数学分析 之 课程讲义

（丘成桐数学英才班试用）

起箸雍阉茂尽上章困敦凡二年

清华大学数学系 及 丘成桐数学科学中心

2020 年 12 月 27 日

前言之二

较之二零年六月的版本，讲义增加了数学分析第三学期的讲义[1]，主要研究 Schwartz 的分布理论以及关于缓增分布的 Fouirer 变换。分布理论是一门重要的分析语言，尽管这套对函数的描述方式是上个世纪 40 年代才被引入的，但是掌握分布理论所需要的基础知识却并不多：线性代数和多元微积分就足够了！本课程所研究的具体问题，大多来自于物理学，譬如引力理论、电磁学、热传导理论等，在解决这些问题过程中，貌似抽象的分布语言只是提供了叙述上的便捷，实实在在的仍然是微积分各种技术和想法。这是我们选择数学分析第三学期的内容的初衷：借此学习前两个学期中漏掉的内容与技巧并巩固那些贯穿始终的思想。

课程所用到的主要参考书是我之前在 École Polytechnique 读书时所用的教材：Théories des distributions（由 G. Lebeau 和 J.-Y.Chemin 编写）以及 J.-M. Bony 教授所撰写的课本

- Cours d'analyse: Théorie des distributions et analyse de Fourier.

以 Laurent Schwartz 命名的数学中心占据了 École Polytechnique 建筑群中不起眼的一角，我还依稀记得一次 Bony 教授在那里叼着烟斗提起这些教学的材料从 Schwartz 本人开始辗转经过多位数学家教授最终成为 Polytechnique 分析的标准教材。我们在课上还经常借鉴另一本法语的教材：

- Éléments de distributions et d'équations aux dérivées partielles - Cours et problèmes résolus. Claude Zuily

比如对 Weyl 的特征值渐近公式的证明就是采取这本书的处理方法[2]。另外，这本教材有一系列精心挑选的习题，这对于讲课或者听课的人都是难得的财富。

学期结束，从教学效果上来说课程还是有很大欠缺，特别是最后两周的课程内容可能略显多余（介绍了波前集的概念，并对于线性变系数微分算子证明了微局部的椭圆正则性定理和微局部正则性的传播定理）。如果可以再次教授这门课程，我想更好的选择可能是更为经典的话题，比如说继续介绍波动方程有关的理论，或者研究调和函数作为课上对 Dirichlet 问题一部分内容的补充，或者利用 Littlewood-Paley 理论来刻画已有的几种函数空间。当然，这都是将来的事情了。

2020 年 12 月于静斋

[1] 也增加了更多的打印错误

[2] 我曾在 Patrick Gérard 教授的指导下以特征值的 Weyl 定律为主题做过一次课后的深入学习 (étude d'approfondissement)，按照 C. Zuily 在书中的讲法，他也受到了 Gérard 教授的影响。

前言之一

清华大学目前数学分析课程共分三个学期，每学期分三堂课分别由三位老师讲授。这份讲义是其中一堂课的关于数学分析一和二课堂笔记，该课程是丘成桐数学英才班的实验课，与传统数学分析教学相比，我们尝试在不动根本的前提下在内容与设计上探索新的可能。目前已公认数学分析为数学系学生最重要的基础课之一，各类书目文献汗牛充栋，其中不乏有颇具口碑同时又经得住时间考验的优秀教材。这些教材在内容安排上大都以下面的内容为范本：数列极限，函数连续性，一元函数导数，一元函数 Riemann 积分，多元函数连续性，多元函数微分，多元函数 Riemann 积分理论。在这样的构架上，不同教材的作者会搭建各具特色的补充，比如有的在函数列级数的理论上详细展开，也有的在含参数积分上进行引申，有的更注重概念的进阶于是增加了拓扑空间的概念或者干脆引入 Lebesgue 积分。不同的作者有不同的视角，我们的课程也尽力表达个人的见解。正如对积分论作出了不可磨灭贡献的法国数学家 Henri-Léon Lebesgue 所说的:[3]对数学对象有多少种不同的理解就有多少种不同的教授方式。

课程内容

在课程知识点的选择上，有如下四个内容需要略加解释：

1) 距离空间与赋范线性空间。

   我们选择在课程伊始便引入距离空间与赋范线性空间。收敛是贯穿分析始终的概念，而大多数的收敛可以统一地用距离空间的语言叙述。尽早引入这些概念使得课程后续的诸多命题陈述变得自然明了。实际教学也表明讲授实数收敛理论的难度与距离空间收敛理论的难度无论对教师还是学生是无差别的，同学们对此接受程度很高。之后到多元函数的连续性、函数列一致收敛乃至 L^p-空间上收敛性的过渡变得水到渠成。

2) Riemann 积分的定义。

   我们采取用简单函数逼近的方式来定义一元函数的 Riemann 积分。作为推论，传统的 Riemann 和与 Darboux 上下和的定义方式可以用这种语言清楚地表达。这种积分的处理方式应用范围更广，它容许对在完备赋范线性空间中取值的映射定义积分。这也是对引入抽象测度空间上的积分做铺垫。

3) 子流形理论。

   隐函数定理的应用见著于数学的各个分支，其重要程度无论如何强调都不过分，然而它向来都是多元微分学中较难讲解和理解的课题，甚至即使只要求“背过”定理的陈述也不是轻而易举的事情（至少笔者记不住）。然而，几何的陈述通常带来直观的视觉记忆：隐函数定理说如果一族函数满足非退化条件，那么它们的公共零点集是 $\mathbb{R}^n$ 的子流形。当然，引入子流形概念不仅仅是方便形象地陈述和记忆隐函数定理，在学习曲线与曲面的积分理论时，子流形的

[3]Il y a plusieurs manières de comprendre les mathématiques donc il y a plusieurs manières de les enseigner.

概念完全无法避免。尤为重要的是，在子流形上也可以做微分，这种在弯曲空间上的微分学可以更清晰地揭示微分 df 的含义，众多经典的理论例如 Lagrange 乘子法也变得一目了然。子流形理论是 $\mathbb{R}^n$ 上微分学最根本最有启发性的内容，低年级数学系同学通过这里的学习也能认识到引入正确概念是近代数学的重要秉性。

4) 抽象测度理论

传统上，多变量函数积分的学习把一维 Riemann 积分理论作为原型，通过把区间替换为长方体，高维的 Riemann 和就可以用来定义 Riemann 积分。然而，我们用简单函数逼近的办法来定义一维 Riemann 积分，类比到高维情况，如果 $A \subset \mathbb{R}^n$ 的面积（测度）有定义，就可以对其示性函数 $\mathbf{1}_A$ 定义积分。所以，一旦明确了哪些子集可以定义面积，我们就能对它们的示性函数的线性组合（即简单函数）进行积分，然后就对一切可以被简单函数逼近的函数进行积分（在第一学期我们特意讲授了 Stieltjes 积分作为这个想法的另一明证）。

从计算的角度我们也可以做一番事后诸葛亮的讨论。为了计算积分，我们需要（且只需要）两个基本工具：第一，Fubini 定理，它把高维积分（乘积空间上）转化为低维的来计算；第二，换元积分公式，它把不规则区域上的积分转化为乘积型区域上的积分，从而可运用 Fubini 定理。不夸张地说，相当一部分的积分计算只是运用这两个分析工具的代数运算。只要在抽象积分的框架里建立这两个公式，抽象意义下定义的积分与 Riemann 积分在计算上完全一致。

抽象积分理论优势更明显：它囊括大部分可能的积分与求和，比如说级数的求和与概率空间上的积分等；它拥有 Lebesgue 控制收敛等有效工具，所以在讨论级数的收敛、积分与求导数可交换或者 Fourier 级数时非常地方便。还有一点常常被忽略：在技术层次上，抽象积分要比 Riemann 积分更容易学习。Riemann 积分伟大之处在于能够对由长方体构成的 Jordan 代数／可测集进行刻画，然而其技术细节比抽象积分要困难，因为在抽象的场合人们不需要关心集合具体的几何从而更自由。

另外，不少同学和同事认为这部分内容与实分析课程过度重合，我们认为在抽象层次上讨论积分实际上可以尽量避免使用真正的实变技术，这对实分析影响甚微。

课程作业与习题

作业是课堂所教授内容的补充与支撑，与课堂内容同等重要。在题目选择方面，首要考虑的它们在数学和科学中的重要性，其次是它们与其他数学方向的关联，再其次是问题背后想法和机制的优美程度以及困难性。每学期我们有十二次作业，每次六十道以上的习题，平均对每位同学至少是六到八小时的工作量。我们不提倡偏题难题，不为解题而解题。大家也都顽固地坚持老派的信条，坚信数学至少相当的一部分内容是技术性的，必须经过反复练习才能领会贯通。所以，这门课程坚持足够的作业量和一定质量的习题，在做习题的过程中同学自然就可以引导思考启发原创。

作业题目通常有三种形式：

1) 课堂内容的补充。

比如说，课上证明连续函数的某个判断准则，习题就要求对在赋范线性空间取值的情形进行证明。这种题目通常是直接推广，适当的练习有利于更好的理解课堂内容和熟悉新的概念与术语。再比如说，某些复杂的计算，课程上跳过的细节可以有相应的题目进行补充。还有某些技术性结果其论证本身并没有启发性（单位分解的构造），恰当设计习题可以由同学验证细节已保证课程的完整。在作业中，我们也鼓励同学查阅教科书、网络或者相互讨论，因为解决这些问题的目标不是为了得到答案而是为了掌握基本的知识，这些手段也更接近于学科研究的前沿。

2) 传统教科书上某些习题。

现有教科书中有很多有意思的问题，我们从中广泛取材，尤其是一些比较技巧的问题，比如说 L'Hôpital 法则（多练无益）或者定积分的计算等，这是基本训练，数学上不应该回避技巧的提升。

3) 经典结果和定理的证明。

这是明显具有法国风格的习题，具有明显的特点：通常将某个著名的结果适当拆分成若干步，问题从而包含很多小问，每一小问题的难度适当。这样的问题也可以来自研究论文。所以，同学们在解答过程中步步为营，逐渐接近目标，最终证明一个很大的结果，非常鼓舞士气。比如说，某一套习题证明所谓的 Borel 引理，这个引理说任意给定数列我们都可以构造光滑函数 f 使得 f 在 0 处的各阶导数恰为此序列，这个引理在拟微分算子理论中很有用处，但是不值得在课堂上大费周章地讲解，用这种方式出题很合适；再比如说，八字班的期末考试题目的背景如下：1922 年，19 岁还在莫斯科大学读本科的 Kolmogorov 在论文 *Une série de Fourier-Lebesgue divergente presque partout* 中构造一个 L^1 函数，这个函数的 Fourier 级数的部分和几乎处处是发散的。期末考试重现这个著名反例的构造，一共有 20 个小问题。

其他

以上无论是课程知识点还是习题的选择，着眼点从来都不是对已有的教材或者教学系统优点的综合或者缺点的躲避。我们更看重这些内容在将来分析学的学习（实分析、复分析、调和分析、偏微分方程以及物理学等）以及研究中的重要和基本程度。从历史渊源来看，数学分析被单独从分析学中剥离出成为独立的课程应该是很近的事情。我们希望能尽量还原历史上那些有启发性的例子，把课程放到相对动态的大背景里，以便同学们能意识到当下的学习与后来的课程、数学的历史和以及现行的研究不是割裂的。在教学上内容的跳跃也是课程的一个显著特（缺）点，我们相信同学能体会到跳跃学习是更应该尝试的和鼓励的，这个和研究更接近。事实上，我们所接触的大多国内优秀同学都直接或者间接地默认数学学习一定顺序，须得按部就班，层层递进，才能扎实地进步。这种想法本身没有错，但是数学的发展绝对不是线性的，知识也不存在先天的顺序。一个普遍被提及的例证就是从 Newton 引入积分（反流数术）到 Cauchy 的极限理论被广泛接受中间至少有半个世纪的光景，那个优秀的数学家层出不穷，利用还没有被严格化的积分理论作出了很多享誉后世的工作。数学学习不应该被课程设置（甚至被其名称）切割。比如说，所谓的曲线或者曲面的

几何学、数学物理方程都可以称为（多元）微积分学习的好例题，很多微积分知识就是为了解决这样的问题应运而生的。微积分课程就应该学习和研究这样的例子，不能简单地将它们归结为是微分几何或者微分方程的内容而忽略。仅仅为了教学而刻意构造的各种例题长远来看有目光短浅之嫌。这种做法通常还会从心理上折射出胆怯的影子，以至于有同学误以为课程的顺序和难度有关联，“后面”的课程比“前面”的课程困难。唯有正视，方能前行。我们可以参阅一些相对古老的经典教科书／讲义，这些教科书大多数历史上真正有名望并作出过不朽工作的数学家写的：

- André-Marie Ampère 和 G. Vincens 在 École Polytechnique 的分析学与力学讲义，1823-1824 年左右出版，从函数的微分学开始定义，微分方程和曲面几何一应俱全，从书（手写）的长度推测，这应该是那时一学期的课堂笔记。

- Charles Hermite 在 École Polytechnique 的分析学讲义的第一卷开篇不久就仔细地研究了曲面的几何，反而 Riemann 积分没有太多严格讨论，之后花了大量的篇幅进行计算并应用于几何的研究。

- Camille Jordan 的分析学教程是 1893 年发表的，在当时欧洲很有影响力，很多现代教材大都直接或者间接的从这份讲义上取过经。翻开目录，我们就会看到关于平面代数曲线，复分析，椭圆函数，偏微分方程等内容。

- Fritz John 和 Richard Courant 的 Introduction to Calculus and Analysis，这两位是 Courant 研究所代表人物，在偏微分方程和应用数学上很有成就，他们的课本包含了基本的复分析以及微积分在几何和物理上的各种应用，知识之间没有停顿，每个例题都值得玩味。

近年来较流行的数学分析教材，菲赫金哥尔茨或者卓里奇这些俄国作者仍然注重应用。其余如 Rudin 或者国内的种种著名的课本，在细节和清晰程度上下足了功夫，可读性很强，数学分析课程作为一门基础课的诞生与成长与这些优秀的课本密不可分。然而，很多现代数学课本越来越讲究深入浅出，对读者感受的呵护有时凌驾于数学，殊不知读书与写文章往往穷而后工，走过弯路才更能领会定理的要义。所以，在课堂上我们只建议学生选择自己觉得合适的参考书而已，我们能坚持的是以课堂为主，教材只是辅助的，强调每个人都要记笔记，这是建立学术自信而不盲从任何参考书（包括这份讲义）的正路。然而，得失之间其实很难说清楚，特别当把学习在更宏大的尺度下衡量的时候。无论怎样，可能只有上过课程的同学才能议论这一年的课程是否偏离了它的初衷：在宏观上把握分析整体图像，理解几个提纲挈领的想法；在微观上获得基本的计算能力，了解分析中为数不多的核心例子以及它们的历史渊源。

2020 年 6 月于静斋

课程所使用的参考书目

我们在讲义和习题的编写中参考（≈ 抄）了很多文献，很多例子和掌故也是道听途说，来自师长、同学或者同事。以下是主要参考书目，部分课题和习题选自这些教材：

- Richard Courant; Fritz John, *Introduction to calculus and analysis. Vol. I.* Reprint of the 1965 edition. Springer-Verlag, New York, 1989. xxiv+661 pp.
- Richard Courant; Fritz John, *Introduction to calculus and analysis. Vol. II.* Reprint of the 1974 edition. Springer-Verlag, New York, 1989. xxvi+954 pp.
- Edmond Ramis; Claude Deschamps; Jacques Odoux, *Cours de mathématiques spéciales. 3. Topologie et éléments d'analyse.* Masson, Paris, 1982. viii+362 pp.

  法国预科学校的经典教材，我们 Riemann 积分的处理完全跟随这里的节奏。
- Edmond Ramis; Claude Deschamps; Jacques Odoux, *Cours de mathématiques spéciales. 4. Séries et équations différentielles.* Masson, Paris, 1988. viii+328 pp.

  关于高维 Riemann 积分的习题课材料来自最后一部分。
- Walter Rudin, *Principles of Mathematical Analysis.* McGraw-Hill Book Company, Inc., New York-Toronto-London, 1953. ix+227 pp.

  关于 Stieltjes 积分部分的讲义来自这里。
- Vladimir A. Zorich, *Mathematical Analysis. I.* Translated from the 2002 fourth Russian edition by Roger Cooke. Universitext. Springer-Verlag, Berlin, 2004. xviii+574 pp.
- Vladimir A. Zorich, *Mathematical Analysis. II.* Translated from the 2002 fourth Russian edition by Roger Cooke. Universitext. Springer-Verlag, Berlin, 2004. xvi+681 pp.
- 陈天权，数学分析讲义，第一册，北京大学出版社，2009 年 8 月.
- 陈天权，数学分析讲义，第二册，北京大学出版社，2010 年 3 月.
- 常庚哲，史济怀，数学分析教程，上册，高等教育出版社，2003 年 5 月.

声明

- 这是清华大学丘成桐数学英才班教学改革与尝试的一小步。
- 感谢清华大学数学系八字班参与本课程的同学，他们是第一批试验品。
- 感谢清华大学数学系九字班参与本课程的同学，讲义前半部分可读性较高，这要归功于同学们课前课后指出了大量的错误。
- 感谢李梦妮、刘明昶、林汛、林天润、薛宇皓诸位博士生助教，他们批改了几万页的作业。
- 感谢荆文甲和罗天文两位老师，他们为这门课程花了大量心血。
- 这不是课本，内容与课堂黑板笔记完全同步（字体远比板书漂亮）。
- 我们对讲义中数学的原创性没有贡献。
- 讲义包含很多笔误，请慎重使用。讲义的后半部分笔误明显增多，一方面是笔者的粗心与懒惰之过，一方面因为同学们在系统地学习后声称也学会了如何自动纠正笔误（所以没有必要改了），我们姑且信之。
- 如果在阅读之后能多多找出笔误的修改意见，我们在清华备有拿铁和茗茶聊表心意。

但卑意欲少年为学者，每一书皆作数过尽之。书富如入海，百货皆有，人之精力，不能兼收尽取，但得其所欲求者尔。故愿学者每次作一意求之。如欲求古今兴亡治乱、圣贤作用、但作此意求之，勿生余念。又别作一次，求事迹故实典章文物之类，亦如之。他皆仿此。此虽迂钝，而他日学成，八面受敌，与涉猎者不可同日而语也。

——— 苏 轼，又答王庠书

目录

1 实数的公理化描述 **19**

2 区间套公理与确界原理，距离空间 **25**

3 实数的构造：Dedekind 分割 **30**

3.1 作业：可数与不可数，Schroeder-Bernstein 定理 36

4 极限，级数，Cauchy 列，距离空间中的收敛 **40**

5 Cauchy 判别准则，向量序列的收敛，Bolzano-Weierstrass 的列紧性定理，数列和级数收敛的 Cauchy 判断，e 的定义，绝对收敛，收敛的控制判别法 **48**

6 指数函数与三角函数的构造，指数函数与三角函数的代数性质，双指标级数的求和 **59**

6.1 作业：Riemann 重排，Cesàro 求和，Banach-Mazur 游戏 65

7 乘积级数与 Riemann ζ-函数，振荡级数的收敛判断（Dirichlet 与 Abel 判别法），完备距离空间，完备赋范线性空间，Picard 不动点定理／压缩映像定理，等价距离／范数，矩阵的指数映射 **71**

7.1 作业：素数的倒数和，Basel 问题的 Euler“证明” 81

8 函数的连续性 **86**

9 距离空间之间的连续映射，介值定理，初等函数的构造 **92**

10 开集，闭集，紧集，闭包，聚点，连续函数的拓扑刻画 **102**

10.1 作业 107

11 紧性与开覆盖，Heine-Borel 定理，Lebesgue 数，一致连续，函数列的逐点收敛与一致收敛，闭区间上连续函数空间 $\big(C([a,b]),\|\cdot\|_\infty\big)$ **111**

12 利用级数收敛来构造连续函数，距离空间的完备化 **119**

12.1 作业：有无穷多素数的拓扑证明 126

12.2 期中考试：连续函数环的极大理想 132

13 导数的定义，初等函数导数的计算 **136**

14 Leibniz 公式，Faà di Bruno 公式，导数的基本性质，在赋范线性空间中取值的函数的导数，求极值，Rolle 与 Lagrange 中值定理，处处连续处处不可微函数的构造 **141**

15 $x \mapsto \exp(xA)$ 的导数，中值定理的应用：线性常微分方程，Darboux 介值定理，Cauchy 中值定理，三角函数与微分方程，π 的定义 151
15.1 作业：高木贞治函数 158

16 空间填充曲线，L'Hôpital 法则，Taylor 展开 166

17 凸函数，利用 Jensen 不等式证明常见的不等式 175
17.1 作业：Émile Borel 引理，Peano 的证明 183

18 Riemann 积分的定义：区间的分划，简单函数，Riemann 可积函数 189

19 Riemann 和，Darboux 上下和，上下积分 199
19.1 作业：Sturm-Louville 理论的一个例子 204

20 紧区间上 Riemann 可积函数的刻画，Newton-Leibniz 公式，分部积分公式，变量替换公式，原函数的计算，圆面积的计算：证明 π 的几何意义 210
20.1 作业：Dini 定理，多项式逼近与 Weierstrass-Stone 定理 219

21 振幅，零测集，Lebesgue 定理 224

22 积分的基本性质，积分余项的 Taylor 公式，反常积分，反常积分的收敛性：控制收敛，面积法，Euler 常数，Wallis 积分与 Stirling 公式 229

23 历史注记：Newton 和 Leibniz 时代的微积分，Leibniz 对 $\frac{\pi}{4}$ 的级数表示，微分与积分的交换定理，利用含参积分的计算积分 239
23.1 作业：$\zeta(2)$ 的无理性 246

24 常微分方程解的存在唯一性，Kepler 三大定律的证明，变分法：两点之间线段最短 250

25 变分法：最速降线问题，Huygens 等时定理，积分第一中值定理 262
25.1 作业：可写成两个完全平方数的和的整数的密度 267

26 第二积分中值定理，Stieltjes 积分 275

27 Stieltjes 积分的中值定理，利用 Stieltjes 积分的中值定理证明第二积分中值定理 283
27.1 作业：振荡积分 290

28 一元微积分拾遗：Baire 纲定理及其应用，原函数的初等函数表示定理：Liouville 定理 296

29 一元微积分拾遗：振荡与衰减，Riemann-Lebesgue 引理，van der Corput 型估计 306
29.1 建议阅读：第一学期课程没有覆盖的内容 313

29.2 第一学期期末考试：Gauss 的一个椭圆积分的计算，$\int_0^\infty \frac{1}{1+x^\alpha}dx$ 的计算，整系数多项式的 Chudnovsky 逼近 313
29.3 寒假作业 320

30 方向导数，偏导数和微分，利用偏导数连续判断微分存在，导数与极值 335

31 映射的微分，利用偏导数表示微分，Jacobi 矩阵，微分的链式法则，逆映射微分的计算，矩阵代数中指数映射的微分 344
31.1 作业：齐次函数与 Euler 公式 352

32 微分同胚，坐标变换，Clairaut-Schwarz 定理（偏导数可交换），多元函数的 Taylor 展开，$\mathbb{R}^n$ 中光滑子流形的定义 357

33 子流形用坐标表示，子流形作为零点集来实现，反函数定理及证明，反函数定理与坐标变换之间的关系 368
33.1 习题课：拓扑空间 375
33.2 作业：反函数和隐函数定理 378

34 隐函数定理，隐函数定理的子流形叙述，子流形的判定：光滑曲线，光滑曲面超曲面的例子，子流形的参数化表示，Möbius 带的研究 384

35 子流形的原像定理，切空间的定义，超曲面的切空间与法向量 394
35.1 作业：隐函数与反函数定理，隐函数定理在多项式和矩阵上的一个重要应用，经典群的子流形结构 401

36 子流形之间的光滑映射，子流形之间光滑映射的微分，切向量场的定义（$\neq$ 向量值函数），切丛，切丛是光滑子流形，子流形上的微分学与极值问题：Lagrange 乘子法，子流形上的反函数定理和隐函数定理 406

37 多元函数 Hesse 矩阵，极值点的二阶导数判定，凸函数 418
37.1 习题课：球极投影 425
37.2 作业：Lagrange 乘子法，Morse 引理，横截相交性 428

38 σ-代数，由某些子集生成的 σ-代数，Borel-代数，σ-代数的张量积，$\mathbb{R}^2$ 上的 Borel 代数与 $\mathbb{R}^1$ 上的 Borel 代数之间的关系，单调类，单调类 + 代数 $\Rightarrow\sigma$-代数，可测空间，可测映射，σ-代数的拉回，可测性的生成元判据，到乘积空间可测性的判据，距离空间的乘积空间上的 σ-代数，可测函数的性质 433

39 测度，测度空间，σ-有限性，测度的基本性质，Carathéodory 扩张定理：如何利用代数定义外测度导出 σ-代数上的测度 449

40 测度的像（推出），Lebesgue 测度的构造，平移不变性刻画 Lebesgue 测度，Lebesgue 测度在尺度变换下的变换，测度空间的完备化，简单／阶梯函数的定义 457

41 简单函数的积分，测度空间上积分的定义，零测集，几乎处处，Beppo Levi 定理 470

41.1 作业：子流形与零测集，Stieltjies 测度的构造，Borel-Cantelli 定理和无理数的逼近 478

42 Lebesgue 积分与 Riemann 积分之间的关联，可积函数空间，积分的线性，L^1-空间，Fatou 引理，Lebesgue 控制收敛定理，积分对参数的连续依赖性 484

43 积分与求导数的可交换性，乘积测度的构造 496

43.1 作业：Lebesgue 控制收敛，十进制小数的研究 503

43.2 习题课：硬币空间的测度理论 509

44 乘积测度的构造，Fubini 公式的证明和应用，$\mathbb{R}^n$ 上积分的降维计算：Arichimedes 的墓碑 512

44.1 作业：Archimedes 对抛物线面积的计算，Gauss 积分 520

45 抽象换元积分公式，Borel 测度的正则性引理，微分同胚下换元积分公式的证明 524

46 常用的换元积分，行列式的几何含义，子流形上的测度与积分：第一型曲线／曲面积分，函数图像上的积分公式 537

46.1 期中考试：非 Borel 集的构造 547

47 子流形上积分的计算：n-维球的体积，Archimedes 公式. 截断函数的构造，周期的单位分解，有界带边光滑区域，单位外法向量，在函数图像上的计算，Stokes 公式的第一个证明 553

48 Sard 型引理，Stokes 公式的微分拓扑证明 565

48.1 作业：曲面曲线积分的计算 573

48.2 习题课：Riemann 积分的定义 1 578

49 子流形上的积分的计算，第一／二型子流形／曲线／曲面积分，向量场的运算，散度定理，Green 公式，Gauss-Ostrogradsky 公式，散度的几何/物理解释 585

50 二维的 Brouwer 不动点定理的证明，Hilbert 空间，赋范线性空间完备性的级数判定，Fischer-Riesz 定理 595

51 偶数维球面上的切向量场，L^2 与 L^∞ 的完备性，连续线性算子，线性映射的延拓，卷积与函数逼近 602

51.1 作业：Stokes 公式的应用 612

51.2 习题课：Riemann 积分的定义 2 618

52 用光滑函数的逼近 L^1 函数，可分的 Hilbert 空间，Hilbert 基，Fourier 级数的定义，Fourier 级数的 L^2 理论 623

53 Fourier 级数的 L^2-理论，高维 Fourier 级数的 L^2-理论，绝对收敛的 Fourier 级数 632

53.1 作业：波动方程的局部能量估计 640

53.2 习题课：Riemann 积分的定义 3 644

54 物理空间的光滑性与频率空间的衰减，Riemann-Lebesgue 引理，Dirichlet 核，Féjer 核，Féjer 核的收敛理论，局部化引理 647

55 du Bois-Reymond 反例，Fourier 级数收敛的 Dirichlet 定理，Fourier 级数收敛的 Dini 定理，Hölder 连续函数 Fourier 级数收敛理论，有界变差函数的 Jordan 定理 Fourier 级数收敛理论 660

55.1 作业：Fourier 级数的计算，三角函数与球谐函数 667

56 Bernstein 定理，Fourier 级数在等分布问题上的应用 674

57 Roth 三项等差数列定理 684

57.1 作业：Fourier 级数几乎处处发散的 L^1-函数 698

57.2 期末考试：Maass 波函数的展开 704

58 分布的定义与例子：Radon 测度，局部可积函数到分布的嵌入，主值部分 710

59 分布的操作：限制，求导数，与微分同胚复合，链式法则. Stokes 公式的分布形式 717

60 1 维的跳跃公式，Cauchy 积分公式，Cauchy-Riemann 算子的基本解，单位分解 726

61 分布的局部刻画，分布的支集，支集在一点处的分布的结构定理 735

61.1 作业：齐次分布，Hadamard 有限部分，分布除以多项式 742

62 分布的卷积：分布与微分和积分的交换，分布与紧支集分布的卷积，Dirac 分布与平移，卷积与求导数的交换 747

63 卷积的连续性，基本解，Cauchy-Riemann 方程的椭圆正则性，热方程的基本解 756

63.1 作业：分布的例子，Laplace 算子、位势方程与分布 765

64 可卷集，$\mathbb{R}^3$ 上的波动算子的基本解，$\mathbb{R}^3$ 上的波动方程 772

65 选读部分：复分析相关知识简介，极大值原理，Laurent 展开，留数定理（在 Fourier 变换的计算中要用） 781

65.1 Fourier 分析：$L^1(\mathbb{R}^n)$ 上的 Fourier 变换，Riemann-Lebesgue 引理，频率空间与物理空间的基本观点，Gauss 函数的 Fourier 变换，Fourier 变换的逆变换 790

66 L^2 上的 Fourier 变换，Planchrel 公式，Schwartz 空间，Schwartz 函数的例子，试验函数在 Schwartz 空间中的稠密性，Schwartz 空间上的 Fourier 变换，缓增分布的定义798

66.1 作业：Fourier 逆变换的另一个计算，一个分布扩张的问题，分布的张量积 805

67 缓增分布的 Fourier 变换：定义与基本例子的计算 811

68 缓增分布的 Fourier 变换与卷积 819

69 分布理论与 Fourier 变换在数学物理方程上的应用:求波动方程与热方程的基本解. Sobolev 空间的定义，Sobolev 空间的映射性质 824

70 Sobolev 空间的基本性质，Fourier 乘子，Sobolev 空间的稠密子空间，有紧支集分布的结构定理，Sobolev 嵌入定理，高指数的 Sobolev 空间是一个代数 830

70.1 作业：Fourier 变换的计算，Heisenberg 测不准原理，分数次 Sobolev 空间的物理空间刻画，1 维的 Poisson 求和公式 837

71 Riesz 表示定理，Hilbert 空间中的正交投影，Sobolev 空间的对偶性，Sobolev 空间的限制定理（迹定理） 842

72 有界区域上的 Sobolev 空间，1 维区间上的 Sobolev 空间的经典表述，椭圆边值问题，Poincaré 不等式 851

72.1 期中测验 858

73 Dirichlet 问题的初步研究：存在性与解算子，用全空间上的连续函数逼近上半空间上的 H^1 函数，H^1 函数从半空间到全空间的扩张 859

74 限制（迹）定理的连续性，半空间上 Sobolev 空间的限制（迹）定理，限制的正合列 869

75 高指数 Sobolev 空间从半空间到全空间的扩张，Sobolev 函数的局部刻画，从 $H^1(\Omega)$ 到 $H^1(\mathbb{R}^n)$ 扩张，用全空间上的光滑函数逼近的 $H^1(\Omega)$，曲面上的 Sobolev 空间的定义 876

75.1 作业：二维波动方程的基本解，Airy 函数与线性 KdV 方程 885

76 子流形上的 Sobolev 空间的定义与等价定义，有界区域上 Sobolev 空间的限制（迹）定理以及限制正合列，Dirichlet 问题的边值的确切含义以及 Dirichlet 问题的解决，位势方程的边值问题的高阶椭圆正则性定理 889

76.1 习题（利用变分与 Riesz 表示定理解微分方程）：一个弹性力学的模型 899

77 完备内积空间上的紧算子，自伴算子与弱收敛理论 900

78 紧算子的谱理论：**Hilbert-Schmidt** 定理，**Laplace** 算子的谱分解定理，正方形区域上 **Laplace** 算子的特征函数与特征值的计算，正方形区域上上 **Laplace** 算子特征值的增长率**907**

79 **Laplace** 算子的特征函数的应用：利用特征函数刻画函数空间，各阶特征值的变分表述，相互包含区域的特征值比较，弱化版本的 **Weyl** 渐近公式：一般区域上的 **Laplace** 算子的特征值的增长率 **917**

80 $H_0^1(\Omega) \cap H^2(\Omega)$ 的刻画，保证分布落在 $H_0^1(\Omega) \cap H^m(\Omega)$ 中的必要条件，特征函数到边界的连续性（光滑边界），利用特征函数构造热核（频率空间的观点） **924**

81 利用热方程刻画热核（物理空间的观点），热方程的极大值原理，热核的正性与对称性，区域上热核与全空间热核的比较定理 **936**

82 热核的渐近展开，热核在对角线上的积分，**Karamata** 的 **Tauber** 型渐近公式，**Weyl** 渐近公式的证明，波前集的定义 **947**

83 波前集以及其等价定义，微局部光滑性意味着局部光滑性，一个非驻相法的引理（理论的技术核心），波前集在微分同胚下的变换 **956**

84 微局部的椭圆正则性定理，拟解（**paramatrix**）的构造，双特征曲线（**bicharacteristics**），奇性传播定理的叙述 **968**

85 奇性传播定理的证明 **977**

86 分布理论期末复习题 **988**

86.1 第一套 988

86.2 第二套 992

86.3 第三套 994

简介

数学分析一这门课程的内容包含五个部分：

0) [4] 实数理论：

- 实数的公理化表述；
- Dedekind 构造；
- 空间的概念：线性空间与度量／距离空间。

1) 数列与级数的收敛性：

- 实数序列的收敛理论：$\varepsilon - N$ 语言与收敛的基本判定；
- 收敛的例子：级数收敛与具体实数的构造；
- 度量空间中点列的收敛与完备性；

2) 函数与映射的连续性：

- 一元函数的连续性：$\varepsilon - \delta$ 语言与连续性的判定；
- 连续函数的基本性质；
- 连续函数空间与连续函数的构造；
- 度量空间上的连续函数与映射。
- 紧性。

4) 导数与微分：

- 一元函数的导数与可微函数的基本性质；
- 多元函数的偏导数；
- 微分的定义；
- Taylor 展开与其它应用。

3) [5] 积分理论：

- Riemann 积分：几种等价定义与反常积分；
- Newton-Leibniz 公式与积分的计算；

[4]实数理论从知识体系完备性的角度而言是非常重要的，然而对于理解数学分析的基本思想并不是不可或缺的。

[5]积分和微分的理论体系是完全独立的，连结它们的纽带是 Newton-Leibniz 公式。Riemann 积分的定义比微分的定义更为自然，从现代分析学的观点来看，先教授积分学是更自然的做法。

- 积分的性质：中值公式与 Lebesgue 定理等；
- Stieltjes 积分。

除了理论部分，课程的真正核心是围绕着所谓的 Euler 公式展开的：

$$e^{i\pi} = -1.$$

我们课程一个主要的目标是正确的定义 e，π 以及指数函数 e^z 并严格的证明上面的公式。（如果你可以完整地说清楚这些对象是什么并且给出 Euler 公式的严格证明的话，数学分析一就算完全学明白了）

1 实数的公理化描述

二零一九年九月九日，星期一，多云

实数的公理体系

实数理论这一章节的目标是从基本的集合论出发（只假设同学们熟知有理数和自然数），通过**严格**和**完备**的推理，证明若干论题作为基本的工具。这些通过严格论证得来的工具将会支撑着我们直观的图像，使得我们可以形象地思考和解决问题。

$\mathbb{R}$ 是一个集合，我们假设 x，y，z 等都是它的元素。

- $\mathbb{R}$ 上面配有两个操作：
  - 加法. $+:\mathbb{R}\times\mathbb{R}\to\mathbb{R},(x,y)\mapsto x+y$；
  - 乘法. $\cdot:\mathbb{R}\times\mathbb{R}\to\mathbb{R},(x,y)\mapsto x\cdot y$。
- $\mathbb{R}$ 上面还有一个序关系 $\leqslant$：$x\leqslant y$（也记为 $y\geqslant x$）。

我们要求四元组 $(\mathbb{R},+,\cdot,\leqslant)$ 表示满足下面四套公理：

(F) 域公理：$\mathbb{R}$ 是一个域

(F1) 加法结合律：$x+(y+z)=(x+y)+z$。

(F2) 加法交换律：$x+y=y+x$。

(F3) 存在加法单位元：存在 $0\in\mathbb{R}$，使得对任意 $x\in\mathbb{R}$，$0+x=x$ 成立。

(F4) 加法逆元的存在性：对任意 $x\in\mathbb{R}$，存在 $-x\in\mathbb{R}$，使得 $x+(-x)=0$。

练习. 证明，加法逆元 $-x$ 是唯一的，即如果 $x'\in\mathbb{R}$ 也满足 $x+x'=0$，那么 $x'=-x$。

注记. 如果假定 *(F1)-(F4)* 成立，按照通行的代数学的概念，$(\mathbb{R},+)$ 被称作是一个（交换）群。我们强调 $-x$ 目前只是一个记号。根据俗套的约定，我们把 $x+(-y)$ 简写成 $x-y$。

(F5) 乘法结合律：$x\cdot(y\cdot z)=(x\cdot y)\cdot z$。

(F6) 乘法交换律：$x\cdot y=y\cdot x$。

(F7) 存在乘法单位元：存在 $1\in\mathbb{R}$，使得 $1\neq 0$ 并且对任意 $x\in\mathbb{R}$，$1\cdot x=x$ 成立。我们还要求 $1\neq 0$，从而 $\mathbb{R}$ 中至少有两个元素。

(F8) 乘法逆元的存在性：对任意 $x\in\mathbb{R}-\{0\}$，存在 $x^{-1}\in\mathbb{R}$，使得 $x\cdot x^{-1}=1$。

练习. 证明，$x\neq 0$ 乘法逆元 x^{-1} 是唯一的，即若 $x'\in\mathbb{R}$ 也满足 $x\cdot x'=1$，那么 $x'=x^{-1}$。

注记. *(F5)-(F9)* 这四条公理表明 $(\mathbb{R}^\times := \mathbb{R}-\{0\}, \cdot)$ 是一个（交换）群。我们强调 x^{-1} 只是一个记号。根据约定，我们把 $x\cdot y^{-1}$ 也记作 $\dfrac{x}{y}$。另外，有时候我们还省略掉 $\cdot$ 把 $x\cdot y$ 写成 xy。

(F9) 乘法分配律：$x\cdot(y+z)=x\cdot y+x\cdot z$。

注记. 假定 *(F1)-(F7)* 以及 *(F9)* 这八条，我们就称 $(\mathbb{R},+,\cdot)$ 是一个（交换）环；满足这九条公理的 $(\mathbb{R},+,\cdot)$ 被称作是一个域。

注记 (空间的概念.)**.** 上面的定义具有下述的模式，首先固定一个集合 $X=\mathbb{R}$，然后在 X 上加上额外的结构，比如说加法结构 $+: X\times X\to X$，当然，我们对这个额外的加法结构也可以做一些要求，比如说满足 *(F1)-F(4)* 等；我们还可以加更多的结构，比如还要求 X 上有乘法结构 $\cdot: X\times X\to X$ 并且对这些结构之间的关系也有限定（比如说 *(F9)*）。在数学中，所谓的**空间**通常指的是一个配备了某些结构的集合 X。

对于正整数 n，我们还约定 $x^n=\underbrace{x\cdot x\cdot\cdots\cdot x}_{n个}$，$nx=\underbrace{x+x+\cdots+x}_{n个}$。类似的，对于 $n\in\mathbb{Z}_{<0}$，我们约定 $x^n=\underbrace{x^{-1}\cdot x^{-1}\cdot\cdots\cdot x^{-1}}_{|n|个}$，$nx=\underbrace{(-x)+(-x)+\cdots+(-x)}_{|n|个}$。我们规定 $0x=0$，$x^0=1$。特别地，我们定义了以整数 $n\in\mathbb{Z}$ 为幂的幂函数：

$$\mathbb{R}\to\mathbb{R},\quad x\mapsto x^n.$$

不难验证，对任意的 $m,n\in\mathbb{Z}$，$(n+m)x=nx+mx$，$x^{n+m}=x^n\cdot x^m$。特别地，$nx+(-n)x=0$，$x^n\cdot x^{-n}=1$。

练习. *1)* 证明，对任意的 x,y，如果 $b\neq 0$，我们有

$$x+a=y+a\Rightarrow x=y;\quad x\cdot b=y\cdot b\Rightarrow x=y.$$

2) 证明，对任意的 x,y,z,w，如果 $y\neq 0$，$w\neq 0$，那么我们有

$$\frac{x}{y}+\frac{z}{w}=\frac{xw+zx}{yw},\quad \frac{x}{y}\cdot\frac{z}{w}=\frac{x\cdot z}{y\cdot w}.$$

3) 证明，对任意非零的 x 和 y，我们有

$$\left(\frac{x}{y}\right)^{-1}=\frac{y}{x}.$$

4) 证明，$(-1)\cdot x=-x$。据此，进一步证明 $(-x)\cdot y=-(xy)$，$(-x)\cdot(-y)=xy$。
（提示：利用 $x+(-1)\cdot x=1\cdot x+(-1)\cdot x$）

从此之后，我们可以不计后果地使用这些公式（熟知的公式是**所谓的直观**的一部分）。

(O) 序公理：$\mathbb{R}$ 是有序域

(O1) 序的传递性：$x \leqslant y, y \leqslant z \Rightarrow x \leqslant z$。

(O2) 序可以决定元素：$x \leqslant y, y \leqslant x \Rightarrow x = y$。

(O3) 全序关系：对任意的 x 和 y，$x \leqslant y$ 或者 $y \leqslant x$，二者必居其一（可以都成立）。

注记. 我们给出大于号和小于号的定义：如果 $x \leqslant y$ 且 $x \neq y$，我们就说 $x < y$；类似地，可以定义 $y < x$。所以任意给定 $x, y \in Q$，那么如下三种情形（这三种情形是互斥的）必居其一：

$$x < y, \quad x = y, \quad x > y.$$

我们在数学分析中偏爱 $\leqslant$ 和 $\geqslant$，这是因为这两种符号在后来的去极限的运算下是封闭的，这是后话。

另外，如果 $x > 0$，我们就称 x 是**正实数**并且称它的符号是正的，记做 $\mathrm{sign}(x) = +$；如果 $x > 0$，我们就称 x 是**负实数**并且称它的符号是负的，记做 $\mathrm{sign}(x) = -$。换句话说，我们定义了一个映射

$$\mathrm{sign} : \mathbb{R}^{\times} \to \{+, -\}.$$

我们在课程中会强调集合之间的映射这个概念。习惯上我们把正实数，记作 $\mathbb{R}_{>0}$，我们还将 $\geqslant 0$ 的实数称作非负实数，记作 $\mathbb{R}_{\geqslant 0}$；类似地，我们可以定义非正或者负实数并且“望文生义”地定义符号 $\mathbb{R}_{<0}$ 和 $\mathbb{R}_{\leqslant 0}$。

注记 (区间的定义（也请参见之后所谓的几何定义）). 假设 $a, b \in \mathbb{R}$，满足 $a < b$。我们定义**开区间**，**闭区间**和**半开半闭的区间**如下

$$[a, b] := \{x \in \mathbb{R} | a \leqslant x \leqslant b\}, \quad (a, b) := \{x \in \mathbb{R} | a < x < b\},$$
$$[a, b) := \{x \in \mathbb{R} | a \leqslant x < b\}, \quad (a, b] := \{x \in \mathbb{R} | a < x \leqslant b\}.$$

另外，作为约定，我们还采取下面的记号

$$[a, +\infty) := \{x \in \mathbb{R} | x \geqslant a\}, \quad (a, +\infty) := \{x \in \mathbb{R} | x > a\},$$
$$(-\infty, b] := \{x \in \mathbb{R} | x \leqslant b\}, \quad (-\infty, b) := \{x \in \mathbb{R} | x < b\}.$$

其中，a 和 b 分别被称作是这些区间的**左端点**或者**右端点**。

用数学分析中的黑话来说，符号 ∞ 被称作是**无穷**，$\pm\infty$ 被称作是**正／负无穷**，它们目前完全不具有具体的含义。

(O4) 与加法相容：$x \leqslant y \Rightarrow x + z \leqslant y + z$。

(O5) 与乘法相容：$x \geqslant 0, y \geqslant 0 \Rightarrow xy \geqslant 0$。

(A) Achimedes 公理：$\mathbb{R}$ **是 Archimedes 有序域**，即

对任意 $x > 0$ 和 y，总存在正整数 n，使得 $n \cdot x \geqslant y$。

思考题. 假设 $a>0$，你是否能证明开区间 $(0,a)$ 是非空的。通过反证法，你就可以看到所谓的 *Dedekind* 分割的影子。*Dedekind* 分割是构造实数的一种手段，我们在后面的课程会讨论。在思考的过程中你也会发现，真正的困难在于，基于目前的公理，我们不清楚 $\mathbb{R}$ 中都有什么样子的元素（目前我们只知道 $0,1,-1\in\mathbb{R}$），所以，能否大量的构造实数是一个很重要的问题。我们课程的一个关键点就是构造 $\sqrt{2}$，e 和 π，这听起来有些无聊，但是中学数学教学从来都没有给出这些数的具体定义（只要求大家必须接受它们的某些性质）。

练习. *1)* 证明，$x\geqslant 0$ 等价于 $-x\leqslant 0$；$y>1$ 可以推出 $0<\dfrac{1}{y}<1$。进一步证明，$x\geqslant y$ 等价于 $-x\leqslant -y$。

2) 证明，$1>0$，$-1\neq 1$。（提示：如若不然，那么，$-1>0$，利用 *(O5)* 就得到了矛盾）

3) 证明，如果 $x\leqslant y$，$a\leqslant 0$，那么 $a\cdot x\geqslant a\cdot y$。

4) 证明，如果 $a\leqslant b$，$x\leqslant y$，那么 $a+x\leqslant b+y$ 并且 $=$ 成立当且仅当 $a=b$，$x=y$；再证明，如果 $0<a\leqslant b$，$0<x\leqslant y$，那么 $ax\leqslant by$ 并且 $=$ 成立当且仅当 $a=b$，$x=y$。

5) 给定 $x,y\in\mathbb{R}$，如果对任意的 $a<x$ 都能推出 $a<y$，证明，$x\leqslant y$。

6) 证明，对任意的 $x\in\mathbb{R}$，我们有 $x^2\geqslant 0$。

7) 证明，如果 $a^2<a$，那么 $0<a<1$。

8) 如果非零实数 x 和 y 的符号相同，证明，$(x+y)^2>(x-y)^2$。

命题 1. $\mathbb{R}$ 包含所有有理数，即存在单射 $\iota:\mathbb{Q}\to\mathbb{R}$，使得对任意的 $x,y\in\mathbb{Q}$，我们有

$$\iota(x+_{\mathbb{Q}}y)=\iota(x)+\iota(y),\quad \iota(x\cdot_{\mathbb{Q}}y)=\iota(x)\cdot\iota(y),$$

其中 $+_{\mathbb{Q}}$ 和 $\cdot_{\mathbb{Q}}$ 分别为有理数 $\mathbb{Q}$ 上的乘法和加法。映射 $\iota:\mathbb{Q}\to\mathbb{R}$ 还保持序关系，即对任意的 $x,y\in\mathbb{Q}$，如果 $x\leqslant_{\mathbb{Q}}y$，那么 $\iota(x)\leqslant\iota(y)$，其中 $\leqslant_{\mathbb{Q}}$ 是有理数上的序。

证明: 首先对于整数 n 定义映射 $\iota(n)\in\mathbb{R}$。我们令

$$\iota(n)=\begin{cases}\overbrace{1+1+\cdots+1}^{n个}, & \text{如果 } n>0;\\ 0, & \text{如果 } n>0;\\ \underbrace{(-1)+(-1)+\cdots+(-1)}_{-n个}, & \text{如果 } n<0.\end{cases}$$

不难验证（请同学参考课堂笔记），对任意的 $m,n\in\mathbb{Z}$，我们都有

$$\iota(m+_{\mathbb{Q}}n)=\iota(m)+\iota(n),\ \iota(-n)=-\iota(n),\ \iota(m\cdot_{\mathbb{Q}}n)=\iota(m)\cdot\iota(n).$$

据此，为了验证 ι 是单射，只要说明如果 $\iota(m)=\iota(n)$，则 $m=n$。我们有

$$\iota(m-n)=\iota(m)+\iota(-n)=\iota(m)-\iota(n)=0,$$

按照定义，就有 $k=m-n$（因为对任意的 $k>0, \overbrace{1+1+\cdots+1}^{k个}>0, \overbrace{(-1)+1+\cdots+(-1)}^{k个}<0$）。对于有理数 $x=\frac{p}{q}\in\mathbb{Q}$，其中 $p,q\in\mathbb{Z}$，$q\neq 0$，我们定义

$$\iota(x)=\frac{\iota(p)}{\iota(q)},$$

其中，因为 p 和 q 是整数，所以 $\iota(p)$ 和 $\iota(q)$ 已经有了定义。当然，x 可以表示为其他的整数的商的形式，比如说，$x=\frac{s}{t}$，其中 $s,t\in\mathbb{Z}$，为了说明 $\iota(x)$ 是良好定义的，我们就要说这两种表示所给出的 $\mathbb{R}$ 中的元素是一样的，即证明

$$\frac{\iota(p)}{\iota(q)}=\frac{\iota(s)}{\iota(t)}.$$

根据整数情况已经证明的结论，我们知道上式等价于

$$\iota(p)\cdot\iota(t)=\iota(q)\cdot\iota(s)\Leftrightarrow\iota(pt)=\iota(qs)\Leftrightarrow pt=qs,$$

其中最后一个等价性用到了 ι 在整数集上是单射，所以 $\iota(x)$ 是良好定义的，从而我们得到了映射

$$\iota:\mathbb{Q}\to\mathbb{R}.$$

为了说明 ι 是单射，考虑 $\iota(\frac{p}{q})=\iota(\frac{s}{t})$，其中 p,q,s,t 是整数。此时，按照定义，我们有

$$\frac{\iota(p)}{\iota(q)}=\frac{\iota(s)}{\iota(t)}\Rightarrow pt=qs\Rightarrow\frac{p}{q}=\frac{s}{t}.$$

最后，我们把 ι 保持序关系，即说明如果 $\frac{p}{q}>\frac{s}{t}$，那么 $\frac{\iota(p)}{\iota(q)}>\frac{\iota(s)}{\iota(t)}$，其中，我们总能假设 $q>0\,t>0$。由于 $\iota(q)>0$，$\iota(t)>0$（因为对任意的 $k>0$，$\overbrace{1+1+\cdots+1}^{k个}>0$），所以 $\frac{\iota(p)}{\iota(q)}>\frac{\iota(s)}{\iota(t)}$ 等价于

$$[\iota(p)\cdot\iota(t)>\iota(q)\cdot\iota(s)\Leftrightarrow\iota(pt-qs)>0.$$

后者是显然的，因为 $pt-qs>0$。 □

注记. 我们之后将 $\mathbb{Q}$ 和 ι 的像等同。自此往后，我们就在这个意义下认为有理数 $\mathbb{Q}$ 是 $\mathbb{R}$ 的子集。换句话说，我们用 n 来表示 $\mathbb{R}$ 中的 $\overbrace{1+1+\cdots+1}^{k个}$，$\frac{p}{q}$ 表示 $\frac{\overbrace{1+1+\cdots+1}^{p个}}{\underbrace{1+1+\cdots+1}_{q个}}$ 等。习惯上，我们将 $\mathbb{R}-\mathbb{Q}$ 的元素称为是**无理数**（这是一个在中文世界里被广泛接受的“无理的”术语）。

练习. 有了以上的各种准备和经验，我们就不难证明

1) 证明，利用 $\mathbb{R}$ 中的 n 的定义，我们有 $n\cdot x=nx$，$n\cdot x=nx$。（先证明在 $\mathbb{R}$ 中，我们有 $n\cdot x=\overbrace{x+x+\cdots+x}^{n个}$，其中 $n>0$ 是 $\mathbb{R}$ 中的一个自然数）。

2) 证明，对于任意的 $a<b$，(a,b) 有无限多个元素。（提示：我们可以考虑 $\frac{1}{2}(a+b)$ 并利用 $0<\frac{1}{2}<1$ 这个事实）

3) 如果 $\mathbb{R}$ 中存在元素 $o>0$，使得对任意的 $x>0$，我们都有 $o<x$，我们就称 o 是无穷小元。证明，$\mathbb{R}$ 中没有无穷小元。（有一种专门研究含有无穷小的数的分析，叫做非标准分析）

(I) 区间套公理

给定有限（即要求下面的 a_n 和 b_n 均为实数）闭区间的序列 $\{I_n=[a_n,b_n]\}_{n=1,2,\cdots}$，如果这个序列是下降的，即 $I_1\supset I_2\supset I_3\supset\cdots$（等价于对任意的 $n\geqslant 1$ 都有 $a_n\leqslant a_{n+1}$ 并且 $b_{n+1}\geqslant b_n$），那么它们的交集非空，即

$$\lim_{n\to\infty} I_n:=\bigcap_{n\geqslant 1} I_n\neq\emptyset.$$

定义 2. 我们将满足上述四条公理系统 **(F)**，**(O)**，**(A)**，**(I)** 的四元组 $(\mathbb{R},+,\cdot,\leqslant)$ 称作是**实数**。

注记. 这个定义目前并不是**良好定义的**：我们完全不知道这样的实数理论是不是**唯一的**；我们甚至没有证明满足四条的四元组 $(\mathbb{R},+,\cdot,\leqslant)$ 是存在的。另外，除了有理数之外，我们并没有证明无理数（比如说 $\sqrt{2}$）是存在的。

我们注意到，如果就强行要求 $\mathbb{R}$ 是全体有理数的集合并且配备了 $+$，$\cdot$ 和 $\leqslant$ 这几种有理数上的结构，我们所得到的四元组是满足 **(F)**，**(O)**，**(A)** 这三条公理系统的，所以，要想真的得到我们中学所熟悉的实数（比如说存在 $\sqrt{2}$），区间套公理是必不可缺的。

2 区间套公理与确界原理，距离空间

二零一九年九月十二日，星期四，多云

实数的基本性质

假设存在 $(\mathbb{R}, +, \cdot, \leqslant)$ 满足 **(F)**，**(O)**，**(A)**，**(I)** 四套公理，除了上次课练习题中已证明的结论，实数还有许多为人们所"熟知的"性质。我们强调的是这些性质需要在以上公理化的体系下被证明。我们补充几个这样的性质，它们证明比较简单，我们仍以课堂习题的方式给出，细节留给同学在整理笔记的时候给出：

练习. *1)* $x \leqslant y$ 和 $y < z$ 可以推出 $x < z$；$x < y$ 和 $y \leqslant z$ 可以推出 $x < z$。

2) $\mathbb{R}$ 的有限子集都有唯一的最大元和唯一的最小元（我们约定集合中的两个元素是不同的）。特别地，如果 $A \subset \mathbb{R}$ 是有限子集，$n = |A|$，那么可以将 A 中的元素排序，使得

$$A = \{a_1, a_2, \cdots, a_n\}, \quad a_1 < a_2 < \cdots < a_n.$$

3) $x_1, \cdots, x_n$ 和 $y_1, \cdots, y_n$ 是实数，对任意的指标 $1 \leqslant i \leqslant n$，$x_i \leqslant y_i$，那么

$$x_1 + \cdots + x_n \leqslant y_1 + \cdots + y_n.$$

上面的不等式取等号当且仅当对所有的 i，我们都有 $x_i = y_i$。特别地，对于非负实数 $a_1, \cdots, a_n$，它们的和 $\sum_{i=1}^{n} a_i$ 也是非负实数，此和为零当且仅当所有的 a_i 均为零。（这个命题当 $n = 2$ 时前面已经证明）

关于实数有一个重要（但是简单）的事实：存在充分大的正数，也存在充分小的正数，这将是数学分析中反复使用的事实。这个命题的精确说法如下（这是我们第一次用到 $\varepsilon - \delta$ 语言）：

引理 3. 对任意的正实数 A，总存 M，使得 $M > A$；对任意的正实数 a，总存在正实数 ε，使得 $\varepsilon < a$。

证明：我们可以选取 $M = A + 1$，$\varepsilon = \dfrac{a}{2}$。证明的要点在于说明 $\frac{1}{2} < 1$。 □

我们按照如下方式定义**绝对值函数** $|\cdot| : \mathbb{R} \to \mathbb{R}_{\geqslant 0}$：对任意的 $x \in \mathbb{R}$，

$$x \mapsto |x| = \begin{cases} x, & \text{如果} x \geqslant 0; \\ -x, & \text{如果} x < 0. \end{cases}$$

练习. a 是非负实数。证明，$|x| \leqslant a$ 当且仅当 $-a \leqslant x \leqslant a$。特别地，$x = 0$ 当且仅当 $|x| = 0$。

引理 4. 绝对值函数 $|\cdot| : \mathbb{R} \to \mathbb{R}_{\geqslant 0}$ 满足下面的性质：

1) 假设 $a \in \mathbb{R}_{\geqslant 0}$。证明，$|x| \leqslant a$ 当且仅当 $-a \leqslant x \leqslant a$。特别地，$x = 0$ 当且仅当 $|x| = 0$。

2) 对任意实数 x 和 y，我们有

$$|x+y| \leqslant |x|+|y|, \quad \big||x|-|y|\big| \leqslant |x-y|.$$

特别地，我们有 $|x_1+\cdots x_n| \leqslant |x_1|+\cdots+|x_n|$，其中 $\{x_i\}_{1\leqslant i\leqslant n} \subset \mathbb{R}$。

证明: 我们只对 2) 中第一个不等式进行证明，其余的部分是简单的，留给同学自行验证。对 x 和 y 分情况讨论：如果 x 和 y 的符号一致，命题是显然的，所以不妨假设 $x \leqslant 0 \leqslant y$。此时，$x+y \leqslant |x|+|y|$ 是明显的，下面说明 $x+y \geqslant -(|x|+|y|)$，这是因为 $-|x| \leqslant x$，$-|y| \leqslant y$。 □

在实数 $\mathbb{R}$ 上研究分析有一个重要的"几何观点"：尽管这个看法很"自然"，然而从公理化的观点而言，这绝非显而易见：

> 我们将 $\mathbb{R}$ 想象成一条"直线"：每个实数对应直线上的一个点；大小关系可以两点的一左一右来确定；区间就是两点之前的线段，诸如此类。
> 这种形象的看法使得在很多场合下的推理和计算变得容易操作和叙述。然而必须强调的是，在证明或者计算的过程中上述图像只起辅助的作用，一切结论都是严格根据从实数公理出发的所得到的结论通过正确推理而来。（这一如平面几何中画图对证明所起到的作用）
> 分析学会进一步深化这种几何化的看法：我们倾向于 **几何地**考虑问题。比如说，我们会将尽量多的数学对象想象成空间的点并发展相应的理论使得于几何的直观在应用的时候是严格的。

我们注意到，迄今为止，我们还未使用公理 **(I)**（区间套原理），上次课的最后也提到过，如果只用前三套公理体系，那么这种"实数"可能只包含有理数，这自然不是我们想要的实数理论。下面要证明的*确界原理*就依赖于区间套原理。

$X \subset \mathbb{R}$ 是实数的集合，$a, A \in \mathbb{R}$。如果对任意 $x \in X$，都有 $x \leqslant A$，就称 A 是 X 的一个**上界**；如果对任意 $x \in X$，都有 $x \geqslant a$，我们就称 a 是 X 的一个**下界**。如果 X 既有上界又有下界，我们就说 X 是**有界的**。X 有界等价于存在（大的）正实数 M，使得对任意 $x \in X$，我们都有 $|x| \leqslant M$。

定理 5 (确界原理)**.** *假设 $X \subset \mathbb{R}$ 是非空的并且 X 有上界。令 $\overline{\mathcal{M}} = \big\{\overline{M} \in \mathbb{R} \big| \overline{M}$ 是 X 的上界$\big\}$，则 $\overline{\mathcal{M}}$ 有（唯一的）最小元，即存在 $\overline{M_0} \in \overline{\mathcal{M}}$，使得任意的 $\overline{M} \in \overline{\mathcal{M}}$，都有 $\overline{M_0} \leqslant \overline{M}$。*

*我们称 $\overline{M_0}$ 为 X 的***上确界***，记作 $\sup X$。*

证明: 任选 $x \in X$，任取上界 $\overline{M} \in \overline{\mathcal{M}}$ 并不妨假设 $x \notin \overline{\mathcal{M}}$（否则 $x = \sup X$）。对每一个正整数 $n \geqslant 1$，根据 Archimedes 公理，存在正整数 k，使得 $x + k2^{-n} \geqslant \overline{M}$，从而 $x + k2^{-n} \in \overline{\mathcal{M}}$ 是上界。我们令 k_n（$\geqslant 1$ 根据 $x \notin \overline{\mathcal{M}}$）是最小的使得 $x + k_n 2^{-n} \in \overline{\mathcal{M}}$ 是上界的正整数，令 $I_n = [x + (k_n - 1)2^{-n}, x + k_n 2^{-n}] = [a_n, b_n]$（闭区间），那么

a) $I_n \cap X \neq \emptyset$。

如果不然，对任意的 $y \in X$，都有 $y \notin I_n$。很明显，y 落在 I_n 的右端点 b_n 的左边（因为 b_n 是上界），所以 y 只能落在整个 I_n 的左边，这表明 I_n 的左端点 a_n 也是上界。但是，这与 k_n 的选取方式（最小性）矛盾。

b) $I_n \supset I_{n+1}$。

根据区间 I_n 的构造方式，它的右端点 b_n 是上界，左端点 a_n 不是上界（即存在 $y \in X$ 使得 $y > a_n$）。为了说明 $I_n \supset I_{n+1}$，我们要证明下面的不等式即可：

$$a_n \leqslant a_{n+1},\ b_{n+1} \leqslant b_n \Leftrightarrow (k_n - 1)2^{-n} \leqslant (k_{n+1} - 1)2^{-(n+1)},\ k_{n+1}2^{-(n+1)} \leqslant k_n 2^{-n}$$

我们用反证法。

先比较右端点，如果 $k_{n+1}2^{-(n+1)} > k_n 2^{-n}$，那么 $k_{n+1}2^{-(n+1)}$ 比 $k_n 2^{-n}$ 至少大出来 $2^{-(n+1)}$ 这么多，所以 $(k_{n+1} - 1)2^{-(n+1)} \geqslant k_n 2^{-n}$，这表明 I_{n+1} 的左端点 a_{n+1} 在 I_n 的右端点 b_n 的右边（可能重合），即 $a_n \geqslant b_{n+1}$，从而 a_n 也是 X 的上界，矛盾。

其次比较左端点，如果 $(k_{n+1} - 1)2^{-(n+1)} < (k_n - 1)2^{-n}$，和上面一样的推理我们就知道 $k_{n+1}2^{-(n+1)} \leqslant (k_n - 1)2^{-n}$，这表明 I_{n+1} 的右端点 b_{n+1} 在 I_n 的左端点 a_n 的左边（可能重合），所以 a_n 是 X 的上界，矛盾。

根据区间套公理，我们令 $J = \bigcap\limits_{n \geqslant 1} I_n \neq \emptyset$。我们注意到，$J$ 恰好只含有一个元素：如果不然，我们在 J 中选取两个点（我们认为点 = 数）a 和 b 并且不妨假设 $a < b$。按照定义，对任意的 n，$a, b \in I_n$。特别地，我们可以选取 n，使得 $b - a > 2^{-n}$（即 $2^n(b-a) < 1$），此时区间 I_n 的长度小于 a 和 b 之间的距离，它不可能同时包含 a 和 b 这两个点，矛盾！所以，我们假设 $J = \{\overline{M_0}\}$。那么，我们有

- $\overline{M_0} \in \overline{\mathcal{M}}$，即 $\overline{M_0}$ 是 X 的一个上界。

  如若不然，存在 $y \in X$，使得 $y > \overline{M_0}$。根据定义，对每个 n，我们都有 $\overline{M}_0 \in I_n$（因为 J 是所有 I_n 的交集），所以 $b_n - \overline{M}_0 \leqslant 2^{-n}$。我们可以选取很大的 n，使得 $2^{-n} < y - \overline{M_0}$，从而 $y > \overline{M_0} + y \geqslant b_n$，这和 b_n 是 X 的上界矛盾（按照 I_n 的构造方式，它的右端点 b_n 是 X 的上届）。

- $\overline{M_0} = \min\limits_{\overline{M} \in \overline{\mathcal{M}}} \overline{M}$ 是最小的上界。

  如若不然，那么存在 $\widetilde{M} \in \overline{\mathcal{M}}$ 使得 $\widetilde{M} < \overline{M_0}$。由于对每个 n，我们都有 $\overline{M}_0 \in I_n$，所以 $a_n + 2^{-n} \geqslant \overline{M}_0$。我们可以选取很大的 n，使得 $2^{-n} < \overline{M_0} - \widetilde{M}$，从而 $a_n > \widetilde{M}$，这和 a_n 不是 X 的上界矛盾。

综上所述，命题得证。 □

注记. 关于确界，我们有下面几个补充，第三个尤为重要：

1) 对偶的命题：假设 $X \neq \emptyset$ 有下界，令 $\underline{\mathcal{M}} = \{\underline{M} \in \mathbb{R} | \underline{M}$ 是 X 的下界$\}$。那么，$\underline{\mathcal{M}}$ 有（唯一的）最大元，即存在 $\underline{M_0} \in \underline{\mathcal{M}}$，使得任意的 $\underline{M} \in \underline{\mathcal{M}}$，都有 $\underline{M_0} \geqslant \underline{M}$。我们称 $\underline{M_0}$ 为 X 的**下确界**，记作 $\inf X$。

2) $\inf X \leqslant \sup X$。

3) (上确界的刻画) 假设非空集合 $X \subset \mathbb{R}$ 有上界并且实数 M 是 X 的上界。那么，下面两个命题是等价的

- $M = \sup X$。
- 对任意的 $\varepsilon > 0$，都存在 $x \in X$，使得 $x > M - \varepsilon$。

集合的下确界也可以类似地刻画，我们略去叙述。这个刻画的证明我们留做作业题。

4) 假设 $(\mathbb{R}, +, \cdot, \leqslant)$ 满足 **(F)**，**(O)** 和 *Archimedes* 公理 **(A)** 这三套公理。如果我们假设确界定理成立，那么区间套公理 **(I)** 可以被证明：

对任意的闭区间套序列 $\{I_n = [a_n, b_n]\}_{n=1,2,\cdots}$，其中对任意的 n，有 $a_n \leqslant a_{n+1}$，$b_{n+1} \geqslant b_n$。很明显，集合 $A = \{a_n\}_{n\geqslant 1}$ 是有上界的（取 b_1 为上界），根据确界原理，我们令 $a = \sup A$，那么 $a \leqslant b_n$（因为每个 b_n 都是上界而上确界是最小的上界）；$B = \{b_n\}_{n\geqslant 1}$ 是有下界的（取 a 为下界），根据确界定理，我们令 $b = \inf B$，所以，$a \leqslant b$。那么，$\bigcap\limits_{n\geqslant 1} I_n \supset [a, b] \neq \emptyset$。

空间的概念

我们引入度量／距离空间的概念。我们已经讲过，所谓的空间，就是一个集合加上一些附加的结构：

定义 6. X 是集合。如果存在 $d : X \times X \to \mathbb{R}_{\geqslant 0}$ 是 X 上双变量的函数，满足如下三条性质：

a) 对任意的 x 和 y，$d(x, y) \geqslant 0$ 并且取等号当且仅当 $x = y$；

b) 对任意的 x 和 y，$d(x, y) = d(y, x)$；

c) 三角不等式：对任意的 $x, y, z \in X$，我们有 $d(x, z) \geqslant d(x, y) + d(y, z)$。

我们就称二元组 (X, d) 是一个**距离空间**或者**度量空间**。函数 d 被称作该距离空间上的**距离函数**。

注记. 直观上，$d(x, y)$ 衡量空间 X 中两点 x 和 y 的远近（我们可以形象地将 X 想成是中学所熟悉的平面）。

1) 在距离空间的定义中，我们已经用到了实数 $\mathbb{R}$ 的概念。

2) (**重要**) 对于 $x, y \in \mathbb{R}$，定义 $d(x, y) = |x - y|$，那么 $(\mathbb{R}, d)$ 是度量空间（只要验证定义即可）。从此往后，我们就可以将 $\mathbb{R}$ 等同成一个几何对象了。

3) (**重要**) 我们令 $\mathbb{R}^n = \underbrace{\mathbb{R} \times \mathbb{R} \times \cdots \times \mathbb{R}}_{n\text{个}}$，即 $\mathbb{R}^n = \big\{(x_1, \cdots, x_n) \big| x_1 \in \mathbb{R}, \cdots, x_n \in \mathbb{R}\big\}$。对于 $x, y \in \mathbb{R}^n$，我们定义

$$d(x, y) = \sqrt{(x_1 - y_1)^2 + \cdots + (x_n - y_n)^2}, \quad x = (x_1, \cdots, x_n),\ y = (y_1, \cdots, y_n).$$

那么，$(\mathbb{R}^n, d)$ 是度量空间（验证三角不等式时候需要用到所谓的 *Cauchy-Schwarz* 不等式）。注意到，我们目前并没有定义开方运算！

4) 考虑复数域 $\mathbb{C}$。对于 $z_1, z_2 \in \mathbb{C}$，定义 $d(z_1, z_2) = |z_1 - z_2|$，那么 $(\mathbb{C}, d)$ 是度量空间。

5)（子空间）假设 $Y \subset X$，我们定义 Y 上的距离函数：

$$d_Y : Y \times Y \to \mathbb{R}, (y_1, y_2) \mapsto d_Y(y_1, y_2) = d(y_1, y_2).$$

那么，(Y, d_Y) 是度量空间。我们称 d_Y 是 d 在 Y 上的**诱导度量**，(Y, d_Y) 称作是 (X, d) 的**子（度量）空间**。

我们之后会经常用到稠密性的概念。给定度量空间 (X, d)，$Y \subset X$ 是子集。如果对任意的 $x \in X$ 和任意（小）的 $\varepsilon > 0$，都存在 $y \in Y$，使得 $d(y, x) < \varepsilon$，我们就称 Y 在 X 中是**稠密的**。直观上，Y 在 X 中稠密说的是对于 X 的每个点都有一个 Y 中的点和它离得要多近就有多近。对这个概念的理解有助于学习 $\varepsilon - \delta$ 语言。我们有如下经典的习题：

练习. 证明，有理数和无理数在实数（这是一个度量空间）中都是稠密的。

我们还可以引入其他的几何概念，比如说区间的长度（高维数时称为面积／体积或者测度）的概念：对于区间 $I = [a, b]$ 或者 (a, b) 或者相应的半开半闭的形式，定义 I 的**长度**为 $|I| = |b - a|$。我们不系统地引入所有和几何相关的概念，而是这个概念真的必要时再进行讨论。

最终，我们引用线性／向量空间的概念，作为另一个例子来阐述所谓的空间的概念：

定义 7. $\mathbb{F} = \mathbb{R}$ 是实数域（可以是别的域，比如说有理数或者复数），V 是集合。我们假设存在两种运算：

- 加法运算。$+ : V \times V \to V, \quad (v, w) \mapsto v + w$；
- 数乘运算。$\cdot : \mathbb{F} \times V \to V, \quad (\lambda, v) \mapsto \lambda \cdot v$。

我们假设三元组 $(V, +, \cdot)$ 满足如下的八条公理（其中 $\lambda\, \mu \in \mathbb{F}$，$u, v, w \in V$ 是任意选取的）：

1) 加法结合律：$u + (v + w) = (u + v) + w$；

2) 加法交换律：$u + v = v + u$；

3) 存在加法单位元：存在 $0 \in V$（被称作是 V 的**原点**），使得对任意 $v \in V$，$0 + v = v$ 成立；

4) 加法逆元的存在性：对任意 $v \in V$，存在 $-v \in \mathbb{R}$，使得 $v + (-v) = 0$；

5) 数乘的结合律：$\lambda \cdot (\mu \cdot v) = (\lambda \cdot_{\mathbb{F}} \mu) \cdot v$；

6) 数乘的分配律之一：$(\lambda +_{\mathbb{F}} \mu) \cdot v = \lambda \cdot v + \mu \cdot v$；

7) 数乘的分配律之二：$\lambda \cdot (u + v) = \lambda \cdot u + \lambda \cdot v$；

8) 乘法单位元：$1 \cdot v = v$。

我们就称三元组 $(V, +, \cdot)$ 是 $\mathbb{F}$ 上的一个**线性空间**或者**向量空间**，或者称作 $\mathbb{F}$-线性空间。

练习. 试在 $\mathbb{R}^n$ 定义 $+$ 和 $\cdot$ 的结构使得它成为一个 $\mathbb{R}$-线性空间。

3 实数的构造：Dedekind 分割

二零一九年九月十六日，星期一，晴天

Dedekind 分割与实数的构造

定义 8 (Dedekind 分割)**.** $X \subset \mathbb{Q}$ 是有理数的子集，令 $X' = \mathbb{Q} - X$。如果下面三条性质都成立：

1) $X \neq \emptyset$，$X' \neq \emptyset$；

2) 对任意 $x \in X$，$x' \in X'$，都有 $x < x'$；

3) X 中没有最大元，

我们就称 X 或 $X \cup X'$ 是 $\mathbb{Q}$ 的一个 **Dedekind 分割**。我们用 $\mathcal{R}$ 表示所有 *Dedekind* 分割组成的集合。

注记. 我们可以重新解读后两条性质：

- 第二条 *2)* 等价于说，如果 $x_1 \in X$，那么对任意的 $x_2 < x_1$，我们就有 $x_2 \in X$。这也表明，如果 $x_1' \in X$，那么对任意的 $x_2' > x_1'$，我们就有 $x_2' \in X'$。
- 第二条 *3)* 指的是对任意 $x \in X'$，总存在 $x' \in X$，使得 $x' > x$。

例子. 我们给出两个 *Dedekind* 分割的例子（课堂练习：直接验证定义）：

a) 假设 $\frac{p}{q}$ 是有理数，其中 $p \in \mathbb{Z}$，$q \in \mathbb{Z}_{\geqslant 1}$ 并且 p 和 q 互素，我们定义

$$X_{\frac{p}{q}} = \{x \in \mathbb{Q} \big| x < \frac{p}{q}\}.$$

这个例子将给出所有的有理数。

b) 我们定义 $X_{\sqrt{2}}$ 如下：

$$X_{\sqrt{2}} = \{x \in \mathbb{Q}_{\geqslant 0} \big| x^2 < 2\} \cup \mathbb{Q}_{\leqslant 0}.$$

这个例子将给出我们所熟悉的 $\sqrt{2}$。

序结构

现在，我们要在 $\mathcal{R}$ 上定义序关系，加法和乘法，使得 $(\mathcal{R}, +, \cdot, \leqslant)$ 满足四套公理。
我们先定义序关系：

定义 (序关系)**.** 对任意的 $X, Y \in \mathcal{R}$，作为 $\mathbb{Q}$ 的子集合，如果

- $X = Y$，我们称 $X = Y$；

- $X \subset Y$ 且 $X \neq Y$，我们称 $X < Y$（也记做 $Y > X$）。

我们首先来验证这个序关系满足公理 (O1)，(O2) 和 (O3)：

(O3). $\leqslant$ 是一个全序关系，即对任意的 $X, Y \in \mathcal{R}$，$X \subset Y$ 和 $Y \subset X$ 必居其一：

我们假设 $X \not\subset Y$，只要说明 $Y \subset X$ 即可。根据假设，存在 $x \in X$，使得 $x \notin Y$，根据序的定义中的 2)，对任意的 $y \in Y$，都有 $y \leqslant x$（否则 $x < y$，$y \in Y$，就可以推出 $x \in Y$，矛盾!），从而 $y \in X$（因为 $x \in X$），所以 $Y \subset X$。

(O1). 对任意的 $X, Y, Z \in \mathcal{R}$，如果 $X < Y$，$Y < Z$，那么 $X < Z$。

这就是集合关系的传递性的重新叙述。

(O2). 对任意的 $X, Y \in \mathcal{R}$，$X < Y$，$X = Y$ 或者 $X > Y$ 三者恰有一种情形成立。

用集合的包含关系来看，这是显然的。

一个有趣的事情是，我们现在就可以证明确界原理了（只需要用到序的定义!）。为此，我们先做一番准备。假设 $\mathcal{X} \subset \mathcal{R}$（这是一个由若干 Dedekind 分割所组成的集合）有上界，即存在 Dedekind 分割 M，使得对任意 $X \in \mathcal{X}$，$X \leqslant M$，这里我们假设 $\mathcal{X} \neq \emptyset$。我们现在定义 $\sup \mathcal{X}$ 的候选：

$$M_0 = \bigcup_{X \in \mathcal{X}} X.$$

这是有理数的子集，我们首先要说明 $M_0 \in \mathcal{R}$，即 M_0 是 Dedekind 分割。我们逐条验证定义：

1) $M_0 \neq \emptyset$，$M_0' \neq \emptyset$。

$M_0 \neq \emptyset$ 是显然的。为了说明 $M_0' \neq \emptyset$，我们考虑 $\mathcal{X}$ 的一个上界 M，也就是说对于任意的 $X \in \mathcal{X}$，我们都有 $X \subset M$。此时，我们考虑任意的一个 $m' \in M'$，这是 M_0' 中的一个元素。

2) 对每一个 $x_1 \in M_0$，如果 $x_2 < x_1$，那么一定有 $x_2 \in M_0$。

实际上，由于 $x_1 \in M_0$，所以存在 Dedekind 分割 $X \in \mathcal{X}$，使得 $x_1 \in X$，所以 $x_2 \in X \subset M_0 = \bigcup_{Y \in \mathcal{X}} Y$。

3) M_0 中没有最大元。

如若不然，我们可以找到 $x \in M_0$ 是最大元。根据 M_0 的定义，存在 Dedekind 分割 $X \in \mathcal{X}$，使得 $x \in X$，那么 X 必须是 X 的最大元（因为 $X \subset M_0$），这与 X 是 Dedekind 分割矛盾。

我们下面说明，M_0 是 $\mathcal{X}$ 最小的上界，即若定义

$$\mathcal{M} = \{M \in \mathcal{R} \mid M \text{是 } \mathcal{X} \text{ 的上界}\},$$

那么，我们有 $M_0 \in \mathcal{M}$ 并且如果 $M \in \mathcal{M}$，那么 $M_0 \leqslant M$：

$M_0 \in \mathcal{M}$ 是明显的。为了说明 $M_0 \leqslant M$，其中 M 是上界，我们用定义：由于 M 是上界，所以对任何的 $X \in \mathcal{X}$，我们都有 $X \subset M$，所以它们的并 $\bigcup_{X\in\mathcal{X}} X$ 也是 M 的子集，即 $M_0 \subset M$，这等价于 $M_0 \leqslant M$。

这就证明了确界原理。

加法结构

我们现在定义加法运算：

定义 (加法). 对任意的 $X, Y \in \mathcal{R}$，我们定义

$$X + Y := \{x + y | x \in X, y \in Y\}.$$

我们还定义零元素 $\overline{0}$ 为

$$\overline{0} = X_0 = \{x \in \mathbb{Q} | x < 0\}.$$

注记. 我们首先说明加法运算

$$\mathbb{R} \times \mathbb{R} \xrightarrow{+} \mathbb{R},\ (X, Y) \mapsto X + Y$$

是良好定义的，即说明上面所定义的 $X+Y$ 的确是 *Dedekind* 分割（先验地来看，$X+Y$ 只是 $\mathbb{Q}$ 的一个子集）。为此，我们只需要依次验证 *Dedekind* 分割的定义中的三条性质：

1) $X+Y$ 及其补集 $(X+Y)'$ 是非空的。

由于 X 和 Y 都非空，$X+Y$ 自然非空；我们随意选定 $x_0' \in X'$，$y_0' \in Y'$，那么对任意的 $x \in X$，$y \in Y$，我们都有 $x < x_0'$，$y < y_0'$。所以，对任意的 $x \in X$，$y \in Y$，我们都有 $x + y < x_0' + y_0'$，这就表明 $x_0' + y_0'$ 比 $X+Y$ 中所有数都大。特别地，$x_0' + y_0' \notin X + Y$，也就是说 $x_0' + y_0' \in (X+Y)'$。

2) 对任意的 $x + y \in X + Y$，其中 $x \in X$，$y \in Y$，如果有理数 $z < x + y$，那么 $z \in X + Y$。

为此，我们注意到 $z - x < y$，根据，$y \in Y$，我们知道 $z - x \in Y$，所以

$$z = \underbrace{x}_{\in X} + \underbrace{z - x}_{\in Y} \in X + Y.$$

3) $X+Y$ 中没有最大元。

如果不然，假设 $x_0 + y_0 \in X + Y$ 是最大元，其中 $x_0 \in X$，$y_0 \in Y$。根据定义，由于 X 和 Y 没有最大元，所以存在 $x \in X$，$y \in Y$，使得 $x_0 < x$，$y_0 < y$，从而 $x_0 + y_0 < x + y$。但是 $x + y \in X + Y$，这与 $x_0 + y_0$ 是最大元这个假设矛盾。

我们下面需要依次验证和加法相关的几条公理。如果不加其他的陈述，我们总假设 $X, Y, Z \in \mathcal{R}$ 是任意选定的 Dedekind 分割。那么，我们有

(F1) $(X+Y)+Z=X+(Y+Z)$。

根据定义，这是显然的。

(F2) $X+Y=Y+X$。

根据定义，这也是显然的。

(O4) 如果 $X<Y$，那么 $X+Z<Y+Z$。

这个性质的证明不是一句话就可以完成的，我们先准备一个技术工具

引理 9. *X 是 Dedekind 分割。那么，对任意的正整数 n，总存在 $x\in X$，$x'\in X'$，使得*

$$0<x'-x<\frac{1}{n}.$$

证明想法可以追溯到庄周：一尺之棰，日取其半，万世不竭。—《庄子 · 天下》

证明: 任意选定 $x_0\in X, x_0'\in X'$，我们通过归纳的方式来构造一些列的 (x_k,x_k')，其中 $k\geqslant 0$，$x_k\in X$，$x_k'\in X'$。假设 (x_k,x_k') 已经构造好了，那么我们考虑 $y=\dfrac{1}{2}(x_k+x_k')$。有两种可能的情况：如果 $y\in X$，那么定义 $(x_{k+1},x_{k+1}')=(y,x_k')$；如果 $y\in X'$，那么定义 $(x_{k+1},x_{k+1}')=(x_k,y)$。很明显，我们有

$$|x_{k+1}-x_{k+1}'|=\frac{1}{2}|x_k-x_k'|\quad\Rightarrow\quad |x_k-x_k'|=\frac{1}{2^k}|x_0-x_0'|.$$

上面的等式对任意的 k 都对。特别地，我们选取 k_0，使得 $2^{k_0}>n|x_0-x_0'|$，从而 $x=x_{k_0}\in X$ 和 $x'=x_{k_0}'\in X'$ 满足要求。 □

回到 (O4) 的证明：首先注意到不等式 $X+Z\leqslant Y+Z$ 是显然的，所以只要说明 $X+Z\neq Y+Z$ 即可。我们选取 $y,\tilde{y}\in Y-X$，使得 $y>\tilde{y}$（先取 $\tilde{y}$，由于 Y 没有最大元，可以再选 y）。现在选取自然数 n，使得 $\dfrac{1}{n}<y-\tilde{y}$（在有理数中这一点总能做到）。根据引理，我们再取 $z\in Z$，$z'\in Z'$，使得 $z'-z<\dfrac{1}{n}$。

我们现在说明 $y+z\notin X+Z$：如若不然，那么存在 $x\in X$，$z_1\in Z$，使得 $y+z=x+z_1$。按照定义，$x<\tilde{y}$，所以 $z_1<z'$。这样，我们得到如下两个不等式：

$$\begin{aligned}x&<(\tilde{y}-y)+y,\\ z_1&<(z'-z)+z<\frac{1}{n}+z<(y-\tilde{y})+z.\end{aligned}$$

将这两个不等式相加，我们得到 $x+z_1<y+z$，矛盾。

(F3) 对任意的 $X\in\mathcal{R}$，我们都有 $\overline{0}+X=X$。

根据 $\overline{0}$ 的定义，$\overline{0}+X=\{x+(-a)|a\in\mathbb{Q}_{>0}\}$。假设 $x-a\in\overline{0}+X$ 是任意给定的一个元素，其中 $x\in X$，$a\in\mathbb{Q}_{>0}$，那么，根据 $x-a<x$，我们知道 $x-a\in X$，这说明 $\overline{0}+X\subset X$；假设

$x \in X$ 是任意给定的一个元素，那么一定存在 $x' \in X$，使得 $x > x'$，我们令 $a = x' - x \in \mathbb{Q}_{>0}$，从而，$x = x' + (-a)$，这个分解表明 $x \in \overline{0} + X$，所以 $X \subset \overline{0} + X$。

(F4) 对任意的 Dedekind 分割 X 和 Y，存在唯一的 $Z \in \mathcal{R}$，使得 $X + Z = Y$。$Y = \overline{0}$ 的情况就是公理 (F5)。

$X, Y \in \mathcal{R}$ 给定，我们按照下面的方式定义 Z：

$$Z = \left\{ y - x' \middle| y \in Y, x' \in X' \right\}.$$

我们首先要说明 $Z \in \mathcal{R}$，其推理与之前验证 $X + Y \in \mathcal{R}$ 的方式类似：

- $Z \neq \emptyset$，$Z' \neq \emptyset$。

  Z 显然不是空集；为了说明 Z' 不是空集，我们选定 $y_0' \in Y, x_0 \in X$，只要说明 $y_0' - x_0 \notin Z$ 即可：如若不然，存在 $y \in Y$ 和 $x' \in X'$，使得 $y - x' = y_0' - x_0$，即 $y + x_0 = y_0' + x'$，这与 $y < y_0'$ 以及 $x_0 < x'$ 矛盾。

- 对任意的 $y - x' \in Z$，其中 $y \in Y, x' \in X'$，如果 $z < y - x'$，就一定有 $z \in Z$。

  我们只需要把 z 分解为 $z = y - (y - z)$，根据 $z < y - x'$，我们知道 $y - z > x'$，所以 $y - z \in X'$，所以 z 可以写成 Y 中元素与 X' 中元素的差。

- Z 中的元素没有最大元。

  证明是平凡的。

现在来说明 $X + Z = Y$。

由于 $X + Z$ 中的元素形如 $x + (y - x')$，其中 $x \in X$，$y \in Y$，$x' \in X$，根据 $x < x'$，所以这个元素必然 $< y$，这就说明该元素落在 Y 中，所以 $X + Z \subset Y$；对任意 y，由于 Y 没有最大元，我们可以选取 $\tilde{y} \in Y$，使得 $y < \tilde{y}$。根据上面的技术性引理，我们还可以选取 $x \in X$ 和 $x' \in X'$，使得 $x' - x < \tilde{y} - y$，那么 $\tilde{\tilde{y}} = y + (x' - x) < y$，所以 $\tilde{\tilde{y}} \in Y$。从而，

$$y = \underbrace{x}_{\in X} + \underbrace{\tilde{\tilde{y}} - x'}_{\in Z} \in X + Z$$

这说明 $Y \subset X + Z$。综合上述，$X + Z = Y$。

我们再进一步研究刚才所以定义的加法的结构。注意到，根据刚刚证明的 (F4)，我们可以定义相反数的运算：

$$- : \mathcal{R} \to \mathcal{R}, X \mapsto -X.$$

其中 $-X$ 是使得 $X + Z = \overline{0}$ 的（唯一的）那一个 Z。按照定义，我们有 $-X = \{y - x' | y \in \overline{0}, x \in X'\}$。根据公理 (F1)-(F4)（目前已经证明），这个 Z 是唯一的并且 $-(-X) = X$。我们还可以说明，这里定义的负号运算和本来有理数上的负号的运算是一致的，即若 $\dfrac{p}{q}$ 是有理数，那么 $X_{-\frac{p}{q}} = -X_{\frac{p}{q}}$：为此，我们只需要说明 $X_{-\frac{p}{q}} + X_{\frac{p}{q}} = \overline{0}$ 即可，我们把验证细节的乐趣以及下面的练习题一并留给同学们：

练习 (正负性). 对于 $X \in \mathcal{R}$，如果 $X > \overline{0}$，我们就称 X 是正的；如果 $X < \overline{0}$，我们就称 X 是负的。

1) 证明，X 是正的当且仅当 X 中有正的有理数。

2) 证明，X 是正的当且仅当 $-X$ 是负的。

利用这个习题的第一个结论，我们现在可以证明 Archimedes 公理 **(A)**，即如果 Dedekind 分割 $X > \overline{0}$，那么对任意的 $Y \in \mathcal{R}$，存在正整数 n，使得

$$\underbrace{X + \cdots + X}_{n\text{ 个}} > Y.$$

实际上，根据上面习题的结论，我们先选一个有理数 $x \in X$，并且 $x > 0$。我们任意选取一个有理数 $y' \in Y$。我们当然可以找到 n，使得 $n \cdot x > y'$，此时，$n \cdot x \in \underbrace{X + \cdots + X}_{n\text{ 个}}$ 并且比 Y' 中的一个元素还要大，这说明 $\underbrace{X + \cdots + X}_{n\text{ 个}} > Y$。

乘法结构

我们现在要来定义并研究 $\mathcal{R}$ 上的乘法。为了定义乘法，对于 $X, Y \in \mathcal{R}$，我们分情况来讨论：

$$X \cdot Y = \begin{cases} \overline{0}, & \text{如果 } X = \overline{0} \text{ 或 } Y = \overline{0}; \\ \{x \cdot y | x \geqslant 0, x \in X, y \geqslant 0, y \in Y\} \cup \overline{0}, & \text{如果 } X > 0,\ Y > 0; \\ -\big(X \cdot (-Y)\big), & \text{如果 } X > 0,\ Y < 0; \\ -\big((-X) \cdot Y\big), & \text{如果 } X < 0,\ Y > 0; \\ \big((-X) \cdot (-Y)\big), & \text{如果 } X < 0,\ Y < 0. \end{cases}$$

我们希望（将证明）乘法单位元 $\overline{1}$ 恰好就是

$$\overline{1} = X_1 = \{x \in \mathbb{Q} | x < 1\}.$$

在第一次作业题中，我们将逐条验证域公理 **(F)** 和序公理 **(O)** 所剩下的部分。至此，我们证明了 $(\mathcal{R}, +, \cdot, \leqslant)$ 满足实数的所有公理，从而证明了实数的存在性。关于实数的唯一性，我们将会在后面课程的展开中顺便讨论（比如说在什么意义下是唯一的）。

3.1 作业：可数与不可数，Schroeder-Bernstein 定理

数学分析一作业 1

请用 A4 大小的纸张**正反面**用钢笔，签字笔或者圆珠笔书写，并注明自己的姓名，年级和作业的**总页数**。除定理公式所涉及的人名之外，请使用**中文**。本次作业的提交时间和地点为 **9 月 26 日**上午的课堂上，逾期视作零分。

基本习题

习题 A：课堂内容的补充

A1) 假设非空集合 $X \subset \mathbb{R}$ 有上界并且实数 M 是 X 的上界。证明，如下两个命题等价：

- $M = \sup X$。
- 对任意的 $\varepsilon > 0$，都存在 $x \in X$，使得 $x > M - \varepsilon$。

A2) 证明，每个非空开区间都包含无限多个有理数。

A3) (X, d) 是距离空间，$Y \subset X$。我们定义 Y 上的距离函数:

$$d_Y : Y \times Y \to \mathbb{R}, (y_1, y_2) \mapsto d_Y(y_1, y_2) = d(y_1, y_2).$$

证明，d_Y 是 Y 上的距离函数，从而 (Y, d_Y) 是度量空间。我们称 d_Y 是 d 在 Y 上的**诱导度量**，(Y, d_Y) 称作是 (X, d) 的**子（度量）空间**。（提示：直接验证定义）

A4) $\mathbb{R}^n = \underbrace{\mathbb{R} \times \mathbb{R} \times \cdots \times \mathbb{R}}_{n\text{个}} = \left\{(x_1, \cdots, x_n) \middle| x_1 \in \mathbb{R}, \cdots, x_n \in \mathbb{R}\right\}$。对于 $x, y \in \mathbb{R}^n$，我们定义

$$d(x, y) = \sqrt{(x_1 - y_1)^2 + \cdots + (x_n - y_n)^2}, \quad x = (x_1, \cdots, x_n),\ y = (y_1, \cdots, y_n).$$

证明，$(\mathbb{R}^n, d)$ 是度量空间。（我们默认开方运算和中学的一致，尽管我们现在没有定义开方运算）

A5)（重要）给定距离空间 (X, d)，$Y \subset X$ 是子集。如果对任意的 $x \in X$ 和任意的 $\varepsilon > 0$，都存在 $y \in Y$，使得 $d(y, x) < \varepsilon$，我们就称 Y 在 X 中是**稠密的**。证明，有理数在 $\mathbb{R}$ 中（距离函数由两个数的差的绝对值定义）是稠密的。

A6) 对于 $(x, y) \in \mathbb{R}^2$，如果它的坐标 x 和 y 都是有理数，我们就称这个点是有理点。证明，$(\mathbb{R}^2, d)$（参见习题 A4)）中的有理点是稠密的。

A7) 证明，假定域公理 **(F)** 和序公理 **(O)**，确界原理可以推出 Archimedes 公理 **(A)**。

A8)（无理数的存在性）令 $X = \left\{x \in \middle| x^2 \leqslant 2\right\}$，这是一个有界的集合。令 $\sqrt{2} = \sup X$。证明，$\sqrt{2}$ 不是有理数。

A9) 证明，每个开区间总包含无限多个无理数。

习题 B：可数集和不可数集

（鼓励同学们查找图书或者资料来完成这一部分）令 $\mathbb{N}$ 表示自然数的集合（包括 0）。X 是一个集合，如果存在单射 $f : X \to \mathbb{N}$，我们就称 X 是**可数的**。如果 X 不是可数的，我们就称它是**不可数的**。

B1) 证明，有限集是可数的。

B2) 证明，可数集合的子集是可数的。

B3) 证明，如果 X 是可数集，那么我们总可以将 X 写成 $X = \{x_1, x_2, x_3, \cdots\}$（即可以把 X 中的元素用自然数来标号）。（我们可以从一个开始一个一个的数下去把这些元素都罗列出来，所以叫做可数集）

B4) 证明，有理数 $\mathbb{Q}$ 是可数集。

B5) 证明，可数个可数集合的并集也是可数集，也就是说，如果 $X_1, X_2, \cdots, X_n, \cdots$ 都是可数集合，那么它们的并集 $\bigcup_{n\geqslant 1} X_n$ 也是可数集合。（提示：这是一个需要记住的经典证明，请查阅参考书或者网络）

B6) X 是可数的，映射 $f : X \to Y$ 是满射。证明，Y 是可数的。

B7) 按照以下步骤证明 $\mathbb{R}$ 是不可数的（同学们可以查阅一个用所谓对角线法则的经典证明，我们的证明基于区间套原理）：

B7-1) $J \subset \mathbb{R}$ 是闭区间并且它的长度 $|J| > 0$。证明，对任意的 $x \in \mathbb{R}$，总存在闭区间，使得 $I \subset J$，$|I| > 0$ 且 $x \notin I$。

B7-2) 证明，如果 $\{x_1, x_2, \cdots\}$ 是 $\mathbb{R}$ 的一个可数子集，那么存在闭区间套 $I_1 \supset I_2 \supset \cdots$，使得对任意的 n，$x_n \notin I_n$。

B7-3) 证明，$\mathbb{R}$ 不可数。

B8) 证明，如果 X 是不可数集，A 是 X 的可数子集，那么 $X - A$ 是不可数的。

B9) 证明，任意的长度不为 0 的区间（无论开或闭）都是不可数的。

B10) 证明，复数 $\mathbb{C}$ 是不可数的。

B11) 假设 $\mathcal{I}$ 是 $\mathbb{R}$ 上的某些开区间组成的集合，它满足如下性质：对任意的 $I, J \in \mathcal{I}$，$I \neq J$，那么它们的交集是空集，即 $I \cap J = \emptyset$。证明，$\mathcal{I}$ 是可数集。

课外知识：一个集合论的补充

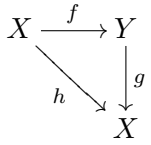
习题 C：Schroeder-Bernstein 定理

假设 X 和 Y 是两个集合，映射 $f: X \to Y$ 和 $g: Y \to X$ 都是单射。我们令 $X' = X - g(Y)$。

C1) 如果 X 是有限集，证明，存在 $\varphi: X \to Y$，使得 φ 是双射。

C2) 如果 X 是可数集，证明，存在 $\varphi: X \to Y$，使得 φ 是双射。

从现在开始，对 X 不加任何的限制。我们令 $h: X \to X$ 是复合映射 $h = g \circ f$:



C3) 考虑 X 的子集的集合 $\mathcal{F} = \{A \subset X | X' \cup h(A) \subset A\}$。证明，$\mathcal{F}$ 是非空。

C4) 证明，如果 $A \in \mathcal{F}$，那么 $X' \cup h(A) \in \mathcal{F}$。

C5) 我们定义

$$A_0 = \bigcap_{A \in \mathcal{F}} A = \{x \in X | \text{对任意的 } A \in \mathcal{F}, \text{ 都有 } x \in A\}.$$

证明，$A_0 \in \mathcal{F}$。

C6) 证明，$X' \cup h(A_0) = A_0$。

C7) 令 $B_0 = X - A_0$。证明，$f(A_0) \cap g^{-1}(B_0) = \emptyset$ 并且 $f(A_0) \cup g^{-1}(B_0) = Y$。

C8) 我们定义映射 $\varphi: X \to Y$：对于 $x \in X$，我们要求

$$\varphi(x) = \begin{cases} f(x), & \text{如果} x \in A_0; \\ g^{-1}(x), & \text{如果} x \in B_0. \end{cases}$$

证明，这是双射。

根据上述，我们证明了

定理 (Schroeder-Bernstein)**.** 如果有单射 $f: X \to Y$ 和单射 $g: Y \to X$，那么存在着两个集合之间的双射 $\varphi: X \to Y$。

习题 D：Dedekind 分割证明的细节

这一部分习题的目标是完成课堂上关于 Dedekind 分割乘法结构的部分，从而给出了实数域的构造完整证明。

D1) 证明，如果 X 和 Y 都是 Dedekind 分割，那么讲义中定义的 $X \cdot Y$ 也是 Dedekind 分割，即乘法

$$\mathbb{R} \times \mathbb{R} \xrightarrow{\cdot} \mathbb{R},\ (X, Y) \mapsto X \cdot Y,$$

是良好定义的。（提示：只要对 $X > 0$，$Y > 0$ 这一种情形证明即可，当然你需要用到课上的练习和结论）

D2) 证明，$(X \cdot Y) \cdot Z = X \cdot (Y \cdot Z)$。（⇒(F5)）

D3) 证明，$X \cdot Y = Y \cdot X$。（⇒(F6)）

D4) 证明，$X \cdot (Y + Z) = X \cdot Y + X \cdot Z$。（⇒(F9)）（提示：略难）

D5) 证明，$\overline{1} \cdot X = X$ 并且 $\overline{1} \neq \overline{0}$。（⇒(F7)）

D6) 证明，如果 $X \cdot Y = \overline{0}$，那么 $X = \overline{0}$ 或者 $Y = \overline{0}$；再证明，如果 $X \geqslant \overline{0}$，$Y \geqslant \overline{0}$，那么 $X \cdot Y \geqslant \overline{0}$。（⇒(O5)）

D7) （技术工具）X 是一个正的 Dedekind 分割。证明，对任意的正整数 n，总存在 $x \in X$，$x' \in X'$，使得

$$1 < \frac{x'}{x} < 1 + \frac{1}{n}.$$

D8) 证明，对任意的 Dedekind 分割 X 和 Y，其中 $Y \neq \overline{0}$，存在唯一的 Dedekind 分割 Z，使得

$$Y \cdot Z = X.$$

我们将 Z 记作 $\dfrac{X}{Y}$。当 $X = \overline{1}$ 时，我们也将它记作 Y^{-1}。（⇒(F8)）

D9) 请检验，我们已经完整的证明了域公理 **(F)** 和序公理 **(O)**。

4 极限，级数，Cauchy 列，距离空间中的收敛

二零一九年九月十九日，星期四，晴天

数列和点列的极限

我们通常用 $\{x_n\}_{n\geqslant 1}$ 来表示一列数（有顺序）$x_1, x_2, \cdots, x_m, \cdots$ 并且将它称为是**数列**。一个数列实际上就是一个映射

$$f : \mathbb{Z}_{\geqslant 1} \to \mathbb{R}, \quad k \mapsto f(k) = a_k.$$

同样的，给定一个映射

$$f : \mathbb{Z}_{\geqslant 1} \to X, \quad k \mapsto f(k) = x_k,$$

其中 (X, d) 是一个距离空间，我们就得到了距离空间中的一个**点列**。如果把 $\mathbb{R}$ 看成是距离空间的话（我们之前已经这样做了），那么数列只是点列的一个特殊情况。

定义 10 (极限的定义，$\varepsilon - N$ 语言)**.** 假设 $\{x_n\}_{n\geqslant 1}$ 是实数序列。如果存在 $x \in \mathbb{R}$，使得对任意的 $\varepsilon > 0$，总存在 $N \geqslant 1$，使得对任意的 $n \geqslant N$，我们都有

$$|x_n - x| < \varepsilon,$$

那么，我们就说数列 $\{x_n\}_{n\geqslant 1}$ **有极限**并把 x 称作是数列 $\{x_n\}$ 的**极限**，记作 $\lim\limits_{n\to\infty} x_n = x$。如果数列 $\{x_n\}$ 有极限，我们还说它是**收敛的**。

注记. 我们直观上总是认为 ε 是非常小的数，N 是非常大的数

1) 上述定义中的 N 是在 ε（任意的）选定之后再选择的，它通常依赖于 ε 的大小，我们有时候也写成 $N = N(\varepsilon)$ 来表示这种依赖性。

2) 上述定义就是想描叙下面的直观：无论 ε 有多么小，总存在很大的 N，使得从数列的第 N 项开始，数列的每个数都和 x 的误差不超过 ε（离着 x 很近）。

把数列的极限的概念推广到距离空间是轻而易举的事情：

定义 11 (度量空间中点列的极限)**.** $\{x_n\}_{n\geqslant 1}$ 是度量空间 (X, d) 中点的序列，如果存在 $x \in X$，使得对任意的 $\varepsilon > 0$，总存在 $N \geqslant 1$，使得对任意的 $n \geqslant N$，我们都有

$$d(x_n, x) < \varepsilon,$$

那么就称点列 $\{x_n\}$ 在 (X, d) 中有**极限**的，x 称作是它的**极限**，记为 $\lim\limits_{n\to\infty} x_n = x$。

为了了解极限的概念，我们从一些重要的例子入手：

例子. 1) 这是一个反面的例子。极限是否存在将是我们的核心课题，下面的例子可以表明这个问题的重要性与困难之处：我们考虑实数数列 $\{x_n\}_{n\geqslant 1}$，其中

$$x_n = 1 - \frac{1}{3} + \frac{1}{5} - \cdots + \frac{(-1)^{n-1}}{2n-1},$$

- 我们想知道这个序列是不是有极限。

  然而，根据极限的定义，如果有极限，我们要先验地知道极限是什么才可以利用定义来证明。事实上，要想猜到这个极限的值是很困难的。通过这个例子，我们发现极限的定义对于判断极限是否存并没有太大的帮助。

- 从概念上而言，极限的定义在哲学意义上有一个先天的缺陷：判断数列收敛与否，应该由数列本身所决定而不需要依赖于外在的信息，比如说事先知道极限是什么。换句话说，我们想要**内蕴地**来考虑极限问题。这种看法和思考方式对数学的学习是大有裨益的。

- 假设我们已经得知上述数列的极限是 $\frac{\pi}{4}$（注意，目前我们还没有定义圆周率 π）。然而，要尝试从定义出发来证明极限就是 $\frac{\pi}{4}$，我们目前还是无能为力，因为我们缺乏基本的工具。我们将学习微分和积分等基本的分析工具，有了这些技术手段的支持，我们才可以真正地计算类似的极限。

- 让我们暂且接受 π 是无理数这个事实。我们注意到，这个数列的每一项都是有理数，但是它在有理数 $\mathbb{Q}$ 中是没有极限的。我们知道，$(\mathbb{Q}, d(x,y)=|x-y|)$ 也是一个距离空间（参见第一次作业题的 A3)）可见，一个点的序列能否收敛与这个点列所生活的空间的性质密切相关。

- 题外话：我们可以把上述极限写为

$$\lim_{n\to\infty}\left(1-\frac{1}{3}+\frac{1}{5}-\cdots+\frac{(-1)^{n-1}}{2n-1}\right)=\frac{\pi}{4}.$$

  等式左边每一项都是用算术来定义的，即通过整数进行加减乘除而来。然而，通过取极限的过程，右边得到了和圆周率有关的信息。简而言之，一个算术的对象和一个几何的对象通过极限联系在一起：可以不夸张地说，极限是从算术世界通往几何世界的秘密通道。

2) 如果 $x_n = x$，那么 $\lim\limits_{n\to\infty} x_n = x$。

我们按照定义来证明：对任意的 $\varepsilon > 0$，我们取 $N = 1$，当 $n \geqslant 1$ 时，有 $|x_n - x| = 0 < \varepsilon$。

3) 我们有极限 $\lim\limits_{n\to\infty}\frac{1}{n} = 0$。我们按照定义来证明：对任意的 $\varepsilon > 0$，可以选取 N，使得 $N > \frac{1}{\varepsilon}$，比如说，$N = \lfloor\frac{1}{\varepsilon}\rfloor + 1$。当 $n \geqslant N$ 时，有 $|x_n - 0| = \frac{1}{n} < \frac{1}{N} < \varepsilon$。

4) 我们有极限 $\lim\limits_{n\to\infty}\frac{1}{\lambda^n} = 0$，其中 $\lambda > 1$。

先陈述一个技术事实：假设 $\lambda = 1 + \delta$，其中，$\delta > 0$。根据二项式的展开（只需要乘法和加法以及各类实数的四则运算就可以证明），我们有

$$\lambda^n = (1+\delta)^n = 1 + n\delta + \cdots \geqslant 1 + n\delta.$$

所以，对任意的一个数 M，总能选到 n，使得 $\lambda^n > N$（*Archimedes* 原理）。

我们现在按照定义来证明：对任意的 $\varepsilon>0$，可以选取 N，使得 $\lambda^N>\dfrac{1}{\varepsilon}$，比如说，$N=\lfloor\dfrac{1}{\varepsilon\delta}\rfloor+1$。当 $n\geqslant N$ 时，有 $|x_n-0|=\dfrac{1}{\lambda^n}<\dfrac{1}{\lambda^N}<\varepsilon$。

有了极限，我们可以定义无限个数求和这一个概念，这也就是级数的概念。不夸张地说，这是极限最重要的一个应用，因为几乎所有有意义的数和函数都是通过级数的方式来构造的。另外，我们还可以仿照极限的定义方式，研究极限是 $+\infty$ 或者 $-\infty$ 的数列。为此，我们对极限的定义进行扩充

定义 12 (极限定义的补充)**.** 我们假定 $\{x_n\}_{n\geqslant1}$ 是实数的序列。

1) 如果对任意的 $M>0$（很大），总存在 $N>0$，使得对任意的 $n\geqslant N$，都有 $x_n>M$，我们就称 x_n **收敛到正无穷**，记作 $\lim\limits_{n\to\infty}x_n=+\infty$。类似地，如果对任意的 $M>0$，总存在 $N>0$，使得对任意的 $n\geqslant N$，都有 $x_n<-M$，我们就称 x_n **收敛到负无穷**，记作 $\lim\limits_{n\to\infty}x_n=-\infty$。如果上述之一发生，我们就称 x_n 是**发散的**。

2) $a_1,a_2,\cdots$ 是一列实数，令 $x_n=\sum\limits_{i=1}^{n}a_i$，如果 $\{x_n\}_{n\geqslant1}$ 有极限，我们就说级数 $a_1+a_2+a_3+\cdots$ **收敛**并把它的极限 $\lim\limits_{n\to\infty}x_n$ 记作 $\sum\limits_{i=1}^{\infty}a_i$；如果 $\{x_n\}_{n\geqslant1}$ 是发散的，我们就称级数 $a_1+a_2+a_3+\cdots$ **发散**。根据之前的定义，我们可以想当然地定义 $\sum\limits_{i=1}^{\infty}a_i=+\infty$ 或者 $\sum\limits_{i=1}^{\infty}a_i=+\infty$ 并称这个级数是收敛到正无穷或者负无穷的。

注记. 为了理解这几个定义，我们再给出几个简单的例子：

1) 一个数列既没有极限也不发散，比如说 $\{x_n\}_{n\geqslant1}$，其中 $x_n=(-1)^{n-1}$。这个数列是没有极限的：如若不然，假设它的极限是 x。根据定义，对于 $\varepsilon=0.1$（因为对与任意的 ε 都成立，特别地，我们就取这个数），存在 N_0，使得当 $n\geqslant N_0$ 时，$|x_n-x|<0.1$。所以，根据三角不等式，

$$2=|x_{N_0}-x_{N_0+1}|\leqslant|x_{N_0}-x|+|x-x_{N_0+1}|<0.1+0.1=0.2,$$

矛盾。

2) 在度量空间中，我们有收敛的概念，但是我们一般而言并没有级数的概念，原因是对于点列 $\{a_k\}_{k\geqslant1}$，求和 $a_1+a_2+\cdots+a_n$ 是不能被定义的（因为我们没有加法运算 $X\times X\to X$）。根据这个讨论，如果 X 还是一个线性空间，那么，我们可以定义级数 $a_1+a_2+\cdots+a_n$。比如说，$X=\mathbb{C}$ 是复数，我们就可以定义复数的级数。

