2010.5.
第05期

目录

谈数学职业

张恭庆*

编者按： 本文[1]是张恭庆院士在2009年4 月中国数学会厦门学术年会上所做的公众报告，并在这次年会上荣获中国数学会“华罗庚数学奖”。张恭庆院士长期工作在数学研究和数学教学的第一线，为国家培养了很多优秀的数学人才。张恭庆院士在报告中谈到了数学，数学的特征，数学职业和青年数学家的历史使命，也谈到了数学教育的重要性。这是张恭庆院士一生从事数学研究和教学工作的认识和感悟，愿与所有的数学爱好者共勉。

1 前言

上个世纪50年代，数学通报刊登了苏联数学家柯尔莫果洛夫的“论数学职业”[1]的译文。我上大学时，这是我们“专业教育”的材料。对于我们这代学数学的人产生了很大的影响。然而半个多世纪过去了，数学的面貌发生了很大的变化，数学的职业也多样化了。今年的华尔街杂志上，发表了一篇文章[2]其中附有一张由CareerCast.com制作的，以工作环境、收入、就业前景、体力要求、体力强度为指标的职业排行榜。在这排行榜中，数学家荣登榜首，保险精算师和统计学家分列第二和第三，后面是生物学家，软件工程师，计算机系统的分析员等等。从这5个指标来看，数学家的收入不算很高，但综合起来还排在第一，可见在其它方面占有优势。

2 数学和它的基本特征

2.1 什么是数学

从中学起，我们就知道数学是研究“空间形式”和“数量关系”的学科。数量

*北京大学数学科学学院教授。1991年当选为中国科学院院士。1994年当选为第三世界科学院院士，并在同年举行的世界数学家大会上作45分钟应邀报告。曾任中国数学会理事长。

[1]此次刊登系转载自《数学通报》2009年第48卷第7期（已获得张恭庆院士和《数学通报》授权）。另感谢周坚老师向《荷思》推荐本文。

关系，简称为“数”，空间形式简称为“形”；“数”的对象比如说：自然数，复数，向量，矩阵，函数，概率等，“形”的对象比如说曲线，图，空间，流形等。

数学实际上是一门形式科学，它研究的是抽象元素之间的“关系”和“运算法则”。这些“相互关系”和“运算法则”构成了数学“结构”。判断数学结论的真伪，主要看其逻辑演绎是否正确，被实践检验的只是构成这些“相互关系”与“运算法则”的“结构”是否与实际相符。

我们举一个例子来说明。大家都知道平面几何中的“平行公设”：在平面上过直线外一点，有且仅有一条直线平行于该直线。这是公设，是假定，可以由此推出平面几何的很多定理。但为什么在平面上过直线外一点有且只有一条直线平行于该直线呢？可不可以没有？可不可以不止一条？就几何学来说，假定只有一条，可以推出一大堆几何命题，例如：三角形三内角之和为180°，这是欧氏几何；假定有不止一条也可以推出另外一大堆命题，例如：三角形三内角之和小于180°，这是双曲几何；假定一条也没有照样还可以推出一大堆命题，这是椭圆几何。双曲几何与椭圆几何都是非欧几何。那么到底哪一种几何的结论是正确的？这要看你把这些几何结论应用在什么范围内，应用到什么问题上去。在以地球为尺度的空间范围内，欧氏几何是适用的，实际上它与非欧几何中的双曲几何的差异不大。当把宇宙作为一个整体来描述时，就要用双曲几何了。这有点像牛顿力学与相对论力学的关系。由此可见，决不能认为凡是数学上证明了的定理就是真理。只能认为这些结论是在它的“结构”中在逻辑上被正确地证明了。至于其是否与实际相符，还要检查它的前提。从这个意义上说，数学只是一个形式体系。

如果把数学的研究对象只用“数”和“形”来概括，那么有些东西还无法概括进去。比如数学语言学，各种计算机的语言都是根据数学原理制做出来的，可语言是“形”？还是“数”呢？看来都不是。又比如在“选举”办法上：

[例] Arrow不可能性定理——(不存在“公平的”选举制度)有一个非常有名的结论——Arrow不可能性定理[3]。Arrow是个经济学家，诺贝尔奖获得者，学数学出身。他证明了一条定理：对于不少于3个候选人的选举按“排序”投票，不存在任何同时遵循以下四个原则的群体决策：

(1) 无限制原则。(任何人可对所有候选人任意排序)。

(2) 一致性原则。(如果每个人的态度都是“A优于B”，那么群体决策结果也应“A优于B”)。

(3) 独立性原则。(添加或减少候选人，“排序”不变)。

(4) 非独裁原则。(不能一个人说了算)。

这也是一条经济学上的定理：没有同时遵从以上四条原则的“社会福利函数”。这是数学在其他领域——政治学、经济学——的一个重要的应用。在这条定理中，哪里有“数”？哪里有“形”？可见“数”和“形”已经不能完全概括数学的研究对象了。现代人们不再限定研究对象是不是“数”和“形”，只要能对其建立数学模型，就能通过模拟计算来研究其中的规律，例如对于社会心理、动物行为等方面的数学分析。所以，数学研究的范围扩大了，现在入们说数学的对象是：“模式”(pattern)、“秩序”(order)、“结构”(structure)。

2.2 纯数学与应用数学

数学又划分为纯数学和应用数学，纯数学在我国又称基础数学。研究数学自身提出的问题，划归纯数学；研究数学之外(特别是现实世界中) 提出来的问题划归应用数学。应该说，这种划分只是大致的，并没有严格的界限。一方面，纯数学中的许多对象，追根溯源还是来自解决其它方面的问题，如天文学、力学、物理学等。比如几何来源于测量：天文测量、大地测量。就在数学已经高度发展了的今天，从外部提出来的数学问题照样可以转化为非常有意义的纯数学的问题，刺激出深刻的数学理论。比如说，Navier-Stokes方程是流体力学中的重要的方程，NP问题是从计算理论中提出来的问题，到现在都还没有被解决，成为“千年七大难题”中的两个。另一方面，纯数学的理论在适当条件下也能在其它科学中放出异彩：群论和几何对物理的贡献是众所周知的。大家都认为数论是很“纯”的数学，但数论在现代密码学中起重要作用，此外如傅里叶分析与通讯，随机过程与金融，几何分析与图像处理等等都是这方面的例子。特别是，许多在应用数学中行之有效的方法都有深刻的纯数学背景，如快速Fourier变换，有限元方法等。

纯数学大致有：数理逻辑、数论、代数、拓扑、几何、分析、组合与图论等分支，它们之间的融合与渗透又产生出许许多多的交叉分支，如代数几何、代数数论、微分几何、代数拓扑、表示理论、动力系统、泛函分析……等以及更多的子分支。

微分方程与概率论是介于纯数学与应用数学之间的分支，它们的理论部分属于纯数学，其余部分则是应用数学。

计算数学与数理统计是应用数学最重要的两个分支。

纯数学对于问题的解答往往只停留在研究解的存在性以及个数上，未必讨论解的具体算法(如代数方程求根)。但实际问题的解答一般总要求具体的数据，单有纯数学的结论不能满足要求，因此还要研究算法，以及如何对待巨大的计算量、存储量、复杂性、精确性、速度、稳定性等等问题。这些就是计算数学要解决的问题。

以概率论为基础的统计学称为数理统计。在日常生活、社会调查、科学实验都积累了大量的数据，如何从这些数据中科学地得到有用的信息？数理统计研究如何通过有效的收集、整理和分析带有随机性的数据，对所考察的问题做出推断、预测乃至决策的方法。

当代的数学已被应用到很多领域。自然科学：物理、化学、生物、天文、地质、气象等，人文学科：经济、金融、精算、语言、考古等。很多管理科学问题也要用到数学。

这么多有用处的数学，表面上看都属于应用数学。然而，正如Borel说的："纯数学和应用数学就像是一座冰山——水面上的是应用数学，因为它有用，大家都看得见；水底下的是纯数学"[4]。没有水底下纯数学的深厚积累，上面的应用数学建立不起来。

2.3 数学的基本特征

因为数学研究的是抽象的对象，所以应用范围必然广泛；又因为它的研究手段不是实验，而是逻辑推理。这就决定了它必须是严密的和精确的。因此数学明显地有如下基本特征：

(1) 高度的抽象性与严密的逻辑性；

(2) 应用的广泛性与描述的精确性；

数学应用的广泛性不仅表现在：它是各门科学和技术的语言和工具，数学的概念、公式和理论早已渗透到其它学科的教科书和研究文献中去了；而且还表现在：许许多多数学方法都已被写成软件，有的还被制成芯片装置在几亿台电脑、以及各种电器设备之中，成为产品高科技含量的核心，还有些数学软件则是作为商品在出售。在这些应用中，我想特别指出：数学是探求知识的重要手段。举一个例子，历史上许多重要的发现，没有数学光靠实验是不够的。现在大家人人用手机，不论多远几秒钟就能通上话，为什么信息能传输得这么快？靠的是电磁波。电磁波是怎样发现的？

[例] 以$\mathbf{E} = (E_1, E_2, E_3)$；$\mathbf{B} = (B_1, B_2, B_3)$表示电场强度和磁感应强度。$\rho$, $\mathbf{j}$, ϵ, μ表示：电荷，传导电流密度，介电常数和磁导率。Maxwell方程组写

作

$$\begin{cases} -\dfrac{\partial \mathbf{B}}{\partial t} = \nabla \times \mathbf{E} & \text{(Faraday定律)} \\ \nabla \cdot \mathbf{B} = 0 & \text{(磁场的Gauss定律)} \\ \dfrac{1}{\mu}\nabla \times \mathbf{B} = \epsilon \dfrac{\partial \mathbf{E}}{\partial t} + \mathbf{j} & \text{(Ampere定律)} \\ \nabla \cdot \mathbf{E} = \dfrac{1}{\epsilon}\rho & \text{(电荷的Gauss定律)} \end{cases}$$

Maxwell在研究电磁现象时注意到原来的Ampere定律：

$$\nabla \times \mathbf{B} = \mu \mathbf{j},$$

中没有$\dfrac{\partial \mathbf{E}}{\partial t}$这一项，有待修正。又因为等式

$$\nabla \cdot (\nabla \times \mathbf{B}) = 0.$$

将导出：

$$\nabla \cdot \mathbf{j} = 0.$$

再由电荷守恒：

$$\nabla \cdot \mathbf{j} + \frac{\partial \rho}{\partial t} = 0,$$

便有$\dfrac{\partial \rho}{\partial t} = 0$。然后由Gauss定律：

$$\nabla \cdot \mathbf{E} = \frac{1}{\epsilon}\rho.$$

一般来说$\dfrac{\partial \rho}{\partial t} \neq 0$，所以他把$\dfrac{\partial \mathbf{E}}{\partial t} = 0$称为“位移电流”，并把它添加到Ampere定律中将其修改为：

$$\frac{1}{\mu}\nabla \times \mathbf{B} = \epsilon \frac{\partial \mathbf{E}}{\partial t} + \mathbf{j}.$$

Maxwell以液体的流动、热的传导和弹性作为模型，认为这是传导电磁波的媒介“以太”[5]。尽管这种解释在物理上是不对的，也讲不清楚，但它的数学形式——Maxwell方程组却是正确的，它奠定了电磁学的基础。这个修正后的方程组是一个波动方程，由此预见了电磁波，后来Hertz在实验上证实了电磁波的存在。

同样地，在量子力学，相对论的理论建立过程中，数学也起了极为重要的作用。

在当今时代，科学计算更是在一定程度上取代了实验。一旦研究对象的机理已经清楚，准确的数学模型已经建立，就可以用模拟计算替代部分试验，如核试验等。

数学的基本特征除以上两条外，还有

(3) 数学研究对象的多样性和内部的统一性。

随着数学研究对象的扩充，数学分支不断增加，方向繁多，内容丰富。同时数学分支之间的内在联系也不断被发现，数学内部的千丝万缕的联系被愈理愈清。Hilbert-Noether-Bourbaki利用数学分支间的这些被清理过的联系和公理化方法，从规定的几条“公理”及其相关的一套演算规则中提炼出数学“结构”，如代数结构、拓扑结构、序结构等。数学的不同分支是由这些不同的“结构”组成的，而这些结构之间的错综复杂、盘根错节的联系又把所有的分支联成一个整体。在这方面反映了数学的统一性。

对统一性追求的意义在于：对于同一个对象可以从不同角度去认识，不同分支的问题可以相互转化，理论和方法可以相互渗透，从而发展出许多新的强有力的工具，解决许多单个分支方法难于解决的重大问题。

回顾以下历史是颇有裨益的。历史上有三大几何难题：倍立方问题，化圆为方问题，三等分角问题，都是“圆规直尺”的作图题。两千多年了，光用几何方法研究，不知有多少人费了多少心血，可就是解决不了！在那些时代，代数学主要研究解方程。后来笛卡尔用解析几何统一了几何与代数。18世纪末到19世纪初，多项式方程可解性的研究继Gauss代数基本定理证明之后应运而生。Gauss研究正多边形的圆规直尺作图就换了一个角度，把它看成一个多项式方程的可解性问题，从几何问题转化到了代数问题。后来Abel，Galois在代数上把方程的可解性研究推向了高峰。

什么样的“数”能被圆规直尺作出来？对于事先给定了的一组实数$\mathbb{Q}$，能从它们“尺规作图”出来的数x，就是通过它们出发，作$\pm, \times, \div$以及开平方$\sqrt{\ }$所能得到的数。也就是说：

尺规作图问题可解 $\Leftrightarrow \exists m, x \in \mathbf{F}_m$, 其中$\mathbf{F}_{i+1} = \{a_i + b_i\sqrt{c_i} \mid a_i, b_i, c_i \in \mathbf{F}_i\}$, $\mathbf{F}_1 = \mathbb{Q}$.

三等分角问题是：对任意给定的角θ，作出一个大小为其三分之一的角等$\dfrac{\theta}{3}$。令$\alpha = \cos\theta$，要作出数$x = \cos\dfrac{\theta}{3}$。$x$是多项式：

$$4x^3 - 3x - \alpha = 0$$

的根。但$x \notin \mathbf{F}_m$, $\forall m$。三等分一个任意角只用圆规与直尺是不可能的！

1837年法国数学家Wantzel证明了以上方程连同倍立方问题对应的方程在$\{\mathbf{F}_m \mid m = 1, 2, 3, \ldots\}$都是不可解的。后来，1882年Lindemann证明了π的超越性从而确立了尺规作图化圆为方也是不可能的。两千年前的三大几何难题就是这样用代数和数论的方法以否定的形式解决了！

现在数学已经发展成为一个庞大的、内部和谐与统一的、充满活力的有机的整体，它是人类文化宝库中一座既宏伟又精致的创造物。

3 当代社会对数学家的要求

当今世界，数学已被应用到几乎一切领域。然而现实情况一方面是，许许多多新的领域要求人们用数学的眼光，数学的理论和方法去探讨；另一方面，科学的发展使人们的分工愈来愈细。18世纪以前的数学家中有不少人同时也是天文学家、力学家、物理学家；在19世纪，许多数学家还可以在数学的几个不同分支上工作；但自Poincaré和Hilbert之后，已经没有一个人能够像他们当年那样通晓数学全貌了。大多数的数学家只能在狭窄的领域内从事研究。这种过于专门化的趋势对于数学学科的发展是十分有害的！这确实是一个矛盾的现象。如果我们不能对当今数学的发展与趋势有一个大致的了解，我们就不知道如何应对，也不知道应该怎样培养学生。

3.1 当代数学发展的趋势

大致有如下几个特点：

3.1.1 数学内部的联系进一步加强

在指数增涨的研究文献中，尽管数学的各个分支的前沿都在不断推进，数学在深度与广度两方面都得到快速发展，然而不同分支之间的融合与相互渗透则是一个重要的特征。这表现为：

原来长期处于纯数学边缘的分支，比如偏微分方程和概率论，现在已经进入了纯数学的核心。

相隔很远的分支间的内在联系不断发现。如de Rham-Hodge定理，Atiyah-Singer指标定理等。

许多困难的问题都需要很多学科的知识综合起来才能解决。例如，庞加莱猜想的提法本来纯粹是拓扑学的，后来转化为几何问题，光从几何也解决不了，最后是综合使用偏微分方程、拓扑、几何的思想、理论和方法才把这一个复杂问题解决了的。

3.1.2 数学与其它科学的交叉形成了许多交叉学科群

比如，科学计算就是数学与物理，化学，材料，生物，力学等很多学科的交叉。

数学与控制论，运筹学交叉形成了系统科学。数学与物理交叉，形成数学物理。此外还有计算机科学，信息科学，数量经济学，金融数学，生物数学，数学化学，数学地质学，数学心理学，数学语言学等等很多的交叉学科。

3.1.3 数学应用的领域空前扩张，成为开发高新技术的主要工具

信息安全、信息传输、图像处理、语音识别、医疗诊断、网络、海量数据处理、网页搜索、商业广告、反恐侦破、遥测遥感，包括当代制造业，成衣业等等都大量应用数学。

3.2 数学家的职业

长期以来人们心目中的数学职业只是限于学术界和教育界：大学、中学教师和科研机构的研究人员。如今逐渐有所改变，有些公司也开始雇佣学数学的人了。在一些发达国家，过去工业界(计算机)和商业界(统计、保险)雇佣一些数学硕士、学士从事计算、统计、程序编制和数据处理工作。随着工业中有兴趣的应用数学问题愈来愈多，近年来吸引了一定比例的数学博士和优秀的数学家。象弗里德曼(Freedman，Michael)，1986年菲尔兹奖获得者，现就职于微软公司。在许多发达国家现在都有专门的机构支持工业应用数学的发展，这标志着数学在这些国家的应用已相当广泛。我查了美国最近几期的《Notices of the American Math. Soc.》，从03年到08年，美国大概每年有800多名数学博士毕业后在美国求职。这800多个人中大约有200人，约占四分之一，到工商业界去；其他的人都就职于各种类型的学校，有研究型的，也有教学型的。不过，从读数学研究生到拿到数学博士学位其人数比大约是四比一，除去其中有些人转到别的学科攻读博士学位外，其余大多数或是直接，或是再读一个其它学位，如统计、精算等之后，都到工商业界和政府部门去工作了，这个数字可是惊人的。

学术界和工商界对数学的要求很不一样。在学术界，要求发表论文，证明定理，推进数学的进展；工商界的要求则是解决问题，尽快给出结果。对学术界来说，研究结果深刻、精确、有新思想的是好工作；工商界则要求有针对性和可用性，如果得到的公式虽然很广泛、很精确，但计算起来太费时费钱，就不一定会被采用。对于学术界的人，做研究可以自由选题，不受限制；但是在工商界，数学家的工作是被指定的，开发的项目也是被指定的。大家在选择自己的出路时要注意这些差别。

3.3 对数学家的要求

主修数学的人在学习过程中提高了抽象能力和逻辑推理能力，思考问题比较严

密，学习那些属于符号分析方面的新知识比较容易入门，这是学数学的人的优势。他们当然也有劣势，比如不擅长做实验等。

(1)到工商界工作的数学家主要从事符号分析，数据处理，建模、编程等方面的工作。然而数学的宝库是非常丰富的，如何采用更有效的理论和方法来解决问题，则要求更多地懂得该工作领域以及数学两方面的知识。要想工作有成绩，就不能只掌握几套现成的方法，而是要加宽数学的知识面。

(2)在交叉学科从事应用数学研究的数学家，更要深入到这个新领域中去，了解研究问题的来龙去脉。这些数学家并不以证明定理为成果的主要表现方式，而是创建好的模型。但创建好的模型正如证明深刻定理一样有意义，它是利用数学工具寻找客观规律的重要手段。实际问题很复杂，如何抓出主要因素，使之既能反映出主要现象，又能在数学上有办法处理。这种抽象、简化以及解决问题的方法是一种数学艺术。

然而在有些数学家中流行一种看法，认为应用数学是搞不了纯数学的人才去搞的。这是极为错误的，也是有害的观点！20世纪不仅有许多极有才华的应用数学家开创了许多应用数学分支，把数学的疆界空前地扩大了，如Turing，Shannon；而且还有些在纯数学领域中有卓越成就的数学家后来都又在应用数学领域做出了极富开创性的贡献，如Von Neumann，Wiener，Thom，Smale，以及2007年获得邵逸夫奖的Mumford和吴文俊等。

(3)做基础理论研究的数学家做纯数学的研究是非功利的. 在这个意义上有点像文学和艺术，也没有统一的评价标准。研究的成果贵在创新。然而这种创新并不是数学家们没有目标的随心所欲的创造。正如柯朗(R. Courant)说过，“只有在以达到有机整体为目标的前提，只有在内在需要的引导下，自由的思维才能做出有科学价值的成果”[6]。整个数学是一个有机整体，学科之间是相互牵连在一起，互相补充，互相促进的。一项工作如果很孤立，和主流上的问题都没有联系，也没有多少新的思想，那么就很难说意义有多大了。

数学分支间的融合与渗透是当代数学发展趋势的一个特征，要想在有意义的问题上做出贡献，知识面一定不能太窄。然而当代大多数数学家工作面过于专门化却是一种普遍现象。这有其内在的原因：数学的体系太庞大，内容又极为丰富，要想在前人工作的基础上有所拓广就很难有精力去了解其它分支；同时也有其外在原因，数量剧增的研究人员产生了大量的研究论文，发表的论文多就逼迫研究人员多读，而且“发表论文的压力”又逼迫他们多写，如此互为因果，也就无暇他顾了。这是当今国际学术界普遍存在的严重问题。然而研究贵在创新！真正的“原创”思想往往来自那些能“精通”看来相距遥远的几个领域，而且能“洞察”到把一个领域的结果用于解决另一个领域问题的途径。那是建立在全面了解、长期思考、过人

功力基础之上的。

(4)数学老师。有人说，我不想做研究，只想当老师教书。不错，本来教书就是学数学的一个重要出路。作为大学、甚至中学的数学老师，对他们所教的学科也不能只掌握教科书上所写的那一点内容，如果那样的话，或者会把书教得枯燥无味，或者不得要领。反之，如果教师的知识渊博，再肯学习新东西教给学生，学生对学习一定会产生很大的兴趣，事实上，只有那些热爱数学，并能把数学看成活生生的、不断发展着的人才能激励起学生的好奇心和求知欲。很多数学家回忆自己走过的道路时，怀念当年的数学老师，正是这些老师把他们引进了数学的殿堂。我们现在正处于数学理论和应用空前大发展的时代，怎么改革数学教育？怎样的师资才能适应大发展的需求？这些都是需要我们认真思考的问题。

4 数学教育的重要性

作为一种“思想的体操”，数学一直是中、小学义务教育的重要组成部分。现在大学理、工、文、法、农、医等科都有数学课，说明了人们认识到数学的重要性。不过，在许多学校，这些数学课的收效并不理想。原因可能是多种多样的，要具体分析。比如某系课程表上规定要上数学课，任课老师未必知道为什么这个系的学生需要开这门课。是作为“语言”的需要?专业课的需要？看书看文献的需要？还是做研究的需要？这是不同层次的要求。不按需要教，就是无的放矢，学生自然没有兴趣，效果也不会好。所以我建议教非数学专业学生的教师首先要了解一下这个专业的需求。

4.1 改善数学教育

几千年数学发展的丰富积累是人类的知识宝库。在知识社会，这个知识宝库是一种重要的资源。怎样能让这些资源共享，就要靠老师们传承给各行各业的人。

如何改善我国现行的数学教育，我认为要综合考虑以下几方面：

(1)知识。既重视基础，也照顾前沿，特别要考虑受教育对象的需要和基础。

(2)能力。“数学是一种普遍适用的，并赋予人以能力的技术”。计算能力(包括使用计算机进行计算的能力)、几何直观能力、逻辑推理能力、抽象能力、把实际问题转化为数学问题的能力。不能只灌输知识，更重要的是培养能力。在教学过程中，通过哪些内容培养哪些能力，或者培养哪几方面的能力，教师要做到心里有数。

(3)修养。数学是一种文化，数学不是一门自然科学，它有文化的层面。受过良好数学教育的人看问题的角度和一般的人不完全一样，它能开阔人的视野，增添人的智慧。一个人是否受过这种文化熏陶，在观察世界、思考问题时会有很大差别。

会不会欣赏数学，怎样欣赏数学，与数学修养有关，就如同欣赏音乐一样，不是人人都能欣赏贝多芬的交响乐的。

然而数学修养不但对数学工作者很重要，对于一般科学工作者也重要。就是有了数学修养的经营者、决策者在面临市场有多种可能的结果，技术路线有多种不同选择的时候，会借助数学的思想和方法，甚至通过计算来做判断，以避免或减少失误。詹姆斯・西蒙斯就是一个最好的例证。在进入华尔街之前，西蒙斯是个优秀的数学家。他和巴菲特的“价值投资”不同，西蒙斯依靠数学模型和电脑管理自己旗下的巨额基金，用数学模型捕捉市场机会，由电脑做出交易决策。他称自己为“模型先生”，认为建立好的模型可以有效地降低风险。在西蒙斯的公司里雇用了大量的数学、统计学和自然科学的博士。

发达国家在大型公共设施建设，管道、网线铺设以及航班时刻表的编排等方面早已普遍应用运筹学的理论和方法，既省钱，省力又提高效率。可惜，运筹学的应用在我国还不普遍.

其实我们不能要求决策者本人一定要懂很多数学，但至少他们要经常想想工作中有没有数学问题需要请数学家来咨询。

数学修养对于国民素质的影响，正如美国国家研究委员会发表的“人人关心数学教育的未来”一书中所说：“除了经济以外，对数学无知的社会和政治后果给每个民主政治的生存提出了惊恐的信号。因为数学掌握着我们的基于信息的社会的领导能力的关键”[8]。

4.2 对于教学改革的几点意见

“十年树木，百年树人”说明教育的成果需要经过相当长的时间才能收获。因此教学改革的效果也不可能立竿见影。这就决定了教学改革只能“渐进”不能“革命”。20世纪中期美国的“新数学运动”以及1958－1960年中国的“教育大革命”的历史教训必须记取!

