中国科学: 数学 2016 年 第 46 卷 第 4 期: 371 ～ 408
SCIENTIA SINICA Mathematica

《中国科学》杂志社
SCIENCE CHINA PRESS

综 述

非标准分析及其应用

金人麟

Department of Mathematics, College of Charleston, Charleston, SC 29424, USA
E-mail: jinr@cofc.edu

收稿日期: 2015-08-12; 接受日期: 2015-12-09; 网络出版日期: 2016-03-08

摘要 本文简要介绍非标准分析基础及其在其他数学分支中的应用, 特别是在组合数论中的应用. 所介绍的基础包含数理逻辑常识、非标准模型构造以及非标准分析常用原理和性质. 所介绍的应用包含随机微分方程强解的存在性、关于局部群的 Hilbert 第五问题、精确大数定律及其在经济学中的应用、组合数论中的和集现象、关于密度的 Plünnecke 类型不等式和 Freiman 逆问题.

关键词 非标准分析 数理逻辑 随机微分方程 Hilbert 第五问题 精确大数定律 组合数论

MSC (2010) 主题分类 03H05, 03H15, 11B05, 11P70, 22E05, 60F15, 91B30, 11U10, 11B13, 22A99, 28A35, 28E05

1 引言

假设读者已学过数学分析课程或实分析课程并有一些关于测度空间的知识.

1.1 非标准分析合理性

在现实世界里, 无穷大量和无穷小量可以被认为不是一种客观存在, 而是理性思维的产物. 在通常的微积分或实分析的教科书中, 为了方便, 常引入记号 ∞ 称为无穷大, $1/\infty$ 称为无穷小. 显然, 通常所介绍的无穷大和无穷小并不是两个数, 因为它们不满足我们对数的通常要求. 例如, 数应该可以参与加减乘除运算, 且满足数域的一些规则, 如消去律等.

非标准分析的最基本特征是引入无穷大数和无穷小数. 但是为什么要在数字系统中引入无穷大数和无穷小数? 引入无穷大数和无穷小数是纯粹的逻辑游戏呢, 还是对数学在客观世界中的应用有所帮助? 微积分里介绍的 ∞ 或 $1/\infty$ 不能被视为数字, 怎么办?

数字系统的扩张是客观世界的需求和内在逻辑双重引导下的产物. 自然数集 $\mathbb{N}$ 的引进 “自然” 应该是牧民点羊、将军点卯、国王点金币等行为的需要, 自然数集上的加法必定会引出减法, 而减法又会引导人们接受零和负整数, 从而组成整数集 $\mathbb{Z}$ 使之成为加法群. 为了简化同数连加的繁琐, 在整数集 $\mathbb{Z}$ 上可引进乘法, 而乘法的引进必定导致对除法的接纳, 而除法必定会引导人们接受分数概念, 从而把数字系统扩充成了有理数域 $\mathbb{Q}$. 同样, 为了简化同数多次相乘, 整数幂运算被引进, 而幂运算的逆运算又不可避免地要求我们接受那些代数无理数, 再加上一些超越无理数的出现, 数字系统又必须被

英文引用格式: Jin R L. Nonstandard analysis and its applications (in Chinese). Sci Sin Math, 2016, 46: 371–408, doi: 10.1360/N012015-00266

扩充. 假设完备性公理成立, 即每个有上界的非空子集必有上确界, 数字系统的扩张即可一步到位, 即扩张成现在通常使用的实数系统 $\mathbb{R}$, 即一个完备的有序域, 称为实数域.

在以上数字系统的扩张过程中, 有一点很重要: 除了在旧系统中加入新的元素, 还要保证在旧系统上定义的各种数学运算能推广到新系统上, 即新系统是旧系统上的相容扩张. 例如, 在有理数域上的加法定义为 $\frac{p}{m}+\frac{q}{n}=\frac{np+mq}{mn}$. 如果 $n=m=1$, 即 $\frac{p}{m}$ 和 $\frac{q}{n}$ 是整数, 那么以上的定义就和整数上的加法一致. 例如, 对任意两个非负有理数 r_1 和 r_2, 不等式 $r_1<r_2$ 推出不等式 $r_1^2<r_2^2$. 所以, 在实数域上, 当 r_1 和 r_2 是任意两个非负实数时, 我们当然也需要有 $r_1<r_2$ 推出 $r_1^2<r_2^2$. 如果要引进无穷大数和无穷小数从而进一步扩张实数域, 我们必须把数学运算和一些性质相容地推广到新的结构上去. 例如, 我们希望扩张了的结构也是一个有序域; 希望通常的三角函数公式和其他类似的公式在新的结构中仍成立等.

在一元微分学中, 我们需要解释 $\frac{dy}{dx}$ 的意义. 直观地讲, 符号 $\frac{dy}{dx}$ 就是两个无穷小数的商. 所以有引进无穷小数的客观需要和逻辑动力. 在经典的微积分教科书中, 无穷小量 dx 和 dy 被解释为两个可比较的极限过程, 而不是数. 这当然能解决问题, 但是也使得数学的推导变得相对复杂. 但是, 如果把 dx 和 dy 定义为无穷小数, 我们就必须使 dx 和 dy 都能与其他实数比较大小, 能参与加减乘除等运算. 另外, 在我们的实数域中, 如果加了一个数 dx, 就必须再加入很多与 dx 相关的数, 如 $2+dx$, $3dx$, $\frac{1}{dx}$, $\sin(dx)$ 和 $(1+dx)^{dx}$ 等. 这会不会导致矛盾, 例如, 违反消去律或其他有序域的公理?

为了证明无穷小数是相容的, 即加入无穷小数不会与现有的实分析产生矛盾, 我们需要构造一个有序数域, 称为非标准实数域, 使得 (1) 通常的实数域是其子域; (2) 非标准实数域包含了非零正无穷小数, 即此类数中每一个数都小于任何通常的实数域中的正实数; (3) 非标准实数域满足绝大部分在通常实数域中成立的性质. 称通常的实数域为标准实数域. 构造这样的非标准实数域, 我们需要使用数理逻辑中模型论的方法.

当然, 标准实数域和非标准实数域不是相同的个体. 它们一定会满足一些不同的性质. 例如, 有理数域和实数域不是相同的结构, 虽然后者也是前者的相容扩张, 但是有理系数方程 $x^2=2$ 在实数域中有解而在有理数域中无解. 有什么性质在标准实数域中成立而在非标准实数域中不成立呢?

若读者已学过实分析课程, 则应该知道实数域的完备性阻止了实数域作为一个有序域的进一步扩张. 如果在实数域中加入无穷小数, 那么在非标准实数域中, 所有无穷小数组成的集合就是非空有上界, 但无上确界. 所以, 完备性公理在非标准实数域中必不再成立. 但这应该不会产生大问题. 一方面, 完备性公理是一条公理, 接受或不接受对数学在现实世界的应用并不产生实质影响; 另一方面, 标准实数域是非标准实数域的子域, 如有必要使用完备性公理, 我们可以退回到标准实数域中去.

作者在和他人讨论非标准分析时, 常听到初学者抱怨无穷大数或无穷小数不是客观存在的反映, 所以不见得会有用. 但这种情形其实在引进无穷小数之前就已存在. 在标准实数域中, 因为完备性, 存在不可数个实数, 但因为语言的限制, 所有可以被我们叙述的实数最多只有可数多个. 所以大部分标准实数域中的实数是我们无法叙述的, 也不是客观存在的反映. 但这并不妨碍完备性所带来的数学推导上的便利. 我们引入无穷小数的目的不仅是为了结构上的完美, 更重要的也是在引进了无穷小数后所带来的数学推导上的优势. 本文的最后两节介绍一些当前非标准分析在其他标准数学分支中的应用, 希望读者能在其中体会到非标准分析方法的有用之处.

1.2 非标准分析简史

本小节简要介绍非标准分析历史. 本文作者不是数学史专家, 本小节中的大部分信息来源于二手

资料. 读者如有兴趣可进一步参见文献 [1–5] 或其他相关著作.

在古希腊时代, Archimedes 曾使用基于直观上的无穷小数解决几何问题. 17 世纪后期, 在微积分被发明的过程中, Newton 和 Leibniz 都把无穷小数作为数字来定义导数 $\frac{dy}{dx}$ 和进行数学推导. 作为物理学家, Newton 对待无穷小数, 较从实用出发. 而作为哲学家, Leibniz 更对无穷小数的本质感兴趣. 所以, 一般都把 Leibniz 作为在严格意义上提出无穷小数概念的第一人. 但是, Newton 和 Leibniz 都没能解决无穷小数的相容性问题.

在以后将近二百年里, 研究微积分的数学家们常使用无穷小数来解决数学问题, 而把无穷小数存在的相容性搁置一边. 例如, 大数学家 Euler 曾用无穷大数来导出指数、对数和三角函数的级数公式 (参见文献 [1, 第 8 页]). Marx 在其数学手稿中也使用无穷小数并对无穷小数的合理性进行了哲学意义上的讨论. 虽然无穷小数有实用价值, 但相容性证明的缺失, 总是一个挥之不去的阴影, 以至微积分在当时受到一些人的攻击, 其中最为代表的是 18 世纪 30 年代英国著名主教 Berkeley 对无穷小数的批评. Berkeley 的观点简单地讲就是, 无穷小数作为一个非零数参与运算, 而在使导出的结果回到标准世界中无穷小数又必须像零一样被忽略, 因此, 无穷小数是介于零和非零之间的怪物, 所以, 无穷小数可能是一个自相矛盾的概念.

在 19 世纪, Weierstrass、Cauchy 和其他一些数学家在共同努力下发展出了一套极限理论, 给出了一个极限的 ϵ-δ 定义, 从而解决了微积分的基础问题. 但是, 解决的方法与 Newton 和 Leibniz 的思想并不一样. 例如, 假设 $y = f(t)$ 是一个运动物体和一个定点之间的一维距离函数, 这物体在时间 t_0 的速度 $\frac{dy}{dt}$ 等于 v_0 的 ϵ-δ 定义是

$$\forall \epsilon > 0, \quad \exists \delta > 0, \quad \forall t \left(0 < |t - t_0| < \delta \to \left| \frac{f(t) - f(t_0)}{t - t_0} - v_0 \right| < \epsilon \right), \tag{1.1}$$

即对于任意正实数 ϵ, 总存在另一个正实数 δ 使得在长度小于 δ 的时间区间上, 物体的平均速度与 v_0 的差总小于 ϵ. 这样一来, $\frac{dy}{dt}$ 就被看作为一个在不断缩小的时间区间上平均速度的极限, 而不是两个无穷小数 dy 和 dt 的商. 极限的 ϵ-δ 定义避免了无穷小数的引进, 是一个划时代的成就. 从那以后一直到 20 世纪 50 年代末, 无穷小数便基本不再出现在数学基础的讨论之中.

但是极限的 ϵ-δ 定义使用了三次交替量词符号 $\forall$, $\exists$ 和 $\forall$. 在数理逻辑中, 交替量词符号在一个逻辑公式中出现的次数是一种衡量公式复杂度的标志. 所以, 在一定意义下, 极限的 ϵ-δ 定义是比较复杂的, 也欠直观, 初学微积分的学生往往会觉得较难理解和掌握. 这作为一个普遍现象大概是因为 ϵ-δ 定义的较高复杂度. 如果 $\frac{dy}{dx}$ 被定义为两个无穷小数的商 (严格来讲, 标准实数域上可微函数 y 对自变量 x 的导数 $\frac{dy}{dx}$ 是两个无穷小数的商的标准化), 就比较简单直观, 牵涉到的量词符号也相对少. 当然, 这所谓的直观和简单是在非标准实数域已经被接受为寻常所使用的数字系统为前提的.

无穷小数相容性的建立和非标准分析的发明归功于 Robinson [6]. Robinson 是一位重要的数理逻辑学家. 在 20 世纪 50 年代末到 60 年代初, Robinson 使用了数理逻辑中紧致性定理成功构造了一个标准实数域 $\mathbb{R}$ 的扩张 ${}^*\mathbb{R}$ 使得 ${}^*\mathbb{R}$ 是个有序域, 包含了非零无穷小数, 并且任意一阶语句 $\varphi(a_1, a_2, \ldots, a_s)$ 在 $\mathbb{R}$ 中为真当且仅当在 ${}^*\mathbb{R}$ 中为真, 这里 $a_1, a_2, \ldots, a_s$ 是 s 个标准实数作为语句 φ 中的参数. 这个性质通常称为转换原理 (transfer principle). Robinson 在构造非标准实数域的同时, 还作了一些应用, 其中最有名的是与 Bernstein 一起用非标准分析方法解决了标准泛函分析中关于 Hilbert 空间中一类闭子空间存在性的一个公开问题.

为了使非标准分析的应用更加广泛, 在 20 世纪 60 年代末, Robinson 和 Zakon [7] 建立了超结构的非标准扩张. 超结构不但包含了实数域, 还包含了更高阶的数学实体, 如实数域或复数域上的函数空

间、函数空间上的算子空间、算子空间上的代数等. 从那以后, 超结构的非标准扩张就成为了大部分非标准分析学者们工作的一个基本框架.

实数结构或超结构的非标准扩张可以用模型论中的超幂构造方法来完成. 超幂构造方法在模型论中已被使用了一段时间, 但被用于构造超结构的非标准扩张则是在 20 世纪 60 年代末由 Luxemburg [8] 第一次进行. Luxemburg 还把模型论中的饱和性引入非标准分析中. 饱和性后来成了非标准分析中非常有用的工具. 在 20 世纪 70 年代初, Loeb [9] 建立了 Loeb 测度空间理论. Loeb 测度空间是非标准分析在所有与测度和概率有关的数学分支中应用的主要工具之一. 自 21 世纪初开始, 本文作者致力于非标准分析在组合数论中的应用, 也得到了很多有趣的结果.

1.3 全文计划

第 2 节简要介绍一阶谓词逻辑基本知识. 我们只涉及与非标准分析有关的逻辑语言, 其目的是为非标准模型扩张提供足够的数理逻辑基础而不使篇幅过长. 第 3 节先用超幂方法构造非标准实数域, 然后简要介绍怎样在非标准实数域中建立单变量微积分, 最后仍用超幂方法构造超结构的非标准扩张. 第 4 节介绍一些非标准分析中常用的工具, 证明一些常用的性质. 第 5 节介绍一些非标准分析在其他数学分支中的应用. 最后一节比较详细地介绍非标准分析在组合数论中的应用.

2 一阶谓词逻辑

由于篇幅的原因, 作者不打算像逻辑学教科书那样严格地介绍数理逻辑语言, 很多地方会借助于直观和一般数学家们都有的共识. 数理逻辑给数学提供了一种形式语言. 每一个数学陈述都可以转化为一个逻辑语句, 用逻辑语言来讨论数学是数学的严格化和精确化的需要.

下面要介绍的逻辑语言可分成两个方面.

2.1 逻辑语法

逻辑语言的符号可分成逻辑符号和非逻辑符号.

(1) 逻辑符号: $\to$, $\neg$, $v_1, v_2, v_3, \ldots$, $\forall$, $=$, $($, $)$.

原则上讲, 逻辑符号就是一些符号, 在介绍它们的语义之前没有含意. 但是为了直观起见, 我们把要给这些符号赋予的语义先说清楚.

$\to$ 的意思是 "推出", $\neg$ 的意思是 "不是", $v_1, v_2, v_3, \ldots$ 是一列变量, $\forall v_i$ 的意思是 "对所有的 $v_i \ldots$", $=$ 的意思是 "等同", $($, $)$ 就是通常的左右括号. $\forall$ 被称为全称量词. 我们通常用 x, y, z 等表示任意变量.

(2) 非逻辑符号 (关于实数域的): $+$, $\cdot$, $<$, 0, 1, $\{F : F \in \mathcal{F}\}$.

符号 $+$, $\cdot$, $\leqslant$, 0, 1 的意思是显然的. $\mathcal{F}$ 是一个包含关系符号的有限集使得每一个 $\mathcal{F}$ 中的符号都代表一个实数域上的有限维关系 (有序域 K 上的 n 维关系就是 K^n 的一个子集). 事实上, $+$, $\cdot$, $\leqslant$, 0 和 1 都可被看作 $\mathcal{F}$ 中的元素, 例如, $+$ 是一个二维函数, 或三维关系. 这里之所以要把它们列出是为了给读者更多的直观. 在数理逻辑中, 非逻辑符号只是符号而已, 其具体意义只有在具体数学结构中讨论时才被赋予. 严格地讲, 一个关系符号与其在所代表的具体关系是不同的, 但为了直观简便, 也因为我们只限于考虑标准实数域和它的扩张, 我们就不对它们加以区别. 假设 $\mathcal{F}$ 中的元素足够多, 即在一个特定的数学讨论中涉及的每一个 $\mathbb{R}$ 上有限维关系都被 $\mathcal{F}$ 中一个元素代表.

可以看出, 非逻辑符号是针对特定数学分支的, 而逻辑符号则独立于具体数学分支.

逻辑公式或简称公式可以用递归的方法来定义.

(关于有序域的) 公式 (formula):

(1) 设 F 是 $\mathbb{R}$ 上任意一个 s 维关系, $x_1, x_2, \ldots, x_s$ 是任意 s 个变量或 0, 1. 符号串 $F(x_1, x_2, \ldots, x_s)$ 称为原子公式. 注意, $=$ 也是一个二元关系, 所以, $x_1 = x_2$ 也是一个原子公式.

(2) 如果 ϕ 和 η 是两个公式, 那么 $\neg\phi$, $\phi \to \eta$, $\forall x\phi$ 也是公式.

如果一个符号串是通过使用有限步上述步骤 (1) 和 (2) 构造而成, 这符号串就是一个公式.

虽然公式的定义有些繁琐, 从直观上讲, 公式就是一个有数学意义的符号串, 例如, "$\forall x\ (x < y \to \neg\, y < x)$" 是一个公式, 而 "$\to x \forall \neg$" 不是一个公式.

在公式 $\forall x\ (x < y \to \neg\,(y < x))$ 中, 变量 x 称为有界变量 (即 x 被 $\forall$ 界定了), 而变量 y 称为自由变量. 公式

$$\forall x\ \forall y\ \forall z\ (\neg\,(x < y \to \neg\,(y < z)) \to x < z) \tag{2.1}$$

没有自由变量, 这是因为所有变量 x, y, z 都被 (全称) 量词 $\forall$ 界定了. 没有自由变量的公式称为语句 (sentence). 以上这个语句的含义什么呢?

为了使逻辑语言更接近我们的直觉, 我们还要引入一些可以用已有符号来表示的符号. 它们是 $\vee$ (表示 "或")、$\wedge$ (表示 "和")、$\leftrightarrow$ (表示 "当且仅当") 和 $\exists$ (表示 "存在"). 符号 $\exists$ 称为存在量词. 所以, $\phi \vee \eta$ 是 $\neg\phi \to \eta$ 的另一种形式, $\phi \wedge \eta$ 是 $\neg\,(\neg\,\phi \vee \neg\,\eta)$ 的另一种形式, $\phi \leftrightarrow \eta$ 是 $(\phi \to \eta) \wedge (\eta \to \phi)$ 的另一种形式, $\exists x\phi$ 是 $\neg\forall x\,\neg\,\phi$ 的另一种形式. 使用新引入的符号, 语句 (2.1) 可重写成

$$\forall x\ \forall y\ \forall z\ ((x < y) \wedge (y < z) \to x < z).$$

现在此语句的含义清楚了, 它叙述了序的传递性.

有时还会用到以下的简化: 设 F 是一维关系符号, ϕ 是一公式. 符号串 $(\forall x \in F)\ \phi$ 是公式 $\forall x\ (F(x) \to \phi)$ 的简化. 符号串 $(\exists x \in F)\ \phi$ 是公式 $\exists x\,(F(x) \wedge \phi)$ 的简化. 我们也常用 $(\forall x > a)\ \phi$ 来表示 $\forall x\ (x > a \to \phi)$, 常用 $(\exists x > a)\ \phi$ 来表示 $\exists x\ (x > a \wedge \phi)$. 在 ϵ-δ 定义 (1.1) 中, 我们使用了这样的简化, 即 $\forall \epsilon > 0, \exists \delta > 0, \phi$ 是 $\forall \epsilon\,(\epsilon > 0 \to \exists\,\delta\ (\delta > 0 \wedge \phi))$ 的简化.

2.2　逻辑语义

令 $\mathcal{R} = (\mathbb{R}; +, \cdot, <, 0, 1, \{F \in \mathcal{F}\})$ 为我们所考虑的实数域系统. 现在定义一个 (关于实数域的) 公式在 $\mathcal{R}$ 中的真假值.

设 $\phi(x_1, x_2, \ldots, x_s)$ 为一个公式, 其中 $x_1, x_2, \ldots, x_s$ 为 $\phi(x_1, x_2, \ldots, x_s)$ 中的所有自由变量. 公式 ϕ 在 $\mathcal{R}$ 中的真假值依赖于变量 $x_1, x_2, \ldots, x_s$ 在 $\mathbb{R}$ 中所取参数. 假设变量 $x_1, x_2, \ldots, x_s$ 在 $\mathbb{R}$ 中所取参数分别为 $r_1, r_2, \ldots, r_s \in \mathbb{R}$, 称 $\phi(r_1, r_2, \ldots, r_s)$ 为含参数 $r_1, r_2, \ldots, r_s$ 的语句. 我们用递归方法定义含参数语句在 $\mathcal{R}$ 中的真假值.

假设 ϕ 是原子公式 $F(x_1, x_2, \ldots, x_s)$, 其中 F 是 $\mathbb{R}$ 上的 s 维关系, $r_1, r_2, \ldots, r_s \in \mathbb{R}$. $F(r_1, r_2, \ldots, r_s)$ 在 $\mathbb{R}$ 中为真当且仅当 $(r_1, r_2, \ldots, r_s) \in F$, 如果 ϕ 是 $x_1 = x_2$, 则 $r_1 = r_2$ 为真当且仅当 r_1 和 r_2 是同一个数.

对于非原子公式 $\phi := \phi(r_1, r_2, \ldots, r_s)$, 如果 ϕ 是 $\neg\,\eta$, 那么 ϕ 在 $\mathcal{R}$ 中为真当且仅当 η 在 $\mathcal{R}$ 中为假. 如果 ϕ 是 $\eta \to \chi$, 那么 ϕ 在 $\mathbb{R}$ 中为真当且仅当 η 在 $\mathcal{R}$ 中为假或者 χ 在 $\mathcal{R}$ 中为真. 如果 ϕ 是 $\forall x\eta(x, r_1, r_2, \ldots, r_s)$, 那么 ϕ 在 $\mathcal{R}$ 中为真当且仅当对于所有 $\mathbb{R}$ 中元素 r, 含参数公式 $\eta(r, r_1, r_2, \ldots, r_s)$ 在 $\mathcal{R}$ 中为真.

通过对于公式的长度进行归纳, 容易证明任何一个公式 $\phi(x_1,x_2,\ldots,x_s)$ 和 $\mathbb{R}$ 中任意参数 $r_1,r_2,\ldots,r_s$, $\phi(r_1,r_2,\ldots,r_s)$ 在 $\mathcal{R}$ 中的真假值是确定的. 作为推论, 每一个语句 ϕ, 即没有自由变量的公式, 在 $\mathcal{R}$ 中的真假值是确定的.

虽然一个公式在实数域中的真假值的定义较繁琐, 但在很多情形下, 有数学能力的读者就是不看定义, 也对语句真假值的确定性没有怀疑.

2.3 何谓 "一阶"

以上所介绍的逻辑称为一阶谓词逻辑. 名称中 "谓词" 的来源是因为关系符号 F 有时也称为谓词符号. 那么 "一阶" 是什么意思呢? 我们可以用下面的语句来叙述一个实数域中的区间 $[0,1]$ 上的完备性:

$$\forall X\subseteq[0,1]\ (\exists x\in X\to\exists z\ (\forall x\in X(x\leqslant z)\wedge\forall y\ (y<z\to\exists x\in X(x>y)))).\tag{2.2}$$

语句 (2.2) 不是一阶逻辑的语句. 在这语句的开头有一个全称量词 $\forall X$, 这里变量 X 取值为区间 $[0,1]$ 的子集而不是实数域中的单个元素. 称一个公式 ϕ 是一阶的, 指的是在决定 ϕ 在一个数学结构中的真假值时, 对 ϕ 中出现的所有量词 $\forall x$ 或 $\exists x$, 变量 x 只可取此数学结构中的元素而不是子集或其他更为复杂的实体.

在定义一般非逻辑符号时, 通常还要引入另外的函数符号和常数符号. 为了简便, 我们仅引入关系符号. 事实上, 一个 s 维函数是一个 $s+1$ 维关系. 一个常数符号也称作 0 维关系.

以上对于实数域的讨论可推广到任意数学结构和数学分支. 一般来讲, 一个数学结构就是由一个集合和集合上的一些关系 (和常数、函数) 组成的. 这些关系可以用符号来代表. 这些符号组成了非逻辑符号组. 逻辑公式可以用与以上相似的递归方式来定义. 一个含参数语句在一个数学结构中的真假值也可用与以上相似的递归方式来定义. 读者如想了解更严格的数理逻辑基础知识, 可参考一些数理逻辑教科书, 如文献 [10].

