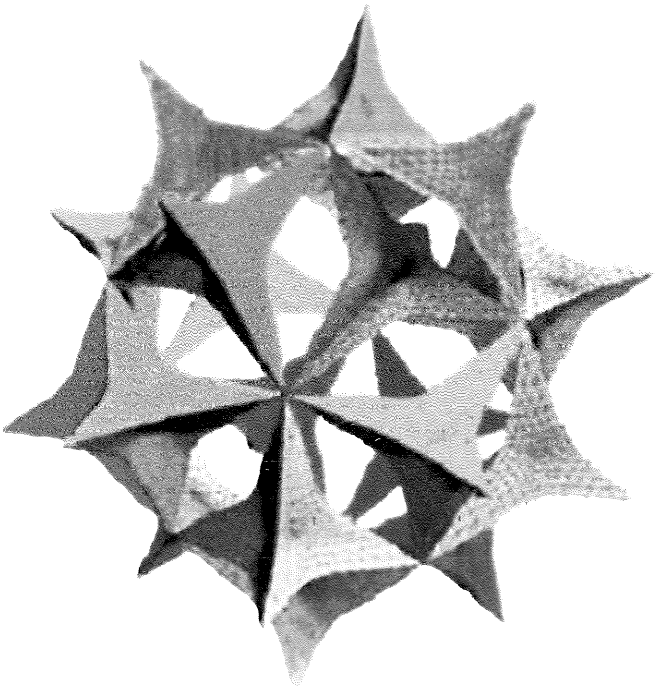
蛙鸣

第62期



中国科学技术大学数学系　编

2008年5月

刊首寄语

是学长对学弟的鼓励，
是同学之间的相互勉励，
是我们一步步成长的脚印，
是科大数学系精神火炬的传递。
这就是当一本接一本的《蛙鸣》
出版时我们所看到的。

如今，
又一本新的《蛙鸣》向你呈现。
不知你是否心怀期盼。
或许我们现在拥有的，
只是一些稚嫩的语言。
但是正是通过这些实践，
我们才最终能够写出美丽的诗篇。
我们自豪我们这样幼稚而勇敢地走过，
并坚信所有曾付出的
努力都是值得的。

目 录

刊 首 寄 语 1

大师睿语

数学的创造 …… 庞加莱(H.Poincaré) 3

研究探讨

Deligne's Integrality Theorem …… 0401 薛航 8

超现实数 …… 0401 洪继展 13

用内点法通过马氏决策过程解决汉密尔顿圈问题 …… 0401 王岚晖 18

蛙声一片

有关Cayley-Hamilton定理的证明 …… 0401 张伟哲 29

非线性优化中的免梯度算法 …… 0401 蔡亦青 32

关于量子化的两个定理的证明 …… 0401 官永辉, 洪继展, 李盈伟, 王常亮 38

Sites and Group Cohomology …… 0401 薛航 48

新生园地

线性空间与Hilbert空间维数定义的合理性 …… 0501 姚成建 56

凸函数两种定义之间的关系 …… 0701 刘博睿, 申国桢, 袁嘉辰 58

数学文摘

Medal and Awards of ICCM and the Background …… 62

数学的创造

庞加莱(H.Poincaré)

数学是怎样造出来的？能够做出数学命题和系统的头脑是怎样的头脑？几何学家或代数学家的智力活动比之音乐家、诗人、画家和棋手又怎么样？在数学的创造中哪些是关键因素？是直觉吗？是敏锐的时空感吗？是计算机似的精确性吗？是特强的记忆力吗？还是追随复杂的逻辑次序时可敬畏的技巧？或者是极高度的用心集中吗？

下面的短文是20世纪初年在巴黎心理学会的一次讲演，是努力描绘数学家的头脑里进行些什么的最著名的作品。它的作者，昂利.庞加莱($HenriPoincaré$)是政治家雷蒙.庞加莱的堂兄，特别适于做这项工作。他是在一切时代中都居于最前列的数学家之一，是无可比拟的分析学家和数学物理学家。庞加莱也以科学哲学的杰出而流畅的讲解者著称。这些作品对于科学家说来是头等重要的专业著作，而对于爱思考的外行在很大程度上又是可以接受和理解的。 —— 纽曼 序

数学的创造的创生过程是一个心理学家应该特别有兴趣的问题。它是这样一种活动：人的心智在此间取自外界的绝少，而是或者好象是只靠它自身起作用，也只作用于其自身。所以我们希望通过研究几何思维的过程就能达到人的心智的最本质的东西。

如果我们还不习惯于数学思维，这是应该使我们吃惊，或者将要使我们吃惊的第一件事，怎么还会有不懂数学的人呢？如果数学只援引所有心智正常的人都接受的逻辑规则；如果它的根据只是基于所有人所共有的原则，谁要想否认这些规则，想必是疯子，那么怎么还会有这么多人觉得数学那么难对付呢？

说是并非凡人均会发明，当非怪事。要说并非人人都能记住自己学过的证明，也还说得过去。但要说在对数学推理已作了解释后，仍然并非人人能懂，这一点想起来确是怪事。然而，很难跟得上数学推理的人确实是多数：这无法否认，根据中学教师的经验，这一点确实无法否认。

进一步还要问：数学中怎么还会有错误？一个清醒的头脑本不会出逻辑上的错误，然而还会有那么多很聪明的头脑，对日常生活中那种简短的推理本不会失足，却不能无误的跟上或重复校长的数学证明，而这些证明归根结底不过是将简单推理汇集起来而已。这些简短推理又与他们游刃有余的推理完全类似。我们还需要多问一下，数学家自己也不是不犯错误……

至于我自己，我必须承认，我绝做不到加法不出错误。……我的记忆里并不差，但却不足以成为一个好棋手。那么为什么一些困难的数学推理难不倒我，却使绝大多数棋手无能为力？显然这是因为数学是由一个总的推理进程来指导的。数学证明并不是三段论法的简单的堆积，而是三段论法按一定次序放置。放置他们的次序比这些三段论法本身重要得多。如果我对这种次序有一种感觉，或者说是知觉，能够一下子看到推理的整体，我就不必害怕忘记

了某一个三段论法，因为他们每一个都放在某一种排列的指定位置上，而这完全不必自己费力去记忆。

我们知道并不是每个人都具有这种对数学次序的感觉或直觉，而正是这才使我们能预见到深藏着的和谐与关系。有些人既没有这种很难说清楚的感觉，又没有超常的记忆力与集中的注意力，所以就决不能懂得高等的数学。这是多数人。另一些人只是稍微有这种感觉，但是赋有不寻常的记忆力与注意力，他们也能一步步地记得推理的细节；她们能够懂得数学，有时也能应用，但是他们不能创造。最后，另外的人或多或少有上面说到的种种特殊的直觉，他们不但能动的懂的数学，虽然他们的记忆力并不超常，他们却能够成为创造者，而且根据他们的这种直觉发展的程度而能有或大或小的发明。

事实上数学创造究竟是怎么回事？它不在于把已知的数学试题拼拼凑凑。谁都会做这件事，但是这样的组合为数无穷，其中绝大多数毫无意义。创造恰好在于不做无用的组合而只作有用的，但其为数甚少。发明就是鉴别和选择。

挖得更深入一些看一看数学家的灵魂里究竟发生了什么事，现在是时候了。我相信，为此最好是追溯我的回忆。但我将限于说明我是怎样写出我关于福克斯（L. Fuchs）函数的第一篇论文的。我要请读者原谅：我会用一些专业名词，但是读者不必害怕，因为他不必去弄懂它们。例如，我会说到，我在什么情景下找到了一个什么定理的证明。这个定理有一个许多人都不熟悉的怪名字，但那并不重要；一个心理学家感兴趣的并不是定理本身而是情景。

整整15天我努力去证明不会有任何函数像我后来称之为福克斯函数的那样；我每天坐在书桌前，一两个钟头，实验着各种各样的组合而毫无结果。一天晚上，我违反了平日的习惯喝了一杯浓咖啡而不能入睡。各种想法纷至沓来；我觉得这些想法在碰撞，终于一对对的联接起来，可以说是成了一个稳定的组合。一夜下来我已经确定了一类福克斯函数的存在性，即那些源于超几何级数的福克斯函数。清晨，我立刻开始把这些结果写出来，而且仅用了几个小时。

接着我想用两个级数的商来表示这些函数；这种想法完全是有意识的思考，引导我的是它们与椭圆函数的类比。我问自己，如果这样的级数存在，他们应当具有什么特性，于是轻而易举我就成功地获得了我后来所称的0 -福克斯函数。

正好这时，我离开驻地卡昂（Caen）去参加矿业学院主办的一次地质考察旅行。我们不断从一个目的地到另一个目的地，数学完全让我置之脑后了。到达库斯坦（Coutances）后,有一次我们又要乘大汽车出发到什么地方去，当我的一只脚踏上汽车的那一瞬间，一种想法突然涌入脑际：我用来定义福克斯函数的那些变换不就是非欧几何的变换吗？而这种想法事先在我头脑中并没有什么征兆。当时我没有足够的时间证实这种想法，只感到它是完全正确的，上了汽车我很快加入了人们已经开始的谈话，结束旅行回到卡昂我抓紧时间证实了这一结果，总算了结了一桩心愿。

后来，我的兴趣又转向某些数论问题，表面上看，与我从前的研究方向无关而且没有什么重要进展，失败使我大为扫兴，我决定到海边休整几天不去想它。一天，正当我在一处峭壁前散步时，突然一个念头穿入脑海。它同上次一样，鲜明、突兀，我一瞬间就断定了：不定三元二次型的算术变换就是非欧几何的变换。

返回卡昂我推敲这一结果并推导出一系列有关结论。二次型的例子预示着，一定还存在新的福克斯群，它们不同于那些与超几何级数对应的福克斯群；我知道我可以把0 -福克斯级数理论用于它们。于是我果真得出存在不同于我当时仅知的与超几何级数对应的福克斯函

数。接着我自然着手构造所有这类函数。我开始向它们发起系统攻势，接二连三的攻克了所有外围工事，但仍有一处攻不下来，然而只有攻克它才能占领整个阵地。所有努力似乎只是更清晰地向我展示确实有分量的困难。所有工作完全都是在一种有意识状态下进行的。

（图一）

（图二）

（图三）

（图四）

创造的过程，如庞加莱所描述的，可能是在有意识的工作时期之间下意识地进行的。图一是研究一个问题；图二是庞加莱去做地质旅行。当他踏上大汽车时，得到解法（图三），后来才写出来（图四）。

于是，我为了一些兵役中的任务去了百莱里山；所以我做的事完全不同了。有一天，当我在街上走时，那个让我停顿不前的困难的答案突然浮现在我眼前，我并没有马上就深入下去，一直等到事情办完才重新拿起这个问题。所有材料都有了，只需要编排起来就行了。这样我就一下子最后写出了我关于福克斯函数的论文而毫无困难。

我只讲这一个例子；再多讲也没有用了……

首先最令人激动的就是出现了顿悟，他是此前进行了长时间的不自觉工作的明显的征兆。在数学发明中这种不自觉工作的作用是无可争辩的。不自觉的工作在别的不太明显的情况下，其实也可以找到它的踪迹。时常当一个人做一个困难问题时，第一次着手什么东西也做不出来。然后这个人可以休息一下，或长或短，然后再坐下来做这件事。开始半小时和以前完全一样，什么也找不到，可是一个决定性的想法会突然涌上心头。……

关于这种无意识的工作的条件还有一点要说明：如果在它之前与在它之后都进行了一段时间有意识的工作，这种无意识的工作就可能有，而且肯定会有结果。除非是做了好几天自觉地努力，这种突然的灵感是不会发生的（上面的例子就证明了这一点），它曾经看来绝无成果，似乎由此什么好东西也得不到，所走的路看来完全是山穷水尽。然后，这些努力绝不如人所想的毫无成果；这些努力启动了无意识工作的机器，否则这个机器就不会运转，不会产生出任何东西。……

其实就是这样;现在讲一讲它让我有了什么想法，无意识的自我，或者成为“下意识的自我”（Subliminal Self），在数学创造中起重要作用。但是通常都把下意识自我看成完全自发

的。我们已经看到，数学工作并非简单机械的，再完美的机器也做不了数学工作。它不只是应用某些规则的问题，不只是按照一些固定的规则作出绝大多数可能的组合的问题。这样得出的组合将是多得不可胜数，毫无用处而又繁琐冗长。发明者真正的工作是从中选择以消除无用的组合，或干脆就避免去作出这种组合，指导这种选择的规则将是极为微妙的。几乎不可能准确的陈述他们；他们只可意会不可言传，在这种情况下，怎么能把它们想象为可以机械地使用的筛子呢？

现在有第一个假设：下意识的自我并不低于真正的自我；它并不纯粹是自发的；它能作鉴别；它机智而且细致；它知道怎样选择，怎样预卜。我说的是什么？它比有意识的自我更会预卜，因为在后者失败了的地方它却成功了。总而言之，难道下意识的自我不是高于有意识的自我吗？你们都能理解到这个问题的全部重要性。……

难道这个肯定的回答就是我前面讲的事是必然给出的吗？我承认，就我而言，我并不想接受它。那么，重新检查一下这些事实，看一看它们是否与另一种解释必不相容。

可以肯定，在经过一个相当长的自觉工作以后才以顿悟形式出现在我们心灵中的那些组合，一般都是有用的，有成果的。从最初的印象似乎正会得出这个结论。由此能得出什么结论，是下意识的自我通过一种很微妙的直觉而预卜到这些组合会是有用的，从而只得到这些组合呢？或者它也得出了许多没有意义的组合，而任其仍然停留在下意识中呢？

按第二种看法，由于下意识自我的自发作用得到了各种各样的组合，而只有其中有意义的才进入到有意识的范围之中。而这仍然是很神秘的。无意识活动的成千产物中只有一些被召跨过了门槛，而其余的仍在其下，这里的原因是什么呢？授予这个特权的，是否简单就是机遇呢？显然不是；例如，对于我们的感官的所有刺激中，只有最强的才引起我们注意，除非还有其它原因把这些刺激引向我们。更一般地说，那些很特殊的能够变为有意识的无意识现象，必是那些直接或间接地最深刻地影响我们情感上的敏感性的。

看到恰好是情感的敏感性应数学证明之召而来，这似乎是令人惊奇的。数学证明本来只对理智才有兴趣。产生这种惊奇是因为忘记了数学的美，数和形的和谐以及几何的优雅。这是所有真正的数学家都知道的真正的美感，它肯定属于情感的敏感性范围。

那么，是什么样的数学实体赋予了这种美和优雅的性格，而且能在我们中间引起一种美的感情？这就是那些起元素是和谐的安排着的东西，是我们的心灵在观察到细节的同时，还能不费力地掌握整体的东西。这种和谐既能满足我们审美的需要，又能有助于我们的心智，使能持续，使能得到引导。同时，在一个有秩序的整体放置于我们眼下时，就使我们能预见到一个数学规律性，……所以，正是这种特殊的对于美的敏感性起着我上面说过的筛子的作用，这足以解释缺少这种美的敏感性的人永不会成为一个创造者。

然而并非所有的困难就此消失。意识本身受到很紧的限制，而对于下意识的自我，我们却不知道它有什么界限，这就是为什么我们可以不太勉强的假设：下意识的自我在短短一段时间里可以做出的不同的组合为数比一个有意识的人一生做出的还要多，然而这种界限由又确实是存在的。难道很可能它确可做出一切可能的组合，为数之多让想象力也吓住？尽管如此，似乎必需这样想，因为如果它只能做出这些组合的一小部分，而且是随机做出的，则它从种种组合中选出真的好的、我们应该选择的组合，其机会也就很小了。

或者我们可以在预先的时期中的自觉工作中寻找解释，在有成果的无意识的工作之前必定有这样的时期。让我们把这些组合中将要出现的元素想象为伊壁鸠鲁的带钩的原子。当心灵完全休止时，这些原子也完全不动，可以说，好像是挂在墙上一样。……

另一方面，在一个看起来休止而实际上已有下意识工作之时，它们中有一些就从墙上脱落下来开始活动了。它们在所居的空间中(我几乎想说是在房间里）各个方向如闪电飞逝，好像一圈小蚊虫，或者换一个更有学术味的比喻，好像气体运动论的分子一样。它们的互相撞击就可能产生新的组合。

预先的自觉工作起什么作用？很显然是在驱动一些原子，把它们从墙上摘下来，让它们振动。我们想，我们并没有作什么有益的事，因为我们已经把它们用一千种不同方式动起来，努力去组合它们，而终未找到满意的组合。但是我们的意志又去动摇它们以后，这些原子就再也不会回到原始的平静状态，它们自由地继续舞蹈。

现在，我们的意志并不是随机地选择它们；它追求一个完全确定的目标。所以，驱动起来的原子并不是随便什么原子，而是我们可以合理的期望由此可能得到所求解答的原子，这些动起来了的原子就经历着碰撞，或者它们彼此碰撞而进入组合，或者它们在行程中撞上了其它静止着的原子。我要再次请求原谅，我的比喻是很粗糙的，但是我确实不知道有没有别的方法使我的思想得到理解。

然而说不定是这样的，只有那种其中至少有一个元素是我们的意志自由选择的那种组合才是可能形成的组合。现在很显然正是这样找到的组合是我所说的好的组合。可能这是在我们原来提出的假说中减少悖理之处的方法。……

我最后还要指出：当我在上面讲自己的个人看法时，我讲的只是我自己工作中的一个令人激动的夜晚发生的事。这种情况是常有的，脑的这种非正常的活动并不一定是如我上面提到的那样由某种物理刺激产生的。在这种情况下，似乎是一个人处于无意识工作之中，他已经部分地觉察到过度兴奋的意识的活动，却不能完全改变其本质。于是我们模糊地理解了这两种机制的不同，或者如果你愿意的话，也可以说是两个自我的工作方法的不同。我迄今所能做的心理学的观察，在我看来，与我提出的观点在要点上是相符的。

这些观点肯定还需要[证实]，因为尽管作了这一切，它们仍然是非常假设性的。这个问题的趣味性如此之大，所以我不揣冒昧地提给了读者们。

Deligne's Integrality Theorem

04001 Xue Hang *†

Abstract

This is the seminar talk I gave at the Morningside Center of Mathematics in Beijing on Oct.19th 2007 on Deligne's Integrality Theorem ([4] Exposé XXI Theorem 5.2.2).

§1 Notations and Statements of the Main Theorem

1.1 Let p be a prime number, q a power of p, k a finite field with q elements and l a prime distinct from p. We fix an algebraic closure $\overline{\mathbb{Q}}_l$ of $\mathbb{Q}_l$. An endomorphism of a finite dimensional $\overline{\mathbb{Q}}_l$-vector space is called integral if all its eigenvalues are algebraic integers.

1.2 Let X be a scheme of finite type over k and $\mathcal{G}$ be a $\overline{\mathbb{Q}}_l$-sheaf on X. In this talk, we keep the notations and conventions of [6]. Take a closed point x of X, and let $\overline{x}$ be the geometric point of X localized at x. We define the geometric Frobenius $F_x : a \mapsto a^{1/|k(x)|}$, where $|k(x)|$ is the order of residue field $k(x)$ at x. For a $\overline{\mathbb{Q}}_l$–sheaf $\mathcal{G}$, F_x acts on the stalk $\mathcal{G}_{\overline{x}}$ of $\mathcal{G}$ at $\overline{x}$.

Definition 1.3. *One says that $\mathcal{G}$ is integral if for each closed point x of X, and each geometric point $\overline{x}$ localized at x, F_x is an integral endomorphism of $\mathcal{G}_{\overline{x}}$. If $\mathcal{G}$ is an object in $D_c^b(X, \overline{\mathbb{Q}}_l)$, then $\mathcal{G}$ is called integral if $\mathscr{H}^i\mathcal{G}$ is integral for all i.*

Remark 1.4. For a sheaf $\mathcal{G}$ on X, the property of being integral is stable under taking sub-object, quotient-object, and extension. If we denote the full subcategory of $D_c^b(X, \overline{\mathbb{Q}}_l)$ consisting of all integral objects by $D_c^b(X, \overline{\mathbb{Q}}_l)_{int}$, then $D_c^b(X, \overline{\mathbb{Q}}_l)_{int}$ is a triangulated subcategory of $D_c^b(X, \overline{\mathbb{Q}}_l)$.

The main theorem of this talk is

Theorem 1.5 (Deligne, [4] Exposé XXI). *Let X and Y be schemes of finite type and separated over k, and $f : X \longrightarrow Y$ be a k-morphism. Then $Rf_!$ sends $D_c^b(X, \overline{\mathbb{Q}}_l)_{int}$ to $D_c^b(Y, \overline{\mathbb{Q}}_l)_{int}$.*

*Address: Department of Mathematics, University of Science and Technology of China, Hefei, 230026
†E-mail: hhh8511@mail.ustc.edu.cn

§2 Proof of the Main Theorem

2.1 If X is a scheme separated of finite type over k, $a_X : X \longrightarrow \mathrm{Spec}(k)$ being the structure morphism, and $\mathcal{G}$ is a $\overline{\mathbb{Q}}_l$-sheaf on X, one then denotes the sheaves $R^i a_{X*}\mathcal{G}$ and $R^i a_{X!}\mathcal{G}$ on $\mathrm{Spec}(k)$ by $\mathcal{H}^i(X,\mathcal{G})$ and $\mathcal{H}^i_c(X,\mathcal{G})$ respectively. If $\overline{k}$ is an algebraic closure of k and $\overline{X} = X \times_{\mathrm{Spec}(k)} \mathrm{Spec}(\overline{k})$, then the stalk of $\mathcal{H}^i(X,\mathcal{G})$ at $\overline{k}$ gets identified with $H^i(\overline{X},\overline{\mathcal{G}})$, where $\overline{\mathcal{G}}$ is the inverse image of $\mathcal{G}$ under the morphism $\overline{X} \longrightarrow X$. The same holds for cohomology with compact support.

In order to prove theorem 1.5, we need the following lemma.