要“改革”就可能有成功也可能有失败，而且成败未必就那么容易察觉，有时很可能所得之处就含有所失，所以做改革实验之前必须考虑到可能出现的问题与补救方法。

我们应当鼓励实验的多样化．事实上每个教师都可以通过自己的教学实践对具体教学内容进行改革，这是应当受到鼓励的。所以教学改革的关键在教师，特别是教师的学术水平和知识视野。

我对于数学教学改革的具体意见是：正确处理好数学理论教学中：一般与特殊、抽象与具体、形式与实质的关系。特别在讲述中，要避免过分形式化。大多数

人学习数学并不是为了从事专门的纯数学研究，形式化的教学会使人或如堕云雾，或如隔靴搔痒，甚至令人望而生畏。即使是培养专门的纯数学研究人才，形式化方法有时虽有其直接了当、逻辑清晰的优点，但过于形式化也不利于更深刻的理解。

我们不仅要关注主修数学学科学生的教学改革，也要关心其它学科的数学课程改革。事实上，数学在其它学科中应用的新的生长点往往首先是由该学科的研究者开始的，而且要使数学家能够进入这个领域工作，也必须有该学科的研究者的帮助与支持。在这个意义上说其它学科数学课程的改革和数学学科的课程改革一样重要。

4.3 人人学好数学

我们不必过分夸大数学需要特殊的才能。数学特别难的印象往往是由于数学的书和文献在表达中过于形式化的缘故。如果课堂教学是干巴巴的“定义一定理一推理”形式地讲，自学时也是亦步亦趋地跟着复习，那么必然会感到枯燥乏味。但如果喜爱数学，而且“教”与“学”都得法，普通中等才能的人照样可以学好数学，顺利地完成大学数学的学业。然而学习方法很重要，每个人要根据学习的不同阶段，来调整自己的学习方法。不断认识自己，明确目标，不断改进学习方法。

5 中国青年数学家的使命

5.1 “中国要成为数学大国”

中国没有理由不能成为数学大国。

第一，中国有辉煌的古代数学一祖冲之、刘徽等都遥遥领先于他们的同辈西方学者。只是由于我国的封建社会太长，有很长一段时间不鼓励科学发展，才落后于西方。

第二，老一辈数学家在20世纪初才从西方引进近代数学的“火种”。在不到100年这段期间，还经历过八年抗战和十年“文化革命”的灾难，几代数学家艰苦奋斗，承上启下，终于以ICM’2002在北京召开为标志，登上了世界数学舞台。

然而怎样才算“数学大国”呢？我认为：

第一，在基础研究方面能在有重大意义的问题上，做原创性的、有自己特色的工作。或者是对数学的有机整体作出贡献，或者是在交叉学科中独辟蹊径。我们要逐渐改变跟在别人后面走的状态，争取引领潮流，逐渐形成中国自己的学派。

第二，在应用研究方面，中国数学家要为自己的国家，包括科学技术、国防建设、经济建设等各个方面做贡献，使数学真正扎根在我国自己的土地上。

我们在这方面确实还有相当长的路要走。过去我国自主创新的产品与我国的经济状况很不适应。许多在发达国家工商业界早已应用成熟的数学理论和方法在我国还没有需求，也应用不上。因此我国和世界强国在研究基金和数学毕业生就业方面差别很大。以美国为例，美国数学研究基金除NSF外，还来自海军、空军、陆军、国家安全局、高技术局、宇航局、能源部、健康医疗(NIH)等很多方面。除此之外，在美国，不仅传统的科技领域，而且金融、保险、医药、信息、交通运输、材料等等行业也大量应用数学。所以学数学的学生出路很广，除了大学和研究机构外，还有许许多多大大小小的公司雇用数学家。不管经济好坏，不大会有拿了数学博士学位而没有职业的情况。这是因为：数学已经成为他们社会发展的需要。

现在我国经济的发展已经到了提高GDP中科技含量的阶段，对于我国青年数学家来说这是一个空前的机会，也一定是大有作为的！真正用数学来提高我国的科技、国防、经济、管理各方面的水平是我们大家共同努力的方向。

5.2 “团结自信”，抗拒“诱惑”，“锲而不舍”

青年人要有充分的自信。“数学是年青人的学问”。大家都知道天才的Abel，Galois在很年轻的时候就做出了划时代的贡献。如今尽管数学的内容已经如此丰富，体系如此庞大，研究人员如此众多，然而真正有能力的青年数学家照样可以脱颖而出！每四年一次的Fields奖就是奖给40岁以下青年数学家的。从历届Fields奖得主的成就来看“数学是年青人的学问”这句话至今依然未变。

我国当今青年一代数学家享有中国历史上最好的学习条件和工作条件. 包括图书资料、网络信息和学术交流等方面都与发达国家相差无几了。因此没有理由说在中国不能做出第一流的成果。问题在于当今我们的学术环境不理想：急功近利，虚夸浮躁，正在腐蚀人们的思想，败坏我们的学风。中国有志气的青年数学家要自觉抗拒各种“诱惑”、抵制学术不端行为；要继承优良学术传统，要脚踏实地，不畏艰难，锲而不舍，团结奋斗；这样就一定能够实现中国的数学大国—强国之梦。

参考文献

[1] 柯尔莫果洛夫，关肇直译。论数学职业。数学通报，1953，5

[2] Doing the Math to Find the Good Jobs, *Wall Street Journal*, Jan.6th2009, www.CareerCast.com

[3] Arrow, K J., *Social Choice and Individual Values*, John Wiley and Sons, 1951(中译本：社会选择与个人价值，成都，四川大学出版社。1957)

[4] Dreifuss, R., *Speech at ICM'94, Proc.of ICM'94*. Zurich, Birkhauser, 1995, pp.24-27

[5] Kline, M., *Mathematics and the search for knowledge*, Oxford UniversityPress, 1986(中译本：数学与知识的探求，上海复旦大学出版社，2005)

[6] Courant, R., Robbins, H., *What is Mathematics*, (中译本：什么是数学——对思想和方法的基本研究，上海：复旦大学出版社，2005)

[7] 胡作玄，邓明立。大有作为的数学。石家庄：河北教育出版社，2006

[8] 美国国家研究委员会。人人关心数学教育的未来。北京：世界图书出版公司，1993

****

Mathematical Quotations

Dieudonné, Jean Alexandre Eugéne (1906 - 1992)

一般来说，数学倾向的觉醒大约在16岁，……可是，与通常流行的见解相反，最早显露出创造性的时期很少在20岁到25岁之前。

——数学家与数学发展，1978

方阵 AB 与 BA 的关系研究（一）

李奇芮 [1]

一、　摘要

设 A、B 是两个同阶方阵，通常有 AB≠BA，但是，AB 与 BA 的特征多项式一定相等。受此现象启发，本文研究 AB 与 BA 之间有多大相似度。

符号说明：为行文方便，文中记方阵 A^k 的秩为 $r(A)^k$。

本文的研究主要有以下几方面的结论：

1、线性变换 AB 与线性变换 BA 具有相同的特征多项式。因而具有相同的特征根（重数计入）且有相同维数的根子空间。

2、线性变换 AB 与 BA 在相同非零特征根 λ 的根子空间上的限制相似。

3、AB 与 BA 的最小多项式最多相差一个因子 λ 。

定义**幂秩函数** $f_A(k)=r(A)^k$。

定义**幂秩降速** $v_A(k)=f_A(k)-f_A(k+1)$。

4、AB 与 BA 相似当且仅当它们的幂秩函数相同。

5、幂秩降速 $v_A(k)$ 单调不增。

6、双重限制关系成立：

$$f_{AB}(k+1)\le \min\{f_{AB}(k),f_{BA}(k)\}$$

$$v_{AB}(k+1)\le \min\{v_{AB}(k),v_{BA}(k)\}$$

二、　正文

（一）引论

结论 1、如果 $f(x)$ 是一个没有 0 根的多项式，那么有如下结论：

$$null(f(AB))=null(f(BA)) \qquad \det(f(AB))=\det(f(BA))$$

结论 2、弗洛比尼乌斯不等式：

$$r(AB)-r(ABC)\le r(B)-r(BC)$$

简单的证明：

现在我们来证明结论一。

[1] 基科 91

取$P=\begin{pmatrix} I & -B \\ & I \end{pmatrix}$ $Q=\begin{pmatrix} I & \\ A & I \end{pmatrix}$则有：$PQ\begin{pmatrix} I & \\ & I-AB \end{pmatrix}P^{-1}Q^{-1}=\begin{pmatrix} I-BA & \\ & I \end{pmatrix}$

于是我们有 $null(I\pm AB))=null(I\pm BA)$，$\det(I\pm AB))=\det(I\pm BA)$

首先，对任意方阵 A 和 B，有 $null(I\pm AB))=null(I\pm BA)$，因此：

$$
\begin{aligned}
&null(f(AB))=null(cI+ABg(AB))=null(cI+Bg(AB)A)=null(cI+BAg(BA)) \\
&=null(f(BA))
\end{aligned}
$$

其中 c≠0。第二个结论 $\det(f(AB))=\det(f(BA))$ 证明如法炮制。对于弗洛比尼乌斯不等式，证明请参考《高等代数学》（张贤科，许甫华）112 页，这里不再赘述。

AB 与 BA 在非零根子空间上限制必相似，因为引论，设 λ_0 为其一非零的特征根，于是我们有，对任意的 k，$null(\lambda_0 I-AB)^k=null(\lambda_0 I-BA)^k$，这样就证明了 AB 和 BA 的形如 $(\lambda-\lambda_0)^k$ 的初等因子完全相同，因而其在根子空间上的限制相似。

我们发现，引理在当 f(0)=0 的时候不一定成立，这样就说明 AB 与 BA 在 0 根子空间上的限制的方阵表示可能不相似，因此，AB 与 BA 在空间结构上的差异，全部反映在 0 根子空间内。

为了研究零根子空间内的结构，我们只需要研究 AB 和 BA 的幂秩关系，即 $r(AB)^k$ 与 $r(BA)^k$，若对任意的 k 满足 $r(AB)^k=r(BA)^k$，则 AB 与 BA 相似。因此，AB 与 BA 在零根子空间上的差异全部反映在这两个秩的关系上。

（二）幂秩函数以及其性质

对一个 n 阶方阵 A，**我们定义非负整值函数** $f_A(k)=r(A)^k$，把这个函数叫做 A 的**"幂秩函数"**，k 的定义域为所有非负整数。

我们定义 A 的幂秩函数的一阶差分 $v_A(k)=f_A(k)-f_A(k+1)$ 叫做 A 的**幂秩降速**。

我们有如下性质：

性质一、$v_A(k)$ 非负且单调不增。

这个性质文中称为“幂秩降减速性”。

证明：

由于 $r(A)^{k+1}=r(A\times A^k)\le r(A)^k$ 因而 $v_A(k)=r(A)^k-r(A)^{k+1}\ge 0$，故 $v_A(k)$ 非负。

由弗洛比尼乌斯不等式，有 $r(A)^{k+1}-r(A)^{k+2}\le r(A)^k-r(A)^{k+1}$ 即 $v_A(k+1)\le v_A(k)$，于是 $v_A(k)$ 单调不增。证毕。

定义：设 $K=\{k\mid f_A(k)=f_A(k+1)\}$，取 t 为集合 K 中的最小元，那么我们把 t 的值叫做 A 的秩稳定点。而把这一点的取值 $f_A(t)$ 叫做 A 的秩稳定值。

定理 1：设 t 是方阵 A 的秩稳定点，那么对于任意的 s≥t，均有： $f_A(s)=f_A(t)$

证明：因为幂秩降减速性，对任意的 k≥t 由于 $0\le v_A(k)\le v_A(t)=0$，故有 $v_A(k)=0$，而 $f_A(t)-f_A(s)=\sum_{k=t}^{s-1}v_A(k)=0$，证毕。

定理二：设 t 是 A 的秩稳定点，则 t≤n

证明：反设 t>n，于是对任意的 k<n，一定有 $v_A(k)>0$ 即 $v_A(k)\ge 1$ 而如果这样，则 $r(A)^{n+1}=n-\sum_{k=0}^{n}v_A(k)\le n-(n+1)=-1$ 矛盾。

引理 1：A 的秩稳定值等于 A 的所有非零根子空间的维数之和。

证明：

只要说明 A^n 的秩等于 A 的所有非零根子空间的维数之和即可。

我们把 A 化成 Jordan 标准型，易知，特征非零的准素块都是满秩的，而特征为 0 的准素块一定是幂零方阵，所以，A^n 的秩等于 A^n 的所有准素块的秩的和，因为 A^n 的特征为 0 的准素块是 0 方阵，而特征非零的准素块是满秩的，所以 A^n 的秩等于所有特征非零准素块的阶数之和，从而等于 A 的所有非零根子空间的维数之和。

接下来我们来看 AB 与 BA 的幂秩关系。

注意，下文中，我们设 A 和 B 分别是 m×n 和 n×m 的矩阵。

（三）AB 与 BA 的幂秩以及其双重限制性

（**幂秩的双重限制性**）AB 与 BA 的幂秩函数和幂秩降速满足如下关系：

$$f_{AB}(k+1) \le \min\{f_{AB}(k), f_{BA}(k)\}$$
$$v_{AB}(k+1) \le \min\{v_{AB}(k), v_{BA}(k)\}$$

证明：因为$r(AB)^{k+1} = r(A(BA)^k B) \le r((BA)^k B) \le r(BA)^k$，则我们有：

$f_{AB}(k+1) \le \min\{f_{AB}(k), f_{BA}(k)\}$；

由弗洛比尼乌斯不等式，我们有

$r(AB)^{k+1} - r(AB)^{k+2} = r(A(BA)^k B) - r(A(BA)^{k+1} B)$
$\le r((BA)^k B) - r((BA)^{k+1} B) \le r(BA)^k - r(BA)^{k+1}$

于是$v_{AB}(k+1) \le \min\{v_{AB}(k), v_{BA}(k)\}$

我们看一些简单的推论：

AB 与 BA 有相同的秩稳定值。

因为 AB 与 BA 在非零根子空间上的限制相似，于是 AB 与 BA 做准素分解后，非零的准素块的阶数对应相等，按照前文的引理 1，AB 与 BA 有相同的秩稳定值。

AB 与 BA 的最小多项式最多只差一个因子λ

不妨设 AB 的最小多项式次数比 BA 的最小多项式次数高，设 AB 的最小多项式是 $f(\lambda)$，BA 的最小多项式是$g(\lambda)$。

我们知道，AB 与 BA 在非零根子空间上的限制是相似的，于是其最小多项式的差异只能在因子λ上。

设 BA 的最小多项式λ因子的次数是 k，因此，BA 在 k 点便到达秩稳定点，则$v_{BA}(k) = 0$，由于幂秩的双重限制性，我们有$0 \le v_{AB}(k+1) \le v_{BA}(k) = 0$于是 AB 必然在 k+1 点到达秩稳定点。因此 $f(\lambda) = \lambda g(\lambda)$。

S^1的$C^{(\infty)}$类结构同构的一个证明

苏桃*

致谢: 特别衷心感谢杨鑫、杜升华两位学长对本稿的修改、帮助. 他们对本文进行了仔细认真地审阅, 并提出了很多宝贵的修改意见, 使本文较初稿有较大改观, 其中不少错误、不妥善之处均由他们提出并改进.

1 问题

文 [1] 中的一道习题引发了笔者的一些思考, 发现这与我以前想过的一个问题非常有联系. 以下是得到的相关内容:

命题1. 圆周 S^1 (一维球面)上, 任意两个$C^{(\infty)}-$结构同构.

这是 [1] 文中的一道习题 ($Page$ 295, 3), 本文将对此进行证明.

大致思路: 从 $\mathbb{R}$ 的情形获得启发: $\mathbb{R}$ 的两个 $C^{(\infty)}-$结构图册 (等价图册的并) A_R, B_R, 若有简单图 $\varphi \in A_R$, $\psi \in B_R$, 则直接取同构映射 $f: \mathbb{R}_A^1 \to \mathbb{R}_B^1$, $f = \psi \circ \varphi^{-1}$ 即可. 所以大体思路是将两个局部图拼成一个图, 对 $\mathbb{R}$ 的情形这样做就已经可以了. 对 S^1 的情形, 这样还不够, 问题在于一张图不可能盖住 S^1 , 想法是用两张图来覆盖, 且它们在两部分边界均光滑衔接, 所以证明的关键是如何处理使最后找到的两张图的边界光滑衔接, 这就是下面将要证明的引理 (2) .

2 引理

为证命题, 我们先做一些准备:

引理1. 任给两个实数列 $\{a_n\}_{(n\geq 0)}$, $\{b_n\}_{(n\geq 0)}$ 及闭区间 $[a,b] (a < b)$, 存在函数 $f \in C^{(\infty)}[a,b]$, 使得 $f^{(n)}(a) = a_n$, $f^{(n)}(b) = b_n$ $(\forall\, n \geq 0)$.

证明. 证明见文 [2] $Page$ 44. □

*基科81

引理2. 任给两个实数列$\{a_n\}_{n\geq 0}$, $\{b_n\}_{n\geq 0}$, $a_0<b_0$, $a_1>0$, $b_1>0$, 及 $[a,b]\subset\mathbb{R}(a<b)$, 存在函数 $f:\mathbb{R}\to\mathbb{R}$ 为 $C^{(\infty)}$ 类微分同胚, 使 $f^{(i)}(a)=a_i, f^{(i)}(b)=b_i(i\geq 0)$.

证明. 利用已有的引理 1 来证明引理 2 :

考虑 $\widetilde{f}=\dfrac{f((b-a)t+a)-a_0}{b_0-a_0}$, 有 $\widetilde{f}(0)=0,\ \widetilde{f}(1)=1$. 故只需证明 $a=0,\ b=1,\ a_0=0,\ b_0=1$ 的情形.

只需证: 存在 $\widetilde{f}$ (以后仍记为f) : $[0,1]\to[0,1]$ 为 $C^{(\infty)}$ 类微分同胚, 满足引理 2 中相应条件. $(*)$

因为可分别构造 $f_n:[n,n+1]\to[n,n+1]$ 为 $C^{(\infty)}$ 类微分同胚, 使相邻两个函数在共同端点处取相同的各阶导数值, 将这些函数首尾衔接即可.

若 f 存在, 取 $g=f'$, 有

$g^{(i)}(0)=a_{i+1},\ g^{(i)}(1)=b_{i+1}\ (i\geq 0)$, 且 $\int_0^1 g(t)dt=1,\ g(t)\geq 0\ (\forall t\in[0,1])$.

反过来, 若存在 $g>0$ 满足上述条件, 则取 $f(x)=\int_0^x g(t)dt\ (x\in[0,1])$ 满足 $(*)$ 条件 .

下面只需构造符合条件的 $g(g>0)$: $(**)$

由引理 1 可取 $g_0\in C^{(\infty)}(([0,1]),\mathbb{R})$, 使

$$g_0^{(i)}(0)=a_{i+1},\ g_0^{(i)}(1)=b_{i+1}(i\geq 0).$$

注意到若 $\varphi\in C^{(\infty)}(\mathbb{R},\mathbb{R})$ 满足:

$$(\varphi^{(i)}(a_1))=(a_1,1,0,\ldots),\ (\varphi^{(i)}(b_1))=(b_1,1,0,\ldots).$$

(以后记 $(h^{(i)}(x))$ 为 h 在 x 处的各阶 (含 0 阶) 导数值序列). 则验证知:

$$(\varphi\circ g_0)^{(i)}(0)=\varphi'(a_1)\cdot g_0^{(i)}(0)=a_{i+1}(i\geq 1),\ \varphi\circ g_0(0)=a_1;$$
$$(\varphi\circ g_0)^{(i)}(1)=g_0^{(i)}(1)=b_{i+1}(i\geq 0).$$

即 $\varphi\circ g_0$ 保持 g_0 的性质.

同理可证对 $\psi\in C^{(\infty)}([0,1],[0,1])$, 若

$$(\psi^{(i)}(0))=(0,1,0,\ldots),\ (\psi^{(i)}(1))=(1,1,0,\ldots).$$

则 $g_0\circ\psi$ 仍保持 g_0 的性质.

目标找合适的 φ, ψ 使 $g=\varphi\circ g_0\circ\psi$ 满足条件.

引理3. 对任意 $a>0$, 存在 $f\in C^{(\infty)}([0,1],[0,a])$ 为单调函数, 且 $(f^{(i)}(0))=(0,\ldots)$, $(f^{(i)}(1))=(a,1,0,\ldots)$.

证明. 由引理 1 可构造 $f_1 \in C^{(\infty)}(\mathbb{R},\mathbb{R}),\ supp\, f_1 \subset [0,1],\ f_1 \geq 0$ 而不恒为 0. (先构造 $h \in C^{(\infty)}(\mathbb{R},\mathbb{R}),\ supp\, h \subset [0,1]$, 且不恒为 0. 再取$f_1 = h^2$).

令 $f_2(x) = \int_{-\infty}^{x} f_1(t)dt\ (\forall x \in \mathbb{R})$, 知

$$f_2 \in C^{(\infty)}(\mathbb{R},\mathbb{R}),\ f_2(x) = 0\ (\forall x \leq 0),\ f_2(1) = \int_0^1 f_1(t)dt > 0.$$

完全可取 f_1 使 $f_2(1) = 1$ (用 Af_1 (适当的 $A > 0$) 代替 f_1 即可).

令 $F_1(x) = \int_{-\infty}^{x}(f_2(t) + \lambda f_1(t))dt\ (\lambda \geq 0$ 待定), 有

$$F_1 \in C^{(\infty)}(\mathbb{R},\mathbb{R}),\ F_1^{(i)}(0) = f_2^{(i-1)}(0) + \lambda f_1^{(i-1)}(0) = 0 (i \geq 1),$$

而

$$F_1(1) = \int_0^1 (f_2(t) + \lambda f_1(t))dt = \int_0^1 f_2(t)dt + \lambda,\ F_1'(1) = 1,$$
$$F_1^{(i)}(1) = f_2^{(i-1)}(1) + \lambda f_1^{(i-1)}(1) = 0 (i \geq 2).$$

可取 $\lambda \geq 0$ 足够大, 使 $F_1(1) > a$.

令 $F(x) = \int_{-\infty}^{x} f_2(t)dt$, 同样

$F^{(i)}(0) = 0\ (i \geq 0),\ F(1) = \int_0^1 f_2(t)dt > 0,\ F'(1) = 1,\ F^{(i)}(1) = 0\ (i \geq 2).$

取 $F_2(x) = b \cdot F(1 + \dfrac{x-1}{b}),\ 0 < b < 1$, 有

$F_2(0) = b \cdot F(1 - \dfrac{1}{b}) = 0,\ F_2^{(i)}(0) = b \cdot (\dfrac{1}{b})^i \cdot F^{(i)}(1 - \dfrac{1}{b}) = 0\ (i \geq 1),$

而

$$F_2(1) = b \cdot F(1),\ F_2'(1) = F'(1) = 1,\ F_2^{(i)}(1) = b \cdot (\frac{1}{b})^i \cdot F^{(i)}(1) = 0\ (i \geq 2).$$

取 $0 < b < 1$ 足够小, 使 $0 < F_2(1) < a$.

这样得到了两个函数 $F_1,\ F_2$ 单调增, $F_1(1) = a_1 > a,\ F_2(1) = a_2 < a$. 令

$$f(x) = \frac{a - a_2}{a_1 - a_2} F_1(x) + \frac{a_1 - a}{a_1 - a_2} F_2(x).$$

验证知 f 满足引理 3 的条件, 引理 3 得证.

注: 通过平移、对称、伸缩等操作引理 3 对以下情形亦成立:

a) 对 0, 1 处的导数值要求互换 (作图可以非常直观).

b) 闭区间 $[0,1],\ [0,a]$ 换成任何闭区间.

可设 $h : [0,1] \to [0, \dfrac{d-c}{b-a}]$, 满足引理 3 条件, 则取 $f : [a,b] \to [c,d]$ 为

$$f(x) = c + (b-a) \cdot h(\frac{x-a}{b-a}).$$

知: $(f^{(i)}(a)) = (c, 0, \ldots),\ (f^{(i)}(b)) = (d, 1, 0, \ldots)$.

c) 改成相应的单调减情况. □

回到引理 2 :

先设 $a_1 \neq b_1$, 可不妨设 $a_1 < b_1$. 由 g_0 的连续性, 存在 $t_0 \in [0,1]$ 使 $g_0(t_0) = \dfrac{a_1+b_1}{2}$.