3 非标准模型构造

构造非标准实数域或非标准超结构需要一个无限集 X 上的非主超滤子. 为了简单起见, 设 X 是自然数集 $\mathbb{N}$. $\mathbb{N}$ 上非主超滤子的存在依赖于选择公理. 这里将使用选择公理的的一个等价命题, 即 Zorn 引理, 来构造 $\mathbb{N}$ 上的超滤子. 关于集合论的知识可参见文献 [11].

引理 3.1 (Zorn 引理) 任给一个偏序集 $(P,<)$, 如果 P 中任意一个全序子集都有一个上界, 那么 P 中就一定有一个极大元素 b, 即 P 中没有比 b 更大的元素.

对于任意一个集合 X, 我们用 $\mathcal{P}(X)$ 表示 X 中所有子集组成的集合. $\mathcal{P}(X)$ 也称作 X 的幂集合.

定义 3.2 设 $\mathcal{U}\subseteq\mathcal{P}(\mathbb{N})$, 即 $\mathcal{U}$ 是 $\mathbb{N}$ 的幂集合的子集合. 考虑以下性质:

(1) $\emptyset\notin\mathcal{U}$;

(2) 对任意 $\mathbb{N}$ 的子集 A 和 B, $A\in\mathcal{U}$ 和 $A\subseteq B$ 推出 $B\in\mathcal{U}$;

(3) 对任意 $\mathbb{N}$ 的子集 A 和 B, $A\in\mathcal{U}$ 和 $B\in\mathcal{U}$ 推出 $A\cap B\in\mathcal{U}$;

(4) 对任意 $\mathbb{N}$ 的子集 A, 一定有 $A\in\mathcal{U}$ 或者 $\mathbb{N}\setminus A\in\mathcal{U}$, 这里 $\mathbb{N}\setminus A$ 是 A 在 $\mathbb{N}$ 中的补集;

(5) 每个 $A\in\mathcal{U}$ 都是 $\mathbb{N}$ 的无限子集.

如果 $\mathcal{U}$ 满足性质 (1)–(3), 则 $\mathcal{U}$ 称为 $\mathbb{N}$ 上的一个滤子 (filter). 如果 $\mathcal{U}$ 满足性质 (1)–(4), 则 $\mathcal{U}$ 称为 $\mathbb{N}$ 上

的一个超滤子 (ultrafilter). 如果 $\mathcal{U}$ 满足性质 (1)–(5), 则 $\mathcal{U}$ 称为 $\mathbb{N}$ 上的一个非主超滤子 (non-principal ultrafilter).

例 3.3　设 $\mathcal{U}_0 := \{\mathbb{N}\setminus A : A$ 是 $\mathbb{N}$ 中的有限子集$\}$, 则 $\mathcal{U}_0$ 是 $\mathbb{N}$ 上的非平凡滤子, 称为 Fréchet 滤子. 如果 $\mathcal{U}$ 是包含了 Fréchet 滤子的滤子, 那么 $\mathcal{U}$ 也一定是非平凡的. Fréchet 滤子不是超滤子. 如果 E 是 $\mathbb{N}$ 中所有偶数的集合, 那么 E 和 $\mathbb{N}\setminus E$ 都不是 $\mathcal{U}_0$ 中的元素, 所以 $\mathcal{U}_0$ 不满足上述性质 (4).

直观地讲, 一个非主超滤子 $\mathcal{U}$ 给出了 $(\mathbb{N};\mathcal{P}(\mathbb{N}))$ 上的一个有限可加 $\{0,1\}$- 值测度. 考虑每个 $A\subseteq\mathbb{N}$, 如果 $A\in\mathcal{U}$, 则定义 A 的测度为 1; 如果 $A\notin\mathcal{U}$, 则定义 A 的测度为 0. 因此, 以上性质 (3) 说的就是两个测度为 1 的集相交, 测度还是 1. 性质 (5) 说的就是 $\mathbb{N}$ 上的有限子集测度为 0. 另外, 由性质 (3) 不难推出 $\mathcal{U}$ 中任意有限个元的交还是 $\mathcal{U}$ 中的元. 由性质 (4) 也不难推出, 如果 $\mathbb{N}$ 是有限个互不相交的集合的并, 那么这些集合中总有唯一的一个属于 $\mathcal{U}$.

定理 3.4　假设 Zorn 引理成立, 则 $\mathbb{N}$ 上存在一个非主超滤子.

证明　首先定义一个偏序集 $(P,<)$. 设 $\mathcal{U}_0$ 是 Fréchet 滤子. 令

$$P := \{\mathcal{U}\subseteq\mathcal{P}(\mathbb{N}) : \mathcal{U} \text{ 是滤子并且 } \mathcal{U}\supseteq\mathcal{U}_0\}.$$

对于任意 $\mathcal{U}_1,\mathcal{U}_2\in P$, 定义 $\mathcal{U}_1\leqslant\mathcal{U}_2$ 当且仅当 $\mathcal{U}_1\subseteq\mathcal{U}_2$. 显然, $(P,<)$ 是一个偏序集. 如果 $\mathcal{L}=\{\mathcal{U}_i : i\in I\}$ 是 P 中的一个全序子集, 不难证明 $\mathcal{U}'=\bigcup_{i\in I}\mathcal{U}_i$ 是 $\mathcal{L}$ 在 $(P,<)$ 中的一个上界. 所以, 由 Zorn 引理可推出 $(P,<)$ 中存在一个极大元 $\mathcal{U}$. 现在只需证明 $\mathcal{U}$ 满足性质 (4).

假设存在 $E\subseteq\mathbb{N}$ 使得 $E\notin\mathcal{U}$ 并且 $E^C=\mathbb{N}\setminus E\notin\mathcal{U}$. 显然, E 和 E^C 都是无限集. 如果 $\mathcal{U}$ 中存在元素 A 使得 $X=E\cap A$ 是个有限集, 那么 $\mathbb{N}\setminus X\in\mathcal{U}_0$, 并且由性质 (3) 可知, $(\mathbb{N}\setminus X)\cap A\in\mathcal{U}$. 因为 $(\mathbb{N}\setminus X)\cap A\subseteq E^C$, 由性质 (2), $E^C\in\mathcal{U}$. 这与 $E^C\notin\mathcal{U}$ 矛盾. 所以可以假设对 $\mathcal{U}$ 中任意元素 A, $A\cap E$ 一定是无限集. 现在定义

$$\mathcal{U}' := \{X\subseteq\mathbb{N} : \text{存在 } A\in\mathcal{U} \text{ 使得 } A\cap E\subseteq X\}.$$

因为 $E\in\mathcal{U}'$, 容易证明 $\mathcal{U}'$ 是比 $\mathcal{U}$ 大的滤子, 而这与 $\mathcal{U}$ 是 P 中极大元的假设矛盾. 定理 3.4 得证.　□

如果要构造一个其他无限集合 X 上的非主超滤子, 以上方法仍然有效, 只需把 $\mathbb{N}$ 换成 X 即可.

3.1　实数域的超幂扩张

为简便起见, 我们将用 $\{a_n\}$ 来表示数列 $\{a_n : n=1,2,\ldots\}$. 用 $\{r\}$ 表示常数数列 $\{a_n=r : n=1,2,\ldots\}$.

根据 Cauchy 实数构造方法, 每个实数可以被看成有理数 Cauchy 数列的一个等价类. 设 $\{a_n\}$ 和 $\{b_n\}$ 是有理数 Cauchy 数列. 定义 $\{a_n\}$ 和 $\{b_n\}$ 等价当且仅当 $\lim_{n\to\infty}(b_n-a_n)=0$. 记 $[c_n]$ 为包含有理数 Cauchy 数列 $\{c_n\}$ 的等价类, 定义 $[a_n]+[b_n]:=[a_n+b_n]$, $[a_n]\cdot[b_n]:=[a_n\cdot b_n]$ 和 $[a_n]<[b_n]$ 当且仅当 $\lim_{n\to\infty}a_n<\lim_{n\to\infty}b_n$, 则所有有理数 Cauchy 数列的等价类组成了实数域. 所有有理数常数数列的等价类就是它的有理数子域. 我们将用类似的方法构造非标准实数域.

设 $\mathbb{R}$ 为所有标准实数的集合, $\mathbb{R}^{\mathbb{N}}$ 表示所有实数数列的集合. 令 $\mathcal{U}$ 为 $\mathbb{N}$ 上的一个给定的非主超滤子.

定义 3.5　(1) 对于任意两个实数数列 $\{r_n\},\{s_n\}\in\mathbb{R}^{\mathbb{N}}$, 令 $\{r_n\}\simeq\{s_n\}$ 当且仅当

$$\{n\in\mathbb{N} : r_n=s_n\}\in\mathcal{U};$$

(2) 对于任意实数数列 $\{r_n\} \in \mathbb{R}^{\mathbb{N}}$, 令 $[r_n] := \{\{s_n\} \in \mathbb{R}^{\mathbb{N}} : \{s_n\} \simeq \{r_n\}\}$;

(3) 令 $\mathbb{R}^{\mathbb{N}}/\mathcal{U} := \{[r_n] : \{r_n\} \in \mathbb{R}^{\mathbb{N}}\}$.

利用 $\mathcal{U}$ 的滤子性质不难验证 $\simeq$ 是 $\mathbb{R}^{\mathbb{N}}$ 上的一个等价关系. 所以, $\mathbb{R}^{\mathbb{N}}/\mathcal{U}$ 就是 $\mathbb{R}^{\mathbb{N}}$ 上所有等价类的集合. 我们所要构造的非标准实数集就是 $\mathbb{R}^{\mathbb{N}}/\mathcal{U}$. 为了更直观, 我们再给它一个记号, ${}^*\mathbb{R} := \mathbb{R}^{\mathbb{N}}/\mathcal{U}$. 接下来要在 ${}^*\mathbb{R}$ 上定义加法、乘法和序.

定义 3.6 设 $[r_n], [s_n], [t_n] \in {}^*\mathbb{R}$,

(1) $[r_n] + [s_n] = [t_n]$ 当且仅当 $\{n \in \mathbb{N} : r_n + s_n = t_n\} \in \mathcal{U}$, 或简单地, $[r_n] + [s_n] := [r_n + s_n]$;

(2) $[r_n] \cdot [s_n] = [t_n]$ 当且仅当 $\{n \in \mathbb{N} : r_n \cdot s_n = t_n\} \in \mathcal{U}$, 或简单地, $[r_n] \cdot [s_n] := [r_n \cdot s_n]$;

(3) $[r_n] < [s_n]$ 当且仅当 $\{n \in \mathbb{N} : r_n < s_n\} \in \mathcal{U}$.

以上定义 3.6(3) 定义了 $\mathbb{R}^{\mathbb{N}}/\mathcal{U}$ 上的一个全序. 这是因为对任意 $[a_n], [b_n] \in \mathbb{R}^{\mathbb{N}}/\mathcal{U}$, 三个集合

$$A = \{n \in \mathbb{N} : a_n < b_n\}, \quad B = \{n \in \mathbb{N} : a_n = b_n\}, \quad C = \{n \in \mathbb{N} : a_n > b_n\}$$

中有且只有一个属于 $\mathcal{U}$.

以下还可以把 $\mathbb{R}$ 上的所有有限维关系推广到 ${}^*\mathbb{R}$ 上去. 因为 s 维函数可被看作为特定的 $s+1$ 维关系, 以下的推广也适用于有限维实函数.

定义 3.7 设 F 是 $\mathbb{R}$ 上的 s 维关系. 定义

$${}^*F := \{([a_n^{(1)}], [a_n^{(2)}], \ldots, [a_n^{(s)}]) \in {}^*\mathbb{R}^s : \{n \in \mathbb{N} : (a_n^{(1)}, a_n^{(2)}, \ldots, a_n^{(s)}) \in F\} \in \mathcal{U}\}.$$

显然, ${}^*F$ 是 ${}^*\mathbb{R}$ 上的 s 维关系. 注意, 本文不区分关系符号和该符号表示的关系. 严格地讲, 在 ${}^*\mathbb{R}$ 上的加法应该写成 ${}^*+$ 而不是 $+$. 但是为了直观, 就不作那样的区别了. 对 $\cdot$, $<$, 0, 1 也是同样.

结构 ${}^*\mathcal{R} = ({}^*\mathbb{R}; +, \cdot, <, 0, 1, \{{}^*F : F \in \mathcal{F}\})$ 称为 $\mathcal{R}$ 的超幂 (ultrapower) 扩张, 这正是我们所要的非标准实数域系统.

现在定义关于 ${}^*\mathcal{R}$ 的公式、语句及其在 ${}^*\mathcal{R}$ 中的真假值.

关于 ${}^*\mathcal{R}$ 的原子公式仍是 ${}^*F(x_1, x_2, \ldots, x_s)$, 这里 ${}^*F$ 是 $\mathbb{R}$ 上的 s 维关系符号 F 的非标准推广. 对于任意 $([a_n^{(1)}], \ldots, [a_n^{(s)}]) \in {}^*\mathbb{R}^s$, ${}^*F([a_n^{(1)}], \ldots, [a_n^{(s)}])$ 在 ${}^*\mathcal{R}$ 中为真当且仅当 $\{n : (a_n^{(1)}, \ldots, a_n^{(s)}) \in F\} \in \mathcal{U}$.

对于任意关于 $\mathcal{R}$ 的公式 $\phi(x_1, x_2, \ldots, x_s)$, 记关于 ${}^*\mathcal{R}$ 的公式 ${}^*\phi(x_1, x_2, \ldots, x_s)$ 为公式 $\phi(x_1, x_2, \ldots, x_s)$ 的非标准化, 即 ${}^*\phi(x_1, x_2, \ldots, x_s)$ 是把所有 $\phi(x_1, x_2, \ldots, x_s)$ 中的关系符号 F 都换成 ${}^*F$ 而获得的. 例如, 关于 ${}^*\mathcal{R}$ 的原子公式 ${}^*F(x_1, x_2, \ldots, x_s)$ 就是关于 $\mathcal{R}$ 的原子公式 $F(x_1, x_2, \ldots, x_s)$ 的非标准化.

对于任意一个 ${}^*\mathcal{R}$ 的公式 ${}^*\phi(x_1, x_2, \ldots, x_s)$, 任意的 $[a_n^{(1)}], \ldots, [a_n^{(s)}] \in {}^*\mathbb{R}$, 我们可以按照上一节中同样的递归方法来定义 ${}^*\phi([a_n^{(1)}], \ldots, [a_n^{(s)}])$ 在 ${}^*\mathcal{R}$ 中的真假值. 注意公式 $\forall x {}^*\phi(x, [a_n^{(1)}], \ldots, [a_n^{(s)}])$ 在 ${}^*\mathcal{R}$ 中为真当且仅当对于所有 $x \in {}^*\mathbb{R}$, ${}^*\phi(x, [a_n^{(1)}], \ldots, [a_n^{(s)}])$ 在 ${}^*\mathcal{R}$ 中为真.

如果把每一个标准实数 r 等同于 ${}^*\mathbb{R}$ 中包含常数数列 $\{r\}$ 的等价类 $[r]$, 则 $\mathbb{R}$ 就是 ${}^*\mathbb{R}$ 的一个子集. 而且定义 3.6 在 $\mathbb{R}$ 上的限制就是 $\mathbb{R}$ 上原来的加法、乘法和序. 不难证明对所有 $F \in \mathcal{F}$, 我们有 $F \subseteq {}^*F$.

不难看出 $[n]$ 是 ${}^*\mathbb{R}$ 中的无穷大数, 即对任意一个标准实数 $[r]$, 因为 $\{n \in \mathbb{N} : r < n\} \in \mathcal{U}$, 所以 $[n] > [r]$, 并且 $[\frac{1}{n}]$ 是 ${}^*\mathbb{R}$ 中的非零正无穷小数, 即对任意标准正实数 $[r]$, 因为 $\{n \in \mathbb{N} : \frac{1}{n} < r\} \in \mathcal{U}$, 所以 $0 < [\frac{1}{n}] < [r]$. 以下要证明结构 $({}^*\mathbb{R}; +, \cdot, <, 0, 1)$ 是一个有序域. 事实上, 我们将证明一个强得多的命题, 即 Łoś 定理的一个特例.

命题 3.8 对于任意关于 $\mathcal{R}$ 的一阶公式 $\phi(x_1,\dots,x_s)$, 任意的 $[a_n^{(1)}],\dots,[a_n^{(s)}]\in {}^*\mathbb{R}$, 公式 ${}^*\phi([a_n^{(1)}],\dots,[a_n^{(s)}])$ 在 ${}^*\mathcal{R}$ 中为真当且仅当

$$\{n\in\mathbb{N}:\phi(a_n^{(1)},\dots,a_n^{(s)})\ \text{在}\ \mathcal{R}\ \text{中为真}\}\in\mathcal{U}.$$

证明 命题的证明将利用对于公式 ϕ 的复杂度进行归纳来完成. 一个公式 ϕ 或者是原子公式, 或者是 $\neg\psi$, $\eta\to\psi$, 或 $\forall x\psi(x)$, 其中 ψ 和 η 都比 ϕ 简单.

(1) 假设 ϕ 为原子公式 $F(x_1,\dots,x_s)$, 则由 ${}^*F$ 的定义推出命题.

(2) 假设 $\phi(x_1,\dots,x_s)$ 是 $\neg\psi(x_1,\dots,x_s)$, 则 ${}^*\phi([a_n^{(1)}],\dots,[a_n^{(s)}])$ 在 ${}^*\mathcal{R}$ 中为真当且仅当 ${}^*\psi([a_n^{(1)}],\dots,[a_n^{(s)}])$ 在 ${}^*\mathcal{R}$ 中为假, (由归纳假设) 当且仅当

$$\{n\in\mathbb{N}:\psi(a_n^{(1)},\dots,a_n^{(s)})\ \text{在}\ \mathcal{R}\ \text{中为真}\}\notin\mathcal{U}$$

当且仅当

$$\{n\in\mathbb{N}:\neg\psi(a_n^{(1)},\dots,a_n^{(s)})\ \text{在}\ \mathcal{R}\ \text{中为真}\}\in\mathcal{U}.$$

(3) 假设 $\phi(x_1,\dots,x_s)$ 是 $\psi(x_1,\dots,x_s)\to\eta(x_1,\dots,x_s)$, 则 ${}^*\phi([a_n^{(1)}],\dots,[a_n^{(s)}])$ 在 ${}^*\mathcal{R}$ 中为真当且仅当 ${}^*\psi([a_n^{(1)}],\dots,[a_n^{(s)}])$ 在 ${}^*\mathcal{R}$ 中为假或 ${}^*\eta([a_n^{(1)}],\dots,[a_n^{(s)}])$ 在 ${}^*\mathcal{R}$ 中为真, (由归纳假设) 当且仅当

$$\{n\in\mathbb{N}:\psi(a_n^{(1)},\dots,a_n^{(s)})\ \text{在}\ \mathcal{R}\ \text{中为假}\}\in\mathcal{U}$$

或

$$\{n\in\mathbb{N}:\eta(a_n^{(1)},\dots,a_n^{(s)})\ \text{在}\ \mathcal{R}\ \text{中为真}\}\in\mathcal{U}.$$

因为

$$\begin{aligned}\{n\in\mathbb{N}:\phi(a_n^{(1)},\dots,a_n^{(s)})\ \text{在}\ \mathcal{R}\ \text{中为真}\}=\{&n\in\mathbb{N}:\psi(a_n^{(1)},\dots,a_n^{(s)})\ \text{在}\ \mathcal{R}\ \text{中为假}\}\\ &\cup\{n\in\mathbb{N}:\eta(a_n^{(1)},\dots,a_n^{(s)})\ \text{在}\ \mathcal{R}\ \text{中为真}\},\end{aligned}$$

所以有 ${}^*\phi([a_n^{(1)}],\dots,[a_n^{(s)}])$ 在 ${}^*\mathcal{R}$ 中为真当且仅当

$$\{n\in\mathbb{N}:\phi(a_n^{(1)},\dots,a_n^{(s)})\ \text{在}\ \mathcal{R}\ \text{中为真}\}\in\mathcal{U}.$$

(4) 假设 $\phi(x_1,\dots,x_s)$ 是 $\forall x\psi(x,x_1,\dots,x_s)$.

假设 $\forall x{}^*\psi(x,[a_n^{(1)}],\dots,[a_n^{(s)}])$ 在 ${}^*\mathcal{R}$ 中为真但 $I=\{n\in\mathbb{N}:\forall x\psi(x,a_n^{(1)},\dots,a_n^{(s)})$ 在 $\mathcal{R}$ 中为真$\}\notin\mathcal{U}$. 因为 $I\notin\mathcal{U}$, 所以有 $I^C=\mathbb{N}\setminus I\in\mathcal{U}$. 设 $n\in I^C$, 则有 $\forall x\psi(x,a_n^{(1)},\dots,a_n^{(s)})$ 在 $\mathcal{R}$ 中为假. 所以存在一个实数 c_n 使得 $\psi(c_n,a_n^{(1)},\dots,a_n^{(s)})$ 在 $\mathcal{R}$ 中为假. 如果 $n\in I$, 令 $c_n=0$. 现在有了一个数列 $\{c_n\}$ 使得 $I^C\subseteq\{n\in\mathbb{N}:\psi(c_n,a_n^{(1)},\dots,a_n^{(s)})$ 在 $\mathcal{R}$ 中为假$\}\in\mathcal{U}$. 因为 ψ 的复杂度低于 ϕ, 由归纳假设, ${}^*\psi([c_n],[a_n^{(1)}],\dots,[a_n^{(s)}])$ 在 ${}^*\mathcal{R}$ 中为假. 但这与 $\forall x{}^*\psi(x,[a_n^{(1)}],\dots,[a_n^{(s)}])$ 在 ${}^*\mathcal{R}$ 中为真相矛盾.

假设 $\forall x{}^*\psi(x,[a_n^{(1)}],\dots,[a_n^{(s)}])$ 在 ${}^*\mathcal{R}$ 中为假, 但 $I=\{n\in\mathbb{N}:\forall x\psi(x,a_n^{(1)},\dots,a_n^{(s)})$ 在 $\mathcal{R}$ 中为真$\}\in\mathcal{U}$, 那么就存在一个 $[c_n]\in{}^*\mathbb{R}$ 使得 ${}^*\psi([c_n],[a_n^{(1)}],\dots,[a_n^{(s)}])$ 在 ${}^*\mathcal{R}$ 中为假. 由归纳假设, 有 $\{n\in\mathbb{N}:\psi(c_n,a_n^{(1)},\dots,a_n^{(s)})$ 在 $\mathcal{R}$ 中为假$\}\in\mathcal{U}$. 因为 $\{n\in\mathbb{N}:\psi(c_n,a_n^{(1)},\dots,a_n^{(s)})$ 在 $\mathcal{R}$ 中为假$\}$ 是 $I^C=\{n\in\mathbb{N}:\forall x\psi(x,a_n^{(1)},\dots,a_n^{(s)})$ 在 $\mathcal{R}$ 中为假$\}$ 的子集, 所以有 $I^C\in\mathcal{U}$. 但这与 $I\in\mathcal{U}$ 矛盾. 命题得证. □

定理 3.9 (转换原理) 对于任意关于实数域的一阶公式 $\phi(x_1,\dots,x_s)$, 其中 $x_1,\dots,x_s$ 是所有在 ϕ 中的自由变量, 对于任意 $r_1,\dots,r_s\in\mathbb{R}$, $\phi(r_1,\dots,r_s)$ 在 $\mathcal{R}$ 中为真当且仅当 ${}^*\phi([r_1],\dots,[r_s])$ 在 ${}^*\mathcal{R}$ 中为真.

证明 在上一节中, 对于任意标准实数 r, 我们定义了 $[r]$ 为常数数列 $\{r\}$ 的等价类. 所以, $\phi(r_1, \ldots, r_s)$ 在 $\mathcal{R}$ 中为真推出 $\{n \in \mathbb{N} : \phi(r_1, \ldots, r_s)$ 在 $\mathcal{R}$ 中为真$\} = \mathbb{N} \in \mathcal{U}$. 由命题 3.8 有 ${}^*\phi([r_1], \ldots, [r_s])$ 在 ${}^*\mathcal{R}$ 中为真. 反之, ${}^*\phi([r_1], \ldots, [r_s])$ 在 ${}^*\mathcal{R}$ 中为真推出 $\{n \in \mathbb{N} : \phi(r_1, \ldots, r_s)$ 在 $\mathcal{R}$ 中为真$\} \in \mathcal{U}$, 继而推出 $\phi(r_1, \ldots, r_s)$ 在 $\mathcal{R}$ 中为真. 定理得证. □

转换原理有一个直接推论, 即任意一个关于 $\mathcal{R}$ 的 (不含参数的) 一阶语句 ϕ 在 $\mathcal{R}$ 中为真当且仅当 ${}^*\phi$ 在 ${}^*\mathcal{R}$ 中为真. 因为所有有序域的公理都是一阶语句, 所以都在 ${}^*\mathcal{R}$ 中为真, 即 $({}^*\mathbb{R}; +, \cdot, <, 0, 1)$ 是一个有序域. 另外, 每一个三角公式 (如 $\forall x\,(\sin^2 x + \cos^2 x = 1)$) 都是 $\mathcal{R}$ 中的真语句, 所以此公式的非标准形式 (如 $\forall x\,({}^*\!\sin^2 x + {}^*\!\cos^2 x = 1)$) 是 ${}^*\mathcal{R}$ 中的真语句, 即这个三角函数公式 (如 ${}^*\!\sin^2 x + {}^*\!\cos^2 x = 1$) 在非标准实数域中成立. 总而言之, $\mathcal{R}$ 和 ${}^*\mathcal{R}$ 满足相同的一阶语句.