Lemma 2.2. *Suppose X/k is a scheme of finite type over k of dimension $\leqslant n$, and $\mathcal{G}$ is a $\overline{\mathbb{Q}}_l$-sheaf on X.*

(1). There exists a closed subset X' of X of dimension 0, such that the natural morphism $\mathcal{H}^0(X,\mathcal{G}) \longrightarrow \mathcal{H}^0(X',\mathcal{G})$ is injective;

(2). If X is separated, and U an open set of X with complement Z of dimension $< n$, then there exists a closed subset U' of U of dimension 0 and a surjective morphism $\mathcal{H}^0(U',\mathcal{G})(-n) \longrightarrow \mathcal{H}^{2n}_c(X,\mathcal{G})$.

Proof. Let us prove (1). If X is geometrically connected and $\mathcal{G}$ is lisse, then we can take X' to be just one closed point of X. In fact we have the following cartesian diagrams:

$$\begin{tikzcd}
x \arrow[d, "i"'] \arrow[dr, phantom, "\square"] & y \arrow[l] \arrow[d] \\
X \arrow[d, "a"'] \arrow[dr, phantom, "\square"] & \overline{X} \arrow[l] \arrow[d] \\
\mathrm{Spec}(k) & \mathrm{Spec}\,\overline{k} \arrow[l]
\end{tikzcd}$$

where y is a bunch of points $y = \coprod y_i$. We have to show $(R^0 a_* \mathcal{G})_{\overline{k}} \longrightarrow (R^0(ai)_*\mathcal{G})_{\overline{k}}$ is injective.

Consider the composition

$$H^0(\overline{X},\mathcal{G}) \longrightarrow H^0(y,\mathcal{G}) \longrightarrow H^0(y_i,\mathcal{G})$$

Since y_i is a geometric point and X is connected, $H^0(\overline{X},\mathcal{G}) \simeq \mathcal{G}_{y_i}^{\pi(X,y_i)}$, thus the morphism $H^0(\overline{X},\mathcal{G}) \longrightarrow H^0(y_i,\mathcal{G})$ is injective and so is $H^0(\overline{X},\mathcal{G}) \longrightarrow H^0(y,\mathcal{G})$.

For the second assertion, first note that we can replace X by X_{red} and shrink U to ensure U is smooth and $\mathcal{G}|U$ is lisse. Then since $\dim Z < n$, we have an exact sequence

$$0 = \mathcal{H}^{2n-1}_c(Z,\mathcal{G}) \longrightarrow \mathcal{H}^{2n}_c(X,\mathcal{G}) \longrightarrow \mathcal{H}^{2n}_c(U,\mathcal{G}) \longrightarrow \mathcal{H}^{2n}_c(Z,\mathcal{G}) = 0$$

The morphism in the middle is thus an isomorphism. By Poincaré duality, we have

$$\mathcal{H}_c^{2n}(U,\mathcal{G}) \simeq \mathcal{H}^0(U,\mathcal{G}^\vee(n))^\vee$$

By (1), we have an open subset of U and an injective morphism $\mathcal{H}^0(U,\mathcal{G}^\vee(n)) \hookrightarrow \mathcal{H}^0(U',\mathcal{G}^\vee(n))$, hence a surjective morphism $\mathcal{H}^0(U',\mathcal{G}^\vee(n))^\vee \twoheadrightarrow \mathcal{H}^0(U,\mathcal{G}^\vee(n))^\vee$ But the left hand side is just $\mathcal{H}^0(U',\mathcal{G})(-n)$. $\square$

2.3 Now let us prove theorem 1.5.

We may assume $\mathcal{G}$ is a $\overline{\mathbb{Q}}_l$-sheaf on X. Moreover we may assume $Y = \mathrm{Spec}(k)$. Indeed, once the theorem is verified under this assumption, we can apply it to each closed fiber of f and the theorem follows from the *Proper Base Change Theorem.*

We proceed by induction on $n = \dim X$

The case $n = 0$ is trivial, and the case that $i = 0$ or $i = 2n$ follows from lemma 2.2. For the case $n = 1$, we consider the zeta function of X. We have

$$Z(X,\mathcal{G},t) = \prod_{x\in|X|} \det(1 - F_x t^{\deg(x)}, \mathcal{G}_{\overline{x}})^{-1} \in \overline{\mathbb{Q}}_l[[t]]$$

where $|X|$ is the closed points of X. By Grothendieck's theorem ([3] XV 3.2), we know that

$$Z(X,\mathcal{G},t) = \prod_i \det(1 - Ft, \mathcal{H}_c^i(X,\mathcal{G}))^{(-1)^{i+1}}$$

where F is the geometric Frobenius of $\mathrm{Spec}(k)$. Since $\dim X = 1$, $\mathcal{H}_c^i(X,\mathcal{G})$ vanishes for $i > 2$ and the theorem has already been proved for $i = 0$ and $i = 2$. By hypothesis the coefficients of the power series $Z(X,\mathcal{G},t)$ are algebraic integers, so are those of the product

$$Z(X,\mathcal{G},t)\cdot\det(1 - Ft, \mathcal{H}_c^0(X,\mathcal{G}))\cdot\det(1 - Ft, \mathcal{H}_c^2(X,\mathcal{G}))$$

However, this product is nothing but $\det(1-Ft,\mathcal{H}_c^1(X,\mathcal{G}))$, which yields the assertion for $i = 1$.

Assume the theorem holds for $\dim X < n$ and we prove it for $\dim X = n$. Suppose U is an affine open subset of X whose complement Z is of dimension $< n$. By Noether Normalization Lemma, there is a finite morphism $u : U \longrightarrow \mathbb{A}_k^n$. The following diagram

$$\begin{tikzcd}
U \arrow[r, "u"] \arrow[rdddd] & \mathbb{A}_k^n \arrow[d] \\
& \mathbb{A}_k^{n-1} \arrow[d] \\
& \vdots \arrow[d] \\
& \\
& \mathbb{A}_k^1
\end{tikzcd}$$

where the vertical morphisms are the projections from $\mathbb{A}_k^r$ to $\mathbb{A}_k^{r-1}$, gives a morphism $g : U \longrightarrow W(= \mathbb{A}_k^r)$ whose target W is a curve and whose fibers are all of dimension $< n$, then we get a diagram

$$\begin{tikzcd}
U \arrow[r, "j"] \arrow[d, "g"] \arrow[dd, bend right, "a_U"'] & X & Z \arrow[l, "i"'] \\
W \arrow[d, "a_W"] & & \\
\mathrm{Spec}(k) & &
\end{tikzcd}$$

From the exact sequence

$$0 \longrightarrow j_! j^* \mathcal{G} \longrightarrow \mathcal{G} \longrightarrow i_* i^* \mathcal{G} \longrightarrow 0$$

we have an exact sequence

$$\mathcal{H}_c^i(U, \mathcal{G}) \longrightarrow \mathcal{H}_c^i(X, \mathcal{G}) \longrightarrow \mathcal{H}_c^i(Z, \mathcal{G})$$

So it suffices to show that $\mathcal{H}_c^i(U, \mathcal{G})$ is integral for $\mathcal{G}$ integral on U.

Applying the inductive hypothesis to each closed fiber of g, one concludes that $Rg_!\mathcal{G}$ is integral. Then $Ra_{W!}Rg_!\mathcal{G}$ is integral by the case $n = 1$. But $Ra_{W!}Rg_!\mathcal{G} = R(a_W g)_!\mathcal{G}$, which completes the proof.

§3 Complement

In [4] Exposé XXI, Deligne shows the following refinement of theorem 1.5.

Theorem 3.1 (Deligne [4] XXI Theorem 5.2.2). *Suppose X/k is a scheme of finite type and separated over k and $\dim X = n$. If $\mathcal{G}$ is a $\overline{\mathbb{Q}}_l$ integral sheaf on X, then both F and $q^{n-i}F$ on $\mathcal{H}_c^i(X, \mathcal{G})$ are integral.*

Proof. The fact that F is integral is a special case of the above theorem. For the other assertion, we can mimic the proof above and proceed by induction on $n = \dim X$. First note that $q^{i-n}F$ being a integral endomorphism of $\mathcal{H}_c^i(X, \mathcal{G})$ is equivalent to $\mathcal{H}_c^i(X, \mathcal{G})(n-i)$ being integral. For the case $n = 1$, only the case $i = 0$ needs proof from lemma 2.2. Assume the theorem for $\dim X < n$ and let us prove it for $\dim X = n$. Instead of the derived-category-theoretic argument in the last part of the proof, we apply the Leray Spectral Sequence, which in this case reads

$$E_2^{i,j} = \mathcal{H}_c^i(W, R^j g_! \mathcal{G})(i+j-n) \Longrightarrow E^{i+j} = \mathcal{H}_c^{i+j}(U, \mathcal{G})(i+j-n)$$

Since the fibers of g are all of dimension $< n$, the inductive hypothesis indicates that $R^j g_! \mathcal{G}(n-1-i-j)$ is integral and then the case $n = 1$ ensures that $\mathcal{H}_c^i(W, R^j g_! \mathcal{G})(i+j-n)$ is integral,and so is $\mathcal{H}_c^{i+j}(U, \mathcal{G})(i+j-n)$. □

Acknowledgement. The speaker thanks Prof. L. Illusie, Prof. Y. Ouyang and W. Zheng for their generous help during the preparation of this talk. In particular, the speaker offers special thanks to Prof. Y. Ouyang for his translation of [4] Exposé XXI Appendice into English.

References

[1] [SGA 4] *Théorie des topos et cohomologie étale des schémas*, Séminaire de géométrie algébrique du Bois-Marie 1963-1964, dirigé par M. Artin, A. Grothendieck, J.-L. Verdier, SLN 269, 270, 305, Springer-Verlag, 1972, 1973.

[2] [SGA $4\frac{1}{2}$] *Cohomologie étale*, par P. Deligne, SLN 569, Springer-Verlag, 1977.

[3] [SGA 5] *Cohomologie ℓ-adique et fonctions L*, Séminaire de géométrie algébrique du Bois-Marie 1965-1966, dirigé par A. Grothendieck, SLN 589, Springer-Verlag, 1977.

[4] [SGA 7] *Groupes de monodromie en géométrie algébrique*, Séminaire de géometrie algébraique du Bois-Marie 1967-1969,I, dirigé par A. Grothendieck, II par P. Deligne et N. Katz, SLN 288, 340, Spinger-Verlag, 1972-1973.

[5] *Topics in Algebraic Geometry*, Luc Illusie, lecture notes in Tsinghua University, 2005.

[6] *La Conjecture de Weil*, P. Deligne, Pub. Math. IHES 52 (1980), 137-152.

超现实数

0401　洪继展

§0　闲话少许

1970年左右，剑桥大学的有一个叫作J. H. Conway 的人发现（发明）了一种新的数.这种数的全体并不是一个集合，而是一个类，是一个既包含了实数，又包含了序数的真类（proper class）.他的好友D. E. Knuth 为这种新的数取名叫做"surreal number". Knuth 还专门为此写了一本小说，据说写的是关于两个学生如何发现这种surreal number 的历程.但我没机会去看一看. 我在此将这种数用中文来命名，叫它们是"超现实数". [1] 本文之中，我将对超现实数理论做一个简介。首先先谈谈我在类上建立分析学的想法. 一直以来，我都在想着，如何才能实现任意多个数，特别是任意多个实数的加法.这种加法，当然，应当是要比较有意义的才行. 当我学到公理集合论中的基数的算术时，了解到了任意多个的基数的加法是很容易做的，例如：

$$\sum_{\alpha<\lambda}\alpha=\lambda,\quad \lambda\in\mathbf{Cn}\backslash\mathbb{N}$$

如果能够将实数与基数合并成为一个类，或许能在这个类上实现任意多个实数的加法.更进一步，还有可能实现在这个类上的分析学，亦即可以在这个类上做四则运算，做序列的极限，做函数的积分与微分.还可以有其他的，例如类上的拓扑学，几何学，代数学，大部分数学或许都可以搬到类上面来做. 后来我发现了Conway 的这些工作以及其他一些学者的结果，觉得很有可能在超现实数类上来做这些事情，虽然Conway 本人在他的书中对于将超现实数应用于非标准分析并不抱太大的希望.[2] 实际上，现在有一些学者如N. L. Alling，Antongiulio Fornasiero 对于其上的一些分析学的研究已经颇有成果.

§1　超现实数

对于超现实数，实际上它是有两种定义方式的.我们首先采用的方法有点类似Dedekind分割.

定义1.1. 设L,R是两个超现实数集，且L中不存在$\geq$在R之中的数，[3] 那么(L,R)是一个超现实数，且所有的超现实数都是这样构造的.

在这里，我们将用Conway 本人所采用的符号，即$\{L|R\}$ 来代替(L,R) .有时候，对于$x\equiv\{L|R\}$，我们也将它写成$x\equiv\{x^L|x^R\}$，其中的x^L指的是L中的某个元素，x^R是某个R中的元素.你可以认为这种写法中的x^L,x^R是分别取遍L与R中的元素. 请注意我们这里用了$x\equiv$

[1] 之所以不用"超实数"来命名是因为非标准分析里面有这个名词，避免混淆.

[2] 参见[1]

[3] 去掉这个条件实际得到了更广泛的一种"数"，叫做game

(L,R)来表示x就是有序对(L,R)，与后面定义的"$=$"要相互区别. 例如$x\equiv\{a,b,c,\ldots|d,e,f,\ldots\}$，指的是$x\equiv(L,R)$且这时$L\equiv\{a,b,c,\ldots\}$, $R\equiv\{d,e,f,\ldots\}$.

定义1.2. 定义超现实数的序如下：定义$x\geqslant y$成$(x^R\nleqslant y)\bigwedge(x\nleqslant y^L)$. 定义$x\leqslant y$为$y\geqslant x$. 其中$x\nleqslant y$指的是$\neg(x\leqslant y)$.定义$x=y$为$(x\geqslant y)\bigwedge(y\geqslant x)$. 定义$x>y$为$(x\geqslant y)\bigwedge(y\ngeqslant x)$. 定义$x<y$为$y>x$.

定义1.3. 定义超现实数的加法为：

$$x+y\equiv\{x^L+y,x+y^L|x^R+y,x+y^R\}$$

定义1.4. 定义超现实数的取负操作为：

$$-x\equiv\{-x^R|-x^L\}$$

定义1.5. 超现实数的乘法定义成：

$$xy\equiv\{x^Ly+xy^L-x^Ly^L,x^Ry+xy^R-x^Ry^R|x^Ly+xy^R-x^Ly^R,x^Ry+xy^L-x^Ry^L\}$$

其中的例如$x^Ly+xy^L-x^Ly^L$指的是同一个x^L与同一个y^L.

根据定义1.1我们就可以从空集ϕ出发开始构造超现实数了。令：

$$0\equiv\{\phi|\phi\}$$

我们又经常将其记为$\{|\}$. 由于此时的$0^L,0^R$都是ϕ，故超现实数的定义的条件被满足，因此，0是一个超现实数. 也容易验证$\{0|\},\{|0\}$都是超现实数，[4] 将它们分别记为$1,-1$.

命题1.1. $-1<0<1$.

我们如果令$2\equiv\{1,0|\}$，则有$-2\equiv\{|-1,0\}$；如果令$\frac{1}{2}\equiv\{0|1\}$，则有$-\frac{1}{2}\equiv\{-1|0\}$. 容易验证以下的结论是成立的。

命题1.2.

$$\begin{aligned}
-2&=\{|-1\}=\{|-1,0\}=\{|-1,1\}=\{|-1,0,1\}\\
-1&=\{|0\}=\{|0,1\}\\
-\frac{1}{2}&=\{-1|0\}=\{-1|0,1\}\\
0&=\{|\}=\{-1|\}=\{|1\}=\{-1|1\}\\
\frac{1}{2}&=\{0|1\}=\{-1,0|1\}\\
1&=\{0|\}=\{-1,0|\}\\
2&=\{1|\}=\{0,1|\}=\{-1,1|\}=\{-1,0,1|\}
\end{aligned}$$

且有

$$-2<-1<-\frac{1}{2}<0<\frac{1}{2}<1<2$$

[4] 0在这里作为"数"而非集。我们下面的写法除非特别说明指的都是$x\equiv\{x^R|x^L\}$的写法.

另外

$$\begin{aligned} 1+1&=2\\ \frac{1}{2}+\frac{1}{2}&=1 \end{aligned}$$

下面我们来定义一下在这种超现实数的意义底下的序数：

定义1.6. 序数是形如$\alpha \equiv \{\beta < \alpha|\}$的超现实数，其中的$\beta$也是序数.

同样的，从空集出发，我们得到0是一个序数，然后$1 \equiv \{0|\}$也是一个序数，再往下$2 \equiv \{1,0|\}$也是一个序数，记$3 \equiv \{2,1|\}, 4 \equiv \{3,2,1|\}, \ldots, n \equiv \{n-1,\ldots,1|\}$，则它们都是序数。我们再令

$$\omega \equiv \{0,1,\ldots,n,\ldots|\}$$

则ω 也是一个序数. 如此的构造，实际上与Cantor 的序数构造法是一样的. 接下去，我们还可以有：

$$\omega+1=\{0,1,\ldots,n,\ldots,\omega|\}$$

由此可知，在这个由超现实数全集构成的这样一个类当中，包含了与序数类**On**同构的一个真子类，故而可知超现实数全体构成一个真类，我们记之为**No**. 当然，在**No**当中，我们有更多的运算，例如我们令$x=\{0,1,\ldots|\omega\}$ 那么我们可以验证$x+1=\omega$，此即$x=\omega-1$. 类似地

$$\omega-n=\{0,1,2,\ldots|\omega,\omega-1,\ldots,\omega-(n-1)\}$$

我们再多想一点点，下面这个数应该是什么呢？

$$\{0,1,2,\ldots|\omega,\omega-1,\omega-2,\ldots\}$$

实际上这个是$\omega/2$. 我们还可以对这些无限数做取倒数，开根号等等的运算。

定理1.1. **No** 是一个真类.

我们通常在讨论代数结构的时候，都要求我们的对像是一个集合，在这里，我们不妨把这样的一个限制取消，对于对像是一个真类，但其上的运算仍满足那些抽像结构的性质的，我们在其通常的名称之前加一个“超”。[5] 例如，对于一个真类C，若C上有一个满足结合律的二元运算，我们就称之为“超半群”。类似有“超群”,“超环”等等。我们有以下的

定理1.2. **No** 是一个超序域。

这里比较值得一说的是**No**上的除法：

命题1.3. 对于任意一个正的x 都有表示形式，使得0是一个x^L 且其余的x^L 都是正的。

在命题1.3 下，我们可以找到y，使得$xy=1$：

$$y=\left\{0,\frac{1+(x^R-x)y^L}{x^R},\frac{1+(x^L-x)y^R}{x^L}\middle|\frac{1+(x^L-x)y^L}{x^L},\frac{1+(x^R-x)y^R}{x^R}\right\}$$

[5]之所以不用“类域”这个名称是因为在其他分支有这个名称但意义不同，可参见“类域论”等有关专题.

定义1.7. 称x 是一个无相整数，[6] 如果$x = \{x-1|x+1\}$. 所有的无相整数组成一个真类，用**Oz** 表示.

可以证明，$\mathbf{Oz} \supset \mathbf{On}$ 且**No** 是**Oz** 的超商域.

定义1.8. 说x是实数，指的是存在n使得$-n<x<n$且有

$$x = \{x-(1/n)|x+(1/n)\}_{n>0}$$

当然这里的n是取遍所有的正整数，即小于ω的序数。

这样，我们就在**No** 里面找到了实数的对应。

对于超现实数的构造，还有另外一种方法。考虑以下映射的全体，这些映射是从某个序数（不同的映射可以是不从的序数）到某个二元集的，如$\{+,-\}$. 定义这样的任意的两个映射的$x \le y$，如果x 与y 满足如下的关系之一：

1. $x=y$
2. $x(\gamma)$ 未定义且$y(\gamma)=+$
3. $x(\gamma)=-$ 且$y(\gamma)=+$
4. $x(\gamma)=-$ 且$y(\gamma)$ 未定义

其中$\gamma = \min\{\beta \in \mathbf{On} : x(\beta) \ne y(\beta)\}$. 有一个叫做Elwyn Berlekamp 的人发现了这样一个规则，使得我们够从表示某个实数的一序列的符号串读出它的值，做法是首先我们总能假定这样的一个序列是从“+”开始的，否则我们可以取负一下，使整串符号都改变符号；如果这一串符就是n 个正号，那它就是数n ；否则将第一个负号与其前面最近的正号用括号括起来，括号前的正号的个数表示的十进制数的整数部分的值，括号后面的符号表示的是小数部分的二进制展开，正号表示的是1，而负号则表示0；如果发现这个序列是有限的话，那就在小数部分的二进制展开后面再添加一个1. 举个例子如$++++-+--$ 加入括号成为$+++(+-)+--$，所以它表示的是$3.1001 = 3\frac{9}{16}$. 再如$+--+$ 表示的是$\frac{3}{4}$.

§2 ω-映射与正则形式

定义2.1. 对于任给的一个序数α 我们定义超现实数集合M_α，$x \in M_\alpha$ 当且仅当

$$x^L, x^R \in O_\alpha := \bigcup_{\beta<\alpha} M_\beta$$

再记$N_\alpha = M_\alpha - O_\alpha$.

由于可以证明对于每个超现实数x 都有惟一的一个N_α 使得$x \in N_\alpha$，我们就可以给每个超现实数x 定义一个“生日”，即序数α. 对于两个超现实数x 与y，称x 比y “早”或者“简单”，如果x 的生日比y小。

定义2.2. 两个正数x, y 称为是相当的，[7] 如果存在一个正自然数n 使得$x<ny, y<nx$.

---

[6]其英文为omnific integer.

[7]其英文为commensurate

定义2.3. $\omega^x = \{0, r\omega^{x^L} | r\omega^{x^R}\}_{r\in\mathbb{R}_{>0}}$

定理2.1. 任意一个正数都与某个ω^y 相当.