首先来构造 φ_1, ψ_1 使 $g_1 = \varphi_1 \circ g_0 \circ \psi_1$ 满足: $\int_0^1 g_1(t)dt > 1$, $g_1 > 0$:
$\forall \epsilon > 0$, 取 $M > 0$ 使 $g_0 < M$. 由引理 3 :
可取 $\varphi_{11} \in C^{(\infty)}([0,a_1],[\epsilon,a_1])$ 单调增, 且

$$(\varphi_{11}^{(i)}(0)) = (\epsilon, 0, \ldots),\ (\varphi_{11}^{(i)}(a_1)) = (a_1, 1, 0, \ldots);$$

取 $\varphi_{12} \in C^{(\infty)}([a_1, \dfrac{a_1+b_1}{2}],[a_1,M])$ 单调增, 且

$$(\varphi_{12}^{(i)}(a_1)) = (a_1, 1, 0, \ldots),\ (\varphi_{12}^{(i)}(\frac{a_1+b_1}{2})) = (M, 0, \ldots);$$

取 $\varphi_{13} \in C^{(\infty)}([\dfrac{a_1+b_1}{2}, \dfrac{a_1}{4} + \dfrac{3b_1}{4}],[\epsilon,M])$ 单调减, 且

$$(\varphi_{13}^{(i)}(\frac{a_1+b_1}{2})) = (M, 0, \ldots),\ (\varphi_{13}^{(i)}(\frac{3b_1}{4})) = (\epsilon, 0, \ldots);$$

取 $\varphi_{14} \in C^{(\infty)}([\dfrac{3b_1}{4}, b_1],[\epsilon,b_1])$单调增, 且

$$(\varphi_{14}^{(i)}(\frac{3b_1}{4})) = (\epsilon, 0, \ldots),\ (\varphi_{14}^{(i)}(b_1)) = (b_1, 1, 0, \ldots);$$

取 $\varphi_{15} \in C^{(\infty)}([b_1, b_1+\epsilon],[b_1, b_1+\epsilon])$ 单调增,且

$$(\varphi_{15}^{(i)}(b_1)) = (b_1, 1, \ldots),\ (\varphi_{15}^{(i)}(b_1+\epsilon)) = (b_1+\epsilon, 0, \ldots);$$

最后取 φ_1 :

$$\varphi_1(x) = \begin{cases} \epsilon & x \in]-\infty, 0] \\ \varphi_{11} & x \in [0, a_1] \\ \varphi_{12} & x \in [a_1, \frac{a_1+b_1}{2}] \\ \varphi_{13} & x \in [\frac{a_1+b_1}{2}, \frac{a_1}{4} + \frac{3b_1}{4}] \\ \varphi_{14} & x \in [\frac{3b_1}{4}, b_1] \\ \varphi_{15} & x \in [b_1, b_1+\epsilon] \\ b_1+\epsilon & x \in [b_1+\epsilon, +\infty[\end{cases}$$

有 $\varphi_1 \in C^{(\infty)}(\mathbb{R},\mathbb{R})$, $\varphi_1 > 0$, $((\varphi_1^{(i)})(a_1)) = (a_1, 1, 0, \ldots)$, $(\varphi_1^{(i)}(b_1)) = (b_1, 1, 0, \ldots)$.

再取 $0 < \delta < \frac{1}{2}$. 类似存在 $\psi_1 \in C^{(\infty)}([0,1],[0,1])$ 单调非减
$(\psi_1^{(i)}(0)) = (0, 1, 0, \ldots)$; $(\psi_1^{(i)}(1)) = (1, 1, 0, \ldots)$; $(\psi_1^{(i)}(\delta)) = (t_0, 0, \ldots)$;
$(\psi_1^{(i)}(1-\delta)) = (t_0, 0, \ldots)$, 且 $\psi_1|_{[\delta,1-\delta]} = t_0$.

考虑 $g_1=\varphi_1\circ g_0\circ\psi_1\in C^{(\infty)}([0,1],\mathbb{R})$, 有 $g_1>0$, 且

$$\int_0^1 g_1(t)dt=\int_0^1\varphi_1\circ g_0\circ\psi_1(t)dt\ge\int_\delta^{1-\delta}\varphi_1\circ g_0\circ\psi_1(t)dt$$

$$=\int_\delta^{1-\delta}\varphi_1\circ g_0(t_0)dt=\int_\delta^{1-\delta}\varphi_1(\frac{a_1+b_1}{2})dt=M(1-2\delta).$$

取 $M>0$ 足够大而 $\delta>0$ 足够小, 使 $\int_0^1 g_1(t)dt>1$.

再来构造 φ_2, ψ_2, 使 $g_2=\varphi_2\circ g_0\circ\psi_2\in C^{(\infty)}([0,1],\mathbb{R})$, $g_2>0$, $0,1$ 处导数值同 g_1, 且 $0<\int_0^1 g_2(t)dt<1$：

可取 $\epsilon>0$ 使 $\epsilon<a_1<b_1$, 类似 φ_1 的构造可取 $\varphi_2\in C^{(\infty)}(\mathbb{R},\mathbb{R})$, 使

$$0<\varphi_2\le 2M,\ (\varphi_2^{(i)}(a_1))=(a_1,1,0,\ldots),$$

$$(\varphi_2^{(i)}(b_1))=(b_1,1,0,\ldots),\ (\varphi_2^{(i)}(\frac{a_1+b_1}{2}))=(\epsilon,0,\ldots).$$

而取 $\psi_2=\psi_1$.

令 $g_2=\varphi_2\circ g_0\circ\psi_2\in C^{(\infty)}([0,1],\mathbb{R})$, 有 $g_2>0$, $0,1$ 处导数值同 g_1. 而

$$\begin{aligned}\int_0^1 g_2(t)dt &= \int_0^1\varphi_2\circ g_0\circ\psi_2(t)dt\\ &= \int_0^\delta\varphi_2\circ g_0\circ\psi_2(t)dt+\int_{1-\delta}^1\varphi_2\circ g_0\circ\psi_2(t)dt\\ &+ \int_\delta^{1-\delta}\varphi_2\circ g_0\circ\psi_2(t)dt\\ &\le 2M\cdot 2\delta+\int_\delta^{1-\delta}\varphi_2\circ g_0\circ\psi_2(t)dt\\ &= 4M\delta+\int_\delta^{1-\delta}\varphi_2(\frac{a_0+b_0}{2})dt\\ &= 4M\delta+\epsilon(1-2\delta).\end{aligned}$$

可取 $\delta>0$, $\epsilon>0$ 足够小, 使 $0<\int_0^1 g_2(t)dt<1$.

最后, 设 $\int_0^1 g_1(t)dt=A_1$, $\int_0^1 g_2(t)dt=A_2$, $0<A_2<1<A_1$. 令

$$g=\frac{1-A_2}{A_1-A_2}g_1+\frac{A_1-1}{A_1-A_2}g_2.$$

验证知: g 满足 $(**)$ 的所有条件, 从而 $(**)$ 对 $a_1\ne b_1$ 成立.

对 $a_1=b_1$, 可取 $\dfrac{1}{2}$ 及 $c>0,\ c\neq a_1$. 由已证的结论可构造
$\widetilde{g_1}\in C^{(\infty)}([0,\dfrac{1}{2}],\mathbb{R}),\ \widetilde{g_1}>0$, 且

$$(\widetilde{g_1}^{(i)}(0))=(a_{i+1})_{(i\geq 0)},\ (\widetilde{g_1}^{(i)}(\frac{1}{2}))=(c,0,\ldots),\ \int_0^{\frac{1}{2}}\widetilde{g_1}(t)dt=\frac{1}{2}.$$

$\widetilde{g_2}\in C^{(\infty)}([\dfrac{1}{2},1],\mathbb{R}),\ \widetilde{g_2}>0$, 且

$$(\widetilde{g_2}^{(i)}(\frac{1}{2}))=(c,0,\ldots),\ (\widetilde{g_2}^{(i)}(1))=(b_{i+1}),\ \int_{\frac{1}{2}}^{1}\widetilde{g_2}(t)dt=\frac{1}{2}.$$

取 $g\in C^{(\infty)}([0,1],\mathbb{R}),\ g|_{[0,\frac{1}{2}]}=\widetilde{g_1},\ g|_{[\frac{1}{2},1]}=\widetilde{g_2}$. 则 g 满足 $(**)$ 条件.

综上, 引理 2 证毕. □

3 命题的证明

断言 (1) : 对 S^1 的任一个 $C^{(\infty)}-$结构 M (等价图册类), 设其结构图册为 A_M, 即

$$A_M=\bigcup_{B\in M}B$$

则存在 $\varphi:\mathbb{R}\mapsto S^1$ 为局部同胚, $\varphi(\mathbb{R})\supset S^1$ 为包含 S^1 的一条弧线 (有重叠部分), 设为 $\widehat{ad}$, 随自变量的增大而在 S^1 上沿逆时针方向运动 (以后称为递增的). 而设 $\widehat{aa'}$, $\widehat{d'd}$ 为 $\widehat{ad}$ 上的重合部分不自交, 即 $\widehat{aa'}=\widehat{d'd}\subset S^1$ (以后形如 $\widehat{AB}$ 的弧均表示沿逆时针方向: $A\to B$) (见图1). 有如下特征:

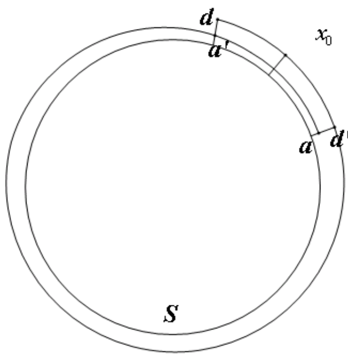


图 1

(a) : 任一点 $x\in\widehat{aa'}$, $\widehat{xx}=S^1\setminus\{x\}$, 存在 $]a,b[\subset\mathbb{R},\ \varphi:]a,b[\to S^1\setminus\{x\}$ 为拓扑同胚. 记 $\overline{\varphi}=\varphi|_{]a,b[}$, 则对任意 $(\widetilde{U},\widetilde{\varphi})\in A_M$, $\overline{\varphi}^{-1}\circ\widetilde{\varphi},\ \widetilde{\varphi}^{-1}\circ\overline{\varphi}$ 在相应定义域上为 $C^{(\infty)}$ 类函数, 即 $\overline{\varphi}$ 与 $\widetilde{\varphi}$ 等价.

(b) : 存在 $[t_1, t_2]$ $(t_1 < t_2) \subset \mathbb{R}$, 使得 $\varphi(t_1) = \varphi(t_2) = x_0 \in \widehat{aa'}$. 且:
$\varphi :]t_1, t_2[\mapsto S^1 \setminus \{x_0\}$ 为拓扑同胚, 而存在 t_1, t_2 的邻域 $O(t_1)$, $O(t_2) \subset \mathbb{R}$, 使 $\varphi_1 = \varphi|_{O(t_1)}$, $\varphi_2 = \varphi|_{O(t_2)}$ 满足: $\varphi_2^{-1} \circ \varphi_1 :\ O(t_1) \to O(t_2)$ 为 $C^{(\infty)}$ 类微分同胚.

(c) : 在 (b) 中, φ 还满足:

$((\varphi_2^{-1} \circ \varphi_1)^{(i)}(t_1)) = (t_2, 1, 0, \ldots)$, $((\varphi_1^{-1} \circ \varphi_2)^{(i)}(t_2)) = (t_1, 1, 0, \ldots)$.

注: 以后若有 φ 满足除 (b), (c) 外断言 (1) 中的其余条件, 则称 φ 与 A_M 中元素等价. (重要!)

若断言 (1) 成立, 则对 S^1 的任两个 $C^{(\infty)}-$结构 M_1, M_2, 流形分别记为 S^1_A, S^1_B, 设对应结构图册为 A_M, B_M.

可取 φ 对应于 A_M, ψ 对应于 B_M. φ, ψ 分别满足断言 (1) 中条件, 且对应断言中开区间 $]t_1, t_2[$, $]t_1', t_2'[$.

设 $\varphi(]t_1, t_2[) = S \setminus \{x_1\}$, $\psi(]t_1', t_2'[) = S \setminus \{x_2\}$. 令

$$T(t) = \frac{t_2' - t_1'}{t_2 - t_1}(t - t_1) + t_1' \ (t \in \mathbb{R})$$

知 $T : \mathbb{R} \to \mathbb{R}$ 为 $C^{(\infty)}$ 类微分同胚, 且 $T(t_1) = t_1'$, $T(t_2) = t_2'$.

在这样的条件下, 构造映射: $f :\ S^1_A \to S^1_B$.

$$f(x) = \begin{cases} \overline{\psi} \circ T \circ \overline{\varphi}^{-1}(x) & x \neq x_1 \\ x_2 & x = x_1 \end{cases}$$

(其中 $\overline{\varphi} = \varphi|_{]t_1, t_2[}$, $\overline{\psi} = \psi|_{]t_1', t_2'[}$).
下面证明: 这样得到的 $f :\ S^1_A \to S^1_B$ 是 $C^{(\infty)}$ 类微分同胚, 从而 M_1, M_2 同构.

证明. 事实上, $\forall\ (\widetilde{U}, \widetilde{\varphi}) \in A_M, (\widetilde{V}, \widetilde{\psi}) \in B_M$, $\widetilde{\psi}^{-1} \circ f \circ \widetilde{\varphi}$ 的定义域设为 $W = \widetilde{\varphi}^{-1}(f^{-1}(\widetilde{V}) \cap \widetilde{U})$ (开集). 任取 $s_1 \in W$.

(i). 若 $\widetilde{\varphi}(s_1) \neq x_1$, 则存在邻域 $O(s_1) \subset W$, 使 $x_1 \notin \widetilde{\varphi}(O(s_1))$.
此时在 $O(s_1)$ 上:

$$\widetilde{\psi}^{-1} \circ f \circ \widetilde{\varphi}(s) = \widetilde{\psi}^{-1} \circ \overline{\psi} \circ T \circ \overline{\varphi} \circ \widetilde{\varphi}(s) = (\widetilde{\psi}^{-1} \circ \overline{\psi}) \circ T \circ (\overline{\varphi} \circ \widetilde{\varphi})(s).$$

由 ψ, φ 的性质即知, $\widetilde{\psi}^{-1} \circ f \circ \widetilde{\varphi}$ 在 $O(s_1)$ 上为 $C^{(\infty)}$ 类函数;

(ii). 若 $\widetilde{\varphi}(s_1) = x_1$, 则 $\exists \delta > 0$ 使 $]s_1 - \delta, s_1 + \delta[\subset W$, 且满足:
$\overline{\varphi}^{-1} \circ \widetilde{\varphi}|_{]s_1 - \delta, s_1]} = \varphi_1^{-1} \circ \widetilde{\varphi}|_{]s_1 - \delta, s_1]}$ 或 $\overline{\varphi}^{-1} \circ \widetilde{\varphi}|_{]s_1 - \delta, s_1]} = \varphi_2^{-1} \circ \widetilde{\varphi}|_{]s_1 - \delta, s_1]}$.
而限制在 $[s_1, s_1 + \delta[$ 上则相反. (其中 φ_1, φ_2 为断言 (1) 中对应 φ 的两个局部映射, ψ_1, ψ_2 对应 ψ). 可不妨设

$\overline{\varphi}^{-1}\circ\widetilde{\varphi}|_{]s_1-\delta,s_1]}=\varphi_1^{-1}\circ\widetilde{\varphi}|_{]s_1-\delta,s_1]}$, $\overline{\varphi}^{-1}\circ\widetilde{\varphi}|_{[s_1,s_1+\delta]}=\varphi_2^{-1}\circ\widetilde{\varphi}|_{[s_1,s_1+\delta]}$.
相应的有:

$$\begin{aligned}\widetilde{\psi}^{-1}\circ f\circ\widetilde{\varphi}|_{]s_1-\delta,s_1]}&=\widetilde{\psi}^{-1}\circ\psi_1\circ T\circ\varphi_1^{-1}\circ\widetilde{\varphi}|_{]s_1-\delta,s_1]}\\ \widetilde{\psi}^{-1}\circ f\circ\widetilde{\varphi}|_{[s_1,s_1+\delta[}&=\widetilde{\psi}^{-1}\circ\psi_2\circ T\circ\varphi_2^{-1}\circ\widetilde{\varphi}|_{[s_1,s_1+\delta[}.\end{aligned}$$

现在我们只需证明:

$$(\varphi_1^{-1}\circ\widetilde{\varphi})^{(i)}(s_1)=(\varphi_2^{-1}\circ\widetilde{\varphi})^{(i)}(s_1);\ (\widetilde{\psi}^{-1}\circ\psi_1)^{(i)}(t_1')=(\widetilde{\psi}^{-1}\circ\psi_2)^{(i)}(t_2')(i\geq 0)$$

则由复合函数求导及归纳证明可知 $(\widetilde{\psi}^{-1}\circ f\circ\widetilde{\varphi})^{(i)}(s_1)$ $(\forall i\geq 0)$ 存在. 从而 $\widetilde{\psi}^{-1}\circ f\circ\widetilde{\varphi}$ 在定义域中为 $C^{(\infty)}$ 类函数.

而 $\varphi_2^{-1}\circ\widetilde{\varphi}=(\varphi_2^{-1}\circ\varphi_1)\circ(\varphi_1^{-1}\circ\widetilde{\varphi})$, 由断言 (1) 中性质 (b) 及复合求导可得:

$$\begin{aligned}&(\varphi_2^{-1}\circ\varphi_1)\circ(\varphi_1^{-1}\circ\widetilde{\varphi})^{(i)}(s_1)=(\varphi_2^{-1}\circ\varphi_1)'(t_1)\cdot(\varphi_1^{-1}\circ\widetilde{\varphi})^{(i)}(s_1)\\ &=(\varphi_1^{-1}\circ\widetilde{\varphi})^{(i)}(s_1)\ (i\geq 0).\end{aligned}$$

同样可得: $(\widetilde{\psi}^{-1}\circ\psi_1)^{(i)}(t_1')=(\widetilde{\psi}^{-1}\circ\psi_2)^{(i)}(t_2')(i\geq 0)$. 于是上述需证结论成立.
同理可证 $\widetilde{\varphi}^{-1}\circ f\circ\widetilde{\psi}$ 在定义域中亦为 $C^{(\infty)}$ 类函数.

这样, $f:\ S_A^1\to S_B^1$ 为 $C^{(\infty)}$ 类同构映射. □

下面只需证明断言 (1) 成立:

断言 (1) 的证明. $\forall(U_1,\varphi_1),(U_2,\varphi_2)\in A_M$, 因 $(U,\varphi)\in A_M\Rightarrow(U,\varphi_-)\in A_M(\varphi_-(t)=\varphi(-t))$. 可取 φ_1, φ_2 均递增.[1] 若 U_1, U_2 互不包含, 且 $U_1\cap U_2\neq\emptyset$, 设 $x_0\in U_1\cap U_2$, $U_1=\widehat{ab}$, $U_2=\widehat{cd}$, 据条件可设 $\widehat{cb}\subset\widehat{ab}$, $\widehat{cd}$, 取 $x_0\in\widehat{cb}$.
(见图2). 可设 $\varphi_1^{-1}(\widehat{ax_0})=]-\infty,t_1]$, $\varphi_2^{-1}(\widehat{x_0d})=[t_2,+\infty[$.

断言 (2): 存在 $\varphi:\mathbb{R}\to\widehat{ad}$ 与 A_M 中元素等价[2] (若 $\widehat{ad}$ 不自交则是通常意义下的等价).

断言 (2) 的证明. 取 $\varphi:\mathbb{R}\to\widehat{ad}$:

$$\varphi(t)=\begin{cases}\varphi_1(t) & t\in]-\infty,t_1]\\ \varphi_2\circ\chi(t) & t\in[t_1,+\infty[\end{cases}$$

其中 $\chi:\ [t_1,+\infty[\to[t_2,+\infty[$ 为 $C^{(\infty)}$ 类微分同胚, 满足

$$\chi^{(i)}(t_1)=(\varphi_2^{-1}\circ\varphi_1)^{(i)}(t_1)(i\geq 0).$$

[1] 见断言 (1) 的陈述.
[2] 见断言 (1) 后的注.

注意到 $\varphi_2^{-1}\circ\varphi_1$ 是局部坐标的 $C^{(\infty)}$ 类微分同胚变换, 且单调增, 有
$(\varphi_2^{-1}\circ\varphi_1)'(t_1)>0$, 故由引理 2 知这样的 χ 存在.

我们来验证这样得到的 φ 满足题意: 这时要证明 $\forall\,(\widetilde{U},\widetilde{\varphi})\in A_M$ 与 $(\widehat{ad},\varphi)$ 等价.

考虑到 φ 的构造, 我们实际上只需对 φ_1, φ_2 证明即可.

这是因为在局部上可以考虑 φ^{-1}, 有
$\widetilde{\varphi}^{-1}\circ\varphi=(\widetilde{\varphi}^{-1}\circ\varphi_1)\circ(\varphi_1^{-1}\circ\varphi)$ 或 $(\widetilde{\varphi}^{-1}\circ\varphi_2)\circ(\varphi_2^{-1}\circ\varphi)$
$\varphi^{-1}\circ\widetilde{\varphi}=(\varphi^{-1}\circ\varphi_1)\circ(\varphi_1^{-1}\circ\widetilde{\varphi})$ 或 $(\varphi^{-1}\circ\varphi_2)\circ(\varphi_2^{-1}\circ\widetilde{\varphi})$
由 A_M 的定义, $\widetilde{\varphi}^{-1}\circ\varphi_1,\widetilde{\varphi}^{-1}\circ\varphi_2$ 均为 $C^{(\infty)}$ 类的.
若对 φ_1, φ_2 成立, 即有 $\varphi_1^{-1}\circ\varphi$, $\varphi_2^{-1}\circ\varphi$ 为局部的 $C^{(\infty)}$ 类微分同胚, 则由复合求导 $\widetilde{\varphi}^{-1}\circ\varphi$ 亦然.

对 φ_1,φ_2, 由 φ 的定义, 只需验证在 x_0 处的局部坐标变换为 $C^{(\infty)}$ 次的.

考虑 $\varphi^{-1}(U\cap U_1)$ (定义域) 中 t_1 附近, 有

$$\varphi_1^{-1}\circ\varphi(t)=\begin{cases} t=(\varphi_1^{-1}\circ\varphi_2)\circ(\varphi_2^{-1}\circ\varphi_1)(t) & t\in]-\infty,t_1] \\ \varphi_1^{-1}\circ\varphi_2\circ\chi(t) & t\in[t_1,+\infty[\end{cases}$$

由 χ 的条件: $\chi^{(i)}(t_1)=(\varphi_2^{-1}\circ\varphi_1)^{(i)}(t_1)$, 复合求导知: $\varphi_1^{-1}\circ\varphi$ 在 t_1 处 $C^{(\infty)}$ 次可微.
反过来, 在 $\varphi_1^{-1}(U\cap U_1)$ 中 t_1 附近, 有:

$$\varphi^{-1}\circ\varphi_1(t)=\begin{cases} t=(\varphi_1^{-1}\circ\varphi_2)\circ(\varphi_2^{-1}\circ\varphi_1)(t) & t\in]-\infty,t_1] \\ \chi^{-1}\circ\varphi_2^{-1}\circ\varphi_1(t) & t\in[t_1,+\infty[\end{cases}$$

由局部上 $\chi\circ\chi^{-1}(t)=t$, $\psi\circ\psi^{-1}(t)=t$ $(\psi=\varphi_2^{-1}\circ\varphi_1)$ 求导并归纳可得: $(\chi^{-1})^{(i)}(t_2)=(\psi^{-1})^{(i)}(t_2)$ $(i\geq 0)$. 于是同前, 可证 $\varphi^{-1}\circ\varphi_1$ 在 t_1 处 $C^{(\infty)}$ 次可微.
对 φ_2 同理可证. 于是断言 (2) 成立. □

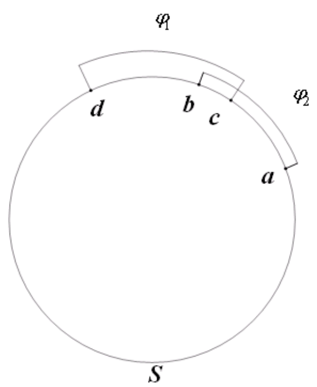


图 2

回到断言 (1) :

由 S^1 的紧性可从 A_M 中取出有限子覆盖: $\{U_i\}_{(1\leq i\leq m)}$ $((U_i,\varphi_i)\in A_M)$, 且 φ_i 均单调增. 可不妨设覆盖 $\{U_i\}$ $(1\leq i\leq m)$ 不可再减少.

取 $(V_1,\psi_1)=(U_1,\varphi_1)$, 存在 U_{i_1} $(i_1\neq 1)$, 设为 U_2, 使 $U_1\cap U_2\neq\emptyset$, 且 U_2 含 U_1 的右端点 (沿逆时钟方向的终点).

由断言 (2) 存在 $\psi_2:\ \mathbb{R}\to V_2$ (对应断言 (2) 中的 $\widehat{ad}$), 与 A_M 中元素等价[3], 且 U_1, $U_2\subset V_2$.

若 V_2 不自交, 则 $(V_2,\psi_2)\in {A_M}$[4]. 则同样存在 U_{i_2}, 设为 U_3, 使 $U_3\cap V_2\neq\emptyset$, U_3 含 V_2 的右端点. 由断言 (2) 可取 $\psi_3:\mathbb{R}\to V_3$, 与 A_M 中元素等价, 且 V_2, $U_3\subset V_3$.

……

如此下去, 经有限步, 必能找到 $\psi:\ \mathbb{R}\to V$, (V,ψ) 与 A_M 中元素等价, 且 V 自交, 这时自然 $S^1\subset V$.

由 ψ 的构造 (A_M中的两个局部图的衔接)知道:

存在 t_1, $t_2\in\mathbb{R}$ $(t_1<t_2)$, 使 $\psi(t_1)=\psi(t_2)=x_1\in S^1$, 且 $(S^1\setminus\{x_1\},\psi|_{]t_1,t_2[})\in A_M$. 同时存在邻域 $O(t_1)$, $O(t_2)$, 若记 $\psi_1=\psi|_{O(t_1)}$, $\psi_2=\psi|_{O(t_2)}$, 则 $\psi_2^{-1}\circ\psi_1:\ O(t_1)\to O(t_2)$ 为 $C^{(\infty)}$ 类微分同胚, $\psi_2^{-1}\circ\psi_1(t_1)=t_2$.

验证知 ψ 已满足断言 (1) 中除 (c) 外的所有条件. (见图3.)

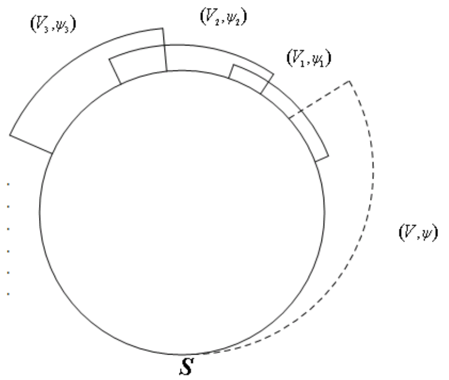


图 3

取 $\chi:\ \mathbb{R}\to\mathbb{R}$ 为 $C^{(\infty)}$ 类微分同胚, 使 $\chi(t_1)=\chi(t_1)$, $\chi(t_2)=t_2$.