为了方便, 从这里开始, 我们把 ${}^*\mathcal{R}$ 中的标准实数 $[r]$ 直接就写成 r.

命题 3.10 如果 $F \subseteq \mathbb{R}$ 是有限集, 那么 ${}^*F = F$.

证明 假设 $F = \{r_1, r_2, \ldots, r_s\}$ 为一有限集. 因为一阶语句

$$\phi(r_1, r_2, \ldots, r_s) := \forall x \in F\ (x = r_1 \vee x = r_2 \vee \cdots \vee x = r_s)$$

在 $\mathcal{R}$ 中为真, 由转换原理, 语句 ${}^*\phi(r_1, r_2, \ldots, r_s)$ 在 ${}^*\mathcal{R}$ 中也为真. 所以, ${}^*F = \{r_1, r_2, \ldots, r_s\}$. □

3.2 简略版一元微积分

在非标准实数域系统 ${}^*\mathcal{R}$ 的框架下可以利用无穷小数而不是 ϵ-δ 定义来建立微积分理论. 以下简略介绍怎样定义标准数列极限、标准一元函数的连续性、导数和积分. 如需加强版, 可参见文献 [1,3,5] 等. 本小节记 r 为一标准或非标准实数, 而用希腊字母 α 和 β 等表示标准实数.

定义 3.11 对于一个实数 $r \in {}^*\mathbb{R}$,

(1) 如果对所有标准实数 α 都有 $r > \alpha$ $(r < \alpha)$, 则称 r 为无穷大 (负无穷大);

(2) 如果对所有标准正实数 $\alpha > 0$ 都有 $|r| < \alpha$, 则称 r 为无穷小, 记为 $r \approx 0$;

(3) 如果存在一个标准正实数 β 使得 $|r| < \beta$, 则称 r 为标准有界的;

(4) 给定两个实数 $r_1, r_2 \in {}^*\mathbb{R}$, 如果 $r_1 - r_2 \approx 0$, 则称 r_1 无限接近于 r_2, 并记为 $r_1 \approx r_2$. 注意, $\approx$ 是一等价关系.

例如, 如果 r 是个非零无穷小数, 那么 $5 + r$ 就是一个标准有界的非标准实数.

命题 3.12 对于任意一个标准有界的实数 r, 总存在唯一的标准实数 α 使得 $r \approx \alpha$.

证明 设 β 为标准正实数使得 $|r| < \beta$. 令 $S = \{x \in \mathbb{R} : x < r\}$. 集合 S 是 $\mathbb{R}$ 中非空 (因为 $-\beta \in S$) 有上界 (因为 S 中任何元素都小于 β) 子集, 所以 S 在 $\mathbb{R}$ 中有上确界 $\alpha \in \mathbb{R}$. 如果存在标准正实数 $\epsilon > 0$ 使得 $|\alpha - r| > \epsilon$, 那么 $\alpha - \epsilon > r$ 或者 $\alpha + \epsilon < r$. 但 $\alpha - \epsilon > r$ 与 α 是 S 的上确界矛盾, $\alpha + \epsilon < r$ 与 α 是 S 的上界矛盾. 这证明了 $r \approx \alpha$.

如果 $\alpha' \in \mathbb{R}$ 使得 $r \approx \alpha'$, 那么 $\alpha - \alpha'$ 是无穷小, 所以 $\alpha = \alpha'$. 唯一性得证. □

定义 3.13 对于任意一个标准有界实数 $r \in {}^*\mathbb{R}$, 记 $\mathrm{st}(r)$ 为 $\mathbb{R}$ 中唯一实数 α 使得 $r \approx \alpha$, 即 $\mathrm{st}(r) = \alpha$, 称 st 为标准化映照 (standard part map).

在以下四个定义中, 我们利用无穷小数来定义数列极限、连续性、导数和积分. 这些定义与标准分析中的相应定义是等价的. 有兴趣的读者不妨自己来试一下证明这些等价性. 主要使用的方法是转换原理.

定义 3.14 设 $s=\{a_n : n\in\mathbb{N}\}$ 为一个标准实数数列, 且设 $\alpha\in\mathbb{R}$. 实质上, s 是从 $\mathbb{N}$ 到 $\mathbb{R}$ 的一个函数. 设 ${}^*s=\{a_n : n\in {}^*\mathbb{N}\}$ 为此数列在非标准实数域中的扩张. 如果对于任意无穷大整数 N 都有 $a_N\approx\alpha$, 就称数列 $\{a_n : n\in\mathbb{N}\}$ 收敛于极限 $\alpha\in\mathbb{R}$.

注意, 定义 3.14 把一个标准实数数列看成一个从 $\mathbb{N}$ 到 $\mathbb{R}$ 的函数, 再考虑其在非标准实数域上的非标准扩张. 而在定义 3.5, 一个实数数列被用来生成一个等价类, 再把这等价类看成非标准实数域中的一个元素. 我们希望读者能体会这两者之间的差别, 避免把它们混淆.

定义 3.15 (1) 设 $f:(\alpha,\beta)\to\mathbb{R}$ 为一标准函数, $\gamma\in(\alpha,\beta)$. 如果对所有 $r\in^*(\alpha,\beta)$ 都有

$$r\approx\gamma\to^* f(r)\approx f(\gamma),$$

则称 f 在 γ 连续; (2) 设 $E\subseteq\mathbb{R}$ 并且 $f:E\to\mathbb{R}$ 为一标准函数. 如果对所有 $r,r'\in^* E$ 都有

$$r\approx r'\to^* f(r)\approx^* f(r'),$$

则称 f 在 E 上一致连续.

定义 3.16 设 $f:(\alpha,\beta)\to\mathbb{R}$ 为一标准函数, $\gamma\in(\alpha,\beta)$. 如果 δ 是一标准实数使得对所有非零无穷小 dx 都有

$$\frac{{}^*f(\gamma+dx)-f(\gamma)}{dx}\approx\delta,$$

则称 f 在 γ 可导, 并称 δ 为 f 在 γ 的导数, 记为 $f'(\gamma)=\delta$.

定义 3.17 设 $f:[\alpha,\beta]\to\mathbb{R}$ 为一标准分段连续函数, 令 H 为无穷大整数, 并记 $x_i=\alpha+i(\beta-\alpha)/H$ 和 $\Delta x=(\beta-\alpha)/H$. 以下定义了 f 在 $[\alpha,\beta]$ 上的定积分:

$$\int_\alpha^\beta f(x)dx:=\mathrm{st}\biggl(\sum_{i=1}^{H}{}^*f(x_i)\Delta x\biggr).$$

读者可能会说以上四个定义和标准分析中的相应定义没有本质差别. 这的确如此. 但我们还需耐心, 非标准分析方法的真正优点会在最后两节中予以介绍. 下面用非标准方法来证明几个微积分中的经典定理. 证明中所用的主要思想是把实数轴离散化.

定理 3.18 在 $\mathbb{R}$ 中的有界数列一定包含一个收敛子列.

证明 设 $s=\{a_n\}$ 为 $[\alpha,\beta]$ 中的一个实数数列, 令 ${}^*s$ 是 s 的非标准扩张. 如果存在一标准实数 δ 使得 $\{n\in{}^*\mathbb{N} : a_n=\delta\}$ 是个无限集, 那么一组一阶语句 ${}^*\phi_t(\delta)=(\forall n_1,\ldots,n_t\in{}^*\mathbb{N})\ (\exists y\in{}^*\mathbb{N})\ (\bigwedge_{i=1}^t y\neq n_i\wedge a_y=\delta)$, 对所有 $t\in\mathbb{N}$, 在 ${}^*\mathcal{R}$ 中为真, 所以由转换原理, 所有 $\phi_t(\delta)$ 在 $\mathcal{R}$ 中为真. 这些语句 $\phi_t(\delta)$ 其实在 $\mathcal{R}$ 中表示了 $\{a_n\}$ 包含一个常数子数列 $\{\delta\}$.

现在可以假设对任意给定标准实数 δ, 数列 ${}^*s=\{a_n : n\in{}^*\mathbb{N}\}$ 中最多只有有限个项等于 δ. 现在要证 ${}^*s$ 中有至少有一项 a_N 不在 $\mathbb{R}$ 中.

取自然数 m, 令 $x_i^{(m)}=\alpha+i(\beta-\alpha)/m$. 注意, $\mathbb{N}$ 是 $\mathcal{R}$ 上的一维关系. 因为一阶语句

$$\phi:=(\forall m\in\mathbb{N})\ (\exists i\in\mathbb{N}\cap[1,m])\ (\{n\in\mathbb{N} : a_n\in[x_i^{(m)},x_{i+1}^{(m)}]\}\ \text{在 }\mathbb{N}\text{ 中无界})$$

在 $\mathcal{R}$ 中为真, 所以 ${}^*\phi$ 在 ${}^*\mathcal{R}$ 中为真. 取一无穷大整数 M 和正整数 $i\leqslant M$ 使得 $\{n\in{}^*\mathbb{N} : a_n\in[x_i^{(M)}, x_{i+1}^{(M)}]\}$ 在 ${}^*\mathbb{N}$ 中无界. 因为区间 $[x_i^{(M)},x_{i+1}^{(M)}]$ 的长度是无穷小, 其中最多只有一个标准实数, 所以一定有一个 $a_N\in[x_i^{(M)},x_{i+1}^{(M)}]$ 使得 $a_N\notin\mathbb{R}$.

设 $\mathrm{st}(a_N)=\eta$ 且令 S 表示 s 中所有项的集合. 任给标准实数 $\epsilon>0$, 如果 $F=S\cap(\eta-\epsilon,\eta+\epsilon)$ 是有限的, 那么 $F={}^*F={}^*S\cap^*(\eta-\epsilon,\eta+\epsilon)$. 因为 $[x_i^{(M)},x_{i+1}^{(M)}]\subseteq^*(\eta-\epsilon,\eta+\epsilon)$, 所以 $a_N\in^* F=F\subseteq\mathbb{R}$. 这与 $a_N\notin\mathbb{R}$ 矛盾. 因此, $S\cap(\eta-\epsilon,\eta+\epsilon)$ 一定是无限的. 由 ϵ 的任意性, 我们推出 η 是 $\{a_n\}$ 中一子列的极限. 证毕. □

定理 3.19 设 $f:[\alpha,\beta]\to\mathbb{R}$ 为标准连续函数, 则存在 $\gamma\in[\alpha,\beta]$ 使得 $f(\gamma)$ 是函数值中的最大值.

证明 因为 $\phi:=(\forall n\in\mathbb{N})\,(\exists i\in\mathbb{N}\cap[0,n])\,(\forall j\in\mathbb{N}\cap[0,n])\,(f(\alpha+j(\beta-\alpha)/n)\leqslant f(\alpha+i(\beta-\alpha)/n))$ 在 $\mathcal{R}$ 中为真, 所以 ${}^*\phi$ 在 ${}^*\mathcal{R}$ 中为真. 令 H 为一无穷大整数, 则在非标准实数域中存在自然数 $i_0\leqslant H$ 使得对任何 $j=0,1,\ldots,H$, 都有 ${}^*f(\alpha+j(\beta-\alpha)/H)\leqslant^* f(\alpha+i_0(\beta-\alpha)/H)$. 令 $\gamma=\mathrm{st}(\alpha+i_0(\beta-\alpha)/H)$. 因为 $\alpha+i_0(\beta-\alpha)/H\approx\gamma$, 由 f 在 γ 的连续性, 有 $f(\gamma)\approx^* f(\alpha+i_0(\beta-\alpha)/H)$. 对任意标准实数 $x\in[\alpha,\beta]$, 存在自然数 $j\leqslant H$ 使得 $x\approx\alpha+j(\beta-\alpha)/H$. 因此有 $f(x)\approx^* f(\alpha+j(\beta-\alpha)/H)\leqslant^* f(\alpha+i_0(\beta-\alpha)/H)\approx f(\gamma)$, 即 $f(x)\leqslant f(\gamma)$. 证毕. □

以下给出一个求导的简单例子.

例 3.20 设 $f(x)=x^2$, 并且 dx 为一非零无穷小, 则对任意 $x\in\mathbb{R}$, 有

$$f'(x)=\mathrm{st}\left(\frac{(x+dx)^2-x^2}{dx}\right)=\mathrm{st}\left(\frac{2xdx+dx^2}{dx}\right)=\mathrm{st}(2x+dx)=2x.$$

虽然以上例子极其简单, 但可以从中观察到, 在非标准分析中发展一元微分理论的基本要点. 读者不妨可以自己来试一下怎样利用无穷小数来证明中值定理和 Taylor 定理等.

定理 3.21 设 f 在 $[\alpha,\beta]$ 上连续, F 在 $[\alpha,\beta]$ 上连续, 在 (α,β) 上可导并且 $F'(x)=f(x)$, 则

$$\int_\alpha^\beta f(x)dx=F(\beta)-F(\alpha).$$

证明 令 H 为一个无穷大整数, $x_i=\alpha+i(\beta-\alpha)/H$, $\Delta x=x_{i+1}-x_i$. 由中值定理, 对任意 $i=0,1,\ldots,H-1$ 都有 $x_i'\in[x_i,x_{i+1}]$ 使得 ${}^*F(x_{i+1})-{}^*F(x_i)={}^*F'(x_i')\Delta x={}^*f(x_i')\Delta x$. 因为 $x_i'\approx x_i$, 所以 ${}^*f(x_i)-{}^*f(x_i')=\iota_i$ 是无穷小数, 并且

$$\left|\sum_{i=0}^{H-1}\iota_i\Delta x\right|\leqslant\Big(\max_{0\leqslant i<H}\iota_i\Big)\sum_{i=0}^{H-1}\Delta x=\Big(\max_{0\leqslant i<H}\iota_i\Big)(\beta-\alpha)\approx 0.$$

因此,

$$\begin{aligned}\int_\alpha^\beta f(x)dx&=\mathrm{st}\left(\sum_{i=0}^{H-1}{}^*f(x_i)\Delta x\right)=\mathrm{st}\left(\sum_{i=0}^{H-1}({}^*f(x_i')+\iota_i)\Delta x\right)\\&=\mathrm{st}\left(\sum_{i=0}^{H-1}{}^*F'(x_i')\Delta x+\sum_{i=0}^{H-1}\iota_i\Delta x\right)\\&=\mathrm{st}\left(\sum_{i=0}^{H-1}({}^*F(x_{i+1})-{}^*F(x_i))+\sum_{i=0}^{H-1}\iota_i\Delta x\right)\\&=F(\beta)-F(\alpha).\end{aligned}$$

证毕. □

3.3 超结构的超幂扩张

利用无穷小数来发展微积分只是非标准分析应用中的一小部分. 为了在其他数学领域中的应用, 只考虑实数域的扩张是不够的. 例如, 实数上的 Lebesgue 测度就无法在实数域的扩张中来讨论. 现在要介绍一种结构称为超结构 (superstructure), 使得大部分数学分支中的数学实体都能在其中被找到. 超结构的思想取自于集合论.

定义 3.22 定义 $V_0 = \mathbb{R}$. 对于任意自然数 $n \in \mathbb{N}$, 定义 $V_{n+1} = V_n \cup \mathcal{P}(V_n)$. 显然, 对任意 $n \in \mathbb{N}$, 都有 $V_{n-1} \subseteq V_n$. 给定一个足够大的自然数 N, 定义

$$V(\mathbb{R}) = \bigcup_{n=0}^{N} V_n.$$

令 $\in$ 为 $V(\mathbb{R})$ 上的从属关系, 即 $a \in b$ 表示 a 是集合 b 中的元素, 则结构 $\mathcal{V} = (V(\mathbb{R}); \in)$ 称为超结构. 如果 $V(\mathbb{R})$ 中的一个元素 a 属于 V_n 但不属于 V_{n-1}, 称 a 是 $\mathcal{V}$ 中的第 n 层元素, 记为 $l(a) = n$.

显然, 对于每个实数 $r \in \mathbb{R}$, 都有 $l(r) = 0$. 对实数集 $\mathbb{R}$, 则有 $l(\mathbb{R}) = 1$. 注意, 利用集合论的符号, 一个有序对 (a, b) 可以看作是集合 $\{\{a\}, \{a, b\}\}$. 所以,

$$l((a,b)) = l(\{\{a\},\{a,b\}\}) = \max\{l(\{a\}), l(\{a,b\})\} + 1 = \max\{l(a), l(b)\} + 2.$$

一个实数上的二维关系 $P \subseteq \mathbb{R}^2$ 是一个有序实数对的集合, 每个有序实数对属于第二层, 所以 $P \subseteq V_2$, 这推出 $P \in V_3$. 如果 $E \subseteq \mathbb{R}$, $f : E \to \mathbb{R}$ 是一函数, 则 f 可看作是 $E \times \mathbb{R}$ 的子集. 所以 $E \times \mathbb{R} \in V_3$ 推出 $f \in V_4$. 至此我们可以看出实数域 $\mathcal{R}$ 是 $V(\mathbb{R})$ 中的一个元素. 因为每一个特定的数学讨论只涉及有限个数学实体, 所以只要 N 取得足够大, 这些数学实体都可以在 $V(\mathbb{R})$ 中找到.

在超结构理论中, 关系只有一个, 即二维关系 $\in$. 在定义超结构的一阶语言时, 逻辑符号不变, 但非逻辑符号就只有一个, 即 $\in$. 所有元素项就只有逻辑变量 $v_1, v_2, v_3, \ldots$ 一阶谓词逻辑公式可以递归定义, 即原子公式为 $x = y$ 或 $x \in y$, 这里 x 和 y 可以是相同或不同的变量符号, 如果 ψ 和 η 是一阶公式, 则 $\neg\psi$, $\psi \to \eta$ 和 $\forall x \psi$ 也是一阶公式.

对于任意关于超结构的一阶公式 $\phi(x_1, x_2, \ldots, x_s)$ 和 $V(\mathbb{R})$ 中元素 $a_1, a_2, \ldots, a_s$, 含参数语句 $\phi(a_1, a_2, \ldots, a_s)$ 在 $\mathcal{V}$ 中的真假值也可递归定义: $a_1 \in a_2$ 和 $a_1 = a_2$ 在 $\mathcal{V}$ 中为真的定义是自明的. 如果 $\phi(a_1, a_2, \ldots, a_s)$ 是 $\neg\psi(a_1, a_2, \ldots, a_s)$, 则 $\phi(a_1, a_2, \ldots, a_s)$ 在 $\mathcal{V}$ 中为真当且仅当 $\psi(a_1, a_2, \ldots, a_s)$ 在 $\mathcal{V}$ 中为假. 如果 $\phi(a_1, a_2, \ldots, a_s)$ 是 $\psi(a_1, a_2, \ldots, a_s) \to \eta(a_1, a_2, \ldots, a_s)$, 则 $\phi(a_1, a_2, \ldots, a_s)$ 在 $\mathcal{V}$ 中为真当且仅当 $\psi(a_1, a_2, \ldots, a_s)$ 在 $\mathcal{V}$ 中为假或 $\eta(a_1, a_2, \ldots, a_s)$ 在 $\mathcal{V}$ 中为真. 如果 $\phi(a_1, a_2, \ldots, a_s)$ 是 $\forall x \psi(x, a_1, a_2, \ldots, a_s)$, 则 $\phi(a_1, a_2, \ldots, a_s)$ 在 $\mathcal{V}$ 中为真当且仅当对所有 $a \in V(\mathbb{R})$, 含参数语句 $\psi(a, a_1, a_2, \ldots, a_s)$ 在 $\mathcal{V}$ 中为真. 注意, 在定义语句 $\forall x \psi(x, a_1, a_2, \ldots, a_s)$ 在 $\mathcal{V}$ 中的真假值时, "$\forall x \cdots$" 中的变量 x 所取的值为 $V(\mathbb{R})$ 中的元素, 这是一阶谓词逻辑所要求的. 但是 $V(\mathbb{R})$ 中的元素也可能是含有元素的集合, 例如, $\mathbb{N}$ 是 V_0 的子集, 也是 V_1 中的元素. 所以, 我们有必要强调从逻辑的角度来讲, $V(\mathbb{R})$ 中的每个个体都是元素, 而 $\in$ 仅是 $V(\mathbb{R})$ 上元素之间的一个二维关系.

一般非标准分析的书上都把超结构定义为 $(\bigcup_{n\in\mathbb{N}} V_n; \in)$, 即超结构包含了无限多层元素. 这样的定义有其便利之处. 但如果采用这样的定义, 在以后介绍非标准扩张中的转换原理时就需要对逻辑公式进行修正, 即转换原理只应用于有界公式. 本文尽量不把精力用在逻辑讨论上. 所以, 我们的定义就是这种思想指导下的产物.

其实在定义超结构时, V_0 不是一定要设成 $\mathbb{R}$. 在讨论有些数学分支时, 如点集拓扑学, 其中的数

学实体可能有很大的基数, 所以这些数学实体在 $V(\mathbb{R})$ 中有可能找不到. 所以, 一般来讲, V_0 可以设定为 $X\cup\mathbb{R}$, 这里 X 可以是任何一个集合. 在本文中, V_0 被设定为 $\mathbb{R}$ 已经足够满足我们的需要了.

以下建立超结构的超幂扩张. 设 $\mathcal{U}$ 仍是 $\mathbb{N}$ 上的一个非主超滤子. 一个 $V(\mathbb{R})$ 中的元素序列可表示为 $\{a_n\}:=\{a_n:n=0,1,\ldots\}$, 这里 $a_n\in V(\mathbb{R})$. 设 $V(\mathbb{R})^{\mathbb{N}}$ 为所有 $V(\mathbb{R})$ 中的元素组成的序列的集合.

定义 3.23 (1) 对于任意两个元素序列 $\{a_n\},\{b_n\}\in V(\mathbb{R})^{\mathbb{N}}$, 定义 $\{a_n\}\simeq\{b_n\}$ 当且仅当

$$\{n\in\mathbb{N}:a_n=b_n\}\in\mathcal{U};$$

(2) 对于任意元素序列 $\{a_n\}\in V(\mathbb{R})^{\mathbb{N}}$, 设 $[a_n]:=\{\{b_n\}\in V(\mathbb{R})^{\mathbb{N}}:\{a_n\}\simeq\{b_n\}\}$;

(3) 定义 $V(\mathbb{R})^{\mathbb{N}}/\mathcal{U}:=\{[a_n]:\{a_n\}\in V(\mathbb{R})^{\mathbb{N}}\}$.

根据以上定义, 有 $[a_n]=[b_n]$ 当且仅当 $\{n\in\mathbb{N}:a_n=b_n\}\in\mathcal{U}$. 设 a 为 $V(\mathbb{R})$ 中一元素, 设 $\{a\}$ 为常项 a 的序列, 则 $[a]$ 是 $V(\mathbb{R})/\mathcal{U}$ 中的一元素. 如果 a 和 b 是 $V(\mathbb{R})$ 中两个不同元素, 则 $[a]\neq[b]$. 所以 $V(\mathbb{R})$ 通过映照 $a\mapsto[a]$ 可以被嵌入到 $V(\mathbb{R})^{\mathbb{N}}/\mathcal{U}$ 中去.

定义 3.24 设 $[a_n],[b_n]\in V(\mathbb{R})^{\mathbb{N}}/\mathcal{U}$, 定义 $[a_n]E[b_n]$ 当且仅当 $\{n\in\mathbb{N}:a_n\in b_n\}\in\mathcal{U}$.

例 3.25 令 $A_n=[0,n]\subseteq\mathbb{R}$ 并且 $r\in\mathbb{N}$, 则有 $[r]E[A_n]$. 这是因为 $\{n\in\mathbb{N}:r\in A_n\}=\{n\in\mathbb{N}:n\geqslant r\}\in\mathcal{U}$.

如果 $a,b\in V(\mathbb{R})$, 则显然有 $[a]E[b]$ 当且仅当 $a\in b$, 所以以上定义给出的 $V(\mathbb{R})^{\mathbb{N}}/\mathcal{U}$ 上的二维关系 E 可以被看作是 $V(\mathbb{R})^{\mathbb{N}}/\mathcal{U}$ 上的 "属于" 关系. $V(\mathbb{R})^{\mathbb{N}}/\mathcal{U}$ 中的元素也可以被赋予一个在 E 的意义下的层次, 即如果 $vE[V_0]$, 则 $l(v)=0$; 如果 $vE[V_n]$ 但没有 $vE[V_{n-1}]$, 则 $l(v)=n$. 不难证明, $[V_n]$ 是 $V(\mathbb{R})^{\mathbb{N}}/\mathcal{U}$ 中所有层次小于等于 n 的元素的集合.

定义 3.26 设 $*:V(\mathbb{R})\to V(\mathbb{R})^{\mathbb{N}}/\mathcal{U}$ 为一函数使得对任意 $a\in V(\mathbb{R})$, $*(a)=[a]$.

