剖析一下「变率」这个概念的直观内含。

令 $y = f_1(x)$ 和 $y = f_2(x)$ 为两个函数，它们在 $x = a$ 点的值相同，即 $f_1(a) = f_2(a)$ 。设

$$\begin{cases} \text{当 } a - \delta < x < a, & f_1(x) < f_2(x) \\ \text{但当 } a < x < a + \delta, & f_1(x) > f_2(x) \end{cases}$$



[图 3-1]

亦即在 $x = a$ 点时 $f_1(x)$ 「从後赶上」 $f_2(x)$ 而且超越之。在这个情况下， $f_1(x)$ 在 $x = a$ 的「变率」显然是不可能小于 $f_2(x)$ 在 $x = a$ 的「变率」的。假若不然，则我们在讨论的东西和「变率」的直观意义不符；换句话说，若所讨论的概念并不满足上述比较条件，则它绝不会是我们想要研究的「变率」！

或者同学们试想想下述一个实际例子：设有甲、乙两人在一条公路上作自行车竞赛，分别以 $f_1(x)$ 和 $f_2(x)$ 表示甲、乙在 x -秒後和起点的距离。设甲在 $x = a$ 时从後赶上乙而且超越之，亦即在 x 略小于 a 时，甲在乙之後；但是在 x 略大于 a 时，则甲在乙之前。则甲在 $x = a$ 时的「速率」显然不能小于乙者，要不然，则甲是不可能于 $x = a$ 时达成後来居上的，是不？

把上述直观上十分明显的事实，改用函数框架叙述之，即为下述刻划变率的直观内含的比较原则：

【变率的比较原则】(Comparison principle of rate of change)：

设有两个函数 $f_1(x)$ 和 $f_2(x)$ 满足 $f_1(a) = f_2(a)$ ，而且在 $x = a$ 的邻近具有下述大小关系，即对于一个足够小的 $\delta > 0$,

$$(3.2) \quad \begin{cases} \text{当 } a - \delta < x < a, & f_1(x) < f_2(x) \\ \text{当 } a < x < a + \delta, & f_1(x) > f_2(x) \end{cases}$$

则 $f_1(x)$ 在 $x = a$ 的变率 不小于 $f_2(x)$ 在 $x = a$ 的变率。

根据上述比较原则和已知的线性函数变率，我们将会证明多项式函数 $y = f(x)$ 在 $x = a$ 的变率就是 $Df(a)$ 。

【例一】：且让我们先以 $f(x) = cx^3$ 为例，用上述比较原则去研讨它在 $x = a$ 点的变率应该是什麼。因为一次函数的变率乃是已知者，我们可以把 $f(x)$ 和 $g(x) = ca^3 + m(x - a)$ (注意： $f(a) = g(a)$) 来比较。令 $x = a + \delta$ ，则有

$$\begin{aligned} f(x) &= c(a + \delta)^3 = ca^3 + 3ca^2\delta + 3ca\delta^2 + c\delta^3 \\ g(x) &= ca^3 + m\delta \\ f(x) - g(x) &= (3ca^2 - m)\delta + \delta^2(3ca + c\delta) \end{aligned}$$

由此不难看到，当 $(3ca^2 - m) \neq 0$ 时，只要把 $|\delta|$ 取得足够小，则 $(f(x) - g(x))$ 和 $(3ca^2 - m)\delta$ 同号。所以由比较原则 (3.2) 即有

$$\begin{aligned} m > 3ca^2 &\Rightarrow f(x) \text{ 在 } x = a \text{ 点的变率} \leq m \\ m < 3ca^2 &\Rightarrow f(x) \text{ 在 } x = a \text{ 点的变率} \geq m \end{aligned}$$

所以「 $f(x)$ 在 $x = a$ 点的变率」必须小于任何大于 $3ca^2$ 的 m ，又必须大于任何小于 $3ca^2$ 的 m ，所以唯有把它定义为 $3ca^2$ 才合乎上述比较原则。

【定理 3.1】：一个多项式函数 $y = f(x)$ 在 $x = a$ 的变率的唯一合理定义是 $Df(a)$ 。

证明：其实，只要把[例一]的方法的直接推广，就可以用来证明[定理 3.1]。给定一个实数 m ，令

$$(3.3) \quad g(x) = f(a) + m(x - a)$$

另一方面，由 $f(x)$ 的泰勒展开式可得

$$(3.4) \quad f(x) = f(a) + Df(a)(x-a) + \sum_{k=2}^n \frac{1}{k!} D^k f(a)(x-a)^k$$

由此可见，令 $\delta = (x-a)$ ，则有

$$(3.5) \quad f(x) - g(x) = (Df(a) - m)\delta + \delta^2 \left(\sum_{k=2}^n \frac{1}{k!} D^k f(a) \delta^{k-2} \right)$$

同理，当 $(Df(a) - m) \neq 0$ 时，只要把 $|\delta|$ 取得足够小，则 $(f(x) - g(x))$ 和 $(Df(a) - m)\delta$ 同号。所以由比较原则 (3.2) 即有

$$\begin{aligned} m > Df(a) &\Rightarrow f(x) \text{ 在 } x=a \text{ 点的变率} \leq m \\ m < Df(a) &\Rightarrow f(x) \text{ 在 } x=a \text{ 点的变率} \geq m \end{aligned}$$

所以「 $f(x)$ 在 $x=a$ 点的变率」的唯一合理定义就是 $Df(a)$ 。□

[注]：在上一章中，运算符号 D 只是形式上的运算，我们引进它来简洁地把二项定理推广成泰勒公式。而上述[定理 3.1]则说明了 $Df(x)$ 在分析学上的意义——它刚好就是记录著 $f(x)$ 在每一点上的变率。这种变率的概念其实可以推广至适用于更广泛的函数类，记录著一个一般函数的变率之函数我们称为该函数的「导函数」(derivative function)，而求出导函数的运算过程则称为「微分」(differentiation)。由此可见，在上一章的形式代数运算其实就是「多项式的微分」。至于如何把泰勒公式推广成能应用于更为广泛的函数类型，这正是整个微分学的基本骨干。

插值问题的推广：

在第一章中所讨论的插值问题，乃是要求待求公式必须在某些特定地方取某些特定值。但在实际的应用问题上，有时候还要求待求公式的图象在该点有特定变率、特定曲率等等。例如：设有一个 4 次多项式 $f(x)$ ，它在 $x = \pm 1$ 和 $x = 2$ 三点的值已给定，而且其变率在 $x = \pm 1$ 时也给定，问能否唯一地确定 $f(x)$ ？换句话说，能否由下述五个给值条件唯一地确定一个 4 次多项式 $f(x)$ ？

$$f(-1), Df(-1), f(1), Df(1), f(2)$$

且以上述数值分别为 9, -6, 1, 6, 27 为例说明其解法。

首先, 在 $x = -1, 1, 2$ 分别取值 9, 1, 27 的 2 次多项式 $g(x)$ 是可以用插值公式唯一地确定之, 即

$$g(x) = 10x^2 - 4x - 5$$

留意 $f(x) - g(x)$ 在 $x = -1, 1, 2$ 这三点皆为 0, 亦即 $f(x) - g(x)$ 必定具有因式 $(x+1)(x-1)(x-2)$ 。由于 $f(x) - g(x)$ 乃是 4 次多项式, 易见有

$$f(x) - g(x) = (x+1)(x-1)(x-2)(ax+b)$$

其中 a, b 为两个待定的系数, 亦即

$$f(x) = (ax+b)(x+1)(x-1)(x-2) + 10x^2 - 4x - 5$$

对于上述 $f(x)$ 求导, 即有

$$Df(x) = (ax+b)(x-1)(x-2) + (ax+b)(x+1)(x-2) + 20x - 4 + \dots$$

代入条件式 $Df(-1) = -6, Df(1) = 6$, 得

$$-6 = Df(-1) = (-a+b)(-2)(-3) + 0 - 20 - 4 + 0 \Rightarrow -a+b = 3$$

$$6 = Df(1) = 0 + (a+b)(2)(-1) + 20 + 4 + 0 \Rightarrow a+b = 5$$

易解得 $a = 1, b = 4$ 。所以

$$f(x) = (x+4)(x^2-1)(x-2) + g(x) = x^4 + 2x^3 + x^2 - 6x + 3$$

【习题】：

- (1) 试求出藏于半径为 r 的球体内的圆柱体之最大体积。
- (2) 试求出藏于半径为 r 的球体内的圆锥体之最大体积。
- (3) 试找出一个不高于 5 次的多项式 $f(x)$, 它满足下述条件：

$$f(-1) = 1$$

$$f(0) = 1, Df(0) = -1, D^2f(0) = 8$$

$$f(1) = 5, Df(1) = 14$$

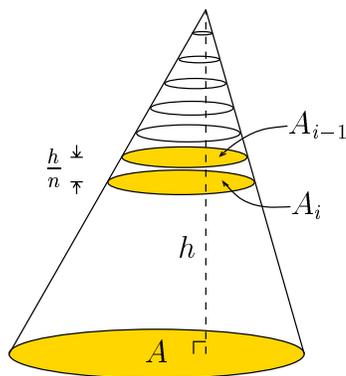
3.2 总和与积分

我们首先描述一个在历史上具有重要意义的例子，它应用了求和公式来解决了一个有基本重要性的几何问题。

假如我们用一个圆锥形的杯装水倒进一个同底同高的圆柱形的杯时，会发现需要三满杯圆锥形杯的水才可注满一个圆柱形杯。相信在古代从类似的实验也自然地引导古代中国和古希腊的几何学家发现了圆锥体的体积公式，亦即圆锥体体积为同底同高的圆柱体体积之三分之一。它是三角形面积公式的「高一维」推广，因为三角形面积就是同底同高的平行四边形的面积之一半。在二维的情况我们很容易就可以把一个平行四边形切割成两个全等的三角形，但是在三维的情况我们并没有这种简单的切割重组方法！相信有很多古希腊的几何学家曾努力尝试寻找类似的切割重组方法求证锥体的体积公式，但所有尝试皆终归失败，直至最后由 Eudoxus 开创了一套崭新的方法锥体体积公式才能成功得证。这其实是最早的一个具有基本重要性的积分，也是现代积分学之源起。

锥体体积公式的 Eudoxus 证明：

如 [图 3-2] 所示，一个底面积为 A 、高度为 h 的锥体可以用平行于底面的平面切割成 n 块均一厚度的薄片：



[图 3-2]

由顶层向下数的第 i 块薄片，其底面积为 A_i ，顶面积为 A_{i-1} 。由相似

形定理，古希腊的几何学家已知

$$(3.6) \quad A_i : A = \left(\frac{i}{n} h : h \right)^2 = \left(\frac{i}{n} \right)^2, \quad \text{即 } A_i = \frac{i^2}{n^2} A$$

由此易知其第 i 块薄片之体积是介乎于 $\frac{h}{n} \cdot A_{i-1}$ 和 $\frac{h}{n} \cdot A_i$ 两者之间，亦即

$$(3.7) \quad \frac{h}{n} A_{i-1} = \frac{(i-1)^2}{n^3} hA < V_i < \frac{i^2}{n^3} hA = \frac{h}{n} A_i, \quad 1 \leq i \leq n$$

把上述的不等式整合起来，便得出

$$(3.8) \quad \frac{hA}{n^3} \sum_{i=1}^n (i-1)^2 < \sum_{i=1}^n V_i = V < \frac{hA}{n^3} \sum_{i=1}^n i^2$$

接著 Eudoxus 便应用已知的 $\sum i^2$ 求和公式把上式重写成

$$(3.9) \quad \frac{hA}{3} \left(1 - \frac{1}{n} \right) \left(1 - \frac{1}{2n} \right) < V < \frac{hA}{3} \left(1 + \frac{1}{n} \right) \left(1 + \frac{1}{2n} \right)$$

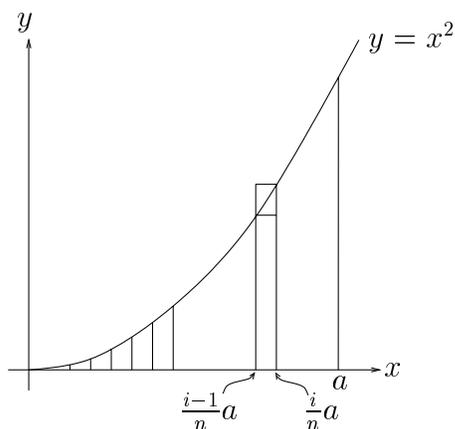
至此，Eudoxus 发现上述不等式对所有正整数 n 都成立，但无论 n 如何增大， V 的值（锥体体积）是不变的！而当 n 无限地增大时，上式的上限和下限之差别为

$$(3.10) \quad \frac{hA}{3} \left(1 + \frac{1}{n} \right) \left(1 + \frac{1}{2n} \right) - \frac{hA}{3} \left(1 - \frac{1}{n} \right) \left(1 - \frac{1}{2n} \right) = \frac{hA}{n}$$

这个量会无限地缩小（即可小于任意给出的正实数）。由此可见，唯一可以对所有 n 能满足上述不等式的量必定是 $\frac{hA}{3}$ ，因此得证

$$(3.11) \quad V = \frac{1}{3} hA$$

随后，古希腊的几何学家也用同样的方法来求得曲线 $y = x^2$ 底下的面积：



[图 3-3]

如 [图 3-3] 所示, 图中被曲线 $y = x^2$ 、直线 $x = a$ 和 x -轴所包围的区域可被切割成 n 条均一宽度为 $\frac{1}{n}a$ 的窄条。由左数起第 i 条窄条的面积显然小于其外包矩形之面积, 但大于其内含矩形之面积, 亦即

$$(3.12) \quad \frac{a}{n} \left(\frac{i-1}{n} a \right)^2 = \frac{a^3}{n^3} (i-1)^2 < A_i < \frac{a^3}{n^3} i^2 = \frac{a}{n} \left(\frac{i}{n} a \right)^2, \quad 1 \leq i \leq n$$

再把上述 n 个不等式整合起来, 即得

$$(3.13) \quad \frac{a^3}{n^3} \sum_{i=1}^n (i-1)^2 < \sum_{i=1}^n A_i = A < \frac{a^3}{n^3} \sum_{i=1}^n i^2$$

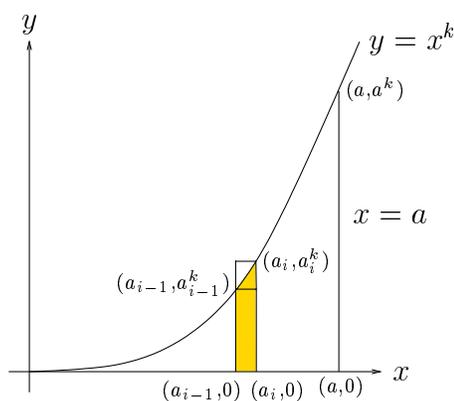
因此, 运用同一求和公式可得

$$(3.14) \quad \frac{a^3}{3} \left(1 - \frac{1}{n} \right) \left(1 - \frac{1}{2n} \right) < A < \frac{a^3}{3} \left(1 + \frac{1}{n} \right) \left(1 + \frac{1}{2n} \right)$$

再以同样的论证, 即得那个能满足所有以上不等式的 A 必定是 $\frac{a^3}{3}$, 这样就证明了

$$(3.15) \quad A = \frac{a^3}{3}$$

显然这个方法是可以直接推广到求曲线 $y = x^k$ 、直线 $x = a$ 和 x -轴所包围的区域之面积, 其值为 $\frac{a^{k+1}}{(k+1)}$ 。



[图 3-4]

证明：如 [图 3-4] 所示，由左边数起第 i 条窄条是夹于该两个矩形之间，其面积满足

$$(3.16) \quad \frac{a}{n} \left(\frac{i-1}{n} a \right)^k = \frac{a^{k+1}}{n^{k+1}} (i-1)^k < A_i < \frac{a^{k+1}}{n^{k+1}} i^k = \frac{a}{n} \left(\frac{i}{n} a \right)^k, \quad 1 \leq i \leq n$$

把上述不等式整合起来，得

$$(3.17) \quad \frac{a^{k+1}}{n^{k+1}} \sum_{i=1}^n (i-1)^k < \sum_{i=1}^n A_i = A < \frac{a^{k+1}}{n^{k+1}} \sum_{i=1}^n i^k$$

现在，我们需要运用第一章的求和公式来求得上述两个总和，即 $\sum_{i=1}^n (i-1)^k$ 和 $\sum_{i=1}^n i^k$ 。令

$$(3.18) \quad x^k = c_k g_k(x) + c_{k-1} g_{k-1}(x) + \dots$$

由比较 x^k 的系数即可得出 $c_k = k!$ ，因此

$$(3.19) \quad \begin{aligned} \sum_{i=0}^{(n-1)} i^k &= (k!) \cdot g_{k+1}(n) + c_{k-1} g_k(n) + \dots \\ &= \frac{1}{k+1} n^{k+1} + \text{至多是 } n \text{ 的 } k \text{ 次的项} \end{aligned}$$

所以

$$(3.20) \quad \frac{1}{n^{k+1}} \sum_{i=0}^{(n-1)} i^k = \frac{1}{k+1} + \text{只含 } n \text{ 的负指数的项}$$

当 n 无限地增大时, 那些负指数的项自然会无限地缩小, 由此可得

$$(3.21) \quad A = \frac{a^{k+1}}{k+1}$$

用现今的积分记号, 上式即可写成下述基本的多项式函数积分公式:

$$(3.22) \quad \int_0^a x^k dx = \frac{a^{k+1}}{k+1}$$

【习题】:

- (1) 运用 $\sum_{i=0}^{(n-1)} i^3$ 的求和公式写出 $\int_0^a x^3 dx = \frac{a^4}{4}$ 的详细证明。
- (2) 由曲线 $y = x^2 + x$ 、直线 $x = a$ 和 x -轴所包围的区域之面积为多少?
- (3) 由曲线 $y = 4x^3 + 2x^2 + 4x$ 、直线 $x = a$ 和 x -轴所包围的区域之面积为多少?
- (4) 由曲线 $y = 3x^4 + 2x^2$ 、直线 $x = a$ 、 $x = b$ 和 x -轴所包围的区域之面积为多少?

第四章

线性方程组与行列式的基础理论

我们在前三章已经和同学们简明扼要地讨论了一元高次多项式的基础理论。现在让我们改弦更张，转为讨论多元代数表达式的基础理论。其中最为简朴者，当然就是多元一次的表达式；而其方程之求解，乃是同学们在小学和初中时期常常碰见的求解线性方程组 (system of linear equations) 是也。

求解线性方程组乃是代数学中的基本课题和必用、常用的基本功。在很多的代数解题方法中，例如待定系数法，我们必需通过一些联立方程来解出待定系数的值，而其中比较简单基本者，则往往是可以设法用线性方程组来表达之。同学们在小学和初中的阶段已熟习了运用代入法 (substitution) 和消元法 (elimination) 求解一些简单的联立方程，所以我们将先从这些简明的方法入手，作为研讨线性方程组的解法和基础理论的起点。

本质上， n -阶行列式乃是一个特定的 n^2 元、 n 次多项式。它乃是在研讨 n 元线性方程组的基础理论中自然而然地发现者，而且在其中扮演著主角。本章将采取返璞归真的处理方式，引领同学们由二元、三元、四元逐步归纳、分析探索线性方程组和行列式的基础理论，作一次归纳发现、归纳定义和归纳论证，顺理成章的三结合。其实，在各式各样的多元高次多项式之中，大多数是无用者或是难以应用者；唯独有行列式，因为它所独有的优良性质，使得在不同的数学范畴的研讨中（包括代数、分析、几何），它都是一个不可或缺、精简好用的重要工具。

4.1 代入法和消元法

让我们先来研讨一个线性方程组的解集究竟会出现那些可能性。若所涉及的线性方程组只是一个一元一次方程，即 $ax = b$ ，则结论非常简单：当 $a \neq 0$ 时，我们有唯一解 $x = b/a$ ；当 $a = 0$ 但 $b \neq 0$ 时，方程无解；而当 $a = 0$ 和 $b = 0$ 时，则任何 x 都是解，亦即有无穷多解。

而同学们在小学中所碰见的「鸡兔同笼」问题，其实就是一种二个二元一次方程的求解。例如：笼中有鸡、兔共 36 只，而脚数共有 108 只，问鸡和兔各有多少只？设鸡数为 x ，兔数为 y ，则有

$$\begin{cases} x + y = 36 \\ 2x + 4y = 108 \end{cases}$$

由第一式解得 $x = 36 - y$ ，再代入第二式中，得

$$\begin{aligned} 2(36 - y) + 4y &= 108 \\ 72 + 2y &= 108 \\ y &= 18 \end{aligned}$$

再将 $y = 18$ 代回第一式中求得 $x = 18$ 。所以鸡、兔各有 18 只。

既然「鸡兔同笼」可解，我们很自然会想到其它如「鸡鸭同笼」、「龟兔同笼」等等问题是否也同样可解？当我们尝试解一些实例时，便会发现事实并非如此，例如：若笼中有鸡、鸭共 36 只，而脚数共有 108 只，问鸡和鸭各有多少只？设鸡数为 x ，鸭数为 y ，则有

$$\begin{cases} x + y = 36 \\ 2x + 2y = 108 \end{cases}$$

当我们由第一式得 $x = 36 - y$ ，而代入第二式时，却得出一个矛盾：

$$\begin{aligned} 2(36 - y) + 2y &= 108 \\ 72 &= 108 \quad !!! \end{aligned}$$

所以问题无解。又如笼中有龟、兔共 27 只，而脚数共有 108 只，问龟和兔各有多少只？设龟数为 x ，兔数为 y ，则有

$$\begin{cases} x + y = 27 \\ 4x + 4y = 108 \end{cases}$$

当我们由第一式得 $x = 27 - y$ ，而代入第二式时，这次得出

$$4(27 - y) + 4y = 108$$

$$108 = 108$$

所以第二式基本上是没有加上额外的限制，换句话说，只需要满足第一式就是方程组的解。这样，我们便会有很多解：如 $x = 0, y = 27$ 或 $x = 12, y = 15$ 等等都是解。

经过上述的讨论，我们可以看到在求解二个二元一次方程时仍然出现三种情况，即有唯一解、无解或有无穷多解；而且究竟是属于那种情况又和方程组的系数有著某种关系。

至于同学们需要对三个三元一次方程求解，一般会在高中时于讨论圆的坐标方程时遇到。例如：求过平面上给定三点 $(4, 2), (3, 5), (2, -2)$ 的圆的坐标方程。通常的解法是：我们设所求的方程为 $x^2 + y^2 + Dx + Ey + F = 0$ （其实这就是一种待定系数法），则由所设可写出下列方程：

$$(4.1) \quad 16 + 4 + 4D + 2E + F = 0$$

$$(4.2) \quad 9 + 25 + 3D + 5E + F = 0$$

$$(4.3) \quad 4 + 4 + 2D - 2E + F = 0$$

现在我们可以用消元法来消去其中一元。例如，用 (4.2)-式减去 (4.1)-式及用 (4.2)-式减去 (4.3)-式後，我们便得出下述两式：

$$(4.4) \quad 14 - D + 3E = 0$$

$$(4.5) \quad 26 + D + 7E = 0$$

这样便消去了其中一元 F ，变成二个二元一次方程的情形。再将上述两式相加（目的是消去 D ），便得出：

$$(4.6) \quad 40 + 10E = 0$$

由此解得 $E = -4$ 。代入 (4.5)-式得 $26 + D - 28 = 0$ ，即 $D = 2$ 。最後，把 $E = -4, D = 2$ 代入 (4.3)-式得 $8 + 4 + 8 + F = 0$ ，求得 $F = -20$ 。因此该圆的坐标方程为 $x^2 + y^2 + 2x - 4y - 20 = 0$ 。

若所给定的(相异)三点并列在某一直线时,则由几何直观知问题应该无解。若所给定三点有其中两点相重时,则易见会有无穷多解。现在让我们再用另一个例子来研讨三个三元一次方程的解。试求下述三个平面在空间内的交集:

$$(4.7) \quad x + 2y - z = 1$$

$$(4.8) \quad 3x - y + 2z = 4$$

$$(4.9) \quad x - 5y + 4z = 3$$

用(4.8)-式减去(4.7)-式的3倍及用(4.8)-式减去(4.9)-式的3倍后,我们便得出下述两式:

$$(4.10) \quad -7y + 5z = 1$$

$$(4.11) \quad 14y - 10z = -5$$

易见上述 y, z 的方程组无解。若把原来方程组的第三式(4.9)的常数项改为2,即

$$(4.12) \quad x + 2y - z = 1$$

$$(4.13) \quad 3x - y + 2z = 4$$

$$(4.14) \quad x - 5y + 4z = 2$$

则相应的 y, z 方程为

$$(4.15) \quad -7y + 5z = 1$$

$$(4.16) \quad 14y - 10z = -2$$

上述的 y, z 方程组会有无穷多解,因此原本的方程组也会有无穷多解:例如 $x = \frac{6}{7}, y = \frac{4}{7}, z = 1$ 和 $x = \frac{9}{7}, y = -\frac{1}{7}, z = 0$ 等等都是其解。

所以,三个三元一次方程组也是有唯一解、无解和有无穷多解这三种情况,而且究竟是属于那一种情况又和方程组的系数集有著某种关系。再者,在方程组有唯一解的情况,该唯一解与方程组系数集之间理当有一种特殊的函数关联加以表达之;至于这种特殊的函数表达式究竟是如何,乃是我们将会在下一节详细探讨者也。

实给系数的多元一次方程组的求解：

其实，当一个多元一次方程组的系数乃是实给者，我们并不需要什么特别的理论，就可以用代入消元法求解之，其原理简单明了，即：

- (i) 选用其一，解得其中系数非零之一个变元，以其他诸元的线性表达式表达之；
- (ii) 用它代入其他各式即可得方程个数和元数各减其一的多元一次方程组。

例如，设第一条方程 $a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n = c_1$ 满足 $a_{11} \neq 0$ ，由此即可解得

$$x_1 = \frac{1}{a_{11}}(c_1 - a_{12}x_2 - \dots - a_{1n}x_n)$$

代入第 i -式 ($i \geq 2$) 之所得就是：

$$(a_{i2} - \frac{a_{i1}}{a_{11}}a_{12})x_2 + \dots + (a_{in} - \frac{a_{i1}}{a_{11}}a_{1n})x_n = (c_i - \frac{a_{i1}}{a_{11}}c_1)$$

亦即将第 i -式换成：

$$(\text{第 } i\text{-式}) - \frac{a_{i1}}{a_{11}} \times (\text{第 } 1\text{-式})$$

如此逐步代入消元，就可能会有下列两种蜕化 (degenerate) 情形发生，即：

其一：某一式变成两侧之系数都为 0；

其二：某一式变成左侧系数全为零，但是右侧常数项则不为 0。

前者出现的实质意义是：该式乃是那些选用来解元代入的方程式的线性组合，所以可以略去不计；而後者出现的实质意义则是：该式和那些选用来解元代入者是互相矛盾的，因此所给方程组是无解者也！所以後者一出现即可断言所给方程组无解，自然就不必再做徒劳无解的虚功。反之，设後者一直不出现，则所给方程组有解。

为了便于下面的叙述，我们不妨先行将方程和变元的编号作一调整，使得 $a_{11} \neq 0$ ；而且用第 1-式解得 x_1 代入消元之後的第 2-式的 x_2 的系数也不为零，即

$$a'_{22} = (a_{22} - \frac{a_{21}}{a_{11}}a_{12}) \neq 0$$

再者，以它解得 x_2 代入消元之後的第 3-式的 x_3 的系数又不为零。以此类推，在第 k 次代入消元後所得之 x_{k+1} 的系数依然不为零。如此逐步代入消元之所得者，乃是下述简单的方程组，它和原给者具有同解：

$$\begin{aligned}
 & a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n = c_1 \\
 & a'_{22}x_2 + \dots + a'_{2n}x_n = c'_2 \\
 & a''_{33}x_3 + \dots = c''_3 \\
 & \qquad \qquad \qquad \vdots \qquad \qquad \qquad \vdots
 \end{aligned}
 \tag{*}$$

若最後一个方程是 $a^*_{nn}x_n = c^*_n$ ，则可以用它解得 $x_n = c^*_n/a^*_{nn}$ ，然後逐步反代而求解 $x_{n-1}, x_{n-2}, \dots, x_1$ 之唯一解。若最後一个方程是：

$$a^*_{m,m}x_m + \dots + a^*_{m,n}x_n = c^*_m, \quad (m < n)$$

则可以用它解得

$$x_m = \frac{1}{a^*_{m,m}}(c^*_m - a^*_{m,m+1}x_{m+1} - \dots - a^*_{m,n}x_n)$$

它是将 x_m 表达成 (x_{m+1}, \dots, x_n) 的线性函数的一个表达式。将它逐步反代，即可解得 x_{m-1}, \dots, x_1 也表成 (x_{m+1}, \dots, x_n) 的线性函数的各别表式，亦即

$$\begin{cases}
 x_1 = \ell_1(x_{m+1}, \dots, x_n) \\
 x_2 = \ell_2(x_{m+1}, \dots, x_n) \\
 \vdots \qquad \qquad \qquad \vdots \\
 x_m = \ell_m(x_{m+1}, \dots, x_n)
 \end{cases}$$

乃是所给方程组之通解，其中 (x_{m+1}, \dots, x_n) 可以取任何之值，所以是有无穷多个解。

总结上述简短的讨论，即得下述对于多元一次线性方程组的基本认识：

1. 任何实给系数的线性方程组，都可以用代入消元法，逐步消元然後再逐步反代，直截了当地求解。
2. 线性方程组之解有三种情形，即唯一解、无穷多个解和无解，而它们都会在代入消元法中自然地确定之。

3. 代入消元法的具体运算，其实就是有系统地用所给方程组作适当的「线性组合」，把它转换成十分简单易解的 (*)-型式。其计算可以用系数的行、列之间的相应计算（亦即分离系数法）加以处理 [这也就是所谓高斯消元法 (gaussian elimination)，其实此法中国在宋元时代即以采用]。

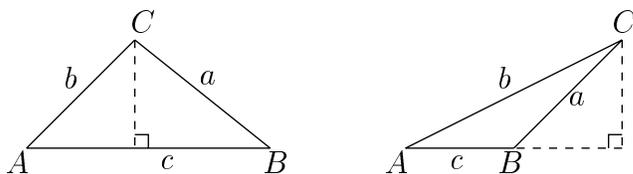
【习题】：

- (1) 在一个三角形 $\triangle ABC$ 中，易证其内角 A, B, C 和对边边长满足下述方程组

$$c = b \cos A + a \cos B$$

$$b = a \cos C + c \cos A$$

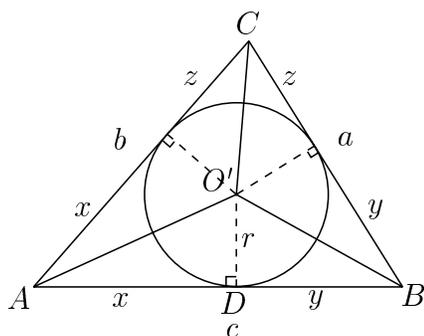
$$a = c \cos B + b \cos C$$



[图 4-1]

试由上述方程组解出 $\cos A, \cos B, \cos C$ 。

- (2) 设 O 点为三角形 $\triangle ABC$ 的内心（内切圆圆心）， x, y, z 如图所示为 O 在三边的垂足到 $\triangle ABC$ 顶点的距离，试以 a, b, c 表达 x, y, z 。



[图 4-2]

(3) 求解下述线性方程组 (如果有解) :

$$\begin{cases} x + \frac{1}{2}y + \frac{1}{3}z = 1 \\ \frac{1}{2}x + \frac{1}{3}y + \frac{1}{4}z = 1 \\ \frac{1}{3}x + \frac{1}{4}y + \frac{1}{5}z = 1 \end{cases}, \quad \begin{cases} x + 2y + 3z = 1 \\ 2x + 3y + 4z = 2 \\ 3x + 4y + 5z = 3 \end{cases}, \quad \begin{cases} x + 3y + 5z = 1 \\ 2x + 4y + 6z = 4 \\ 3x + 5y + 7z = 9 \end{cases}$$

(4) 在平面上四点 $(6, 4)$, $(7, 1)$, $(2, 6)$, $(-1, -3)$ 是否共圆 (cocircular)? 若是共圆的情形, 求该圆之方程。

(5) 由一个二元一次方程式的所有解点为坐标之点构成一条直线。试讨论分别由两个二元一次方程式

$$\begin{cases} a_1x + b_1y = c_1 \\ a_2x + b_2y = c_2 \end{cases}$$

的解点组成的两条直线的交截的各种可能性。

4.2 二阶和三阶行列式

要找出前一节所述的唯一解与系数集之间的函数关联, 我们便需要用抽象的方法去讨论, 即假设方程组的系数为一些任给者, 并探讨在什麼情况下该方程组会有唯一解, 由此设法找出加于这些普遍系数身上的限制和特性。由于在方程组的系数中含有待定者, 所以往常的代入法便无从入手, 因不知何时某系数会蜕化为 0, 那样便不能用其他变元来表达该变元了; 同理在消元法中我们也要避免用到除法, 所以在上一节所提到的「高斯消元法」者也不能在这里直接使用。

开宗明义, 让我们首先明确所要研讨的基本问题, 即:

线性方程组的基本问题:

「在什麼条件之下, 一组 n 个 n 元线性方程具有唯一解? 再者, 在满足唯一解条件的情形, 试求以方程组系数表达其唯一解组之公式。」

虽然在 $n = 1$ 时的解答是很简单, 但它提供了我们一个起点, 所以我们先把这时的结果写出如下:

$$(4.17) \quad ax = b \text{ 当 } a \neq 0 \text{ 时有唯一解, 其解之公式为 } x = \frac{b}{a}$$

接著让我们讨论 $n = 2$ 的情形。设方程组为

$$(4.18) \quad a_1x + b_1y = c_1$$

$$(4.19) \quad a_2x + b_2y = c_2$$

若要求出 x 的解，我们首先要设法消去 y 。这可以用下述方法做到：
 $b_2 \times (4.18) - b_1 \times (4.19)$ ；即有

$$\begin{aligned} (b_2a_1x + b_2b_1y) - (b_1a_2x + b_1b_2y) &= b_2c_1 - b_1c_2 \\ \Rightarrow (a_1b_2 - a_2b_1)x &= c_1b_2 - c_2b_1 \end{aligned}$$

所以，如果 $(a_1b_2 - a_2b_1) \neq 0$ ，则 x 的解是唯一的。若 $(a_1b_2 - a_2b_1) = 0$ ，则 $a_1 : b_1 = a_2 : b_2$ ，即 (4.18) 和 (4.19) 的左方只相差一个常数倍，所以是属于无解或是有无穷解的情形。至于 y 又怎样呢？我们可用同样方法，先消去 x 然後再看看 y 的系数，即由

$a_2 \times (4.18) - a_1 \times (4.19)$ 可得：

$$\begin{aligned} (a_2a_1x + a_2b_1y) - (a_1a_2x + a_1b_2y) &= a_2c_1 - a_1c_2 \\ \Rightarrow (a_2b_1 - a_1b_2)y &= a_2c_1 - a_1c_2 \end{aligned}$$

由此可见，同样的条件 $(a_1b_2 - a_2b_1) \neq 0$ 也可以保证 y 的解是唯一的，亦即方程组有唯一解的充要条件就是 $(a_1b_2 - a_2b_1) \neq 0$ 。而当这个条件成立时， x 和 y 的解是可以分别用下述公式表达之：

$$(4.20) \quad x = \frac{c_1b_2 - c_2b_1}{a_1b_2 - a_2b_1}, \quad y = \frac{a_1c_2 - a_2c_1}{a_1b_2 - a_2b_1}$$

这个就是我们想探求者在 $n = 2$ 时的情形。我们发现方程组的解的唯一性是取决于某一个量是否 不为 0，亦即

$$(4.21) \quad \begin{cases} a_1x + b_1y = c_1 \\ a_2x + b_2y = c_2 \end{cases} \quad \text{当 } (a_1b_2 - a_2b_1) \neq 0 \text{ 时有唯一解。}$$

上述结果和 $n = 1$ 的情形 (4.17) 很类似。此点是否偶然？还是能够有其普遍推广呢？此事当然有待继续探讨，逐步由 $n = 3, 4, \dots$ 归纳研究才有定论。为了方便以後的讨论，我们在此先引入二阶行列式的记号：

$$(4.22) \quad \begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix} = a_1b_2 - a_2b_1$$

上述代数表达式称之为该方程组系数的「二阶行列式」(determinant of order 2), 同学们暂且可以把它想成是 $a_1b_2 - a_2b_1$ 的另一种写法 (我们也定义相应于 $ax = b$ 的「一阶行列式」为 a 本身)。当这个二阶行列式不为 0 时, x, y 就会有唯一解, 而且其唯一解之公式就可以用上述符号如下表达之, 亦即:

$$(4.23) \quad x = \frac{c_1b_2 - c_2b_1}{a_1b_2 - a_2b_1} = \frac{\begin{vmatrix} c_1 & b_1 \\ c_2 & b_2 \end{vmatrix}}{\begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix}}, \quad y = \frac{a_1c_2 - a_2c_1}{a_1b_2 - a_2b_1} = \frac{\begin{vmatrix} a_1 & c_1 \\ a_2 & c_2 \end{vmatrix}}{\begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix}}$$

接著让我们考虑 $n = 3$ 的情形。设方程组为

$$(4.24) \quad a_1x + b_1y + c_1z = d_1$$

$$(4.25) \quad a_2x + b_2y + c_2z = d_2$$

$$(4.26) \quad a_3x + b_3y + c_3z = d_3$$

若 $c_2 \neq 0$, 则可如下先消去 z :

$$c_2 \times (4.24) - c_1 \times (4.25):$$

$$(4.27) \quad (a_1c_2 - a_2c_1)x + (b_1c_2 - b_2c_1)y = d_1c_2 - d_2c_1$$

$$c_3 \times (4.25) - c_2 \times (4.26):$$

$$(4.28) \quad (a_2c_3 - a_3c_2)x + (b_2c_3 - b_3c_2)y = d_2c_3 - d_3c_2$$

再如下消去 y :

$$(b_2c_3 - b_3c_2) \times (4.27) - (b_1c_2 - b_2c_1) \times (4.28):$$

$$(4.29) \quad \begin{aligned} & \{(a_1c_2 - a_2c_1)(b_2c_3 - b_3c_2) - (a_2c_3 - a_3c_2)(b_1c_2 - b_2c_1)\}x \\ & = \{(d_1c_2 - d_2c_1)(b_2c_3 - b_3c_2) - (d_2c_3 - d_3c_2)(b_1c_2 - b_2c_1)\} \end{aligned}$$

当我们展开上式中 x 的系数时, 发现所得的 8 项之中只有 2 项是不含有 c_2 者, 而且它们刚好互相抵消。同样情况亦出现在右方的常数项, 所以我们可以提出公因子 c_2 , 并且把 (4.29) 重写成:

$$(4.30) \quad \begin{aligned} & c_2(a_1b_2c_3 - a_1b_3c_2 + a_2b_3c_1 - a_2b_1c_3 + a_3b_1c_2 - a_3b_2c_1)x \\ & = c_2(d_1b_2c_3 - d_1b_3c_2 + d_2b_3c_1 - d_2b_1c_3 + d_3b_1c_2 - d_3b_2c_1) \end{aligned}$$

为何 c_2 会成为公因子呢？原因是当初我们用了两次 (4.25) 来消去 z ，而其中 c_2 是乘了两次。跟著我们可把不等于 0 的 c_2 约掉，并得出 x 有唯一解的充分条件为

$$(4.31) \quad a_1 b_2 c_3 - a_1 b_3 c_2 + a_2 b_3 c_1 - a_2 b_1 c_3 + a_3 b_1 c_2 - a_3 b_2 c_1 \neq 0$$

上述 (4.31)-式可重写成

$$(4.32) \quad a_1(b_2 c_3 - b_3 c_2) - a_2(b_1 c_3 - b_3 c_1) + a_3(b_1 c_2 - b_2 c_1) \neq 0$$

同学们请注意上式那些 $(b_2 c_3 - b_3 c_2)$ 其实是 y, z 某部分系数的二阶行列式。例如 $(b_2 c_3 - b_3 c_2)$ 是把 a_1 所在的行 ((4.24)-式) 和列 (x 的系数) 删掉而计算余下系数的行列式。因此，由「後见之明」可以看到了一个更好的做法，即我们只需直接考虑：

$$(b_2 c_3 - b_3 c_2) \times (4.24) - (b_1 c_3 - b_3 c_1) \times (4.25) + (b_1 c_2 - b_2 c_1) \times (4.26)$$

则 y, z 就可以一蹴而成地全被消去，并且可以直接得出

$$(4.33) \quad \left(a_1 \begin{vmatrix} b_2 & c_2 \\ b_3 & c_3 \end{vmatrix} - a_2 \begin{vmatrix} b_1 & c_1 \\ b_3 & c_3 \end{vmatrix} + a_3 \begin{vmatrix} b_1 & c_1 \\ b_2 & c_2 \end{vmatrix} \right) x \\ = d_1 \begin{vmatrix} b_2 & c_2 \\ b_3 & c_3 \end{vmatrix} - d_2 \begin{vmatrix} b_1 & c_1 \\ b_3 & c_3 \end{vmatrix} + d_3 \begin{vmatrix} b_1 & c_1 \\ b_2 & c_2 \end{vmatrix}$$

(在此不需假设 $c_2 \neq 0$ 。)

上述有效的消元法让我们坚信 (4.32) 是比 (4.31) 更能突出其要点，而用 (4.32) 来定义「三阶行列式」要比用 (4.31) 来得顺当自然，即：

$$(4.34) \quad \begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix} = a_1 \begin{vmatrix} b_2 & c_2 \\ b_3 & c_3 \end{vmatrix} - a_2 \begin{vmatrix} b_1 & c_1 \\ b_3 & c_3 \end{vmatrix} + a_3 \begin{vmatrix} b_1 & c_1 \\ b_2 & c_2 \end{vmatrix}$$

用上述三阶行列式符号来重写 (4.33)，即有：

$$(4.35) \quad \begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix} \cdot x = \begin{vmatrix} d_1 & b_1 & c_1 \\ d_2 & b_2 & c_2 \\ d_3 & b_3 & c_3 \end{vmatrix}$$

而相对于 y 和 z 时的情形, 我们同样可得 (证明留作习题) :

$$(4.36) \quad \begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix} \cdot y = \begin{vmatrix} a_1 & d_1 & c_1 \\ a_2 & d_2 & c_2 \\ a_3 & d_3 & c_3 \end{vmatrix}, \quad \begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix} \cdot z = \begin{vmatrix} a_1 & b_1 & d_1 \\ a_2 & b_2 & d_2 \\ a_3 & b_3 & d_3 \end{vmatrix}$$

上述 (4.35)-式和 (4.36)-式便是著名的 Cramer's rules 在三个三元一次方程的情形。总结上述对于三元的讨论, 即有:

$$(4.37) \quad \begin{cases} a_1x + b_1y + c_1z = d_1 \\ a_2x + b_2y + c_2z = d_2 \\ a_3x + b_3y + c_3z = d_3 \end{cases} \quad \text{当} \quad \begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix} \neq 0 \quad \text{时有唯一解。}$$

而且这个唯一解可以用 (4.35) 和 (4.36) 表达之。再者,

$$\text{若} \quad \begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix} = 0, \quad \text{则 (4.35)-式和 (4.36)-式就是}$$

$$(4.38) \quad 0 \cdot x = \begin{vmatrix} d_1 & b_1 & c_1 \\ d_2 & b_2 & c_2 \\ d_3 & b_3 & c_3 \end{vmatrix}, \quad 0 \cdot y = \begin{vmatrix} a_1 & d_1 & c_1 \\ a_2 & d_2 & c_2 \\ a_3 & d_3 & c_3 \end{vmatrix}, \quad 0 \cdot z = \begin{vmatrix} a_1 & b_1 & d_1 \\ a_2 & b_2 & d_2 \\ a_3 & b_3 & d_3 \end{vmatrix}$$

所以当右侧不为零时方程组 (4.24)–(4.26) 显然无解。再者, 当右侧也为零时, 则方程组 (4.24)–(4.26) 这三个之中, 其中之一其实是另外二

个的推论, 所以不是无解就是有无穷多解。总之, 在 $\begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix} = 0$ 时, 方程组不是无解就是有无穷多解。

【习题】:

(1) 试用类似 (4.33) 的消元法证明 (4.36)-式。

(2) 试验证下述等式:

$$\begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix} = - \begin{vmatrix} a_1 & c_1 & b_1 \\ a_2 & c_2 & b_2 \\ a_3 & c_3 & b_3 \end{vmatrix}, \quad \begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix} = - \begin{vmatrix} c_1 & b_1 & a_1 \\ c_2 & b_2 & a_2 \\ c_3 & b_3 & a_3 \end{vmatrix}$$

- (3) 试证明当三阶行列式 (4.34) 为零时，方程组 (4.24)–(4.26) 左方其中之一是另外二个的线性组合。
- (4) 试用行列式写出在平面上三点 (x_0, y_0) , (x_1, y_1) , (x_2, y_2) 为共线 (collinear) 的代数条件式。由此再写出过 (x_1, y_1) 和 (x_2, y_2) 两点的直线方程。
- (5) 试用行列式写下在平面上三条直线 $l_i: A_i x + B_i y + C_i = 0, i = 1, 2, 3$ 共交于一点 (concurrent) 的代数条件式。

4.3 四阶行列式

在讨论三元线性方程组时所发现的三阶行列式和「有效消元法」是否一般地推广呢？让我们再进一步去考虑四元线性方程组和四阶行列式的情形。在一般的中学课程中通常是不讨论四阶行列式的，但是行列式的重要性质要在这个阶段才明显地展现出来，其实应该要再进一步讨论四阶的情形。讨论四阶行列式的繁处在于其定义公式展开後共有 24 项，所以直接使用展开公式来讨论是颇为繁琐的。因此我们将改弦更张，放弃直接使用公式而采用归纳法探究其性质的研讨方法。唯有这样，才能顺理成章地建立起 n -阶行列式的基础理论。

总之，让我们接著考虑四个四元线性方程：

$$(4.39) \quad a_1 x + b_1 y + c_1 z + d_1 w = e_1$$

$$(4.40) \quad a_2 x + b_2 y + c_2 z + d_2 w = e_2$$

$$(4.41) \quad a_3 x + b_3 y + c_3 z + d_3 w = e_3$$

$$(4.42) \quad a_4 x + b_4 y + c_4 z + d_4 w = e_4$$

我们当然可以直接运用消元法来先消去 w ，再消去 z ，然後消去 y ；这时， x 的系数便是所求的量。但是在 $n = 3$ 的情形我们已发现了一个更为有效的方法可以一蹴而成地消去其他变元，所以不妨依样画葫芦地

来试一试其法是否依然可行，即：

$$(4.43) \quad \begin{aligned} & \begin{vmatrix} b_2 & c_2 & d_2 \\ b_3 & c_3 & d_3 \\ b_4 & c_4 & d_4 \end{vmatrix} \times (4.39) - \begin{vmatrix} b_1 & c_1 & d_1 \\ b_3 & c_3 & d_3 \\ b_4 & c_4 & d_4 \end{vmatrix} \times (4.40) \\ & + \begin{vmatrix} b_1 & c_1 & d_1 \\ b_2 & c_2 & d_2 \\ b_4 & c_4 & d_4 \end{vmatrix} \times (4.41) - \begin{vmatrix} b_1 & c_1 & d_1 \\ b_2 & c_2 & d_2 \\ b_3 & c_3 & d_3 \end{vmatrix} \times (4.42) \end{aligned}$$

如上述方法是可行的，则 x 的系数便是我们所求的「四阶行列式」，亦即

$$(4.44) \quad \begin{vmatrix} a_1 & b_1 & c_1 & d_1 \\ a_2 & b_2 & c_2 & d_2 \\ a_3 & b_3 & c_3 & d_3 \\ a_4 & b_4 & c_4 & d_4 \end{vmatrix} = a_1 \cdot \begin{vmatrix} b_2 & c_2 & d_2 \\ b_3 & c_3 & d_3 \\ b_4 & c_4 & d_4 \end{vmatrix} - a_2 \cdot \begin{vmatrix} b_1 & c_1 & d_1 \\ b_3 & c_3 & d_3 \\ b_4 & c_4 & d_4 \end{vmatrix} \\ + a_3 \cdot \begin{vmatrix} b_1 & c_1 & d_1 \\ b_2 & c_2 & d_2 \\ b_4 & c_4 & d_4 \end{vmatrix} - a_4 \cdot \begin{vmatrix} b_1 & c_1 & d_1 \\ b_2 & c_2 & d_2 \\ b_3 & c_3 & d_3 \end{vmatrix}$$

上述依样试用的消元法所得的 y, z 和 w 的系数分别如下，所以此法是否依然可以一蹴而成地消去 y, z 和 w ，则当然就取决于它们是否都自然而然地恒为 0！亦即

$$(4.45) \quad \begin{cases} b_1 \begin{vmatrix} b_2 & c_2 & d_2 \\ b_3 & c_3 & d_3 \\ b_4 & c_4 & d_4 \end{vmatrix} - b_2 \begin{vmatrix} b_1 & c_1 & d_1 \\ b_3 & c_3 & d_3 \\ b_4 & c_4 & d_4 \end{vmatrix} + b_3 \begin{vmatrix} b_1 & c_1 & d_1 \\ b_2 & c_2 & d_2 \\ b_4 & c_4 & d_4 \end{vmatrix} - b_4 \begin{vmatrix} b_1 & c_1 & d_1 \\ b_2 & c_2 & d_2 \\ b_3 & c_3 & d_3 \end{vmatrix} \stackrel{?}{=} 0 \\ c_1 \begin{vmatrix} b_2 & c_2 & d_2 \\ b_3 & c_3 & d_3 \\ b_4 & c_4 & d_4 \end{vmatrix} - c_2 \begin{vmatrix} b_1 & c_1 & d_1 \\ b_3 & c_3 & d_3 \\ b_4 & c_4 & d_4 \end{vmatrix} + c_3 \begin{vmatrix} b_1 & c_1 & d_1 \\ b_2 & c_2 & d_2 \\ b_4 & c_4 & d_4 \end{vmatrix} - c_4 \begin{vmatrix} b_1 & c_1 & d_1 \\ b_2 & c_2 & d_2 \\ b_3 & c_3 & d_3 \end{vmatrix} \stackrel{?}{=} 0 \\ d_1 \begin{vmatrix} b_2 & c_2 & d_2 \\ b_3 & c_3 & d_3 \\ b_4 & c_4 & d_4 \end{vmatrix} - d_2 \begin{vmatrix} b_1 & c_1 & d_1 \\ b_3 & c_3 & d_3 \\ b_4 & c_4 & d_4 \end{vmatrix} + d_3 \begin{vmatrix} b_1 & c_1 & d_1 \\ b_2 & c_2 & d_2 \\ b_4 & c_4 & d_4 \end{vmatrix} - d_4 \begin{vmatrix} b_1 & c_1 & d_1 \\ b_2 & c_2 & d_2 \\ b_3 & c_3 & d_3 \end{vmatrix} \stackrel{?}{=} 0 \end{cases}$$

现在让我们尝试验证 (4.45) 的第一式。用三阶行列式的定义，把左方展开成：

$$(4.46) \quad \begin{aligned} & b_1(b_2c_3d_4 - b_2c_4d_3 + b_3c_4d_2 - b_3c_2d_4 + b_4c_2d_3 - b_4c_3d_2) \\ & - b_2(b_1c_3d_4 - b_1c_4d_3 + b_3c_4d_1 - b_3c_1d_4 + b_4c_1d_3 - b_4c_3d_1) \\ & + b_3(b_1c_2d_4 - b_1c_4d_2 + b_2c_4d_1 - b_2c_1d_4 + b_4c_1d_2 - b_4c_2d_1) \\ & - b_4(b_1c_2d_3 - b_1c_3d_2 + b_2c_3d_1 - b_2c_1d_3 + b_3c_1d_2 - b_3c_2d_1) \end{aligned}$$

耐心地展开上式，便会发觉项与项之间真的能逐对正负相消而恒等于 0。虽然我们可以同样地验证 (4.45) 的其余两式，但预见过程也是同样麻烦，再者到了讨论五、六阶行列式时这种做法是愈来愈繁而不胜其繁者也。因此我们需要寻求的出路是对行列式的性质作深一层了解；亦即不再停留于其展开公式，而是对它的本质作深入探讨。

像 (4.45) 这类公式能够普遍地成立，我们相信此事决非纯属偶然，肯定应有其本质性的原因！因此我们应当去把其中隐藏的原因发掘出来。首先我们看到 (4.45) 的左方其实是比较特别的四阶行列式，用定义式 (4.44) 我们可以将它们分别改写成下述模样：

$$(4.47) \quad \begin{vmatrix} b_1 & b_1 & c_1 & d_1 \\ b_2 & b_2 & c_2 & d_2 \\ b_3 & b_3 & c_3 & d_3 \\ b_4 & b_4 & c_4 & d_4 \end{vmatrix}, \quad \begin{vmatrix} c_1 & b_1 & c_1 & d_1 \\ c_2 & b_2 & c_2 & d_2 \\ c_3 & b_3 & c_3 & d_3 \\ c_4 & b_4 & c_4 & d_4 \end{vmatrix}, \quad \begin{vmatrix} d_1 & b_1 & c_1 & d_1 \\ d_2 & b_2 & c_2 & d_2 \\ d_3 & b_3 & c_3 & d_3 \\ d_4 & b_4 & c_4 & d_4 \end{vmatrix}$$

可见它们全是有两列 (columns) 完全一样的四阶行列式。所以，若我们能够证明任何有两列全同的四阶行列式的值恒为 0（这种特性称之为「交错性」(alternating property)），则 (4.45) 就会普遍成立。可惜这个方法涉及寻找四阶行列式的确实数值，所以仍受著繁复公式的牵制；我们不妨避重就轻，改为研讨下述的「特性」：

「把一个行列式中的两列互换时，它的新旧值是否只相差一个负号？」

若上述特性（称之为「斜对称性」(skew-symmetric property)）能够普遍成立，则只需把两个全同之列互换，例如

$$(4.48) \quad \begin{vmatrix} b_1 & b_1 & c_1 & d_1 \\ b_2 & b_2 & c_2 & d_2 \\ b_3 & b_3 & c_3 & d_3 \\ b_4 & b_4 & c_4 & d_4 \end{vmatrix} = - \begin{vmatrix} b_1 & b_1 & c_1 & d_1 \\ b_2 & b_2 & c_2 & d_2 \\ b_3 & b_3 & c_3 & d_3 \\ b_4 & b_4 & c_4 & d_4 \end{vmatrix}$$

即得这个四阶行列式的值必为 0。同理，其余两值也为 0，于是 (4.45)-式普遍成立，亦即「有效消元法」是可行的。现在让我们著手验证上述斜对称性。

[注]：交错性与斜对称性在一般抽象环、域的行列式理论中是有少许差别的，但在这里我们可以把它们当作同一样的概念。

对于二阶行列式，由其定义式 (4.22) 即知：

$$\begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix} = a_1 b_2 - a_2 b_1 = -(b_1 a_2 - b_2 a_1) = - \begin{vmatrix} b_1 & a_1 \\ b_2 & a_2 \end{vmatrix}$$

所以斜对称性在 $n = 2$ 时成立。

对于三阶行列式，有 3 种互换方法要考虑，如互换第一列和第二列：

$$\begin{aligned} \begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix} &= a_1 b_2 c_3 - a_1 b_3 c_2 + a_2 b_3 c_1 - a_2 b_1 c_3 + a_3 b_1 c_2 - a_3 b_2 c_1 \\ &= -(b_1 a_2 c_3 - b_1 a_3 c_2 + b_2 a_3 c_1 - b_2 a_1 c_3 + b_3 a_1 c_2 - b_3 a_2 c_1) \\ &= - \begin{vmatrix} b_1 & a_1 & c_1 \\ b_2 & a_2 & c_2 \\ b_3 & a_3 & c_3 \end{vmatrix} \end{aligned}$$

同理可证余下的两种互换方法，即互换第一列和第三列，及互换第二列和第三列。斜对称性在 $n = 3$ 时亦成立。

对于四阶行列式，上述的直接验证方法便很费时了。在此，我们将改用性质来证明。首先回顾四阶行列式的定义式：

$$(4.44') \quad D = \begin{vmatrix} a_1 & b_1 & c_1 & d_1 \\ a_2 & b_2 & c_2 & d_2 \\ a_3 & b_3 & c_3 & d_3 \\ a_4 & b_4 & c_4 & d_4 \end{vmatrix} = a_1 D_{1,1} - a_2 D_{2,1} + a_3 D_{3,1} - a_4 D_{4,1}$$

在这里我们引入了简约符号 $D_{i,j}$ ，它代表把原来的四阶行列式内的第 i 行 (i -th row) 和第 j 列 (j -th column) 删去後，所得的三阶行列式。例如

$D_{2,1}$ 就是：

$$D_{2,1} = \begin{vmatrix} a_1 & b_1 & c_1 & d_1 \\ \hline a_2 & b_2 & c_2 & d_2 \\ a_3 & b_3 & c_3 & d_3 \\ a_4 & b_4 & c_4 & d_4 \end{vmatrix} = \begin{vmatrix} b_1 & c_1 & d_1 \\ b_3 & c_3 & d_3 \\ b_4 & c_4 & d_4 \end{vmatrix}$$

由上述重写的定义式 (4.44') 可以看到，若在原来的四阶行列式中所互换的两列并不涉及第一列，则以上每一个 $D_{i,1}$ 也有相应的换列发生；由已知的 $n=3$ 的情形得知每个 $D_{i,1}$ 也转作 $-D_{i,1}$ ，因此原来的四阶行列式 D 也转为 $-D$ 。余下只需验证涉及第一列的换列。

稍加分析後，我们只需集中讨论第一列和第二列的互换。例如，第一列和第三列的互换可用下述一串换列来得出：

0. 起始： A—B—C—D
1. 互换第二列和第三列： A—C—B—D
2. 互换第一列和第二列： C—A—B—D
3. 互换第二列和第三列： C—B—A—D

由于第 1 步和第 3 步都引入一个负号，所以其作用互相抵消；若可验证第 2 步也同样地引入一个负号，则涉及第一列的换列之验证便可完成。

在行列式的定义式 (4.44') 中，第一列是有别于其余各列的，因为用来写定义式的系数是取于第一列者，而其他列的系数则全都统括在那些低一阶的子行列式内。所以，当第一列和第二列互换後，骤看起来定义式便会变得面目全非、难以处理。因此，我们先要做一些准备功夫，把那些三阶行列式再进一步分解下去，将原来四阶行列式的第二列系数也提出来，这样第二列和第一列系数便可以处于同等地位。例如：

$$\begin{aligned} D_{1,1} &= \begin{vmatrix} b_2 & c_2 & d_2 \\ b_3 & c_3 & d_3 \\ b_4 & c_4 & d_4 \end{vmatrix} = b_2 \begin{vmatrix} c_3 & d_3 \\ c_4 & d_4 \end{vmatrix} - b_3 \begin{vmatrix} c_2 & d_2 \\ c_4 & d_4 \end{vmatrix} + b_4 \begin{vmatrix} c_2 & d_2 \\ c_3 & d_3 \end{vmatrix} \\ &= b_2 D_{12,12} - b_3 D_{13,12} + b_4 D_{14,12} \end{aligned}$$

其中 $D_{14,12}$ 代表把原来的四阶行列式内的第 1 行和第 4 行、以及第一列和第二列删去後，所得的二阶子行列式。其他如 $D_{12,12}$ 和 $D_{13,12}$ 也类

似地定义之。当我们把在 (4.44') 中的四个三阶子行列式全都如前述分解後，便得出：

$$\begin{aligned}
 & \text{原来的四阶行列式, } D \\
 &= a_1(b_2D_{12,12} - b_3D_{13,12} + b_4D_{14,12}) - a_2(b_1D_{12,12} - b_3D_{23,12} + b_4D_{24,12}) \\
 &\quad + a_3(b_1D_{13,12} - b_2D_{23,12} + b_4D_{34,12}) - a_4(b_1D_{14,12} - b_2D_{24,12} + b_3D_{34,12}) \\
 &= (a_1b_2 - a_2b_1)D_{12,12} + (a_3b_1 - a_1b_3)D_{13,12} + (a_1b_4 - a_4b_1)D_{14,12} \\
 &\quad + (a_2b_3 - a_3b_2)D_{23,12} + (a_4b_2 - a_2b_4)D_{24,12} + (a_3b_4 - a_4b_3)D_{34,12}
 \end{aligned}$$

由上式我们立即可看到若把 D 的第一列和第二列互换（即把 a_i 和 b_i 互换），则每个 $D_{ij,12}$ 保持不变，但其系数却都引进了负号，所以整体上 D 便转为 $-D$ 。斜对称性得证，所以「有效消元法」在 $n=4$ 时依然可行。

跟著我们再用「有效消元法」来从 (4.39)–(4.42) 一次地消去 x, z, w :

$$\begin{aligned}
 (4.49) \quad & \begin{vmatrix} a_2 & c_2 & d_2 \\ a_3 & c_3 & d_3 \\ a_4 & c_4 & d_4 \end{vmatrix} \times (4.39) - \begin{vmatrix} a_1 & c_1 & d_1 \\ a_3 & c_3 & d_3 \\ a_4 & c_4 & d_4 \end{vmatrix} \times (4.40) \\
 & + \begin{vmatrix} a_1 & c_1 & d_1 \\ a_2 & c_2 & d_2 \\ a_4 & c_4 & d_4 \end{vmatrix} \times (4.41) - \begin{vmatrix} a_1 & c_1 & d_1 \\ a_2 & c_2 & d_2 \\ a_3 & c_3 & d_3 \end{vmatrix} \times (4.42)
 \end{aligned}$$

此时， x 的系数就是：

$$a_1 \cdot \begin{vmatrix} a_2 & c_2 & d_2 \\ a_3 & c_3 & d_3 \\ a_4 & c_4 & d_4 \end{vmatrix} - a_2 \cdot \begin{vmatrix} a_1 & c_1 & d_1 \\ a_3 & c_3 & d_3 \\ a_4 & c_4 & d_4 \end{vmatrix} + a_3 \cdot \begin{vmatrix} a_1 & c_1 & d_1 \\ a_2 & c_2 & d_2 \\ a_4 & c_4 & d_4 \end{vmatrix} - a_4 \cdot \begin{vmatrix} a_1 & c_1 & d_1 \\ a_2 & c_2 & d_2 \\ a_3 & c_3 & d_3 \end{vmatrix}$$

但由四阶行列式定义，上述表达式其实是：

$$\begin{vmatrix} a_1 & a_1 & c_1 & d_1 \\ a_2 & a_2 & c_2 & d_2 \\ a_3 & a_3 & c_3 & d_3 \\ a_4 & a_4 & c_4 & d_4 \end{vmatrix}$$

由斜对称性知其值必为 0，所以 x 的确被消去。同理可证 z, w 也被消去，而余下 y 的系数则是：

$$\begin{aligned} & b_1 \cdot \begin{vmatrix} a_2 & c_2 & d_2 \\ a_3 & c_3 & d_3 \\ a_4 & c_4 & d_4 \end{vmatrix} - b_2 \cdot \begin{vmatrix} a_1 & c_1 & d_1 \\ a_3 & c_3 & d_3 \\ a_4 & c_4 & d_4 \end{vmatrix} + b_3 \cdot \begin{vmatrix} a_1 & c_1 & d_1 \\ a_2 & c_2 & d_2 \\ a_4 & c_4 & d_4 \end{vmatrix} - b_4 \cdot \begin{vmatrix} a_1 & c_1 & d_1 \\ a_2 & c_2 & d_2 \\ a_3 & c_3 & d_3 \end{vmatrix} \\ &= \begin{vmatrix} b_1 & a_1 & c_1 & d_1 \\ b_2 & a_2 & c_2 & d_2 \\ b_3 & a_3 & c_3 & d_3 \\ b_4 & a_4 & c_4 & d_4 \end{vmatrix} = - \begin{vmatrix} a_1 & b_1 & c_1 & d_1 \\ a_2 & b_2 & c_2 & d_2 \\ a_3 & b_3 & c_3 & d_3 \\ a_4 & b_4 & c_4 & d_4 \end{vmatrix} \end{aligned}$$

所以这个四阶行列式的值不为 0 也是 y 有唯一解的充分条件。类似地可得出 z 或 w 有唯一解的同一充分条件。由此即得下述结论：

$$\begin{cases} a_1x + b_1y + c_1z + d_1w = e_1 \\ a_2x + b_2y + c_2z + d_2w = e_2 \\ a_3x + b_3y + c_3z + d_3w = e_3 \\ a_4x + b_4y + c_4z + d_4w = e_4 \end{cases} \quad \text{当} \quad \begin{vmatrix} a_1 & b_1 & c_1 & d_1 \\ a_2 & b_2 & c_2 & d_2 \\ a_3 & b_3 & c_3 & d_3 \\ a_4 & b_4 & c_4 & d_4 \end{vmatrix} \neq 0 \quad \text{时有唯一解。}$$

而证明中的要点就是四阶行列式满足斜对称性：「若互换两列，则其值变号」。

再者，易见在上述四阶行列式非零的情形，(4.39)–(4.42) 的唯一解的公式如下：

$$(4.50) \quad x = \frac{1}{D} \begin{vmatrix} e_1 & b_1 & c_1 & d_1 \\ e_2 & b_2 & c_2 & d_2 \\ e_3 & b_3 & c_3 & d_3 \\ e_4 & b_4 & c_4 & d_4 \end{vmatrix}, \quad y = \frac{1}{D} \begin{vmatrix} a_1 & e_1 & c_1 & d_1 \\ a_2 & e_2 & c_2 & d_2 \\ a_3 & e_3 & c_3 & d_3 \\ a_4 & e_4 & c_4 & d_4 \end{vmatrix}, \quad \text{等等}$$

在 $D = 0$ 的情形，方程组不是无解就是有无穷多解的讨论则类似于 $n = 3$ 的情形，在此留作习题。

【习题】：

- (1) 证明上式 (4.50)，亦即 Cramer's rules 在 $n = 4$ 的情形。
- (2) 试证明当四阶行列式 (4.44) 为零时，方程组 (4.39)–(4.42) 左方其中之一是另外三个的推论。

- (3) 试用行列式写出在平面上四点 $(x_0, y_0), (x_1, y_1), (x_2, y_2), (x_3, y_3)$ 为共圆 (cocircular) 的代数条件式。
- (4) 试用行列式写出在空间中四点 $(x_0, y_0, z_0), (x_1, y_1, z_1), (x_2, y_2, z_2), (x_3, y_3, z_3)$ 为共面 (coplanar) 的代数条件式。由此再写出过 $(x_1, y_1, z_1), (x_2, y_2, z_2)$ 和 (x_3, y_3, z_3) 三点的平面方程。
- (5) 试用行列式写下在空间中四个平面 $\Pi_i : A_i x + B_i y + C_i z + D_i = 0, i = 1, 2, 3, 4$ 共交于一点的代数条件式。

第五章

行列式的基本性质与应用范例

在前一章中我们采用归纳法逐步发现三阶及四阶行列式的归纳定义式及验证其斜对称性，至此，我们当然也可以顺理成章，归纳地由 $(n-1)$ -阶行列式类似地定义 n -阶行列式，然后再归纳地论证如此归纳定义所得者也具有列、列互换其值变号的斜对称性。

【历史的注记】：虽然直接给出 n -阶行列式的展开式来作定义也是可行的，但是这样便需要对排列群 (permutation group) (或称作对称群, symmetric group) 有一定的认识才可以妥善地做好。因为五阶行列式展开後共有 120 项，六阶行列式展开後共有 720 项，而一般地 n -阶行列式展开後共有 $n!$ 项，所以对于初学者或是对排列群认识不深的人来说，直接使用展开式来作讨论是难以胜任自如的。想当年 Cramer 和 Vandermonde 乃是非常优秀的代数学家的，而且他们对于排列群又有著深刻的认识，所以他们才直截了当地处理那个巨大的展开公式；但对于一般人来说，这样讨论行列式肯定是吃力难懂的。本章所采用的归纳法，返璞归真地来讨论行列式，既顺理成章地表现其本质，又简朴、易懂、好用。

5.1 n -阶行列式的归纳定义

总结前一章的三阶和四阶行列式的讨论，我们得见下述两点：

- (i) 行列式可以用第一列的 n 个系数 (即 $a_{i1}, 1 \leq i \leq n$)，以交错的形式分别乘上其相应的低一阶的子行列式 (即 $D_{i,1}, 1 \leq i \leq n$) 後

，以其总和来加以归纳地定义；（这正是「有效消元法」的具体化）

(ii) 行列式具有斜对称性，即互换两列时，其值变号。

所以，现在让我们来正式地定义 n -阶行列式。考虑下述 n 个 n 元一次的线性方程组（ $n \geq 3$ ）：

$$(5.1) \quad \begin{cases} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n = b_2 \\ \qquad \qquad \qquad \qquad \qquad \qquad \qquad \qquad \qquad \vdots \qquad \qquad \qquad \qquad \qquad \qquad \qquad \qquad \qquad \vdots \\ a_{n1}x_1 + a_{n2}x_2 + \dots + a_{nn}x_n = b_n \end{cases}$$

由它的系数集所得的 n -阶行列式 D 则定义为

$$(5.2) \quad D = \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix} = a_{11}D_{1,1} - a_{21}D_{2,1} + \dots + (-1)^{n+1}a_{n1}D_{n,1}$$

$$= \sum_{i=1}^n (-1)^{i+1} a_{i1} D_{i,1}$$

其中每一个 $D_{i,1}$, $1 \leq i \leq n$, 都是低一阶的子行列式，它们分别是把原来 n -阶行列式内的第 i 行和第 1 列删去，所得的 $(n-1)$ -阶子行列式。例如， $D_{2,1}$ 就是如下把 D 的第 2 行和第 1 列删去，在式子最右方所示的 $(n-1)$ -阶行列式，即：

$$(5.3) \quad D_{2,1} = \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \hline a_{21} & a_{22} & \cdots & a_{2n} \\ \hline a_{31} & a_{32} & \cdots & a_{3n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix} = \begin{vmatrix} a_{12} & \cdots & a_{1n} \\ a_{32} & \cdots & a_{3n} \\ \vdots & \ddots & \vdots \\ a_{n2} & \cdots & a_{nn} \end{vmatrix}$$

注意上述是「归纳地」定义 n -阶行列式，亦即借助业已定义的 $(n-1)$ -阶行列式来定义 n -阶行列式。其实在绪论中已和同学们提及，一个完整的归纳法应该可以分为下述三个步骤：

- 一、归纳地发现具有某种有用特性的事物 (inductive discovery)；
- 二、归纳地定义该事物 (inductive definition)；
- 三、归纳地证明上述归纳定义者的确具有所应有之特性 (inductive proof)。

而其中的「归纳发现」则是最有意思也是最为重要的起步。若我们再看一看上一章对于四阶行列式的讨论，就会发现行列式所需要的「归纳定义」，以及其所需的「归纳证明」其实已经在「归纳发现」的过程中呼之欲出地展现著。

现在让我们来归纳地证明 (5.2) 所定义的 n -阶行列式就是我们所需求者，要点在于验证其斜对称性。归纳假设所有不高于 $(n-1)$ -阶的行列式 ($n \geq 3$) 业已具有斜对称性，我们把 n -阶行列式 D 的换列分成下述两种情形来考虑：

一、不涉及第一列的换列。

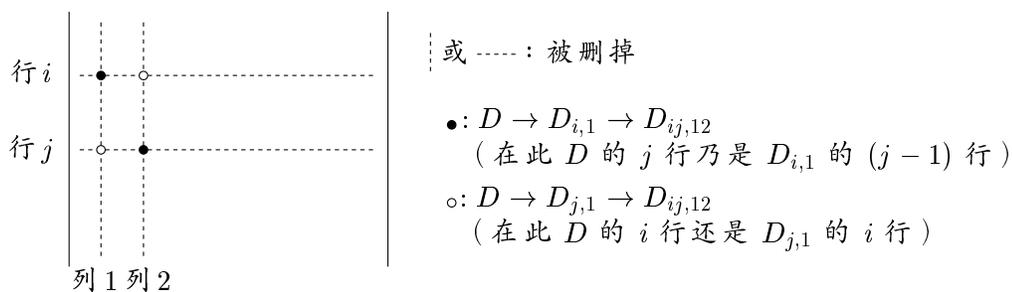
从定义式 (5.2) 可以直接看到，其中每一个 $D_{i,1}$ 也都有相应的换列，由归纳假设知道这种换列使得每一个 $D_{i,1}$ 变为 $-D_{i,1}$ ，而 a_{i1} 则全部不变，易见这种换列使得 D 变为 $-D$ 。

二、第一列和第 k 列的互换。

这个情况和在讨论四阶行列式的斜对称性证明时一样，我们只需集中讨论第一列和第二列的互换。而验证的方法则是把那些 $D_{i,1}$ 分别再用 $(n-1)$ -阶行列式的归纳定义式展开成一些 $(n-2)$ -阶子行列式 (即 $D_{ij,12}$) 的倍数和，然後研讨最後所得的展开式中每一个 $D_{ij,12}$ 的系数在上述换列後的变化。 $D_{ij,12}$ 的定义是把 D 的第 i 行和第 j 行、以及第 1 列和第 2 列删去，所得的 $(n-2)$ -阶子行列式。易见可以有下述两种途径得出含有 $D_{ij,12}$ 者：

- (i) 从 $D_{i,1}$ 中删掉相对于 D 的第 j 行和第 2 列；
- (ii) 从 $D_{j,1}$ 中删掉相对于 D 的第 i 行和第 2 列。

可参考下述图示：



[图 5-1]

我们不妨假设 $i < j$ 。在方法 (i) 中得出的 $D_{ij,12}$ ，其系数（在 D 的展开）为 $(-1)^{i+1}a_{i1} \cdot (-1)^j a_{j2}$ ；而在方法 (ii) 中得出的 $D_{ij,12}$ ，其系数为 $(-1)^{j+1}a_{j1} \cdot (-1)^{i+1}a_{i2}$ 。于是在 D 的展开中， $D_{ij,12}$ 是以下述形式出现：

$$(5.4) \quad D = \dots + (-1)^{i+j+1}(a_{i1}a_{j2} - a_{j1}a_{i2})D_{ij,12} + \dots$$

由此可见，当 D 的第一列和第二列互换后（即 $a_{i1} \leftrightarrow a_{i2}, a_{j1} \leftrightarrow a_{j2}$ ），其展开中 $D_{ij,12}$ 的系数的改变为：

$$(5.5) \quad a_{i1}a_{j2} - a_{j1}a_{i2} \rightarrow a_{i2}a_{j1} - a_{j2}a_{i1} = -(a_{i1}a_{j2} - a_{j1}a_{i2})$$

总括来说，当 D 的第一列和第二列互换后，在它的上述表成 $(n-2)$ -阶子行列式 $D_{ij,12}$ 的展开中，每个 $D_{ij,12}$ 保持不变，但其系数则变号，所以整体上对 D 的影响乃是 $D \rightarrow -D$ 。斜对称性得证。□

为了便于讨论 n -阶行列式的其他性质，在此引入列向量记号。令

$$(5.6) \quad \mathbf{a}_1 = \begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{n1} \end{bmatrix}, \mathbf{a}_2 = \begin{bmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{n2} \end{bmatrix}, \dots, \mathbf{a}_n = \begin{bmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{nn} \end{bmatrix}, \mathbf{b} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix}$$

而列向量的加和倍积运算是直接用其分量加以定义的，即

$$(5.7) \quad \mathbf{a}_1 + \mathbf{a}_2 = \begin{bmatrix} a_{11} + a_{12} \\ a_{21} + a_{22} \\ \vdots \\ a_{n1} + a_{n2} \end{bmatrix}, \quad \lambda \mathbf{a}_1 = \begin{bmatrix} \lambda a_{11} \\ \lambda a_{21} \\ \vdots \\ \lambda a_{n1} \end{bmatrix}$$

在这种符号体系下，我们把 D 重新写成：

$$(5.8) \quad D = \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix} = \det(\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n)$$

于是，当我们用下述「有效消元法」尝试从 (5.1) 中消去 x_2, \dots, x_n 时：

$$(5.9) \quad D_{1,1} \times (5.1.1) - D_{2,1} \times (5.1.2) + \dots + (-1)^{n+1} D_{n,1} \times (5.1.n)$$

则得出 (5.9) 的左方为：

$$\begin{aligned} & (a_{11}D_{1,1} - a_{21}D_{2,1} + \dots + (-1)^{n+1}a_{n1}D_{n,1})x_1 \\ & + (a_{12}D_{1,1} - a_{22}D_{2,1} + \dots + (-1)^{n+1}a_{n2}D_{n,1})x_2 + \dots + \\ & + (a_{1n}D_{1,1} - a_{2n}D_{2,1} + \dots + (-1)^{n+1}a_{nn}D_{n,1})x_n \\ = & \det(\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n)x_1 + \det(\mathbf{a}_2, \mathbf{a}_2, \dots, \mathbf{a}_n)x_2 + \dots + \det(\mathbf{a}_n, \mathbf{a}_2, \dots, \mathbf{a}_n)x_n \\ = & \det(\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n)x_1 + 0x_2 + \dots + 0x_n \quad (\text{用 } \det \text{ 的斜对称性}) \\ = & \det(\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n)x_1 \end{aligned}$$

而 (5.9) 的右方则为：

$$b_1D_{1,1} - b_2D_{2,1} + \dots + (-1)^{n+1}b_nD_{n,1} = \det(\mathbf{b}, \mathbf{a}_2, \dots, \mathbf{a}_n)$$

所以，我们得出在一般情况下相对于 x_1 的 Cramer's rule：

$$(5.10) \quad \det(\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n)x_1 = \det(\mathbf{b}, \mathbf{a}_2, \dots, \mathbf{a}_n)$$

而余下的 Cramer's rules 亦可以类似地导出。至此，我们可以作下述结论：

$$(5.11) \quad \begin{cases} a_{11}x_1 + \dots + a_{1n}x_n = b_1 \\ a_{21}x_1 + \dots + a_{2n}x_n = b_2 \\ \vdots \\ a_{n1}x_1 + \dots + a_{nn}x_n = b_n \end{cases} \quad \text{当 } D \neq 0 \text{ 时有唯一解。}$$

而当 $D = \det(\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n) \neq 0$ 时, 其唯一解是可以由下述 Cramer's rules 表示之, 即

$$(5.12) \quad x_i = \frac{1}{D} \det(\dots, \mathbf{b}_i, \dots), \quad 1 \leq i \leq n$$

其中 $(\dots, \mathbf{b}_i, \dots)$ 表示把 $(\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n)$ 的第 i 列换成 \mathbf{b} 。

在此我们还得到一个常用、好用的推论, 即

【推论】: n 个 n 元齐线性方程组

$$(5.13) \quad \begin{cases} a_{11}x_1 + \dots + a_{1n}x_n = 0 \\ a_{21}x_1 + \dots + a_{2n}x_n = 0 \\ \vdots \\ a_{n1}x_1 + \dots + a_{nn}x_n = 0 \end{cases}$$

具有非零解的条件是

$$(5.14) \quad \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix} = 0$$

至于当 $\det(\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n) = 0$ 时 (5.1) 不是无解就是有无穷多解的证明和以往完全一样, 在此留作习题。我们在前一章第二节开始时所研讨的基本问题, 至此业已完全解决。

5.2 斜对称多线性函数与行列式的界定定理

一般来说, 当我们对于某一问题或某一事物由表及里作深入探讨时, 要点在于精益求精地掌握其本质和精要所在, 所以在进一步研讨行列式时, 我们的中心课题就是要对于行列式的各种各样性质作系统的整理。行列式的两个基本性质在于其斜对称性和多线性。

由行列式的定义式 (5.2) 易见它对于第一列是线性 (linear) 的, 即:

$$(5.15) \quad \begin{aligned} & \det(\lambda \mathbf{a}_1 + \mu \mathbf{a}_1^*, \mathbf{a}_2, \dots, \mathbf{a}_n) \\ &= \lambda \det(\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n) + \mu \det(\mathbf{a}_1^*, \mathbf{a}_2, \dots, \mathbf{a}_n) \end{aligned}$$

运用其斜对称性，就可以把线性条件推广至其他各列，例如第二列：

$$\begin{aligned} \det(\mathbf{a}_1, \lambda \mathbf{a}_2 + \mu \mathbf{a}_2^*, \mathbf{a}_3, \dots, \mathbf{a}_n) \\ &= -\det(\lambda \mathbf{a}_2 + \mu \mathbf{a}_2^*, \mathbf{a}_1, \mathbf{a}_3, \dots, \mathbf{a}_n) \\ &= -\lambda \det(\mathbf{a}_2, \mathbf{a}_1, \dots, \mathbf{a}_n) - \mu \det(\mathbf{a}_2^*, \mathbf{a}_1, \dots, \mathbf{a}_n) \\ &= \lambda \det(\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n) + \mu \det(\mathbf{a}_1, \mathbf{a}_2^*, \dots, \mathbf{a}_n) \end{aligned}$$

上述对于每列都是线性的特性称之为「多线性」(multilinear)，一般有以下定义：

【定义】：设 $f(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n)$ 为一个以向量为变元的多元函数。若对于每一个 j ($1 \leq j \leq n$)，当 $\mathbf{x}_1, \dots, \mathbf{x}_{j-1}, \mathbf{x}_{j+1}, \dots, \mathbf{x}_n$ 全部固定不动时， f 恒为 \mathbf{x}_j 的线性函数，则称 f 为多线性函数。

所以行列式就是其 n 个列向量的斜对称多线性函数。我们再引入下述符号：

【定义】：令下述 n 个 n -维列向量为标准基底列向量：

$$(5.16) \quad \mathbf{e}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \quad \mathbf{e}_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \quad \dots, \quad \mathbf{e}_j = \begin{bmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{bmatrix} \leftarrow j \text{ 处}, \quad \dots, \quad \mathbf{e}_n = \begin{bmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix}$$

[注]：在不同大小的 n 之下，我们仍沿用 $\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n$ 等同一符号，只要不会产生混乱便可。

现在我们证明下述定理，它说明了斜对称性和多线性两者基本上已经界定了行列式的基本性质：

【行列式界定定理】：设 $f(\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n)$ 为一个含有 n 个 n -维向量变元的函数，并满足斜对称性及多线性，则

$$(5.17) \quad f(\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n) = f(\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n) \cdot \det(\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n)$$

证明：令

$$F(\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n) = f(\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n) - f(\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n) \cdot \det(\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n)$$

易见 $F(\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n)$ 也满足斜对称性及多线性, 而且 $F(\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n) = 0$ 。我们要用上述三点推论 $F(\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n) \equiv 0$, 其证明如下:

我们可以用标准基底列向量, 把每个 \mathbf{a}_j 写成它们的线性组合, 即:

$$\begin{aligned}
 \mathbf{a}_1 &= a_{11}\mathbf{e}_1 + a_{21}\mathbf{e}_2 + \dots + a_{n1}\mathbf{e}_n = \sum_{i_1=1}^n a_{i_1 1}\mathbf{e}_{i_1} \\
 \mathbf{a}_2 &= a_{12}\mathbf{e}_1 + a_{22}\mathbf{e}_2 + \dots + a_{n2}\mathbf{e}_n = \sum_{i_2=1}^n a_{i_2 2}\mathbf{e}_{i_2} \\
 &\dots\dots\dots \\
 \mathbf{a}_j &= a_{1j}\mathbf{e}_1 + a_{2j}\mathbf{e}_2 + \dots + a_{nj}\mathbf{e}_n = \sum_{i_j=1}^n a_{i_j j}\mathbf{e}_{i_j} \\
 &\dots\dots\dots \\
 \mathbf{a}_n &= a_{1n}\mathbf{e}_1 + a_{2n}\mathbf{e}_2 + \dots + a_{nn}\mathbf{e}_n = \sum_{i_n=1}^n a_{i_n n}\mathbf{e}_{i_n}
 \end{aligned}
 \tag{5.18}$$

运用 F 的多线性, 逐步把 F 如下的展开:

$$\begin{aligned}
 F(\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n) &= F\left(\sum_{i_1=1}^n a_{i_1 1}\mathbf{e}_{i_1}, \sum_{i_2=1}^n a_{i_2 2}\mathbf{e}_{i_2}, \dots, \sum_{i_n=1}^n a_{i_n n}\mathbf{e}_{i_n}\right) \\
 &= \sum_{i_1=1}^n a_{i_1 1} F\left(\mathbf{e}_{i_1}, \sum_{i_2=1}^n a_{i_2 2}\mathbf{e}_{i_2}, \dots, \sum_{i_n=1}^n a_{i_n n}\mathbf{e}_{i_n}\right) \\
 &= \sum_{i_1=1}^n a_{i_1 1} \sum_{i_2=1}^n a_{i_2 2} F\left(\mathbf{e}_{i_1}, \mathbf{e}_{i_2}, \dots, \sum_{i_n=1}^n a_{i_n n}\mathbf{e}_{i_n}\right) \\
 &\quad \vdots \\
 &= \sum_{i_1=1}^n a_{i_1 1} \sum_{i_2=1}^n a_{i_2 2} \dots \sum_{i_n=1}^n a_{i_n n} F(\mathbf{e}_{i_1}, \mathbf{e}_{i_2}, \dots, \mathbf{e}_{i_n})
 \end{aligned}$$

在上述看来具有 n^n 项的展开式之中其实每一项都是 0, 所以其总和还是 0。其理由如下: 一方面当足标 (i_1, i_2, \dots, i_n) 之中有任何两个相同时, 则由 F 的斜对称性可知其值必为 0。另一方面在 n 个足标都相异时, 再由其斜对称性可得 $F(\mathbf{e}_{i_1}, \mathbf{e}_{i_2}, \dots, \mathbf{e}_{i_n}) = \pm F(\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n) = 0$ 。□

行列式的定义公式是取第一列的系数来展开, 这是因为我们前述的有效消元法目的是消去第二至第 n 变元; 若要保留第二变元而消去其他变元, 我们用的消元法就应该以第二列系数来展开行列式 (参考第四章第 70 页)。一般来说, 我们可借助斜对称性, 先把第一列和第 j

列交换后再展开，由此导出从第 j 列展开行列式的公式。例如 $j = 2$ ：

$$\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix} = - \begin{vmatrix} a_{12} & a_{11} & \cdots & a_{1n} \\ a_{22} & a_{21} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n2} & a_{n1} & \cdots & a_{nn} \end{vmatrix} \\ = -a_{21}D_{2,1} + a_{22}D_{2,2} - \cdots + (-1)^{n+2}a_{n2}D_{n,2}$$

及一般 j ：

$$\begin{vmatrix} a_{11} & \cdots & a_{1j} & \cdots & a_{1n} \\ a_{21} & \cdots & a_{2j} & \cdots & a_{2n} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nj} & \cdots & a_{nn} \end{vmatrix} = - \begin{vmatrix} a_{1j} & \cdots & a_{11} & \cdots & a_{1n} \\ a_{2j} & \cdots & a_{21} & \cdots & a_{2n} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ a_{nj} & \cdots & a_{n1} & \cdots & a_{nn} \end{vmatrix} \\ = -a_{1j} \begin{vmatrix} a_{22} & \cdots & a_{21} & \cdots & a_{2n} \\ a_{32} & \cdots & a_{31} & \cdots & a_{3n} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ a_{n2} & \cdots & a_{n1} & \cdots & a_{nn} \end{vmatrix} + a_{2j} \begin{vmatrix} a_{12} & \cdots & a_{11} & \cdots & a_{1n} \\ a_{32} & \cdots & a_{31} & \cdots & a_{3n} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ a_{n2} & \cdots & a_{n1} & \cdots & a_{nn} \end{vmatrix} - \cdots + \\ + (-1)^{n+2}a_{nj} \begin{vmatrix} a_{12} & \cdots & a_{11} & \cdots & a_{1n} \\ a_{22} & \cdots & a_{21} & \cdots & a_{2n} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ a_{n-1,2} & \cdots & a_{n-1,1} & \cdots & a_{n-1,n} \end{vmatrix} \\ = (-1)^{1+j}a_{1j}D_{1,j} + (-1)^{2+j}a_{2j}D_{2,j} + \cdots + (-1)^{n+j}a_{nj}D_{n,j}$$

由于 D 的第一列系数是在那些 $(n-1)$ -阶子行列式的第 $(j-1)$ 列中，所以可利用 $(j-2)$ 次换列把它们全都移回各自子行列式的第一列，因而得出上述的等式。总括来说，把行列式对第 j 列展开有下述公式：

$$(5.19) \quad D = (-1)^{1+j}a_{1j}D_{1,j} + (-1)^{2+j}a_{2j}D_{2,j} + \cdots + (-1)^{n+j}a_{nj}D_{n,j}$$

5.3 行列式的常用基本性质

本质上, n -阶行列式乃是一个非常特出的多项式函数。它含有 n^2 个变元, 其展开公式共计有 $n!$ 项, 每项的次数是 n , 亦即它是一个齐 n 次 n^2 -元多项式函数。一般的这种多项式函数个数多过恒河沙数, 绝大部分都是无用者, 唯行列式这个齐 n 次 n^2 -元多项式在整个数学的基础理论中扮演重要的角色。究其原因, 乃是它所具有的种种简洁美好的性质。在所有 n^2 -元的多项式之中, 行列式可以说是丽质天成, 独一无二的精要所在。总之, 行列式之所以重要、有用乃是它所独具的优良性质, 所以学习行列式的要点当然也就是要抓好行列式的基本性质及其用法。

在基础数学中, 行列式有两个自然的源起。其一就是我们在前面所研讨的 n -阶线性方程组的基础理论, 其二则是我们即将在后面研讨的 n -维平行体 (2-维的平行四边形和 3-维的平行六面体的自然推广) 的高维有向体积 (high dimensional oriented volume)。而上述两者的结合又是多变量微积分基础理论之所基。由此可见, 行列式不但具有简洁好用的基本性质, 而且也自然地是代数、几何、分析这数学三大支柱的基石所在。由此可见行列式广泛而且基本的用场的梗概。

从上面行列式的两个自然出处, 易见它的 n^2 个变元是自然而然地具有行、列的编排的 (这也就是行列式这个名称的来由)。因此, 在行列式基本性质的研讨中, 自然应该把每一列 (或每一行) 的 n 个变元看成一个整体。我们将用 $\mathbf{x}_j, 1 \leq j \leq n$, 表示 j -列向量; $\bar{\mathbf{x}}_i, 1 \leq i \leq n$, 表示 i -行向量, 亦即

$$\mathbf{x}_j = \begin{bmatrix} x_{1j} \\ x_{2j} \\ \vdots \\ x_{nj} \end{bmatrix}, \quad \bar{\mathbf{x}}_i = (x_{i1}, x_{i2}, \dots, x_{in})$$

在前面的讨论中, 我们把行列式看成其 n 个列向量的函数, 即 $\det(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n)$ 。由行列式的界定定理可见, 下述三点业已构成行列式的一组特征性质:

(i) 斜对称性: $\det(\dots \mathbf{x}_i \dots \mathbf{x}_j \dots) = -\det(\dots \mathbf{x}_j \dots \mathbf{x}_i \dots)$

$$(ii) \text{ 多线性: } \det(\lambda \mathbf{x}_1 + \mu \mathbf{y}_1, \mathbf{x}_2, \dots, \mathbf{x}_n) \\ = \lambda \det(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n) + \mu \det(\mathbf{y}_1, \mathbf{x}_2, \dots, \mathbf{x}_n)$$

以及对于其他各列同此。

$$(iii) \det(\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n) = 1$$

我们当然也可以把行列式看做它的 n 个行向量 $\{\bar{\mathbf{x}}_1, \bar{\mathbf{x}}_2, \dots, \bar{\mathbf{x}}_n\}$ 的多项式函数, 且以符号 $f(\bar{\mathbf{x}}_1, \bar{\mathbf{x}}_2, \dots, \bar{\mathbf{x}}_n)$ 记号之。我们很自然要问它是否也同样地具有上述三点基本性质? 假如具有, 则 $f(\bar{\mathbf{x}}_1, \bar{\mathbf{x}}_2, \dots, \bar{\mathbf{x}}_n) = \det(\bar{\mathbf{x}}_1, \bar{\mathbf{x}}_2, \dots, \bar{\mathbf{x}}_n)$ 。换言之, 亦即行列式是否具有在 $x_{ij} \leftrightarrow x_{ji}$ 互换之下保持不变的行、列对称性? 不难看到, 在 $n = 1, 2, 3$ 的情形是具有这种行、列互换, 保持不变的对称性的。由此, 我们当然又会想到用归纳法去研讨它是否普遍成立? 由前面的经验可见, 我们只要能够验证行列式也可以对于其第一行有下述展开公式, 即

$$(5.20) \quad D = |x_{ij}| \stackrel{?}{=} \sum_{j=1}^n (-1)^{j+1} x_{1j} D_{1,j}$$

假若上式得证, 则可用完全同样的论证, 归纳地证明 $f(\bar{\mathbf{x}}_1, \bar{\mathbf{x}}_2, \dots, \bar{\mathbf{x}}_n)$ 的斜对称性和多线性, 而 $f(\bar{\mathbf{e}}_1, \bar{\mathbf{e}}_2, \dots, \bar{\mathbf{e}}_n) = 1$ 是显然的。兹验证 (5.20)-式如下:

(5.20)-式之证明: 设 $|a_{ij}| \neq 0$, 由行列式的斜对称性和其归纳定义式 (5.2) 可见

$$(5.21) \quad \det(\mathbf{a}_1, \dots, \mathbf{a}_{j-1}, \mathbf{e}_1, \mathbf{a}_{j+1}, \dots, \mathbf{a}_n) = (-1)^{j+1} D_{1,j}$$

再者, 由方程组 (5.1) 在 $\mathbf{b} = \mathbf{e}_1$ 时唯一解的公式, 即有

$$(5.22) \quad x_j = \frac{1}{D} \det(\dots, \mathbf{e}_1, \dots) = \frac{1}{D} (-1)^{j+1} D_{1,j}$$

将它代入 (5.1) 的第一式, 即有

$$(5.23) \quad \sum_{j=1}^n a_{1j} x_j = \frac{1}{D} \sum_{j=1}^n (-1)^{j+1} a_{1j} D_{1,j} = 1$$

亦即

$$(5.24) \quad \sum_{j=1}^n (-1)^{j+1} a_{1j} D_{1,j} = D$$

对于任给 $|a_{ij}| \neq 0$ 的取值恒成立。因此 (5.20) 乃是 n^2 -变元的恒等式。在这里，我们用到一个多项式代数中简单有用的事实，即一个 m 元多项式 $P(x_1, \dots, x_m)$ ，若它在每个变元皆可相互无关地选用无穷多个取值使得 $P(x_1, \dots, x_m) = 0$ ，则 P 恒等于 0。此事乃是一个非零单元多项式至多只有有限个根的直接推论。□

【推论】：行列式对于行向量也有同样的斜对称性和多线性。再者，它也有下述对于第 i 行的展开式

$$(5.25) \quad D = |x_{ij}| = \sum_{j=1}^n (-1)^{i+j} a_{ij} D_{i,j}$$

5.4 矩阵的乘法公式和行列式的乘法公式

在研讨矩阵运算 (matrix operations) 之前，我们先要提纲挈领地谈一谈什么是抽象化？它的基本想法和做法为何？大体上来说，人类理性认知大自然的逐步过程不外乎实事求是地认识问题和解决问题，从此精益求精地寻求大自然至精至简的结构与规律，然后再用其所得的精简来以简御繁。可以说这也就是整个理性文明的一个简朴的概括。而在世代相承理性认知自然的途径中，抽象化和数理分析一直是行之有效的两种基本的方法论。当我们研讨某些基本问题时，确切认识该问题的本质和精要所在，当然是能否有效解决该问题的关键所在。其实，不但要由表及里，抓好问题的本质和精要所在（亦即择其精要），而且还要把它们妥加组织（例如各种各样数理模型 (mathematical models)）；这种把抽象所得的精要妥加组织之所得也就是原给问题的抽象化 (abstraction)。对于各种各样定量型的问题，其抽象化就往往是一种数理模型。所以抽象化往往和数理分析结合使用，而抽象化则是能够有效运用数理分析对于各种各样问题进行系统、有效的深入研讨的基础所在。显然，举例说明乃是解说抽象化这种方法论 (methodology) 的不二法门。本节所要讨论的矩阵运算也就是一个「抽象化」的实例。

在单元函数中，齐一次函数 $f(x) = c \cdot x$ 乃是最为简单基本的一种类型。数系的分配律在这种函数的表现就是：

「所有齐一次（单元）函数 $f(x)$ 皆满足下述两点，即：

$$(5.26) \quad (i) f(a+b) = f(a) + f(b), \quad (ii) f(k \cdot a) = kf(a)$$

反之，任何一个具有上述两个性质的函数也一定是齐一次的。」

亦即上述 (i), (ii) 两点乃是齐一次单元函数的特征性质。在 n -元的情形，一个齐一次函数可以写成

$$(5.27) \quad f(x_1, x_2, \dots, x_n) = c_1x_1 + c_2x_2 + \dots + c_nx_n$$

它的变域乃是所有 n -数串 (n -tuples) 所组成的集合，而它本身又被上述 n 个系数所组成的 n -数串 (c_1, c_2, \dots, c_n) 所唯一确定。为了配合线性方程组的系统研讨，我们把一个横向排列的 n -数串想成是一个行向量，用一个简写符号

$$\bar{c} = (c_1, c_2, \dots, c_n)$$

表达之；而把一个纵向排列的 n -数串想成是一个列向量，用一个简写符号

$$\mathbf{a} = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix}, \quad \mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$

表达之。而且用 \mathbf{x} 表示以所有纵向 n -数串的集合为其变域的向量变元 (vector variable)。在此不难看到，假如我们在纵向 n -数串（亦即 n -维列向量）之间引进下述「加法」和「倍积」，即：

$$(5.28) \quad \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix} + \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix} = \begin{bmatrix} a_1 + b_1 \\ a_2 + b_2 \\ \vdots \\ a_n + b_n \end{bmatrix} \quad k \cdot \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix} = \begin{bmatrix} ka_1 \\ ka_2 \\ \vdots \\ ka_n \end{bmatrix}$$

则前述单元齐一次函数的特征性质就可以得到同样形式的推广。亦即

【引理】： $f(\mathbf{x})$ 是一个齐一次函数的充要条件是它满足下述两点，即

$$(5.26') \quad (i) f(\mathbf{a} + \mathbf{b}) = f(\mathbf{a}) + f(\mathbf{b}), \quad (ii) f(k\mathbf{a}) = kf(\mathbf{a})$$

[证明]：必要性：设 $f(\mathbf{x}) = \sum_{i=1}^n c_i x_i$ 。则

$$\begin{aligned} f(\mathbf{a} + \mathbf{b}) &= \sum_{i=1}^n c_i(a_i + b_i) = \sum_{i=1}^n c_i a_i + \sum_{i=1}^n c_i b_i \\ &= f(\mathbf{a}) + f(\mathbf{b}) \\ f(k\mathbf{a}) &= \sum_{i=1}^n c_i(ka_i) = k \sum_{i=1}^n c_i a_i = kf(\mathbf{a}) \end{aligned}$$

充分性：设 $f(\mathbf{x})$ 满足 (i), (ii) 则

$$\begin{aligned} f(\mathbf{x}) &= f\left(\sum_{i=1}^n x_i \mathbf{e}_i\right) = \sum_{i=1}^n f(x_i \mathbf{e}_i) \\ &= \sum_{i=1}^n x_i f(\mathbf{e}_i) = \sum_{i=1}^n c_i x_i, \quad c_i = f(\mathbf{e}_i), \quad 1 \leq i \leq n \quad \square \end{aligned}$$

[注]：(一) 上述简短的论证之中，主要是数系分配律的运用。

(二) 在数学中，我们把上述 (5.26')-式中的两点作为「线性函数」的定义条件式。再者，一个含有多个向量元的函数若对于其中每一个向量元皆各别满足上述两个条件式，则称之为「多线性函数」。例如 $\det(\mathbf{x}_1, \dots, \mathbf{x}_n)$ 就是一个多线性函数。

(三) 上述简朴自然的「妥加组织」(即定义 n -数串之间的加法和倍积)使得分配律在线性函数上的表现更加简明、易用，乃是一种自然而然的抽象化。

5.4.1 矩阵运算

其实，我们还可以进一步定义一对横向 n -数串和纵向 n -数串之间的乘法，即

$$(5.29) \quad \bar{\mathbf{c}} \cdot \mathbf{a} = \sum_{i=1}^n c_i a_i$$

则一个 n 元齐一次函数（亦即线性函数）就可以表达成

$$f(\mathbf{x}) = \bar{\mathbf{c}} \cdot \mathbf{x}$$

这种十分精简的形式。而它的特征性质（亦即 (5.26')-式）又变成：

$$\bar{\mathbf{c}} \cdot (\mathbf{a} + \mathbf{b}) = \bar{\mathbf{c}} \cdot \mathbf{a} + \bar{\mathbf{c}} \cdot \mathbf{b}, \quad \bar{\mathbf{c}} \cdot (k\mathbf{a}) = k\bar{\mathbf{c}} \cdot \mathbf{a}$$

这种惯用、好用的运算律。

我们可以把上述「乘法」想成是一个 $1 \times n$ -矩阵和一个 $n \times 1$ -矩阵之间的乘法。是否还可以进一步地定义一种 $m \times n$ -矩阵和 $n \times \ell$ -矩阵之间的乘法呢？首先让我们看一看 $\ell = 1$ 的情形。一个 $m \times n$ -矩阵 (c_{ij}) , $1 \leq i \leq m, 1 \leq j \leq n$, 可以想成是 m 个 n 元线性函数的系数矩阵，它的 m 个行向量 $\bar{\mathbf{c}}_i, 1 \leq i \leq m$, 就是它们各别的系数 n -数串。由此可见， m 个 n 元线性函数组可以简约地写成

$$(5.30) \quad (c_{ij}) \cdot (\mathbf{x}) = \begin{bmatrix} \bar{\mathbf{c}}_1 \cdot \mathbf{x} \\ \bar{\mathbf{c}}_2 \cdot \mathbf{x} \\ \vdots \\ \bar{\mathbf{c}}_m \cdot \mathbf{x} \end{bmatrix}$$

由此可以想到一个 $m \times n$ -矩阵 (c_{ij}) 和一个 $n \times 1$ -矩阵 \mathbf{x} 之间乘法可以很自然地定义如下，即

$$(5.31) \quad (c_{ij}) \cdot \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} \sum_{j=1}^n c_{1j}x_j \\ \sum_{j=1}^n c_{2j}x_j \\ \vdots \\ \sum_{j=1}^n c_{mj}x_j \end{bmatrix}$$

以上述定义为基础，我们就可以用下述思路去探索一般情形， $m \times n$ -矩阵和 $n \times \ell$ -矩阵的乘法应该如何赋予自然好用的定义。令

$$A = (a_{ij}), \quad 1 \leq i \leq m, 1 \leq j \leq n$$

$$B = (b_{jk}), \quad 1 \leq j \leq n, 1 \leq k \leq \ell$$

分别是 $m \times n$ -矩阵和 $n \times \ell$ -矩阵； \mathbf{x} 和 \mathbf{t} 分别是 $n \times 1$ 和 $\ell \times 1$ 的变元列向量。则

$$(5.32) \quad A \cdot \mathbf{x} = \begin{bmatrix} \sum_{j=1}^n a_{1j}x_j \\ \sum_{j=1}^n a_{2j}x_j \\ \vdots \\ \sum_{j=1}^n a_{mj}x_j \end{bmatrix}, \quad B \cdot \mathbf{t} = \begin{bmatrix} \sum_{k=1}^{\ell} b_{1k}t_k \\ \sum_{k=1}^{\ell} b_{2k}t_k \\ \vdots \\ \sum_{k=1}^{\ell} b_{nk}t_k \end{bmatrix}$$

分别是由 m 个 n 元线性函数和 n 个 ℓ 元线性函数纵列而成的 $m \times 1$ 和 $m \times 1$ 矩阵。设想我们令 $\mathbf{x} = B \cdot \mathbf{t}$ 并以此代入 $A \cdot \mathbf{x}$ ，即得 m 个 ℓ 元线性函数，用

$$C = (c_{ik}), \quad 1 \leq i \leq m, \quad 1 \leq k \leq \ell$$

表示如此所得的线性函数串的系数矩阵，亦即有

$$(5.33) \quad A \cdot \mathbf{x} = A \cdot (B \cdot \mathbf{t}) = C \cdot \mathbf{t}$$

直接代入，用分配律展开集项即可算得

$$(5.34) \quad \sum_{j=1}^n a_{ij}x_j = \sum_{j=1}^n a_{ij} \left(\sum_{k=1}^{\ell} b_{jk}t_k \right) = \sum_{k=1}^{\ell} \left(\sum_{j=1}^n a_{ij}b_{jk} \right) \cdot t_k$$

所以

$$(5.35) \quad c_{ik} = \sum_{j=1}^n a_{ij}b_{jk}$$

由此可以想到 $A \cdot B$ 是以 (5.35) 所表达的 c_{jk} 为其 (i, k) -系数的 $m \times \ell$ 矩阵。这也就是我们所探讨的矩阵乘法公式。请注意，上述矩阵乘法仅仅在 $m \times n$ -矩阵和 $n \times \ell$ -矩阵之间才有定义！再者，在同一类型的矩阵之间，当然还可以赋予下述加法和倍积的定义。即

$$(5.36) \quad \begin{aligned} (a_{ij}) + (b_{ij}) &= (a_{ij} + b_{ij}) \\ \lambda(a_{ij}) &= (\lambda a_{ij}) \quad 1 \leq i \leq m, \quad 1 \leq j \leq n \end{aligned}$$

容易验证上述三种矩阵运算满足下列运算律，即

$$\begin{aligned} k \cdot (A_1 + A_2) &= kA_1 + kA_2 \\ A \cdot (B_1 + B_2) &= A \cdot B_1 + A \cdot B_2 \\ (A_1 + A_2) \cdot B &= A_1 \cdot B + A_2 \cdot B \\ (kA) \cdot B &= kA \cdot B = A \cdot (kB) \end{aligned}$$

5.4.2 行列式乘法公式

当 A, B 都是 $n \times n$ -矩阵时, $A \cdot B$ 也是 $n \times n$ -矩阵而且有下述公式

【定理】: $\det(A \cdot B) = \det(A) \cdot \det(B)$

证明: 现在我们先任意选定矩阵 A , 然后考虑 $\det(AB)$ 为 B 的 n 个列向量 $\mathbf{b}_1, \dots, \mathbf{b}_n$ 的函数。由以上矩阵乘积的表达式可看到,

$$\begin{aligned} & \det(A [\mathbf{b}_1 \cdots (\lambda \mathbf{b}_j + \mu \mathbf{b}_j^*) \cdots \mathbf{b}_n]) \\ &= \det [A\mathbf{b}_1 \cdots A(\lambda \mathbf{b}_j + \mu \mathbf{b}_j^*) \cdots A\mathbf{b}_n] \\ &= \det [A\mathbf{b}_1 \cdots (\lambda A\mathbf{b}_j + \mu A\mathbf{b}_j^*) \cdots A\mathbf{b}_n] \\ &= \lambda \det [A\mathbf{b}_1 \cdots A\mathbf{b}_j \cdots A\mathbf{b}_n] + \mu \det [A\mathbf{b}_1 \cdots A\mathbf{b}_j^* \cdots A\mathbf{b}_n] \\ &= \lambda \det (A [\mathbf{b}_1 \cdots \mathbf{b}_j \cdots \mathbf{b}_n]) + \mu \det (A [\mathbf{b}_1 \cdots \mathbf{b}_j^* \cdots \mathbf{b}_n]) \end{aligned}$$

因此 $\det(AB)$ 对列向量 $\mathbf{b}_1, \dots, \mathbf{b}_n$ 是具有多线性的。再者, 如有某些 $\mathbf{b}_k = \mathbf{b}_\ell$, 则 $A\mathbf{b}_k = A\mathbf{b}_\ell$, 所以

$$(5.37) \quad \det(AB) = \det [A\mathbf{b}_1 \cdots A\mathbf{b}_k \cdots A\mathbf{b}_\ell \cdots A\mathbf{b}_n] = 0$$

亦即 $\det(AB)$ 对列向量 $\mathbf{b}_1, \dots, \mathbf{b}_n$ 是具有斜对称性。由行列式界定定理得常数 K , 使得:

$$(5.38) \quad \det(AB) = \det (A [\mathbf{b}_1 \cdots \mathbf{b}_n]) = K \det [\mathbf{b}_1 \cdots \mathbf{b}_n]$$

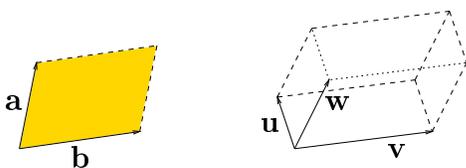
对所有的列向量 $\mathbf{b}_1, \dots, \mathbf{b}_n$ 恒成立, 其中 $K = \det A([\mathbf{e}_1 \cdots \mathbf{e}_n]) = \det A$ 。所以

$$(5.38') \quad \det(AB) = (\det A)(\det B)$$

上述等式若直接用行列式的展开公式验证乃是很繁复和费劲的!

5.4.3 平行体体积与行列式

在平面上, 两个不共线的向量 \mathbf{a}, \mathbf{b} 可以「张成」(span) 一个平行四边形; 而在空间中, 三个不共面的向量 $\mathbf{u}, \mathbf{v}, \mathbf{w}$ 可以张成一个平行六面体, 如下图所示:



[图 5-2]

由简单的几何分析容易求得上述平行四边形的面积和平行六面体的体积，即

$$(5.39) \quad \text{面积} = |\mathbf{a}| \cdot |\mathbf{b}| \cdot |\sin \angle \mathbf{a}, \mathbf{b}|, \quad \text{体积} = |\mathbf{u} \times \mathbf{v} \cdot \mathbf{w}|$$

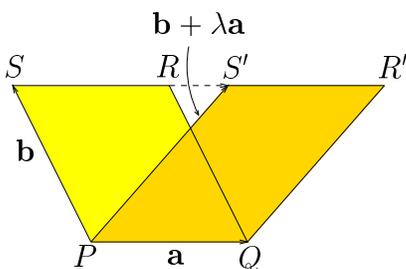
在定义 \times -积或写出平行体体积公式时，还会引入行列式记号：

$$(5.40) \quad \mathbf{u} \times \mathbf{v} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ u_1 & u_2 & u_3 \\ v_1 & v_2 & v_3 \end{vmatrix}, \quad \mathbf{u} \times \mathbf{v} \cdot \mathbf{w} = \begin{vmatrix} w_1 & w_2 & w_3 \\ u_1 & u_2 & u_3 \\ v_1 & v_2 & v_3 \end{vmatrix}$$

能用以上记号去表达体积是纯属碰巧，或是由于有某些本质上的原因呢？若我们只停留在讨论表面公式，是难以体认其本质的。现在，我们对行列式已有著一些本质上的认识，所以跟著也要对平行四边形面积、平行六面体体积找寻一些本质上的认识。

在平面上一个由 \mathbf{a} , \mathbf{b} 所张成的平行四边形，我们以 $\tilde{A}(\mathbf{a}, \mathbf{b})$ 表示其通常（恒取正值）的面积。不难看出 $\tilde{A}(\mathbf{a}, \mathbf{b})$ 是具有以下的基本性质：

- (i) 对称性： $\tilde{A}(\mathbf{a}, \mathbf{b}) = \tilde{A}(\mathbf{b}, \mathbf{a})$
- (ii) 拟比性： $\tilde{A}(\lambda \mathbf{a}, \mathbf{b}) = |\lambda| \tilde{A}(\mathbf{a}, \mathbf{b}) = \tilde{A}(\mathbf{a}, \lambda \mathbf{b})$
- (iii) 斜移不变性： $\tilde{A}(\mathbf{a} + \lambda \mathbf{b}, \mathbf{b}) = \tilde{A}(\mathbf{a}, \mathbf{b}) = \tilde{A}(\mathbf{a}, \mathbf{b} + \mu \mathbf{a})$



[图 5-3]

[$\square PQRS$ 和 $\square PQR'S'$ 之间的变化只是将 \overline{SR} 作平行滑动。从几何观点来看，两者为同底同高的平行四边形，所以面积应该相等。]

若在平面上先选定一个正角之转向，亦即其上的一个定向 (orientation)，而且定义 \mathbf{a}, \mathbf{b} 所张成的平行四边形的有向面积 (oriented area) 为

$$(5.41) \quad A(\mathbf{a}, \mathbf{b}) = |\mathbf{a}| \cdot |\mathbf{b}| \cdot \sin \angle \mathbf{a}, \mathbf{b}$$

上述面积在由 \mathbf{a} 到 \mathbf{b} 的转向为正时取正值，为负时取负值。易见有向面积 $A(\mathbf{a}, \mathbf{b})$ 满足下列基本性质，即

(i) 斜对称性： $A(\mathbf{a}, \mathbf{b}) = -A(\mathbf{b}, \mathbf{a})$

(ii) 比性： $A(\lambda \mathbf{a}, \mathbf{b}) = \lambda A(\mathbf{a}, \mathbf{b}) = A(\mathbf{a}, \lambda \mathbf{b})$

(iii) 斜移不变性： $A(\mathbf{a} + \lambda \mathbf{b}, \mathbf{b}) = A(\mathbf{a}, \mathbf{b}) = A(\mathbf{a}, \mathbf{b} + \mu \mathbf{a})$

把上述两种面积函数相比较，骤看起来，(i) 略优于 (i) 而 (ii) 则略优于 (ii)，似乎是半斤八两。但细加分析，则 $A(\mathbf{a}, \mathbf{b})$ 远比 $\tilde{A}(\mathbf{a}, \mathbf{b})$ 优越！此事可以由下述结果充分说明其优越性：

$$(5.42) \quad A(\mathbf{a}, \mathbf{b} + \mathbf{c}) = A(\mathbf{a}, \mathbf{b}) + A(\mathbf{a}, \mathbf{c})$$

证明：令 \mathbf{e} 为在平面中使得 $\angle \mathbf{a}, \mathbf{e} = \pi/2$ 的单位长向量，则可把 \mathbf{b} 和 \mathbf{c} 分解成垂直于 \mathbf{a} 和平行于 \mathbf{a} 的分向量之和如下：

$$\mathbf{b} = \lambda_1 \mathbf{e} + \mu_1 \mathbf{a}, \quad \mathbf{c} = \lambda_2 \mathbf{e} + \mu_2 \mathbf{a}$$

$$\mathbf{b} + \mathbf{c} = (\lambda_1 + \lambda_2) \mathbf{e} + (\mu_1 + \mu_2) \mathbf{a}$$

由此再连续使用性质 (ii), (iii) 即有

$$\begin{aligned} A(\mathbf{a}, \mathbf{b} + \mathbf{c}) &= A(\mathbf{a}, (\lambda_1 + \lambda_2) \mathbf{e}) && [\text{用 (iii)}] \\ &= (\lambda_1 + \lambda_2) A(\mathbf{a}, \mathbf{e}) && [\text{用 (ii)}] \\ &= \lambda_1 A(\mathbf{a}, \mathbf{e}) + \lambda_2 A(\mathbf{a}, \mathbf{e}) \\ &= A(\mathbf{a}, \lambda_1 \mathbf{e}) + A(\mathbf{a}, \lambda_2 \mathbf{e}) && [\text{用 (ii)}] \\ &= A(\mathbf{a}, \mathbf{b}) + A(\mathbf{a}, \mathbf{c}) && [\text{用 (iii)}] \end{aligned}$$

因此， $A(\mathbf{a}, \mathbf{b})$ 是以向量 \mathbf{a}, \mathbf{b} 为变元的函数，并满足斜对称性和多线性。所以由行列式界定定理得有常数 K ，使得：

$$(5.43) \quad A(\mathbf{a}, \mathbf{b}) = K \det(\mathbf{a}, \mathbf{b})$$

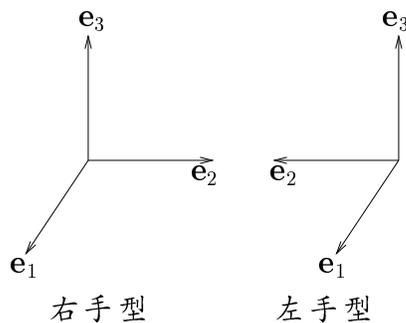
对所有列向量 \mathbf{a}, \mathbf{b} 恒成立，其中 $K = A(\mathbf{e}_1, \mathbf{e}_2) = 1$ ，亦即

$$(5.43') \quad A(\mathbf{a}, \mathbf{b}) = \det(\mathbf{a}, \mathbf{b})$$

现在我们已看出平行四边形的面积和行列式的确有著本质上的直接关联。至于在平行六面体体积时的情况，亦能类似地把这个本质上的关联找出来。首先我们要讨论空间的定向问题。

空间的定向和平行六面体的定向体积：

一条直线只有两个方向，其定向就是选定其一为正向；一个平面上的转角方向也只有两个，其定向就是选取其一为正的转角方向。现在让我们来分析一下空间的「定向」应该如何定义。空间中的所有正交标架可以分成两类，即如下图所示的右手型和左手型，两个同型者可以用适当的平移和转轴的组合同样相互变换，而不同型者则无法如此相互变换。



[图 5-4]

再者，一个反射对称则把两种类型的正交标架互换。通常的做法是：空间的定向乃是在上述两种类型中选定其一为正向者，而一个由 $(\mathbf{u}, \mathbf{v}, \mathbf{w})$ 所张的平行六面体的定向体积之正负则由标架 $(\mathbf{u}, \mathbf{v}, \mathbf{w})$ 是正向型或负向型而定。在这样的空间的定向概念下，定向平行六面体的有向体积是一个满足下列基本性质的函数 $V(\mathbf{u}, \mathbf{v}, \mathbf{w})$ ：

$$(i) \text{ 斜对称性: } V(\mathbf{u}, \mathbf{v}, \mathbf{w}) = V(\mathbf{v}, \mathbf{w}, \mathbf{u}) = V(\mathbf{w}, \mathbf{u}, \mathbf{v}) = -V(\mathbf{u}, \mathbf{w}, \mathbf{v}) = -V(\mathbf{v}, \mathbf{u}, \mathbf{w}) = -V(\mathbf{w}, \mathbf{v}, \mathbf{u})$$

$$(ii) \text{ 比性: } V(\lambda\mathbf{u}, \mathbf{v}, \mathbf{w}) = \lambda V(\mathbf{u}, \mathbf{v}, \mathbf{w}) \text{ 等}$$

$$(iii) \text{ 斜移不变性: } V(\mathbf{u} + \lambda\mathbf{v} + \mu\mathbf{w}, \mathbf{v}, \mathbf{w}) = V(\mathbf{u}, \mathbf{v}, \mathbf{w}) \text{ 等}$$

再者，我们也可以同样地由 (ii) 和 (iii) 推导出下述结果：

$$(5.44) \quad V(\mathbf{u}_1 + \mathbf{u}_2, \mathbf{v}, \mathbf{w}) = V(\mathbf{u}_1, \mathbf{v}, \mathbf{w}) + V(\mathbf{u}_2, \mathbf{v}, \mathbf{w})$$

证明：若 \mathbf{v}, \mathbf{w} 所张成的平行四边形不是 2-维的，则上述三个体积都是 0，所以上述等式当然成立。

若 \mathbf{v}, \mathbf{w} 所张成的是一个非蜕化的平行四边形，在其所在平面的两个单位长法向量中选定其一为 \mathbf{e} ，把 $\mathbf{u}_1, \mathbf{u}_2$ 分别分解成 $\mathbf{e}, \mathbf{v}, \mathbf{w}$ 的倍积之和，即有：

$$\begin{aligned} \mathbf{u}_1 &= u_1 \mathbf{e} + \lambda_1 \mathbf{v} + \mu_1 \mathbf{w} \\ \mathbf{u}_2 &= u_2 \mathbf{e} + \lambda_2 \mathbf{v} + \mu_2 \mathbf{w} \\ \mathbf{u}_1 + \mathbf{u}_2 &= (u_1 + u_2) \mathbf{e} + (\lambda_1 + \lambda_2) \mathbf{v} + (\mu_1 + \mu_2) \mathbf{w} \end{aligned}$$

由此再连续使用性质 (ii), (iii) 即有

$$\begin{aligned} V(\mathbf{u}_1 + \mathbf{u}_2, \mathbf{v}, \mathbf{w}) &= V((u_1 + u_2) \mathbf{e}, \mathbf{v}, \mathbf{w}) && \text{[用 (iii)]} \\ &= (u_1 + u_2) V(\mathbf{e}, \mathbf{v}, \mathbf{w}) && \text{[用 (ii)]} \\ &= u_1 V(\mathbf{e}, \mathbf{v}, \mathbf{w}) + u_2 V(\mathbf{e}, \mathbf{v}, \mathbf{w}) \\ &= V(u_1 \mathbf{e}, \mathbf{v}, \mathbf{w}) + V(u_2 \mathbf{e}, \mathbf{v}, \mathbf{w}) && \text{[用 (ii)]} \\ &= V(\mathbf{u}_1, \mathbf{v}, \mathbf{w}) + V(\mathbf{u}_2, \mathbf{v}, \mathbf{w}) && \text{[用 (iii)]} \end{aligned}$$

因此， $V(\mathbf{u}, \mathbf{v}, \mathbf{w})$ 亦是以向量 $\mathbf{u}, \mathbf{v}, \mathbf{w}$ 为变元的斜对称多线性函数，所以存在常数 K ，使得

$$(5.45) \quad V(\mathbf{u}, \mathbf{v}, \mathbf{w}) = K \det(\mathbf{u}, \mathbf{v}, \mathbf{w})$$

对所有 $\mathbf{u}, \mathbf{v}, \mathbf{w}$ 恒成立，其中 $K = V(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3) = 1$ ，亦即

$$(5.45') \quad V(\mathbf{u}, \mathbf{v}, \mathbf{w}) = \det(\mathbf{u}, \mathbf{v}, \mathbf{w})$$

所以，在一个 n -维的空间内，由 n 个向量 $\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n$ 所张成的平行多面体的 n -维有向体积便自然而然地定义为

$$(5.46) \quad V(\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n) = \det(\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n)$$

5.5 行列式的几个应用范例

5.5.1 几何图形的坐标方程

在绪论中, 我们讨论到用待定系数法把过平面上两点 $(x_1, y_1), (x_2, y_2)$ 的直线方程, 转化为下述线性方程组有非零解 $\{A, B, C\}$ 的条件式:

$$(5.47) \quad \begin{cases} x_1 \cdot A + y_1 \cdot B + 1 \cdot C = 0 \\ x_2 \cdot A + y_2 \cdot B + 1 \cdot C = 0 \\ x \cdot A + y \cdot B + 1 \cdot C = 0 \end{cases}$$

由行列式的理论即得上述方程组有非零解的充要条件为:

$$(5.48) \quad \begin{vmatrix} x_1 & y_1 & 1 \\ x_2 & y_2 & 1 \\ x & y & 1 \end{vmatrix} = 0$$

同理, 在空间中过三点 $(x_1, y_1, z_1), (x_2, y_2, z_2), (x_3, y_3, z_3)$ 的平面坐标方程 $Ax + By + Cz + D = 0$ 满足

$$(5.49) \quad \begin{cases} Ax_1 + By_1 + Cz_1 + D = 0 \\ Ax_2 + By_2 + Cz_2 + D = 0 \\ Ax_3 + By_3 + Cz_3 + D = 0 \end{cases}$$

把上述方程组和 $Ax + By + Cz + D = 0$ 想成以 A, B, C, D 为变元的线性方程组, 亦即

$$(5.50) \quad \begin{cases} x_1 \cdot A + y_1 \cdot B + z_1 \cdot C + 1 \cdot D = 0 \\ x_2 \cdot A + y_2 \cdot B + z_2 \cdot C + 1 \cdot D = 0 \\ x_3 \cdot A + y_3 \cdot B + z_3 \cdot C + 1 \cdot D = 0 \\ x \cdot A + y \cdot B + z \cdot C + 1 \cdot D = 0 \end{cases}$$

则该平面的坐标方程为

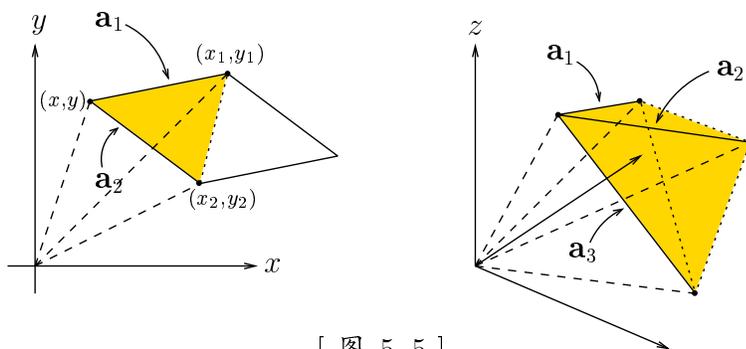
$$(5.51) \quad \begin{vmatrix} x_1 & y_1 & z_1 & 1 \\ x_2 & y_2 & z_2 & 1 \\ x_3 & y_3 & z_3 & 1 \\ x & y & z & 1 \end{vmatrix} = 0$$

从另一个角度来看, 上述直线、平面方程亦可想成是三点共线、四点共面的条件; 而容许满足条件的第三点 (x, y) 或第四点 (x, y, z) 随意走动, 则可得出代表该直线或平面的方程。

[注]: 上述所得的行列式其实和该三点所定的三角形面积 (或该四点所定的四面体体积) 有关。例如, 把 (5.48)-式里的行列式的第一行和第二行分别减去第三行, 再从第三列展开可得

$$\begin{vmatrix} x_1 & y_1 & 1 \\ x_2 & y_2 & 1 \\ x & y & 1 \end{vmatrix} = \begin{vmatrix} x_1 - x & y_1 - y & 0 \\ x_2 - x & y_2 - y & 0 \\ x & y & 1 \end{vmatrix} = \begin{vmatrix} x_1 - x & y_1 - y \\ x_2 - x & y_2 - y \end{vmatrix}$$

而上述的二阶行列式就是向量 $\mathbf{a}_1 = \begin{bmatrix} x_1 - x \\ y_1 - y \end{bmatrix}$ 和 $\mathbf{a}_2 = \begin{bmatrix} x_2 - x \\ y_2 - y \end{bmatrix}$ 所张成的平行四边形面积, 如下图 (左) 所示:



[图 5-5]

因此,

$$\begin{vmatrix} x_1 & y_1 & 1 \\ x_2 & y_2 & 1 \\ x & y & 1 \end{vmatrix} = \begin{matrix} \text{以 } (x, y), (x_1, y_1), (x_2, y_2) \text{ 为顶点} \\ \text{的三角形的有向面积的二倍。} \end{matrix}$$

所以, 三点共线的条件就是此三角形面积为 0。同理亦可求得

$$\begin{vmatrix} x_1 & y_1 & z_1 & 1 \\ x_2 & y_2 & z_2 & 1 \\ x_3 & y_3 & z_3 & 1 \\ x & y & z & 1 \end{vmatrix} = \begin{matrix} \text{以 } (x, y, z), (x_1, y_1, z_1), (x_2, y_2, z_2), (x_3, y_3, z_3) \\ \text{为顶点的四面体的有向体积的六倍。} \end{matrix}$$

亦即四点共面的条件就是此四面体的体积为 0。而过五点的二次曲线的

方程 (亦即六点共在同一二次曲线上的条件式) 可以写成:

$$(5.52) \quad \begin{vmatrix} x_1^2 & x_1 y_1 & y_1^2 & x_1 & y_1 & 1 \\ x_2^2 & x_2 y_2 & y_2^2 & x_2 & y_2 & 1 \\ x_3^2 & x_3 y_3 & y_3^2 & x_3 & y_3 & 1 \\ x_4^2 & x_4 y_4 & y_4^2 & x_4 & y_4 & 1 \\ x_5^2 & x_5 y_5 & y_5^2 & x_5 & y_5 & 1 \\ x^2 & xy & y^2 & x & y & 1 \end{vmatrix} = 0$$

5.5.2 行列式与插值公式

在第一章的讨论中, 我们已求得 Lagrange 插值公式, 即在相异位置 $\{a_i, 0 \leq i \leq n\}$ 取值 $\{b_i, 0 \leq i \leq n\}$ 而次数不高于 n 的多项式 $f(x)$ 为:

$$(5.53) \quad f(x) = b_0 f_0(x) + b_1 f_1(x) + \cdots + b_n f_n(x) = \sum_{i=0}^n b_i f_i(x)$$

其中

$$\begin{aligned} f_0(x) &= \frac{(x-a_1)(x-a_2)\cdots(x-a_n)}{(a_0-a_1)(a_0-a_2)\cdots(a_0-a_n)} = \prod_{i=1}^n (x-a_i) / \prod_{i=1}^n (a_0-a_i) \\ &\dots\dots\dots \\ f_j(x) &= \prod_{i \neq j} (x-a_i) / \prod_{i \neq j} (a_j-a_i) \\ &\dots\dots\dots \\ f_n(x) &= \prod_{i=0}^{n-1} (x-a_i) / \prod_{i=0}^{n-1} (a_n-a_i) \end{aligned}$$

上述公式稍嫌繁复, 难于处理。现在不妨采用前述的待定系数法以行列式写下 $f(x)$ 系数集 $\{c_k, 0 \leq k \leq n\}$ 的条件方程组

$$(5.54) \quad \begin{aligned} -y + \sum_{k=0}^n x^k c_k &= 0 \\ -v_j + \sum_{k=0}^n a_j^k c_k &= 0, \quad 0 \leq j \leq n \end{aligned}$$

把它们想成以下述 $(n+2) \times (n+2)$ 方阵为其系数矩阵的齐次线性方程组

$$(5.54') \quad \begin{pmatrix} -y & 1 & x & \dots & x^n \\ -v_0 & 1 & a_0 & \dots & a_0^n \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ -v_j & 1 & a_j & \dots & a_j^n \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ -v_n & 1 & a_n & \dots & a_n^n \end{pmatrix}$$

则它具有非零解 $(-1, c_0, \dots, c_n)$ 的充要条件就是上述方阵的行列式等于 0, 亦即

$$(5.55) \quad \begin{vmatrix} -y & 1 & x & \dots & x^n \\ -v_0 & 1 & a_0 & \dots & a_0^n \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ -v_j & 1 & a_j & \dots & a_j^n \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ -v_n & 1 & a_n & \dots & a_n^n \end{vmatrix} = 0$$

把上述行列式以第一列 (first column) 展开, 遍除以 $(-y)$ 的系数子行列式, 再用熟知的 Vandermonde 行列式公式即得回 Lagrange 公式

$$(5.53') \quad -y + \sum_{k=0}^n v_k f_k(x) = 0$$

所以 (5.55) 和 (5.53) 是相通的两种插值公式 (验算留作习题)。

插值问题的推广:

在第三章中 (见第 46 页) 讨论的广义插值问题, 即除了要求 $f(x)$ 在给定位置取某些给定值之外, 还要求它在给定位置有特定变率、特定曲率等等, 我们是用插值公式先行求得一个在前述给定位置取给定值的 $g(x)$, 然后用它来简化待求公式 $f(x)$ 的样子。因为 $f(x) - g(x)$ 已经在前述的给定位置取零为值, 所以它必定具有某些特定的因式, 由此再用给定的变率、曲率等值去确定余下的未知数的值。

从上述用行列式的解题方法可以看到, 其实我们不必分开处理给定值、给定变率、给定曲率等等的条件, 只需对应于每一个给定条件写

下一个条件式，然後要求所得的条件方程组具有非零的系数解，亦即其系数方阵的行列式等于 0，就可以写下一个 $f(x)$ 的公式。例如，求一个次数不高于 4 的多项式 $f(x)$ ，它满足

$$f(-1) = 9, Df(-1) = -6, f(1) = 1, Df(1) = 6, f(2) = 27$$

设 $f(x)$ 的样子是 $-y + c_0 + c_1x + c_2x^2 + c_3x^3 + c_4x^4 = 0$ ，则 $Df(x)$ 的样子就是 $-Dy + c_1 + 2c_2x + 3c_3x^2 + 4c_4x^3 = 0$ 。即有下述条件方程组：

$$(5.56) \quad \begin{pmatrix} -y & 1 & x & x^2 & x^3 & x^4 \\ -9 & 1 & -1 & (-1)^2 & (-1)^3 & (-1)^4 \\ 6 & 0 & 1 & 2(-1) & 3(-1)^2 & 4(-1)^3 \\ -1 & 1 & 1 & 1^2 & 1^3 & 1^4 \\ -6 & 0 & 1 & 2(1) & 3(1)^2 & 4(1)^3 \\ -27 & 1 & 2 & 2^2 & 2^3 & 2^4 \end{pmatrix} \begin{pmatrix} -1 \\ c_0 \\ c_1 \\ c_2 \\ c_3 \\ c_4 \end{pmatrix} = \mathbf{0}$$

上述方程组具有非零解 $(-1, c_0, c_1, c_2, c_3, c_4)$ 的充要条件是其系数行列式等于 0，亦即

$$(5.57) \quad \begin{vmatrix} -y & 1 & x & x^2 & x^3 & x^4 \\ -9 & 1 & -1 & (-1)^2 & (-1)^3 & (-1)^4 \\ 6 & 0 & 1 & 2(-1) & 3(-1)^2 & 4(-1)^3 \\ -1 & 1 & 1 & 1^2 & 1^3 & 1^4 \\ -6 & 0 & 1 & 2(1) & 3(1)^2 & 4(1)^3 \\ -27 & 1 & 2 & 2^2 & 2^3 & 2^4 \end{vmatrix} = 0$$

把上述行列式展开後即有

$$(5.58) \quad -y + 3 - 6x + x^2 + 2x^3 + x^4 = 0$$

亦即 $f(x) = 3 - 6x + x^2 + 2x^3 + x^4$ 。

特别地，当所有条件式都集中在某一个特定位置时，例如

$$(5.59) \quad f(0) = v_0, Df(0) = v_1, \dots, D^n f(0) = v_n$$

则对应的行列式条件式就是

$$(5.60) \quad \begin{vmatrix} -y & 1 & x & x^2 & \dots & x^n \\ -v_0 & 1 & 0 & 0^2 & \dots & 0^n \\ -v_1 & 0 & 1 & 2(0) & \dots & n(0)^{n-1} \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ -v_j & 0 & \dots & j! & \dots & \frac{n!}{(n-j)!}(0)^{n-j} \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ -v_n & 0 & 0 & \dots & 0 & n! \end{vmatrix} = 0$$

把上述行列式展开後所得者其实就是泰勒公式。

5.5.3 圆锥截线交点与求解四次方程

在一般位置的圆锥截线 (conic sections) 的坐标方程可以写成下述二次曲线的方程

$$Ax^2 + 2Bxy + Cy^2 + 2Dx + 2Ey + F = 0$$

因此，求解任给两个圆锥截线的交点一般来说便涉及求解四次方程，所以并不是一件容易的工作。再者，怎样求解四次方程呢？下面将会讨论用行列式来把上述问题简化为求解三次方程，因为三次方程的公式解是可以明确地写下来。

我们先讨论怎样解一般的三次方程。设 $x^3 + bx^2 + cx + d = 0$ 为任给的三次方程。先作代换 $y = x + \frac{b}{3}$ ，简化後得

$$(5.61) \quad y^3 + py + q = 0, \quad \text{其中} \quad p = c - \frac{b^2}{3}, \quad q = d + \frac{2b^3}{3^3} - \frac{bc}{3}$$

再代入

$$(5.62) \quad y = u + v, \quad -p = 3uv$$

展开後得出 $u^3 + v^3 + q = 0$ 。整式遍乘以 u^3 ，再代入假设 $-p = 3uv$ 便得出

$$(5.63) \quad (u^3)^2 + q(u^3) - \frac{p^3}{27} = 0$$

因此可用二次方程式解写下 u^3 的解。引进复数 $\omega = e^{\frac{2\pi i}{3}}$ (即 $\omega^3 = 1$)，便可以将 y 的解 (在复数系内) 明确地写下来：

$$\begin{aligned} y_1 &= \sqrt[3]{-\frac{q}{2} + \sqrt{\frac{q^2}{4} + \frac{p^3}{27}}} + \sqrt[3]{-\frac{q}{2} - \sqrt{\frac{q^2}{4} + \frac{p^3}{27}}} \\ y_2 &= \sqrt[3]{-\frac{q}{2} + \sqrt{\frac{q^2}{4} + \frac{p^3}{27}}} \omega + \sqrt[3]{-\frac{q}{2} - \sqrt{\frac{q^2}{4} + \frac{p^3}{27}}} \omega^2 \\ y_3 &= \sqrt[3]{-\frac{q}{2} + \sqrt{\frac{q^2}{4} + \frac{p^3}{27}}} \omega^2 + \sqrt[3]{-\frac{q}{2} - \sqrt{\frac{q^2}{4} + \frac{p^3}{27}}} \omega \end{aligned}$$

下述引理就是求解圆锥截线交点的要点所在：

【引理】：设 $f(x, y) = Ax^2 + 2Bxy + Cy^2 + 2Dx + 2Ey + F$ ，则 $f(x, y)$ 可被分解为两个线性因式的乘积，亦即

$$(5.64) \quad f(x, y) \equiv \ell_1(x, y) \cdot \ell_2(x, y)$$

的充要条件就是

$$(5.65) \quad \begin{vmatrix} A & B & D \\ B & C & E \\ D & E & F \end{vmatrix} = 0$$

证明： $A = C = 0$ 的情形是容易处理的 (留作习题)。现在不妨假设 $A = 1$ 。我们将用待定系数法寻找条件使得

$$(5.66) \quad f(x, y) \equiv (x + \alpha_1 y + \beta_1)(x + \alpha_2 y + \beta_2)$$

比较恒等式两边的 xy 和 y^2 的系数，得

$$\alpha_1 + \alpha_2 = 2B$$

$$\alpha_1 \cdot \alpha_2 = C$$

由此解出 $\alpha_1 = B + \sqrt{B^2 - C}$, $\alpha_2 = B - \sqrt{B^2 - C}$ 。再比较 x 的系数和常数项，得

$$\beta_1 + \beta_2 = 2D$$

$$\beta_1 \cdot \beta_2 = F$$

由此解出 $\beta_1 = D + \sqrt{D^2 - F}$, $\beta_2 = D - \sqrt{D^2 - F}$ 。然後比较 y 的系数得

$$(5.67) \quad \alpha_1\beta_2 + \alpha_2\beta_1 = 2E$$

代入先前已得的 $\alpha_1, \alpha_2, \beta_1, \beta_2$ 表达式, 则

$$\begin{aligned} 2E &= (B + \sqrt{B^2 - C})(D - \sqrt{D^2 - F}) + (B - \sqrt{B^2 - C})(D + \sqrt{D^2 - F}) \\ &= BD - B\sqrt{D^2 - F} + D\sqrt{B^2 - C} - \sqrt{B^2 - C}\sqrt{D^2 - F} \\ &\quad + BD + B\sqrt{D^2 - F} - D\sqrt{B^2 - C} - \sqrt{B^2 - C}\sqrt{D^2 - F} \\ &= 2BD - 2\sqrt{B^2 - C}\sqrt{D^2 - F} \end{aligned}$$

从而

$$\begin{aligned} (B^2 - C)(D^2 - F) &= (E - BD)^2 \\ B^2D^2 - B^2F - CD^2 + CF &= E^2 - 2BDE + B^2D^2 \\ CF + 2BDE - CD^2 - E^2 - B^2F &= 0 \end{aligned}$$

上述条件其实就是

$$(5.68) \quad \begin{vmatrix} 1 & B & D \\ B & C & E \\ D & E & F \end{vmatrix} = 0 \quad \square$$

现假设 $f_1(x, y) = A_1x^2 + 2B_1xy + C_1y^2 + 2D_1x + 2E_1y + F_1 = 0$ 和 $f_2(x, y) = A_2x^2 + 2B_2xy + C_2y^2 + 2D_2x + 2E_2y + F_2 = 0$ 为两个圆锥截线的方程, 而在一般情况下它们会有四个交点, 而所有过这四个交点的圆锥截线的方程 (除 $f_2(x, y) = 0$ 外) 可以写成

$$f_1(x, y) + kf_2(x, y) = 0$$

亦即

$$\begin{aligned} (A_1 + kA_2)x^2 + 2(B_1 + kB_2)xy + (C_1 + kC_2)y^2 \\ + 2(D_1 + kD_2)x + 2(E_1 + kE_2)y + (F_1 + kF_2) = 0 \end{aligned}$$

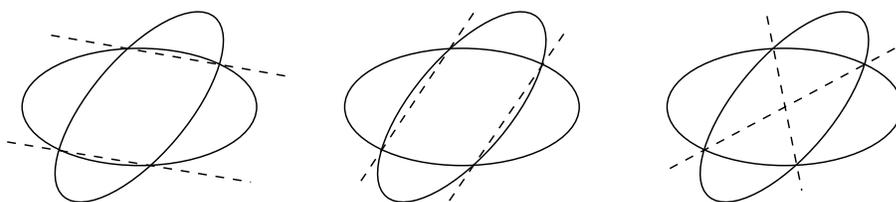
如果上述方程是代表著蜕化圆锥截线的情形, 即可被分解成两个直线方程

$$\ell_1(x, y) \cdot \ell_2(x, y) = 0$$

则求交点的问题便简化成求上述两直线与圆锥截线 $f_1(x, y) = 0$ (或 $f_2(x, y) = 0$) 的交点, 这是很常见易解的二次方程问题。所以, 我们希望能找出适当的 k 使得该方程就是代表著蜕化圆锥截线。由刚刚所证的引理知道, k 必需满足

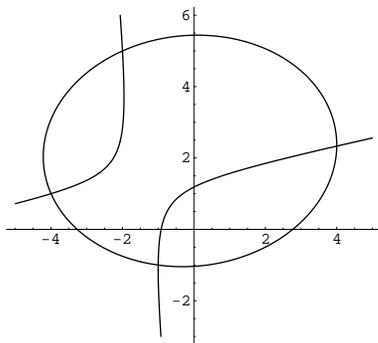
$$\begin{vmatrix} A_1 + kA_2 & B_1 + kB_2 & D_1 + kD_2 \\ B_1 + kB_2 & C_1 + kC_2 & E_1 + kE_2 \\ D_1 + kD_2 & E_1 + kE_2 & F_1 + kF_2 \end{vmatrix} = 0$$

这一个三次方程。运用三次方程公式解, 我们求得 k 的三个值使得 $f_1(x, y) + k_1 f_2(x, y) = 0$ 、 $f_1(x, y) + k_2 f_2(x, y) = 0$ 、 $f_1(x, y) + k_3 f_2(x, y) = 0$ 都代表著蜕化圆锥截线。其实这三个情况分别就是



[图 5-6]

【例子】：令 $f_1(x, y) = 86x^2 - 11xy + 138y^2 + 43x - 608y - 778 = 0$ (椭圆) 和 $f_2(x, y) = 26x^2 - 121xy - 12y^2 + 233x - 148y + 192 = 0$ (双曲线) 为两个圆锥截线的方程, 求两者的四个交点的坐标。



[图 5-7]

解：先求出适当的 k 使得 $f_1(x, y) + k f_2(x, y) = 0$ 为蜕化圆锥截线，

亦即求解方程

$$\begin{vmatrix} A_1 + kA_2 & B_1 + kB_2 & D_1 + kD_2 \\ B_1 + kB_2 & C_1 + kC_2 & E_1 + kE_2 \\ D_1 + kD_2 & E_1 + kE_2 & F_1 + kF_2 \end{vmatrix} \\ = \begin{vmatrix} 86 + 26k & -\frac{11}{2} - \frac{121}{2}k & \frac{43}{2} + \frac{233}{2}k \\ -\frac{11}{2} - \frac{121}{2}k & 138 - 12k & -304 - 74k \\ \frac{43}{2} + \frac{233}{2}k & -304 - 74k & -778 + 192k \end{vmatrix} = 0$$

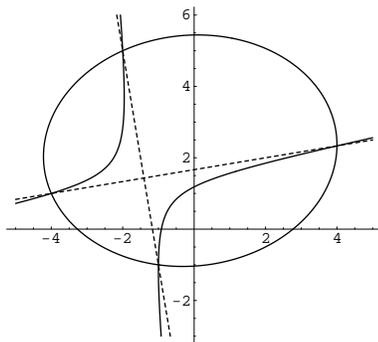
展开上述行列式，得出

$$5280(57k^3 + 866k^2 - 939k - 3248) = 0$$

其中一个解为 $k = -16$ (另外两解为 $k = -\frac{29}{19}$, $k = \frac{7}{3}$)。因此，前述引理保证了 $f_1(x, y) - 16f_2(x, y)$ 必可分解，亦即：

$$\begin{aligned} f_1(x, y) - 16f_2(x, y) &= -55(6x^2 - 35xy - 6y^2 + 67x - 32y + 70) \\ &= -55(x - 6y + 10)(6x + y + 7) \end{aligned}$$

而对应的图象为



[图 5-8]

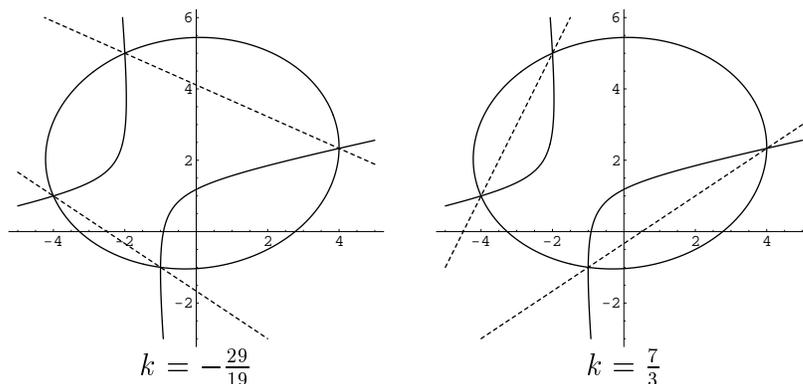
用 $f_1(x, y) = 86x^2 - 11xy + 138y^2 + 43x - 608y - 778 = 0$ 和 $x - 6y + 10 = 0$ 可以解出其中两个交点为

$$(-4, 1), \quad \left(4, \frac{7}{3}\right)$$

而用 $f_1(x, y) = 0$ 和 $6x + y + 7 = 0$ 可以解出另外两个交点为

$$(-2, 5), \quad (-1, -1)$$

[注]: 当 k 为 $-\frac{29}{19}$ 和 $k = \frac{7}{3}$ 时图象分别为



[图 5-9]

求解四次方程:

一般的四次方程可写成 $x^4 + \alpha x^3 + \beta x^2 + \gamma x + \delta = 0$ 。把上式遍除以 x^2 , 然后把 x 的负指数项代入 $y = \frac{1}{x}$, 即得 $x^2 + \alpha x + \beta + \gamma y + \delta y^2 = 0$ 。所以, 上述求解四次方程问题便转化成求解下述联立方程:

$$\begin{cases} f_1(x, y) = x^2 + \delta y^2 + \alpha x + \gamma y + \beta = 0 \\ f_2(x, y) = xy - 1 = 0 \end{cases}$$

亦即求解上式中的两个二次曲线的四个相交点之 x -坐标。

这样, 我们可以用前述找圆锥截线交点时所引入的方法, 先用行列式找出适当的 k 使得

$$f_1(x, y) + k f_2(x, y) = x^2 + \delta y^2 + \alpha x + \gamma y + \beta + k(xy - 1) = \ell_1(x, y) \cdot \ell_2(x, y)$$

为代表蜕化二次曲线的方程, 亦即

$$\begin{vmatrix} 1 & \frac{k}{2} & \frac{\alpha}{2} \\ \frac{k}{2} & \delta & \frac{\gamma}{2} \\ \frac{\alpha}{2} & \frac{\gamma}{2} & \beta - k \end{vmatrix} = \frac{1}{4}(k^3 - \beta k^2 + (\alpha\gamma - 4\delta)k + (4\beta\delta - \alpha^2\delta - \gamma^2)) = 0$$

然后再分别用 $\ell_1(x, y)$ 和 $xy - 1 = 0$ 及 $\ell_2(x, y)$ 和 $xy - 1 = 0$ 求得四个交点之 x -坐标, 亦即四次方程 $x^4 + \alpha x^3 + \beta x^2 + \gamma x + \delta = 0$ 的四个解。

【例子】: 求解 $x^4 + 2x^3 - 5x^2 + 6x + 2 = 0$ 。

上述方程所对应的联立二次曲线方程乃是

$$\begin{cases} f_1(x, y) = x^2 + 2y^2 + 2x + 6y - 5 = 0 \\ f_2(x, y) = xy - 1 = 0 \end{cases}$$

而使得 $f_1(x, y) + kf_2(x, y)$ 为代表蜕化二次曲线方程的 k 所满足的方程为

$$\begin{vmatrix} 1 & \frac{k}{2} & 1 \\ \frac{k}{2} & 2 & 3 \\ 1 & 3 & -5 - k \end{vmatrix} = \frac{k^3}{4} + \frac{5k^2}{4} + k - 21 = 0$$

上述三次方程的其中一个解为 $k = 3$ (另外两个解为 $k = -4 + 2\sqrt{3}i$ 和 $k = -4 - 2\sqrt{3}i$)。所以, 前述引理保证了 $f_1(x, y) + 3f_2(x, y)$ 必可分解, 亦即

$$\begin{aligned} f_1(x, y) + 3f_2(x, y) &= x^2 + 3xy + 2y^2 + 2x + 6y - 8 \\ &= (x + y + 4)(x + 2y - 2) \end{aligned}$$

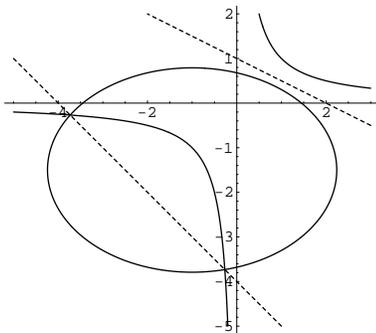
用 $x + y + 4 = 0$ 和 $xy - 1 = 0$ 可以解出其中两个交点的 x -坐标为

$$-2 - \sqrt{3}, \quad -2 + \sqrt{3}$$

而用 $x + 2y - 2 = 0$ 和 $xy - 1 = 0$ 可以解出另外两个交点的 x -坐标为

$$1 - i, \quad 1 + i$$

(即在实平面上没有交点。)



[图 5-10]

附录： 域上的线性空间与域的代数扩张

位移向量的加： $\mathbf{a} + \mathbf{b}$ 和 倍积： $k \cdot \mathbf{a}$ 直接反映著空间的线性结构，而且其所满足的运算律如加的交换律： $\mathbf{a} + \mathbf{b} = \mathbf{b} + \mathbf{a}$ 和 倍积的分配律： $k \cdot (\mathbf{a} + \mathbf{b}) = k \cdot \mathbf{a} + k \cdot \mathbf{b}$ 等则又能简明扼要地掌握了空间在平行和相似上的基本定理，把它们转化成简朴易用、直截了当的代数计算。这种由加和倍积结合而成的代数体系是十分好算易用。其实，所用到的计算就是用分配律展开和集项，而且这种代数又具有简朴的几何直观内涵，那就是空间的线性结构。概括地说，线性代数就是加和倍积的分配律的代数，而这种代数体系也就是空间线性结构的代数化表达形式。总之，它是融合几何和代数的精简体系，既易算好用又富有几何直观性，是数学中最为简朴而且广泛有用的精华，是基础数学重要基石所在。

I. 系数域和给定域上的向量空间

域 (field) 的概念其实就是将我们常用的数系四则运算推广至更广泛类型的代数结构。简单来说，域就是可以让我们如常进行加减乘除运算的代数结构。一个域 F 的定义可写成以下六点：

1. 存在加、乘运算（ $+$ 和 \times ），可从任给的 $a, b \in F$ 唯一地确定 $a + b \in F$ 和 $a \times b \in F$ 。
2. 加、乘运算满足结合律和交换律。
3. 存在 $0, 1 \in F$ ($0 \neq 1$) 使得对于任意 $a \in F$ ，有 $0 + a = a$ 和 $1 \cdot a = a$ 。
4. 对于任意 $a \in F$ ，存在 $-a \in F$ 使得 $a + (-a) = 0$ 。

5. 对于任意 $a \in F, a \neq 0$, 存在 a^{-1} 使得 $a \times a^{-1} = 1$ 。

6. 满足乘对于加的分配律： $a \times (b + c) = (a \times b) + (a \times c)$ 。

常见常用的域有 \mathbb{Q} (有理数域)、 \mathbb{R} (实数域) 和 \mathbb{C} (复数域)；而其他常用的域有：

$$(1) \mathbb{Q}(\sqrt{2}) = \{a + b\sqrt{2} : a, b \in \mathbb{Q}\}$$

加减乘除等同实数域的运算，即

$$\text{加减法：}(a + b\sqrt{2}) \pm (c + d\sqrt{2}) = (a \pm c) + (b \pm d)\sqrt{2}。$$

$$\text{乘法：}(a + b\sqrt{2}) \cdot (c + d\sqrt{2}) = (ac + 2bd) + (bc + ad)\sqrt{2}。$$

值得注意的是除法运算仍是「封闭」的，因除法可利用

$$(a + b\sqrt{2}) \cdot \left(\frac{a}{a^2 - 2b^2} - \frac{b}{a^2 - 2b^2}\sqrt{2} \right) = 1$$

来定义倒数 $(a + b\sqrt{2})^{-1}$ 之后再加以定义。

$$(2) \mathbb{Q}(\sqrt[3]{2}) = \{a + b\sqrt[3]{2} + c\sqrt[3]{2^2} : a, b, c \in \mathbb{Q}\}$$

同样地，加减乘除等同实数域的运算。而除法是封闭的验证我们将以一个实例来表达一般做法：

如取 $\alpha = 1 + 2\sqrt[3]{2} - \sqrt[3]{2^2}$ 。由于 $x^3 - 2$ 是没有有理数根，所以 $x^3 - 2$ 是在有理系数之下不可约的多项式，亦即它和 $1 + 2x - x^2$ 的最高公因式为 1。运用辗转相除法，我们求得

$$\begin{aligned} x^3 - 2 &= -(x + 2)(1 + 2x - x^2) + 5x \\ 1 + 2x - x^2 &= \frac{2 - x}{5}(5x) + 1 \end{aligned}$$

然後利用所得等式将最高公因式 1 重写成 $x^3 - 2$ 和 $1 + 2x - x^2$ 的组合：

$$\begin{aligned} 1 &= (1 + 2x - x^2) - \frac{2 - x}{5}(5x) \\ &= (1 + 2x - x^2) - \frac{2 - x}{5}((x^3 - 2) + (x + 2)(1 + 2x - x^2)) \\ &= (1 - (x + 2)\left(\frac{2 - x}{5}\right))(1 + 2x - x^2) - \frac{2 - x}{5}(x^3 - 2) \end{aligned}$$

代入 $x = \sqrt[3]{2}$ 得

$$1 = (1 - (\sqrt[3]{2} + 2)\left(\frac{2 - \sqrt[3]{2}}{5}\right))\alpha = \frac{1 + \sqrt[3]{2^2}}{5}\alpha$$

$$\text{所以 } \alpha^{-1} = \frac{1 + \sqrt[3]{2^2}}{5}.$$

【定义】：设 F 是一个给定的域。一个满足下面的加法和 F -倍积运算律的代数结构 $V = (V, +, \cdot)$ 称为一个在 F 上的向量空间，而 F 称为 V 的系数域，简称为「 V/F 」。即对于任给元素 $\mathbf{a}, \mathbf{b} \in V$ 和 $\lambda \in F$ ，唯一确定元素 $\mathbf{a} + \mathbf{b} \in V$ 和 $\lambda \cdot \mathbf{a} \in V$ 并满足向量运算的基本性质：

$$\text{加：} \begin{cases} \text{结合律：} \mathbf{a} + (\mathbf{b} + \mathbf{c}) = (\mathbf{a} + \mathbf{b}) + \mathbf{c} \\ \text{交换律：} \mathbf{a} + \mathbf{b} = \mathbf{b} + \mathbf{a} \\ \text{可逆性：对于任给 } \mathbf{a}, \mathbf{b} \in V, \text{ 唯一存在 } \mathbf{x} \text{ 使得 } \mathbf{a} + \mathbf{x} = \mathbf{b} \end{cases}$$

$$\text{倍积：} \begin{cases} 1 \cdot \mathbf{a} = \mathbf{a}, (\lambda + \mu) \cdot \mathbf{a} = \lambda \cdot \mathbf{a} + \mu \cdot \mathbf{a}, \\ \lambda \cdot (\mathbf{a} + \mathbf{b}) = \lambda \cdot \mathbf{a} + \lambda \cdot \mathbf{b}, \\ \lambda \cdot (\mu \cdot \mathbf{a}) = (\lambda \cdot \mu) \cdot \mathbf{a} \end{cases}$$

【例子】：

- (1) 当然，在空间内的位移向量和其加、倍积运算组成 \mathbb{R} 上的一个向量空间。
- (2) 复数域 \mathbb{C} 包含实数域 \mathbb{R} ，我们可以取其加法和其乘法中只限于 $\lambda \in \mathbb{R}$ 和 $\alpha \in \mathbb{C}$ 相乘（看成倍积），则可以把 \mathbb{C} 看成是 \mathbb{R} 上的一个向量空间。
- (3) 设以 $F[x]$ 表示以 F 中元数为系数的单元多项式所组成的集合，则其多项式加法和其乘法中只限于 $\lambda \in F$ 和 $f(x) \in F[x]$ 的相乘（看成倍积）显然满足上述向量空间的运算律，所以也是一个 F 上的向量空间的实例。
- (4) 设以 $F_n[x]$ 表示在 $F[x]$ 中次数不高于 n 的多项式集合（包括零多项式）。在多项式加法和上述局限于 $\lambda \in F$ 和 $f(x) \in F_n[x]$ 的乘法之下， $F_n[x]$ 也是一个 F 上的向量空间。

(5) 系数取值于 F 的元素的 $m \times n$ -矩阵在其加法与倍积之下构成一个 F 上的向量空间。通常我们用 F^n 表示 $n \times 1$ -矩阵所组成者，以 $(F^n)^*$ 表示由 $1 \times n$ -矩阵所组成者。

(6) 设 α 为一个在有理系数中不可约 n 次多项式的根。则

$$\mathbb{Q}(\alpha) = \{c_0 + c_1\alpha + \cdots + c_{n-1}\alpha^{n-1} : c_0, c_1, \cdots, c_{n-1} \in \mathbb{Q}\}$$

为 \mathbb{Q} 上的向量空间（其实它也是一个域）。

II. 线性组合、子空间和其生成系

向量空间所具有的「加」与「倍积」这两种运算，合称之为线性运算（所以向量空间也常常称为线性空间）。从一组给定向量 $\mathbf{a}_1, \mathbf{a}_2, \cdots, \mathbf{a}_n$ 出发，重复地运用线性运算，所得的结果都是能写成下述形式的元素，即

$$\lambda_1 \cdot \mathbf{a}_1 + \lambda_2 \cdot \mathbf{a}_2 + \cdots + \lambda_n \cdot \mathbf{a}_n, \quad \lambda_i \in F, 1 \leq i \leq n$$

所有能写成上述形式的元素都叫做是 $\mathbf{a}_1, \mathbf{a}_2, \cdots, \mathbf{a}_n$ 的线性组合 (linear combination)，例如

$$\mathbf{a}_1 = 1 \cdot \mathbf{a}_1 + 0 \cdot \mathbf{a}_2 + \cdots + 0 \cdot \mathbf{a}_n$$

$$\mathbf{a}_n = 0 \cdot \mathbf{a}_1 + 0 \cdot \mathbf{a}_2 + \cdots + 1 \cdot \mathbf{a}_n$$

$$\mathbf{x} = \mathbf{a}_1 - \mathbf{a}_2 + \mathbf{a}_3 + \cdots + (-1)^n \cdot \mathbf{a}_n$$

等等都是 $\mathbf{a}_1, \mathbf{a}_2, \cdots, \mathbf{a}_n$ 这一组向量的线性组合。再者，设

$$\mathbf{b} = \sum_{i=1}^n \lambda_i \cdot \mathbf{a}_i, \quad \mathbf{c} = \sum_{i=1}^n \mu_i \cdot \mathbf{a}_i$$

则不难用向量空间的运算律得出

$$\lambda \cdot \mathbf{b} + \mu \cdot \mathbf{c} = \lambda \cdot \sum_{i=1}^n \lambda_i \cdot \mathbf{a}_i + \mu \cdot \sum_{i=1}^n \mu_i \cdot \mathbf{a}_i = \sum_{i=1}^n (\lambda \lambda_i + \mu \mu_i) \cdot \mathbf{a}_i$$

所以 $\lambda \cdot \mathbf{b} + \mu \cdot \mathbf{c}$ 对于任给 $\lambda, \mu \in F$ 都又是 $\mathbf{a}_1, \mathbf{a}_2, \cdots, \mathbf{a}_n$ 的线性组合。

【子空间与生成系】：设 U 是向量空间 V 的一个子集，若满足条件：

$$\mathbf{a}, \mathbf{b} \in U \Rightarrow \lambda \cdot \mathbf{a} + \mu \cdot \mathbf{b} \in U, \quad \text{对于任何 } \lambda, \mu \in F \text{ 皆成立}$$

则称 U 为 V 的一个子空间 (subspace)。

再者，设 $\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n$ 是子空间 U 的一组元素，若 U 中任何元素都可以表成 $\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n$ 的一个线性组合，则称 $\{\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n\}$ 为子空间 U 的一组生成系 (generator system)。

【例子】：

- (1) $\{0\}$ 和 V 本身都是 V 的子空间。
- (2) 前述的不高于 n 次多项式集 $F_n[x]$ 是所有多项式集 $F[x]$ 的子空间，而 $\{1, x, x^2, \dots, x^n\}$ 则构成它的一组生成系。
- (3) $\{1, i\}$ 构成 \mathbb{C}/\mathbb{R} 的一组生成系。
- (4) 设有三个域 E, M, F 具有关系 $E \supset M \supset F$ ，则在把 E 看成 F -向量空间的观点之下， M 是 E 的子空间。如 $\mathbb{C} \supset \mathbb{R} \supset \mathbb{Q}$ 就是一个例子。
- (5) 设 $\{\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n\} \subset V$ ，令

$$\langle \mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n \rangle = \left\{ \sum_{i=1}^n \lambda_i \cdot \mathbf{a}_i; \lambda_i \in F \right\}$$

是所有 $\{\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n\}$ 的线性组合所组成的子集，则它是 V 的一个子空间，而 $\{\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n\}$ 就当然是它的一组生成系。我们称它为由 $\{\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n\}$ 所张的子空间。

设 $\{\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n\}$ 是子空间 U 的一组生成系，一个很自然的问题是：这一组生成系是否有多余的元素可去掉？假如 \mathbf{a}_n 是多余的，即 $\{\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_{n-1}\}$ 依然是 U 的一组生成系，则 \mathbf{a}_n 这个属于 U 的元素当然要能表成 $\{\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_{n-1}\}$ 的一个线性组合。反之，设 \mathbf{a}_n 能表成 $\{\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_{n-1}\}$ 的线性组合，即有

$$\mathbf{a}_n = \alpha_1 \cdot \mathbf{a}_1 + \dots + \alpha_{n-1} \cdot \mathbf{a}_{n-1}, \quad \alpha_1, \dots, \alpha_{n-1} \in F$$

则任何 $\mathbf{a}_1, \dots, \mathbf{a}_{n-1}, \mathbf{a}_n$ 的线性组合

$$\sum_{i=1}^n \lambda_i \mathbf{a}_i, \quad \lambda_1, \dots, \lambda_{n-1}, \lambda_n \in F$$

都可以改写成

$$\sum_{i=1}^n \lambda_i \mathbf{a}_i = \sum_{i=1}^{n-1} \lambda_i \mathbf{a}_i + \lambda_n \cdot \left(\sum_{i=1}^{n-1} \alpha_i \mathbf{a}_i \right) = \sum_{i=1}^{n-1} (\lambda_i + \lambda_n \alpha_i) \mathbf{a}_i$$

所以 $\{\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_{n-1}\}$ 就已经组成 U 的一组生成系了，亦即 \mathbf{a}_n 是多馀的。

【定义】：一组向量 $\{\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n\}$ 之中若有一个向量可以表达成其余 $(n-1)$ 个向量的线性组合，则称 $\{\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n\}$ 为线性相关 (linearly dependent)；反之，则称之为线性无关 (linearly independent)。

设 U 的一组生成系 $\{\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n\}$ 中有可供去掉者，则称 $\{\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n\}$ 为可约生成系；反之，则称之为不可约生成系。

【定理】：

- (i) $\{\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n\}$ 为线性相关的充要条件是存在一组不全为 0 的 $\alpha_i \in F, 1 \leq i \leq n$ ，使得 $\sum_{i=1}^n \alpha_i \mathbf{a}_i = \mathbf{0}$ 。
- (ii) $\{\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n\}$ 为线性无关的充要条件是 $\sum_{i=1}^n \alpha_i \mathbf{a}_i = \mathbf{0}$ 只有在 α_i 全部为 0 的情形才成立！
- (iii) 设 $\{\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n\}$ 为 U 的一组生成系，则它是可约（或不可约）的充要条件是 $\{\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n\}$ 线性相关（或线性无关）。

证明：上述命题中，(iii) 是定义的直接推论，而 (i) \Rightarrow (ii) 也是显然的，所以我们只需证明 (i) 如下：

设 $\{\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n\}$ 线性相关，即存在其中之一元素， \mathbf{a}_{i_0} ，它是其余 $(n-1)$ 个元素的线性组合，亦即存在 $\alpha_i \in F, 1 \leq i \leq n, i \neq i_0$ ，使得

$$\mathbf{a}_{i_0} = \sum_{i \neq i_0} \alpha_i \mathbf{a}_i$$

令 $\alpha_{i_0} = -1$ ，即有

$$\sum_{i=1}^n \alpha_i \mathbf{a}_i = \sum_{i \neq i_0} \alpha_i \mathbf{a}_i - \mathbf{a}_{i_0} = \mathbf{0}$$

其中至少有 $\alpha_{i_0} \neq 0$ 。反之，设有全不为 0 的 $\alpha_i \in F$ ， $1 \leq i \leq n$ ，使得

$$\sum_{i=1}^n \alpha_i \mathbf{a}_i = \mathbf{0}$$

设 $\alpha_{i_0} \neq 0$ ，则可将 $\alpha_{i_0} \mathbf{a}_{i_0}$ 移项，然後遍乘以 $-\frac{1}{\alpha_{i_0}}$ ，即得

$$\alpha_{i_0} = -\frac{1}{\alpha_{i_0}} \sum_{i \neq i_0} \alpha_i \cdot \mathbf{a}_i = \sum_{i \neq i_0} \left(-\frac{\alpha_i}{\alpha_{i_0}}\right) \cdot \mathbf{a}_i$$