定理2.2. $\omega^0 = 1, \omega^{-x} = 1/\omega^x, \omega^{x+y} = \omega^x\omega^y$

定理2.3. 对于任何一个超现实数x 我们总能定义x 的正则形式（$= x$）$\sum_{\beta<\alpha}\omega^{r_\beta}r_\beta$，其中$r_\beta$ 均是不为0实数，α 是某个序数，y_β 成为一个递减序列. 不同的x 的正则序列是不同的.

还可以定义形式无穷和$\sum_{y\in\mathbf{No}}\omega^y r_y$，即我们赋给每个$y$ 一个对应的实数r_y，当y 不在某个下降序列$\{y_\beta : \beta < \alpha\}$ 时，r_y 总取为0. 这样上述形式和实际上就是$\sum_{\beta<\alpha}\omega^{y_\beta}r_{y_\beta}$. 然后可以证明，一个超现实数是无相数当且仅当它的正则形式$\sum_{y\in\mathbf{No}}\omega^y r_y$ 当$y < 0$ 时$r_y = 0$ 且$r_0 \in \mathbb{Z}$.

关于超现实数的代数性质以及分析性质的研究现在还在进行当中，有很多学者将赋值论运用到超现实数类上面，获得了很不错的结果，超现实数类与数论方面也有一定的联系. 读者如果对此有兴趣，可参见参考文献。

参考文献

[1] *On Numbers and Games*, J. H. Conway, Academic Press Inc. (London) LTD. 1976

[2] *Recursive Definitions on Surreal Numbers*, Antongiulio Fornasiero, arXiv:math/0612234v1 [math.LO] 9 Dec 2006

[3] *Integration on Surreal Number*, Antongiulio Fornasiero, PhD thesis, University of Edinburgh, 2004.

[4] *An Introduction to Conway's Games and Numbers*, Dierk Schleicher and Michael Stoll, arXiv:math.CO/0410026 v2 30 Sep 2005

[5] 1989, P.M. Cohn's book review on *Foundations of Analysis over Surreal Number Fields* written by N. L. Alling, 1987

用内点法通过马氏决策过程解决汉密尔顿圈问题*

0401 王岚晖

摘 要

我们在一个单摄动马氏决策过程（MDP）中考虑汉密尔顿圈问题（HCP）。我们将看到在频率空间下汉密尔顿圈（HC）对应一个不定二次规划问题的最小值。接着我们可以用近似凸的二次函数取代原二次函数来逼近原问题。最后我们用内点法解决改造过的二次规划问题，再根据其解消去某些弧来简化原图，直至找到HC。

关键字. 汉密尔顿圈（HC），内点法，马氏决策过程，非凸优化。

§1 引言

过图中每点恰好一次的圈称为哈密顿圈。哈密顿圈问题是图论最古老的研究课题之一，是至今未解决的世界难题，在许多领域有着重要应用。

许多经典的寻找HC的方法着眼于防止子圈的形成，我们的方法是在单摄动马氏决策过程中定位汉密尔顿圈。

更确切地，我们这个寻找HC动态的、随机的方法，是假想在有向图G上由函数（决策）f控制行动的运动物体，其中函数f是点集V=V（G）=$\{0,1,\ldots,N\}$到弧集A=A（G）的映射。我们设弧集为马氏链Σ=Σ（G）的状态空间，在每个状态（结点）i，行动空间A（i）：=$\{a|$（i，a）$\in$ A$\}$与从i出发的弧集是一一对应的。

说明. 考虑完全图G5 （有5个结点，无自圈）。自然地，HC c_1：$0\to1\to2\to3\to4\to0$ 与确定性策略f_1：=$\{0,1,2,3,4\}\to\{1,2,3,4,0\}$对应，其中$f_1$（2）=3表示在状态2以概率1选取弧（2，3）。由f_1导出的马氏链是由不可约的0−1 转移矩阵P（f_1）给出的。另一方面，两个子圈的合并：$0\to1\to2\to0$ 和$3\to4\to3$ 与策略f_2：=$\{0,1,2,3,4\}\to\{1,2,0,4,3\}$对应。它的转移矩阵P（$f_2$）（见下图）含有两个不同的遍历类。这个MDP有多链遍历结构，会对分析造成困难。但是，这种多链遍历结构可以在单摄动的帮助下“伪装”起来而不会完全失去原来的性质。比如，我们可以简单的用
Pε (f_2)代替P(f_2):

*这篇文章为Vladimir Ejov , Jerzy Filar and Jacek Gondzio (2003),An Interior Point Heuristic for the Hamiltonian Cycle Problem via Markov Decision Processes,Journal of Global Optimization,Volume 29, Number 3 / 2004.7.的主要内容和扩充

$$P(f_2)=\begin{pmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 \end{pmatrix}$$

和

$$P_{\varepsilon}(f_2)=\begin{pmatrix} 0 & 1-\varepsilon^2 & \varepsilon^2 & \varepsilon^2 & \varepsilon^2 \\ \varepsilon & 0 & 1-\varepsilon & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ \varepsilon & 0 & 0 & 0 & 1-\varepsilon \\ \varepsilon & 0 & 0 & 1-\varepsilon & 0 \end{pmatrix}$$

上诉摄动将P（f_2）变为不可约的马氏链Pε (f_2)，从而改变了遍历结构，是单摄动。

现在我们就能在前面提到过的频率空间上讨论问题了。我们摄动的功效是，对能够描绘出一个HC的策略，访问状态0（结点0）的平均频率能被计算出为

$$x_0=1/d_{1+N}(\varepsilon)$$

（详见命题3）。这可以为频率空间引入一个约束。我们将看见HC与二次规划

$$\min \overrightarrow{x}^T Q \overrightarrow{x}$$

$$\overrightarrow{x}\in X_{\varepsilon} \& x_0=1/d_{1+N}(\varepsilon).$$

的最小值对应。更进一步，使目标函数达到最小值0的策略是确定性策略。并注意到Q是不定的，但有很多特殊的结构。

这篇文章的基本思想是考虑一个“更凸”的目标函数。所以我们构造函数

$$\mathrm{h}_{\alpha}=S(\overrightarrow{x})-\alpha s(\overrightarrow{x})$$

其中，α>0足够小，并且

$$S(\overrightarrow{x})=\sum\nolimits_i\left(\sum\nolimits_a x_{ia}\right)^2$$

$$s(\overrightarrow{x})=\sum\nolimits_i\sum\nolimits_a x_{ia}^2$$

x_{ia} 是在状态i选择弧a的频率。尽管h_{α}还是非凸的，但是导致它非凸的因素$\alpha s(\overrightarrow{x})$可以被足够小的$\alpha$控制。我们将要证明$h_{\alpha}$的全局最小值不是在HC发生，便在HC的一个小邻域内发生。并且，我们有理由相信当α充分小时，对凸二次规划十分有效的内点法同样对上述问题十分有效。

我们用内点法寻找h_{α}的极小值后，可以确定一条弧是否在从同一点发出的所有弧中占“主导”地位。若是，则这条弧便会保留下来被认为有可能在HC中，而从同一点发出的其他弧则会被去掉。这样，就简化了原来的图。在很多情况下，这种简化作用是很显著的。

若没有“主导”的弧，我们就把从这点发出的每一条弧都作为“主导”弧继续计算，直到找到HC为止。

§2　用摄动后的MDP寻找HC

为保证我们的MDP是不可约的，我们对Σ引入摄动的转移矩阵来构造一个新的ε-perturbed process Σ_ε（$0<\varepsilon<1$）我们定义

$$p_\varepsilon[j;i,a]=\begin{cases}1-(N-1)\varepsilon^2 & \text{if i=0 and a=j}\\ \varepsilon^2 & \text{if i=0 and a}\neq\text{j}\\ 1 & \text{if } i>0 \text{ and a=j=0}\\ \varepsilon & \text{if } i>0 \text{ and a}\neq\text{j=0}\\ 1-\varepsilon & \text{if } i>0 \text{ and } a=j>0\\ 0 & \text{in all remaining cases}\end{cases}$$

这里p_ε[j;i,a]表示选择行动（i,a）后由结点i运动到结点j的概率，所以有

$$\sum_{j=0}^{N}p_\varepsilon\left[j;i,a\right]=1,\forall(i,a)\in A$$

注意到随着上述的摄动，对每一对对应于“确定性弧”（i,a）的结点i，a（不等于0），我们用权重分别为ε和（1−ε）的“随机弧”（i,1）和（i,a）来代替它。这种摄动保证了只会生成不可约的由平稳策略定义的马氏链。

一个策略是由（1+N）×（1+N）的随机矩阵f定义的，其中若a∈A（i），则f[i,a]＝访问状态i时选择行动a的概率；否则若a∉A（i），则则f[i,a]＝0．由这些策略构成的空间叫作策略空间，用$\mathbf{F}_S$表示。若f[i,a]∈{0,1}，∀i,a，则它叫做确定性策略，我们可以写作f（i）＝a，由这些策略构成的空间用$\mathbf{F}_D$表示。因为决策人的决定不依赖于时间，所以这些策略可以称为平稳策略。任意一个平稳策略f可以生成一个由（1+N）×（1+N）维的概率转移矩阵Pε (f)，其中Pε (f)定义为

$$P_\varepsilon\left(f\right)\left[i,j\right]:=p_\varepsilon\left(j;i,f\right):=\sum_{a=0}^{N}p_\varepsilon\left[j;i,a\right]f\left[i,a\right]$$

上述式子表明Pε (f)也是一个随机矩阵。先考虑一个线性规划问题（LP1）[1]

$$\max\sum\nolimits_i\sum\nolimits_a r_{ia}x_{ia}$$

s.t.

$$\sum_{i=0}^{n}\sum_{a\in A(i)}(\delta_{ij}-P\left[j;i,a\right])x_{ia}=0,j\in V$$

$$\sum_{i=0}^{n}\sum_{a\in A(i)}x_{ia}=1$$

$$x_{ia}\geq 0,i\in V,a\in A(i)$$

（记为X）我们知道当P不可约时，方程组

$$\overrightarrow{q}P=\overrightarrow{q}$$

[1]这是将HCP与二次规划问题结合起来的铺垫，基本思路是由J.A.Filar and D.Krass, Hamiltonian Cycles and Markov Chains这篇文章得到的启发。

$$\vec{q}\,\vec{1} = \vec{q}$$

有唯一解$\vec{q} = (q_0, \ldots, q_N) > 0$ 为平稳分布。

$$x_{ia}(f) = q_i(f)\, f[i,a]$$

表示访问状态i并选择弧a的平均频率。

$$x_i(f) := \sum_{a \in A(i)} x_{ia}(f) = q_i(f)$$

为访问状态i的平均频率。其中最后一个等式成立是因为

$$\sum\nolimits_{a \in A(i)} f[i,a] = 1$$

再定义

$$\vec{x}_i(f) := \left(x_{ia_1^i}(f), x_{ia_2^i}(f), \ldots, x_{ia_{m_i}^i}(f)\right)^T$$

其中，$a_j^i \in A(i)$，m_i是A（i）中的总弧数。令

$$m = \sum_{i=0}^{N} m_i$$

定义映射

$$M : F_S \to R^m$$

$$M(f) := \vec{x}(f) = \begin{pmatrix} \vec{x_0}(f) \\ \vdots \\ \vec{x_N}(f) \end{pmatrix}$$

单摄动中的ε^2部分保证了$x_i(f) > 0, \forall i$ 从而，我们可以定义M的逆映射

$$M^{-1} : R^m \to F_S$$

$$M^{-1}(\vec{x})[i,a] = f_{\vec{x}}[i,a] := \frac{x_{ia}}{x_i}$$

其中$M^{-1}(\vec{x})[i,a]$ 表示将m维列向量相应地分为N组后，它的第i组的第a个元素。

我们有

定理1 （1）若$\vec{x}$在（LP1）中是最优的，则$f_{\vec{x^0}} = M^{-1}(\vec{x^0})$ 是最优的。反之，若f^0是最优的，则$M(f^0) = x(f^0)$是（LP1）中最优的。

（2）$\forall \vec{x} \in X, M(M^{-1}(\vec{x})) = \vec{x}$，但是，一般说来，$M^{-1}(M(f)) \neq f$。

（3）若L（S）$=\{\vec{x}|f \in F_S\}$，则X=L（S），并且X的极点为满足$f_{\vec{x}} \in F_D$的$\vec{x}$。

反之，一般来说，若$f_{\vec{x}} \in F_D$,则$\vec{x}(f)$不一定是X的极点。

将下述约束记为X_ε:

(i). $\sum\limits_{i=0}^{N} \sum\limits_{a \in A(i)} \left(\delta_{ij}(i,j) - P_\varepsilon[j;i,a]\right) x_{ia} = 0; j \in V$

(ii). $\sum_{i=0}^{N}\sum_{a\in A(i)} x_{ia}=1$

(iii). $x_{ia}\geq 1; a\in A(i), i\in V$

可以证明

命题1 若$f\in F_D$,且f 描绘了一个HC，则$\overrightarrow{x}(f)$是X_ε的一个极点。

命题2 令$\overrightarrow{x}\in X_\varepsilon s.t. f_{\overrightarrow{x}}=M^{-1}(\overrightarrow{x})\in F_D$,则$f_{\overrightarrow{x}}$描绘了一个HC.

我们现在回忆[14]中F_D的划分.我们定义G的子图G_f

$$arc(i,a)\in G_f \Leftrightarrow f(i)=a$$

我们用

$$c_{1+k}^0=\{(i_0=0,i_1),(i_1,i_2),\ldots,(i_k,i_{1+k}=0)\};k=2,3,\ldots,N.$$

表示从0出发长度为1＋k的简单圈. 所以c_{1+N}^0是HC. 若G_f含有圈c_{1+k}^0我们写作$G_f\supset c_{1+k}^0$对$k=2,3,\ldots,N$,定义集合

$$C_{1+k}^0:=\{f\in F_D|G_f\supset c_{1+k}^0\}$$

为描绘出从0出发长度为1＋k的简单圈的策略集. 所以，$\bigcup_{k=1}^N C_{1+k}^0$ 包含了所有从0出发且第一次返回的结点是0的策略.用B表示$\bigcup_{k=1}^N C_{1+k}^0$在$F_D$中的余集，则B包含了所有从0出发且第一次返回的结点不是0的策略.于是

$$F_D=\left[\bigcup_{m=2}^N C_m\right]\bigcup B$$

接下来的命题的证明思路同[9]，[11]

命题3 令$\varepsilon\in(0,\frac{1}{\sqrt{N-1}}), f\in F_D$ 并且$\overrightarrow{x}(f)=M(f)$，则有访问开始结点0的平均频率为

$$x_0(f)=\sum_{a\in A(0)}x_{0a}(f)=\begin{cases}\frac{1}{d_{1+N}(\varepsilon)}=\frac{1}{1+N}+O(\varepsilon) & \text{if } f\in C_{1+N}\\ \frac{1}{d_{1+k}(\varepsilon)}+O(\varepsilon)=\frac{1}{1+k}+O(\varepsilon) & \text{if } f\in C_{1+k}, k=1,2,3,\ldots,N-1\\ \frac{\varepsilon}{1+\varepsilon}+O(\varepsilon) & \text{if } f\in B\end{cases}$$

其中

$$d_{1+k}(\varepsilon)=\frac{1-(1-\varepsilon)^k}{\varepsilon}+\varepsilon+(1+(k-1)\varepsilon)(1-\varepsilon)^k=1+k+O(\varepsilon).$$

推论3[14] G中的HC与满足下述两个条件的X_ε的点是一一对应的:

(1)$x_0=\sum_{a\in A(0)}x_{0a}=\frac{1}{d_{1+N}(\varepsilon)}$

(2)$\forall i\in V, x_i=\sum_{a\in A(i)}x_{ia}>0$并且$\frac{x_{ia}}{x_i}\in\{0,1\},\forall a\in A(i), i\in V$

令

$$S(\vec{x}) := \sum_i x_i^2 = \sum_i \left(\sum_a x_{ia}\right)^2$$

$$s(\vec{x}) := \sum_i \sum_a x_{ia}^2$$

考虑二次型

$$Q(\vec{x}) := S(\vec{x}) - s(\vec{x}) = \sum_i ((\sum_{a\in A(i)} x_{ia})^2 - \sum_{a\in A(i)} x_{ia}^2) = \vec{x}^T Q \vec{x}$$

其中Q为第i块对角块为对角元素为0，非对角元素为1的m_i*m_i子矩阵的准对角阵.
易证有

引理1 $\vec{x} \in X_\varepsilon$

$$\frac{x_{ia}}{x_i} \in \{0,1\}, \forall i \in V, a \in A(i) \tag{1}$$

$$\Leftrightarrow \vec{x}^T Q \vec{x} = 0$$

下面的定理2说明了HCP是如何与不定二次规划问题紧密联系的.

定理2 1. 若f 是G中的HC，则$\vec{x}(f)$是

$$\min \vec{x}^T Q \vec{x} \tag{2}$$

s.t.

$$G(\varepsilon)\begin{cases} (i) & \text{if } \vec{x} \in X_\varepsilon \\ \\ (ii) & (1)x_0 = \sum_{a\in A(0)} x_{0a} = \frac{1}{d_{1+N(\varepsilon)}} \end{cases}$$

的全局最小值，并且$\vec{x}^T Q \vec{x} = 0$

2. 反之，若$\vec{x}^*$是$G(\varepsilon)$ 的全局最小值，并且有$(\vec{x}^*)^T Q \vec{x}^* = 0$,则$f_{\vec{x}^*} = M^{-1}(\vec{x}^*) \in F_D$，并且$f_{\vec{x}^*}$描绘出一个HC.

证明： 注意到$\vec{x}^T Q \vec{x} \geq 0, \forall \vec{x} \in G(\varepsilon)$，并且因$G(\varepsilon)$是有界闭的，故其最小值存在.

1. 由命题1，3，$\vec{x}(f) \in G(\varepsilon)$, 故由引理1，$\vec{x}(f)^T Q \vec{x}(f) = 0$.

2. 若有$\vec{x}^* \in G(\varepsilon) s.t. (\vec{x}^*)^T Q \vec{x}^* = 0$，由引理1，$f_{\vec{x}^*} \in F_D$，再由命题2得证.

证毕.

在矩阵表示中X_ε可以写作

$$X_\varepsilon = \{\vec{x} | W_\varepsilon \vec{x} = \vec{0}, \vec{1}\vec{x} = 1, \vec{x} \geq \vec{0}\}$$

其中$\vec{x}$, $\vec{1}$都是m维列向量（$\vec{1}$的每个元素均为1)，并且W_ε 是（1+N）×m 的矩阵

$$w_\varepsilon[j,(i,a)] := \delta(i,j) - p_\varepsilon[j;i,a]$$

从X_ε的构造可以看出W_ε与incidence（1＋N）×m matrix D（G）的联系

$$D(G)[j,(i,a)] := \begin{cases} 1 & \text{if } i=j \\ -1 & \text{if } j=a \\ 0 & \text{otherwise} \end{cases}$$

易证有

引理2 矩阵W_ε是一个（ε－quadratically）perturbed incidence matrix D.

我们注意到W_i的每一行的行和为1，从而W_ε的每一行的行和为0. 所以，rank（W_ε）$\leq$ N.事实上，因rank（D（G））＝N（由[6]），故由引理2知当ε充分小时，rank（W_ε）＝N.

我们再注意到（4）中的约束可以被表示为

$$A_\varepsilon(G)\overrightarrow{x} = \overrightarrow{b}, \overrightarrow{x} \geq \overrightarrow{0} \tag{3}$$

其中$(i)A_\varepsilon(G)$是（N＋3）×m矩阵. $A_\varepsilon(G) = \begin{pmatrix} W_\varepsilon \\ \overrightarrow{a_1} \\ \overrightarrow{a_2} \end{pmatrix}$，$\overrightarrow{a_1}$的元素全为1，$\overrightarrow{a_2}$在对应于从0出发的弧的位置（0，a）为1，其余为0. $(ii)\overrightarrow{b} = \left(0,\ldots,0,1,\frac{1}{d_{1+N}(\varepsilon)}\right)^T_{1\times(3+N)}$

系统（5）有最大的秩对数值计算有重要意义.事实上，因rank（W_ε）＝N，故rank（$A_\varepsilon(G)$）$\leq$ N+2.下面的引理保证了连通图G中rank（$A_\varepsilon(G)$）＝N+2. 显然，含HC的图一定是连通图，而我们只要花多项式时间就可以验证一个图是否连通图.

引理3 若图G是连通图，则当$\varepsilon \in \left(0,\frac{1}{\sqrt{N-1}}\right)$，有rank（$A_\varepsilon(G)$）＝N+2.

上述引理可以证明.现在我们的目标函数$Q(\overrightarrow{x})$不足以充分利用我们的方法.我们需要一个目标函数族

$$h_\alpha(\overrightarrow{x}) := S(\overrightarrow{x}) - \alpha s(\overrightarrow{x}) = \alpha Q(\overrightarrow{x}) + (1-\alpha)S(\overrightarrow{x}), 0 < \alpha \leq 1$$

我们将看到h_α的全局最小值不是在HC发生，便在HC的一个小邻域内发生. 我们用

$$\arg\min(h_\alpha(\overrightarrow{x}))$$

表示在X_ε中h_α达到最小值的位置.

引理4 若G含有HC，则$\forall\delta > 0, \exists\varepsilon_0(\delta,\alpha) > 0, s.t. \forall\varepsilon \in (0,\varepsilon_0(\delta,\alpha)), \arg\min(h_\alpha(\overrightarrow{x}))$在G中一个HC的$\delta-$邻域中.