易见 $\psi\circ\chi$ 将仍保持 ψ 的已有性质, 只需再找 χ 使 $\varphi=\psi\circ\chi$ 满足断言 (1) 的 (c) .

同 ψ, 设 φ 对应 φ_1, φ_2, $\varphi_1=\psi\circ\chi|_{\chi^{-1}(O(t_1))}$, $\varphi_2=\psi\circ\chi|_{\chi^{-1}(O(t_2))}$. φ 只需满足:

$$((\varphi_2^{-1}\circ\varphi_1)^{(i)}(t_1))=(t_2,1,0,\ldots).$$

[3]注意这里等价的更一般意义, 以后同样如此.

[4]注意 A_M 作为结构图册的性质.

则由 $\varphi_2^{-1}\circ\varphi_1$, $\varphi_1^{-1}\circ\varphi_2$ 局部上互为反函数, 复合求导可得:

$$((\varphi_1^{-1}\circ\varphi_2)^{(i)}(t_2))=(t_1,1,0,\ldots).$$

而 $\varphi_2^{-1}\circ\varphi_1=\chi^{-1}\circ\psi_2^{-1}\circ\psi_1\circ\chi\Leftrightarrow\chi\circ\varphi_2^{-1}\circ\varphi_1=\psi_2^{-1}\circ\psi_1\circ\chi.$ $\qquad(i)$

简记 $\varphi_2^{-1}\circ\varphi_1=h$, $\psi_2^{-1}\circ\psi_1=k$. (i) 左边的 χ 记为 χ_2, 右边的 χ 记为 χ_1. 则 $(i)\Leftrightarrow$ $\chi_2\circ h=k\circ\chi_1$ $\qquad(ii)$

而记

$$h^{(i)}=\frac{d^ih}{dt^i}|_{t=t_1},\ k^{(i)}=\frac{d^ik}{dt^i}|_{t=t_1},\ \chi_1^{(i)}=\frac{d^i\chi_1}{dt^i}|_{t=t_1},\ \chi_2^{(i)}=\frac{d^i\chi_2}{dt^i}|_{t=t_2}.$$

若取 χ 使:

$$(\chi^{(i)}(t_1))=(\chi_1^{(i)})=(t_1,1,0,\ldots).$$

而 h 应满足: $(h^{(i)})=(t_2,1,0,\ldots)$.

故由 (ii) 对 $t=t_1$ 求各阶导数, 归纳得到:

$(\chi_2\circ h)^{(i)}(t_1)=\chi_2^{(i)}\cdot(h^{'})^i=\chi_2^{(i)}$, $(k\circ\chi_1)^{(i)}(t_1)=k^{(i)}\cdot(\chi_1^{'})^i=k^{(i)}(i\geq0)$

$\Rightarrow\chi_2^{(i)}=k^{(i)}$.

即 χ 应满足:

$$(\chi^{(i)}(t_1))=(t_1,1,0,\ldots),\ (\chi^{(i)}(t_2))=(k^{(i)}).$$

注意到 $k=\psi_2^{-1}\circ\psi_1$, 有 $k(t_1)=t_2>t_1$, $k^{'}(t_1)>0$. 于是由引理 2 知这样的 χ 存在.

反过来, 若 χ 满足上述条件, 由 (ii) 求导归纳知道 $h=\varphi_2^{-1}\circ\varphi_1$ 满足断言 (1) 中的 (c), 从而 φ 满足断言 (1), 断言 (1) 成立. $\square$

综上, 所证命题成立.

参考文献

[1] B.A.Zorich 著; 蒋铎等译, 数学分析, 第二卷 (第四版), 高等教育出版社, 北京, 2006.

[2] 荀思, 2009.12, 第04期.

代数整数是单位根的一个刻画

林洁*

1 问题

在代数数论中有一个有趣的小练习:

命题 1 一个代数整数 α 是单位根当且仅当它所有的共轭模为 1.

用初等的语言描述，命题 1 即为:

命题 1′ $p(x)$ 是一个首一整系数多项式, α 是它的根. 若 p 所有的根模都为 1, 则 α 为单位根. 特别地, 若还已知 p 不可约, 则 p 是分圆多项式.

而用矩阵的语言, 该命题又等价于:

命题 1″ r 是一个正整数, $g \in GL_r(\mathbb{Z})$. 若 g 有 r 个不同的复特征根 $\lambda_1, \ldots, \lambda_r$, 且 $|\lambda_1| \le 1, \ldots, |\lambda_r| \le 1$, 则存在正整数 N, 使得 $g^N = 1$. 从而 $\lambda_1^N = \ldots = \lambda_r^N = 1$.

下面我们给出该命题的三种证明, 其中证明 1, 2 需要简单的代数数论知识, 而证明 3 只需要数学分析和高等代数知识.

2 证明

证明 1: (采用命题1的记号)

记 α 在 $\mathbb{Q}$ 上的最小多项式为 $P(x) = x^r + c_1 x^{r-1} + \ldots + c_r$, 并记 $Gal(\mathbb{Q}(\alpha)/\mathbb{Q}) = \sigma_1, \ldots, \sigma_r$, 其中 $\sigma_1(\alpha) = \alpha$. 则由条件, $\forall\ 1 \le i \le r$, $|\sigma_i(\alpha)| = 1$. 从而 $\forall\ n \in \mathbb{N}$, $\forall\ 1 \le i \le r$, 有 $|\sigma_i(\alpha^n)| = 1$.

记 $P_n(x) = \prod_{i=1}^{r}(x - \sigma_i(\alpha^n)) = x^r + c_{1n}x^{r-1} + \ldots + c_{rn}$. 则

$$
\begin{aligned}
|c_{kn}| &= |\sum_{\substack{1 \le i_1 < \ldots \\ < i_k \le r}} \sigma_{i_1}(\alpha^n) \ldots \sigma_{i_k}(\alpha^n)| \\
&\le \sum_{\substack{1 \le i_1 < \ldots \\ < i_k \le r}} |\sigma_{i_1}(\alpha^n) \ldots \sigma_{i_k}(\alpha^n)| = \binom{r}{k},
\end{aligned}
$$

*基数 71

$\forall\ n \in \mathbb{N},\ \forall\ 1 \le k \le r$.

注意到 P_n 是整系数多项式, 故对任意给定的 n, $c_{nk}(1 \le k \le r)$ 的选择都是有限的. 从而 $\{P_n\}_{n=0}^{\infty}$ 是有限集. 进一步地, 由 $P_n(\alpha^n) = 0$ 知 $\{\alpha^n\}_{n=0}^{\infty}$ 也是有限集.

故 $\exists\ n_1 < n_2,\ s.t.\ \alpha^{n_1} = \alpha^{n_2}$, 即 $\alpha^{n_2-n_1} = 1$, α 是单位根.

证明 2: (沿用证明 1 中记号)

假设 α 不是单位根, 则 $\{\alpha^n\}_{n=1}^{\infty}$ 在单位圆周上稠密 (请读者自行证明). 故 $\exists\ n \in \mathbb{N},\ s.t.\ |1-\alpha^n| < \frac{1}{2^r}$.

注意到 $\prod_{i=1}^{r}(1-\sigma_i(\alpha^n)) \neq 0$ (否则 $\exists\ i$, $\sigma_i(\alpha)$ 是单位根, 那么 α 也是单位根, 与假设矛盾), 从而 $\prod_{i=1}^{r}|1-\sigma_i(\alpha^n)| \ge 1$. 但 $\prod_{i=1}^{r}|1-\sigma_i(\alpha^n)| = |1-\alpha^n|\prod_{i=2}^{r}|1-\sigma_i(\alpha^n)| \le |1-\alpha^n|\prod_{i=2}^{r}(1+|\sigma_i(\alpha)|) = 2^{r-1}|1-\alpha^n| < 1/2$, 矛盾. 故假设不成立, 原命题得证.

在继续证明三之前, 我们先回顾一下范数等价定理. 这是数学分析大二上的内容, 这里就不证明了.

定理 2 $\mathbb{R}^n$(或 $\mathbb{C}^n$)中的范数都是等价的. 从而在任何范数下, $\mathbb{R}^n$(或 $\mathbb{C}^n$)的单位闭球总是紧的.

而在下面的情况中, 两个范数等价只需简单的不等式放缩即可证明.

证明 3:(采用命题 $1''$ 中记号)

首先, 由于 $\lambda_1, \dots, \lambda_r$ 互不相同知存在可逆方阵 T, 使得 $T^{-1}gT = diag(\lambda_1, \dots, \lambda_r)$.

我们在 $M_n(\mathbb{C})$ 中定义范数如下: 对 $x \in M_n(\mathbb{C})$, 以 x_{ij} 表示其第 ij 位分量, 定义

$$\|x\| = \|T^{-1}xT\|_{\infty} := sup_{1 \le i,j \le r}|(T^{-1}xT)_{ij}|,$$

则 $\|\ \|$ 是 $M_n(\mathbb{C})$ 中范数 (请读者自行验证).

一方面, $\forall\ n \in \mathbb{N}$, $\|g^n\| = \|T^{-1}g^nT\|_{\infty} = \|diag(\lambda_1^n, \dots, \lambda_r^n)\|_{\infty} \le 1$. 故 $\{g^n\}_{n=0}^{\infty} \subseteq B(0,1)$, 其中 $B(0,1)$ 是 $\|\ \|$ 下的单位闭球, 是一个紧集.

另一方面, $\{g^n\}_{n=0}^{\infty}$ 是 $\|\ \|$ 下的离散集. 事实上, 更一般地, $GL_r(\mathbb{Z})$ 是 $M_r(\mathbb{Z})$ 在 $\|\ \|$ 下的离散集. 这是由于 $GL_r(\mathbb{Z})$ 是 $M_r(\mathbb{Z})$ 在 $\|\ \|_{\infty}$ 下的离散集, 而 $\|\ \|_{\infty}$ 与 $\|\ \|$ 是等价的.

综合两方面, 离散集 $\{g^n\}_{n=0}^{\infty}$ 包含在一个紧集内, 故它是有限集. 即 $\exists\ n_1 < n_2,\ s.t.\ g^{n_1} = g^{n_2}$, 由 g 可逆得到 $g^{n_2-n_1} = 1$. 得证.

参考文献

[1] 冯克勤 著; 代数数论, 科学出版社, 北京, 2000.

[2] 巴黎高师考题, 2006年

纤维丛和Hopf纤维化*†

高挺然‡

切丛是微分流形理论中非常基本的概念。设M为一个n维光滑流形，$\mathscr{F}$是其图册。记TM_x为M上x点处的切空间，则可以考虑

$$TM = \bigcup_{x\in M} TM_x$$

并在其上附加一个由$\mathscr{F}$诱导得到的光滑结构，从而将其看成$2n$维流形。TM上还有一个自然投影

$$\pi\colon TM \longrightarrow M$$

$$\xi \longmapsto x \qquad \xi \in TM_x, x \in M$$

任取$\mathscr{F}$的一个图卡(U,φ)，U为开集。可以想见，由于每个点x上的切空间都是同构的，应有$\pi^{-1}(U) = \bigcup_{x\in U} TM_x \cong U \times TM_{x_0}$，其中$x_0 \in U$是$U$中任意一个固定的点。因此，局部地看，$TM$就像是一个乘积空间。

纤维丛的概念是上述切丛概念的推广。借助于纤维丛，我们可以更好地理解Hopf在1931年作出的一个非常精巧的观察 [1]，现在一般称之为Hopf纤维化(Hopf fibration)。它在$\pi_3(S^2)$的计算中起到重要的作用 [2] [5]，而且具有令人印象深刻的几何直观 [4]。更令人注目的是，它在物理学的许多地方都有着令人意想不到的应用 [8]。

这则笔记首先介绍纤维丛的概念，由此引出Hopf纤维化的构造，最后简单提及Hopf纤维化的几个应用。其中将仅讨论初等的内容，避免使用代数拓扑的工具。对相关及更深入的代数拓扑背景感兴趣的读者将会发现 [5]的价值。

1　纤维丛

1. 启发性的考虑和例子

*根据2009年4月29日学术沙龙报告整理

†感谢毛天一为插图付出的辛勤劳动

‡基数 61

暂时先不考虑在其上装备坐标系统，如果我们希望推广切丛的构造，那么一个纤维丛的几何图象至少应该包含哪些因素？实践表明，我们至少可以要求一个“纤维丛”$\mathscr{B}$含有以下成分：

(i) 一个拓扑空间B，起到TM的作用，称之为**丛空间(bundle space)**或**丛 (bundle)**；

(ii) 一个拓扑空间X，起到M的作用，称之为**底空间(base space)**；

(iii) 一个连续映射（在适当的拓扑下）$\mathrm{p}\colon B \longrightarrow X$，且映满$X$，称之为**投影(projection)**，就像切丛上的自然映射π；

(iv) 一个“模板”Y，称之为**纤维(fibre)**；并记$Y_x := \mathrm{p}^{-1}(x), \forall x \in X$，称之为**$x$上的纤维(fibre over the point x of X)**。就像对任意的$x \in M$，TM_x都同构（装备了适当的拓扑以后甚至可以是同胚）那样，对任意的$x \in X$，要求Y_x同胚于Y。对每个固定的$x \in X$，从Y_x到Y的同胚映射不必唯一，但要求它们至多只相差某个群G中的一个元素，这里群G是作用在Y上的一个从Y到Y的变换群，且G与具体的x的选择无关。我们称群G为此纤维丛的**丛结构群**或**丛群(group of the bundle)**；

(v) $\forall x \in X$，存在x的邻域V以及同胚$\phi\colon V \times Y \longrightarrow \mathrm{p}^{-1}(V)$且满足

$$\mathrm{p}\phi(x', y) = x', \quad \forall x' \in V, y \in Y.$$

上面的(i)-(v)都是通过直观类比得到基本要求。其中第五条要求体现出我们强烈地希望B至少在局部上如同乘积空间一样易于考察。

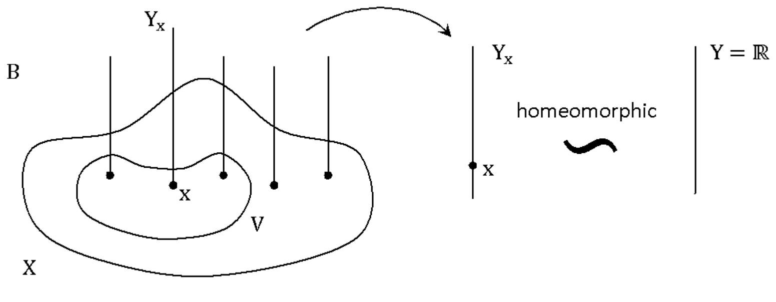


图 1: 取$Y = \mathbb{R}$。在线丛的情形，B就像是一个带有头发的皮球。

例 1.1 (乘积丛) X, Y为两个拓扑空间，$B = X \times Y$，$\mathrm{p}(x, y) = x$, $Y_x = \{x\} \times Y$。Y_x与Y之间的同胚映射由$(x, y) \mapsto y$给出。为满足(v)，可取$V = X$，ϕ为从$V \times Y$到$\mathrm{p}^{-1}(V) = V \times Y$的恒等映射。

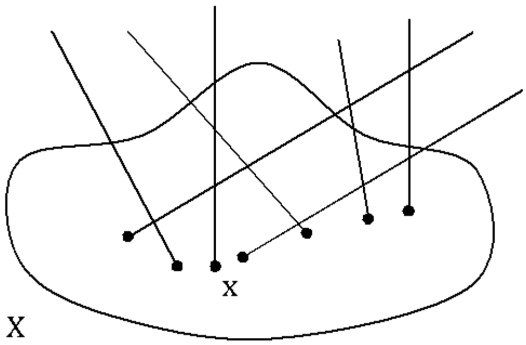
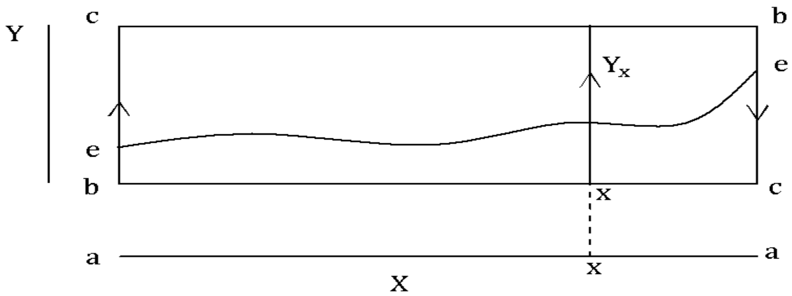
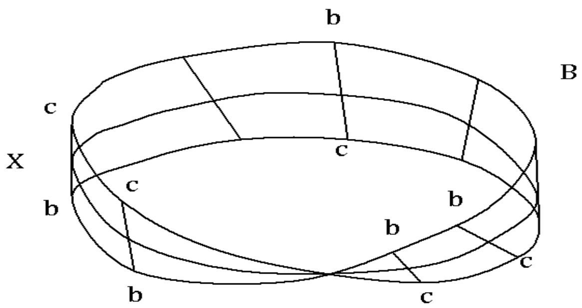


图 2: 若没有第五条要求，B可能非常混乱，并非如我们所愿。

例 1.2 (Möbius带) 取底空间X为一条线段L粘合其两端而得到的圆周。纤维Y取为一条线段。B则由$L\times Y$将两端的线段反向粘合得到。

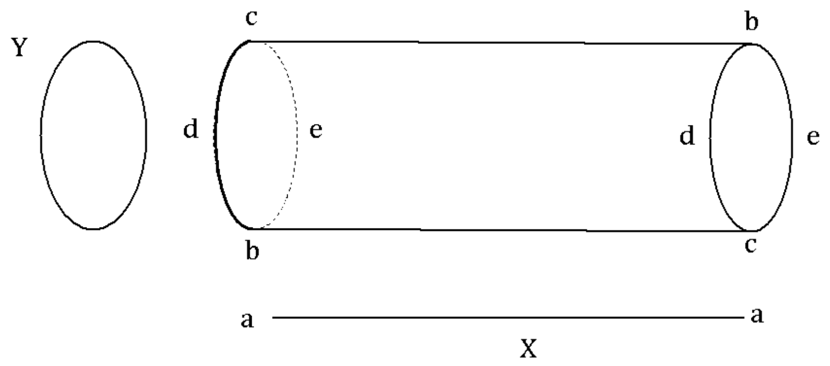




投影p: $B\to X$可以看成自然投影π: $L\times Y\to L$在上述粘合下诱导的映射，直观上将其理解为将Möbius带上各点投影至腰圆上是非常方便的。可取$V=X$，并将ϕ按与p类似的方式理解。从B上的一个**截面(cross section)**是指一个连续映射$f\colon X\to B$满足p$f(x)=x,\forall x\in X$。上图中连结矩形左端e点和矩形右端e点的曲线即是Möbius带上的一个截面的例子。注意在前述的粘合下，两端的e点是

重合的。

例 1.3 (Klein瓶) 将例2中的纤维Y替换为一个圆周，并将$L \times Y$两端的圆周绕一条直径$\overline{de}$反射后粘合得到B。这种粘合方式下，B是一个Klein瓶。



例 1.4 (陪集空间) 取B为一个李群，X为一个流形，而且B作为一个变换群传递地作用在X上。为定义投影p，任意固定X上一点x_0，并定义p: $B \to X$为$b \mapsto b(x_0)$。纤维Y可取为B的固定x_0的子群；对每个$x \in X$，由于B在X上的作用是传递的，故存在$\beta_x \in B$使得$\beta_x(x_0) = x$。而Y_x就是B中包含β_x的陪集$\beta_x Y$。容易验证此Y_x的定义不依赖于β_x的特定选取。进一步注意到，每一个$b \in Y_x$都给出了从Y到Y_x的同胚映射：$y \mapsto by, y \in Y$，而且任意两个这样的同胚映射$b: Y \to Y_x, b': Y \to Y_x$之间只相差一个$Y$上的左乘$b'b^{-1}$的平移。易见$b'b^{-1} \in Y$。因此在这个例子里丛结构群就是$Y$,而且以左平移的方式作用在$Y$上。在这样的一个丛$B$上找一个截面即是构造一个$B$中的单传递连续变换群。

通过以上几个例子，我们可以感觉到纤维丛和乘积空间有很大的相似性，纤维丛的概念是乘积空间概念的推广。事实上，对空间X, Y以及映射$f: X \to Y$的研究等同于对乘积空间$X \times Y$、两个自然投影（$\pi_1: X \times Y \to X$和$\pi_2: X \times Y \to Y$）以及映射f的图象的研究。纤维丛的概念将$X \times Y$推广到丛空间B，将自然投影π_2替换为对每个固定的$x \in X$指定的一族从Y_x到Y的同胚，其中任意两个同胚只相差一个作用在Y上的变换群G中的元素；$f: X \to Y$的图象则被推广到丛上的截面。

2. 坐标丛和纤维丛的精确定义

反思我们在前面做过的事情。在例4中，丛结构群G具有一个拓扑结构！这是我们非常希望任何一个“纤维丛”都能具有的。因此在上一节中我们并不急于给出纤维丛的精确定义，虽然所有的要素都已经齐备。本节首先借助坐标在丛群G上给出一个拓扑，再通过在坐标丛之间引入一个等价关系来摆脱坐标，最终达到我们的目的。引入坐标也将给我们日后的工作提供很大的便利。

一个坐标丛$\mathscr{B}$是以下要素的集合：

(1) 一个拓扑空间B，称为**丛空间(bundle space)**；

(2) 一个拓扑空间X，称为**底空间(base space)**；

(3) 一个连续满映射$\mathrm{p}\colon B \to X$，称为**投影(projection)**；

(4) 一个拓扑空间Y，称为**纤维(fibre)**；

(5) 一个Y上的有效(effective)拓扑变换群G，称为**丛结构群**或**丛群(group of the bundle)**（G**有效**是指对任意的$g \in G$，$gy = y, \forall y \in G \leftrightarrow g = e$）；

(6) X的一个开覆盖$\{V_j\}_{j\in J}$，其中每一个V_j称为一个**坐标邻域(coordinate neighborhood)**；

(7) 对任意的$j \in J$，存在一个同胚$\phi_j\colon V_j \times Y \longrightarrow \mathrm{p}^{-1}(V_j)$，称为一个**坐标函数(coordinate function)**（或一个**局部平凡化(local trivialization)**）满足以下条件(8)-(10)；

(8) $\mathrm{p}\phi_j(x, y) = x, \quad \forall x \in V_j, y \in Y$；

(9) 若定义映射$\phi_{j,x}\colon Y \longrightarrow \mathrm{p}^{-1}(x)$为$\phi_{j,x}(y) = \phi_j(x, y)$，则对任意$i, j \in J$，以及任意$x \in V_i \cap V_j$，同胚$\phi_{j,x}^{-1}\phi_{i,x}\colon Y \longrightarrow Y$与$G$中一个元素在$Y$上的作用相同（在我们的定义下这个元素必定唯一，因为G是有效的）；

(10) $\forall i, j \in J$，由$g_{ji}(x) = \phi_{j,x}^{-1}\phi_{i,x}$定义的映射$g_{ji}\colon V_i \cap V_j \longrightarrow G$是连续的。

不难发现，若没有(5)，(9)，(10)则这个纤维丛的概念和§1中的定义是一样的。条件(9)将群G与纤维丛的结构本质地联系了起来，而条件(10)则刻画了群G的拓扑。

(10)中定义的映射g_{ji}称为纤维丛$\mathscr{B}$的坐标变换。这个定义的直接推论是，对任意的$i, j, k \in J$

(11) $$g_{kj}(x)g_{ji}(x) = g_{ki}(x) \qquad \forall x \in V_i \cap V_j \cap V_k$$

特别地，若$i = j = k$，则

(12) $$g_{ii}(x) = \text{群G的单位元} \qquad \forall x \in V_i$$

若在(11)中令$i=k$，并运用(12)，可得

$$(13)\qquad g_{jk}(x)=[g_{kj}(x)]^{-1}\qquad \forall x\in V_j\cap V_k$$

为方便计，定义映射$\mathrm{p}_j\colon \mathrm{p}^{-1}(V_j)\to Y$为$\mathrm{p}_j(b)=\phi_{j,x}^{-1}(b)$，$x=\mathrm{p}(b)$，则$\mathrm{p}_j$满足恒等式

$$\mathrm{p}_j\phi_j(x,y)=y,\quad \phi_j(\mathrm{p}(b),\mathrm{p}_j(b))=b,\quad g_{ji}(\mathrm{p}(b))\mathrm{p}_i(b)=\mathrm{p}_j(b)\quad \forall p(b)=x\in V_i\cap V_j$$

我们称两个坐标丛$\mathscr{B}$和$\mathscr{B}'$在严格意义上等价，若它们有相同的丛空间、底空间、投影、纤维、丛结构群，而且它们的坐标函数$\{\phi_j\},\{\phi'_j\}$满足下述条件：$\forall x\in V_j\cap V'_k$，映射$\overline{g}_{kj}(x)=\phi'^{-1}_{k,x}\phi_{j,x}$与$G$中一个元素在$Y$上的作用相同，而且由此得到的映射$\overline{g}_{kj}\colon V_j\cap V'_k\longrightarrow G$是连续的；简言之，两组坐标函数之并仍是一组坐标函数。