所以 $\{\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n\}$ 是线性相关的。 \square

【推论】：设 $\{\mathbf{a}_1, \dots, \mathbf{a}_{n-1}\}$ 线性无关，但是 $\{\mathbf{a}_1, \dots, \mathbf{a}_{n-1}, \mathbf{a}_n\}$ 则转变成线性相关，则 \mathbf{a}_n 必定是一个 $\mathbf{a}_1, \dots, \mathbf{a}_{n-1}$ 的线性组合。

证明：因为 $\{\mathbf{a}_1, \dots, \mathbf{a}_{n-1}, \mathbf{a}_n\}$ 是线性相关的，所以存在不全为 0 的 $\alpha_i \in F$ ， $i = 1, \dots, n$ ，使得 $\sum_{i=1}^n \alpha_i \mathbf{a}_i = \mathbf{0}$ ，其中 α_n 必须不等于 0。要不然，即 $\alpha_n = 0$ ；则 $\alpha_1, \dots, \alpha_{n-1}$ 之中必有不为 0 者，而且 $\sum_{i=1}^n \alpha_i \mathbf{a}_i = \sum_{i=1}^{n-1} \alpha_i \mathbf{a}_i + 0 \cdot \mathbf{a}_n = \sum_{i=1}^{n-1} \alpha_i \mathbf{a}_i = \mathbf{0}$ ，这是和假设 $\{\mathbf{a}_1, \dots, \mathbf{a}_{n-1}\}$ 线性无关相矛盾的，因此 α_n 必须不为 0，即有

$$\mathbf{a}_n = \sum_{i=1}^{n-1} \left(-\frac{\alpha_i}{\alpha_n}\right) \cdot \mathbf{a}_i$$

【例子】：

(1) $\{1, x, x^2, \dots, x^n\}$ 是线性无关的多项式。

(2) $g_0(x) = 1$, $g_k(x) = \frac{1}{k!} x(x-1)\cdots(x-k+1)$, $k = 1, 2, \dots, n$ 也是 $(n+1)$ 个线性无关的多项式。

(3) 设 $\{1, x, x^2, \dots, x^n, f(x)\}$ 是线性相关的多项式，则 $f(x)$ 必定是一个次数不高于 n 的多项式。

(4) $\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n$ 在 F^n 或 $(F^n)^*$ 中是线性无关的。

III. 基底和维数

从向量空间的代数结构来看，我们可以在一个给定的向量空间 V 中选取一个不可约生成系，以后称之为「基底」， $\{\mathbf{b}_i; 1 \leq i \leq m\}$ ，这样就可以把 V 中的任一元素唯一地表成 $\{\mathbf{b}_i\}$ 的线性组合，从而把 V 中的向量运算，归于线性组合的系数之间的运算，这也就是向量空间的「坐标化」。一组选定的基底也就是 V 中一个选定的坐标系，而选取坐标系（亦即基底）的好处是可以利用它来简洁地把向量空间的结构「数量化」；但是上述坐标系的选定是含有很大的「任意性」的，所以坐标系本身并非向量空间的结构本质。因此，在采取坐标化的方法研究向量空间时，一定要牢牢记著，只有那种不依赖于特殊坐标系而取定的事物，亦即在任何坐标系中均普遍成立的事物，才是向量空间的本质性事物。例如我们将要证明的：「 V 中任一基底中所含的元素个数相等」便是 V 的结构不变量，以后叫做 V 的「维数」。

【定理】：设 $\{\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n\}$ 是子空间 U 的一组生成系而 $\{\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_m\}$ 则是 U 中的一组线性无关向量组，则 $n \geq m$ ，而且存在适当的 $\{\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n\}$ 中恰好含有 m 个元素的子集 A_m ，使得

$$\{\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_m\} \cup (\{\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n\} \setminus A_m)$$

构成 U 的另一组生成系。

证明：我们将对 m 作归纳法来证明上述定理，先证 $m = 1$ 的情形。由假设， \mathbf{b}_1 是 $\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n$ 的一个线性组合，即

$$\mathbf{b}_1 = \sum_{i=1}^n \beta_i \mathbf{a}_i$$

其中 β_i 不可能全为 0。设 $\beta_{i_1} \neq 0$ ，则可由上式解得

$$\mathbf{a}_{i_1} = -\frac{1}{\beta_{i_1}} \cdot \left(\sum_{i \neq i_1} \beta_i \mathbf{a}_i - \mathbf{b}_1 \right)$$

所以

$$\{\mathbf{b}_1\} \cup (\{\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n\} \setminus \{\mathbf{a}_{i_1}\})$$

当然构成一组 U 的生成系，因为 U 中的任何一个元素都可以写成

$$\begin{aligned}\sum_{i=1}^n \lambda_i \mathbf{a}_i &= \sum_{i \neq i_1} \lambda_i \mathbf{a}_i + \left(-\frac{\lambda_{i_1}}{\beta_{i_1}}\right) \left(\sum_{i \neq i_1} \beta_i \mathbf{a}_i - \mathbf{b}_1\right) \\ &= \frac{\lambda_{i_1}}{\beta_{i_1}} \mathbf{b}_1 + \sum_{i \neq i_1} \left(\lambda_i - \frac{\lambda_{i_1}}{\beta_{i_1}} \beta_i\right) \cdot \mathbf{a}_i\end{aligned}$$

现在再由归纳假设上述定理对于 $(m-1)$ 业已成立，进而证明定理对于 m 也成立：由 $\{\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_m\}$ 的线性无关，显然有 $\{\mathbf{b}_2, \dots, \mathbf{b}_m\}$ 这 $(m-1)$ 个向量也线性无关，对它运用归纳假设，即得一 $(m-1)$ 个元素的子集 $A_{m-1} = \{\mathbf{a}_{i_1}, \dots, \mathbf{a}_{i_{m-1}}\}$ ，使得

$$\{\mathbf{b}_2, \dots, \mathbf{b}_m\} \cup (\{\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n\} \setminus A_{m-1})$$

依然构成 U 的一组生成系，所以 \mathbf{b}_1 是上述 n 个向量的线性组合，即

$$\mathbf{b}_1 = \lambda_2 \mathbf{b}_2 + \dots + \lambda_m \mathbf{b}_m + \sum_{i \neq i_1, \dots, i_{m-1}} \mu_i \mathbf{a}_i$$

上式中的 μ_i 不可能全为 0，否则就和 $\{\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_m\}$ 是线性无关相矛盾。设 $\mu_{i_m} \neq 0$ ，则 $A_m = A_{m-1} \cup \{\mathbf{a}_{i_m}\}$ ，不难由前段对 $m=1$ 的证明看出

$$\{\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_m\} \cup (\{\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n\} \setminus A_m)$$

构成 U 的一组生成系。 □

【定义】：一个具有有限个元素所组成的生成系的向量空间叫做有限维向量空间。

【定义】：一个有限维向量空间 V 的一组不可约生成系叫做 V 的一组基底（亦即一组线性无关的生成系）。

【推论一】：一个有限维向量空间 V 的任给二组基底所含有的元素个数必定相同，这个相等的个数叫做 V 的维数，以 $\dim V$ 记之（若 $V = \{\mathbf{0}\}$ ，则定义 $\dim V = 0$ ）。

证明：设 $\{\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n\}$ 和 $\{\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_m\}$ 是 V 的两组基底，即它们都是线性无关的生成系。由前面定理即得 $n \leq m$ 及 $m \leq n$ ，所以 $n = m = \dim V$ 。

【推论二】：设 V 为有限维，则其中任何一个线性无关向量组 $\{\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_m\}$ 都可以扩充成一组 V 的基底。

证明：取 $\{\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n\}$ 为 V 的一组基底，则由前面定理及其推论一即得一组含有 $\{\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_m\}$ 的 V 的基底。

【推论三】：设 V 为有限维的，则它的任给子空间 U 也必定是有限维的，而且 $\dim U \leq \dim V$ 。

证明：若 $U = \{\mathbf{0}\}$ ，则 $\dim U = 0$ ，推论三显然成立。不然，任取 U 中一个向量 $\mathbf{b}_1 \neq \mathbf{0}$ ，若 $U = \langle \{\mathbf{b}_1\} \rangle$ ，则 $\dim U = 1$ ，推论三依然成立。不然，即 $U \setminus \langle \{\mathbf{b}_1\} \rangle \neq \emptyset$ 。任取 $\mathbf{b}_2 \in U \setminus \langle \{\mathbf{b}_1\} \rangle$ ，则显然有

$$U \supseteq \langle \{\mathbf{b}_1, \mathbf{b}_2\} \rangle, \quad \mathbf{b}_1, \mathbf{b}_2 \text{ 线性无关}$$

如此逐步在 U 中选取线性无关向量组，因为 $U \subseteq V$ ，这种线性无关向量组的所含的个数不可能比 $\dim V$ 大，所以到了某一步时，必定会出现

$$U = \langle \{\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_m\} \rangle$$

的情形， $m \leq \dim V$ 。

【例子】：设 α 为一个在有理系数内不可约 n 次多项式 $f(x)$ 的根，则 $\dim_{\mathbb{Q}} \mathbb{Q}(\alpha) = n = \deg f(x)$ （详见第 V 节）。

IV. 线性方程组的基础理论

在第四章中，我们以 n -阶线性方程组（亦即 n 个 n 元线性方程）的唯一解问题为中心课题，经由逐阶的归纳研讨，发现了行列式，并且完满地解答了上述问题，亦即系数行列式不为零是一个 n 阶线性方程组具有唯一解的充要条件，而且其唯一解可以用行列式加以表达（即 Cramer's 公式）。但是一般线性方程组的基础理论，则有待于上述域上线性空间的基本概念和基础理论才能给以完满、系统的解答。

让我们先来研讨一般的齐线性方程组。设 $A = (a_{ij}), 1 \leq i \leq m, 1 \leq j \leq n$ ，是由 m 个 n 元线性方程所组成者的系数矩阵， \mathbf{x} 是以 F^n 为其变域的向量变元。则方程组可以简约地写成

$$(*) \quad A \cdot \mathbf{x} = \mathbf{0}$$

【引理 1】：令 $S = \{\mathbf{x} \in F^n, A \cdot \mathbf{x} = \mathbf{0}\}$ 是所有解向量所组成的集合。则 S 乃是 F^n 中的一个子空间。

证明：若 $\mathbf{x}_1, \mathbf{x}_2 \in S, k \in F$ ，则有

$$\begin{aligned} A \cdot (\mathbf{x}_1 + \mathbf{x}_2) &= A \cdot \mathbf{x}_1 + A \cdot \mathbf{x}_2 = \mathbf{0} + \mathbf{0} = \mathbf{0} \\ A \cdot (k \cdot \mathbf{x}_1) &= k \cdot A \cdot \mathbf{x}_1 = k \cdot \mathbf{0} = \mathbf{0} \end{aligned}$$

所以 $\mathbf{x}_1 + \mathbf{x}_2$ 和 $k\mathbf{x}_1$ 也都属于 S ，亦即 S 是一个 F^n 中的子空间。□

既然上述解向量的集合乃是一个子空间，则它有其维数 $\dim S$ 。所以一个自然的问题是如何确定 $\dim S$ 。例如当 $m = n$ 而且 $\det(A) \neq 0$ 时， S 只含有 $\mathbf{0}$ （零向量），亦即 $\dim S = 0$ 。在一般的情形，我们也要从系数矩阵的子行列式去探讨。让我们先来看一看下述特殊情形，即

$$A = (a_{ij}), \quad 1 \leq i \leq m, 1 \leq j \leq n, m < n$$

而且 $\det(\mathbf{a}_1, \dots, \mathbf{a}_m) \neq 0$ 。令 A' 是由 A 的前 m 列所构成的 $m \times m$ 方阵，并以 $x_{m+1} = 1, x_j = 0, m+1 < j \leq n$ 代入方程组 $A \cdot \mathbf{x} = \mathbf{0}$ 。即得

$$(*) \quad A' \mathbf{x}' = -\mathbf{a}_{m+1}, \quad \mathbf{x}' = \begin{bmatrix} x_1 \\ \vdots \\ x_m \end{bmatrix}$$

由所设 $\det A' \neq 0$ ，即可用 Cramer's 公式解得

$$x_j^{(m+1)} = \frac{-1}{\det A'} \det(\dots, \mathbf{a}_{m+1}, \dots), \quad 1 \leq j \leq m$$

其中 $\det(\dots, \mathbf{a}_{m+1}, \dots)$ 表示把 \mathbf{a}_j 换成 \mathbf{a}_{m+1} 者。由此即得

$$\mathbf{b}_1 = (x_1^{(m+1)}, \dots, x_m^{(m+1)}, 1, 0, \dots, 0)$$

乃是原给方程组 (*) 的一个解向量。同理可得 $(n - m)$ 个 (*) 的解向量，即

$$\mathbf{b}_{\ell-m} = (x_1^{(\ell)}, \dots, x_m^{(\ell)}, 0, \dots, 1, \dots, 0), \quad m+1 \leq \ell \leq n$$

其中

$$x_j^{(\ell)} = \frac{-1}{\det A'} \det(\dots, \mathbf{a}_\ell, \dots), \quad (\mathbf{a}_\ell \text{ 在第 } j \text{ 的位置})$$

【引理 2】：上述 $(n-m)$ 个解向量业已构成 $A \cdot \mathbf{x} = \mathbf{0}$ 的解子空间 S 的一组基底（亦即线性无关的生成系）。

证明：上述 $\{\mathbf{b}_k, 1 \leq k \leq n-m\}$ 显然线性无关。因为

$$\lambda_1 \mathbf{b}_1 + \lambda_2 \mathbf{b}_2 + \cdots + \lambda_{n-m} \mathbf{b}_{n-m}$$

的最後 $(n-m)$ 个分量显然就是 $\{\lambda_1, \lambda_2, \cdots, \lambda_{n-m}\}$ 。所以上述线性组合等于 $\mathbf{0}$ 的充要条件就当然是所有 $\{\lambda_k, 1 \leq k \leq n-m\}$ 都是 0。

再者，设 \mathbf{b} 是 S 中的任给向量。令 $\beta_{m+k}, 1 \leq k \leq n-m$ 为 \mathbf{b} 的最後 $(n-m)$ 个分量，则

$$\mathbf{b}^* = \mathbf{b} - \sum_{k=1}^{n-m} \beta_{m+k} \mathbf{b}_k \in S$$

而且 \mathbf{b}^* 的最後 $(n-m)$ 个分量都等于 0。由此可见 \mathbf{b}^* 的前 m 个分量所构成的 $m \times 1$ 矩阵 \mathbf{b}' 满足 $A' \cdot \mathbf{b}' = \mathbf{0}$ 。但是 $\det A' \neq 0$ ，所以 $\mathbf{b}' = \mathbf{0}$ 。亦即 $\mathbf{b}^* = \mathbf{0}$ ， $\mathbf{b} = \sum_{k=1}^{n-m} \beta_{m+k} \mathbf{b}_k$ 。□

接著，让我们再来看一下另一种特殊情形，即 $m = n$ ， $\det A = 0$ 但是 $D_{n,n} \neq 0$ 。令 $\tilde{A} \cdot \mathbf{x} = \mathbf{0}$ 是原给方程组 $A \cdot \mathbf{x} = \mathbf{0}$ 中略去最後一个之所得者。由前段的讨论（即 $m = n-1$ 的情形）可知 $\tilde{A} \cdot \mathbf{x} = \mathbf{0}$ 的解子空间就是 \mathbf{b}_1 的所有倍积，而且不难看出

$$\mathbf{b}_1 = \frac{(-1)^{n+1}}{D_{n,n}} \begin{bmatrix} D_{n,1} \\ -D_{n,2} \\ D_{n,3} \\ \vdots \\ (-1)^n D_{n,n-1} \\ (-1)^{n+1} D_{n,n} \end{bmatrix}$$

是否上述 $\tilde{A} \cdot \mathbf{x} = \mathbf{0}$ 的解向量也满足原给的第 n 个线性方程呢？亦即 $\bar{\mathbf{a}}_n \cdot \mathbf{b}$ 是否也等于 0 呢？直接代入即有

$$\bar{\mathbf{a}}_n \cdot \mathbf{b}_1 = \frac{1}{D_{n,n}} \sum (-1)^{n+j} a_{n,j} D_{n,j} = \frac{\det A}{D_{n,n}}$$

所以由所设 $\det A = 0$ 得出 $\bar{\mathbf{a}}_n \cdot \mathbf{b}_1 = 0$ 。

上述计算直截了当地证明了在所设的特殊情形， $\tilde{A} \cdot \mathbf{x} = \mathbf{0}$ 的解也就是 $A \cdot \mathbf{x} = \mathbf{0}$ 的解。

在矩阵 $A = (a_{ij})$ 的 m 行 n 列中各取定其中 k 行、 k 列，然后删去其余各行、各列，则所剩者组成一个 $k \times k$ 方阵。我们把如此所得的 $k \times k$ -方阵的行列式叫做 A 的一个 k -阶子行列式。

【定义】：若 A 至少具有一个不等于零的 k -阶子行列式，而 A 的所有 $(k+1)$ -阶行列式都是零（或不定义），亦即 k 乃是 A 的非零子行列式的最高阶数，则称 k 为 A 的秩 (rank of A)，以 $\text{rk}(A)$ 记之。

【定理】：齐线性方程组 $A \cdot \mathbf{x} = \mathbf{0}$ 的解子空间的维数等于 $n - \text{rk}(A)$ 。

证明：作适当的行、行或列、列互换，不妨设 k 阶子行列式 $|a_{ij}| \neq 0$, $1 \leq i, j \leq k$ 。令 \hat{A} 是由 A 的前 k 行所构成的 $k \times n$ -矩阵。将引理 2 用到线性方程组 $\hat{A} \cdot \mathbf{x} = \mathbf{0}$ ，可得 $\hat{A} \cdot \mathbf{x} = \mathbf{0}$ 的解子空间的一组基底 $\{\mathbf{b}_j; 1 \leq j \leq n - k\}$ 。显然， $\hat{A} \cdot \mathbf{x} = \mathbf{0}$ 的解子空间包含 $A \cdot \mathbf{x} = \mathbf{0}$ 的解子空间。所以我们只要在 $m > k$ 的情形再证明每一个 \mathbf{b}_j 其实也满足

$$\bar{\mathbf{a}}_i \cdot \mathbf{b}_j = 0, \quad k + 1 \leq i \leq m.$$

当然，我们可以对于每个 i 逐个验证上述等式。不难看出，每次所要验证者，其实就是业已验证的特殊情形（亦即 $n = k + 1$ 的情形）。□

F^n 和 $(F^n)^*$ 的对偶性与齐线性方程组：

F^n 和 $(F^n)^*$ 分别是 F 上的 $n \times 1$ -矩阵和 F 上 $1 \times n$ 矩阵所构成的线性空间。它们的维数都等于 n ，而且两者的元素可以用矩阵乘法相乘，即

$$\bar{\mathbf{a}} \in (F^n)^*, \quad \mathbf{b} \in F^n, \quad \bar{\mathbf{a}} \cdot \mathbf{b} = \sum_{i=1}^n a_i b_i$$

设 $A \cdot \mathbf{x} = \mathbf{0}$ 是一个给定的齐线性方程组， $\{\bar{\mathbf{a}}_i, 1 \leq i \leq m\}$ 是 A 的 m 个行向量。令

$$U = \langle \bar{\mathbf{a}}_1, \dots, \bar{\mathbf{a}}_m \rangle$$

是其所张的子空间，而 S 则是 $A \cdot \mathbf{x} = \mathbf{0}$ 的解子空间。请注意， U 和 S 分别是 $(F^n)^*$ 和 F^n 中的子空间。

【定理】： $\dim U + \dim S = n$

证明：令 $k = \text{rk}(A)$ ，不妨设 $|a_{ij}|, 1 \leq i, j \leq k$ 是 A 的一个非零 k -阶子行列式。则 $\{\bar{\mathbf{a}}_i, 1 \leq i \leq k\}$ 显然线性无关，所以 $\dim U \geq k$ 。再者，由前述定理可知 $\dim S = n - k$ ，所以我们只须证明 $\dim U$ 不可能大于 k 。

以前述定理中的 $\{\mathbf{b}_j, 1 \leq j \leq n - k\}$ 为列向量构成一个 $n \times (n - k)$ 矩阵 B ，然后再将行列转置成一个 $(n - k) \times n$ 矩阵 B^t 。因为它们的后 $(n - k)$ 列所成的方阵就是 $\begin{pmatrix} 1 & & 0 \\ & \ddots & \\ 0 & & 1 \end{pmatrix}$ ，所以 $\text{rk}(B^t) = (n - k)$ ，因此 $B^t \cdot \mathbf{x} = \mathbf{0}$ 的解子空间的维数是 k 。再者，由

$$\bar{\mathbf{a}}_i \cdot \mathbf{b}_j = 0 \Leftrightarrow \mathbf{b}_j^t \cdot \bar{\mathbf{a}}_i^t = 0$$

可见 $\{\bar{\mathbf{a}}_i^t, 1 \leq i \leq m\}$ 都是 $B^t \cdot \mathbf{x} = \mathbf{0}$ 的解向量。所以 $\{\bar{\mathbf{a}}_i^t, 1 \leq i \leq m\}$ 之中最多只有 k 个线性无关者，亦即 $\dim U \leq k$ 。这也就证明 $\dim U = k$ ，

$$\dim U + \dim S = n$$

□

【推论】： $\text{rk}(A)$ 等于其行向量所张的子空间的维数，也等于其列向量所张的子空间的维数。

现在让我们进而讨论非齐次的线性方程组，亦即 $A \cdot \mathbf{x} = \mathbf{b}$, $\mathbf{b} \neq \mathbf{0}$ ，的情形。不难看到，我们所要研讨的要点有二，其一是方程组 $A \cdot \mathbf{x} = \mathbf{b}$ 有解的充要条件为何？其二是方程组 $A \cdot \mathbf{x} = \mathbf{b}$ 满足有解条件者，其解向量所组成的集合如何描述？

令 \tilde{A} 是将 $m \times n$ 系数矩阵 A 再加以 $-\mathbf{b}$ 作为其第 $(n + 1)$ 列的 $m \times (n + 1)$ -矩阵。

【定理】：方程组 $A \cdot \mathbf{x} = \mathbf{b}$ 有解的充要条件是 $\text{rk}(\tilde{A}) = \text{rk}(A)$ 。

证明：令 $\tilde{\mathbf{x}} = \begin{pmatrix} \mathbf{x} \\ x_{n+1} \end{pmatrix}$ ， S, S_0 和 \tilde{S} 分别是

$$A \cdot \mathbf{x} = \mathbf{b}, \quad A \cdot \mathbf{x} = \mathbf{0}, \quad \text{和} \quad \tilde{A} \cdot \tilde{\mathbf{x}} = \mathbf{0}$$

的解集。则显然有

$$A \cdot \mathbf{x} = \mathbf{b} \Leftrightarrow \tilde{A} \cdot \begin{pmatrix} \mathbf{x} \\ 1 \end{pmatrix} = \mathbf{0}$$

$$A \cdot \mathbf{x} = \mathbf{0} \Leftrightarrow \tilde{A} \cdot \begin{pmatrix} \mathbf{x} \\ 0 \end{pmatrix} = \mathbf{0}$$

令 $(F^n, 0)$ 和 $(F^n, 1)$ 表示 F^{n+1} 中其第 $(n+1)$ 分量分别是 0 和 1 的子集。则上述关系可以用下述集合关系式表述之，即

$$\tilde{S} \cap (F^n, 0) = (S_0, 0)$$

$$\tilde{S} \cap (F^n, 1) = (S, 1)$$

由此可见， S 非空的充要条件是

$$\tilde{S} \supsetneq \tilde{S} \cap (F^n, 0) \Leftrightarrow \dim \tilde{S} > \dim S_0$$

再者，由前述定理

$$\dim \tilde{S} = (n+1) - \text{rk}(\tilde{A}),$$

$$\dim S_0 = n - \text{rk}(A)$$

而且显然有 $\text{rk}(\tilde{A}) \geq \text{rk}(A)$ 。所以

$$S \neq \phi \Leftrightarrow \text{rk}(\tilde{A}) = \text{rk}(A). \quad \square$$

【引理】：设 $A \cdot \mathbf{x} = \mathbf{b}$ 有解，而 \mathbf{x}_0 是其中一个解。则 $A \cdot \mathbf{x} = \mathbf{b}$ 的解集可以写成 $\mathbf{x}_0 + S_0$ ，其中 S_0 是 $A \cdot \mathbf{x} = \mathbf{0}$ 的解集。

证明：设 $A \cdot \mathbf{x}_0 = \mathbf{b}$, $A \cdot \mathbf{y} = \mathbf{0}$ ，则有

$$A(\mathbf{x}_0 + \mathbf{y}) = A \cdot \mathbf{x}_0 + A \cdot \mathbf{y} = \mathbf{b} + \mathbf{0} = \mathbf{b}$$

反之，设 $A \cdot \mathbf{x}_0 = \mathbf{b}$, $A \cdot \mathbf{x} = \mathbf{b}$ ，则有

$$A(\mathbf{x} - \mathbf{x}_0) = A \cdot \mathbf{x} - A \cdot \mathbf{x}_0 = \mathbf{b} - \mathbf{b} = \mathbf{0} \quad \square$$

V. 域的扩张与维数

【定义】：设 F 是包含在一个域 E 中的子集，而且在和 E 同一的运算下也满足域的性质，则称 F 是 E 的子域，而 E 则称为 F 的扩张域。

例如， \mathbb{R} 是 \mathbb{C} 的一个子域， \mathbb{C} 是 \mathbb{R} 的一个扩张域； \mathbb{Q} 是 \mathbb{R} 的一个子域， \mathbb{R} 是 \mathbb{Q} 的一个扩张域，当然 \mathbb{Q} 也是 \mathbb{C} 的一个子域，所以 \mathbb{C} 也是 \mathbb{Q} 的一个扩张域。在第一节例子中，我们已指出 \mathbb{C} 是 \mathbb{R} 上的一个向量空间， \mathbb{R} 是 \mathbb{Q} 上的一个向量空间；同样地，当 E 是 F 的一个扩张域时， E 也是 F 上的一个向量空间（其加法和倍积其实就是 E 中域的加法和域的乘法限制在 $F \times E$ 部份）。

【定义】：设 E 是 F 的扩张域，令 $\dim_F E$ 为把 E 看成 F 上的向量空间的维数。

【定理】：设 $E \supset F$ 是 F 一个扩张域， $\alpha \in E$ 是某一个 $F[x]$ 中的 n 次不可约多项式 $f(x)$ 的根，亦即 $f(\alpha) = 0$ 。令 $F(\alpha)$ 是 E 中含有 $\{F$ 和 $\alpha\}$ 的最小的子域，亦即由 F 和 α 所生成的子域，则 $\dim_F F(\alpha) = n$ 。

证明：令

$$\left. \begin{aligned} I(\alpha) &= \{g(x) \in F[x], g(\alpha) = 0\} \\ m(x) &\text{ 是 } I(\alpha) \text{ 中一个最低次的非零元素。} \end{aligned} \right\}$$

我们要先证明 (i) $m(x) = \lambda \cdot f(x)$, $\lambda \neq 0 \in F$ 。假若不然，则有带余除式

$$f(x) = q(x) \cdot m(x) + r(x), \quad r(x) \neq 0, \deg r(x) < \deg m(x)$$

以 α 代入上式，即得

$$r(\alpha) = f(\alpha) - q(\alpha) \cdot m(\alpha) = 0 \quad \Rightarrow \quad r(x) \in I(\alpha)$$

这和 $m(x)$ 是 $I(\alpha)$ 中最低次非零元素的假设矛盾，所以 $r(x) = 0$, $q(x) = \lambda \in F$ （因为 $f(x)$ 是假设为不可约的）。

(ii) $\{1, \alpha, \alpha^2, \dots, \alpha^{n-1}\}$ 是 F -向量空间 E 中的线性无关向量组。假若不然，则存在一组不全为 0 的 $\lambda_i \in F$, $i = 0, 1, \dots, n-1$, 使得

$$\sum_{i=0}^{n-1} \lambda_i \alpha^i = 0 \quad \Rightarrow \quad g(x) = \sum_{i=0}^{n-1} \lambda_i x^i \in I(\alpha)$$

是和 (i) 中所证者相矛盾的。

(iii) 最後让我们来证明 $\{1, \alpha, \alpha^2, \dots, \alpha^{n-1}\}$ 构成 $F(\alpha)$ 的一组生成系。换句话说， $\langle 1, \alpha, \alpha^2, \dots, \alpha^{n-1} \rangle = \{\sum_{i=0}^{n-1} \lambda_i \alpha^i; \lambda_i \in F\} = M$ 业已构成一个域，其证明如下：

关系式 $f(\alpha) = 0$ 就是说 α^n 可以用 $\{1, \alpha, \alpha^2, \dots, \alpha^{n-1}\}$ 的线性组合表达。由此，不难归纳地证明 $\alpha^N \in M$ 对任何 N 皆成立。换句话说， M 已经在乘法之下封闭的了。再者，设

$$\sum_{i=0}^{n-1} \lambda_i \alpha^i \in M$$

是任给的 M 中一个非零元素，令 $g(x) = \sum_{i=0}^{n-1} \lambda_i x^i$ ，则显然有 $(g(x), f(x)) = 1$ (即互素)，所以存在有适当的多项式 $A(x), B(x)$ 使得

$$A(x) \cdot g(x) + B(x) \cdot f(x) \equiv 1$$

以 α 代入上式，即得

$$A(\alpha) \cdot g(\alpha) + B(\alpha) \cdot 0 = A(\alpha) \cdot g(\alpha) = 1$$

这也就是说 M 中的任给非零元素的倒数依然在 M 之中！ □

【定理】：设有三个域 F, M, E 具有关系

$$F \subset M \subset E, \quad \dim_F M = m, \quad \dim_M E = n$$

(亦即 M 是 F 上的 m 维向量空间， E 是 M 上的 n 维向量空间) 则

$$\dim_F E = m \cdot n = \dim_F M \cdot \dim_M E$$

证明：在 M 中取定一组 F 上基底 $\{a_1, a_2, \dots, a_m\}$ ， E 中取定一组 M 上的基底 $\{b_1, b_2, \dots, b_n\}$ 。我们要证明

$$\{a_i \cdot b_j; i = 1, 2, \dots, m; j = 1, 2, \dots, n\}$$

恰好构成 E 在 F 上的一组基底。先证它是一组 F 上的生成系：设 ξ 是 E 中任给元素，则有

$$\begin{aligned}\xi &= \sum_{j=1}^n \mu_j b_j, \quad \mu_j \in M \\ \mu_j &= \sum_{i=1}^m \lambda_{ij} a_i, \quad \lambda_{ij} \in F\end{aligned}$$

以第二式代入第一式，即得

$$\xi = \sum_{i=1, j=1}^{m, n} \lambda_{ij} (a_i \cdot b_j), \quad \lambda_{ij} \in F$$

这也就说明了 $\{a_i \cdot b_j; i = 1, 2, \dots, m; j = 1, 2, \dots, n\}$ 构成 E 在 F 上的一组生成系。

再者，设有线性关系

$$\sum_{i=1, j=1}^{m, n} \lambda_{ij} (a_i \cdot b_j) = 0, \quad \lambda_{ij} \in F$$

则可将上式改写成

$$\sum_{j=1}^n \left(\sum_{i=1}^m \lambda_{ij} a_i \right) \cdot b_j = 0, \quad \sum_{i=1}^m \lambda_{ij} a_i \in M$$

因为 $\{b_1, b_2, \dots, b_n\}$ 是在 M 上线性无关，所以对于任一 j ，都有

$$\sum_{i=1}^m \lambda_{ij} a_i = 0, \quad j = 1, 2, \dots, n$$

再由 $\{a_1, a_2, \dots, a_m\}$ 在 F 上的线性无关性，即得所有 λ_{ij} 都必须全部为 0，这也就证明了

$$\dim_F E = m \cdot n = \dim_F M \cdot \dim_M E$$

运用上述公式，我们就可以完满地解决在绪论中所提出的轨尺作图问题。

基础数学讲义之二

《基础几何学之一》

定性与定量平面几何，立体几何基础论

项武义

香港科技大学数学系

目录

引言	v
0.1 基本概念和基本结构	vi
0.2 平面上的次序与分隔	ix
0.3 对称性	x
0.4 平行性	xii
一 连结、分隔与对称——定性平面几何	1
1.1 等腰三角形的特征性质	1
1.2 定性平面几何中的常用基本事实	3
1.3 例题和习题	15
二 平行性与定量平面几何基础理论	19
2.1 平行性和三角形内角和	19
2.2 平行性、平行四边形和面积公式	21
2.3 中国古代的定量几何	26
2.4 不可公度性的发现与克服	29
2.5 例题和习题	36
三 圆与三角学	47
3.1 正弦、余弦函数的基本性质	48
3.2 三角定律	52
3.3 习题	60
四 空间中的平行与垂直	63
4.1 平直性与平行	63

4.2	对称性与垂直	69
4.2.1	垂直平分线与平面上的反射对称	69
4.2.2	立体几何中的作图题	72
4.2.3	空间反射对称性与垂直投影	76

引言——空间的基本概念与基本结构

几何学乃是人类理性文明，对于我们和大自然中的万物万象共存于其中的空间的「认识论」。宇宙中的所有事物皆存在于其中、发生于其内，当然也永远受著空间本质的制约与蕴育。几何学的课题也就是去研究、理解空间的本质，它是我们研讨大自然、理解大自然的天然起点和基石所在；它也是整个自然科学的启蒙者和奠基者，是理所当然的第一科学。不论在自然科学的发展顺序上，或在全局的基本重要性上，几何学都是当之无愧的先行者与奠基者，也是种种科学思想和方法论的自然发祥地。它源远流长，历经数千年世代相承精益求精的研究和逐步逐阶的进展，至今依然根深干粗，蓬勃拙壮。在现今廿一世纪，它会继续是开拓新知的有力工具，而自然科学的拓展又必然对于空间几何学的理解深度和广度提出新的要求和问题。总之自然科学和几何学的进展是密切相关、相辅相成的。伽利略 (Galileo) 曾说：「上帝必定是一个几何学家 (God must be a geometer)。」其所指也许就是上述自然科学和几何学之间的自然结合。

自古到今，几何学的研究在方法论上大体可以划分成下述几个阶段：

- (1) 实验几何：用归纳实验去发现空间之本质。
- (2) 推理几何：以实验几何之所得为基础，改用演绎法，以逻辑推理去探索新知，并对于已知的各种各样空间本质，精益求精地作系统化和深刻的分析。在这方面，古希腊文明获得了辉煌的成就，它也是全人类理性文明中的重大篇章。
- (3) 坐标解析几何：笛卡儿 (Descartes) 和费玛 (Fermat) 通过坐标系的建立，把当代数学中的两大主角——几何学和代数学——简明有力

地结合起来，开创了近代数学的先河。其自然而然的结果是微积分的产生和大量地运用解析法研讨自然现象。

- (4) 向量几何：从现代的观点来看，空间最为根本而且控制全局的本质乃是由它的所有保长变换所构成的变换群 (transformation group)，通常又称之为 3-维的欧氏群 (Euclidean group of \mathbb{E}^3)，而几何学所研究者就是这个变换群的不变量理论 (invariant theory)。因为所有几何量都根源于长度，所以必然在保长变换之下保持不变；反之，任何在保长变换群的作用之下保持不变的（亦即不变量—*invariants*）也都具有几何意义，而且也一定根源于长度。向量几何在本质上乃是坐标解析几何的返璞归真，它的最大优越性在于向量运算的正交不变性 (orthogonal invariance)。可以说，向量几何乃是不依赖于坐标系的解析几何 (coordinate-free analytical geometry)，它自然而然地化解了原先在坐标解析几何中，由坐标系的选取所引入的各种各样（非几何的）非不变量的困扰！Hamilton 和 Grassmann 分别是 3-维和高维的向量代数的创始者。

0.1 基本概念和基本结构

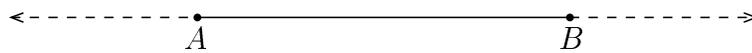
位置 (location) 是空间的基本概念之中最为原始者。空间本身其实就是宇宙之中所有可能的位置的总体。在几何学的讨论中，通常用点 (point) 来标记位置，所以点其实就是位置的抽象化 (abstraction)。当一个动点 (moving point) 由一个位置移动到另一位置，其所经过的点组成这个运动的通路 (path)。连结于空间各地之间的通路则是空间基本概念中第二个最原始者。再者，光线的普遍存在和我们的视觉很自然地启示我们，并促使我们认识到空间的基本结构乃是：

「连结给定两点之间的所有通路之中，有一条唯一的最短通路——它就是连结两点的直线段。」

这也就是在我们日常生活的大气层内，或者在太空中，光由一点射向另一点所经过的通路，亦即我们常见的光线 (light rays)。再者，我们日常的经验是：若不受阻碍，光线是会一直向前无限延伸的。射线 (ray) 这个基本几何概念就是上述这种可以无限向前延伸的光线的抽象化，

而空间中给定相异两点 $\{A, B\}$ 所确定的直线则是由 A 射向 B 的射线和由 B 射向 A 的射线的和集 (union)。总结上述的讨论，空间的基本结构可以描述如下：

【基本几何结构】：对于空间给定相异两点 $\{A, B\}$ 存在有唯一连结于 A, B 之间的最短通路，称之为连结 A, B 的直线段 (interval)，将以符号 \overline{AB} 表示之。再者由 A 到 B 的最短通路可以向前无限延伸，称之为由 A 射向 B 的射线，将以符号 \overrightarrow{AB} 表示之。而该线段向两端无限延伸的通路，亦即 $\overrightarrow{AB} \cup \overrightarrow{BA}$ ，则称之为由 $\{A, B\}$ 所唯一确定的直线 (straight line)，将以符号 AB 表示之。



[图 0-1]

[注]：点是最为原始的几何事物 (geometric object)，所有其他的几何事物都是由点组合而成的。直线段和直线则是第二种最为原始的几何事物，所有其他的几何结构和性质都是由它们所表达的基本结构来刻划和表述的。

直线和直线段之间，显然有下述基本关系：

直线上的次序与分隔：

- (i) \overline{AB} 是 AB 的一个子集。若 $C \in \overline{AB}$ 而且 $C \neq A$ 或 B ，则称 C 位于 A, B 之间。再者，若 $C, D \in \overline{AB}$ 则 $\overline{CD} \subseteq \overline{AB}$ 。
- (ii) 直线 l 上任给一点 P 把直线 l 分割成两段，称为 P 的两侧。属于同侧的两点 A_1, A_2 其直线段 $\overline{A_1A_2}$ 不包含 P ；而属于异侧的两点 A, B 其直线段 \overline{AB} 则包含 P 。

设 $\{A, B, C\}$ 是 $l \setminus \{P\}$ 中的相异三点，而且 $P \in \overline{AB}$ ，则 \overline{AC} 和 \overline{BC} 中有一且仅有一包含 P 。再者，由相异两点定一直线段（或一直线）这两种密切相关的空间基本结构，就可以自然地定义下述两种和两者各别相容的子集。

【定义】：空间中的子集 S 若满足性质：

$$\text{若 } A, B \in S, A \neq B \text{ 则 } \overline{AB} \subset S$$

则称 S 为一个凸子集 (convex subset)。

【定义】：空间中的子集 S 若满足性质：

$$\text{若 } A, B \in S, A \neq B \text{ 则 } AB \subset S$$

则称 S 为一个平直子集 (straight or rectilinear subset)。

显然，所有平直子集也都是凸子集。但是，反之则不然。例如直线段 \overline{AB} 是一个凸子集，但是它并非平直子集，而直线 AB 本身则当然是一个平直子集和凸子集。

再者，由上述定义，易见凸子集的交集 (intersection) 还是凸子集，平直子集的交集还是平直子集。由此可见，对于空间给定的点集 S ，在所有包含 S 的凸子集之中有一个唯一的最小者，它其实就是所有包含 S 的凸子集的交集是也，通常叫做 S 的凸包 (convex hull of S)，我们将以 $C(S)$ 表示之。同样地，所有包含 S 的平直子集的交集乃是那个包含 S 的平直子集中的最小者，通常叫做 由 S 所张的平直子集 (the rectilinear subset spanned by S)，我们将以 $\langle S \rangle$ 表示之。

注意：我们将把空集合 ϕ 和单点集合 $\{P\}$ 想成凸子集和平直子集的特例。(因为它们根本不会含有相异两点，所以其检验条件无从用起！)

【例子】：

(1) 当 $S = \{A\}$ 只含有单个点者，则

$$C(\{A\}) = \{A\}, \quad \langle \{A\} \rangle = \{A\}$$

(2) 当 $S = \{A, B\}$ 是由相异两点组成者，则

$$C(\{A, B\}) = \overline{AB}, \quad \langle \{A, B\} \rangle = AB$$

(3) 当 $S = \{A, B, C\}$ 而且 A, B, C 不共线，则

$$C(\{A, B, C\}) = \triangle ABC, \\ \langle \{A, B, C\} \rangle \text{ 是由不共线三点所张的平面。}$$

(4) 当 $S = \{A, B, C, D\}$ 是由不共面四点所组成者，则

$C(\{A, B, C, D\})$ 是以 A, B, C, D 为其顶点的四面体，
它的四个面就是 $\triangle ABC, \triangle BCD, \triangle CDA$ 和 $\triangle DAB$ ，
而其所张的平直子集已经是全空间了！

$$\langle \{A, B, C, D\} \rangle = \text{全空间}。$$

所以空间中的平直子集只有五种，即空集合 ϕ ，单点子集 $\{P\}$ ，直线，平面和全空间。

由此可见，平面乃是仅次于全空间的平直子集，它是一种介乎于直线和全空间之间，而又具有连点直线段和直线这种空间基本结构的子空间。所以，平面乃是一种既比空间简单而又保有空间基本结构的几何结构。平面几何学的课题就是研究平面上所保有的空间基本结构和所反映的各种性质。它是进而研讨空间（立体）几何学的自然而且非常理想的中途站。

0.2 平面上的次序与分隔

类似于点和直线之间的关系，在平面 Π 中的任给一条直线 l 也把平面 Π 切成两片，称之为 l 的两侧。居于同侧的两点 A_1, A_2 ，其直线段 $\overline{A_1A_2}$ 和 l 不相交；而居于异侧的两点 A, B ，其直线段 \overline{AB} 和 l 相交。设 $\{A, B, C\}$ 是 $\Pi \setminus l$ 中的相异三点，而且 $\overline{AB} \cap l \neq \phi$ ，则 $\overline{AC} \cap l$ 和 $\overline{BC} \cap l$ 之中有一且仅有一为非空的。

上面所讨论的是平面在连结（亦即 $\{A, B\} \rightarrow \overline{AB}$ 和 AB ）和次序分隔上的基本结构和基本性质。现在让我们再来探讨平面在这种基础之上所具有的进一层的本质和基本性质，例如常见常用的长度、角度、大小、形状等等。

由平面的分隔，即一个平面 Π 被其上的一条直线分割成两个半面，亦即 $\Pi \setminus l = H_\ell^+ \cup H_\ell^-$ ，我们将称之为对于 l 的两个开半面 (open half-plane with respect to l)，而 $H_\ell^+ \cup l$ 和 $H_\ell^- \cup l$ 则称之为对于 l 的闭半面 (close half-plane with respect to l)，易证它们都是凸子集。

[证明留作习题]

设 $\{A, B, C\}$ 是不共线三点， Π 是其所张的平面。令

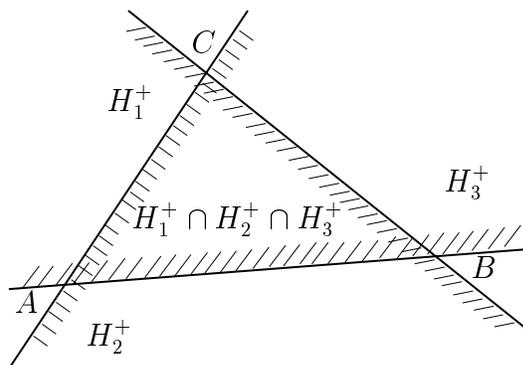
H_1^+ 是关于直线 BC 的闭半面而且含有 A 者，

H_2^+ 是关于直线 CA 的闭半面而且含有 B 者，

H_3^+ 是关于直线 AB 的闭半面而且含有 C 者；

则凸子集 $H_1^+ \cap H_2^+ \cap H_3^+$ (如 [图 0-2] 所示) 就是一个以 A, B, C 为其顶点的三角形，通常以 $\triangle ABC$ 表示之。

[习题：试证它正是 $\{A, B, C\}$ 的凸包。]



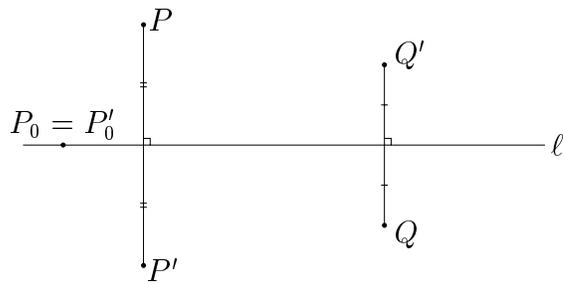
[图 0-2]

在几何学的研讨中，三角形是仅次于线段和直线的基本几何图形，而空间的大部分基本性质都已经在三角形的几何性质中充分体现。三角形之所以成为古希腊几何学所研讨的主角，其原因也就是：三角形既简单而又能充分反映空间的本质。

大体上来说，空间的本质最为基本者就是前面已经讨论的连结、分隔再加上对称性，平行性和连续性。

0.3 对称性

在平面几何的范畴来说，平面对于其上每一条直线皆成反射对称。用现代的术语来说，平面 Π 上对于直线 l 的反射对称是一个从 Π 到 Π 的映射 (亦称为变换) $\mathfrak{R}_l: \Pi \rightarrow \Pi$ ，它把 l 的点固定不动，把不在 l 的点如 P 点映射到 P' 点使得 $\overline{PP'}$ 和 l 正交于 $\overline{PP'}$ 的中点，如 [图 0-3] 所示。

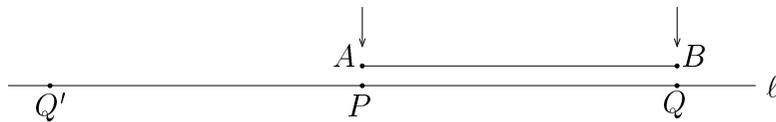


[图 0-3]

假如把一张纸想成是平面的局部，则在上述反射对称 \mathfrak{R}_ℓ 之下相互对应的点 P, P' 也就是把纸张沿 ℓ 折摺时相互叠合者。而二个相互叠合的直线段（或角区）则显然是等长（或等角）的。由此可见这种反射对称乃是一种保长、保角的变换。

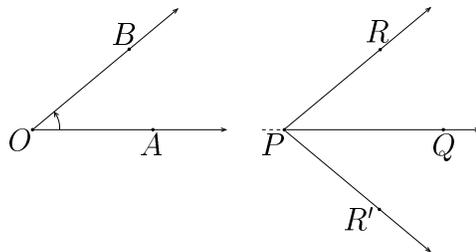
[注]：古希腊的几何学家肯定是认识到上述反射对称性及其保长保角性的。也许他们认为这种描述还不够初等，所以他们改用下述三条叠合公理来描述空间对称性这种本质。

1. 直线段的叠合公理：设 \overline{AB} 是一个给定直线段， P 是给定直线上一个给定点。则在 P 的两侧各有唯一一点 Q, Q' 使得 \overline{AB} 和 $\overline{PQ}, \overline{PQ'}$ 能够叠合，亦即等长。



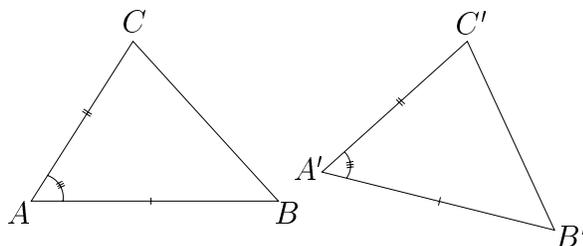
[图 0-4]

2. 角区的叠合公理：设 $\angle AOB$ 是一个给定角， \overrightarrow{PQ} 是一个给定射线，则在直线 PQ 的两侧各有唯一的射线 \overrightarrow{PR} 和 $\overrightarrow{PR'}$ ，使得 $\angle AOB$ 和 $\angle QPR$ 或 $\angle QPR'$ 能够叠合，亦即等角。



[图 0-5]

3. 三角形的叠合公理：两个三角形 $\triangle ABC$ 和 $\triangle A'B'C'$ 能够相互叠合（亦称之为全等）的充要条件是它有相应的两边及其夹角 (S.A.S.) 能够彼此叠合（亦即对应等长和等角）。



[图 0-6]

例如 $\triangle ABC$ 和 $\triangle A'B'C'$ 之间有 \overline{AB} 和 $\overline{A'B'}$ 等长， \overline{AC} 和 $\overline{A'C'}$ 等长，而且 $\angle BAC$ 和 $\angle B'A'C'$ 等角；则必有 \overline{BC} 和 $\overline{B'C'}$ 等长， $\angle ABC$ 和 $\angle A'B'C'$ 等角和 $\angle ACB$ 和 $\angle A'C'B'$ 等角。

[注]：上述三条叠合公理之中，第一条和第二条其实是用来确立等长和等角这两个基本概念，而第三条才是真正的反映著空间对称性这种本质。再者，他们用上述公理推导的第一个定理就是等腰三角形 $\triangle ABC$ （即 \overline{AB} 和 \overline{AC} 等长）是对于其顶角平分线成反射对称的，然後以等腰三角形为工具去探索空间对称性各种各样广泛而且深远的影响。

0.4 平行性

平面上的一条射线表达了一个方向，而一条直线则是具有两个相反的方向。再者两条共起点的射线 \overrightarrow{AB} 和 \overrightarrow{AC} 在方向上的差别也就是 $\angle BAC$ 的角度所表达者，亦即角度乃是其方向差的度量是也。在平面几何中，两个射线同向平行的直观内含是两者所表达的方向相同。以下让我们来分析一下，这种「方向相同」的概念究竟应该如何定义才算是合理。

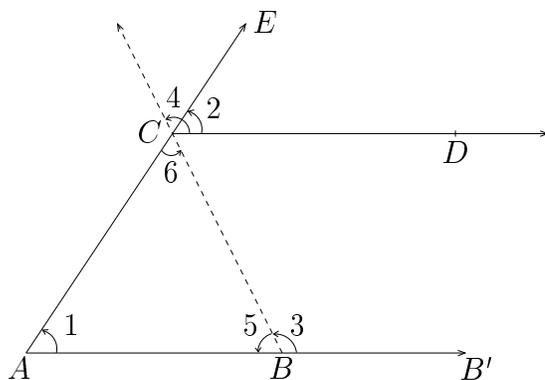
【分析】：

1. 若 $\overrightarrow{AB}, \overrightarrow{CD}$ 共在一条直线上，则易见两者同向的充要条件是

$$\overrightarrow{AB} \cap \overrightarrow{CD} = \left\{ \begin{array}{l} \overrightarrow{AB}, \text{ 或} \\ \overrightarrow{CD}. \end{array} \right.$$

[在两者反向时 $\overrightarrow{AB} \cap \overrightarrow{CD}$ 的可能性是 ϕ ，单点或一个有限长线段。]

2. 在 $\overrightarrow{AB}, \overrightarrow{CD}$ 不共线的情形，连结 \overrightarrow{AC} 。如 [图 0-7] 所示， \overrightarrow{AC} 和 \overrightarrow{CE} 共线同向，而 $\angle 1$ 和 $\angle 2$ 则分别度量著 \overrightarrow{AB} 和 \overrightarrow{AC} ， \overrightarrow{CD} 和 \overrightarrow{CE} 的方向差。由此可见， $\angle 1 = \angle 2$ 应该是检验 \overrightarrow{AB} 和 \overrightarrow{CD} 是否方向相等的合理条件，因为「等向」减「等向」应该还是相等者，是不？



[图 0-7]

3. 假如我们采取上述「合理」的条件来定义 $\overrightarrow{AB}, \overrightarrow{CD}$ 是否「等向」，即 $\angle 1 = \angle 2$ 表示 $\overrightarrow{AB}, \overrightarrow{CD}$ 「等向」。令 $\overrightarrow{BB'}$ 是和 \overrightarrow{AB} 共线同向者，则由

$$\overrightarrow{AB}, \overrightarrow{BB'} \text{ 「等向」 和 } \overrightarrow{AB}, \overrightarrow{CD} \text{ 「等向」}$$

当然也应该有 $\overrightarrow{BB'}$ 和 \overrightarrow{CD} 「等向」。因此如 [图 0-7] 所示的 $\angle 3$ 应该等于 $\angle 4$ 。再者，由 [图 0-7] 所示

$$\angle 4 = \angle 6 + \angle 2, \quad \angle 3 + \angle 5 = \text{平角} (\pi)$$

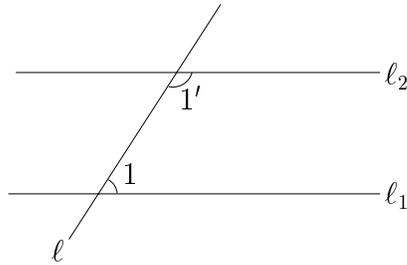
所以

$$\angle 1 = \angle 2 \text{ 和 } \angle 3 = \angle 4$$

$$\Rightarrow \angle 1 + \angle 6 + \angle 5 = \angle 2 + \angle 6 + \angle 5 = \angle 3 + \angle 5 = \pi$$

上面三点分析说明了平面上的「等向」概念是否能够合理地定义而又不产生矛盾，那就要看三角形的内角和是否恒等于一个平角(π)。换句话说，平面的平行性和三角形的内角和恒等于 π 其实是同一件事的两种表现。

[注]：在古希腊的欧几里得的“Elements”中，以下述平行公理来描述平行性，即设平面上两条直线 l_1, l_2 和第三条直线 l 相交：



[图 0-8]

若同旁内角 $\angle 1, \angle 1'$ 之和小于 π ，则 l_1, l_2 必定相交于 l 的 $\angle 1, \angle 1'$ 所在的那一侧。

平行性在平面几何中所扮演的角色是它使得定量几何中的各种公式都大大的简化。例如三角形的面积公式是「底 \times 高 $\div 2$ 」，直角三角形的三边满足勾方加股方等于弦方，以及相似三角形定理等等都是必须依赖于平行性的！

[注]：空间的连续性在直观上业已由一条直线乃是连续不断的，但是它又是「一剪即断」，亦即一条直线 l 略去其中任给一点 P 後，即已分割成两断。但是上述连续性的深入理解和深远影响则有待往後在适当的地方再详加研讨。

第一章

连结、分隔与对称——定性平面几何

定性平面几何所要研讨的主题是「全等形」和「平行性」。在本质上，前者乃是平面对于任给直线的反射对称性的具体反映，而後者则是三角形的内角和恒等于一个平角 (π) 所表达的「平直性」。在本章中，我们将以一个平面上连结与分隔的基本结构和三个叠合公理为基础，把定性平面几何中常用的基本事实，作一次简明扼要的逻辑推导和系统化整理。而关于平行性的讨论，则会留待下一章再作系统整理；亦即在本章的论证中，我们将会完全避免用平行性，所以所证得的结果，在非欧面也同样成立！

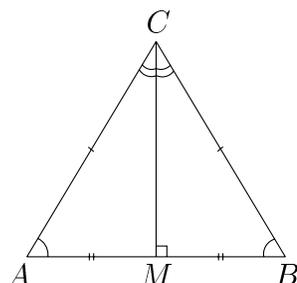
为了简化叙述起见，往後将会用 \overline{AB} 同时表示该直线段子集及其长度， $\angle ABC$ 同时表示该角区及其角度（亦即 \overrightarrow{BA} 和 \overrightarrow{BC} 的方向差）等等。而一些集合的表述方式如 $S = \langle l_1 \cup l_2 \cup \{A\} \rangle$ 和 $\{P\} = AB \cap CD$ 将会简化为 $S = \langle l_1, l_2, A \rangle$ 和 $P = AB \cap CD$ 等等。

1.1 等腰三角形的特徵性质

在各种各样的平面图形之中，三角形乃是最为简单者；而在各种各样的三角形之中，最为基本者则首推等腰三角形。究其原因，就是等腰三角形所具有的轴对称能够具体而微地反映著平面的反射对称性，所以它们乃是研讨平面几何之中对称性的种种表现与推论的基本工具。所以定性平面几何的首要之务，就是推导等腰三角形的各种各样的特徵性质，亦即：

等腰三角形的特征性质：(如 [图 1-1] 所示)

- (i) $\overline{CA} = \overline{CB}$ (定义)；
- (ii) $\angle A = \angle B$ ；
- (iii) $\angle C$ 的分角线 \overline{CM} 垂直底边 \overline{AB} ；
- (iv) 中线 \overline{CM} 垂直底边；
- (v) 垂线 \overline{CM} 平分顶角。



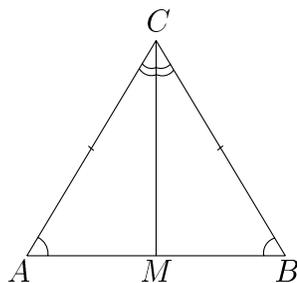
[图 1-1]

上述特征性质的系统推导如下：

【定理 1.1】：设 $\triangle ABC$ 的两边为等腰，即 $\overline{CA} = \overline{CB}$ ，则其顶角 $\angle C$ 的分角线垂直平分底边，而且其两底角相等，即有 $\angle A = \angle B$ 。

证明：设 \overline{CM} 平分顶角，由所设 $\triangle MCA$ 和 $\triangle MCB$ 满足 S.A.S. 全等条件。所以

$$\angle A = \angle B, \overline{AM} = \overline{MB}, \angle CMA = \angle CMB = \frac{\pi}{2} \quad \square$$



[图 1-2]

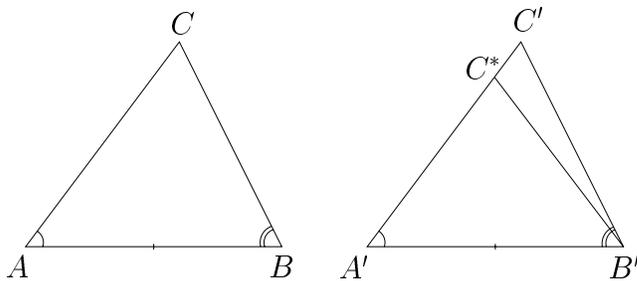
[注]：如上所证， \overline{CM} 乃是等腰三角形 $\triangle ABC$ 的对称轴。等腰三角形乃是平面图形中具有轴对称的最简单者，因此它也是最便于用来分析平面对称性在各种各样几何问题中的作用的「工具」。上述定理已证明由 (i) 可推得余下各个性质；反之，由 (iii), (iv) 或 (v) 即有 $\triangle MCA \cong \triangle MCB$ [证明留作习题]，即得回 (i)。所以现在还需验证余下的 (ii) \Rightarrow (i)，其证明可以由下述一个熟知的叠合条件 A.S.A. 来推导而得：

【定理 1.2】：(A.S.A.) 若 $\triangle ABC$ 和 $\triangle A'B'C'$ 满足 $\angle A = \angle A'$, $\angle B = \angle B'$ 而且 $\overline{AB} = \overline{A'B'}$ ，则 $\triangle ABC \cong \triangle A'B'C'$ 。

证明：若 $\overline{AC} = \overline{A'C'}$ ，则两者已经满足 S.A.S. 全等条件。不然，不妨设 $\overline{A'C'} > \overline{AC}$ 。在 $\overline{A'C'}$ 上取 C^* 点使得 $\overline{A'C^*} = \overline{AC}$ ，则有 $\triangle ABC \cong \triangle A'B'C^*$ (S.A.S.)，所以

$$\angle ABC = \angle A'B'C^* < \angle A'B'C'$$

显然和所设 $\angle ABC = \angle A'B'C'$ 矛盾。 □



[图 1-3]

【定理 1.3】：若 $\triangle ABC$ 的两底角相等，即 $\angle A = \angle B$ ，则 $\triangle ABC$ 必为等腰，亦即 $\overline{CA} = \overline{CB}$ 。

证明：由所设 $\triangle ABC$ 和 $\triangle ACB$ 满足 A.S.A. 条件，用[定理 1.2]即得 $\triangle ABC \cong \triangle ACB$ ，所以 $\overline{CA} = \overline{CB}$ 。 □

1.2 定性平面几何中的常用基本事实

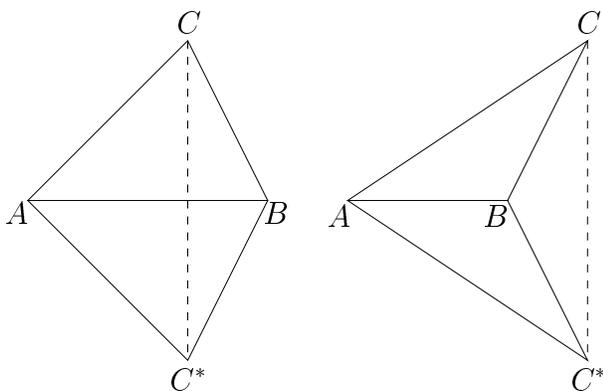
在定性平面几何中还有其他的叠合条件、作图题和不等式，它们都可以用等腰三角形的特征性质来系统地推导而得。运用这些基本的事实，我们还可以证明三角形内角和不大于一个平角（详见本章之末）。

【定理 1.4】：(S.S.S.) 若 $\triangle ABC$ 和 $\triangle A'B'C'$ 的三边对应等长，则 $\triangle ABC \cong \triangle A'B'C'$ 。

证明：如 [图 1-4] 所示，我们可以在 \overline{AB} 的另一侧作 $\triangle ABC^*$ 使得 $\angle BAC^* = \angle B'A'C'$, $\overline{AC^*} = \overline{A'C'}$ ，则有 $\triangle ABC^* \cong \triangle A'B'C'$ (S.A.S.)，所以 $\overline{BC^*} = \overline{B'C'} = \overline{BC}$ 。连结 $\overline{CC^*}$ ，由所设 $\triangle ACC^*$ 和 $\triangle BCC^*$ 皆为等腰。再由 [定理 1.1]，即有

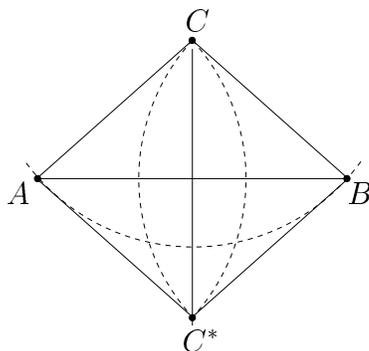
$$\begin{aligned} \angle ACC^* &= \angle AC^*C, \angle BCC^* = \angle BC^*C \\ \Rightarrow \angle ACB &= \angle AC^*B \\ \Rightarrow \triangle ABC &\cong \triangle ABC^* \quad (\text{S.A.S.}) \\ &\cong \triangle A'B'C' \quad (\text{所作}) \end{aligned}$$

□



[图 1-4]

【基本作图题 1.1】：作一个给定角的分角线。



[图 1-5]

[作法] 以给定角顶点 C 为圆心，任取一半径在其角边之两条射线上分别截取 $\overline{CA} = \overline{CB} = r$ ，如 [图 1-5] 所示。在 AB 的另一侧有一个 C^* ，

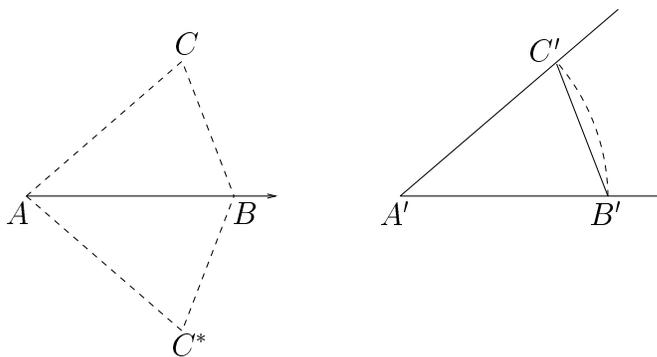
使得

$$\overline{AC^*} = \overline{AC} = \overline{BC^*} = \overline{BC} = r$$

而 C^* 可以由以 r 为半径分别以 A, B 为圆心的交截而得 (C 和 C^* 分别是这两个圆的交点)。连结 $\overline{CC^*}$, 即为所求之 $\angle C$ 的分角线。

证明: 由所作 $\triangle ACC^* \cong \triangle BCC^*$ (S.S.S.), 所以 $\angle ACC^* = \angle BCC^*$ 而且 A, B 分居于 CC^* 之两侧。所以 $\overline{CC^*}$ 平分 $\angle C$ 。□

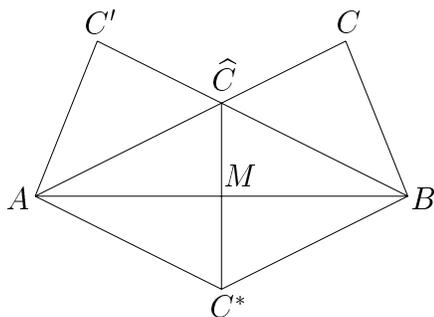
【基本作图题 1.2】: 在射线 \overrightarrow{AB} 之一侧作另一射线 \overrightarrow{AC} 使得 $\angle BAC$ 等于一个给定角 $\angle A'$ 。



[图 1-6]

[作法] 以 \overline{AB} 为半径, A' 为圆心, 分别在 $\angle A'$ 的两个角边上截取 $\overline{A'B'} = \overline{A'C'} = \overline{AB}$ 。再以 A 为圆心, \overline{AB} 为半径和以 B 为圆心, $\overline{B'C'}$ 为半径各作一圆而得两个分居于 AB 两侧的交点 C 和 C^* 。由所作即有 $\triangle ABC \cong \triangle A'B'C' \cong \triangle ABC^*$ (S.S.S.)。□

【基本作图题 1.3】: 作一给定线段 \overline{AB} 的垂直平分线。



[图 1-7]

[作法] 如 [图 1-7] 所示, 在 AB 线外任取一点 C 。若 $\overline{AC} = \overline{BC}$, 则可由 [基本作图 1.1] 的作法求得居于另一侧的 C^* , 使得 $\overline{AC^*} = \overline{BC^*}$ 。则 CC^* 就是 \overline{AB} 的垂直平分线。

若 $\overline{AC} \neq \overline{BC}$, 则可在 AB 同侧再作 C' 点使得 $\triangle BAC' \cong \triangle ABC$ 。它们的两对对应边, 即

$$\{\overline{AC}, \overline{BC'}\} \text{ 和 } \{\overline{BC}, \overline{AC'}\}$$

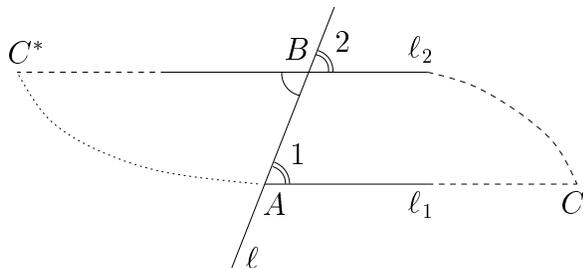
之中必有一对相交于一点 \widehat{C} (如 [图 1-7] 所示)。由 [定理 1.3], $\widehat{AC} = \widehat{BC}$, 所以可以用前面的作法求得 AB 的另一侧的 C^* 点, 使得

$$\overline{AC^*} = \overline{A\widehat{C}} = \overline{BC^*} = \overline{B\widehat{C}}$$

则 $\widehat{CC^*}$ 就是所求作的垂直平分线。

证明: 由所作, $\triangle ABC\widehat{C}$ 是等腰, 而且 $\widehat{CC^*}$ 平分其顶角。由此易见 $\triangle A\widehat{C}M \cong \triangle B\widehat{C}M$ (S.A.S.), 所以 $\overline{AM} = \overline{BM}$, $\angle AMC\widehat{C} = \angle BMC\widehat{C} =$ 直角。□

[定理 1.5]: 设有相异直线 l_1, l_2 分别和第三条直线 l 相交, 若同位角相等 (如 [图 1-8] 所示, $\angle 1 = \angle 2$), 则 l_1, l_2 不相交。



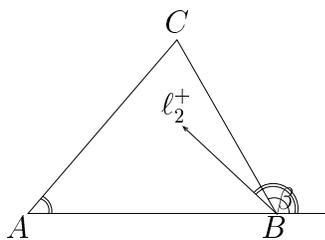
[图 1-8]

证明: 我们用反证法, 设 $C = l_1 \cap l_2$ 。在 l_2 上取 $\overline{BC^*} = \overline{AC}$, $\{C, C^*\}$ 分居 l 的两侧, 连结 $\overline{C^*A}$ 。则由所作 $\triangle ABC^*$ 和 $\triangle BAC$ 满足 S.A.S., 所以两者全等。由此可得

$$\angle BAC^* + \angle 1 = \angle ABC + \angle 2 = \pi \quad (\text{平角})$$

因此 $C^*A = AC = l_1$, 而 l_1, l_2 相交于分居 l 的两侧的 C 和 C^* 。这显然和二点确定唯一一条直线相矛盾。所以 l_1 和 l_2 是不可能相交的! □

【定理 1.6】：三角形的任一外角大于其任一内对角。即如 [图 1-9] 所示， $\angle\beta$ 大于 $\angle A$ 和 $\angle C$ 。



[图 1-9]

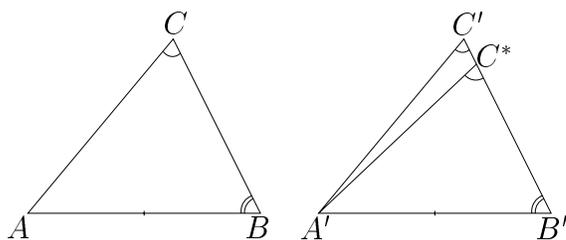
证明：这是[定理 1.5]的直接推论。兹用反证法证之如下：设 B 点的外角 $\angle\beta$ 小于 $\angle A$ ，则过 B 点作直线 l_2 ，使得它和 $l_1 = AC$ 具有相等的同位角。由[定理 1.5]， l_2 和 l_1 不相交。但是由假设 $\angle\beta < \angle A =$ 同位角，射线 l_2^+ 是夹在 \overrightarrow{BA} 和 \overrightarrow{BC} 之间的。所以 A, C 分居于 l_2 的两侧，故此 l_2 和 \overline{AC} 必须相交，亦即和[定理 1.5]相矛盾。这也就证明了外角 $\angle\beta$ 必须都大于其二个内对角 $\angle A, \angle C$ 。□

【定理 1.7】：(A.A.S.) 若 $\triangle ABC$ 和 $\triangle A'B'C'$ 满足 $\overline{AB} = \overline{A'B'}$, $\angle B = \angle B'$ 和 $\angle C = \angle C'$ ，则 $\triangle ABC \cong \triangle A'B'C'$ 。

证明：若两者还有 $\overline{BC} = \overline{B'C'}$ ，则两者其实已经满足 S.A.S.。假若不然，可设 $\overline{BC} < \overline{B'C'}$ 。在 $\overline{B'C'}$ 上取 C^* 点使得 $\overline{B'C^*} = \overline{BC}$ ，则有

$$\triangle ABC \cong \triangle A'B'C^* \quad (\text{S.A.S.})$$

所以 $\angle A'C^*B' = \angle C = \angle C'$ ，亦即 $\triangle A'C^*C'$ 的外角等于其中一个内对角，这是和[定理 1.6] (或者[定理 1.5]) 相矛盾的。所以 \overline{BC} 和 $\overline{B'C'}$ 必须等长，亦即 $\triangle ABC \cong \triangle A'B'C'$ 。□

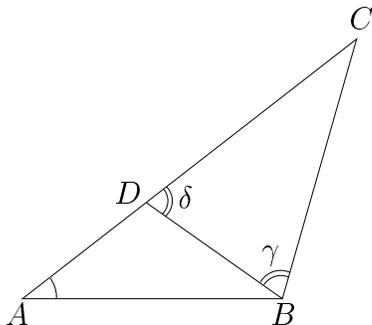


[图 1-10]

[注]：上述的证明中没有用到平直性。若有平直性，即三角形内角和恒为平角，则 A.A.S. 是 A.S.A. 的直接推论。

【定理 1.8】：大边对大角，大角对大边。

证明：先证大边对大角。设 $\triangle ABC$ 的边长中 $\overline{AC} > \overline{BC}$ ，我们要证明 $\angle ABC > \angle BAC$ 。



[图 1-11]

证明：如 [图 1-11] 所示，在 \overline{AC} 上取 D 点使得 $\overline{DC} = \overline{BC}$ ，则有等腰三角形 $\triangle BCD$ 的两底角 $\angle \gamma = \angle \delta$ ，而 $\angle \delta$ 是 $\triangle ABD$ 的 D 点外角，所以即有

$$\angle ABC > \angle \gamma = \angle \delta > \angle BAC$$

[大角对大边的证明留作习题] □

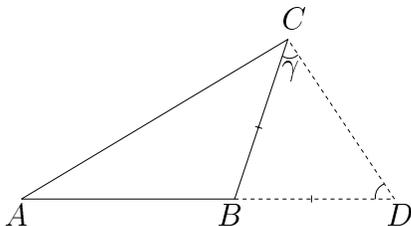
【定理 1.9】：三角形的两边之和大于第三边，即

$$\overline{AB} + \overline{BC} > \overline{AC}$$

证明：如 [图 1-12] 所示，在 \overline{AB} 的延线上取 $\overline{BD} = \overline{BC}$ ，则等腰三角形 $\triangle BDC$ 的两底角相等，所以在 $\triangle ADC$ 中 $\angle ACD > \angle \gamma = \angle ADC$ ，即有

$$\overline{AB} + \overline{BC} = \overline{AD} > \overline{AC}$$

□



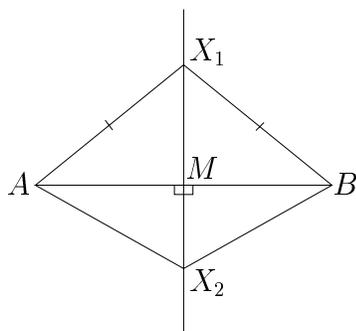
[图 1-12]

【定理 1.10】：在平面 Π 上给定两个相异点 A, B ，其等距的点所成的子集

$$S = \{X \in \Pi, \overline{AX} = \overline{BX}\}$$

乃是 \overline{AB} 的垂直平分线。

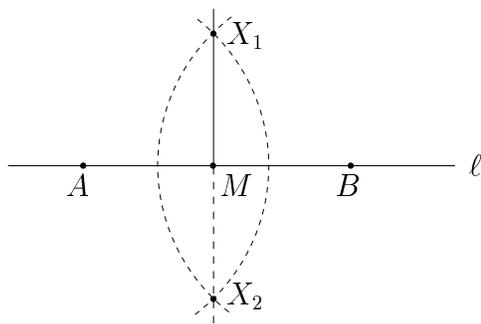
证明：令 M 为 \overline{AB} 的中点， X_1 为 S 中任给一点，则有 $\triangle X_1AB$ 等腰，所以 $XM \perp AB$ 。反之，若 X_2 是 \overline{AB} 中垂线上任给一点，则 $\triangle X_2MA \cong \triangle X_2MB$ (S.A.S.)，所以 $\overline{X_2A} = \overline{X_2B}$ ，亦即 $X \in S$ 。 \square



[图 1-13]

【基本作图题 1.4】：过直线 l 上一点 M ，作其垂直线。

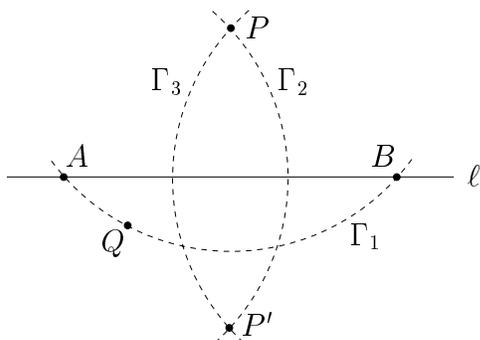
[作法] 在 l 上 M 点的两侧各取 A, B 使得 M 是 \overline{AB} 的中点，再用圆规作两个分别以 A, B 为圆心，以相同但是大于 \overline{AM} 为半径的圆。则两圆会交于 X_1, X_2 两点，而 X_1X_2 即为所求作者。[证明留作习题]



[图 1-14]

【基本作图题 1.5】：设 P 是直线 l 之外一点，作过 P 点而且和 l 垂直的直线。

[作法] 在平面上取一点 Q , 使得 P, Q 分居 l 线的两侧, 如 [图 1-15] 所示。以 \overline{PQ} 为半径, 先以 P 点为圆心作 Γ_1 , 与 l 交于 A, B 两点。然後分别以 A, B 为圆心作 Γ_2, Γ_3 , 得 $\Gamma_2 \cap \Gamma_3$ 的两交点 P, P' 。则 $\overline{PP'}$ 被 l 所垂直平分。[证明留作习题]



[图 1-15]

【定义】：对于平面上一条给定直线 l , 线外两点 P, P' 若满足 $\overline{PP'}$ 被 l 垂直平分, 则称 P, P' 对于 l 成反射对称。

[上述作图题说明如何去由 P 作出 P' 。]

【定义】：平面对于给定直线 l 的反射对称 \mathfrak{R}_l 是平面到自身的一个映射。它把线上的点映射到自己, 线外的点映射到其反射对称点 P , 即

$$\begin{cases} \text{若 } P \in l, \text{ 则 } \mathfrak{R}_l(P) = P; \\ \text{若 } P \notin l, \text{ 则 } \mathfrak{R}_l(P) = P', \overline{PP'} \text{ 被 } l \text{ 垂直平分。} \end{cases}$$

【定理 1.11】： \mathfrak{R}_l 是平面上的一个保长变换, 即：

$$\overline{\mathfrak{R}_l(P)\mathfrak{R}_l(Q)} = \overline{PQ}$$

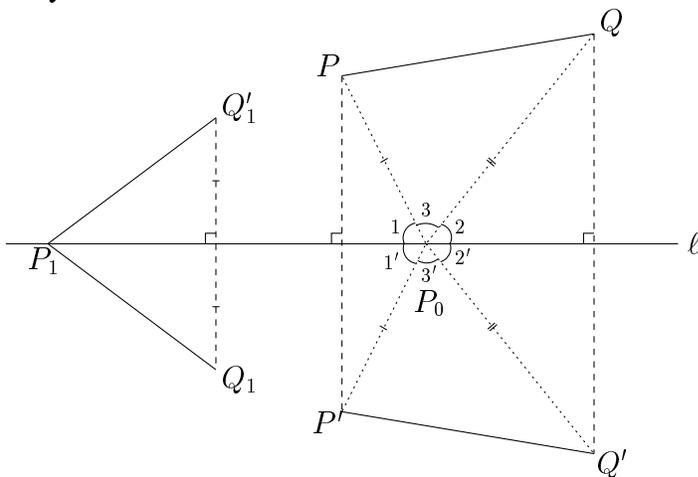
对于平面上任给两点 P, Q 皆成立。

证明：在此, 将采用简约符号 P' 表示 $\mathfrak{R}_l(P)$ 。

(i) 若 $P, Q \in l$, 则 $P' = P, Q' = Q$, 即 $\overline{P'Q'}$ 和 \overline{PQ} 是同一直线段。

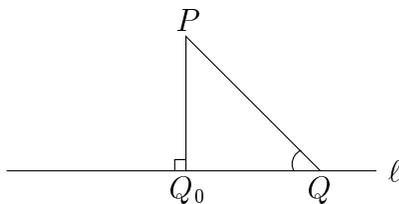
(ii) 若 $P \in l, Q \notin l$, 则 $\triangle PQQ'$ 的底边 $\overline{QQ'}$ 被 l 所垂直平分, 即有 $\triangle PQM \cong \triangle PQ'M$ 。所以 $\overline{P'Q'} = \overline{PQ'} = \overline{PQ}$ 。

- (iii) 若 $P, Q \notin l$, 则在 l 上取定一点 P_0 。由 (ii) 之所证, $\triangle P_0PP'$ 和 $\triangle P_0QQ'$ 都是等腰三角形, 因此 l 乃是它们在顶角的平分线, 由此可见 $\angle PP_0Q = \angle P'P_0Q'$ 。所以 $\triangle PP_0Q \cong \triangle P'P_0Q'$ (S.A.S.), $\overline{P'Q'} = \overline{PQ}$ 。□



[图 1-16]

【定理 1.12】：从直线 l 之外一点 P , 到线上各点的距离以垂直线所给出者为唯一极小。



[图 1-17]

证明：设 $\overline{PQ_0}$ 和 l 正交于 Q_0 , Q 为 l 上任给另外一点, 则 $\triangle PQ_0Q$ 在 Q_0 点的外角是 $\frac{\pi}{2}$ 。由 [定理 1.6], 即得 $\angle Q < \frac{\pi}{2} = \angle Q_0$ 。再由 [定理 1.8] 即得

$$\overline{PQ} > \overline{PQ_0} \quad \square$$

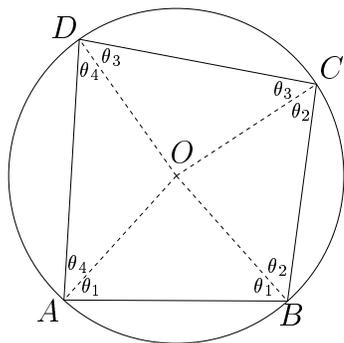
[注]：在这里没有用到勾股定理（勾股定理是依赖于平行性的）。

【定理 1.13】：设 $\square ABCD$ 为圆内接四边形, 亦即四个顶点 A, B, C, D 共圆。则有

$$\angle A + \angle C = \angle B + \angle D$$

证明：令 O 为 $\square ABCD$ 的外接圆圆心，则 $\triangle OAB, \triangle OBC, \triangle OCD, \triangle ODA$ 皆为等腰，所以其底角分别相等。设其有向角为 $\theta_1, \theta_1, \theta_2, \theta_2, \theta_3, \theta_3, \theta_4, \theta_4$ ，则有

$$\angle A + \angle C = \theta_1 + \theta_2 + \theta_3 + \theta_4 = \angle B + \angle D \quad \square$$



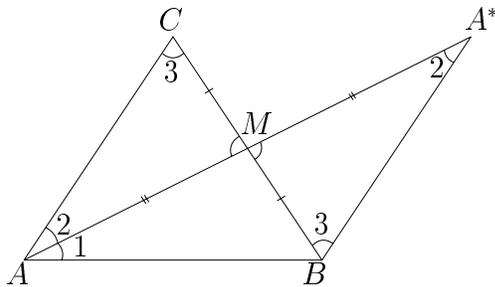
[图 1-18]

【定理 1.14】：三角形的内角和不大于 π 。

证明：我们将用反证法，亦即设存在有一个内角和大于 π 的三角形 $\triangle ABC$ ，即有

$$\angle A + \angle B + \angle C = \pi + \varepsilon, \quad \varepsilon > 0.$$

我们要仅仅用叠合公理，连结与分隔去得出矛盾。设 $\angle A$ 是三个内角中的最小者， M 是 \overline{BC} 的中点。如 [图 1-19] 所示，连结 \overline{AM} ，延长一倍而得 A^* ：



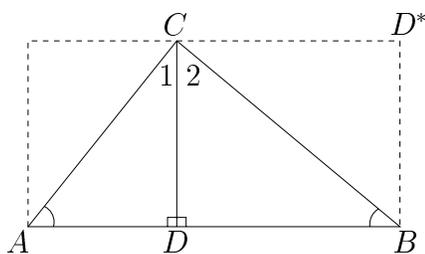
[图 1-19]

由所作易见 $\triangle MBA^* \cong \triangle MCA$ (S.A.S.)，所以 $\triangle ABA^*$ 的三个内角和也等于 $\pi + \varepsilon$ ，而且 $\angle 1 + \angle A^* = \angle A$ 。由此可见 $\triangle ABA^*$ 中的最小内角至

多只是原先 $\triangle ABC$ 的最小内角 ($\angle A$) 的一半。如此逐步构造, 所得的三角形的内角和一直保持是 $\pi + \varepsilon$, 而且其最小内角的大小每次至少减半。所以只要作足够多次, 则其最小内角就肯定要比 ε 还要小! 亦即其另外两个内角之和已经大于 π ! 这显然是和外角大于内对角 (定理 1.6) 相矛盾的。这也是证明了这种内角和大于平角的三角形其实是不可能存在的, 亦即任何三角形的内角和不大于 π 。□

【定理 1.15】: 若存在一个三角形其内角和等于 π , 则任何三角形的内角和也必须等于 π 。

证明: 设有一个三角形 $\triangle ABC$ 其内角和等于 π , 则由其大角 $\angle C$ 到其对边 \overline{AB} 作垂线 \overline{CD} 。

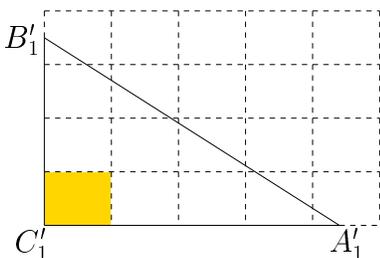


[图 1-20]

即有

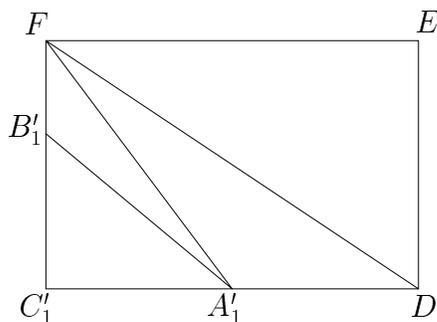
$$\begin{aligned} (\angle 1 + \angle A + \frac{\pi}{2}) + (\angle 2 + \angle B + \frac{\pi}{2}) &= \angle A + \angle B + \angle C + \pi = 2\pi \\ \angle 1 + \angle A + \frac{\pi}{2} &\leq \pi, \quad \angle 2 + \angle B + \frac{\pi}{2} \leq \pi \quad (\text{定理 1.14}) \end{aligned}$$

所以直角三角形 $\triangle CDA$ 和 $\triangle CDB$ 的内角和也都等于 π 。用其中之一即可得一个四内角皆为 $\frac{\pi}{2}$ 的「矩形」, 如 [图 1-20] 所示的 $\square CDBD^*$ 。将这个矩形逐一堆砌即可得出长和宽都可以任意大的矩形, 如 [图 1-21] 所示:



[图 1-21]

现在我们要用[定理 1.14]来证明任何直角三角形的内角和都必然等于 π 。设 $\triangle A_1B_1C_1$ 是一个任给直角三角形， $\angle C_1 = \frac{\pi}{2}$ ，我们可以 [图 1-21] 所作的那个足够大的矩形构造 $\triangle A'_1B'_1C'_1 \cong \triangle A_1B_1C_1$ ，使得其直角边 $\overline{A'_1C'_1}$ 和 $\overline{B'_1C'_1}$ 都包含在该矩形的两个直角边之内。而由 [图 1-22] 所示：



[图 1-22]

$$\begin{aligned}
 \triangle C'_1DF \text{ 的内角和} &= \pi \\
 \Rightarrow \triangle C'_1A'_1F \text{ 的内角和} + \triangle A'_1DF \text{ 的内角和} &= 2\pi \\
 \Rightarrow \triangle C'_1A'_1F \text{ 的内角和} &= \pi \quad (\text{定理 1.14}) \\
 \Rightarrow \triangle A'_1B'_1C'_1 \text{ 的内角和} + \triangle A'_1B'_1F \text{ 的内角和} &= 2\pi \\
 \Rightarrow \triangle A'_1B'_1C'_1 \text{ 的内角和} &= \pi \quad (\text{定理 1.14})
 \end{aligned}$$

这样就证明了任何直角三角形的内角和皆等于 π 。而任何三角形都可以像 [图 1-20] 那样分割成两个直角三角形，所以它的内角和也必然等于 π 。□

[注]：[定理 1.14]和[定理 1.15]证明了在任何满足连结、分隔和叠合（对称性）的几何之中，三角形的内角和不是恒等于 π 就是恒小于 π 。前者是欧氏几何，而后者则是非欧几何。在后者的情形，我们还可以证明下述角亏 (angle defect)

$$\pi - (\angle A + \angle B + \angle C)$$

和 $\triangle ABC$ 的面积成比例。

1.3 例题和习题

【例题】：

(1) 光的反射定律与极小性：光的反射定律是

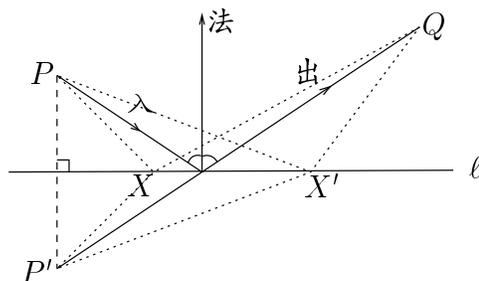
「入射线、反射线和平面在反射点的法线三线共面，而且两者和法线的夹角相等。」

上述定律的几何意义乃是光反射之途径是在所有下述通路

$$\overline{PX} + \overline{XQ}, \quad X \in \ell$$

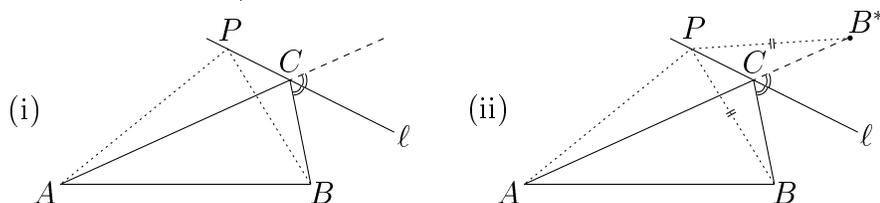
之中取极小值。如 [图 1-23] 所示，

$$\begin{aligned} \overline{PA} + \overline{AQ} &= \overline{P'A} + \overline{AQ} = \overline{P'Q} \\ &\leq \overline{PX} + \overline{XQ} = \overline{P'X} + \overline{XQ} \end{aligned}$$



[图 1-23]

(2) 给定 $\triangle ABC$ ，如 [图 1-24(i)] 所示令 ℓ 为 C 点外角的分角线。设 P 是 ℓ 上一点， $P \neq C$ ，则恒有 $\overline{AP} + \overline{PB} > \overline{AC} + \overline{CB}$ 。



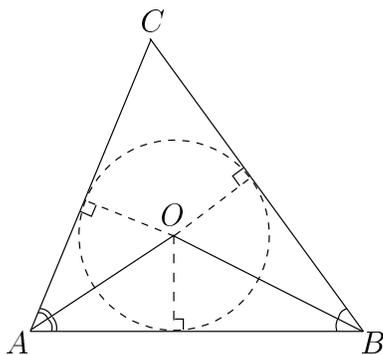
[图 1-24]

如[图 1-24(ii)]所示, 令 B^* 为 B 相对于 l 的反射对称点, 所以即有 $\overline{PB} = \overline{PB^*}$ 。由此可见,

$$\begin{aligned}\overline{AP} + \overline{PB} &= \overline{AP} + \overline{PB^*} \\ &> \overline{AB^*} = \overline{AC} + \overline{CB^*} \\ &= \overline{AC} + \overline{CB}\end{aligned}$$

- (3) 内切圆作图: 对于一个给定的 $\triangle ABC$, 唯一存在一个和其三边相切的圆, 称之为 $\triangle ABC$ 的内切圆 (如 [图 1-25] 所示)。其作图法如下:

用[基本作图 1.1], 分别作 $\angle A$ 和 $\angle B$ 的角平分线。则两线的交点 O' 乃是具有和三边等距的唯一之点, 所以它就是所求作的内心 (内切圆圆心)。



[图 1-25]

- (4) 设 $\triangle ABC$ 和 $\triangle A'BC$ 具有相同的外接圆, 则

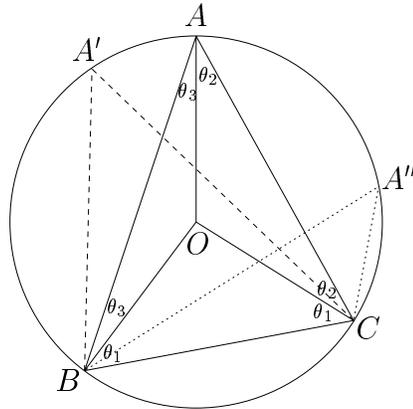
$$\angle ABC + \angle ACB - \angle BAC = \angle A'BC + \angle A'CB - \angle BA'C$$

证明: 如 [图 1-26] 所示, $\triangle OBC$, $\triangle OCA$, $\triangle OAB$ 皆为等腰, 所以其底角各别相等, 即 [图 1-26] 所示之 $\theta_1, \theta_2, \theta_3$ 。再者,

$$\begin{aligned}\angle ABC + \angle ACB - \angle BAC &= (\theta_1 + \theta_3) + (\theta_1 + \theta_2) - (\theta_2 + \theta_3) \\ &= 2\theta_1\end{aligned}$$

由此可见, 上式之角度只和 \widehat{BC} 的大小有关, 而和 \widehat{AC} , \widehat{AB} 的大小无关。

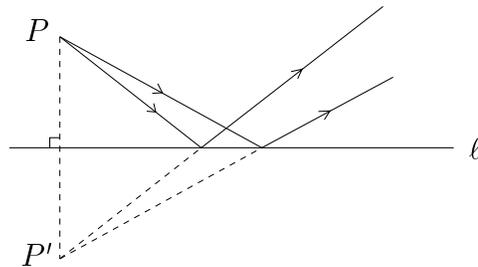
[注意：上式中的 $\theta_1, \theta_2, \theta_3$ 乃是有向角。如在 A'' 的情形， θ_3'' 是负向角。]



[图 1-26]

【习题】：

- (1) 镜子成象的几何原理：令 P' 是 P 对于 l 的反射对称点。试证起始于 P 点的反射线的延长线共交于 P' 点（如 [图 1-27] 所示）。



[图 1-27]

- (2) 试证明在[定理 1.10]中的「大角对大边」部分。
- (3) 试证明在[基本作图题 1.4]中的 X_1X_2 乃是过 l 上 M 点的垂线。
- (4) 试证明在[基本作图题 1.5]中的 PP' 乃是垂直于 l 的直线。
- (5) 设直线 l 和圆 Γ 仅交于一点 P ，试证 $\overline{OP} \perp l$ 。
- (6) 过圆 Γ 上一点 P 作其切线 l 。

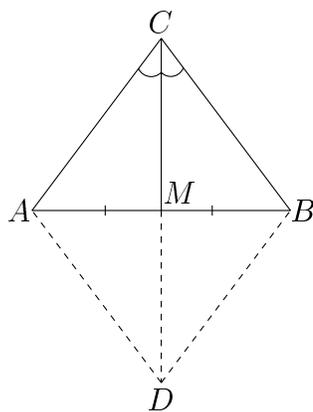
(7) 设四边形 $\square ABCD$ 的两对对边各别等长，试证：

- (i) 其对角线互相平分；
- (ii) 其两对对角各别相等。

(8) 设四边形 $\square ABCD$ 的两对角线互相平分，试证：

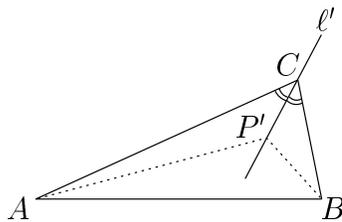
- (i) 其两对对边各别等长；
- (ii) 其两对对角各别相等。

(9) 设 $\triangle ABC$ 的 $\angle C$ 分角线等分对边 \overline{AB} ，试证 $\overline{AC} = \overline{BC}$ 。[延长 \overline{CM} 到 $\overline{CD} = 2\overline{CM}$ (如 [图 1-28] 所示)，则可运用习题 (8) 之结果。]



[图 1-28]

(10) 给定 $\triangle ABC$ ，如 [图 1-29] 所示令 ℓ' 为 $\angle C$ 的分角线， P' 是 ℓ' 上一点， $P' \neq C$ ，试证明 $\overline{AP'} - \overline{P'B} < \overline{AC} - \overline{CB}$ 。



[图 1-29]

第二章

平行性与定量平面几何基础理论

一般来说，我们对于事物的研究，大体上都先作定性的探讨，然後再进而作定量的分析。这是一种由表及里、逐步深入、精益求精的自然进展，平面几何学的研究当然也遵循著这样一种自然的顺理成章的途径。由上一章的讨论中可以看到，在定性地探讨几何中的「等」与「不等」时，我们可以完全不用平行性；但是在定量的平面几何中，我们要对于不等长的两个线段，不同大小的两个角区或不同大小的两个区域，赋以两者之间定量的比值去度量 (measure) 两者之间的差别。在这个时候，平行性扮演著一个举足轻重的「要角」，其作用是大大简化了定量几何的基础理论和基本公式。换句话说，在定量几何中，三角形内角和是恒等于平角还是恒小于平角这两种几何开始有了重大的差别。前者的基本公式要比後者的基本公式简单得多！在前者有简朴易用的基本定理如矩形面积公式、勾股定理和相似三角形定理；而在後者所相应者，不是根本没有，就是要复杂得多。

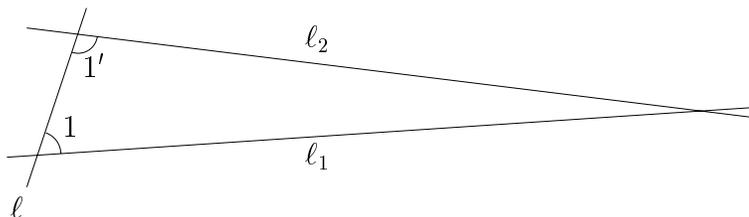
在本章中，我们将会简明扼要地讨论平行性如何反映在平面几何的基本定理之上。再者，我们还会比较分析一下古代中国和古希腊对于定量平面几何的治学方法。

2.1 平行性和三角形内角和

在欧几里德 (Euclid) 的原著《几何原本》 (“Elements”) 中，平面的平行性是用下述第五公设 (fifth postulate)，也就是我们通常称之为

平行公理 (parallel axiom) 者，来加以刻划的：

【第五公设】：设 l_1, l_2 和 l 是平面上三条相异直线，若 l 和 l_1, l_2 相交的同旁内角（如 [图 2-1] 所示之 $\angle 1$ 和 $\angle 1'$ ）之和小于平角，则 l_1, l_2 必定在 $\angle 1$ 和 $\angle 1'$ 的同侧相交。



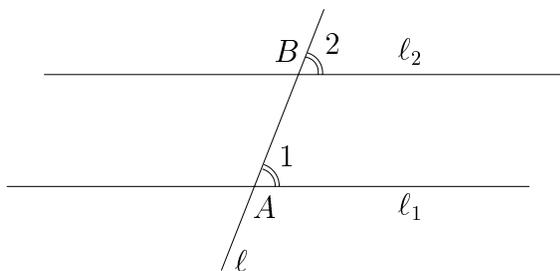
[图 2-1]

亦即

$$\angle 1 + \angle 1' < \text{平角} \Rightarrow l_1, l_2 \text{ 相交于 } \angle 1, \angle 1' \text{ 之同侧}$$

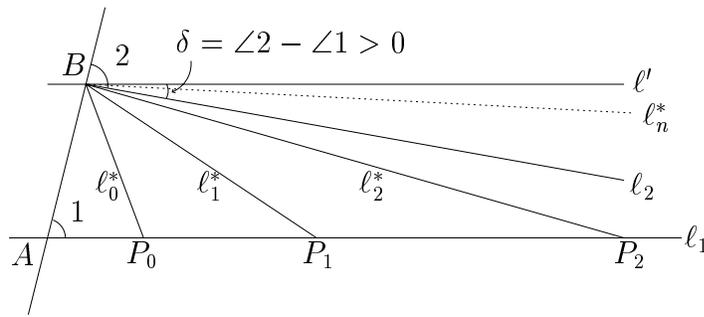
另一方面，在上一章中我们只需用到 S.A.S. 和「两点定唯一直线」就可以证明下述[定理 1.5]：

【定理 1.5】：若 $\angle 1 = \angle 2$ ，则 l_1, l_2 不相交。



[图 1-8']

由此可见，上述的第五公设也就是以「公设」的方式宣称 $\angle 1 + \angle 1' = \pi$ 是 l_1, l_2 不相交的唯一可能性。换句话说，在平面上过直线 l_1 外的一个给定点 P ，而且和 l_1 不相交的直线 l_2 的存在性乃是已证的[定理 1.5]，而其唯一性则就是上述第五公设，亦即是使得同旁内角之和 $\angle 1 + \angle 1' = \pi$ 的那一条 l_2 乃是唯一和 l_1 不相交者。用这个唯一性之所设，就不难推导平直性：三角形内角和恒为一平角（其证明留作习题）。其实，反之亦可用平直性来证明上述第五公设，其证法如下：



[图 2-2]

证明：如 [图 2-2] 所示， l' 是使得同位角 $\angle 2' = \angle 1$ 者，由所设 $\delta = \angle 2 - \angle 2' = \angle 2 - \angle 1 > 0$ 。令 $\{P_0, P_1, \dots, P_n, \dots\}$ 是在 l 上 A 点右侧之点列，满足

$$\overline{P_0P_1} = \overline{BP_0}, \quad \overline{P_1P_2} = \overline{BP_1}, \quad \dots, \quad \overline{P_nP_{n+1}} = \overline{BP_n}$$

亦即

$$\triangle BP_0P_1, \quad \triangle BP_1P_2, \quad \dots, \quad \triangle BP_nP_{n+1}, \quad \dots$$

都是等腰三角形。由平直性和等腰三角形的底角相等，即有 l' 和 $l_n^* = BP_n$ 之间的夹角逐次减半。所以在 n 足够大时，其值小于 $\delta = \angle 2 - \angle 1 > 0$ ，亦即当

$$2^n > \frac{\angle AP_0B}{\delta}$$

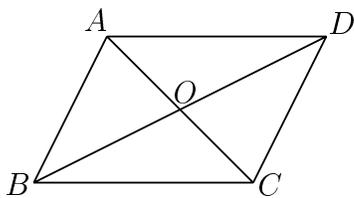
时， l_n^* 业已夹在 l_2 和 l' 之间。但是由所作 l_n^* 和 l_1 相交于 P_n ，所以 l_2 必然和 l_1 相交于线段 $\overline{AP_n}$ 之内。□

[注意，上述证明之中，三角形内角和恒等于一个平角扮演著必不可缺的角色！所以上述论证只是证明了第五公设和平直性的逻辑等价性。]

2.2 平行性、平行四边形和面积公式

在上一章对于对称性的讨论中，得知等腰三角形就是那些具有轴对称性的三角形。而具有两对对边各别等长的四边形则是那些具有心对称的四边形（参看上一章习题）。它的对称中心就是其两条互相平分的对角线的交点，而且它的两对对角也各别相等。证明要点在于每一

条对角线把该四边形切为两个全等的三角形 (S.S.S.)，而两条对角线则把它切为两对互相全等的三角形 (A.S.A.)。



[图 2-3]

上述的论证只用到叠合条件，毋须平行性的帮助的。假如我们把平行性再用上去分析上述四边形的几何结构时，则由对角相等

$$\angle A = \angle C, \angle B = \angle D$$

和四内角和等于 2π 合起来就得出

$$\angle A + \angle B = \angle A + \angle D = \angle C + \angle B = \angle C + \angle D = \pi$$

亦即其对边互相平行：

$$AB \parallel DC, \quad AD \parallel BC$$

所以把这种四边形叫做平行四边形 (parallelogram)。我们把平行四边形的各种特征性质总结如下，叙述为本章的[定理 2.1]。它是研讨平行性在几何学中各种各样的反映的主要工具。[它在平行性的研讨上所扮演的角色一如等腰三角形在对称的研讨上所扮演的角色——都是到处有用和好用的基本工具。]

【定理 2.1】：下列各特性皆为平行四边形的特征性质。

- (i) 两对对边互相平行 (定义)。
- (ii) 两对对边各别等长。
- (iii) 两对对角各别等角。
- (iv) 两条对角线互相平分。
- (v) 有一对对边平行而且等长。

[上述五点的逻辑等价性大部分已经证过，而且证明的要点也在上面说明了，所以其逐一验证则留作习题。]

【定义】：四个内角皆相等（亦即都是直角）的四边形叫做矩形 (rectangle)，亦称为长方形。

【定义】：四个边长皆相等的四边形叫做菱形。

矩形的面积公式：

「矩形的面积等于长乘宽」，这是一个自古以来就在中外古今所熟知和惯用的公式。在中国的古算中，把它当做「显然成立者」，并用来作为推导其他定量几何的公式的起点和基点（详见下一节）。但是古希腊的几何学则进一步去追究这个看来相当明显的面积公式的真正涵意何在，并设法论证之。长话短说，让我们在此简明扼要地回顾一下古希腊当年在定量几何基础论上的探索历程：

几何学是古希腊文明最辉煌的成就，它是以古埃及和古巴比伦文明的几何知识为基础，集希腊的精英，历经好几世纪世代相承、精益求精的研究创造而成者，乃是人类文明中第一个趋于成熟的科学。古希腊的学者把几何学的研讨作为理解宇宙的基础学科，所以治学十分严谨，高度注重其基本概念的确性和推理论证上的严格性。它不但是其他自然科学的基础所在，而且也是整个自然科学在思想上、方法论上治学的典范。古希腊几何学的进程是先研究定性平面几何，其研讨主题是全等形和平行性，然后再进而研讨定量平面几何，而且一开始他们就认识到直线段长度的度量 (measurement of length) 乃是定量几何研讨的起点和基础所在。通常的做法是先取定（或约定）一个单位长 (unit of length)，它可以是公尺、市尺、英尺，或光在真空走一秒的长度「光秒」，也可以是「光年」，然后把一个给定直线段的长度定义为它和单位长之间的「比值」 (ratio)。由此可见，长度度量这个基本概念的关键在于上述「比值」的明确定义。

设给定线段 a 恰好可以等分成 m 段和单位长 u 等长者首尾相接而成，则 a 和 u 之间的比值当然就是 m ，而称这样的直线段 a 的长度是 m 单位。反之，若单位长 u 恰好可以等分成 n 段和 a 等长者首尾相接而成，则 a 和 u 之间的比值应该等于 $\frac{1}{n}$ 。再者，若 a 恰好可以等分成

m 段和 $\frac{1}{n}u$ 等长者连接而成, 则 a 和 u 之间的比值应该定义为分数 $\frac{m}{n}$, 称 a 的长度是 $\frac{m}{n}$ 单位, 以 $a = \frac{m}{n}u$ 表达之。

大约在纪元之前五世纪前後, 古希腊的几何学学界 (例如毕氏及其门人) 关于长度度量提出下述概念及论断:

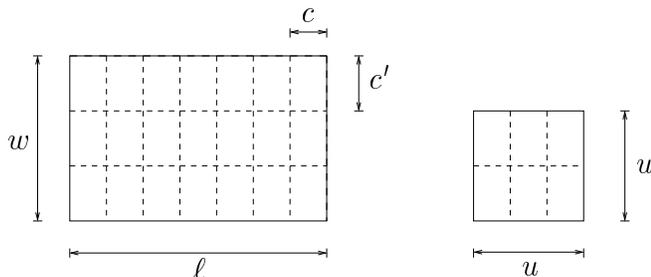
【可公度性】(Commensurability): 对于两个直线段 a, b 若存在一个公尺度 (common yardstick) c 恰能同时整量 a, b , 亦即 a, b 都是 c 的整数倍时: $a = m \cdot c, b = n \cdot c$, 则称 a, b 为可公度 (commensurable), 而 a, b 的长度比值就定义为分数 $\frac{m}{n}$ 。另外一个等价的说法就是: 若存在适当的整数 m, n 使得 $n \cdot a$ 和 $m \cdot b$ 恰好等长, 则称 a, b 为可公度, 而这样的 a, b 长度的比值就定义为 $\frac{m}{n}$ 。

然後他们主观地论断: 任何两个直线段总是可公度的, 亦即可公度性是普遍成立的 (universal validity of commensurability)。并且以此作为他们当年所致力构筑的定量几何基础论的「头号公理」(Principle Axiom), 亦即以「可公度性的普遍成立」为依据、为基石 (foundation), 给定量几何中的基本定理如矩形面积公式、毕氏定理、相似三角定理给出其证明。大致上, 下面所述就是他们当年对于矩形面积公式的证明。

设矩形的长和宽分别是 l 和 w , 而 u 则是取定的单位长度。由可公度性普遍成立的「公设」即分别有 $\{l, u\}$ 和 $\{w, u\}$ 的公尺度 c, c' , 使得它们分别是 c, c' 的整数倍, 亦即

$$\begin{aligned} l &= m \cdot c, & u &= n \cdot c \\ w &= p \cdot c', & u &= q \cdot c' \end{aligned}$$

如 [图 2-4] 所示, $\square(l, w)$ 和 $\square(u, u)$ 分别可以用平行线分割成 $m \cdot p$ 和 $n \cdot q$ 个 $\square(c, c')$ 。



[图 2-4]

由此可见

$$\square(\ell, w) = (m \cdot p)\square(c, c')$$

$$\square(u, u) = (n \cdot q)\square(c, c')$$

所以

$$\begin{aligned}\square(\ell, w) : \square(u, u) &= \frac{mp}{nq} = \frac{m}{n} \cdot \frac{p}{q} \\ &= (\ell : u) \cdot (w : u)\end{aligned}$$

这也就是矩形的面积等于长乘宽的真正涵意。 \square

总之，他们当年基于「可公度性普遍成立」这个「公设」，对于定量平面几何的重要公式如毕氏定理、相似三角形边长比例式等等，都给以严格的证明，建立起洋洋大观的定量平面几何基础论。其中毕氏学派的贡献良多，引以自豪。

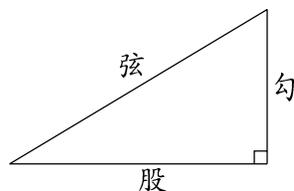
【历史的注记】：假如任何两个直线段真的总是可公度的，则上述证明已经完整无缺地证明了矩形面积公式。但是在毕氏本人百年之後不久，其门徒 Hippasus 却有一个石破天惊的发现，那就是一个正五边形的对角线长和其边长乃是不可公度的！因此当年用来建立定量几何基础论的「头号公设」根本是错误的！亦即可公度性并非普遍成立（随後他也证明了正方形的对角线长和边长也是不可公度的）。由此可见，上述证明只是证明了矩形两个边长 a, b 都是和单位长 u 可公度时这种特殊情形的矩形面积公式，在一般不可公度的情形还得加以补证！

另一点值得在此一提的是上述证明中很关键地用了平行分割，所以矩形面积公式和平行性是必然相关的。当然在三角形内角和恒小于平角的几何中，任何四边形的四个内角和恒小于 2π ，所以根本没有四个内角均为直角的四边形。但是「矩形」在那种几何中其实还是有自然的「推广」者，那就是两对对边各别等长而且其两条对角线也等长的那种四边形，它的面积是其两个边长的函数，可是其公式要比 $a \cdot b$ 复杂得多！

在讨论古希腊的几何学家如何克服上述不可公度性的问题之前，我们先来看看定量平面几何学在中国古代是如何建立和处理的。

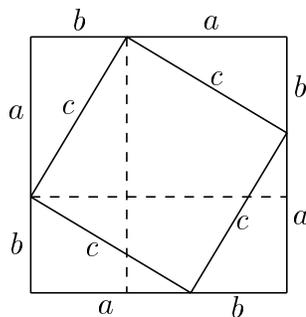
2.3 中国古代的定量几何

大概远在战国时代，定量几何知识乃是以一套简朴实用的测量公式在工匠和水利「工程师」之间流传，如公输般、墨子、西门豹、李冰、李二郎等等很可能就是中国古文明中几何知识的创建者和继承者。在中国古算中，一个直角三角形的两个直角边分别叫做「勾」和「股」，而斜边则叫做「弦」：



[图 2-5]

中国古代几何的独到灼见是善用面积，以矩形面积等于长乘宽为基础，推导直角三角形的面积等于底乘高之半，然後再用下述图解简洁利落地证明了勾股弦公式（或称作勾股定理）和相似直角三角形的比例式。



[图 2-6]

如 [图 2-6] 所示，一个以 $a+b$ 为边长的正方形可以有如实线和虚线所示的两种分割：前者把它分割成一个以 c 为边长的正方形和 4 个以 a, b 为直角边的直角三角形；而後者则它分割成两个以 a, b 为边长的矩形和两个分别以 a, b 为边长的正方形。用以上述两种分割法去计算其总面积，即得下述等式

$$c^2 + 4 \cdot \frac{1}{2}ab = a^2 + b^2 + 2ab$$

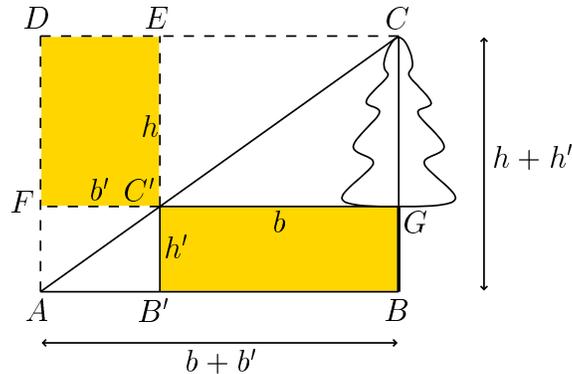
$$\Rightarrow a^2 + b^2 = c^2 \quad (\text{亦即勾方加股方等于弦方})$$

出入相补原理：

在中国的古算测量术中，其所用的基本工具就是上述勾股弦公式和下述出入相补原理。

如 [图 2-7] 所示，在一个给定的矩形的对角线上任取一点 C' ，再过 C' 点作平行于两边的直线段（在实际测量中的水平线和垂线），则有：

$$\triangle ABC \cong \triangle ADC, \triangle AB'C' \cong \triangle AFC', \triangle C'GC \cong \triangle C'EC$$



[图 2-7]

所以 [图 2-7] 所示的两个矩形面积相等，亦即

$$\begin{aligned} \Rightarrow b \cdot h' &= \square BBGC' = \triangle ABC - \triangle AB'C' - \triangle C'GC \\ &= \triangle ADC - \triangle AFC' - \triangle C'EC = \square FC'ED = b' \cdot h \\ \Rightarrow \frac{b}{b'} &= \frac{h}{h'} \Rightarrow \frac{b+b'}{b'} = \frac{h+h'}{h'} \\ \text{亦即 } \overline{AB} : \overline{AB'} &= \overline{BC} : \overline{B'C'} \end{aligned}$$

亦即相似直角三角形 $\triangle ABC$ 和 $\triangle AB'C'$ 的对应直角边边长比例式。由它再加上勾股弦公式，就可以推导 $\overline{AC} : \overline{AC'}$ 也等于上述比值。再者，两个一般的相似三角形总可以用垂线分割成两对相似的直角三角形，所以一般的相似三角形定理又可以直截了当地归于相似直角三角形的对应边比例式去推导。

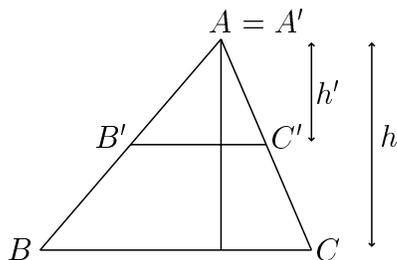
若用现代定量平面几何的知识来分析，上述矩形和直角三角形的面积公式，以及勾股弦和出入相补比例式其实业已构成一组完备的定量平面几何基础。它不但简明扼要，而且用面积公式直截了当地一以贯之。这种处理方式易学好用，至今依然是定量平面几何入门的捷径。

再者，上述讨论也启示我们相似三角形定理本身应该也可以用中国古法，以简简单单的面积计算来加以证明：

【相似三角形定理】：设 $\triangle ABC$ 和 $\triangle A'B'C'$ 的三个对应角相等，则有其三个对应边长成比例，即

$$\frac{\overline{AB}}{\overline{A'B'}} = \frac{\overline{BC}}{\overline{B'C'}} = \frac{\overline{AC}}{\overline{A'C'}} (=k)$$

证一：如 [图 2-8] 所示，我们不妨设 $A' = A$, $\overline{B'C'} \parallel \overline{BC}$ 。



[图 2-8]

效法中国古法，我们用两种办法去计算梯形 $\triangle BCC'B'$ 的面积：

$$\triangle BCC'B' = \triangle ABC - \triangle AB'C' = \frac{1}{2}ah - \frac{1}{2}a'h'$$

$$\triangle BCC'B' = \triangle BCC' + \triangle BC'B' = \frac{1}{2}(h - h')(a + a')$$

由两式相减，即得

$$0 = \frac{1}{2}(ah' - a'h) \Rightarrow \frac{a}{a'} = \frac{h}{h'}$$

$$\Rightarrow \frac{\triangle ABC}{\triangle A'B'C'} = \frac{\frac{1}{2}ah}{\frac{1}{2}a'h'} = \left(\frac{a}{a'}\right) \left(\frac{h}{h'}\right) = \left(\frac{a}{a'}\right)^2$$

同理可得 $\triangle ABC : \triangle A'B'C'$ 也等于 $\left(\frac{b}{b'}\right)^2$ 和 $\left(\frac{c}{c'}\right)^2$ 。即

$$\frac{\triangle ABC}{\triangle A'B'C'} = \left(\frac{a}{a'}\right)^2 = \left(\frac{b}{b'}\right)^2 = \left(\frac{c}{c'}\right)^2$$

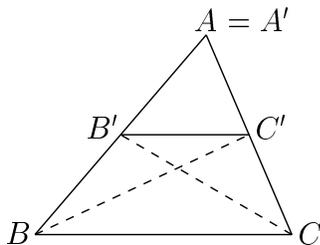
而两个正数的平方相等时，其本身也相等，所以

$$\frac{a}{a'} = \frac{b}{b'} = \frac{c}{c'}$$

□

证二：如 [图 2-9] 所示， $\triangle BCB'$ 和 $\triangle CC'B'$ 是同底等高的，所以它们的面积相等。因此 $\triangle BC'A$ 和 $\triangle CAB'$ 的面积也相等。再者 $\triangle ABC$ 和 $\triangle ABC'$ 是同高的，所以

$$\triangle ABC : \triangle ABC' = \overline{AC} : \overline{AC'}$$



[图 2-9]

同理亦有 $\triangle ABC : \triangle AB'C = \overline{AB} : \overline{AB'}$ 。因为 $\triangle ABC'$ 和 $\triangle AB'C$ 等面积，所以

$$\overline{AC} : \overline{AC'} = \overline{AB} : \overline{AB'}$$

这也就是我们所要证的相似比例式。 □

2.4 不可公度性的发现与克服——Hippasus 和 Eudoxus 对于人类理性文明的重大贡献

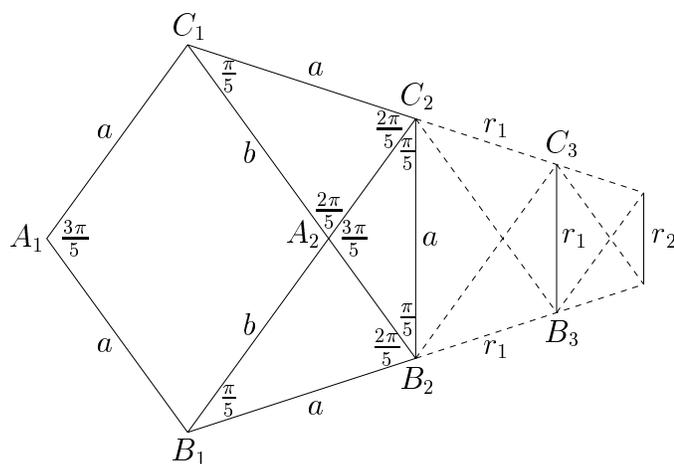
话说当年，毕氏的门徒 Hippasus 对于当年定量几何学的头号公设，亦即「直线段之间可公度性普遍成立」还一直在锲而不舍地钻研。在当年，至少他已认识到下述两个给定的可公度线段 a, b 的最长公尺度的几何求法：

【辗转丈量法】：设 $a < b$ ，我们用 a 为尺去丈量 b 。若恰能整量，即 $b = n_1 a$ ，则显然 a 本身就是 $\{a, b\}$ 的最长公尺度。不然， $b = n_1 a + r_1$ ， $r_1 < a$ 。再用 r_1 为尺丈量 a 。若恰能整量，则 r_1 即为 $\{a, b\}$ 的最长公尺度。不然， $a = n_2 r_1 + r_2$ 。再用 r_2 为尺丈量 r_1 ，如此辗转丈量，一直到 r_k 恰能整量 r_{k-1} 为止。则 r_k 即为所求的最长公尺度。

【历史的注记】：在 $\{a, b\}$ 可公度的情形，即有 $a = mc, b = nc$ 。相应于 $\{a, b\}$ 的辗转丈量，即有 $\{m, n\}$ 的辗转相除求最大公因数的算法，而

在 $r_k = d \cdot c$ 式当中, d 就是 m, n 的最大公因数。因为这种辗转相除算法写在欧几里得的“Elements”中, 所以通常称之为欧氏算法 (Euclidean Algorithm)。但是在古希腊极可能是先有辗转丈量求最长公尺度, 因为这是当年学者们钻研的定量几何基本问题。总之由其一自然也就可直接对应而有其另一。所以欧氏算法显然在欧氏的“Elements”之前二百年即已为 Hippasus 所知和所用。

话说当年, Hippasus 在沙盘上用芦苇杆画了一个大致如 [图 2-10] 所示的正五边形, 然後开始用当时业已熟知的等腰三角形定理和三角形内角和定理来作下述分析, 即 (i) 三角形内角和恒等于 π (平角); 和 (ii) 等腰三角形的两底角相等, 反之, 两底角相等的三角形必为等腰。



[图 2-10]

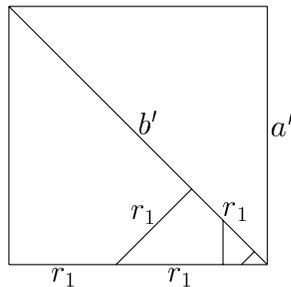
五边形的内角和恒等于 3π , 所以上述正五边形 $\square A_1B_1B_2C_2C_1$ 的每个内角都等于 $\frac{3\pi}{5}$ 。由此可见等腰三角形 $\triangle B_1B_2C_2$ 和 $\triangle B_2C_2C_1$ 的两底角皆为 $\frac{\pi}{5}$ 。令对角线 $\overline{B_1C_2}$ 和 $\overline{B_2C_1}$ 的交点为 A_2 , 则有 $\triangle C_1A_2C_2$ 的两底角皆为 $\frac{2\pi}{5}$, 而 $\triangle A_2B_2C_2$ 的两底角皆为 $\frac{\pi}{5}$, 所以它们都是等腰的。

上述看来不起眼的几何分析却使得 Hippasus 大为震惊! 为什么呢? 若以上述五边形的边长 a 去丈量其对角线长 b , 则其余段 r_1 就是等腰 $\triangle A_2B_2C_2$ 的等边边长。若将 $\overline{C_1C_2}$ 延长一段 $\overline{C_2C_3} = r_1$, $\overline{B_1B_2}$ 延长一段 $\overline{B_2B_3} = r_1$, 则易证 $\square A_2B_2B_3C_3C_2$ 又是一个正五边形, 而它的边长是 r_1 , 对角线长则是 a 。

因此当我们再用 r_1 去丈量 a 时, 在本质上又是用一个正五边形的边

长去丈量其对角线长。同理，所得的餘段 r_2 又是一个更小一号的正五边形的边长而其对角线长则为 r_1 。如此辗转丈量，每一次所做者在本质上总是用一个正五边形的边长去丈量其对角线长，只是那个正五边形逐次缩小吧了。这样就理论上证明了 $\{a, b\}$ 的辗转丈量必然是永无止休的！因此 $\{a, b\}$ 必然是不可公度的 (non-commensurable)！此事焉能叫他不吃惊！?这个惊人的发现事实胜于雄辩地证明了当年定量几何基础论的头号基石——「可公度性的普遍成立」其实根本是一个错误的「公设」！

Hippasus 接著还用下述图解证明正方形的边长和对角线长 $\{a', b'\}$ 之间的辗转丈量也是永无止休的，所以也是不可公度的。



[图 2-11]

由 [图 2-11] 可以看出 $\{a', b'\}$ 的辗转丈量所得的逐步算式是

$$\begin{aligned} b' &= a' + r_1 \\ a' &= 2r_1 + r_2 \\ r_1 &= 2r_2 + r_3 \\ &\dots\dots\dots \\ &\text{往後的表式总是一样的} \\ r_{k-1} &= 2r_k + r_{k+1} \end{aligned}$$

所以是永无止休的。[细节的证明留作习题]

【历史的注记】：Hippasus 的伟大发现，是人类理性文明的重要里程碑，有如发现了一个理念上的新大陆，它不单对于定量几何学有根本的重要性，其实对于整个自然科学都有深远的影响。但是当年古希腊几何学界，特别是 Hippasus 本人所在的毕氏学派对于这个伟大发现的反应，却是全然无理性的。据某些现在已不可详考的记载，Hippasus 反而因为这个重大发现而丧生于同门之手。

其实，不可公度性的存在，并不是全面否定了当年古希腊几何学在定量几何基础论上的成就。它只是说原本以为已经完整无缺的证明其实只是在可公度的情形的证明，而在一般不可公度的情形，则尚有待补证！这个亟待补证的任务对于当年整个古希腊几何学界是一个严峻而且迫切的挑战。大约经历半个世纪的努力，才促使 Eudoxus 开创了影响无比深远广阔的逼近法和逼近原理而得以完美成功。可以这么说，Eudoxus 的思想和方法提供了研讨和理解 Hippasus 所发现的新大陆的基础。

Eudoxus 的逼近法和逼近原理：

(Method and Eudoxus principle of approximation)

当 $\{a, b\}$ 不可公度 时，「 $a : b$ 」不是一个分数。它是一种有待理解的新兴事物，不管你如何称呼它，反正是一种当时尚未了解有待研究的「新量」。例如 $\{a, b\}$ 和 $\{a', b'\}$ 是两对不可公度的直线段，Eudoxus 认识到「 $a : b$ 」和「 $a' : b'$ 」这两个「新量」之间的大小或相等关系都还有待定义！但是当 $\{a, b\}$ 不可公度而 $\{a', b'\}$ 可公度的情形下：

$$a : b \text{ 和 } a' : b' = \frac{m}{n}$$

它们之间的大小关系却又是相当清楚的，亦即：

$$a : b \begin{cases} > \frac{m}{n} \\ < \frac{m}{n} \end{cases} \Leftrightarrow \begin{cases} a \text{ 比 } \frac{m}{n} \cdot b \text{ 长} \\ a \text{ 比 } \frac{m}{n} \cdot b \text{ 短} \end{cases} \Leftrightarrow \begin{cases} n \cdot a > m \cdot b \\ n \cdot a < m \cdot b \end{cases}$$

这也就是 Eudoxus 在研究这种「新量」时第一个认识到的：

Eudoxus 比较原则：

$$a : b \begin{cases} > \frac{m}{n} \\ < \frac{m}{n} \end{cases} \text{ 的充要条件就是 } \begin{cases} n \cdot a > m \cdot b \\ n \cdot a < m \cdot b \end{cases}$$

亦即 $a : b$ 和 $\frac{m}{n}$ 的比较大小可以由 $n \cdot a$ 和 $m \cdot b$ 之间的比较长短而判定之。而後者是极为初等而且其几何意义乃是一目了然的。

在 $\{a:b\}$ 和 $\{a':b'\}$ 都是不可公度的情形，若有分数 $\frac{m}{n}$ 使得

$$a:b > \frac{m}{n} > a':b' \quad \text{亦即} \quad \lceil na > mb \quad \text{但是} \quad na' < mb' \rceil$$

则显然应该定义前者大于后者。反之，若有分数 $\frac{m}{n}$ 使得 $a:b < \frac{m}{n} < a':b'$ ，则应该定义前者小于后者。

再者，假如这种间于 $a:b$ 和 $a':b'$ 之间的分数是不存在的情形，亦即对于任何分数 $\frac{m}{n}$ ， $a:b$ 和 $a':b'$ 与 $\frac{m}{n}$ 之间的大小关系总是同步同样的，理当就可以作为「 $a:b = a':b'$ 」的定义。

这也就是 Eudoxus 当年对于两个不可公度的比值之间的大、小及相等关系的定义，即

$$\text{【定义】} : a:b > a':b' \Leftrightarrow \text{存在分数 } \frac{m}{n} \text{ 使得 } a:b > \frac{m}{n} > a':b'$$

$$a:b < a':b' \Leftrightarrow \text{存在分数 } \frac{m}{n} \text{ 使得 } a:b < \frac{m}{n} < a':b'$$

$$a:b = a':b' \Leftrightarrow \text{对于任何分数 } \frac{m}{n} \text{ 皆有相同的大小关系。}$$

$$\text{亦即：} \quad na \begin{cases} > \\ < \end{cases} mb \Leftrightarrow na' \begin{cases} > \\ < \end{cases} mb'$$

为了论证上述定义的必然性，Eudoxus 开创了影响极为深远的逼近法 (Method of Approximation)。首先，他提出下述直观上极为明显的「公设」作为其论证的依据：

任给两个直线段 a, b ，不论前者有多短而后者有多长，总有足够大的整数 N 使得 $N \cdot a$ 比 b 长。

【定理】：设 $\{a, b\}$ 是不可公度者，对于任给正整数 n ，恒存在 m 使得

$$\frac{m}{n} < a:b < \frac{m+1}{n}$$

证明：由上述公设，必有足够大的 N 使得 $\frac{1}{n}b$ 的 N 倍要比 a 长。令 $m+1$ 为这种 N 之中的最小者，则有

$$m\left(\frac{1}{n}b\right) < a < (m+1)\left(\frac{1}{n}b\right)$$

亦即

$$\frac{m}{n} < a : b < \frac{m+1}{n} \quad \square$$

[注]: 因为 n 是可以任意大的, 所以上述左、右夹逼 $a : b$ 的两个分数之间的差额 $\frac{1}{n}$ 是可以小到任意小的。[用现代的术语, 即对于任给正数 $\varepsilon > 0$, 皆有足够大的 n 使得 $\frac{1}{n} < \varepsilon$ 。] 所以 $a : b$ 和 $\frac{m}{n}$ 或 $\frac{m+1}{n}$ 之间的差别当然也可以小到任意小。基于上述定理, 就可以进一步说明前述不可公度的「比值」之间大、小、相等关系的定义的必然性 :

设 $a : b$ 和 $a' : b'$ 对于任给分数恒具有相同的大小关系, 则对于任给 n , 都有相应的 m , 使得

$$\frac{m}{n} < a : b, a' : b' < \frac{m+1}{n}$$

因此 $a : b$ 和 $a' : b'$ 之间的差别要比所有 $\frac{1}{n}$ 都小。不论上述差别是那一种新量, 它是一个固定的量而它又比所有 $\frac{1}{n}$ 都小, 所以唯一的可能者就是零, 亦即 $a : b = a' : b'$ 。再者, 在 $a : b$ 和 $a' : b'$ 不等的情形, 则有一个分数, 它和两者的大小关系是不同的, 这也就是前述比较大小的定义。

有了上述思想和逼近法, 再进而重建当年希腊的定量几何学, 乃是顺理成章之事, 其要点在于原先仅仅对于可公度的情形具有证明的各种各样定理和公式, 作出其在不可公度的情形的「补证」。例如下述矩形公式 :

$$\square(a, b) : \square(u, u) = (a : u)(b : u)$$

在 $a : u$ 和 $b : u$ 都是分数时业已证明, 而在 $a : u$ 和 $b : u$ 至少有一个不是分数 (亦即不可公度) 时, 需要补证。

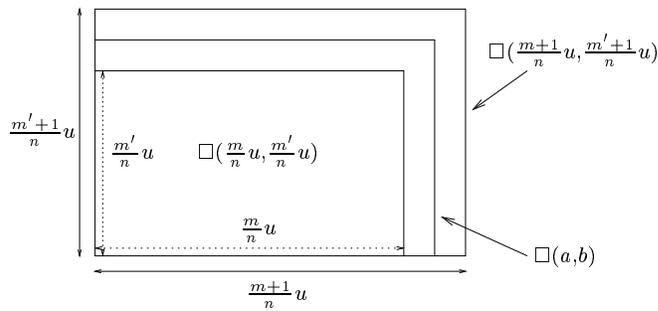
Eudoxus 对于上述矩形面积公式所作的补证, 大致如下 :

对于任给正整数 n , 不论它有多大, 皆有 m 和 m' 使得

$$\frac{m}{n} < a : u < \frac{m+1}{n}, \frac{m'}{n} < b : u < \frac{m'+1}{n}$$

亦即

$$\frac{m}{n}u < a < \frac{m+1}{n}u, \frac{m'}{n}u < b < \frac{m'+1}{n}u$$



[图 2-12]

如 [图 2-12] 所示， $\square(a, b)$ 包含 $\square(\frac{m}{n}u, \frac{m'}{n}u)$ ，而且它又包含于 $\square(\frac{m+1}{n}u, \frac{m'+1}{n}u)$ 之中，由此可得

$$\frac{mm'}{n^2} < \square(a, b) : \square(u, u) < \frac{(m+1)(m'+1)}{n^2}.$$

再者，由前述不等式的相乘，亦有

$$\frac{mm'}{n^2} < (a : u)(b : u) < \frac{(m+1)(m'+1)}{n^2}.$$

因此， $\square(a, b) : \square(u, u)$ 和 $(a : u)(b : u)$ 之间的差别（假如有的话）必然要比同时左、右夹逼两者的两个分数之间的差别要更小，即小于

$$\frac{(m+1)(m'+1)}{n^2} - \frac{mm'}{n^2} = \frac{m+m'+1}{n^2} = \frac{1}{n} \left(\frac{m}{n} + \frac{m'+1}{n} \right)$$

在 n 无限增大时，它是可以小到任意小的。所以 $\square(a, b) : \square(u, u)$ 和 $(a : u)(b : u)$ 之间不可能有任何差别，亦即所要补证的矩形面积公式

$$\square(a, b) : \square(u, u) = (a : u)(b : u) \quad \square$$

长话短说，Eudoxus 所创的逼近法不但把当年轻仅在可公度的特殊情形下具有其证明的各种各样定理和公式，加以明确简洁的「补证」，使得它们不论在可公度或不可公度的情形皆普遍成立，从而彻底重建了定量几何基础论。再者，他有鉴于曾经采用错误的「公设」作为几何学的论证依据的惨痛教训，决心下功夫彻底检查当代的几何学，尽其所知所能把其论证的依据，精简压缩到「至精至简」；流传至今的《欧氏几何学》(Euclidean Geometry) 其中绝大部分来自 Eudoxus 的

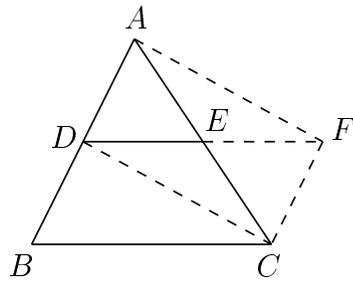
著作。所以「公理化」治学的典范和人类理性文明中的第一科学的初阶集其大成，实乃 Eudoxus（而并非 Euclid）的伟大贡献。不但此也，Eudoxus 的逼近原理和方法论不但重建了定量几何基础论，而且也是分析学 (Analysis) 的发祥地和基本方法。他本人就把它用来证明锥体体积等于三分之一底面积乘高这个立体几何基本公式，他的证法以及随后 Archimedes 把它拓展到球面面积公式和球体体积公式的论证乃是积分学的雏形和范例。

如今回顾反思，将中、西古文明的定量平面几何作一比较分析：两者所得的基本公式大致相同，亦即矩形的面积公式，勾股弦公式（亦即毕氏定理）和相似三角形的边长比例式，但是在基调和格局上则两者是迥然不同的。中国古代的工程师研讨几何是为了致用，是唯用是尚的；他们在基本测量公式的推导上善用面积，的确有其独到的长处，但是在对于空间本质理解的深度上，比之于古希腊几何学是的确瞠乎其后的了。究其原因，相信并非是在聪明才智上有任何差别，而是在格调、志趣和气概上有所分野！例如「可公度性」乃是一个纯理论性的问题；在实用的度量中，在力所能及的准确度之下的微量根本没有其实质意义，所以不存在不可公度这种问题。由此可见，在唯用是尚的格局下，根本是不会有此一问的，当然也不会有 Hippiasus 这种深深触及空间的连续性的发现和历经半世纪的奋斗才结晶而得出的 Eudoxus 逼近原理和方法论，是不？由此反思，同学们应该体认到局限中国古代几何的发展因素乃是：「唯用是尚，则难见精深，所及不远」；而古希腊几何学上的成功给全人类的启示与鼓舞则是：「若以理解大自然为志趣，并能世代相承、精益求精，则宇宙基本结构的至精至简、至善至美是可望可及的」。

2.5 例题和习题

【例题】：

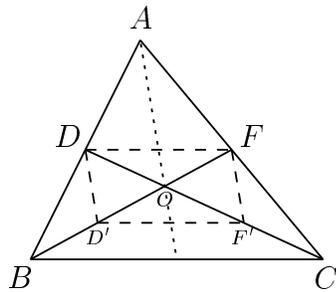
- (1) 令 D, E 分别为 $\triangle ABC$ 两边 $\overline{AB}, \overline{AC}$ 的中点，试证 $\overline{DE} = \frac{1}{2}\overline{BC}$ 而且 $\overline{DE} \parallel \overline{BC}$ 。



[图 2-13]

证明：如 [图 2-13] 所示，延长 \overline{DE} 至 \overline{DF} 使得 $\overline{EF} = \overline{DE}$ ，则 E 点平分 \overline{DF} 和 \overline{AC} ，所以 $\square ADCF$ 乃是一个平行四边形。由此可得 \overline{FC} 与 \overline{AD} 为平行等长，即 \overline{FC} 与 \overline{DB} 也为平行等长，因此 $\square BCFD$ 亦是一个平行四边形。所以 \overline{DF} 与 \overline{BC} 为平行等长，即得所求证。

- (2) 试证 $\triangle ABC$ 的三条中线共交于一点（称之为重心），而该点把每条中线均分成 2:1 的两段。



[图 2-14]

证明：令中线 \overline{BF} 和 \overline{CD} 之交点为 O 。取 D', F' 分别为 \overline{OB} 和 \overline{OC} 之中点，则有（见例题 (1)）：

$$\overline{D'F'} = \frac{1}{2}\overline{BC} \quad \text{而且} \quad \overline{D'F'} \parallel \overline{BC}$$

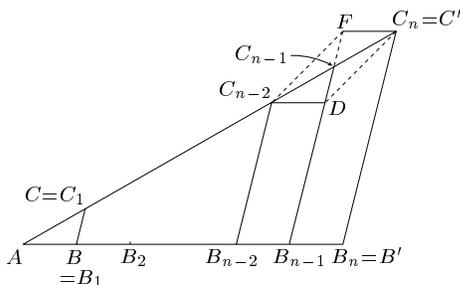
所以 $\overline{D'F'}$ 与 \overline{DF} 为平行等长，即 $\square D'F'FD$ 是一个平行四边形。所以

$$\overline{DO} = \overline{OF'}, \quad \overline{D'O} = \overline{OF}$$

即得所求证。

(3) 相似三角形定理在整数比时的古希腊证法：

如 [图 2-15] 所示， $B'C'$ 是分别位于 \overline{AB} 和 \overline{AC} 的延长线上之点，满足 $\overline{AB'} = n\overline{AB}$ ， $B'C' \parallel BC$ ， n 为正整数。试证 $\overline{AC'} = n\overline{AC}$ ， $\overline{B'C'} = n\overline{BC}$ 。



[图 2-15]

归纳证明：当 $n = 1$ 时结果是显然的。当 $n = 2$ 时乃是例题 (1) 之所证。兹对于 $n \geq 3$ 时作归纳证明如下：

以 \overline{AB} 之长度等分 $\overline{AB'}$ 为 n 段，令其等分点为

$$B = B_1, B_2, \dots, B_{n-2}, B_{n-1}, B' = B_n$$

过 B_{n-2} 及 B_{n-1} 作两条平行于 BC 的直线

$$B_{n-2}C_{n-2} \parallel BC, \quad B_{n-1}C_{n-1} \parallel BC$$

其中 C_{n-2} 及 C_{n-1} 为 AC' 上两点；再过 C_{n-2} 及 C' 作两条平行于 AB 的直线

$$C_{n-2}D \parallel AB, \quad FC' \parallel AB$$

其中 D 及 F 为 $B_{n-1}C_{n-1}$ 上两点，如 [图 2-15] 所示。

由所作易见 $\square B_{n-2}B_{n-1}DC_{n-2}$ 和 $\square B_{n-1}B'C'F$ 都是平行四边形，所以

$$\begin{aligned} \overline{C_{n-2}D} & \text{ 与 } \overline{B_{n-2}B_{n-1}} & \text{ 为平行等长,} \\ \overline{FC'} & \text{ 与 } \overline{B_{n-1}B'} & \text{ 为平行等长;} \end{aligned}$$

由所作亦有 $\overline{B_{n-2}B_{n-1}} = \overline{B_{n-1}B'}$ ，所以 $\overline{C_{n-2}D}$ 与 $\overline{FC'}$ 也为平行等长，即 $\square C_{n-2}DC'F$ 亦是平行四边形。因为平行四边形的两条对

角线互相平分，所以

$$\begin{aligned}\overline{C_{n-1}C'} &= \overline{C_{n-2}C_{n-1}}, \quad \text{和} \\ \overline{C_{n-1}F} &= \overline{DC_{n-1}}\end{aligned}$$

现用归纳假设可得

$$\begin{aligned}\overline{AC_{n-2}} &= (n-2)\overline{AC}, & \overline{B_{n-2}C_{n-2}} &= (n-2)\overline{BC} \\ \overline{AC_{n-1}} &= (n-1)\overline{AC}, & \overline{B_{n-1}C_{n-1}} &= (n-1)\overline{BC}\end{aligned}$$

所以

$$\begin{aligned}\overline{C_{n-2}C_{n-1}} &= \overline{AC_{n-1}} - \overline{AC_{n-2}} = (n-1)\overline{AC} - (n-2)\overline{AC} = \overline{AC} \\ \overline{DC_{n-1}} &= \overline{B_{n-1}C_{n-1}} - \overline{B_{n-1}D} = \overline{B_{n-1}C_{n-1}} - \overline{B_{n-2}C_{n-2}} \\ &= (n-1)\overline{BC} - (n-2)\overline{BC} = \overline{BC}\end{aligned}$$

亦即

$$\begin{aligned}\overline{AC'} &= \overline{AC_{n-1}} + \overline{C_{n-1}C'} = \overline{AC_{n-1}} + \overline{C_{n-2}C_{n-1}} \\ &= (n-1)\overline{AC} + \overline{AC} = n\overline{AC}\end{aligned}$$

和

$$\begin{aligned}\overline{B'C'} &= \overline{B_{n-1}F} = \overline{B_{n-1}C_{n-1}} + \overline{C_{n-1}F} \\ &= (n-1)\overline{BC} + \overline{DC_{n-1}} = (n-1)\overline{BC} + \overline{BC} = n\overline{BC}\end{aligned}$$

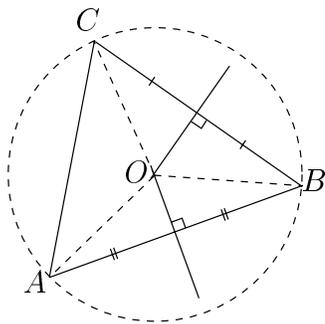
定理证毕。

- (4) 外接圆作图：对于一个给定的 $\triangle ABC$ ，唯一存在一个过其三顶点的圆，称之为 $\triangle ABC$ 的外接圆（如 [图 2-16] 所示）。其作图法如下：用 [基本作图 1.3]，分别作 \overline{AB} 和 \overline{BC} 的垂直平分线。则两线的交点 O 乃是具有和三顶点等距，即

$$\overline{OA} = \overline{OB} = \overline{OC}$$

的唯一之点，所以它就是所求作的外心（外接圆圆心）。

[注]: \overline{AB} 和 \overline{BC} 的垂直平分线只有在 A, B, C 三点共线时才不相交。

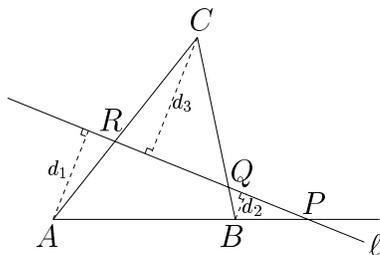


[图 2-16]

- (5) 试证 Menelous 定理: 设直线 l 与 $\triangle ABC$ 三边所在之直线 AB, BC 和 CA 分别相交于 P, Q, R (相异) 三点, 则下述有向长度比乘积条件式恒成立:

$$\frac{\overrightarrow{AP}}{\overrightarrow{PB}} \cdot \frac{\overrightarrow{BQ}}{\overrightarrow{QC}} \cdot \frac{\overrightarrow{CR}}{\overrightarrow{RA}} = -1$$

[Menelous 逆定理亦成立。证明留作习题。]



[图 2-17]

证明: 令 d_1, d_2, d_3 分别是顶点 A, B, C 到直线 l 之垂直距离, 则由相似三角形定理可得

$$\frac{\overrightarrow{AP}}{\overrightarrow{PB}} = -\frac{d_1}{d_2}, \quad \frac{\overrightarrow{BQ}}{\overrightarrow{QC}} = \frac{d_2}{d_3}, \quad \frac{\overrightarrow{CR}}{\overrightarrow{RA}} = \frac{d_3}{d_1}$$

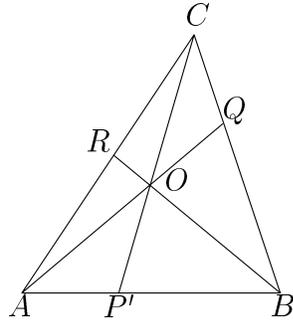
所以

$$\frac{\overrightarrow{AP}}{\overrightarrow{PB}} \cdot \frac{\overrightarrow{BQ}}{\overrightarrow{QC}} \cdot \frac{\overrightarrow{CR}}{\overrightarrow{RA}} \equiv -1$$

- (6) 试证 Ceva 定理：设 O 为 $\triangle ABC$ 内部一点， P', Q, R 分别为 AB 与 CO , BC 与 AO , CA 与 BO 之交点，则下述有向长度比乘积条件式恒成立：

$$\frac{\overrightarrow{AP'}}{\overrightarrow{P'B}} \cdot \frac{\overrightarrow{BQ}}{\overrightarrow{QC}} \cdot \frac{\overrightarrow{CR}}{\overrightarrow{RA}} = 1$$

[Ceva 逆定理亦成立。证明留作习题。]



[图 2-18]

证明：令 $\triangle OBC, \triangle OCA, \triangle OAB$ 的面积分别为 $\Delta_1, \Delta_2, \Delta_3$ 。因为 $\triangle CAP$ 与 $\triangle CPB$ 同高，所以其面积之比等于其底边边长之比，亦即

$$\frac{\Delta_{CAP}}{\Delta_{CPB}} = \frac{\overrightarrow{AP}}{\overrightarrow{PB}}$$

同理

$$\frac{\Delta_{OAP}}{\Delta_{OPB}} = \frac{\overrightarrow{AP}}{\overrightarrow{PB}}$$

所以

$$\frac{\Delta_2}{\Delta_1} = \frac{\Delta_{OAC}}{\Delta_{OBC}} = \frac{\overrightarrow{AP}}{\overrightarrow{PB}}$$

类似地可得

$$\frac{\Delta_3}{\Delta_2} = \frac{\overrightarrow{BQ}}{\overrightarrow{QC}}, \quad \frac{\Delta_1}{\Delta_3} = \frac{\overrightarrow{CR}}{\overrightarrow{RA}}$$

将三式相乘即为所求证。

[定义]：在直线上之四点列 P, A, P', B 称之为调和点列，记以

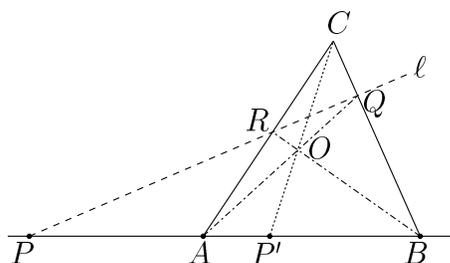
$(PP'; AB) = -1$, 若满足下述有向长度比的条件式:

$$\frac{\overrightarrow{PA}}{\overrightarrow{PB}} = -\frac{\overrightarrow{P'A}}{\overrightarrow{P'B}} \Leftrightarrow \frac{\overrightarrow{AP}}{\overrightarrow{PB}} \cdot \frac{\overrightarrow{BP'}}{\overrightarrow{P'A}} = -1$$

亦即 P, P' 两点以同等比例分割线段 \overline{AB} (一在外、一在内)。



- (7) 调和点列作图法: 对给定已共线三点 P, A, B , 求作 P' 点使得 P, A, P', B 成调和点列。



[图 2-19]

[作法] 用已给线段 \overline{AB} 为一边, 任选线外一点 C 构作三角形 $\triangle ABC$ 。过 P 点作任意的直线 l , 使得 l 与 $\overline{BC}, \overline{CA}$ 分别交于 Q, R 两点。连结 AQ, BR , 设两者相交于 O 点, 则 CO 与 \overline{AB} 就会交于所求作之 P' 点。

证明: 直接运用 Menelous 定理和 Ceva 定理!

$$\begin{aligned} \text{Menelous 定理: } & \frac{\overrightarrow{AP}}{\overrightarrow{PB}} \cdot \frac{\overrightarrow{BQ}}{\overrightarrow{QC}} \cdot \frac{\overrightarrow{CR}}{\overrightarrow{RA}} = -1 \\ \text{Ceva 定理: } & \frac{\overrightarrow{AP'}}{\overrightarrow{P'B}} \cdot \frac{\overrightarrow{BQ}}{\overrightarrow{QC}} \cdot \frac{\overrightarrow{CR}}{\overrightarrow{RA}} = 1 \end{aligned}$$

两式相除即得所需的有向长度比例式。

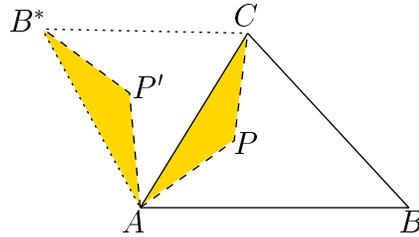
- (8) Steiner 点: 设 $\triangle ABC$ 的三内角皆小于 120° (即小于 $\frac{2\pi}{3}$), 则

$$\overline{AP} + \overline{BP} + \overline{CP}$$

有一个唯一的极小值，其解点 P_0 乃是使得

$$\angle AP_0B = \angle BP_0C = \angle CP_0A = 120^\circ$$

之点。



[图 2-20]

如[图 2-20]所示，作 B^* 使得 $\triangle ACB^*$ 为等边三角形。设 P 是平面上任给一点，令 $\triangle AP'B^*$ 为 $\triangle APC$ 绕 A 点旋转 60° 之所得者，则有 $\triangle APP'$ 为等边三角形，而且

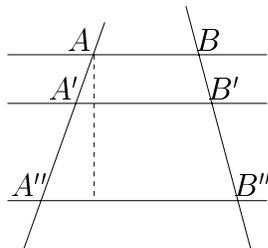
$$\overline{CP} + \overline{AP} + \overline{BP} = \overline{B^*P'} + \overline{P'P} + \overline{PB}$$

由此可见，上述总长在 P, P' 皆位于 $\overline{BB^*}$ 之上为极小。类似地定义 A^* 和 C^* 点，则 P_0 应该就是 $\overline{AA^*}, \overline{BB^*}, \overline{CC^*}$ 的共交点，而所求证者则显而易见了。

【习题】：

- (1) 以矩形面积公式为基础，试证：平行四边形面积 = 底乘高。
- (2) 试证三角形面积公式为 $A(\triangle) = \frac{1}{2}bh$ 。
- (3) 试证平行线的截割保持线段之比；即如 [图 2-21] 所示，求证：

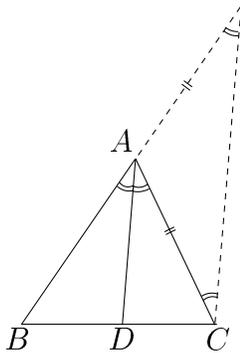
$$\text{当 } AB \parallel A'B' \parallel A''B'', \text{ 则 } \frac{\overline{AA'}}{\overline{A'A''}} = \frac{\overline{BB'}}{\overline{B'B''}}$$



[图 2-21]

(4) 如 [图 2-22] 所示, AD 乃是 $\angle A$ 的分角线。试证

$$\frac{\overline{BD}}{\overline{DC}} = \frac{\overline{AB}}{\overline{AC}}$$



[图 2-22]

(5) 试将一个给定线段 \overline{AB} 等分为 n 段, n 为某一正整数。

(6) 试证明相似三角形定理在分数比的情形。

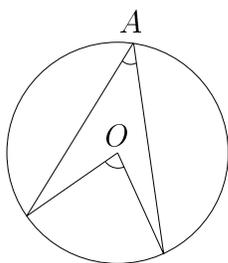
(7) 试证 Menelous 逆定理: 设 (相异) 三点 P, Q, R 分别在 $\triangle ABC$ 三边 AB, BC, CA 之上并满足下述有向长度比的条件式, 则 P, Q, R 三点共线。

$$\frac{\overrightarrow{AP}}{\overrightarrow{PB}} \cdot \frac{\overrightarrow{BQ}}{\overrightarrow{QC}} \cdot \frac{\overrightarrow{CR}}{\overrightarrow{RA}} = -1$$

(8) 试证 Ceva 逆定理: 设 P', Q, R 分别为 $\triangle ABC$ 的三边 $\overline{AB}, \overline{BC}, \overline{CA}$ 上的三点并满足下述有向长度比的条件式, 则 CP', AQ, BR 三线共点。

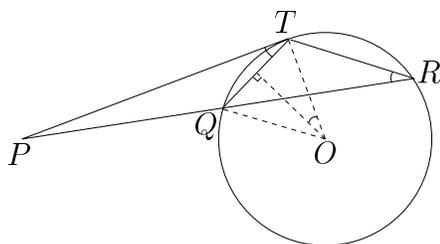
$$\frac{\overrightarrow{AP'}}{\overrightarrow{P'B}} \cdot \frac{\overrightarrow{BQ}}{\overrightarrow{QC}} \cdot \frac{\overrightarrow{CR}}{\overrightarrow{RA}} = 1$$

(9) 试证圆心角为圆周角之两倍: $\angle O = 2\angle A$ 。



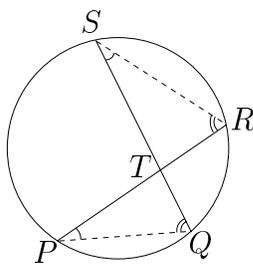
[图 2-23]

- (10) 在 [图 2-24] 中, PQR 为一直线, PT 与圆相切于 T 点。试证:
 $\overline{PQ} \cdot \overline{PR} = \overline{PT}^2$ 。



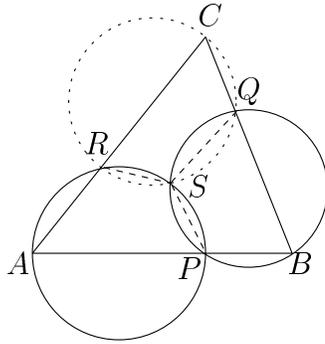
[图 2-24]

- (11) 设 P, Q, R, S 为圆上四点, 而 \overline{PR} 与 \overline{QS} 则相交于圆内 (或圆外) T 点。试证: $\overline{PT} \cdot \overline{TR} = \overline{QT} \cdot \overline{TS}$ 。



[图 2-25]

- (12) 试证 Miquel 定理: 在 $\triangle ABC$ 的三边 \overline{AB} , \overline{BC} , \overline{CA} 上分别取三点 P, Q, R , 过 P, B, Q 及 P, R, A 分别作圆, 令两圆交点为 S (与 P), 则 S, Q, C, R 四点共圆。



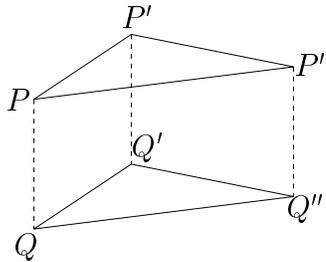
[图 2-26]

(13) 若平面中的一个变换 τ 满足条件

$$\overrightarrow{P\tau(P)} \text{ 和 } \overrightarrow{Q\tau(Q)} \text{ 恒为同向平行且等长}$$

则称 τ 为平面上一个平移 (translation)。试证平移之组合仍是平移，即验证：

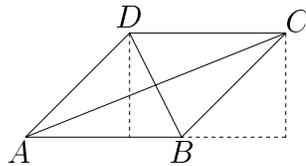
若 $\overrightarrow{PP'}$ 和 $\overrightarrow{QQ'}$ 为同向平行且等长，
 和 $\overrightarrow{P'P''}$ 和 $\overrightarrow{Q'Q''}$ 为同向平行且等长；
 则必有 $\overrightarrow{PP''}$ 和 $\overrightarrow{QQ''}$ 亦为同向平行且等长。



[图 2-27]

(18) 试证广义勾股定理：设 \overline{AB} 与 \overline{CD} 为平行等长，则：

$$\overline{AC}^2 + \overline{BD}^2 = \overline{AB}^2 + \overline{BC}^2 + \overline{CD}^2 + \overline{DA}^2$$



[图 2-28]

第三章

圆与三角学

在各种各样的平面形之中，圆是最为完美对称者，而三角形则是最为简单者。所以在平面几何的研讨中，圆和三角形理所当然地是其精要之所在。例如定量平面几何中的基本定理，首推三角形的面积公式、相似三角形定理和勾股定理，即

- 面积公式：三角形面积 = $\frac{1}{2}$ 底 \times 高
- 相似三角形定理：设 $\triangle ABC$ 和 $\triangle A'B'C'$ 的三内角对应相等，则其三对对应边成比例，即

$$\frac{\overline{AB}}{\overline{A'B'}} = \frac{\overline{AC}}{\overline{A'C'}} = \frac{\overline{BC}}{\overline{B'C'}} = k \quad (k: \text{相似比})$$

- 勾股定理：直角三角形的边长满足

$$\overline{AC}^2 + \overline{BC}^2 = \overline{AB}^2 \quad (\text{亦即毕氏定理})$$

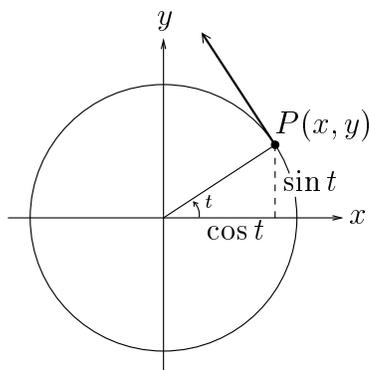
(勾方加股方等于弦方)

本章将以上述三者为基础，研讨圆与三角形的解析几何，其所得之基础理论也就是三角函数的基本性质和三角定律。正弦、余弦函数是一起源于圆周运动，密切配合的周期函数，它们是解析几何学和周期函数的分析学中最为基本和重要的函数；而正弦、余弦函数的基本性质乃是圆的几何性质（主要是其对称性）的直接反映。

三角学 (Trigonometry) 所讨论的课题是三角形的各种各样几何量之间的函数关联。由此可见，三角学其实就是三角形的解析几何，可以说是具体而微的解析几何；它是整个平面解析几何的基础所在，也是用解析法系统研究几何的基本公具。

3.1 正弦、余弦函数的基本性质

如 [图 3-1] 所示，设 $P(x, y)$ 是在单位圆上，以 $(1, 0)$ 为起点作逆时针方向的单位速率运动的动点，则它的 x, y 坐标乃是时间 t 的函数，分别定义为余弦函数 $\cos t$ 和正弦函数 $\sin t$ 。



[图 3-1]

其实， $x = \cos t$ 和 $y = \sin t$ 乃是单位圆的自然的动态（解析）描述。由此可以想到，正弦、余弦函数的基本性质乃是圆的几何性质（主要是对称性）的解析表述。例如

$$1. \overline{OP}^2 = 1 \Leftrightarrow \cos^2 t + \sin^2 t = 1 \quad (\text{勾股定理})$$

2. 圆周周长 $= 2\pi \Leftrightarrow$ 周期性：

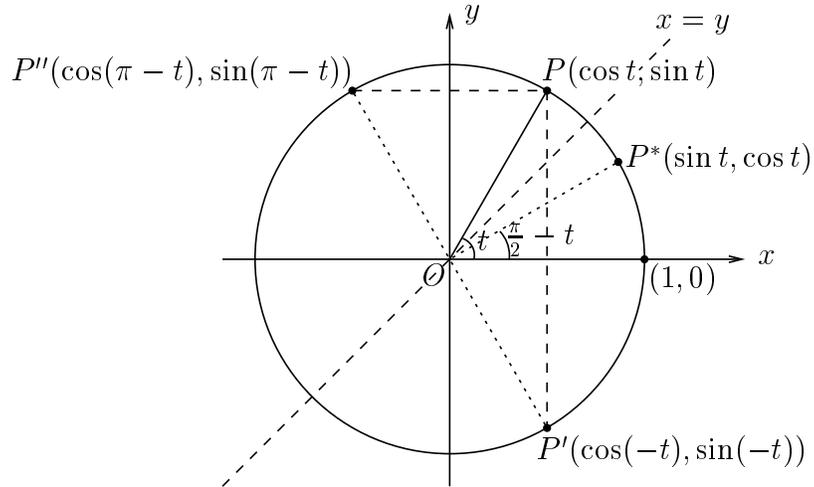
$$(3.1) \quad \begin{cases} \cos(2\pi + t) = \cos t \\ \sin(2\pi + t) = \sin t \end{cases}$$

3. 对于 x -轴（或 y -轴）的反射对称性（参看 [图 3-2]）

$$(3.2) \quad \Leftrightarrow \begin{cases} \cos(-t) = \cos t, & \sin(-t) = -\sin t \\ \cos(\pi - t) = -\cos t, & \sin(\pi - t) = \sin t \end{cases}$$

4. 对于直线 $x = y$ 的反射对称性 (参看 [图 3-2]) : $(x, y) \leftrightarrow (y, x)$

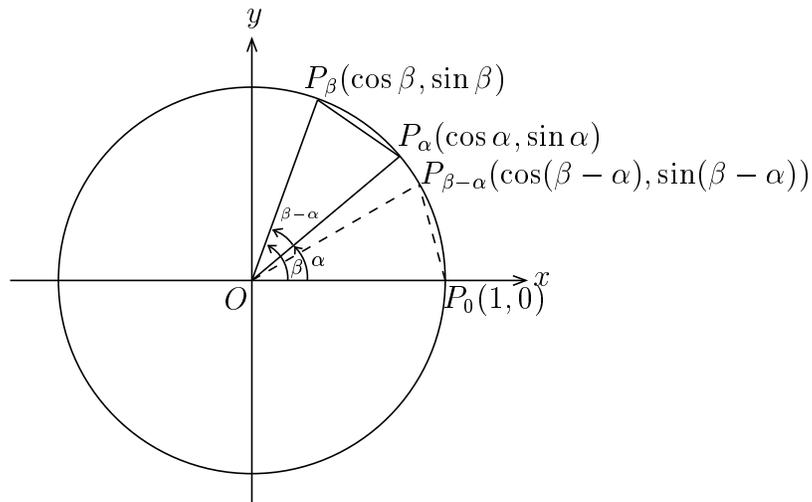
$$(3.3) \quad \Leftrightarrow \sin\left(\frac{\pi}{2} - t\right) = \cos t, \quad \cos\left(\frac{\pi}{2} - t\right) = \sin t$$



[图 3-2]

5. 圆的旋转对称性 \Leftrightarrow 复角公式 :

$$(3.4) \quad \cos(\beta - \alpha) = \cos \beta \cos \alpha + \sin \beta \sin \alpha$$



[图 3-3]

如 [图 3-3] 所示, $\triangle OP_\alpha P_\beta$ 乃是 $\triangle OP_0 P_{\beta-\alpha}$ 旋转 α 角之所得, 所以当然有 $\overline{P_\alpha P_\beta}^2 = \overline{P_0 P_{\beta-\alpha}}^2$, 即有

$$(\cos \beta - \cos \alpha)^2 + (\sin \beta - \sin \alpha)^2 = (\cos(\beta - \alpha) - 1)^2 + \sin^2(\beta - \alpha)$$

亦即

$$\begin{aligned} & \cos^2 \beta + \sin^2 \beta + \cos^2 \alpha + \sin^2 \alpha - 2(\cos \beta \cos \alpha + \sin \beta \sin \alpha) \\ (3.5) \quad & = \cos^2(\beta - \alpha) + \sin^2(\beta - \alpha) + 1 - 2 \cos(\beta - \alpha) \\ & \Rightarrow \cos(\beta - \alpha) = \cos \beta \cos \alpha + \sin \beta \sin \alpha \end{aligned}$$

把 (3.4)-式和 (3.2)-式结合, 即得

$$(3.6) \quad \begin{aligned} \cos(\beta + \alpha) &= \cos(\beta - (-\alpha)) \\ &= \cos \beta \cos \alpha - \sin \beta \sin \alpha \end{aligned}$$

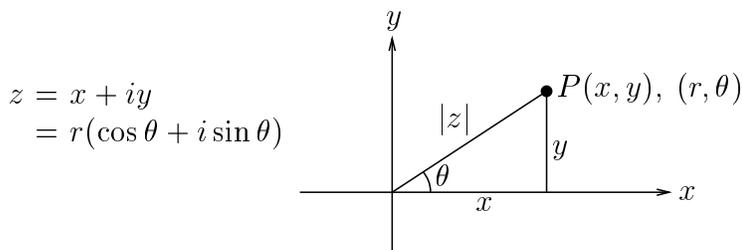
再把 (3.6)-式和 (3.3)-式相结合, 即得

$$(3.7) \quad \begin{aligned} \sin(\alpha + \beta) &= \cos\left[\left(\frac{\pi}{2} - \alpha\right) - \beta\right] \\ &= \sin \alpha \cos \beta + \cos \alpha \sin \beta \end{aligned}$$

我们还可以把 (3.6)-式和 (3.7)-式用复数的乘法组合成下述更加整齐的复值形式, 即

$$(3.8) \quad (\cos \alpha + i \sin \alpha)(\cos \beta + i \sin \beta) = \cos(\alpha + \beta) + i \sin(\alpha + \beta)$$

再者, 我们可以把 $z = x + iy$ 想为平面上 $P(x, y)$ 点的复数坐标 (complex coordinate)。如 [图 3-4] 所示, P 点的极坐标 $\overline{OP} = r = \sqrt{x^2 + y^2}$ 和 θ 分别就是 z 的绝对值和幅角。



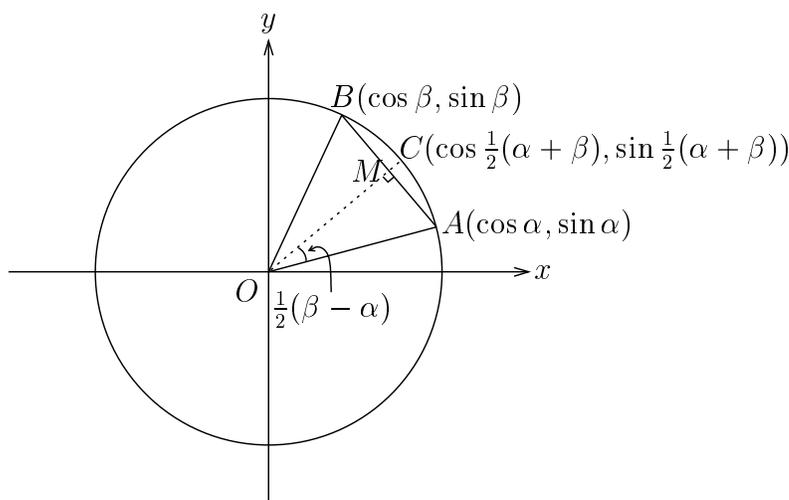
[图 3-4]