证明： 令

$$\overrightarrow{x^0} := \overrightarrow{x^0}(\alpha) = \arg\min(h_\alpha(\overline{x}))$$

并令$\overrightarrow{x^0} := \overrightarrow{x^0}(\alpha) = \arg\min(h_\alpha(\overline{x}))$为HC的频率向量.因$Q(\overrightarrow{x_{HC}})$=0，并且由命题3，$S(\overrightarrow{x_{HC}}) = \frac{1}{1+N} + O(\varepsilon)$，则有

$$h_\alpha(\overrightarrow{x^0}(\alpha)) \leq h_\alpha(\overrightarrow{x_{HC}}) = \frac{1-\alpha}{1+N} + O(\varepsilon)$$

因当$x_i = \frac{1}{1+N}, i = 0, \ldots, 1+N$，$S(\overrightarrow{x})$达到最小值，故

$$S(\overrightarrow{x}) \geq \frac{1}{1+N}, \forall \overrightarrow{x} \in X_\varepsilon$$

从而

$$S(\overrightarrow{x^0}) = \frac{1}{1+N} + O(\varepsilon)$$

并且，$Q(\overrightarrow{x^0})$应为$O(\varepsilon)$，这说明了$\overrightarrow{x^0}$在X_ε的极点$\overrightarrow{x^*}$ 的一个小邻域内.因为M建立了X_ε的极点与$\sum_\varepsilon$ 中F_D 确定性策略的一一对应，所以$f_{\overrightarrow{x^*}} := M^{-1}(\overrightarrow{x^*})$ 是确定性策略.由命题3，唯一满足$x_0(f_{\overrightarrow{x^*}}) = \frac{1}{1+N} + O(\varepsilon)$ 的$f_{\overrightarrow{x^*}}$ 是$f_{\overrightarrow{x^*}}$ =HC.

证毕.

§3　用内点法解二次规划问题

我们已经很好地理解了[2], [3]中如何用内点法解优化问题.这种方法对大规模的问题十分有效. 我们先尝试用内点法来解非凸二次规划问题（4）.为此运用[2, 5]中的HOPDM solver.下面我们讨论一般的运用方法.

Primal

$$\begin{aligned} &\min \overrightarrow{c}^T\overrightarrow{x} + \frac{1}{2}\overrightarrow{x}^TQ\overrightarrow{x} \\ &s.t. A\overrightarrow{x} = \overrightarrow{b}, \\ &\overrightarrow{x} \geq \overrightarrow{0} \end{aligned}$$

Dual

$$\begin{aligned} &\max \overrightarrow{b}^T\overrightarrow{y} - \frac{1}{2}\overrightarrow{x}^TQ\overrightarrow{x} \\ &s.t. A^T\overrightarrow{y} + \overrightarrow{s} - Q\overrightarrow{x} = \overrightarrow{c}, \\ &\overrightarrow{y}\, free, \overrightarrow{x}, \overrightarrow{s} \geq 0, \end{aligned}$$

其中$A \in R^{m\times n}, Q \in R^{n\times n}, \overrightarrow{x}, \overrightarrow{s}, \overrightarrow{c} \in R^n \quad \overrightarrow{y}, \overrightarrow{b} \in R^m$

算法的重点是计算Prmal－dual 牛顿方向. 这需要解下面的线性系统.

$$\begin{bmatrix} A & 0 & 0 \\ -Q & A^T & 0 \\ S & 0 & X \end{bmatrix} \begin{bmatrix} \Delta_x \\ \Delta_y \\ \Delta_s \end{bmatrix} = \begin{bmatrix} \xi_p \\ \xi_d \\ \xi_\mu \end{bmatrix} \tag{4}$$

其中

$$\xi_p = \overrightarrow{b} - A\overrightarrow{x}, \xi_d = \overrightarrow{c} - A^T\overrightarrow{y} - \overrightarrow{s} + Q\overrightarrow{x}, \xi_\mu = \mu e - XSe,$$

X和S是分别以$\overrightarrow{x}, \overrightarrow{s}$的元素为对角元素的n×n对角矩阵.

因

$$\Delta_s = X^{-1}\xi_\mu - X^{-1}S\Delta_x$$

牛顿系统被简化为

$$\begin{bmatrix} -Q - X^{-1}S & A^T \\ A & 0 \end{bmatrix} \begin{bmatrix} \Delta_x \\ \Delta_y \end{bmatrix} = \begin{bmatrix} \xi_d - X^{-1}\xi_\mu \\ \xi_p \end{bmatrix} \tag{5}$$

（7）中的矩阵是对称的担是是不定的.为了提高效率，在HOPDM implementation [2]，在简化了的牛顿系统中的矩阵被diagonal terms R_p 和R_d规则化

$$H=\begin{bmatrix} -Q-X^{-1}S & A^T \\ A & 0 \end{bmatrix}\begin{bmatrix} -R_p & 0 \\ 0 & R_d \end{bmatrix} \tag{6}$$

来获得一个半定矩阵[19]. 这允许了Cholesky－like factorization的使用，分解LDL^T中D是对角元素有正有负的对角阵. primal－dual regularization（8）的使用保证了有对角阵D的Cholesky－like factorization的存在，避免了使用需另外方式求得2×2 pivots来分解不定矩阵的需要[5, 8].

直接用内点法解（5）面临着两个难题. 第一，Q中对角方块的密度导致矩阵H的对称分解LDL^T非常密集，这会严重影响内点法的效率. 第二，矩阵Q是不定的，并且二次规划问题（4）有很多局部最小值. 接下来我们来讨论解决这两个难题的方法.

Q中对角方块可以表示对角阵的rank－one corrections.事实上，对于

$$Q_i(\overrightarrow{x_i}):=(\sum_{a\in A(i)} x_{ia})^2-\sum_{a\in A(i)} x_{ia}^2=\overrightarrow{x_i}^T Q_i \overrightarrow{x_i}$$

其中

$$Q_i=\begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix}\begin{bmatrix} 1 & 1 & \cdots & 1 \end{bmatrix}-I=\begin{bmatrix} 0 & 1 & \cdots & 1 \\ 1 & 0 & \cdots & 1 \\ \vdots & \vdots & \ddots & \vdots \\ 1 & 1 & \cdots & 0 \end{bmatrix}$$

我们引入一个附加变量

$$x_i=\sum_{a\in A(i)} x_{ia} \tag{7}$$

我们把二次型$Q_i(\overrightarrow{x_i})$变成对角的形式

$$\widetilde{Q_i}(\overrightarrow{x_i},x_i)=x_i^2-\sum_{a\in A(i)} x_{ia}^2$$

当然，我们对每个结点要添加新的约束（9）和新的变量x_i到重建的二次规划问题中去.所以，原来（4）中的N＋3个约束和$|A|$个变量被替换为含2N＋4个约束和$|A|$＋N＋1个变量的新二次规划问题.然而，难题一已迎刃而解了.

二次规划问题的非凸性意味着在系统（7）左上方的矩阵$-Q-X^{-1}S$不一定负定.因此便不能保证小的primal and dual regularization R_p 和R_d足够使得（8）中的H半定. 我们的方法是考虑一个“更凸”的目标函数.所以我们构造函数

$$\mathrm{h}_\alpha=S(\overrightarrow{x})-\alpha s(\overrightarrow{x})$$

从而有

$$\widetilde{Q_i}(\overrightarrow{x_i},x_i)=x_i^2-\alpha\sum_{a\in A(i)} x_{ia}^2$$

这个函数还是非凸的.但是，它的非凸程度随着α减小而下降. 因此，一个较弱的regularization足以凸化原问题. 引理4为h_α的运用奠定了理论基础.

§4 基于二次规划问题的解的算法

作为主干问题，在（4）中我们很难直接解出使目标函数值达到0的点. 更一般的情况找到是（4）中的一个局部最小值.这样的解存在结点i含有两个或两个以上$x_{ia}, a \in A(i)$非负.所以这个解无法对应HC.然而，这个局部解提供了很多有用的信息，并且允许我们使用基于x_{ia}含义的方法来寻找HC.

消弧法（ARC ELIMINATION）

对于二次规划问题的局部最小值，我们计算

$$f[i,a] = \frac{x_{ia}}{x_i}, a \in A(i)$$

若$f[i,a]$是可忽略的，即对某一特定的δ，$f[i,a] < \delta$，那么弧（i,a）会被认为不可能在HC中出现，从而从图中消除. 这样接着解对应于简化后图的新的二次规划问题并继续分析.重复几次简化和解二次规划问题后，我们会发现再没有弧满足消除准则.如果最后的解对任意i，只有一个$f[i,a]=1$，那么我们找到了一个HC. 然而，如果有两个或两个以上的$f[i,a]$非负，我们就用下面的分支法.

分支法（BRANCHING）

Branching是再整数规划中广泛使用的技术[20].若对应于简化后图的新的二次规划问题中，存在结点i有两个或两个以上的$f[i,a] \geq \delta$，我们用$|A(i)|$个新问题代替原问题.它们中的每个对应于不同的图，这个图中由i出发的弧只有（i,a），其他原图中由i出发的弧均被消除.

这种Branching无疑增加了假设的复杂度.于是我们将这种Branching结合大大简化原图的消弧法（ARC ELIMINATION）

参考文献

[1] Alspach, B. (1983), The classi cation of Hamiltonian generalized Petersen graphs, Journal of Combinatorial Theory ser.B 34, 293-312.

[2] Altman A. and Gondzio J. (1999), Regularized symmetric inde nite systems in interior point methods for linear and quadratic optimization, Optimization Methods and Software 11-12, 275-302.

[3] Andersen, E.D., Gondzio, J.,(1996), Implementation of interior point methods for large scale linear programming.In:Terlaky, T.(ed.), Interior Point Methods in Mathematical Programming, Kluwer Academic Publishers, Dordrecht, pp.189-252.

[4] Andramonov, M., Filar, J.A., Rubinov, A. and Pardalos, P. (2000), Dordrecht Hamiltonian cycle problem via Markov chains and min-type approaches.In:Pardalos, P.M.(ed.), Approx-imation and Complexity in Numerical Optimization:Continuous and Discrete Problems, Kluwer Academic Publishers, Dordrecht, pp.31-47.

[5] Bunch, J.R. and Parlett, B.N. (1971), Direct methods for solving symemtric inde nite systems of linear equations, SIAM Journal on Numerical Analysis, 8, 639-655.

step1.开始新的迭代：取$x_t = \arg\min_{x\in I} f(x)$。若$f(x_c) - f(x_t) > 0.1\epsilon_\Delta\Delta$，取$x_c = x_t$（$f(x_t)$充分小则改变$x_c$为$x_t$，否则保持$x_c$不变）

step2.收敛测试：若$\Delta < \epsilon_\Delta$（I的半径足够小），或$f(x_c) < f_l + \epsilon_f$（$f(x_c)$足够小），停止并返回$x_c$作为近似解

step3.计算高次模型：若$|I| = p, i.e. \forall x \in J, \|x - x_c\| \leqslant \Delta$，构造高次模型$m(x_c + s) = f(x_c) + g^T s + \frac{1}{2}s^T Hs$（利用lagrange插值），否则，$|I| < p, i.e. \exists x \in J, s.t. \|x - x_c\| > \Delta$

1.构造$m(x_c + s)$形式的模型，$s.t. m(x) = f(x) \forall x \in I$，且最小化$\|g\|^2 + \|H\|^2$，若无法求得，利用算法A1改进，进入step7

2.构造子模型$m(x)$

step4. 解信赖域问题$\min\limits_{\|s\|\leqslant\Delta} m(x_c+s)$，若$m(x_c) - m(x_c+s) < \max[\epsilon_\Delta\Delta, \frac{\eta_a}{2}, \frac{\eta_r}{2}|f(x_c+s)|]$（预测模型减少量大小），则改进模型的几何性质（用A1）进入step7；否则令$M = M \cup \{x_c + s\}$，计算$f(x_c + s)$及$\rho = \dfrac{f(x_c) - f(x_c + s)}{m(x_c) - m(x_c + s)}$

step5.若$x_c + s$成功，即$\rho \geqslant 0.9$，计算
$Theta = \min[1000\Delta, \max\|x_i - x_c\|]$，其中$x_i \in M\backslash I(\|x_i - x_c\| \geqslant \Delta), s.t. f(x_i) - f(x_c) \geqslant 0.85(m(x_i - m(x_c))$（$x_c + s$就是其中一个）

若$\Theta > (1+r)\|s\|$，求解$d, s.t. \min\limits_{\|d\|\leqslant\Theta} m(x_c+d)$，若$\|d\| > (1+r)\|s\|$，令$M = M\cup\{x_c+d\}$且计算$f(x_c + d)$，重新定义$s = d, \rho = \rho d$

step6.计算I，J中能替代$x_c + s$的最佳点。$x_I = \arg\max\limits_{x\in I\backslash\{x_c\}} S(x_c + s, x), x_J = \arg\max\limits_{x\in J} S(x_c + s, x)$，定义$S_I = S(x_c + s, x_I), S_J = S(x_c + s, x_J)$。

若$\rho \geqslant 0.05$，令$I = (I\backslash\{x_I\}) \cup \{x_c + s\}$，若$S_I > 2S_J$，$I = I \cup \{x_c + s\}$，且$J = J\backslash\{x_j\}$，若$S_I \leqslant S_J$；

否则，若$\max(S_I, S_J) > 1$，令$I = (I\backslash\{x_I\}) \cup \{x_c + s\}$，若$S_I > S_J$，$I = I \cup \{x_c + s\}, J = J\backslash\{x_J\}$，若$S_I \leqslant S_J$；

若以上两者均不符合，或者$\rho < 0.15$，利用A1改进模型。

step7. 若$\rho \geqslant 0.75$，令$\Delta = \min[r\Delta, \max(\Delta, r\|s\|)]$；

反之若$\Delta > \epsilon_\Delta r$，且step3未进入问题状态，令$\Delta = \max(\epsilon_\Delta r, \frac{\Delta}{r})$。

若上两种都失败且A1部嫩改进几何形状，令$\Delta = \frac{\Delta}{r}$。

回到step1开始新的循环。

§3 newton插值算法

接下来考虑用newton插值多项式作为近似模型。

n元d次多项式插值对应的点集为Y，$\#Y=p=2^n$，分为$d+1$块：$Y^o, Y^1, \cdots, Y^d$，分别含有$1, n, \cdots, \binom{d+n-1}{d}$个点，对应排列为$y_1, \cdots, y_p$；对应于点的分块，将newton基多项式也分块为$\{N_i^0(\cdot)\}, \{N_i^1(\cdot)\}, \cdots, \{N_i^d(\cdot)\}$，分别含有$1, n, \cdots, \binom{d+n-1}{d}$个多项式，其中$N_i^l$为$l$次多项式，且满足条件$N_i(y_i)=1, N_i(y_j)=0$ 当且仅当y_j与y_i位于同一块中或y_j在y_i后面的块中，可以用矩阵表示为

$$\delta(I)=\begin{pmatrix} 1 & 0 & \cdots & 0 \\ * & I_n & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ * & * & \cdots & I_{\binom{d+n-1}{d}} \end{pmatrix}$$

其中$\delta(I)$中第i行第j列表示$N_j(y_i)$的值（$N_1, \cdots, N_p$为$\{N_i^0\}, \cdots, \{N_i^d\}$的依次排列）。

要获得这样的newton基可以通过Gramm-Schimidt正交化完成。首先任取n维d次多项式空间的基，例如$\{1, x_1, \cdots, x_n, \cdots, x_1^d, \cdots, x_n^d\}$。

算法：令$Y_{temp}=\Phi$，

For $l=0, \cdots, d$

For $i=1, \cdots, |Y^l|$，取$y_i^l \in Y \backslash Y_{temp}, s.t.\ N_i^l(y_i^l) \neq 0$。

若这样的y_i^l不存在，则令$Y=Y_{temp}$，停止算法。

$Y_{temp}=Y_{temp} \cup \{y_i^l\}$

单位化：$N_i^l(x)=\dfrac{N_i^l(x)}{|N_i^l(y_i^l)|}$

对所有的l块及以上块中的多项式处理

$N_j^l(x)=N_j^l(x)-N_j^l(y_i^l)N_i^l(x)\ (j \neq i, j=1, \cdots, |Y^l|)$

$N_j^k(x)=N_j^k(x)-N_j^k(y_i^l)N_i^l(x)\ (j=1, \cdots, |Y^k|, k=l+1, \cdots, d)$

end

关于newton插值多项式有如下结论：令$\pi(x)=(y_0^0, y_1^1, \cdots, y_d^d, y_{d+1}^{d+1}=x)$，其中$y_i^i \in Y^i, i=0, \cdots, d$，$x \notin Y$为任意点。令$\Pi(x)=\{\pi(x)\}$。

设$m(x)$为$f(x)$的newton插值多项式，则有

$$|f(x)-m(x)| \leqslant \frac{n^{d+1}\|f^{(d)}\|}{(d+1)!} \sum_{\pi(x) \in \Pi(x)} [\prod_{i=0}^{d} \|y_{i+1}^{i+1}-y_i^i\| |N_i^i(y_{i+1}^{i+1})|]$$

意即当Y落在一个很小的范围内，例如以δ为半径的球内，则$f(x)$与$m(x)$的差别很小，即可以用$m(x)$很好地逼近$f(x)$。

有了以上的结论，我们可以给出如下的算法：给定常数 $0<\eta_0 \leqslant \eta_1<1, 0<\gamma_0 \leqslant \gamma_1<1 \leqslant \gamma_2, \epsilon_g>0, 1 \leqslant \mu, \mathcal{A}_k=\{y \in \mathbb{R}^n : \|y-x_k\| \leqslant \mu\|g_k\|\}$

step0.初始化：给定$x_s, f(x_s)$，选定初始插值集合$Y, x_s \in Y$，且$Y \backslash \{x_s\} \neq \Phi$，取$x_0 \in Y, s.t.\ f(x_0) = \min\limits_{y_i \in Y} f(y_i)$，选定初始信赖域半径$\Delta_0>0$。令$k=0$

step1.建立模型（用newton插值Y中的点）：使用插值点集合Y，建立模型$m_k(x_k+s)$，若$\|g_k\| \leqslant \epsilon_g$，且Y的几何性质不够好，则改进，直至存在某个$\delta_k \in (0, \mu\|g_k\|)$，使得Y在$\mathcal{B}_k(\delta_k)$中有好的性质，改进方法：取$y_i^l \in Y$，且$y_i \notin \mathcal{B}_k(\delta_k)$，将它用点$z$替代，其中$z=\arg\max\limits_{x \in \mathcal{B}_k(\delta_k)} N_i^l(x)$。

step2.在信赖域内最小化模型：计算x_k+s_k, $s.t.\ m_k(x_k+s_k)=\min_{x\in\mathcal{B}_k} m_k(x)$，计算$f(x_k+s_k)$，及$\rho_k := \dfrac{f(x_k)-f(x_k+s_k)}{m_k(x_k)-m_k(x_k+s_k)}$

step3. 改进插值点集

若$\rho_k\geqslant\eta_1$，用x_k+s_k代替Y中的某一个点；

若$\rho_k<\eta_1$，且Y的几何性质不够好，则改进

step4. 改进半径

若$\rho_k\geqslant\eta_1$，取$\Delta_{k+1}\in[\Delta_k,\gamma_2\Delta_k]$；

若$\rho_k<\eta_1$，且Y的几何性质不够好，取$\Delta_{k+1}\in[\gamma_0\Delta_k,\gamma_1\Delta_k]$；

若不然，取$\Delta_{k+1}=\Delta_k$

step5.改进当前点：取$f(\overline{x_k})=\min_{y_i\in Y,y_i\neq x_k} f(y_i)$，

若$\overline{\rho_k} := \dfrac{f(x_k)-f(\overline{x_k})}{m_k(x_k)-m_k(x_k+s_k)}\geqslant\eta_0$，令$x_{k+1}=\overline{x_k}$；否则令$x_{k+1}=x_k$。

返回step1。

§4 两种算法比较

1、lagrange插值多项式的次数由插值点数$\#Y$唯一决定；而newton插值多项式的次数可以任意指定，或根据插值精度的要求给定；

2、Lagrange算法比较复杂，经常需要处理一些最大值问题；而newton算法比较简单，只需简单的计算以及比较大小；

3、newton算法的收敛性是得到保证的，这是由于前文中所给出的newton插值误差范围可以人为地进行控制；然而lagrange算法并没有如此优良的性质，因此不能保证它的收敛性；

因此在一般使用时，倾向于使用newton算法。

更一般地，在与其它迭代方法进行比较时，newton法也具有其优势：如该算法要求得到的函数值相对较少，而且对于函数值小范围震荡较不敏感，这些优点都使得它能够被广泛使用。

参考文献

[1] A.R.Conn,K.Scheinber,Ph.L.Toint:Recent progress in unconstrained nonlinear optimization without derivatives.Mathematical Programming,1997,79:397-414

[2] A.R.Conn,Ph.L.Toint:An algorithm using quadratic interpolation for unconstrained derivative free optimization. Di Pillo G,Gianessi F,eds.Nonlinear Optimization and Applications.New York:Plenum Publishing,1996.27-47

[3] K. Scheinberg. Derivative free optimization method. Technical report, IBM T.J. Watson Research Center, Newyork, 2000. Tech. Rep. Report 98/11

[4] A. R. Conn, K. Scheinberg, and P.L. Toint. On the convergence of derivative- free methods for unconstrained optimization. In A. Iserles and M. Buhmann, editors, Approximation Theory and Optimization: Tribute to M.J.D. Powell, pages 83108, Cambridge, UK, 1997. Cambridge University Press.

[5] M.J.D. Powell, On trust region methods for unconstrained minimization without derivatives, Math. Programming, Vol. 97, pp. 605 - 623, (2003).

[6] M.J.D. Powell, A view of unconstrained minimization algorithms that do not require derivatives,ACM Transactions on Mathematical Software 1 (2) (1975) 97 - 107.

THE 2008 WOLF FOUNDATION PRIZE WINNERS

Pierre R. Deligne Institute for Advanced Study Princeton, New Jersey, USA for his work on mixed Hodge theory; the Weil conjectures; the Riemann-Hilbert correspondence; and for his contributions to arithmetic.

Phillip A. Griffiths Institute for Advanced Study Princeton. New Jersey, USA for his work on variations of Hodge structures; the theory of periods of abelian integrals; and for his contributions to complex differential geometry.