3) 调和级数 $1+\dfrac{1}{2}+\dfrac{1}{3}+\cdots+\dfrac{1}{n}+\cdots$ 是最经典的发散级数的例子。

调和级数是发散的证明分析中最经典的证明之一：考虑 $n=2^k$，那么，我们有

$$\sum_{j=1}^{n-1}\frac{1}{j}=\underbrace{\frac{1}{1}}_{1\text{个，}\geqslant\frac{1}{2}}+\underbrace{\left(\frac{1}{2}+\frac{1}{3}\right)}_{2\text{个，}\geqslant\frac{1}{4}}+\underbrace{\left(\frac{1}{4}+\frac{1}{5}+\frac{1}{6}+\frac{1}{7}\right)}_{4\text{个，}\geqslant\frac{1}{8}}+\cdots+\underbrace{\left(\frac{1}{2^{k-1}}+\frac{1}{2^{k-1}+1}+\cdot+\frac{1}{2^k-1}\right)}_{2^{k-1}\text{个，}\geqslant\frac{1}{2^k}}$$
$$>1\times\frac{1}{2}+2\times\frac{1}{4}+4\times\frac{1}{8}+\cdots+2^{k-1}\times\frac{1}{2^k}=\frac{k}{2}.$$

所以，对任意的 M，我们总可以选取很大的 n，使得 $a_n\geqslant M$，这说明调和级数是发散的。我们发现证明从技术上仍然和庄子的二分法相似，究其根本，是因为二分之后求和变得容易计算。可计算性将在很多数学问题中都是头等重要的事情，我们后面会数次遇到类似地事情。

我们收集与极限定义相关的一些简单性质，我们对距离空间 (X,d) 中的点列来陈述这些性质，它们自然的对实数的序列也成立：

命题 13. 给定距离空间 (X,d)，$\{x_n\}_{n\geqslant1}$ 是 X 中点的序列，我们有

1) 如果 $\{x_n\}_{n\geqslant1}$ 收敛，那么它的极限唯一，即若 $\lim\limits_{n\to\infty}x_n=x$ 并且 $\lim\limits_{n\to\infty}x_n=y$，那么 $x=y$。

2) 任意改变序列中有限项之后，不会改变序列的极限，即若 $\lim\limits_{n\to\infty}x_n=x$，$\{y_n\}_{n\geqslant1}$ 是另外一个点列，其中当 $n\geqslant N_0$ 时，$y_n=x_n$，那么 $\{y_n\}_{n\geqslant1}$ 也收敛并且 $\lim\limits_{n\to\infty}y_n=x$。

3) 假设 $1\leqslant n_1<n_2<\cdots<n_k<\cdots$ 是一列上升的指标，对 $k\geqslant1$，$y_k=x_{n_k}$，那么点列 $\{y_k\}_{k\geqslant1}$ 称作是 $\{x_n\}$ 的**子列**，子列也经常写成 $\{x_{n_k}\}_{k\geqslant1}$。一个收敛点列的子列也收敛到同一个极限，即若序列 $\{x_n\}_{n\geqslant1}$ 是收敛的，那么子列 $\{x_{n_k}\}_{k\geqslant1}$ 也收敛并且 $\lim\limits_{n\to\infty}x_n=\lim\limits_{k\to\infty}x_{n_k}$。

4) 如果实数数列 $\{x_n\}_{n\geqslant1}$ 是收敛的，那么它有界的。(在距离空间中我们也可以定义有界性，但是这里我们只对实数陈述这个事实)

5) (逆否命题的叙述，重要) 假设 $\{x_n\}_{n\geqslant1}$ 的极限不是 x（可以不收敛），这个命题用数学的语言说就是：存在 $\varepsilon>0$，对任意的 $N>0$，总存在 $n>N$，使得 $d(x_n,x)\geqslant\varepsilon$。

证明：我们逐条的进行论证。这些证明是简单的，但是对于第一次接触极限的同学，应该坚持把证明写完整以熟悉这种语言。

1) 任意给定 $\varepsilon>0$，根据极限的定义，存在 $N_1\geqslant1$，使得对任意的 $n\geqslant N_1$，我们都有 $d(x_n,x)<\varepsilon$，也存在 $N_2\geqslant1$，使得对任意的 $n\geqslant N_2$，我们都有 $d(x_n,y)<\varepsilon$。那么，对于 $n_0=\max(N_1,N_2)$，我们有 $d(x_{n_0},x)<\varepsilon$ 和 $d(x_{n_0},y)<\varepsilon$ 同时成立。从而，根据三角不等式，我们有

$$d(x,y)<d(x_{n_0},x)+d(x_{n_0},y)<2\varepsilon.$$

由于 ε 是任意选取的，所以 $d(x,y)=0$，所以 $x=y$。

2) 对任意给定的 $\varepsilon > 0$，我们想要找一个 $N \geqslant 1$，使得对任意的 $n \geqslant N$，我们都有 $d(y_n, x) < \varepsilon$。根据极限的定义，对于这个 ε，存在 $N_1 \geqslant 1$，使得对任意的 $n \geqslant N_1$，我们都有 $d(x_n, x) < \varepsilon$。我们只需要令 $N = \max(N_0, N_1)$，此时，当 $n \geqslant N$ 时，我们有

$$d(y_n, x) = d(x_n, x) < \varepsilon.$$

3) 我们将 $\{x_n\}_{n\geqslant 1}$ 的极限记做 x，为了证明 $\lim_{k\to\infty} x_{n_k} = x$，我们任选 $\varepsilon > 0$。根据 $\{x_n\}_{n\geqslant 1}$ 的极限的定义，存在 $N \geqslant 1$，使得对任意的 $n \geqslant N$，$d(x_n, x) < \varepsilon$。那么，对任意的 $k \geqslant N$，根据 $n_k \geqslant k \geqslant N$，我们知道 $d(x_{n_k}, x) < \varepsilon$，从而 $\lim_{k\to\infty} x_{n_k} = x$。

4) 假设 $\lim\limits_{n\to\infty} x_n = x$，根据定义，对 $\varepsilon = 1$，存在 N，使得当 $n \geqslant N$ 时，我们有 $|x_n - x| \leqslant 1$。特别地，对一切满足 $n \geqslant N$ 的指标 n，x_n 都落在区间 $[x-1, x+1]$ 中，这是有界的。对于 $n = 1, 2, \cdots, N-1$，这是有限多个数，它们自然有界，所以数列本身有界。

最后一个命题 5) 是基本的（但是别扭）逻辑。 □

注记. 对于 4)，我们注意到如果一个数列是有界的，那么它不一定是收敛的，比如说 $x_n = (-1)^n$ 所给出的数列。但是，（后面会证明）如果一个数列是有界的，那么它一定包含着收敛的子列。

极限是我们在高等数学学习中所最先接触的几个数学对象之一。在数学中，对于每一个数学对象（例如极限），我们会例行公事般地考虑它的一些常见的性质。比如说，这个对象最基本的例子是什么，这种对象是否存在，如果存在的话它是否具有唯一性，它的子对象和商对象（如果有的话）都具有什么性质（比如说遗传了原来的对象的什么性质），这个对象的可计算性以及在特定映射下的行为等等。作为例子，我们刚刚见到一个点列的子对象（即子列）遗传了点列的收敛性（和有界性）。尽管这是一种八股文一般的讨论方式（*Bourbaki* 学派是这种方式最忠实的实践者），但是是非常有效率的一种学习和记忆方式，我们在课程上会尽量的按照这种习惯来学习。特别要强调的是，每一个定义大家都应该搞清楚最基本的例子是什么。

在实数 $\mathbb{R}$ 上，我们有四则运算和序关系，我们自然要研究它们与极限的关系

命题 14 (四则运算与序关系的交换性)**.** 假设 $\{x_n\}_{n\geqslant 1}$ 和 $\{y_n\}_{n\geqslant 1}$ 是实数数列，那么我们有

1) 数列 $\{x_n + y_n\}_{n\geqslant 1}$ 收敛并且 $\lim\limits_{n\to\infty}(x_n + y_n) = \lim\limits_{n\to\infty} x_n + \lim\limits_{n\to\infty} y_n$。

2) 数列 $\{x_n - y_n\}_{n\geqslant 1}$ 收敛并且 $\lim\limits_{n\to\infty}(x_n - y_n) = \lim\limits_{n\to\infty} x_n - \lim\limits_{n\to\infty} y_n$。

3) $\{x_n \cdot y_n\}_{n\geqslant 1}$ 收敛并且 $\lim\limits_{n\to\infty}(x_n \cdot y_n) = \lim\limits_{n\to\infty} x_n \cdot \lim\limits_{n\to\infty} y_n$。特别地，对任意实数 $\lambda \in \mathbb{R}$，$\lim\limits_{n\to\infty} \lambda \cdot x_n \cdot y_n = \lambda \cdot \lim\limits_{n\to\infty} x_n$。

4) 如果 $\lim\limits_{n\to\infty} y_n \neq 0$，那么数列 $\{\dfrac{x_n}{y_n}\}_{n\geqslant N}$ 收敛（$\lim\limits_{n\to\infty} y_n \neq 0$ 表明存在 N，使得当 $n \geqslant N$ 时，$y_n \neq 0$）并且 $\lim\limits_{n\to\infty} \dfrac{x_n}{y_n} = \dfrac{\lim\limits_{n\to\infty} x_n}{\lim\limits_{n\to\infty} y_n}$。

5) 如果对足够大的 n（即存在 N，使得 $n \geqslant N$ 时），有 $x_n \leqslant y_n$，那么 $\lim\limits_{n\to\infty} x_n \leqslant \lim\limits_{n\to\infty} y_n$。

证明： 这几个命题的证明不困难，但是对初学者而言，细节的训练是重要的。我们假设 $\lim\limits_{n\to\infty} x_n = x$，$\lim\limits_{n\to\infty} y_n = y$。

1) 任意给定 $\varepsilon > 0$，我们要证明存在 N，使得对所有的 $n \geqslant N$，我们都有 $|(x_n+y_n)-(x+y)| < \varepsilon$。因为 $x_n \to x$，对于正数 $\dfrac{\varepsilon}{2}$ 而言，所以存在 N_1，使得对所有的 $n \geqslant N_1$，我们都有 $|x_n-x| < \dfrac{1}{2}\varepsilon$；因为 $y_n \to y$，所以存在 N_2，使得对所有的 $n \geqslant N_1$，我们都有 $|y_n - y| < \dfrac{1}{2}\varepsilon$。此时，我们可以选取 $N = \max(N_1, N_2)$，所以当 $n \geqslant N$ 时，我们有

$$|(x_n + y_n) - (x + y)| \leqslant |x_n - x| + |y_n - y| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

2) 把上面证明中的适当位置的加号换成减号即可。

3) 由于两个数列都收敛，所以它们都是有界的，所以存在一个常数 C，使得对任意的 $n \geqslant 1$，我们都有 $|x_n| \leqslant C$，$|y_n| \leqslant C$。为了证明极限存在，我们先做计算

$$\begin{aligned} |x_n y_n - xy| &= |x_n(y_n - y) + (x_n - x)y| \leqslant |x_n||y_n - y| + |x_n - x||y| \\ &\leqslant C\big(|y_n - y| + |x_n - x|\big). \end{aligned}$$

任意给定 ε，我们可以选取 N，使得当 $n \geqslant N$ 时，$|x_n - x| < \varepsilon$ 和 $|y_n - y| < \varepsilon$ 同时成立，此时，对于 $n \geqslant N$，我们就有

$$|x_n y_n - xy| \leqslant C(\varepsilon + \varepsilon) = 2C\varepsilon.$$

由于 ε 可以任意选取，命题就得到了证明（实际上，为了把证明写的和定义完全一致，我们应该回头把原来的 ε 替换成 $\dfrac{\varepsilon}{2C}$）。

4) 这一条性质的证明我们留成作业。

5) 我们可以利用反证法：如若不然，我们有 $x > y$，选取 $\varepsilon = \dfrac{|x - y|}{2}$，根据极限的定义，我们可以选取 N，使得当 $n \geqslant N$ 时，$|x_n - x| < \varepsilon$ 和 $|y_n - y| < \varepsilon$ 同时成立，从而对于任意一个这样的 n，我们都有

$$x_n = (x_n - x) + x > x - \varepsilon = \frac{x + y}{2}, \quad y_n = (y_n - y) + y < y + \varepsilon = \frac{x + y}{2},$$

从而，我们得到 $x_n > y_n$，矛盾！

□

注记. *我们对命题本身稍加解释：*

a) 在命题的 5) 中出现的“对足够大的 n”的讲法，指的是“存在 N，使得 $n \geqslant N$ 时”，我们将经常使用这种分析中的“黑话”。

b) 命题 1) 的等号左边说我们先逐项求和然后取极限，右边我们可以先求极限再求和，所以交换了求和和取极限这两种操作之后，我们得到了同样的结果。其他几个命题都是将极限符号和四则运算可以交换。在数学和物理的很多分支中，两种操作的顺序是否可交换的是很重要的，比如说在线性代数里两个矩阵的乘积何时可交换是很核心的话题，时空的弯曲与此有关，量子力学中的 *Heisenberg* 测不准原理也是。

c) 命题的 5) 说的是 lim 与 $\leqslant$ 可以交换。如果我们要求对足够大的 n，有 $x_n < y_n$，那么取极限之后 $x < y$ 未必成立（试举一个反例，举反例是很重要的练习，尤其是接触新概念时，这不失为理解抽象概念的最好方式之一），即严格大小关系和极限不交换。另外，将 $\leqslant$ 换成 $\geqslant$ 结论也成立，课程后面这种平行的命题我们一般只选择证明其中的一条。

d) 在命题的 3) 的证明中，我们不需要区别 ε 和 $2C\varepsilon$ 或者 10000ε，这是因为 ε 可以任意选取。这是数学分析的行规：$\varepsilon \,'\!=\!' \, 10000\varepsilon$。

e) 这个性质最重要的应用是用于计算！通过交换极限和四则运算，很多复杂极限问题可以化成简单的四则运算问题（四则运算我们更熟悉）比说，我们要计算 $\lim\limits_{n\to\infty} \dfrac{n^2+3n+2}{n^2+1}$，我们可以按照下面的步骤来做：

$$\begin{aligned}\lim_{n\to\infty} \frac{n^2+3n+2}{n^2+1} &= \lim_{n\to\infty} \frac{1+\frac{3}{n}+\frac{2}{n^2}}{1+\frac{1}{n^2}} = \frac{\lim\limits_{n\to\infty}\left(1+\frac{3}{n}+\frac{2}{n^2}\right)}{\lim\limits_{n\to\infty}\left(1+\frac{1}{n^2}\right)} \\ &= \frac{\lim\limits_{n\to\infty} 1+\lim\limits_{n\to\infty}\frac{3}{n}+\lim\limits_{n\to\infty}\frac{2}{n^2}}{\lim\limits_{n\to\infty} 1+\lim\limits_{n\to\infty}\frac{1}{n^2}} = \frac{1+0+0}{1+0} = 1.\end{aligned}$$

我们现在回到最本质的一个问题：极限的存在性。我们有一种内蕴的方式来判断极限的存在性：

定理 15 (Cauchy 列与 Caucy 判别准则)**.** 假设 $\{x_n\}_{n\geqslant 1}$ 是实数的数列，那么如下命题等价：

1) $\{x_n\}_{n\geqslant 1}$ 收敛；

2) 对任意的 $\varepsilon > 0$，存在 $N > 0$，使得对任意的 $m, n > N$，都有 $|x_n - x_m| < \varepsilon$。

满足第二条性质的数列被称作是 **Cauchy 列**。

为了证明这个重要的定理，我们先做一些准备工作。如果实数数列 $\{x_n\}_{n\geqslant 1}$ 满足 $x_1 \leqslant x_2 \leqslant x_3 \leqslant \cdots$，就称之为**单调上升的**或者**递增的**；如果它满足 $x_1 < x_2 < x_3 < \cdots$，就称之为**严格单调上升的**或者**严格递增的**。类似地可定义**（严格）单调下降的**或者**（严格）递减的**的序列。

定理 16. 如果单调上升的实数序列 $\{x_n\}_{n\geqslant 1}$ 是有界的，那么 $\{x_n\}_{n\geqslant 1}$ 收敛并且 $\lim\limits_{n\to\infty} x_n = \sup\limits_{n\geqslant 1} x_n$。(单调下降的序列也满足类似的结论)

证明：根据确界原理，我们令 $s=\sup_{n\geqslant 1} x_n$。我们将证明对任意给定的 ε，存在 N，使得当 $n\geqslant N$ 时，$|x_n-s|<\varepsilon$。

在第二课确界原理证明之后的评注中，我们知道可以如下地刻画上确界：对 $\varepsilon>0$，存在 x_{n_0}，使得 $s-\varepsilon<x_{n_0}$。我们选取 $N=n_0$，根据单调性，对任意的 $n\geqslant N=n_0$，我们有 $x_n\geqslant x_{n_0}>x-\varepsilon$。另外，根据上确界的定义，我们还有 $x_n\leqslant x$。所以，对任意的 $n\geqslant N$，都有 $|x_n-x|\leqslant\varepsilon$。证毕。 □

作为应用，我们证明级数 $\displaystyle\sum_{k=1}^{\infty}\frac{1}{k^2}$ 是收敛的：

例子. 级数 $\dfrac{1}{1^2}+\dfrac{1}{2^2}+\dfrac{1}{3^2}+\dfrac{1}{4^2}+\cdots$ 是收敛的。我们定义部分和 $x_n=\displaystyle\sum_{k=1}^{n}\frac{1}{k^2}$，所谓的级数收敛指的就是 $\{x_n\}_{n\geqslant 1}$ 这个数列有极限。这是一个单调上升的数列，根据上面的定理，我们只需要说明 $\{x_n\}_{n\geqslant 1}$ 有界即可：

$$\begin{aligned}
0<x_n&=\frac{1}{1^2}+\frac{1}{2^2}+\frac{1}{3^2}+\frac{1}{4^2}+\cdots+\frac{1}{(n-1)^2}+\frac{1}{n^2}\\
&\leqslant 1+\frac{1}{1\times 2}+\frac{1}{2\times 3}+\frac{1}{3\times 4}+\cdots+\frac{1}{(n-1)\times n}+\frac{1}{n\times(n+1)}\\
&=1+\left(\frac{1}{1}-\frac{1}{2}\right)+\left(\frac{1}{2}-\frac{1}{3}\right)+\left(\frac{1}{3}-\frac{1}{4}\right)+\cdots+\left(\frac{1}{n-1}-\frac{1}{n}\right)+\left(\frac{1}{n}-\frac{1}{n+1}\right)\\
&=2-\frac{1}{n+1}<2.
\end{aligned}$$

5 Cauchy 判别准则，向量序列的收敛，Bolzano-Weierstrass 的列紧性定理，数列和级数收敛的 Cauchy 判断，e 的定义，绝对收敛，收敛的控制判别法

二零一九年九月二十三日，星期一，晴天

我先对前面的课程做适当的补充：

1) 收敛级数的逐项加减法与逐项数乘法：假设实数项（对于复数和 $\mathbb{R}^n$ 取值的级数也成立）的 $\sum\limits_{k=0}^{\infty} a_k$ 和 $\sum\limits_{k=0}^{\infty} b_k$ 收敛，$\lambda \in R$，那么 $\sum\limits_{k=0}^{\infty}(a_k \pm b_k)$ 和 $\sum\limits_{k=0}^{\infty} \lambda \cdot a_k$ 都收敛

$$\sum_{k=0}^{\infty}(a_k \pm b_k) = \sum_{k=0}^{\infty} a_k \pm \sum_{k=0}^{\infty} b_k, \quad \sum_{k=0}^{\infty} \lambda \cdot a_k = \lambda\big(\sum_{k=0}^{\infty} a_k\big).$$

这两个事实的证明是极限四则运算法则的直接应用（用部分和来表示级数）。

2)（向量值序列的极限问题）对于 $\mathbb{R}^n = \big\{(x_1, \cdots, x_n)\big| x_1 \in \mathbb{R}, \cdots, x_n \in \mathbb{R}\big\}$，我们采取如下定义的距离函数

$$d(x, y) = \sqrt{(x_1 - y_1)^2 + \cdots + (x_n - y_n)^2}, \quad x = (x_1, \cdots, x_n),\ y = (y_1, \cdots, y_n).$$

对于复数域 $\mathbb{C}$，我们通过 $z = x + iy$ 这种表示，可以认为 $\mathbb{C} = \mathbb{R}^2$。对于 $z_1, z_2 \in \mathbb{C}$，我们自然有 $d(z_1, z_2) = |z_1 - z_2|$，其中 $|\cdot|$ 是取复数的模长。

假设 $\{x^{(k)}\}_{k\geqslant 1}$ 是 $\mathbb{R}^n$ 中的点列，其中，$x^{(k)} = (x_1^{(k)}, x_2^{(k)}, \cdots, x_n^{(k)})$。那么，$\{x^{(k)}\}_{k\geqslant 1}$ 在 $\mathbb{R}^n$ 中收敛当且仅当它的每个分量都是收敛的实数数列，即对任意的 $i = 1, 2, \cdots, n$，$\{x_i^{(k)}\}_{k\geqslant 1}$ 在 $\mathbb{R}$ 中收敛。

我们把这个性质留在第二次作业题中来证明，请注意这是非常重要的习题，我们会在今后经常考虑在线性空间中收敛的问题。

3) 不同的距离函数的问题。我们之前的课上提到过一个距离空间上可以有几个不同的距离函数，比如在 $\mathbb{R}$ 上，还可以定义 $d(x, y) = 2|x - y|$。对于这个问题的理解可以进一步加深对收敛的理解。

假设 d_1 和 d_2 均为集合 X 上距离函数，如果存在常数 $C_1 > 0$ 和 $C_2 > 0$，使得对任意的 $x, y \in X$，都有

$$d_2(x, y) \leqslant C_1 d_1(x, y), \quad d_1(x, y) \leqslant C_2 d_2(x, y),$$

那么就称这两个距离函数 d_1 和 d_2 是**等价的**。上面的叙述也等价于存在常数 $c > 0$ 和 $C > 0$，使得对任意的 $x, y \in X$，都有

$$c d_1(x, y) \leqslant d_2(x, y) \leqslant C d_1(x, y),$$

比如说，我们在 $\mathbb{R}^n$ 上面可以定义三种距离函数：

$$d_1(x,y)=\sum_{j=1}^{n}|x_j-y_j|,$$
$$d_2(x,y)=\sqrt{\sum_{j=1}^{n}(x_j-y_j)^2},$$
$$d_\infty(x,y)=\sup_{i=1,2,\cdots,n}|x_i-y_i|.$$

这三个距离自然是等价的：

$$nd_\infty(x,y)\geqslant d_1(x,y)\geqslant d_2(x,y)\geqslant d_\infty(x,y).$$

我们还有一个重要的的例子：我们可以把 $n\times n$ 的矩阵的全体 $\mathbf{M}_n(\mathbb{R})$ 视作是 $\mathbb{R}^{n^2}$（加法就是逐个分量相加），那么 $\mathbf{M}_n(\mathbb{R})$ 也有三个距离函数：对任意的 $n\times n$ 的矩阵 $A=(A_{ij})$ 和 $B=(B_{ij})$，其中 $1\leqslant i,j\leqslant n$，我们有

$$d_1(A,B)=\sum_{1\leqslant i,j\leqslant n}|A_{ij}-B_{ij}|,$$
$$d_2(A,B)=\sqrt{\sum_{1\leqslant i,j\leqslant n}(A_{ij}-B_{ij})^2},$$
$$d_\infty(A,B)=\sup_{1\leqslant i,j\leqslant n}|A_{ij}-B_{ij}|.$$

回到抽象的场合：如果 d_1 和 d_2 是 X 上两个等价的距离函数，那么它们所定义的收敛的概念是一致的，即对于任意的点列 $\{x_n\}_{n\geqslant 1}\subset X$，当 $n\to\infty$，在 d_1 这个距离下 $x_n\to x\in X$ 等价于在 d_2 这个距离下 $x_n\to x\in X$。证明是平凡的：假设 $x_n\xrightarrow{d_1}x$，那么对任意的 $\varepsilon>0$，存在 $N>0$，使得当 $n\geqslant N$ 时，我们有 $d_1(x_n,x)<\varepsilon$，所以根据距离的等价性，我们有 $d_2(x_n,x)<C_1\varepsilon$，这表明 $x_n\xrightarrow{d_2}x$。

特别地，数学分析通常在 $\mathbb{R}^n$ 中研究各种收敛的问题，除非我们对距离有特定的要求，我们假定度量可以是上面的任意一种。

4) 对于复数和 $n\times n$ 的矩阵，我们也可以谈论乘法。此时，乘法和除法也和极限交换。我们只给出矩阵的版本，复数的版本的叙述和证明都是一样的。假设 $\{A_k\}_{k\geqslant 1}$ 和 $\{B_k\}_{k\geqslant 1}$ 均为 $n\times n$ 的矩阵的序列并且收敛，那么

   - 序列 $\{A_k\cdot B_k\}_{k\geqslant 1}$ 收敛并且 $\lim\limits_{k\to\infty}(A_k\cdot B_k)=\lim\limits_{k\to\infty}A_k\cdot\lim\limits_{k\to\infty}B_k$。
   - 如果 $\lim\limits_{k\to\infty}B_k$ 是可逆矩阵，那么序列 $\{A_n\cdot(B_n)^{-1}\}_{B\geqslant N}$ 收敛（$\lim\limits_{n\to\infty}B_k$ 可逆表明存在 N，使得当 $k\geqslant N$ 时，B_k 均可逆）并且 $\lim\limits_{k\to\infty}A_k\cdot(B_k)^{-1}=\lim\limits_{k\to\infty}A_k\cdot\big(\lim\limits_{k\to\infty}B_k\big)^{-1}$。

   我们将在第二次作业中证明复数的情形。

上次课的最后我们证明了单调上升并且有界的实数序列有极限，这个极限恰好是它的上确界。当然，如果 $\{x_n\}_{n\geqslant 1}$ 单调上升的但是无界，我们有 $\lim\limits_{n\to\infty} x_n = +\infty$。利用这个结果，我们可以证明

定理 17 (Bolzano-Weierstrass 的列紧性定理)**.** *任意有界的实数序列 $\{x_n\}_{n\geqslant 1}$ 必有收敛的子列。*

我们只要证明如下引理即可：

引理 18. *对任意实数数列 $\{x_n\}_{n\geqslant 1}$，我们总能找到一个单调的（上升或下降）子序列。*

证明： 考虑下面的集合 $X \subset \{x_1, x_2, \cdots, x_n, \cdots\}$：

$$X = \left\{x_k \middle| \text{对任意的 } \ell \geqslant k, \text{ 都有 } x_k \geqslant x_\ell\right\}.$$

分两种情况讨论：

- 如果 X 是无限集，那么将 X 中元素按照下标从小到大排列，就得到了一个递减的子序列。
- 如果 X 是有限集，可以假设 $X = \left\{x_{i_1}, \cdots, x_{i_\ell}\right\}$，我们令 $y_1 = x_{i_\ell+1}$，然后归纳地定义 y_{n+1}：我们假设 $y_n = x_{s(n)}$ 其中 $s(n) > i_\ell$。根据 X 的定义以及由于 $y_n \notin X$，我们知道存在 $x_{s(n+1)}$，其中 $s(n+1) > s(n)$ 并且 $x_m > y_n$，那么我们令 $y_{n+1} = x_{s(n+1)}$。那么，$\{y\}_{n\geqslant 1}$ 是单调上升的序列。

综合两种情况，我们总有单调的子序列。 □

在证明 Cauchy 判别准则之前，我们先回忆一下 Cauchy 列的定义：一个距离空间 (X, d) 中的点列被称作是 **Cauchy 列**，指的是对任意的 $\varepsilon > 0$，存在 $N > 0$，使得对任意的 $m, n > N$，都有 $d(x_n, x_m) < \varepsilon$。

定理 (Cauchy 判别准则)**.** *$\{x_n\}_{n\geqslant 1}$ 是实数的数列。那么，$\{x_n\}_{n\geqslant 1}$ 收敛当且仅当 $\{x_n\}_{n\geqslant 1}$ 是 Cauchy 列。*

证明： 如果 $\{x_n\}_{n\geqslant 1}$ 收敛，我们假设 $x_n \to x$。此时，根据极限的定义，对于 $\dfrac{1}{2}\varepsilon$ 而言，存在 N，使得对于任意的 $n, m \geqslant N$，我们都有 $d(x, x_n) < \dfrac{1}{2}\varepsilon$，$d(x, x_m) < \dfrac{1}{2}\varepsilon$。所以，利用三角不等式，我们就有

$$d(x_n, x_m) < d(x, x_n) + d(x, x_m) < \frac{1}{2}\varepsilon + \frac{1}{2}\varepsilon = \varepsilon.$$

这表明收敛的序列（在任意的距离空间中）一定是 Cauchy 列。

为了说明 Cauchy 列必然收敛，我们先证明两个有用的引理：

引理 19. *Cauchy 列必有界。*

证明： 假设 $\{x_n\}_{n\geqslant 1}$ 是 Cauchy 列，我们先说明它是有界的。令 $\varepsilon = 1$，那么存在 N，使得对任意的 m, n，我们都有 $|x_n - x_m| < 1$。特别地，我们令 $m = 1$，这表明对所有的 $n \geqslant M$，都有 $|x_n - X_N| \leqslant 1$，所以 $\{x_n\}_{n\geqslant N}$ 是有界的。再加上前面的 $x_1, \cdots, x_{N-1}$，这还是一个有界集合。 □

引理 20. *如果一个 Cauchy 列的子列收敛，那么这个 Cauchy 列也收敛。*

证明: 假设 $\{x_n\}_{n\geqslant 1}$ 是 Cauchy 列，$\{x_{i_k}\}_{k\geqslant 1}$ 是其子列并且 $\lim\limits_{k\to\infty} x_{i_k} = x$，我们要证明在 (X,d) 中，$x_n \to x$。任意选取 $\varepsilon > 0$。首先，根据 $x_{i_k} \to x$，我们可以找到 $N_1 > 0$，使得对任意的 $k > N_1$，我们都有 $|x_{i_k} - x| < \dfrac{\varepsilon}{2}$；另外，根据 Cauchy 列的定义，我们有 $N_2 > 0$，使得对任意的 $n, m > N_2$，我们都有 $|x_n - x_m| < \dfrac{\varepsilon}{2}$。令 $N = \max\big(\min\limits_{i_k\geqslant N_1}\{i_k\}, N_2\big)$，所以当 $n \geqslant N$ 时，我们有

$$d(x_n, x) \leqslant d(x_n, dx_{i_k}) + d(x_n, dx_{i_k}) < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon,$$

其中 $k \geqslant N_1$ 可以任意选取。 □

根据上面第二个引理，我们只要构造一个收敛的子列即可。根据第一个引理，我们能找到一个有界的子列，再利用 Bolzano-Weierstrass 的列紧性，这个有界子列有一个收敛的子列。证毕。 □

注记. *1)* 相比于极限定义本身，利用 *Cauchy* 判别准则证明极限存在的优势在于不需要先验地知道极限的值。

比如说，我们有

$$1 - \frac{1}{2} + \frac{1}{3} - \cdots + \frac{(-1)^{n+1}}{n} + \cdots = \frac{1}{4}\pi.$$

我们可以证明上面的级数是收敛的（而不需要知道最终是 $\dfrac{1}{4}\pi$）：考虑部分和 $x_n = 1 - \dfrac{1}{2} + \dfrac{1}{3} - \cdots + \dfrac{(-1)^{n+1}}{n}$，不妨假设 $m > n$，那么我们有

$$x_m - x_n = \frac{(-1)^n}{n+1} + \frac{(-1)^{n+1}}{n+2} + \cdots + \frac{(-1)^{m+1}}{m}$$

如果 n 是偶数，那么，我们知道

$$x_m - x_n = \frac{1}{n+1} - \left(\frac{1}{n+2} - \frac{1}{n+3}\right) + \cdots + \frac{(-1)^{m+1}}{m}$$

从第二项开始每两个数一组，最后如果剩下一个数可以单独一组，我们发现这些组都是负数，从而 $x_m - x_n < \dfrac{1}{n}$。类似地，我们有

$$x_m - x_n = \frac{1}{n+1} - \frac{1}{n+2} + \left(\frac{1}{n+3} - \frac{1}{n+4}\right) + \cdots + \frac{(-1)^{m+1}}{m}$$

即从第三项开始每两个数一组，最后如果剩下一个数可以单独一组，我们发现这些组都是正数，从而 $x_m - x_n > \dfrac{1}{n} - \dfrac{1}{n+1}$。对 n 是奇数的情况类似讨论，我们可以得到

$$\frac{1}{n(n+1)} < |x_m - x_n| < \frac{1}{n}.$$

所以，对任意的 $\varepsilon > 0$，任取 $N > \frac{1}{\varepsilon}$，从而当 $m > n \geqslant N$ 时，我们有 $|x_m - x_n| < \varepsilon$，这就得到一个 *Cauchy* 列。利用同样的想法，我们可以证明下面的命题（需要记住结论）：

练习. 假设 $\{a_n\}_{n\geqslant 1}$ 是递减的正实数的数列并且 $\lim\limits_{n\to\infty} a_n = 0$，那么，级数

$$a_1 - a_2 + a_3 - a_4 + \cdots + (-1)^{n-1}a_n + \cdots$$

是收敛的。

我们把这个练习留做本次的作业。有一个和这个习题相关联的有趣的姊妹问题，证明和结论都值得大家研究：

命题 21. $\{a_n\}_{n\geqslant 1}$ 是递减的正实数的数列并且 $\lim\limits_{n\to\infty} a_n = 0$。假设级数 $\sum\limits_{k=1}^{\infty} a_k$ 发散。那么，对任意 $x \in \mathbb{R}$，我们可以选取一组正负号 $\iota_k \in \{\pm 1\}$，使得级数 $\sum_{k=1}^{\infty} \iota_k a_k$ 收敛并且

$$\sum_{k=1}^{\infty} \iota_k a_k = x.$$

证明：因为 $\sum\limits_{k=1}^{\infty} a_k = +\infty$，所以对任意的 $M > 0$，任意的 n，从存在唯一一个 $k \geqslant 0$，使得

$$a_n + a_{n+1} + \cdots + a_{n+k-1} \leqslant M \ \text{ 且 } \ a_n + a_{n+1} + \cdots + a_{n+k-1} + a_{n+k} > M.$$

我们不妨假设 $x > 0$，根据上面的观察，我们能找到 n_1，使得

$$a_1 + a_2 + \cdots + a_{n_1-1} \leqslant x \ \text{ 但是 } \ a_1 + a_2 + \cdots + a_{n_1-1} + a_{n_1} > x.$$

特别地，我们知道对于前面的 n_1 项的和，我们有

$$|a_1 + a_2 + \cdots + a_{n_1-1} + a_{n_1} - x| < a_{n_1}.$$

由于前 n_1 项的和已经超过了 x，我们现在在后面的项前面加上负号。令 $\delta_1 = a_1 + a_2 + \cdots + a_{n_1-1} + a_{n_1} - x$ 为多出来的部分，那么 $\delta_1 < a_{n_1}$。根据上面的观察，一定存在 n_2，使得

$$a_{n_1+1} + a_{n_1+2} + \cdots + a_{n_2-1} \leqslant \delta_1 \ \text{ 且 } \ a_{n_1+1} + a_{n_1+2} + \cdots + a_{n_2} > \delta_1.$$

我们现在要求级数的前 n_2 项为

$$a_1 + a_2 + \cdots + a_{n_1} - a_{n_1+1} - a_{n_1+2} - \cdots - a_{n_2}.$$

很明显，对于 $k < n_2$，我们有

$$0 < a_1 + a_2 + \cdots + a_{n_1} - a_{n_1+1} - a_{n_1+2} - \cdots - a_k < \delta_1.$$

令 $-\delta_2 = a_1 + a_2 + \cdots + a_{n_1} - a_{n_1+1} - a_{n_1+2} - \cdots - a_{n_2} - x$，这是一直到 n_2 为止，所求和比 x 少的部分，按照定义，我们有 $\delta_2 < a_{n_2}$。然后，我们再选取 n_3，使得

$$a_{n_2+1} + a_{n_2+2} + \cdots + a_{n_3-1} \leqslant \delta_2 \ \text{ 且 } \ a_{n_2+1} + a_{n_2+2} + \cdots + a_{n_3} > \delta_2.$$

我们现在要求级数的前 n_3 项为

$$a_1+a_2+\cdots+a_{n_1}-a_{n_1+1}-a_{n_1+2}-\cdots-a_{n_2}+a_{n_2+1}+a_{n_2+2}+\cdots+a_{n_3}.$$

按照这种方式，最新得到的部分和与 x 的差距都不超过 $\delta_3<a_{n_3}$。如此重复归纳地选取正负号即可。 □

2) Cauchy 判别准则是对实数的序列来陈述的。如果在一般的距离空间 (X,d) 上，Cauchy 列未必收敛（试举出一个反例），即收敛性和序列所生活的空间的性质是密切相关的。

3) 对复数和 $\mathbb{R}^n$ 也成立。

推论 22 (级数收敛的 Cauchy 判别准则)**.** *实数项的级数 $\sum\limits_{k=0}^{\infty}a_k$ 收敛的充分必要条件是对任意 $\varepsilon>0$，存在 $N>0$，对任意自然数 $n\geqslant N$ 和任意自然数 $p\geqslant 0$，我们都有*

$$\Big|\sum_{n\leqslant k\leqslant n+p}a_k\Big|<\varepsilon.$$

证明: 对任意的 n，令 $S_n=a_1+\cdots+a_n$ 为部分和。级数收敛等价于说 $\{S_n\}_{n\geqslant 1}$ 收敛。根据 Cauchy 判别准则，对任意的 $\varepsilon>0$，存在 $N>0$，对任意自然数 $n\geqslant N$ 和 $m\geqslant N$，我们都有 $|S_n-S_m|<\varepsilon$。我们不妨假设 $m\geqslant n$，令 $m=n+p$ 即得到了推论所要求的形式。 □

注记. *（复数或 $\mathbb{R}^n$ 中也成立）。另外，假设 $\sum\limits_{k=0}^{\infty}a_k$ 收敛，取 $p=0$，上述判别法说明 $a_n\to 0$。反之则未必成立，比如考虑调和级数。*

对于一般的实数数列，尽管极限不一定存在，但是我们总能定义它的**上极限**和**下极限**：任意给定实数数列 $\{x_n\}_{n\geqslant 1}$，对任意 $n\geqslant 1$，我们令

$$\overline{x}_n=\sup_{\ell\geqslant n}x_\ell,\quad \underline{x}_n=\inf_{\ell\geqslant n}x_\ell.$$

很明显，$\{\overline{x}\}_{n\geqslant 1}$ 是单调下降的序列，$\{\underline{x}_n\}_{n\geqslant 1}$ 是单调上升的序列，所以 $\{\overline{x}\}_{n\geqslant 1}$ 和 $\{\underline{x}_n\}_{n\geqslant 1}$ 都有极限（极限可以是无穷）。据此，我们定义数列 $\{x_n\}_{n\geqslant 1}$ 的上极限和下极限为

$$\overline{\lim}x_n=\limsup_{n\to\infty}x_n=\lim_{n\to\infty}\overline{x}_n=\lim_{n\to\infty}(\sup_{\ell\geqslant n}x_\ell),$$

$$\underline{\lim}x_n=\liminf_{n\to\infty}x_n=\lim_{n\to\infty}\underline{x}_n=\lim_{n\to\infty}(\inf_{\ell\geqslant n}x_\ell).$$

根据定义以及极限保持不等号的性质，我们有 $\limsup\limits_{n\to\infty}x_n\geqslant\liminf\limits_{n\to\infty}x_n$。我们如下命题：

命题 23. *$\{x_n\}_{n\geqslant 1}$ 是实数数列。那么，$\{x_n\}_{n\geqslant 1}$ 收敛的充分必要条件是 $\limsup\limits_{n\to\infty}x_n=\liminf\limits_{n\to\infty}x_n$。*

我们会在作业中证明这个命题。

在进一步讨论如何判断极限是否存在之前，我们再研究两个极为重要的极限：

例子. $\lim\limits_{n\to\infty} n^{\frac{1}{n}} = 1$。（开 n 次方目前并未定义，我们先假设自己懂（按照中学的理解））

证明: 首先，我们显然有 $n^{\frac{1}{n}} \geqslant 1$。其次，根据算术-几何平均值不等式[6]（我们课程后面会严格证明这个不等式）可以得到

$$n^{\frac{1}{n}} = (1\cdot 1\cdot 1\cdots 1\cdot\sqrt{n}\cdot\sqrt{n})^{\frac{1}{n}} \leqslant \frac{1}{n}\big(1+\cdots+1+\sqrt{n}+\sqrt{n}\big) = 1-\frac{2}{n}+\frac{2}{\sqrt{n}}.$$

从而，$0 \leqslant n^{\frac{1}{n}} - 1 \leqslant \dfrac{2}{n}+\dfrac{2}{\sqrt{n}}$。只需要选择比较大的 N，就可以使得对 $n \geqslant N$ 的自然数 n，有 $\dfrac{2}{n}+\dfrac{2}{\sqrt{n}} < \varepsilon$。

□

例子 (Euler 常数 e 的构造)**.** 极限 $\lim\limits_{n\to\infty}(1+\dfrac{1}{n})^n$ 是存在的，我们把它记做是 e。e 还有如下的级数表达式:

$$e = \sum_{k=0}^{\infty}\frac{1}{k!}.$$

为了说明 $\lim\limits_{n\to\infty}(1+\dfrac{1}{n})^n$ 的存在性，我们只需要说明 $x_n = \big(1+\dfrac{1}{n}\big)^n$ 是单调上升的有界序列。为了说明有界性，我们注意到（利用数学归纳法）对任意的 $k \geqslant 2$，

$$k! \geqslant 2^k.$$

根据二项式展开，我们有

$$\begin{aligned}(1+\frac{1}{n})^n &= \sum_{k=0}^{n}\binom{n}{k}\frac{1}{n^k} = \sum_{k=0}^{n}\frac{n!}{k!(n-k)!}\frac{1}{n^k}\\ &= \sum_{k=0}^{n}\frac{1}{k!}\big(1-\frac{1}{n}\big)\big(1-\frac{2}{n}\big)\cdots\big(1-\frac{k-1}{n}\big) \leqslant \sum_{k=0}^{n}\frac{1}{k!}.\end{aligned}$$

根据 $k! \geqslant 2^k$ $(k \geqslant 2)$，我们就有

$$x_n = (1+\frac{1}{n})^n \leqslant 1+1+\sum_{k=1}^{n}\frac{1}{2^k} = 3.$$

其中，我们用到了如下的事实:

$$\sum_{k=1}^{n}\frac{1}{2^k} = 1.$$

[6]给定 n 个正实数 $a_1, a_2, \cdots, a_n$，它们的算术平均值 A_M 和几何平均值 G_M 分别定义为

$$A_M = \frac{a_1+a_2+\cdots+a_n}{n},\quad G_M = \big(a_1\cdot a_2\cdot\cdots\cdot a_n\big)^{\frac{1}{n}}.$$

那么，$G_M \leqslant A_M$ 并且如果等号成立当且仅当 $a_1 = a_2 = \cdots = a_n$。

下面证明 $x_n \leqslant x_{n+1}$:

$$\begin{aligned}(1+\frac{1}{n})^n &= \sum_{k=0}^{n}\frac{1}{k!}(1-\frac{1}{n})(1-\frac{2}{n})\cdots(1-\frac{k-1}{n})\\ &< \sum_{k=0}^{n}\frac{1}{k!}(1-\frac{1}{n+1})(1-\frac{2}{n+1})\cdots(1-\frac{k-1}{n+1})\\ &< \sum_{k=0}^{n+1}\frac{1}{k!}(1-\frac{1}{n+1})(1-\frac{2}{n+1})\cdots(1-\frac{k-1}{n+1}) = (1+\frac{1}{n+1})^{n+1}.\end{aligned}$$

为了证明 e 的级数表达式，我们先说明级数 $\sum_{k=2}^{\infty}\frac{1}{k!}$ 是收敛的（这个证明很具有一般性）：注意到级数 $\sum_{k=1}^{\infty}\frac{1}{2^k}$ 是收敛的，所以，根据 *Cauchy* 判别准则，对于任意的 $\varepsilon>0$，存在 N，使得当 $n\geqslant N$ 并且 $m\geqslant N$ 时，前 n 项的部分和与前 m 项的部分和的差满足

$$\frac{1}{2^{n+1}}+\frac{1}{2^{n+2}}+\cdots+\frac{1}{2^{n+m}}<\varepsilon.$$

我们利用 2^{-k} 控制 $(k!)^{-1}$，从而对于级数 $\sum_{k=2}^{\infty}\frac{1}{k!}$ 而言，对于上述任意选定的 ε 和对应的 N，前 n 项的部分和与前 m 项的部分和的差有如下的控制

$$\frac{1}{(n+1)!}+\frac{1}{(n+2)!}+\cdots+\frac{1}{(n+m)!}\leqslant\frac{1}{2^{n+1}}+\frac{1}{2^{n+2}}+\cdots+\frac{1}{2^{n+m}}<\varepsilon.$$

这就证明了级数的收敛性（*Cauchy* 判别准则）。

最终来证明这个级数的值就是 e。

一方面，对于任意的 n，我们有

$$\begin{aligned}(1+\frac{1}{n})^n &= \sum_{k=0}^{n}\binom{n}{k}\frac{1}{n^k} = \sum_{k=0}^{n}\frac{1}{k!}(1-\frac{1}{n})(1-\frac{2}{n})\cdots(1-\frac{k-1}{n})\\ &\leqslant \sum_{k=0}^{n}\frac{1}{k!}\leqslant\sum_{k=0}^{\infty}\frac{1}{k!}.\end{aligned}$$

即 $x_n\leqslant\sum_{k=0}^{\infty}\frac{1}{k!}$。通过对 n 取极限，我们得到 $e\leqslant\sum_{k=0}^{\infty}\frac{1}{k!}$。

另一方面，我们要利用 x_n 是单调上升的这个性质。先任意选取 n 和 n_0，使得 $n\geqslant n_0$（这两个数是待定的），我们有

$$\begin{aligned}e=\lim_{m\to\infty}(1+\frac{1}{m})^m &\geqslant (1+\frac{1}{n})^n = \sum_{k=0}^{n}\frac{1}{k!}(1-\frac{1}{n})(1-\frac{2}{n})\cdots(1-\frac{k-1}{n})\\ &\geqslant \sum_{k=0}^{n_0}\frac{1}{k!}(1-\frac{1}{n})(1-\frac{2}{n})\cdots(1-\frac{k-1}{n}).\end{aligned}$$

上面的不等式对任意的 n 和 n_0 都成立。我们先固定住 n_0，令 $n\to\infty$，就得到 $e\geqslant\sum_{k=0}^{n_0}\frac{1}{k!}$。再令 $n_0\to\infty$，我们就有 $e\geqslant\sum_{k=0}^{\infty}\frac{1}{k!}$。

综上所述，

$$e=\lim_{n\to\infty}(1+\frac{1}{n})^n=\sum_{k=0}^{\infty}\frac{1}{k!}.$$

注记. *1)* 利用 e 的级数表达式很容易相对精确的计算 e 的大小。7! 是一个 *5* 位数，8! 是一个 *6* 位数，所以只要算 6 项（前两项整数部分不算）就已经可以精确到小数点后 5 位了：$e=2.71828\cdots!$

2) e 是一个无理数。如若不然，我们假设 $e=\frac{m}{n}$，其中 m 和 n 都是正整数，那么 $n!\times\big(e-\sum_{k=0}^{n}\frac{1}{k!}\big)$ 应该是整数，然而

$$\begin{aligned}
&n!\times\big(e-\sum_{k=0}^{n}\frac{1}{k!}\big)\\
&=\frac{1}{(n+1)}+\frac{1}{(n+1)(n+2)}+\frac{1}{(n+1)(n+2)(n+3)}+\frac{1}{(n+1)(n+2)(n+3)(n+4)}+\cdots\\
&<\frac{1}{(n+1)}+\frac{1}{(n+1)(n+2)}+\frac{1}{(n+2)(n+3)}+\frac{1}{(n+3)(n+4)}+\cdots\\
&=\frac{1}{(n+1)}+\left(\frac{1}{(n+1)}-\frac{1}{(n+2)}\right)++\left(\frac{1}{(n+2)}-\frac{1}{(n+3)}\right)+\cdots\\
&=\frac{2}{n+1}<1.
\end{aligned}$$

这表明 $n!\times\big(e-\sum_{k=0}^{n}\frac{1}{k!}\big)$ 0，矛盾。

3) e 是怎么来的：历史上，e 和自然对数（我们尚未定义）是在 *16* 世纪计算银行存款的利息时自然出现的，有兴趣的同学可以去参考 *wiki* 上的叙述。我们给出一个简化版本的例子（有趣）：大汉银行每年的利息是 100%，数学系的司马迁同学有存款 100 元，打算一年之后买一辆 272 元的须臾牌自行车。他算了一下，存钱一年之后总资产为 200 元。几天后，大汉银行规定改了，可以每半年结算一下并且半年的利息就是 $100\%\div 2=50\%$，司马迁发现上半年刚到的时候就提钱出来，然后再存进去，这样子前半年的利息也有利息，所以他就可以多赚一点，这样一来，半年之后，他就有 $100\times(1+0.5)$ 元钱，一年之后，再乘以 $1+0.5$ 的利息，他一共有 $100\times(1+0.5)\times(1+0.5)=100\times(1+0.5)^2=225$ 元，总收益又多了 25 元。后来，大汉银行容许每个月都结算，并且每个月的利率是 $100\%\div 12=\frac{1}{12}$，司马迁同学本着尽量赚利息的利息的原则算了一下，每个月都去提出钱来然后再次存进去，他一年之后的资产变成了 $100\times(1+\frac{1}{12})^{\frac{1}{12}}\approx 261$ 元，这只差 11 块钱就可以买车了。他高兴地发现只要不厌其烦地多存一次，最终得到的钱就会变多（$x_{n+1}\geqslant x_n$）。大汉银行最终容许每天都可以存钱取钱一次，司马迁同学每天坚持去银行，一年之后还是没有买得起自行车。

我们现在列举出极限收敛的几种常见的判别方式。尽管它们的形式并不统一，但是背后的想法却是一致的：我们需要找一个所谓的**控制序列**！

命题 24. *我们有如下的判断收敛的方法：*

1) *(双边控制) 假设有三个实数序列 $\{a_n\}_{n\geqslant 1}$，$\{x_n\}_{n\geqslant 1}$ 和 $\{b_n\}_{n\geqslant 1}$，对任意的 $n\geqslant 1$，都有 $a_n\leqslant x_n\leqslant b_n$（即 x_n 在左右两边分别被 a_n 和 b_n 控制）。如果 $\{a_n\}_{n\geqslant 1}$ 和 $\{b_n\}_{n\geqslant 1}$ 都收敛并且 $\lim\limits_{n\to\infty} a_n=\lim\limits_{n\to\infty} b_n$，那么 $\{x_n\}_{n\geqslant 1}$ 收敛并且*

$$\lim_{n\to\infty} x_n=\lim_{n\to\infty} a_n=\lim_{n\to\infty} b_n.$$

2) *(上界控制)假设有非负实数序列 $\{x_n\}_{n\geqslant 1}$ 和 $\{y_n\}_{n\geqslant 1}$，对任意的 $n\geqslant 1$，都有满足 $0\leqslant x_n\leqslant y_n$（即 x_n 被 y_n 控制）。如果 $\lim\limits_{n\to\infty} y_n=0$，那么 $\{x_n\}_{n\geqslant 1}$ 也收敛并且 $\lim\limits_{n\to\infty} x_n=0$。*

证明： 第二个命题是第一个的推论。现在证明第一个命题。注意到

$$\limsup_{n\to\infty} x_n\leqslant\limsup_{n\to\infty} b_n=\lim_{n\to\infty} b_n,$$

和

$$\liminf_{n\to\infty} x_n\geqslant\liminf_{n\to\infty} a_n=\lim_{n\to\infty} a_n,$$

所以

$$\limsup_{n\to\infty} x_n\leqslant\liminf_{n\to\infty} x_n.$$

这说明 $\{x_n\}_{n\geqslant 1}$ 收敛。特别地，上面不等式也给出了 $\lim\limits_{n\to\infty} x_n=\lim\limits_{n\to\infty} a_n=\lim\limits_{n\to\infty} b_n$。 □

数学中很多的极限都以级数的形式出现，我们给出上述命题的级数版本：

命题 25 (控制收敛定理与绝对收敛的概念)**.** （重要!）

1) $\sum\limits_{k=0}^{\infty} a_k$ *是正项级数（即 $a_k\geqslant 0$，其中 $k\geqslant 1$），那么 $\sum\limits_{k=0}^{\infty} a_k$ 收敛当且仅当存在常数 M，使得每个部分和 $S_n=\sum\limits_{k=0}^{n} a_k\leqslant M$。*

2) *(正项级数的控制收敛定理)* $\sum\limits_{k=0}^{\infty} a_k$ *和* $\sum\limits_{k=0}^{\infty} b_k$ *是正项级数。假设对任意 $k\geqslant 0$，都有 $a_k\leqslant b_k$（即 b_k 控制了 a_k）。如果 $\sum\limits_{k=0}^{\infty} b_k$ 收敛，那么 $\sum\limits_{k=0}^{\infty} a_k$ 也收敛（等价的表述是若 $\sum\limits_{k=0}^{\infty} a_k$ 发散，那么 $\sum\limits_{k=0}^{\infty} b_k$ 也发散）。*

3) *(绝对收敛的概念) 考虑实数项的级数 $\sum\limits_{k=0}^{\infty} a_k$。如果 $\sum\limits_{k=0}^{\infty}|a_k|$ 收敛，那么 $\sum\limits_{k=0}^{\infty} a_k$ 也收敛。此时，我们称这个级数是* **绝对收敛的** *（即加了绝对值之后收敛）。*

证明：第一个命题的证明是单调递增的有界数列的必有极限的应用；第二个命题是第一个命题的直接推论。为了证明第三个命题，我们用级数收敛的 Cauchy 判别法：由于 $\sum_{k=0}^{\infty}|a_k|$ 收敛，所以对任意的 $\varepsilon>0$，存在 N，使得对任意的 $n\geqslant N$ 和任意的 $p\geqslant 0$，我们有 $\Big|\sum_{n\leqslant k\leqslant n+p}|a_k|\Big|<\varepsilon$。根据三角不等式，我们就有

$$\Big|\sum_{n\leqslant k\leqslant n+p}a_k\Big|<\Big|\sum_{n\leqslant k\leqslant n+p}|a_k|\Big|<\varepsilon,$$

所以，$\sum_{k=0}^{\infty}a_k$ 收敛。 □

注记. 上面的 2) 和 3) 结合在一起非常好用：为了证明一个级数 $\sum_{k=0}^{\infty}a_k$ 收敛，很多情况下只要说明它绝对收敛就可以了，此时，再找一个收敛的正项级数 $\sum_{k=0}^{\infty}b_k$ 控制 $\sum_{k=0}^{\infty}|a_k|$ 即可。

用较大的收敛级数来控制较小的级数从而证明较小级数是收敛的，这是分析中最基本一个技术和想法。

然而，这个想法貌似存在着不合理的地方：直观上，证明更大的级数收敛是比证明原来的小一点的级数收敛更难的事情。真正的解释是理解这个想法的核心（这在定理叙述中无法看出来）：通过适当选取较大的级数应**容易计算**。

例子. 我们给出上述命题的几个简单应用：

1) 交错项的调和级数 $\sum_{k=1}^{\infty}\frac{(-1)^k}{k}$ 是收敛的但是不绝对收敛。

2) 级数 $\sum_{k=1}^{\infty}\frac{1}{k^2}$ 是收敛的：我们可以用 $\sum_{k=1}^{\infty}\frac{1}{(k-1)k}$ 作为控制级数，此时，通过将后一个级数中的单项写成 $\frac{1}{k-1}-\frac{1}{k}$ 的形式很容易算出部分和（望远镜求和法）。

3) 级数 $(e=)\sum_{k=0}^{\infty}\frac{1}{k!}$ 收敛：我们可以用 $\sum_{k=2}^{\infty}\frac{1}{2^k}$ 作为控制级数，因为等比数列更容易求和。

练习. 绝对收敛的概念对于复数项的级数或者在线性空间中（包括矩阵）取值的级数也成立，我们这里只考虑复数的情形，其余的我们会有更为一般的讨论：

考虑复数项的级数 $\sum_{k=0}^{\infty}a_k$。如果 $\sum_{k=0}^{\infty}|a_k|$ 收敛，证明，$\sum_{k=0}^{\infty}a_k$ 也收敛，其中 $|\cdot|$ 是取复数的模长。

6 指数函数与三角函数的构造，指数函数与三角函数的代数性质，双指标级数的求和

二零一九年九月二十六日，星期四，晴天

在讲解指数函数的构造之前，我们先讲解两道经典的极限习题：

例子. 当 $\alpha>1$ 时，级数 $\displaystyle\sum_{k=1}^{\infty}\frac{1}{k^{\alpha}}$ 是收敛的。

这是一个正项的级数，我们利用我们之前的分析：找一个能够控制这个级数并且同时它比较方便计算，为此我们再次利用调和级数中庄子的二分法的想法。对于 $n=2^k-1$，我们有

$$\sum_{k=1}^{n-1}\frac{1}{j^{\alpha}}=\underbrace{\frac{1}{1}}_{1\text{个},\ \leqslant 1}+\underbrace{\left(\frac{1}{2^{\alpha}}+\frac{1}{3^{\alpha}}\right)}_{2\text{个},\ \leqslant\frac{1}{2^{\alpha}}}+\underbrace{\left(\frac{1}{4^{\alpha}}+\frac{1}{5^{\alpha}}+\frac{1}{6^{\alpha}}+\frac{1}{7^{\alpha}}\right)}_{4\text{个},\ \leqslant\frac{1}{2^{2\alpha}}}+\cdots+\underbrace{\left(\frac{1}{2^{(k-1)\alpha}}+\frac{1}{\left(2^{k-1}+1\right)^{\alpha}}+\cdot+\frac{1}{\left(2^{k}-1\right)^{\alpha}}\right)}_{2^{k-1}\text{个},\ \leqslant\frac{1}{2^{(k-1)\alpha}}}$$

$$\leqslant\sum_{j=1}^{k}\frac{1}{2^{(j-1)\alpha}}\times 2^{j-1}=\sum_{j=1}^{k}2^{-j(\alpha-1)}.$$

由于 $\alpha-1>0$，根据等比数列的求和公式，上面的求和是有限的。

例子. $\{x_n\}_{n\geqslant 1}$ 是由非负实数构成数列。假设对任意的自然数 m 和 n，都有 $x_{m+n}\leqslant x_m+x_n$。证明，数列 $\left\{\dfrac{x_n}{n}\right\}_{n\geqslant 1}$ 有极限。

证明: 先固定一个自然数 k。对任意的 $n=\ell k+r$，其中 $r=0,1,\cdots,k-1$。根据题目中所给的不等式，我们有

$$\frac{a_n}{n}\leqslant\frac{\ell a_k+a_r}{n}=\frac{\ell a_k+a_r}{\ell k+r}.$$

据此可知，数列 $\left\{\dfrac{x_n}{n}\right\}_{n\geqslant 1}$ 有界。我们可以选取 $n_i\to\infty$（这是本次作业的一个题目），使得

$$\lim_{i\to\infty}\frac{a_{n_i}}{n_i}\to\limsup\frac{a_n}{n}.$$

根据上述的不等式（把 n 换成 n_i，ℓ 换成 ℓ_i，r 换成 r_i），令 $i\to\infty$，即 $n_i\to\infty$，从而，$\dfrac{\ell_i a_k+a_{r_i}}{\ell_i k+r_i}\to\frac{a_k}{k}$。所以，我们得到

$$\limsup_{n\to\infty}\frac{a_n}{n}\leqslant\frac{a_k}{k}.$$

现在允许 k 变化，对上面左右两边同时取下极限，我们得到

$$\limsup_{n\to\infty}\frac{a_n}{n}\leqslant\liminf_{k\to\infty}\frac{a_k}{k}.$$

所以所研究的极限存在。 □

指数函数的构造

目前，我们在实数 $\mathbb{R}$ 上能够定义的函数非常有限：根据 $\mathbb{R}$ 上的乘法和加法结构，我们目前只能定义多项式函数

$$\mathbb{R} \to \mathbb{R}, \quad x \mapsto a_d x^d + a_{d-1}x^{d-1} + \cdots + a_1 x + a_0,$$

其中 $a_0, \cdots, a_d$ 是实数。

利用级数（取极限的概念），我们可以定义

定义 26. 我们定义**指数函数**$\exp : \mathbb{R} \to \mathbb{R}$ 如下：

$$x \mapsto \exp(x) = e^x = \sum_{k=0}^{\infty} \frac{x^k}{k!}.$$

我们必须要说明指数函数是良好定义，即对任意的 $x \in \mathbb{R}$，级数 $\sum\limits_{k=0}^{\infty} \frac{x^k}{k!}$ 是收敛的：这是因为对任意给定的 $x \in \mathbb{R}$，我们假设 $|x| = M$，此时，一定存在 $N > 0$，使得 $k \geqslant N$ 时，我们有（请证明这一点）

$$\left|\frac{x^k}{k!}\right| \leqslant \frac{1}{2^k} (\Leftrightarrow (2M)^k \leqslant k!, \text{ 如果 } k \geq M)$$

此时，我们可以用 $\sum\limits_{k \geqslant N} \frac{1}{2^k}$ 来控制 $\sum\limits_{k \geq N} \frac{x^k}{k!}$，这表明级数 $\sum\limits_{k=0}^{\infty} \frac{x^k}{k!}$ 收敛。

练习. （重要!）证明，我们可以我们 $\mathbb{C}$ 上的指数函数：

$$\exp : \mathbb{C} \to \mathbb{C}, \quad z \mapsto \exp(z) = e^z = \sum_{k=0}^{\infty} \frac{z^k}{k!}.$$

为了进一步了解 e^x 的性质，比如说对任意的 $x \in \mathbb{R}$，我们都有 $e^x > 0$，我们需要关于级数的进一步的性质。事实上，为了证明 $\exp : \mathbb{R} \to \mathbb{R}_{>0}$，我们需要证明 exp 的一个代数性质：对任意的 $x, y \in \mathbb{R}$，我们都有 $e^{x+y} = e^x \cdot e^y$（即 $\exp : (\mathbb{R}, +) \to (\mathbb{R}_{>0}, \times)$ 是群同态）。

我们需要研究双指标的数列。所谓的双指标的数列 $\{x_{i,j}\}_{i,j \geqslant 1}$（我们目前先假设这是实数序列即可，其余的情况可以简单的推广过去），就是一个映射

$$F : \mathbb{Z}_{\geqslant 1} \times \mathbb{Z}_{\geqslant 1} \to \mathbb{R}, \quad (i, j) \mapsto F(i, j) = x_{i,j}.$$

一个双指标序列的重排 $\{y_n\}_{n \geqslant 1}$ 直观上说的我们要求当每个 $a_{i,j}$（位置不同的时候视作是不同的）都在数列 $\{y_n\}_{n \geqslant 1}$ 中出现且只出现一次并且 $\{y_n\}_{n \geqslant 1}$ 中不再出现别的项。我们用映射的语言严格定义重排的概念：所谓的**重排**指的是一个双射 $\Phi : \mathbb{Z}_{\geqslant 1} \to \mathbb{Z}_{\geqslant 1} \times \mathbb{Z}$，从而对任意的 $n \geqslant 1$，我们令 $y_n = x_{\Phi(n)}$，其中 $\Phi(n) \in \mathbb{Z}_{\geqslant 1} \times \mathbb{Z}$ 是一个双指标。

我们用双指标的序列来研究两个级数的乘积：

命题 27. $\sum\limits_{k=1}^{\infty} a_k$ 和 $\sum\limits_{k=1}^{\infty} b_k$ 是收敛的正项（实数）级数，$\{c_n\}_{n\geqslant 1}$ 是 $\{a_i b_j\}_{i,j\geqslant 1}$ 的一个重排，那么，级数 $\sum\limits_{n=1}^{\infty} c_n$ 收敛并且（无论采取何种重排）有

$$\sum_{n=1}^{\infty} c_n = \big(\sum_{k=1}^{\infty} a_k\big)\big(\sum_{k=1}^{\infty} b_k\big).$$

特别地（通过选取 45° 线的重排），级数 $\sum\limits_{k=2}^{\infty}\big(\sum\limits_{j=1}^{k} a_j b_{k-j}\big)$ 是收敛的，并且

$$\sum_{k=2}^{\infty}\big(\sum_{j=1}^{k} a_j b_{k-j}\big) = \big(\sum_{k=1}^{\infty} a_k\big)\big(\sum_{k=1}^{\infty} b_k\big).$$

进一步，如果 $\sum\limits_{k=1}^{\infty} a_k$ 和 $\sum\limits_{k=1}^{\infty} b_k$ 是绝对收敛的实数级数（未必是正项的），那么上面的结论仍然成立。

证明: 先处理正项级数的情形：对任意 $N>0$，一定存在 N_1 和 N_2，使得部分和

$$C_N = \sum_{n=1}^{N} c_n \leqslant \Big(\sum_{k=1}^{N_1} a_k\Big)\Big(\sum_{\ell=1}^{N_1} b_\ell\Big) \leqslant \Big(\sum_{k=1}^{\infty} a_k\Big)\Big(\sum_{k=1}^{\infty} b_k\Big).$$

即部分和是有界的，从而 $\sum\limits_{n=1}^{\infty} c_n$ 收敛。

其次，我们先任意给定 $\varepsilon>0$。对任意 ε_1 和 ε_2（这两个数待定，将由 ε 决定），按照定义，我们可以选取 N_1 和 N_2 使得

$$|\sum_{k=1}^{N_1} a_k - \sum_{k=1}^{\infty} a_k| < \varepsilon_1,\ |\sum_{k=1}^{N_2} b_k - \sum_{k=1}^{\infty} b_k| < \varepsilon_2.$$

从而，我们有

$$\begin{aligned}
&\big|(\sum_{k=1}^{N_1} a_k)(\sum_{k=1}^{N_2} b_k) - (\sum_{k=1}^{\infty} a_k)(\sum_{k=1}^{\infty} b_k)\big| \\
\leqslant &\big|(\sum_{k=1}^{N_1} a_k)(\sum_{k=1}^{N_2} b_k) - (\sum_{k=1}^{N_1} a_k)(\sum_{k=1}^{\infty} b_k)\big| + \big|(\sum_{k=1}^{N_1} a_k)(\sum_{k=1}^{\infty} b_k) - (\sum_{k=1}^{\infty} a_k)(\sum_{k=1}^{\infty} b_k)\big| \\
= &\big|\sum_{k=1}^{N_1} a_k\big|\big|\sum_{k=1}^{N_2} b_k - \sum_{k=1}^{\infty} b_k\big| + \big|\sum_{k=1}^{N_1} a_k - \sum_{k=1}^{\infty} a_k\big|\big|\sum_{k=1}^{\infty} b_k\big| \\
\leqslant &\varepsilon_1\big|\sum_{k=1}^{\infty} a_k\big| + \varepsilon_2\big|\sum_{k=1}^{\infty} b_k\big|
\end{aligned}$$

所以，可以适当的选取 ε_1 和 ε_2（比如 $\varepsilon_1=\varepsilon_2=\dfrac{\varepsilon}{2}\big(\sum\limits_{k=1}^{\infty}a_k+\sum\limits_{k=1}^{\infty}b_k\big)^{-1}$）使得上面式子的右端小于 ε。最后，再选取 N 足够大，使得 $\{c_1,\cdots,c_N\}$ 已经囊括了所有的 a_ib_j 形式的项，其中 $1\leqslant i\leqslant N_1$，$1\leqslant j\leqslant N_2$，此时，

$$\Big|C_N-\big(\sum_{k=1}^{\infty}a_k\big)\big(\sum_{k=1}^{\infty}b_k\big)\Big|<\varepsilon.$$

这就完成了正项级数情形的证明。

为了证明绝对收敛的情形，我们要把级数分拆成两个部分。首先对于任意的实数其中给定实数 x，我们定义

$$x^+=\begin{cases}x, & \text{如果 } x\geqslant 0;\\ 0, & \text{如果 } x\leqslant 0,\end{cases}\qquad x^-=\begin{cases}0, & \text{如果 } x\geqslant 0;\\ -x, & \text{如果 } x\leqslant 0。\end{cases}$$

所以，$x=x^+-x^-$。从而，我们将 $\sum\limits_{k=1}^{\infty}a_k$ 分拆为

$$\sum_{k=1}^{\infty}a_k=\sum_{k=1}^{\infty}a_k^+-\sum_{k=1}^{\infty}a_k^-.$$

由于 $|x^{\pm}|\leqslant|x|$，根据绝对收敛性，$\sum\limits_{k=1}^{\infty}a_k^{\pm}$ 均收敛。类似地讨论对 $\sum\limits_{k=1}^{\infty}b_k$ 也成立，其中我们将它分解为

$$\sum_{k=1}^{\infty}b_k=\sum_{k=1}^{\infty}b_k^+-\sum_{k=1}^{\infty}b_k^-.$$

对于数列 $\{c_n\}_{n\geqslant1}$ 的任意一项有 $c_n=a_ib_j$，我们有

$$c_n=\underbrace{a_i^+b_j^+}_{c_n^{++}}+\underbrace{a_i^-b_j^-}_{c_n^{--}}-(\underbrace{a_i^+b_j^-}_{c_n^{+-}}+\underbrace{a_i^-b_j^+}_{c_n^{-+}}).$$

所以，我们将 $\sum\limits_{k=1}^{\infty}c_n$ 分拆为四个级数的和（差）：

$$\sum_{k=1}^{\infty}c_n=\sum_{k=1}^{\infty}c_n^{++}+\sum_{k=1}^{\infty}c_n^{--}-\sum_{k=1}^{\infty}c_n^{+-}-\sum_{k=1}^{\infty}c_n^{-+}.$$

根据刚证明的关于正项级数的结论，上面右端的每个级数都收敛，所以 $\sum\limits_{k=1}^{\infty}c_n$ 也收敛。另外，利

用关于正项级数的结论我们还有

$$\begin{aligned}\sum_{k=1}^{\infty} c_n &= (\sum_{k=1}^{\infty} a_k^+)(\sum_{k=1}^{\infty} b_k^+) + (\sum_{k=1}^{\infty} a_k^-)(\sum_{k=1}^{\infty} b_k^-) - (\sum_{k=1}^{\infty} a_k^+)(\sum_{k=1}^{\infty} b_k^-) - (\sum_{k=1}^{\infty} a_k^-)(\sum_{k=1}^{\infty} b_k^+) \\ &= (\sum_{k=1}^{\infty} a_k^+ + \sum_{k=1}^{\infty} a_k^-)(\sum_{k=1}^{\infty} b_k^+ + \sum_{k=1}^{\infty} b_k^-) \\ &= (\sum_{k=1}^{\infty} a_k)(\sum_{k=1}^{\infty} b_k).\end{aligned}$$

命题得到了证明。 □

注记. 上述绝对收敛部分定理对于复数的情形也成立。证明的过程中，我们需要进一步将一个复数再分解成实部和虚部之后，再对它们分别作上述分析。

练习. 试证明级数形式的 *Funibi* 定理：假设级数 $\sum_{k=2}^{\infty}(\sum_{j=1}^{k} a_j b_{k-j})$ 是绝对收敛的（复数）级数，那么，我们有

$$\sum_{i=1}^{\infty}(\sum_{j=1}^{\infty} a_i b_j) = (\sum_{k=1}^{\infty} a_k)(\sum_{k=1}^{\infty} b_k) = \sum_{j=1}^{\infty}(\sum_{i=1}^{\infty} a_i b_j).$$

定理 28. 指数函数 $\exp : \mathbb{C} \to \mathbb{C}$ 满足：对任意的 $z_1, z_2 \in \mathbb{C}$，我们都有

$$e^{z_1+z_2} = e^{z_1} \cdot e^{z_2}.$$

特别地，对任意的 $z \in \mathbb{C}$，我们有 $\exp(z) \neq 0$；对任意的 $x \in \mathbb{R}$，我们有 $\exp(x) > 0$。（用代数的语言讲，$\exp : (\mathbb{C}, +) \to (\mathbb{C} - \{0\}, \times)$ 是群同态）

证明: 我们已经知道 $e^z = \sum_{k=0}^{\infty} \frac{z^k}{k!}$ 是绝对收敛的级数，利用上面的性质（保证了下面的所有操作都合法），我们有

$$\begin{aligned}e^{z_1} \cdot e^{z_2} &= (\sum_{k=0}^{\infty} \frac{z_1^k}{k!})(\sum_{k=0}^{\infty} \frac{z_2^k}{k!}) = \sum_{k=0}^{\infty} \sum_{i+j=k} \frac{z_1^i}{i!} \frac{z_2^j}{j!} \\ &= \sum_{k=0}^{\infty} \frac{1}{k!} \sum_{i+j=k} \frac{k!}{i!j!} z_1^i \cdot z_2^j = \sum_{k=0}^{\infty} \frac{1}{k!}(z_1 + z_2)^k = e^{z_1+z_2}.\end{aligned}$$

我们注意到，为了上面的红色等号成立，我们已经隐含的用到了 $z_1 \cdot z_2 = z_2 \cdot z_1$。

另外，我们注意到一个很简单但是很重要的事实（用定义）：

$$e^0 = 1.$$

为了说明 $e^z \neq 0$，我们只需要观察到 $e^z \cdot e^{-z} = e^0 = 1$ 即可；为了说明 $e^x > 0$，其中 $x \in \mathbb{R}$，我们首先用定义说明如果 $x > 0$，那么 $e^x > 0$，进一步对于 $-x$ $(x > 0)$，我们有 $e^{-x} = \left(e^x\right)^{-1} > 0$。 □

练习. 证明，函数 $\exp(x)$ 在 $\mathbb{R}$ 上是严格递增的。特别地，$\exp : \mathbb{R} \to \mathbb{R}_{>0}$ 是单射。

实际上，$\exp : \mathbb{R} \to \mathbb{R}_{>0}$ 是满射（从而可以定义 log 函数），我们还需要等待必要的工具来完成这个证明。

三角函数的定义

利用指数映射 $\exp(z)$，我们可以定义正弦和余弦函数：

$$\cos z = \frac{e^{iz}+e^{-iz}}{2} = \sum_{k=0}^{\infty} \frac{(-1)^k x^{2k}}{(2k)!}, \quad \sin z = \frac{e^{iz}-e^{-iz}}{2i} = \sum_{k=0}^{\infty} \frac{(-1)^k x^{2k+1}}{(2k+1)!}.$$

换句话说，目前 $\sin(z)$ 和 $\cos(z)$ 使用级数来定义的。特别地，我们有 **Euler 公式**：

$$e^{iz} = \cos z + i \sin z.$$

注记. *1) 另外，关于正切函数余切函数等，我们仍然和中学一样，通过正弦和余弦的商的定义。*

3 中学数学里我们是这样定义三角函数的：假设 $\triangle ABC$ 是直角三角形，$\angle B$ 是直角，$\angle A = \theta$，那么 $\sin(\theta) = \dfrac{|AB|}{|AC|}$。我们将证明我们所定义的三角函数也满足这个性质。这将是一个很深刻的定理，因为我们的函数只用极限就可以定义，目前没有任何的证据表明它实际上还可以几何地被定义。

现在是一个绝佳的实际来展示函数代数性质的应用，利用 $e^{z_1+z_2} = e^{z_1} \cdot e^{z_2}$，我们可以证明有关三角函数的熟知的性质：

1) 对任意的 $z \in C$，都有 $(\cos z)^2 + (\sin z)^2 = 1$。

   按照定义，

   $$(\cos z)^2 + (\sin z)^2 = \frac{\left(e^{iz}+e^{-iz}\right)^2}{4} - \frac{\left(e^{iz}-e^{-iz}\right)^2}{4} = 1.$$

2) 对任意 $x \in \mathbb{R}$，都有 $\left|e^{ix}\right| = 1$，其中 $|\cdot|$ 是取复数的模长。特别地，对于 $x \in \mathbb{R}$，我们有 $|\sin(x)| \leqslant 1$ 和 $|\cos(x)| \leqslant 1$。

   这是前一公式的推论，因为如果 $x \in \mathbb{R}$，那么 $\sin(x) \in \mathbb{R}$，$\cos(x) \in \mathbb{R}$。

3) 三角函数的和差化积公式：对任意的实数 x 和 y，

   $$\cos(x+y) = \cos x \cos y - \sin x \sin y, \quad \sin(x+y) = \sin x \cos y + \cos x \sin y.$$

   根据 exp 的性质，

   $$\cos(x+y) + i \sin(x+y) = e^{ix} \cdot e^{iy} = \big(\cos(x) + i \sin(x)\big)\big(\cos(y) + i \sin(y)\big).$$

   展开比较实部和虚部即可（因为此时 $\sin(x)$ 和 $\cos(x)$ 都是实数）。

注记. *中学中所熟知的关于三角函数的性质都是由以上几个性质通过代数运算得来的。所以，我们之后仍然可以熟练使用所熟悉的关于三角函数的公式。*

6.1 作业：Riemann 重排，Cesàro 求和，Banach-Mazur 游戏

清华大学 19-20 秋季学期，数学分析一，作业 2

请用 A4 大小的纸张**正反面**用钢笔，签字笔或者圆珠笔书写，并注明自己的姓名，年级和作业的**总页数**。除定理公式所涉及的人名之外，请使用**中文**。本次作业的提交时间和地点为 **10 月 10 日**上午的课堂上，逾期视作零分。

基本习题

习题 A：课堂内容的补充及 $\varepsilon - N$ 语言的训练

A1) $\{x_n\}_{n\geqslant 1}$ 是有界的实数数列。证明，这个数列有子列 $\{x_{n_i}\}_{n\geqslant 1}$ 使得 $\lim\limits_{i\to\infty} x_{n_i}$ 存在并且

$$\lim_{i\to\infty} x_{n_i} = \limsup_{n\to\infty} x_n.$$

A2) $\{x_n\}_{n\geqslant 1}$ 是实数数列。证明，$\{x_n\}_{n\geqslant 1}$ 收敛的充分必要条件是 $\limsup\limits_{n\to\infty} x_n = \liminf\limits_{n\to\infty} x_n$。

A3) $\{x^{(k)}\}_{k\geqslant 1}$ 是 $\mathbb{R}^n$ 中的点列，其中，$x^{(k)} = (x_1^{(k)}, x_2^{(k)}, \cdots, x_n^{(k)})$。那么，$\{x^{(k)}\}_{k\geqslant 1}$ 在 $\mathbb{R}^n$ 中收敛当且仅当它的每个分量都是收敛的实数数列，即对任意的 $i = 1, 2, \cdots, n$，$\{x_i^{(k)}\}_{k\geqslant 1}$ 在 $\mathbb{R}$ 中收敛。（你在作业中只需要对 $n = 2$ 证明即可）

A4)（复数数列的四则运算）假设 $\{z_n\}_{n\geqslant 1}$ 和 $\{w_n\}_{n\geqslant 1}$ 是两个收敛的复数的数列。证明（你在作业中只需要证明第三条即可），

- 数列 $\{z_n \pm w_n\}_{n\geqslant 1}$ 收敛并且 $\lim\limits_{n\to\infty}(z_n \pm w_n) = \lim\limits_{n\to\infty} z_n \pm \lim\limits_{n\to\infty} w_n$；
- $\{z_n \cdot w_n\}_{n\geqslant 1}$ 收敛并且 $\lim\limits_{n\to\infty}(z_n \cdot w_n) = \lim\limits_{n\to\infty} z_n \cdot \lim\limits_{n\to\infty} w_n$；
- 如果 $\lim\limits_{n\to\infty} w_n \neq 0$，那么数列 $\{\dfrac{z_n}{w_n}\}_{n\geqslant N}$ 收敛（$\lim\limits_{n\to\infty} w_n \neq 0$ 表明存在 N，使得当 $n \geqslant N$ 时，$w_n \neq 0$）并且 $\lim\limits_{n\to\infty} \dfrac{z_n}{w_n} = \dfrac{\lim\limits_{n\to\infty} z_n}{\lim\limits_{n\to\infty} w_n}$。

A5) 假设 $\{a_n\}_{n\geqslant 1}$ 是递减的正实数的数列并且 $\lim\limits_{n\to\infty} a_n = 0$。证明，级数

$$a_1 - a_2 + a_3 - a_4 + \cdots + (-1)^{n-1} a_n + \cdots$$

是收敛的。

A6) $\sum\limits_{k=0}^{\infty} a_k$ 是复数项的级数。如果 $\sum\limits_{k=0}^{\infty} |a_k|$ 收敛，证明，$\sum\limits_{k=0}^{\infty} a_k$ 也收敛，其中 $|\cdot|$ 是取复数的模长。

A7) 证明，我们可以在 $\mathbb{C}$ 上定义指数函数：

$$\exp : \mathbb{C} \to \mathbb{C}, \quad z \mapsto \exp(z) = e^z = \sum_{k=0}^{\infty} \frac{z^k}{k!}.$$

A8) $\{a_n\}_{n\geqslant 1}$ 是复数的数列，我们假设对任意的 $n, a_n \neq 0$。令 $P_n = a_1 \cdot a_2 \cdots \cdots a_n$，如果数列 $\{P_n\}_{n\geqslant 1}$ 的极限存在并且该极限不是 0，我们就称无限乘积 $\prod\limits_{n\geqslant 1} a_n$ 收敛并且记 $\prod\limits_{n\geqslant 1} a_n = \lim\limits_{n\to\infty} P_n$。证明 Cauchy 判别准则：$\prod\limits_{n\geqslant 1} a_n$ 收敛当且仅当对任意的 $\varepsilon > 0$，存在 N，使得对任意的 $n \geqslant N$，任意的 $p \geqslant 0$，我们都有

$$\left|a_n \cdot a_{n+1} \cdot \cdots \cdot a_{n+p} - 1\right| < \varepsilon.$$

A9) 证明，函数 $\exp(x)$ 在 $\mathbb{R}$ 上是严格递增的。

A10) （基本的增长速度比较，记住）假设 $P(X)$ 是一个 n 次多项式（实数系数），$Q(X)$ 是一个 m 次多项式（实数系数），$m > n$，证明：

$$\lim_{k\to\infty} \frac{P(k)}{Q(k)} = 0, \quad \lim_{k\to\infty} \frac{Q(k)}{e^k} = 0.$$

习题 B 计算下面的极限（包含发散的情形）：

(1) $\lim\limits_{n\to\infty} \dfrac{n+10}{2n-1}$　(2) $\lim\limits_{n\to\infty} \dfrac{\sqrt{n}+10}{2\sqrt{n}-1}$　(3) $\lim\limits_{n\to\infty} 0.\underbrace{999\cdots 99}_{n个}$

(4) $\lim\limits_{n\to\infty} \dfrac{1}{n(n+3)}$　(5) $\lim\limits_{n\to\infty} \dfrac{\cos(n)}{n}$　(6) $\lim\limits_{n\to\infty} \dfrac{2^n}{n!}$

(7) $\lim\limits_{n\to\infty} \dfrac{n!}{n^n}$　(8) $\lim\limits_{n\to\infty} \sqrt{n+10} - \sqrt{n+1}$　(9) $\lim\limits_{n\to\infty} \dfrac{1+2+\cdots+n}{n^2}$

(10) $\lim\limits_{n\to\infty} \dfrac{1^2+2^2+\cdots+n^2}{n^3}$　(11) $\lim\limits_{n\to\infty} a^{\frac{1}{n}}$，其中$a > 0$　(12) $\lim\limits_{n\to\infty} \dfrac{n^{10000}}{a^n}$，其中$a > 1$

(13) $\lim\limits_{n\to\infty} \dfrac{2^n+n}{3^n+n^2}$　(14) $\lim\limits_{n\to\infty} \dfrac{3^n+2^n}{3^n+n^2}$　(15) $\lim\limits_{n\to\infty} \sqrt{n}(\sqrt{n+1}-\sqrt{n})$

(16) $\lim\limits_{n\to\infty} \dfrac{n^{10000}}{a^n}$，其中$a > 1$,　(17) $\lim\limits_{n\to\infty} \left(1-\dfrac{1}{n}\right)^n$　(18) $\lim\limits_{n\to\infty} \left(1-\dfrac{1}{5n}\right)^{n+2019}$

(19) $\lim\limits_{n\to\infty} (n^3+n^2+9n+1)^{\frac{1}{n}}$　(20) $\lim\limits_{n\to\infty} (2018^n+2019^n)^{\frac{1}{n}}$

习题 C（Riemann 重排定理）

Riemann 证明下面有趣的定理：如果实数项级数 $\sum\limits_{n=1}^{\infty} a_n$ 收敛但不绝对收敛，那么可以将级数重新排列，使得重排后的级数可以收敛到任意事先指定的 $\alpha \in \mathbb{R}\cup\{-\infty, +\infty\}$。假设 $\varphi : \mathbb{Z}_{\geqslant 1} \to \mathbb{Z}_{\geqslant 1}$ 是正整数到自身的双射，令 $b_k = a_{\varphi(k)}$，序列 $\{b_k\}_{k\geqslant 1}$ 被称为是 $\{a_n\}_{n\geqslant 1}$ 的一个**重排**，级数 $\sum\limits_{k=1}^{\infty} b_k$ 被称为是级数 $\sum\limits_{n=1}^{\infty} a_n$ 的一个**重排**。

我们将 $\{a_n\}_{n\geqslant 1}$ 中的非负项（$\geqslant 0$）的全体按照它们在 $\{a_n\}_{n\geqslant 1}$ 中的先后次序排列得到序列 $c_1, c_2, c_3, \cdots$；类似地，将 $\{a_n\}_{n\geqslant 1}$ 中的负项（< 0）按原顺序排列得到序列 $d_1, d_2, d_3, \cdots$。

C1) 证明，$\lim\limits_{n\to\infty} c_n = 0$，$\lim\limits_{n\to\infty} d_n = 0$。

C2) 证明，$\sum_{n=1}^{\infty} c_n = +\infty$，$\sum_{n=1}^{\infty} d_n = -\infty$。

C3) 证明: 对任意 $\alpha \in \mathbb{R}$，存在级数 $\sum_{n=1}^{\infty} a_n$ 的一个重排 $\sum_{k=1}^{\infty} b_k$，使得 $\sum_{k=1}^{\infty} b_k = \alpha$。

C4) 证明: 存在级数 $\sum_{n=1}^{\infty} a_n$ 的一个重排 $\sum_{k=1}^{\infty} x_k$，使得 $\sum_{k=1}^{\infty} x_k = +\infty$。

习题 D（Cesàro 求和极限）

设 $\{a_n\}_{n\geqslant 1}$ 为实数序列，我们定义算数平均值序列 $\sigma_n = \dfrac{a_1 + a_2 + \cdots + a_n}{n}$，其中 $n = 1, 2, 3, \cdots$。

D1) 假设 $\lim_{n\to\infty} a_n = a$。证明，$\lim_{n\to\infty} \sigma_n = a$。

D2) 构造一个不收敛的序列 $\{a_n\}_{n\geqslant 1}$，使得 $\lim_{n\to\infty} \sigma_n = 0$。

D3) 是否存在序列 $\{a_n\}_{n\geqslant 1}$, 使得对任意 $n \geqslant 1$, 都有 $a_n > 0$ 并且 $\limsup_{n\to\infty} a_n = \infty$ 然而 $\lim_{n\to\infty} \sigma_n = 0$?

D4) 对 $k \geqslant 1$, 记 $b_k = a_{k+1} - a_k$。证明，对任意的 $n \geqslant 2$，都有 $a_n - \sigma_n = \dfrac{1}{n}\sum_{k=1}^{n-1} k b_k$。

D5) 设 $\lim_{k\to\infty} k b_k = 0$ 并且 $\{\sigma_n\}_{n\geqslant 1}$ 收敛。证明，$\{a_n\}$ 也收敛。
注意，这是 1) 在条件 $n|a_{n+1} - a_n| \to 0$ 这一额外条件下的逆命题。

D6) 把 5) 的条件减弱为: $\{kb_k\}_{k\geqslant 1}$ 是有界的并且 $\lim_{n\to\infty} \sigma_n = \sigma$。证明，$\lim_{n\to\infty} a_n = \sigma$。

课后补充

习题 E（$\sqrt[n]{x}$ 和 b^x 的定义）

这个问题的目表是在 $\mathbb{R}$ 上定义初等函数 $\sqrt[n]{x}$ 和 b^x，比如说我们自然希望定义 $\sqrt[n]{x}$ 为满足 $y^n = x$ 唯一的正实数。

E1) 给定正整数 n 和实数 $x > 0$，证明: 如果正实数 y_1 和 y_2 满足 $y_1^n = x = y_2^n$，那么 $y_1 = y_2$。

E2) 证明，如果 $x > 0$，那么集合 $E(x) = \{t \in \mathbb{R} | t^n < x\}$ 是非空的并且有上界。

E3) 证明，$y = \sup E(x)$ 满足 $y^n = x$ 并且 $y > 0$。

E4) 证明，映射 $\sqrt[n]{\cdot} : \mathbb{R}_{>0} \to \mathbb{R}_{>0}$，$x \mapsto \sqrt[n]{x} = y$ 是良好定义的。我们也记 $\sqrt[n]{x}$ 为 $x^{\frac{1}{n}}$。

E5) 证明，$\sqrt[n]{\cdot} : \mathbb{R}_{>0} \to \mathbb{R}_{>0}$ 是双射。

E6) a, b 是正实数，n 为正整数。证明，$(ab)^{\frac{1}{n}} = a^{\frac{1}{n}} b^{\frac{1}{n}}$.

在接下来的问题里，我们给定实数 $b>1$ 来定义以 b 为底的指数函数 $x\mapsto b^x$。

E7) 设 $m,n,p,q\in\mathbb{Z}$，其中 $n>0$，$q>0$。令 $r=\dfrac{m}{n}=\dfrac{p}{q}$ 为有理数 r 的两种表示。证明：

$$(b^m)^{\frac{1}{n}}=(b^p)^{\frac{1}{q}}.$$

E8) 证明，对任意的有理数 r，函数 $r\mapsto b^r$ 是良好定义的。

E9) 证明，对任意的有理数 r,s，我们有 $b^{r+s}=b^rb^s$。

E10) 对于 $x\in\mathbb{R}$，令 $B(x)=\{b^t|t\in\mathbb{Q},t\leqslant x\}$。证明，$B(x)$ 非空且有上界。我们定义 $b^x=\sup B(x)$，这就定义了映射 $x\mapsto b^x$。

E11) 证明，如果 r 是有理数，那么

$$b^r=\sup B(r),\qquad \forall r\in\mathbb{Q}.$$

我们定义的指数映射在 $r\in\mathbb{Q}$ 时和之前是一致的。

E12) 证明，E11) 中定义的映射满足，对任意的 $x,y\in\mathbb{R}$，都有 $b^{x+y}=b^xb^y$。

E13*) 证明，当 $b=e$ 时，这样定义的函数和课程中定义的 e^x 是一致的。

习题 F（根式的逼近）

给定正实数 α 和初始值 $x_1>\sqrt{\alpha}$，我们归纳地定义序列 $\{x_n\}_{n\geqslant1}$：

$$x_{n+1}=\frac{1}{2}\left(x_n+\frac{\alpha}{x_n}\right),\qquad n=1,2,\cdots. \tag{A1}$$

F1) 证明，$\{x_n\}_{n\geqslant1}$ 是递减的并且 $\lim\limits_{n\to\infty}x_n=\sqrt{\alpha}$（在问题 E 中已经定义）。这表明，从任意大于 $\sqrt{\alpha}$ 的初值出发，可用上述迭代公式近似地计算（逼近）$\sqrt{\alpha}$。

F2) 定义逼近的误差项 $\varepsilon_n=x_n-\sqrt{\alpha}$。证明，$\varepsilon_{n+1}=\dfrac{\varepsilon_n^2}{2x_n}<\dfrac{\varepsilon_n^2}{2\sqrt{\alpha}}$。

F3) 证明，如果 $\beta=2\sqrt{\alpha}$，那么 $\varepsilon_{n+1}<\beta\left(\dfrac{\varepsilon_1}{\beta}\right)^{2^n}$。这表明，迭代公式 (A1) 收敛速度非常快。

F4) 设 $\alpha=3$，$x_1=2$，验证 $\dfrac{\varepsilon_1}{\beta}<0.1$，继而 $\varepsilon_5<4\cdot10^{-16}$，$\varepsilon_6<4\cdot10^{-32}$。

F5) 试用纸和笔，计算 $\sqrt{3}$ 的精确到小数点 5 位的近似值。

接下来，我们换另外一个迭代公式。固定 $\alpha>1$ 和 $y_1>\sqrt{\alpha}$，我们归纳地定义

$$y_{n+1}=\frac{\alpha+y_n}{1+y_n}=y_n+\frac{\alpha-y_n^2}{1+y_n}\qquad n=1,2,\cdots \tag{A2}$$

F6) 证明，$\{y_{2k-1}\}_{k\geqslant 1}$ 为递减序列。

F7) 证明，$\{y_{2k}\}_{k\geqslant 1}$ 为递增序列。

F8) 证明，$\lim\limits_{n\to\infty} y_n = \sqrt{\alpha}$。

F9) 试讨论迭代公式 (A2) 逼近 $\sqrt{\alpha}$ 的收敛速度并与 (A1) 的比较。

思考题 (不交作业)

问题 G（Banach-Mazur）

老王和王老饭后玩一个 $\mathbb{R}$ 上的区间套游戏决定谁来付饭钱：老王先选取一个闭区间 W_1，然后王老选一个 W_1 的子区间 L_1，但是要求 L_1 的长度不能超过 W_1 的一半；然后老王再选一个 L_1 的子区间 W_2，又轮到王老再选一个 W_2 的子区间 L_2，同样要求 L_2 的长度不超过 W_2 的一半；如此下去，第 n 步，老王选一个 L_{n-1} 的子区间 W_n，又轮到王老再选一个 W_n 的子区间 L_n，但是要求 L_n 的长度不超过 W_n 的一半。他们两个人得到一个区间套：

$$W_1 \supset L_1 \supset W_2 \supset L_2 \supset \cdots \supset W_n \supset L_n \supset \cdots .$$

老王和王老发现 $\bigcap\limits_{n\geqslant 1} W_n = \bigcap\limits_{n\geqslant 1} L_n = \{x\}$ 是一个实数。他们规定，如果 x 是有理数那么老王赢，如果 x 是无理数就是王老赢，试问最后谁会付钱?

问题 H

考虑数列的集合 $\mathcal{P} = \left\{\{p_n\}_{n\geqslant 1} \middle| p_n \in \mathbb{Z}, p_1 \geqslant 2, p_{n+1} \geqslant (p_n)^2\right\}$。

H1) 对任意 $p = \{p_n\}_{n\geqslant 1} \in \mathcal{P}$，我们定义数列

$$a_n = \prod_{k=1}^{n} (1 + \frac{1}{p_n}).$$

证明，a_n 有极限（记 $f(p) = \lim\limits_{n\to\infty} a_n$）并且 $f(p) \in (1, 2]$。

H2) 证明，$f : \mathcal{P} \to (1, 2]$ 是双射。

H3) 证明，$\mathcal{P}$ 是不可数的。

问题 I（2-进制展开）

考虑数列的集合 $\mathcal{S} = \left\{\{s_n\}_{n\geqslant 0} | s_n \in \{-1, 1\}\right\}$。

I1) 对任意 $s = \{s_n\}_{n\geqslant 1} \in \mathcal{S}$，我们定义数列

$$c_n = \sum_{k=0}^{n} \frac{s_1 s_2 \cdots s_k}{2^k}.$$

证明，c_n 有极限（记 $h(s) = \lim\limits_{n\to\infty} c_n$）并且 $h(s) \in [-2, 2]$。

I2) 证明，$h : \mathcal{S} \to (1, 2]$ 是满射。请问这是单射么?

I3) 对于 $s = \{s_n\}_{n\geqslant 1} \in \mathcal{S}$，证明，

$$2\sin(\frac{\pi}{4}c_n) = s_0\sqrt{2 + s_1\sqrt{2 + s_2\sqrt{2 + \cdots + s_{n-1}\sqrt{2 + s_n}}}}$$

I4) 计算极限

$$\lim_{n\to\infty} \underbrace{\sqrt{2 + \sqrt{2 + \sqrt{2 + \cdots + \sqrt{2}}}}}_{n\ 个\ 2}.$$

问题 J*

$k \geqslant 2$ 是一个给定的正整数。数列由如下的方式归纳地定义:

$$a_0 > 0 \text{ 预先给定，} a_{n+1} = a_n + \frac{1}{\sqrt[k]{a_n}}, \quad n \geqslant 0.$$

证明，$\lim\limits_{n\to\infty} \dfrac{(a_n)^{k+1}}{n^k}$ 存在并计算它的值。



国庆节快乐!