将 (3.8)-式用来表达复数的乘法，即有

$$(3.9) \quad \begin{aligned} z_1 \cdot z_2 &= |z_1|(\cos \theta_1 + i \sin \theta_1) \cdot |z_2|(\cos \theta_2 + i \sin \theta_2) \\ &= |z_1| \cdot |z_2|(\cos(\theta_1 + \theta_2) + i \sin(\theta_1 + \theta_2)) \end{aligned}$$

亦即两个复数 z_1, z_2 相乘，其绝对值相乘而其幅角则相加。此事在研讨复数时具有基本的重要性。

6. 和化积公式和反射对称性



[图 3-5]

如 [图 3-5] 所示，等腰三角形 $\triangle OAB$ 对于 OM 成反射对称。所以 $\angle AOM = \frac{1}{2}(\beta - \alpha)$, $\angle xOC = \frac{1}{2}(\alpha + \beta)$ 。即有

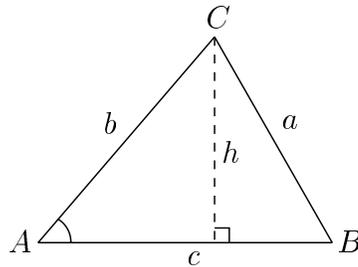
$$(3.10) \quad \begin{aligned} M &= \left(\frac{1}{2}(\cos \alpha + \cos \beta), \frac{1}{2}(\sin \alpha + \sin \beta)\right) \\ C &= \left(\cos \frac{1}{2}(\alpha + \beta), \sin \frac{1}{2}(\alpha + \beta)\right) \\ \overline{OM} &= \cos \frac{1}{2}(\alpha - \beta) \end{aligned}$$

$$(3.11) \quad \Rightarrow \begin{cases} \frac{1}{2}(\cos \alpha + \cos \beta) = \cos \frac{1}{2}(\alpha - \beta) \cos \frac{1}{2}(\alpha + \beta) \\ \frac{1}{2}(\sin \alpha + \sin \beta) = \cos \frac{1}{2}(\alpha - \beta) \sin \frac{1}{2}(\alpha + \beta) \end{cases}$$

3.2 三角定律

一个三角形 $\triangle ABC$ 含有各种各样的几何量，例如它的三边边长、三个内角的角度、面积、外径（外接圆的半径）和内径（内切圆的半径）等等。而它们之间，又存在着各种各样的函数关系。本节所要研究者，乃是它们之间的基本函数关系，通称之为三角定律。

1. 三角形面积公式与正弦定律



[图 3-6]

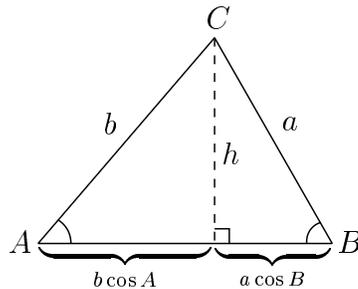
如 [图 3-6] 所示，我们将以 a, b, c 分别表示角 A, B, C 的对边边长， Δ 表示其面积。易见 $h = b \sin A$ ，所以

$$(3.12) \quad \Delta = \frac{1}{2}c \cdot h = \frac{1}{2}bc \sin A$$

同理： $\Delta = \frac{1}{2}ca \sin B = \frac{1}{2}ab \sin C$ 。由此即得下述正弦定律：

$$(3.13) \quad \frac{\sin A}{a} = \frac{\sin B}{b} = \frac{\sin C}{c} = \frac{2\Delta}{abc}$$

2. 垂直投影与余弦定律



[图 3-7]

由 [图 3-7] 和 Cosine 的定义，即有

$$(3.14) \quad \left. \begin{array}{l} b \cos A + a \cos B = c \\ \text{同理：} \quad c \cos B + b \cos C = a \\ \quad \quad \quad a \cos C + c \cos A = b \end{array} \right\}$$

由上述 $\{\cos A, \cos B, \cos C\}$ 的线性方程组即可解得

$$(3.14') \quad \left\{ \begin{array}{l} \cos A = \frac{b^2 + c^2 - a^2}{2bc} \\ \cos B = \frac{c^2 + a^2 - b^2}{2ac} \\ \cos C = \frac{a^2 + b^2 - c^2}{2ab} \end{array} \right. \quad (\text{余弦定律})$$

[注]：S.S.S. 叠合条件的几何意义是 $\triangle ABC$ 的三边边长业已唯一地确定了它的三个内角。换句话说，其三个内角分别是它的三边边长的函数。上述余弦定律给出了它们的具体表达式，亦即

$$A = \cos^{-1} \left\{ \frac{b^2 + c^2 - a^2}{2bc} \right\}, \quad \text{等等}$$

同样的，三角形的一组叠合条件如 S.A.S., A.S.A. 的几何意义其实也就是三角形的其他变量都可以用这样所给的一组自变元加以表达。[参看习题 (8) 和 (9)。]

3. 正弦定律之第二证法

我们也可以用余弦定律来推导正弦定律，即

$$(3.15) \quad \begin{aligned} \frac{\sin^2 A}{a^2} &= \frac{1 - \cos^2 A}{a^2} \\ &= \frac{1}{4a^2b^2c^2} \{4b^2c^2 - (b^2 + c^2 - a^2)^2\} \\ &= \frac{1}{4a^2b^2c^2} \{2(a^2b^2 + b^2c^2 + c^2a^2) - (a^4 + b^4 + c^4)\} \end{aligned}$$

因为上式右侧是 a, b, c 的对称式，所以

$$(3.16) \quad \frac{\sin^2 B}{b^2} = \frac{\sin^2 C}{c^2} = \frac{\sin^2 A}{a^2}$$

而 $\frac{\sin A}{a}, \frac{\sin B}{b}, \frac{\sin C}{c}$ 都是恒正的, 所以由 (3.15)-式和 (3.16)-式可以推论

$$(3.16') \quad \frac{\sin A}{a} = \frac{\sin B}{b} = \frac{\sin C}{c} = \frac{\{2\sum a^2b^2 - \sum a^4\}^{\frac{1}{2}}}{2abc}$$

其中 $\sum a^2b^2 = a^2b^2 + b^2c^2 + c^2a^2$, $\sum a^4 = a^4 + b^4 + c^4$ 乃是常用的简约写法。再将 (3.16')-式和 (3.13)-式相对比, 即得

$$(3.17) \quad 16\Delta^2 = 2\sum a^2b^2 - \sum a^4$$

其实, 上式之右侧是可以分解成四个一次因式的乘积者, 而且此事可以从一个简单的几何常识推论而知, 亦即三角形的三边边长中, 若有一为其他两者之和, 则其面积为零。亦即

$$\begin{aligned} a + b - c = 0 \quad \text{或} \quad a - b + c = 0 \quad \text{或} \quad -a + b + c = 0 \\ \Rightarrow \quad 2\sum a^2b^2 - \sum a^4 = 0 \end{aligned}$$

将上述事实 and 余式定理相结合, 即可推论它含有因式

$$(a + b - c)(a - b + c)(-a + b + c)$$

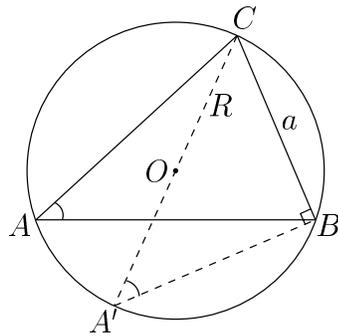
至此即可直接验证

$$(3.17') \quad 16\Delta^2 = (a + b + c)(a + b - c)(a - b + c)(-a + b + c)$$

通常把它改写成

$$(3.17'') \quad \Delta = \sqrt{s(s-a)(s-b)(s-c)}, \quad s = \frac{1}{2}(a + b + c)$$

此式在西方远在公元前三世纪已由 Archimedes 所求得, 但是因为讹传而称之为 Heron's formula. 而在中国南宋时期, 秦九韶也独立地求得此一公式。

4. 正弦定律的第三证法

[图 3-8]

如 [图 3-8] 所示， O 是 $\triangle ABC$ 的外接圆之圆心， R 是其半径， $\overline{CA'}$ 则是一条直径。由熟知的圆周角等于圆心角之半可见 $\angle A' = \angle A$ 而且 $\triangle A'BC$ 是直角三角形。所以

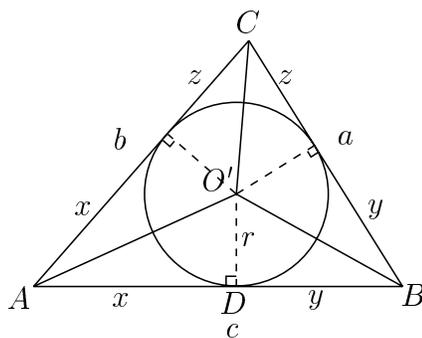
$$(3.18) \quad \sin A = \sin A' = \frac{a}{2R}$$

亦即

$$(3.19) \quad \frac{\sin A}{a} = \frac{\sin B}{b} = \frac{\sin C}{c} = \frac{1}{2R}$$

将 (3.19)-式和 (3.13)-式相比，即得外径 R 的公式

$$(3.20) \quad R = \frac{abc}{4\Delta}$$

5. 内切圆半径和半角公式

[图 3-9]

如 [图 3-9] 所示, O' 是 $\triangle ABC$ 的内切圆之圆心, 它是三个内角分角线的共交之点。再者 $\triangle O'AB$, $\triangle O'BC$ 和 $\triangle O'CA$ 的高都是内径 r 。即有

$$\begin{aligned} \triangle ABC &= \triangle O'AB + \triangle O'BC + \triangle O'CA \\ (3.21) \quad &= \frac{1}{2}rc + \frac{1}{2}ra + \frac{1}{2}rb = r \cdot s \end{aligned}$$

由此即得

$$(3.22) \quad r = \frac{\Delta}{s} = \sqrt{\frac{(s-a)(s-b)(s-c)}{s}}$$

再者, 如 [图 3-9] 所示 (用熟知的切线长相等)

$$(3.23) \quad x + y = c, \quad y + z = a, \quad z + x = b$$

解之即得

$$(3.23') \quad x = s - a, \quad y = s - b, \quad z = s - c$$

再由直角三角形 $\triangle ADO'$ 就得出正切的半角公式:

$$(3.24) \quad \tan \frac{A}{2} = \frac{r}{s-a} = \sqrt{\frac{(s-b)(s-c)}{s(s-a)}}$$

同理亦有

$$(3.24') \quad \begin{aligned} \tan \frac{B}{2} &= \sqrt{\frac{(s-a)(s-c)}{s(s-b)}} \\ \tan \frac{C}{2} &= \sqrt{\frac{(s-a)(s-b)}{s(s-c)}} \end{aligned}$$

在此, 当可顺水推舟, 轻而易举地推导正弦、余弦的半角公式如下。由

$$(3.25) \quad \tan \frac{A}{2} = \frac{\sin \frac{A}{2}}{\cos \frac{A}{2}} = \sqrt{\frac{(s-b)(s-c)}{s(s-a)}}$$

即得

$$(3.25') \quad \frac{\sin \frac{A}{2}}{\sqrt{(s-b)(s-c)}} = \frac{\cos \frac{A}{2}}{\sqrt{s(s-a)}} = k \quad (\text{亦即令其为 } k)$$

再用面积公式

$$\begin{aligned}
 \Delta &= \frac{1}{2}bc \sin A = bc \sin \frac{A}{2} \cos \frac{A}{2} \\
 (3.26) \quad &= bc \cdot k^2 \sqrt{(s-b)(s-a)} \cdot \sqrt{s(s-a)} = bc \cdot k^2 \cdot \Delta \\
 &\Rightarrow k^2 = \frac{1}{bc}, \quad k = \frac{1}{\sqrt{bc}}
 \end{aligned}$$

所以即得

$$(3.27) \quad \sin \frac{A}{2} = \sqrt{\frac{(s-b)(s-c)}{bc}}, \quad \cos \frac{A}{2} = \sqrt{\frac{s(s-a)}{bc}}$$

同理亦有

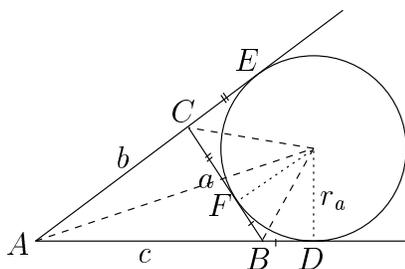
$$\begin{aligned}
 (3.27') \quad \sin \frac{B}{2} &= \sqrt{\frac{(s-c)(s-a)}{ca}}, \quad \cos \frac{B}{2} = \sqrt{\frac{s(s-b)}{ca}} \\
 \sin \frac{C}{2} &= \sqrt{\frac{(s-a)(s-b)}{ab}}, \quad \cos \frac{C}{2} = \sqrt{\frac{s(s-c)}{ab}}
 \end{aligned}$$

6. 傍心圆及其半径公式

傍心圆是于 $\triangle ABC$ 外与三角形的一边及另外两边的延线相切的圆，如 [图 3-10] 所示。其半径 r_a 可由下述公式求得：

$$r_a = s \cdot \tan \frac{A}{2}$$

[其余两者也可类似地求得。]



[图 3-10]

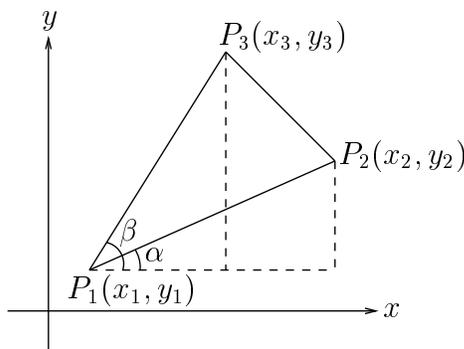
因 \overline{AD} 与 \overline{AE} 同是由 A 点到傍心圆之切线, 易见它们应该等长; 同理可知 $\overline{BD} = \overline{BF}$, $\overline{CE} = \overline{CF}$ 。考虑:

$$\begin{aligned} c + \overline{BD} &= \overline{AD} = \overline{AE} = b + \overline{CE}; \\ 2\overline{AD} &= c + \overline{BD} + b + \overline{CE} \\ &= b + c + a = 2s \end{aligned}$$

所以 $\overline{AD} = s$, 即有:

$$\tan \frac{A}{2} = \frac{r_a}{s}, \quad r_a = s \cdot \tan \frac{A}{2}$$

7. 三角形面积的坐标公式



[图 3-11]

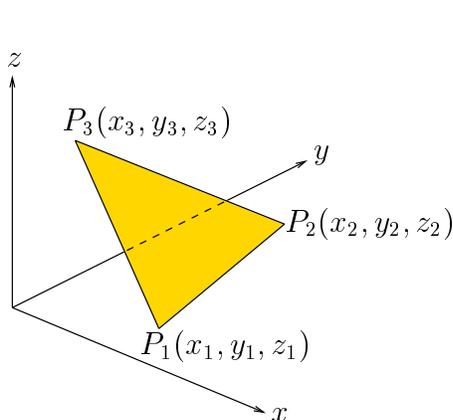
如 [图 3-11] 所示

$$(3.28) \quad \begin{aligned} \overline{P_1P_2} \cos \alpha &= x_2 - x_1, & \overline{P_1P_2} \sin \alpha &= y_2 - y_1 \\ \overline{P_1P_3} \cos \beta &= x_3 - x_1, & \overline{P_1P_3} \sin \beta &= y_3 - y_1 \end{aligned}$$

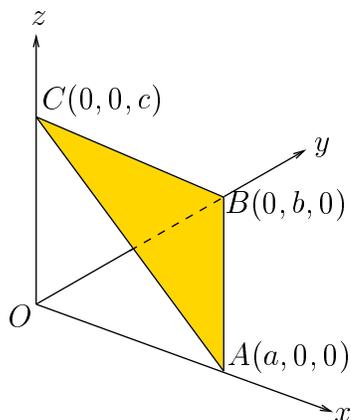
再者, $\triangle P_1P_2P_3$ 的定向面积

$$\begin{aligned} \Delta P_1P_2P_3 &= \frac{1}{2} \overline{P_1P_2} \cdot \overline{P_1P_3} \sin(\beta - \alpha) \\ &= \frac{1}{2} \overline{P_1P_2} \cdot \overline{P_1P_3} (\sin \beta \cos \alpha - \cos \beta \sin \alpha) \\ &= \frac{1}{2} \left\{ (x_2 - x_1)(y_3 - y_1) - (x_3 - x_1)(y_2 - y_1) \right\} \\ &= \frac{1}{2} \begin{vmatrix} x_2 - x_1 & y_2 - y_1 \\ x_3 - x_1 & y_3 - y_1 \end{vmatrix} = \frac{1}{2} \begin{vmatrix} 1 & x_1 & y_1 \\ 1 & x_2 & y_2 \\ 1 & x_3 & y_3 \end{vmatrix} \end{aligned}$$

在此，我们不妨继续探讨空间三点 $P_i(x_i, y_i, z_i)$, $1 \leq i \leq 3$, 所定的三角形面积应该如何去用其 9 个坐标 $\{x_i, y_i, z_i, 1 \leq i \leq 3\}$ 加以表达呢？



[图 3-12]

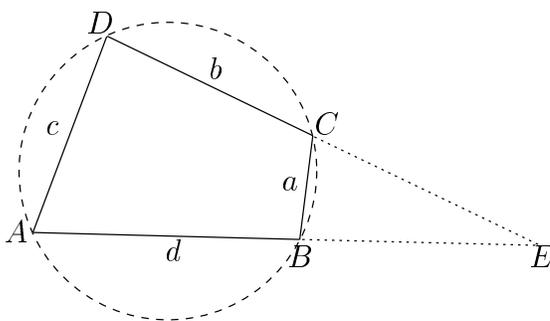


[图 3-12']

此事也许应该从某种特殊情形的研讨著手。例如 [图 3-12'] 所示者，我们将 $(\triangle ABC)^2$ 的计算留作习题，并试将所算得的结果给以几何解释。

8. 圆内接四边形面积公式：令 $\square ABCD$ 为圆内接四边形，其四边边长分别为 a, b, c, d 。令 $s = \frac{1}{2}(a + b + c + d)$ ，则 $\square ABCD$ 面积为

$$\sqrt{(s-a)(s-b)(s-c)(s-d)}$$



[图 3-13]

若 $\square ABCD$ 是平行四边形，则它必定是矩形，上述面积公式显然成立。现不妨假设 AB 交 DC 于 E 点，则易见有 $\triangle AED \sim \triangle CEB$ ，所以有面积比

$$\square ABCD : \triangle AED = \frac{c^2 - a^2}{c^2}$$

再者，令 $\overline{DE} = b'$, $\overline{AE} = d'$ ，则由相似三角形定理即有下述比例式：

$$\frac{b'}{c} = \frac{d' - d}{a}, \quad \frac{d'}{c} = \frac{b' - b}{a}$$

由两式之和及差即有

$$b' + d' = \frac{c}{c - a}(b + d), \quad b' - d' = \frac{c}{c + a}(b - d)$$

再者，将下述 $\triangle AED$ 的秦九韶公式

$$\triangle AED \text{ 面积} = \frac{1}{4} \sqrt{(b' + d' + c)(b' + d' - c)(b' - d' + c)(-b' + d' + c)}$$

代换入上面所得的 $b' + d'$ 和 $b' - d'$ 式然後简化：

$$\begin{aligned} b' + d' + c &= \frac{c}{c - a}(b + d + c - a) = \frac{c}{c - a}(2s - 2a) \\ b' + d' - c &= \frac{c}{c - a}(b + d - c + a) = \frac{c}{c - a}(2s - 2c) \\ b' - d' + c &= \frac{c}{c + a}(b - d + c + a) = \frac{c}{c + a}(2s - 2d) \\ -b' + d' + c &= \frac{c}{c + a}(-b + d + c + a) = \frac{c}{c + a}(2s - 2b) \end{aligned}$$

由此可得

$$\triangle AED \text{ 面积} = \frac{c^2}{c^2 - a^2} \sqrt{(s - a)(s - b)(s - c)(s - d)}$$

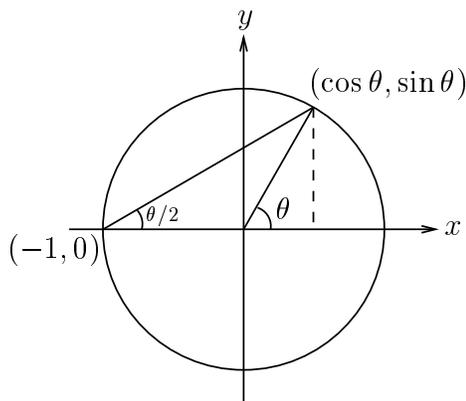
即得所求证者。

3.3 习题

(1) 证用正弦、余弦的复角公式推导出下述积化和公式，即

$$\begin{aligned} \cos A \cdot \cos B &= \frac{1}{2}(\cos(A + B) + \cos(A - B)) \\ \sin A \cdot \sin B &= \frac{1}{2}(\cos(A - B) - \cos(A + B)) \\ \sin A \cdot \cos B &= \frac{1}{2}(\sin(A + B) + \sin(A - B)) \end{aligned}$$

- (2) 在上式中, 令 $A+B = \alpha$, $A-B = \beta$ 。易见 $A = \frac{1}{2}(\alpha+\beta)$, $B = \frac{1}{2}(\alpha-\beta)$, 用之即可把习题 (1) 中的积化和公式转化成和化积公式。再者, 试用类似于 [图 3-5] 者, 说明它们和圆的反射对称之间的关系。
- (3) 试求用 $\cos \theta$ 分别表达 $\cos \frac{\theta}{2}$ 和 $\sin \frac{\theta}{2}$ 的公式 (亦即半角公式)。
- (4) 试用下述图解求
- (i) 用 $\cos \theta$ 和 $\sin \theta$ 表达 $\tan \frac{\theta}{2}$ 的公式。
- (ii) 用 $\tan \frac{\theta}{2}$ 表达 $\cos \theta$ 和 $\sin \theta$ 的公式。



[图 3-14]

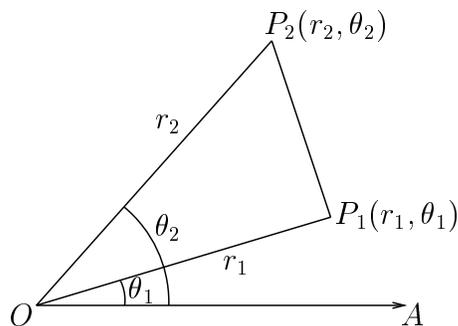
- (5) 试用二项定理和

$$(\cos \theta + i \sin \theta)^n = \cos n\theta + i \sin n\theta$$

求得正弦、余弦 n -倍角公式。

- (6) 试求 $z^n - 1 = 0$ 的所有复数解。
- (7) 试求 $z^n - (1+i) = 0$ 的所有复数解。
- (8) 试求用 $\triangle ABC$ 的两边一夹角 $\{a, b, C\}$ 表达其他一边和两角 (即 c 和 A, B) 之公式。
- (9) 试求用 $\triangle ABC$ 的两角一夹边 $\{A, B, c\}$ 表达其他一角和两边 (即 C 和 a, b) 之公式。

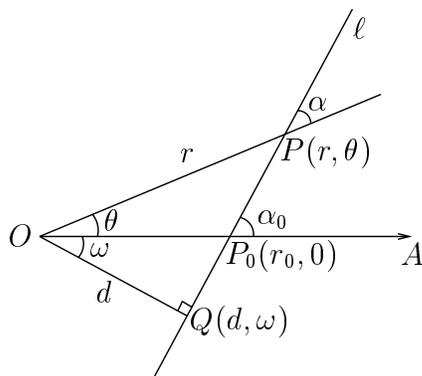
- (10) 如 [图 3-15] 所示, P_1, P_2 的极坐标分别是 (r_1, θ_1) 和 (r_2, θ_2) 。试求 $\overline{P_1P_2}^2$ 和 $\triangle OP_1P_2$ 的有向面积的极坐标公式。



[图 3-15]

- (11) 如 [图 3-16] 所示, α 是直线 l 和矢径 \overrightarrow{OP} 之间的夹角, α_0 是 l 和基准方向 \overrightarrow{OA} 之间的夹角, ω 是 l 的法线 (normal line) 的方向角, d 是原点和 l 之间的距离。试证

$$\begin{cases} r \cdot \sin \alpha = r_0 \sin \alpha_0 = d, \\ \sin \alpha = \cos(\theta - \omega) \end{cases}$$



[图 3-16]

- (12) 计算 [图 3-12'] 中三角形 $\triangle ABC$ 的面积平方 $(\triangle ABC)^2$, 并试将所算的结果给以几何解释。

第四章

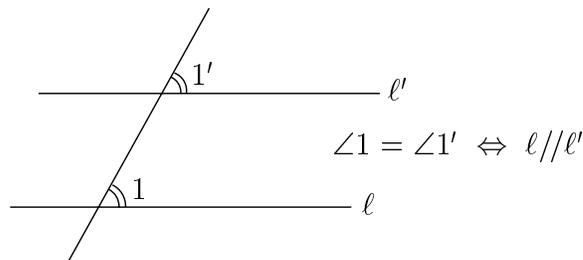
空间中的平行与垂直

——平直性与对称性的交互作用

本章将进而研讨定量空间（亦即立体）几何的基础理论。在空间的种种性质之中，最为基本而且影响无比深远者，首推对称性 (Symmetry) 和平直性 (Flatness 或 Rectilinearity)。两者在三角形上的表述分别就是「S.A.S. 叠合条件」和「三内角和恒为一个平角」。我们将会分别对于平直性和对称性在立体几何中的表现作一番返璞归真的分析，而其所展现者，乃是空间中的「平行」与「垂直」以及两者之间的密切关联。其实平行与垂直乃是整个定量立体几何的基础所在，当然也就是同学们学习立体几何的起点与要点所在。

4.1 平直性与平行

在平面几何中，同学们熟用如 [图 4-1] 所示的同位角相等来检验或构造平行线：



[图 4-1]

在引言和第二章中已经和同学们讨论过，上述验证条件的可行性其实和「三角形内角和恒等于一个平角」是同一件事的两种表现。而欧几里德的著名「第五公设」也就是以公设形式宣称上述条件就是 l_1, l_2 不相交的唯一可能性，亦即以公设形式假设了三角形内角和恒等于一个平角。

开宗明义，在空间中点、直线、平面之间的连结与交截，具有下述三点基本性质（通常也称之为公理）：

其一：相异两点定一直线，不共线三点定一平面。

其二：平面 Π 上相异两点所定的直线会完全包含于 Π 之内。

其三：设 $\Pi_1 \cap \Pi_2 \neq \phi$ （亦即非空），则 Π_1, Π_2 相交于一条直线或相重（亦即两面之交集不可能仅仅是一个点）。

空间中平行性的定义：

(i) 面、面平行之定义

$$\Pi_1 // \Pi_2 \Leftrightarrow \Pi_1 \cap \Pi_2 = \begin{cases} \phi \text{ 或} \\ \Pi_1 = \Pi_2 \end{cases}$$

(ii) 线、面平行之定义

$$l // \Pi \Leftrightarrow l \cap \Pi = \begin{cases} \phi \text{ 或} \\ l \end{cases}$$

(iii) 线、线平行之定义

$$l_1 // l_2 \Leftrightarrow l_1, l_2 \text{ 共面而且 } l_1 \cap l_2 = \begin{cases} \phi \text{ 或} \\ l_1 = l_2 \end{cases}$$

注意：线、线平行关系要求两者必须共面。两条不共面的直线的关系则称之为不交线 (skew lines)。

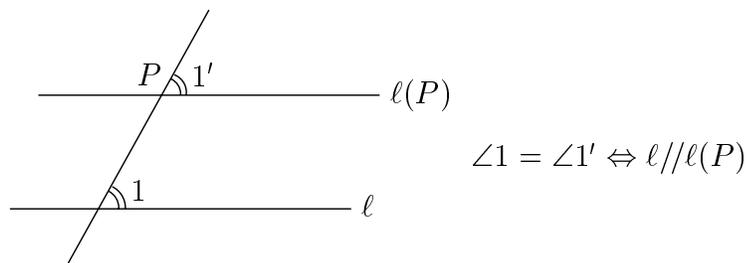
【引理 4.1】：设 $\Pi_1 // \Pi_2, \Pi \cap \Pi_1 = l_1, \Pi \cap \Pi_2 = l_2$ ，则 $l_1 // l_2$ 。

证明：若 $\Pi_1 = \Pi_2$ 则显然有 $l_1 = l_2$ 。若 $\Pi_1 \cap \Pi_2 = \phi$ ，则 l_1, l_2 共在 Π 之中而且

$$\begin{aligned} l_1 \cap l_2 &= (\Pi \cap \Pi_1) \cap (\Pi \cap \Pi_2) \\ &= \Pi \cap \Pi_1 \cap \Pi_2 \\ &= \Pi \cap \phi = \phi \end{aligned}$$

□

在平面几何中关于平行线的一个基本结果是过平面上任何给定点 P ，存在唯一一条直线 $l(P)$ ，它和给定直线 l 互相平行。再者，如 [图 4-1'] 所示，同位角相等乃是 $l // l(P)$ 的一个特征性质。

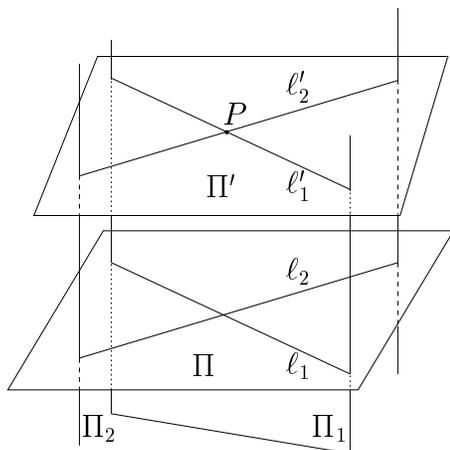


[图 4-1']

同样的，在立体几何中我们也有下述定理：

【定理 4.1】：对于一个给定平面 Π 和一个给定点 P ，存在唯一的一个平面 $\Pi(P)$ ，它过 P 点而且 $\Pi(P) // \Pi$ 。

证明：我们将用解作图题的手法和精神去证明上述 $\Pi(P)$ 的唯一存在性。



[图 4-2]

当 $P \in \Pi$ 时，显然有 $\Pi(P) = \Pi$ ，所以只需讨论 $P \notin \Pi$ 的情形。我们所要论证者乃是一个包含 P 而且和 Π 不相交的平面是唯一存在的。先证其唯一性：设 Π' 是一个这样的平面，如 [图 4-2] 所示，在 Π 上

任取两条相交的直线 l_1, l_2 。令

$$\begin{aligned}\Pi_1 &= \langle \{P\} \cup l_1 \rangle, \quad \Pi_2 = \langle \{P\} \cup l_2 \rangle; \\ \ell'_1 &= \Pi_1 \cap \Pi', \quad \ell'_2 = \Pi_2 \cap \Pi'\end{aligned}$$

由[引理 4.1]得知

$$\ell'_1 // l_1 \text{ (在 } \Pi_1 \text{ 之中)}, \quad \ell'_2 // l_2 \text{ (在 } \Pi_2 \text{ 之中)}$$

再者, 由平面几何中平行线的唯一性可知 ℓ'_1, ℓ'_2 分别在 Π_1, Π_2 之中是唯一存在的。因此唯一地确定了 $\Pi' = \langle \ell'_1 \cup \ell'_2 \rangle$ 。

其实上述论证业已明确地指出这样一个 Π' 的构造作图法。现在直截了当地用构造来证明其存在性:

第一步: 在 Π 上任取相交的两条直线 l_1, l_2 。令 $\Pi_1 = \langle \{P\} \cup l_1 \rangle$, $\Pi_2 = \langle \{P\} \cup l_2 \rangle$ 。

第二步: 用平面几何所熟知的作图法分别在 Π_1, Π_2 中作过 P 点而且分别平行于 l_1, l_2 的直线 ℓ'_1, ℓ'_2 。令 $\Pi(P) = \langle \ell'_1 \cup \ell'_2 \rangle$, 则 $\Pi(P) \cap \Pi = \phi$ 。不然, 则 $\Pi(P)$ 和 Π 的交线 ℓ 至少和 l_1, l_2 中之一交于一点 Q 。则有

$$\begin{aligned}l_i \cap \ell &= (\Pi_i \cap \Pi) \cap (\Pi(P) \cap \Pi) \\ &= (\Pi_i \cap \Pi) \cap (\Pi_i \cap \Pi(P)) \\ &= l_i \cap \ell'_i\end{aligned}$$

其中至少有一是非空的, 和所作 $l_1 // \ell'_1, l_2 // \ell'_2$ 相矛盾。 \square

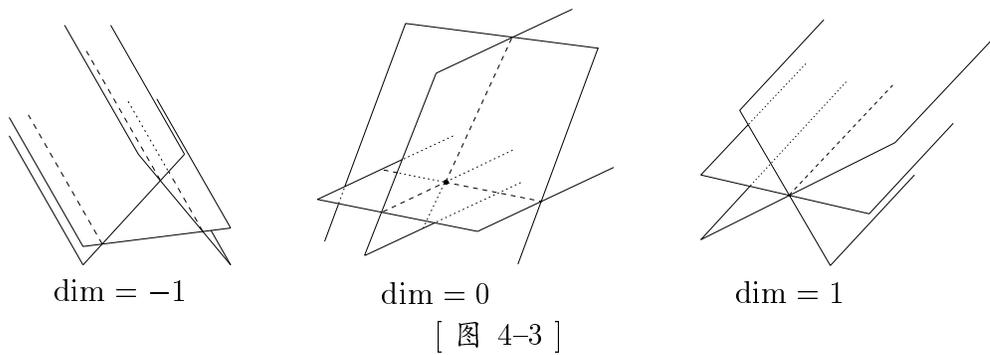
【推论】: 设 $\Pi // \Pi'$ 和 $\Pi' // \Pi''$, 则 $\Pi // \Pi''$ 。

证明: 设 $\Pi \cap \Pi'' \neq \phi$ 。令 P 为 $\Pi \cap \Pi''$ 中一点, 则 Π 和 Π'' 都是过 P 点而且和 Π' 平行者。所以由[定理 4.1]的唯一性得知 $\Pi = \Pi''$, 亦即

$$\Pi \cap \Pi'' = \begin{cases} \phi \\ \Pi = \Pi'' \end{cases} \Leftrightarrow \Pi // \Pi''$$

【引理 4.2】: 设 $\Pi_i, 1 \leq i \leq 3$, 是三个两两相交于一直线的平面。令 $\Pi_i \cap \Pi_j = l_k, (i, j, k) \sim (1, 2, 3)$, 则其交截有下列三种可能性, 即

$$\dim \Pi_1 \cap \Pi_2 \cap \Pi_3 = \begin{cases} -1 \\ 0 \\ 1 \end{cases} \Leftrightarrow \begin{cases} l_1 // l_2 // l_3 \\ l_1 \cap l_2 \cap l_3 = \{P\} \\ l_1 = l_2 = l_3 \end{cases}$$



证明：

$$l_1 \cap l_2 = (\Pi_2 \cap \Pi_3) \cap (\Pi_1 \cap \Pi_3) = \Pi_1 \cap \Pi_2 \cap \Pi_3$$

同理亦有

$$l_2 \cap l_3 = l_1 \cap l_3 = \Pi_1 \cap \Pi_2 \cap \Pi_3$$

由此可见，在 $\Pi_1 \cap \Pi_2 \cap \Pi_3$ 是 1-维时三线相重，而当其为 0-维时则三线共交于一点，即 $\{P\} = \Pi_1 \cap \Pi_2 \cap \Pi_3$ 。再者，当 $\Pi_1 \cap \Pi_2 \cap \Pi_3 = \phi$ 时，则有 $\{l_1, l_2, l_3\}$ 两两共面而且不相交，亦即

$$l_1 // l_2, \quad l_2 // l_3, \quad l_3 // l_1 \quad \square$$

【定理 4.2】：设 $l_1 // l_2, l_2 // l_3$ ，则 $l_1 // l_3$ 。

证明：若三线中有两线相重，则上述命题显然成立。所以只需讨论三线相异的情形如下：令

$$\langle l_1 \cup l_2 \rangle = \Pi_3, \quad \langle l_2 \cup l_3 \rangle = \Pi_1$$

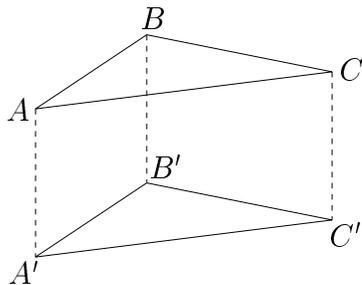
在 l_3 上任取一点 P ，令

$$\langle \{P\} \cup l_1 \rangle = \Pi_2, \quad l'_3 = \Pi_1 \cap \Pi_2$$

由[引理 4.2]即有 $\{l_1, l_2, l'_3\}$ 两两互相平行。所以 l_3 和 l'_3 都是 Π_1 中过 P 点而且和 l_2 平行者。由平面几何中平行线的唯一性得知 $l_3 = l'_3$ 。这就证明了 $l_1 // l_3$ 。 \square

【推论】：设 $\triangle ABC$ 和 $\triangle A'B'C'$ 是空间中两个三角形。若 $\{\overline{AB}, \overline{A'B'}\}$ 和 $\{\overline{BC}, \overline{B'C'}\}$ 分别同向平行而且等长，则 $\{\overline{AC}, \overline{A'C'}\}$ 也同向平行而且等长。

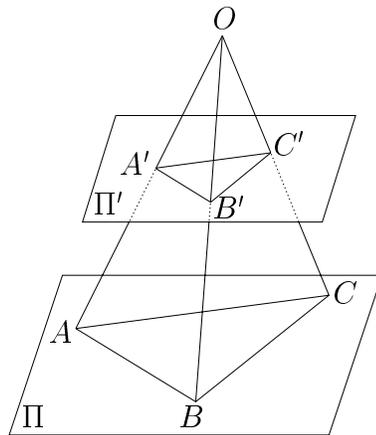
证明：连结 $\overline{AA'}$, $\overline{BB'}$ 和 $\overline{CC'}$ 。由所设和平面几何中平行四边形的特征性质定理，可见 $\square ABB'A'$ 和 $\square BCC'B'$ 都是平行四边形。因此 $\overline{AA'} // \overline{BB'}$, $\overline{BB'} // \overline{CC'}$ 而且等长。由[定理 4.2], $\overline{AA'} // \overline{CC'}$ 而且等长。所以 $\square ACC'A'$ 也是平行四边形。这就证明了 \overline{AC} 和 $\overline{A'C'}$ 也是同向平行而且等长 (参看 [图 4-4])。□



[图 4-4]

【习题】：

- (1) 若 $\Pi // \Pi'$, $l \subset \Pi$, 则 $l // \Pi'$ 。试证之。
- (2) 设 $l_1 \cap l_2 = \{P\}$, $\langle l_1 \cup l_2 \rangle = \Pi$ 。若 $l_1 // \Pi'$, $l_2 // \Pi'$, 则 $\Pi // \Pi'$ 。试证之。
- (3) 设 $l \cap \Pi = \phi$, $\Pi_1 \cap \Pi_2 = l$, $\Pi_1 \cap \Pi = l_1$, $\Pi_2 \cap \Pi = l_2$ 。试证 $l_1 // l_2$ 。
- (4) 设 $\Pi_1 \cap \Pi_2 = l'$ 而且 $l // \Pi_1$, $l // \Pi_2$ 。试证 $l // l'$ 。
- (5) 设 $\Pi \cap \Pi' = \phi$, $\triangle ABC$ 和 $\triangle A'B'C'$ 分别位于 Π 和 Π' 之中。若 AA' , BB' , CC' 三线共交于一点 O (如 [图 4-5] 所示) 则 $\triangle ABC \sim \triangle A'B'C'$ (为相似三角形)。试证之。
- (6) 如 [图 4-5] 所示, $\Pi = \langle \{A, B, C\} \rangle$, $\Pi' = \langle \{A', B', C'\} \rangle$ 而且 AA' , BB' 和 CC' 三线共交于 O 点。若 $\triangle OAB \sim \triangle OA'B'$ 和 $\triangle OBC \sim \triangle OB'C'$, 则 $\Pi // \Pi'$ 。试证之。
- (7) 如 [图 4-5] 所示和上题所设, 若 AB 和 $A'B'$, AC 和 $A'C'$, BC 和 $B'C'$ 各自交于一点, 则上述三个交点共线。试证之。



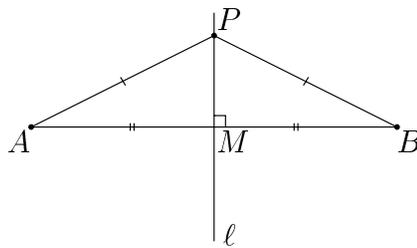
[图 4-5]

4.2 对称性与垂直

4.2.1 垂直平分线与平面上的反射对称

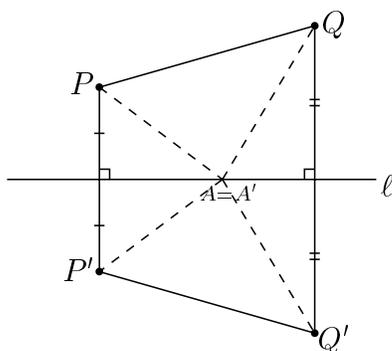
让我们先来温习一下平面的情形：

- 平面中和给定两点 A, B 等距的点集（亦称之为轨迹）是 \overline{AB} 的垂直平分线



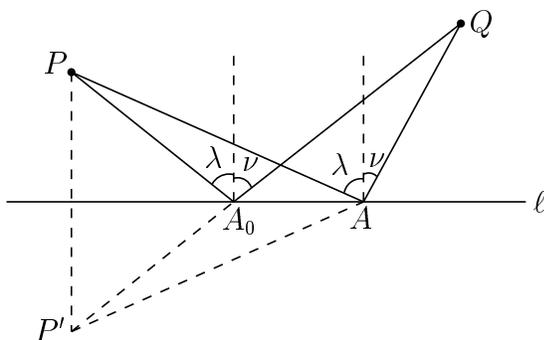
[图 4-6]

- 平面对于其中一条给定直线 l 成反射对称； P, P' 对于 l 互为对称点的充要条件是 $\overline{PP'}$ 被 l 所垂直平分或 $P = P' \in l$ （如 [图 4-7] 所示，试证 \overline{PQ} 和 $\overline{P'Q'}$ 恒为等长。）



[图 4-7]

【习题】：



[图 4-8]

如 [图 4-8] 所示, 设 A 是 l 上的任给一点, λ 和 ν 是 \overline{PA} 和 \overline{AQ} 与 l 的垂线之间的夹角。试证 $\overline{PA} + \overline{AQ}$ 在 $\lambda = \nu$ 时为极小 (亦即 $\overline{QA_0}$ 的延长线过 P 的对称点 P' 者)。

上述平面上的反射对称和垂直平分线之间的简洁关系在空间的推广乃是立体几何极为基本的要点。首先, 我们要研讨空间之中和给定两点 A, B 等距的点集, 即

【定理 4.3】：设 A, B 是给定相异两点, S 是空间中所有和 A, B 等距的点所构成者。则 S 是一个过 \overline{AB} 中点 M 的平面, 而且 S 上任何过 M 点的直线都和 \overline{AB} 垂直 (亦称正交)。

[一个简约的描述是： S 乃是直线段 \overline{AB} 的垂直平分面。]

证明：论证的要点在于证明 S 的平直性。设 P_1, P_2 是 S 之中的相异两点, 而 P 是直线 P_1P_2 上任给一点。我们所要证者是 P 点也必然

属于 S ，亦即由所设

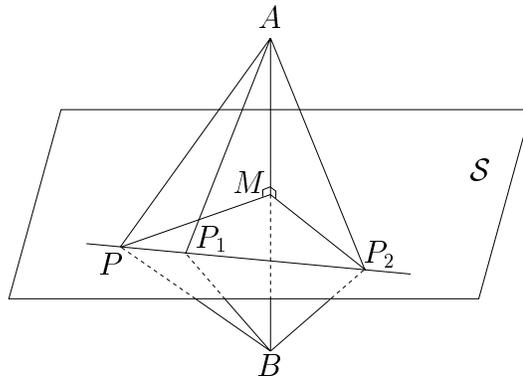
$$\overline{AP_1} = \overline{BP_1}, \quad \overline{AP_2} = \overline{BP_2}, \quad P \in P_1P_2$$

推论 $\overline{AP} = \overline{BP}$ 。兹证之如下：

由所设 $\triangle AP_1P_2$ 和 $\triangle BP_1P_2$ 满足 S.S.S. 全等条件，所以

$$\angle PP_1A = \angle PP_1B, \quad \angle PP_2A = \angle PP_2B$$

由此可见， $\triangle PP_1A$ 和 $\triangle PP_1B$ 满足 S.A.S. 全等条件，即有 $\triangle PP_1A \cong \triangle PP_1B$ ，所以 $\overline{AP} = \overline{BP}$ 。（同理也有 $\triangle PP_2A \cong \triangle PP_2B$ ，但只用其一已经证得 $\overline{AP} = \overline{BP}$ 。）所以 S 是一个平直子集，它显然包含 \overline{AB} 的中点 M ，但不是包含 A, B 。而且对于任给一个 S 中相异于 M 的点 P ， $S \cap \langle \{A, B, P\} \rangle$ 等于平面 $\langle \{A, B, P\} \rangle$ 中 \overline{AB} 的垂直平分线，所以 S 必定是一个平面。它其实是由所有这种垂直平分线所组合成者（参看 [图 4-9]）。



[图 4-9]

【推论】：设 $l \cap \Pi = \{M\}$ ，而且 l 和 Π 上两条相异的 MP_1, MP_2 正交，则 l 和 Π 上任给过 M 点的直线皆正交。

证明：在直线 l 上任取一直线段 \overline{AB} 以 M 为其中点，令 S 为 \overline{AB} 的垂直平分面。由所设 M, P_1, P_2 是 S 和 Π 上不共线三点。所以 $S = \Pi$ 。

【线、面垂直之定义】：若 $l \cap \Pi = \{M\}$ 而且 l 和 Π 上所有过 M 点的直线皆为正交，则称 l 垂直于 Π ，以 $l \perp \Pi$ 记之。

[上述推论则说，只要检验其中两条在 Π 上过交点 M 和 l 的正交性，即可得知 $l \perp \Pi$ 。]

4.2.2 立体几何中的作图题

在讨论立体几何的作图题之前，让我们先分析一下它和平面几何的作图题在实践上的基本差别：

- (i) 一张纸、一块黑板都是一个平面的局部，因此常用的平面几何作图是真的可以在一张纸或黑板上逐步用直尺、圆规去画出来的。但是立体几何的作图所要用到的是 3-维的「纸」或「黑板」才能实地执行之。而这种 3-维的「纸」和「黑板」是不实用而且难以供应的。
- (ii) 在平面几何作图中，可以用直尺去画出纸上两点所决定的那条直线的局部，但是在立体几何作图中并没有一个简单的工具能够利落简洁地画出不共线三点所决定的那个平面的局部。
- (iii) 由此可见，立体几何作图在本质上乃是「理念作图」。我们真正所要做的是把某种唯一存在、简单而且基本的立体几何事物，在理念中逐步分解成某些特定的平面上的平面几何作图的组合而加以明确的刻划。这种理念作图在训练如何把立体几何中的基本图象归于平面几何作图来加以分析。唯有通过这种理念作图的练习，才能学会如何有效运用平面几何所学者去进而理解空间的本质。我们与生俱来的视觉是具有相当好的空间想象能力的，但是要把它提升到对于空间图象及其所蕴含的空间本质的洞察力，这种训练是不可缺的！

[例如在[定理 4.2]的存在性的证明中，所做的就是这种理念作图。]

【基本作图题 4.1】：过直线 l 上的给定点 M ，作其垂面。

作法：在 l 外取两点 P_1, P_2 使得

$$\Pi_1 = \langle l \cup \{P_1\} \rangle \text{ 和 } \Pi_2 = \langle l \cup \{P_2\} \rangle \text{ 相异。}$$

用平面几何基本作图分别作 Π_1 和 Π_2 中过 M 点而且和 l 垂直的直线 l_1, l_2 。则 $\Pi = \langle l_1 \cup l_2 \rangle$ 即为所求作的垂面。[证明留作习题]

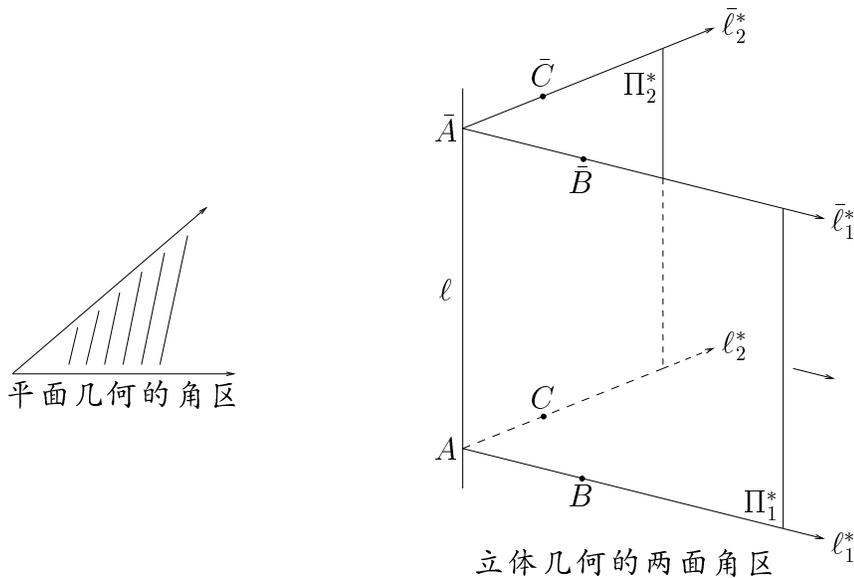
【基本作图题 4.2】：过平面 Π 上一点 P ，作其垂线。

作法：在 Π 上任取一条过 P 点的直线 l 。过 P 点用[基本作图 4.1]作 l 的垂面 Π_1 。令 $l_1 = \Pi_1 \cap \Pi$ ，再在 Π_1 中作过 P 点而且垂直于 l_1 的直线，即为所求作的直线。[证明留作习题]

【基本作图题 4.3】：设 P 是平面 Π 之外的给定点。求作过 P 点而且垂直于 Π 的直线。

作法：先在 Π 上任取一点 Q 。用[基本作图 4.2]作和 Π 正交于 Q 的直线 l_1 。若 l_1 恰好也过 P 点，则 l_1 即为所求作者。不然，令 $\Pi_1 = \langle \{P\} \cup l_1 \rangle$ ，再用平面几何基本作图在平面 Π_1 中作过 P 点而且和 l_1 平行的直线 $l(P)$ ，则 $l(P)$ 即为所求作的垂线。[证明留作习题]

在平面几何中，一个角区的边界由两条共顶点的半线（亦即射线）所组成。相类似地，空间的一个两面角区的边界是由两个共顶棱的半平面所组成（参看 [图 4-10]）。



[图 4-10]

【引理 4.3】：如 [图 4-10] 所示，设 A, \bar{A} 是 l 上任取两点。 $\Pi, \bar{\Pi}$ 分别是过 A, \bar{A} 而且和 l 正交的平面。令

$$\begin{aligned} l_1^* &= \Pi \cap \Pi_1^*, & l_2^* &= \Pi \cap \Pi_2^* \\ \bar{l}_1^* &= \bar{\Pi} \cap \bar{\Pi}_1^*, & \bar{l}_2^* &= \bar{\Pi} \cap \bar{\Pi}_2^* \end{aligned}$$

则有 $\angle(\ell_1^*, \ell_2^*) = \angle(\bar{\ell}_1^*, \bar{\ell}_2^*)$ 。

证明：如 [图 4-10] 所示，在射线 $\ell_1^*, \ell_2^*, \bar{\ell}_1^*, \bar{\ell}_2^*$ 上分别取 B, C, \bar{B}, \bar{C} 使得 $\overline{AB} = \overline{\bar{A}\bar{B}}, \overline{AC} = \overline{\bar{A}\bar{C}}$ 。由所作易见 \overline{AB} 和 $\overline{\bar{A}\bar{B}}$ ； \overline{AC} 和 $\overline{\bar{A}\bar{C}}$ 皆为同向平行而且等长。再由 [定理 4.2] 的推论即有 \overline{BC} 和 $\overline{\bar{B}\bar{C}}$ 也是同向平行而且等长的。所以 $\triangle ABC$ 和 $\triangle \bar{A}\bar{B}\bar{C}$ 满足 S.S.S. 全等条件， $\angle(\ell_1^*, \ell_2^*) = \angle A = \angle \bar{A} = \angle(\bar{\ell}_1^*, \bar{\ell}_2^*)$ 。□

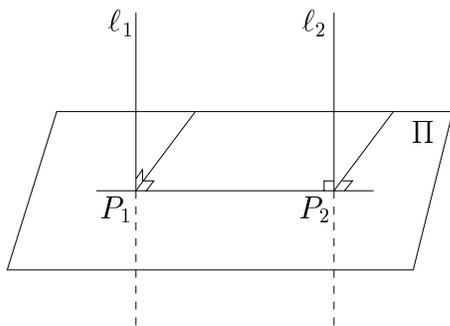
【两面角的定义】： $\angle(\Pi_1^*, \Pi_2^*)$ 的大小定义为 $\angle(\ell_1^*, \ell_2^*)$ ，因为后者是和 l 上 A 点的选取无关的！再者，在 $\angle(\Pi_1^*, \Pi_2^*) = \frac{\pi}{2}$ 时，则称两面正交，以符号 $\Pi_1 \perp \Pi_2$ 记之。

【基本作图题 4.4】：设 $l \subset \Pi$ ， \mathfrak{U}_Π^* 是以平面 Π 为边界的半空间， Π_ℓ^* 是 Π 中以 l 为边界的半平面。求作 \mathfrak{U}_Π^* 中那个和 Π_ℓ^* 共以 l 为顶棱的半平面 $\bar{\Pi}_\ell^*$ ，使得 $\angle(\Pi_\ell^*, \bar{\Pi}_\ell^*)$ 等于一个给定角。

作法：在 l 上任取一点 A ，过 A 点作 l 的垂面 Π_1 ，令 $\ell_1^* = \Pi_1 \cap \Pi_\ell^*$ 。再用平面几何基本作图在 Π_1 中的半面 $\Pi_1 \cap \mathfrak{U}_\Pi^*$ 中作那条半线 ℓ_2^* 使得 $\angle(\ell_1^*, \ell_2^*)$ 等于给定角。则 $\langle l \cup \ell_2^* \rangle \cap \mathfrak{U}_\Pi^*$ 就是所求作者。[由上述定义，可见证明是显然的。]

【引理 4.4】：设 $l_1 \perp \Pi, l_2 \perp \Pi$ ，则 $l_1 // l_2$ 。

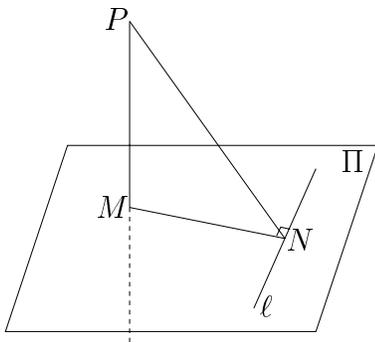
证明：令 $P_1 = l_1 \cap \Pi, P_2 = l_2 \cap \Pi$ 。若 $P_1 = P_2$ ，则 $l_1 = l_2$ 。若 $P_1 \neq P_2$ ，令 $\Pi' = \langle l_1 \cup \{P_2\} \rangle$ ， ℓ_2' 是 Π' 中过 P_2 点而且和 l_1 平行的直线。由 $l_1 \perp \Pi$ 可见 $\angle(\Pi, \Pi') = \frac{\pi}{2}$ 。再由 [引理 4.3]，可证 $\ell_2' \perp \Pi$ 。所以 $\ell_2' = l_2$ ，亦即 $l_1 // l_2$ （参看 [图 4-11]）。



[图 4-11]

【习题】：

- (1) 证明[基本作图 4.1]的作法所得者乃是唯一合乎所作条件的条件者。
- (2) 证明[基本作图 4.2]的作法所得者乃是唯一合乎所作条件的条件者。
- (3) 证明[基本作图 4.3]的作法所得者乃是唯一合乎所作条件的条件者。
- (4) 设 $\Pi \perp l, \Pi' \perp l$, 则 $\Pi // \Pi'$ 。试证之。
- (5) 设 $\Pi \perp l$ 而且 $\Pi // \Pi'$, 则 $l \perp \Pi'$ 。试证之。
- (6) 设 $l_1 \perp \Pi, l_1 // l_2$, 则 $l_2 \perp \Pi$ 。试证之。
- (7) 设 $l_1 \cap l_2 = \{P\}$ 而且 $l_1 // \Pi, l_2 // \Pi$, 则 $\langle l_1 \cup l_2 \rangle // \Pi$ 。试证之。
- (8) 设 $l \perp \Pi, \Pi' \supset l$, 则 $\Pi' \perp \Pi$ 。试证之。
- (9) 设 $\Pi_1 \cap \Pi_2 = l, \Pi_1 \perp \Pi, \Pi_2 \perp \Pi$, 则 $l \perp \Pi$ 。试证之。
- (10) 设 $\Pi_1 \cap \Pi_2 = \phi, l_1, l_2$ 是两条垂直于 Π_1, Π_2 的直线。令 $P_1 = l_1 \cap \Pi_1, P_2 = l_1 \cap \Pi_2, Q_1 = l_2 \cap \Pi_1, Q_2 = l_2 \cap \Pi_2$, 试证 $\overline{Q_1 Q_2}$ 和 $\overline{P_1 P_2}$ 等长。
[上述和公垂线选取无关的长度, 将定义为 Π_1, Π_2 这一对平行面之间的距离, 将以符号 $d(\Pi_1, \Pi_2)$ 记之。]
- (11) 如 [图 4-12] 所示, PM 和 Π 正交于 $M, l \subset \Pi, MN$ 和 l 正交于 N , 则 PN 和 l 正交。试证之。



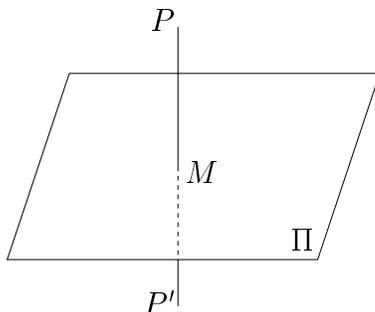
[图 4-12]

[提示：要点在于证明 $\langle \{P, M, N\} \rangle$ 和 l 是正交的。]

(12) 【基本作图题 4.5】过直线 l 之外一点 P 作其垂面 $\Pi(P)$ 。

4.2.3 空间反射对称性与垂直投影

对于空间中一个给定平面 Π ，每个点 P 有一个唯一的对称点 P' 。若 $P \in \Pi$ 则其对称点 P' 就是 P 本身。若 $P \notin \Pi$ ，则 $\overline{PP'}$ 以 Π 为其垂直平分面，如 [图 4-13] 所示。易见 P 也就是 P' 的对称点。



[图 4-13]

【定义】：上述空间中 $P \mapsto P'$ 这个变换 (transformation) 叫做空间对于给定平面 Π 的反射对称，将以符号 \mathfrak{R}_Π 记之，亦即

$$\text{当 } \begin{cases} P \in \Pi \\ P \notin \Pi \end{cases} \text{ 时, } \begin{cases} \mathfrak{R}_\Pi(P) = P, \\ \overline{P\mathfrak{R}_\Pi(P)} \text{ 被 } \Pi \text{ 所垂直平分。} \end{cases}$$

显然有 $\mathfrak{R}_\Pi(\mathfrak{R}_\Pi(P)) = P$ 恒成立，亦即 \mathfrak{R}_Π^2 乃是空间的恒等变换。

【定理 4.4】： \mathfrak{R}_Π 是空间的一个保长变换，亦即直线段 \overline{PQ} 和 $\overline{\mathfrak{R}_\Pi(P)\mathfrak{R}_\Pi(Q)}$ 恒为等长。

证明：在此我们采用简约符号 P' 表示 $\mathfrak{R}_\Pi(P)$ ，而且把证明分为下述四种情形：

- (i) 若 $P, Q \in \Pi$ 时， $P' = P, Q' = Q$ 。 $\overline{P'Q'}$ 就是 \overline{PQ} 本身。
- (ii) 若 P, Q 中其一个属于 Π ，设其为 P 。则 $P' = P$ ，而 $\overline{QQ'}$ 被 Π 所垂直平分。所以 $\triangle PQQ'$ 的中线 \overline{PM} 和底边 $\overline{QQ'}$ 正交，亦即 $\triangle PQQ'$ 是等腰三角形，即已证得 $\overline{PQ} = \overline{PQ'} = \overline{P'Q'}$ 。
- (iii) 若 $PQ \cap \Pi = \phi$ ，令 $\overline{PP'}$ 和 $\overline{QQ'}$ 的中点分别为 M, N 。则易证 $\square PMNQ$ 和 $\square P'MNQ'$ 都是矩形，所以即有 $\overline{PQ} = \overline{MN} = \overline{P'Q'}$ 。

(iv) 若 $PQ \cap \Pi = \{A\}$, 则 (ii) 之所证可见 $\triangle APP'$ 和 $\triangle AQQ'$ 都是等腰的, 亦即

$$\overline{AP} = \overline{AP'}, \quad \overline{AQ} = \overline{AQ'} \quad \Rightarrow \quad \overline{PQ} = \overline{P'Q'} \quad \square$$

[注]: 保长变换的组合当然还是保长变换。现在让我们来分析一下, 对于两个(相异的)平面 Π_1, Π_2 的反射对称的组合 $\mathfrak{R}_{\Pi_2} \circ \mathfrak{R}_{\Pi_1}(P) = \mathfrak{R}_{\Pi_2}(\mathfrak{R}_{\Pi_1}(P))$ 究竟是一种怎样的保长变换。为此, 我们将采用简约符号

$$P' = \mathfrak{R}_{\Pi_1}(P), \quad P'' = \mathfrak{R}_{\Pi_2}(P') = \mathfrak{R}_{\Pi_2} \circ \mathfrak{R}_{\Pi_1}(P)$$

下面将分成 $\Pi_1 \cap \Pi_2 = \phi$ 和 $\Pi_1 \cap \Pi_2 = \ell$ 这两种情形来加以分析:

情况一: $\Pi_1 \cap \Pi_2 = \phi$ (亦即 $\Pi_1 // \Pi_2$)

令 $\ell(P)$ 是过 P 点而且和 Π_1, Π_2 正交的直线, 易见 P', P'' 点都在 $\ell(P)$ 上。令 $Q_1 = \ell(P) \cap \Pi_1, Q_2 = \ell(P) \cap \Pi_2$, 并以 $\overrightarrow{Q_1Q_2}$ 为 $\ell(P)$ 上的正向, 则有下列有向长度关系式, 即

$$\overrightarrow{PQ_1} = \overrightarrow{Q_1P'}, \quad \overrightarrow{P'Q_2} = \overrightarrow{Q_2P''}, \quad \overrightarrow{Q_1P'} + \overrightarrow{P'Q_2} = \overrightarrow{Q_1Q_2}$$

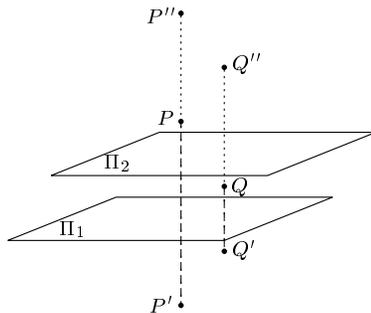
所以

$$\overrightarrow{PP''} = \overrightarrow{PP'} + \overrightarrow{P'P''} = 2\overrightarrow{Q_1P'} + 2\overrightarrow{P'Q_2} = 2\overrightarrow{Q_1Q_2}$$

我们把上述分析之所得叙述为[定理 4.5]:

【定理 4.5】: 设 $\Pi_1 \cap \Pi_2 = \phi$, 则 $\mathfrak{R}_{\Pi_2} \circ \mathfrak{R}_{\Pi_1}$ 把空间中每一点 P 沿著垂直于 Π_1, Π_2 的方向(亦即同向平行于 $\overrightarrow{Q_1Q_2}$ 者)向前移动 $2d(\Pi_1, \Pi_2)$ 。

[这种保长变换叫做平移(translation)。]



[图 4-14]

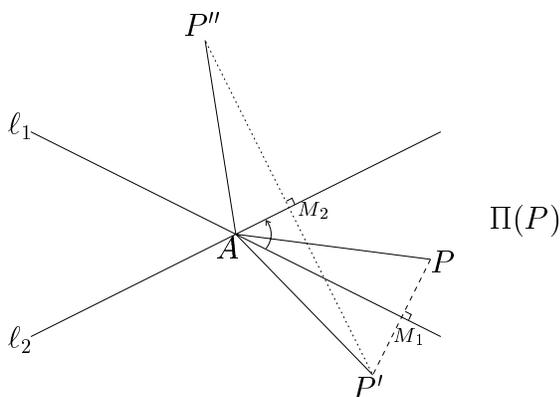
情况二： $\Pi_1 \cap \Pi_2 = \ell$ (相交于直线 ℓ)

令 $\Pi(P)$ 是过 P 点而且和 ℓ 垂直的平面 ([基本作图 4.5], 亦即习题 (12)), 则 $\Pi(P) \perp \Pi_1, \Pi(P) \perp \Pi_2$ 。由此易证 P' 和 P'' 点都和 P 点共在 $\Pi(P)$ 之中 (证明留作习题)。令 $\ell_1 = \Pi(P) \cap \Pi_1, \ell_2 = \Pi(P) \cap \Pi_2, A = \ell \cap \Pi(P)$ 。在 $\Pi(P)$ 上选取正向转角使得

$$0 < \angle \Pi_1, \Pi_2 = \angle \ell_1, \ell_2 \leq \frac{\pi}{2}$$

则有下列 $\Pi(P)$ 中定向角的关系式, 即如 [图 4-15] 所示

$$\angle PAM_1 = \angle M_1AP', \quad \angle P'AM_2 = \angle M_2AP''$$



[图 4-15]

所以

$$\begin{aligned} \angle PAP'' &= \angle PAP' + \angle P'AP'' \\ &= 2\angle M_1AP' + 2\angle P'AM_2 \\ &= 2\angle M_1AM_2 \\ &= 2\angle \Pi_1, \Pi_2 \end{aligned}$$

上述分析已经证得下述定理, 即

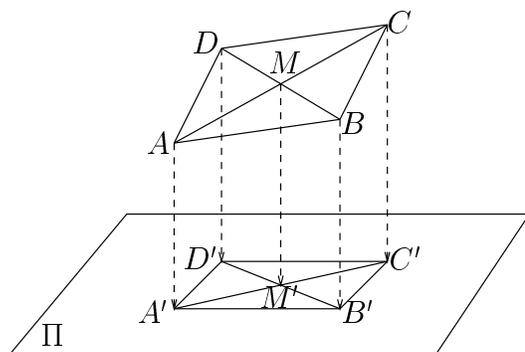
【定理 4.6】: 设 $\Pi_1 \cap \Pi_2 = \ell, \theta = \angle \Pi_1, \Pi_2$, 则 $\mathfrak{R}_{\Pi_2} \circ \mathfrak{R}_{\Pi_1}$ 保持任何一个 ℓ 的垂面 $\Pi(P)$ 为不变子集 (invariant subset), 而 ℓ 本身则是它的定点子集 (subset of fixed points)。再者, 把 $\mathfrak{R}_{\Pi_2} \circ \mathfrak{R}_{\Pi_1}$ 限制到任一不变平面 $\Pi(P)$ 上的作用乃是一个 2θ -角的旋转 (rotation)。

[这种保长变换叫做空间绕 ℓ -轴的旋转。]

垂直投影 (orthogonal projection) :

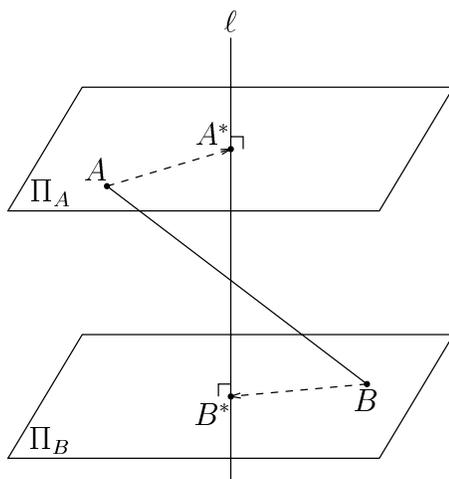
对于空间中一个给定平面 Π ，过每一点 P 可作唯一一条直线 $l(P)$ 和 Π 正交。对于空间一条给定直线 l ，过每一点 P 可作唯一一个平面 $\Pi(P)$ 和 l 正交。运用上述两个事实就可以定义空间的两种垂直投影：

【定义】：空间对于一个给定平面 Π 的垂直投影把 P 点映射到 $l(P) \cap \Pi$ 。由此可见，它把每条垂直于 Π 的直线上的所有点都映射到该直线和 Π 的交点，将以符号 \mathfrak{P}_{Π} 记之。



[图 4-16]

【定义】：空间对于一个给定直线 l 的垂直投影把 P 点映射到 $\Pi(P) \cap l$ 。由此可见，它把每个和 l 垂直的平面上的所有点，全都映射到它和 l 的交点，将以符号 \mathfrak{P}_l 记之。



[图 4-17]

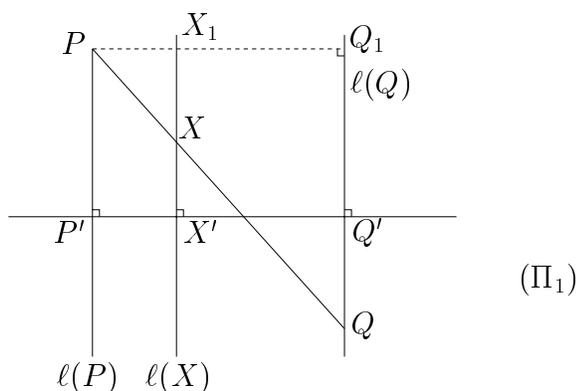
【定理 4.7】：以简约符号 P' 表示 $\mathfrak{P}_{\Pi}(P)$ 。设 $P' \neq Q'$ ，则有 $X \in PQ \Rightarrow X' \in P'Q'$ 而且

$$\frac{\overrightarrow{PX}}{\overrightarrow{XQ}} = \frac{\overrightarrow{P'X'}}{\overrightarrow{X'Q'}} \quad (\text{有向比相等})$$

证明：令 $\Pi_1 = \langle \ell(P) \cup \ell(Q) \rangle, (\ell(P) // \ell(Q))$ 。由所设

$$X \in PQ \text{ 和 } \ell(X) // \ell(P) // \ell(Q) \Rightarrow \ell(X) \subset \Pi_1$$

由此可见 $X' \in P'Q'$ (整个 Π_1 在 \mathfrak{P}_{Π} 下的象点点集)。



[图 4-18]

再者，如 [图 4-18] 所示和平面几何中的相似三角形定理，即得

$$\frac{\overrightarrow{PX}}{\overrightarrow{XQ}} = \frac{\overrightarrow{PX_1}}{\overrightarrow{X_1Q_1}} = \frac{\overrightarrow{P'X'}}{\overrightarrow{X'Q'}} \quad \square$$

【推论】：若 $\square ABCD$ 是一个平行四边形而且 A', B', C', D' 不共线，则 $\square A'B'C'D'$ 也是一个平行四边形 (见 [图 4-16])。

证明：由所设 \overline{AC} 和 \overline{BD} 互相平分。由 [定理 4.7]，即得 $\overline{A'C'}$ 和 $\overline{B'D'}$ 也互相平分，所以 $\square A'B'C'D'$ 也是一个平行四边形。 \square

【定理 4.8】：以简约符号 P^* 表示 $\mathfrak{P}_{\ell}(P)$ 。设 $P^* \neq Q^*$ ， $X \in PQ$ ，则同样地也有

$$\frac{\overrightarrow{PX}}{\overrightarrow{XQ}} = \frac{\overrightarrow{P^*X^*}}{\overrightarrow{X^*Q^*}}$$

证明：过 P 点作 l 的平行线 l' 。则 l' 和 $\Pi(P), \Pi(Q), \Pi(X)$ 也都是正交的。令

$$\Pi_1 = \langle PQ \cup l' \rangle, \quad X_1 = l' \cap \Pi(X), \quad Q_1 = l' \cap \Pi(Q)$$

则易见

$$\overrightarrow{P^*X^*} = \overrightarrow{PX_1}, \quad \overrightarrow{X^*Q^*} = \overrightarrow{X_1Q_1}$$

而同样的图形和相似三角形定理即得

$$\frac{\overrightarrow{PX}}{\overrightarrow{XQ}} = \frac{\overrightarrow{P^*X^*}}{\overrightarrow{X^*Q^*}} \quad \square$$

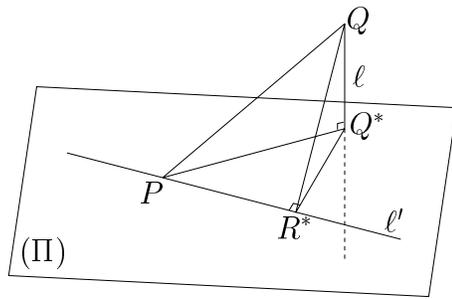
【定理 4.9】：设 $\{l_1, l_2, l_3\}$ 是共交于一点 O 而且两两正交的三条直线。若以简约符号 P_i 表示 $\mathfrak{P}_{l_i}(P)$ 则有

$$\overline{PQ}^2 = \overline{P_1Q_1}^2 + \overline{P_2Q_2}^2 + \overline{P_3Q_3}^2$$

[这也就是勾股定理在空间几何中的表达式]

证明：令 Π 是过 P 点和 l_3 正交的平面， l 是过 Q 点和 l_3 平行的直线， l' 是过 P 点而且和 l_1 平行的直线。如 [图 4-19] 所示，令 $Q^* = l \cap \Pi$ ，并作 Q^*R^* 和 l' 正交于 R^* 则有

$$\overline{QQ^*} \perp \overline{PQ^*} \quad (\text{亦有 } \overline{QQ^*} \perp \overline{R^*Q^*})$$



[图 4-19]

所以由勾股定理即有

$$\begin{aligned} \overline{PQ}^2 &= \overline{PQ^*}^2 + \overline{QQ^*}^2 \\ &= \overline{PR^*}^2 + \overline{R^*Q^*}^2 + \overline{QQ^*}^2 \end{aligned}$$

再者，由所作易见

$$\overline{PR^*} = \overline{P_1Q_1}, \quad \overline{R^*Q^*} = \overline{P_2Q_2}, \quad \overline{QQ^*} = \overline{P_3Q_3}$$

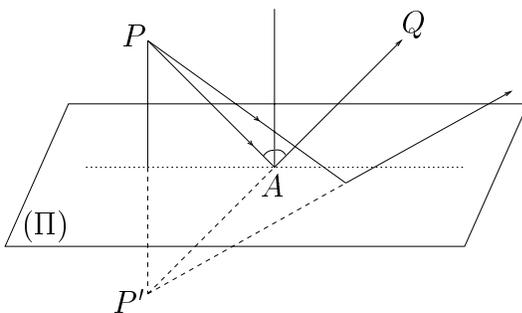
所以即已证得

$$\overline{PQ}^2 = \overline{P_1Q_1}^2 + \overline{P_2Q_2}^2 + \overline{P_3Q_3}^2 \quad \square$$

【习题】：

光的反射定律：一个平滑的镜面可以想成是一个平面的局部。远在古代希腊时代，即已认识到下述光的反射定律：

「入射线、反射线和平面在反射点的法线三线共面，而且两者和法线的夹角相等。」



[图 4-20]

- (1) 令 P' 是 P 对于镜面 Π 的反射对称点。试证起始于 P 点的反射线的延长线共交于 P' 点。[这也就是视觉中镜象的几何原由。]
- (2) 如 [图 4-20] 所示，起始于 P 点的光线，先走直线到达镜面 Π 上一点 A 然后再反射到 Q 点。试证这种走法是由 P 点经过 Π 上一点然后再到 Q 点的唯一最短通路。（亦即其他经过 Π 上一点由 P 到 Q 的通路都比 $\overline{PA} + \overline{AQ}$ 要长！）
- (3) 设 Π_1, Π_2, Π_3 是共交于 O 点而且两两垂直的三个平面。令 $P^\# = \mathfrak{R}_{\Pi_3} \circ \mathfrak{R}_{\Pi_2} \circ \mathfrak{R}_{\Pi_1}(P)$ 。试问 $P, O, P^\#$ 三点的几何关系是什麼？

设 T 是空间的一个保长变换（例如是好些个反射对称的组合），并且用简约符号 \hat{P} 记号 $T(P)$ 。

- (4) 试证 $X \in \overline{PQ} \Rightarrow \hat{X} \in \overline{\hat{P}\hat{Q}}$ 。
- (5) 设 $\triangle PQR$ 为直角三角形，试证 $\triangle \hat{P}\hat{Q}\hat{R}$ 也是一个直角三角形。
- (6) 试证 $\triangle PQR$ 和 $\triangle \hat{P}\hat{Q}\hat{R}$ 必定是全等的。
- (7) 设 A, B 是空间任给两点，是否存在一个适当的平移 T 使得 $T(A) = B$ ？试用反射对称之组合构造之。
- (8) 上述平移是否是唯一存在的？试证明你的观点。

基础数学讲义之三

《基础几何学之二》

向量几何，解析几何与球面几何

项武义

香港科技大学数学系

目录

五 向量几何和向量代数	85
5.1 位移向量的基本性质	85
5.2 位移向量的运算律	87
5.2.1 相似三角形定理和位移向量的倍积	88
5.2.2 勾股定理和位移向量的内积	90
5.2.3 面积的勾股定理和位移向量的 \times -积	95
5.3 结语	103
5.4 例题、习题和思考题	105
六 坐标解析几何简介	111
6.1 正交坐标系和平面（或空间）的坐标化	112
6.2 直线和圆，平面和球	116
6.3 圆的反射对称；共轴圆系和共轭等轴圆系	123
七 球面几何和球面三角学	131
7.1 单位球面的基本性质	131
7.2 球面三角学	139
八 圆锥截线的故事	159
8.1 圆柱截线和圆锥截线	159
8.2 圆锥截线的光学性质	162
8.3 圆锥截线和二次曲线	164
8.4 坐标变换和不变量	165
8.5 五点定一「二次曲线」和六点共在一「二次曲线」的条件	171
8.6 Pascal 定理和 Pappus 定理	172

8.7 Kepler 行星三定律	179
8.8 由 Kepler 定律到牛顿万有引力定律	181
8.9 圆锥截线例题，极与极线	189

第五章

向量几何和向量代数 ——空间结构的系统代数化

在空间中由 A 到 B 的有向直线段 \overrightarrow{AB} 其本身就是 A, B 两点所标记的两个位置之间的差别 (difference between the locations of A and B) 的具体化描述；而位移向量 (displacement vector) 则是将这种「位置差别」加以定量化所定义的基本几何量。它的本质内含是 \overrightarrow{AB} 的方向和长度。换句话说，当两个有向直线段 \overrightarrow{AB} 和 $\overrightarrow{A'B'}$ 为同向平行而且等长时，两者所表达的位移向量定义为相等。另一方面，空间中的平移变换把空间中每一个点作同向等距的移动，由此可见位移向量本质上就是平移的另一种表达。总之，位移向量和平移是同一事物的两种观点，前者着眼于其所表达的位置移动，而后者则着眼于它是空间一种特殊的保长变换。

5.1 位移向量的基本性质

【定义】：若空间中的一个变换 τ 满足条件

$$\overrightarrow{P\tau(P)} \text{ 和 } \overrightarrow{Q\tau(Q)} \text{ 恒为同向平行而且等长}$$

则称 τ 为空间的一个平移 (translation)。

【定理 5.1】：设 A, B 为空间中任给两点，则存在一个唯一的平移 τ 使得 $\tau(A) = B$ (我们将以 $\tau(A, B)$ 记之)。

证明：先证存在性：

若 $A = B$ ，则恒等变换即为所求者。

若 $A \neq B$ ，令 Π_1, Π_2 分别是和直线段 \overline{AB} 正交于 A 和 M (M 是 \overline{AB} 的中点) 的平面，则 $\mathfrak{R}_{\Pi_2} \circ \mathfrak{R}_{\Pi_1}$ 这个反射对称的组合即为所求者。

再证唯一性：

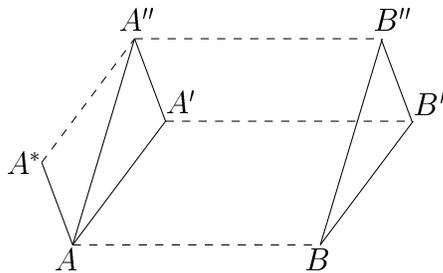
设 P 是空间中任给一点， τ 是上述所作的平移，而 τ' 则是一个将 A 映射到 B 的平移，我们所要证明者是 $\tau'(P) = \tau(P)$ 。由所设 $\overrightarrow{P\tau'(P)}$, $\overrightarrow{P\tau(P)}$ 都和 \overline{AB} 同向平行而且等长。因为过 P 点和 \overline{AB} 平行的直线是唯一存在的，所以 $\tau'(P) = \tau(P)$ 。由于 P 是任给的，所以 τ' 和 τ 是同一个变换。□

【定理 5.2】：设 τ_1, τ_2 是空间的两个平移，则 $\tau_2 \circ \tau_1$ 也是一个平移，而且 $\tau_2 \circ \tau_1 = \tau_1 \circ \tau_2$ 。

证明：在此我们将采用简约符号 $A' = \tau_1(A)$, $A'' = \tau_2(A') = \tau_2 \circ \tau_1(A)$ 。由所设，对于空间任给两点 A, B 恒有

$$\overrightarrow{AA'} \text{ 和 } \overrightarrow{BB'}, \quad \overrightarrow{A'A''} \text{ 和 } \overrightarrow{B'B''}$$

各别同向平行而且等长，而我们所要证明者则是 $\overrightarrow{AA''}$ 和 $\overrightarrow{BB''}$ 也必然同向平行而且等长。其证明如下：



[图 5-1]

如 [图 5-1] 所示，连结 \overrightarrow{AB} , $\overrightarrow{A'B'}$ 和 $\overrightarrow{A''B''}$ 。由平面几何中熟知的平行四边形特征性质可见，

$$\square ABB'A' \text{ 和 } \square A'B'B''A''$$

两者都是平行四边形，所以

$$\overrightarrow{AB} \text{ 和 } \overrightarrow{A'B'}, \quad \overrightarrow{A'B'} \text{ 和 } \overrightarrow{A''B''}$$

都是同向平行而且等长，即 \overrightarrow{AB} 和 $\overrightarrow{A''B''}$ 也是同向平行而且等长的，因此 $\square ABB''A''$ 乃是一个平行四边形。所以就证明了 $\overrightarrow{AA''}$ 和 $\overrightarrow{BB''}$ 也是同向平行而且等长的，亦即 $\tau_2 \circ \tau_1$ 也是一个平移。

再证： $\tau_1 \circ \tau_2 = \tau_2 \circ \tau_1$

令 $A^* = \tau_2(A)$ ，则有 $\overrightarrow{AA^*}$ 和 $\overrightarrow{A'A''}$ 同向平行而且等长。同理， $\square AA'A''A^*$ 也是一个平行四边形，即有 $\overrightarrow{A^*A''}$ 和 $\overrightarrow{AA''}$ 同向平行而且等长，所以

$$\tau_1 \circ \tau_2(A) = \tau_1(A^*) = A'' = \tau_2 \circ \tau_1(A)$$

而 A 是空间任给一点，亦即 $\tau_1 \circ \tau_2$ 和 $\tau_2 \circ \tau_1$ 是相同的变换。□

【定义与符号】：定义上述和次序无关的平移组合 $\tau_1 \circ \tau_2 = \tau_2 \circ \tau_1$ 为 τ_1, τ_2 之和 (sum)，将改用符号 $\tau_1 + \tau_2$ 表示之。再者，往後我们将改用粗体小写字母如 $\mathbf{a}, \mathbf{b}, \mathbf{c}, \mathbf{x}, \mathbf{y}$ 等表示位移向量。采取这种符号， $\tau(B, C) \circ \tau(A, B) = \tau(A, B) \circ \tau(B, C) = \tau(A, C)$ 就可以写成 $\mathbf{a} + \mathbf{b} = \mathbf{c}$ ，或者 $\overrightarrow{AB} + \overrightarrow{BC} = \overrightarrow{AC}$ 。

[注]：在加法符号下，我们把恒等变换这个特殊的平移 τ_0 改用 $\mathbf{0}$ 来表示，因为对于任何平移 τ ，恒有 $\tau_0 \circ \tau = \tau \circ \tau_0 = \tau$ ，而这个事实在加法符号下即为

$$\mathbf{0} + \mathbf{a} = \mathbf{a} + \mathbf{0} = \mathbf{a}$$

再者 \overrightarrow{AB} (亦即 $\tau(A, B)$) 的逆变换就是 \overrightarrow{BA} (亦即 $\tau(B, A)$)，而且有

$$\tau(A, B) \circ \tau(B, A) = \tau_0 = \tau(B, A) \circ \tau(A, B)$$

$$(\overrightarrow{AB} + \overrightarrow{BA} = \overrightarrow{AA} = \overrightarrow{BB} = \overrightarrow{BA} + \overrightarrow{AB})$$

所以记 $-\overrightarrow{AB} = \overrightarrow{BA}$ (即 $-\tau(A, B) = \tau(B, A)$)，亦即

$$(-\mathbf{a}) + \mathbf{a} = \mathbf{0} = \mathbf{a} + (-\mathbf{a})$$

5.2 位移向量的运算律

在上一节所定义的位移向量的加法运算，显然具有下述熟悉的运算律：

交换律： $\mathbf{a} + \mathbf{b} = \mathbf{b} + \mathbf{a}$

结合律： $(\mathbf{a} + \mathbf{b}) + \mathbf{c} = \mathbf{a} + (\mathbf{b} + \mathbf{c})$

[注]：因为一般的变换组合都是满足结合律的，而位移向量的加法是定义为平移的组合，当然也会满足结合律。再者，由

$$(\overrightarrow{AB} + \overrightarrow{BC}) + \overrightarrow{CD} = \overrightarrow{AC} + \overrightarrow{CD} = \overrightarrow{AD}$$

$$\overrightarrow{AB} + (\overrightarrow{BC} + \overrightarrow{CD}) = \overrightarrow{AB} + \overrightarrow{BD} = \overrightarrow{AD}$$

亦可以直接验证位移向量的加法结合律。

零和可逆性：以 $\mathbf{0}$ 表示恒等变换这个特殊的平移， $(-\mathbf{a})$ 表示和 \mathbf{a} 互逆的平移，则有

$$\mathbf{0} + \mathbf{a} = \mathbf{a} + \mathbf{0} = \mathbf{a} \quad \text{和} \quad \mathbf{a} + (-\mathbf{a}) = (-\mathbf{a}) + \mathbf{a} = \mathbf{0} \quad \text{恒成立}$$

[注]：平移和[定理 5.2]的证明都和空间中的「平行性」(parallelism) 以及平行四边形定理密切相关的。而交换律 $\mathbf{a} + \mathbf{b} = \mathbf{b} + \mathbf{a}$ 更可以想成是平行四边形定理的向量表述形式。由此可见，往後我们每次运用向量加法交换律，其实也就是对于所研讨的几何问题用了一次平行四边形定理。

5.2.1 相似三角形定理和位移向量的倍积

一个数 a 的整数倍 $n \cdot a$ 其实就是 n 个 a 相加的总和。同样我们也自然地把 n 个 (位移向量) \mathbf{a} 相加的总和定义为倍积 $n \cdot \mathbf{a}$ ，亦即：

$$n \cdot \mathbf{a} = \underbrace{\mathbf{a} + \mathbf{a} + \dots + \mathbf{a}}_{n \text{ 个}}, \quad (n+1) \cdot \mathbf{a} = n \cdot \mathbf{a} + \mathbf{a}$$

再者 $(-n) \cdot \mathbf{a} = n \cdot (-\mathbf{a}) = -(n \cdot \mathbf{a})$ 。

由上述位移向量的整数倍的定义，容易直接验证下列运算律，即：

$$(i) \quad m \cdot \mathbf{a} + n \cdot \mathbf{a} = (m+n) \cdot \mathbf{a}$$

$$(ii) \quad m \cdot (n \cdot \mathbf{a}) = (mn) \cdot \mathbf{a}$$

$$(iii) \quad n \cdot (\mathbf{a} + \mathbf{b}) = n \cdot \mathbf{a} + n \cdot \mathbf{b}$$

对于任给整数 m, n 和任给位移向量 \mathbf{a}, \mathbf{b} 恒成立。

[习题：试用归纳法验证 (i), (ii) 和 (iii)。]

设 $\mathbf{a} = \overrightarrow{AB} \neq \mathbf{0}$ ，我们可以把 \overrightarrow{AB} 作 n 等分，令 $\{B_i, 1 \leq i \leq (n-1)\}$ 为其等分点，则有

$$\overrightarrow{AB_1} = \overrightarrow{B_1B_2} = \dots = \overrightarrow{B_iB_{i+1}} = \dots = \overrightarrow{B_{n-1}B}, \quad n \cdot \overrightarrow{AB_1} = \overrightarrow{AB}$$

由此可见，我们应该把 $\frac{1}{n} \cdot \mathbf{a}$ 定义为 $\overrightarrow{AB_1}$ ，因为它是那个满足 $n \cdot \mathbf{x} = \mathbf{a}$ 的唯一解。再者 $\frac{m}{n} \cdot \mathbf{a}$ 的定义应该就是

$$\frac{m}{n} \cdot \mathbf{a} = m \cdot \left(\frac{1}{n} \cdot \mathbf{a} \right) = \frac{1}{n} (m \cdot \mathbf{a})$$

这样，就可以把位移向量的倍积由整数倍扩张到有理数倍。而且上述扩张法是唯一能够使得下列运算律依然成立者，即

$$(i') \quad \frac{m}{n} \cdot \mathbf{a} + \frac{p}{q} \cdot \mathbf{a} = \left(\frac{m}{n} + \frac{p}{q} \right) \cdot \mathbf{a}$$

$$(ii') \quad \frac{m}{n} \cdot \left(\frac{p}{q} \cdot \mathbf{a} \right) = \left(\frac{m}{n} \cdot \frac{p}{q} \right) \cdot \mathbf{a}$$

$$(iii') \quad \frac{m}{n} \cdot (\mathbf{a} + \mathbf{b}) = \frac{m}{n} \cdot \mathbf{a} + \frac{m}{n} \cdot \mathbf{b}$$

最後一步，让我们来分析一下位移向量的实数倍应该如何定义。设 λ 是一个非比实数（亦称为无理数）， $\mathbf{a} = \overrightarrow{AB}$ 。令 B^* 是直线 AB 上那个唯一的点使得有向长度之比

$$\overrightarrow{AB^*} : \overrightarrow{AB} = \lambda$$

则 $\lambda \cdot \mathbf{a}$ 应该定义为 $\overrightarrow{AB^*}$ ，因为它是唯一能够使得下述比较原则成立者，即

$$\left[\text{设 } \lambda \text{ 介于两个有理数 } \frac{m}{n} \text{ 和 } \frac{p}{q} \text{ 之间而且 } \overrightarrow{AB'} = \frac{m}{n} \overrightarrow{AB}, \right. \\ \left. \overrightarrow{AB''} = \frac{p}{q} \overrightarrow{AB}, \text{ 则 } B^* \text{ 亦必介于 } B', B'' \text{ 之间。} \right]$$

而运用上述比较原则和 Eudoxus 夹逼原理即可验证上述所定义的实数倍的倍积也满足同样的运算律, 即

$$(i'') \quad \lambda \cdot \mathbf{a} + \mu \cdot \mathbf{a} = (\lambda + \mu) \cdot \mathbf{a}$$

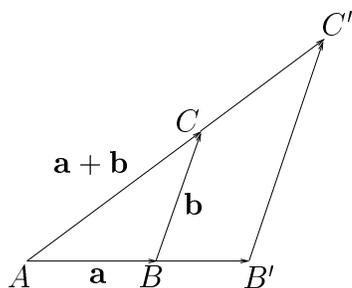
$$(ii'') \quad \lambda \cdot (\mu \cdot \mathbf{a}) = (\lambda\mu) \cdot \mathbf{a}$$

$$(iii'') \quad \lambda \cdot (\mathbf{a} + \mathbf{b}) = \lambda \cdot \mathbf{a} + \lambda \cdot \mathbf{b}$$

对于任给实数 λ, μ 和位移向量 \mathbf{a}, \mathbf{b} 恒成立。

[注]: 放大、缩小这种相似变换是空间中常见常用者, 而平面几何中的相似三角形定理则是关于相似变换的基本定理。在此, 值得注意的是倍积分配律 $k \cdot (\mathbf{a} + \mathbf{b}) = k \cdot \mathbf{a} + k \cdot \mathbf{b}$ 的本质就是上述基本定理的代数化形式 (参看 [图 5-2])。

令 $\mathbf{a} = \overrightarrow{AB}$, $\mathbf{b} = \overrightarrow{BC}$, 则 $\mathbf{a} + \mathbf{b} = \overrightarrow{AC}$ 。如 [图 5-2] 所示,



[图 5-2]

$\triangle ABC \sim \triangle A'B'C'$, $A = A'$ 而且 k 是其相似比, 则 $\overrightarrow{A'B'} = k \cdot \mathbf{a}$, $\overrightarrow{B'C'} = k \cdot \mathbf{b}$, $\overrightarrow{A'C'} = k \cdot \overrightarrow{AC} = k \cdot (\mathbf{a} + \mathbf{b})$ 。

5.2.2 勾股定理和位移向量的内积

一个位移向量 $\overrightarrow{AB} = \tau(A, B)$ 含有方向和长度这样两种本质, 我们将用符号 $|\mathbf{a}|$ 表示其长度, 以 $\angle \mathbf{a}, \mathbf{b}$ 表示两者的方向之差, 亦即两者之间的夹角。在平面几何学的研讨中, 三角形是既精且简的基本图形, 用向量来表达三角形, 则它的三个有向边就可以分别表达成 \mathbf{a}, \mathbf{b} 和 $\mathbf{a} + \mathbf{b}$ 。由平面几何中所熟知的 S.S.S. 叠合条件可见夹角 $\angle \mathbf{a}, \mathbf{b}$ 业已被其三边

边长 $|\mathbf{a}|$, $|\mathbf{b}|$, $|\mathbf{a} + \mathbf{b}|$ 所唯一确定。再者，中国古算中的勾股定理（即古希腊的毕氏定理）则可以改写成

$$\text{若 } \angle \mathbf{a}, \mathbf{b} = \text{直角}, \text{ 则有 } |\mathbf{a} + \mathbf{b}|^2 = |\mathbf{a}|^2 + |\mathbf{b}|^2$$

而在一般 \mathbf{a}, \mathbf{b} 并非互相垂直的情形则 $|\mathbf{a} + \mathbf{b}|^2 - |\mathbf{a}|^2 - |\mathbf{b}|^2 \neq 0$ 。例如当 $\mathbf{a} = \mathbf{b}$ 的特殊情形，则有

$$|\mathbf{a} + \mathbf{b}|^2 - |\mathbf{a}|^2 - |\mathbf{b}|^2 = 4|\mathbf{a}|^2 - |\mathbf{a}|^2 - |\mathbf{a}|^2 = 2|\mathbf{a}|^2$$

总之，对于任给两个位移向量 \mathbf{a}, \mathbf{b} ，下述函数

$$(5.1) \quad f(\mathbf{a}, \mathbf{b}) = \frac{1}{2} \{ |\mathbf{a} + \mathbf{b}|^2 - |\mathbf{a}|^2 - |\mathbf{b}|^2 \}$$

是一个值得研讨的几何量，例如 $f(\mathbf{a}, \mathbf{b}) = 0$ 乃是 \mathbf{a}, \mathbf{b} 互相垂直的充要条件，而 $f(\mathbf{a}, \mathbf{a}) = |\mathbf{a}|^2$ 。所以它显然是一个和 \mathbf{a}, \mathbf{b} 的长度、夹角都密切相关的几何量。但是归根究底 (5.1)-式所定义的几何量是否真正有用、好用，还得要看它是否具有简洁好用的优良性质。它显然具有对称性，即 $f(\mathbf{a}, \mathbf{b}) = f(\mathbf{b}, \mathbf{a})$ ，而详加研讨的结果会发现它其实还具有下述简洁易算的性质，即

$$(5.2) \quad f(\mathbf{a}, \mathbf{b} + \mathbf{c}) = f(\mathbf{a}, \mathbf{b}) + f(\mathbf{a}, \mathbf{c})$$

若以 $f(\mathbf{a}, \mathbf{b})$ 的定义（即 (5.1)-式）代入 (5.2)-式，即得所需证者，实乃下述含有三个任意向量的恒等式，亦即

$$(5.2') \quad \begin{aligned} & 2\{f(\mathbf{a}, \mathbf{b} + \mathbf{c}) - f(\mathbf{a}, \mathbf{b}) - f(\mathbf{a}, \mathbf{c})\} \\ & := |\mathbf{a} + \mathbf{b} + \mathbf{c}|^2 - |\mathbf{a} + \mathbf{b}|^2 - |\mathbf{b} + \mathbf{c}|^2 - |\mathbf{c} + \mathbf{a}|^2 \\ & \quad + |\mathbf{a}|^2 + |\mathbf{b}|^2 + |\mathbf{c}|^2 \stackrel{?}{=} 0 \end{aligned}$$

要证明上述对于任给三个向量 $\{\mathbf{a}, \mathbf{b}, \mathbf{c}\}$ 都普遍成立的恒等式之前，自然要看一看是否有一种对于任给二个向量 $\{\mathbf{u}, \mathbf{v}\}$ 都普遍成立的恒等式呢？若有，则一来肯定比较容易证明，二来说不定还可以把「後者」用来证明「前者」。要把上述想法付诸实践，当然就得有一个「後者」究竟是怎麼样的恒等式的「猜想」才能进而证明之，是不？

在此，我们自然要用上反推法去「按图索骥」。亦即假想 (5.2)-式成立的话，应该会有那种对于任何一对向量都普遍成立的恒等式？若用 (5.2)-式反推，则有

$$\begin{aligned} |\mathbf{u} + \mathbf{v}|^2 &= f(\mathbf{u} + \mathbf{v}, \mathbf{u} + \mathbf{v}) && (\text{用 (5.2)-式展开}) \\ &= |\mathbf{u}|^2 + |\mathbf{v}|^2 + f(\mathbf{u}, \mathbf{v}) + f(\mathbf{v}, \mathbf{u}) \\ |\mathbf{u} - \mathbf{v}|^2 &= f(\mathbf{u} - \mathbf{v}, \mathbf{u} - \mathbf{v}) && (\text{用 (5.2)-式展开}) \\ &= |\mathbf{u}|^2 + |\mathbf{v}|^2 - f(\mathbf{u}, \mathbf{v}) - f(\mathbf{v}, \mathbf{u}) \end{aligned}$$

由此可见，应该有恒等式（称之为广义勾股定理）

$$(5.3) \quad |\mathbf{u} + \mathbf{v}|^2 + |\mathbf{u} - \mathbf{v}|^2 \equiv 2|\mathbf{u}|^2 + 2|\mathbf{v}|^2$$

请注意，上面这一小段反推法分析只是说明：假如 (5.2)-式恒成立，则 (5.3)-式也恒成立。而我们真正要做的是先用几何直接证明 (5.3)-式恒成立，然后再设法用它来证明 (5.2)-式（亦即 (5.2')-式）恒成立。反推法的分析其实只是让我们想到 (5.3)-式恒成立这个待证的猜想。在论证上述猜想之前，不妨先对几个特别简单的情形，看一看它是否成立，亦即在 \mathbf{u}, \mathbf{v} 之间的夹角是 $0, \frac{\pi}{2}$ 和 π 这三种情形：

当 $\angle \mathbf{u}, \mathbf{v} = 0$ 时， $|\mathbf{u} \pm \mathbf{v}| = ||\mathbf{u}| \pm |\mathbf{v}||$ ，所以

$$|\mathbf{u} + \mathbf{v}|^2 + |\mathbf{u} - \mathbf{v}|^2 = (|\mathbf{u}| + |\mathbf{v}|)^2 + (|\mathbf{u}| - |\mathbf{v}|)^2 = 2|\mathbf{u}|^2 + 2|\mathbf{v}|^2$$

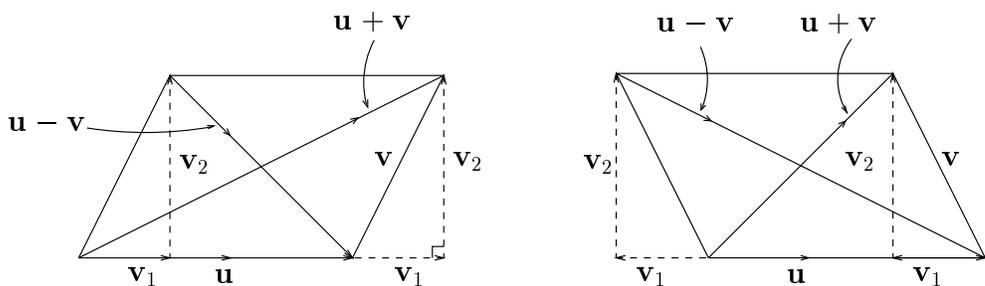
当 $\angle \mathbf{u}, \mathbf{v} = \pi$ 时， $|\mathbf{u} \pm \mathbf{v}| = ||\mathbf{u}| \mp |\mathbf{v}||$ ，所以

$$|\mathbf{u} + \mathbf{v}|^2 + |\mathbf{u} - \mathbf{v}|^2 = (|\mathbf{u}| - |\mathbf{v}|)^2 + (|\mathbf{u}| + |\mathbf{v}|)^2 = 2|\mathbf{u}|^2 + 2|\mathbf{v}|^2$$

当 $\mathbf{u} \perp \mathbf{v}$ 时，由勾股定理，即有

$$\begin{aligned} |\mathbf{u} + \mathbf{v}|^2 &= |\mathbf{u} - \mathbf{v}|^2 = |\mathbf{u}|^2 + |\mathbf{v}|^2 \\ \Rightarrow |\mathbf{u} + \mathbf{v}|^2 + |\mathbf{u} - \mathbf{v}|^2 &= 2|\mathbf{u}|^2 + 2|\mathbf{v}|^2 \end{aligned}$$

上述对于三种简单特例的验证，其实也提供了下述把一般的情形的证明归于上述业已验证的三种特殊情形来加以推导的思路，如 [图 5-3] 所示，我们可以用垂直投影把 \mathbf{v} 分解成 $\mathbf{v}_1 + \mathbf{v}_2$ ，其中 \mathbf{v}_2 和 \mathbf{u} 垂直而 \mathbf{v}_1 和 \mathbf{u} 同向（或反向）。



[图 5-3]

这样，就可以把 (5.3)-式的证明归于上述三种业已验证的情形作如下推导：

$$\begin{aligned}
 (5.4) \quad & |\mathbf{u} + \mathbf{v}|^2 = |\mathbf{u} + \mathbf{v}_1|^2 + |\mathbf{v}_2|^2 \\
 & |\mathbf{u} - \mathbf{v}|^2 = |\mathbf{u} - \mathbf{v}_1|^2 + |\mathbf{v}_2|^2 \\
 & |\mathbf{u} + \mathbf{v}|^2 + |\mathbf{u} - \mathbf{v}|^2 = |\mathbf{u} + \mathbf{v}_1|^2 + |\mathbf{u} - \mathbf{v}_1|^2 + 2|\mathbf{v}_2|^2 \\
 & = 2|\mathbf{u}|^2 + 2|\mathbf{v}_1|^2 + 2|\mathbf{v}_2|^2 = 2|\mathbf{u}|^2 + 2|\mathbf{v}|^2
 \end{aligned}$$

现在让我们再用刚才证明的 (5.3)-式纯代数地去推导 (5.2')-式的普遍成立：

令 $\mathbf{u} = \mathbf{a} + \mathbf{b}$, $\mathbf{v} = \mathbf{c}$, 即有

$$(i) \quad |\mathbf{a} + \mathbf{b} + \mathbf{c}|^2 + |\mathbf{a} + \mathbf{b} - \mathbf{c}|^2 - 2|\mathbf{a} + \mathbf{b}|^2 - 2|\mathbf{c}|^2 = 0$$

令 $\mathbf{u} = \mathbf{a}$, $\mathbf{v} = \mathbf{b} - \mathbf{c}$, 即有

$$(ii) \quad -|\mathbf{a} + \mathbf{b} - \mathbf{c}|^2 - |\mathbf{a} - \mathbf{b} + \mathbf{c}|^2 + 2|\mathbf{a}|^2 + 2|\mathbf{b} - \mathbf{c}|^2 = 0$$

令 $\mathbf{u} = \mathbf{a} + \mathbf{c}$, $\mathbf{v} = \mathbf{b}$, 即有

$$(iii) \quad |\mathbf{a} - \mathbf{b} + \mathbf{c}|^2 + |\mathbf{a} + \mathbf{b} + \mathbf{c}|^2 - 2|\mathbf{a} + \mathbf{c}|^2 - 2|\mathbf{b}|^2 = 0$$

令 $\mathbf{u} = \mathbf{b}$, $\mathbf{v} = \mathbf{c}$, 即有

$$(iv) \quad -2|\mathbf{b} + \mathbf{c}|^2 - 2|\mathbf{b} - \mathbf{c}|^2 + 4|\mathbf{b}|^2 + 4|\mathbf{c}|^2 = 0$$

将上述四个恒等式相加後再遍乘以 $\frac{1}{2}$, 即得恒等式 (5.2'), 亦即 (5.2)-式普遍成立。

总结上面的讨论，我们以勾股定理为基础，证明几何量 $f(\mathbf{a}, \mathbf{b}) = \frac{1}{2}\{|\mathbf{a} + \mathbf{b}|^2 - |\mathbf{a}|^2 - |\mathbf{b}|^2\}$ 具有 (5.2)-式所表达既简且精的性质，它将是用向量去研讨几何广泛有用的有力工具。在向量代数中，我们索兴把

它想成是一种由两个向量求得一个数值的一种乘积，叫做向量的内积 (inner product) 并改用符号 $\mathbf{a} \cdot \mathbf{b}$ 表达之，亦即以

$$(5.1') \quad \mathbf{a} \cdot \mathbf{b} = \frac{1}{2} \{ |\mathbf{a} + \mathbf{b}|^2 - |\mathbf{a}|^2 - |\mathbf{b}|^2 \}$$

为向量内积的定义式。这样做的基本原由就是使得性质 (5.2) 可以直截了当地改写成

$$(5.2'') \quad \mathbf{a} \cdot (\mathbf{b} + \mathbf{c}) = \mathbf{a} \cdot \mathbf{b} + \mathbf{a} \cdot \mathbf{c}$$

这种分配律的形式，使得它运用起来能够更加得心应手。

【定义】：(向量内积) 位移向量 \mathbf{a}, \mathbf{b} 的内积 $\mathbf{a} \cdot \mathbf{b}$ 定义为

$$\mathbf{a} \cdot \mathbf{b} = \frac{1}{2} \{ |\mathbf{a} + \mathbf{b}|^2 - |\mathbf{a}|^2 - |\mathbf{b}|^2 \}$$

内积的运算律：

$$(i) \quad \mathbf{a} \cdot \mathbf{b} = \mathbf{b} \cdot \mathbf{a}$$

$$(ii) \quad \mathbf{a} \cdot (\mathbf{b} + \mathbf{c}) = \mathbf{a} \cdot \mathbf{b} + \mathbf{a} \cdot \mathbf{c}$$

$$(iii) \quad (k\mathbf{a}) \cdot \mathbf{b} = \mathbf{a} \cdot (k\mathbf{b}) = k(\mathbf{a} \cdot \mathbf{b})$$

[当 k 是整数或有理数时，(iii) 是 (i) 和 (ii) 的推论。当 k 是非比实数时，则可用倍积的比较原则和 Eudoxus 原理加以推导。]

内积的几何意义：

$$(i) \quad \mathbf{a} \cdot \mathbf{a} = |\mathbf{a}|^2$$

$$(ii) \quad \mathbf{a} \cdot \mathbf{b} = 0 \Leftrightarrow \mathbf{a} \perp \mathbf{b} \quad (\text{亦即 } \angle \mathbf{a}, \mathbf{b} = \pm \frac{\pi}{2})$$

(iii) 在 $\angle \mathbf{a}, \mathbf{b} = \theta$ 的一般情形

$$\mathbf{a} \cdot \mathbf{b} = |\mathbf{a}| \cdot |\mathbf{b}| \cos \theta$$

[(iii) 的证明] 先验证 $\theta = 0$ 或 $\theta = \pi$ 这两种特殊情形。

若 $\theta = 0$ ，则有 $|\mathbf{a} + \mathbf{b}| = |\mathbf{a}| + |\mathbf{b}|$ 。所以

$$\mathbf{a} \cdot \mathbf{b} = \frac{1}{2} \{ (|\mathbf{a}| + |\mathbf{b}|)^2 - |\mathbf{a}|^2 - |\mathbf{b}|^2 \} = |\mathbf{a}| \cdot |\mathbf{b}| = |\mathbf{a}| \cdot |\mathbf{b}| \cos 0$$

若 $\theta = \pi$ ，则有 $|\mathbf{a} + \mathbf{b}| = \left| |\mathbf{a}| - |\mathbf{b}| \right|$ ，所以

$$\mathbf{a} \cdot \mathbf{b} = \frac{1}{2} \{ (|\mathbf{a}| - |\mathbf{b}|)^2 - |\mathbf{a}|^2 - |\mathbf{b}|^2 \} = -|\mathbf{a}| \cdot |\mathbf{b}| = |\mathbf{a}| \cdot |\mathbf{b}| \cos \pi$$

在一般的情形，可将 \mathbf{b} 分解成

$$\mathbf{b} = \mathbf{b}_1 + \mathbf{b}_2, \quad \mathbf{b}_1 \perp \mathbf{b}_2$$

而 $\angle(\mathbf{a}, \mathbf{b}_1) = 0$ 或 π ，则有

$$\begin{aligned} \mathbf{a} \cdot \mathbf{b} &= \mathbf{a} \cdot (\mathbf{b}_1 + \mathbf{b}_2) = \mathbf{a} \cdot \mathbf{b}_1 + \mathbf{a} \cdot \mathbf{b}_2 \\ &= \mathbf{a} \cdot \mathbf{b}_1 = \pm |\mathbf{a}| \cdot |\mathbf{b}_1| \end{aligned}$$

而 $\pm |\mathbf{b}_1| = |\mathbf{b}| \cos \theta$ 。 □

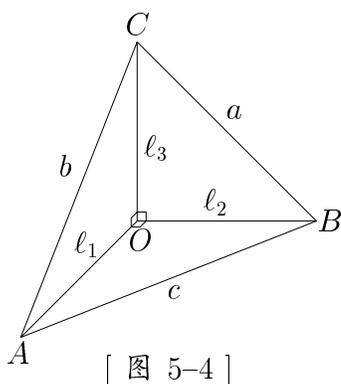
[注]：上述公式提供了用内积表达两个非零向量的夹角余弦的公式，即

$$\cos \theta = \frac{\mathbf{a} \cdot \mathbf{b}}{|\mathbf{a}| \cdot |\mathbf{b}|} = \frac{\mathbf{a} \cdot \mathbf{b}}{\sqrt{\mathbf{a} \cdot \mathbf{a}} \sqrt{\mathbf{b} \cdot \mathbf{b}}}$$

其实，上述公式就是平面几何中熟悉的余弦定律。由此可见，长度和角度都可以用向量内积去有效计算，而内积本身又具有一套十分简明有力的运算律，特别是分配律。在本质上，内积分配律乃是勾股定理的提升和精简之所得，也可以说是勾股定理代数化的最佳形式。

5.2.3 面积的勾股定理和位移向量的 \times -积

四面体是三角形的三维推广。而具有三个棱正交于一点的四面体则是直角三角形的推广，我们将称之为正交四面体，如 [图 5-4] 所示：



面积的勾股定理：

设 \overline{OA} , \overline{OB} , \overline{OC} 正交于 O ，它共交于 O 点的三个三角形 $\triangle OAB$, $\triangle OBC$, $\triangle OCA$ 互相垂直，而 $\triangle ABC$ 则是和其他三个面斜交者。是否也有类似于勾股定理的公式说明上述正交四面体的四个面积之间的关系呢？例如斜面面积的平方是否恒等于其他三个面的面积平方之和呢？这就是我们接著所要论证者。

【定理 5.3】：（面积的勾股定理）一个正交四面体的斜面面积的平方恒等于其他三个面的面积平方之和，亦即如 [图 5-4] 所示

$$(\triangle ABC)^2 = (\triangle OAB)^2 + (\triangle OBC)^2 + (\triangle OCA)^2$$

证明：令正交四面体的三个正交于 O 点的棱长分别是 l_1, l_2, l_3 ，则由勾股定理即得其斜面 $\triangle ABC$ 的三边边长分别是

$$a = \sqrt{l_2^2 + l_3^2}, \quad b = \sqrt{l_1^2 + l_3^2}, \quad c = \sqrt{l_1^2 + l_2^2}$$

易见其三个互相正交的三角形的面积分别为

$$\triangle OAB = \frac{1}{2}l_1l_2, \quad \triangle OBC = \frac{1}{2}l_2l_3, \quad \triangle OCA = \frac{1}{2}l_3l_1$$

再者，由平面几何中的秦九韶公式，其斜面面积平方是

$$\begin{aligned}
 & (\triangle ABC)^2 \\
 &= \frac{1}{16} \{2(a^2b^2 + b^2c^2 + c^2a^2) - (a^4 + b^4 + c^4)\} \\
 &= \frac{1}{16} \left\{ 2[(\ell_2^2 + \ell_3^2)(\ell_1^2 + \ell_3^2) + (\ell_1^2 + \ell_3^2)(\ell_1^2 + \ell_2^2) + (\ell_1^2 + \ell_2^2)(\ell_2^2 + \ell_3^2)] \right. \\
 &\quad \left. - [(\ell_2^2 + \ell_3^2)^2 + (\ell_1^2 + \ell_3^2)^2 + (\ell_1^2 + \ell_2^2)^2] \right\} \\
 &= \frac{1}{16} \{4\ell_1^2\ell_2^2 + 4\ell_2^2\ell_3^2 + 4\ell_3^2\ell_1^2\} \\
 &= (\triangle OAB)^2 + (\triangle OBC)^2 + (\triangle OCA)^2 \quad \square
 \end{aligned}$$

在坐标几何中，勾股定理的重要推论是下述距离公式

$$|\overline{P_1P_2}|^2 = (x_2 - x_1)^2 + (y_2 - y_1)^2 + (z_2 - z_1)^2$$

其几何意义是：空间中直线段 $\overline{P_1P_2}$ 的长度平方等于它在三个互相垂直的直线上的各别垂直投影的长度平方之和。再者，设向量 \mathbf{a} 和 \mathbf{b} 在三个坐标轴上的垂直投影的有向长度分别是 (a_1, a_2, a_3) 和 (b_1, b_2, b_3) ，则有内积坐标计算公式

$$\mathbf{a} \cdot \mathbf{b} = a_1b_1 + a_2b_2 + a_3b_3,$$

当 $\mathbf{a} = \mathbf{b}$ 时，即有

$$|\mathbf{a}|^2 = \mathbf{a} \cdot \mathbf{a} = a_1^2 + a_2^2 + a_3^2$$

由此可见上述内积公式实乃距离公式的推广。

[定理 5.3] 的几何意义是：空间中一片平面的面积平方等于它在三个互相垂直的平面（例如一个正交坐标系的三个坐标面）上的垂直投影的面积平方之和。由此可以猜想到，我们也应该试著将面积的勾股定理的本质进一步转化为在空间中一片平面和另一片平面之间的「内积」的适当定义。

假如我们把一个有向线段想成一种有向的 1-维基本几何事物，定向平行四边形 $\|(\mathbf{a}, \mathbf{b})$ 想成一种有向的 2-维基本几何事物，就自然会想到是否也可以定义一种 $\|(\mathbf{a}, \mathbf{b})$ 和 $\|(\mathbf{c}, \mathbf{d})$ 之间的「内积」使得

$$\|(\mathbf{a}, \mathbf{b}) \cdot \|(\mathbf{a}, \mathbf{b}) = \|(\mathbf{a}, \mathbf{b}) \text{ 的面积平方}$$

而且 $\iint(\mathbf{a}, \mathbf{b})$ 和 $\iint(\mathbf{c}, \mathbf{d})$ 的「内积」也可以有类似于上述的坐标计算公式。

让我们先来看看 1-维的情形。设 \overrightarrow{AB} 和 \overrightarrow{CD} 共线，则两者的内积就等于它们的有向长度的乘积，即：

$$\overrightarrow{AB} \cdot \overrightarrow{CD} = \begin{cases} |\overrightarrow{AB}| \cdot |\overrightarrow{CD}| & \text{若两者同向,} \\ -|\overrightarrow{AB}| \cdot |\overrightarrow{CD}| & \text{若两者异向.} \end{cases}$$

若 \overrightarrow{AB} 和 \overrightarrow{CD} 不共线，令 $\overrightarrow{C'D'}$ 是 \overrightarrow{CD} 在 AB 上的垂直投影，则

$$\overrightarrow{AB} \cdot \overrightarrow{CD} = \overrightarrow{AB} \text{ 和 } \overrightarrow{C'D'} \text{ 的有向长度之乘积。}$$

现在让我们试著对于 $\iint(\mathbf{a}, \mathbf{b})$ 和 $\iint(\mathbf{c}, \mathbf{d})$ 的内积作类同的定义如下，即

- (i) 当 $\iint(\mathbf{a}, \mathbf{b})$ 和 $\iint(\mathbf{c}, \mathbf{d})$ 共面时，定义其内积为两者的定向面积的乘积；
- (ii) 当 $\iint(\mathbf{a}, \mathbf{b})$ 和 $\iint(\mathbf{c}, \mathbf{d})$ 不共面时，则定义其内积为 $\iint(\mathbf{a}, \mathbf{b})$ 和 $\iint(\mathbf{c}', \mathbf{d}')$ 的定向面积的乘积，其中 $\iint(\mathbf{c}', \mathbf{d}')$ 乃是 $\iint(\mathbf{c}, \mathbf{d})$ 在 $\iint(\mathbf{a}, \mathbf{b})$ 所在平面上的垂直投影。

且以符号 $\langle \iint(\mathbf{a}, \mathbf{b}), \iint(\mathbf{c}, \mathbf{d}) \rangle$ 表示上面所定义的量。

【定理 5.4】：

$$\langle \iint(\mathbf{a}, \mathbf{b}), \iint(\mathbf{c}, \mathbf{d}) \rangle = \begin{vmatrix} \mathbf{a} \cdot \mathbf{c} & \mathbf{b} \cdot \mathbf{c} \\ \mathbf{a} \cdot \mathbf{d} & \mathbf{b} \cdot \mathbf{d} \end{vmatrix}$$

证明：(i) 设 $\iint(\mathbf{a}, \mathbf{b})$ 和 $\iint(\mathbf{c}, \mathbf{d})$ 共在一个平面 Π 之内。在 Π 上取定一组向量 $\{\mathbf{e}_1, \mathbf{e}_2\}$ ， $|\mathbf{e}_1| = |\mathbf{e}_2| = 1$ ， $\angle(\mathbf{e}_1, \mathbf{e}_2) = \frac{\pi}{2}$ 。令 $\mathbf{a} \cdot \mathbf{e}_i = a_i$ ， $\mathbf{b} \cdot \mathbf{e}_i = b_i$ ， $\mathbf{c} \cdot \mathbf{e}_i = c_i$ ， $\mathbf{d} \cdot \mathbf{e}_i = d_i$ ， $i = 1, 2$ ，则由上述定义和行列式乘法公式即有

$$\begin{aligned} \langle \iint(\mathbf{a}, \mathbf{b}), \iint(\mathbf{c}, \mathbf{d}) \rangle &= \begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix} \cdot \begin{vmatrix} c_1 & d_1 \\ c_2 & d_2 \end{vmatrix} \\ &= \begin{vmatrix} a_1c_1 + a_2c_2 & b_1c_1 + b_2c_2 \\ a_1d_1 + a_2d_2 & b_1d_1 + b_2d_2 \end{vmatrix} \\ &= \begin{vmatrix} \mathbf{a} \cdot \mathbf{c} & \mathbf{b} \cdot \mathbf{c} \\ \mathbf{a} \cdot \mathbf{d} & \mathbf{b} \cdot \mathbf{d} \end{vmatrix} \end{aligned}$$

(ii) 设 $\text{//}(\mathbf{a}, \mathbf{b})$ 和 $\text{//}(\mathbf{c}, \mathbf{d})$ 不共面，而 $\text{//}(\mathbf{c}', \mathbf{d}')$ 则是 $\text{//}(\mathbf{c}, \mathbf{d})$ 垂直投影到 $\text{//}(\mathbf{a}, \mathbf{b})$ 的所在平面 Π 的影象，由定义和上述所已证者，即有

$$\begin{aligned} \langle \text{//}(\mathbf{a}, \mathbf{b}), \text{//}(\mathbf{c}, \mathbf{d}) \rangle &= \langle \text{//}(\mathbf{a}, \mathbf{b}), \text{//}(\mathbf{c}', \mathbf{d}') \rangle \\ &= \begin{vmatrix} \mathbf{a} \cdot \mathbf{c}' & \mathbf{b} \cdot \mathbf{c}' \\ \mathbf{a} \cdot \mathbf{d}' & \mathbf{b} \cdot \mathbf{d}' \end{vmatrix} \end{aligned}$$

再者， $\mathbf{c} - \mathbf{c}'$ 和 $\mathbf{d} - \mathbf{d}'$ 乃是垂直于 Π 的向量，而 \mathbf{a}, \mathbf{b} 则是位于 Π 之内者，所以

$$\begin{aligned} \mathbf{a} \cdot (\mathbf{c} - \mathbf{c}') &= 0, & \mathbf{b} \cdot (\mathbf{c} - \mathbf{c}') &= 0 \\ \mathbf{a} \cdot (\mathbf{d} - \mathbf{d}') &= 0, & \mathbf{b} \cdot (\mathbf{d} - \mathbf{d}') &= 0 \end{aligned}$$

亦即

$$\mathbf{a} \cdot \mathbf{c}' = \mathbf{a} \cdot \mathbf{c}, \quad \mathbf{b} \cdot \mathbf{c}' = \mathbf{b} \cdot \mathbf{c}, \quad \mathbf{a} \cdot \mathbf{d}' = \mathbf{a} \cdot \mathbf{d}, \quad \mathbf{b} \cdot \mathbf{d}' = \mathbf{b} \cdot \mathbf{d}$$

即已证得

$$\langle \text{//}(\mathbf{a}, \mathbf{b}), \text{//}(\mathbf{c}, \mathbf{d}) \rangle = \begin{vmatrix} \mathbf{a} \cdot \mathbf{c} & \mathbf{b} \cdot \mathbf{c} \\ \mathbf{a} \cdot \mathbf{d} & \mathbf{b} \cdot \mathbf{d} \end{vmatrix} \quad \square$$

[定理 5.4] 的公式充分显示这种内积的 2-维推广肯定就是我们所想要者，例如

$$\begin{aligned} \langle \text{//}(\mathbf{a}, \mathbf{b}), \text{//}(\mathbf{a}, \mathbf{b}) \rangle &= \begin{vmatrix} \mathbf{a} \cdot \mathbf{a} & \mathbf{b} \cdot \mathbf{a} \\ \mathbf{a} \cdot \mathbf{b} & \mathbf{b} \cdot \mathbf{b} \end{vmatrix} \\ &= |\mathbf{a}|^2 \cdot |\mathbf{b}|^2 - (|\mathbf{a}| \cdot |\mathbf{b}| \cos \theta)^2 \\ &= (|\mathbf{a}| \cdot |\mathbf{b}| \sin \theta)^2 \\ &= \text{//}(\mathbf{a}, \mathbf{b}) \text{ 的面积平方} \end{aligned}$$

再者，设 $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$ 分别是一个取定正交坐标系在 x, y, z 轴上的单位向量。令

$$\mathbf{a} = a_1 \mathbf{e}_1 + a_2 \mathbf{e}_2 + a_3 \mathbf{e}_3$$

$$\mathbf{b} = b_1 \mathbf{e}_1 + b_2 \mathbf{e}_2 + b_3 \mathbf{e}_3$$

$$\mathbf{c} = c_1 \mathbf{e}_1 + c_2 \mathbf{e}_2 + c_3 \mathbf{e}_3$$

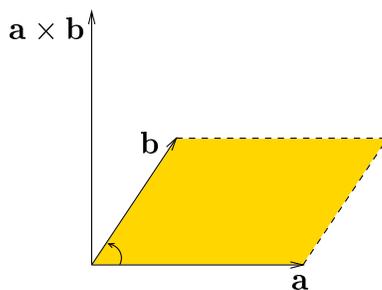
$$\mathbf{d} = d_1 \mathbf{e}_1 + d_2 \mathbf{e}_2 + d_3 \mathbf{e}_3$$

不难用直接计算去验证：

$$\begin{aligned}
 & \langle \mathbb{A}(\mathbf{a}, \mathbf{b}), \mathbb{A}(\mathbf{c}, \mathbf{d}) \rangle \\
 &= \begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix} \cdot \begin{vmatrix} c_1 & d_1 \\ c_2 & d_2 \end{vmatrix} + \begin{vmatrix} a_2 & b_2 \\ a_3 & b_3 \end{vmatrix} \cdot \begin{vmatrix} c_2 & d_2 \\ c_3 & d_3 \end{vmatrix} \\
 &\quad + \begin{vmatrix} a_3 & b_3 \\ a_1 & b_1 \end{vmatrix} \cdot \begin{vmatrix} c_3 & d_3 \\ c_1 & d_1 \end{vmatrix} \\
 &= \langle \mathbb{A}(\mathbf{a}, \mathbf{b}), \mathbb{A}(\mathbf{e}_1, \mathbf{e}_2) \rangle \cdot \langle \mathbb{A}(\mathbf{c}, \mathbf{d}), \mathbb{A}(\mathbf{e}_1, \mathbf{e}_2) \rangle \\
 &\quad + \langle \mathbb{A}(\mathbf{a}, \mathbf{b}), \mathbb{A}(\mathbf{e}_2, \mathbf{e}_3) \rangle \cdot \langle \mathbb{A}(\mathbf{c}, \mathbf{d}), \mathbb{A}(\mathbf{e}_2, \mathbf{e}_3) \rangle \\
 &\quad + \langle \mathbb{A}(\mathbf{a}, \mathbf{b}), \mathbb{A}(\mathbf{e}_3, \mathbf{e}_1) \rangle \cdot \langle \mathbb{A}(\mathbf{c}, \mathbf{d}), \mathbb{A}(\mathbf{e}_3, \mathbf{e}_1) \rangle
 \end{aligned}$$

向量的 \times -积：

在一个已经定向的空间之中（通常约定以右手型为所选之正向），我们可以用一个唯一的向量 $(\mathbf{a} \times \mathbf{b})$ 去充分表达空间中一个（定向）平行四边形 $\mathbb{A}(\mathbf{a}, \mathbf{b})$ 的方向和面积，如 [图 5-5] 所示。它是一个和 \mathbf{a}, \mathbf{b} 皆为正交，长度等于 $\mathbb{A}(\mathbf{a}, \mathbf{b})$ 的面积而 $(\mathbf{a}, \mathbf{b}, \mathbf{a} \times \mathbf{b})$ 是右手型者，称之为 \mathbf{a}, \mathbf{b} 的 \times -积。



[图 5-5]

由上述 \times -积的定义易见

$$\mathbf{a} \times \mathbf{b} = -(\mathbf{b} \times \mathbf{a})$$

而且对于任给 \mathbf{c} ，

$$(\mathbf{a} \times \mathbf{b}) \cdot \mathbf{c} = V(\mathbf{a}, \mathbf{b}, \mathbf{c})$$

其实上式就是平行六面体的体积等于底面积乘高的向量表达式。

【定理 5.5】：向量的 \times -积满足下列运算律：

$$(i) \mathbf{a} \times \mathbf{b} = -(\mathbf{b} \times \mathbf{a})$$

$$(ii) (k\mathbf{a}) \times \mathbf{b} = \mathbf{a} \times (k\mathbf{b}) = k(\mathbf{a} \times \mathbf{b})$$

$$(iii) (\mathbf{a} \times \mathbf{b}) \cdot \mathbf{c} = \mathbf{a} \cdot (\mathbf{b} \times \mathbf{c}) = V(\mathbf{a}, \mathbf{b}, \mathbf{c})$$

$$(iv) \mathbf{a} \times (\mathbf{b}_1 + \mathbf{b}_2) = \mathbf{a} \times \mathbf{b}_1 + \mathbf{a} \times \mathbf{b}_2$$

$$(v) (\mathbf{a} \times \mathbf{b}) \cdot (\mathbf{c} \times \mathbf{d}) = \begin{vmatrix} \mathbf{a} \cdot \mathbf{c} & \mathbf{b} \cdot \mathbf{c} \\ \mathbf{a} \cdot \mathbf{d} & \mathbf{b} \cdot \mathbf{d} \end{vmatrix}$$

$$(vi) (\mathbf{a} \times \mathbf{b}) \times \mathbf{c} - \mathbf{a} \times (\mathbf{b} \times \mathbf{c}) = (\mathbf{a} \cdot \mathbf{b})\mathbf{c} - (\mathbf{b} \cdot \mathbf{c})\mathbf{a}$$

证明 :

$$(ii) [(k\mathbf{a}) \times \mathbf{b} - k(\mathbf{a} \times \mathbf{b})] \cdot \mathbf{c} = V(k\mathbf{a}, \mathbf{b}, \mathbf{c}) - kV(\mathbf{a}, \mathbf{b}, \mathbf{c}) = 0, \quad \forall \mathbf{c} \\ \Rightarrow k\mathbf{a} \times \mathbf{b} - k(\mathbf{a} \times \mathbf{b}) = \mathbf{0}$$

$$(iii) (\mathbf{a} \times \mathbf{b}) \cdot \mathbf{c} = V(\mathbf{a}, \mathbf{b}, \mathbf{c}),$$

$$\mathbf{a} \cdot (\mathbf{b} \times \mathbf{c}) = (\mathbf{b} \times \mathbf{c}) \cdot \mathbf{a} = V(\mathbf{b}, \mathbf{c}, \mathbf{a}) = V(\mathbf{a}, \mathbf{b}, \mathbf{c})$$

$$(iv) [\mathbf{a} \times (\mathbf{b}_1 + \mathbf{b}_2) - \mathbf{a} \times \mathbf{b}_1 - \mathbf{a} \times \mathbf{b}_2] \cdot \mathbf{c}$$

$$= V(\mathbf{a}, \mathbf{b}_1 + \mathbf{b}_2, \mathbf{c}) - V(\mathbf{a}, \mathbf{b}_1, \mathbf{c}) - V(\mathbf{a}, \mathbf{b}_2, \mathbf{c}) = 0, \quad \forall \mathbf{c} \\ \Rightarrow \mathbf{a} \times (\mathbf{b}_1 + \mathbf{b}_2) - \mathbf{a} \times \mathbf{b}_1 - \mathbf{a} \times \mathbf{b}_2 = \mathbf{0}$$

(v) 乃是[定理 5.4]和 \times -积定义的直接结合。

$$(vi) [(\mathbf{a} \times \mathbf{b}) \times \mathbf{c} - \mathbf{a} \times (\mathbf{b} \times \mathbf{c}) - (\mathbf{a} \cdot \mathbf{b})\mathbf{c} + (\mathbf{b} \cdot \mathbf{c})\mathbf{a}] \cdot \mathbf{d}$$

$$= [(\mathbf{a} \times \mathbf{b}) \times \mathbf{c}] \cdot \mathbf{d} - [\mathbf{a} \times (\mathbf{b} \times \mathbf{c})] \cdot \mathbf{d} - (\mathbf{a} \cdot \mathbf{b})(\mathbf{c} \cdot \mathbf{d}) + (\mathbf{b} \cdot \mathbf{c})(\mathbf{a} \cdot \mathbf{d})$$

$$= (\mathbf{a} \times \mathbf{b}) \cdot (\mathbf{c} \times \mathbf{d}) - (\mathbf{d} \times \mathbf{a}) \cdot (\mathbf{b} \times \mathbf{c}) - (\mathbf{a} \cdot \mathbf{b})(\mathbf{c} \cdot \mathbf{d}) + (\mathbf{b} \cdot \mathbf{c})(\mathbf{a} \cdot \mathbf{d})$$

$$= \begin{vmatrix} \mathbf{a} \cdot \mathbf{c} & \mathbf{b} \cdot \mathbf{c} \\ \mathbf{a} \cdot \mathbf{d} & \mathbf{b} \cdot \mathbf{d} \end{vmatrix} - \begin{vmatrix} \mathbf{d} \cdot \mathbf{b} & \mathbf{a} \cdot \mathbf{b} \\ \mathbf{d} \cdot \mathbf{c} & \mathbf{a} \cdot \mathbf{c} \end{vmatrix} - (\mathbf{a} \cdot \mathbf{b})(\mathbf{c} \cdot \mathbf{d}) + (\mathbf{b} \cdot \mathbf{c})(\mathbf{a} \cdot \mathbf{d})$$

$$= 0 \quad \forall \mathbf{d}$$

$$\Rightarrow (\mathbf{a} \times \mathbf{b}) \times \mathbf{c} - \mathbf{a} \times (\mathbf{b} \times \mathbf{c}) - (\mathbf{a} \cdot \mathbf{b})\mathbf{c} + (\mathbf{b} \cdot \mathbf{c})\mathbf{a} = \mathbf{0}$$

四元数——时空的代数：

时间是一维的，空间是三维的，所以时空组合的总体是四维的，亦即

$$\mathbb{R} \oplus \mathfrak{U} = \{(t, \mathbf{a}); t \in \mathbb{R}, \mathbf{a} \in \mathfrak{U}\}$$

Hamilton 的创见赋予时空一种既自然又美妙的代数结构，这也就是著名的四元数 (quaternions)。其加、乘运算的定义如下：

$$\begin{aligned}(t_1, \mathbf{a}_1) + (t_2, \mathbf{a}_2) &= (t_1 + t_2, \mathbf{a}_1 + \mathbf{a}_2) \\ (t_1, \mathbf{a}_1) \cdot (t_2, \mathbf{a}_2) &= (t_1 t_2 - \mathbf{a}_1 \cdot \mathbf{a}_2, t_1 \mathbf{a}_2 + t_2 \mathbf{a}_1 + \mathbf{a}_1 \times \mathbf{a}_2)\end{aligned}$$

上述加、乘运算除了乘法的交换律之外，满足所有其他各种惯用的运算律，例如加法的交换律和结合律，加、乘结合的分配律和乘法的结合律等等。其中验起来比较繁复的是乘法结合律，兹证之如下：运用分配律，可以把所要验证的要点归结到

$$[(0, \mathbf{a}) \cdot (0, \mathbf{b})] \cdot (0, \mathbf{c}) \stackrel{?}{=} (0, \mathbf{a}) \cdot [(0, \mathbf{b}) \cdot (0, \mathbf{c})]$$

这个最不平凡的情形。由前述四元数的定义，即有

$$\begin{aligned}[(0, \mathbf{a}) \cdot (0, \mathbf{b})] \cdot (0, \mathbf{c}) &= (-\mathbf{a} \cdot \mathbf{b}, \mathbf{a} \times \mathbf{b}) \cdot (0, \mathbf{c}) \\ &= (-V(\mathbf{a}, \mathbf{b}, \mathbf{c}), -(\mathbf{a} \cdot \mathbf{b})\mathbf{c} + (\mathbf{a} \times \mathbf{b}) \times \mathbf{c}) \\ (0, \mathbf{a}) \cdot [(0, \mathbf{b}) \cdot (0, \mathbf{c})] &= (0, \mathbf{a}) \cdot (-\mathbf{b} \cdot \mathbf{c}, \mathbf{b} \times \mathbf{c}) \\ &= (-V(\mathbf{a}, \mathbf{b}, \mathbf{c}), -(\mathbf{b} \cdot \mathbf{c})\mathbf{a} + \mathbf{a} \times (\mathbf{b} \times \mathbf{c}))\end{aligned}$$

由此可见，用[定理 5.5]的公式 (vi) 即有

$$[(0, \mathbf{a}) \cdot (0, \mathbf{b})] \cdot (0, \mathbf{c}) - (0, \mathbf{a}) \cdot [(0, \mathbf{b}) \cdot (0, \mathbf{c})] \equiv (0, \mathbf{0})$$

再者，我们也可以类同于复数 (complex number) 的情形定义共轭和绝对值，即

$$\begin{aligned}\overline{(t, \mathbf{a})} &= (t, -\mathbf{a}) \\ |(t, \mathbf{a})| &= \sqrt{t^2 + \mathbf{a} \cdot \mathbf{a}}\end{aligned}$$

则有

$$(t, \mathbf{a}) \cdot \overline{(t, \mathbf{a})} = (t^2 + \mathbf{a} \cdot \mathbf{a}, \mathbf{0})$$

$$(t, \mathbf{a})^{-1} = \frac{1}{|(t, \mathbf{a})|^2} \cdot (t, -\mathbf{a})$$

总之，四元数乃是研讨时空的精简有效的数学工具，它也是研究学习电磁学、狭义相对论的基本工具。

5.3 结语

总结上述对于空间的保长变换和向量代数的讨论，我们再提纲挈领地吧所得的结果和认识作一概括性的系统列述：

1. 反射对称是空间中最为简单的保长变换，而空间的所有其他保长变换又都可以由它们的组合而得之，所以它们又是最为基本者。再者，等腰三角形是一种最为初等简朴的反射对称图形；在古希腊的几何学中，能够仅用等腰三角形的各种特征性质之间的逻辑转换来分析空间对称性在几何学中的各种各样反映，究其原因其实也就是反射对称性的组合，已经含盖了保长变换的全体这个基本事实。
2. 将两个反射对称 $\{\mathfrak{R}_{\Pi_1}, \mathfrak{R}_{\Pi_2}\}$ 加以组合，其所得的保长变换以 Π_1, Π_2 是否相交而分成两种：当 $\Pi_1 \cap \Pi_2 = \phi$ 时 $\mathfrak{R}_{\Pi_2} \circ \mathfrak{R}_{\Pi_1}$ 是一个平移，它把空间每一点在 Π_1, Π_2 的公垂线上由 Π_1 往 Π_2 的方向向前移动 $2d(\Pi_1, \Pi_2)$ ；当 $\Pi_1 \cap \Pi_2 = \ell$ 时， $\mathfrak{R}_{\Pi_2} \circ \mathfrak{R}_{\Pi_1}$ 保持其交线 ℓ 上的每一点固定不动，而其他各点 P 则在其 ℓ 的垂面 $\Pi(P)$ 中以 $\Pi(P) \cap \ell$ 为中心作 $2\angle\Pi_1, \Pi_2$ 的旋转。
3. 两个平移 τ_1, τ_2 的组合还是一个平移，而且 $\tau_1 \circ \tau_2 = \tau_2 \circ \tau_1$ 。用群的术语来说，空间所有平移组成的是保长变换群的一个正规子群；从几何的本质来看，一个平移把空间每一点都作了一个同向等距的位置移动，所以它是位移的自然「量化」，称之为位移向量。总之，平移和位移向量是同一事物的两种观点；从变换观点来看叫做平移，从几何的本质来看则是位移向量，这是同一事物的两面

观，各有所长。我们用前者来定义其加法和倍积，但是在讨论长度、角度、面积、体积等等的几何内含时，则自要采取位移、有向线段这种几何的观点。在第二节中的讨论中，这种观点的自然转换是十分明显的。

4. 位置是空间最为基本原始的概念，由此可见位移向量理所当然地是空间的最为基本原始的几何量。第二节中所讨论的向量代数基础理论也就是要把其他各种各样基本几何量归结到位移向量去表达、计算；由此自然地产生了内积和 \times -积运算，但是这种顺理成章、返璞归真的探索的成果，不但求得用位移向量去表达、计算其他基本几何量的精而简的公式，而且还把定量几何学中的基本定理如相似三角形定理、勾股定理和面积勾股定理等等有系统地转换成向量代数中的运算律！例如：

(i) 向量加法的定义（即[定理 5.1]）植基于空间的平直性（亦即平行性或三角形内角和恒为平角）。在古典几何学中关于平行的基本定理就是平行四边形的各种特征性质之间的转换，而平行四边形定理所转换而得者就是向量加法的交换律！

(ii) 相似放大缩小是（欧氏）空间的特色，这也就是向量的倍积的来源。而关于相似形的基本定理——相似三角形定理——用倍积来表达就是倍积分配律：

$$k \cdot (\mathbf{a} + \mathbf{b}) = k\mathbf{a} + k\mathbf{b}$$

(iii) 关于长度和角度的基本定理——勾股定理及广义勾股定理——经过一番分析和整理之后又可以简化、优化而成为简单易用的内积分配律：

$$\mathbf{a} \cdot (\mathbf{b} + \mathbf{c}) = \mathbf{a} \cdot \mathbf{b} + \mathbf{a} \cdot \mathbf{c}$$

(iv) [定理 5.4]和[定理 5.5]所总结的 \times -积运算律乃是空间几何学中关于面积、体积、两面角等等的基本定理是也！但是这些基本定理在古希腊时期尚未能发现。

由上述分析可见向量运算不仅提供了表达计算各种各样基本几何量的有效能算的代数公式，而向量运算律本身其实就是一套完美

精简的几何基本定理，且其中重要的都是分配律或多线性函数这种简单易用的形式表达。由此可见，向量代数乃是空间结构的全面而且美妙的代数，而其运算律则是空间本质（亦即基本定理所表达者）的一种至精至简的表达。

5. 空间的基本结构在于任给两点 A, B 之间的唯一最短通路——直线段 \overline{AB} ，而空间的基本本质主要的就是对称性和平直性。在古希腊几何学中，用叠合公理来描述对称性，而用平行公理来描述平直性；在现代的几何学中，我们把空间的对称性的总体赋予自然而且全局的结构，称之为保长变换群，而其中的平移子群则系统表述了空间的平直性，从而把空间几何学的研讨提升到保长变换群的不变量理论。再者，向量运算都是在正交变换之下协变的 (covariant)，所以用向量运算所得的数量 (scalar) 都自然而然是不变的 (invariant)。

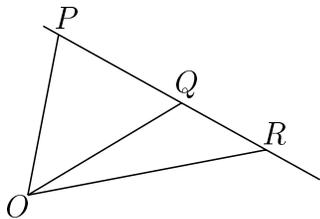
总结上述五点分析，我们认识到用向量代数研讨几何可以寓不变量理论于向量运算之中，而空间的基本性质和基本定理的运用则转化为其运算律的系统运用。这就是学习向量几何，并用以探索大自然所要达到的境界！

5.4 例题、习题和思考题

【例题】：

- (1) P, Q, R 三点共线的向量条件式：

设 P, Q, R 三点共线， O 是任选之原点，即如 [图 5-6] 所示：



[图 5-6]

则存在实数 k , 使得 $\overrightarrow{PR} = k\overrightarrow{PQ}$, 亦即

$$\begin{aligned}\overrightarrow{OR} - \overrightarrow{OP} &= k(\overrightarrow{OQ} - \overrightarrow{OP}) \\ \Rightarrow (k-1)\overrightarrow{OP} + (-k)\overrightarrow{OQ} + \overrightarrow{OR} &= \mathbf{0}\end{aligned}$$

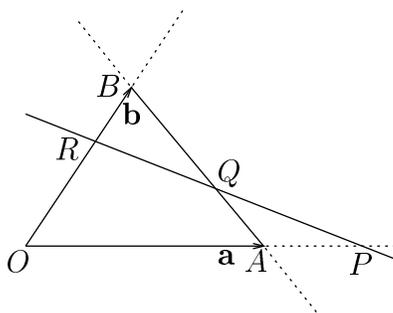
换言之, 存在有不全为零的系数 α, β, γ 满足

$$\begin{cases} \alpha\overrightarrow{OP} + \beta\overrightarrow{OQ} + \gamma\overrightarrow{OR} = \mathbf{0} \\ \alpha + \beta + \gamma = 0 \end{cases} \quad (\text{设 } \gamma \neq 0)$$

反之, 若存在满足上述条件的 α, β, γ (设 $\gamma \neq 0$), 即有

$$\beta(\overrightarrow{OQ} - \overrightarrow{OP}) + \gamma(\overrightarrow{OR} - \overrightarrow{OP}) = \mathbf{0} \Rightarrow \overrightarrow{PR} = \left(-\frac{\beta}{\gamma}\right)\overrightarrow{PQ}$$

(2) 用向量代数证明 Menelous 定理:



[图 5-7]

令 $\overrightarrow{OR} = k_1\mathbf{a}$, $\overrightarrow{OQ} = k_2\mathbf{b}$, $\overrightarrow{OQ} = x\mathbf{a} + y\mathbf{b}$ 。首先, 由 A, B, Q 三点共线的向量条件式, 即有

$$\begin{aligned}x \cdot \overrightarrow{OA} + y \cdot \overrightarrow{OB} + (-1)\overrightarrow{OQ} &= \mathbf{0} \\ \Rightarrow x + y - 1 &= 0\end{aligned}$$

再者, 由 P, Q, R 三点共线的向量条件式, 即有

$$\begin{aligned}\frac{x}{k_1} \cdot \overrightarrow{OP} + \frac{y}{k_2} \cdot \overrightarrow{OR} + (-1)\overrightarrow{OQ} &= \mathbf{0} \\ \Rightarrow \frac{x}{k_1} + \frac{y}{k_2} - 1 &= 0\end{aligned}$$

由此解得

$$x = \frac{k_1 k_2 - k_1}{k_2 - k_1}, \quad y = \frac{k_1 k_2 - k_2}{k_1 - k_2}$$

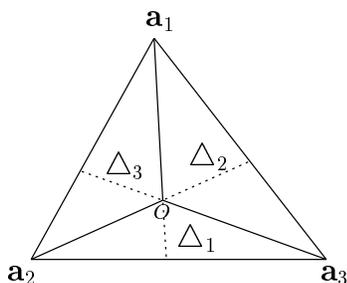
所以

$$\frac{\overrightarrow{OP}}{\overrightarrow{PA}} = \frac{k_1}{k_1 - 1}, \quad \frac{\overrightarrow{BR}}{\overrightarrow{RO}} = \frac{1 - k_2}{k_2}$$

即得

$$\frac{\overrightarrow{AQ}}{\overrightarrow{QB}} = \frac{y}{x} \Rightarrow \frac{\overrightarrow{OP}}{\overrightarrow{PA}} \cdot \frac{\overrightarrow{AQ}}{\overrightarrow{QB}} \cdot \frac{\overrightarrow{BR}}{\overrightarrow{RO}} = -1$$

(3) 重心和 Ceva 定理：



[图 5-8]

设 $\mathbf{a}_i = \overrightarrow{OP}_i$ 分别为质量 μ_i 的三个质点的位置向量，则其重心的位置向量就是

$$\frac{1}{\mu_1 + \mu_2 + \mu_3} (\mu_1 \mathbf{a}_1 + \mu_2 \mathbf{a}_2 + \mu_3 \mathbf{a}_3) = m_1 \mathbf{a}_1 + m_2 \mathbf{a}_2 + m_3 \mathbf{a}_3$$

其中 m_i 是 P_i 的质量百分比（亦即 $m_1 + m_2 + m_3 = 1$ ）。若改用重心作为原点，则 $\{\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3\}$ 满足条件式

$$m_1 \mathbf{a}_1 + m_2 \mathbf{a}_2 + m_3 \mathbf{a}_3 = \mathbf{0}$$

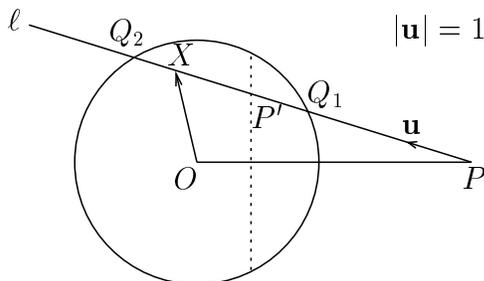
由 $(m_1 \mathbf{a}_1 + m_2 \mathbf{a}_2 + m_3 \mathbf{a}_3) \times \mathbf{a}_3 = \mathbf{0}$ 得

$$\begin{aligned} m_1 \mathbf{a}_1 \times \mathbf{a}_3 + m_2 \mathbf{a}_2 \times \mathbf{a}_3 &= \mathbf{0} \\ \Rightarrow -m_1 \Delta_2 + m_2 \Delta_1 &= 0 \\ \Rightarrow \Delta_1 : \Delta_2 : \Delta_3 &= m_1 : m_2 : m_3 \\ \Rightarrow \Delta_i &= m_i \Delta \end{aligned}$$

这其实乃是 Ceva 定理在「力学」中的表现。

(4) 圆幂定理的向量证法：

l 为过 P 点的直线，与圆交于 Q_1, Q_2 两点。令 X 为 l 上的一个动点， \mathbf{u} 是沿 l 方向的单位长向量。



[图 5-9]

易见有

$$\overrightarrow{OX} = \overrightarrow{OP} + \overrightarrow{PX} = \overrightarrow{OP} + x\mathbf{u}$$

由此可见，直线 l 上的动点 X 恰好位于以 O 为圆心，半径为 R 的圆上的条件式就是

$$\begin{aligned} R^2 &= \overrightarrow{OX} \cdot \overrightarrow{OX} = (\overrightarrow{OP} + x\mathbf{u}) \cdot (\overrightarrow{OP} + x\mathbf{u}) \\ &= \overrightarrow{OP} \cdot \overrightarrow{OP} + 2x\overrightarrow{OP} \cdot \mathbf{u} + x^2 \end{aligned}$$

上述二次方程式的两个根 λ_1, λ_2 也就是 $\overrightarrow{PQ_1}$ 和 $\overrightarrow{PQ_2}$ 的有向长度。所以

$$\begin{cases} \lambda_1 \cdot \lambda_2 = \overrightarrow{OP} \cdot \overrightarrow{OP} - R^2 \\ \lambda_1 + \lambda_2 = -2\overrightarrow{OP} \cdot \mathbf{u} \end{cases}$$

第一式的几何意义就是

$$\overrightarrow{PQ_1} \cdot \overrightarrow{PQ_2} = |\overrightarrow{OP}|^2 - R^2 \quad (\text{圆幂定理})$$

第二式的几何意义是什么呢？

令 P' 是 l 上使得 (PP', Q_1Q_2) 成调和点列之点，亦即

$$\frac{\overrightarrow{Q_1P'}}{\overrightarrow{P'Q_2}} \cdot \frac{\overrightarrow{Q_2P'}}{\overrightarrow{P'Q_1}} = -1$$

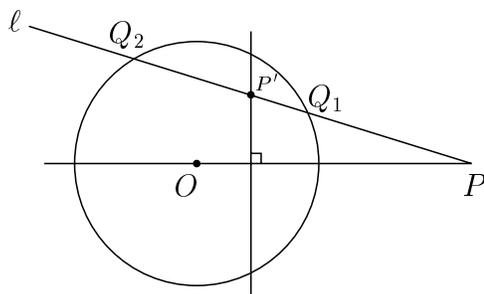
设 $\overrightarrow{PP'} = k\mathbf{u}$, 则有

$$\frac{k - \lambda_1}{\lambda_2 - k} \cdot \frac{-\lambda_2}{\lambda_1} = -1 \quad \Rightarrow \quad k = \frac{2\lambda_1\lambda_2}{\lambda_1 + \lambda_2}$$

所以

$$\begin{aligned} \overrightarrow{PP'} \cdot \overrightarrow{OP} &= k \cdot \mathbf{u} \cdot \overrightarrow{OP} = \frac{\lambda_1\lambda_2}{\lambda_1 + \lambda_2} 2\mathbf{u} \cdot \overrightarrow{OP} \\ &= -\lambda_1\lambda_2 = R^2 - |\overrightarrow{OP}|^2 \end{aligned}$$

上述向量代数式的几何意义就是： $\overrightarrow{PP'}$ 在直线 OP 上的垂直投影是一个固定的向量，亦即和 l 无关。



[图 5-10]

【习题】：

- (1) 试验证在例题 (1) 中当 $\gamma \neq 0$ 的条件换成 $\alpha \neq 0$ 或 $\beta \neq 0$ 时，同样结论亦成立。

设 l_1, l_2 是空间中两条（相异）直线， $A_1, B_1 \in l_1, A_2, B_2 \in l_2$ 为线上给定的（相异）四点。令 X_1, X_2 分别为 l_1, l_2 上的动点，则容易把 $\overrightarrow{X_1X_2}$ 改写成下述形式：

$$\overrightarrow{X_1X_2} = \overrightarrow{A_1A_2} + x_1\overrightarrow{A_1B_1} + x_2\overrightarrow{A_2B_2}$$

- (2) 问 $\overrightarrow{X_1X_2}$ 同时垂直于 l_1, l_2 的代数条件式是什么？

- (3) 试讨论在什么的情况下这个公垂线段 $\overline{X_1X_2}$ 是

(i) 唯一存在的，而且其长度不为 0；

- (ii) 唯一存在的，但是其长度为 0；
- (iii) 并非唯一存在的。
- (4) 令 $\overline{P_1P_2}$ 为题 (3)(i) 情形中的公垂线段。试证这是所有 $\overline{X_1X_2}$ 中的最短者。

第六章

坐标解析几何简介

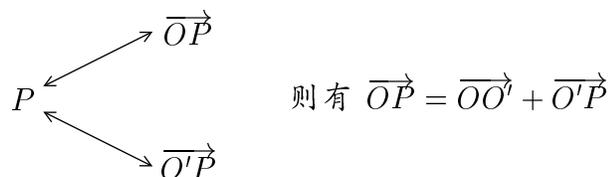
几何学在古希腊文明蓬勃进展，获得辉煌的成就，奠定了定性和定量几何的基础理论。在西方，由古希腊文明一直到十五、六世纪的文艺复兴，理性文明历经了上千年的暗淡时期。其间战乱不断，希腊文明的成果，仅赖东罗马 (Byzantine Empire) 和阿刺伯文明得以保存其大部。长话短说，1453 年东罗马帝国的灭亡迫使大批希腊学者和书稿流入意大利，和当时经由阿刺伯传入的中国印刷术，促成了希腊数理名著如 Euclid 的 “Elements”，Apollonius 的 “Conic Sections”，Diophantus 的 “Arithmetica” 等等的拉丁文版在意大利相继问世。希腊文明的硕果终于熬过了上千年的暗淡岁月，浴火重生；在欧洲大陆重新扎根、萌芽、茁壮，发展成影响全世界的近代科技文明。数学和天文学的进展，又再一次扮演著先行者、奠基者的角色，此点在近代科学的创始者如 Copernicus, Brahe, Kepler, Galileo, Vieta, Descartes, Fermat, Newton, Leibniz 等等的贡献中可见其梗概。而其中影响无比深远者则首推坐标解析几何和微积分，它们奠定了对于各种各样自然现象作深刻的数理分析的基本工具。

Descartes 和 Fermat 在坐标解析几何学上的创建，把当代新兴的代数学，系统地用来研究几何。他们的工作都直接得益于 Vieta 在代数学上的著作。总之，几何学和代数学从此得以密切结合、相辅相成、相得益彰。它不但促进了两者的大幅度进展，而且也使得微积分学的展现变得水到渠成。所以坐标解析几何学的创见，乃是整个人类理性文明发展史上的一件大事。从几何学的发展史来看，向量代数的产生乃是十九世纪中叶的事，足足比坐标解析几何的产生晚了两个世纪。如

今回顾、反思，可见向量代数其实乃是坐标几何的返璞归真和精益求精，它使得几何和代数结合得更加真切自然、直截了当。在此，我们将改用向量代数为工具回头来对于坐标解析几何作一简明扼要的介绍。这种讲法虽然是历史发展的逆序，但是这种返璞归真的向量讲法可以把解析几何的精要体现得更加简明朴实。

6.1 正交坐标系和平面（或空间）的坐标化

在平面（或空间）中取定一点 O 为基点（称之为原点），则其中任给一点 P 的位置就可以用位移向量 \overrightarrow{OP} （亦即平移 $\tau(O, P)$ ）加以唯一确定。改用集合的术语来说，即平面（或空间）中所有点所成的集合和平面（或空间）的所有位移向量（亦即平移）所构成的代数体系之间具有上述自然的一、一对应，即 $P \leftrightarrow \overrightarrow{OP}$ （亦即 $\tau(O, P)$ ）。这个一、一对应显然和原点的选取有关。例如

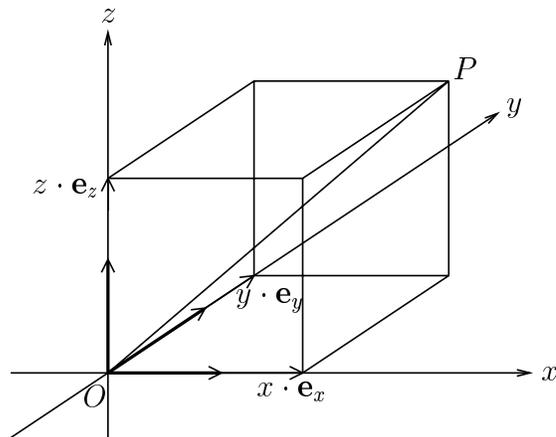


$$P \begin{cases} \nearrow \overrightarrow{OP} \\ \searrow \overrightarrow{O'P} \end{cases} \quad \text{则有 } \overrightarrow{OP} = \overrightarrow{OO'} + \overrightarrow{O'P}$$

再者，在平面上我们可以取用两个互相垂直的射线 \overrightarrow{Ox} 和 \overrightarrow{Oy} ，而在空间之中则可以取用三个互相垂直的 \overrightarrow{Ox} , \overrightarrow{Oy} 和 \overrightarrow{Oz} 。令 $\mathbf{e}_x, \mathbf{e}_y$ 和 \mathbf{e}_z 分别是以 $\overrightarrow{Ox}, \overrightarrow{Oy}$ 和 \overrightarrow{Oz} 为其方向的单位长向量，则有

$$\begin{aligned} \overrightarrow{OP} &= x \cdot \mathbf{e}_x + y \cdot \mathbf{e}_y, \quad \text{其中 } x = \overrightarrow{OP} \cdot \mathbf{e}_x, y = \overrightarrow{OP} \cdot \mathbf{e}_y \\ (\text{或 } \overrightarrow{OP} &= x \cdot \mathbf{e}_x + y \cdot \mathbf{e}_y + z \cdot \mathbf{e}_z \\ \text{其中 } x &= \overrightarrow{OP} \cdot \mathbf{e}_x, y = \overrightarrow{OP} \cdot \mathbf{e}_y, z = \overrightarrow{OP} \cdot \mathbf{e}_z) \end{aligned}$$

由此可见 P 点在平面（或空间）中的位置又可以用有序数组 (x, y) （或 (x, y, z) ）加以唯一确定，称之为 P 点在平面坐标系 $\{O; \mathbf{e}_x, \mathbf{e}_y\}$ （或空间坐标系 $\{O; \mathbf{e}_x, \mathbf{e}_y, \mathbf{e}_z\}$ ）中的坐标。如 [图 6-1] 所示， x, y, z 的几何意义乃是位移向量 \overrightarrow{OP} （亦即有向线段）在 x, y, z -轴上的垂直投影的有号长度是也。



[图 6-1]

在平面上（或空间中）选取一个正交坐标系，不但可以有系统地用有序数偶 (x, y) （或三数组 (x, y, z) ）来标记平面上（或空间中）点的位置，而且也可以用下述有序数偶（或三数组）来表达平面上（或空间中）的向量，即

$$\mathbf{a} \leftrightarrow (a_x, a_y) : a_x = \mathbf{a} \cdot \mathbf{e}_x, a_y = \mathbf{a} \cdot \mathbf{e}_y$$

$$(\text{或 } \mathbf{a} \leftrightarrow (a_x, a_y, a_z) : a_x = \mathbf{a} \cdot \mathbf{e}_x, a_y = \mathbf{a} \cdot \mathbf{e}_y, a_z = \mathbf{a} \cdot \mathbf{e}_z)$$

亦即

$$\mathbf{a} = a_x \cdot \mathbf{e}_x + a_y \cdot \mathbf{e}_y$$

$$(\text{或 } \mathbf{a} = a_x \cdot \mathbf{e}_x + a_y \cdot \mathbf{e}_y + a_z \cdot \mathbf{e}_z)$$

通常把 a_x, a_y, a_z 分别叫做向量 \mathbf{a} 在 x, y, z 方向的分量 (components)。

再者，用向量 $\{\mathbf{e}_x, \mathbf{e}_y, \mathbf{e}_z\}$ 的下述内积关系

$$\mathbf{e}_x \cdot \mathbf{e}_x = \mathbf{e}_y \cdot \mathbf{e}_y = \mathbf{e}_z \cdot \mathbf{e}_z = 1$$

$$\mathbf{e}_x \cdot \mathbf{e}_y = \mathbf{e}_y \cdot \mathbf{e}_z = \mathbf{e}_z \cdot \mathbf{e}_x = 0$$

和向量运算律即可推导改用分量来表达的向量运算公式，即

$$\mathbf{a} \leftrightarrow (a_x, a_y, a_z), \quad \mathbf{b} \leftrightarrow (b_x, b_y, b_z)$$

$$(i) \quad \mathbf{a} + \mathbf{b} \leftrightarrow (a_x + b_x, a_y + b_y, a_z + b_z)$$

$$(ii) \quad k \cdot \mathbf{a} \leftrightarrow (ka_x, ka_y, ka_z)$$

$$(iii) \mathbf{a} \cdot \mathbf{b} = a_x b_x + a_y b_y + a_z b_z$$

(iv) 若 $\{\mathbf{e}_x, \mathbf{e}_y, \mathbf{e}_z\}$ 是右手型则有

$$\mathbf{a} \times \mathbf{b} \leftrightarrow \left(\begin{vmatrix} a_y & b_y \\ a_z & b_z \end{vmatrix}, \begin{vmatrix} a_z & b_z \\ a_x & b_x \end{vmatrix}, \begin{vmatrix} a_x & b_x \\ a_y & b_y \end{vmatrix} \right)$$

兹验证 (iv)-式如下, 其他三式的验证留作习题:

由所设 $\{\mathbf{e}_x, \mathbf{e}_y, \mathbf{e}_z\}$ 的正交性和右手型, 即有

$$\mathbf{e}_x \times \mathbf{e}_y = -\mathbf{e}_y \times \mathbf{e}_x = \mathbf{e}_z$$

$$\mathbf{e}_y \times \mathbf{e}_z = -\mathbf{e}_z \times \mathbf{e}_y = \mathbf{e}_x$$

$$\mathbf{e}_z \times \mathbf{e}_x = -\mathbf{e}_x \times \mathbf{e}_z = \mathbf{e}_y$$

再用分配律, 即有

$$\begin{aligned} \mathbf{a} \times \mathbf{b} &= (a_x \mathbf{e}_x + a_y \mathbf{e}_y + a_z \mathbf{e}_z) \times (b_x \mathbf{e}_x + b_y \mathbf{e}_y + b_z \mathbf{e}_z) \\ &= (a_x b_y - a_y b_x) \mathbf{e}_x \times \mathbf{e}_y + (a_y b_z - a_z b_y) \mathbf{e}_y \times \mathbf{e}_z + (a_z b_x - a_x b_z) \mathbf{e}_z \times \mathbf{e}_x \\ &= \begin{vmatrix} a_y & b_y \\ a_z & b_z \end{vmatrix} \mathbf{e}_x + \begin{vmatrix} a_z & b_z \\ a_x & b_x \end{vmatrix} \mathbf{e}_y + \begin{vmatrix} a_x & b_x \\ a_y & b_y \end{vmatrix} \mathbf{e}_z \end{aligned}$$

将 (iii)-式和 (iv)-式相结合, 即得

$$(\mathbf{a} \times \mathbf{b}) \cdot \mathbf{c} = \begin{vmatrix} a_x & a_y & a_z \\ b_x & b_y & b_z \\ c_x & c_y & c_z \end{vmatrix}$$

由此可见, 选取正交坐标系所达成者乃是把四种向量运算, 有系统地归于向量的分量加以计算, 而这些运算公式都是正交性和分配律的直接推论。

距离公式: 直线段的长度是空间的各种各样几何量之中的最、最基本者。所以坐标几何的头号基本公式理当是给定两点 P_1, P_2 之间的距离的坐标表达式, 即

$$d(P_1, P_2)^2 = \overline{P_1 P_2} \cdot \overline{P_1 P_2} = (x_2 - x_1)^2 + (y_2 - y_1)^2 + (z_2 - z_1)^2$$

[若 P_1, P_2 共在平面 $z = 0$ 之上, 则 $z_2 - z_1 = 0$ 。]

夹角公式：设 P_0, P_1, P_2 是给定三点，则 $\overrightarrow{P_0P_1}$ 和 $\overrightarrow{P_0P_2}$ 之间的夹角 θ 的余弦可以用它们的内积公式计算之，即

$$\begin{aligned}\cos \theta &= \frac{\overrightarrow{P_0P_1} \cdot \overrightarrow{P_0P_2}}{\sqrt{\overrightarrow{P_0P_1} \cdot \overrightarrow{P_0P_1}} \sqrt{\overrightarrow{P_0P_2} \cdot \overrightarrow{P_0P_2}}} \\ &= \frac{(x_1 - x_0)(x_2 - x_0) + (y_1 - y_0)(y_2 - y_0) + (z_1 - z_0)(z_2 - z_0)}{\sqrt{(x_1 - x_0)^2 + (y_1 - y_0)^2 + (z_1 - z_0)^2} \sqrt{(x_2 - x_0)^2 + (y_2 - y_0)^2 + (z_2 - z_0)^2}}\end{aligned}$$

三角形面积公式：设 P_0, P_1, P_2 是平面上（或空间中）的给定三点，则 $\triangle P_0P_1P_2$ 的定向面积（或不定向面积）可以分别计算如下：

$$\begin{aligned}\Delta P_0P_1P_2 &= \frac{1}{2} \begin{vmatrix} 1 & 1 & 1 \\ x_0 & x_1 & x_2 \\ y_0 & y_1 & y_2 \end{vmatrix} \\ |\Delta P_0P_1P_2|^2 &= \frac{1}{4} \left(\overrightarrow{P_0P_1} \times \overrightarrow{P_0P_2} \right) \cdot \left(\overrightarrow{P_0P_1} \times \overrightarrow{P_0P_2} \right) \\ &= \frac{1}{4} \left\{ \begin{vmatrix} 1 & 1 & 1 \\ x_0 & x_1 & x_2 \\ y_0 & y_1 & y_2 \end{vmatrix}^2 + \begin{vmatrix} 1 & 1 & 1 \\ y_0 & y_1 & y_2 \\ z_0 & z_1 & z_2 \end{vmatrix}^2 + \begin{vmatrix} 1 & 1 & 1 \\ z_0 & z_1 & z_2 \\ x_0 & x_1 & x_2 \end{vmatrix}^2 \right\}\end{aligned}$$