David B. Mumford Brown University Providence, Rhode Island, USA for his work on algebraic surfaces; on geometric invariant theory; and for laying the foundations of the modern algebraic theory of moduli of curves and theta functions.

Central to modern algebraic geometry is the theory of moduli, i.e., variation of algebraic or analytic structure. This theory was traditionally mysterious and problematic. In critical special cases, i.e., curves, it made sense, i.e., the set of curves of genus greater than one had a natural algebraic structure. In dimensions greater than one, there was some sort of structure locally, but globally everything remained mysterious. The two main (and closely related) approaches to moduli were invariant theory on the one hand and periods of abelian integrals on the other. This key problem was tackled and greatly elucidated by Deligne, Griffiths, and Mumford.

关于量子化的两个定理的证明*

0401　官永辉　　洪继展　　李盈伟　　王常亮

§1　简介

在Thomas Henry Coates的论文《Riemann-Roch Theorems in Gromov-Witten Theory》中，他引用了两个定理.在本文中，我们将对这两个定理给出一个部分的证明.为了方便符号上的统一，我们将用到的一些定义在本节中叙述.

给定一个h维的，带有一个非退化的，对称的，双线性复函数$\langle\cdot,\cdot\rangle$的复向量空间$V$，取定其基$\phi_1,\phi_2,\ldots,\phi_h$，并记$g_{\alpha\beta}=\langle\phi_\alpha,\phi_\beta\rangle$，$\phi^\alpha=g^{\alpha\beta}\phi_\beta$，其中$g^{\alpha\beta}$为$(g_{\alpha\beta})$的逆的矩阵元.令$\mathbb{H}$为$V$上的Laurent 级数所构成的集合.在$\mathbb{H}$上可以定义一个$\Omega$，如下：

$$\Omega\left(f,g\right)=\frac{1}{2\pi i}\oint\left\langle f(-z),g(z)\right\rangle\,dz$$

其中积分路径延绕原点的闭曲线.按这个定义，Ω就是$\mathbb{H}$上的一个非退化的，斜对称的辛形式，$(\mathbb{H},\Omega)$也就成为一个辛向量空间. $\mathbb{H}$中的任意一个$f(z)$都可以表示成以下的方式：

$$f(z)=\sum_{k=0}^{\infty}\mathbf{p}_k\frac{1}{(-z)^{k+1}}+\sum_{k=0}^{\infty}\mathbf{q}_k z^k$$

其中，使用求和约定，$\mathbf{p}_k=p_k^\alpha\phi^\alpha,\mathbf{q}_k=q_k^\alpha\phi_\alpha$.定义$\mathbb{H}$上的无穷小辛变换为满足如下条件的$A$：

$$\langle(Af)(-z),g(z)\rangle+\langle f(-z),(Ag)(z)\rangle=0,\forall f,g\in\mathbb{H}$$

此时易得：

$$\Omega(Af,g)+\Omega(f,Ag)=0$$

定义$h_A(f)=\frac{1}{2}\Omega(Af,f)$，则$h_A(f)$可以表示成关于$p_k^\alpha,q_k^\alpha$的二次型.在此基础上我们再定义$\widehat{q_k^\alpha q_l^\beta}$，$\widehat{q_k^\alpha p_l^\beta},\widehat{p_k^\alpha q_l^\beta},\widehat{p_k^\alpha p_l^\beta}$分别为$q_k^\alpha q_l^\beta,q_k^\alpha p_l^\beta,p_k^\alpha q_l^\beta,p_k^\alpha p_l^\beta$的量子化：

$$\begin{aligned}\widehat{q_k^\alpha q_l^\beta}&=\frac{q_k^\alpha q_l^\beta}{\hbar}\\ \widehat{q_k^\alpha p_l^\beta}&=q_k^\alpha\frac{\partial}{\partial q_l^\beta}=\widehat{p_l^\beta q_k^\alpha}\\ \widehat{p_k^\alpha p_l^\beta}&=\hbar\frac{\partial}{\partial q_k^\alpha}\frac{\partial}{\partial q_l^\alpha}\end{aligned}$$

从而，我们可以用线性扩充的方法定义A的量子化$\widehat{A}=\widehat{h_A(f)}$.我们再考虑形如$S(z)=\exp\left\{Bz^m\right\}$，$B\in\mathrm{End}(V)$，其中$Bz^m$是一个无穷小辛变换，定义这种变换$S$的量子化为$\widehat{S}=\exp\left\{\widehat{Bz^m}\right\}$.令$\mathfrak{F}$是

*衷心感谢中科院数学与系统科学研究所潘建中老师的指导

以$q_k^\alpha, k=0,1,\ldots.\alpha=1,\ldots,h$为未定元的形式幂级数，对于$\mathfrak{F}$中的元素$\mathcal{G}$，定义$\mathcal{G}(\mathbf{r})$是将$\mathcal{G}$中的$q_k^\alpha$替换成相应的$r_k^\alpha$后得到的幂级数，其中，$\mathbf{r}\in\mathbb{H}$且$\mathbf{r}=\mathbf{r}_0+\mathbf{r}_1z^1+\cdots$ 对于$A=\sum_m B_m z^m$, $\mathbf{q}=\sum_k \mathbf{q}_k z^k$，定义$A\mathbf{q}=\sum_{m,k}B_m\mathbf{q}_k z^{m+k}$ 对于$\forall B\in \mathrm{End}(V)$,定义

$$\begin{aligned}
B\phi_\alpha &= B^\beta{}_\alpha\phi_\beta\\
B^\alpha{}_\beta &= \langle B\phi_\beta,\phi^\alpha\rangle\\
B^{\alpha\beta} &= \langle B\phi^\beta,\phi^\alpha\rangle\\
B_{\alpha\beta} &= \langle B\phi_\beta,\phi_\alpha\rangle\\
\partial_{\alpha,k} &= \frac{\partial}{\partial q_k^\alpha}
\end{aligned}$$

引理1. 若Bz^m是一个无穷小辛变换，则：

$$\begin{aligned}
B^{\alpha\beta} &= (-1)^{m+1}B^{\beta\alpha}\\
B_{\alpha\beta} &= (-1)^{m+1}B_{\beta\alpha}\\
B^\alpha{}_\beta &= (-1)^{m+1}B^\beta{}_\alpha\\
B\phi^\alpha &= (-1)^{m+1}B^\alpha{}_\beta\phi^\beta
\end{aligned}$$

证明： 对于前三个公式，这里只证$B_{\alpha\beta}=(-1)^{m+1}B_{\beta\alpha}$，其余的类似可证.由于$Bz^m$为一个无穷小辛变换，故

$$\langle B(-z)^m\phi_\alpha,\phi_\beta\rangle+\langle\phi_\alpha,Bz^m\phi_\beta\rangle=0$$

对于最后一个，设

$$B\phi^\alpha=X_\beta\phi^\beta$$

则

$$\langle B\phi^\alpha,\phi_\tau\rangle=X_\tau$$

故

$$B\phi^\alpha=\langle B\phi^\alpha,\phi_\beta\rangle\phi^\beta=(-1)^{m+1}\langle B\phi_\beta,\phi^\alpha\rangle\phi^\beta=(-1)^{m+1}B^\alpha{}_\beta\phi^\beta$$

证毕. □

定理1. 设无穷小辛变换$A=Bz^m$，那么

当$m<0$时：

$$\widehat{A}=\frac{1}{2\hbar}\sum_{k=0}^{-m-1}(-1)^{k+m}B_{\alpha\beta}q_k^\beta q_{-m-1-k}^\alpha-\sum_{k=-m}^{\infty}B^\alpha{}_\beta q_k^\beta\partial_{\alpha,k+m} \tag{1.1}$$

当$m\geq 0$时：

$$\widehat{A}=-\sum_{k=0}^{\infty}B^\alpha{}_\beta q_k^\beta\partial_{\alpha,k+m}+\frac{\hbar}{2}\sum_{k=0}^{m-1}(-1)^kB^{\alpha\beta}\partial_{\beta,k}\partial_{\alpha,m-1-k} \tag{1.2}$$

证明： 在$m<0$的情形下，设

$$f(z)=\sum_{k=0}^{\infty}q_k^\alpha\phi_\alpha z^k+\sum_{l=0}^{\infty}p_l^\beta\phi^\beta(-z)^{-1-l}\in\mathbb{H} \tag{1.3}$$

则可得：

$$(Af)(-z)=\sum_{k=0}^{\infty}q_k^\alpha B\phi_\alpha(-z)^{k+m}+(-1)^m\sum_{l=0}^{\infty}p_l^\beta B\phi^\beta z^{-1-l+m}$$

故，可求得$\langle(Af)(-z),f(z)\rangle$在原点处的留数.考虑

$$\left\langle q_k^\alpha\phi_\alpha z^k,(-1)^m p_l^\beta B\phi^\beta z^{-1-l+m}\right\rangle=(-1)^m q_k^\alpha p_l^\beta\left\langle B\phi^\beta,\phi_\alpha\right\rangle z^{-1-l+m+k}$$

取$l=k+m$得$-\sum_{k=-m}^{\infty}q_k^\alpha p_{k+m}^\beta B^\beta{}_\alpha$一项，再考虑

$$\left\langle q_k^\alpha B\phi_\alpha(-z)^{k+m},p_l^\beta\phi^\beta(-z)^{-1-l}\right\rangle=q_k^\alpha p_l^\beta\left\langle B\phi_\alpha,\phi^\beta\right\rangle(-z)^{-1-l+m+k}$$

取$l=k+m$得$-\sum_{k=-m}^{\infty}q_k^\alpha p_{k+m}^\beta B^\beta{}_\alpha$又一项，还有

$$\left\langle q_l^\beta\phi_\beta z^l,q_k^\alpha B\phi_\alpha(-z)^{k+m}\right\rangle=(-1)^{k+m}q_l^\beta q_k^\alpha\left\langle\phi_\beta,B\phi_\alpha\right\rangle z^{k+m+l}$$

取$l+k+m=-1$得$\sum_{k=0}^{-1-m}(-1)^{k+m}B_{\beta\alpha}q_k^\alpha q_{-1-k-m}^\beta$一项，故知所求留数为：

$$-2\sum_{k=0}^{\infty}q_k^\alpha p_{k+m}^\beta B^\beta{}_\alpha+\sum_{k=0}^{-1-m}(-1)^{k+m}B_{\beta\alpha}q_k^\alpha q_{-1-k-m}^\beta$$

从而求得在$m<0$的时候，$\widehat{A}$如(1.1)式. 在$m>0$的情形下，仍取$f(z)$如(1.3)式，这里只算与$m<0$的情形不同的一项

$$\left\langle(-1)^m p_l^\beta B\phi^\beta z^{-1-l-m},p_k^\theta\phi^\theta(-z)^{-1-k}\right\rangle=(-1)^{m-1-k}p_k^\beta p_k^\theta B^{\theta\beta}z^{-2-l+m-k}$$

取$l=m-1-k$得$\sum_{k=0}^{m-1}(-1)^k p_k^\beta p_{m-1-k}^\alpha B^{\alpha\beta}$一项.故得(1.2)式. 在$m=0$时(1.2)式的第二项取为0.

证毕。□

§2 几个定理

定理2[1]. 对于形如$S(z)=\exp\{Bz^m\},B\in\mathrm{End}(V)$，其中$Bz^m$是一个无穷小辛变换且$m<0$，之$S$满足：[1]

$$(\widehat{S}^{-1}\mathcal{G})(\mathbf{q})=\exp\left\{\frac{W_S(\mathbf{q})}{2\hbar}\right\}\mathcal{G}([S\mathbf{q}]_+) \tag{2.1}$$

其中，$[S\mathbf{q}]_+$定义为$S(z)\mathbf{q}$中z的负次幂项的系数取为0后的Laurent级数，

$$\sum_{k,l}\frac{W_{kl}}{w^k z^l}=\frac{S^*(w)S(z)-I}{z^{-1}+w^{-1}}$$

[1]文献[1]与文献[2]中的叙述均有错

$$\begin{aligned} W_S(\mathbf{q}) &= \sum_{k,l}\langle \mathbf{q}_k, W_{kl}\mathbf{q}_l\rangle \\ S^*(-z)S(z) &= I \\ \mathbf{q} &= \sum_{k=0}^{\infty}\mathbf{q}_k z^k \end{aligned}$$

定理3[1]. 对于形如$R(z)=\exp\{Bz^m\}, B\in\mathrm{End}(V)$，其中$Bz^m$是一个无穷小辛变换且$m\geq 0$，之$R$满足：

$$(\widehat{R}\mathcal{G})(\mathbf{q}) = \left(\exp\left\{\frac{\hbar V_R(\partial_{\mathbf{q}})}{2}\right\}\mathcal{G}\right)(R^{-1}\mathbf{q}) \tag{2.2}$$

其中，$V_R(\partial_{\mathbf{q}})$指的是将$V_R(\mathbf{p})$的表达式中的$p_k^\alpha$换成$\partial_{\alpha,k}$

$$\begin{aligned} \sum_{k,l}(-1)^{k+l}V_{kl}w^k z^l &= \frac{R^*(w)R(z)-I}{z+w} \\ V_R(\mathbf{p}) &= \sum_{k,l}\langle \mathbf{p}_k, V_{kl}\mathbf{p}_l\rangle \\ R^*(-z)R(z) &= I \\ \mathbf{q} &= \sum_{k=0}^{\infty}\mathbf{q}_k z^k \\ \mathbf{p} &= \sum_{k=0}^{\infty}\mathbf{p}_k(-z)^{-1-k} \end{aligned}$$

引理2. 设∂_x是一个关于实变量y的一阶微分算子，则对于任意关于y的形式幂级数$f(y),g(y)$有

$$\exp\{\partial_x\}fg = (\exp\{\partial_x\}f)(\exp\{\partial_x\}g)$$

证明： 注意到$\partial_x(fg)=f\partial_x g+g\partial_x f$

$$\begin{aligned} \exp\{\partial_x\}fg &= \sum_{n=0}^{\infty}\frac{1}{n!}\partial_x^n(fg) \\ &= fg + \sum_{n=1}^{\infty}\frac{1}{n!}\partial_x^{n-1}(f\partial_x g+g\partial_x f) \\ &= fg + (f\partial_x g+g\partial_x f) + \sum_{n=2}^{\infty}\frac{1}{n!}\partial_x^{n-2}\left(f\partial_x^2 g+2\partial_x f\partial_x g+g\partial_x f\right) \end{aligned}$$

利用数学归纳法易证：

$$\begin{aligned} \exp\{\partial_x\}fg &= \sum_{n=0}^{\infty}\frac{1}{n!}\sum_{k=0}^{n}\binom{n}{k}\left(\partial_x^k f\right)\left(\partial_x^{n-k}g\right) \\ &= \sum_{n=0}^{\infty}\sum_{k=0}^{n}\frac{\partial_x^k f}{k!}\frac{\partial_x^{n-k}g}{(n-k)!} \\ &= (\exp\{\partial_x\}f)(\exp\{\partial_x\}g) \end{aligned}$$

证毕.□

引理3. 各符号约定同引理2，定义形式幂级数$f(y)$作用在$g(y)$上为$f(g)=fg$，则有：

$$\partial_x^n \circ f = \sum_{k=0}^{n} \binom{n}{k} \left(\partial_x^k f\right) \partial_x^{n-k}$$

证明：

$$\begin{aligned}
\partial_x^n \circ f(g) &= \partial_x^{n-1}\partial_x(fg) \\
&= \partial_x^{n-1}\left(g\partial_x f + f\partial_x g\right) \\
&= \partial_x^{n-2}\left(f\partial_x^2 g + 2\partial_x f\partial_x g + g\partial_x^2 f\right) \\
&= \cdots \\
&\cdots \\
&= \sum_{k=0}^{n} \binom{n}{k} \left(\partial_x^k f\right)\left(\partial_x^{n-k} g\right) \\
&= \left(\sum_{k=0}^{n} \binom{n}{k} \left(\partial_x^k f\right) \partial_x^{n-k}\right) g
\end{aligned}$$

证毕.□

定理2的证明： 以下分成三个步骤来证明：

(i.) (2.1)式两端关于$\hbar^{-1}$的零次方项相等.

(2.1)式的左边关于$\hbar^{-1}$的零次方项应该是$\exp\{\partial_A\}\mathcal{G}(\mathbf{q})$，其中$\partial_A=\sum_{k=-m}^{\infty}B^{\alpha}{}_{\beta}q_k^{\beta}\cdot \cdot\partial_{\alpha,k+m}$.根据引理2，对于$\forall\mathcal{G}_1,\mathcal{G}_2\in\mathfrak{F}$，总有$\exp\{\partial_A\}\mathcal{G}_1\mathcal{G}_2(\mathbf{q})=\exp\{\partial_A\}\mathcal{G}_1(\mathbf{q})\exp\{\partial_A\}\cdot \cdot\mathcal{G}_2(\mathbf{q})$，而右端的零次项满足$(\mathcal{G}_1\mathcal{G}_2)([S\mathbf{q}]_+)=\mathcal{G}_1([S\mathbf{q}]_+)\mathcal{G}_2([S\mathbf{q}]_+)$故只需证对于$\mathcal{G}\equiv 1$及$\mathcal{G}\equiv q_k^{\alpha}$，两端的零次项相等即可.而对于前一种情况，结论是显然的.以下对任意的q_k^{α} 来证明$\mathcal{G}\equiv q_k^{\alpha}$时，结论成立.

$$\begin{aligned}
\exp\{\partial_A\}q_k^{\alpha} &= \sum_{n=0}^{\infty}\frac{{\partial_A}^n}{n!}q_k^{\alpha} \\
&= q_k^{\alpha} + \sum_{n=1}^{\infty}\frac{{\partial_A}^{n-1}}{n!}B^{\alpha}{}_{\alpha_1}q_{k-m}^{\alpha_1} \\
&= q_k^{\alpha} + B^{\alpha}{}_{\beta}q_{k-m}^{\beta} + \sum_{n=2}^{\infty}\frac{{\partial_A}^{n-2}}{n!}B^{\alpha}{}_{\alpha_1}B^{\alpha_1}{}_{\alpha_2}q_{k-2m}^{\alpha_2} \\
&= \cdots \\
&\cdots \\
&= \sum_{n=0}^{\infty}\frac{B^{\alpha}{}_{\alpha_1}B^{\alpha_1}{}_{\alpha_2}\cdots B^{\alpha_{n-1}}{}_{\alpha_n}q_{k-mn}^{\alpha_n}}{n!}
\end{aligned}$$

考虑

$$[S\mathbf{q}]_+ = \left[\sum_{n=0}^{\infty}\frac{(Bz^m)^n}{n!}\sum_{t=0}^{\infty}q_t^{\beta}\phi_{\beta}z^t\right]_+$$

$$= \left[\sum_{n=0}^{\infty}\sum_{t=0}^{\infty}\frac{1}{n!}q_t^{\beta}B^n\phi_\beta z^{t+mn}\right]_+$$
$$= \left[\sum_{n=0}^{\infty}\sum_{t=0}^{\infty}\frac{1}{n!}q_t^{\beta}B^{\alpha_n}{}_{\alpha_{n-1}}B^{\alpha_{n-1}}{}_{\alpha_{n-2}}\cdots B^{\alpha_1}{}_{\beta}\phi_{\alpha_n}z^{t+mn}\right]_+$$

可以求z^k的系数为

$$\sum_{n=0}^{\infty}\frac{q_{k-mn}^{\beta}B^{\alpha}{}_{\alpha_{n-1}}B^{\alpha_{n-1}}{}_{\alpha_{n-2}}\cdots B^{\alpha_1}{}_{\beta}}{n!}\phi_\alpha$$

将β换成α_n，α_{n-1}换为α_1，α_{n-2}换为α_2，…，α_1换为α_{n-1}即可知

$$\mathcal{G}([S\mathbf{q}]_+) = \sum_{n=0}^{\infty}\frac{q_{k-mn}^{\alpha_n}B^{\alpha}{}_{\alpha_1}B^{\alpha_1}{}_{\alpha_2}\cdots B^{\alpha_{n-1}}{}_{\alpha_n}}{n!}$$

故知结论成立.