我们常把 $*(a)$ 写成 ${}^*a$, 把 $V(\mathbb{R})^{\mathbb{N}}/\mathcal{U}$ 写成 ${}^*V$, 把 E 写成 ${}^*\!\in$, 把 $({}^*V;{}^*\!\in)$ 写成 ${}^*\mathcal{V}$. 当然, ${}^*\!\in$ 可能不是 ${}^*V$ 上真正的集合论意义上的 "属于" 关系. 例如, $1\in\mathbb{R}$ 当且仅当 $[1]\,{}^*\!\!\in{}^*\mathbb{R}$. 但是作为一个集合论意义上的集合, ${}^*\mathbb{R}=*(\mathbb{R})=[\mathbb{R}]$ 是一个与常数序列 $\{\mathbb{R}\}$ 等价的所有序列组成的等价类, 它的元素是所有 $\{A_n\}$ 使得 $\{\mathbb{R}\}\simeq\{A_n\}$. 所以, $[1]$ 不是 ${}^*\mathbb{R}$ 作为一个集合论意义下集合中的元素.

不过也可把 $({}^*V;{}^*\!\in)$ 转换成一个与之同构的结构 $(L({}^*V);\in)$ 使得二维关系 ${}^*\!\in$ 被转换成真正的属于关系 $\in$. 对于任意 $a\in{}^*V_0$, 令 $L(a)=a$, 对于 $n=0,1,\ldots,N-1$, 递归定义 L 在 ${}^*V_{n+1}\setminus{}^*V_n$ 中元素上的值使得对任意 $a\in{}^*V_{n+1}\setminus{}^*V_n$, $L(a)=\{L(x):x\,{}^*\!\!\in a\}$. 令 $L({}^*V)=\{L(a):a\in{}^*V\}$. 容易证明 L 是从 $({}^*V,{}^*\!\in)$ 到 $(L({}^*V);\in)$ 上的同构, 即 L 是一一映上的, 并且对任意的 $a,b\in{}^*V$, 有 $a\,{}^*\!\!\in b$ 当且仅当 $L(a)\in L(b)$. 结构 $(L({}^*V);\in)$ 称为结构 $({}^*V;{}^*\!\in)$ 的 Lévy 塌缩, 而 $L\circ*:V(\mathbb{R})\to L({}^*V)$ 则是从 $V(\mathbb{R})$ 到 $L({}^*V)$ 中的嵌入. 所以, $(L({}^*V);\in)$ 是超结构 $(V(\mathbb{R});\in)$ 的非标准扩张. 为了直观, 本文仍使用 ${}^*\mathcal{V}=({}^*V;{}^*\!\in)$ 为超结构的非标准扩张而不是使用其 Lévy 塌缩. 但是在 ${}^*\mathcal{V}$ 中, 我们常把 ${}^*\!\in$ 想象成属于关系. 例如, 如果 $a,b\in{}^*V$ 使得 $a\,{}^*\!\!\in b$, 我们常把 a 称为 b 的元素, 如果对于任意 ${}^*V$ 中的元素 x, 由 $x\,{}^*\!\!\in a$ 可以推出 $x\,{}^*\!\!\in b$, 我们就称 a 是 b 的一个子集. 所以常把 ${}^*\!\in$ 写成 $\in$.

注意, 如果 $r\in V_0=\mathbb{R}$, 我们可以把 r 和 ${}^*r=[r]$ 等同起来, 这是因为 r 在 $\mathcal{V}$ 中和 $[r]$ 在 ${}^*\mathcal{V}$ 中都是原子, 即不含元素的集合. 但是, 如果 A 是 $V(\mathbb{R})$ 中层次高于 0 的元素, 那么 ${}^*A$ 和 $\{[a]:a\in A\}$ 就可能不相同. 例如, $\mathbb{R}\in V_1$, $\mathbb{R}=\{[r]:r\in\mathbb{R}\}\neq{}^*\mathbb{R}$.

命题 3.27 对于所有 $A\in V(\mathbb{R})\setminus V_0$, ${}^*A=\{{}^*a:a\in A\}$ 当且仅当 A 是一个有限集.

证明 充分性的证明与命题 3.10 的证明基本一样. 下面证必要性. 如果 A 是无限集, 则 A 包含了无穷多个元素 $a^{(m)}$, $m=0,1,\ldots$ 设 $\{b_n\}$ 是一个序列使得对所有自然数 $n=0,1,\ldots$, $b_n=a^{(n)}$, 则

$\{n\in\mathbb{N}: b_n\in A\}=\mathbb{N}\in\mathcal{U}$. 因此 $[b_n]\in^* A$. 如果 $a\in A$, 则 $|\{n\in\mathbb{N}: b_n=a\}|\leqslant 1$. 所以对任意一个 $a\in A$, 都有 $[b_n]\neq[a]$, 这证明了 $^*A\neq\{^*a: a\in A\}$. □

以上命题也说明 $^*\mathcal{V}$ 是 $\mathcal{V}$ 的真扩张. 更具体一些, $^*\mathcal{V}$ 中的实数域是标准实数域的真扩张.

定义 3.28 (1) 如果 $A\in V(\mathbb{R})$, 那么 $^*A$ 称为标准集 (standard set);

(2) 如果 $B\in {}^*V$, 那么 B 称为内集 (internal set);

(3) 如果 $C\subseteq {}^*V$, 但 C 不是内集, 那么 C 称为外集 (external set);

(4) 无穷大自然数称为超整数 (hyperfinite integer);

(5) 如果 H 是一个超整数, $\mathcal{H}=\{1,2,\dots,H\}$, A 是一个内集, 并且存在一个从 $\mathcal{H}$ 到 A 的一一对应 b 使得 $b\in {}^*V$, 则 A 称为超有限集 (hyperfinite set). 如果一个内集是有限的或超有限的, 则称为 $*$ 有限的.

例 3.29 $^*\mathbb{R}$ 是标准集. 如果对每个自然数 n, 设 $A_n=[0,n]\subseteq\mathbb{R}$, 则 $[A_n]$ 是内集但不是标准集. 集合 $\mathbb{R}$ 作为 $^*\mathbb{R}$ 的子集是外集而不是内集. 容易证明 $\mathbb{R}\subseteq[A_n]\subseteq {}^*\mathbb{R}$.

$^*\mathbb{N}$ 是标准集, 但不是超有限的. 对每个 $n\in\mathbb{N}$, 设 $B_n=[0,n]\cap\mathbb{N}$, 则 $[B_n]$ 是超有限集但不是标准集. 集合 $\mathbb{N}$ 作为 $^*\mathbb{N}$ 的子集是外集而不是内集. 我们有 $\mathbb{N}\subseteq[B_n]\subseteq {}^*\mathbb{N}$.

因为在 $\mathbb{N}$ 中每个有限集都有一最大元, $^*\mathbb{N}$ 中的超有限集也一定包含一最大元. 这可以从转换原理 (定理 3.31) 中看出. 换一个角度, 我们也可以说一个超有限集 A 是一个内集, 所以, $A=[A_n]$ 使得每个 A_n 都是 $\mathbb{N}$ 中的有限集, 因此, A_n 有最大元 a_n. 容易证明 $[a_n]$ 是 A 中的最大元.

注意, 内集的元素是内集但内集的子集有可能是外集.

对于任意关于超结构的公式 $\phi(x_1,x_2,\dots,x_s)$ 使得 $x_1,x_2,\dots,x_s$ 是 ϕ 中所有自由变量, 对于任意 $[a_n^{(1)}],[a_n^{(2)}],\dots,[a_n^{(s)}]\in {}^*V$, 含参数语句 $\phi([a_n^{(1)}],[a_n^{(2)}],\dots,[a_n^{(s)}])$ 在 $^*\mathcal{V}$ 中的真假值也可依照公式的长度来递归定义. 对于原子公式 ϕ, $[a_n]=[b_n]$ 在 $^*\mathcal{V}$ 中为真当且仅当 $\{n\in\mathbb{N}: a_n=b_n\}\in\mathcal{U}$, $[a_n]\in[b_n]$ 在 $^*\mathcal{V}$ 中为真当且仅当 $\{n\in\mathbb{N}: a_n\in b_n\}\in\mathcal{U}$. 如果 ϕ 是 $\neg\psi$, 则 ϕ 在 $^*\mathcal{V}$ 中为真当且仅当 $\neg\psi$ 在 $^*\mathcal{V}$ 中为假. 如果 ϕ 是 $\psi\to\eta$, 则 ϕ 在 $^*\mathcal{V}$ 中为真当且仅当 ψ 在 $^*\mathcal{V}$ 中为假或者 η 在 $^*\mathcal{V}$ 中为真. 如果 $\phi(x_1,x_2,\dots,x_s)$ 是 $\forall x\psi(x,x_1,x_2,\dots,x_s)$, 则 $\phi([a_n^{(1)}],[a_n^{(2)}],\dots,[a_n^{(s)}])$ 在 $^*\mathcal{V}$ 中为真当且仅当对任意 $[a]\in {}^*V$, $\psi([a],[a_n^{(1)}],[a_n^{(2)}],\dots,[a_n^{(s)}])$ 在 $^*\mathcal{V}$ 中为真. 注意, 因为超结构的语言中只有一个关系符号 $\in$, 并且我们还规定可以把 $^*\in$ 写成 $\in$, 所以对于关于 $\mathcal{V}$ 的公式 ϕ, 我们没有必要对它进行翻译, 写成 $^*\phi$.

命题 3.30 设 $\phi(x_1,x_2,\dots,x_s)$ 为一个关于超结构的公式, 使得 $x_1,x_2,\dots,x_s$ 是 ϕ 中所有自由变量, 设 $[a_n^{(1)}],[a_n^{(2)}],\dots,[a_n^{(s)}]\in {}^*V$, 则有 $\phi([a_n^{(1)}],[a_n^{(2)}],\dots,[a_n^{(s)}])$ 在 $^*\mathcal{V}$ 中为真当且仅当

$$\{n\in\mathbb{N}:\phi(a_n^{(1)},a_n^{(2)},\dots,a_n^{(s)})\text{ 在 }\mathcal{V}\text{ 中为真}\}\in\mathcal{U}.$$

以上命题也是 Łoś 定理的特例, 其证明与命题 3.8 的证明基本一样.

以下是超结构的非标准扩张转换原理.

定理 3.31 (转换原理) 设 $\phi(x_1,x_2,\dots,x_s)$ 为一个关于超结构的公式, 使得 $x_1,x_2,\dots,x_s$ 是 ϕ 中所有自由变量, $a^{(1)},a^{(2)},\dots,a^{(s)}\in V(\mathbb{R})$, 则 $\phi(a^{(1)},a^{(2)},\dots,a^{(s)})$ 在 $\mathcal{V}$ 中为真当且仅当 $\phi(^*a^{(1)},{}^*a^{(2)},\dots,{}^*a^{(s)})$ 在 $^*\mathcal{V}$ 中为真.

Robinson 最初的非标准扩张使用了模型论中的紧致性定理. 那是一种比较抽象的扩张方法. 用超幂方法构造的非标准扩张比较具体, 每一个 $^*V$ 中元素 a 都有一个具体的表示, 即 $a=[a_n]$. 但这可能也有负面作用, 因为非标准分析方法的优点之一是可以使有些数学问题变得比较直观. 如果每次考虑非标准元素时, 都把它想象成一个序列, 那么复杂度就会增加, 就比较不容易获得直观. 这与实数情形有些类似, 采用 Cauchy 方法, 每个标准实数可以看成是一个有理数 Cauchy 数列的一个等价类, 但是

在使用实数时, 我们并不把它们看成序列, 而是把它们看成数轴上的点. 想象一下, 如果每次碰到实数时, 都把它们想象成一个有理数的 Cauchy 数列, 学习微积分会有多烦. 所以, 读者在初学时可以经常把非标准元素或内集看成一个序列的等价类, 等到熟悉后, 就要努力训练自己把非标准元素想象成一些点, 把一个内集想象成一个集合论意义下通常的集合.

4 非标准分析中的常用性质

本节首先介绍非标准分析中常用的一些性质并讨论一些有关标准集、内集和外集之间的差别, 然后介绍怎样用这些性质构造 Loeb 测度空间.

4.1 不依赖非标准模型的性质

本小节中的命题或性质不依赖于非标准模型的构造. 如有时在证明中使用了超幂的结构, 也只是为了方便而已.

以下这个命题常称为溢出 - 溢入原理.

命题 4.1 设 $U \subseteq {}^*\mathbb{N}$ 为 ${}^*\mathbb{N}$ 中的非空前截 (initial segment), 即满足 $a, b \in {}^*\mathbb{N}$, $a < b$ 和 $b \in U$ 推出 $a \in U$, U 在 ${}^*\mathbb{N}$ 中有上界, U 没有最大元 (所以 U 是外集). 设 $A \subseteq {}^*\mathbb{N}$ 为一非空内集, 则

(1) 如果对于 U 中任意元素 u, 集合 A 中总有元素 $a > u$, 则 A 一定不是 U 的子集;

(2) 如果对于任意不在 U 中的元素 $v \in {}^*\mathbb{N}$, 集合 A 中总有元素 $a < v$, 则 $A \cap U \neq \emptyset$.

证明 令 $A = [A_n]$. 先证明 (1). 如果 A 在 ${}^*\mathbb{N}$ 中有上界, 那么 A 就有最大元 $[a_n]$. 注意, $[a_n]$ 一定不在 U 里面, 这是因为, 如果 $[a_n]$ 在 U 里面, A 就变成在 U 中有界了. (2) 的证明与 (1) 的证明基本相同, 只是把最大元换成最小元即可. 证毕. □

一个内集 $A \in {}^*V$ 显然可以被一个含参数的关于超结构的一阶公式在 ${}^*\mathcal{V}$ 中来定义的. 这是因为 A 本身就可以是一个参数, 即 $\phi(x, A) := \forall y\, (y \in x \leftrightarrow y \in A)$ 在 ${}^*\mathcal{V}$ 中定义了集合 A. 现在来证明其另一面, 即用内集来定义的集合是内集.

命题 4.2 设 $\phi(x, y_1, y_2, \ldots, y_s)$ 为一个关于超结构的一阶公式, $x, y_1, y_2, \ldots, y_s$ 是其所有自由变量. 设 $A^{(0)}, A^{(1)}, A^{(2)}, \ldots, A^{(s)}$ 为内集, 则集合

$$B = \{b \in A^{(0)} : \phi(b, A^{(1)}, A^{(2)}, \ldots, A^{(s)}) \text{ 在 } {}^*\mathcal{V} \text{ 中为真}\}$$

是一个内集. 如果 $A^{(0)}, A^{(1)}, A^{(2)}, \ldots, A^{(s)}$ 都是标准集, 则集合 B 也是标准集.

证明 设 $m \leqslant N$ 足够大使得 $A^{(0)} \in {}^*V_m$. 令 $A_n^{(0)} \in V_m$ 使得 $A^{(0)} = [A_n^{(0)}]$. 我们需要找到 $B_n \subseteq A_n^{(0)}$ 使得 $B = [B_n] \in {}^*V_m$. 对 $i = 1, 2, \ldots, s$, 令 $A^{(i)} = [A_n^{(i)}]$. 任给 $n \in \mathbb{N}$, 令

$$B_n = \{x \in A_n^{(0)} : \phi(x, A_n^{(1)}, A_n^{(2)}, \ldots, A_n^{(s)}) \text{ 在 } \mathcal{V} \text{ 中为真}\}.$$

我们需要证明 $B = [B_n]$. 给定一个 $[a_n] \in A^{(0)}$, 则 $[a_n] \in B$ 当且仅当

$$\phi([a_n], A^{(1)}, A^{(2)}, \ldots, A^{(s)}) \text{ 在 } {}^*\mathcal{V} \text{ 中为真},$$

当且仅当 (由命题 3.30)

$$\{n \in \mathbb{N} : \phi(a_n, A_n^{(1)}, A_n^{(2)}, \ldots, A_n^{(s)}) \text{ 在 } \mathcal{V} \text{ 中为真}\} \in \mathcal{U}$$

当且仅当 $\{n\in\mathbb{N}: a_n\in B_n\}\in\mathcal{U}$, 当且仅当 $[a_n]\in[B_n]$. 所以 B 是内集.

如果集合 $A^{(0)},A^{(1)},A^{(2)},\ldots,A^{(s)}$ 都是标准集, 则所定义的所有集合 B_n 都是相同的 $\mathcal{V}$ 中的集合. 所以 B 是标准集. 证毕. □

命题 4.3 设 $A\subseteq{}^*\mathbb{N}$ 是一内集. 如果 $0\in A$ 并且 $n\in A$ 推出 $n+1\in A$, 则 $A={}^*\mathbb{N}$.

证明 假设 $A\neq{}^*\mathbb{N}$, 则 $A^C={}^*\mathbb{N}\setminus A$ 是一个非空内集. 在 $\mathcal{V}$ 中, $\mathbb{N}$ 的每一个非空子集都有最小元. 由转换原理, 在 ${}^*\mathbb{N}$ 中, 每个非空内集有最小元. 所以 A^C 有最小元 a. 因为 $0\in A$, 所以 $a>0$ 并且 $a-1\in A$. 这与 $a-1\in A$ 推出 $a\in A$ 矛盾. 证毕. □

在第 2 节的公式 (2.2) 中, 我们介绍了一个关于实数完备性的二阶语句. 在那个语句中, 有一量词 $\forall X\subseteq[0,1]$. 变量 X 的值是集合而不是元素, 而这正是语句 (2.2) 不能在非标准扩张中成立的根本原因. 在超结构中, 集合可以是元素. 所以, $\forall X\subseteq[0,1]$ 可写成 $\forall x\in\mathcal{P}([0,1])$, 这里 $\mathcal{P}([0,1])$ 是 $[0,1]$ 的幂集. 设

$$\phi:=\forall x\in\mathcal{P}([0,1])\,(\exists y\in x\rightarrow\exists z\in[0,1]\,(\forall u\in x\,(u\leqslant z)\wedge\forall y\,(y<z\rightarrow\exists u\in x\,(u>y)))).$$

语句 ϕ 是一阶的, 叙述了实数的完备性且在 $\mathcal{V}$ 中为真. 这是不是与非标准扩张 ${}^*\mathcal{V}$ 和 $\mathcal{V}$ 满足相同的一阶语句且 ${}^*\mathcal{V}$ 中的非标准实数域不完备相矛盾? 其实是不矛盾的, 这里的要点是集合 $\mathcal{P}({}^*[0,1])$ 和 ${}^*\mathcal{P}([0,1])$ 的不同. 因为 $x=\mathcal{P}([0,1])$ 当且仅当

$$\phi(x,[0,1]):=\forall y\,(y\in x\leftrightarrow\forall z\,(z\in y\rightarrow z\in[0,1]))$$

在 $\mathcal{V}$ 中为真, 所以, 公式

$$\phi(x,{}^*[0,1]):=\forall y\,(y\in x\leftrightarrow\forall z\,(z\in y\rightarrow z\in^*[0,1]))$$

在 ${}^*\mathcal{V}$ 中定义了一个集合, 记为 ${}^*\mathcal{P}([0,1])$. 由命题 4.2 可知, ${}^*\mathcal{P}([0,1])$ 是一个内集. 再从 ${}^*\mathcal{P}([0,1])$ 的定义和 "$\forall y$" 在 ${}^*\mathcal{V}$ 中的解释来看, ${}^*\mathcal{P}([0,1])$ 中的元素 y 必须取值于 ${}^*V$ 中的元, 所以 ${}^*\mathcal{P}([0,1])$ 是 ${}^*[0,1]$ 中所有内集的集合族, 而 $\mathcal{P}({}^*[0,1])$ 则是 ${}^*[0,1]$ 中所有集合 (内集或外集) 的集合族. 例如, 令 I 为 $[0,1]$ 中所有无穷小数的集合, 则 $I\in\mathcal{P}({}^*[0,1])$ 但 I 没有上确界, 所以 $I\notin^*\mathcal{P}([0,1])$. 显然, I 是一个外集. 可以证明所有在 ${}^*\mathbb{R}$ 中有上界但无上确界的子集一定是个外集.

4.2 可数饱和性

可数饱和性 (countable saturation) 依赖于非标准模型的构造. 本文所用的超幂构造, 满足可数饱和性. 如果采用模型论中的紧致性定理, 则可以构造出满足或不满足可数饱和性的非标准模型. 一般来讲, 对于任意基数 κ, 我们都可以用超幂方法或紧致性定理来构造满足 κ 饱和性的非标准模型. 但为了可读性, 本文只讨论可数饱和性. 可数饱和性在本文涉及的应用中也足够了.

定义 4.4 设 $\mathcal{F}$ 为一集合族. 如果对于 $\mathcal{F}$ 中的任意有限子族 $\mathcal{F}_0$, 都有 $\bigcap_{F\in\mathcal{F}_0}F\neq\emptyset$, 则称集合族 $\mathcal{F}$ 满足有限交性质 (finite intersection property).

命题 4.5 (可数饱和性) 如果 $\mathcal{F}$ 是一个可数集合族使得 $\mathcal{F}$ 中元素都是 ${}^*V$ 中内集, 且 $\mathcal{F}$ 满足有限交性质, 则有

$$\bigcap_{F\in\mathcal{F}}F\neq\emptyset.$$

证明 设 $\mathcal{F}=\{[F_n^{(m)}]:m\in\mathbb{N}\}$. 我们需要找到一个 $[a_n]\in{}^*V$ 使得对于任何一个 $m\in\mathbb{N}$, 都有 $[a_n]\in[F_n^{(m)}]$. 不失一般性, 可假设 $[F_n^{(0)}]\supseteq[F_n^{(1)}]\supseteq[F_n^{(2)}]\supseteq\cdots$. 任取一元 $[a_n^{(m)}]\in[F_n^{(m)}]$. 对于任一自然数 m, 令

$$U_m:=\{n\in\mathbb{N}:n>m,a_n^{(m)}\in F_n^{(m)},\text{ 并且 }F_n^{(m)}\subseteq F_n^{(m-1)}\subseteq\cdots\subseteq F_n^{(0)}\}.$$

由命题 3.30, $U_m\in\mathcal{U}$. 显然, $U_0\supseteq U_1\supseteq\cdots$ 并且 $\bigcap_{m\in\mathbb{N}}U_m=\emptyset$. 现在定义所要的序列 $\{a_n\}$. 给定 $n\in\mathbb{N}$, 令 $m_n:=\max\{m:n\in U_m\}$. 注意, 因为所有 U_m 的交是空集, 所以 m_n 一定存在. 显然, $n\in U_{m_n}$. 令

$$a_n:=a_n^{(m_n)}.$$

现在只需证明对任意的 $m\in\mathbb{N}$, 有 $[a_n]\in[F_n^{(m)}]$. 由命题 3.8, 这等价于 $U=\{n\in\mathbb{N}:a_n\in F_n^{(m)}\}\in\mathcal{U}$. 当 $n\in U_m$ 时, 有 $m\leqslant m_n$, 这是因为 m_n 是这些 m 中的最大值. 因为 $n\in U_{m_n}$, 所以 $a_n^{(m_n)}\in F_n^{(m_n)}$ 并且 $F_n^{(m)}\supseteq F_n^{(m_n)}$, 这推出 $a_n=a_n^{(m_n)}\in F_n^{(m)}$. 因此 $n\in U$. 这证明了 $U_m\subseteq U$. 由滤子的定义, 有 $U\in\mathcal{U}$. 证毕. □

推论 4.6 设 A 是一内集, $s=\{a_n:n\in\mathbb{N}\}$ 是 A 中一个序列 (s 本身可能是外集), 则存在一个超整数 H 和一个内集 $S=\{b_n\}_{n=0}^H\subseteq A$ 使得对每个 $n\in\mathbb{N}$, 有 $b_n=a_n$, 即 S 是 s 在 A 中的延伸.

证明 对每个 $m\in\mathbb{N}$, 定义一个集合 $F^{(m)}$ 使得

$$F^{(m)}:=\{\{b_n\}_{n=0}^k:k\in{}^*\mathbb{N},k>m,\forall n\leqslant k(b_n\in A),b_m=a_m\}.$$

不难看出 $F^{(m)}$ 的定义可以用一阶公式来表示. 命题 4.2 推出 $F^{(m)}$ 是一个内集. 对于任意有限 $m\in\mathbb{N}$, 有限序列 $\{a_n\}_{n=0}^m$ 是内集并且属于 $\bigcap_{i=0}^m F^{(i)}$, 所以 $\{F^{(m)}:m\in\mathbb{N}\}$ 满足有限交性质, 所以由命题 4.5, 存在内集 $S\in\bigcap_{m\in\mathbb{N}}F^{(m)}$. 显然, $S=\{b_n\}_{n=0}^H$ 就是命题中所需要的. 证毕. □

4.3 Loeb 测度空间

Loeb 测度空间是非标准分析在标准数学中应用的最重要工具之一. Loeb 测度空间是用非标准分析的方法在一内集上构造一个标准测度空间, 使得非标准方法和标准方法两边的优势可以联系起来并应用. Loeb 测度空间可以从最广泛的定义开始, 但为了可读性, 我们只介绍超有限集上的由均权生成的 Loeb 概率测度空间, 而这对本文涉及的应用足够了.