四面体体积公式：设 P_0, P_1, P_2, P_3 是空间中给定四点，则其所张的四面体 $T(P_0, P_1, P_2, P_3)$ 的定向体积可以用下述公式表达之，即

$$V(T(P_0, P_1, P_2, P_3)) = \frac{1}{6} \begin{vmatrix} 1 & 1 & 1 & 1 \\ x_0 & x_1 & x_2 & x_3 \\ y_0 & y_1 & y_2 & y_3 \\ z_0 & z_1 & z_2 & z_3 \end{vmatrix}$$

[试用 $V(T(P_0, P_1, P_2, P_3)) = \frac{1}{6}(\overrightarrow{P_0P_1} \times \overrightarrow{P_0P_2}) \cdot \overrightarrow{P_0P_3}$ 验证上式。]

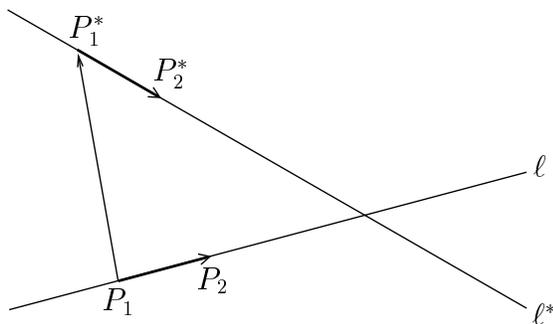
【习题】：

- (1) 试求 $\triangle P_0P_1P_2$ 和 $\triangle P_0P_1P_3$ 之间的两面角的余弦用 P_i 的坐标 $(x_i, y_i, z_i), 0 \leq i \leq 3$, 的表达式。
- (2) 试求平面上三点 $\{P_i(x_i, y_i), 0 \leq i \leq 2\}$ 共线的条件式。
- (3) 试求空间中四点 $\{P_i(x_i, y_i, z_i), 0 \leq i \leq 3\}$ 共面的条件式。

- (4) 试求空间中三点 $\{P_i(x_i, y_i, z_i), 0 \leq i \leq 2\}$ 共线的条件式。
- (5) 如 [图 6-2] 所示, 设 l, l^* 是空间中两条不交线 (亦即不共面者), $\overline{P_1P_2}$ 和 $\overline{P_1^*P_2^*}$ 分别是 l, l^* 上单位长线段。试证 $\{d(P, P^*), P \in l, P^* \in l^*\}$ 的极小值 (称之为 l, l^* 之间的距离) 等于下述比值:

$$\begin{vmatrix} 1 & 1 & 1 & 1 \\ x_1 & x_2 & x_1^* & x_2^* \\ y_1 & y_2 & y_1^* & y_2^* \\ z_1 & z_2 & z_1^* & z_2^* \end{vmatrix} : \left| \overrightarrow{P_1P_2} \times \overrightarrow{P_1^*P_2^*} \right|$$

的绝对值。



[图 6-2]

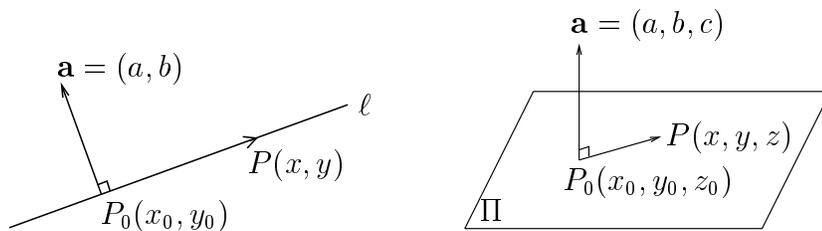
6.2 直线和圆, 平面和球

直线与圆是平面上最为精简的一维子集, 而平面与球则是空间中最为精简的二维子集。在解析几何中它们都可以用一个简单的方程式去刻画之。

【定理 6.1】: 在平面上 (或空间内) 一条直线 (或一个平面) 的点的坐标满足一个 $\{x, y\}$ (或 $\{x, y, z\}$) 的一次方程式。反之, 任给一个二元 (或三元) 一次方程式的所有「解点」构成的子集乃是 (x, y) -平面上 (或 (x, y, z) -空间内) 的一条直线 (或一个平面)。

证明: 设 $P_0(x_0, y_0)$ (或 $P_0(x_0, y_0, z_0)$) 是直线 l (或平面 Π) 的取定一点, 而 $P(x, y)$ (或 $P(x, y, z)$) 则是 l (或 Π) 上的任意点。再者, 如

[图 6-3] 所示，令 $\mathbf{a} = (a, b)$ (或 $\mathbf{a} = (a, b, c)$) 是平面上 (或空间内) 和 ℓ (或 Π) 垂直的向量：



[图 6-3]

则有

$$0 = \mathbf{a} \cdot \overrightarrow{P_0P} = a(x - x_0) + b(y - y_0)$$

$$(\text{或 } 0 = \mathbf{a} \cdot \overrightarrow{P_0P} = a(x - x_0) + b(y - y_0) + c(z - z_0))$$

亦即 $P(x, y) \in \ell$ (或 $P(x, y, z) \in \Pi$) 乃是上述二元 (或三元) 一次方程的「解点子集」。

反之，设 S (或 S^*) 分别是下述二元 (或三元) 一次方程的解点子集，即

$$ax + by + c = 0$$

$$(\text{或 } ax + by + cz + d = 0)$$

令 $P_0(x_0, y_0)$ (或 $P_0(x_0, y_0, z_0)$) 是 S (或 S^*) 中取定一点而 $P(x, y)$ (或 $P(x, y, z)$) 则是 S (或 S^*) 中任意一点，则有

$$ax + by + c = 0 \text{ 和 } ax_0 + by_0 + c = 0$$

$$(\text{或 } ax + by + cz + d = 0 \text{ 和 } ax_0 + by_0 + cz_0 + d = 0)$$

两者相减，即得

$$a(x - x_0) + b(y - y_0) = 0 \Leftrightarrow (a, b) \perp \overrightarrow{P_0P}$$

$$(\text{或 } a(x - x_0) + b(y - y_0) + c(z - z_0) = 0 \Leftrightarrow (a, b, c) \perp \overrightarrow{P_0P})$$

由此可见 S (或 S^*) 乃是平面上 (或空间中) 过 P_0 点而且和向量 $\mathbf{a} = (a, b)$ (或 $\mathbf{a} = (a, b, c)$) 垂直的那条直线 (或那个平面) 是也。□

[注]: 上述论证中所建立的直线 (或平面) 和二元 (或三元) 一次方程式之间的相互对应, 显示了方程式中变元的系数组 (亦即 (a, b) 或 (a, b, c)) 乃是直线 (或平面) 的一个法向量的分量。一条直线 (或一个平面) 在平面上 (或空间中) 的所有非零法向量之间只差一个 k -倍。由此可见, 一条直线 (或一个平面) 所对应的二元 (或三元) 一次方程式之间, 也只差一个 k -倍。

【定理 6.2】(点、线或点、面的距离公式): 设 $P_1(x_1, y_1)$ (或 $P_1(x_1, y_1, z_1)$) 是平面上 (或空间内) 给定一点, 而 ℓ (或 Π) 则是以

$$ax + by + c = 0$$

(或 $ax + by + cz + d = 0$)

为其方程的直线 (或平面), 则有下列距离公式:

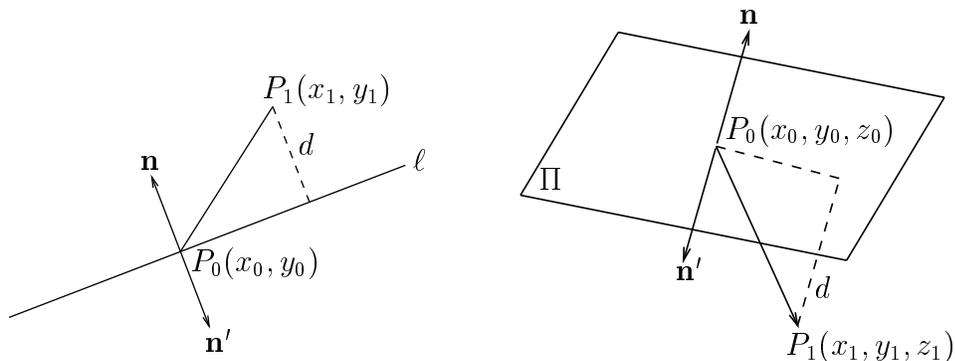
$$d(P_1, \ell) = \left| \frac{ax_1 + by_1 + c}{\sqrt{a^2 + b^2}} \right|$$

$$d(P_1, \Pi) = \left| \frac{ax_1 + by_1 + cz_1 + d}{\sqrt{a^2 + b^2 + c^2}} \right|$$

证明: 如 [图 6-4] 所示

$$\pm \frac{1}{\sqrt{a^2 + b^2}}(a, b) \quad (\text{或} \quad \pm \frac{1}{\sqrt{a^2 + b^2 + c^2}}(a, b, c))$$

乃是 ℓ (或 Π) 的两个单位长法向量 \mathbf{n} 和 \mathbf{n}' , 而 P_0 则是其上任取一点



[图 6-4]

则有

$$d(P_1, \ell) \text{ 或 } d(P_1, \Pi) = \mathbf{n} \cdot \overrightarrow{P_0P_1} \text{ 或 } \mathbf{n}' \cdot \overrightarrow{P_0P_1} \quad \square$$

【习题】：

- (1) 试写下过平面上（相异）两点 $P_1(x_1, y_1), P_2(x_2, y_2)$ 的直线的方程式。
- (2) 试写下过空间中（不共线）三点 $P_1(x_1, y_1, z_1), P_2(x_2, y_2, z_2), P_3(x_3, y_3, z_3)$ 的平面的方程式。
- (3) 试写下过平面上一点 $P_0(x_0, y_0)$ 并垂直于向量 $\mathbf{n} = (a, b)$ 的直线的方程式。
- (4) 试写下过空间中一点 $P_0(x_0, y_0, z_0)$ 并垂直于向量 $\mathbf{n} = (a, b, c)$ 的平面的方程式。
- (5) 试写下两条直线的两条交角平分线的方程式。
- (6) 试写下两个平面的两个两面角平分面的方程式。
- (7) 问在平面上两线（或空间中两面）平行的条件为何？

【定理 6.3】：平面上（或空间中）以 $P_0(x_0, y_0)$ （或 $P_0(x_0, y_0, z_0)$ ）为圆心（或球心）以 R 为半径的圆（或球）的方程式如下，即

$$(x - x_0)^2 + (y - y_0)^2 = R^2$$

$$(\text{或 } (x - x_0)^2 + (y - y_0)^2 + (z - z_0)^2 = R^2)$$

证明：是距离公式和圆与球的定义的直接推论。 □

【推论】：平面上（或空间中）一个圆（或球）所相应的方程式皆可写成下述形式，即

$$k[x^2 + y^2 + 2Dx + 2Ey + F] = 0$$

$$(\text{或 } k[x^2 + y^2 + z^2 + 2Dx + 2Ey + 2Fz + G] = 0)$$

其中 $D^2 + E^2 - F$ (或 $D^2 + E^2 + F^2 - G$) 乃是其半径之平方, 所以必须非负。反之, 任给满足 $D^2 + E^2 - F \geq 0$ (或 $D^2 + E^2 + F^2 - G \geq 0$) 的上述二元 (或三元) 二次方程式的解点子集乃是一个圆 (或球)。

证明: 令 $(x_0, y_0) = (-D, -E)$ (或 $(x_0, y_0, z_0) = (-D, -E, -F)$), 则易见上述二元 (或三元) 二次方程式可以改写成:

$$k[(x - x_0)^2 + (y - y_0)^2 - (D^2 + E^2 - F)] = 0$$

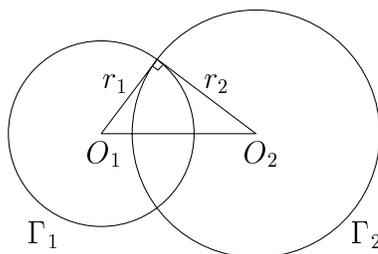
$$(\text{或 } k[((x - x_0)^2 + (y - y_0)^2 + (z - z_0)^2 - (D^2 + E^2 + F^2 - G)] = 0)$$

□

【引理】: 设 $\Gamma_i: x^2 + y^2 + 2D_i x + 2E_i y + F_i = 0, i = 1, 2$, 则 Γ_1, Γ_2 互相正交的充要条件是其系数满足下列条件:

$$2(D_1 D_2 + E_1 E_2) - (F_1 + F_2) = 0$$

证明: 由上面[推论]易见 $\Gamma_i, i = 1, 2$, 的圆心 O_i 的坐标为 $(-D_i, -E_i)$; 而其半径平方则为 $r_i^2 = D_i^2 + E_i^2 - F_i$ 。



[图 6-5]

再者, 由 [图 6-5] 中可见两圆正交的充要条件就是 $\overline{O_1 O_2}^2 = r_1^2 + r_2^2$, 即

$$(D_2 - D_1)^2 + (E_2 - E_1)^2 - (D_1^2 + E_1^2 - F_1) - (D_2^2 + E_2^2 - F_2) = 0$$

亦即

$$-2(D_1 D_2 + E_1 E_2) + (F_1 + F_2) = 0 \quad \square$$

【定理 6.4】: 设 Γ 和两个圆

$$\Gamma_i: x^2 + y^2 + 2D_i x + 2E_i y + F_i = 0, \quad i = 1, 2$$

都互相正交，则 Γ 和由 Γ_1, Γ_2 所生成的圆系，即

$$k_1(x^2 + y^2 + 2D_1x + 2E_1y + F_1) + k_2(x^2 + y^2 + 2D_2x + 2E_2y + F_2) = 0$$

其中每一个圆都是正交的。

证明：设 Γ 的方程式为

$$\Gamma: x^2 + y^2 + 2Dx + 2Ey + F = 0$$

则由所设的正交性即有

$$\begin{cases} 2(DD_1 + EE_1) - (F + F_1) = 0 & \dots\dots\dots (1) \\ 2(DD_2 + EE_2) - (F + F_2) = 0 & \dots\dots\dots (2) \end{cases}$$

再者，由 Γ_1, Γ_2 所生成的圆系中的任给一圆的方程式可以改写成标准式为

$$x^2 + y^2 + 2\left(\frac{k_1D_1 + k_2D_2}{k_1 + k_2}\right)x + 2\left(\frac{k_1E_1 + k_2E_2}{k_1 + k_2}\right)y + \left(\frac{k_1F_1 + k_2F_2}{k_1 + k_2}\right) = 0$$

由 (1), (2) 易得

$$2\left(D \cdot \frac{k_1D_1 + k_2D_2}{k_1 + k_2} + E \cdot \frac{k_1E_1 + k_2E_2}{k_1 + k_2}\right) - \left(F + \frac{k_1F_1 + k_2F_2}{k_1 + k_2}\right) = 0$$

所以它也是和 Γ 正交的。 □

【习题】：

- (1) 试写下过平面上（不共线）三点 $P_1(x_1, y_1), P_2(x_2, y_2), P_3(x_3, y_3)$ 的圆的方程式。
- (2) 试写下过空间中（不共面）四点 $P_1(x_1, y_1, z_1), P_2(x_2, y_2, z_2), P_3(x_3, y_3, z_3), P_4(x_4, y_4, z_4)$ 的球的方程式。
- (3) 试求分别以下述方程式表达的直线与圆相交于两点，相切和不相交的系数条件式：

$$l: ax + by + c = 0$$

$$\Gamma: x^2 + y^2 + 2Dx + 2Ey + F = 0, \quad (D^2 + E^2 - F > 0)$$

- (4) 试求分别以下述方程式表达的平面与球相交于一圆，相切于一点和不相交的系数条件式：

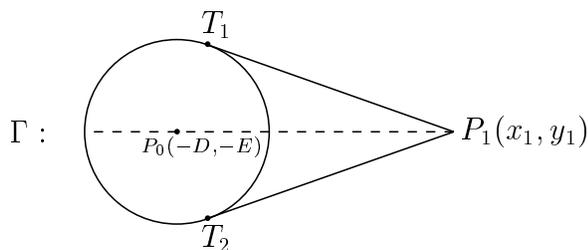
$$\Pi: ax + by + cz = 0 = 0$$

$$\Sigma: x^2 + y^2 + z^2 + 2Dx + 2Ey + 2Fz + G = 0, \quad (D^2 + E^2 + F^2 - G > 0)$$

- (5) 设 $P_1(x_1, y_1)$ 是位于圆 Γ 外的一点，而 Γ 的方程式是

$$\Gamma: x^2 + y^2 + 2Dx + 2Ey + F = 0$$

试求如 [图 6-6] 所示由 P_1 到 Γ 的切线长平方的公式（其值叫做 P_1 点对于 Γ 的幂）。



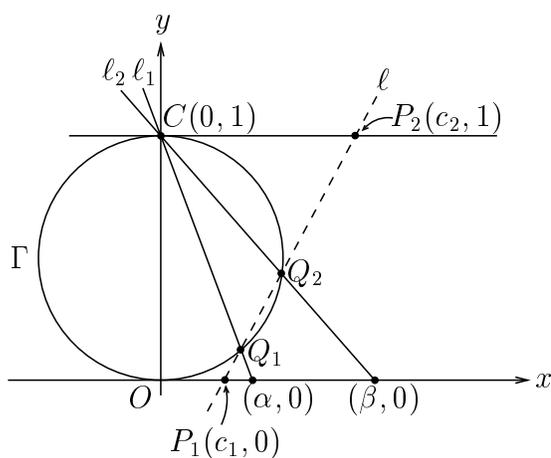
[图 6-6]

- (6) 等幂轴：设 Γ_1, Γ_2 是相异两圆，其方程分别为

$$\Gamma_i: x^2 + y^2 + 2D_i x + 2E_i y + F_i = 0, \quad i = 1, 2$$

和 Γ_1, Γ_2 的切线长相等的点集叫做它们的等幂轴，试求其方程式。

- (7) 如 [图 6-7] 所示， $\Gamma: x^2 + y^2 - y = 0$ 乃是一个以 \overline{OC} 为直径的圆。令 $A(\alpha, 0), B(\beta, 0)$ 为 x -轴上给定两点，连结 $l_1 = AC$ 和 $l_2 = BC$ 使得两线分别相交 Γ 于 Q_1, Q_2 两点。



[图 6-7]

- (i) 写下 l_1, l_2 的方程。令 g_λ 为下述组合所得的二次曲线

$$l_1 \cdot l_2 - \lambda \Gamma$$

注意 C, Q_1, Q_2 都在 g_λ 之上。试求一个适当的值 $\lambda = \lambda_0$ 使得 g_{λ_0} 可以被 $(y - 1)$ 所整除。

- (ii) 上述特殊的 λ_0 令 g_{λ_0} 分解成两个一次方程式的乘积，即

$$g_{\lambda_0} = (y - 1) \cdot \ell$$

其中 $\ell = P_1P_2$ 。求 P_1, P_2 的坐标（亦即 c_1, c_2 的值）。

- (iii) 试用上面所得者来描述一个几何作图方法来求解二次方程 $x^2 - px + q = 0$ 。

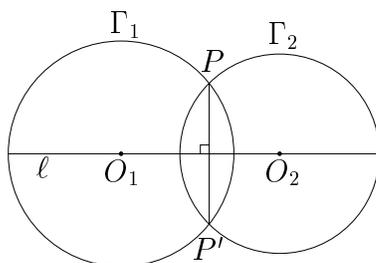
6.3 圆的反射对称：共轴圆系和共轭等轴圆系

所有欧氏平面的保长变换当然都是保圆保角的，但是除了它们之外是否还有不保长的保圆保角变换呢？再者，在欧氏平面的保长变换之中，对于一条给定直线的反射对称则是其中的至精至简者；所有其他保长变换都可以由它们组合而得，在保圆保角变换的范畴中是否也有同样的精简事物呢？上述对于直线的反射对称在保圆保角变换的范畴

中是否可以适当地推广成对于一个圆的反射对称呢？为此，让我们先来做一些温故知新的工作，改用圆与角的观点再来分析一下对于一条直线的反射对称的几何特征。再者，在以后的讨论中，我们将把直线看做是圆的特例，亦即半径是无穷大的圆。这样做可以使得往后的讨论和叙述变得更加整齐划一、更具统一性，例如原先的「不共线三点定一圆」在新观点之下就变成一律的「三点定一圆」。

【分析】：

(i) 如 [图 6-8] 所示， P, P' 是关于直线 l 成反射对称的任给一对点偶，亦即 l 是线段 $\overline{PP'}$ 的垂直平分线。由此可见，任何过 P, P' 点的圆，如 [图 6-8] 所示的 Γ_1, Γ_2 等，其圆心总是位于 l 之上，因此都是和 l 互相正交的。反之，若过相异两点 P, P' 的任何一个圆总是和 l 正交，则 P, P' 必定和 l 成反射对称。

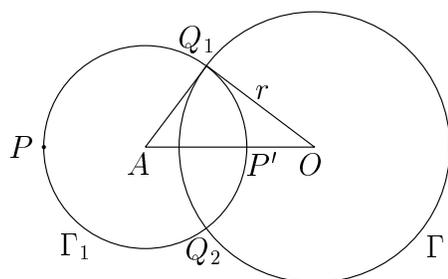


[图 6-8]

由此可问：对于一个定圆 Γ 和不在圆上的一点 P 是否会存在一个点 P' 使得任何过 P, P' 点的圆总是和圆 Γ 互相正交呢？

(ii) 假设果真存在有这样的一个点 P' ，则直线 PP' 乃是过 P, P' 的圆的一个特例，所以应该是和 Γ 互相垂直的，因此这个假设中的 P' 应该在 O, P 连线之上（我们且先把 P 点和 Γ 的圆心 O 点相重的情形除外）。再者，设 A 点是 $\overline{PP'}$ 的中点， Γ_1 是以 $\overline{PP'}$ 为直径的圆，亦即以 A 点为圆心， $\frac{1}{2}|\overline{PP'}|$ 为半径者也。由所设它和 Γ 正交于两点 Q_1, Q_2 （如 [图 6-9] 所示）。由此可见， AQ_i 是圆 Γ 的切线，而 OQ_i 则是 Γ_1 的切线。再由圆幂定理即得

$$\overline{OP} \cdot \overline{OP'} = r^2$$



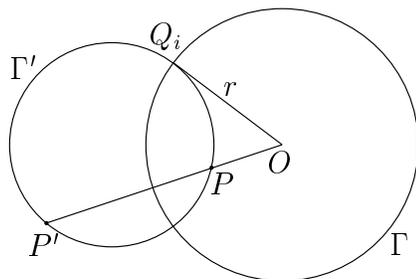
[图 6-9]

上述分析说明了：设 $P \neq O$ 而且不在圆 Γ 之上，假设存在有它的「对称点」 P' ，则它必须就是那个在射线 \overrightarrow{OP} 之上而且 $\overrightarrow{OP} \cdot \overrightarrow{OP'} = r^2$ 的点。下述引理则证明这样的 P' 点，的确是具有所要求的性质者。

【引理】：设 Γ 是一个以 O 点为圆心， r 为半径的圆， O, P, P' 三点共线而且 $\overrightarrow{OP} \cdot \overrightarrow{OP'} = r^2$ ，则任何过 P, P' 的圆都是和 Γ 互相正交的。

证明：设 Γ' 是任给一个过 P, P' 的圆，交 Γ 于 $Q_i, i = 1, 2$ 。由所设，即有

$$\overrightarrow{OP} \cdot \overrightarrow{OP'} = r^2 = \overrightarrow{OQ_i}^2$$



[图 6-10]

再者，由圆幂定理，

$$\overrightarrow{OQ_i}^2 = \overrightarrow{OP} \cdot \overrightarrow{OP'} = \text{由 } O \text{ 到 } \Gamma' \text{ 的切线长平方}$$

亦即 $\overrightarrow{OQ_i}$ 和切线等长，所以 Q_i 就是直线 OQ_i 和 Γ' 的唯一交点，即 OQ_i 本身就是一条由 O 到 Γ' 的切线，所以 Γ 和 Γ' 正交于 Q_i 。□

【定义】：对于给定圆 $\Gamma = \odot(O, r)$ 和任意点 $P \neq O$ ，那个在射线 \overrightarrow{OP} 之上而且满足 $\overrightarrow{OP} \cdot \overrightarrow{OP'} = r^2$ 的唯一的 P' 点叫做 P 点对于 Γ 的反射对称点。

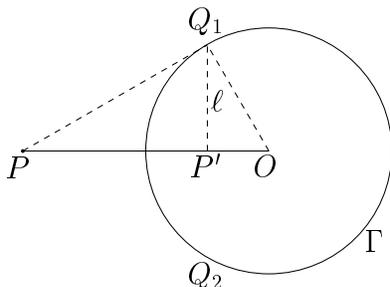
[注]: (i) 当 P 点位于圆 Γ 之上时, 其对称点就是其本身, 亦即 $P = P'$; 反之亦然, 即 $P = P' \Leftrightarrow P \in \Gamma$ 。

(ii) 在平面 Π 上, 圆心 O 点对于 Γ 的对称点暂无定义。若我们把平面 Π 添加一个无穷远点 ∞ , 而且把它定义为和圆心互相对称之点, 则一条直线可以当作是过 ∞ 点的圆。在这种看法之下, 我们可以证明上述对于 Γ 的反射对称在这个加点平面上是既保圆又保角的。

【例 1】: 对称点的作图法

设 P, P' 是一对对于 Γ 成反射对称之点。则 P, P' 和 Γ 之间有如图 6-11 所示的几何结构, 即有 $Q_1 \in \Gamma, OQ_1 \perp PQ_1, Q_1P' \perp OP$ 。由此可见求作对称点的下述作图法:

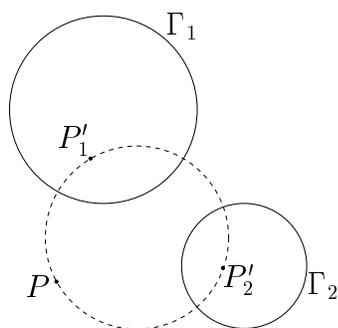
- (i) 若 P 点在圆外: 连结 OP , 以 \overline{OP} 为直径作圆 Γ_1 , 交 Γ 于 Q_1, Q_2 点, 连结 $\overline{Q_1Q_2}$ 交 \overline{OP} 于 P' 点, P' 点即为所求作的 P 点对于 Γ 的对称点。



[图 6-11]

- (ii) 若 P' 点在圆内: 连结 OP' , 过 P' 点作 OP' 垂线 l , 交 Γ 于 Q_1, Q_2 点, 过 Q_1 点作 OQ_1 的垂线, 交 OP' 于 P 点, 它即为所求作的对称点。

【例 2】: 设 Γ_1, Γ_2 是两个给定的圆, $P \notin \Gamma_1 \cup \Gamma_2$, 试作一圆 Γ 它过 P 点而且和 Γ_1, Γ_2 都正交。

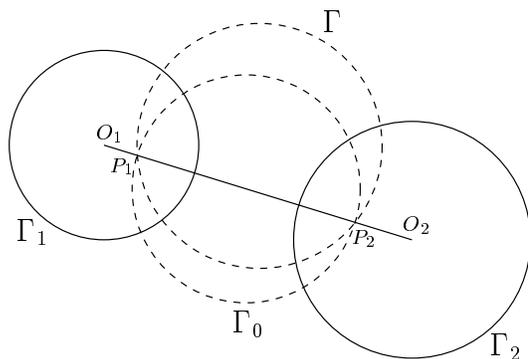


[图 6-12]

解：令 P'_1 和 P'_2 分别是 P 点对于 Γ_1, Γ_2 的反射对称点。则由 P, P'_1, P'_2 所定的圆 Γ 即为所求者。[试证上述 Γ 是过 P 点且和 Γ_1, Γ_2 正交的唯一圆。]

【例 3】：设 Γ_1, Γ_2 是两个相离或相切的圆，亦即 $\Gamma_1 \cap \Gamma_2 = \phi$ ，试问在所有和 Γ_1, Γ_2 都正交的圆系中是否有一个唯一的最小者，亦即其半径为极小者？若有，则试讨论其作图法。

解：先讨论相离的情形：



[图 6-13]

如 [图 6-13] 所示，设 Γ 是任给一个正交于 Γ_1, Γ_2 的圆，连结 $\overline{O_1O_2}$ ，它和 Γ 相交于 P_1, P_2 点，则 P_1, P_2 对于 Γ_1 和 Γ_2 都是互相对称的。即有

$$\overline{O_1P_1} \cdot \overline{O_1P_2} = r_1^2, \quad \overline{O_2P_1} \cdot \overline{O_2P_2} = r_2^2$$

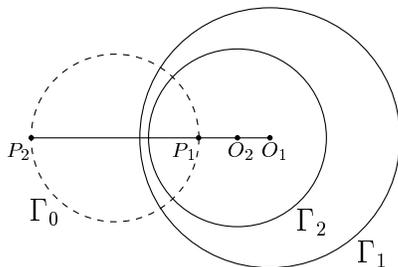
令 $\overline{O_1O_2} = \ell, \overline{O_1P_1} = x, \overline{O_1P_2} = y$ ，则上述条件就可以改用下述代数联立

方程式加以表达，亦即

$$\begin{cases} xy = r_1^2 \\ (\ell - x)(\ell - y) = r_2^2 \end{cases} \Leftrightarrow \begin{cases} xy = r_1^2 \\ x + y = \frac{\ell^2 + r_1^2 - r_2^2}{\ell} \end{cases}$$

由此可见，上述交点 P_1, P_2 是唯一确定的（和 Γ 的选取无关）而且可以用轨尺作图求得者。再者，和 Γ_1, Γ_2 正交的圆系也就是过 P_1, P_2 点的圆系，其中半径最小者显然就是那个以 $\overline{P_1 P_2}$ 为直径者。

相含情形基本上和上述相离的的情形相似，所不同者只是 P_1, P_2 和 O_1, O_2 之间的相对位置改为 [图 6-14] 所示的情况，其讨论则留作习题。



[图 6-14]

共轴圆系：

设 Γ_1, Γ_2 是给定两个相异的圆，其方程式分别为

$$\Gamma_i: x^2 + y^2 + 2D_i x + 2E_i y + F_i = 0, \quad i = 1, 2$$

则以

$$k_1(x^2 + y^2 + 2D_1 x + 2E_1 y + F_1) + k_2(x^2 + y^2 + 2D_2 x + 2E_2 y + F_2) = 0$$

为其方程式之圆乃是随著比值 $k_1:k_2$ 之不同而相异。它们组成一个圆系 (family of circles)，称之为由 Γ_1, Γ_2 所产生的共轴圆系 (co-axial family of circles)，易见它包含一条直线，亦即 $k_1:k_2 = -1$ 者，其方程式就是

$$2(D_1 - D_2)x + 2(E_1 - E_2)y + (F_1 - F_2) = 0$$

【习题】：

(1) 试验证上述共轴圆系中任何相异的两个圆的等幂轴就是直线

$$2(D_1 - D_2)x + 2(E_1 - E_2)y + (F_1 - F_2) = 0$$

[这也就是「共轴」这个名称的来由。]

(2) 设 Γ_1 和 Γ_2 相交于 P, Q 两点，试证上述共轴圆系也就是所有过 P, Q 点的圆所组成者（包括直线 PQ ，它其实是就是它们所共有之等幂轴）。

(3) 设 Γ_1 和 Γ_2 相切于 A 点，试证上述共轴圆系也就是共切于 A 点的圆所组成者，而它们的公切线也就是其所共有之等幂轴。

由此可见，共轴圆系是可以分成三类的。第一类是共交于两相异点者，其中不含有点圆；第二类是共切于一点者，其中仅含有一个点圆，即其共切之点；而第三类是各不相交者，其中含有两个点圆。上一节的[定理 6.4]证明了若 Γ 和共轴圆系中两个圆正交，则 Γ 和它们所产生的共轴圆系中每一个圆皆为正交。其实，我们还可以将结果再作推广如下：

【定理 6.5】：所有和两个定圆 Γ_1, Γ_2 都正交的圆构成一个共轴圆系。

证明：设 $P \notin \Gamma_1 \cup \Gamma_2$ ，令 P', P'' 分别是它对于 Γ_1, Γ_2 的反射对称点，则过 P, P', P'' 的圆就是和 Γ_1, Γ_2 都是正交的。由此可见，和 Γ_1, Γ_2 都正交的圆显然有无穷多个。设 Γ, Γ' 是其中两个相异的成员，由[定理 6.4]得知，由 Γ, Γ' 所生成的共轴圆系中的每一成员都是和 Γ_1, Γ_2 正交的，最后，我们还要说明不可能有上述共轴圆系之外的圆和 Γ_1, Γ_2 都正交。假若可能，设 Γ'' 是一个系外者，则所有其方程能表成下述形式者也都和 Γ_1, Γ_2 正交，即

$$\Gamma + \lambda\Gamma' + \mu\Gamma'', \quad \lambda, \mu \in \mathbb{R}$$

这是和三元一次联立方程式

$$\begin{cases} 2(DD_1 + EE_1) - (F + F_1) = 0 \\ 2(DD_2 + EE_2) - (F + F_2) = 0 \end{cases}$$

中解组 (D, E, F) 中只含有一个任意参数相矛盾。由此可见，对于一个共轴圆系，唯一存在另一个共轴圆系，每一个前者的成员和每一个後者的成员都是互相正交的。这样一对互相正交的共轴圆系称为是互相共轭的 (conjugate)。

在例题 §5.4 (4) 的圆幂定理的向量证法中，我们不但证明了 $\overline{PQ_1} \cdot \overline{PQ_2} = |\overline{OP}|^2 - R^2 = \text{切线长的平方}$ ，而且也证明了使得 $(PP'; Q_1Q_2)$ 成调和点列的 P' 恒位于 OP 的一条垂线之上，它和圆心 O 点的距离等于 R^2/\overline{OP} 。现在让我们改用坐标几何的观点，由圆 Γ 的方程式和 P 点的坐标 (x_0, y_0) 去求得上述直线（称之为圆 Γ 对于 P 点的极线）的方程式。

【定理 6.6】：设圆 Γ 的方程式为

$$x^2 + y^2 + 2Dx + 2Ey + F = 0$$

而 P 点的坐标为 (x_0, y_0) ，则圆 Γ 对于 P 点的极线方程式为

$$x_0x + y_0y + D(x_0 + x) + E(y_0 + y) + F = 0$$

证明：设 $X(x, y)$ 为上述直线 ℓ 上的任给一点，则由圆心 $O(-D, -E)$ 到 X 的向量 $\overrightarrow{OX} = (x + D, y + E)$ ，而 $\overrightarrow{OP} = (x_0 + D, y_0 + E)$ 。再者， ℓ 和 O, P 的几何关系可以用 $\overrightarrow{OP} \cdot \overrightarrow{OX} = R^2$ 表达之，即得其代数条件式：

$$\overrightarrow{OP} \cdot \overrightarrow{OX} = (x_0 + D)(x + D) + (y_0 + E)(y + E) = D^2 + E^2 - F$$

亦即

$$x_0x + y_0y + D(x_0 + x) + E(y_0 + y) + F = 0$$

同理亦有球幂定理和球对于 P 点的极面，其有关的讨论则留作习题。

第七章

球面几何和球面三角学

球面乃是空间中最完美匀称的曲面。两个半径相等的球面可以用一个平移把它们叠合起来，而两个半径不相等的球面所相差者就是放大或缩小这种相似变换，由此可见本质性的球面几何可以归纳到单位半径的球面来研讨。再者，在古典天文学的研讨中，观察星星的方向可以用单位球面上的一个点来标记它，而两个方向之间的角度（亦即方向差）则相应于单位球面上两点之间的球面距离 (spherical distance)。这也就是为什么古希腊天文学和几何学总是合为一体的，而且古希腊的几何学家对于球面三角学 (spherical trigonometry) 的投入程度要远远超过他们对于平面测量学的兴趣，因为「量天的学问」才是他们所致力去理解者；它的确比丈量土地、计量财产等更引人入胜，是不？

从现代的观点来看，球面几何乃是空间几何中蕴含在正交子群的部分，而向量几何则是空间几何中蕴含在平移子群的部分，而且两者又密切相关、相辅相成，例如向量运算都是正交协变的 (orthogonal covariant)，所以向量代数又是研讨球面几何的简明有力的利器。

7.1 单位球面的基本性质

设 O 为球面的心，而单位球面 $S^2(1)$ 则是空间 \mathcal{U} 中所有和 O 点的距离为 1 的点所成的点集，即：

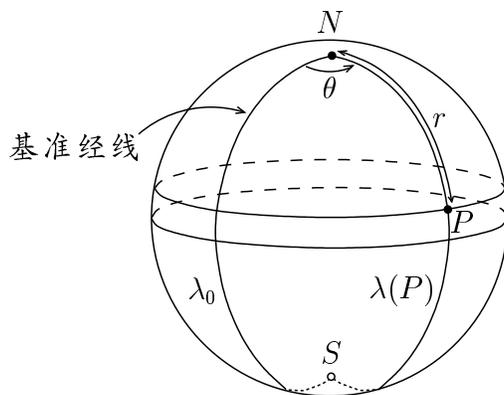
$$S^2(1) = \{A \in \mathcal{U}, \overline{OA} = 1\}$$

它是以 O 为其定点的正交子群的一个轨道 (orbit)。

- (i) 反射对称性：设 Π 是一个过球心 O 点的平面，则显然有 \mathfrak{R}_{Π} 保持 O 点不动。由 \mathfrak{R}_{Π} 的保长性可见它把和 O 点相距为 1 的点变换成和 O 点相距为 1 之点，所以 $\mathfrak{R}_{\Pi}(S^2(1)) = S^2(1)$ 。再者， \mathfrak{R}_{Π} 在 $S^2(1)$ 上的定点子集就是 $S^2(1) \cap \Pi$ 这一个大圆 (great circle)，我们将把 \mathfrak{R}_{Π} 限制在 $S^2(1)$ 上的变换叫做以大圆 $S^2(1) \cap \Pi$ 为定点子集的球面反射对称。
- (ii) 旋转对称性：设 l 是一条过球心 O 点的直线，它和球面 $S^2(1)$ 的交点是球面上的两个互为对顶的点 A, A' (一如南、北两极)；换言之，球面上两点 A, A' 互为对顶 (antipodal) 的条件是 $\overline{AA'}$ 以球心为其中点。在空间以 l 为轴的旋转之下，球心 $O \in l$ 是固定不动的；同理可见 $S^2(1)$ 也是它的一个不变子集，而它限制在球面上的变换乃是一个以对顶点 $\{A, A'\}$ 为其定点子集的球面旋转对称 (如日常地球所作者就是一个以南北极为其定点子集的旋转)。

球面极坐标：

设 $\{N, S\}$ 是单位球面上给定的两个互相对顶之点，在以 $\{N, S\}$ 为定点子集的球面旋转之下，每点的「纬度」保持不变，而其「经度」则随著转角而增加，如 [图 7-1] 所示。设 P 是球面上相异于两个极点者，令 $\lambda(P)$ 是过 P 点的那条经线 (longitude arc)， λ_0 是选定的基准经线。设 r 为 N 到 P 的球面距离，亦即 \widehat{NP} 这一段「经弧」的弧长， θ 是 λ_0 转到 $\lambda(P)$ 的 (有向) 转角，则称 (r, θ) 为 P 点对于以 N 为基点的球面极坐标 (spherical polar coordinates)。



[图 7-1]

若在空间选取正交坐标系，以球心为原点，以 \overrightarrow{ON} 为 z -轴的方向，以 \overrightarrow{OE} 为 x -轴的方向，其中 E 点乃是基准经线 λ_0 的中点，则有：

$$x = \sin r \cos \theta, \quad y = \sin r \sin \theta, \quad z = \cos r$$

[注]：由直接的微分计算可得

$$ds^2 = dx^2 + dy^2 + dz^2 = dr^2 + \sin^2 r d\theta^2$$

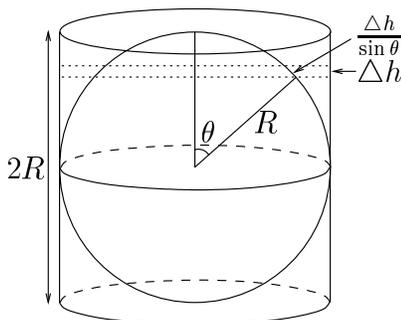
用上述弧长的微分式，不难证明经弧 \widehat{NP} 乃是球面上连结 N, P 两点的最短曲线（亦称测地线 (geodesics)）。

【阿基米德定理】：半径为 R 的球面面积等于

$$4\pi R^2$$

[注]：阿基米德 (Archimedes, 287–212 B.C.) 是公认的古希腊时代伟大的科学家和几何学家，他一生有很多卓越的贡献；而他最引以自豪者，首推上述定理及其简洁的证明，这也就是遵照他本人的遗嘱刻在他的墓碑上者。

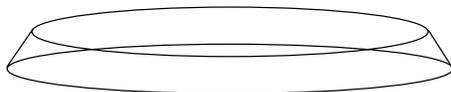
证明：其证明的要点在于论证一个半径为 R 的球面面积和一个高为 $2R$ ，半径为 R 的圆柱面面积相等。而在他的墓碑上所刻划的，就是如 [图 7-2] 所示把两者放在相切同高的位置。



[图 7-2]

设想用一系列和柱面正交的平行平面，把两个面都细分成很窄很窄的一圈圈。设相邻两个平行面之间的距离是 Δh ，则柱面上的窄条（或圈）的面积等于 $2\pi R \Delta h$ ，而在球面上的相应窄圈，虽然其宽度和长度

会随著 θ 而改变，但在 Δh 非常、非常小的时候，它可以看成如 [图 7-3] 所示的圆台之侧面：



[图 7-3]

其中环长度是 $2\pi R \sin \theta$ ，亦即其环长的平均值是 $2\pi R \sin \theta$ ，而侧面的宽度则为 $\frac{\Delta h}{\sin \theta}$ ，所以其面积的高度近似值也是 $2\pi R \Delta h$ （亦即可能的误差肯定在 Δh^2 这种量级）。由此他就用 Eudoxus 所创的逼近原理证明了两者的面积必然相等，而後者的面积显然等于高为 $2R$ ，长为 $2\pi R$ 的长方形面积，亦即 $4\pi R^2$ 。□

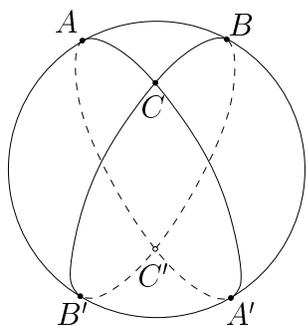
球面三角形面积公式：

设 A, B, C 是球面上任取三点但不含对顶者，令 $\widehat{AB}, \widehat{BC}, \widehat{CA}$ 为连结于点与点之间的测地线，称之为球面三角形 $\triangle ABC$ 的三个边。我们将采用和平面三角学中相同的符号体系，以 A, B, C 表示 $\triangle ABC$ 在三个顶点的内角，及以 a, b, c 表示 $\triangle ABC$ 的各角对边边长。在平面几何中，一个三角形的三个内角和恒等于一个平角，这是逻辑等价于平行公理的基本事实，也是平面的平直性的一种基本表达；在球面三角形的情形下，三内角之和则恒大于一个平角，而下述 [定理 7.1] 证明在单位球面上的球面三角形，其内角和与 π 的差额（称之为「角盈」）其实恰好等于其面积。

【定理 7.1】：在单位球面上，一个球面三角形 $\triangle ABC$ 的面积就是

$$\Delta ABC = A + B + C - \pi$$

证明：如 [图 7-1] 所示，由二个夹角为 θ 的经线所围成的球面部分，其面积显然和 θ 成正比（这是球面对以 N, S 为定点的旋转对称性的直接推论）。再者，当 $\theta = 2\pi$ 时，其面积等于 4π （阿基米德定理）！所以上述以 θ 为夹角者（称之为 spherical lune）的面积等于 2θ 。



[图 7-4]

如 [图 7-4] 所示，令 A', B', C' 分别是 A, B, C 的对顶者。用上述 spherical lune 的面积公式即得：

$$\begin{aligned} \triangle ABC + \triangle A'BC &= 2A, \\ \triangle ABC + \triangle AB'C &= 2B, \\ \triangle ABC + \triangle ABC' &= 2C, \\ \triangle ABC' &= \triangle A'B'C, \quad (\text{互为对顶者, 当然等面积}) \\ \triangle ABC + \triangle A'BC + \triangle AB'C + \triangle A'B'C &= 2\pi, \quad (\text{半球面积}) \end{aligned}$$

由此可得

$$2\triangle ABC + 2\pi = 2A + 2B + 2C$$

亦即

$$\triangle ABC = A + B + C - \pi \quad \square$$

[注]：上述具有基本重要性的球面三角形面积公式其实就是阿基米德球面面积公式的局部化和精细化。

球面三角形的叠合条件及等腰三角形定理：

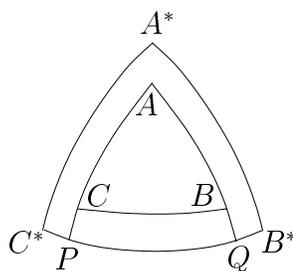
设 A, B 是球面上任给两点。在空间中和 A, B 等距的点集是直线段 \overline{AB} 的垂直平分面 Π_1 ，它当然包含球心 O ，所以和 A, B 等距的球面上之点乃是 $\Pi_1 \cap S^2(1)$ 这个大圆，而球面对于这个大圆的反射对称将 A, B 互换。用上述球面上的反射对称即可推导出：

- (i) S.A.S. 也是球面三角形的叠合条件；
- (ii) 球面等腰三角形的两底角相等；反之，两底角相等的球面三角形亦必为等腰。

再者，由上述两点还可以同样地推导出球面三角形也具有其他如 S.S.S. 和 A.S.A. 等叠合条件。在此值得一提的是 A.A.A. 也是球面三角形的一个叠合条件，我们可以用球面三角形中所特有的对偶关系来推导它也是一个叠合条件。设 A, A' 互为对顶，则和 A, A' 等距的球面上的点集就是和 A, A' 的距离是 $\pi/2$ 的那个大圆，将以 Γ_A 记之。设 $\triangle ABC$ 是一个任给球面三角形，在下述三对对顶点偶（即 $\Gamma_B \cap \Gamma_C, \Gamma_C \cap \Gamma_A, \Gamma_A \cap \Gamma_B$ ）之中，分别取其靠近 A, B, C 者，以 A^*, B^*, C^* 记之，则称 $\triangle A^*B^*C^*$ 为 $\triangle ABC$ 的对偶球面三角形（ $\triangle ABC$ 也是 $\triangle A^*B^*C^*$ 的对偶球面三角形）。

【引理 7.1】：令 a, b, c 和 a^*, b^*, c^* 分别是 $\triangle ABC$ 和 $\triangle A^*B^*C^*$ 的各角对边边长，则有：

$$\begin{aligned} a^* &= \pi - A, & b^* &= \pi - B, & c^* &= \pi - C \\ a &= \pi - A^*, & b &= \pi - B^*, & c &= \pi - C^* \end{aligned}$$



[图 7-5]

证明：我们只需要证明其中之一，其余各式皆可同理类推。由 [图 7-5] 所示，在大圆 Γ_A 上 $\widehat{B^*P} = \widehat{C^*Q} = \pi/2$, $\widehat{PQ} = A$ ，故有

$$a^* = \widehat{B^*P} + \widehat{C^*Q} - \widehat{PQ} = \pi - A \quad \square$$

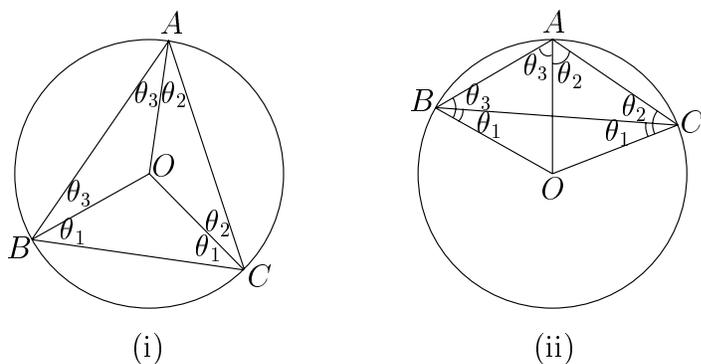
【推论】：A.A.A. 也是一种球面三角形的叠合条件。

证明：设 $\triangle A_1B_1C_1$ 和 $\triangle A_2B_2C_2$ 的三角内角对应相等，由 [引理 7.1] 得知它们的对偶球面三角形 $\triangle A_1^*B_1^*C_1^*$ 和 $\triangle A_2^*B_2^*C_2^*$ 的三个边长对应等长，所以是全等的，因此当然有三个对应内角相等。再用 [引理 7.1]，即得 $\triangle A_1B_1C_1$ 和 $\triangle A_2B_2C_2$ 满足 S.S.S. 全等条件。 \square

【引理 7.2】：设 $\triangle ABC$ 和 $\triangle A'BC$ 的顶点共圆而且 A, A' 同在 \widehat{BC} 的一侧，则

$$\angle ABC + \angle ACB - \angle A = \angle A'BC + \angle A'CB - \angle A'$$

再者，上述之逆命题也成立。



[图 7-6]

证明：如 [图 7-6(i)] 所示， $\triangle OAB, \triangle OBC, \triangle OCA$ 皆为等腰，所以其底角相等，设其分别是 $\theta_1, \theta_2, \theta_3$ 。则有

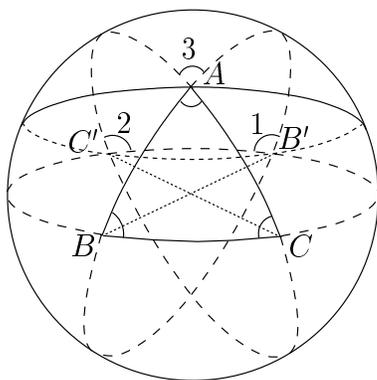
$$\angle ABC + \angle ACB - \angle A = 2\theta_1$$

同理亦有

$$\angle A'BC + \angle A'CB - \angle A' = 2\theta_1$$

[图 7-6(ii)] 的情况和逆命题的证明留作习题。 □

【定理 7.2】(Lexell)：设球面三角形 $\triangle A_1BC$ 和 $\triangle A_2BC$ 具有相等的定向面积，而 B', C' 分别是 B, C 的对顶点，则 B', C', A_1, A_2 四点共圆。



[图 7-7]

证明：如 [图 7-7] 所示：

$$\angle 1 = \pi - B, \angle 2 = \pi - C, \angle 3 = A$$

所以

$$\angle 1 + \angle 2 - \angle 3 = \pi - \triangle ABC$$

分别取 $A = A_1$ 和 A_2 ，再对 $\triangle A_1 B' C'$ 和 $\triangle A_2 B' C'$ 运用 [引理 7.2] 的逆命题，即得 B', C', A_1, A_2 共圆。□

【习题】：

- (1) 设 P_1, P_2 的球面极坐标分别是 $(r_1, 0)$ 和 $(r_2, 0)$ ，而 $\gamma(t) = (f(t), g(t))$ 是一条一阶可微曲线， $a \leq t \leq b$ ， $\gamma(a) = (r_1, 0)$ ， $\gamma(b) = (r_2, 0)$ 。试证其长度至少等于 $|r_1 - r_2|$ 。
- (2) 若 $\triangle ABC$ 是一个半径为 R 的球面三角形，试问 $A + B + C - \pi$ 和其面积之间的关系是什么？并试证你的主张。
- (3) 设 $\triangle A_1 B_1 C_1$ 和 $\triangle A_2 B_2 C_2$ 是满足 S.A.S. 条件的两个球面三角形，例如 $A_1 = A_2, b_1 = b_2, c_1 = c_2$ 。试构造一系列球面上的反射对称，它们的组合恰好把 $\triangle A_1 B_1 C_1$ 变换到 $\triangle A_2 B_2 C_2$ 。
- (4) 试用球面的反射对称性证明等腰三角形的底角相等，而顶角平分线垂直平分底边。

- (5) 试用上述 (3), (4) 所证得者, 证明 S.S.S. 也是球面三角形的一种全等条件。
- (6) 设 O 为一个球面的心, A 为球面上任给一点, Π 为过 A 点而且和 \overline{OA} 垂直的平面。试证 Π 和球面仅仅交于 $\{A\}$ 点。
- (7) 设 Γ 是极坐标下「 $r = \text{常数}$ 」所构成的纬圆。试求 Γ 上任一点 P 的切平面和直线 ON 的交点 V (亦即确定 \overline{OV} 的长度)。

7.2 球面三角学

球面三角学研讨球面三角形的各种各样几何量如边长、角度、面积、外接圆和内切圆的半径等等的相互关系。远古希腊时代, 球面三角学即已倍加重视。Menelous 所著的“Sphaerica”和 Ptolemy 所著的“Almagest”总结了当年在球面三角学上的研究成果和它们在天文学上的应用。大体上, 他们已经充分理解了直角球面三角形的各种几何量之间的相互关系; 然後一直到十八世纪, 球面三角学的研究才又得以蓬勃开展。

在本节的讨论中, 将以 $\mathbf{a}, \mathbf{b}, \mathbf{c}$ 等等表示单位球面上给定点 A, B, C 等等的位置向量, 亦即 $\mathbf{a} = \overrightarrow{OA}, \mathbf{b} = \overrightarrow{OB}, \mathbf{c} = \overrightarrow{OC}$ 等等, 它们当然都是单位长的向量。由此可见, 从向量几何的观点来看, 球面几何其实也就是单位长向量的几何。

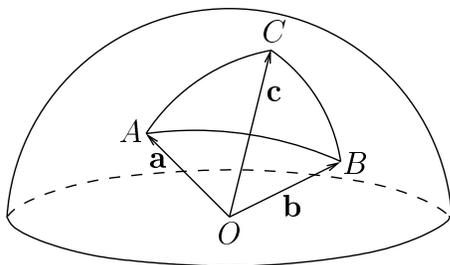
由向量运算的几何内含, 即有 (参看 [图 7-8]) :

- (i) $\mathbf{b} \cdot \mathbf{c} = \cos a, \mathbf{c} \cdot \mathbf{a} = \cos b, \mathbf{a} \cdot \mathbf{b} = \cos c$;
- (ii) $\angle(\mathbf{b}, \mathbf{c}), \angle(\mathbf{c}, \mathbf{a}), \angle(\mathbf{a}, \mathbf{b})$ 的面积, 亦即 $|\mathbf{b} \times \mathbf{c}|, |\mathbf{c} \times \mathbf{a}|, |\mathbf{a} \times \mathbf{b}|$, 分别等于 $\sin a, \sin b, \sin c$;
- (iii) 球面三角形 $\triangle ABC$ 的三个内角 A, B, C 等于 $\angle(\mathbf{a}, \mathbf{b})$ 和 $\angle(\mathbf{a}, \mathbf{c})$, $\angle(\mathbf{b}, \mathbf{a})$ 和 $\angle(\mathbf{b}, \mathbf{c})$, $\angle(\mathbf{c}, \mathbf{a})$ 和 $\angle(\mathbf{c}, \mathbf{b})$ 之间的两面角 ;
- (iv) 设 $V(\mathbf{a}, \mathbf{b}, \mathbf{c}) = (\mathbf{a} \times \mathbf{b}) \cdot \mathbf{c} > 0$, 以後将以 D 表示之。由行列式的乘

法公式即有：

$$D^2 = \begin{vmatrix} \mathbf{a} \cdot \mathbf{a} & \mathbf{b} \cdot \mathbf{a} & \mathbf{c} \cdot \mathbf{a} \\ \mathbf{a} \cdot \mathbf{b} & \mathbf{b} \cdot \mathbf{b} & \mathbf{c} \cdot \mathbf{b} \\ \mathbf{a} \cdot \mathbf{c} & \mathbf{b} \cdot \mathbf{c} & \mathbf{c} \cdot \mathbf{c} \end{vmatrix} = \begin{vmatrix} 1 & \cos c & \cos b \\ \cos c & 1 & \cos a \\ \cos b & \cos a & 1 \end{vmatrix}$$

$$= 1 + 2 \cos a \cos b \cos c - \cos^2 a - \cos^2 b - \cos^2 c$$



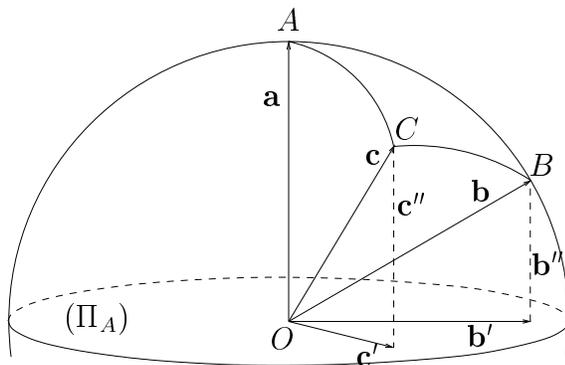
[图 7-8]

【定理 7.3】（球面三角正弦定律）：

$$\frac{\sin A}{\sin a} = \frac{\sin B}{\sin b} = \frac{\sin C}{\sin c} = \frac{D}{\sin a \sin b \sin c}$$

证明：令 Π_A 为过球心 O 点而和 \overrightarrow{OA} 垂直的平面， \mathbf{b}' 和 \mathbf{c}' 是 \mathbf{b} , \mathbf{c} 在 Π_A 上的垂直投影，亦即：

$$\mathbf{b} = \mathbf{b}' + \mathbf{b}'', \quad \mathbf{c} = \mathbf{c}' + \mathbf{c}''$$

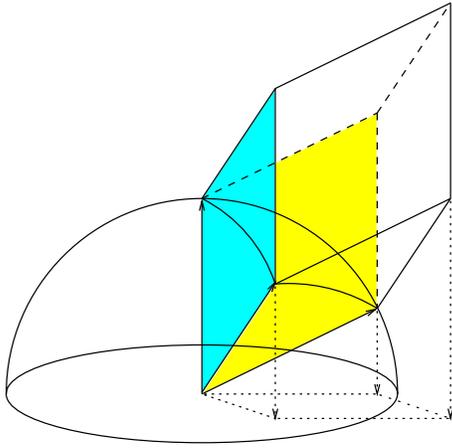


[图 7-9]

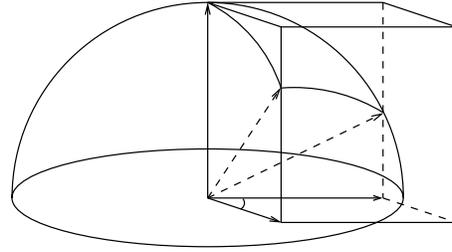
其中 \mathbf{b}' , \mathbf{c}' 和 \mathbf{a} 垂直而 \mathbf{b}'' 和 \mathbf{c}'' 则为 \mathbf{a} 的倍积，所以由内积和 \times -积的分配律，得：

$$\begin{aligned} D &= \mathbf{a} \cdot (\mathbf{b} \times \mathbf{c}) = \mathbf{a} \cdot (\mathbf{b}' \times \mathbf{c}') + \text{为 } 0 \text{ 之三项} \\ &= |\mathbf{b}' \times \mathbf{c}'| = |\mathbf{b}'| \cdot |\mathbf{c}'| \cdot \sin A \\ &= \sin c \sin b \sin A \end{aligned}$$

上述所作的垂直投影其实是把由 \mathbf{a} , \mathbf{b} , \mathbf{c} 所张的平行六面体沿 \mathbf{a} 的方向滑动，最後得出由 \mathbf{a} , \mathbf{b}' , \mathbf{c}' 所张的长方体，如下图所示：



[图 7-10(i)]



[图 7-10(ii)]

因为体积是斜移不变的，由此亦可以看到

$$D = \mathbf{a} \cdot (\mathbf{b} \times \mathbf{c}) = \mathbf{a} \cdot (\mathbf{b}' \times \mathbf{c}') = \sin c \sin b \sin A$$

由此易见

$$\frac{\sin A}{\sin a} = \frac{\sin B}{\sin b} = \frac{\sin C}{\sin c} = \frac{D}{\sin a \sin b \sin c} \quad \square$$

【定理 7.4】（球面三角余弦定律）：

$$\sin b \sin c \cos A = \cos a - \cos b \cos c$$

$$\sin c \sin a \cos B = \cos b - \cos c \cos a$$

$$\sin a \sin b \cos C = \cos c - \cos a \cos b$$

证明：由面积的勾股定理，即有：

$$\begin{aligned} (\mathbf{c} \times \mathbf{a}) \cdot (\mathbf{b} \times \mathbf{a}) &= \begin{vmatrix} \mathbf{c} \cdot \mathbf{b} & \mathbf{a} \cdot \mathbf{b} \\ \mathbf{c} \cdot \mathbf{a} & \mathbf{a} \cdot \mathbf{a} \end{vmatrix} = \begin{vmatrix} \cos a & \cos c \\ \cos b & 1 \end{vmatrix} \\ &= \cos a - \cos b \cos c \end{aligned}$$

再者，由内积 \times -积的几何意义，以及 A 等于 $\angle(\mathbf{c}, \mathbf{a})$ 和 $\angle(\mathbf{b}, \mathbf{a})$ 之间的两面角，即有：

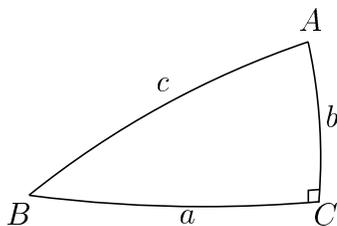
$$\begin{aligned} (\mathbf{c} \times \mathbf{a}) \cdot (\mathbf{b} \times \mathbf{a}) &= |\mathbf{c} \times \mathbf{a}| \cdot |\mathbf{b} \times \mathbf{a}| \cos A \\ &= \sin b \sin c \cos A \end{aligned} \quad \square$$

球面三角余弦定律的另一证法：

$$\begin{aligned} \sin c \sin b \cos A &= \mathbf{b}' \cdot \mathbf{c}' \\ &= (\mathbf{b} - (\mathbf{b} \cdot \mathbf{a})\mathbf{a}) \cdot (\mathbf{c} - (\mathbf{c} \cdot \mathbf{a})\mathbf{a}) \\ &= \mathbf{b} \cdot \mathbf{c} - (\mathbf{b} \cdot \mathbf{a})(\mathbf{c} \cdot \mathbf{a}) - (\mathbf{c} \cdot \mathbf{a})(\mathbf{b} \cdot \mathbf{a}) + (\mathbf{b} \cdot \mathbf{a})(\mathbf{c} \cdot \mathbf{a}) \cdot 1 \\ &= \cos a - \cos b \cos c \end{aligned} \quad \square$$

【推论 1】：在 $C = \frac{\pi}{2}$ （亦即直角球面三角形）时，则有：

$$\begin{aligned} \text{(i)} \quad \cos c &= \cos a \cos b, \quad \sin A = \frac{\sin a}{\sin c}, \quad \sin B = \frac{\sin b}{\sin c}; \\ \text{(ii)} \quad \cos A &= \frac{\tan b}{\tan c}, \quad \cos B = \frac{\tan a}{\tan c}, \quad \tan A = \frac{\tan a}{\sin b}, \quad \tan B = \frac{\tan b}{\sin a}. \end{aligned}$$



[图 7-11]

证明：由所设 $C = \frac{\pi}{2}$ 即有 $\cos C = 0, \sin C = 1$ 。所以 (i)-式乃是正、余弦定律的直接结论。再者，

$$\begin{aligned}\sin b \sin c \cos A &= \cos a - \cos b \cos c \\ &= \cos a - \cos^2 b \cos a = \sin^2 b \cos a\end{aligned}$$

所以

$$\cos A = \frac{\sin b \cos a}{\sin c} = \frac{\tan b \cos b \cos a}{\sin c} = \frac{\tan b \cos c}{\sin c} = \frac{\tan b}{\tan c}$$

[其他三式的证明留作习题。]

半角公式：

在平面三角学中，我们有下述易算好用的半角公式，即令 $s = \frac{1}{2}(a + b + c)$ ，则有：

$$\begin{aligned}\cos \frac{A}{2} &= \sqrt{\frac{s(s-a)}{bc}}, \text{ etc.} \\ \sin \frac{A}{2} &= \sqrt{\frac{(s-b)(s-c)}{bc}}, \text{ etc.} \\ \tan \frac{A}{2} &= \sqrt{\frac{(s-b)(s-c)}{s(s-a)}}, \text{ etc.} \\ \text{面积 } \Delta &= \sqrt{s(s-a)(s-b)(s-c)} \\ \text{内切圆半径 } r &= \frac{\Delta}{s} = \sqrt{\frac{(s-a)(s-b)(s-c)}{s}}\end{aligned}$$

在球面三角学中，也有类似的半角公式，即：

【推论 2】（球面三角半角公式）：

$$\cos \frac{A}{2} = \sqrt{\frac{\sin s \sin(s-a)}{\sin b \sin c}}, \text{ etc.}$$

$$\sin \frac{A}{2} = \sqrt{\frac{\sin(s-b) \sin(s-c)}{\sin b \sin c}}, \text{ etc.}$$

$$\tan \frac{A}{2} = \sqrt{\frac{\sin(s-b) \sin(s-c)}{\sin s \sin(s-a)}}, \text{ etc.}$$

$$\tan r = \sqrt{\frac{\sin(s-a) \sin(s-b) \sin(s-c)}{\sin s}}$$

证明：以 $\cos A = 2 \cos^2 \frac{A}{2} - 1$ （或 $1 - 2 \sin^2 \frac{A}{2}$ ）代入余弦定律，即得：

$$\begin{aligned} 2 \sin b \sin c \cos^2 \frac{A}{2} &= \cos a - \cos b \cos c + \sin b \sin c \\ &= \cos a - \cos(b+c) \\ &= 2 \sin \frac{1}{2}(a+b+c) \sin \frac{1}{2}(-a+b+c) \end{aligned}$$

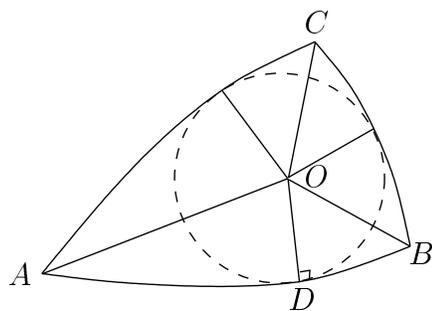
或

$$\begin{aligned} 2 \sin b \sin c \sin^2 \frac{A}{2} &= -\cos a + \cos b \cos c + \sin b \sin c \\ &= \cos(b-c) - \cos a \\ &= 2 \sin \frac{1}{2}(a+b-c) \sin \frac{1}{2}(a-b+c) \end{aligned}$$

这也就证明了 (i) 和 (ii)，而 (iii) 则是 (i), (ii) 的直接推论。兹证 (iv)-式如下：

如 [图 7-12] 所示， $\triangle ADO$ 是直角球面三角形， $\widehat{AD} = s-a$ ， $\angle OAD = \frac{A}{2}$ ，所以

$$\frac{\tan r}{\sin(s-a)} = \tan \frac{A}{2} = \sqrt{\frac{\sin(s-b) \sin(s-c)}{\sin s \sin(s-a)}} \Rightarrow \text{(iv)} \quad \square$$



[图 7-12]

阿基米德定理以及它的局部化——球面三角形面积公式：

$$\Delta = A + B + C - \pi$$

是球面几何中至关重要的基本定理。从纯几何的观点，上述面积公式已经是十分简洁完美的了；但是从向量代数的不变量理论来看，我们还需要把三角形面积和 $\{\mathbf{a}, \mathbf{b}, \mathbf{c}\}$ 的基本正交不变量，亦即 $\{\mathbf{a} \cdot \mathbf{b}, \mathbf{b} \cdot \mathbf{c}, \mathbf{c} \cdot \mathbf{a}\} = \{\cos c, \cos a, \cos b\}$ 之间整理出一个简洁、整体的关系式。当然，我们可以用球面三角余弦定律，即

$$\sin a \sin b \cos C = \cos c - \cos a \cos b \quad \text{等等}$$

得出

$$C = \cos^{-1} \left\{ \frac{\cos c - \cos a \cos b}{\sin a \sin b} \right\} \quad \text{等等}$$

所以这个用向量内积的面积公式当然就可以写成：

$$\begin{aligned} \Delta = & \cos^{-1} \left\{ \frac{\cos a - \cos b \cos c}{\sin b \sin c} \right\} + \cos^{-1} \left\{ \frac{\cos b - \cos c \cos a}{\sin c \sin a} \right\} \\ & + \cos^{-1} \left\{ \frac{\cos c - \cos a \cos b}{\sin a \sin b} \right\} - \pi \end{aligned}$$

但是这样一个繁复的表式显然不好用，因此有必要去探讨上述球面三角形面积的内积表达式背后的精简形式。这种精益求精的所得就是：

【定理 7.5】： $\tan \frac{\Delta}{2} = \frac{D}{u}$, $D = \mathbf{a} \cdot (\mathbf{b} \times \mathbf{c}) > 0$, $u = 1 + \mathbf{a} \cdot \mathbf{b} + \mathbf{b} \cdot \mathbf{c} + \mathbf{c} \cdot \mathbf{a}$.

证明：由球面三角正弦、余弦定律（亦即[定理 7.3]、[定理 7.4]）即有

$$\sin A = \frac{D}{\sin b \sin c}, \quad \cos A = \frac{\cos a - \cos b \cos c}{\sin b \sin c}$$

等等直接代换和代数计算可得：

$$\begin{aligned} \sin \Delta &= \sin(A + B + C - \pi) = -\sin(A + B + C) \\ &= \frac{D}{\prod \sin^2 a} \{D^2 + \sum \cos a \cos^2 b - \sum \cos a \cos b - (u - 1) \cdot \prod \cos a\} \end{aligned}$$

上式之分子为

$$\prod(1 - \cos^2 a) = \prod(1 + \cos a) \prod(1 - \cos a)$$

而一个令人惊喜的事实是括号内 $\{D^2 + \dots\}$ 的代数表式可以简化成 $u \cdot \prod(1 - \cos a)$ 。所以即得：

$$\sin \Delta = \frac{Du}{\prod(1 + \cos a)} = \frac{2Du}{D^2 + u^2}$$

同样的代数计算可得

$$\cos \Delta = \frac{u^2 - D^2}{u^2 + D^2}$$

所以

$$\tan \frac{\Delta}{2} = \frac{\sin \Delta}{1 + \cos \Delta} = \frac{D}{u} \quad \square$$

[注]：在直角球面三角形，即 $C = \frac{\pi}{2}$ 时，尚有下列特殊公式，即：

$$\begin{aligned} \cos \Delta &= \cos \left(A + B - \frac{\pi}{2} \right) = \sin(A + B) \\ &= \frac{1}{\sin c \tan c} \{ \sin a \tan a + \sin b \tan b \} = \frac{\cos a + \cos b}{1 + \cos c} \end{aligned}$$

【推论 1】：若将 $\triangle ABC$ 的两边 a, b 固定而让第三边 c 变动，令

$$c_1 = \cos a, \quad c_2 = \cos b, \quad x = 1 + \cos c, \quad \varphi = \angle C$$

则有

$$\begin{aligned}\frac{d\Delta}{dx} &= \frac{c_1 + c_2 - x}{xD} \\ \frac{d\Delta}{d\varphi} &= \frac{x - c_1 - c_2}{x}\end{aligned}$$

证明：由上所设，

$$\begin{aligned}u &= c_1 + c_2 + x \\ D^2 &= -(c_1 + c_2)^2 + 2(c_1c_2 + 1)x - x^2\end{aligned}$$

将 $\tan \frac{\Delta}{2} = \frac{D}{u}$ 对于 x 求微分，即

$$\begin{aligned}\frac{d\Delta}{dx} &= \frac{d}{dx} \left(2 \tan^{-1} \frac{D}{u} \right) \\ &= \frac{uD' - D}{\frac{1}{2}(u^2 + D^2)} = \frac{uDD' - D^2}{(1 + c_1)(1 + c_2)xD}\end{aligned}$$

在这里，有趣的是分子也含有 $(1 + c_1)(1 + c_2)$ 因式。约分後即得

$$\frac{d\Delta}{dx} = \frac{c_1 + c_2 - x}{xD}$$

再者，将下述余弦定律

$$\sin a \sin b \cos \varphi = x - 1 - c_1c_2$$

对于 x 微分，即有

$$-\sin a \sin b \sin \varphi \frac{d\varphi}{dx} = -D \frac{d\varphi}{dx} = 1$$

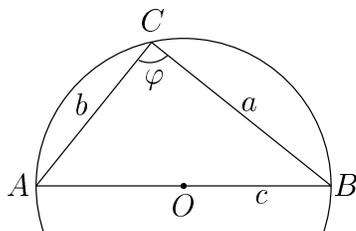
所以

$$\frac{d\Delta}{d\varphi} = \frac{d\Delta}{dx} \frac{dx}{d\varphi} = \frac{x - c_1 - c_2}{x} \quad \square$$

[注]：当 φ 从 0 变到 π ， $\Delta = \Delta(\varphi)$ 的变化有下述三种情形，即：

- (i) 若 $a + b < \pi$, 则 $c_1 + c_2 > 0$, 而其对边 c 则从 $|a - b|$ 变到 $a + b$, 函数值 $\Delta(\varphi)$ 由 0 增加到其在 $x = c_1 + c_2$ 时的唯一极大值, 然后再递减到 0。

[注]: $x = c_1 + c_2$, 即 $1 + \cos c = \cos a + \cos b$ 的几何意义乃是 $\triangle ABC$ 的外接圆圆心位于 \widehat{AB} 之上, 如 [图 7-13] 所示。其证明在讨论球面四边形时便会详细说明。



[图 7-13]

- (ii) 若 $a + b > \pi$, 则 $c_1 + c_2 < 0$; 其第三边 $c(\varphi)$ 则由 $|a - b|$ 增加到 $2\pi - (a + b)$ 。因为 $\frac{d\Delta}{d\varphi}$ 一直是正的, $\Delta(\varphi)$ 由 0 递增到 2π , 亦即 $\triangle ABC$ 以半个球面为其上限。
- (iii) 若 $a + b = \pi$, 则 $c_1 + c_2 = 0$; 其第三边 $c(\varphi)$ 则由 $|a - b|$ 增加到 π , 而 $\Delta(\varphi)$ 则由 0 递增到 π , 亦即 $\triangle ABC$ 以四分之一球面为其上限。

【推论 2】: 设球面三角形 $\triangle ABC$ 的三边边长中 c 和 $a + b$ 保持不变, 则 $\triangle ABC$ 的面积是 $|a - b|$ 的递减函数。

证明: 令 $A = \cos \frac{a+b}{2}$, $K = 1 + \cos c$, $t = \cos \frac{a-b}{2}$, 则有:

$$\begin{aligned}
 u &= \cos a + \cos b + \cos c + 1 = 2At + K \\
 D^2 + u^2 &= 2(1 + \cos a)(1 + \cos b)(1 + \cos c) = 2K(A + t)^2 \\
 D^2 &= (K - 2A^2)(2t^2 - K) \\
 \frac{d\Delta}{dt} &= \frac{D'u - Du'}{\frac{1}{2}(u^2 + D^2)} = \frac{DD'u - D^2u'}{\frac{1}{2}(u^2 + D^2)D} = \frac{2(K - 2A^2)}{D(A + t)} > 0
 \end{aligned}$$

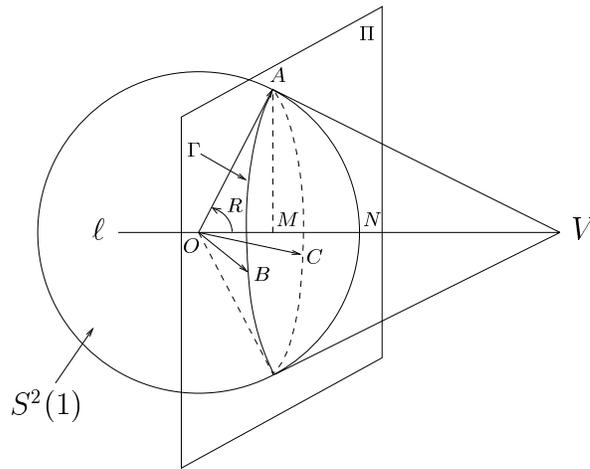
上式中 $\frac{d\Delta}{dt} > 0$ 是因为

$$\begin{aligned} K - 2A^2 &= 1 + \cos c - 2 \cos^2 \frac{a+b}{2} = \cos c - \cos(a+b) \\ &= 2 \sin \frac{1}{2}(a+b+c) \sin \frac{1}{2}(a+b-c) > 0, \\ A+t &= \cos \frac{a+b}{2} + \cos \frac{a-b}{2} = 2 \cos \frac{a}{2} \cos \frac{b}{2} > 0 \end{aligned}$$

由此可见，面积 $\Delta(t)$ 是 t 的递增函数，亦即是 $|a-b|$ 的递减函数。□

球面三角形的外接圆：

一个球面三角形 $\triangle ABC$ 的外接圆是由 A, B, C 三点所定的平面 Π 和球面的交集，而它也是平面三角形 $\triangle'ABC$ 在平面 Π 中的外接圆，如 [图 7-14] 所示：



[图 7-14]

令 M 为在 Π 上 Γ 的圆心， N 为在球面上 Γ 的圆心， $l = OM$ 。易证 OM 和 Π 垂直，而且 $N \in l$ 。[证明留作习题]

由此可见，平面 Π 和圆 Γ 是在以 l 为轴的旋转之下的不变子集 (invariant subsets)，设球面在 A 点的切面 T_A 交 l 于 V 点，由上述旋转不变性可见球面在 Γ 上任给一点 P 的切面 T_P 也过 V 点。

【定理 7.6】：

$$\overrightarrow{OV} = \frac{1}{D}(\mathbf{a} \times \mathbf{b} + \mathbf{b} \times \mathbf{c} + \mathbf{c} \times \mathbf{a})$$

证明：设 $T_A \cap \ell = \{V\}$ 。由所设 ℓ 过球心 O 而且和 Π 正交，所以 $S^2(1)$ 和 Π 在以 ℓ 为轴的旋转下都是不变子集，因此 $\Gamma = S^2(1) \cap \Pi$ 也当然是不变子集。设 P 为 Γ 上任给一点，则有一个 ℓ -轴旋转 ρ ，它把 A 点移动到 P 点，易见 $\rho(T_A) = T_P$ ，而 ρ 是保持 ℓ 每一点固定不动的，所以 $T_P \cap \ell = \rho(T_A \cap \ell) = \rho(V) = V$ 。

再者， $T_A \cap T_B \cap T_C$ 只含有一个点，所以 V 其实就是 $\{T_A, T_B, T_C\}$ 的共交点。由此可见：

$$\overrightarrow{OV} \cdot \mathbf{a} = 1, \overrightarrow{OV} \cdot \mathbf{b} = 1, \overrightarrow{OV} \cdot \mathbf{c} = 1$$

而且 \overrightarrow{OV} 是唯一能够满足上述三个等式的向量。由直接验算可知：

$$\begin{aligned} \frac{1}{D}(\mathbf{a} \times \mathbf{b} + \mathbf{b} \times \mathbf{c} + \mathbf{c} \times \mathbf{a}) \cdot \mathbf{a} &= \frac{1}{D}(\mathbf{b} \times \mathbf{c}) \cdot \mathbf{a} = 1 \\ \frac{1}{D}(\mathbf{a} \times \mathbf{b} + \mathbf{b} \times \mathbf{c} + \mathbf{c} \times \mathbf{a}) \cdot \mathbf{b} &= \frac{1}{D}(\mathbf{c} \times \mathbf{a}) \cdot \mathbf{b} = 1 \\ \frac{1}{D}(\mathbf{a} \times \mathbf{b} + \mathbf{b} \times \mathbf{c} + \mathbf{c} \times \mathbf{a}) \cdot \mathbf{c} &= \frac{1}{D}(\mathbf{a} \times \mathbf{b}) \cdot \mathbf{c} = 1 \end{aligned}$$

所以它必然是 \overrightarrow{OV} 。 □

【推论 1】：单位球面上四点 $\{A, B, C, D\}$ 共圆的充要条件是它们的位置向量 $\{\mathbf{a}, \mathbf{b}, \mathbf{c}, \mathbf{d}\}$ 满足

$$(\mathbf{a} \times \mathbf{b}) \cdot \mathbf{c} + (\mathbf{a} \times \mathbf{c}) \cdot \mathbf{d} - (\mathbf{a} \times \mathbf{b}) \cdot \mathbf{d} - (\mathbf{b} \times \mathbf{c}) \cdot \mathbf{d} = 0$$

证明：由上述讨论， D 位于 $\triangle ABC$ 的外接圆的充要条件乃是 T_D 也过 V 的，即 $\overrightarrow{OV} \cdot \mathbf{d} = 1$ ，亦即：

$$(\mathbf{a} \times \mathbf{b}) \cdot \mathbf{c} - (\mathbf{a} \times \mathbf{b} + \mathbf{b} \times \mathbf{c} + \mathbf{c} \times \mathbf{a}) \cdot \mathbf{d} = 0 \quad \square$$

【推论 2】：设 R 为球面上 $\triangle ABC$ 的外接圆半径，亦即 [图 7-14] 中之 \widehat{AN} 的弧长，则有：

$$\tan^2 R = \frac{2}{D^2}(1 - \cos a)(1 - \cos b)(1 - \cos c)$$

证明：如 [图 7-14] 所示，

$$\begin{aligned}\tan^2 R &= \overline{OV} \cdot \overline{OV} - 1 \\ &= \frac{1}{D^2} \{4(1+w) - u^2 - D^2\} \\ &= \frac{2}{D^2} (1 - \cos a)(1 - \cos b)(1 - \cos c)\end{aligned}$$

其中 $w = \cos a \cos b + \cos b \cos c + \cos c \cos a$ 。 □

球面四边形 (spherical quadrilaterals)：

一如在平面的情形，两个三角形若其三边对应等长则为全等，但是两个四边形若其四边对应等长仍然是可能不全等的；即使在凸的情况，还是需要加上其一个对应对角线等长，才能完全确定其叠合性。设 $\{\mathbf{a}, \mathbf{b}, \mathbf{c}, \mathbf{d}\}$ 是一个球面四边形的顶点的位置向量，它们共有六个交叉内积，亦即是其四个边长和两个对角线长的余弦，它们之间具有一个函数关系，由此可以从给定其中五个之值去确定其第六个之值。我们可以用平行六面体定向体积（亦即行列式）的乘法公式来求得这种函数关系式，即：

$$[V(\mathbf{a}, \mathbf{b}, \mathbf{c})]^2 \cdot [V(\mathbf{a}, \mathbf{c}, \mathbf{d})]^2 = [V(\mathbf{a}, \mathbf{b}, \mathbf{c}) \cdot V(\mathbf{a}, \mathbf{c}, \mathbf{d})]^2$$

两边分别用行列式乘法公式计算，即得：

$$\begin{vmatrix} \mathbf{a} \cdot \mathbf{a} & \mathbf{b} \cdot \mathbf{a} & \mathbf{c} \cdot \mathbf{a} \\ \mathbf{a} \cdot \mathbf{b} & \mathbf{b} \cdot \mathbf{b} & \mathbf{c} \cdot \mathbf{b} \\ \mathbf{a} \cdot \mathbf{c} & \mathbf{b} \cdot \mathbf{c} & \mathbf{c} \cdot \mathbf{c} \end{vmatrix} \cdot \begin{vmatrix} \mathbf{a} \cdot \mathbf{a} & \mathbf{c} \cdot \mathbf{a} & \mathbf{d} \cdot \mathbf{a} \\ \mathbf{a} \cdot \mathbf{c} & \mathbf{c} \cdot \mathbf{c} & \mathbf{d} \cdot \mathbf{c} \\ \mathbf{a} \cdot \mathbf{d} & \mathbf{c} \cdot \mathbf{d} & \mathbf{d} \cdot \mathbf{d} \end{vmatrix} = \begin{vmatrix} \mathbf{a} \cdot \mathbf{a} & \mathbf{b} \cdot \mathbf{a} & \mathbf{c} \cdot \mathbf{a} \\ \mathbf{a} \cdot \mathbf{c} & \mathbf{b} \cdot \mathbf{c} & \mathbf{c} \cdot \mathbf{c} \\ \mathbf{a} \cdot \mathbf{d} & \mathbf{b} \cdot \mathbf{d} & \mathbf{c} \cdot \mathbf{d} \end{vmatrix}^2$$

若以

$$\mathbf{a} \cdot \mathbf{a} = \mathbf{b} \cdot \mathbf{b} = \mathbf{c} \cdot \mathbf{c} = \mathbf{d} \cdot \mathbf{d} = 1$$

$$\mathbf{a} \cdot \mathbf{b} = \cos l_1, \mathbf{b} \cdot \mathbf{c} = \cos l_2, \mathbf{c} \cdot \mathbf{d} = \cos l_3, \mathbf{d} \cdot \mathbf{a} = \cos l_4$$

$$\mathbf{a} \cdot \mathbf{c} = \cos d_1, \mathbf{b} \cdot \mathbf{d} = \cos d_2$$

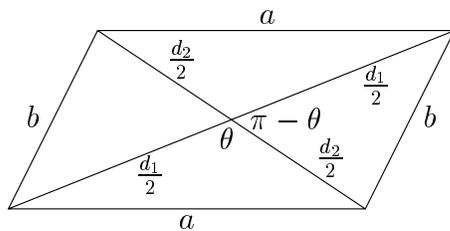
代入即为：

$$\begin{vmatrix} 1 & \cos l_1 & \cos d_1 \\ \cos l_1 & 1 & \cos l_2 \\ \cos d_1 & \cos l_2 & 1 \end{vmatrix} \cdot \begin{vmatrix} 1 & \cos d_1 & \cos l_4 \\ \cos d_1 & 1 & \cos l_3 \\ \cos l_4 & \cos l_3 & 1 \end{vmatrix} = \begin{vmatrix} 1 & \cos l_1 & \cos d_1 \\ \cos d_1 & \cos l_2 & 1 \\ \cos l_4 & \cos d_2 & \cos l_3 \end{vmatrix}^2$$

上述是一般情形的普遍关系式，若在 $l_1 = l_3 (= a)$, $l_2 = l_4 (= b)$ ，亦即二对对边各别等长的特殊情形，则上述关系式可以大为简化，其结果为

$$(1 + \cos d_1)(1 + \cos d_2) = (\cos a + \cos b)^2$$

[注]：在平面几何中，两对对边各别等长的四边形就是平行四边形，因此这种球面四边形也可以看成平行四边形在球面几何中的推广（虽然「平行」这个名称实在已经名不符实的了！）。其实，这种球面四边形除了没有两对对边互相平行这个性质之外，也具有平行四边形的其他性质，例如其对角线互相平分、两对对角各别相等 [证明留作习题]。再者，我们也可以用对角线互相平分来直接验证上述关系式：



[图 7-15]

如 [图 7-15] 所示，即有余弦定律公式：

$$\begin{aligned} \sin \frac{d_1}{2} \sin \frac{d_2}{2} \cos \theta &= \cos a - \cos \frac{d_1}{2} \cos \frac{d_2}{2} \\ \sin \frac{d_1}{2} \sin \frac{d_2}{2} \cos(\pi - \theta) &= \cos b - \cos \frac{d_1}{2} \cos \frac{d_2}{2} \end{aligned}$$

两式相加，即得

$$0 = (\cos a + \cos b) - 2 \cos \frac{d_1}{2} \cos \frac{d_2}{2}$$

亦即

$$(\cos a + \cos b)^2 = (1 + \cos d_1)(1 + \cos d_2)$$

再者，上述四边形顶点共圆的充要条件是 $d_1 = d_2 = d$ ，亦即

$$\cos a + \cos b = 1 + \cos d$$

由此可见，[定理 7.5]的[推论 1]中 $\frac{d\Delta}{d\varphi} = 0$ 的条件式：

$$c_1 + c_2 - x = \cos a + \cos b - (1 + \cos c) = 0$$

其几何意义是 $\triangle ABC$ 的外接圆圆心位于 \widehat{AB} 之上。[在平面几何中， \overline{AB} 是外接圆的直径和 $\angle C = \pi/2$ 是等价的，显然也是 a, b 边固定时 \triangle 面积的极大。在此值得一提者是：二边边长给定的三角形，在第三边为其外接圆的直径时面积极大在非欧几何中亦成立。]

球面四边形的面积：

在欧氏平面几何中，矩形面积公式扮演著重要的角色。再者，矩形乃是平行四边形的特例，亦即对角线等长的平行四边形（或者说，是四顶点共圆的平行四边形）。所以，在球面几何中，矩形的自然推广就是上述球面「平行四边形」在四顶点共圆（或对角线等长）的情形。由[定理 7.5]的[推论 1]可见，给定两对对边边长（设为 a, b ）的「平行四边形」之中，其面积以矩形为极大。

【定理 7.7】：（球面矩形面积公式）在单位球面上以 a, b 为其两对对边长的矩形，其面积公式为

$$\square(a, b) = 4 \tan^{-1} \left\{ \frac{(1 - \cos a)(1 - \cos b)}{2(\cos a + \cos b)} \right\}^{\frac{1}{2}}$$

证明：令其对角线长为 c （两者等长）。则有

$$1 + \cos c = \cos a + \cos b$$

$$u = 2(\cos a + \cos b)$$

$$D^2 = 2(\cos a + \cos b)(1 - \cos a)(1 - \cos b)$$

亦即

$$\square(a, b) = 4 \tan^{-1} \frac{D}{u} = 4 \tan^{-1} \left\{ \frac{(1 - \cos a)(1 - \cos b)}{2(\cos a + \cos b)} \right\}^{\frac{1}{2}} \quad \square$$

其实，顶点共圆时面积为极大并不只限于「平行四边形」，详见下述[定理 7.8]：

【定理 7.8】：在四边边长给定为 $\{\ell_i, 1 \leq i \leq 4\}$ 而且总长少于 2π 的四边形中，其面积以顶点共圆时为唯一的极大值，而且远离共圆的变形使得面积愈来愈小。

证明: 令 $c_i = \cos l_i$, $x = 1 + \cos d$, 其中 $d = \widehat{AC}$ (在 $\square ABCD$ 非凸时, 我们选定 \widehat{AC} 是位于内部的那条对角线), 以 Δ_1, Δ_2 分别表示 $\triangle ABC$ 和 $\triangle ACD$ 的面积, 则有:

$$A(\square ABCD) = \Delta_1 + \Delta_2 = 2 \tan^{-1} \frac{D_1}{u_1} + 2 \tan^{-1} \frac{D_2}{u_2}$$

其中

$$\begin{aligned} u_1 &= c_1 + c_2 + x, & u_2 &= c_3 + c_4 + x \\ D_1^2 &= -(c_1 + c_2)^2 + 2(c_1 c_2 + 1)x - x^2 \\ D_2^2 &= -(c_3 + c_4)^2 + 2(c_3 c_4 + 1)x - x^2 \end{aligned}$$

[选取顶点之顺序使得 $D_1 = \mathbf{a} \cdot (\mathbf{b} \times \mathbf{c})$ 和 $D_2 = \mathbf{a} \cdot (\mathbf{c} \times \mathbf{d})$ 都是正的。]

由[定理 7.5] [推论 1]的微分计算, 即有

$$A'(x) = \Delta_1'(x) + \Delta_2'(x) = \frac{c_1 + c_2 - x}{xD_1} + \frac{c_3 + c_4 - x}{xD_2}$$

现在让我们用向量代数来分析一下上式右侧的几何意义: 令 $l_1 = T_A \cap T_C$, $P_0 = T_A \cap T_C \cap (OAC)$, 则有 $\overrightarrow{OP_0}$ 是 \mathbf{a}, \mathbf{c} 的线性组合而且 $\overrightarrow{OP_0} \cdot \mathbf{a} = \overrightarrow{OP_0} \cdot \mathbf{c} = 1$, 所以

$$\overrightarrow{OP_0} = \frac{1}{1 + \mathbf{a} \cdot \mathbf{c}} (\mathbf{a} + \mathbf{c})$$

再者, 设 P 是 l_1 上任给一点, 则由 $l \perp (OAC)$ 可见 $\overrightarrow{P_0P}$ 乃是 $\mathbf{a} \times \mathbf{c}$ 的一个倍积, 所以

$$\overrightarrow{OP} = \frac{1}{1 + \mathbf{a} \cdot \mathbf{c}} (\mathbf{a} + \mathbf{c}) + k(\mathbf{a} \times \mathbf{c})$$

令 V_1, V_2 分别是 T_B, T_D 和 l_1 的交点, 亦即

$$V_1 = T_A \cap T_C \cap T_B, \quad V_2 = T_A \cap T_C \cap T_D$$

则有

$$\begin{aligned} \overrightarrow{OV_1} &= \frac{1}{1 + \mathbf{a} \cdot \mathbf{c}} (\mathbf{a} + \mathbf{c}) + k_1(\mathbf{a} \times \mathbf{c}), & \overrightarrow{OV_1} \cdot \mathbf{b} &= 1 \\ \overrightarrow{OV_2} &= \frac{1}{1 + \mathbf{a} \cdot \mathbf{c}} (\mathbf{a} + \mathbf{c}) + k_2(\mathbf{a} \times \mathbf{c}), & \overrightarrow{OV_2} \cdot \mathbf{d} &= 1 \end{aligned}$$

直接计算可得

$$\begin{aligned}\frac{1}{x}(c_1 + c_2) - k_1 D_1 = 1 &\Rightarrow k_1 = \frac{c_1 + c_2 - x}{x D_1} \\ \frac{1}{x}(c_3 + c_4) + k_2 D_2 = 1 &\Rightarrow k_2 = \frac{x - c_3 - c_4}{x D_2}\end{aligned}$$

由此可见

$$\overrightarrow{V_2 V_1} = (k_1 - k_2) \mathbf{a} \times \mathbf{c} = \left(\frac{c_1 + c_2 - x}{x D_1} + \frac{c_3 + c_4 - x}{x D_2} \right) \mathbf{a} \times \mathbf{c} = A'(x) \mathbf{a} \times \mathbf{c}$$

所以在 $V_1 = V_2$ 时 $A(x)$ 为极大，而 $V_1 = V_2$ 也就是顶点共圆。 \square

【例题 1】：设 x_0 为四边形面积函数的极值点，即

$$A'(x_0) = \frac{c_1 + c_2 - x_0}{x_0 D_1(x_0)} + \frac{c_3 + c_4 - x_0}{x_0 D_2(x_0)} = 0$$

若 $c_1 + c_2 = c_3 + c_4$ ，则显然 $x_0 = c_1 + c_2 = c_3 + c_4$ 是所求之解，再者在这种情形， \widehat{AC} 乃是外接圆的直径而 $\triangle_1(x_0)$ 和 $\triangle_2(x_0)$ 各别都是极大。

若 $c_1 + c_2 \neq c_3 + c_4$ （可以设 $c_1 + c_2 > c_3 + c_4$ ），由条件式

$$\frac{c_1 + c_2 - x_0}{x_0 D_1} + \frac{c_3 + c_4 - x_0}{x_0 D_2} = 0$$

即

$$\frac{c_1 + c_2 - x_0}{x_0 D_1} = \frac{x_0 - c_3 - c_4}{x_0 D_2} \quad (> 0)$$

所以有

$$\frac{c_1 + c_2 - x_0}{x_0 - c_3 - c_4} = \frac{D_1(x_0)}{D_2(x_0)}$$

令其为 λ ，则有

$$\begin{aligned}\lambda^2 &= \frac{D_1(x_0)^2}{D_2(x_0)^2} = \frac{(c_1 + c_2 - x_0)^2}{(c_3 + c_4 - x_0)^2} \\ &= \frac{D_1(x_0)^2 + (c_1 + c_2 - x_0)^2}{D_2(x_0)^2 + (c_3 + c_4 - x_0)^2} \\ &= \frac{(1 - c_1)(1 - c_2)}{(1 - c_3)(1 - c_4)} \\ \Rightarrow \lambda &= \sqrt{\frac{(1 - c_1)(1 - c_2)}{(1 - c_3)(1 - c_4)}}\end{aligned}$$

再由

$$\frac{c_1 + c_2 - x_0}{x_0 - c_3 - c_4} = \lambda = \sqrt{\frac{(1 - c_1)(1 - c_2)}{(1 - c_3)(1 - c_4)}}$$

即求得

$$x_0 = \frac{(c_1 - c_2)\sqrt{(1 - c_3)(1 - c_4)} + (c_3 - c_4)\sqrt{(1 - c_1)(1 - c_2)}}{\sqrt{(1 - c_3)(1 - c_4)} + \sqrt{(1 - c_1)(1 - c_2)}}$$

注意: 在 $c_1 + c_2 = c_3 + c_4$ 时, 上述公式即为

$$x_0 = c_1 + c_2 = c_3 + c_4$$

所以上述由 $\{c_i = \cos l_i, 1 \leq i \leq 4\}$ 表达 $x_0 = 1 + \cos d$ 的公式是普遍成立的!

【例题 2】: 设四边形的四个边长依序取定为 $\{l_i, 1 \leq i \leq 4\}$, 令 φ 为 $\angle B$ 的角度, 则其面积为 φ 的函数, 亦即 $A(x), x = x(\varphi)$ 。由余弦定律, 即

$$\sin l_1 \sin l_2 \cos \varphi = (x - 1) - \cos l_1 \cos l_2$$

对 x 求微分, 即得

$$\frac{d\varphi}{dx} = \frac{1}{D_1} = \frac{1}{(\mathbf{a} \times \mathbf{c}) \cdot \mathbf{b}}$$

再者, 原先由[定理 8]的证明已得

$$\overrightarrow{V_2 V_1} = \frac{dA}{dx} \mathbf{a} \times \mathbf{c}$$

所以

$$\begin{aligned} \overrightarrow{V_2 V_1} \cdot \mathbf{b} &= \frac{dA}{dx} (\mathbf{a} \times \mathbf{c}) \cdot \mathbf{b} \\ &= \frac{dA}{dx} \cdot \frac{dx}{d\varphi} = \frac{dA}{d\varphi} \end{aligned}$$

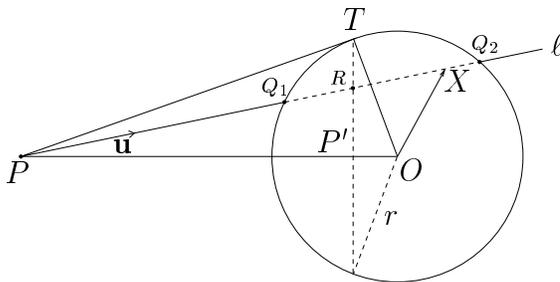
【习题】:

- (1) 试问球面上一个保长变换的定点子集有那些可能性? 并举例说明你所说的那种可能性是的确可能的。

(2) 设 $\triangle ABC$ 是一个直角球面三角形, $\angle C = \frac{\pi}{2}$ 。试证

$$\cos B = \frac{\tan a}{\tan c}, \quad \tan A = \frac{\tan a}{\sin b}, \quad \tan B = \frac{\tan b}{\sin a}$$

(3) 设 $S^2(r)$ 是一个以 O 为球心, 半径为 r 的球面。 P 是球外一个给定点 (如 [图 7-16] 所示) :



[图 7-16]

设 ℓ 为过 P 点而且交 $S^2(r)$ 于 Q_1, Q_2 的直线。试证恒有

$$\overline{PQ_1} \cdot \overline{PQ_2} = |\overline{OP}|^2 - r^2$$

[提示: 设 \mathbf{u} 是直线 ℓ 上的单位长向量, $\overline{PQ_1} = k_1$, $\overline{PQ_2} = k_2$ 而 X 是 ℓ 上的动点, 则有 $\overline{OX} = \overline{OP} + \overline{PX} = \overline{OP} + k\mathbf{u}$, 其中 k 是 \overline{PX} 的有向长度, 而 $X \in S^2(r)$ 的条件式则是 $\overline{OX} \cdot \overline{OX} = r^2$ 。]

- (4) 设 PT_i 是和 $S^2(r)$ 相切于 T_i 的那条切线, $i = 1, 2$ 。试证 $\overline{PT_1}$ 和 $\overline{PT_2}$ 等长, 并描述所有过 P 点和 $S^2(r)$ 相切的切点所组成的点集。
- (5) 令 P' 是位于直线 OP 之上而且 $\overline{OP'} \cdot \overline{OP} = r^2$ 的点, Π 是和 OP 正交于 P' 点的平面, 令 $R = \ell \cap \Pi$ 。试证 $(PR; Q_1Q_2)$ 成调和点列, 亦即

$$\frac{\overline{Q_1P}}{\overline{PQ_2}} \cdot \frac{\overline{Q_2R}}{\overline{RQ_1}} = -1$$

第八章

圆锥截线的故事

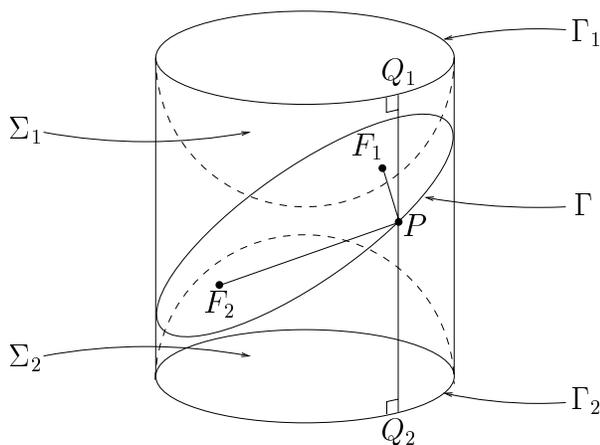
在现代的中学数学课程中，通常是在初等解析几何中学到圆锥截线，亦即椭圆、双曲线和抛物线。圆锥截线的发现和研究起始于古希腊。Euclid, Archimedes, Apollonius, Pappus 等几何学大师都热衷于圆锥截线的研究，而且都有专著论述其几何性质，其中以 Apollonius 所著的八册《圆锥截线论》集其大成，可以说是古希腊几何学一个登峰造极的精湛之作。当时对于这种既简朴又完美的曲线的研究，乃是纯粹从几何学的观点，研讨和圆密切相关的这种曲线；它们的几何乃是圆的几何的自然推广，在当年这是一种纯理念的探索，并不寄望也无从预期它们会真的在大自然的基本结构中扮演著重要的角色。此事一直到十六、十七世纪之交，Kepler 行星运行三定律的发现才知道行星绕太阳运行的轨道，乃是一种以太阳为其一焦点的椭圆。Kepler 三定律乃是近代科学开天辟地的重大突破，它不但开创了天文学的新纪元，而且也是牛顿万有引力定律的根源所在。由此可见，圆锥截线不单单是几何学家所爱好的精简事物，它们也是大自然的基本规律中所自然选用的精要之一。

在本章将以圆锥截线的来龙去脉为中心课题，简明扼要地叙述这一段由古希腊几何学到牛顿天体力学引人入胜、发人深思的篇章。

8.1 圆柱截线和圆锥截线

常见的一段竹杆，大体上是一个圆柱。它的正切截线是一个圆，但是其斜切截线则不再是圆的，这也许就是「椭圆」的一种自然出处。

圆的几何特性乃是它有一个圆心，和其上各点等距；自然会问这种由斜截圆柱所得的「椭圆」是否也具有类似的几何特性呢？古希腊几何学家在上述问题的探讨中获得令人鼓舞的简洁答案，亦即一个椭圆具有两个焦点 F_1, F_2 使得其上任给一点到两者的距离之和为一定长（其实，这也就是通常在初等解析几何中椭圆的定义），我们用 [图 8-1] 来解说当年对于这种圆柱截线的基本特性的证法。设 Γ 是一个半径为 R 的圆柱面和一个斜截平面 Π 的交集，我们可以用两个半径为 R 的球面 Σ_1, Σ_2 由上、下两端，沿著柱面向截面 Π 滑动，一直到分别和 Π 相切于 F_1, F_2 的位置（如 [图 8-1] 所示）。



[图 8-1]

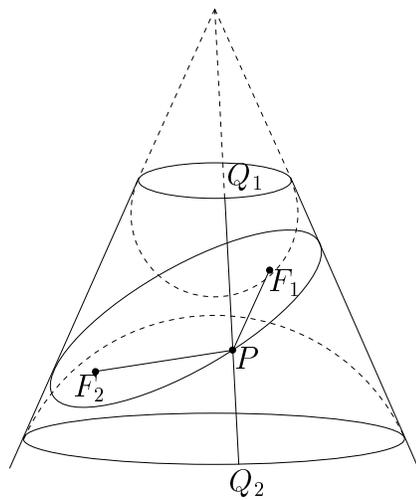
令 Γ_1, Γ_2 分别是上、下球面 Σ_1, Σ_2 和柱面相切的圆。设 P 是椭圆 Γ 上任给一点， $\overline{Q_1Q_2}$ 是柱面上过 P 点的那一条直线段， $Q_1 \in \Gamma_1, Q_2 \in \Gamma_2$ 。则有

$$\overline{PF_1} = \overline{PQ_1}, \quad \overline{PF_2} = \overline{PQ_2} \quad (\text{定点到一个球面的切线长相等})$$

$$\overline{PF_1} + \overline{PF_2} = \overline{Q_1Q_2} \quad (\text{定长})$$

大体上，这也就是当年古希腊几何学家运用圆柱和球面的简朴特性所得出的「圆柱斜截线」的几何特性及其证明。这的确是一个令人鼓舞的杰作！

后来又发现上述简洁精警的证明其实可以稍加推广，亦即把圆柱面更换成圆锥面依然成立，如 [图 8-2] 所示。

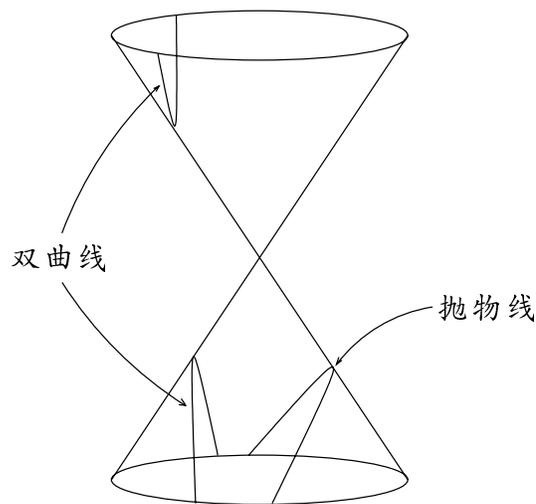


[图 8-2]

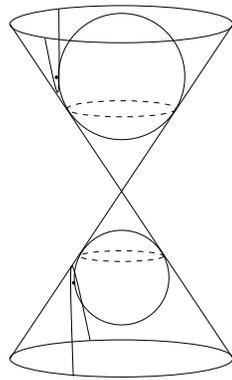
再者，如 [图 8-3] 所示，平面和圆锥面的交截还可以产生另外两种曲线，亦即现在叫做双曲线和抛物线者也。如 [图 8-4]、[图 8-5] 所示，双曲线也有两个焦点，而抛物线则只有一个焦点，而且也可以用类似的几何论证，证明双曲线和抛物线的几何特性分别如下，即：

双曲线： $|\overline{PF_1} - \overline{PF_2}| = \text{定值}$ ；

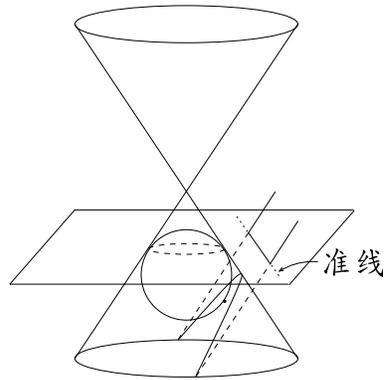
抛物线： $\overline{PF} = d(P, \ell)$ ，其中 ℓ 是准线。



[图 8-3]



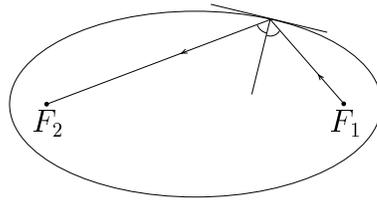
[图 8-4]



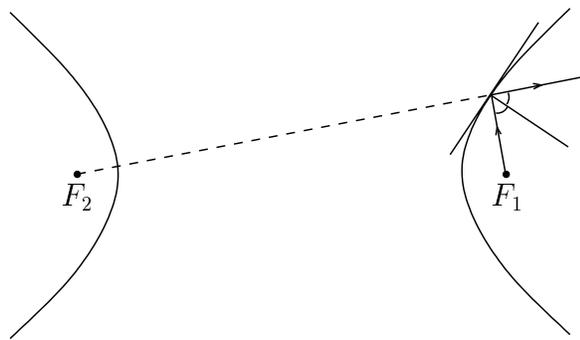
[图 8-5]

8.2 圆锥截线的光学性质

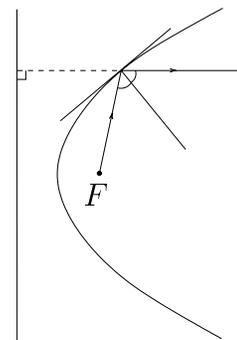
为什么把上面所说的特殊点 F_1, F_2 和抛物线的 F 叫做焦点呢？这种名称根源于椭圆、双曲线和抛物线的下述光学性质，即如 [图 8-6]、[图 8-7]、[图 8-8] 所示。



[图 8-6]



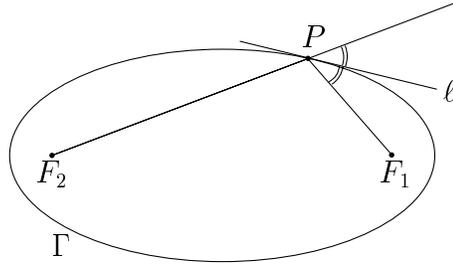
[图 8-7]



[图 8-8]

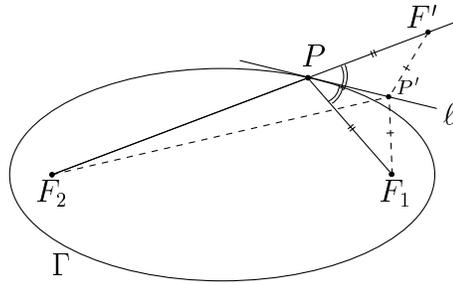
以上各点在 Apollonius 的论著中都有详细的几何证明。兹简述在椭圆情况的证明如下（其余两者的证明则留作习题）：

令 P 为椭圆 Γ 上任意给定点。如 [图 8-6'] 所示，过 P 点作 F_2PF_1 外角的角平分线 ℓ ：



[图 8-6']

现证明 ℓ 和 Γ 相切于 P 点，亦即 ℓ 上任何相异于 P 之点 P' 皆在椭圆 Γ 之外，即 $\overline{F_2P'} + \overline{P'F_1} > \overline{F_2P} + \overline{PF_1}$ 。兹证之如下：



[图 8-6'']

如 [图 8-6''] 所示，在 F_2P 线上取 F' 使得 $\overline{PF'} = \overline{PF_1}$ ，则有 $\triangle P'PF_1 \cong \triangle P'PF'$ (S.A.S.)。由此即得

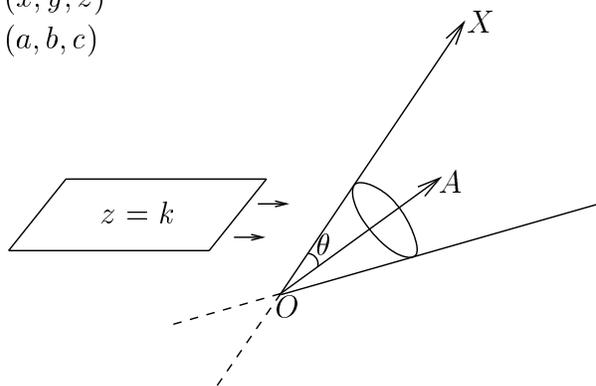
$$\begin{aligned} \overline{F_2P} + \overline{PF_1} &= \overline{F_2P} + \overline{PF'} \\ &< \overline{F_2P'} + \overline{P'F'} \\ &= \overline{F_2P'} + \overline{P'F_1} \end{aligned}$$

由此易证 F_1P, F_2P 两者和在 P 点的法线 (normal line) 所成的角度必定相等（亦即：入射角 = 反射角）。

8.3 圆锥截线和二次曲线

到了十七世纪，几何学的研究方法出现了革命性的突破，那就是笛卡儿 (Descartes)、费玛 (Fermat) 所创导的解析几何学。而这种新方法的牛刀小试就自然是把这种新引进的坐标解析法用来研讨圆锥截线的几何，以期能够温故知新。首先，我们要来看一看一个圆锥截线的方程式究竟为何？

$$\begin{aligned}\overrightarrow{OX} &= (x, y, z) \\ \overrightarrow{OA} &= (a, b, c)\end{aligned}$$



[图 8-9]

$$\overrightarrow{OX} \cdot \overrightarrow{OA} = |\overrightarrow{OX}| \cdot |\overrightarrow{OA}| \cdot \cos \theta \quad (\text{圆锥的方程式})$$

$$\text{亦即} \quad (ax + by + cz)^2 = (x^2 + y^2 + z^2)(a^2 + b^2 + c^2) \cos^2 \theta .$$

由此可见上述圆锥和 $z = k$ 的交截的方程式就是：

$$(ax + by + ck)^2 = (x^2 + y^2 + k^2)(a^2 + b^2 + c^2) \cos^2 \theta$$

总之，上述直截了当的计算说明了任何一个圆锥截线的方程式都是一个二元二次方程式，其一般形式可以写成

$$Ax^2 + 2Bxy + Cy^2 + 2Dx + 2Ey + F = 0$$

概括地说，所有圆锥截线都是二次曲线。很自然地我们会反问，是否任何二次曲线也都是圆锥截线呢？

为了要解答上述逆问题，就自然地会认识到下面这种「坐标变换不变量」的基本思想。

8.4 坐标变换和不变量

空间的一个重要本质是它完美的对称性和均匀性，它在坐标解析几何中的反映就是所有正交坐标系（亦称之为笛卡儿坐标系）之间的互换等价性。当我们用坐标解析法去研讨几何事物时，一定要铭记在心，只有那种和坐标系的选取无关者才具有本质的几何内涵。例如当我们在研讨二次曲线时，同一条二次曲线在两个不同的正交坐标系 $\{O; x, y\}$ 和 $\{O'; x', y'\}$ 中的方程式各异。设其分别为

$$Ax^2 + 2Bxy + Cy^2 + 2Dx + 2Ey + F = 0$$

和

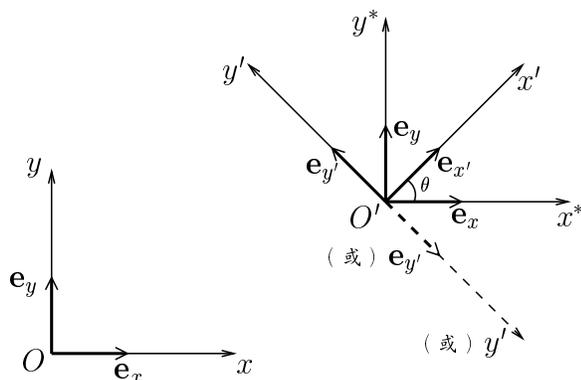
$$A'x'^2 + 2B'x'y' + C'y'^2 + 2D'x' + 2E'y' + F' = 0$$

所以上述两个方程式的系数比例组即 $\{A : B : C : D : E : F\}$ 和 $\{A' : B' : C' : D' : E' : F'\}$ 其本身并不具有本质的几何意义，而只有在坐标变换之下保持不变的那种系数组合才会具有本质性的几何意义。由此可见，要研讨一般二次曲线的几何就自然得从其方程式系数的不变组合（通常叫做方程式的坐标变换不变量）的系统研讨著手。长话短说，二次方程式在正交变换和平移变换之下具有下述三个基本不变量，即恒有：

$$\begin{aligned} A + C &= A' + C' \\ B^2 - AC &= B'^2 - A'C' \\ \begin{vmatrix} A & B & D \\ B & C & E \\ D & E & F \end{vmatrix} &= \begin{vmatrix} A' & B' & D' \\ B' & C' & E' \\ D' & E' & F' \end{vmatrix} \end{aligned}$$

通常分别用符号 H, δ, Δ 表示上述三个基本不变量。再者，所有其他的不变量都可以表达成上述三个基本不变量 $\{H, \delta, \Delta\}$ 的函数。在此且先说明 H, δ 和 Δ 在坐标变换之下的不变性：

1. 平面上的任给保长变换都可以由平移和正交（转轴）变换组合而成。相应地，平面上任给两个正交坐标系 $\{O; x, y\}$ 和 $\{O'; x', y'\}$ 之间的坐标变换也可以分解成平移和正交坐标变换的组合，如 [图 8-10] 所示。



[图 8-10]

2. 设 P 是平面上任给一点, 它在坐标系 $\{O; x, y\}$, $\{O'; x^*, y^*\}$ 和 $\{O'; x', y'\}$ 中的坐标分别为 (x, y) , (x^*, y^*) 和 (x', y') , 即有

$$\begin{aligned}\overrightarrow{OP} &= x\mathbf{e}_x + y\mathbf{e}_y \\ \overrightarrow{O'P} &= x^*\mathbf{e}_x + y^*\mathbf{e}_y = x'\mathbf{e}_{x'} + y'\mathbf{e}_{y'}\end{aligned}$$

设 $\overrightarrow{OO'} = h\mathbf{e}_x + k\mathbf{e}_y$, $\angle(\mathbf{e}_x, \mathbf{e}_{x'}) = \theta$, $\angle(\mathbf{e}_y, \mathbf{e}_{y'}) = \theta$ (或 $\pi + \theta$), 则有

$$\begin{aligned}\overrightarrow{OP} &= \overrightarrow{OO'} + \overrightarrow{O'P} = (h + x^*)\mathbf{e}_x + (k + y^*)\mathbf{e}_y \\ \Rightarrow x &= x^* + h, \quad y = y^* + k \\ \Rightarrow \begin{cases} \mathbf{e}_{x'} = \cos \theta \mathbf{e}_x + \sin \theta \mathbf{e}_y \\ \mathbf{e}_{y'} = -\sin \theta \mathbf{e}_x + \cos \theta \mathbf{e}_y \quad (\text{或 } \sin \theta \mathbf{e}_x - \cos \theta \mathbf{e}_y) \end{cases} \\ \Rightarrow \begin{cases} x^* = x' \cos \theta - y' \sin \theta \\ y^* = x' \sin \theta + y' \cos \theta \end{cases} \\ \Rightarrow \begin{cases} x = x' \cos \theta - y' \sin \theta + h \\ y = x' \sin \theta + y' \cos \theta + k \end{cases}\end{aligned}$$

3. 将上式代入一个二次曲线在 $\{O; x, y\}$ 中的方程式

$$Ax^2 + 2Bxy + Cy^2 + 2Dx + 2Ey + F = 0$$

展开, 集项而得者乃是它在 $\{O'; x', y'\}$ 中的方程式。由此, 比较系数即

得

$$\begin{cases} A' = A \cos^2 \theta + C \sin^2 \theta + B \sin 2\theta \\ B' = (C - A) \sin \theta \cos \theta + B \cos 2\theta \\ C' = A \sin^2 \theta + C \cos^2 \theta - B \sin 2\theta \\ D' = (Ah + Bk + D) \cos \theta + (Bh + Ck + E) \sin \theta \\ E' = (Bh + Ck + E) \cos \theta - (Ah + Bk + D) \sin \theta \\ F' = Ah^2 + 2Bhk + Ck^2 + 2Dh + 2Ek + F \end{cases}$$

4. 由上述二次曲线的系数的坐标变换式，即可直接验证 H 和 δ 的不变性如下，即：

$$A' + C' = A(\cos^2 \theta + \sin^2 \theta) + C(\sin^2 \theta + \cos^2 \theta) = A + C$$

而用直接展开可得

$$\begin{aligned} B'^2 &= (C - A)^2 \sin^2 \theta \cos^2 \theta + \underline{B(C - A) \cos 2\theta \sin 2\theta} + B^2 \cos^2 2\theta \\ -A'C' &= -A^2 \cos^2 \theta \sin^2 \theta - AC \cos^4 \theta + \underline{AB \cos^2 \theta \sin 2\theta} \\ &\quad - AC \sin^4 \theta - C^2 \cos^2 \theta \sin^2 \theta + \underline{BC \sin^2 \theta \sin 2\theta} \\ &\quad - \underline{AB \sin^2 \theta \sin 2\theta} - \underline{BC \sin 2\theta \cos^2 \theta} + B^2 \sin^2 2\theta \end{aligned}$$

在上式中， $B^2 \cos^2 2\theta$ 和 $B^2 \sin^2 2\theta$ 合起来当然就是 B^2 ，而 B'^2 中有底线的项又刚好和 $-A'C'$ 中有底线的项相消，所以即得

$$\begin{aligned} B'^2 - A'C' &= B^2 + (C^2 - 2AC + A^2) \sin^2 \theta \cos^2 \theta - A^2 \cos^2 \theta \sin^2 \theta \\ &\quad - C^2 \cos^2 \theta \sin^2 \theta - AC(\cos^4 \theta + \sin^4 \theta) \\ &= B^2 - AC(\cos^4 \theta + \sin^4 \theta + 2\sin^2 \theta \cos^2 \theta) \\ &= B^2 - AC(\cos^2 \theta + \sin^2 \theta)^2 = B^2 - AC \end{aligned}$$

5. 我们当然可以用同样的直接代换方法来验证 Δ 的不变性，但是其中的运算颇为繁复。下面将提供一个较佳的方法，而且亦可以突出矩阵运算的优越性。一个二次曲线可以用下述矩阵形式表达：

$$Ax^2 + 2Bxy + Cy^2 + 2Dx + 2Ey + F = (x \ y \ 1) \begin{pmatrix} A & B & D \\ B & C & E \\ D & E & F \end{pmatrix} \begin{pmatrix} x \\ y \\ 1 \end{pmatrix} = 0$$

另一方面，前述的坐标变换式在矩阵形式下亦可以写成

$$\begin{pmatrix} x \\ y \\ 1 \end{pmatrix} = \begin{pmatrix} \cos \theta & -\sin \theta & h \\ \sin \theta & \cos \theta & k \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x' \\ y' \\ 1 \end{pmatrix}$$

$$(x \ y \ 1) = (x' \ y' \ 1) \begin{pmatrix} \cos \theta & \sin \theta & 0 \\ -\sin \theta & \cos \theta & 0 \\ h & k & 1 \end{pmatrix}$$

所以由直接代换可得

$$\begin{aligned} & (x \ y \ 1) \begin{pmatrix} A & B & D \\ B & C & E \\ D & E & F \end{pmatrix} \begin{pmatrix} x \\ y \\ 1 \end{pmatrix} \\ &= (x' \ y' \ 1) \begin{pmatrix} \cos \theta & \sin \theta & 0 \\ -\sin \theta & \cos \theta & 0 \\ h & k & 1 \end{pmatrix} \begin{pmatrix} A & B & D \\ B & C & E \\ D & E & F \end{pmatrix} \begin{pmatrix} \cos \theta & -\sin \theta & h \\ \sin \theta & \cos \theta & k \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x' \\ y' \\ 1 \end{pmatrix} \\ &= (x' \ y' \ 1) \begin{pmatrix} A' & B' & D' \\ B' & C' & E' \\ D' & E' & F' \end{pmatrix} \begin{pmatrix} x' \\ y' \\ 1 \end{pmatrix} \end{aligned}$$

亦即

$$\begin{pmatrix} A' & B' & D' \\ B' & C' & E' \\ D' & E' & F' \end{pmatrix} = \begin{pmatrix} \cos \theta & \sin \theta & 0 \\ -\sin \theta & \cos \theta & 0 \\ h & k & 1 \end{pmatrix} \begin{pmatrix} A & B & D \\ B & C & E \\ D & E & F \end{pmatrix} \begin{pmatrix} \cos \theta & -\sin \theta & h \\ \sin \theta & \cos \theta & k \\ 0 & 0 & 1 \end{pmatrix}$$

由此可见，用行列式的乘法公式即有

$$\begin{aligned} \begin{vmatrix} A' & B' & D' \\ B' & C' & E' \\ D' & E' & F' \end{vmatrix} &= \begin{vmatrix} \cos \theta & \sin \theta & 0 \\ -\sin \theta & \cos \theta & 0 \\ h & k & 1 \end{vmatrix} \cdot \begin{vmatrix} A & B & D \\ B & C & E \\ D & E & F \end{vmatrix} \cdot \begin{vmatrix} \cos \theta & -\sin \theta & h \\ \sin \theta & \cos \theta & k \\ 0 & 0 & 1 \end{vmatrix} \\ &= \begin{vmatrix} A & B & D \\ B & C & E \\ D & E & F \end{vmatrix} \end{aligned}$$

将上述基本不变量的理解再和坐标变换结合运用，就可以有系统地把二次曲线加以几何分类。例如当 $\delta \neq 0$ 的情形，即可求得一个平移变换：

$$\begin{cases} x = x^* + h \\ y = y^* + k \end{cases}$$

使得 $D^* = E^* = 0$ ，亦即由下述线性方程组

$$\begin{cases} Ah + Bk + D = 0 \\ Bh + Ck + E = 0 \end{cases}$$

即可求得上述平移中的 $\{h, k\}$ 。这样，原给曲线在坐标系 $\{O'; x^*, y^*\}$ 中的方程式就变为（注意：平移变换不会改变 A, B, C 的值）

$$Ax^{*2} + 2Bx^*y^* + Cy^{*2} + F^* = 0$$

其中 $F^* = -\frac{\Delta}{\delta}$ 。接著我们还可以用适当的转轴变换，即

$$\begin{cases} x^* = x' \cos \theta - y' \sin \theta \\ y^* = x' \sin \theta + y' \cos \theta \end{cases}$$

使得 $B' = (C - A) \sin \theta \cos \theta + B \cos 2\theta = 0$ ，亦即

$$\tan 2\theta = \frac{2B}{A - C} \Rightarrow \theta = \frac{1}{2} \tan^{-1} \frac{2B}{A - C}$$

（若取 $0 < \theta \leq \frac{\pi}{4}$ 或 $\frac{3\pi}{4} \leq \theta < \pi$ 者，则可以保持 $A' - C'$ 和 $A - C$ 同号，试证之。）则曲线在坐标系 $\{O'; x', y'\}$ 中的方程式为

$$A'x'^2 + C'y'^2 + F' = 0, \quad (F' = F^* = -\frac{\Delta}{\delta})$$

而上述 $\{A', C'\}$ 是可以由方程组 $\begin{cases} A' + C' = H = (A + C) \\ A'C' = -\delta = AC - B^2 \end{cases}$ 直接求得，

亦即 $\{A', C'\} = \{\frac{H}{2} \pm \sqrt{(\frac{H}{2})^2 + \delta}\}$ 。

[注]：若一条二次曲线在变换 $(x, y) \mapsto (-x, -y)$ 之下保持不变，则称之为具有「心对称」者。把这个变换代入二次曲线的方程式中，即有

$$Ax'^2 + 2Bx'y' + Cy'^2 - 2Dx' - 2Ey' + F = 0$$

由此可见一条二次曲线具有心对称的充要条件就是 $D = E = 0$ 。当 $\delta \neq 0$ 时，我们可以求得平移变换使得 $D^* = E^* = 0$ ，所以在这个情况二次曲线是具有心对称者；但若果 $\delta = 0$ ，则情况会是怎样？我们可以用下面的分析来研讨：

设 $\delta = 0$ ，我们可以先用一个转轴变换

$$\begin{cases} x = x^* \cos \theta - y^* \sin \theta \\ y = x^* \sin \theta + y^* \cos \theta \end{cases}$$

使得 $B^* = 0$ 。因为 $0 = \delta = B^{*2} - A^*C^* = -A^*C^*$ ，所以 A^* 和 C^* 之中必有一个为零，不妨由 θ 的适当选取（即使得 $A - C, A^* - C^*$ 同号或异号的 θ ）使得 $C^* = 0$ 。所以经过上述转轴变换之后，二次曲线在坐标系 $\{O; x^*, y^*\}$ 的方程式为

$$A^* x^{*2} + 2D^* x^* + 2E^* y^* + F^* = 0$$

再用一个平移变换 $x^* = x' - \frac{D^*}{A^*}, y^* = y'$ 即有

$$A' x'^2 + 2E' y' + F' = 0, \quad (A' = A^*, E' = E^*)$$

上式中的 $\{A', E'\}$ 可以用不变量求得： $H = A', \Delta = -A'E'^2$ ，即 $E' = \pm \sqrt{-\frac{\Delta}{H}}$ 。当 $E' \neq 0$ ，亦即 $\Delta \neq 0$ 时，这是一条抛物线，而且还可以用平移变换 $x' = x'', y' = y'' - \frac{F'}{2E'}$ 将方程式简化为 $A' x''^2 + 2E' y'' = 0$ 。当 $E' = 0$ ，亦即 $\Delta = 0$ 时，则二次曲线会蜕化至一对直线（实或虚）的情形。

总结上述分析，我们可以用不变量 H, δ, Δ 把一般的二次曲线作系统分类如下：

第一种 $\delta \neq 0$ ：在这种情形，二次曲线的方程式经过适当的平移、转轴变换后就可以写成

$$A' x'^2 + C' y'^2 + F' = 0, \quad (F' = F^* = -\frac{\Delta}{\delta})$$

- (1) 若 A', C' 同号（即 $\delta = -A'C' < 0$ ）而且与 F' 异号（即 $H \cdot \Delta < 0$ ），则属于椭圆的类别 $\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$ ；
- (2) 若 A', C', F' 同号（即 $\delta < 0, H \cdot \Delta > 0$ ），则属于虚椭圆的类别 $\frac{x^2}{a^2} + \frac{y^2}{b^2} = -1$ ；
- (3) 若 A', C' 同号（即 $\delta < 0$ ）而且 $F' = 0$ （即 $\Delta = 0$ ），则二次曲线会蜕化为一点 $\frac{x^2}{a^2} + \frac{y^2}{b^2} = 0$ （或可当做交于一实点的一对虚直线）；

- (4) 若 A', C' 异号 (即 $\delta > 0$) 而且 $F' \neq 0$ (即 $\Delta \neq 0$)，则属于双曲线的类别 $\frac{x^2}{a^2} - \frac{y^2}{b^2} = \pm 1$ ；
- (5) 若 A', C' 异号 (即 $\delta > 0$) 而且 $F' = 0$ (即 $\Delta = 0$)，则属于一对相交直线的类别 $\frac{x^2}{a^2} - \frac{y^2}{b^2} = 0$ ；

第二种 $\delta = 0$ ：在这种情形，二次曲线的方程式经过适当的转轴、平移变换後就可以写成

$$A'x'^2 + 2E'y' + F' = 0$$

- (6) 若 $E' \neq 0$ (即 $\Delta = -A'E'^2 \neq 0$)，则属于抛物线的类别 $y = ax^2$ ；
- (7) 若 $E' = 0$ (即 $\Delta = 0$)，则属于一对 (平行，相重或虚) 直线的类别 $x^2 = a^2$, $x^2 = 0$ 或 $x^2 = -a^2$ 。

[注]：由上述分类可以直接看到，一条二次曲线会蜕化为一对 (平行，相重，相交或虚) 直线的充要条件就是 $\Delta = 0$ 。

8.5 五点定一「二次曲线」和六点共在一「二次曲线」的条件

在初等几何中有两个熟知的事实，即两点定一直线和不共线三点定一圆。若改用解析观点来看，上述几何事实与直线和圆的方程式中系数比的个数是密切相对应的。亦即

$$\text{直线方程：} Ax + By + C = 0$$

中具有两个相互独立的系数比 $\{A:B, A:C\}$ ；

$$\text{圆方程：} A(x^2 + y^2) + 2Dx + 2Ey + F = 0$$

中具有三个相互独立的系数比 $\{A:D, A:E, A:F\}$ 。再者，运用线性方程组的基础理论，我们还可以用行列式直接写下三点共线和四点共圆的坐标条件式，即

$P_i(x_i, y_i)$, $1 \leq i \leq 3$, 三点共线的条件是

$$\begin{vmatrix} x_1 & y_1 & 1 \\ x_2 & y_2 & 1 \\ x_3 & y_3 & 1 \end{vmatrix} = 0$$

$P_i(x_i, y_i)$, $1 \leq i \leq 4$, 四点共圆的条件是

$$\begin{vmatrix} x_1^2 + y_1^2 & x_1 & y_1 & 1 \\ x_2^2 + y_2^2 & x_2 & y_2 & 1 \\ x_3^2 + y_3^2 & x_3 & y_3 & 1 \\ x_4^2 + y_4^2 & x_4 & y_4 & 1 \end{vmatrix} = 0$$

同样理由, 我们也有下述六点 $\{P_i(x_i, y_i), 1 \leq i \leq 6\}$ 共在一个二次曲线上的代数条件是

$$\begin{vmatrix} x_1^2 & x_1 y_1 & y_1^2 & x_1 & y_1 & 1 \\ x_2^2 & x_2 y_2 & y_2^2 & x_2 & y_2 & 1 \\ x_3^2 & x_3 y_3 & y_3^2 & x_3 & y_3 & 1 \\ x_4^2 & x_4 y_4 & y_4^2 & x_4 & y_4 & 1 \\ x_5^2 & x_5 y_5 & y_5^2 & x_5 & y_5 & 1 \\ x_6^2 & x_6 y_6 & y_6^2 & x_6 & y_6 & 1 \end{vmatrix} = 0$$

因为它就是存在有不全为零的系数组 $\{A, B, C, D, E, F\}$ 使得

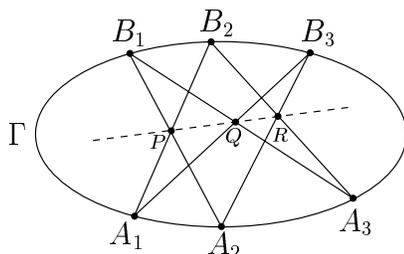
$$Ax_i^2 + Bx_i y_i + Cy_i^2 + Dx_i + Ey_i + F = 0, \quad 1 \leq i \leq 6$$

的充要条件。假如我们把 $\{(x_i, y_i), 1 \leq i \leq 5\}$ 想成是取定者, 而把 (x_6, y_6) 想成是变动者, 而且改用符号 (x, y) , 则上述六阶行列式也就是过 $\{P_i(x_i, y_i), 1 \leq i \leq 5\}$ 这五点的唯一二次曲线的方程式。这也就是「五点定一二次曲线」的代数表述。

8.6 Pascal 定理和 Pappus 定理

上面所讨论的是六点共在一二次曲线上的坐标代数条件式。很自然地我们还应该研讨六点共在一二次曲线的几何条件究竟是什麼? 这也就是著名的 Pascal 定理。Blaise Pascal (1623–1662) 英年早逝, 他在少年时代曾著有一本圆锥截线的小册子, 可惜当年印份极少, 很早就连孤本也找

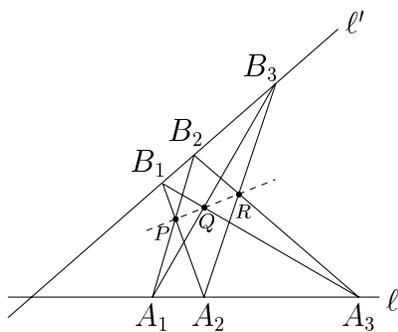
不到了。但是 Leibniz 说他早年曾经拜读过，还记得其中有下列这个美妙的定理，即在一个圆锥截线 Γ 上任取六点 $\{A_1, A_2, A_3; B_1, B_2, B_3\}$ ，如 [图 8-11] 所示，令 $P = A_1B_2 \cap A_2B_1$ ， $Q = A_1B_3 \cap A_3B_1$ ， $R = A_2B_3 \cap A_3B_2$ ，则恒有 $\{P, Q, R\}$ 三点共线。



[图 8-11]

但是 Leibniz 说他已经不记得 Pascal 在书中所给的证明了，所以我们现在只能从 Pascal 当代的几何认知环境来作一种合情合理的探讨与推测。

早在纪元三世纪末，Pappus 的几何著作中即已出现下述命题，即如 [图 8-12] 所示，设 $\{A_1, A_2, A_3\}$ ， $\{B_1, B_2, B_3\}$ 是分别在 ℓ 和 ℓ' 上的三点组。令 $P = A_1B_2 \cap A_2B_1$ ， $Q = A_1B_3 \cap A_3B_1$ ， $R = A_2B_3 \cap A_3B_2$ ，则 $\{P, Q, R\}$ 三点共线。



[图 8-12]

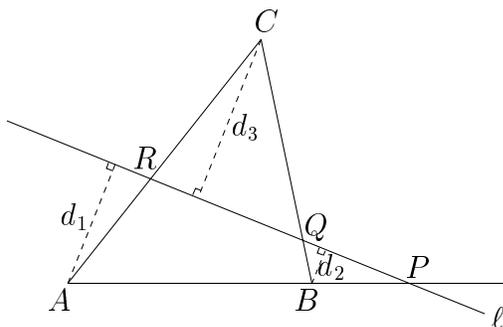
可惜这样美妙的命题乃是仅仅以一个习题形式出现而并没有在书中给以论证，显然 Pappus 和当代的几何学家们是知道如何证明上述命题的。他们当时究竟是怎麼证的呢？却又是一个有待探讨与推测的「证法考古题」。

现在让我们先来探讨 Pappus 定理的古证理当如何。由于这个命题是以习题形式出现，我们相信它是能够以当年熟知的几何定理推导出来的。回顾当年有关的几何定理，其中就只有 Menelous 逆定理是以「三点共线」为其结论者，所以我们可以想到当年大概是要运用 Menelous 逆定理来证明 Pappus 定理。我们先重温 Menelous 定理和其逆定理（详见第二章例题）：

【Menelous 定理和逆定理】：设直线 l 与 $\triangle ABC$ 三边所在之直线 AB, BC, CA 分别相交于 P, Q, R 相异三点，则下述有向长度比之乘积恒等于 -1 ：

$$\frac{\overrightarrow{AP}}{\overrightarrow{PB}} \cdot \frac{\overrightarrow{BQ}}{\overrightarrow{QC}} \cdot \frac{\overrightarrow{CR}}{\overrightarrow{RA}} \equiv -1$$

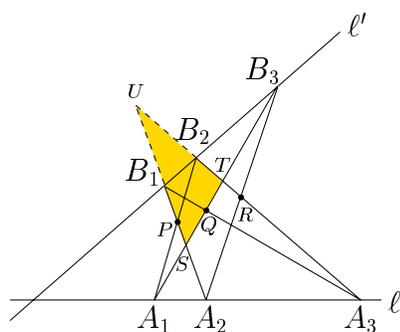
反之，若相异三点 P, Q, R 分别在 $\triangle ABC$ 三边 AB, BC, CA 之上并满足上面有向长度比的条件式，则 P, Q, R 三点共线。



[图 8-13]

现在让我们运用 Menelous 定理和逆定理来证明 Pappus 定理。

就命题的图形所给的八条直线，我们可以就地取材，选取其中三条分别过 P, Q, R 者来构成一个三角形，然后再对于其余五条直线运用 Menelous 定理。如 [图 8-14] 所示，我们选取了 A_1B_3, A_2B_1, A_3B_2 来构成 $\triangle STU$ 。



[图 8-14]

对 A_1B_2 运用 Menelous 定理 :

因 $ST \cap A_1B_2 = \{A_1\}$, $TU \cap A_1B_2 = \{B_2\}$, $US \cap A_1B_2 = \{P\}$, 所以 :

$$\frac{\overrightarrow{SA_1}}{\overrightarrow{A_1T}} \cdot \frac{\overrightarrow{TB_2}}{\overrightarrow{B_2U}} \cdot \frac{\overrightarrow{UP}}{\overrightarrow{PS}} = -1$$

对 A_2B_3 运用 Menelous 定理 :

因 $ST \cap A_2B_3 = \{B_3\}$, $TU \cap A_2B_3 = \{R\}$, $US \cap A_2B_3 = \{A_2\}$, 所以 :

$$\frac{\overrightarrow{SB_3}}{\overrightarrow{B_3T}} \cdot \frac{\overrightarrow{TR}}{\overrightarrow{RU}} \cdot \frac{\overrightarrow{UA_2}}{\overrightarrow{A_2S}} = -1$$

对 A_3B_1 运用 Menelous 定理 :

因 $ST \cap A_3B_1 = \{Q\}$, $TU \cap A_3B_1 = \{A_3\}$, $US \cap A_3B_1 = \{B_1\}$, 所以 :

$$\frac{\overrightarrow{SQ}}{\overrightarrow{QT}} \cdot \frac{\overrightarrow{TA_3}}{\overrightarrow{A_3U}} \cdot \frac{\overrightarrow{UB_1}}{\overrightarrow{B_1S}} = -1$$

对 l 运用 Menelous 定理 (然後颠倒) :

因 $ST \cap l = \{A_1\}$, $TU \cap l = \{A_3\}$, $US \cap l = \{A_2\}$, 所以 :

$$\frac{\overrightarrow{A_1T}}{\overrightarrow{SA_1}} \cdot \frac{\overrightarrow{A_3U}}{\overrightarrow{TA_3}} \cdot \frac{\overrightarrow{A_2S}}{\overrightarrow{UA_2}} = -1$$

对 l' 运用 Menelous 定理 (然後颠倒) :

因 $ST \cap l' = \{B_3\}$, $TU \cap l' = \{B_2\}$, $US \cap l' = \{B_1\}$, 所以 :

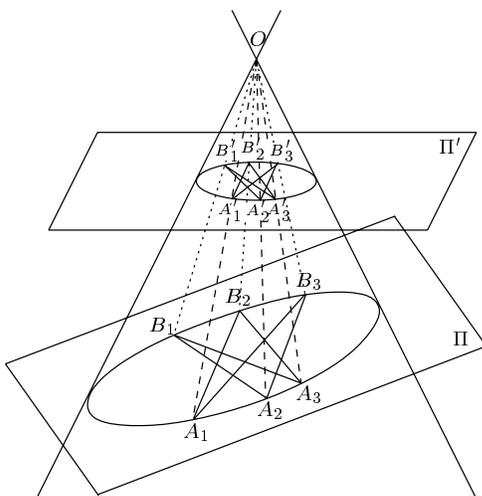
$$\frac{\overrightarrow{B_3T}}{\overrightarrow{SB_3}} \cdot \frac{\overrightarrow{B_2U}}{\overrightarrow{TB_2}} \cdot \frac{\overrightarrow{B_1S}}{\overrightarrow{UB_1}} = -1$$

现将上面所得的五条等式相乘起来，便会发现有很多项能够互相抵消，最後便可得出下面等式：

$$\frac{\overrightarrow{SQ}}{\overrightarrow{QT}} \cdot \frac{\overrightarrow{TR}}{\overrightarrow{RU}} \cdot \frac{\overrightarrow{UP}}{\overrightarrow{PS}} = -1$$

运用 Menelous 逆定理，即知 P, Q, R 三点共线，Pappus 定理证毕。

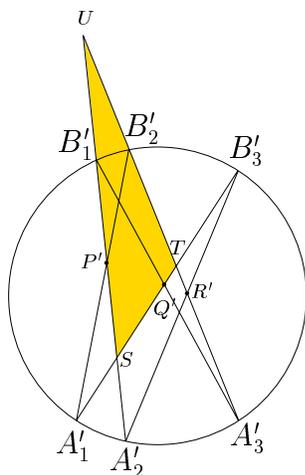
接著让我们来探讨 Pascal 当年的证明究竟有那种合情合理的可能性。首先，他当然熟知 Pappus 定理，而且也知道对于 Pappus 定理的上述证法。再者，他所要证明者，其实就是 Pappus 定理在非蜕化锥线的推广。如 [图 8-15] 所示， Σ, Σ' 分别是圆锥 Γ 和两个平面 Π, Π' 的截线，而且 Σ' 是一个圆。令 $\{A_1, A_2, A_3; B_1, B_2, B_3\}$ 是 Σ' 上的相应六点使得 $\{O, A_i, A'_i\}$ 和 $\{O, B_i, B'_i\}$ 皆为共线三点组。



[图 8-15]

令 $P' = A_1B_2 \cap A_2B_1$, $Q' = A_1B_3 \cap A_3B_1$, $R' = A_2B_3 \cap A_3B_2$ 。不难看到 $\{O, P, P'\}$, $\{O, Q, Q'\}$, $\{O, R, R'\}$ 也都是共线三点组。所以 $\{P', Q', R'\}$ 三点共线的充要条件乃是 $\{P, Q, R\}$ 三点共线。由此可以想到，Pascal 当年想要证明者，其实可以归于圆的情形去加以证明的。换句话说，只要能够把它对圆的「特殊情形」加以证明，则「一般情形」的证明就可以运用上述透视对应加以推论而得！按照 Pascal 当代的几何认知来判断，上面这种运用圆锥截线「与生俱来」的透视对应把 Pascal 定理的

证明归于圆的情形加以论证的想法是相当自然的。总之，只要有了这个想法，就自然会同样地去构造在 Pappus 定理的古证中那个分别含有 P', Q', R' 于其三边之上的那个三角形，如 [图 8-16] 所示的 $\triangle STU$ 。



[图 8-16]

把 Pappus 定理中所用的 [图 8-16] 和 [图 8-14] 相比较，其差别是前者除 $\triangle STU$ 的三边之外尚有五条直线而后者则尚有三条直线和一个圆。我们当然可以先对这三条直线运用 Menelous 定理，即得

$$\begin{aligned} \frac{\overrightarrow{SA'_1}}{\overrightarrow{A'_1T}} \cdot \frac{\overrightarrow{TB'_2}}{\overrightarrow{B'_2U}} \cdot \frac{\overrightarrow{UP'}}{\overrightarrow{P'S}} &= -1 \\ \frac{\overrightarrow{SB'_3}}{\overrightarrow{B'_3T}} \cdot \frac{\overrightarrow{TR'}}{\overrightarrow{R'U}} \cdot \frac{\overrightarrow{UA'_2}}{\overrightarrow{A'_2S}} &= -1 \\ \frac{\overrightarrow{SQ'}}{\overrightarrow{Q'T}} \cdot \frac{\overrightarrow{TA'_3}}{\overrightarrow{A'_3U}} \cdot \frac{\overrightarrow{UB'_1}}{\overrightarrow{B'_1S}} &= -1 \end{aligned}$$

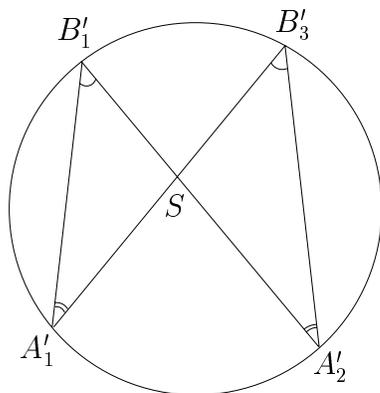
相乘之後再重组，其所得为

$$\frac{\overrightarrow{SA'_1}}{\overrightarrow{A'_2S}} \cdot \frac{\overrightarrow{SB'_3}}{\overrightarrow{B'_1S}} \cdot \frac{\overrightarrow{TB'_2}}{\overrightarrow{B'_3T}} \cdot \frac{\overrightarrow{TA'_3}}{\overrightarrow{A'_1T}} \cdot \frac{\overrightarrow{UA'_2}}{\overrightarrow{B'_2U}} \cdot \frac{\overrightarrow{UB'_1}}{\overrightarrow{A'_3U}} \cdot \frac{\overrightarrow{SQ'}}{\overrightarrow{Q'T}} \cdot \frac{\overrightarrow{TR'}}{\overrightarrow{R'U}} \cdot \frac{\overrightarrow{UP'}}{\overrightarrow{P'S}} = -1$$

把上式和 Menelous 逆定理中的条件相比，上式左侧多了一个

$$\frac{\overrightarrow{SA'_1}}{\overrightarrow{A'_2S}} \cdot \frac{\overrightarrow{SB'_3}}{\overrightarrow{B'_1S}} \cdot \frac{\overrightarrow{TB'_2}}{\overrightarrow{B'_3T}} \cdot \frac{\overrightarrow{TA'_3}}{\overrightarrow{A'_1T}} \cdot \frac{\overrightarrow{UA'_2}}{\overrightarrow{B'_2U}} \cdot \frac{\overrightarrow{UB'_1}}{\overrightarrow{A'_3U}}$$

在此，自然会想到应该运用圆的几何性质去证明它恒等于 1。其实，如 [图 8-17] 所示，由圆周角相等易见



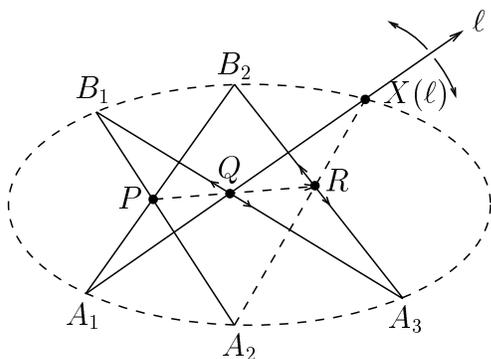
[图 8-17]

$\triangle SA'_1B'_1$ 和 $\triangle SA'_2B'_3$ 相似，所以 $\frac{\overrightarrow{SA'_1}}{\overrightarrow{A'_2S}} \cdot \frac{\overrightarrow{SB'_3}}{\overrightarrow{B'_1S}} = 1$ 。同理即得

$$\frac{\overrightarrow{TB'_2}}{\overrightarrow{B'_3T}} \cdot \frac{\overrightarrow{TA'_3}}{\overrightarrow{A'_1T}} = 1, \quad \frac{\overrightarrow{UA'_2}}{\overrightarrow{B'_2U}} \cdot \frac{\overrightarrow{UB'_1}}{\overrightarrow{A'_3U}} = 1$$

这样就可以证得 P', Q', R' 三点共线，再用透视对应即可证得 P, Q, R 三点也共线。由上面这一段分析和探讨来看，我们相信 Pascal 当年的证法，大致应该如此。

[注]：(一) Pascal 定理不但提供了六点共在一锥线的充要条件的几何表述，而且也提供了描绘那个五点所定的锥线的点线作图法。



[图 8-18]

(二) 锥线的几何本质乃是圆的透视投影。由此可见，关于圆的任何一个在「透视投影之下保持不变」的几何性质，当然也是任给锥线所共有的几何性质。上述 Pascal 定理只是其中之一，当然它是其中极为重要的锥线特征性质；其他各种各样的锥线的透视投影不变的共性，都可以用它来加以推导。

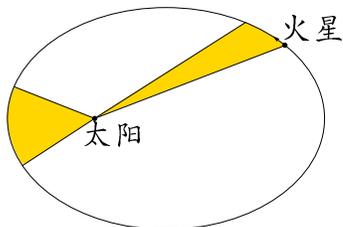
(三) Pappus 定理当年在 Pappus 的著作中，仅仅以习题形式出现，究其原因，大概是当代的几何学家们认为它只是一个美妙的命题而已；它在「当时」的几何体系中，并没有基本的重要性。其实 Pappus 定理的深刻蕴涵，一直到了十九世纪後半在射影几何学的研讨中，才展现出它的基本重要性。它乃是整个射影几何的基本定理 (fundamental theorem of projective geometry)。

8.7 Kepler 行星三定律

日月和四季的变化，具有明确的周期性和规律性；而且它们都全面地影响著大地上的一切现象和活动。所以自古以来各古文明都十分重视天象的观察也都认识到各种各样的天文知识。例如中国、埃及、巴比伦、希腊等古文明都发现漫天斗星之中，几乎所有的星象之间的相对位置都是固定不动，可以把它们想象成固定在一个旋转不休的「天球」上的「恒星」(fixed stars)。唯独有很少几颗相当明亮的星星，它们和其他星星的相对位置则是一直在变动，乃是在天球上不停游走的「行星」(planets：其意义是漫游者)。这几颗特出的行星，自古以来就吸引著世世代代的星象观察者。在古代的天文学中，观察和研讨日、月、行星在天空中运行的规律一直是最主要的中心课题。此事一直到十六、十七世纪之交，Kepler 行星三定律的发现才克竟其功，这也就是开创近代天文学的重大突破和奠基者。长话短说，Kepler 三定律乃是以 Tyco de Brahe 数十年对于行星运行夜以继夜，从不间断的详尽、致细的观察数据为基础。Kepler 再花了十多年功夫，用各种各样几何模型的数理分析去探索那个能够和 Tyco de Brahe 的观察数据充分吻合的行星运行模式。其间历经艰辛，屡败不馁才在 1609 年终于完成了火星 (Mars) 的运行模式，这也就是他在《新天文学》(Astronomia Nova, 1609) 中所发表的第一与第二定律，即：

【第一定律】：火星绕太阳以一个以太阳位于其焦点之一的椭圆运行。

【第二定律】：在单位时间内，火星和太阳连线所扫过的面积恒相等，如 [图 8-19] 所示。

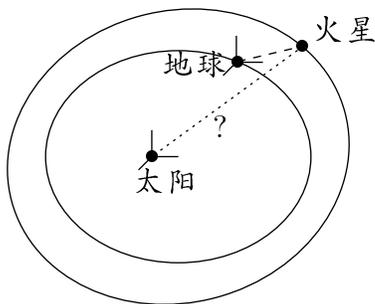


[图 8-19]

他接著又花了十年苦功去研究其他各行星的运行规律，它们不但也满足同样的第一、第二定律，而且在当年所知的六个行星（包括地球）的椭圆轨道的长径和其周期之间还满足下述美妙的第三定律，发表在《宇宙的和谐》(Harmonices Mundi, 1619)：

【第三定律】：各个行星不但也满足同样的第一、第二定律，而且其轨道长径的立方和周期的平方之间的比值皆相等。

首先，让我们以後见之明，再来回顾、分析一下当年 Kepler 在发现行星三定律所经历的艰难苦功究竟是那一种数理分析？此事的细节说来话长，而 Kepler 所发表的《新天文学》其实乃是一本把他当年所经历的困难、失败与成功的实况实录，在此只简要地概述其中主要的困难所在。以後见之明来看，地球和火星都绕著太阳作椭圆的非等速运动，大约如 [图 8-20] 所示。



[图 8-20]

要注意，当年 Tycho de Brahe 以及自古以来所有的观察数据都是由地球向火星夜以继夜的观察，而 Kepler 所要探索的运行规律则是火

星和地球究竟是如何绕太阳作相对运动。因此他先得要作出一种猜想 (conjecture) 然後根据它来计算在各个时刻地球和火星之间的方向，并且把计算所得和 Tyco de Brahe 的实测方向逐一比较，看看是否逐一吻合？（亦即偏差在观察误差之内。）由此可见，其所涉及的计算在本质上乃是个不断在作相对运动的坐标系之间的坐标变换。在此得提醒一个现代读者的，Kepler 在作这种计算的年代比笛卡儿的坐标几何的诞生 (1637) 还要早四十年！总之，这是一种十分艰巨的「科学猜谜」。因为我们所能实测者乃是由地球到火星的观测数据，而 Kepler 所要探索者则是地球和火星对于太阳的相对运动的规律！他在这个无比艰巨的科学猜谜的征程上屡猜屡败，但他坚毅不屈，奋斗了十多年才成功地发现了第一、第二定律。然後，又继续努力了整整十年，才发现了第三定律。这种艰苦奋斗，屡败不馁的经历，在 Kepler 的名著 *Astronomia Nova* 中都有详细的纪录。他不但告诉我们後之来者那辉煌的最後成功，他也忠实地记录下他每次失败的经验。其中最令人钦佩也最值得後学後进效法的是他在第 19 章之末所写的一段话。在那一章中他对于又一个猜想进行计算，发现它和 Tyco de Brahe 的火星实测数据都吻合到 2 分（亦即 $\frac{1}{30}$ 度）而这正是上述观测的确保准确度。但是在他正要为大功终于告成而欢心鼓舞的时刻，突然记起还有另外四个实测数据忘了验算。一算之下发现误差是 8 分。在这个时刻，他写下这样一段话¹：「上帝赐与我们 Tyco de Brahe 的实测数据，是确保准到 2 分的。8 分的误差是不能接受的，让我们重新开始吧。」Kepler 就是以这种高洁的科学精神，不得真相永不休的奋斗才终于发现了 Kepler 行星运行三定律的。

8.8 由 Kepler 定律到牛顿万有引力定律

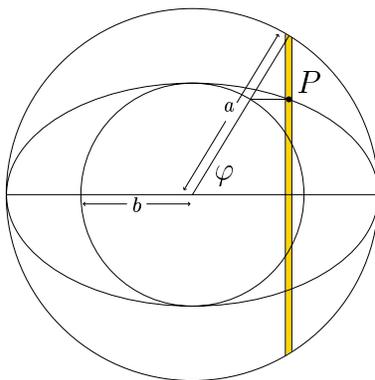
Kepler 三定律清楚地描述了行星运行的简单模式。一个自然的问题就是：「为什麼行星会以太阳为其一焦点的椭圆轨道运行？」牛顿对这个问题给了一个直接的答案，即行星与太阳之间存在满足平方反比

¹原文英译：“Since the divine benevolence has vouchsafed us Tyco Brahe, a most diligent observer, from whose observations the 8' error in this Ptolemaic computation is shown, Now, because they could not have been ignored, these eight minutes alone will have led the way to the reformation of all of astronomy,”

定律² (inverse square law) 的引力。而且这种引力同样地存在于任何两个物体之间，不论是火星与太阳或是苹果与地球，都是同一种引力；这就是牛顿著名的万有引力定律。现在让我们以现今的符号体系来重新看看牛顿这个对後世科学发展有著深远影响的工作。首先我们需要对椭圆的几何性质有一定的了解。

(一) 椭圆面积公式：椭圆面积 = πab

设椭圆的长、短径分别为 $2a, 2b$ 。现构造两个圆，半径分别为 a, b ，并把椭圆夹于两者之间，如 [图 8-21] 所示。



$$\begin{aligned} P &= P(x, y) \\ x &= a \cos \varphi \\ y &= b \sin \varphi \end{aligned}$$

[图 8-21]

考虑图中那条过 $P(x, y)$ 的窄条面积。由于 $x = a \cos \varphi, y = b \sin \varphi$ ，在椭圆内的窄条面积和整条窄条面积的比率约为：

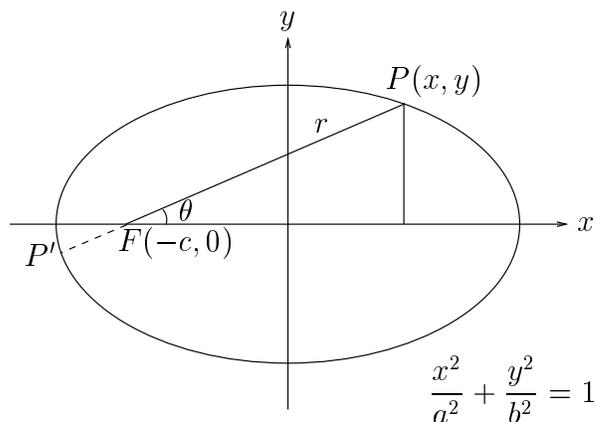
$$\frac{2b \sin \varphi \Delta x}{2a \sin \varphi \Delta x} = \frac{b}{a}$$

这是一个与 $P(x, y)$ 位置无关的常数。当把所有这种窄条的面积加起来时，便得

$$\text{椭圆面积} = \pi a^2 \cdot \frac{b}{a} = \pi ab$$

²注：在这里有一段小插曲。平方反比引力的猜测并不是源自牛顿，而是 Hooke 在研究自由落体轨迹时的猜想。但 Hooke 本身是一位实验科学家，他不懂怎样用数学方法去证明这个猜想。虽然他和牛顿之间曾因光是波动或粒子学说引起争论，他仍然写信给牛顿，希望牛顿能以数理分析的方法来给与证明。

(二) 椭圆的极坐标方程式



[图 8-22]

如 [图 8-22] 所示，椭圆上一点 P 有 (x, y) 和 (r, θ) 两种表示方法，而 (x, y) 和 (r, θ) 之间的转换可以用

$$r \cos \theta = x + c, \quad r \sin \theta = y$$

来达成。把上面转换方法代入熟悉的椭圆坐标方程式 $\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$ ，即得

$$\begin{aligned} b^2(r \cos \theta - c)^2 + a^2(r \sin \theta)^2 - a^2b^2 &= 0 \\ b^2(r^2 \cos^2 \theta - 2cr \cos \theta + c^2) + a^2r^2(1 - \cos^2 \theta) - a^2b^2 &= 0 \\ (b^2 - a^2)r^2 \cos^2 \theta - 2b^2cr \cos \theta + b^2c^2 + a^2r^2 - a^2b^2 &= 0 \\ -c^2r^2 \cos^2 \theta - 2b^2cr \cos \theta - b^4 + a^2r^2 &= 0 \quad (a^2 = b^2 + c^2) \\ a^2r^2 - (rc \cos \theta + b^2)^2 &= 0 \\ [ar - (rc \cos \theta + b^2)][ar + (rc \cos \theta + b^2)] &= 0 \end{aligned}$$

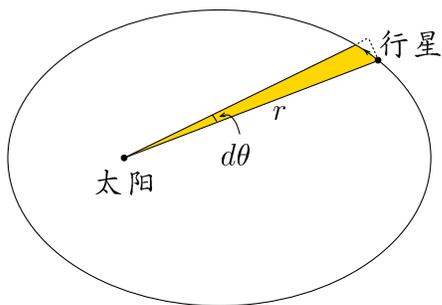
因此 $r = \frac{b^2}{a - c \cos \theta}$ ($r = \frac{-b^2}{a + c \cos \theta}$ 是负值，它会以 P' 来描绘出椭圆)

。为了方便以后的计算，我们取其颠倒式为椭圆极坐标方程式：

$$\frac{1}{r} = \frac{a - c \cos \theta}{b^2}$$

(三) 第二定律的数理分析

以太阳（焦点）为中心，极坐标 (r, θ) 表达行星位置。当 θ 增大到 $\theta + d\theta$ 时，太阳与行星的连线所扫过的面积为 $dA \approx \frac{1}{2}r^2 d\theta$ ，如 [图 8-23] 所示



[图 8-23]

运用第二定律，这个面积的改变速率为常数，即：

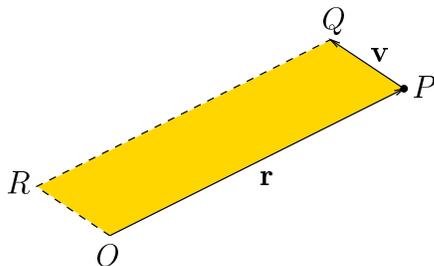
$$\frac{dA}{dt} = \frac{1}{2}r^2 \frac{d\theta}{dt} = \frac{1}{2}r^2 \omega = k, \quad (k: \text{常数}, \omega: \text{角速度})$$

$$\text{椭圆面积} = \pi ab = \int_0^T dA = \int_0^T k dt = kT \quad (T: \text{周期})$$

$$r^2 \omega = 2k = \frac{2\pi ab}{T}$$

注意在上面我们只是用了微积分的记号和想法，并没有用到深奥技巧。

[注]：从物理学观点来看，第二定律是有物理意义的。如 [图 8-24] 所示， \mathbf{r} 是动点 P 的位置向量， \mathbf{v} 是其速度向量：



[图 8-24]

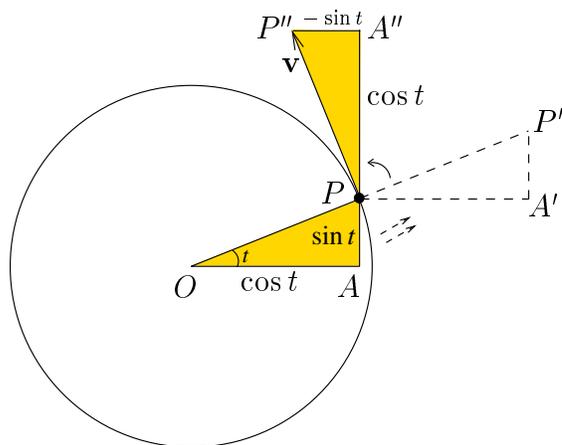
$|\mathbf{r} \times \mathbf{v}|$ 是平行四边形 $OPQR$ 的面积。但从物理学观点来看， $\mathbf{r} \times m\mathbf{v}$ 是物体相对于 O 点的角动量，因为是平面（椭圆）运动，此向量是恒垂直于平面的，所以由第二定律亦可得知角动量在行星运行中是不变的（这也是人类理性文明中首次接触到角动量守恒律）。再者，

$$\begin{aligned} \mathbf{0} &= \frac{d}{dt}(\mathbf{r} \times m\mathbf{v}) = \mathbf{v} \times m\mathbf{v} + \mathbf{r} \times m\mathbf{a} \\ &= \mathbf{r} \times \mathbf{F} \end{aligned}$$

所以引力 \mathbf{F} 的作用方向是和 \mathbf{r} 反向平行。

(四) 温习： $\cos \theta$ 和 $\sin \theta$ 的微分

从圆的简单几何性质和简单的物理观念，我们很容易便得出 $\cos \theta$ 和 $\sin \theta$ 的微分。其简单的推导如下：



[图 8-25]

如 [图 8-25] 所示，动点 P 在单位圆上作单位速率运动。用熟知的圆的参数表示方法， P 的坐标可写成 $(\cos t, \sin t)$ 。另一方面，从几何观点得知速度向量应是垂直于半径，所以把图中的 $\triangle OAP$ 平移至 $\triangle PA'P'$ 再旋转 $\frac{\pi}{2}$ 至 $\triangle PA''P''$ 。这样，速度向量 \mathbf{v} 在 \vec{Ox} , \vec{Oy} 的分量为：

$$v_x = -\sin t, \quad v_y = \cos t$$

但从物理学观点来说，速度向量 \mathbf{v} 正是：

$$\mathbf{v} = \left(\frac{d}{dt} \cos t \right) \mathbf{e}_x + \left(\frac{d}{dt} \sin t \right) \mathbf{e}_y$$

因此即得下面熟悉的公式：

$$\frac{d}{dt} \cos t = -\sin t, \quad \frac{d}{dt} \sin t = \cos t$$

(五) 向心加速的公式： $-(r\omega^2 + \frac{d^2r}{dt^2})$

上述公式只需对坐标 $x = r \cos \theta$, $y = r \sin \theta$ 直接微分便可得出。计算过程大致如下：

$$\begin{aligned} \begin{cases} x = r \cos \theta \\ y = r \sin \theta \end{cases} &\Rightarrow \begin{cases} \dot{x} = \dot{r} \cos \theta - r\omega \sin \theta \\ \dot{y} = \dot{r} \sin \theta + r\omega \cos \theta \end{cases} \quad (\omega = \dot{\theta}) \\ \Rightarrow \begin{cases} \ddot{x} = \ddot{r} \cos \theta - r\omega^2 \cos \theta - 2\dot{r}\omega \sin \theta - r\dot{\omega} \sin \theta \\ \ddot{y} = \ddot{r} \sin \theta - r\omega^2 \sin \theta + 2\dot{r}\omega \cos \theta + r\dot{\omega} \cos \theta \end{cases} \end{aligned}$$

因此

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

由(四)知引力的方向是平行于位置方向 $(\cos \theta, \sin \theta)$ ，所以 $(2\dot{r}\omega + r\dot{\omega})$ 应该是 0，并只余下向心加速 $-(r\omega^2 + \ddot{r})$ 。