(ii.) $T_1(1) = W_S(\mathbf{q})/(2\hbar)$，其中$T_n$为$\widehat{S}^{-1}$展开式中关于$\hbar^{-1}$的$n$次项

$$\begin{aligned} T_1 &= \sum_{k=1}^{\infty}\frac{1}{k!}\sum_{s=0}^{k-1}\partial_A^s C\partial_A^{k-1-s} \\ C &= -\frac{1}{2\hbar}\sum_{k=0}^{-m-1}(-1)^{k+m}B_{\alpha\beta}q_k^{\beta}q_{-m-1-k}^{\alpha} \end{aligned}$$

化一下$T_1(1)$.

$$T_1(1) = \frac{1}{2\hbar}\sum_{n=1}^{\infty}\frac{1}{n!}\sum_{k=0}^{n-1}\partial_A^{n-1-k}C\partial_A^k(1) = \frac{1}{2\hbar}\sum_{n=0}^{\infty}\frac{\partial_A^n(C)}{(n+1)!}$$

$$\begin{aligned} \sum_{k,l}\frac{W_{kl}}{w^k z^l} &= \frac{S^*(w)S(z)-I}{w^{-1}+z^{-1}} \\ &= \frac{\exp\left\{Bz^m+(-1)^{m+1}Bw^m\right\}-I}{w^{-1}+z^{-1}} \\ &= \sum_{n=1}^{\infty}\frac{1}{n!}\frac{B^n(z^m+(-1)^{m+1}w^m)^n}{w^{-1}+z^{-1}} \\ &= \sum_{n=1}^{\infty}\frac{B^n(z^m+(-1)^{m+1}w^m)^{n-1}}{n!}\frac{z^m+(-1)^{m+1}w^m}{w^{-1}+z^{-1}} \\ &= \sum_{n=1}^{\infty}\frac{B^n}{n!}\sum_{s=0}^{n-1}\binom{n-1}{s}z^{ms}w^{m(n-1-s)}(-1)^{(m+1)(n-1-s)}\cdot \\ &\quad \sum_{k=0}^{-1-m}(-1)^{k+m+1}z^{-k}w^{k+m+1} \\ &= \sum_{n=1}^{\infty}\frac{1}{n!}\sum_{k=0}^{-1-m}\sum_{s=0}^{n-1}\binom{n-1}{s}B^n z^{ms-k}w^{m(n-s)+k+1}(-1)^{(m+1)(n-s)+k} \end{aligned}$$

所以可得$W_S(\mathbf{q})$的表达式为

$$W_S(\mathbf{q})=\sum_{n=0}^{\infty}\sum_{k=0}^{-1-m}\sum_{s=0}^{n}\frac{1}{(n+1)!}\binom{n}{s}(-1)^{(m+1)(n+1-s)+k}(\mathbf{q}_{-m(n+1-s)-k-1},B^{n+1}\mathbf{q}_{k-ms})$$

故欲证$T_1(1)=W_S(\mathbf{q})/(2\hbar)$，只需证

$$\partial_A^n(C)=\frac{1}{2\hbar}\sum_{k=0}^{-1-m}\sum_{s=0}^{n}\binom{n}{s}(-1)^{(m+1)(n+1-s)+k}(\mathbf{q}_{-m(n+1-s)-k-1},B^{n+1}\mathbf{q}_{k-ms})\triangleq a_{n+1} \tag{2.3}$$

下面用数学归纳法证之：当$n=0$时，$\partial_A^0(C)=C$，此时(2.3)式最右端化为：

$$-\frac{1}{2\hbar}\sum_{k=0}^{-1-m}(-1)^{m+k}q_{-1-m-k}^{\alpha}q_k^{\alpha_1}B_{\alpha\alpha_1}$$

故可知当$n=0$时(2.3)式成立. 假设结论已经对$n-1$成立，即有：

$$\partial_A^{n-1}(C)=\frac{1}{2\hbar}\sum_{k=0}^{-1-m}\sum_{s=0}^{n-1}\binom{n-1}{s}(-1)^{(m+1)(n-2)+k}\left\langle B^n\mathbf{q}_{k-ms},\mathbf{q}_{-m(n-s)-k-1}\right\rangle=a_n$$

则要证$\partial_A^n(C)=a_{n+1}$

$$\begin{aligned}\partial_A^n(C)&=\partial_A(\partial_A^{n-1})\\&=\frac{1}{2\hbar}\sum_{k=0}^{-1-m}\sum_{s=0}^{n-1}\binom{n-1}{s}\Big[(-1)^{(m+1)(n-s)+k}\left\langle B^{n+1}\mathbf{q}_{k-(m+1)s},\mathbf{q}_{-m(n-s)-k-1}\right\rangle+\\&\quad+(-1)^{(m+1)(n+1-s)+k}\left\langle B^{n+1}\mathbf{q}_{k-ms},\mathbf{q}_{-m(n+1-s)-k-1}\right\rangle\Big]\end{aligned} \tag{2.4}$$

在(2.4)式中令$t=s+1$便得到

$$\partial_A^n(C)=\frac{1}{2\hbar}\sum_{k=0}^{\infty}\sum_{s=0}^{n}\binom{n}{s}(-1)^{(m+1)(n+1-s)+k}\left\langle B^{n+1}\mathbf{q}_{k-ms},\mathbf{q}_{-m(n+1-s)-k-1}\right\rangle$$

(iii.) $T_n=(T_1(1))^nT_0/(n!)$

右边化为

$$\left(\sum_{k=0}^{\infty}\frac{\partial_A^k(C)}{(k+1)!}\right)^n\left(\sum_{s=0}^{\infty}\frac{\partial_A^s}{s!}\right)\frac{1}{n!}$$

于是，形如$\partial_A^{r_1}(C)\cdots\partial_A^{r_n}(C)\partial_A^{r_{n+1}}$的系数为（注：$r_1,\ldots,r_n$可以轮换）

$$n!\frac{1}{(r_1+1)!}\cdots\frac{1}{(r_n+1)!}\frac{1}{r_{n+1}}\frac{1}{n!}=\frac{1}{(r_1+1)!\cdots(r_n+1)!r_{n+1}!}$$

左边化为

$$\sum_{k=n}^{\infty}\sum_{\substack{\alpha_1+\cdots+\alpha_{n+1}\\=n+k}}\frac{\partial_A^{\alpha_1}C\partial_A^{\alpha_2}C\cdots C\partial_A^{\alpha_{n+1}}}{k!}$$

利用引理3得

$$\partial_A^{\alpha_1}C\partial_A^{\alpha_2}C\cdots C\partial_A^{\alpha_{n+1}}=\sum_{s_1=0}^{\alpha_1}\binom{\alpha_1}{s_1}\partial_A^{\alpha_1-s_1}(C)\partial_A^{s_1+\alpha_2}C\cdots C\partial_A^{\alpha_{n+1}}$$

$$= \sum_{s_1-0}^{\alpha_1} \sum_{s_2=0}^{s_1+\alpha_2} \cdots \sum_{s_n=0}^{s_{n-1}+\alpha_n} \binom{\alpha_1}{s_1} \binom{s_1+\alpha_2}{s_2} \cdots \binom{s_{n-1}+\alpha_n}{s_n} \cdot$$
$$\cdot \partial_A^{\alpha_1-s_1}(C) \partial_A^{s_1+\alpha_2-s_2}(C) \cdots \partial_A^{s_{n-1}+\alpha_n-s_n}(C) \partial_A^{s_n+\alpha_n+1}$$

同样，我们找形如$\partial_A^{r_1}(C)\cdots\partial_A^{r_n}(C)\partial_A^{r_{n+1}}$ 的系数（同样，$r_1,\ldots,r_n$可轮换），记

$$\begin{aligned} \alpha_1 - s_1 &= r_1 \\ s_1 + \alpha_2 - s_2 &= r_2 \\ &\cdots \\ s_n + \alpha_{n+1} &= r_{n+1} \end{aligned}$$

则可得

$$\begin{aligned} s_1 &= \alpha_1 - r_1 \\ s_2 &= \alpha_1 + \alpha_2 - (r_1 + r_2) \\ &\cdots \\ s_n &= (\alpha_1 + \cdots + \alpha_n) - (r_1 + \cdots + r_n) \end{aligned}$$

且

$$\alpha_1 + \cdots + \alpha_{n+1} = r_1 + \cdots + r_{n+1} = k - n, \qquad n = k - (r_1 + \cdots + r_{n+1})$$

故对应的$\partial_A^{r_1}(C)\cdots\partial_A^{r_n}(C)\partial_A^{r_{n+1}}$的固定$r_1,\ldots,r_n$的一种排列后系数为

$$I = \sum_{\alpha_1+\cdots+\alpha_{n+1}=r_1+\cdots+r_{n+1}} \frac{1}{n+r_1!+\cdots+r_{n+1}!} \binom{\alpha_1}{\alpha_1-r_1} \cdot$$
$$\cdot \binom{\alpha_2+\alpha_1-r_1}{\alpha_1+\alpha_2-r_1-r_2} \cdots \binom{\alpha_1+\cdots+\alpha_n-r_1-\cdots-r_{n-1}}{\alpha_1+\cdots+\alpha_n-r_1-\cdots-r_n}$$

由于$s_2 = s_1 - r_2 + \alpha_2$，所以$s_1 - (r_2 + s_2) = -\alpha_2 \le 0$，故$s_1 \le r_2 + s_2$.同理可知$s_2 \le r_3 + s_3$，…，$s_{n-1} \le r_n + s_n$.所以

$$I = \sum_{s_n=0}^{r_{n+1}} \sum_{s_{n-1}=0}^{r_n+s_n} \cdots \sum_{s_1=0}^{r_2+s_2} \frac{1}{(n+r_1+\cdots+r_{n+1})!} \binom{r_1+s_1}{r_1} \binom{r_2+s_2}{r_2} \cdots \binom{r_n+s_n}{r_n}$$
$$= \frac{1}{(n+r_1+\cdots+r_{n+1})!} \sum_{s_n=0}^{r_{n+1}} \sum_{s_{n-1}=0}^{r_n+s_n} \cdots \sum_{s_1=0}^{r_2+s_2} \binom{r_1+s_1}{r_1} \cdots \binom{r_n+s_n}{r_n}$$
$$= \frac{1}{(n+r_1+\cdots+r_{n+1})!} \sum_{s_n=0}^{r_{n+1}} \cdots \sum_{s_2=0}^{r_3+s_3} \binom{r_1+r_2+s_2+1}{r_1+1} \binom{r_2+s_2}{r_2} \cdots \binom{r_n+s_n}{r_n}$$
$$= \frac{1}{(n+r_1+\cdots+r_{n+1})!} \sum_{s_n=0}^{r_{n+1}} \cdots \sum_{s_2=0}^{r_3+s_3}$$
$$\binom{r_1+r_2+s_2+1}{r_1+1} \binom{r_2+s_2}{r_2} \binom{r_3+s_3}{r_3} \cdots \binom{r_n+s_n}{r_n}$$
$$= \frac{1}{(n+r_1+\cdots+r_{n+1})!} \sum_{s_n=0}^{r_{n+1}} \cdots \sum_{s_2=0}^{r_3+s_3}$$

$$
\begin{aligned}
&\binom{r_1+r_2+1}{r_2}\binom{r_1+r_2+1+s_2}{r_1+r_2+1}\binom{r_3+s_3}{r_3}\cdots\binom{r_n+s_n}{r_n}\\
&=\frac{1}{k!}\sum_{s_n=0}^{r_{n+1}}\cdots\sum_{s_3=0}^{r_4+s_4}\binom{r_1+r_2+1}{r_2}\binom{r_1+r_2+r_3+s_3+2}{r_1+r_2+2}\binom{r_3+s_3}{r_3}\cdots\binom{r_n+s_n}{r_n}\\
&=\cdots\\
&\cdots\\
&=\frac{1}{k!}\binom{r_1+r_2+1}{r_2}\binom{r_1+r_2+r_3+2}{r_3}\cdots\binom{r_1+\cdots+r_{n+1}+n}{r_{n+1}}\\
&=\frac{1}{(n+r_1+\cdots+r_{n+1})!}\frac{(n+r_1+\cdots+r_{n+1}+n)!}{r_2!\cdots r_{n+1}!(r_1+1)!}\frac{(r_1+r_2+1)!\cdots(r_1+\cdots+r_n+n-1)!}{(r_1+r_2+2)!\cdots(r_1+r_2+\cdots+r_n+n)!}\\
&=\frac{1}{r_2!\cdots r_{n+1}!(r_1+1)!}\frac{1}{(r_1+r_2+2)\cdots(r_1+\cdots+r_n+n)}\\
&=\frac{(r_2+1)\cdots(r_n+1)}{(r_1+1)!\cdots(r_n+1)!r_{n+1}!\cdots(r_1+\cdots+r_n+n)}
\end{aligned}
$$

再考虑到$r_1,\dots,r_n$可以轮换，故$\partial_A^{r_1}(C)\cdots\partial_A^{r_n}(C)\partial_A^{r_{n+1}}$的系数为

$$
\sum_{\substack{r_1,\dots,r_n\\ \text{的排列}}}\frac{1}{(r_1+1)!\cdots(r_n+1)!r_{n+1}!}\frac{(r_2+1)\cdots(r_n+1)}{(r_1+r_2+2)\cdots(r_1+\cdots+r_n+n)}
$$

与右边的结果对比，只需证

$$
\sum_{\substack{r_1,\dots,r_n\\ \text{的排列}}}\frac{(r_2+1)\cdots(r_n+1)}{(r_1+r_2+2)\cdots(r_1+\cdots+r_n+n)}=1
$$

当$n=2$时结论成立.设结论对$n-1$成立，则对于左边，按固定在第n个位置的数的不同情况分别求和，有

$$
\sum_{\substack{r_1,\dots,r_n\\ \text{的排列}}}\frac{(r_2+1)\cdots(r_n+1)}{(r_1+r_2+2)\cdots(r_1+\cdots+r_n+n)}=\sum_{\substack{r_n\\ \text{的取法}}}\frac{r_n+1}{r_1+\cdots+r_n+n}=1
$$

证毕.□

定理3的大致证明： 首先$m=0$时，$B\equiv\mathbf{0},R\equiv I$，结论显然成立，以下分成三个步骤对$m>0$的情况进行证明.

(i.) (2.2)式两端关于$\hbar$的零次项相等.

再做一些定义

$$
\partial_A\triangleq\sum_{k=0}^{\infty}B^{\alpha}{}_{\beta}q_k^{\beta}\partial_{\alpha,k+m}\qquad \partial_H\triangleq\sum_{k=0}^{m-1}(-1)^kB^{\alpha\beta}\partial_{\beta,k}\partial_{\alpha,m-1-k}\qquad \partial_{\alpha,k}\triangleq\frac{\partial}{\partial q_k^{\alpha}}
$$

(ii.) (2.2)式中两端关于$\hbar$的一次项相等.

可得$V_R(\partial_{\mathbf{q}})$的表达式

$$
V_R(\partial_{\mathbf{q}})=\sum_{n=0}^{\infty}\sum_{k=0}^{m-1}\sum_{i=0}^{n}(-1)^{k-(m+1)i+n-1-m}\frac{1}{(n+1)!}\binom{n}{i}B^{\alpha_2}{}_{\alpha_1}\cdots B^{\alpha_{n+2}}{}_{\alpha_{n+1}}.
$$

$$\cdot g^{\alpha_1\alpha_0}\frac{\partial}{\partial q^{\alpha_{n+2}}_{mn-mi+k}}\frac{\partial}{\partial q^{\alpha_0}_{mi+m-1-k}}$$

下面我们记：

$$\partial_{R,s}=\sum_{k=0}^{m-1}\sum_{i=0}^{s}(-1)^{k-(m+1)i+s-1-m}\binom{s}{i}B^{\alpha_2}{}_{\alpha_1}\cdots B^{\alpha_{s+2}}{}_{\alpha_{s+1}}\cdot g^{\alpha_0\alpha_1}\frac{\partial}{\partial q^{\alpha_{s+2}}_{ms-mi+k}}\frac{\partial}{\partial q^{\alpha_0}_{mi+m-1-k}}$$

则有$V_R(\partial_{\mathbf{q}})=\sum_{s=0}^{\infty}\partial_{R,s}/[(s+1)!]$

引理4. $\partial_{R,s}(-\partial_A)=(-\partial_A)\partial_{R,s}+\partial_{R,s+1}$

引理5 $\partial_H(-\partial_A)^k=\sum_{s=0}^{k}\binom{k}{s}(-\partial_A)^{k-s}\partial_{R,s}$

(iii.) (2.2)中两端关于$\hbar$的n次项相等. 证明方法与$m<0$的情况大致相同。

另外，上述两个定理中的A对和式的特殊形式来讲也是对的，我们能够证明在定理2之中的A若形如$\sum_{m<0}B_m z^m$，其中的$B_m z^m$均是无穷小变换，那也是对的。在定理3之中的A若形如$\sum_{m\geq 0}B_m z^m$，其中的$B_m z^m$也都是无穷小变换，则其亦真。这后面两种情况的证明方法是与前面的大致相同的，这里就不再细述。

参考文献

[1] Thomas Henry Coates *Riemann-Roch Theorems in Gromov-Witten Theory*, A dissertation submitted in partial satisfaction of the requirements for the degree of Doctor of Philosophy in Mathematics in the GRADUATE DIVISION of the UNIVERSITY of CALIFORNIA at BERKELEY, Spring 2003

[2] Alexander B. Givental. *Gromov-Witten invariants and quantization of quadratic Hamiltonians*, Mosc. Math. J. 1(4):551-568,645,2001

Sites and Group Cohomology

0401 Xue Hang *†

Sheaf theory is perversive and indispensable in modern mathematics. A sheaf on a topological space can be viewed as a contravariant functor from the category of open sets to the category of sets (or abelian groups) satisfying certain gluing conditions. But sheaf theory can be formulated in a much larger context, i.e. the site. The concept of site, initiated by Grothendieck, grasps the essence of the topological space in which the sheaf theory works. In this note, the basic materials of sites and sheaves on it is presented, emphasizing the cohomology theory of sheaves.

Group cohomology is an important machinery in mathematics with a wide application to number theory, topology, abstract group theory and so on. We treated group cohomology as an example of the general theory of sites. In this sense, sheaf cohomology on sites can be viewed as a common generalization of several different aspect of mathematics. We thus see the width and power of the concept of sites.

The content of this note is as follows. Firstly, after a brief review of group cohomology, we introduce the Grothendieck's notion of topology, i.e. the site. Then we develop some cohomology theory of sheaves on a site. We apply these theories to group cohomology and eventually deduce the celebrated Lydon-Serre-Hochschild spectral sequence.

This note is by no means my own work, and everything here can be found in [2], chapter 1, thus I do not claim the source of these materials. We assume some familiarities with homological algebra over abelian categories. The reader can refer to [1] for the backgrounds.

§1 A quick review of classical group cohomology

1.1 Let G be a group and A be an abelian group. G acts on A so that A becomes a G-module (in fact, it is $\mathbb{Z}[G]$-module). Let G acts on $\mathbb{Z}$ trivially, and define the cohomology groups of G with coefficients in A to be $\mathrm{Ext}^i_{\mathbb{Z}[G]}(\mathbb{Z}, A)$, denoted by $H^i(G, A)$.

Lemma 1.2. *$H^0(G,A) = A^G$, and $H^i(G,-)$ are the right derived functors of $-^G$. Let H be a normal subgroup of G, then the group A^H is a G/H-group. Let's denote the forgetful functor from ${}_{G/H}\mathscr{M}od$ to ${}_G\mathscr{M}od$ by ρ. Then ρ has $-^H$ as its right adjoint functor.*

The proof can be found in [1] lemma 6.8.1.

One of the most important theorem in group cohomology is the following existence theorem of spectral sequence.

*Address: Department of Mathematics, University of Science and Technology of China, Hefei 230026
†E-mail: hhh8511@mail.ustc.edu.cn

Theorem 1.3. *For every normal subgroup of H of a group G, and an abelian group A, there is a first quadrant spectral sequence*

$$E_2^{pq} = H^p(G/H, H^q(H,A)) \Longrightarrow H^{p+q}(G,A) \tag{1.1}$$

The low degree terms of spectral sequence 1.1 is

$$0 \to H^1(G/H, A^H) \xrightarrow{inf} H^1(G,A) \xrightarrow{res} H^1(H,A)^{G/H} \xrightarrow{d} H^2(G/H, A^H) \xrightarrow{inf} H^2(G,A) \tag{1.2}$$

The proof of this theorem can be found in [1] theorem 6.8.2 and 6.8.3.

§2 Grothendieck's notion of topology: the site

Definition 2.1. *A* site T *consists of a category $cat(T)$ and a collection $cov(T)$ of* coverings *which we shall call the* topology on the underlying category, *i.e. families $\{U_i \to U\}_{i\in I}$ of morphisms in $cat(T)$, such that the following properties hold:*

(1) For $\{U_i \to U\}$ in $cov(T)$ and a morphism $V \to U$ in $cat(T)$, all fibre products $U_i \times V$ exist and $\{U_i \times_U V \to V\}$ is again in $cov(T)$.

(2) Given $\{U_i \to U\} \in cov(T)$, and a family $\{V_{ij} \to U_i\} \in cov(T)$ for all $i \in I$, the family $\{V_{ij} \to U\}$, obtained by composition of morphisms, also belongs to $cov(T)$.

(3) If $\phi : U' \to U$ is an isomorphism in $cat(T)$, then $\{U' \to U\} \in cov(T)$.

Definition 2.2. *A* morphism $f : T \to T'$ *of sites is a functor $f : cat(T) \to cat(T')$, of the underlying categories with the following two properties:*

(1) $\{U_i \to U\} \in cov(T)$ implies $\{f(U_i) \to f(U)\} \in cov(T')$.

(2) For $\{U_i \to U\} \in cov(T)$ and a morphism $V \to U$ in $cat(T)$ the canonical morphism

$$f(U_i \times_U V) \to f(U_i) \times_{f(U)} f(V)$$

is an isomorphism for all i.

2.3 Let T be a site and $\mathscr{Ab}$ be the category of abelian groups. An (abelian) *presheaf* on T is a contravariant functor $\mathcal{F} : T \to \mathscr{Ab}$. A morphism of presheaves is a morphism of contravariant functors.

A presheaf $\mathcal{F}$ is a *sheaf* on T if for every covering $\{U_i \to U\}$ in T the diagram

$$\mathcal{F}(U) \longrightarrow \prod_i \mathcal{F}(U_i) \rightrightarrows \prod_{i,j} \mathcal{F}(U_i \times U_j)$$

is exact in $\mathscr{Ab}$. Morphism of sheaves are defined to be morphism of presheaves.

2.4 Example Let X be a topological space. The category of open sets of X together with the usual coverings, i.e. the families $\{U_i \to U\}$ with $U = \cup U_i$, defines a site. The presheaves and sheaves on this site coincide with the usual notion of presheaves and sheaves.

§3 The category of abelian sheaves

3.1 In this section, we consider the category of abelian sheaves. The assertions are easy to understand, but the proofs are quite technical which need some machinery from homological algebra. Therefore I omit all of them, except for those straight forward ones.

First we consider the category of presheaves (on T) and denote it by $\mathcal{P}$.

Proposition 3.2. *(1) $\mathcal{P}$ is an abelian category and has enough injectives.*

(2) A sequence $\mathcal{F}' \to \mathcal{F} \to \mathcal{F}''$ of abelian sheaves is exact in $\mathcal{P}$ if and only if the sequence $\mathcal{F}'(U) \to \mathcal{F}(U) \to \mathcal{F}''(U)$ of abelian groups is exact for all $U \in cat(T)$.