利用G作为一个拓扑群的性质，容易知道上面定义的“在严格意义上等价”确实是一个等价关系。利用这个等价关系，我们可以给出如下的定义：

定义 1 *一个**纤维丛**是指坐标丛在上述等价关系下的一个等价类。*

实用中一类非常重要的纤维丛是**主丛(principal bundle)**，其纤维Y即是丛结构群G自身，且G在Y上的作用是左平移。

作为本节的结束，我们指出有如下重要的基本定理成立：

定理 1 (存在性定理) *若G是Y的一个有效拓扑变换群，$\{V_j\}_{j\in J}$和$\{g_{ij}\}_{i,j\in J}$是X的一族坐标变换，则存在一个纤维丛$\mathscr{B}$，以X为底空间，以Y为纤维，以G为丛结构群，以$\{g_{ij}\}_{i,j\in J}$为坐标变换。而且任何两个这样的纤维丛都是等价的。*

证明略去，可参阅文献 [3]§3.2。

2 Hopf纤维化

1. **Hopf映射$S^3\longrightarrow S^2$**

   将球面S^3和S^2按如下方式参数化：

$$S^3=\left\{(x^1,x^2,x^3,x^4)\in\mathbb{R}^4\,\middle|\,(x^1)^2+(x^2)^2+(x^3)^2+(x^4)^2=1\right\}$$
$$S^2=\left\{(\xi^1,\xi^2,\xi^3)\in\mathbb{R}^3\,\middle|\,(\xi^1)^2+(\xi^2)^2+(\xi^3)^2=1\right\}$$

Hopf [1]犀利地指出，S^3是S^2上以U(1)为纤维的一个丛。U(1)即是S^1，因此S^3即是“the fibering of spheres by spheres”的一个最简单的非平凡的例子。（将S^1看作以S^1为底空间、以S^0为纤维的丛，我们就看到了一个平凡的例子。）

为$\mathbb{R}^4$引入复坐标$z^0 = x^1 + \mathrm{i}x^2, z^1 = x^3 + \mathrm{i}x^4$，则$S^3$可以写成

$$S^3 = \left\{(z^0, z^1) \in \mathbb{C}^2 \middle| |z^0|^2 + |z^1|^2 = 1\right\}$$

定义Hopf映射$h\colon S^3 \longrightarrow S^2$如下：

$$\begin{aligned}
\xi^1 &= 2(x^1x^3 + x^2x^4) = z^0\overline{z^1} + \overline{z^0}z^1 = 2\,\mathrm{Re}\,z^0\overline{z^1} \\
\xi^2 &= 2(x^2x^3 - x^1x^4) = -i(z^0\overline{z^1} - \overline{z^0}z^1) = 2\,\mathrm{Im}\,z^0\overline{z^1} \\
\xi^3 &= (x^1)^2 + (x^2)^2 - (x^3)^2 - (x^4)^2 = |z^0|^2 - |z^1|^2
\end{aligned}$$

容易验证h确实将S^3映成S^2：

$$(\xi^1)^2 + (\xi^2)^2 + (\xi^3)^2 = [(x^1)^2 + (x^2)^2 + (x^3)^2 + (x^4)^2]^2 = 1$$

而且容易验证h是满射。这个映射h就是我们要找的从丛空间S^3到底空间S^2的投影。

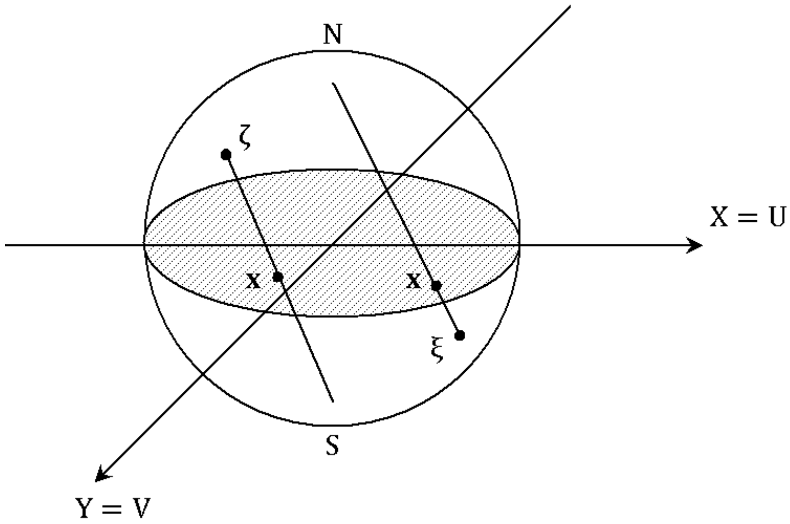
为构造坐标函数，首先考虑S^2的球极投影。取U_N为S^2的上半球面与赤道的并，取U_S为S^2的下半球面与赤道的并，则$U_N \cup U_S = S^2$，$U_N \cap U_S$为赤道。在赤道平面上装备复坐标，则$\forall \xi \in U_S$，可以考虑北极与ξ的连线与赤道平面的交点（我们考虑这种球极投影，因为我们希望所得到的投影坐标在单位圆内），其复坐标为

$$Z = \frac{\xi^1 + \mathrm{i}\xi^2}{1 - \xi^3} = \frac{z^0\overline{z^1}}{|z^1|^2} = \frac{z^0}{z^1}$$

注意到对任意$\lambda \in \mathrm{U}(1) = S^1$，$Z$在旋转$(z^0, z^1) \mapsto (\lambda z^0, \lambda z^1)$上不变，而且$\frac{z^0}{z^1}$在此球极投影下的原像都是形如$(\lambda z^0, \lambda z^1)$的点，因此纤维即是$S^1$！又由于$|\lambda| = 1, (\lambda z^0, \lambda z^1) \in S^3$，可见每条纤维都是$S^3$上的大圆。类似地，$\forall \zeta \in U_N$

$$Z = \frac{\zeta^1 - \mathrm{i}\zeta^2}{1 - \zeta^3} = \frac{\overline{z^0}z^1}{|z^0|^2} = \frac{z^1}{z^0}$$

在赤道$U_N \cap U_S$上，坐标变换为$z \mapsto \dfrac{1}{z}$。



借助球极投影，下面可以给出S^3作为S^2上纤维丛的坐标函数：

$$\begin{aligned}\phi_S^{-1}: h^{-1}(U_S) &\longrightarrow U_S \times S^1 \\ (z^0, z^1) &\longmapsto (\frac{z^0}{z^1}, \frac{z^1}{|z^1|}) \\ \phi_N^{-1}: h^{-1}(U_N) &\longrightarrow U_N \times S^1 \\ (z^0, z^1) &\longmapsto (\frac{z^1}{z^0}, \frac{z^0}{|z^0|})\end{aligned}$$

赤道上$\xi^3 = 0$，$|z^0| = |z^1| = \dfrac{1}{\sqrt{2}}$，对每个固定的$\dfrac{z^0}{z^1} \in U_S \cap U_N$，$\phi_N$与$\phi_S$相差一个坐标变换

$$t_{NS}(\xi) = \frac{\sqrt{2}z^0}{\sqrt{2}z^1} = \frac{z^0}{z^1} \in \mathrm{U}(1) = S^1$$

因此丛结构群即是S^1。我们可以验证以上构造满足坐标丛定义的所有要求，而且纤维和丛结构群都是S^1。这即是说S^3不但是S^2上以S^1为纤维的一个纤维丛，而且是一个主丛!

2. **Hopf纤维化$S^3 \longrightarrow S^2$的几何图象**

需要指出的是，S^3作为S^2上的S^1丛这个事实远不是平凡的。换言之，S^3并不同胚于$S^2 \times S^1$，这只要考察它们各自的高阶同伦群就可以得知。这个非平凡性还体现在如下的有趣事实中：任意两条纤维S^1都不是两个不相干的圆周，它们在球极投影下是相互套在一起的(linked)！它们构成了最简单的一种link，

一般称之为Hopf link。为了说明这一点，我们需要考察将S^3映入$\mathbb{R}^3$的球极投影。

记其中一个投影函数为

$$
\begin{aligned}
s : S^3 \backslash (1,0,0,0) &\longrightarrow \mathbb{R}^3 \\
(w,x,y,z) &\longmapsto (\frac{x}{1-w}, \frac{y}{1-w}, \frac{z}{1-w})
\end{aligned}
$$

容易验证，s将S^3中不经过$(1,0,0,0)$的圆周映为$\mathbb{R}^3$中的圆周，而将S^3中经过$(1,0,0,0)$的圆周映为$\mathbb{R}^3$中的直线。我们的观察分为两步：

(1) 直接计算可以发现，纤维$h^{-1}((0,0,1)) = \{(z^0,z^1) \in \mathbb{C}^2 \big| |z^0| = 1\}$被$s$映为$\mathbb{R}^3$中的$x$轴，而纤维$h^{-1}((0,0,-1)) = \{(z^0,z^1) \in \mathbb{C}^2 \big| |z^1| = 1\}$被$s$映为$\mathbb{R}^3$中的$yz$平面上的单位圆。

除此之外，对任意的点$P = (p_1,p_2,p_3) \in S^2, P \neq (0,0,\pm 1)$，考虑$s \circ h^{-1}(P)$与$yz$平面的交点，即要解方程组

$$
\begin{cases}
2(x^1x^3 + x^2x^4) = p_1 \\
2(x^2x^3 - x^1x^4) = -p_2 \\
(x^1)^2 + (x^2)^2 - (x^3)^2 - (x^4)^2 = p_3 \\
(x^1)^2 + (x^2)^2 + (x^3)^2 + (x^4)^2 = 1 \\
(p^1)^2 + (p^2)^2 + (p^3)^2 = 1 \\
x^2 = 0
\end{cases}
\Leftrightarrow
\begin{cases}
x^1x^3 = \dfrac{p_1}{2} \\
x^1x^4 = \dfrac{p_2}{2} \\
(x^1)^2 = (x^3)^2 + (x^4)^2 = \dfrac{1+p_3}{2}
\end{cases}
$$

一般地，只要$(p_1,p_2,p_3) \neq (0,0,\pm 1)$，交点$(\frac{x^3}{1-x^1}, \frac{x^4}{1-x^1})$可以有而且仅有两个不同的位置，分别对应于取正值和负值的x^1。注意到

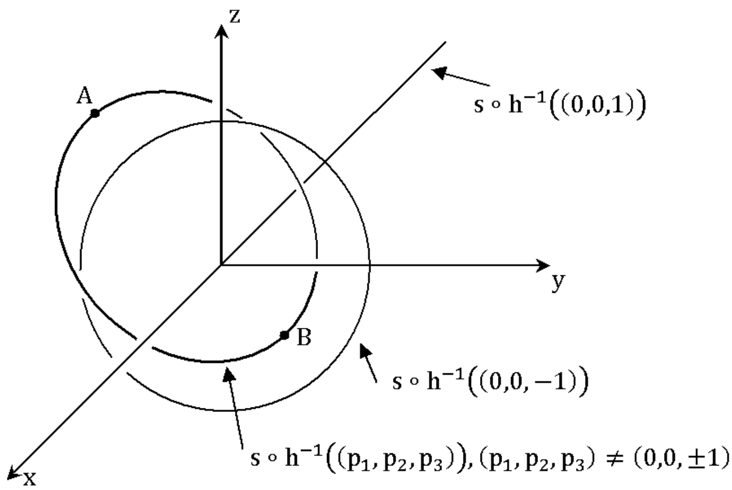
$$
\Big(\frac{x^3}{1-x^1}\Big)^2 + \Big(\frac{x^4}{1-x^1}\Big)^2 = \frac{1-(x^1)^2}{(1-x^1)^2} = \frac{1+x^1}{1-x^1}, \quad x^1 = \pm\sqrt{\frac{1+p_3}{2}} \in (-1,1)
$$

故若取 $x^1 = \sqrt{\dfrac{1+p_3}{2}}$, $x^1 \in (0,1)$, $\dfrac{1+x^1}{1-x^1} > 1$，则交点在单位圆外；

若取 $x^1 = -\sqrt{\dfrac{1+p_3}{2}}$, $x^1 \in (-1,0)$, $\dfrac{1+x^1}{1-x^1} < 1$，则交点在单位圆内。

由此即可得知，$s \circ h^{-1}(P)$和$s \circ h^{-1}((0,0,-1))$是套着的，而且x轴必定不在$s \circ h^{-1}(P)$所在的平面内，因为S^2中不同点上的纤维是彼此不交的。

(2) 对任意两根纤维$h^{-1}(P_1)$和$h^{-1}(P_2)$（假设$(0,0,\pm 1)$不在上述两根纤维的任何一根上），我们要说明它们在球极投影下的像是套着的。为此首先考虑一个S^3上的旋转ρ，它将$h^{-1}(P_1)$转到S^3上一个新的位置，使得$(0,0,-1)$位

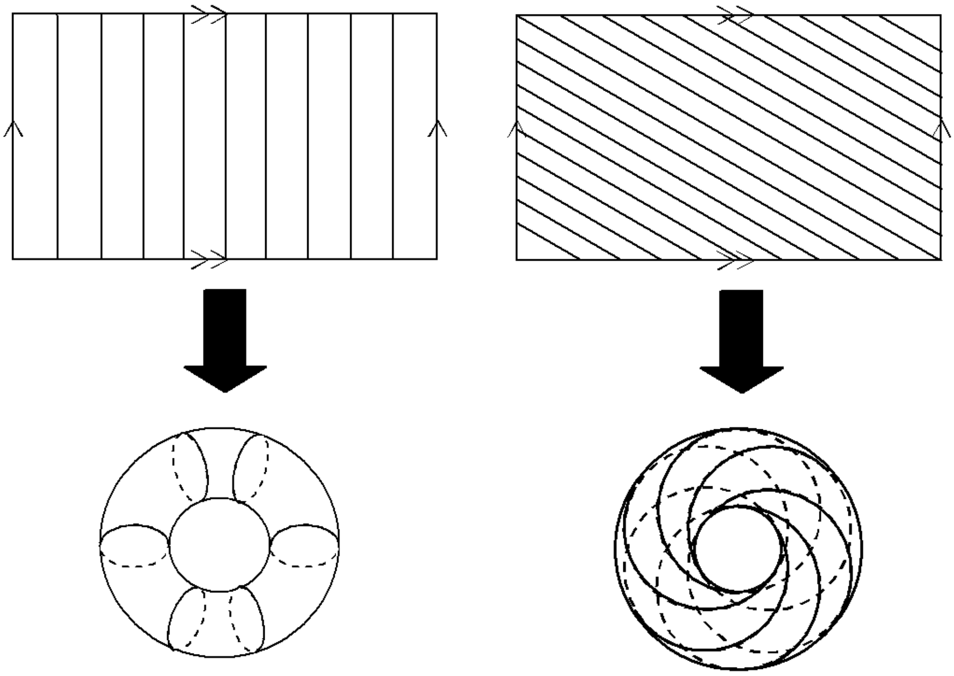


于$\rho \circ h^{-1}(P_1)$上。接着考虑(1)中的球极投影，由(1)容易知道$\rho \circ h^{-1}(P_1)$与$\rho \circ h^{-1}(P_2)$在此球极投影下是套着的。但现在这两个圆环是$s \circ h^{-1}(P_1)$和$s \circ h^{-1}(P_2)$在映射$s \circ \rho \circ s^{-1}$下的像，由于旋转ρ是同痕于恒等映射的，$s \circ \rho \circ s^{-1}$必然也是同痕于恒等映射的。由此可以断言，$s \circ h^{-1}(P_1)$和$s \circ h^{-1}(P_2)$也是套着的，因为一个Hopf link对应的Jones多项式是$\sqrt{t}(1+t^2)$，两个不套着的圆环对应的Jones多项式则是$-(\frac{1}{\sqrt{t}} + \sqrt{t})$，而且我们知道Jones多项式具有同痕不变性(isotopic invariance)（参见 [6]）。

统合(1)(2)，我们注意到，借助从S^3到$\mathbb{R}^3$的球极投影，我们将$\mathbb{R}^3$分解成了一条直线和许多圆环的并，而且非常不平凡地，其中任意两个圆环都是套着的，同时直线穿过所有圆环的内部。我们也可以将这种分解视为$\mathbb{R}^3$的一个“纤维化”(fibration)。

特别有趣的一个情形是，若我们考察S^2上的一个纬圆C（不经过极点，落在某个球极投影s的定义域里），容易发现$h^{-1}(C)$同胚于$S^1 \times S^1$，从而$s \circ h^{-1}(C)$是$\mathbb{R}^3$的一个环面，它被一族圆环$\left\{s \circ h^{-1}(P) \middle| P \in C \subset S^2\right\}$纤维化，而且其中任意两个圆环都是套住的！这样的一族圆环历史上被称为Villarceau circles，它是S^1以非平凡的方式纤维化T^2的著名例子。

作为一个重要的注记，在本节的最后我们指出，不但S^3上的任意两根纤维在球极投影下的像是套在一起的，他们本身在S^3上也是套在一起的。关于这一点，可以参阅 [5]§17, pp.227-239。（h的Hopf不变量为+1，即是说$\forall P_1, P_2 \in S^2$，



(a) T^2的S^1纤维化：平凡的方式　　(b) T^2的S^1纤维化：Villarceau circles

$h^{-1}(P_1)$和$h^{-1}(P_2)$在S^3中是套在一起的。）

3. **其他Hopf映射**

我们可以换一种方式来考虑Hopf映射$S^3 \to S^2$，从而得到它的一些推广。将S^3看作一维复球面：

$$S^3 \cong S^1_{\mathbb{C}} = \left\{(z^0, z^1) \in \mathbb{C}^2 \middle| \, |z^0|^2 + |z^1|^2 = 1\right\}$$

将S^2看作一维复射影空间：

$$S^2 \cong \mathbb{CP}^1 = \left\{[(z^0, z^1)] \middle| \, (z^0, z^1) \in \mathbb{C}^2, \ [(z^0, z^1)] = \{\lambda(z^0, z^1) | \, \lambda \in \mathbb{C} \backslash \{0\}\}\right\}$$

考虑如下映射：

$$\begin{aligned} h\colon S^1_{\mathbb{C}} &\longrightarrow \mathbb{CP}^1 \\ (z^0, z^1) &\longmapsto [(z^0, z^1)] \end{aligned}$$

在此映射下，$S^1_{\mathbb{C}}$上所有形如$\lambda(z^0, z^1), |\lambda| = 1$的点都被映为$\mathbb{CP}^1$中相同的点，而且$h^{-1}([(z^0, z^1)])$必定由所有形如$\lambda(z^0, z^1), |\lambda| = 1$的点构成。容易验证，这个映射$h$确实和之前定义的Hopf映射$S^3 \to S^2$相一致。

推广的基本想法是将数域$\mathbb{C}$换为其他代数。仔细检查Hopf映射$S^3 \to S^2$的构造，可以注意到我们并没有用到$\mathbb{C}$的全部性质，而仅要求这个代数满足：(1)是一个可除代数(division algebra)；(2)其上的范数是可乘的($\forall z_1, z_2 \in \mathbb{C}, |z_1 z_2| = |z_1||z_2|$)。注意(2)可以保证这个代数是一个无零因子环。由于$\mathbb{C}$可看作是一个实二维可除代数，因此我们推广的方向就成为：寻找所有装备了可乘范数的实可除代数。注意这里的代数不但未必是交换的，甚至可以是非结合的。

1878年，Frobenius [9]证明了所有的结合实可除代数只能是实数$\mathbb{R}$，复数$\mathbb{C}$和四元数$\mathbb{H}$；1958年，Bott和Milnor [10]证明了所有的有限维实可除代数只能是$\mathbb{R}, \mathbb{C}, \mathbb{H}$和八元数$\mathbb{O}$。八元数代数$\mathbb{O}$又称为Cayley numbers，不但不是域、不交换，甚至不是结合的。若再要求范数可乘，则不可能有满足要求的无限维代数，Hurwitz定理 [11]指出$\mathbb{R}, \mathbb{C}, \mathbb{H}$和$\mathbb{O}$就是我们所有可能的选择。这四种情形下相应的Hopf映射分别为：

$$\begin{aligned}
\mathbb{R}: &\quad S^0 \hookrightarrow S^1 \xrightarrow{h} S^1 \\
\mathbb{C}: &\quad S^1 \hookrightarrow S^3 \xrightarrow{h} S^2 \\
\mathbb{H}: &\quad S^3 \hookrightarrow S^7 \xrightarrow{h} S^4 \\
\mathbb{O}: &\quad S^7 \hookrightarrow S^{15} \xrightarrow{h} S^8
\end{aligned}$$

这就是我们所可能拥有的所有Hopf纤维化。对$\mathbb{R}$，$\mathbb{C}$和$\mathbb{H}$的情形可以直接套用上面对Hopf映射$S^3 \to S^2$的等价描述，只需要选取不同的代数即可，而且得到的纤维丛都是主丛。$\mathbb{O}$的情形有一些小的困难，因为它是非结合的，故$\mathbb{C}\mathbb{O}^1$未必能定义好；而且S^7不是一个群，从而得到的纤维丛必不是主丛。但用另外的方式仍能看出Hopf映射$S^{15} \to S^8$上的纤维丛结构，详细的讨论可参见[3]§20, pp.105-110。

值得一提的是，上述四个代数在物理中都有广泛而深刻的应用。

3 Hopf纤维化的各种应用

可以想见，既然Hopf纤维化研究的是球面这样基本的几何对象的拓扑性质，它应能在许多场合下出现。

1. **力学中的两个例子 [13]**

   (1) **1-1共振(The One-to-one Resonance)**

考虑两个谐振子的共振系统时，人们常常遇到如下形式的Hamilton函数：

$$H = \frac{1}{2}(q_1^2 + p_1^2) + \frac{\lambda}{2}(q_2^2 + p_2^2) + \text{高阶项}$$

若$\lambda = 1$，则此Hamilton函数的二次型部分描述的即是被称为“1-1共振”的力学系统。考虑临界情形

$$H_0 = \frac{1}{2}(q_1^2 + p_1^2 + q_2^2 + p_2^2)$$

等能量面$H_0 \equiv \mathrm{const}$即是$S^3 \subset \mathbb{R}^4$。引入复坐标$z_1 = q_1 + \mathrm{i}p_1, z_2 = q_2 + \mathrm{i}p_2$，则$H_0 = \frac{1}{2}(|z_1|^2 + |z_2|^2)$。易见$H_0$在SU(2)的作用下是左不变的，相应的守恒量为

$$W_1 = 2(q_1q_2 + p_1p_2)$$
$$W_2 = 2(q_2p_1 - q_1p_2)$$
$$W_3 = q_1^2 + p_1^2 - q_2^2 - p_2^2$$

且它们满足$4H_0^2 = W_1^2 + W_2^2 + W_3^2$。因此上述映射$j\colon (q_1, p_1, q_2, p_2) \mapsto (W_1, W_2, W_3)$是一个从$S^3$到$S^2$的连续映射，不难发现这就是Hopf映射$S^3 \to S^2$。因此，在相空间中1-1共振系统的轨道$(q_1, p_1, q_2, p_2)$就是$S^3$上的大圆。（这个为人所熟知的经典例子在很多地方都被提及，例如 [7], pp.24）

(2) 刚体的运动

刚体的构形空间是$SO(3) \cong S^3/\mathbb{Z}_2$。若固定一个正交矩阵$A$，将沿固定在刚体上的坐标系的某条坐标轴的单位向量$\vec{k}$映到$A\vec{k}$，则我们构造了一个从SO(3)到$S^2$的投影（本质上是一个动量映射的限制，参阅 [13] §1.3, §1.10）。将此投影与从$SU(2) \cong S^3$到$SO(3) \cong S^3/\mathbb{Z}_2$的商映射复合，我们将再次见到Hopf映射$S^3 \to S^2$。 [13] §1.10中提供了大量相关介绍和历史注记。

2. **磁单极子(Magnetic Monopole)的势场 [14] [8]**

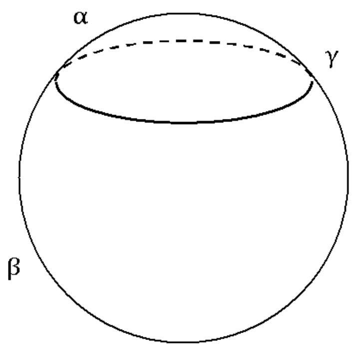
考虑一个强度为$g \neq 0$的磁单极子。为描述一个在它所产生的磁场中运动的电子的波函数，我们需要找该磁场的一个矢量势$\vec{A}$。1931年，Dirac [15]描述了这样的一个势$\vec{A}$，但它不是光滑的，其奇点排列在某些线上。

一个这样的光滑的矢量势是不存在的，为此只需考虑以这个磁单极子为中心，R为半径的一个球面，在其上取一条闭合曲线γ，γ将球面分成两个区域α和β。

利用Stokes定理可以计算出通过球冠α和β的磁通量：

$$\Omega_\alpha = \oint_\gamma A_\mu \mathrm{d}x^\mu$$

$$\Omega_\beta = \oint_\gamma A_\mu \mathrm{d}x^\mu$$



因此穿过整个球面的总磁通量为$\Omega_\alpha - \Omega_\beta = 0$，与由Gauss定理给出的总磁通量应为$4\pi g \neq 0$矛盾。

然而，一个磁单极子激发的磁场应是没有奇点的，因此$\overrightarrow{A}$的这些奇点仅仅是数学上而不是真实物理上的困难。克服这一困难的方法是用两张图卡覆盖上面所说的球面，并分别在它们上面定义磁矢势，就像我们经常在流形上做的那样。记R_a为半径为R的球面挖去南极点的部分，R_b为半径为R的球面挖去北极点的部分，它们组成球面的一个开覆盖。在R_a上定义矢势$(\overrightarrow{A})^a$：

$$(A_r)^a = (A_\theta)^a = 0, \quad (A_\phi)^a = \frac{g}{r\sin\theta}(1-\cos\theta)$$

在R_b上定义矢势$(\overrightarrow{A})^b$：

$$(A_r)^b = (A_\theta)^b = 0, \quad (A_\phi)^b = \frac{-g}{r\sin\theta}(1+\cos\theta)$$