7 乘积级数与 Riemann ζ-函数，振荡级数的收敛判断（Dirichlet 与 Abel 判别法），完备距离空间，完备赋范线性空间，Picard 不动点定理／压缩映像定理，等价距离／范数，矩阵的指数映射

二零一九年九月三十日，星期一，晴天

例子 (无限乘积)**.** $\{a_n\}_{n\geqslant 1}$ 是复数的数列，我们假设对任意的 n，$a_n\neq 0$。令 $P_n=a_1\cdot a_2\cdot\cdots\cdot a_n$，如果数列 $\{P_n\}_{n\geqslant 1}$ 的极限存在并且该极限不是 0，我们就称无限乘积 $\prod\limits_{n\geqslant 1}a_n$ 收敛并且记 $\prod\limits_{n\geqslant 1}a_n=\lim\limits_{n\to\infty}P_n$。作为练习，我们有这种情况下的 *Cauchy* 判别准则：

练习. 证明，$\prod\limits_{n\geqslant 1}a_n$ 收敛当且仅当对任意的 $\varepsilon>0$，存在 N，使得对任意的 $n\geqslant N$，任意的 $p\geqslant 0$，我们都有

$$\left|a_n\cdot a_{n+1}\cdot\cdots\cdot a_{n+p}-1\right|<\varepsilon.$$

证明: 假设 $\lim\limits_{n\to\infty}P_n=P\neq 0$。特别地，存在 N_0，使得当 $n\geqslant N_0$ 时，$|P_n|\geqslant\dfrac{P}{2}$；根据 Cauchy 判别准则，存在 N_1，使得对任意的 $n\geqslant N_1$，任意的 $p\geqslant 0$，我们都有 $|P_{n-1}-P_{n+p}|<\varepsilon$。从而，当 $n\geqslant\max(N_0,N_1)$ 时，我们有

$$|P_{n-1}-P_{n+p}|<\varepsilon\ \Leftrightarrow\ |P_{n-1}|\left|a_n\cdot a_{n+1}\cdot\cdots\cdot a_{n+p}-1\right|<\varepsilon.$$

从而，

$$\left|a_n\cdot a_{n+1}\cdot\cdots\cdot a_{n+p}-1\right|<\frac{\varepsilon}{P_{n-1}}<\frac{2\varepsilon}{|P|}.$$

这完成了命题的一个方面。

反之，假设对任意的给定的 $\varepsilon>0$，存在 N，使得对任意的 $n\geqslant N$，任意的 $p\geqslant 0$，我们都有

$$\left|a_n\cdot a_{n+1}\cdot\cdots\cdot a_{n+p}-1\right|<\varepsilon.$$

我们要证明 $\{P_n\}_{n\geqslant 1}$ 是 Cauchy 列。此时，我们注意到 $\{P_n\}_{n\geqslant 1}$ 首先是有界的序列，这因为如果我们取 $\varepsilon=1$，那么存在 N，使得对任意的 $n\geqslant N$，任意的 $p\geqslant 0$，我们都有

$$\left|a_n\cdot a_{n+1}\cdot\cdots\cdot a_{n+p}-1\right|<1\Rightarrow|a_n\cdot a_{n+1}\cdot\cdots\cdot a_{n+p}|<2.$$

所以，对任意的 p，我们都有 $P_{N+p}<2P_N$，所以整个序列有界，假设 $M=\sup\limits_{n\geqslant 1}|P_n|$。

此时，我们对于选取好的 ε，有

$$|P_{n-1}-P_{n+p}|=|P_{n-1}|\left|a_n\cdot a_{n+1}\cdot\cdots\cdot a_{n+p}-1\right|<M\varepsilon.$$

所以，$\{P_n\}_{n\geqslant 1}$ 是 Cauchy 列，从而 $\lim\limits_{n\to\infty}P_n$ 存在。

最终，我们要说明 $\lim\limits_{n\to\infty}P_n\neq 0$。如果不然，那么 $\lim\limits_{n\to\infty}P_n\neq 0$，首先取 $\varepsilon=\frac{1}{2}$，根据刚才的证明，存在 N，当 $n\geqslant N$ 时，我们有

$$\left|\frac{P_n}{P_N}-1\right|<\frac{1}{2}.$$

固定 N，$\lim\limits_{n\to\infty}P_n\neq 0$，在上面不等式中令 $n\to\infty$ 就得到了矛盾。 □

我们首先证明关于正数序列的存在定理：

引理 29. $\{a_n\}_{n\geqslant 1}$ 是正数的数列，即对任意的 n，$a_n>0$。那么，如下两个命题是等价的：

(P) 无限乘积 $\prod_{n=1}^{\infty}(1+a_n)$ 收敛（极限必然不是零）；

(S) 级数 $\sum_{n=1}^{\infty}a_n$ 收敛。

推论 30 (绝对收敛的类比)**.** $\{a_n\}_{n\geqslant 1}$ 是复数的数列，如果 $\prod_{n=1}^{\infty}(1+|a_n|)$ 收敛，那么，$\prod_{n=1}^{\infty}(1+a_n)$ 收敛（极限不是零）。特别地，如果级数 $\sum_{n=1}^{\infty}a_n$ 绝对收敛，那么 $\prod_{n=1}^{\infty}(1+a_n)$ 收敛。

证明: (P)⇒(S) 是显然的，这是因为部分和有界：

$$\sum_{k=1}^{n}a_k\leqslant\prod_{k=1}^{n}(1+a_k)\leqslant\prod_{n=1}^{\infty}(1+a_n).$$

现在证明 (S)⇒(P)，我们要用到指数函数的 exp 的性质：$1+x<e^x$，其中 $x>0$。对于部分乘积，我们有

$$\prod_{k=1}^{n}(1+a_k)\leqslant\prod_{k=1}^{n}e^{a_k}\leqslant\exp\big(\sum_{k=1}^{n}a_k\big)\leqslant\exp\big(\sum_{k=1}^{\infty}a_k\big).$$

这也是有界的而部分乘积 $\{P_n=\prod_{k=1}^{n}(1+a_k)\}_{n\geqslant 1}$ 是单调上升的序列，所以有极限。 □

推论的证明. 只要证明第一个论断即可：我们有如下的不等式

$$\big|\prod_{k=n}^{n+p}(1+a_k)-1\big|\leqslant\prod_{k=p}^{n+p}(1+|a_k|)-1.$$

利用 Cauchy 判别准则立得结论。 □

我们还有一个有趣（有用）的习题：

练习. 数列 $\{a_n\}_{n\geqslant 1}$ 满足对任意的 n，$1>a_n>0$。特别地，无限乘积 $\prod_{n=1}^{\infty}(1-a_n)$ 的部分乘积总是收敛的（递减）。证明，$\prod_{n=1}^{\infty}(1-a_n)$ 收敛（$\neq 0$）当且仅当级数 $\sum_{n=1}^{\infty}a_n$ 收敛。

证明: 如果 $\prod_{n=1}^{\infty}(1-a_n)$ 收敛但是级数 $\sum_{n=1}^{\infty}a_n$ 发散，那么我们有

$$(1-a_1)\cdots(1-a_n)\leqslant\frac{1}{(1+a_1)\cdots(1+a_n)}\leqslant\frac{1}{1+a_1+\cdots a_n}\to 0.$$

从而，$\prod_{n=1}^{\infty}(1-a_n)=0$，矛盾。所以，级数 $\sum_{n=1}^{\infty}a_n$ 收敛。反过从级数收敛证明乘积收敛是显然的。 □

我们现在假定大家熟悉算数基本定理，即每个正整数都可以唯一地写成若干个素数的乘积。

定理 31 (Euler 乘积公式)**.** 假设 $\mathcal{P}$ 是所有的素数组成的集合，我们可以将它们从小到大排列为 $2=p_1<p_2<p_3<\cdots$。对于 $s>1$，ζ-函数

$$\zeta(s)=\sum_{n=1}^{\infty}\frac{1}{n^s}$$

是良好定义的。那么，我们有

$$\zeta(s)=\prod_{p\in\mathcal{P}}\frac{1}{1-p^{-s}},$$

其中上面的乘积是收敛的（我们可以假定是按照素数从小到大的顺序来做乘积的）。

（我们将在连续函数一节定义幂函数 x^s，这里我们不妨假定它满足中学大家学过的基本性质）

证明: 我们已经证明过，当 $s>1$ 时，$\displaystyle\sum_{n=1}^{\infty}\frac{1}{n^s}$ 收敛。我们先来研究乘积。令 $\displaystyle P_n=\prod_{k\leqslant n}\frac{1}{1-p_k^{-s}}$。首先，我们注意到，我们有

$$\frac{1}{1-(p_k)^{-s}}=\sum_{e=1}^{\infty}\frac{1}{p_k^{es}},$$

这是一个绝对（正项）收敛级数，所以，根据我们已经证明的关于绝对收敛级数乘积的公式（不依赖于重排），我们有

$$P_N=\prod_{k\leqslant N}\frac{1}{1-p_k^{-s}}=\sum_{e_1,e_2,\cdots,e_N=1}^{\infty}\frac{1}{p_1^{e_1s}p_2^{e_2s}\cdots p_N^{e_Ns}}=\sum_{e_1,e_2,\cdots,e_N=1}^{\infty}\frac{1}{\left(p_1^{e_1}p_2^{e_2}\cdots p_N^{e_N}\right)^s}.$$

所以，如果令 $\mathbf{N}_N=\left\{n\in\mathbb{Z}_{\geqslant 1}\middle| n\text{的素因子分解中出现的素数}\leqslant p_N\right\}$，我们就有

$$P_N=\prod_{k\leqslant n}\frac{1}{1-p_k^{-s}}=\sum_{n\in\mathbf{N}_N}\frac{1}{n^s}.$$

从而，我们有

$$0\leqslant\zeta(s)-P_N=\sum_{n\text{含有}>p_N\text{的素因子}}\frac{1}{n^s}<\sum_{n>p_N}\frac{1}{n^s}.$$

（注意到如果只有有限个素数，那么上式的右端是零，所以证明就结束了）由于 $\zeta(s)$ 是收敛的，所以当 $N\to\infty$ 时，我们有 $\lim\limits_{N\to\infty}(\zeta(s)-P_N)=0$，这就完成了证明。 □

根据庄子的想法（再一次！这就为什么调和级数发散的证明我们需要好好理解），我们对于比较接近于 1 的 $s>1$，给出 $\zeta(s)$ 的下界估计：

$$\begin{aligned}\sum_{k=1}^{\infty}\frac{1}{n^s}&=\underbrace{\frac{1}{1^s}}_{1\text{个},\ \geqslant\frac{1}{2^s}}+\underbrace{\left(\frac{1}{2^s}+\frac{1}{3^s}\right)}_{2\text{个},\ \geqslant\frac{1}{4^s}}+\cdots+\underbrace{\left(\frac{1}{2^{(k-1)s}}+\frac{1}{\left(2^{k-1}+1\right)^s}+\cdot+\frac{1}{\left(2^k-1\right)^s}\right)}_{2^{k-1}\text{个},\ \geqslant\frac{1}{2^{ks}}}+\cdots\\&\geqslant\sum_{k=1}^{\infty}\frac{1}{2^{ks}}\times 2^{k-1}=\frac{1}{2}\frac{1}{1-\frac{1}{2^{s-1}}}.\end{aligned}$$

假设我们知道当 $s \to 1$ 时，$2^{s-1} \to 1$（这个在后面学习连续函数的时候就会证明），从而当 s 离着 1 很近的时候，$\zeta(s)$ 的值可以很大（$\to \infty$）。根据乘积公式，如果 $\mathcal{P}$ 时有限集合，那么，选取一串 $s_\ell \to 1$，其中 $s_\ell > 1$，那么，我们有

$$\zeta(s_\ell) = \prod_{p\in\mathcal{P}} \frac{1}{1-p^{-s_\ell}} \overset{\ell\to\infty}{\longrightarrow} \prod_{p\in\mathcal{P}} \frac{1}{1-p^{-1}} = 1.$$

这给出了矛盾。所以有无限多个素数。这种方法的进一步发展可以给出 $\leqslant n$ 的素数的个数的渐进公式（第 n 个素数的大小大约应该是 $n\log(n)$），这就是著名的素数定理，首先由 *Hadamard* 和 *de la Vallée Poussin* 分别于 *1896* 年独立证明。

振荡型的级数收敛判别

我们目前接触的数列或级数收敛的判定的基本模式都是“控制”，即用某个较大且收敛的对象来作为上界从而说明本来问题的收敛性。我们下面将看到另一种模式的收敛判别准则，这对应着积分理论中振荡积分的想法。作为例子，我们回忆在第二次的作业中已经证明的结论（Leibniz 的结果）：

$\{a_n\}_{n\geqslant 1}$ 是递减的正数数列并且 $\lim\limits_{n\to\infty} a_n = 0$。那么，交错级数 $a_1 - a_2 + a_3 - a_4 + \cdots$ 收敛。

通过这个例子，我们有这样的直观：如果一个级数有正有负（**振荡的**：把它们画成 (n, a_n) 就可以看到这个数列对于 x 轴是上下波动的），那么正负可以相互抵消很多，这样求和的时候仍然可以收敛。

我们先回忆一下经典的 **Abel 求和公式**：

引理 32. $\{a_k\}_{k\geqslant 1}$ 和 $\{b_k\}_{k\geqslant 1}$ 是复数（矩阵）的序列，那么

$$\sum_{k=1}^{n} a_k b_k = S_n b_n + \sum_{k=1}^{n-1} S_k(b_k - b_{k+1}),$$

其中我们用 $S_n = a_1 + a_2 + \cdots + a_n$ 表示数列 $\{a_k\}_{k\geqslant 1}$ 的部分和。

证明: 证明是平凡的：在和式 $\sum\limits_{k=1}^{n} a_k b_k$ 中，将 a_k 替换成 $S_k - S_{k-1}$（我们要求 $S_0 = 0$），然后按照 S_k 的指标合并同类项即可。 □

Abel 求和公式是离散版本的分部积分公式，它在（振荡的）级数的收敛理论中有着非凡的应用：

定理 33 (Dirichlet 判别法)**.** 给定实数数列 $\{a_k\}_{k\geqslant 1}$ 和 $\{b_k\}_{k\geqslant 1}$，用 S_n 表示 $\{a_k\}_{k\geqslant 1}$ 的部分和。假设它们满足如下条件：

1) $\{b_k\}_{k\geqslant 1}$ 是单调数列并且 $\lim\limits_{k\to\infty} b_k = 0$；

2) 存在 M，使得对任意的 $n \geqslant 1$，$|S_n| \leqslant M$。

那么，级数 $\sum_{k=1}^{\infty} a_k b_k$ 收敛。

上面的第二个条件说的是 $\{a_k\}_{k\geqslant 1}$ 是振荡并且正负相互抵消很多！

证明：我们利用级数收敛的 Cauchy 判别准则：任取 n, m，其中 $m \geqslant n$，那么，根据 Abel 求和法

$$\sum_{n+1\leqslant k\leqslant m} a_k b_k = S_m b_m - S_n b_n + \sum_{k=n}^{m-1} S_k (b_k - b_{k+1}).$$

上式中，前两项 $S_m b_m$ 和 $S_n b_n$ 在 $n, m \to \infty$ 时自然趋于 0，这是因为 S_n 有界，$b_n \to 0$。为了处理上式中最后一项的和式，我们需要利用单调性。我们不妨假设 $\{b_n\}_{n\geqslant 1}$ 是单调下降的，那么

$$\begin{aligned} \Big| \sum_{k=n}^{m-1} S_k (b_k - b_{k+1}) \Big| &\leqslant \sum_{k=n}^{m-1} |S_k| (b_k - b_{k+1}) \\ &\leqslant M \sum_{k=n}^{m-1} (b_k - b_{k+1}) = M(b_n - b_m). \end{aligned}$$

很明显，当 $n, m \to \infty$ 时，这一项的极限也是零。 □

作为 Dirichlet 判别法的推论，我们有

定理 34 (Abel 判别法)**.** 实数数列 $\{a_k\}_{k\geqslant 1}$ 和 $\{b_k\}_{k\geqslant 1}$ 满足如下条件：

1) $\{b_k\}_{k\geqslant 1}$ 是单调有界数列；

2) 级数 $a_1 + a_2 + a_3 + \cdots$ 收敛。

那么，级数 $\sum_{k=1}^{\infty} a_k b_k$ 收敛。

证明：令 $b = \lim_{k\to\infty} b_k$，那么，我们可以将 $\sum_{k=1}^{\infty} a_k b_k$（形式上）写成：

$$\sum_{k=1}^{\infty} a_k b_k = \sum_{k=1}^{\infty} a_k (b_k - b) + b \sum_{k=1}^{\infty} a_k.$$

上述等式右边的第一个级数满足 Dirichlet 判别法的要求，所以是收敛的，从而 $\sum_{k=1}^{\infty} a_k b_k$ 收敛。 □

作为上判别法的应用，我们考虑一个（最）经典的级数收敛问题（这里我们假设我们已经知道 $\sin(x)$ 和 $\cos(x)$ 在 $\mathbb{R}$ 上的零点的位置，这个事实我们会在本学期后面的课程给出证明）：

例子. 研究级数 $\sum_{n=1}^{\infty} \frac{\cos(nx)}{n}$ 的收敛性问题，其中 $x \in \mathbb{R}$。（我们先假设 $\cos(x)$ 满足我们中学学习过的各种性质）。

我们这里可以选取 $a_n=\cos(nx)$，$b_n=\dfrac{1}{n}$，先计算部分和：

$$S_N=\sum_{1\leqslant n\leqslant N}\cos(nx)=\frac{\sin(N+\frac{1}{2})x-\sin\frac{x}{2}}{2\sin\frac{x}{2}}.$$

我们这里假设 $x\notin 2\pi\mathbb{Z}$，所以 $|S_N|\leqslant|\dfrac{1}{\sin(\frac{x}{2})}|$ 是有界的，根据 Dirichlet 判别法，级数是收敛的。如果 $x\in 2\pi\mathbb{Z}$，我们级数变成了 $\displaystyle\sum_{n=1}^{\infty}\frac{1}{n}$，即调和级数，自然不收敛。

练习. 证明，如下两个级数都是收敛的：

$$\sum_{b=1}^{\infty}\frac{(-1)^n}{n}\left(1+\frac{1}{n}\right)^n,\quad\sum_{n=1}^{\infty}\frac{(-1)^n}{n}\cos(\frac{1}{n}).$$

复习与进阶：距离空间和赋范线性空间（矩阵空间）上的收敛与极限

给定距离空间 (X,d)，我们也有 Cauchy 列的概念：如果点列 $\{x_n\}_{n\geqslant 1}$ 使得对任意的 $\varepsilon>0$，总存在 N，对任意的 $n,m\geqslant N$，都有 $d(x_m,x_n)<\varepsilon$，就称该点列为 **Cauchy 列**。然而，在一般的距离空间中，Cauchy 列未必收敛，我们举两个例子：

1) $(\mathbb{Q},d(x,y)=|x-y|)$ 不是完备的度量空间。我们可以选取 $x_n=\displaystyle\sum_{k=0}^{n}\frac{1}{k!}$，这是一个 Cauchy 列，我们已经证明过这个序列在 $\mathbb{Q}$ 中不收敛（我们用的是反证法，并且那个证明不需要事先知道 e 的存在性）。

2) $X=\mathbb{R}-\{0\}$，$d(x,y)=|x-y|$。选取数列 $\{\dfrac{1}{n}\}_{n\geqslant 1}$，它是 Cauchy 列但是不收敛。

3)* X=[0,1] 上的实系数多项式函数, $d(f,g)=\sup\limits_{x\in[0,1]}|f(x)|$。这个空间不完备。在闭区间上，多项式可以收敛到任意的连续函数（Weierstrass 逼近定理）。我们将在后来的作业中证明这个著名的定理。

定义 35. (X,d) 是距离空间，如果 (X,d) 中的每个 *Cauchy* 列都收敛，那么我们就称 X 是**完备的**距离空间。

很明显，假设 (X,d) 是完备的距离空间，如果将 d 换成与它相等价的距离 d'，那么 (X,d') 也是完备的。

定理 36. $(\mathbb{R}^n,d_2)$ 是完备的距离空间，其中 $d_2(x,y)=\sqrt{\displaystyle\sum_{j=1}^{n}(x_j-y_j)^2}$。

证明: 假设 $\{x^{(k)}\}_{k\geqslant 1}$ 是 $\mathbb{R}^n$ 中的 Cauchy 列。先固定指标 $j\in\{1,\cdots,n\}$，那么 $x^{(k)}$ 的第 j 个坐标 $x_j^{(k)}$ 显然满足 $|x_j^{(m_1)}-x_j^{(m_2)}|\leqslant d(x_{m_1},x_{m_2})$。所以，作为实数数列 $\{x_j^{(k)}\}_{k\geqslant 1}$ 也是 Cauchy 列，根据 $\mathbb{R}$ 的完备性，这个数列收敛。我们已经证明过（作业二的 A3)）$\mathbb{R}^n$ 的序列收敛当且仅当其所有坐标的序列收敛，这就完成了证明。 □

定理 37 (压缩映像定理／ Banach 或 Picard 不动点定理)**.** (X,d) 是完备的 距离空间。假设映射 $T:X\to X$ 是**压缩映射**，即存在常数 $0<\gamma<1$，使得对任意的 $x,x'\in X$，我们都有

$$d(T(x),T(x'))\leqslant\gamma d(x,x').$$

那么，T 必有唯一的不动点，即存在唯一的 $x_*\in X$，使得 $T(x_*)=x_*$。

证明: 任意选取 $x_0\in X$，考虑点列 $\{x_n\}_{n\geqslant 1}$，其中 $x_n=T^n(x)$。我们说明它是 Cauchy 列：任选自然数 n 和 p，我们有

$$\begin{aligned}d(x_n,x_{n+p})&=d(T^n(x_0),T^{n+p}(x_0))\leqslant\gamma^n d(x_0,T^p(x_0))\\&\leqslant\gamma^n\big(d(x_0,T(x_0))+d(T(x_0),T^2(x_0))+\cdots+d(T^{p-1}(x_0),T^p(x_0))\big)\\&\leqslant\gamma^n\big(d(x_0,T(x_0))+\gamma d(x_0,T(x_0))+\cdots+\gamma^{p-1}d(x_0,T(x_0))\big)\\&\leqslant\frac{\gamma^n}{1-\gamma}d(x_0,x_1).\end{aligned}$$

由于 $\gamma<1$，所以当 $n\to\infty$ 时，上式可以任意小，所以 $\{x_n\}_{n\geqslant 1}$ 是 Cauchy 列。根据 (X,d) 的完备性，存在 $x_*\in X$，使得 $\lim\limits_{n\to\infty}x_n=x_*$。

现在说明 $T(x_*)=x_*$。任意选取 $\varepsilon>0$，存在 m，使得 $d(x_m,x_*)<\varepsilon$ 并且 $d(x_{m+1},x_m)<\varepsilon$。根据三角不等式，我们有

$$\begin{aligned}d(T(x_*),x_*)&\leqslant d(T(x_*),T(x_m))+d(T(x_m),x_m)+d(x_m,x_*)\\&\leqslant\gamma d(x_*,x_m)+d(x_{m+1},x_m)+d(x_m,x_*)\leqslant(\gamma+2)\varepsilon.\end{aligned}$$

所以，$d(T(x_*),x_*)=0$，即 $T(x_*)=x_*$。

最终我们来证明不动点是唯一的：假设 x_* 和 $x_\bullet$ 均为不动点，那么

$$0\leqslant d(x_*,x_\bullet)=d(T(x_*),T(x_\bullet))\leqslant\gamma d(x_*,x_\bullet).$$

由于 $0<\gamma<1$，所以 $d(x_*,x_\bullet)=0$，从而 $x_*=x_\bullet$。 □

练习. *1)* (X,d) 是完备的度量空间，$T:X\to X$ 是映射，$r\in\mathbb{Z}_{\geqslant 1}$ 是正整数。假设 T^r 是压缩映射，证明，T 必有唯一的不动点。

2) 考虑映射 $f:\mathbb{R}\to\mathbb{R}$, $x\mapsto e^{-x}$。证明，f 不是压缩映像但是 $f\circ f$ 是压缩映像。

在分析学习中，最经常用到的距离空间的实际上都是赋范线性空间，我们现在就引入这个概念。

定义 38 (赋范线性空间)**.** V 是 $\mathbb{R}$ (或者 $\mathbb{C}$) 上的线性空间 (可以是无限维的)。假设 $\|\cdot\|: V \to \mathbb{R}_{\geqslant 0}$ 是 V 上的满足如下三条性质的函数:

1) 对任意的 $x \in V$, $\lambda \in \mathbb{R}$, 有 $\|\lambda x\| = |\lambda|\|x\|$;

2) $\|x\| = 0$ 当且仅当 $x = 0$;

3) 对任意的 $x, y \in V$, $\|x + y\| \leqslant \|x\| + \|y\|$。

我们称 $\|\cdot\|$ 是 V 上的一个**范数**并且称 $(V, \|\cdot\|)$ 是一个**赋范线性空间**。

直观上, $\|v\|$ 就是计算向量 v 的某种长度。

注记. $(V, \|\cdot\|)$ 是赋范线性空间, 那么 $(V, d(x, y) = \|x - y\|)$ 是一个距离空间。如果我们不另加注释的话, 我们总是以这种方式把一个赋范线性空间看作是一个距离空间。

特别地, 在赋范线性空间 $(V, \|\cdot\|)$ 上, 点列 $\{x_n\}_{n\geqslant 1}$ 收敛到 $x \in V$ 的含义是 $\lim\limits_{n\to\infty} \|x_n - x\| = 0$。

我们有如下常见的赋范线性空间 (请注意, 它们都是有限维的, 最重要的一个无限维的例子需要等到我们引入了连续函数才介绍):

例子. *1)* 我们在 $\mathbb{R}^n$ 或者 $\mathbb{C}^n$ 上面可以定义三种范数:

$$\|x\|_1 = \sum_{j=1}^{n} |x_j|,$$

$$\|x\|_2 = \sqrt{\sum_{j=1}^{n} (x_j)^2},$$

$$\|x\|_\infty = \sup_{i=1,2,\cdots,n} |x_i|.$$

其中, $x = (x_1, \cdots, x_n)$。

2) 我们在 $\mathbf{M}_n(\mathbb{R})$ 或者 $\mathbf{M}_n(\mathbb{C})$ (视作是 $\mathbb{R}^{n^2}$ 或者 $\mathbb{C}^{n^2}$) 上也有三个范数: 对任意的 $n \times n$ 的矩阵 $A = (A_{ij})$, 其中 $1 \leqslant i, j \leqslant n$, 我们有

$$\|A\|_1 = \sum_{1\leqslant i,j\leqslant n} |A_{ij}|,$$

$$\|A\|_2 = \sqrt{\sum_{1\leqslant i,j\leqslant n} (A_{ij})^2},$$

$$\|A\|_\infty = \sup_{1\leqslant i,j\leqslant n} |A_{ij}|.$$

这三种范数都是上面线性空间情况直接搬过来的。实际上, 利用矩阵的知识, 我们还可以定义其他更有几何意义的范数, 比如说我们可以利用矩阵的特征值来定义一些范数, 限于知识范围, 我们不再展开。

假设 $\|\cdot\|_1$ 和 $\|\cdot\|_2$ 均为线性空间 V 上的范数，如果存在常数 $C_1>0$ 和 $C_2>0$，使得对任意的 $x\in X$，都有

$$\|x\|_2\leqslant C_1\|x\|_1,\quad \|x\|_1\leqslant C_2\|x\|_1,$$

那么就称这两个范数 $\|x\|_1$ 和 $\|x\|_2$ 是**等价的**。根据上面的注记，等价的范数自然给出等价的距离函数，所以它们定义出的收敛和极限的概念是一致的。

如果一个赋范线性空间是完备的距离空间，我们就称它为**完备的赋范线性空间**或者是 **Banach 空间**，我们上面看到的例子（$\mathbb{R}^n$ 以及矩阵的空间）都是完备的赋范线性空间。

我们已经在矩阵空间上 $\mathbf{M}_n(\mathbb{R})$ 建立了距离结构并说明这是一个完备的空间，作为应用，我们要定义在矩阵空间上的指数函数。为此，我们要在这个场合下把我们对于实数数列熟知的结论再按照之前学习的顺序验证一下。

除了定义之外（根据完备性，极限的判断可以用 Cauchy 判别准则），我们对极限所研究的第一个有意义的就问题是求极限这个操作是否和代数操作交换。事实上，假设 $(V,\|\cdot\|)$ 是赋范线性空间，那么作为线性空间 V 中的两个点能做加法（特别地，我们还可以通过用加法加上取极限的操作来定义级数），我们自然要问：

$$\lim_{n\to\infty}(x_n+y_n)=\lim_{n\to\infty}x_n+\lim_{n\to\infty}y_n,\quad \lim_{n\to\infty}\lambda x_n=\lambda\lim_{n\to\infty}x_n,$$

是否成立。这个问题非常简单：只要把数列的证明过程中的绝对值符号 $|\cdot|$ 换成 $\|\cdot\|$ 即可！

某些赋范线性空间上是可以定义乘法的，比如说 $n\times n$ 的矩阵空间 $\mathbf{M}_n(\mathbb{R})$，这样的空间中的两个点能相乘，所以在这种情况下，我们想知道下面的公式是否成立：

$$\lim_{n\to\infty}(x_n\cdot y_n)=\lim_{n\to\infty}x_n\cdot\lim_{n\to\infty}y_n.$$

不妨设 $\lim\limits_{n\to\infty}x_n=x$ 和 $\lim\limits_{n\to\infty}y_n=y$。在数列的情况下，我们的证明如下：

$$|x_n\cdot y_n-x\cdot y|\leqslant|(x_n-x)y_n|+|x(y_n-y)|\leqslant|x_n-x||y_n|+|x||y_n-y|.$$

然后，利用 y_n 的有界性，我们用一个常数 M 来控制 x 和 y_n，就有

$$|x_n\cdot y_n-x\cdot y|\leqslant M(|x_n-x|+|y_n-y|)\to 0.$$

证明的第一个不等号用到了三角不等式，这对任意赋范线性空间的范数 $\|\cdot\|$ 都成立；第二个不等号用到了绝对值和乘法之间的关系 $|x\cdot y|=|x||y|$；第三步用到了 y_n 是有界的。

在矩阵的空间里，$|x\cdot y|=|x||y|$ 可能不成立，比如说，我们用 $\|\cdot\|_\infty$-范数，那么，

$$\|A\cdot B\|_\infty=\sup_{i,j\leqslant n}\Big|\sum_{k=1}^{n}A_{ik}B_{kj}\Big|\leqslant\sum_{k=1}^{n}\|A\|_\infty\|B\|_\infty=n\|A\|_\infty\|B\|_\infty.$$

即对任意的 $x,y\in X$，我们都有

$$\|x\cdot y\|\leqslant C\|x\|\|y\|,$$

其中 C 不依赖于 x 和 y 的选取，所以仿照数列的证明，我们就有

$$\begin{aligned} d(x_n \cdot y_n, x \cdot y) = \|x_n \cdot y_n - x \cdot y\| &\leqslant \|(x_n - x)y_n\| + \|x(y_n - y)\| \\ &\leqslant C\big(\|x_n - x\|\|y_n\| + \|x\|\|y_n - y\|\big) \\ &\leqslant CM\big(\|x_n - x\| + \|y_n - y\|\big). \end{aligned}$$

其中 $\|y_n\| \leqslant M$ 这个有界性是明显的。据此，我们知道乘法和极限是交换的。

注记. 我们（自豪地）发现，上述做法根本没涉及矩阵的分量，我们可以将矩阵视为一个整体（不需要考虑分量）进行操作，就像对数字的序列一样。

最终，我们来定义 e^A，其中，A 是一个 $n \times n$ 的实（复）系数矩阵。令

$$\exp : \mathbf{M}_n(\mathbb{R}) \to \mathbf{M}_n(\mathbb{R}), \quad A \mapsto e^A := \sum_{k=0}^{\infty} \frac{A^k}{k!}.$$

为了说明 e^A 是良好定义的，我们验证级数 $\displaystyle\sum_{k=0}^{\infty} \frac{A^k}{k!}$ 是收敛的（即部分和在这个赋范线性空间或者距离空间里收敛）。我们也想仿照数列的情况利用**绝对收敛的级数**的概念。所以，对于级数 $\displaystyle\sum_{k=0}^{\infty} x_k$，我们定义，如果 $\displaystyle\sum_{k=0}^{\infty} \|x_k\|$ 是收敛的（加范数后变成了我们熟悉的数项级数），就称级数 $\displaystyle\sum_{k=0}^{\infty} x_k$ 是**绝对收敛的**。有了这样的动机，我们自然要证明下面的命题：

命题 39. 假设 $\{x_k\}_{k\geqslant 1}$ 是完备的赋范线性空间 $(V, \|\cdot\|)$ 中的点列。如果级数 $\displaystyle\sum_{k=0}^{\infty} x_k$ 绝对收敛，那么它收敛。

证明和数列的情况完全一致，我们把它留做作业题。

为了证明绝对收敛性，我们做如下的估计：

$$\sum_{k=0}^{\infty} \|\frac{A^k}{k!}\| \leqslant \sum_{k=0}^{\infty} \frac{C^{k-1}\|A\|^k}{k!} \leqslant \sum_{k=0}^{\infty} \frac{C^k\|A\|^k}{k!} = e^{C\|A\|}.$$

其中，我们用到了 $\|A^k\| \leqslant C^{k-1}\|A\|^k$（不妨设 $C \geqslant 1$）。

注记. 假设 A 和 B 是 $n \times n$ 的（复）矩阵，那么 $e^{A+B} \overset{?}{=} e^A \cdot e^B$ 可能未必成立，因为如果 $A \cdot B \neq B \cdot A$，那么上面的红色等号可能就不成立了。请参考 $e^{z_1+z_2} = e^{z_1}e^{z_2}$ 的证明。

练习. 如果 A 是 $n \times n$ 的（复）矩阵，那么对任意的 $s, t \in \mathbb{C}$，我们都有 $e^{sA} \cdot e^{tA} = e^{(s+t)A}$。特别地，证明，$e^{-A} = \big(e^A\big)^{-1}$。

我们还可以仿照数的情形定义 $\cos(A)$ 和 $\sin(A)$，当然，此时这些函数所要表达的几何意义已经不再清楚（但是为什么一定要清楚呢）。最后，我们要说明，e^A 在我们后面的课程中大有用处。

至此，我们关于极限部分的介绍结束，下一部分我们要研究函数的连续性。

7.1 作业：素数的倒数和，Basel 问题的 Euler“证明”

清华大学 19-20 秋季学期，数学分析一，作业 3

请用 A4 大小的纸张**正反面**用钢笔，签字笔或者圆珠笔书写，并注明自己的姓名，年级和作业的**总页数**。除定理公式所涉及的人名之外，请使用**中文**。本次作业的提交时间和地点为 **10 月 17 日**上午的课堂上，逾期视作零分。

基本习题

习题 A：连续函数的定义和基本性质

A1)（函数极限的 Cauchy 判别准则）给定函数 $f:(a,x_0)\cup(x_0,b)\to\mathbb{R}$。那么，$f$ 在 x_0 处有极限当且仅当对任意的 $\varepsilon>0$，存在 $\delta>0$，使得对任意的 $x_1,x_2\in(x_0-\delta,x_0+\delta)$，都有 $|f(x_1)-f(x_2)|<\varepsilon$。

A2) 假设 I 是一个非空的区间并且 I 不是一个点，证明，I 上的连续函数 $C(I)$ 所构成的 $\mathbb{R}$-线性空间是无限维的。

A3)（重要）假设 (X,d_X) 和 (Y,d_Y) 是距离空间，$f:X\to Y$ 是映射。假设 $x_0\in X$，如果对任意 X 中的点列 $\{x_n\}_{n\geqslant 1}$，$x_n\xrightarrow{d_X}x_0$，我们都有 $f(x_n)\xrightarrow{d_Y}f(x_0)$，我们就称 f 在 x_0 处是连续的。如果 f 在一切 $x\in X$ 处均连续，那么我们就称 f 是距离空间之间的**连续映射**。假设 (X,d_X)，(Y,d_Y) 和 (Z,d_Z) 是三个距离空间，$f:X\to Y$，$g:Y\to Z$ 均为连续映射。证明，它们的复合 $g\circ f:X\to Z$ 也是连续映射。

A4) 假设 (X,d_X) 和 (Y,d_Y) 是距离空间，$f:X\to Y$ 是连续映射。如果 d'_X 是与 d_X 等价的距离，d'_Y 是与 d_Y 等价的距离，那么，对于 (X,d'_X) 和 (Y,d'_Y) 而言，f 也是连续映射。

A5) 在 $\mathbb{R}^n$ 上我们配有常见的距离，比如说 d_2（请参考之前的讲义）；(X,d_X) 是距离空间。我们将映射 $f:X\to\mathbb{R}^n$ 写成分量的形式：

$$f:X\to\mathbb{R}^n,\quad x\mapsto f(x)=(f_1(x),f_2(x),\cdots,f_n(x)),$$

其中 $f_i:X\to\mathbb{R}$ 是函数。证明，f 是连续映射当且仅当对所有的 $i=1,\cdots,n$，f_i 是连续函数。

A6)（重要）假设 (X,d_X) 是距离空间，$(V,\|\cdot\|)$ 是赋范线性空间，$f:X\to V$ 和 $g:X\to V$ 是连续映射。证明，它们的和与差（自然的定义）$f\pm g:X\to V$ 也是连续映射。如果 $V=\mathbb{C}$（或者 $n\times n$ 的矩阵构成的赋范线性空间），那么 $f\cdot g:X\to\mathbb{C}$ 是连续映射。如果 $V=\mathbb{C}$ 并且对任意的 $x\in X$，$g(x)\neq 0$，那么 $\dfrac{f}{g}:X\to\mathbb{C}$ 是连续映射。你可以选择上面的一个情况来证明。

A7) 试找出 $\mathbb{R}$ 上定义的函数

$$f(x)=\begin{cases}\frac{1}{q}, & \text{如果 } x=\dfrac{p}{q} \text{ 是有理数，其中 } p\in\mathbb{Z},\ q\in\mathbb{Z}_{\geqslant 1} \text{ 且 } p \text{ 和 } q \text{ 互素;}\\ 0, & \text{如果 } x \text{ 是无理数.}\end{cases}$$

的所有不连续点。

A8) 计算 $\lim\limits_{x\to 0}\dfrac{e^x-1}{x}$。

A9) 计算 $\lim\limits_{x\to+\infty}(1+\dfrac{1}{x})^x$。

A10) 计算 $\lim\limits_{x\to-\infty}(1+\dfrac{1}{x})^x$.

习题 B（级数的基本计算技巧和收敛判断）

B1) 试计算下列级数（计算技术的训练）

(1) $\displaystyle\sum_{n=1}^{\infty}\frac{1}{n(n+1)}$　(2) $\displaystyle\sum_{n=1}^{\infty}\frac{1}{4n^2-1}$　(3) $\displaystyle\sum_{n=1}^{\infty}\frac{1}{n(n+1)(n+2)}$

(4) $\displaystyle\sum_{n=1}^{\infty}\arctan\frac{1}{n^2+n+1}$　(5) $\displaystyle\sum_{n=0}^{\infty}\frac{(-1)^n+2}{3^n}$　(6) $\displaystyle\sum_{n=1}^{\infty}\frac{n}{3^n}$

(7) $\displaystyle\sum_{n=1}^{\infty}\frac{(-1)^{n-1}}{2^{n-1}}$　(8) $\displaystyle\sum_{n=1}^{\infty}\frac{2n-1}{2^n}$　(9) $\displaystyle\sum_{n=0}^{\infty}\frac{2n+1}{n^2(n+1)^2}$

(10) $\displaystyle\sum_{n=1}^{\infty}(\sqrt{n+2}-2\sqrt{n+1}+\sqrt{n})$　(11) $\displaystyle\sum_{n=1}^{\infty}\log\left(\frac{n(2n+1)}{(n+1)(2n-1)}\right)$　(12) $\displaystyle\sum_{n=1}^{\infty}\frac{1}{n(n+m)}$　$(m\in\mathbb{N})$.

B2) 试判断下列级数的收敛性。

(1) $\displaystyle\sum_{n=1}^{\infty}(\sqrt{n+1}-\sqrt{n})$　(2) $\displaystyle\sum_{n=1}^{\infty}\frac{\sqrt{n+1}-\sqrt{n}}{n}$　(3) $\displaystyle\sum_{n=2}^{\infty}(\sqrt[n]{n}-1)^n$

(4) $\displaystyle\sum_{n=1}^{\infty}\frac{1}{1+x^n}$　$(x\neq -1)$　(5) $\displaystyle\sum_{n=1}^{\infty}\frac{1}{n2^n}$　(6) $\displaystyle\sum_{n=1}^{\infty}\left(\frac{n^2}{3n^2+1}\right)^n$

(7) $\displaystyle\sum_{n=1}^{\infty}\frac{1}{n^{1+\frac{1}{n}}}$　(8) $\displaystyle\sum_{n=2}^{\infty}\frac{1}{(\log n)^{\log n}}$　(9) $\displaystyle\sum_{n=1}^{\infty}\frac{n^{n+\frac{1}{n}}}{(n+\frac{1}{n})^n}$

(10) $\displaystyle\sum_{n=1}^{\infty}(-1)^{n-1}\frac{\sqrt{n}}{n+1}$　(11) $\displaystyle\sum_{n=1}^{\infty}\frac{(-1)^{n-1}}{\sqrt[n]{n}}$　(12) $\displaystyle\sum_{n=1}^{\infty}\left(1+\frac{1}{2}+\cdots+\frac{1}{n}\right)\frac{\sin nx}{n}$.

B3) 试判断下列级数的收敛性并确定它们是否绝对收敛。

(1) $\displaystyle\sum_{n=2}^{\infty}(-1)^n\frac{1}{n\log n}$　(2) $\displaystyle\sum_{n=2}^{\infty}\frac{\sin\frac{n\pi}{4}}{\log n}$

3) $\displaystyle\sum_{n=1}^{\infty}(-1)^n\frac{n-1}{n+1}\frac{1}{\sqrt[3]{n}}$　(4) $\displaystyle\sum_{n=1}^{\infty}(-1)^{\frac{n(n-1)}{2}}\frac{n^{10}}{a^n}$　$(a>1)$

习题 C （递减正项级数的凝聚检验法）假设正整数 $b \geqslant 2$，$f : (1, \infty) \to \mathbb{R}_{>0}$ 是递减（未必严格递减）函数。

C1) 证明下面的不等式：

$$(b-1)b^{k-1}f(b^k) \leqslant \sum_{j=b^{k-1}}^{b^k-1} f(j) \leqslant (b-1)b^{k-1}f(b^{k-1}).$$

C2) 证明：级数

$$\sum_{n=1}^{\infty} f(n) \qquad \text{和} \qquad \sum_{n=1}^{\infty} b^n f(b^n)$$

同时收敛或者同时发散 (我们称通过后者的收敛来判断前者收敛的方法为**凝聚检验法**)。

C3) 证明，$\displaystyle\sum_{n=2}^{\infty} \frac{1}{n \log n}$ 是发散的。

C4) 证明，$\displaystyle\sum_{n=100}^{\infty} \frac{1}{(n \log n)(\log(\log n))}$ 是发散的。

C5) 证明，如果 $s > 1$，那么 $\displaystyle\sum_{n=1}^{\infty} \frac{1}{n^s}$ 是收敛的；如果 $0 < s < 1$，那么 $\displaystyle\sum_{n=1}^{\infty} \frac{1}{n^s}$ 是发散的。

C6) 假设 $s > 1$。证明，$\displaystyle\sum_{n=2}^{\infty} \frac{1}{n(\log n)^s}$ 和 $\displaystyle\sum_{n=10}^{\infty} \frac{1}{(n \log n)(\log(\log n))^s}$ 都是的收敛。

注记. *上述的几个收敛的结论（C3)-C6)）是标准并且重要的，记住这几个结论对于收敛以及函数大小的理解很有帮助。这里的证明尽管巧妙，但是等我们接触到积分的时候，我们就可以用统一的、简单的、更本质的也更容易记忆的方法来证明这些结论。*

习题 D 实数列 $\{a_n\}_{n\geqslant 1}$ 的上下极限的刻画。先补充几个定义：

- $\alpha \in \mathbb{R}$，如果对任意 $\varepsilon > 0$, 存在无穷多 n，使得 $a_n \in (\alpha - \varepsilon, \alpha + \varepsilon)$，我们称 α 为 $\{a_n\}_{n\geqslant 1}$ 的一个**聚点**；
- 如果对任意 $M > 0$，存在无穷多个 n，使得 $a_n \in (M, \infty)$，我们就称 $+\infty$ 为 $\{a_n\}_{n\geqslant 1}$ 的一个**聚点**；
- 如果对任意 $M > 0$，存在无穷多个 n，使得 $a_n \in (-\infty, -M)$，我们就称 $-\infty$ 为 $\{a_n\}_{n\geqslant 1}$ 的一个**聚点**。

D1) 证明，$\alpha \in \mathbb{R}$ 是 $\{a_n\}_{n\geqslant 1}$ 的聚点当且仅当该数列有子序列 $\{a_{n_k}\}_{k\geqslant 1}$ 收敛于 α。

D2) 证明，$+\infty$ 是 $\{a_n\}_{n\geqslant 1}$ 的聚点当且仅当该数列有子序列 $\{a_{n_k}\}_{k\geqslant 1}$ 使得 $\lim\limits_{k\to\infty} a_{n_k} = +\infty$。

D3) 令 $E = \{\alpha \in \mathbb{R} \cup \{\pm\infty\} \,|\, \alpha \text{ 是}\{a_n\}_{n\geqslant 1} \text{ 的聚点}\}$ 为 $\{a_n\}_{n\geqslant 1}$ 的所有聚点所组成的集合（它是 $\mathbb{R} \cup \{\pm\infty\}$ 的子集）。证明，$E \neq \emptyset$。

D4) 证明：$E \subset \mathbb{R}$ 当且仅当数列 $\{a_n\}_{n\geqslant 1}$ 有界。

D5) 假设 $\{a_n\}_{n\geqslant 1}$ 有界。试证明，$\sup E = \limsup\limits_{n\to\infty} a_n$，$\inf E = \liminf\limits_{n\to\infty} a_n$。

D6) 假设 $\{a_n\}_{n\geqslant 1}$ 有界。令 $a^* = \limsup\limits_{n\to\infty} a_n$，证明如下两个命题：

i) $a^* \in E$（所以 $\sup E \in E$）；

ii) 对任意 $x > a^*$，存在 $N \in \mathbb{Z}_{\geqslant 1}$，使得对任意 $n > N$，都有 $a_n < x$。

D7) 试举一个数列 $\{a_n\}_{n\geqslant 1}$ 作为例子，使得 $E \cap \mathbb{R} \neq \emptyset$ 且 $E \not\subseteq \mathbb{R}$。

D8) 试举一个数列 $\{a_n\}_{n\geqslant 1}$ 作为例子，使得 E 为无穷集。

思考题 (不交作业)

问题 E（素数的倒数和）根据第七次课的内容，对于 $s > 1$，我们可以定义 ζ-函数

$$\zeta(s) = \sum_{n=1}^{\infty} \frac{1}{n^s}$$

并且证明了 Euler 乘积公式：

$$\zeta(s) = \prod_{p\in\mathcal{P}} \frac{1}{1-p^{-s}},$$

其中 $\mathcal{P}$ 是全体素数的集合。据此证明，级数

$$\sum_{p\in\mathcal{P}} \frac{1}{p^s}$$

对于 $s > 1$ 是收敛的；对于 $0 < s \leqslant 1$ 是发散的（这给出了有无穷多个素数的另一个证明）。

问题 F（Basel 问题的 Euler“证明”）对任意的 $\theta \in \mathbb{R}$ 和 $n \in \mathbb{Z}_{\geqslant 1}$，试证明恒等式

$$\frac{\sin\big((2n+1)\theta\big)}{(2n+1)\sin\theta} = \prod_{k=1}^{n} \left(1 - \frac{\sin^2(\theta)}{\sin^2\left(\frac{k\pi}{2n+1}\right)}\right).$$

据此证明，对任意的 $x \in \mathbb{R}$，我们都有

$$\frac{\sin(\pi x)}{\pi x} = \prod_{n=1}^{\infty} \left(1 - \frac{x^2}{n^2}\right).$$

（如果 $x = 0$，我们将左边定义为极限 $\lim\limits_{x\to 0} \dfrac{\sin(\pi x)}{\pi x}$）

注记. 根据上面的公式，形式上，我们就有

$$\sin(\pi x)=\pi x\prod_{n=1}^{\infty}\left(1-\frac{x^2}{n^2}\right)=\pi x\left(1-\left(\sum_{n=1}^{\infty}\frac{1}{n^2}\right)x^2+(\cdots)x^4+(\cdots)x^6+\cdots\right).$$

所以，右边 x^3 的系数就是 $-\pi\sum_{n=1}^{\infty}\frac{1}{n^2}$。根据我们对于 sin 的定义

$$\sin z=\frac{e^{iz}-e^{-iz}}{2i}=\sum_{k=0}^{\infty}\frac{(-1)^k z^{2k+1}}{(2k+1)!}.$$

所以，$\sin(\pi x)$ 中展开式的 x^3 项的系数是 $-\frac{\pi^3}{6}$。比较系数，我们得到所谓的 *Basel* 问题的解：

$$\sum_{n=1}^{\infty}\frac{1}{n^2}=\frac{\pi^2}{6}.$$

这是 *Euler* 原始的想法，他就是利用了上面的乘积公式来猜测最终的极限是 $\frac{\pi^2}{6}$。有趣的是，他观察到 $\frac{\sin(\pi x)}{\pi x}$ 的零点恰好是 $x=\pm n$，所以如果 $\frac{\sin(\pi x)}{\pi x}$ 的行为与多项式类似的话，那么这个函数应该是单项式的乘积，所以 *Euler* 认为

$$\frac{\sin(\pi x)}{\pi x}=c\prod_{n\in\mathbb{Z}-\{0\}}\left(1-\frac{x}{n}\right)=c\prod_{n=1}^{\infty}\left(1-\frac{x^2}{n^2}\right),$$

其中 c 是一个待定的常数，通过取 $x\to 0$ 的极限，他计算出 $c=1$。

8 函数的连续性

二零一九年十月十日，星期四，晴天

函数的连续性

在开始讨论函数之前，我们先澄清和回忆几个概念：

- 当谈论一个函数 f 的时候，我们坚持原则：要说清楚 f 是定义在哪里又是在哪里取值的，即我们用映射 $f: X \to Y$ 的观点来看函数 f，通常用的 $f(x)$ 不是一个好的记号。
- 给定两个函数 $f: X \to \mathbb{R}$ 和 $g: X \to \mathbb{R}$（或者 $\mathbb{C}$，或者是其他的赋范线性空间或者距离空间），其中 X 是某个空间（集合），那么函数 $f+g$ 指的是映射

$$f+g: X \to \mathbb{R}, \quad x \mapsto f(x)+g(x).$$

  类似地，我们也可以定义函数 $f+g$，$f \cdot g$ 等，这里不在赘述。

- 如果函数 f 在 $\mathbb{R}$ 或者 $\mathbb{R}$ 的子集上面定义并且在 $\mathbb{R}$ 中取值，利用 $\mathbb{R}$ 上的序关系，我们可以定义**单调性**，即递增或者递减的函数，精确的定义留给同学自己叙述。

函数的连续性有两个等价的定义，一种用数列的语言的来描述，一种用所谓的 $\varepsilon-\delta$ 语言（在数学分析中，ε 和 δ 通常代表两个很小（任意小）的正数）。

定义 40 (函数的左右极限，左右连续性和连续性；数列的语言)**.** 假设 $a<b$ 是实数，$I=(a,b)\subset\mathbb{R}$ 是开区间，$x_0 \in I$。

1) 给定函数 $f:(a,x_0)\cup(x_0,b)\to\mathbb{R}$。

如果存在 $y_0\in\mathbb{R}$，使得对任意的 序列 $\{x_n^-\}_{n\geqslant 1}\subset(a,x_0)$，$\lim\limits_{n\to\infty}x_n^-=x_0$，都有 $\lim\limits_{n\to\infty}f(x_n^-)=y_0$，我们就称 f 在 x_0 处有**左极限**y_0，并记作 $\lim\limits_{x\to {x_0}^-}f(x)=y_0$。（我们注意到 y_0 不依赖于序列 $\{x_n^-\}_{n\geqslant 1}$ 的选取）

类似地，如果存在 $y_0\in\mathbb{R}$，使得对任意的序列 $\{x_n^+\}_{n\geqslant 1}\subset(x_0,b)$，$\lim\limits_{n\to\infty}x_n^+=x_0$，都有 $\lim\limits_{n\to\infty}f(x_n^+)=y_0$，我们就称 f 在 x_0 处有**右极限**y_0，并记作 $\lim\limits_{x\to {x_0}^+}f(x)=y_0$。

如果存在 $y_0\in\mathbb{R}$，使得对任意的序列 $\{x_n\}_{n\geqslant 1}\subset(a,x_0)\cup(x_0,b)$，$\lim\limits_{n\to\infty}x_n=x_0$，都有 $\lim\limits_{n\to\infty}f(x_n)=y_0$，我们就称 f 在 x_0 处有**极限**y_0，并记作 $\lim\limits_{x\to x_0}f(x)=y_0$。

2) 给定函数 $f:I\to\mathbb{R}$。如果对任意的序列 $\{x_n\}_{n\geqslant 1}\subset I$，$\lim\limits_{n\to\infty}x_n=x_0$，都有 $\lim\limits_{n\to\infty}f(x_n)=f(x_0)$，我们就称 f 在 x_0 处**连续**。如果 f 在每个 $x_0\in I$ 处都连续，我们就称 f 是（I 上的）**连续函数**。如果 $J=[a,b)$，$f:J\to\mathbb{R}$ 在 a 处的右极限存在并且恰好等于 $f(a)$，我们就称 f 在 a 处连续；类似地，我们可以定义在闭区间 $[a,b]$ 的端点处连续的函数进而定义在闭区间上的连续函数。

注记. 我们还可以"望文生义地"定义左连续或者右连续的概念：给定函数 $f: I \to \mathbb{R}$，如果 f 在 x_0 处有左极限并且左极限为 $f(x_0)$，那么我们称 f 在 x_0 处是**左连续的**。右连续性可以类似定义。

练习. *1)* 给定函数 $f: (a, x_0) \cup (x_0, b) \to \mathbb{R}$。证明，$f$ 在 x_0 有极限当且仅当 f 在 x_0 处的左右极限都存在并且相等。

2) 给定函数 $f: I \to \mathbb{R}$。证明，f 在 $x_0 \in I$ 处连续当且仅当 f 在 x_0 处的即是左连续的也是右连续的。

我们可以用这个练习来熟悉一下连续性的第一种定义，证明只需要用数列收敛的概念即可。

定义 41 (函数的极限和连续性; $\varepsilon-\delta$-语言)**.** 假设 $a < b$ 是实数，$I = (a, b) \subset \mathbb{R}$ 是开区间，$x_0 \in I$。考虑在 $I - \{x_0\}$ 上定义的实数（可以是复数）值函数 f，即 $f: (a, x_0) \cup (x_0, b) \to \mathbb{R}$，如果存在 $y_0 \in \mathbb{R}$，使得对**任意的** $\varepsilon > 0$（我们总是默认 ε 很小），存在 $\delta > 0$（通常 ε 总是取的很小），使得对**任意**满足 $|x - x_0| < \delta$ 的 x，都有 $|f(x) - y_0| < \varepsilon$，我们就称 f 在 x_0 处有**极限** y_0，并记作 $\lim\limits_{x \to x_0} f(x) = y_0$。如果假设 $f: I \to \mathbb{R}$ 在 x 处有极限并且 $\lim\limits_{n \to \infty} f(x_n) = f(x_0)$，我们就称 f 在 x_0 处**连续**。

注记. 我们类似地可以用 $\varepsilon - \delta$ 语言定义左右极限以及左右连续性，同学们应该自己尝试着来做这一点（基本是语言的游戏）或者查阅任何一本数学分析的参考书，这里不再赘述。

定理 42 (Heine)**.** 上述两种连续性的定义是等价的。

证明: 首先，假设 f 在 x_0 是在 $\varepsilon - \delta$-语言意义下连续的，我们现在证明，对任意的 $\{x_n\}_{n \geqslant 1} \subset I$，$x_n \to x_0$，我们都有 $f(x_n) \to f(x_0)$：事实上，对任意的 $\varepsilon > 0$，根据定义，存在 $\delta > 0$，使得当 $|x - x_0| < \delta$ 时，我们有 $|f(x) - f(x_0)| < \varepsilon$；根据 $x_n \to x_0$ 的定义，对于这个 δ，存在 N，使得当 $n \geqslant N$ 时，我们有 $|x_n - x_0| < \delta$。所以，当 $n \geqslant N$ 时，$|f(x_n) - f(x_0)| < \varepsilon_0$，从而 $f(x_n) \to f(x_0)$。

反过来，我们假设 f 在 x_0 在点列语言的意义下是连续的，即假设对任意的 序列 $\{x_n\}_{n \geqslant 1} \subset I$，$x_n \to x_0$，都有 $\lim\limits_{n \to \infty} f(x_n) = f(x_0)$。我们用反证法来证明 f 在 x_0 是在 $\varepsilon - \delta$-语言意义下连续的，即考虑逆否命题：如果不然，那么存在某个 $\varepsilon > 0$，对任意的 δ，特别地对于 $\delta = \dfrac{1}{n}$，存在 $|x_n - x_0| < \dfrac{1}{n}$，使得 $|f(x_n) - f(x_0)| \geqslant \varepsilon$。这样，我们就得到一个序列 $\{x_n\}$，使得 $x_n \to x_0$ 但是 $f(x_n) \not\to f(x_0)$，矛盾。 □

另外，关于函数在一点处的极限的存在性也有 $\varepsilon - \delta$ 版本的 Cauchy 判别准则：

定理 43 (Cauchy 判别准则)**.** 给定函数 $f: (a, x_0) \cup (x_0, b) \to \mathbb{R}$。那么，$f$ 在 x_0 处有极限当且仅当对任意的 $\varepsilon > 0$，存在 $\delta > 0$，使得对任意的 $x_1, x_2 \in (x_0 - \delta, x_0 + \delta)$，都有 $|f(x_1) - f(x_2)| < \varepsilon$。

证明: 证明是例行公事，只需要利用定义！我们将它留作本次的作业题之一。 □

注记. 我们还可以模拟上面的定义来定义函数 f 在 $x \to \pm\infty$ 时的极限：给定函数 $f: (a, +\infty) \to \mathbb{R}$，如果存在 $y_0 \in \mathbb{R}$，使得对任意的 $\varepsilon > 0$，存在 $M > 0$，使得对任意的 $x \in (M, +\infty)$，都有

$|f(x)-y_0|<\varepsilon$，我们就称 f 在 $+\infty$ 处的极限是 y_0，并记作 $\lim\limits_{x\to+\infty} f(x)=y_0$；类似地可以定义函数在 $-\infty$ 处的极限。

有了连续函数这个对象，我们就可以讨论数学（分析）中很重要的一个观点：

注记 (局部与整体). 局部，整体和一个点在连续函数的定义中是要分清楚的：

- 连续性是一个局部的概念：要搞清楚 f 在 x_0 处是否连续，只需要对任意小的 ε，理解 f 在 x_0 的一个小邻域（局部）$(x_0-\varepsilon, x_0+\varepsilon)$ 上的行为就足够了。这里，ε 可以任意得小，但不能是 0。
- 关于函数在整个区间 I 上是否连续的是整体的概念，因为我们需要知道 f 在每个点 附近的信息。
- f 在区间 I 上是否连续这个整体的性质是由局部性质决定的，即如果 f 在每个局部上连续，那么在整个区间 I 上就连续。数学上很多概念都是这样的：局部决定了整体，比如说函数如果局部上单调那么整体上就是单调函数；多项式（或者解析函数）是更极端的例子：根据代数基本定理，只要知道多项式 f 在一个点附近的值就可以完全确定这个多项式。当然，有很多数学对象的局部上完全不能决定整体，比如说用来描述经典力学的辛几何。
- 连续性是局部的概念但不是一个点处的概念：只知道函数在一个点 x_0 处的值是完全不可以判定函数在这个点处的连续性的！

通过上面局部与整体的讨论，我们可以很自然地引出所谓的开覆盖的概念：

1) (整体连续性意味着局部连续性) 假设 $J\subset I$ 是子区间，那么，如果 $f: I\to\mathbb{R}$ 在 I 上连续，f 在 J 上的限制 $f\big|_J$ 也连续，其中

$$f\big|_J: J\to\mathbb{R},\quad x\mapsto f(x).$$

2) (局部决定了整体) 假设 $\{I_k\}_{k\in K}$ 一族开区间，其中 K 是指标的集合，如果 $I\subset\bigcup\limits_{k\in K} I_k$，我们就称开区间族 $\{I_k\}_{k\in K}$ 覆盖了 I，我们也说 $\{I_k\}_{k\in K}$ 是 I 的一族**开覆盖**。假设对每个 $k\in K$，函数 f_k 在 I_k 上都有定义并且取实数值，如果任意的 $j,k\in K$，都有相容性条件：

$$f_k\big|_{I_j\cap I_k}=f_j\big|_{I_j\cap I_k},$$

那么下面的 f 在 I 上是良好定义的：由于对任意的 $x\in I$，x 一定属于某个 I_k，我们令 $f(x)=f_k(x)$。我们来说明 $f(x)$ 的定义不依赖于指标 k 的选取：如果 x_0 属于另一个 I_j，我们需要说明 $f_j(x)=f_k(x)$，这是因为 $x\in I_j\cap I_k$，相容性条件保证了这一点。

通过上面的构造方式，我们把在局部的小片 I_k 上定义的一族函数粘成了整体定义在 I 上的一个函数。

如果对任意的 k，f_k 在 I_k 连续，那么 f 在 I 上也连续，这因为对任意的 $x_0\in I$，x_0 一定属于某个 I_k，根据 f_k 在 I_k 上的连续性，f 就在 x_0 处连续，所以 f 连续。

最简单的连续函数的例子如下（请验证定义）：

例子. *1)* 常值函数 $f(x)=c$ 和 $f(x)=x$ 是连续函数。

2) 常值函数 $f(x)=c$ 和 $f(x)=x$ 是连续函数。

3) 指数函数 e^x 是 $\mathbb{R}\to\mathbb{R}$（或 $\mathbb{C}\to\mathbb{C}$）上的连续函数。

我们用 $\varepsilon-\delta$ 语言来证明：任意选定的 $x_0\in\mathbb{R}$。对任意的 $\varepsilon>0$，令 $\delta=\min\left(\dfrac{\varepsilon}{2e^{x_0+1}},1\right)$（这是一个事后诸葛亮的决定），当 $|x-x_0|<\delta$ 时，根据

$$e^x-e^{x_0}=e^{x_0}(e^{x-x_0}-1)=e^{x_0}\sum_{k=1}^{\infty}\frac{(x-x_0)^k}{k!},$$

我们有

$$\begin{aligned}|e^x-e^{x_0}|&\leqslant e^{x_0}\sum_{k=1}^{\infty}\frac{|x-x_0|^k}{k!}<e^{x_0}\sum_{k=1}^{\infty}\frac{\delta^k}{k!}\\&=\delta e^{x_0}\sum_{k=1}^{\infty}\frac{\delta^{k-1}}{k!}\leqslant\delta\big(e^{x_0}\sum_{k=0}^{\infty}\frac{1^{k-1}}{(k-1)!}\big)=\delta e^{x_0+1}<\varepsilon.\end{aligned}$$

这表明 exp 在 x_0 处连续。

要想构造更多的连续函数，我们就要利用关于连续函数的代数运算：

命题 44 (四则运算与序关系的交换性)**.** 假设实（复）值函数 f 和 g 在 x_0 附近定义（比方说在 $(x_0-\delta,x_0)\cup(x_0,x_0+\delta)$ 上定义）$x_0\in I$ 处有极限，那么

1) $f\pm g$ 在 $x_0\in I$ 处有极限并且 $\lim\limits_{x\to x_0}(f\pm g)(x)=\lim\limits_{x\to x_0}f(x)\pm\lim\limits_{x\to x_0}g(x)$。

2) $f\cdot g$ 在 $x_0\in I$ 处有极限并且 $\lim\limits_{x\to x_0}(f\cdot g)(x)=\lim\limits_{x\to x_0}f(x)\cdot\lim\limits_{x\to x_0}g(x)$。

2) 如果在 x_0 附近，$g(x)\neq0$，那么 $\dfrac{f}{g}$ 在 $x_0\in I$ 处有极限并且 $\lim\limits_{x\to x_0}\big(\dfrac{f}{g}\big)(x)=\dfrac{\lim\limits_{x\to x_0}f(x)}{\lim\limits_{x\to x_0}g(x)}$。

4) 如果对任意的 x，都有 $f(x)\leqslant g(x)$，那么 $\lim\limits_{x\to x_0}f(x)\leqslant\lim\limits_{x\to x_0}g(x)$。

证明: 连续性的表述有两种，我们这里选用连续性的数列语言，从而上面的性质只是数列相应性质的重新表述。 □

推论 45 (连续函数的四则运算)**.** 用 $C(I)$ 表示区间 I 上的连续函数的全体所构成的集合，那么 $\big(C(I),+,\cdot\big)$ 是一个环（连续函数环），即对任意的 $f,g\in C(I)$，我们有 $f\pm g,f\cdot g\in C(I)$。特别地，$C(I)$ 是 $\mathbb{R}$-线性空间（用常数函数 $f(x)\equiv c$ 作为数乘）。另外，如果对任意 $x\in I$，$g(x)\neq0$，那么 $\dfrac{f}{g}\in C(I)$。

注记. 我们注意到 $\big(C(I),+,\cdot\big)$ 与实数 $(\mathbb{R},+,\cdot)$ 有很多相似的地方（代数结构）。另外，除非 I 是一个点，$C(I)$ 是无限维的 $\mathbb{R}$-线性空间，我们把证明留成做本次的作业。

推论 46 (绝对值)**.** 如果 $f\in C(I)$，那么 $|f|\in C(I)$，其中 $|f|:I\to\mathbb{R}_{\geqslant 0}$，$x\mapsto|f(x)|$。

证明*:* 利用连续性的数列的描述立得。 □

定理 47 (函数的复合)**.** 给定 $f\in C(I;J)$（即 $f:I\to\mathbb{R}$ 是连续函数并且 f 的值域落在区间 J 中）和 $g\in C(J)$，那么复合函数 $(f\circ g)(x)=f(g(x))\in C(I)$。

证明*:* 我们利用数列的语言来证明，即证明当 $n\to\infty$ 时，对任意的 $x_n\to x_0$，我们有 $(f\circ g)(x_n)=f(g(x_n))\to f(g(x_0))$。令 $y_n=f(x_n)\in J$，其中 $n\geqslant 0$，由于 f 在 x_0 处连续，所以 $y_n\to y_0$；由于 g 在 y_0 处连续，所以 $g(y_n)\to g(y_0)$，这就是要证明的结论。 □

注记. 利用数列的语言进行证明是最直接最简单的，然而我们后来会意识到简单干净的证明可能未必是最好的。

利用上述四则运算以及函数的复合，我们可以构造一大类的连续函数：

例子. 假设 $I\subset\mathbb{R}$ 是一个给定的非空的区间，其中 $|I|\neq 0$，即它不是一个点。那么

1) 对任意的多项式 $P\in\mathbb{R}[X]$，即 $P(X)=a_dX^d+a_{d-1}X^{d-1}+\cdots+d_0$，其中 $a_i\in\mathbb{R}$，$0\leqslant i\leqslant d$，$a_d\neq 0$，d 是它的次数（我们要强调多项式 $\neq$ 多项式函数），我们可以将 P 视作是 I 上的函数，即对任意的 $x\in I$，我们令 $P(x)=a_dx^d+a_{d-1}x^{d-1}+\cdots+d_0$。那么，多项式函数 $P\in C(I)$。另外，如果 $Q\neq P\in\mathbb{R}[X]$ 是另一个多项式，那么在 $C(I)$ 中，作为函数 $Q\neq P$。在这个意义下，我们可以将多项式的全体 $\mathbb{R}[X]$ 视作是连续函数的子空间，即有单射映射 $\iota:\mathbb{R}[X]\hookrightarrow C(I)$ 并且这个映射保持两边的四则运算（环结构）：比如说，如果我们用 $\bullet$ 表示多项式的乘法，那么 $\iota(P\bullet Q)=\iota(P)\cdot\iota(Q)$，其中此式右边的乘法是函数的乘法。

2) 如果多项式 Q 在 I 上没有零点，那么有理函数 $\dfrac{P(x)}{Q(x)}\in C(I)$。

3) 三角函数 $\sin(x)$ 和 $\cos(x)$ 是连续函数，此时，我们隐含的用到了取复数值的指数函数具有连续性这个性质，我们后面会处理更一般的情况。

4) 利用复合函数保持连续性，我们知道函数 e^{x^2} 是 $\mathbb{R}$（或者 $\mathbb{C}$）上的连续函数。

注记. 不夸张的说，我们在数学中遇到（几乎）一切连续函数都是通过两种手段构造的：第一，通过连续函数的复合和四则运算；第二，通过逼近的方式，特别是级数的方式来定义，比如说 exp 的构造。这种逼近的方式是最值得我们注意的，我们很快会发现，$C(I)$ 这个空间和实数 $\mathbb{R}$ 很相似，构造无理数就是通过有理数逼近的方式。

更具体一点，我们会在 $C(I)$ 上面给定一个范数 $\|\cdot\|_\infty$ 并且证明这样得到的赋范线性空间是完备的。此时，任给 $f\in C(I)$，我们可以仿照实数的情况定义

$$e^f:=\sum_{n=0}^{\infty}\frac{f^k}{n!}.$$

在完备的赋范线性空间中，我们只要全盘照抄实数的情况就可以证明上面的（函数）级数收敛，从而 e^f 是良好定义的并且是连续函数。特别地，我们可以通过这种方式定义 e^x（把 x 看成是函数 f）而且说明这和我们最初定义的 $\exp(x)$ 是一码事。

在这种类比下，我们就可以利用对实数的直观来研究函数空间，从而得到很多关于函数的深刻结果。在课程后面的学习中，我们会遇到很具体的例子，比方说存在处处连续但是处处都不能微分的函数，我们就是通过构造函数的级数来实现的。

9 距离空间之间的连续映射，介值定理，初等函数的构造

二零一九年十月十四日，星期一，晴天

我们首先来复习／学习一下连续性的定义：

定义 48 (距离空间之间的连续映射)**.** 假设 (X,d) 和 (Y,d_Y) 是两个距离空间，$f:X\to Y$ 是这两个距离空间之间的映射。假设 $x_0\in X$，$y_0=f(x_0)\in Y$。如果对**任意的** $\varepsilon>0$，总存在 $\delta>0$，使得对**任意**满足 $d(x,x_0)<\delta$ 的 $x\in X$，都有 $d_Y\big(f(x),f(x_0)\big)<\varepsilon$，我们就称 f 在 x_0 处**连续**。如果 f 在 X 的每个点处都连续，那么我们就称 f 是**连续映射**。

注记. 当 Y 为 $\mathbb{R}$ 或者 $\mathbb{C}$ 的时候，我们就称 f 为连续函数。另外，我们有

1) 第三次作业中我们用点列的方式定义了连续映射，我们可以仿照 *Heine* 定理的证明，说明这个两个定义方式是等价的。我们将在第四次作业题中证明这个结论。

2) 假设 $X'\subset X$ 是子集，我们用 d' 表示 d 在 X 上诱导出来的距离函数，从而 (X',d') 是距离空间（参见作业一 *A3)*)。我们考虑限制映射

$$f\big|_{X'}:X'\to Y,\quad x'\in X'\mapsto f(x').$$

那么，f' 是 (X',d') 和 (Y,d_2) 之间的连续映射。简而言之，连续映射的限制仍是连续映射。

3) (距离空间的乘积与连续映射) 假设 (Y,d_Y) 和 (Z,d_Z) 是距离空间，我们定义 $Y\times Z$ 上的距离函数

$$d_{Y\times Z}:\big(Y\times Z\big)\times\big(Y\times Z\big)\to\mathbb{R}_{\geqslant 0},\quad \big((y_1,z_1),(y_2,z_2)\big)\mapsto\sqrt{d(y_1,y_2)^2+d(z_1,z_2)^2}.$$

不难验证，$d_{Y\times Z}$ 是 $Y\times Z$ 上的距离函数（当然，我们有很多其他方式定义新的距离函数，这里我们选取和勾股定理类似的一种选择)。此时，我们有两个自然的投影映射

$$\pi_Y:Y\times Z\to Y,\ (y,z)\mapsto y;\ \pi_Z:Y\times Z\to Z,\ (y,z)\mapsto z.$$

那么，π_Y 和 π_Z 都是连续映射。

另外，给定距离空间 (X,d) 和 $(Y\times Z,d_{Y\times Z})$ 之间的映射 $F:X\to Y\times Z$，那么，F 连续当且仅当两个复合映射 $\pi_Y\circ F:X\to Y$ 和 $\pi_Z\circ F:X\to Z$ 都连续（即到乘积空间的映射是连续的当且仅当其分量是连续的)。我们将在作业中证明这个性质。

作为例子，我们知道如果在 $\mathbb{R}^n$ 上配有距离 d_2（请参考之前的讲义)，那么

$$(\mathbb{R}^n,d_2)=\underbrace{(\mathbb{R},d_2)\times(\mathbb{R},d_2)\times\cdots\times(\mathbb{R},d_2)}_{n\text{个}}$$

假设 (X,d_X) 是距离空间，我们可以将映射 $f:X\to\mathbb{R}^n$ 写成分量的形式：

$$f:X\to\mathbb{R}^n,\quad x\mapsto f(x)=(f_1(x),f_2(x),\cdots,f_n(x)).$$

我们在第 3 次作业中已经证明了 f 连续当且仅当每个 $f_i : X \to \mathbb{R}$ 都连续。这当然是上面关于到乘积空间的映射的连续性的推论。

然而，我们要给出下面著名的反例来说明多个变量的函数如果对每个固定的变量都连续并不能说明函数本身连续：

$$f : \mathbb{R}^2 \to \mathbb{R}, \quad (x, y) \mapsto f(x, y) = \begin{cases} \frac{xy}{x^2+y^2}, & \text{如果 } (x, y) \neq (0, 0); \\ 0, & \text{如果 } (x, y) = (0, 0). \end{cases}$$

首先，对于任意的 x_0 固定，$f(x_0, y) : \mathbb{R} \to \mathbb{R}$ 作为 y 的函数是连续的；对于任意的 y_0 固定，$f(x, y_0) : \mathbb{R} \to \mathbb{R}$ 作为 x 的函数是连续的。这是显然的。另外，我们说明 f 在 $(0, 0)$ 处不连续：对任意的 $\lambda \neq 0$，我们可以取 $(x_k, y_k) = (\frac{1}{k}, \frac{\lambda}{k}) \to (0, 0)$，此时，$f(x_k, y_k) \equiv \frac{\lambda}{1+\lambda^2} \not\to 0$。

作为例子，我们研究矩阵空间上的映射：

例子. 映射 $\exp : \mathbf{M}_n(\mathbb{C}) \to \mathbf{M}_n(\mathbb{C})$。

首先回忆一下，$\mathbf{M}_n(\mathbb{C})$ 可以看成是一个 $2n^2$-维的 $\mathbb{R}$-线性空间，对于 $A = (A_{ij}) \in \mathbf{M}_n(\mathbb{C})$，我们有范数

$$\|A\|_2 = \sqrt{\sum_{i,j=1}^{n} |A_{ij}|^2},$$

从而 $\big(\mathbf{M}_n(\mathbb{C}), \|\cdot\|_2\big)$ 是一个赋范线性空间。作为距离空间，两个矩阵 A 和 B 之间的距离由 $d_2(A, B) = \|A - B\|_2$ 给出。这个赋范线性空间是**完备的**（我们证明过 $\big(\mathbb{R}^N, \|\cdot\|_2\big)$ 完备，即 *Cauchy* 列都收敛）。

我们已经定义过映射

$$\exp : \mathbf{M}_n(\mathbb{C}) \to \mathbf{M}_n(\mathbb{C}), \quad A \mapsto \sum_{k=0}^{\infty} \frac{A^k}{k!}.$$

我们现在证明这个映射是连续的。为此，我们首先回忆之前对于 $\exp : \mathbb{R} \to \mathbb{R}$ 的连续性的证明，我们用到了

$$e^x - e^{x_0} = e^{x_0}(e^{x-x_0} - 1) = e^{x_0} \sum_{k=1}^{\infty} \frac{(x - x_0)^k}{k!}.$$

然而，对于矩阵的情形，我们评论过 $e^{A+B} = e^A e^B$ 可能并不成立，所以上述第一步可能并不正确。然而不管怎么样，我们先证明 exp 在 $A = 0$ 处连续（这和实数的情形没有差别！），其中 $e^0 = 1$，其中 1 代表 $\mathbf{I}_{n\times n}$：

$$\|e^A - 1\|_2 \leqslant \sum_{k=1}^{\infty} \frac{\|A\|_2^k}{k!} = \|A\|_2 \sum_{k=1}^{\infty} \frac{\|A\|_2^{k-1}}{k!} \leqslant \|A\|_2 e^{\|A\|_2} < e\|A\|_2.$$

最后我们（不妨）假设了 $\|A\|_2 < 1$。所以当 $A \to 0$ 时或者 $\|A\|_2 \to 0$ 时，我们有 $\|e^A - 1\|_2 \to 0$，这就说明 exp 在 $A = 0$ 处是连续的。

在一般的情形，假设 $A \to A_0$，即 $B = A - A_0 \to 0$，我们计算

$$e^A - e^{A_0} = \sum_{k=1}^{\infty} \frac{A^k - A_0^k}{k!} = \sum_{k=1}^{\infty} \frac{(A_0 + B)^k - A_0^k}{k!}.$$

关键的观察如下（我们用到存在常数 c，对任意的 A 和 B，我们有 $\|A \cdot B\|_2 \leqslant c\|A\|_2\|B\|_2$（$c$ 是一个不依赖于 A 和 B）。）：

$$\begin{aligned}\|(A_0 + B)^2 - A_0^2\|_2 = \|A_0B + BA_0 + BB\|_2 &\leqslant \|A_0B\|_2 + \|BA_0\|_2 + \|BB\|_2 \\ &\leqslant c\|A_0\|_2\|B\|_2 + \|B\|_2\|A_0\|_2 + \|B\|_2\|B\|_2 \\ &= c\left[\left(\|A_0\|_2 + \|B\|_2\right)^2 - \|A_0\|_2^2\right].\end{aligned}$$

类似地，通过把作用范数，我们可以借此消除不交换性而得到

$$\|(A_0 + B)^k - A_0^k\|_2 \leqslant= c^{k-1}\left[\left(\|A_0\|_2 + \|B\|_2\right)^k - \|A_0\|_2^k\right].$$

从而，

$$\begin{aligned}\|e^A - e^{A_0}\|_2 &\leqslant \sum_{k=1}^{\infty} \frac{\|(A_0 + B)^k - A_0^k\|_2}{k!} \leqslant \sum_{k=1}^{\infty} \frac{c^{k-1}\left[\left(\|A_0\|_2 + \|B\|_2\right)^k - \|A_0\|_2^k\right]}{k!} \\ &= \frac{1}{c}\sum_{k=1}^{\infty} \frac{\left[\left(c\|A_0\|_2 + c\|B\|_2\right)^k - \left(c\|A_0\|_2\right)^k\right]}{k!} \\ &= \frac{1}{c}\left(e^{c\|A_0\|_2 + c\|B\|_2} - e^{c\|A_0\|_2}\right).\end{aligned}$$

现在根据在实数上指数函数的连续性，我们就知道当 $\|B\|_2 \to 0$ 时，右边是连续的，从而 exp 在 A_0 处连续。

我们再研究几个有代表性的例子（请同学们自己把细节写清楚，借此可以练习一下 $\varepsilon - \delta$ 语言）：

例子 (连续函数和不连续函数的例子)**.** 先从不连续的例子开始：

1) $X \subset \mathbb{R}$ 是一个子集，由这个子集所定义的**示性函数** $\mathbf{1}_X : \mathbb{R} \to \mathbb{R}$ 指的是：

$$\mathbf{1}_X(x) = \begin{cases} 1, & x \in X; \\ 0, & x \notin X. \end{cases}$$

假设 $X = [a, b]$ 是一个有限闭区间，$\mathbf{1}_X(x)$ 在 a 处是右连续的但是不是左连续（从而不连续），这个函数在 $x \neq a, b$ 处都是连续的。如果 $X = \mathbb{Q}$，那么 $\mathbf{1}_{\mathbb{Q}}(x)$ 在任何一个点处都是不连续的（函数 $\mathbf{1}_{\mathbb{Q}}(x)$ 通常被称作是 *Dirichlet* 函数）。

2) 我们考虑函数

$$f(x)=\begin{cases}\frac{1}{x}, & x\neq 0;\\ 0, & x=0.\end{cases}$$

这个函数在 0 处是不连续性的（我们可以说函数在 0 处的右极限是正无穷大，在 0 处的左极限是负无穷大)。

3) 我们考虑函数

$$f(x)=\begin{cases}\frac{1}{q}, & \text{如果 } x=\frac{p}{q} \text{ 是有理数，其中 } p\in\mathbb{Z},\ q\in\mathbb{Z}_{\geqslant 1} \text{ 且 } p \text{ 和 } q \text{ 互素;}\\ 0, & \text{如果 } x \text{ 是无理数.}\end{cases}$$

我们在第 *3* 次的作业中已经证明了 f 在有理数上不连续但是在无理数上连续!

练习. 是否存在 $\mathbb{R}$ 上的函数 f，它在任和一个点处都不连续，但是 $|f|$ 是连续的?。

我们现在给出两个最具有代表性的例子，建议大家将课堂笔记整理清楚并搞懂细节，通过这样的方式，也可以加深对连续性的认识。

例子. 下面的两个函数在 0 处都是连续的:

1) 我们定义函数

$$f(x)=\begin{cases}\dfrac{\sin(x)}{x}, & x\neq 0;\\ 1, & x=0.\end{cases}$$

为了说明 f 在 0 处连续，当 $x\neq 0$ 时，利用 $\sin(x)$ 的定义，我们有

$$\frac{\sin(x)}{x}=\sum_{k=0}^{\infty}\frac{(-1)^k x^{2k}}{(2k+1)!}.$$

从而，

$$\left|\frac{\sin(x)}{x}-1\right|\leqslant\sum_{k=1}^{\infty}\frac{|x|^{2k}}{(2k+1)!}=|x|^2\sum_{k=1}^{\infty}\frac{|x|^{2k-2}}{(2k+1)!}\leqslant|x|^2e^{|x|}.$$

从而，$\lim\limits_{x\to 0}\dfrac{\sin(x)}{x}=1$。

2) 我们考虑函数

$$f(x)=\begin{cases}e^{-\frac{1}{|x|}}, & x\neq 0;\\ 0, & x=0.\end{cases}$$

为了说明 $f(x)$ 在 0 处连续，对任意的 $\varepsilon>0$，我们只要选取 δ 使得 $e^{-\frac{1}{\delta}}<\varepsilon$ 即可，这个很容易做到。

另外，根据与数列版本的类比，我们不难想象出各种版本的对函数连续性进行判断的命题，比如如下两边进行控制的版本:

命题 49. *g_1 和 g_2 是区间 I 上定义的 $\mathbb{R}$-值函数，它们在 x_0 处有极限并且 $\lim\limits_{x\to x_0} g_1(x) = \lim\limits_{x\to x_0} g_2(x)$。如果 $f : I \to \mathbb{R}$ 使得对任意的 $x \in I$，都有 $g_1(x) \leqslant f(x) \leqslant g_2(x)$，那么，$f$ 在 x_0 处有极限并且*

$$\lim_{x\to x_0} f(x) = \lim_{x\to x_0} g_1(x) = \lim_{x\to x_0} g_2(x).$$

利用函数收敛的数列版本的定义（根据不同的情形选择需要的形式），这个性质的证明显而易见的。我们只要看到这一点，这个命题也容易记住了。另外，我们指出，这样的命题可能在某些场合（比如说解题目）有用，但是就本身而言，（我觉得）可能不是很有价值。

连续函数的基本性质

我们现在来了解连续函数的一些基本性质并利用它们做两件基本的事情：第一，研究 e^x 的性质并定义所有的初等函数；第二，利用连续函数来研究空间的几何性质。

定理 50. *I 是区间（可开可闭），$f : I \to \mathbb{R}$ 是 I 上的单调函数，那么*

1) 对任意 $x_0 \in I$，f 在 x_0 处的左极限 $\lim\limits_{x\to x_0^-} f(x)$ 和右极限 $\lim\limits_{x\to x_0^+} f(x)$ 都存在。

2) f 在 I 上的不连续点的集合是可数的。(这表明单调函数 f 在“大部分”点处都是连续的)

证明: 我们不妨假设 $I = \mathbb{R}$，f 是单调上升的函数。我们逐个证明这两个论断：

1) 只证明左极限的情况即可，因为右极限可以类似地证明。我们要说明存在 y_0，使得对任意的数列 $\{x_n\}_{n\geqslant 1}$，$x_n < x_0$，$x_n \to x_0$，$\lim\limits_{n\to\infty} f(x_n) = y_0$。实际上，我们令 $y_0 = \sup\limits_{x\in(-\infty,x_0)} f(x)$。根据 y_0 的定义，我们有 $f(x_n) \leqslant y_0$。另外，根据上确界的性质，对任意的 $\varepsilon > 0$，存在 $x' < x_0$，使得 $0 \leqslant y_0 - f(x') < \varepsilon$。由于 $x_n \to x_0$，所以存在 N，使得当 $n \geqslant N$ 时，有 $x_n > x'$，从而 $f(x_n) \geqslant f(x') \geqslant y_0 - \varepsilon$。综合上述，对任意的 $\varepsilon > 0$，存在 N，使得当 $n \geqslant N$ 时，$y_0 \geqslant f(x_n) \geqslant y_0 - \varepsilon$，所以 $\lim\limits_{n\to\infty} f(x_n) = y_0$。作为证明的推论，我们有

   推论 51. *假设 $f : I \to \mathbb{R}$ 是区间 I 上的单调递增的函数，对于 $x_0 \in I$，f 在 x_0 处的左右极限由下面的公式给出：*

$$\lim_{x\to x_0^-} f(x) = \sup_{x\in(-\infty,x_0)} f(x), \quad \lim_{x\to x_0^+} f(x) = \inf_{x\in(x_0,+\infty)} f(x).$$

   特别地，$\lim\limits_{x\to x_0^-} f(x) \leqslant \lim\limits_{x\to x_0^+} f(x)$。

2) 这是一个值得大家记住的经典证明。考虑 f 的不连续点的集合

$$Y = \big\{y \in I \big| f \text{ 在 } y \text{ 处不连续}\big\}.$$

   对于任意 $y \in Y$，按照定义，$\lim\limits_{x\to x_0^-} f(x) \neq \lim\limits_{x\to x_0^+} f(x)$。再利用上面推论中的结论，我们知道，对任意 $y \in Y$，我们有

$$\lim_{x\to y^-} f(x) < \lim_{x\to y^+} f(x).$$

据此，对于任意 $y \in Y$，都唯一地确定了一个非空的开区间 $I_y = (\lim_{x\to y^-} f(x), \lim_{x\to y^+} f(x))$，即我们构造了映射

$$Y \to \{\mathbb{R} \text{ 上的全体非空开区间}\},\ \ y \mapsto I_y.$$

利用单调性，我们首先说明对任意的 $y_1, y_2 \in Y$，$y_1 \neq y_2$，$I_{y_1} \cap I_{y_2} = \emptyset$：

不妨假设 $y_1 < y_2$，根据上面的推论的结论，我们可以选取单调下降的数列 $\{x_k\}_{k\geqslant 1}$，使得 $x_k \downarrow y_1$（从右边逼近）并且 $\lim\limits_{k\to\infty} f(x_k)$ 为 I_{y_1} 的右端点；类似地，我们选取单调上升的数列 $\{z_k\}_{k\geqslant 1}$，$z_k \uparrow y_2$（从左边逼近），使得 $\lim\limits_{k\to\infty} f(z_k)$ 为 I_{y_2} 的左端点。由于 $y_1 < y_2$，我们可以假设对任意 k 都有 $x_k < z_k$，所以 $\lim\limits_{k\to\infty} f(x_k) \leqslant \lim\limits_{k\to\infty} f(z_k)$，也就是说 I_{y_1} 的右端点要么在 I_{y_2} 的左端点的左边要么重合，由于这两个区间都是开区间，所以它们的交集是空集。

这样，我们就得到了一组 $\mathbb{R}$ 上两两不相交的非空开区间 $\{I_y | y \in Y\}$，我们在每个开区间 I_y 里选定一个有理数 q_y，由于这些开区间互不相交，这些有理数 q_y 也决定了这些开区间。从而，我们得到了单射

$$Y \to \mathbb{Q},\ \ y \mapsto q_y.$$

由于有理数是可数的，所以 Y 可数。

□

我们下面证明著名的介值定理，这个定理有很多其它的证明，都比较有启发性，建议大家查阅资料，比如陈天权，Zorich 或者科大的教材。

定理 52 (介值定理)**.** $a < b$，$f : [a, b] \to \mathbb{R}$ *是连续函数。*

1) *如果* $f(a) < 0$，$f(b) > 0$，*那么一定存在* $c \in [a, b]$，*使得* $f(c) = 0$。

2) *如果* $f(a) \neq f(b)$，*那么对任意介于* $f(a)$ *和* $f(b)$ *之间的数* y_0，*总存在* $c \in (a, b)$，*使得* $f(c) = y_0$。

证明: 2) 是 1) 的推论，不妨假设 $f(a) < f(b), y \in \big(f(a), f(b)\big)$，我们考察连续函数 $F(x) = f(x) - y_0$，此时，$F(a) < 0$，$F(b) > 0$，从而有 $c \in (a, b)$，使得 $F(c) = 0$，这等价于 $f(c) = y_0$。

我们现在利用反证法来证明 1)。如若不然，我们利用采用庄子的二分法：令 $I_0 = [a, b]$。考虑 $\dfrac{a+b}{2}$，根据反证假设，$f(\dfrac{a+b}{2}) \neq 0$。如果 $f(\dfrac{a+b}{2}) > 0$，我们就令 $I_1 = [a, \dfrac{a+b}{2}]$；如果 $f(\dfrac{a+b}{2}) > 0$，我们就令 $I_1 = [\dfrac{a+b}{2}, b]$。假设我们已经有了区间 $I_k = [a_k, b_k]$，区间 I_{k+1} 的构造如下：考虑 $\dfrac{a_k+b_k}{2}$，根据反证假设，$f(\dfrac{a_k+b_k}{2}) \neq 0$。如果 $f(\dfrac{a_k+b_k}{2}) > 0$，我们就令 $I_{k+1} = [a_k, \dfrac{a_k+b_k}{2}]$；如果 $f(\dfrac{a_k+b_k}{2}) > 0$，我们就令 $I_{k+1} = [\dfrac{a_k+b_k}{2}, b_k]$。

通过上述构造，我们得到闭区间套 $I_0 \supset I_1 \supset \cdots$，其中 f 在 I_k 的左端点处取值是负的，在右端点处取值是正的，并且当 $k \to \infty$ 时，$|I_k| \to 0$，所以 $\bigcap\limits_{k\geqslant 0} I_k = \{c\}$ 是单点集。很明显，$c \in [a, b]$。我们现在说明 $f(c) = 0$（从而得到矛盾命题得证）：如若不然，不妨设 $f(c) > 0$，那么按照函数

连续性的 $\epsilon-\delta$ 语言的定义，对于 $\varepsilon=\dfrac{f(c)}{2}$，存在 $\delta>0$，使得对任意 $x\in(c-\delta,c+\delta)\cap I$，$|f(x)-f(c)|<\dfrac{1}{2}f(c)$。特别地，对于任意的 $x\in(c-\delta,c+\delta)\cap I$，$f(x)>0$。然而，根据 $c\in I_k$，当 k 很大的时候，必然有 $I_k\subset(c-\delta,c+\delta)\cap I$，但是 f 在 I_k 的左端点处的取值是负的，矛盾。 □

注记. 直观连续函数的图像是不会断掉的，这是我们对连续性最朴素的理解。介值定理就是这个直观的数学表达。

另外，我们要指出一个值得注意和讨论的例子：假设函数 $f:I_1\cup I_2\to\mathbb{R}$ 定义在两个不相交的开区间的并集上，按照定义，如果 f 在每个点上都连续，f 就是在 $I_1\cup I_2$ 上连续的。尽管函数图像在 $I_1\cup I_2$ 是“断开的”，这个函数仍然是连续的（因为连续性本质上是个局部性质，在 I_1 和 I_2 上分别连续即可）。此时，介值定理可能并不成立，因为我们要求 f 的定义域不是断掉的（“连同的”，这是一个拓扑学的概念）。

另外，对于对有理数上定义的连续函数介值定理一般并不成立的，比如说，

$$f:\mathbb{Q}\cap[0,3]\to\mathbb{R},\ x\mapsto f(x)=x-e.$$

这和有理数不完备有关系。

连续函数的介值定理有很多经典的应用，特别是在证明方程解的存在性方面。我们给出几个有代表性的例子，在第 4 次作业中有相应的练习。

例子 (经典应用). *1)* $f:[0,1]\to[0,1]$ 是连续映射，那么，f 有不动点，即存在 $x\in[0,1]$，使得 $f(x)=x$。（请比较压缩映像定理）

事实上，考虑函数 $F(x)=f(x)-x$，按照 f 的要求，$F(0)\geqslant 0$，$F(1)\leqslant 1$，所以根据介值定理，存在 $x\in[0,1]$，使得 $F(x)=0\Leftrightarrow f(x)=x$。

对于高维的情形，令 $I^n=[0,1]^n=\big\{(x_1,\cdots,x_n)\big|0\leqslant x_1,\cdots,x_n\leqslant 1\big\}$。那么，连续映射 $f:I_n\to I_n$ 一定有不动点。这是一个深刻的结论，叫做 *Brouwer* 不动点定理，我们在后面的课程（作业）中会给出证明。

2) 实系数奇数次的多项式一定有实根。

通过乘一个常数，我们不妨假设 $P(x)=x^{2m+1}+\sum\limits_{0\leqslant k\leqslant 2m}a_kx^k$，很明显，当 $x\to+\infty$ 时，$P(x)>0$；当 $x\to-\infty$ 时，$P(x)<0$。利用介值定理，我们就可以找到一个实根。

这个性质在很多地方有应用，比如说可以用来证明三维空间的旋转一定有旋转轴（从而可以视作是二维的旋转），还可以与 *Galois* 理论结合证明代数基本定理。

3) 映射 $f:\mathbb{R}_{\geqslant 0}\to\mathbb{R}_{\geqslant 0},\ x\mapsto x^2$ 是双射。

首先，根据序的性质（大小关系），我们知道 f 是单调递增的函数，从而为单射；另外，任意给定 $y_0\in\mathbb{R}_{>0}$，由于 $f(0)=0<y_0$，$f(y_0+1)>2y_0+1>y_0$，根据介值定理，就存在（唯一的）$x_0\in[0,y_0+1]$，使得 $f(x_0)=y_0$，这表明 f 为双射。

根据这个双射的结论，我们可以定义 $f^{-1}:\mathbb{R}_{\geqslant 0}\to\mathbb{R}_{\geqslant 0}$，这就是我们所熟悉的开根号映射，习惯上，我们把它记做 $f^{-1}(x)=\sqrt{x}$。这是一个有启发性的例子，比如下面的 log 的定义就是机械地重复一下这个例子的想法。请同学们也与第 2 次作业习题 E 做比较。

再者，我们还关心函数 $\sqrt{x}$ 的连续性，我们有更一般性的定理来处理这一点。

定理 53. $f:[a,b]\to\mathbb{R}$ 是在闭区间上定义的连续函数，那么 f 有界的并且能取到最大最小值，即存在 $x_1,x_2\in[a,b]$，使得 $f(x_1)=\inf\limits_{x\in[a,b]}f(x)$，$f(x_2)=\sup\limits_{x\in[a,b]}f(x)$。

练习. 定理叙述中的两个条件“闭区间”和“连续”缺一不可，试举出反例。

证明: 用 I 表示闭区间 $[a,b]$。我们用反证法证明 f 是有界的：如若不然，那么存在数列 $\{x_n\}_{n\geqslant 1}\subset I$，使得 $f(x_n)\to+\infty$（不妨设是正无穷）。通过选取 $\{x_n\}_{n\geqslant 1}$ 的子列，我们可以进一步假设 $x_n\to x$。证明的关键点在于 $x\in I$，这是由 I 是闭区间保证的：因为 $x_n\in I\Leftrightarrow a\leqslant x_n\leqslant b$，所以通过取极限（极限和 $\leqslant$ 以及 $\geqslant$ 交换），$a\leqslant x\leqslant b\Leftrightarrow x\in I$。（这是对词语“闭”的基本理解，极限点不能跑到外面去，被封闭在里面了），根据 f 在 x 处的连续性，所以 $f(x_n)\to f(x)\neq\infty$，这不可能。

下面证明 $A=\sup\limits_{x\in I}f(x)$ 一定能被某个点 x_2 所实现，即存在 $x_2\in I$ 使得，$f(x_2)=A$。我们仍然使用反证法：如若不然，对任意 $x\in I$，$f(x)\neq A$，所以，对任意的 $x\in I$，$f(x)<A$。我们考虑正值的函数

$$g(x)=\frac{1}{A-f(x)}.$$

这个函数自然是良好定义的（因为分母不是零）并且是连续的。然而，由于 A 是 $f(x)$ 取值的上确界，所以存在 $x\in I$，使得 $f(x)$ 可以无限地接近 A，从而 $g(x)$ 是无界的：对任意的 $M>0$，存在 $x\in I$，使得 $A-f(x)<\dfrac{1}{M}$，从而 $g(x)>M$。这和已经证明的有界性矛盾。 □

注记. 上面用来证明最大值能被某个 x_2 实现的方法（在复变函数课程中学习 *Liouville* 定理的应用时）也可以用来证明代数基本定理。

定理 54. $f:[a,b]\to\mathbb{R}$ 是严格递增（或者递减）的连续函数，那么 f 是从 $[a,b]$ 到 $[f(a),f(b)]$ 的双射并且其逆映射 $f^{-1}:[f(a),f(b)]\to[a,b]$ 是连续的。

注记. 这是 1-元函数的特殊性质，究其根本，是因为 1 维空间上面 $\mathbb{R}$ 有序关系 $\leqslant$ 而在 $\mathbb{R}^n$ 上是没有的。

证明: 首先，我们知道 f 是单射（因为严格递增）并且对任意的 $x\in[a,b]$，$f(x)\in[f(a),f(b)]$。另外，根据介值定理，f 是满射。所以，我们可以定义其逆 $f^{-1}:[f(a),f(b)]\to[a,b]$。我们注意到，$f^{-1}$ 也是严格递增的（否则假设存在 $[f(a),f(b)]$ 中两个点 $y_1<y_2$，使得 $x_1=f^{-1}(y_1)\geqslant x_2=f^{-1}(y_2)$，根据 f 是递增的，$f(x_1)=y_1\geqslant f(x_2)=y_2$，矛盾）。

为了证明 f^{-1} 是连续的，我们用反证法：如若不然，假设 $y_0\in[f(a),f(b)]$ 是一个不连续点（假设 $y_0=f(x_0)$，$x_0\in[a,b]$）。根据定理50 关于单调函数不连续点是可数的证明，我们对于 y_0 可以分配一个非空的开区间：

$$I_{y_0}=(x_1,x_2),\quad x_1=\sup_{z<y_0}f(z)=\lim_{z\to y_0^-}f^{-1}(z),\ x_2=\inf_{z>y_0}f^{-1}(z)=\lim_{z\to y_0^+}f^{-1}(z).$$

特别地，$x_1 < x_2$。根据单调性，我们知道 $x_1 \leqslant x_0 \leqslant x_2$。另外，对于任意的 $y > y_0$，$f^{-1}(y) \geqslant \inf\limits_{z>y_0} f^{-1}(z) = x_2$；对于任意的 $y < y_0$，$f^{-1}(y) \leqslant \inf\limits_{z<y_0} f^{-1}(z) = x_1$。所以，$(x_1, x_2)$ 之间的数（无限多个）除了可能 x_0 之外，都不落在 f^{-1} 的值域里面，然而它们自然在 f 的定义域里面（按照定义就在 f^{-1} 的值域里），矛盾。 □

第 4 次作业中我们将会解答一个很有意思的问题：

练习. *假设连续函数 $f : [a, b] \to \mathbb{R}$ 是单射。如果 $f(a) < f(b)$，证明，f 是严格递增的的。*

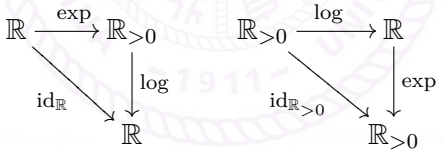
连续函数性质的应用：初等函数的构造

定理 55. *指数函数 $\exp : \mathbb{R} \to \mathbb{R}_{>0}$，$x \mapsto e^x$ 是双射。我们用 $\log : \mathbb{R}_{>0} \to \mathbb{R}$ 表示 $\exp$ 的反函数并称其为***以 e 为底的对数函数***，它满足：*

1) 对任意 $x, y > 0$，我们有 $\log(xy) = \log(x) + \log(y)$。

2) $\log(x)$ 是 $\mathbb{R}_{>0}$ 上的连续的单调递增函数。

证明： 只需要说明 $\exp : \mathbb{R} \to \mathbb{R}_{>0}$ 是满射（单调性意味着是单射）即可，其余的都是前面几个定理的直接推论。对任意的 $y_0 \in \mathbb{R}_{>0}$，我们可以选取 N，使得 $e^{-N} < y_0 < e^N$（比如说，令 $N = y_0 + \dfrac{1}{y_0} + 1$ 即可，请自行验证），所以根据介值定理，存在 $x_0 \in [-N, N]$，使得 $\exp(x_0) = y_0$，这表明 $\exp$ 是满射。至此，我们可以构造 $\exp$ 的逆映射 $\log : \mathbb{R}_{>0} \to \mathbb{R}$。用交换图标的语言（交换图的说法值得学习，尤其对于后来学习代数和拓扑），我们将 $\log$ 和 $\exp$ 之间的关系写成：



为了说明 $\log$ 是连续的，我们只需要在 $[e^{-n}, e^n]$ 这个区间上证明 $\log$ 连续即可（连续是局部性质），其中 $n \in \mathbb{Z}$ 是任取的。根据定理54，$\exp : [-n, n] \to [e^{-n}, e^n]$ 是严格递增的连续映射，所以它的逆 $\log$ 连续。

另外，根据双射性质，$\log(xy) = \log(x) + \log(y)$ 等价于 $\exp\big(\log(xy)\big) = \exp\big(\log(x) + \log(y)\big)$，即 $xy = \exp\big(\log(x)\big)\exp\big(\log(y)\big) = xy$，所以对任意 $x, y > 0$，$\log(xy) = \log(x) + \log(y)$ 成立。 □

利用 $\exp$ 和 $\log$，我们终于可以定义一般的对数函数和幂函数了：

定理 56. *对于 $\alpha \in \mathbb{R}$ 和 $x \in \mathbb{R}_{>0}$，我们定义***幂函数***$x^\alpha = e^{\alpha \log(x)}$，那么 x^α 满足*

1) 对任意 $x, y > 0$ 和 α, β，我们有 $(xy)^\alpha = x^\alpha y^\alpha$，$(x^\alpha)^\beta = x^{\alpha\beta}$。

2) x^α 是 $\mathbb{R}_{>0}$ 上的连续函数。

3) 当 $\alpha \in \mathbb{Z}_{\geqslant 0}$ 时，这个定义和经典的定义是一致的，即 $e^{n\log(x)} = \underbrace{x \cdot x \cdot \cdots \cdot x}_{n\text{个}}$。

对于 $a, x \in \mathbb{R}_{>0}$，我们定义**对数函数**$\log_a(x) = \dfrac{\log x}{\log a}$ 和**指数**$a^x = e^{\log(a)x}$，那么它们都是连续函数并且满足下面的性质：

1) $a^{\log_a x} = x$。

2) $a^{x+y} = a^x a^y$，$\log_a(x \cdot y) = \log_a(x) + \log_a(y)$（要求 $x > 0$，$y > 0$）。

证明： 定理所有的性质都是 exp 和 log 的性质在复合映射下的直接推论。为了验证当 α 为非负整数时，这个定义和经典的一致：我们有 $x^n = e^{n\log(x)} = e^{\log(x)+\cdots+\log(x)} = e^{\log(x)} \cdots e^{\log(x)} = x \cdots x$，所以恰好是 n 个 x 乘起来。对于其余的性质证明不放心的同学，可以在第四次作业中安心地验证这些细节。 □

我们讲一些相关的题外话。中学的数学学习中我们实际上从未定义 $x^{\frac{1}{n}}$（只是想当然的认为这个函数存在），今天我（你）们第一次正确地定义了开方这个运算并且验证了我们之前认为是正确的各种性质（所以我们之后会不假思索地继续应用这些中学熟知的性质）。中学我们对 e 的了解很少，所以和 e 相关的对象反而变得容易理解，因为我们只需要验证关于 e 的很少的几个已经知道的性质即可；然而，对于 π 以及相关的 $\sin x$ 等三角函数，我们了解很多它们的性质，这反而给我们新发展的函数理论带来了很大的挑战：我们在定义这些对象的同时要能够证明它们满足所知的所有性质。实际上，我们必须建立了整个微积分的理论之后才能够做到这一点。

日本京都大学的数学家望月新一在他的宇宙际 Techimüller 理论的论文（他在这一系列论文中声称他证明了 abc 猜想，目前还没有得到承认）里说：Unlike many mathematical papers, which are devoted to verifying properties of mathematical objects that are either well-known or easily constructed from well-known mathematical objects, in the present series of papers, most of our efforts will be devoted to constructing new mathematical objects. 我们这里构造的幂函数就是他所说的“well-known or easily constructed from well-known mathematical objects”。

10 开集，闭集，紧集，闭包，聚点，连续函数的拓扑刻画

二零一九年十月十七日，星期四，阴有小雨

对于上次课定义的对数函数和幂函数等，我们对它们的连续性做一下补充，这可以加深我们作为从映射的观点看函数的理解。以对数函数为例子：

给定 $a>0$，$a\neq 1$，$x>0$，我们定义了 $\log_a x=\dfrac{\log x}{\log a}$。当 $a>0$ 固定的时候，作为 x 的函数 $x\mapsto \log_a x$，我们已经说明它是连续函数。现在，我们把 a 看成变量，从而我们定义函数

$$\mathrm{LOG}:\mathbb{R}_{>0,\neq 1}\times\mathbb{R}_{>0}\to\mathbb{R},\quad (a,x)\mapsto \log_a x,$$

其中 $\mathbb{R}_{>0,\neq 1}=(0,1)\cup(1,\infty)$。

在 $\mathbb{R}_{>0,\neq 1}\times\mathbb{R}_{>0}\to\mathbb{R}$ 有自然的距离函数（作为 $\mathbb{R}^2$ 的子集），我们要说明 LOG 是连续函数（作为双变量的函数）！

首先考虑映射如下两个映射

$$f_1:\mathbb{R}_{>0,\neq 1}\times\mathbb{R}_{>0}\longrightarrow\mathbb{R}_{\neq 0},\quad (a,x)\mapsto\log(a);$$
$$f_2:\mathbb{R}_{>0,\neq 1}\times\mathbb{R}_{>0}\longrightarrow\mathbb{R},\quad (a,x)\mapsto\log(x).$$

根据实数上定义的 log 的连续性，我们很容易验证 f_1 和 f_2 连续（请自行验证），根据我们上次课关于在乘积空间上取值的函数的连续性的判断，映射

$$F:\mathbb{R}_{>0,\neq 1}\times\mathbb{R}_{>0}\longrightarrow\mathbb{R}_{\neq 0}\times\mathbb{R},\quad (a,x)\mapsto\big(\log(a),\log(x)\big),$$

是连续的。再考虑映射

$$D:\mathbb{R}_{>0,\neq 1}\times\mathbb{R}_{>0}\longrightarrow\mathbb{R},\quad (y,x)\mapsto\frac{y}{x}.$$

我们很容易验证 D 是连续的（请自行验证）。从而，LOG 是 D 和 F 的复合：

$$\begin{array}{ccc}\mathbb{R}_{>0,\neq 1}\times\mathbb{R}_{>0} & \xrightarrow{F} & \mathbb{R}_{\neq 0}\times\mathbb{R}\\ & \searrow^{\mathrm{LOG}} & \downarrow D\\ & & \mathbb{R}\end{array},$$

自然是连续的。以上，我们把一个复杂的函数拆成若干个相对简单的连续映射的复合来证明连续性。

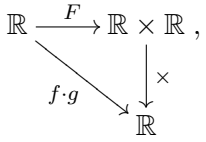
我们再举一个例子来说明这个想法的应用，假设 $f(x)$ 和 $g(x)$ 是 $\mathbb{R}$ 上的连续函数，我们想证明它们的乘积 $f\cdot g$ 也是连续的。为此，我们考虑

$$F:\mathbb{R}\longrightarrow\mathbb{R}\times\mathbb{R},\quad x\mapsto\big(f(x),g(x)\big),$$

它两个分量都是连续的，所以 F 连续。另外，再考虑乘法映射：

$$\times:\mathbb{R}\times\mathbb{R}\longrightarrow\mathbb{R},\quad (x,y)\mapsto x\cdot y.$$

我们容易验证 M 是连续的（请自行验证）。从而，$f\cdot g$ 可以看作是连续映射 $\times$ 和 F 的复合：



从而是连续的。

我们再做一点补充，关于连续函数在一点上的值的计算：假设 $\mathbb{R}$ 上定义函数 $f(x)$ 是连续的，如果我们知道 f 在有理数上的取值，那么对任意的无理数 x，我们任取一列有理数 r_i，使得 $r_i\to x$，连续性保证了 $f(x)=f(\lim\limits_{i\to\infty} r_i)=\lim\limits_{i\to\infty} f(r_i)$，所以 $f(x)$ 可以确定。注意到 $\mathbb{Q}$ 是 $\mathbb{R}$ 中“很小”的子集，由连续性，它已经可以确定 f 了！这是分析学最重要的精神之一：如何从“局部”到整体！

我们现在给出这个补充的数学表述：首先回忆一下我们在第一次作业的题目 A 中引入的定义：

定义 57. *给定距离空间 (X,d)，$X'\subset X$ 是子集。如果对任意的 $x\in X$ 和任意的 $\varepsilon>0$，都存在 $x'\in X$，使得 $d(x',x)<\varepsilon$，我们就称 X' 在 X 中是*稠密的。

根据定义，对任意给定的 $x\in X$，令 $\varepsilon=\dfrac{1}{n}$，我们选取 $x'_n\in X'$ 使得 $d(x'_n,x)<\dfrac{1}{n}$，从而我们得到点列 $\{x'_n\}_{n\geqslant 1}\subset X'$，使得 $x'_n\to x$，即 x 中的每个点都可以用 X' 中的点来逼近，这个表述是很有用的。

定理 58. *(X,d) 和 (Y,d_Y) 是距离空间，$X'\subset X$ 是一个稠密的子集，我们仍然用 d 表示 X' 上的距离。对任意的连续映射 $F_1:X'\to Y$ 和 $F_2:X'\to Y$，$F_1\big|_{X'}=F_2\big|_{X'}$，那么，$F_1=F_2$，换而言之，连续映射被它在一个稠密子集上的限制所决定。*

证明： 假设 $x\in X$，我们任意选取 $\{x'_n\}_{n\geqslant 1}\subset X'$，使得 $x'_n\to x$。因为 F_1 和 F_2 在 X 上连续，所以

$$F_1(x)=\lim_{n\to\infty} F_1(x'_n)=\lim_{n\to\infty} F_2(x'_n)=F_2(x).$$

命题得证。 □

紧性

我们要讨论函数**一致连续**的概念，为此，我们要采取一个迂回的方式而不是直接利用目前已经熟悉的数列收敛的工具。这个繁琐方法的优点可以引进所谓的拓扑的概念并给出连续映射的第三种等价的刻画（之前我们可以用点列或者 $\varepsilon-\delta$ 语言来谈论连续性）。在开始介绍拓扑的概念之前，我们引用 John von Neumann 的一段话：

Young man, in mathematics you don't understand things. You just get used to them.