Denote the category of sheaves (on T) by $\mathcal{S}$. Then there is a forgetful functor $i : \mathcal{S} \to \mathcal{P}$, and the left adjoint functor of it exists, denoted by ${}^{ad}i$ or $\#$.

We have

Theorem 3.3. *(1) $\mathcal{S}$ is an abelian category and has enough injectives.*

(2) The functor $i : \mathcal{S} \to \mathcal{P}$ is left exact and $\# : \mathcal{P} \to \mathcal{S}$ is exact.

Remark 3.4. The point here is that in general, i is not exact. The exactness of abelian sheaves are not defined to be the exactness of abelian presheaves.

3.5 Let T and T' be two sites, and let $f : T \to T'$ be a morphism of topologies. Let $\mathcal{P}$ (resp. $\mathcal{S}$) and $\mathcal{P}'$ (resp. $\mathcal{S}'$) be the categories of abelian presheaves (resp. sheaves) on T and T' respectively.

Given an abelian presheaf $\mathcal{F}'$ on T', we obtain an abelian presheaf $f^p\mathcal{F}'$ on T by $f^p\mathcal{F}'(U) = \mathcal{F}'(f(U))$ for U in $cat(T)$. Moreover, for each morphism $v' : \mathcal{F}' \to \mathcal{G}'$ of presheaves on T', we obtain a morphism $f^pv' : f^p\mathcal{F}' \to f^p\mathcal{G}'$ of presheaves on T by defining $f^p\mathcal{F}'(U) = \mathcal{F}'(f(U)) \to \mathcal{G}'(f(U)) = f^p\mathcal{G}'(U)$. The resulting functor $f^p : \mathcal{P}' \to \mathcal{P}$ is additive and exact and commutes with inductive limits.

Theorem 3.6. *The functor f^p has a left adjoint functor $f_p : \mathcal{P} \to \mathcal{P}'$, which is right exact, additive and commutes with inductive limits.*

3.7 Let's consider this kind of morphisms for sheaves. Define

$$\begin{cases} f^s : \mathcal{S}' \to \mathcal{S} \\ f_s : \mathcal{S} \to \mathcal{S}' \end{cases}$$

in an obvious notation as compositions $f^s := \# \circ f^p \circ i'$ and $f_s := \#' \circ f_p \circ i$ (cf. theorem 3.3). It is easily checked that for $\mathcal{F}' \in \mathcal{S}'$ the image $f^p(i'(\mathcal{F}))$ is already a sheaf, so that $f^s = f^p \circ i'$.

Proposition 3.8. *(1) f_s is left adjoint to f^s.*

(2) f^s is left exact.

(3) f_s is right exact and commutes with inductive limits.

The question now is when f_s is exact. Here is a general result towards this direction.

Proposition 3.9. *Let T and T' be sites, and that the underlying categories possess final objects and finite fibre products. Let $f : T \to T'$ be a morphism of sites which respects final objects and finite fibre products. Then $f_s : \mathcal{S} \to \mathcal{S}'$ is exact.*

§4 Cohomology of presheaves

4.1 Since the category $\mathcal{P}$ has enough injectives, homological algebra tells us that each left exact functor from $\mathcal{P}$ to $\mathscr{A}b$ has right derived functors.

4.2 The section functor $\Gamma_U : \mathcal{P} \to \mathscr{A}b$, defined by $\mathcal{F} \mapsto \mathcal{F}(U)$, is exact by proposition 3.2, so the right derived functors $R^q\Gamma_U = 0$ for all $q \geq 1$.

Let $\{U_i \to U\}$ be a covering in $cov(T)$. We consider the functor

$$H^0(\{U_i \to U\}, -) : \mathcal{P} \to \mathscr{A}b$$

which assigns to each abelian presheaf $\mathcal{F}$ on T the abelian group

$$H^0(\{U_i \to U\}, \mathcal{F}) = \ker(\prod_i \mathcal{F}(U_i) \rightrightarrows \prod_{i,j} \mathcal{F}(U_i \times U_j))$$

This functor is left exact, but in general not exact.

Definition 4.3. *For an abelian sheaf $\mathcal{F}$ on T, the q^{th}* Čech cohomology group with values in $\mathcal{F}$, associated with *$\{U_i \to U\}$ is defined as*

$$H^q(\{U_i \to U\}, \mathcal{F}) := R^q H^0(\{U_i \to U\}, -)(F)$$

We want to define the Čech cohomology which only depends on U, but not the covering of it. This times, we appeal to the inductive limit. We first need the following concept.

Definition 4.4. *A* refinement map of coverings of U

$$\{U'_j \to U\}_{j \in J} \longrightarrow \{U_i \to U\}_{i \in I}$$

consists of a map $\varepsilon : J \to I$ of index sets and of a family $(f_j)_{j \in J}$ of U-morphisms $f_j : U'_j \to U_{\varepsilon(j)}$.

Definition 4.5. *For $q \geq 0$ the group*

$$\check{H}^q(U, \mathcal{F}) = \varinjlim_{\mathcal{J}_U} H^q(\{U_i \to U\}, \mathcal{F})$$

where the category $\mathcal{J}_U$ is the coverings of U ordered by refinement of coverings, is the q^{th} Čech cohomology group of U with values in $\mathcal{F}$.

Theorem 4.6. *The functor $\mathcal{F} \to \check{H}^0(U, \mathcal{F})$ from the category $\mathcal{P}$ to $\mathscr{A}b$ is left exact and additive. The right derivative of this functor are given by the Čech cohomology groups.*

§5 Cohomology of sheaves

5.1 For the category of abelian sheaves $\mathcal{S}$, there are enough injectives, therefore the right derived functors exists for left exact functors.

Let's consider for a fixed $U \in cat(T)$, the section functor

$$\Gamma_U : \mathcal{S} \longrightarrow \mathscr{A}b$$

is left exact. This functor can be viewed as a composition of the left exact functor $i : \mathcal{S} \to \mathcal{P}$ and the exact section functor of presheaves.

Definition 5.2. *For an abelian sheaf $\mathcal{F}$ on T we define* the q^{th} cohomology group of U with values in $\mathcal{F}$ *by*

$$H^q(U, \mathcal{F}) := R^q\Gamma_U(\mathcal{F})$$

5.3 The section functor of abelian sheaves can be factorized as

$$\Gamma_U = H^0(\{U_i \to U\}, -) \circ i = \check{H}^0(U, -) \circ i$$

Now let's turn to the left exact functor $i : \mathcal{P} \to \mathcal{S}$. Define $\mathcal{H}^q(-) := R^q i$. For each $\mathcal{F} \in \mathcal{S}$, we obtain an presheaf $\mathcal{H}^q(\mathcal{F})$, in particular $\mathcal{H}^0(\mathcal{F}) = \mathcal{F}$, viewed as an abelian presheaf.

Proposition 5.4. *(1) For each abelian sheaf $\mathcal{F}$ and each $U \in cat(T)$, we have a canonical isomorphism*

$$\mathcal{H}^q(\mathcal{F})(U) \simeq H^q(U, \mathcal{F}) \qquad \forall q \geq 0$$

(2)For each abelian sheaf $\mathcal{F}$ and each $U \in cat(T)$, we have

$$\check{H}^0(U, \mathcal{H}^q(\mathcal{F})) = 0 \qquad \forall q > 0$$

The relations between so many derived functors are described in the following spectral sequence.

Theorem 5.5. *(1) Let $\{U_i \to U\}$ be a covering in $cov(T)$. For each $\mathcal{F} \in \mathcal{S}$ there is a spectral sequence*

$$E_2^{pq} = H^p(\{U_i \to U\}, \mathcal{H}^q(\mathcal{F})) \Longrightarrow E^{p+q}(U, \mathcal{F})$$

(2) For each $\mathcal{F} \in \mathcal{S}$ there is a spectral sequence

$$E_2^{pq} = \check{H}^p(U, \mathcal{H}^q(\mathcal{F})) \Longrightarrow E^{p+q}(U, \mathcal{F})$$

Proof. Factorize Γ_U as

$$\mathcal{S} \xrightarrow{i} \mathcal{P} \xrightarrow{H^0(\{U_i \to U\}, -)} \mathscr{A}b$$

$$\mathcal{S} \xrightarrow{i} \mathcal{P} \xrightarrow{\check{H}^0(U, -)} \mathscr{A}b$$

i maps injective object of $\mathcal{S}$ to injective objects to $\mathcal{P}$ since it has an exact left adjoint functor. Now the result follows from an application of Grothendieck spectral sequence (cf. [1] theorem 5.8.3). □

§6 The Larey spectral sequence

6.1 Let $f : T \to T'$ be a morphism of topologies. Recall that the additive functor $f^s : \mathcal{S}' \to \mathcal{S}$ is left exact and its right derived functor exists.

Proposition 6.2. *For each abelian sheaf $\mathcal{F}' \in \mathcal{S}'$ there is an isomorphism*

$$R^q f^s(\mathcal{F}') \simeq (f^p \mathcal{H}^q(\mathcal{F}'))^\#$$

Proof. Split the functor f^s as

$$\mathcal{S}' \xrightarrow{i'} \mathcal{P}' \xrightarrow{\# \circ f^p} \mathcal{S}$$

Therefore there is a spectral sequence

$$E_2^{pq} = R^p(\# \circ f^p)(\mathcal{H}^q(\mathcal{F}')) \Longrightarrow E^{p+q} = R^{p+q} f^s(\mathcal{F}')$$

Now $E_2^{pq} = 0$ for $p > 0$, since $\# \circ f^p$ is exact. Hence the edge morphism

$$E_2^{0,q} = (f^p \mathcal{H}^q(\mathcal{F}'))^\# \longleftarrow E^q = R^q f^s(\mathcal{F}')$$

are isomorphisms for all q. □

6.3 Consider the composition $T'' \xrightarrow{g} T \xrightarrow{f} T'$ of morphisms of topologies, the following formulae are immediate from the definition

$$\begin{cases} (fg)^s = g^s f^s \\ (fg)_s = f_s g_s \end{cases}$$

It can be shown that $f^s \mathcal{F}'$ is g^s-acyclic for an injective objects in $\mathcal{S}'$ via the concept of *flasque sheaf* (but we shall not present the proof here). So the general existence theorem for spectral sequence is applicable, and we obtain

Theorem 6.4. *Keep the notations as above. For all abelian sheaves $\mathcal{F}'$ on T' there is a spectral sequence*

$$E_2^{pq} = R^p g^s(R^q f^s(\mathcal{F}')) \Longrightarrow E^{p+q} = R^{p+q}(fg)^s(\mathcal{F}') \tag{6.1}$$

which is functorial in $\mathcal{F}'$.

6.5 Let $f : T \to T'$ be a morphism and let $U \in cat(T)$. Consider the site whose underlying category consists of only one object and one arrow. Denote this site by P and call it *discrete site*. Let $g : P \to T$ be the unique morphism from the discrete site P to T which send the single object in P to U. Then we have $g^s = \Gamma_U$ and $(fg)^s = \Gamma_{f(U)}$, hence

$$\begin{aligned} R^n g^s(\mathcal{F}) &= H^n(U, \mathcal{F}) && \forall \mathcal{F} \in \mathcal{S} \\ R^n (fg)^s(\mathcal{F}') &= H^n(f(U), \mathcal{F}') && \forall \mathcal{F}' \in \mathcal{S}' \end{aligned}$$

Therefore the spectral sequence 6.1 transforms into

$$E_2^{pq} = H^p(U, R^q f^s(\mathcal{F}')) \Longrightarrow E^{p+q} = H^{p+q}(f(U), \mathcal{F}') \tag{6.2}$$

This is called *Larey spectral sequence.*

§7 Applications to group cohomology

7.1 We first investigate how to endow a category C with fibre products a topology. Recall that a morphism $U \to V$ is called an epimorphism if the map $\mathrm{Hom}(V,Z) \to \mathrm{Hom}(U,Z)$ is injective for each $Z \in C$. A family of morphisms $\{U_i \to V\}$ is called *a family of effective epimorphisms* if the diagram

$$\mathrm{Hom}(V,Z) \longrightarrow \prod \mathrm{Hom}_i(U_i, Z) \rightrightarrows \prod \mathrm{Hom}_{i,j}(U_i \times_V U_j, Z)$$

is exact for each $Z \in C$. It is called *universally effective epimorphism* if $\{U_i \times_V V' \to V'\}$ is a family of effective epimorphisms for each morphism $V' \to V$ in C.

On the category C the *canonical topology* is now defined by taking for the set of coverings the collection of all families $\{U_i \to U\}$ of universally effective epimorphisms in C. The resulting site is called the *canonical site.* It can be checked directly from the definition that it is indeed a site.

7.2 Now for a group G, consider the category of (left) G-modules. It Denote the canonical site of this category by T_G. It is immediate to check that the family of morphisms $\{U_i \xrightarrow{\varphi_i} U\}$ in the category of G-modules is universally effective epimorphic if and only if $U = \cup_i \varphi_i(U_i)$. Note that a homomorphism $G \to H$ of groups induces naturally on each left H-modules the structure of G-modules, and this operation is a morphism $T_H \to T_G$ of sites.

The most important property of the canonical site of G-modules is the following

Theorem 7.3. *The category of left G-modules is equivalent to the category of abelian sheaves on the canonical site T_G. The equivalence is given by the mutually quasi-inverse functors $A \mapsto \mathrm{Hom}_G(-,A)$ and $F \mapsto F(G)$.*

7.4 Let G be a group which acts on $\mathbb{Z}$ trivially and A be an G module. For an abelian sheaf $\mathrm{Hom}_G(-,A)$ belonging to A, we obtain

$$\Gamma_{\mathbb{Z}}(\mathrm{Hom}_G(-,A)) = \mathrm{Hom}_G(\mathbb{Z}, A) = A^G$$

We see that under the equivalence of categories the section functor $\Gamma_{\mathbb{Z}}$ gets identified with the functor $A \mapsto A^G$. Therefore we obtain for the q^{th} cohomology group of $\mathbb{Z}$ with values in $\mathrm{Hom}_G(-,A)$

$$H^q(\mathbb{Z}, \mathrm{Hom}_G(-,A)) = H^q(G,A)$$

where the righthand side is the usual group cohomology.

Note also that for a normal subgroup H of G the application of section functor $\Gamma_{G/H}$ to $\mathrm{Hom}_G(-,A)$ results in the following

$$\Gamma_{G/H}(\mathrm{Hom}_G(-,A)) = \mathrm{Hom}_G(G/H, A) = A^H$$

Hence we obtain

$$H^q(G/H, \mathrm{Hom}_G(-,A)) = H^q(H,A)$$

where the righthand side is again the usual group cohomology (view A as both a G module and an H module).

7.5 Let $\pi : G' \to G$ be a homomorphism of groups. As we indicated in 7.2, π induces a morphism $\tilde{\pi} : T_G \to T_{G'}$ of sites, and hence the functors

$$\begin{cases} \pi_* : \mathcal{S}_{G'} \longrightarrow \mathcal{S}_G \\ \pi^* : \mathcal{S}_G \longrightarrow \mathcal{S}_{G'} \end{cases}$$

between the categories $\mathcal{S}_G$ and $\mathcal{S}_{G'}$ of left G-modules or left G'-modules respectively. Here we use the notation that $\pi_* = \tilde{\pi}^s$ and $\pi^* = \tilde{\pi}_s$. More explicitly, we have:

For a left G'-module A' the left G-module $\pi_* A'$ equals $\mathrm{Hom}_{G'}(G, A')$, where G acts on $\pi_* A$ by $(ga)(h) = a(hg)$, $g, h \in G$, $a \in \pi_* A'$. If A is a left G-module, we have $\pi^* A = A$ with the left G'-module structure induced by $\pi : G' \to G$.

7.6 Keep the notation as above, the Larey spectral sequence 6.2 yields the following spectral squence in group cohomology:

For all left G'-modules A' there is a spectral sequence

$$E_2^{pq} = H^p(G, R^q\pi_*(A')) \Longrightarrow E^{p+q} = H^{p+q}(G', A') \tag{7.1}$$

which is functorial in A'.

As a special case let's take $\pi : G \to G/H$ where H is a normal subgroup of G. For each G-module A, we have

$$\pi_* A = \mathrm{Hom}_G(G/H, A) = A^H$$

hence

$$R^q\pi_*(A) = H^q(H, A)$$

The above spectral sequence 7.1 thus reads

$$E_2^{pq} = H^p(G/H, H^q(H, A)) \Longrightarrow H^{p+q}(G, A)$$

This is precisely the spectral sequence 1.1.

As another special case we consider $\pi : H \to G$ of a subgroup of H of G. It is easy to see that π_* is exact, and hence $E_2^{pq} = 0$ for $q > 0$. This yields *Shapiro's Lemma*, namely that the edge morphisms $E_2^{p,0} \to E^p$ are isomorphisms (cf. [1] lemma 6.3.2).

References

[1] *An Introduction to Homological Algebra*, Charles A. Weibel, Cambridge University Press, Cambridge 1994.

[2] *Introduction to Étale Cohomology*, Güter Tamme, Translated by Manfred Kolster, Spring-Verlag, Berlin Heidelberg 1994.

[3] *Grothendieck Topologies*, Mimeographed notes, Harvard University 1962

[4] *Categories and Homological Algebra*, Lecture note in University of Paris VI, Paris 2006

线性空间与Hilbert空间维数定义的合理性*

0501　姚成建

摘要

本文讨论了线性空间的极大无关组的基数问题，从而说明了将极大无关组的基数定义为它的代数维数（dim）是合理的。讨论了Hilbert空间的规范的完备正交集的基数定义为它的内积维数（Dim）是合理的。

§1　Hamel基与Schauder基

定义1.1. 线性空间$\mathbb{X}$的任一极大线性无关组称为$\mathbb{X}$的一个*Hamel*基。

定义1.2. *Hilbert*空间$\mathbb{H}$的任一规范的完备正交集称为$\mathbb{H}$的一个*Schauder*基。

线性空间$\mathbb{X}$的Hamel基的存在性与Hilbert空间$\mathbb{H}$的Schauder基的存在性都可以通过Zorn引理证明。

$\mathbb{K}^n$（$\mathbb{K}$为$\mathbb{C}$或者$\mathbb{R}$）的Hamel基$\{e_1, e_2, \cdots, e_n\}$同时也是Schauder基。

$P[x] = \{p(x)$为复系数多项式$\}$的Hamel基为$\{1, x, x^2, \cdots, x^n, \ldots\}$。

§2　主要定理

命题2.1. A, B为$k-$线性空间$\mathbb{X}$的两个*Hamel*基，则$\overline{\overline{A}} = \overline{\overline{B}}$。

证明. 1. 若$\overline{\overline{A}} < \aleph_0$，则$\overline{\overline{B}} < \aleph_0$，否则因只需$B$中有限个元素即可表出$A$中所有元素，将$B$中余下的元素任取一个添加到$A$中，便得到一个真包含$A$的线性无关组，这与$A$的极大性矛盾。设$A = \{e_1, e_2, \cdots, e_n\}, B = \{f_1, f_2, \cdots, f_m\}$，$A$可由$B$线性表出，所以$n \leq m$。从而$n = m$。

2. 若$\overline{\overline{A}} \geq \aleph_0$。下面建立映射$\varphi$: $B \to \bigcup_{n=1}^{\infty} A^n$。$\forall y \in B$，$\exists! x_1, x_2, \cdots, x_n \in A$，$y = a_1x_1 + \cdots + a_nx_n$，$a_i \in k$，$a_i \neq 0$，$y \mapsto (x_1, x_2, \cdots, x_n)$。显然，对于每一个$(x_1, x_2, \cdots, x_n) \in \bigcup_{n=1}^{\infty}$至多有$n$个原像，所以$\overline{\overline{B}} \leq \overline{\overline{\bigcup_{n=1}^{\infty} A^n \times \{1, 2, \cdots, n\}}} = \sum_{i=1}^{\infty} n\overline{\overline{A^n}} = \sum_{i=1}^{\infty} \overline{\overline{A}} = \overline{\overline{A}}$。

□

上述命题指出了线性空间的Hamel基均有相同的基数，从而可以把这个基数定义为此线性空间的代数维数，用dim表示。

对于一般的Hilbert空间，对应于Schauder基是否有类似于命题2.1的结论呢？我没能得到一般的结论，但是对于可分的Hilbert空间得到了完整的结论。

*感谢杨扬同学给我的勇气和信心，以及对本文热忱的审阅。

命题2.2. *1.* 可分的*Hilbert*空间$\mathbb{H}$同构于$\mathbb{K}^n$或l^2。

2. 若$\mathbb{H} \cong \mathbb{K}^n$，则*Hamel*基与*Schauder*基可取成相同的。

3. 若$\mathbb{H} \cong l^2$，则$\dim\mathbb{H} > \aleph_0$。

4. 若$\mathbb{H} \cong l^2$，则$\mathbb{H}$的任一*Schauder*基具有可数势。

证明. 1. 由[1]第63页的定理得。

2. 由于$\mathbb{K}^n$的正交规范基$\{e_1, e_2, \cdots, e_n\}$也是$\mathbb{K}^n$的Hamel基即得。

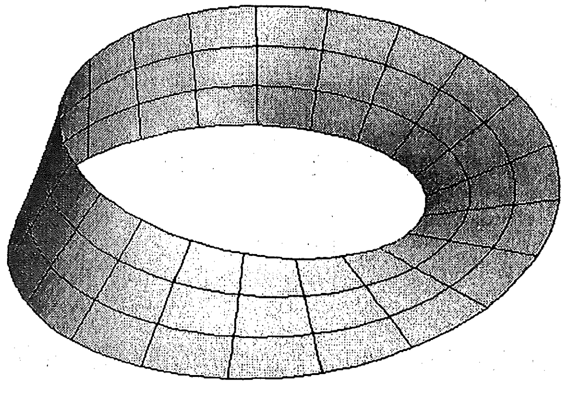
3. 利用反证法，假设$\dim\mathbb{H} = \aleph_0$。考虑$[0,1]$空间上的实系数多项式$P[0,1]$，在其中定义$(f,g) = \int_0^1 f(x)g(x)\,dx$，也即把$P[0,1]$看成$L^2[0,1]$的一个子空间，将$\{1, x, x^2, \cdots\}$进行Gram-Schmidt正交化得到一个内积空间$P[0,1]$的正交规范集$\{p_0, p_1, \cdots, p_n, \cdots\}$。设$\mathbb{H}$的一个Hamel基为$\{\tilde{e_0}, \tilde{e_1}, \cdots, \tilde{e_n}, \cdots\}$，进行Schmidt正交化得一规范正交集$\{e_0, e_1, \cdots, e_n, \cdots\}$，同时也是正交规范基。映射$\varphi$: $\mathbb{H} \to P[0,1]$，$e_i \mapsto p_i$，$\sum_{i=1}^n \lambda_i e_i \mapsto \sum_{i=1}^n \lambda_i p_i$。显然$\varphi$为线性同构，而且保持内积，从而$\mathbb{H}$的完备性也对应于$P[0,1]$的完备性。实际上由于$P[0,1]$的完备化空间为$L^2[0,1]$，从而$P[0,1]$不完备，矛盾。

4. 显然任意一个有限集不会成为l^2的规范正交基，否则l^2作为线性空间是有限维的，与3矛盾。假设$\mathbb{H}$的一个规范正交基S不可数，$E = \{e_1, e_2, \cdots\}$是$\mathbb{H}$的规范正交基，$S_i = \{f \in S | (e_i, f) \neq 0\} (i = 1, 2, \cdots)$，因$S_i$可数，所以$S \backslash \bigcup_{i=1}^{\infty} S_i \neq \emptyset$。取$f \in S \backslash \bigcup_{i=1}^{\infty} S_i \neq \emptyset$，则$(f, e_i) = 0 (i = 1, 2, \cdots)$。但$f \neq 0$，这与$E$的完备性矛盾。

□

参考文献

[1] 泛函分析讲义（上），张恭庆，林源渠，北京大学出版社，2006年



Möbius strip is a nontrivial line bundle.