它们在各自的定义域上都是没有奇点的。不难验证它们都给出磁单极子g激发的磁场，故在$R_a \cap R_b$上它们相差一个梯度项。直接的计算表明

$$(A_\mu)^a - (A_\mu)^b = \partial_\mu \alpha, \quad \alpha = 2g\phi \quad (\text{其中}\phi\text{是方位角})$$

于是磁单极子的磁场中电子的Schrödinger方程为

$$\Big[\frac{1}{2m}(p-eA_a)^2 + V\Big]\psi_a = E\psi_a \quad \text{in} \quad R_a$$

$$\Big[\frac{1}{2m}(p-eA_b)^2 + V\Big]\psi_b = E\psi_a \quad \text{in} \quad R_a$$

其中ψ_a,ψ_b就是电子分别在R_a和R_b中的波函数。由于两个方程中的矢量势相差一个梯度，由规范场的理论可知ψ_a和ψ_b相差一个相位因子的变换

$$\psi_a = S\psi_b, \quad S = \mathrm{e}^{\mathrm{i}e\alpha}$$

或写为

$$\psi_a = \mathrm{e}^{2\mathrm{i}q\phi}\psi_b, \quad q = eg$$

考虑球面上的赤道，它落在$R_a \cap R_b$中，由ψ_a在R_a中的单值性以及ψ_b在R_b中的单值性可知，绕赤道一周后回到出发点的波函数ψ_a,ψ_b应和它们出发前相同，这就导出了著名的Dirac量子化条件

$$2q = 2eg = n, \quad n \in \mathbb{Z}$$

这里的波函数ψ_a,ψ_b已经不是通常意义上的S^2上的函数，而是S^2上的截面，它定义在S^2上的一个非平凡的纤维丛上。这是一个怎样的纤维丛呢？这就由上面的Dirac量子化条件中的整数n决定。整数n（又称为**拓扑量子数(topological quantum number)**）是R_a和R_b两个图卡之间的坐标变换映射沿赤道的**环绕数(winding number)**，在相差一个同构的意义下唯一地刻画了S^2上的一个复线丛，或附加于其上单位向量的一个球丛（参阅 [5], pp.297-302）。特别地，当$n = 1$时，我们得到的就是S^2上的一个S^1丛，也就是Hopf纤维化$S^3 \to S^2$中给出的纤维丛结构。$n = -1$的丛被称为anti-Hopf的。

3. **两态量子系统的Bloch球面(Bloch Sphere)表示 [16] [8]**

一个两态量子系统可以用一个二维Hilbert空间$\mathbb{C}^2$描述，其上装备标准的Hermite内积。一个量子比特(qubit)可以写为

$$|\psi\rangle = \alpha|0\rangle + \beta|1\rangle, \quad \alpha,\beta \in \mathbb{C}, \quad |\alpha|^2 + |\beta|^2 = 1$$

系统的态由密度矩阵（迹为1的Hermite正定2×2矩阵）给出，而系统的观测量则由Hermite矩阵给出。用矩阵A表示的观测量在密度矩阵为ρ的态上的期望由实数$\mathrm{Tr}\,A\rho$给出。当ρ的秩为1时，它被称为一个纯态(pure state)，可以用向量$z \in \mathbb{C}^2$表示为$\rho = \dfrac{zz^\dagger}{z^\dagger z}$，其中$z^\dagger w = \langle z, w\rangle = \overline{z_1}w_1 + \overline{z_2}w_2$；而观测量$A$在其上的期望为$\dfrac{z^\dagger A z}{z^\dagger z}$。注意到相差一个复因子的向量$z$决定同一个纯态。

现在考虑以Pauli矩阵$\sigma_1,\sigma_2,\sigma_3$为观测量的期望，其中

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \qquad \sigma_2 = \begin{pmatrix} 0 & -\mathrm{i} \\ \mathrm{i} & 0 \end{pmatrix} \qquad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

将它们写成一个矩阵向量$\overrightarrow{\sigma} = (\sigma_1, \sigma_2, \sigma_3)$，它在态$\rho$上的期望可以写为$\overrightarrow{R} = \mathrm{Tr}\,\rho\overrightarrow{\sigma} \in \mathbb{R}^3$。容易验证这个关系可以逆过来写成$\rho = \frac{1}{2}(1 + \overrightarrow{R} \cdot \overrightarrow{\sigma})$，其中$\rho$的正定性蕴含$\overrightarrow{R}^2 \leq 1$，且仅当$\rho$是纯态时等号成立。这即是说，$\overrightarrow{R}(z) := \dfrac{z^\dagger \overrightarrow{\sigma} z}{z^\dagger z}$满足$\overrightarrow{R}^2(z) = 1$，亦即密度矩阵$\rho$可以用$\mathbb{R}^3$中单位闭球表示，而纯态都相应于$\mathbb{R}^3$中的单位向量。这个单位闭球面历史上称之为Bloch sphere（这个单位闭球则相应地可以被称为Bloch ball），因为F. Bloch曾用它解释了磁自旋共振现象。

具体地写出来，态集合$e^{i\phi}|\psi\rangle(\phi \in [0, 2\pi))$被映到$S^2 \subset \mathbb{R}^3$中的一个点，其坐标为

$$X = \langle\sigma_1\rangle_\psi = 2\,\mathrm{Re}\,\overline{\alpha}\beta$$
$$Y = \langle\sigma_2\rangle_\psi = 2\,\mathrm{Im}\,\overline{\alpha}\beta$$
$$Z = \langle\sigma_3\rangle_\psi = |\alpha|^2 - |\beta|^2$$

Bloch球面坐标和纯态密度矩阵$\rho_{|\psi\rangle}$的关系为

$$\rho_{|\psi\rangle} = \rho_{e^{i\phi}|\psi\rangle} = |\psi\rangle\langle\psi| = \frac{1}{2}\begin{pmatrix} 1+Z & X-iY \\ X+iY & 1-Z \end{pmatrix}$$

对混合态(mixed state)的情形，密度矩阵与Bloch sphere内部的点一一对应。显然，(X, Y, Z)就是前面给出的Hopf映射$S^3 \to S^2$。

完全类似地，我们可以考虑一个two-qubit Hilbert空间，其中一个2-量子比特纯态可以写为

$$|\Psi\rangle = \alpha|00\rangle + \beta|01\rangle + \gamma|10\rangle + \delta|11\rangle, \quad \alpha, \beta, \gamma, \delta \in \mathbb{C}, \quad |\alpha|^2 + |\beta|^2 + |\gamma|^2 + |\delta|^2 = 1.$$

借助四元数代数$\mathbb{H}$，我们可以考虑一个类似的“Bloch球面表示”，其中用到Hopf映射$S^7 \to S^4$。而这里一个有趣的事实是：S^7的Hopf纤维化对纠缠态(entangled state)是“敏感”的(entanglement sensitive)，并且因此为2-量子比特态提供了某种“分层”(“stratification”)（参阅 [16]中的详细论述）。

4. **更多的应用例子**

[8]中提及了更多物理中应用Hopf纤维化的例子。

4 一点历史注记

Heinz Hopf最早为了计算$\pi_3(S^2)$而提出的Hopf纤维化 [1]是拓扑学中的重要发现，也在李群理论的发展中起到了不容忽视的作用。 [1]和 [2]无疑是早期同伦论发展的光辉成就。

P. A. M. Dirac对磁单极子的磁场所作的研究推广了传统波动力学的框架，具有重要意义。直到20世纪70年代，人们才意识到Dirac所描述的东西在数学上相当于纤维丛、截面和联络的语言 [17]。类似于数学上研究函数的Hilbert空间，人们发现需要研究截面的Hilbert空间 [18]。

"Dimensions"是一个非常著名的非商业数学系列科普短片，在各种视频网站上都可以容易地找到，其中使用大量直观的图象和动态视频演示了几何学的许多研究对象。它的第7、8两集深入地解释了纤维丛和Hopf纤维化的构造，基本涵盖了这则笔记的核心内容，观看之将给人非常大的启发。

参考文献

[1] Hopf, H. *Über die Abbildungen der dreidimensionalen Sphäre auf die Kugelfläche.* Math. Ann. 104(1931), 637-665

[2] Hopf, H. *Über die Abbildungen von Sphären auf Sphären neidrigerer Dimension.* Fund. Math. 25(1935), 427-440

[3] Steenrod, N. *The Topology of Fibre Bundles.* Princeton Univ. Press, 1974

[4] Lyons, D. *An Elementary Introduction to the Hopf Fibration.* Mathematics Magazine. 76(2003), no.2, 87-98

[5] Bott, R. and Tu, L. W. *Differential Forms in Algebraic Topology.* Springer, New York, 1982 (GTM82)

[6] Lickorish, W. B. R. *An Introduction to Knot Theory.* Springer, New York, 1997 (GTM175)

[7] Arnold, V.I. *Mathematical Methods of Classical Mechanics.* 2nd edition, Springer, New York, 1989 (GTM60)

[8] Urbantke, H. K. *The Hopf Fibration—seven times in physics.* J. Geom. Physics. 46(2003), no.2, 125-160

[9] Mishchenko, A. and Solovyov, Y. *Quaternions.* Quantum 11(2000), 4-7 and 18

[10] Bott, R. and Milnor, J. *On the Parallelizability of the Spheres.* Bull. Amer. Math. Soc. 64(1958), 87-89

[11] Kantor, I. L. and Solodnikov, A. S. *Hypercomplex Numbers, an elementary introduction to algebras.* Springer-Verlag, New York, 1989

[12] Nakahara, M. *Geometry, Topology and Physics.* Institute of Physics Publishing, Philadelphia, 1990

[13] Marsden, J. and Ratiu,T. *Introduction to Mechanics and Symmetry.* Springer-Verlag, New York, 1994

[14] Yang, C. N. *Fibre Bundles and the Physics of the Magnetic Monopole.* The Chern Symposium, Springer-Verlag, 1979, 247-254

[15] Dirac, P. A. M. *Quantised Singularities in the Electromagnetic Field.* Proc. R. Soc. London. A133(1931), 60-72

[16] Mosseri, R. and Dandoloff, R. *Geometry of Entangled States, Bloch Sphere and Hopf Fibrations.* J. Phys. A: Math. Gen. 34(2001), 10243-10252

[17] Wu, T. T. and Yang, C. N. *Concept of Nonintegrable Phase Factors and Global Formulation of Gauge Fields.* Phys. Rev. D12(1975), 3845-3857

[18] Wu, T. T. and Yang, C. N. *Dirac Monopole Without Strings: Monopole Harmonics.* Nuclear Phys. B107(1976), 365.

Four Equivalent Statements Concerning Baire-1 Functions*

Shi Xiaojie†

编者按:本文假定读者熟悉点集拓扑中的基本概念并知道 Tietze 扩张定理、单位分解定理及度量空间都是仿紧的。

In this article, I will give four equivalent statements concerning Baire-1 functions. Let's start with some definitions. (X will be a metric space if not specified otherwise.)

Definition 1 *Let f be a function defined on X, we say f is a* Baire-1 function *if there exist continuous functions $\{f_n\}_{n=1}^{\infty}$ such that $\lim_{n\to\infty} f_n(x) = f(x), \forall\ x \in X$.*[1] *The set of all Baire-1 functions defined on X is denoted by $\mathscr{B}_1(X)$.*

Definition 2 *A set V is called F_σ if $V = \bigcup_{n=1}^{\infty} V_n$ for some sequence of closed sets V_n.*

Definition 3 [2]*In set theory, two ordered sets X, Y are said to have the same* order type *when there exists a bijection $f : X \to Y$ such that both f and its inverse are monotone (order preserving).*

An ordinal number*, or just* ordinal*, is the order type of a well-ordered set. The finite ordinals are the natural numbers: 0,1,2,.... The least infinite ordinal is ω which is identified with the cardinal number $\aleph_0$. The set of all countable ordinals constitutes the first uncountable ordinal ω_1 which is identified with the cardinal $\aleph_1$.*

---

*本文系 NUS Undergraduate Research Opportunity Program 中的一个项目。指导老师 Denny H. Leung, National University of Singapore. 入选本刊时有删改.

†Department of Mathematics, National University of Singapore

[1]在通常的定义中 Baire-1 函数不包含连续函数，不过在本文中可以包含。

[2]编者加。参考 Wikipedia 中 Order type, Ordinal number, Successor ordinal, Limit ordinary 等词条。

The successor *of an ordinal number α is the smallest ordinal number greater than α. An ordinal number that is a successor is called a* successor ordinal. *A* limit ordinal *is an ordinal number which is neither zero nor a successor ordinal.*

Definition 4 *Let P be a closed subset of X, $f : X \to \mathbb{R}$, $\epsilon > 0$. $P(f, \epsilon) := \{x \in P :$ for all open neighborhood U of x in X, $\exists\, x_1, x_2 \in P \cap U$, such that $|f(x_1) - f(x_2)| \geq \epsilon\}$. Define $P^0(f, \epsilon) = P(f, \epsilon)$; $P^\alpha(f, \epsilon) = (P^{\alpha-1}(f, \epsilon))(f, \epsilon)$ if α is a successor ordinal; $P^\alpha(f, \epsilon) = \bigcap_{\gamma<\alpha} P^\gamma(f, \epsilon)$ if α is a limit ordinal.*

Define

$$\beta(f, \epsilon) = \begin{cases} \text{the smallest } \alpha < \omega_1 \text{ such that } P^\alpha(f, \epsilon) = \emptyset & \text{if such } \alpha \text{ exists} \\ \omega_1 & \text{otherwise.} \end{cases}$$

Define $\beta(f) = \sup_{\epsilon>0} \beta(f, \epsilon)$.

Note that $P(f, \epsilon)$ is a closed subset of P. If $0 < \epsilon_1 < \epsilon_2$, then $P(f, \epsilon_1) \supset P(f, \epsilon_2)$ and $\beta(f, \epsilon_1) \geq \beta(f, \epsilon_2)$. $\sup_{\epsilon>0} \beta(f, \epsilon) = \sup_{\epsilon_n} \beta(f, \epsilon_n)$ for any sequence ϵ_n goes to 0. If $\beta(f, \epsilon) < \omega_1, \forall\, \epsilon > 0$, then $\beta(f) < \omega_1$.

Definition 5 *A space X is said to be a* Baire space *if the following condition holds: Given any countable collection $\{A_n\}$ of closed subsets of X, if each of them has empty interior in X, then their union $\bigcup A_n$ also has empty interior in X.*

Theorem 6 (Baire Category Theorem) *If X is a complete metric space, then X is a Baire space.*

Proof: Let $\{A_n\}$ be as defined in Definition 5, we prove this theorem by showing that any nonempty open set U_0 in X contains a point x that does not lie in any set A_n, thus the interior of $\bigcup A_n$ must be empty, which is what we want.

Since $\text{Int}A_1 = \emptyset$, there exists $y_1 \in U_0$ such that $y_1 \notin A_1$. Then there exists $r > 0$ such that $B(y_1, r) \cap A_1 = \emptyset$ since A_1 is closed. Since U_0 is open, there exists $s > 0$ such that $B(y_1, s) \subset U_0$. Let $U_1 = B(y_1, \min\{s, t\}/4)$, then $\overline{U}_1 \subset U_0$ and $\overline{U}_1 \cap A_1 = \emptyset$.

Now U_1 is a nonempty open set in X, we can find $y_2 \in U_1$ such that $y_2 \notin A_2$ since $\text{Int}A_2 = \emptyset$. Do the same thing as above, there is an open ball U_2 such that $\overline{U}_2 \subset U_1$ and $\overline{U}_2 \cap A_2 = \emptyset$, we may shrink U_2 (if necessary) so that its radius is not more than $s/8$. Inductively, there is a sequence of open sets $\{U_n\}$ such that $\overline{U}_n \subset U_{n-1}$, $\overline{U}_n \cap A_n = \emptyset$ and the radius of U_n is not more than $s/2^{n+1}$.

Consider $\bigcap \overline{U}_n$, claim that $(\bigcap \overline{U}_n) \cap (\bigcup A_n) = \emptyset$; $\bigcap \overline{U}_n \subset U_0$; $\bigcap \overline{U}_n \neq \emptyset$. The first two statements are obvious since $\overline{U}_n \cap A_n = \emptyset$ and $\overline{U}_n \subset U_{n-1}$, respectively. For the third one, notice that if $z_1, z_2 \in U_n$, then $d(z_1, z_2) \leq s/2^n$. Therefore the sequence $\{y_n\}$ we construct is a Cauchy sequence. Since X is complete, there exists $x \in X$ such that $\lim_{n\to\infty} y_n = x$. For each n, since $\overline{U}_n$ is closed and $y_m \in U_n, \forall\, m \geq n+1$, we have $x \in \overline{U}_n$. Hence $x \in \bigcap \overline{U}_n$.

The point $x \in U_0$ does not lie in any A_n and thus is what we want. □

Theorem 7 *Let X be a complete separable metric space, $f : X \to \mathbb{R}$. The following four statements are equivalent (The first 3 are called Baire Characterization Theorem):*

1. *$f \in \mathscr{B}_1(X)$;*
2. *For all open set U of $\mathbb{R}$, $f^{-1}(U)$ is F_σ;*
3. *For all closed subset F of X, $f|_F : F \to \mathbb{R}$ has a point of continuity;*
4. *$\beta(f) < \omega_1$.*

Proof: (1)⇒(2)

For any open set U of $\mathbb{R}$, we can write $U = \bigcup_{i=1}^\infty V_i$, where $V_i = (a_i, b_i)$ are pairwise disjoint intervals. So we only need to show for any $a < b$, $f^{-1}(a,b)$ is F_σ.

Take arbitrary $c, d \in \mathbb{Q}$ with $a < c < d < b$, consider the set $\bigcap_{n=N}^\infty f_n^{-1}[c,d]$, where $\{f_n\}$ are as defined in Definition 1. Since f_n is continuous for each n, $f_n^{-1}[c,d]$ is closed, thus $\bigcap_{n=N}^\infty f_n^{-1}[c,d]$ is closed. Take

$$A = \bigcup_{\substack{a<c<d<b \\ c,d\in\mathbb{Q}}} \bigcup_{N=1}^\infty \bigcap_{n=N}^\infty f_n^{-1}[c,d].$$

A is F_σ since $\mathbb{Q}$ is countable. Now it suffices to prove $A = f^{-1}(a,b)$.

On one hand, for any $x \in f^{-1}(a,b)$, $a < f(x) < b$. Take $c, d \in \mathbb{Q}$ such that $a < c < f(x) < d < b$. Since $\lim_{n\to\infty} f_n(x) = f(x)$, there exists N such that for $n \geq N$, $f_n(x) \in [c,d]$, hence $x \in \bigcap_{n=N}^\infty f_n^{-1}[c,d] \subset A$.

On the other hand, for any $x \in A$, there exist $c, d \in \mathbb{Q}$, $N \in \mathbb{N}$ with $a < c < d < b$ such that $x \in \bigcap_{n=N}^\infty f_n^{-1}[c,d]$. Thus for $n \geq N$, $f_n(x) \in [c,d] \subset (a,b)$. Since $\lim_{n\to\infty} f_n(x) = f(x)$, we have $f(x) \in (a,b)$, hence $x \in f^{-1}(a,b)$.

(2)⇒(3)

Let F be a closed subset of X, define $F_0 = F$, then F_0 is a Baire space since F_0 is closed (hence complete) and metric. With F_{n-1} defined, we construct E_n, F_n and x_n as follows: (Note that we will use the information about F_n: F_n is closed in F_0 and $\mathrm{Int}_{F_0}(F_n)$ is not empty. For F_0 those two are clearly true, for other F_n we will prove them after construction.)

Use countable internals (a_i, b_i) with length $1/2^n$ to cover $\mathbb{R}$. For each (a_i, b_i), $f|_{F_0}^{-1}(a_i, b_i) = \bigcup_{j=1}^{\infty} V_{i,j}$, where each $V_{i,j}$ is closed in F_0. Let $T_{i,j} = V_{i,j} \cap F_{n-1}$, then $T_{i,j}$ is closed in F_0 since F_{n-1} does so. It is easy to show $F_{n-1} = \bigcup_{i=1}^{\infty} \bigcup_{j=1}^{\infty} T_{i,j}$, hence $\mathrm{Int}_{F_0}(\bigcup_{i=1}^{\infty} \bigcup_{j=1}^{\infty} T_{i,j})$ $=\mathrm{Int}_{F_0}(F_{n-1}) \neq \emptyset$. By the Baire Category Theorem, there exists some $T_{i,j}$ such that $\mathrm{Int}_{F_0}(T_{i,j}) \neq \emptyset$. Let $E_n = T_{i,j}$, then there exists $x_n \in \mathrm{Int}_{F_0}(E_n)$. Since $\mathrm{Int}_{F_0}(E_n)$ is open in F_0, there exists $B(x_n, \delta_n) \subset \mathrm{Int}_{F_0}(E_n)$ where $\delta_n > 0$.

Let

$$F_n = \overline{B(x_n, \min\{\delta_n/2, 1/2^n\})} \cap F_0,$$

then we verify the two conditions we assume about F_n. For the first one, F_n is certainly closed in F_0. For the second one, since $x_n \in B(x_n, \min\{\delta_n/2, 1/2^n\})$, we have $x_n \in \mathrm{Int}_{F_0}(F_n)$, therefore $\mathrm{Int}_{F_0}(F_n)$ is not empty.

For each n, it is easy to check that $F_n \subset \mathrm{Int}_{F_0}(E_n) \subset E_n \subset F_{n-1}$ and for $x, y \in E_n$, $|f(x) - f(y)| < 1/2^n$.

Consider the sequence $\{x_n\}$. For $n, m > N$, $x_n, x_m \in F_N$, hence $d(x_n, x_m) \leq 1/2^{N-1}$, therefore $\{x_n\}$ is a Cauchy sequence. Since F_0 is complete, there exists $x \in F_0$ such that $x = \lim_{n\to\infty} x_n$. For each n, since F_n is closed, we have $x \in F_n$, thus $x \in \bigcap_{n=1}^{\infty} F_n \subset \mathrm{Int}_{F_0}(E_n)$ for any n.

Claim that for this $x \in F$, $(x, f(x))$ is a point of continuity. For any $\epsilon > 0$, let $1/2^n < \epsilon$, then for any $y \in \mathrm{Int}_{F_0}(E_n)$, $|f(x) - f(y)| < 1/2^n < \epsilon$. Therefore, $(x, f(x))$ is indeed a point of continuity.

(3)⇒(4)

First we state a theorem.

Theorem 8 (Cantor-Baire Stationary Principle) *Let X be a separable metric space. If there exist $\{F_\alpha\}, \alpha < \omega_1$, such that F_α is a closed subset of X for each α and $\alpha_1 < \alpha_2$ implies $F_{\alpha_2} \subset F_{\alpha_1}$, then there exists $\beta < \omega_1$ such that $F_\alpha = F_\beta,\ \forall\ \alpha \geq \beta$.*

Proof: Suppose there does not exist such β, then for any $\alpha_1 < \omega_1$, we can find α_2 such that $\alpha_2 > \alpha_1$ and $F_{\alpha_2} \subsetneqq F_{\alpha_1}$. So without loss of generality, we can assume that $\alpha_1 < \alpha_2$ actually implies $F_{\alpha_2} \subsetneqq F_{\alpha_1}$. If not, we can construct a subset of $\{F_\alpha : \alpha < \omega_1\}$ namely F'_α $(\alpha < \omega_1)$ with $\alpha_2 > \alpha_1$ implies $F'_{\alpha_2} \subsetneqq F'_{\alpha_1}$.

Since X is separable, let $X = \overline{\{x_i\}}$. Consider

$$\{B(x_n, q) : x_n \in \{x_i\}, q \in \mathbb{Q}^+\}$$

The set is countable, we denote it by $\{O_n\}$.

For each $\alpha < \omega_1$, we show that there exists $O_j \in \{O_n\}$ such that $O_j \cap F_\alpha \neq \emptyset$ and $O_j \cap F_{\alpha+1} = \emptyset$. Since $F_{\alpha+1} \subsetneqq F_\alpha$, there exists $x \in F_\alpha \setminus F_{\alpha+1}$. Since $F_{\alpha+1}$ is closed, there exists $B(x, \delta)$ such that $B(x, \delta) \cap F_{\alpha+1} = \emptyset$ and $\delta \in \mathbb{Q}^+$. Since $\{x_n\}$ is dense, there exists $x_i \in \{x_n\}$ such that $x_i \in B(x, \delta/4)$. Then consider $B(x_i, \delta/4)$, we have $x \in B(x_i, \delta/4) \cap F_\alpha$ and $B(x_i, \delta/4) \cap F_{\alpha+1} = \emptyset$. So $B(x_i, \delta/4)$ is just the O_j we want to find.

For each $\alpha < \omega_1$, we can find a corresponding O_α and $O_\alpha \neq O_\beta$ for $\alpha \neq \beta$. This contradicts with $\{O_n\}$ is countable. So we conclude that there exists $\beta < \omega_1$ such that $F_\alpha = F_\beta$, $\forall\ \alpha \geq \beta$. $\Box$

Then we use Cantor-Baire Stationary Principle to complete the proof of (3)$\Rightarrow$(4).

Fix an $\epsilon > 0$, first show that $\beta(f, \epsilon) < \omega_1$. Suppose not, look at $P^\alpha(f, \epsilon)$ $(\alpha < \omega_1)$. It satisfies the assumption of Cantor-Baire Stationary Principle, so there exists $\beta < \omega_1$ such that $P^\alpha(f, \epsilon) = P^\beta(f, \epsilon)$, $\forall\ \alpha \geq \beta$.

Consider $P^\beta(f, \epsilon)$, it is a closed subset of X and is nonempty. So it has a point of continuity, namely x. By definition, $x \notin P^{\beta+1}(f, \epsilon)$. We get a contradiction.

So for any $\epsilon > 0$, $\beta(f, \epsilon) < \omega_1$. Thus, $\beta(f) = \sup_{\epsilon>0} \beta(f, \epsilon) < \omega_1$.