(六) 平方反比定律的证明 (第一、第二定律的综合分析)

要证明引力是满足平方反比定律，我们只需验证 $r^2 \left(\frac{d^2r}{dt^2} - r\omega^2 \right)$ 是否为一常数。先对椭圆的极坐标方程式微分：

$$\begin{aligned} \frac{1}{r} &= \frac{a - c \cos \theta}{b^2} \\ -\frac{1}{r^2} \frac{dr}{dt} &= \frac{c}{b^2} \sin \theta \omega \\ \Rightarrow \frac{dr}{dt} &= -\frac{c}{b^2} \sin \theta r^2 \omega = -\frac{2\pi ac}{bT} \sin \theta \end{aligned}$$

在上式用了(三) $r^2 \omega = \frac{2\pi ab}{T}$ 。同样地在下面的计算中，我们尽可能分

离出 $r^2\omega$ 这一项，然後换成常数 $\frac{2\pi ab}{T}$ 。

$$\begin{aligned} \frac{d^2r}{dt^2} &= -\frac{2\pi ac}{bT} \cos\theta \omega \\ \Rightarrow r^2 \frac{d^2r}{dt^2} &= -\frac{2\pi ac}{bT} \cdot \frac{2\pi ab}{T} \cos\theta \\ &= -\frac{4\pi^2 a^2}{T^2} c \cos\theta \\ r^2(-r\omega^2) &= -r^3\omega^2 = -\frac{1}{r}(r^2\omega)^2 = -\frac{1}{r} \frac{4\pi^2 a^2 b^2}{T^2} \\ &= -\frac{4\pi^2 a^3}{T^2} + \frac{4\pi^2 a^2}{T^2} c \cos\theta \end{aligned}$$

因此 $r^2 \left[-r\omega^2 + \frac{d^2r}{dt^2} \right] = -\frac{4\pi^2 a^3}{T^2}$ 。

(七) 由行星引力到万有引力

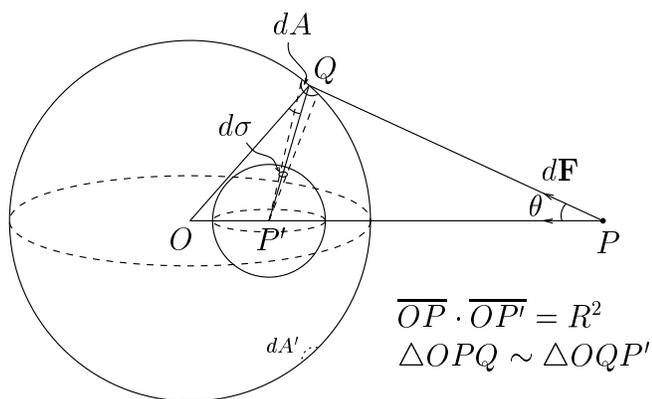
当牛顿想再进一步把行星与太阳之间的引力推广到任何物体与物体之间的引力，他遇到一个困难，使这位科学史上的巨人困扰了数年。由于行星与太阳之间的距离很大，所以在计算中可把行星和太阳当作两个质点，即可以假设质点集中了整个球体的质量；但当推广至任何物体与物体之间的情形，如苹果与地球，则便不可随便地把地球当作一个质点了。牛顿遇到的困难，就是他不能证明的确可以把地球当作为质点的猜想。即使在 1684 年他的好友 Halley (哈雷) 力邀牛顿发表已得的结果，他仍不愿意³发表。到了 1686 年，他终于成功地证明了上述猜想，即一个密度只随著到球心距离而变化的球体，在吸引球外一个质点时，所作用的力就像假设全部质量都集中在球心一样。在这年他写信给 Halley 表示同意写出他的工作，这就是在次年 (1687) 出版的科学巨著《自然哲学⁴的数学原理》(Philosophiae Naturalis Principia Mathematica)。

牛顿在书中所给的证明是很繁复的。在这里，我们给出另一个证明，它巧妙地运用了球的几何特性而大大简化了计算过程。

³其後牛顿在给 Halley 的信中提到，直至 1685 年，他还在怀疑这个猜想是错的。

⁴在当年，哲学其实是指科学。自然哲学亦即是物理学。

对于一个球面的最自然、最对称的点当然就是球心。但是在研讨球面与球外一点的互相作用时，从几何观点来看，最自然、最对称的点就不再是球心，而是 [图 8-26] 的 P' 点（这是 P 相对于球面的反射对称点）。

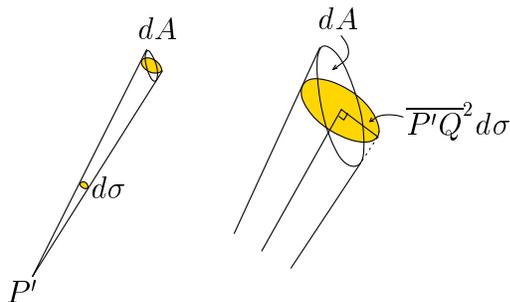


[图 8-26]

设球体其中一层薄壳的半径为 R ，面密度为 ρ ，薄壳质量为 $M = 4\pi R^2 \rho$ ，球外质点 P 的质量为 m 。考虑在薄壳上的一小片面积 dA 作用于 P 的引力 $d\mathbf{F}$ 。因薄壳对于 OP 是旋转对称， $d\mathbf{F}$ 垂直于 OP 的分量会被对称小片 dA' 所作用力抵消，所以只需考虑 $d\mathbf{F}$ 在 OP 方向的分量：

$$|d\mathbf{F}| \cos \theta = G \frac{\rho dA \cdot m}{PQ^2} \cos \theta$$

在直线段 \overline{OP} 上选 P' 使得 $\overline{OP} \cdot \overline{OP'} = R^2$ ，并以 P' 为心构造一个单位球面。令 P' 连向 dA 的射线在这单位球面上的影象为 $d\sigma$ 。



[图 8-27]

如 [图 8-27] 所示， dA 和 $d\sigma$ 之间有一个简单的关系：

$$dA \cdot \cos \theta = \overline{P'Q}^2 \cdot d\sigma$$

因此，整个薄壳作用于 P 的力就是

$$\begin{aligned} |\mathbf{F}| &= \sum |d\mathbf{F}| \cdot \cos \theta = \sum Gm\rho \frac{dA \cdot \cos \theta}{PQ^2} = \sum Gm\rho \frac{\overline{P'Q}^2 \cdot d\sigma}{PQ^2} \\ &= \sum Gm\rho \frac{R^2}{OP^2} d\sigma = Gm\rho \frac{R^2}{OP^2} \sum d\sigma \\ &= Gm\rho \frac{R^2}{OP^2} \cdot 4\pi = G \frac{Mm}{OP^2} \end{aligned}$$

因此这层薄壳作用于 P 的力就相等于将全部质量 M 集中于 O 而作用于 P 的力。再将所有薄壳作用的力加起来，便得所需之公式。

这也就是由 Kepler 行星运行定律的数理分析自然而然地推导出牛顿万有引力定律的一个简朴详尽的叙述。它就是牛顿的科学巨著 *Philosophiae Naturalis Principia Mathematica* 中所讨论的主要结果。它也自然是後学後进应当心领神会，并从此体会人类理性文明世代相承，继往开来的精要和精神。

8.9 圆锥截线例题，极与极线

【例一】：圆锥截线方程的矩阵表达式：在 §8.4 中我们用到以矩阵来表达圆锥截线的方程式

$$\Gamma: Ax^2 + 2Bxy + Cy^2 + 2Dx + 2Ey + F = 0$$

若引入矩阵记号

$$M = \begin{pmatrix} A & B & D \\ B & C & E \\ D & E & F \end{pmatrix}, \quad \mathbf{u} = \begin{pmatrix} x \\ y \\ 1 \end{pmatrix}$$

则 Γ 的方程可简洁地重新写成下面模式：

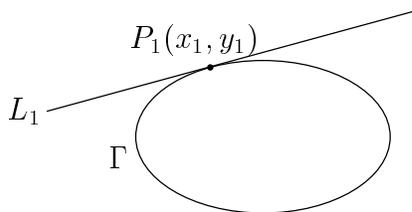
$$\Gamma: \mathbf{u}^T M \mathbf{u} = (x \ y \ 1) \begin{pmatrix} A & B & D \\ B & C & E \\ D & E & F \end{pmatrix} \begin{pmatrix} x \\ y \\ 1 \end{pmatrix} = 0$$

其中 \mathbf{u}^T 表示把矩阵的行列互相转置 (transpose), 这个运算满足下述简单公式:

$$\text{设 } A, B \text{ 为适当大小的矩阵, 则 } (A \cdot B)^T = B^T \cdot A^T$$

【例二】: 圆锥截线切线方程: 令 $\Gamma: Ax^2 + 2Bxy + Cy^2 + 2Dx + 2Ey + F = 0$ 为一圆锥截线, $P_1(x_1, y_1)$ 为 Γ 上任意的一点。在初等解析几何中我们熟知过 P_1 点的切线方程可以写成

$$L_1: Ax_1x + 2B\left(\frac{x_1y + y_1x}{2}\right) + Cy_1y + 2D\left(\frac{x + x_1}{2}\right) + 2E\left(\frac{y + y_1}{2}\right) + F = 0$$



[图 8-28]

这个公式的证明一般来说需要用到较为繁复的验算。现在我们尝试改用矩阵表达式和矩阵运算, 把其验算过程大大的简化如下:

证明: 不妨假设 Γ 是非蜕化的情况。若改用矩阵表达式, 则 Γ 和 L_1 可以分别写成

$$\Gamma: \mathbf{u}^T M \mathbf{u} = 0, \quad L_1: \mathbf{u}_1^T M \mathbf{u} = 0$$

其中 $\mathbf{u}_1^T = (x_1 \ y_1 \ 1)$ 。若 Γ 与 L_1 只交于 P_1 这一点, 则 L_1 即为 Γ 在 P_1 点的切线; 假若不然, 即 Γ 与 L_1 至少交于 P_1 和 P_2 两点:

$$P_1(x_1, y_1): \mathbf{u}_1 = \begin{pmatrix} x_1 \\ y_1 \\ 1 \end{pmatrix}, \quad P_2(x_2, y_2): \mathbf{u}_2 = \begin{pmatrix} x_2 \\ y_2 \\ 1 \end{pmatrix}$$

首先, 由所设 P_1, P_2 在 Γ 之上, 即有

$$\mathbf{u}_1^T M \mathbf{u}_1 = 0 \quad \text{及} \quad \mathbf{u}_2^T M \mathbf{u}_2 = 0$$

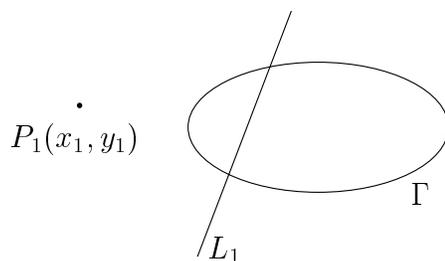
再者, 我们也假设了 P_2 在 L_1 之上, 即 $\mathbf{u}_1^T M \mathbf{u}_2 = 0$ 。

现于直线 $L_1 = P_1P_2$ 上任取一点 P ，则可选取适当的实数 λ 使得 $P = (\lambda x_1 + (1-\lambda)x_2, \lambda y_1 + (1-\lambda)y_2)$ 。考虑 P 点的矩阵表达式： $\mathbf{v} = \lambda \mathbf{u}_1 + (1-\lambda)\mathbf{u}_2$ ，则有

$$\begin{aligned} & \mathbf{v}^T M \mathbf{v} \\ &= (\lambda \mathbf{u}_1 + (1-\lambda)\mathbf{u}_2)^T M (\lambda \mathbf{u}_1 + (1-\lambda)\mathbf{u}_2) \\ &= \lambda^2 \mathbf{u}_1^T M \mathbf{u}_1 + \lambda(1-\lambda)\mathbf{u}_1^T M \mathbf{u}_2 + \lambda(1-\lambda)\mathbf{u}_2^T M \mathbf{u}_1 + (1-\lambda)^2 \mathbf{u}_2^T M \mathbf{u}_2 \\ &= 0 \end{aligned}$$

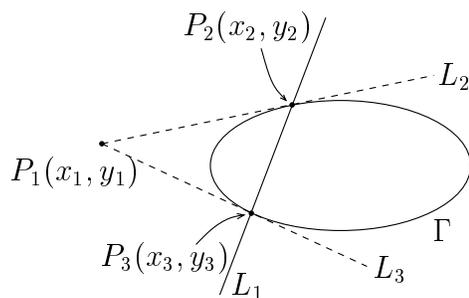
所以 P 点也同时在 Γ 之上，亦即 Γ 已包含整条直线 L_1 ，这自然是和 Γ 是非蜕化的假设互相矛盾。由此可见， Γ 与 L_1 只可能交于 P_1 这一点，亦即 L_1 乃是 Γ 在 P_1 的切线。

【例三】：当 $P_1 \in \Gamma$ 时，已知 $L_1: \mathbf{u}_1^T M \mathbf{u} = 0$ 乃是 Γ 在 P_1 的切线；但若 P_1 不在 Γ 之上（如 [图 8-29] 所示）时， L_1 应该代表著什么呢？



[图 8-29]

[解答] L_1 乃是由 P_1 到 Γ 的两条切线 L_2, L_3 的切点 P_2, P_3 所定的那条直线，如 [图 8-30] 所示。



[图 8-30]

证明：由前面的结果已知 L_2 和 L_3 的方程分别是

$$L_2: \mathbf{u}_2^T M \mathbf{u} = 0, \quad L_3: \mathbf{u}_3^T M \mathbf{u} = 0$$

再者，由于两者都过 $P_1(x_1, y_1)$ 点，即

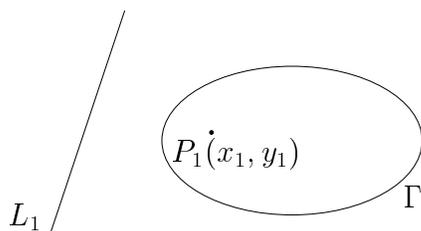
$$\mathbf{u}_2^T M \mathbf{u}_1 = 0 \quad \text{及} \quad \mathbf{u}_3^T M \mathbf{u}_1 = 0$$

经转置 (transpose) 後即有

$$\mathbf{u}_1^T M \mathbf{u}_2 = 0 \quad \text{及} \quad \mathbf{u}_1^T M \mathbf{u}_3 = 0$$

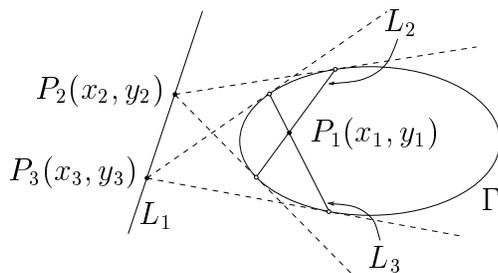
这就是说 P_2, P_3 在 $L_1: \mathbf{u}_1^T M \mathbf{u} = 0$ 之上。

【例四】：其实 [图 8-29] 并不是 $P_1 \notin \Gamma$ 的唯一情况，我们还须考虑如 [图 8-31] 的情况，即 $P_1(x_1, y_1)$ 在 Γ 之内：



[图 8-31]

[解答] L_1 乃是由过 P_1 点的弦（如 [图 8-32] 所示的 L_2, L_3 等），其两端点切线的交点（如 P_2, P_3 等）所组成的点集。



[图 8-32]

证明：由前面的讨论已知：

$$L_2 \text{ 的方程是: } \mathbf{u}_2^T M \mathbf{u} = 0$$

$$L_3 \text{ 的方程是: } \mathbf{u}_3^T M \mathbf{u} = 0$$

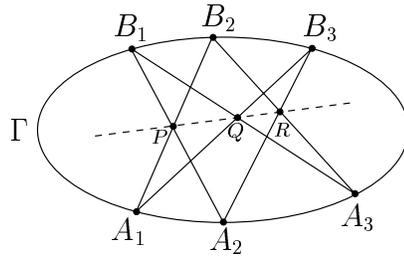
由于 L_2, L_3 都是过 P_1 点的弦，即

$$\mathbf{u}_2^T M \mathbf{u}_1 = 0 \quad \text{及} \quad \mathbf{u}_3^T M \mathbf{u}_1 = 0$$

亦即 $\mathbf{u}_1^T M \mathbf{u}_2 = 0$ 及 $\mathbf{u}_1^T M \mathbf{u}_3 = 0$ 。所以 P_2, P_3 皆在 L_1 之上。

[注]：由上面的例子我们可以看出 P 和 L 的密切（对偶）关系：给出 P 点，我们可以如前述例子般构作对应的 L ；反之，给出直线 L ，我们也可以把上面的程序倒过来构作 P 。在射影几何中， P 与 L 分别称为「极 (pole)」与「极线 (polar)」。

【例五】：Pascal 定理的纯代数证明：在 §8.6 中，我们运用了几何透视对应方法来证明 Pascal 定理，即在一个圆锥截线 Γ 上任取六点 $\{A_1, A_2, A_3; B_1, B_2, B_3\}$ ，如 [图 8-11] 所示，令 $P = A_1B_2 \cap A_2B_1$ ， $Q = A_1B_3 \cap A_3B_1$ ， $R = A_2B_3 \cap A_3B_2$ ，则恒有 $\{P, Q, R\}$ 三点共线。



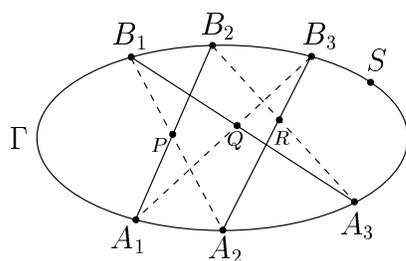
[图 8-11]

现在我们改用纯代数的方法来再次证明这个定理。这样做的好处一来是反映出代数理论在几何学上的不平凡应用；二来这种做法亦提供了一种途径来研究高维、高次的代数曲线。设

$$\begin{aligned} A_1B_2 : \ell_1 = 0, & \quad A_2B_1 : \ell'_1 = 0 \\ A_2B_3 : \ell_2 = 0, & \quad A_3B_2 : \ell'_2 = 0 \\ A_3B_1 : \ell_3 = 0, & \quad A_1B_3 : \ell'_3 = 0 \end{aligned}$$

令 $F(k) : \ell_1 \cdot \ell_2 \cdot \ell_3 + k \cdot \ell'_1 \cdot \ell'_2 \cdot \ell'_3 = 0$ ，则易见对于任给的实数 k ， $A_1, A_2, A_3, B_1, B_2, B_3$ 这六点必然在三次曲线 $F(k)$ 之上，而且 $F(k)$ 也包含了 P, Q, R 这三点。不妨假设 Γ 为非蜕化圆锥截线。现于 Γ 上取第七点 S （但不在已给的六点组 $\{A_1, A_2, A_3; B_1, B_2, B_3\}$ 之内），然後选取适当的 $k = k_0$ 使得 S 同时也在 $F(k_0)$ 之上，亦即 Γ 与 $F(k_0)$ 相交于至少七点：

$$A_1, A_2, A_3, B_1, B_2, B_3, S \in \Gamma \cap F(k_0)$$



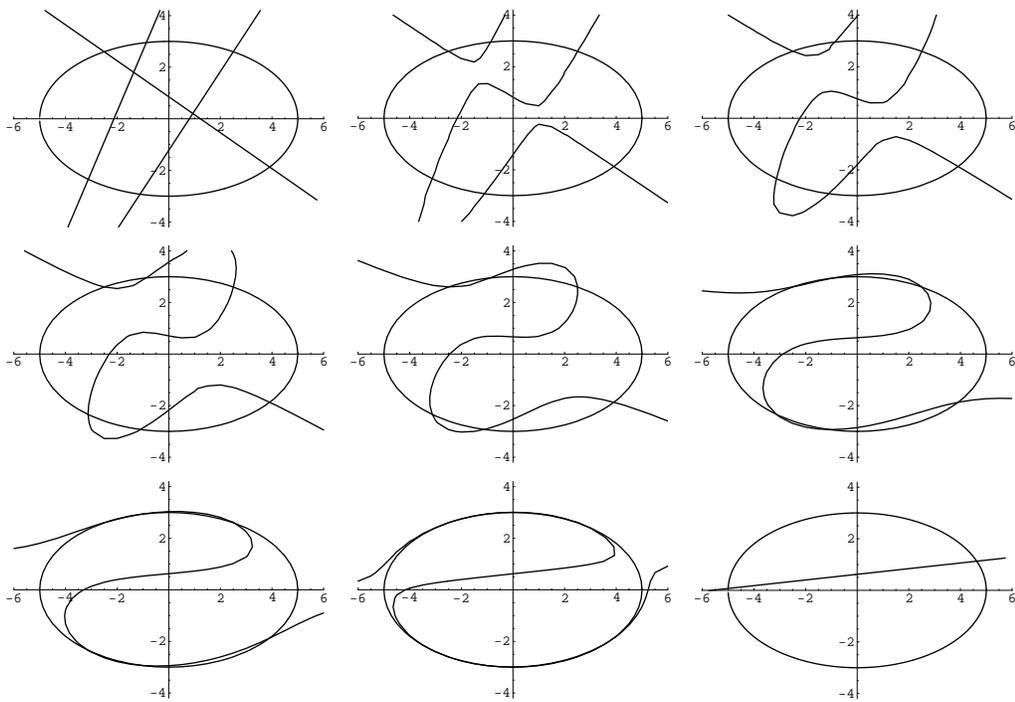
[图 8-33]

因为 Γ 是一个非蜕化的二次曲线, $F(k_0)$ 是一个三次曲线, 一般来说它们最多只有六个交点, 除非 Γ 本身就是 $F(k_0)$ 的一个因式。由于我们的构造使它们两者至少有七个交点, 因此 Γ 必然可以整除 $F(k_0)$, 即

$$F(k_0) : (Ax^2 + 2Bxy + Cy^2 + 2Dx + 2Ey + F) \cdot (Lx + My + N) = 0$$

显然 P, Q, R 均不在 Γ 之上, 所以它们必然全都在直线 $Lx + My + N = 0$ 之上, Pascal 定理证毕。

当 k 由 $k=0$ 渐渐趋向 $k=k_0$, $F(k)$ 的变化就如 [图 8-34] 所示:



[图 8-34]

基础数学讲义之四

《基础分析学之一》

单元微积分学

项武义

香港科技大学数学系

目录

引子	v
导论	ix
0.1 自然数系	x
0.2 整数数系	xiv
0.3 有理数系	xvi
0.4 实数系	xviii
0.5 复数系	xix
一 实数系和函数的连续性	1
1.1 实数系的连续性	1
1.2 连续函数的基本概念	4
1.3 多项式函数	11
1.3.1 多项式的唯一性定理与插值公式	11
1.3.2 单元多项式的除法与辗转相除求公因式	15
1.3.3 Sturm 定理	17
1.3.4 代数基本定理	19
二 微积分	29
2.1 变率与微分	29
2.2 总和与积分	38
2.3 微积分基本定理与均值定理	43
三 指数及对数函数	59
3.1 指数、对数函数的定义与基本性质	59

3.2	指数函数与对数函数的微分	64
3.3	自然对数表的计算法	67
3.4	复变数指数函数和三角函数	69
3.5	复利与指数函数	71
四	初等函数及其应用举例	75
4.1	多项式函数	75
4.1.1	n -阶密切多项曲线	76
4.1.2	高阶局部逼近与不定式之极限	79
4.1.3	插值问题的推广	80
4.2	三角函数与反三角函数	81
4.2.1	圆的对称性与正弦、余弦函数的基本性质	82
4.2.2	三角定律与极坐标	84
4.2.3	等速圆周运动与正弦、余弦的微分	85
4.2.4	等周问题 (Isoperimetric Problem)	87
4.2.5	Kepler 行星运行三定律及其数理分析	91
4.2.6	三角函数的积分计算	96
4.2.7	反三角函数及 π 的近似值计算	97
4.3	常系数常微分方程	101
4.3.1	算子符号	101
4.3.2	$p(D) = (D - \lambda)^k$ 的情形, $\lambda \in \mathbb{C}$	102
4.3.3	$p(D)$ 是一般的情形	104
4.3.4	$p(D)y = g(x)$ 的解法	106
五	欧氏几何、球面几何和非欧几何的统一理论	109
5.1	非欧几何的发现过程及其历史意义	110
5.2	发现非欧几何学的思路与突破点	112
5.3	欧氏、球面与非欧三角定律的统一理论	115
5.4	旋转面的解析几何	122
5.5	旋转面的 Gauss 曲率和 Gauss-Bonnet 公式	135
5.6	结语	143
5.7	思考题与习题	145
	结语	151

引子

人类生存的地球，比之于无限的宇宙实乃无比的渺小；而自古至今人类的理性文明只不过数千年，比之于地球上几十亿年的生命历史又是无比的短暂。但是人类得天独厚，具有几十亿年逐步进化而得——奇妙超群的脑力；再者，人类还能善用天赋的脑力，创造语言与文字，使得全人类的聪明才智不但能够群策群力、集思广益，而且还能世代相承、精益求精，创造了博大精深的理性文明。作为一个现代人，我们不但具有天赋的脑力，而且还承继了数千年来全人类的聪明才智，世代相承探讨研究的成果，总称之为理性文明 (Civilization of Rational Mind)。它使得每一位肯学、肯想的人生，大大地拓展了其能够理解的时空。

综观自古至今的理性文明，历代的先智先贤用来探索宇宙的基本思想和方法其实是既自然又朴实的，大体上可以简述如下：

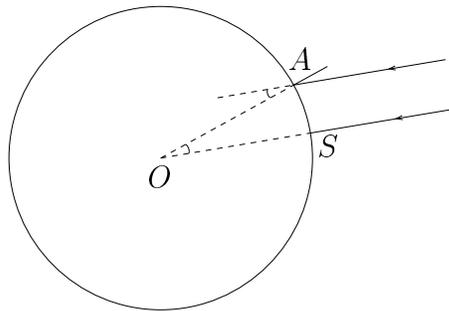
先用归纳法，师法自然，实事求是。例如观测天象，纪录昼夜之长短，四季的变化等等。然後再用解析思维把实验探索所得的各种各样认知加以综合、演绎，逐步由定性到定量，由表象到内在结构，精益求精地研讨大自然中万象万物的基本结构和基本原理。

总之，实验归纳和综合演绎乃是两种相辅相成交互进展的科学方法，而数理分析 (Mathematical Analysis) 则是後者的基本方法和有力工具，分析学的基础理论也就是如此产生的。我们将在往後章节中简明扼要地讨论数理分析在人类理性文明中产生的过程 and 其所扮演的角色。

在我们开始逐章逐节作系统研讨之前，且以一个既清晰简朴又具有历史意义的范例，来对于数理分析的本质和要点作一初步解说。那就

是远在纪元前三世纪 Eratosthenes (约 276–194 B.C.) 所达成者：人类对于地球大小的首次定量估算。

话说当年，Eratosthenes 乃是古希腊文明後期中心所在地 Alexandria 的图书馆馆长，他是一位精通当时的天文地理的博学之士。首先，当年业已由月蚀，航海的观察得知我们所居住的大地乃是一个大球；而且太阳和地球之间的距离要比地球和月亮之间的距离还要大许多倍。总之，同时照射到地球上各地的太阳光基本上就是平行光线。再者，他还知得人们在 Alexandria 正南方尼罗河出山口的 Syene (现今的 Aswan) 的一口深井，观察到夏至正午的太阳光可以直射井底，不留任何井壁的阴影。换言之，在 Syene 夏至正午的太阳光乃是垂直于该井的水平面者也。但是在 Alexandria 的夏至正午的阳光，则和井壁大约有 $\frac{1}{50}$ 周角 (亦即 7.2 度) 的夹角。当年 Eratosthenes 大概就用如下图所示的图解，把上述几个天文地理的知识加以综合分析。



他以一个圆表示过 Alexandria 和 Syene 的经圆 (longitudinal circle)，以两条平行线表达分别照射在 Alexandria 和 Syene 的夏至正午的阳光。从上述简明扼要的抽象表述，就可以用熟知的几何知识推论以下几点：其一是过 Syene 的光线垂直于球面，所以其延长线乃是过球心 O 者也。其二是由平行线和 Alexandria-球心联线的内错角相等得知 $\angle AOS$ 也等于 $\frac{1}{50}$ 周角。由此可得整个经圆的周长大约是 Alexandria 和 Syene 之间的距离的 50 倍！

Eratosthenes 这位老先生并没有去实地测量上述两地之间的距离。他只是去市集向往来于 Alexandria 和 Syene 之间的骆驼商队，询问他们一共要走几天才能由 Syene 走到 Alexandria，而且每天大约能走多少希腊里 (Stadia)。他就将询问所得的天数 (50 天) 和每天的里程 (100 Stadia)

得出下列经圆圆周长度的估计，亦即

$$50 \times 50 \times 100 = 250,000 \quad (\text{Stadia})$$

这就是人类对于其所生存的地球大小的第一次定量估计。

假如要把它和现代测算的地球大小作个比较，当然就要把希腊里 (Stadia) 和公里 (kilometer) 的长度作一换算。在这里有一个小小的疑案：Stadia 者，乃是当年运动场 (Stadium) 的跑道的长度是也。但是目下遗留的古希腊运动场的跑道却有长短不一的长度。总之，以较长者来换算则其估计比实测要大些，而以较短者来换算则其估计比实测要小些，而相去都在 10% 之内。

如今回顾 Eratosthenes 在两千多年前，能够把当时所知道的天文、地理和几何知识，用简朴的图解加以综合分析，一举而得数量上相当准确的地球大小的估计，它实在是人类以数理分析去理解大自然的一个杰出的典范。世世代代理性文明的继承者，不但应该以怀古之情瞻仰它，而且更要从中学习其用法，获得数理分析的启蒙启发。

导论——数系和数系的扩张

数理分析是我们由表及里，定量地深入研讨大自然的重要方法，而分析学 (Analysis) 也就是为了有效运用数理分析去理解大自然而发展起来的一门数学。当所要研讨的事物还是相当简单初等者，其相应的各种各样定量问题往往是初等的代数问题和几何问题，例如复利的计算、平面测量等等。但是当所要研讨的问题逐渐深入，逐步拓展，就自然而然地超出了初等代数和初等几何力所能及的范畴，例如天体的运行、曲线曲面的研究、弦的振动等等。其实，分析学乃是以初等代数和初等几何为基础，把两者结合起来，更上层楼而发展起来的；它是对于各种各样动态事物作数理分析的重要工具，而其本身的基础理论则在于函数的微分、积分和连续性的明确及其基本定理。这也就是为什么通常把分析学基础论叫做微积分的原由。

代数和几何的基础，归根研底乃是建筑在数系和空间这两种基本结构之上。空间乃是我们和宇宙万物共存于其中者，是大自然所赋予的；而数系则是人类理性文明为了更加精确地定量研讨事物所构造的「计量」体系，可以说是人的创造。在定量几何学定量地研讨空间本质的讨论之中，前者和后者则自然而然地结合在一起，相辅相成。

最原始的数系就是我们用来数个数 (counting) 的自然数系 (system of natural numbers)，然後逐步扩充而得整数系 (system of integers)、有理数系 (system of rational numbers)、实数系 (system of real numbers) 和复数系 (system of complex numbers)。通常分别以 \mathbb{N} 、 \mathbb{Z} 、 \mathbb{Q} 、 \mathbb{R} 、 \mathbb{C} 表示上述数系，则上述逐步扩充 (张) (extensions) 就可以用下述符号

$$\mathbb{N} \subset \mathbb{Z} \subset \mathbb{Q} \subset \mathbb{R} \subset \mathbb{C}$$

简约地表达之。

我们首先将对上述逐步扩张而构造的一连串数系作一次归本究源的结构分析。唯有对数系结构的来龙去脉了然于心，才能在运用它们作各种各样定量分析、探索自然时得心应手。

0.1 自然数系

自然数系是人类为了数个数 (counting) 这样一种原始而且基本的「定量化」而创造的体系。例如有一位牧羊人要知道其羊群的个数，或当古人发现月亮的圆缺变化是一种周而复始的事情，自然就想统计一下其周期的天数等等。虽然各古文明所用的符号和体系不同，但是其本质都是一串逐一相连的符号体系，例如

$$\begin{array}{ccccccc} | & || & ||| & |||| & \dots & & \\ \text{—} & \text{—} & \text{—} & \text{—} & \dots & & \\ 1 & 2 & 3 & 4 & \dots & & \end{array}$$

其中第一个符号表示「单元」，是一只羊，一个人或一棵树的抽象化，而後继符号则表示比前述所表达者再增添多一个单元，亦即

$$2 = 1 + 1, 3 = 2 + 1, 4 = 3 + 1, 5 = 4 + 1, \dots$$

$$101 = 100 + 1, 102 = 101 + 1, \dots, \text{如此类推}$$

由此可见，自然数系最为原始、基本的结构就是「+1」运算。在自然数系这一串顺序排列的符号体系中，後继者就是前者「+1」之所得，所以任何一个自然数都可以由 1 起始，逐步「+1」而得之（其实这就是我们构造自然数系的方法）。

把上述事实改用「数学化」的集合用语来描述，即为：

【数学归纳法原理】(Principle of Mathematical Induction)：一个自然数系的子集 $S \subseteq \mathbb{N}$ 若满足下述两个条件

$$(0.1) \quad (i) 1 \in S, \quad (ii) n \in S \Rightarrow n + 1 \in S$$

则 S 其实就等于 \mathbb{N} 。

自然数系之所以有用、好用是因为它具有满足交换律、结合律和分配律的加法和乘法运算，这都是大家所熟知常用者。假如有人问你或

者你自问：「为什麼那些运算律总是成立的呢？」此问显然不能用：「实例的计算从来没碰到不成立的情形，所以应该是成立的」，或者是人云亦云地：「大家都说它们总是成立，所以我也相信它们是对的」等等作为解释，因为自然数的个数是无限，而实例计算所能验算者只是很小、很小的一部分，所以加、乘运算的运算律的普遍成立是需要加以论证的！这也就是我们所要讨论的首个课题。此事当然得从加法和乘法的根源（亦即本质）说起。

加法的本质乃是「+1」运算的复合。例如「+2」就是把「+1」做两次（符号“2”所代表的意义）的结果，「+3」就是把「+1」做三次的结果，「+(n+1)」就是比「+n」再多做一次「+1」者也。由此可见

$$(0.2) \quad a + (n + 1) = (a + n) + 1$$

其实就是加法的归纳定义式。再者下述的加法结合律所描述者：

$$(0.3) \quad (a + m) + n = a + (m + n)$$

就是对 a 先作 m 次「+1」然後接著再作 n 次「+1」，其结果也就是对 a 作 $(m+n)$ 次「+1」，由此可见加法结合律的普遍成立是直观明显的。把它改用数学化的用语与格式说清楚，则是下述归纳论证 (proof by induction)。

【定理 0.1】（加法结合律）： $(a + m) + n = a + (m + n)$ 。

证明：对于任给 a, m ，归纳证明上式对于所有自然数 n 皆成立。

(i) $n = 1$ 时， $(a + m) + 1 = a + (m + 1)$ 就是加法的归纳定义式。

(ii) 由归纳假设 $(a + m) + n = a + (m + n)$ ，则有

$$\begin{aligned} (a + m) + (n + 1) &= [(a + m) + n] + 1 && \text{[加法归纳定义式]} \\ &= [a + (m + n)] + 1 && \text{[归纳假设]} \\ &= a + [(m + n) + 1] && \text{[加法归纳定义式]} \\ &= a + [m + (n + 1)] && \text{[加法归纳定义式]} \quad \square \end{aligned}$$

接著让我们再用归纳法来证明加法交换律。

【定理 0.2】（加法交换律）： $a + b = b + a$ 。

证明：先用归纳法证明 (i) $a + 1 = 1 + a$ ：

(1) $a = 1$ 时， $1 + 1 = 1 + 1$ 是显然的。

(2) 由归纳假设 $a + 1 = 1 + a$ ，则有

$$\begin{aligned} (a + 1) + 1 &= (1 + a) + 1 && \text{[归纳假设]} \\ &= 1 + (a + 1) && \text{[结合律]} \end{aligned}$$

然後再由归纳假设 $a + b = b + a$ 证明

$$(ii) \quad a + (b + 1) = (b + 1) + a$$

其证明如下：

$$\begin{aligned} a + (b + 1) &= (a + b) + 1 && \text{[结合律]} \\ &= (b + a) + 1 && \text{[归纳假设]} \\ &= b + (a + 1) && \text{[结合律]} \\ &= b + (1 + a) = (b + 1) + a && \square \end{aligned}$$

现在让我们来分析一下乘法的本质。乘法其实是自相加的缩写：

「 $m \cdot a$ 」就是 m 个 a 自相加的总和（所以 $1 \cdot a = a$ ）

因此 $(m + 1) \cdot a$ 所表达者就是比 $m \cdot a$ 再多加一个 a 。由此可见，乘法的归纳定义式就是

$$(0.4) \quad 1 \cdot a = a, \quad (m + 1) \cdot a = m \cdot a + a$$

再者，乘法的左分配律：

$$(0.5) \quad m \cdot a + n \cdot a = (m + n) \cdot a$$

就是说把 m 个 a 自相加的「总和」和 n 个 a 自相加的「总和」再加起来，其实就等于 $(m + n)$ 个 a 自相加的总和。此事亦为直观上极为明显者，下面我们给它作一次归纳的证明。

[注意]：在尚未证明乘法交换律之前，分配律是有左、右之分别的。其实，乘法交换律的证明是要在左、右分配律都证得之後才能证得者。

【定理 0.3】（乘法的左分配律）： $m \cdot a + n \cdot a = (m + n) \cdot a$ 。

证明：对于任给 m, a ，归纳证明上式对于所有自然数 n 皆成立。

(i) $n = 1$ 时，由定义式知 $m \cdot a + 1 \cdot a = (m + 1) \cdot a$ 。

(ii) 由归纳假设 $m \cdot a + n \cdot a = (m + n) \cdot a$ ，则有

$$\begin{aligned} m \cdot a + (n + 1) \cdot a &= m \cdot a + (n \cdot a + a) && \text{[乘法归纳定义式]} \\ &= (m \cdot a + n \cdot a) + a && \text{[加法结合律]} \\ &= (m + n) \cdot a + a && \text{[归纳假设]} \\ &= [(m + n) + 1] \cdot a && \text{[乘法归纳定义式]} \\ &= [m + (n + 1)] \cdot a && \square \end{aligned}$$

【定理 0.4】（乘法的右分配律）： $m \cdot (a + b) = m \cdot a + m \cdot b$ 。

证明：任取 a, b ，对 m 作归纳证明如下：

(i) $m = 1$ 时是显然成立的。

(ii) 由归纳假设 $m \cdot (a + b) = m \cdot a + m \cdot b$ 成立，则有

$$\begin{aligned} (m + 1) \cdot (a + b) &= m \cdot (a + b) + (a + b) && \text{[乘法归纳定义式]} \\ &= (m \cdot a + m \cdot b) + (a + b) && \text{[归纳假设]} \\ &= (m \cdot a + a) + (m \cdot b + b) && \text{[加法结合律、交换律]} \\ &= (m + 1) \cdot a + (m + 1) \cdot b && \square \end{aligned}$$

【定理 0.5】（乘法交换律）： $a \cdot b = b \cdot a$ 。

证明：留作习题。

【定理 0.6】（乘法结合律）： $(a \cdot b) \cdot c = a \cdot (b \cdot c)$ 。

证明：留作习题。

指数符号：在本质上，乘方乃是自相乘的缩写。我们常用的指数符号，其归纳定义式也就是：

$$(0.6) \quad a^1 = a, \quad a^{(m+1)} = a^m \cdot a$$

【习题】：

(1) 试证指数定则 $a^m \cdot a^n = a^{m+n}$ 。

(2) 试证指数定则 $(a \cdot b)^m = a^m \cdot b^m$ 。

(3) 试证指数定则 $(a^m)^n = a^{n \cdot m}$ 。

0.2 整数数系

在自然数系内，我们定义减法为加法的逆算： $(a-b)$ 就是加上 b 之後就等于 a 的那个唯一的数，亦即方程式

$$(0.7) \quad x + b = a$$

的唯一解就是 $(a-b)$ 。但是在自然数系中，上述方程式只有在 a 大于 b 时才能「有解」！在此，一个自然而然的想法是：上式在 $a \leq b$ 时在自然数系中「无解」乃是由于自然数系的范畴太狭窄了，因此可以把它作适当的扩充来消除这种不理想的缺陷。换句话说，我们由自然数去构造一些新的数，使得在扩张後的数系中 $x + b = a$ 总是有唯一解。例如设

$$(0.8) \quad x + 1 = 1$$

的解是一个新的数「0」，则由下述归纳论据

$$(0.9) \quad 0 + n = n \Rightarrow 0 + (n + 1) = (0 + n) + 1 = n + 1$$

可见这个新数「0」应当满足 $0 + a = a$ (a 为自然数)。再者，我们也需要引进

$$(0.10) \quad x + a = 0$$

的解，以符号「 $-a$ 」表之。由此易见在 $a > b$ 时， $x + b = a$ 的解是一个自然数 $(a-b)$ ，而在 $a < b$ 时 $x + b = a$ 的解就应该是那个新数 $-(b-a)$ ，因为理应有

$$(0.11) \quad \begin{aligned} -(b-a) + b &= -(b-a) + [(b-a) + a] \\ &= [- (b-a) + (b-a)] + a \\ &= 0 + a = a \end{aligned}$$

总结上述简短分析，可见一个包含自然数系的扩张数系，若要满足同样的加乘运算律，而且使得减法通行无阻，亦即 $x + b = a$ 对于任给 a, b 恒有解，则它至少包含一个「0」和每个自然数 a 的负数「 $-a$ 」。这样一个数系就是我们常用、好用的整数系：

$$(0.12) \quad \mathbb{Z} = \mathbb{N} \cup \{0\} \cup \{-a, a \in \mathbb{N}\}$$

当然，我们还得要适当地定义整数系中的加、乘和指数运算。在这个关键性的步骤上，定义的适当性的检验准则是在扩张的数系中依然保有各个运算律，因为那些运算律乃是数系有用、好用的根本所基。其实上述保有运算律的准则业已唯一地确定了整数系中加、乘、指数运算的适当定义的必然性。兹分析如下：

【分析】：

(1) $0 + 0 = 0, 0 + a = a, (-a) + a = 0$ 乃是「0」和「 $-a$ 」的定义式。

(2) 对于任给自然数 a, b ,

$$\begin{cases} (-a) + b = b + (-a) = b - a \\ (-a) + b = b + (-a) = -(a - b) \end{cases} \quad \text{若} \quad \begin{cases} b > a \\ a > b \end{cases}$$

上述定义之必然性：

若 $b > a$ ，则

$$\begin{aligned} (-a) + b &= (-a) + [a + (b - a)] \\ &= [(-a) + a] + (b - a) \\ &= 0 + (b - a) = b - a \end{aligned}$$

若 $a > b$ 则

$$\begin{aligned} [(-a) + b] + (a - b) &= (-a) + [b + (a - b)] \\ &= (-a) + a = 0 \end{aligned}$$

所以 $(-a) + b = -(a - b)$ 。

(3) 对于任给自然数 a ，皆有

$$0 \cdot a = a \cdot 0 = 0, \quad (-1) \cdot a = a \cdot (-1) = -a$$

上述定义之必然性：

$$\begin{aligned} 0 \cdot a &= (0 + 0) \cdot a = 0 \cdot a + 0 \cdot a \Rightarrow 0 \cdot a = 0 \\ 0 &= 0 \cdot a = [1 + (-1)] \cdot a = 1 \cdot a + (-1) \cdot a = a + (-1) \cdot a \\ &\Rightarrow (-1) \cdot a = -a \end{aligned}$$

[特例： $(-1) \cdot 1 = 1 \cdot (-1) = -1$]

$$(4) (-1) \cdot (-1) = 1$$

定义之必然性：

$$\begin{aligned} 0 &= (-1) \cdot 0 = (-1) \cdot [(-1) + 1] = (-1) \cdot (-1) + (-1) \cdot 1 \\ &= (-1) \cdot (-1) + (-1) \Rightarrow (-1) \cdot (-1) = 1 \end{aligned}$$

$$(5) \text{ 符号定则：} (-a) \cdot b = -a \cdot b, (-a)(-b) = a \cdot b$$

定义之必然性：

$$\begin{aligned} (-a) \cdot b &= [(-1) \cdot a]b = (-1) \cdot (a \cdot b) = -a \cdot b \\ (-a) \cdot (-b) &= [(-1) \cdot a][(-1) \cdot b] = [(-1)(-1)](a \cdot b) \\ &= 1 \cdot (a \cdot b) = a \cdot b \end{aligned}$$

总结上面几点分析，可见我们业已熟习的整数系中的正、负数和 0 之间的计算法则，其实乃是能使得整数系的运算依然保持原先的那一套运算律的「唯一可能定义法」（必然性）。当然，我们还得去验证这种唯一的可能定义法是真的能够保有各个运算律的！这一点是很好的习题题材。

【习题】：逐一证明上述所定义的整数系的加乘运算满足加、乘的交换律、结合律和分配律。

0.3 有理数系

除法乃是乘法的逆算： $(a \div b)$ 就是那个乘以 b 後等于 a 的唯一解 ($b \neq 0$)，亦即方程式

$$x \cdot b = a, \quad (b \neq 0)$$

的唯一解定义为 $(a \div b)$ 。但是在整数系中，上述方程式只有在 b 是 a 的因数（亦即 a 是 b 的倍数）的情形才能「有解」。由此可见，整数系还得要加以适当的扩充才能消除这种不理想的「缺陷」。这也就是由整数系到有理数系的扩张。

【定义】：对于任给整数 a 和 b ($b \neq 0$) 定义有理数「 $\frac{a}{b}$ 」为方程式

$$x \cdot b = a$$

的唯一解；易见它也是 $x \cdot bk = ak$ 的解，所以 $\frac{a}{b} = \frac{ak}{bk}$ 。

若 $ad = cb$ ($b, d \neq 0$) 则 $\frac{a}{b} = \frac{ad}{bd} = \frac{cb}{db} = \frac{c}{d}$ ，反之若 $\frac{a}{b} = \frac{c}{d}$ ，亦即 $x \cdot b = a$ 和 $x \cdot d = c$ 具有同解，所以 $ad = x \cdot bd = (xd)b = cb$ 。由此可见「 $\frac{a}{b} = \frac{c}{d}$ 的充要条件」就是 $ad = cb$ 。

接著让我们来分析一下有理数之间的加法和乘法应该如何定义才能确保有理数系的加乘运算依然保有交换、结合和分配律。

$$(i) \quad \left(\frac{a}{b} + \frac{a'}{b'}\right) \cdot bb' = \left(\frac{a}{b}\right) \cdot bb' + \left(\frac{a'}{b'}\right) b'b \\ = ab' + a'b$$

所以应该如下定义加法：

$$\frac{a}{b} + \frac{a'}{b'} = \frac{ab' + a'b}{bb'} \quad (b, b' \neq 0)$$

$$(ii) \quad \left(\frac{a}{b} \cdot \frac{a'}{b'}\right) bb' = \left[\left(\frac{a}{b}\right) \cdot b\right] \left[\left(\frac{a'}{b'}\right) \cdot b'\right] = aa'$$

所以应该如下定义乘法：

$$\left(\frac{a}{b}\right) \left(\frac{a'}{b'}\right) = \frac{aa'}{bb'} \quad (b, b' \neq 0)$$

采用上述有理数的加法、乘法定义式和两个有理数的相等检验条件，就容易把有理系的加乘运算律归于整数系的加乘运算律来加以逐一验证。[此事留作习题]

指数定义的推广：设 a, b 是非零整数， n 为正整数，则 $\left(\frac{a}{b}\right)^0$ 定义为 1， $\left(\frac{a}{b}\right)^{-n}$ 定义为 $\left(\frac{b}{a}\right)^n$ 。

上述定义是使得指数定则在指数等于整数时恒成立的唯一定义法，
 即有

$$\begin{aligned} \left(\frac{a}{b}\right)^m \cdot \left(\frac{a}{b}\right)^n &= \left(\frac{a}{b}\right)^{m+n}, \quad m, n \in \mathbb{Z} \\ \left[\left(\frac{a}{b}\right)^m\right]^n &= \left(\frac{a}{b}\right)^{m \cdot n}, \quad m, n \in \mathbb{Z} \\ \left[\left(\frac{a}{b}\right)\left(\frac{c}{d}\right)\right]^m &= \left(\frac{a}{b}\right)^m \cdot \left(\frac{c}{d}\right)^m, \quad m \in \mathbb{Z} \end{aligned}$$

[注]：以有理数作为指数的定义需要用到实数系的连续性，我们将会
 在往後的章节系统地研讨其本质。

0.4 实数系

在概念上和研讨方法上，由有理数系到实数系是一个大幅度的跃进。
 在人类理性文明发展史上，这个跃进发生于纪元前五、四世纪古希腊
 几何学家在定量几何基础理论的深入研究中，由长度度量而产生的可
 公度性的问题。在这里我们把 Eudoxus 所创的逼近原理和逼近法，
 改用现代通用的术语，把从有理数系到实数系的扩张的精要之点，作
 一简朴明确的叙述。（参看基础几何学之一的 §2.4，这是理性文明一个
 重大篇章，是引人入胜、启发人深思的一段史话。）

1. 不可公度直线段的发现 (Hippasus)，事实胜于雄辩地证明了简单、
 初等的有理系是不足以表达任给两个直线段之间的比值的；足以用
 来表达、研讨长度的度量的数系肯定要包含许许多多那些不可公度的
 长度比值的非比实数 (irrational numbers)。总之，它是一个比有理数系
 更大的数系，我们称之为实数系。

2. Eudoxus 的比较原则和逼近原理明确了实数系和有理数系之间的
 关系，以逼近法用有理数逼近非比数，从而提供了研讨实数系的有效
 途径。

3. 在当年 Eudoxus 所要研讨的长度度量问题上，比值 $a:b$ 是「原
 给者」，对于每个 n ，他证明存在由 $a:b$ 而定的 m 使得

$$r_n = \frac{m}{n} < a:b < \frac{m+1}{n} = s_n$$

从而构造得 $a:b$ 的一对左、右夹逼数列 $\{r_n\}$ 和 $\{s_n\}$ ，所以根本没有「存不存在一个被它们所左、右夹逼的实数？」这种存在性问题，因为它就是原给的「 $a:b$ 」！但是在近代数学中，特别是分析学，我们会用到以种种方式构造而得的数列，因而讨论它们的极限的存在性问题就自然而然地成为很有其必要了。归根研底，在近代数学中各种各样的存在性定理之所本就是被上述左、右夹逼数列所夹逼者的存在性！

总结上面三点，实数系的发现和理解都和长度度量问题密切相关，而实数系中任何一对左、右夹逼数列都存在有（唯一）一个被它们所夹逼的实数（亦即它们的共同极限）则是直线连续不断的解析描述，称之为实数系的连续性，而它又是近代数学中各种各样存在性定理（例如 Sturm 定理、代数基本定理等等）的证明之所据。再者，实数系是有理数系的一种自然扩张，任何一个实数都能用一对左、右夹逼的有理数数列去唯一地描述它；反之，任给一对左、右夹逼的有理数数列也都描述著一个实数。这样不但简明扼要地刻划了有理数系和实数系之间的关系，而且也提供了用有理数系去研讨实数系的有效途径和方法。

0.5 复数系

在数系的逐步扩张中，不论在概念的跳跃上和所涉及的方法论和技术性上，从有理数系到实数系这一步都是跨得最大也是所涉最为艰巨者。所幸者，实数系和直线段长度度量问题紧密相关，所以其直观性极强，而且是定量几何学上必须充分理解的基础。这也就是为什麼实数系的理解在 Eudoxus 重建定量几何基础理论时业已大体完成。

相比起来，由实数系到复数系这一步实在要简单得多。本质上只是引进一个平方为 -1 的「虚数」单位 i (unit of imaginary number)，然後用简单自然的代数公式定义所有能够表示成 $a + bi$, $a, b \in \mathbb{R}$ 的复数 (complex numbers) 之间的加法和乘法，即

$$(a + bi) + (c + di) = (a + c) + (b + d)i$$
$$(a + bi) \cdot (c + di) = (ac - bd) + (ad + bc)i$$

很容易用直接验算证明这种扩张的数系 \mathbb{C} 依然满足加、乘的交换律，结合律和分配律。再者，当 a, b 不全为零时

$$(a + bi) \cdot \frac{a - bi}{a^2 + b^2} = 1$$

所以复数系 \mathbb{C} 在代数方面具有和有理数系 \mathbb{Q} 、实数系 \mathbb{R} 同样易算好用的运算律。

【历史的注记】：复数的引入，起始于解一元二次方程式：

$$ax^2 + bx + c = 0$$

在判别式 $b^2 - 4ac < 0$ 时，上述二次方程式无实根。若引进复数，则上述二次方程式的解恒可以公式

$$x = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$$

表示之。但是这种做法并非真正有其必要而且的确是的确有其心理障碍的，这也就是为什麼把「 i 」叫做虚数的原因。但是到後来发现解一元三次方程式的公式解中，即使其三个根都是相异实数时，其公式必须通过复数然後再由于表达式中两部分的虚部相消而得其三个实根。这是第一次在数学中显示复数引入的某种必要性。随著心理障碍的逐渐克服和对于复数理解的逐渐增加，它在数学中的用场和重要性也就开始有了长足的进展，例如代数基本定理的证明，复变函数论在十九世纪的蓬勃开展，以及它在电磁学、量子力学所扮演的自然而且基本的角色。复数系已经是理解大自然必不可缺的基本数系了。

第一章

实数系和函数的连续性——连续函数的基本性质和定理举例

直线是连续不断的，但是一剪即断，这是几何直观上至为明显者；但是几何直观是不能用来计算或用作分析的。Eudoxus 所创的逼近法和逼近理论，亦即任给左、右夹逼数列总是（唯一）存在着为其所夹逼的实数，则把上述直观几何概念转化为便于讨论的「直线连续性的解析描述」，从而重建定量几何基础理论及创积分法之雏形。本章将以直线（亦即实数系）的连续性为基础，进而研讨函数的连续性 (continuity of functions)。

1.1 实数系的连续性

当年古希腊几何学家们基于「可公度性普遍成立」这个「公设」，对于定量平面几何的重要公式，例如矩形面积公式、毕氏定理和相似三角形定理都给以严格的证明，从而建立起洋洋大观的定量平面几何基础初论。不可公度性的发现 (Hippasus)，使得古希腊的整个几何学界经历了近半世纪的严峻挑战；从全人类的理性文明发展史来看，Hippasus 的发现乃是人类在理解大自然的进程上，第一次触及空间本质中的连续性，它对于当年的几何基础初论而言，则是其所有论证的基本假设的全面否定。因此其所证者，其实只是在可公度的特殊情形的证明，而不可公度的一般情形，则整个理论亟待补证。当年 Eudoxus 就是在这样一个迫切任务的挑战下，促使他首创逼近法 (method of approximation)

来达成几何基础论的重建工程。逼近法思想简朴精到，它不但是理解连续性和重建几何基础论的有效途径（如今反观，它其实也是唯一的途径），而且其用途广泛深远，乃是整个分析学和数理分析的基本方法，所以 Eudoxus 当年在重建几何基础论的成就，其实也就是分析学的基础和源起之所在！

改用现代的术语来说，Hippasus 的发现使得我们认识到有理数系是不足以表达直线的长度者，亦即由长度的度量 (measurement of length) 所自然产生的数系——称之为实数系——乃是一个包含有理数系为其真子集的数系，而 Eudoxus 所创的逼近法则是用有理数去理解实数的有效途径。用通常的语句来说，乃是一种以「已知」去理解「未知」、以「简」去御「繁」的具体实践，亦即任给实数皆可用有理数列去逼近之，从而把实数系的研讨，归于其逼近有理数列的研讨。

为了便于往後分析学的研讨，让我们改用现代的数列术语来叙述 Eudoxus 的逼近原理和逼近法：

(一) Eudoxus 比较原则：

当年 Eudoxus 在运用其所创的逼近法去克服不可公度量在当年定量几何基础论所产生的困难时，他充分认识到不可公度的线段比和有理数之间比较大小的原则必须先行明确。前者是有待明确的「未知量」，而後者则是业已熟悉的「已知量」，两者当然不会相等，所以乃是或大、或小的不等关系。由此可见，唯有明确了两者之间比较大小的直观内含之後，才能够运用这种大小关系去以已知研讨未知。此事乃是运用逼近法去研讨各种各样未知量的首务之要。其实，也只有确定了体现其正确的直观内涵的比较原则，才能确保基于这种比较原则，用逼近法所逼近者乃是具有正确直观内含的待定者。例如在研讨不可公度线段比 $\overline{AB} : \overline{A'B'}$ 时，Eudoxus 明确其比较原则为

$$(1.1) \quad \overline{AB} : \overline{A'B'} \begin{cases} > \\ < \end{cases} \frac{m}{n} \Leftrightarrow n\overline{AB} \begin{cases} > \\ < \end{cases} m\overline{A'B'}$$

(二) Eudoxus 逼近原理：

对于一个给定的不可公度线段比 $\overline{AB} : \overline{A'B'}$ 和任给正整数 N ，恒有一个非负整数 m ，使得

$$(1.2) \quad \frac{m}{N} < \overline{AB} : \overline{A'B'} < \frac{m+1}{N}$$

改用现代术语，亦即对于任给一个正的非分数 (positive irrational number) λ ，恒有一对左、右夹逼的分数数列 $\{a_N\}$ 和 $\{b_N\}$ ，满足

$$(1.3) \quad a_N < \lambda < b_N, \quad b_N - a_N = \frac{1}{N}$$

(三) Eudoxus 等量定义和唯一性：

当年 Eudoxus 对于两对不可公度的线段比 $\lambda = \overline{AB} : \overline{A'B'}$ 和 $\mu = \overline{AC} : \overline{A'C'}$ 相等的定义乃是两者和所有分数 $\{\frac{m}{n}\}$ 都具有同样的大小关系。改用现代术语表达，亦即下述熟知的唯一性：若 λ 的上述左、右夹逼数列也是 μ 的左、右夹逼数列，即

$$(1.4) \quad a_N < \left\{ \begin{array}{l} \lambda \\ \mu \end{array} \right\} < b_N, \quad b_N - a_N = \frac{1}{N}$$

对于所有正整数 N 恒成立，则 $\lambda = \mu$ 。上述唯一性的要点在于 $b_N - a_N$ 在 N 无限增大时可以小到任意小，而 λ 和 μ 所可能有的差别，即 $|\lambda - \mu|$ ，总是要比 $b_N - a_N$ 要小者也，所以它必须是零（亦即 $\lambda = \mu$ ）。以现代的观点和术语来描述：任给实数 λ 皆可用一对左、右夹逼分数数列 $\{a_N\}$ 和 $\{b_N\}$ 来加以刻划，它们之间的关系是

$$(1.5) \quad a_1 \leq \dots \leq a_N \leq a_{N+1} \leq \dots \leq \lambda \leq \dots \leq b_{N+1} \leq b_N \leq \dots \leq b_1$$

而且 $(b_N - a_N) \rightarrow 0$ （小到任意小）。通常以符号 $a_N \rightarrow \lambda \leftarrow b_N$ 表达之。再者，因为 $(b_N - a_N) \rightarrow 0$ 所以被满足上述条件的一对数列 $\{a_N\}$ 和 $\{b_N\}$ 所左、右夹逼者是唯一的，亦即

$$(1.6) \quad a_N \rightarrow \lambda \leftarrow b_N \text{ 和 } a_N \rightarrow \mu \leftarrow b_N \Rightarrow \lambda = \mu$$

(四) 左、右夹逼数列的「被夹逼者」的存在性和直线的连续性：

相对于一对左、右夹逼数列（亦即 $\{a_N\}$ 递增， $\{b_N\}$ 递减，而且 $a_N < b_N$, $(b_N - a_N) \rightarrow 0$ 者）之「被夹逼者」的唯一性，当然还可以探讨其存在性的意义何在？在当年 Eudoxus 为了重建几何基础论的研讨中，其所用的夹逼数列乃是为了夹逼一个已给的不可公度线段比而造构者，所以是「有的造矢」，其存在性当然不成问题，而他所要用者则是其「唯一性」。到了现代的分析学，逼近法的应用的范畴和

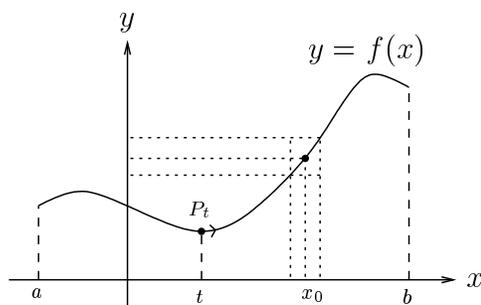
用法大为扩展和深化，其中就自然而然会遇到确立某种构造而得的夹逼数列的「被夹逼者」的存在性问题。归根究底，这种存在性的确立（亦即论证之）当然要有所本（证明是不可能无中生有的！）；而分析学中各种各样存在性定理之所本就是下述最为简朴基本的存在性，亦即：

「任给一对左、右夹逼数列 $\{a_N\}$ 和 $\{b_N\}$ ，
恒存在著被其所夹逼的实数。」

在此，也许有人会问：「为什么？」对于上述问题的恰当回答并非「设法加以论证」。因为这种「挖空心思」的论证，其所本者很可能还比不上所证者的直观性和简朴性，岂非画蛇添足！? 总之，上述问题的恰当回答乃是直截了当地明确上述存在性的直观内含乃是直线是连续不断的连续性 (the continuity of a straight line)。我们对于一条直线的几何直观是：它本身是连续不断的，但是一剪即断。一条直线 l 上任给一点 P 把直线分隔成不连通的两段。由此可见，一条直线是不能缺掉其中任何一点的，否则直线就不是连续不断的了，是不？归根究底，上述存在性的「否定」的直观内含就是缺掉夹逼于其间的那个点（亦即直线具有切断点是也），所以上述存在性的肯定乃是直线的连续性的解析描述（亦即数量化，确切化），通常把它叫做实数系的连续性 (continuity of real number system)。它是整个分析学的一个重要基石，是所有分析学的存在性定理的论证依据。我们将在往后的章节中逐步逐样解说逼近法和连续性在数理分析中既深且广的用场。

1.2 连续函数的基本概念

首先，让我们先来分析一下函数的连续性的直观内含和应该如何赋以明确的定义。用函数 $y = f(x)$ 的图象的几何直观来说， $f(x)$ 在一个区间 $[a, b]$ 上的连续性也就是其图象乃是一条「连续不断」的曲线。如 [图 1-1] 所示，它可以看成一个动点 $P(t, f(t))$ 在时间由 $t = a$ 到 $t = b$ 所经过的轨迹，当 $x = t$ 作微小的变动时， $y = f(t)$ 也只能作微小的变动。



[图 1-1]

把上述直观内含局部化到一个取定点 x_0 的微小邻近来看，亦即：
「 $f(x)$ 和 $f(x_0)$ 的差别是可以小到任意小的，只要 x 和 x_0 的差别小到足够小」。改用数学语句来说，即为：

「对于任给 $\varepsilon > 0$ ，恒有 $\delta > 0$ 使得

$$|x - x_0| < \delta \Rightarrow |f(x) - f(x_0)| < \varepsilon$$

其实，这也就是 $\lim_{x \rightarrow x_0} f(x) = f(x_0)$ 的定义是也。

【定义】（函数连续性）：函数 $f(x)$ 在 $x = x_0$ 点连续的定义就是 $\lim_{x \rightarrow x_0} f(x) = f(x_0)$ ； $f(x)$ 在区间 $[a, b]$ 之上连续的定义乃是 $f(x)$ 在 $[a, b]$ 之中的每一点 x_0 都是连续的（亦即在 $[a, b]$ 上到处连续）。

我们也可以改用数列术语来叙述 $f(x)$ 在 $[a, b]$ 之上的连续性，即：

设 $\{s_n\}$ 是一个取值于 $f(x)$ 的定义区间 $[a, b]$ 之中的数列，而且 $\lim_{n \rightarrow \infty} s_n = \lambda$ ，则恒有

$$(1.7) \quad \lim_{n \rightarrow \infty} f(s_n) = f(\lambda), \quad \text{亦即} \quad \lim_{n \rightarrow \infty} f(s_n) = f\left(\lim_{n \rightarrow \infty} s_n\right)$$

[注一]：一个函数 $f(x)$ 在一个给定点 x_0 的连续性乃是上述极限式对于所有以 x_0 为其极限值的数列皆成立。由此可见，一个函数 $f(x)$ 在 $[a, b]$ 之上的连续性乃是上述局部化的连续性对于 $[a, b]$ 之中的每一点皆成立，其所包含的条件是非常庞大的。

[注二]：自然界各种各样的动态事物，其变动的常态是逐渐的改变，亦即连续地改变，所以描述它们的参变量之间的函数往往自然而然地是

连续函数。当然也会有某些特殊的临界情况，会出现不连续的改变，其相应的函数就会出现不连续的「奇点」(singularity)。

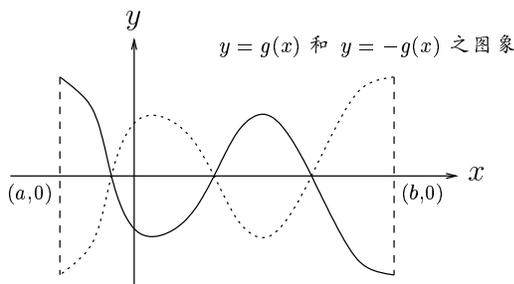
接著让我们研讨在闭区间 $[a, b]$ 之上的连续函数所具有的某些基本性质，它们在分析学的基础理论中扮演著重要的角色。

【定理 1.1】(中间值定理)：设 $f(x)$ 是在 $[a, b]$ 之上连续的函数，而 c 是介乎于 $f(a)$ 和 $f(b)$ 之间的任给实数，则至少存在一个 a, b 之间的点 x_0 使得 $f(x_0) = c$ 。

证明：令 $g(x) = f(x) - c$ ，则由上所设

$$(1.8) \quad g(a) \cdot g(b) = (f(a) - c) \cdot (f(b) - c) < 0$$

而我们所要证明者就是存在 $a < x_0 < b$ 使得 $g(x_0) = 0$ 。由 $g(x)$ 的图象来看，它是一条端点（即 $(a, g(a))$ 和 $(b, g(b))$ ）分居于 x -轴的两侧的一条连续曲线。从几何直观上是可以想到它必须和 x -轴至少有一个交点。



[图 1-2]

首先，我们要认清上述存在性的证明自然要用到实数系连续性的解析描述，即那个左、右夹逼数列「所夹逼者」的存在性。而下述证法也就是利用 $g(x)$ 在 $[a, b]$ 之上的连续性和 $g(a) \cdot g(b) < 0$ 去构造一对左、右夹逼数列 $\{a_n\} \rightarrow \leftarrow \{b_n\}$ ，使得其所夹逼者 x_0 就是所求者，亦即 $g(x_0) = 0$ 。

以 $[a, b]$ 的中点 $\frac{1}{2}(a+b)$ 把 $[a, b]$ 二等分成左、右两段。若 $g(\frac{1}{2}(a+b)) = 0$ 则定理业已得证。不然，则 $g(\frac{1}{2}(a+b)) \neq 0$ 必然和异号的 $\{g(a), g(b)\}$ 中之一异号，因此必有一个半段 $[a_1, b_1]$ 依然满足 $\{g(a_1), g(b_1)\}$ 异号者。如此逐次二等分（不妨设每次的分点都不是 $g(x)$ 的零点），每次总是

由 $[a_n, b_n]$ 取其半段 $[a_{n+1}, b_{n+1}]$ 使得 $\{g(a_{n+1}), g(b_{n+1})\}$ 依然保持异号。这样就构造而得一对左、右夹逼数列 $\{a_n\} \rightarrow \leftarrow \{b_n\}$ ，它具有

$$(1.9) \quad g(a_n) \cdot g(b_n) < 0 \quad \text{对于所有 } n \text{ 成立}$$

在此，由实数系的连续性即得它们所夹逼者的存在性，令其为 x_0 ，亦即

$$(1.10) \quad \lim_{n \rightarrow \infty} a_n = x_0 = \lim_{n \rightarrow \infty} b_n$$

再者，由 $g(x)$ 的连续性和 (1.9)-式即有

$$(1.11) \quad \begin{aligned} 0 &\geq \lim_{n \rightarrow \infty} g(a_n) \cdot g(b_n) = \lim_{n \rightarrow \infty} g(a_n) \cdot \lim_{n \rightarrow \infty} g(b_n) = g(x_0) \cdot g(x_0) \\ &\Rightarrow g(x_0) = 0 \end{aligned} \quad \square$$

【定义】：一个定义于区间 $[a, b]$ 上的函数 $f(x)$ ，若其值恒小于或等于一个给定常数 K ，则称之为上界者 (bounded above)，而 K 则是 $\{f(x), a \leq x \leq b\}$ 的一个上界，亦即

$$(1.12) \quad f(x) \leq K, \quad a \leq x \leq b$$

若把上式的“ $\leq K$ ”改为“ $\geq K$ ”恒成立，则称 K 为 $\{f(x), a \leq x \leq b\}$ 的一个下界，而 $f(x)$ 则为下界者 (bounded below)。

上、下均有界者，通常简称为有界者 (bounded)。

【引理 1.1】：设 $f(x)$ 为 $[a, b]$ 上的一个连续函数，则 $f(x)$ 是有界者。

证明：上有界性和下有界性的证明本质完全一样，我们将用反证法来证明其上有界性。

设 $f(x)$ 在 $[a, b]$ 之上没有上界。将 $[a, b]$ 等分为两个半段，则 $f(x)$ 至少在一个半段上没有上界。如此逐次二等分，每次取其半段，依然保持 $f(x)$ 在其上是没有上界者，这样就构造而得一对左、右夹逼数列 $\{a_n\} \rightarrow \leftarrow \{b_n\}$ ，使得 $f(x)$ 在 $[a_n, b_n]$ 之上是没有上界的。在此，再用实数系的连续性得知存在被其所夹逼的 x_0 ，亦即

$$(1.13) \quad \lim_{n \rightarrow \infty} a_n = x_0 = \lim_{n \rightarrow \infty} b_n$$

亦即在 x_0 的任意小的邻近 $(x_0 - \delta, x_0 + \delta)$, $\delta > 0$, 总是包含着一个 n 足够大的 $[a_n, b_n]$ 。

一方面由 $f(x)$ 在 x_0 点的局部化连续性得知当 δ 足够小时

$$(1.14) \quad |x - x_0| < \delta \Rightarrow |f(x) - f(x_0)| < \varepsilon$$

所以 $f(x)$ 在 $(x_0 - \delta, x_0 + \delta)$ 之上显然是有界的, 所以在它的子区间 $[a_n, b_n]$ 之上当然也是有界的。但是另一方面, 由上述 $\{a_n\} \rightarrow \leftarrow \{b_n\}$ 的构造法, 我们一直保持著 $f(x)$ 在 $[a_n, b_n]$ 上的无上界性。两者显然矛盾, 这也就证明原本的假设和 $f(x)$ 在 $[a, b]$ 之上的连续性是互相矛盾的, 所以 $f(x)$ 在 $[a, b]$ 之上是有界的。□

【引理 1.2】: 设 S 是一个有上界的非空实数子集, 则在其所有上界之中, 必有其最小者, 称之为 S 的极小上界。同理, 设 S 是一个有下界的非空实数子集, 则在其所有下界之中, 必有其最大者, 称之为 S 的极大下界。

证明: 若 S 中有一个极大者 s_0 (或极小者 s'_0) 则它显然就是 S 的极小上界 (或极大下界)。所以我们不妨在下述证明中设 S 中没有极大者 (或极小者)。

设 K_1 是 S 的一个上界, k_1 是 S 的非上界 (亦即至少具有一个大于 k_1 的 $s \in S$)。将 $[k_1, K_1]$ 等分为两段, 选取其半段 $[k_2, K_2]$ 满足 K_2 为 S 的上界而 k_2 则并非 S 的上界。如此逐次两分, 每次由 $[k_n, K_n]$ 选取其半段 $[k_{n+1}, K_{n+1}]$, 使得 K_{n+1} 和 k_{n+1} 分别保持其 S 的上界性和非上界性。如此即得一对左、右夹逼数列 $\{k_n\} \rightarrow \leftarrow \{K_n\}$ 。不难验证它们所夹逼者, 就是 S 极小上界。□

【定理 1.2】(极小值、极大值存在定理): 设 $f(x)$ 为 $[a, b]$ 之上的一个连续函数, 则 $[a, b]$ 之中存在 \hat{x}_0 和 \check{x}_0 使得

$$(1.15) \quad f(\check{x}_0) \leq f(x) \leq f(\hat{x}_0)$$

对于所有 $a \leq x \leq b$ 恒成立。

证明: 由 [引理 1.1] 得知 $S = \{f(x), a \leq x \leq b\}$ 是有界的。再由 [引理 1.2] 得知 S 分别具有其极小上界和极大下界, 以 M 和 m 记之。而

本定理所要证明者乃是在 $[a, b]$ 之中分别存在 \hat{x}_0 和 \check{x}_0 使得 $f(\hat{x}_0) = M$, $f(\check{x}_0) = m$ 。其证法依然是用二分法去分别构造两对左、右夹逼数列 $\{a_n\} \rightarrow \leftarrow \{b_n\}$ 和 $\{a'_n\} \rightarrow \leftarrow \{b'_n\}$ ，其所夹逼者分别是所求证其存在性的 \hat{x}_0 和 \check{x}_0 。其逐步归纳构造法则分别如下：

设 $S = S' \cup S''$ 是一个有界实数集，则在 S' 和 S'' 之中至少有其一，它的极小上界（或极大下界）和 S 的极小上界（或极大下界）相同。基于上述简单事实我们可以逐步归纳地由 $[a_n, b_n]$ 和 $[a'_n, b'_n]$ 分别取其半段为 $[a_{n+1}, b_{n+1}]$ （及 $[a'_{n+1}, b'_{n+1}]$ ）使得 $f(x)$ 在 $[a_{n+1}, b_{n+1}]$ 上（及 $[a'_{n+1}, b'_{n+1}]$ 上）的极小上界（及极大下界）依然是 M （及 m ）。

令 \hat{x}_0 （及 \check{x}_0 ）分别是 $\{a_n\} \rightarrow \leftarrow \{b_n\}$ （及 $\{a'_n\} \rightarrow \leftarrow \{b'_n\}$ ）所左、右夹逼者。不难用 $f(x)$ 在 \hat{x}_0 （及 \check{x}_0 ）的局部化连续性结合上述所构造的 $[a_n, b_n]$ 和 $[a'_n, b'_n]$ 的性质验证 $f(\hat{x}_0) = M$ 及 $f(\check{x}_0) = m$ 。
[其验证留作习题] □

均匀连续性：

在一个给定函数 $f(x)$ 的局部化逐点连续性的解析描述中，那个足够小的 $\delta > 0$ 使得

$$(1.16) \quad |x - x_0| < \delta \Rightarrow |f(x) - f(x_0)| < \varepsilon$$

是随著 $\varepsilon > 0$ 和给定点 x_0 的取定而定的。一般来说，对于同一 $\varepsilon > 0$ ，在 x_0 点业已足够小的 δ 对另外一点 x'_0 就可能不够小。是否能够选取一个仅仅随著 $\varepsilon > 0$ 而定的 $\delta > 0$ ，它对于 $f(x)$ 在 $[a, b]$ 之上每一点的局部化连续性都足够小呢？亦即是否对于任给 $\varepsilon > 0$ ，总是有一个足够小的 $\delta > 0$ ，使得

$$(1.17) \quad |x_1 - x_2| < \delta \Rightarrow |f(x_1) - f(x_2)| < \varepsilon$$

能够对于所有 $x_1, x_2 \in [a, b]$ 均成立呢？这种比较划一的连续性叫做均匀连续性 (uniform continuity)。

【定理 1.3】：设 $f(x)$ 是一个在 $[a, b]$ 上的连续函数，则它也是在 $[a, b]$ 上均匀连续的。

证明：我们将用反证法，亦即由上述命题的「否定命题」出发，运用 $f(x)$ 在 $[a, b]$ 上的连续性和实数系连续性去推导矛盾。为此，我们得先将其否定命题的逻辑内含加以明确：

$f(x)$ 在 $[a, b]$ 上并非均匀连续者，乃是存在有一个 $\varepsilon_0 > 0$ 使得不论 $\delta > 0$ 有多麼小，在 $[a, b]$ 之中总会有一对相距小于 δ 的 $\{x_1, x_2\}$ ，使得 $|f(x_1) - f(x_2)| > \varepsilon_0$ 。

我们把上述命题以「 $P(\varepsilon_0)$ 在 $[a, b]$ 上成立」简称之。

把 $[a, b]$ 等分为两段，则其中至少有一个半段 $[a_1, b_1]$ 使得 $P(\varepsilon_0)$ 依然在 $[a_1, b_1]$ 上成立。假若不然，亦即有一个足够小的 δ 使得

$$(1.18) \quad |x_1 - x_2| < \delta \Rightarrow |f(x_1) - f(x_2)| < \varepsilon_0$$

对于 $\{x_1, x_2\}$ 同在前半段或同在後半段时皆成立。再者，由 $f(x)$ 在分点 $\frac{1}{2}(a+b)$ 的连续性，即有另一个足够小的 $\delta' > 0$ 使得

$$(1.19) \quad \left| x - \frac{1}{2}(a+b) \right| < \delta' \Rightarrow \left| f(x) - f\left(\frac{1}{2}(a+b)\right) \right| < \frac{\varepsilon_0}{2}$$

不妨设 $\delta \leq \delta'$ 。因此当 x_1, x_2 分居于前、後半段时，亦有

$$\begin{aligned} & \left| x_1 - \frac{1}{2}(a+b) \right| < \delta \leq \delta', \quad \left| x_2 - \frac{1}{2}(a+b) \right| < \delta \leq \delta' \\ \Rightarrow & |f(x_1) - f(x_2)| \leq \left| f(x_1) - f\left(\frac{1}{2}(a+b)\right) \right| + \left| f(x_2) - f\left(\frac{1}{2}(a+b)\right) \right| \\ & < \frac{\varepsilon_0}{2} + \frac{\varepsilon_0}{2} = \varepsilon_0 \end{aligned}$$

所以和 $P(\varepsilon_0)$ 在 $[a, b]$ 上成立的假设相矛盾。由此可见 $P(\varepsilon_0)$ 至少在一个半段 $[a_1, b_1]$ 上成立，而且我们可以逐次二等分，每次选取其中半段 $[a_n, b_n]$ ，依然保持 $P(\varepsilon_0)$ 在 $[a_n, b_n]$ 成立。这样，就构造而得一对左、右夹逼数列 $\{a_n\} \rightarrow \leftarrow \{b_n\}$ ， $P(\varepsilon_0)$ 在所有 $[a_n, b_n]$ 上皆成立。由实数系连续性，存在著它们所夹逼的 x_0 。由 $f(x)$ 在 x_0 点连续性即有一个足够小的 δ 使得

$$(1.20) \quad |x - x_0| < \delta \Rightarrow |f(x) - f(x_0)| < \frac{\varepsilon_0}{2}$$

再者，当 n 足够大时，显然有 $[a_n, b_n] \subset (x_0 - \delta, x_0 + \delta)$ 。因此

$$\begin{aligned} x_1, x_2 \in [a_n, b_n] &\Rightarrow |x_1 - x_0| < \delta, \quad |x_2 - x_0| < \delta \\ \Rightarrow |f(x_1) - f(x_2)| &\leq |f(x_1) - f(x_0)| + |f(x_2) - f(x_0)| \\ &< \frac{\varepsilon_0}{2} + \frac{\varepsilon_0}{2} = \varepsilon_0 \end{aligned}$$

显然和所构造的 $P(\varepsilon_0)$ 在 $[a_n, b_n]$ 成立相矛盾。归根研底，上述矛盾证明了原始假设 $P(\varepsilon_0)$ 在 $[a, b]$ 上成立乃是不能成立的。这也就证明了 $f(x)$ 在 $[a, b]$ 上的均匀连续性。 \square

1.3 多项式函数

在各种各样函数之中，最为简朴者首推多项式函数。一个给定的多项式函数 $f(x)$ ，其在 a 点的函数值可以由所给的代数式直接计算而得。再者，一个多项式 $f(x)$ 在 a 点邻近的局部性质的研讨是非常简单明了的。我们可以用替换 $x = a + t$ 把它改写成 t 的升幂多项式，亦即

$$(1.21) \quad f(x) = f(a + t) = f(a) + c_1 t + c_2 t^2 + \dots + c_n t^n$$

其中 $|t|$ 是足够小者，因而右侧各项的绝对值乃是随著次数的升高而大幅缩小，因此它们的局部影响力显然是高次项要远小于低次项，可以说乃是「阶段分明，一目了然」者也。由此可见，要研究一般函数在 a 点邻近的局部性质的一个好办法就是设法用多项式函数去局部逼近它，从而把所给函数在 a 点的局部性质的研讨，归于其局部逼近多项式的局部性质而研讨之。其实，这也就是「微分学」的基本思想！本节先将对多项式的常用基本性质作一简要论述。（参看基础代数学第一、二、三章的讨论。）

1.3.1 多项式的唯一性定理与插值公式

由熟知的简易公式

$$(1.22) \quad (x^k - a^k) = (x - a)(x^{k-1} + ax^{k-2} + \dots + a^{k-1})$$

易证 $f(x) - f(a)$ 含有 $(x - a)$ 因式。设

$$(1.23) \quad f(x) = \sum_{k=0}^n c_k x^k$$

则有

$$(1.24) \quad \begin{aligned} f(x) - f(a) &= \sum_{k=1}^n c_k (x^k - a^k) \\ &= (x - a) \sum_{k=1}^n c_k (x^{k-1} + ax^{k-2} + \dots + a^{k-1}) \\ &= (x - a)q(x, a) \end{aligned}$$

[注]：由 (1.24)-式易见

$$(1.25) \quad \lim_{x \rightarrow a} \frac{f(x) - f(a)}{x - a} = \lim_{x \rightarrow a} q(x, a) = q(a, a) = \sum_{k=1}^n c_k (ka^{k-1})$$

【定理 1.4】：一个 n 次多项式 $f(x)$ 至多只有 n 个相异的根。再者，一个 n 次多项式由其在（相异的） $(n + 1)$ 点之值所唯一确定。

证明：(i) 用归纳法易证：当 $\{a_i, 1 \leq i \leq \ell\}$ 是 $f(x)$ 的 ℓ 个相异之根，则

$$f(x) = (x - a_1) \dots (x - a_\ell) Q(x)$$

所以当 $\ell = n$ 时，

$$f(x) = c_n (x - a_1) \dots (x - a_n)$$

显然任何和 $\{a_i, 1 \leq i \leq n\}$ 都不同的 a_{n+1} 不可能 使得 $f(a_{n+1}) = 0$ ，因为

$$f(a_{n+1}) = c_n (a_{n+1} - a_1) \dots (a_{n+1} - a_n)$$

乃是 $(n + 1)$ 个非零者的乘积，也一定是非零的。

(ii) 设 $f(x)$ 和 $g(x)$ 都是次数至多为 n 的多项式，而且它们在相异的 $(n + 1)$ 个点 $\{a_j, 0 \leq j \leq n\}$ 上取相同之值，则 $f(x)$ 和 $g(x)$ 必须是同一个多项式。假若不然，则 $(f(x) - g(x))$ 乃是一个次数至多为 n 的多项

式，而由所设它至少有 $(n+1)$ 个相异的根 $\{b_j, 0 \leq j \leq n\}$ ，此事和 (i) 中所证矛盾。□

插值问题和插值公式：

既然一个次数至多为 n 的多项式 $f(x)$ 业已由它在 $(n+1)$ 个取定点 $\{a_j, 0 \leq j \leq n\}$ 之值所唯一确定，自然就会问下述存在性问题：

【插值问题】：设 $\{a_j, 0 \leq j \leq n\}$ 是任给 $(n+1)$ 个相异之点， $\{v_j, 0 \leq j \leq n\}$ 是任给（不一定相异）之值。是否存在一个次数至多为 n 的多项式 $f(x)$ 使得

$$(1.26) \quad f(a_j) = v_j, \quad 0 \leq j \leq n$$

由中国古算中韩信点兵法的启示，我们应该先行求解 $\{v_j\}$ 之中只有一个为 1 而其他皆为零的情形，例如 $v_0 = 1, v_1 = v_2 = \dots = v_n = 0$ 。由 [定理 1.4]，这个所求的 $f_0(x)$ 可以表达成

$$(1.27) \quad f_0(x) = c(x - a_1)(x - a_2) \dots (x - a_n)$$

而且

$$(1.27') \quad 1 = f_0(a_0) = c(a_0 - a_1)(a_0 - a_2) \dots (a_0 - a_n)$$

所以

$$(1.28) \quad f_0(x) = \prod_{j=1}^n (x - a_j) / \prod_{j=1}^n (a_0 - a_j)$$

同理，那个满足 $f_k(a_k) = 1$ 而且 $f_k(a_j) = 0, j \neq k$ ，者即为

$$(1.29) \quad f_k(x) = \prod_{j \neq k} (x - a_j) / \prod_{j \neq k} (a_k - a_j)$$

至此，我们所求的一般公式业已呼之欲出了了。只要用分配律就可看出

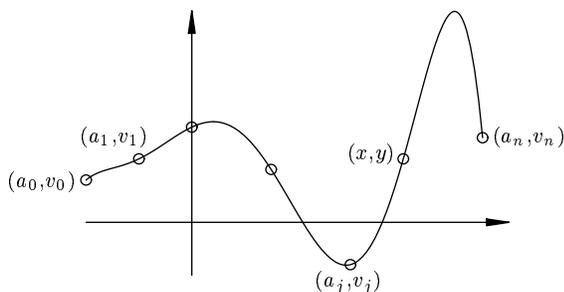
$$(1.30) \quad f(x) = \sum_{k=0}^n v_k f_k(x)$$

显然就是那个所求者, 因为 $f_k(a_j)$ 中只有 $f_j(a_j) = 1$, 其他皆为 0, 亦即

$$(1.31) \quad f(a_j) = \sum_{k=0}^n v_k f_k(a_j) = v_j f_j(a_j) = v_j$$

这也就是 Lagrange 插值公式, 它其实就是韩信点兵法的直接推广。

让我们再用函数的图象的观点来研讨上述插值问题, 则所求者乃是 $(n+1)$ 个待定系数 $\{c_k, 0 \leq k \leq n\}$ 使得 $y = f(x) = \sum_{k=0}^n c_k x^k$ 的图象过 $(n+1)$ 个给定点 $\{(a_j, v_j), 0 \leq j \leq n\}$ 。



[图 1-3]

设 (x, y) 是 $y = f(x)$ 图象曲线上任意一点, 则有下列 $\{c_k, 0 \leq k \leq n\}$ 的条件方程组

$$(1.32) \quad \begin{aligned} -y + \sum_{k=0}^n x^k c_k &= 0 \\ -v_j + \sum_{k=0}^n a_j^k c_k &= 0, \quad 0 \leq j \leq n \end{aligned}$$

把它们想成以下述 $(n+2) \times (n+2)$ 方阵为其系数矩阵的齐次线性方程组

$$(1.32') \quad \begin{pmatrix} -y & 1 & x & \dots & x^n \\ -v_0 & 1 & a_0 & \dots & a_0^n \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ -v_j & 1 & a_j & \dots & a_j^n \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ -v_n & 1 & a_n & \dots & a_n^n \end{pmatrix}$$

则它具有非零解 $(-1, c_0, \dots, c_n)$ 的充要条件就是上述方阵的行列式等于 0，亦即

$$(1.33) \quad \begin{vmatrix} -y & 1 & x & \dots & x^n \\ -v_0 & 1 & a_0 & \dots & a_0^n \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ -v_j & 1 & a_j & \dots & a_j^n \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ -v_n & 1 & a_n & \dots & a_n^n \end{vmatrix} = 0$$

把上述行列式以第一列 (first column) 展开，遍除以 $(-y)$ 的系数子行列式，再用熟知的 Vandermonde 行列式公式即得回 Lagrange 公式

$$(1.30) \quad -y + \sum_{k=0}^n v_k f_k(x) = 0$$

所以 (1.33) 和 (1.30) 是相通的两种插值公式。

1.3.2 单元多项式的除法与辗转相除求公因式

设 k 是一个在四则运算上满足和有理数系 (或实数系、复数系) 同样的运算律的体系，通常称之为域 (field)。令 $k[x]$ 是所有以 k 中的元素为系数的多项式所组成的代数体系，通常称之为 k 上的单元多项式环 (polynomial ring over k)，它具有熟知的加法和乘法，而且满足常用的运算律例如加、乘的结合律与交换律，乘对于加的分配律等等。总之， $k[x]$ 在加和乘上和整数系在加和乘上具有十分相似的运算律。再者， $k[x]$ 也具有和整数系类似的带余除式，亦即

【引理 1.3】：设 $f(x), g(x)$ 是 $k[x]$ 中任给一对多项式， $g(x) \neq 0$ ，则存在唯一的商式 $q(x)$ 和余式 $r(x)$ 使得

$$(1.34) \quad f(x) = q(x) \cdot g(x) + r(x)$$

而且 $r(x)$ 的次数低于 $g(x)$ 的次数，或 $r(x) \equiv 0$ 。

[注意]：非零常数的次数为 0，但是零多项式的次数则定义为 $-\infty$ 。唯有如此定义，通用的公式

$$(1.35) \quad \deg(p(x) \cdot q(x)) = \deg p(x) + \deg q(x)$$

才得以普遍成立。

证明：设 $f(x)$ 和 $g(x)$ 的次数分别是 n 和 m 。

若 $n < m$ ，则 $q(x) = 0, r(x) = f(x)$ 。

若 $n = m$ ，则 $q(x) = \frac{a_0}{b_0}, r(x) = f(x) - \frac{a_0}{b_0}g(x)$ ，其中 a_0x^n 和 b_0x^m 分别是 $f(x)$ 和 $g(x)$ 的首项。

若 $n > m$ ，则易见

$$(1.36) \quad f_1(x) = f(x) - \frac{a_0}{b_0}x^{n-m} \cdot g(x)$$

的次数低于 n ，所以即可对 n 归纳证之。 \square

【定理 1.5】：设 $d(x)$ 是 $f(x)$ 和 $g(x)$ 的最高公因式，则存在 $k[x]$ 中适当的 $A(x)$ 和 $B(x)$ 使得

$$(1.37) \quad d(x) = A(x)f(x) + B(x)g(x)$$

证明：在整数系的情形，我们有熟知的辗转相除求两个给定整数 m, n 的最大公约数。同样的，在 $k[x]$ 中也有同样的辗转相除求两个给定多项式 $f(x)$ 和 $g(x)$ 的最高公因式，(1.37)-式乃是这样求法的直接推论。因为逐次所得的余式总是可以表达成 $f(x)$ 和 $g(x)$ 的倍式之和者也，例如

$$(1.38) \quad \begin{aligned} r_1(x) &= 1 \cdot f(x) + (-q_1(x)) \cdot g(x) \\ r_2(x) &= 1 \cdot g(x) + (-q_2(x)) \cdot r_1(x) \\ &= (-q_2(x))f(x) + (1 + q_1(x)q_2(x))g(x) \quad \text{等等} \quad \square \end{aligned}$$

【定义】：若 $f(x)$ 含有因式 $(x - a)^2$ ，则称 a 为 $f(x)$ 的一个重根 (multiple root)。

易见 a 是 $f(x)$ 的重根的充要条件是 $f(a) = 0$ 而且 $Df(a) = 0$ 。

【定义】：若 $f(x)$ 和 $g(x)$ 的最高公因式 $d(x)$ 次数是 0，则称 $f(x)$ 和 $g(x)$ 是互素的 (relatively prime)。

【推论】：设 $f(x)$ 和 $g(x)$ 是互素的，而且

$$(1.39) \quad \deg h(x) < \deg f(x) + \deg g(x)$$

则存在唯一的一组 $p(x)$ 和 $q(x)$, 使得

$$(1.40) \quad \frac{h(x)}{f(x) \cdot g(x)} = \frac{p(x)}{f(x)} + \frac{q(x)}{g(x)}$$

而且 $\deg p(x) < \deg f(x)$, $\deg q(x) < \deg g(x)$ 。

证明: 由所设, 1 可以表达成 $f(x)$ 和 $g(x)$ 的倍式之和, 即有 $A(x)$ 和 $B(x)$ 使得

$$(1.41) \quad 1 = A(x)f(x) + B(x)g(x)$$

两边同乘以 $h(x)$, 即得

$$(1.42) \quad h(x) = (h(x)A(x))f(x) + (h(x)B(x))g(x)$$

令 $p(x)$ 为 $h(x)B(x)$ 被 $f(x)$ 除的余式, 亦即

$$(1.43) \quad h(x)B(x) = C(x)f(x) + p(x), \quad \deg p(x) < \deg f(x)$$

令 $q(x) = h(x)A(x) + C(x)g(x)$, 则有

$$(1.44) \quad h(x) = q(x)f(x) + p(x)g(x), \quad \deg q(x) < \deg g(x)$$

这也就证明了满足 (1.40)-式的 $\{p(x), q(x)\}$ 的存在性。再者, 它们的唯一性乃是 $f(x)$ 和 $g(x)$ 互素和 $\deg p(x) < \deg f(x)$, $\deg q(x) < \deg g(x)$ 的直接推论。□

1.3.3 Sturm 定理

设 $f(x)$ 是一个实系数多项式, $[a, b]$ 是一个给定区间。是否有一个有效能算的判定法, 可以确定 $f(x)$ 在 $[a, b]$ 中实根的个数 (不计重数)? 这是在整个实系数多项式理论及应用上一个至关重要的基本问题。本节所证明的 Sturm 定理, 提供了上述基本问题的完美解答。

若 $f(x)$ 含有重根, 则其重根都含于 $f(x)$ 和 $Df(x)$ 的最高公因式 $d(x)$ 之中, 而且 $\tilde{f}(x) = f(x)/d(x)$ 含有和 $f(x)$ 同样的根, 但是不含重根。所以在研讨上述问题时不妨设 $f(x)$ 和 $Df(x)$ 互素。令 $f_0(x) = f(x)$,

$f_1(x) = Df(x)$ 然後由 $f_0(x), f_1(x)$ 的辗转相除, 求得下述一串「餘式」 $f_2(x), f_3(x), \dots, f_k(x)$, 即有

$$\begin{aligned} f_0(x) &= q_1(x)f_1(x) - f_2(x) \\ f_1(x) &= q_2(x)f_2(x) - f_3(x) \\ &\vdots \quad \quad \quad \vdots \\ f_{k-2}(x) &= q_{k-1}(x)f_{k-1}(x) - f_k(x) \end{aligned}$$

注意: 上述除式中的餘式带有负号, 所以 $f_2(x), f_3(x), \dots, f_k(x)$ 乃是通常的辗转相除中所得的餘式再乘以 (-1) 。再者由所设 $f_0(x)$ 和 $f_1(x)$ 互素, 所以 $f_k(x)$ 其实是一个非零常数。

上述函数串 $\{f_0(x), f_1(x), f_2(x), \dots, f_k(x)\}$ 叫做 Sturm 函数串。

【定理 1.6】(Sturm 定理): $f(x)$ 在区间 (a, b) 之内的根的个数等于 $V(a) - V(b)$, 其中 $V(a)$ 和 $V(b)$ 分别是下列数串

$$(1.45) \quad \{f_0(a), f_1(a), f_2(a), \dots, f_k(a)\} \text{ 和 } \{f_0(b), f_1(b), f_2(b), \dots, f_k(b)\}$$

的变号数。

[注]: 对于任给一数 c , 不可能有相邻的 $f_i(c)$ 和 $f_{i+1}(c)$ 同时为 0。不然, 则由上述除式即可逐步推导得 $f_j(c) = 0, i \leq j \leq k$, 但是 $f_k(x)$ 是一个非零常数! 再者, 设有 $f_i(c) = 0, 0 < i < k$, 则由除式可见 $f_{i-1}(c) = -f_{i+1}(c) \neq 0$ 。所以不论把 $f_i(c)$ 想成 + 或 -, 数串 $\{f_i(c), 0 \leq i \leq k\}$ 的变号数是一样的。由此可见在 c 不是 $f(x)$ 的一个根时, 即使 $\{f_i(c)\}$ 中含有为零者 $V(c)$ 也是唯一确定的。

证明: 若区间 $[c_1, c_2]$ 中不含有任何 $f_i(x)$ 的根 $0 \leq i \leq k$, 则由中间值定理可知 $f_i(c_1)$ 和 $f_i(c_2), 0 \leq i \leq k$, 每一对皆为同号。因此 $V(c_1)$ 当然和 $V(c_2)$ 相同。

(i) 设 c 是其中一个 $f_i(x), 0 < i < k$ 的根。则有 $f_{i-1}(c) = -f_{i+1}(c) \neq 0$ 是异号的。再者, 在 $\delta > 0$ 取得足够小时, $f_{i-1}(c \pm \delta)$ 和 $f_{i-1}(c)$ 同号, $f_{i+1}(c \pm \delta)$ 和 $f_{i+1}(c)$ 同号, 所以 $f_{i-1}(c \pm \delta)$ 和 $f_{i+1}(c \pm \delta)$ 异号, 如下图

表所示

	$c - \delta$	c	$c + \delta$
f_{i-1}	±	±	±
f_i	+ 或 -	0	+ 或 -
f_{i+1}	∓	∓	∓

不论 $f_i(c \pm \delta)$ 的正、负，在 $\{f_{i-1}, f_i, f_{i+1}\}$ 这一段所含的变号数总是 1。

(ii) 设 c 是 $f_0(x) = f(x)$ 的一个根。则由无重根之所设 $f_1(c) = Df(c) \neq 0$ ，而且在 $\delta > 0$ 取得足够小时， $f_1(x)$ 在区间 $[c - \delta, c + \delta]$ 上保持其正、负不变，亦即 $f_0(x) = f(x)$ 在 $[c - \delta, c + \delta]$ 上保持其递增、递减性不变。由此可见， $\{f_0, f_1\}$ 在 $[c - \delta, c + \delta]$ 上的符号如下表所示：

	$c - \delta$	c	$c + \delta$
f_0	∓	0	±
f_1	±	±	±

由上表可见 $\{f_0, f_1\}$ 这一段所含的变号数在 c 的左侧为 1 但是在 c 的右侧则是 0。

总结上面对于变号数的逐段局部分析如下：

设区间 $[a, b]$ 中 $f(x)$ 的根的个数为 k ，它们把 $[a, b]$ 分割成 $(k + 1)$ 个区间，在每一分段上变号数 V 保持不变，但是从 $f(x)$ 的一个根的左侧到其右侧，则变号数 V 减 1。这也就证明了

$$V(b) = V(a) - k$$

亦即

$$k = V(a) - V(b) \quad \square$$

1.3.4 代数基本定理

在实数系的范围中，一个像 $x^2 + 1$ 这样简单的多项式已经是没有它的根了。当年把实数系扩充到复数系 $\mathbb{C} = \{x + iy, x, y \in \mathbb{R}\}$ 至少使得所有二次方程式都有复数解，接著发现所有三次、四次方程式的根也都已

经在复数范围之内。很自然会问，是否任何高次方程式的根也都在复数系之内呢？还是会有些高次方程式在复数系中依然无解呢？是耶非耶，这就是当年 Euler 和 Lagrange 他们想解答的一个代数学基本问题。其答案是任何高次复系数多项式都存在有复数根，然後很容易用余式定理归纳地证明它的所有根都已经在复数系之中。这就是高斯 (Gauss) 在 1799 年于他的博士论文中所证明的代数基本定理 (Fundamental Theorem of Algebra)。

这个定理有好几个不同的证法。下面所给者是一个比较初等而且证明中所涉的步骤又是十分简朴直观的证法。当 $f(z)$ 是一次多项式时，定理显然成立，所以我们设 $f(z)$ 的次数最少为 2。

令 $F(z) = |f(z)|^2 = f(z) \cdot \overline{f(z)} (\geq 0)$, $z = x + iy$ 。则 $F(z)$ 可以想成一个二元连续函数，亦即当 $a_n \rightarrow x_0, b_n \rightarrow y_0$ (亦即 $a_n + ib_n \rightarrow x_0 + iy_0$) 则恒有 $F(a_n + ib_n) \rightarrow F(x_0 + iy_0)$ 。

把 F 的变域限制到下述方块之上

$$(1.46) \quad \square(2K, 2K) = \{(x, y), |x| \leq K, |y| \leq K\}$$

我们将用实数系的连续性去证明 F 在 $\square(2K, 2K)$ 的函数值中有一个极小值，亦即存在 $\square(2K, 2K)$ 中一点 (x_0, y_0) 使得

$$(1.47) \quad F(x_0, y_0) \leq F(x, y), \quad (x, y) \in \square(2K, 2K)$$

恒成立。

[设想]：假如 $\square(2K, 2K)$ 业已大到包括我们所要证明其存在的 $f(z)$ 的一个根的话，则上述极小值必须是 0，而 $x_0 + iy_0$ 就是 $f(z)$ 的一个根。由此可见，代数基本定理证明的要点在于论证当 K 足够大时，上述极小值必须等于 0。

【引理 1.4】： F 在 $\square(2K, 2K)$ 上的所有函数值中，存在有一个极小值，亦即存在 $(x_0, y_0) \in \square(2K, 2K)$ 使得

$$(1.48) \quad F(x_0, y_0) \leq F(x, y)$$

对所有 $|x| \leq K, |y| \leq K$ 皆成立。

证明：令 S 为 F 在 $\square(2K, 2K)$ 上所取的所有函数值所成的集合。以 $g.l.b.(S)$ 表示 S 的极大下界。

[注意]：一个具有下界的实数子集 S 总是有一个极大下界的（参看 [引理 1.1] 及 [引理 1.2]），但是这个 S 的极大下界却不一定属于 S ；若属于 S ，则它当然就是 S 的极小者。

把 $\square(2K, 2K)$ 等分成四个 $\square(K, K)$ （每个都包含其边界）。令 S_i 是函数 F 在各别方块上函数值的集合， $i = 1, 2, 3, 4$ 。因为 S 的极大下界等于 S_i 的极大下界之中的最小者，所以至少有一个 S_i 的极大下界等于 S 的极大下界。设其所相应的方块是 \square_1 ，我们又可以把 \square_1 等分成四个方块而选取其中之一 \square_2 ，使得 F 在 \square_2 上的函数值的极大下界等于 F 在 $\square(2K, 2K)$ 上的函数值极大下界。如此逐步四等分，而每次选择其中之一，使得 F 在其上函数值的极大下界依然和 F 在 $\square(2K, 2K)$ 上者相等，这样就得到「方块列」

$$(1.49) \quad \square(2K, 2K) \supset \square_1 \supset \square_2 \supset \cdots \supset \square_n \supset \square_{n+1} \supset \cdots$$

F 在 \square_n 上的函数值的极大下界一直和 F 在 $\square(2K, 2K)$ 上者相等。再者，由上述逐次四等分然後选取其中之一这种几何构造法，可见有两对左、右夹逼实数数列 $\{a_n\} \rightarrow \leftarrow \{b_n\}$ 和 $\{c_n\} \rightarrow \leftarrow \{d_n\}$ 使得

$$(1.50) \quad \square_n = \{(x, y); a_n \leq x \leq b_n, c_n \leq y \leq d_n\}$$

在此，我们运用实数系的连续性即得

$$(1.51) \quad \lim_{n \rightarrow \infty} a_n = x_0 = \lim_{n \rightarrow \infty} b_n \quad \text{和} \quad \lim_{n \rightarrow \infty} c_n = y_0 = \lim_{n \rightarrow \infty} d_n$$

然後，再用 F 的连续性去证明

$$(1.52) \quad F(x_0, y_0) = g.l.b.(S)$$

所以 $F(x_0, y_0)$ 是 S 中的极小者。我们将用反证法来证明 (1.52)-式。假若不然，令

$$(1.53) \quad F(x_0, y_0) - g.l.b.(S) = 2\varepsilon$$

则用 F 在 (x_0, y_0) 点的连续性, 即有一个足够小的 $\delta > 0$, 使得

$$(1.54) \quad |x - x_0| < \delta, |y - y_0| < \delta \Rightarrow |F(x, y) - F(x_0, y_0)| < \varepsilon$$

再者, 当 n 足够大时, 即有

$$(1.55) \quad |a_n - x_0|, |b_n - x_0|, |c_n - y_0|, |d_n - y_0|$$

四者都小于 δ , 亦即整个 \square_n 都包含在

$$(1.56) \quad |x - x_0| < \delta \text{ 和 } |y - y_0| < \delta$$

这个方块之内, 所以 F 在 \square_n 上的函数值都大于

$$(1.57) \quad F(x_0, y_0) - \varepsilon = g.l.b.(S) + \varepsilon$$

这显然和 \square_n 的选取是使得 F 在其上函数值的极大下界等于 $g.l.b.(S)$ 相矛盾。所以 $F(x_0, y_0)$ 必须等于 $g.l.b.(S)$, 亦即 $F(x_0, y_0)$ 就是 F 在 $\square(2K, 2K)$ 上的极小值。□

注意: 在上述证明中, 其实仅仅用到 F 的连续性和 F 的下界性 (即 $F \geq 0$)。

在代数基本定理的证明中, 不妨设 $f(z)$ 的首项系数为 1。

【引理 1.5】: 令 $f(z) = z^n + c_1 z^{n-1} + c_2 z^{n-2} + \cdots + c_n$ 及

$$M = \text{Max} \{|c_i|, 1 \leq i \leq n\}$$

$$K > 2(M + 2)$$

则 $|f(z)|$ 的极小值不能够取值于 $\square(2K, 2K)$ 的边界之上。

证明: 当我们把 z 限制在 $\square(2K, 2K)$ 的边界上时,

$$\begin{aligned} |f(z)| &= |z|^n \cdot \left| 1 + c_1 \frac{1}{z} + \cdots + c_n \frac{1}{z^n} \right| \\ &\geq |z|^n \left\{ 1 - M \left(\frac{1}{|z|} + \cdots + \frac{1}{|z|^n} \right) \right\} \\ &\geq K^n \left\{ 1 - \frac{M}{2M+3} \right\} \\ &\geq [2(M+2)]^n \cdot \frac{M+3}{2M+3} > |c_n| = |f(0)| \end{aligned}$$

由此可见 F 在 $\square(2K, 2K)$ 上的极小点 (x_0, y_0) 不可能位于其边界之上，亦即 (x_0, y_0) 必定是一个内点。 \square

【引理 1.6】：当 $|f(z)|$ 在 $\square(2K, 2K)$ 上的极小值发生于 $\square(2K, 2K)$ 的一个内点 $z_0 = x_0 + iy_0$ 时，即有 $|f(z_0)| = 0$ ，亦即 $f(z_0) = 0$ 。

证明：我们将用反证法，设 $f(z_0) = w_0 \neq 0$ 然後去证明一定有 $\square(2K, 2K)$ 中的另外一点 z_1 使得 $|f(z_1)| < |f(z_0)|$ 。

因为 z_0 是 $\square(2K, 2K)$ 的一个内点。所以以 z_0 为圆心的一个足够小半径的圆依然位于 $\square(2K, 2K)$ 之内。其上之点可以表成

$$z = z_0 + \rho(\cos \theta + i \sin \theta), \quad \rho \text{ 是半径}$$

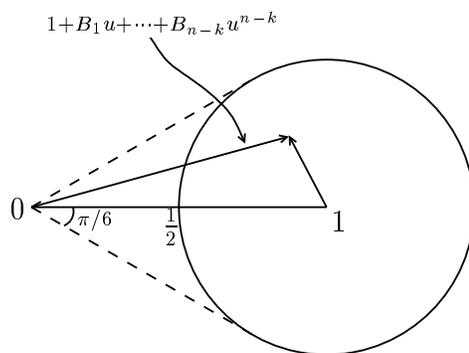
再者，以 $z = z_0 + u$ 代入 $f(z)$ 然後再以 u 的升幂表达，即有

$$(1.58) \quad f(z) = f(z_0) + A \cdot u^k \{1 + B_1 u + B_2 u^2 + \cdots + B_{n-k} u^{n-k}\}$$

令 $M' = \text{Max} \{|B_i|, 1 \leq i \leq n-k\}$ ，取 $\rho < \frac{1}{2(M'+2)}$ ，则有

$$(1.59) \quad \begin{aligned} & |B_1 u + B_2 u^2 + \cdots + B_{n-k} u^{n-k}|, \quad u = \rho(\cos \theta + i \sin \theta) \\ & \leq M' \cdot \{\rho + \rho^2 + \cdots + \rho^{n-k}\} \\ & \leq \frac{M'}{2M'+3} < \frac{1}{2} \end{aligned}$$

由 [图 1-4] 可见 $\{1 + B_1 u + \cdots + B_{n-k} u^{n-k}\}$ 的幅角 φ 必然介于 $\pm \frac{\pi}{6}$ 之间：



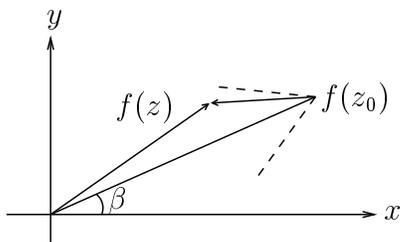
[图 1-4]

设 A 的幅角为 α , 则

$$(1.60) \quad Au^k \cdot \{1 + B_1u + \cdots + B_{n-k}u^{n-k}\}, \quad u = \rho(\cos \theta + i \sin \theta)$$

的幅角等于

$$(1.61) \quad \alpha + k\theta + \varphi, \quad |\varphi| \leq \frac{\pi}{6}$$



[图 1-5]

设 $f(z_0)$ 的幅角为 β 。取

$$(1.62) \quad \theta = \frac{\pi + \beta - \alpha}{k}$$

则有

$$(1.63) \quad \alpha + k\theta + \varphi = \pi + \beta + \varphi, \quad |\varphi| \leq \frac{\pi}{6}$$

由 [图 1-5] 易见 $f(z) - f(z_0)$ 的方向必定夹在 $\pi + \beta \pm \frac{\pi}{6}$ 之间, 所以 $|f(z)| < |f(z_0)|$, 和所设 $|f(z_0)|$ 是极小者相矛盾。这也就证明了 $|f(z_0)|$ 必须是 0, 亦即 z_0 乃是 $f(z)$ 的一个根! \square

【定理 1.7】(代数基本定理, Fundamental Theorem of Algebra): 任何一个非常数复系数的多项式 $f(z)$ 都必定有一个零点, 亦即存在 $z_0 \in \mathbb{C}$ 使得 $f(z_0) = 0$ 。

【推论 1】: 任给一个 n 次复系数多项式 $f(x) \in \mathbb{C}[x]$ 都可以分解成 n 个复系数一次因式之乘积。

【推论 2】: 任给一个 n 次实系数多项式 $f(x) \in \mathbb{R}[x]$ 都可以分解成一次或二次实系数多项式的乘积。

关于连续性教学的探讨：

连续性是整个分析学的基础之所在。但是长期以来，连续性却一直是分析学在学与教上的难点所在，亟待改进。本章对于实数系和函数的连续性先作一次返璞归真、探本究源的分析。实数系的发现与产生，起源于古希腊定量几何基础论对于长度的度量的深入研讨。由 Hippasus 不可公度性的发现，到 Eudoxus 引入长度的比较原则和逼近法达成定量几何基础论之重建这一段引人入胜的史话，理当是每一位学习、理解连续性的後之来者自然的入门和坦途。由此即可顺理成章地认识到左、右夹逼数列的「被夹逼者」的存在性实乃直线「连续不断但一剪就断」这种直观明显的特性的解析描述，故称之为实数的连续性，它是其他各种各样存在定理论证之所本！

有鉴于「理在用中方知妙」这个认知上的规律，学习、理解连续性当然要从它的各种各样的用场和用法中逐步体会，加深认识。本章所讨论的闭线段上的连续函数的几个基本定理和实变多项函数的 Sturm 定理及复变多项函数的代数基本定理，它们不但是连续性的重要用场，而且其证明则又是学会连续性的用法的自然途径。

为了便于读者体认连续性的用法，本章对于上述这几个具有基本重要性的定理的证法，力求简朴、划一；基本上可以归于两种：其一是直截了当地由所设去构造那个所要证明其存在性的左、右夹逼数列，从而由实数的连续性直接推论所要证的存在性（例如[定理 1.1]和[定理 1.2]的证明）。其二是采用反证法，我们由其否定命题出发，去构造一个左、右夹逼数列，使得其所夹逼之点和所设矛盾（例如[引理 1.1]和[定理 1.3]的证明）。再者，在两种情况，其左、右夹逼数列的构造法基本上都可以用二分（或四分）逼近法去达成。

【例题与习题】：I. 二分法与连续性之配合使用：

如上所述，本章在运用实数连续性来证明各种各样存在性定理时，基本上可以分为采取直接证法和反证法这样两种途径。但是不论采用前者或後者，在连续性的运用上总是采用二分逼近法去构造论证所需的左、右夹逼数列。对它们所用的二分逼近构造法略加分析，即可归纳出下述「通法」：若有一种闭线段上的性质 P 具有下述「继承性」

，亦即当把具有性质 P 的一个闭线段等分成两段，则其中至少有一个半线段依然具有性质 P 。由此即可逐次二分，每次取其依然保有性质 P 的半段。这样就按保有性质 P 为准则，构造而得逐次减半的闭线段列 $\{[a_n, b_n]\}$ ，它们的端点 $\{a_n\}$ 和 $\{b_n\}$ 乃是某一个实数 λ 的左、右夹逼数列，即 $a_n \rightarrow \lambda \leftarrow b_n$ 。易见如此所得的 λ 的任意小邻域，即对任给 $\delta > 0$ ， $(\lambda - \delta, \lambda + \delta)$ 总是包含著足够大 n 的闭线段 $[a_n, b_n]$ 于其中。

由此可见，在运用二分逼近法和连续性去论证某种存在性定理时，如何选择适当的具有继承性的 P 乃是要点之所在。

【习题一】：试分析本章中各定理论证中所用的性质 P ，逐一列述之。

II. 逼近和数列极限：

在 Eudoxus 首创逼近法时，他用分数 $\frac{m}{n}$ 和 $\frac{m+1}{n}$ 去左、右夹逼一个给定的不可公度比 $a:b$ 。用现代的术语来说，其所用的 $\frac{m}{n}$ 和 $\frac{m+1}{n}$ 乃是 $a:b$ 的一对近似值 (approximate values)，把所要逼近者 $a:b$ 左、右夹逼于其间。因为两者之差是 $\frac{1}{n}$ ，所以两者和 $a:b$ 的差别显然都小于 $\frac{1}{n}$ 。当 n 无限增大时，上述逐次求得的近似值构成两个数列 $\{a_n\}$ 和 $\{b_n\}$ ，它们始终把所要逼近者夹逼于其间，而 $b_n - a_n$ 又可以小到任意小，从而用夹逼数列唯一地确定了所要逼近者。用现代的术语来说，它就是两个夹逼数列的共同极限值是也。由此可见，逼近和极限乃是同一件事情的两面。若业已给定一个有待明确或研讨者（例如不可公度比 $a:b$ ，或一个给定函数在某点的变率，或某一区间上的积分等等），我们设法用一系列比较简单明了的「近似值」去逼近它，从而化未知为已知，这就是逼近法。反之，若由某种构造业已给出一个数列 $\{a_n\}$ ，我们要去确立是否存在著一个被其所无限逼近者（亦即极限值），这就是极限 (Limit)。所以逼近和极限实乃同一事物的两面，但是出发点和著重点是不同的。前者的要点在于其极限值的唯一性，而後者的要点则在于其极限值的存在性。下述几个习题，乃是关于数列极限的存在性的检验上常用的基本工具。它们的证明当然又得运用实数系连续性，所以它们的证明和第三节的几个证法大同小异，乃是同学们练习体会这种简朴自然的证法的好地方。

数列极限之定义：当 n 无限增大，数列 $\{a_n\}$ 趋于某一定值 λ 为极限乃是当 n 足够大时， $|a_n - \lambda|$ 可以小到任意小。其逻辑叙述即为：对于任

给 $\varepsilon > 0$ 皆有足够大之 N 使得 $n > N$ 时 $|a_n - \lambda| < \varepsilon$.

【习题二】：

- (1) 试证：一个递增上有界的数列 $\{a_n\}$ 的极限恒存在。同理，一个递减下有界的数列 $\{b_n\}$ 的极限恒存在。

[注]：前者的极限值实乃 $S = \{a_n\}$ 的极小上界，而後者的极限值实乃 $S' = \{b_n\}$ 的极大下界。所以上述存在性的证明基本上就是 [引理二] 的证明。

- (2) 试证一个具有极限值的数列 $\{a_n\}$ 必然是有界的。(所以其各项所成的集合 S 具有其极小上界和极大下界。)
- (3) 设 $\{a_n\}$ 是一个上、下有界的数列。令 $S_k = \{a_n, n \geq k\}$ ，而 M_k 和 m_k 分别是 S_k 的极小上界和极大下界。试证 $\{M_k\}$ 和 $\{m_k\}$ 分别是递减和递增的数列，而且

$$m_1 \leq \dots \leq m_k \leq m_{k+1} \leq \dots \leq M_{k+1} \leq M_k \leq \dots \leq M_1$$

- (4) 结合习题 (1) 和 (3) 的结果得知 $\lim_{k \rightarrow \infty} m_k$ 和 $\lim_{k \rightarrow \infty} M_k$ 皆存在而且 $\lim_{k \rightarrow \infty} m_k \leq \lim_{k \rightarrow \infty} M_k$ 。试证 $\lim_{k \rightarrow \infty} m_k = \lim_{k \rightarrow \infty} M_k$ 乃是 $\lim_{n \rightarrow \infty} a_n$ 的存在性的充要条件。

[注]：通常称 $\lim_{k \rightarrow \infty} m_k$ 为数列 $\{a_n\}$ 的 \liminf ，而 $\lim_{k \rightarrow \infty} M_k$ 则称之为 $\{a_n\}$ 的 \limsup 。

- (5) 设 $\{a_n\}$ 是一个上、下有界的数列，则可取其适当的子数列 $\{a_{n_i}\}$ 其极限值为 $\liminf\{a_n\}$ (或 $\limsup\{a_n\}$)。
- (6) 若数列 $\{a_n\}$ 满足条件：

给定实数 $\varepsilon > 0$ ，恒存在正整数 N 使得
当 $n, m \geq N$ 时，恒有 $|a_n - a_m| < \varepsilon$

则称之为满足 Cauchy 条件的数列，或 Cauchy 数列。试证 Cauchy 条件乃是 $\lim_{n \rightarrow \infty} a_n$ 的存在性的充要条件。

- (7) 令 $a_n = (1 + \frac{1}{n})^n$, $b_n = (1 + \frac{1}{n})^{n+1}$ 。试证 $\{a_n\}$ 为递增数列, $\{b_n\}$ 为递减数列, 而且 $\lim_{n \rightarrow \infty} a_n = \lim_{n \rightarrow \infty} b_n$ 。
- (8) 设 a 是一个大于 1 的给定实数, n 是一个大于 1 的给定整数, 试证存在唯一的正实数 x_0 , 记为 $\sqrt[n]{a}$, 它的 n 次方等于 a 。
- (9) 试证 $\lim_{n \rightarrow \infty} \sqrt[n]{a} = 1$ 。
- (10) 试证 $\lim_{n \rightarrow \infty} \sqrt[n]{n} = 1$ 。
- (11) 试用习题 (5) 的结果给出[定理 1.3]的另一种证明。

第二章

微积分——微分与积分的直观内涵 与基础理论

微积分的基础和基本方法就是定量几何基础论中开始认识的实数系连续性和 Eudoxus 所首创的逼近原理与方法。再者，积分的想法及其几何学中的应用的范例，业已在 Eudoxus 求锥体体积，Archimedes 在求球面面积和刘徽、祖暅在求球体体积等工作中充分体现。但微分的想法则一直到了 Kepler 研究行星运行，解析几何中求切线斜率 (Descartes, Fermat) 才逐渐展现。总之，微积分学乃是上千年世代相承，逐步演进而得的成果，及至十七世纪後半，业已水到渠成，由 Newton 和 Leibniz 分别从两种观点集成其初步的体系者也（通常常有 Newton 和 Leibniz 创造微积分这种流传，乃是和史实有相当差距的）。

2.1 变率与微分

从数理分析的观点来说，一个变动的事物或现象，含有各种各样的参变量 (parameters)，而其所含的诸多参变量之间又具有相互关联的关系 (interlocking relationships)。其中最为简单而且基本的关系乃是两个参变量 $\{x, y\}$ ，後者随著前者之取定而唯一确定的关系，我们通常把这种最为简单基本的关系叫做 y 是 x 的一个函数。例如一个圆的面积乃是其半径的函数，即 $A = \pi r^2$ ；一个球的体积乃是其半径的函数，即 $V = \frac{4\pi}{3} r^3$ ，等等。由此可见，微积分学的基本题材就是一个实变数 x 的各种各样函数的基本性质。例如：

- (i) $y = ax^2 + bx + c$, $y = x^3 - 5x + 6$, ... 等等 y 可以用 x 的一个多项式加以表达的多项式函数 (包括常数函数 $y = c$ 为其特例)。
- (ii) $y = \sin x$, $y = \cos x$, $y = \tan x$ 等三角函数。
- (iii) 一个在 x -轴上以等速 c 行进的动点 P , 其 x -坐标乃是时间 t 的函数, 即 $x = a + ct$, 其中 a 是起点的坐标。
- (iv) 以 $(1, 0)$ 为起点在单位圆上作单位角速度的运动可以用下述一对函数表达之, 即

$$x = \cos t, \quad y = \sin t$$

- (v) 一个在空间运动之点 P , 其 x, y, z -坐标分别都是时间 t 的函数, 即

$$x = f_1(t), \quad y = f_2(t), \quad z = f_3(t)$$

由此可见, 一个点在空间的运动可以由上述三个时间的函数所组成者描述之。

概括地说, 一个函数 $y = f(x)$ (或 $x = g(t)$) 所描述者, 乃是一个变数的值如何随著另一个变数的变化而变化的方式。例如例-(iii) 中的 $x = a + ct$, 当 t 由 t_1 变到 t_2 时, x 的值由 $a + ct_1$ 变到 $a + ct_2$, 其中「 t 的改变」是 $(t_2 - t_1)$, 而相应的「 x 的改变」则是

$$(a + ct_2) - (a + ct_1) = c \cdot (t_2 - t_1)$$

两者的比值就是上述直线等速运动的速度 c 。从函数的观点来说, 上述简单的事实就是:

一个一次多项式函数 $y = a + mx$, 其应变数 y 的改变恒等于其自变数 x 的改变的 m 倍, 亦即

$$\frac{y_2 - y_1}{x_2 - x_1} = \frac{(a + mx_2) - (a + mx_1)}{x_2 - x_1} \equiv m$$

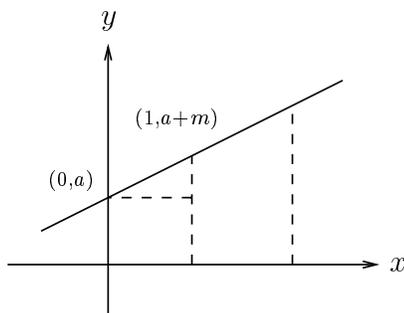
称之为: 其变率恒等于 m 。反之, 设 $y = f(x)$ 的变率恒等于 m , 亦即

$$\frac{f(x) - f(0)}{x - 0} = m$$

对于任何非零的 x 恒成立，则有

$$y = f(x) = f(0) + mx$$

若用图象来表现上述一次多项式函数，则其图象乃是一条斜率为 m ， y -轴的截距为 $a = f(0)$ 的直线：



[图 2-1]

当 $y = f(x)$ 不是一次多项式函数时（例如 $y = x^2$, $y = \sin x$ 等等）， $(f(x_2) - f(x_1))$ 和 $(x_2 - x_1)$ 的比值是随著 (x_2, x_1) 的不同而有所不同的。我们应该如何把此类函数在某一点的「变率」赋以恰当的定义呢？在此，我们大可效法 Eudoxus 的做法，即运用逼近思想来达成以简御繁，以「已知」去理解「未知」，亦即以已知的一次函数的变率去夹逼一个给定的非一次函数在某一给定点的有待定义的「变率」。首先我们得探讨一个合乎变率的直观内含的比较原则。

设有甲、乙两人在一条公路上作自行车竞赛，分别以 $f(t)$ 和 $g(t)$ 表示甲、乙在 t -秒後和起点的距离。设甲在 $t = t_0$ 时从後赶上乙而且超越之，亦即在 t 略小于 t_0 时，甲在乙之後，但是在 t 略大于 t_0 时，则甲在乙之前。则甲在 $t = t_0$ 时的「速率」显然不能小于乙者。[要不然，则甲是不可能于 $t = t_0$ 时达成後来居上的，是不？]

把上述直观上十分明显的事实，改用函数框架叙述之，即为下述刻画变率的直观内含的比较原则：

【函数变率的比较原则】(Comparison principle of rate of change)：

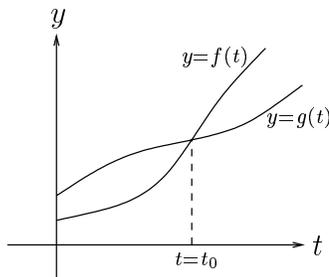
设有两个函数 $f(t)$ 和 $g(t)$ ，在 $t = t_0$ 的邻近具有下述大小关系，即对于任给足够小的 $\delta > 0$,

$$f(t_0 - \delta) < g(t_0 - \delta)$$

$$f(t_0 + \delta) > g(t_0 + \delta) \quad (\text{参见 [图 2-2]})$$

$$\text{而且 } f(t_0) = g(t_0)$$

则 $f(t)$ 在 t_0 点的「变率」就不小于 $g(t)$ 在 t_0 点的「变率」。



[图 2-2]

【例 1】：让我们且以 $f(t) = ct^n$, $n \geq 2$ 为例，用上述比较原则去研讨它在 $t = a$ 点的变率应该是什么。因为一次函数的变率乃是已知者，我们可以把 $f(t)$ 和 $g(t) = ca^n + m(t - a)$ 来作比较。令 $t = a + \delta$ ，则有

$$f(t) = c(a + \delta)^n = ca^n + nca^{n-1}\delta + \frac{n(n-1)}{2}ca^{n-2}\delta^2 + \dots + c\delta^n$$

$$g(t) = ca^n + m\delta$$

$$f(t) - g(t) = (nca^{n-1} - m)\delta + \delta^2 \left[\frac{n(n-1)}{2}ca^{n-2} + \dots + c\delta^{n-2} \right]$$

由此不难看到，当 $(nca^{n-1} - m) \neq 0$ 时，只要把 $|\delta|$ 取得足够小，则 $(f(t) - g(t))$ 和 $(nca^{n-1} - m)\delta$ 同号。所以由比较原则即有

$$m > nca^{n-1} \Rightarrow f(t) \text{ 在 } t = a \text{ 点的变率} < m$$

$$m < nca^{n-1} \Rightarrow f(t) \text{ 在 } t = a \text{ 点的变率} > m$$

所以「 $f(t)$ 在 $t = a$ 点的变率」唯有定义为 nca^{n-1} 才合乎上述比较原则。

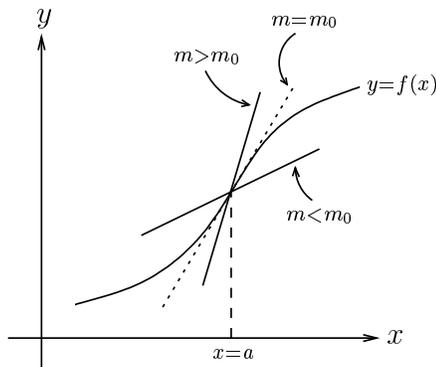
【例 2】：设有函数 $y = f(x)$ ，它在 $x = a$ 的邻近和一次函数 $g(x) = f(a) + m(x - a)$ 相比，满足下列不等式，即存在有一个数 m_0 ，当

$m > m_0$ (及 $m < m_0$)， δ 是足够小的正值时分别有下述比较关系 (如 [图 2-3] 所示)：

$$m > m_0 : f(a - \delta) > g(a - \delta), f(a + \delta) < g(a + \delta)$$

$$m < m_0 : f(a - \delta) < g(a - \delta), f(a + \delta) > g(a + \delta)$$

则由变率比较原则可得 $f(x)$ 在 $x = a$ 的变率不大于任何比 m_0 大者，而又不小于任何比 m_0 小者，因此它应该等于 m_0 ！



[图 2-3]

【变率的定义】：当函数 $y = f(x)$ 在 $x = a$ 的邻近满足[例 2]和一次函数的比较关系时，我们定义其在 $x = a$ 点的变率为 m_0 。不然，则称其在 $x = a$ 点的变率是不定义 (undefined)。

【符号】：把一个给定函数 $y = f(x)$ 在其变率有定义之点的变率逐一记录，即得一个由 $f(x)$ 引导而得的新函数，通常以 $y' = f'(x)$ 记号之，称其为 $y = f(x)$ 的导函数 (derivative)。[注意，导函数 $f'(x)$ 的定义域乃是 $f(x)$ 的定义域之中，那些变率有定义之点所构成的子集。它有可能是一个真子集，甚至在某些例子是一个空集。]

【例 3】：当 $f(x) = cx^n$ 时， $f'(x) = ncx^{n-1}$ 。

变率与极限：

对于一个给定函数 $f(x)$ 在某一定点 $x = a$ 的变率这个微分学的基本概念，我们采取了概念上探本究源的分析。先认识到一次函数乃是变率恒为常数的「简朴已知者」，然後再由变率的直观内涵的分析认清了变率的比较原则，最後再把两者结合起来，得出具有直观上的必然

性的「变率定义」。通过这样一种返璞归真的分析来探讨「变率」这个基本概念所应有的定义，其好处是一来使得定义的直观内含充分展现，唯有这样做才能真正体现定义的必然性和自然性，二来也使得读者有一次在基本概念层面上探索求知的体验。但是这种纯概念性的变率的定义，还要略加转换，使得它便于计算。

[分析]：

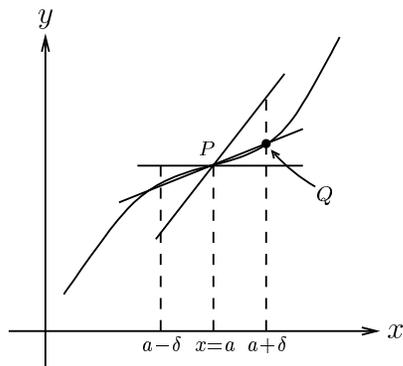
设 $y = f(x)$ 是一个满足[例 2]中所述的和一次函数的比较关系的函数， $\varepsilon > 0$ 是一个任给的正数。则存在一个足够小的 $\delta > 0$ ，使得 $y = f(x)$ 在 $|x - a| < \delta$ 之上的图象夹逼于过 P 点而斜率分别为 $m_0 \pm \varepsilon$ 的直线之间，如 [图 2-4] 所示。亦即

$$m_0 - \varepsilon < \frac{f(a+h) - f(a)}{h} < m_0 + \varepsilon$$

对于任给 $|h| < \delta$ 皆成立。改用极限的术语，也就是

$$\lim_{h \rightarrow 0} \frac{f(a+h) - f(a)}{h} = m_0$$

这也就是便于计算的函数变率的极限定义式。



[图 2-4]

由 [图 2-4] 可见， $\frac{f(a+h)-f(a)}{h}$ 的几何意义乃是割线 PQ 的斜率，而 m_0 则是图象 $y = f(x)$ 在 P 点切线的斜率。

【例 4】： $f(x) = cx^n$

$$\frac{f(a+h) - f(a)}{h} = nca^{n-1} + h \left[\frac{n(n-1)}{2} ca^{n-2} + \dots + ch^{n-2} \right]$$

所以 $\lim_{h \rightarrow 0} \frac{f(a+h) - f(a)}{h} = nca^{n-1}$ 。

【例 5】： $f_1(x) = \sin x$, $f_2(x) = \cos x$ ，则有

$$f_1'(x) = \cos x, \quad f_2'(x) = -\sin x$$

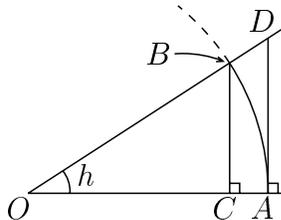
证明：由 $\sin x$, $\cos x$ 的和角公式易得

$$\begin{aligned} \frac{\sin(a+h) - \sin a}{h} &= \frac{\sin a(\cos h - 1)}{h} + \frac{\cos a \sin h}{h} \\ \frac{\cos(a+h) - \cos a}{h} &= \frac{\cos a(\cos h - 1)}{h} - \frac{\sin a \sin h}{h} \end{aligned}$$

因此，我们只需要在此补证下述两个基本极限，即

【引理 2.1】： $\lim_{h \rightarrow 0} \frac{\sin h}{h} = 1$, $\lim_{h \rightarrow 0} \frac{\cos h - 1}{h} = 0$

证明：



[图 2-5]

如 [图 2-5] 所示，扇形 OAB 的面积是 $\frac{1}{2}h$ ， $\triangle OCB$ 的面积是 $\frac{1}{2}\sin h \cos h$ ， $\triangle OAD$ 的面积是 $\frac{1}{2}\tan h$ 。所以有下述不等式：

$$\frac{1}{2}\sin h \cos h < \frac{1}{2}h < \frac{1}{2}\tan h = \frac{1}{2}\frac{\sin h}{\cos h}$$

由此即得：

$$\cos h < \frac{\sin h}{h} < \frac{1}{\cos h} = \sec h$$

当 $h \rightarrow 0$ 时， $\cos h \rightarrow 1$, $\sec h \rightarrow 1$ 。所以上述始终夹逼于 $\cos h$ 和 $\sec h$ 之间者，即 $\frac{\sin h}{h}$ ，也必然趋于 1 为其极限，亦即已证得：

$$\lim_{h \rightarrow 0} \frac{\sin h}{h} = 1$$

再者，

$$\begin{aligned}\frac{\cos h - 1}{h} &= \frac{\cos^2 h - 1}{h(\cos h + 1)} = \frac{-\sin^2 h}{h(\cos h + 1)} \\ &= -\frac{\sin h}{h} \cdot \frac{\sin h}{1 + \cos h}\end{aligned}$$

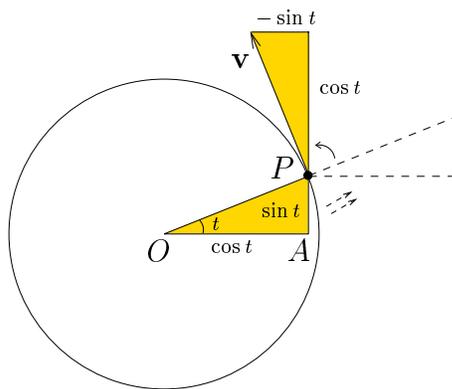
所以

$$\begin{aligned}\lim_{h \rightarrow 0} \frac{\cos h - 1}{h} &= \lim_{h \rightarrow 0} \left(-\frac{\sin h}{h} \right) \cdot \lim_{h \rightarrow 0} \frac{\sin h}{1 + \cos h} \\ &= -1 \cdot 0 = 0\end{aligned}$$

□

【例 6】：在[例 6]的两个公式，其实还可以用一个具有物理意义的方法来证明的。当动点 P 在单位圆 $x^2 + y^2 = 1$ 上作单位速率运动时，其 x, y -坐标表示式分别为：

$$x = \cos t, \quad y = \sin t$$



[图 2-6]

由几何的熟知事实，其速度向量是垂直于 \overrightarrow{OP} 的单位长向量，易见其 x, y 分量分别是 $-\sin t$ 和 $\cos t$ 。所以应该有：

$$\frac{d}{dt} \cos t = -\sin t, \quad \frac{d}{dt} \sin t = \cos t$$

$$\begin{aligned} \text{【定理 2.1】} : \quad & [f(x) + g(x)]' = f'(x) + g'(x) \\ & [f(x) \cdot g(x)]' = f'(x) \cdot g(x) + f(x) \cdot g'(x) \\ & \left[\frac{1}{f(x)} \right]' = -\frac{f'(x)}{[f(x)]^2} \end{aligned}$$

[注] 上述公式的意义是当 $f'(x)$ 和 $g'(x)$ 在该点有定义时, 则上述公式中的左侧函数在该点的变率也就有定义, 而且其值等于右侧所给者。

证明: 由定义的极限式易证:

$$\begin{aligned} & \lim_{h \rightarrow 0} \frac{(f(a+h) + g(a+h)) - (f(a) + g(a))}{h} \\ &= \lim_{h \rightarrow 0} \frac{f(a+h) - f(a)}{h} + \lim_{h \rightarrow 0} \frac{g(a+h) - g(a)}{h} \\ &= f'(a) + g'(a) \\ & \lim_{h \rightarrow 0} \frac{(f(a+h)g(a+h)) - (f(a)g(a))}{h} \\ &= \lim_{h \rightarrow 0} \frac{f(a+h) - f(a)}{h} \cdot \lim_{h \rightarrow 0} g(a+h) + f(a) \cdot \lim_{h \rightarrow 0} \frac{g(a+h) - g(a)}{h} \\ &= f'(a) \cdot g(a) + f(a) \cdot g'(a) \end{aligned}$$

再者, 因为

$$\frac{1}{h} \left\{ \frac{1}{f(a+h)} - \frac{1}{f(a)} \right\} = -\frac{1}{f(a+h) \cdot f(a)} \cdot \frac{1}{h} (f(a+h) - f(a))$$

所以

$$\begin{aligned} & \lim_{h \rightarrow 0} \frac{1}{h} \left\{ \frac{1}{f(a+h)} - \frac{1}{f(a)} \right\} \\ &= -\lim_{h \rightarrow 0} \frac{1}{f(a+h)f(a)} \cdot \lim_{h \rightarrow 0} \frac{f(a+h) - f(a)}{h} \\ &= -\frac{1}{f(a)^2} f'(a) \end{aligned}$$

【定理 2.2】: 设 $y = f(u)$ 和 $u = g(x)$ 在 $u_0 = g(x_0)$ 和 x_0 点皆为可微, 则有复合函数 $y = f(g(x))$ 在 x_0 点亦为可微, 而且其变率为 $f'(u_0)g'(x_0)$, 亦即对于复合函数 $f(g(x))$ 的微分, 有下述常用好用的公式, 即

$$f(g(x))' = f'(g(x)) \cdot g'(x)$$

证明：令 $\Delta u = g(x_0 + \Delta x) - g(x_0)$, $\Delta y = f(g(x_0 + \Delta x)) - f(g(x_0))$. 由所设易见当 $\Delta x \rightarrow 0$ 时, $\Delta u \rightarrow 0$, $\Delta y \rightarrow 0$. 让我们先讨论 $g'(x_0) \neq 0$ 的情形。当 Δx 足够小而且非零时, Δu 恒为非零, 所以即有

$$\frac{\Delta y}{\Delta x} = \frac{\Delta y}{\Delta u} \cdot \frac{\Delta u}{\Delta x} \Rightarrow \lim_{\Delta x \rightarrow 0} \frac{\Delta y}{\Delta x} = \lim_{\Delta u \rightarrow 0} \frac{\Delta y}{\Delta u} \cdot \lim_{\Delta x \rightarrow 0} \frac{\Delta u}{\Delta x} = f'(g(x_0)) \cdot g'(x_0)$$

请注意, 在 $g'(x_0) = 0$ 的情形, 当 Δx 足够小而且非零时, 是不能保证 Δu 恒为非零的。而在 $\Delta u = 0$ 的时候当然不能把 $\frac{\Delta y}{\Delta x}$ 改写成 $\frac{\Delta y}{\Delta u} \cdot \frac{\Delta u}{\Delta x}$ (不能除以 0!)。其实在这种特殊情形 Δy 根本就是 0, 又何必画蛇添足去改写它呢? 总之, 不论 Δu 是否为零,

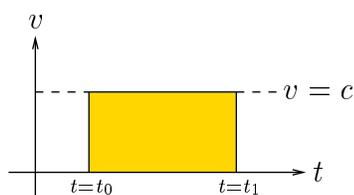
$$\lim_{\Delta x \rightarrow 0} \frac{\Delta y}{\Delta x} = 0 = f'(g(x_0)) \cdot g'(x_0)$$

在 $g'(x_0) = 0$ 的情形是依然成立的。□

2.2 总和与积分

从计算的角度来看, 上一节所讨论的微分 (亦即由给定函数 $f(x)$ 去求那个记录其逐点之变率的导函数 $f'(x)$) 是相当简单的一种比值为极限; 而积分则是一种无限分割求和的极限计算, 它显然要比微分计算要来得复杂得多。但是在人类理性文明发展史上, 积分的出现要比微分的起始早两千年。前者的首现就是在 Eudoxus 对于锥体体积公式的论证, 而後者则在 Kepler 研究行星运行和 Galileo 研究自由落体中才出现其雏形 (prototype)。如今回顾, 此事决非偶然, 耐人寻思。究其原因, 大概是因为那些激发无限分割求和的几何问题既自然、重要而且其难度上又富有挑战性; 如今反思, 要克服这种几何问题, 其实别无他途! 所以积分的想法, 自然而然地起源于几何学, 实非偶然。

要给积分求和这个基本概念赋以明确的定义, 自然又得用 Eudoxus 逼近原理。首先, 让我们来看一下那一种函数的「求和」是简单而且基本的。例如一列以 c 为其等速行进的火车在 $t = t_0$ 到 $t = t_1$ 这一段时间所走的总里程显然就是 $c \cdot (t_1 - t_0)$ 。若用函数的图象给以几何化的表达, 亦即常函数 $v = c$ 在区间 $[t_0, t_1]$ 上方如 [图 2-7] 所示的面积等于 $c \cdot (t_1 - t_0)$:

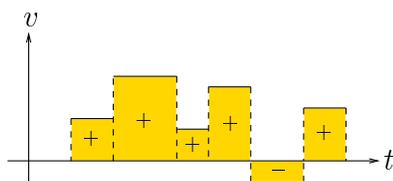


[图 2-7]

由此易见，一个常数函数 $f(t) = c$ 在区间 $[t_0, t_1]$ 上的积分乃是 $c(t_1 - t_0)$ 。而它相应的几何量，就是 [图 2-7] 所示的区域的面积。当然，我们还可以把这种最为简单的例子稍加推广，即

【例 7】：设 $v = f(t)$ 是一个分段常数函数，亦即

$$f(t) = \begin{cases} c_1, & t_0 \leq t < t_1 \\ c_2, & t_1 \leq t < t_2 \\ \vdots & \\ c_k, & t_{k-1} \leq t \leq t_k \end{cases}$$



[图 2-8]

则可以分段用常数函数求和而得其总和为

$$c_1(t_1 - t_0) + c_2(t_2 - t_1) + \dots + c_k(t_k - t_{k-1})$$

为了往後符号上的统一，我们将以积分符号表达上述分段求和之值，即当 $f(t)$ 是分段常数函数时，

$$\int_{t_0}^{t_k} f(t) dt = \sum_{i=1}^k \int_{t_{i-1}}^{t_i} c_i dt = \sum_{i=1}^k c_i (t_i - t_{i-1})$$

[因为分段常数函数的图象有如高高低低的阶梯，所以通常称之为阶梯函数 (step functions)。这种函数的积分就是上述简单明了的分段地以常

数函数求和，而它们也就是我们将要在积分层面，用来以简御繁的基本函数。]

由积分求和的直观内含，显然应该有下列比较原则，即

【积分求和的比较原则】(Comparison principle of sum of total effect)：

设 $f(x)$ 在区间 $[a, b]$ 上恒不小于 $g(x)$ ，则 $f(x)$ 在 $[a, b]$ 上的积分当然也不小于 $g(x)$ 在 $[a, b]$ 上的积分。以算式表达，即为

$$f(x) \geq g(x) \quad \forall a \leq x \leq b \Rightarrow \int_a^b f(x)dx \geq \int_a^b g(x)dx$$

在明确了积分求和的比较原则之后，自然又可以效法 Eudoxus 运用逼近原理，用一系列妥为选取的阶梯函数去上、下夹逼一个给定函数，从而确立它在一个给定区间 $[a, b]$ 上的积分所应有之值。这样也就可以给相当广泛的各种各样函数的积分，赋予严格的定义。其具体做法，述之如下：

设 $f(x)$ 是一个定义于有限闭区间 $[a, b]$ 上的函数。若存在有上、下夹逼阶梯函数列 $\{G_n(x)\}$ 和 $\{g_n(x)\}$ ，满足下列条件，即

$$\begin{aligned} g_1(x) \leq \dots \leq g_n(x) \leq g_{n+1}(x) \leq \dots \leq f(x) \leq \dots \\ \leq G_{n+1}(x) \leq G_n(x) \leq \dots \leq G_1(x) \end{aligned}$$

对于所有 $a \leq x \leq b$ 成立；而且

$$\int_a^b G_n(x)dx - \int_a^b g_n(x) \rightarrow 0$$

由积分的比较原则，我们所要确定的 $\int_a^b f(x)dx$ 乃是一个分别被下述递增及递减数列

$$\left\{ \int_a^b g_n(x)dx \right\} \quad \text{和} \quad \left\{ \int_a^b G_n(x)dx \right\}$$

夹逼于其间者，所以它必然就是那个唯一满足对于所有 n 皆有下述夹逼不等式者也，即

$$\int_a^b g_n(x)dx \leq \int_a^b f(x)dx \leq \int_a^b G_n(x)dx$$

【定义】：满足上述条件的函数 $f(x)$ 定义为：在区间 $[a, b]$ 上的可积函数 (integrable function)，而由上述夹逼数列所唯一确定之值则定义为 $f(x)$ 在区间 $[a, b]$ 上的积分，以符号 $\int_a^b f(x)dx$ 记之。

【定理 2.3】：设 $f(x)$ 是 $[a, b]$ 上的一个单调递增（或递减）函数，则它是可积的。

证明：递增和递减这两种情形的论证在本质上是一样的。兹讨论前者的证明如下：

将 $[a, b]$ 等分成 2^n 段，设其分点为 $\{a_i, 1 \leq i \leq 2^n - 1\}$, $a_0 = a$, $a_{2^n} = b$ 。令 $g_n(x)$ （及 $G_n(x)$ ）为在 $[a_{i-1}, a_i)$ 分段上取常数值 $f(a_{i-1})$ （及 $f(a_i)$ ）的阶梯函数。由 $f(x)$ 的单调递增性可知

$$g_n(x) \leq f(x) \leq G_n(x)$$

对于所有 $a \leq x \leq b$ 恒成立，而且有

$$\begin{aligned} \int_a^b G_n(x)dx - \int_a^b g_n(x)dx &= \left(\frac{b-a}{2^n}\right) \sum_{i=1}^{2^n} f(a_i) - \left(\frac{b-a}{2^n}\right) \sum_{i=0}^{2^n-1} f(a_i) \\ &= \frac{b-a}{2^n} (f(b) - f(a)) \rightarrow 0 \end{aligned}$$

这也就证明了 $f(x)$ 在 $[a, b]$ 上的可积性。 \square

【推论】：设 $f(x)$ 是 $[a, b]$ 上的分段单调函数，亦即存在有限个分点 $\{a_i, 1 \leq i \leq k-1\}$ ，使得 $f(x)$ 在每一个分区间 $[a_{i-1}, a_i]$ 之上都是单调函数，则 $f(x)$ 是可积的。

证明：由所设和 [定理 2.3]， $f(x)$ 在每一分区间 $[a_{i-1}, a_i]$ 之上都是可积的。把它们的可积条件之中的上、下夹逼阶梯函数列 $\{G_{i,n}(x)\}$ 和 $\{g_{i,n}(x)\}$, $1 \leq i \leq k$ ，组合起来，亦即当 $a_{i-1} \leq x < a_i$ 时

$$G_n(x) = G_{i,n}(x), \quad g_n(x) = g_{i,n}(x)$$

则易见 $\{G_n(x)\}$ 和 $\{g_n(x)\}$ 构成了 $f(x)$ 在 $[a, b]$ 之上的上、下夹逼阶梯函数列。所以 $f(x)$ 在全区间之上也是可积的，而且有

$$\int_a^b f(x)dx = \sum_{i=1}^k \int_{a_{i-1}}^{a_i} f(x)dx \quad (a_0 = a, a_k = b) \quad \square$$

[注]: 很多常用的函数都是分段单调的, 所以它们都是可积的。但是上述简单的论证仅仅证明其可积性, 亦即其在给定区间上的积分的唯一存在性。至于这种由两列项数无限增大的和式所左、右夹逼者的积分值通常是很难用这种和式来直接计算其极限值的。

【例 8】: 设 $f_1(x)$ 和 $f_2(x)$ 都是在 $[a, b]$ 之上可积者, c_1, c_2 是任给常数, 则 $c_1 f_1(x) + c_2 f_2(x)$ 也是在 $[a, b]$ 之上可积者而且

$$\int_a^b [c_1 f_1(x) + c_2 f_2(x)] dx = c_1 \int_a^b f_1(x) dx + c_2 \int_a^b f_2(x) dx$$

[证明留作习题]

【例 9】: 设 $f(x)$ 是分别在 $[a, b]$ 和 $[b, c]$ 之上可积者, 则 $f(x)$ 在 $[a, c]$ 之上也是可积者, 而且

$$\int_a^c f(x) dx = \int_a^b f(x) dx + \int_b^c f(x) dx$$

[证明留作习题]

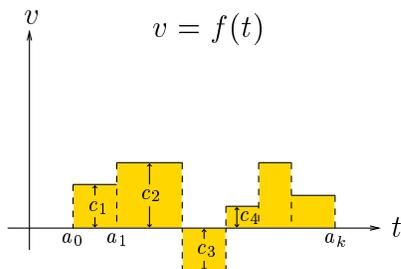
【例 10】: $\int_0^b x dx = \frac{1}{2} b^2$, $\int_0^b x^2 dx = \frac{1}{3} b^3$, $\int_0^b x^3 dx = \frac{1}{4} b^4$

[证明留作习题] (参看基础代数学第三章。)

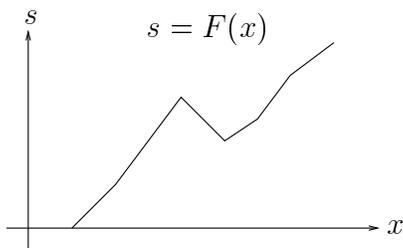
【例 11】: 设 $f(t)$ 是一个阶梯函数, 它在分区间 $[a_{i-1}, a_i)$ 上取常数值 c_i , $1 \leq i \leq k$ 。令

$$F(x) = \int_{a_0}^x f(t) dt, \quad a_0 \leq x \leq a_k$$

则 $F(x)$ 乃是一个分段一次函数, 亦即其图象为一折线, 它在分区间 $[a_{i-1}, a_i]$ 上者乃是一个斜率等于 c_i 的直线段。 $v = f(t)$ 和 $s = F(x)$ 的图象分别如 [图 2-9] 和 [图 2-10] 所示。



[图 2-9]



[图 2-10]

若用算式表达，则有

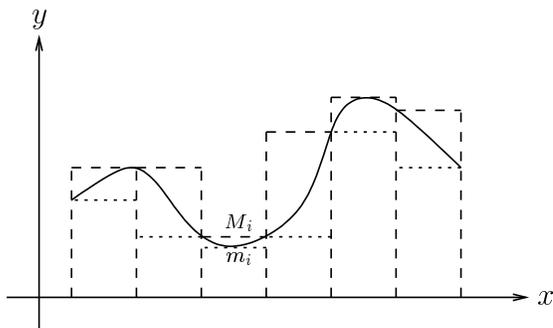
$$\begin{aligned}
 F(x) &= c_1(x - a_0), & a_0 \leq x \leq a_1 \\
 F(x) &= c_1(a_1 - a_0) + c_2(x - a_1), & a_1 \leq x \leq a_2 \\
 &\dots\dots\dots \\
 F(x) &= \sum_{j=1}^{i-1} c_j(a_j - a_{j-1}) + c_i(x - a_{i-1}), & a_{i-1} \leq x \leq a_i \\
 &\dots\dots\dots
 \end{aligned}$$

2.3 微积分基本定理与均值定理

【定理 2.4】：设 $f(x)$ 是在 $[a, b]$ 上的连续函数，则 $f(x)$ 是在 $[a, b]$ 上可积的。

证明：把 $[a, b]$ 等分为 2^n 个子区间。令 M_i, m_i 分别是 $f(x)$ 在第 i 个子区间 $[a_{i-1}, a_i]$ 上的极大、极小值。定义 $G_n(x)$ 和 $g_n(x)$ 为分别在 $[a_{i-1}, a_i]$ 上取常数值 M_i 和 m_i 的阶梯函数，则有

$$\begin{aligned}
 g_n(x) \leq g_{n+1}(x) \leq \dots \leq f(x) \leq \dots \leq G_{n+1}(x) \leq G_n(x) \\
 \int_a^b G_n(x) dx - \int_a^b g_n(x) dx = \frac{b-a}{2^n} \sum_{i=1}^{2^n} (M_i - m_i)
 \end{aligned}$$



[图 2-11]

再者，设 $\varepsilon > 0$ 是一个任给正数。由 $f(x)$ 的均匀连续性得知存在一个足够小的 $\delta > 0$ ，使得

$$|x_1 - x_2| < \delta \Rightarrow |f(x_1) - f(x_2)| < \varepsilon$$

因此, 只要把 n 取得足够大使得 $2^n > \frac{b-a}{\delta}$, 则所有 $(M_i - m_i)$ 都小于 ε 。亦即

$$\int_a^b G_n(x)dx - \int_a^b g_n(x)dx < \varepsilon \cdot (b-a)$$

所以 $f(x)$ 是在 $[a, b]$ 上可积的, 而 $\int_a^b f(x)dx$ 就是被左、右夹逼数列 $\{\int_a^b g_n(x)dx\}$ 和 $\{\int_a^b G_n(x)dx\}$ 所夹逼者也。□

【定理 2.5】(微积分基本定理, Fundamental Theorem of Calculus) :

设 $f(x)$ 为 $[a, b]$ 上的连续函数。令

$$F(x) = \int_a^x f(t)dt$$

则 $F'(x) = f(x)$ 。

证明: 设 x_0 为 $[a, b]$ 中任取一点。令 $M(x_0, h)$ 和 $m(x_0, h)$ 分别是 $f(x)$ 在 $[x_0 - h, x_0 + h]$ 上的极大、极小值, 其中 h 是一个足够小的正数 (在 $x = a$ 或 b 的特殊情形, 则上述邻域改用 $[a, a + h]$ 或 $[b - h, b]$)。则有

$$hm(x_0, h) \leq \left\{ \begin{array}{l} \int_{x_0}^{x_0+h} f(x)dx \\ \int_{x_0-h}^{x_0} f(x)dx \end{array} \right\} \leq hM(x_0, h)$$

亦即

$$m(x_0, h) \leq \left\{ \begin{array}{l} \frac{1}{h}(F(x_0 + h) - F(x_0)) \\ \frac{1}{-h}(F(x_0 - h) - F(x_0)) \end{array} \right\} \leq M(x_0, h)$$

再者, 由 $f(x)$ 在 x_0 点的局部连续性, 易见

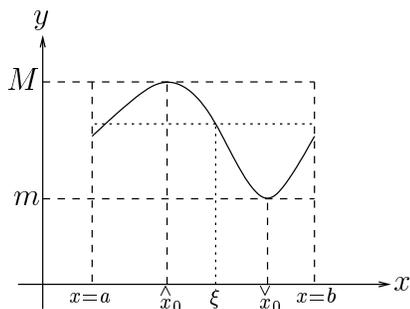
$$\lim_{h \rightarrow 0} m(x_0, h) = f(x_0) = \lim_{h \rightarrow 0} M(x_0, h)$$

所以 $F'(x_0) = f(x_0)$ 对于任给 $x_0 \in [a, b]$ 皆成立。□

【定理 2.6】(积分均值定理) : 设 $f(x)$ 为在 $[a, b]$ 上连续者, 则存在 a, b 之间的一点 ξ 使得

$$\frac{1}{b-a} \int_a^b f(x)dx = f(\xi)$$

证明：令 M, m 分别是 $f(x)$ 在 $[a, b]$ 上的极大、极小值。



[图 2-12]

如 [图 2-12] 所示，

$$m(b-a) \leq \int_a^b f(x)dx \leq M(b-a)$$

亦即

$$(f(\check{x}_0) =) m \leq \frac{1}{b-a} \int_a^b f(x)dx \leq M (= f(\hat{x}_0))$$

再用中间值定理即得 \check{x}_0 和 \hat{x}_0 之间的一点 ξ 使得

$$f(\xi) = \frac{1}{b-a} \int_a^b f(x)dx \quad \square$$

设 $G(x)$ 是一个以 $f(x)$ 为其导函数者，则上述积分均值定理又可以改写成 $G(x)$ 的微分均值定理，即

$$G(b) - G(a) = \int_a^b f(x)dx = (b-a)G'(\xi)$$

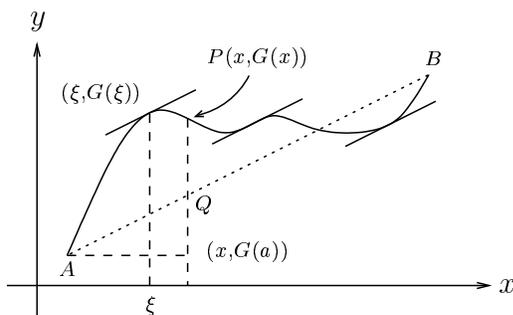
亦即

$$G(b) = G(a) + (b-a)G'(\xi)$$

在此，我们假设了 $G'(x) = f(x)$ 在 $[a, b]$ 上的连续性。有鉴于上述微分均值定理的基本重要性，我们在此再给以独立的证明，而且这样做的一个优点是它只需要 $G'(x)$ 在 (a, b) 中的每点皆能定义，从而摆脱了 $G'(x)$ 在 $[a, b]$ 上的连续这个过强的要求。

【定理 2.7】（微分均值定理）：设 $G(x)$ 在 $[a, b]$ 上连续，而且 $G'(x)$ 在 (a, b) 中的每点皆能定义（亦即可微），则在 (a, b) 中存在一点 ξ 使得

$$G(b) = G(a) + (b - a)G'(\xi)$$



[图 2-13]

证明：如 [图 2-13] 所示，上述公式的几何意义乃是 $G(x)$ 的图象在 $(\xi, G(\xi))$ 点的切线和连结 $A(a, G(a))$ 和 $B(b, G(b))$ 两点的割线互相平行。从几何直观来看，不难想到那些使得 \overrightarrow{QP} 极大或极小之点，其切线应该和 \overline{AB} 平行。再者，不难看出

$$\overrightarrow{QP} = G(x) - G(a) - \frac{G(b) - G(a)}{b - a}(x - a)$$

令其为 $H(x)$ ，则有

$$H(a) = H(b) = 0$$

而且 $H(x)$ 也是在 $[a, b]$ 上连续，在 (a, b) 中每点皆为可微者。再者，除非 $H(x) \equiv 0$ （亦即 $G'(x) \equiv \frac{G(b) - G(a)}{b - a}$ ）， $H(x)$ 必然有其极大或极小值不等于 0。设其取非零的极值之点为 ξ ，即有

$$H'(\xi) = G'(\xi) - \frac{G(b) - G(a)}{b - a} = 0$$

亦即

$$G(b) = G(a) + (b - a)G'(\xi) \quad \square$$

【推论 1】：设 $f(x)$ 为 $[a, b]$ 上的连续函数而且 $f'(x) \equiv 0$ ，则 $f(x)$ 为一个常数函数。

证明：令 c 为 $(a, b]$ 中任何一点，则由微分均值定理知道存在一个 $\xi \in (a, c)$ 使得

$$f(c) = f(a) + (c - a)f'(\xi) = f(a)$$

由此可见 $f(x)$ 必为常数函数。 \square

【推论 2】：设 $f(x)$ 为 $[a, b]$ 上的连续函数而 $G'(x) = f(x)$ ，则

$$\int_a^b f(t)dt = G(b) - G(a)$$

证明：由所设

$$\left[G(x) - \int_a^x f(t)dt \right]' = G'(x) - f(x) \equiv 0$$

所以 $G(x) - \int_a^x f(t)dt$ 乃是一个常数函数，亦即存在一个常数 C 使得

$$\begin{aligned} G(x) &= \int_a^x f(t)dt + C \Rightarrow C = G(a) \\ \Rightarrow \int_a^b f(t)dt &= G(b) - C = G(b) - G(a) \end{aligned} \quad \square$$

【推论 3】（部分积分公式）：设 $f'(x)$ 和 $g'(x)$ 都是在 $[a, b]$ 上连续者，则有

$$\int_a^b f(x)g'(x)dx = [f(x)g(x)]_a^b - \int_a^b f'(x)g(x)dx$$

$$\begin{aligned} \text{证明} : \int_a^b f(x)g'(x)dx + \int_a^b f'(x)g(x)dx \\ = \int_a^b [f(x)g(x)]'dx = f(b)g(b) - f(a)g(a) \end{aligned} \quad \square$$

【定理 2.8】（高阶微分均值定理，亦即泰勒公式）：

设 $f(x)$ 的 k -阶导函数 $f^{(k)}(x)$ 在 $[a, b]$ 上到处连续，则存在一个介于 a, b 之间的 ξ ，使得

$$\begin{aligned} f(b) &= f(a) + (b - a)f'(a) + \frac{(b - a)^2}{2!}f''(a) + \dots \\ &\quad + \frac{(b - a)^{k-1}}{(k - 1)!}f^{(k-1)}(a) + \frac{(b - a)^k}{k!}f^{(k)}(\xi) \end{aligned}$$

证明：由

$$f(b) = f(a) + \int_a^b f'(x)dx$$

起始，逐次用部分积分公式，即有

$$\begin{aligned} f(b) &= f(a) + (b-a)f'(a) + \int_a^b (b-x)f''(x)dx \\ &= f(a) + (b-a)f'(a) + \frac{1}{2!}(b-a)^2 f''(a) + \int_a^b \frac{(b-x)^2}{2!} f'''(x)dx \\ &\dots\dots\dots \\ &= f(a) + (b-a)f'(a) + \frac{1}{2!}(b-a)^2 f''(a) + \dots \\ &\quad + \frac{(b-a)^{k-1}}{(k-1)!} f^{(k-1)}(a) + \int_a^b \frac{(b-x)^{k-1}}{(k-1)!} f^{(k)}(x)dx \end{aligned}$$

令 M, m 分别是 $f^{(k)}$ 在 $[a, b]$ 上的极大值和极小值。由比较原则，即有

$$m \int_a^b \frac{(b-x)^{k-1}}{(k-1)!} dx \leq \int_a^b \frac{(b-x)^{k-1}}{(k-1)!} f^{(k)}(x) dx \leq M \int_a^b \frac{(b-x)^{k-1}}{(k-1)!} dx$$

亦即

$$m \frac{(b-a)^k}{k!} \leq \int_a^b \frac{(b-x)^{k-1}}{(k-1)!} f^{(k)}(x) dx \leq M \frac{(b-a)^k}{k!}$$

由上述不等式再用中间值定理，即得一个介于 a, b 之间的 ξ ，使得

$$\int_a^b \frac{(b-x)^{k-1}}{(k-1)!} f^{(k)}(x) dx = \frac{(b-a)^k}{k!} f^{(k)}(\xi) \quad \square$$

【例题与习题】：

本章所讨论的微积分基础理论乃是整个分析学的基础和精要之所在。它广泛的应用和深厚的发展可以说是无限的。下面所列述的例、习题只不过略述其一、二，以供读者之初步练习者也。

I. 高阶密切多项曲线：

一个函数 $y = f(x)$ 在 x_0 点可微 (differentiable) 的几何化描述乃是其图象在 $(x_0, f(x_0))$ 点具有切线，其切线方程式即为：

$$y = f(x_0) + f'(x_0)(x - x_0)$$

它乃是图象上两点 $P_i(x_i, f(x_i))$, $i = 0, 1$, 的割线 P_0P_1 在 $x_1 \rightarrow x_0$ 的极限位置。

设 $y = f(x)$ 在 x_0 点的邻近 k -阶连续可微 (亦即 $f^{(k)}(x)$ 在 x_0 点邻近存在而且连续)，则由泰勒公式可见下述多项函数

$$g(x) = f(x_0) + f'(x_0)(x - x_0) + \frac{f''(x_0)}{2!}(x - x_0)^2 + \dots + \frac{f^{(k)}(x_0)}{k!}(x - x_0)^k$$

和 $f(x)$ 在 $|x - x_0|$ 相当小时，其差额乃是一个比 k -阶微量 $|x - x_0|^k$ 还要小得多者也，亦即

$$|f(x) - g(x)| = \frac{(x - x_0)^k}{k!} |f^{(k)}(\xi) - f^{(k)}(x_0)|,$$

$$\lim_{x \rightarrow x_0} |f^{(k)}(\xi) - f^{(k)}(x_0)| = 0$$

由此可见， $f(x)$ 在 x_0 点邻近 k -阶连续可微的实质内涵乃是 $f(x)$ 在 x_0 点邻近可以用上述多项函数 $g(x)$ 逼近到 k -阶。

再者，设 $\{x_i, 1 \leq i \leq k\}$ 是 x_0 的邻近 k 点。由多项式的插值公式可得一个唯一的次数至多为 k 的多项式，它在 $\{x_i, 0 \leq i \leq k\}$ 的值也就是 $f(x_i)$ 。试问上述多项曲线在 $\{x_i, 1 \leq i \leq k\}$ 都趋于 x_0 时，是否也和 $k = 1$ 的情形一样具有其极限多项曲线？若然，则它乃是切线的推广，称之为曲线 $y = f(x)$ 在 $(x_0, f(x_0))$ 点的 k -阶密切多项曲线。兹以 $k = 2$ 的情形为例，证明其唯一存在性，及得出其方程式：

【例 1】：由 (1.33) 可见过 $\{P_i(x_i, f(x_i)), 0 \leq i \leq 2\}$ 的二次多项曲线的方程式可以用行列式表达之，即

$$\begin{vmatrix} -y & 1 & x & x^2 \\ -f(x_0) & 1 & x_0 & x_0^2 \\ -f(x_1) & 1 & x_1 & x_1^2 \\ -f(x_2) & 1 & x_2 & x_2^2 \end{vmatrix} = 0$$

但是上式是不能直接求其在 $x_1 \rightarrow x_0$ 之极限式者，因为如此所得者乃是第二，第三行相同者也，方程式蜕化成无用的 $0 = 0$ 。所以我们要先用行列式的性质和微分均值定理，消除其蜕化的原由，然后再求其 $x_1 \rightarrow x_0$ 之极限式如下：将第三行减去第二行然后再用 $f(x_1) - f(x_0) = f'(\xi)(x_1 - x_0)$ ，即可把上式改写为

$$(x_1 - x_0) \begin{vmatrix} -y & 1 & x & x^2 \\ -f(x_0) & 1 & x_0 & x_0^2 \\ -f'(\xi) & 0 & 1 & x_1 + x_0 \\ -f(x_2) & 1 & x_2 & x_2^2 \end{vmatrix} = 0$$