首先谈论 $\mathbb{R}^n$ 上的开闭集，这些概念是我们熟知的开区间和闭区间的类比：

(1) 如果 $U \subset \mathbb{R}$ 是开区间的并，即 $U = \bigcup_{\alpha \in \mathcal{A}} I_\alpha$，其中 I_α 是开区间（指标集 $\mathcal{A}$ 是任意的），就称 X 是**开集**。

按照定义，U 是开集当且仅当对任意 $x \in U$，存在包含 x 的开区间，使得该开区间完全包含在 U 中。另外一个等价的描述就是，对任意的 $x \in U$，存在 $\delta_x > 0$，当 $|y - x| < \delta$ 时，$y \in U$，即和 x 邻近的点都落在 U 中。

按照定义，$\mathbb{R}$ 是开集。我们规定 $\emptyset$ 也是开集。对于 $x \in \mathbb{R}$，我们习惯上将包含 x 的一个开集称作是 x 的一个开邻域。

2) 如果 $F \subset \mathbb{R}$ 的补集是开集，我们就称 F 是**闭集**。

按照定义，$\mathbb{R}$，$\emptyset$，闭区间 $[a, b]$，$[a, \infty)$ 以及 $(-\infty, a]$ 都是闭集。

我们还有一种方式来判断一个集合是否是闭集（闭的意思可以解释为对这个集合中的点列取极限是封闭的）：

- F 是闭集当且仅当对任意数列 $\{x_n\}_{n \geqslant 1} \in F$，如果 $\lim\limits_{n\to} x_n = x$，那么 $x \in F$。

证明分两个方面：

如果 F 是闭集，任选 $\{x_n\}_{n \geqslant 1} \in F$，$\lim\limits_{n\to} x_n = x$，我们用反正法假设 $x \notin F$ 来推出矛盾：按照定义，$\mathbb{R} - F$ 是开集，所以存在 $\delta > 0$，使得 $[x - \delta, x + \delta] \cap F = \emptyset$，从而对任意的 $x_n \in F$，$|x_n - x| \geqslant \delta$，这与 $\lim\limits_{n\to} x_n = x$ 矛盾。

反过来，我们假设对任意数列 $\{x_n\}_{n \geqslant 1} \in F$，$\lim\limits_{n\to} x_n = x$ 能推出来 $x \in F$。我们需要证明 $\mathbb{R} - F$ 为开集：如若不然，存在 $x \notin F$，使得对任意的包含 x 的开区间 I，$I \not\subset \mathbb{R} - F$，或者说 $I \cap F \neq \emptyset$，特别地，我们依次选取 $I = (x - \frac{1}{n}, x + \frac{1}{n})$，由于 $I \cap F \neq \emptyset$，我们选取 $x_n \in (x - \frac{1}{n}, x + \frac{1}{n}) \cap F$，从而 $x_n \to x$，但是 $x \notin F$，矛盾。

利用上面的结论也很容易判断闭区间等是闭集。

我们有下面证明简单但是意义重大的命题：

命题 59 ($\mathbb{R}$ 上标准拓扑的概念)**.** *我们用 $\mathcal{T}$ 表示 $\mathbb{R}$ 上的开集的全体，它满足如下的性质：*

1) $\emptyset \in \mathcal{T}$，$X \in \mathcal{T}$。

2) 对任意开集的集合 $\{U_\alpha\}_{\alpha \in \mathcal{A}}$，其中 $\mathcal{A}$ 为指标集合，我们有 $\bigcup\limits_{\alpha \in \mathcal{A}} U_\alpha \in \mathcal{T}$。

3) 对任意有限个开集 $U_1, U_2, \cdots, U_m \in \mathcal{T}$，我们有 $\bigcap\limits_{1 \leqslant i \leqslant m} U_i \in \mathcal{T}$。

对于闭集，我们有如下对偶的性质：

1) $\emptyset$ 和 $\mathbb{R}$ 都是闭集。

2) 任意多闭集的交集还是闭集。

3) 有限个闭集的并集还是闭集。

证明: 给定子集 $S\subset\mathbb{R}$，我们用 S^{c} 表示 S 在 $\mathbb{R}$ 中的补集。对任意的指标集合 $\mathcal{A}$，我们知道

$$\left(\bigcup_{\alpha\in\mathcal{A}}S_\alpha\right)^{\mathrm{c}}=\bigcap_{\alpha\in\mathcal{A}}(S_\alpha)^{\mathrm{c}},\quad \left(\bigcap_{\alpha\in\mathcal{A}}S_\alpha\right)^{\mathrm{c}}=\bigcup_{\alpha\in\mathcal{A}}(S_\alpha)^{\mathrm{c}}.$$

利用这个结论，我们可以用开集的结论立即推出闭集的结论。

关于开集的结论中的前两条是平凡的，现在来证明 3)：任意选取 $x\in\bigcap\limits_{1\leqslant i\leqslant m}U_i$，所以对每个 $i=1,\cdots,m$，$x\in U_i$，从而存在开区间 I_i，使得 $I_i\subset U_i$。我们注意到 $I=\bigcap\limits_{1\leqslant i\leqslant m}I_i$ 也是开区间并且是 $\bigcap\limits_{1\leqslant i\leqslant m}U_i$ 的子集，所以 $\bigcap\limits_{1\leqslant i\leqslant m}U_i$ 是开集。 □

我们可以更精确地用开区间刻画 $\mathbb{R}$ 上的开集。请记住，这是 $\mathbb{R}$ 的特殊性质，不能推广到更一般维数:

命题 60. 假设 $U\subset\mathbb{R}$ 是开集，那么 U 是可数个两两相互不交的开区间的并，即 $U=\coprod\limits_{i=1}^{\infty}I_i$，其中 I_i 是开区间（包括空集)。

证明: 根据开集的定义，对任意 $x\in U$，$\mathcal{I}_x=\left\{I\middle|I\text{是开区间},x\in I,I\subset U\right\}$ 非空。据此，我们定义 $\mathbf{I}_x=\bigcup\limits_{I\in\mathcal{I}_x}I$，这是落在 U 中并且包含 x 的所有的开区间的并。特别地，这是一个开集并且 $x\in\mathbf{I}_x\subset U$。我们要证明 $\mathbf{I}_x\in\mathcal{I}_x$，即 $\mathbf{I}_x$ 也是开区间。

假设 $x_1,x_2\in\mathbf{I}_x$ 并且 $x_1<x_2$。按定义，存在开区间 I_1 和 I_2，使得 $x_1\in I_1$，$x_2\in I_2$，从而 $x_1,x_2\in I_1\cup I_2$。由于 $x\in I_1\cap I_2$，所以 $I_1\cup I_2$ 是开区间（它们相交)，从而 $(x_1,x_2)\in\mathcal{I}_x$ 中，所以 $(x_1,x_2)\subset\mathbf{I}_x$。令 $b=\sup\mathbf{I}_x$，$a=\inf\mathbf{I_x}$（可以是正负无穷大)。按照定义，$\mathbf{I}_x\subset[a,b]$。我们现在可以证明 $(a,b)\subset\mathbf{I}_x$，这是因为对于任意的 $c\in(a,b)$，利用上下确界的定义，存在 $x_1,x_2\in\mathbf{I}_x$ 并且 $a<x_1<c<x_2<b$，上面的推理表明 $(x_1,x_2)\subset\mathbf{I}_x$，所以 $c\in\mathbf{I}_x$。

综合上述，我们知道 $(a,b)\subset\mathbf{I}_x\subset[a,b]$，根据 $\mathbf{I}_x$ 是开集，只有 $\mathbf{I}_x$ (a,b)。至此，我们说明了存在包含 x 的最大的开区间 $\mathbf{I}_x\subset U$。

根据区间的最大性，对于 $x,y\in U$，要么 $\mathbf{I}_x=\mathbf{I}_y$，要么 $\mathbf{I}_x\cap\mathbf{I}_y=\emptyset$。所以，$U$ 可以写成不同的 $\mathbf{I}_x$ 的无交集并。我们在研究单调函数的不连续点的时候已经证明了 $\mathbb{R}$ 上不交的开区间只有可数个，参见定理50。 □

我们现在给出连续函数的第三个刻画:

定理 61 (连续函数的拓扑表示)**.** 给定函数 $f:\mathbb{R}\to\mathbb{R}$，那么 f 是连续函数当且仅当对任意开集的逆像是开集，即对任意的 $U\in\mathcal{T}$，$f^{-1}(U)\in\mathcal{T}$。

证明：先假设 f 是连续函数，我们证明 $f^{-1}(U)$ 是开集，其中 U 是开集：任选 $x_0 \in f^{-1}(U)$，令 $y_0 = f(x_0) \in U$，由于 U 是开集，存在 $\delta > 0$，是的对任意满足 $|y - y_0| < \delta$ 的 y，我们都有 $y \in U$。由于 f 连续，所以存在 $\varepsilon > 0$，使得当 $|x - x_0| < \varepsilon$ 时，$|f(x) - y_0| < \delta$，这表明当 $|x - x_0| < \varepsilon$ 时，$f(x) \in U \Leftrightarrow x \in f^{-1}(U)$，所以 $f^{-1}(U)$ 是开集。

现在假设对任意的开集 U，$f^{-1}(U)$ 是开集，我们来证明 f 是连续的：给定 $x_0 \in \mathbb{R}$，令 $y_0 = f(x_0)$，考虑 y_0 处的开集 $B(y_0, \delta)$（以 y_0 为中心半径为 δ 的小球，即对任意的 $x \in \mathbb{R}$，$a > 0$，我们令 $B(x, a) = (x - a, x + a)$）。由于 $f^{-1}\big(B(y_0, \delta)\big)$ 为开集且 $x_0 \in f^{-1}\big(B(y_0, \delta)\big)$，所以存在 $\varepsilon > 0$，使得 $B(x_0, \varepsilon) \subset f^{-1}\big(B(y_0, \delta)\big)$，从而，$f\big(B(x_0, \varepsilon)\big) \subset B(y_0, \delta)$，这说明对任意的 δ，存在 ε，使得当 $x - x_0| < \varepsilon$ 时，我们有 $|f(x) - f(x_0)| < \delta$，从而 f 在 x_0 处连续。 □

还有一个有用的概念，叫做一个集合的闭包：

定义 62. 给定集合 $X \subset \mathbb{R}$，我们称 $\overline{X}$ 为 X 的**闭包**，其中 $\overline{X}$ 为包含 X 的所有闭集交，即包含 X 的最小闭集。对于 $x \in \mathbb{R}$，如果存在数列 $\{x_k\}_{k \geqslant 1} \subset X$，使得 $\lim\limits_{k \to \infty} x_k = x$，我们就称 x 是 X 的一个**聚点**。

X 的聚点是 X 中子列能收敛到的点。X 中的点都是 X 的聚点。

命题 63. $\overline{X}$ 恰为 X 的聚点所组成的集合。特别地，F 是闭集当且仅当 F 的聚点都在 F 中。

证明：首先证明 $\overline{X}$ 包含 X 的一切聚点：假设 $x \in \mathbb{R}$ 是 X 的聚点，即存在数列 $\{x_k\}_{k \geqslant 1} \subset X$，使得 $\lim\limits_{k \to \infty} x_k = x$，根据 $\overline{X}$ 是闭区间，所以 $x \in \overline{X}$。

其次，我们要证明对任意的 $x \in \overline{X}$，存在数列 $\{x_k\}_{k \geqslant 1} \subset X$，使得 $\lim\limits_{k \to \infty} x_k = x$。

不妨假设 $x \notin X$。那么，对任意的任意 $n > 0$，$B(x, \frac{1}{n}) \cap X \neq \emptyset$，其中，$B(x, \frac{1}{n}) = (x - \dfrac{1}{n}, x + \dfrac{1}{n})$。（如果不然，那么 $B(x, \frac{1}{n}) \cap X = \emptyset$，所以 $X \subset \mathbb{R} - B(x, \frac{1}{n})$。然而 $\mathbb{R} - B(x, \frac{1}{n})$ 是闭集并且包含 X，所以 $\mathbb{R} - B(x, \frac{1}{n}) \supset \overline{X}$，这与 $x \notin \mathbb{R} - B(x, \frac{1}{n})$ 相矛盾）。所以，我们可以选取 $x_k \in B(x, \frac{1}{k}) \cap X$，这个数列的极限就是 x。 □

10.1 作业

清华大学 19-20 秋季学期，数学分析一，作业 4

请用 A4 大小的纸张**正反面**用钢笔，签字笔或者圆珠笔书写，并注明自己的姓名，年级和作业的**总页数**。除定理公式所涉及的人名之外，请使用**中文**。本次作业的提交时间和地点为 **10 月 24 日**上午的课堂上，逾期视作零分。

基本习题

习题 A：距离空间上的拓扑与连续性

A1)（距离空间上的 Heine 定理）假设 (X,d_X) 和 (Y,d_Y) 是距离空间，$f:X\to Y$ 是映射。我们有两种方式来定义连续映射（参见第 3 次作业题的 A3) 和第 9 次课）

- 假设 $x_0\in X$，如果对任意 X 中的点列 $\{x_n\}_{n\geqslant 1}$，$x_n\xrightarrow{d_X}x_0$，我们都有 $f(x_n)\xrightarrow{d_Y}f(x_0)$，我们就称 f 在 x_0 处是连续的。如果 f 在一切 $x\in X$ 处均连续，那么我们就称 f 是距离空间之间的**连续映射**。
- 假设 $x_0\in X$，$y_0=f(x_0)\in Y$。如果对**任意的** $\varepsilon>0$，总存在 $\delta>0$，使得对**任意**满足 $d_X(x,x_0)<\delta$ 的 $x\in X$，都有 $d_Y\big(f(x),f(x_0)\big)<\varepsilon$，我们就称 f 在 x_0 处**连续**。如果 f 在 X 的每个点处都连续，那么我们就称 f 是**连续映射**。

证明，上面两个对连续映射的定义是等价的。

A2) (X,d) 是距离空间。对任意的点 $x\in X$，$r>0$，我们称 $B(x,r)=\big\{y\in X|d(y,x)<r\big\}$ 为以 x 为中心以 r 为半径的开球。证明，对任意的点 $x\in X$，$r>0$，如果 $x'\in B(x,r)$，那么存在 $r'>0$，使得 $B(x',r')\subset B(x,r)$。

如果 $U\subset X$ 是若干开球的并，即 $U=\bigcup\limits_{\alpha\in\mathcal{A}}B(x_\alpha,r_\alpha)$（指标集 $\mathcal{A}$ 是任意的），就称 X 是距离空间 (X,d) 中的**开集**。证明，$U\subset X$ 是开集当且仅当对任意的 $x\in U$，存在 $\delta_x>0$，使得 $B(x,\delta_x)\subset U$。

A3)（距离空间上的标准拓扑）我们用 $\mathcal{T}$ 表示距离空间 (X,d) 上的开集的全体，其中，我们规定 $\emptyset$ 和 X 都是开集。证明，它们满足

1) $\emptyset\in\mathcal{T}$，$X\in\mathcal{T}$。
2) 对任意开集的集合 $\{U_\alpha\}_{\alpha\in\mathcal{A}}$，其中 $\mathcal{A}$ 为指标集合，我们有 $\bigcup\limits_{\alpha\in\mathcal{A}}U_\alpha\in\mathcal{T}$。
3) 对任意有限个开集 $U_1,U_2,\cdots,U_m\in\mathcal{T}$，我们有 $\bigcap\limits_{1\leqslant i\leqslant m}U_i\in\mathcal{T}$。

A4) (X,d) 是距离空间。如果 $F\subset X$ 的补集是开集，我们就称 F 是**闭集**。证明，F 是闭集当且仅当对任意点列 $\{x_n\}_{n\geqslant 1}\in F$，如果 $\lim\limits_{n\to\infty}x_n=x$，那么 $x\in F$。

A5) 证明，距离空间 (X,d) 上的闭集，

1) $\emptyset$ 和 X 都是闭集。

2) 任意多闭集的交集还是闭集。

3) 有限个闭集的并集还是闭集。

A6) 假设 (X,d_X) 和 (Y,d_Y) 是距离空间，$f:X\to Y$ 是映射。那么，如下三个叙述是等价的：

1) f 是连续映射。

2) 对任意 Y 中的开集 U，其逆像 $f^{-1}(U)$ 为 X 中的开集。

2) 对任意 Y 中的闭集 F，其逆像 $f^{-1}(F)$ 为 X 中的闭集。

A7) (X,d) 是距离空间，$A\subset X$ 是子集，我们将包含 A 的所有闭集交 $\overline{A}$ 为 A 的**闭包**，根据上题，这是闭集，所以是包含 A 的最小闭集。对于 $x\in X$，如果存在点列 $\{a_k\}_{k\geqslant 1}\subset A$，使得 $\lim\limits_{k\to\infty} a_k=x$，我们就称 x 是 A 的一个**聚点**。A 的聚点是 A 中子列能收敛到的点，这些点不一定在 A 中的，但是 A 中的点都是 A 的聚点，即 $A\subset\overline{A}$。

证明，$\overline{A}$ 恰为 A 的聚点所组成的集合。特别地，F 是闭集当且仅当 F 的聚点都在 F 中，即 $F=\overline{F}$。

A8)（距离空间的乘积与连续映射）假设 (Y,d_Y) 和 (Z,d_Z) 是距离空间，我们定义 $Y\times Z$ 上的距离函数

$$d_{Y\times Z}:\big(Y\times Z\big)\times\big(Y\times Z\big)\to\mathbb{R}_{\geqslant 0},\quad \big((y_1,z_1),(y_2,z_2)\big)\mapsto\sqrt{d(y_1,y_2)^2+d(z_1,z_2)^2}.$$

证明，$d_{Y\times Z}$ 是 $Y\times Z$ 上的距离函数。证明，两个自然的投影映射是连续的：

$$\pi_Y:Y\times Z\to Y,\ (y,z)\mapsto y;\ \pi_Z:Y\times Z\to Z,\ (y,z)\mapsto z.$$

证明，给定距离空间 (X,d) 和 $(Y\times Z,d_{Y\times Z})$ 之间的映射 $F:X\to Y\times Z$，那么，F 连续当且仅当两个复合映射 $\pi_Y\circ F:X\to Y$ 和 $\pi_Z\circ F:X\to Z$ 都连续。

A9) 证明，加法映射 $+$ 和乘法映射 $\times$ 都是连续映射，其中

$$+:\mathbb{R}^2\to\mathbb{R},\ \ (x,y)\mapsto x+y;\qquad \times:\mathbb{R}^2\to\mathbb{R},\ \ (x,y)\mapsto x\cdot y.$$

A10) 证明，矩阵上的加法映射 $+$ 和乘法映射 $\bullet$ 都是连续映射，其中

$$+:\mathbf{M}_n(\mathbb{R})\times\mathbf{M}_n(\mathbb{R})\to\mathbf{M}_n(\mathbb{R}),\ (A,B)\mapsto A+B;\quad \bullet:\mathbf{M}_n(\mathbb{R})\times\mathbf{M}_n(\mathbb{R})\to\mathbf{M}_n(\mathbb{R}),\ (A,B)\mapsto A\cdot B.$$

A11) 证明，$\mathbf{M}_n(\mathbb{R})$ 上的可逆矩阵的全体 $\mathbf{GL}_n(\mathbb{R})$ 是 $\mathbf{M}_n(\mathbb{R})$ 中的开集。（提示：构造一个连续映射使得该集合是连续映射的逆像）

A12) 证明，取逆映射 $\mathrm{Inv}:\mathbf{GL}_n(\mathbb{R})\to\mathbf{GL}_n(\mathbb{R}),\ \ A\mapsto A^{-1}$ 是连续映射。

函数的连续性

习题 B 用 $\varepsilon-\delta$ 语言证明以下极限等式，其中 n 是正整数：

B1) $\lambda>0$，$\lim\limits_{x\to\infty}\dfrac{x^n}{e^{\lambda x}}=0$。

B2) 给定 $\alpha>0$，我们有

$$\lim_{x\to\infty}x^{\alpha}\log\left(1+\frac{1}{x}\right)=\begin{cases}\infty, & \text{如果}\,\alpha>1\\ 1, & \text{如果}\,\alpha=1\\ 0, & \text{如果}\,0<\alpha<1\end{cases}$$

B3) 右极限 $\lim\limits_{x\to0^+}\dfrac{1}{x^n}e^{-\frac{1}{x^2}}=0$。

B4) 在第九次讲义中我们证明了 $\lim\limits_{x\to0}\dfrac{\sin x}{x}=1$，据此计算并用 $\varepsilon-\delta$ 语言证明这两个结果：

$$\lim_{x\to0}\frac{\cos x-1}{x}\qquad\text{和}\qquad\lim_{x\to0}\frac{\cos x-1}{x^2/2}.$$

习题 C

C1) 证明，$x^3+2x-1=0$ 恰有一个根并且落在 $(0,1)$ 内。

C2) 设 $0\leqslant\lambda<1$，$b>0$，判断方程 $x-\lambda\sin x=b$ 是否有根。

C3) 证明，$\sin x=\dfrac{1}{x}$ 有无穷多个根。

C4) 假设 $f\in C([0,2])$ 并且 $f(0)=f(2)$。证明，$f(x)-f(x+1)=0$ 在 $[0,1]$ 上有根。

C5) 证明，方程 $x^3+3=e^x$ 在 $\mathbb{R}$ 上一定有解。

C6) 假设 $f:[0,2]\to\mathbb{R}$ 是连续函数并且 $f(0)=f(2)$，那么存在 $x\in[1,2]$，使得 $f(x)=f(x-1)$。

C7) $f:\mathbb{R}\to\mathbb{R}$ 是函数，对 $c\in\mathbb{R}$，我们定义 $f^{-1}(c)=\{x\in\mathbb{R}|f(x)=c\}$。证明，如果对任意 $c\in\mathbb{R}$，我们都有 $|f^{-1}(c)|=2$，那么 f 不是连续函数。

C8) 假设连续函数 $f:[a,b]\to\mathbb{R}$ 是单射。如果 $f(a)<f(b)$，证明，f 是严格递增的的。

习题 D 试计算下面函数的极限，其中 m 和 n 是正整数（题号除以 5 余 1 的请写出解答过程，其

余可以只给答案)：

(1) $\lim\limits_{x\to+\infty}\dfrac{a_0x^n+a_1x^{n-1}+\cdots+a_n}{b_0x^m+b_1x^{m-1}+\cdots+b_m}$, $\quad a_0\neq0, b_0\neq0$

(2) $\lim\limits_{x\to+\infty}\dfrac{x^b}{a^x}$, $\quad a>1, b>0$

(3) $\lim\limits_{x\to+\infty}\dfrac{\log x}{x^a}$, $\quad a>0$

(4) $\lim\limits_{x\to0^+}x^a\log x$, $\quad a>0$

(5) $\lim\limits_{x\to\infty}\left(\dfrac{x^2+1}{x^2-2}\right)^{x^2}$

(6) $\lim\limits_{x\to+\infty}(x-\sqrt{x^2-a})$

(7) $\lim\limits_{x\to+\infty}\sqrt{x+1}-\sqrt{x-1}$

(8) $\lim\limits_{x\to0}\dfrac{(1+x)(1+2x)(1+3x)-1}{x}$

(9) $\lim\limits_{x\to1}\dfrac{x+x^2+\cdots+x^n-n}{x-1}$

(10) $\lim\limits_{x\to1}\dfrac{x^{100}-2x+1}{x^{50}-2x+1}$

(11) $\lim\limits_{x\to1}\left(\dfrac{m}{1-x^m}-\dfrac{n}{1-x^n}\right)$,

(12) $\lim\limits_{x\to0}\dfrac{(1+x)^a-1}{x}$

(13) $\lim\limits_{x\to1}\dfrac{x^a-1}{x^b-1}$

(14) $\lim\limits_{x\to+\infty}(\log x)^{1/x}$

(15) $\lim\limits_{x\to0}\left(\dfrac{a^x+b^x}{2}\right)^{1/x}$, $\quad a>0, b>0$

(16) $\lim\limits_{x\to+\infty}\left(\sqrt[k]{(x+a_1)(x+a_2)\cdots(x+a_k)}-x\right)$

(17) $\lim\limits_{x\to0}\dfrac{(\sqrt{1+x^2}+x)^n-(\sqrt{1+x^2}-x)^n}{x}$

(18) $\lim\limits_{x\to\frac{\pi}{2}}(\sin x)^{\tan x}$

(19) $\lim\limits_{x\to\infty}\left(\sin\dfrac{1}{x}+\cos\dfrac{1}{x}\right)^x$

(20) $\lim\limits_{x\to\infty}\dfrac{\sqrt{x+\sqrt{x+\sqrt{x}}}}{x^\alpha}, \alpha>0$

(21) $\lim\limits_{x\to0+}\dfrac{\sqrt{x+\sqrt{x+\sqrt{x}}}}{x^\alpha}, \alpha>0.$

思考题 (不交作业)

问题 E 假设 $A\subset\mathbb{R}$ 是一个可数集。证明，存在单调函数 $f:\mathbb{R}\to\mathbb{R}$，使得 f 的不连续点的集合恰好是 A。

问题 F 函数 $f:[0,1]\to[0,1]$ 是递增的函数，证明，f 有不动点。

问题 G（关于 $[0,1]$ 上同胚的共轭问题）考虑 $[0,1]$ 到自身的自同胚，即

$$\mathrm{Homeo}([0,1])=\big\{f:[0,1]\to[0,1]\big|f\text{ 为连续的双射}\big\}$$

我们知道对于任意的 $f\in\mathrm{Homeo}([0,1])$, $f^{-1}\in\mathrm{Homeo}([0,1])$。假设 $f\in\mathrm{Homeo}([0,1])$ 并且 0 和 1 是它仅有的不动点，$g\in\mathrm{Homeo}([0,1])$ 并且 0 和 1 也是它仅有的不动点，证明，存在 $h\in\mathrm{Homeo}([0,1])$，使得

$$h^{-1}\circ f\circ h=g.$$

11 紧性与开覆盖，Heine-Borel 定理，Lebesgue 数，一致连续，函数列的逐点收敛与一致收敛，闭区间上连续函数空间 $\big(C\big([a,b]\big),\|\cdot\|_\infty\big)$

二零一九年十月二十一日，星期一，晴天

我们先来回顾一个例子：对多元函数 $f(x)=f(x_1,\cdots,x_n)$，如果将 $x_2,\cdots,x_n$ 固定而将 f 视作是 x_1 的函数，f 此时对 x_1 连续，类似地，f 对其它变量也连续的，但是这样不能推导出 f 做为多元函数 $\mathbb{R}^n\to\mathbb{R}$ 是连续的。我们考察了函数：

$$f(x,y)=\begin{cases}\frac{xy}{x^2+y^2}, & (x,y)\neq(0,0)\\ 0, & (x,y)=(0,0).\end{cases}$$

其中，我们用 x 和 y 来表示 $\mathbb{R}^2$ 上面常用的坐标系。这个函数对两个变量分别连续，但是我们考虑了收敛到 $(0,0)$ 点列 $\left\{\frac{1}{n}(1,\lambda)\right\}_{n\geqslant 1}$ 并发现

$$f\big(\frac{1}{n}(1,\lambda)\big)=\frac{\lambda}{1+\lambda^2}\not\to 0.$$

从而，f 在 $(0,0)$ 处不连续。

我们还可以讲这个函数用极坐标系 (r,ϑ) 来表达：

$$f(r,\vartheta)=\begin{cases}\frac{1}{2}\sin(2\vartheta), & r\neq 0\\ 0, & r=0.\end{cases}$$

很明显，f 对 r 这个变量不连续。

我们强调过，尽管习惯如此，但是我们不愿意将 f 写成 $f()$ 的形式，因为坐标系只是对这个对象 f 的一种描述方式，而这个 f 的连续性是不依赖于坐标系选择的（只依赖于定义域和值域上的距离的定义）。

在第四次作业中，我们已经将开集和闭集的概念推广到了一般的距离空间，特别地，我们可以在 $\mathbb{R}^n$ 讨论开集和闭集的概念。我们做一下简单的回顾：(X,d) 是距离空间。对任意的点 $x\in X$，$r>0$，我们称 $B(x,r)=\big\{y\in X|d(y,x)<r\big\}$ 为以 x 为中心以 r 为半径的开球。如果 $U\subset X$ 是若干开球的并，即 $U=\bigcup\limits_{\alpha\in\mathcal{A}}B(x_\alpha,r_\alpha)$（指标集 $\mathcal{A}$ 是任意的），就称 X 是距离空间 (X,d) 中的**开集**。证明，$U\subset X$ 是开集当且仅当对任意的 $x\in U$，存在 $\delta_x>0$，使得 $B(x,\delta_x)\subset U$。我们用 $\mathcal{T}$ 表示距离空间 (X,d) 上的开集的全体，并且强行规定 $\emptyset$ 和 X 都是开集。$\mathcal{T}$ 满足（请比较命题 59）：

1) $\emptyset\in\mathcal{T}$，$X\in\mathcal{T}$。

2) 对任意开集的集合 $\{U_\alpha\}_{\alpha\in\mathcal{A}}$，其中 $\mathcal{A}$ 为指标集合，我们有 $\bigcup\limits_{\alpha\in\mathcal{A}}U_\alpha\in\mathcal{T}$。

3) 对任意有限个开集 $U_1,U_2,\cdots,U_m\in\mathcal{T}$，我们有 $\bigcap\limits_{1\leqslant i\leqslant m}U_i\in\mathcal{T}$。

如果 $F \subset X$ 的补集是开集，我们就称 F 是**闭集**。类似于 $\mathbb{R}$ 的情况，F 是闭集当且仅当对任意点列 $\{x_n\}_{n\geqslant 1} \in F$，如果 $\lim\limits_{n\to\infty} x_n = x$，那么 $x \in F$。我们还知道，任意多闭集的交集还是闭集，有限个闭集的并集还是闭集。特别地，两个距离空间 (X, d_X) 和 (Y, d_Y) 之间的映射 $f : X \to Y$ 是连续的当且仅当对任意 Y 中的开集 U，其逆像 $f^{-1}(U)$ 为 X 中的开集。

练习. (X, d) 是距离空间，证明，一个点所构成的集合是紧集。

我们现在引入紧集的概念：

定义 64 (开覆盖与紧性)**.** (X, d) 是距离空间，$S \subset X$ 是子集。如果 (X, d) 中开集的集合 $\mathcal{U} = \{U_\alpha\}_{\alpha\in A}$ 满足 $S \subset \bigcup\limits_{\alpha\in A} U_\alpha$，我们就把 $\mathcal{U} = \{U_\alpha\}_{\alpha\in A}$ 称作是 S 的一个开覆盖。

我们考虑 $\mathcal{U}$ 的子集 $\mathcal{U}'$，即 $\mathcal{U}' = \{U_{\alpha'}\}_{\alpha'\in A'}$，其中 $A' \subset A$。如果 $X \subset \bigcup\limits_{\alpha'\in A'} U_{\alpha'}$，我们就把 $\mathcal{U}' = \{U_{\alpha'}\}_{\alpha'\in A'}$ 称作是 $\mathcal{U} = \{U_\alpha\}_{\alpha\in A}$ 的一个子覆盖。

$K \subset X$ 是子集，如果对 K 的任意开覆盖 $\mathcal{U} = \{U_\alpha\}_{\alpha\in A}$，都能找到一个有限的子覆盖 $\mathcal{U}' = \{U_{\alpha'}\}_{\alpha'\in A'}$，即 A' 是有限集，我们就称 K 是**紧集**。

我们最关心的例子自然是 $\mathbb{R}$（和 $\mathbb{R}^n$）上的紧集，我们将证明，有界的闭区间 $[a, b]$ 是紧集（这是一个大定理）。

命题 65. 紧集在连续映射下被保持，即若 $f : (X, d_X) \to (Y, d_Y)$ 是距离空间之间的连续映射，如果 $K \subset X$ 是紧集，那么 $f(K) \subset Y$ 也是紧集。

证明: 考虑 $\mathcal{U} = \{U_\alpha\}_{\alpha\in A}$ 是 Y 中 $f(K)$ 的开覆盖，那么根据 $K \subset f^{-1}\big(f(K)\big)$，$f^{-1}\big(\mathcal{U}\big) = \{f^{-1}(U_\alpha)\}_{\alpha\in A}$ 是 X 中 K 的开覆盖，从而有有限的子覆盖 $\{f^{-1}(U_{\alpha'})\}_{\alpha'\in A'}$，从而 $\{U_{\alpha'}\}_{\alpha'\in A'}$ 覆盖了 $f(K)$，证毕。 □

注记. 我们知道开集和闭集在连续映射的逆下被保持，但是通常不被连续映射保持，请举出反例。

为了刻画 $\mathbb{R}$ 和 $\mathbb{R}^n$ 上的紧集，我们先证明引理：

命题 66 (Lebesgue 数)**.** 假设 $K \subset \mathbb{R}$ 是有界闭集，$\mathcal{U} = \{U_\alpha\}_{\alpha\in A}$ 是 K 的开覆盖。那么，存在 $\delta > 0$（习惯上被称作是开覆盖 $\mathcal{U}$ 的一个 *Lebesgue* 数），使得对任意的 $x, y \in K$，如果 $|x - y| < \delta$，那么存在 $\alpha \in A$，使得 $[x, y] \cap K \subset U_\alpha$。

证明: 我们利用反证法：如果不然，那么每个 $\delta = \dfrac{1}{n}$，存在 $x_n, y_n \subset K$（不妨假设 $x_n < y_n$），使得 $|x_n - y_n| \leqslant \dfrac{1}{n}$（特别地，$\lim\limits_{n\to\infty} |x_n - y_n| = 0$），但是对任意的 $\alpha \in A$，U_α 不能完整的覆盖住 $[x_n, y_n] \cap K$，即 $[x_n, y_n] \cap K \not\subset U_\alpha$。

由于 K 是有界的，所以数列 $\{x_n\}_{n\geqslant 1}$ 和 $\{y_n\}_{n\geqslant 1}$ 是有界的，通过选取子序列，我们可以假设当 $n \to \infty$ 时，$x_n \to x$，$y_n \to y$。根据 $\lim\limits_{n\to\infty} |x_n - y_n| = 0$，自然有 $x = y$。又因为 K 是闭集，所以 $x = y \in K$。

由于 $\mathcal{U}$ 是 K 的开覆盖，$x\in K$，所以存在 $U_\alpha\in\mathcal{U}$，使得 $x\in U_\alpha$。根据 U_α 是开集，那么存在开区间 $(x-\varepsilon,x+\varepsilon)\subset U_\alpha$，此时，根据 $x_n\to x$，$y_n\to x$，选取很大的 N，使得 $[x_N,y_N]\subset U_\alpha$，所以 $[x_N,y_N]\cap K\subset U_\alpha$，矛盾。 □

定理 67 (Heine-Borel)**.** *假设 $K\subset\mathbb{R}$。那么，K 是紧集当且仅当 K 是有界闭集。*

推论 68. *闭区间 $[a,b]$ 是紧集。特别地，任意选取 $[a,b]$ 的一个开区间 覆盖 $\mathcal{U}=\{I_\alpha\}_{\alpha\in A}$，其中 $I_\alpha=(a_\alpha,b_\alpha)$ 是开区间，$[a,b]\subset\bigcup\limits_{\alpha\in A}I_\alpha$，我们都能找到有限个开区间 $I_k=(a_k,b_k)(k=1,2,\cdots,N)$，使得 $[a,b]\subset\bigcup\limits_{k=1}^{N}I_k$。*

Heine-Borel 定理的证明. 首先假设 K 是紧集，我们分两步证明 K 是有界闭集。

- K 是有界的：由于 $\bigcup\limits_{n=1}^{\infty}(-n,n)=\mathbb{R}$，所以 $\bigcup\limits_{n=1}^{\infty}(-n,n)\supset K$，据此，我们有 K 的开覆盖 $\mathcal{U}=\big\{(-n,n)\big|n=1,2,\cdots\big\}$。根据 K 的紧性，可以找到有限个的开区间 $(-n_1,n_1),\cdots,(-n_k,n_k)$，使得它们的并集包含 K。不妨假设 $n_1<n_2<\cdots<n_k$。很明显，$K\subset(-n_k,n_k)$，所以有界。
- K 是闭集：利用反证法，如若不然，存在序列 $\{x_i\}_{i\geqslant1}\subset K$，$\lim\limits_{i\to\infty}x_i=x$ 但是 $x\notin K$。通过选取子序列，我们还可以进一步假设对任意的 $i\geqslant1$，$|x_i-x|<\dfrac{1}{i}$。

  考虑下降的闭区间序列 $F_n=[x-\dfrac{1}{n},x+\dfrac{1}{n}]$，我们自然有 $\bigcap\limits_{n\geqslant1}F_n=\{x\}$。我们注意到 $U_n=\mathbb{R}-F_n$ 是开集并且 $\bigcup\limits_{n\geqslant1}U_n\supset K$（因为 $x\notin K$）！据此，我们得到据此 K 的开覆盖 $\mathcal{U}=\big\{U_n\big|n=1,2,\cdots\big\}$，所以 K 的紧性意味着存在 $U_{n_1},U_{n_2},\cdots,U_{n_k}$，使得 $K\subset\bigcup\limits_{i=1}^{k}U_{n_i}$。不妨假设 $n_1<n_2<\cdots<n_k$，所以 $K\subset U_{n_k}$。根据 F_{n_k} 的定义，x 与任意一个 K 中的点的距离至少是 $\dfrac{1}{n_k}$，这与 $\lim\limits_{i\to\infty}x_i=x$ 矛盾。

其次，在 K 是有界闭集的假设下证明 K 是紧集。

任意给定 K 的开覆盖 $\mathcal{U}=\{U_\alpha\}_{\alpha\in A}$，根据前一命题，我们可以选取该覆盖的一个 Lebesgue 数 δ，通过将 δ 适当缩小，我们不妨假设 $\delta=\dfrac{1}{N}$，其中 $N\in\mathbb{Z}_{>0}$。我们将 K 砍成若干长度不超过 $\dfrac{1}{N}$ 小段 $K=\bigcup\limits_{k\in\mathbb{Z}}K_k$，其中 $K_k=K\cap[\dfrac{k}{N},\dfrac{k+1}{N}]$。根据有界性，只有有限个 K_k 是非空的，所以我们有 $K=\bigcup\limits_{|k|\leqslant k_0}K_k$。利用 Lebesgue 数的定义，对每个 $|k|\leqslant k_0$，存在 $U_k\in\mathcal{U}$，使得 $K_k\subset U_k$，所以 $K\subset\bigcup\limits_{|k|\leqslant k_0}U_k$，这就给出了有限的子覆盖。 □

推论 69 (紧性和列紧性的等价性)**.** *$K\subset\mathbb{R}$ 是子集，如果对任意 K 中的子序列 $\{x_k\}_{k\geqslant1}\subset K$，都存在收敛的子序列 $\{x_{k_j}\}_{j\geqslant1}$，使得 $\lim\limits_{j\to\infty}x_{k_j}\in K$，我们就称 X 是**列紧的**。*

那么，K 是紧集当且仅当 K 是列紧的。

证明：假设 K 是紧集，那么它是有界的，所以对任意的 K 中的子序列 $\{x_k\}_{k\geqslant 1}\subset K$，存在收敛的子序列 $\{x_{k_j}\}_{j\geqslant 1}$，又因为 K 是闭集，所以这个极限仍然在 K 中，从而 K 是列紧的。

假设 K 是列紧的，一个重要的观察是上面关于 Lebesgue 数的证明仍然成立（只用到了列紧性），据此，我们可以原封不动地重复 Heine-Borel 定理中第二步关于 K 是紧集的证明即可。（细节留给不放心的同学揣摩） □

我们可以讲上面的推理加以简单的改造从而得到关于距离空间上紧性的若干结论，但是我们需要先预警一下，不是每个结论都可以推广，比如说在一般的距离空间上，紧性可以推出有界闭性，但是两者不等价。我们这些命题整理为如下几条，其中 (X,d) 是距离空间：

1) 假设 $K\subset X$ 是紧集，那么 K 是有界闭集。其中，K 在距离空间 X 中有界指的是，存在 $x_0\in X$ 和 $R>0$，使得 $K\subset B(x_0,R)$。

   先证明 K 是有界的：任意选定 $x\in X$，由于 $\bigcup\limits_{n=1}^{\infty}B(x,n)=X$，所以 $\bigcup\limits_{n=1}^{\infty}B(x,n)\supset K$，据此，我们有 K 的开覆盖 $\mathcal{U}=\big\{B(x,n)\big|n=1,2,\cdots\big\}$。根据 K 的紧性，可以找到有限个开球 $B(x,n_1),\cdots,B(x,n_k)$，使得它们的并集包含 K。不妨假设 $n_1<n_2<\cdots<n_k$。很明显，$K\subset B(x,n_k)$，所以有界。

   再证明 K 是闭集：利用反证法，如若不然，存在序列 $\{x_i\}_{i\geqslant 1}\subset K$，$\lim\limits_{i\to\infty}x_i=x$ 但是 $x\notin K$。通过选取子序列，我们还可以进一步假设对任意的 $i\geqslant 1$，$d(x_i,x)<\dfrac{1}{i}$。

   考虑下降的序列 $F_n=\{y\in X|d(y,x)\leqslant\dfrac{1}{n}\}$，首先注意到

   $$f:X\to\mathbb{R},\quad y\mapsto d(y,x),$$

   是连续映射，所以 $F_n=f^{-1}\big([-\dfrac{1}{n},\dfrac{1}{n}]\big)$ 是闭集。我们自然有 $\bigcap\limits_{n\geqslant 1}F_n=\{x\}$。注意到 $U_n=X-F_n$ 是开集并且 $\bigcup\limits_{n\geqslant 1}U_n\supset K$（因为 $x\notin K$）！据此，我们得到据此 K 的开覆盖 $\mathcal{U}=\big\{U_n\big|n=1,2,\cdots\big\}$，所以 K 的紧性意味着存在 $U_{n_1},U_{n_2},\cdots,U_{n_k}$，使得 $K\subset\bigcup\limits_{i=1}^{k}U_{n_i}$。不妨假设 $n_1<n_2<\cdots<n_k$，所以 $K\subset U_{n_k}$。根据 F_{n_k} 的定义，x 与任意一个 K 中的点的距离至少是 $\dfrac{1}{n_k}$，这与 $\lim\limits_{i\to\infty}x_i=x$ 矛盾。

2) **定理** 假设 (X,d) 列紧的度量空间（即如果对任意 X 中的点列 $\{x_k\}_{k\geqslant 1}$，都存在收敛的子序列 $\{x_{k_j}\}_{j\geqslant 1}$，即 $\lim\limits_{j\to\infty}x_{k_j}$ 存在），$\mathcal{U}=\{U_\alpha\}_{\alpha\in A}$ 是 X 的开覆盖。那么，存在 $\delta>0$（习惯上被称作是开覆盖 $\mathcal{U}$ 的一个 Lebesgue 数），使得对任意的 $x\in X$，存在 $\alpha\in A$，使得 $B(x,\delta)\subset U_\alpha$。

   我们利用反证法：如若不然，那么每个 $\delta=\dfrac{1}{n}$，存在 $x_n\in X$，使得 $B(x_n,\dfrac{1}{n})$ 不被任意一个开集所包含，即对任意的 $\alpha\in A$，$B(x_n,\dfrac{1}{n})\not\subset U_\alpha$。根据列紧性，我们可以选取子列，使得 $k\to\infty$ 时，$x_{n_k}\to x$。由于 $\mathcal{U}$ 是开覆盖，所以存在 $U_\alpha\in\mathcal{U}$，使得 $x\in U_\alpha$。根据 U_α 是

开集，那么存在开球 $B(x,r)\subset U_\alpha$。由于 $x_{n_k}\to x$，选取很大的 N，当 $k\geqslant N$ 时，使得 $B(x_{n_k},\dfrac{1}{n_k})\subset U_\alpha$，矛盾。

3) **定理** 假设 (X,d) 度量空间，它是列紧的当且仅当它是紧的。

首先证明，如果 X 是列紧的，那么 X 必然是紧的：任意给定 X 的开覆盖 $\mathfrak{U}=\{U_\alpha\}_{\alpha\in A}$，根据 2)，我们选取该覆盖的一个 Lebesgue 数 δ。我们通过归纳的方式构造一个点列（可以是有限点列）：任选 x_0，那么存在某个 $U_0\in\mathfrak{U}$，使得 $U_0\supset B(x_0,\delta)$。假定 x_k 已经选定，那么存在某个 $U_k\in\mathfrak{U}$，使得 $U_k\supset B(x_k,\delta)$。现在分两种情况：

- 如果 $B(x_0,\delta)\cup B(x_1,\delta)\cup\cdots\cup B(x_k,\delta)=X$，那么我们已经找到了有限子覆盖 $U_1,\cdots,U_k\in\mathfrak{U}$，这个过程到此结束。
- 如果 $B(x_0,\delta)\cup B(x_1,\delta)\cup\cdots\cup B(x_k,\delta)\neq X$，那么我们就任选 $x_{k+1}\notin B(x_0,\delta)\cup B(x_1,\delta)\cup\cdots\cup B(x_k,\delta)$，然后继续上面的过程。我们注意到这一步选取的 x_{k+1} 使得 $d(x_{k+1},x_j)\geqslant\delta$，其中 $j=0,1,2,\cdots,k$。

当然，上面的第二种选择不可能无限地进行，否则我们得到一个点列 $\{x_k\}_{k\geqslant1}$，使得任意两个点之间的距离都不小于 δ，这与列紧性矛盾。所以，到某一步我们就在第一种选择上结束了，这就给出了有限的子覆盖。

其次证明，如果 X 是紧的，那么 X 必然是列紧的：假设 $\{x_k\}_{k\geqslant1}$ 是一列点，如果存在点 $x\in X$，使得对任意的 $\delta>0$，总存在 $B(x,\delta)$，使得 $B(x,\delta)\cap\{x_k\}_{k\geqslant1}\neq\emptyset$，那么 $\{x_k\}_{k\geqslant1}$ 有收敛的子列：因为我们对 $\delta=\dfrac{1}{i}$，取 $x_{k_i}\in B(x,\delta)\cap\{x_k\}_{k\geqslant1}$ 即可。

我们用反证法：如果 $\{x_k\}_{k\geqslant1}$ 没有收敛子列，那么对任意的 $x\in X-\{x_k\}_{k\geqslant1}$，存在 $\delta>0$，使得 $B(x,\delta)\cap\{x_k\}_{k\geqslant1}=\emptyset$，这表明 $\{x_k\}_{k\geqslant1}$ 是 X 中的一个闭集。同样的推理表明，$\{x_k\}_{k\geqslant m}$ 也是 X 中的闭集，所以 $U_m=X-\{x_k\}_{k\geqslant m}$ 是一族开集。很明显，$\bigcup\limits_{m\geqslant1}U_m=X$，利用紧性，我们有 $\bigcup\limits_{m\leqslant m_0}U_m=X$，这表明 $\{x_k\}_{k\geqslant1}$ 是有限的点集，自然收敛。

推论 70 (Heine-Borel)**.** *假设 $K\subset\mathbb{R}^n$。那么，K 是紧集当且仅当 K 是有界闭集。*

证明: 首先，假设 K 是紧集，在一般的距离中我们已经证明了 K 是有界闭集。

其次，假设 K 是有界闭集（在 $\mathbb{R}^n$ 中显然是列紧的，因为我们可以看每个坐标），我们要证明 K 是紧集。

任意给定 K 的开覆盖 $\mathfrak{U}=\{U_\alpha\}_{\alpha\in A}$，列紧性表明我们可以选取该覆盖的一个 Lebesgue 数 δ，通过将 δ 适当缩小，我们不妨假设 $\delta=\dfrac{1}{N}$，其中 $N\in\mathbb{Z}_{>0}$。假设 K 落在方体 $[-M,M]\times[-M,M]\times\cdots\times[-M,M]$，其中 M 是正整数。通过将我们将这个方体分解成 $\left(\dfrac{2M}{N}\right)^n$ 个小方体 $\{Q_i\}$，K 砍成有限个小块 $Q_i\cap K$。利用 Lebesgue 数的定义，每个小块 $Q_i\cap K$ 都包含在某个 U_i 中，所以 $K\subset\bigcup U_i$，这就给出了有限的子覆盖。 □

一致连续性与一致收敛

定义 71 (一致收敛性)**.** 假设 $f: X \to \mathbb{R}$ 是连续函数，如果对任意的 $\varepsilon > 0$，存在 $\delta > 0$，使得对任意的 $x, y \in X$，只要 $d(x,y) < \delta$，就有 $|f(x) - f(y)| < \varepsilon$，我们就称 f 在 X 上**一致连续**。

例子. 考虑 $(0,1)$ 上的函数 $f(x) = \dfrac{1}{x}$，我们希望研究 $f(x)$ 在 x_0 处的连续性。

按照定义，f 在 x_0 处连续，指的是：对任意的 $\varepsilon > 0$，存在 $\delta > 0$（注意到，这个 δ 可能依赖 x_0），使得当 $|x - x_0| < \delta$ 时，我们就有 $|f(x) - f(x_0)| < \varepsilon$。

对于 $f(x) = \dfrac{1}{x}$，我们不妨先假设 x 已经离着 x_0 很近：$\dfrac{1}{2}x_0 \leqslant x \leqslant 2x_0$，那么

$$\varepsilon > \left|\frac{1}{x} - \frac{1}{x_0}\right| = \frac{|x - x_0|}{x_0 x} \geqslant \frac{|x - x_0|}{2(x_0)^2}.$$

所以，要想 $|f(x) - f(x_0)| < \varepsilon$，我们需要 $\delta < 2(x_0)^2\varepsilon$。由此可见，$x_0$ 越小，需要选取的 δ 就越小，这就给出了 δ 对 x_0 的依赖性。然而，在一致连续的概念中，δ 的选取不依赖于 x_0，所以 $(0,1)$ 区间上定义的 $\dfrac{1}{x}$ 不是一致连续的。（$[0.1, 1]$ 区间上定义的 $\dfrac{1}{x}$ 是一致连续的!）

在作业题中，我们会见到很多连续和一致连续的例子。

定理 72. $I = [a,b]$ 是有界闭区间，每个 $f \in C(I)$ 都是一致连续的函数。

证明: 给定 $f \in C(I)$，通过定义

$$f(x) = \begin{cases} f(a), & x \leqslant a; \\ f(x), & a \leqslant x \leqslant b; \\ f(b), & x \geqslant b. \end{cases}$$

任意给定 $\varepsilon > 0$，对每个 $x \in I$，存在开区间 $I_x = (x - \delta_x, x + \delta_x)$，使得对任意的 $y \in I_x$（等价于说 $|y - x| < \delta_x$），都有 $|f(y) - f(x)| < \dfrac{1}{2}\varepsilon$。特别地，对任意的 $y, z \in I_x$，我们都有

$$|f(y) - f(z)| \leqslant |f(y) - f(x)| + |f(z) - f(x)| < \varepsilon.$$

我们得到 I 一个开覆盖 $\{I_x\}_{x \in I}$。令 δ 为这个开覆盖的 Lebesgue 数，对于 $|x - y| < \delta$，按照 Lebesgue 数的定义，x 和 y 落在同一个 I_{x_i} 里面，从而 $|f(x) - f(y)| < \varepsilon$。这就证明了一致连续性。 □

我们现在引入关于函数的两种收敛的概念：(X,d) 是距离空间，$\{f_n : I \to \mathbb{R}\}_{n \geqslant 1}$ 和 $f : I \to \mathbb{R}$ 是函数的序列。我们定义：

- **逐点收敛.** 如果对每个点 $x \in X$，函数值 $f_n(x) \to f(x)$，即 $\lim\limits_{n \to \infty} |f_n(x) - f(x)| = 0$，我们就称 f_n 在 X 上逐点收敛到 f。
- **一致收敛.** 如果 $\lim\limits_{n \to \infty} \sup\limits_{x \in X} |f_n(x) - f(x)| = 0$，我们就称 f_n 在 X 上一致收敛到 f。

注记. 逐点收敛指的是对任意的 $x \in I$，对任意的 $\varepsilon > 0$，存在 N（可能依赖于 ε 和 x），使得当 $n \geqslant N$ 时，我们有 $|f_n(x) - f(x)| < \varepsilon$。

一致收敛指的是对任意的 ε，存在 N（只依赖于 ε 不依赖于点 x），使得当 $n \geqslant N$ 时，对任意的 $x \in I$，我们都有 $|f_n(x) - f(x)| < \varepsilon$。换句话说，$N$ 的选取对于 x 是一致的（即不依赖于 x）。

特别地，函数列的一致收敛能推出逐点收敛。

我们先看几个例子：

1) $I = (0,1)$，$f_n(x) = \dfrac{1}{nx}$，对于每个 x，$\lim\limits_{n\to\infty} f_n(x) = 0$，所以 $f_n(x)$ 逐点收敛到 0（函数）；然而，对任意的 $\sup\limits_{x\in I} |f_n(x) - 0| = \infty$，所以 $\{f_n\}_{n\geqslant 1}$ 不一致收敛。

2) $I = [-2,2]$，$f_n(x) = \begin{cases} -1, & x \leqslant -\dfrac{1}{n} \\ nx, & -\dfrac{1}{n} \leqslant x \leqslant \dfrac{1}{n} \\ 1, & x \geqslant \dfrac{1}{n}. \end{cases}$

f_n 逐点收敛到函数 $f(x) = \begin{cases} -1, & x < 0; \\ 0, & x = 0; \\ 1, & x > 0. \end{cases}$ 我们注意到极限函数并不连续。不难看出，f_n 并不一致收敛。

有了以上的准备工作，我们现在研究闭区间上实数（或者复数）值连续函数空间 $C\big([a,b]\big)$，其中 $a < b$（否则不是很有意思）。这是一个无限维的 $\mathbb{R}$-线性空间。我们在 $C\big([a,b]\big)$ 上定义一个范数

$$\|\cdot\|_\infty : C\big([a,b]\big) \to \mathbb{R}_{\geqslant 0}, \quad f \mapsto \|f\|_\infty = \sup_{x\in[a,b]} |f(x)|.$$

我们注意到，$|f|$ 也是闭区间 $[a,b]$ 上的连续函数，所以有界，从而 $\sup\limits_{x\in[a,b]} |f(x)| < \infty$ 是良好定义的。这个范数定义了 $C\big([a,b]\big)$ 上的距离 $d_\infty(f,g) = \sup\limits_{x\in I} |f(x) - g(x)|$。

我们应该将 $\Big(C\big([a,b]\big), +, \cdot, \|\cdot\|_\infty\Big)$ 和实数 $(\mathbb{R}, +, \cdot, |\cdot|)$ 做类比。

另外，在一致收敛的概念中，我们要求 $\lim\limits_{n\to\infty} \sup\limits_{x\in X} |f_n(x) - f(x)| = 0$，也就是说，$f_n$ 是按照距离函数 d_∞ 在距离空间 $\Big(C\big([a,b]\big), d_\infty\Big)$ 里收敛。所以，把函数视为点，所谓一致收敛的概念变成了我们熟悉的点列收敛的概念。

定理 73. $\big(C([a,b]), \|\cdot\|_\infty\big)$ 是完备赋范线性空间，即 $\big(C([a,b]), d_\infty\big)$ 是完备的距离空间。

注记. $C(I)$ 的完备性是关于连续函数最重要的性质之一。另外，我们强调这个性质和之前关于连续函数看法完全不同：这不是关于一个函数的性质而是关于一族（或者所有）连续函数的性质。

从证明的角度而言，一致连续性将起重要的作用，而逐点地考虑这个问题是徒劳的，因为有无限多个点。由于一致连续性的证明依赖于紧性，我们可以理解为什么函数所定义的空间的紧性（几何性质）非常关键：这个概念提供了从无限到有限的途径！

证明：假设 $\{f_n\}_{n\geqslant 1}$ 是 $\big(C([a,b]),\|\cdot\|_\infty\big)$ 中 Cauchy 列，即对任意的 $\varepsilon>0$，存在 N，使得当 $n,m\geqslant N$ 时，我们有 $d_\infty(f_n,f_m)<\varepsilon$。我们的目标是构造 $f\in C\big([a,b]\big)$，使得 $f_n\xrightarrow{d_\infty} f$。

首先定义函数 f：对任意 $x\in I$，按照定义，对任意的 m 和 n，我们的都有 $|f_n(x)-f_m(x)|\leqslant d(f_n,f_m)$，所以 $\{f_n(x)\}_{n\geqslant 1}$ 是 Cauchy 数列，据此，可以定义

$$f(x)=\lim_{n\to\infty} f_n(x),\quad \forall x\in I.$$

问题的关键在于证明 $f(x)$ 是连续的。

任选 $x_0\in I$，我们证明 f 在 x_0 处连续：对任意的 $\varepsilon>0$，先选取 N，当 $n,m\geqslant N$ 时，有 $d(f_n,f_m)<\dfrac{1}{9}\varepsilon$。特别地，$|f(x_0)-f_N(x_0)|\leqslant\dfrac{1}{9}\varepsilon$。由于 f_N 在 I 上一致连续，所以存在 $\delta>0$，使得当 $|x-y|<\delta$ 时，我们有 $|f_N(x)-f_N(y)|<\dfrac{1}{9}\varepsilon$。从而，对任意的 $n\geqslant N$，当 $|x-x_0|<\delta$ 时，我们有

$$\begin{aligned}|f_n(x)-f_n(x_0)|&\leqslant|f_n(x)-f_N(x)|+|f_N(x)-f_N(x_0)|+|f_n(x_0)-f_N(x_0)|\\&\leqslant d(f_n,f_N)+\frac{1}{9}\varepsilon+d(f_n,f_N)<\frac{1}{9}\varepsilon+\frac{1}{9}\varepsilon+\frac{1}{9}\varepsilon=\frac{1}{3}\varepsilon.\end{aligned}$$

此时，对任意满足 $|x-x_0|<\delta$ 的 x，存在 $n\geqslant N$，使得 $|f_n(x)-f(x)|<\dfrac{1}{3}\varepsilon$，从而，我们有

$$\begin{aligned}|f(x)-f(x_0)|&\leqslant|f(x)-f_n(x)|+|f_n(x)-f_n(x_0)|+|f_n(x_0)-f(x_0)|\\&<\frac{1}{3}\varepsilon+\frac{1}{3}\varepsilon+\frac{1}{3}\varepsilon\leqslant\varepsilon.\end{aligned}$$

这就证明了 f 是连续函数。

最终，我们说明 f_n 一致收敛到 f，即 $\lim\limits_{n\to\infty} d_\infty(f_n,f)=0$。首先，根据 $\{f_n(x)\}_{n\geqslant 1}$ 是 Cauchy 数列，存在 N_0，使得当 $n,m\geqslant N_0$ 时，我们有 $\sup\limits_{x\in[a,b]}|f_n(x)-f_m(x)|<\dfrac{1}{2}\varepsilon$。

对任意的 $x\in[a,b]$，根据逐点收敛性，存在 $N_x\geqslant N_0$，使得当 $n\geqslant N_x$ 时，我们有 $|f_n(x)-f(x)|<\dfrac{1}{4}\varepsilon$，利用连续性，存在开区间 I_x，使得对任意的 $y\in I_x$，我们都有 $|f_n(y)-f(y)|<\dfrac{1}{2}\varepsilon$。根据上面 N_0 的选取，我们知道对任意的 $m\geqslant N_x$，任意的 $y\in I_x$，我们都有

$$|f_m(x)-f(x)|\leqslant|f_m(y)-f_{N_x}(y)|+|f_{N_x}(y)-f(y)|<\varepsilon.$$

这样的开区间的集合 $\{I_x\}_{x\in[a,b]}$ 自然是 $[a,b]$ 的一个开覆盖，根据紧性，我们可以选取有限的开覆盖 $I_{x_1},\cdots,I_{x_\ell}$。在每个 I_{x_j} 上，当 $n\geqslant N_{x_j}$ 时，对任意的 $y\in I_{x_j}$，$|f_n(y)-f(y)|<\varepsilon$。所以，只要取 $N=\max(N_{x_1},N_{x_2},\cdots,N_{x_\ell})$，当 $n\geqslant N$ 时，在每个每个 I_{x_j} 上，都有

$$|f_n(y)-f(y)|<\varepsilon,\quad \forall y\in I_{N_j}.$$

由于 $I_{x_1},\cdots,I_{x_\ell}$ 覆盖了 $[a,b]$，我们知道 $d_\infty(f_n,f)<\varepsilon$。 □

12 利用级数收敛来构造连续函数，距离空间的完备化

二零一九年十月二十四日，星期四，小雨

级数收敛的应用：连续函数的构造

我们现在通过 $\Big(C\big([a,b]\big),+,\cdot,\|\cdot\|_\infty\Big)$ 和 $(\mathbb{R},+,\cdot,|\cdot|)$ 的类比来构造连续函数。回忆一下，级数的构造一个很重要的应用就是用来构造实数，比如说，e 可以（定义）用级数 $\sum\limits_{k=0}^{\infty}\dfrac{1}{k!}$ 来实现。对于一般的数项级数，我们还有绝对收敛的概念，即对于数项级数 $\sum\limits_{k=1}^{\infty}a_k$ 而言，如果 $\sum\limits_{k=1}^{\infty}|a_k|$ 收敛，那么级数本身也收敛。我们对于闭区间的连续函数的空间也有类似的技术手段：

命题 74. *$a<b$ 是实数，给定函数列 $\{f_1\}_{n\geqslant 1}\subset C\big([a,b]\big)$，我们考虑函数级数 $\sum\limits_{k=1}^{\infty}f_k$：如果部分和 $S_n=\sum\limits_{k=1}^{n}f_k$ 所定义的函数序列 $\{S_n\}_{n\geqslant 1}$ 在 $\Big(C\big([a,b]\big),\|\cdot\|_\infty\Big)$ 中收敛，我们就称函数级数 $\sum\limits_{k=1}^{\infty}f_k$ 在 $\Big(C\big([a,b]\big),\|\cdot\|_\infty\Big)$ 中**收敛**。如果数项级数 $\sum\limits_{k=1}^{\infty}\|f_k\|_\infty$ 收敛，我们就称函数级数 $\sum\limits_{k=1}^{\infty}f_k$ 在 $\Big(C\big([a,b]\big),\|\cdot\|_\infty\Big)$ 中**绝对收敛**（即加了绝对值之后收敛）。*

如果函数级数 $\sum\limits_{k=1}^{\infty}f_k$ 在 $\Big(C\big([a,b]\big),\|\cdot\|_\infty\Big)$ 中绝对收敛，那么函数级数 $\sum\limits_{k=1}^{\infty}f_k$ 在 $\Big(C\big([a,b]\big),\|\cdot\|_\infty\Big)$ 中收敛。特别地，$\sum\limits_{k=1}^{\infty}f_k$ 给出了一个 $[a,b]$ 上定义的连续函数。

证明： 我们应该对比命题25的证明。根据 $C\big([a,b]\big)$ 的完备性，利用 Cauchy 判别法：由于 $\sum\limits_{k=0}^{\infty}\|f_k\|_\infty$ 收敛，所以对任意的 $\varepsilon>0$，存在 N，使得对任意的 $n\geqslant N$ 和任意的 $p\geqslant 0$，我们有 $\sum\limits_{n\leqslant k\leqslant n+p}\|f_k\|_\infty<\varepsilon$。为了判断部分和 S_n 是否是 Cauchy 序列，我们注意到对任意的 $x\in[a,b]$，我们就有

$$\begin{aligned}\big|S_{n+p}(x)-S_{n-1}(x)\big|&=\Big|\sum_{n\leqslant k\leqslant n+p}f_k(x)\Big|\leqslant\sum_{n\leqslant k\leqslant n+p}\big|\big|f_k(x)\big|\big|\\&\leqslant\sum_{n\leqslant k\leqslant n+p}\|f_k\|_\infty<\varepsilon.\end{aligned}$$

所以，$\{S_n\}_{n\geqslant 1}$ 在 $\Big(C\big([a,b]\big),\|\cdot\|_\infty\Big)$ 中是 Cauchy 列从而收敛。 □

作为应用，我们现在重新构造 $\mathbb{R}$ 上的指数函数。我们首先在任意的 $[-N,N]$ 这个闭区间上构造，因为函数 $f(x)=x$ 是 $[-N,N]$ 上的连续函数，我们就考虑

$$\widetilde{\exp}(x):=\sum_{k=0}^{\infty}\frac{f^k}{k!}.$$

由于 $\|\frac{f^k}{k!}\|_\infty \leqslant \frac{N^k}{k!}$，容易看出上面这个函数级数是绝对收敛的，从而我们在 $[-N,N]$ 上定义了 $\widetilde{\exp}(x)$。我们当然需要验证，如果 $N'>N$，那么在 $[-N',N']$ 上按照如上方式定义的 $\widetilde{\exp}(x)$ 在 $[-N,N]$ 上的限制给出的就是按照如上方式定义的 $\widetilde{\exp}(x)$，这就构造出了 $\widetilde{\exp}:\mathbb{R}\to\mathbb{R}$。

练习. 证明，如上定义的指数函数 $\widetilde{\exp}$ 恰好是之前所构造的 exp。

$\Big(C\big([a,b]\big),\|\cdot\|_\infty\Big)$ 的很多性质都可以平行的推广到紧的距离空间上去，并且这些性质在将来分析和几何的学习中会起到很重要的作用。我们把若干这样的性质整理在下面的命题里：

命题 75. *(X,d) 是紧的距离空间，那么*

1) (X,d) 是完备的。

2) 如果 $f:X\to\mathbb{R}$ 是连续函数，那么存在 $x_1\in X$，使得 $f(x_1)=\sup\limits_{x\in X}f(x)$；存在 $x_2\in X$，使得 $f(x_2)=\inf\limits_{x\in X}f(x)$。

3) 用 $C(X)$ 表示 X 上的连续函数的空间（实数值或者复数值）。对每个 $f\in C(X)$，$\|f\|_\infty=\sup\limits_{x\in X}|f(x)|$ 是 X 上的一个范数，那么，$\big(C(X),\|\cdot\|_\infty\big)$ 是完备的赋范线性空间。

我们在讲义中不给出这个命题的证明了。

至此，数学分析一这门课程的第一部分全部完成，这一部分的主旨是收敛和连续。作为这些理论的应用，我们再回过头来研究实数的构造。

距离空间的完备化（阅读）

我们先定义集合上的等价关系。X 是集合，$\sim$ 是 X 上的一个**等价关系**，这指的是对任意 $x,y\in X$，我们要么有 $x\sim y$（此时我们说 x 和 y 等价），要么有 $x\not\sim y$，二者必（且仅）居其一，并且 $\sim$ 满足：

1)（自反性）$x\sim x$；

2)（对称性）$x\sim y$ 等价于 $y\sim x$；

3)（传递性）如果 $x\sim y$，$y\sim z$，那么 $x\sim z$。

假定在集合 X 上给定了等价关系 $\sim$。对任意的 $x\in X$，X 的子集 $[x]=\big\{y\in X\big|y\sim x\big\}$ 被称作是 x（所在）的**等价类**。我们用 $X/\sim$ 表示**等价类的集合**，这是 X 的子集的集合。根据 $\sim$ 的上述三条性质，如果 $x'\in[x]$，那么 $[x']=[x]$；如果 $x'\notin[x]$，那么 $[x']\cap[x]=\emptyset$。据此，我们可以将 X 分成两两不交的子集的并集：

$$X=\bigsqcup_{[x]\in X/\sim}[x].$$

这称作是 X 的等价类的划分。一个等价关系 $\sim$ 实际上等价于将 X 分为若干个子集的无交并，直观上说我们将等价的元素看成是一样的。我们注意到对任意的 $x', x'' \in [x]$，我们都有 $[x'] = [x''] = [x]$，我们把 x, x' 或者 x'' 都称作是等价类 $[x]$ 的**代表元**。代表元的选取并不唯一。

作为例子，我们考虑整数的集合 $\mathbb{Z}$，p 是一个素数，对于 $m, n \in \mathbb{Z}$，如果 m 和 n 除 p 的余数相同，我们就说它们等价，并记做 $m \sim n$。我们知道，$\mathbb{Z}/\sim$ 恰好有 p 个元素。当 $p = 2$ 时，这就是把整数分为奇数和偶数两个等价类。

我们回到距离空间 (X, d) 的场景，它未必是完备。比如说有理数 $\mathbb{Q}$，但是我们知道存在 $\mathbb{R}$，使得 $\mathbb{Q} \subset \mathbb{R}$，而 $\mathbb{R}$ 是完备的。我们的目标是将 (X, d) 类比为 $\mathbb{Q}$，然后构造相应的 $(\overline{X}, d)$。

给定距离空间 (X, d)，我们令

$$\mathcal{C} = \left\{ \{x_n\}_{n\geqslant 1} \subset X \middle| \{x_n\}_{n\geqslant 1} \text{是 Cauchy 列} \right\}.$$

这是距离空间 (X, d) 中所有 Cauchy 的集合。按照定义，对 $\{x_n\}_{n\geqslant 1} \in \mathcal{C}$，对任意的 $\varepsilon > 0$，存在 N，使得当 $m, n \geqslant N$ 时，我们有 $d(x_m, x_n) < \varepsilon$。

我们在 $\mathcal{C}$ 上定义等价关系：

$$\text{对于 } \{x_n\}_{n\geqslant 1}, \{y_n\}_{n\geqslant 1} \in \mathcal{C},\ \{x_n\}_{n\geqslant 1} \sim \{y_n\}_{n\geqslant 1} \text{ 当且仅当 } \lim_{n\to\infty} d(x_n, y_n) = 0。$$

最终，我们令 $\overline{X} = \mathcal{C}/\sim$。

我们现在在 $\overline{X}$ 上定义一个距离函数 $\overline{d}$：

$$\overline{d} : \overline{X} \times \overline{X} \to \mathbb{R},\quad ([\{x_n\}_{n\geqslant 1}], [\{y_n\}_{n\geqslant 1}]) \mapsto \lim_{n\to\infty} d(x_n, y_n),$$

其中 $[\{x_n\}_{n\geqslant 1}], [\{y_n\}_{n\geqslant 1}] \in \overline{X}$。注意到，$\{x_n\}_{n\geqslant 1}$ 只是等价类 $[\{x_n\}_{n\geqslant 1}]$ 的一个代表，而代表的选取并不唯一，所以，为了说明距离函数

$$\overline{d}\big([\{x_n\}_{n\geqslant 1}], [\{y_n\}_{n\geqslant 1}]\big) = \lim_{n\to\infty} d(x_n, y_n)$$

是良好定义的，我们需要验证如下两个性质：

- 极限 $\lim\limits_{n\to\infty} d(x_n, y_n)$ 存在。

  只要证明 $\{d(x_n, y_n)\}_{n\geqslant 1}$ 是 Cauchy 列即可，这是因为

  $$\begin{aligned} |d(x_n, y_n) - d(x_m, y_m)| &= |d(x_n, y_n) - d(x_m, y_n) + d(x_m, y_n) - d(x_m, y_m)| \\ &\leqslant |d(x_n, y_n) - d(x_m, y_n)| + |d(x_m, y_n) - d(x_m, y_m)| \\ &\leqslant d(x_n, x_m) + d(y_n, y_m). \end{aligned}$$

  根据 $[\{x_n\}_{n\geqslant 1}$ 和 $\{y_n\}_{n\geqslant 1}$ 是 X 中的 Cauchy 列，我们知道 $\{d(x_n, y_n)\}_{n\geqslant 1}$ 是 $\mathbb{R}$ 中的 Cauchy 列。

- 如果我们选取其它的代表元 $\{x'_n\}_{n\geqslant 1}\in[\{x_n\}_{n\geqslant 1}]$，$\{y'_n\}_{n\geqslant 1}\in[\{y_n\}_{n\geqslant 1}]$，我们必须说明

$$\lim_{n\to\infty}d(x_n,y_n)=\lim_{n\to\infty}d(x'_n,y'_n).$$

首先，根据三角不等式，我们有

$$\big|d(x_n,y_n)-d(x'_n,y_n)\big|\leqslant d(x_n,x'_n).$$

令 $n\to\infty$，我们得到

$$\lim_{n\to\infty}d(x_n,y_n)=\lim_{n\to\infty}d(x'_n,y_n).$$

类似地，我们有

$$\lim_{n\to\infty}d(x'_n,y_n)=\lim_{n\to\infty}d(x'_n,y'_n).$$

综合这两个等式即可。

为了说明 $(\overline{X},\overline{d})$ 是度量空间，我们需要说明 $\overline{d}$ 满足距离定义中的三个条件：首先，

$$\overline{d}\big([\{x_n\}_{n\geqslant 1}],[\{y_n\}_{n\geqslant 1}]\big)\geqslant 0,\quad \overline{d}\big([\{x_n\}_{n\geqslant 1}],[\{y_n\}_{n\geqslant 1}]\big)=\overline{d}\big([\{y_n\}_{n\geqslant 1}],[\{x_n\}_{n\geqslant 1}]\big),$$

明显成立。为了证明三角不等式，任取 $\overline{X}$ 中三个点 $[\{x_n\}_{n\geqslant 1}]$，$[\{y_n\}_{n\geqslant 1}]$ 和 $[\{z_n\}_{n\geqslant 1}]$，我们有

$$d(x_n,y_n)+d(y_n,z_n)\geqslant d(x_n,z_n)$$

令 $n\to\infty$，我们就得到（因为在取极限的运算下 $\geqslant$ 被保持）

$$\overline{d}\big([\{x_n\}_{n\geqslant 1}],[\{y_n\}_{n\geqslant 1}]\big)+\overline{d}\big([\{y_n\}_{n\geqslant 1}],[\{z_n\}_{n\geqslant 1}]\big)\geqslant\overline{d}\big([\{x_n\}_{n\geqslant 1}],[\{z_n\}_{n\geqslant 1}]\big).$$

所以，三角不等式成立。最终，我们要证明

$$\overline{d}([\{x_n\}]_{n\geqslant 1},[\{y_n\}]_{n\geqslant 1})=0\Rightarrow[\{x_n\}]_{n\geqslant 1}=[\{y_n\}]_{n\geqslant 1}\Leftrightarrow\{x_n\}_{n\geqslant 1}\sim\{y_n\}_{n\geqslant 1}.$$

这因为按照定义 $\lim\limits_{n\to\infty}d(x_n,y_n)=0$，即 Cauchy 列是等价的。从这里，我们可以看到之所以取 Cauchy 列的等价类也是为了保证 $\overline{d}$ 是距离函数。

利用上面的构造，我们通过 (X,d) 以一种确定的方式得到了距离空间 $(\overline{X},\overline{d})$。我们现在说明，可以将 (X,d) 视作是 $(\overline{X},\overline{d})$ 的子空间。对于 $x\in X$，我们可以将它视作是 Cauchy 列 $\{x_n=x\}_{n\geqslant 1}$（的等价类），我们得到映射

$$\iota:X\to\overline{X},\quad x\mapsto\{x_n=x\}_{n\geqslant 1},$$

当然，如果 $y\in X$，$y\neq x$，那么作为 Cauchy 列，$\{x_n=x\}_{n\geqslant 1}$ 和 $\{y_n=y\}_{n\geqslant 1}$ 不等价，这表明 ι 是单射。据此，我们可以认为 $X\subset\overline{X}$ 是一个子集，也可以认为是 $\overline{X}$ 是在 X 的基础上添加了一些东西。

另外，对于 $\{x_n=x\}_{n\geqslant 1}$ 和 $\{y_n=y\}_{n\geqslant 1}$，我们有

$$\overline{d}\big(\{x_n=x\}_{n\geqslant 1},\{y_n=y\}_{n\geqslant 1}\big)=\lim_{n\to\infty}d(x,y)=d(x,y)\quad\Leftrightarrow\quad\overline{d}\big(\iota(x),\iota(y)\big)=d(x,y),$$

这说明 $\overline{d}$ 在 X 上的诱导距离也恰好是 d。在这个意义下，我们将 (X,d) 看做是 $(\overline{X},\overline{d})$ 的子（距离）空间。

类似于 $\mathbb{Q}\subset\mathbb{R}$ 是稠密的，我们有

引理 76. X 是 $(\overline{X},\overline{d})$ 的稠密子集。

证明: 任意给定 $\{x_n = x\}_{n\geqslant 1} \in \mathcal{C}$，存在 N，使得当 $x \geqslant N$ 时，$|x_n - x_N| < \varepsilon$。所以，对于 $[\{x_n\}_{n\geqslant 1}] \in \overline{X}$，我们选取 $x_N \in X$，那么，$\overline{d}(\iota(x_N), \{x_n\}_{n\geqslant 1}) \leqslant \varepsilon$。这表明 X 是稠密的。 □

最终，我们证明 $(\overline{X},\overline{d})$ 是完备的距离空间：我们用 $\overline{x} \in \overline{X}$ 表示某个 Cauchy 列的等价类，即 $\overline{x} = [\{x_n\}_{n\geqslant 1}]$，目标是证明如果 $\{\overline{x}^{(a)}\}_{a\geqslant 1}$ 是 $(\overline{X},\overline{d})$ 中的 Cauchy 序列，那么 $\{\overline{x}^{(a)}\}_{a\geqslant 1}$ 在距离函数 $\overline{d}$ 下收敛。换而言之，由 (X,d) 中的 Cauchy 列的等价类构成的 Cauchy 列是收敛的!

我们直接构造这个极限点。根据 X 在 $\overline{X}$ 中的稠密性，对每个 $\overline{x}^{(a)}$（其中 $a \geqslant 1$），存在 $x_a \in X$，使得

$$d(\overline{x}^{(a)}, \iota(x_a)) = d(\overline{x}^{(a)}, [\{x_a, x_a, \cdots, x_a, \cdots\}]) < \frac{1}{a}.$$

令 $\overline{x} = [\{x_a\}_{a\geqslant 1}]$，我们验证如下两个性质

1) $\{x_a\}_{a\geqslant 1}$ 是 (X,d) 中的 Cauchy 列，从而 $[\{x_a\}_{a\geqslant 1}] \in \overline{X}$。

   这是因为

$$\begin{aligned} d(x_a, x_b) = \overline{d}\big(\iota(x_a), \iota(x_b)\big) &\leqslant \overline{d}\big(\iota(x_a), x^{(a)}\big) + \overline{d}\big(x^{(a)}, x^{(b)}\big) + \overline{d}\big(\iota(x_b), x^{(b)}\big) \\ &\leqslant \frac{1}{a} + \frac{1}{b} + \overline{d}\big(x^{(a)}, x^{(b)}\big). \end{aligned}$$

   由于 $\{\overline{x}^{(a)}\}_{a\geqslant 1}$ 是 $(\overline{X},\overline{d})$ 中的 Cauchy 序列，据此可见 $\{x_a\}_{a\geqslant 1}$ 是 (X,d) 中的 Cauchy 列。

2) 令 $\overline{x} = [\{x_a\}_{a\geqslant 1}] \in \overline{X}$，那么在 $(\overline{X},\overline{d})$ 中，我们有 $\lim\limits_{a\to\infty} \overline{x}^{(a)} = \overline{x}$。

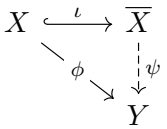
   按照定义，我们有

$$\overline{d}\big(\overline{x}^{(a)}, \overline{x}\big) = \lim_{n\to\infty} d(\overline{x}_n^{(a)}, x_n) = \lim_{n\to\infty} d(\iota\big(\overline{x}_n^{(a)}\big), \iota(x_n))$$

定理 77 (完备化的万有性质)**.** (X,d) 是距离空间。那么，存在完备的距离空间 $(\overline{X},\overline{d})$ 和等距嵌入 $\iota : (X,d) \to (\overline{X},\overline{d})$（即 $\overline{d}(\iota(x),\iota(y)) = d(x,y)$），使得

1) ι 的像 $\iota(X) \subset \overline{X}$ 是稠密的;

2) 对任意完备的距离空间 (Y,d_Y) 以及一致连续的映射 $\phi : X \to Y$，



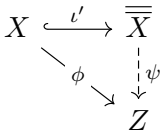
存在唯一的 $\psi : \overline{X} \to Y$（连续映射的扩张），使得 $\phi = \psi \circ \iota$。

另外，上述 $(\overline{X},\overline{d})$ 在如下的意义下是唯一的:

如果存在另外一个完备的距离空间 $(\overline{\overline{X}},\overline{\overline{d}})$ 和等距嵌入 $\iota' : (X,d) \to (\overline{\overline{X}},\overline{\overline{d}})$，使得

1) ι' 的像 $\iota'(X)$ 在 $\overline{\overline{X}}$ 中稠密；

2) 对任意完备的距离空间 (Z, d_Z) 以及一致连续的映射 $\phi : X \to Z$，

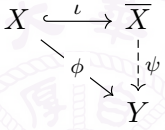


存在唯一的 $\psi : \overline{X} \to Z$（连续映射的扩张），使得 $\phi = \psi \circ \iota'$。

那么存在等距的双射 $\Phi : \overline{X} \to \overline{\overline{X}}$（即对任意的 $\overline{x}_1, \overline{x}_1 \in \overline{X}$，$\overline{\overline{d}}(\Phi(\overline{x}_1), \Phi(\overline{x}_2)) = \overline{d}(\overline{x}_1, \overline{x}_2)$）。

注记. 每个等距嵌入 $\iota : (X, d) \to (\overline{X}, \overline{d})$ 都是单射：因为如果 $\iota(x) = \iota(y)$，那么 $\overline{d}(\iota(x), \iota(y)) = d(x, y) = 0$，从而 $x = y$。

证明: 我们已经构造了 $(\overline{X}, \overline{d})$，所以只需要验证第二条：对任意完备的距离空间 (Y, d_Y) 以及连续映射 $\phi : X \to Y$，

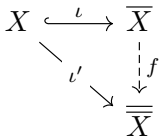


存在唯一的 $\psi : \overline{X} \to Y$（连续映射的扩张），使得 $\phi = \psi \circ \iota$。实际上，对任意的 $\overline{x} = [\{x_n\}_{n \geqslant 1}] \in \overline{X}$，其中 $\{x_n\}_{n \geqslant 1}$ 是 X 中的 Cauchy 序列，我们定义

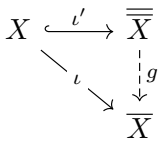
$$\psi(\overline{x}) = \lim_{n \to \infty} \phi(x_n).$$

在这个定义中，我们用到了两个基本的事实：一致连续的连续映射把 Cauchy 列映为 Cauchy 列；$\overline{X}$ 是完备的。另外，如果我们选取 $\overline{x}$ 的另一个可能的代表元 $\{x'_n\}_{n \geqslant 1}$，上述的定义 $\psi(\overline{x}) = \lim\limits_{n \to \infty} \phi(x'_n)$ 是不依赖于代表元素选取的，这个只是例行公事的检验（利用一致连续性）。

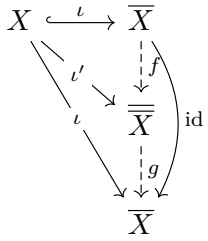
现在来证明唯一性，假设存在第二个 $(\overline{\overline{X}}, \overline{\overline{d}})$，由于等距嵌入 ι' 是一致连续的，所以利用关于 $(\overline{X}, \overline{d})$ 的第二条性质，我们有



其中，$f \circ \iota = \iota'$；类似地，利用关于 $(\overline{\overline{X}}, \overline{\overline{d}})$ 的第二条性质，我们还有



其中，$g \circ \iota' = \iota$。通过映射的符合，我们得到



此时，我们发现 $g \circ f$ 和单位映射 id 都可以被视作是 ι（往右下走的箭头）的扩张，根据唯一性，我们有 $g \circ f = \mathrm{id}$，类似地 $f \circ g = \mathrm{id}$。这就证明了 $f : \overline{X} \to \overline{\overline{X}}$ 是双射。由于 f 在 X 上的限制是等距映射，所以这也是一个等距映射。 □

我们习惯上把上述构造称作是 $(\overline{X}, \overline{d})$ 称作是(X, d) **的完备化**，并把它记做 $(\widehat{X}, \widehat{d})$，它在等距的意义下是唯一的。

如果对 $(X, d) = (\mathbb{Q}, d(x, y) = |x - y|)$ 用上述的结论，由于 $\mathbb{R}$ 中每个数都可以用有理数的 Cauchy 序列逼近，所以 $(\mathbb{R}, d(x, y) = |x - y|)$，是 $\mathbb{Q}$ 的完备化。在这个意义下，实数是唯一的（我们必须指出，实数的存在性是我们利用 Dedekind 分割事先已经构造好的对象）。

如果我们在 $\mathbb{Q}$ 上选取别的距离函数，比如说，给定素数 p，对于任意的有理数 x，我们可以将它唯一的写成 $x = p^v \dfrac{m}{n}$ 的形式，其中 m 和 n 是和 p 互素的整数，$v(x) \in \mathbb{Z}$。对于任意两个有理数 x 和 y，我们定义

$$d_p(x, y) = \left(\frac{1}{p}\right)^{v(x-y)}.$$

这是所谓的 p-adic 距离，有理数 $\mathbb{Q}$ 配有这个距离函数并不完备，它的完备化是所谓的 p-adic 数 $\mathbb{Q}_p$。这个数域和 $\mathbb{R}$ 类似，但是又有很多不同的几何性质，它在数论的研究中不可或缺。

12.1 作业：有无穷多素数的拓扑证明

清华大学 19-20 秋季学期，数学分析一，作业 5

请用 A4 大小的纸张**正反面**用钢笔，签字笔或者圆珠笔书写，并注明自己的姓名，年级和作业的**总页数**。除定理公式所涉及的人名之外，请使用**中文**。本次作业的提交时间和地点为 **10 月 31 日**上午的课堂上，逾期视作零分。

基本习题

习题 A：连续，一致连续和一致收敛

A1) 证明，函数 e^x 在 $\mathbb{R}$ 上不是一致连续的而在 $(-\infty, 0]$ 上是一致连续的。

A2) 证明，幂函数映射 $\mathbb{R}_{>0} \times \mathbb{R}$，$(x, \alpha) \mapsto x^\alpha$ 是 $\mathbb{R}_{>0} \times \mathbb{R}$ 上的连续函数。

A3) 证明，第九次课中定义的幂函数满足如下的性质：对任意 $x, y > 0$ 和 α, β，我们有 $(xy)^\alpha = x^\alpha y^\alpha$，$(x^\alpha)^\beta = x^{\alpha\beta}$；$a^{\log_a x} = x$；如果 $x > 0$，$y > 0$，那么 $a^{x+y} = a^x a^y$，$\log_a(x \cdot y) = \log_a(x) + \log_a(y)$。

A4) 在 $[0, 1)$ 区间上考虑连续函数的序列 $\{f_n(x)\}_{n\geqslant 1}$，其中 $f_n(x) = x^n$。证明，对任意的 $a < 1$，$\{f_n(x)\}_{n\geqslant 1}$ 在 $[0, a]$ 上一致收敛到 0 这个函数；但是 $\{f_n(x)\}_{n\geqslant 1}$ 在 $[0, 1)$ 上不一致收敛。

A5) 在 $\mathbb{R}$ 上考虑连续函数的序列 $\{f_n(x)\}_{n\geqslant 1}$，其中 $f_n(x) = \dfrac{nx}{1 + n^2x^2}$。证明，$\{f_n(x)\}_{n\geqslant 1}$ 在 $\mathbb{R}$ 上逐点收敛到 0 这个函数但是 $\{f_n(x)\}_{n\geqslant 1}$ 在 $\mathbb{R}$ 上不一致收敛。

A6) 在 $\mathbb{R}$ 上考虑连续函数的序列 $\{f_n(x)\}_{n\geqslant 1}$，其中 $f_n(x) = \begin{cases} \dfrac{nx^2}{1+nx}, & x > 0; \\ \dfrac{nx^3}{1+nx^2}, & x \leqslant 0 \end{cases}$. 试研究 $\{f_n(x)\}_{n\geqslant 1}$ 在 $\mathbb{R}$ 上逐点收敛性和一致收敛性。

A7) 给定连续函数 $\varphi : \mathbb{R}_{\geqslant 0} \to \mathbb{R}$，满足 $\varphi(0) = 0$，$\lim\limits_{x\to\infty} \varphi(x) = 0$ 并且 φ 不恒为零。证明，$\mathbb{R}_{\geqslant 0}$ 上的连续函数的序列 $\{f_n(x)\}_{n\geqslant 1}$ 和 $\{f_n(x)\}_{n\geqslant 1}$ 逐点收敛到 0 这个函数但是不一致收敛，其中 $f_n(x) = \varphi(nx)$，$g_n(x) = \varphi\left(\dfrac{x}{n}\right)$。

A8)* （一致连续性的应用：积分的定义）$[a, b]$ 是有限闭区间，$f \in C\big([a, b]\big)$ 是实数值的函数。给定 $n \geqslant 1$，我们将 $[a, b]$ 均分为 n 份：

$$[a, b] = [a_1, b_1] \cup [a_2, b_2] \cup \cdots \cup [a_n, b_n], \quad a_1 = a, b_k = a_{k+1}\ (k = 1, \cdots, n-1), b_n = b,$$

其中对于 $k = 1, 2, \cdots, n$，$a_k = a + \dfrac{k-1}{n}(b - a)$。我们定义

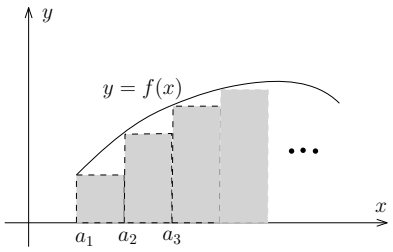
$$S_n = \sum_{k=1}^{n} \frac{b-a}{n} f(a_k).$$

证明，$\{S_n\}_{n\geqslant 1}$ 收敛，我们用 $\int_a^b f$ 这个符号来记极限 $\lim\limits_{n\to\infty} S_n$。进一步证明，映射

$$\int_a^b : C\big([a,b]\big) \to \mathbb{R}, \quad f \mapsto \int_a^b f$$

是线性映射并且如果我们在 $C\big([a,b]\big)$ 用距离函数 d_∞ 来考虑，那么这是连续映射。

注记. 和式 S_n 直观上计算的是下图中阴影柱形表的面积：



所以，如果 f 是正函数，我们认为 $\int_a^b f$ 表示的是 f 的图像下的面积。

另外，就算是选取 $f(x)=x^m$ 这样简单的函数，$\int_a^b f$ 的计算都很复杂，课程后面关于积分的理论的一个重点是如何计算积分。

（提示：为了证明 $\{S_n\}_{n\geqslant 1}$ 是 Cauchy 列，我们可以将 S_n 和 S_m 都与 S_{nm} 做比较）

A9)* 证明 Cauchy 的一个定理: 任给函数 $f:[a,+\infty)\to\mathbb{R}$，我们假设 f 任意闭子区间 $[a,b]$ 上有界 (上界可能依赖于 b)，那么下面两个式自当等号右边极限存在时成立：

$$\lim_{x\to+\infty}\frac{f(x)}{x}=\lim_{x\to+\infty}[f(x+1)-f(x)],$$

$$\lim_{x\to+\infty}[f(x)]^{1/x}=\lim_{x\to+\infty}\frac{f(x+1)}{f(x)},\ \text{进一步假设对任意的 } x\in[a,\infty),\ f(x)\geqslant c>0.$$

习题 B 研究下列函数 f 在区间 I 上的一致连续性：

B1) $f(x)=\sqrt[3]{x}$，$I=(0,\infty)$。

B2) $f(x)=\log x$，$I=(0,1)$。

B3) $f(x)=\cos\dfrac{1}{x}$，$I=(0,1)$。

B4) $f(x)=x\cos\dfrac{1}{x}$，$I=(0,\infty)$。

习题 C 研究下列函数极限的存在性：

C1) $\lim\limits_{x\to 1}\dfrac{\ln x}{(x-1)^{\alpha}}$，$\alpha>0$。

C2) $\lim\limits_{x\to 1}\dfrac{e^{x}-e}{(x-1)^{\alpha}}$，$\alpha>0$。

C3) $\lim\limits_{x\to 1}\dfrac{x^{x}-1}{(x-1)^{\alpha}}$，$\alpha>0$。

C4) $\lim\limits_{x\to 1}\dfrac{\sqrt[3]{1-\sqrt{x}}}{(x-1)^{\alpha}}$，$\alpha>0$。

C5) $\lim\limits_{x\to 0}\dfrac{\sqrt{1+x^{2}}-1}{1-\cos x}$。

C6) $\lim\limits_{x\to 0}\dfrac{\sqrt{1+x^{4}}-1}{1-\cos^{2} x}$。

C7) $\lim\limits_{x\to 1}\dfrac{(x-1)^{\alpha}}{\sin(\pi x)}$, $\alpha>0$。

习题 D（和一致连续性相关的问题）

D1) 如果连续函数 f 在开区间（有限或无限）$I\subset\mathbb{R}$ 上是单调并且有界的，那么 f 在 I 上一致连续。

D2) I 是长度有限的区间（不一定是闭的）。证明，I 上的实值函数 f 一致连续的*充分必要条件*是 f 把 Cauchy 列映成 Cauchy 列（即如果 $\{x_n\}_{n\geqslant 1}\subset I$ 是 Cauchy 列，那么 $\{f(x_n)\}_{n\geqslant 1}$ 也是 Cauchy 列）。

D3) f 在 $\mathbb{R}$ 上一致连续。证明，存在 $a,b\in\mathbb{R}_{>0}$，使得对任意的 $x\in\mathbb{R}$，我们都有

$$|f(x)|\leqslant a|x|+b.$$

D4) 假设函数 f 在 $[0,\infty)$ 上一致连续并且对任意的 $x\in[0,1]$，我们都有 $\lim\limits_{n\to+\infty} f(x+n)=0$（这里 n 是整数）。证明，

$$\lim_{x\to+\infty} f(x)=0.$$

如果我们将条件减弱为 f 在 $[0,\infty)$ 连续，结论是否依然成立? 证明或举出反例。

D5) 假设 X 是区间，$f:X\to\mathbb{R}$ 是连续函数。如果存在正常数 L，使得对任意的 $x,y\in X$，都有

$$|f(x)-f(y)|\leqslant L|x-y|,$$

我们就称 f 在 X 上满足 Lipschitz 条件。

(a) 证明，f 在 X 上满足 Lipschitz 条件是 f 在 X 一致连续的充分条件。

(b) 判断上述条件是否是必要条件？证明或举出反例。

(c) 假设 f 在 $[a,+\infty)$ 上满足 Lipschitz 条件，其中 $a>0$，试证明 $\dfrac{f(x)}{x}$ 在 $[a,+\infty)$ 一致连续。

习题 E 利用连续性研究方程根的存在性（在以下方程中，x 为未知实数）：

E1) 证明，方程 $x^3+2x-1=0$ 在 $\mathbb{R}$ 上有唯一的根且此根在 $(0,1)$ 内。

E2) 设 $0\leqslant\lambda<1$，$b>0$，试判断方程 $x-\lambda\sin x=b$ 的实根的存在性。

E3) 证明，方程 $\sin x=\dfrac{1}{x}$ 有无穷多个实根。

E4) 假设实数值函数 $f\in C([0,2])$ 并且 $f(0)=f(2)$。证明，方程 $f(x)-f(x+1)=0$ 在 $[0,1]$ 上有一个根。

习题 F 试计算下面函数的极限：

$$F1)\ \lim_{x\to\pi}\frac{\sin mx}{\sin nx},\quad m,n\in\mathbb{Z}_{\geqslant 0},\qquad F2)\ \lim_{x\to 0}\frac{1-\cos x\sqrt[2]{\cos 2x}\sqrt[3]{\cos 3x}}{x^2}$$

$$F3)\ \lim_{x\to+\infty}\sin\sqrt{1+x}-\sin\sqrt{x}\qquad F4)\ \lim_{x\to 0}\frac{\sqrt{1+x\sin x}-1}{e^{x^2}-1}$$

$$F5)\ \lim_{n\to\infty}\underbrace{\sin\sin\cdots\sin}_{n个}x$$

思考题 (不交作业)

问题 G 连续函数 $f:\mathbb{R}\to\mathbb{R}$ 满足如下性质：对任意的 $\delta>0$，我们都有

$$\lim_{n\to\infty}f(n\delta)=0.$$

证明，$\lim\limits_{x\to+\infty}f(x)=0$。

问题 H 连续函数 $\varphi:\mathbb{R}\to\mathbb{R}$ 满足如下两个性质：

1) 它在无穷远的行为如下：$\lim\limits_{x\to+\infty}\big(\varphi(x)-x\big)=+\infty$。

2) φ 的不动点集 $\big\{x\in\mathbb{R}\big|\varphi(x)=x\big\}$ 是非空的有限集。

证明，如果 $f:\mathbb{R}\to\mathbb{R}$ 是连续函数并且满足 $f\circ\varphi=f$，那么 f 一定是常值函数。

问题 I 连续函数 $f:\mathbb{R}_{\geqslant 0}\to\mathbb{R}$ 满足 $\lim\limits_{x\to+\infty}\dfrac{f(x)}{x}=0$。假设 $\{a_n\}_{n\geqslant 1}$ 是一列非负实数的数列并且数列 $\{\dfrac{a_n}{n}\}_{n\geqslant 1}$ 是有界的，证明，$\lim\limits_{n\to\infty}\dfrac{f(a_n)}{n}=0$

课外补充：拓扑空间的定义

本节也不交作业。

定义 78 (拓扑空间)**.** *X 是集合，$\mathcal{T}=\{U|U\subset X\}$ 是 X 的某些子集所组成的集合。如果下面三个条件成立*

1) $\emptyset\in\mathcal{T}$，$X\in\mathcal{T}$。

2) 对任意的 $\{U_\alpha\}_{\alpha\in\mathcal{A}}\subset\mathcal{T}$，*其中* $\mathcal{A}$ *为指标集合，我们有* $\bigcup\limits_{\alpha\in\mathcal{A}}U_\alpha\in\mathcal{T}$。

3) 对任意有限个 $U_1,U_2,\cdots,U_m\in\mathcal{T}$，*我们有* $\bigcap\limits_{1\leqslant i\leqslant m}U_i\in\mathcal{T}$。

*我们就称 $\mathcal{T}$ 是 X 上的一个***拓扑***，每个 $U\in\mathcal{T}$ 都被称作是（拓扑 $\mathcal{T}$ 下的）***开集***。我们把二元组 $(X,\mathcal{T})$ 称作是一个***拓扑空间***。*

给定拓扑空间 $(X,\mathcal{T})$，考虑子集 $F\subset X$，如果其补集 $X-F$ 是开集（即 $X-F\in\mathcal{T}$），我们就称 F 是在拓扑 $\mathcal{T}$ 下的）**闭集**。对于闭集，我们有如下的性质（简单的练习）：

1) $\emptyset$ 和 X 都是闭集。

2) 任意多闭集的交集是闭集。

3) 有限个闭集的并集是闭集。

娱乐问题 素数是无限多的拓扑证明（Furstenberg，1955）

考虑 $X=\mathbb{Z}$ 为全体整数的集合，对于 $a,b\in\mathbb{Z}$，其中 $a\geqslant 1$，我们定义 $U_{a,b}=\{ka+b|k\in\mathbb{Z}\}$（就是以 b 为某一项的双边的等差数列）。我们定义 X 的子集的集合 $\mathcal{T}$：如果 $U\in\mathcal{T}$，要么 $U=\emptyset$，要么 U 是某些 $U_{a,b}$ 的并集。

1) $U\in\mathcal{T}$。证明，$x\in U$ 当且仅当存在 $a\geqslant 1$，使得 $U_{a,x}\in\mathcal{T}$。

2) 证明，$\mathbb{Z}\in\mathcal{T}$。

3) 证明，对于任意的 $\{U_i\}_{i\in I}\subset\mathcal{T}$，那么 $\bigcup U_i\in\mathcal{T}$。

4) 证明，如果 $U,V\in\mathcal{T}$，那么 $U\cap V\in\mathcal{T}$。

5) 证明，任何有限集合都不是开集；任何补集是有限的集合都不是闭集。

6) 证明，$U_{a,b}$ 即时开集也是闭集。

7) 证明，$\mathbb{Z}-\{-1,1\}=\bigcup\limits_{p\text{ 是素数}}U_{p,0}$。

8) 用反证法证明，有无限多个素数。

As for everything else, so for a mathematical theory: beauty can be perceived but not explained.

– A. Cayley

12.2 期中考试：连续函数环的极大理想

二零一九年十月二十八日，星期一，晴天

数学分析一 期中考试

考试说明

请在清华大学考试专用纸上**正反面**用钢笔，签字笔或者圆珠笔书写，并注明自己的姓名和年级。除定理公式所涉及的人名之外，请使用**中文**。

考试时间为 上午 9:50 至 中午 12:15。考试结束请交答题纸，草稿纸和试题请带走。两道大题之间是相互独立的。考试中后面的问题可以使用前面问题的结论（无论答题人是否已经得到正确的证明或者答案）。

题目 A 基本概念和技巧的考察（共 8 小题）

A1) 试用 $\varepsilon - N$ 语言证明：$\lim\limits_{n\to\infty} \dfrac{n}{n^2+1} = 0$。

A2) 试用 $\varepsilon - \delta$ 语言说明函数 $\mathbb{R}$ 上的函数 $f(x) = \begin{cases} \sin(\frac{1}{x}), & x \neq 0; \\ 0, & x = 0 \end{cases}$ 在 $x = 0$ 处不连续。

A3) 计算极限：$\lim\limits_{n\to\infty} \left(1^2 + 2^2 + 3^2 + \cdots + n^2\right)^{\frac{1}{n}}$。

A4) 证明，级数 $\sum\limits_{n=1}^{\infty} \dfrac{2n+1}{n^2(n+100)^2}$ 收敛。

A5) 给定实数的序列 $\{a_n\}_{n\geqslant 1}$，假设级数 $\sum\limits_{n=1}^{\infty} a_n$ 收敛。证明，对任意的实数 $x \in (-1, 1)$，级数 $\sum\limits_{n=1}^{\infty} a_n x^n$ 也收敛。（提示：利用 Abel 判别法）

A6) 假设 f 是非空开区间 (a, b) 上的连续函数，证明，对任意的 $x_1, x_2 \in (a, b)$，存在 $x_0 \in (a, b)$，使得

$$f(x_0) = \frac{1}{2}\big(f(x_1) + f(x_2)\big).$$

A7) 假设函数 $f : \mathbb{R} \to \mathbb{R}$ 是以 1 为周期的连续函数，即对任意的 $x \in \mathbb{R}$，我们有 $f(x+1) = f(x)$。证明，f 有界并且能取到其最大值，即存在 $x_0 \in \mathbb{R}$，使得 $f(x_0) = \sup\limits_{x\in\mathbb{R}} f(x)$。

A8) 证明，$f(x) = \sqrt{x+1}$ 作为 $\mathbb{R}_{\geqslant 0}$ 上的函数是一致连续的。

题目 B

首先回忆一下二项式展开：$(1+z)^n = \sum\limits_{k=0}^{n} \binom{n}{k} z^k$。特别地，当 $z = 1$ 时，我们有 $2^n = \sum\limits_{k=0}^{n} \binom{n}{k}$。

给定一个复数的数列 $\{a_k\}_{k\geqslant 0}$，我们定义新的数列 $\{a_n^*\}_{n\geqslant 0}$，其中

$$a_n^* = \frac{1}{2^n}\sum_{k=0}^{n}\binom{n}{k}a_k.$$

第一部分：等比数列

这一部分中，我们假设对任意的 $k\geqslant 0$，$a_k = z^k$，其中 z 为复数。

B1) 证明，如果 $|z|<1$，那么级数 $\displaystyle\sum_{k=0}^{\infty}a_k$ 收敛。我们用 $A(z)$ 表示这个极限。

B2) 证明，如果 $|z|<1$，那么级数 $\displaystyle\sum_{k=0}^{\infty}a_k^*$ 收敛。我们用 $A^*(z)$ 表示这个极限。

B3) 证明，如果 $|z|\geqslant 1$，那么级数 $\displaystyle\sum_{k=0}^{\infty}a_k$ 不收敛。

B4) 试找出一个 $z\in\mathbb{C}$，$|z|>1$，级数 $\displaystyle\sum_{k=0}^{\infty}a_k^*$ 收敛。

B5) 证明，如果 $|z|=1$ 并且 $z\neq\pm 1$，那么级数 $\displaystyle\sum_{k=0}^{\infty}a_k^*$ 收敛。

第二部分：$\{a_k\}_{k\geqslant 0}$ 与 $\{a_k^*\}_{k\geqslant 0}$ 的收敛性比较，$\displaystyle\sum_{k=0}^{\infty}a_k$ 与 $\displaystyle\sum_{k=0}^{\infty}a_k^*$ 的收敛性比较

这一部分中，我们假设 $\{a_k\}_{k\geqslant 0}$ 为实数的序列。

B6) 证明，当 $k\in\mathbb{Z}_{\geqslant 0}$ 固定的时候，我们有

$$\lim_{n\to\infty}\binom{n}{k}\Big/\frac{n^k}{k!}=1,\quad \lim_{n\to\infty}\binom{n}{k}\Big/2^n=0.$$

B7) 任意给定非负整数 $n>q$，我们定义

$$a_{n,q}^* = \frac{1}{2^n}\sum_{k=0}^{q}\binom{n}{k}a_k.$$

对每个固定的 q，计算 $\displaystyle\lim_{n\to\infty}a_{n,q}^*$。

B8) 如果 $\displaystyle\lim_{n\to\infty}a_n=0$，证明，$\displaystyle\lim_{n\to\infty}a_n^*=0$。

B9) 如果 $\displaystyle\lim_{n\to\infty}a_n$ 存在，证明，$\displaystyle\lim_{n\to\infty}a_n^*$ 存在并且恰好等于 $\displaystyle\lim_{n\to\infty}a_n$。

B10) 如果 $\displaystyle\lim_{n\to\infty}a_n^*$ 存在，$\displaystyle\lim_{n\to\infty}a_n$ 是否一定存在?

在后面的问题中，对任意的 $n \geqslant 0$，我们定义部分和

$$S_n = \sum_{k=0}^{n} a_k, \quad S_n^* = \sum_{k=0}^{n} a_k^*, \quad U_n = 2^n S_n^*.$$

B11) 证明，对任意的 $n \geqslant 0$，U_n 都可以写成 $S_0, S_1, \cdots, S_n$ 的整系数线性组合：

$$U_n = \sum_{k=0}^{n} \binom{n+1}{k+1} S_k.$$

B12) 证明，如果级数 $\sum\limits_{k=0}^{\infty} a_k$ 收敛，那么 $\sum\limits_{k=0}^{\infty} a_k^*$ 也收敛。

题目 C：$C\big([a,b]\big)$ 的极大理想

假设 $a < b$ 是实数，我们研究有界闭区间上的实数值连续函数的空间 $C\big([a,b]\big)$。对于子集 $I \subset C\big([a,b]\big)$，如果它满足如下三个条件：

1) $I \neq \emptyset$，$I \neq C\big([a,b]\big)$；

2) 对任意的 $\varphi \in I$，$\psi \in I$，我们有 $\varphi + \psi \in I$；

3) 对任意的 $\varphi \in I$，$f \in C\big([a,b]\big)$，我们有 $\varphi \cdot f \in I$.