设 A_n 为有限集, 并且 $A=[A_n]$ 为一个超有限集, 我们可定义 A 的内基数 $|A|:=[|A_n|]$. 在本小节中, 设 Ω 为一个超有限集并且 $|\Omega|=H$, 即 H 是一个超整数并且 $\Omega=\{\omega_1,\omega_2,\dots,\omega_H\}$. 注意, 基数函数 $|\cdot|$ 只对内集有意义,

命题 4.7 对任一元素 $\omega\in\Omega$, 令 $w(\omega)=\frac{1}{H}$, 即 Ω 中每一个元素的权都为无穷小数 $1/H$, 令 $\Sigma_0:={}^*\mathcal{P}(\Omega)$, 即 Σ_0 是所有 Ω 的内子集的集合族. 对于任意一个 $A\in\Sigma_0$, 定义

$$\mu_0(A)=\sum_{a\in A}w(a)=\frac{1}{H}|A|,$$

则 $(\Omega;\Sigma_0,\mu_0)$ 是一个在非标准意义下的 $*$ 有限可加概率测度空间.

证明 显然, Σ_0 是一个 Boole 代数, 即对交 $\cap$、并 $\cup$、和差 $\setminus$ 是闭合的. 又显然 $\mu_0(\Omega)=1$. 给定 $n\in\mathbb{N}$, 令 Ω_n 为一个 n 元有限集, 令 $\Sigma_{0,n}=\mathcal{P}(\Omega_n)$. 对于任意 $A\in\Omega_n$, 定义 $\mu_{0,n}(A)=|A|/n$, 则 $(\Omega_n;\Sigma_{0,n},\mu_{0,n})$ 是一个有限可加概率测度空间. 所以由转换原理, 命题得证. □

命题 4.8 令 $(\Omega;\Sigma_0,\mu_0)$ 为命题 4.7 中所定义. 对于任意 $A\in\Sigma_0$, 定义

$$\mu(A)=\mathrm{st}(\mu_0(A)),$$

则 $(\Omega;\Sigma_0,\mu)$ 是一个在标准意义下的有限可加无原子 (atomless) 概率测度空间.

证明 注意, 如果 $A_1,A_2\in\Sigma_0$ 并且 $A_1\cap A_2=\emptyset$, 则有 $\mu(A_1\cup A_2)=\mathrm{st}(\mu_0(A_1\cup A_2))=\mathrm{st}(\mu_0(A_1)+\mu_0(A_2))=\mathrm{st}(\mu_0(A_1))+\mathrm{st}(\mu_0(A_2))=\mu(A_1)+\mu(A_2)$. 如果 $A\in\Sigma_0$, $\mu(A)>0$, 则 $|A|$ 一定是一超整数. 令 $A=[A_n]$, 并选取 $B_n\subseteq A_n$ 使得 $|B_n|=\lceil\frac{1}{2}|A_n|\rceil$, 则 $B=[B_n]\subseteq A$ 并且 $\mu(B)=\frac{1}{2}\mu(A)$. 证毕. □

定义 4.9 设 $(\Omega;\Sigma_0,\mu)$ 为命题 4.8 中所定义. 对于 Ω 中任意子集 (内集或外集) X, 定义 X 的内测度 $\underline{\mu}(X)$ 和外测度 $\overline{\mu}(X)$ 分别为

$$\underline{\mu}(X):=\sup\{\mu(A):A\in\Sigma_0,A\subseteq X\},\quad \overline{\mu}(X):=\inf\{\mu(A):A\in\Sigma_0,A\supseteq X\},$$

这里 sup 和 inf 都是标准意义下的算子. 令

$$\Sigma:=\{X\subseteq\Omega:\underline{\mu}(X)=\overline{\mu}(X)\}.$$

对于每个 $X\in\Sigma$, 定义 $\mu_L(X)=\overline{\mu}(X)$.

根据以上定义, 如果 $A\in\Sigma_0$, 则 $\underline{\mu}(A)=\overline{\mu}(A)=\mu(A)$. 所以 $\Sigma_0\subseteq\Sigma$.

命题 4.10 以上定义的空间 $(\Omega;\Sigma,\mu_L)$ 满足下列性质: 设 $X\in\Sigma$,

(1) 对任意标准实数 $\epsilon>0$, 存在内集 $A,B\subseteq\Omega$ 使得 $A\subseteq X\subseteq B$ 并且 $\mu_L(B\setminus A)<\epsilon$.

(2) 如果 $Y\in\Sigma$, 那么 $X\setminus Y\in\Sigma$.

(3) 如果 $X_n\in\Sigma$, 那么 $\bigcup_{n\in\mathbb{N}}X_n\in\Sigma$; 如果 $\{X_n:n\in\mathbb{N}\}\subseteq\Sigma$ 是两两互不相交的, 那么 $\mu_L(\bigcup_{n\in\mathbb{N}}X_n)=\sum_{n=0}^{\infty}\mu_L(X_n)$.

(4) 存在内集 A_n 和 B_n 使得对每一个 $n\in\mathbb{N}$, 都有 $A_n\subseteq A_{n+1}\subseteq X\subseteq B_{n+1}\subseteq B_n$, $\mu_L(\bigcup_{n\in\mathbb{N}}A_n)=\mu_L(X)=\mu_L(\bigcap_{n\in\mathbb{N}}B_n)$.

(5) 存在内集 $A\subseteq\Omega$ 使得 $\mu_L(\Delta(A,X))=0$, 这里 $\Delta(A,X)$ 是 A 和 X 的对称差, 即 $\Delta(A,X)=(A\setminus X)\cup(X\setminus A)$.

(6) $\mu_L(X)=0$ 和 $Y\subseteq X$ 推出 $Y\in\Sigma$ 和 $\mu_L(Y)=0$.

所以, $(\Omega;\Sigma,\mu_L)$ 是一个无原子的、可数可加的、完备的概率测度空间.

证明 以下关于概率空间的基本性质是显然的, 即 (i) $\mu_L(\Omega)=1$; (ii) 对任意 $X\in\Sigma$, 有 $0\leqslant\mu_L(X)\leqslant 1$; (iii) 如果 $X,Y\subseteq\Omega$ 和 $X\subseteq Y$, 则可推出 $\mu_L(X)\leqslant\mu_L(Y)$. 接下来证明上述各性质:

(1) 由 sup 和 inf 的定义, 存在内集 A 和 B 使得 $A\subseteq X\subseteq B$, $\mu(A)>\underline{\mu}(X)-\frac{1}{2}\epsilon=\mu_L(X)-\frac{1}{2}\epsilon$ 和 $\mu(B)<\overline{\mu}(X)+\frac{1}{2}\epsilon=\mu_L(X)+\frac{1}{2}\epsilon$. 由 μ 的有限可加性, 有

$$\mu_L(B\setminus A)=\mu(B\setminus A)=\mu(B)-\mu(A)<\epsilon.$$

(2) 给定任意标准正实数 ϵ 和 $i=1,2$, 令 $A_i,B_i\in\Sigma$ 使得 $A_1\subseteq X\subseteq B_1$, $A_2\subseteq Y\subseteq B_2$, $\mu_L(B_1\setminus A_1)<\epsilon/2$ 和 $\mu_L((B_1\setminus A_1)<\epsilon/2$, 则有 $A_1\setminus B_2\subseteq X\setminus Y\subseteq B_1\setminus A_2$. 所以,

$$\begin{aligned}\overline{\mu}(X\setminus Y)-\underline{\mu}(X\setminus Y)&\leqslant\mu(B_1\setminus A_2)-\mu(A_1\setminus B_2)=\mu((B_1\setminus A_2)\setminus(A_1\setminus B_2))\\&\leqslant\mu((B_1\setminus A_1)\cup(B_2\setminus A_2))<\epsilon.\end{aligned}$$

因为 ϵ 可任意小, 所以 $\underline{\mu}(X\setminus Y)=\overline{\mu}(X\setminus Y)$.

(3) 由 (2) 可假设 X_n 两两不相交. 任给标准正实数 ϵ, 令 $A_n, B_n \in \Sigma$ 使得 $A_n \subseteq X_n \subseteq B_n$ 和 $\mu(B_n \setminus A_n) \leqslant \epsilon/2^{n+2}$. 因为

$$\lim_{n\to\infty} \sum_{i=0}^{n} \mu(A_n) = \lim_{n\to\infty} \mu\left(\bigcup_{i=0}^{n} A_i\right) \leqslant \underline{\mu}\left(\bigcup_{n=1}^{\infty} X_n\right) \leqslant 1,$$

所以存在 n_0 使得 $\sum_{i=n_0+1}^{\infty} \mu(A_n) < \epsilon/2$. 又由命题 4.6 可知, 存在超整数 K 使得 $\{(A_n, B_n) : n = 0, 1, \ldots, K\}$ 是内集且是 $\{(A_n, B_n) : n \in \mathbb{N}\}$ 的延伸, 并满足 $\sum_{i=n_0+1}^{K} \mu(A_n) < \epsilon/2$. 现在有

$$\overline{\mu}\left(\bigcup_{n\in\mathbb{N}} X_n\right) - \underline{\mu}\left(\bigcup_{n\in\mathbb{N}} X_n\right) \leqslant \mu\left(\bigcup_{i=0}^{K} B_i\right) - \mu\left(\bigcup_{i=0}^{n_0} A_i\right) \leqslant \sum_{i=0}^{K} \mu(B_i \setminus A_i) + \sum_{i=n_0+1}^{K} \mu(A_i) < \epsilon.$$

因为 ϵ 可以任意小, 所以, $\overline{\mu}(\bigcup_{n\in\mathbb{N}} X_n) = \underline{\mu}(\bigcup_{n\in\mathbb{N}} X_n)$. 从以上不等式还可看出,

$$\sum_{i=0}^{n_0} \mu(A_i) \leqslant \mu_L\left(\bigcup_{n\in\mathbb{N}} X_n\right) \leqslant \sum_{n=0}^{\infty} \mu_L(X_n) \leqslant \sum_{i=0}^{K} \mu(B_i) \leqslant \sum_{i=0}^{n_0} \mu(A_i) + \epsilon.$$

所以, $\mu_L(\bigcup_{n\in\mathbb{N}} X_n) = \sum_{n=0}^{\infty} \mu_L(X_n)$.

由 (2) 和 (3) 可知, Σ 是一个 σ- 代数, 并且测度 μ_L 是可数可加的.

(4) 对每个 $n \in \mathbb{N}$, 选取 A'_n 和 B'_n 使得 $A'_n \subseteq X \subseteq B'_n$ 并且 $\mu_L(B'_n \setminus A'_n) < \frac{1}{n}$. 令 $A_n = \bigcup_{i=0}^{n} A'_n$ 和 $B_n = \bigcap_{i=0}^{n} B'_n$, 则 $A_n \subseteq A_{n+1} \subseteq X \subseteq B_{n+1} \subseteq B_n$ 并且

$$\mu_L\left(\bigcap_{n\in\mathbb{N}} B_n\right) - \mu_L\left(\bigcup_{n\in\mathbb{N}} A_n\right) \leqslant \mu_L(B_n) - \mu_L(A_n) < \frac{1}{n}.$$

令 $n \to \infty$, 所要结论成立.

(5) 设 $\mathcal{S}_n = \{C \in \Sigma_0 : A_n \subseteq C \subseteq B_n\}$. 由命题 4.2 知, 集合 $\mathcal{S}_n$ 是一个内集. 因为 $m > n$ 推出 $A_m \in \mathcal{S}_n$, 所以, $\{\mathcal{S}_n : n \in \mathbb{N}\}$ 满足有限交性质. 由命题 4.5 可知, 存在 $A \in \bigcap_{n\in\mathbb{N}} \mathcal{S}_n$. 显然, A 是内集, $\bigcup_{n\in\mathbb{N}} A_n \subseteq A \subseteq \bigcap_{n\in\mathbb{N}} B_n$, $\Delta(A, X) \subseteq (\bigcap_{n\in\mathbb{N}} B_n \setminus \bigcup_{n\in\mathbb{N}} A_n)$, 而这推出 $\mu_L(\Delta(A, X)) \leqslant \frac{1}{n} \to 0$.

(6) 因为 $0 \leqslant \underline{\mu}(Y) \leqslant \overline{\mu}(Y) \leqslant \overline{\mu}(X) = 0$, 所以 $\underline{\mu}(Y) = \overline{\mu}(Y) = 0$.

因为 (6), 空间 $(\Omega; \Sigma, \mu_L)$ 称为完备的. 如果 $X \in \Sigma$ 使得 $\mu_L(X) > 0$, 则一定有一个内集 $A \subseteq X$, 使得 $\mu(A) > 0$. 由命题 4.8 的证明, 存在内集 $B \subseteq A$ 使得 $\mu(B) = \frac{1}{2}\mu_L(A)$. 所以空间 $(\Omega; \Sigma, \mu_L)$ 没有原子. 命题证毕. □

注 4.11 $(\Omega; \Sigma, \mu_L)$ 称为 Loeb 概率测度空间或简称 Loeb 空间, 而 μ_L 则称为 Ω 上的 Loeb 测度, Σ 中元素称为 Loeb 可测集.

5 非标准分析在一些数学分支的应用

本节介绍三个非标准分析在其他数学分支中应用的结果. 其实还有很多值得介绍的结果, 之所以选了以下三个, 只是因为篇幅所限. 当然这些结果的证明还需读者自己去阅读原始文章或其他相关书籍.

5.1 随机微分方程强解的存在性定理

首先介绍 Anderson (一维) 随机游动 (random walk). Anderson 的随机游动的标准化即是 Wiener 过程 (Wiener process) 或称 Brown 运动 (Brownian motion). 所以, Anderson 的随机游动给出 Brown 运动的一个具体表示方式.

定义 5.1 设 $(\Omega;\lambda)$ 为一概率空间, $T=[0,\infty)$ 为时间轴. 如果一个随机过程 $b:\Omega\times T\to\mathbb{R}$ 满足以下条件:

(1) 对任意 $\omega\in\Omega$, $b(\omega,0)=0$;

(2) 对 λ- 几乎所有 $\omega\in\Omega$, 函数 $b(\omega,\cdot):T\to\mathbb{R}$ 是一连续函数;

(3) 对任意 $0=t_0<t_1<\cdots<t_m<\infty$, 随机变量集合 $\{b(\cdot,t_{i+1})-b(\cdot,t_i):i=0,1,\dots,m-1\}$ 中的随机变量是相互独立的;

(4) 对任意 $0\leqslant s<t<\infty$, 随机变量 $b(\cdot,t)-b(\cdot,s)$ 有正态分布, 且均值 (mean) 为 0, 标准差 (standard deviation) 为 $\sqrt{t-s}$,

则该随机过程称为 Brown 运动.

定义 5.2 设 H 为一个超整数, $T:=\{0,\frac{1}{H},\frac{2}{H},\dots,\frac{H^2-1}{H},H\}$, $t_i:=\frac{i}{H}$, $\Delta t:=\frac{1}{H}$. 我们可以把 T 看成离散时间轴, 使得最小时间区间的长度为无穷小 Δt, 并且总长度在标准意义下是无限的.

设 Ω 是所有从 T 到 $\{-1,1\}$ 的内函数的集合. 注意, Ω 是内集且 $|\Omega|=2^{H^2+1}$. 在 Ω 上, 我们可以定义 Loeb 测度 μ_L. 注意, Ω 中每一个点 ω 是一个从超有限集 T 到 $\{-1,1\}$ 的内函数.

定义 $B:\Omega\times T\to{}^*\mathbb{R}$ 使得

$$B(\omega,t):=\sum_{s<t}\omega(s)\sqrt{\Delta t},$$

则 $B(\omega,t)$ 称为 Anderson 随机游动.

以下定理由 Anderson [12] 第一次提出并证明, 读者也可在其他非标准分析教课书或参考书中找到, 如文献 [4,5,13].

定理 5.3 令 Ω,μ_L,T 和 $B(\omega,t)$ 为定义 5.2 中所定义. 对任意标准实数 $t\in(0,\infty)$, 令 $t^-=\max\{s\in T:s<t\}$, 设 $b(\omega,0)=0$, 并且对 $t\in(0,\infty)$,

$$b(\omega,t):=\mathrm{st}(B(\omega,t^-)),$$

则存在 μ_L- 测度为 1 的集合 $X\subseteq\Omega$ 使得 $b(\omega,t):X\times[0,\infty)\to\mathbb{R}$ 为 Brown 运动.

因为在非标准结构里, Anderson 随机游动就可看作是离散化的 Brown 运动, 有时在解涉及 Brown 运动的随机微分方程时, 就可以利用其离散化而得到一些运算上的便利. 利用可数饱和性还可证明, Ω 上的任意两个 Brown 运动都是同构的, 所以, 每个 Brown 运动都可以用 Anderson 随机游动生成的 Brown 运动来代替. Anderson 随机游动已成为非标准分析中解随机微分方程的一个常用工具.

以下是 Keisler 和其他一些非标准分析学者证明的一系列随机微分方程强解存在性定理中的一个. 定理的证明可在文献 [14] 中找到, 有兴趣者还可参见文献 [15–17] 等.

定理 5.4 设 $b:\Omega\times[0,\infty)\to\mathbb{R}^d$ 为 d- 维 Brown 运动, $g:[0,1]\times\mathbb{R}^d\to\mathcal{M}_d^k$ 为平方可积函数, 这里 $\mathbb{R}^d$ 上的测度为正规测度, $\mathcal{M}_d^k$ 是所有 $d\times d$ 矩阵使得矩阵中的数取自 $[-k,k]$, 并且 $\det(gg^{\mathrm{T}})\geqslant 1/k$, 则方程

$$x(\omega,t)=\int_0^t g(s,x(\omega,s))db(\omega,s)$$

的解 x 存在并且 x 是 $\Omega\times[0,\infty)$ 上的 d 维连续鞅 (continuous martingale).

以上随机微分方程的解称为强解, 指的是 Brown 运动所在空间和解 x 所在空间相同, 都是 Loeb 概率空间 Ω. 这在标准数学中一般无法做到, 在标准数学中所得到的解为弱解, 即 Brown 运动所在的空间和解所在的空间不是同一个空间, 读者可参见文献 [18].

5.2 关于局部群的 Hilbert 第五问题

如果 $(G;\cdot,1_G)$ 是一个群, $(G;\mathcal{T})$ 是一个 Hausdorff 拓扑空间, 并且群的逆运算 $g\mapsto g^{-1}$ 是 $(G;\mathcal{T})$ 上的连续函数, 乘法运算 $(g,h)\mapsto gh$ 是 $(G\times G;\mathcal{T}\times\mathcal{T})$ 上的连续函数, 则结构 $(G;\cdot,1_G,\mathcal{T})$ 称为拓扑群.

在一个拓扑群 $(G;\cdot,1_G,\mathcal{T})$ 中, 如果存在一个 1_G 的开邻域 $U\subseteq G$ 使得 U 和有限维 Euclid 空间 $\mathbb{R}^n$ 同胚, 则该拓扑群称为局部 Euclid 群. 如果在一个局部 Euclid 群中, 逆运算和乘法运算都是无穷次可微的, 则该局部 Euclid 群称为 Lie 群.

Hilbert 第五问题 是否每个局部 Euclid 群一定是 Lie 群.

Gleason [19], Montgomery 和 Zippin [20] 在 1952 年解决了 Hilbert 第五问题, 给出了肯定的回答.

1957 年, Jacoby [21] 给出了一个 Hilbert 第五问题关于局部 Euclid 局部群的推广. 注意, 一个局部群可能不是群.

定义 5.5 在一个结构 $(G;\iota,p,1_G,\mathcal{T})$ 中, 如果 $(G;\mathcal{T})$ 是 Hausdorff 拓扑空间, $1_G\in G$, 存在 1_G 的开邻域 $U\subseteq G$ 和 $(\{1_G\}\times G)\cup(G\times\{1_G\})$ 的开邻域 $V\subseteq G\times G$ 使得 $\iota:U\to G$ 和 $p:V\to G$ 是连续函数, 并且对所有 $x,y,z\in G$,

(1) $p(1_G,x)=p(x,1_G)=x$;

(2) 如果 $x\in U$, 那么 $(x,\iota(x)),(\iota(x),x)\in V$ 并且 $p(x,\iota(x))=1$, $p(\iota(x),x)=1$;

(3) 如果 $(x,y),(y,z),(p(x,y),z),(x,p(y,z))\in V$, 那么 $p(p(x,y),z)=p(x,p(y,z))$,

则该结构称为局部群.

在以上定义中, ι 是只在 U 中有定义的逆运算, p 是只在 V 中有定义的乘法. 以上条件 (1) 表示 1_G 是单位元, 条件 (2) 表示 ι 是逆运算, 条件 (3) 表示乘法 p 满足群的结合律.

定义 5.6 设 $(G;\iota,p,1_G,\mathcal{T})$ 为一局部群, U 和 V 为以上所定义, 则

(1) 对于包含 1_G 的开集 W, G 在 W 上的限制 $G\mid W$ 是一个局部群 $(W;\iota_W,p_W,1_G,\mathcal{T}\mid W)$, 这里 $U_W:=W\cap U\cap\iota^{-1}(U)$, $V_W:=V\cap W\times W\cap p^{-1}(W)$, ι_W 是 ι 在 U_W 上的限制, p_W 是 p 在 V_W 上的限制, $\mathcal{T}\mid W$ 是 $\mathcal{T}$ 在 W 上的限制;

(2) 如果存在一个 1_G 的开邻域 W 使得 W 和 $\mathbb{R}^n$ 同胚, 则称 G 为局部 Euclid 的局部群;

(3) 如果 ι 和 p 是无穷次可微的, 则称 G 为局部 Lie 群.

Jacoby 的 Hilbert 第五问题关于局部 Euclid 局部群的推广指的是, 证明了对每一个局部 Euclid 的局部群, 存在一单位元的开邻域 W 使得该局部群在 W 上的限制是一个局部 Lie 群. 可惜文献 [21] 中的证明是有瑕疵的, 所以不能被大家所承认, 读者可参见文献 [22].

1990 年, Hirschfeld [23] 使用非标准分析方法给出了 Hilbert 第五问题解的证明. 与最初的证明相比, 简明了许多. 2010 年, Goldbring [24] 沿着 Hirschfeld 的思路, 采用非标准分析方法, 最先解决了关于局部群的 Hilbert 第五问题. 以下是所证明的定理.

定理 5.7 如果 G 是一个局部 Euclid 局部群, 那么就存在 1_G 的一个开邻域 W, 使得 G 在 W 上的限制是一个局部 Lie 群.

5.3 精确大数定律及在经济学中的应用

保险公司售出的每一份保单都包含了金融风险, 但如果该公司售出了同一险种的大量保单, 并且保单和保单之间的金融风险是相互独立的, 这些风险就会相互抵消使得总体的风险大大降低. 怎样来建立针对这个经济现象的合理的数学模型?

一份保单 i 的风险可看作是一概率空间 $(\Omega;\mathcal{F},P)$ 上的随机变量 f_i. 设 I 为所有保单的集合, 则某情形 $\omega\in\Omega$ 下的某一份保单 i 的风险就是一个在 $I\times\Omega$ 上的二元函数 $f(i,\omega)$, 这里 $f_i(\omega)=f(i,\omega)$. 在现实生活中, 所有保单的集合 I, 无论多大, 一定是有限集, 但是为了能更有效地分析总体风险, 经济学家往往假设一种理想情形, 即设 I 是无限的. 集合 I 可以被假设为自然数集, 但更多时候 I 被设为概率空间 $(I;\mathcal{I},\lambda)$. 很自然地, 风险函数 $f(i,\omega)$ 应设为乘积概率空间 $(I\times\Omega;\sigma(\mathcal{I}\times\mathcal{F}),\lambda\times P)$ 上的可测函数, 或称为联合可测函数 (jointly measurable function), 这里 $\sigma(\mathcal{I}\times\mathcal{F})$ 表示由 $\mathcal{I}\times\mathcal{F}$ 生成的 σ- 代数. 同时还应假设风险随机变量集 $\{f_i:i\in I\}$ 中的随机变量是几乎两两互相独立的. 可惜这两个假设对绝大多数有意义的函数 $f(i,\omega)$ 是有矛盾的. 以下命题的证明可在文献 [5, 命题 8.3.3] 中找到.

命题 5.8 [25] 假设 $f:I\times\Omega\to\mathbb{R}$ 是乘积空间 $(I\times\Omega;\sigma(\mathcal{I}\times\mathcal{F}),\lambda\times P)$ 上的联合可测函数, 使得对 λ- 几乎所有的 $i\in I$ 和 λ- 几乎所有的 $s\in I$, $f_i(\omega):=f(i,\omega)$ 和 $f_s(\omega):=f(s,\omega)$ 作为 Ω 上的随机变量是互相独立的, 则对 λ- 几乎所有的 $i\in I$, f_i 是 Ω 上的一个常数随机变量.

注意, f_i 是常数随机变量指的是 f_i P- 几乎处处等于一个常数, 而几乎处处等于常数的随机变量根本不值得研究. 在以上命题中, 函数 $f(i,\omega)$ 的值域可以设为可分的完备度量空间.

命题 5.8 说明, 两个看起来很自然的性质, 即 $f(i,\omega)$ 的联合可测性和对 λ- 几乎所有的随机变量 $f(i,\cdot)$ 之间的相互独立性, 除了平凡情形, 是互不相容的. 这个问题给经济学中很多模型的建立造成了困难. 究其原因, 以上不相容性的产生是因为乘积可测集代数 $\sigma(\mathcal{I}\times\mathcal{F})$ 不是一个 $\mathcal{I}\times\mathcal{F}$ 的合适代数扩张. 但是什么才是合适的代数扩张呢?