(4)$\Rightarrow$(1)

First we state two lemmas.

Lemma 9 *Let X be a metric space. For a bounded function $f : X \to \mathbb{R}$, define the norm $||f|| = \sup_{x\in X} |f(x)|$. Then for a sequence of bounded functions $\{f_n\}$, if $\sum_{n=1}^{\infty} ||f_n(x)|| < \infty$ and each f_n is Baire-1, then $f(x) = \sum_{n=1}^{\infty} f_n(x)$ exists and f is also Baire-1.*

Proof: For any x, $\sum_{n=1}^{\infty} |f_n(x)| < \infty$ since $\sum_{n=1}^{\infty} ||f_n|| < \infty$. Then for any $\epsilon > 0$, there exists N such that for $m > N$, $\sum_{n=m}^{\infty} |f_n(x)| < \epsilon$. Therefore, $f(x) = \sum_{n=1}^{\infty} f_n(x)$ exists for each $x \in X$.

For each f_i, let $M_i = \sum_{x\in X}\{f_i(x)\}$ and $N_i = \inf_{x\in X}\{f_i(x)\}$. Since f_i is Baire-1, there exists a sequence of continuous functions $\{f_{ij} : j \in \mathbb{N}\}$ which converges to f_i. For each f_{ij}, define f'_{ij} by

$$f'_{ij}(x) = \max\{N_i, \min\{M_i, f_{ij}(x)\}\},$$

then f'_{ij} is continuous, $||f'_{ij}|| \le ||f_i||$ and the sequence $\{f'_{ij} : j \in \mathbb{N}\}$ also converges to f_i.

Let $g_i(x) = \sum_{n=1}^{\infty} f'_{ni}(x)$, it is well-defined since

$$\sum_{n=1}^{\infty} ||f'_{ni}|| \le \sum_{n=1}^{\infty} ||f_n|| < \infty.$$

We have $g_i(x) = \lim_{m\to\infty} \sum_{n=1}^{m} f'_{ni}(x)$ and $\sum_{n=1}^{m} f'_{ni}(x)$ uniformly converges to $g_i(x)$. Since $\sum_{n=1}^{m} f'_{ni}(x)$ is continuous for each m, $g_i(x)$ is continuous for each $i \in \mathbb{N}$.

Let $F_m(x) = \sum_{n=1}^{m} f_n(x)$, $F_{mi}(x) = \sum_{n=1}^{m} f'_{ni}(x)$. For any $\epsilon > 0$, there exists $N \in \mathbb{N}$ such that $\sum_{n=N}^{\infty} ||f_n|| < \epsilon/3$. Therefore $||F_n - f|| < \epsilon/3$, for $n \ge N$. Hence $|F_n(x) - f(x)| < \epsilon/3$ for any $n \ge N$ and $x \in X$. Also we have for $n \ge N$, $x \in X$,

$$|F_{ni}(x) - g_i(x)| \le \sum_{n=N}^{\infty} |f'_{ni}(x)| \le \sum_{n=N}^{\infty} ||f_n|| < \epsilon/3.$$

Then for a fixed $x \in X$, since $F_N(x) = \lim_{i\to\infty} F_{Ni}(x)$, there exists an integer M such that for $i \le M$, $|F_{Ni}(x) - F_N(x)| < \epsilon/3$. Therefore $|g_i(x) - f(x)| \le |g_i(x) - F_{Ni}(x)| + |F_{Ni}(x) - F_N(x)| + |F_N(x) - f(x)| < \epsilon/3 + \epsilon/3 + \epsilon/3 = \epsilon$, this implies that $\lim_{i\to\infty} g_i(x) = f(x)$, for any $x \in X$. Thus $\{g_i : i \in \mathbb{N}\}$ is a sequence of continuous functions which converges to f, hence f is Baire-1. $\square$

Lemma 10 *If a sequence of Baire-1 functions $\{f_n\}$ uniformly converges to f, then f is also Baire-1.*

Proof: Since $\{f_n\}$ uniformly converges to f, then for any $\epsilon > 0$, there exists an integer N_ϵ such that for $n, m \ge N_\epsilon$, $||f_n - f_m|| < \epsilon$. Let $M_k = N_{1/2^k}$ for $k \in \mathbb{N}$. It is clear we can choose M_k such that $M_1 < M_2 < \cdots$. Take

$$g_n = \begin{cases} f_{M_1} & if n = 1 \\ f_{M_n} - f_{M_{n-1}} & if n \ge 2, \end{cases}$$

then g_n is bounder for $n \geq 2$. We have $\sum_{n=2}^{\infty} ||g_n|| < \infty$, thus by Lemma9, $\sum_{n=2}^{\infty} g_n$ is Baire-1. Therefore $f = f_{M_1} + \sum_{n=2}^{\infty} g_n$ is Baire-1 since f_{M_1} is Baire-1. $\square$

We now come back to prove (4)$\Rightarrow$(1).

If $\beta(f) < \omega_1$, then $\beta(f, \epsilon) < \omega_1$ for any $\epsilon > 0$. Fix arbitrary $\epsilon > 0$. Let X^0 be X; X^α be $X^{\alpha-1}(f, \epsilon)$ if α is a successor ordinal; X^α be $\bigcap_{\gamma<\alpha} X^\gamma(f, \epsilon)$ if α is a limit ordinal. ($P(f, \epsilon)$ is only defined when P is a closed subset of X. So to show each X^α is well-defined, we need to show X^α is closed in X for any α, we do this by induction. $X^0 = X$ is closed. Assume for any $\gamma < \alpha$, X^γ is closed in X. If α is a successor ordinal, $X^\alpha = X^{\alpha-1}(f, \epsilon)$ so a closed subset of $X^{\alpha-1}$, therefore X^α is closed in X since $X^{\alpha-1}$ is closed in X. If α is a limit ordinal, $X^\alpha = \bigcap_{\gamma<\alpha} X^\gamma(f, \epsilon)$, this is a intersection of closed sets in X, thus X^α is closed in X. So we can conclude for any α, X^α is closed in X.)

Since $\beta(f, \epsilon) < \omega_1$, we have

$$X = \bigcup_{\alpha<\beta(f,\epsilon)} X^\alpha \setminus X^{\alpha+1}.$$

Consider on $X^\alpha \setminus X^{\alpha+1}$ ($\alpha < \beta(f, \epsilon)$), for each $x \in X^\alpha \setminus X^{\alpha+1}$, there exists $\delta_x > 0$ such that for any $y, z \in B_{X^\alpha}(x, \delta_x)$, $|f(y) - f(z)| < \epsilon$. The balls form an open covering Λ of $X^\alpha \setminus X^{\alpha+1}$. Since $X^\alpha \setminus X^{\alpha+1}$ is a metric space hence paracompact, there exists a partition of unity Γ subordinated to the open covering Λ. For any $K_\beta \in \Gamma$, we can find a $U_\beta \in \Lambda$ such that $\mathrm{supp} K_\beta \subset U_\beta$. Those U_β also form an open covering Λ_0 of $X^\alpha \setminus X^{\alpha+1}$. For each $K_\beta \in \Gamma$, fix a $x_\beta \in \mathrm{supp} K_\beta$. Then for any $x \in X^\alpha \setminus X^{\alpha+1}$, there exists $r_x > 0$ such that $B_{X^\alpha \setminus X^{\alpha+1}}(x, r_x)$ only intersects finitely many $\mathrm{supp} K_\beta$. Therefore x only belongs to finitely many $\mathrm{supp} K_\beta$.

Let

$$g^\alpha(x) = \sum_\beta f(x_\beta) K_\beta(x),$$

$K_\beta(x) \neq 0$ for only finitely many β. Claim that for any $x \in X^\alpha \setminus X^{\alpha+1}$, $|f(x) - g^\alpha(x)| < \epsilon$ and also g^α is continuous on $X^\alpha \setminus X^{\alpha+1}$. In fact, by the definition of partition of unity, $\sum_\beta K_\beta(x) = 1$, therefore $f(x) = f(x) \sum_\beta K_\beta(x) = \sum_\beta f(x) K_\beta(x)$, $|f(x) - g^\alpha(x)| = |\sum_\beta (f(x) - f(x_\beta)) K_\beta(x)| \leq \sum_\beta |f(x) - f(x_\beta)| K_\beta(x)$. Since $x_\beta \in B_{X^\alpha}(x, \delta_x)$ for any β, $|f(x) - f(x_\beta)| < \epsilon$, thus $|f(x) - g^\alpha(x)| < \epsilon \sum_\beta K_\beta(x) = \epsilon$. For any $x \in X^\alpha \setminus X^{\alpha+1}$, on $B_{X^\alpha \setminus X^{\alpha+1}}(x, r_x)$, only finitely many $\mathrm{supp} K_\beta$ intersect the ball. Thus, $\forall\ y \in B_{X^\alpha \setminus X^{\alpha+1}}(x, r_x)$, $g^\alpha(y) = \sum_\beta f(x_\beta) K_\beta(y)$ with only those finitely many K_β. Since each K_β is continuous on $X^\alpha \setminus X^{\alpha+1}$, we have g^α is continuous at x. So we conclude $g^\alpha(x)$ is continuous on $X^\alpha \setminus X^{\alpha+1}$.

Define $F_\epsilon(x) = g^\alpha(x)$ for $x \in X^\alpha \setminus X^{\alpha+1}$, $(0 \leq \alpha < \beta(f,\epsilon))$. For any $x \in X$, $x \in X^\alpha \setminus X^{\alpha+1}$ for some $0 \leq \alpha < \beta(f,\epsilon)$, thus $|F_\epsilon - f(x)| = |g^\alpha(x) - f(x)| < \epsilon$. On each $X^\alpha \setminus X^{\alpha+1}$, $F_\epsilon(x) = g^\alpha(x)$, therefore $F_\epsilon(x)$ is continuous on any $X^\alpha \setminus X^{\alpha+1}$. Define $f_n(x) = F_{1/2^n}(x)$, then f_n converges to f uniformly.

Fix $\epsilon > 0$, let X^α $(\alpha < \beta(f,\epsilon))$ be as defined before. Let $T_{r,\alpha} = X^\alpha \setminus N_X(X^{\alpha+1}, r)$ where $N_X(X^{\alpha+1}, r) = \bigcup_{x \in X^{\alpha+1}} B_X(x,r)$. Notice that $T_{r,\alpha}$ is closed in X since we have shown that X^α is closed in X for any α. We also define $V_r = \bigcup_{\alpha < \beta(f,\epsilon)} T_{r,\alpha}$. Claim that for any α,

$$N_X(T_{r,\alpha}, r) \cap (V_r \setminus T_{r,\alpha}) = \emptyset,$$

where $N_X(T_{r,\alpha}, r) = \bigcup_{x \in T_{r,\alpha}} B_X(x,r)$. In fact, it suffices to show $N_X(T_{r,\alpha}, r) \cap T_{r,\beta} = \emptyset$ for any $\alpha \neq \beta$. Without loss of generality, assume $\alpha > \beta$. For any $x \in N_X(T_{r,\alpha}, r)$, $x \in B_X(y,r)$ for some $y \in T_{r,\alpha} \subset X^\alpha$, thus $x \in N_X(X^\alpha, r)$. Since $\alpha > \beta$, $\alpha \geq \beta + 1$, we have $X^\alpha \subset X^{\beta+1}$. Therefore $x \in N_X(X^\alpha, r) \subset N_X(X^{\beta+1}, r)$, $x \notin T_{r,\beta}$. Hence $N_X(T_{r,\alpha}, r) \cap T_{r,\beta} = \emptyset$ for any $\alpha \neq \beta$.

Claim that V_r is closed in X for any r. In fact, for a sequence $\{x_n\}$ in V_r, if it is convergent in X, there exists $N \in \mathbb{N}$ such that for $n, m \geq N$, $d(x_n, x_m) < r/2$. Since $x_N \in V_r$, we have $x_N \in T_{r,\alpha}$ for some α. For any $n > N$, $x_n \in B_X(x_N, r) \subset N_X(T_{r,\alpha}, r)$. Since $N_X(T_{r,\alpha}, r) \cap (V_r \setminus T_{r,\alpha}) = \emptyset$, we have $x_n \notin V_r \setminus T_{r,\alpha}$. Hence $x_n \in T_{r,\alpha}$ for $x_n \in V_r$. So we have for $n \geq N$, $x_n \in T_{r,\alpha}$ for a fixed α. Since $T_{r,\alpha}$ is closed in X, $\{x_n\}$ is convergent in X, we have $\lim_{n\to\infty} x_n \in T_{r,\alpha} \subset V_r$. So V_r is closed in X for any r.

We have shown that F_ϵ is continuous on each $X^\alpha \setminus X^{\alpha+1}$, so continuous on each $T_{r,\alpha}$. Then we show that F_ϵ is continuous on V_r. For any $x \in V_r$, $x \in T_{r,\alpha}$ for some α. Since F_ϵ is continuous on $T_{r,\alpha}$, we have for any $\Delta > 0$, there exists $\delta_1 > 0$ such that for $y \in B_{T_{r,\alpha}}(x, \delta_1)$, $|F_\epsilon(y) - F_\epsilon(x)| < \Delta$. Take $\delta_2 = \min\{r, \delta_1\}$. Since $N_X(T_{r,\alpha}, r) \cap (V_r \setminus T_{r,\alpha}) = \emptyset$, $B_{T_{r,\alpha}}(x, \delta_2)$ is actually $B_{V_r}(x, \delta_2)$. So we have for $y \in B_{V_r}(x, \delta_2)$, $|F_\epsilon(y) - F_\epsilon(x)| < \Delta$. Thus F_ϵ is continuous at x, F_ϵ is continuous on V_r.

Let $F_\epsilon|_{V_r}$ to be $F_{\epsilon,r}$. By Tietze Extension Theorem, $F_{\epsilon,r}$ can be extended from V_r to a continuous map of all of X, namely $\psi_{\epsilon,r}$. Let $\phi_{\epsilon,n}$ to be $\psi_{\epsilon,1/2^n}$. Claim that $\phi_{\epsilon,n}$ converges to F_ϵ. In fact, for any $x \in X$, $x \in X^\alpha \setminus X^{\alpha+1}$ for some α. Since $X^{\alpha+1}$ is closed in X, $x \notin X^{\alpha+1}$, then there exists $\delta > 0$ such that $B_X(x, \delta) \cap X^{\alpha+1} = \emptyset$. Take $1/2^N < \delta$, we have $x \notin N_X(X^{\alpha+1}, 1/2^N)$. Therefore $x \in X^\alpha \setminus N_X(X^{\alpha+1}, 1/2^N) \subset V_{1/2^N}$. For $n > N$, $\phi_{\epsilon,n}(x) = \psi_{\epsilon,1/2^n}(x)$. Since $x \in V_{1/2^N} \subset V_{1/2^n}$, we have $\psi_{\epsilon,1/2^n}(x) = F_{\epsilon,1/2^n}(x) = F_\epsilon(x)$. Thus $\phi_{\epsilon,n}(x)$ converges to $F_\epsilon(x)$ for any x. Since each $\phi_{\epsilon,n}$ is continuous, F_ϵ is Baire-1. Therefore each f_n we defined before is Baire-1. Since f_n converges to f uniformly, by Lemma10, f is Baire-1, which completes the proof. □

参考文献

[1] James Dugundji, *Topology*

[2] James R.Munkres, *Topology*

[3] Thomas Jech, *Set Theory*

****

数学家逸事

有一天，波兰数学家Sierpinski要搬家，他的夫人把行李拿出来以后对他说："我去叫辆出租车，你在这儿看好行李，总共有10个箱子。" 过一会儿，他的夫人回来了，他对夫人说道："刚才你说有10个箱子，可是我数了只有9个箱子。"

"不对，肯定是10个。"

"说什么呢，我再数一遍，0，1，2，3……"

不变集合和极值原理

马力*

引言: 这个小短文基本上是来自我在讨论班上的演讲材料。目的是通过阅读这些内容，读者可以看到不变集合，极值原理和LIUOVILLE型定理的密切关系。

最近BRENDLE-SCHOEN用HAMILTON的RICCI流（它是几何分析中的基本研究对象）解决了微分球定理。其中的基本手段就是找这个流的不变集合。我们利用这个机会谈微分方程的不变集合，它的特殊情况就是极值原理。我们谈到的主要内容有BONY的极值原理。这是一个深刻的数学定理，所以值得在这里好好谈谈。

由于几何分析的研究对象是来自几何和物理中的模型，大家自然要感兴趣这些模型的不变集合。对于热方程，不变集合就是极值原理，利用这个观念，Hamilton得到张量方程组的极值原理。但是对于波动方程来说，不变集合的观念需要有更多的思考。比如在新近我们研究的流形上的非线性SCHRÖDINGER方程，我们就是引入了新的不变集合，从而可以解决一些前人不能解决的问题。

我在这里挑选这么一个话题来讲给大家，主要的目的是希望同学们能深入理解微分学中的基本概念，这样在以后的数学研究中才可以做出更深刻的结果出来。

1 基本的微分不等式

很多同学都知道微分方程的比较原理。其实，这个原理是和（偏）微分方程的不变集合密切联系着的。微分方程的不变集合，它的特殊情况就是极值原理。大家都知道，利用极值原理可以得到解的梯度估计，然后就可以容

*清华数学系基础数学所教授

易的得到LIOUVILLE定理。这样就可以自然的得到一个代数基本定理的分析证明。

我们先来看怎么对LIPSCHITZ函数微分的问题[3]。

假定$h = h(t)$ 是一个LIPSCHITZ函数。我们定义它的DINI导数如下：

$$D^+h(t) = \lim\sup_{s\to 0^+} \frac{h(t+s)-h(t)}{s}$$

和

$$D^-h(t) = \lim\inf_{s\to 0^+} \frac{h(t+s)-h(t)}{s}$$

容易看出关系

$$D^+(e^{Ct}h(t)) = e^{Ct}(Ch(t) + D^+h(t))$$

这里C是一个常数。

这样我们就可以得到下面的引理

引理1： 假设$h(0) \le 0$和$D^+h(t) \le 0$，其中这里的t满足$0 \le t \le T$, 于是$h(T) \le 0.$

证明：

取$\epsilon > 0$，我们来证明有

$$h(t) \le \epsilon t$$

利用条件，

$$D^+h(t) \le 0$$

我们有小时间段$[0, d]$, 使得

$$h(t) \le \epsilon t$$

定义T_m为这种关系成立的最大区间。

利用连续性，我们有

$$h(t) \le \epsilon t$$

对于$t \in [0, T_m]$成立。

如果$T_m < T$, 那么根据条件，

$$D^+h(t) \le 0$$

我们得到$\delta > 0$,

$$h(t) \le \epsilon t$$

对于$t \in [0, T_m + \delta)$成立。

所以有

$$h(t) \le \epsilon t$$

对于$t \in [0, T)$成立。 **证毕。**

这个结果有好几个有意思的推论。

推论1. 假设$h(0) \ge 0$和$D^- h(t) \ge 0$，其中这里的$0 \le t \le T$, 于是$h(T) \ge 0$.

推论2. 假设$h(0) \le 0$和$D^+ h(t) \le Ch(t)$，其中这里的$0 \le t \le T$, 于是$h(T) \le 0$.

这个推论的证明只需要一个变换，也就是我们要考虑$g = e^{-Ct}h$。

如果有$h = f - g$, 特别的，我们有下面的比较原理。

推论3. 假设$f(0) \le g(0)$和$D^+ f(t) \le D^- g(t)$，其中这里的$0 \le t \le T$, 于是$f(T) \le g(T)$.

2 BONY的不变集合原理

假设D是R^n上的区域，$X : D \to R^n$ 是一个连续的向量场。F是D中的闭集合。我们研究微分方程$x'(t) = X(x(t))$ 解的轨迹。

我们说集合F是X的不变集合是说，对$x(0) \in F$,我们的解$(x(t)), t > 0$就一直在F里面。数学语言上说就是$x(t) \in F$, $t > 0$.

对于F中的点y来说，如果有一个球S使得y在S上，但是F中的其他点都不在S的内部的话，我们就可以定义F在y点的法向量为

$$\nu(y) = x - y$$

这里x是球S的球心。

定理（Bony）[1]：假设X和F满足以下条件，(1) 存在正常数L使得

$$|X(x) - X(y)| \le L|x - y|$$

(2) 对F在$y \in F$处的法向量$\nu(y)$ 我们有

$$\nu(y) \cdot X(y) \le 0$$

那么F是向量场X的不变集合。

证明；

如果这个结论不是真的，那么我们可以定义

$$h(t) = dist(x(t), F), \quad t \geq 0$$

这里$h(0) = 0$, $h(t) > 0$, $\ t > 0$. 固定$t > 0, s > 0$, $h(t) = dist(x(t), y)$, $y \in F$, 我们考察

$$h(t+s) = dist(x(t+s), F) \leq dist(x(t+s), y) = |x(t+s) - y|.$$

根据

$$|x(t+s) - y| - |x(t) - y| = \frac{|x(t+s) - y|^2 - |x(t) - y|^2}{|x(t+s) - y| + |x(t) - y|}$$

我们有

$$h(t+s) - h(t) \leq \frac{|x(t+s) - y|^2 - |x(t) - y|^2}{|x(t+s) - y| + |x(t) - y|}.$$

再根据

$$x(t+s) = x(t) + sX(x(t)) + o(s)$$

和

$$|x(t+s) - y|^2 - |x(t) - y|^2 = 2s(x(t) - y) \cdot X(x(t)) + o(s)$$

令$s \to 0^+$ 我们有

$$D^+h(t) \leq \frac{(x(t) - y) \cdot X(x(t))}{|x(t) - y|} \leq \frac{(x(t) - y) \cdot (X(x(t)) - X(y))}{|x(t) - y|}.$$

于是有

$$D^+h(t) \leq Lh(t).$$

利用上面的引理1的推论，我们就知道$h(t) = 0$.

这样就证明BONY的不变集合定理。

3 LIOUVILLE型定理和极值原理

下面的极值原理在几何分析中是非常有用的。所谓极值原理就是数学分析中的基本关系：如果$x \in D$是$u : D \to R$函数的极大值点，那么我们有x处的关系$\nabla u(x) = 0$ 和

$$D^2u(x) \leq 0.$$

特别的，我们有$\Delta u(x)=trD^2u(x)\le 0.$

我们先来看一个简单的应用，这是一个LIOUVILLE型定理。我们这里的论证用到REDHEFFER的一个想法。

定理: 假如我们有有界光滑函数$u=u(t), t\in R$满足

$$u_{tt}=f(u)$$

这里$f:R\to R$也光滑并且$f'>0$. 那么u是常数。

证明：给定任意点x和$\epsilon>0$，定义

$$w(t)=u(t)-u(x)+\epsilon-\epsilon|t-x|^2.$$

于是$w(x)=\epsilon>0$ 和$w(t)\to-\infty$对$|t|\to\infty$. 所以有w的最大值点y，在这点处，$w(y)>0$和$w_{tt}(y)\le 0$,从而$u(y)>u(x)-\epsilon$和$u_{tt}(y)\le 2\epsilon$。再利用

$$u_{tt}(y)=f(u(y))\ge f(u(x)-\epsilon)$$

我们得到

$$f(u(x)-\epsilon)\le 2\epsilon.$$

令$\epsilon\to 0$, 我们得到$f(u(x))\le 0.$

定义

$$v(t)=u(t)-u(x)-\epsilon+\epsilon|t-x|^2.$$

于是$v(x)=-\epsilon<0$ 和$v(t)\to+\infty$对$|t|\to\infty$. 所以有v的最小值点z，在这点处，$v(z)<0$和$v_{tt}(z)\ge 0$,从而$u(z)<u(x)+\epsilon$和$u_{tt}(z)\ge -2\epsilon$。再利用

$$u_{tt}(z)=f(u(z))\le f(u(x)+\epsilon)$$

我们得到

$$f(u(x)+\epsilon)\ge -2\epsilon.$$

令$\epsilon\to 0$, 我们得到$f(u(x))\ge 0.$

这样就有$f(u(x))=0.$从而u是常数。

证毕。

最后我们给出

BONY的极值原理：假设D 是R^n的开集。假设X 是D上的光滑向量场。设$u=u(x), x\in D$ 是一个非负的光滑函数并满足

$$D^2u(x)(X,X)\le -K\inf_{|\xi|\le 1}(D^2u)(x)(\xi,\xi)+Ku,\ x\in D$$

这里K是一个正常数。定义$F = \{x \in D; u(x) = 0\}$。设$\gamma : [0,1] \to D$ 是一个光滑曲线使得$\gamma(0) \in F$ 和

$$\gamma'(t) = f(t)X(\gamma(t))$$

这里$f(t)$ 是光滑函数。于是有$\gamma(t) \in F$, $t > 0$.

由于这个定理论证比较长，我们这里就不证明了。读者可以去看文献中[2]的证明。

参考文献

[1] J.M.Bony, *principle du maximum, inegalite de harnack et unicite du probleme de Cauchy...*, Ann. Inst.Fourier (Grenoble), 19:1(1969)277-304.

[2] S.Brendle, R.Schoen, *Classification of manifolds with weakly 1/4-pinched curvatures*, Acta mathematica, 200(2008)1-13.

[3] R.Hamilton, *Four-manifolds with positive curvature operator*, Journal of Diff. geometry, 24(1986)153-179.