我们就称 I 是 $C\big([a,b]\big)$ 的一个**理想**。假设 $\mathfrak{m}$ 是 $C\big([a,b]\big)$ 的理想并且不存在其它的包含 $\mathfrak{m}$ 的理想，我们就称 $\mathfrak{m}$ 是 $C\big([a,b]\big)$ 的一个**极大理想**（即若理想 $I \supset \mathfrak{m}$，那么 $I = \mathfrak{m}$）。

C1) 对任意的子集 $A \subset [a,b]$，令 $I(A) = \big\{f \in C\big([a,b]\big) \big| \text{对任意的 } x \in A,\ f(x) = 0\big\}$。证明，$I(A)$ 是 $C\big([a,b]\big)$ 的理想。$I([a,b])$ 是什么？证明，如果有两个子集 $A \subset B \subset [a,b]$，那么 $I(A) \supset I(B)$。是否存在 $A \subset [a,b]$ 为真子集，使得 $I(A) = \{0\}$？

C2) 证明，如果 $I \subset C\big([a,b]\big)$ 是理想，那么常值函数 $1 \notin I$。进一步证明，如果 $I \subset C\big([a,b]\big)$ 是理想，那么对任意的 $f \in I$，f 在 $[a,b]$ 上一定有零点（即 $f(x) = 0$ 在 $[a,b]$ 上有解）。

C3) 对于 $f \in C\big([a,b]\big)$，证明，集合 $V(f) = \big\{x \in [a,b] \big| f(x) = 0\big\}$ 是闭集。进一步证明，对于理想 $I \subset C\big([a,b]\big)$，集合 $V(I) = \big\{x \in [a,b] \big| \text{对任意的 } f \in I,\ f(x) = 0\big\}$ 是闭集。如果理想 $I \subset C\big([a,b]\big)$ 使得 $V(I)$ 为全空间 $[a,b]$，你是否能够确定 I？

C4) 对任意的点 $x \in [a,b]$，我们令 $A = \{x\}$ 并记 $\mathfrak{m}_x = I(A) = I\big(\{x\}\big)$，即

$$\mathfrak{m}_x = \big\{f \in C\big([a,b]\big) \big| f(x) = 0\big\}.$$

证明，$\mathfrak{m}_x$ 是极大理想。

C5) 证明，如果 $\mathfrak{m}$ 是 $C\big([a,b]\big)$ 的极大理想，那么存在 $x \in [a,b]$ 使得 $\mathfrak{m} = \mathfrak{m}_x$。（提示：利用 $[a,b]$ 是紧的）

C6)** 假设 A 是闭集，证明，$V(I(A)) = A$。

注记. 上面的结论可以推广到紧的距离空间的情形：假设 (X, d) 是紧的距离空间，$C(X)$ 是 X 上复数值的连续函数的全体，那么我们有如下的一一对应：

$$X \longrightarrow \big\{C(X) \text{ 的极大理想}\big\},\ \ x \mapsto \mathfrak{m}_x,$$

其中 $\mathfrak{m}_x = \big\{f \in C\big(X\big)\big|f(x) = 0\big\}$。据此，我们把可以把环 $C(X)$ 中的极大理想想象成空间的点，从而通过 $C(X)$ 中的代数对象来研究 X 上的几何，这是代数几何的开端。

13 导数的定义，初等函数导数的计算

二零一九年十月三十一日，星期四，小雨

函数和映射的微分学

定义 79. 假设 f 是定义在区间 $I \subset \mathbb{R}$ 上的实值函数，$x_0 \in I$ 是一个给定的点。如果极限

$$\lim_{h\to 0} \frac{f(x_0+h)-f(x_0)}{h}$$

存在，我们就称 f 在 x_0 处**可导**或者**可微**（可微分）并用用 $f'(x_0)$ 记这个极限。如果 f 在任意 $x \in I$ 处都可微，我就说 f 是 I 上的**可微函数**。进一步，如果 $f'(x) \in C(I)$，我们就说 f 是**连续可微的**并且将所有 I 上连续可微的函数记作 $C^1(I)$。

当导数存在时，我们也用下面的极限来写导数：

$$f'(x_0) = \lim_{x\to 0} \frac{f(x)-f(x_0)}{x-x_0}.$$

注记. 我们注意到，如果 f 在 x_0 处可微，那么 f 在 x_0 处连续：因为

$$\lim_{x\to x_0} \big(f(x)-f(x_0)\big) = \lim_{x\to x} \frac{f(x)-f(x_0)}{x-x_0} \lim_{x\to x_0}(x-x_0) = 0.$$

但是，并不是连续的函数都可微分，比如说函数 $f(x)=|x|$：在 $x_0=0$ 处，当 $x>0$ 时，

$$\lim_{x\to 0} \frac{x-0}{x-0} = 1.$$

当 $x<0$ 时，我们有

$$\lim_{x\to 0} \frac{-x-0}{x-0} = -1.$$

这说明 $f'(x_0) = \lim_{x\to 0} \frac{f(x)-f(x_0)}{x-x_0}$ 不存在。直观上，连续但是不可微的东西有“尖儿”。

例子. 根据定义，我们可以计算如下两个函数的导数：

1) $f(x)=c$，其中 c 为常数。那么，$(c)'=0$。

2) $f(x)=x$，其中 c 为常数。那么，$(x)'=1$。

练习. 标准的微积分教材都有左导数和右导数的概念，请“望文生义”地定义它们。进一步证明，f 在 x_0 处的导数存在当且仅当 f 在 x_0 处的左右导数都存在并且相等。

注记 (多重导数与记号)**.** 假设 f 是可微函数并且其 f' 也可微，我们就可以继续计算 f' 的导数 $(f')'$。(如果可能的话) 我们可以重复这个过程来求 k-次导数 $\overbrace{((\cdots((f')'\cdots)')'}^{k\text{-个 }'}$。习惯上，人们用符号 $f^{(k)}$ 表示 f 的 k-次导数而两次或三次导数经常记作 $f''(x)$ 和 $f'''(x)$。

另外，如果 f 在 I 上可微分，也用 $\dfrac{df}{dx}$ 来表示 f'。这种写法的好处是我们可以用算子／映射的观点来写微分，一如之前对函数所强调的。比方说，我们有如下的映射（算子：把函数映射为函数／数的映射）：

$$\frac{d}{dx} : C^1(I) \to C(I).$$

我们还用 $\dfrac{d^k f}{dx^k}$ 来表示 $f^{(k)}$，即用算子 $\dfrac{d^k}{dx^k}$ 表示求了 k-次导数。如果要强调 $f^{(k)}$ 在一个点 x_0 处的值，更标准的记号是 $\dfrac{d^k f}{dx^k}\Big|_{x=x_0}$。比方说，$\dfrac{df}{dx}\Big|_{x=x_0} = f'(x_0)$。如果 f 可以在区间 I 上求 k 次导数并且得到的 $f^{(k)}$ 也是连续的，我们就称 f 是 k-次连续可微的并将这样函数的全体记作 $C^k(I)$。记号 $C^\infty(I)$ 的意思很明显：它代表 I 上无限次连续可微函数的全体。

我们注意到 $C^\infty(I) \neq \emptyset$，比如说常数函数 $1 \in C^\infty(I)$。

注记 (无穷小量与导数)**.** 假设 f 在 x_0 处可维，我们令 $r(h) = \dfrac{f(x_0+h)-f(x_0)}{h} - f'(x_0)$，其中 h 定义在某个区间 $(-\varepsilon, \varepsilon)$ 上面。按照导数的定义，当 $h \to 0$ 时，有 $r(h) \to 0$。在分析学中，我们通常说 $r(h)$ 是（$h \to 0$ **时的）无穷小量**，记作 $r(h) = o(1)$。从而，我们经常说（记）$\dfrac{f(x_0+h)-f(x_0)}{h} - f'(x_0) = o(1)$。在形式上，人们还用如下的写法：

$$f(x_0+h) = f(x_0) + f'(x_0)h + o(h),$$

其中我们暂且认为 $o(h) = o(1)h$。这都是数学分析中的黑话，只是为了说话方便或者为了形式上便于记忆。严格地讲，我们有

$$g(h) = f(x_0+h) - f(x_0) - f'(x_0)h,$$

其中 $\lim\limits_{h\to 0} \dfrac{g(h)}{h} = 0$。

我们首先要解决的问题是导数的计算法则。由于导数是用极限来定义的，极限的运算法则可以很容易地翻译成导数的运算法则：

命题 80 (四则运算)**.** 假设函数 f 和 g 在 (a,b) 上有定义并且在 $x_0 \in (a,b)$ 处可微，那么

1) $f \pm g$ 在 x_0 处可微并且 $(f \pm g)'(x_0) = f'(x_0) \pm g'(x_0)$。

2) $f \cdot g$ 在 x_0 处可微并且 $(f \cdot g)'(x_0) = f'(x_0)g(x_0) + f(x_0)g'(x_0)$。

3) 如果 $g(x_0) \neq 0$，那么 $\dfrac{f}{g}$ 在 x_0 处可微并且 $\Big(\dfrac{f}{g}\Big)'(x_0) = \dfrac{f'(x_0)g(x_0) - f(x_0)g'(x_0)}{g(x_0)^2}$。特别地，我们有 $\dfrac{1}{g}$ 在 x_0 处可微并且 $\Big(\dfrac{1}{g}\Big)'(x_0) = -\dfrac{g'(x_0)}{g(x_0)^2}$。

证明: 我们证明相对困难的第三条，其余的留作作业来验证。我们有如下的等式：

$$\begin{aligned}\frac{1}{h}\Big(\frac{f(x_0+h)}{g(x_0+h)} - \frac{f(x_0)}{g(x_0)}\Big) &= \frac{f(x_0+h)g(x_0) - f(x_0)g(x_0+h)}{hg(x_0+h)g(x_0)} \\ &= \frac{\frac{f(x_0+h)-f(x_0)}{h}g(x_0) - f(x_0)\frac{g(x_0+h)-g(x_0)}{h}}{g(x_0+h)g(x_0)}.\end{aligned}$$

我们按照连续函数求极限的四则运算法则对 $h \to 0$ 求极限即可。 □

现在再处理复合函数的导数：

命题 81. （链式法则）I 和 J 是 $\mathbb{R}$ 上的区间，$f: I \to J$ 和 $g: J \to \mathbb{R}$ 是实值函数。如果 f 在 x_0 处可微，g 在 $f(x_0)$ 处可微，那么复合函数 $g \circ f$ 在 x_0 处可微，并且

$$(g \circ f)'(x_0) = g'\big(f(x_0)\big) f'(x_0).$$

证明： 按照导数的定义，我们有

$$\begin{aligned} f(x_0 + h) &= f(x_0) + f'(x_0)h + \delta(h), \\ g\big(f(x_0) + \ell\big) &= g\big(f(x_0)\big) + g'\big(f(x_0)\big)\ell + \Delta(\ell), \end{aligned}$$

其中 $\lim\limits_{h\to 0} \dfrac{\delta(h)}{h} = \lim\limits_{\ell \to 0} \dfrac{\Delta(\ell)}{\ell} = 0$。我们按照导数的定义计算 $f \circ g$ 的导数：

$$\begin{aligned} \frac{g(f(x_0+h)) - g(f(x_0))}{h} &= \frac{g\big(f(x_0) + \overbrace{f'(x_0)h + \delta(h)}^{\ell}\big) - g\big(f(x_0)\big)}{h} \\ &= \underbrace{\frac{g'(f(x_0))[f'(x_0)h + \delta(h)]}{h}}_{\text{当 } h \to 0 \text{ 时，极限}= g'(f(x_0))f'(x_0)} + \frac{\Delta(f'(x_0)h + \delta(h))}{h}. \end{aligned}$$

我们只要证明最后一项当 $h \to$ 时，极限为 0 即可：

根据 $\lim\limits_{\ell \to 0} \dfrac{\Delta(\ell)}{\ell} = 0$，我们定义 $\mu(\ell) = \dfrac{\Delta(\ell)}{\ell}$（提醒一下同学们，尽管函数 μ 在 0 点处没有定义，这并不影响我们研究它在 0 处的极限），那么 $\mu(\ell) = o(1)$，即 $\lim\limits_{\ell \to 0} \mu(\ell) = 0$（我们也把 Δ 写成 $\Delta(\ell) = o(1)\ell$）。应用这些记号，我们有

$$\frac{\Delta(f'(x_0)h + \delta(h))}{h} = \mu(f'(x_0)h + \delta(h))) = \mu(\ell).$$

然而，当 $h \to 0$ 时，$\ell = f'(x_0)h + \delta(h) \to 0$，所以上式的极限是 0。 □

根据上面的运算法则，我们可以计算更多的函数的导（函）数。我们需要熟练记忆如下几个例子的结论和计算技巧。这些例子（实质上）是我们仅有的可以用解析表达式写下来的函数：

1) $f(x) = x^n$，那么，$(x^n)' = nx^{n-1}$，其中 n 为整数。

当 $n \geqslant 0$ 时，根据乘积的导数的计算方法，我们有

$$\big(x^n\big)' = \big(x^{n-1} \cdot x\big)' = \big(x^{n-1}\big)'x + x^{n-1}.$$

据此，利用归纳法立即得到结论。当然，我们也可以直接计算：

$$\begin{aligned} f'(x_0) &= \lim_{h\to 0} \frac{1}{h} \left(\sum_{k \geqslant 0} \binom{n}{k} h^k x_0^{n-k} - x_0^n) \right) \\ &= \lim_{h \to 0} \sum_{k \geqslant 1} \binom{n}{k} h^{k-1} x_0^{n-k} = n x_0^{n-1}. \end{aligned}$$

当 $n < 0$ 时，我们利用倒数的导数计算即可。

2) $f(x)=e^x$（实数值的函数），那么 $(e^x)'=e^x$（这表明指数函数是微分算子 $\dfrac{d}{dx}$ 的一个不动点）。

这个问题明显要复杂，因为 e^x 本身就是通过极限／级数来定义的。我们可以想象一个“理想”的计算：

$$\Big(\sum_{k=0}^{\infty}\frac{x^k}{k!}\Big)'=\sum_{k=0}^{\infty}\Big(\frac{x^k}{k!}\Big)'=\sum_{k=1}^{\infty}k\frac{x^{k-1}}{k!}=\sum_{k=1}^{\infty}\frac{x^{k-1}}{(k-1)!}=e^x.$$

这个计算（目前）是不对的，因为我们不能够随便交换求导数运算和（无限）求和运算的顺序。在导数的研究中，我们将要想办法让上面的计算成立从而使得大部分直观上应该成立的运算都成立，这样子我们就可以很舒服地做很多运算。这是后话。

我们先计算 e^x 在 $x=0$ 处的导数：

$$\lim_{x\to0}\frac{e^x-1}{x}=\lim_{x\to0}\sum_{k=1}^{\infty}\frac{x^{k-1}}{k!}=1+\lim_{x\to0}\sum_{k=2}^{\infty}\frac{x^{k-1}}{k!}$$

然而，

$$\Big|\sum_{k=2}^{\infty}\frac{x^{k-1}}{k!}\Big|\leqslant|x|\Big|\sum_{k=2}^{\infty}\frac{x^{k-2}}{k!}\Big|\leqslant|x|\Big|\sum_{k=2}^{\infty}\frac{x^{k-2}}{(k-1)!}\Big|=|x|e^{|x|}.$$

所以，$\lim\limits_{x\to0}\sum\limits_{k=2}^{\infty}\dfrac{x^{k-1}}{k!}=0$。这表明

$$\frac{d}{dx}\Big|_{x=0}\exp=1.$$

函数 e^x 在 x_0 处的导数现在就完全由 e^x 的代数性质决定了：

$$\lim_{h\to0}\frac{e^{x_0+h}-e^{x_0}}{h}=e^{x_0}\lim_{h\to0}\frac{e^h-1}{h}=e^{x_0}.$$

这是非常有意义的计算，因为我们可以借此看到代数结构是如何在计算中起作用的。

3) $f(x)=\log x$ 的导数，$\log(x)'=\dfrac{1}{x}$。

为了计算 $\log x$，我们自然要研究反函数的可微性质，因为 $\log x$ 就是被定义成反函数，除此之外，我们没有关于 log 的其它信息（事实上，关于 $\log x$ 的进一步性质也都是通过反函数得来的）。关于复合函数求导的命题有下面的推论：

推论 82. *I 和 J 是 $\mathbb{R}$ 上的区间，$f:I\to J$ 是实值函数并且它的逆 $f^{-1}:J\to I$ 存在。假设 f 是可微函数。如果 $f'(x_0)\neq0$，那么 f^{-1} 在 x_0 处可微分并且*

$$\big(f^{-1}\big)'(f(x_0))=\frac{1}{f'(x_0)}.$$

注记. 这个推论事先假设了 f^{-1} 存在。从局部的观点来看，这个条件可以去掉，这是微分学中最重要的定理（之一）：反函数定理。我们很快就会学习反函数定理。

另外，关于复合函数的记法，我们不推荐 $f(x^2)$ 或者 $f(-x)$ 这样的记号（尽管我们会经常这么写）。这些记号都应该被理解为函数的复合。

根据命题，我们显然有 $\big(\log x\big)' = \dfrac{1}{x}$。

我们也可以直接利用定义来计算 $(\log x)'$:

$$\frac{\log(x+h)-\log(x)}{h} = \frac{\log(1+\frac{h}{x})}{h} = \frac{\log\big((1+\frac{h}{x})^{\frac{x}{h}}\big)}{x}$$

根据 $\lim\limits_{y\to\infty}(1+\dfrac{1}{y})^y = e$，上面的式子就给出所要的结论。

4) $f(x) = x^\alpha$，其中 $x>0$，此时 $(x^\alpha)' = \alpha x^{\alpha-1}$。

按照定义，$x^\alpha = e^{\alpha\log x}$ 是如下函数的复合:

$$\mathbb{R}_{>0} \xrightarrow{\log} \mathbb{R} \xrightarrow{\alpha\cdot} \mathbb{R} \xrightarrow{\exp} \mathbb{R}.$$

所以，利用链式法则逐步计算:

$$\Big(x^\alpha\Big)' = \Big(e^{\alpha\log x}\Big)' = e^{\alpha\log x}\big(\alpha\log x\big)' = x^\alpha\alpha(\log x)' = \alpha x^{\alpha-1}.$$

5) $(\sin x)' = \cos x$，$(\cos x)' = -\sin x$。

根据之前的例子，我们有三种方式来计算三角函数的导数:

- 由于 $\sin x$ 是用级数来定义的，仿照 e^x 的导数计算的讨论，我们希望能够逐项求导数，这种做法现在并不严格;
- 注意到 $\sin x$ 和 $\cos x$ 是通过复数值的函数来定义的（$\cos x = \dfrac{1}{2}\big(e^{ix}+e^{-x}\big)$），所以我们必须离开实数转而研究在复数（或者向量空间）中取值的函数的导数。然而，在证明函数的四则运算和复合函数时，我们并没有用到 $\mathbb{R}$ 的具体性质，所有的证明对于复数域仍然成立（我们将会在作业中见到类似的证明），所以这些技术都是成立的，从而:

$$(\cos x)' = \frac{1}{2}\big(e^{ix}+e^{-ix}\big)' = \frac{1}{2}\big(ie^{ix}-ie^{-ix}\big) = -\frac{1}{2i}\big(e^{ix}-e^{-ix}\big) = -\sin x.$$

- 用定义直接计算。我们仿照 e^x 的情形（利用三角函数的代数性质）首先证明 $\sin'(0)=1$ 和 $\cos'(0)=0$，然后对 $\sin(x+h)-\sin x$ 利用和差化积公式来证明一般 x 的情形。我们会在作业中按照这种方式来完成证明。

14 Leibniz 公式，Faà di Bruno 公式，导数的基本性质，在赋范线性空间中取值的函数的导数，求极值，Rolle 与 Lagrange 中值定理，处处连续处处不可微函数的构造

二零一九年十一月四日，星期一，晴天

先证明上次课堂上没有证明的一个重要推论：

I 和 J 是 $\mathbb{R}$ 上的区间，$f : I \to J$ 是实值函数并且它的逆 $f^{-1} : J \to I$ 存在。假设 f 是可微函数。如果 $f'(x_0) \neq 0$，那么 f^{-1} 在 x_0 处可微分并且

$$\big(f^{-1}\big)'(f(x_0)) = \frac{1}{f'(x_0)}.$$

证明： 实际上，只要证明 f^{-1} 在 $y_0 = f(x_0)$ 处可微就可以了，因为一旦我们知道 f^{-1} 是可微的，那么我们就可以对 $f \circ f^{-1}$ 来运用复合函数求导的法则了。

我们不妨假设 f 是严格递增的（因为 f 连续并且可逆，根据介值定理，f 是单调的），那么 f^{-1} 是连续的。特别地，在 J 中，当 $y \to y_0$ 时，我们有 $f^{-1}(y) \to f^{-1}(y_0)$。根据介值定理，当 y 从左边到右边遍历 $[-\varepsilon + y_0, y_0 + \varepsilon]$ 时，令 $x = f^{-1}(y)$，从而 x 也恰好从左边到右边遍历了区间 $[x_0 - \delta_1, x_0 + \delta_2]$，其中 $x_0 - \delta_1 = f^{-1}(y_0 - \varepsilon)$，$x_0 + \delta_2 = f^{-1}(y_0 + \varepsilon)$。特别地，在这个区间里面，$x \to x_0$ **等价于**$y \to y_0$。所以，

$$\begin{aligned}\lim_{y \to y_0} \frac{f^{-1}(y) - f^{-1}(y_0)}{y - y_0} &= \lim_{y \to y_0} \Big(\frac{y - y_0}{f^{-1}(y) - f^{-1}(y_0)}\Big)^{-1} \\ &= \lim_{y \to y_0} \Big(\frac{f(x) - f(x_0)}{x - x_0}\Big)^{-1} = \lim_{x \to x_0} \Big(\frac{f(x) - f(x_0)}{x - x_0}\Big)^{-1}.\end{aligned}$$

上式最后一个不起眼的等号是证明的核心，我们把 $y \to y_0$ 的信息转化为 $x \to x_0$ 的信息。根据极限的四则运算法则，这个极限显然存在。特别地，上述的证明直接给出了反函数求导的公式。 □

我们给出两个著名的公式：

命题 83 (Leibniz 公式)**.** *假设 f 和 g 是 $\mathbb{R}$ 上定义的两个 n-次可导的（实值或复值）函数，那么*

$$(f \cdot g)^{(n)}(x) = \sum_{k=0}^{n} \binom{n}{k} f^{(k)}(x) g^{(n-k)}(x).$$

证明：我们对 n 进行归纳。当 $n=1$ 时，这就是四则运算法则。假设对于 n 命题成立，那么

$$\begin{aligned}(f\cdot g)^{(n+1)} &= \frac{d}{dx}(f\cdot g)^{(n)} = \Big(\sum_{k=0}^{n}\binom{n}{k}f^{(k)}(x)g^{(n-k)}(x)\Big)'\\ &= \sum_{k=0}^{n}\binom{n}{k}f^{(k+1)}(x)g^{(n-k)}(x) + f^{(k)}(x)g^{(n-k+1)}(x)\\ &= \sum_{k=0}^{n+1}\Big(\binom{n}{k-1}+\binom{n}{k}\Big)f^{(k)}(x)g^{(n+1-k)}(x)\\ &= \sum_{k=0}^{n+1}\binom{n+1}{k}f^{(k)}(x)g^{(n+1-k)}(x)\end{aligned}$$

最后一步我们利用了组合数的基本性质：从 $n+1$ 个数中选取 k 个数的方式由两种可能，如果第 $n+1$ 个数在这 k 个数中出现，那么剩下的 $k-1$ 个数要在前 n 个数中选取，一共有 $\binom{n}{k-1}$ 种选取方式；如果第 $n+1$ 个数不在这 k 个数中出现，那么这 k 个数要在前 n 个数中选取，一共有 $\binom{n}{k}$ 种选取方式。 □

我们对于复合函数的 n 次导数也有公式：

命题 84 (Faà di Bruno 公式 **)**.**

$$(f\circ g)^{(n)}(x) = \sum_{m=1}^{n}\sum_{(k_1,\cdots,k_n)\in\Gamma_m}\frac{n!}{k_1!k_2!\cdots k_n!}f^{(m)}(g(x))\Big(\frac{g^{(1)}(x)}{1!}\Big)^{k_1}\Big(\frac{g^{(2)}(x)}{2!}\Big)^{k_2}\cdots\Big(\frac{g^{(n)}(x)}{n!}\Big)^{k_n},$$

其中，集合 Γ_m 的定义如下

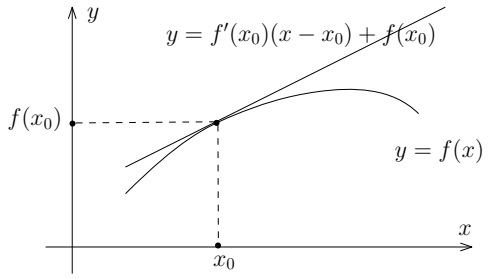
$$\Gamma_m = \big\{(k_1,\cdots,k_n\big|k_1,\cdots,k_n\in\mathbb{Z}_{\geqslant 0}, k_1+k_2+\cdots+k_n=m, k_1+2k_2+\cdots+nk_n=n\big\}.$$

为了对上面的公式有个直观的认识，这次的作业里有对 $n=2$ 或 3 的验证，这是练习链式法则的好例子。由于命题证明的实质是组合数学而和分析学没有更进一步的联系，我们略去，请感兴趣的同学查阅互联网或者其他书籍。

用导数来研究函数的性质是数学分析中的重要课题，在学习这一部分知识之前，我们重新来审视一下导数的“直观意义”：

首先，导数的精确定义是通过公式 $f(x_0)=\lim\limits_{x\to x}\dfrac{f(x)-f(x_0)}{x-x_0}$ 给的。通常有如下几种“直观的”解释：

1) 几何的看法，即把导数解释为函数图像的斜率。



2) 物理的若干看法：$f(t)$ 是一个质点的位置时，导数是瞬间的速度，两次导数是加速度；当 f 代表的是一段线段的质量时，导数是一点处的密度。

下面强调的是**逼近的**解释，我们说过这是分析学中贯穿始终的看问题的方法。

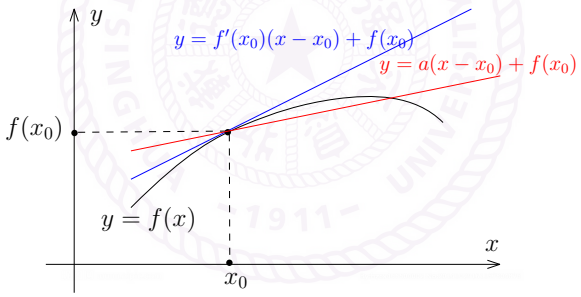
假设 f 在 x_0 处可微，我们定义**线性**映射：

$$\ell : \mathbb{R} \to \mathbb{R}, \quad h \mapsto f'(x_0)h,$$

即令 $\ell(h) = f'(x_0)h$。再令 $\delta(h) = f(x_0+h) - f(x_0) - f'(x_0)h = \big(f(x_0+h) - f(x_0)\big) - \ell(h)$，按照导数的定义，有 $\lim\limits_{h\to 0} \dfrac{\delta(h)}{h} = 0$，换而言之，$h \to 0$ 时，$\delta(h) = o(1)h$。我们把这些讨论写成下面的样子：

$$f(x) - f(x_0) = \ell(x - x_0) + o(1)(x - x_0), \quad x \to x_0.$$

上面最后一项 $o(1)(x - x_0)$ 应该理解为是一个相对于 $x - x_0$ 很小的项，那么直观上，当 $x \to x_0$ 时，$\ell(x - x_0)$ 已经可以比较好的刻画 $f(x) - f(x_0)$（从而刻画 $f(x)$），因为误差应该相对于 $x - x_0$ 还小很多。换句话说，我们直观上认为线性函数 $\ell(x - x_0) + f(x_0)$ 在 x_0 的附近（很小的领域）应该是 f 的很好的逼近。实际上，这是 f 的最好的线性逼近（下图中的蓝色线比红色线更好的“贴近”了本来函数的图象）：



考虑另外线性映射 $\ell'(h) = ah$，其中，$\ell' \neq \ell$，即 $f'(x_0) \neq a$。我们用 ℓ 和 ℓ' 在 x_0 的局部上（即要求 $x - x_0$ 比较小）来逼近 $f(x) - f(x_0)$，也就是说，

$$f(x) - f(x_0) - \ell(x - x_0) = \underbrace{o(1)h}_{E_1(h),\text{第一个误差}},$$

$$f(x) - f(x_0) - \ell'(x - x_0) = \underbrace{(f'(x_0) - a)h + o(1)h}_{E_2(h),\text{第二个误差}}.$$

如果令 $h = x - x_0$，这两个误差项有下面的比较

$$\lim_{h\to 0} \frac{E_2(h)}{E_1(h)} = \lim_{h\to 0} \frac{f'(x_0) - a)h + o(1)h}{h} = +\infty.$$

作为总结，我们有

注记. 如果要用一个线性映射 $L:\mathbb{R}\to\mathbb{R}$ 所定义的 $L(x-x_0)$ 在 x_0 附近来逼近 $f(x)-f(x_0)$，那么 $h\mapsto f'(x_0)h$ 是最好的线性逼近。这里有观点上的重要转变，如果用

$$df(x_0):\mathbb{R}\to\mathbb{R},\quad h\mapsto f'(x_0)h,$$

来表示这个线性映射（请注意，$df(x_0)$ 只是一个记号，请不要急于赋予它任何含义），我们把线性映射 $df(x_0):\mathbb{R}\to\mathbb{R}$ 的地位摆的高于了导数 $f'(x_0)$（尽管它是用导数定义出来的）。另外，我们强调过，每写下一个映射都要说清楚它是从哪里映射到何方的，所以，我们把 $df(x_0)$ 的定义域所对应的线性空间 $\mathbb{R}$ 记为 $T_{x_0}\mathbb{R}$，把 $df(x_0)$ 的值域所对应的线性空间 $\mathbb{R}$ 记为 $T_{f(x_0)}\mathbb{R}$，此时，我们可以把上述映射写为更为标准的形式：

$$df(x_0):T_{x_0}\mathbb{R}\to T_{f(x_0)}\mathbb{R},\quad h\mapsto f'(x_0)h.$$

当我们研究高维甚至是无限维空间（定义域和值域都可以是任意的）的时候，这个新的观点（用线性映射在一点附近来逼近一个映射）有着非常自然的推广。从这个观点来看，我们将要研究的偏导数和微分之间的关系也会更明朗。

定义 85 (函数在一点处的微分)**.** 我们把上述定义的 $df(x_0):T_{x_0}\mathbb{R}\to T_{f(x_0)}\mathbb{R}$ 称作是 f 在 x_0 处的**微分**。

在进一步探究可微函数更为精细的结构之前，如果对比收敛和连续性部分内容，我们很自然会问是否可以将导数的概念推广到其他的空间（值域）？比如说，$f:\mathbb{R}\to\mathbb{C}$ 或者 $f:\mathbb{R}\to\mathbb{R}^N$ 的导数怎么定义？回顾导数的定义：

$$f(x_0)=\lim_{x\to x}\frac{f(x)-f(x_0)}{x-x_0}.$$

为了定义的右边，我们用到了 $f:\mathbb{R}\to V$ 的**值域** V 的如下性质：

1) 两个值可以减（加）；

2) 可以能除以实数。

所以，只要 V 是实线性空间上述就可以满足。另外，由于要取极限，所以 V 最好是距离空间，因此，如果 V 是赋范线性空间，那么导数就应该能定义了（当然需要要求极限存在）

定义 86. 假设 $(V,\|\cdot\|)$ 是赋范线性空间，$I\subset\mathbb{R}$ 是区间，$f:\mathbb{R}\to\|\cdot\|$ 是一个映射。对任意的 $x_0\in I$，如果极限

$$\lim_{x\to x}\frac{f(x)-f(x_0)}{x-x_0}$$

存在，我们就称 f 在 x_0 处**可导**并记此时 $f'(x_0)=\lim_{x\to x}\dfrac{f(x)-f(x_0)}{x-x_0}$。按照定义，$f'(x_0)\in V$ 是 V 中的向量。

例子. 考虑 $V=\mathbb{R}^n$（即选定有限维线性空间的 V 的一组基）以及任一个你喜欢的范数，比如 $\|\cdot\|_2$。我们考虑映射

$$f:\mathbb{R}\to\mathbb{R}^n,\quad x\mapsto f(x)=(f_1(x),\cdots,f_n(x)).$$

我们已经证明过，向量值的函数取极限等价于对每个分量都取极限，所以，如果 f 在 x_0 处的导数存在当且仅当对每个分量函数 f_k，它在 x_0 处的导数都存在并且

$$f'(x)=\big(f_1'(x),\cdots,f_n'(x)\big).$$

特别地，复值函数的导数就是对实部和虚部分别求导数。作为应用，如果要求 $\sin x$ 和 $\cos x$ 的导数，只要对 e^{ix} 求导数就好。为此，我们迫切希望能有复合函数求导数法则（$z\mapsto iz$）。但是，这要求我们要对定义域是 $\mathbb{C}$ 的函数求导数。这样一来，我们就不得不研究多元函数的微分。这是后话。

另外，我们可以按定义来计算 e^{ix} 的导数是什么，请见本次作业。

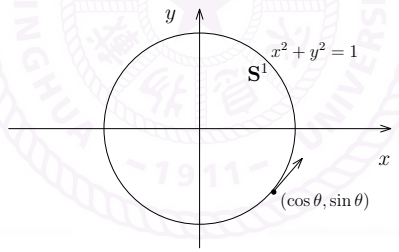
我们定义映射

$$E:\mathbb{R}\to\mathbb{C}=\mathbb{R}^2,\quad \theta\mapsto(\cos\theta,\sin\theta).$$

考虑

$$\mathbf{S}^1=\big\{(x,y)\in\mathbb{R}^2\big|x^2+y^2=1\big\}.$$

这是 $\mathbb{R}^2$ 上用 $\|\cdot\|_2$ 来看长度为 1 的向量的全体，我们把它称作是**单位园**。



试证明 $E(\mathbb{R})=\mathbf{S}^1$（即单位圆上的每个点都可以写成 $(\sin\theta,\cos\theta)$ 的形式，这不是显然的）。对任意的 $\theta\in\mathbb{R}$，试计算 $E'(\theta)$（这一般被视为是单位圆在一点处的切向量）。这是本次作业的一道题目，是非常有意义的练习。

可微函数的性质

导数是局部定义的，我们用导数研究函数的局部性质：

引理 87. *$I\subset\mathbb{R}$ 是开区间，$f:I\to\mathbb{R}$ 在 x_0 处可微并且 $f(x_0)\neq 0$。我们假设 $f'(x_0)>0$。那么，存在 x_0 的开临域 $U=(x_0-\varepsilon,x_0+\varepsilon)$，使得*

1) 对任意的 $x\in U$，$x>x_0$，有 $f(x)>f(x_0)$；

2) 对任意的 $x\in U$，$x<x_0$，有 $f(x)<f(x_0)$。

如果假设 $f'(x_0)<0$ 有类似的结论。

证明: 根据微分以及极限的定义，由于 $f'(x_0)>0$，所以对于 $\delta=\frac{1}{2}f'(x_0)$，存在 $\varepsilon>0$，使得对任意 $|x-x_0|<\varepsilon$，即 $x\in U=(x_0-\varepsilon,x_0+\varepsilon)$，我们有

$$\left|\frac{f(x)-f(x_0)}{x-x_0}-f'(x_0)\right|<\frac{1}{2}f'(x_0)\ \Rightarrow\ \frac{f(x)-f(x_0)}{x-x_0}>\frac{1}{2}f'(x_0)>0.$$

所以 $f(x)-f(x_0)$ 与 $x-x_0$ 同号，从而命题得证。 □

注记. 在所谓的临界情形，$f'(x_0)\geqslant 0$，我们不能判断 f 在 x_0 左右的取值和 $f(x_0)$ 之间的大小。试举出例子。

推论 88. $I\subset\mathbb{R}$ 是开区间，$f:I\to\mathbb{R}$ 在 I 上可微并且对任意的 $x\in I$，$f'(x)>0$。那么，f 是 I 上严格递增的函数。

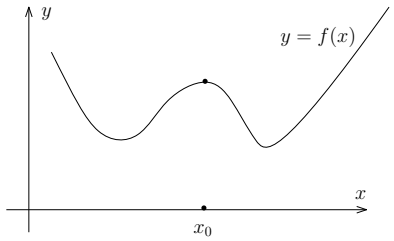
证明: 这是因为严格递增是一个局部性质。 □

不意味着函数单调。

注记. 在所谓的临界情形，$f'(x_0)\geqslant 0$，我们不能判断 f 在 x_0 左右的取值和 $f(x_0)$ 之间的大小。试举出例子。然而，如果对任意的 $x\in I$，$f'(x)\geqslant 0$。那么，f 是 I 上递增（未必严格）函数。我们将用中值定理证明这个性质。

另外，即使 $f'(x_0)>0$，我们也无法说明存在 x_0 的开临域 $U=(x_0-\varepsilon,x_0+\varepsilon)$，使得 f 在整个小领域 U 上是单调上升的，你能否给出这样的反例？（作业）

上面引理有一个重要的推论（然而简单），它可以帮助我们寻找函数的最大最小值（在非数学领域中，这个定理是可能被应用的最多的，比如说经济学和工程里）。为此，我们引入局部极大值和局部极小值的概念。假设 $f:I\to\mathbb{R}$ 是区间 $I\subset\mathbb{R}$ 上定义的函数，$x_0\in I$，如果存在 x_0 的小邻域 $U\subset I$（这是局部的含义），使得 $f(x_0)$ 是函数 f 在 U 上的最大值，我们就称 x_0 是 f 的一个**局部极大值**；类似地，我们可以定义**局部极小值**。x_0 是 f 的局部极大值并不意味着 f 在 x_0 处取到它的（整体）最大值，比如说下面的例子：

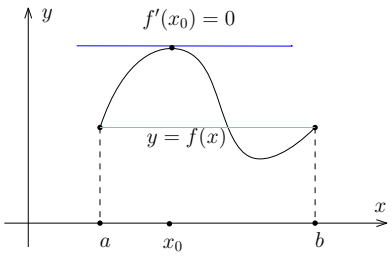


定理 89. 假设 f 是 I 上的可微函数并且 f 在 $x_0\in I$ 处是局部极大(或者极小)值，那么 $f'(x_0)=0$。作为应用，为了找 $f(x)$ 的最大值，我们希望研究 $f'(x)$ 的零点。

证明: 如若不然，不妨假设 $f'(x_0)>0$，那么 f 在 x_0 右边的附近的点的取值比 $f(x_0)$ 要大，所以，x_0 就不可能是局部极大值，矛盾。 □

利用这个定理，证明两个出名的定理（在这两个定理之上有一大类有意思或者困难的习题，然而这两个定理是一元微分学中的结果，在高维空间没有特别有意义的推广）：

定理 90 (Rolle 中值定理)**.** *假设实值函数 $f \in C^0([a,b])$ 并且在 (a,b) 上可微。如果 $f(a) = f(b)$，那么存在 $x_0 \in (a,b)$，使得 $f'(x_0) = 0$。*



证明： 如果 f 是常值函数，那么不证自明。如果 f 不是常值函数，不妨设有 $x_1 \in (a,b)$，使得 $f(x_1) > f(a) = f(b)$。由于连续函数在闭区间上有最大值，我们假设 x_0 是 $f(x)$ 的最大值。所以，$f(x_0) \geqslant f(x_1) > f(a) = f(b)$，这表明 $x_0 \in (a,b)$。另外，x_0 自然是局部极大的，所以 $f'(x_0) = 0$。 □

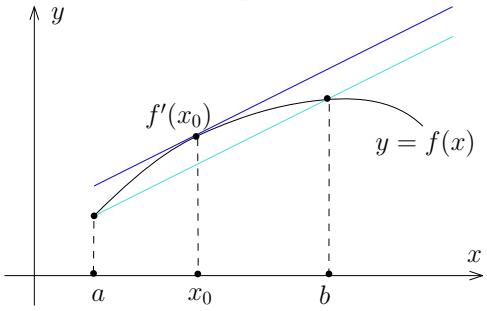
注记. *Rolle 中值定理对向量值的函数不成立，比如说，我们考虑（先假设 π 的存在性以及 2π 是 $\sin x$ 和 $\cos x$ 的周期，我们不久就会证明这个性质）：*

$$E : [0.2\pi] \to \mathbb{R}^2, \quad x \mapsto (\cos x, \sin x).$$

我们知道，$E(0) = E(2\pi)$，但是，对任意的 $x \in [2, 2\pi]$，$E'(x) \neq 0$。

定理 91 (Lagrange 中值定理)**.** *假设实值函数 $f \in C^0([a,b])$ 并且在 (a,b) 上可微。那么，存在 $x_0 \in (a,b)$，使得 $f'(x_0) = \dfrac{f(b) - f(a)}{b - a}$。*

最容易记住这个定理的方式就是搞明白下面的图讲了什么样的几何意义：



证明： 考虑函数

$$g(x) = f(x) - \frac{f(b) - f(a)}{b - a}(x - a).$$

我们知道，$g(a) = g(b)$ 都等于 $f(a)$，所以可以用 Rolle 中值定理，存在 $x_0 \in (a,b)$，使得

$$g'(x_0) = 0 \quad \Leftrightarrow \quad f'(x_0) - \frac{f(b) - f(a)}{b - a} = 0.$$

命题得证。 □

我们证明 1 维版本的反函数定理（由于 $\mathbb{R}^1$ 的几何比 $\mathbb{R}^n$ 简单很多，我们可以利用 $\mathbb{R}^1$ 的特殊性来“投机取巧”，所以这个证明比我们将来要学习的高维版本简单很多（所以也不能推广））。另外，这是我们第一次来体会连续可微和可微这两个概念之间的细微（巨大）差别。

定理 92 (反函数定理（C^1 版本）). *$I \subset \mathbb{R}$ 是开区间，$f \in C^1(I;\mathbb{R})$，即连续可微的实值函数。如果 $f'(x_0) \neq 0$，那么 f 在 x_0 的一个邻域上是 C^1-同胚，即存在 x_0 的邻域 $(-\varepsilon + x_0, x_0 + \varepsilon)$ 和 $f(x_0)$ 的领域 $(f(x_0) - \delta_1, f(x_0) + \delta_2)$，使得 f 在 $(-\varepsilon + x_0, x_0 + \varepsilon)$ 的限制给出的*

$$f\big|_{(-\varepsilon+x_0,x_0+\varepsilon)} : (-\varepsilon + x_0, x_0 + \varepsilon) \longrightarrow f(-\varepsilon + x_0, x_0 + \varepsilon) = (f(x_0) - \delta_1, f(x_0) + \delta_2)$$

是双射并且它的逆

$$f^{-1} : (f(x_0) - \delta_1, f(x_0) + \delta_2) \longrightarrow (-\varepsilon + x_0, x_0 + \varepsilon)$$

也是连续可微的。

证明: 不妨假设 $f'(x_0) > 0$。由于 f' 连续，所以存在 x_0 的邻域 $U = (-\varepsilon + x_0, x_0 + \varepsilon)$，使得 f' 在 U 上的取值都是正的（这里用到了 f' 的连续性!）。所以，f 是 U 上严格递增的函数。在本次课的一开始关于关于反函数导数的结论中，我们证明了 $f : U \to f(U)$ 是双射（去掉端点）并且 $f^{-1} : f(U) \to U$ 是可微的。剩下只要说明 f^{-1} 是连续可微的即可，这因为

$$\left(f^{-1}\right)'(y) = \frac{1}{f'(f^{-1}(y))}$$

是连续函数的复合。 □

推论 93 (反函数定理-C^∞ 版本). *在上述的定理中，如果我们进一步要求 f 是光滑的，即无限次连续可微（$f \in C^\infty I$），那么它逆 f^{-1} 也是光滑的。*

证明: 定理已经说明 f^{-1} 是可微的并且 $\left(f^{-1}\right)'(y) = \dfrac{1}{f'(f^{-1}(y))}$。根据可谓函数的复合仍然可微，所以 $(f^{-1})'$ 还可微的并且可以计算它的导数:

$$\left(\left(f^{-1}\right)'\right)'(y) = -\frac{(f^{-1})'(y) f''(f^{-1}(y))}{(f'(f^{-1}(y)))^2}.$$

据此，f^{-1} 的二次导数也可微。我们可用归纳的方式继续求导，值得注意的是分母上只有 $f'(f^{-1}(y))$ 出现，它永远不会是零。 □

注记. *反函数定理是一个纲领性的定理，凡是用到了微积分的课程它总会出现。定理的本意是如何正确地参数化一个几何对象，由于在一维空间上的结构简单，问题的解决可以依赖于一维空间的特殊性，所以我们体会可能不深。反函数定理在高维空间的情形会以最自然最朴素的方式登场。*

另外，这个定理已经包含了所谓的椭圆正则性（偏微分方程中的黑话）的雏形：定理告诉我们，只要知道是 C^1 的，我们就可以“赢得”无限多个导数!

处处不可微的连续函数

如果 f 是 $\mathbb{R}$ 上的可微函数，我们知道 f 是连续函数。然而，连续性不能退出可微性。历史上有一个著名例子说存在 $\mathbb{R}$ 上的实值连续函数 f，它在 $\mathbb{R}$ 的每个点处的导数都不存在，这就是 Weierstrass 所构造的函数：

$$f(x)=\sum_{k=1}^{\infty}a^k\cos(b^k\pi x).$$

通过这个例子，我们还可以加深对闭区间上连续函数所构成的空间的认识。

我们要求 $a\in(0,1)$，$b\in\mathbb{Z}_{\geqslant 1}$ 是奇数并且 $ab>M_0$，其中 M_0 是个很大的数，目前待定，在证明结束的时候我们就可以确定它的大小。

首先，对任意固定的 $N>0$，在连续函数空间 $C([-N,N])$ 中来研究上面函数级数。回忆一下，在 $C([-N,N])$ 上，我们用如下的范数 $\|g\|_\infty=\sup\limits_{x\in[-M,M]}|g(x)|$。由于

$$\sum_{k=1}^{\infty}\|a^k\cos(b^k\pi x)\|_\infty\leqslant\sum_{k=1}^{\infty}a^k<\infty,$$

也就是说级数是绝对收敛的，从而，级数是收敛的。特别地，$f(x)\in C^0([-N,N])$。令 $N\to\infty$，我们就知道 $f(x)\in C(\mathbb{R})$。这表明，f 是 $\mathbb{R}$ 上的连续函数。

为了说明 f 在任何一点处都没有导数，我们先谈一下如下直观的感受：对每个基本的单位 $a^k\cos(b^k\pi x)$ 而言，如果 b^k 很大，它就振荡得很厉害。特别地，它的导数在某些地方（这些点会越来越密）的大小是 $(ab)^k\pi$。这样，我们每次都加上一个振荡很大的基本单位，希望最终得到的函数振荡变成无穷大！

任意固定 $y_0\in\mathbb{R}$，为了说明 f 在 y_0 处不可微，我们需要找一列点 $\{y_n\}_{n\geqslant q}$，使得 $y_n\to y_0$ 而

$$\lim_{y_n\to y_0}\frac{f(y_n)-f(y_0)}{y_n-y_0}$$

不存在。

首先，对任意的正整数 n，存在唯一的整数 z_n，使得 $b^ny_0-z_n\in[0.1,1.1)$，我们令 $y_n=\dfrac{z_n}{b^n}$。很明显，我们有 $y_n\to y_0$。

其次，我们将要计算的极限拆为两项：

$$\begin{aligned}\frac{f(y_n)-f(y_0)}{y_n-y_0}&=\sum_{k=1}^{\infty}a^k\frac{\cos(b^k\pi y_n)-\cos(b^k\pi y_0)}{y_n-y_0}\\&=\underbrace{\sum_{k=1}^{n-1}(ab)^k\frac{\cos(b^k\pi y_n)-\cos(b^k\pi y_0)}{b^k(y_n-y_0)}}_{S_1}+\underbrace{\sum_{k=0}^{\infty}a^{n+k}\frac{\cos(b^{n+k}\pi y_n)-\cos(b^{n+k}\pi y_0)}{y_n-y_0}}_{S_2}\end{aligned}$$

用 Lagrange 中值定理来估计 S_1：

$$\begin{aligned}|S_1| &\leqslant \sum_{k=1}^{n-1}(ab)^k\Big|\frac{\cos(b^k\pi y_n)-\cos(b^k\pi y_0)}{b^k(y_n-y_0)}\Big| \\ &\leqslant \sum_{k=1}^{n-1}(ab)^k\pi|\cos(\theta_k)| \leqslant \sum_{k=1}^{n-1}\pi(ab)^k \\ &\leqslant \pi\frac{(ab)^n}{ab-1}.\end{aligned}$$

再来估计 S_2。首先，由于 z_k 是整数，b 是奇数，按照 y_n 的定义，我们有所以

$$\cos(b^{n+k}\pi y_n)=\cos(b^k\pi z_n)=(-1)^{z_n}.$$

另外，如果令 $\vartheta_k=b^ky_0-z_k\in[0.1,1.1)$，我们有

$$\cos(b^{n+k}\pi y_0)=\cos(b^ny_0\cdot b^k\pi)=\cos(b^k\pi z_n+b^k\pi\vartheta_n)=(-1)^{z_n}\cos(b^k\pi\vartheta_n).$$

所以，根据 $y_n-y_0=-\dfrac{\vartheta_n}{b^n}$，我们得到

$$\begin{aligned}S_2&=\sum_{k=0}^{\infty}(-1)^{z_n}a^{n+k}\frac{1-\cos(b^k\pi\vartheta_n)}{y_n-y_0} \\ &=\sum_{k=0}^{\infty}(-1)^{z_n+1}(ab)^na^k\frac{1-\cos(b^k\pi\vartheta_n)}{\vartheta_n} \\ &=(-1)^{z_n+1}(ab)^n\sum_{k=0}^{\infty}a^k\frac{1-\cos(b^k\pi\vartheta_n)}{\vartheta_n}.\end{aligned}$$

利用 $\vartheta_n\geqslant 1$（注意 ϑ_k 的选取），我们知道

$$\Big|\sum_{k=0}^{\infty}a^k\frac{1-\cos(b^k\pi\vartheta_n)}{\vartheta_n}\Big|=\sum_{k=0}^{\infty}a^k\frac{1-\cos(b^k\pi\vartheta_n)}{\vartheta_n}\geqslant\frac{1-\cos(\pi\vartheta_n)}{\vartheta_n},$$

其中，在上面的不等式中，我们就保留了第一项。

根据 $\vartheta_n\in[0.1,1.1)$，我们知道 $1-\cos(\pi\vartheta_n)\geqslant\delta_0$，并且 $\vartheta_n\geqslant 0.1$，所以

$$\Big|\sum_{k=0}^{\infty}a^k\frac{1-\cos(b^k\pi\vartheta_n)}{\vartheta_n}\Big|\geqslant 1.1^{-1}\delta_0=\frac{\delta_0}{10}.$$

从而，

$$|S_2|\geqslant\frac{\delta_0}{10}(ab)^n.$$

综合上述，我们有

$$|S_1+S_2|\geqslant|S_2|-|S_1|\geqslant(ab)^n\left(\frac{\delta_0}{10}-\frac{\pi}{ab-1}\right).$$

为了要求 S_1+S_2 变的足够大，我们要求

$$ab>\frac{10\pi}{\delta_0}+1=M_0.$$

此时，当 $n\to\infty$ 时，我们有 $\lim\limits_{y_n\to y_0}\dfrac{f(y_n)-f(y_0)}{y_n-y_0}=+\infty$，这就完成了构造。

15 $x \mapsto \exp(xA)$ 的导数，中值定理的应用：线性常微分方程，Darboux 介值定理，Cauchy 中值定理，三角函数与微分方程，π 的定义

二零一九年十一月七日，星期四，晴天

$x \mapsto \exp(xA)$ 的导数及应用

按照导数的定义，如果 $(V, \|\cdot\|)$ 是赋范线性空间，$f : \mathbb{R} \to V$ 是映射，我们可以定义导数

$$f'(x_0) = \lim_{x \to x} \frac{f(x) - f(x_0)}{x - x_0}.$$

（如果极限存在的话）特别地，$f'(x_0) \in V$。我们来研究我们最爱的例子：exp。令 $V = \mathbf{M}_n(\mathbb{R})$ 为矩阵的空间（复数系数的矩阵的证明是一致的），A 是一个给定的 $n \times n$ 的矩阵，考察映射

$$f : \mathbb{R} \to \mathbf{M}_n(\mathbb{R}),\ x \mapsto e^{xA}.$$

我们计算 $f'(x_0)$。按定义，我们有

$$\frac{f(x_0 + h) - f(x_0)}{h} = \frac{e^{(x_0+h)A} - e^{x_0}}{h} = e^{x_0 A} \frac{e^{hA} - 1}{h},$$

其中我们用到了矩阵 $x_0 A$ 与 hA 可交换。由此可见（在证明 $\left(e^x\right)' = e^x$ 的时候也做了同样的事情），只要计算 $f'(0)$ 即可。利用定义，我们有

$$\frac{e^{hA} - 1}{h} = \sum_{k=1}^{\infty} \frac{1}{k!} h^{k-1} A^k = A + \sum_{k=1}^{\infty} \frac{1}{(k+1)!} h^k A^{k+1}.$$

根据矩阵乘法对于范数的关系（$\|A \cdot B\| \leqslant c\|A\|\|B\|$，其中 $\|\cdot\|$ 是任意一个指定的范数），我们有

$$\| \sum_{k=1}^{\infty} \frac{1}{(k+1)!} h^k A^{k+1} \| \leqslant h \sum_{k=1}^{\infty} \frac{1}{(k+1)!} c^k \|A\|^{k+1} \leqslant \frac{h}{c} e^{c\|A\|}.$$

从而，$f'(0) = A$。进一步，我们知道

$$f'(x) = Ae^{xA} = Af(x).$$

上一次最后提到了 Lagrange 中值定理的一个推论：

推论 94 (常微分方程解的存在唯一性-儿童版本)**.** *$f \in C^0([a, b])$ 并且 f 在 (a, b) 上可微，如果 $f'(x) \equiv 0$ 并且 $f(a) = c$，那么 $f(x) \equiv c$。换而言之，如下的常微分方程*

$$\begin{cases} f'(x) = 0, \\ f\big|_{x=a} = c, \end{cases}$$

存在唯一的解。

证明：用反证法：如果存在某个 $x_1 \in (a,b)$ 使得 $f(x_1) \neq c$，那么，在 $[a,x_1]$ 上使用 Lagrange 中值定理，就有 $x_0 \in (a,x_1)$，使得

$$f'(x_0) = \frac{f(x_1)-c}{x_1-a} \neq 0.$$

矛盾。 □

这个推论实际上对于在赋范线性空间中取值的映射也成立：

命题 95. 给定映射 $f \in C^0([a,b];V)$，f 在 (a,b) 上可微。如果 $f'(x) \equiv 0$，$f(a) = c$，那么对任意的 $x \in [a,b]$，$f(x) = c$。

我们满足于 $V = \mathbb{R}^n$ 是有限维的情形，此时，$f : \mathbb{R} \to \mathbb{R}^n$ 可以用分量 $\big(f_1(x), \cdots, f_n(x)\big)$ 表达并且

$$f'(x) = \big(f_1'(x), \cdots, f_n'(x)\big) = c \quad \Leftrightarrow \quad f_k'(x) \equiv c_k, \text{ 任意的 } k = 1, \cdots, n.$$

所以，利用 1 维的情况就可以得到结论。

如果 V 是无限维的赋范线性空间，比如说 $\Big(C\big([0,1]\big), \|\cdot\|_\infty\Big)$，我们一方面没有 Lagrange 中值定理，所以不能用本来的证明；一方面想把问题约化成 1 维的情况的做法不是很容易推广（我们的确能做到这一点）。然而，我们不打算在这个问题上继续深入（除非我们想仔细的研究一下微分方程理论），有限维的情况对我们就足够用了。

作为应用，我们仍然回到 exp 这个例子：考虑可微映射 $f : \mathbb{R} \to \mathbf{M}_n(\mathbb{R})$，假设它满足如下的微分方程

$$\begin{cases} f'(x) = A \cdot f(x), \\ f\big|_{x=a} = \mathbf{I}_n, \end{cases}$$

其中，$\mathbf{I}_n$ 为 $n \times n$ 的矩阵。我们知道，$x \mapsto e^{xA}$ 是这个方程的一个解。我们现在说明，这是唯一的解。实际上，我们有

$$e^{-xA}(f' - Af) = 0 \Rightarrow e^{-xA}f' + \big(e^{-xA}\big)'f = 0.$$

如果我们有 Leibeniz 法则的话（两个函数相乘之后求导数的法则），那么就有

$$\big(e^{-xA}f\big)' = 0.$$

这是把一个表达式写成全微分的技巧。根据唯一性的命题，我们知道 $e^{-xA}f(x) \equiv \mathbf{I}$，所以我们就证明了 $f(x) = e^{xA}$（因为 $e^A e^{-A} = \mathbf{I}_n$）。

回到 Leibniz 法则的问题，如果 f 和 g 是 $n \times n$ 矩阵中取值的函数，我们有

$$\begin{aligned} (f\cdot g)'(x_0) &= \lim_{h\to 0} \frac{f(x_0+h)g(x_0+h) - f(x_0)g(x_0)}{h} \\ &= \lim_{h\to 0} \frac{f(x_0+h)\big(g(x_0+h) - g(x_0)\big)}{h} + \lim_{h\to 0} \frac{\big(f(x_0+h) - f(x_0)\big)g(x_0)}{h} \\ &= f(x_0)g'(x_0) + f'(x_0)g(x_0). \end{aligned}$$

这个证明和最基本的函数版本的证明没有任何的区别。

利用 Lagrange 中值定理，我们可以进一步联系单调性和导数：

推论 96. *假设实值函数 $f \in C^0([a,b])$ 并且 f 在 (a,b) 上可微。那么，下面的两条性质是等价的：*

1) $f'(x) \geqslant 0$。

2) $f(x)$ 是递增的。

特别地，如果对任意 x，$f'(x) > 0$，那么函数是严格递增的（反之不然）。

证明： 2）⇒ 1) 用导数的定义即可：由于 $f(x)$ 递增，所以 $f(x+h) - f(x)$ 与 h 的符号是一致的，从而

$$\frac{f(x+h) - f(x)}{h} \geqslant 0.$$

令 $h \to 0$ 即可。

1）⇒ 2) 用反证法：如若不然，存在 $x_1 < x_2$，使得 $f(x_1) > f(x_2)$，根据 Lagrange 中值定理，存在 $x_0 \in [x_1, x_2]$，使得

$$f'(x_0) = \frac{f(x_2) - f(x_1)}{x_2 - x_1} < 0.$$

矛盾。

如果 $f'(x) > 0$，为了说明函数是严格递增的，我们仍然用反证法：如若不然，存在 $x_1 < x_2$，使得 $f(x_1) \geqslant f(x_2)$，根据 Lagrange 中值定理，存在 $x_0 \in [x_1, x_2]$，使得

$$f'(x_0) = \frac{f(x_2) - f(x_1)}{x_2 - x_1} \leqslant 0.$$

矛盾。 □

我们现在证明导函数的介值定理。

推论 97 (Darboux)**.** *假设 f 在 $[a,b]$ 上可微，$f'(a) < f'(b)$。那么，对于任意的 $c \in (f'(a), f'(b))$，都存在 $x_0 \in (a,b)$，使得 $f'(x_0) = c$。*

注记. *存在在 $[a,b]$ 上可微的函数 f，它的导数不连续（因此不能直接使用连续函数的介值定理）：*

$$f(x) = \begin{cases} x^2 \sin(\frac{1}{x}), & x \neq 0; \\ 0, & x = 0 \end{cases}$$

很明显，我们有

$$f'(x) = \begin{cases} 2x \sin(\frac{1}{x}) - \cos(\frac{1}{x}), & x \neq 0; \\ 0, & x = 0 \end{cases}$$

我们来说明 $f'(x)$ 在 0 处不连续。实际上，我们考虑两个点列 $\{x_n\}_{n\geqslant 1}$ 和 $\{y_n\}_{n\geqslant 1}$，其中

$$x_n = \frac{1}{2n\pi}, \quad y_n = \frac{1}{(2n+1)\pi}.$$

我们有 $\lim_{n\to\infty} x_n = \lim_{n\to\infty} y_n = 0$，但是，

$$\lim_{n\to\infty} f'(x_n) = -1, \quad \lim_{n\to\infty} f'(y_n) = 1.$$

证明: 通过把 f 换为 $f(x) - cx$，我们不妨假设 $c = 0$，$f'(a) < 0 < f'(b)$。在这种假设下，f 连续但是不是单调的函数（因为导数是改变符号的），根据介值定理，f 不能是单射。所以，存在 $x_1, x_2 \in [a, b]$，使得 $f(x_1) = f(x_2)$，那么，Rolle 中值定理就给出了 c 的存在性。 □

另一个版本的中值定理叫做 Cauchy 中值定理，描述了两个不同函数之间关系：

命题 98 (Cauchy 中值定理)**.** *假设有实值函数 $f, g \in C^0([a, b])$ 并且 f 和 g 均在 (a, b) 上可微。假设对任意的 $x \in (a, b)$，$g'(x) \neq 0$。那么，存在 $x_0 \in (a, b)$，使得*

$$\frac{f'(x_0)}{g'(x_0)} = \frac{f(b) - f(a)}{g(b) - g(a)}.$$

我们注意到，结合 Darboux 的定理，$g'(x) \neq 0$ 这个条件意味着 $g'(x)$ 要么恒为正，要么恒为负，所以函数是严格单调的。特别地，这说明 $g(a) \neq g(b)$，从而上式右边是良好定义的。

尽管这个命题可能在一些较难的习题中大显身手，但就本身的深刻程度而言不过是 Rolle 中值定理的一个无关痛痒的推广。然而，我们给出两个不同的证明并且解释怎么理解这个命题。对一个命题有（编）一个感性的认识在学习数学中是非常重要的。

证明: 1. 我们定义如下的函数：

$$F(x) = f(x) - f(a) - \frac{f(b) - f(a)}{g(b) - g(a)}\big(g(x) - g(a)\big).$$

因为 $F(b) = F(a) = 0$，所以 Rolle 中值定理可用。这是一个技术性的证明，想法和 Lagrange 中值定理的证明如出一辙。（很多习题册上的题目都是根据这个想法来编的，作业中我们会见到几个例子（比较有技巧性，但其实没有多大意思））。

2. 这个证明比第一个证明复杂很多，然而在概念的层次上要更清晰也包含了更多的理解。

首先，我们观察到如果 $g(x) = x$，那么这个定理就是 Lagrange 中值定理。证明的主旨是将一般的 g 的情况化为这个已知情况，只需要微分几何中一些最朴素的想法。

利用 $g'(x) \neq 0$，我们知道 $g : I = [a, b] \to J = [g(a), g(b)]$ 是同胚，即连续的双射并且逆映射也是连续的。我们需要将 g 想成用 I 来重新参数化 $J = [g(a), g(b)]$。

注记. *比方说，给定 $\mathbb{R}$ 上的区间 $J = [0, 1]$，我们假设 $\mathbb{R}$ 上用的坐标是 y，那么，我们可以用 y 来表示（参数化）I 中的点。考虑映射*

$$g : [0, 2] \to [0, 1], \quad x \mapsto \frac{1}{2}.$$

此时，我们可以用通过 g 用 $x \in [0, 2]$ 来（参数化）描述 J 中的点：$x = 1$ 对应的是 J 中的 0.5 这个点。

我们现在认为 $J=[g(a),g(b)]$ 是基本的几何对象。本来 f 是 I 上的函数，我们可以通过上述参数化将 f（通过复合）视作是 $[g(a),g(b)]$ 上的函数：

$$\begin{array}{ccc} J & \xrightarrow{g^{-1}} & I \\ & \searrow^{f\circ g^{-1}} & \downarrow f \\ & & Y \end{array}$$

我们要考虑 $f(g^{-1})(y)=f\circ g^{-1}$ 这个函数。根据 Lagrange 中值定理（对 $f\circ g^{-1}:J\to\mathbb{R}$ 来用），我们有

$$\frac{f\circ g^{-1}(g(a))-f\circ g^{-1}(g(b))}{g(a)-g(b)}=\big(f\circ g^{-1}\big)'(c).$$

不难认出，这就是要证明的等式。这个证明表明，Cauchy 中值定理实际上是 Lagrange 中值定理在不同的参数化下的表述（从这个角度看，f 和 g 的位置不是对等的）。

□

另外，我们会问这个定理直观上说了什么（有没有比较容易记住的方式）？这里可以看出向量值函数的威力：考虑映射

$$F:[a,b]\to\mathbb{R}^2,\ x\mapsto\begin{pmatrix}f(x)\\ g(x)\end{pmatrix}.$$

这个映射的像是 $\mathbb{R}^2$ 上的一条曲线段。Cauchy 中值定理说的是存在曲线上的点 x_0，使得其切线方向 $F'(x_0)$ 和两个端点的连线 $\binom{f(b)}{g(b)}-\binom{f(a)}{g(a)}$ 是同方向的。在图上看起来这与 Rolle/Lagrange 中值定理的几何直观是一致的。事实上，我们还可以定义

$$\widehat{F}:[a,b]\to\mathbf{S}^1,\quad x\mapsto\frac{\big(f'(x),g'(x)\big)}{\sqrt{f'(x)^2+g'(x)^2}}.$$

这个是切线的方向的函数，由于 $\binom{f(b)}{g(b)}-\binom{f(a)}{g(a)}$ 也决定了 $\mathbf{S}^1$ 中的一个方向，这个定理的叙述变得类似于这种情况下的介值定理（$\mathbf{S}^1$ 是 1 维的）。此时，我们很自然地看到为什么需要研究在其它空间上取值的函数（映射）。

三角函数的研究

由于有了导数作为新的工具，我们可以来研究 $\sin x$ 和 $\cos x$ 的周期性了。三角函数是解析的方式来定义的，即用级数来定义：

$$\cos x=\sum_{k=0}^{\infty}\frac{(-1)^kx^{2k}}{(2k)!},\quad \sin x=\sum_{k=0}^{\infty}\frac{(-1)^kx^{2k+1}}{(2k+1)!}.$$

它们具有周期性是一个很深刻的事实。我们令

$$F:\mathbb{R}\to\mathbb{R}^2,\quad x\mapsto F(x)=\begin{pmatrix}\sin x\\ \cos x\end{pmatrix}.$$

用 J 表示矩阵 $J=\begin{pmatrix}0&1\\-1&0\end{pmatrix}$。那么，根据 $(\cos x)'=-\sin x$ 和 $(\sin x)'=\cos x$，我们发现 F 满足如下的微分方程

$$\begin{cases} F'=JF,\\ f\big|_{x=0}=\begin{pmatrix}0\\1\end{pmatrix}. \end{cases}$$

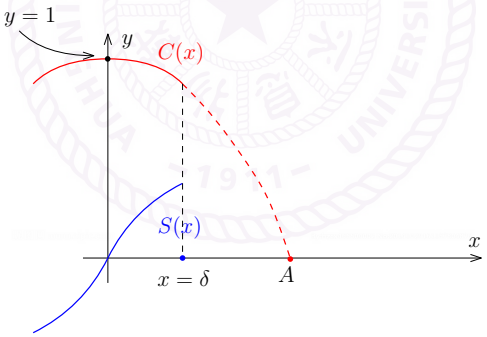
我们还把 $F(x)$ 写成

$$F(x)=\big(S(x),C(x)\big),\quad S(x)=\sin(x),\quad C(x)=\cos(x).$$

注记. 如果你采取其它方式定义 $\sin x$ 和 $\cos x$ 的话，比如说按照中学的方式定义的 $\sin x$ 和 $\cos x$，你应该也能证明 $\sin x$ 和 $\cos x$ 满足上述的关系，根据解的唯一性，我们就知道这两种定义方式是一致的。

三角函数的解析表达式没有给出它们更多的信息，我们利用函数的微分来研究 $\cos x$ 和 $\sin x$ 在 $\mathbb{R}_{\geqslant 0}$ 上的行为：

由于 $S'(0)=C(0)=1$，根据连续性，S 在某个 $[-\delta,\delta]$ 上面是严格递增的（δ 是较小的正数）。在区间 $(0,\delta)$ 上，由于 $C'(x)=-S(x)<0$，从而，$C(x)$ 是严格递减。；在区间 $(-\delta,0)$ 上，由于 $S(x)<0$，从而 $C(x)$ 是严格递增的。这表明 0 是 $C(x)$ 的一个局部最大点（这不意外，因为 $|C(x)|\leqslant 1$ 且 $C(0)=1$）。这一段推导可以用如下的图像来表示：



我们下面要证明说一定能找到一个 $A>0$，使得在 $[0,A]$ 上，$C(x)$ 是单调递减的并且在 A 处 $C(x)$ 变成了 0。考虑使得 C 在 $[0,A)$ 上都 >0 的最大可能的 A。

我们先说明 $A\neq+\infty$。如若不然，如果 $S'=C>0$，所以 S 是 $(0,+\infty)$ 是严格递增的函数。在 $x=\delta$，我们令 $s=S(\delta)>0$，从而，当 $x\geqslant\delta$ 时，$S(x)\geqslant s>0$。我们考虑如下的代数变形（此时，如果有积分的语言的话会更自然）：

$$S(x)\geqslant s\ \Leftrightarrow \big(C(x)+sx\big)'<0.$$

这是把一个表达式写成全微分的技巧。这表明函数 $C(x)+sx$ 在 $[\delta,\infty)$ 上递减，所以，对任意的 $x\geqslant\delta$，由于我们假设了 $C(x)>0$，所以我们有

$$C(\delta)+s\delta\geqslant C(x)+sx>sx.$$

令 $x \to +\infty$，这自然是不可能的!

此时，上面所描述的 $A < \infty$ 是存在。根据连续性，在 A 处，有 $C(A) = 0$。根据 $C\big|_{[0,A]} \geqslant 0$，我们知道 S 在 $[0, A]$ 上上升，所以 $S(A) > 0$。我们现在计算 $S(A)$ 的值。为此，我们对任意满足上述微分方程的 $S(x)$ 和 $C(x)$，我们都有 $\big(S(x)\big)^2 + \big(C(x)\big)^2 \equiv 1$。证明是利用导数判别：一个函数是常数当且仅当它的导数恒为零。

$$\Big(\big(S(x)\big)^2 + \big(C(x)\big)^2\Big)' = 2S(x)S'(x) + 2C(x)C'(x) = 0.$$

而 $\Big(\big(S(x)\big)^2 + \big(C(x)\big)^2\Big)\Big|_{x=0} = 1$。据此，我们知道 $S(A) = 1$。

作为上面推导的总结，我们知道存在常数 A，使得 $S(0) = 1$，$S(A) = 1$ 并且 S 在 $[0, A]$ 上严格递增；$C(0) = 1$，$C(A) = 0$ 并且 C 在 $[0, A]$ 上严格递减。

重复上面的推导（请同学自己补充细节），我可以找到 $B > 0$，使得 $S(A) = 1$，$S(A + B) = 0$ 并且 f 在 $[A, A + B]$ 上严格递减；$C(A) = 0$，$C(A + B) = -1$ 并且 g 在 $[0, A]$ 上严格递减。

我们定义 $\varpi = A + B$，按照上面的推导，它是 $S(x) = \sin x$ 在正实轴上第一个零点。

此时，考虑函数 $\begin{pmatrix}\overline{S}(x)\\ \overline{C}(x)\end{pmatrix} = -\begin{pmatrix}S(x+\pi)\\ C(x+\pi)\end{pmatrix}$。很明显，$\begin{pmatrix}\overline{S}\\ \overline{C}\end{pmatrix}$ 和 $\begin{pmatrix}S\\ C\end{pmatrix}$ 解一样的微分方程（同样的初值），所以根据唯一性，我们得到

$$S(x) = \overline{S}(x) = -S(x + \varpi), \quad C(x) = \overline{C}(x) = -C(x + \varpi).$$

从而，

$$S(x) = S(x + 2\varpi), \quad C(x) = C(x + 2\varpi).$$

这说明 2ϖ 是 S 和 C 的周期。另外，S 在 $[0.\varpi]$ 上面为正，在 $[\varpi, 2\varpi]$ 上面为负表明 2ϖ 是 S 的最小周期；类似地，2ϖ 是 C 的最小周期。

15.1 作业：高木贞治函数

清华大学 19-20 秋季学期，数学分析一，作业 6

请用 A4 大小的纸张**正反面**用钢笔，签字笔或者圆珠笔书写，并注明自己的姓名，年级和作业的**总页数**。除定理公式所涉及的人名之外，请使用**中文**。本次作业的提交时间和地点为 **10 月 14 日**上午的课堂上，逾期视作零分。

基本习题

习题 A：导数的定义和计算

A1) $\mathbb{R}^n$ 上配有范数 $\|(x_1,\cdots,x_n)\|_2=\sqrt{(x_1)^2+\cdots+(x_n)^2}$，我们考虑映射

$$f:\mathbb{R}\to\mathbb{R}^n,\quad x\mapsto f(x)=(f_1(x),\cdots,f_n(x)).$$

证明，f 在 x_0 处的导数存在当且仅当对每个分量函数 f_k，它在 x_0 处的导数都存在并且

$$f'(x)=\big(f_1'(x),\cdots,f_n'(x)\big).$$

A2) 考虑函数 e^{ix}，将它视作是映射

$$f:\mathbb{R}\to\mathbb{C},\quad x\mapsto e^{ix}.$$

利用定义证明，$f'(0)=i\ (\in\mathbb{C})$。仿照课堂上的做法证明 $\big(e^{ix}\big)'=ie^{ix}$。

A3) 利用上面两个问题的结论计算 $\sin x$ 和 $\cos x$ 的导数。

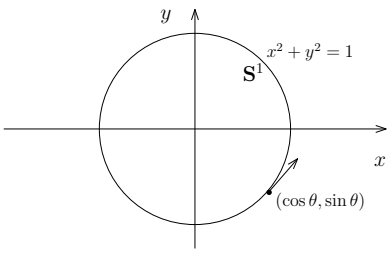
A4) 对 $n=3$ 来验证 Faà di Bruno 公式。

A5) 定义映射

$$E:\mathbb{R}\to\mathbb{C}=\mathbb{R}^2,\quad \theta\mapsto(\cos\theta,\sin\theta).$$

我们用下面符号表示平面上的单位圆：

$$\mathbf{S}^1=\big\{(x,y)\in\mathbb{R}^2\big|x^2+y^2=1\big\}.$$



证明，$\mathbf{S}^1$ 上的点可以写成 $(\sin\theta, \cos\theta)$ 的形式，即 $E(\mathbb{R}) = \mathbf{S}^1$。试计算 $E'(\theta)$ 并验证对于这个映射（$\mathbb{R}^2$ 上取值），Rolle 中值定理并不成立。

A6) 求下列函数的导数 (如果 f 在 x_0 不可微，但单侧左导数或右导数存在时，求出左导数或右导数)

(1) $f(x) = a^x, \quad a > 0.$

(2) $f(x) = \arcsin x.$

(3) $f(x) = \arctan x$

(4) $f(x) = x^{x^x}, \quad x > 0$

(5) $f(x) = \log\Big(\log\big(\log x\big)\Big).$

(6) $f(x) = \sqrt{x + \sqrt{x + \sqrt{x}}}.$

(7) $f(x) = |x|.$

(8) $f(x) = \log|x|.$

(9) $f(x) = \begin{cases} x^n \sin\frac{1}{x}, & x \neq 0, \\ 0, & x = 0. \end{cases}$ $\qquad n = 1, 2, \cdots.$

A7) 求下列函数的 3-阶导数：

(1) $f(x) = \log(1 + x).$

(2) $f(x) = x^{-1}\log x.$

(3) $f(x) = \dfrac{x^m}{1 - x}, \ m \in \mathbb{Z}_{\geqslant 0}.$

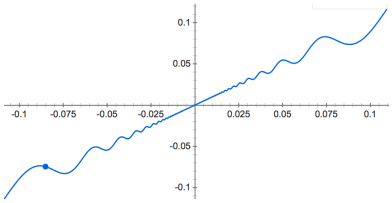
(4) $f(x) = x^m e^x, \ m \in \mathbb{Z}_{\geqslant 0}.$

(5) $f(x) = e^{ax}\sin(bx), \ a, b \in \mathbb{R}.$

(6) $f(x) = e^{-x^2}.$

A8) 如果函数 f 在点 x_0 处的导数 > 0，不能推出存在该点的领域 U，使得 f 在这个邻域上是递增的：考虑函数

$$f(x) = \begin{cases} x + 2x^2 \sin\left(\frac{1}{x}\right), & x \neq 0; \\ 0, & x = 0. \end{cases}$$



证明，f 在 0 处导数存在且大于零，但是对任意的 $\varepsilon > 0$，f 在 $(-\varepsilon, \varepsilon)$ 上的限制都不是单调函数。

A9) （重要）$\mathbf{I}_n$ 是 $n \times n$ 的单位矩阵，$A \in \mathbf{M}_n(\mathbb{R})$，计算 $\dfrac{d}{dx}\Big|_{x=0} \det(\mathbf{I} + xA)$，即 $\det(\mathbf{I} + xA)$ 在 $x = 0$ 处的导数。

A10) 证明，（可微的）奇函数的导数是偶函数，（可微的）偶函数的导数是奇函数。

A11) 证明，Riemann 函数 $f(x)=\begin{cases}1/q, & x=\dfrac{p}{q}\in\mathbb{Q}, q\geqslant 1, p \text{ 和 } q \text{ 互素};\\ 0, & x \text{ 是无理数.}\end{cases}$ 在 $\mathbb{R}$ 上处处不可微.

习题 B

B1) 定义双曲函数:

$$\sinh x=\frac{e^x-e^{-x}}{2},\quad \cosh x=\frac{e^x+e^{-x}}{2},\quad \tanh x=\frac{\sinh x}{\cosh x}.$$

(a) 证明，

(1) $\cosh^2 x-\sinh^2 x=1$.

(2) $\sinh(x+y)=\sinh x\cosh y+\cosh x\sinh y$.

(3) $\cosh(x+y)=\cosh x\cosh y+\sinh x\sinh y$.

(4) $\tanh(x+y)=\dfrac{\tanh x+\tanh y}{1+\tanh x\tanh y}$.

(b) 求导数 $\sinh'(x)$，$\cosh'(x)$ 和 $\tanh'(x)$。

(c) 证明，存在 $\sinh:\mathbb{R}\to\mathbb{R}$ 的反函数 $\operatorname{arcsinh}:\mathbb{R}\to\mathbb{R}$ 并计算 $\operatorname{arcsinh}'(x)$。

B2) 设 $a,b\in\mathbb{R}$, $b>0$, 考察函数 $f:[-1,1]\to\mathbb{R}$，其中

$$f(x)=\begin{cases}x^a\sin(x^{-b}), & \text{如果} x\neq 0,\\ 0, & \text{如果} x=0.\end{cases}$$

证明，下述关于 f 的结论都成立:

(a) $f\in C([-1,1])$ 当且仅当 $a>0$;

(b) f 在 0 处可微当且仅当 $a>1$;

(c) f' 在 $[-1,1]$ 上有界当且仅当 $a\geqslant 1+b$;

(d) $f\in C^1([-1,1])$ 当且仅当 $a>1+b$;

(e) f' 在 0 处可微当且仅当 $a>2+b$;

(f) f'' 在 $[-1,1]$ 上有界当且仅当 $a\geqslant 2+2b$;

(g) $f\in C^2([-1,1])$ 上有界当且仅当 $a>2+2b$。

习题 C（无穷小量与无穷大量的阶的比较）

如果函数 f 在 x_0 的附近（即某个 x_0 的开邻域去掉 x_0）满足 $\lim\limits_{x\to x_0}f(x)=0$，我们就称 f 是 $x\to x_0$ 时的**无穷小量**；类似地，如果 $\lim\limits_{x\to x_0}f(x)=+\infty$ 或者 $\lim\limits_{x\to x_0}f(x)=-\infty$（其注意我们用的“或者”这个词的含义），我们就称 f 是 $x\to x_0$ 时的**无穷大量**。

现在假设 f,g 都是 $x\to x_0$ 时的无穷小量并且 $g(x)$ 在 x_0 的附近不取零值，我们现在引进记号:

- 如果 $\lim\limits_{x\to x_0}\dfrac{f(x)}{g(x)}=0$，我们就称 f 是比 g 高阶的无穷小，记作 $f(x)=o(g(x)),x\to x_0$；
- 如果 $\lim\limits_{x\to x_0}\dfrac{f(x)}{g(x)}=\ell,\ell\neq 0$，我们就称 f 是与 g 同阶的无穷小；
- 特别地，如果 f 与 g 同阶并且 $\ell=1$，我们就称 f 是与 g 等价的无穷小，记作 $f(x)\sim g(x),x\to x_0$；
- 如果 $\limsup\limits_{x\to x_0}\dfrac{|f(x)|}{|g(x)|}<+\infty$，我们将这种情况记作 $f(x)=O(g(x)),x\to x_0$。

类似地，我们可以定义无穷大量的阶之间的比较。这是通用的术语，同学们可参考任一本参考书或者网络。

C1) 假设 $x\to x_0$ 时，函数 $a(x)$ 满足 $a=o(1)$。试证明：

(1) $o(a)+o(a)=o(a)$　　(2) $o(ca)=co(a),\quad c\in\mathbb{R}$

(3) $(o(a)^k)=o(a^k)$,　　(4) $\dfrac{1}{1+a}=1-a+o(a)$.

C2) 假设 $f(x),g(x)$ 是 $x\to x_0$ 时的无穷小，那么

(a) 证明，$f(x)\sim g(x)\iff f(x)-g(x)=o(g(x)),\quad x\to x_0$。

(b) 如果把 g 作为基本的比较单位（小量），我们可以将另外的无穷小量与 $g(x)^k$（k 是正整数）进行比较。如果 $f(x)\sim cg(x)^k$，我们就称 cg^k 是 f 的主部（这里 c 是常数）。试定出下列无穷小或无穷大的主部（与 $x-x_0$ 或者 x 比较）：

(1) $\dfrac{1}{\sin\pi x},\quad x\to 1.$　　(2) $\sqrt{1+x}-\sqrt{1-x},\quad x\to 0.$

(3) $\sin\left(\sqrt{1+\sqrt{1+\sqrt{x}}-\sqrt{2}}\right),\quad x\to 0^+.$　(4) $\sqrt{1+\tan x}-\sqrt{1-\sin x},\quad x\to 0.$

(5) $\sqrt{x+\sqrt{x+\sqrt{x}}},\quad x\to 0^+.$　　(6)$\sqrt{x+\sqrt{x+\sqrt{x}}},\quad x\to+\infty.$

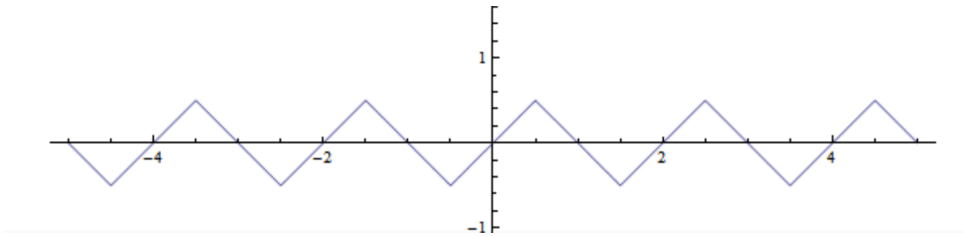
(c) 我们假设 $f(x)\sim cx^k,x\to 0$（即 $f(x)=cx^k+o(x^k)$）。如果 $f(x)-cx^k$ 可以再分出主部 $c'x^{k'}$，其中 $k'>k$，那么我们就将它写为 $f(x)=cx^k+c'x^{k'}+o(x^{k'})$。试将下列无穷小展开到 $o(x^2)$：

(1) $\sqrt{1+x}-1,$　　(2) $\sqrt[m]{1+x}-1,m\in\mathbb{Z}_{\geqslant 1}.$

思考题 (不交作业)：处处连续处处不可微的 Takagi(高木贞治) 函数

习题 T（Takagi(高木贞治) 函数，1903 年）我们先在 $[0,1]$ 区间上定义 $\psi(x)=\begin{cases}x, & 0\leqslant x<\dfrac{1}{2};\\ 1-x, & \dfrac{1}{2}\leqslant x\leqslant 1.\end{cases}$

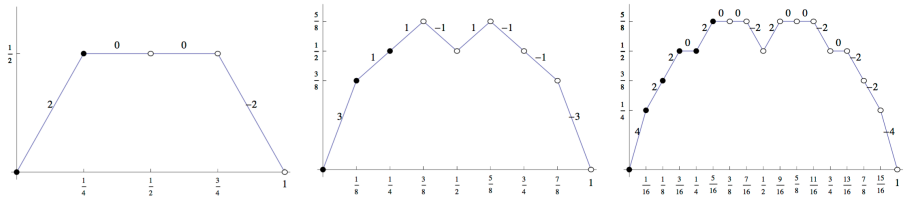
接下来，以 1 为周期，我们可以将 ψ 延拓成 $\mathbb{R}$ 上的周期函数（连续）并且仍然将它记作 ψ，它的函数图像好像是锯齿一般：



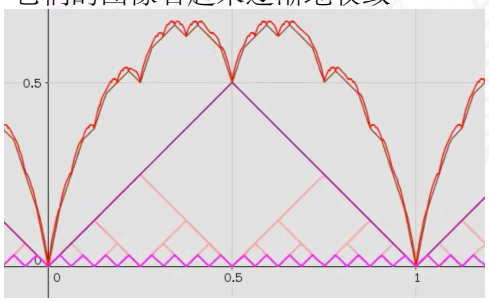
我们定义 Takagi 函数 $T:\mathbb{R}\to\mathbb{R}$ 如下：

$$T(x)=\sum_{k=0}^{\infty}\frac{1}{2^k}\psi(2^k x).$$

我们实际上可以考虑部分和 $T_n(x)=\sum_{k=0}^{n}\frac{1}{2^k}\psi(2^k x)$。当 $n=1,2,3$ 时，它们的图像如下：

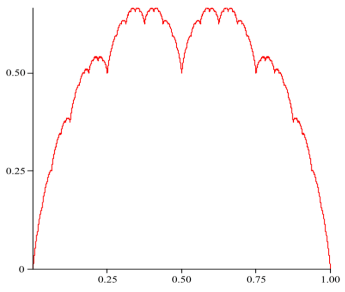


当 n 越来越大的时候，它们的图像看起来逐渐地收敛：



这个习题的目标是粗略地研究 Takagi 函数的性质。

T1) 证明，$T(x)$ 是 $\mathbb{R}$ 上良好定义的有界连续函数。它图像貌似：



T2) 对于 $x\in[0,1]$，假设 $x=\sum_{n=1}^{\infty}\frac{a_n}{2^n}$ 是 x 的 2-进制展开，其中 $a_n=0$ 或者 1。我们令 $v_n=\sum_{k=1}^{n}a_k$。

我们定义函数 $\sigma_n(y) = a_n + (1-2a_n)y$，其中 $y = 0$ 或者 1。证明，

$$\psi(2^m x) = \sum_{n=1}^{\infty} \frac{\sigma_{m+1}(a_{m+n})}{2^n}.$$

T3) 对于 $x \in [0,1]$，假设 $x = \sum_{n=1}^{\infty} \frac{a_n}{2^n}$ 是 x 的 2-进制展开。证明，

$$T(x) = \sum_{n=1}^{\infty} \frac{(1-a_n)v_n + a_n(n-v_n)}{2^n}.$$

T4) 假设 $x_0 = \frac{k_0}{2^{m_0}} \in (0.1]$，其中 $k_0 \in \mathbb{Z}_{\geqslant 1}$ 是奇数，$m_0 \in \mathbb{Z}_{\geqslant 0}$。令 $h_n = \frac{1}{2^n}$，其中 $n \in \mathbb{Z}_{\geqslant m_0}$。证明，数列 $\left\{\frac{T(x+h_n)-T(x)}{h_n}\right\}_{n \geqslant m_0}$ 不收敛。

T5) f 是定义在非空的开区间 I 上实数值函数。如果 f 在 a 处可导，证明，

$$\lim_{\substack{(h,h')\to(0,0),\\ h>0,h'>0}} \frac{f(a+h)-f(a-h')}{h+h'} = f'(a),$$

这里极限 $\lim\limits_{\substack{(h,h')\to(0,0),\\ h>0,h'>0}}$ 的意义指的是任意的序列 $(h_n, h'_n) \to (0,0), h_n > 0, h'_n > 0$ 所对应的序列都收敛。

T6) f 是定义在非空的开区间 I 上实数值函数。假设 $f \in C^1(I)$（连续可微），$a \in I$，证明，

$$\lim_{\substack{(h,h')\to(0,0),\\ h+h'\neq 0}} \frac{f(a+h)-f(a-h')}{h+h'} = f'(a).$$

T7) 假设 $x \in [0,1]$，使得对任意的正整数 n，$2^n x$ 都不是整数。对于每个正整数 n，我们用下面的公式定义序列 $\{h_n\}_{n\geqslant 1}$ 和 $\{h'_n\}_{n\geqslant 1}$：

$$\lfloor 2^n x \rfloor = 2^n(x - h'_n), \quad \lfloor 2^n x \rfloor + 1 = 2^n(x + h_n),$$

其中函数 $\lfloor y \rfloor$ 按照定义是不超过 y 的最大的整数（即 y 的整数部分（如果 $y \geqslant 0$））。证明，对每个给定的 n，$h_n + h'_n = 2^{-n}$ 并且对每个整数 $1 \leqslant \ell \leqslant n-1$，开区间 $\left(2^\ell(x-h'_n), 2^\ell(x+h_n)\right)$ 中不包含任何的整数和半整数。

T8) 假设 $x \in [0,1]$，使得对任意的正整数 n，$2^n x$ 都不是整数，我们沿用 E7) 中的符号，证明，数列 $\left\{\frac{T(x+h_n)-T(x-h'_n)}{h_n+h'_n}\right\}_{n\geqslant 1}$ 不收敛。

T9) 证明，$T(x)$ 是 $\mathbb{R}$ 上处处连续处处不可微的函数。

T10) 证明，我们有如下的函数方程：

$$T(x)=\begin{cases}2x+\dfrac{T(4x)}{4}, & 0\leqslant x<\dfrac{1}{4};\\ \dfrac{1}{2}+\dfrac{T(4x-1)}{4}, & \dfrac{1}{4}\leqslant x<\dfrac{1}{2};\\ \dfrac{1}{2}+\dfrac{T(4x-2)}{4}, & \dfrac{1}{2}\leqslant x<\dfrac{3}{4};\\ 2-2x+\dfrac{T(4x-3)}{4}, & \dfrac{3}{4}\leqslant x\leqslant 1.\end{cases}$$

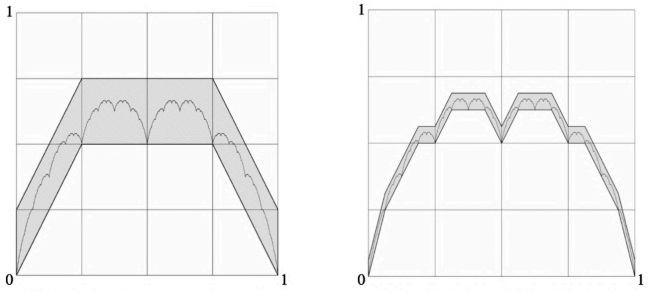
T11) （Takagi 函数图像的自相似性）令 $\Gamma=\big\{(x,T(x))|0\leqslant x\leqslant 1\big\}\subset\mathbb{R}^2$ 是 T 在区间 $[0,1]$ 上的函数图像。我们定义如下四个仿射变换 $\Phi_i:\mathbb{R}^2\to\mathbb{R}^2$：

$$\Phi_0\begin{pmatrix}x\\y\end{pmatrix}=\begin{pmatrix}\frac{1}{4}&0\\\frac{1}{2}&\frac{1}{4}\end{pmatrix}\begin{pmatrix}x\\y\end{pmatrix},\qquad \Phi_1\begin{pmatrix}x\\y\end{pmatrix}=\begin{pmatrix}\frac{1}{4}&0\\0&\frac{1}{4}\end{pmatrix}\begin{pmatrix}x\\y\end{pmatrix}+\begin{pmatrix}\frac{1}{4}\\\frac{1}{2}\end{pmatrix},$$

$$\Phi_2\begin{pmatrix}x\\y\end{pmatrix}=\begin{pmatrix}\frac{1}{4}&0\\0&\frac{1}{4}\end{pmatrix}\begin{pmatrix}x\\y\end{pmatrix}+\begin{pmatrix}\frac{1}{2}\\\frac{1}{2}\end{pmatrix},\quad \Phi_3\begin{pmatrix}x\\y\end{pmatrix}=\begin{pmatrix}\frac{1}{4}&0\\-\frac{1}{2}&\frac{1}{4}\end{pmatrix}\begin{pmatrix}x\\y\end{pmatrix}+\begin{pmatrix}\frac{3}{4}\\\frac{1}{2}\end{pmatrix}.$$

证明，Φ_i 恰好把 Γ 变成 T 在区间 $[\dfrac{i}{4},\dfrac{i+1}{4}]$ 上的图像，其中 $i=0,1,2,3$。

T12) 令 $S_0=\{(x,y)\in\mathbb{R}^2|0\leqslant x\leqslant 1,0\leqslant y\leqslant 1\}$ 是 $\mathbb{R}^2$ 上的闭方块。对每个 $n\geqslant 0$，我们定义 $S_{n+1}=\bigcup\limits_{k=0}^{3}\Phi_k(S_n)$。证明，$S_n$ 是平面上一列单调下降的紧集并且 $\Gamma=\bigcap\limits_{n\geqslant 0}S_n$。我们有 S_1 和 S_2 的示意图：



T13) 证明，$\sup\limits_{x\in\mathbb{R}}T(x)\leqslant\dfrac{2}{3}$。

T14) 找一个 $c\in[0,1]$，使得 $T(c)=\dfrac{2}{3}$。

T15) （$T^{-1}(\dfrac{2}{3})$ 的 Cantor 集的结构）对于 $x\in[0,1]$，假设 $x=\sum\limits_{n=1}^{\infty}\dfrac{b_n}{4^n}$ 是 x 的 4-进制展开，其中 $b_n=0,1,2,3$。证明，

$$\Big\{x\in[0,1]\Big|T(x)=\frac{2}{3}\Big\}=\Big\{x\in[0,1]\Big|x=\sum_{n=1}^{\infty}\frac{b_n}{4^n},b_n\in\{1,2\}\Big\}.$$

T16) 仿照 T11)，研究 Φ_1 和 Φ_2 在集合 $\left\{(x,T(x))\middle| x\in[0,1], T(x)=\frac{2}{3}\right\}$ 上自相似的作用。（这是一个 Hausdorff 维数为 $\frac{1}{2}$ 的集合）。

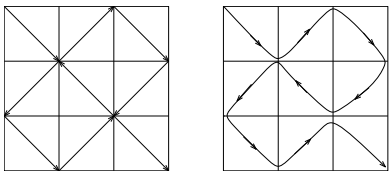
16　空间填充曲线，L'Hôpital 法则，Taylor 展开

二零一九年十一月十一日，星期一，晴天

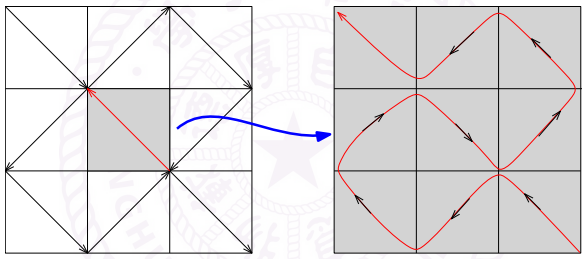
空间填充曲线

利用一致收敛的想法，我们可以构造一个很有趣的（很重要的反例）例子：令 $I=\left\{x\in\mathbb{R}\middle|0\leqslant x\leqslant 1\right\}$，$C=\left\{(x,y)\in\mathbb{R}^2\middle|0\leqslant x\leqslant 1,0\leqslant y\leqslant 1\right\}$，我们按照如下的图像来定义映射

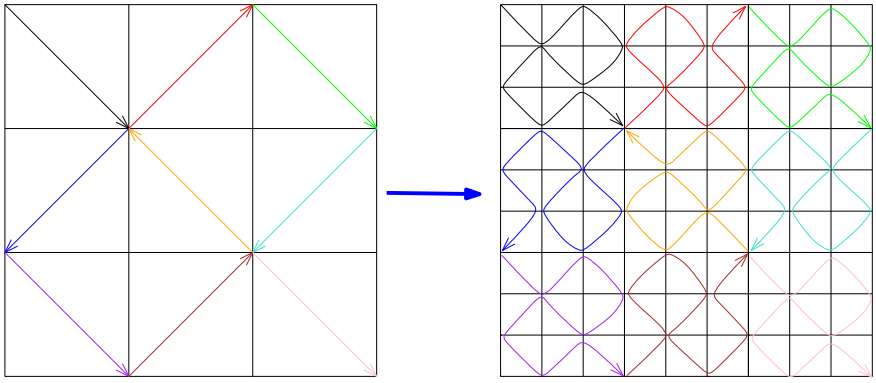
$$f_1:I\to C,$$



我们对每一个小的正方形根据箭头的方向构造同样的映射，以中间这个灰色的为例子，我们要把刚才第一个图中的格子转 180°，这样得到下面的图像：



其中，我们把 I 这个线段映射为图中所画的 9 段折线，折线走的方向如第二个图所示。这样子，我们把 C 分成了 81 个格子，并且构造出一条 81 段的折线：



上面这个图对应着映射：

$$f_2:I\to C,$$

重复上面的操作，我们就得到了一串连续映射：

$$f_n:I\to C,$$

其中 $n \geqslant 1$，并且这是一个 9^n 段的折线。按照这些映射的构造方式，我们知道，对任意的 $N \geqslant 1$，对任意的 $n, m \geqslant N$，f_n 和 f_m 都是对 f_N 的每个边长为 3^{-N} 的方块里面进行修改，特别地，对任意的 $x \in I$，我们就知道 $f_n(x)$ 和 $f_m(x)$ 都落在同一个边长为 3^{-N} 的方块里，所以

$$|f_n(x) - f_m(x)| \leqslant \sqrt{2} \cdot 3^{-N},$$

即

$$\|f_n - f_m(x)\|_{L^\infty} \leqslant \sqrt{2} \cdot 3^{-N},$$

这表明 $\{f_n\}_{n\geqslant 1}$ 是 $\left(C(I;\mathbb{R}^2), \|\cdot\|_\infty\right)$ 中的 Cauchy 列，从而存在连续映射

$$f_\infty : I \to C,$$

使得 $f_n \overset{C(I;\mathbb{R}^2)}{\longrightarrow} f_\infty$。根据 f_n 的构造，对任意的 $p \in C$，p 一定落在 f_n 所对应的 9^n 个小方块中的某一个，所以有 $x \in I$，使得 p 和 $f(x)$ 的距离不超过 $\sqrt{2}\cdot 3^{-n}$，据此，我们知道 f_∞ 的像 $f_\infty(I) \subset C$ 是稠密的。由于 I 是紧集，所以它在连续映射下的像是紧的，从而是闭的，再用稠密性，我们就知道 $f_\infty(I) = C$。最终，我们得到连续的满射：

$$f_\infty : I \to C.$$

我们回到导数的学习，上次课证明了 Cauchy 中值定理：实值函数 $f, g \in C^0([a,b])$ 并且 f 和 g 均在 (a,b) 上可微，若对任意的 $x \in (a,b)$，$g'(x) \neq 0$。那么，存在 $x_0 \in (a,b)$，使得

$$\frac{f'(x_0)}{g'(x_0)} = \frac{f(b) - f(a)}{g(b) - g(a)}.$$

Cauchy 中值定理的重要应用是用来证明 L'Hôpital 法则：

命题 99 (L'Hôpital 法则)**.** *假设 f 和 g 是区间 (a,b) 上的可微实值函数，我们假设即 $f(x), g(x) = o(x-a)$，即*

$$\lim_{x\to a^+} f(x) = 0, \ \lim_{x\to a^+} g(x) = 0.$$

我们假设对任意的 $x \in (a,b)$，$g'(x) \neq 0$。如果极限 $\lim\limits_{x\to a^+} \dfrac{f'(x)}{g'(x)}$ 存在（可以是 $\pm\infty$），那么

$$\lim_{x\to a^+} \frac{f(x)}{g(x)} = \lim_{x\to a^+} \frac{f'(x)}{g'(x)}.$$

证明： 根据 $f(x), g(x) = o(x-a)$，我们知道对任意的 $x \in (a,b)$，f 和 g 是在区间 $[a,x]$ 上连续并且在 (a,x) 上可微。所以，利用 Cauchy 中值定理，存在 $\xi(x) \in (a,x)$，使得

$$\frac{f(x)}{g(x)} = \frac{f(x) - f(a)}{g(x) - g(a)} = \frac{f'(\xi(x))}{g'(\xi(x))}.$$

特别地，当 $x \to a^+$ 时，由于 $a < \xi(x) < x$，所以 $\xi(x) \to a^+$。对上面的式子取极限，我们就得到：

$$\lim_{x\to a^+} \frac{f(x)}{g(x)} = \frac{f(x) - f(a)}{g(x) - g(a)} = \lim_{x\to a^+} \frac{f'(\xi(x))}{g'(\xi(x))} = \lim_{x\to a^+} \frac{f'(x)}{g'(x)}.$$

命题得证。 □

我们还有一个版本 L'Hôpital 法则：

推论 100 (L'Hôpital 法则)**.** *实值函数 f 和 g 在区间 $(a,+\infty)$ 上可微并且*

$$\lim_{x\to+\infty} f(x)=0,\ \lim_{x\to+\infty} g(x)=0.$$

假设对任意的 $x\in(a,+\infty)$，$g'(x)\neq 0$。如果极限 $\lim\limits_{x\to+\infty}\dfrac{f'(x)}{g'(x)}$ 存在（可以是 $\pm\infty$），那么我们有

$$\lim_{x\to+\infty}\frac{f(x)}{g(x)}=\lim_{x\to+\infty}\frac{f'(x)}{g'(x)}.$$

证明： 不妨假设 $a>0$，考虑坐标变换：

$$\varphi:(0,\frac{1}{a})\to(a,\infty),\quad x\mapsto\frac{1}{x}.$$

从而，$\widetilde{f}=f\circ\varphi$ 是 $(0,\frac{1}{a})$ 上的函数，即 $\widetilde{f}(x)=f\left(\dfrac{1}{x}\right)$；类似地，我们定义 $(0,\frac{1}{a})$ 上的函数 $\widetilde{g}=g\circ\varphi$ 是。由于 $x\to+\infty$ 等价于 $\varphi(x)\to 0^+$，所以

$$\lim_{x\to0^+}\widetilde{f}(x)=0,\ \lim_{x\to0^+}\widetilde{g}(x)=0.$$

此时，我们有

$$\lim_{x\to0^+}\frac{\widetilde{f}'(x)}{\widetilde{g}'(x)}=\lim_{x\to+\infty}\frac{-\frac{f'(x)}{x^2}}{-\frac{g'(x)}{x^2}}=\lim_{x\to+\infty}\frac{f'(x)}{g'(x)}.$$

所以，根据前一版本的 L'Hôpital 法则，我们有

$$\lim_{x\to+\infty}\frac{f(x)}{g(x)}=\lim_{x\to0^+}\frac{\widetilde{f}(x)}{\widetilde{g}(x)}=\lim_{x\to+\infty}\frac{f'(x)}{g'(x)}.$$

这就证明了命题。 □

推论 101. *$n\geqslant 1$ 是整数，f 和 g 是区间 (a,b) 上 n-次可微的实值函数。假设对任意的 $0\leqslant k\leqslant n-1$，都有*

$$\lim_{x\to a^+} f^{(k)}(x)=0,\quad \lim_{x\to a^+} g^{(k)}(x)=0,$$

并且极限 $\lim\limits_{x\to a^+}\dfrac{f^{(n)}(x)}{g^{(n)}(x)}$ 存在（可以是 $\pm\infty$）。如果对 $x\in(a,b)$，$g^{(n)}(x)\neq 0$，那么

$$\lim_{x\to a^+}\frac{f(x)}{g(x)}=\lim_{x\to a^+}\frac{f^{(n)}(x)}{g^{(n)}(x)}.$$

证明： 我们 n 用归纳法就立即得到了证明，但是每次都要检验导数不为零的条件：由于 $f^{(n)}(x)\neq 0$，根据 Darboux 介值定理，$f^{(n)}(x)$ 恒正或者恒负，所以函数严格单调。由于 $f^{(n-1)}(a)=0$，所以 $f^{(n-1)}(x)$ 恒为正或者恒为负。以此类推，所有导数都非零。 □

我们还有其它两个类型的 L'Hôpital 法则：

推论 102 (L'Hôpital 法则)**.** 假设 f 和 g 是区间 (a,b) 上的可微实值函数，我们假设

$$\lim_{x\to a^+}|f(x)|=\infty,\ \lim_{x\to a^+}|g(x)|=\infty.$$

我们假设对任意的 $x\in(a,b)$，$g'(x)\neq 0$。如果极限 $\lim\limits_{x\to a^+}\dfrac{f'(x)}{g'(x)}$ 存在（可以是 $\pm\infty$），那么

$$\lim_{x\to a^+}\frac{f(x)}{g(x)}=\lim_{x\to a^+}\frac{f'(x)}{g'(x)}.$$

推论 103 (L'Hôpital 法则)**.** 实值函数 f 和 g 在区间 $(a,+\infty)$ 上可微并且

$$\lim_{x\to+\infty}|f(x)|=0,\ \lim_{x\to+\infty}|g(x)|=0.$$

假设对任意的 $x\in(a,+\infty)$，$g'(x)\neq 0$。如果极限 $\lim\limits_{x\to+\infty}\dfrac{f'(x)}{g'(x)}$ 存在（可以是 $\pm\infty$），那么我们有

$$\lim_{x\to+\infty}\frac{f(x)}{g(x)}=\lim_{x\to+\infty}\frac{f'(x)}{g'(x)}.$$

这两个推论的证明我们留成本次作业。

L'Hôpital 法则可以用来计算极限：

例子. 我们举几个例子：

1) 计算 $\lim\limits_{x\to 0}\dfrac{\sin x}{x}$：

$$\lim_{x\to 0}\frac{\sin x}{x}=\lim_{x\to 0}\frac{\cos x}{1}=1.$$

2) 计算 $\lim\limits_{x\to 0}\dfrac{\cos x}{x^2}$：

$$\lim_{x\to 0}\frac{\cos x}{x^2}=\lim_{x\to 0}\frac{-2\cos x\sin x}{x}=\lim_{x\to 0}\frac{-2(\cos x)^2+2(\sin x)^2}{1}=-\frac{1}{2}.$$

3) 计算 $\lim\limits_{x\to 0}\dfrac{e^x-x-1}{x^2}$：

$$\lim_{x\to 0}\frac{e^x-x-1}{x^2}=\lim_{x\to 0}\frac{e^x-1}{2x}=\lim_{x\to 0}\frac{e^x}{2}=\frac{1}{2}.$$

4) 证明 $\lim\limits_{x\to\infty}\dfrac{x^n}{e^x}=0$：

$$\lim_{x\to\infty}\frac{x^n}{e^x}=\lim_{x\to 0}\frac{nx^{n-1}}{e^x}=\lim_{x\to\infty}\frac{n(n-1)x^{n-2}}{e^x}=\cdots=\lim_{x\to\infty}\frac{n!}{e^x}=0.$$

5) 计算 $\lim\limits_{x\to\infty}\dfrac{x}{\sqrt{x^2+1}}$。

第一次运用 *L'Hôpital* 法则，我们就会有

$$\lim_{x\to\infty}\frac{x}{\sqrt{x^2+1}}=\lim_{x\to\infty}\frac{1}{\frac{x}{\sqrt{x^2+1}}}=\lim_{x\to\infty}\frac{\sqrt{x^2+1}}{x}.$$

再用一次就得到

$$\lim_{x\to\infty}\frac{\sqrt{x^2+1}}{x}=\lim_{x\to\infty}\frac{x}{\sqrt{x^2+1}}.$$

所以一直不会停止。对于这个例子 *L'Hôpital* 法则并不好用。

6) 计算 $\lim\limits_{x\to0}\dfrac{\sin x}{e^x}$。如果我们不检验 $f(x)$ 和 $g(x)$ 在 ∞ 处是否是零而直接运用 *L'Hôpital* 法则，我们就会有

$$\lim_{x\to0}\frac{\sin x}{e^x}=\lim_{x\to0}\frac{\cos x}{e^x}=1.$$

这个结论自然是错误的！

L'Hôpital 法则只是一种计算极限的方法，它之所以有用（更多是做习题的时候）主要因为它可以把求极限这种分析上的操作转化为求导数的问题，而求导数的操作一般而言都是代数操作（因为我们可以**背过**很多导数）。然而，对于微积分的学习，这个法则似乎无关主旨，我们应该尽量早的学习 Taylor 展开的技术，这才是真正要紧的东西：

定理 104 (Taylor 展开公式：用多项式逼近)**.** *我们给出 Taylor 展开的三种不同余项的叙述：*

1) **Peano 余项.** 假设函数 $f:[a,b]\to\mathbb{R}$（或者 $\mathbb{C}$）在 a 处的一直到 n-次导数 $f'(a),\cdots,f^{(n)}(a)$ 都存在。那么，当 $x\to a^+$ 时，我们有

$$f(x)=\sum_{k=0}^{n}\frac{f^{(k)}(a)}{k!}(x-a)^k+o\big((x-a)^n\big).$$

2) **Lagrange 余项.** 假设函数 $f\in C^n\big([a,b]\big)$（在 $\mathbb{R}$ 或 $\mathbb{C}$ 中取值），特别地，f 在 a 处的 n-次导数 $f'(a)$，$\cdots$，$f^{(n)}(a)$ 都存在。如果 f 在 (a,b) 上 $n+1$ 次可导。那么，我们有

$$f(x)=\sum_{k=0}^{n}\frac{f^{(k)}(a)}{k!}(x-a)^k+R_n(x),$$

其中 $R_n(x)=\dfrac{f^{(n+1)}(\xi)}{(n+1)!}(x-a)^{n+1}$，$\xi\in[a,x]$ 由 x 决定（未必唯一）。

2) **Cauchy 余项.** 假设函数 $f\in C^n\big([a,b]\big)$（在 $\mathbb{R}$ 或 $\mathbb{C}$ 中取值），特别地，f 在 a 处的 n-次导数 $f'(a)$，$\cdots$，$f^{(n)}(a)$ 都存在。如果 f 在 (a,b) 上 $n+1$ 次可导。那么，我们有

$$f(x)=\sum_{k=0}^{n}\frac{f^{(k)}(a)}{k!}(x-a)^k+\overline{R}_n(x),$$

其中 $\overline{R}_n(x)=\dfrac{f^{(n+1)}(\xi)}{(n+1)!}(x-\xi)^n(x-a)$，$\xi\in[a,x]$ 由 x 决定（未必唯一）。

注记. *Peano* 余项的公式只在 a 的附近成立，而 *Lagrange* 和 *Taylor* 的情况是整体的公式。

我们注意到当 $n=1$ 时，*Peano* 余项的公式就是导数的定义。

当 $n=0$ 时，*Lagrange* 余项的公式就是 *Lagrange* 中值定理。

另外，如果要求 f 是 n-次连续可微的并且 $n+1$ 次导数存在，那么 *Lagrange* 余项（或者 *Taylor* 余项）的公式成立，此时，根据连续性，*Peano* 余项的公式明显成立。

证明: 1) Peano 余项公式等价于证明

$$\lim_{x\to a^+}\frac{f(x)-\sum_{k=0}^{n}\frac{f^{(k)}(a)}{k!}(x-a)^k}{(x-a)^n}=0.$$

我们利用 L'Hôpital 法则逐次求导数（容易验证该法则所要求的条件）：

$$\begin{aligned}\lim_{x\to a^+}\frac{f(x)-\sum_{k=0}^{n}\frac{f^{(k)}(a)}{k!}(x-a)^k}{(x-a)^n}&=\cdots=\lim_{x\to a^+}\frac{f^{(n-1)}(x)-f^{(n-1)}(a)-f^{(n)}(a)(x-a)}{n!(x-a)}\\&=\frac{1}{n!}\lim_{x\to a^+}\Big(\frac{f^{(n-1)}(x)-f^{(n-1)}(a)}{x-a}-f^{(n)}(a)\Big)\\&=0.\end{aligned}$$

最后一步是利用导数的定义。

2) 将 x 视作是固定的，我们定义

$$F(t)=\sum_{k=0}^{n}\frac{f^{(k)}(t)}{k!}(x-t)^k=f(t)+\sum_{k=1}^{n}\frac{f^{(k)}(t)}{k!}(x-t)^k.$$

我们对 t 求导数得到：

$$\begin{aligned}F'(t)&=f'(t)+\sum_{k=1}^{n}\frac{f^{(k+1)}(t)}{k!}(x-t)^k+\sum_{k=1}^{n}\frac{(-1)f^{(k)}(t)}{(k-1)!}(x-t)^{k-1}\\&=f'(t)+\sum_{k=1}^{n}\frac{f^{(k+1)}(t)}{k!}(x-t)^k-\sum_{\ell=0}^{n-1}\frac{f^{(\ell+1)}(t)}{\ell!}(x-t)^{\ell}.\end{aligned}$$

所以，

$$F'(t)=\frac{f^{(n+1)}(t)}{n!}(x-t)^n.$$

现在考虑另一个关于 t 的函数（定义在 $[a,x]$ 上）：

$$G(t)=\Big(\frac{x-t}{x-a}\Big)^{n+1}$$

很明显，$G'(t)$ 在 (a,x) 上不是零。根据 Cauchy 中值定理，存在 $\xi\in(a,x)$，使得

$$\frac{F'(\xi)}{G'(\xi)}=\frac{F(x)-F(a)}{G(x)-G(a)}.$$

即

$$\frac{\frac{f^{(n+1)}(\xi)}{n!}(x-\xi)^n}{\frac{-(n+1)(x-\xi)^n}{(x-a)^{n+1}}} = \frac{F(x)-F(a)}{-G(a)} = F(x)-F(a) = f(x) - \sum_{k=0}^{n} \frac{f^{(k)}(a)}{k!}(x-a)^k.$$

整理即得 Lagrange 余项的公式。

3) 为了证明 Cauchy 余项的公式，我们同样考虑上述的 $F(t)$，但是我们将选取一个不同的 G：

$$G(t) = \frac{x-t}{x-a}$$

很明显，$G'(t)$ 在 (a,x) 上不是零。根据 Cauchy 中值定理，存在 $\xi \in (a,x)$，使得

$$\frac{F'(\xi)}{G'(\xi)} = \frac{F(x)-F(a)}{G(x)-G(a)}.$$

即

$$\frac{\frac{f^{(n+1)}(\xi)}{n!}(x-\xi)^n}{\frac{-1}{x-a}} = \frac{F(x)-F(a)}{-G(a)} = F(x)-F(a) = f(x) - \sum_{k=0}^{n} \frac{f^{(k)}(a)}{k!}(x-a)^k.$$

整理即得 Cauchy 余项的公式。

□

上面 Lagrange 余项的证明很有技巧性，（我）很难理解。如果有了积分作工具，我们可以给出一个最自然的证明。我们现在给出另一个证明：我们知道，当 $n=1$ 时 Lagrange 余项的公式是 Lagrange 中值定理，我们下面利用中值定理证明的方法，来给出一个相对自然的（容易记住）证明。为此，首先推广 Rolle 定理到高阶导数的情形：

引理 105. *假设 $f \in C^n\big([a,b]\big)$ 并且在 (a,b) 上 $n+1$ 次可导。如果 f 在 a 处的 n-次导数全为零，即 $f'(a)=0,\cdots,f^{(n)}(a)=0$ 并且 $f(a)=f(b)$，那么存在 $x_0 \in (a,b)$，使得 $f^{(n+1)}(c)=0$。*

证明: 我们只要不停地用 Rolle 中值定理即可：由于 $f(a)=f(b)$，根据 Rolle 中值定理，存在 $x_1 \in (a,b)$，使得 $f'(x_1)=0$；由于 $f'(a)=f'(x_1)=0$，再用 Rolle 中值定理，我们就找到 $x_2 \in (a,x_1)$，使得 $f''(x_2)=0$；如此下去，我们得到 $f''(x_2)=f'''(x_3)=\cdots=0$。最后一步，就得到了 $f^{(n+1)}(x_{n+1})=0$。选取 $x_0=x_{n+1}$ 即可。 □

我们仿照 Lagrange 中值定理的证明：先构造多项式 $P(x)$，使得 $P(a)=f(a)$，$P'(a)=f'(a)$，$\cdots$，$P^{(n)}(a)=f^{(n)}(a)$，比如，我们取

$$P_n(x) = \sum_{k=0}^{n} \frac{f^{(k)}(a)}{k!}(x-a)^k.$$

练习. *证明，如果多项式 $P(x)$ 的次数 $\leqslant n$，那么满足上面条件的多项式是唯一的。*

为了应用高阶导数的 Rolle 定理，我们对 $P_n(x)$ 略加改造：选取 $\lambda \in \mathbb{R}$（存在且唯一），使得 $P(b) = f(b)$, 其中 $P(x) = P_n(x) + \lambda(x-a)^{n+1}$, $\lambda = \dfrac{f(b) - P_n(b)}{(b-a)^{n+1}}$。我们现在考虑函数 $f(x) - P(x)$, 它满足高阶导数 Rolle 定理的条件，所以存在 $c \in (a, b)$，使得

$$f^{(n+1)}(c) - P^{(n+1)}(c) = 0.$$

利用 λ 的表达式，我们得到

$$f^{(n+1)}(c) - (n+1)!\frac{f(b) - P_n(b)}{(b-a)^{n+1}} = 0.$$

如果改写为 $c = \xi$，$b = x$，这就是 Lagrange 余项的 Taylor 公式。

注记. *满足 Peano 余项的 Taylor 展开公式是唯一的，即若假设函数 $f : [a, b] \to \mathbb{R}$（或者 $\mathbb{C}$）在 a 处的一直到 n-次导数 $f'(a), \cdots, f^{(n)}(a)$ 都存在。那么，如果存在次数不超过 n 的多项式 $P(x)$, 使得当 $x \to a^+$ 时，我们有*

$$f(x) = P(x) + o\big((x-a)^n\big).$$

那么，

$$P(x) = \sum_{k=0}^{n} \frac{f^{(k)}(a)}{k!}(x-a)^k.$$

令 $Q(x) = P(x) - \sum\limits_{k=0}^{n} \dfrac{f^{(k)}(a)}{k!}(x-a)^k$，实际上按照 Peano 余项的 Taylor 展开公式，我们必然有

$$\lim_{x \to a^+} \frac{Q(x)}{(x-a)^n} = \lim_{x \to a^+} \frac{P(x) - \sum_{k=0}^{n} \frac{f^{(k)}(a)}{k!}(x-a)^k}{(x-a)^n} = 0.$$

由于 $\deg Q \leqslant n$，所以 $Q = 0$。

由此可见，如果限定的多项式的次数 $\leqslant n$，那么 $\sum\limits_{k=0}^{n} \dfrac{f^{(k)}(a)}{k!}(x-a)^k$ 是在 a 附近对 $f(x)$ 最佳的逼近。

例子. *我们有两个比较极端的例子：*

1) 正弦函数 $\sin x$：我们可以在 $x = 0$ 处计算其导数（偶数次的导数都是零），从而得到它的 Peono 展开为：

$$\sin x = \sum_{k=0}^{n} \frac{(-1)^k x^{2k+1}}{(2k+1)!} + o(|x|^{2n+2}).$$

这和 $\sin x$ 的解析表达式之间似乎是一致的（我们暂时不研究这一点）。

2) 我们考虑如下的函数

$$f(x) = \begin{cases} e^{-\frac{1}{x^2}}, & x > 0; \\ 0, & x \leqslant 0 \end{cases}$$

那么函数在 0 点处的所有导数都是 0，从而对任意的 $n \geqslant 1$，

$$f(x) = o(|x|^n).$$

当然，$f(x)$ 不是零。

17　凸函数，利用 Jensen 不等式证明常见的不等式

二零一九年十一月十四日，星期四，晴天

微分的应用：凸函数的性质和常用不等式

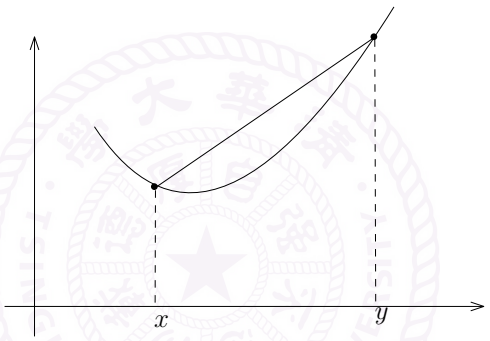
我们给出凸函数的五种刻画：

定理 106. *I 是 $\mathbb{R}$ 上的区间，给定 I 上定义的实值函数 $f: I \to \mathbb{R}$。那么，五个性质是等价的：*

1) 对任意的 $x, y \in I$，任意的 $t \in [0,1]$，我们有

$$f\big(tx + (1-t)y\big) \leqslant tf(x) + (1-t)f(y).$$

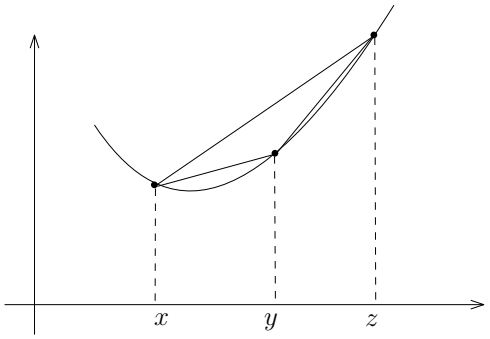
从图形上来看，任给 f 的图像上两个点，这两个点之间的函数图像落在这两个点所连线段的下方。



2) 对任意的 $x, y, z \in I$，如果 $x < y < z$，那么

$$\frac{f(y)-f(x)}{y-x} \leqslant \frac{f(z)-f(x)}{z-x} \leqslant \frac{f(z)-f(y)}{z-y}.$$

我们可以通过函数图像形象的记忆这一串不等式：



3) 对任意 $a \in I$，如下定义的函数

$$I - \{a\} \to \mathbb{R},\ x \mapsto \frac{f(x)-f(a)}{x-a},$$

是变量 x 的增函数。

4) 对任意的 $x,y,z\in I$，如果 $x<y<z$，那么

$$\det\begin{pmatrix}1 & 1 & 1\\ x & y & z\\ f(x) & f(y) & f(z)\end{pmatrix}\geqslant 0.$$

5) 集合 $\Gamma_{\geqslant f}=\Big\{(x,y)\Big|x\in I,y\geqslant f(x)\Big\}\subset\mathbb{R}^2$（即函数图像上方的部分）是凸集

如果函数 f 满足上述条件之一，我们就称 f 是**凸函数**。如果 $-f$ 是凸函数，我们就称 f 是**凹函数**。

注记. V 是 $\mathbb{R}$-线性空间，$C\subset V$ 是子集。如果对任意的 $x,y\in C$，任意的 $t\in[0,1]$，$tx+(1-t)y\in C$，即 x 和 y 所连线段落在 C 中，我们就称 C 是**凸集**。

注记. 上述对凸性的刻画本质上是几何的，叙述中并没有涉及到函数的连续性等。有意思的是，凸性的几何定义可以推出函数的若干分析性质，比如连续性、可微性之类的。

证明: 我们转一圈来证明前面四个命题之间的等价性：

1) $\Rightarrow$ 2) 由于 $y\in(x,z)$，所以存在 $t\in(0,1)$，使得 $y=tx+(1-t)z$，其中 $t=\dfrac{y-z}{x-z}$。所以，

$$f(y)\leqslant tf(x)+(1-t)f(z)=\frac{y-z}{x-z}f(x)+\frac{x-y}{x-z}f(z).$$

这个不等式等价于 2) 中的不等式（直接计算）。

2) $\Leftrightarrow$ 3) 2) 中左侧的不等式讲的就是这个函数是递增的。

3) $\Rightarrow$ 4) 我们注意到

$$\det\begin{pmatrix}1 & 1 & 1\\ x & y & z\\ f(x) & f(y) & f(z)\end{pmatrix}=(z-x)(y-x)\Big(\frac{f(z)-f(x)}{z-x}-\frac{f(y)-f(x)}{y-x}\Big).$$

利用 3) 中的单调性，这个值明显是非负的。

4) $\Rightarrow$ 1) 我们不妨设 $x\leqslant y$，我们对下面的式子展开：

$$\det\begin{pmatrix}1 & 1 & 1\\ x & tx+(1-t)y & y\\ f(x) & f(tx+(1-t)y) & f(y)\end{pmatrix}\geqslant 0.$$

这等价于

$$(y-x)\big((1-t)\big(f(y)-f(x)\big)-\big(f(tx+(1-t)y)-f(x)\big)\big)\geqslant 0.$$