设 I 和 Ω 都是超有限集, 并且 $(I;\mathcal{I},\lambda)$ 和 $(\Omega;\mathcal{F},P)$ 都是 Loeb 概率空间. 在 $I\times\Omega$ 上有两个自然的乘积代数的扩张. 一个是在通常标准数学中的扩张 $\sigma(\mathcal{I}\times\mathcal{F})$, 还有一个称为 Fubini 扩张. 因为 $I\times\Omega$ 是一个超有限集, 所以可以按照第 4.3 小节中的方法构造一个 Loeb 概率测度 (记为 $\lambda\boxtimes P$) 和一个 Loeb 可测集代数 (记为 $\mathcal{I}\boxtimes\mathcal{F}$). 可以证明 $\sigma(\mathcal{I}\times\mathcal{F})\subseteq\mathcal{I}\boxtimes\mathcal{F}$, 并且如果 $X\in\sigma(\mathcal{I}\times\mathcal{F})$, 那么 $\lambda\boxtimes P(X)=\lambda\times P(X)$. Loeb 空间 $(I\times\Omega;\mathcal{I}\boxtimes\mathcal{F},\lambda\boxtimes P)$ 称为 $(I\times\Omega;\mathcal{I}\times\mathcal{F},\lambda\times P)$ 的 Fubini 扩张. 可以证明 Fubini 扩张满足以下 Fubini 性质: 任给 $I\times\Omega$ 上的 $\lambda\boxtimes P$- 可积函数 f,

(1) 对 λ- 几乎所有的 $i\in I$, $f(i,\cdot)$ 是 Ω 上的 P- 可积函数, 对 P- 几乎所有的 $\omega\in\Omega$, $f(\cdot,\omega)$ 是 I 上的 λ- 可积函数;

(2) $\int_I f(i,\cdot)d\lambda$ 是 Ω 上的可积函数, $\int_\Omega f(\cdot,\omega)dP$ 是 I 上的可积函数, 并且

$$\int_{I\times\Omega} fd\lambda\boxtimes P=\int_I\int_\Omega f(\cdot,\omega)dPd\lambda=\int_\Omega\int_I f(i,\cdot)d\lambda dP.$$

以下定理 5.10 是称为 Fubini 扩张的精确大数定律中的几个形式之一, 参见文献 [25].

定义 5.9 假设 $(I;\mathcal{I},\lambda)$ 和 $(\Omega;\mathcal{F},P)$ 为两个概率空间, $\{f_i:i\in I\}$ 是 Ω 上的随机变量集合. 如果对 λ- 几乎所有的 $i\in I$, λ- 几乎所有的 $s\in I$ 和任何 $S\in\mathcal{F}$, 都有

$$\int_S f_i(\omega)f_s(\omega)dP=\int_S f_i(\omega)dP\int_S f_s(\omega)dP,$$

那么, $\{f_i:i\in I\}$ 称为基本两两相互独立的 (essentially pairwise independent).

定理 5.10 设 I 和 Ω 为两个超有限集, $(I;\mathcal{I},\lambda)$ 和 $(\Omega;\mathcal{F},P)$ 为两个 Loeb 空间. 设 $f:I\times\Omega\to\mathbb{R}$ 为一个 $\lambda\boxtimes P$- 平方可积函数, 则以下两个命题等价:

(1) $\{f(i,\cdot):i\in I\}$ 是基本两两相互独立的;

(2) 任给一个 $B\in\mathcal{I}$ 使得 $\lambda(B)>0$, 则对 P- 几乎所有的 $\omega\in\Omega$,

$$\int_B f(i,\omega)d\lambda=\int_B\int_\Omega f(i,\omega)dPd\lambda.$$

注意, 以上定理直观地讲就是, 给定足够多个互相独立的随机变量, 这些随机变量在几乎每个样本上的均值都等于所有这些随机变量在概率空间 Ω 上均值的均值.

从 Fubibi 扩张的定义与乘积 Loeb 空间的定义的关系中可以看出, 非标准分析在这里的应用是很难用标准方法来替代的.

精确大数定理在经济学中有广泛的应用, 如文献 [25–29]. 在此只列出一个在文献 [25] 中关于保险理论的定理. 定理中所涉及的各个概念还需由感兴趣的读者自己去探究.

定理 5.11 风险函数集 $\{f_i, i \in I\}$ 是可保险的 (insurable) 当且仅当它们是基本两两相互独立的.

6 非标准分析在组合数论中的应用

作者在过去二十多年里致力于非标准分析在组合数论中的应用, 获得了一些有趣的结果. 组合数论对于初学者来讲, 门槛比较低, 问题比较容易解释清楚. 所以, 非标准分析在此领域中的应用一旦有成, 就比较能使不了解非标准分析的学者对非标准分析的实用性产生信心.

非标准分析在讨论组合数论中的问题时会有哪些优点呢? 作者认为有两个方面值得一提.

第一, 非标准分析能简化一些证明的复杂度. 在非标准分析中, 一个标准的极限过程可以被简化成一个非标准的简单步骤. 例如, 一个数列 $\{a_n\}$ 包含一个收敛于 a 的子列, 可以叙述为存在一个超自然数 N 使得 a_N 无穷接近 a; 一个无限集合 $A \subseteq \mathbb{N}$ 在一个基数趋向无穷的有限区间序列上的局部密度趋向于 α, 可看成 ${}^*A$ 在一个超有限区间上的局部密度等于 α. 在组合数论中, 很多定理的证明都很长而繁琐, 如果在这类证明中非标准分析带来的简化不断地累积, 那么就能提供更清晰的思路和直观, 乃至关键突破. 在组合数论中有一类问题涉及自然数序列的各种密度, 而多种密度都是用极限来定义的, 所以, 非标准分析之所以在密度问题上得到应用不是偶然的.

第二, 非标准模型中的超整数或超有限集概念对同一问题提供了有限和无限两方面的工具. 从标准分析的角度看, 一个超有限集是无限集, 所以利用标准化映照很多连续性数学的工具可以被引入. 例如, Loeb 空间就是一例. 从非标准分析的角度看, 一个超有限集是有限的, 利用转换原理, 很多离散数学的结果就可以被应用, 所以超有限集的这两方面的特点相辅相成. 如果交替重复利用外部的无限性和内部的 $*$ 有限性, 非标准分析方法的内在潜力就可以展现出来. 在以下的讨论中, 特别是在第 6.3 节, 这个思想被不断应用于证明中. 曾有人在介绍非标准分析时说过, 为了能更好地研究无限集, 最好先把它扩充成 (非标准意义下的) 有限集.

现在介绍几个常用的渐近密度概念. 令 $\mathbb{N}$ 为包含 0 的所有自然数的集合, 令 $\mathbb{N}_+$ 为 $\mathbb{N} \setminus \{0\}$. 对任意的 $a, b \in \mathbb{N}$, 令 $A(a,b) := |A \cap [a,b]|$.

定义 6.1 设 $A \subseteq \mathbb{N}$ 为一个无限集.

(1) A 的上渐近密度 (upper asymptotic density)

$$\overline{d}(A) := \limsup_{n\to\infty} \frac{A(0,n)}{n+1};$$

(2) A 的下渐近密度 (lower asymptotic density)

$$\underline{d}(A) := \liminf_{n\to\infty} \frac{A(0,n)}{n+1};$$

(3) A 的上 Banach 密度 (upper Banach density)

$$\overline{\mathrm{BD}}(A) := \lim_{n\to\infty} \sup_{k\in\mathbb{N}} \frac{A(k,k+n)}{n+1};$$

(4) A 的下 Banach 密度 (lower Banach density)

$$\underline{\mathrm{BD}}(A) := \lim_{n\to\infty} \inf_{k\in\mathbb{N}} \frac{A(k,k+n)}{n+1}.$$

容易看出, 对任意的 $A \subseteq \mathbb{N}$,

$$0 \leqslant \underline{\mathrm{BD}}(A) \leqslant \underline{d}(A) \leqslant \overline{d}(A) \leqslant \overline{\mathrm{BD}}(A) \leqslant 1.$$

上下渐近密度是很自然的定义, 关于它们的讨论常能看到. 关于上下 Banach 密度的讨论相对比较少, 特别是下 Banach 密度. 但上 Banach 密度是在自然数集上建立 Banach 空间的重要概念, 读者可参见文献 [30].

在数论中还有一个常用的不很涉及极限的密度概念称为 Shnirel'man 密度, 有必要介绍一下, 但其不是我们要讨论的重点, 参见文献 [31].

定义 6.2 设 $A \subseteq \mathbb{N}$ 为无限集. A 的 Shnirel'man 密度 (Shnirel'man density)

$$\sigma(A) := \inf_{n\geqslant 1} \frac{A(1,n)}{n}.$$

容易看出, 对任意的 $A \subseteq \mathbb{N}$, $\sigma(A) \leqslant \underline{d}(A)$. 注意, 如果 $1 \notin A$, 那么 $\sigma(A) = 0$.

以下三个小节分别讨论三个不同的专题, 由于篇幅有限, 我们不能把所有牵涉到的定理都给出详细的证明. 我们将证明第 6.1 小节中的所有定理和第 6.2 小节中的部分定理. 第 6.3 小节中定理的证明较长, 我们只能给出粗略的解释. 好在所有的证明都能在所引的文献中找到.

6.1 和集现象

在标准实分析中, 一个 Lebesgue 测度为零的实数子集是一个小集合. 大家知道实数集 $\mathbb{R}$ 一定不是可数个零测集的并. 除了 Lebesgue 测度之外还有一个衡量实数子集大小的概念, 即第一纲/第二纲集. 令 X 为一个实数子集, 如果在每一个非空开区间 (a,b) 中总有一个非空子开区间 $(c,d) \subseteq (a,b)$ 使得 $(c,d) \cap X = \emptyset$, 集合 X 就称为无处稠密集. 如果一个实数子集 X 可以表示为最多可数个无处稠密集的并, 则 X 称为第一纲集. Baire 定理指出, 实数集 $\mathbb{R}$ 一定不是可数个第一纲集的并, 所以, 第一纲集是在序拓扑意义下的一种小集合. 但是在测度意义下的小集合概念和在序拓扑意义下的小集合概念是不可比较的. 这是因为由构造 Cantor 集的方法, 可以构造一个第一纲集 $A \subseteq \mathbb{R}$, 使得 $\mathbb{R} \setminus A$ 的 Lebesgue 测度为零. 所以, A 在测度意义下是一个大集合而在序拓扑意义下是一个小集合, 同时 $\mathbb{R} \setminus A$ 在测度意义下是小集合而在序拓扑意义下是个大集合.

当考虑和集时, 情形就不同了. 如果 A 和 B 是一个加法群中的两个子集, 定义

$$A \pm B := \{a \pm b : a \in A, b \in B\}.$$

以下是一个众所周知的定理.

定理 6.3 设 λ 为 $\mathbb{R}$ 上的 Lebesgue 测度, $A, B \subseteq \mathbb{R}$ 使得 $\lambda(A) > 0$ 和 $\lambda(B) > 0$, 则 $A + B$ 一定包含了一个非空开区间.

粗略地讲, 定理 6.3 表明了, 如果 A 和 B 都不是测度意义下的小集合, 则 $A + B$ 就不是序拓扑意义下的小集合. 此定理可用 Lebesgue 稠点定理来证明. 令 a 是一个实数子集 A 中的元素, 如果对任意的 $\epsilon > 0$ 总存在 $\delta > 0$ 使得对任意 $0 < \delta' < \delta$, $\lambda(A \cap (a - \delta', a + \delta'))/(2\delta') > 1 - \epsilon$, 则 a 称为 A

的稠点 (density point). Lebesgue 稠点定理指的是, 如果 A 是一个 Lebesgue 正测集, 那么几乎所有 A 中的点都是 A 的稠点. 如果 A 和 B 都是正测集, 令 a 和 b 分别是 A 和 B 的稠点, 再令 $\delta > 0$ 使得 $\lambda(A\cap(a-\delta,a+\delta))/(2\delta) > 2/3$ 和 $\lambda(B\cap(b-\delta,b+\delta))/(2\delta) > 2/3$. 用鸽巢原理 (pigeonhole principle) 就可证明 $A+B$ 包含了以 $a+b$ 为中心的非空开区间.

文献 [32] 第一次提出了以下和集现象 (sumset phenomenon): 假设在一个有序加群 G 上可定义一种类似测度的度量, 如果两个 G 的子集 A 和 B 都不是在这类测度意义下的小集合, 则 $A+B$ 就不是序拓扑意义下的小集合. 定理 6.3 是和集现象的一个证据. 是否还有其他证据?

在自然数集上, 密度概念与测度概念很相像, 但因为自然数的序拓扑是离散拓扑, 所以没有多大意义. 文献 [30] 提到了关于 $\mathbb{N}$ 的无限子集的一个概念称为 syndeticity, 似乎与稠密集的概念很接近. 让我们暂把 syndeticity 翻译成链接.

定义 6.4 设 $A\subseteq\mathbb{N}$ 是无限集.

(1) 如果存在一个自然数 k 使得对任意 $m\in\mathbb{N}$, $A\cap[m,m+k]\neq\emptyset$, 则 A 称为链接的 (syndetic), 即 A 在 $\mathbb{N}$ 中的缝隙长度 (gap size) 是有界的.

(2) 如果 A 包含了任意长的有限自然数区间, 则 A 称为厚实的 (thick).

(3) 如果 A 是一个链接集和一个厚实集的交, 则 A 称为局部链接的 (piecewise syndetic), 即存在一区间序列 $[a_n,b_n]$ 使得 $b_n-a_n\to\infty$, 并且 A 在 $\bigcup_{n\in\mathbb{N}}[a_n,b_n]$ 中的缝隙长度是有界的.

例如, 所有能被 10 整除的正整数集合是一链接集, 所有整数区间 $[2^n,2^n+n]$ 的并是一厚实集, 而所有在集合 $\bigcup_{n\in\mathbb{N}}[2^n,2^n+n]$ 中的 10 的整数倍形成了一个局部链接集. 从直观上讲, 链接集可对应于实数上的稠密集, 厚实集对应于实数上的开区间, 而局部链接集则对应于实数上的非无处稠密集.

定理 6.5 [30] 设 $A\subseteq\mathbb{N}$. 如果 $\overline{\mathrm{BD}}(A)>0$, 那么 $(A-A)\cap\mathbb{N}$ 是一个链接集.

注意, 利用类似于构造 Cantor 集的方法, 我们可以构造一个无限集 A 使得 $\underline{d}(A)>\frac{1}{2}$ 但 A 不是局部链接的.

虽然我们可以用模糊的语言来叙述关于自然数集的和集现象, 但如果能把这些概念加以严格处理应该是很有意义的. 特别是关于序拓扑, 我们怎样来严格定义一种拓扑来替代序拓扑呢? 另外, 在以上定理中, 集合 $A-A$ 能否被 $A+B$ 替代呢?

非标准分析方法是解决以上问题的合适方法. Keisler 和 Leth [33] 定义的一个类似于序拓扑的 U-拓扑正好是我们所需要的.

从现在起, 记号 $[a,b]$ 特指整数区间, 即 $[a,b]$ 是所有整数 c 的集合使得 $a\leqslant c\leqslant b$. 当 a 或 b 不是整数时此规定仍然有效.

定义 6.6 设 $U\subseteq{}^*\mathbb{N}$ 是非空前截, 即 $a\in U$ 推出 $[0,a]\subseteq U$, 且满足 $1\in U$ 和 $U+U\subseteq U$, 则 U 称为一个加法分割 (additive cut).

加法分割在文献 [33] 中就有应用, 如文献 [34].

加法分割是一个加法半群, 且把 ${}^*\mathbb{N}$ 分割成前截 U 和后截 ${}^*\mathbb{N}\setminus U$. 自然数集 ${}^*\mathbb{N}$ 上的加法分割有点像实分析中的 Dedekind 分割. 按照定义, ${}^*\mathbb{N}$ 本身是一个加法分割. 如果一个加法分割 $U\neq{}^*\mathbb{N}$, 那么 U 一定是一个外集. 这是因为 U 在 ${}^*\mathbb{N}$ 中有上界但没有最大元. 另外, $\mathbb{N}$ 是一个加法分割. 令 H 为一个超整数, 则

$$U_H := \bigcap_{n\in\mathbb{N}_+}\left[0,\frac{H}{n}\right] \tag{6.1}$$

也是一个加法分割. 容易看出 U_H 是 $[0,H]$ 中最大的加法分割, 而 $\mathbb{N}$ 是 $[0,H]$ 中最小的加法分割. 为了直观上的方便, 常把 $a\in U$ 写成 $a<U$, 并且把 $a\in{}^*\mathbb{N}\setminus U$ 写成 $a>U$.

定义 6.7 设 H 为一个超整数, $U \subseteq [0,H]$ 是一个加法分割. 如果一个集合 $O \subseteq [0,H]$ 满足性质: 对任意的 $a \in O$, 总存在 $r > U$ 使得 $[a-r, a+r] \cap [0,H] \subseteq O$, 则 O 称为一个 U- 开集.

所有 U- 开集组成 $[0,H]$ 上的 U- 拓扑.

定义 6.8 设 U 为 $[0,H]$ 中的一个加法分割. 对每个 $a,b \in [0,H]$, 定义

(1) $a \sim b$ 当且仅当 $|a-b| < U$;

(2) $[a] := \{c \in [0,H] : c \sim a\}$;

(3) $[0,H]/\!\sim \; := \{[a] : a \in [0,H]\}$;

(4) $[a] < [b]$ 当且仅当 $a < b$ 并且 $a \not\sim b$.

利用 U 的可加性质, 容易证明 $\sim$ 是 $[0,H]$ 上的一个等价关系, 每个 $[a]$ 都是一个等价类, 称为包含 a 的单子 (monad), 并且 $[0,H]/\!\sim$ 是一个自稠的全序集, 即任意两个单子之间存在另一个单子. 所以, $[0,H]/\!\sim$ 上的序拓扑就不是平凡的. 设 $\varphi(a) = [a]$ 是从 $[0,H]$ 到 $[0,H]/\!\sim$ 上的商映照, 则可以看出, 由 φ 引导出的 $[0,H]$ 上的商拓扑就是 $[0,H]$ 上的 U- 拓扑. 当然, U- 拓扑不满足 Hausdorff 性质, 但它是真正序拓扑的合适替代. 以下介绍 $[0,H]$ 上关于和集现象的定理.

定理 6.9 设 H 是一个超整数, U 是 $[0,H]$ 中的一个加法分割. 如果 A 和 B 是 $[0,H]$ 中的 Loeb 正测集, 那么, $A+B$ 一定不是 $[0,2H]$ 上的 U- 无处稠密集.

定理 6.9 可用多种方式证明, 例如, 可先证明 $[0,H]$ 上的和 Lebesgue 稠点定理平行的 Loeb 稠点定理, 再用其证明定理 6.9, 参见文献 [35]. 但这里仍用文献 [32] 中的方法, 因为它比较简单直观.

定理 6.9 的证明 首先, 集合 A 和 B 可以设为内集, 这是因为每个 Loeb 正测集包含一个 Loeb 正测内集.

假设对特定的 H 和 U, 定理 6.9 不成立, 我们将导出矛盾. 令 μ_H 表示 $[0,H]$ 上的 Loeb 测度. 为了方便, 如果 C 不是 $[0,H]$ 的子集, 就规定 $\mu_H(C) := \mu_H(C \cap [0,H])$. 设

$$\begin{aligned}\alpha := \sup\{&\mu_H(A) : \exists\, H > U, A \subseteq [0,H], \exists\, B \subseteq [0,H], A \text{ 和 } B \text{ 是内集},\\ &\mu_H(B) > 0, A+B \text{ 是 } [0,2H] \text{ 中 } U\text{- 无处稠密集}\}.\end{aligned}$$

因为假设定理不成立, 所以存在一个反例. 因此 $\alpha > 0$. 设

$$\begin{aligned}\beta := \sup\bigg\{&\mu_H(B) : \exists\, H > U, B \subseteq [0,H], \exists\, A \subseteq [0,H], A \text{ 和 } B \text{ 是内集},\\ &\mu_H(A) > \frac{9}{10}\alpha, A+B \text{ 是 } [0,2H] \text{ 中 } U\text{- 无处稠密集}\bigg\}.\end{aligned}$$

同样 $\beta > 0$. 因为 α 定义在先, 所以 $\beta \leqslant \alpha$. 由 α 和 β 的定义, 我们可以取定一个超整数 $H > U$, 取定内集 $A, B \subseteq [0,H]$ 使得 $\mu_H(A) > \frac{9}{10}\alpha$, $\mu_H(B) > \frac{9}{10}\beta$, 并且 $A+B$ 是 $[0,2H]$ 中的 U- 无处稠密集.

如果 $\alpha+\beta > \frac{10}{9}$, 那么 $\mu_H(A)+\mu_H(B) > 1$. 由鸽巢原理, 可推出 $A+B$ 包含了 $[0,2H]$ 中的单子 $[H]$. 由命题 4.1, 可以找到 $r > U$ 使得 $[H-r, H+r] \subseteq A+B$. 这与假设 $A+B$ 在 $[0,2H]$ 中 U- 无处稠密相矛盾. 所以可假设 $\alpha+\beta \leqslant \frac{10}{9}$ 和 $\beta \leqslant \frac{5}{9}$.

给定任意的 $m \in U$, 集合 $A+B+[0,m]$ 仍在 $[0,2H]$ 中 U- 无处稠密, 所以, 由 β 的定义有 $\mu_H(B+[0,m]) \leqslant \beta$. 设 $\mathcal{M} := \{m \in [0,H] : \frac{|B+[0,m]|}{H+1} < \frac{11}{10}\beta\}$. 由命题 4.2 可知, 集合 $\mathcal{M}$ 是内集. 由命题 4.1 可知, 存在 $m_0 > U$ 使得 $[0,m_0] \subseteq \mathcal{M}$.

现在取定 $K > U$ 使得 $K < \frac{1}{100}(H+1)$, $2K < m_0$, 并且存在 $x_0 \in [0, H-K]$ 使得 $\frac{1}{K}A(x_0, x_0+K-1) > \frac{9}{10}\alpha$. K 的存在是因为存在标准正实数 ϵ 使得 A 的元素多于 $\frac{9+\epsilon}{10}\alpha(H+1)$, 所以把区间 $[0,H]$

切割成长度为 K 的子区间 (最后的子区间可能长度小于 K), 并且 K 足够小的话, 其中必能找到一个 $[x_0, x_0+K-1]$ 使得 A 在其中的元素多于 $\frac{9}{10}\alpha K$.

令 $\mathcal{I} := \{[iK, (i+1)K-1] : i = 0, 1, \ldots, \lfloor H/K \rfloor\}$. 如果 $\mathcal{I}$ 中至少三分之二的区间包含 B 中的元, 那么 $B+[0, m_0]$ 包含了 $\mathcal{I}$ 中至少百分之六十的区间, 所以, $\mu_H(B+[0, m_0]) > 0.6 - 0.01 > 0.58 > \beta$, 这与 $\beta \leqslant \frac{9}{5}$ 矛盾. 所以可设 $\mathcal{I}$ 中至少三分之一的区间不含 B 中的元素, 即 B 中的元素都集中在 $\mathcal{I}$ 中最多三分之二的区间之中, 因此, 这至少三分之二的区间中必定有一个 $[y_0, y_0+K-1]$ 使得 $B \cap [y_0, y_0+K-1]$ 包含了至少 $\frac{3}{2}\beta K$ 的元素.

设 $A' := A \cap [x_0, x_0+K-1] - x_0$, $B' := B \cap [y_0, y_0+K-1] - y_0$. 因为 $\mu_{K-1}(A') > \frac{9}{10}\alpha$, $\mu_{K-1}(B') \geqslant \frac{3}{2}\beta$, 所以, 由 β 的定义可知, $A'+B'$ 不是在 $[0, 2K]$ 中 U- 无处稠密的, 因此, $A+B$ 也不是在 $[0, 2H]$ 中 U- 无处稠密的, 这与 A 和 B 的选择矛盾. 定理得证. □

下面来看为什么定理 6.3 是定理 6.9 的简单推论. 设 A 和 B 是实数单位区间 I 上的两个 Lebesgue 正测集. 我们要证明 $A+B$ 包含一个非空开区间. 给定一个超整数 H, 定义映照 $f : [0, 2H] \to \mathbb{R}$ 使得 $f(x) := \mathrm{st}(x/H)$. 令 $U = U_H$, 这里 U_H 是在 (6.1) 中所定义的加法分割. 容易证明, 对任意 $a \in [0, 2H]$, $f^{-1}(f(a)) = [a]$. 事实上, $[0, H]/\sim$ 是与标准实数单位区间 I 序同构的. 另外也可证明 $f^{-1}(A)$ 和 $f^{-1}(B)$ 是 Loeb 可测集, 并且其 Loeb 测度分别等于 A 和 B 的 Lebesgue 测度 (可以先设 A 和 B 为开集或闭集, 然后推广到 A 和 B 为 Borel 集, 再推广到 Lebesgue 可测集). 设 $A' \subseteq f^{-1}(A)$ 和 $B' \subseteq f^{-1}(B)$ 使得 A' 和 B' 是 $[0, H]$ 中的 Loeb 正测内集.