这种计算显示了上式在 $x_1 \rightarrow x_0$ 时蜕化为 $0 = 0$ 的原由乃是它其实含有 $(x_1 - x_0)$ 因子。所以只要除去它之后再求极限就不会蜕化了。其所得为

$$\begin{vmatrix} -y & 1 & x & x^2 \\ -f(x_0) & 1 & x_0 & x_0^2 \\ -f'(x_0) & 0 & 1 & 2x_0 \\ -f(x_2) & 1 & x_2 & x_2^2 \end{vmatrix} = 0$$

它乃是和 $y = f(x)$ 相切于 $P_0(x_0, f(x_0))$ 点而且过 $P_2(x_2, f(x_2))$ 点的二次多项式曲线。现在再让 $x_2 \rightarrow x_0$ ，并求其极限式。当然，我们又得在求极限之前，先作类似的准备工作去排除其中隐含的 $(x_2 - x_0)$ 因子。要不然肯定会蜕化成 $0 = 0$ ，是不？将第四行减去第二行和 $(x_2 - x_0) \times$ 第三行，再用泰勒公式即得

$$(x_2 - x_0)^2 \begin{vmatrix} -y & 1 & x & x^2 \\ -f(x_0) & 1 & x_0 & x_0^2 \\ -f'(x_0) & 0 & 1 & 2x_0 \\ -\frac{f''(\eta)}{2} & 0 & 0 & 1 \end{vmatrix} = 0$$

所以只要除去上式中的 $(x_2 - x_0)^2$ 因子之后再求其 $x_2 \rightarrow x_0$ 之极限，即得所求之 2-阶密切多项式曲线的方程式乃是

$$\begin{vmatrix} -y & 1 & x & x^2 \\ -f(x_0) & 1 & x_0 & x_0^2 \\ -f'(x_0) & 0 & 1 & 2x_0 \\ -\frac{f''(x_0)}{2} & 0 & 0 & 1 \end{vmatrix} = 0$$

【习题 1】：试证上式之展开式为

$$-y + f(x_0) + f'(x_0)(x - x_0) + \frac{f''(x_0)}{2}(x - x_0)^2 = 0$$

【习题 2】：试用同样的计算求得 $y = f(x)$ 的 3-阶密切多项曲线的方程式。

【习题 3】：设 Γ 是一条由下述 2-阶连续可微的参数式

$$x = f(t), \quad y = g(t)$$

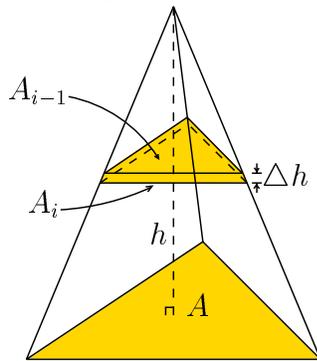
所描述的平面曲线， $P_i(f(t_i), g(t_i))$, $0 \leq i \leq 2$, 是其上相异三点。令 $C(P_0, P_1, P_2)$ 是由上述三点所定之圆。试用类似于[例 1]的计算证明在 t_1, t_2 都趋于 t_0 时， $C(P_0, P_1, P_2)$ 具有唯一的极限圆，而且求得它的方程式。（它叫做曲线 Γ 在 P 点的密切圆。）

【习题 4】：试用 $f'(t_0), g'(t_0)$ 和 $f''(t_0), g''(t_0)$ 表达上述密切圆的半径。

II. 几个历史性的积分范例：

【例 2】：锥体体积公式 (Eudoxus)： $V = \frac{1}{3}hA$

如 [图 2-14] 所示，一个底面积为 A 、高度为 h 的锥体可以用平行于底面的平面切割成 n 块均一厚度的薄片：



[图 2-14]

由顶层向下数的第 i 块薄片，其底面积为 A_i ，顶面积为 A_{i-1} 。由相似形定理，古希腊的几何学家已知

$$A_i : A = \left(\frac{i}{n}h : h \right)^2 = \left(\frac{i}{n} \right)^2, \quad \text{即 } A_i = \frac{i^2}{n^2}A$$

由此易知其第 i 块薄片之体积是介乎于 $\frac{h}{n} \cdot A_{i-1}$ 和 $\frac{h}{n} \cdot A_i$ 两者之间，亦即

$$\frac{h}{n} A_{i-1} = \frac{(i-1)^2}{n^3} hA < V_i < \frac{i^2}{n^3} hA = \frac{h}{n} A_i, \quad 1 \leq i \leq n$$

把上述的不等式整合起来，便得出

$$\frac{hA}{n^3} \sum_{i=1}^n (i-1)^2 < \sum_{i=1}^n V_i = V < \frac{hA}{n^3} \sum_{i=1}^n i^2$$

接著 Eudoxus 便应用已知的 $\sum i^2$ 求和公式把上式重写成

$$\frac{hA}{3} \left(1 - \frac{1}{n}\right) \left(1 - \frac{1}{2n}\right) < V < \frac{hA}{3} \left(1 + \frac{1}{n}\right) \left(1 + \frac{1}{2n}\right)$$

至此，Eudoxus 发现上述不等式对所有正整数 n 都成立，但无论 n 如何增大， V 的值（锥体体积）是不变的！而当 n 无限地增大时，上式的上限和下限之差别为

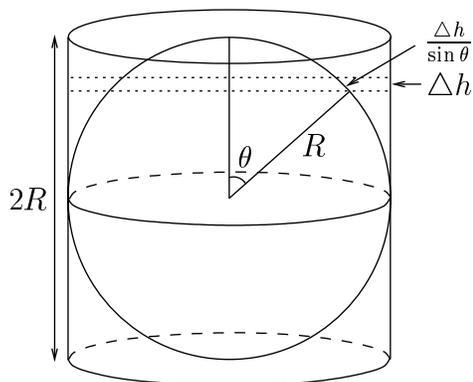
$$\frac{hA}{3} \left(1 + \frac{1}{n}\right) \left(1 + \frac{1}{2n}\right) - \frac{hA}{3} \left(1 - \frac{1}{n}\right) \left(1 - \frac{1}{2n}\right) = \frac{hA}{n}$$

这个量会无限地缩小（即可小于任意给出的正实数）。由此可见，唯一可以对所有 n 能满足上述不等式的量必定是 $\frac{hA}{3}$ ，因此得证

$$V = \frac{1}{3} hA$$

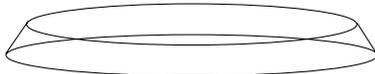
【例 3】：球面面积公式 (Archimedes)： $A = 4\pi R^2$

Archimedes (287–212 B.C.) 是公认的古希腊时代伟大的科学家和几何学家，他一生有很多卓越的贡献；而他最引以自豪者，首推上述公式及其简洁的证明，这也就是遵照他本人的遗嘱刻在他的墓碑上者。证明的要点在于论证一个半径为 R 的球面面积和一个高为 $2R$ ，半径为 R 的圆柱面面积相等。在他的墓碑上所刻划的，就是如 [图 2-15] 所示把两者放在相切同高的位置。



[图 2-15]

设想用一系列和柱面正交的平行面，把两个面都细分成很窄很窄的一圈圈。设相邻两个平行平面之间的距离是 Δh ，则柱面上的窄条（或圈）的面积等于 $\Delta h 2\pi R$ ，而在球面上的相应窄圈其宽度和长度随著 θ 而改变，在 Δh 非常、非常小的时候，它可以看成如 [图 2-16] 所示的圆台之侧面：



[图 2-16]

其中环长度是 $2\pi R \sin \theta$ ，亦即其环长的平均值是 $2\pi R \sin \theta$ ，而侧面的高度则为 $\frac{\Delta h}{\sin \theta}$ ，所以其面积的高度近似值也是 $\Delta h 2\pi R$ （亦即可能的误差肯定在 Δh^2 这种量级）。由此他就用 Eudoxus 所创的逼近原理证明了两者的面积必然相等，而後者的面积显然等于高为 $2R$ ，长为 $2\pi R$ 的长方形面积，亦即是 $4\pi R^2$ 。

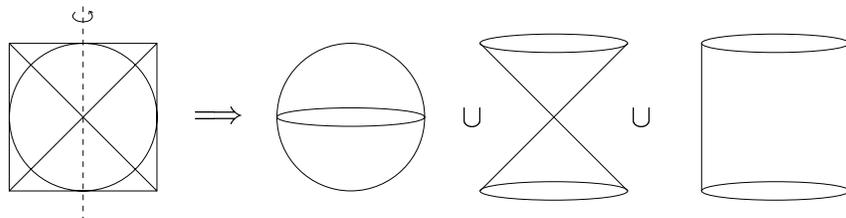
【例 4】：球体体积公式 (祖暅原理之应用)： $V = \frac{4}{3}\pi R^3$

在古代中国，几何学研究乃是基于致用之所需，用来解决各种各样的建筑及测量的实际问题。他们成功地利用矩形面积公式为基础，以割补重组方法证明了平面几何上的基本定理如勾股定理、相似三角形定理等。及至东汉刘徽，开始进而研究体积问题（如求解球体体积公式等等）就发现不能单靠割补重组去求公式了。这种困难也自然而然激发了刘徽、祖冲之、祖暅他们世代相承地对于体积的理论的探讨。一直到了祖暅总结而得下述原理才成功地解答了球体体积公式。

【祖暅原理】：幂势既同，则积不容殊。

亦即是说：如有两个物体，它们沿某一方向的所有截面面积（幂势）皆相等，则两物体的体积必定相等。

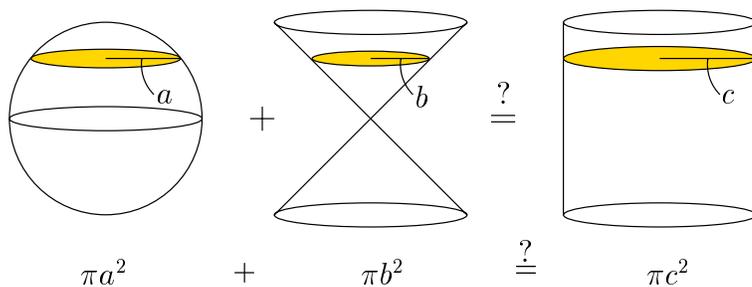
让我们以球体体积公式为例子，说明祖暅原理的应用：首先，将下面的图形以虚线为轴作旋转，便可得出一个球体、两个对顶的圆锥体和一个圆柱体：



[图 2-17]

运用祖暅原理，我们若可证明球体与对顶圆锥体的横截面面积之和等于圆柱体的横截面面积，则有

球体体积 + 2 × 圆锥体体积 = 圆柱体体积

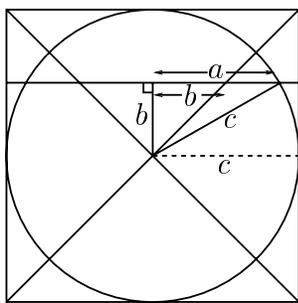


[图 2-18]

亦即

$$\begin{aligned} \text{球体体积} &= \text{圆柱体体积} - 2 \times \text{圆锥体体积} \\ &= \pi R^2 \cdot (2R) - 2 \times \frac{1}{3} \pi R^2 \cdot R \\ &= 2\pi R^3 - \frac{2}{3} \pi R^3 = \frac{4}{3} \pi R^3 \end{aligned}$$

其实，由 [图 2-19] 就可立即看出为何 $a^2 + b^2 = c^2$ （勾股定理！）



[图 2-19]

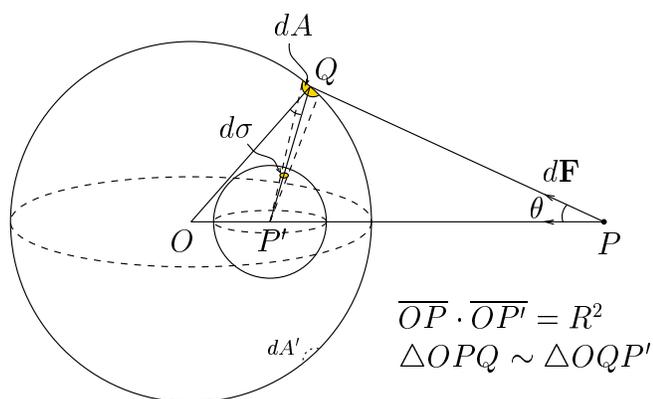
【例 5】：地心引力 (Newton's Formula)

当年牛顿应用数理分析技巧来研究 Kepler 行星运行三定律时，其所得的结果为：行星和太阳之间的引力与两者间距离的平方满足反比的关系（亦即平方反比定律 (inverse square law)）。然后牛顿继而想再进一步把行星与太阳之间的引力推广到任何物体与物体之间的引力（亦即所谓的万有引力），他遇到一个困难，使这位科学史上的巨人困扰了数年。由于行星与太阳之间的距离很大，所以在计算中可把行星和太阳当作两个质点，即可以假设质点集中了整个球体的质量；但当推广至任何物体与物体之间的情形，如苹果与地球，则便不可以随便把地球当作为一个质点了。牛顿所遇到的困难，就是他不能证明的确也可以把地球当作为一个质点。即使在 1684 年他的好友 Halley（哈雷）力邀牛顿发表已得的结果，他仍不愿意发表。到了 1686 年，他终于成功地证明了上述结果，即一个密度只随著到球心距离而变化的球体，在吸引球外一个质点时，所作用的力就像假设全部质量都集中在球心一样。在这年他写信给 Halley 表示同意写出他的工作¹，这就是在次年 (1687) 出版的科学巨著 *Philosophiae Naturalis Principia Mathematica*。

牛顿在书中所给的证明是很繁复的。在这里，我们给出另一个证明，它巧妙地运用了球的几何特性而大大简化了计算过程。

对于一个球面的最自然、最对称的点当然就是球心。但是在研讨球面与球外一点 P 的互相作用时，从几何观点来看，最自然、最对称的点就不再是球心 O ，而是 [图 2-20] 的 P' 点（这是 P 相对于球面的反射对称点）。

¹在信中牛顿还提到直至 1685 年，他还在怀疑这个结果是错的。

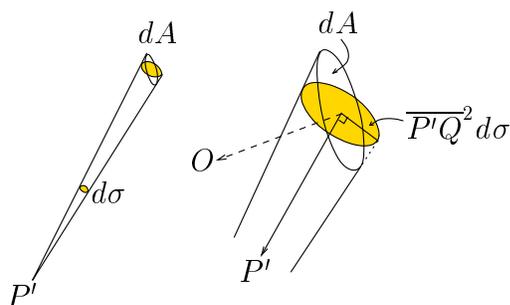


[图 2-20]

设球体其中一层薄壳的半径为 R ，面密度为 ρ ，球外质点 P 的质量为 m 。考虑在薄壳上的一小片面积 dA 作用于 P 的引力 $d\mathbf{F}$ 。因薄壳对于 OP 是旋转对称， $d\mathbf{F}$ 垂直于 OP 的分量会被对称小片 dA' 所抵消，所以只需考虑 OP 方向的分量：

$$|d\mathbf{F}| \cos \theta = G \frac{\rho dA \cdot m}{PQ^2} \cos \theta$$

构造 P' 使得 $\overline{OP} \cdot \overline{OP'} = R^2$ ，并以 P' 为心构造一个单位球面。令 P' 连向 dA 的射线在该单位球面上的影象为 $d\sigma$ 。



[图 2-21]

如 [图 2-21] 所示， dA 和 $d\sigma$ 之间有一个简单的关系：

$$dA \cdot \cos \theta = \overline{P'Q}^2 \cdot d\sigma$$

因此，整个薄壳作用于 P 的力就是

$$\begin{aligned} |\mathbf{F}| &= \sum |d\mathbf{F}| \cdot \cos \theta = \sum Gm\rho \frac{dA \cdot \cos \theta}{PQ^2} = \sum Gm\rho \frac{\overline{P'Q}^2 \cdot d\sigma}{PQ^2} \\ &= \sum Gm\rho \frac{R^2}{OP^2} d\sigma = Gm\rho \frac{R^2}{OP^2} \sum d\sigma \\ &= Gm\rho \frac{R^2}{OP^2} \cdot 4\pi = G \frac{Mm}{OP^2} \end{aligned}$$

因此这层薄壳作用于 P 的力等于将其全部质量集中于 O 而作用于 P 的力。再将所有薄壳作用的力加起来，便得所需之公式。

第三章

指数及对数函数

3.1 指数、对数函数的定义与基本性质

设 $a > 0, a \neq 1$ 是一个给定实数。它的 n 乘方 a^n 乃是 n 个 a 相乘之积，所以显然有指数定则

$$a^m \cdot a^n = a^{m+n}, \quad (a^m)^n = a^{nm}, \quad m, n \in \mathbb{N}$$

我们可以逐步把上述指数的定义推广到整数、有理数和实数指数，使得 a^x 对于所有实数 x 皆有其自然的定义，称之为以 a 为基底的指数函数 (exponential function)。指数定义的逐步自然推广如下：

(i) 0 和负整数指数的定义：

$$a^0 := 1, \quad a^{-n} := \frac{1}{a^n}$$

(ii) 有理指数的定义：

设 $a > 1$ ，则易见 $f(x) = x^n - a$ 是一个递增连续函数。由所设即有 $f(1) < 0, f(a+1) > 0$ ，所以存在唯一的 $1 < \xi < (a+1)$ 使得 $f(\xi) = 0$ ，称之为 a 的 n 方根，以 $\sqrt[n]{a}$ 记之。而 a 的有理数指数定义如下：

$$a^{\frac{m}{n}} := (\sqrt[n]{a})^m = \sqrt[n]{a^m}$$
$$a^{-\frac{m}{n}} := 1/a^{\frac{m}{n}} = \left(\frac{1}{\sqrt[n]{a}}\right)^m$$

[当 $0 < a < 1$ 时, 则有 $\frac{1}{a} > 1$, 所以可以用 $a^{\pm \frac{m}{n}} = (\frac{1}{a})^{\mp \frac{m}{n}}$ 把它归于前者定义之。]

不难验证上述有理数指数的定义乃是能够保有下列指数定则和单调性的唯一者, 即当 $r, s \in \mathbb{Q}$

$$a^r \cdot a^s = a^{r+s}, \quad (a^r)^s = a^{sr}$$

$$r > s \Rightarrow \begin{cases} a^r > a^s & (a > 1) \\ a^r < a^s & (0 < a < 1) \end{cases}$$

(iii) 现在让我们定义实数指数。首先要证明下述引理:

【引理 3.1】: 设 $a > 0$, 则 $\lim_{n \rightarrow \infty} a^{\frac{1}{n}} = 1$ 。

证明: 不妨设 $a > 1$ 。则 $\{a^{\frac{1}{n}} = \sqrt[n]{a}\}$ 乃是一个递减数列, 所以其极限值就是它的极大下界。1 显然是它的一个下界, 所以只要证明任何大于 1 者不可能是它的下界。亦即任给 $\varepsilon_0 > 0$, 当 n 足够大时, $a^{\frac{1}{n}} < 1 + \varepsilon_0$ 。其实, 当 $n > \frac{a-1}{\varepsilon_0}$ 时, 由二项式即有

$$(1 + \varepsilon_0)^n > 1 + n\varepsilon_0 > 1 + (a - 1) = a \Rightarrow 1 + \varepsilon_0 > a^{\frac{1}{n}}$$

当 $0 < a < 1$ 时, 其证明基本上是一模一样的。 □

【推论】: 设 $a > 1$, 而且 $\{r_n\}$ 和 $\{s_n\}$ 是一对左、右夹逼正有理数数列, 即

$$0 < r_1 \leq r_2 \leq \cdots \leq r_n \leq r_{n+1} \leq \cdots \leq s_{n+1} \leq s_n \leq \cdots \leq s_2 \leq s_1$$

而且 $(s_n - r_n) \rightarrow 0$ (亦即可以小到任意小), 则 $\{a^{r_n}\}$ 和 $\{a^{s_n}\}$ 也是一对左、右夹逼数列。

证明: 由 $a > 1$ 和 $\frac{m}{n} > \frac{p}{q} \Rightarrow a^{\frac{m}{n}} > a^{\frac{p}{q}}$ 可见 $\{a^{r_n}\}$ 是递增的而 $\{a^{s_n}\}$ 则是递减的。所以唯一还需要验证者乃是

$$(a^{s_n} - a^{r_n}) \text{ 可以小到任意小}$$

兹证之如下:

设 $\varepsilon > 0$ 是任给正实数, 另取一个正实数 $\varepsilon' > 0$ 使得 $a^{s_1} \cdot \varepsilon' < \varepsilon$ 。由上述引理即有一个足够大的 N , 使得

$$a^{\frac{1}{n}} < 1 + \varepsilon', \quad n \geq N$$

再者，由所设 $(s_n - r_n) \rightarrow 0$ ，即有足够大的 n 使得 $(s_n - r_n) < \frac{1}{N}$ 。由此即得

$$\begin{aligned} a^{s_n} - a^{r_n} &= a^{r_n} (a^{(s_n - r_n)} - 1) \\ &< a^{s_1} \cdot (a^{\frac{1}{N}} - 1) \\ &< a^{s_1} \cdot \varepsilon' < \varepsilon \end{aligned} \quad \square$$

【定义】：设 $a > 1$ ， λ 是一个给定正实数， $\{r_n\}$ 和 $\{s_n\}$ 是 λ 的一对左、右夹逼正分数数列。则定义 a^λ 为左、右夹逼数列 $\{a^{r_n}\}$ 和 $\{a^{s_n}\}$ 所夹逼的那个正实数。

定义之合理性：我们必须验证上述定义和 λ 的左、右夹逼正分数列的选取无关，唯有这样，上述定义才是合理的。为此，设 $\{r'_n\}$ 和 $\{s'_n\}$ 是另一 λ 的左、右夹逼正分数数列，则有

$$\{r_n\} \text{ 和 } \{s'_n\} \quad \text{以及} \quad \{r'_n\} \text{ 和 } \{s_n\}$$

也都是 λ 的左、右夹逼数列。所以

$$\{a^{r_n}\} \text{ 和 } \{a^{s'_n}\} \quad \text{以及} \quad \{a^{r'_n}\} \text{ 和 } \{a^{s_n}\}$$

也都是 a^λ 的左、右夹逼数列。因此

$$\begin{aligned} \lim_{n \rightarrow \infty} a^{r_n} &= \lim_{n \rightarrow \infty} a^{s'_n} \\ \lim_{n \rightarrow \infty} a^{r'_n} &= \lim_{n \rightarrow \infty} a^{s_n} \end{aligned}$$

由此可见上述四个极限都相等，所以定义它们的共同极限为 a^λ 是完全合理而且自然的！

【定义】： $a^{-\lambda}$ 定义为 $\frac{1}{a^\lambda}$ 。

[当 $0 < a < 1$ 时，同样可以用 $a^\lambda = (\frac{1}{a})^{-\lambda}$ 把它归于前者定义之。]

总结上面由一个给定 a 的整数指数到它的实数指数的逐步推广，所用到者基本上就是实数系的连续性，和正有理数指数的单调性和夹逼性，亦即

$$\begin{aligned} \frac{m}{n} > \frac{p}{q} &\Rightarrow a^{\frac{m}{n}} > a^{\frac{p}{q}} \\ (s_n - r_n) \rightarrow 0 &\Rightarrow (a^{s_n} - a^{r_n}) \rightarrow 0 \end{aligned}$$

再者，把 $x \mapsto a^x$ 想成一个实变数 x 的函数，亦即定义 $f(x) = a^x$ ，称之为以 a 为基底的指数函数 (the exponential function with a as the base)，它具有下述特征性质：

【定理 3.1】：设 $a > 1$ (或 $0 < a < 1$)，则指数函数 $f(x) = a^x$, $x \in \mathbb{R}$ 是一个单调递增 (或递减) 函数，而且满足

$$f(x_1 + x_2) = f(x_1) \cdot f(x_2)$$

反之，任何一个满足上式的单调递增 (或递减) 函数 $f(x)$ 总是一个指数函数，亦即

$$f(x) = a^x, \quad a = f(1)$$

证明：其实，[定理 3.1]就是上述指数逐步推广所得的总结。至于其唯一性的验证的主要步骤如下：由假设 $f(x_1 + x_2) = f(x_1) \cdot f(x_2)$ 即得

$$f(x) = f(0 + x) = f(0) \cdot f(x) \Rightarrow f(0) = 1$$

再由其单调 (递增) 性即有 $f(1) > f(0) = 1$ 。令 $a = f(1)$ ，则有 $f(\frac{1}{n}) > f(0) = 1$ ，而且

$$\begin{aligned} [f(\frac{1}{n})]^n &= f(\frac{1}{n} + \cdots + \frac{1}{n}) = f(1) = a \\ \Rightarrow f(\frac{1}{n}) &= \sqrt[n]{a} = a^{\frac{1}{n}} \\ \Rightarrow f(\frac{m}{n}) &= [f(\frac{1}{n})]^m = (\sqrt[n]{a})^m = a^{\frac{m}{n}} \\ f(-\frac{m}{n}) \cdot f(\frac{m}{n}) &= f(0) = 1 \\ \Rightarrow f(-\frac{m}{n}) &= [f(\frac{m}{n})]^{-1} = [a^{\frac{m}{n}}]^{-1} = a^{-\frac{m}{n}} \end{aligned}$$

再由单调性即得

$$f(x) = a^x$$

对于所有实数 x 皆成立。 □

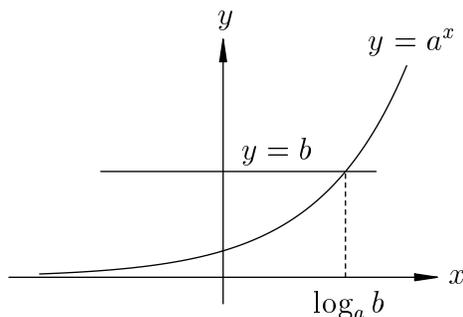
[注意]：在上述由分数指数到实数指数的推广之中，我们所使用者乃是所要定义的 a^x 的单调性。由此可见单调性乃是实变指数函数 $y = a^x$ 定义之所基，而其连续性则是如此定义之所得！

设 $a > 1$ (或 $0 < a < 1$)，则 $f(x) = a^x$ 是一个定义于全实数系的单调递增 (或递减) 连续函数，而且

$$\lim_{x \rightarrow -\infty} a^x = 0, \quad \lim_{x \rightarrow +\infty} a^x = +\infty$$

$$(\text{或 } \lim_{x \rightarrow -\infty} a^x = +\infty, \quad \lim_{x \rightarrow +\infty} a^x = 0)$$

所以对于任给正实数 b , $f(x) = b$ 具有唯一之解，将以符号 $\log_a b$ 记之，叫做 b 以 a 为底的对数 (如 [图 3-1] 所示)。



[图 3-1]

【定义】： $\log_a y = x \Leftrightarrow y = a^x$

不难验证对数函数具有下述基本性质，即

(i) $\log_a y_1 \cdot y_2 = \log_a y_1 + \log_a y_2$

(ii) 单调性

(iii) 连续性

[对数函数乃是指数函数的反函数，所以其基本性质乃是指数函数相应的基本性质的反映。由此即可逐个验证上述三点。]

设 a, b 都是不等于 1 的两个正实数，则有

$$b^x = (a^{\log_a b})^x = a^{(\log_a b)x}$$

$$a^x = (b^{\log_b a})^x = b^{(\log_b a)x}$$

由此可见，不同基底的指数函数所差别者仅仅是 $x \mapsto kx$ 这种常数倍的变换而已。

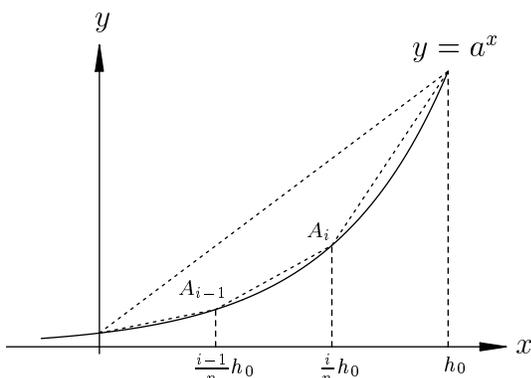
3.2 指数函数与对数函数的微分

设 $a > 1$ 为取定基底, $f(x) = a^x$ 。

【引理 3.2】: $f(x) = a^x$ 在 $x = 0$ 点是可微的, 亦即

$$\lim_{h \rightarrow 0} \frac{a^h - 1}{h}$$

是存在的, 且以 K_a 记之。



[图 3-2]

证明: 若把某一取定的 $h_0 > 0$ 等分成 n 段, 则其第 i 段上的平均变率 (即 [图 3-2] 中 $\overline{A_{i-1}A_i}$ 的斜率) 等于其第一段的斜率的 $a^{\frac{i-1}{n}h_0}$ -倍, 亦即

$$\begin{aligned} m_i(h_0) &= \frac{n}{h_0} (a^{\frac{i}{n}h_0} - a^{\frac{i-1}{n}h_0}) \\ &= a^{\frac{i-1}{n}h_0} \frac{n}{h_0} (a^{\frac{h_0}{n}} - 1) = a^{\frac{i-1}{n}h_0} m_1(h_0) \end{aligned}$$

由此可见其逐段的斜率是逐次以 $a^{\frac{h_0}{n}}$ -倍增加的。用它就不难推导当 $h' < h$ 和 h', h 可公度时

$$m(h') = \frac{a^{h'} - 1}{h'} < \frac{a^h - 1}{h} = m(h)$$

然後便可以用 Eudoxus 逼近方法来证明 $m(h)$, $h > 0$ 乃是一个 h 的单调递增函数, 所以 $\lim_{h \searrow 0} m(h)$ 是存在的, 它就是 $\{m(h), h > 0\}$ 的极大下

界。再者

$$\lim_{h \searrow 0} \frac{a^{(-h)} - 1}{(-h)} = \lim_{h \searrow 0} a^{(-h)} \cdot m(h) = 1 \cdot \lim_{h \searrow 0} m(h)$$

这也就证明了 $f(x) = a^x$ 在 $x = 0$ 是可微的。 \square

[注]: 当 $0 < a < 1$ 时, 亦会有同样结果, 试证之。

【定理 3.2】: $f(x) = a^x$ 是到处可微的, 而且

$$f'(x) = K_a f(x)$$

证明: 对于任给 x_0 点, 皆有

$$\lim_{h \rightarrow 0} \frac{f(x_0 + h) - f(x_0)}{h} = \lim_{h \rightarrow 0} a^{x_0} \cdot \frac{a^h - 1}{h} = a^{x_0} \cdot K_a \quad \square$$

【推论】: $K_b = (\log_a b) K_a$

证明: $b^x = a^{(\log_a b) \cdot x}$, 所以

$$K_b \cdot b^x = (b^x)' = [a^{(\log_a b)x}]' = \log_a b K_a \cdot a^{(\log_a b)x} = (K_a \log_a b) b^x$$

亦即 $K_b = (\log_a b) K_a$ 。 \square

由上述推论, 可见存在一个特殊的基底 e 使得 $K_e = 1$, 即 $e = a^{\frac{1}{K_a}}$ 。这个特殊的基底 e 叫做自然对数的基底 (base of natural logarithm), 乃是 Euler 所发现者, 并把它命名为 e , 我想意在要后人不要忘了这是他所发现者也。

【定理 3.3】: $f(x) = ce^x$, $c \neq 0$, 满足微分方程

$$f'(x) = f(x)$$

反之, 任何满足上述微分方程的非零函数都是 e^x 的常数倍。

证明: 由 $K_e = 1$, 可知

$$[ce^x]' = c[e^x]' = ce^x$$

反之, 设 $f'(x) = f(x)$, 则有

$$[e^{-x} \cdot f(x)]' = -e^{-x} \cdot f(x) + e^{-x} \cdot f(x) \equiv 0$$

所以 $e^{-x} \cdot f(x) = c$ (常数), 亦即 $f(x) = ce^x$ 。 □

同理, 微分方程

$$f'(x) = kf(x)$$

的解都可以表成 $c \cdot e^{kx}$, 其中 $c = f(0)$ 。

【定理 3.4】: e^x 可以用以下幂级数表达, 即

$$e^x = 1 + x + \frac{1}{2!}x^2 + \dots + \frac{1}{n!}x^n + \dots = \sum_{n=0}^{\infty} \frac{1}{n!}x^n$$

证明: 由 $[e^x]' = e^x$ 易见其 n -阶导数函数还是 e^x 。所以泰勒公式中的余项为

$$R_n = \frac{x_0^{n+1}}{(n+1)!} e^{\xi} \quad (x_0 \text{ 取定, } 0 < \xi < x_0)$$

显然有 $\lim_{n \rightarrow \infty} R_n = 0$ 。 □

【推论】: $e = 1 + 1 + \frac{1}{2!} + \dots + \frac{1}{n!} + \dots = \sum_{n=0}^{\infty} \frac{1}{n!}$

$$e = \lim_{n \rightarrow \infty} \left(1 + \frac{1}{n}\right)^n$$

证明: 不难验证 $s_n = \left(1 + \frac{1}{n}\right)^n$ 是一个递增数列, 而且

$$\lim_{n \rightarrow \infty} \left(1 + \frac{1}{n}\right)^n = \sum_{m=0}^{\infty} \frac{1}{m!} \quad \square$$

符号: 通常用 $\ln y$ 表示 $\log_e y$ 。

【定理 3.5】: $[\ln y]' = \frac{1}{y}$

证明：令 $x = \ln y$ ，则有 $y = e^x$ 。再者

$$\frac{dx}{dy} = \frac{1}{\frac{dy}{dx}} = \frac{1}{e^x} = \frac{1}{y}$$

所以 $[\ln y]' = \frac{1}{y}$ 。 □

【定理 3.6】：当 $|y| < 1$ 时，

$$\ln(1+y) = y - \frac{y^2}{2} + \frac{y^3}{3} \mp \dots = \sum_{n=1}^{\infty} (-1)^{n+1} \frac{y^n}{n}$$

证明：由熟知 $\frac{1}{1+t}$ 的几何级数展开式：

$$\frac{1}{1+t} = 1 - t + t^2 - t^3 + \dots + (-1)^{n-1} t^{n-1} + \frac{(-1)^n t^n}{1+t}$$

运用微积分基本定理，左方的积分就是 $\ln(1+y)$ ；而当 $|y| < 1$ 时，右方余项的积分就会有下述估计：

$$\left| \int_0^y \frac{(-1)^n t^n}{1+t} dt \right| \leq \left| \int_0^y \frac{t^n}{1-|y|} dt \right| = \frac{|y|^{n+1}}{(n+1)(1-|y|)}$$

所以这个余项的绝对值会随著 n 的增大而趋于 0，因此得出：

$$\ln(1+y) = y - \frac{y^2}{2} + \frac{y^3}{3} \mp \dots = \sum_{n=1}^{\infty} (-1)^{n+1} \frac{y^n}{n} \quad \square$$

3.3 自然对数表的计算法

因为 $\ln(xy) = \ln x + \ln y$ ，所以我们只需求得所有质数 p 的对数值便可以由此算得其他整数的对数值。由于下述 $\ln(1+x)$ 的幂级数展开式

$$\ln(1+x) = x - \frac{x^2}{2} + \frac{x^3}{3} - \frac{x^4}{4} + \frac{x^5}{5} \pm \dots$$

是一个正负交错的级数，所以我们不妨把上式中的 x 换以 $-x$ 然后将两式相减，即得下述简化了的公式：

$$\ln\left(\frac{1+x}{1-x}\right) = 2\left(x + \frac{x^3}{3} + \frac{x^5}{5} + \dots\right)$$

令 $x = \frac{1}{2p^2 - 1}$, 则有

$$\begin{aligned}\ln\left(\frac{1+x}{1-x}\right) &= \ln\frac{2p^2 - 1 + 1}{2p^2 - 1 - 1} = \ln\frac{p^2}{(p+1)(p-1)} \\ &= 2\ln p - \ln(p+1)(p-1)\end{aligned}$$

注意当质数 $p > 2$ 时, 在 $(p-1)$ 和 $(p+1)$ 中出现的质因数皆小于 p , 所以如果我们已经求得小于 p 的质数的对数高度近似值, 就可以用上式求得 $\ln p$ 的高度近似值:

$$2\ln p = \ln\left(\frac{1+x}{1-x}\right) + \ln(p+1) + \ln(p-1)$$

而未知的 $\ln\left(\frac{1+x}{1-x}\right)$ 是能够有效计算的, 因为 $x = \frac{1}{2p^2-1}$ 的绝对值很小。

例如: 若我们想构造一个 8 位对数表, 则可以依次序地求 $2, 3, 5, \dots$ 的对数值。因为后面的对数值会用到前面所得的对数值来计算, 所以在开始时 $\ln 2$ 要算得准确一些:

$$\begin{aligned}\ln 2 &= \ln\left(\frac{1+\frac{1}{3}}{1-\frac{1}{3}}\right) \\ &\approx 2\left(\frac{1}{3} + \frac{\left(\frac{1}{3}\right)^3}{3} + \frac{\left(\frac{1}{3}\right)^5}{5} + \dots + \frac{\left(\frac{1}{3}\right)^{21}}{21}\right) \\ &= 0.6931471805589\dots\end{aligned}$$

这个和实际数值 $\ln 2 = 0.693147180559945\dots$ 相比其精确度已到达第 11 个位。接著便是要计算 $\ln 3$ 。令 $x = \frac{1}{2 \cdot 3^2 - 1} = \frac{1}{17}$, 则有

$$\ln\left(\frac{1+\frac{1}{17}}{1-\frac{1}{17}}\right) \approx 2\left(\frac{1}{17} + \frac{\left(\frac{1}{17}\right)^3}{3} + \frac{\left(\frac{1}{17}\right)^5}{5} + \frac{\left(\frac{1}{17}\right)^7}{7}\right) = 0.117783035654504\dots$$

所以

$$\ln 3 \approx \frac{1}{2}(0.11778303565\dots + \ln 4 + \ln 2) = 1.098612288635\dots$$

这个和实际数值 $\ln 3 = 1.09861228866811\dots$ 相比其精确度也到达第 10 个位。

【习题】：

- (1) 试证 a^x 的连续性。
- (2) 试计算 $\ln 5, \ln 7, \ln 11$ ，其精确度要有 8 个位。

3.4 复变数指数函数和三角函数

最後我们将指数函数的定义域扩张到复数的领域。由实数系扩张到复数系，所作者在本质上是用一对实数 a, b 和虚数单元 i 组成一个复数

$$z = a + bi$$

而指数函数的最重要性质乃是 (i) 其指数定则的普遍成立 (见[定理 3.1])，(ii) 为微分方程 $f'(x) = kf(x)$ 的通解 (见[定理 3.3])。所以我们可以从这两个性质去探讨复指数的适当定义。

【分析】：(i) 设 e 为自然对数的基底，亦即 Euler 所发现的

$$e = \lim_{n \rightarrow \infty} \left(1 + \frac{1}{n}\right)^n$$

则有 $[e^x]' = e^x$, $[e^{kx}]' = ke^{kx}$, $k \in \mathbb{R}$ 。

(ii) 对于复变数 $z = x + iy$ 我们应该定义「 e^z 」= $e^x \cdot e^{iy}$ ，所以要点在于如何合理地定义「 e^{iy} 」。我们所要探索的函数「 e^{iy} 」应该是一个实变数复值函数，亦即

$$e^{iy} = f(y) + ig(y)$$

而它的微分应该就是

$$[e^{kx}]' = ke^{kx}$$

推广到 $k = i$ 的情形，亦即

$$f'(y) + ig'(y) = i(f(y) + ig(y)) = -g(y) + if(y)$$

所以我们所探求的乃是一对实变、实值函数 $\{f(y), g(y)\}$ ，它们具有关系

$$f'(y) = -g(y), \quad g'(y) = f(y)$$

而且具有初值条件： $f(0) = 1, g(0) = 0$ ，亦即

$$f(0) + ig(0) = e^0 = 1$$

【引理 3.3】：设 $\{f(y), g(y)\}$ 为一对连续可微函数并满足

$$f'(y) = -g(y), g'(y) = f(y) \quad \text{和} \quad f(0) = 1, g(0) = 0$$

则 $f(y) = \cos y, g(y) = \sin y$ 。

证明：由直接微分可得

$$\frac{d}{dy}(f(y) + ig(y))(\cos y - i \sin y) \equiv 0$$

所以 $f(y) + ig(y) = z_0(\cos y + i \sin y)$, $z_0 \in \mathbb{C}$ 为复常数。再用初值条件 $f(0) = 1, g(0) = 0$ 即可证得 $z_0 = 1$ 。□

所以 e^{x+iy} 应该定义为

$$e^{x+iy} = e^x(\cos y + i \sin y)$$

再者，由

$$\begin{cases} e^{iy} = \cos y + i \sin y \\ e^{-iy} = \cos y - i \sin y \end{cases}$$

即可解得

$$\begin{aligned} \cos y &= \frac{1}{2}\{e^{iy} + e^{-iy}\} \\ \sin y &= \frac{1}{2i}\{e^{iy} - e^{-iy}\} \end{aligned}$$

这也就是著名的 Euler 公式。

上述 Euler 公式亦提供了一种途径来推广三角函数的定义域，亦即当 $z \in \mathbb{C}$ ，我们定义

$$\begin{aligned} \cos z &= \frac{1}{2}\{e^{iz} + e^{-iz}\} \\ \sin z &= \frac{1}{2i}\{e^{iz} - e^{-iz}\} \end{aligned}$$

采取上述定义，我们可以逐一验证常见常用的三角公式，例如

$$\begin{aligned}\cos^2 z + \sin^2 z &= 1 \\ \cos(z - w) &= \cos z \cos w + \sin z \sin w\end{aligned}$$

等等。

3.5 复利与指数函数

指数函数还有一个自然的出处，那就是无限细分的复利计算。设本金为 C ，某种约定的单位时期的利率是 x 。若以单利计算则在单位时期後之所值就是 $C(1+x)$ 。若把该单位时期等分为 n 个分期，每次以利率 $\frac{x}{n}$ 为分期之利率来计算其复利，则其所值就是

$$P_n(x) = C\left(1 + \frac{x}{n}\right)^n$$

而且显然有 $P_{2n}(x) > P_n(x)$ 。[其实也不难验证 $P_{n+1}(x) > P_n(x)$ ，留作习题。]由此可见，当 n 无限增大，其极限值乃是一种无限细分的连续复利。不难结合前面的讨论证明

$$\lim_{n \rightarrow \infty} P_n(x) = C \cdot e^x$$

【习题】：

(1) 设 $k < n$ 是两个给定正整数，试证

$$\lim_{x \rightarrow \infty} \frac{x^k}{P_n(x)} = 0$$

(2) 对于任给正整数 k ，试证

$$\lim_{x \rightarrow \infty} \frac{x^k}{e^x} = 0$$

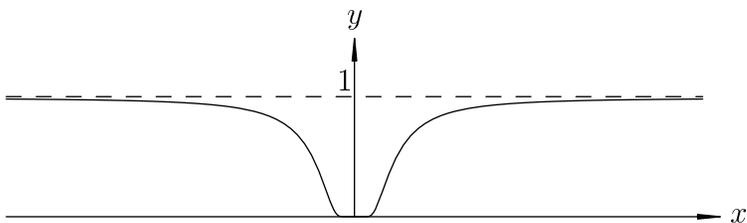
(3) 对于任给正整数 k , 试证

$$\lim_{x \rightarrow \infty} \frac{\ln x}{\sqrt[k]{x}} = 0$$

(4) 对于任给的多项式 $f(x)$, 试证

$$\lim_{x \rightarrow \infty} \frac{f(x)}{e^x} = 0$$

(5) 令 $f(x) = e^{-\frac{1}{x^2}}$ ($x \neq 0$), 而且 $f(0) = 0$. 试证 $f(x)$ 是到处无限连续可微的函数而且对于所有正整数 n 皆有 $f^{(n)}(0) = 0$, 其图象如 [图 3-3] 所示。



[图 3-3]

(6) 令

$$g(x) = \begin{cases} e^{-\frac{1}{x^2}}, & x > 0 \\ 0, & x \leq 0 \end{cases}$$

试证 $g(x)$ 也是到处无限连续可微的函数。

(7) 令 $h(x) = g(x) \cdot g(1-x)$, 则 $h(x)$ 是一个在区间 $[0, 1]$ 之外皆恒等于 0 的到处无限连续可微的函数, 试证之。

(8) 设 $a_1 < a_2 < b_2 < b_1$. 设构造一个在 $[a_1, b_1]$ 之外恒等于 0 而在 $[a_2, b_2]$ 之内恒等于 1 的到处无限连续可微的函数。[提示: 试利用上述 $h(x)$ 的积分函数构造之。]

下面让我们来讨论几个重要而且彼此相关的积分: 对于任给正实数 $n > 0$, Euler 引入下述重要的 Γ -函数, 即

$$\Gamma(n) = \int_0^{\infty} e^{-x} x^{n-1} dx = \lim_{y \rightarrow \infty} \int_0^y e^{-x} x^{n-1} dx$$

(9) 试用部分积分公式证明当 $n > 1$ 时

$$\Gamma(n) = (n-1)\Gamma(n-1)$$

(10) 当 n 是正整数时, 试证

$$\Gamma(n) = (n-1)!$$

(11) 试证

$$\int_{-\infty}^{\infty} e^{-x^2} dx = \Gamma\left(\frac{1}{2}\right)$$

[上述积分是概率论 (probability) 中具有基本重要性的数值, 其答案是 $\sqrt{\pi}$, 而 $\frac{1}{\sqrt{\pi}}e^{-x^2}$ 就是极为重要的高斯分布。下面所要讨论的就是如何巧妙地运用指数函数的特性和几何对称性相结合, 求解上述积分求值问题。我们把它编列成几个习题以利读者从中体认其中之妙处。]

(12) 试用积分求和的定义和分配律说明

$$\int_{-K}^K e^{-x^2} dx \cdot \int_{-K}^K e^{-y^2} dy$$

的几何意义就是在 $\square(2K, 2K)$ 之上介于 x - y 坐标面和曲面 $z = e^{-(x^2+y^2)}$ 之间的体积。由此可见

$$\int_{-\infty}^{\infty} e^{-x^2} dx \cdot \int_{-\infty}^{\infty} e^{-y^2} dy \quad \left(= \Gamma\left(\frac{1}{2}\right)^2 \right)$$

的几何意义就是那个介于 x - y 坐标面和曲面 $z = e^{-(x^2+y^2)}$ 之间的总体积 V 。

(13) 试用上述曲面对于 z -轴的旋转对称性, 采用以 z 为对称轴的圆柱薄片之分割, 说明其体积为

$$\begin{aligned} V &= \int_0^{\infty} e^{-r^2} 2\pi r dr \\ &= \pi \int_0^{\infty} e^{-u} du, \quad u = r^2 \\ &= \pi \end{aligned}$$

由此即得证 $\int_{-\infty}^{\infty} e^{-x^2} dx = \Gamma(\frac{1}{2}) = \sqrt{\pi}$ 。

[我们还可以把上述技巧用来求解 n -维欧氏空间中的单位球面的 $(n-1)$ -维体积之公式。令其为 ω_n ，则已知有 $\omega_2 = 2\pi$, $\omega_3 = 4\pi$ (Archimedes 定理)。

(14) 令 (x_1, \dots, x_n, z) 为 $(n+1)$ -维欧氏空间的笛氏正交坐标。试说明

$$(\sqrt{\pi})^n = \int_{-\infty}^{\infty} e^{-x_1^2} dx_1 \cdot \int_{-\infty}^{\infty} e^{-x_2^2} dx_2 \cdots \int_{-\infty}^{\infty} e^{-x_n^2} dx_n$$

的几何意义就是介于 $z = 0$ 的 n -维坐标面和 n -维超曲面 $z = e^{-(x_1^2 + \dots + x_n^2)}$ 之间的 $(n+1)$ -维体积。

(15) 再用上述超曲面的球对称性说明

$$\begin{aligned} \pi^{\frac{n}{2}} &= \int_0^{\infty} e^{-r^2} (\omega_n r^{n-1}) dr \\ &= \frac{1}{2} \omega_n \int_0^{\infty} e^{-u} u^{\frac{n}{2}-1} du = \frac{1}{2} \omega_n \Gamma\left(\frac{n}{2}\right) \\ \Rightarrow \omega_n &= \frac{2\pi^{\frac{n}{2}}}{\Gamma\left(\frac{n}{2}\right)} \end{aligned}$$

再用[习题 9]和[习题 13]即可写下 ω_n 的明确公式 (分成 n 是偶奇这样两种情形)。

第四章

初等函数及其应用举例

函数的范畴是极其多样广泛的。常见常用的多项式函数、三角函数、反三角函数、指数函数和对数函数乃是其中极小部分。因为它们的性质简单，易学好用，所以称之为初等函数。其实它们不但常用好用，而且也是用来研讨其他各种各样有用的非初等函数的基本工具。

4.1 多项式函数

在各种各样函数之中，最为简朴者首推多项式函数。一个给定的多项式函数 $f(x)$ ，其在 a 点的函数值可以由所给的代数式直接计算而得。再者，一个多项式 $f(x)$ 在 a 点邻近的局部性质的研讨是非常简单明了的。我们可以用替换 $x = a + t$ 把它改写成 t 的升幂多项式，亦即

$$(4.1) \quad f(x) = f(a + t) = f(a) + c_1 t + c_2 t^2 + \dots + c_n t^n$$

其中 $|t|$ 是足够小者，因而右侧各项的绝对值乃是随著次数的升高而大幅缩小，因此它们的局部影响力显然是高次项要远小于低次项，可以说乃是「阶段分明，一目了然」者也。由此可见，要研究一般函数在 a 点邻近的局部性质的一个好办法就是设法用多项式函数去局部逼近它，从而把所给函数在 a 点的局部性质的研讨，归于其局部逼近多项式的局部性质而研讨之。其实，这也就是「微分学」的基本思想！当所给函数 $f(x)$ 在 a 的邻近是 n -阶连续可微时，泰勒定理所证者，就是

$f(x)$ 在 a 点邻近和下述多项式函数

$$(4.2) \quad f(a) + f'(a)(x-a) + \frac{f''(a)}{2!}(x-a)^2 + \dots + \frac{f^{(n)}(a)}{n!}(x-a)^n$$

的差别乃是一个比之于 $|x-a|^n$ 更加微小者也, 亦即上述多项式乃是 $f(x)$ 在 a 点的邻近的 n -阶局部逼近。其实, 这也就是一个函数 $f(x)$ 是 n -次连续可微和它在给定点 a 的 $\{f^{(k)}(a), 1 \leq k \leq n\}$ 的实质内涵。所以微分的真谛在于「局部多项式化」。本节将对多项式的常用基本性质和某些初步应用作一简述。其中关于多项式的唯一性定理与插值公式以及单元多项式的除法与辗转相除法求最高公因式这两点业已在 §1.3.1 和 §1.3.2 这两小节讨论过了。再者, 在 §1.3.3 节所证的 Sturm 定理和 §1.3.4 节所证的代数基本定理则分别是实变多项式和复变多项式的基本定理, 它们是实数连续性在代数学中的重要应用。

4.1.1 n -阶密切多项曲线

设 $g(x)$ 是一个在 a 点连续可微的函数, 则其在 $A = (a, g(a))$ 点的切线乃是过曲线上邻近一点 $B = (b, g(b))$ 的割线 AB 在 $b \rightarrow a$ 的极限, 即

$$(4.3) \quad y = g(a) + g'(a)(x-a)$$

现设 $g(x)$ 是一个在 a 点邻近 n -阶连续可微者。在 a 点邻近取相异之 $(n+1)$ 点 $\{a_0 = a, a_j, 1 \leq j \leq n\}$, 则唯一存在著一条次数至多为 n 的多项曲线, 它过曲线 $y = g(x)$ 上的 $(n+1)$ 点 $\{A_j = (a_j, g(a_j)), 0 \leq j \leq n\}$ 。试问当 $\{a_j, 1 \leq j \leq n\}$ 都趋于 $a_0 = a$ 时, 上述多项曲线是否有唯一的极限曲线? 若有, 则可称之谓曲线 $y = g(x)$ 在 $A = (a, g(a))$ 点的 n -阶密切多项曲线, 它是切线的自然高阶推广。在 $n=1$ 的情形, 割线的方程式就是

$$(4.4) \quad y = g(a) + g'(\xi)(x-a), \quad a < \xi < b \text{ (或 } b < \xi < a)$$

所以当 $b \rightarrow a$ 时, 显然那个夹逼于 a, b 之间的 ξ 也趋于 a 为其极限, 再由 $g'(x)$ 在 a 的连续性即有 $g'(\xi) \rightarrow g'(a)$ 。由此可见, $n=1$ 的情形乃是微分均值定理的直接推论。由此可以想到上述高阶密切多项曲线的唯一存在性, 理当是高阶微分均值定理 (亦即泰勒公式) 的推论。

为了便于论证，我们不妨设 $a = 0$ 而不失其所证的一般性 (without loss of generality)。由公式 (1.33) 那个过 $\{A_j = (a_j, g(a_j)), 0 \leq j \leq n\}$ 的多项曲线的方程式乃是

$$(4.5) \quad \begin{vmatrix} -y & 1 & x & \dots & x^n \\ -g(0) & 1 & a_0 & \dots & a_0^n \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ -g(a_j) & 1 & a_j & \dots & a_j^n \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ -g(a_n) & 1 & a_n & \dots & a_n^n \end{vmatrix} = 0 \quad (a_0 = 0)$$

让我们先考虑 $a_1 \rightarrow 0$ 而其他 $a_j, 2 \leq j \leq n$, 暂且维持不动的极限曲线。当我们在 (4.5)-式中把第三行 (third row) 换成其极限值 $(-g(0), 1, 0, \dots, 0)$ 时，行列式的第二、第三行变成相同者，所以 (4.5)-式蜕化成 $0 \equiv 0$ 。因此在求 $a_1 \rightarrow 0$ 的极限曲线时，我们必须先行找出使得 (4.5)-式蜕化的原由而先行排除之，然後才能求其极限而不致于落空成全然无用的「 $0 \equiv 0$ 」。将均值定理和行列式的基本性质作如下结合，就可以把 (4.5)-式中的第三行先改换成

$$(4.6) \quad (-g(a_1) + g(0), 0, a_1, \dots, a_1^n)$$

然後再用均值定理把它改写成

$$(4.6') \quad (-a_1 g'(\xi_1), 0, a_1, \dots, a_1^n)$$

其中 ξ_1 乃是夹逼于 0 和 a_1 之间者。因此我们就成功地找出 (4.5)-式在 $a_1 \rightarrow 0$ 之下蜕化的原由，因为它含有 a_1 因子。所以我们只要先把 a_1 先行除去，然後再求 $a_1 \rightarrow 0$ 的极限，即可得其极限曲线之方程式，它就是把 (4.5)-式中的第三行换成

$$(19)_3 \quad (-g'(0), 0, 1, 0, \dots, 0)$$

者也。有了上述初步成功的经验，我们当然还可以把上述所得的 (19)-式中的第四行先行改换成

$$(4.8) \quad (-g(a_2) + g(0) + a_2 g'(0), 0, 0, a_2^2, \dots, a_2^n)$$

然後再用 2-阶均值定理把它改写成

$$(4.9) \quad \left(-\frac{a_2^2}{2!}g''(\xi_2), 0, 0, a_2^2, \dots, a_2^n\right)$$

其中 ξ_2 乃是夹逼于 0 和 a_2 之间者。如此所得者可以提出公因子 a_2^2 ，所以我们只要先行把它除去然後再求 $a_2 \rightarrow 0$ 的极限，就可以得其极限曲线之方程式，它就是把 (19)-式中的第四行换成

$$(22)_4 \quad \left(-\frac{1}{2!}g''(0), 0, 0, 1, 0, \dots, 0\right)$$

者也。如此逐次运用行列式的基本性质和高阶均值定理，计算让 a_k , $3 \leq k \leq n$, 依次趋于 0 的极限方程式，即可得 n -阶密切多项曲线的方程式如下，即

$$(4.11) \quad \begin{vmatrix} -y & 1 & x & x^2 & \dots & \dots & x^n \\ -g(0) & 1 & 0 & 0 & \dots & \dots & 0 \\ -g'(0) & 0 & 1 & 0 & \dots & \dots & 0 \\ \vdots & \vdots & \ddots & \ddots & \ddots & \dots & 0 \\ -\frac{g^{(k)}(0)}{k!} & 0 & \dots & 0 & 1 & 0 \dots & 0 \\ \vdots & \vdots & \ddots & \vdots & \ddots & \ddots & \vdots \\ -\frac{g^{(n)}(0)}{n!} & 0 & \dots & 0 & \dots & \dots & 0 & 1 \end{vmatrix} = 0$$

将 (4.11)-式的行列式按第一列展开，即得

$$(4.11') \quad -y + g(0) + g'(0)x + \frac{g''(0)}{2!}x^2 + \dots + \frac{g^{(k)}(0)}{k!}x^k + \dots + \frac{g^{(n)}(0)}{n!}x^n = 0$$

上述计算证明了曲线 $y = g(x)$ 在 a 点的 n -阶密切多项曲线的唯一存在性，而且它的方程式就是下述 (4.11'') (只要 $g(x)$ 在 a 点邻近是 n 次连续可微的)。

$$(4.11'') \quad y = g(a) + g'(a)(x-a) + \frac{g''(a)}{2!}(x-a)^2 + \dots \\ + \frac{g^{(k)}(a)}{k!}(x-a)^k + \dots + \frac{g^{(n)}(a)}{n!}(x-a)^n$$

上式的右侧也就是 $g(x)$ 在 a 点邻近的 n -阶局部逼近多项式。

4.1.2 高阶局部逼近与不定式之极限

当一个多项式 $f(x)$ 含有 $(x-a)$ 因式，则显然有 $f(a) = 0$ 。而上述命题的逆命题也成立则是初等代数中一个易证而且常用的「余式定理」。上述基本而且好用的事实不难通过高阶可微函数的多项式高阶局部逼近，推广到高阶可微函数的范畴，亦即

一个局部高阶可微函数 $f(x)$ 含有 $(x-a)^k$ 因子的充要条件是 $f(a) = f'(a) = \dots = f^{(k-1)}(a) = 0$ ，而且它的局部表达式可以写成

$$(4.12) \quad f(x) = (x-a)^k \frac{f^{(k)}(\xi)}{k!}$$

其中 ξ 是一个适当的介于 a, x 之数值。

这其实就是 k -阶 Taylor 公式的特殊情形。再者，若 $f^{(k)}(a) \neq 0$ ，则 $(x-a)^k$ 乃是局部函数 $f(x)$ 所含有的 $(x-a)^\alpha$ 因子的最高次者。我们将称 $f(x)$ 为一个在 a 的邻近的 k -阶微量。

设 $f(x)$ 和 $g(x)$ 分别是在 a 的邻近的 k -阶和 ℓ -阶微量。在研讨两者比值在 $x \rightarrow a$ 的极限值时，自然要先行约去两者所共有的 $(x-a)$ 因子，亦即

$$(4.13) \quad \frac{f(x)}{g(x)} = \begin{cases} (x-a)^{k-\ell} \frac{\ell! f^{(k)}(\xi)}{k! g^{(\ell)}(\eta)}, & k > \ell \\ \frac{f^{(k)}(\xi)}{g^{(\ell)}(\eta)}, & k = \ell \\ \frac{1}{(x-a)^{\ell-k}} \frac{\ell! f^{(k)}(\xi)}{k! g^{(\ell)}(\eta)}, & k < \ell \end{cases}$$

由此易见

$$(4.14) \quad \lim_{x \rightarrow a} \frac{f(x)}{g(x)} = \begin{cases} 0, & k > \ell \\ \frac{f^{(k)}(a)}{g^{(k)}(a)}, & k = \ell \end{cases}$$

而在 $k < \ell$ 的情形则上述极限不存在。这也就是通常在初等微积分中用 l'Hospital Rule 求不定式之极限之所基，乃是 Taylor 公式的直接推论。

【例子】：

$$1. \lim_{x \rightarrow 0} \frac{\sin x}{x} = 1$$

$$2. \lim_{x \rightarrow 0} \frac{\cos x - 1}{x^2} = -\frac{1}{2}$$

$$3. \lim_{x \rightarrow 0} \frac{x - \sin x}{x^3} = \frac{1}{6}$$

$$4. \lim_{x \rightarrow 0} \frac{\tan x}{\sin x} = 1$$

4.1.3 插值问题的推广

设 $y = f(x)$ 是一个连续可微函数的图象，而 $P_i(x_i, f(x_i))$, $1 \leq i \leq \ell$ ，则是其上给定的相异 ℓ 个点。试求一个次数至多为 $2\ell - 1$ 的多项曲线 $y = g(x)$ ，它和 $y = f(x)$ 相切于上述 ℓ 点，亦即有

$$(4.15) \quad g(x_i) = f(x_i), \quad g'(x_i) = f'(x_i), \quad 1 \leq i \leq \ell$$

[因为上述给定条件式的个数是 2ℓ ，而次数至多为 $2\ell - 1$ 的多项式的维数也是 2ℓ 。所以上述求解问题理当是有解的。]

其实，我们还可以把上述问题更加推广如下：

设 $y = f(x)$ 是一个 k -阶连续可微函数的图象而 $P_i(x_i, f(x_i))$, $1 \leq i \leq \ell$ ，则是其上给定的相异 ℓ 个点。试求一个次数至多为

$$\sum_{i=1}^{\ell} k_i + (\ell - 1), \quad 0 \leq k_i \leq k$$

的多项曲线 $y = g(x)$ ，它和 $y = f(x)$ 在 P_i 点分别是 k_i -阶密切相交，亦即有

$$(4.16) \quad g(x_i) = f(x_i), \dots, g^{(k_i)}(x_i) = f^{(k_i)}(x_i) \quad (1 \leq i \leq \ell)$$

[上述条件式的个数乃是 $\sum_{i=1}^{\ell} (k_i + 1)$ 。]

【分析】：

上述问题乃是插值问题和密切多项曲线问题的自然结合；而密切 k_i -阶则乃是把 $(k_i + 1)$ 个交点缩合到一个点的极限情形。由此可见，上述问题是可以由下述途径去求解的，即先在每个 P_i 的邻近再取 $y = f(x)$ 上之 k_i 点，令其为 $\{P_{i,j} = (x_{i,j}, f(x_{i,j})), 1 \leq j \leq k_i\}$ 。我们可以用插值公式 (1.33) 写下过上述总共有 $\sum_{i=1}^{\ell} (k_i + 1)$ 个点的多项曲线的方程式，然后再用 §1.1 节求密切多项曲线的办法，结合行列式的性质和 Taylor 公式去求得 $\{P_{i,j}, 1 \leq j \leq k_i\}$ 都趋于 P_i 为极限的极限方程式。兹以 $i = 3, k_1 = 2, k_2 = 1, k_3 = 0$ 的情形写下其所得的极限方程式如下：

$$(4.17) \quad \begin{vmatrix} -y & 1 & x & x^2 & x^3 & x^4 & x^5 \\ -f(x_1) & 1 & x_1 & x_1^2 & x_1^3 & x_1^4 & x_1^5 \\ -f'(x_1) & 0 & 1 & 2x_1 & 3x_1^2 & 4x_1^3 & 5x_1^4 \\ -f''(x_1) & 0 & 0 & 2 & 6x_1 & 12x_1^2 & 20x_1^3 \\ -f(x_2) & 1 & x_2 & x_2^2 & x_2^3 & x_2^4 & x_2^5 \\ -f'(x_2) & 0 & 1 & 2x_2 & 3x_2^2 & 4x_2^3 & 5x_2^4 \\ -f(x_3) & 1 & x_3 & x_3^2 & x_3^3 & x_3^4 & x_3^5 \end{vmatrix} = 0$$

【习题】：试验证上述多项曲线满足下列条件，即

$$y' \Big|_{x_1} = f'(x_1), \quad y'' \Big|_{x_1} = f''(x_1), \quad y' \Big|_{x_2} = f'(x_2)$$

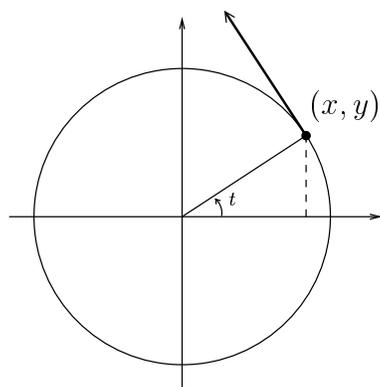
而且过 $P_i(x_i, f(x_i)), 1 \leq i \leq 3$, 这三点。

4.2 三角函数与反三角函数

正弦、余弦函数是定量研讨几何不可或缺的基本函数，它们是解析描述圆与角，三角形的角边关系中自然产生者。用现代的观点来说，如 [图 4-1] 所示

$$x = \cos t, \quad y = \sin t$$

就是单位圆上单位等速运动的动态描述。

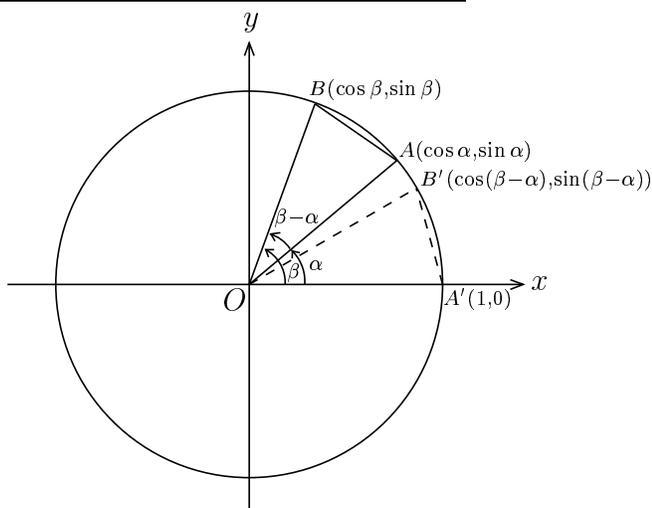


[图 4-1]

正弦、余弦函数乃是一对用来描述单位圆的函数。而圆的本质则在于其完美的对称性，亦即其旋转对称性和对于每一个直径的反射对称性。是以正弦、余弦函数的基本性质，其实也就是圆的对称性的解析表现。

4.2.1 圆的对称性与正弦、余弦函数的基本性质

(i) 圆的对称性与正弦、余弦函数的和角公式：



[图 4-2]

如 [图 4-2] 所示，由圆的旋转对称性可见 $A = (\cos \alpha, \sin \alpha)$ 和 $B = (\cos \beta, \sin \beta)$ 之间的距离之平方仅仅和 $(\beta - \alpha)$ 有关，亦即

$$\begin{aligned}\overline{AB}^2 &= (\cos \beta - \cos \alpha)^2 + (\sin \beta - \sin \alpha)^2 \\ &= (\cos(\beta - \alpha) - 1)^2 + \sin^2(\beta - \alpha) = \overline{A'B'}^2\end{aligned}$$

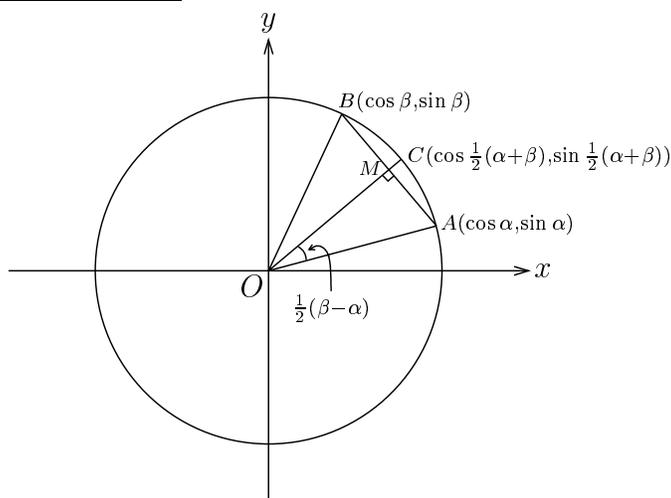
$$\Rightarrow \cos(\beta - \alpha) = \cos \alpha \cos \beta + \sin \alpha \sin \beta$$

$$\Rightarrow \cos(\beta + \alpha) = \cos \alpha \cos \beta - \sin \alpha \sin \beta$$

$$\begin{aligned}\Rightarrow \sin(\beta - \alpha) &= \cos\left(\frac{\pi}{2} - \beta + \alpha\right) \\ &= \cos \alpha \cos\left(\frac{\pi}{2} - \beta\right) - \sin \alpha \sin\left(\frac{\pi}{2} - \beta\right) \\ &= \cos \alpha \sin \beta - \sin \alpha \cos \beta\end{aligned}$$

$$\Rightarrow \sin(\beta + \alpha) = \cos \alpha \sin \beta + \sin \alpha \cos \beta$$

(ii) 反射对称性与和化积公式



[图 4-3]

如 [图 4-3] 所示，等腰三角形 $\triangle OAB$ 对于 OM 成反射对称：

$$\overrightarrow{OM} = \frac{1}{2}(\overrightarrow{OA} + \overrightarrow{OB}) = \left(\frac{1}{2}(\cos \alpha + \cos \beta), \frac{1}{2}(\sin \alpha + \sin \beta)\right)$$

$$\overrightarrow{OM} = \cos \frac{1}{2}(\alpha - \beta) \cdot \overrightarrow{OC}, \quad \overrightarrow{OC} = \left(\cos \frac{1}{2}(\alpha + \beta), \sin \frac{1}{2}(\alpha + \beta)\right)$$

所以即有

$$\begin{aligned}\frac{1}{2}(\cos \alpha + \cos \beta) &= \cos \frac{1}{2}(\alpha - \beta) \cos \frac{1}{2}(\alpha + \beta) \\ \frac{1}{2}(\sin \alpha + \sin \beta) &= \cos \frac{1}{2}(\alpha - \beta) \sin \frac{1}{2}(\alpha + \beta)\end{aligned}$$

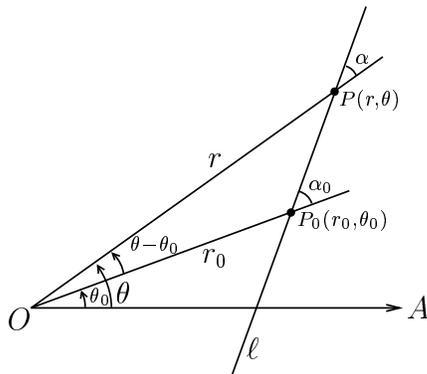
这也就是和化积公式。

4.2.2 三角定律与极坐标

一个三角形的三个角度和三个边长乃是它的六个基本几何量，而三角形的诸多叠合条件则说明六者之间是具有相互关联的函数关系的。这也就是熟知的正弦定律和余弦定律，即

$$\begin{cases} \frac{\sin A}{a} = \frac{\sin B}{b} = \frac{\sin C}{c} = \frac{2\Delta}{abc} \\ \cos A = \frac{b^2 + c^2 - a^2}{2bc} \\ \cos B = \frac{a^2 + c^2 - b^2}{2ac} \\ \cos C = \frac{a^2 + b^2 - c^2}{2ab} \end{cases}$$

极坐标的直线方程与距离公式：



[图 4-4]

如 [图 4-4] 所示，过 $P_0(r_0, \theta_0)$ 点而且和矢径 $\overline{OP_0}$ 的夹角为 α_0 的直

线方程式，可以用 $\triangle OP_0P$ 的正弦定律求之如下：

$$\frac{\sin(\pi - \alpha_0)}{r} = \frac{\sin \alpha}{r_0}, \quad \alpha_0 = \alpha + \theta - \theta_0$$

$$\Rightarrow r \sin(\alpha_0 + \theta_0 - \theta) = r_0 \sin \alpha_0$$

再者 $P_0(r_0, \theta_0)$ 和 $P(r, \theta)$ 之间的距离平方则可以用 $\triangle OP_0P$ 的余弦定律计算如下，即

$$\overline{P_0P}^2 = r^2 + r_0^2 - 2rr_0 \cos(\theta - \theta_0)$$

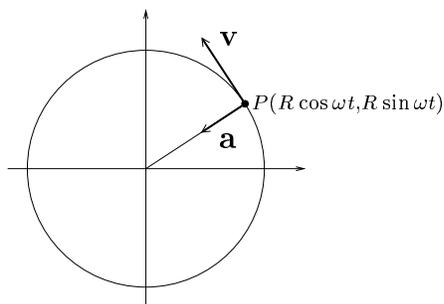
4.2.3 等速圆周运动与正弦、余弦的微分

$$x = R \cos \omega t, \quad y = R \sin \omega t$$

所描述者乃是一个以原点为圆心，半径为 R 的圆上的运动，其角速度等于常数 ω 。其速度向量和加速度向量分别为

$$\mathbf{v} = \left(\frac{dx}{dt}, \frac{dy}{dt} \right) = \omega R (-\sin \omega t, \cos \omega t)$$

$$\mathbf{a} = \left(\frac{d^2x}{dt^2}, \frac{d^2y}{dt^2} \right) = -\omega^2 R (\cos \omega t, \sin \omega t)$$



[图 4-5]

由此可见，其向心加速度的大小等于 $R\omega^2$ 。令 $z(t) = \cos t + i \sin t$ ，则易见它满足微分方程

$$\frac{d}{dt}z(t) = -\sin t + i \cos t = iz(t)$$

显然它的（复）常数倍， $(A + iB)z(t)$ ，也满足同样的微分方程。

【定理 4.1】：设复值非零实变函数 $w(t) = f(t) + ig(t)$ 满足常微分方程

$$\frac{d}{dt}w(t) = iw(t)$$

则存在适当的 $(A + iB)$ 使得

$$\begin{aligned} w(t) &= (A + iB)(\cos t + i \sin t) \\ &= \sqrt{A^2 + B^2} \left(\cos\left(t + \tan^{-1} \frac{B}{A}\right) + i \sin\left(t + \tan^{-1} \frac{B}{A}\right) \right) \end{aligned}$$

证明：令 $\varphi(t) = w(t)(\cos t - i \sin t)$ ，由所设即有

$$\begin{aligned} \frac{d}{dt}\varphi(t) &= \left[\frac{d}{dt}w(t) \right] (\cos t - i \sin t) + w(t)(-\sin t - i \cos t) \\ &= iw(t)(\cos t - i \sin t) - iw(t)(\cos t - i \sin t) \\ &\equiv 0 \end{aligned}$$

亦即 $\varphi(t)$ 乃是一个 (复) 常数 $(A + iB)$ 。 □

【定理 4.2】：正弦、余弦函数可以用下述幂级数表达，即有

$$\begin{aligned} \sin x &= x - \frac{1}{3!}x^3 + \frac{1}{5!}x^5 \mp \dots \\ &= \sum_{n=0}^{\infty} (-1)^n \frac{1}{(2n+1)!} x^{2n+1} \\ \cos x &= 1 - \frac{1}{2!}x^2 + \frac{1}{4!}x^4 \mp \dots \\ &= \sum_{n=0}^{\infty} (-1)^n \frac{1}{(2n)!} x^{2n} \end{aligned}$$

或者可以将上述两个公式合写成

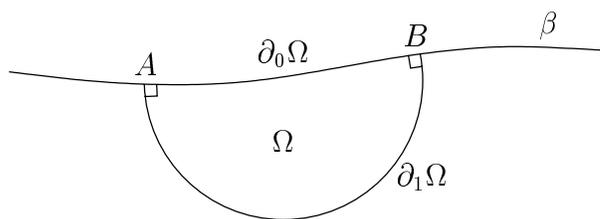
$$\cos x + i \sin x = \sum_{n=0}^{\infty} \frac{1}{n!} (ix)^n$$

证明：当 $f(x)$ 是 $\sin x$ 或 $\cos x$ 时，易见 $f^{(4)}(x)$ 就是 $f(x)$ 本身，而且其泰勒公式中的余项 R_n 在 $n \rightarrow \infty$ 时显然趋于 0。 □

4.2.4 等周问题 (Isoperimetric Problem)

设 Ω 是平面 (或半球面, 或非欧面) 上的一个区域, $\partial\Omega$ 是它的周边, 通常以 $|\Omega|$ 和 $|\partial\Omega|$ 分别表示其面积和长度。等周问题所研讨者乃是在所有周长取某一定值的各种各样区域之中, 以那种区域的面积为极大? 其实, 还有另外两种等周问题, 亦即下述具有给定边界的等周问题和具有给定边界而且固定其端点或端点之一的等周问题, 而原前者则称之为无边界的等周问题。

【具有给定边界的等周问题】: 设 β 是一条给定的边界, 如 [图 4-6] 所示, Ω 是一个位于 β 的给定一侧, 借用 β 为其部份周边的区域。[例如当年希腊神话中的 Queen Dido, 在北非以地中海的一段海岸为自然屏障构筑城墙所围成的城区。]



[图 4-6]

令 $\partial_0\Omega$ 和 $\partial_1\Omega$ 分别是 $\partial\Omega$ 中属于和不属于 β 者。求解 $|\partial_1\Omega|$ 取某一定值的区域之中, $|\Omega|$ 取极大值者。

【具有给定边界和端点的等周问题】: 假如在上述极值问题之中, 我们还加以 $\partial_1\Omega$ 的两个端点 (或其中之一) 为某两 (或某一) 取定点的限制, 则称之为具有给定边界和端点的等周问题。[设想当年 Queen Dido 所在的北非地中海海岸, 恰好有一个或两个相距适中的小山丘。则基于有利于防守的考量, 很自然会选用它们作为构筑城墙 (亦即 $\partial_1\Omega$) 的端点, 是不?]

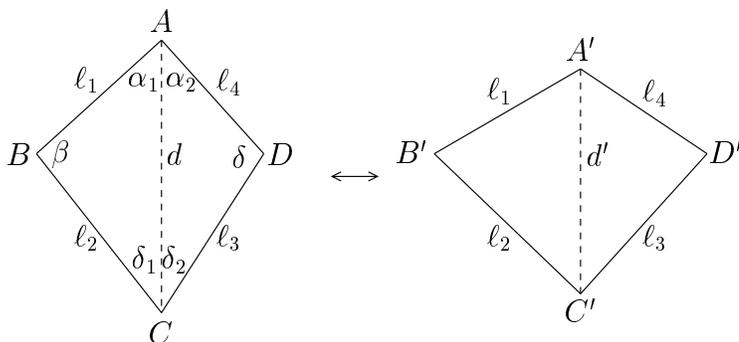
总之, 等周问题很自然地有三种情形, 亦即平面、半球面和非欧面这三种情形; 而且在每一种情形下, 又具有三个类型, 亦即无边界, 具有给定边界和具有给定边界和端点这样三种提法。由此可见, 总共有九个类同但是各异的等周问题, 它们都是几何学中自然而且重要的极值问题。而它们的解答就是:

Ω 是第一类 (或第二类, 第三类) 等周问题之解的充要条件
 乃是 $\partial\Omega$ (或 $\partial_1\Omega$) 是给定周长的圆 (或给定长度的圆弧),
 而且第二类的等周解 $\partial_1\Omega$ 还必须和给定边界 β 正交。

自古以来, 等周问题一直是一个引人入胜的课题, 所以有不少数学家对它提出各种各样的证明, 不胜枚举。而且它们所适用的广度和所需要的假设也各有不同。但是大家所追寻的「皓的」, 乃是那种既简朴初等又具有普遍适用性的统一证法。以下所讨论的初等证明, 乃是一个普遍适用于上述九种等周问题的统一证法; 在这里我们暂且只集中讨论在欧氏平面的情形。这个证明的关键就是下述十分初等的极值问题之解的共圆性。

一个初等的基本引理 :

一个三角形的三边边长业已构成其一组叠合条件, 所以三角形的面积当然业已由其三边边长所唯一确定, 而用其边长去表达面积的公式也就是秦九韶公式 (亦即海伦公式)。但是给定四边边长的四边形, 还可以如 [图 4-7] 所示, 作那种变更其对角线长的变形, 所以它们的面积并非由给定边长所唯一确定, 而是由给定边长和一条选定的对角线长所唯一确定 (例如 [图 4-7] 所示的 $d = \overline{AC}$)。



[图 4-7]

由此可见, 上述四边边长取定不变的四边形的面积乃是其一条选定的对角线长的函数。这个函数在什麼情形之下得极大值呢? 这乃是欧氏平面几何学中一个既初等又基本的极值问题。

【基本引理】: 在四边边长给定的四边形中, 其面积在内接于圆时为极大 (亦即其四个顶点共圆者也)。

证明：令 Δ_1 和 Δ_2 分别为 $\triangle ABC$ 和 $\triangle ACD$ 的面积， β 和 δ 分别为 B 和 D 点的内角。再者，以 $x = d^2$ 为自变数，则有

$$(4.18) \quad \begin{aligned} 2\ell_1\ell_2 \cos \beta &= \ell_1^2 + \ell_2^2 - x \\ 2\ell_3\ell_4 \cos \delta &= \ell_3^2 + \ell_4^2 - x \\ \Delta_1 &= \frac{1}{2}\ell_1\ell_2 \sin \beta, \quad \Delta_2 = \frac{1}{2}\ell_3\ell_4 \sin \delta \end{aligned}$$

将上列各式分别对于 x 求微分，即得

$$(4.19) \quad \begin{aligned} 2\ell_1\ell_2 \sin \beta \frac{d\beta}{dx} &= 1 \quad \Rightarrow \quad \frac{d\beta}{dx} = \frac{1}{2\ell_1\ell_2 \sin \beta} \\ 2\ell_3\ell_4 \sin \delta \frac{d\delta}{dx} &= 1 \quad \Rightarrow \quad \frac{d\delta}{dx} = \frac{1}{2\ell_3\ell_4 \sin \delta} \\ \frac{d\Delta_1}{dx} + \frac{d\Delta_2}{dx} &= \frac{1}{2}\ell_1\ell_2 \cos \beta \frac{d\beta}{dx} + \frac{1}{2}\ell_3\ell_4 \cos \delta \frac{d\delta}{dx} \\ &= \frac{1}{4}(\cot \beta + \cot \delta) \end{aligned}$$

由此即见 $\Delta_1 + \Delta_2$ 得极大值的条件乃是

$$(4.20) \quad \cot \beta + \cot \delta = 0 \quad \Longleftrightarrow \quad \beta + \delta = \pi$$

亦即四边形 $\square ABCD$ 内接于圆。 □

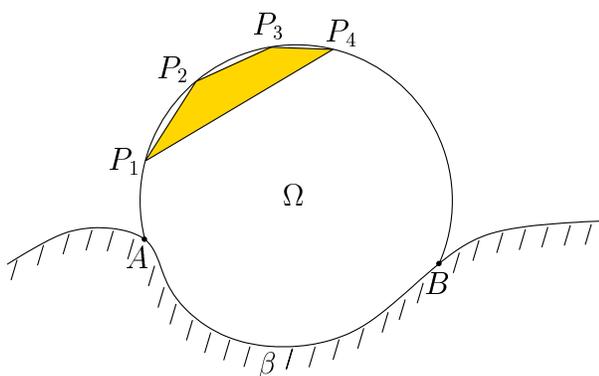
等周问题的一个初等证明（欧氏平面情形）：

如上所述，等周问题共有九种类型，而相应的九种等周解，它们都具有简洁完美的共性，那就是其边界的「非给定」部份（亦即 $\partial\Omega$ 或 $\partial_1\Omega$ ）乃是给定长度的圆弧。现在我们要讨论的欧氏平面等周解的「圆性」(circularity) 的证明，其实含有下述唯一性和存在性这两个方面，即

唯一性：在 Ω 是某一种的等周解的假设之下，证明 $\partial\Omega$ （或 $\partial_1\Omega$ ）必须是给定长度的圆弧。

存在性：证明某一种的等周问题的极大面积解是的确存在的。

唯一性的证明：如 [图 4-8] 所示，设 Ω 乃是欧氏平面中给定边界和端点的等周解。



[图 4-8]

在 $\partial_1\Omega$ 上任取邻近四点 $\{P_i, 1 \leq i \leq 4\}$ 。易见四边形内接于 Ω 。假若不然，则可用反射对称把位于四边形内部的 $\partial_1\Omega$ 之弧段，改用其对称弧段。如此所得的区域 Ω' 显然有 $|\partial_1\Omega'| = |\partial_1\Omega|$ ，但是 $|\Omega'| > |\Omega|$ ，所以是和 Ω 本身乃是一个等周解之所设矛盾！再者， $\{P_i, 1 \leq i \leq 4\}$ 四点必须是共圆的！假若不然，则可以把 Ω 作下述变形，它把 Ω 中位于 $\square P_1P_2P_3P_4$ 之外的部份保持原样，而把 $\square P_1P_2P_3P_4$ 改用具有同样边长但是其四顶点共圆的 $\square P'_1P'_2P'_3P'_4$ 。如此所得的 Ω' 显然保持 $|\partial_1\Omega'| = |\partial_1\Omega|$ ，但是由基本引理，又有

$$(4.21) \quad |\Omega'| - |\Omega| = \square P'_1P'_2P'_3P'_4 - \square P_1P_2P_3P_4 > 0$$

这是和 Ω 本身乃是等周解之所设相矛盾的。再者，因为 $\{P_i, 1 \leq i \leq 4\}$ 是 $\partial_1\Omega$ 上任选的邻近四点，所以它们总是共圆只有在 $\partial_1\Omega$ 本身就是一个圆弧才可能。这也就是用基本引理，直截了当地证明了等周解的唯一性。 \square

存在性的证明：让我们先来讨论具有给定边界和端点的等周解的存在性的证明。设 \mathcal{F} 是所有满足条件 $|\partial_1\Omega| = l_0$ 而且 $\partial_1\Omega$ 以 β 上的给定点 A, B 为其端点的区域 Ω 所组成的集合。令

$$(4.22) \quad M = l.u.b. \{|\Omega|; \Omega \in \mathcal{F}\}$$

存在性的证明要点在于论证 \mathcal{F} 中的确存在著一个 Ω^* 其面积等于上述极小上界 M 。我们将再用前述的基本引理来证上述存在性如下：

由 M 的定义， \mathcal{F} 中显然存在一个序列 $\{\Omega_n\}$ 使得

$$(4.23) \quad \lim_{n \rightarrow \infty} |\Omega_n| = M$$

再者，由所设 $|\partial_1 \Omega_n| = \ell_0$ 。我们可以把 $\partial_1 \Omega_n$ 用其上的 $2^n - 1$ 点 $\{P_{n,i}, 1 \leq i \leq 2^n - 1\}$ 分隔成长度为 $\frac{\ell_0}{2^n}$ 的分段。在 n 相当大时，上述分段是十分微短者，所以我们可以用上述唯一性的证明中所用的变形，把原给的 Ω_n 变换成上述 2^n -等分点共圆的区域 Ω'_n ，它显然还是属于 \mathcal{F} 者，而且 $|\Omega'_n| \geq |\Omega_n|$ （等号只有在 $\Omega'_n \cong \Omega_n$ 时才成立）。总之，我们将改用序列 $\{\Omega'_n\}$ ，显然依然有 $\lim_{n \rightarrow \infty} |\Omega'_n| = M$ 。再者，令 Γ_n 为以 A, B 点为其端点而且和 Ω'_n 相交于其所有 2^n -等分点的圆弧，不难看到

$$(4.24) \quad \lim_{n \rightarrow \infty} |\Gamma_n| = \ell_0$$

再者， $\{\Gamma_n\}$ 趋于那个弧长为 ℓ_0 ，以 A, B 为端点，而且和 Ω'_n 居于 β 的同侧的唯一圆弧 Γ^* 为其极限。令 Ω^* 为以 Γ^* 为 $\partial_1 \Omega^*$ 的区域，易见 $|\Omega^*| = M$ 。这也就证明了第三类形的等周解的存在性。□

其实，上面这种证法是很容易直接推广成第一、第二类等周解的存在性的证明的。例如我们只要把 β 取成一条测地线， A, B 取成十分靠近的两点，然後再把上述业已证明其存在性的第三类等周解，求其在 A, B 的距离趋于 0 的极限，即可得出第一类等周解的存在性。再者，第二类等周解的存在性是有赖于边界 β 的某种几何性质的。此事可以用第三类等周解，再让端点 A, B 在 β 上任意变动，去求解其极大值。这也就是 $\partial_1 \Omega$ 和 β 正交这个必要条件的来由。

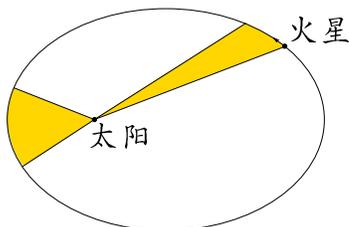
【总结】：上述证明把三种类型等周问题的论证，简明扼要地归结到所证的基本引理，亦即给定边长的四边形以内接于圆者面积为极大。上述简短的论证，直截了当地说明了等周解周边的「圆性」乃是上述初等极值解的顶点共圆性的直接推论。至于在半球面和非欧面的情形，我们也有同样的基本引理，所以其等周解周边的「圆性」也同样是这个基本引理的直接推论。

4.2.5 Kepler 行星运行三定律及其数理分析

Kepler 行星运行三定律清楚地描述了在当年所知的六个行星（包括地球）运行的简单模式：

【第一定律】：火星绕太阳以一个以太阳位于其焦点之一的椭圆运行。

【第二定律】：在单位时间内，火星和太阳连线所扫过的面积恒相等，如 [图 4-9] 所示。



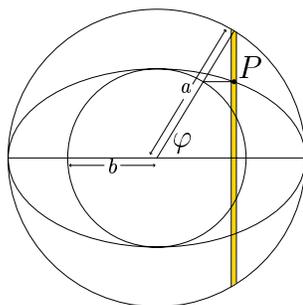
[图 4-9]

【第三定律】：各个行星不但也满足同样的第一、第二定律，而且其椭圆轨道长径的立方和周期的平方之间的比值皆相等。

一个自然的问题就是：「为什么行星会以太阳为其一焦点的椭圆轨道运行？」牛顿对这个问题给了一个直接的答案，即行星与太阳之间存在满足平方反比定律(inverse square law)的引力。而且这种引力同样地存在于任何两个物体之间，不论是火星与太阳或是苹果与地球，都是同一种引力；这就是牛顿著名的万有引力定律。现在让我们以现今的符号体系来重新看看牛顿这个对后世科学发展有著深远影响的工作。首先我们需要对椭圆的几何性质有一定的了解。

(一) 椭圆面积公式：椭圆面积 = πab

设椭圆的长、短径分别为 $2a, 2b$ 。现构造两个圆，半径分别为 a, b ，并把椭圆夹于两者之间，如 [图 4-10] 所示。



$$\begin{aligned} P &= P(x, y) \\ x &= a \cos \varphi \\ y &= b \sin \varphi \end{aligned}$$

[图 4-10]

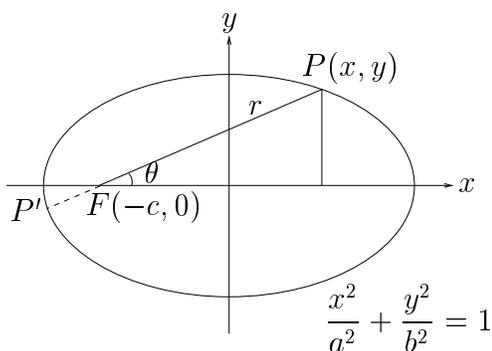
计算在 [图 4-10] 中那条过 $P(x, y)$ 的窄条面积。由于 $x = a \cos \varphi, y = b \sin \varphi$ ，在椭圆内的窄条面积和整条窄条面积的比率约为：

$$\frac{2b \sin \varphi \Delta x}{2a \sin \varphi \Delta x} = \frac{b}{a}$$

这是一个与 $P(x, y)$ 位置无关的常数。当把所有这种窄条的面积加起来时，使得

$$\text{椭圆面积} = \pi a^2 \cdot \frac{b}{a} = \pi ab$$

(二) 椭圆的极坐标方程式



[图 4-11]

如 [图 4-11] 所示，椭圆上任意一点 P 有 (x, y) 和 (r, θ) 两种表示方法，而 (x, y) 和 (r, θ) 之间的转换可以用

$$r \cos \theta = x + c, \quad r \sin \theta = y$$

来达成。把上面转换方法代入熟知的椭圆坐标方程式 $\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$ ，即得

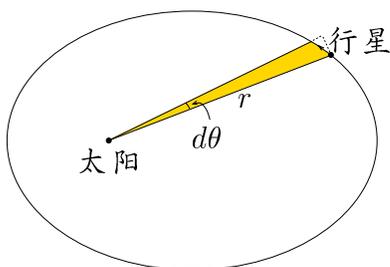
$$\begin{aligned} b^2(r \cos \theta - c)^2 + a^2(r \sin \theta)^2 - a^2b^2 &= 0 \\ b^2(r^2 \cos^2 \theta - 2cr \cos \theta + c^2) + a^2r^2(1 - \cos^2 \theta) - a^2b^2 &= 0 \\ (b^2 - a^2)r^2 \cos^2 \theta - 2b^2cr \cos \theta + b^2c^2 + a^2r^2 - a^2b^2 &= 0 \\ -c^2r^2 \cos^2 \theta - 2b^2cr \cos \theta - b^4 + a^2r^2 &= 0 \quad (a^2 = b^2 + c^2) \\ a^2r^2 - (rc \cos \theta + b^2)^2 &= 0 \\ [ar - (rc \cos \theta + b^2)][ar + (rc \cos \theta + b^2)] &= 0 \end{aligned}$$

因此 $r = \frac{b^2}{a - c \cos \theta}$ (注: $r = \frac{-b^2}{a + c \cos \theta}$ 是负值, 它会以 P' 来描绘出椭圆)。为了方便以后的计算, 我们取其颠倒式为椭圆极坐标方程式:

$$\frac{1}{r} = \frac{a - c \cos \theta}{b^2}$$

(三) 第二定律的数理分析

以太阳(焦点)为中心, 极坐标 (r, θ) 表达行星位置。当 θ 增大到 $\theta + d\theta$ 时, 太阳与行星的连线所扫过的面积为 $dA \approx \frac{1}{2}r^2 d\theta$, 如图 4-12 所示:



[图 4-12]

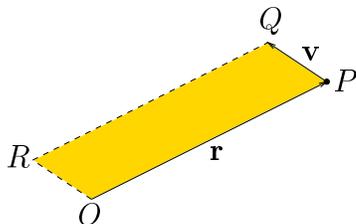
运用第二定律, 这个面积的变率为常数, 即:

$$\frac{dA}{dt} = \frac{1}{2}r^2 \frac{d\theta}{dt} = \frac{1}{2}r^2 \omega = k, \quad (k: \text{常数}, \omega: \text{角速度})$$

$$\text{椭圆面积} = \pi ab = \int_0^T dA = \int_0^T k dt = kT \quad (T: \text{周期})$$

$$r^2 \omega = 2k = \frac{2\pi ab}{T}$$

[注]: 从物理学观点来看, 第二定律是有物理意义的。如图 4-13 所示, \mathbf{r} 是位置向量, \mathbf{v} 是速度向量:



[图 4-13]

$|\mathbf{r} \times \mathbf{v}|$ 是平行四边形 $\square OPQR$ 的面积。但 $\mathbf{r} \times m\mathbf{v}$ 是物体相对于 O 点的角动量，而且因为是平面运动，此向量是恒垂直于平面的，所以由第二定律亦可推出角动量是不变的（这也是人类理性文明中首次接触到角动量守恒定律）。再者，

$$\begin{aligned}\mathbf{0} &= \frac{d}{dt}(\mathbf{r} \times m\mathbf{v}) = \mathbf{v} \times m\mathbf{v} + \mathbf{r} \times m\mathbf{a} \\ &= \mathbf{r} \times \mathbf{F}\end{aligned}$$

所以引力 \mathbf{F} 的作用方向是和 \mathbf{r} 反向平行，亦即是向心的。

(四) 向心加速的公式： $-(r\omega^2 + \frac{d^2r}{dt^2})$

上述公式只需对极坐标 $x = r \cos \theta$, $y = r \sin \theta$ 直接微分便可得出。计算过程大致如下：

$$\begin{aligned}\begin{cases} x = r \cos \theta \\ y = r \sin \theta \end{cases} &\Rightarrow \begin{cases} \dot{x} = \dot{r} \cos \theta - r\omega \sin \theta \\ \dot{y} = \dot{r} \sin \theta + r\omega \cos \theta \end{cases} \quad (\omega = \dot{\theta}) \\ \Rightarrow \begin{cases} \ddot{x} = \ddot{r} \cos \theta - r\omega^2 \cos \theta - 2\dot{r}\omega \sin \theta - r\dot{\omega} \sin \theta \\ \ddot{y} = \ddot{r} \sin \theta - r\omega^2 \sin \theta + 2\dot{r}\omega \cos \theta + r\dot{\omega} \cos \theta \end{cases}\end{aligned}$$

因此

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

由(三)知引力的方向是和位置方向 $(\cos \theta, \sin \theta)$ 相反，所以 $(2\dot{r}\omega + r\dot{\omega})$ 应该是 0，并只馀下向心加速 $-(r\omega^2 + \ddot{r})$ 。

(五) 平方反比定律的证明 (第一、第二定律的综合分析)

要证明引力是满足平方反比定律，我们只需验证 $r^2 \left(\frac{d^2r}{dt^2} - r\omega^2 \right)$ 是否为一常数。先对椭圆的极坐标方程式微分：

$$\begin{aligned}\frac{1}{r} &= \frac{a - c \cos \theta}{b^2} \\ -\frac{1}{r^2} \frac{dr}{dt} &= \frac{c}{b^2} \sin \theta \omega \\ \Rightarrow \frac{dr}{dt} &= -\frac{c}{b^2} \sin \theta r^2 \omega = -\frac{2\pi ac}{bT} \sin \theta\end{aligned}$$

在上式用了 (三) $r^2\omega = \frac{2\pi ab}{T}$ 。同样地在下面的计算中, 我们尽可能分离出 $r^2\omega$ 这一项, 然後换成常数 $\frac{2\pi ab}{T}$ 。

$$\begin{aligned} \frac{d^2r}{dt^2} &= -\frac{2\pi ac}{bT} \cos \theta \omega \\ \Rightarrow r^2 \frac{d^2r}{dt^2} &= -\frac{2\pi ac}{bT} \cdot \frac{2\pi ab}{T} \cos \theta \\ &= \frac{-4\pi^2 a^2}{T^2} c \cos \theta \\ r^2(-r\omega^2) &= -r^3\omega^2 = -\frac{1}{r}(r^2\omega)^2 = -\frac{1}{r} \frac{4\pi^2 a^2 b^2}{T^2} \\ &= -\frac{4\pi^2 a^3}{T^2} + \frac{4\pi^2 a^2}{T^2} c \cos \theta \end{aligned}$$

因此 $r^2 \left[-r\omega^2 + \frac{d^2r}{dt^2} \right] = -\frac{4\pi^2 a^3}{T^2}$ (常数)。

(六) 由行星引力到万有引力 (Newton's Formula)

在这一步, 均匀密度的球面对于球外质点的引力公式乃是关键之所在。请参看第二章之末之证明。

这也就是由 Kepler 行星运行定律的数理分析自然而然地推导出牛顿万有引力定律的一个简朴详尽的叙述。它就是牛顿的科学巨著 *Philosophiae Naturalis Principia Mathematica* 中所讨论的主要结果。它也自然是後学後进应当心领神会, 并从此体会人类理性文明世代相承, 继往开来的精要和精神。

4.2.6 三角函数的积分计算

由 $(\sin x)' = \cos x$ 和 $(\cos x)' = -\sin x$ 这两个基本的结果, 我们可以运用微积分基本定理反过来计算 $\sin x$ 和 $\cos x$ 的积分, 亦即

$$\int_a^x \sin t dt = \cos a - \cos x, \quad \int_a^x \cos t dt = \sin x - \sin a$$

再配合正弦、余弦的半角、和化积等公式, 我们还可以计算一些较为复杂的积分。

【例一】： $\int_a^x \sin^2 t dt = ?$

$$\int_a^x \sin^2 t dt = \int_a^x \left(\frac{1}{2} - \frac{1}{2} \cos 2t \right) dt = \frac{x-a}{2} - \frac{\sin 2x - \sin 2a}{4}$$

【例二】： $\int_{-\pi}^{\pi} \sin mx \cos nx dx = ?$

因为 $\sin mx \cos nx = \frac{1}{2}(\sin(m+n)x - \sin(m-n)x)$ ，所以当 $m \neq n$ 时

$$\begin{aligned} \int_{-\pi}^{\pi} \sin mx \cos nx dx &= \frac{1}{2(m+n)} \{ \cos(m+n)(-\pi) - \cos(m+n)\pi \} \\ &\quad - \frac{1}{2(m-n)} \{ \cos(m-n)(-\pi) - \cos(m-n)\pi \} = 0 \end{aligned}$$

而当 $m = n$ 时， $\sin mx \cos nx = \frac{1}{2} \sin 2mx$ ，所以

$$\int_{-\pi}^{\pi} \sin mx \cos nx dx = \frac{1}{4m} \{ \cos 2m(-\pi) - \cos 2m\pi \} = 0$$

亦即在任何情况之下 $\int_{-\pi}^{\pi} \sin mx \cos nx dx = 0$ 。

4.2.7 反三角函数及 π 的近似值计算

由反函数的微分公式

$$\frac{dx}{dy} = 1 / \frac{dy}{dx}$$

我们可以求出反三角函数的导函数如下：

【例三】： $\frac{d}{dx} \sin^{-1} x = \frac{1}{\sqrt{1-x^2}}$ ， $\frac{d}{dx} \cos^{-1} x = \frac{-1}{\sqrt{1-x^2}}$

令 $y = \sin^{-1} x$ ，则 $x = \sin y$ 。对 y 求导後即得

$$\frac{dx}{dy} = \cos y$$

亦即

$$\frac{dy}{dx} = \frac{1}{\cos y} = \frac{1}{\sqrt{1-x^2}}$$

同理 $\frac{d}{dx} \cos^{-1} x = \frac{-1}{\sqrt{1-x^2}}$ 。

【例四】： $\frac{d}{dx} \tan^{-1} x = \frac{1}{1+x^2}$

令 $y = \tan^{-1} x$ ，则 $x = \tan y$ 。对 y 求导後即得 $\frac{dx}{dy} = \sec^2 y = 1 + \tan^2 y$ ，亦即

$$\frac{dy}{dx} = \frac{1}{1+x^2}$$

在上述 $\tan^{-1} x$ 的导函数表达式中， $\frac{1}{1+x^2}$ 有下述几何级数展开式：

$$\frac{1}{1+x^2} = 1 - x^2 + x^4 - x^6 + \dots + (-1)^{n-1} x^{2(n-1)} + \frac{(-1)^n x^{2n}}{1+x^2}$$

运用微积分基本定理，可得

$$\begin{aligned} \tan^{-1} x &= \int_0^x \frac{dt}{1+t^2} = x - \frac{x^3}{3} + \frac{x^5}{5} - \frac{x^7}{7} + \dots + \\ &\quad (-1)^{n-1} \frac{x^{2n-1}}{2n-1} + \int_0^x \frac{(-1)^n t^{2n}}{1+t^2} dt \end{aligned}$$

当 $|x| \leq 1$ 时：

$$\left| \int_0^x \frac{t^{2n}}{1+t^2} dt \right| \leq \left| \int_0^x t^{2n} dt \right| = \frac{|x|^{2n+1}}{2n+1}$$

这个余项会随著 n 的增大而趋于 0，因此得出：

$$\tan^{-1} x = \int_0^x \frac{dt}{1+t^2} = x - \frac{x^3}{3} + \frac{x^5}{5} - \frac{x^7}{7} \pm \dots \quad -1 \leq x \leq 1$$

这就是 James Gregory 在 1671 年时已得到的 $\tan^{-1} x$ 的幂级数展开式。

当 $x = 1$ 时，左方的 $\tan^{-1} 1 = \frac{\pi}{4}$ ，所以：

$$\frac{\pi}{4} = \tan^{-1} 1 = 1 - \frac{1}{3} + \frac{1}{5} - \frac{1}{7} \pm \dots$$

上述级数提供了一个计算 π 的方法。可惜这个级数的收敛速度实在是太慢了，即使计算了 1,000,000 项，还不能准确至小数点後第七位；但

如果我们能够选取一些较小的 x 代入 $\tan^{-1} x$ 的级数之中，则可以大大提高收敛速度。这也就是 John Machin 在 1706 年提出的公式，而他自己也运用下述公式把 π 计算至 100 个小数位：

$$\frac{\pi}{4} = 4 \tan^{-1} \frac{1}{5} - \tan^{-1} \frac{1}{239}$$

转换方法：基于 $\tan x$ 的和角公式：

$$\tan(\alpha + \beta) = \frac{\tan \alpha + \tan \beta}{1 - \tan \alpha \tan \beta}$$

令 $\alpha = \tan^{-1} A$, $\beta = \tan^{-1} B$ ，则上述和角公式可改写成

$$\alpha + \beta = \tan^{-1} A + \tan^{-1} B = \tan^{-1} \left(\frac{A + B}{1 - AB} \right)$$

所以，若我们可以用较小的 A, B 把 1 重写成 $\frac{A+B}{1-AB}$ ，则因为 $\tan^{-1} A$ 和 $\tan^{-1} B$ 的收敛速度比较快，这样便可以大为改善 $\tan^{-1} 1$ 的收敛速度。

例如：取 $A = \frac{1}{2}$ ，则

$$1 = \frac{\frac{1}{2} + B}{1 - \frac{1}{2}B} \Rightarrow B = \frac{1}{3}$$

亦即 $\tan^{-1} 1 = \tan^{-1} \frac{1}{2} + \tan^{-1} \frac{1}{3}$ 。当然，我们还可以把 $\tan^{-1} \frac{1}{2}$ 和 $\tan^{-1} \frac{1}{3}$ 继续分解下去，便可以进一步提高收敛速度：

$$\begin{aligned} \tan^{-1} \frac{1}{2} &= \tan^{-1} \frac{1}{3} + \tan^{-1} \frac{1}{7} \\ \tan^{-1} \frac{1}{3} &= \tan^{-1} \frac{1}{7} + \tan^{-1} \frac{2}{11} \quad \text{等等} \end{aligned}$$

现在让我们尝试用 Machin 的公式，把 π 计算至小数点后 15 个位。首先我们要决定所需要计算的项数。

若所要求的近似值要准确到小数点后第 15 个位，则需要余项的总和小于 5×10^{-16} ，即

$$\text{余项总和} \lesssim \sum_{k=n+1}^{\infty} \left(\frac{16 \left(\frac{1}{5}\right)^{2k-1}}{2k-1} \right) < 5 \times 10^{-16}$$

由此得出 n 的粗略估计为 $n \geq 11$, 即需要计算 11 项。计算过程大致如下:

$$\begin{aligned}\tan^{-1} \frac{1}{5} &\approx \frac{1}{5} - \frac{(\frac{1}{5})^3}{3} \pm \dots + \frac{(\frac{1}{5})^{21}}{21} = 0.19739555984988076189 \\ \tan^{-1} \frac{1}{239} &\approx \frac{1}{239} - \frac{(\frac{1}{239})^3}{3} \pm \dots + \frac{(\frac{1}{239})^{21}}{21} = 0.0041840760020747238645\end{aligned}$$

由此便可得出 π 的近似值

$$\pi = 16 \tan^{-1} \frac{1}{5} - 4 \tan^{-1} \frac{1}{239} \approx 3.1415926535897932947$$

这个近似值和 π 的实际数值相比

$$\pi = 3.14159265358979323846 \dots$$

其精确度已到达第 16 个位了。

[注] 在往後约二百多年时间, 人们都是采用 Machin 公式或类似的公式来计算 π 的高度精确值。而其中最为注目的就是 William Shanks 在 1853 年发表的 “Rectification of the circle”, 他把 π 计算至 607 个小数位。其实 Shanks 是计算了 707 个位的, 但是他只发表前 607 个值。可惜他在计算过程中出错, 使得第 527 位是错的。不过这个错误要到了 1945 年才被发现出来。

【习题】:

(1) 计算 $\int_{-\pi}^{\pi} \sin mx \sin nx dx$ 。

(2) 计算 $\int_{-\pi}^{\pi} \cos mx \cos nx dx$ 。

(3) 试证明 Machin 公式。

(4) 试用 $\frac{\pi}{4} = 2 \tan^{-1} \frac{1}{3} + \tan^{-1} \frac{1}{7}$ 来计算 π 的 12 个小数位的近似值。

4.3 常系数常微分方程

在解各种各样代数问题时，我们经常要运用到或归于代数方程之求解。相类似的，在各种各样的数理分析，理解自然之中，也经常要运用到或归于解微分方程。而在各种各样微分方程之中，最为简单而且好用者，莫过于常系数常微分方程。这也就是本节所要研讨的课题。

4.3.1 算子符号

本节将以 x 表示自变数， y 表示待求的函数， D 表示微分运算 $\frac{d}{dx}$ ，亦即

$$(4.25) \quad Df(x) = f'(x), D^2f(x) = f''(x), \dots, D^k f(x) = f^{(k)}(x)$$

再者，我们将用 $p(D)$ 表示一个以算子 D 为「不定元」的多项式，即

$$(4.26) \quad p(D) = c_0 D^n + c_1 D^{n-1} + \dots + c_k D^{n-k} + \dots + c_n$$

其中 $c_i \in \mathbb{C}$ 。把 $p(D)$ 想成由所有无穷阶连续可微实变（复值）函数所组成的线性空间 \mathcal{V} 上的算子，亦即

$$(4.27) \quad \begin{aligned} p(D)f(x) = & c_0 D^n f(x) + c_1 D^{n-1} f(x) + \dots \\ & + c_k D^{n-k} f(x) + \dots + c_n f(x) \end{aligned}$$

显然有

$$(4.28) \quad p(D)(k_1 f_1(x) + k_2 f_2(x)) = k_1 p(D)f_1(x) + k_2 p(D)f_2(x)$$

所以 $p(D)$ 乃是一个线性算子。再者，设 $p(D)$ 和 $q(D)$ 是两个这样的线性算子， $p(D) \cdot q(D)$ 是将它们用 D 的多项式乘法所得的算子，则容易验证

$$(4.29) \quad (p(D) \cdot q(D))f(x) = p(D)(q(D)f(x))$$

对于任给 $f(x) \in \mathcal{V}$ 皆成立，亦即线性算子 $p(D)$ 和 $q(D)$ 的组合就是 $p(D) \cdot q(D)$ 。其实，(4.29)-式所表达者——算子组合和多项式乘法之间的

「同构」关系——乃是引进算子符号的原由和优点所在。它使得我们可以有效地运用多项式的代数知识来达成微分方程

$$(4.30) \quad p(D)y = 0$$

的求解。

采用研讨线性空间和线性算子的通用符号，我们将用 $\text{Ker}(p(D))$ 表示 \mathcal{V} 中所有满足

$$(4.31) \quad p(D)f(x) \equiv 0, \quad f(x) \in \mathcal{V}$$

所成的子集，亦即微分方程 (4.30) 的解函数的集合。由 $p(D)$ 的线性（亦即 (4.28)-式）可见 $\text{Ker}(p(D))$ 乃是 \mathcal{V} 中的一个线性子空间，亦即

$$(4.32) \quad f_1(x), f_2(x) \in \text{Ker}(p(D)) \Rightarrow k_1 f_1(x) + k_2 f_2(x) \in \text{Ker}(p(D))$$

4.3.2 $p(D) = (D - \lambda)^k$ 的情形， $\lambda \in \mathbb{C}$

让我们先来研讨两个简单的特殊情形，即其一是 $\lambda = 0$ 者，其二是 $k = 1$ 者。

【例一】： $f(x) \in \text{Ker}(D^k)$ 的充要条件是： $f(x)$ 乃是一个次数至多为 $k - 1$ 的多项式函数。

证明：设 $f(x)$ 是一个次数至多为 $k - 1$ 的多项式函数，则显然有 $D^k f(x) = f^{(k)}(x) \equiv 0$ 。反之，设 $f^{(k)}(x) \equiv 0$ ，则泰勒公式中的 $R_k \equiv 0$ ，亦即

$$(4.33) \quad f(x) = f(0) + f'(0)x + \dots + \frac{f^{(k-1)}(0)}{(k-1)!}x^{k-1} + R_k$$

而 $R_k \equiv 0$ ，所以 $f(x)$ 乃是一个次数至多为 $k - 1$ 的多项式函数。□

【例二】：设 $\lambda = k + i\omega \neq 0$ 。则 $\text{Ker}(D - \lambda)$ 乃是由 $e^{kx}(\cos \omega x + i \sin \omega x)$ 的常数倍所组成的一维子空间。

证明：设 $f(x) \in \text{Ker}(D - \lambda)$ 是一个满足

$$(4.34) \quad Df(x) = \lambda f(x)$$

的非零函数。令

$$(4.35) \quad g(x) = e^{-kx}(\cos \omega x - i \sin \omega x)f(x)$$

则有

$$(4.36) \quad \begin{aligned} Dg(x) &= (-\lambda) \cdot e^{-kx}(\cos \omega x - i \sin \omega x)f(x) \\ &\quad + e^{-kx}(\cos \omega x - i \sin \omega x) \cdot \lambda f(x) \equiv 0 \end{aligned}$$

亦即 $g(x)$ 乃是一个复值常数。所以 $f(x)$ 乃是 $e^{kx}(\cos \omega x + i \sin \omega x)$ 的复常数倍是也。 \square

【定理 4.3】： $f(x) \in \text{Ker}(D - \lambda)^k$ 的充要条件是

$$(4.37) \quad f(x) = e^{kx}(\cos \omega x + i \sin \omega x)(c_0 + c_1x + \dots + c_{k-1}x^{k-1})$$

其中 c_i 是复常数。

证明：一如[例二]令

$$(4.35) \quad g(x) = e^{-kx}(\cos \omega x - i \sin \omega x)f(x)$$

则有

$$(4.38) \quad \begin{aligned} Dg(x) &= (-\lambda)e^{-kx}(\cos \omega x - i \sin \omega x)f(x) \\ &\quad + e^{-kx}(\cos \omega x - i \sin \omega x)Df(x) \\ &= e^{-kx}(\cos \omega x - i \sin \omega x)(D - \lambda)f(x) \end{aligned}$$

如此逐步计算，即得：

$$(4.39) \quad D^k g(x) = e^{-kx}(\cos \omega x - i \sin \omega x)(D - \lambda)^k f(x) \equiv 0$$

由[例一]即得 $g(x)$ 乃是一个次数至多为 $(k-1)$ 的多项式函数，亦即

$$(4.40) \quad f(x) = e^{kx}(\cos \omega x + i \sin \omega x)g(x)$$

乃是如 (4.37)-式所表达者。 \square

4.3.3 $p(D)$ 是一般的情形

在 $p(D)$ 是一般的情形, 由代数基本定理可见 $p(D)$ 可以表达成下述形式, 即

$$(4.41) \quad p(D) = \prod_{j=1}^{\ell} (D - \lambda_j)^{\alpha_j}$$

其中 $\{\lambda_j = k_j + i\omega_j, 1 \leq j \leq \ell\}$ 是相异的。显然有

$$(4.42) \quad \text{Ker}(p(D)) \supseteq \text{Ker}(D - \lambda_j)^{\alpha_j}, \quad 1 \leq j \leq \ell$$

所以当然也包含

$$(4.43) \quad \sum_{j=1}^{\ell} e^{k_j x} (\cos \omega_j x + i \sin \omega_j x) g_j(x)$$

其中 $g_j(x)$ 是任给次数至多是 $(\alpha_j - 1)$ 的多项式函数。下述[定理 4.4]证明任何 $\text{Ker}(p(D))$ 中的函数都可以唯一地表达成 (4.43)-式所给者。

【定理 4.4】：设 $f(x) \in \text{Ker}(p(D))$, $p(D)$ 如 (4.41)-式所给者, 则 $f(x)$ 可以唯一地表达成 (4.43)-式所给者。

用下述引理, 则可把上述[定理 4.4]归于[定理 4.3]直接推导之。

【引理】：设 $d(X)$ 和 $m(X)$ 分别是 $p(X)$ 和 $q(X)$ 的最高公因式和最低公倍式, 则

$$(4.44) \quad \text{Ker}(d(D)) = \text{Ker}(p(D)) \cap \text{Ker}(q(D))$$

$$(4.45) \quad \text{Ker}(m(D)) = \text{Ker}(p(D)) + \text{Ker}(q(D))$$

证明：先证 (4.44)-式：显然有

$$(4.46) \quad \text{Ker}(d(D)) \subseteq \text{Ker}(p(D)) \cap \text{Ker}(q(D))$$

所以我们只需再证 (4.46)-式的逆向包含式, 亦即要由 $p(D)f(x) = 0$ 和 $q(D)f(x) = 0$ 去推导 $d(D)f(x) = 0$ 。由[定理 1.5]得知存在多项式的表达式

$$(4.47) \quad d(X) = A(X) \cdot p(X) + B(X) \cdot q(X)$$

所以

$$(4.48) \quad d(D)f(x) = A(D)p(D)f(x) + B(D)q(D)f(x) = 0 + 0 = 0$$

再证 (4.45)-式：显然有

$$(4.49) \quad \text{Ker}(m(D)) \supseteq \text{Ker}(p(D)) + \text{Ker}(q(D))$$

所以我们只需再证 (4.49)-式的逆向包含式，亦即要由 $m(D)f(x) = 0$ 去求得 $f(x) = f_1(x) + f_2(x)$ 而且 $p(D)f_1(x) = 0$, $q(D)f_2(x) = 0$ 的分解。由所设

$$(4.50) \quad p(X) = d(X) \cdot p_1(X), \quad q(X) = d(X) \cdot q_1(X)$$

其中 $p_1(X)$ 和 $q_1(X)$ 互素。再者

$$(4.51) \quad m(X) = q(X)p_1(X) = p(X)q_1(X)$$

在此，再用[定理 1.5]和所设 $p_1(X)$ 与 $q_1(X)$ 的最高公因式等于 1，即有

$$(4.52) \quad 1 = A_1(X)p_1(X) + B_1(X)q_1(X)$$

所以

$$(4.53) \quad \begin{aligned} f(x) &= 1 \cdot f(x) = A_1(D)p_1(D)f(x) + B_1(D)q_1(D)f(x) \\ &= f_2(x) + f_1(x) \end{aligned}$$

易见

$$(4.54) \quad \begin{aligned} p(D)f_1(x) &= B_1(D)p(D)q_1(D)f(x) = B_1(D)m(D)f(x) = 0 \\ q(D)f_2(x) &= A_1(D)q(D)p_1(D)f(x) = A_1(D)m(D)f(x) = 0 \end{aligned}$$

这也就证明了 $\text{Ker}(m(D))$ 中的任给 $f(x)$ 都属于 $\text{Ker}(p(D)) + \text{Ker}(q(D))$ 。□

【推论】：当 $p(D)$ 如 (4.41)-式所给者，则有

$$(4.55) \quad \text{Ker}(p(D)) = \bigoplus_{j=1}^{\ell} \text{Ker}(D - \lambda_j)^{\alpha_j}$$

将上述推论和[定理 4.3]结合，即得[定理 4.4]。

4.3.4 $p(D)y = g(x)$ 的解法

若 $f(x)$ 和 $f_1(x)$ 都是 $p(D)y = g(x)$ 的解, 则易见

$$p(D)(f(x) - f_1(x)) = p(D)f(x) - p(D)f_1(x) = g(x) - g(x) \equiv 0$$

亦即 $f_0(x) = f(x) - f_1(x)$ 是 $p(D)y = 0$ 的解。所以 $p(D)y = g(x)$ 的通解可以写成下述模样

$$f(x) = f_1(x) + f_0(x), \quad f_0(x) \in \text{Ker } p(D)$$

亦即我们只需要找到一个特殊解 $f_1(x)$ 就可以写下 $p(D)y = g(x)$ 的通解。如果我们可以从某些途径 (如解题的实际经验) 知道这个特殊解的样子时, 当然就可以用待定系数法求之; 但一般来说, 我们所能辨认的 $g(x)$ 的样子还是十分有限, 大多数是 $P(x)e^{kx}$ 的组合。下面我们将会讨论一个有系统的方法来寻找这些特殊解, 且以一个简单例子来具体说明其做法。

【例三】: 求解 $(D^2 + 1)f(x) = \tan x, 0 < x < \frac{\pi}{2}$ 。

解: 易见 $(D^2 + 1)f_0(x) = 0$ 的通解为

$$f_0(x) = c_1 \cos x + c_2 \sin x, \quad c_1, c_2 \text{ 为常数}$$

而寻找 $f_1(x)$ 的方法是把上述 $f_0(x)$ 的通解中那些常数 c_1 和 c_2 当作为待定函数, 并设法寻找适当的 $c_1(x)$ 和 $c_2(x)$ 来构造出 $g(x)$ 。令

$$f_1(x) = c_1(x) \cos x + c_2(x) \sin x$$

我们需要有 2 个条件式才可以唯一地确定 $c_1(x)$ 和 $c_2(x)$ 。当然, $f_1(x)$ 首先要满足 $(D^2 + 1)f(x) = g(x)$, 所以我们还需要额外一个条件式。这个条件式自然要选得愈便于计算愈好, 若把它选得适当使我们不用考虑 $c_1(x), c_2(x)$ 的高阶微分者当然就是最好的! 先对 $f_1(x)$ 微分:

$$f_1'(x) = c_1'(x) \cos x - c_1(x) \sin x + c_2'(x) \sin x + c_2(x) \cos x$$

由此可见, 我们应该令 $c_1'(x) \cos x + c_2'(x) \sin x = 0$, 使得 $f_1'(x) = -c_1(x) \sin x + c_2(x) \cos x$ 。继续对 $f_1'(x)$ 微分, 得

$$\begin{aligned} f_1''(x) &= -c_1'(x) \sin x - c_1(x) \cos x + c_2'(x) \cos x - c_2(x) \sin x \\ &= -c_1' \sin x + c_2' \cos x - f_1(x) \end{aligned}$$

$$\Rightarrow f_1''(x) + f_1(x) = -c_1'(x) \sin x - c_2'(x) \cos x = \tan x$$

所以， $c_1(x)$ 和 $c_2(x)$ 是会满下述联立一阶微分方程：

$$\begin{cases} c_1'(x) \cos x + c_2'(x) \sin x = 0 \\ -c_1'(x) \sin x + c_2'(x) \cos x = \tan x \end{cases}$$

易见其解为 $c_1'(x) = -\sin x \tan x = \cos x - \sec x$, $c_2'(x) = \sin x$ ，亦即 $c_1(x) = \sin x - \ln(\sec x + \tan x)$, $c_2(x) = -\cos x$ 。所以，我们找到一个特殊解

$$f_1(x) = -\cos x \ln(\sec x + \tan x)$$

而 $(D^2 + 1)f(x) = \tan x$ 的通解为

$$f(x) = c_1 \cos x + c_2 \sin x - \cos x \ln(\sec x + \tan x)$$

易见上述方法可推广到处理一般的常系数常微分方程 $p(D)y = g(x)$ ，其解法我们称之为「参数变分法」(method of variation of parameters)。

第五章

欧氏几何、球面几何和非欧几何的统一理论

自古到今，人类理性文明世代相承、精益求精地对于我们和万物万象共存于其中的「空间」的本质的研究，总称之为几何学 (Geometry)。由此可见，几何学也就是空间本质的认识论。因此，几何学理所当然地是各种各样自然科学的基础所在，自然而然地是整个自然科学的先行者和科学方法论的发祥地与典范。在人类理性文明的发展史上，几何学在古希腊时代获得长足的进展，其辉煌的成就有赖于 Euclid 所著的几何原理 (Elements) 集其大成，流传後世，所以现在通常把它称之为欧氏几何学 (Euclidean Geometry)。

Euclid 的 Elements 对于当代研讨所得的几何认知，采用「公理化」(axiomatic) 的论述方式。[其实，公理化的论述方式并非 Euclid 所首创，Euclid 的体系大概可以追溯到 Eudoxus 的论著；可惜 Eudoxus 的原著早已失传，我们已无缘得见其原貌了。] 概括地说，公理化的论述方式乃是选定一组简朴而且基本的性质作为起点，有系统地逐步推演所要研讨的事物（例如几何学所要研讨的空间）的其他各种各样有用、常用的性质。当年把所选用的基础性性质（起点）叫做公理，把那些由公理演绎推导而得者叫做定理。对于欧氏几何学的公理体系和公理化治学的要点，在 1899 年 Hilbert 所著的几何基础论 (Grundlagen der Geometrie) 中有详尽的讨论，可供有意深究者参阅。大体上来说，公理化的方法论也就是用逻辑推理 (logical deduction) 来对于某一种知识体系构造一

种性质层面上的以简御繁。

长话短说，我们所在的空间的各种各样本质之中，最为简朴基本者可以归为下述四种，即

1. 点、线、面的连结与分隔（基本结构）；
2. 连续性：直线的连续不断 \longleftrightarrow 实数系的连续性；
3. 对称性：空间对于平面的反射对称性 \longleftrightarrow 叠合公理 (S.A.S.) 等；
4. 平直性：平行公理 \longleftrightarrow 三角形内角和恒等于一平角。

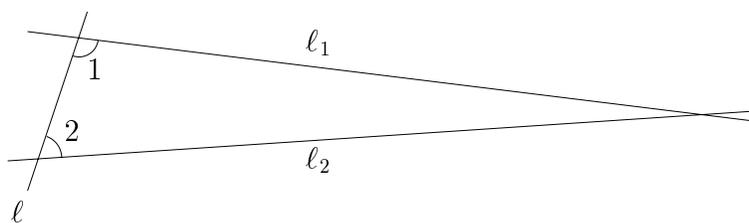
从几何直观来看，前三者的直观内涵都十分明确、自然，唯独平直性的直观内涵令人难以理解其精要何在？此事一直到十九世纪初才真相大白，这也就是为什麼从古希腊到十八世纪末，二千多年来很多几何学家都对于平行公理感到不自在，因而都想（用其他公理来）证之而後快的原由。一直到十九世纪初 (Bolyai, Lobachevsky, Gauss) 非欧几何学的发现才一方面知道平行公理是不能用其他公理来加以论证的，而另一方面也搞清楚平行公理的直观内涵乃是空间的「本质性曲率」等于零（亦即平直性）。

本章将以微积分为工具，对于欧氏、球面与非欧这三种几何，建立其求同存异的统一理论，作为本书所研讨的基础分析学的一个重要应用之范例。

5.1 非欧几何的发现过程及其历史意义

在欧氏所著的 Elements 中所论述的几何学公理体系之中，其第五公设也就是我们通常称之为平行公理者，即平面上任给两条直线 l_1, l_2 ，若其为另一直线 l 交截的同傍内角之和少于平角，则 l_1, l_2 必定相交于 l 的该侧，亦即如 [图 5-1] 所示：

$$\angle 1 + \angle 2 < \text{平角} \quad \Rightarrow \quad l_1, l_2 \text{ 相交于 } \angle 1, \angle 2 \text{ 之同侧}$$



[图 5-1]

若我们对于整个欧氏几何学的逻辑体系再作一次详细的分析，就会发现上述第五公设在所有定量几何的基本定理（如三角形内角和定理、面积公式、相似三角形定理、毕氏定理等等）的论证之中都扮演著基本重要的角色。但是这个第五公设所反映的空间本质的直观内含究竟是什麼呢？却又是一直困扰著古希腊到十九世纪初非欧几何学的发现之前世世代代的几何学家的老大疑问。他们对于上述几何基础论上的疑难的研究大致上可以分成三个阶段：

【第一阶段】：试著改用其他几何命题来替代第五公设，例如当然可以用三角形内角和恒等于平角来替换平行公理。但是这种尝试并不成功，因为他们用来替代的几何命题的直观内含依然难以捉摸。

【第二阶段】：试著用欧氏公理体系中其他的公理去证明第五公设。假若真的能够达成这样一个证明的话，则第五公设就可以根本从公理体系中省却！岂不使得整个几何公理体系更加简洁明了？但是此事两千多年来，世世代代的几何学家们，屡试皆败，事与愿违。

【第三阶段】：从世世代代试证第五公设的教训中，促使十八、十九世纪的某些几何学家们去探索另一种几何体系的可能性，亦即是否存在著另一种满足唯独在第五公设上和欧氏体系相异的几何体系？假如存在的话，一方面使我们充分理解为什麼试证平行公理屡试必败！而另一方面也可从欧氏几何和这种大同小异的「非欧几何」的本质的异同比较之中，真正认识到第五公设的直观内涵究竟是什麼！在这方面的探讨终于在十九世纪初获得了突破性的成功，这就是当年 Bolyai, Lobachevsky 和 Gauss 在非欧几何学的发现与建立上的重大成就。

长话短说，如今我们已经知道这种和欧氏几何大同小异（亦即仅仅在三角形内角和上与欧氏几何相异而在其他各方面完全和欧氏几何相同）的非欧几何体系是存在的。现在我们来回顾这一段历史性的进

展，似乎宜于下述两点著手研讨之，即

- (一) 导致非欧几何学之发现的思路和重要的突破点是那些？
- (二) 欧氏几何、球面几何和非欧几何乃是三种大同小异的几何体系，其大同在于它们都具有同样的连续性和对称性，而其小异则在于它们的三角形内角和分别是等于、大于和小于一个平角。是否可以求大同存小异，建立起一种统一这三种几何学的理论呢？

这也就是我们要接著详加研讨的中心课题。

5.2 发现非欧几何学的思路与突破点

骤看起来，非欧几何学的发现当然是一种存在性问题，其所要探索者乃是某种数理模型（亦即体系）是否存在的问题。大体上来说，在各种各样的数理模型的研讨之中，存在性问题和唯一性问题乃是两个既相对又相关的基本问题。在这里有一个耐人寻思、富有哲理的想法，那就是存在性的探索往往要先从唯一性的研讨著手。其实，後者经常是达成前者的捷径。这种想法骤听起来难免感到太玄了些，但是稍加分析，就不难初步看出它的合理性。因为要凭空构造出一种具有各种各样基本性质的数理模型有如大海捞针，往往是徒劳无成的。而唯一性方面的研讨正好就是逐步缩小范畴，也就自然而然地把所要构造的事物逐步加以明确。

把上述想法运用到「非欧几何」的探索上，就是要深入研讨那些不依赖于第五公设就能证明的几何定理。例如我们在《基础几何学之一》中所证明的下述定理就是其中一个重要的突破点，即

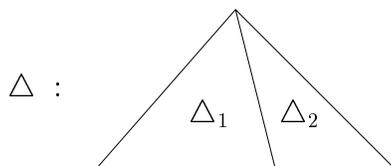
【定理 5.1】：三角形的内角和不大于一个平角。

【定理 5.2】：若有一个三角形的内角和等于一个平角，则所有三角形的内角和皆恒等于一个平角。

由此可见，我们所要探索者，乃是那种三角形内角和恒小于一个平角的几何体系。很自然我们可以把平角和一个三角形的内角和的差额定义为该三角形的「角亏」(defect)；当我们把一个三角形分割成两个三角形

，如 [图 5-2] 所示，易见它的角亏乃是两个子三角形的角亏之和，亦即

$$\delta(\Delta) = \delta(\Delta_1) + \delta(\Delta_2)$$



[图 5-2]

显然，三角形的面积也有同样的分割可加性，即

$$A(\Delta) = A(\Delta_1) + A(\Delta_2)$$

再者，角亏和面积当然都是叠合不变量，而且它们也具有同样的「包含不等式」，即

$$\Delta \supseteq \Delta' \Rightarrow \begin{cases} A(\Delta) \geq A(\Delta') \\ \delta(\Delta) \geq \delta(\Delta') \end{cases}$$

用上述三点共同的性质就可证得

【定理 5.3】：在一个给定的非欧几何体系中，任给三角形的角亏和面积之间的比值乃是一个本质性的常数，亦即存在一个常数 K 使得

$$\delta(\Delta) = KA(\Delta)$$

对于所有三角形恒成立。

三角形乃是最为简朴的二维几何事物 (2-dimensional geometric object)，三角形的几何性质则自然而然整个几何学的基础所在。在欧氏几何体系中，三角形的内角和恒等于一个平角和 S.A.S. 叠合条件分别反映着空间的两大基本性质，亦即平直性和对称性。在非欧几何体系中，上述[定理 5.3]乃是前者的变化，而 S.A.S. 的三角形叠合公理则完全同样成立，这也就是非欧几何体系和欧氏几何体系具有完全同样的对称性的另一种表达方式。从定量几何的观点来看，S.A.S. 叠合条件乃是说三角形的六个基本量（三边边长和三角角度）之间，其他三者乃是选定的两边一夹角的函数，而欧氏三角学 (Euclidean Trigonometry) 中的正弦定律和余弦定律则是把任意三角形的角边之间的函数关系加以

具体明确的定量几何基本定理，它们又是进而用解析法研讨几何的基础工具。可以说欧氏三角学乃是欧氏三角形的解析几何，而它又是整个欧氏解析几何学的基石所在。由此可见对于所要探索的非欧几何体系的唯一性研讨的重大突破点，理当在于其中任意三角形的三边三角之间的函数关系的确定，亦即我们所要进而研讨的要点在于其中任意三角形的「正弦、余弦定律」。

首先，为了便于讨论，我们可以适当地选定长度的单位使得[定理 5.3]中的常数 $K = 1$ 。就好像在研讨球面几何时，半径为 R 的球面上，三角形的角盈和面积之间的关系式乃是

$$\varepsilon(\Delta) = \frac{1}{R^2}A(\Delta)$$

假如我们改用其半径为长度单位，则上述关系式就可以简化为 $\varepsilon(\Delta) = A(\Delta)$ 。

长话短说，Bolyai, Lobachevsky 和 Gauss 各别在十九世纪初叶所达成的重大突破也就是确立了非欧空间中的正弦、余弦定律，即

$$\text{非欧正弦定律} : \frac{\sin A}{\sinh a} = \frac{\sin B}{\sinh b} = \frac{\sin C}{\sinh c}$$

$$\text{非欧余弦定律} : \begin{cases} \sinh a \sinh b \cos C = \cosh a \cosh b - \cosh c \\ \sinh b \sinh c \cos A = \cosh b \cosh c - \cosh a \\ \sinh c \sinh a \cos B = \cosh c \cosh a - \cosh b \end{cases}$$

[注]：双曲正弦、余弦函数 $\sinh t$ 和 $\cosh t$ 的定义式是

$$\sinh t = \frac{1}{2}(e^t - e^{-t}), \quad \cosh t = \frac{1}{2}(e^t + e^{-t})$$

其命名之原由是 $x = \cosh t, y = \sinh t$ 所表达之曲线乃是双曲线 $x^2 - y^2 = 1$ 之一支。再者，由 §3.4 节的 Euler 公式可见：

$$\sin(it) = i \sinh t, \quad \cos(it) = \cosh t$$

5.3 欧氏、球面与非欧三角定律的统一理论

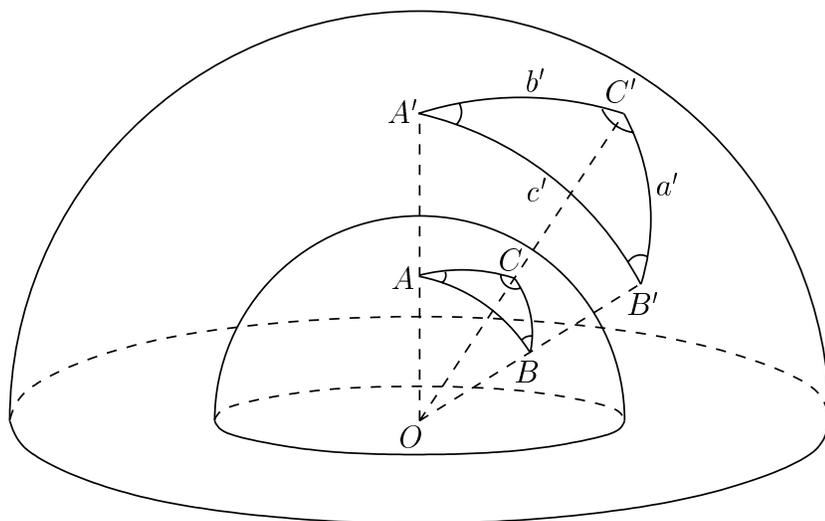
欧氏、球面与非欧乃是三种具有同样的连续性和对称性的几何体系，例如其中的三角形都满足 S.A.S. 叠合公理，而在三种几何学中各别的正弦余弦定律则是三种解析几何的基础所在。假如我们把三种几何的正弦、余弦定律加以比较分析，就会看到三者的正弦定律在形式上十分相似，即

$$\text{欧氏正弦定律: } \frac{\sin A}{a} = \frac{\sin B}{b} = \frac{\sin C}{c}$$

$$\text{球面正弦定律: } \frac{\sin A}{\sin a} = \frac{\sin B}{\sin b} = \frac{\sin C}{\sin c}$$

$$\text{非欧正弦定律: } \frac{\sin A}{\sinh a} = \frac{\sin B}{\sinh b} = \frac{\sin C}{\sinh c}$$

但是欧氏余弦定律和其他两种在形式上则相差甚远，倒是球面和非欧余弦定律则十分相似。



[图 5-3]

如 [图 5-3] 所示，把单位球放大 R -倍，即得半径为 R 的球面。令 $\widehat{\triangle}ABC$ 和 $\widehat{\triangle}A'B'C'$ 分别是相应的球面三角形，则有 $A' = A$, $B' = B$, $C' = C$, $a' = Ra$, $b' = Rb$, $c' = Rc$ 。由此可见半径为 R 的球面上的正、

余弦定律分别是

$$(5.1) \quad \frac{\sin A'}{\sin \frac{a'}{R}} = \frac{\sin B'}{\sin \frac{b'}{R}} = \frac{\sin C'}{\sin \frac{c'}{R}}$$

$$\sin \frac{a'}{R} \sin \frac{b'}{R} \cos C' = \cos \frac{c'}{R} - \cos \frac{a'}{R} \cos \frac{b'}{R} \quad \text{等等}$$

我们可以把上述半径为 R 的球面上的正、余弦定律略作变化，即改写为

$$(5.1)_R \quad \frac{\sin A'}{R \sin \frac{a'}{R}} = \frac{\sin B'}{R \sin \frac{b'}{R}} = \frac{\sin C'}{R \sin \frac{c'}{R}}$$

$$R^2 \sin \frac{a'}{R} \sin \frac{b'}{R} \cos C' = R^2 \left(\cos \frac{c'}{R} - \cos \frac{a'}{R} \cos \frac{b'}{R} \right) \quad \text{等等}$$

从几何的观点，很自然地可以把欧氏平面想成是球面无限放大的局部形式。因此，欧氏正、余弦定律应该可以表达成 R -球面上的正、余弦定律在 $R \rightarrow \infty$ 的极限形式。用上述改写后的形式就很容易看到这一点，即有

$$(5.2) \quad \lim_{R \rightarrow \infty} R \sin \frac{a'}{R} = a' \lim_{\frac{a'}{R} \rightarrow 0} \frac{\sin \frac{a'}{R}}{\frac{a'}{R}} = a' \cdot 1 = a',$$

$$\lim_{R \rightarrow \infty} R^2 \left(1 - \cos \frac{c'}{R} \right) = c'^2 \lim_{\frac{c'}{R} \rightarrow 0} \frac{1 - \cos \frac{c'}{R}}{\left(\frac{c'}{R}\right)^2} = \frac{1}{2} c'^2,$$

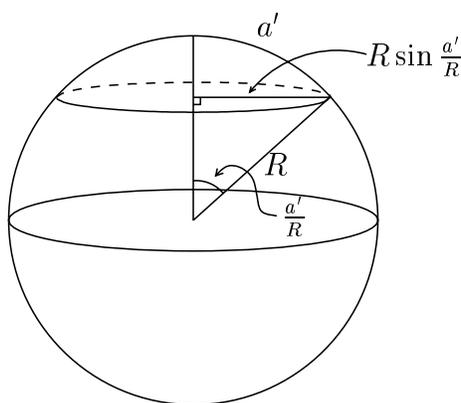
$$\lim_{R \rightarrow \infty} R^2 \left(1 - \cos \frac{a'}{R} \cos \frac{b'}{R} \right) = \frac{1}{2} (a'^2 + b'^2).$$

所以 (5.1)_R-式在 $R \rightarrow \infty$ 的极限形式就是

$$(5.1)_\infty \quad \frac{\sin A'}{a'} = \frac{\sin B'}{b'} = \frac{\sin C'}{c'}$$

$$a'b' \cos C' = \frac{1}{2} (a'^2 + b'^2 - c'^2), \quad \text{等等}$$

这也就是欧氏正、余弦定律。其实，在 R -球面正弦定律中所出现的函数 $R \sin \frac{a'}{R}$ 还有其自然的几何意义，即如 [图 5-4] 所示



[图 5-4]

R -球面上一个半径为 a' 的圆周周长等于 $2\pi R \sin \frac{a'}{R}$ 。由此可见， R -球面的正弦定律还可以进一步改写成下述富有几何意义的形式，即

$$(5.3) \quad \frac{\sin A}{\odot a} = \frac{\sin B}{\odot b} = \frac{\sin C}{\odot c}$$

其中 $\odot r$ 表示半径为 r 的圆周周长。

相信当年（1825 年前后）Bolyai 很可能就是从类似的分析而想到在非欧空间的正弦定律也应该是上述形式，而且他当时业已研讨得非欧空间中半径为 r 的圆周周长公式乃是

$$(5.4) \quad \odot r = \frac{2\pi}{\sqrt{K}} \sinh \sqrt{K} r$$

其中本质性常数 K 就是 $\delta(\Delta) : A(\Delta)$ 的那个比值常数。

[注]：角亏 $\delta(\Delta)$ 乃是负的角盈 $\varepsilon(\Delta)$ 。在 R -球面上，三角形的面积公式可写成

$$\varepsilon(\Delta) = \frac{1}{R^2} A(\Delta)$$

亦即

$$\varepsilon(\Delta) = K A(\Delta), \quad K = \frac{1}{R^2}$$

而 R -球面中的 $\odot r$ 函数则可以改写成

$$\odot r = \frac{2\pi}{\sqrt{K}} \sin \sqrt{K} r$$

再者，在非欧空间的面积公式则是

$$\varepsilon(\Delta) = -\delta(\Delta) = -KA(\Delta)$$

而上述圆周周长公式 (5.4) 则可以改写成

$$(5.4') \quad \odot r = \frac{2\pi}{\sqrt{-K}} \sin \sqrt{-K} r \quad \left(= \frac{2\pi}{i\sqrt{K}} (i \sinh \sqrt{K} r) \right)$$

总结上面的分析所得，对于欧氏、球面和非欧几何的本质有下述统一全局的几个要点：

(一) 三角形内角和与面积之间的关联：

$$\varepsilon(\Delta) = A + B + C - \pi = KA(\Delta)$$

其中本质性常数 K 在欧氏空间为零，球面空间为正而非欧空间为负。

(二) 圆周周长公式：

$$\odot r = \frac{2\pi}{\sqrt{K}} \sin \sqrt{K} r$$

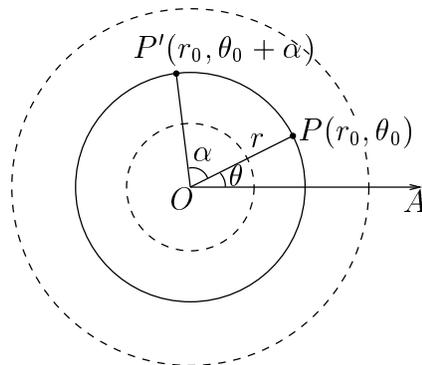
(三) 正弦定律：

$$\frac{\sin A}{\odot a} = \frac{\sin B}{\odot b} = \frac{\sin C}{\odot c} \quad (\text{Bolyai})$$

从几何结构 (geometric structures) 和空间模型 (space forms) 的观点来看，欧氏、球面和非欧这三种空间模型所共有的基本结构就是连结于各点之间的最短通路（亦即欧氏、非欧空间的直线段和球面的大圆圆弧）而它们所共有的基本性质则是它们所具有的反射对称性（亦即 S.A.S. 三角形叠合公理）。对于这样的空间，其中最为简单基本的几何事物 (geometric objects) 首推直线段、三角形和圆，而最为基本重要的几何量 (geometric quantities) 则是直线段的长度，三角形的边长、角度和面积。再者，上述三点当然就是三种几何学中的基本定理的统一形式！由此可见，对于这三种具有最为丰富的对称性的几何的统一理论当然就要从上述三点的统一论证著手研讨之。

首先，让我们对于欧氏和球面的正弦定律的几何内涵再来下一番温故知新的功夫。

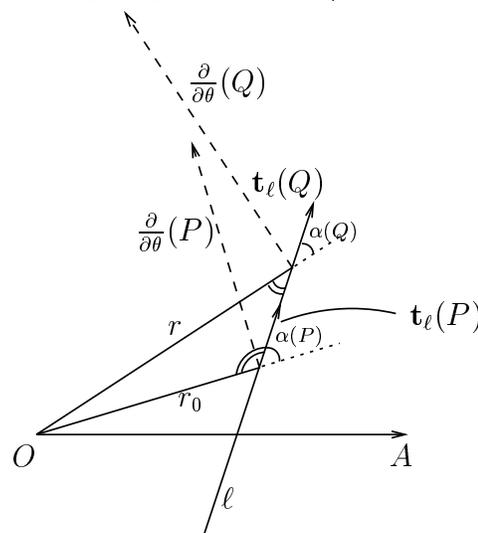
(i) 欧氏平面对于任给一点 O 的旋转对称可以用其上以 O 点为原点的极坐标 (polar coordinate system) 加以简洁描述, 如 [图 5-5] 所示。



[图 5-5]

一个绕 O 点 α 角度的旋转把极坐标为 (r_0, θ_0) 的 P 点映射到极坐标为 $(r_0, \theta_0 + \alpha)$ 的 P' 点。再者, 在绕 O 点的旋转运动之下, $P(r_0, \theta_0)$ 的轨道就是以原点为圆心的圆, 亦即其极坐标方程式就是 $r = r_0$ 。

(ii) 对于上述极坐标, 射线 \overrightarrow{OP} 的方程就是 $\theta = \theta_0$, 而一条过原点 O 的直线的方程式则是 $\theta = \theta_0$ 或 $\theta_0 + \pi$ 。现在让我们再来看一下其他那些不过原点的直线的极坐标方程式应该如何表达? 如 [图 5-6] 所示, 设 ℓ 是一条过 $P(r_0, \theta_0)$ 点但是不过原点的直线。



[图 5-6]

如 [图 5-6] 所示, 设 Q 是 l 上任给一点。令 $\alpha(P), \alpha(Q)$ 分别是 l 和矢径 $\overrightarrow{OP}, \overrightarrow{OQ}$ 之间的夹角。由 $\triangle OPQ$ 的正弦定律, 即得

$$(5.5) \quad \frac{\sin(\pi - \alpha(P))}{r} = \frac{\sin \alpha(Q)}{r_0}$$

亦即

$$(5.5') \quad r \sin \alpha(Q) = r_0 \sin(\pi - \alpha(P)) = r_0 \sin \alpha(P)$$

所以那些不过原点的直线的极坐标方程式乃是

$$(5.6) \quad r \sin \alpha = \text{常数}$$

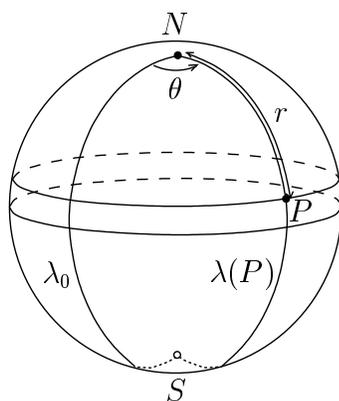
(iii) 设想全平面对于 O 点作单位角速度的旋转运动, 则它在每一点 $P(r_0, \theta_0)$ 的速度向量就是那个圆周 $r = r_0$ 在 P 点的切向量, 其长度等于半径 r_0 (亦即 [图 5-6] 中所示之 $\frac{\partial}{\partial \theta}(P)$, 这是往後在多元分析学中通用的符号, 它所表达者乃是 只有 θ -坐标作单位变率的速度向量)。再者, 沿著直线 l 的单位速率的速度向量当然就是直线 l 在该点的单位长切向量, 分别以 $\frac{\partial}{\partial \theta}(P)$ 和 $\mathbf{t}_l(P)$ 表示上述两种速度向量, 则直线方程 (5.6) 的几何意义就是:

$$(5.7) \quad \begin{aligned} \frac{\partial}{\partial \theta}(P) \cdot \mathbf{t}_l(P) &= \left| \frac{\partial}{\partial \theta}(P) \right| \cdot |\mathbf{t}_l(P)| \cdot \cos\left(\frac{\pi}{2} - \alpha\right) \\ &= r \sin \alpha = \text{常数} \end{aligned}$$

[注]: 若用运动学的观点来说, 上述方程式也就是直线等速运动的角动量守恒定律是也。

接著, 让我们再用同样的观点来分析球面上的情形如下:

(i) 有鉴于球面的均匀性 (homogeneity), 不妨设给定的原点位于北极 N , 则单位球面以北极为原点的极坐标为 [图 5-7] 所示:



[图 5-7]

在球面对于北、南极的旋转运动之下，其轨道乃是「 $r = \text{常数}$ 」的纬圆，而过北、南极的大圆则是经线。

(ii) 同样的，我们以 $\frac{\partial}{\partial \theta}(P)$ 表示上述单位角速度的旋转运动在 P 点的速度向量，则它就是纬圆在 P 点的切向量，其长度为 $\sin r$ 。

另一方面由球面正弦定律也可以得出球面上的大圆的极坐标方程式乃是

$$(5.8) \quad \sin r \sin \alpha = \text{常数}$$

而且上述测地线（亦即大圆）方程的几何意义也同样的

$$(5.9) \quad \frac{\partial}{\partial \theta}(P) \cdot \mathbf{t}_\ell(P) = \text{常数}$$

亦即是球面测地线等速运动的角动量守恒定律。

由上面这样一段「温故」，我们认识到欧氏、球面正弦定律其实就是测地线上等速运动的角动量守恒定律，亦即

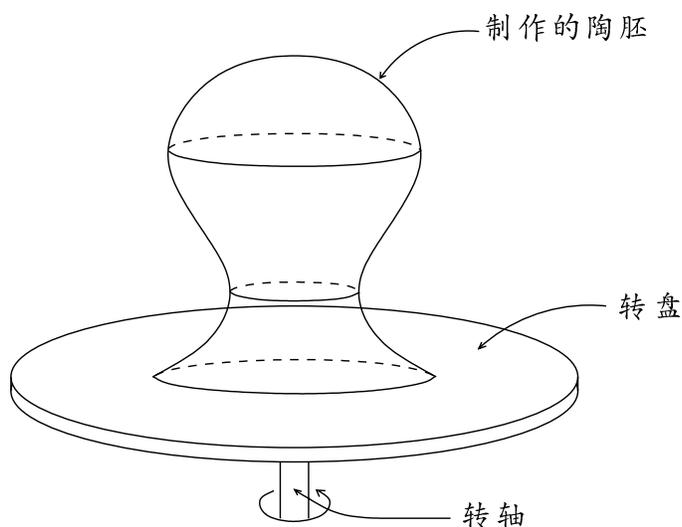
$$\frac{\partial}{\partial \theta}(P) \cdot \mathbf{t}_\ell(P) = \text{常数}$$

由此自然会想到非欧正弦定律也一定就是这种角动量守恒定律，而且三种角动量守恒定律应该具有一种适度的推广加以统一，通常这种更加广泛成立的定律的证明往往会比原先各别论证其特殊情形的证法反而更加简明朴素。这也就是我们在下一节所要研讨的课题。

5.4 旋转面的解析几何

从前面的逐层剖析和一路探讨,使得我们认识到旋转面的解析几何乃是能够把三种几何的三角学统一起来的自然范畴,因此三种几何中的正弦定律应该可以推广成旋转面上的测地线方程的普遍形式而加以统一!