整理立得。

最终，我们来证明 5) $\Leftrightarrow$ 1)：假设 $\Gamma_{\geqslant f}$ 是凸集，那么对任意的 $x, y \in I$，我们有 $(x, f(x)), (y, f(y)) \in \Gamma_{\geqslant f}$，所以，对任意的 $t \in [0,1]$，我们有

$$(tx + (1-t)y, tf(x) + (1-t)f(y)) \in \Gamma_{\geqslant f}.$$

按照 $\Gamma_{\geqslant f}$ 的定义，我们有 $f(tx + (1-t)y) \leqslant tf(x) + (1-t)f(y))$。

反之，假设 1) 成立，任选 $(x_1, y_1), (x_2, y_2) \in \Gamma_{\geqslant f}$，按照定义 $y_1 \geqslant f(x_1)$，$y_2 \geqslant f(x_2)$。考虑这两个点所连线段上的任意一个点 $(x, y) = t(x_1, y_1) + (1-t)(x_2, y_2)$，其中 $t \in [0,1]$。根据 1)，我们知道

$$f(x) \leqslant tf(x_1) + (1-t)f(x_2) \leqslant ty_1 + (1-t)y_2 = y,$$

所以 $(x, y) \in \Gamma_{\geqslant f}$。 □

根据等价定义中的第一条，我们可以证明：

命题 107 (Jensen 不等式)**.** *假设 $f : I \to \mathbb{R}$ 是凸函数，那么对任意的 $x_1, \cdots, x_n \in I$ 和任意的 $t_1, \cdots, t_n \in [0,1]$，其中 $t_1 + t_2 + \cdots + t_n = 1$，我们有*

$$f(t_1x_1 + \cdots + t_nx_n) \leqslant t_1f(x_1) + \cdots + t_nf(x_n).$$

证明: 我们对 n 进行归纳。$n = 2$ 时，这是定义；假设对 n 成立，考虑 $n+1$ 的情形。对任意给定的 $x_1, \cdots, x_{n+1} \in I$，$t_1, \cdots, t_{n+1} \in [0,1]$，其中 $t_1 + t_2 + \cdots + t_{n+1} = 1$，我们有

$$f(t_1x_1 + \cdots + t_nx_n + t_{n+1}x_{n+1}) = f\big((1-t_{n+1})\frac{t_1x_1 + \cdots + t_nx_n}{1-t_{n+1}} + t_{n+1}x_{n+1}\big).$$

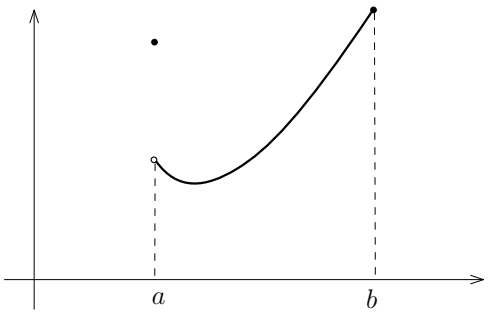
不难看出，$\dfrac{t_1x_1 + \cdots + t_nx_n}{1-t_{n+1}} \in I$，此时我们用等价定义中的第一条，就得到

$$\begin{aligned} f(t_1x_1 + \cdots + t_nx_n + t_{n+1}x_{n+1}) &\leqslant (1-t_{n+1})f\big(\frac{t_1}{1-t_{n+1}}x_1 + \cdots + \frac{t_n}{1-t_{n+1}}x_n\big) + t_{n+1}f(x_{n+1}) \\ &\leqslant (1-t_{n+1})\left(\frac{t_1}{1-t_{n+1}}f(x_1) + \cdots + \frac{t_n}{1-t_{n+1}}f(x_n)\right) + t_{n+1}f(x_{n+1}) \\ &= t_1f(x_1) + \cdots + t_{n+1}f(x_{n+1}). \end{aligned}$$

这就完成了证明。 □

注记. *我们会利用这个命题来证明很多经典的不等式。*

一个凸函数可以不连续，比如说下图所示的函数



它在左端点处不连续。然而，这是一个凸函数可能不连续的唯一方式：

定理 108. *$I \subset \mathbb{R}$ 是区间，$f : I \to \mathbb{R}$ 是凸函数，我们令 $\mathring{I}$ 为 I 的内部（对于 $I = [a,b], [a,b), (a,b]$ 或 (a,b)，那么 $\mathring{I} = (a,b)$），那么*

1) $f \in C(\mathring{I})$。

2) 对任意的 $x_0 \in \mathring{I}$，f 在 x_0 处的左导数 $f'^-(x_0)$ 和右导数 $f'^+(x_0)$ 存在。进一步，它们满足

$$f'^-(x_0) \leqslant f'^+(x_0).$$

3) 对任意的 $x, y \in \mathring{I}$，如果 $x < y$，那么我们有

$$f'^-(x) \leqslant \frac{f(y) - f(x)}{y - x} \leqslant f'^+(y).$$

4) 左右导数 $f'^{\pm} : \mathring{I} \to \mathbb{R}$ 是递增的函数。

证明： 任意选定 $x_0 \in \mathring{I}$。存在 $h_1, h_2 > 0$，$x_0 - h_1, x_0 + h_2 \in \mathring{I}$，对三个点 $x_0 - h_1, x_0$ 和 $x_0 + h_2$ 用凸性的第二个等价定义，我们得到

$$\frac{f(x_0) - f(x_0 - h_1)}{h_1} \leqslant \frac{f(x_0 + h_2) - f(x_0)}{h_2},$$

根据凸性的第三个等价定义，由于上述不等式的右边是 h_2 的增函数，令 $h_2 \to 0$，我们立即得到右导数的存在性；类似地，如果令 $h_1 \to 0$，我们就得到左右导数存在性。既然左右导数都存在，那么，f 自然在 $\mathring{I}$ 上连续。

为了证明 3)，我们先令 $h_2 \to 0$，从而，

$$\frac{f(x_0) - f(x_0 - h_1)}{h_1} \leqslant f'^+(x_0),$$

令 $x_0 = y$，$x_0 - h_1 = x$，我们就得到了 3) 中不等式的右边；左边类似可得。

为了证明 4)，我们考虑 3) 中的不等式

$$f'^-(x) \leqslant \frac{f(x) - f(y)}{x - y} \leqslant \frac{f(x) - f(z)}{x - z},$$

其中 $x < z < y$，后一个不等式是运用凸性的第三个等价定义。再令 $z \to x$（$z < x$）就证明了结论。 □

上述分析性质基本上刻画了凸函数：

定理 109. *假设 I 是开区间，$f : I \to \mathbb{R}$ 是函数。如下两个性质等价：*

1) f 是凸函数；

2) f 是连续函数，f 的右导数处处存在并且是增函数。

证明: 只要说明 2)⇒1) 即可：我们固定 I 中的三个点 $x<y<z$，对任意的 $t\in[y,z]$，我们有 $f'^{+}(y)\leqslant f'^{+}(t)\leqslant f'^{+}(z)$。我们定义函数

$$g:[y,z]\to\mathbb{R},\quad t\mapsto f(t)-tf'^{+}(y).$$

那么，对于 $t\in(y,z)$，g 在 t 处的右导数 $g'^{+}(t)=f'^{+}(t)-f'^{+}(y)\geqslant 0$。把这个和导数的性质做类比，我们想说明 g 是 $t\in(y,z)$ 的增函数：

引理 110. 假设 $g:[a,b]\to\mathbb{R}$ 是连续函数并且右导数在每个点处都定义。如果对任意的 $t\in(a,b)$，我们都有 $g'^{+}(t)\geqslant 0$，那么 $g(a)\leqslant g(b)$。

先假设引理是成立的，那么我们有 $g(z)\geqslant g(y)$，经过整理，我们得到

$$f'^{+}(y)\leqslant\frac{f(z)-f(y)}{z-y}.$$

类似地，我们可以证明

$$f'^{+}(y)\geqslant\frac{f(y)-f(x)}{y-x}.$$

从而，

$$\frac{f(y)-f(x)}{y-x}\leqslant\frac{f(z)-f(y)}{z-y}.$$

所以，我们有

$$\det\begin{pmatrix}1&1&1\\x&y&z\\f(x)&f(y)&f(z)\end{pmatrix}=\det\begin{pmatrix}1&0&0\\x&y-x&z-y\\f(x)&f(y)-f(x)&f(z)-f(y)\end{pmatrix}\geqslant 0.$$

这表明 f 是凸函数。 □

引理的证明.（证明的重要想法：退 ε-步海阔天空）任意选定 $\varepsilon>0$，令

$$\varphi_\varepsilon:[a,b]\to\mathbb{R},x\mapsto-(g(x)-g(a))-\varepsilon(x-a).$$

此时，$\varphi_\varepsilon'^{+}\leqslant-\varepsilon<0$，我们实质上想证明这是一个递减的函数，如果成立的话，那么

$$\varphi_\varepsilon(a)\geqslant\varphi_\varepsilon(b)\ \Rightarrow\ \underbrace{0\geqslant\varphi_\varepsilon(b)}_{\text{过渡的结论}}\ \Rightarrow\ g(b)-g(a)\geqslant-\varepsilon(b-a).$$

令 $\varepsilon\to 0$ 就得到了要证明的结论。

我们现在直接证明上述过渡的结论。为此，定义

$$X_\varepsilon=\left\{x\in[a,b]\middle|\varphi_\varepsilon(x)\leqslant\varepsilon\right\}$$

由于 $X=[a,b]\cap(\varphi_\varepsilon)^{-1}\big((-\infty,\varepsilon]\big)$ 是闭集，所以它的上确界 $c=\sup X_\varepsilon\in X_\varepsilon$。我们想证明 $b\in X$（从而，$\varepsilon\geqslant\varphi_\varepsilon(b)$，令 $\varepsilon\to 0$，过渡性结论就成立了），即要证明 $c=b$。

因为 $\varphi_\varepsilon(a)=0$，所以 $a\in X_\varepsilon$。特别地，根据 $\varphi_\varepsilon(x)$ 的连续性，存在 $\delta>0$，使得 $[a,a+\delta)\subset X_\varepsilon$。所以 $X_\varepsilon\neq\emptyset$ 并且 $c>a$。

我们用反证法证明 $c=b$：如若不然，我们假设 $c<b$。先说明 $\varphi_\varepsilon(c)=\varepsilon$，这因为按照定义 $\varphi_\varepsilon(c)\leqslant\varepsilon$，如果等号不成立，那么 $\varphi_\varepsilon(c)<\varepsilon$，根据 $\varphi_\varepsilon(x)$ 的连续性，存在比 c 略大的数使得其值仍然不超过 ε，这就与 $c=\sup X_\varepsilon$ 矛盾。

其次根据右导数的信息，$\varphi_\varepsilon'^{+}(c)\leqslant-\varepsilon$，利用右导数的定义，存在 $h_0>0$，使得当 $h\in[0,h_0)$ 时，我们有

$$\frac{\varphi_\varepsilon(c+h)-\varphi_\varepsilon(c)}{h}<-\frac{1}{2}\varepsilon.$$

从而，

$$\varphi_\varepsilon(c+h)<\varepsilon-\frac{1}{2}\varepsilon h<\varepsilon.$$

这表明 $[c,c+h_0)\subset X_\varepsilon$，这和 c 是上确界矛盾。 □

推论 111. *假设 $I\subset\mathbb{R}$ 是开区间，f 是 I 上二次可微的实值函数。如下两个陈述是等价的：*

1) f 是凸函数；

2) $f''\geqslant 0$。

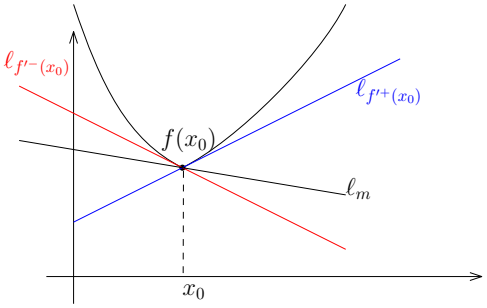
证明： 这是上面定理的直接推论。 □

注记. *这个推论是最常用的来判断凸性的工具，因为函数二阶导数通常容易计算。*

推论 112 (函数图像的支撑直线)**.** *$f:I\to\mathbb{R}$ 是区间 I 上的凸函数，$x_0\in\mathring{I}$。考虑 $\mathbb{R}^2$ 上的过 $(x_0,f(x_0))$ 直线*

$$\ell_m=\{(x,y)\in\mathbb{R}^2|y=m(x-x_0)+f(x_0)\},$$

其中斜率 $m\in\mathbb{R}$。那么，ℓ_m 在 f 的图像之下当且仅当 $f'^{-}(x_0)\leqslant m\leqslant f'^{+}(x_0)$。



证明： 定理的证明过程表明，当 $x>x_0$ 时，我们有

$$f'^{+}(x_0)\leqslant\frac{f(x)-f(x_0)}{x-x_0}\ \Rightarrow\ f(x)\geqslant f(x_0)+(x-x_0)f'^{+}(x_0).$$

反过来，当 $x<x_0$ 时，我们有

$$f(x)\geqslant f(x_0)+(x-x_0)f'^{-}(x_0).$$

据此，当 $m \leqslant f'^{+}(x_0)$ 时，在 x_0 的右边，该直线在 f 的图像之下；当 $m \geqslant f'^{-}(x_0)$ 时，在 x_0 的做边，该直线在 f 的图像之下。

反之，假设 ℓ_m 在 f 的图像之下，在 x_0 处的右边的时候，我们有 $f(x) \geqslant m(x-x_0)+f(x_0)$，从而

$$f'^{+}(x_0) = \lim_{x \to x_0^+} \frac{f(x)-f(x_0)}{x-x_0} \geqslant m.$$

类似地，我们可以证明另一边的不等式。 □

用凸函数证明常见的不等式

我们给出两个例子：Minkowski 不等式和算术-几何平均值不等式。

1) 假设 $p \geqslant 1$，对于 $x \in [0,1]$，函数 $f(x) = (1-x^{\frac{1}{p}})^p$ 是凸函数，这因为

$$f'(x) = -px^{\frac{1}{p}-1}\big(1-x^{\frac{1}{p}}\big)^{p-1}$$

是递增的函数（求两次导数的表达式不是很简单）。假设 $a_i, b_i \in \mathbb{R}_{\geqslant 0}$ 并且假设 $a_i + b_i \neq 0$，其中 $i = 1, 2, \cdots, n$。我们令

$$x_i = \frac{a_i^p}{(a_i+b_i)^p}, \quad t_i = \frac{(a_i+b_i)^p}{\sum\limits_{k=1}^{n}(a_k+b_k)^p}.$$

Jensen 不等式给出了如下所谓的 Minkowski 不等式：

$$\Big(\sum_{k=1}^{n}(a_k+b_k)^p\Big)^{\frac{1}{p}} \leqslant \Big(\sum_{k=1}^{n}a_k^p\Big)^{\frac{1}{p}} + \Big(\sum_{k=1}^{n}b_k^p\Big)^{\frac{1}{p}}.$$

作为推论，对于线性空间 $V = \mathbb{R}^n$，我们定义

$$\|x\|_p = \Big(\sum_{k=1}^{n}x_k^p\Big)^{\frac{1}{p}},$$

其中，$x = (x_1, \cdots, x_p)$。此时，Minkowski 不等式讲的是

$$\|x\|_p + \|y\|_p \geqslant \|x+y\|_p.$$

这表明 $\|x\|_p$ 是范数（$p=2$ 或者 1 我们更熟悉）。另外，我们还可以采取如下的范数：

$$\|x\|_\infty = \sup_{k \leqslant n}|x_k|.$$

2) 函数 $-\log x$ 是 $\mathbb{R}_{>0}$ 上的凸函数，这因为

$$(-\log x)'' = \frac{1}{x^2} \geqslant 0$$

假设 $x_1, \cdots, x_n$ 是正数，根据 Jensen 不等式，我们有

$$-\log\left(\frac{1}{n}x_1 + \cdots + \frac{1}{n}x_n\right) \leqslant -\frac{1}{n}\log(x_1) - \cdots - \frac{1}{n}\log(x_n).$$

这个等价于算术-几何平均值不等式：

$$\frac{1}{n}(x_1 + \cdots + x_n) \geqslant (x_1 \cdots x_n)^{\frac{1}{n}}.$$

17.1 作业：Émile Borel 引理，Peano 的证明

清华大学 19-20 秋季学期，数学分析一，作业 7

请用 A4 大小的纸张**正反面**用钢笔，签字笔或者圆珠笔书写，并注明自己的姓名，年级和作业的**总页数**。除定理公式所涉及的人名之外，请使用**中文**。本次作业的提交时间和地点为 **11 月 21 日**上午的课堂上，逾期视作零分。

基本习题

习题 A：中值定理和 Taylor 展开

如果不额外说明，f 总代表一个区间 I 上定义的函数。

A1) 设 f 在点 x 处有二阶导数。证明，下面的极限给出了二阶导数：

$$f''(x) = \lim_{h\to 0} \frac{f(x+h)+f(x-h)-2f(x)}{h^2}$$

A2) (Taylor 展开式的唯一性，Peano 余项)。假设 f 在 x_0 附近的函数，并且当 $x \to x_0$ 是，满足

$$\begin{aligned} f(x) &= a_0 + a_1(x-x_0) + a_2(x-x_0)^2 + \cdots + a_n(x-x_0)^n + o(|x-x_0|^n) \\ &= b_0 + b_1(x-x_0) + b_2(x-x_0)^2 + \cdots + b_n(x-x_0)^n + o(|x-x_0|^n) \end{aligned}$$

其中 $a_i, b_i, i = 0, \cdots, n$ 是实数，那么，对任意的 i，我们都有 $a_i = b_i$。

A3) 假设 f 在 0 处有 n 阶导数。证明，如果 $f(x)$ 是偶函数（奇函数），f 在 0 处的 Taylor 展开式（Peano 余项）中只有 x 的偶次项（奇数项）。

A4) (Rolle 定理的简单推广)设函数在有限或无穷的区间 (a,b) 上可微并且 $\lim\limits_{x\to a^+} f(x) = \lim\limits_{x\to b^-} f(x)$。证明，存在 $x_0 \in (a,b)$，使得 $f'(x_0) = 0$。

A5) 设函数 $f \in C^0([a,b])$ 并且在 (a,b) 上可微。证明，f 在 $[a,b]$ 上严格递增的充分必要条件是对任意 $x \in (a,b)$，$f'(x) \geqslant 0$ 并且在任意子区间 $(c,d) \subset (a,b)$ 上，$f'(x)$ 不恒等于 0。

A6) $g \in C(\mathbb{R})$ 在 $\mathbb{R}$ 上可微。假设存在常数 M，使得 $\sup\limits_{x\in\mathbb{R}} |g'(x)| \leqslant M$。对任意的 $\varepsilon > 0$，我们定义

$$f_\varepsilon(x) = x + \varepsilon g(x).$$

证明，存在仅依赖于 M 的常数 $\delta = \delta(M) > 0$，使得当 $\varepsilon < \delta$ 时，$f : \mathbb{R} \to \mathbb{R}$ 是双射。

A7) 设函数 f 在 $[a,b]$ 上有两阶导数并且 $f'(a) = f'(b) = 0$。证明，存在 $c \in (a,b)$ 使得

$$|f''(c)| \geqslant \frac{4}{(b-a)^2} |f(b) - f(a)|$$

A8) 假设 f 在 $\mathbb{R}$ 上二次可导，对 $k=0,1,2$，我们假设 $M_k=\sup\limits_{x\in\mathbb{R}}|f^{(k)}(x)|$ 都是有限的。证明，$M_1^2\le 2M_0M_2$。

A9) 假设 f 在 $(0,+\infty)$ 上二次可导，f'' 在 $(0,+\infty)$ 上有界并且 $\lim\limits_{x\to+\infty}f(x)=0$。证明，$\lim\limits_{x\to+\infty}f'(x)=0$。

习题 B 用 L'Hôpital 法则求极限：

(1) $\lim\limits_{x\to+\infty}\dfrac{\log x}{x^a},\quad a>0$　(2) $\lim\limits_{x\to+\infty}\dfrac{x^a}{b^x},\quad a>0,b>1$　(3) $\lim\limits_{x\to0}\dfrac{e^{ax}-e^{bx}}{\sin ax-\sin bx}$

(4) $\lim\limits_{x\to0}\dfrac{\tan x-x}{x-\sin x}$　(5) $\lim\limits_{x\to0}\dfrac{1-\cos x^2}{x^2\sin x^2}$　(6) $\lim\limits_{xt01}\dfrac{\sqrt{2x-x^4}-\sqrt[3]{x}}{1-x^{4/3}}$

(7) $\lim\limits_{x\to1^-}\log(x)\log(1-x)$　(8) $\lim\limits_{x\to0^+}\dfrac{\log\sin ax}{\log\sin bx},\quad a,b>0$　(9) $\lim\limits_{x\to0^+}x^x$

(10) $\lim\limits_{x\to1}x^{\frac{1}{1-x}}$　(11) $\lim\limits_{x\to1}\left(\dfrac{1}{\log x}-\dfrac{1}{x-1}\right)$　(12) $\lim\limits_{x\to0^+}(\sin x)^x$

(13) $\lim\limits_{x\to0}\left(\dfrac{\sin x}{x}\right)^{1/(1-\cos x)}$　(14) $\lim\limits_{x\to a}\dfrac{a^x-x^a}{x-a},\quad a>0$　(15) $\lim\limits_{x\to+\infty}\dfrac{(1+1/x)^x-e}{1/x}$

(16) $\lim\limits_{x\to+\infty}\dfrac{x^{\log x}}{(\log x)^x}$　(17) $\lim\limits_{x\to+\infty}\left[(x+a)^{1+1/x}-x^{1+\frac{1}{x+a}}\right]$

(18) $\lim\limits_{x\to+\infty}\left[\sqrt[3]{x^3+x^2+x+1}-\sqrt{x^2+x+1}\cdot\dfrac{\log(e^x+x)}{x}\right]$

习题 C 求下列函数在给定区间上的最大值和最小值：

1) $f(x)=x^4-2x^2+5$，$x\in[-2,2]$；　2) $f(x)=\dfrac{2x}{1+x^2}$，$x\in\mathbb{R}$；

3) $f(x)=\arctan x-\dfrac{1}{2}\log(1+x^2)$；　4) $f(x)=x\log x$，$x\in(0,+\infty)0$；

5) $f(x)=\sqrt{x}\log x$，$x\in(0,+\infty)$；　6) $f(x)=2\tan x-\tan^2 x$，$x\in[0,\frac{\pi}{2})$。

习题 D（极大值的判定，重要）f 在 (a,b) 上可微。假设对于 $x_0\in(a,b)$，我们有 $f'(x_0)=0$。

D1) 证明，f 在 x_0 处取局部极大值的一个充分条件是：存在某一邻域 $(x_0-\delta,x_0+\delta)\subset(a,b)$，使得

$$f'(x)=\begin{cases}<0,\ \text{对所有的 } x\in(x_0-\delta,x_0);\\ >0,\ \text{对所有的 } x\in(x_0,x_0+\delta).\end{cases}$$

D2)（最重要的判别方法，有很多应用） 证明，如果 $f''(x_0)$ 存在并且 $f''(x_0)<0$，那么 f 在 x_0 处取局部极大值。

D3) 假定 f 在 x_0 处有 n 阶导数，$f'(x_0)=\cdots=f^{(n-1)}(x_0)=0$ 并且 $f^{(n)}(x_0)\neq0$，试讨论 f 在 x_0 处取局部极大值的条件（对 n 分奇偶讨论）。

习题 E（提示：利用中值定理和 n-次多项式的至多（恰好）有 n 个根）

E1) 证明，如果实系数多项式 $P_n(x)=a_nx^n+a_{n-1}x^{n-1}+\cdots+a_0$ 的根都是实数（不妨设 $a_n\neq 0$），那么它的逐次导函数 $P_n'(x)$，$P_n''(x)$，$\cdots$，$P_n^{(n-1)}(x)$ 的根也都是实数。

E2) 证明，Legendre 多项式 $P_n(x)=\dfrac{1}{2^n n!}\dfrac{d^n}{dx^n}(x^2-1)^n$ 的根都是实数并且包含于区间 $(-1,1)$ 中。

E3) 证明，$L_n(x)=\dfrac{e^x}{n!}\dfrac{d^n}{dx^n}(e^{-x}x^n)$ 是多项式（被称作是 Laguerre 多项式）并且它所有的根都是正实数。

E4) 证明，$H_n(x)=(-1)^n e^{x^2}\dfrac{d^n}{dx^n}(e^{-x^2})$ 是多项式（被称作是 Hermite 多项式）并且它所有的根都是实数。

习题 F（Émile Borel 引理）

第一部分：截断函数的构造

F1) 定义函数 $\phi:\mathbb{R}\to\mathbb{R}$:

$$\phi(x)=\begin{cases}e^{-\frac{1}{x^2}}, & x>0;\\ 0, & x\leqslant 0.\end{cases}$$

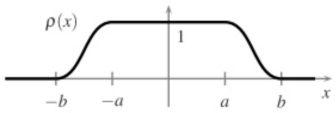
证明，$\phi\in C^\infty(\mathbb{R})$。

F2) 定义函数 $\chi:\mathbb{R}\to\mathbb{R}$:

$$\chi(x)=\frac{\phi(2-|x|)}{\phi(2-|x|)+\phi(|x|-1).}$$

证明，$\chi(x)\in C^\infty(\mathbb{R})$ 并且 $\chi\big|[-1,1]\equiv 1$，$\chi\big|_{(-\infty,-2]\cup[2,\infty)}\equiv 0$，$0\leqslant\chi(x)\leqslant 1$ 并且是偶函数。

F3) 证明，对任意的 $0<a<b$，存在光滑偶函数 $\rho(x)\in C^\infty(\mathbb{R})$，使得 $\rho\big|[-a,a]\equiv 1$，$\rho\big|_{(-\infty,-b]\cup[b,\infty)}\equiv 0$，$0\leqslant\rho(x)\leqslant 1$。



F4) 证明，存在偶函数 $\psi\in C^\infty(\mathbb{R}^n)$，使得 $\psi\big|_{\{x||x|\leqslant 1\}}\equiv 1$，$\chi\big|_{\{x||x|\geqslant 2\}}\equiv 0$，$0\leqslant\psi(x)\leqslant 1$。

第二部分：逐项求导定理

$I=[a,b]$ 是闭区间，$\{f_k\}_{k\geqslant 0}$ 是 $C^1(I)$ 的一列函数，我们假设 $\sum\limits_{k=0}^{\infty}f_k$ 在 I 上逐点收敛，即对任意的 $x\in I$，$\sum\limits_{k=0}^{\infty}f_k(x)$ 收敛，我们记 $f(x)=\sum\limits_{k=0}^{\infty}f_k(x)$。

F5) 我们假设函数级数 $\sum_{k=0}^{\infty} f_k'(x)$ 在 I 上绝对收敛，即数项的级数 $\sum_{k=0}^{\infty} \|f_k'\|_\infty$ 收敛，其中 $\|f\|_\infty = \sup_{x\in I}|f(x)|$。证明，$f$ 是可导的并且 $f'(x) = \sum_{k=0}^{\infty} f_k'(x)$。（提示：设法将求和拆 $\sum_{k=0}^{\infty} = \sum_{k=0}^{N} + \sum_{k=N+1}^{\infty}$ ）

F6)（逐项求导定理）$I=[a,b]$ 是闭区间，$\{f_k\}_{k\geqslant 0}$ 是 $C^1(I)$ 的一列函数，我们假设 $\sum_{k=0}^{\infty} f_k$ 在 I 上逐点收敛。如果函数级数 $\sum_{k=0}^{\infty} f_k'(x)$ 在 I 上一致收敛，那么 f 是可导的并且 $f'(x) = \sum_{k=0}^{\infty} f_k'(x)$。

F7) 试利用逐项求导定理计算 e^x 的导数。

第三部分：Borel 引理的证明

我们现在任意给定一个数列 $\{a_k\}_{k\geqslant 0}$。

F8) 对任意给定的正数 $\lambda_k > 0$，试计算函数 $f_k(x) = \frac{a_k}{k!}x^k\chi(t_k x)$ 在 $x=0$ 处的任意阶导数（包括零阶）。

F9) 证明，当 $k \geqslant 2n$ 时，我们有

$$f_k^{(n)}(x) = a_k \sum_{\ell=0}^{n} \binom{n}{\ell} \frac{t_k^{n-\ell}}{(k-\ell)!} x^{k-\ell} \chi^{(n-\ell)}(t_k x).$$

F10)（Borel 引理）任意给定一个数列 $\{a_k\}_{k\geqslant 0}$，证明，存在 $\mathbb{R}$ 上的光滑函数 f，使得对任意的 $k\geqslant 0$，$f^{(k)}(0) = a_k$。（提示：略难。）

第四部分：Peano 对 Borel 引理的证明 （选做部分，不交作业）

F11) $\{c_k\}_{k\geqslant 0}$ 和 $\{b_k\}_{k\geqslant 0}$ 是两个序列，其中 b_k 都是正数。证明，

$$\Big(\frac{c_k x^k}{1+b_k x^2}\Big)^{(n)}(0) = \begin{cases} n!(-1)^j c_{n-2j} b_{n-2j}^j, & \text{若 } k = n-2j, j\in\mathbb{Z}_{\geqslant 0}; \\ 0, & \text{其它情形.} \end{cases}$$

F12) 证明，存在常数 C，对任意的 $k\geqslant n+2$，对任意的 x，我们有

$$\Big|\Big(\frac{c_k x^k}{1+b_k x^2}\Big)^{(n)}(x)\Big| \leqslant C(n+1)!\frac{|c_k|k!}{b_k}|x|^{k-n-2}.$$

F13) 证明，当 $\{c_k\}_{k\geqslant 0}$ 给定的时候，我们可以选取 b_k，使得 b_k 仅依赖于 c_k 的选取，并且函数级数 $f(x) = \sum_{k=0}^{\infty} \frac{c_k x^k}{1+b_k x^2}$ 是无限次可微分的。

F14) 证明，$f(0)=c_0$，$f'(0)=c_1$，并且当 $n\geqslant 2$ 时，我们有

$$\frac{f^{(n)}(0)}{n!}=d_n+\sum_{j=1}^{\lfloor\frac{n}{2}\rfloor}(-1)^j c_{n-2j}b_{n-2j}^j.$$

F15) 证明，我们可以通过恰当的选取 $\{c_k\}_{k\geqslant 0}$ 和 $\{b_k\}_{k\geqslant 0}$ 来证明 Borel 引理。

以下为上学期期中考试的 B 部分，供同学参考，不交作业

题目 B 考虑在整个实数上定义的函数 $f:\mathbb{R}\to\mathbb{R}$。

- 如果存在正实数 A（不是无穷），使得对任意的 $x\in\mathbb{R}$，都有 $|f(x)|\leqslant A$，我们就称 f 是有界的。我们将 $\mathbb{R}$ 上定义的所有有界函数的集合记作 $\mathcal{B}$。
- 如果存在正实数 B（不是无穷），使得对任意的 $x,y\in\mathbb{R}$，都有 $|f(x)-f(y)|\leqslant B|x-y|$，我们就称 f 是 Lipschitz 函数。我们将 $\mathbb{R}$ 上定义的所有 Lipschitz 函数的集合记作 $\mathcal{L}$。

假设 $a,\lambda\in\mathbb{R}$，$f\in\mathcal{B}\cap\mathcal{L}$，这个问题的目的是找到一个函数 $F\in\mathcal{L}$ 来解下面的函数方程：

$$F(x)-\lambda F(x+a)=f(x),\quad x\in\mathbb{R}.\quad\cdots\cdots(\star)$$

第一部分：Lipschitz 函数的基本性质

B1) 证明，如果 $f,g\in\mathcal{B}\cap\mathcal{L}$，那么它们的乘积 $fg\in\mathcal{L}$。

B2) 证明，如果 f 是 $\mathbb{R}$ 上的可微函数并且 $f\in\mathcal{L}$，那么 $f'\in\mathcal{B}$。

B3) 证明，如果 f 是 $\mathbb{R}$ 上的可微函数并且 $f'\in\mathcal{B}$，那么 $f\in\mathcal{L}$。

B4) 如果 $f\in\mathcal{B}$ 并且存在正实数 B，使得对任意的 $x,y\in\mathbb{R}$，$|x-y|\leqslant 1$，我们都有 $|f(x)-f(y)|\leqslant B|x-y|$，证明，$f\in\mathcal{L}$。

第二部分：当 $|\lambda|<1$ 时，方程 $(\star)$ 的解，其中 $f\in\mathcal{B}\cap\mathcal{L}$。

我们在这一部分假设 $|\lambda|<1$。

B5) 假设 F 满足 $(\star)$，证明，对任意的 $x\in\mathbb{R}$ 和 $n\in\mathbb{Z}_{\geqslant 1}$，我们都有

$$F(x)=\lambda^n F(x+na)+\sum_{k=0}^{n-1}\lambda^k f(x+ka)\ \text{ 和 }\ F(x)=\lambda^{-n}F(x-na)-\sum_{k=0}^{n-1}\lambda^{-k}f(x-ka).$$

请任选其一来证明即可。

B6) 证明，对任意的 $x \in \mathbb{R}$，级数 $\sum_{k\geqslant 0} \lambda^k f(x+ka)$ 是收敛的。

B7) 根据上述，对每个 $x \in \mathbb{R}$，我们定义 $F(x) = \sum_{k\geqslant 0} \lambda^k f(x+ka)$。证明，$F \in \mathcal{L}$ 并且解方程 $(\star)$。

B8) 证明，如果 $F_1, F_2 \in \mathcal{L}$ 都满足方程 $(\star)$，那么 $F_1 = F_2$。

B9) 在 $(\star)$ 中取函数 $f(x) \equiv 1$，试求方程的解 F_1；

在 $(\star)$ 中取函数 $f(x) = \cos(x)$，证明，$F_2(x) = \dfrac{\cos(x) - \lambda\cos(x-a)}{1 - 2\lambda\cos(a) + \lambda^2}$ 是 $(\star)$ 的解。

第三部分：当 $|\lambda| > 1$ 时，方程 $(\star)$ 的解，其中 $f \in \mathcal{B} \cap \mathcal{L}$。

B10) 参照第二部分的结论，试陈述当 $|\lambda| > 1$ 时应当如何来解方程。你的陈述不要超过 100 个汉字。

B11) 当 $|\lambda| > 1$ 时，在 $(\star)$ 中取函数 $f(x) \equiv 1$，试求方程的解 F_1；当 $|\lambda| > 1$ 时，在 $(\star)$ 中取函数 $f(x) = \cos(x)$，试求 $(\star)$ 的解 $F_2(x)$。

第四部分：当 $|\lambda| = 1$ 时的情形。

B12) 假设 $\lambda = 1$。证明，存在非零的函数 $F \in \mathcal{L}$，使得对任意的 x，我们都有 $F(x) - F(x+a) = 0$。

B13) 假设 $\lambda = 1$。在 $(\star)$ 中取函数 $f(x) = \cos(x)$。证明，如果 $\cos(a) \neq 1$，那么存在 $F \in \mathcal{L}$ 是 $(\star)$ 的解。进一步阐述这个解是否唯一?

B14) 假设 $\lambda = 1$。在 $(\star)$ 中取函数 $f(x) = \cos(x)$。证明，如果 $a = 2\pi$，那么我们不能在 $\mathcal{L}$ 中找到 F，使得 F 是 $(\star)$ 的解。

B15) 假设 $\lambda = -1$。证明，存在非零的函数 $F \in \mathcal{L}$，使得对任意的 x，我们都有 $F(x) + F(x+a) = 0$。

B16) 假设 $\lambda = -1$，$a = 1$，$f \in \mathcal{L}$，f 是递减的并且 $\lim_{x\to+\infty} f(x) = 0$，$f$ 可微并且 f' 是递增的。证明，存在 $F \in \mathcal{L}$，解如下方程：

$$F(x) + F(x+1) = f(x), \quad x \in \mathbb{R}.$$

进一步，如果我们要求 $F \in \mathcal{L}$ 并且 $\lim_{x\to+\infty} F(x) = 0$，那么这样的解是存在唯一的。

18 Riemann 积分的定义：区间的分划，简单函数，Riemann 可积函数

二零一九年十一月十七日，星期一，晴天

一维的 Riemann 积分

假设 $a<b$ 是实数，$I=[a,b]$ 是一个闭区间，我们定义所谓的分划的概念：选取 $n+1$ 个实数 $a_0,a_1,a_2,\cdots,a_n$，使得 $a=a_0<a_1<\cdots<a_{n-1}<a_n=b$。这些有序的数将 I 分成了 n 份，

$$I=[a_0,a_1]\cup[a_1,a_2]\cup\cdots\cup[a_{n-1},a_n]$$

我们将上面的对象（这 $n+1$ 个有序实数）称为是 I 的一个有限）**分划**。我们把 I 上所有分划所组成的集合记作 $\mathcal{S}(I)$，为了简单起见，我们通常把它写成 $\mathcal{S}$。给定一个分划 $\sigma\in\mathcal{S}$，假设它对应着上述的 $n+1$ 个数 $a=a_0<a_1<\cdots<a_{n-1}<a_n=b$，我们把下面的量称作是分划 σ 的**步长**：

$$|\sigma|=\max_{0\leqslant i\leqslant n-1}|a_i-a_{i+1}|.$$

我们把 $\{a_0,a_1,a_2,\cdots,a_n\}$ 称作是分划 σ 的**分割点**。很明显，给定 I 的一个分划 $\sigma\in\mathcal{S}$ 等价于给定包含 I 两个端点的（I 的）有限子集（= 分割点的集合）。

例子. 我们可以将 I 均分为 n 份：$a_k=a+\dfrac{k}{n}(b-a)$，其中 $k=0,1,\cdots,n$。这是 I 的一个步长为 $\dfrac{b-a}{n}$ 的分划。

考虑 I 的两个分划 $\sigma,\sigma'\in\mathcal{S}$，如果 σ 的分割点的集合是 σ' 的分割点的集合，我们就称 σ' **比 σ 细**并记作 $\sigma'\prec\sigma$。对于 $\sigma'\prec\sigma$，我们还说 σ' 是 σ 的**加细**。特别地，任给两个 σ_1 和 σ_2，我们用 $\sigma_1\cup\sigma_2$ 表示把它们两个分割点放到一起所对应的分划，这是 σ_1 和 σ_2 共同的加细。和明显，$(\mathcal{S},\prec)$ 满足下面的三条性质：

1) 如果 $\sigma\prec\sigma'$，$\sigma'\prec\sigma$，那么 $\sigma=\sigma'$；

2) 如果 $\sigma\prec\sigma'$，$\sigma'\prec\sigma''$，那么 $\sigma\prec\sigma'$；

3) 对任意的 $\sigma,\sigma\in\mathcal{S}$，一定存在 σ''，使得 $\sigma''\prec\sigma$ 并且 $\sigma''\prec\sigma'$。

在定义积分之前，我们先做如下的注解：

注记. 在下面关于积分的构造过程中，尽管我们只考虑实值函数，但是大部分的理论对函数 $f:I\to V$ 都成立，其中 V 是一个赋范线性空间。在应用的时候，$V=\mathbb{C}$ 或者 $\mathbf{M}_n(\mathbb{R})$ 是重要的。另外，我们注意到 $\mathbf{M}_n(\mathbb{R})$ 的时候两个函数还可以相乘，此时和乘积有关的定理也都成立。

定义 113. 给定函数 $f:I\to\mathbb{R}$，如果存在一个分划 $\sigma=\{a=a_0<a_1<\cdots<a_{n-1}<a_n=b\}$，使得 f 在每个开区间 (a_i,a_{i+1}) 上面都是常值，我们就称 f 是**阶梯函数**或者**简单函数**。我们将 I 上阶梯函数的全体记作 $\mathcal{E}(I)$。

注记. 首先，函数在分割点 a_i 的值可以任意选取，这对于后面定义积分是无关紧要的。

其次，对给定的 $f\in\mathcal{E}(I)$，可能存在另一个分划 $\sigma'=\{a=a'_0<a'_1<\cdots<a'_{m-1}<a'_m=b\}$，使得 f 在每个开区间 (a'_j,a'_{j+1}) 上面都是常值。

引理 114. 给定有界闭区间 I，它上面的阶梯函数空间 $\mathcal{E}(I)$ 满足如下的性质：

1) $\mathcal{E}(I)$ 是 $\mathbb{R}$-线性空间。（如果 f 在一个 $\mathbb{C}$-线性空间中取值，那么 $\mathcal{E}(I)$ 是 $\mathbb{C}$-线性空间）

2) 对任意的 $f,g\in\mathcal{E}(I)$，$fg\in\mathcal{E}(I)$。（如果 $V=\mathbb{C}$ 或者 $\mathbf{M}_n(\mathbb{C})$，这个结论仍然成立）

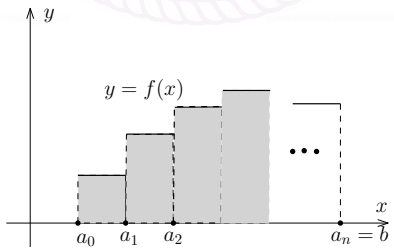
3) 任意给定映射 $\varphi:\mathbb{R}\to\mathbb{R}$，那么复合函数 $\varphi\circ f$ 是阶梯函数。（对于在赋范线性空间 V 中取值的阶梯函数 f 和任意的赋范线性空间之间的映射 $V\varphi:V\to V'$，它们的复合仍然是阶梯函数）

4) 假设 $f\in\mathcal{E}(I)$，那么 $|f|:x\mapsto|f(x)|$ 是阶梯函数。（对于在赋范线性空间 $(V,\|\cdot\|)$ 中取值的阶梯函数 f，我们就考虑 $\|f\|$）

证明: 4) 是 3) 的推论（与绝对值函数复合），而 3) 证明是显然的。我们现在来证明 1)，2) 的证明如出一辙，我们将会略去。

为了证明 $\mathcal{E}(I)$ 是 $\mathbb{R}$-线性空间，我们说明如果 $f_1,f_2\in\mathcal{E}(I)$，那么 $f_1+f_2\in\mathcal{E}(I)$，其余的关于线性空间的公理类似可以证明：假设 σ_1 和 σ_2 是和 f_1 以及 f_2 相对应的分划，我们用 $\sigma=\sigma_1\cup\sigma_2$ 表示把它们两个分割点放到一起所对应的分划并假设 $\sigma=\{a=a_0<a_1<\cdots<a_{n-1}<a_n=b\}$。很明显，$f_1$ 和 f_2 在每一个 (a_i,a_{i+1}) 上都是常数，所以 f_1+f_2 在每一个 (a_i,a_{i+1}) 上都是常数，其中 $i=0,1,\cdots,n-1$。 □

我们现在来定义 $f\in\mathcal{E}(I)$ 的函数图像所围出的面积（可以有符号）：



假设 f 是与 $\sigma\in\mathcal{S}$ 相容的阶梯函数，其中 $\sigma=\{a=a_0<a_1<\cdots<a_{n-1}<a_n=b\}$ 令 $f\big|_{(a_i,a_{i+1})}\equiv f_i$，我们按照直观来定义：

$$S_\sigma(f)=(a_1-a_0)f_1+(a_2-a_1)f_2+\cdots+(a_n-a_{n-1})f_n.$$

然而，可能存在另外一个分划 $\sigma'\in\mathcal{S}$，σ' 与 f 也相容。如果 $\sigma'=\{a=a'_0<a;_1<\cdots<a'_{m-1}<a'_m=b\}$，那么 $f\big|_{(a'_i,a'_{i+1})}\equiv f'_i$，所以我们还可以如下地定义面积：

$$S_{\sigma'}(f)=(a'_1-a'_0)f_1+(a'_2-a'_1)f'_2+\cdots+(a'_m-a'_{m-1})f'_m.$$

为了说明我们的面积是良好定义的，就需要说明 $S_\sigma(f) = S_{\sigma'}(f)$。实际上，考虑这两个分划共同的加细 $\sigma \cup \sigma'$，f 与这个新的分划也相容，所以只要说明当 $\overline{\sigma} \prec \sigma$ 时，$S(f) = S'(f)$ 即可：因为 $\sigma \cup \sigma' \prec \sigma$，$\sigma \cup \sigma' \prec \sigma'$，所以

$$S_\sigma(f) = S_{\sigma\cup\sigma'}(f) = S_{\sigma'}(f).$$

这个简单的推理在之后会重复出现。

引理 115. *对于给定的阶梯函数 $f \in \mathcal{E}(I)$，假设 σ 和 σ' 都是与 f 相容的分划并且 $\sigma' \prec \sigma$，那么 $S_\sigma(f) = S_{\sigma'}(f)$。*

证明： 我们为 σ' 的分割点按照如下的方式从小到大编号：

$$a_{0,0}, a_{0,1}, \cdots, a_{0,m_0-1}; a_{1,0}, a_{1,1}, \cdots, a_{1,m_1-1}; \cdots; a_{n-1,0}, a_{n-1,1}, \cdots, a_{n-1,m_{n-1}-1}; a_{n,0},$$

其中，$a_{0,0}, a_{1,0}, \cdots, a_{n,0}$ 恰好是 σ 的分割点。我们假设 $f\big|_{(a_{k,0},a_{k+1,0})} = f_k$，其中 $k = 0, 1, \cdots, n-1$，那么

$$S_{\sigma'}(f) = \sum_{k=0}^{n-1}\sum_{\ell=0}^{m_k-1} f_k(a_{k,\ell+1} - a_{k,\ell}) = \sum_{k=0}^{n-1} f_k(a_{k+1,0} - a_{k,0}) = S_\sigma(f).$$

这就完成了证明。 □

上面的讨论，表明映射

$$\int_a^b : \mathcal{E}(I) \to \mathbb{R}, \quad f \mapsto S(f),$$

是良好定义的。我们将 $S(f)$ 记作 $\displaystyle\int_I f$ 或者 $\displaystyle\int_a^b f$ 并称它为 f 的**积分**，其中 f 是阶梯函数。

关于阶梯函数的积分，我们有如下的性质：

定理 116. *积分 $\displaystyle\int_a^b : \mathcal{E}(I) \to \mathbb{R}$ 是 $\mathbb{R}$-线性映射。进一步，它满足*

1) 对于 $f \in \mathcal{E}(I)$，我们有 $\left|\displaystyle\int_a^b f\right| \leqslant \int_a^b |f|$。

2) (区间可加性) 假设 $a < c < b$，那么对于任意的 $f \in \mathcal{E}(I)$，我们有 f 在 $[a, c]$ 和 $[c, b]$ 上的限制都是阶梯函数，并且

$$\int_a^b f = \int_a^c f + \int_c^b f.$$

证明： 我们先证明积分的线性，只需要证明如果 $f, g \in \mathcal{E}(I)$，那么

$$\int_a^b f + g = \int_a^b f + \int_a^b g$$

即可，其余性质可以类似地验证。为此，我们取 $\sigma \in \mathcal{S}$，使得 σ 与 f 和 g 都相容。假设 $\sigma = \{a = a_0 < a_1 < \cdots < a_{n-1} < a_n = b\}$，$f\big|_{(a_i,a_{i+1})} \equiv f_i$，$g\big|_{(a_i,a_{i+1})} \equiv g_i$，其中 $0 \leqslant i \leqslant n-1$，那么，

$$\begin{aligned} S_\sigma(f+g) &= \sum_{i=0}^{n-1} (f_i + g_i)(a_{i+1} - a_i) = \sum_{i=0}^{n-1} f_i(a_{i+1} - a_i) + \sum_{i=0}^{n-1} g_i(a_{i+1} - a_i) \\ &= S_\sigma(f) + S_\sigma(g) \end{aligned}$$

这就证明了线性。

为了证明 1)，我们利用距离的三角不等式：

$$\big|S_\sigma(f)\big| = \Big|\sum_{i=0}^{n-1} (a_{i+1} - a_i) f_i\Big| \leqslant \sum_{i=0}^{n-1} (a_{i+1} - a_i)\big|f_i\big| = S_\sigma\big(|f|\big).$$

为了证明 2)，我们可以选取分划 $\sigma = \{a = a_0 < a_1 < \cdots < a_{n-1} < a_n = b\}$，使得 $c = a_{i_0}$ 为某一个分割点，此时，

$$\begin{aligned} S_\sigma(f) &= \sum_{i=0}^{n-1} f_i(a_{i+1} - a_i) = \sum_{i=0}^{i_0-1} f_i(a_{i+1} - a_i) + \sum_{i=i_0}^{n-1} f_i(a_{i+1} - a_i) \\ &= \int_a^c f + \int_c^b f. \end{aligned}$$

命题成立。 □

关于阶梯函数的积分，我们还有如下的性质：

命题 117. 对于 $f \in \mathcal{E}(I)$，如果除去有限个点之外，$f \geqslant 0$，我们就称 f 是正的阶梯函数。我们有如下的性质：

1) 假设 $f \in \mathcal{E}(I)$ 是正的阶梯函数，那么 $\displaystyle\int_a^b f \geqslant 0$。

2) 假设 $f, g \in \mathcal{E}(I)$ 使得 $f \geqslant g$，那么 $\displaystyle\int_a^b f \geqslant \int_a^b g$。

3) 对任意的 $f \in \mathcal{E}(I)$，我们有如下的估计：

$$\Big\| \int_a^b f \Big\| \leqslant |b-a| \|f\|_{L^\infty(I)},$$

其中任取与 f 相容的分划 $\sigma = \{a = a_0 < a_1 < \cdots < a_{n-1} < a_n = b\}$，假设 $f\big|_{(a_{i-1},a_i)} = f_i$，$i = 1, \cdots, n$，我们定义

$$\|f\|_{L^\infty(I)} = \sup_{1 \leqslant i \leqslant n} |f_i|.$$

特别地，如果我们只改动 f 在有限个点处的值，那么 $\|f\|_{L^\infty(I)}$ 不发生变化。

证明: 按照定义，1) 是显然的；2) 是 1) 和积分线性的推论。为了证明 3)，我们可以选取分划 $\sigma = \{a = a_0 < a_1 < \cdots < a_{n-1} < a_n = b\}$ 与 f 相容，那么

$$\begin{aligned} |S_\sigma(f)| &= \Big|\sum_{i=0}^{n-1} f_i(a_{i+1} - a_i)\Big| \leqslant \sum_{i=0}^{n-1} \|f\|_{L^\infty(I)}(a_{i+1} - a_i) \\ &= \|f\|_{L^\infty(I)}(b - a). \end{aligned}$$

证明完毕。 □

我们对阶梯函数这一类函数定义了积分。这样的函数比较特殊，我们想尽量扩大可以定义积分的函数的类，比如说，要包含连续函数类，使得我们仍然能够定义它们图像像下的面积。最基本的想法是利用阶梯函数来逼近这些可以积分的函数。能够被阶梯函数在好的意义下逼近的函数将会被称作是 Riemann 可积的函数。我们的处理方式和传统的直接用 Riemann 和或 Darboux 上下和的定义方式有所差别（我们会证明两者的等价性），然而，整个套路上和我们下学期要定义的抽象积分可以一一对应，很容易做推广。实际上，如果我们允许分划更一般一些（不仅仅是分成若干个闭区间的并），这些更一般的分划所对应的阶梯函数也会更一般一些，同样的处理方式（逼近）就给出了 Lebesgue 的积分理论。另外，我们指出，上面关于阶梯函数的定义并不依赖于所谓的面积（目前我们还没有定义什么叫做面积）。

为了定义 Riemann 积分，我们需要一个技术性的引理（定义）：

引理 118. *$I = [a, b]$ 是有界闭区间，$f : I \to \mathbb{R}$ 是函数，如下命题是等价的：*

1) 对任意的 $\varepsilon > 0$，存在两个阶梯函数 $F_\varepsilon : I \to \mathbb{R}$ 和 $\Psi_\varepsilon : I \to \mathbb{R}$，使得对任意的 $x \in I$，都有

$$\big|f(x) - F_\varepsilon(x)\big| < \Psi_\varepsilon(x),$$

并且

$$\int_I \Psi_\varepsilon < \varepsilon.$$

2) 存在两个阶梯函数的序列 $\{f_n\}_{n\geqslant 1} \subset \mathcal{E}(I)$ 和 $\{\psi_n\}_{n\geqslant 1} \subset \mathcal{E}(I)$，使得对任意的 $x \in I$，我们都有

$$\big|f(x) - f_n(x)\big| < \psi_n(x),$$

并且

$$\lim_{n\to\infty} \int_I \psi_n = 0.$$

用 $\varepsilon - \delta$ 语言描述函数在一点的连续性与用序列来描述函数在一点的连续性是等价的，这个引理的描述与此相似。

证明: 1)⇒2) 是显然的，因为对每个 $\varepsilon = \dfrac{1}{n}$，我们可以选取阶梯函数 $f_n = F_\varepsilon : I \to \mathbb{R}$ 和 $\psi_n = \Psi_\varepsilon : I \to \mathbb{R}$，使得对任意的 $x \in I$，都有

$$\big|f(x) - f_n(x)\big| < \Psi_\varepsilon(x)$$

并且

$$\int_I \psi_n < \varepsilon.$$

反过来，假设 2) 成立，我们证明 1)：按照定义，对于任意给定的 ε，存在 N，使得当 $n \geqslant N$ 时，我们有

$$\int_I \psi_n < \varepsilon.$$

我们就选取 $F_\varepsilon = f_N$，$\Psi_\varepsilon = \psi_N$。 □

定义 119. 如果函数 f 满足上述引理中的条件之一，我们通常称它可以被阶梯函数或简单函数逼近，我们就说 f 是**区间 I 上 Riemann 可积的函数**。我们用 $\mathcal{R}(I)$ 表示区间 I 上 *Riemann* 可积函数的全体。

注记. 如果我们在这里考虑向量值的函数 $f : I \to V$，我们通常需要假设 V 是完备的赋范线性空间以避免各种不收敛的因素。

对于 $f \in \mathcal{R}(I)$，根据定义，我们任意选取上述引理中的一列逼近函数 $\{f_n\}_{n\geqslant 1}$。我们定义它的积分为：

$$\int_a^b f = \lim_{n\to\infty} \int_a^b f_n.$$

我们首先证明上面的极限存在：根据

$$|f_n(x) - f_m(x)| \leqslant |f(x) - f_n(x)| + |f(x) - f_m(x)| \leqslant |\psi_m(x)| + |\psi_n(x)|.$$

根据阶梯函数的积分性质，我们有

$$|\int_a^b f_n - \int_a^b f_m| \leqslant \int_a^b |f_n - \int_a^b f_m| \leqslant \int_a^b \psi_n + \int_a^b \psi_m \to 0.$$

从而，$\{\int_a^b f_n\}_{n\geqslant 0}$ 是 Cauchy 列，所以极限存在。

为了证明 $\int_a^b f$ 是良好定义的，我们再来说明它实际上不依赖于逼近序列 $\{(f_n, \psi_n)\}_{n\geqslant 1}$ 的选取。我们假设另有 $f_n' : I \to \mathbb{R}$ 和 $\psi_n' : I \to \mathbb{R}$，使得对任意的 $x \in I$，$|f(x) - f_n(x)| < \psi_n'(x)$ 并且 $\lim\limits_{n\to\infty} \int_I \psi_n' = 0$，那么，

$$|f_n(x) - f_n'(x)| \leqslant |f(x) - f_n(x)| + |f(x) - f_{n'}(x)| \leqslant \psi_n(x) + \psi_n'(x).$$

从而，

$$|\lim_{n\to\infty} \int_a^b f_n - \lim_{n\to\infty} \int_a^b f_n'| \leqslant \lim_{n\to\infty} \int_a^b \psi_n + \int_a^b \psi_n' = 0.$$

定义 120 (Riemann 积分的定义)**.** 根据上面的证明，我们可以定义积分：

$$\int_I = \int_a^b : \mathcal{R}(I) \to V, \quad f \mapsto \lim_{n\to\infty} \int_a^b f_n.$$

我们来研究 Riemann 可积函数空间的 $\mathcal{R}(I)$ 的基本性质：

1) $\mathcal{E}(I) \subset \mathcal{R}(I)$。

对于阶梯函数 $f \in \mathcal{E}(I)$，我们可以选取 $f_n = f$，$\psi_n \equiv 0$，从而满足 Riemann 可积分函数定义中的要求。特别地，它的 Riemann 积分就是 f 作为阶梯函数的积分。

2) 假设 $f \in \mathcal{R}(I)$，那么 f 是有界函数。

利用 Riemann 可积函数中的等价定义 1)：取 $\varepsilon = 1$，此时存在两个阶梯函数 $F_1 : I \to \mathbb{R}$ 和 $\Psi_1 : I \to \mathbb{R}$，使得对任意的 $x \in I$，都有

$$\big|f(x) - F_1(x)\big| < \Psi_1(x).$$

由于阶梯函数都有界，所以

$$\big|f(x)\big| \leqslant |f(x) - F_1(x)\big| + \big|F_1(x)\big| < \Psi_1(x) + \big|F_1(x)\big|.$$

是有界的。

3) $C(I) \subset \mathcal{R}(I)$。

假设 $f \in C(I)$，根据 f 的一致连续性，对任意的 $\varepsilon > 0$，存在 $n \in \mathbb{Z}_{\geqslant 1}$，使得对任意的 $x, y \in I$，当 $|x - y| < \dfrac{1}{n}$ 时，我们有

$$|f(x) - f(y)| < \frac{\varepsilon}{b-a}.$$

此时，我们令

$$F(x) = \sum_{k=1}^{n} f\big(a + k\frac{b-a}{n}\big)\mathbf{1}_{[a+(k-1)\frac{b-a}{n}, a+k\frac{b-a}{n}]}(x),$$

其中，$\mathbf{1}_{[a+(k-1)\frac{b-a}{n}, a+k\frac{b-a}{n}]}$ 是示性函数。这个 $F(x)$ 显然是阶梯函数。由一致连续性，我们知道

$$\big|f(x) - F(x)\big| < \Psi_(x) \equiv \frac{\varepsilon}{b-a}.$$

所以，$\displaystyle\int_a^b \Psi_x = \varepsilon$。从而，$f$ 是 Riemann 可积的函数。

4) $\mathcal{R}(I)$ 是 $\mathbb{R}$-线性空间。

我们来证明如果 $f, g \in \mathcal{R}(I)$，那么 $f + g \in \mathcal{R}(I)$，其余的关于线性空间的性质可以类似地验证。按照定义，存在阶梯函数的序列 (f_n, ψ_n) 和 (g_n, χ_n)，使得

$$|f(x) - f_n(x)| < \psi_n(x), \quad \lim_{n\to\infty} \int_I \psi_n = 0,$$
$$|g(x) - g_n(x)| < \chi_n(x), \quad \lim_{n\to\infty} \int_I \chi_n = 0.$$

现在考虑阶梯函数序列 $(f_n+g_n)_{n\geqslant 1}$，由于 $f(x)$ 有界，我们有

$$\begin{aligned}\big|\big(f(x)+g(x)\big)-\big(f_n(x)+g_n(x)\big)\big| &\leqslant \big|f(x)-f_n(x)\big|+\big|g_n(x)-g(x)\big|\\ &\leqslant \psi_n(x)+\chi_n(x).\end{aligned}$$

很明显，$\displaystyle\int_I \psi_n+\chi_n\to 0$。所以我们可以选取 $(f_n+g_n,\psi_n+\chi_n)$ 作为 $f+g$ 的逼近序列。

5) 如果 $f,g\in\mathcal{R}(I)$，那么 $fg\in\mathcal{R}(I)$（对于在 V 中取值的 Riemann 可积函数也成立，其中 $V=\mathbb{C}$ 或者 $\mathbf{M}_n(\mathbb{C})$）

这个性质的证明与上面的类似，但是需要做技术上的改动：对于 $f,g\in\mathcal{R}(I)$，存在阶梯函数的序列 (f_n,ψ_n) 和 (g_n,χ_n)，使得

$$\begin{aligned}&|f(x)-f_n(x)|<\psi_n(x),\quad \lim_{n\to\infty}\int_I\psi_n=0,\\ &|g(x)-g_n(x)|<\chi_n(x),\quad \lim_{n\to\infty}\int_I\chi_n=0.\end{aligned}$$

现在考虑阶梯函数序列 $(f_ng_n)_{n\geqslant 1}$。由于 f 是 Riemann 可积的函数，所以它有界，假设对任意的 $x\in I$，我们有 $|f(x)|\leqslant M$，从而

$$\begin{aligned}|f(x)g(x)-f_n(x)g_n(x)| &\leqslant |f(x)-f_n(x)||g_n(x)|+|f(x)||g_n(x)-g(x)|\\ &\leqslant |\psi_n(x)||g_n(x)|+M|\chi_n(x)|.\end{aligned}$$

我们想选取阶梯函数

$$\vartheta_n(x)=|\psi_n(x)||g_n(x)|+M|\chi_n(x)|$$

作为逼近中的控制函数。然而，由于 $g_n(x)$ 还是依赖于 n 的，它不容易被控制住。我们观察到，如果是一开始就知道 $|g_n(x)|\leqslant N$，那么 $\vartheta_n(x)$ 的积分就会有如下的控制：

$$\int_I\vartheta_n\leqslant\int_I N\psi_n(x)+M\chi_n(x)\to 0.$$

从而，命题得到证明。

为了保证 g_n 有界，我们在一开始选择逼近序列的 $\{g_n\}_{n\geqslant 1}$ 时候就要加以限制：令 $N=\sup\limits_{x\in I}|g(x)|$，定义

$$\overline{g_n}(x)=\begin{cases}g_n(x), & \text{如果}\chi_n(x)\leqslant N;\\ 0, & \text{如果}\chi_n(x)>N.\end{cases}$$

这是一列阶梯函数（容易验证）。此时，我们定义

$$\overline{\chi_n}(x)=\begin{cases}\chi_n(x), & \text{如果}\chi_n(x)\leqslant N;\\ N, & \text{如果}\chi_n(x)>N.\end{cases}.$$

很明显，我们有 $|g_n(x)| \leqslant 2N$ 并且 $|g(x) - \overline{g_n}(x)| \leqslant \overline{\chi_n}(x)$。由于 $\overline{\chi_n} \leqslant \chi_n$，所以 $\int_I \overline{\chi_n} \to 0$（利用关于阶梯函数的积分的不等式）。我们用 $(\overline{g_n}, \overline{\chi_n})$ 来代替原来的 (g_n, χ_n)，其中 $\overline{g_n}$ 是一致有界的。

（如果 $f : I \to \mathbb{R}$，$g : I \to V$ 是 Riemann 可积的函数，其中 V 是某个赋范线性空间，上面的推理也成立）

6) 假设 $f \in \mathcal{R}(I)$，那么 $|f| : x \mapsto |f(x)|$ 也是 Riemann 可积的函数。

对于 $f \in \mathcal{R}(I)$，存在阶梯函数的序列 (f_n, ψ_n) 使得

$$|f(x) - f_n(x)| < \psi_n(x), \quad \lim_{n\to\infty} \int_I \psi_n = 0.$$

我们令 $F_n(x) = |f_n(x)|$，此时，

$$\big||f(x)| - F_n(x)\big| \leqslant |f(x) - f_n(x)| < \psi_n(x).$$

所以，我们选取阶梯函数的序列 (F_n, ψ_n) 来逼近 $|f|$ 即可。

7) 我们考虑 f 在有限维线性空间中取值的情况：$f : I \to \mathbb{R}^n, x \mapsto (f_1(x), \cdots, f_n(x))$，其中 f_i 是 f 的每个分量。那么，f 是 Riemann 可积分的当且仅当对每个分量 f_i，$f_i \in \mathcal{R}(I)$，其中 $i = 1, 2, \cdots, n$。在此情形下，我们还有

$$\int_a^b f = \left(\int_a^b f_1, \cdots, \int_a^b f_n\right).$$

特别地，当 $f : \mathbb{R} \to \mathbb{C}$ 是复值函数时，我们有

$$\int_a^b f = \int_a^b \Re f + i \int_a^b \Im f,$$

其中 $\Re f$ 和 $\Im f$ 分别为 f 的实部和虚部。

定理 121. *积分 $\int_a^b : \mathcal{R}(I) \to \mathbb{R}$ 满足下面的性质：*

1) $\int_a^b : \mathcal{R}(I) \to V$ *是线性映射。（显然，直接对逼近序列进行线性操作即可）*

2) 对于 $f \in \mathcal{R}(I)$，假设 $\varphi : V \to V'$ 是连续线性映射，那么 $\int_a^b \varphi \circ f = \varphi(\int_a^b f)$。

3)（三角不等式）对于 $f \in \mathcal{R}(I)$，我们有 $\| \int_a^b f\| \leqslant | \int_a^b f|$。

4)（对区间的可加性）假设 $a < c < b$，那么对于任意的 $f \in \mathcal{R}(I)$，我们有 f 在 $[a, c]$ 和 $[c, b]$ 上的限制都是阶梯函数，并且

$$\int_a^b f = \int_a^c f + \int_c^b f.$$

证明: 2) 目前我们可以假设 V 和 V' 都是有限维的，φ 是线性映射，那么对于任意的 $v \in V$，我们都有 $\|\varphi(v)\| \leqslant M\|v\|$（为什么?)。此时，假设 (f_n, ψ_n) 是 f 的逼近序列，我们只要取 $(\varphi \circ f_n, M\psi_n)$ 即可。 □

最后，大家可以通过设想如何定义高维的积分仔细体会 Riemann 积分的定义。我们首先要定义所谓的阶梯函数，这个依赖于如何定义最基本的分划：1 维的时候我们用闭区间来分割区域，2 维的时候，我们可能可以用小长方体来分割整个区域。下一步，我们要求阶梯函数在这样的长方体上面是常数（在 Lebesgue 积分的理论中，长方体将会被换成是可测集合)。然而，在 2 维的时候，没有几个区域可以被分解为有限个方体的并，这是和 1 维积分不一样的。高维几何的复杂程度导致了积分理论的困难。

19 Riemann 和，Darboux 上下和，上下积分

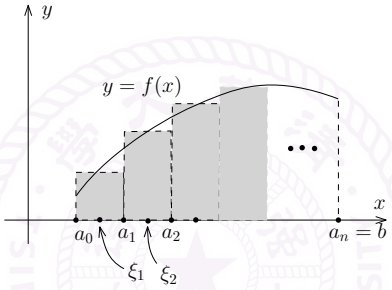
二零一九年十一月二十一日，星期四，阴

我们介绍如何利用经典的 Riemann 和以及 Darboux 上下和的观点来研究 Riemann 可积的函数。值得强调的是，Riemann 和（以及上节课用阶梯函数来定义积分的方式）对于在赋范线性空间中取值的函数可以定义，然而 Darboux 上下和只对实数值的函数可行（原因是我们需要用到实数域上的序结构）。

任意给定分划 $\sigma \in \mathcal{S}$，假设 $\sigma = \{a = a_0 < a_1 < \cdots < x_n - 1 < a_n = b\}$，我们在每一小段上面选取 $\xi = (\xi_1, \cdots, \xi_n)$，使得 $\xi_i \in [a_{i-1}, a_i]$。为了简单起见，我们用 ξ 表示这些 ξ_i，用 (σ, ξ) 表示给定了分划 σ 并在每一段中选好 ξ。我们将 (σ, ξ) 的全体记作 $\mathcal{S}' = \mathcal{S}'(I)$。

给定函数 $f : I \to \mathbb{R}$ 和 $(\sigma, \xi) \in \mathcal{S}'$，我们定义所谓的 Riemann 和

$$S(f; \sigma, \xi) = (a_1 - a_0)f(\xi_i) + (a_2 - a_1)f(\xi_i) + \cdots + (a_n - a_{n-1})f(\xi_n).$$



我们证明，对于 Riemann 可积的函数 f，在正确的意思下，有

$$\lim_{|\sigma| \to 0} S(f; \sigma, \xi) = \int_a^b f.$$

定理 122. *假设 $I \subset \mathbb{R}$ 是有界闭区间，$f \in \mathcal{R}(I)$。对任意的 $\varepsilon > 0$，存在 $\delta > 0$，使得对任意的 $(\sigma, \xi) \in \mathcal{S}'(I)$，如果其步长 $|\sigma| < \delta$，那么*

$$\left| S(f; \sigma, \xi) - \int_a^b f \right| < \varepsilon.$$

证明： 根据 Riemann 可积的定义，我们选定阶梯函数 $F \in \mathcal{E}(I)$ 和 $\psi \in \mathcal{E}(I)$，使得 $|f(x) - F(x)| \leqslant \psi(x)$ 并且 $\displaystyle\int_a^b \psi(x) < \frac{\varepsilon}{4}$。我们进一步假设 F 和 ψ 都对应着分划 $\sigma_0 = \{b_0 < b_1 < \cdots < b_m\}$，即它们在每一段 (b_i, b_{i+1}) 上的限制都是常数。

现在，任意选取分划 $\sigma = \{a = a_0 < a_1 < \cdots < x_n - 1 < a_n = b\}$，我们要求 $|\sigma|$ 比 σ_0 中最短区间的长度还要小。利用积分的线性，我们有

$$\left| S(f; \sigma, \xi) - \int_a^b f \right| \leqslant \sum_{i=1}^{n} \left| \int_{a_{i-1}}^{a_i} f(x) - f(\xi_i) \right|$$

我们想逐项地控制上面右端的的每一个积分项 $\left| \displaystyle\int_{a_{i-1}}^{a_i} f(x) - f(\xi_i) \right|$。为此，我们需要讨论两个分划 σ_0 和 σ 之间的相对关系。对每个 i，我们有两种情况：

- 情形 1：$[a_{i-1}, a_i]$ 中不包含 σ_0 的任何一个分割点 b_k。

  此时，ψ 和 F 在 (a_{i-1}, a_i) 都是常数。所以，对任意的 $x \in (a_{i-1}, a_i)$ 上，我们有

$$\begin{aligned} |f(x) - f(\xi_i)| &\leqslant |f(x) - F(x)| + |F(x) - f(\xi_i)| \\ &= |f(x) - F(x)| + |F(\xi_i) - f(\xi_i)| \\ &\leqslant \psi(x) + \psi(\xi_i) = 2|\psi(x)|. \end{aligned}$$

  所以，我们有

$$\Big| \int_{a_{i-1}}^{a_i} f(x) - f(\xi_i) \Big| \leqslant 2 \int_{a_{i-1}}^{a_i} \psi.$$

- 情形 2：$[a_{i-1}, a_i]$ 中包含某个 σ_0 的某一个分割点 b_{k_0}。（我们注意到，由于 $|\sigma|$ 比 σ_0 中的最短区间的长度还小，所以 $[a_{i-1}, a_i]$ 恰好只包含 b_{k_0} 的这一个 σ_0 中的分割点。）这样的区间的个数不会超过 m 个，因为 σ_0 的分割点一共就 m 个。

  此时，我们用 $\|f\|_\infty$ 来控制 $f(x)$，从而有

$$\begin{aligned} \Big| \int_{a_{i-1}}^{a_i} f(x) - f(\xi_i) \Big| &\leqslant |a_i - a_{i-1}| \|f\|_\infty + |a_i - a_{i-1}| |f(\xi_i)| \\ &\leqslant 2|\sigma| \|f\|_\infty. \end{aligned}$$

综合上面的讨论，当 $|\sigma| < |\sigma_0|$ 时，我们有

$$\begin{aligned} \Big| S(f;\sigma,\xi) - \int_a^b f \Big| &\leqslant \underbrace{2 \sum \int_{a_{i-1}}^{a_i} \psi}_{\text{第一种情况}} + \underbrace{2m|\sigma| \|f\|_\infty}_{\text{第二种情况}} \\ &\leqslant 2 \int_a^b \psi + 2m|\sigma| \|f\|_\infty \\ &\leqslant \frac{1}{2}\varepsilon + 2m|\sigma| \|f\|_\infty. \end{aligned}$$

由于 m 是固定的，当 $|\sigma|$ 足够小的时候，上面的式子明显小于 ε，证明完毕。 □

注记. 我们们通常将定理写作 $\lim\limits_{|\sigma| \to 0} S(f;\sigma,\xi) = \int_a^b f$，其中 $f \in \mathcal{R}(I)$。实际上，任意选取一个序列 $\{(\sigma_n, \xi_n)\}_{n \geqslant 1} \subset \mathcal{S}'(I)$，上面的定理表明只要当 $|\sigma_n| \to 0$，我们就有 $\lim_{n\to} S(f;\sigma_n,\xi_n) = \int_a^b f$。根据这个定理，我们可以选择方便计算的 *Riemann* 和来计算积分。另外，这个定理还提供了一个很重要的观点用来计算极限：如果某个极限过程是某个 *Riemann* 和取极限的话，那么我们可以利用积分算极限（当然，这需要我们有足够好的手段来计算积分）。

下面来定义所谓的上积分和下积分的概念。假设 $f : I \to \mathbb{R}$ 是在 $\mathbb{R}$ 中取值的有界函数，我们要利用 $\mathbb{R}$ 中序关系。定义

$$\begin{aligned} \overline{\mathcal{E}_f}(I) &= \{\varphi \in \mathcal{E}(I) | \varphi(x) \geqslant f(x)\}, \\ \underline{\mathcal{E}_f}(I) &= \{\varphi \in \mathcal{E}(I) | \varphi(x) \leqslant f(x)\}. \end{aligned}$$

由于 f 是有界函数，所以这两个集合是非空的：假设对任意的 $x \in I$，$-M \leqslant f(x) \leqslant M$，那么常值函数 $M \in \overline{\mathcal{E}_f}(I)$，常值函数 $-M \in \underline{\mathcal{E}_f}(I)$。据此，我们定义

$$\overline{\int_a^b} f = \inf_{\varphi \in \overline{\mathcal{E}_f}(I)} \int_a^b \varphi, \quad \underline{\int_a^b} f = \sup_{\varphi \in \underline{\mathcal{E}_f}(I)} \int_a^b \varphi.$$

由于对任意的 $\varphi \in \overline{\mathcal{E}_f}(I)$，$\varphi(x) \geqslant -M$，所以 $\int_a^b \varphi \geqslant -M(b-a)$，这说明 $\overline{\int_a^b} f$ 是良好定义的；类似地，$\underline{\int_a^b} f$ 也是良好定义的。这两个式子分别被称作是有界函数 f 的**上积分**和**下积分**。按照定义，我们有如下平凡的不等式：

$$\overline{\int_a^b} f \geqslant \underline{\int_a^b} f.$$

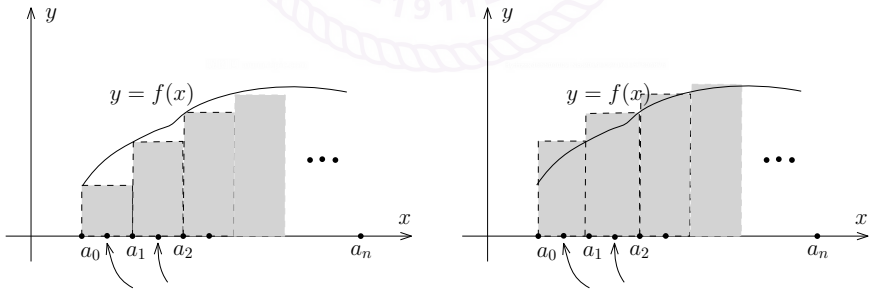
现在引入 Darboux 上下和的概念：假设 $f : I \to \mathbb{R}$ 是有界函数，$\sigma \in \mathcal{S}(I)$ 是一个分划，其中 $\sigma = \{a = a_0 < a_1 < \cdots < a_{n-1} < a_n = b\}$。根据 f 有界，对每个 $i = 1, 2, \cdots, n$，我们令

$$M_i = \sup_{x \in [a_{i-1}, a_i]} f(x), \quad m_i = \inf_{x \in [a_{i-1}, a_i]} f(x),$$

据此，可以定义

$$\overline{S}(f; \sigma) = \sum_{i=1}^n (a_i - a_{i-1}) M_i, \quad \underline{S}(f; \sigma) = \sum_{i=1}^n (a_i - a_{i-1}) m_i.$$

我们把这两个和式称作是有界函数 f 对于分划 σ 的 **Darboux 上和**和 **Darboux 下和**，我们可以用下面的示意图来表示：



上图中，左边代表的是 Darboux 下和，右边代表的是 Darboux 上和。很明显，我们有

$$\overline{S}(f; \sigma) \geqslant \underline{S}(f; \sigma).$$

实际上，我们有更强的结论：

引理 123. *对于任意两个分划 $\sigma, \sigma' \in \mathcal{S}(I)$，我们有*

$$\underline{S}(f; \sigma) \leqslant \overline{S}(f; \sigma').$$

证明：事实上，我们只需要证明当 $\sigma \prec \sigma'$ 时，我们有

$$\underline{S}(f;\sigma) \leqslant \overline{S}(f;\sigma'), \quad \underline{S}(f;\sigma') \leqslant \overline{S}(f;\sigma),$$

即可。我们把这个性质的证明留作作业。 □

另外，当 $\sigma \prec \sigma'$ 时，我们有

$$\overline{S}(f;\sigma) \leqslant \overline{S}(f;\sigma'), \quad \underline{S}(f;\sigma) \geqslant \underline{S}(f;\sigma').$$

这个的证明也是显然的。

我们定义

$$\overline{D}(f) = \inf_{\sigma\in\mathcal{S}(I)} \overline{S}(f;\sigma), \quad \underline{D}(f) = \sup_{\sigma\in\mathcal{S}(I)} \underline{S}(f;\sigma).$$

下面的命题将 Darboux 上下和与上下积分联系起来：

命题 124. *假设 $f : I \to \mathbb{R}$ 是有界函数，那么*

$$\overline{D}(f) = \overline{\int_a^b} f, \quad \underline{D}(f) = \underline{\int_a^b} f.$$

证明：根据对称性，我们证明 $\underline{D}(f) = \underline{\int_a^b} f$ 即可。证明分为两步。

- 按照定义，对任意的 σ，我们定义函数 $\phi\big|_{[a_{i-1},a_i)} = m_i$ $(1 \leqslant i \leqslant n)$，$\phi(b) = m_n$，从而 $\phi \in \underline{\mathcal{E}_f}(I)$。另外，我们有 $\underline{S}(f;\sigma) = \int_a^b \phi$。按照下积分的定义，我们有 $\int_a^b \phi \leqslant \underline{\int_a^b} f$，从而 $\underline{S}(f;\sigma) \leqslant \underline{\int_a^b} f$。再用 $\underline{D}(f)$ 的定义（对 σ 取上确界），我们得到

  $$\underline{D}(f) \leqslant \underline{\int_a^b} f.$$

- 其次，对任意选定的 ε，按照下极限的定义，我们可以选取 $\psi \in \underline{\mathcal{E}_f}(I)$（其中 $\sigma_0 = \{b_0 < b_1 < \cdots < b_\ell\}$ 是相应的分划），使得

  $$\underline{\int_a^b} f \leqslant \varepsilon + \int_a^b \psi.$$

  按照定义，我们知道 $\psi(x) \leqslant f(x)$。我们现在任意选取 I 的分划 $\sigma = \{a_0 < a_1 < \cdots < a_n\}$，我们想要比较 $\int_a^b \psi$ 和 $\underline{S}(f;\sigma)$：

  $$\int_a^b \psi - \underline{S}(f;\sigma) = \sum_{i=1}^n \int_{a_{i-1}}^{a_i} (\psi - m_i)$$

  现在有两个分划 σ_0 和 σ，我们需要讨论它们之间的相对关系。假设 i 是固定的，那么我们有两种情况：

- 情形 1：$[a_{i-1},a_i]$ 中不包含 σ_0 的任何一个分割点 b_k。此时，ψ 在 $[a_{i-1},a_i]$ 的取值是常数。特别地，根据 $\psi(x)\leqslant f(x)$，我们知道 $\psi(x)\leqslant \inf\limits_{[a_{i-1},a_i]} f(x)=m_i$，从而

$$\int_{a_{i-1}}^{a_i}(\psi(x)-m_i)\leqslant 0.$$

- 情形 2：$[a_{i-1},a_i]$ 中包含某个 σ_0 的（恰好一个，如果我们假设 $|\sigma|$ 足够小）一个分割点 b_{k_0}。此时，我们用 $M=\sup\limits_{x\in I} f(x)$ 来控制 $\psi(x)$，从而有

$$\Big|\int_{a_{i-1}}^{a_i}\psi(x)-m_i\Big|\leqslant (a_i-a_{i-1})(M-m)\leqslant |\sigma|(M-m).$$

  其中 $m=\inf\limits_{x\in I} f(x)$。

综合上述两种情况，我们得到

$$\int_a^b \psi-\underline{S}(f;\sigma)\leqslant \underbrace{0}_{\text{第一种情况}}+\underbrace{2\overbrace{\ell}^{\text{第二种区间的个数}}|\sigma|(M-m)}_{\text{第二种情况}}.$$

只要取得 $|\sigma|$ 足够小，我们就得到

$$\int_a^b\psi\leqslant \underline{S}(f;\sigma)+\varepsilon\leqslant \underline{D}(f)+\varepsilon,$$

即

$$\underline{\int_a^b} f\leqslant \underline{D}(f)+2\varepsilon.$$

令 $\varepsilon\to 0$，我们得到

$$\underline{D}(f)\geqslant \underline{\int_a^b} f.$$

命题得到了证明。 □

注记. 课后，有同学讲第二步可以很简单地证明：

对任意的 $\varphi\in\underline{\mathcal{E}_f}(I)$，选取与 φ 相容的分划 $\sigma=\{a_0<a_1<\cdots<a_n\}$，此时，由于 $\varphi\big|_{(a_{i-1},a_i)}\equiv\varphi_i$ 是常数并且 $\varphi\leqslant f$，所以，$\varphi_i\leqslant m_i$。据此，我们知道

$$\int_a^b\varphi=\sum_{i=1}^n\varphi_i(a_i-a_{i-1})\leqslant\sum_{i=1}^n(a_i-a_{i-1})m_i=\underline{S}(f;\sigma).$$

从而，对任意的 $\varphi\in\underline{\mathcal{E}_f}(I)$

$$\int_a^b\varphi\leqslant\sup_{\sigma\in\mathcal{S}(I)}\underline{S}(f;\sigma)=\underline{D}(f).$$

再对所有可能的 φ 取上确界即可。

19.1 作业：Sturm-Louville 理论的一个例子

清华大学 19-20 秋季学期，数学分析一，作业 8

请用 A4 大小的纸张**正反面**用钢笔，签字笔或者圆珠笔书写，并注明自己的姓名，年级和作业的**总页数**。除定理公式所涉及的人名之外，请使用**中文**。本次作业的提交时间和地点为 **11 月 28 日**上午的课堂上，逾期视作零分。

基本习题

凸函数的研究

如果不另加说明，f 总代表某区间 I 上实值函数。如果对任意的 $x, y \in I$ 和 $t \in (0,1)$，都有 $f(tx+(1-t)y) < tf(x)+(1-t)f(y)$，我们就称 f 为**严格凸的**。仿照课堂上凹函数的定义，我们可以类似地定义严格凹的函数（即 $-f$ 是严格凸的）。

习题 A（凸函数的基本性质）

A1) 试判断下列函数的凸 (凹) 或者严格凸 (凹) 性：

(1) $f(x) = |x|$, $I = \mathbb{R}$ (2) $f(x) = x^p$, $p \in \mathbb{R}$, $I = \mathbb{R}_{>0}$, (3) $f(x) = \sin x$, $I = [0, \pi]$,

(4) $f(x) = x \log x$, $I = \mathbb{R}_{>0}$ (5) $f(x) = \begin{cases} 1, & x = 0, 1 \\ 0, & 0 < x < 1. \end{cases}$ $I = [0,1]$

A2) 试证明凸函数的如下基本性质：

(a) 如果 f, g 是区间 I 上的凸函数，那么 $f+g$ 也是凸函数。

(b) 如果 f, g 是区间 I 上的单调递增的非负的凸函数, 那么 fg 是凸函数。

(c) 如果 f 是区间 I 上的凸函数，g 是区间 $J \supset f(I)$ 上的单调递增凸函数, 那么 $g \circ f$ 是凸函数。

(d) 如果 f, g 是区间 I 上的凸函数, 那么 $h(x) = \max\{f(x), g(x)\}$ 是凸函数。

A3) 假设 $f \in C\big((a,b)\big)$。如果对任意的 $x, y \in (a,b)$，都有 $f\left(\dfrac{x+y}{2}\right) \leqslant \dfrac{f(x)+f(y)}{2}$，证明，$f$ 是凸函数。

A4) f 是 $[a,b]$ 上的凸函数。如果存在 $c \in (a,b)$，使得 $f(c) \geqslant \max\{f(a), f(b)\}$。试证明，$f$ 是常值函数。

A5) f 是 $\mathbb{R}$ 上的凸函数。如果 f 在 $\mathbb{R}$ 上有界，证明，f 是常值函数。

A6) 设 f 是区间 I 上的严格凸函数。假设 $f(x_0) \in I$ 是 f 的局部极小值，证明，x_0 是 f 唯一的整体极小值点，即对任意的 $x \in I - \{x_0\}$，我们有 $f(x_0) < f(x)$。

A7) I 是开区间。证明，f 是凸函数等价于对任意点 $x_0 \in I$，存在实数 $a \in \mathbb{R}$，使得对任意的 $x \in I$，我们都有 $f(x) \geqslant a(x-x_0)+f(x_0)$。

习题 B（凸函数与不等式）

B1) 证明下面的不等式：

$$(1)x-\frac{x^2}{2}<\log(1+x)<x,\quad x>0;\quad (2)(x^\alpha+y^\alpha)^{1/\alpha}>(x^\beta+y^\beta)^{1/\beta},\quad x,y>0,\beta>\alpha>0$$

$$(3)x-\frac{x^3}{6}<\sin x<x,\quad x>0;\quad (4)\left(\frac{1+x}{2}\right)^p+\left(\frac{1-x}{2}\right)^p\leqslant\frac{1}{2}(1+x^p),\quad p\geqslant 2,x\in[0,1].$$

B2) 试求所有的正数 a，使得不等式 $a^x \geqslant x^a$ 对任意的 $x>0$ 都成立。

B3) 证明如下不等式并给出等号成立的条件：对任意正数 x_i 和 t_i（$i=1,\cdots,n$），$\sum\limits_{i=1}^n t_i=1$，都有

$$\left(\sum_{i=1}^n t_i x_i\right)^{\sum\limits_{i=1}^n t_i x_i}\leqslant\prod_{i=1}^n x_i^{t_i x_i}.$$

B4) 证明，Young 不等式并给出等号成立的条件：对于任意正数 a,b，任意的实数 p,q，其中 $\frac{1}{p}+\frac{1}{q}=1$（我们要求 p 和 q 不是 0 或者 1），我们有

$$ab\leqslant\frac{a^p}{p}+\frac{b^q}{q},\ \text{如果 } p>1;\quad ab\geqslant\frac{a^p}{p}+\frac{b^q}{q},\ \text{如果 } p<1.$$

B5) 证明 Hölder 不等式并给出等号成立的条件：设 $x_i,y_i\geqslant 0$，其中 $i=1,\cdots,n$，$p,q\neq 0,1$，其中 $\frac{1}{p}+\frac{1}{q}=1$，我们有

$$\sum_{i=1}^n x_i y_i\leqslant\left(\sum_{i=1}^n x_i^p\right)^{1/p}\left(\sum_{i=1}^n y_i^q\right)^{1/q},\ \text{如果 } p>1;\ \sum_{i=1}^n x_i y_i\geqslant\left(\sum_{i=1}^n x_i^p\right)^{1/p}\left(\sum_{i=1}^n y_i^q\right)^{1/q},\ \text{如果 } p<1.$$

在上述不等式中，如果 $p<0$，我们假设 $x_i>0$。（这个不等式的结论比证明重要得多，同学们可以参考其它资料来写下证明）

导数的性质和应用：复习

习题 C（利用（高阶）导数的信息刻画函数（低阶导数））在本习题中，我们把函数 f 也记为 $f^{(0)}$。

C1) 我们假设 $f\in C\big([0,1]\big)$，g 在 $[0,1]$ 上可导且 $g(0)=0$。如果存在常数 $\lambda\neq 0$，使得对任意的 $x\in[0,1]$，都有 $|g(x)f(x)+\lambda g'(x)|\leqslant|g(x)|$, 证明，$g(x)\equiv 0$。

C2) f 在 $(-1,1)$ 上二阶可导，$f(0)=f'(0)=0$。如果对任意的 $x\in(-1,1)$，都有 $|f''(x)|\leqslant|f(x)|+|f'(x)|$，证明，$f(x)\equiv 0$。

C3) n 是正整数，f 在 $\mathbb{R}$ 上 n 阶可导，$f(0)=f'(0)=\cdots=f^{(n-1)}(0)=0$。如果 $C\in\mathbb{R}_{>0}$ 和 $\ell\in\mathbb{Z}_{\geqslant 0}$，使得对任意 $x\in\mathbb{R}$ 都有 $|f^{(n)}(x)|\leqslant C|f^{(\ell)}(x)|$，证明，$f(x)\equiv 0$。

C4) n 是正整数，证明，多项式 $P(x)=\sum\limits_{k=0}^{n+1}C_{n+1}^{k}(-1)^k(x-k)^n$ 是 0。

C5) $f\in C^\infty(\mathbb{R})$。假设存在正实数 C，使得对任意 $n\in\mathbb{Z}_{\geqslant 0}$ 和任意的 $x\in\mathbb{R}$，我们都有 $|f^{(n)}(x)|\leqslant C$。

(i) 证明，给定任意的 $x_0\in\mathbb{R}$，我们可以 (以 x_0 为中心) 将 $f(x)$ 在 $\mathbb{R}$ 上展开为无穷 Taylor 级数，即
$$f(x)=\sum_{k=0}^{\infty}\frac{f^{(k)}(x_0)}{k!}(x-x_0)^k,\qquad \forall x\in\mathbb{R}.$$

(ii) $E\subset\mathbb{R}$ 是一个无穷集并且有界。证明，如果 f 在 E 上的取值都是零，那么 $f\equiv 0$。

C6) 假设 $f\in C^2\big((0,1)\big)$，$\lim\limits_{x\to 1^-}f(x)=0$。如果存在 $C>0$，使得对任意 $x\in(0,1)$，我们都有不等式 $(1-x)^2|f''(x)|\leqslant C$。证明，$\lim\limits_{x\to 1^-}(1-x)f'(x)=0$.

习题 D（利用导数求极值）求下列函数在给定集合 I 上的极值 $\sup\limits_{x\in I}f(x)$ 和 $\inf\limits_{x\in I}f(x)$，其中 m,n 为正整数。

D1) $f(x)=\dfrac{(\log x)^2}{x}$，$I=\mathbb{R}_{>0}$; D2) $f(x)=|x(x^2-1)|$，$I=\mathbb{R}$;

D3) $f(x)=\dfrac{x(x^2+1)}{x^4-x^2+1}$，$I=\mathbb{R}$; D3) $f(x)=x^{\frac{1}{3}}(1-x)^{\frac{2}{3}}$，$I=(0,1)$

D5) $f(x)=\left(1+x+\dfrac{x^2}{2!}+\cdots+\dfrac{x^n}{n!}\right)e^{-x}$，$I=\mathbb{R}$; D6) $f(x)=\sin^{2m}x\cos^{2n}x$，$I=\mathbb{R}$。

习题 E 利用导数比较函数（或实数）大小。

E1) $f(x)=e^x$ 和 $g(x)=1+xe^x$，其中 $x>0$; E2) $f(x)=xe^{\frac{x}{2}}$ 和 $g(x)=e^x-1$，其中 $x>0$;

E3) $f(x)=\left(\frac{x+1}{2}\right)^{x+1}$ 和 $g(x)=x^x$，其中 $x>0$; E4) $2^{\sqrt{2}}$ 和 e;

E5) $f(x)=\log(1+\sqrt{1+x^2})$ 和 $g(x)=\frac{1}{x}+\log x$，其中 $x>0$; E6) $\log_e 8$ 和 2。

习题 F 如果 f 在 x_0 的一个邻域内满足 $f(x)=(x-x_0)^r g(x)$，其中 $r\in Z_{\geqslant 0}$，g 在 x_0 处连续并且 $g(x_0)\neq 0$，那么我们就称 x_0 为 f 的 r-重根。我们注意到，0-重根并非根。

F1) 假设 x_0 是 f 的 r-重根，其中 $r\geqslant 1$。证明，如果 $g(x)=\dfrac{f(x)}{(x-x_0)^r}$ 可微，那么 x_0 是 f' 的 $(r-1)$-重根。

F2) 假设 f 为 $\mathbb{R}$ 上 n 阶的可微函数。证明，如果 $f(x)=0$ 有 $n+1$ 个不同的实根，那么 $f^{(n)}(x)=0$ 至少有一个实根。

F3) f 为 $\mathbb{R}$ 上的可微函数。假设 $f(x)=0$ 按重数计算恰有 r 个实根，也就是说，$f(x)=0$ 有 s 个相异的实根 $x_1,x_2,\cdots,x_s$，它们的重数分别为 $r_1,r_2,\cdots,r_s$ 并且 $r_1+r_2+\cdots+r_s=r$。证明，$f'(x)=0$ 按重数计至少有 $r-1$ 个根。

F4) 假设 f 为 $\mathbb{R}$ 上 n 阶的可微函数。证明，如果 $f(x)=0$ 按重数计恰有 $n+1$ 个实根，那么 $f^{(n)}(x)=0$ 至少有一个实根。

习题 G 设 $a\in\mathbb{R}$, $f:(a,\infty)\to\mathbb{R}$ 在 (a,∞) 上二阶可微，并且

$$M_0:=\sup_{x\in(a,\infty)}|f(x)|,\qquad M_1:=\sup_{x\in(a,\infty)}|f'(x)|,\qquad \text{和 } M_2:=\sup_{x\in(a,\infty)}|f''(x)|$$

均为实数。

G1) 证明：$M_1^2\leqslant 4M_0M_2$. (**提示**：对任意的 $h>0$ 和任意的 $x\in(a,\infty)$, 证明, 有 $\xi\in(x,x+2h)$, 使得

$$f'(x)=\frac{1}{2h}[f(x+2h)-f(x)]-hf''(\xi)$$

成立，继而用 M_0,M_2 来控制 M_1.)

G2) 令 $a=-1$, 考察函数 $f:(-1,\infty)\to\mathbb{R}$:

$$f(x)=\begin{cases}2x^2-1 & x\in(-1,0),\\ \frac{x^2-1}{x^2+1} & x\in[0,\infty),\end{cases}$$

验证 f 二阶可微，验证 $M_0=1,M_1=4,M_2=4$. 因此，在不等式 1) 中，等号可能取到。

G3) 假设 $\mathbf{f}:(a,\infty)\to\mathbb{R}^n$ 是 (a,∞) 上二阶可微的在 $\mathbb{R}^n$ 中取值的向量值函数，记 M_0,M_1 和 M_2 分别为 $|\mathbf{f}|$, $|\mathbf{f}'|$ 和 $|\mathbf{f}''|$ 的上确界。证明：$M_1\leq 4M_0M_2$.