由定理 6.9, $A'+B'$ 不是在 U- 无处稠密的, 所以存在区间 $[a, b] \subseteq [0, 2H]$ 使得 $b-a > U$, 并且 $[a, b]$ 中任意长度大于 U 的区间都含有 $A'+B'$ 中的元素. 因为 $A'+B'$ 是内集, 所以, $A'+B'$ 在 $[a, b]$ 中最大空隙存在且长度小于 U. 现在容易证明 $f(a) < f(b)$ 并且 $(f(a), f(b)) \subseteq f(A'+B') \subseteq A+B$. 所以, $A+B$ 包含了非空开区间 $(f(a), f(b))$.

在以上证明中, 我们使用了在 (6.1) 中定义的加法分割 U_H. 如果使用其他的加法分割 (如 $U = \mathbb{N}$) 会有什么结果呢? 下面将使用 $U = \mathbb{N}$ 来证明自然数集上的和集现象. 首先需要上 Banach 密度和局部链接性的的非标准分析等价条件.

引理 6.10 设 $A \subseteq \mathbb{N}$ 且 $\alpha \in \mathbb{R}$, 则 $\overline{\mathrm{BD}}(A) \geqslant \alpha$ 当且仅当存在长度为超整数的区间 $[a, b]$ 使得 $\mathrm{st}({}^*A(a, b)/(b-a+1)) \geqslant \alpha$.

证明 必要性. 由上 Banach 密度的定义, 对每个 $k \in \mathbb{N}_+$ 能找到 $[a_k, b_k] \subseteq \mathbb{N}$ 使得 $b_k - a_k > k$, $A(a_k, b_k)/(b_k - a_k + 1) > \alpha - 1/k$. 由溢出原理, 存在一个超整数 K 满足 $b_K - a_K > K$ 和 ${}^*A(a_K, b_K)/(b_K - a_K + 1) > \alpha - 1/K \approx \alpha$. 显然, $[a, b] = [a_K, b_K]$ 满足要求.

充分性. 任给 $k \in \mathbb{N}$. 在 ${}^*\mathcal{V}$ 中存在 $a, b \in {}^*\mathbb{N}$ 满足 $b - a > k$ 和 ${}^*A(a, b)/(b-a+1) > \alpha - 1/k$. 由转换原理, 在 $\mathcal{V}$ 中存在 $a_k, b_k \in \mathbb{N}$ 满足 $b_k - a_k > k$ 和 $A(a_k, b_k)/(b_k - a_k + 1) > \alpha - 1/k$, 而这两性质推出 $\overline{\mathrm{BD}}(A) \geqslant \alpha - 1/k$. 因为 k 是任意的, 所以 $\overline{\mathrm{BD}}(A) \geqslant \alpha$. 引理证毕. □

引理 6.11 设 $A \subseteq \mathbb{N}$.

(1) 集合 A 是厚实的当且仅当存在长度为超整数的区间 $[a, b] \subseteq^* A$;

(2) 集合 A 是局部链接的当且仅当存在长度为超整数的区间 $[a, b]$ 和有限自然数 $k \in \mathbb{N}$ 使得 ${}^*A$ 在 $[a, b]$ 中的缝隙的长度小于 k, 即 $g = \max\{d - c : [c, d] \subseteq [a, b] \setminus {}^*A\} < k$.

证明 (1) 只需转换原理的简单应用. 我们只证明 (2).

必要性. 设 A 是链接集 B 和厚实集 C 的交, 则 ${}^*A = {}^*B \cap {}^*C$. 因为 B 在 $\mathbb{N}$ 中缝隙的长度小于一个自然数 k, 所以, 由转换原理, ${}^*B$ 在 ${}^*\mathbb{N}$ 中缝隙的长度小于 k. 因为 C 是厚实集, 由 (1) 知, 存在一个超有限区间 $[a, b] \subseteq^* C$. 显然, ${}^*A$ 在 $[a, b]$ 中缝隙的长度小于 k.

充分性. 设 $g = \max\{d - c : [c,d] \subseteq [a,b] \setminus {}^*A\}$. 因为在 ${}^*\mathcal{V}$ 中, 整数 g 的定义只涉及有限集或超有限集, 所以 g 存在. 当然, 由引理的假设, g 一定是有限的. 对任意的 $m \in \mathbb{N}$, 语句 "存在区间 $[a,b]$ 满足 $b - a > m$, 并且 ${}^*A$ 在 $[a,b]$ 中的缝隙的长度 $\leqslant g$" 在 ${}^*\mathcal{V}$ 中成立, 由转换原理, 存在区间 $[a_m, b_m] \subseteq \mathbb{N}$ 满足 $b_m - a_m > m$, 并且 A 在 $[a_m, b_m]$ 中缝隙的长度 $\leqslant g$. 令 $C = A \cup \bigcup_{m\in\mathbb{N}}[a_m, b_m]$ 和 $B = (\mathbb{N} \setminus \bigcup_{m\in\mathbb{N}}[a_m, b_m]) \cup A$. 显然, B 是链接集, C 是厚实集, 并且 $A = B \cap C$. 引理证毕. □

定理 6.12 设 $A, B \subseteq \mathbb{N}$ 使得 $\overline{\mathrm{BD}}(A) > 0$ 和 $\overline{\mathrm{BD}}(B) > 0$, 则 $A + B$ 一定是局部链接的.

证明 由引理 6.10, 存在超有限区间 $[x, x + K_1]$ 和 $[y, y + K_2]$ 使得 $\mathrm{st}({}^*A(x, x + K_1)/(K_1 + 1)) \geqslant \overline{\mathrm{BD}}(A)$ 和 $\mathrm{st}({}^*B(y, y + K_2)/(K_2 + 1)) \geqslant \overline{\mathrm{BD}}(B)$. 事实上, 我们可假设 $K_1 = K_2 = K$, 因为不然的话, 取 $K = \lceil\sqrt{\min\{K_1, K_2\}}\rceil$, 再把 $[x, x + K_1]$ 和 $[y, y + K_2]$ 分割成长度为 $K + 1$ 的子区间 (最后一个区间的长度可能小于 $K + 1$), 然后在这些子区间中可找到 $[x', x' + K]$ 和 $[y', y' + K]$ 使得 $\mathrm{st}({}^*A(x', x' + K)/(K + 1)) \geqslant \overline{\mathrm{BD}}(A)$ 和 $\mathrm{st}({}^*B(y', y' + K)/(K + 1)) \geqslant \overline{\mathrm{BD}}(B)$, 再用 K 替代 K_1 和 K_2. 由定理 6.9, 集合 ${}^*(A + B)$ 在 $[x + y, x + y + 2K]$ 中不是 U- 无处稠密的, 所以存在超有限区间 $[a,b] \subseteq [x + y, x + y + 2K]$ 和有限自然数 $k \in \mathbb{N}$ 使得 ${}^*(A + B)$ 在 $[a,b]$ 中缝隙的长度都小于 k. 由引理 6.11, $A + B$ 是局部链接的. 证毕. □

注意, 定理 6.12 是对定理 6.5 的一个充实. 在定理 6.12 的结论中, 局部链接不能被改成链接, 这是加法和减法本质上的差别.

在过去十几年中, 定理 6.12 被不断推广或重证. 例如, 文献 [36, 37] 把定理 6.12 推广至 amenable 群并建立了等价关系; 文献 [38] 用超滤子作为工具重证了定理 6.12; 文献 [39] 用组合方法重证了定理 6.12 并给出了一些量化信息. 文献 [35] 中的一个定理把以上定理 6.12 推广至上渐近密度和下渐近密度.

在发现定理 6.12 的过程中非标准分析无疑发挥了重要的作用. 只有通过定理 6.9, 连续情形和离散情形下的和集现象才能被严格地统一起来. 在定理 6.9 的应用中, 我们只用到了 $U = U_H$ 和 $U = \mathbb{N}$ 两种情形. 对于其他类型加法分割的应用, 还有待于进一步探究.

6.2 Plünnecke 类型不等式

设 $A \subseteq \mathbb{N}$. 在定义 6.2 中, 我们定义了 A 的 Shnirel'man 密度, 记为 $\sigma(A)$. 注意, 我们要求 $0 \in \mathbb{N}$, 但 $\sigma(A)$ 的值不依赖于 A 是否包含元素 0.

定义 6.13 对于一个包含了 0 的集合 $B \subseteq \mathbb{N}$, 如果对任何 $A \subseteq \mathbb{N}$, $0 < \sigma(A) < 1$ 推出

$$\sigma(A + B) > \sigma(A),$$

我们就称 B 为本质部件 (essential component).

什么集合是本质部件呢? 以下定理称为 Shnirel'man 定理, 其证明可在文献 [31, 40] 中找到.

定理 6.14 如果 $0 \in A \cup B$, 那么 $\sigma(A + B) \geqslant \sigma(A) + \sigma(B)(1 - \sigma(A))$.

由定理 6.14 可以看出, 如果 $0 \in B$ 并且 $\sigma(B) > 0$, 那么 B 就是一个本质部件. 但一个本质部件不一定有正 Shnirel'man 密度. 例如, $S := \{m^2 : m \in \mathbb{N}\}$ 是一个本质部件 (参看以下的 Erdős 定理), 但是 $\sigma(S) = 0$.

定义 6.15 设 $h \in \mathbb{N}_+$. 对一个集合 $B \subseteq \mathbb{N}$,

(1) 如果

$$hB := \underbrace{B + B + \cdots + B}_{h} = \mathbb{N},$$

我们就称 B 为 h 维基集 (basis of order h); 如果 B 是有限维基集, 我们就简称 B 为基集.

(2) 如果存在 $m \in \mathbb{N}$ 使得

$$hB \supseteq \mathbb{N} \setminus [0, m],$$

我们就称 B 为 h 维尾基集 (asymptotic basis of order h); 如果 B 是有限维尾基集, 我们就简称 B 为尾基集.

注意, 由著名的 Lagrange 定理, 以上集合 $S = \{m^2 : m \in \mathbb{N}\}$ 是 4 维基集.

在 20 世纪 30 年代, Erdős 证明了以下定理. 定理的证明可在文献 [40] 中找到.

定理 6.16 如果 $B \subseteq \mathbb{N}$ 是 h 维基集, 那么对任意 $A \subseteq \mathbb{N}$,

$$\sigma(A+B) \geqslant \sigma(A) + \frac{1}{2h}\sigma(A)(1-\sigma(A)).$$

在 20 世纪 70 年代, Plünnecke 改进了 Erdős 的结果, 在文献 [41] 中证明了以下定理.

定理 6.17 如果 $B \subseteq \mathbb{N}$ 是 h 维基集, 那么对任意 $A \subseteq \mathbb{N}$,

$$\sigma(A+B) \geqslant \sigma(A)^{1-\frac{1}{h}}.$$

利用一元微分方法就可证明, 如果 $0 < \sigma(A) < 1$, 那么 $\sigma(A)^{1-\frac{1}{h}} > \sigma(A) + \frac{1}{h}\sigma(A)(1-\sigma(A))$, 其证明只需考虑函数 $f(x) = x^{1-\frac{1}{h}} - x - \frac{1}{h}x(1-x)$ 在 $(0,1)$ 上的单调性即可.

Plünnecke 定理的重要性不单是改进了 Erdős 定理, 更重要的是引进了一种全新的方法, 即引入了称为 Plünnecke 图的工具. 对 Plünnecke 理论感兴趣的读者可参见文献 [42]. 这里所需要的是 Plünnecke 理论中一个重要的不等式, 其证明可在文献 [42] 中找到.

引理 6.18 给定集合 $A, B \subseteq \mathbb{N}$ 和 $n, h \in \mathbb{N}_+$ 使得 $A(0,n) > 0$, 则一定存在非空集 $A' \subseteq A \cap [0,n]$ 满足

$$\frac{(A+B)(0,n)}{A(0,n)} \geqslant \left(\frac{(A'+hB)(0,n)}{A'(0,n)}\right)^{1/h}.$$

在定理 6.17 中, Shnirel'man 密度是否可以被其他密度来替换呢? 为了回答这个问题, 我们首先要探讨在其他密度下基集的概念应该作怎么样的调整. 注意, 当 $0 \in B$ 时, B 是一个 h 维基集当且仅当 $\sigma(hB) = 1$.

定义 6.19 设 $B \subseteq \mathbb{N}$, $h > 0$.

(1) 若 $\overline{d}(hB) = 1$, 则 B 称为 h 维上渐近基集 (upper asymptotic basis of order h);

(2) 若 $\underline{d}(hB) = 1$, 则 B 称为 h 维下渐近基集 (lower asymptotic basis of order h);

(3) 若 $\overline{\mathrm{BD}}(hB) = 1$, 则 B 称为 h 维上 Banach 基集 (upper Banach basis of order h);

(4) 若 $\underline{\mathrm{BD}}(hB) = 1$, 则 B 称为 h 维下 Banach 基集 (lower Banach basis of order h).

例 6.20 令 $P := \{p \in \mathbb{N} : p \text{ 是素数}\}$, $C := \{m^3 : m \in \mathbb{N}\}$. Estermann, Chudakov 和 van der Corput 证明了集合 P 是 3 维下渐近基集 (参见文献 [43]) 和 4 维尾基集. 如果 Goldbach 猜想成立, 则 P 是 3 维尾基集, 集合 C 是 4 维下渐近基集和 7 维尾基集 (参见文献 [44]).

以下定理是文献 [45, 46] 中结果的细微推广.

定理 6.21 给定任意 $A, B \subseteq \mathbb{N}$ 和 $h > 0$, 则

(1) $\underline{d}(A+B) \geqslant \underline{d}(A)^{1-\frac{1}{h}}\underline{d}(hB)^{\frac{1}{h}}$;

(2) $\overline{\mathrm{BD}}(A+B) \geqslant \overline{\mathrm{BD}}(A)^{1-\frac{1}{h}}\overline{\mathrm{BD}}(hB)^{\frac{1}{h}}$;

(3) $\underline{\mathrm{BD}}(A+B) \geqslant \underline{\mathrm{BD}}(A)^{1-\frac{1}{h}}\overline{\mathrm{BD}}(hB)^{\frac{1}{h}}$.

我们只证明定理 6.21(1). 定理 6.21(2) 和 6.21(3) 的证明参见文献 [45]. 在证明定理 6.21(1) 之前, 我们先证明一个下渐近密度的非标准等价条件.

引理 6.22 设 $A \subseteq \mathbb{N}$ 和 $\alpha \in \mathbb{R}$, 则 $\underline{d}(A) \geqslant \alpha$ 当且仅当对任意超整数 H,

$$\mathrm{st}\left(\frac{{}^*A(0,H)}{H+1}\right) \geqslant \alpha.$$

证明 必要性. 给定一个超整数 H. 对任意的 $k \in \mathbb{N}_+$, 存在 $N \in \mathbb{N}$ 使得语句 "对任意的 $m \geqslant N$, 有 $A(0,m)/(m+1) > \alpha - 1/k$" 在 $\mathcal{V}$ 中为真. 所以, 由转换原理, 语句 "对任意的 $m \geqslant N$, 有 ${}^*A(0,m)/(m+1) > \alpha - 1/k$" 在 ${}^*\mathcal{V}$ 中为真. 显然, $H > N$, 所以 ${}^*A(0,H)/(H+1) > \alpha - 1/k$. 因为 H 不依赖于 k 的选择, 所以 $\mathrm{st}({}^*A(0,H)/(H+1)) \geqslant \alpha$.

充分性. 给定任意的 $k \in \mathbb{N}_+$, 语句 "存在 $N \in {}^*\mathbb{N}$ 使得对任意 $m \geqslant N$, ${}^*A(0,m)/(m+1) > \alpha - 1/k$" 在 ${}^*\mathcal{V}$ 中为真. 由转换原理, "存在 $N \in \mathbb{N}$ 使得对任意 $m \geqslant N$, $A(0,m)/(m+1) > \alpha - 1/k$" 在 $\mathcal{V}$ 中为真. 而这表示 $\underline{d}(A) \geqslant \alpha$. 证毕. □

定理 6.21 的证明 设 $\alpha = \underline{d}(A)$ 和 $\beta = \underline{d}(hB)$. 如果 $\alpha = 0$, 则不等式显然成立, 所以可假设 $\alpha > 0$. 任给超整数 H, 由引理 6.22, 我们只需证明

$$\mathrm{st}\left(\frac{{}^*(A+B)(0,H)}{H+1}\right) \geqslant \alpha^{1-\frac{1}{h}}\beta^{\frac{1}{h}}.$$

由引理 6.22, 有 $\mathrm{st}({}^*A(0,H)/(H+1)) \geqslant \alpha$. 令 $K = H - \lceil\sqrt{H}\rceil$. 显然, $K/H \approx 1$, $\mathrm{st}({}^*A(0,K)/(H+1)) \geqslant \alpha$. 取 $C_0 :=^* A \cap [0,K]$. 我们用内递归定义一个 C_0 中递减的子集内序列 $C_0 \supseteq C_1 \supseteq \cdots \supseteq C_H$ 满足

$$C_{k+1} = \begin{cases} C_k, & \text{如果 } C_k(H-k,H) \leqslant \alpha(k+1), \\ C_k \setminus \{H-k\}, & \text{如果 } C_k(H-k,H) > \alpha(k+1). \end{cases}$$

令 $A_0 = C_H$. 接下来验证 $\mathrm{st}(A_0(0,H)/(H+1)) \geqslant \alpha$, 并且对任意的 $z \in [0,H]$, 有

$$\frac{A_0(z,H)}{H-z+1} \leqslant \alpha.$$

假设存在一个标准正实数 ϵ 使得 $A_0(0,H)/(H+1) < \alpha - \epsilon$, 定义 $k_0 = \max\{k \in [0,H] : C_{k+1} \neq C_k\}$, 则整数 k_0 存在. 这是因为, 否则的话有 $A_0 = C_0$, 而这与 $\mathrm{st}(C_0(0,K)/(H+1)) \geqslant \alpha$ 矛盾. 因为 $C_0 \cap [H-K+1,H] = \emptyset$, 所以有 $H - k_0 < K$. 由 k_0 的定义有 $A_0 = C_{k_0+1}$, $A_0 \cap [0, H-k_0-1] = C_0 \cap [0, H-k_0-1]$ 和 $C_{k_0}(H-k_0,H) > \alpha(k_0+1)$.

假设 $H - k_0 \in \mathbb{N}$, 则

$$\begin{aligned}\mathrm{st}\left(\frac{A_0(0,H)}{H+1}\right) &= \mathrm{st}\left(\frac{A_0(0,H-k_0-1)}{H+1}\right) + \mathrm{st}\left(\frac{A_0(H-k_0,H)}{H+1}\right) \\ &= \mathrm{st}\left(\frac{C_{k_0+1}(H-k_0,H)}{k_0+1}\cdot\frac{k_0+1}{H+1}\right) = \mathrm{st}\left(\frac{C_{k_0}(H-k_0,H)-1}{k_0+1}\right) \geqslant \alpha.\end{aligned}$$

假设 $H - k_0$ 是超整数, 则

$$\begin{aligned}\mathrm{st}\left(\frac{A_0(0,H)}{H+1}\right) &= \mathrm{st}\left(\frac{A_0(0,H-k_0-1)}{H+1}\right) + \mathrm{st}\left(\frac{A_0(H-k_0,H)}{H+1}\right) \\ &\geqslant \mathrm{st}\left(\frac{C_0(0,H-k_0-1)}{H-k_0}\cdot\frac{H-k_0}{H+1}\right) + \mathrm{st}\left(\frac{C_{k_0}(H-k_0,H)-1}{k_0+1}\cdot\frac{k_0+1}{H+1}\right)\end{aligned}$$

$$\geqslant \alpha \cdot \mathrm{st}\left(\frac{H-k_0}{H+1}\right) + \alpha \cdot \mathrm{st}\left(\frac{k_0+1}{H+1}\right) = \alpha.$$

因为以上两种情形都与 $A_0(0,H)/(H+1) < \alpha - \epsilon$ 相矛盾, 所以, $\mathrm{st}(A_0(0,H)/(H+1)) \geqslant \alpha$ 一定成立.

假设存在 $z \in [0,H]$ 使得 $A_0(z,H) > \alpha(H-z+1)$. 令

$$z_0 := \max\{z \in [0,H] : A_0(z,H) > \alpha(H-z+1)\}.$$

因为 $A_0(K+1,H) \leqslant C_0(K+1,H) = 0$, 所以 $z_0 \leqslant K$. 因为 $C_{H-z_0}(z_0,H) \geqslant A_0(z_0,H) > \alpha(H-z_0+1)$, 所以由 C_{H-z_0+1} 的定义, 有 $z_0 \notin C_{H-z_0+1}$. 因此 $z_0 \notin A_0$. 而这推出

$$\frac{A_0(z_0+1,H)}{H-z_0} = \frac{A_0(z_0,H)}{H-z_0} > \frac{A_0(z_0,H)}{H-z_0+1} > \alpha.$$

这与 z_0 的极大性矛盾, 所以, 对任意的 $z \in [0,H]$, 都有

$$\frac{A_0(z,H)}{h-z+1} \leqslant \alpha.$$

由引理 6.18 中的不等式, 我们可以找到非空集和 $A' \subseteq A_0$ 使得 $z = \min A' \leqslant K$ 满足

$$\begin{aligned}
\mathrm{st}\left(\frac{{}^*(A+B)(0,H)}{H+1}\right) &\geqslant \mathrm{st}\left(\frac{(A_0 +^* B)(0,H)}{A_0(0,H)}\right) \cdot \mathrm{st}\left(\frac{A_0(0,H)}{H+1}\right) \\
&\geqslant \mathrm{st}\left(\left(\frac{(A' + h^*B)(0,H)}{A'(0,H)}\right)^{\frac{1}{h}}\right) \cdot \mathrm{st}\left(\frac{A_0(0,H)}{H+1}\right) \\
&\geqslant \mathrm{st}\left(\left(\frac{(z + h^*B)(z,H)}{A'(z,H)}\right)^{\frac{1}{h}}\right) \cdot \alpha \\
&\geqslant \mathrm{st}\left(\left(\frac{(h^*B)(0,H-z)/(H-z+1)}{A_0(z,H)/(H-z+1)}\right)^{\frac{1}{h}}\right) \cdot \alpha \\
&\geqslant \left(\frac{\beta}{\alpha}\right)^{\frac{1}{h}} \cdot \alpha = \alpha^{1-\frac{1}{h}} \beta^{\frac{1}{h}},
\end{aligned}$$

这里 $\mathrm{st}((h^*B)(0,H-z)/(H-z+1)) \geqslant \beta$, 是因为引理 6.22 和 $H - z \geqslant H - K$ 是一个超整数. 至此, 由引理 6.22, 定理得证. □

虽然定理 6.21(2) 和 6.21(3) 的证明需要一些不一样的技巧, 但总体思路与定理 6.21(1) 的证明类似.

定理 6.21(1) 在文献 [47] 中也涉及但无证明.

在定理 6.21(3) 的不等式中, 有趣的一点是最后一项不是 $\underline{\mathrm{BD}}(hB)^{\frac{1}{h}}$ 而是 $\overline{\mathrm{BD}}(hB)^{\frac{1}{h}}$.

在定理 6.21 中, 如果 B 是 h 维下渐近基集或上 Banach 基集, 则相应的最后一项 $\underline{d}(hB)^{\frac{1}{h}}$ 或 $\overline{\mathrm{BD}}(hB)^{\frac{1}{h}}$ 等于 1, 所以就变成了 Plünnecke 类型的不等式.

定理 6.21 中没有关于上渐近密度的不等式, 这是因为关于上渐近密度的 Plünnecke 不等式不成立. 在文献 [45] 中有一个例子, 即存在一个二维上渐近基集 B 和一集合 A 使得

$$\overline{d}(A+B) = \overline{d}(A) = \frac{1}{2}.$$

定理 6.21(1) 有以下推论. 令 P 表示所有素数的集合, C 表示所有自然数的三次方的集合.

推论 6.23　任给 $A \subseteq \mathbb{N}$,

(1) $\underline{d}(A+P) \geqslant \underline{d}(A)^{2/3}$;

(2) $\underline{d}(A+C) \geqslant \underline{d}(A)^{3/4}$.

除了以上所介绍的密度外, 还有对数上渐近密度和对数下渐近密度. 对于这两种密度, 我们不知道 Plünnecke 类型不等式是否成立.

6.3　Freiman 逆问题

因为本小节中所要介绍的结果其证明都太长, 所以就只能介绍结果而不给证明. 有兴趣的读者可以在所引文献中找到这些证明.

在可加数论中, 人们一般假设集合 A 和 B 满足一定条件, 然后讨论集合 $A \pm B$ 或 hA 的性质. 例如, Goldbach 猜想就是预测, 如果 A 是所有素数的集合, 那么 $A+A$ 就包含了所有大于 3 的偶数. 再例如, Waring 问题就是, 假设 A 是所有自然数的 k 次方, 然后讨论是否存在 h 或者 h 有多大使得 $hA = \mathbb{N}$.

可加数论中的逆问题从相反的方向来考虑. 任给集合 (有限或无限) A 和 B, 如果 $A \pm B$ 满足一定的性质, 我们希望能获得一些关于 A 和 B 的性质.

在 20 世纪 50 年代末 60 年代初, Freiman 获得了一系列的结果, 印证了一种现象, 即所称的 Freiman 逆现象: 如果 $A+A$ 是一个小集合, 那么集合 A 就有一些算术结构 (arithmetic structure). 以上是个模糊的叙述, 在其中有两个概念有待进一步说明:

(1) "小" 的标准是什么?