凸函数两种定义之间的关系

0701 刘博睿 申国桢 袁嘉辰

摘 要

文献[1]中的268和269页给出了凸函数的几种不同的定义，其中有两种(即：$\forall x_1,x_2\in I$，$\forall t\in(0,1)$，$f(tx_1+(1-t)x_2)\le tf(x_1)+(1-t)f(x_2)$与$\forall x_1,x_2\in I$，$f(\frac{x_1+x_2}{2})\le\frac{f(x_1)+f(x_2)}{2}$)比较常见。但他们并不等价，本文深入研究了他们之间的关系，得出一个结论：后一种定义下的凸函数，如果不是前一种定义下的凸函数，那么必然在I的任意子空间中无上界。

关键词：凸函数 连续 上界

文献[1]中的268和269页给出了凸函数的几种不同的定义，其中有两种比较常见，它们分别是

定义1.1. *设I为区间，如果$\forall x_1,x_2\in I$，$\forall t\in(0,1)$，$f(tx_1+(1-t)x_2)\le tf(x_1)+(1-t)f(x_2)$，则称$f(x)$在$I$上为凸函数。*

定义1.2. *设I为区间，如果$\forall x_1,x_2\in I$，$f(\frac{x_1+x_2}{2})\le\frac{f(x_1)+f(x_2)}{2}$，则称$f(x)$在$I$上为凸函数。*

文献[1]中还指出，现代数学常采用定义1.1作为凸函数的定义，并且在271–272页给出了下面两个结论：

命题1.3. *若 f 在I°中连续，则定义1.1与定义1.2等价。*

命题1.4. *在定义1.2成立的条件下，$\forall x_1,x_2\in I$，$\forall t\in\mathbb{Q}\cap(0,1)$，$f(tx_1+(1-t)x_2)\le tf(x_1)+(1-t)f(x_2)$。*

文献[1]的第275页又证明了下面一个关键性的结论：

命题1.5. *在定义1.1成立的条件下，f在I°上连续，也就是说，*

$$\text{定义1.1}\Leftrightarrow\begin{cases}\text{定义1.2}\\ f\text{在}I^\circ\text{上连续}\end{cases}\tag{1.1}$$

在其他一些文献中也有对此问题的研究，如文献[2]的第83页给出了强于命题1.3的结论：

命题1.6. *设I为区间，f在I上连续，如果$\forall x,y\in I$，$\exists t\in(0,1)$，使得$f(tx+(1-t)y)\le tf(x)+(1-t)f(y)$，则$f(x)$在$I$上是（定义1.1中的）凸函数；反之亦真。*

很明显，在命题1.6中，令t恒为$\frac{1}{2}$，就得到了命题1.3，因此，命题1.3是命题1.6极为特殊的情况。

本文继续上述问题的结果进行研究，发现1.1式中f在I°上连续这个条件可以减得很弱。

为了叙述方便，我们就把定义1.1作为“凸函数”的定义，而把定义1.2中定义的“凸函数”叫做“2分-凸函数”。而且在以下论述中，I恒表示一个给定的区间。

我们首先引入一种结构，它有很多良好的性质：

定义1.3. $\forall x_1,x_2\in I$，$x_1\neq x_2$，我们把由x_1,x_2确定的集合$\{tx_1+(1-t)x_2|t\in\mathbb{Q}\}\cap I$定义为$E(x_1,x_2)$。

它的性质如下：

性质1.1. $\forall x_1,x_2\in I$，$x_1\neq x_2$，$E(x_1,x_2)=E(x_2,x_1)$，且$x_1,x_2\in E(x_1,x_2)$。

性质1.2. $\forall x_1,x_2\in I$，$x_1\neq x_2$，设$m\in E(x_1,x_2)$，$m\neq x_1,x_2$，则$E(m,x_1)=E(m,x_2)=E(x_1,x_2)$。

性质1.3. $\forall x_1,x_2\in I$，$x_1\neq x_2$，设$m,n\in E(x_1,x_2)$，$m\neq n$，则$E(m,n)=E(x_1,x_2)$。

性质1.4. $\forall x_1,x_2,x_3,x_4\in I$，$x_1\neq x_2$，$x_3\neq x_4$，若$E(x_1,x_2)\neq E(x_3,x_4)$，则$E(x_1,x_2)\cap E(x_3,x_4)$ 中至多只有一个元素。

性质1.5. $\forall x_1,x_2\in I$，$x_1\neq x_2$，确定$E(x_1,x_2)$，则$\forall x\in I^{\circ}$，x是$E(x_1,x_2)$的双侧聚点。

在引入了这样一个结构之后，笔者发现，2分-凸函数在任何一个给定集合$E(x_1,x_2)$中的性质，和凸函数在I中的性质完全相同。我们有：

引理1.1. 设f在I上为2分-凸函数，$m,n\in I$，$m\neq n$，则$\forall x_1,x_2,x_3\in E(m,n)$，$x_1<x_2<x_3$，有$\frac{f(x_1)-f(x_2)}{x_1-x_2}\leq\frac{f(x_1)-f(x_3)}{x_1-x_3}\leq\frac{f(x_2)-f(x_3)}{x_2-x_3}$。

证明. 由$x_2\in E(m,n)\overset{\text{性质1.3}}{=}E(x_1,x_3)$，及$x_2=\frac{x_3-x_2}{x_3-x_1}x_1+(1-\frac{x_3-x_2}{x_3-x_1})x_3$，故$\frac{x_3-x_2}{x_3-x_1}\in(0,1)\cap\mathbb{Q}$，由命题1.4，有$f(x_2)=f(\frac{x_3-x_2}{x_3-x_1}x_1+(1-\frac{x_3-x_2}{x_3-x_1})x_3)\leq\frac{x_3-x_2}{x_3-x_1}f(x_1)+(1-\frac{x_3-x_2}{x_3-x_1})f(x_3)$，在上式两端分别减去$f(x_1)$，化简就得$\frac{f(x_2)-f(x_1)}{x_2-x_1}\leq\frac{f(x_3)-f(x_1)}{x_3-x_1}$。再分别减去$f(x_3)$，化简就得$\frac{f(x_2)-f(x_3)}{x_2-x_3}\geq\frac{f(x_1)-f(x_3)}{x_1-x_3}$。于是即得：

$$\frac{f(x_2)-f(x_1)}{x_2-x_1}\leq\frac{f(x_3)-f(x_1)}{x_3-x_1}\leq\frac{f(x_2)-f(x_3)}{x_2-x_3}$$

□

引理1.2. 设f在I上为2分-凸函数，则$\forall x_0\in I^{\circ}$，取定$m\in I$，$m\neq x_0$，设$A=E(m,x_0)$，那么f关于A在x_0处连续，即$\lim_{\substack{x\to x_0\\ x\in A}}f(x)=f(x_0)$

证明. $\forall x_1,x_2\in A$，$x_1<x_2<x_0$，取定$n\in A$，$n>x_0$，则由引理1.1，有$\frac{f(x_1)-f(x_0)}{x_1-'x_0}\leq\frac{f(x_2)-f(x_0)}{x_2-x_0}\leq\frac{f(n)-f(x_0)}{n-x_0}$，故$g(x)=\frac{f(x)-f(x_0)}{x-x_0}$在$A\cap(\inf I,x_0)$中为增函数且有上

界$\dfrac{f(n)-f(x_0)}{n-x_0}$。故极限$\lim_{\substack{x\to x_0^-\\x\in A}} g(x)$即$\lim_{\substack{x\to x_0^-\\x\in A}} \dfrac{f(x)-f(x_0)}{x-x_0}$存在有限，因此，$\lim_{\substack{x\to x_0^-\\x\in A}}(f(x)-f(x_0))=\lim_{\substack{x\to x_0^-\\x\in A}}\left(\dfrac{f(x)-f(x_0)}{x-x_0}(x-x_0)\right)=0$，即$f$关于$A$在$x_0$处左连续。同理，$f$关于$A$在$x_0$处右连续，所以$f$关于$A$在$x_0$处连续。 □

利用引理1.1，引理1.2，可以得到下面的：

引理1.3. *设f在I上为2分-凸函数，$x_0\in I^\circ$，若$\exists\delta>0$，使f在$U(x_0,\delta)=(x_0-\delta,x_0+\delta)\subset I$中有界，则$f$在$x_0$处连续。*

证明. 取定$m\in A$，$m\neq x_0$，记$A=E(m,x_0)$，由引理1.2，有$\lim_{\substack{x\to x_0\\x\in A}} f(x)=f(x_0)$。设$f$在$U(x_0,\delta)$中的一个上界为$M\in\mathbb{R}$，则上极限$\limsup_{x\to x_0^-} f(x)$满足：$f(x_0)=\lim_{\substack{x\to x_0^-\\x\in A}} f(x)\le \limsup_{x\to x_0^-} f(x)\le M$，从而$\limsup_{x\to x_0^-} f(x)\in\mathbb{R}$，记为$L$。即$\exists\{c_n\}$，$c_n\in(x_0-\delta,x_0)$，$\forall n\in\mathbb{N}$，$c_n\to x_0^-(n\to\infty)$，$\lim_{n\to\infty} f(c_n)=L$。反设$f$在$x_0$处左极限不存在，则$\exists\{a_n\}$，$a_n\in(x_0-\delta,x_0)$，$\forall n\in\mathbb{N}$，$a_n\to x_0^-(n\to\infty)$，$\lim_{n\to\infty} f(a_n)=a<L$。

1. 若$a\in\mathbb{R}$，知$\exists N\in\mathbb{N}$，使得$\forall n>N$，$f(a_n)<a+\dfrac{L-a}{5}<L-\dfrac{L-a}{5}<f(c_n)$，取$k_1>N$，有$f(c_{k_1})>L-\dfrac{L-a}{5}$。由$a_n\to x_0^-(n\to\infty)$知$\exists l_1>N$，使得$c_{k_1}<a_{l_1}<x_0$，有$f(a_{l_1})<a+\dfrac{L-a}{5}$。令$b_1=2c_{k_1}-a_{l_1}$，则$2c_{k_1}-x_0<b_1<x_0$，由$f(c_{k_1})=f(\dfrac{a_{l_1}+b_1}{2})\le\dfrac{f(a_{l_1})+f(b_1)}{2}$知$f(b_1)\ge 2f(c_{k_1})-f(a_{l_1})>2(L-\dfrac{L-a}{5})-(a+\dfrac{L-a}{5})=L+\dfrac{2}{5}(L-a)$。类似构造$b_2,b_3,\dots$，得$\{c_{k_n}\},\{a_{l_n}\},\{b_n\}$满足：$k_1<k_2<\cdots<k_n$，$l_1<l_2<\cdots<l_n$，$x_0-\delta<c_{k_1}<a_{l_1}<c_{k_2}<a_{l_2}<\cdots<x_0$，$\forall n\in\mathbb{N}$，$2c_{k_n}-x_0<b_n<x_0$，$f(b_n)>L+\dfrac{2}{5}(L-a)$。由$\lim_{n\to\infty}(2c_{k_n}-x_0)=2x_0-x_0=x_0$及夹逼原理知$b_n\to x_0^-(n\to\infty)$，又$\{b_n\}$必有有极限的子列$\{b_{s_n}\}$，有$b_{s_n}\to x_0^-(n\to\infty)$，且$\lim_{n\to\infty} f(b_{s_n})\ge L+\dfrac{2}{5}(L-a)>L$，这与$L=\limsup_{x\to x_0^-} f(x)$矛盾。

2. 若$a=-\infty$，类似于1构造$\{b_n\}$推出矛盾。

所以，f在x_0处的左极限存在，且$\lim_{x\to x_0^-} f(x)=\lim_{\substack{x\to x_0^-\\x\in A}} f(x)=f(x_0)$，即$f$在$x_0$处左连续。同理，$f$在$x_0$处右连续。故$f$在$x_0$处连续。 □

引理1.4. *设f在I上为2分-凸函数，f在$x_0\in I^\circ$处连续，则$\forall x_1,x_2\in I$，$x_1<x_2<x_0$或$x_0<x_1<x_2$，有$\dfrac{f(x_1)-f(x_0)}{x_1-x_0}\le\dfrac{f(x_2)-f(x_0)}{x_2-x_0}$。*

证明. 若$x_1<x_2<x_0$，记$A=E(x_1,x_2)$，由引理1.1，$\forall x\in A\cap(x_2,x_0)$，有$\dfrac{f(x_1)-f(x_0)}{x_1-x_0}\le\dfrac{f(x_2)-f(x_0)}{x_2-x_0}$，在上式两端取$x\to x_0^-(x\in A)$时的极限，有

$$\frac{f(x_1)-f(x_0)}{x_1-x_0}=\lim_{\substack{x\to x_0^-\\x\in A}}\frac{f(x_1)-f(x)}{x_1-x}\le\lim_{\substack{x\to x_0^-\\x\in A}}\frac{f(x_2)-f(x)}{x_2-x}\le\frac{f(x_2)-f(x_0)}{x_2-x_0}$$

即得。若$x_0<x_1<x_2$，同理可得。 □

有了这4个引理之后，可以得到如下重要结论：

定理1.1. *设f在I上为2分-凸函数，若$\exists x_0 \in I^\circ$，使得f在x_0处连续，则f在I°上连续，从而推得f为I上的凸函数。*

证明. $\forall x \in (\inf I, x_0)$，取$n \in (\inf I, x)$，取$\delta > 0$，使得$U(x,\delta) \subset (n, x_0)$。则$\forall y \in U(x,\delta)$，由引理1.4，$\dfrac{f(n)-f(x_0)}{n-x_0} \le \dfrac{f(y)-f(x_0)}{y-x_0}$，从而$f(y) = f(y) - f(x_0) + f(x_0) \le \dfrac{f(n)-f(x_0)}{n-x_0}(n - x_0) + f(x_0) = \dfrac{y-x_0}{n-x_0}f(n) + (1 - \dfrac{y-x_0}{n-x_0})f(x_0) \le \max\{f(n), f(x_0)\}$，即$f$在$U(x,\delta)$中有上界。由引理1.3，$f$在$x$处连续。同理，$\forall x' \in (x_0, \sup I)$，$f$在$x'$处连续。故$f$在$I^\circ$上连续。 □

定理1.2. *设f在I上为2分-凸函数，若存在$(\alpha,\beta) \subset I$，使得f在(α,β)上有界，则f在I°上连续，从而f在I上为凸函数。*

证明. 取$x_0 \in (\alpha,\beta)$，则$x_0 \in I^\circ$且$\exists\delta > 0$，使得f在$U(x_0,\delta) \subset (\alpha,\beta)$上有上界，由引理1.3，$f$在$x_0$处连续。由定理1.1，$f$在$I^\circ$上连续。 □

以上两个定理揭示了一个惊人的结论：一个2分-凸函数，如果不是凸函数，那么在定义域的任一子区间中无上界。但是，反例仍存在：

在$\mathbb{R}$中找一族数$S = \{x_\lambda | \lambda \in \Lambda\}$使$\forall x \in \mathbb{R}$，存在唯一的$a_\lambda \in \mathbb{Q}$（$\lambda \in \Lambda, a_\lambda$中仅有有限个非零），使得$x = \sum_{\lambda\in\Lambda} a_\lambda x_\lambda$，令$f(x) = \sum_{\lambda\in\Lambda} a_\lambda f(x_\lambda) = -\sum_{\lambda\in\Lambda} a_\lambda x_\lambda^2$且$\forall x \in \mathbb{R}$，$x = \sum_{\lambda\in\Lambda} a_\lambda x_\lambda$，令$f(x) = \sum_{\lambda\in\Lambda} a_\lambda f(x_\lambda) = -\sum_{\lambda\in\Lambda} a_\lambda x_\lambda^2$，则$f(x)$为满足“2分-凸函数”定义的不连续函数，从而不是凸函数。

虽然如此，本文仍具有重要意义：首先，是2分-凸函数但不是凸函数的函数极不常见（需要在定义域的任一子区间中无上界）；其次，有了本文定理之后，人们可以大胆地使用2分-凸函数的定义来证明凸函数了（因为绝大多数函数都会在一点处连续或在定义域的某一子区间中有上界），这就为上述验证工作提供一个理论依据。

参考文献

[1] 数学分析中的典型问题与方法（第二版），裴礼文，高等教育出版社，2006年

[2] 数学分析习题及其解答，邹应，武汉大学出版社，2001年

Medal and Awards of ICCM and the Background

At each International Congress of Chinese Mathematicians (ICCM), the winners of several prestigious awards are announced at the opening ceremony. These awards include the Morningside Medal of Mathematics, the Chern Prize in Mathematics, and the ICCM International Cooperation Award. At ICCM 2007, two new prizes will be introduced: the New World Mathematics Awards and the Taikang Mathematics Awards.

Morningside Medal of Mathematics The MorningsideMedal of Mathematics is awarded to exceptional mathematicians of Chinese descent under the age of forty-five for their seminal achievements in mathematics and applied mathematics. The winners of the Morningside Medal of Mathematics are traditionally announced at the opening ceremony of the triennial International Congress of Chinese Mathematicians. The inaugural medals were presented in 1998. Each Morningside Medalist receives a certificate, a medal, and cash award of US 25, 000 for a gold medal or US10,000 for a silver medal.

Chern Prize in Mathematics The Chern Prize in Mathematics was established in 2001 in honor of Professor Shing-Shen Chern, one of the greatest geometers and Chinese mathematicians of the twentieth century. The Chern Prize is presented every three years at the International Congress of Chinese Mathematicians to mathematicians of Chinese descent who have made exceptional contributions to mathematical research or to public service activities in support of mathematics.

ICCM International Cooperation Award The ICCM International Cooperation Award is presented to an individual who has promoted the development of mathematics in the Chinese mainland, Hong Kong, and Taiwan through collaboration, teaching, and support of Chinese mathematicians. The inaugural award was presented at the 3rd International Congress of Chinese Mathematicians in 2004.

New World Mathematics Awards The New World Mathematics Awards recognize outstanding doctoral, master and undergraduate thesis written by mathematicians of Chinese descent who have graduated from universities and institutes in the past three years. The purpose is to provide encouragement to talented Chinese mathematicians and to promote creativity and innovation in mathematics. The inaugural New World Mathematics Awards will be presented at a dinner on December 17, 2007.

Taikang Mathematics Awards The Taikang Mathematics Awards recognize excellence in mathematics re- search projects among high school students of Chinese descent throughout the world. The goal is to identify gifted mathematicians at a young age and foster their interest in this field of study. An official ceremony marking the establishment of the Taikang Mathematics Awards will take place on December 18, 2007. The first awards will be presented in October 2008 in Beijing.

✐100囚徒帽子问题

在一个监狱里，有100个犯人，一天，召开全体囚徒大会。国王大赦，给大家一个机会。条件：会有人来给这100个人分别戴上帽子,帽子只有黑和白两种颜色，颜色的选择是同等概率随机的（比如用抛硬币的方法决定黑色还是白色），犯人们都不知道自己帽子是什么颜色。犯人都有机会看见所有其他犯人的脸和帽子的颜色,但自己看不见自己帽子的颜色,而且帽子带好后犯人之间不许交谈和传递信息.犯人会依次被叫到典狱长办公室里，在办公室里典狱长让每个囚犯猜自己帽子的颜色，只能回答说"黑色"或者"白色"。然后犯人被带走，下一个犯人再被叫出询问。如此这般，直到所有人都被叫出来一次为止。注意：犯人是不知道前面其他犯人的回答的。机会：最后典狱长统计一下所有犯人的回答。如果猜对自己帽子颜色的犯人数达到或超过50，那么就释放所有犯人。如果不足一半，每个犯人都只好把牢继续坐下去。

问题1： 囚徒大会后给大家20分钟时间讨论，囚徒们能找到必胜的方法么？

问题2： 试证明如果囚徒们找到了必胜的方法,那么按此方法他们中一定恰好有50个人猜对.

主编：中国科学技术大学2004级数学系
编委：管 枫 洪继展 薛 航 张伟哲
审稿：李 果 罗世森 孙 祥 王岚晖
设计：洪继展 叶 晗