思考题 (8 字班期末考试题之一，不交作业)

题目 S（Sturm-Louville 理论的一个例子）

在这个问题中，我们仿照 $\sin x$ 和 $\cos x$ 的构造来研究一类微分方程的解的零点。我们假设如下的唯一性定理成立：

定理 假设 $a(t)\in C^1(\mathbb{R})$，$t_0\in\mathbb{R}$。如果 $x(t),y(t)\in C^2(\mathbb{R})$，满足同样的方程

$$x''(t)+a(t)x(t)=0,$$
$$y''(t)+a(t)y(t)=0.$$

并且它们的初始值 $(x(t_0),x'(t_0))=(y(t_0),y'(t_0))$ 是一样的，那么 $x(t)\equiv y(t)$。

这个定理说的是在任意一个点 $t_0 \in \mathbb{R}$ 处，一个解 x 和它的导数 x' 在这个点的值决定了 x。作为定理的一个应用，假设

$$x''(t) + a(t)x(t) = 0.$$

如果 $x(0) = 0$ 并且 $x'(0) = 0$，那么 $x(t) \equiv 0$

对任意的函数 $f : \mathbb{R} \to \mathbb{R}$，$t \geqslant 0$，我们用 $\mathrm{Z}_t(f)$ 表示 f 在区间 $[0, t]$ 上的零点的个数，即

$$Z_t(f) = \Big|\big\{x \in [0, t]\big|f(x) = 0\big\}\Big|.$$

第一部分

在此部分中，函数 $a(t), b(t) \in C^1(\mathbb{R})$ 并且对任意的 $t \in \mathbb{R}$，都有 $a(t) \leqslant b(t)$。假设 $x(t), y(t) \in C^2(\mathbb{R})$，满足如下两个方程

$$x''(t) + a(t)x(t) = 0,$$
$$y''(t) + b(t)y(t) = 0.$$

我们还假设 $x(t)$ 和 $y(t)$ 都**不恒为** 0。

S1) 假设 t_1 是 x 的一个零点（即 $x(t_1) = 0$），如果存在 $t > t_1$，使得 $x(t) = 0$，证明，一定存在 $t_2 > t_1$，使得 $x(t_2) = 0$ 并且 x 在 (t_1, t_2) 上无零点。

我们将这样的两个零点 t_1 和 t_2 称作是 x 的两个**相邻的零点**。

（提示：利用我们证明过的微分方程解的唯一性：如果 $x''(t) + a(t)x(t) = 0$ 并且 $x(t_0) = x'(t_0) = 0$，那么 $x(t) \equiv 0$。）

S2) 假设 $t_2 > t_1$ 是 x 的两个相邻的零点。证明，y 在 $(t_1, t_2]$ 上有零点。

（提示：考察函数 $h(t) = x'(t)y(t) - x(t)y'(t)$）

S3) 证明，对任意的 $t \geqslant 0$，我们都有 $Z_t(y) \geqslant Z_t(x) - 1$。

S4) 假设 $t_2 > t_1$ 使得 $x(t_1) = 0$，$x'(t_2) = 0$。证明（任选其一），

- 如果 $y(t_1) = 0$，那么存在 $t_3 \in [t_1, t_2]$，使得 $y'(t_3) = 0$。
- 如果 $y'(t_2) = 0$，那么存在 $t_4 \in [t_1, t_2]$，使得 $y(t_4) = 0$。

第二部分

在此部分中，$p(t) \in C^1(\mathbb{R})$ 是正函数，即对任意的 $t \in \mathbb{R}$，都有 $p(t) > 0$。$x(t), y(t) \in C^2(\mathbb{R})$ 都不恒为 0 并且满足

$$x''(t) + p(t)x(t) = 0,$$
$$y''(t) + p(t)y(t) = 0.$$

S5) 证明，对任意的 $t \geqslant 0$，我们都有 $\left|Z_t(x) - Z_t(y)\right| \leqslant 1$。

S6) 证明下面的两个命题（任选其一）：

- 如果 t_1 和 t_2 是 x 的两个相邻的零点，那么存在唯一的 $t_3 \in [t_1, t_2]$，使得 $x'(t_3) = 0$。
- 如果 t_1' 和 t_2' 是 x' 的两个相邻的零点，那么存在唯一的 $t_3' \in [t_1, t_2]$，使得 $x(t_3') = 0$。

S7) 证明下面的两个命题（任选其一）：

- t_0 是 $|x(t)|$ 的局部最大值当且仅当 $x'(t_0) = 0$。
- t_0' 是 $|x'(t)|$ 的局部最大值当且仅当 $x(t_0') = 0$。

第三部分

在此部分中，我们假设 $p(t) \in C^1(\mathbb{R})$ 是递减的函数并且 $\lim\limits_{t\to+\infty} p(t) > 0$。我们记

$$p(\infty) := \lim_{t\to+\infty} p(t).$$

$x(t) \in C^2(\mathbb{R})$ 不恒为 0 并且满足

$$x''(t) + p(t)x(t) = 0.$$

*S8) 证明，$\lim\limits_{t\to+\infty} \dfrac{Z_t(x)}{t}$ 存在并计算这个极限。

S9) 假设 $0 \leqslant t_1 < t_2 < t_3 < \cdots$ 是 $x(t)$ 在 $[0, +\infty)$ 上所有的零点，$0 \leqslant t_1' < t_2' < t_3' < \cdots$ 是 $x'(t)$ 在 $[0, +\infty)$ 上所有的零点。证明，数列 $\left\{|x'(t_k)|\right\}_{k\geqslant 1}$ 是递减的而数列 $\left\{|x(t_k')|\right\}_{k\geqslant 1}$ 是递增的，并且

$$\lim_{k\to\infty} |x'(t_k)| = \sqrt{p(\infty)} \lim_{k\to\infty} |x(t_k')|.$$

（提示：考察函数 $E(t) = p(t)x(t)^2 + x'(t)^2$ 和 $F(t) = x(t)^2 + \dfrac{1}{p(t)}x'(t)^2$。）

*S10) 假设 $0 \leqslant \widetilde{t}_1 < \widetilde{t}_2 < \widetilde{t}_3 < \cdots$ 是 $x(t)x'(t)$ 在 $[0, +\infty)$ 上所有的零点。证明，数列 $\left\{\widetilde{t}_{k+1} - \widetilde{t}_k\right\}_{k\geqslant 1}$ 是递增的并计算它的极限。

20 紧区间上 Riemann 可积函数的刻画，Newton-Leibniz 公式，分部积分公式，变量替换公式，原函数的计算，圆面积的计算：证明 π 的几何意义

二零一九年十一月二十五日，星期一，晴天

我们现在可以对紧区间 I 上的 Riemann 可积函数做全面的刻画：

定理 125. 假设 $I=[a,b]$（$-\infty<a<b<+\infty$）是有界的闭区间，$f:I\to\mathbb{R}$ 是函数，那么如下四条性质是等价的：

1) $f\in\mathcal{R}(I)$；

2) 存在常数 $\mathbf{I}$，对任意的 $\varepsilon>0$，存在 $\delta>0$，使得对任意的 $(\sigma,\xi)\in\mathcal{S}'(I)$，只要步长 $|\sigma|<\delta$，就有

$$\left|S(f;\sigma,\xi)-\mathbf{I}\right|<\varepsilon.$$

3) f 是有界函数并且 $\underline{D}(f)=\overline{D}(f)$；

4) f 是有界函数并且 $\underline{\int_a^b}f=\overline{\int_a^b}f$。

当上面的任意一个条件成立时，我们就有

$$\underline{D}(f)=\overline{D}(f)=\underline{\int_a^b}f=\overline{\int_a^b}f=\int_a^b f.$$

证明：1)⇒2) 就是定理122，其中，$\mathbf{I}$ 取为 $\int_a^b f$。

2)⇒3) 任意给定 $\varepsilon>0$，按照定义，存在 δ，当 $|\sigma|<\delta$ 时，对任意的 (σ,ξ) 我们都有 $|S(f;\sigma,\xi)-\mathbf{I}|<\varepsilon$。我们选定一个 $\sigma_0=\{b_0<b_1<\cdots<b_0\}$，使得 $|\sigma_0|<\delta$。

我们令 $\xi_i=a_i, i\neq i_0$，$\xi_{i_0}=x\in[a_{i_0-1},a_{i_0}]$ 可以任意选择。根据 $|S(f;\sigma,\xi)|<\mathbf{I}+\varepsilon$，我们知道

$$\begin{aligned}|(a_{i_0}-a_{i_0-1})f(x)|&\leqslant|\sum_{i\neq i_0}(a_i-a_{i-1})|f(\xi_i)|+|S(f;\sigma,\xi)|\\&\leqslant|\sum_{i\neq i_0}(a_i-a_{i-1})|f(a_i)|+\mathbf{I}+\varepsilon.\end{aligned}$$

除了 x 之外，上面的式子中每个值都是给定的，这就证明了 f 在 $[a_{i_0-1},a_{i_0}]$ 上是有界的，从而在 I 上有界。

现在令 $m_i=\inf\limits_{x\in[a_{i-1},a_i]}f(x)$，按照定义，存在 $\zeta_i\in[a_{i-1},a_i]$，使得 $m_i\leqslant f(\zeta_i)\leqslant m_i+\dfrac{\varepsilon}{b-a}$。这样定义的 $S(f;\sigma,\zeta)$ 很明显满足

$$S(f;\sigma,\zeta)\leqslant\underline{S}(f;\sigma)+\varepsilon\leqslant\underline{D}(f)+\varepsilon.$$

从而

$$\mathbf{I} \leqslant \underline{D}(f) + 2\varepsilon.$$

类似地，我们有

$$\mathbf{I} \geqslant \overline{D}(f) - 2\varepsilon.$$

上面两个式子一起就给出了 $\underline{D}(f) = \overline{D}(f)$，其中，最终我们令 $\varepsilon \to 0$。

3)⇔4) 就是前一个命题。

4)⇒1) 根据上下极限的定义，对任意的 $\varepsilon > 0$，存在两个阶梯函数 $\overline{f}(x)$ 和 $\underline{f}(x)$，使得 $\overline{f}(x) \geqslant f(x) \geqslant \underline{f}(x)$，并且

$$0 \leqslant \int_a^b \overline{f}(x) - \int_a^b \underline{f}(x) < \varepsilon.$$

从而，我们令 $F(x) = \underline{f}(x)$ 和 $\psi(x) = \overline{f}(x) - \underline{f}(x) \in \mathcal{E}(I)$，我们自然有

$$|f(x) - F(x)| \leqslant \psi(x), \quad \int_a^b \psi < \varepsilon.$$

这说明 $f \in \mathcal{R}(I)$。另外，我们自然有

$$\int_a^b \underline{f} \leqslant \int_a^b f \leqslant \int_a^b \overline{f},$$

这就证明定理中要求的积分的等式。 □

我们已经建立了比较完备的积分理论，下面我们要发展一个最基本的工具来联系积分与微分并以此进行计算。先引入原函数的概念：假设 f 是定义在 I 上的函数（I 可以是任意区间，f 可以在任意赋范线性空间中取值），如果存在 I 上定义的函数 F，使得 F 可微分并且 $F' = f$，我们就称 F 是 f 的**原函数**。假设 F 是 f 的一个原函数，那么 f 的所有原函数的集合是 $\{F + c | c \in \mathbb{R}\}$，这是因为一个函数的导数恒为零当且仅当这个函数是常值函数。

命题 126. *对任意的 $f \in \mathcal{R}(I)$，对任意的 $x \in I$，我们定义 $F(x) = \int_a^x f$，即 f 在 $[a,x]$ 上的积分。那么，$F(x)$ 是 $I = [a,b]$ 上的连续函数。进一步，如果 f 在 x_0 处是右连续的，那么 F 在 x_0 的右导数存在并且等于 $f(x_0^+) = \lim_{h \to 0, h > 0} f(x_0 + h)$。特别地，当 $f \in C(I)$ 时，$F(x) = \int_a^x f(y)dy$ 是 f 的原函数。*

证明： 首先证明 $F \in C(I)$。假设 $|f| \leqslant M$。对任意的 $x \in I$，对任意的 $h > 0$，我们有

$$\begin{aligned} |F(x-h) - F(x)| &= \Big| - \Big(\int_a^x f - \int_a^{x-h} f \Big) \Big| \\ &= \Big| \int_{x-h}^x f \Big| \leqslant |h| \times M = O(h). \end{aligned}$$

从而，F 在 x 处左连续。类似地，可以证明右连续性，所以 F 在 x 处连续。特别地，这个证明表明 $F(a) = 0$。

进一步假设 f 在 x_0 处右连续，令 $f(x_0^+)$ 为 f 在 x_0 处的右极限，即 $f(x_0^+)=\lim\limits_{h\to 0^+} f(x_0+h)$。按定义，对任意的 ε，存在 $\delta>0$，使得当 $h\in(0,\delta)$ 时，我们有 $|f(x_0+h)-f(x_0^+)|<\varepsilon$。据此，我们计算

$$\begin{aligned}|F(x_0+h)-F(x_0)-f(x_0^+)h| &= \left|\left(\int_a^{x_0+h} f-\int_a^{x_0} f\right)-f(x_0^+)h\right| \\ &= \left|\int_{x_0}^{x_0+h}\big(f(x)-f(x_0^+)\big)dx\right| \leqslant \varepsilon h.\end{aligned}$$

由于 ε 是任意选取的，所以

$$F(x_0+h)-F(x_0)-f(x_0^+)h=o(h).$$

这表明，$f(x_0^+)$ 是 F 在 x_0 处的右导数。 □

推论 127. *假设 $f\in C(I)$，那么 $F(x)=\int_a^x f\in C^1(I)$ 并且 $F'=f$（即 F 是 f 的原函数）。*

定理 128 (微积分基本定理：Newton-Leibniz 公式)**.** *假设 I 是有界闭区间，$f\in\mathcal{R}(I)$。如果 f 有原函数 F，那么*

$$\int_a^b f=F(x)\Big|_a^b.$$

其中，$F(x)\Big|_a^b=F\Big|_{x=a}^{x=b}=F(b)-F(a)$

证明: 我们注意到，如果 $f\in C(I)$ 是连续函数，上面命题的推论表明

$$F(b)=\int_a^b f,\quad F'=f.$$

另外，$F(a)=0$，所以 Newton-Leibniz 公式此时成立（在 $F\in C^1(I)$ 的时候）。

现在考虑一般情况：我们要证明，对任意的 $\varepsilon>0$，都有

$$|F(b)-F(a)-\int_a^b f|<\varepsilon.$$

此时，由于 f 可能在某些点处不连续，所以 $\int_a^x f(y)dy$ 未必可微分。特别地，我们未必有 $F(x)=\int_a^x f(y)dy$。实际上，一个可微函数的导数未必是 Rieman 可积的：比如说，在区间 $[-1,1]$ 上定义的函数

$$[-1,1]\to\mathbb{R},\quad x\mapsto\begin{cases}x^2\sin\left(\frac{1}{x^2}\right), & x\neq 0;\\ 0, & x=0.\end{cases}$$

这个函数是可微的但是它的导数是无界的，所以不是 Riemann 可积分的。

根据 f 是 Riemann 可积的，我们可以选取 $g,\psi\in\mathcal{E}(I)$，使得

$$|f(x)-g(x)|\leqslant\psi(x),\quad\forall x\in I;\qquad\int_a^b\psi<\varepsilon.$$

我们再选取分划 $\sigma = \{a = a_0 < \cdots < a_n = b\}$，使得该分划与 g 和 ψ 两个阶梯函数相容。按照定义，g 和 ψ 在每个区间 (a_{i-1}, a_i) 上的限制是常值函数：我们令

$$y_i = g\Big|_{(a_{i-1},a_i)}, \quad z_i = \psi\Big|_{(a_{i-1},a_i)}.$$

为了控制 $\left|F(b) - F(a) - \int_a^b f\right|$，我们自然考虑它较好的一个逼近：

$$\begin{aligned} F(b) - F(a) - \int_a^b g &= \sum_{i=1}^n \big(F(a_i) - F(a_{i-1})\big) - \sum_{i=1}^n \int_{a_{i-1}}^{a_i} g \\ &= \sum_{i=1}^n \Big(F(a_i) - F(a_{i-1}) - (a_i - a_{i-1})y_i\Big) \\ &= \sum_{i=1}^n \Big(\big(F(a_i) - a_i y_i\big) - \big(F(a_{i-1}) - a_{i-1}y_i\big)\Big). \end{aligned}$$

在每个区间 $[a_{i-1}, a_i]$ 上，我们考虑函数 $x \mapsto F(x) - y_i x$。根据 Lagrange 中值定理，存在 $\xi_i \in [a_{i-1}, a_i]$，使得

$$\begin{aligned} &\Big|\big(F(a_i) - a_i y_i\big) - \big(F(a_{i-1}) - a_{i-1}y_i\big)\Big| \\ &= \big|F'(\xi_i) - y_i\big|(a_i - a_{i-1}) \\ &= \big|f(\xi_i) - g(\xi_i)\big|(a_i - a_{i-1}) \\ &\leqslant \psi(\xi_i)(a_i - a_{i-1}) \\ &= z_i(a_i - a_{i-1}). \end{aligned}$$

所以，我们得到

$$|F(b) - F(a) - \int_a^b g| \leqslant \sum_{i=1}^n z_i(a_i - a_{i-1}) = \int_a^b \psi < \varepsilon.$$

最终，我们有

$$\begin{aligned} |F(b) - F(a) - \int_a^b f| &\leqslant |F(b) - F(a) - \int_a^b g| + |\int_a^b f - g| \\ &\leqslant \varepsilon + \int_a^b \psi < 2\varepsilon \end{aligned}$$

令 $\varepsilon \to 0$，我们就证明了微积分基本定理。 □

注记. *Newton-Leibniz 公式最常用的形式是：$f \in C^1(I)$，我们有*

$$\int_a^b f' = f(x)\Big|_a^b.$$

给定一个函数 f，我们用 $\int f$ 来表示它（某个）原函数，我们注意到，按照原函数的定义，一个函数的原函数在差一个常数的意义下是才是良好定义的，但是这种差别不是本质的，没有必要纠结，所以，我们之后谈论原函数的时候，总是随便取一个就好，尽管这是从概念上而言不严格。另外，即便 f 是 Riemann 可积的，它未必具有原函数，这是因为如果 f 有原函数 F，那么 $F'=f$，根据 Darboux 定理，f 具有介值性，而我们随便选取的阶梯函数大都不具有这个性质。

我们罗列几个基本的原函数：

$$\int x^\alpha = \frac{1}{\alpha+1}x^{\alpha+1},\ \alpha\neq -1; \qquad \int \frac{1}{x} = \log x; \quad \int \log x = x\log x - x;$$
$$\int \sin x = -\cos x; \qquad \int \cos x = \sin x; \qquad \int e^x = e^x;$$
$$\int \frac{1}{a^2+x^2} = \frac{1}{a}\arctan\frac{x}{a}; \quad \int \frac{1}{\sqrt{a^2-x^2}} = \arcsin\frac{x}{a}; \quad \int \tan ax = -\frac{1}{a}\log\cos ax.$$

还有几个稍微复杂但是也会见到的：

$$\int \sec x = \log|\sec x+\tan x| = 2\tanh^{-1}(\tan\frac{x}{2});$$
$$\int \sec^2 ax = \frac{\tan ax}{a}; \quad \int \frac{1}{\cos x} = \log|\tan\frac{x}{2}|.$$

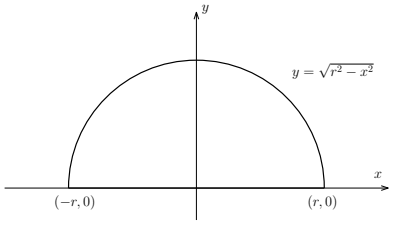
注记. 我们将会看到，关于原函数的计算（如果可以用初等函数表示的话），本质上是代数的变形，所以，我们需要记忆并熟练的运用以上几个例子。

这些例子的证明是平凡的，我们只需要对原函数求导来验证它恰好给出了所求的函数即可。利用 Newton-Leibniz 公式和上面的原函数，我们可以计算很多积分了：

1) 单项式的积分：

$$\int_a^b x^n = \int_a^b \frac{1}{n+1}x^{n+1} = \frac{1}{n+1}x^{n+1}\Big|_a^b = \frac{1}{n+1}(b^{n+1}-a^{n+1}).$$

2) $\overline{\pi}=\pi$ 的第一个证明。我们计算平面上半圆盘的面积（我们期待它应该是 $\frac{r^2}{2}\pi$）：



这就是函数 $f(x)=\sqrt{r^2-x^2}$ 下的面积，其中，$x\in[-r,r]$。换而言之，我们要计算如下的积分：

$$\int_{-r}^r f = \int_{-r}^r \sqrt{1-x^2}.$$

我们发现，计算 $\sqrt{1-x^2}$ 的原函数并不简单（至少不是一目了然）。我们需要发展一点工具来计算原函数（积分）然后再来计算这个积分。

作为 Newton-Leibniz 公式应用，我们现在来证明分部积分公式和变量替换公式。

命题 129 (分部积分公式)**.** 假设 $f, g \in C^1(I)$。那么，我们有

$$\int_a^b f' \cdot g = (f \cdot g)\Big|_a^b - \int_a^b f \cdot g'.$$

证明: 对 $(f \cdot g)'$ 用 Newton-Leibniz 公式，所以

$$(f \cdot g)\Big|_a^b = \int_a^b \big(f \cdot g\big)' = \int_a^b f' \cdot g + \int_a^b f \cdot g'.$$

这自然等价于分部积分公式。 □

注记. 我们回忆一下所谓的 *Abel* 求和公式：假设 $\{a_k\}_{k\geqslant 1}$ 和 $\{b_k\}_{k\geqslant 1}$ 是复数（矩阵）的序列，那么

$$\sum_{k=1}^{n} a_k b_k = S_n b_n + \sum_{k=1}^{n-1} S_k (b_k - b_{k+1}),$$

其中我们用 $S_n = a_1 + a_2 + \cdots + a_n$ 表示数列 $\{a_k\}_{k\geqslant 1}$ 的部分和。这和分部积分公式非常相似：我们将 S_k 和 b_k 看成在 $\mathbb{Z}_{\geqslant 1}$ 定义的函数，那么，如下的对应

$$S_k \longrightarrow f, \quad b_k \longrightarrow f, \quad \mathbb{Z}_{\geqslant 1} \longrightarrow [a, b],$$

就可以把 *Abel* 求和公式看成分部积分。(我们下个学期会学习抽象的积分理论把这个对应说清楚)

命题 130 (变量替换公式)**.** 假设 $\Phi : I \to J$ 是两个有界闭紧区间之间的连续的单调递增的可微映射，其中 $I = [a, b]$，$J = [c, d]$。$f \in C(J)$，那么对于 I 中任意的 $\alpha < \beta$，都有

$$\int_\alpha^\beta \big((f \circ \Phi) \cdot \Phi'\big) = \int_{\Phi(\alpha)}^{\Phi(\beta)} f.$$

证明: 我们先假设 $\Phi(\beta) > \Phi(\alpha)$. 根据 $f \in C(I)$ 和 $\Phi \in C^1(I)$，我们知道上面两个积分的被积函数都是连续的，所以都是 Riemann 可积的。根据复合函数求导公式，令 $F(x)$ 为 $f(x)$ 的一个原函数，那么我们有

$$F(\Phi(x))' = f(\Phi(x))\Phi'(x) \quad \Leftrightarrow \quad \big(F \circ \Phi\big)' = (f \circ \Phi) \cdot \Phi'.$$

所以利用 Newton-Leibniz 公式，我们有

$$\int_\alpha^\beta (f \circ \Phi) \cdot \Phi' = \int_\alpha^\beta \big(F \circ \Phi\big)' = F \circ \Phi\Big|_\alpha^\beta = F(\Phi(\beta)) - F(\Phi(\alpha)).$$

另外，由于 $f(x)$ 是连续的，所以我们可以选取 F 为

$$F(x) = \int_a^x f.$$

从而，

$$F(\Phi(\beta)) - F(\Phi(\alpha)) = \int_a^{\Phi(\beta)} f - \int_a^{\Phi(\alpha)} f = \int_{\Phi(\alpha)}^{\Phi(\beta)} f.$$

命题得证。 □

注记. 注意到，我们假设了 $\Phi(\alpha)<\Phi(\beta)$ 是递增的，这是人为的原因的：对于区间 $I=[a,b]$，最好的积分的符号应该是 $\int_I f$。然而，由于传统的原因，我们经常使用 $\int_a^b f$，这里，我们已经默认了 $a>b$ 的假设。从这角度看，$\int_a^b$ 不是一个太好的符号。然而，重新观察证明，我们发现等式

$$\int_\alpha^\beta (f\circ\Phi)\cdot\Phi'=F(\Phi(\beta))-F(\Phi(\alpha))$$

是永远成立的，所以，当 $\Phi(\beta)<\Phi(\alpha)$ 的时候，我们需要将 $\int_{\Phi(\alpha)}^{\Phi(\beta)} f$ 解释为

$$\int_{\Phi(\alpha)}^{\Phi(\beta)} f=-\int_{\Phi(\beta)}^{\Phi(\alpha)} f=-\left(F(\Phi(\alpha))-F(\Phi(\beta))\right)$$

即可。换而言之，作为约定，我们规定

$$\int_a^b f=-\int_b^a f.$$

作为换元积分公式的推论，我们证明两个基本的命题。尽管这两个命题的证明不值一提，但是由于下面的平移和相似变换是 $\mathbb{R}$ 上最基本的对称性（共形变换群），所以它们占有特殊的位置：

- 假设 $f\in\mathcal{R}(I)$，我们定义函数 $g(x)=f(x+x_0)$，即函数 $g=f\circ\tau_{x_0}:[a-x_0,b-x_0]\to\mathbb{R}$，其中 $\tau_{x_0}:[a-x_0,b-x_0]\to[a,b]$ 是平移 $x\mapsto x+x_0$。此时，我们有
$$\int_{a-x_0}^{b-x_0} g=\int_a^b f.$$

- 假设 $f\in\mathcal{R}(I)$，$\lambda>0$，我们定义函数 $g(x)=f(\lambda x)$，即函数 $g=f\circ m_\lambda:[\frac{a}{\lambda},\frac{b}{\lambda}]\to\mathbb{R}$，其中 $m_\lambda:[\frac{a}{\lambda},\frac{b}{\lambda}]\to[a,b]$ 是相似变化 $x\mapsto\lambda x$。此时，我们有
$$\lambda\int_{\frac{a}{\lambda}}^{\frac{b}{\lambda}} g=\int_a^b f.$$
当 $\lambda<0$ 时，我们要将 $\int_{\frac{a}{\lambda}}^{\frac{b}{\lambda}}$ 换做 $\int_{\frac{b}{\lambda}}^{\frac{a}{\lambda}}$。

注记 (符号的解释)**.** 根据变量替换公式的形式，我们倾向于将积分写成传统的形式：$\int_a^b f(x)dx$。

为了讲的更清楚，我们回忆微分（不是导数!）的定义。给定一个两个区间之间的可微的映射 $\Phi:X\to Y$，对于每个点 $x_0\in X$，我们给它分配一个线性空间 $T_{x_0}X\simeq\mathbb{R}$；对于 $y_0\in Y$，我们给它分配一个线性空间 $T_{y_0}Y\simeq\mathbb{R}$。那么，映射 Φ 在 x_0 处的微分定义为：

$$d\Phi(x_0):T_{x_0}X\to T_{\Phi(x_0)}Y,\quad h\mapsto\Phi'(x_0)h.$$

假设 $X = Y$，我们选取一个特殊的 Φ，即恒同映射 $x \mapsto x$。我们就把这个映射记作 x。特别地，此时 $\Phi' = 1$。所以，微分现在可以被写为

$$dx(x_0) : T_{x_0}X \to T_{\Phi(x_0)}X, \quad h \mapsto h.$$

比较上面两个映射，我们知道，作为 $T_{x_0}X \to T_{\Phi(x_0)}Y$ 之间的映射，我们可以把上面的映射写成

$$d\Phi(x_0) = \Phi'(x_0)dx(x_0).$$

忽略掉上面式子对点的依赖性，我们就得到了

$$d\Phi = \Phi'(x)dx.$$

所以，如果我们用传统的积分记号，换元积分公式就可以想象成用变元 y 代替 $\Phi(x)$，即 $y = \Phi(x)$，公式可以写成

$$\int_\alpha^\beta f\big(\Phi(x)\big) \underbrace{\Phi'(x)dx}_{d\Phi(x)} = \int_{\Phi(\alpha)}^{\Phi(\beta)} f(y)dy.$$

我们强调，目前这种写法只是形式的记法，到下个学期我们用正确的测度的观点或者微分形式的观点，就可以完美地解释这种记号。大家大可不必纠结你是否理解这个记法，因为这目前是一个帮助记忆的方式，仅此而已。

我们可以看一下如何利用分部积分和换元积分公式进行积分的计算。在一定的意义上，这已经不属于分析的范畴了（不用取极限或者 ε 语言），而算是组合或者是代数上的操作了：

例子. 我们先来完成关于圆的面积的计算：

1) 我们用如下的变量替换 $x = r\sin\theta$，即考虑 $\Phi : [-\overline{\pi}, \overline{\pi}] \to [-r, r]$。从而，

$$\begin{aligned}
\int_{-r}^{r} \sqrt{r^2 - x^2}dx &= \int_{-\frac{\overline{\pi}}{2}}^{\frac{\overline{\pi}}{2}} \sqrt{r^2 - (r\sin\theta)^2} r\cos\theta d\theta \\
&= r^2 \int_{-\frac{\overline{\pi}}{2}}^{\frac{\overline{\pi}}{2}} (\cos\theta)^2 d\theta = \frac{r^2}{2} \int_{-\frac{\overline{\pi}}{2}}^{\frac{\overline{\pi}}{2}} (\cos(2\theta) + 1)\, d\theta \\
&= \frac{r^2}{2} \left(= \int_{-\frac{\overline{\pi}}{2}}^{\frac{\overline{\pi}}{2}} \cos(2\theta)d\theta + \overline{\pi} \right) \\
&\overset{\vartheta = 2\theta}{=} \frac{r^2}{2} \left(\frac{1}{2} = \int_{-\overline{\pi}}^{\overline{\pi}} \cos(\vartheta)d\vartheta + \overline{\pi} \right)
\end{aligned}$$

另外，根据 $\sin x$ 的周期性，我们知道 $\sin(\overline{\pi}) = \sin(-\overline{\pi})$，所以

$$\int_{-\overline{\pi}}^{\overline{\pi}} \cos(\vartheta)d\vartheta = \sin\vartheta\Big|_{-\overline{\pi}}^{\overline{\pi}} = 0.$$

所以，

$$\int_{-r}^{r} \sqrt{r^2 - x^2}dx = \frac{r^2}{2}\overline{\pi}.$$

所以，半径为 r 的圆盘的面积是 $\overline{\pi}r^2$。至此我们验证了 $\overline{\pi} = \pi$，其中 π 是大家想象中的 π。

2) 计算 $\int_0^{\frac{3\pi}{2}} \sin^2 x \cos x dx$。我们将积分分为几块然后用换元积分公式：

$$\begin{aligned}\int_0^{\frac{3\pi}{2}} \sin^2 x \cos x dx &= \left(\int_0^{\frac{\pi}{2}} + \int_{\frac{\pi}{2}}^{\pi} + \int_{\pi}^{\frac{3\pi}{2}}\right) \sin^2 x d\sin x \\ &= \int_0^1 y^2 dy - \int_0^1 y^2 dy + \int_0^1 y^2 dy \\ &= \frac{1}{3} - \frac{1}{3} + \frac{1}{3} = 1.\end{aligned}$$

请尤其注意第二项的换元积分，因为这时候对应着之前 $\Phi(\beta) < \Phi(\alpha)$ 的情形。

3) 计算 $\int_0^{\frac{\pi}{2}} x \sin x \cos x dx$。这个积分的计算需要用到分部积分公式来除掉 x：

$$\begin{aligned}\int_0^{\frac{\pi}{2}} x \sin x \cos x dx &= \frac{1}{2}\int_0^{\frac{\pi}{2}} x \sin(2x) dx = -\frac{1}{4}\int_0^{\frac{\pi}{2}} x\big(\cos(2x)\big)' dx \\ &= \frac{1}{4}\int_0^{\frac{\pi}{2}} \cos(2x) dx - \frac{1}{4} x\cos(2x)\Big|_0^{\frac{\pi}{2}} \\ &= \frac{1}{8}\int_0^{\frac{\pi}{2}} \big(\sin(2x)\big)' dx + \frac{1}{8}\pi = \frac{1}{8}\pi.\end{aligned}$$

4) 计算 $\int_1^2 \frac{1}{x\sqrt{1+x^2}} dx$。在这个计算中，我们将 x 换成 $\frac{1}{y}$：

$$\int_1^2 \frac{1}{x\sqrt{1+x^2}} dx = \int_{\frac{1}{2}}^1 \frac{1}{\sqrt{1+y^2}} dy.$$

为了计算上面的积分，我们首先声明 $\int \frac{1}{\sqrt{1+x^2}} = \log(x + \sqrt{1+x^2})$，据此，

$$\int_{\frac{1}{2}}^1 \frac{1}{\sqrt{1+y^2}} dy = \log\frac{2+2\sqrt{2}}{2\sqrt{5}}.$$

很明显，上面的不定积分的计算是有难度的，这样的技巧我们留到后面来谈。我们给一个更透明的计算，想法是注意到 $1+x^2$ 的形式让我们联想到双曲三角函数，我们做变量替换 $y = \sinh z$，从而

$$\int_{\frac{1}{2}}^1 \frac{1}{\sqrt{1+y^2}} dy. = \int_\alpha^\beta \frac{1}{\cosh z} \cosh z dz = \beta - \alpha,$$

其中，$\sinh(\alpha) = \frac{1}{2}$，$\sinh(\beta) = 1$。通过解二次方程，我们就可以确定 α 和 β。

20.1 作业：Dini 定理，多项式逼近与 Weierstrass-Stone 定理

清华大学 19-20 秋季学期，数学分析一，作业 9

请用 A4 大小的纸张**正反面**用钢笔，签字笔或者圆珠笔书写，并注明自己的姓名，年级和作业的**总页数**。除定理公式所涉及的人名之外，请使用**中文**。本次作业的提交时间和地点为 **12 月 5 日**上午的课堂上，逾期视作零分。

基本习题

习题 A（积分定义的补充与扩展） 我们总假设 $I=[a,b]\subset\mathbb{R}$ 是一个有界闭区间，V 是一个赋范线性空间。

A1) $\sigma_1,\sigma_2\in\mathcal{S}$ 是两个分划。证明，对任意的 $\varepsilon>0$，总存在分划 σ，使得 $\sigma\prec\sigma_1$，$\sigma\prec\sigma_2$ 并且它的步长 $|\sigma|<\varepsilon$。

A2) 考虑在 V 中取值的阶梯／简单函数的空间 $\mathcal{E}(I)$，证明，这是 $\mathbb{R}$-线性空间并且积分算子 $\int_a^b:\mathcal{E}(I)\to V$ 是良好定义的（不依赖于分划的选取）并且是线性映射，其中积分的定义方式与函数在 $\mathbb{R}$ 中取值时的方式一致。据此，用阶梯函数逼近的方式定义在 V 中取值的 Riemann 可积的函数。你不需要写下细节但是你应该对照笔记研究原来证明的每一步。

A3) 假设 $f:I\to\mathbb{R}^n$，其中 f_i 是 f 的每个分量。那么 $f\in\mathcal{R}(I)$ 当且仅当对每个分量 f_i 我们都有 $f_i\in\mathcal{R}(I)$。特别地，当 $f:\mathbb{R}\to\mathbb{C}$ 时，我们有 $\int_a^b f=\int_a^b\Re f+i\int_a^b\Im f$，其中 $\Re f$ 和 $\Im f$ 分别为 f 的实部和虚部。

A4) 试证明积分的区间可加性：假设 $a<c<b$，那么对于任意的 $f\in\mathcal{R}(I)$，我们有 f 在 $[a,c]$ 和 $[c,b]$ 上的限制都是阶梯函数，并且

$$\int_a^b f=\int_a^c f+\int_c^b f.$$

A5) 证明，对于任意两个分划 σ 和 σ'，它们所对应的 Darboux 上下和满足

$$\underline{S}(f;\sigma)\leqslant\overline{S}(f;\sigma').$$

据此证明，如果 $f\in\mathcal{R}(I)$，就有 $\lim\limits_{|\sigma|\to0}|\underline{S}(f;\sigma)-\overline{S}(f;\sigma)|=0$，即对任意的 $\varepsilon>0$，一定存在 $\delta>0$，对任意的分划 σ，只要 $|\sigma|<\delta$，我们就有

$$|\underline{S}(f;\sigma)-\overline{S}(f;\sigma)|<\varepsilon.$$

A6) $f\in\mathcal{R}(I)$。证明，改变 f 在有限个点上的取值所得到的函数仍是 Riemann 可积的并且积分与 f 的相同。

A7) $f \in C([a,b])$。假设对任意的 $x \in I$，我们都有 $f(x) \geqslant 0$ 并且存在点 $x_0 \in I$ 使得 $f(x_0) > 0$。证明，$\int_a^b f > 0$。

A8)（不定积分的分部积分公式：对计算不定积分有用）假设 $f, g \in C^1(I)$，那么，我们有

$$\int f' \cdot g = f \cdot g - \int f \cdot g'.$$

A9)（不定积分的变量替换公式：对计算不定积分有用）假设 $\Phi : \mathbb{R} \to \mathbb{R}$ 可微，f 是连续函数，那么

$$\int (f \circ \Phi)\Phi' = \int f.$$

习题 B(不定积分的计算) 试计算下列不定积分

(1) $\int \frac{x^5}{1+x}$　(2) $\int \sqrt{x\sqrt{x\sqrt{x}}}$　(3) $\int \left(\frac{1+x}{1-x} + \frac{1-x}{1+x}\right)$

(4) $\int \frac{e^{3x}+1}{1+e^x}$　(5) $\int \sqrt{1-\sin(2x)}$　(6) $\int \frac{\cos(2x)}{\cos x - \sin x}$

(7) $\int \tan^2 x$　(8) $\int |x|$　(9) $\int e^{-|x|}$

(10) $\int \frac{x^2}{(1-x)^{2018}}$　(11) $\int |x-1|$　(12) $\int \frac{1}{\sqrt{b^2+x^2}}$

(13) $\int \frac{1}{\sqrt{x}(1+x)}$　(14) $\int \frac{x^4}{(1-x^5)^4}$　(15) $\int \left(\frac{1}{\sqrt{3-x^2}} + \frac{1}{1-3x^2}\right)$

(16) $\int \frac{2x-3}{x^2-3x+8}$　(17) $\int \frac{1}{\sin^2(2x+\frac{\pi}{4})}$　(18) $\int \frac{1}{1+\cos x}$

(19) $\int \frac{1}{x^2}\sin\frac{1}{x}$　(20) $\int \cos^5 x$　(21) $\int \cos(ax)\sin(bx)$

(22) $\int \frac{1}{a\cos x + b\sin x}$　(23) $\int \frac{\sin(2x)}{a^2\cos^2 x + b^2\sin^2 x}$　(24) $\int \frac{1}{2-\sin^2 x}$

(25) $\int \frac{1}{x(\ln x)(\ln\ln x)}$　(26) $\int \frac{\ln x}{x\sqrt{1+\ln x}}$　(27) $\int \frac{\cos x + \sin x}{\sqrt[3]{\sin x - \cos x}}$

(28) $\int e^{\sqrt{x}}$　(29) $\int \frac{x^{\frac{n}{2}}}{1+x^{n+2}}$　(30) $\int \frac{\sqrt{x}}{1-\sqrt[3]{x}}$

(31) $\int \frac{1}{(x^2+a^2)^{\frac{3}{2}}}$　(32) $\int \frac{1}{\cos^4 x}$　(33) $\int (\arcsin x)^2$

(34) $\int x\arcsin x$　(35) $\int x\arctan x$　(36) $\int \frac{\arctan x}{x^2}$

(37) $\int x^2 \sin x$　(38) $\int \frac{x}{\cos^2 x}$　(39) $\int \ln(x+\sqrt{1+x^2})$

(40) $\int \sin(\ln x)$　(41) $\int \sqrt{x^2+a^2}$　(42) $\int \frac{x^2}{\sqrt{x^2-a^2}}$

(43) $\int \frac{x\ln(x+\sqrt{1+x^2})}{\sqrt{1+x^2}}$　(44) $\int \frac{1}{\sqrt{x^2+a^2}}$　(45) $\int \frac{xe^x}{(1+x)^2}$

(46) $\int \arctan(1+\sqrt{x})$　(47) $\int \left(1-\frac{2}{x}\right)^2 e^x$　(48) $\int \sqrt{2+\tan^2 x}$

(49) $\int \frac{1}{1+x^3}$　(50) $\int \frac{x^7}{x^4+2}$　(51) $\int \frac{2x^2+1}{(x+3)(x-1)(x-4)}$

(52) $\int \frac{1+x^2}{1+x^4}$　(53) $\int \frac{x}{\sqrt{x+1}+\sqrt[3]{x+1}}$　(54) $\int \frac{1}{\sqrt{x+x^2}}$

(55) $\int \frac{1-r^2}{1-2r\cos x+r^2}, r\in(0,1)$　(56) $\int \frac{1}{x\sqrt{1+x^2}}$　(57) $\int \frac{1}{5-3\cos x}$

(58) $\int \frac{1}{2+\sin^2 x}$　(59) $\int \frac{\sin^3 x}{\cos^4 x}$　(60) $\int \frac{1}{\sin x\cos^4 x}$.

思考题（重要结论，不交作业，鼓励讨论!）

习题 W（多项式逼近和 Weierstrass-Stone 定理）我们要证明如下著名的定理：任意给定有界闭区间 $I=[a,b]$ 上的连续函数，我们总是可以用一个多项式来足够好地逼近它。更精确地说，给定 $f\in C([a,b])$，对任意的 $\varepsilon>0$，存在多项式 P_ε，使得 $\sup\limits_{x\in[a,b]}|f(x)-P_\varepsilon(x)|<\varepsilon$，即若用 $P([a,b])$ 表示 $[a,b]$ 上多项式函数组成的空间，则 $P([a,b])$ 在度量空间 $C([a,b])$ 中稠密（$C([a,b])$ 上用的范数是 $\|f\|_\infty=\sup\limits_{x\in[a,b]}|f(x)|$）。

第一部分：Dini 定理及应用

W1)（Dini 定理）假设 $K\subset\mathbb{R}^n$ 是紧子集，$f_n:K\to\mathbb{R}$ 是一列连续函数，它们逐点地收敛到连续函数 $f:K\to\mathbb{R}$，即对每个 $x\in K$，我们都有 $\lim\limits_{n\to\infty}f_n(x)=f(x)$。证明，如果 $\{f_n\}_{n\geqslant 1}$ 是上升的函数列（即对任意 $x\in K$ 和 n，我们都有 $f_n(x)\leqslant f_{n+1}(x)$），那么 f_n 一致收敛到 f。（参考荆公子的某次习题课）

W2) 考虑区间 $I=[-1,1]$。我们通过归纳的方式定义一族多项式函数：

$$P_0(x)=0,\quad P_{n+1}(x)=P_n(x)+\frac{1}{2}(x^2-P_n^2(x)).$$

证明，对任意的 n 和 x，我们都有 $0\leqslant P_n(x)\leqslant P_{n+1}(x)\leqslant|x|$。

W3) 证明，绝对值函数 $|x|$ 在 $I=[-1,1]$ 上可以被多项式一致地逼近，即对任意的 $\varepsilon>0$，存在某个多项式函数 $P_\varepsilon(x)$，使得 $\sup\limits_{x\in[-1,1]}\big||x|-P_\varepsilon(x)\big|<\varepsilon$。

第二部分：区间上的情形

这一部分中，我们假设 $I=[0,1]$，n 是正整数。

W4) 对任意的 $0\leqslant k\leqslant n$，我们定义 $p_{n,k}(x)=\binom{n}{k}x^k(1-x)^{n-k}$。证明，$\sum\limits_{0\leqslant k\leqslant n}p_{n,k}(x-\dfrac{k}{n})^2=\dfrac{x(1-x)}{n}$。

W5) 任意给定 $f\in C([0,1])$，我们定义 $B_{f,n}(x)=\sum\limits_{0\leqslant k\leqslant n}f(\dfrac{k}{n})\binom{n}{k}x^k(1-x)^{n-k}$。对 $x\in[0,1]$，证明，

$$|f(x)-B_{f,n}(x)|\leqslant\sum_{k=0}^{n}|f(x)-f(\frac{k}{n})|p_{n,k}(x).$$

W6)（用 Bernstein 多项式逼近连续函数）任意给定 $f\in C([0,1])$，证明，对任意 $\varepsilon>0$，总存在 n，使得 $\|f-B_{f,n}\|_\infty<\varepsilon$。

注记. 用概率论的观点，$x\in[0,1]$ 给定，*Bernstein* 多项式逼近的的方法讲的是概率测度 $\mu_x=\sum\limits_{k=0}^{n}\binom{n}{j}x^k(1-x)^{n-k}\delta_{\frac{k}{n}}$ 的极限是 *Dirac* 测度 δ_x。

第三部分：紧集上的情形

从此往后，我们假设 $K\subset\mathbb{R}^n$ 是紧集，$C(K)$ 是 K 上的实数值连续函数所构成的线性空间，用 $P(K)$ 表示 K 上多项式函数组成的空间（即形如 $\sum\limits_{|\alpha|\leqslant m}c_\alpha x^\alpha$ 的函数，其中 m 是正整数，α 是多重指标，请参见关于多重偏导数的课堂笔记）。

W7) 假设 $\mathcal{A}\subset C(K)$ 是非零的线性子空间。如果对任意的 $f,g\in\mathcal{A}$，它们的乘积 $f\cdot g\in\mathcal{A}$，我们就把 $\mathcal{A}$ 称作是 $C(K)$ 的一个**子代数**。证明，$P(K)$ 是 $C(K)$ 的子代数。

W8) 非零的线性子空间 $\mathcal{A}\subset C(K)$ 是**闭子代数**，也就是说如果 $\{f_k\}_{k\geqslant1}\subset\mathcal{A}$，$f_k$ 一致收敛到 f，那么 $f\in\mathcal{A}$。假设常值函数 $1\in\mathcal{A}$，证明，如果 $f\in\mathcal{A}$，那么 $|f|\in\mathcal{A}$。（提示：利用 W3)）

W9) 假设 $\mathcal{A}\subset C(K)$ 是子集，如果对任意的 $x,x'\in K$，$x\neq x'$，都存在 $f\in\mathcal{A}$，使得 $f(x)\neq f(x')$，我们就称 $\mathcal{A}$ 是**能够区分点的**。证明，$P(K)$ 是能够区分点的。

W10) 假设 $f,g\in C(K)$。证明，函数 $f\wedge g(x)=\min\{f(x),g(x)\}$ 和 $f\vee g(x)=\max\{f(x),g(x)\}$ 都是连续的。

W11) 假设 $\mathcal{A}\subset C(K)$ 是子集，如果对任意的 $f,g\in\mathcal{A}$，$f\wedge g,f\vee g\in\mathcal{A}$，我们就称 $\mathcal{A}$ 是**$\wedge\vee$-封闭的**。证明，$\overline{P(K)}$ 是 $\wedge\vee$-封闭的，其中 $\overline{P(K)}=\big\{f\in C(K)|\text{存在}\{f_k\}_{k\geqslant1}\subset P(K),f_k\text{一致收敛到}f\big\}$ 为 $P(K)$ 在 $C(K)$ 中的闭包。

W12) 我们现在假设 $\mathcal{A} \subset C(K)$ 是 $\wedge\vee$-封闭的，并且对任意的 $\alpha, \beta \in \mathbb{R}$，任意的 $x, y \in K$，$x \neq y$（我们此时假设 K 至少含有两个点）都存在函数 $\varphi \in A$，使得 $\varphi(x) = \alpha$，$\varphi(y) = \beta$。

我们通过下面的步骤来证明 $\mathcal{A} \subset C(K)$ 是稠密的（即对任意给定的连续函数 $f \in C(K)$，对任意的 $\varepsilon > 0$，总存在 $\varphi \in \mathcal{A}$，使得 $\|\varphi - f\|_\infty < \varepsilon$）：

- 固定 $x_0 \in X$，对任意的 $y \in K$，我们选取 $\varphi_y \in \mathcal{A}$，使得 $\varphi_y(x_0) = f(x_0)$ 并且 $\varphi_y(y) = f(y)$，这样我们得到一族 $\{f_y \in \mathcal{A} | y \in K\}$，据此，对每个 $y \in K$，可以定义 $U_y = \{z \in K | \varphi_y(z) > f(z) - \varepsilon\}$。证明，存在有限个 $y_1, \cdots, y_m \in K$，使得 $K \subset U_{y_1} \cup \cdots \cup U_{y_m}$。
- 证明，$\varphi_{x_0} = \sup\limits_{x \in K}\{\varphi_{y_1}(x), \cdots, \varphi_{y_m}(x)\} \in \mathcal{A}$ 并且对任意的 $x \in K$，我们都有 $\varphi_{x_0}(x) - f(x) > -\varepsilon$。
- 证明，存在 $\varphi \in \mathcal{A}$，使得对任意的 $x \in K$，都有 $|\varphi(x) - f(x)| < \varepsilon$。

W13) 证明，如果子代数 $\mathcal{A} \subset C(K)$ 是能够区分点的，$\wedge\vee$-封闭的并且包含所有的常数值函数，那么 $\mathcal{A} \subset C(K)$ 是稠密的。

W14) （Weierstrass-Stone 定理）如果 $\mathcal{A} \subset C(K)$ 是一个能区分点的子代数并且包含常值函数 1，那么 $\mathcal{A} \subset C(K)$ 是稠密的。

注记. 我们还有一种复值函数的 *Weierstrass-Stone* 定理，在很多的场合有着重要的应用，有兴趣的同学可以自己查阅。

W15) 多项式函数 $P(K)$ 在连续函数空间 $C(K)$ 中是稠密的。

W16) 给定以 2π 为周期的连续函数 $f \in C(\mathbb{R})$。证明，任给 $\varepsilon > 0$，总存在一个有限的三角级数

$$T(x) = \sum_{-N \leqslant k \leqslant N} a_k \cos(kx) + \sum_{-N \leqslant k \leqslant N} b_k \sin(kx),$$

其中 N 是正整数，a_k, b_k 是实数，使得对任意的 $x \in \mathbb{R}$，我们都有 $|f(x) - T(x)| < \varepsilon$。（提示：考虑 $C[0, 2\pi]$ 和它的某个子代数）

21 振幅，零测集，Lebesgue 定理

二零一九年十一月二十八日，星期四，晴天

我们这一次给出 Riemann 积分的最后一个刻画，这就是所谓的 Lebesgue 定理。

假设 $f:\mathbb{R}\to\mathbb{R}$ 是一个函数，$x_0\in\mathbb{R}$，我们定义 f 在 x 出的振幅 $\omega(f,x_0)$ 为

$$\omega(f,x_0)=\inf_{\delta>0}\left(\sup_{\substack{|x-x_0|<\delta,\\ |y-x_0|<\delta}}|x-y|\right)=\lim_{\delta\to 0}\sup_{\substack{|x-x_0|<\delta,\\ |y-x_0|<\delta}}|x-y|.$$

这个概念可以被推广到距离空间的范畴上，假设 $f:X\to Y$ 是两个距离空间 (X,d_X) 和 (Y,d_Y) 之间的映射，对于给定的 $x_0\in X$，我们定义 f 在 x 处的**振幅**为：

$$\omega(f,x_0)=\inf_{\substack{x_0\in U\\ U\text{是开集}}}\operatorname{diam}\big(f(U)\big),$$

其中 $\operatorname{diam}V=\sup\limits_{y_1,y_2\in V}d_Y(y_1,y_2)$ 为 V 的**直径**。

例子. 我们有一个在一点处振幅很大的函数：

$$f(x)=\begin{cases}\sin\left(\frac{1}{x}\right), & x\neq 0;\\ 0, & x=0.\end{cases}$$

很明显，$\omega(f,0)=2$。我们自然知道 f 在 0 处是不连续的。

振幅消失实际上是连续性的刻画：

引理 131. *$f:X\to Y$ 是距离空间 (X,d_X) 和 (Y,d_Y) 之间的映射。f 在 x_0 处连续当且仅当 $\omega(f,x_0)=0$。*

证明: 假设 f 在 x_0 处连续，那么，对任意的 $\varepsilon>0$，存在 $\delta>0$，当 $d(x,x_0)<\delta$ 时（$\Leftrightarrow x\in B(x_0,\delta)\subset X$），我们有 $d\big(f(x),f(x_0)\big)<\varepsilon$（$\Leftrightarrow f(x)\in B\big(f(x_0),\varepsilon\big)\subset Y$）。所以，对于开集合 $U=B(x_0,\delta)$ 而言，$f(U)\subset B\big(f(x_0),\varepsilon\big)\subset Y$，从而 $\operatorname{diam}f(U)<2\varepsilon$。由于 ε 是任意选取的，所以 $\omega(f,x_0)=0$。

反过来，假设 $\omega(f,x_0)=0$，即 $\inf\limits_{\substack{x_0\in U\\ U\text{是开集}}}\operatorname{diam}\big(f(U)\big)=0$，所以对于任意的 $\varepsilon>0$，存在包含 x_0 的开集 $U\subset X$，使得

$$\operatorname{diam}\big(f(U)\big)<\delta.$$

由于 U 是开集，所以存在 $\delta>0$，使得 $B(x_0,\delta)\subset U$，从而 $\operatorname{diam}f\big(B(x_0,\delta)\big)<\varepsilon$，所以 $d(x,x_0)<\delta$ 时（$\Leftrightarrow x\in B(x_0,\delta)$），$d_Y\big(f(x),f(x_0)\big)\leqslant\operatorname{diam}f\big(B(x_0,\delta)\big)<\varepsilon$。这表明 f 在 x_0 处连续。 □

对于不连续点，我们有如下的刻画：

引理 132. $f:X\to Y$ 是距离空间 (X,d_X) 和 (Y,d_Y) 之间的（任意）映射。对任意 $\varepsilon>0$，集合 $\Omega_\varepsilon(f)=\{x\in X|\omega(f,x)\geqslant\varepsilon\}$ 是 X 中的闭集。

证明: 按照闭集的定义，只需要说明它的补集 $X-\Omega_\varepsilon(f)=\{x\in X|\omega(f,x)<\varepsilon\}$ 为开集即可。首先，我们注意到 $x_0\in X-\Omega_\varepsilon(f)$ 等价于存在包含 x_0 的开集 U，使得 $\mathrm{diam}f(U)<\varepsilon$。据此，我们知道，对于任意的 $x\in U$，利用 $\mathrm{diam}f(U)<\varepsilon$，所以 $x\in X-\Omega_\varepsilon(f)$，即 $U\subset X-\Omega_\varepsilon(f)$。这说明 $X-\Omega_\varepsilon(f)$ 是开集。 □

我们现在在 $\mathbb{R}$ 上定义所谓的**测度为零的**集合（你可以认为这是所谓的长度为零的集合）。首先，给定一个有限区间 I，我们定义 $|I|$ 为其长度，即右端点减掉左端点的值。

定义 133. $X\subset\mathbb{R}$ 是子集，如果对任意的 $\varepsilon>0$，总存在可数个有界闭区间 $\{I_k\}_{k\geqslant1}$，使得 $X\subset\bigcup\limits_{k\geqslant1}I_k$ 并且 $\sum\limits_{k=1}^{\infty}|I_k|<\varepsilon$，我们就称 X 是一个**零测集**。

注记. 我们可以将定义中的有界闭区间换成有界开区间，这样定义出的零测集与上述定义的是一致的。实际上，如果接受第二种定义，假设对任意的 $\varepsilon>0$，存在开区间 $\{I_k\}_{k\geqslant1}$，使得

$$X\subset\bigcup_{k\geqslant1}I_k,\quad\sum_{k=1}^{\infty}|I_k|<\varepsilon.$$

此时，我们可以选取闭区间 $\{\overline{I_k}\}_{k\geqslant1}$，我们自然有

$$X\subset\bigcup_{k\geqslant1}\overline{I_k},\quad\sum_{k=1}^{\infty}|I_k|=\sum_{k=1}^{\infty}|\overline{I_k}|<\varepsilon.$$

反之，如果接受第一种定义，假设对任意的 $\varepsilon>0$，存在有界闭区间 $\{K_k\}_{k\geqslant1}$，使得

$$X\subset\bigcup_{k\geqslant1}K_k,\quad\sum_{k=1}^{\infty}|K_k|<\varepsilon.$$

假设 $K_k=[a_k,b_k]$，我们令 $I_k=(a_k-\dfrac{1}{2^{k+1}}\varepsilon,b_k+\dfrac{1}{2^{k+1}}\varepsilon)$，那么开区间 I_k 的长度为 $|K_k|+\dfrac{\varepsilon}{2^k}$，所以

$$\sum_{k=1}^{\infty}|I_k|<\sum_{k=1}^{\infty}|K_k|+\sum_{k=1}^{\infty}\frac{\varepsilon}{2^k}=2\varepsilon.$$

所以，我们可以选取开区间 $\{\overline{I_k}\}_{k\geqslant1}$ 来覆盖 X。

命题 134. 可数个零测集的并集仍然是零测集。特别地，可数集是零测集。

证明: 假设 $\{X_k\}_{k\geqslant1}$ 是零测集，按照定义，对于每个 X_k 而言，对任意的 $\dfrac{\varepsilon}{2^k}$，存在有界闭区间的集合 $\{I_{k,i}\}_{i\geqslant1}$，使得

$$X_k\subset\bigcup_{i\geqslant1}I_{k,i},\quad\sum_{i=1}^{\infty}|I_{k,i}|<\frac{\varepsilon}{2^k}.$$

所以，这一些 $\{I_{k,i}\}_{k,i\geqslant 1}$ 可以作为 $\bigcup\limits_{k\geqslant 1} X_k$ 的覆盖的闭区间

$$\bigcup_{k\geqslant 1} X_k \subset \bigcup_{k\geqslant 1}\bigcup_{i\geqslant 1} I_{k,i}.$$

它们的总长度满足

$$\sum_{i,k=1}^{\infty} |I_{k,i}| < \sum_{k=1}^{\infty} \frac{\varepsilon}{2^k} = \varepsilon.$$

所以，$\bigcup\limits_{k\geqslant 1} X_k$ 是零测集。 □

我们用上面的工具来研究 Riemann 可积函数（f 可以在一个赋范线性空间中取值）：

引理 135. （关键的引理）假设 $f \in \mathcal{R}([a,b])$，对任意的 $\varepsilon > 0$，

$$\Omega_\varepsilon(f) = \big\{x \in X \;\big|\; \omega(f,x) \geqslant \varepsilon\big\}$$

是零测集。

证明: 任意给定正常数 $\delta > 0$，根据 Riemann 积分的定义，我们找两个阶梯函数 F 和 ψ，使得它们都对应着相同的分划 $\sigma = \{a_0 < a_1 < \cdots < a_n\}$ 并且

$$|f(x) - F(x)| \leqslant \psi(x), \quad \int_a^b \psi < \delta\varepsilon.$$

我们将区间 $[a_i, a_{i+1}]$ 分成两类：

- ψ 在 $[a_i, a_{i+1}]$ 上的取值小于 $\dfrac{1}{2}\varepsilon$。

  在这一类区间上面，我们来研究 f 的振幅，其中 $x, y \in [a_i, a_{i+1}]$：

  $$\begin{aligned} |f(x) - f(y)| &\leqslant |f(x) - F(x)| + |f(y) - F(y)| + \underbrace{|F(x) - F(y)|}_{=0} \\ &\leqslant |\psi(x)| + |\psi(y)| < \varepsilon. \end{aligned}$$

  据此，我们知道 $\Omega_\varepsilon(f)$ 与此类区间的交集为空集。

- ψ 在 $[a_i, a_{i+1}]$ 上的取值大于等于 $\dfrac{1}{2}\varepsilon$。

  我们将这些小区间记作 $I_1, \cdots, I_m$。按照上面的推理，这一类区间覆盖了 $\Omega_\varepsilon(f)$。根据不等式 $\displaystyle\int_a^b \psi < \delta\varepsilon$，我们知道

  $$\sum_{j\leqslant m} \frac{1}{2}\varepsilon |I_j| \leqslant \int_a^b \psi < \delta\varepsilon$$

  从而，$|I_j|$ 的长度总和小于 2δ。

由于 δ 是任意选取的并且 $\Omega_\varepsilon(f)\subset\bigcup_{i=1}^{n}I_m$，所以 $\Omega_\varepsilon(f)$ 是零测集。 □

由于 f 在 x 处连续当且仅当 $\omega(f,x)=0$，从而 f 的不连续点具有如下的刻画：

$$\big\{x\in I\ \big|\ f\text{在}x\text{处不连续}\big\}=\bigcup_{n\geqslant 1}\Omega_{\frac{1}{n}}(f).$$

所以，如果 $f\in\mathcal{R}(I)$，那么 f 的不连续点的集合是零测集（因为可数个零测集的并集还是零测集）。

我们现在来证明 Lebesgue 定理，它给出了 Riemann 可积函数和连续函数之间的基本关联：

定理 136 (Lebesgue)**.** *$f\in\mathcal{R}\big([a,b]\big)$ 当且仅当 f 有界并且其不连续点所构成的集合是零测集。*

证明： 我们做如下的准备工作：

- 选取 $M>0$，使得对任意的 $x\in[a,b]$，都有 $|f(x)|\leqslant M$。
- 令

$$A=\Omega_{\frac{\varepsilon}{2(b-a)}}(f)=\big\{x\in[a,b]\big|\omega(f,x)\geqslant\frac{\varepsilon}{2(b-a)}\big\}.$$

那么，A 是一个闭集也是一个零测集。由于 A 有界，它还是一个紧集。

对于 A，由于它是零测集，所以可以选取可数个开区间 $I_{ii\geqslant 1}$ 作为 A 的一个开覆盖，并且它们的总长度小于 $\dfrac{\varepsilon}{2M}$，即

$$\sum_{i=1}^{\infty}|I_i|<\frac{\varepsilon}{2M}.$$

根据 A 的紧性，存在有限个小区间 $I_1,\cdots,I_m$，使得 $A\subset U=I_1\cup\cdots\cup I_m$，我们自然还有

$$\sum_{i=1}^{m}|I_i|<\frac{\varepsilon}{2M}.$$

令 $K=I-U=I\cap(X-U)$。这是两个闭集的交集，所以是闭集。另外，K 是有界的，所以 K 是紧集。对任意的 $y\in K$，按照振幅的定义，都存在包含 y 的开区间 J_y，使得对任意的 $t_1,t_2\in J_y$，我们都有

$$|f(t_1)-f(t_2)|\leqslant\frac{\varepsilon}{2(b-a)}.$$

利用紧性，我们能找到有限个小区间 $J_{y_1},\cdots,J_{y_\ell}$，使得 $K\subset V=J_{y_1}\cup\cdots\cup J_{y_\ell}$。

我们现在将 $I_1,\cdots,I_m$ 和 $J_{y_1},\cdots,J_{y_\ell}$ 的端点按照大小顺序排成一列，加上 a 和 b，就得到了 $[a,b]$ 的一个分划 $a_0<a_1<\cdots<a_{n-1}<a_n$。所以，每个小区间 $[a_i,a_{i+1}]$ 要么完全落在某个 I_i 中，要么完全落在某个 I_{y_j} 中。我们现在构造阶梯函数 F 和 ψ 来逼近 f：

$$F(x)=\begin{cases}0, & x=\text{某个}a_i;\\ 0, & x\in(a_i,a_{i+1})\text{并且}(a_i,a_{i+1})\text{包含在某个}I_j\text{中};\\ f(\frac{a_i+a_{i+1}}{2}), & x\in(a_i,a_{i+1})\text{但是}(a_i,a_{i+1})\text{不包含在任何}I_j\text{中}.\end{cases}$$

以及

$$\psi(x)=\begin{cases}M, & x=\text{某个}a_i;\\ M, & x\in(a_i,a_{i+1})\text{并且}(a_i,a_{i+1})\text{包含在某个}I_j\text{中};\\ \frac{\varepsilon}{2(b-a)}, & x\in(a_i,a_{i+1})\text{但是}(a_i,a_{i+1})\text{不包含在任何}I_j\text{中}.\end{cases}$$

按照定义，我们有 $|f(x)-F(x)|\leqslant\psi(x)$。为此，我们只需要分情况讨论，在上面前两种情况下，这是显然的；如果 $x\in(a_i,a_{i+1})$但是(a_i,a_{i+1})不包含在任何I_j中，那么，存在 j_0，使得 $x\in J_{j_0}$，特别地，$x,\frac{a_i+a_{i+1}}{2}\in J_{j_0}$，从而，根据 J-型区间的定义，我们有

$$\left|f\left(\frac{a_i+a_{i+1}}{2}\right)-f(x)\right|<\frac{\varepsilon}{2(b-a)}.$$

这就是 $|f(x)-F(x)|\leqslant\psi(x)$。最终，我们验证 $\int_a^b\psi\leqslant\varepsilon$：

$$\begin{aligned}\int_a^b\psi&\leqslant\sum_{i=1}^m|I_i|\times M+(b-a)\times\frac{\varepsilon}{2(b-a)}\\&\leqslant\frac{\varepsilon}{2}+\frac{\varepsilon}{2}=\varepsilon.\end{aligned}$$

这就证明了 f 是 Riemann 可积的。 □

22 积分的基本性质，积分余项的 Taylor 公式，反常积分，反常积分的收敛性：控制收敛，面积法，Euler 常数，Wallis 积分与 Stirling 公式

二零一九年十二月二日，星期一，晴天

给定有界闭区间 $[a,b]$，我们知道连续函数的空间 $C\big([a,b]\big)$ 是 $\mathcal{R}\big([a,b]\big)$ 的线性子空间。我们在 $C\big([a,b]\big)$ 上配备范数 $\|\cdot\|_\infty$，即

$$\|f\|_\infty = \sup_{x\in[a,b]} |f(x)|.$$

此时，我们有映射

$$\int_a^b : C\big([a,b]\big) \longrightarrow \mathbb{R}, \quad f \mapsto \int_a^b f.$$

这是一个连续线性映射：因为对任意的 $f,g \in C\big([a,b]\big)$，我们有

$$\Big|\int_a^b f - \int_a^b g\Big| \leqslant \int_a^b |f-g| \leqslant \int_a^b \|f-g\|_\infty = (b-a)d_\infty(f,g).$$

特别地，如果 $f_n \to f$ 是 $C\big([a,b]\big)$ 一致收敛的序列，那么

$$\lim_{n\to\infty}\int_a^b f_n = \int_a^b f.$$

练习. *存在 $f_n, f \in C\big([a,b]\big)$，使得对任意的 $x\in[a,b]$，当 $n\to\infty$ 时，$f_n(x)\to f(x)$，但是*

$$\lim_{n\to\infty}\int_a^b f_n \neq \int_a^b f.$$

我们上一次课证明了关于 Riemann 积分的 Lebesgue 定理：$f\in\mathcal{R}\big([a,b]\big)$ 当且仅当 f 有界并且其不连续点所构成的集合是零测集。作为应用，我们有

推论 137 (对值域进行复合)**.** *$f: I\to J\subset\mathbb{R}$ 是可积的，$g: J\to\mathbb{R}$（或某个赋范线性空间 V）是连续映射，那么 $g\circ f$ 是可积的。特别地，$f: I\to\mathbb{C}$ 是可积的并且 $|f(x)|\geqslant\delta>0$，那么 $\dfrac{1}{f}$ 是可积的。*

证明: 这是因为如果 f 在 x_0 处连续，那么 $g\circ f$ 在 x_0 处连续。所以，$g\circ f$ 的不连续点的集合是 f 的不连续点的子集，所以仍然是零测集。

对于 $f: I\to\mathbb{C}$，我们按照实部和虚部分解，有 $f=f_1+if_2$，此时，f_1 和 f_2 都是可积的。那么

$$\frac{1}{f} = \frac{f_1}{(f_1)^2+(f_2)^2} - i\frac{f_2}{(f_1)^2+(f_2)^2},$$

是可积的。 □

推论 138. $f:[a,b]\to\mathbb{R}_{\geqslant 0}$ 是 *Riemann* 可积的。那么，$\int_a^b f=0$ 当且仅当 $\{x\in[a,b]|f(x)\neq 0\}$ 是零测集。

证明： 如果 $\int_a^b f=0$，为了说明 $\big\{x\in[a,b]\big|f(x)\neq 0\big\}$ 是零测集，我们证明在任意连续点 x_0 处，$f(x_0)=0$：因为 Lebesgue 定理，$\big\{x\in[a,b]\big|f(x)\neq 0\big\}$ 只能是不连续点，所以测度为零。我们用反证法：假设 f 在 x_0 处连续，但是 $f(x_0)\neq 0$，所以 $f(x_0)>0$。特别地，存在正数 δ 和 ε，使得对任意的 $x\in(x_0-\delta,x_0+\delta)$，$f(x)\geqslant\varepsilon$。我们现在构造阶梯函数 $\varphi(x)$ 使得

$$\varphi(x)=\begin{cases}\varepsilon, & x\in(x_0-\delta,x_0+\delta),\\ 0, & x\notin(x_0-\delta,x_0+\delta).\end{cases}$$

由于 $f:[a,b]\to\mathbb{R}_{\geqslant 0}$，所以 $f\geqslant\varphi$，从而 $\int_a^b f\geqslant\int_a^b\varphi=\varepsilon\delta$，矛盾。

反过来，我们假设 $\big\{x\in[a,b]\big|f(x)\neq 0\big\}$ 是零测集。根据 $f\geqslant 0$，存在阶梯函数 $\varphi(x)\geqslant 0$，使得 $\varphi(x)\leqslant f(x)$ 并且 $\int_a^b(f-\varphi)\leqslant\varepsilon$：我们只需要用足够细的 Darboux 下和所对应的阶梯函数即可！

由于 $0\leqslant\varphi(x)\leqslant f(x)$，所以 $\big\{x\in[a,b]\big|\varphi(x)\neq 0\big\}\subset\big\{x\in[a,b]\big|f(x)\neq 0\big\}$，从而是零测集。由于 φ 是阶梯函数，所以 $\big\{x\in[a,b]\big|\varphi(x)\neq 0\big\}$ 是有限集合，所以 $\int_a^b\varphi=0$。根据 φ 和 f 之间的关系，我们有

$$\int_a^b f=\int_a^b(f-\varphi)+\int_a^b\varphi\leqslant\varepsilon+0.$$

令 $\varepsilon\to 0$，我们就证明了 $\int_a^b f=0$。 □

另外一个推论讲的是在一个零测集上改变一个函数的值不会改变这个函数的积分：

推论 139. $f,g\in\mathcal{R}\big([a,b]\big)$ 是 *Riemann* 可积的，除去可数个点之外，它们是相同的，即

$$\big\{x\in[a,b]\big|f(x)\neq g(x)\big\}$$

是可数集，那么，$\int_a^b f=\int_a^b g$。

证明： 根据线性，我们只需要证明，如果 $\big\{x\in[a,b]\big|f(x)\neq 0\big\}$ 是零测集，那么，$\int_a^b f=0$。为此，我们可以用上下积分的定义（它们都等于函数的积分）：对任意的 x，存在阶梯函数 φ，使得

$$\varphi(x)\geqslant f(x),\ \ \forall x\in[a,b];\qquad \varepsilon+\int_a^b f\geqslant\int_a^b\varphi.$$

根据 f 的性质，我们知道 $\big\{x\in[a,b]\big|\varphi(x)<0\big\}$ 是零测集。由于 $\varphi(x)$ 是阶梯函数，所以 $\big\{x\in[a,b]\big|\varphi(x)<0\big\}$ 是零测集只能是有限点集，从而存在一个分划 σ，其中 $\sigma=\{a_0<a_1<\cdots<a_n\}$，

使得对每个 $i \geqslant 1$，$\varphi\Big|_{(a_{i-1},a_i)}$ 是非负的常数。特别地，我们有

$$\int_a^b \varphi \geqslant 0 \quad \Rightarrow \int_a^b f \geqslant -\varepsilon.$$

令 $\varepsilon \to 0$，我们就得到 $\int_a^b f \geqslant 0$；同理，我们有 $\int_a^b f \leqslant 0$。这就证明了命题。 □

我们现在离开 Lebesgue，再回到 Newton-Leibniz 公式的场合。利用这个公式，我们可以给出积分余项的 Taylor 展开公式：

命题 140 (积分余项的 Taylor 公式)**.** *考虑区间 $I = [a, b]$ 上的 $m+1$ 次连续可微函数 $f \in C^{m+1}(I)$，其中 $m \in \mathbb{Z}_{\geqslant 0}$（函数可以在赋范线性空间中取值），我们有*

$$f(b) = \sum_{k=0}^{m} \frac{(b-a)^k}{k!} f^{(k)}(a) + \int_a^b \frac{(b-x)^m}{m!} f^{(m+1)}(x)dx.$$

证明： 对 $m = 0$，这就是 Newton-Leibniz 公式；对于一般的 m，我们进行归纳：假设命题对 m 是成立的，即

$$f(b) = \sum_{k=0}^{m} \frac{(b-a)^k}{k!} f^{(k)}(a) + \int_a^b \frac{(b-x)^m}{m!} f^{(m+1)}(x)dx,$$

所以，利用 Newton-Leibniz 公式，我们有

$$\begin{aligned} f(b) &= \sum_{k=0}^{m} \frac{(b-a)^k}{k!} f^{(k)}(a) + \int_a^b \frac{(b-x)^m}{m!} f^{(m+1)}(x)dx \\ &= \sum_{k=0}^{m} \frac{(b-a)^k}{k!} f^{(k)}(a) + \int_a^b \frac{(b-x)^m}{m!} \left(f^{(m+1)}(a) + \int_a^x f^{(m+2)}(y)dy \right) dx \\ &= \sum_{k=0}^{m+1} \frac{(b-a)^k}{k!} f^{(k)}(a) + \int_a^b \frac{(b-x)^m}{m!} \left(\int_a^x f^{(m+2)}(y)dy \right) dx. \end{aligned}$$

我们对最后一项用分部积分公式（注意到边界项是零）：

$$\begin{aligned} f(b) &= \sum_{k=0}^{m+1} \frac{(b-a)^k}{k!} f^{(k)}(a) - \int_a^b \left(\frac{(b-x)^{m+1}}{(m+1)!} \right)' \left(\int_a^x f^{(m+2)}(y)dy \right) dx \\ &= \sum_{k=0}^{m+1} \frac{(b-a)^k}{k!} f^{(k)}(a) + \int_a^b \frac{(b-x)^{m+1}}{(m+1)!} \left(\int_a^x f^{(m+2)}(y)dy \right)' dx \end{aligned}$$

由于 $f^{(m+2)}(x)$ 是连续函数，所以 $\int_a^x f^{(m+2)}(y)dy$ 是 $f^{(m+2)}$ 的原函数，从而，$\left(\int_a^x f^{(m+2)}(y)dy \right)' = f^{(m+2)}(x)$，这就完成了归纳证明。 □

反常积分的概念

如果一个函数不是定义在一个有界闭区间上，我们也可以定义积分，这就是所谓的反常积分。给定函数 $f:[a,b)\to\mathbb{R}$，其中 b 可以是 $+\infty$。假设对任意的 $b^-\in[a,b)$，f 是区间 $[b,b^-]$ 上的 Riemann 可积函数。如果极限

$$\lim_{\substack{b^-\to b\\ b^-<b}}\int_a^{b^-}f(x)dx$$

存在，我们就称 f 在 $[a,b)$ 上 **(反常) 可积**并记

$$\int_a^b f(x)dx=\lim_{b^-\to b}\int_a^{b^-}f(x)dx.$$

我们经常说 f 的（反常）积分在 $[a,b)$ 上收敛。我们必须指出，如果 $b<\infty$，这个符号是有欺骗性的，因为从符号本身看不出来这是一个反常积分，所以，我们每次研究这样的问题的时候要先搞清楚函数的定义域。类似地，我们可以 $(a,b]$ 上定义反常积分。

在开区间 (a,b) 上定义反常积分需要额外的小心（其中 a 和 b）可以取正负无穷：假设对任意的 $c\in(a,b)$，使得 f 在 $(a,c]$ 和 $[c,b)$ 都反常可积，那么我们就称 f 在开区间 (a,b) 上**反常可积**。

注记. *根据极限的线性，我们知道在一个区间 I 上的反常可积的函数构成一个 $\mathbb{R}$-线性空间。*

我们对定义用如下的几个例子加以解释：

1) 考虑函数 $f(x)=\sin x$ 在 $[0,\infty)$ 上的反常积分的收敛性。首先，如果我们选取 $x_n=2n\pi$，那么，容易看到，

$$\lim_{n\to\infty}\int_a^{x_n}f(x)dx=\lim_{n\to\infty}0=0.$$

然而，如果我们换一个收敛到无穷的子列 $y_n=2n\pi+\dfrac{\pi}{2}$，那么

$$\lim_{n\to\infty}\int_a^{y_n}f(x)dx=\lim_{n\to\infty}1=1.$$

所以，$\lim\limits_{b^-\to\infty}\int_a^{b^-}f(x)dx$ 并不存在，因为函数极限的存在要求所算出来的极限值不依赖于子列的选取。

2) 考虑函数 $f(x)=\sin x$ 在 $(-\infty,\infty)$ 上的可积性。它自然不可积分，因为前面的一个例子已经说明了这一点。然而，

$$\lim_{R\to\infty}\int_{-R}^{R}f(x)dx=\lim_{R\to\infty}0=0.$$

所以，在开区间上定义积分时 (a,b) 采取类似于 $\lim\limits_{\varepsilon\to0^+}\int_{a+\varepsilon}^{b-\varepsilon}f(x)dx$ 的方式是不可取的。

我们来研究几个最为经典的例子，它们展现了函数衰减／增长的速度对（反常）积分收敛性的影响。大家应该熟记这些计算和结论：

例子. *1)* $\int_1^\infty \frac{1}{x^\alpha}$，讨论 α 的范围。

当 $\alpha \neq -1$ 时，对任意的 $M > 0$，我们有

$$\int_1^M \frac{1}{x^\alpha} = \frac{1}{\alpha+1}\left(M^{\alpha+1} - 1\right).$$

从而，为了要求 $M \to \infty$ 极限存在，需要 $\alpha < -1$（$\alpha = -1$ 的情况可以类似地的排除)。

2) $\int_0^1 \frac{1}{x^\alpha}$，讨论 α 的范围，其中反常积分的积分区域是 $(0, 1]$。

当 $\alpha \neq -1$ 时，对任意的 $\varepsilon > 0$，我们有

$$\int_\varepsilon^1 \frac{1}{x^\alpha} = \frac{1}{\alpha+1}\left(1 - \varepsilon^{\alpha+1}\right).$$

从而，为了要求 $\varepsilon \to 0$ 极限存在，需要 $\alpha > -1$（$\alpha = -1$ 的情况可以类似地的排除)。

3) $\int_2^\infty \frac{1}{x\log x} dx$。

对任意的 $M > 0$，我们有

$$\int_2^M \frac{1}{x\log x} dx = \int_2^M \frac{1}{\log x} d\log x = \log\log M - \log\log 2.$$

当 $M \to \infty$ 时，上式的极限是无穷大，从而这个反常积分不收敛。

4) $\int_{100}^\infty \frac{1}{x\log x\log\log x}$。对任意的 $M > 0$，我们有

$$\int_{100}^M \frac{1}{x\log x\log\log x} dx = \int_2^M \frac{1}{\log\log x} d\log\log x = \log\log\log M - \log\log\log(100).$$

当 $M \to \infty$ 时，上式的极限是无穷大，从而这个反常积分不收敛。

5) $\int_1^\infty \frac{1}{x\log^\alpha(x)}$，其中 $\alpha > 1$。

对任意的 $M > 0$，我们有

$$\int_2^M \frac{1}{x\log^\alpha(x)} dx = \int_2^M \frac{1}{\log^\alpha(x)} d\log x = \frac{1}{-\alpha+1}\left(\frac{1}{(\log M)^{\alpha-1}} - \frac{1}{(\log 2)^{\alpha-1}}\right).$$

当 $M \to \infty$ 时，上式的极限是无穷大，从而这个反常积分不收敛。

6) 在假设函数 f 在 $[0, \infty)$ 上可积分，这并不意味着当 $x \to \infty$ 时，$f \to 0$。

比如说，我们构造如下的函数：对任意的 $n \in \mathbb{Z}_{\geqslant 1}$，我们要求

$$f(x) = \frac{1}{n}, \quad x \in [n - \frac{1}{n^3}, n].$$

在其余的地方，$f(x) \equiv 0$。此时，$\int_0^\infty f = \sum_{n=1}^\infty \frac{1}{n^2}$ 是收敛的。很明显，$\lim_{x\to\infty} f(x) \neq 0$。

关于不定积分，我们有如下的收敛判别法（不令人惊讶）：

引理 141. *f 和 F 在区间 I 上定义，即 $f: I \to \mathbb{R}$，$F: I \to \mathbb{R}$ 并且对任意的有界闭区间 $J \subset I$，f 和 F 均为 J 上的 Riemann 可积函数。假设对任意的 $x \in I$，我们都有*

$$\big|f(x)\big| \leqslant F(x).$$

如果 F 在区间 I 上的反常积分收敛，那么 f 在区间 I 上的反常积分也收敛。

证明: 我们把它留成作业题。 □

面积法

作为积分的基本应用，我们用所谓的面积方法来研究级数的大小。我们大多假设 $f(x)$ 是一个单调的函数（有时候它不单调，我们就需要更细致地分析，但是思路是一致的），比如说是递增的，我们考虑 f 在区间 $[1, n]$ 上的积分。我们可以构造两个阶梯函数：

$$\underline{f} = \sum_{k=1}^{n-1} f(k)\mathbf{1}_{[k,k+1]}(x), \quad \overline{f} = \sum_{k=1}^{n-1} f(k+1)\mathbf{1}_{[k,k+1]}(x).$$

很明显，$\underline{f} \leqslant f \leqslant \overline{f}$，所以

$$\int_1^n \underline{f} \leqslant \int_1^n f \leqslant \int_1^n \overline{f}.$$

这表明

$$\sum_{k=1}^{n-1} f(k) \leqslant \int_1^n f, \quad \sum_{k=2}^{n} f(k) \geqslant \int_1^n f.$$

这就可以用积分 $\displaystyle\int_1^n$ 给出级数 $\displaystyle\sum_{k=1}^{\infty} f(k)$ 的部分和的一个估计。由于上面两个阶梯函数的积分我们形象地将它们看成是直方图下的面积，所以我们也称这个方法为面积法。

注记. *这个方法的核心是***用一个方便计算的积分来逼近级数***。*

我们来看几个经典的例子：

例子. *1)* $\displaystyle\sum_{1\leqslant n\leqslant N} n^\alpha = \frac{N^{\alpha+1}-1}{\alpha+1} + O(N^\alpha)$，*其中 $\alpha > 0$。也就是说，存在常数 $M > 0$，使得*

$$\left|\frac{\sum_{1\leqslant n\leqslant N} n^\alpha - \frac{N^{\alpha+1}-1}{\alpha+1}}{N^\alpha}\right| \leqslant M.$$

我们用 $f(x) = x^\alpha$ 的积分来控制 $\displaystyle\sum_{1\leqslant n\leqslant N} n^\alpha$：

$$\sum_{1\leqslant n\leqslant N} n^\alpha \leqslant \int_2^{N+1} x^\alpha dx = \frac{(N+1)^\alpha - 2^\alpha}{\alpha+1}.$$

这表明

$$\sum_{1\leqslant n\leqslant N} n^\alpha - \frac{N^{\alpha+1}-1}{\alpha+1} \leqslant (N+1)^{\alpha+1} - N^{\alpha+1} + O(1).$$

根据 *Lagrange* 中值定理，我们有 $(N+1)^{\alpha+1} - N^\alpha \leqslant CN^\alpha$，其中 C 是一个常数。我们还有

$$1 + \sum_{2\leqslant n\leqslant N} n^\alpha \geqslant 1 + \int_1^N x^\alpha dx = 1 + \frac{(N-1)^\alpha - 1}{\alpha+1}.$$

这表明

$$\sum_{1\leqslant n\leqslant N} n^\alpha - \frac{N^{\alpha+1}-1}{\alpha+1} \leqslant N^{\alpha+1} - (N-1)^{\alpha+1} + O(1) = O(N^\alpha).$$

这就证明了命题。

2) 关于 $\sum_{k=1}^{n} \log k$ 的增长估计。

利用函数 $f(x) = \log x$。我们首先有

$$\sum_{k=1}^{n} \log k \geqslant \int_1^n \log x = n\log n - n + 1.$$

我们还有

$$\sum_{k=1}^{n} \log k \leqslant \int_2^{n+1} \log x = (n+1)\log(n+1) - n + 1 - 2\log 2.$$

我们现在说明 $(n+1)\log(n+1) - n + 1 - \log 4 \leqslant n\log n - n + 1 + \log n$，通过代数变形，这等价于 $\left(1+\frac{1}{n}\right)^{n+1} \leqslant 4$，这对于比较大的 n 自然成立（对所有的 n 其实都成立）。综上所述，我们有

$$n\log n - n + 1 \leqslant \sum_{k=1}^{n} \log k \leqslant n\log n - n + 1 + \log n.$$

这个等价于

$$e \leqslant \frac{e^n n!}{n^n} \leqslant en.$$

我们马上就证明所谓的 *Stirling* 公式，这将给出更精细的估计。

3) Euler 常数 γ。

定义数列 $a_n = 1 + \frac{1}{2} + \frac{1}{3} + \cdots + \frac{1}{n} - \log n$，一个不平凡的事实是这个数列的极限是存在的，我们把这个极限定义为 *Euler* 常数：

$$\gamma = \lim_{n\to\infty} \left(1 + \frac{1}{2} + \frac{1}{3} + \cdots + \frac{1}{n} - \log n\right).$$

对 x^{-1} 这个函数用面积法，我们有

$$1 + \frac{1}{2} + \frac{1}{3} + \cdots + \frac{1}{n} \geqslant \int_1^{n+1} \frac{dx}{x} = \log(n+1).$$

我们还有

$$1+\frac{1}{2}+\frac{1}{3}+\cdots+\frac{1}{n}\leqslant 1+\int_2^{n+1}\frac{dx}{x}=\log(n+1)-\log 2+1.$$

由此可见 $1+\dfrac{1}{2}+\dfrac{1}{3}+\cdots+\dfrac{1}{n}$ 的增长和 $\log(n)$ 是一致的。实际上，上面两个不等式表明

$$\log\left(1+\frac{1}{n}\right)\leqslant 1+\frac{1}{2}+\frac{1}{3}+\cdots+\frac{1}{n}-\log n\leqslant \log\left(1+\frac{1}{n}\right)-\log 2+1$$

然而，这不足以说明 γ 的存在性，所以，我们对 a_n 要做更精确的表述：

$$a_n=\sum_{k=1}^{\infty}\left(\frac{1}{k}-\int_k^{k+1}\frac{dx}{x}\right)=\sum_{k=1}^{\infty}\int_k^{k+1}\left(\frac{1}{k}-\frac{1}{x}\right)dx.$$

由此可见，这是一个单调递增的序列，上面的证明已经说明 $\{a_n\}_{n\geqslant 1}$ 是有界的，所以极限存在。

我们将在作业题中证明

$$a_n-\gamma=O(\frac{1}{n}).$$

根据这个计算，我们还知道

$$1-\frac{1}{2}+\frac{1}{3}-\frac{1}{4}+\cdots+(-1)^{n-1}\frac{1}{n}+\cdots=\log 2.$$

注记. 人们猜想 γ 应该是无理数（超越数），这个猜想到今天还没有被证明。$\log 2$ 是无理数? 历史。

Wallis 积分与 Stirling 公式：三角函数积分的一个应用

历史上，Wallis 研究过如下的定积分

$$I_n=\int_0^{\frac{\pi}{2}}\sin^n x dx.$$

很明显，$I_0=\dfrac{\pi}{2}$，$I_1=1$。由于 $\sin x\leqslant 1$，我们知道 $\{I_n\}_{n\geqslant 1}$ 是单调递减的。我们利用积分的基本技术来推理数列 $\{I_n\}_{n\geqslant 1}$ 的递归关系：

$$\begin{aligned}I_{n+2}-I_n&=\int_0^{\frac{\pi}{2}}\sin^n x(-\cos^2 x)dx=-\frac{1}{n+1}\int_0^{\frac{\pi}{2}}\cos x d\big(\sin^{n+1}x\big)\\&=-\frac{\sin^{n+1}(x)}{n+1}\cos x\Big|_0^{\frac{\pi}{2}}-\frac{1}{n+1}\underbrace{\int_0^{\frac{\pi}{2}}\sin^{n+2}x dx}_{I_{n+2}}.\end{aligned}$$

从而，我们得到递推公式

$$I_{n+2}=\frac{n+1}{n+2}I_n.$$

据此，我们就可以得到

$$I_{2p}=\frac{(2p)!}{2^{2p}(p!)^2}\frac{\pi}{2},\quad I_{2p+1}=\frac{2^{2p}(p!)^2}{(2p+1)!}.$$

由于 $\{I_n\}_{n\geqslant 1}$ 是单调递减的，所以

$$\frac{n+1}{n+2}=\frac{I_{n+2}}{I_n}\leqslant 1.$$

从而，$\lim\limits_{n\to\infty}\dfrac{I_{n+1}}{I_n}=1$。另外，根据上面的计算，我们还知道

$$I_{2p}I_{2p+1}=\frac{\pi}{2(2p+1)}$$

根据这些结果，我们考虑序列 $\{x_n\}_{n\geqslant 1}$，其中

$$x_n=\sqrt{\frac{2n}{\pi}}I_n.$$

所以，$\dfrac{x_{n+2}}{x_n}=\dfrac{n+1}{\sqrt{n(n+2)}}\geqslant 1$，所以这是一个在奇数项或者偶数项递增的序列。再根据

$$x_{2p}x_{2p+1}=\frac{2p}{2p+1},$$

我们很容易得到 $\lim\limits_{n\to\infty}x_n=1$，所以

$$I_n\sim\sqrt{\frac{\pi}{2n}},\quad \text{即}\quad \lim_{n\to\infty}\frac{I_n}{\sqrt{\frac{\pi}{2n}}}=1.$$

这就是所谓的 **Wallis 积分的渐进公式**。

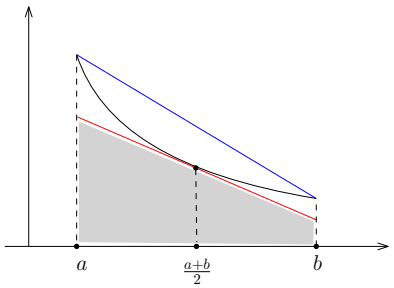
我们现在用 Wallis 积分的计算来推导 Stirling 公式：考虑数列 $\{a_n\}_{n\geqslant 1}$，其中

$$a_n=\frac{e^n n!}{n^{n+\frac{1}{2}}}.$$

我们要证明 $\lim\limits_{n\to\infty}a_n=\sqrt{2\pi}$。我们先比较相邻的两项：

$$\log\frac{a_n}{a_{n+1}}=\log\left(\frac{1}{e}\left(1+\frac{1}{n}\right)^{n+\frac{1}{2}}\right)=\left(n+\frac{1}{2}\right)\log\left(1+\frac{1}{n}\right)-1.$$

我们要来证明 $\log\left(\dfrac{a_n}{a_{n+1}}\right)>0$。为此，考虑 $\dfrac{1}{x}$ 在 $[a,b]$ 上的图像 Γ。我们注意到这是凸函数（算二阶导数），所以 Γ 在它在 $\frac{a+b}{2}$ 处的切线的上方，并且 Γ 在 a 和 b 处两点的连线的下方：



通过观察面积，我们知道

$$\frac{1}{2}(\frac{1}{a}+\frac{1}{b})(b-a)>\int_a^b\frac{1}{x}>\frac{1}{\frac{a+b}{2}}(b-a).$$

令 $b=n+1$，$a=n$，后一个不等式给出了 $\log\dfrac{a_n}{a_{n+1}}>0$，从而数列 $\{a_n\}_{n\geqslant 1}$ 是递减。特别地，它有极限。根据 Wallis 积分的计算，我们有

$$\sqrt{2(2p+1)}I_{2p+1}=\sqrt{2(2p+1)}\frac{2^{2p}(p!)^2}{(2p+1)!}\to\sqrt{\pi}.$$

从而，

$$\sqrt{\pi}=\lim_{n\to\infty}\frac{(n!)^2 2^{2n}}{(2n)!\sqrt{n}}=\lim_{n\to\infty}\frac{a_n^2}{a_{2n}\sqrt{2}}.$$

由于 a_n 的极限存在，所以

$$\lim_{n\to\infty}a_n=\lim_{n\to\infty}\frac{e^n n!}{n^{n+\frac{1}{2}}}=\sqrt{2\pi}.$$

我们得到 Stirling 公式

$$n!\sim\sqrt{2\pi n}\left(\frac{n}{e}\right)^n.$$

23 历史注记：Newton 和 Leibniz 时代的微积分，Leibniz 对 $\frac{\pi}{4}$ 的级数表示，微分与积分的交换定理，利用含参积分的计算积分

二零一九年十二月五日，星期四，晴天

上次课的开始，我们陈述并证明了下述积分算子：

$$\int_a^b : C\big([a,b]\big) \longrightarrow \mathbb{R}, \quad f \mapsto \int_a^b f.$$

是连续线性映射。特别地，如果 $f_n \to f$ 是 $C\big([a,b]\big)$ 一致收敛的序列，那么

$$\lim_{n\to\infty} \int_a^b f_n = \int_a^b f.$$

我们来讲一下这个命题的历史渊源（参见史记－牛顿莱布尼兹列传）。为此，我们考虑级数版本的陈述：假设 $\{f_n\}_{n\geqslant 1} \subset C\big([a,b]\big)$，考虑 $C\big([a,b]\big)$ 中的函数项级数

$$f = f_1 + f_2 + f_3 + \cdots$$

我们假设这个级数在 $\|\cdot\|_\infty$ 所定义的距离下是收敛到 f 的，即部分函数和

$$S_n = f_1 + f_2 + \cdots + f_n$$

一致收敛到 f。比如说，如果 $f = f_1 + f_2 + f_3 + \cdots$ 绝对收敛，即级数 $\sum\limits_{k=1}^{\infty} \|f_k\|_\infty$，那么这个性质是满足的。此时，我们有

$$\int_a^b \left(\sum_{k=1}^{\infty} f_k(x) \right) dx = \sum_{k=1}^{\infty} \left(\int_a^b f_k(x) dx \right).$$

Newton 对于微积分理论的很大贡献来自于他对（广义）二项式展开的研究，用今天的语言来讲，他研究可以用无穷幂级数的形式所给出的函数（他大概相信所有的函数都可以写长这样的形式）。由于时代的限制，他醉心于形如

$$f(x) = a_0 + a_1 x + a_2 x^2 + a_3 x^3 + \cdots$$

的函数的流数术（微分）与反流数术（积分）。他首先发现了形如 $(P + PQ)^{\frac{m}{n}}$ 型函数的二项式展开公式，用今天的话说，我们有

$$(1+x)^{\frac{m}{n}} = 1 + \sum_{k=1}^{\infty} \frac{\frac{m}{n}\left(\frac{m}{n}-1\right)\cdots\left(\frac{m}{n}-k\right)}{(k+1)!} x^k.$$

Newton 本人后来还推导了正弦三角函数 $\sin x$ 的级数表达式：

$$\sin x = x - \frac{1}{6}x^3 + \frac{1}{120}x^5 - \cdots.$$

他还能够将 $f(x)=a_0+a_1x+a_2x^2+a_3x^3+\cdots$ 的反函数用级数表达，姑且不论其严格性，这在当时是非常了不起的成就。由于当时他知道如何计算形如 x^m 的函数的微分与原函数，所以利用这些函数的级数表达，Newton 可以自由地计算一个给定函数的流数与反流数：

$$f'(x)=a_1+2a_2x^1+3a_3x^2+\cdots,\quad \int_0^1 f=a_1+\frac{a_2}{2}+\frac{a_3}{3}+\cdots.$$

当然，从现在的观点看，Newton 假设了对一个级数可以逐项的求微分或者求积分（我们需要某种一致收敛性），这自然不严格。无论怎样，Newton 的成就在于通过级数的理论可以对非常一般的函数进行微分和积分！另外，我们也看到了，级数在微积分的历史上是如何以优雅的方式在极限的概念出现之前登场！这是我们为什么在讲极限的理论的时候一定要同时研究级数（因为这是历史上最自然的极限过程）。

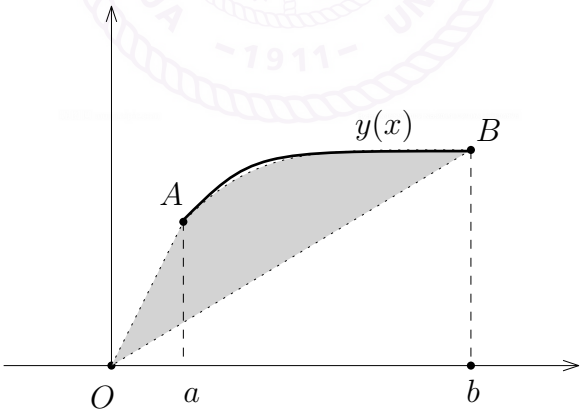
有了这个背景，我们来谈论一下 Leibniz 的某些贡献，他不仅引入了 dx 和 $\int$ 的符号，还和 Newton 分别证明了不同形式的 Newton-Leibniz 公式。假设 $y=y(x)$ 是一个（为了方便起见，递增的）连续可微的函数，利用我们所学的分部积分的知识，我们有

$$\int_a^b y(x)dx=-\int_a^b xy'(x)dx+by(b)-ay(a).$$

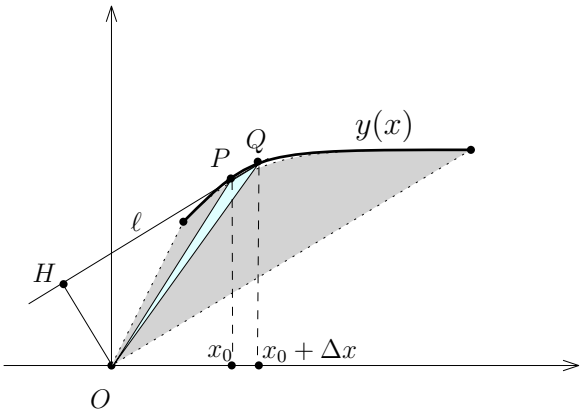
Leibniz 证明了分部积分公式是这个公式的代数变形：

$$\int_a^b y(x)dx=\frac{1}{2}\int_a^b z(x)dx+\frac{1}{2}\big(by(b)-ay(a)\big),\quad z(x)=y(x)-xy'(x).$$

他的公式有很清晰的几何意义：



函数 $y(x)$ 在 $[a,b]$ 区间上所对应的面积是 $\int_a^b y(x)dx$，表达式 $\frac{1}{2}\big(by(b)-ay(a)\big)$ 是三角形 $\triangle ObB$ 与三角形 $\triangle OaA$ 的面积之差，而 $\frac{1}{2}\int_a^b z(x)dx$ 是灰色区域的面积。为了说明 $\frac{1}{2}\int_a^b z(x)dx$ 是面积，我们观察下面的图（由于历史的原因，这个方法有时候被称作是微元法，实质上就是用 Riemann 和来定义积分）：



点 $P=\big(x_0,y(x_0)\big)$ 落在函数图像上，过这一个点的切线 ℓ 的方程为

$$y-y(x_0)=y'(x_0)(x-x_0).$$

从而，O 到 ℓ 的距离为 $|OH|=\dfrac{|y(x_0)-y'(x_0)x_0|}{\sqrt{y'(x_0)^2+1}}=\dfrac{|z(x_0)|}{\sqrt{y'(x_0)^2+1}}$。假设 $Q=\big(x_0+\Delta x,y(x_0+\Delta x)\big)$，当 Δx 很小的时候，我们可以认为上面淡蓝色三角形的底边 PQ 和 ℓ 是同一条直线，它高是 OH。由于 $|PQ|=\sqrt{y'(x_0)^2+1}\Delta x$，所以，

$$|\triangle OPQ|=\frac{1}{2}\times\frac{|z(x_0)|}{\sqrt{y'(x_0)^2+1}}\times\sqrt{y'(x_0)^2+1}\Delta x=\frac{1}{2}|z(x_0)|\Delta x.$$

将所有这样的小三角形加起来自然是灰色区域的面积。

Leibniz 的时代自然知道如何计算圆的面积，所以，四分之一圆弧

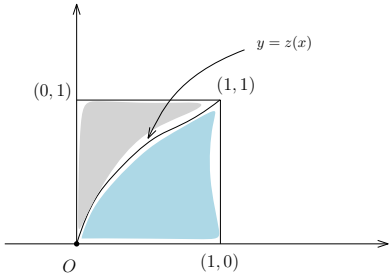
$$\big\{(x,y)\big|(x-1)^2+y^2=1,x\in[0,1],y\geqslant 0\big\}$$

下的面积是 $\dfrac{1}{4}\pi$。此时，函数 $y=\sqrt{2x-x^2}$，所以我们有

$$\frac{1}{\pi}=\int_0^1 y(x)dx=\int_0^1\sqrt{2x-x^2}dx.$$

我们注意到，$\sqrt{2x-x^2}$ 的 Taylor 展开或者级数展开并不容易计算。Leibniz 利用他的分部积分公式得到

$$\frac{\pi}{4}=\int_0^1 y(x)dx=\frac{1}{2}\int_0^1 z(x)dx+\frac{1}{2},\quad z(x)=y(x)-xy'(x)=\sqrt{\frac{x}{2-x}}.$$



为了计算 $\int_0^1 z(x)dx$，我们注意到这加上灰色区域的面积恰好是整个正方形的面积（这只是直观解释，我们将要在作业里证明下面的等式），即

$$\int_0^1 z(x)dx + \int_0^1 x(z)dz = 1.$$

所以，

$$\frac{\pi}{4} = \int_0^1 y(x)dx = 1 - \frac{1}{2}\int_0^1 x(z)dz = 1 - \int_0^1 \frac{z^2}{1+z^2}dz.$$

我们可以想象在 Leibniz 那个时代应该没有换元积分公式等技术，所以这种直观的几何方法反而是最自然最简单的思考方式。这种变形所得到的被积函数容易展开成级数的形式，所以 Newton 的反流数术可以派得上用场：

$$\frac{\pi}{4} = \int_0^1 \frac{1}{1+z^2}dz = \int_0^1 \sum_{k=0}^{\infty}(-1)^k z^{2k}dz = \sum_{k=0}^{\infty}\frac{(-1)^k}{2k+1}.$$

这就证明我们在第四次课中提到的重要极限：

$$1 - \frac{1}{3} + \frac{1}{5} - \cdots + \frac{(-1)^{n-1}}{2n-1} + \cdots = \frac{\pi}{4}.$$

这是 Leibniz 的原始证明。当然，为了使得最后一步推理严格，我们注意到，对任意的 ε，在 $[0, 1-\varepsilon]$ 上，级数 $\sum_{k=0}^{\infty}(-1)^k z^{2k}$ 一致收敛到 $\frac{1}{1+z^2}$，所以，我们有

$$\int_0^{1-\varepsilon} \frac{1}{1+z^2}dz = \int_0^{1-\varepsilon} \sum_{k=0}^{\infty}(-1)^k z^{2k}dz = \sum_{k=0}^{\infty}\frac{(-1)^k(1-\varepsilon)^{2k+1}}{2k+1}.$$

我们需要证明当 $\varepsilon \to 0$ 时，两边的极限可以取到，我们将这一步留作作业。Huygens 对 Leibniz 的这个公式大加赞许，认为这个公式 *will be celebrated among mathematicians forever*。然而，由于同一时代的英国数学家 James Gregory 研究的了 $\arctan x$ 的级数展开（你是否能证明）：

$$\arctan x = x - \frac{1}{3}x^3 + \frac{1}{5}x^5 - \cdots,$$

这个公式也为后面 Leibniz 与 Newton 在微积分发明权上的争执埋下了伏笔。

一些无法计算原函数的定积分的计算

命题 142 (微分与积分的交换定理，Riemann 积分的版本)**.** *假设二元函数*

$$f(x,t) : [a,b] \times [t_0 - c, t_0 + c] \to \mathbb{R}$$

是连续函数。对于每个 $x \in [a,b]$，我们假设 $\frac{df}{dt}(x,t)$ 存在并且

$$\frac{df}{dt} : [a,b] \times [t_0 - c, t_0 + c] \to \mathbb{R}$$

是连续函数，那么对于每个固定的 $x \in [a,b]$，映射

$$[t_0-c, t_0+c] \to \mathbb{R}, \quad t \mapsto \int_a^b f(x,t)dx$$

是关于变量 t 的可微函数并且有如下的等式：

$$\frac{d}{dt}\Big|_{t=t_0}\Big(\int_a^b f(x,t)dx\Big) = \int_0^1 \frac{df}{dt}(x,t_0)dx.$$

证明： 固定 $x \in [a,b]$，我们令 $F(t) = \int_a^b f(x,t)dx$。根据 Lagrange 中值定理，我们有

$$\frac{F(t+h)-F(t)}{h} = \int_a^b \frac{df}{dt}\big(x, t_0+\theta(x,h)h\big),$$

其中 $\theta(x,h) \in [0,1]$。从而，

$$\frac{F(t+h)-F(t)}{h} - \int_a^b \frac{df}{dt}(x,t_0)dx = \int_a^b \left(\frac{df}{dt}\big(x,t_0+\theta(x,h)h\big) - \frac{df}{dt}(x,t_0)\right).$$

按照要求，$\dfrac{df}{dt}(x,t)$ 是紧集合 $[a,b]\times[t_0-c,t_0+c]$ 上的连续函数，从而是一致连续的函数。按照定义，对任意的 ε，存在 $\delta>0$，只要 $|(x,t)-(x',t')|<\delta$，就有 $|f(x,t)-f(x',t')|<\varepsilon$。据此，当 $h<\delta$ 时，我们有

$$\Big|\frac{F(t+h)-F(t)}{h} - \int_a^b \frac{df}{dt}(x,t_0)dx\Big| < \int_a^b \varepsilon = (b-a)\varepsilon.$$

即 $\lim\limits_{h\to 0}\Big|\dfrac{F(t+h)-F(t)}{h} - \int_a^b \dfrac{df}{dt}(x,t_0)dx\Big| = 0$，这就证明命题。 □

在这个证明中，最核心的技术是利用如下的引理

引理 143. *假设 $f:(X,d)\to\mathbb{R}$ 是紧的距离空间上的连续函数，那么 f 是一致连续的。*

证明： 任选 $\varepsilon>0$，按照定义，对任意的 x，存在 $\delta_x>0$，使得对任意的 $y\in B(x,2\delta_x)$，我们都有 $|f(y)-f(x)|<\dfrac{1}{2}\varepsilon$。很明显，开球的集合 $\big\{B(x,\delta_x)\big|x\in X\big\}$ 是 X 的开覆盖，根据紧性，存在 $B(x_1,\delta_1)$，$B(x_2,\delta_2)$，$\cdots$，$B(x_n,\delta_n)$（其中 $\delta_k=\delta_{x_k}$），它们构成了 X 的一个有限开覆盖。令 $\delta=\min(\delta_1,\cdots,\delta_n)$。此时，任选 $y,y'\in X$，使得 $d(y,y')<\delta$。存在某个 i_0，使得 $y\in B(x_{i_0},\delta_{i_0})$，从而，$y'\in B(x_{i_0},2\delta_{i_0})$。据此，我们知道：

$$|f(y)-f(x)|<\frac{1}{2}\varepsilon,\ |f(y')-f(x)|<\frac{1}{2}\varepsilon \quad\Rightarrow\quad |f(y)-f(y')|<\varepsilon.$$

这就证明了 f 是一致连续的。 □

通过变动参数 t，我们可以用连续变换的思想来研究积分的值（依赖于 t）：

例子. （梗概）

1) Euler 关于 $n!$ 的积分公式:

$$\int_0^\infty x^n e^{-x} dx = n!.$$

我们自然可以通过分部积分将积分化为 $\int_0^\infty e^{-x} dx = 1$ 来进行计算。我们要注意到，一般而言，分部积分公式可能对反常积分并不成立。为此，我们需要用极限来表达反常积分才可以，我们在此略去细节。

另外一个有趣的方法是利用积分和导数可以交换的性质：通过换元 $x \mapsto tx$，我们有

$$\int_0^\infty e^{-tx} dx = \frac{1}{t}.$$

我们对 t 求 n 次导数，得到

$$\int_0^\infty x^n e^{-tx} dx = n! \frac{1}{t^n}.$$

令 $t = 1$ 即可。当然，我们并不知道上面的关于积分和导数可以交换的性质在此情况下是否能够运用，但是，不难看出，每次求导数之后，函数 $x^n e^{-tx}$ 是关于 $(x, t) \in [0, \infty) \times [\frac{1}{2}, \frac{3}{2}]$ 的一致连续的函数，所以命题仍然成立。

2) (Gauss 积分) $\int_{-\infty}^\infty e^{-x^2} dx = \sqrt{\pi}$ (应该熟记这个结论)。

我们定义

$$I = \int_0^\infty e^{-\frac{x^2}{2}} dx, \quad F(t) = \int_0^\infty \frac{e^{-\frac{t^2(1+x^2)}{2}}}{1+x^2} dx.$$

由于积分因子在 $[1, \infty)$ 上被 $\frac{1}{x^2}$ 控制，所以 $F(t)$ 收敛。我们对 t 求导数，它可以与积分号交换（和上面一个道理），从而

$$F'(t) = -te^{-\frac{t^2}{2}} \int_0^\infty e^{-\frac{t^2 x^2}{2}} dx = -Ie^{-\frac{t^2}{2}},$$

其中。我们注意到 $F(0) = \frac{\pi}{2}$。另外，我们有

$$F(\infty) = \lim_{t \to \infty} F(t) = 0.$$

这因为

$$\int_0^\infty \frac{e^{-\frac{t^2(1+x^2)}{2}}}{1+x^2} \leqslant \int_0^\infty \frac{2}{t^2(1+x^2)^2} = \frac{1}{t^2} \times \int_0^\infty \frac{dx}{(1+x^2)^2} = O(\frac{1}{t^2}).$$

我们有

$$\frac{\pi}{2} = \int_0^\infty F'(t) dt = (-I)^2.$$

据此，我们就算得了 *Gauss* 积分。

2) 我们研究积分 $\int_0^\infty e^{-tx} \frac{\sin x}{x} dx$，其中 $t \in [0, \infty)$。

我们注意到

- 当 $t>0$ 时，

为此，我们令 $F(t)=\int_0^\infty e^{-tx}\frac{\sin x}{x}dx$，那么（为什么），$F'(t)=-\int_0^\infty e^{-tx}\sin x dx$。我们可以计算不定积分

$$\int e^{ax}\sin x dx=\frac{a\sin x-\cos x}{1+a^2}e^{ax}.$$

所以，

$$F'(t)=\frac{t\sin x+\cos x}{1+t^2}e^{-tx}\Big|_{x=0}^{x=\infty}=-\frac{1}{1+t^2}.$$

从而，$F(t)=-\arctan t+C$，其中 C 为常数。为了确定 C，我们令 $t\to\infty$ 即可。特别地，令 $t=0$，我们得到

$$\int_0^\infty \frac{\sin x}{x}dx=\frac{\pi}{2}.$$

$\int_0^\infty e^{-tx}\frac{\sin x}{x}dx=\frac{\pi}{2}-\arctan t$。

23.1 作业：$\zeta(2)$ 的无理性

清华大学 19-20 秋季学期，数学分析一，作业 10

请用 A4 大小的纸张**正反面**用钢笔，签字笔或者圆珠笔书写，并注明自己的姓名，年级和作业的**总页数**。除定理公式所涉及的人名之外，请使用**中文**。本次作业的提交时间和地点为 **12 月 12 日**上午的课堂上，逾期视作零分。

基本习题

习题 A（课堂补充）

A1) 试构造连续函数 $f_n, f \in C\big([0,1]\big)$，使得对任意的 $x \in [0,1]$，当 $n \to \infty$ 时，$f_n(x) \to f(x)$，但是

$$\lim_{n\to\infty} \int_0^1 f_n \neq \int_0^1 f.$$

A2) $\alpha \in \mathbb{R}_{\geqslant 0}$。证明，反常积分 $\displaystyle\int_{100}^{\infty} \frac{dx}{x \log^{\alpha}(x)}$ 收敛当且仅当 $\alpha > 1$。

A3) （反常积分的控制判别法）f 和 F 在区间 I 上定义，即 $f : I \to \mathbb{R}$，$F : I \to \mathbb{R}$ 并且对任意的有界闭区间 $J \subset I$，f 和 F 均为 J 上的 Riemann 可积函数。假设对任意的 $x \in I$，我们都有

$$\big|f(x)\big| \leqslant F(x).$$

如果 F 在区间 I 上的反常积分收敛，那么 f 在区间 I 上的反常积分也收敛。

A4) 利用控制判别法，证明如下反常积分的收敛性：

$$(1)\ \int_0^{\infty} e^{-x^2} \qquad (2)\ \int_0^1 \frac{1}{\sqrt{1-x^3}} \qquad (3)\ \int_0^{\infty} \frac{(\log x)^2}{1+x(\log x)^5}.$$

A5) 利用控制判别法和面积法，证明如下级数的收敛性：

$$(1)\ \sum_{n=1}^{\infty} e^{-n}(n^2 + \log n) \qquad (2)\ \sum_{n=1}^{\infty} \frac{\log n}{1+n(\log n)^3}.$$

A6) 利用积分的定义计算：

$$\lim_{n\to\infty} \sum_{k=1}^{n} \frac{k^{\alpha}}{n^{\alpha+1}},\ \alpha > -1.$$

A7) （$\dfrac{22}{7}$）计算 $\displaystyle\int_0^1 \frac{x^4(1-x)^4}{1+x^2}$。你是否能借此证明 $\pi = 3.14...$，即给出 π 到小数点两位的逼近。

A8) （π 是无理数）假定 a, b 和 n 是正整数，定义函数

$$f_{a,b;n}(x) = \frac{x^n(a-bx)^n}{n!}.$$

- 证明：对于 $k=0,1,2,\cdots,2n$，函数 $f_{a,b;n}(x)$ 的 k 阶导数 $f_{a,b;n}^{(k)}(x)$ 在 $x=0$ 和 $x=\dfrac{a}{b}$ 处的取值都是整数。
- 假定圆周率 $\pi=\frac{a}{b}$ 是有理数，其中 a 和 b 是正整数。证明，对任意的正整数 n，积分

$$\int_0^{\pi} f_{a,b;n}(x)\sin(x)dx$$

是整数。
- 证明：圆周率 π 不是有理数。

A9) （完成 Wallis 积分的渐进公式）利用第 23 次讲义中的知识，对于 Wallis 积分 $I_n=\displaystyle\int_0^{\frac{\pi}{2}}\sin^n x dx$，证明，

$$I_n\sim\sqrt{\frac{\pi}{2n}}.$$

A10) 证明当年 Leibniz 用的面积公式：假设 $f:[0,1]\to[0,1]$ 是双射，$g=f^{-1}:[0,1]\to[0,1]$ 是它的逆映射，并且 f 和 g 是连续可微的。那么，

$$\int_0^1 f(x)dx+\int_0^1 g(y)dy=1.$$

A11) 证明讲义中 Leibniz 级数的计算中所用的极限：

$$\lim_{\varepsilon\to 0}\sum_{k=0}^{\infty}\frac{(-1)^k(1-\varepsilon)^{2k+1}}{2k+1}=\sum_{k=0}^{\infty}\frac{(-1)^k}{2k+1}.$$

A12) 对任意的连续函数 $f:[a,b]\times[c,d]\to\mathbb{R},(x,y)\to f(x,y)$，其中 $a,b,c,d\in\mathbb{R}$，证明，f 在 $[a,b]\times[c,d]$ 上一致连续。（其参考讲义的一般情形）

习题 B（关于 $\zeta(2)$）

第一部分：数列 $\left\{\displaystyle\sum_{k=1}^{n}\frac{1}{k^p}\right\}_{\geqslant 1}$

我们定义序列 $S_n(p)=\displaystyle\sum_{k=1}^{n}\frac{1}{k^p}$，其中 $p\in\mathbb{Z}_{\geqslant 1}$。

B1) 证明，对任意的 $k\in\mathbb{Z}_{\geqslant 1}$，我们有

$$\frac{1}{(k+1)^p}\leqslant\int_k^{k+1}\frac{1}{x^p}dx\leqslant\frac{1}{k^p}.$$

B2) 证明，对任意的 $n\in\mathbb{Z}_{\geqslant 2}$，我们有

$$S_n(p)-1\leqslant\int_1^n\frac{1}{x^p}dx\leqslant S_{n-1}(p).$$

B3) 我们假设 $p \in \mathbb{Z}_{\geqslant 1}$。证明，函数 $x \mapsto \dfrac{1}{x^p}$ 在 $[1, +\infty)$ 上反常可积当且仅当 $p \geqslant 2$。

B4) 证明，数列 $\{S_n(p)\}_{n\geqslant 1}$ 收敛当且仅当 $p \geqslant 2$。对于 $p \geqslant 2$，我们下面记

$$\zeta(p) = \lim_{n\to\infty} S_n(p) = \sum_{k=1}^{\infty} \frac{1}{k^p}.$$

第二部分：计算 $\zeta(2)$

定义函数 $h(t) = \dfrac{t^2}{2\pi} - t$，定义 $[0, \pi]$ 上的函数 $\varphi : [0, \pi] \to \mathbb{R}$:

$$\varphi(x) = \begin{cases} -1, & x = 0; \\ \dfrac{h(x)}{2\sin(\frac{x}{2})}, & 0 < x \leqslant \pi. \end{cases}$$

B5) 证明，$\varphi \in C^1([0, \pi])$。

B6) 对所有的整数 $k \geqslant 1$，计算如下积分的值：

$$\int_0^{\pi} h(x)\cos(kx)dx.$$

B7) 证明，存在常数 λ（给出它的值），使得对任意的 $x \in (0, \pi]$，我们都有

$$\sum_{k=1}^{n} \cos(kx) = \frac{\sin\big((n + \frac{1}{2})x\big)}{2\sin(\frac{x}{2})} - \lambda.$$

B8) 证明，对于任意的函数 $\psi(x) \in C^1([0, \pi])$，我们都有如下的极限：

$$\lim_{n\to+\infty} \int_0^{\pi} \psi(x)\sin\big((n + \frac{1}{2})x\big)dx = 0.$$

B9) 证明如下的等式：

$$\zeta(2) = \frac{1}{6}\pi^2.$$

（提示：利用分部积分）

第三部分：$\zeta(2)$ 是无理数

我们采取反证法。在这一部分中，我们假设

$$\pi^2 = \frac{a}{b}$$

其中 a，b 是正整数。我们要推出矛盾。

B10) 定义多项式的序列 $f_n(x)=\dfrac{x^n(1-x)^n}{n!}$，其中 $n\in\mathbb{Z}_{\geqslant 1}$。证明，对任意的整数 k，上述函数的高阶导数 $f_n^{(k)}(0)$ 和 $f_n^{(k)}(1)$ 都是整数。

B11) 定义函数的序列

$$F_n(x)=b^n(\pi^{2n}f_n(x)-\pi^{2n-2}f_n^{(2)}(x)+\pi^{2n-4}f_n^{(4)}(x)-\cdots+(-1)^n f_n^{(2n)}(x)).$$

证明，$F_n(0)$ 和 $F_n(1)$ 都是整数。

B12) 对于 $n\geqslant 1$，定义函数序列 $\{g_n\}_{n\geqslant 1}$ 和序列 $\{A_n\}_{n\geqslant 1}$ 如下：

$$g_n(x)=F_n'(x)\sin(\pi x)-\pi F_n(x)\cos(\pi x),\quad A_n=\pi\int_0^1 a^n f_n(x)\sin(\pi x)dx.$$

证明等式：

$$g_n'(x)=\pi^2 a^n f_n(x)\sin(\pi x)$$

以及 A_n 是正整数。

B13) 证明，存在正整数 n，使得对任意的 $x\in[0,1]$，我们都有

$$a^n f_n(x)<\frac{1}{2}.$$

B14) 证明，存在正整数 n，使得 $A_n\in(0,1)$。这和 B12) 矛盾。

习题 C(定积分的计算) 试计算下列定积分（常数 $a\neq 0$，$b\neq 0$。）

(1) $\displaystyle\int_0^{\pi}\sin^3 x$　(2) $\displaystyle\int_{-\pi}^{\pi}x^2\cos x$　(3) $\displaystyle\int_0^3\frac{x}{1+\sqrt{1+x}}$

(4) $\displaystyle\int_0^{\sqrt{3}}x\arctan(x)$　(5) $\displaystyle\int_{-1}^0(2x+1)\sqrt{1-x-x^2}$　(6) $\displaystyle\int_{\frac{1}{e}}^e|\log x|$

(7) $\displaystyle\int_0^a x^2\sqrt{a^2-x^2}$　(8) $\displaystyle\int_0^{\log 2}\sqrt{e^x-1}$　(9) $\displaystyle\int_1^2 x^{100}\log x$

(10) $\displaystyle\int_0^a\log(x+\sqrt{x^2+a^2})$　(11) $\displaystyle\int_0^{\frac{\pi}{2}}\frac{\cos x\sin x}{a^2\sin^2 x+b^2\cos^2 x}$　(12) $\displaystyle\int_0^{\frac{\pi}{2}}\frac{\cos^2 x}{\cos x+\sin x}$

(13) $\displaystyle\int_0^{\frac{\pi}{2}}\frac{\sin^2 x}{\cos x+\sin x}$　(14) $\displaystyle\int_0^{\frac{\pi}{4}}\log(1+\tan x)$　(15) $\displaystyle\int_0^4\frac{|x-1|}{|x-2|+|x-3|}$

(16) $\displaystyle\int_0^{\pi}\frac{x\sin x}{1+\cos^2 x}$　(17) $\displaystyle\int_0^{\frac{\pi}{4}}\frac{\sin x}{\sin x+\cos x}$　(18) $\displaystyle\int_0^{\frac{\pi}{2}}\frac{\sin 2019x}{\sin x}$

(19) $\displaystyle\int_2^4\frac{\log\sqrt{9-x}}{\log\sqrt{9-x}+\log\sqrt{x+3}}$　(20) $\displaystyle\int_0^1\frac{1}{\sqrt{1+x^2}+\sqrt{1-x^2}}$　(21) $\displaystyle\int_0^1\sqrt{x+\sqrt{1+x}}$

(22) $\displaystyle\int_{-1}^1\frac{\sin\Big(\sin\big(\sin(x)\big)\Big)}{x^{800}+1}$

24 常微分方程解的存在唯一性，Kepler 三大定律的证明，变分法：两点之间线段最短

二零一九年十二月九日，星期一，雾霾

回忆一下，在第七次课中，我们证明如下压缩映像定理：

(X,d) 是完备的距离空间。假设映射 $T:X\to X$ 是压缩映射，即存在常数 $0<\gamma<1$，使得对任意的 $x,x'\in X$，我们都有

$$d(T(x),T(x'))\leqslant\gamma d(x,x').$$

那么，T 必有唯一的不动点，即存在唯一的 $x_*\in X$，使得 $T(x_*)=x_*$。

不动点定理与微积分基本定理结合，可以证明常微分方程解的存在唯一性定理：

定理 144 (Cauchy-Lipschitz)**.** 给定完备的赋范线性空间 V（通常我们假设 $V=\mathbb{R}^n$），$\Omega\subset V$ 是开集，$I\subset\mathbb{R}$ 是开区间，$f:\Omega\times I\to V$ 是给定的连续函数（映射）。假设存在 $C>0$，对任意的 $t\in I$，映射

$$\Omega\to V,x\mapsto f(x,t)$$

是 C-*Lipschitz* 映射，即对任意的 $x,y\in\Omega$，有如下估计

$$|f(x,t)-f(y,t)|\leqslant C|x-y|.$$

那么，对任意的（初始值）$(x_0,t_0)\in\Omega\times I$，存在 $\delta>0$（可能依赖于 (x_0,t_0)），存在唯一的映射 $x(t):(t_0-\delta,t_0+\delta)\to\Omega$，满足如下的常微分方程：

$$\begin{cases}x'(t)=f(x(t),t),\\ x(t_0)=x_0.\end{cases}$$

假设我们想解常微分方程（以 $V=\mathbb{R}$ 为例子）

$$x'(t)=f(x(t),t).$$

这个定理讲的是对任意给定的初值 x_0，即要求 $x(t_0)=x_0$，总能存在比较小的的一个时间窗口 $(t_0-\delta,t_0+\delta)$，使得我们能找到唯一一个函数

$$x:(t_0-\delta,t_0+\delta)\to\mathbb{R}$$

满足方程和初值条件。换句话说，这样的常微分方程系统局部上总是有解的，当然，我们必须要求 $f(x,t)$ 对 x 有比较好的光滑性假设（Lipschitz）。

注记. 定理 *Lipschitz* 的假设是重要的：

1) 如果只假设 f 对于 x 变量是连续的，即不进一步要求 *Lipschitz* 条件 $|f(x,t)-f(y,t)|\leqslant C|x-y|$ 成立，那么定理的结论并不成立。实际上，我们仍然可以证明上面的常微分方程系统局部上有解（*Peano* 定理），但是解可能并不唯一。

2) 在应用这个定理的时候，$f(t,x)$ 通常都是非常光滑的函数，比如说对任意的 $t\in I$，映射 $x\mapsto f(t,x)$ 是 C^1 的（$x\in\mathbb{R}$）。我们进一步假设（这个条件在应用的时候一般也都成立）函数 $(t,x)\mapsto \dfrac{df}{dx}(t,x)$ 是连续映射。所以，通过适当的缩小 Ω 和 I，就存在 $C>0$，使得

$$\sup_{(t,x)\in I\times\Omega}\Big|\frac{df}{dx}(t,x)\Big|\leqslant C.$$

根据 *Lagrange* 中值定理，我们有

$$|f(x,t)-f(y,t)|=\Big|\frac{df}{dx}(t,x+\theta(y-x))(x-y)\Big|\leqslant C|x-y|.$$

此时，定理的条件是成立的。

证明*:* 我们首先利用 Newton-Leibniz 公式将问题转化为一个积分方程的问题。假设 $x'(t)$ 是方程

$$\begin{cases} x'(t)=f(x(t),t),\\ x(t_0)=x_0. \end{cases}$$

的解，这个解自然是连续的。根据 Newton-Leibniz 公式，我们必然有

$$\boxed{x(t)}=x_0+\int_{t_0}^{t}f(\boxed{x(\tau)},\tau)d\tau,\quad t\in I.$$

如果我们将 $x(t)$ 视为某个空间中的点，那么上面的积分等式可以被理解为一个不动点的表达式。我们可以精确地描述这个问题：令 $X=\Big(C\big([t_0-\delta,t_0+\delta]\big),\|\cdot\|_\infty\Big)$，这是一个完备的距离空间，其中 δ 待定。我们定义 X 到自身的映射：

$$T:C\big([t_0-\delta,t_0+\delta]\big)\to C\big([t_0-\delta,t_0+\delta]\big),\quad x(t)\mapsto x_0+\int_{t_0}^{t}f(\boxed{x(\tau)},\tau)d\tau.$$

所以，积分形式的问题可以表述为

$$Tx=x.$$

为了证明不动点的存在性，我们需要验证压缩映像的条件：对任意的 $x(t),y(t)\in C\big([t_0-\delta,t_0+\delta]\big)$，对给定的 $t\in[t_0-\delta,t_0+\delta]$，我们有

$$\begin{aligned}|T\big(x(t)\big)-T\big(y(t)\big)|&\leqslant\int_{t_0}^{t}|f(x(\tau),\tau)-f(y(\tau),\tau)|d\tau\\ &\leqslant\int_{t_0}^{t}C|x(\tau)-y(\tau)|d\tau\leqslant C\sup_{x\in[t_0-\delta,t_0+\delta]}\int_{t_0}^{t}d\tau\\ &\leqslant C\delta d_\infty\big(x(t),y(t)\big).\end{aligned}$$

由于这对任意的 $t \in [t_0-\delta, t_0+\delta]$ 都成立，所以

$$d_\infty\Big(T\big(x(t)\big), T\big(y(t)\big)\Big) \leqslant C\delta d_\infty\big(x(t), y(t)\big).$$

只要 $\delta C \leqslant \frac{1}{2}$，我们就有 $d_\infty\Big(T\big(x(t)\big), T\big(y(t)\big)\Big) \leqslant \dfrac{1}{2} d_\infty\big(x(t), y(t)\big)$，根据不动点定理，存在唯一的 $x(t) \in X$，满足

$$x(t) = x(t_0) + \int_{t_0}^{t} f(x(\tau), \tau) d\tau.$$

很明显，$x(t_0) = x_0$ 由于 $x(t) \in C\big([t_0-\delta, t_0+\delta]\big)$，所以右边的积分项定义的函数是连续可微的，从而 $x(t)$ 连续可微。通过求微分，我们得到

$$x'(t) = f(x(t), t).$$

命题得到了证明。 □

在应用的时候，我们经常会遇到二阶的微分方程

$$\begin{cases} x''(t) = f(x'(t), x(t), t), \\ \big(x(t_0), x'(t_0)\big) = (x_0, t_0) \end{cases}$$

这样的系统可以约化为一阶的系统：定义 $X(t) = \binom{x(t)}{x'(t)}$，那么，上述方程就变成了

$$\begin{cases} X'(t) = f(X(t), t), \\ X(t_0) = \binom{x_0}{v_0}. \end{cases}$$

从而，仍然局部有解。当然，为了解这个方程，我们需要一开始给定两个初始值。比如说，考虑方程

$$f'' + f = 0.$$

如果初始值给的是 $f(0) = 0$，$f'(0) = 1$，那么就唯一的锁定了 $\sin x$ 作为上述方程的解；如果初始值给的是 $f(0) = 1$，$f'(0) = 0$，那么就得到 $\cos x$ 作为上述方程的解。上面从二阶方程到一阶方程的转变过程也体现了向量值函数的微积分理论的威力。

练习. 你是否能写出下面方程的解：

$$\begin{cases} f'' + f = 0, \\ (f(0), f'(0)) = (c_0, c_1). \end{cases}$$

行星运动的 Kepler 定律：Newton 的数学推导

我们中学学习过经典力学中 Newton 的三个运动定律：

1) 任何物体都保持静止或匀速直线运动的状态，直到受到其它物体的作用力迫使它改变这种状态为止。

2) 物体在受到合外力的作用会产生加速度，加速度的方向和合外力的方向相同，加速度的大小正比于合外力的大小与物体的惯性质量成反比。

3) 两个物体之间的作用力和反作用力，在同一条直线上，大小相等，方向相反。

我们再回忆 Newton 的万有引力定律：

给定两个质点，它们之间引力的大小正比于每个质点的质量，反比于质量的平方：

$$F = \frac{Gm_1m_2}{r^2}, \quad G = 6.67 \times 10^{-11} \frac{\mathrm{N} \cdot \mathrm{m}^2}{\mathrm{kg}^2},$$

其中 m_1 和 m_2 是两个质点的质量，r 是他们之间的距离。其中，根据标准的单位制，$\mathrm{N} = \frac{\mathrm{kg} \cdot \mathrm{m}}{\mathrm{s}^2}$。

对于太阳系中行星的运动，我们可以做如下理想的数学假设：

1) 把太阳和行星都想象成质点，这因为星球本身的半径和运动轨道的尺度相比较可以忽略；

2) 假设行星之间的并没有相互引力，这因为太阳的质量大概占了太阳系质量的 99.8% 以上，其它行星的质量贡献太小可以忽略不计。

行星运动的空间我们用 $\mathbb{R}^3$ 来描述，通过选取坐标系，可以假设太阳（质量 $= M$）在原点 $(0,0,0) \in \mathbb{R}^3$ 处。假设在时刻 t 时，行星（质量 $= m$）所处的位置是 $x(t) \in \mathbb{R}^3$，所谓行星运动的刻画就是描述映射 $t \mapsto x(t)$，比如说，知道 2019 年 12 月 9 日早上 9:50 数学分析课开始的时候地球在太阳系的位置 $x(2019)$ 和速度 $x'(2019)$，我们要知道 1750000000 年后地球的位置（据说地球上的生物还有机会生活至少 17.5 亿年）。

根据 Newton 第二定律（速度的导数是加速度），我们有

$$F = mx''(t) = -\frac{GMm}{r^2}\frac{x(t)}{r},$$

其中 $r(t) = |x(t)|$ 是行星到太阳的距离，$-\dfrac{x(t)}{r}$ 是它受力的方向。所以说，行星的运动轨迹 $x(t)$ 满足如下的常微分方程：

$$x''(t) = -GM\frac{x(t)}{r^3}.$$

我们注意到上面的方程和行星本身的质量 m 没有关系，只和太阳的质量 M 有关！

这个问题本身很明显是不依赖于坐标系的选择的，我们为了方便才引入的坐标系。所以，通过适当的选取坐标系统，我们假设在开始的时刻 $x(0) = x_0$ 和 $x'(0) = v_0$ 都落在坐标平面 $z = 0$ 上，即 x_0 和 v_0 的第三个坐标是零。

我们来观察运动方程 $x''(t) = -GM\dfrac{x(t)}{r^3}$ 中第三个坐标 $z(t)$ 满足的方程：

$$\begin{cases} z''(t) = -\frac{GM}{r^3}z(t), \\ (z(0), z'(0)) = (0,0). \end{cases}$$

我们可以将 r 看做是已经给定的函数（如果有解的话）并且假设 $r \neq 0$（如果 $r = 0$ 的话行星就撞上太阳了!）。那么，$z(t) \equiv 0$ 是上述方程的一个解。根据常微分方程的解的存在唯一性，我们知道

$$z(t) \equiv 0.$$

。所以，如果初始时刻行星的位置和速度处于一个平面，那么自始至终，行星都在这个平面上运动。根据这些推到，我们可以假设行星只做平面运动，即 $x(t) \in \mathbb{R}^2$。

注记. 上面的论证可以在物理上用角动量守恒来证明，我们走了捷径。下面的推导可以通过能量守恒来考虑（在常微分方程的理论中，这被称作是首次积分），为了篇幅和时间的考虑，我们不在追究背后的这些深层次的原因。

用平面上的极坐标系改写 Newton 的运动方程（$x(t) = r(t)e^{i\theta(t)}$ 用复数的形势表达），那么 $r(t)$ 和 $\theta(t)$ 也是可微的（为什么）。此时，我们有

$$x''(t) = -\Big[r''(t) - r(t)|\theta'(t)|^2 + i\big(2r'(t)\theta'(t) + r(t)\theta''(t)\big)\Big]e^{i\theta(t)}.$$

代入方程，比较系数就得到

$$\begin{cases} 2r'(t)\theta'(t) + r(t)\theta''(t) = 0, \\ r''(t) - r(t)|\theta'(t)|^2 = -\frac{GM}{r(t)^2}. \end{cases}$$

第一个方程等价于 $\big(r^2\theta'(t)\big)' = 0$，所以，存在常数 c，使得

$$r(t)^2\theta(t)' = c.$$

如果一开始 $c = 0$，那么只能是 $\theta(t)' = 0$，即 $\theta(t)$ 恒为常数，所以行星做直线运动，这样它或者远离太阳或者冲向太阳，我们不关心这种情况。

我们假设 $c > 0$（通过对换 x 轴和 y 轴，我们总可以假设行星逆时针运动）。特别地，我们 θ 局部上是 t 的增函数，所以局部上我们可以考虑 $t \mapsto \theta(t)$ 的逆映射，也就是说我们认为 t（至少在局部上）是 θ 的函数。再定义函数 $s(t) = \dfrac{1}{r(t)}$。转换一下视角：我们现在用 θ 作为变量，来描述 $s(t) = s(\theta)$ 的运动。

按照定义，第一个方程可以写成

$$\theta'(t) = c \cdot s(t)^2.$$

根据链式法则，我们有

$$r'(t) = -\frac{s'(t)}{s^2(t)} = -\frac{c}{s'(t)}\theta'(t) = -c\frac{ds}{d\theta}(t).$$

所以，我们有

$$r''(t) = \big(-c\frac{ds}{d\theta}(\theta(t))\big)' = -c^2 s(t)^2\frac{d^2s}{d^2\theta}(\theta(t)) \quad \Leftrightarrow \quad r''(t) = -c^2 s^2\frac{d^2s}{d^2\theta}.$$

代入第二个方程，我们得到

$$-c^2s^2\ddot{s} - c^2s^3 = -GMs^2,$$

其中 $\dot{s}$ 代表对 θ 求导数。经过整理，我们得到

$$\ddot{s} + s = \frac{GM}{c^2}.$$

这个方程的解不容易一下子猜出来。但是我们观察到，如果右边没有常数 $\dfrac{GM}{c^2}$，解这个方程就是在解关于 sin 和 cos 的方程。从而，我们知道方程的解一定形如

$$s(\theta) = A\cos\theta + B\sin\theta + \frac{GM}{c^2},$$

这里 A 和 B 都是待定的常数。

我们选取行星近地点（即离着太阳最近的时候）作为初始时刻：数学上来说，就是 $\theta = 0$ 时，s 最大，所以

$$s'(0) = 0, \quad s''(0) \leqslant 0$$

带入 $s(\theta)$ 的待定表达式，我们得到

$$s(\theta) = B\cos\theta + \frac{GM}{c^2}, \quad B \geqslant 0.$$

从而，我们得到

$$r = \frac{\frac{c^2}{GM}}{1 + \frac{Bc^2}{GM}\cos\theta}.$$

这是圆锥曲线方程在极坐标下的表达。令 $e = \dfrac{Bc^2}{GM}$，$\ell = \dfrac{c^2}{GM}$，方程变化为

$$r = \frac{\ell}{1 + e\cos\theta},$$

其中 $e \geqslant 0$ 是圆锥曲线的离心率。按照离心率的大小，我们有如下的分类

- $e = 0$。此时，行星的运动轨迹恰好是一个圆，半径为 $a = \ell$。
- $0 < e < 1$。此时，行星的运动轨迹恰好是一个椭圆。我们回忆一下，椭圆的半短轴 b 和半长轴 a 的比例为 $\sqrt{1-e^2} = \dfrac{b}{a}$，那么 $\ell = a(1-e^2) = \dfrac{b^2}{a}$。
- $e = 1$。此时，行星的运动轨迹恰好是一个抛物线。
- $e > 1$。此时，行星的运动轨迹恰好是一个双曲线的一支。

我们回忆一下，Kepler 通过研究前人的观测数据，发现和总结了关于行星运动的三个定律：

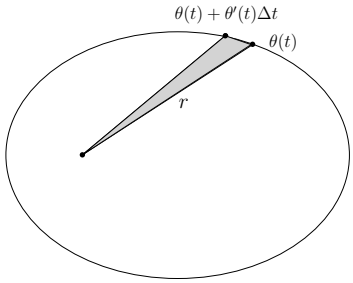
(1) 每个行星都沿各自的椭圆轨道环绕太阳运行，太阳处在椭圆的一个焦点上。

(2) 在相等时间内，太阳和运动着的行星的连线所扫过的面积都是相等的。

(3) 行星绕太阳公转周期的平方和它们的椭圆轨道的半长轴的立方成正比。

至此，我们已经严格从数学上论证了 Kepler 的第一定律!