(2) 什么是 "算术结构"?

如果存在实数 a、正实数 d 和正整数 k 使得集合 $A = \{a, a+d, a+2d, \ldots, a+(k-1)d\}$, 则 A 称为等差为 d、长度为 k 的等差数列 (等差数列有时也称算术数列). 如果 $A = \{a, a+d, a+2d, \ldots\}$ 没有最大元, 我们也称 A 为等差数列. 因为在本小节中只考虑整数, 所以, a 和 d 一般都是整数.

关于 Freiman 逆现象的最简单事实是, 设 A 是一个有限整数集合, $|A| = k$, 如果 $|2A| = 2k-1$, 则可以推出 A 是等差数列 (即 A 有算术结构). 注意, 对任意包含 k 个整数的集合 A, $|2A| \geqslant 2k-1$ 一定成立. 所以, $2k-1$ 是和集 $2A$ 可能达到的最小基数. 同样, 如果 $|A+B| = |A|+|B|-1$, 则可以推出 A 和 B 都是等差数列且具有相同的等差. 如果把 $|2A|$ 的上界放松, 如 $|2A| \leqslant 2k+3$, A 会是个什么样的集合呢? 又例如, 给定一个正数 C, 如果集合 A 足够大且满足 $|2A| < Ck$, A 又会是个什么样的集合呢? 注意, $2A$ 的最大可能基数是 k 的二次函数, 即 $|2A| \leqslant \frac{1}{2}k(k+1)$.

定义 6.24　设 a 是一个整数, d 是一个正整数, $\{d_i : i = 1, 2, \ldots, d\}$ 和 $\{k_i : i = 1, 2, \ldots, d\}$ 是正整数集合. 集合

$$P := \left\{a + \sum_{i=1}^{d} d_i x_i : 0 \leqslant x_i < k_i, i = 1, 2, \ldots, d\right\}$$

称为 d 维广义等差数列 (generalized arithmetic progression).

以下定理 6.25 是 Freiman 众多定理中最著名的一个, 称为大双倍常数 (large doubling constant) 的 Freiman 定理. 其证明可在文献 [42, 48, 49] 中找到.

定理 6.25　任给实数 $C > 2$, 存在正实数 c 和 K, 使得对任意有限整数集合 A, 如果 $|A| > K$ 和 $|2A| < C|A|$, 则 A 是一个 $\lfloor C-1 \rfloor$ 维广义等差数列 P 的子集, 并且 $|A| > c|P|$.

定理 6.25 中常数 C 称为双倍常数.

定理 6.25 作为主要工具之一被用在了 Fields 奖获奖者 Gowers 的著名文献 [50] 中. 在过去将近二十年中, 人们在常数 c 的优化上做了很多工作.

定义 6.26 如果 I 和 J 是两个具有相同等差的等差数列使得三个集合 $2I$, $I+J$ 和 $2J$ 两两互不相交, 则集合 $I \cup J$ 称为双重等差数列 (bi-arithmetic progression).

注意, 双重等差数列和二维广义等差数列完全不是一回事.

在本小节中, Freiman 大双倍常数问题不是我们主要的兴趣所在. 我们主要的兴趣是 Freiman 小双倍常数 (small doubling constant) 问题. 下列定理 6.27 常称为 Freiman 小双倍常数定理.

定理 6.27 设 A 是有限整数集.

(1) 如果 $|A| > 2$ 并且 $|2A| = 2|A| - 1 + b < 3|A| - 3$, 则 A 是一个等差数列 I 的子集并且 $|I| \leqslant |A| + b$;

(2) 如果 $|A| > 6$ 并且 $|2A| = 3|A| - 3$, 则 A 是一个等差数列 I 的子集并且 $|I| \leqslant 2|A| - 1$, 或 A 是一个双重等差数列.

以上两个等差数列 I 长度的上界都是最优的.

例 6.28 设 $k \geqslant c+4$ 和 $A := [0, k-2] \cup \{k+c\}$, 则 $|A| = k$ 和 $|2A| = 2k - 1 + (c+1) < 3k - 3$. 包含 A 的最短等差数列是 $I = [0, k+c]$, 其长度为 $k + (c+1)$.

设 $k \geqslant 7$ 和 $A := [0, k-3] \cup \{k-1, 2k-2\}$, 则 $|A| = k$ 和 $|2A| = 3k - 3$. 集合 A 不是一个双等差数列, 并且包含 A 的最短等差数列是 $I = [0, 2k-2]$, 其长度为 $2k - 1$.

Freiman 的以上两个定理都说明, 如果和集 $2A$ 的基数相对较小, 那么 A 就是一个等差或广义等差数列的相对较大的子集 (即 A 有算术结构). 在定理 6.25 中, 对和集 $2A$ 的基数的限制较弱, 所得到的关于 A 的结构的信息就比较少; 而在定理 6.27 中, 对和集 $2A$ 的基数限制比较强, 所以得到的关于 A 的结构的信息就比较多, 比较精确.

Freiman 的逆现象在无限自然数集上也有体现. 如果 A 是个无限集, 用基数来衡量 A 的大小显然不合适. 所以, 我们就用密度来代替. 以下定理, 称为 Kneser 定理, 解决了一个在下渐近密度意义下的 Freiman 逆问题, 其证明可在文献 [40] 中找到.

定理 6.29 设 $A, B \subseteq \mathbb{N}$ 满足 $\underline{d}(A+B) < \underline{d}(A) + \underline{d}(B)$, 则存在 $g \in \mathbb{N}_+$ 和集合 $G \subseteq [0, g-1]$ 使得

(1) $A + B \subseteq G + \{gn : n \in \mathbb{N}\}$;

(2) $(G + \{gn : n \in \mathbb{N}\}) \setminus (A+B)$ 是有限的;

(3) $\underline{d}(A+B) \geqslant \underline{d}(A) + \underline{d}(B) - \frac{1}{g}$.

定理 6.29 与以下定理等价. 等价性证明可在文献 [51] 中找到.

定理 6.30 设 $A, B \subseteq \mathbb{N}$ 满足 $\underline{d}(A+B) < \underline{d}(A) + \underline{d}(B)$, 则存在 $g \in \mathbb{N}_+$ 和集合 $F, F' \subseteq [0, g-1]$ 使得 $A \subseteq F + \{gn : n \in \mathbb{N}\}$, $B \subseteq F' + \{gn : n \in \mathbb{N}\}$, 并且

$$\underline{d}(A) + \underline{d}(B) > \frac{|F| + |F'| - 1}{g}.$$

定理 6.30 似乎比定理 6.29 更接近于 Freiman 逆现象, 这是因为定理 6.30 分别叙述了集合 A 和 B 的等差数列结构. 注意, $X + \{gn : n \in \mathbb{N}\}$ 是 $|X|$ 个具有相同等差的等差数列之并.

接下来介绍作者的几个关于 Freiman 逆问题的结果. 这些结果的证明都依赖于非标准分析方法.

首先有关于上渐近密度的 Freiman 逆问题的结果, 其证明可在文献 [52] 中找到.

定理 6.31 设 $A \subseteq \mathbb{N}$ 满足 $0 \in A$, $\gcd(A) = 1$, $0 < \overline{d}(A) < \frac{1}{2}$ 和 $\overline{d}(2A) = \frac{3}{2}\overline{d}(A)$, 则下列两种情形之一必成立.

(1) 存在 $g > 4$ 和 $a \in [1, g-1]$ 使得 $2a \neq g$, $A \subseteq \{0, a\} + \{gn : n \in \mathbb{N}\}$, 并且 $\overline{d}(A) = \frac{2}{g}$.

(2) 存在 $0 \leqslant c_n \leqslant b_n \leqslant h_n$ 使得对所有 n, $A \cap [c_n, b_n] = \emptyset$ 并且

(i) $\lim_{n\to\infty} \frac{A(0,h_n)}{h_n+1} = \overline{d}(A)$;

(ii) $\lim_{n\to\infty} \frac{c_n}{h_n} = 0$;

(iii) $\lim_{n\to\infty} \frac{A(b_n,h_n)}{h_n-b_n+1} = 1$.

在以上定理中, 条件 $0 \in A$ 的作用只是为了使定理的叙述更方便一些.

例 6.32 设 $A := \{0, 1\} + \{10n : n \in \mathbb{N}\}$, 则 $\overline{d}(A) = \frac{1}{5}$ 和 $\overline{d}(2A) = \frac{3}{2}\overline{d}(A)$ 成立.

设 $A := \{0\} \cup \bigcup_{n\in\mathbb{N}} [3 \cdot 2^{2^n}, 4 \cdot 2^{2^n}]$, 则 $\overline{d}(A) = \frac{1}{4}$ 和 $\overline{d}(2A) = \frac{3}{2}\overline{d}(A)$ 成立.

设 $A := \{0\} \cup \bigcup_{n\in\mathbb{N}} ([2^{2^n}, 2^{2^n} + n] \cup [2 \cdot 2^{2^n}, 4 \cdot 2^{2^n}])$, 则 $\overline{d}(A) = \frac{1}{2}$, $\overline{d}(2A) = \frac{3}{2}\overline{d}(A)$, 但是 A 不满足定理 6.31 中所叙述的结构. 所以, 在定理 6.31 中的条件 $\overline{d}(A) < \frac{1}{2}$ 是必要的.

下一个定理是关于上 Banach 密度的 Freiman 逆问题的结果, 其证明可在文献 [53] 中找到.

定理 6.33 设 $A, B \subseteq \mathbb{N}$ 满足 $\overline{\mathrm{BD}}(A) = \alpha > 0$, $\overline{\mathrm{BD}}(B) = \beta > 0$, $\overline{\mathrm{BD}}(A + B) < \alpha + \beta$, 则存在正整数 g 和集合 $G \subseteq [0, g-1]$ 使得

(1) $\overline{\mathrm{BD}}(A + B) \geqslant \alpha + \beta - \frac{1}{g}$;

(2) $A + B \subseteq G + \{gn : n \in \mathbb{N}\}$;

(3) 对于 $i = 1, 2$, 如果 $\{[a_n^{(i)}, b_n^{(i)}] : n \in \mathbb{N}\}$ 是两个正整数区间序列满足

$$
\begin{gathered}
\lim_{n\to\infty} (b_n^{(1)} - a_n^{(1)}) = \infty, \quad \lim_{n\to\infty} (b_n^{(2)} - a_n^{(2)}) = \infty, \\
\lim_{n\to\infty} \frac{A(a_n^{(1)}, b_n^{(1)})}{b_n^{(1)} - a_n^{(1)} + 1} = \alpha, \quad \lim_{n\to\infty} \frac{B(a_n^{(2)}, b_n^{(2)})}{b_n^{(2)} - a_n^{(2)} + 1} = \beta, \\
0 < \liminf_{n\to\infty} \frac{b_n^{(1)} - a_n^{(1)}}{b_n^{(2)} - a_n^{(2)}} \leqslant \limsup_{n\to\infty} \frac{b_n^{(1)} - a_n^{(1)}}{b_n^{(2)} - a_n^{(2)}} < \infty,
\end{gathered}
$$

则对所有的 $n \in \mathbb{N}$ 和 $i = 1, 2$, 存在整数区间 $[c_n^{(i)}, d_n^{(i)}] \subseteq [a_n^{(i)}, b_n^{(i)}]$ 使得

$$
\begin{aligned}
&\lim_{n\to\infty} \frac{d_n^{(1)} - c_n^{(1)}}{b_n^{(1)} - a_n^{(1)}} = 1, \quad \lim_{n\to\infty} \frac{d_n^{(2)} - c_n^{(2)}}{b_n^{(2)} - a_n^{(2)}} = 1, \\
&(A + B) \cap [c_n^{(1)} + c_n^{(2)}, d_n^{(1)} + d_n^{(2)}] = (G + \{gn : n \in \mathbb{N}\}) \cap [c_n^{(1)} + c_n^{(2)}, d_n^{(1)} + d_n^{(2)}].
\end{aligned}
$$

读者可在文献 [53] 中找到一些例子说明定理 6.33 中条件的必要性. 定理 6.33 可看作是定理 6.29 的平行定理. 事实上, 其证明的基本思想来源于一个称为 “买一送一” 现象, 即每证明一个关于下渐近密度或 Shnirel'man 密度的定理, 就一定会有一个关于上 Banach 密度的平行定理, 参见文献 [54].

以上介绍的作者所获得的结果都是关于密度的. 下面将介绍一个不是关于密度的结果.

Freiman 小双倍常数定理在 $|2A| \leqslant 3|A| - 3$ 时刻画了集合 A 的算术结构. 在文献 [49] 中, 定理 6.27(2) 的证明相当长, 在文献 [42] 中, 定理 6.27(2) 甚至没被包含进去. 其实在文献 [49] 中, 还介绍了一个定理是关于在 $|2A| = 3|A| - 2$ 的条件下刻画 A 的算术结构. 但由于篇幅限制没给证明. 可惜这定理有缺陷, 反例可参见文献 [51]. 定理 6.27 还有一些推广, 例如, 在文献 [55, 56] 中, 定理 6.27(1) 被推广至两个不同集合的和, 在文献 [57] 中, 定理 6.27(2) 被推广至集合 A 满足条件 $|A \pm A| = 3|A| - 3$. 但是, 在文献 [51] 之前没有出现任何当 $|2A| > 3|A| - 3$ 时关于集合 A 的算术结构的结果.

以下定理的证明可在文献 [51] 中找到.

定理 6.34 存在标准正实数 ϵ 和 K 使得对任意有限集合 $A \subseteq \mathbb{N}$, 如果 $|A| > K$ 和 $3|A|-3 \leqslant |2A| = 3|A| - 3 + b \leqslant 3|A| + \epsilon|A|$, 则 A 是一个长度不超过 $2|A| - 1 + 2b$ 的等差数列的子集或 A 是一个总长度不超过 $|A| + b$ 的双等差数列的子集.

在定理 6.34 中, 等差数列或双等差数列的长度上界是最优的, 以下是例子.

例 6.35 设 $k \in \mathbb{N}$ 足够大.

(1) 令 $A := [0, k-3] \cup \{k+10, 2k+20\}$, 则 $|A| = k$, $|2A| = 3k - 3 + 11$, A 不是一个有合理长度的双等差数列的子集, 并且包含 A 的最短等差数列的长度为 $2k - 1 + 2 \times 11$.

(2) 令 $A := [0, k-3] \cup \{3k, 3k+12\}$, 则 $|A| = k$, $|2A| = 3k - 3 + 11$, A 不是一个有合理长度的等差数列的子集, 并且包含 A 的最短双等差数列的长度为 $k + 11$.

虽然定理 6.34 是一个有限组合问题, 但是, 我们可以把它提升到非标准扩张里进行讨论, 从而获得一个分析上的突破口. 定理证明的主要思想如下:

假设定理不成立. 利用转换原理, 我们可以获得一个定理的超有限集反例, 即存在一个超有限集合 A 使得 $|2A|/3|A| \approx 1$, 但 A 不满足定理的结论. 可以假设 $\{0, H\} \subseteq A \subseteq [0, H]$ 并且 $\mathrm{st}(|A|/H)$ 达到所有超有限反例中的最大值. 接下去的任务是展示 A 一定有定理结论中所叙述的结构. 令 U_H 为在 (6.1) 中定义的加法分割. 因为 U_H 和 $\mathbb{N}$ 有很多共同点, 所以, 我们可以利用与定理 6.30 类似的定理来建立 $A' = A \cap U_H$ 的算术结构. 如果 A' 的双倍常数很小, 则 A' 就有所需的算术结构, 然后利用溢出原理, A' 上的结构可以延伸到整个 A 上去. 如果 A' 的双倍常数较大, 则存在一个 $K \in [0, H]$ 使得 $A \cap [0, K]$ 的双倍常数较大, 而这能导致 $A \cap [K, H]$ 的双倍常数较小, 从而推出 $A \cap [K, H]$ 有所需算术结构. 接着就容易证明集合 A 有所需的算术结构, 从而推出矛盾.

从以上讨论中可看出, 在定理 6.34 的证明中, 加法分割 U_H 提供了一个两分法并使得定理 6.30 的思想可在 U_H 上得到应用. 这个方法在标准分析中没有.

如果有读者认为非标准分析方法对标准数学意义不大, 认为所有用非标准分析方法证明了的定理都可用标准方法重新证明, 那么就可以来试一试不用非标准分析而证明定理 6.34. 当然所谓不用非标准分析的证明应该采用本质上不同的思路, 长度上也不能比非标准分析证明长太多.

从理论上讲, 每个非标准分析步骤一定都可以人为地转换成一些不涉及非标准扩张的步骤. 但这样的转换往往会导致证明过程不必要地长, 且会失去对问题直观上的理解.

致谢 感谢审稿人的细心审阅, 他们指出本文原稿中的很多文字错误和打印错误, 极大增强了本文的可读性.

参考文献

1 Goldblatt R. Lectures on the Hyperreals—An Introduction to Nonstandard Analysis. New York: Springer, 1998
2 Henson C W. Foundations of Nonstandard Analysis: A Gentle Introduction to Nonstandard Extension. Nonstandard Analysis: Theory and Applications. New York: Kluwer Academic Publishers, 1997
3 Keisler H J. Elementary Calculus. Boston: Prindle, Weber & Schmidt, 1976
4 Lindstrom T. An Invitation to Nonstandard Analysis, Nonstandard Analysis and its Application. Cambridge: Cambridge University Press, 1988
5 Loeb P, Wolff M. Nonstandard Analysis for the Working Mathematician. 2nd ed. New York: Springer, 2012
6 Robinson A. Nonstandard Analysis. Amsterdam: North-Holland, 1966
7 Robinson A, Zakon E. A set-theoretical characterization of enlargements. In: Applications of Model Theory to Algebra, Analysis, and Probability. New York: Holt, Rinehart, and Winston, 1969, 109–122
8 Luxemburg W A J. A general theory of monads: Applications of model theory to algebra, analysis, and probability. J Econom Soc Hist Orient, 1969, 56: 119–121

9 Loeb P. Conversion from nonstandard to standard measure spaces and applications in probability theory. Trans Amer Math Soc, 1975, 21: 113–122
10 Enderton H. A Mathematical Introduction to Logic. 2nd ed. Amsterdam: Academic Press, 2001
11 Jech T. Set Theory. 3rd ed. New York: Springer, 2002
12 Anderson R M. A nonstandard representation for Brownian motion and Itô integration. Israel J Math, 1976, 25: 15–46
13 Cutland N J, Henson C W, Arkeryd L. Nonstandard Analysis: Theory and Applications. New York: Kluwer Academic Publishers, 1997
14 Keisler H J. An Infinitesimal Approach to Stochastic Analysis. Providence: Amer Math Soc, 1984
15 Fajardo S, Keisler H J. Existence theorems in probability theory. Adv Math, 1996, 120: 191–257
16 Hoover D, Keisler H J. Adapted probability distributions. Trans Amer Math Soc, 1984, 286: 159–201
17 Hoover D N, Perkins E. Nonstandard Constructions of the Stochastic Integral and Applications to Stochastic Differential Equations, I; II. Trans Amer Math Soc, 1983, 275: 1–36; 1983, 275: 37–58
18 Barlow M T. One dimensional stochastic differential equations with no strong solutions. Proc Lond Math Soc (3), 1982, 26: 335–347
19 Gleason A M. Groups without small subgroups. Ann Math, 1952, 56: 193–212
20 Montgomery D, Zippin L. Small subgroups of finite dimensional groups. Ann Math, 1952, 56: 213–241
21 Jacoby R. Some theorems on the structure of locally compact local groups. Ann Math, 1957, 66: 36–69
22 Plaut C. Associativity and the Local Version of Hilbert's Fifth Problem, Notes. Knoxiville: University of Tennessee, 1993
23 Hirschfeld J. The nonstandard treatment of Hilbert's Fifth problem. Trans Amer Math Soc, 1990, 321: 379–400
24 Goldbring I. Hilbert's fifth problem for local groups. Ann Math, 2010, 172: 1269–1314
25 Sun Y. The exact law of large numbers via Fubini extension and characterization of insurable risks. J Economic Theory, 2006, 126: 31–69
26 Duffie D, Sun Y N. The exact law of large numbers for independent random matching. J Econom Theory, 2012, 147: 1105–1139
27 Khan M A, Sun Y N. Non-cooperative games on hyperfinite Loeb spaces. J Math Econom, 1999, 31: 455–492
28 Sun Y. A theory of hyperfinite processes: The complete removal of individual uncertainty via exact LLN. J Math Econom, 1998, 29: 419–503
29 Sun Y. The complete removal of individual uncertainty: Multiple optimal choices and random exchange economies. Econom Theory, 1999, 14: 507–544
30 Furstenberg H. Recurrence in Ergodic Theory and Combinatorial Number Theory. Princeton: Princeton University Press, 1981
31 Nathanson M B. Additive Number Theory—The Classical Bases. New York: Springer, 1996
32 Jin R. Sumset phenomenon. Proc Amer Math Soc, 2002, 130: 855–861
33 Keisler H J, Leth S. Meager sets on the hyperfinite time line. J Symbolic Logic, 1991, 56: 71–102
34 Li B, Zhang J. On the Dedekind completion of $^*\mathbb{R}$. Systems Sci Math Sci, 1988, 1: 29–39
35 Di Nasso M, Goldbring I, Jin R, et al. High density piecewise syndeticity of sumsets. Adv Math, 2015, 278: 1–33
36 Beiglböck M, Bergelson V, Fish A. Sumset phenomenon in countable amenable groups. Adv Math, 2010, 223: 416–432
37 Bergelson V, Furstenberg H, Weiss B. Piecewise-Bohr sets of integers and combinatorial number theory. In: Topics in Discrete Mathematics. Berlin: Springer, 2006, 13–37
38 Beiglböck M. An ultrafilter approach to Jin's theorem. Israel J Math, 2011, 185: 369–374
39 Di Nasso M. An elementary proof of Jin's theorem with a bound. Electron J Combin, 2014, 21: 1–7
40 Halberstam H, Roth K F. Sequences. Oxford: Oxford University Press, 1966
41 Plünnecke H. Eine zahlentheoretische Anwendung der Graphentheorie. J Reine Angew Math, 1970, 234: 171–183
42 Nathanson M B. Additive Number Theory—Inverse Problems and the Geometry of Sumsets. New York: Springer, 1996
43 Estermann T. On Goldbach's problem: Proof that almost all even positive integers are sums of two primes. Proc Lond Math Soc (3), 1938, 44: 307–314
44 Davenport H. On Waring's problem for cubes. Acta Math, 1939, 71: 123–143
45 Jin R. Plünnecke's Theorem for other densities. Trans Amer Math Soc, 2011, 363: 5059–5070
46 Jin R. Density Versions of Plünnecke Inequality: Epsilon-Delta Approach. Combinatorial and Additive Number Theory. New York: Springer, 2014
47 Ruzsa I Z. Sumsets and structure, combinatorial number theory and additive group theory. In: Advanced Courses in Mathematics, CRM Barcelona. Basel: Birkhäuser, 2009, 87–210

48 Bilu Y. Structure of sets with small sumset. Astérisque, 1999, 258: 77–108
49 Freiman G A. Foundations of a Structural Theory of Set Addition. Translated from the Russian. Translations of Mathematical Monographs, vol 37. Providence: Amer Math Soc, 1973
50 Gowers T. A new proof of Szemerédi's theorem. Geom Funct Anal, 2001, 11: 465–588
51 Jin R. Freiman's inverse problem with small doubling property. Adv Math, 2007, 216: 711–752
52 Jin R. Solution to the inverse problem for upper asymptotic density. J Reine Angew Math, 2006, 59: 121–166
53 Jin R. Nonstandard methods for upper Banach density problems. J Number Theory, 2001, 91: 20–38
54 Jin R. Standardizing nonstandard methods for upper Banach density problems. DIMACS Ser Discrete Math Theoret Comput Sci, 2004, 64: 109–124
55 Lev V, Smeliansky P Y. On addition of two distinct sets of integers. Acta Arith, 1995, 70: 85–91
56 Stanchescu Y V. On addition of two distinct sets of integers. Acta Arith, 1996, 75: 191–194
57 Hamidoune Y O, Plagne A. A generalization of Freiman's $3k-3$ theorem. Acta Arith, 2002, 103: 147–156

Nonstandard analysis and its applications

JIN RenLin

Abstract In this paper, we briefly introduce the foundation of nonstandard analysis and its applications to other fields of mathematics, especially to combinatorial number theory. The foundation part includes the basic knowledge of mathematical logic, the construction of nonstandard structures, and some principles and properties commonly used in nonstandard analysis. The application part includes theorems on the existence of strong solutions for some stochastic differential equations, solution to the Hilbert's fifth problem for local groups, exact law of large numbers and its applications to mathematical economics, and sumset phenomenon, Plünnecke type inequalities for densities, and Freiman's inverse problems in combinatorial number theory.

Keywords nonstandard analysis, mathematical logic, stochastic differential equation, Hilbert's fifth problem, exact law of large numbers, combinatorial number theory

MSC(2010) 03H05, 03H15, 11B05, 11P70, 22E05, 60F15, 91B30, 11U10, 11B13, 22A99, 28A35, 28E05

doi: 10.1360/N012015-00266