****

数学家趣闻

▲华裔数学家钟开莱（1917-2009）原本就读于清华大学（西南联大）物理系，当时西南联大理学院院长是吴有训先生。吴有训开设光学课，钟开莱听了几次课以后觉得他讲授的内容书上全都有，自己看书自学就可以，就开始逃课。可当时理学院听课的学生只有寥寥十来人，走了一个很现形，吴有训很快发觉钟开莱逃课的事实，大发雷霆。钟开莱担心今后的日子混不下去，所以转到了数学系。

钟开莱给人的印象那是牛逼哄哄的。甫列华氏门墙，一天华老爷子啰里八嗦讲了一大堆东西，钟开莱不服气回家折腾了十页的纲领性文字，第二天上课的时候扔给华罗庚，意思是你拿去看吧。华老爷子很不爽，说你不就是个毛头孩子吗，跟我玩这套，华罗庚也折腾了一晚上，硬是把这10页压缩成3页，回敬给钟开莱。当时（西南联大时期）学习氛围可见一斑。

刘思齐老师访谈录

编者按：本文由刘思齐老师于 2010 年 4 月 8 日“走近理学院”活动中与数学系同学交流的记录整理而成。作为清华数学系自己培养的年轻教师的代表之一，刘思齐老师与学生们分享了他的学习和科研经历，并回答了同学们的一些问题。我们将整个互动过程一并整理发表，以飨读者。整理时，部分字句有所改动。

（最开始是漫谈部分）

我其实当年是靠数学竞赛进来的。进来之后，发现自己跑到物理系去了。所以大三分流的时候，我就毫不犹豫地选择了数学系。当时我对物理也很有兴趣，所以就选了个数学物理，看起来好像跟物理有关的(方向)。但进来之后发现，其实跟物理关系也不是很大。到了数学系之后，我发现我的数学基础不是非常的好。当时小基科学的这个数学课程跟物理系学的数学课程是差不多的。跟数学系相比还是很不一样。相对来说我数学成绩比较差。当时没有意识到这个问题，有讨论班就跟着上，导师要读点什么就读点什么，总之当时也没有发现。一直到我当助教，教数学分析或者高等代数，我才发现当年学得很差，但是这个时候我已经来不及了。不过事实上数学这个东西总是这样的，你当时不管学多好，过几年再回头看，总是会发现当时有一些东西不懂。作为弥补，以后要用到这些东西的话，你会逼着自己重新去学一遍，这样慢慢地回忆起来。当时，我是以一个不太好的数学基础跑到数学系跟张友金老师学可积系统。谈到这个东西，第一个问题肯定就是“什么是可积系统”。这个问题没有一个很好的答案——大家也不知道什么是可积系统。我在系里做过一个报告，讲的是什么是一个好的数学领域。不知道你们有没有学过，我觉得拓扑学就是一个比较好的领域，比较成熟的数学领域。为什么说比较成熟呢？拓扑学和代数——例如李代数——之类的课本，一般书的第一章第一页都会告诉你这个学科研究的是什么。例如拓扑学第一页就会告诉你什么是拓扑空间，拓扑空间就是一个集合加上开集条件。李代数呢，就是一个线性空间加上一个双线性运算满足一些恒等式。像这些成熟的学科，第一件事情就告诉你我们这个学科要研究什么。就好象一个国家建立起来之后得划分国界线，我们这个学科以后就研究这个国家以内的事情，国界线以外的事情就不管了。但是可积系统这个东西很不一样。随便拿一本可积系统的书，你从第一页翻到最后一页也找不到可积系统的定义。所以这个就是不太成熟的一个学科，随便拿本书，发现它主要都在讲例子，说这个东西是可积系统，那个东西是可积系统。或者讲一些特殊的性质，比如孤子可积系统，Lax 可积系统等等，但是并没有一个统一的定义。打个比方，就好比我们还没有建立一个国家，在这个范围内有好多游牧民族正在打架。大家各自定义一个可积的概念，但这些概念是矛盾的。一个系统可能满足这个性质却不满足那个性质。大体情况就是这样。我当时数学虽然学得

不多，但是我觉得很不舒服。我之所以跑到数学系，是因为我觉得物理系的数学太差了。看广义相对论还好，看量子场论就觉得里边的数学太烂了。于是我就转到数学系去。但是到这边也有这个问题。我要做研究的话，就得先把可积系统这个概念搞清楚。所以接下来一段时间做的研究都是关于这方面的。当时也是刚入门，做的都是一些小问题。我是 01 年秋天开始做 seminar 的。用一个学期的时间学了些基本的知识，第二学期就开始琢磨可积性的定义，做了一年也没有做出来。当时发现了一些东西，但是想要最终解决这个问题还是解决不了，就卡在一个地方过不去。到 03 年的时候，张友金老师给了我一个题目。他这个题目的好处也可以用前面的比方来说明。成熟的领域虽然一上来就划了国界线，但是具体到研究的问题上还是在处理一些比较好的东西。比如拓扑学研究的东西不是一般的拓扑空间，而是加一些比较好的拓扑的条件，比如加一些紧性的条件，研究的都是比较好的对象。像李代数这个东西，你要学过这门课就知道半单李代数的分类。加上一个半单的条件就得到一些好的结果。张老师给我的题目就是一类特别好的可积系统。我的工作就是研究这类系统在坐标变换下的分类。那个东西我从 03 年开始做，二三月份的时候张老师把题目给我，到七八月份的时候我觉得我做出来了。03 年正好是我本科毕业的那年，所以我的毕业论文就是把这个问题的细节说全。当时是把这个要分类的东西展开成无穷级数，把它的坐标变换也展开成无穷级数，然后要证明坐标变换后的某些性质。当时作为本科生，对级数的前两阶给出了证明。毕业之后，我觉得我可以写出一般的证明。后面的东西就非常复杂，总是有些很难算的东西。后来张老师发现一个漏洞，指出我的证明是过不去的。花了很长时间也没能修补，就这样搁置了三个月。到了 10 月份的时候才再次拿起来。当时这个过程对我以后的研究很有启发。一个东西做不出来，不要钻牛角尖，可以先放一放，去做别的事情。培养出一种陌生感之后，能让你跳出这个圈子，看到新的突破口。当时我三个月没有看什么数学的东西。到了 10 月 11 月回过头来看，发现一个更简单的办法，比我 8 月份走的路子要简单得多。然后这就成了我的第一篇文章。解决之后，我又接下来考虑一个相关的、更难的问题。这些问题研究的都是有若干分量的偏微分方程。大概到 04 年一二月份的时候对单变量的情形就证明了，但是推广到一般的多变量，大概花了 10 多个月的时间才把这个工作做完。我发现我每做一个东西大概都要花 10 个月。这些问题都是可积系统中的分类问题。当时我已经是博士了，需要参加一些资格考试，还要做社会实践。05 年我又回过头来看当时 02 年做的东西。这里有一个很有意思的经历。你们中应该有些人学过常微分方程，知道一阶拟线性偏微分方程可以归结为常微分方程来解，用特征线法找一些守恒量，然后就可以解出来。当时我解的是一个无穷多个变量的这种方程。02 年的时候我算出了前面若干个守恒量，但我不知道后面是什么东西，当时也猜不出来。然后 04 年开始，我做微积分助教，需要做一些习题，还要给学生改和讲作业。有一个很简单的事实大家微积分都学过，就是反函数的导数等于原来函数导数的倒数。一般人不知道的是，反函数的二阶导数，算出来是原来函数的二阶导数除以原来函数一阶导数的三次方，

可能还有个负号。接下去可以算出三阶导数四阶导数是什么，结果越来越长，很难写出个一般的表达式。这时候我才发现，02 年算出来的那些守恒量其实是反函数的高阶导数，发现这一点之后，02 年那个问题立刻就解决了。这已经是 05 年接近 06 年的事了。所以这个事情也说明事情放到一边，回头再看会带来新的想法。这也就是我的第三篇文章。前面两篇文章都是研究的一些比较好的东西，05 年过后开始做的就是一些不那么好的东西了。我之前的两个东西就好比是在国家的首都里搞建设，把城区规划得有秩序一点。而 05 年做的东西就像是在国界线上，把国界线划清楚一点。当时做的都是单变量的情形，最近这几个月开始考虑多变量的情形。类比是有的，但是在一些细节上的困难，现在还在研究。我学数学的经历大概就是如此。

（可积系统这个学科在世界上的发展情况如何？）
现代数学的一大趋势就是好的数学都是相互联系的，我这个学科跟其它的学科关系也很大。现在很难说一个学科的发展怎么样，因为所有学科都是连在一起的。我们做的这个非常基础的东西，就是分类问题。我觉得任何一个学科的目标，最重要的就是分类问题。也就是说，把这个国界线划清楚之后，还要做一下人口普查。分类问题做好之后，可以很快地用到其它的领域。具体一点，我们的出发点是一个不变量。我想你们可能不知道这些细节，我讲点历史。90、91 年 Witten 研究二维量子引力。这个东西有两种做法，一种是把它归结为矩阵模型，然后可以算出一些性质来；还有一种就是归结为稳定曲线模空间上的积分，转化为代数几何的问题。但是这个世界上只有一种量子引力，所以用代数几何方法算出来的积分的生成函数应该满足矩阵模型里的某些微分方程，这些微分方程就是可积系统里的 KdV 方程。这个问题解决了，之后又推广到更一般的情况。二维量子引力可以看作一个曲面到一个点的映射，然后对其进行积分。如果把那个作为靶子的点换成一般空间的话，就能得到一般的东西，这就是 Gromov-Witten 不变量，在物理里面大概叫拓扑弦。不管是数学还是物理，有一点是很像的，同样的东西，可能从不同的角度去看，以发现一些联系。所以之后这个不变量都是很多人研究的方向，最近 10 年来都是很热的。我们做这个东西就试图用可积系统的方法把这个东西确定下来。这些不变量的生成函数满足一族微分方程，我们就可以通过解微分方程把它们解出来。但是我们对可积系统了解得太少了，任给一个靶空间，我们根本就不知道对应的可积系统长什么样子。所以就需要一种分类，首先要把所有的可积系统都找出来，这些系统间有坐标变换，将对应的等价关系模掉。至于你说的进展，我只能给一些个人的偏见了。我觉得（可积系统领域中的）其他问题和分类问题相比都不是很大的问题。其他的问题还有研究求解，研究对称以及一些具体的应用，比如流体力学等等。那些当然也是很好的东西，只是我个人不太关心这方面。

（想做一些东西，其中要用的东西发现自己没学过，遇到这种问题怎么办）

这是很常见的，因为现在要学的东西太多了，就现学吧。但是现学你要注意，你们现在本科阶段，课程学的东西都是将来要用到的基础。到了研究阶段，就得抓住重点。用到哪部分，就把需要的这部分看过，会用就行了。有时候不能完全看懂，也不能把相关的东西都学完——都学完就没时间做研究了。到以后你会发现，做研究和好好学东西两个方面是矛盾的。因为时间是有限的，如果一件事情占去，另一件事就没有时间了，总要找一个平衡。你得记住你的目标是什么。数学最终的目标终究还是要做研究而不是学以前的东西，现在读书也是为了以后做研究作准备。以后做研究需要学什么东西的话，学到基础的就可以了，不用完全搞彻底。大家都学过数学分析吧，数学分析一开始就讲集合论。集合论可以写一本很厚的书来讨论。特别朴素的集合论是数学的基础，也有很多很难的问题。但对以后的研究并没有什么太多的帮助。

（没有学卓里奇对以后的研究有没有什么影响？）

你们除了卓里奇还学什么书？（学生答：科大的徐书）很难说啊。卓里奇当然是很好的，但是有些人不适应。科大的书呢可读性比较高，不像卓里奇，至少它的习题是可以做出来的。说到最后还是看人啦，有的人学这个比较好，有的人学那个比较好。两本书都挺不错的。

（大三你选择这个方向的时候是凭什么想法选择的，还是凭感觉的）

我刚才说过我对物理比较感兴趣，看一些广义相对论或者量子场论的东西。我当时想的是把数学学好了，回头来看量子场论，数学基础会好一点。所以我就想跑到数学系来学跟物理有关的数学。当时数学系的方向中有一个是张友金、周坚，还有另外三个。我当时有个同学是来学数论的，就找了别的老师。这个同学也是我们中做得比较好的一个。总之当时最初的想法是为了学物理，但学着学着才发现这方面跟物理的关系还是比较远的。当然还是有很深的物理在里面，但总而言之，不太用物理的理论，像利用的不变量都是数学的想法。比如上同调场论，和物理某个领域做的东西一样，都用到一些变换。虽然做的东西很相似，但基本的想法还是不一样的。所以要学物理的话，还是按物理的思路从下往上学上去比较好，而不是从数学这边往物理跳，跳不过去的。物理有物理的想法。所以现在的话我也不能再去从事物理了。以前跟学堂班的同学座谈的时候，他们有个问题，就是量子力学和分析力学应该选哪个。我当时说，这两个东西不是二选一的问题，必须先学分析力学，然后才能学量子力学。学量子力学，对数学的很多方面也是有用的，因为数学现在有很多东西都是量子化的。

（你们解决科研问题的话一般都是发明新的方法吗？）

做问题肯定是要有新的东西，要不然你是做不出来的。但是具体到我们，这里其实有不同的风格。比如那些很牛的人，为了解决一个问题，他就造了一个很大的机器。看起来他写的书和文章，每句话都是废话，但是这些废话连起来就是

有用。比如 Grothendieck 和他的学生、同事花 20 年时间写了一万页的书，把一个问题解决掉。他的一个很重要的想法是要自然，每个部分都不应该有人工的痕迹，就是最自然地做。举个例子，就像吃核桃，一锤子砸开，然后挖出核来吃。他会觉得这样很暴力，而选择把核桃放在水里，等壳泡软了用手掰开吃。他做事的风格就是崇尚自然，一定要理解到最自然的程度。但是这件事是需要花时间的。理解一个东西，首先本人需要有很深刻的认识，才能把一个问题理解透彻。另一种风格就是，还是拿核桃作比方，就用一个小针把核桃壳一块一块撬下来。或者说，前面那个人杀鸡用牛刀，而这种人杀牛用小杀鸡刀，最后也能一刀刀把牛给杀掉。这种方式，可能也会引入一些概念，但都只是为了叙述的方便，不会那么多废话。具体做起来，也是把最基础的东西解决了，没有什么深入的探索。这个跟我自己做的方向有一定关系，因为方程总是比较现实的。当然我们也希望有一些工具。用一把小刀固然是把牛砍死了，但是过程很漫长，就像我们现在的计算非常复杂。但如果真的要有了强有力的工具的话，效率会大大提高。所以我们现在也试图发明一些好的工具。至少现在我们还在用最土的办法，用小刀一刀刀杀牛。其实我想数学界的人是更愿意用前一种方法去杀的，但大多数人使用的还是后一种方式。为什么你们现在要学好数学分析和高等代数，因为以后的很多问题都归结为两类思想——要么是分析的思想，要么是代数的思想。别的东西和结构其实都是一种语言，思考的时候还是用到数学分析和高等代数。不过这些语言的好处是简化一下表达，但本质上都还是你大一学的东西。

（你们在学术前沿的交流主要是靠什么，发 paper 和做学术报告吗）

以前都是发在期刊上，当然发在期刊上有个问题，就是比较慢。我做了一个东西，发到期刊上还要等很长的时间。另一个问题就是优先权。比如我做的一个东西和其他人撞车了，我比他快一点，然后投给期刊。但期刊的负责人可能正好是我的竞争对手，于是就把我的稿子压几个月。而他和他的合作者把他们的文章发到另一个比较快的期刊上，这样他们就成了先发表的。这种事情以前是比较多的。后来有一群做理论物理的人对这个现象很不满意，他们就架了一个服务器，让大家把文章都发上面。这样第一大家能够第一时间方便地看到他人的成果，顺便也解决了优先权的问题。这个办法让很多做数学和物理的人都感到很满意，于是这个服务器越做越大，就成了一个很重要的交流平台：arXiv。现在大多数做数学、物理、生物和统计的人都会把自己做好的文章放到这个地方。而且这个地方它有一些规则。你当然可以修改你的文章，但你以前的版本要一直留在那里。这个地方也没有谁管，文章都随便下载。但是上传的时候个人的身份是要留证据的，不是什么都能往上放。一般来说，邮箱域名是学术机构的用户上传文件更方便，而使用公共邮箱服务的人，文章往往只能放在“一般”分类里了，除非得到学术机构人员的背书。总之这个地方就是交流文章的。在其他的地方讨论也有，像我们每逢假期的时候都会往外跑。平时开学的时候在这里上课，放假的时候，做研究的人都不会在家呆着，而是跑到外面开会。上课的时间要开会的话挺麻烦，

要耽误课程。用 email 讨论的话往往不太好，还不如面对面的交流更有效率。

（在网上，一般通过什么方式来获得需要的资料）

主要是数据库。数学类的话有 AMS 美国数学会的 Math Review。所有数学的文章、书和其他出版物，它这里都会有记录。然后别的人对这些记录做评论。这里的东西通常会滞后一点，也就是说有的文章已经出版了，但是还没有入库。看看文章的评论，你会知道这个文章大概讲了些什么东西。相比摘要，有个其他人的评述会了解得更清楚。当然这里的评论赞美的居多，但要遇到真正写得不好的，批评也相当严厉。一般没有评论的文章都表示“这篇文章不值得评论”，价值很可能比较低。除了这个之外，还有一个德国的数据库也是收录各种数学文章的。名字我忘了，是一个 Z 打头的德文（编者按：Zentralblatt MATH），大家去搜一下“zmath”就知道了。跟 AMS 比的话，Z-Math 的东西要少一点。除了这些，有的时候我们也会看一些很一流的期刊，比如 Annals。因为 AMS 的收录是比较慢的，而且它得花时间写评论，而文章早就发在期刊上了。所以有的时候直接去看期刊会比较快。比如大家都觉得可能计算数学现在不是主流，很难想象计算数学的文章能发表在第一流的期刊上。但是前几年就有一篇研究牛顿法解多项式方程的文章发在 annals 上。牛顿法的关键在于初始点的选择，如果初始点选得不好的话，就会陷入循环，找不到解。这篇文章就证明了在平面上有某组点集，对所有方程都可以找到确定的解。所以这篇文章还是很有价值的。杂志的话，最好的两个是 Annnals of Mathematics 和 Inventiones Mathematicae。这些都是综合性的杂志，也有一些偏向某个方向的，比如偏向偏微分方程的 Communications on Pure and Applied Mathematics。判断一个期刊的影响力可以看影响因子，它是利用期刊间相互引用的情况和发表文章的数量来计算的。当然这个因子是可以通过人为手段来作弊增加的，所以这也不是唯一的标准。但如果你不了解这个研究方向，也可以作为参考。

（和大数学家接触感受是什么样子的？）

GTM 你们知道吗？GTM 就是 Springer 出版社出的一套黄皮的数学教材。GTM 的教材都还是比较好的，其中有一套叫《现代几何》，不知道你们看过没有，作者之一是 Dubrovin。我导师张友金 90 年左右在俄国待过一年，结识了 Dubrovin。当时苏联正在解体，他就回来了。后来 Dubrovin 去了意大利的 SISSA 研究所，于是我们也就经常往意大利跑，因为现在我们做的很多问题都是跟 Dubrovin 合作的。我和 Dubrovin 主要还是数学上的交流，有的时候也有一些跟数学无关的话题，比如一起喝酒。Dubrovin 也有很多别的爱好，他上次去西藏，就背个大包，带了很专业的登山装备去爬喜马拉雅山。今年有一个会议就是庆祝他 60 岁生日。

（学生接着问：像他这种年纪很大的数学家还在自己做问题吗？还是基本上是手

下的人做？）

他们这种年纪的人一般不做太细的东西了，就像刚才说的用小刀杀牛。他们一般给学生指一个方向。跟这种人交流是有很大收获的，他们往往能预见一些重要的东西。比如刚才我说我 03 年做的那个课题，一开始写得很复杂。我们要做分类，找到一些不变量，这些不变量可以用来区分方程，也就是说，如果两个方程对应的不变量不同，则方程就不是同一类的。Dubrovin 对我们这个不变量的选取不满意。我当时是通过系数和标准形来定义的不变量。他说：“你这个东西能不能给个公式啊？”后来果然找到了公式，这篇文章就变得很好看了。今后要计算这个不变量，只需直接代入计算就可以了。他能让我们把这些很繁的东西变得非常精致。我们在年轻的时候，尽量透支一下自己的生命。我们应该尽量刻苦，年轻的时候多做一些麻烦的事情，会有一些收获。

（我们应该做什么方向，什么方向比较热门）

这个问题太大了。比如你做数论，很可能完全无法下手。但是你如果做一个问题做不出来了，你可以做一些相关的东西。你要知道自己的终极目标并朝着目标努力，当然途中也有一些小的目标。一个最简单的办法是看一下这个领域最牛的人在做什么。比如你看每年的菲尔兹奖得主，他们的文章很可能看不懂。数学部分不看，可以看看语言，了解一下问题的背景，还可以查一下文章的引用。最牛的文章引用的东西通常是有代表性的。通过这些，你可以了解你要做的方向。这里有一个潜在的问题，就是要不要去做最热门的东西。好的数学总是相互联系的，有可能一个领域新作出的结果很有价值，于是就吸引了其他人，然后这个领域发展成很热门的。数学中有很多问题，每个人都想去解决它，比如说黎曼猜想。但是这些问题往往是不能做的。我们要考虑一些现实的东西，比如我做这个东西，我花几年时间应该能把它做好，但这样我就不能发文章，赚不了什么钱，活下去比较难。所以为了活下去，可能需要做一些让自己很失望的工作。做费马猜想的 Wiles 做了 7 年，七年间也发一些小文章。这种方式对大家来说不是很可取。有一种评论认为，Wiles 当年要早点把最初的想法公布出来，参加会议，可能一两年内就有人解决了。但他不公布，偷偷摸摸地做，就花了比较长的时间，虽然最后这个结果是归功于他了。有一个很鲜明的对比，也是一个做数论的人叫做 N. Katz。如果去看费马定理的历史，也会看到这个人。这个人大家对他评价就很高。别人有了问题都和他讨论，他帮人解决了之后也不独占。这就是两种不同的习惯。

（怎么看待本科生的论文，本科生搞的研究）

（先讲了 Dantzig 和单纯形法的故事。）在新生领域可能会有年轻人能解决一些问题，但在成熟的领域，这样的事情不太可能发生。比如 Dantzig 十个问题就能解决好几个，但在数论中 100 个问题能解决一个就很不错了。和其他的工作不同，数学最大的特点是你几乎每天都在失败，很可能不知道要怎么去做一个问

题，总是做不出来。然后突然有一天成功了，就获得了最终的成功。而其他的工作是每天都在成功，事务性的工作，至少每天都有一点进展。但是一旦出了错，往往很严重。所以要培养一种感觉，先要失败，才能成功。你不要相信自己是很幸运的，要相信自己是不幸的。不会做了，出去走走，回来再做。首先你不要对自己失望，其次你要坚定地去做。我们学的东西越多，就越觉得这些东西不是人类能想出来的。但是做起研究来才明白，东西不是一下想到的，而是试了这样不行，那样不行，失败了多次，最终才找到一个可行的办法。数学研究就是这样，和学习很不一样的。

（你们在开学期间要呆在学校上课，放假还得到处跑去参加学术会议，是不是没有什么休闲娱乐的时间啊？）

你不一定每天工作十四个小时，一般工作累了就休息，看看电影，看看书，然后休息累了就继续工作（大家笑）。计划真的很难订。因为数学这个东西是不知道结果的。你只能说，我要做一些东西，但是做多长时间是不知道的，不像学习可以做计划。有时候你算一个东西，算到一半你计划的时间到了，就能放下来么？这是很难受的事情。比如我经常半夜做一个东西，算到十一点多，没有算完。去睡觉吗？睡不着的。必须把它算完才能安心。有的时候，算完发现天已经亮了。

（中国的数学水平在世界上怎么样？）

这个要看哪一方面了。（学生说：几何分析吧）中国的微分几何还是很强的。你们有机会的话，还是该到国外去学习一下，交流是很重要的。不过清华现在有了数学中心，会请国外的大师过来讲课，这也很好。我读书的时候，当时 01、02 年，数学系跟法国数论界的关系很好，联系得比较多，从法国请了一批大神级的人物来讲课，借这个机会培养出了一批人才。像获奖的田一超、郑维喆等等。清华数学中心就是要进一步发展当年那种模式，搞一些长期的课程，让学生们受益。利用这种合作的模式，培养出很多年轻的人才。这个课程一定要长，至少要一个月，三个月，或者半年一年，开出一学期长的课程。这样学生和老师能够有时间熟悉，是一件很好的事。

（那英语的问题老师有什么建议吗）

英语很重要，需要花点时间去学。你口语要说得好一点，听力也很重要。因为数学重在交流。看别人文章能够知道一些东西，但是很慢，你要是把作者抓过来讨论一番，就会很快知道这篇文章到底是在干什么，这是很有好处的一件事情。所以英语是必须要学好的。

（毛天一，余成龙整理）

征稿启事

《荷思》是清华大学数学系学生自主创办的数学学术刊物，面向各个院系中对数学感兴趣的本科生及研究生。

本刊欢迎全校师生任何与数学有关的投稿，无论是长篇的论述，还是精彩的小品，抑或学习/教学的心得、习题的妙解。在原作者的允许下，推荐他人的作品也同样欢迎。

为了编辑方便，建议投稿者能够提供电子版，并且采用 Word 或 LaTeX 排版。来稿请注明作者，联系方式。

本刊创办以来，一直得到老师和同学们的大力帮助。在基科 9 字班“数学分析”课程上，卢旭光老师提出捐出本门课程的 730 元讲义费来支持我们的刊物。这已经是卢老师第二次为我们的刊物捐献讲义费。在此，我们对卢老师和“数学分析”课程的全体同学致以最真的谢意！也对一直以来所有热心支持本刊的读者表示由衷的感谢！

投稿请寄： THUmath@googlegroups.com

欢迎访问荷思网站：http://mathmu.cn/hesi

《荷思》编辑部

2010-5

联系本刊：THUmath@googlegroups.com

北京市 海淀区 清华大学 紫荆 9 号楼 501A

010-515-31892

我们所做的事可能是渺小的，
但它具有某些永恒的性质。

——［英］G.H.哈代

清华大学
数学科学系