旋转面在人类的历史文明中,在远古的陶器中即已经常出现。在很多博物馆展览的石器时代的出土陶器,经常可以见到造型柔美、旋转对称的陶器。如今以怀古之情,想像远古先民在一个转盘上制作它们的情景;对称性的运用在人类文明中的久远与广泛,令人神往。



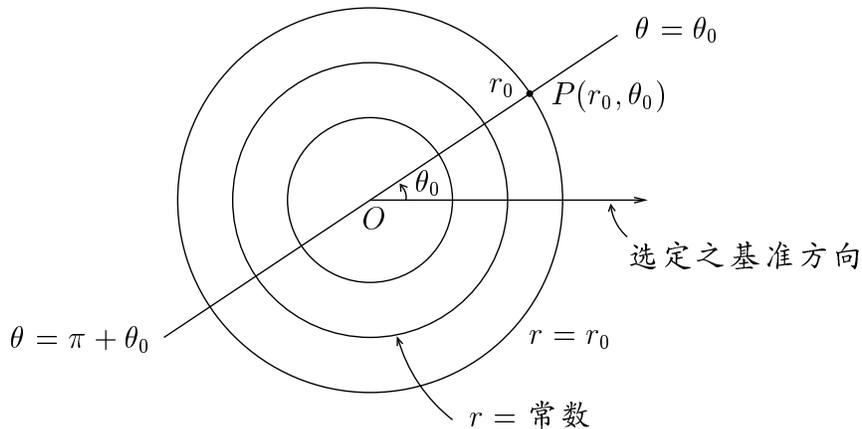
[图 5-8]

旋转面的极坐标表达方式

旋转面乃是那种具有旋转对称性的三维几何模型。例如欧氏平面、球面、[图 5-8]所示的陶胚面和非欧面等等都是旋转面的例子。我们现在要用解析法去研讨一个旋转面的几何,很自然地应该采用极坐标的表达方式。

设 M 是一个对于基点 O 旋转对称的旋转面。对于其上另外一点 P ,令 r 为 O, P 之间的距离,不妨设 O, P 之间的最短通路(亦即测地线)唯一存在(因为只有像在球面上 O, P 互相对顶的这种特殊情形,

其间最短通路才会不唯一)，我们将以符号 \overline{OP} 表达之。再者，令 θ 为 \overline{OP} 和一个选定的基准方向之间的定向夹角，则 (r, θ) 就叫做 P 点的极坐标，如 [图 5-9] 所示。



[图 5-9]

由极坐标的定义，可见「 $r = \text{常数}$ 」所代表者乃是以基点为圆心的同心圆系，它们也就是 M 在旋转对称的转动之下的轨道 (orbits of the rotation of M around O)。再者，「 $\theta = \text{常数}$ 」所表达者，则是其他各点到基点的最短通路，例如欧氏平面中的射线和球面上连结南北极的经线。所以那些通过基点 O 的测地线方程也就是 $\theta = \theta_0$ 或 $\theta = \theta_0 + \pi$ 。

对于一个旋转面 M 上的解析几何，所要研讨的一个首要课题就是那些不过基点的测地线应该如何表达？其研讨之所得即为下述 [定理 5.4]:

【定理 5.4】：设 M 是一个给定的旋转面， (r, θ) 是其上以基点 O 为原点的极坐标系，则 M 上的测地线方程乃是

$$(5.10) \quad \odot r \sin \alpha = \text{常数}$$

其中 $\odot r_0$ 是 $r = r_0$ 的圆周周长， $\alpha(P)$ 是测地线在 P 点和矢径 \overrightarrow{OP} 的夹角。

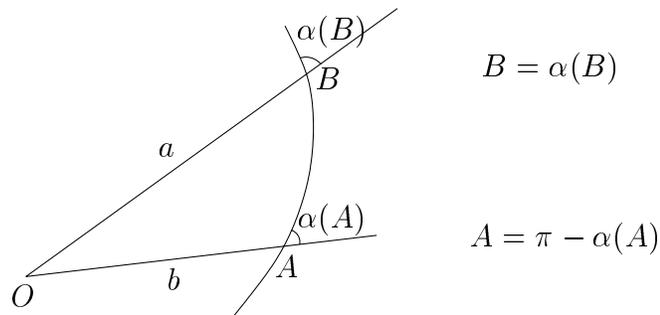
上述 [定理 5.4] 的几何意义有二，其一是它乃是欧氏、球面和非欧这三种几何的正弦定律的共同推广，其二则是旋转面上等速测地运动的角动量守恒定律的普遍成立，即

【推论 1】（广义正弦定律）：设 M 是对于基点 O 旋转对称者， $\triangle OAB$ 是以基点为其顶点之一的三角形，则有

$$(5.11) \quad \frac{\sin A}{\odot a} = \frac{\sin B}{\odot b}$$

其中 a, b 分别是 A, B 的对边边长。

[如 [图 5-10] 所示，它是[定理 5.4]的直接推论。]



$$\odot b \sin A = \odot b \sin \alpha(A) = \odot a \sin \alpha(B) = \odot a \sin B$$

[图 5-10]

【推论 2】（ M 上对于旋转的角动量守恒定律）：

$$(5.12) \quad \frac{\partial}{\partial \theta}(P) \cdot \mathbf{t}_\ell(P) = \frac{1}{2\pi} \odot r \sin \alpha(P) = \text{常数}$$

[定理 5.4]还有一个重要的推论，即：

【推论 3】（广义余弦定律）：令 c 为 $\triangle OAB$ 的第三边边长，亦即 $c = \overline{AB}$ ，则

$$(5.13) \quad c = \int_b^a \frac{f(r)dr}{\pm \sqrt{f(r)^2 - (f(b) \sin A)^2}}$$

其中 $f(r) = \frac{1}{2\pi} \odot r$ ，而上述分母的符号在 $\alpha(P) \leq \frac{\pi}{2}$ 时取正号，在 $\frac{\pi}{2} < \alpha(P) < \pi$ 时取负号。

[由[定理 5.4]推导[推论 3]的证明将在证明[定理 5.4]之后再做。]

弧长元素和弧长的第一变分公式：

要研究一个旋转面中的测地线，当然首先得明确其中一条给定曲线（亦即通路）的长度的定义以及它的解析表达式，然后再进而研究连结两个给定点之间的某种曲线族的长度的变化规律。

从解析几何的观点，一条曲线乃是一个动点的轨迹，它的位置（亦即坐标）可以用「时间参数」 t 的函数加以解析描述，叫做该曲线的参数表达式。例如， $\{x = a \cos t, y = b \sin t; 0 \leq t \leq 2\pi\}$ 乃是椭圆的一种参数表式。再者，一条曲线的长度则可以用积分加以表达。例如欧氏平面上一条以 $\{x = f(t), y = g(t), a \leq t \leq b\}$ 为其笛氏坐标参数表式的曲线，其弧长的解析表式乃是 $\int_a^b [f'(t)^2 + g'(t)^2]^{\frac{1}{2}} dt$ 。上述弧长积分公式也可以用下述弧长元素的微分公式简洁地加以概括，亦即

$$(5.14) \quad ds^2 = dx^2 + dy^2$$

上述微分公式的几何意义就是说曲线的「微段」 $\{(x(t), y(t)), t_i \leq t \leq t_{i+1} = t_i + \Delta t\}$ 的弧长和直线段 $\overline{P(t_i)P(t_{i+1})}$ 的长度几乎相等，亦即至多相差一个 Δt 的高阶微量。若用算式表达，即有

$$(5.15) \quad \Delta_i s^2 \equiv \overline{P(t_i)P(t_i + \Delta t)}^2 = \Delta_i x^2 + \Delta_i y^2 \pmod{\Delta t^3}$$

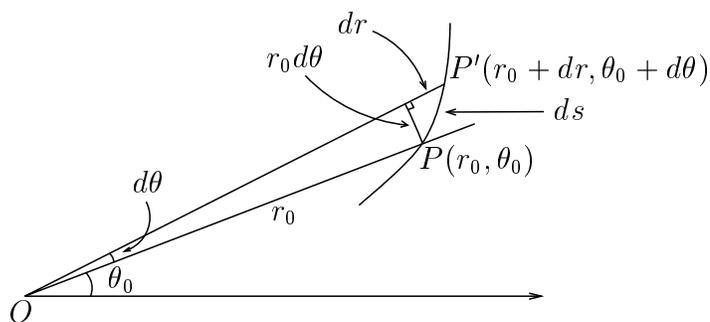
其实这也就是(5.14)-式的实质内涵，而它也就是勾股定理的微分化形式是也！若改用欧氏平面的极坐标来表达，则有

$$(5.16) \quad \begin{aligned} x &= r \cos \theta, \quad y = r \sin \theta \\ dx &= dr \cos \theta - r \sin \theta d\theta, \quad dy = dr \sin \theta + r \cos \theta d\theta \end{aligned}$$

代入(5.14)-式即得

$$(5.14') \quad ds^2 = (dr \cos \theta - r \sin \theta d\theta)^2 + (dr \sin \theta + r \cos \theta d\theta)^2 = dr^2 + r^2 d\theta^2$$

这也就是欧氏平面的极坐标弧长元素的微分式。如[图 5-11]所示，(5.14')其实也是勾股定理的微分化形式。



[图 5-11]

在研讨一般的旋转面的极坐标弧长元素的微分式之前，让我们再来看一看单位球面这个常见而且重要的实例。设 $P(r, \theta)$ 是以北极为基点的球面极坐标，而 $P(x, y, z)$ 则是 P 点的笛氏坐标。如 [图 5-12] 所示，两者之间有下述坐标变换式，即

$$(5.17) \quad x = \sin r \cos \theta, \quad y = \sin r \sin \theta, \quad z = \cos r$$

将 (5.17')-式微分即得

$$(5.18) \quad \begin{cases} dx = \cos r \cos \theta dr - \sin r \sin \theta d\theta \\ dy = \cos r \sin \theta dr + \sin r \cos \theta d\theta \\ dz = -\sin r dr \end{cases}$$

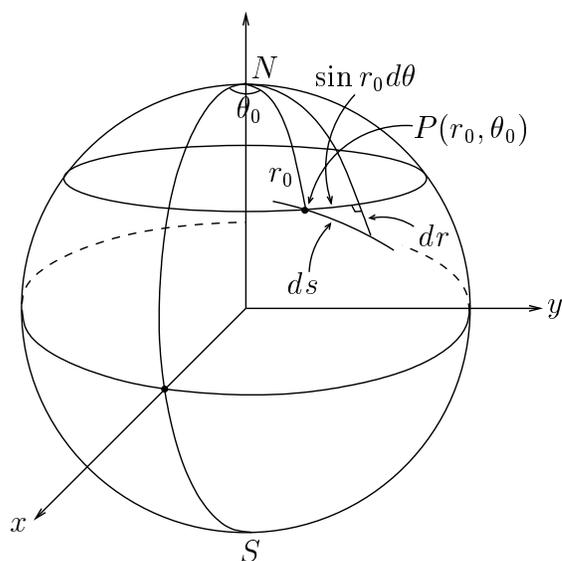
将 (5.18')-式代入三维笛氏坐标所熟知的弧长元素微分式

$$(5.19) \quad ds^2 = dx^2 + dy^2 + dz^2$$

即可计算而得下述球面坐标的弧长元素微分式：

$$(5.20) \quad ds^2 = dr^2 + \sin^2 r d\theta^2$$

如 [图 5-12] 所示，(5.20')-式也是勾股定理的微分形式。



[图 5-12]

从上面对于曲线弧长的实例分析，我们可以看到下述几点基本认识：

(i) 一条参数曲线的弧长本质上乃是它的「细分线段」的弧长之总和，所以它乃是一个积分，亦即弧长元素的积分 ($\int ds$)。

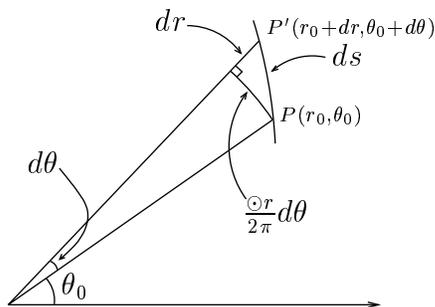
(ii) 弧长元素的几何本质就是那种细分微段的长度的线性（亦即一阶）逼近 (linear approximation)，因为在 $\int ds$ 的积分求和之中，任何比 dt 高阶的微量是完全不影响其积分值者。所以“ ds ”的实质乃是微段弧长的线性（亦即一阶）部分。换言之，也就是微段弧长的线性化 (linearization)。

(iii) 一个曲面 M 在其上一点 P 的切平面 (tangent plane) 就是 M 在 P 点的邻近微片的线性逼近。所以上述切平面上的几何，其实就是 M 在 P 点邻近的几何的局部线性化 (local linearization) 者也；而切平面上的几何乃是欧氏平面的几何，其中的线段长度满足勾股定理，切向量的内积则是表达这种局部线性化所得的切平面几何的基本结构。

将上述对于曲面上曲线弧长概念的剖析所得的三点认识，用到旋转面上就可以得到它对于极坐标的弧长元素微分式，即

$$(5.21) \quad ds^2 = dr^2 + \left(\frac{\odot r}{2\pi}\right)^2 d\theta^2$$

用 [图 5-13] 来说明, 上述公式也就是微小直角三角形的勾股公式微分形式。



[图 5-13]

旋转面及其弧长元素:

总结上面的讨论, 一个旋转面 M 在对于基点 O 的旋转运动下的轨道组成以 O 点为圆心的同心圆系, 它们的圆周周长乃是描述 M 的几何本质函数 $\odot r$ 。令 $f(r) = \frac{1}{2\pi} \odot r$, 则 M 上的弧长元素可以用极坐标以下述微分式表达之, 即

$$(5.21') \quad ds^2 = dr^2 + f(r)^2 d\theta^2$$

再者, M 在 O 点的局部线性逼近乃欧氏平面, 所以 $f(r)$ 在 $r=0$ 的邻近的线性逼近就是 r , 亦即 $f(0) = 0, f'(0) = 1$ 。

反之, 上述弧长元素微分式也完全确定了 M 本身的几何结构, 所以我们可以反过来, 对于任给满足 $f(0) = 0$ 和 $f'(0) = 1$ 的二阶连续可微函数 $f(r)$, 定义一个以 $f(r)$ 为其本质函数的旋转面, $M^2(f)$, 它具有极坐标和 (5.21)-式所给的弧长元素微分式。这也就是我们要研讨旋转面的解析几何, 自然而且恰如其广的范畴。在 $M^2(f)$ 上的任给一条平滑曲线 Γ , 可以用其极坐标参数式 $\{r = \varphi(t), \theta = \psi(t), a \leq t \leq b\}$ 加以描述, 而它的长度则可以用下述积分式表达之, 即

$$(5.22) \quad \begin{aligned} L(\Gamma) &= \int ds = \int [dr^2 + f(r)^2 d\theta^2]^{\frac{1}{2}} \\ &= \int_a^b [\varphi'(t)^2 + f(\varphi(t))^2 \psi'(t)^2]^{\frac{1}{2}} dt \end{aligned}$$

有了上述明确的概念性基础，就可以用微积分的计算去研究 $M^2(f)$ 上的弧长变分，从而证明[定理 5.4]及其[推论 3]。

$M^2(f)$ 上的弧长变分及其第一变分公式：

设 A, B 是 $M^2(f)$ 上的给定两点。连结于 A, B 之间的曲线是非常多种多样的；它们之间变动的自由度是无穷维的。由此可以想到，要用微积分去研究 A, B 两点之间的最短通路是无法像有限元函数的极值问题那样，直截了当地运用微分为零条件式去求解的。为了研究像弧长积分这种定义于无穷维事物（例如连结于 A, B 之间的平滑曲线之总体）的「函数」的极值问题，所发展的学科叫做变分学 (variation theory)，而上述最短通路问题的研讨也就是变分学的发祥地。克服这种由于「变动自由度太大」在极值问题的研讨上所产生的困难的基本想法如下：

尽管连结于 A, B 之间的曲线在总体上有无穷维的变动自由度，但是归根究底，它们都可以想成是各种各样的「单维变动」的组合所构成的总体。由此可以想到，只要能够把「单维变动」的极值条件研究得一清二楚，其实也就足以用来探讨原来的极值问题。这种返璞归真，以简御繁的想法就是变分学的基本想法和出发点。对于我们所要研讨的弧长极值问题的具体做法如下：

曲线之变分和弧长第一变分公式：

$M^2(f)$ 上一条连结 A, B 之间的平滑曲线 Γ 可以用它的一个极坐标参数式表达之，即

$$(5.23) \quad \Gamma = \{\Gamma(t) = (\varphi(t), \psi(t)), a \leq t \leq b\}$$

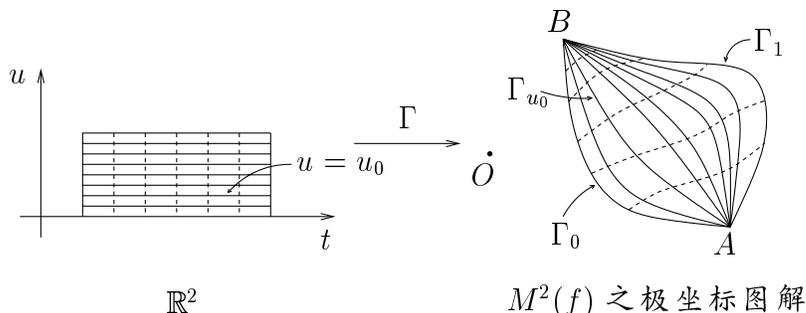
其中 $\varphi(t), \psi(t)$ 是二阶连续可微函数， $A = (\varphi(a), \psi(a)), B = (\varphi(b), \psi(b))$ 。

设 $\varphi(t, u)$ 和 $\psi(t, u)$ 是一对二阶连续可微的函数，其变域为 $a \leq t \leq b, 0 \leq u \leq 1$ 而且 $(\varphi(a, u), \psi(a, u))$ 和 $(\varphi(b, u), \psi(b, u)), 0 \leq u \leq 1$ ，分别是给定点 A, B 的极坐标。则对于任一给定之 u ，

$$(5.24) \quad \Gamma_u = \{\Gamma_u(t) = (\varphi(t, u), \psi(t, u)), a \leq t \leq b\}$$

乃是一条连结于 A, B 之间的平滑曲线，而 $\{\Gamma_u, 0 \leq u \leq 1\}$ 则构成一个「单维变动」的平滑曲线系，其中 $0 \leq u \leq 1$ 就是用来标记该曲线系的

「变动参数」(parameter of variation), 我们把它简称为一个曲线之变分 (a variation of curves)。一个曲线之变分的图像表达如 [图 5-14] 之图解所示。



[图 5-14]

令 $L(u)$ 为 Γ_u 的弧长, 则 $L(u)$ 乃是一个定义于 $0 \leq u \leq 1$ 的可微函数, 而其函数值则可以用弧长积分式表达之, 即

$$(5.25) \quad L(u) = \int_a^b \left[\left(\frac{\partial \varphi}{\partial t} \right)^2 + f(\varphi)^2 \left(\frac{\partial \psi}{\partial t} \right)^2 \right]^{\frac{1}{2}} dt$$

我们所要研讨者乃是 $\left. \frac{dL}{du} \right|_{u=0}$ 的积分表达式。因为 $\varphi(t, u)$ 和 $\psi(t, u)$ 的几何意义分别就是 $\Gamma_u(t)$ 的极坐标 r 和 θ , 所以在下面的计算中我们将根本改用 $r(t, u)$ 和 $\theta(t, u)$ 记之。

[注]: 在这里, 同学们遇到在纯为单元的微积分中尚未接触的偏微分符号如 $\frac{\partial \varphi}{\partial t}$, $\frac{\partial \varphi}{\partial u}$ 等等。例如 $\varphi(t, u)$ 乃是一个二元函数, 其值是随着 t 和 u 的变动而改变的。但是在 u 取定为 u_0 的局限之下, $\varphi(t, u_0)$ 其实就是一个单元函数, 它对于 t 的微分叫做 $\varphi(t, u)$ 对于 t 的偏微分, 以 $\frac{\partial \varphi}{\partial t}$ 记之, 所以它本质上乃是单元函数的微分。在此初见, 其实乃是旧知新用。

$$(5.26) \quad \begin{aligned} \left. \frac{dL}{du} \right|_{u=0} &= \int_a^b \left. \frac{\partial}{\partial u} \left[\left(\frac{\partial r}{\partial t} \right)^2 + f(r)^2 \left(\frac{\partial \theta}{\partial t} \right)^2 \right]^{\frac{1}{2}} \right|_{u=0} dt \\ &= \int_a^b \left[\left(\frac{\partial r}{\partial t} \right)^2 + f^2 \left(\frac{\partial \theta}{\partial t} \right)^2 \right]^{-\frac{1}{2}} \left\{ \frac{\partial r}{\partial t} \frac{\partial^2 r}{\partial u \partial t} + f f' \frac{\partial r}{\partial u} \left(\frac{\partial \theta}{\partial t} \right)^2 + f^2 \frac{\partial \theta}{\partial t} \frac{\partial^2 \theta}{\partial u \partial t} \right\} \Big|_{u=0} dt \end{aligned}$$

在上述积分式中所出现的 $\left[\quad \right]^{-\frac{1}{2}} \Big|_{u=0}$ 因式, 在计算上是一个令人头痛者, 而我们可以运用下述简朴的几何思想把这个令人头痛者化解成恒

等于 1：一条曲线的长度当然和它的参数表式中参数选取无关。假如我们把上述参数 t 改用曲线 Γ_0 上的弧长参数 s ，则有

$$\left[\left(\frac{\partial r}{\partial s} \right)^2 + f(r)^2 \left(\frac{\partial \theta}{\partial s} \right)^2 \right]^{\frac{1}{2}} \Big|_{u=0} \equiv 1$$

请注意： s 只是 Γ_0 上的弧长参数，一般来说它并非其他的 Γ_u 上的弧长参数，亦即 (5.26)-式中的 $[\quad]^{-\frac{1}{2}}$ 只有在代入 $u=0$ 之后才恒等于 1；但是这也正是我们在简化 (5.26) 中所需要者。总之，在下述计算中，我们将改用上述 Γ_0 的弧长参数 s ，而 (5.26)-式则简化成

$$(5.26') \quad \frac{dL}{du} \Big|_{u=0} = \int_0^{L(0)} \left\{ \frac{\partial r}{\partial s} \frac{\partial^2 r}{\partial u \partial s} + f f' \frac{\partial r}{\partial u} \left(\frac{\partial \theta}{\partial s} \right)^2 + f^2 \frac{\partial \theta}{\partial s} \frac{\partial^2 \theta}{\partial u \partial s} \right\} \Big|_{u=0} ds$$

在 (5.26')-式之中所含的两个交差偏微分 $\frac{\partial^2 r}{\partial s \partial u} \Big|_{u=0}$ 和 $\frac{\partial^2 \theta}{\partial s \partial u} \Big|_{u=0}$ 的几何意义间接且不甚明确，宜用部分积分公式 (integration by parts) 把它们更换成其几何意义更加明确的 $\frac{\partial^2 r}{\partial s^2} \Big|_{u=0}$ 和 $\frac{\partial^2 \theta}{\partial s^2} \Big|_{u=0}$ 。这样的计算就可以把 (5.26')-式中的第一、第三项分别作下述改写：

$$(5.27) \quad \int_0^{L(0)} \frac{\partial r}{\partial s} \frac{\partial^2 r}{\partial s \partial u} \Big|_{u=0} ds = \left[\frac{\partial r}{\partial s} \frac{\partial r}{\partial u} \Big|_{u=0} \right]_0^{L(0)} - \int_0^{L(0)} \frac{\partial^2 r}{\partial s^2} \frac{\partial r}{\partial u} \Big|_{u=0} ds$$

$$\int_0^{L(0)} f^2 \frac{\partial \theta}{\partial s} \frac{\partial^2 \theta}{\partial s \partial u} \Big|_{u=0} ds = \left[f^2 \frac{\partial \theta}{\partial s} \frac{\partial \theta}{\partial u} \Big|_{u=0} \right]_0^{L(0)} - \int_0^{L(0)} \left\{ 2f f' \frac{\partial r}{\partial s} \frac{\partial \theta}{\partial s} \frac{\partial \theta}{\partial u} + f^2 \frac{\partial^2 \theta}{\partial s^2} \frac{\partial \theta}{\partial u} \right\} \Big|_{u=0} ds$$

而且由于 Γ_u 的始终两点是固定不动的，所以

$$\frac{\partial r}{\partial u} \Big|_{s=0} = \frac{\partial r}{\partial u} \Big|_{s=L(0)} = \frac{\partial \theta}{\partial u} \Big|_{s=0} = \frac{\partial \theta}{\partial u} \Big|_{s=L(0)} = 0$$

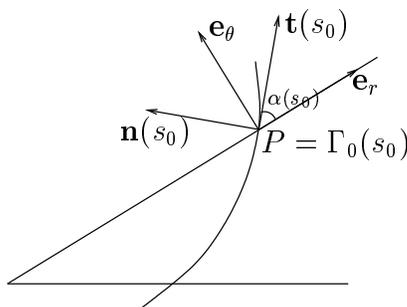
所以 (5.27') 的两式中的第一项皆为 0。将 (5.27')-式取代 (5.26)-式中的

第一、第三项，即得

$$(5.28) \quad \left. \frac{dL}{du} \right|_{u=0} = - \int_0^{L(0)} \left\{ \frac{\partial r}{\partial u} \frac{\partial^2 r}{\partial s^2} - f f' \frac{\partial r}{\partial u} \left(\frac{\partial \theta}{\partial s} \right)^2 + (2f f' \frac{\partial r}{\partial s} \frac{\partial \theta}{\partial s} + f^2 \frac{\partial^2 \theta}{\partial s^2}) \frac{\partial \theta}{\partial u} \right\} \Big|_{u=0} ds$$

现在，我们还要进一步研讨 (5.28)-式的表式整体的几何意义何在。此事当然得从起始曲线 Γ_0 在每一点 s 的局部几何中去探索之。

设 $P = \Gamma_0(s_0) = (r(s_0, 0), \theta(s_0, 0))$ 为起始曲线 Γ_0 上任给一点。如 [图 5-15] 所示， $M^2(f)$ 在 P 的切平面中具有两组自然给有的正交基 (orthonormal bases)，其一为曲线 Γ_0 在 P 点的单位切向量 $\mathbf{t}(s_0)$ 和单位法向量 $\mathbf{n}(s_0)$ ，其二则是极坐标系在 P 点坐标曲线单位切向量 $\mathbf{e}_r(P)$ 和 $\mathbf{e}_\theta(P)$ 。



[图 5-15]

令 $\alpha(s_0)$ 为它们之间相差的转角，则有

$$(5.29) \quad \begin{aligned} \mathbf{t}(s_0) &= \cos \alpha(s_0) \mathbf{e}_r + \sin \alpha(s_0) \mathbf{e}_\theta \\ \mathbf{n}(s_0) &= -\sin \alpha(s_0) \mathbf{e}_r + \cos \alpha(s_0) \mathbf{e}_\theta \end{aligned}$$

再者，由 [图 5-15] 可见

$$(5.30) \quad \begin{aligned} \cos \alpha(s_0) &= \left. \frac{dr}{ds} \right|_{s_0}, \quad \sin \alpha(s_0) = f(r) \left. \frac{d\theta}{ds} \right|_{s_0} \\ \mathbf{e}_r &= \frac{\partial}{\partial r}, \quad \mathbf{e}_\theta = \frac{1}{f(r)} \frac{\partial}{\partial \theta} \end{aligned}$$

其中 $\frac{\partial}{\partial r}$ 和 $\frac{\partial}{\partial \theta}$ 分别表示沿著坐标曲线「 $\theta = \text{常数}$ 」和「 $r = \text{常数}$ 」而且分别以 r 和 θ 为其参变数的速度向量。再对 (5.30)-(i) 微分即得

$$(5.31) \quad -\sin \alpha \frac{d\alpha}{ds} = \frac{d^2 r}{ds^2}, \quad \cos \alpha \frac{d\alpha}{ds} = f'(r) \frac{dr}{ds} \frac{d\theta}{ds} + f(r) \frac{d^2 \theta}{ds^2}$$

用上式来消去 (5.28)-式中的 $\frac{d^2 r}{ds^2}$ 和 $\frac{d^2 \theta}{ds^2}$ ，可得

$$(5.32) \quad \begin{aligned} \frac{dL}{du} \Big|_{u=0} &= - \int_0^{L(0)} \left(\frac{d\alpha}{ds} + f'(r) \frac{d\theta}{ds} \right) \cdot \left\{ -\sin \alpha \frac{\partial r}{\partial u} \Big|_{u=0} + \cos \alpha f(r) \frac{\partial \theta}{\partial u} \Big|_{u=0} \right\} ds \\ &= - \int_0^{L(0)} \left(\frac{d\alpha}{ds} + f'(r) \frac{d\theta}{ds} \right) \mathbf{n}(s) \cdot \mathbf{v}(s) ds \end{aligned}$$

其中

$$\mathbf{v}(s) = \frac{\partial r}{\partial u} \Big|_{u=0} \frac{\partial}{\partial r} + \frac{\partial \theta}{\partial u} \Big|_{u=0} \frac{\partial}{\partial \theta} = \frac{\partial r}{\partial u} \Big|_{u=0} \mathbf{e}_r + f(r) \frac{\partial \theta}{\partial u} \Big|_{u=0} \mathbf{e}_\theta$$

乃是沿著以 u 为参变数，曲线 $\{(r(s, u), \theta(s, u)), 0 \leq u \leq 1\}$ 在 $(r(s, 0), \theta(s, 0))$ 点的速度向量（参见 [图 5-15] 之所示）。通常我们把 $\{\mathbf{v}(s), 0 \leq s \leq L(0)\}$ 叫做变分向量场，把公式

$$(5.33) \quad \frac{dL}{du} \Big|_{u=0} = - \int_0^{L(0)} \left(\frac{d\alpha}{ds} + f'(r) \frac{d\theta}{ds} \right) \mathbf{n}(s) \cdot \mathbf{v}(s) ds$$

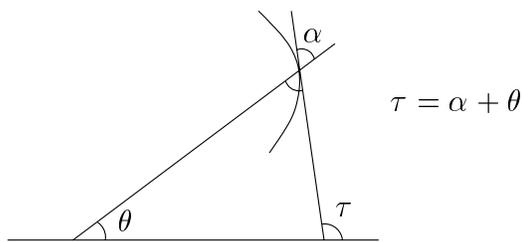
叫做 $M^2(f)$ 上的弧长第一变分公式。

[注]：上述 $M^2(f)$ 上的弧长第一变分公式的证明，对于初读者也许难以一读就懂。所以得逐步自行验算，逐步检查每步之要点。其实所用到的计算每步都是形式的微分、部分积分和代换化简。

在 $f(r) = r$ ，亦即 $M^2(f)$ 是欧氏平面的情形。如 [图 5-16] 所示，弧长第一变分公式之中所出现之

$$\frac{d\alpha}{ds} + f'(r) \frac{d\theta}{ds} = \frac{d\alpha}{ds} + \frac{d\theta}{ds} = \frac{d\tau}{ds}$$

乃是欧氏平面之中曲线的曲率 (curvature)。所以在 $M^2(f)$ 之中， $\frac{d\alpha}{ds} + f'(r) \frac{d\theta}{ds}$ 乃是曲线曲率的推广。



[图 5-16]

【定义】： $M^2(f)$ 中曲线的测地曲率 (geodesic curvature) 定义为

$$(5.34) \quad \kappa_g(s) = \frac{d\alpha}{ds} + f'(r) \frac{d\theta}{ds}$$

$M^2(f)$ 中测地曲率恒为 0 的曲线叫做 $M^2(f)$ 中的测地线 (geodesic curves)。

[定理 5.4]之证明：

$$(5.35) \quad \begin{aligned} \frac{d}{ds}(f(r) \sin \alpha) &= f(r) \cos \alpha \frac{d\alpha}{ds} + f'(r) \frac{dr}{ds} \sin \alpha \\ &= f(r) \frac{dr}{ds} \left(\frac{d\alpha}{ds} + f'(r) \frac{d\theta}{ds} \right) = 0 \end{aligned}$$

所以 $f(r) \sin \alpha$ 沿著一条测地线乃是一个常数。

[推论 3]之证明：

沿著测地线 \overline{AB} , $f(r) \sin \alpha = f(b) \sin A$ (常数), 所以

$$(5.36) \quad \begin{aligned} \frac{dr}{ds} = \cos \alpha &= \pm \sqrt{1 - \sin^2 \alpha} = \pm \sqrt{1 - \left(\frac{f(b) \sin A}{f(r)} \right)^2} \\ &= \frac{\pm \sqrt{f(r)^2 - (f(b) \sin A)^2}}{f(r)} \end{aligned}$$

亦即

$$(5.37) \quad \overline{AB} = \int_b^a ds = \int_b^a \frac{f(r) dr}{\pm \sqrt{f(r)^2 - (f(b) \sin A)^2}}$$

[注]：在 $f(r)$ 是一般的二阶连续可微函数的情形， $M^2(f)$ 中对于 $\triangle OAB$ 的广义余弦定律就是上述积分公式。但是在 $f(r)$ 是 r , $\sin r$ 或 $\sinh r$ 这

三种特殊而且重要的情形，则积分公式 (5.37) 可以用初等函数加以表达，这也就分别是欧氏、球面和非欧几何中的余弦定律。[其验算乃是积分计算的一个有意思的习题。]

5.5 旋转面的 Gauss 曲率和 Gauss-Bonnet 公式

在一般的旋转面 $M^2(f)$ 中，§5.4 中所证明的广义正弦、余弦定律只有在三角形的一个顶点位于旋转对称的基点，亦即 $\triangle OAB$ ，这种情形才成立。而且正弦、余弦定律分别只有一个公式，即

$$\frac{\sin A}{f(a)} = \frac{\sin B}{f(b)}$$

$$\overline{AB} = \int_b^a \frac{f(r)dr}{\pm \sqrt{f(r)^2 - (f(b) \sin A)^2}}$$

只有在 $M^2(f)$ 对于其中任给一点皆为旋转对称的情形，才会对于任何三角形都有其正弦、余弦定律，因为其任选一点皆可作为旋转对称之基点也！换句话说，只有在 $M^2(f)$ 是匀称 (homogeneous) 的情形，正弦、余弦定律才能够对于任给三角形普遍成立！

在这里自然会想到去研讨下述问题，即究竟有那些旋转面是匀称的呢？亦即对于其中每一点都是旋转对称者。欧氏平面、半径为 R 的球面以及非欧面都是匀称的旋转面，它们的本质函数 $f(r)$ 分别是

$$(5.38) \quad f(r) = \begin{cases} r & (\text{欧氏}) \\ \frac{1}{\sqrt{K}} \sin \sqrt{K}r & (\text{球面}, K = \frac{1}{R^2}) \\ \frac{1}{\sqrt{K'}} \sinh \sqrt{K'}r & (\text{非欧}, \delta(\Delta) = K'A(\Delta)) \end{cases}$$

从解析的观点来看，上述三者分别满足

$$(5.39) \quad f''(r) + Kf(r) = 0, \quad f(0) = 0 \quad \text{和} \quad f'(0) = 1$$

其中 $K = 0$ (欧氏), $K > 0$ (球面), $K < 0$ (非欧), 亦即 (5.39)-式中的 K 也就是使得

$$(5.39') \quad \varepsilon(\Delta) = KA(\Delta)$$

的那个本质性常数。

本节将证明下述匀称旋转面的唯一性定理，即

【定理 5.5】： $M^2(f)$ 是一个匀称旋转面的充要条件就是 $f(r)$ 满足 (5.39)。

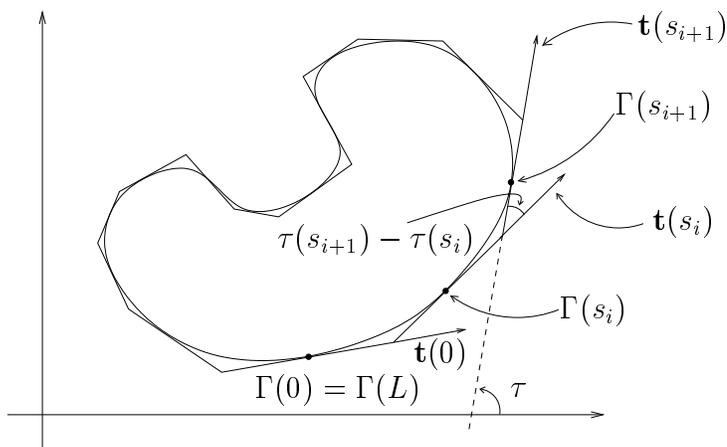
再者，从上述简短的分析，还可以想到[定理 5.5]的证明，自然地会和欧氏、球面以及非欧面的三角形内角和定理在 $M^2(f)$ 上的推广有关。首先，让我们先来分析一下三角形内角和定理本身在欧氏平面和球面上的自然推广应该是什么形式。

【分析】：

(一) 欧氏平面三角形内角和定理的自然推广：

由三角形内角和恒为一个平角，很容易推广成 n 边形的内角和恒等于 $(n-2)$ 个平角，因为一个 n 边形可以分割成 $(n-2)$ 个三角形。再者，上述 n 边形的内角和定理又可以转化为一个统一的外角和定理，即任何多边形的外角和恒等于一个周角，亦即 2π 。

设 Γ 是欧氏平面中一条不自相交的平滑闭曲线。 $0 \leq s \leq L$ 是 Γ 的弧长参数， $\kappa(s)$ 是它在 $\Gamma(s)$ 点的曲率， $\kappa(s) = \frac{d\tau}{ds}$ 。



[图 5-17]

如 [图 5-17] 所示，对于 Γ 的每一小段皆有

$$(5.40) \quad \int_{s_i}^{s_{i+1}} \kappa(s) ds = \int_{s_i}^{s_{i+1}} d\tau = \tau(s_{i+1}) - \tau(s_i)$$

所以由前述外角和定理即得

$$(5.41) \quad \int_0^L \kappa(s) ds = \sum_{i=0}^N \int_{s_i}^{s_{i+1}} \kappa(s) ds = \sum_{i=0}^N (\tau(s_{i+1}) - \tau(s_i)) = 2\pi$$

在一般的情形，设 Ω 是一个单连通的区域，其边界 $\partial\Omega$ 是分段平滑的曲线，而 α_i 则是 $\partial\Omega$ 在其有限个角点的外角，则不难看到

$$(5.42) \quad \int_{\partial\Omega} \kappa(s) ds + \sum \alpha_i = 2\pi$$

这也就是多边形外角和恒等于一个周角的推广。

(二) R -球面上三角形内角和的自然推广：

当 Δ 是一个 R -球面上的三角形时，即有

$$(5.43) \quad \begin{aligned} 2\pi - (\alpha_1 + \alpha_2 + \alpha_3) &= 2\pi - [(\pi - A) + (\pi - B) + (\pi - C)] \\ &= A + B + C - \pi = \varepsilon(\Delta) = \frac{1}{R^2} A(\Delta) \end{aligned}$$

当 Ω 是一个 R -球面上的多边形，则可以把 Ω 分割成三角形再逐个运用 (5.43)，即得

$$(5.44) \quad 2\pi - \sum \alpha_i = \frac{1}{R^2} A(\Omega)$$

由此可以想到在 R -球面上一个边界是分段平滑曲线的单连通区域 R ，相应于 (5.42)-式的推广公式应该是：

$$(*) \quad 2\pi - \int_{\partial\Omega} \kappa_g(s) ds - \sum \alpha_i = \frac{1}{R^2} A(\Omega)$$

其中 $\sum \alpha_i$ 乃是 Ω 在其各个角点的外角之和，而 $\kappa_g(s) = \frac{d\alpha}{ds} + \cos\left(\frac{r}{R}\right) \frac{d\theta}{ds}$ 是 R -球面上的测地曲率。至少，在 $\partial\Omega$ 是分段测地线的情形，(*) 的正确性业已成立，即 (5.44)。现在让我们再来验证一个 Ω 是一个半径为 r_0 的圆的特殊情形。如 [图 5-18] 所示， $\partial\Omega$ (亦即 $r = r_0$) 在每一点的 α 都等于 $\frac{\pi}{2}$, $r = r_0$ 。所以

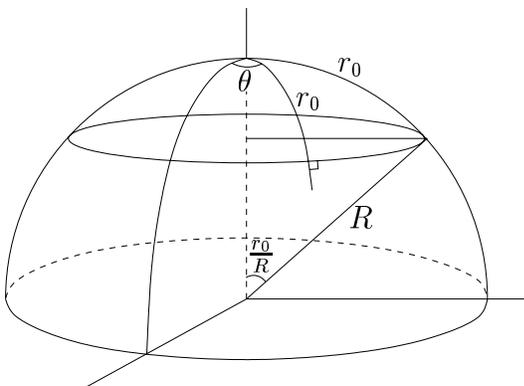
$$(5.45) \quad \begin{aligned} \kappa_g(s) &= \frac{d\alpha}{ds} + f'(r) \frac{d\theta}{ds} = \cos \frac{r_0}{R} \frac{d\theta}{ds} = \frac{1}{R} \cot \frac{r_0}{R} \\ \int_{\partial\Omega} \kappa(s) ds &= \frac{1}{R} \cot \frac{r_0}{R} \cdot 2\pi R \sin \frac{r_0}{R} = 2\pi \cos \frac{r_0}{R} \end{aligned}$$

另一方面，由 Archimedes 的球面面积公式的证明，易见

$$(5.46) \quad A(\Omega) = R^2 \cdot 2\pi \left(1 - \cos \frac{r_0}{R}\right)$$

这样也就验证了 (*)-式在上述特殊情形的正确性，即

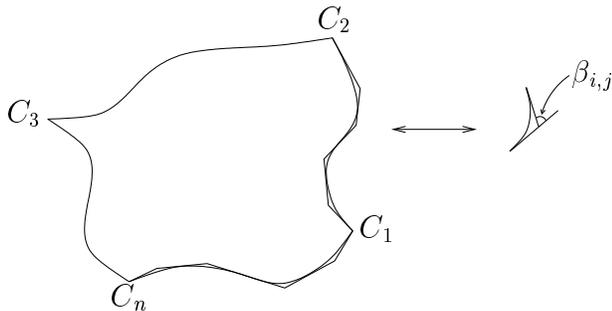
$$(5.47) \quad 2\pi - \int_{\partial\Omega} \kappa_g(s) ds = 2\pi \left(1 - \cos \frac{r_0}{R}\right) = \frac{1}{R^2} A(\Omega)$$



[图 5-18]

现在让我们进而验证 (*)-式在 R -球面上的普遍正确性：我们将采用一种简朴自然的想法，把上述验证归于前面两种业已得证的特殊情形（亦即 $\partial\Omega$ 是多边形和圆）的适当组合来加以论证。

设 Ω 是一个 R -球面上的单连通区域，而且其边界 $\partial\Omega$ 是分段平滑的。令 $\{C_i\}$ 是 $\partial\Omega$ 的角点， $\{\alpha_i\}$ 是各个角点的外角。在 $\partial\Omega$ 的第 i 段平滑曲线 $\widehat{C_i C_{i+1}}$ 选取其 n -等分点，然后再在每点作其相切于 $\partial\Omega$ 的测地线（亦即大圆圆弧）。如 [图 5-19] 所示，它们围成一个其各边和 $\partial\Omega$ 相切于那些等分点而且和 $\partial\Omega$ 在角 $\{C_i\}$ 具有同一外角的多边形 Ω_n 。



[图 5-19]

对于上述所构造的多边形 Ω_n ，它除了在那些和 Ω 共有的角点 $\{C_i\}$ 的外角是 α_i ，还有那种如局部放大插图所示的外角 $\beta_{i,j}$ 。所以有

$$(5.48) \quad 2\pi - \sum_i \alpha_i - \sum_{i,j} \beta_{i,j} = \frac{1}{R^2} A(\Omega_n)$$

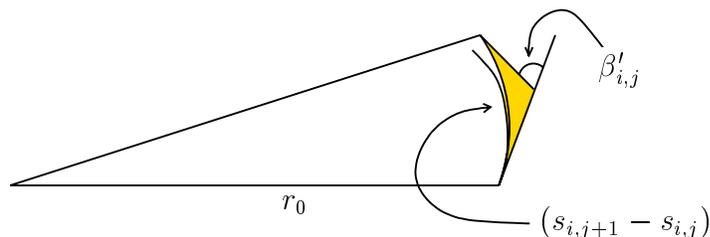
把 (5.48)-式和所要证明的 (*) 相比较，不难看到我们所要证明者，可以归于下述极限式去验证之，亦即 (**) 是否成立？

$$(**) \quad \lim_{n \rightarrow \infty} \left(\int_{s_i}^{s_{i+1}} \kappa_g(s) ds - \sum_{j=1}^n \beta_{i,j} \right) = 0$$

再者，由积分的定义，(**)-式的验证又可以归于

$$(***) \quad \left| \kappa_g(s_{i,j}) \cdot (s_{i,j+1} - s_{i,j}) - \beta_{i,j} \right|$$

是否是微量 $(s_{i,j+1} - s_{i,j})$ 的高阶微量（亦即是至大和 $(s_{i,j+1} - s_{i,j})^2$ 相当者）的验证。由于 $\kappa_g(s)$ 是一个连续可微函数，所以 (***) 的验证又可以归于 $\kappa_g(s)$ 在微段 $[s_{i,j}, s_{i,j+1}]$ 上取常数值 $\kappa_g(s_{i,j})$ 的情形验算之，亦即 $\partial\Omega$ 在上述微段乃是一段圆弧的情形。由 [图 5-20] 所示



[图 5-20]

由 (5.45), (5.46) 的局部形式可得

$$(5.49) \quad \kappa_g(s_{i,j})(s_{i,j+1} - s_{i,j}) - \beta'_{i,j} = \frac{1}{R^2} (\triangleright)$$

而 (5.49)-式右侧的面积显然是弧长 $(s_{i,j}, s_{i,j+1})$ 的二阶微量。这样，也就验证了 (*)-式在 R -球面上是普遍成立的。

从上面对于欧氏平面和 R -球面这种简朴而且基本的情形，实事求是的探索和验证，使得我们认识到

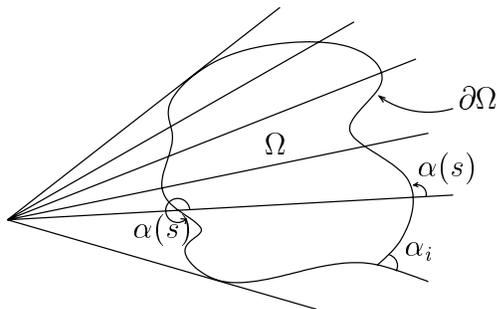
$$(*) \quad 2\pi - \int_{\partial\Omega} \kappa_g(s) ds - \sum \alpha_i$$

乃是一个值得深究的基本几何量，因为它乃是多边形外角和的自然推广，而且在欧氏和 R -球面的情形简洁地反映著它们的「本质性二维曲率」。有了这样一个明确的目标，就不难实事求是去研究它和区域 Ω 的那一种几何本质密切相关。这种研讨之所得就是下述重要的几何公式，即

【定理 5.6】（旋转面的 Gauss-Bonnet 公式）：

$$(5.50) \quad 2\pi - \int_{\partial\Omega} \kappa_g(s) ds - \sum \alpha_i = \iint_{\Omega} \left(-\frac{f''(r)}{f(r)} \right) dA$$

证明：若将所给的 Ω 分割成 Ω_1, Ω_2 ，则易见 (5.50)-式的两侧都可以分别表成它在 Ω_1, Ω_2 的几何量之和。由此可见[定理 5.6]的证明可以用适当的分割，把一般的情形归于 Ω 如 [图 5-21] 所示的情形加以论证，亦即 Ω 位于 $\theta_0 \leq \theta \leq \theta_1$ 的范围之内。



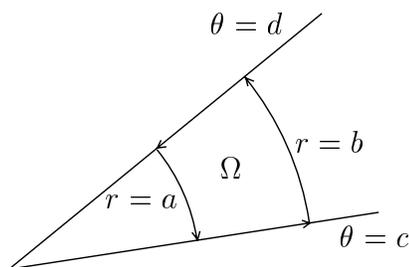
[图 5-21]

对于这样的区域 Ω , $2\pi - \int_{\partial\Omega} d\alpha - \sum \alpha_i = 0$ 。所以

$$\begin{aligned}
 & 2\pi - \int_{\partial\Omega} \kappa_g(s) ds - \sum \alpha_i \\
 &= (2\pi - \int_{\partial\Omega} d\alpha - \sum \alpha_i) - \int_{\partial\Omega} f'(r) d\theta \\
 (5.51) \quad &= - \int_{\partial\Omega} f'(r) d\theta = - \iint_{\Omega} f''(r) dr \wedge d\theta \quad (\text{Green's 定理}) \\
 &= \iint_{\Omega} \left(-\frac{f''(r)}{f(r)} \right) f(r) dr \wedge d\theta = \iint_{\Omega} \left(-\frac{f''(r)}{f(r)} \right) dA
 \end{aligned}$$

□

[注]: $\int_{\partial\Omega} f'(r)d\theta = \iint_{\Omega} f''(r)dr \wedge d\theta$ 是二元微积分中 Green's 定理的特殊情形。其证明可以归于 Ω 是坐标矩形 (亦即 $\Omega = \{(r, \theta), a \leq r \leq b, c \leq \theta \leq d\}$) 的情形组合而得。再者, 在 Ω 是坐标矩形的情形, 如 [图 5-22] 所示:



[图 5-22]

$$\begin{aligned}
 \int_{\partial\Omega} f'(r)d\theta &= f'(b)(d-c) + f'(a)(d-c) \\
 (5.52) \quad &= (d-c)(f'(b) - f'(a)) = (d-c) \int_a^b f''(r)dr \\
 &= \int_c^d \left[\int_a^b f''(r)dr \right] d\theta = \iint_{\Omega} f''(r)dr \wedge d\theta
 \end{aligned}$$

乃是单元微积分基本定理的直接推论。

【推论 1】: $-\frac{f''(r)}{f(r)}$ 乃是一个具有本质性的局部 (intrinsic local) 几何量, 它就是 $M^2(f)$ 在 $P(r, \theta)$ 点的高斯曲率。

证明: 当我们在 $P(r_0, \theta_0)$ 点任选一个邻域 Ω , 则由 (5.50)-式的左侧可见 $\iint_{\Omega} \left(-\frac{f''(r)}{f(r)} \right) dA$ 是一个本质性的局部几何量。再者当 $A(\Omega) \rightarrow 0$

$$(5.53) \quad -\frac{f''(r_0)}{f(r_0)} = \lim_{A(\Omega) \rightarrow 0} \frac{\iint_{\Omega} \left(-\frac{f''(r)}{f(r)} \right) dA}{A(\Omega)}$$

所以 $-\frac{f''(r)}{f(r)}$ 本身乃是一个具有本质性的局部几何量。 \square

【推论 2】： $M^2(f)$ 是匀称 (homogeneous) 的必要条件是 $-\frac{f''(r)}{f(r)}$ 等于一个常数。

证明： $M^2(f)$ 是一个匀称的旋转面，亦即 $M^2(f)$ 中任给两点 P, Q 都有一个 $M^2(f)$ 的保长变换 (isometry) 把 P 点映射到 Q 点。由此可见 P, Q 两点具有完全相同的局部几何性质，所以 $\left(-\frac{f''(r)}{f(r)}\right)_P = \left(-\frac{f''(r)}{f(r)}\right)_Q$ ，亦即 $-\frac{f''(r)}{f(r)}$ 是一个常数。□

【推论 3】（定理 5.5）： $M^2(f)$ 是匀称的必要条件是

$$(5.54) \quad f(r) = \begin{cases} r & (K = 0) \\ \frac{1}{\sqrt{K}} \sin \sqrt{K}r & (K > 0) \\ \frac{1}{\sqrt{-K}} \sinh \sqrt{-K}r & (K < 0) \end{cases}$$

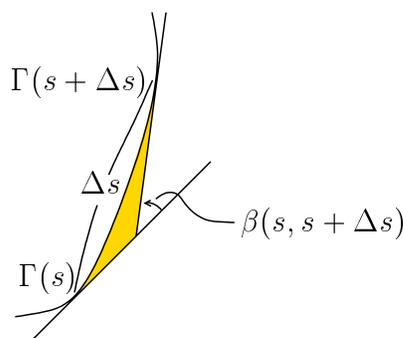
证明：由 [推论 2] 得知 $-\frac{f''(r)}{f(r)} = \text{常数} = K$ ，亦即 $f(r)$ 乃是下述二阶常微分方程

$$(5.55) \quad f''(r) + Kf(r) = 0$$

满足初值条件 $f(0) = 0$ 和 $f'(0) = 1$ 的解。易证这种解是唯一的，而 (5.54) 所给者就是那个唯一解。□

【推论 4】：设 Γ 是一条平滑曲线， s 是其弧长参数。如 [图 5-23] 所示， $\beta(s, s + \Delta s)$ 乃是 Γ 在 $\Gamma(s)$ 和 $\Gamma(s + \Delta s)$ 这两点的测地线切线相交的外角，则有

$$(5.56) \quad \lim_{\Delta s \rightarrow 0} \frac{\beta(s, s + \Delta s)}{\Delta s} = \kappa_g(s)$$



[图 5-23]

证明：令 Ω 为 [图 5-23] 中阴影所示的区域。由公式 (5.50)，即有

$$(5.57) \quad 2\pi - \int_{s+\Delta s}^s \kappa_g(s) ds - 2\pi - \beta(s, s + \Delta s) = \iint_{\Omega} \left(-\frac{f''(r)}{f(r)} \right) dA$$

而上式右侧之量显然是和 $A(\Omega)$ 相当者，所以是和 Δs^2 相当者。因此

$$(5.58) \quad \lim_{\Delta s \rightarrow 0} \frac{\beta(s, s + \Delta s)}{\Delta s} = \lim_{\Delta s \rightarrow 0} \int_s^{s+\Delta s} \kappa_g(s) ds = \kappa_g(s) \quad \square$$

[注]：(5.56)-式也就说明了一条曲线的测地曲率的几何内涵及其名称的来由。

5.6 结语

总结本章的研讨，我们对于欧氏、球面和非欧这三种具有高度对称性的几何体系，获得了下述几点统一全局的认知：

(一) 三者具有同样的连续性和对称性（例如三角形 S.A.S. 叠合性），但是三者的三角形内角和则分别是等于，大于和小于一个平角。

(二) 对于对称性的深入定量分析所得的基本定理就是三种几何学中的正弦、余弦定律。而且通过旋转面的解析几何，我们认识到正弦定律的本质乃是沿著测地线的等速运动对于旋转对称的角动量守恒定律。这种统一的观点使得我们一方面可以把原来只是在欧氏、球面和非

欧几何中的正弦、余弦定律推广成所有旋转面皆有的广义正弦、余弦定律，而且也因而把原先各别论证的基本定理给以统一的证明。

(三) 对于欧氏、球面和非欧几何中的三角形内角和，我们有一个形式上统一的定理，亦即

$$\varepsilon(\Delta) = KA(\Delta)$$

其中 K 分别是等于，大于和小于零的常数。把它扩展而且深化之所得，即为下述 Gauss-Bonnet 公式：

$$2\pi - \int_{\partial\Omega} \kappa_g(s) ds - \sum \alpha_i = KA(\Omega)$$

而上述公式又可以推广成任给旋转面 $M^2(f)$ 上的 Gauss-Bonnet 公式，即

$$2\pi - \int_{\partial\Omega} \kappa_g(s) ds - \sum \alpha_i = \iint_{\Omega} \left(-\frac{f''(r)}{f(r)} \right) dA$$

(四) 上述公式的[推论 2]和[推论 3]确立了具有欧氏、球面和非欧几何同样的高度对称性的几何体系是唯一存在的，对于每一个 K 只有一个，它就是 (5.54)-式所给定的 $f(r)$ 所组成的 $M^2(f)$ ！

(五) 上述唯一性定理的一个直接推论是：在欧氏公理体系中，若将第五公设改用任何「一小点」在欧氏空间成立但是在非欧空间不成立的假设，皆足以完全刻划欧氏几何。这也就是为什麼两千多年来有很多杰出的几何学家，误认为他们业已找到第五公设的证明的原由。因为在任何地方不自觉地渗入「一小点」欧氏几何中显然成立的「奥援」(implicit help)，第五公设就可以从而得以证之。从 Gauss 关于这方面的书信中，可以看到他在试图证明第五公设中，肯定有过多地渗入杂念因而犯错的私下体验。

(六) 由本章的讨论，可以看到欧氏、球面和非欧这三种几何，不但大同小异，而且还存在著大量深刻的共同定理。这种三者皆成立的共性几何的系统研究，其实是非常有意义而且有待我们致力拓展的领域。当年 Bolyai 在他开创性的短文（总共只有 26 页），业已创导这样一种几何学，他称之为绝对几何 (absolute geometry)。Bolyai 正弦定律，即

$$\frac{\sin A}{\odot a} = \frac{\sin B}{\odot b} = \frac{\sin C}{\odot c}$$

就是其中所得的一个辉煌成就。本章的研讨方式可以说是直接继承他的创见者。其实，也唯有对于绝对几何学作深入研究，我们才能真正深刻地理解对称性的既深且广的影响与威力！

5.7 思考题与习题

I. 旋转面上测地线的参数式

极坐标乃是一个具有旋转对称的二维空间的简洁表述，它简明扼要地描述了任给一点 P 和基点 O 之间的距离和方向 (r, θ) ，而且弧长元素： $ds^2 = dr^2 + f(r)^2 d\theta^2$ 之中的 $f(r)$ 则简明扼要地总结了所给旋转面的整体本质 (complete intrinsic invariant)。再者，测地线 乃是空间之中最为简朴基本的几何事物，在极坐标中，它们可以用下述二阶常微分方程加以刻划，即

$$(*) \quad \frac{d\alpha}{ds} + f'(r) \frac{d\theta}{ds} = 0, \quad \cos \alpha = \frac{dr}{ds}, \quad \sin \alpha = f(r) \frac{d\theta}{ds}$$

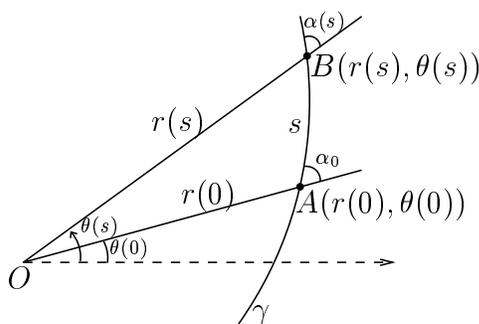
再者，上述二阶常微分方程乃是下述含有一个任意常数的一阶常微分方程的微分所得，即

$$(**) \quad f(r) \sin \alpha = C_1 \quad (\text{常数})$$

通常把 (**) 叫做 (*) 的第一积分 (first integral)。当然，在一个给定旋转面 $M^2(f)$ 上的一条曲线最为明确的描述乃是直接给出它的极坐标参数式，即

$$(***) \quad r = \varphi(s), \quad \theta = \psi(s) \quad (s: \text{弧长参数})$$

为此，我们得由 (**) 再做一次积分：



[图 5-24]

如 [图 5-24] 所示, γ 乃是以 $A(r(0), \theta(0))$ 为起点, 以 α_0 为其起始方向的测地线。令 $C_1 = f(r(0)) \sin \alpha_0$, 再由

$$(5.59) \quad \frac{dr}{ds} = \cos \alpha = \pm \sqrt{1 - \sin^2 \alpha} = \frac{\pm \sqrt{f(r)^2 - C_1^2}}{f(r)}$$

积分可得

$$(***)_1 \quad s = \pm \int_{r(0)}^{r(s)} \frac{f(r) dr}{\sqrt{f(r)^2 - C_1^2}}$$

亦即所求的 r 的参数表式 $\varphi(s)$ 乃是上述积分式所表述的函数关系的逆函数是也。这也就是广义余弦定律。

再者,

$$(5.60) \quad C_1 = f(r) \sin \alpha = f(r)^2 \frac{d\theta}{ds}$$

积分可得

$$(***)_2 \quad \theta(s) - \theta(0) = \int_0^s \frac{C_1}{f(r)^2} ds$$

对于一般的 $f(r)$, 测地线的参数式只能写成上述两个积分表达的形式, 即 $(***)_1$ 和 $(***)_2$ 。但是在 $f(r) = r$, 或 $\sin r$, 或 $\sinh r$ 这三种既特别又重要的情形, 上述积分表达式是可以利用初等函数直接表达的, 这也就是我们要研讨的一系列习题, 它们可以看做微积分中的积分技巧 (techniques of integration) 的一种自然而且重要的应用。

【习题】：

- (1) 在 $f(r) = r$ 的情形（亦即欧氏平面），求解初值为 $\{r(0) = b, \theta(0) = 0, \alpha_0 = \frac{\pi}{2}\}$ 的测地线的参数式。
- (2) 在 $f(r) = \sin r$ 的情形（亦即单位球面），求解初值为 $\{r(0) = b, \theta(0) = 0, \alpha_0 = \frac{\pi}{2}\}$ 的测地线的参数式。
- (3) 定义 $\cosh x = \frac{1}{2}(e^x + e^{-x})$, $\sinh x = \frac{1}{2}(e^x - e^{-x})$ 。验证下列等式：

$$\begin{aligned} \cosh^2 x - \sinh^2 x &= 1 \\ \cos(ix) &= \cosh x, \quad \sin(ix) = i \sinh x \\ \cosh(x+y) &= \cosh x \cosh y + \sinh x \sinh y \\ \sinh(x+y) &= \sinh x \cosh y + \cosh x \sinh y \\ \frac{d}{dx} \cosh x &= \sinh x, \quad \frac{d}{dx} \sinh x = \cosh x \end{aligned}$$

[注]：Euler 公式： $\cos x = \frac{1}{2}(e^{ix} + e^{-ix})$, $\sin x = \frac{1}{2i}(e^{ix} - e^{-ix})$ 。

- (4) 在 $f(r) = \sinh r$ 的情形（亦即曲率为 -1 的非欧面），求解初值为 $\{r(0) = b, \theta(0) = 0, \alpha_0 = \frac{\pi}{2}\}$ 的测地线的参数式。

II. $M^2(\sinh r)$ 的匀称性：

在[推论 3]业已证明了 $M^2(f)$ 只有在

$$f(r) = \begin{cases} r & (K = 0) \\ \frac{1}{\sqrt{K}} \sin \sqrt{K} r & (K > 0) \\ \frac{1}{\sqrt{-K}} \sinh \sqrt{-K} r & (K < 0) \end{cases}$$

这三种情形才可能是匀称者。但是上述三种 $M^2(f)$ 是否真的是匀称者？则当然还得实事求是地加以研讨之。在第一种情形 $M^2(f)$ 其实就是欧氏平面，而在第二种情形 $M^2(f)$ 其实就是半径为 $\frac{1}{\sqrt{K}}$ 的球面，所以它们的匀称性乃是熟知的事实。因此，唯一尚须详加研究者乃是第三种情形。再者，由于匀称性显然在整体的放大、缩小之下保持不变，所以只需要对于其中任给一个取定的 $K < 0$ 加以研讨（例如 $K = -1$ ），亦即 $M^2(\sinh r)$ 是否匀称？上述问题乃是一个重要的基本问题！我们可以从正、反两面来分析其重大的意义如下：

(一) 假如 $M^2(\sinh r)$ 是匀称的, 则它就是一个几何体系, 它和欧氏平面的唯一差别在于其三角形内角和恒小于平角。所以平行公理当然不可能由其他的欧氏几何公理推导而得, 这也就证明了非欧几何体系的存在性。

(二) 反之, 假如 $M^2(\sinh r)$ 「并非」匀称, 则欧氏平面乃是唯一可定向 (orientable) 而且两点定一测地线的匀称旋转面 (球面上过对顶两点有无穷多个大圆, 而射影平面则是不可定向者也)。由此可见, 假如真的能够「证明」 $M^2(\sinh r)$ 是不匀称的话, 这个证明其实也就构成了一个由其他的欧氏几何公理推导出平行公理的一个论证, 是不? 当然, 此事业已在十九世纪真相大白, $M^2(\sinh r)$ 是的的确确匀称的。而我们下面所要和同学研讨者, 乃是如何实事求是地去逐步证明 $M^2(\sinh r)$ 的匀称性。

【分析】:

(一) 对于一般的旋转面 $M^2(f)$, 我们所证的广义正弦、余弦定律, 仅仅对于 $\triangle OAB$ 才成立, 而且仅仅各有一个等式, 即

$$\frac{\sin A}{f(a)} = \frac{\sin B}{f(b)} \quad (\text{广义正弦定律})$$

$$\overline{AB} = \pm \int_b^a \frac{f(r)dr}{\sqrt{f(r)^2 - (f(b) \sin A)^2}} \quad (\text{广义余弦定律})$$

假如 $M^2(f)$ 是匀称的, 则不但上述广义正弦、余弦定律对于任何三角形 $\triangle ABC$ 都普遍成立, 而且有下述两组等式, 即

$$\frac{\sin A}{f(a)} = \frac{\sin B}{f(b)} = \frac{\sin C}{f(c)}$$

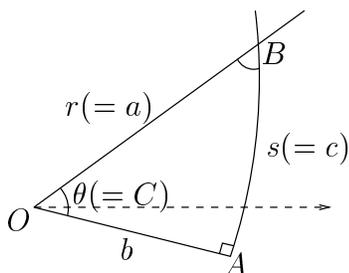
$$\begin{cases} \overline{AB} = \pm \int_b^a \frac{f(r)dr}{\sqrt{f(r)^2 - (f(b) \sin A)^2}} \\ \overline{BC} = \pm \int_c^b \frac{f(r)dr}{\sqrt{f(r)^2 - (f(c) \sin B)^2}} \\ \overline{CA} = \pm \int_a^c \frac{f(r)dr}{\sqrt{f(r)^2 - (f(a) \sin C)^2}} \end{cases}$$

从上面这一段分析，可以想到 $M^2(\sinh r)$ 的匀称性的验证的自然起步，应该就是推导 $\triangle OAB$ 的其他几个等式，例如

$$\frac{\sin \angle AOB}{f(\overline{AB})} = \frac{\sin A}{f(a)}, \quad \text{等等}$$

【习题】：

- (5) 首先考虑 $f(r) = \sin r$ (即单位球面) 的情形。设 $\angle A = \frac{\pi}{2}$ ，如 [图 5-25] 所示：



[图 5-25]

试验证余下的正弦、余弦定律：

- (i) $\frac{\sin A}{\sin a} \stackrel{?}{=} \frac{\sin C}{\sin c}$ (注意： $\sin A = 1$)
 (ii) $\cos B \stackrel{?}{=} \frac{\tan c}{\tan a}$
 (iii) $\cos C \stackrel{?}{=} \frac{\tan b}{\tan a}$

- (6) 考虑 $f(r) = \sinh r$ (亦即曲率为 -1 的非欧面)。设 $\angle A = \frac{\pi}{2}$ ，写下并验证余下的正弦、余弦定律。
- (7) (讨论) 上述对于直角非欧三角形成立 $\triangle OAB$ 的正弦、余弦定律是否能够推广至任意的非欧三角形？再者，能否以此为根据来解答非欧面的匀称性 (亦即存在性) 问题？

[注]：上面这一段讨论，乃是基于匀称旋转面的唯一性的认知，按图索骥地去验证 $M^2(\sinh r)$ 乃是匀称者，从而证明非欧面之存在性。当然也可以用其他数学知识，设法去构造一个相异于球面或欧氏平面的匀称旋

转面，则由[推论 3]可知它必然是 $K < 0$ 者，这样也就证明了 $M^2(\sinh r)$ 的匀称性，例如在匀称空间 $SO(2,1)/SO(2)$ 上就可以直接构造之（但是得用到李群的概念）。

结语

本册所讨论的单元微积分的基础理论和初步应用乃是整个分析学的雏形与基础之所在。分析学所研讨者乃是变量数学，它是我们对于千变万化的大自然，由表及里、精益求精地作数理分析的主要工具。其实，在一般所要研讨的事物和现象中，其所涉及的参变量当然不可能只是简简单单地两个参变量（即一个自变量和一个因变量），而是多个参变量而且它们之间又具有多个相互关联的函数关系者也。例如三角形就有三个边长，三个角度，面积，外径，内径等等参变量，而它们之间又满足正弦、余弦定律，面积公式，外径、内径公式等等。由此可见，本册所研讨之课题，乃是一种最为简化的基本情况：一个单变函数的分析，它是我们进而研讨多元多关系的分析学的起点和基本功。唯有先把它学得扎实，懂得清楚，学习分析学、学会运用数理分析去理解大自然才有了好的开始，也就不难登堂入室了。在此且再对本册所讨论的单元微积分的精要作一简短的回顾与总结，作为本册的结语。

（一）函数的各种各样的性质之中最为基本者有四，即单调性、连续性，变率与总和，其前两者是定性者而後两者则是定量者。总之，对于上述四者的研讨、理解与运用乃是单元微积分的中心课题和基础理论之所在。

（二）单调性是十分简单而且易用好用的。再者，绝大部分常见常用的函数都是单调或分段单调的，所以读者切勿因为它简单而忽视它，而是要注意到「有则善用之」。例如在第三章研讨指数函数的定义中，善用其单调性和保有指数定则乃是要点之所在。

（三）一个函数 $y = f(x)$ 在某一区间 $[a, b]$ （或 (a, b) ）上的连续性所包含的条件乃是无比庞大的极限式（亦即 $\lim f(a_n) = f(\lim a_n)$ 对于所有

其中收斂的數列恒成立)。由此可見函數連續性的理解與運用是的不簡單的。首先，我們先得理解直線（和它相應的實數系）的連續性，懂得如何把其所含的幾何直觀轉換成左、右夾逼數列的存在性這種解析描述，然後再通過把它用來研討閉線段上連續函數的基本性，證明如 Sturm 定理，代數基本定理，Cauchy 條件的極限存在定理等等才能逐步逐樣地體認其用場與用法。

（四）歸本究源，Hippasus 對於「不可公度線段比」的發現乃是人類理性文明首次觸及空間連續性這個深蘊的本質，而 Eudoxus 所創的比較原則和逼近法不但成功地補救了當年定量幾何基礎論的重大欠缺，而且也提供了理解連續性並進而研討整個分析學的基本方法。這一段引人入勝、發人深思的史話當然也就是每一位學習分析學的後之來者最自然的啟蒙、奠基的佳園。

（五）本冊第二章所討論者，也就是效法 Eudoxus 所創的途徑，把它用來研討函數的兩個定量型的基本性質：變率與總和。其具體做法是先行分析其直觀內含，從而確立其比較原則，然後再用逼近法去達成它們各別的解析定義。這也就是微分求變率，積分求總和的來由，其本身乃是分析學基礎理論的初步階段。接著把它們和函數連續性（第一章所研討者）相結合，就不難順理成章地證明各種各樣基本定理（參看第三章）。

基础数学讲义之五

《基础分析学之二》

多元微积分学

项武义

香港科技大学数学系

目录

引子	v
一 多元函数的连续性与微分	1
1.1 多元函数的连续性	1
1.2 多元函数的微分	6
二 多元多关系的微分	19
2.1 隐函数定理	20
2.2 坐标变换	24
2.3 极大, 极小的微分条件式	27
三 高维勾股定理与 Grassmann 代数	33
3.1 向量代数与平行体的有向体积	33
3.1.1 平面的定向与平行四边形的有向面积	34
3.1.2 三维空间的定向和平行六面体的有向体积	35
3.2 向量内积与勾股定理的高维推广	38
3.3 格氏代数	47
四 外微分与多元积分	51
4.1 多元函数的多重积分	51
4.1.1 多重积分的定义	51
4.1.2 多重积分与坐标变换	56
4.2 线积分、面积分及其高维推广	61
4.2.1 线积分	62
4.2.2 曲面积分	64

4.2.3	例子	66
4.2.4	习题	76
4.3	外微分和微积分基本定理的高维推广	79
4.3.1	外微分和广义 Stoke's 定理	87
4.3.2	例子	90
	编後语	103

引子

本册将以上一册研讨单元微积分所得之基础理论为基本，进而研讨多元微积分。一如在上一册的结语中所提及的，在各种各样数理分析中所遇到的问题，通常都是多元、多关系的体系而不是只有一个自变元者。总之，多元微积分才是普遍可用者，而单元微积分则仅仅是在理论上提供了简朴的雏型和基础。把它进而推广到多元、多关系的范畴，一来是十分自然的顺理成章，二来也是迫切亟需的；这是分析学必然的进程。

第一章

多元函数的连续性与微分

1.1 多元函数的连续性

设 $u = f(x_1, \dots, x_n)$ 是一个给定的 n 个自变元 (x_1, \dots, x_n) 的函数， n 数组 (x_1, \dots, x_n) 的变域是 \mathbb{R}^n 中的一个区域 D ，而应变元 u 之值则由 n 数组 (x_1, \dots, x_n) 之给定而唯一确定者也。它在某一给定点 (a_1, \dots, a_n) 的「局部连续性」的定义乃是在单元者的直接推广，亦即

【定义】：定义于 D 上的 $u = f(x_1, \dots, x_n)$ 在 $A(a_1, \dots, a_n)$ 点的连续的条件是对于任给 $\varepsilon > 0$ ，恒有足够小的 $\delta > 0$ 使得

$$\begin{aligned} |x_i - a_i| < \delta, \quad 1 \leq i \leq n, \quad (x_1, \dots, x_n) \in D \\ \Rightarrow |f(x_1, \dots, x_n) - f(a_1, \dots, a_n)| < \varepsilon \end{aligned}$$

我们也可以改用数列与极限，把上述局部连续性重述如下：

对于任给 D 中以 A 为其极限的点列 $\{P_k\}$ ，亦即

$$P_k(x_{1,k}, \dots, x_{n,k}) \in D, \quad \lim_{k \rightarrow \infty} x_{i,k} = a_i, \quad 1 \leq i \leq n$$

皆有

$$\lim_{k \rightarrow \infty} f(x_{1,k}, \dots, x_{n,k}) = f(a_1, \dots, a_n)$$

由此可见 $f(x_1, \dots, x_n)$ 在其定义域 D 上到处连续的条件就是对于任给 D 中其极限依然在 D 中的点列 $\{P_k\}$ ，皆有

$$\lim_{k \rightarrow \infty} f(P_k) = f(\lim_{k \rightarrow \infty} P_k)$$

亦即对于任给 n 个数列 $\{x_{i,k}\}, 1 \leq i \leq n$, 若有

$$(x_{1,k}, \dots, x_{n,k}) \in D, \lim_{k \rightarrow \infty} x_{i,k} \text{ 存在}, \quad 1 \leq i \leq n$$

$$\text{而且 } (\lim_{k \rightarrow \infty} x_{1,k}, \dots, \lim_{k \rightarrow \infty} x_{n,k}) \in D$$

则恒有

$$\lim_{k \rightarrow \infty} f(x_{1,k}, \dots, x_{n,k}) = f(\lim_{k \rightarrow \infty} x_{1,k}, \dots, \lim_{k \rightarrow \infty} x_{n,k})$$

[注]: (i) 从上述多元函数的连续性的定义的基本面 (basic feature) 来看, 多元函数的情形和单元函数的情形如出一辙, 只是每个变元 x_i 皆有其趋于 a_i 的数列 $\{x_{i,k}\}$, 所以在上述连续性的条件式中涉及 n 个收敛的数列而不再是仅仅一个收敛数列。

(ii) 其实多元的连续性和单元的连续性两者相比, 前者的确要比后者来得复杂多样, 其主要原因有二: 其一是 \mathbb{R}^n 中的区域 D 要远比 \mathbb{R}^1 中的区间多样得多; 其二是上述极限条件式要比单元者强得很多 (在单元的情形一个收敛数列基本上只有从左、右两个方向去逼近其极限值; 但是在多元的情形则可以从无穷多个方向逼近其极限点), 所以多元函数的不连续性要远比单元函数的不连续性复杂多样。由此可以想到, 在多元的数理分析中, 连续性就变得更加重要了。很多常用好用的公式和定理, 往往都有赖于所涉及的函数的连续性!

(iii) 多元函数的连续性在本质上是一个非常强的条件, 但是它也是一种非常自然的条件。所以在很多多元的数理分析问题中, 所涉及的函数往往都自然而然地在大部分定义域上连续。总之, 在对于某一问题作数理分析中, 对于其所涉及的函数的连续性成立的范围务必小心检查; 而在连续性不成立的地方 (奇点), 当然就得格外用心, 另行设法区别处理之。

(iv) 单元微积分中, 在一个闭线段上到处连续的函数具有好些优良的基本性质 (参看第四册第一章), 而且它们又在单元微积分的基础理论中扮演重要的角色。很自然地我们在此要问: 在多元的各种各样变域中, 究竟那些才是「闭线段」的适当推广呢? 换句话问, 在那一种变域上到处连续的多元函数依然保有在闭线段上到处连续的单元函数所具有的那些优良性质呢?

【分析】：

要解答上述问题，自然又得对于上一册第一章对于那些优良性质的论证，再做一次温故知新，分析一下在论证中，所用到的闭线段的本质究竟是什麼？不难看到下述三点是显然必要的：其一是有界性，其二是连通性而其三则是极限封闭性，亦即闭线段中的任给收敛数列的极限点依然位于其中。

由此可见，要真正掌握多元函数的连续性，就必须对 \mathbb{R}^n 中的区域的几何本质先下一番功夫，明确其有界性、连通性、极限封闭性等等的实质内含。

有界性：将 \mathbb{R}^1 中一个子集 S 的有界性直接推广，即得 \mathbb{R}^n 中的一个子集 S 的有界性之定义，即：存在一个足够大的 K 使得所有 S 中之点的坐标皆满足 $|x_i| \leq K, 1 \leq i \leq n$ 。

例如下述 n -维方块：

$$\square^{(n)}(2K) := \{(x_1, \dots, x_n); |x_i| \leq K, 1 \leq i \leq n\}$$

本身乃是 \mathbb{R}^n 中的一个有界子集，而任何有界子集都是足够大的 $\square^{(n)}(2K)$ 一个子集是也。

连通性：设 $\{\varphi_i(t), 1 \leq i \leq n\}$ 乃是定义于 $[a, b]$ 上的 n 个连续函数，则参数式

$$x_i = \varphi_i(t), \quad 1 \leq i \leq n$$

所描述者乃是 \mathbb{R}^n 中连结 $A(\varphi_1(a), \dots, \varphi_n(a))$ 和 $B(\varphi_1(b), \dots, \varphi_n(b))$ 的一条连续曲线。

【定义】：若 \mathbb{R}^n 中的子集 S ，其中任给两点 A, B 皆能有完全包含在 S 之内的连续曲线连结之，则称之为连通子集。

极限封闭性： \mathbb{R}^n 中的子集 S ，若满足其中任给收敛点列的极限点依然是 S 中之点，则称 S 为 \mathbb{R}^n 中的一个闭子集 (close set)。

例如上述 $\square^{(n)}(2K)$ 是一个闭子集，但将其中任何一点略去，则就不再是闭子集了。

相对于闭子集，下述开子集 (open set) 也是在分析学中极为基本的概念，即

【定义】：对于给定点 $A(a_1, \dots, a_n)$ 和 $\delta > 0$,

$$U_\delta(A) := \{(x_1, \dots, x_n); |x_i - a_i| < \delta, 1 \leq i \leq n\}$$

叫做 A 点的 δ -邻域 (δ -neighborhood)。

【定义】： \mathbb{R}^n 中的子集 \mathcal{S} ，若对于其中任给一点 A 皆有其 (足够小的) δ -邻域 $U_\delta(A) \subset \mathcal{S}$ ，则称 \mathcal{S} 为一开子集。

注意：空集合应该看做开子集 (或闭子集) 的特例。

令

$$\square_0^{(n)}(2K) := \{(x_1, \dots, x_n); |x_i| < K, 1 \leq i \leq n\}$$

则不难验证 $\square_0^{(n)}(2K)$ 乃是 \mathbb{R}^n 中的开子集。

【习题】：

- (1) 任给一组闭子集的交集也是一个闭子集，而任给一组有限个开子集的交集也是一个开子集，试证之。
- (2) 任给一组开子集的和集也是一个开子集，而任给一组有限个闭子集的和集也是一个闭子集，试证之。
- (3) 试证一个开子集的补集乃是一个闭子集。
- (4) 试证一个闭子集的补集乃是一个开子集。
- (5) 令 $f(x, y)$ 为定义于 \mathbb{R}^2 的二元函数：

$$f(x, y) = \begin{cases} \frac{xy}{\sqrt{x^2 + y^2}}, & (x, y) \neq (0, 0) \\ 0, & (x, y) = (0, 0) \end{cases}$$

试验证 $f(x, y)$ 在 $(0, 0)$ 点连续。

- (6) 令 $f(x, y)$ 为定义于 \mathbb{R}^2 的二元函数：

$$f(x, y) = \begin{cases} \frac{xy}{x^2 + y^2}, & (x, y) \neq (0, 0) \\ 0, & (x, y) = (0, 0) \end{cases}$$

试验证 $f(x, y)$ 在 $(0, 0)$ 点不连续。[试把 $f(x, y)$ 局限于变域 $x = y$ ，并求当 $x \rightarrow 0$ 时其极限。]

(6) 令 $f(x, y)$ 为定义于 \mathbb{R}^2 的二元函数：

$$f(x, y) = \begin{cases} \frac{xy^2}{x^2 + y^4}, & (x, y) \neq (0, 0) \\ 0, & (x, y) = (0, 0) \end{cases}$$

问 $f(x, y)$ 在 $(0, 0)$ 点是否连续？

(8) 设 $f(x_1, \dots, x_n)$ 和 $g(x_1, \dots, x_n)$ 是两个同一变域 D 上到处连续的函数，试证

$$\begin{aligned} (f + g)(x_1, \dots, x_n) &:= f(x_1, \dots, x_n) + g(x_1, \dots, x_n) \\ (f \cdot g)(x_1, \dots, x_n) &:= f(x_1, \dots, x_n) \cdot g(x_1, \dots, x_n) \end{aligned}$$

也是 D 上到处连续的函数。

(9) 试证任给 n 元多项式函数都是 \mathbb{R}^n 上到处连续的函数。

(10) 设 $f(x_1, \dots, x_n)$ 是一个连通的区域 D 上到处连续的函数。 A, B 是 D 中任给两点， c 是介于 $f(A)$ 和 $f(B)$ 之间的值。试证 D 中总是存在一点 P 使得 $f(P) = c$ 。 [单元函数中间值定理之推广]

(11) 设 D 是 \mathbb{R}^n 中一个有界、连通闭子集而 $f(x_1, \dots, x_n)$ 是 D 上到处连续的函数。试证其函数值所构成者乃是 \mathbb{R} 中的一个闭线段（亦即 \mathbb{R}^1 中的有界、连通闭子集是也）。 [提示：参考上册第一章中 [定理 1.2] 和代数基本定理的证明之中 $|f(z)|^2$ 在 $\square(2K)$ 上的极小值之存在性之论证。]

(12) 设

$$u_i = f_i(x_1, \dots, x_n), \quad 1 \leq i \leq m$$

都是 D 上的连续函数，而 $y = g(u_1, \dots, u_m)$ 则是 \mathbb{R}^m 上的连续函数。试证复合函数

$$y = g(f_1(\mathbf{x}), f_2(\mathbf{x}), \dots, f_m(\mathbf{x})), \quad \mathbf{x} = (x_1, \dots, x_n)$$

也是 D 上的连续函数。

(13) (思考题) 试问应该如何定义由 \mathbb{R}^n 中的一个区域 D 到 \mathbb{R}^m 中的一个区域 D' 的映射的连续性？在适当的定义之下，连续映射的组合应该还是连续映射，试说明之。

1.2 多元函数的微分

归根究底，一个单元函数 $y = f(x)$ 在 $x = a$ 「可微」的实质就是在该点的微小 δ -邻域上， $f(x)$ 具有一个线性逼近 (linear approximation) $\ell(x)$ (亦即 $f(x) - \ell(x)$ 在足够小的 δ -邻域上是一个高于一阶的微量)，它就是 $f(a) + f'(a)(x - a)$ ，而 $y = f(a) + f'(a)(x - a)$ 也就是 $y = f(x)$ 在 $(a, f(a))$ 点的切线方程式。由此可以想到，一个多元函数 $y = f(x_1, \dots, x_n)$ 在 $A(a_1, \dots, a_n)$ 点可微的定义也就是它在 A 点的一个足够小的 δ -邻域， $U_\delta(A)$ ，上具有一个线性逼近，亦即

【定义】：若存在一个线性函数

$$\ell(x_1, \dots, x_n) = f(a_1, \dots, a_n) + \sum_{i=1}^n c_i(x_i - a_i)$$

使得

$$\frac{1}{\sqrt{\sum_{i=1}^n (x_i - a_i)^2}} \left| f(x_1, \dots, x_n) - \ell(x_1, \dots, x_n) \right|, \quad |x_i - a_i| < \delta$$

在 $\delta \rightarrow 0$ 时的极限值为 0，则称 $f(x_1, \dots, x_n)$ 在 $A(a_1, \dots, a_n)$ 点是可微的 (differentiable)。

[注]：若 f 在 A 点可微，则 f 必须在该点连续，其证明留作习题。

当我们把上述可微性的条件式局限到 $x_i = a_i, 2 \leq i \leq n$ ，的特殊情形之下，它就简化成单元函数 $f(x_1, a_2, \dots, a_n)$ 在 $x_1 = a_1$ 点的可微的条件式。再者，由此易见上述线性逼近中的系数 c_1 必须等于 $f(x_1, a_2, \dots, a_n)$ 在 $x_1 = a_1$ 点的变率。同理 c_2 必须等于 $f(a_1, x_2, a_3, \dots, a_n)$ 在 $x_2 = a_2$ 点的变率， \dots ， c_n 必须等于 $f(a_1, \dots, a_{n-1}, x_n)$ 在 $x_n = a_n$ 点的变率。

偏微分的定义与符号：设 x_i 的单元函数 $f(a_1, \dots, a_{i-1}, x_i, a_{i+1}, \dots, a_n)$ 在 $x_i = a_i$ 点可微，则其在该点的变率定义为函数 $f(x_1, \dots, x_n)$ 在 $A(a_1, \dots, a_n)$ 点对于 x_i 的偏导数 (partial derivative)，将以符号 $\left. \frac{\partial f}{\partial x_i} \right|_A$ 记号之，亦即

$$\left. \frac{\partial f}{\partial x_i} \right|_A = \lim_{x_i \rightarrow a_i} \frac{f(\dots, x_i, \dots) - f(\dots, a_i, \dots)}{x_i - a_i}$$

总结上述简短的讨论，即有下述可微性的一个必要条件，即

【引理1】：若 $f(x_1, \dots, x_n)$ 在 $A(a_1, \dots, a_n)$ 点可微，则它在 $A(a_1, \dots, a_n)$ 点对于 $x_i, 1 \leq i \leq n$ ，的偏导数皆存在（亦即其定义的极限存在），而且它在 A 点邻近的线性逼近就是

$$\ell(x_1, \dots, x_n) = f(a_1, \dots, a_n) + \sum_{i=1}^n \left. \frac{\partial f}{\partial x_i} \right|_A (x_i - a_i)$$

[注]：(i) 一般来说，上述 n 个偏导数的存在，只是可微性的必要条件（并非充要条件，见下面的例子）！

(ii) 把偏导数 $\left. \frac{\partial f}{\partial x_i} \right|_A$ 有定义之点的值逐一纪录，即得另一个 n 元函数（其定义域可能要比 $f(x_1, \dots, x_n)$ 的变域要小些）称之为 f 对于 x_i 的偏导函数，将以符号 $\frac{\partial f}{\partial x_i}$ 记之。

(iii) 由一个给定函数 $f(x_1, \dots, x_n)$ 去求它的各个偏导函数 $\frac{\partial f}{\partial x_i}$ 的计算叫做偏微分 (partial differentiation)，它本质上其实就是单元函数的微分，只是在计算 $\frac{\partial f}{\partial x_i}$ 时，除了 x_i 之外的 $(n-1)$ 个变元都要当做固定不变的常数来看待。总之，偏微分运算实乃原先业已熟习的单元函数的微分，亦即那个只让 x_i 变动而其他 $(n-1)$ 个变元则都暂且固定不动，如此简化而得的单元函数的微分是也。读者只要稍作练习，即可充分掌握。

【例子】：令

$$f(x, y) = \begin{cases} \frac{xy}{x^2 + y^2} & (x, y) \neq (0, 0) \\ 0 & (x, y) = (0, 0) \end{cases}$$

则由偏导函数的定义即可求得

$$\begin{aligned} \left. \frac{\partial f}{\partial x} \right|_{(0,0)} &= \lim_{h \rightarrow 0} \frac{f(h, 0) - f(0, 0)}{h} = \lim_{h \rightarrow 0} \frac{0 - 0}{h} = 0 \\ \left. \frac{\partial f}{\partial y} \right|_{(0,0)} &= \lim_{h' \rightarrow 0} \frac{f(0, h') - f(0, 0)}{h'} = \lim_{h' \rightarrow 0} \frac{0 - 0}{h'} = 0 \end{aligned}$$

但是 $f(x, y)$ 在 $(0, 0)$ 点不可微，因为它根本在 $(0, 0)$ 点不连续！由此可见，若只是要求偏导函数的存在是不足以用来有效研讨问题的。若要能够有效研讨问题，就必须加上偏导函数的连续性！

在单元微积分中，当导函数（或高阶导函数）也具有连续性时，均值（或高阶均值）定理就成为十分好用的有力工具。其实，在多元微积分中，偏导函数的连续性就变得更加重要和必要了。有鉴于此，在往後的讨论中，除非另加申明，我们总是设所涉及的函数的定义域是一个开子集 D ，而且它所涉及的偏导函数（或高阶偏导函数）也都是 D 上的连续函数。首先，偏导函数的连续性足以保证原给函数的可微性，即：

【定理 1.1】：设 $f(x_1, \dots, x_n)$ 的偏导函数 $\frac{\partial f}{\partial x_i}$, $1 \leq i \leq n$, 都在 A 点的一个 δ -邻域 $U_\delta(A)$ 上是连续的，则 $f(x_1, \dots, x_n)$ 在 A 点是可微的，其线性逼近为

$$\ell(x_1, \dots, x_n) = f(a_1, \dots, a_n) + \sum_{i=1}^n \frac{\partial f}{\partial x_i} \Big|_A (x_i - a_i)$$

证明：先证 $n=2$ 的情形，而一般情形的证法其实是和 $n=2$ 者完全相同（只是叙述和符号上稍加繁复）。用单元的均值定理，即得

$$\begin{aligned} & f(x_1, x_2) - f(a_1, a_2) \\ &= \left(f(x_1, x_2) - f(a_1, x_2) \right) + \left(f(a_1, x_2) - f(a_1, a_2) \right) \\ &= \frac{\partial f}{\partial x_1}(\xi_1, x_2)(x_1 - a_1) + \frac{\partial f}{\partial x_2}(a_1, \xi_2)(x_2 - a_2) \end{aligned}$$

再由所设之 $\frac{\partial f}{\partial x_1}$ 和 $\frac{\partial f}{\partial x_2}$ 的连续性，即有

$$f(x_1, x_2) - f(a_1, a_2) = \left(\frac{\partial f}{\partial x_1} \Big|_A + \varepsilon_1 \right) (x_1 - a_1) + \left(\frac{\partial f}{\partial x_2} \Big|_A + \varepsilon_2 \right) (x_2 - a_2)$$

而且当 $(x_1, x_2) \rightarrow (a_1, a_2)$ 时， $\varepsilon_1, \varepsilon_2 \rightarrow 0$ 。由此易见

$$\begin{aligned} & \frac{\left| f(x_1, x_2) - f(a_1, a_2) - \frac{\partial f}{\partial x_1} \Big|_A (x_1 - a_1) - \frac{\partial f}{\partial x_2} \Big|_A (x_2 - a_2) \right|}{\sqrt{(x_1 - a_1)^2 + (x_2 - a_2)^2}} \\ & \leq |\varepsilon_1| + |\varepsilon_2| \rightarrow 0 \end{aligned}$$

至此不难看出，上述证法是可以直截了当地推广到 n 是一般的情形的

◦ 先将 $f(x_1, \dots, x_n) - f(a_1, \dots, a_n)$ 改写成 n 个差之和，即

$$\begin{aligned} & \left(f(x_1, x_2, \dots, x_n) - f(a_1, x_2, \dots, x_n) \right) \\ & + \left(f(a_1, x_2, x_3, \dots, x_n) - f(a_1, a_2, x_3, \dots, x_n) \right) \\ & + \dots \dots \dots \\ & + \left(f(a_1, \dots, a_{n-1}, x_n) - f(a_1, \dots, a_{n-1}, a_n) \right) \end{aligned}$$

然後再对于每个差运用单元的均值定理和各个偏导函数的连续性即可证得。□

全微分和 Leibnitz 符号：

关于多元函数的微分，当年 Leibnitz 所采用的符号既简洁又好用，这也就是现在大家所通用者：

当 $x_i, 1 \leq i \leq n$, 是自变元时，我们用 dx_i 表示它的一个任给的微小「增量」，亦即设想自变元 x_i 由 a_i 改变为 $a_i + dx_i$ 。在上述变动之下， $f(x_1, \dots, x_n)$ 所相应的「增量」乃是：

$$\Delta f = f(a_1 + dx_1, \dots, a_n + dx_n) - f(a_1, \dots, a_n)$$

在 $f(x_1, \dots, x_n)$ 满足上述[定理 1.1]的条件时，则 Δf 有其局部线性逼近，我们将以符号 $df|_A$ 表示它，即

$$df|_A = \sum_{i=1}^n \left. \frac{\partial f}{\partial x_i} \right|_A dx_i$$

称之为 f 在 A 点的全微分 (total differential of f at A)。

在本质上，每个 dx_i 乃是一个取值微小的自变元，所以上述 $df|_A$ 乃是 $\{dx_i, 1 \leq i \leq n\}$ 这 n 个微小的自变元的一个齐线性函数。再者， A 乃是 $f(x_1, \dots, x_n)$ 的变域 D (设其为开子集而 $\frac{\partial f}{\partial x_i}, 1 \leq i \leq n$, 都在 D 上到处连续) 任给一点。由此可见，上述全微分当 A 点在 D 中任意变动时，实乃下述 $2n$ 个变元的函数，即 $\{x_i\}$ 和 $\{dx_i\}$

$$df = \sum_{i=1}^n \frac{\partial f}{\partial x_i} dx_i$$

它对于 $\{dx_i\}$ 这 n 个微小变元来说，总是齐线性的；它也就是一个 D 上到处可微的函数 $f(x_1, \dots, x_n)$ 在各处的局部线性逼近的总体表述。

【例子】：求 $f(x, y) = e^{ax} \cos by$ 的全微分。

$$\begin{aligned} df &= \left(\frac{\partial f}{\partial x}\right)dx + \left(\frac{\partial f}{\partial y}\right)dy \\ &= (ae^{ax} \cos by)dx - (be^{ax} \sin by)dy \end{aligned}$$

接著，让我们来看一看复合函数的全微分：

设 $f(u_1, \dots, u_m)$ 是 $u_j, 1 \leq j \leq m$, 的函数而每个 u_j 则是 n 个自变元 $x_i, 1 \leq i \leq n$, 的函数，即 $u_j = g_j(x_1, \dots, x_n)$ ，而且它们都满足上述定理的条件，亦即 $\frac{\partial f}{\partial u_j}, \frac{\partial g_j}{\partial x_i}$ 都是在一个开的定义域 D 上到处连续者。

在 D 中任取一点 (x_1, \dots, x_n) 及其任给邻近一点 $(x_1 + dx_1, \dots, x_n + dx_n)$ ，令

$$\Delta u_j = g_j(x_1 + dx_1, \dots, x_n + dx_n) - g_j(x_1, \dots, x_n), \quad 1 \leq j \leq m$$

$$\Delta f = f(u_1 + \Delta u_1, \dots, u_m + \Delta u_m) - f(u_1, \dots, u_m)$$

其中 $u_j = g_j(x_1, \dots, x_n), 1 \leq j \leq m$ 。和[定理 1.1]的证明中同样的计算，即可用均值定理和所涉及的偏导函数的连续性得出

$$\begin{aligned} \Delta f &= \sum_{j=1}^m \left(\frac{\partial f}{\partial u_j}(u_1, \dots, u_m) + \varepsilon_j \right) \Delta u_j \\ \Delta u_j &= \sum_{i=1}^n \left(\frac{\partial g_j}{\partial x_i}(x_1, \dots, x_n) + \delta_{i,j} \right) dx_i \end{aligned}$$

其中 ε_j 和 $\delta_{i,j}$ 都是在 $\rho = \sqrt{\sum_{i=1}^n dx_i^2} \rightarrow 0$ 之下也 $\rightarrow 0$ 者也。由此可见 Δf 在 (x_1, \dots, x_n) 的局部线性逼近乃是

$$\begin{aligned} df &= \sum_{j=1}^m \left(\frac{\partial f}{\partial u_j}(u_1, \dots, u_m) \sum_{i=1}^n \frac{\partial g_j}{\partial x_i}(x_1, \dots, x_n) dx_i \right) \\ &= \sum_{j=1}^m \frac{\partial f}{\partial u_j} du_j \end{aligned}$$

所以，在算式上来看，一个（一阶）连续可微的函数 $f(u_1, \dots, u_m)$ 的全微分，不论其变元是自变元或是因变元，恒有

$$(*) \quad df = \sum_{j=1}^m \frac{\partial f}{\partial u_j} du_j$$

[注]：把上式中以 $du_j = \sum_{i=1}^n \frac{\partial u_j}{\partial x_i} dx_i$ 代入，即得

$$df = \sum_{i=1}^n \left(\sum_{j=1}^m \frac{\partial f}{\partial u_j} \frac{\partial u_j}{\partial x_i} \right) dx_i$$

由此可见复合函数 $f(g_1(x_1, \dots, x_n), \dots, g_m(x_1, \dots, x_n))$ 对于 x_i 的偏导函数 $\frac{\partial f}{\partial x_i}$ 也就是上述表式中 dx_i 的系数，亦即

$$\frac{\partial f}{\partial x_i} = \sum_{j=1}^m \frac{\partial f}{\partial u_j} \frac{\partial u_j}{\partial x_i}$$

这也就是对于复合函数求其偏导函数时的 Chain Rule。总之，上述（连续）可微函数一以贯之的全微分公式业已自然而然地把 Chain Rule 包含于其中。再者，在多元数理分析中，全微分才是真正有用的主角，偏导函数的计算乃是求全微分的表式中的系数函数者也。而且在实际的计算中，只要逐次去求所涉及的变元的全微分（而它们都可以用（*）-式一以贯之），然後只要再用展开、集项等简单的代数运算（亦即分配律）即可得其所求。

【例子】：令 $f = f(x, y)$, $x = r \cos \theta$, $y = r \sin \theta$, 其中 $\{r, \theta\}$ 是相互无关的自变元。则有

$$\begin{aligned} \frac{\partial f}{\partial r} &= \frac{\partial f}{\partial x} \frac{\partial x}{\partial r} + \frac{\partial f}{\partial y} \frac{\partial y}{\partial r} = \frac{\partial f}{\partial x} (\cos \theta) + \frac{\partial f}{\partial y} (\sin \theta) \\ \frac{\partial f}{\partial \theta} &= \frac{\partial f}{\partial x} \frac{\partial x}{\partial \theta} + \frac{\partial f}{\partial y} \frac{\partial y}{\partial \theta} = \frac{\partial f}{\partial x} (-r \sin \theta) + \frac{\partial f}{\partial y} (r \cos \theta) \end{aligned}$$

由上述等式即可解出

$$\begin{aligned} \frac{\partial f}{\partial x} &= \frac{\partial f}{\partial r} (\cos \theta) + \frac{\partial f}{\partial \theta} \left(-\frac{1}{r} \sin \theta\right) \\ \frac{\partial f}{\partial y} &= \frac{\partial f}{\partial r} (\sin \theta) + \frac{\partial f}{\partial \theta} \left(\frac{1}{r} \cos \theta\right) \end{aligned}$$

方向导数 (directional derivative) :

假如我们把 $f(x_1, \dots, x_n)$ 的自变元限制到其变域 D 中的一条可微参数曲线 γ , 亦即

$$\gamma: \{x_i = \varphi_i(t), 1 \leq i \leq n\}$$

则有 $dx_i = \varphi_i'(t)dt$

$$df = \sum_{i=1}^n \frac{\partial f}{\partial x_i} \Big|_{\gamma} \varphi_i'(t)dt$$

亦即复合函数 $f(\varphi_1(t), \dots, \varphi_n(t))$ 对于单元参变数 t 的导数乃是

$$\frac{df}{dt} = \sum_{i=1}^n \frac{\partial f}{\partial x_i} \Big|_{\gamma} \cdot \varphi_i'(t)$$

再者, 设 $\{a_i = \varphi_i(0), 1 \leq i \leq n\}$ 是 γ 的起始点坐标, 则

$$\frac{df}{dt} \Big|_{t=0} = \sum_{i=1}^n \frac{\partial f}{\partial x_i} \Big|_A \varphi_i'(0)$$

叫做 $f(x_1, \dots, x_n)$ 在 A 点对于 γ 的初速方向的方向导数, 仅仅和 γ 的初速方向 $\mathbf{v} = (\varphi_1'(0), \dots, \varphi_n'(0))$ 有关。我们将以 $D_{\mathbf{v}}f|_A$ 表示之, 亦即

$$D_{\mathbf{v}}f|_A = \sum_{i=1}^n \frac{\partial f}{\partial x_i} \Big|_A v_i$$

其中 v_i 乃是在 A 点的速度向量 \mathbf{v} 的分量。

[注]: f 在 A 点的可微性也是必须的。例如, 令

$$f(x, y) = \begin{cases} \frac{xy^2}{x^2 + y^4} & (x, y) \neq (0, 0) \\ 0 & (x, y) = (0, 0) \end{cases}$$

则易证 $\frac{\partial f}{\partial x} \Big|_{(0,0)} = 0 = \frac{\partial f}{\partial y} \Big|_{(0,0)}$ 。但是当 $\varphi_1(t) = t, \varphi_2(t) = t$ 时, $(\varphi_1'(0), \varphi_2'(0)) = (1, 1)$ 而且

$$\frac{df(\varphi_1(t), \varphi_2(t))}{dt} \Big|_{t=0} = \lim_{t \rightarrow 0} \frac{\frac{t^3}{t^2+t^4} - 0}{t} = 1$$

显然在这个情况

$$\left. \frac{df(\varphi_1(t), \varphi_2(t))}{dt} \right|_{t=0} \neq \left. \frac{\partial f}{\partial x} \right|_{(0,0)} \varphi_1'(0) + \left. \frac{\partial f}{\partial y} \right|_{(0,0)} \varphi_2'(0)$$

【习题】：

- (1) 试证若 f 在 A 点可微，则 f 必需在 A 点连续。
- (2) 试验证上一节习题 (5) 中的连续函数 $f(x, y)$ 在 $(0, 0)$ 点不可微。
- (3) 求下述函数的全微分

(i) $f(x, y, z) = x^2 + y^2 + z^2 + 2xy + 2xz + 2yz$

(ii) $f(x, y, z) = \log(x^2 + y^2 + z^2), \quad x^2 + y^2 + z^2 > 0$

(iii) $f(x, y) = \sin x \cos y$

(iv) $f(x, y) = e^{(x^2+y^2)} \sin(x^2 + y^2)$

(4) 令 $f = f(ax + by)$ ，试证 $b \frac{\partial f}{\partial x} - a \frac{\partial f}{\partial y} = 0$ 。

(5) 设 $\varphi(t)$ 为一阶连续可微函数。令 $f(x, y) = x^n \varphi\left(\frac{y}{x^2}\right)$ ，试验证

$$x \frac{\partial f}{\partial x} + 2y \frac{\partial f}{\partial y} = n \cdot f$$

多元的泰勒公式与局部高阶逼近：

在单元微积分中，泰勒公式说明了一个局部高阶连续可微的单元函数具有多项式函数的局部高阶逼近。同样的，一个高阶连续可微的多元函数也具有其多元多项式函数的局部高阶逼近。这也就是我们接着所要研讨者——泰勒公式的多元推广。它也就是多元微分学中的高阶均值定理。

【定义】：设某一给定多元函数 $f(x_1, \dots, x_n)$ ，不但其偏导函数 $\left\{ \frac{\partial f}{\partial x_i} \right\}$ 都是在其（开的）定义域 D 上到处连续，而且它们的偏导函数 $\frac{\partial^2 f}{\partial x_j \partial x_i} = \frac{\partial}{\partial x_j} \left(\frac{\partial f}{\partial x_i} \right)$, $1 \leq i, j \leq n$ ，也都是 D 上到处连续，则称 f 为 D 上二阶连

续可微的函数。再者，若所有二阶偏导函数 $\{\frac{\partial^2 f}{\partial x_j \partial x_i}\}$ 也都是在 D 上到处连续可微，则称 f 为 D 上三阶连续可微的函数。以此逐步推进，若 f 的所有 k -阶偏导函数在 D 上都是到处连续，则称 f 为 D 上 k -阶连续可微的函数。

首先，当 $i \neq j$ 时，二阶偏导函数 $\frac{\partial^2 f}{\partial x_j \partial x_i}$ 和 $\frac{\partial^2 f}{\partial x_i \partial x_j}$ 是有可能在某些点两者都有其确定值但是彼此相异者也（见下面例子）。但是在 $\frac{\partial^2 f}{\partial x_j \partial x_i}$ 和 $\frac{\partial^2 f}{\partial x_i \partial x_j}$ 都是到处连续的情形，则它们在每点之值恒相等，亦即两者乃是 D 上的同一个函数，即下述引理

【引理 2】：若 $\frac{\partial^2 f}{\partial x_j \partial x_i}$ 和 $\frac{\partial^2 f}{\partial x_i \partial x_j}$ 在 D 上都是到处连续，则它们乃是 D 上的同一个函数。

证明：因为在求偏导函数 $\frac{\partial^2 f}{\partial x_j \partial x_i}$ 和 $\frac{\partial^2 f}{\partial x_i \partial x_j}$ 时，除了 x_i 和 x_j 之外的其他变元都是暂且固定不变的，所以上述引理的证明显然可以归于 $n = 2$ 的情形验证之，而 $\{i, j\}$ 则为 $\{1, 2\}$ 。兹证之如下：

对于 D 中任给一点 (x_1, x_2) 和足够小的 (h, k) ，由均值定理，即得

$$\begin{aligned} & f(x_1 + h, x_2 + k) - f(x_1 + h, x_2) - f(x_1, x_2 + k) + f(x_1, x_2) \\ &= h\phi'(x_1 + \theta h) \quad [\phi(x) = f(x, x_2 + k) - f(x, x_2)] \\ &= h\left[\frac{\partial f}{\partial x_1}(x_1 + \theta h, x_2 + k) - \frac{\partial f}{\partial x_1}(x_1 + \theta h, x_2)\right] \\ &= hk\frac{\partial^2 f}{\partial x_2 \partial x_1}(x_1 + \theta h, x_2 + \theta' k), \quad 0 \leq \theta, \theta' \leq 1 \end{aligned}$$

同理也可以说明存在适当的 $0 \leq \theta_1, \theta'_1 \leq 1$ 使得

$$\begin{aligned} & f(x_1 + h, x_2 + k) - f(x_1 + h, x_2) - f(x_1, x_2 + k) + f(x_1, x_2) \\ &= hk\frac{\partial^2 f}{\partial x_1 \partial x_2}(x_1 + \theta_1 h, x_2 + \theta'_1 k) \end{aligned}$$

亦即存在适当的 $0 \leq \theta, \theta', \theta_1, \theta'_1 \leq 1$ 使得

$$\frac{\partial^2 f}{\partial x_2 \partial x_1}(x_1 + \theta h, x_2 + \theta' k) = \frac{\partial^2 f}{\partial x_1 \partial x_2}(x_1 + \theta_1 h, x_2 + \theta'_1 k)$$

由此可见，当 $\frac{\partial^2 f}{\partial x_2 \partial x_1}$ 和 $\frac{\partial^2 f}{\partial x_1 \partial x_2}$ 连续的情形

$$\frac{\partial^2 f}{\partial x_2 \partial x_1} = \frac{\partial^2 f}{\partial x_1 \partial x_2}$$

因为它们分别是前述等式两端在 $h, k \rightarrow 0$ 之下的极限值。 \square

[注]: 在[引理 2]中二阶偏导数函数的连续性的条件是需要。试考虑下述定义于 \mathbb{R}^2 的函数 $f(x, y)$:

$$f(x, y) = \begin{cases} \frac{xy(x^2 - y^2)}{x^2 + y^2}, & (x, y) \neq (0, 0) \\ 0, & (x, y) = (0, 0) \end{cases}$$

则当 $(x, y) \neq (0, 0)$ 时, 易求得:

$$\begin{aligned} \frac{\partial f}{\partial x} \Big|_{(x,y)} &= \frac{x^4 y + 4x^2 y^3 - y^5}{(x^2 + y^2)^2} \\ \frac{\partial f}{\partial y} \Big|_{(x,y)} &= \frac{x^5 - 4x^3 y^2 - x^4 y}{(x^2 + y^2)^2} \end{aligned}$$

另一方面, 当 $(x, y) = (0, 0)$ 时, 由偏导数的定义极限式易求得 $\frac{\partial f}{\partial x} \Big|_{(0,0)} = 0 = \frac{\partial f}{\partial x} \Big|_{(0,0)}$ 。由此可见

$$\frac{\partial^2 f}{\partial x \partial y} \Big|_{(0,0)} = \frac{\partial}{\partial x} \left(\frac{\partial f}{\partial y} \right) \Big|_{(0,0)} = \lim_{h \rightarrow 0} \frac{\frac{\partial f}{\partial y} \Big|_{(h,0)} - \frac{\partial f}{\partial y} \Big|_{(0,0)}}{h} = \lim_{h \rightarrow 0} \frac{h^5/h^4}{h} = 1$$

类似地, 亦可算得

$$\frac{\partial^2 f}{\partial y \partial x} \Big|_{(0,0)} = -1$$

所以一般来说 $\frac{\partial^2 f}{\partial y \partial x} \neq \frac{\partial^2 f}{\partial x \partial y}$ 。

【例子】: 设 $f(x, y)$ 为二阶连续可微函数。令 $x = r \cos \theta$, $y = r \sin \theta$, 其中 r, θ 为相互无关的独立自变元。试验证

$$\frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} = \frac{\partial^2 f}{\partial r^2} + \frac{1}{r} \frac{\partial f}{\partial r} + \frac{1}{r^2} \frac{\partial^2 f}{\partial \theta^2}$$

解: 先求一阶偏导函数 $\frac{\partial f}{\partial r}, \frac{\partial f}{\partial \theta}$:

$$\begin{aligned} \frac{\partial f}{\partial r} &= \frac{\partial f}{\partial x} \cos \theta + \frac{\partial f}{\partial y} \sin \theta \\ \frac{\partial f}{\partial \theta} &= -\frac{\partial f}{\partial x} y + \frac{\partial f}{\partial y} x \end{aligned}$$

再求二阶偏导函数

$$\begin{aligned}\frac{\partial^2 f}{\partial r^2} &= \frac{\partial}{\partial x} \left(\frac{\partial f}{\partial r} \right) \frac{\partial x}{\partial r} + \frac{\partial}{\partial y} \left(\frac{\partial f}{\partial r} \right) \frac{\partial y}{\partial r} \\ &= \left\{ \frac{\partial^2 f}{\partial x^2} \cos \theta + \frac{\partial f}{\partial x} \frac{\partial(\cos \theta)}{\partial x} + \frac{\partial^2 f}{\partial x \partial y} \sin \theta + \frac{\partial f}{\partial y} \frac{\partial(\sin \theta)}{\partial x} \right\} \frac{\partial x}{\partial r} \\ &\quad + \left\{ \frac{\partial^2 f}{\partial y \partial x} \cos \theta + \frac{\partial f}{\partial y} \frac{\partial(\cos \theta)}{\partial y} + \frac{\partial^2 f}{\partial y^2} \sin \theta + \frac{\partial f}{\partial y} \frac{\partial(\sin \theta)}{\partial y} \right\} \frac{\partial y}{\partial r} \\ &= \frac{\partial^2 f}{\partial x^2} \cos^2 \theta + \frac{\partial^2 f}{\partial y^2} \sin^2 \theta + 2 \frac{\partial^2 f}{\partial x \partial y} \sin \theta \cos \theta\end{aligned}$$

其中用上了 $\frac{\partial^2 f}{\partial x \partial y} = \frac{\partial^2 f}{\partial y \partial x}$ 和

$$\begin{aligned}\frac{\partial(\cos \theta)}{\partial x} \frac{\partial x}{\partial r} + \frac{\partial(\cos \theta)}{\partial y} \frac{\partial y}{\partial r} &= \frac{\partial(\cos \theta)}{\partial r} = 0 \\ \frac{\partial(\sin \theta)}{\partial x} \frac{\partial x}{\partial r} + \frac{\partial(\sin \theta)}{\partial y} \frac{\partial y}{\partial r} &= \frac{\partial(\sin \theta)}{\partial r} = 0\end{aligned}$$

同理

$$\begin{aligned}\frac{\partial^2 f}{\partial \theta^2} &= \frac{\partial}{\partial x} \left(\frac{\partial f}{\partial \theta} \right) \frac{\partial x}{\partial \theta} + \frac{\partial}{\partial y} \left(\frac{\partial f}{\partial \theta} \right) \frac{\partial y}{\partial \theta} \\ &= \left\{ -\frac{\partial^2 f}{\partial x^2} y + \frac{\partial^2 f}{\partial x \partial y} x + \frac{\partial f}{\partial y} \right\} \frac{\partial x}{\partial \theta} + \left\{ -\frac{\partial^2 f}{\partial y \partial x} y - \frac{\partial f}{\partial x} + \frac{\partial^2 f}{\partial^2 y} x \right\} \frac{\partial y}{\partial \theta} \\ &= \frac{\partial^2 f}{\partial x^2} (r^2 \sin^2 \theta) + \frac{\partial^2 f}{\partial y^2} (r^2 \cos^2 \theta) - 2 \frac{\partial^2 f}{\partial x \partial y} r^2 \sin \theta \cos \theta \\ &\quad + \frac{\partial f}{\partial y} (-r \sin \theta) - \frac{\partial f}{\partial x} (r \cos \theta)\end{aligned}$$

由此可见

$$\begin{aligned}\frac{\partial^2 f}{\partial r^2} + \frac{1}{r^2} \frac{\partial^2 f}{\partial \theta^2} &= \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} - \frac{1}{r} \left(\frac{\partial f}{\partial y} \sin \theta + \frac{\partial f}{\partial x} \cos \theta \right) \\ &= \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} - \frac{1}{r} \frac{\partial f}{\partial r}\end{aligned}$$

即得所求证者。

设 $f(x_1, \dots, x_n)$ 是一个开子集 D 上的 k -阶连续可微函数, $A(a_1, \dots, a_n)$ 和 $B(a_1 + dx_1, \dots, a_n + dx_n)$ 是其中取定的邻近两点, 而且连结 A, B 两点的直线段完全位于 D 之内, 亦即对于所有 $0 \leq t \leq 1$,

$$(a_1 + tdx_1, \dots, a_n + tdx_n) \in D$$

令 $F(t) = f(a_1 + tdx_1, \dots, a_n + tdx_n)$ (因为 a_i 和 dx_i 都是取定之常数, 所以它乃是一个 t 的单元函数), 不难看到 $F(t)$ 是在 $[0, 1]$ 上 k -阶连续可微的。所以, 我们可以直接用单元的泰勒公式得出

$$\begin{aligned} f(a_1 + dx_1, \dots, a_n + dx_n) &= F(1) \\ &= F(0) + F'(0) + \frac{1}{2!}F''(0) + \dots + \frac{1}{(k-1)!}F^{(k-1)}(0) + \frac{1}{k!}F^{(k)}(0) \end{aligned}$$

由此可见, 我们只要去计算 $F'(0), \dots, F^{(k-1)}(0)$ 和 $F^{(k)}(0)$ 即可得出多元 k -阶连续可微函数的泰勒公式, 例如

$$F'(0) = \sum_{i=1}^n \left. \frac{\partial f}{\partial x_i} \right|_A dx_i = df|_A$$

而 $F''(0)$ 等等的逐步计算其本身乃是一个自然的好习题。

【习题】：

- (1) 设 $f(x, y, z)$ 为二阶连续可微函数。令 $x = \rho \sin \theta \cos \varphi, y = \rho \sin \theta \sin \varphi, z = \rho \cos \theta$, 其中 ρ, θ, φ 为相互无关的独立自变元。试验证

$$\frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} + \frac{\partial^2 f}{\partial z^2} = \frac{\partial^2 f}{\partial \rho^2} + \frac{2}{\rho} \frac{\partial f}{\partial \rho} + \frac{1}{\rho^2} \frac{\partial^2 f}{\partial \theta^2} + \frac{\cos \theta}{\rho^2 \sin \theta} \frac{\partial f}{\partial \theta} + \frac{1}{\rho^2 \sin^2 \theta} \frac{\partial^2 f}{\partial \varphi^2}$$

[提示：先把自变元 (x, y, z) 转换为 (r, ϕ, z) , 其中 $x = r \cos \phi, y = r \sin \phi$, 由例题结论即得

$$\frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} + \frac{\partial^2 f}{\partial z^2} = \frac{\partial^2 f}{\partial r^2} + \frac{1}{r} \frac{\partial f}{\partial r} + \frac{1}{r^2} \frac{\partial^2 f}{\partial \phi^2} + \frac{\partial^2 f}{\partial z^2}$$

然後再用 $r = \rho \sin \theta, \phi = \varphi, z = \rho \cos \theta$ 转换上述等式右方。]

(2) 试验证在上面多元泰勒展开式中

$$F''(0) = \sum_{i,j=1}^n \frac{\partial^2 f}{\partial x_i \partial x_j} \Big|_A dx_i dx_j$$

(3) 试写出下列函数在 $(0, 0)$ 点的泰勒展开式 (直到二阶微量为止) :

(i) $f(x, y) = \sin(x + y)$

(ii) $f(x, y) = e^{xy}$

(iii) $f(x, y) = e^x \cos y$

第二章

多元多关系的微分

在上一章中，我们讨论了多元函数的连续性与微分。一个在某个开子集 D 上连续可微的函数 f ，它的全微分 df 简洁地总括了它在 D 上每点的局部线性逼近。这就是一个连续可微函数用微分法去分析之所得的总体和精要。再者全微分的计算用来用去就是

$$(*) \quad df = \sum_{i=1}^m \frac{\partial f}{\partial u_i} du_i$$

的逐步运用和用代数分配律的展开、集项，是十分简朴、易算、好用的。它是多元微分的初步与基础。

但是在各种各样数理分析中所遇到的问题，通常乃是多元而且多关系的，亦即所涉及的数理体系含有多个参变量，而这些参变量之间又具有多个函数关系相互关连著。例如三角形的三边边长和三内角角度及其正弦、余弦定律；一个球面上各点的三个坐标和它们所满足的方程式等等。本章将进而研讨这种多元、多关系的数理体系在微分分析上的基础理论。其要点在于隐函数定理、坐标变换和定义在这种体系上的极值问题。

2.1 隐函数定理

设有 n 个变元 $\mathbf{x} = (x_1, x_2, \dots, x_n)$ 在 \mathbf{x}_0 点之邻域之内满足一组函数关系：

$$(*) \quad \{f_i(\mathbf{x}) = 0, 1 \leq i \leq m < n\}, \quad (f_i \text{ 一阶连续可微})$$

试问在什麼条件之下存在著上述方程组的局部「解函数组」？它将其中的 m 个变元（不妨设为 x_1, \dots, x_m ）表达成其余 $(n-m)$ 个变元（不妨设为 x_{m+1}, \dots, x_n ）的连续函数，亦即存在连续函数

$$x_i = \varphi_i(x_{m+1}, \dots, x_n), \quad \varphi_i(x_{m+1,0}, \dots, x_{n,0}) = x_{i,0}$$

使得

$$f_i(\varphi_1, \dots, \varphi_m, x_{m+1}, \dots, x_n), \quad 1 \leq i \leq m$$

在 $(x_{m+1,0}, \dots, x_{n,0})$ 的适当邻域内恒等于 0。

在 $f_i(\mathbf{x}) = 0, 1 \leq i \leq m$, 都是线性方程这种最为简单的情形，亦即 $\frac{\partial f_i}{\partial x_j}$ 都等于常数的情形，则显然有熟知的代数条件式，即当（常数行列式）

$$(2.1) \quad \begin{vmatrix} \frac{\partial f_1}{\partial x_1} & \cdots & \frac{\partial f_1}{\partial x_m} \\ \vdots & & \vdots \\ \frac{\partial f_m}{\partial x_1} & \cdots & \frac{\partial f_m}{\partial x_m} \end{vmatrix} \neq 0$$

时，存在唯一的线性解函数 $x_i = \ell_i(x_{m+1}, \dots, x_n), 1 \leq i \leq m$ 。往後我们将用 Jacobi 所创的符号以

$$(2.2) \quad \frac{\partial(f_1, \dots, f_m)}{\partial(x_1, \dots, x_m)} \text{ 表示 } \left| \frac{\partial f_i}{\partial x_j}(\mathbf{x}) \right|$$

$$\frac{\partial(f_1, \dots, f_m)}{\partial(x_1, \dots, x_m)} \Big|_{\mathbf{x}_0} \text{ 表示 } \left| \frac{\partial f_i}{\partial x_j}(\mathbf{x}_0) \right|$$

这类由一阶偏导函数组成的行列式通称之为 Jacobians。在一般的情形，若

$$(2.3) \quad \frac{\partial(f_1, \dots, f_m)}{\partial(x_1, \dots, x_m)} \Big|_{\mathbf{x}_0} \neq 0$$

则原方程组 $\{f_i = 0\}$ 在 \mathbf{x}_0 点的局部线性逼近方程组

$$(\star_0) \quad df_i \Big|_{\mathbf{x}_0} = \sum_{j=1}^n \frac{\partial f_i}{\partial x_j}(\mathbf{x}_0) dx_j, \quad dx_j = (x_j - x_{j,0})$$

具有下述唯一线性解函数

$$(2.4) \quad dx_i = \sum_{j=m+1}^n c_{ij} dx_j, \quad 1 \leq i \leq m$$

其中

$$(2.4') \quad c_{ij} = - \frac{\frac{\partial(f_1, \dots, f_m)}{\partial(\dots, x_j, \dots)} \Big|_{\mathbf{x}_0}}{\frac{\partial(f_1, \dots, f_m)}{\partial(x_1, \dots, x_m)} \Big|_{\mathbf{x}_0}}$$

而 $\partial(\dots, x_j, \dots)$ 表示将 x_i 改为 x_j 者。

【隐函数定理】：设 $\{f_i(\mathbf{x}), 1 \leq i \leq m\}$ 在 \mathbf{x}_0 的邻域皆为一阶连续可微，而且 $\frac{\partial(f_1, \dots, f_m)}{\partial(x_1, \dots, x_m)} \Big|_{\mathbf{x}_0} \neq 0$ ，则方程组

$$(\star) \quad f_i(\mathbf{x}) = 0, \quad 1 \leq i \leq m$$

存在一组唯一的连续（局部）解函数

$$(*) \quad x_i = \varphi_i(x_{m+1}, \dots, x_n), \quad \varphi_i(x_{m+1,0}, \dots, x_{n,0}) = x_{i,0}$$

使得

$$f_i(\varphi_1, \dots, \varphi_m, x_{m+1}, \dots, x_n), \quad 1 \leq i \leq m$$

在 $(x_{m+1,0}, \dots, x_{n,0})$ 的适当邻域内恒等于 0。再者， $\{\varphi_i, 1 \leq i \leq m\}$ 都是一阶连续可微者，而且

$$(*)' \quad d\varphi_i = - \sum_{j=m+1}^n \left(\frac{\frac{\partial(f_1, \dots, f_m)}{\partial(\dots, x_j, \dots)}}{\frac{\partial(f_1, \dots, f_m)}{\partial(x_1, \dots, x_m)}} \right) dx_j, \quad 1 \leq i \leq m$$

证明：我们将先证 $m = 1$ 的情形，然后归纳地论证其一般的情形。

$m = 1$ ：亦即由 $\frac{\partial f_1}{\partial x_1}(\mathbf{x}_0) \neq 0$ 论证 $x_1 = \psi_1(x_2, \dots, x_n)$ 之存在性。由所设 $f_1(\mathbf{x})$ 的一阶连续可微性和 $\frac{\partial f_1}{\partial x_1}(\mathbf{x}_0) \neq 0$ ，只要 $|\mathbf{x} - \mathbf{x}_0|$ 足够小，恒有

$$(2.5) \quad \left| \frac{\partial f_1}{\partial x_1}(\mathbf{x}) \right| > \frac{1}{2} \left| \frac{\partial f_1}{\partial x_1}(\mathbf{x}_0) \right| \quad (\text{令其为 } a)$$

$$\text{Max} \left\{ \left| \frac{\partial f_1}{\partial x_j}(\mathbf{x}) \right|, 2 \leq j \leq n \right\} < K$$

设 $\tilde{\mathbf{x}} = (x_2, \dots, x_n)$ 是 $\tilde{\mathbf{x}}_0 = (x_{2,0}, \dots, x_{n,0})$ 的足够小邻域中的取定一点，亦即 $|x_j - x_{j,0}| < \delta$, δ 足够小。由一阶均值定理，即有

$$(2.6) \quad \Delta f_1 = \frac{\partial f_1}{\partial x_1}(\xi_1)(x_1 - x_{1,0}) + \sum_{j=2}^n \frac{\partial f_1}{\partial x_j}(\xi_j)(x_j - x_{j,0})$$

若将 $(x_1 - x_{1,0})$ 分别取用 $\pm \varepsilon_0$ ，而将 δ 取成小于 $\frac{a\varepsilon_0}{(n-1)K}$ ，则由 (2.6)-式显然有

$$f_1(x_{1,0} + \varepsilon_0, \tilde{\mathbf{x}}) \quad \text{和} \quad f_1(x_{1,0} - \varepsilon_0, \tilde{\mathbf{x}}) \text{ 异号}$$

所以由中间值定理和单调性即有唯一的 $x_1 = \psi_1(\tilde{\mathbf{x}})$ 使得

$$f_1(\psi_1(\tilde{\mathbf{x}}), \tilde{\mathbf{x}}) = 0$$

再者，由 $\Delta f_1 = 0$ 易见 $\psi_1(\tilde{\mathbf{x}})$ 一阶可微而且

$$\frac{\partial f_1}{\partial x_1}(\psi_1(\tilde{\mathbf{x}}), \tilde{\mathbf{x}}) d\psi_1 + \sum_{j=2}^n \frac{\partial f_1}{\partial x_j}(\psi_1(\tilde{\mathbf{x}}), \tilde{\mathbf{x}}) dx_j = 0$$

$$\Rightarrow d\psi_1 = - \sum_{j=2}^n \frac{\frac{\partial f_1}{\partial x_j}(\psi_1(\tilde{\mathbf{x}}), \tilde{\mathbf{x}})}{\frac{\partial f_1}{\partial x_1}(\psi_1(\tilde{\mathbf{x}}), \tilde{\mathbf{x}})} dx_j$$

一般情形的归纳论证：

亦即归纳假设定理在 $m - 1$ 时业已成立，推论其在 m 时也成立。不妨设 $\frac{\partial f_1}{\partial x_1}(\mathbf{x}_0) \neq 0$ 。由上述之所证，存在 $\psi_1(\tilde{\mathbf{x}})$ 使得 $\psi_1(\tilde{\mathbf{x}}_0) = x_{1,0}$, $f_1(\psi_1(\tilde{\mathbf{x}}), \tilde{\mathbf{x}}) \equiv 0$ 。令

$$g_i(\tilde{\mathbf{x}}) = f_i(\psi_1(\tilde{\mathbf{x}}), \tilde{\mathbf{x}}), \quad 2 \leq i \leq m$$

即有

$$\frac{\partial g_i}{\partial x_j} = \frac{\partial f_i}{\partial x_1} \cdot \frac{\partial \psi_1}{\partial x_j} + \frac{\partial f_i}{\partial x_j} = \frac{\partial f_i}{\partial x_1} \left(-\frac{\frac{\partial f_1}{\partial x_j}}{\frac{\partial f_1}{\partial x_1}} \right) + \frac{\partial f_i}{\partial x_j}$$

由此可见

$$0 \neq \frac{\partial(f_1, f_2, \dots, f_m)}{\partial(x_1, x_2, \dots, x_m)} \Big|_{(\psi_1(\bar{x}_0), \bar{x}_0)} = \frac{\partial f_1}{\partial x_1} \cdot \frac{\partial(g_2, \dots, g_m)}{\partial(x_2, \dots, x_m)} \Big|_{(\psi_1(\bar{x}_0), \bar{x}_0)}$$

所以

$$\frac{\partial(g_2, \dots, g_m)}{\partial(x_2, \dots, x_m)} \Big|_{\bar{x}_0} \neq 0$$

由归纳假设，存在唯一的一组连续函数

$$\varphi_i(x_{m+1}, \dots, x_n), \quad \varphi_i(x_{m+1,0}, \dots, x_{n,0}) = x_{i,0}, \quad 2 \leq i \leq m$$

使得 $g_i(\varphi_2, \dots, \varphi_m, x_{m+1}, \dots, x_n) \equiv 0$ ，而且 φ_i 都是一阶连续可微者。

$$d\varphi_i = - \sum_{j=m+1}^n \left(\frac{\frac{\partial(g_2, \dots, g_m)}{\partial(\dots, x_j, \dots)}}{\frac{\partial(g_2, \dots, g_m)}{\partial(x_2, \dots, x_m)}} \right) dx_j$$

令 $\varphi_1 = \psi_1(\varphi_2, \dots, \varphi_m, x_{m+1}, \dots, x_n)$ ，不难验证 $\{\varphi_i, 1 \leq i \leq m\}$ 即为所求证者。 \square

注意：上述证明仅仅论证其解函数的唯一存在性。一般来说，除非在极为简单的方程组（例如线性方程组），这种唯一存在的解函数是无法求得其明确表达式 (explicit expressions) 的。所以它乃是由方程组所唯一确定的隐函数 (implicit functions)。但是这组隐函数的线性逼近式却又是具有 (\star') 这种简明的表达式者。这里，再一次说明微分法（亦即线性逼近）的妙处与用场。

总之，隐函数定理乃是一个概念性的基本定理。让我们再用几何观点来对于隐函数定理的内含作一剖析：

(i) 设 $f_i(\mathbf{x})$ 是开子集 D 上的连续可微函数，则条件式 $f_i(\mathbf{x}) = 0$ 所描述者，乃是 D 中的一个 $(n-1)$ -维子集，通常称之为超曲面 (hypersurface)。

设 \mathbf{x}_0 是其上一点 (即有 $f_i(\mathbf{x}_0) = 0$) 而且 $\left\{ \frac{\partial f_i}{\partial x_j} \Big|_{\mathbf{x}_0}, 1 \leq j \leq n \right\}$ 之中至少有一非零, 则

$$0 = df_i \Big|_{\mathbf{x}_0} = \sum_{j=1}^n \frac{\partial f_i}{\partial x_j} \Big|_{\mathbf{x}_0} dx_j, \quad dx_j = (x_j - x_{j,0})$$

乃是 \mathbb{R}^n 中过 \mathbf{x}_0 点和向量 $(\frac{\partial f_i}{\partial x_1} \Big|_{\mathbf{x}_0}, \frac{\partial f_i}{\partial x_2} \Big|_{\mathbf{x}_0}, \dots, \frac{\partial f_i}{\partial x_n} \Big|_{\mathbf{x}_0})$ 垂直的那个超平面, 它就是 $f_i(\mathbf{x}) = 0$ 所描述的超曲面的切平面是也。

[注]: 通常我们把上述法向量叫做 f_i 在 \mathbf{x}_0 的梯度向量 (the gradient vector of f_i at \mathbf{x}_0), 并以符号 $\nabla f_i \Big|_{\mathbf{x}_0}$ 表示之。

(ii) 隐函数定理的条件式, 亦即存在有 $1 \leq i_1 < i_2 < \dots < i_m \leq n$ 使得

$$\frac{\partial(f_1, \dots, f_m)}{\partial(x_{i_1}, \dots, x_{i_m})} \Big|_{\mathbf{x}_0} \neq 0$$

其几何意义乃是它们的梯度向量 $\{\nabla f_i \Big|_{\mathbf{x}_0}, 1 \leq i \leq m\}$ 是线性无关的 (linearly independent)。所以它们在 \mathbf{x}_0 点的切平面的交集乃是一个 $(n-m)$ -维的平面, 亦即那个过 \mathbf{x}_0 点而且和 $\{\nabla f_i \Big|_{\mathbf{x}_0}, 1 \leq i \leq m\}$ 都垂直的 $(n-m)$ -维平面 [通常把这种交截叫做超曲面 $\{f_i(\mathbf{x}) = 0, 1 \leq i \leq m\}$ 在 \mathbf{x}_0 点横截 (transversal intersection)]。

(iii) 隐函数定理的几何意义就是证明了: 若 m 个超曲面互相横截于 \mathbf{x}_0 , 则其交集在 \mathbf{x}_0 点的邻近乃是一个 $(n-m)$ -维曲面, 而且可以选用适当的 $(n-m)$ 个坐标为参数, 把其余的 m 个坐标表达为前者的函数 (亦即所证之隐函数)。

(iv) 往後我们称这种横截的交点 \mathbf{x}_0 为这种多元多关系所描述的子集的规则点 (regular point), 隐函数定理保证了规则点的局部具有上述简单的参数表达。其所用的 $(n-m)$ 个坐标也就可以作为所给子集在规则点邻近的数理分析的自变元。

2.2 坐标变换

设 $\{f_i(\mathbf{x}), 1 \leq i \leq n\}$ 是 n 个一阶连续可微的 n 元函数。若

$$\frac{\partial(f_1, \dots, f_n)}{\partial(x_1, \dots, x_n)} \Big|_{\mathbf{x}} \neq 0$$

对于所有 $|\mathbf{x} - \mathbf{x}_0| < \delta$ 之点皆成立，令

$$y_i = f_i(\mathbf{x}), \quad 1 \leq i \leq n$$

而且把它们看成下述 n 个 $2n$ 变元之方程组，即

$$F_i(\mathbf{x}, \mathbf{y}) = f_i(\mathbf{x}) - y_i = 0, \quad 1 \leq i \leq n$$

则有

$$\frac{\partial(F_1, \dots, F_n)}{\partial(x_1, \dots, x_n)} = \frac{\partial(f_1, \dots, f_n)}{\partial(x_1, \dots, x_n)} \Big|_{\mathbf{x}} \neq 0$$

由隐函数定理，可知唯一存在 \mathbf{y} 的函数组

$$x_i = \varphi_i(\mathbf{y})$$

使得 $F_i(\varphi_1, \dots, \varphi_n, y_1, \dots, y_n) = f_i(\varphi_1, \dots, \varphi_n) - y_i \equiv 0$

亦即

$$\{x_i = \varphi_i(\mathbf{y})\} \longleftrightarrow \{y_i = f_i(\mathbf{x})\}$$

是相互确定的互逆函数关系。再者，不难验证

$$\frac{\partial(\varphi_1, \dots, \varphi_n)}{\partial(y_1, \dots, y_n)} \Big|_{\mathbf{y}} = \left\{ \frac{\partial(f_1, \dots, f_n)}{\partial(x_1, \dots, x_n)} \Big|_{\mathbf{x}} \right\}^{-1}$$

【例一】：极坐标 (r, θ) 与笛氏坐标 (x, y)

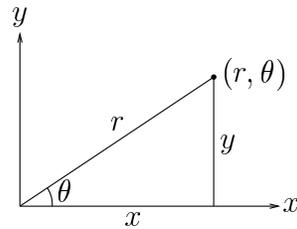
$$x = r \cos \theta$$

$$y = r \sin \theta$$

$$dx = \cos \theta dr - r \sin \theta d\theta$$

$$dy = \sin \theta dr + r \cos \theta d\theta$$

$$\frac{\partial(x, y)}{\partial(r, \theta)} = \begin{vmatrix} \cos \theta & -r \sin \theta \\ \sin \theta & r \cos \theta \end{vmatrix} = r$$

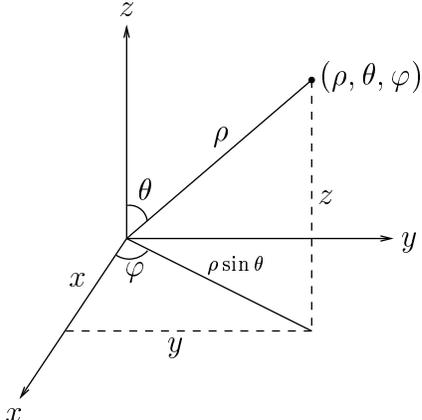


[图 2-1]

【例二】：球面坐标 (ρ, θ, φ) 与笛氏坐标 (x, y, z)

$$\begin{aligned} x &= \rho \sin \theta \cos \varphi \\ y &= \rho \sin \theta \sin \varphi \\ z &= \rho \cos \theta \end{aligned}$$

$$\frac{\partial(x, y, z)}{\partial(\rho, \theta, \varphi)} = \begin{vmatrix} \sin \theta \cos \varphi & \rho \cos \theta \cos \varphi & -\rho \sin \theta \sin \varphi \\ \sin \theta \sin \varphi & \rho \cos \theta \sin \varphi & \rho \sin \theta \cos \varphi \\ \cos \theta & -\rho \sin \theta & 0 \end{vmatrix}$$

$$= \rho^2 \sin \theta$$


[图 2-2]

【例三】：设 (u_1, \dots, u_n) , (y_1, \dots, y_n) 和 (x_1, \dots, x_n) 是三种可以相互变换的局部坐标系，亦即

$$\frac{\partial(u_1, \dots, u_n)}{\partial(y_1, \dots, y_n)} \neq 0, \quad \frac{\partial(y_1, \dots, y_n)}{\partial(x_1, \dots, x_n)} \neq 0$$

(在某一邻域中到处不为零)，则有

$$\frac{\partial(u_1, \dots, u_n)}{\partial(x_1, \dots, x_n)} = \frac{\partial(u_1, \dots, u_n)}{\partial(y_1, \dots, y_n)} \cdot \frac{\partial(y_1, \dots, y_n)}{\partial(x_1, \dots, x_n)}$$

验证：

$$du_i = \sum_{j=1}^n \frac{\partial u_i}{\partial y_j} dy_j$$

$$dy_j = \sum_{k=1}^n \frac{\partial y_j}{\partial x_k} dx_k$$

将 dy_j 的表式代入 du_i 的表式中，即得

$$du_i = \sum_{k=1}^n \left(\sum_{j=1}^n \frac{\partial u_i}{\partial y_j} \cdot \frac{\partial y_j}{\partial x_k} \right) dx_k$$

由此即有

$$\left(\frac{\partial u_i}{\partial x_k}\right) = \left(\frac{\partial u_i}{\partial y_j}\right) \cdot \left(\frac{\partial y_j}{\partial x_k}\right)$$

所以由行列式乘法公式即得

$$\begin{aligned} \frac{\partial(u_1, \dots, u_n)}{\partial(x_1, \dots, x_n)} &= \left| \frac{\partial u_i}{\partial x_k} \right| = \left| \frac{\partial u_i}{\partial y_j} \right| \cdot \left| \frac{\partial y_j}{\partial x_k} \right| \\ &= \frac{\partial(u_1, \dots, u_n)}{\partial(y_1, \dots, y_n)} \cdot \frac{\partial(y_1, \dots, y_n)}{\partial(x_1, \dots, x_n)} \end{aligned}$$

2.3 极大, 极小的微分条件式

设 $y = f(\mathbf{x})$ 在 \mathbf{x}_0 点是局部极大 (或极小), 即

$$f(\mathbf{x}_0) \geq f(\mathbf{x}) \quad (\text{或 } f(\mathbf{x}_0) \leq f(\mathbf{x}))$$

对于所有 $|\mathbf{x} - \mathbf{x}_0| < \varepsilon$ (适当小的正数) 皆成立。若 $f(\mathbf{x})$ 在上述邻域一阶可微, 则显然有下述必要条件, 即

$$\frac{\partial f}{\partial x_i}(\mathbf{x}_0) = 0, \quad 1 \leq i \leq n$$

再者, 若 $f(\mathbf{x})$ 在上述邻域二阶可微而且二次型

$$\sum_{i,j=1}^n \frac{\partial^2 f}{\partial x_i \partial x_j}(\mathbf{x}_0) \xi_i \xi_j$$

是正定 (或负定) 的, 即在 ξ_i 不全为零时恒取正 (或负) 值, 则 $f(\mathbf{x}_0)$ 乃是局部极小 (或极大) 值。

上述简明的结果乃是 Taylor 均值定理的直接推论。

现在让我们进而研究一个一阶连续可微函数 $y = f(\mathbf{x})$ 其中变元 $\mathbf{x} = (x_1, \dots, x_n)$ 在 \mathbf{x}_0 邻近并非自由变动, 而是局限于某些条件式之下的极大、极小的必要条件。设所给的附加条件式是

$$(*) \quad g_i(\mathbf{x}) = 0, \quad 1 \leq i \leq m$$

而且存在 x_{i_1}, \dots, x_{i_m} 使得

$$\frac{\partial(g_1, \dots, g_m)}{\partial(x_{i_1}, \dots, x_{i_m})} \Big|_{\mathbf{x}_0} \neq 0 \quad (\text{为了便于叙述, 不妨设 } \frac{\partial(g_1, \dots, g_m)}{\partial(x_1, \dots, x_m)} \Big|_{\mathbf{x}_0} \neq 0)$$

一种自然的想法是用隐函数定理把上述问题归于原先那种 不含有 局限条件的情形, 其具体做法如下:

令 $x_i = \varphi_i(x_{m+1}, \dots, x_n)$, $1 \leq i \leq m$, 是由局限条件式 $\{g_i(\mathbf{x}) = 0, 1 \leq i \leq m\}$ 所唯一确定的 隐函数。用来代入 $y = f(\mathbf{x})$, 即得

$$y = f(\varphi_1, \dots, \varphi_m, x_{m+1}, \dots, x_n) = F(x_{m+1}, \dots, x_n)$$

它乃是 $\tilde{\mathbf{x}} = (x_{m+1}, \dots, x_n)$ 在 $\tilde{\mathbf{x}}_0 = (x_{m+1,0}, \dots, x_{n,0})$ 邻近 $(n-m)$ 个自由变元的函数, 所以它在 $\tilde{\mathbf{x}}_0$ 点取极大或极小的必要条件乃是

$$(\star_0) \quad \frac{\partial F}{\partial x_j}(\tilde{\mathbf{x}}_0) = 0, \quad m+1 \leq j \leq n$$

亦即

$$(\star) \quad \frac{\partial F}{\partial x_j}(\tilde{\mathbf{x}}_0) = \sum_{i=1}^m \frac{\partial f}{\partial x_i}(\mathbf{x}_0) \frac{\partial \varphi_i}{\partial x_j}(\tilde{\mathbf{x}}_0) + \frac{\partial f}{\partial x_j}(\mathbf{x}_0) = 0, \quad m+1 \leq j \leq n$$

而其中 $\{\frac{\partial \varphi_i}{\partial x_j}(\tilde{\mathbf{x}}_0), 1 \leq i \leq m\}$ 则是下述线性方程组所唯一确定者也, 即

$$(\star') \quad \frac{\partial g_k}{\partial x_j}(\mathbf{x}_0) = \sum_{i=1}^n \frac{\partial g_k}{\partial x_i}(\mathbf{x}_0) \frac{\partial \varphi_i}{\partial x_j}(\tilde{\mathbf{x}}_0) + \frac{\partial g_k}{\partial x_j}(\mathbf{x}_0) = 0, \quad 1 \leq k \leq m$$

让我们先来看一下 $m=1$ 的情形, 则 (\star') 即为

$$\begin{aligned} & \frac{\partial g}{\partial x_1}(\mathbf{x}_0) \frac{\partial \varphi}{\partial x_j}(\tilde{\mathbf{x}}_0) + \frac{\partial g}{\partial x_j}(\mathbf{x}_0) = 0 \\ \Rightarrow & \frac{\partial \varphi}{\partial x_j}(\tilde{\mathbf{x}}_0) = -\frac{\partial g}{\partial x_j}(\mathbf{x}_0) / \frac{\partial g}{\partial x_1}(\mathbf{x}_0) \end{aligned}$$

以它代入 (\star) 即为

$$-\frac{\partial f}{\partial x_1}(\mathbf{x}_0) \cdot \frac{\partial g}{\partial x_j}(\mathbf{x}_0) / \frac{\partial g}{\partial x_1}(\mathbf{x}_0) + \frac{\partial f}{\partial x_j}(\mathbf{x}_0) = 0, \quad 2 \leq j \leq n$$

Lagrange 指出，上述代数条件式可以想成是下述含有参数 λ 的代数方程组

$$(\#) \quad \frac{\partial f}{\partial x_i}(\mathbf{x}_0) + \lambda \frac{\partial g}{\partial x_i}(\mathbf{x}_0) = 0, \quad 1 \leq i \leq n$$

用第一式解得 λ ，代入其余 $(n-1)$ 式者也。其实，在一般的情形，我们可以引入 m 个参数 $\{\lambda_i, 1 \leq i \leq m\}$ ，则下述方程组

$$(\#) \quad \frac{\partial}{\partial x_i}(f + \lambda_1 g_1 + \dots + \lambda_m g_m) = \frac{\partial f}{\partial x_i} + \sum_{k=1}^m \lambda_k \frac{\partial g_k}{\partial x_i} = 0, \quad 1 \leq i \leq n$$

可以用前面 m 式解得 $\{\lambda_k, 1 \leq k \leq m\}$ （因为 $\frac{\partial(g_1, \dots, g_m)}{\partial(x_1, \dots, x_m)}(\mathbf{x}_0) \neq 0$ ），代入其余 $(n-m)$ 式之所得也就是 (\star) 。亦即方程组 $(\#)$ 和方程组 $\{(\star), (\star')\}$ 其实是代数等价者也。这也就是 Lagrange's method of multipliers. 兹总结如下：

【定理】(Lagrange)：设 $f(\mathbf{x})$ 和 $\{g_k(\mathbf{x}), 1 \leq k \leq m\}$ 都是一阶连续可微，而且存在 (i_1, \dots, i_m) 使得

$$\frac{\partial(g_1, \dots, g_m)}{\partial(x_{i_1}, \dots, x_{i_m})}(\mathbf{x}_0) \neq 0$$

则 $f(\mathbf{x})$ 在局限条件 $\{g_k(\mathbf{x}) = 0, 1 \leq k \leq m\}$ 之下，在 \mathbf{x}_0 点取极大或极小值的必要条件是

$$(\#) \quad \begin{cases} \frac{\partial f}{\partial x_i}(\mathbf{x}_0) + \sum_{k=1}^m \lambda_k \frac{\partial g_k}{\partial x_i}(\mathbf{x}_0) = 0, & 1 \leq i \leq n \\ g_k(\mathbf{x}_0) = 0 & 1 \leq k \leq m \end{cases}$$

其中 $\{\lambda_k, 1 \leq k \leq m\}$ 是待定参数。

[一般来说，上述含有待定参数的条件式 $(\#)$ 总是比原先用隐函数的想法所得者 $\{(\star), (\star')\}$ 要来得简朴。]

下述是它的一个简单但是具有基本重要性的应用。

【例子】：设 (a_{ij}) 是一个对称方阵，亦即 $a_{ij} = a_{ji}$ 。令

$$f(\mathbf{x}) = (x_1, \dots, x_n)(a_{ij}) \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \sum_{i,j=1}^n a_{ij}x_ix_j$$

$$g(\mathbf{x}) = \sum_{i=1}^n x_i^2 - 1$$

则 $f(\mathbf{x})$ 在 $g(\mathbf{x}) = 0$ 的局限条件之下的极值问题的 Lagrange 必要条件式乃是

$$(\#) \quad \begin{cases} \frac{1}{2} \frac{\partial}{\partial x_i} (f - \lambda g) = \sum_{j=1}^n (a_{ij} - \lambda \delta_{ij}) x_j = \ell_i(\mathbf{x}) = 0, & 1 \leq i \leq n \\ \sum x_i^2 - 1 = 0 \end{cases}$$

其中 δ_{ij} 在 $i \neq j$ 时为零， $i = j$ 时为 1，而 λ 则是一个待定参数。因为 $\{(\#)_i, 1 \leq i \leq n\}$ 是齐线性的，所以只有在其系数行列式 $|a_{ij} - \lambda \delta_{ij}| = 0$ 时，才能有非零解，而 $(\#)_{n+1}$ 显然要求 $\mathbf{x} \neq 0$ 。由此可见方程组 $(\#)$ 中的待定参数 λ 必须是方阵 (a_{ij}) 的特征值，亦即

$$|a_{ij} - \lambda \delta_{ij}| = 0$$

的根。再者，设 λ 是一个特征值， \mathbf{x}_0 是它的一个单位长特征向量，亦即

$$\begin{cases} \ell_i(\mathbf{x}_0) = \sum_{j=1}^n (a_{ij} - \lambda \delta_{ij}) x_{j,0} = 0 \\ |\mathbf{x}_0|^2 = \sum_{j=1}^n x_{j,0}^2 = 1 \end{cases}$$

$$\Rightarrow \sum_{i=1}^n x_{i,0} \ell_i(\mathbf{x}_0) = \sum_{i,j=1}^n a_{ij} x_{i,0} x_{j,0} - \lambda \sum_{i=1}^n x_{i,0}^2 = 0$$

$$\Rightarrow f(\mathbf{x}_0) = \lambda$$

由此可见 $f(\mathbf{x})$ 在单位超球面上的极大（或极小）值乃是 (a_{ij}) 的特征值中的极大（或极小）者是也！而其所相应的单位特征向量也就是单位超球面上达成 $f(\mathbf{x}_0)$ 极大（或极小）值之点。

【推论】：二次型 $Q(\mathbf{x}) = \sum_{i,j=1}^n a_{ij}x_i x_j$ ($a_{ij} = a_{ji}$) 是正定（或负定，或不定）的充要条件是其特征值都是正的（或负的，或有正、有负的）。

【习题】：求解下列有局限条件的极值问题：

(1) 求解下列有局限条件的极值问题：

(i) $f(x, y) = x^2 + y^2$, 条件： $x^2 + 4y^2 = 4$

(ii) $f(x_1, \dots, x_n) = \sum_{i=1}^n a_i x_i$, 条件： $\sum_{i=1}^n x_i^p - 1 = 0, p \geq 1$

(2) 试用 (1)(ii) 之结果，证明当 $a_i, b_i, 1 \leq i \leq n$, 为非负实数， $p, q \geq 1$ 满足 $\frac{1}{p} + \frac{1}{q} = 1$ 时有

(i) $\sum_{i=1}^n a_i b_i \leq (\sum_{i=1}^n a_i^p)^{\frac{1}{p}} (\sum_{i=1}^n b_i^q)^{\frac{1}{q}}$

(ii) $\left\{ \sum_{i=1}^n (a_i + b_i)^p \right\}^{\frac{1}{p}} \leq (\sum_{i=1}^n a_i^p)^{\frac{1}{p}} + (\sum_{i=1}^n b_i^p)^{\frac{1}{p}}$

第三章

高维勾股定理与 Grassmann 代数

在多元、多关系数理体系的分析中，我们要用到各种各样的积分如线积分、面积分、体积分、多重积分等等，以及它们之间的相互关联。在它们的研讨中，当然要用到各种维数的定向体积 (oriented volumes of various dimensions) 的有效计算与相互关联。因此，高维度量几何中的基本定理——高维勾股定理——和把它们妥为组织而成的代数体系——Grassmann 代数——乃是多元积分理论的几何基础和精简好用的代数工具。

在各种各样的几何量中，长度乃是一切的基本，其他如角度、面积、体积等等都可以由所涉及的长度加以确定和计算。其实，这也就是定量几何学的基础理论的主要部分。在基础几何学的第五章中，我们把欧氏度量几何的基础理论，精简、优化成向量代数来加以表达，把几何的基本定理转化成向量运算的运算律。例如关于长度、角度的基本定理——勾股定理——就转化、优化成简洁、有力的内积分配律。本章将把该章对于三维欧氏几何的研讨，推广到一般的 n 维欧氏几何。

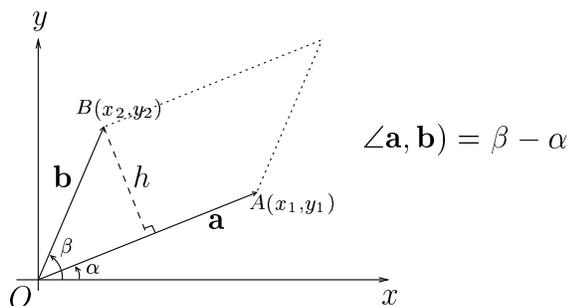
3.1 向量代数与平行体的有向体积

开宗明义，一个 n 维欧氏空间中的平移构成一个 \mathbb{R} 上的 n -维向量空间。其长度同样的满足勾股定理，所以也可以同样地定义内积，满足同样的运算律。本节将以这样一个具有内积的 \mathbb{R} 上 n 维向量空间为出发点，研讨各维平行体的有向体积 (oriented volume)。

3.1.1 平面的定向与平行四边形的有向面积

一条直线有两个相反的方向, 所以其定向乃是取定其一为正向, 而一个有向线段 \overrightarrow{AB} 的有向长度的正负则取决于它的方向和取定之正向的同异。在二维的情形, 一个平面上有两个相反的角度转向, 所以其定向乃是取定其一为正的转向, 而一个由其上两个向量 \mathbf{a}, \mathbf{b} 所张的平行四边形 $\text{//}(\mathbf{a}, \mathbf{b})$ 的有向面积的正、负则取决于由 \mathbf{a} 转向 \mathbf{b} 的角度 $\angle \mathbf{a}, \mathbf{b}$ 之正负。如 [图 3-1] 所示, $\text{//}(\mathbf{a}, \mathbf{b})$ 的有向面积:

$$(3.1) \quad A(\mathbf{a}, \mathbf{b}) = |\mathbf{a}| \cdot |\mathbf{b}| \cdot \sin \angle \mathbf{a}, \mathbf{b}$$



[图 3-1]

再者, 在平面上选取正向的笛氏坐标系, 令 $\mathbf{a} = \overrightarrow{OA}$ 和 $\mathbf{b} = \overrightarrow{OB}$ 的坐标 (亦即 A, B 的坐标) 分别是 (x_1, y_1) 和 (x_2, y_2) , 则有

$$(3.2) \quad \begin{aligned} x_1 &= |\mathbf{a}| \cos \alpha, & y_1 &= |\mathbf{a}| \sin \alpha \\ x_2 &= |\mathbf{b}| \cos \beta, & y_2 &= |\mathbf{b}| \sin \beta \\ A(\mathbf{a}, \mathbf{b}) &= |\mathbf{a}| |\mathbf{b}| \sin(\beta - \alpha) \\ &= |\mathbf{a}| |\mathbf{b}| (\sin \beta \cos \alpha - \cos \beta \sin \alpha) \\ &= x_1 y_2 - x_2 y_1 = \begin{vmatrix} x_1 & y_1 \\ x_2 & y_2 \end{vmatrix} \end{aligned}$$

令 $\mathbf{e}_x, \mathbf{e}_y$ 分别是 x, y 方向的单位长向量, 则 $\text{//}(\mathbf{e}_x, \mathbf{e}_y)$ 当然等于 1。所以上述公式还可以改写成下述形式, 即

$$(3.2') \quad A(\mathbf{a}, \mathbf{b}) \cdot A(\mathbf{e}_x, \mathbf{e}_y) = \begin{vmatrix} \mathbf{a} \cdot \mathbf{e}_x & \mathbf{a} \cdot \mathbf{e}_y \\ \mathbf{b} \cdot \mathbf{e}_x & \mathbf{b} \cdot \mathbf{e}_y \end{vmatrix}$$

上述公式是否可以推广成一般情形呢? 亦即

$$A(\mathbf{a}, \mathbf{b}) \cdot A(\mathbf{c}, \mathbf{d}) \stackrel{?}{=} \begin{vmatrix} \mathbf{a} \cdot \mathbf{c} & \mathbf{a} \cdot \mathbf{d} \\ \mathbf{b} \cdot \mathbf{c} & \mathbf{b} \cdot \mathbf{d} \end{vmatrix}$$

此事不难再用 (3.2)-式结合行列式的乘法公式验证之，即

$$\begin{aligned}
 A(\mathbf{a}, \mathbf{b}) \cdot A(\mathbf{c}, \mathbf{d}) &= \begin{vmatrix} x_1 & y_1 \\ x_2 & y_2 \end{vmatrix} \cdot \begin{vmatrix} u_1 & v_1 \\ u_2 & v_2 \end{vmatrix} \\
 (3.3) \qquad &= \begin{vmatrix} x_1 & y_1 \\ x_2 & y_2 \end{vmatrix} \cdot \begin{vmatrix} u_1 & u_2 \\ v_1 & v_2 \end{vmatrix} \\
 &= \begin{vmatrix} x_1 u_1 + y_1 v_1 & x_1 u_2 + y_1 v_2 \\ x_2 u_1 + y_2 v_1 & x_2 u_2 + y_2 v_2 \end{vmatrix} = \begin{vmatrix} \mathbf{a} \cdot \mathbf{c} & \mathbf{a} \cdot \mathbf{d} \\ \mathbf{b} \cdot \mathbf{c} & \mathbf{b} \cdot \mathbf{d} \end{vmatrix}
 \end{aligned}$$

3.1.2 三维空间的定向和平行六面体的有向体积

接著让我们进而探讨三维空间应该如何定向？易见在平面上两个同向的有序正交基总可以由其一旋转到其另一，但是异向的两个则不然。设 $\{\mathbf{a}, \mathbf{b}, \mathbf{c}\}$ 和 $\{\mathbf{a}', \mathbf{b}', \mathbf{c}'\}$ 是空间的任给两组有序正交基。易见我们可以通过一个以和 \mathbf{a}, \mathbf{a}' 都垂直的轴的适当旋转 (rotation) R_1 使得 $R_1(\mathbf{a}) = \mathbf{a}'$ 。因此 $\{R_1(\mathbf{b}), R_1(\mathbf{c})\}$ 和 $\{\mathbf{b}', \mathbf{c}'\}$ 乃是位于和 \mathbf{a}' 垂直的平面中的两组有序正交基。若它们同向，则有一个以 \mathbf{a}' 为轴的适当旋转 R_2 使得

$$R_2 R_1(\mathbf{a}) = R_2(\mathbf{a}') = \mathbf{a}', \quad R_2 R_1(\mathbf{b}) = \mathbf{b}', \quad R_2 R_1(\mathbf{c}) = \mathbf{c}'$$

亦即有旋转的组合 $R_2 R_1$ ，它把 $\{\mathbf{a}, \mathbf{b}, \mathbf{c}\}$ 分别映射到 $\{\mathbf{a}', \mathbf{b}', \mathbf{c}'\}$ 。由上述讨论不难看到，空间两组同是右手型（或同是左手型）的正交基是可以有旋转的组合相互映射的。但是各别是左、右手型的两组正交基则无法有这种相互映射。由此可见，三维空间的定向乃是在上述二类有序正交基中取定其一为正向者是也（例如通常约定取右手型者为正向）。

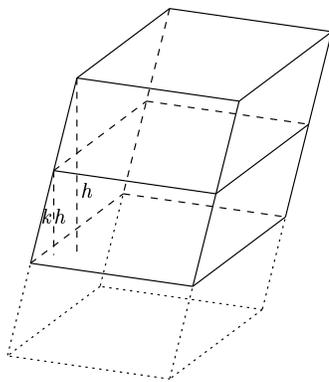
同样地，有序的三个线性无关的向量组也可以分成左手型和右手型，所以对于业已选定何者为正向型之後，一个有序的三个向量 $\{\mathbf{a}, \mathbf{b}, \mathbf{c}\}$ 所张的平行六面体（且以 $\|(\mathbf{a}, \mathbf{b}, \mathbf{c})\|$ 记之）的有向体积的正负也就取决于它是否和选定者同属一型。我们将以符号 $V(\mathbf{a}, \mathbf{b}, \mathbf{c})$ 表示 $\|(\mathbf{a}, \mathbf{b}, \mathbf{c})\|$ 的有向体积 (oriented volume)。不难验证 $V(\mathbf{a}, \mathbf{b}, \mathbf{c})$ 具有下述几何性质：

1. 当 $\{\mathbf{a}, \mathbf{b}, \mathbf{c}\}$ 是正向正交基时， $V(\mathbf{a}, \mathbf{b}, \mathbf{c}) = 1$ ；
2. $V(\mathbf{a}, \mathbf{b}, \mathbf{c})$ 是其向量变元的斜对称函数；
3. $V(k\mathbf{a}, \mathbf{b}, \mathbf{c}) = kV(\mathbf{a}, \mathbf{b}, \mathbf{c})$ ；

$$4. V(\mathbf{a} + \beta\mathbf{b} + \gamma\mathbf{c}, \mathbf{b}, \mathbf{c}) = V(\mathbf{a}, \mathbf{b}, \mathbf{c}).$$

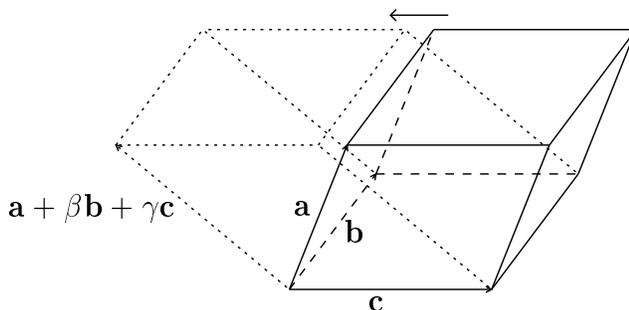
性质 1 是显然的。性质 2 则是因为定向的正负在两个向量变元对换下变为其反向。

如 [图 3-2] 所示，在 $k > 0$ (或 $k < 0$) 时 $\{k\mathbf{a}, \mathbf{b}, \mathbf{c}\}$ 和 $\{\mathbf{a}, \mathbf{b}, \mathbf{c}\}$ 同向 (或异向)；而且 $\|(k\mathbf{a}, \mathbf{b}, \mathbf{c})$ 和 $\|(\mathbf{a}, \mathbf{b}, \mathbf{c})$ 具有同样的底面 (亦即 $\|(\mathbf{b}, \mathbf{c})$) 而前者的高乃是後者的 $|k|$ -倍。



[图 3-2]

再者，如 [图 3-3] 所示， $\|(\mathbf{a} + \beta\mathbf{b} + \gamma\mathbf{c}, \mathbf{b}, \mathbf{c})$ 和 $\|(\mathbf{a}, \mathbf{b}, \mathbf{c})$ 的底面都是 $\|(\mathbf{b}, \mathbf{c})$ ，而前者的顶面乃是将後者的顶面在两者所共在的平面作一滑动，所以两者是同高而且同向者也。这也就说明了性质 4 (通常把上述变形叫做斜切 (shearing)，所以性质 4 也就叫做斜切不变性 (shearing invariance))。



[图 3-3]

[注]：在基础几何学的第五章中，我们业已用 \times -积说明了 $V(\mathbf{a}, \mathbf{b}, \mathbf{c}) = \mathbf{a} \cdot (\mathbf{b} \times \mathbf{c}) = \det(\mathbf{a}, \mathbf{b}, \mathbf{c})$ 。但是在这里我们将直接用上述有向体积的基本

性质去证明 $V(\mathbf{a}, \mathbf{b}, \mathbf{c}) = \det(\mathbf{a}, \mathbf{b}, \mathbf{c})$ 。因为这种证法可以直接推广到高维，证明高维有向体积也都是高阶行列式。其实，我们所要论证之点，仅仅是下列引理。

【引理 1】： $V(\mathbf{a}_1 + \mathbf{a}_2, \mathbf{b}, \mathbf{c}) = V(\mathbf{a}_1, \mathbf{b}, \mathbf{c}) + V(\mathbf{a}_2, \mathbf{b}, \mathbf{c})$

证明：不妨设 \mathbf{b}, \mathbf{c} 线性无关。因为在 \mathbf{b}, \mathbf{c} 线性相关的情形，上式三项都显然等于 0，所以 $0 = 0 + 0$ 当然成立。

令 \mathbf{e} 是垂直于 $\text{span}(\mathbf{b}, \mathbf{c})$ 的单位法向量， $\{\mathbf{e}, \mathbf{b}, \mathbf{c}\}$ 组成三维空间的一组基底。令

$$\mathbf{a}_1 = \alpha_1 \mathbf{e} + \beta_1 \mathbf{b} + \gamma_1 \mathbf{c}, \quad \mathbf{a}_2 = \alpha_2 \mathbf{e} + \beta_2 \mathbf{b} + \gamma_2 \mathbf{c}$$

由性质 3, 4 即有

$$\begin{aligned} V(\mathbf{a}_1 + \mathbf{a}_2, \mathbf{b}, \mathbf{c}) &= V((\alpha_1 + \alpha_2)\mathbf{e}, \mathbf{b}, \mathbf{c}) = (\alpha_1 + \alpha_2)V(\mathbf{e}, \mathbf{b}, \mathbf{c}) \\ V(\mathbf{a}_1, \mathbf{b}, \mathbf{c}) &= V(\alpha_1 \mathbf{e}, \mathbf{b}, \mathbf{c}) = \alpha_1 V(\mathbf{e}, \mathbf{b}, \mathbf{c}) \\ V(\mathbf{a}_2, \mathbf{b}, \mathbf{c}) &= V(\alpha_2 \mathbf{e}, \mathbf{b}, \mathbf{c}) = \alpha_2 V(\mathbf{e}, \mathbf{b}, \mathbf{c}) \end{aligned}$$

由此即得

$$V(\mathbf{a}_1 + \mathbf{a}_2, \mathbf{b}, \mathbf{c}) = V(\mathbf{a}_1, \mathbf{b}, \mathbf{c}) + V(\mathbf{a}_2, \mathbf{b}, \mathbf{c})$$

【推论 1】： $V(\mathbf{a}, \mathbf{b}, \mathbf{c}) = \det(\mathbf{a}, \mathbf{b}, \mathbf{c}) \cdot \det(\mathbf{e}_x, \mathbf{e}_y, \mathbf{e}_z) = \det(\mathbf{a}, \mathbf{b}, \mathbf{c})$

[因为 $V(\mathbf{a}, \mathbf{b}, \mathbf{c})$ 业已满足 $\det(\mathbf{a}, \mathbf{b}, \mathbf{c})$ 的所有特征性质。]

【推论 2】： $V(\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3) \cdot V(\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3)$

$$= \begin{vmatrix} \mathbf{a}_1 \cdot \mathbf{b}_1 & \mathbf{a}_1 \cdot \mathbf{b}_2 & \mathbf{a}_1 \cdot \mathbf{b}_3 \\ \mathbf{a}_2 \cdot \mathbf{b}_1 & \mathbf{a}_2 \cdot \mathbf{b}_2 & \mathbf{a}_2 \cdot \mathbf{b}_3 \\ \mathbf{a}_3 \cdot \mathbf{b}_1 & \mathbf{a}_3 \cdot \mathbf{b}_2 & \mathbf{a}_3 \cdot \mathbf{b}_3 \end{vmatrix}$$

现在让我们再来研讨一般的 k -维子空间中的定向和 k -维平行体的有向体积。由上面对于一、二、三维的讨论，我们可以看出，一个欧氏空间之中的「定向」和其保长变换之间的关系如下，即：

(i) 定向在最为基本的反射对称之下互换，所以定向乃是在那种由偶数个反射对称组合而成的保长变换之下保持者也；

(ii) 因为任何两个反射对称的组合不是平移就是旋转，所以易见由所有偶数个反射对称之复合所构成之子集（子群）乃是连通的 (connected)；

(iii) 一个 k -维欧氏空间中的所有正交坐标系和其中所有正交标架 (orthonormal frames) 显然成一、一对应。两个正交坐标系 (或正交标架) 之间存在著唯一的保长变换将其一映射到其另一。若这个唯一存在的保长变换乃是偶数 (或奇数) 个反射对称之复合, 则这两个正交坐标系 (亦即正交标架) 具有同样的 (相异的) 定向。

(iv) 不难验证 (用归纳法), 任给一个空间非正交的标架都可以用适当的斜切变形 (shearing deformations) 把它变形成为互相正交者, 然後只要把它们逐一乘以其长度之倒数, 即可变形成为一个正交标架。定向 (orientation) 乃是在上述这种变形之下保持者也。

总结上述几点, 可见一个 k -维欧氏空间的定向, 乃是在其上正交标架的两个等价类之中取定其一为正向, 其另一则为负向。再者, 非正交的标架之定向乃是等同于它在 (iv) 中描述之变形所得的正交标架之定向。由此即定义了一个定向的 k -维欧氏空间之中的 k -维平行体 $//(\mathbf{a}_1, \dots, \mathbf{a}_k)$ 的有向体积 (oriented k -dim volume), 我们将以 $V(\mathbf{a}_1, \dots, \mathbf{a}_k)$ 记之。再者, 这样定义的 $V(\mathbf{a}_1, \dots, \mathbf{a}_k)$ 显然也同样具有一如在三维情形的四个基本性质, 及其推论——[引理 1]。所以也有下述重要公式, 即

【定理 3.1】: 一个定向 k -维欧氏空间中的 k -维平行体 $//(\mathbf{a}_1, \dots, \mathbf{a}_k)$ 的有向体积 $V(\mathbf{a}_1, \dots, \mathbf{a}_k)$ 具有下述基本公式, 即

$$V(\mathbf{a}_1, \dots, \mathbf{a}_k) = \det(\mathbf{a}_1, \dots, \mathbf{a}_k) \quad (\text{亦即 } |\mathbf{a}_i \cdot \mathbf{e}_j|)$$

$$V(\mathbf{a}_1, \dots, \mathbf{a}_k) \cdot V(\mathbf{b}_1, \dots, \mathbf{b}_k) = |\mathbf{a}_i \cdot \mathbf{b}_j|$$

其中 $\{\mathbf{e}_1, \dots, \mathbf{e}_k\}$ 是一个任选正向正交标架。

上述[定理 3.1]业已在 $k = 1, 2, 3$ 时详为说明, 其归纳论证则留作习题。

3.2 向量内积与勾股定理的高维推广

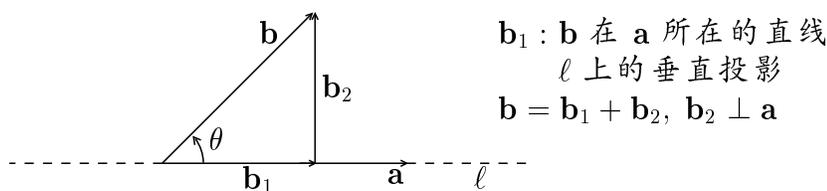
向量内积和勾股定理是密切相关的。内积的定义植根于勾股定理, 而内积分配律则是勾股定理的优化形式; 它是用向量代数研讨涉及长度、夹角的各种各样问题的至精至简。本节将探讨向量内积和勾股定

理在高维度量（亦即面积，体积和高维体积）上的推广。首先，让我们再来分析一下向量内积的定义及其几何本质。

我们在基础几何学的第五章讨论过，向量内积乃是从下述三角形三边边长的特定函数的深入分析来引入的：

$$f(\mathbf{a}, \mathbf{b}) = \frac{1}{2} \left\{ |\mathbf{a} + \mathbf{b}|^2 - |\mathbf{a}|^2 - |\mathbf{b}|^2 \right\}$$

亦即发现它乃是一个美好的双线性函数，所以顺理成章地把它定义为 \mathbf{a} , \mathbf{b} 的内积，则其双线性就自然转化成简洁易用的内积分配律了。再者，如 [图 3-4] 所示，将分配律和垂直投影相结合，即得向量内积的下述几何内含，即



[图 3-4]

$$\begin{aligned} \mathbf{a} \cdot \mathbf{b} &= \mathbf{a} \cdot (\mathbf{b}_1 + \mathbf{b}_2) = \mathbf{a} \cdot \mathbf{b}_1 + \mathbf{a} \cdot \mathbf{b}_2 \\ &= \mathbf{a} \cdot \mathbf{b}_1 = |\mathbf{a}| \cdot |\mathbf{b}| \cos \theta \end{aligned}$$

其实，我们也可以采用上述几何内含作为内积的定义式，然后再用垂直投影的「线性」来推导如此定义的内积的分配律。

总之，内积和垂直投影乃是两个密切相关的事物：内积的分配律相当于垂直投影的线性，由其一即可推导其另一。回顾在该章证明 $f(\mathbf{a}, \mathbf{b})$ 的双线性中，垂直投影在广义勾股定理的证明中扮演重要角色，但是并不需要用到垂直投影的线性。所以这种处理方式的好处是：它不依赖于垂直投影的线性，而且也给垂直投影的线性这个欧氏几何的要点提供了简洁的证明。这也就是采取这种处理方式的好处和原由。

【分析】：

(i) 当我们进而探索如何把内积推广到高维的情形，首先得认识到高维的情形和一维的情形具有下述两点差别：其一是对于两个给定 k -维平行体 $\parallel(\mathbf{a}_1, \dots, \mathbf{a}_k)$ 和 $\parallel(\mathbf{b}_1, \dots, \mathbf{b}_k)$ 是无法构造另一个 k -维平行体，可以自然地定义为它们的和 (sum)。由此可见，原先一维向量采用

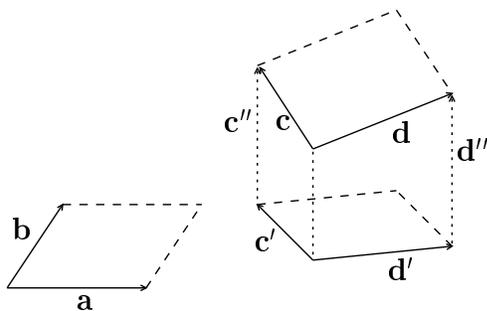
$\mathbf{a} \cdot \mathbf{b} := \frac{1}{2}\{|\mathbf{a} + \mathbf{b}|^2 - |\mathbf{a}|^2 - |\mathbf{b}|^2\}$ 来定义其内积的做法乃是此路不通者也。其二是垂直投影的线性业已由一维向量内积的分配律推导而得，在此正好把它用来作为定义高维向量内积的有力工具。基于上述两点，很自然要设法把 $\mathbf{a} \cdot \mathbf{b} = |\mathbf{a}| \cdot |\mathbf{b}| \cos \theta$ 这种定义方式推广到高维。

(ii) 有鉴于 $\mathbf{a} \cdot \mathbf{b} = |\mathbf{a}| \cdot |\mathbf{b}| \cos \theta$ 的实质内含乃是 \mathbf{a} 和 \mathbf{b}_1 这两个共线的有向线段的有向长度的乘积，其中 \mathbf{b}_1 则是 \mathbf{b} 在 \mathbf{a} 所在的直线上的垂直投影（参看 [图 3-4]）。设 $\parallel(\mathbf{a}, \mathbf{b})$ 和 $\parallel(\mathbf{c}, \mathbf{d})$ 是空间中两个给定的平行四边形。若它们共在某一平面 Π 之内，我们可以取定 Π 上的定向，则有两者的有向面积的乘积公式

$$A(\mathbf{a}, \mathbf{b}) \cdot A(\mathbf{c}, \mathbf{d}) = \begin{vmatrix} \mathbf{a} \cdot \mathbf{c} & \mathbf{a} \cdot \mathbf{d} \\ \mathbf{b} \cdot \mathbf{c} & \mathbf{b} \cdot \mathbf{d} \end{vmatrix}$$

它是和 Π 上定向的选取无关的。在两者不共面的情形，令 \mathbf{c}' 和 \mathbf{d}' 分别是 \mathbf{c}, \mathbf{d} 在 $\parallel(\mathbf{a}, \mathbf{b})$ 所在的平面 Π 上的垂直投影。如 [图 3-5] 所示，

$$\begin{aligned} \mathbf{c} &= \mathbf{c}' + \mathbf{c}'', & \mathbf{d} &= \mathbf{d}' + \mathbf{d}'' \\ \parallel(\mathbf{c}', \mathbf{d}') & \text{和 } \parallel(\mathbf{a}, \mathbf{b}) & \text{共在 } \Pi & \text{之内} \\ \mathbf{c}'' \text{ 和 } \mathbf{d}'' & \text{则和 } \Pi & \text{垂直, 亦即 } & \mathbf{c}'' \cdot \mathbf{a} = \mathbf{c}'' \cdot \mathbf{b} = \mathbf{d}'' \cdot \mathbf{a} = \mathbf{d}'' \cdot \mathbf{b} = 0 \end{aligned}$$



[图 3-5]

由此可见共面的 $\parallel(\mathbf{a}, \mathbf{b})$ 和 $\parallel(\mathbf{c}', \mathbf{d}')$ 的有向面积的乘积还是可以用同样的公式表达，即

$$\begin{vmatrix} \mathbf{a} \cdot \mathbf{c}' & \mathbf{a} \cdot \mathbf{d}' \\ \mathbf{b} \cdot \mathbf{c}' & \mathbf{b} \cdot \mathbf{d}' \end{vmatrix} = \begin{vmatrix} \mathbf{a} \cdot \mathbf{c} & \mathbf{a} \cdot \mathbf{d} \\ \mathbf{b} \cdot \mathbf{c} & \mathbf{b} \cdot \mathbf{d} \end{vmatrix}$$

因为有 $\mathbf{a} \cdot \mathbf{c} = \mathbf{a} \cdot \mathbf{c}' + \mathbf{a} \cdot \mathbf{c}'' = \mathbf{a} \cdot \mathbf{c}'$ 等等。

由此可见，我们应该把空间中两个平行四边形 $//(a, b)$ 和 $/(c, d)$ 的内积定义为共面的 $/(a, b)$ 和 $/(c', d')$ 的有向平面之乘积，其中 $/(c', d')$ 乃是 $/(c, d)$ 在 $/(a, b)$ 所在的平面上的垂直投影，将以符号 $\langle //(\mathbf{a}, \mathbf{b}), //(\mathbf{c}, \mathbf{d}) \rangle$ 记之。则有内积公式：

$$\langle //(\mathbf{a}, \mathbf{b}), //(\mathbf{c}, \mathbf{d}) \rangle = \begin{vmatrix} \mathbf{a} \cdot \mathbf{c} & \mathbf{a} \cdot \mathbf{d} \\ \mathbf{b} \cdot \mathbf{c} & \mathbf{b} \cdot \mathbf{d} \end{vmatrix}$$

它是对于四个向量变元都是线性的！

(iii) 我们可以把上述定义同样地推广到 k -维的情形，即有

【定义】：设 $/(a_1, \dots, a_k)$ 和 $/(b_1, \dots, b_k)$ 是两个任给的 k -维平行体， Π 是 $/(a_1, \dots, a_k)$ 所在的 k -维子空间， b'_i 分别是 b_i 在 Π 上的垂直投影， $1 \leq i \leq k$ 。则 $/(a_1, \dots, a_k)$ 和 $/(b'_1, \dots, b'_k)$ 乃是共在一个 k -维子空间 Π 上的 k -维平行体，两者的有向 k -维体积的乘积其实是和 Π 上的定向选取无关的（因为负负得正），定义为 $/(a_1, \dots, a_k)$ 和 $/(b_1, \dots, b_k)$ 的内积，将以符号 $\langle //(\mathbf{a}_1, \dots, \mathbf{a}_k), //(\mathbf{b}_1, \dots, \mathbf{b}_k) \rangle$ 记之。

将上述内积的定义和[定理 3.1]相结合，即有

$$\text{【定理 3.2】} : \langle //(\mathbf{a}_1, \dots, \mathbf{a}_k), //(\mathbf{b}_1, \dots, \mathbf{b}_k) \rangle = \det(\mathbf{a}_i \cdot \mathbf{b}_j)$$

$$\left(\text{亦即} \begin{vmatrix} \mathbf{a}_1 \cdot \mathbf{b}_1 & \mathbf{a}_1 \cdot \mathbf{b}_2 & \dots & \mathbf{a}_1 \cdot \mathbf{b}_k \\ \mathbf{a}_2 \cdot \mathbf{b}_1 & \mathbf{a}_2 \cdot \mathbf{b}_2 & \dots & \mathbf{a}_2 \cdot \mathbf{b}_k \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{a}_k \cdot \mathbf{b}_1 & \mathbf{a}_k \cdot \mathbf{b}_2 & \dots & \mathbf{a}_k \cdot \mathbf{b}_k \end{vmatrix} \right)$$

证明：由定义 $\langle //(\mathbf{a}_1, \dots, \mathbf{a}_k), //(\mathbf{b}_1, \dots, \mathbf{b}_k) \rangle$ 等于共在 k -维子空间 Π 上的 $/(a_1, \dots, a_k)$ 和 $/(b'_1, \dots, b'_k)$ 的有向 k -维体积之乘积，其 b'_j 分别是 b_j 在 Π 上的垂直投影。由[定理 3.1]即有

$$\langle //(\mathbf{a}_1, \dots, \mathbf{a}_k), //(\mathbf{b}_1, \dots, \mathbf{b}_k) \rangle = \det(\mathbf{a}_i \cdot \mathbf{b}'_j)$$

再者，由于 b'_j 乃是 b_j 在 $/(a_1, \dots, a_k)$ 所在的 k -维子空间 Π 上的垂直投影，即有

$$\begin{aligned} \mathbf{b}_j &= \mathbf{b}'_j + \mathbf{b}''_j, \quad \mathbf{b}'_j \text{ 在 } \Pi \text{ 上而 } \mathbf{b}''_j \perp \Pi \\ \Rightarrow \mathbf{a}_i \cdot \mathbf{b}''_j &= 0 \quad \text{对于所有 } i, j \text{ 皆成立} \\ \Rightarrow \mathbf{a}_i \cdot \mathbf{b}_j &= \mathbf{a}_i \cdot (\mathbf{b}'_j + \mathbf{b}''_j) = \mathbf{a}_i \cdot \mathbf{b}'_j + \mathbf{a}_i \cdot \mathbf{b}''_j = \mathbf{a}_i \cdot \mathbf{b}'_j \end{aligned}$$

对于所有 $1 \leq i, j \leq k$ 皆成立。所以

$$\langle \langle (\mathbf{a}_1, \dots, \mathbf{a}_k), \langle \langle (\mathbf{b}_1, \dots, \mathbf{b}_k) \rangle \rangle \rangle = \det(\mathbf{a}_i \cdot \mathbf{b}'_j) = \det(\mathbf{a}_i \cdot \mathbf{b}_j) \quad \square$$

[注]：上述公式的右侧乃是 $\{\mathbf{a}_1, \dots, \mathbf{a}_k; \mathbf{b}_1, \dots, \mathbf{b}_k\}$ 的多线性函数，而且对于 $\{\mathbf{a}_i\}$ 和 $\{\mathbf{b}_i\}$ 各别是斜对称的；乃是十分简洁、易算、好用者也。下面所讨论者，实乃上述多线性和其斜对称性的直接应用，亦即直截了当的展开计算是也。且先以我们身在其中的三维空间中的二维内积为例：

设 (x_1, x_2, x_3) 是空间中一个选定的正向笛氏坐标系， $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ 是其相应的正交标架，则有

$$\begin{aligned} \langle \langle (\mathbf{a}, \mathbf{b}), \langle \langle (\mathbf{e}_1, \mathbf{e}_2) \rangle \rangle \rangle &= \begin{vmatrix} \mathbf{a} \cdot \mathbf{e}_1 & \mathbf{a} \cdot \mathbf{e}_2 \\ \mathbf{b} \cdot \mathbf{e}_1 & \mathbf{b} \cdot \mathbf{e}_2 \end{vmatrix} \\ \langle \langle (\mathbf{a}, \mathbf{b}), \langle \langle (\mathbf{e}_2, \mathbf{e}_3) \rangle \rangle \rangle &= \begin{vmatrix} \mathbf{a} \cdot \mathbf{e}_2 & \mathbf{a} \cdot \mathbf{e}_3 \\ \mathbf{b} \cdot \mathbf{e}_2 & \mathbf{b} \cdot \mathbf{e}_3 \end{vmatrix} \\ \langle \langle (\mathbf{a}, \mathbf{b}), \langle \langle (\mathbf{e}_3, \mathbf{e}_1) \rangle \rangle \rangle &= \begin{vmatrix} \mathbf{a} \cdot \mathbf{e}_3 & \mathbf{a} \cdot \mathbf{e}_1 \\ \mathbf{b} \cdot \mathbf{e}_3 & \mathbf{b} \cdot \mathbf{e}_1 \end{vmatrix} \end{aligned}$$

由内积的定义，上述三者分别就是 $\langle \langle (\mathbf{a}, \mathbf{b}) \rangle \rangle$ 在三个坐标面上的垂直投影对于分别以 $\{\mathbf{e}_1, \mathbf{e}_2\}$, $\{\mathbf{e}_2, \mathbf{e}_3\}$ 和 $\{\mathbf{e}_3, \mathbf{e}_1\}$ 为其上正向的有向面积。

令 a_i (或 b_i, c_i, d_i) 分别是向量 \mathbf{a} (或 $\mathbf{b}, \mathbf{c}, \mathbf{d}$) 和 \mathbf{e}_i 的内积，亦即

$$\mathbf{a} = a_1 \mathbf{e}_1 + a_2 \mathbf{e}_2 + a_3 \mathbf{e}_3$$

$$\mathbf{b} = b_1 \mathbf{e}_1 + b_2 \mathbf{e}_2 + b_3 \mathbf{e}_3$$

$$\mathbf{c} = c_1 \mathbf{e}_1 + c_2 \mathbf{e}_2 + c_3 \mathbf{e}_3$$

$$\mathbf{d} = d_1 \mathbf{e}_1 + d_2 \mathbf{e}_2 + d_3 \mathbf{e}_3$$

【推论 1】： $\langle \langle (\mathbf{a}, \mathbf{b}), \langle \langle (\mathbf{c}, \mathbf{d}) \rangle \rangle \rangle$

$$= \begin{vmatrix} a_1 & a_2 \\ b_1 & b_2 \end{vmatrix} \cdot \begin{vmatrix} c_1 & c_2 \\ d_1 & d_2 \end{vmatrix} + \begin{vmatrix} a_2 & a_3 \\ b_2 & b_3 \end{vmatrix} \cdot \begin{vmatrix} c_2 & c_3 \\ d_2 & d_3 \end{vmatrix} + \begin{vmatrix} a_3 & a_1 \\ b_3 & b_1 \end{vmatrix} \cdot \begin{vmatrix} c_3 & c_1 \\ d_3 & d_1 \end{vmatrix}$$

【推论 1'】： $\langle \langle (\mathbf{a}, \mathbf{b}), \langle \langle (\mathbf{a}, \mathbf{b}) \rangle \rangle \rangle$

$$= \begin{vmatrix} a_1 & a_2 \\ b_1 & b_2 \end{vmatrix}^2 + \begin{vmatrix} a_2 & a_3 \\ b_2 & b_3 \end{vmatrix}^2 + \begin{vmatrix} a_3 & a_1 \\ b_3 & b_1 \end{vmatrix}^2$$

[注]: 这也就是二维 (亦即面积) 的勾股定理。

证明: 由二维内积的多线性和各别的斜对称性, 对于 $\mathbf{a}, \mathbf{b}, \mathbf{c}, \mathbf{d}$ 这四个向量变元逐一展开, 即有

$$\langle \langle \mathbf{a}, \mathbf{b} \rangle, \langle \mathbf{c}, \mathbf{d} \rangle \rangle = \sum_{1 \leq i_1, i_2, j_1, j_2 \leq 3} \langle \langle a_{i_1} \mathbf{e}_{i_1}, b_{i_2} \mathbf{e}_{i_2} \rangle, \langle c_{j_1} \mathbf{e}_{j_1}, d_{j_2} \mathbf{e}_{j_2} \rangle \rangle$$

易见在上述各项中当 $i_1 = i_2$ 或 $j_1 = j_2$ 时显然为 0, 而且在 $\{i_1, i_2\}$ 和 $\{j_1, j_2\}$ 相异时也显然为 0 (因为 $\langle \mathbf{e}_{i_1}, \mathbf{e}_{i_2} \rangle$ 和 $\langle \mathbf{e}_{j_1}, \mathbf{e}_{j_2} \rangle$ 乃是互相垂直者也)。略去上述显然为 0 之项, 不难验算所剩之项等于

$$\begin{aligned} & \begin{vmatrix} a_1 & a_2 \\ b_1 & b_2 \end{vmatrix} \cdot \begin{vmatrix} c_1 & c_2 \\ d_1 & d_2 \end{vmatrix} \langle \langle \mathbf{e}_1, \mathbf{e}_2 \rangle, \langle \mathbf{e}_1, \mathbf{e}_2 \rangle \rangle \\ + & \begin{vmatrix} a_2 & a_3 \\ b_2 & b_3 \end{vmatrix} \cdot \begin{vmatrix} c_2 & c_3 \\ d_2 & d_3 \end{vmatrix} \langle \langle \mathbf{e}_2, \mathbf{e}_3 \rangle, \langle \mathbf{e}_2, \mathbf{e}_3 \rangle \rangle \\ + & \begin{vmatrix} a_3 & a_1 \\ b_3 & b_1 \end{vmatrix} \cdot \begin{vmatrix} c_3 & c_1 \\ d_3 & d_1 \end{vmatrix} \langle \langle \mathbf{e}_3, \mathbf{e}_1 \rangle, \langle \mathbf{e}_3, \mathbf{e}_1 \rangle \rangle \end{aligned}$$

其实, $\mathbf{a}, \mathbf{b}, \mathbf{c}, \mathbf{d}$ 乃是四个向量变元, 在取定坐标系中, 它们各别的坐标 (即 $\{a_i, b_i, c_i, d_i, 1 \leq i \leq 3\}$) 乃是 12 个自变元, 可以把它们排列成一个 3×4 的矩阵

$$\begin{pmatrix} a_1 & b_1 & c_1 & d_1 \\ a_2 & b_2 & c_2 & d_2 \\ a_3 & b_3 & c_3 & d_3 \end{pmatrix}$$

而

$$\langle \langle \mathbf{a}, \mathbf{b} \rangle, \langle \mathbf{c}, \mathbf{d} \rangle \rangle = \begin{vmatrix} \mathbf{a} \cdot \mathbf{c} & \mathbf{a} \cdot \mathbf{d} \\ \mathbf{b} \cdot \mathbf{c} & \mathbf{b} \cdot \mathbf{d} \end{vmatrix}$$

则是一个它们的 4 次齐次多项式, 对于每一列是线性的, 而且对于 $\{1, 2\}$ -列和 $\{3, 4\}$ -列是斜对称的。由上面解说可见这个 4 次多项式可以自然地分解成三个多项式之和, 各别只含有其中两行的那 8 个变元。由此可见, 我们只要把其中一行的 4 个变元皆以 0 代入, 即可得出那个只含有其余两行者。例如令第三行为 0, 则 $\mathbf{a}, \mathbf{b}, \mathbf{c}, \mathbf{d}$ 都是位于 $(1, 2)$ -坐标面之中, 直接由内积的定义即得其值为 $\begin{vmatrix} a_1 & a_2 \\ b_1 & b_2 \end{vmatrix} \cdot \begin{vmatrix} c_1 & c_2 \\ d_1 & d_2 \end{vmatrix}$ 。

□

【定理 3.3】(高维勾股定理)：设 $\langle\langle \mathbf{a}_1, \dots, \mathbf{a}_k \rangle\rangle$ 和 $\langle\langle \mathbf{b}_1, \dots, \mathbf{b}_k \rangle\rangle$ 是 n -维欧氏空间的两个任给 k -维平行体， $\{\mathbf{e}_1, \dots, \mathbf{e}_n\}$ 是其中任选的一组正向正交基。令

$$a_{ij} = \mathbf{e}_i \cdot \mathbf{a}_j, \quad b_{ij} = \mathbf{e}_i \cdot \mathbf{b}_j, \quad 1 \leq i \leq n, 1 \leq j \leq k$$

则有

$$\begin{aligned} & \langle\langle \mathbf{a}_1, \dots, \mathbf{a}_k \rangle\rangle, \langle\langle \mathbf{b}_1, \dots, \mathbf{b}_k \rangle\rangle \\ &= \sum_{1 \leq i_1 < \dots < i_k \leq n} \langle\langle \mathbf{a}_1, \dots, \mathbf{a}_k \rangle\rangle, \langle\langle \mathbf{e}_{i_1}, \dots, \mathbf{e}_{i_k} \rangle\rangle \cdot \langle\langle \mathbf{b}_1, \dots, \mathbf{b}_k \rangle\rangle, \langle\langle \mathbf{e}_{i_1}, \dots, \mathbf{e}_{i_k} \rangle\rangle \\ &= \sum_{1 \leq i_1 < \dots < i_k \leq n} \begin{vmatrix} a_{i_1,1} & \dots & a_{i_1,k} \\ \vdots & \ddots & \vdots \\ a_{i_k,1} & \dots & a_{i_k,k} \end{vmatrix} \cdot \begin{vmatrix} b_{i_1,1} & \dots & b_{i_1,k} \\ \vdots & \ddots & \vdots \\ b_{i_k,1} & \dots & b_{i_k,k} \end{vmatrix} \end{aligned}$$

证明：由[定理 3.2]可见上述 k -维内积乃是 $2k$ 个向量变元 $\{\mathbf{a}_j, \mathbf{b}_j, 1 \leq j \leq k\}$ 的多线性函数，而且它对于 $\{\mathbf{a}_j\}$ 和 $\{\mathbf{b}_j\}$ 都是斜对称的。在选用正向正交坐标系 $\{\mathbf{e}_i\}$ 之下，它就可以表达成 $\{a_{ij}, b_{ij}; 1 \leq i \leq n, 1 \leq j \leq k\}$ 这样 $2kn$ 个自变元的 $2k$ 次齐次多项式，它们可编组成一个 $n \times 2k$ (或两个 $n \times k$) 的矩阵，即

$$(a_{ij}, b_{ij}) \quad (\text{或 } (a_{ij}) \text{ 和 } (b_{ij}))$$

用其多线性展开，再用 $\{\mathbf{e}_i\}$ 的正交性可见

$$\langle\langle \mathbf{e}_{i_1}, \dots, \mathbf{e}_{i_k} \rangle\rangle, \langle\langle \mathbf{e}_{i'_1}, \dots, \mathbf{e}_{i'_k} \rangle\rangle$$

在足标子集 $\{i_1, \dots, i_k\}$ 和 $\{i'_1, \dots, i'_k\}$ 不相同恒为 0。由此即得上述 $2k$ 次齐次多项式的下述分解，亦即它乃是 $\binom{n}{k}$ 个多项式 $\{P_{i_1, \dots, i_k}; 1 \leq i_1 < \dots < i_k \leq n\}$ 之和，其中 P_{i_1, \dots, i_k} 只含有 $\{i_1, \dots, i_k\}$ 这 k 行的变元。由此可见，当我们把其他 $(n-k)$ 行的所有变元都代之以 0 时，即得

$$P_{i_1, \dots, i_k} = \langle\langle \mathbf{a}'_1, \dots, \mathbf{a}'_k \rangle\rangle, \langle\langle \mathbf{b}'_1, \dots, \mathbf{b}'_k \rangle\rangle$$

其中 \mathbf{a}'_j 和 \mathbf{b}'_j , $1 \leq j \leq k$, 乃是 \mathbf{a}_j 和 \mathbf{b}_j 在 $\{\mathbf{e}_{i_1}, \dots, \mathbf{e}_{i_k}\}$ 所张的 k -维坐标面上的垂直投影是也。所以由 k -维内积的定义可见

$$\begin{aligned} & P_{i_1, \dots, i_k} \\ &= \langle \langle \mathbf{a}_1, \dots, \mathbf{a}_k \rangle, \langle \mathbf{e}_{i_1}, \dots, \mathbf{e}_{i_k} \rangle \rangle \cdot \langle \langle \mathbf{b}_1, \dots, \mathbf{b}_k \rangle, \langle \mathbf{e}_{i_1}, \dots, \mathbf{e}_{i_k} \rangle \rangle \\ &= \sum_{1 \leq i_1 < \dots < i_k \leq n} \begin{vmatrix} a_{i_1, 1} & \dots & a_{i_1, k} \\ \vdots & \ddots & \vdots \\ a_{i_k, 1} & \dots & a_{i_k, k} \end{vmatrix} \cdot \begin{vmatrix} b_{i_1, 1} & \dots & b_{i_1, k} \\ \vdots & \ddots & \vdots \\ b_{i_k, 1} & \dots & b_{i_k, k} \end{vmatrix} \quad \square \end{aligned}$$

[注]：(i) 在 $\langle \mathbf{a}_1, \dots, \mathbf{a}_k \rangle$ 和 $\langle \mathbf{b}_1, \dots, \mathbf{b}_k \rangle$ 相同的特殊情形，即有 $\langle \mathbf{a}_1, \dots, \mathbf{a}_k \rangle$ 的 k -维体积之平方等于它在各个 k -维坐标面（共有 $\binom{n}{k}$ 个）的垂直投影的 k -维体积平方之总和，这也就是 k -维体积的勾股定理。而[定理 3.3]本身则是高维勾股定理更为普遍好用的推广。

(ii) 我们可以把 $\langle \mathbf{a}_1, \dots, \mathbf{a}_k \rangle$ 在每个 k -维坐标面（共有 $\binom{n}{k}$ 个）上之垂直投影的有向 k -维体积看做它的坐标（一如把 $\mathbf{e}_i \cdot \mathbf{a}$ 看做 \mathbf{a} 的第 i 个坐标），则上述定理所证者，乃是 k -维平行体的内积也具有和常用的向量内积坐标化公式同样形式的坐标化公式。

(iii) 在一维的情形，我们有下述重要的 Schwarz 不等式，即

$$(\mathbf{a} \cdot \mathbf{b})^2 \leq (\mathbf{a} \cdot \mathbf{a})(\mathbf{b} \cdot \mathbf{b})$$

而且等号当且仅当 \mathbf{a}, \mathbf{b} 线性相关时才成立。下述推论乃是 Schwarz 不等式在高维内积的推广。

【推论】： $\langle \langle \mathbf{a}_1, \dots, \mathbf{a}_k \rangle, \langle \mathbf{b}_1, \dots, \mathbf{b}_k \rangle \rangle^2$

$$\leq \langle \langle \mathbf{a}_1, \dots, \mathbf{a}_k \rangle, \langle \mathbf{a}_1, \dots, \mathbf{a}_k \rangle \rangle \cdot \langle \langle \mathbf{b}_1, \dots, \mathbf{b}_k \rangle, \langle \mathbf{b}_1, \dots, \mathbf{b}_k \rangle \rangle$$

而且等号当且仅当 $\langle \mathbf{a}_1, \dots, \mathbf{a}_k \rangle = \langle \mathbf{b}_1, \dots, \mathbf{b}_k \rangle$ 才成立。

证明：令 $1 \leq \alpha \leq N = \binom{n}{k}$ 为标记 $\binom{n}{k}$ 个 k -维坐标面的足标， ξ_α 和 η_α 分别是 $\langle \mathbf{a}_1, \dots, \mathbf{a}_k \rangle$ 和 $\langle \mathbf{b}_1, \dots, \mathbf{b}_k \rangle$ 在第 α 个 k -维坐标面的

的垂直投影的有向 k -维体积。由[定理 3.3]即有

$$\begin{aligned} \langle //(\mathbf{a}_1, \dots, \mathbf{a}_k), //(\mathbf{b}_1, \dots, \mathbf{b}_k) \rangle &= \sum_{\alpha=1}^N \xi_{\alpha} \eta_{\alpha} \\ \langle //(\mathbf{a}_1, \dots, \mathbf{a}_k), //(\mathbf{a}_1, \dots, \mathbf{a}_k) \rangle &= \sum_{\alpha=1}^N \xi_{\alpha}^2 \\ \langle //(\mathbf{b}_1, \dots, \mathbf{b}_k), //(\mathbf{b}_1, \dots, \mathbf{b}_k) \rangle &= \sum_{\alpha=1}^N \eta_{\alpha}^2 \end{aligned}$$

由此可见，上式不等式乃是原来的 Schwarz 不等式，即

$$\left(\sum_{\alpha=1}^N \xi_{\alpha} \eta_{\alpha} \right)^2 \leq \left(\sum_{\alpha=1}^N \xi_{\alpha}^2 \right) \left(\sum_{\alpha=1}^N \eta_{\alpha}^2 \right)$$

的直接结论，而且等号只有在 $\{\xi_{\alpha}\}$ 和 $\{\eta_{\alpha}\}$ 成比例时才成立。在这种情形，我们可以选用正交坐标系使得 $//(\mathbf{a}_1, \dots, \mathbf{a}_k)$ 位于第一个坐标面之内，则 $\{\xi_{\alpha}\}$ 中只有 ξ_1 不为 0，因此 $\{\eta_{\alpha}\}$ 中也只有 η_1 不为 0，亦即 $//(\mathbf{b}_1, \dots, \mathbf{b}_k)$ 也位于第一个坐标面之内。所以等号只有在两者「平行」才能成立，亦即 $\langle \mathbf{a}_1, \dots, \mathbf{a}_k \rangle = \langle \mathbf{b}_1, \dots, \mathbf{b}_k \rangle$ ，反之亦然。因为在两者平行的情形，即有线性关系

$$\mathbf{b}_i = \sum_{j=1}^k \beta_{ij} \mathbf{a}_j, \quad 1 \leq i \leq k$$

由垂直投影的线性则有它们在任一 k -维坐标面上的垂直投影 $\{\mathbf{b}'_i\}$ 和 $\{\mathbf{a}'_j\}$ 也有同样的线性关系，亦即

$$\mathbf{b}'_i = \sum_{j=1}^k \beta_{ij} \mathbf{a}'_j \quad \Rightarrow \quad \eta_{\alpha} = \det(\beta_{ij}) \cdot \xi_{\alpha}, \quad 1 \leq \alpha \leq N \quad \square$$

[注]：在具有上述线性关系而且 $\det(\beta_{ij}) = 1$ 时，则称 $//(\mathbf{a}_1, \dots, \mathbf{a}_k)$ 和 $//(\mathbf{b}_1, \dots, \mathbf{b}_k)$ 平行而且等积。由此可见，两个 k -维平行体平行而且等积的充要条件乃是它们在 $\binom{n}{k}$ 个 k -维坐标面上的垂直投影都具有相等的有向 k -维体积，亦即 $\xi_{\alpha} = \eta_{\alpha}, 1 \leq \alpha \leq \binom{n}{k}$ 。

3.3 格氏代数 (Grassmann Algebra)

抽象化的基本想法是对于所要研讨的事物择其精要妥加组织，从而获得既精且简的体系，成为以简御繁，解答各种各样问题的有力工具。格氏代数乃是在度量几何的研讨上既精且简的基本工具。它就是把前两节研讨所得的度量几何学的精要：如定向面积， k -维平行体的内积和 k -维勾股定理等等再作妥加组织而得的代数体系；乃是高维度量几何，微分几何和多元微积分的有力基本工具。

【定义】：我们将先把互相平行等积的 k -维平行体等价类看做一种 k -维的有向几何量，称之为 k -vector，并以符号 $\mathbf{a}_1 \wedge \dots \wedge \mathbf{a}_k$ 表示包含 $\|(\mathbf{a}_1, \dots, \mathbf{a}_k)$ 于其中的那个等价类，亦即

$$\begin{aligned} & \mathbf{a}_1 \wedge \dots \wedge \mathbf{a}_k = \mathbf{b}_1 \wedge \dots \wedge \mathbf{b}_k \\ \Leftrightarrow & \|(\mathbf{a}_1, \dots, \mathbf{a}_k) \text{ 和 } \|(\mathbf{b}_1, \dots, \mathbf{b}_k) \text{ 平行而且等积} \\ \Leftrightarrow & \mathbf{b}_i = \sum_{j=1}^k \beta_{ij} \mathbf{a}_j, \quad 1 \leq i \leq k \text{ 而且 } \det(\beta_{ij}) = 1 \end{aligned}$$

倍积之定义：设 $\|(\mathbf{a}_1, \dots, \mathbf{a}_k)$ 和 $\|(\mathbf{b}_1, \dots, \mathbf{b}_k)$ 平行，而後者有向 k -维体积乃是前者的 λ 倍，亦即有

$$\mathbf{b}_i = \sum_{j=1}^k \beta_{ij} \mathbf{a}_j, \quad 1 \leq i \leq k, \text{ 而且 } \det(\beta_{ij}) = \lambda$$

则定义 $\mathbf{b}_1 \wedge \dots \wedge \mathbf{b}_k$ 为 $\mathbf{a}_1 \wedge \dots \wedge \mathbf{a}_k$ 之 λ 倍，以符号

$$\mathbf{b}_1 \wedge \dots \wedge \mathbf{b}_k = \lambda \cdot \mathbf{a}_1 \wedge \dots \wedge \mathbf{a}_k$$

记之。

外积之定义：当 $\{\mathbf{a}_1, \dots, \mathbf{a}_k; \mathbf{b}_1, \dots, \mathbf{b}_\ell\}$ 是线性无关时定义

$$(\mathbf{a}_1 \wedge \dots \wedge \mathbf{a}_k) \wedge (\mathbf{b}_1 \wedge \dots \wedge \mathbf{b}_\ell) = \mathbf{a}_1 \wedge \dots \wedge \mathbf{a}_k \wedge \mathbf{b}_1 \wedge \dots \wedge \mathbf{b}_\ell$$

亦即将 $\|(\mathbf{a}_1, \dots, \mathbf{a}_k; \mathbf{b}_1, \dots, \mathbf{b}_\ell)$ 的等价类定义为 $\|(\mathbf{a}_1, \dots, \mathbf{a}_k)$ 的等价类和 $\|(\mathbf{b}_1, \dots, \mathbf{b}_\ell)$ 的等价类的外积。当 $\{\mathbf{a}_1, \dots, \mathbf{a}_k; \mathbf{b}_1, \dots, \mathbf{b}_\ell\}$ 是线性相关时则定义两者的外积为 0。

[注]: 上述外积显然满足结合律。当 $\{\mathbf{a}_1, \dots, \mathbf{a}_k\}$ 是线性无关时, $\wedge(\mathbf{a}_1, \dots, \mathbf{a}_k)$ 所属之等价类则就是 $\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_k$ 的外积; 这也就是 $\mathbf{a}_1 \wedge \dots \wedge \mathbf{a}_k$ 这个符号的原由。

内积的定义: $\langle \mathbf{a}_1 \wedge \dots \wedge \mathbf{a}_k, \mathbf{b}_1 \wedge \dots \wedge \mathbf{b}_k \rangle := \det(\mathbf{a}_i \cdot \mathbf{b}_j)$

[这也就是[定理 3.3]的重新表述。]

至此, 所尚缺如者, 就是如何恰当地定义 k -vectors ($k \geq 2$) 之间的加法了。(也可以说, 我们所要妥为组织的度量几何的代数体系, 几乎业已水到渠成, 只待再去妥为定义「加法」。) 在我们探讨如何定义高维平行体之间的内积时, 业已指出, 在 $k=1$ 和 $k=2$ 的情形有一个根本的差别, 那就是两个有向线段具有一个自然的有向线段作为两者之和; 但是两个 k -维平行体 ($k \geq 2$) 则不再具有一个自然的 k -维平行体可以作为两者之和。由此可见, 在达成妥为定义「加法」之前, 我们还得先下一番探讨分析的功夫。

【分析】:

(i) 假若要把 $\mathbf{c}_1 \wedge \dots \wedge \mathbf{c}_k$ 定义为 $\mathbf{a}_1 \wedge \dots \wedge \mathbf{a}_k$ 和 $\mathbf{b}_1 \wedge \dots \wedge \mathbf{b}_k$ 之和, 则它至少得满足下述内积分配律, 即

$$\begin{aligned} & \langle \mathbf{c}_1 \wedge \dots \wedge \mathbf{c}_k, \mathbf{x}_1 \wedge \dots \wedge \mathbf{x}_k \rangle \\ &= \langle \mathbf{a}_1 \wedge \dots \wedge \mathbf{a}_k, \mathbf{x}_1 \wedge \dots \wedge \mathbf{x}_k \rangle + \langle \mathbf{b}_1 \wedge \dots \wedge \mathbf{b}_k, \mathbf{x}_1 \wedge \dots \wedge \mathbf{x}_k \rangle \end{aligned}$$

对于任给 $\mathbf{x}_1 \wedge \dots \wedge \mathbf{x}_k$ 恒成立。若用[定理 3.3]来验证上述必要条件, 即为

$$\begin{aligned} & \langle \mathbf{c}_1 \wedge \dots \wedge \mathbf{c}_k, \mathbf{e}_{i_1} \wedge \dots \wedge \mathbf{e}_{i_k} \rangle \\ &= \langle \mathbf{a}_1 \wedge \dots \wedge \mathbf{a}_k, \mathbf{e}_{i_1} \wedge \dots \wedge \mathbf{e}_{i_k} \rangle + \langle \mathbf{b}_1 \wedge \dots \wedge \mathbf{b}_k, \mathbf{e}_{i_1} \wedge \dots \wedge \mathbf{e}_{i_k} \rangle \end{aligned}$$

对于任给正交基 $\{\mathbf{e}_i\}$ 和 $1 \leq i_1 < \dots < i_k \leq n$ 恒成立。

(ii) 对于给定的两个 $\mathbf{a}_1 \wedge \dots \wedge \mathbf{a}_k$ 和 $\mathbf{b}_1 \wedge \dots \wedge \mathbf{b}_k$, 假如满足上述必要条件的 $\mathbf{c}_1 \wedge \dots \wedge \mathbf{c}_k$ 是存在的, 则易证它是唯一存在的, 而且应该顺理成章地把它定义为 $\mathbf{a}_1 \wedge \dots \wedge \mathbf{a}_k$ 和 $\mathbf{b}_1 \wedge \dots \wedge \mathbf{b}_k$ 之和。

(iii) 由此可见, 定义加法的症结所在是满足上述必要条件的 $\mathbf{c}_1 \wedge \dots \wedge \mathbf{c}_k$ 是否对于任给两个 $\mathbf{a}_1 \wedge \dots \wedge \mathbf{a}_k$ 和 $\mathbf{b}_1 \wedge \dots \wedge \mathbf{b}_k$ 总是存在的呢? 是耶? 非耶? 此事当然只有实事求是地去探讨。

(iv) 其实，这种「普遍存在性」是根本不成立的！例如在 $k=2, n=4$ 的情形，只要取 $\mathbf{a}_1 \wedge \mathbf{a}_2 = \mathbf{e}_1 \wedge \mathbf{e}_2$, $\mathbf{b}_1 \wedge \mathbf{b}_2 = \mathbf{e}_3 \wedge \mathbf{e}_4$ 就容易证明满足上述必要条件的 $\mathbf{c}_1 \wedge \mathbf{c}_2$ 是根本不存在的。其证明如下：

同反证法设其存在，令

$$\mathbf{c}_1 = \sum_{i=1}^4 a_i \mathbf{e}_i, \quad \mathbf{c}_2 = \sum_{i=1}^4 b_i \mathbf{e}_i$$

代入其所满足的必要条件式，即有

$$\begin{aligned} \langle \mathbf{c}_1 \wedge \mathbf{c}_2, \mathbf{e}_1 \wedge \mathbf{e}_2 \rangle &= \begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix} = 1 \Rightarrow a_1 : b_1 \neq a_2 : b_2 \\ \langle \mathbf{c}_1 \wedge \mathbf{c}_2, \mathbf{e}_3 \wedge \mathbf{e}_4 \rangle &= \begin{vmatrix} a_3 & b_3 \\ a_4 & b_4 \end{vmatrix} = 1 \Rightarrow a_3 : b_3 \neq a_4 : b_4 \\ \langle \mathbf{c}_1 \wedge \mathbf{c}_2, \mathbf{e}_1 \wedge \mathbf{e}_3 \rangle &= \begin{vmatrix} a_1 & b_1 \\ a_3 & b_3 \end{vmatrix} = 0 \Rightarrow a_3 : b_3 = a_1 : b_1 \\ \langle \mathbf{c}_1 \wedge \mathbf{c}_2, \mathbf{e}_2 \wedge \mathbf{e}_3 \rangle &= \begin{vmatrix} a_2 & b_2 \\ a_3 & b_3 \end{vmatrix} = 0 \Rightarrow a_3 : b_3 = a_2 : b_2 \\ \langle \mathbf{c}_1 \wedge \mathbf{c}_2, \mathbf{e}_1 \wedge \mathbf{e}_4 \rangle &= \begin{vmatrix} a_1 & b_1 \\ a_4 & b_4 \end{vmatrix} = 0 \Rightarrow a_4 : b_4 = a_1 : b_1 \\ \langle \mathbf{c}_1 \wedge \mathbf{c}_2, \mathbf{e}_2 \wedge \mathbf{e}_4 \rangle &= \begin{vmatrix} a_2 & b_2 \\ a_4 & b_4 \end{vmatrix} = 0 \Rightarrow a_4 : b_4 = a_2 : b_2 \end{aligned}$$

易见上述诸式是互相矛盾的，这也就证明了 $\mathbf{c}_1 \wedge \mathbf{c}_2$ 的不存在性！

总结上述几点实况的分析，摆在我们面前者只有两种处理方式：其一是鉴于上述不存在性而根本放弃加法之引入，其二则是依然引入加法，但是在 $\mathbf{c}_1 \wedge \dots \wedge \mathbf{c}_k$ 不存在的情形把 $\mathbf{a}_1 \wedge \dots \wedge \mathbf{a}_k$ 和 $\mathbf{b}_1 \wedge \dots \wedge \mathbf{b}_k$ 的和想成一种比 k -维平行体更为广义的事物。显然，我们是不能接受前者的！岂能功败垂成，前功尽弃呢？是不？总之，我们是要义无反顾地完成第二种处理方式的。其基本思想是把 k -维平行体的等价类所组成的集合，适当地扩充成一个具有内积的向量空间，当把内积限制到 k -维平行体的等价类时，它就是 §3.2 中所定义者也。显然这个我们所要寻求的内积空间包含下述 $\binom{n}{k}$ 个 k -维平行体等价类：

$$\{\mathbf{e}_{i_1} \wedge \dots \wedge \mathbf{e}_{i_k}; 1 \leq i_1 < \dots < i_k \leq n\}$$

它们彼此正交而且都是单位长的。再者，它们的所有线性组合所构成的 $\binom{n}{k}$ -维内积空间业已包括所有 k -维平行体的等价类了。由此可见，它就是所要寻求的适当而且自然的扩充，至此业已水到渠成，兹定义格氏代数如下：

【定义】：设 V 为 \mathbb{R} 上的一个 n -维内积空间。令 $\Lambda^0(V) = \mathbb{R}$, $\Lambda^1(V) = V$, $\Lambda^k(V)$, $2 \leq k \leq n$, 是包含所有 k -维平行体的等价类为其子集的 $\binom{n}{k}$ -维内积空间。对于 V 中任给一组正交基 $\{\mathbf{e}_i, 1 \leq i \leq n\}$, 则

$$\{\mathbf{e}_{i_1} \wedge \dots \wedge \mathbf{e}_{i_k}, 1 \leq i_1 < \dots < i_k \leq n\}$$

构成 $\Lambda^k(V)$ 的一组正交基。令

$$G(V) = \bigoplus_{k=0}^n \Lambda^k(V)$$

再者，定义 $G(V)$ 中的外积为下述双线性映射：

$$\wedge : \Lambda^k(V) \times \Lambda^\ell(V) \rightarrow \Lambda^{k+\ell}(V)$$

其限制于平行体的等价类的情形即为前述

$$(\mathbf{a}_1 \wedge \dots \wedge \mathbf{a}_k) \wedge (\mathbf{b}_1 \wedge \dots \wedge \mathbf{b}_\ell) := \mathbf{a}_1 \wedge \dots \wedge \mathbf{a}_k \wedge \mathbf{b}_1 \wedge \dots \wedge \mathbf{b}_\ell$$

[在 $k + \ell > n$ 或 $\{\mathbf{a}_1, \dots, \mathbf{a}_k, \mathbf{b}_1, \dots, \mathbf{b}_\ell\}$ 线性相关时，其外积为 0 。]

第四章

外微分与多元积分

在单元微积分中，微分与积分分别起源于函数的变率与求和这两个定量型的基本性质，是各有所本，各有特色者；但是它们却是互逆的两种基本运算，而这种互逆关系也就是微积分基本定理。

本章将以前三章研讨之所得为基础，进而研讨在多元、多关系的数理分析中常见常用的各种各样积分，以及微积分基本定理在多元分析学中自然而且基本的推广。

4.1 多元函数的多重积分

4.1.1 多重积分的定义

让我们先来看一看下述最为简单的情形：设 $f(x_1, x_2)$ 是定义于一个「坐标矩形」 D ：

$$D = \{(x_1, x_2); a \leq x_1 \leq b, c \leq x_2 \leq d\}$$

上的连续函数。我们可以和单元相同地用二元阶梯函数去上、下夹逼 $f(x_1, x_2)$ ，而且同样地用无限细分和比较原则去确定二重定积分

$$\int_D f(x_1, x_2) dx_1 \wedge dx_2$$

的定义。其具体做法可以把 $a \leq x_1 \leq b$ 和 $c \leq x_2 \leq d$ 分别等分成 N -段, $N = 2^k$, 令其分点为:

$$\begin{aligned} a &= a_0 < a_1 < \dots < a_i < \dots < a_N = b \\ c &= c_0 < c_1 < \dots < c_j < \dots < c_N = d \end{aligned}$$

则纵线 $\{x_1 = a_i, 1 \leq i \leq N-1\}$ 和横线 $\{x_2 = c_j, 1 \leq j \leq N-1\}$ 就把 D 分割成 N^2 个小格, 分别令其为 $\{D_{ij}, 1 \leq i, j \leq N\}$ 。再者, 令 M_{ij} 和 m_{ij} 分别是 $f(x_1, x_2)$ 在 D_{ij} 上所取之极大和极小值, 而 $g_k(x_1, x_2)$ 和 $G_k(x_1, x_2)$ 则是分别在 D_{ij} 之内点取常数值 m_{ij} 和 M_{ij} 的阶梯函数, 显然有

$$g_k(x_1, x_2) \leq f(x_1, x_2) \leq G_k(x_1, x_2)$$

对于 D 中任给 (x_1, x_2) 皆成立。所以由同样的比较原则即有

$$\begin{aligned} \int_D g_k(x_1, x_2) dx_1 \wedge dx_2 &\leq \int_D f(x_1, x_2) dx_1 \wedge dx_2 \leq \int_D G_k(x_1, x_2) dx_1 \wedge dx_2 \\ \frac{(b-a)(d-c)}{N^2} \sum_{1 \leq i, j \leq N} m_{ij} &\qquad \qquad \qquad \frac{(b-a)(d-c)}{N^2} \sum_{1 \leq i, j \leq N} M_{ij} \end{aligned}$$

易见对于任给 $(x_1, x_2) \in D$, $\{g_k(x_1, x_2)\}$ 和 $\{G_k(x_1, x_2)\}$ 分别是 k 的递增和递减数列。再者, 由 $f(x_1, x_2)$ 在 D 上的均匀连续性, 可见当把 k 取得足够大时, 所有 $(M_{ij} - m_{ij})$ 都可以小到足够小。所以在 $k \rightarrow \infty$ 时, 即有

$$\left(\int_D G_k(x_1, x_2) dx_1 \wedge dx_2 - \int_D g_k(x_1, x_2) dx_1 \wedge dx_2 \right) \rightarrow 0$$

由此可见, 上述左、右夹逼数列具有那个唯一被夹逼于其间的值, 这也就是我们所要定义的二重定积分

$$\int_D f(x_1, x_2) dx_1 \wedge dx_2$$

它乃是左、右夹逼数列

$$\left\{ \int_D g_k(x_1, x_2) dx_1 \wedge dx_2 \right\} \quad \text{和} \quad \left\{ \int_D G_k(x_1, x_2) dx_1 \wedge dx_2 \right\}$$

的共同极限。再者，假如我们把上述左、右夹逼之和式作下述改写：

$$\frac{(b-a)(d-c)}{N^2} \sum_{j=1}^N \left(\sum_{i=1}^N m_{ij} \right) \quad \text{和} \quad \frac{(b-a)(d-c)}{N^2} \sum_{j=1}^N \left(\sum_{i=1}^N M_{ij} \right)$$

$$\left[\text{或} \frac{(b-a)(d-c)}{N^2} \sum_{i=1}^N \left(\sum_{j=1}^N m_{ij} \right) \quad \text{和} \quad \frac{(b-a)(d-c)}{N^2} \sum_{i=1}^N \left(\sum_{j=1}^N M_{ij} \right) \right]$$

就可以看出上述共同极限其实也等于下述逐次单元积分之所得，即有

$$\int_D f(x_1, x_2) dx_1 \wedge dx_2 = \int_c^d \left\{ \int_a^b f(x_1, x_2) dx_1 \right\} dx_2$$

$$(\text{或}) = \int_a^b \left\{ \int_c^d f(x_1, x_2) dx_2 \right\} dx_1$$

[注]：在上式右端的 $\int_a^b f(x_1, x_2) dx_1$ (或 $\int_c^d f(x_1, x_2) dx_2$) 所表达者，乃是把 x_2 (或 x_1) 暂且取定之後的单元函数之定积分 (类似于「偏微分」中的处理方式)，所以将称之为「偏积分」。

【例子】：

$$\int_D e^{-(x_1^2+x_2^2)} dx_1 \wedge dx_2 = \int_c^d \left\{ \int_a^b e^{-(x_1^2+x_2^2)} dx_1 \right\} dx_2$$

$$= \int_c^d \left\{ e^{-x_2^2} \int_a^b e^{-x_1^2} dx_1 \right\} dx_2$$

$$= \int_c^d e^{-x_2^2} dx_2 \cdot \int_a^b e^{-x_1^2} dx_1$$

再取 $a, c \rightarrow -\infty$ 而 $b, d \rightarrow +\infty$ 的极限，则有

$$\int_{\mathbb{R}^2} e^{-(x_1^2+x_2^2)} dx_1 \wedge dx_2 = \int_{-\infty}^{\infty} e^{-x_1^2} dx_1 \cdot \int_{-\infty}^{\infty} e^{-x_2^2} dx_2 = \left[\int_{-\infty}^{\infty} e^{-x^2} dx \right]^2$$

接著让我们来讨论一般 n 个自变元的情形。设 $f(x_1, \dots, x_n)$ 是一个定义于下述 n -维坐标矩形 D 上的连续函数

$$D = \{(x_1, \dots, x_n); a_i \leq x_i \leq b_i, 1 \leq i \leq n\}$$

同样地, 我们也可以把区间 $[a_i, b_i]$ N -等分 ($N = 2^k$), 从而把 n -维坐标矩形分割成 N^n 个 n -维小格

$$\{D_{i_1, \dots, i_n}; 1 \leq i_1, \dots, i_n \leq N\}$$

令 m_{i_1, \dots, i_n} 和 M_{i_1, \dots, i_n} 分别是 $f(x_1, \dots, x_n)$ 在 D_{i_1, \dots, i_n} 上取值之极小与极大, 而 $g_k(x_1, \dots, x_n)$ 和 $G_k(x_1, \dots, x_n)$ 则是分别在 D_{i_1, \dots, i_n} 之内点取常数值 m_{i_1, \dots, i_n} 和 M_{i_1, \dots, i_n} 的 n 元阶梯函数。由同样的比较原则, 即有

$$\begin{aligned} \frac{1}{N^n} \prod_{i=1}^n (b_i - a_i) \sum_{1 \leq i_1, \dots, i_n \leq n} m_{i_1, \dots, i_n} &= \int_D g_k(x_1, \dots, x_n) dx_1 \wedge \dots \wedge dx_n \\ &\quad \wedge \\ &\quad \int_D f(x_1, \dots, x_n) dx_1 \wedge \dots \wedge dx_n \\ &\quad \wedge \\ \frac{1}{N^n} \prod_{i=1}^n (b_i - a_i) \sum_{1 \leq i_1, \dots, i_n \leq n} M_{i_1, \dots, i_n} &= \int_D G_k(x_1, \dots, x_n) dx_1 \wedge \dots \wedge dx_n \end{aligned}$$

再用 $f(x_1, \dots, x_n)$ 在 D 上的均匀连续性易证下述左、右夹逼数列

$$\left\{ \int_D g_k(x_1, \dots, x_n) dx_1 \wedge \dots \wedge dx_n \right\} \quad \text{和} \quad \left\{ \int_D G_k(x_1, \dots, x_n) dx_1 \wedge \dots \wedge dx_n \right\}$$

确定了 $\int_D f(x_1, \dots, x_n) dx_1 \wedge \dots \wedge dx_n$ 为其唯一共同极限值。再者, 我们也可以同样地把上述左、右夹逼之和式改写为

$$\begin{aligned} &\frac{1}{N^n} \prod_{i=1}^n (b_i - a_i) \sum_{i_1=1}^N \left\{ \sum_{i_2=1}^N \left\{ \dots \left\{ \sum_{i_n=1}^N m_{i_1, \dots, i_n} \right\} \dots \right\} \right\} \\ \text{和} &\frac{1}{N^n} \prod_{i=1}^n (b_i - a_i) \sum_{i_1=1}^N \left\{ \sum_{i_2=1}^N \left\{ \dots \left\{ \sum_{i_n=1}^N M_{i_1, \dots, i_n} \right\} \dots \right\} \right\} \end{aligned}$$

即可看到上述共同极限其实也等于下述逐次偏积分之所得, 即有

$$\begin{aligned} &\int_D f(x_1, \dots, x_n) dx_1 \wedge \dots \wedge dx_n \\ &= \int_{a_1}^{b_1} dx_1 \left\{ \dots \left\{ \int_{a_{n-1}}^{b_{n-1}} dx_{n-1} \left\{ \int_{a_n}^{b_n} f(x_1, \dots, x_n) dx_n \right\} \right\} \dots \right\} \end{aligned}$$

【例子】：

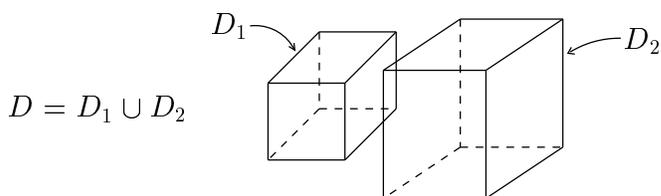
$$\begin{aligned} & \int_D e^{-(x_1^2+\dots+x_n^2)} dx_1 \wedge \dots \wedge dx_n \\ &= \int_{a_1}^{b_1} e^{-x_1^2} dx_1 \left\{ \dots \left\{ \int_{a_{n-1}}^{b_{n-1}} e^{-x_{n-1}^2} dx_{n-1} \left\{ \int_{a_n}^{b_n} e^{-x_n^2} dx_n \right\} \right\} \dots \right\} \\ &= \prod_{i=1}^n \left\{ \int_{a_i}^{b_i} e^{-x_i^2} dx_i \right\} \end{aligned}$$

再取 $a_i \rightarrow -\infty, b_i \rightarrow +\infty, 1 \leq i \leq n$, 的极限, 则有

$$\int_{\mathbb{R}^n} e^{-(x_1^2+\dots+x_n^2)} dx_1 \wedge \dots \wedge dx_n = \prod_{i=1}^n \int_{-\infty}^{\infty} e^{-x_i^2} dx_i = \left[\int_{-\infty}^{\infty} e^{-x^2} dx \right]^n$$

上述简明扼要的讨论, 业已明确了一个多元连续函数 $f(x_1, \dots, x_n)$ 在一个 n -维坐标矩形上的多重积分之定义。接著让我们进而研讨如何把多重积分的区域 D 加以适当的推广。

首先, 由于积分乃是求和之本质, 若区域 D 乃是两个分别都是坐标矩形 D_1, D_2 彼此不相重叠 (non-overlapping) 的和集(如 [图 4-1] 所示):



[图 4-1]

则显然应该把 $f(\mathbf{x})$ 在 D 上的多重积分, 直截了当地定义为它分别在 D_1 和 D_2 上的多重积分之和, 亦即

$$\begin{aligned} & \int_D f(x_1, \dots, x_n) dx_1 \wedge \dots \wedge dx_n \\ &:= \int_{D_1} f(x_1, \dots, x_n) dx_1 \wedge \dots \wedge dx_n + \int_{D_2} f(x_1, \dots, x_n) dx_1 \wedge \dots \wedge dx_n \end{aligned}$$

再者, 若 D 能够分解成有限个 n -维坐标矩形 $\{D_i, 1 \leq i \leq N\}$ 的和集, 而且不含任何重叠, 则 $f(\mathbf{x})$ 在 D 上的多重积分, 就是它在 D_i 上的多

重积分的总和，亦即

$$\int_D f(x_1, \dots, x_n) dx_1 \wedge \dots \wedge dx_n = \sum_{i=1}^N \int_{D_i} f(x_1, \dots, x_n) dx_1 \wedge \dots \wedge dx_n$$

其实，我们还可以师法古希腊即已想到、用到的穷尽原理 (principle of exhaustion)，把多重积分的区域推广到业已非常广泛的范畴。

设 Ω 是一个包含于 n -维坐标矩形 D 之内的区域，即有

$$\Omega \subseteq D = \{(x_1, \dots, x_n); a_i \leq x_i \leq b_i, 1 \leq i \leq n\}$$

分别将区间 $[a_i, b_i]$ N -等分, $N = 2^k$, 从而将 D 分割成 N^n 个 n -维小格

$$\{D_{i_1, \dots, i_n}; 1 \leq i_1, \dots, i_n \leq N\}$$

令 $\check{\Omega}_k$ (及 $\hat{\Omega}_k$) 分别是那些完全包含于 Ω 之内 (及和 Ω 有非空之交集) 的 n -维小格的和集, 则显然有

$$\check{\Omega}_k \subset \check{\Omega}_{k+1} \subset \Omega \subset \hat{\Omega}_{k+1} \subset \hat{\Omega}_k$$

再者, 若有

$$\lim_{k \rightarrow \infty} \text{vol}_n(\check{\Omega}_k) = \lim_{k \rightarrow \infty} \text{vol}_n(\hat{\Omega}_k)$$

则上述共同极限就自然应该定义为 $\text{vol}_n(\Omega)$, 而且

$$\lim_{k \rightarrow \infty} \int_{\check{\Omega}_k} f(\mathbf{x}) dx_1 \wedge \dots \wedge dx_n$$

也就可以作为 $f(\mathbf{x})$ 在这种区域 Ω 上的多重积分的定义, 即

$$\int_{\Omega} f(\mathbf{x}) dx_1 \wedge \dots \wedge dx_n := \lim_{k \rightarrow \infty} \int_{\check{\Omega}_k} f(\mathbf{x}) dx_1 \wedge \dots \wedge dx_n$$

4.1.2 多重积分与坐标变换

在多元、多关系的数理分析中, 自变元的选取乃是要「因地制宜」和「因时制宜」者也, 往往在不同的区域或不同的阶段, 宜于采用不同的变元作为数理分析的自变元, 这也就是数理分析中经常要用到的

局部坐标系的妥为选取和坐标变换。在讨论多重积分的坐标变换公式之前，且让我们再来回顾一下单元积分的情形。

设 $x = \varphi(t)$ 是 t 的一阶连续可微函数，而且 $a = \varphi(c)$, $b = \varphi(d)$ 。则有公式：

$$\int_a^b f(x)dx = \int_c^d f(\varphi(t))\varphi'(t)dt$$

在形式上，这乃是一种直接代换，亦即把 x 代以 $\varphi(t)$ ，把 dx 代以 $\varphi'(t)dt$ ；而求积分的区间则是在对于 t 的积分区间 $[c, d]$ ，相应于对于 x 的积分区间 $[\varphi(c), \varphi(d)]$ 。归根究底，在实质内含上，公式右侧乃是下述和式（在无限细分）的极限，即

$$\sum_{i=1}^N f(\varphi(t_i))\varphi'(t_i)\Delta t_i, \quad \Delta t_i = (t_i - t_{i-1})$$

而公式左侧则是下述和式的极限，即

$$\sum_{i=1}^N f(x_i)\Delta x_i$$

其中 $x_i = \varphi(t_i)$, $\Delta x_i = \varphi(t_i) - \varphi(t_{i-1})$ 。由所设 $\varphi(t)$ 的一阶连续可微性， $\varphi'(t_i)\Delta t_i$ 乃是 Δx_i 的一阶逼近，亦即存在 ε 使得

$$|\Delta x_i - \varphi'(t_i)\Delta t_i| < \varepsilon \Delta t_i$$

而且当 $\Delta t_i \rightarrow 0$ 时 $\varepsilon \rightarrow 0$ 。由此可见，上述两个和式之差别满足下述估计，即

$$\left| \sum_{i=1}^N f(\varphi(t_i))(\Delta x_i - \varphi'(t_i)\Delta t_i) \right| < M\varepsilon \sum_{i=1}^N \Delta t_i = M\varepsilon(d - c)$$

其中 M 是 $|f(\varphi(t))|$ 在 $[c, d]$ 上的极大值。所以在无限细分下其差别趋于 0，这也就是上述积分公式的证明。

接著让我们先来研讨二维的情形。设

$$u_1 = g_1(x_1, x_2), \quad u_2 = g_2(x_1, x_2)$$

这样一对一阶连续可微函数定义了一个将 (x_1, x_2) -平面中的区域 Ω 映射到 (u_1, u_2) -平面中的区域 Ω' 之上, 而 $f(u_1, u_2)$ 则是定义于 Ω' 上的连续函数。先从形式上直接代入, 即有

$$\begin{aligned} du_1 &= \frac{\partial g_1}{\partial x_1} dx_1 + \frac{\partial g_1}{\partial x_2} dx_2, & du_2 &= \frac{\partial g_2}{\partial x_1} dx_1 + \frac{\partial g_2}{\partial x_2} dx_2 \\ du_1 \wedge du_2 &= \left(\frac{\partial g_1}{\partial x_1} \frac{\partial g_2}{\partial x_2} - \frac{\partial g_1}{\partial x_2} \frac{\partial g_2}{\partial x_1} \right) dx_1 \wedge dx_2 = \frac{\partial(g_1, g_2)}{\partial(x_1, x_2)} dx_1 \wedge dx_2 \end{aligned}$$

所以应该有公式:

$$(*)_2 \quad \int_{\Omega'} f(u_1, u_2) du_1 \wedge du_2 = \int_{\Omega} f(g_1(\mathbf{x}), g_2(\mathbf{x})) \frac{\partial(g_1, g_2)}{\partial(x_1, x_2)} dx_1 \wedge dx_2$$

接著再从 $(*)_2$ -式侧的实质内含来探讨其证明。

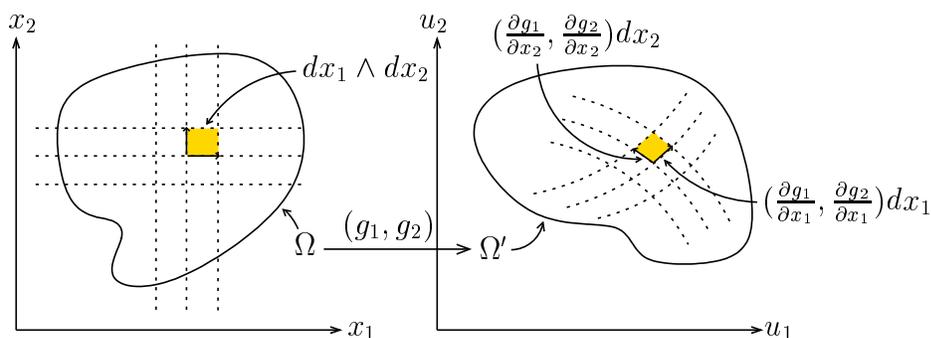
我们先用 (x_1, x_2) -平面上的纵横坐标线把 Ω 分割成很多小格, $dx_1 \wedge dx_2$ 所表达者乃是其中一个典型小格的有向面积。在 (g_1, g_2) 所定义的映射之下 (如 [图 4-2] 所示), 上述典型小格的映象是一个微小的典型四边形。它的有向面积可以用下述微小切向量:

$$\left(\frac{\partial g_1}{\partial x_1}, \frac{\partial g_2}{\partial x_1} \right) dx_1 \quad \text{和} \quad \left(\frac{\partial g_1}{\partial x_2}, \frac{\partial g_2}{\partial x_2} \right) dx_2$$

所张的微小平行四边形的有向面积表达其二阶逼近, 亦即其有向面积和

$$\frac{\partial(g_1, g_2)}{\partial(x_1, x_2)} dx_1 \wedge dx_2$$

的差别乃是比 $dx_1 \wedge dx_2$ 更高阶的微量。



[图 4-2]

不难看到，当我们用细分求和的极限去表达 $(*)_2$ -式之左侧时，采用 $\{u_1, u_2\}$ 的坐标线把 Ω' 细分乃是一种直观简单的分割法，但是此事并非必要！例如我们也可以如 [图 4-2] 所示，改用 (x_1, x_2) 的坐标线在 (g_1, g_2) -映射下所得的两系参数曲线把 Ω' 加以细分，则所得之和式中的典型项乃是

$$\begin{aligned} & f(u_1, u_2) \cdot \text{曲线小格的有向面积} \\ &= f(g_1(\mathbf{x}), g_2(\mathbf{x})) \cdot \frac{\partial(g_1, g_2)}{\partial(x_1, x_2)} dx_1 \wedge dx_2 + \varepsilon dx_1 \wedge dx_2 \end{aligned}$$

其中在无限细分下 $\varepsilon \rightarrow 0$ 。

由此可见，改用上述两系参数曲线把 Ω' 加以细分所得之和式和 $(*)_2$ -式右侧二重积分采用相应的 (x_1, x_2) 坐标线的细分所得的和式，其差别至多是

$$M\varepsilon(\sum dx_1 \wedge dx_2) = M\varepsilon \cdot \Omega \text{ 的面积}$$

其中 M 乃是 $f(g_1(\mathbf{x}), g_2(\mathbf{x}))$ 在 Ω 上的极大值。易见它是在无限细分下趋于 0 的，这也就证明了 $(*)_2$ -式。

上述对于二元的讨论，可以直接推广到 n 元，其结果乃是下述 [定理 4.1]：

【定理 4.1】：设 $u_i = g_i(x_1, \dots, x_n)$, $1 \leq i \leq n$ ，乃是将 (x_i) -空间中的区域 Ω 映射到 (u_i) -空间中的 n -维区域 Ω' 之上的一阶连续可微映射， $f(u_1, \dots, u_n)$ 是定义在 Ω' 上的连续函数，则有

$$\begin{aligned} & \int_{\Omega'} f(u_1, \dots, u_n) du_1 \wedge \dots \wedge du_n \\ (*)_n &= \int_{\Omega} f(g_1(\mathbf{x}), \dots, g_n(\mathbf{x})) \frac{\partial(g_1, \dots, g_n)}{\partial(x_1, \dots, x_n)} dx_1 \wedge \dots \wedge dx_n \end{aligned}$$

证明：上述公式的验证乃是 $(*)_2$ -式的论证的直接推广，读者可先行讨论 $n=3$ 的情形，然后再验证一般的情形，乃是一种促进自己对多重积分的理解的好练习。

【例子】：在 §4.1.1 分节中，由指数定则和 n -维坐标矩形上的多重积分可以归于逐次偏积分相结合，说明

$$\int_{\mathbb{R}^n} e^{-(x_1^2 + \dots + x_n^2)} dx_1 \wedge \dots \wedge dx_n = \left\{ \int_{-\infty}^{\infty} e^{-x^2} dx \right\}^n$$

但是具有基本重要性的单元积分 $\int_{-\infty}^{\infty} e^{-x^2} dx$ 之值的确定, 其实本身依然是一个具有挑战性的问题。其实, 解决上述求值问题的一个巧妙的途径乃是利用公式:

$$\left\{ \int_{-\infty}^{\infty} e^{-x^2} dx \right\}^2 = \int_{\mathbb{R}^2} e^{-(x_1^2+x_2^2)} dx_1 \wedge dx_2$$

然後利用右侧重积分所含有的旋转对称性计算得其值乃是 π !

有鉴于极坐标乃是充分体现 \mathbb{R}^2 上的旋转对称性的坐标系, 所以上式右侧的重积分应当改用极坐标加以表达, 这样才能充分体现它乃是旋转对称的好处。由坐标变换公式

$$x_1 = r \cos \theta, \quad x_2 = r \sin \theta$$

即得

$$\begin{aligned} dx_1 \wedge dx_2 &= (\cos \theta dr - r \sin \theta d\theta) \wedge (\sin \theta dr + r \cos \theta d\theta) \\ &= r dr \wedge d\theta \end{aligned}$$

所以

$$\begin{aligned} \int_{\mathbb{R}^2} e^{-(x_1^2+x_2^2)} dx_1 \wedge dx_2 &= \int_0^{\infty} dr \int_0^{2\pi} e^{-r^2} r d\theta \\ &= 2\pi \int_0^{\infty} e^{-r^2} r dr = \pi \int_0^{\infty} e^{-u} du, \quad u = r^2 \\ &= \pi \end{aligned}$$

由此即得

$$\int_{-\infty}^{\infty} e^{-x^2} = \sqrt{\pi}$$

既然业已求得上述积分之值, 我们又可以再用

$$\int_{\mathbb{R}^n} e^{-(x_1^2+\dots+x_n^2)} dx_1 \wedge \dots \wedge dx_n = \pi^{\frac{n}{2}}$$

和上述多重积分的球对称性反过来求得 \mathbb{R}^n 中的单位球而的 $(n-1)$ -维体积之公式。令其为 ω_n , 不难由多重积分的几何内含看到

$$\begin{aligned} &\int_{\mathbb{R}^n} e^{-(x_1^2+\dots+x_n^2)} dx_1 \wedge \dots \wedge dx_n \\ &= \int_0^{\infty} e^{-r^2} r^{n-1} \omega_n dr = \omega_n \int_0^{\infty} e^{-r^2} r^{n-1} dr \end{aligned}$$

其中 $r^{n-1}\omega_n dr$ 乃是分别以 r 和 $r+dr$ 为半径的同心球面之间的 n -维体积。再者

$$\int_0^\infty e^{-r^2} r^{n-1} dr = \frac{1}{2} \int_0^\infty e^{-u} u^{\frac{n}{2}-1} du = \frac{1}{2} \Gamma\left(\frac{n}{2}\right)$$

而 Euler's Γ -函数具有性质：

$$\Gamma(x) = (x-1)\Gamma(x-1) \quad (\text{参看第四册第三章之习题 9})$$

由此可得当 n 是偶数时

$$\Gamma\left(\frac{n}{2}\right) = \left(\frac{n}{2}-1\right)! \quad (! \text{表示阶乘})$$

而当 n 是奇数时

$$\Gamma\left(\frac{n}{2}\right) = \left(\frac{n}{2}-1\right)\left(\frac{n}{2}-2\right)\cdots\frac{1}{2}\Gamma\left(\frac{1}{2}\right), \quad \Gamma\left(\frac{1}{2}\right) = \sqrt{\pi}$$

【习题】：在 $n=3$ 的情形，试用球面坐标和笛氏坐标的坐标变换公式验证：

$$\begin{aligned} & \int_{\mathbb{R}^3} e^{-(x_1^2+x_2^2+x_3^2)} dx_1 \wedge dx_2 \wedge dx_3 \\ &= 4\pi \int_0^\infty e^{-r^2} r^2 dr \quad (\text{Archimedes' } \omega_3 = 4\pi) \end{aligned}$$

[注]：在 $n \geq 4$ 的一般情形，我们也可以有 n -维欧氏空间的球坐标，它除了矢径之长 r 之外还有 $(n-1)$ 个角度的坐标。当 r 取定为 1 时，它们就是 $(n-1)$ -维单位球面上的一组坐标。在下一节我们就会讨论如何用它们的积分公式表达 ω_n 。

4.2 线积分、面积分及其高维推广

在一般多元、多关系的数理分析中，所涉及的体系往往含有多个参变量（以变元表之），而它们又满足多个函数关系，这也就是在第二章所研讨的模式

$$S = \{\mathbf{x} = (x_1, \dots, x_n), f(\mathbf{x}) = 0, 1 \leq i \leq m\}$$

其中 $f_i(\mathbf{x})$ 都是在某一定义域上至少一阶连续可微的。 S 中的一点 \mathbf{x}_0 ，若有 $1 \leq j_1 < \dots < j_m \leq n$ 使得

$$\left. \frac{\partial(f_1, \dots, f_m)}{\partial(x_{j_1}, \dots, x_{j_m})} \right|_{\mathbf{x}_0} \neq 0$$

则称 \mathbf{x}_0 为 S 中的一个规则点 (regular point)。由隐函数定理，得知在这种规则点 \mathbf{x}_0 的邻近 S 在概念上可以改用 m 个 $(n-m)$ 个自变元的隐函数描述之，亦即

$$x_{j_1} = g_1(\tilde{\mathbf{x}}), \dots, x_{j_m} = g_m(\tilde{\mathbf{x}})$$

其中 $\tilde{\mathbf{x}}$ 表示 $\{x_{j_1}, \dots, x_{j_m}\}$ 之外的那 $(n-m)$ 个变元，其变动范围乃是 $\tilde{\mathbf{x}}_0$ 的一个适当邻域。用几何的语言来说， S 在规则点 \mathbf{x}_0 的邻近乃是 \mathbb{R}^n 之中的一片 $(n-m)$ -维曲面。由此可见，一个足以研讨多元、多关系的数理分析的积分概念，当然还得包括那种以 \mathbb{R}^n 中的一片 k -维曲面为其求积之区域的曲面积分。其中最为简单的情形当然就是 $k=1$ 的线积分。

4.2.1 线积分 (Line Integral)

\mathbb{R}^n 中的一段曲线 γ 若能用一种一阶连续可微的参数式加以表达，亦即

$$\gamma = \{\varphi(t) = (\varphi_1(t), \dots, \varphi_n(t)), a \leq t \leq b\}$$

则称 γ 为一阶可微曲线。我们可以把 φ 看做由区间 $[a, b]$ 到 \mathbb{R}^n 中的一个一阶可微映射，而 γ 则是其象 (image)。显然，同一条曲线 γ 可以有很多不同的参数式加以表达，往後我们将明确其参数式者叫做参数曲线，所以参数曲线其实就是上述映射本身。在概念上，一个映射 φ 和其象点所组成的子集 $\gamma = \{\varphi(t), a \leq t \leq b\}$ 是有本质上的区别的，请读者注意，切勿混淆。例如单位圆乃是一条曲线， $x_1 = \cos kt$, $x_2 = \sin kt$ 所描述者则是参数曲线。当 k 不同时，其所描述者乃是不同的参数曲线。

设 $f_i(\mathbf{x})$, $1 \leq i \leq n$, 是定义于 \mathbb{R}^n 中某一区域 D 上的连续函数，则表式

$$\omega = \sum_{i=1}^n f_i(\mathbf{x}) dx_i$$

叫做 D 上的一个 一次微分形式 (differential 1-form)。

设 $\varphi: [a, b] \rightarrow D \subseteq \mathbb{R}^n$, $\varphi(t) = (\varphi_1(t), \dots, \varphi_n(t))$ 是 D 中的一条一阶可微参数曲线, 我们将以 $\varphi^*(\omega)$ 表示把 ω 中的 x_i 代以 $\varphi_i(t)$ 之所得, 亦即

$$\varphi^*(\omega) = \sum_{i=1}^n f_i(\varphi(t)) \varphi_i'(t) dt$$

再者, 设 $\psi: [c, d] \rightarrow D$, $\psi(s) = (\psi_1(s), \dots, \psi_n(s))$ 是同一条曲线 γ 的另一种参数描述, 亦即

$$\{\varphi(t), a \leq t \leq b\} = \gamma = \{\psi(s), c \leq s \leq d\}$$

是 D 中同一个子集, 则有

$$\psi^*(\omega) = \sum_{i=1}^n f_i(\psi(s)) \psi_i'(s) ds$$

试问是否恒有

$$\int_a^b \varphi^*(\omega) \stackrel{?}{=} \int_c^d \psi^*(\omega)$$

其实, φ 和 ψ 所描述者乃是在同一条轨道 γ 上的两种运动。我们可以把 t 和 s 看做标记 γ 上各点的两种坐标 (所以不妨假设下述 $\lambda'(s)$ 和 $\mu'(t)$ 两者到处不为零), 同一点的 t -坐标和 s -坐标之间的坐标变换可表达为

$$t = \lambda(s), \quad s = \mu(t), \quad \lambda'(s) = \frac{dt}{ds}, \quad \mu'(t) = \frac{ds}{dt}$$

由此可见

$$\begin{aligned} \varphi_i'(t) &= \frac{dx_i}{dt} = \frac{dx_i}{ds} \cdot \frac{ds}{dt} = \psi_i'(s) \mu'(t) \\ \psi_i'(s) &= \frac{dx_i}{ds} = \frac{dx_i}{dt} \cdot \frac{dt}{ds} = \varphi_i'(t) \lambda'(s) \end{aligned}$$

所以

$$\begin{aligned}\int_a^b \varphi^*(\omega) &= \int_a^b \sum_{i=1}^n f_i(\varphi(t)) \varphi'_i(t) dt \\ &= \int_c^d \sum_{i=1}^n f_i(\psi(s)) \varphi'_i(t) \lambda'(s) ds \\ &= \int_c^d \sum_{i=1}^n f_i(\psi(s)) \psi'_i(s) ds = \int_a^b \psi^*(\omega)\end{aligned}$$

上述恒为相等的单元积分就定义为一次微分形式 ω 在曲线 γ 上的线积分, 以符号 $\int_\gamma \omega$ 记之。亦即

$$\int_\gamma \omega := \int_a^b \varphi^*(\omega) \quad \left(= \int_c^d \psi^*(\omega) \right)$$

[注]: 上述线积分的处理方式是通过曲线的参数表示把它归于单元积分, 其关键是得要验证它其实是和参数表式的选取无关的。其好处是把定义和计算法一并解决, 而且这种处理法可以直接推广到 k -维曲面积分。

4.2.2 曲面积分

在 \mathbb{R}^n 中的一片 k -维曲面 Σ 具有其局部坐标系 (t_1, \dots, t_k) , 亦即其上各点的坐标都可以表达为 (t_1, \dots, t_k) 的函数

$$x_i = \varphi_i(t_1, \dots, t_k), \quad 1 \leq i \leq n$$

而且对于 Σ 中每点 \mathbf{x} 都存在适当的 $1 \leq i_1 < \dots < i_k \leq n$ 使得

$$\left. \frac{\partial(\varphi_{i_1}, \dots, \varphi_{i_k})}{\partial(t_1, \dots, t_k)} \right|_{\mathbf{x}} \neq 0$$

用几何的语言来说, $x_i = \varphi_i(t_1, \dots, t_k)$, $1 \leq i \leq n$, 所给者乃是 k -维 \mathbf{t} -空间中的一个区域 Ω 到 n -维 \mathbf{x} -空间中的一片 k -维曲面 Σ 上的一对一可微映射

$$\Phi: \Omega \rightarrow \Sigma \subset \mathbb{R}^n$$

当我们只让 t_j 变动而把其他 $(k-1)$ 个自变元取定，则其象点所描述者乃是 \mathbb{R}^n 中的一条参数曲线，其速度向量为

$$\mathbf{v}_j = \left(\frac{\partial \varphi_1}{\partial t_j}, \dots, \frac{\partial \varphi_n}{\partial t_j} \right)$$

由此可见，上述 Φ 的规则性条件 (regularity conditions) 的几何意义乃是 $\{\mathbf{v}_1, \dots, \mathbf{v}_k\}$ 是到处线性无关的。

再者，设 (s_1, \dots, s_k) 是 Σ 上另一个局部坐标系，亦即

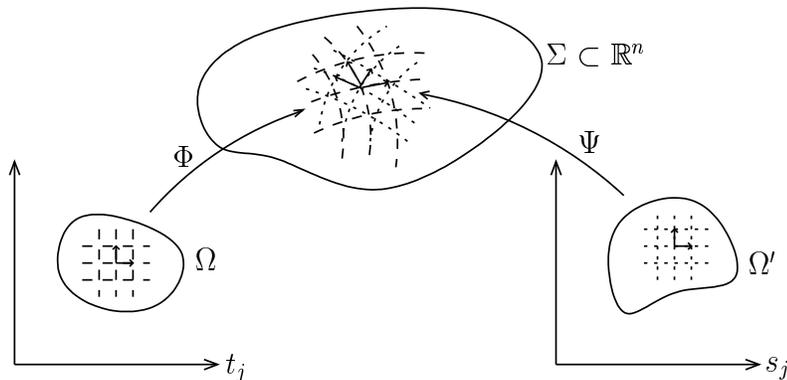
$$\Psi: \Omega' \rightarrow \Sigma \subset \mathbb{R}^n, \quad x_i = \psi_i(s_1, \dots, s_k)$$

是将 k -维 s -空间中的一个区域 Ω' 映射到 Σ 上的一对一可微映射，而且也满足规则性条件。

不难验证

$$\begin{aligned} dx_{i_1} \wedge \dots \wedge dx_{i_k} &= \frac{\partial(\varphi_{i_1}, \dots, \varphi_{i_k})}{\partial(t_1, \dots, t_k)} dt_1 \wedge \dots \wedge dt_k \\ dx_{i_1} \wedge \dots \wedge dx_{i_k} &= \frac{\partial(\psi_{i_1}, \dots, \psi_{i_k})}{\partial(s_1, \dots, s_k)} ds_1 \wedge \dots \wedge ds_k \\ dt_1 \wedge \dots \wedge dt_k &= \frac{\partial(t_1, \dots, t_k)}{\partial(s_1, \dots, s_k)} ds_1 \wedge \dots \wedge ds_k \end{aligned}$$

其中 $\frac{\partial(t_1, \dots, t_k)}{\partial(s_1, \dots, s_k)}$ 乃是 Σ 上两个局部坐标系之间坐标变换式， $t_j = g_j(s)$, $1 \leq j \leq k$, 的 Jacobian.



[图 4-3]

设 $\{f_{i_1, \dots, i_k}(\mathbf{x}), 1 \leq i_1 < \dots < i_k \leq n\}$ 是 $\binom{n}{k}$ 个定义于 \mathbb{R}^n 中某一区域 D 上的连续函数, 则表式

$$\omega = \sum_{1 \leq i_1 < \dots < i_k \leq n} f_{i_1, \dots, i_k}(\mathbf{x}) dx_{i_1} \wedge \dots \wedge dx_{i_k}$$

叫做 D 上的一个 k 次外微分形式 (exterior differential k -form)。

设 $\Phi: \Omega \rightarrow \Sigma \subset D$, 亦即 Σ 是包含在 D 中的一片 k -维曲面。我们将以 $\Phi^*(\omega)$ 表示把 ω 中的 x_i 代以 $\varphi_i(\mathbf{t})$ 之所得, 亦即

$$\Phi^*(\omega) = \sum_{1 \leq i_1 < \dots < i_k \leq n} f_{i_1, \dots, i_k}(\Phi(\mathbf{t})) \frac{\partial(\varphi_{i_1}, \dots, \varphi_{i_k})}{\partial(t_1, \dots, t_k)} dt_1 \wedge \dots \wedge dt_k$$

同样的计算有

$$\Psi^*(\omega) = \sum_{1 \leq i_1 < \dots < i_k \leq n} f_{i_1, \dots, i_k}(\Psi(\mathbf{s})) \frac{\partial(\psi_{i_1}, \dots, \psi_{i_k})}{\partial(s_1, \dots, s_k)} ds_1 \wedge \dots \wedge ds_k$$

由此即可验证

$$\int_{\Omega} \Phi^*(\omega) = \int_{\Omega'} \Psi^*(\omega)$$

恒成立, 其共同积分值就是我们所要定义的 ω 在 Σ 上的积分, 亦即

$$\int_{\Sigma} \omega := \int_{\Omega} \Phi^*(\omega) \quad \left(= \int_{\Omega'} \Psi^*(\omega) \right)$$

4.2.3 例子

(1) 线积分与向量场 :

在线积分中之所积者 (integrand) 乃是一个一次微分形式 $\omega = \sum_{i=1}^n f_i(\mathbf{x}) dx_i$, $\mathbf{x} \in D$ 。我们可以把 $\{f_i(\mathbf{x}), 1 \leq i \leq n\}$ 看做一个向量 $\mathbf{F}(\mathbf{x})$ 的分量, 即令

$$\mathbf{F}(\mathbf{x}) = (f_1(\mathbf{x}), \dots, f_n(\mathbf{x}))$$

则它乃是在 D 中每一点 \mathbf{x} 指定一个于该点 (切空间中) 的一个向量, 称之为 向量场 (vector field)。再者

$$d\mathbf{x} = (dx_1, \dots, dx_n)$$

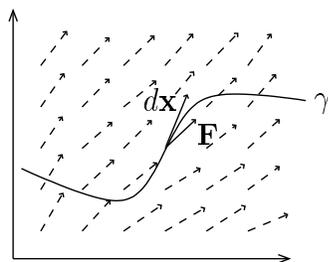
乃是一个将 \mathbf{x} 点移到 $\mathbf{x} + d\mathbf{x}$ 点的微小位移向量，而 ω 实乃 $\mathbf{F}(\mathbf{x})$ 和 $d\mathbf{x}$ 的内积，即

$$\omega = \mathbf{F}(\mathbf{x}) \cdot d\mathbf{x} = \sum_{i=1}^n f_i(\mathbf{x}) dx_i$$

例如在物理学中重要的重力场，电场和磁场的数理表达都是向量场的实例，而其所相应的一次微分形式所表达者就是在这种力场的作用下，微小位移所作的「功」(work)，亦即

$$dW = \mathbf{F}(\mathbf{x}) \cdot d\mathbf{x} \quad (= \omega)$$

由此可见线积分 $\int_{\gamma} \omega$ 的实质内涵也就是上述内积在无限细分之下的总和。当 $\mathbf{F}(\mathbf{x})$ 是表达某种力的向量场时，则 $\int_{\gamma} \omega$ 的物理意义就是沿 γ 的运动所作的功。



[图 4-4]

(2) 设 $U(\mathbf{x})$ 是开区域 D 上的一阶连续可微函数，则向量场

$$\nabla U(\mathbf{x}) := \left(\frac{\partial U}{\partial x_1}(\mathbf{x}), \dots, \frac{\partial U}{\partial x_n}(\mathbf{x}) \right)$$

叫做 U 的梯度向量场 (the gradient vector field of U)。再者， $\nabla U(\mathbf{x})$ 所相应的一次微分形式

$$\omega = \nabla U(\mathbf{x}) \cdot d\mathbf{x} = \sum_{i=1}^n \frac{\partial U}{\partial x_i}(\mathbf{x}) \cdot dx_i$$

其实就是 U 的全微分 (total differential, 亦即 dU)。

设 γ 是 D 中的一条参数曲线，即以

$$\gamma(t) = (\varphi_1(t), \dots, \varphi_n(t)), \quad a \leq t \leq b$$

为其参数式者。令 $U(t)$ 为将 x_i 以 $\varphi_i(t)$ 代入之所得，即

$$U(t) := U(\varphi_1(t), \dots, \varphi_n(t))$$

则有

$$U'(t) = \sum_{i=1}^n \frac{\partial U}{\partial x_i} \Big|_{\gamma(t)} \varphi_i'(t)$$

所以由线积分的定义和微积分基本定理，可得

$$\begin{aligned} \int_{\gamma} \omega &= \int_a^b \varphi^*(\omega) = \int_a^b U'(t) dt = U(b) - U(a) \\ &= U(\varphi_1(b), \dots, \varphi_n(b)) - U(\varphi_1(a), \dots, \varphi_n(a)) \end{aligned}$$

由此可见，当 ω 所相应的向量场乃是一个函数 $U(\mathbf{x})$ 的梯度向量场时（亦即 $\omega = dU$ ），线积分 $\int_{\gamma} \omega$ 的值其实仅仅和曲线 γ 的端点有关，它的值就是 $U(\gamma(b)) - U(\gamma(a))$ 。

反之，设 $\omega = \sum_{i=1}^n f_i(\mathbf{x}) dx_i$ 是一个定义在开的一个连通区域 D 上的一次微分形式，而且 $\int_{\gamma} \omega$ 对于 D 中任给可微曲线 γ 都仅仅和 γ 的端点有关。我们不妨取定 D 中任给一点 $\mathbf{a} = (a_1, \dots, a_n)$ 作为基点，令

$$U(\mathbf{x}) = \int_{\gamma} \omega$$

其中 γ 是任何一条由基点 \mathbf{a} 到 \mathbf{x} 点的可微曲线（由所设，上述线积分之值仅仅和 \mathbf{x} 点有关）。不难验证，如此定义的 $U(\mathbf{x})$ 即有

$$\frac{\partial U}{\partial x_i} \Big|_{\mathbf{x}} = f_i(\mathbf{x}), \quad \text{亦即} \quad \omega = \nabla U(\mathbf{x}) \cdot d\mathbf{x} = dU$$

由此可见，线积分 $\int_{\gamma} \omega$ 仅仅和其端点有关的充要条件乃是 ω 等于某一个一阶连续可微的函数 U 的全微分，亦即

$$\omega = \nabla U(\mathbf{x}) \cdot d\mathbf{x}$$

(3) 对于一个给定的一次微分形式 $\omega = \sum_{i=1}^n f_i(\mathbf{x}) dx_i$ 其线积分 $\int_{\gamma} \omega$ 是否仅仅和其端点有关？上述所讨论的充要条件乃是一个十分难于检验的条

件（除非能够一眼看到某个函数 U ，它满足 $\left. \frac{\partial U}{\partial x_i} \right|_{\mathbf{x}} = f_i(\mathbf{x}), 1 \leq i \leq n$ ）。是否还有另一种易于检验的条件呢？在 $f_i(\mathbf{x})$ 也都是一阶连续可微的情形，假如存在著 $U(\mathbf{x})$ 使得

$$\left. \frac{\partial U}{\partial x_i} \right|_{\mathbf{x}} = f_i(\mathbf{x}), \quad 1 \leq i \leq n$$

则有

$$\left. \frac{\partial f_i}{\partial x_j} \right|_{\mathbf{x}} = \left. \frac{\partial^2 U}{\partial x_j \partial x_i} \right|_{\mathbf{x}} = \left. \frac{\partial^2 U}{\partial x_i \partial x_j} \right|_{\mathbf{x}} = \left. \frac{\partial f_j}{\partial x_i} \right|_{\mathbf{x}}$$

所以

$$\left. \frac{\partial f_i}{\partial x_j} \right|_{\mathbf{x}} = \left. \frac{\partial f_j}{\partial x_i} \right|_{\mathbf{x}}, \quad 1 \leq i, j \leq n$$

乃是一种易于检验的必要条件。但是它是否业已充分呢？假若果真业已充分，此事岂不妙哉！不然则应该去探求某种补充条件，使得上述必要条件再加上它就能够保证 ω 乃是一个全微分。首先，让我们先来看一个简单但是颇具重要性的例子，它说明了上述必要条件在一般情形并非充分。令

$$\omega_0 = \frac{-x_2 dx_1 + x_1 dx_2}{x_1^2 + x_2^2}, \quad D = \mathbb{R}^2 \setminus \{(0, 0)\}$$

亦即

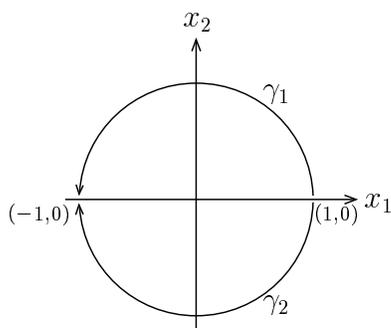
$$f_1 = \frac{-x_2}{x_1^2 + x_2^2}, \quad f_2 = \frac{x_1}{x_1^2 + x_2^2}$$

由此易见

$$\frac{\partial f_1}{\partial x_2} = \frac{\partial f_2}{\partial x_1} = \frac{x_2^2 - x_1^2}{(x_1^2 + x_2^2)^2}$$

但如 [图 4-5] 所示：

$$\begin{aligned} \gamma_1 : x_1 &= \cos t, x_2 = \sin t, \quad 0 \leq t \leq \pi \\ \gamma_2 : x_1 &= \cos t, x_2 = -\sin t, \quad 0 \leq t \leq \pi \\ \int_{\gamma_1} \frac{-x_2 dx_1 + x_1 dx_2}{x_1^2 + x_2^2} &= \int_0^\pi dt = \pi \\ \int_{\gamma_2} \frac{-x_2 dx_1 + x_1 dx_2}{x_1^2 + x_2^2} &= \int_0^\pi -dt = -\pi \end{aligned}$$



[图 4-5]

由此可见, 上述 ω_0 的线积分并非仅仅和积分曲线 γ 的端点有关。在这里, $\int_{\gamma} \omega_0$ 之值究竟和积分曲线的什么才是真正相关的呢? 这就是一个值得实事求是去探索的问题。

有鉴于上述 ω_0 的 f_1 和 f_2 都含有旋转不变的分母 $(x_1^2 + x_2^2)$, 把它改用极坐标 (r, θ) 来表达, 肯定会得到某种简化。即令

$$x_1 = r \cos \theta, \quad x_2 = r \sin \theta$$

代入 ω_0 , 即得

$$\begin{aligned} \omega_0 &= \frac{1}{r^2} \{-r \sin \theta (\cos \theta dr - r \sin \theta d\theta) + r \cos \theta (\sin \theta dr + r \cos \theta d\theta)\} \\ &= \frac{1}{r^2} \{r^2 (\sin^2 \theta + \cos^2 \theta) d\theta\} = d\theta \end{aligned}$$

由此易见 $\int_{\gamma} \omega$ 其实只和积分曲线在 θ 的改变量有关。设 $\gamma(t) = (r(t), \theta(t))$, $a \leq t \leq b$, 是其极坐标的参数式, 则

$$\int_{\gamma} \omega_0 = \int_{\theta(a)}^{\theta(b)} d\theta = \int_a^b \theta'(t) dt = \theta(b) - \theta(a)$$

再者, 设 γ_1 和 γ_2 是两条具有同样的始、终点的曲线, 则 γ_1 和 γ_2 的反向曲线就连接成一条闭曲线 γ_0 , 显然有

$$\int_{\gamma_1} \omega_0 - \int_{\gamma_2} \omega_0 = \int_{\gamma_0} \omega_0 = 2\pi n(\gamma_0)$$

其中 $n(\gamma_0)$ 是闭曲线绕著原点的有向绕数 (winding number)。

其实，我们还可以把上述有意思的现象作下述推广：

设 $U(x_1, x_2)$ 是一个在 $\mathbb{R}^2 \setminus \{(0, 0)\}$ 上二阶连续可微函数。令

$$\omega = c\omega_0 + dU = \left(\frac{-cx_2}{x_1^2 + x_2^2} + \frac{\partial U}{\partial x_1}\right)dx_1 + \left(\frac{cx_1}{x_1^2 + x_2^2} + \frac{\partial U}{\partial x_2}\right)dx_2$$

则显然有：

(i) ω 在 $\mathbb{R}^2 \setminus \{(0, 0)\}$ 上到处满足 $\frac{\partial f_1}{\partial x_2} = \frac{\partial f_2}{\partial x_1}$ 这个必要条件；

(ii) 对于 $\mathbb{R}^2 \setminus \{(0, 0)\}$ 中任给闭曲线 γ_0

$$\int_{\gamma_0} \omega = 2\pi n(\gamma_0)c$$

在此，我们还可以进一步提出下述很值得我们作更加深入的探索的问题：
：是否在 $\mathbb{R}^2 \setminus \{(0, 0)\}$ 上满足 $\frac{\partial f_1}{\partial x_2} = \frac{\partial f_2}{\partial x_1}$ 这个必要条件的一阶连续可微一次微分形式 ω 也都具有

$$\int_{\gamma_0} \omega = 2\pi n(\gamma_0)c \quad (\text{其中 } c \text{ 是一个随 } \omega \text{ 而定的常数})$$

再者，是否满足上述条件的一阶连续可微一次微分形式 ω 其实都可以表达成 $\omega = c\omega_0 + dU$ 这种形式？

(4) 三维欧氏空间的面积分与向量场：

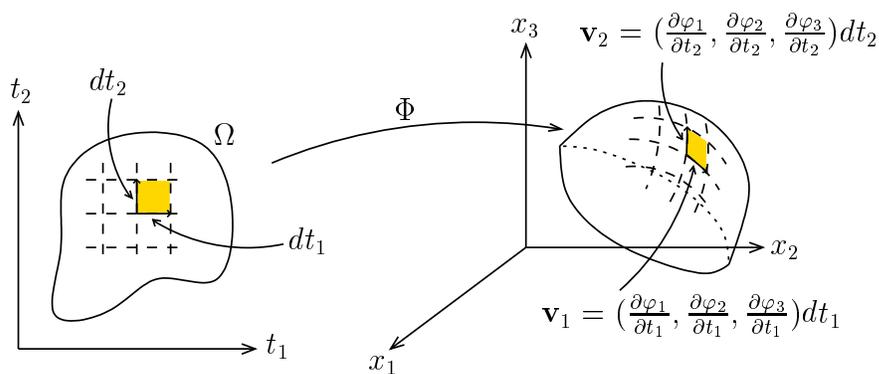
在 \mathbb{R}^3 中某一个连通开子集 D 上定义的一个二次外微分形式，都可以写成

$$\omega = f_1(\mathbf{x})dx_2 \wedge dx_3 + f_2(\mathbf{x})dx_3 \wedge dx_1 + f_3(\mathbf{x})dx_1 \wedge dx_2$$

由此可见它也有一个相应的 D 上的向量场，即

$$\omega \longleftrightarrow \mathbf{F}(\mathbf{x}) = (f_1(\mathbf{x}), f_2(\mathbf{x}), f_3(\mathbf{x}))$$

如 [图 4-6] 所示， Σ 是 D 中的一片曲面， $\Phi: \Omega \rightarrow \Sigma$ 是 Σ 的一个取定参数表达式。



[图 4-6]

以 dt_1 和 dt_2 表示在 \mathbf{t} 点分别和 t_1 -轴和 t_2 -轴平行的微小位移。在映射 Φ 之下，其相应的切向量分别就是

$$\mathbf{v}_1 = \left(\frac{\partial \varphi_1}{\partial t_1}, \frac{\partial \varphi_2}{\partial t_1}, \frac{\partial \varphi_3}{\partial t_1} \right) dt_1 \quad \text{和} \quad \mathbf{v}_2 = \left(\frac{\partial \varphi_1}{\partial t_2}, \frac{\partial \varphi_2}{\partial t_2}, \frac{\partial \varphi_3}{\partial t_2} \right) dt_2$$

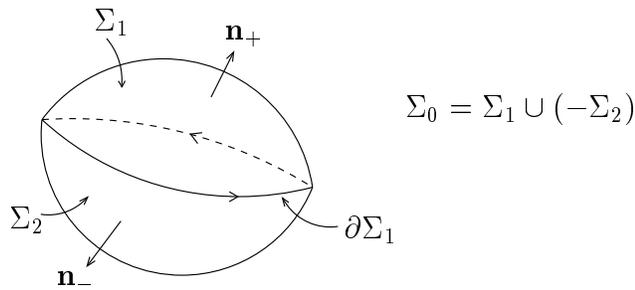
由此可见 $\mathbf{v}_1 \wedge \mathbf{v}_2$ 所表达者乃是曲面 Σ 上相应于 Ω 中的坐标小格（亦即 $\| (dt_1, dt_2) \|$ ）的那个微片的面积和方向的二阶逼近。

设想向量场 $\mathbf{F}(\mathbf{x})$ 乃是空间中一个给定的流动 (flow) 的速度向量场，则它在单位时间内通过 Σ 的上述微片的液体的二阶逼近就等于

$$\begin{aligned} & \mathbf{F}(\mathbf{x}) \wedge \mathbf{v}_1 \wedge \mathbf{v}_2 \\ &= \left\{ f_1(\Phi(\mathbf{t})) \frac{\partial(\varphi_2, \varphi_3)}{\partial(t_1, t_2)} + f_2(\Phi(\mathbf{t})) \frac{\partial(\varphi_3, \varphi_1)}{\partial(t_1, t_2)} + f_3(\Phi(\mathbf{t})) \frac{\partial(\varphi_1, \varphi_2)}{\partial(t_1, t_2)} \right\} dt_1 \wedge dt_2 \end{aligned}$$

由此可见，曲面积分 $\int_{\Sigma} \omega$ 所表达者，乃是以其相应的 $\mathbf{F}(\mathbf{x})$ 为其流速向量场在单位时间内通过曲面 Σ 的总流量是也。

(5) 设 Σ_1 是一个具有逐段平滑的边界的定向曲面，则它在每一点的两个法向 (normal directions) 也就有了其正、负之别。再者，它的边界也就有其相应的定向（如 [图 4-7] 所示）。



[图 4-7]

再者，如 [图 4-7] 所示，设 Σ_2 是另一个定向曲面，其定向边界和 $\partial\Sigma_1$ 相同，则 Σ_1 和 Σ_2 的反向者就组合成一个空间中的定向闭曲面 (oriented closed surface) Σ_0 ，其正法向不是都向外就是都向里。再者，易见

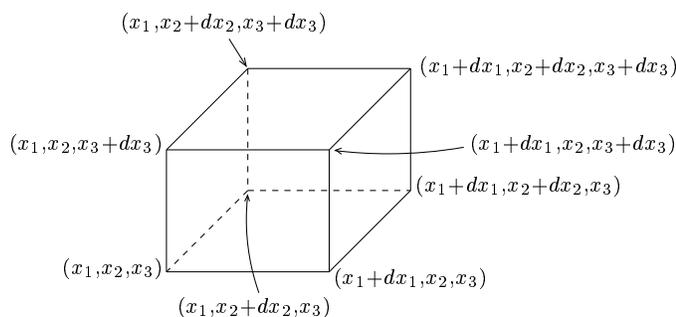
$$\int_{\Sigma_0} \omega = \int_{\Sigma_1} \omega - \int_{\Sigma_2} \omega$$

由此可见，以 $\mathbf{F}(\mathbf{x})$ 为其流速向量场的流动，其单位时间流过 Σ_1 者等于其流过 Σ_2 者的条件也就是流入 Σ_0 者等于流出 Σ_0 者。换句话说，亦即上述流体包含在那个以 Σ_0 为其边界的三维区域之内的液体之总量是守恒的。

有鉴于某一个区域之内的液体之总量乃是其所含的局部各小分块之内的液体之总和，假若流体对于其所含的每一小分块都是守恒的，则它当然对于以 Σ_0 为其边界的这一整块也必然是守恒的。再者，若这种局部化的守恒性普遍成立，则显然也就有全局化的守恒性对于那种其所有内点皆具有局部守恒性的三维区域皆成立。

(6) 局部守恒性的微分条件：

由上可见，局部守恒性乃是一种自然而且有用的条件，而微分法则求得其解析条件式的有效方法，其具体做法乃是分析如 [图 4-8] 所示的坐标小块的守恒性在 dx_1 , dx_2 和 dx_3 都无限缩小的极限式。其实，我们只要计算流体在单位时间流进和流出这个坐标小分块的流量之差的三阶逼近（亦即略去高于三阶之微量）。



[图 4-8]

上述坐标小方块共有上、下；右、左；前、後六个面。其方向和有向面积可以分别用 $dx_1 \wedge dx_2$, $-dx_1 \wedge dx_2$; $dx_2 \wedge dx_3$, $-dx_2 \wedge dx_3$; $dx_3 \wedge dx_1$, $-dx_3 \wedge dx_1$ 表达之。由此可见，其上、下两面出、入之差乃是

$$\begin{aligned} & f_3(x_1, x_2, x_3 + dx_3)dx_1 \wedge dx_2 - f_3(x_1, x_2, x_3)dx_1 \wedge dx_2 \\ & \equiv \left. \frac{\partial f_3}{\partial x_3} \right|_{\mathbf{x}} dx_1 \wedge dx_2 \wedge dx_3 \quad (\text{mod 高于三阶者}) \end{aligned}$$

其右、左两面出、入之差乃是

$$\begin{aligned} & f_1(x_1 + dx_1, x_2, x_3)dx_2 \wedge dx_3 - f_1(x_1, x_2, x_3)dx_2 \wedge dx_3 \\ & \equiv \left. \frac{\partial f_1}{\partial x_1} \right|_{\mathbf{x}} dx_1 \wedge dx_2 \wedge dx_3 \quad (\text{mod 高于三阶者}) \end{aligned}$$

其前、後两面出、入之差乃是

$$\begin{aligned} & f_2(x_1, x_2 + dx_2, x_3)dx_3 \wedge dx_1 - f_2(x_1, x_2, x_3)dx_3 \wedge dx_1 \\ & \equiv \left. \frac{\partial f_2}{\partial x_2} \right|_{\mathbf{x}} dx_1 \wedge dx_2 \wedge dx_3 \quad (\text{mod 高于三阶者}) \end{aligned}$$

所以将出入上述坐标小分块的流量除以其体积元素 $dx_1 \wedge dx_2 \wedge dx_3$ 後再求其在 dx_1, dx_2, dx_3 都趋于零的极限值，其所得即为

$$\left. \left(\frac{\partial f_1}{\partial x_1} + \frac{\partial f_2}{\partial x_2} + \frac{\partial f_3}{\partial x_3} \right) \right|_{\mathbf{x}}$$

称之为所给流动在 \mathbf{x} 点的发散量 (divergence)，而发散量在该点等于 0 则是流动在该点的局部化守恒性的微分条件式，亦即

$$(*) \quad \left. \left(\frac{\partial f_1}{\partial x_1} + \frac{\partial f_2}{\partial x_2} + \frac{\partial f_3}{\partial x_3} \right) \right|_{\mathbf{x}} = 0$$

若所给的流动（或其所相应的一阶可微二次外微分形式 ω ）在某一开域 D 上到处满足上述条件（*），则称之为在 D 上无发散（divergence free）的流动（或二次外微分形式）。

(7) n -维欧氏空间的超曲面积分与向量场

在 (4) 中对于 $n = 3$ 的情形的研讨，可以直截了当地推广到一般的情形。首先，定义于 \mathbb{R}^n 中某个开子集 D 上的一个 $(n-1)$ -次外微分形式，都可以写成

$$\begin{aligned} \omega = & f_1(\mathbf{x})dx_2 \wedge \dots \wedge dx_n - f_2(\mathbf{x})dx_1 \wedge dx_3 \wedge \dots \wedge dx_n \\ & + \dots + (-1)^{n+1}f_n(\mathbf{x})dx_1 \wedge \dots \wedge dx_{n-1}, \quad \mathbf{x} \in D \end{aligned}$$

所以它也同样地有其相应的 D 上的向量场，即为

$$\mathbf{F}(\mathbf{x}) = (f_1(\mathbf{x}), f_2(\mathbf{x}), \dots, f_n(\mathbf{x}))$$

设 Σ 是 D 中的一片超（亦即 $(n-1)$ -维）曲面， $\Phi: \Omega \rightarrow \Sigma$ 是 Σ 的一个取定参数表达式，即

$$x_i = \varphi_i(\mathbf{t}), \quad 1 \leq i \leq n, \quad \mathbf{t} \in \Omega$$

而且满足规则性条件

$$\mathbf{v}_j = \left(\frac{\partial \varphi_1}{\partial t_j}, \dots, \frac{\partial \varphi_n}{\partial t_j} \right) dt_j, \quad 1 \leq j \leq n-1$$

总是线性无关的。令 $//(dt_1, \dots, dt_{n-1})$ 是 Ω 中的坐标小分格，则其在 Φ 之下的映象 $\Phi(//(dt_1, \dots, dt_{n-1}))$ 的 $(n-1)$ -阶（线性）逼近就是 $\mathbf{v}_1 \wedge \dots \wedge \mathbf{v}_{n-1}$ 。所以其相应的 n -维流动 (n -dimensional flow) 在单位时间内通过超曲面 Σ 的上述微片的流量的 $(n-1)$ -阶逼近就等于

$$\begin{aligned} \mathbf{F}(\mathbf{x}) \wedge \mathbf{v}_1 \wedge \dots \wedge \mathbf{v}_{n-1} = & \left\{ f_1(\Phi(\mathbf{t})) \frac{\partial(\varphi_2, \dots, \varphi_n)}{\partial(t_1, \dots, t_{n-1})} \right. \\ & \left. + f_2(\Phi(\mathbf{t})) \frac{\partial(\varphi_1, \varphi_3, \dots, \varphi_n)}{\partial(t_1, \dots, t_{n-1})} + \dots + f_n(\Phi(\mathbf{t})) \frac{\partial(\varphi_1, \dots, \varphi_{n-1})}{\partial(t_1, \dots, t_{n-1})} \right\} dt_1 \wedge \dots \wedge dt_{n-1} \end{aligned}$$

由此可见， $(n-1)$ -维曲面积分 $\int_{\Sigma} \omega$ 所表达者乃是以其相应的 $\mathbf{F}(\mathbf{x})$ 为其流速向量场的 n -维流动在单位时间内通过超曲面 Σ 的总流量是也。

(8) 同样地我们也可以验证 n -维流动的局部守恒性的微分条件乃是其发散量等于 0, 即

$$\left(\sum_{i=1}^n \frac{\partial f_i}{\partial x_i} \right) \Big|_{\mathbf{x}} = 0$$

其验证是 $n=3$ 的情形直截了当的推广, 留作习题。

4.2.4 习题

(1) 由笛氏坐标 (x_1, x_2) 和极坐标 (r, θ) 之间的变换公式, 亦即

$$x_1 = r \cos \theta, \quad x_2 = r \sin \theta \quad (x_1^2 + x_2^2 = r^2)$$

微分即得

$$\begin{cases} dx_1 = \cos \theta dr - r \sin \theta d\theta \\ dx_2 = \sin \theta dr + r \cos \theta d\theta \end{cases}$$

把它想成是 dr 和 $d\theta$ 的线性方程组, 试求解 $d\theta$ 。

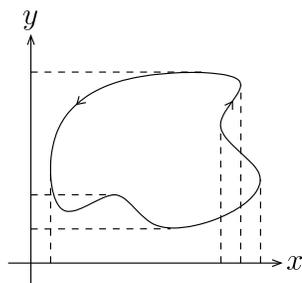
(2) 试构造一个在 $D = \mathbb{R}^2 \setminus \{(\pm 1, 0)\}$ 上一阶连续可微的一次微分形式, 它对于 D 中的任给闭曲线的线积分之值为

$$\int_{\gamma} \omega = 2\pi c_1 n_1 + 2\pi c_2 n_2$$

其中 c_1, c_2 是任给常数而 n_1 和 n_2 则分别是 γ 对于 $(1, 0)$ 和 $(-1, 0)$ 的定向绕数。

(3) 试问是否可以把你的构造法推广到平面上任给有限个点的情形?

(4) 如 [图 4-9] 所示, $\Omega \subset \mathbb{R}^2$ 是平面上的区域, 其边界 $\partial\Omega$ 是一个一阶连续可微的简单闭曲线。



[图 4-9]

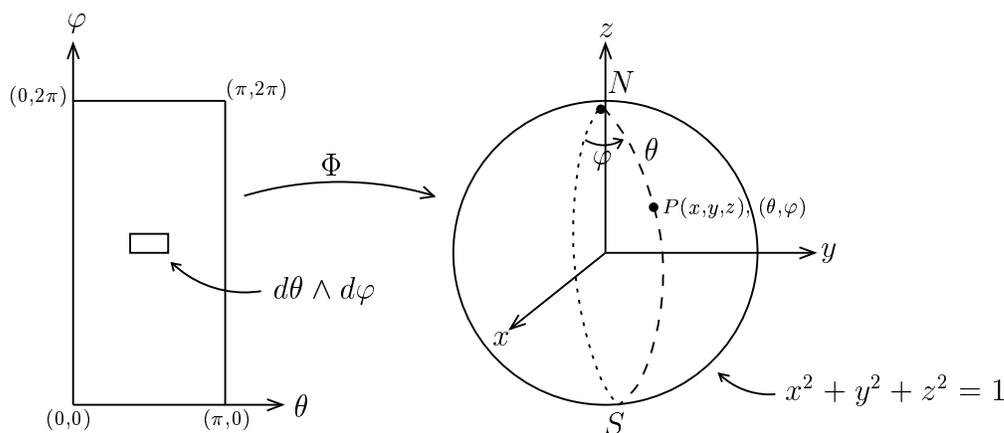
试以图解说明 $\int_{\partial\Omega} -ydx$ 和 $\int_{\partial\Omega} xdy$ 都是 Ω 的定向面积。(注意 $\partial\Omega$ 的定向和 Ω 的定向是密切相关的。如 [图 4-9] 所示者, 是 Ω 取正向的情形。) 由此可见, 也有

$$\int_{\partial\Omega} \frac{1}{2}(xdy - ydx) = \text{Area}(\Omega)$$

(5) 设 Ω 是 $\frac{x^2}{a^2} + \frac{y^2}{b^2} \leq 1$ 所描述的椭圆内部, 则 $\partial\Omega$ 可以用 $x = a \cos \theta$, $y = b \sin \theta$, $0 \leq \theta \leq 2\pi$, 加以参数描述。试用它来计算

$$\text{Area}(\Omega) = \int_{\partial\Omega} \frac{1}{2}(xdy - ydx)$$

(6) 球面上的极坐标与其面积元素 (area element): 球面具有极为丰富的对称性。例如它对于其任给「直径」是旋转对称的, 一如地球的自转乃是对于南北极连线为轴的旋转对称性的具体表现。如 [图 4-10] 所示, 我们取球的半径为单位长, 则球面上任给南北极之外的点 P 可以用 (θ, φ) 作为其坐标, 其中 θ 是 P 点到北极的球面距离而 φ 则是过 P 点的经线和选定的基准经线的夹角。



[图 4-10]

令 (x, y, z) 为 P 点在 \mathbb{R}^3 中的笛氏坐标, 则有

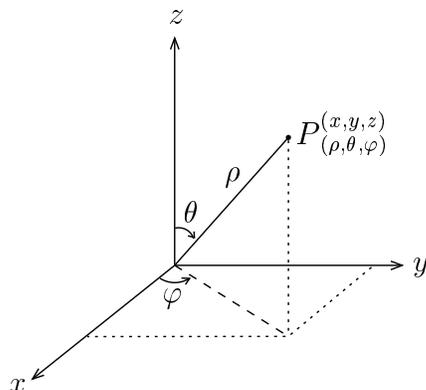
$$\Phi: \quad x = \sin \theta \cos \varphi, \quad y = \sin \theta \sin \varphi, \quad z = \cos \theta$$

试求在 (θ, φ) 坐标面中的小格 $//(d\theta, d\varphi)$ 的 Φ 之映象的球面面积的二阶逼近 (称之为球面坐标的面积元素)。

(7) 试用习题 (6) 中所得的结果求单位球面的面积, 并将你的计算和当年 Archimedes' 的球面面积公式之论证作一比较分析 (参看第四册 §2.3 的例 (3))。

(8) \mathbb{R}^3 的球坐标 (ρ, θ, φ) 和笛氏坐标 (x, y, z) 之间的变换公式是 (如 [图 4-11] 所示) :

$$x = \rho \sin \theta \cos \varphi, \quad y = \rho \sin \theta \sin \varphi, \quad z = \rho \cos \theta$$



[图 4-11]

试改用球坐标计算

$$\int_{\mathbb{R}^3} e^{-(x^2+y^2+z^2)} dx \wedge dy \wedge dz$$

(9) 试求 $dx \wedge dy$, $dy \wedge dz$, $dz \wedge dx$ 改用球坐标 (ρ, θ, φ) 的各自表式。然后把它们想成 $d\rho \wedge d\theta$, $d\rho \wedge d\varphi$, $d\theta \wedge d\varphi$ 的三元联立方程组。用代数方法求解 $\sin \theta d\theta \wedge d\varphi$ (它乃是单位球面上的面积元素, 参考习题 (6)), 其所得乃是定义于 $\mathbb{R}^3 \setminus \{(0, 0, 0)\}$ 上的二次外微分形式, 即有:

$$\omega_0 = f_1(x, y, z) dy \wedge dz + f_2(x, y, z) dz \wedge dx + f_3(x, y, z) dx \wedge dy$$

试说明 ω_0 所相应的向量场的几何意义。

(10) 上述二次微分形式 ω_0 乃是习题 (1) 中 $d\theta$ 的自然推广。试验证其相应的向量场 $\mathbf{F}(x, y, z)$ 乃是无发散量 (divergence free) 者。

(11) 设 Ω 是 \mathbb{R}^3 中的区域, 其边界是由一片片可微曲面组合而成者。试证

$$\int_{\partial\Omega} \omega_0 = \begin{cases} 0 & \text{若 } (0, 0, 0) \notin \Omega \\ 4\pi & \text{若 } (0, 0, 0) \in \Omega \end{cases}$$

(12) 试构造类似于习题 (3) 在三维中，二次微分形式的推广。

4.3 外微分和微积分基本定理的高维推广

在本章的第一节中，我们研讨了一个 n 个自变元的函数 $f(\mathbf{x})$ 在给定的有界闭区域 D 上的积分，即 $\int_D f(\mathbf{x}) dx_1 \wedge \dots \wedge dx_n$ 。接著在第二节中，我们进而把积分推广到多元、多关系的范畴，如线积分、面积分和 k -维的面积分等等。在一般情况，其所要积者 (integrand) 乃是一个定义于 n -维开子集 D 上的 k -次外微分形式 ω ，而其所积之范围 (domain of integration) Σ 则是 D 中的一种定向的 k -维几何事物 (k -dimensional geometric objects)，通常是一个 k -维的定向曲面或由好些片 k -维定向曲面接合 (match up) 而成者。总之，一个 k -次外微分形式 ω 在这种定向的 k -维几何事物 Σ 上的积分可以归于 ω 在其组成之各片 k -维定向曲面上的积分之总和而定义之。再者， D 中一个具有局部坐标化的 k -维定向曲面 Σ ，其实乃是一个由 k -维 \mathbf{t} -空间的一个区域 Ω 到 D 中的映射，即有

$$\begin{array}{ccc} \mathbb{R}^k \supset \Omega & \xrightarrow{\Phi} & \Sigma \subset D \subset \mathbb{R}^n \\ \Downarrow & & \Downarrow \\ & & \mathbf{t} \longmapsto \Phi(\mathbf{t}) = (\varphi_1(\mathbf{t}), \dots, \varphi_n(\mathbf{t})) \end{array}$$

而且满足规则性条件。在上一节我们证明了：

$$\int_{\Omega} \Phi^*(\omega) = \int_{\Sigma} \Psi^*(\omega)$$

而其共同值则定义为 $\int_{\Sigma} \omega$ 。所以在本质上，一般的 k -维积分 $\int_{\Sigma} \omega$ 乃是通过局部坐标映射 $\Phi: \Omega \rightarrow \Sigma$ 把它归于第一节中所讨论的 k 个自变元的函数的积分，亦即

$$\int_{\Sigma} \omega := \int_{\Omega} \Phi^*(\omega)$$

本节的中心课题是如何把微积分基本定理推广到多元、多关系的范畴。在单元微积分的基本定理乃是

$$\int_a^b df = f(b) - f(a)$$

df 乃是一个单元的一次微分形式，而上述积分则是它和一维的有向线段 $[a, b]$ 相结合之所得。在此不妨采用线性代数中讨论对偶空间 (dual space) 所用的符号，即用： $\langle df, [a, b] \rangle$ 表达 $\int_a^b df$ 。采用这种符号意味着我们将试著把一次微分形式和一维有向线段看成互相对偶的事物，而 $\int_a^b df$ 则就是这两种互相对偶的事物相互结合所产生之值。让我们把这种对偶性的想法下推到 0-次和 0-维的情形：一个函数 $f(x)$ 可以看做一个 0-次微分形式，而一个带有正、负号的点集组合则可以想成是一种 0-维定向几何事物。例如有向线段 $[a, b]$ 的边界就应该是 $\{b\} - \{a\}$ ，即

$$\partial[a, b] = \{b\} - \{a\}$$

如此顺理成章，我们应该把一个函数 $f(x)$ 和带有正、负号的点集相结合之所得定义为 $f(x)$ 在这些点上的函数值的代数和。例如

$$\langle f, \partial[a, b] \rangle = f(b) - f(a)$$

采用这种隐含对偶性的符号来表达微积分基本定理，其形式即为

$$\langle df, [a, b] \rangle = \langle f, \partial[a, b] \rangle$$

用线性代数中对偶空间的观点来看，则是求全微分运算“ d ”和求边界运算“ ∂ ”乃是互相对偶的线性映射。

上述讨论其实只是采取一种新的观点来表达原来的微积分基本定理，其好处是促使我们想到在高维的微积分中也应该有一种求微分的运算“ d ”使得

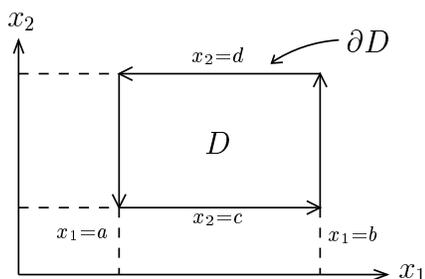
$$\langle "d"\omega, \Sigma \rangle = \int_{\Sigma} "d"\omega = \int_{\partial\Sigma} \omega = \langle \omega, \partial\Sigma \rangle$$

这也就是我们现在要去探索的外微分 (exterior differentiation) 和微积分基本定理的高维推广。

首先，让我们来分析一下 $n = k = 2$ 而且 D 是一个坐标矩形的情形。设 $\omega = f(x_1, x_2)dx_2$ ，而且 $f(x_1, x_2)$ 一阶连续可微。如 [图 4-12] 所示

，因为 ω 不含有 dx_1 ,

$$\begin{aligned} \int_{\partial D} f(x_1, x_2) dx_2 &= \int_c^d f(b, x_2) dx_2 + \int_d^c f(a, x_2) dx_2 \\ &= \int_c^d [f(b, x_2) - f(a, x_2)] dx_2 = \int_c^d \left\{ \int_a^b \frac{\partial f}{\partial x_1} \Big|_{\mathbf{x}} dx_1 \right\} dx_2 \\ &= \int_D \frac{\partial f}{\partial x_1} \Big|_{\mathbf{x}} dx_1 \wedge dx_2 \end{aligned}$$



[图 4-12]

接著，再来看一下 $\omega = g(x_1, x_2) dx_1$ 的情形。因为 ω 不含有 dx_2 ,

$$\begin{aligned} \int_{\partial D} g(x_1, x_2) dx_1 &= \int_a^b g(x_1, c) dx_1 + \int_b^a g(x_1, d) dx_1 \\ &= \int_a^b [g(x_1, c) - g(x_1, d)] dx_1 = \int_a^b dx_1 \left\{ \int_c^d - \frac{\partial g}{\partial x_2} \Big|_{\mathbf{x}} dx_2 \right\} \\ &= \int_D - \frac{\partial g}{\partial x_2} \Big|_{\mathbf{x}} dx_1 \wedge dx_2 \end{aligned}$$

结合上述两种特殊情形的分析，即得下述公式，即

$$\int_{\partial D} f_1(\mathbf{x}) dx_1 + f_2(\mathbf{x}) dx_2 = \int_D \left\{ - \frac{\partial f_1}{\partial x_2} + \frac{\partial f_2}{\partial x_1} \right\} \Big|_{\mathbf{x}} dx_1 \wedge dx_2$$

其中 D 乃是一个坐标矩形。由此可见，我们应该把一个一阶连续可微的一次微分形式 $\omega = \sum_{i=1}^2 f_i(\mathbf{x}) dx_i$ 的外微分定义为

$$d\omega := \left(- \frac{\partial f_1}{\partial x_2} + \frac{\partial f_2}{\partial x_1} \right) \Big|_{\mathbf{x}} dx_1 \wedge dx_2$$

则公式

$$\langle d\omega, D \rangle = \langle \omega, \partial D \rangle$$

至少在 D 是坐标矩形的情形是成立的。再者，易见

$$\begin{aligned} & df_1 \wedge dx_1 + df_2 \wedge dx_2 \\ &= \left(\frac{\partial f_1}{\partial x_1} dx_1 + \frac{\partial f_1}{\partial x_2} dx_2 \right) \wedge dx_1 + \left(\frac{\partial f_2}{\partial x_1} dx_1 + \frac{\partial f_2}{\partial x_2} dx_2 \right) \wedge dx_2 \\ &= \frac{\partial f_1}{\partial x_2} dx_2 \wedge dx_1 + \frac{\partial f_2}{\partial x_1} dx_1 \wedge dx_2 = \left(-\frac{\partial f_1}{\partial x_2} + \frac{\partial f_2}{\partial x_1} \right) dx_1 \wedge dx_2 = d\omega \end{aligned}$$

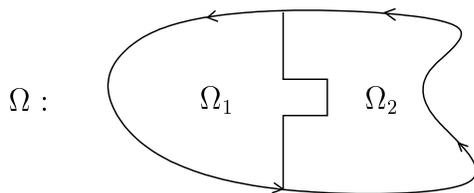
【Green's 定理】：设 $\Omega \subset \mathbb{R}^2$ 是一个有界闭区域，其边界 $\partial\Omega$ 是由逐段可微曲线所组成者；

$$\omega = f_1(\mathbf{x})dx_1 + f_2(\mathbf{x})dx_2$$

是定义于一个包含 Ω 的开区域 D 上的一阶连续可微一次微分形式，则

$$(*) \quad \int_{\partial\Omega} \omega = \int_{\Omega} d\omega, \quad d\omega = \left(-\frac{\partial f_1}{\partial x_2} + \frac{\partial f_2}{\partial x_1} \right) dx_1 \wedge dx_2$$

证明：首先，前述分析业已证明了 Ω 本身就是坐标矩形的特殊情形。再者，设 Ω 可以分解成 Ω_1 和 Ω_2 如 [图 4-13] 所示的接合：



[图 4-13]

则 $\partial\Omega_1$ 和 $\partial\Omega_2$ 在它们接合的共同边界上的定向相反，所以具有下述线积分的可加性

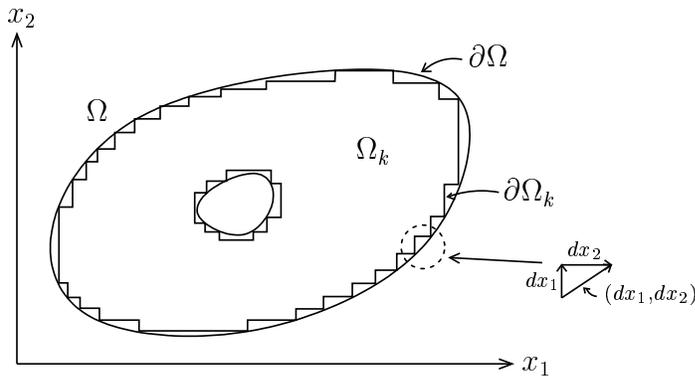
$$\int_{\partial\Omega_1} \omega + \int_{\partial\Omega_2} \omega = \int_{\partial\Omega} \omega$$

而面积分的可加性则是显然的。由此易证 (*)-式在 Ω 乃是由有限个坐标矩形接合而成情形也总是成立的。

如 [图 4-14] 所示, 我们可以用逐步扩大的 Ω 的子区域 $\{\Omega_k\}$ 去穷尽 Ω (参看 §4.1), 而它们本身都是由有限个坐标矩形接合而得者。因为 $\{\Omega \setminus \Omega_k\}$ 的面积在 $k \rightarrow \infty$ 时趋于 0, 则显然有

$$\lim_{k \rightarrow \infty} \int_{\Omega_k} d\omega = \int_{\Omega} d\omega$$

所以我们只需要再去验证 $\lim_{k \rightarrow \infty} \int_{\partial\Omega_k} \omega = \int_{\partial\Omega} \omega$ 。兹证之如下:



[图 4-14]

由所设 $\partial\Omega$ 是分段可微的, 所以它的微段可以用放大插图中的有向线段: (dx_1, dx_2) 为其一阶逼近, 而其相应的 $\partial\Omega_k$ 的微段则为 $(dx_1, 0)$ 和 $(0, dx_2)$ 。设 $\varepsilon > 0$ 是任给正数 (不论有多小), 由 $f_1(\mathbf{x})$ 和 $f_2(\mathbf{x})$ 的连续性可见, 只要把 k 取得足够大, $f_1(\mathbf{x})$ (及 $f_2(\mathbf{x})$) 在 $\partial\Omega$ 的微段和其相应的 $\partial\Omega_k$ 的微段之取值之差总是小于 ε 。由此即得

$$\left| \int_{\partial\Omega} \omega - \int_{\partial\Omega_k} \omega \right| < \ell(\partial\Omega)\varepsilon$$

亦即证得

$$\lim_{k \rightarrow \infty} \int_{\partial\Omega_k} \omega = \int_{\partial\Omega} \omega$$

这样也就证明了

$$\int_{\partial\Omega} \omega = \lim_{k \rightarrow \infty} \int_{\partial\Omega_k} \omega = \lim_{k \rightarrow \infty} \int_{\Omega_k} d\omega = \int_{\Omega} d\omega \quad \square$$

【Gauss 定理】：设 $\Omega \subset \mathbb{R}^3$ 是一个 3-维的有界闭区域，其边界 $\partial\Omega$ 是由逐片可微的曲面所组成者。设

$$\omega = f_1(\mathbf{x})dx_2 \wedge dx_3 + f_2(\mathbf{x})dx_3 \wedge dx_1 + f_3(\mathbf{x})dx_1 \wedge dx_2$$

是定义于一个包含 Ω 的开区域 D 之上的一阶连续可微二次外微分形式，则

$$(\#) \quad \int_{\partial\Omega} \omega = \int_{\Omega} d\omega, \quad d\omega = \left(\sum_{i=1}^3 \frac{\partial f_i}{\partial x_i} \right) dx_1 \wedge dx_2 \wedge dx_3$$

证明：直接验算，即有

$$\begin{aligned} & df_1 \wedge dx_2 \wedge dx_3 + df_2 \wedge dx_3 \wedge dx_1 + df_3 \wedge dx_1 \wedge dx_2 \\ &= \frac{\partial f_1}{\partial x_1} dx_1 \wedge dx_2 \wedge dx_3 + \frac{\partial f_2}{\partial x_2} dx_2 \wedge dx_3 \wedge dx_1 + \frac{\partial f_3}{\partial x_3} dx_3 \wedge dx_1 \wedge dx_2 \\ &= \left(\sum_{i=1}^3 \frac{\partial f_i}{\partial x_i} \right) dx_1 \wedge dx_2 \wedge dx_3 \end{aligned}$$

一如前述 Green's 定理的证明，让我们先来验证 Ω 是坐标矩形的特殊情况，即

$$\Omega_0 = \{\mathbf{x}; a_i \leq x_i \leq b_i, 1 \leq i \leq 3\}$$

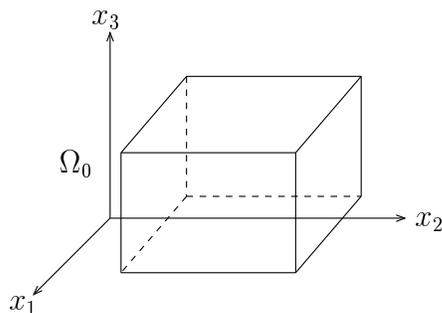
再者，令

$$\omega_1 = f_1(\mathbf{x})dx_2 \wedge dx_3, \quad \omega_2 = f_2(\mathbf{x})dx_3 \wedge dx_1, \quad \omega_3 = f_3(\mathbf{x})dx_1 \wedge dx_2$$

则上述验证显然可以归于

$$\int_{\Omega_0} \omega_i = \int_{\Omega_0} d\omega_i, \quad 1 \leq i \leq 3$$

的各别验证。



[图 4-15]

如 [图 4-15] 所示, 3-维坐标矩形 Ω_0 共有上、下; 右、左; 前、後这六个面。因为 ω_1 只含有 $dx_2 \wedge dx_3$ 项, 所以 $\int_{\partial\Omega_0} \omega_1$ 除了在前、後两面之外, 其他四面之积分皆显然为 0, 即有

$$\int_{\Omega_0} \omega_1 = \int_{D_1} [f_1(b_1, x_2, x_3) - f_1(a_1, x_2, x_3)] dx_2 \wedge dx_3$$

其中 $D_1 = \{(x_2, x_3); a_2 \leq x_2 \leq b_2, a_3 \leq x_3 \leq b_3\}$ 。再以微积分基本定理将 $[f_1(b_1, x_2, x_3) - f_1(a_1, x_2, x_3)]$ 代以

$$\int_a^{b_1} \frac{\partial f_1}{\partial x_1}(\mathbf{x}) dx_1$$

即得

$$\begin{aligned} \int_{\partial\Omega_0} \omega_1 &= \int_{D_1} \left\{ \int_{a_1}^{b_1} \frac{\partial f_1}{\partial x_1} dx_1 \right\} dx_2 \wedge dx_3 \\ &= \int_{\Omega_0} \frac{\partial f_1}{\partial x_1} dx_1 \wedge dx_2 \wedge dx_3 = \int_{\Omega_0} d\omega_1 \end{aligned}$$

同理亦可验证

$$\int_{\partial\Omega_0} \omega_2 = \int_{\Omega_0} d\omega_2 \quad \text{和} \quad \int_{\partial\Omega_0} \omega_3 = \int_{\Omega_0} d\omega_3$$

三者两侧各别相加即得

$$\int_{\partial\Omega_0} \omega = \int_{\Omega_0} d\omega$$

同样地, 当 Ω 可以分割成 Ω_1 和 Ω_2 相接而成者, 则 $\partial\Omega_1$ 和 $\partial\Omega_2$ 的共同面上的定向相反。由此可见面积分对于这种接合的可加性, 即

$$\int_{\partial\Omega} \omega = \int_{\partial\Omega_1} \omega + \int_{\partial\Omega_2} \omega$$

而体积分的接合可加性则是显然的。由此可见 (#)-式在 Ω 乃是有有限个 3-维坐标矩形接合而成者总是成立的。

最後, 我们也同样地可以用逐步扩大的 Ω 的子区域 $\{\Omega_k\}$ 去穷尽 Ω (参看 §4.1), 而它们本身则都是由有限个 3-维坐标矩形接合而得者, 即有

$$\int_{\partial\Omega_k} \omega = \int_{\Omega_k} d\omega$$

对于所有 Ω_k 皆成立, 而且

$$\lim_{k \rightarrow \infty} \int_{\Omega_k} d\omega = \int_{\Omega} d\omega$$

所以我们只需要再验证 $\lim_{k \rightarrow \infty} \int_{\partial\Omega_k} \omega = \int_{\partial\Omega} \omega$ 。兹证之如下:

为此, 我们不妨分别验证 $\lim_{k \rightarrow \infty} \int_{\partial\Omega_k} \omega_i = \int_{\partial\Omega} \omega_i, 1 \leq i \leq 3$, 而它们各别的验证几乎是完全一样的。所以在此仅详述 $i = 1$ 的情形:

由所设 $\partial\Omega$ 是分片可微的。再者, 对于任给 $\delta > 0$ (不论其有多小), 只要把 k 取得足够大, 则 $\partial\Omega_k$ 业已包含在和 Ω 相距小于 δ 的那个薄层之内。对于 $\partial\Omega$ (或 $\partial\Omega_k$) 的每个微片, 在求 ω_1 的面积分时, 只有其在 $dx_2 \wedge dx_3$ 方向的分量 (亦即垂直投影) 才有非零之值。再者, $\partial\Omega$ 上的微片可以分成两类, 第一类是在它的 2δ 邻近具有相应的 $\partial\Omega_k$ 上的薄片, 它和前者具有相等的 $dx_2 \wedge dx_3$ 方向的分量; 而其他部分则为第二类。只要把 k 取得足够大, 属于第二类的那种微片的垂直投影的绝对值之和就可以小到任意小。另一方面, 由 $f_1(\mathbf{x})$ 的均匀连续性易见 ω_1 在所有第一类微片上的积分和与其在 $\partial\Omega_k$ 上相应的积分和两者之差也是可以小到任意小的。这也就验证了

$$\lim_{k \rightarrow \infty} \int_{\partial\Omega_k} \omega_1 = \int_{\partial\Omega} \omega_1$$

同理亦可验证

$$\lim_{k \rightarrow \infty} \int_{\partial\Omega_k} \omega_i = \int_{\partial\Omega} \omega_i, \quad i = 2, 3$$

三者相加, 即得所需证之

$$\lim_{k \rightarrow \infty} \int_{\partial\Omega_k} \omega = \int_{\partial\Omega} \omega$$

亦即有

$$\int_{\partial\Omega} \omega = \lim_{k \rightarrow \infty} \int_{\partial\Omega_k} \omega = \lim_{k \rightarrow \infty} \int_{\Omega_k} d\omega = \int_{\Omega} d\omega \quad \square$$

Gauss 定理的高维推广:

设 $\Omega \subset \mathbb{R}^n$ 是一个有界的 n -维闭区域, 其边界 $\partial\Omega$ 是由逐片可微超曲面所组成者。设

$$\omega = \sum_{i=1}^n (-1)^{i+1} f_i(\mathbf{x}) dx_1 \wedge \dots \wedge dx_{i-1} \wedge dx_{i+1} \wedge \dots \wedge dx_n$$

基础分析学之二

是定义于一个包含 Ω 的开区域 D 之上的一阶连续可微的 $(n-1)$ -次微分形式，则

$$\int_{\partial\Omega} \omega = \int_{\Omega} d\omega, \quad d\omega = \left(\sum_{i=1}^n \frac{\partial f_i}{\partial x_i} \right) dx_1 \wedge \dots \wedge dx_n$$

证明：和 3-维的情形基本相同。读者试自证之。 \square

4.3.1 外微分和广义 Stoke's 定理

一个连续可微函数 $f(\mathbf{x})$ 也可以想成是外微分形式的特例，亦即 0-次者也。在这种最低次的情形其外微分也就是它的全微分，即

$$df := \sum_{i=1}^n \frac{\partial f}{\partial x_i} dx_i$$

设 γ 是 $f(\mathbf{x})$ 的定义域 D 中的一条可微曲线段，其始、终点分别是 \mathbf{x}_0 、 \mathbf{x}_1 。把微积分基本定理用到单元函数 $f(\gamma(t))$ ，即得

$$\int_{\gamma} df = f(\mathbf{x}_1) - f(\mathbf{x}_0)$$

亦即

$$\langle df, \gamma \rangle = \langle f, \partial\gamma \rangle$$

再者，前面所研讨的 Green's 定理和广义 Gauss 定理证明了

$$\langle d\omega, \Omega \rangle = \langle \omega, \partial\Omega \rangle$$

在 Ω 是 n -维， ω 是 $(n-1)$ -次的情形，其中外微分“ d ”的定义是

$$\begin{aligned} & d \left\{ \sum_{i=1}^n f_i dx_1 \wedge \dots \wedge dx_{i-1} \wedge dx_{i+1} \wedge \dots \wedge dx_n \right\} \\ &= \sum_{i=1}^n df_i \wedge dx_1 \wedge \dots \wedge dx_{i-1} \wedge dx_{i+1} \wedge \dots \wedge dx_n \end{aligned}$$

本小节所研讨的课题是如何对于其他次数（亦即 $0 < k < (n-1)$ ）的外微分形式定义其外微分运算，使得

$$\langle d\omega, \Sigma \rangle = \int_{\Sigma} d\omega = \int_{\partial\Sigma} \omega = \langle \omega, \partial\Sigma \rangle$$

普遍成立。[这也就是广义 Stoke's 定理。]

【分析】：

(一) 全微分满足下述两个基本、好用的运算律，即

$$(i) \quad (\text{线性}) \quad d(c_1 f_1 + c_2 f_2) = c_1 df_1 + c_2 df_2 \quad (c_i \text{ 是常数})$$

$$(ii) \quad d(f \cdot g) = df \cdot g + f \cdot dg$$

我们所要探索的外微分乃是全微分的推广，所以它自然也应该有类似的运算律，即

$$(i) \quad d(c_1 \omega_1 + c_2 \omega_2) = c_1 d\omega_1 + c_2 d\omega_2$$

$$(ii)' \quad d(\omega_1 \wedge \omega_2) = d\omega_1 \wedge \omega_2 + (-1)^k \omega_1 \wedge d\omega_2$$

其中 k 是 ω_1 的次数。

[注]：因为外积的斜对称性，即

$$\omega_1 \wedge \omega_2 = (-1)^{k\ell} \omega_2 \wedge \omega_1 \quad (\deg \omega_1 = k, \deg \omega_2 = \ell)$$

(ii)' 中的 $(-1)^k$ 是必须的。例如在 $k = \ell = 1$ 的情形若不加 $(-1)^k$ ，就会产生下述矛盾：

$$\begin{aligned} d(\omega_1 \wedge \omega_2) &= -d(\omega_2 \wedge \omega_1) \\ \Rightarrow d\omega_1 \wedge \omega_2 + \omega_1 \wedge d\omega_2 &= -(d\omega_2 \wedge \omega_1 + \omega_2 \wedge d\omega_1) \\ \Rightarrow 2d(\omega_1 \wedge \omega_2) &= 0 \end{aligned}$$

对于任给一次微分形式 ω_1, ω_2 皆成立，此事显然不对。但是采取 (ii)', 则有

$$\begin{aligned} d(\omega_1 \wedge \omega_2) &= d\omega_1 \wedge \omega_2 - \omega_1 \wedge d\omega_2 \\ -d(\omega_2 \wedge \omega_1) &= -\{d\omega_2 \wedge \omega_1 - \omega_2 \wedge d\omega_1\} \\ &= d\omega_1 \wedge \omega_2 - \omega_1 \wedge d\omega_2 \end{aligned}$$

而不产生矛盾。由此可见，在 (ii)' 中加上 $(-1)^k$ 这种更改乃是外积的斜对称性所使然者也。

(二) 由

$$\langle df, \gamma \rangle = \langle f, \partial\gamma \rangle$$

易见 $\int_{\gamma} df$ 对于任给闭曲线 γ 恒为 0。由此可见，假如要使得广义 Stoke's 定理对于 $k=2$ 的情形成立，则其中所用的外微分“ d ”必须使得 $d(df) \equiv 0$ ，因为

$$\langle d(df), \Sigma \rangle = \langle df, \partial\Sigma \rangle \equiv 0$$

(三) 设“ d ”是满足 (ii)' 和 $d^2g \equiv 0$ 者，则

$$\begin{aligned} & d(fdg_1 \wedge \dots \wedge dg_k) \\ &= df \wedge (dg_1 \wedge \dots \wedge dg_k) + f \cdot d(dg_1 \wedge \dots \wedge dg_k) \\ &= df \wedge dg_1 \wedge \dots \wedge dg_k + f \cdot 0 = df \wedge dg_1 \wedge \dots \wedge dg_k \end{aligned}$$

由此可见，下述定义之外微分运算“ d ”乃是唯一能够具有运算律 (i), (ii)' 和 $d^2f \equiv 0$ 者也。所以顺理成章，即有下述外微分之定义：

$$\begin{aligned} \text{【定义】} : d \left\{ \sum_{1 \leq i_1 < \dots < i_k \leq n} f_{i_1, \dots, i_k} dx_{i_1} \wedge \dots \wedge dx_{i_k} \right\} \\ = \sum_{1 \leq i_1 < \dots < i_k \leq n} df_{i_1, \dots, i_k} \wedge dx_{i_1} \wedge \dots \wedge dx_{i_k} \end{aligned}$$

[注]：在全微分的讨论中，我们业已指出，公式

$$df = \sum_{i=1}^n \frac{\partial f}{\partial x_i} dx_i$$

不但在 $\{x_i\}$ 是自变元时成立，而是不论 x_i 是自变元或应变元总是普遍成立的。此事至关重要，因为在多元、多关系的数理分析中，那些是自变元，那些是应变元乃是一种因地（或因时）制宜的一种选用吧了。所以在分析学中的重要公式，是很有必要把它们从自变元的选取的「依赖性」中解放出来的。在上述外微分的定义式中，骤看起来，好像是用到 $\{x_i\}$ 乃是自变元的。其实，在 $\{x_i\}$ 不是自变元时，上述公式依然成立！因为上述公式乃是性质 (i), (ii)' 和 $d^2x_i \equiv 0$ 的结论。由此可见，外微分运算 d 还有下述重要的性质：

设 Φ 是将 k -维 t -空间中的区域 Ω 映射到一个外微分形式 ω 的定义域 D 之中的一阶可微映射, 即

$$\begin{array}{ccc} \mathbb{R}^k \supset \Omega & \xrightarrow{\Phi} & D \subset \mathbb{R}^n \\ \cup & & \cup \\ & \mathbf{t} \mapsto & \Phi(\mathbf{t}) \\ & & \Phi(\mathbf{t}) = (\varphi_1(\mathbf{t}), \dots, \varphi_n(\mathbf{t})) \end{array}$$

令 $\Phi^*(\omega)$ 是将 ω 中的 x_i 代以 $\varphi_i(\mathbf{t})$ 之所得, 则有

$$d\Phi^*(\omega) = \Phi^*(d\omega)$$

把上述性质应用到 $\Phi: \Omega \rightarrow \Sigma \subset D$, 且满足规则性, 而 ω 则是 $(k-1)$ -次微分形式的情形, 即可证明下述广义 Stoke's 定理:

【广义 Stoke's 定理】:

$$\langle d\omega, \Sigma \rangle = \int_{\Sigma} d\omega = \int_{\partial\Sigma} \omega = \langle \omega, \partial\Sigma \rangle$$

证明: 设 $\Phi: \Omega \rightarrow \Sigma$ 是 k -维曲面的局部坐标化 (亦即参数表示)。则有

$$\int_{\Sigma} d\omega := \int_{\Omega} \Phi^*(d\omega), \quad \int_{\partial\Sigma} \omega := \int_{\partial\Omega} \Phi^*(\omega)$$

再者, 由 k -维 Gauss 定理即有

$$\int_{\partial\Omega} \Phi^*(\omega) = \int_{\Omega} d\Phi^*(\omega) = \int_{\Omega} \Phi^*(d\omega) = \int_{\Sigma} d\omega \quad \square$$

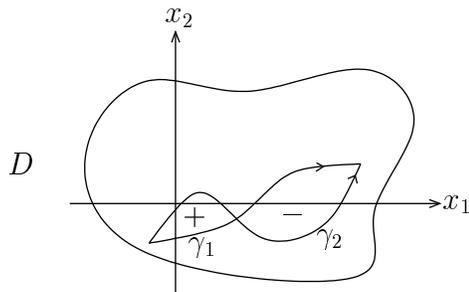
[注]: 上述广义 Stoke's 定理乃是微积分基本定理在多元、多关系的分析学中的自然推广, 其应用之范畴既深且广, 真可以说是妙用无穷者也。本书限于篇幅和尽可能维持所讨论的课题是既基础又初等之选择, 将仅仅以几个直接推论和简单的例题习题作一初步简介。

4.3.2 例子

(1) 设 ω 是一个定义于 \mathbb{R}^2 中一个单连通 (亦即不含任何空洞) 的闭区域 D 上的一阶连续可微的一次微分形式。若 $\frac{\partial f_1}{\partial x_2} \equiv \frac{\partial f_2}{\partial x_1}$ (亦即 $d\omega \equiv 0$)

则存在 D 上一个二阶可微的函数 U ，使得 $\omega = dU$ （亦即 ω 乃是一个全微分）。

证明：如 [图 4-16] 所示，设 γ_1 和 γ_2 是 D 中任给两条具有同样的始、终点的曲线。则 γ_1 和 γ_2 的逆向所组成者乃是 D 中一个定向区域 Ω 的有向边界。



[图 4-16]

由 Green's 定理和所设 $d\omega \equiv 0$ 即有

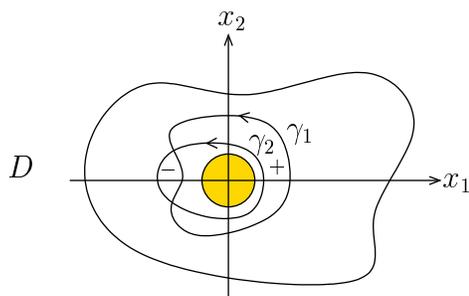
$$\int_{\gamma_1} \omega - \int_{\gamma_2} \omega = \int_{\partial\Omega} \omega = \int_{\Omega} d\omega \equiv 0$$

亦即 $\int_{\gamma} \omega$ 仅仅和 γ 的始、终点有关。所以对于选定起点 \mathbf{a} 的所有曲线 $\int_{\gamma} \omega$ 乃是其终点的函数，即可定义

$$U(\mathbf{x}) = \int_{\gamma} \omega \quad (\gamma \text{ 以 } \mathbf{a} \text{ 和 } \mathbf{x} \text{ 为始、终点})$$

不难验证如此构造的 $U(\mathbf{x})$ 的全微分就是 ω [其验证留作习题]。

(2) 设 ω 是一个定义于如 [图 4-17] 所示的开区域 D 上的一阶连续可微一次微分形式，而且 $d\omega \equiv 0$ 。



[图 4-17]

设 γ_1 和 γ_2 是 D 中两条同向地绕 D 的空洞 (可能仅仅是原点这一个点) 一周的简单闭曲线 (simple closed curve), 则 γ_1 和 γ_2 的逆向所组成者, 乃是 D 中一个定向区域 Ω 的定向边界 $\partial\Omega$ 。同样的, 由 Green's 定理和所设 $d\omega \equiv 0$, 即有

$$\int_{\gamma_1} \omega - \int_{\gamma_2} \omega = \int_{\partial\Omega} \omega = \int_{\Omega} d\omega \equiv 0$$

亦即这样的两条定向闭曲线的线积分恒相等。由此不难稍作推广, 得知 D 中一条定向闭曲线 γ 的线积分 $\int_{\gamma} \omega$ 仅仅和它对于 D 的空洞的定向绕数有关, 亦即有公式:

$$\int_{\gamma} \omega = 2\pi cn(\gamma)$$

其中 c 是一个随 ω 而定的常数。令

$$\omega_0 = \frac{-x_2 dx_1}{x_1^2 + x_2^2} + \frac{x_1 dx_2}{x_1^2 + x_2^2}$$

不难验证

$$\widehat{\omega} = \omega - c\omega_0$$

满足 $d\widehat{\omega} \equiv 0$ 而且对于任给 D 中闭曲线 γ 恒有

$$\int_{\gamma} \widehat{\omega} = \int_{\gamma} \omega - \int_{\gamma} c\omega_0 \equiv 0$$

由此易证 $\widehat{\omega}$ 乃是 D 上的一个二阶连续可微的函数 U 的全微分, 亦即:

$$\omega = c\omega_0 + dU, \quad (\mathbf{x} \in D)$$

【习题】: 试将上述讨论推广到 D 有多个空洞的情形。

(3) 在第四册第五章中, 我们由弧长第一变分公式得出一个抽象旋转面中, 曲线的曲率公式为

$$\kappa_g(s) = \frac{d\alpha}{ds} + f'(r) \frac{d\theta}{ds}$$

再者，设 Ω 是其中一个给定的区域，其边界 $\partial\Omega$ 分段连续可微，则有其 Gauss-Bonnet 公式，即

$$2\pi - \int_{\partial\Omega} \kappa_g(s) ds - \sum \alpha_i = \int_{\Omega} \left(-\frac{f''(r)}{f(r)} \right) dA$$

回顾其证明乃是：

$$\begin{aligned} & 2\pi - \int_{\partial\Omega} \kappa_g(s) ds - \sum \alpha_i \\ &= \left(2\pi - \int_{\partial\Omega} d\alpha - \sum \alpha_i \right) \stackrel{=0}{=} - \int_{\partial\Omega} f'(r) d\theta \\ &= - \int_{\partial\Omega} f'(r) d\theta = - \int_{\Omega} f''(r) dr \wedge d\theta \quad (\text{Green's 定理}) \\ &= \int_{\Omega} \left(-\frac{f''(r)}{f(r)} \right) dA, \quad dA = f(r) dr \wedge d\theta \end{aligned}$$

所以其证明的关键就是运用 Green's 定理。这个公式把边界的曲线曲率和区域中每点的曲面曲率的关系得以明确。

(4) 在 §4.2.4 的习题 (4)，我们要读者以图解说明 $\int_{\partial\Omega} -y dx$, $\int_{\partial\Omega} x dy$ 都等于 Ω 的定向面积。其实，它们乃是 Green's 定理的特例是也。这也就是说，即使像求面积这样简单的问题，Green's 定理依然是有其用武之地的（这当然是一种大才小用）。

[以上是对于 \mathbb{R}^2 中的线积分和 Green's 定理的一些实例的讨论。接著让我们讨论关于 \mathbb{R}^3 中的线积分、面积分的一些实例。因为 \mathbb{R}^3 乃是我们和万物共存于其间者，所以下面所要讨论的实例就更有其现实的意义和重要性。]

(5) 在 \mathbb{R}^3 中的一个连续可微一次形式 ω_1 有其相应的向量场 $\mathbf{F}(\mathbf{x})$ ，亦即

$$\omega_1 = \mathbf{F}_1(\mathbf{x}) \cdot d\mathbf{x}$$

而一个 \mathbb{R}^3 中的一个连续可微的二次形式 ω_2 又有其相对应的向量场，即

$$\begin{aligned} \omega_2 &= g_1(\mathbf{x}) dx_2 \wedge dx_3 + g_2(\mathbf{x}) dx_3 \wedge dx_1 + g_3(\mathbf{x}) dx_1 \wedge dx_2 \\ &\quad \updownarrow \\ \mathbf{F}_2(\mathbf{x}) &= (g_1(\mathbf{x}), g_2(\mathbf{x}), g_3(\mathbf{x})) \end{aligned}$$

假如我们从一个可微向量场

$$\mathbf{F}_1(\mathbf{x}) = (f_1(\mathbf{x}), f_2(\mathbf{x}), f_3(\mathbf{x}))$$

出发, 先求其相应的一次形式, 即

$$\omega_1 = \sum_{i=1}^3 f_i(\mathbf{x}) dx_i$$

然后再求其外微分, 即有

$$\begin{aligned} \omega_2 = d\omega_1 &= \sum_{i=1}^3 df_i \wedge dx_i \\ &= \left(\frac{\partial f_3}{\partial x_2} - \frac{\partial f_2}{\partial x_3} \right) dx_2 \wedge dx_3 + \left(\frac{\partial f_1}{\partial x_3} - \frac{\partial f_3}{\partial x_1} \right) dx_3 \wedge dx_1 \\ &\quad + \left(\frac{\partial f_2}{\partial x_1} - \frac{\partial f_1}{\partial x_2} \right) dx_1 \wedge dx_2 \end{aligned}$$

而它所相应的向量场 $\mathbf{F}_2(\mathbf{x})$ 就是

$$d\omega_1 \longleftrightarrow \mathbf{F}_2(\mathbf{x}) = \left(\left(\frac{\partial f_3}{\partial x_2} - \frac{\partial f_2}{\partial x_3} \right), \left(\frac{\partial f_1}{\partial x_3} - \frac{\partial f_3}{\partial x_1} \right), \left(\frac{\partial f_2}{\partial x_1} - \frac{\partial f_1}{\partial x_2} \right) \right)$$

这样就得到一种由一个向量场 $\mathbf{F}_1(\mathbf{x})$ 得出另一个向量场 $\mathbf{F}_2(\mathbf{x})$ 的运算, 它和外微分相对应。其基本算法就是

$$(f_1, f_2, f_3) \longmapsto \left(\frac{\partial f_3}{\partial x_2} - \frac{\partial f_2}{\partial x_3}, \frac{\partial f_1}{\partial x_3} - \frac{\partial f_3}{\partial x_1}, \frac{\partial f_2}{\partial x_1} - \frac{\partial f_1}{\partial x_2} \right)$$

假如我们把 ∇ 想成一个以 $\frac{\partial}{\partial x_i}$ 为其分量的算子向量, 亦即令

$$\nabla = \left(\frac{\partial}{\partial x_1}, \frac{\partial}{\partial x_2}, \frac{\partial}{\partial x_3} \right)$$

则上述运算就可以表达成

$$\mathbf{F} \longmapsto \nabla \times \mathbf{F}$$

而原先的梯度向量

$$\nabla f = \left(\frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}, \frac{\partial f}{\partial x_3} \right)$$

则又可以看做算子向量乘以函数 f [这就是在向量分析 (vector analysis) 中通用的符号]。

(6) 假如我们把一个可微向量场 $\mathbf{F}(\mathbf{x})$ 对应于其相应的二次形式，即

$$\mathbf{F}(\mathbf{x}) = (f_1(\mathbf{x}), f_2(\mathbf{x}), f_3(\mathbf{x}))$$

$$\updownarrow$$

$$\omega = f_1(\mathbf{x})dx_2 \wedge dx_3 + f_2(\mathbf{x})dx_3 \wedge dx_1 + f_3(\mathbf{x})dx_1 \wedge dx_2$$

则有
$$d\omega = \left(\sum_{i=1}^3 \frac{\partial f_i}{\partial x_i} \right) dx_1 \wedge dx_2 \wedge dx_3$$

所以相当于外微分的运算乃是由一个向量场 \mathbf{F} 去计算它的发散量 (divergence) $\sum_{i=1}^3 \frac{\partial f_i}{\partial x_i}$ ，而它又可以采用 ∇ 这个算子向量，写成

$$\nabla \cdot \mathbf{F}$$

在向量分析中通用之术语与符号是

函数 f 的梯度向量 $\text{grad} : \nabla f$

向量场 \mathbf{F} 的发散量 $\text{div} : \nabla \cdot \mathbf{F}$

向量场 \mathbf{F} 的涡旋量 $\text{curl} : \nabla \times \mathbf{F}$

(7) 在外微分中，由 $d^2f \equiv 0$ 和它的运算律易证 $d^2\omega \equiv 0$ [其验证留作习题]。由此可见，上述关于向量场的运算即有

$$\nabla \times \nabla f \equiv 0, \quad \nabla \cdot \nabla \times \mathbf{F} \equiv 0$$

(8) 从外微分形式来看， $d^2\omega \equiv 0$ 所表达者乃是说：一个外微分形式 $\hat{\omega}$ 是另外一个低一次的外微分形式 ω 的外微分时（亦即存在 ω 使得 $\hat{\omega} = d\omega$ ）则其本身的外微分必然恒等于零，亦即 $d\hat{\omega} = d(d\omega) \equiv 0$ 。例如当一个一次形式 $\hat{\omega}$ 是一个全微分，则必然有 $d\hat{\omega} \equiv 0$ 。换句话说 $d\hat{\omega} \equiv 0$ 乃是存在 ω 使得 $\hat{\omega} = d\omega$ 的必要条件。由前面对于 $n=2, k=1$ 的情形之讨论，可见 $d\hat{\omega} \equiv 0$ 是否业已构成其充要条件乃是取决于其定义域 D 的某种「拓扑性质」(topological property) [例如 D 是否有空洞]。再者，即使在 D 具有有限个空洞的情形， $d\hat{\omega} \equiv 0$ 虽然并不能保证它本身是一个全微分，但是它和全微分之间只相差某种非常简单的一次微分形式的线

性组合 (参看例 (2) 的讨论)。所以即使在一般情况, 虽然 $d\hat{\omega} \equiv 0$ 并非充分, 其实已经是: 「虽不中, 不远矣」。把这种虽不中不远矣的现象赋予有系统的研究, 就可反过来得出定义域的拓扑性质。这就是微分拓扑中的上同调理论 (cohomology theory)。

(9) 设 $D \subseteq \mathbb{R}^3$ 是一个开的实心凸体 (例如长方体或实心球) 它们都可以连续地收缩到其中任给一点 (contractible), 所以都应该看做是那种不含任何非平凡的拓扑结构者, 亦即是拓扑平凡者 (topologically trivial)。由此可以想到, 对于这样的区域上的外微分形式 $\hat{\omega}$, $d\hat{\omega} \equiv 0$ 应该业已构成它是一个低一次的 ω 的外微分的充要条件, 亦即

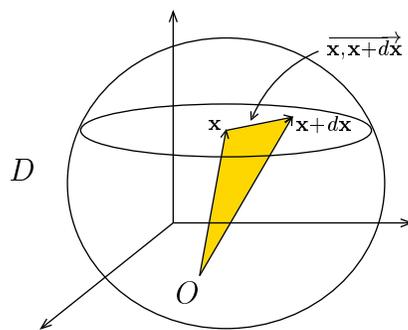
$$“d\hat{\omega} \equiv 0 \Leftrightarrow \exists \omega \text{ 使得 } \hat{\omega} = d\omega”$$

其实, 上述事实的证明乃是建立上面所提及的上同调理论的一个基本引理。[它乃是对于任何维数 n 都成立的。在此先行证明 $n=3$ 的情形。]

在 $n=3$ 的情形, 只有次数 $k=1$ 和 $k=2$ 是非平凡的, 兹分别讨论其验证如下:

$k=1$: 设 ω 是定义于 D 上的一阶连续可微一次微分形式而且 $d\omega \equiv 0$ (恒等于 0)。在 D 中任意取定其基点 O 。令 \overrightarrow{Ox} 为连结基点 O 和 D 中任给一点 x 的有向直线段。令

$$U(x) = \int_{\overrightarrow{Ox}} \omega$$



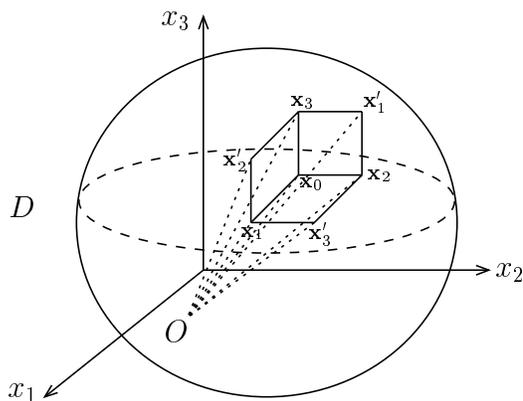
[图 4-18]

由所设 $d\omega \equiv 0$ 和 Stoke's 定理

$$\begin{aligned} U(\mathbf{x} + d\mathbf{x}) - U(\mathbf{x}) &= \int_{\overline{\mathbf{x}, \mathbf{x} + d\mathbf{x}}} \omega \equiv \mathbf{F}(\mathbf{x}) \cdot d\mathbf{x} \\ \Rightarrow dU &= \mathbf{F}(\mathbf{x}) \cdot d\mathbf{x} = \omega \end{aligned}$$

$k=2$: 设 $\hat{\omega}$ 是定义于 D 上的一阶连续可微二次微分形式而且 $d\hat{\omega} \equiv 0$ 。如 [图 4-19] 所示, O 是 D 中任意取定的基点, $\mathbf{x}_0 = (x_1, x_2, x_3)$ 是其中任给一点。令

$$\begin{aligned} \mathbf{x}_1 &= (x_1 + dx_1, x_2, x_3), & \mathbf{x}_2 &= (x_1, x_2 + dx_2, x_3), & \mathbf{x}_3 &= (x_1, x_2, x_3 + dx_3) \\ \mathbf{x}'_1 &= (x_1, x_2 + dx_2, x_3 + dx_3), & \mathbf{x}'_2 &= (x_1 + dx_1, x_2, x_3 + dx_3), \\ \mathbf{x}'_3 &= (x_1 + dx_1, x_2 + dx_2, x_3) \\ \hat{\omega} &= g_1(\mathbf{x})dx_2 \wedge dx_3 + g_2(\mathbf{x})dx_3 \wedge dx_1 + g_3(\mathbf{x})dx_1 \wedge dx_2 \end{aligned}$$



[图 4-19]

令

$$\begin{aligned} f_1(\mathbf{x}_0) &= \lim_{dx_1 \rightarrow 0} \frac{1}{dx_1} \int_{\Delta O \mathbf{x}_0 \mathbf{x}_1} \hat{\omega} \\ f_2(\mathbf{x}_0) &= \lim_{dx_2 \rightarrow 0} \frac{1}{dx_2} \int_{\Delta O \mathbf{x}_0 \mathbf{x}_2} \hat{\omega} \\ f_3(\mathbf{x}_0) &= \lim_{dx_3 \rightarrow 0} \frac{1}{dx_3} \int_{\Delta O \mathbf{x}_0 \mathbf{x}_3} \hat{\omega} \\ \omega &= \sum_{i=1}^3 f_i(\mathbf{x}_0) dx_i \end{aligned}$$

一方面，不难验证下述二阶逼近：

$$\begin{aligned}\int_{\partial \square_{\mathbf{x}_0 \mathbf{x}_1 \mathbf{x}'_3 \mathbf{x}_2}} \omega &\equiv \left(\frac{\partial f_2}{\partial x_1} - \frac{\partial f_1}{\partial x_2} \right) dx_1 \cdot dx_2 \\ \int_{\partial \square_{\mathbf{x}_0 \mathbf{x}_2 \mathbf{x}'_1 \mathbf{x}_3}} \omega &\equiv \left(\frac{\partial f_3}{\partial x_2} - \frac{\partial f_2}{\partial x_3} \right) dx_2 \cdot dx_3 \\ \int_{\partial \square_{\mathbf{x}_0 \mathbf{x}_3 \mathbf{x}'_2 \mathbf{x}_1}} \omega &\equiv \left(\frac{\partial f_1}{\partial x_3} - \frac{\partial f_3}{\partial x_1} \right) dx_3 \cdot dx_1\end{aligned}$$

[其验证留作习题，用微分均值定理。]

而另一方面，由 $f_i(\mathbf{x})$ 的定义， $d\hat{\omega} \equiv 0$ 和 Stoke's 定理即有

$$\begin{aligned}\int_{\partial \square_{\mathbf{x}_0 \mathbf{x}_1 \mathbf{x}'_3 \mathbf{x}_2}} \omega &= \int_{\Delta O \mathbf{x}_0 \mathbf{x}_1} \hat{\omega} + \int_{\Delta O \mathbf{x}_1 \mathbf{x}'_3} \hat{\omega} + \int_{\Delta O \mathbf{x}'_3 \mathbf{x}_2} \hat{\omega} + \int_{\Delta O \mathbf{x}_2 \mathbf{x}_0} \hat{\omega} \\ &\stackrel{\text{Stoke's}}{=} \int_{\square_{\mathbf{x}_0 \mathbf{x}_1 \mathbf{x}'_3 \mathbf{x}_2}} \hat{\omega} \equiv g_3(\mathbf{x}_0) dx_1 \cdot dx_2\end{aligned}$$

同理亦有

$$\begin{aligned}\int_{\partial \square_{\mathbf{x}_0 \mathbf{x}_2 \mathbf{x}'_1 \mathbf{x}_3}} \omega &= \int_{\square_{\mathbf{x}_0 \mathbf{x}_2 \mathbf{x}'_1 \mathbf{x}_3}} \hat{\omega} \equiv g_1(\mathbf{x}_0) dx_2 \cdot dx_3 \\ \int_{\partial \square_{\mathbf{x}_0 \mathbf{x}_3 \mathbf{x}'_2 \mathbf{x}_1}} \omega &= \int_{\square_{\mathbf{x}_0 \mathbf{x}_3 \mathbf{x}'_2 \mathbf{x}_1}} \hat{\omega} \equiv g_2(\mathbf{x}_0) dx_3 \cdot dx_1\end{aligned}$$

这也就验证了 $\hat{\omega} = d\omega$ 。

[注]：(i) 改用向量分析的运算来表达，则上述引理乃是：设 \mathbf{F} 是定义于一个开的实心凸体上的可微向量场，则

$$\begin{aligned}\nabla \times \mathbf{F} = 0 &\Leftrightarrow \exists f \text{ 使得 } \mathbf{F} = \nabla f \\ \nabla \cdot \mathbf{F} = 0 &\Leftrightarrow \exists \mathbf{V} \text{ 使得 } \mathbf{F} = \nabla \times \mathbf{V}\end{aligned}$$

这是在电磁学的场论中极为重要的引理。

(ii) 因为外微分运算是在坐标变换下保持不变的，亦即

$$d\Phi^*(\omega) = \Phi^*(d\omega)$$

由此易见上述引理对于任何和实心凸体可微同胚 (diffeomorphic) 的区域也当然成立。

(iii) 不难把上述构造法和论证推广到一般的 n 和 k 的情形。这也就是在上同调理论中的一个基本引理——Poincaré Lemma。

[接著让我们来分析一下定义域 D 并非实心凸体的情形，看一看上述引理究竟要作何种修改？有鉴于非实心凸体是非常多种多样的，例如一个含有气泡的玻璃球，一个被虫蛀了各种各样的洞的水果等等。基于以简御繁的方法论，我们应该先从一些既简单又典型的实例著手研讨问题的本质。]

(10) 设 D 是一个仅仅含有一个气泡的球体，例如

$$D = \{(x_1, x_2, x_3); r_0^2 < x_1^2 + x_2^2 + x_3^2 < R^2\}$$

在这样的情形，设 ω 是一个 $d\omega \equiv 0$ 的一次微分形式。则它在任给 D 中的闭曲线 γ 上的线积分依然恒等于零。

$$\oint_{\gamma} \omega \equiv 0$$

其理由是 D 中的任给闭曲线 γ 总是可以有一片曲面 Σ 使得 $\gamma = \partial\Sigma$ (证明留作习题)。所以

$$\int_{\gamma} \omega = \int_{\Sigma} d\omega \equiv 0$$

由此可见其线积分 $\int_{\gamma} \omega$ 仅仅和其端点有关，而且 ω 本身乃是某一个 D 上的二阶连续可微函数 U 的全微分，即 $\omega = dU$ 。所以在这种 D 上，引理对于 $k=1$ 的情形依然成立。

接著让我们来看一看 $k=2$ 的情形。设 ω 是 D 上一个其外微分恒等于 0 的二次微分形式，即 $d\omega \equiv 0$ 。设 Σ 是一个定向的闭曲面，若其内部并不包含 D 的那个气泡，亦即 $\Sigma = \partial\Omega$ 而 $\Omega \subset D$ ，则由 Gauss 定理，即有

$$\int_{\Sigma} \omega = \int_{\Omega} d\omega = 0$$

再者，若 Σ 的内部包含那个气泡则我们就无法用 Gauss 定理去得出 $\int_{\Sigma} \omega = 0$ 。例如 §4.2.4 习题 (9) 中所算得的

$$\begin{aligned}\omega_0 &= \sin \theta d\theta \wedge d\varphi \\ &= (x_1^2 + x_2^2 + x_3^2)^{-\frac{3}{2}} \{x_1 dx_2 \wedge dx_3 + x_2 dx_3 \wedge dx_1 + x_3 dx_1 \wedge dx_2\}\end{aligned}$$

则不难验证

$$\int_{\Sigma} \omega_0 = 4\pi$$

但是在这种情形，并非 Gauss 定理就变得「无用武之地」的了。其实它还是可以用来证明

$$\int_{\Sigma_1} \omega = \int_{\Sigma_2} \omega$$

对于两个包含那个气泡于其内部而且都以向外法向量为其定向的 Σ_1, Σ_2 恒成立（证明留作习题）。

再者，令上述共同值为 c （随 ω 而定之常数），则 $\hat{\omega} = \omega - \frac{c}{4\pi}\omega_0$ 就满足

$$\int_{\Sigma} \hat{\omega} = 0$$

对于任给 D 中的闭曲面恒成立。所以也可以用类似的构造法得出一个一次微分形式 ω_1 使得

$$\hat{\omega} = d\omega_1$$

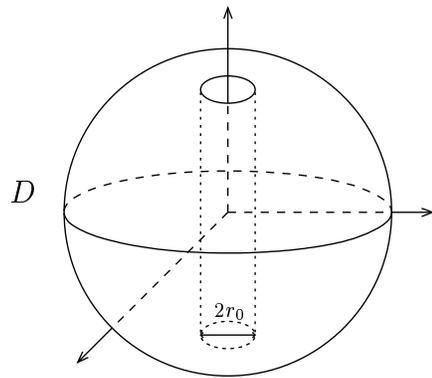
读者试自证之。

【思考题】：如 [图 4-20] 所示 D 是一个在 z -轴邻近被虫蛀了一个圆洞的球形水果，亦即 D 中之点的圆柱坐标满足下述条件，即

$$D = \{(r, \theta, z); r^2 + z^2 < R^2, r_0 < r < R\}$$

试证：(i) D 上任给满足 $d\hat{\omega} \equiv 0$ 的二次微分形式都是另一个一次微分形式的外微分，亦即存在 ω_1 使得 $\hat{\omega} = d\omega_1$ 。

(ii) 设 ω 是一个 D 上满足 $d\omega \equiv 0$ 的一次微分形式，则存在一个由 ω 而定的常数 c ，和适当的 D 上的可微函数 U ，使得 $\omega = dU + cd\theta$ 。



[图 4-20]

编後语

本书共分五册，分别对于基础数学中的：基础代数学；定性与定量平面几何和立体几何基础论；向量几何、解析几何与球面几何；单元微积分和多元微积分做了一遍反璞归真探本究源的研讨与建构。希望它能有助于青年学子们学好基础数学，并从中获得启发与奠基；也希望它能有助于正在教学育人、辛勤耕耘的老师，把基础数学教得更加精简实用、平实近人、引人入胜。这也就是我试著写这一组讲义的本意和心愿。

纵观古今中外，整个人类理性文明的发展史中，基础数学一直扮演著核心的角色。它不但是人类世代相承，对于大自然由表及里精益求精的认知、研究的基本工具，它也是整个自然科学在思想上、在方法论上的先行者和奠基者。由此可见，学习基础数学乃是一个有智而且有志的青年，作为一个人类理性文明的继承者的启蒙与奠基，也是把自己培训、拙壮成一个善于认识问题、善于解决问题、富有理性的人的不二法门。例如几何学是当之无愧的第一科学，它对于我们和宇宙万物共存于其间的空间，其精简美妙的本质精益求精的认知，乃是理性文明发展中贯串古今的重大篇章。在第二、三册的研讨中，读者应该让自己身历其境地去体会先贤们如中国古代善用面积的测量术；Hippasus 不可公度性的发现，Eudoxus 以逼近法对于定量几何基础论的重建与拓展，Archimedes 继承 Eudoxus 穷尽原理的妙用；圆锥截线的故事，解析几何的发明，非欧几何的发现，高维勾股定理与 Grassmann 代数等等；它们乃是理性文明史中，引人入胜、发人深思的诗篇。

代数学的基本思想就是善用运算律去探讨各种各样代数问题的系统解法。在运算律中，最具威力者首推分配律；而在各种各样代数问题之中，线性问题乃是最为简单者。在中国古算中的韩信点兵法（亦即孙子

算法，因为它可考的最早出处是东汉的孙子算经），业已把如何善用分配律去解决线性问题的基本思想，直截了当地展现无遗。从本书各册的研讨中，可以看到线性问题既简单又基本而且广泛有用，例如第三册中所讨论的向量几何：将空间的平移（亦即位移向量）妥加组织，引入自然的代数结构（亦即加法、倍积、内积和 \times -积）则空间的基本定理（如相似三角形定理、勾股定理、面积公式、高维勾股定理）都可以转换成分配律。由此可见，欧氏空间的几何结构可以全面代数化，而其所得的向量代数则是一种线性代数 (Linear Algebra)。再者，微分学的基本思想乃是有系统地运用局部线性逼近去分析各种各样变量问题；而微分和积分运算本身又都是线性运算。读者不但应该认识到线性代数的基本重要性和广泛可用性，而且也应该体会到解决各种各样的线性问题的基本方法，实乃中国古算中的韩信点兵法一以贯之者也。

概括地说，基础代数学的基础理论有二，其一是对于一元多项式的基础理论，例如插值公式，泰勒公式和辗转相除求最高公因式等；其二则是行列式和线性方程组的基础理论。而这些基础理论的发现与论证的基本方法都是归纳法，亦即由低次到高次，由少元到多元，逐步归纳地去探索、研讨和论证。总之，归纳法乃是代数学治学的基本方法，而善用分配律则是代数学的基本想法。

宇宙中各种各样事物和现象，动态乃是其常态，而静态则仅仅是其中极为少见的特例。由此可见，若要从这样无穷无尽，不断变化的万物万象中，由表及里探索其中所蕴含的规律性，当然就需要有一种善于分析变量问题的数学。数理分析 (Mathematical Analysis) 乃是我们由表及里，定量地深入探索大自然的本质的基本方法；而分析学 (Analysis) 也就是如此发展起来的一门数学。本书第四、第五册所研讨的单元、多元微积分则是分析学的基础理论。

当我们要对于某一种变量问题运用数理分析去探索其本质与规律时，分别以变元表达其中所含的各种变量，以函数关系去描述其各个变量之间的相互关联 (interlocking relationship)；而对于这样一个变量问题的数理分析也就是要对于描述它的那些函数关系 (functional relations) 做系统的定量分析。其中最最简单的情形就是那种只含两个变元，而且其一可以表达成其另一之函数（例如 $y = f(x)$ ）。对于这种最简的情形的基础理论就是单元微积分，它是进而研讨一般的多元、多关系的数

理体系的雏形和奠基。在第五册所研讨的多元微积分乃是以单元微积分所得者为基础，进而推广拓展而成者，它才是能够对于一般多元、多关系的变量问题作系统数理分析的基础理论。其中全微分和隐函数定理，外微分和 Stoke's 定理尤为基本重要。

在数理分析中，通常把所要研讨的数理体系所含的诸多变元看做一个点的坐标，则它们总体的变域就是坐标空间中的某种子集。隐函数定理的几何意义是：由一组可微的函数关系所描述的子集，在其描述函数的全微分是线性无关之点（称之为规则点）的邻近乃是一片 $(n-m)$ -维曲面，而且可以选取适当的 $(n-m)$ 个变元作为其局部坐标系。再者，定义在这种子集上（或某种包含它的坐标空间的子集之上）的函数、外微分形式或向量场则是各种各样植根于空间的解析事物 (analytic objects)。在本册第四章所讨论的积分则是空间中的几何事物和上述解析事物之间自然的相互作用所得之值，其关系有如第一册的附录中所讨论的向量空间和其偶空间之间的相互作用。由此可见，多元积分所描述者，乃是空间中的几何事物（即 k -维子集）和外微分形式这种解析事物之间的自然对偶关系，而微积分基本定理则可以推广为广义 Stoke's 定理！它证明了解析的外微分和几何的「求边界」两者之间的对偶性。这的确是一种应用广阔，意义深远的推广，读者宜多读多想它几遍，然后再在它的各种各样的应用之中（例如电磁学、微分几何学、微分拓扑学等等）逐步、逐样地去体会其用场和意义。

项武义

二〇〇三年 秋
于香港科技大学