我们现在证明 Kepler 的第二定律，这等价于说面积的变化率是常数。当然，迄今为止，我们并未定义面积（严格的定义和二重积分密切相关）。我们给出不严格的论证：在上面的计算中，我们已经证明 $r^2(t)\theta'(t) = c$，这就是面积的变化率，因为当 t 从 t 变为 $t+\Delta t$ 时，θ 变为 $\theta+\theta'(t)\Delta t$，而面积的变化大约为 $\dfrac{1}{2}r^2\theta'(t)\Delta t$，请参见下图的灰色三角形：



Kepler 第三定律给出了行星运动的周期，它实际上是第二定律的推论。我们首先计算椭圆的面积：

练习. 给定 $\mathbb{R}^2$ 上由方程 $\dfrac{x^2}{a^2}+\dfrac{y^2}{b^2}=1$ 定义的椭圆，它的面积为 πab。实际上，我们可以用积分来定义这个叙述：

$$4\int_0^a \sqrt{b^2(1-\frac{x^2}{a^2})}dx = \pi ab.$$

于是，周期 $T = \dfrac{\pi ab}{c}$（因为 c 是面积的变化率）。通过之前对各个参数的计算，我们立即得到

$$T^2 = \frac{4\pi^2 a^3}{GM}.$$

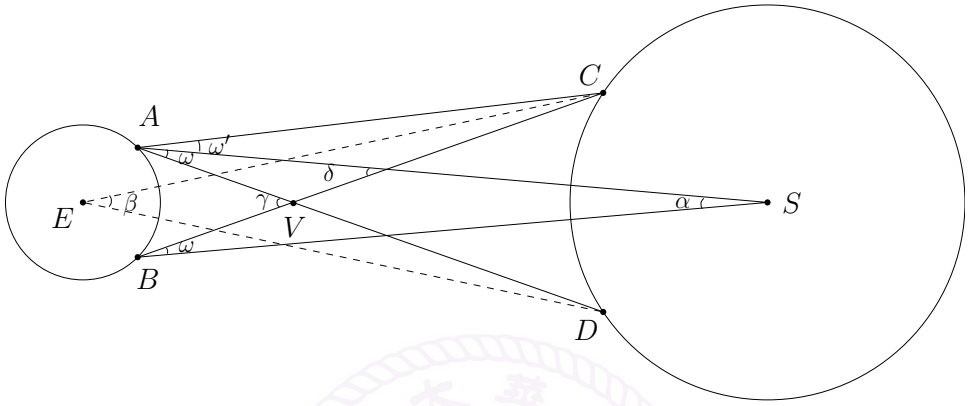
这就是第三定律。我们特别强调，上面的表达式中 M 是太阳的质量。

跑题：太阳质量的测量

根据 Kepler 第三定律的公式，太阳的质量可以用 $M = \dfrac{4\pi^2 a^3}{GT^2}$ 来计算。在这个公式中，π 和 G 是已知的参数。由于我们绕太阳转一周就是一年，所以 T 也已经知道。参数 a（几乎）是地球到太阳的距离，所以如果能测量日地距离（需要一把稍长的尺子），我们就能称太阳的重量！关于日地距离的计算，我们用金星凌日法。所谓的金星凌日就是太阳，地球和金星这三个点在一条直线上的天文现象（我国古人称之为太白犯主：太白主兵，与日合，大不详），历史上有记载的玄武门之变之前有金星凌日！1677 年，Halley 预测了 1761 年的会发生金星凌日，他认为通过在地球上若干个地点的观测再加上对金星运动周期的观测数据，就可能算出太阳的距离。他晚年的时候又提出了利用金星凌日计算日地距离的精确方法。然而，老英雄 Halley 壮志未酬，因为 1656 年生的他注定等不到这一天的来临。1769 年（1761 年那次金星凌日由于技术原因没有成功），通过欧洲的

科学家的观测以及与在 Tahiti 岛的英国航海家 James Cook 船长的观测（英法正处战争，法国政府特别下令要保护 Cook 的船以保证测量的进行），之后不久法国人 Lalande 据此算出了地球与太阳间的距离大约为 1.5 亿公里（$\approx 1.5\times 10^{11}$m）。据此，我们可以算出太阳的质量是 2×10^{30}kg。

我们现在简单描述金星凌日的基本想法。在下图中，我们用 E 表示地球的球心所在的位置，用 S 表示太阳的中心的位置，V 表示金星的位置。相比于地球到太阳的距离 $d_E\approx ES$ 和金星到太阳的距离 $d_V\approx VS$，这三个星球的半径都可以忽略。



在金星凌日的时刻，我们假设在地球上的上半球 A 点有人观测太阳，他将讲看到金星在太阳上的影子在 D 处；同样地，我们不妨假设在下半球对称的位置 B 处有人观测，她将看到金星的影子投在了太阳上的 C 处。根据照片上 C 和 D 的位置以及相机焦距的性质，我们可以算出角度 β，也就是说 β 是通过测量可以知道的。

由于地球和太阳相距特别远，所以我们可以假设 $\omega+\omega'=\angle CAD\approx\angle CED=\beta$。类似地，我们有 $\omega\approx\omega'$。根据三角形外角的性质，我们有

$$\gamma=\omega+\omega+\alpha\approx\omega+\omega'+\alpha=\beta+\alpha.$$

从而，$\gamma\approx\alpha+\beta$。根据正弦定理，我们有

$$\frac{\frac{1}{2}AB}{AV}=\sin\big(\frac{1}{2}\gamma\big)\approx\frac{1}{2}\gamma,\quad \frac{\frac{1}{2}AB}{AS}=\sin\big(\frac{1}{2}\alpha\big)\approx\alpha.$$

上面我们用到了当 x 足够小的时候，$\sin x\approx x$。从而，我们得到如下的近似等式

$$\frac{\gamma}{\alpha}\approx\frac{AS}{AV}\approx\frac{ES}{EV}=\frac{d_E}{d_E-d_V}=\frac{1}{1-\frac{d_V}{d_E}}.$$

又因为 $\gamma=\alpha+\beta$，所以

$$\frac{\alpha+\beta}{\alpha}\approx\frac{1}{1-\frac{d_V}{d_E}}\ \Rightarrow\ \alpha\approx\big(\frac{d_E}{d_V}-1\big)\beta.$$

另外，我们可以通过天文观测来得到地球和金星的运行周期（地球恰好是 1 年!），根据 Kepler 的第三个定律，我们就可以计算 $\dfrac{d_E}{d_V}$（这因为这两个行星的轨道差不多就是圆）。据此，我们可以计算角度 α。我们可以在地球上计算 AB 之间的距离（测量!），根据

$$\frac{\frac{1}{2}AB}{AC}=\sin\big(\frac{1}{2}\alpha\big)\approx\frac{1}{2}\alpha\quad AC\approx d_E,$$

我们得到计算公式

$$d_E \approx \frac{AB}{(\frac{d_E}{d_V}-1)\beta}.$$

据此，我们就可以计算地球到太阳的距离了。

空间曲线的长度

我们引入（空间）曲线和曲线的长度的概念。假设 $I=[a,b]\subset$ 是有界区间，考虑 C^1 的映射

$$\gamma: I \to \mathbb{R}^n,$$

如果 γ 是 C^1 的，我们就将映射 γ 称作是一条 C^1 **的（空间）曲线**。有些书籍的作者会要求 $\gamma'(x)\neq 0$，这和子流形的概念相关，我们会在下个学期进行系统的学习。习惯上，我们倾向于把映射的像 $\gamma([a,b])$ 看做是曲线。在做计算的时候，我们则更关心如何通过一个映射 γ 用一个区间 $I=[a,b]$ 将这个曲线参数化的：假设我们有连续可微的可逆的递增映射 $\varphi: J=[a',b']\to[a,b]$，那么我们可以考虑曲线的另一个参数化 $\gamma\circ\varphi: [a',b']\to\mathbb{R}^n$：

$$\begin{array}{ccc} J & \xrightarrow{\varphi} & I \\ & \searrow^{\gamma\circ\varphi} & \downarrow{\scriptstyle \gamma} \\ & & \mathbb{R}^n \end{array}$$

很明显，这两个参数化给出的曲线的像是一致的，但是我们认为这是两条不同的曲线。

任意给定曲线 $\gamma:[a,b]\to\mathbb{R}^n$，我们定义它的**长度**为（= 速度的积分）：

$$\ell(\gamma)=\int_a^b |\gamma'(t)|dt.$$

我们强调，对于 $\gamma'(t)=(\gamma_1'(t),\cdots,\gamma_n'(t))$，其长度 $|\gamma'(t)|$ 指的是

$$|\gamma'(t)|=\sqrt{\sum_{j=1}^n |\gamma_i'(t)|^2}.$$

如果换了参数化，根据换元积分公式，我们可以做如下的计算（按照分量来计算!）：

$$\begin{aligned}\ell(\gamma\circ\varphi)&=\int_{a'}^{b'}|\varphi(t)\cdot\gamma'(\varphi(t))|dt=\int_{a'}^{b'}|\gamma'(\varphi(t))|\varphi'(t)dt\\ &\overset{\tau=\gamma(t)}{=}\int_a^b|\gamma'(\tau)|d\tau=\ell(\gamma).\end{aligned}$$

这表明一条曲线的长度的定义不依赖于参数化的选取。

注记. 如果我们仅仅和要求 γ 是连续的而不是连续可微的，我们就可以造出很奇怪的曲线，比如说我们已经证明了存在可以填充 $[0,1]\to[0,1]\times[0,1]$ 的曲线。

如果只能学习两条曲线的话，那么我们没有选择：

例子. 我们研究直线和圆：

1) **直线段**。假设 $P=(x_1,\cdots,x_n)$ 和 $Q=(y_1,\cdots,y_n)$ 是 $\mathbb{R}^n$ 中的两点，它们之间的线段用 $\overline{PQ}$ 表示。我们用 $[0,1]$ 区间来参数化它：

$$\gamma:[0,1]\to\mathbb{R}^n,\quad t\mapsto P+t(Q-P)=\big(\cdots,x_j+t(y_j-x_j),\cdots\big).$$

据此，我们有

$$\begin{aligned}\ell(\overline{PQ})&=\int_0^1\sqrt{\sum_{j=1}^n|y_j-x_j|^2}dt\\&=\sqrt{(x_1-y_1)^2+\cdots+(x_n-y_n)^2}\end{aligned}$$

2) **圆周**。我们把平面 $\mathbb{R}^2$ 上到原点的距离为 R 的点所组成的集合 $\big\{(x,y)\big|x^2+y^2=R^2\big\}$ 称作是半径 R 的圆周。根据对称性，我们只考虑 $x\geqslant 0$ 和 $y\geqslant 0$ 的部分（在第一象限的圆周），也就是四分之一个圆周，它可以被下面的映射参数化：

$$[0,R]\to\mathbb{R}^2,x\mapsto(x,\sqrt{R^2-x^2}).$$

我们还可以令 $x=R\sin\theta$，其中 $x\in[0,\frac{\overline{\pi}}{2}]$（我们这里利用了 $\sin:[0,\frac{\overline{\pi}}{2}]\to[0,1]$ 是连续可微的单调的可逆映射），那么就有四分之一个圆周的另一个参数化

$$[0,\frac{\overline{\pi}}{2}]\to\mathbb{R}^2,\theta\mapsto(R\sin(\theta),R\cos(\theta)).$$

我们只利用解析意义下定义的 sin 和 cos 的性质（这个参数化实际上就已经说明了经典意义下所定义的三角函数和我们所定义的是一致的）。根据第二个参数化，我们有

$$\ell=\int_0^{\frac{\overline{\pi}}{2}}\big|(R\cos(\theta),-R\sin(\theta))\big|d\theta=\frac{1}{2}\overline{\pi}.$$

所以圆的周长公式为 $4\times\frac{1}{2}\overline{\pi}=2\overline{\pi}$，这就说明我们定义的 $\overline{\pi}$ 和经典意义下的 π 是一致的。另外，如果利用第一个参数化，我们就有

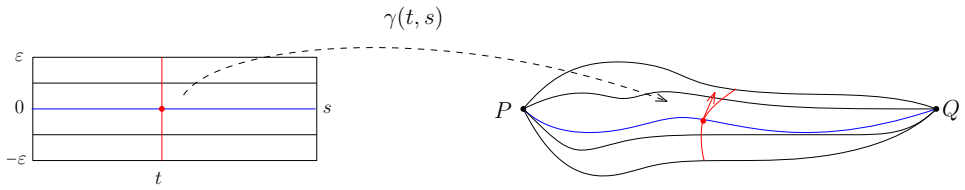
$$\ell=\int_0^R\sqrt{\frac{R^2}{R^2-x^2}}dx.$$

关于这个积分的计算实际上还是要做变量替换 $x=R\cos(\theta)$ 和 $y=R\sin(\theta)$。

我们中学知道两点之间线段最短，然而，只有有了微积分，才可以正确的叙述并证明这个结论：给定 $P,Q\in\mathbb{R}^n$，我们考虑连接 P 和 Q 的一族曲线：

$$\gamma(t,s):[0,1]\times[-\varepsilon,\varepsilon]\to\mathbb{R}^n,$$

其中对于每个固定的 $s\in[-\varepsilon,\varepsilon]$，$t\mapsto\gamma(t,s)$ 都是曲线。我们还要求所有曲线的起点和终点都分别是 P 和 Q，即 $\gamma(0,s)\equiv P$，$\gamma(1,s)\equiv Q$：



我们可以认为这族曲线是曲线 $\gamma(t,0)$ 的附近的一个小的扰动。按照定义，曲线 $\gamma_s(t)=\gamma(t,s)$ 的长度为：

$$\ell(\gamma_s)=\int_a^b\sqrt{\big(\dot\gamma_1(s,t)\big)^2+\big(\dot\gamma_2(s,t)\big)^2+\cdots+\big(\dot\gamma_n(s,t)\big)^2}dt.$$

其中，我们用 $\dot\gamma(t,s)$ 表示对 t 的导数。如果 γ_0 的长度是最短的，那么根据极值的导数判定，我们就应该有

$$\frac{d\ell(\gamma_s)}{ds}\Big|_{s=0}=0.$$

此时，我们假设以上给出的曲线族是足够光滑的，即 $\dot\gamma(t,s)$ 在 $[0,1]\times[-\varepsilon,\varepsilon]$ 上是连续的并且 $\dfrac{d}{ds}\left(\dot\gamma(t,s)\right)$ 也是连续的。这样，我们就可以交换积分和求导，得到

$$\begin{aligned}\frac{d\ell(\gamma_s)}{ds}\Big|_{s=0}&=\int_a^b\frac{\dot\gamma(t,0)}{|\dot\gamma(t,0)|}\cdot\frac{d}{ds}\left(\dot\gamma(t,0)\right)dt\\&=\int_a^b\frac{\dot\gamma(t,0)}{|\dot\gamma(t,0)|}\cdot\frac{d}{dt}\left(\frac{d}{ds}\left(\gamma(t,0)\right)\right)dt\\&=-\int_a^b\frac{d}{dt}\left(\frac{\dot\gamma(t,0)}{|\dot\gamma(t,0)|}\right)\cdot\frac{d}{ds}\left(\gamma(t,0)\right)dt.\end{aligned}$$

最后一步是用了分部积分公式，边界项消失是因为这族曲线在 P 和 Q 处是不动的：我们看看 $\dfrac{d}{ds}\left(\gamma(t,0)\right)$ 的几何意义（它被称作是这一族曲线的**变分向量场**）。在每个点 $\dot\gamma(t_0,0)$ 处，它表示的是曲线 $s\mapsto\gamma(t_0,s)$ 在这个点处的切线，如上面图形中红色向量所示。这个计算中的红色等号成立是因为我们交换了两个算子 $\dfrac{d}{dt}$ 和 $\dfrac{d}{ds}$，这是关于偏导数的基本性质，我们下个学期会证明。

关键的观察是如果 γ_0 的长度是最短，不仅仅对于上面选定的这族曲线，无论对于什么样的曲线族，上面的导数都是 0。特别地，任意给定连续可微的 $X:[0,1]\to\mathbb{R}^n$，下面这一族曲线

$$[0,1]\times[-\varepsilon,\varepsilon]\to\mathbb{R}^n,\quad(t,x)\to\gamma_0(t)+sX(t)$$

它的变分向量场恰好就是 X。所以，对任意的连续可微的向量场 $X(t)$，我们都必须有

$$\int_a^b\frac{d}{dt}\left(\frac{\dot\gamma(t,0)}{|\dot\gamma(t,0)|}\right)\cdot X(t)dt=0.$$

特别地，我们可以取 $X=\frac{d}{dt}\left(\frac{\dot\gamma(t,0)}{|\dot\gamma(t,0)|}\right)$，所以，

$$\int_a^b\Big|\frac{d}{dt}\left(\frac{\dot\gamma(t,0)}{|\dot\gamma(t,0)|}\right)\Big|^2dt=0.$$

这表明存在 $v_0 \in \mathbb{R}^n$，使得对任意的 $t \in [0,1]$

$$\frac{d}{dt}\left(\frac{\dot{\gamma}(t,0)}{|\dot{\gamma}(t,0)|}\right) = 0 \quad \Leftrightarrow \quad \dot{\gamma}(t,0) = f(t)v_0.$$

所以，γ_0 这条曲线为

$$t \mapsto P + \left(\int_0^t f(\tau)d\tau\right) v_0.$$

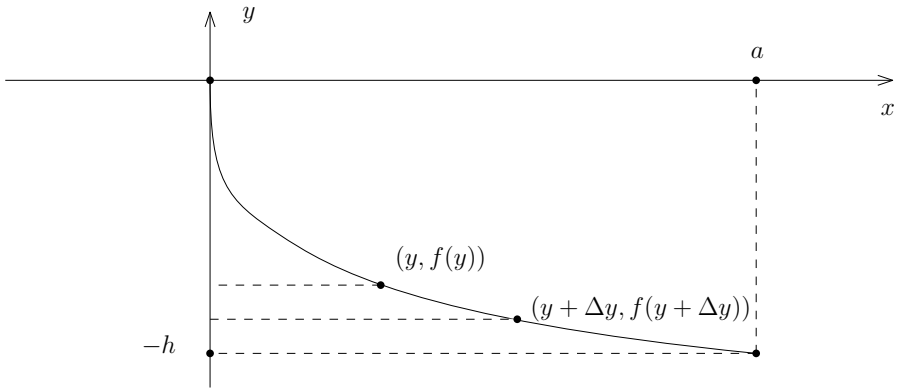
这显然是直线段。

当然，你可能会对上面的证明有质疑，因为两个点之间的最长的距离对这种求导数计算也成立，想一下为什么这个不会发生。实际上，我们还可以计算 ℓ_s 的二阶导数来说明 $\left.\frac{d^2\ell_s}{ds^2}\right|_{s=0} \geqslant 0$ 来说明这是最小值，这里就不再展开讨论了。

25　变分法：最速降线问题，Huygens 等时定理，积分第一中值定理

二零一九年十二月十二日，星期四，晴

最速降线问题



从点 $(x, f(x))$ 到 $(x+\Delta x, f(x+\Delta x))$ 的距离是 $\sqrt{1+f'(x)^2}\Delta x$。根据重力势能到动能的转化（能量守恒），我们知道

$$\frac{1}{2}mv^2 = -mGf(x).$$

在 $(y, f(y))$ 处，小球的速度应该是 $\sqrt{-2Gf(x)}$，所以从 $(x, f(x))$ 到 $\big(x+\Delta x, f(x+\Delta x)\big)$ 所需要的时间是

$$\Delta t = \frac{\sqrt{1+f'(x)^2}\Delta y}{\sqrt{-2Gf(x)}}$$

这样，如果曲线 γ_s 由函数 $f_s(x) = f(x,s)$ 给出，即

$$\gamma_s : [0,a] \to (x, f_s(x)).$$

小球经过整个曲线需要的时间为

$$T(s) = \int_0^a \sqrt{\frac{1+f'(x,s)^2}{-2Gf(x,s)}}dx.$$

其中，f' 指的是对 $f(x,s)$ 的 x 变量求导数。为了保证曲线的两个端点是固定的，我们还要求

$$f(0,s) \equiv 0, \quad f(a,s) \equiv -h.$$

对于最速降线 f_0 而言，我们必然有

$$\frac{d}{ds}T(s)\Big|_{s=0} = 0.$$

为此，我们假设 $L(f,f') = \sqrt{\frac{1+f'(x,s)^2}{-2Gf(x,s)}}$，即在函数

$$L(X,V) = \sqrt{\frac{1+V^2}{-2GX}}$$

中代换变量 $X = f$，$V = f'$。此时，我们有（我们总是假设足够的光滑性使得积分与导数可以交换）：

$$\begin{aligned}\frac{d}{ds}T(s) &= \int_0^a \frac{d}{ds}\big(L(f_s, f_s')\big)dx \\ &= \int_0^a (\frac{dL}{dX})(f_s, f_s')\frac{d}{ds}f(x,s) + (\frac{dL}{dV})(f_s, f_s')\frac{d}{ds}f'(x,s)dx\end{aligned}$$

最后一项中，我们利用两个导数可以交换（下学期），所以 $\dfrac{d}{ds}f'(x,s) = \left(\dfrac{d}{ds}f(x,s)\right)'$。令 $s = 0$ 并记 $f = f_0$，我们就有

$$\begin{aligned}\frac{d}{ds}T(s)\Big|_{s=0} &= \int_0^a (\frac{dL}{dX})(f, f')\frac{d}{ds}f(x,0) + (\frac{dL}{dV})(f, f')\left(\frac{d}{ds}f(x,s)\Big|_{s=0}\right)' dx \\ &= \int_0^a \left((\frac{dL}{dX})(f, f') - \frac{d}{dx}(\frac{dL}{dV})(f, f')\right)\frac{df(x,0)}{ds}dx.\end{aligned}$$

最后一步分部积分的过程中，我们用到了 $f(0,s) \equiv 0$，$f(a,s) \equiv -h$，从而

$$\frac{d}{ds}f(0,s) \equiv 0,\ \frac{d}{ds}f(a,s) \equiv 0.$$

另外，对任意的函数 g（$g(0) = g(a) = 0$），我们通过定义 $f(x,s) = f(x) + sg(x)$ 就可以使得

$$\frac{d}{ds}f(x,s)\Big|_{s=0} = g(x).$$

所以，通过选取

$$g(x) = (\frac{dL}{dX})(f, f') - \frac{d}{dx}(\frac{dL}{dV})(f, f'),$$

我们就得到了 f 必须满足的方程（这被称作是 Euler-Lagrange 方程）：

$$(\frac{dL}{dX})(f, f') - \frac{d}{dx}\left((\frac{dL}{dV})(f, f')\right) = 0.$$

假设上述方程成立，那么

$$\begin{aligned}\frac{d}{dx}\left((V\frac{dL}{dV} - L)(f, f')\right) &= \frac{d}{dx}\left(f'\frac{dL}{dV}(f, f') - L(f, f')\right) \\ &= f''\frac{dL}{dV}(f, f') + f'\frac{d}{dx}\left(\frac{dL}{dV}(f, f')\right) - f'\frac{dL}{dX}(f, f') - f''\frac{dL}{dV}(f, f') \\ &= 0.\end{aligned}$$

这表明，存在常数 C，使得

$$(V\frac{dL}{dV} - L)(f, f') \equiv C$$

按照定义，我们有

$$\frac{dL}{dX} = G\frac{(1+V^2)^{\frac{1}{2}}}{(-2GV)^{\frac{3}{2}}},\quad \frac{dL}{dV} = \frac{V}{\big(-2GX(1+V^2)\big)^{\frac{1}{2}}}.$$

所以，

$$\frac{(f')^2}{\big(-2Gf(1+(f')^2)\big)^{\frac{1}{2}}}-\sqrt{\frac{1+(f')^2}{-2Gf}}=C,\ \Leftrightarrow\ f\big(1+(f')^2\big)=-C_1.$$

我们令 $f(x)=-g(x)$，那么，我们就有

$$g'(x)=\sqrt{\frac{C_1-g}{g}}$$

如果把 g 看作参数，$x=x(g)$，那么，根据反函数的导数的计算，我们有

$$\frac{dx}{dg}=\sqrt{\frac{g}{C_1-g}}$$

从而，作为原函数，我们有

$$x(t)=\int\sqrt{\frac{t}{C_1-t}}dt.$$

为了能积分出这个函数，我们做变量替换

$$t=C_1\sin^2\theta,$$

从而，

$$x(\theta)=2C_1\int\tan\theta\cdot\cos\theta\cdot\sin\theta d\theta=C_1(\theta-\frac{1}{2}\sin(2\theta)).$$

这样子，我们就得到了这条曲线的参数方程（原函数差一个常数 C_2）

$$\theta\mapsto\big(\frac{C_1}{2}(2\theta-\sin(2\theta))+C_2,\frac{1}{2}(1-\cos(2\theta))\big)$$

重新参数化之后，我们得到一条曲线

$$\gamma:[\theta_1,\theta_2]\to\mathbb{R}^2,\ \ \theta\mapsto\big(c(\theta-\sin\theta)+C,-c(1-\cos\theta)\big)$$

这种曲线叫做被称作是摆线或者旋轮线。我们注意到，初始时刻 $y(\theta_1)=0$，所以，$\cos\theta_1=1$，这表明 $\theta_1=2k\pi$，$C=-2ck\pi$，通过再次调整 $\theta\mapsto\theta-2k\pi$，我们可以假设 $\theta_1=0$，$C=0$，所以我们要找的参数曲线为

$$\gamma:[0,\Theta]\to\mathbb{R}^2,\ \ \theta\mapsto\big(c(\theta-\sin\theta),-c(1-\cos\theta)\big)$$

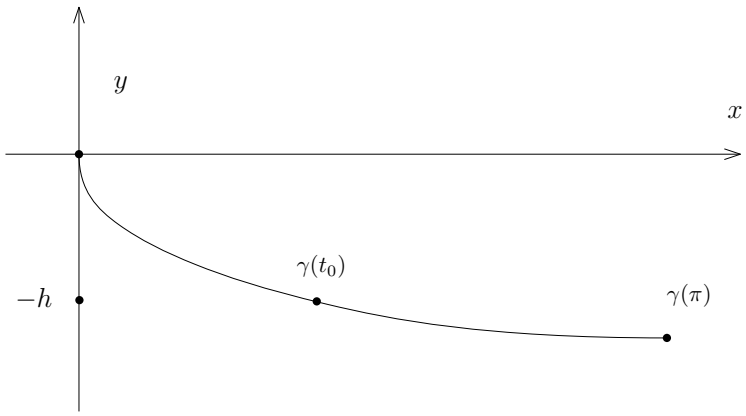
从而，

$$c(\Theta-\sin\Theta)=a,\ \ -c(1-\cos\Theta)=-h$$

可以确定 $\Theta\in[0,\frac{\pi}{2}]$ 和 c。

命题 145 (等时性，Huygens)**.** *考虑摆线*

$$\gamma:[0,\frac{\pi}{2}]\to\mathbb{R}^2,\ \ \theta\mapsto\big(c(\theta-\sin\theta),-c(1-\cos\theta)\big).$$



那么，从任意一点 $\gamma(\theta_0)$ 出发的小球滚动到 $\gamma(\pi)=(c\pi,-2c)$ 的时间都是一样的。

证明: 实际上，小球在 $\gamma(\theta)$ 到 $\gamma(\theta+\Delta\theta)$ 处要经过的距离为:

$$\Delta=|\gamma'(\theta)|\Delta\theta=c\sqrt{2-2\cos\theta}\Delta\theta$$

小球的速度为:

$$v=\sqrt{2Gc(\cos\theta_0-\cos\theta)}.$$

所以，从 $\gamma(\theta_0)$ 滑到底一共需要

$$\begin{aligned}T(\theta_0)&=\sqrt{\frac{c}{G}}\int_{\theta_0}^{\pi}\sqrt{\frac{1-\cos\theta}{\cos\theta_0-\cos\theta}}d\theta\\&=\sqrt{\frac{2c}{G}}\int_{\theta_0}^{\pi}\frac{\sin\left(\frac{\theta}{2}\right)}{\sqrt{\cos^2\left(\frac{\theta_0}{2}\right)-\cos^2\left(\frac{\theta}{2}\right)}}d\theta\end{aligned}$$

我们可以令 $s=\cos\left(\frac{\theta}{2}\right)$ 进行换元积分，据此，得到 $T(\theta_0)\equiv\pi\sqrt{\frac{c}{G}}$。 □

积分中值定理

定理 146 (第一积分中值定理)**.** $[a,b]$ 是有界闭区间，$f,g\in\mathcal{R}([a,b])$ 是实值 *Riemann* 可积函数。我们假设对任意的 $x\in[a,b]$，$g(x)\geqslant 0$。令

$$m=\inf_{x\in I}f(x),\quad M=\sup_{x\in I}f(x).$$

那么，存在 $\ell\in[m,M]$，使得

$$\int_a^b fg=\ell\int_a^b g.$$

特别地，如果进一步要求 f 是连续函数，那么存在 $\xi\in[a,b]$，使得

$$\int_a^b fg=f(\xi)\int_a^b g.$$

证明：由于 $m \leqslant f(x) \leqslant M$，根据 $g \geqslant 0$ 以及积分保持不等号的性质，我们有

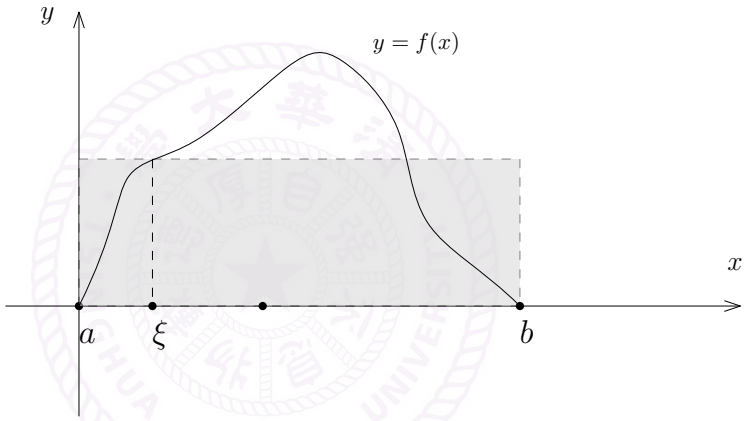
$$m\int_a^b g \leqslant \int_a^b fg \leqslant M\int_a^b g.$$

如果 $\int_a^b g \neq 0$，显然存在 $\ell \in [m, M]$ 使得命题成立；如果 $\int_a^b g = 0$，由于 $g \geqslant 0$，所以 g 在一个零测集之外恒为 0，从而 fg 在一个零测集之外都是 0，从而 $\int_a^b fg = 0$，所以 ℓ 可以任意选。

如果 f 是连续的，所以 f 在 $[a, b]$ 可以取到最大值 M 和最小值 m，即有 $x_1, x_2 \in [a, b]$，使得 $f(x_1) = m$，$f(x_2) = M$，根据连续函数的中值定理，存在 x_1 与 x_2 之间的数 ξ，使得 $f(\xi) = \ell$。 □

推论 147. *假设 f 是连续函数，那么存在 $\xi \in [a, b]$，使得 $\int_a^b f = f(\xi)(b - a)$。*

证明：只要取 $g \equiv 1$ 即可。 □



25.1 作业：可写成两个完全平方数的和的整数的密度

清华大学 19-20 秋季学期，数学分析一，作业 11

请用 A4 大小的纸张**正反面**用钢笔，签字笔或者圆珠笔书写，并注明自己的姓名，年级和作业的**总页数**。除定理公式所涉及的人名之外，请使用**中文**。本次作业的提交时间和地点为 **12 月 19 日**上午的课堂上，逾期视作零分。

基本习题

习题 A（Riemann 积分的概念）

A1) （积分第一中值定理的一个变体）f 是有界闭区间 $[a,b]$ 上的连续函数，$g\in\mathcal{R}([a,b])$ 是 Riemann 可积的并且对任意的 $x\in[a,b]$，$g(x)>0$。证明，存在 $\xi\in(a,b)$，使得

$$\int_a^b fg=f(\xi)\int_a^b g.$$

A2) （单调函数是 Riemann 可积的）不用 Lebesgue 定理证明：f 是 $[a,b]$ 上的单调增函数，那么 $f\in\mathcal{R}(I)$。（提示：利用 Darboux 上下颌）

A3) 证明，函数 $1_{\mathbb{Q}}(x)=\begin{cases}1, & x\in\mathbb{Q}\\ 0, & x\in\mathbb{R}\setminus\mathbb{Q}\end{cases}$ 在 $[0,1]$ 上不是 Riemann 可积的。

A4) 证明: 如果 $f\in\mathcal{R}([a,b])$, 那么 $|f|^p\in\mathcal{R}([a,b])$，其中 $p\geqslant 0$ 是实数。

A5) 我们证明过 Hölder 不等式：对 $x_i,y_i\geqslant 0,p,q>1,\dfrac{1}{p}+\dfrac{1}{q}=1$，有

$$\sum_{i=1}^n x_iy_i\leqslant\Big(\sum_{i=1}^n x_i^p\Big)^{1/p}\Big(\sum_{i=1}^n y_i^q\Big)^{1/q}.$$

试证明积分形式的 Hölder 不等式并给出等号成立的条件：$f,g\in\mathcal{R}([a,b]),p,q>1,\dfrac{1}{p}+\dfrac{1}{q}=1$，那么

$$\Big|\int_a^b fg\Big|\leqslant\Big(\int_a^b|f|^p\Big)^{1/p}\Big(\int_a^b|g|^q\Big)^{1/q}.$$

A6) 证明积分形式的 Minkowski 不等式并给出等号成立的条件: 如果 $f,g\in\mathcal{R}([a,b])$, $p\geq 1$，那么

$$\Big(\int_a^b|f+g|^p\Big)^{1/p}\leqslant\Big(\int_a^b|f|^p\Big)^{1/p}+\Big(\int_a^b|g|^p\Big)^{1/p}.$$

习题 B（凸函数与积分）

B1) 假设 $f\in\mathcal{R}([a,b])$ 并且 f 为凸函数。证明，

$$f\left(\frac{a+b}{2}\right)\leqslant\frac{1}{b-a}\int_a^b f(x)dx\leqslant\frac{f(a)+f(b)}{2}.$$

B2) 假设函数 f 在 $[a,b]$ 上二阶可导并且对任意 x，$f''(x)>0$, $f(x)\leqslant 0$。证明，对任意 x，我们有

$$f(x)\geqslant \frac{2}{b-a}\int_a^b f(y)dy.$$

B3) 假设 f 在 $\mathbb{R}$ 上二阶可导且 $f''(x)\geqslant 0$，$\varphi\in C([a,b])$。证明，我们有

$$\frac{1}{b-a}\int_a^b \big(f\circ\varphi\big)(t)dt\geqslant f\left(\frac{1}{b-a}\int_a^b \varphi(t)dt\right).$$

B4) 假设 $f\in C([a,b])$ 并且对任意 x，$f(x)>0$。证明，

$$\log\left(\frac{1}{b-a}\int_a^b f(x)dx\right)\geqslant \frac{1}{b-a}\int_a^b \log(f(x))dx.$$

B5) 证明积分形式的 Jensen 不等式并给出等号成立的条件：如果 f 是 $\mathbb{R}$ 上的凸函数，$\varphi\in C([0,1])$, 那么

$$f\big(\int_0^1 \varphi\big)\leqslant \int_0^1 f\circ\varphi.$$

习题 C（积分和导数）

C1) 假设 $f\in C^1([0,2])$，$|f'|\leqslant 1$，$f(0)=f(2)=1$。证明，$1\leqslant\displaystyle\int_0^2 f(x)dx\leqslant 3$。

C2) 假设 $f\in C^2([0,1])$。证明，存在 $\xi\in[0,1]$, 使得 $\displaystyle\int_0^1 f(x)dx=f\left(\frac{1}{2}\right)+\frac{1}{24}f''(\xi)$。

C3) 假设 $f\in C^1([a,b])$。证明，$\displaystyle\max_{x\in[a,b]}|f(x)|\leq\frac{1}{b-a}\left|\int_a^b f(x)dx\right|+\int_a^b|f'(x)|dx$。

C4) 设 $f\in C([0,1])$ 而且对任意满足 $g(0)=g(1)=0$ 的 $g\in C([0,1])$，都有 $\displaystyle\int_0^1 f(x)g(x)dx=0$。证明，$f(x)\equiv 0$。

C5) 设 $f\in C([0,1])$ 而且对任意 $n\in\mathbb{Z}_{\geqslant 0}$，都有 $\displaystyle\int_0^1 f(x)x^n dx=0$。证明，$f(x)\equiv 0$。

C6) (Gronwall 不等式) 设 $\varphi\in C([0,T])$ 且对任意 $t\in[0,T]$，有 $|\varphi(t)|\leqslant M+k\displaystyle\int_0^t|\varphi(s)|ds$，其中 M,k 为正实数。试证明，对任意 $t\in[0,T]$，$|\varphi(t)|\leqslant Me^{kt}$。

C7) 假设 $a,b>0$，$f\in C([-a,b])$。如果对任意 $x\in(-a,b)$，$f(x)>0$ 并且 $\displaystyle\int_{-a}^b xf(x)=0$。试证明，$\displaystyle\int_{-a}^b x^2f(x)\leqslant ab\int_{-a}^b f(x)$。

C8) 假设 $f\in C([-1,1])$。证明，$\displaystyle\lim_{\lambda\to 0+}\int_{-1}^1\frac{\lambda}{\lambda^2+x^2}f(x)dx=\pi f(0)$。

C9) 假设 f 在 $[1,\infty)$ 上可导并且 $\int_1^\infty f(x)$ 与 $\int_0^\infty f'(x)$ 都收敛（作为反常积分，请回忆反常积分的定义）。证明，$\lim\limits_{x\to\infty} f(x)=0$.

C10) 试证明，$\int_0^\infty \frac{\sin^2 x}{x^2}=\int_0^\infty \frac{\sin x}{x}$，　　$\int_0^\infty \frac{\cos x}{1+x}=\int_0^\infty \frac{\sin x}{(1+x)^2}$。

习题 D 试计算如下反常积分的数值:

(1) $\int_0^1 \log x dx$

(2) $\int_{-\infty}^{+\infty} \frac{dx}{1+x^2}$

(3) $\int_0^{+\infty} \frac{dx}{1+x^4}$

(4) $\int_0^{+\infty} \frac{1+x^2}{1+x^4}dx$

(5) $\int_{-\infty}^0 xe^x dx$

(6) $\int_0^{+\infty} e^{-\sqrt{x}}dx$

(7) $\int_0^{+\infty} \frac{dx}{(a^2+x^2)^{3/2}}$

(8) $\int_2^{+\infty} \frac{dx}{x^2+x-2}$

(9) $\int_{-\infty}^{+\infty} \frac{dx}{(x^2+x+1)^2}$

(10) $\int_{-1}^1 \frac{dx}{\sqrt{1-x^2}}$

(11) $\int_{-1}^1 \frac{\arcsin x}{\sqrt{1-x^2}}dx$

(12) $\int_{-1}^1 \frac{dx}{(2-x)^2\sqrt{1-x^2}}$

(13) $\int_0^1 \frac{\arcsin\sqrt{x}}{x(1-x)}dx$

(14) $\int_0^1 \frac{(1-x)^n}{\sqrt{x}}dx,\ n\in\mathbb{N}_+$

(15) $\int_0^1 \frac{x^n}{\sqrt{1-x^2}}dx,\ n\in\mathbb{N}_+$

(16) $\int_0^1 x^m(\log x)^n dx, n\in\mathbb{N}_+, m\geqslant 0$

(17) $\int_2^{+\infty} \frac{dx}{x(\log x)^p},\ p>1$

(18) $\int_0^{+\infty} \frac{\log x}{x^2+a^2}dx$

(19) $\int_0^{+\infty} x^n e^{-x}dx, n\in\mathbb{N}_+$

(20) $\int_{-\infty}^{+\infty} \frac{dx}{(ax^2+2bx+c)^n},\ ac-b^2>0$

(21) $\int_0^{+\infty} x^{2n-1}e^{-x^2}dx, n\in\mathbb{N}_+$

(22) $\int_{-\pi}^{\pi} \frac{1-r^2}{1-2r\cos x+r^2}dx,\ 0<r<1$

(23) $\int_0^{+\infty} e^{-ax}\cos bx dx,\ a>0$

(24) $\int_0^{+\infty} e^{-ax}\sin bx dx,\ a>0$

(25) $\int_0^{+\infty} \frac{dx}{x(x+1)\cdots(x+n)}$

(26) $\int_0^{2\pi} \log\sin x dx$

(27) $\int_0^{+\infty} e^{-x^2}dx$

期末考试模拟题（不交作业，鼓励讨论）

题目 E（可写成两个完全平方数的和的整数的密度）在此问题中，$I=(0,\infty)$

第一部分：一个含参数积分的研究

E1) 证明，函数 $\dfrac{e^{-u}}{\sqrt{u}}$ 是 I 上的可积函数（作为反常积分可积），记 $K=\displaystyle\int_0^{\infty}\frac{e^{-u}}{\sqrt{u}}du$。

E2) 证明，对任意的 $x\in I$，$F(x)=\displaystyle\int_0^{\infty}\frac{e^{-u}}{\sqrt{u}(u+x)}du$ 是良好定义的（即此反常积分收敛）。

E3) 证明，$F\in C^1(I)$ 并写下 $F'(x)$ 的积分表达式。

E4) 证明，对任意的 $x\in I$，我们有

$$xF'(x)-(x-\frac{1}{2})F(x)=-K.$$

E5) 通过 $G(x)=\sqrt{x}e^{-x}F(x)$ 定义 $G:I\to\mathbb{R}$。证明，存在实数 C，使得

$$G(x)=C-K\int_0^x\frac{e^{-t}}{\sqrt{t}}dt.$$

E6) 计算 K 的值（提示：计算 G 在 0 和 $+\infty$ 处的极限）。

第二部分：两个函数级数的研究。

我们用级数定义函数

$$f(x)=\sum_{n=1}^{\infty}\frac{e^{-nx}}{\sqrt{n}},\quad g(x)=\sum_{n=0}^{\infty}\sqrt{n}e^{-nx}.$$

E7) 证明，f 和 g 在 I 上是良好定义的并且是 I 上的连续函数。

E8) 证明，对任意 $x\in I$，我们都有

$$\int_1^{\infty}\frac{e^{-ux}}{\sqrt{u}}du\leqslant f(x)\leqslant\int_0^{\infty}\frac{e^{-ux}}{\sqrt{u}}du.$$

E9) 证明，存在常数 C_0，使得

$$\lim_{x\to0^+}\sqrt{x}f(x)=C_0.$$

E10) 我们定义序列 $\{a_n\}_{n\geqslant1}$ 如下：

$$a_n=\Big(\sum_{k=1}^{n}\frac{1}{\sqrt{k}}\Big)-2\sqrt{n}.$$

证明，$\lim\limits_{n\to\infty}a_n$ 存在。

E11) 证明，对任意的 $x\in I$，由级数定义的函数

$$h(x)=\sum_{n\geqslant1}\Big(\sum_{k=1}^{n}\frac{1}{\sqrt{k}}\Big)e^{-nx}$$

是良好定义的（即上面的级数是收敛的）。

E12) 试用 $f(x)$ 来表达 $h(x)$ 并证明存在常数 C_1，使得

$$\lim_{x\to 0^+} x^{\frac{3}{2}} h(x) = C_1.$$

E13) 证明，

$$\lim_{x\to 0^+} x^{\frac{3}{2}} g(x) = \frac{\sqrt{\pi}}{2}.$$

第三部分：由自然数的子集所定义的函数级数的研究。

给定 $A \subset \mathbb{Z}_{\geqslant 0}$，我们可以定义一个数列 $\{a_n\}_{n\geqslant 0}$：

$$a_n = \begin{cases} 1, \text{ 如果} n \in A; \\ 0, \text{ 如果} n \notin A. \end{cases}$$

我们定义集合 $I_A \subset \mathbb{R}_{\geqslant 0}$ 如下：

$$I_A = \big\{ x \geqslant 0 \;\big|\; \text{级数} \sum_{n\geqslant 0} a_n e^{-nx} \text{收敛} \;\big\}.$$

我们定义函数 $f_A : I_A \to \mathbb{R}$ 如下：

$$f_A(x) = \sum_{n\geqslant 0^+} a_n e^{-nx}.$$

我们记 $\Phi(A) = \lim\limits_{x\to 0} x f_A(x)$（如果极限存在的话）并令

$$\mathcal{S} = \big\{ A \subset \mathbb{Z}_{\geqslant 0} \;\big|\; \lim_{x\to 0^+} x f_A(x) \text{存在} \big\}.$$

作为例子，我们令 A_1 表示所有的非零完全平方数，A_2 表示两个非零完全平方数的和：

$$A_1 = \big\{1, 4, 9, 16, \cdots, n^2, (n+1)^2, \cdots \big| n \in \mathbb{Z}_{\geqslant 1}\big\},\ A_2 = \big\{p^2 + q^2 \big| p, q \in \mathbb{Z}_{\geqslant 1}\big\}.$$

E14) 如果 A 是有限集，试确定 I_A；如果 A 是无限集，试确定 I_A。

E15) 给定 $A \subset \mathbb{Z}_{\geqslant 0}$，对于任意 $n \geqslant 0$，我们定义集合 $A_{\leqslant n}$：

$$A_{\leqslant n} = \big\{ \ell \in A | \ell \leqslant n \big\}.$$

我们用 $|A_{\leqslant n}|$ 表示它所包含的元素的个数。证明，对任意的 $x > 0$，级数

$$\sum_{n=0}^{\infty} |A_{\leqslant n}| \cdot e^{-nx}$$

是收敛的，并且满足

$$\sum_{n=0}^{\infty} |A_{\leqslant n}| \cdot e^{-nx} = \frac{f_A(x)}{1 - e^{-x}}.$$

E16) 证明，对任意 $x>0$，都有

$$\frac{f_{A_1}(x)}{1-e^{-x}}=\sum_{n=0}^{\infty}\lfloor\sqrt{n}\rfloor e^{-nx},$$

其中 $\lfloor\sqrt{n}\rfloor$ 代表不超过 $\sqrt{n}$ 的最大整数（即它的整数部分）。

E17) 证明，极限

$$\lim_{x\to 0^+}\sqrt{x}f_{A_1}(x)$$

存在并计算 $\Phi(A_1)$。

E18) 令 $v(n)$ 为集合 $\big\{(p,q)\in\mathbb{Z}_{\geqslant 1}\times\mathbb{Z}_{\geqslant 1}|p^2+q^2=n\big\}$ 的元素个数。证明，对任意的 $x>0$，级数

$$\sum_{n\geqslant 1}v(n)e^{-nx}$$

是收敛的并且

$$\sum_{n\geqslant 1}v(n)e^{-nx}=\big(f_{A_1}(x)\big)^2.$$

E19) 证明，对任意 $x>0$，我们都有

$$f_{A_2}(x)\leqslant\big(f_{A_1}(x)\big)^2$$

据此给出 $\Phi(A_2)$ 的一个上界（我们假设 $\Phi(A_2)$ 存在）。

第四部分：一个 Tauber 型定理及应用。

我们现在假设 $\{a_n\}_{n\geqslant 0}$ 是非负实数组成的序列，使得对任意的 $x>0$，级数

$$S(x)=\sum_{n\geqslant 0}a_ne^{-nx}$$

收敛。我们进一步假设下面的极限存在：

$$\lim_{x\to 0^+}xS(x)=\lim_{x\to 0^+}x\sum_{n\geqslant 0}a_ne^{-nx}=\ell\in[0,+\infty).$$

令 F 为 $[0,1]$ 上的全体实值函数的所构成的线性空间，$E_0=C([0,1])$。令 E 是分段连续的有界函数所组成的空间，即对任意的 $\varphi\in E$，φ 有界并且存在 $\{b_i\}_{i=0,1,\cdots,m}$，使得

$$0=b_0<b_1<b_2<\cdots<b_{m-1}<b_m=1,$$

并且 φ 在每一段 (b_i,b_{i+1}) 上的限制是连续的。我们在 E 上取范数

$$\|\psi\|_\infty=\sup_{x\in[0,1]}|\psi(x)|.$$

E20) 我们定义映射 $L : E \to F$ 如下：

$$\big(L(\psi)\big)(x) = \sum_{n=0}^{\infty} a_n e^{-nx} \psi(e^{-nx}), \quad \psi \in E.$$

证明，L 是良好定义的（上面等式右边的级数对任意的 $x > 0$ 均收敛）并且是线性映射。进一步，如果对任意的 $x \in [0,1]$，$\psi_1(x) \leqslant \psi_2(x)$，那么对任意的 $x \in [0,1]$，证明，

$$\big(L(\psi_1)\big)(x) \leqslant \big(L(\psi_2)\big)(x).$$

E21) 定义 E 的子空间

$$E_1 = \big\{\psi \in E \big| \lim_{x \to 0^+} x\big(L(\psi)\big)(x) \text{存在}\big\}.$$

我们定义线性映射 $\Delta : E_1 \to \mathbb{R}$ 如下：

$$\Delta(\psi) = \lim_{x \to 0^+} x\big(L(\psi)\big)(x), \quad \psi \in E_1.$$

证明，E_1 是 E 的线性子空间并且存在常数 $M > 0$，使得对任意的 $\psi \in E_1$，我们都有

$$|\Delta(\psi)| \leqslant M \|\psi\|_\infty.$$

E22) 对于单项式函数 $P_n(x) = x^n$，证明，$P_n \in E_1$ 并计算 $\Delta(P_n)$。

E23) 证明，$E_0 \subset E_1$ 并对每个 $\psi \in E_0$ 来计算 $\Delta(\psi)$。

E24) 对于 $a \in (0,1)$，$\varepsilon \in \big(0, \min(a, 1-a)\big)$，我们定义函数

$$g_-(x) = \begin{cases} 1, & x \in [0, a-\varepsilon]; \\ \frac{a-x}{\varepsilon}, & x \in (a-\varepsilon, a) \\ 0, & x \in [a,1] \end{cases}, \quad g_+(x) = \begin{cases} 1, & x \in [0, a]; \\ \frac{a+\varepsilon-x}{\varepsilon}, & x \in (a, a+\varepsilon) \\ 0, & x \in [a+\varepsilon, 1] \end{cases},$$

证明，$g_\pm \in E_0$ 并计算 $\Delta(g_\pm)$。进一步证明，示性函数 $\mathbf{1}_{[0,a]} \in E_1$ 并计算 $\Delta(\mathbf{1}_{[0,a]})$。

E25) 证明，$E_1 = E$ 并对 $\psi \in E$ 给出 $\Delta(\psi)$ 的计算公式。

E26) 我们定义函数

$$\psi(x) = \begin{cases} 0, & x \in [0, e^{-1}); \\ \frac{1}{x}, & x \in [e^{-1}, 1]. \end{cases}$$

通过计算 $L(\psi)(\frac{1}{N})$（这里 N 是正整数）证明 Tauber 型公式：

$$\lim_{N \to \infty} \frac{1}{N} \sum_{k=0}^{N} a_k = \ell.$$

E27) 考虑 $A \in \mathcal{S}$（第三部分），试计算

$$\lim_{n\to\infty} \frac{|A_{\leqslant n}|}{n}.$$

这个数值也被称作是 A 在 $\mathbb{Z}_{\geqslant 0}$ 中的自然密度。

E28) 试计算

$$\lim_{n\to\infty} \frac{v(1)+v(2)+\cdots+v(n)}{n},$$

并给出 A_2 的自然密度的一个上界。

God does not care about our mathematical difficulties. He integrates empirically.

Albert Einstein

---

26 第二积分中值定理，Stieltjes 积分

二零一九年十二月十六日，星期一，晴

引理 148 (第二积分中值定理的弱形式，f 和 g 的正则性较好)**.** *假设 $I=[a,b], g\in C(I), f\in C^1(I)$，$f$ 单调递减（未必严格）并且对任意的 $x\in I$，$f(x)\geqslant 0$。那么，一定存在 $c\in[a,b]$，使得*

$$\int_a^b fg = f(a)\int_a^c g.$$

证明： 由于 $g\in C(I)$，我们可以定义它的原函数 $G(x)=\int_a^x g(y)dy$，使得 $G\in C^1(I)$ 并且 $G(a)=0$。再根据 $f\in C^1(I)$，我们可以使用分部积分公式：

$$\int_a^b fg = f(b)G(b)+\int_a^b G(-f')$$

由于 f 是单调下降的，从而 $-f'\geqslant 0$，我们可以利用第一积分中值定理，找到 $\xi\in[a,b]$ 使得

$$\begin{aligned}\int_a^b fg &= f(b)G(b)+G(\xi)\int_a^b(-f')\\ &= f(b)G(b)+G(\xi)(f(a)-f(b)).\end{aligned}$$

通过简单变形（我们明显可以假设 $f(a)\neq 0$），我们得到

$$\frac{1}{f(a)}\int_a^b fg = \frac{f(b)}{f(a)}G(b)+\big(1-\frac{f(b)}{f(a)}\big)G(\xi).$$

这表明 $\dfrac{1}{f(a)}\displaystyle\int_a^b fg$ 是线段 $[G(a),G(\xi)]$ 或者 $[G(\xi),G(a)]$ 上的一个点，对连续函数 G 用中值定理，我们就得到 $c\in[a,\xi]$，使得

$$\frac{1}{f(a)}\int_a^b fg = G(c).$$

这就是所要证明的结论。

我们还可以采取下面的方式来证明：令 $M=\sup\limits_{x\in I}G(x)$，$m=\inf\limits_{x\in I}G(x)$，那么根据

$$\frac{1}{f(a)}\int_a^b fg = \frac{f(b)}{f(a)}G(b)+\big(1-\frac{f(b)}{f(a)}\big)G(\xi).$$

我们得到

$$m\leqslant\frac{1}{f(a)}\int_a^b fg\leqslant M.$$

我们对 $G(x)$ 用介值定理即可。 □

引理 149 (第二积分中值定理)**.** $I=[a,b]\subset\mathbb{R}$ 是闭区间，f 和 g 是 I 上 *Riemann* 可积分的函数。假设 f 是递减的并且对任意的 $x\in I$，$f(x)\geqslant 0$。那么，存在 $c\in[a,b]$，使得

$$\int_a^b fg=f(a)\int_a^c g.$$

注记. 和上一个引理比较，我们不再假设 f 和 g 有较好的正则性（只假设可积性），那么，引理中分部积分的技巧就不能再用了。我们之前提过，所谓的 *Abel* 求和法是分部积分的类比。以下，我们用 *Darboux* 和代替积分，用 *Abel* 求和法代替分部积分，这是一个有启发性并且值得深究的证明。

证明: 我们不妨假设 $f(a)\neq 0$（否则 $f\equiv 0$，此时定理显然成立。我们定义

$$G(x)=\int_a^x g(y)dy.$$

此时，$G(x)$ 是 $[a,b]$ 上面的连续函数（未必可微），而在引理中，这个函数是连续可微的。

根据引理的证明，我们只需要证明下面的不等式即可:

$$mf(a)\leqslant\int_a^b fg\leqslant Mf(a),$$

即可，$M=\sup\limits_{x\in I}G(x)$，$m=\inf\limits_{x\in I}G(x)$，这是因为我们可以利用 $G(x)$ 的介值定理来完成证明，参见上面引理的最后一步。

为了能够最大程度的保持之前的证明，在逼近的意义下，我们要给出如下的分部积分公式

$$\int_a^b fg=f(b)G(b)+\int_a^b G(-f')$$

的替代品。

我们先任意地选取一个分划 $\sigma=\{a=a_0<a_1<\cdots<a_n=b\}$（最终，我们将令它的步长 $|\sigma|\to 0$）。对于每个小区间 $[a_{i-1},a_i]$，我们定义

$$m_i=\inf_{x\in[a_{i-1},a_i]}g(x),\quad M_i=\sup_{x\in[a_{i-1},a_i]}g(x).$$

根据第一中值定理，有 $\ell_i\in[m_i,M_i]$，使得

$$G(a_i)-G(a_{i-1})=\int_{a_{i-1}}^{a_i}g=\ell_i(a_i-a_{i-1}),\quad m_i\leqslant\ell_i\leqslant M_i.$$

特别地，我们知道（注意 $G(a_0)=0$）

$$G(a_k)=\sum_{i=1}^k\ell_i(a_i-a_{i-1}).$$

我们现在要用求和来代替积分 $\int_a^b fg$。任意选取 $\xi_i\in[a_{i-1},a_i]$，如下成立:

$$\Big|\sum_{i=1}^n f(\xi_i)\ell_i(a_i-a_{i-1})-\int_a^b fg\Big|\to 0,\ \text{当 }|\sigma|\to 0\quad\cdots\cdots\ (\star)$$

极限 $(\star)$ 成立并不明显。为此，我们用 Darboux 上下和来逼近积分：当 $|\sigma| \to 0$，我们知道

$$\Big|\sum_{i=1}^{n} f(\xi_i)g(\xi_i)(a_i - a_{i-1}) - \int_a^b fg\Big| \to 0.$$

所以，为了证明 $(\star)$，只需要说明

$$\Big|\sum_{i=1}^{n} f(\xi_i)\ell_i(a_i - a_{i-1}) - \sum_{i=1}^{n} f(\xi_i)g(\xi_i)(a_i - a_{i-1})\Big| \to 0, \quad \text{当 } |\sigma| \to 0.$$

我们有

$$\begin{aligned}
\text{上式左边} &= \Big|\sum_{i=1}^{n} f(\xi_i)\big(\ell_i - g(\xi_i)\big)(a_i - a_{i-1})\Big| \\
&\leqslant \sum_{i=1}^{n} f(\xi_i)(M_i - m_i)(a_i - a_{i-1}) \\
&\leqslant f(a)\left(\sum_{i=1}^{n} M_i(a_i - a_{i-1}) - \sum_{i=1}^{n} m_i(a_i - a_{i-1})\right).
\end{aligned}$$

后者为 Darboux 上下和之差，自然趋向于零，从而 $(\star)$ 得到证明。

为了证明 $mf(a) \leqslant \int_a^b fg \leqslant Mf(a)$，我们利用 Abel 求和法来代替分部积分，从而

$$\begin{aligned}
\sum_{i=1}^{n} f(\xi_i)[\ell_i(a_i - a_{i-1})] &= \sum_{i=1}^{n} f(\xi_i)(G(a_i) - G(a_{i-1})) \\
&= G(b)f(\xi_n) + \sum_{i=1}^{n-1} \big(f(\xi_i) - f(\xi_{i+1})\big)G(a_i)
\end{aligned}$$

根据 f 的性质以及 $G(b), G(a_i) \leqslant M$，我们有

$$\sum_{i=1}^{n} f(\xi_i)[\ell_i(a_i - a_{i-1})] \leqslant Mf(\xi_n) + \sum_{i=1}^{n-1} \big(f(\xi_i) - f(\xi_{i+1})\big)M = f(\xi_1)M.$$

类似地，

$$\sum_{i=1}^{n} f(\xi_i)[\ell_i(a_i - a_{i-1})] \geqslant f(\xi_1)m.$$

另外，ξ_1 的选取是任意的。特别地，我们令 $\xi_1 = a$，这就给出了 $mf(a) \leqslant \int_a^b fg \leqslant Mf(a)$。 $\square$

Stieltjes 积分

给定有界闭区间 $I = [a, b]$ 上递增的函数 $\mu : [a, b] \to \mathbb{R}$。我们重新定义有界**闭**区间 $[c, d] \subset [a, b]$ 的长度：

$$\ell_\mu\big([c, d]\big) = \mu(d) - \mu(c).$$

例子. *1)* 如果 $\mu(x)=x$，那么，上面定义的区间长度就是我们所熟悉的长度。

2) 对任意的 $\rho\in\mathbb{R}\big([a,b]\big)$，$\rho\geqslant 0$，我们令

$$\mu(x)=\int_a^x \rho(x)dx.$$

此时，

$$\ell_\mu\big([c,d]\big)=\int_c^d \rho(x)dx.$$

考虑一个有界函数 $f:[a,b]\to\mathbb{R}$ 和分划 $\sigma\in\mathcal{S}(I)$，其中 $\sigma=\{a=a_0<a_1<\cdots<x_n-1<a_n=b\}$，我们由上面这个新的长度 ℓ_μ 所定义的的 Darboux 上和以及 Darboux 下和。我们 j 记

$$M_i=\sup_{x\in[a_{i-1},a_i]} f(x),\quad m_i=\inf_{x\in[a_{i-1},a_i]} f(x).$$

我们定义

$$\overline{S}_\mu(f;\sigma)=\sum_{i=1}^n M_i\cdot\ell_\mu([a_{i-1},a_i])=\sum_{i=1}^n(\mu(a_i)-\mu(a_{i-1}))M_i,$$
$$\underline{S}_\mu(f;\sigma)=\sum_{i=1}^n m_i\cdot\ell_\mu([a_{i-1},a_i])=\sum_{i=1}^n(\mu(a_i)-\mu(a_{i-1}))m_i.$$

显然，我们有 $\overline{S}_\mu(f;\sigma)\geqslant\underline{S}_\mu(f;\sigma)$。实际上，与 Riemann 积分的情形类似，我们可以证明：

引理 150. *Darboux* 上和与 *Darboux* 下和满足如下的性质：

1) 如果 $\sigma\prec\sigma'$，那么

$$\overline{S}_\mu(f;\sigma)\leqslant\overline{S}_\mu(f;\sigma'),\quad \underline{S}_\mu(f;\sigma)\geqslant\underline{S}_\mu(f;\sigma').$$

2) 对于任意两个分划 $\sigma,\sigma'\in\mathcal{S}(I)$，我们有

$$\underline{S}_\mu(f;\sigma)\leqslant\overline{S}_\mu(f;\sigma').$$

证明：1) 是显然的。为了证明 2)，只需要证明当 $\sigma\prec\sigma'$ 时，我们有

$$\underline{S}_\mu(f;\sigma)\leqslant\overline{S}_\mu(f;\sigma'),\quad \underline{S}_\mu(f;\sigma')\leqslant\overline{S}_\mu(f;\sigma),$$

我们在研究 Riemann 积分的时候就已经证明了类似的结论，这里就不再重复，留作课后的习题。 □

仿照 Riemann 积分的情形，我们还可以定义**上积分**和**下积分**：

$$\overline{\int_a^b} f d\mu=\inf_{\sigma\in\mathcal{S}(I)}\overline{S}_\mu(f;\sigma),\quad \underline{\int_a^b} f d\mu=\sup_{\sigma\in\mathcal{S}(I)}\underline{S}_\mu(f;\sigma).$$

根据上面的引理，我们自然有

$$\overline{\int_a^b} f d\mu\geqslant\underline{\int_a^b} f d\mu.$$

定义 151 (Stieltjes 积分)**.** 给定递增的函数 $\mu:[a,b]\to\mathbb{R}$ 和有界函数 $f:[a,b]\to\mathbb{R}$，如果上述上下积分相等，即

$$\overline{\int_a^b} f d\mu = \underline{\int_a^b} f d\mu,$$

我们就称f **对于** μ **是 Stieltjes 可积的**并将上述数值记作 $\int_a^b f d\mu$。我们用 $\mathcal{R}(I;\mu)$ 表示所有的 *Stieltjes* 可积函数的集合。

注记. 我们注意到如果 $\mu(x)=x$, *Stieltjes* 积分就是 *Riemann* 积分。上述的构造中，我们通过 μ 重新定义了区间的长度（测度），这种新的长度（测度）就给出了新的积分理论。我们将证明 *Stieltjes* 积分满足很多与 *Riemann* 积分类似的性质，比如说 $\mathcal{R}(I;\mu)$ 是 $\mathbb{R}$-线性空间并且

$$\int_a^b \cdot\, d\mu : \mathcal{R}(I;\mu)\to\mathbb{R},\quad f\mapsto \int_a^b f d\mu,$$

是 $\mathbb{R}$-线性映射。

练习. 对任意的 $\rho\in\mathbb{R}\big([a,b]\big)$，$\rho\geqslant 0$，$\mu(x)=\int_a^x \rho(x)dx$。证明，对任意的 $f\in\mathcal{R}\big([a,b]\big)$，都有 $f\in\mathcal{R}\big([a,b];\mu\big)$ 并且

$$\int_a^b f d\mu = \int_a^b f(x)\rho(x)dx.$$

据此，我们也把此时的 $d\mu$ 记作 ρdx 或者 $\rho(x)dx$，ρ 被称作是**密度函数**。特别地，假设 μ 连续可微（μ 默认是递增的），证明，

$$\int_a^b f d\mu = \int_a^b f\mu'.$$

注记. 根据定义，一个函数是 *Stieltjes* 可积的可以用如下的方式判定：$f\in\mathcal{R}(I;\mu)$ 当且仅当对任意的 $\varepsilon>0$，存在分划 $\sigma\in\mathcal{S}(I)$，使得

$$\left|\overline{S}_\mu(f;\sigma)-\underline{S}_\mu(f;\sigma)\right|<\varepsilon.$$

下面的命题表明，我们仍然可以用 Riemann 和来逼近 Stieltjes 积分：

命题 152 (Riemann 和与 Stieltjes 积分)**.** 给定一个 *Stieltjes* 可积的函数 $f\in\mathcal{R}(I;\mu)$，其中 $I=[a,b]$。对任意的 $\varepsilon>0$，存在分划 $\sigma\in\mathcal{S}(I)$，其中 $\sigma=\{a=a_0<a_1<\cdots<x_n-1<a_n=b\}$，使得对任意的 $\xi_i\in[a_{i-1},a_i]$，都有

$$\left|\sum_{i=1}^n f(\xi_i)\cdot\ell([a_{i_1},a_i])-\int_a^b f d\mu\right|<\varepsilon.$$

证明: 第一个不等式的证明是显然，我们只要选取分划 σ，使得

$$\left|\overline{S}_\mu(f;\sigma)-\underline{S}_\mu(f;\sigma)\right|<\varepsilon.$$

此时，因为在每个区间 $[a_{i-1},a_i]$ 上，有 $m_i\leqslant f(\xi_i)\leqslant M_i$，所以

$$\underline{S}_\mu(f;\sigma)\leqslant\sum_{i=1}^n f(\xi_i)\cdot\ell([a_{i_1},a_i])\leqslant\overline{S}_\mu(f;\sigma).$$

这立即就给出了所要的结论。 □

我们来看几类 Stieltjes 可积函数的例子：

例子. 证明下面的函数是 *Stieltjes* 可积的（μ 给定）

1) $[a,b]$ 上的连续函数是 *Stieltjes* 可积的。

用一致连续性来证明：假设 $f\in C([a,b])$，对任意的 $\varepsilon>0$，存在 $\delta>0$，使得当 $|x-y|<\delta$ 时，我们有 $|f(x)-f(y)|<\varepsilon$。特别地，当 $|\sigma|<\delta$ 时，我们有 $M_i-m_i<\varepsilon$，所以

$$\begin{aligned}\left|\overline{S}_\mu(f;\sigma)-\underline{S}_\mu(f;\sigma)\right|&=\sum_{i=1}^n(\mu(a_i)-\mu(a_{i-1}))(M_i-m_i)\\&<\varepsilon\sum_{i=1}^n(\mu(a_i)-\mu(a_{i-1}))=\ell_\mu([a,b])\varepsilon.\end{aligned}$$

2) 如果 μ 是连续的，那么 $[a,b]$ 上的单调函数是 *Stieltjies* 可积的。

假设 f 是 $[a,b]$ 上的递增（不妨假设）函数。根据 μ 的连续性，我们取一个特殊的分划，使得 $\mu(a_i)-\mu(a_{i-1})=\dfrac{\mu(b)-\mu(a)}{n}$。此时，根据 f 是递增的，我们知道 $M_i=f(a_i)$，$m_i=f(a_{i-1})$，所以选取

$$\begin{aligned}\left|\overline{S}_\mu(f;\sigma)-\underline{S}_\mu(f;\sigma)\right|&=\sum_{i=1}^n(\mu(a_i)-\mu(a_{i-1}))(f(a_i)-f(a_{i-1}))\\&<\sum_{i=1}^n\frac{\mu(b)-\mu(a)}{n}(f(a_i)-f(a_{i-1}))\\&=\frac{\mu(b)-\mu(a)}{n}(f(b)-f(a)).\end{aligned}$$

所以，当 $n\to\infty$ 时，我们有 $\left|\overline{S}_\mu(f;\sigma)-\underline{S}_\mu(f;\sigma)\right|\to0$，这就说明 $f\in\mathcal{R}\big([a,b];\mu\big)$。

我们现在研究 Stieltjes 积分的性质。

命题 153. 给定有界闭区间 $I=[a,b]$ 和递增的函数 μ，由此定义的 *Stieltjes* 可积函数具有如下的性质：

1) $\mathcal{R}(I,\mu)$ 是 $\mathbb{R}$-线性空间，积分

$$\int_a^b\cdot\,d\mu:\mathcal{R}(I;\mu)\to\mathbb{R},\quad f\mapsto\int_a^b f\,d\mu$$

是线性映射。

2) 如果对任意的 $x \in I$，我们都有 $f_1(x) \leqslant f_2(x)$，那么，

$$\int_a^b f_1 d\mu \leqslant \int_a^b f_2 d\mu.$$

3)（区间可加性）如果 $f \in \mathcal{R}([a,b];\mu)$，那么对任意的 $c \in [a,b]$，f 在 $[a,c]$ 和 $[c,b]$ 上的限制都是 *Stieltjes* 可积的并且

$$\int_a^b f d\mu = \int_a^c f d\mu + \int_c^b f d\mu.$$

4) 如果 $f \in \mathcal{R}([a,b];\mu)$，那么对于 $|f| \in \mathcal{R}([a,b];\mu)$ 并且

$$\Big|\int_a^b f d\mu\Big| \leqslant \int_a^b |f| d\mu.$$

5) $\lambda > 0$ 是常数，那么

$$\int_a^b f d(\lambda\mu) = \lambda \int_a^b f d\mu.$$

假设 ν 也是 $[a,b]$ 上递增的函数并且 $f \in \mathcal{R}(I;\mu) \cap \mathcal{R}(I;\nu)$，那么 $f \in \mathcal{R}(I;\mu+\nu)$ 并且

$$\int_a^b f d(\mu+\nu) = \int_a^b f d\mu + \int_a^b f d\nu.$$

6) 如果 $f, g \in \mathcal{R}([a,b],\mu)$，那么 $f \cdot g \in \mathcal{R}([a,b],\mu)$。

证明: 证明的想法非常简单：先用适当的 Darboux 上下和的序列逼近积分，对这些序列证明相应的性质，最终取极限。

我们利用 4) 和 5) 来演示一下如何运用上述的想法：对任意的 $\varepsilon > 0$，选取分划 $\sigma \in \mathcal{S}([a,b])$，使得

$$\overline{S}_\mu(f;\sigma) - \underline{S}_\mu(f;\sigma) < \varepsilon.$$

考虑函数 $|f|$ 的 Darboux 上下和

$$\overline{S}_\mu(|f|;\sigma) = \sum_{i=1}^n \tilde{M}_i \cdot \ell_\mu([a_{i-1},a_i]),$$
$$\underline{S}_\mu(|f|;\sigma) = \sum_{i=1}^n \tilde{m}_i \cdot \ell_\mu([a_{i-1},a_i]),$$

其中

$$M_i = \sup_{x\in[a_{i-1},a_i]} f(x), \quad m_i = \inf_{x\in[a_{i-1},a_i]} f(x),$$
$$\tilde{M}_i = \sup_{x\in[a_{i-1},a_i]} |f(x)|, \quad \tilde{m}_i = \inf_{x\in[a_{i-1},a_i]} |f(x)|$$

由于

$$\tilde{M}_i - \tilde{m}_i \leqslant M_i - m_i,$$

所以

$$\overline{S}_\mu(|f|;\sigma)-\underline{S}_\mu(|f|;\sigma)\leqslant\overline{S}_\mu(f;\sigma)-\underline{S}_\mu(f;\sigma)<\varepsilon.$$

这表明，$|f|$ 是 Stieltjes 可积的。另外，对于有限的求和，我们自然有

$$\left|\underline{S}_\mu(f;\sigma)\right|\leqslant\overline{S}_\mu(|f|;\sigma).$$

然后，左右两边分别与 $\left|\int_a^b f d\mu\right|$ 和 $\int_a^b |f|d\mu$ 差不超过 ε，令 $\varepsilon\to 0$，我们就得到了要证明的不等式。

对于 5)，假设 $f\in\mathcal{R}([a,b];\mu)\cap\mathcal{R}([a,b];\nu)$，那么，对于任意给定的 $\varepsilon>0$，存在分划 $\sigma\in\mathcal{S}([a,b])$，使得

$$\overline{S}_\mu(f;\sigma)-\underline{S}_\mu(f;\sigma)<\varepsilon,\ \ \overline{S}_\nu(f;\sigma)-\underline{S}_\nu(f;\sigma)<\varepsilon.$$

另外，根据定义，我们有

$$\begin{aligned}&\overline{S}_{\lambda\mu}(f;\sigma)=\lambda\overline{S}_\mu(f;\sigma),\ \ \overline{S}_{\mu+\nu}(f;\sigma)=\overline{S}_\mu(f;\sigma)+\overline{S}_\nu(f;\sigma),\\&\underline{S}_{\lambda\mu}(f;\sigma)=\lambda\underline{S}_\mu(f;\sigma),\ \ \underline{S}_{\mu+\nu}(f;\sigma)=\underline{S}_\mu(f;\sigma)+\underline{S}_\nu(f;\sigma).\end{aligned}$$

从而，

$$\overline{S}_{\lambda\mu}(f;\sigma)-\underline{S}_{\lambda\mu}(f;\sigma)<\lambda\varepsilon,\ \ \overline{S}_{\mu+\nu}(f;\sigma)-\underline{S}_{\mu+\nu}(f;\sigma)<2\varepsilon.$$

由于 ε 是任意选取的，这就证明了可积性。相应的等式由上述关于 Darboux 上下和的等式取极限立的。

其余几条的证明我们留作作业。 □

27 Stieltjes 积分的中值定理，利用 Stieltjes 积分的中值定理证明第二积分中值定理

二零一九年十二月十九日，星期四，晴

Stieltjes 积分在计数问题中很有用，基本的想法是通过选取适当的 μ，使得和计数有关的求和或者算级数的计算转化为算积分的问题，这样子就可以借鉴微积分中的经验和想法来研究问题。为此，对于 Steiltjes 积分，我们还需要再建立一个基本工具：分部积分。

命题 154 (Stieltjes 积分的分部积分公式)**.** *假设 μ 是闭 $[a,b]$ 区间上的递增函数，$f \in C^1([a,b])$。那么，我们有*

$$\int_a^b f' \cdot \mu = f \cdot \mu\Big|_a^b - \int_a^b f d\mu.$$

证明: 证明的想法和第二积分中值定理的想法如出一辙，我们用求和逼近积分，然后用 Abel 求和来代替分部积分，最终再取极限：

对于任意的分划 $\sigma \in \mathcal{S}(I)$，其中 $\sigma = \{a = a_0 < a_1 < \cdots < x_n - 1 < a_n = b\}$，根据 Lagrange 中值定理，存在 $\xi_i \in [a_{i-1}, a_i]$，使得

$$f(a_i) - f(a_{i-1}) = f'(\xi_i)(a_i - a_{i-1}).$$

我们首先说明，当分划 σ 选取的足够细的时候，我们有

$$\sum_{i=1}^n f'(\xi_i)\mu(a_i)(a_i - a_{i-1}) \to \int_a^b f' \cdot \mu. \quad \cdots \cdots (\star)$$

即对任意的 $\varepsilon > 0$，存在 $\delta > 0$，当分划的步长 $|\sigma| < \delta$ 时，我们有

$$\Big|\sum_{i=1}^n f'(\xi_i)\mu(a_i)(a_i - a_{i-1}) - \int_a^b f' \cdot \mu\Big| < \varepsilon.$$

这样就表明 $\sum_{i=1}^n f'(\xi_i)\mu(a_i)(a_i - a_{i-1})$ 是要证的公式左边的一个精确的逼近。现在证明 $(\star)$：

$$\begin{aligned}
&\Big|\sum_{i=1}^n f'(\xi_i)\mu(a_i)(a_i - a_{i-1}) - \int_a^b f' \cdot \mu\Big| \\
\leqslant &\Big|\sum_{i=1}^n f'(\xi_i)\big(\mu(a_i) - \mu(\xi_i)\big)(a_i - a_{i-1})\Big| + \underbrace{\Big|\sum_{i=1}^n f'(\xi_i)\mu(\xi_i)(a_i - a_{i-1}) - \int_a^b f' \cdot \mu\Big|}_{\text{按照 Riemann 积分的定义，这一项可以很小}=o(1)} \\
\leqslant &\sup_{x\in I}|f'(x)| \underbrace{\sum_{i=1}^n |\mu(a_i) - \mu(\xi_i)|(a_i - a_{i-1})}_{\text{这一项被 } \mu \in \mathcal{R}(I) \text{ 的 Darboux 上下和之差控制}} + o(1).
\end{aligned}$$

在上面的推导中，我们用到了如下的事实（作业已经证明）：单调函数是 Riemann 可积的。当 $|\sigma|$ 足够小的时候，上面最后一个不等号右边的第一项可以任意小，这就证明了 $(\star)$。

现在可以证明分部积分公式了。我们对 $\sum_{i=1}^{n} f'(\xi_i)\mu(a_i)(a_i - a_{i-1})$ 用 Abel 求和法：

$$\begin{aligned}\sum_{i=1}^{n} f'(\xi_i)\mu(a_i)(a_i - a_{i-1}) &= \sum_{i=1}^{n} \mu(a_i)(f(a_i) - f(a_{i-1})) \\ &= \mu(a_n)f(a_n) - \mu(a_1)f(a_0) + \sum_{i=1}^{n-1} f(a_i)(\mu(a_{i+1}) - \mu(a_i)) \\ &= \underbrace{\mu(a_n)f(a_n) - \mu(a_0)f(a_0)}_{=f\cdot\mu\big|_a^b} + \underbrace{\sum_{i=0}^{n-1} f(a_i)(\mu(a_{i+1}) - \mu(a_i))}_{\mathbf{I}}.\end{aligned}$$

上式中

$$\mathbf{I} = \sum_{i=0}^{n-1} f(a_i)\ell_\mu\big([a_i, a_{i+1}]\big) = \int_a^b f d\mu + o(1).$$

我们可以通过选取 σ 使得 f 与 $\int_a^b f d\mu$ 的差距是 $o(1)$，最终取极限就证明了分部积分公式。 □

我们要用 Stieltjes 积分的观点来看级数。为此，先研究一个基本的例子，这既是所谓的 Dirac δ-函数：

例子. $[a, b] \subset \mathbb{R}$ 是有界闭区间，$c \in [a, b]$，我们考虑示性函数

$$\mu_c = \mathbf{1}_{\geqslant c}(x) = \begin{cases} 1, & x \geqslant c; \\ 0, & x < c. \end{cases}$$

这是递增的函数。假设 $f : [a, b] \to \mathbb{R}$ 是有界函数并且 f 在 c 处**连续**，我们来计算 f 在测度 $d\mu_c$ 下的 *Stieltjes* 积分。

根据 μ_c 的构造，我们选取一个分划 $\sigma \in \mathcal{S}([a, b])$，使得

$$a = a_0 < a_1 < \cdots < a_k = c < a_{k+1} < \cdots < a_n = b.$$

此时，相应的 *Darboux* 上下和之差为

$$\begin{aligned}\overline{S}_{\mu_c}(f;\sigma) - \underline{S}_{\mu_c}(f;\sigma) &= \underbrace{\sum_{i=1}^{k-2}(M_i - m_i)\big(\mu(a_i) - \mu(a_{i+1})\big)}_{=0} + (M_k - m_k)(\mu(a_k) - \mu(a_{k-1})) \\ &\quad + \underbrace{\sum_{i=k}^{n}(M_i - m_i)\big(\mu(a_i) - \mu(a_{i+1})\big)}_{=0} \\ &= (M_k - m_k)\big(\mu(a_k) - \mu(a_{k-1})\big).\end{aligned}$$

根据 f 在 c 处的连续性，如果包含 c 的分划越来越小，那么 $M_k - m_k \to 0$，从而上式的极限为零，所以 f 是可积的。有了可积性，我们可以选任意的 *Riemann* 和来计算积分，特别地，我们选取 $\xi_i = a_i$，我们就有如下的 *Riemann* 和：

$$\underbrace{\sum_{i=1}^{k-2} f(a_i)\big(\mu(a_i) - \mu(a_{i+1})\big)}_{=0} + f(a_k)(\mu(a_k) - \mu(a_{k-1})) + \underbrace{\sum_{i=k}^{n} f(a_i)\big(\mu(a_i) - \mu(a_{i+1})\big)}_{=0} = f(c).$$

此时，我们令分划的步长趋于零，从而

$$\int_a^b f d\mu_c = f(c).$$

注记. 假设 f 在 $[a,b]$ 连续可微，即 $f \in C^1([a,b])$，根据分部积分公式，我们有

$$f(b) - \int_a^b f d\mu = \int_a^b f' \cdot \mu = \int_c^b f' = f(b) - f(c).$$

这也能给出上述结论。

引理 155. 假设 $\{\alpha_n\}_{n\geqslant 1}$ 是正实数所组成的序列并且 $\sum_{n=1}^{\infty} \alpha_n$ 收敛，$\{x_n\}_{n\geqslant 1} \subset (a,b)$。定义递增函数

$$\mu = \sum_{n=1}^{\infty} \alpha_n \mu_{x_n} = \sum_{n=1}^{\infty} \alpha_n \mathbf{1}_{\geqslant x_n}.$$

那么，对于 $f \in C([a,b])$，我们有

$$\int_a^b f d\mu = \sum_{i=1}^{\infty} \alpha_i f(x_i).$$

证明: 由于级数 $\sum_{n=1}^{\infty} \alpha_n$ 收敛，对于任意给定的 $\varepsilon > 0$，我们可以选取 N，使得 $\sum_{n\geqslant N+1} \alpha_n < \varepsilon$。我们将 μ 分拆为两个部分（后一部分的贡献很小，是所谓的误差项）：

$$\mu = \underbrace{\sum_{i=1}^{N} \alpha_n \mu_{x_n}}_{\mu_{\leqslant N}} + \underbrace{\sum_{i=N+1}^{\infty} \alpha_n \mu_{x_n}}_{\mu_{\geqslant N+1}}.$$

根据上面计算的例子，我们有

$$\begin{aligned}\int_a^b f d\mu - \sum_{i=1}^{\infty} \alpha_i f(x_i) &= \left(\int_a^b f d\mu_{\leqslant N} - \sum_{i=1}^{N} \alpha_i f(x_i)\right) + \left(\int_a^b f d\mu_{\geqslant N+1} - \sum_{i=N+1}^{\infty} \alpha_i f(x_i)\right) \\ &= \int_a^b f d\mu_{\geqslant N+1} - \underbrace{\sum_{i=N+1}^{\infty} \alpha_i f(x_i)}_{\leqslant \varepsilon \|f\|_\infty}\end{aligned}$$

很明显，根据定义，我们有 $\|\mu_{\geqslant N+1}\|_{L^\infty} < \varepsilon$，所以

$$\begin{aligned}\int_a^b f d\mu_{\geqslant N+1} &\leqslant \|f\|_\infty \int_a^b d\mu_{\geqslant N+1} \\ &= \|f\|_\infty \big(\mu_{\geqslant N+1}(b) - \mu_{\geqslant N+1}(a)\big) \leqslant 2\varepsilon \|f\|_\infty.\end{aligned}$$

从而，

$$\Big| \int_a^b f d\mu - \sum_{i=1}^{\infty} \alpha_i f(x_i) \Big| \leqslant 3\varepsilon \|f\|_\infty.$$

令 $\varepsilon \to 0$，我们就证明了这个命题。 □

注记. 这个命题对于反常积分的情形也成立，为此，我们需要将 *Stieltjes* 积分的概念加以推广。基于 *Riemann* 积分的情况，这些推广实质上只是语言上的简单替换，并没有跟深刻的数学含义，所以我们仅仅给出例子说明。比如说，我们在区间 $[0,\infty)$ 定义 *Stieltjes* 积分，我们假设 f 是有界连续（你当然可以用更弱的假设），我们首先定义

$$\int_0^\infty f d\mu = \lim_{M\to\infty} \int_0^M f d\mu.$$

假设 $\{\alpha_n\}_{n\geqslant 1}$ 是正实数所组成的序列并且 $\sum_{n=1}^{\infty} \alpha_n$ 收敛。定义递增函数

$$\mu = \sum_{n=1}^{\infty} \alpha_n \mu_n = \sum_{n=1}^{\infty} \alpha_n \mathbf{1}_{\geqslant n}.$$

那么，我们有

$$\int_1^\infty f d\mu = \sum_{n=1}^{\infty} \alpha_n f(n).$$

证明和上面的命题是一致的，我们留作作业。

作为 Stieltjes 积分的应用，我们首先来证明第二积分中值定理（上次已经证明了，下面的版本中 g 是连续的）：

命题 156 (第二积分中值定理)**.** 给定闭区间 $I = [a,b] \subset \mathbb{R}$ 和 *Riemann* 可积的函数 $f, g \in R(I)$。假设 f 是递减并且对任意的 $x \in I$，$f(x) \geqslant 0$；函数 g 是连续函数。那么，一定存在 $c \in [a,b]$，使得

$$\int_a^b fg = f(a) \int_a^c g.$$

证明*:* 之前证明的瓶颈在于不能自由地运用分部积分。现在定义

$$\mu = -f, \;\; G(x) = \int_a^x g(y) dy.$$

由于 g 是连续函数，从而 $G(x)$ 是连续可微的。根据 Stieltjes 积分版本的分部积分公式，我们有

$$\int_a^b G'(-f) = G\cdot(-f)\big|_a^b - \int_a^b G d\mu.$$

即

$$\int_a^b fg = G(b)f(b) + \int_a^b G d\mu.$$

令 $M = \sup\limits_{x\in I} G(x)$，$m = \inf\limits_{x\in I} G(x)$，那么，根据上面的式子以及 $\mu = -f$，我们有

$$\int_a^b fg \leqslant f(b)M + M\int_a^b d\mu = f(a)M.$$

类似地，我们有

$$\int_a^b fg \geqslant f(b)m + m\int_a^b d\mu = f(a)m.$$

这就证明了

$$f(a)m \leqslant \int_a^b fg \leqslant f(a)M.$$

根据之前的证明，从这个不等式出发第二中值定理的证明是明显的。 □

命题 157 (第二积分中值定理的第二个版本)**.** *假设 $I=[a,b]\subset\mathbb{R}$ 是有界闭区间，$f\in\mathcal{R}(I)$ 是递增的函数，$g\in C(I)$ 是连续函数。那么，存在 $c\in[a,b]$，使得*

$$\int_a^b fg = f(a)\int_a^c g + f(b)\int_c^b g(t)dt.$$

证明: 我们选取 Stieltjes 积分中的测度 $d\mu$ 为 $\mu = f$。令 $G(x) = \int_a^x g(y)dy$。所要证明的等式的右边等于

$$\begin{aligned} f(a)G(c) + f(b)\big(G(b)-G(c)\big) &= \big(f(a)-f(b)\big)G(c) + f(b)G(b) \\ &= f(b)G(b) - G(c)\mu([a,b]) \end{aligned}$$

根据 Stieltjes 积分的分部积分公式，我们有

$$\begin{aligned} \int_a^b fg = \int_a^b G'\cdot f &= G\cdot f\big|_a^b - \int_a^b G d\mu \\ &= G(b)f(b) - \int_a^b G d\mu. \end{aligned}$$

据此，只需要说明存在 $c\in[a,b]$，使得

$$\int_a^b G d\mu = G(c)\mu([a,b]).$$

很明显，这是 Stieltjes 积分情形下的积分第一中值定理，这是非常容易证明的，我们也留作作业。 □

我们只给出第一中值定理的叙述：

定理 158 (第一积分中值定理)**.** $[a,b]$ 是有界闭区间，μ 是该区间上单调递增的函数，$f,g\in\mathcal{R}([a,b];\mu)$ 是实值 *Riemann* 可积函数。我们假设对任意的 $x\in[a,b]$，$g(x)\geqslant 0$。令

$$m=\inf_{x\in I}f(x),\quad M=\sup_{x\in I}f(x).$$

那么，存在 $\ell\in[m,M]$，使得

$$\int_a^b fgd\mu=\ell\int_a^b gd\mu.$$

特别地，如果进一步要求 f 是连续函数，那么存在 $\xi\in[a,b]$，使得

$$\int_a^b fgd\mu=f(\xi)\int_a^b gd\mu.$$

练习. 你是否可以构造一个 *Stieltjes* 积分来说明 *Abel* 求和法也是 *Stieltjes* 积分意义下的分部积分公式的特例？

我们再看一个例子，这是我们之前提到过的面积方法来研究积分和求和之间的关系。

引理 159. 假设 $a,b\in\mathbb{Z}$，f 是 $[a,b]$ 上的单调函数并且是 C^1 的，那么存在 $\theta=\theta(a,b;f)\in[0,1]$，使得

$$\sum_{a<n\leqslant b}f(n)=\int_a^b f(x)dx+\theta\cdot(f(b)-f(a)).$$

证明：我们定义函数 $\mu(x)=\lfloor x\rfloor$，即任意给定 $x\in\mathbb{R}$，$\mu(x)$ 是不超过 x 的最大的整数。很明显，这是在 $\mathbb{R}$ 上定义的递增的函数。由于 a 和 b 都是整数，所以

$$\sum_{a<n\leqslant b}f(n)=\int_a^b fd\mu.$$

据此，我们有

$$\begin{aligned}\sum_{a<n\leqslant b}f(n)-\int_a^b f(x)dx&=\int_a^b f(x)d\lfloor x\rfloor-\int_a^b f(x)dx\\&=-\int_a^b f'(x)\lfloor x\rfloor dx+f(x)\lfloor x\rfloor\Big|_a^b+\int_a^b f'(x)xdx-f(x)x\Big|_a^b\\&=\int_a^b\{x\}df-f(x)\{x\}\Big|_a^b.\end{aligned}$$

其中，$\{x\}=x-\lfloor x\rfloor$。这个函数在整数处取值为 0，所以

$$\sum_{a<n\leqslant b}f(n)-\int_a^b f(x)dx=\int_a^b\{x\}df.$$

最终，我们注意到 $\{x\}\in[0,1)$，现在对 f 所定义的 Stieltjes 积分用第一中值定理就证明了命题。 □

注记. 我们只需要假设 $f \in C([a,b])$ 并且 f 是递增的即可，上述结论仍然成立。

注记. 我们可以简单总结一下为什么要对积分理论的加以推广。首先要指出的是，推广本身或者理论的抽象本身并不是目标。我们主要是想能够在更广阔的天地里看到微积分理论与其它数学对象的相似性，从而通过类比可以运用微积分中最基本的工具来处理其它的问题，比如说分部积分（这是联系局部和整体的最重要工具)。

在数学上，我们对大部分对象的了解都是基于它所满足的性质（不精确地讲，这既是代数上称之为万有性质的性质）而不是定义本身。我们已经见过这样的例子：比如说三角函数 sin 和 cos，只要知道它们满足和差化积公式、知道它们的导数和原函数，我们完全代数的操作证明了它们的其它性质，不夸张的说这些性质完整刻画了三角函数。

所以，我们可以想象，如果我们要建立一套积分理论，它需要有哪些基本的性质：线性，三角不等式以及分部积分（等)。一旦一个积分理论可以做这些操作，那么我们就可以照搬微积分的做法得到很多很有用的结论。下个学期的学习我们也将一直按照这个思路来进行。

27.1 作业：振荡积分

清华大学 19-20 秋季学期，数学分析一，作业 12

请用 A4 大小的纸张**正反面**用钢笔，签字笔或者圆珠笔书写，并注明自己的姓名，年级和作业的**总页数**。除定理公式所涉及的人名之外，请使用**中文**。本次作业的提交时间和地点为 **12 月 26 日**上午的课堂上，逾期视作零分。

基本习题

习题 A（Stieltjes 积分） 如果不加说明，μ 表示的是有界闭区间 $I=[a,b]$ 上的一个递增的函数。下面题目中 A1)，A3)，A4)，A5)，A6) 是对概念的直接验证，可以不提交作业。

A1) 对于任意两个分划 $\sigma,\sigma'\in\mathcal{S}(I)$，我们有

$$\underline{S}_\mu(f;\sigma)\leqslant\overline{S}_\mu(f;\sigma').$$

A2) 对任意的 $\rho\in C\big([a,b]\big)$，$\rho\geqslant 0$，$\mu(x)=\displaystyle\int_a^x\rho(x)dx$。证明，对任意的 $f\in\mathcal{R}\big([a,b]\big)$，都有 $f\in\mathcal{R}\big([a,b];\mu\big)$ 并且

$$\int_a^b f d\mu=\int_a^b f(x)\rho(x)dx.$$

特别地，假设 μ 连续可微（μ 默认是递增的），我们有

$$\int_a^b f d\mu=\int_a^b f\mu'.$$

A3) 证明，$\mathcal{R}(I,\mu)$ 是 $\mathbb{R}$-线性空间并且积分算子

$$\int_a^b\cdot\,d\mu:\mathcal{R}(I;\mu)\to\mathbb{R},\quad f\mapsto\int_a^b f d\mu$$

是线性映射。

A4) 假设 $f_1,f_2\in\mathcal{R}(I,\mu)$。如果对任意的 $x\in I$，我们都有 $f_1(x)\leqslant f_2(x)$，那么，

$$\int_a^b f_1 d\mu\leqslant\int_a^b f_2 d\mu.$$

A5)（区间可加性）如果 $f\in\mathcal{R}([a,b];\mu)$，那么对任意的 $c\in[a,b]$，f 在 $[a,c]$ 和 $[c,b]$ 上的限制都是 Stieltjes 可积的并且

$$\int_a^b f d\mu=\int_a^c f d\mu+\int_c^b f d\mu.$$

A6) 如果 $f,g\in\mathcal{R}([a,b];\mu)$，那么 $f\cdot g\in\mathcal{R}([a,b];\mu)$。

A7) 我们在区间 $[0,\infty)$ 定义 Stieltjes 积分：假设 $f\in C([0,\infty)$ 是有界连续函数，定义

$$\int_0^\infty f d\mu=\lim_{M\to\infty}\int_0^M f d\mu.$$

假设 $\{\alpha_n\}_{n\geqslant 1}$ 是正实数所组成的序列并且 $\sum\limits_{n=1}^{\infty}\alpha_n$ 收敛，定义递增函数 $\mu=\sum\limits_{n=1}^{\infty}\alpha_n\mathbf{1}_{\geqslant n}$。那么，我们有

$$\int_1^\infty f d\mu=\sum_{n=1}^{\infty}\alpha_n f(n).$$

A8)（第一积分中值定理）$f,g\in\mathcal{R}([a,b];\mu)$ 是实值 Riemann 可积函数。我们假设对任意的 $x\in[a,b]$，$g(x)\geqslant 0$。令

$$m=\inf_{x\in I}f(x),\quad M=\sup_{x\in I}f(x).$$

那么，存在 $\ell\in[m,M]$，使得

$$\int_a^b fgd\mu=\ell\int_a^b gd\mu.$$

特别地，如果进一步要求 f 是连续函数，那么存在 $\xi\in[a,b]$，使得

$$\int_a^b fgd\mu=f(\xi)\int_a^b gd\mu.$$

A9) 你是否可以构造一个 Stieltjes 积分来说明 Abel 求和法也是 Stieltjes 积分意义下的分部积分公式的特例?

习题 B（反常积分收敛的判断）以下记号中，b 可以是 ∞。

B1) (Cauchy 判别法) 假设 $f:[a,b)\to\mathbb{R}$，对任意 $b^-<b$，f 在 $[a,b^-]$ 上可积。证明，反常积分 $\int_a^b f(x)dx$ 存在的充要条件是: 对任意 $\varepsilon>0$, 存在 $b(\varepsilon)\in(a,b)$, 使得对任意 $b',b''>b(\varepsilon)$, $\left|\int_{b'}^{b''}f(x)dx\right|<\varepsilon$。

B2) (比较判别法，已经证明，可不交作业) 如果 $|f|$ 在 $[a,b)$ 上反常可积，就称积分 $\int_a^b f(x)dx$ **绝对收敛**。证明，如果 $|f(x)|\leqslant F(x),x\in[a,b)$ 并且 $\int_a^b F(x)dx$ 收敛，那么积分 $\int_a^b|f(x)|dx$ 收敛。

B3) 证明 **Dirichlet 判别法**：假设 $f,g:[a,\infty)\to\mathbb{R}$ 满足

- f 是连续函数并且存在 $A>0$，使得对任意的 $M\geqslant a$，我们都有

$$\left|\int_a^M f(x)dx\right|\leqslant A.$$

- g 是单调函数并且 $\lim\limits_{x\to\infty} g(x)=0$。

那么反常积分 $\displaystyle\int_a^\infty f(x)g(x)dx$ 收敛。

（你可以查阅文献来证明这个判别法，一个好的开始是先假设 f 和 g 都是连续可微的，一般的情形可以考虑 Stieltjes 积分。请对比级数收敛的 Dirichlet 判别法以及相应的证明，你可以看到 Abel 求和法和分部积分之间的相似性）

B4) 证明 **Abel 判别法**：假设 $f,g:[a,\infty)\to\mathbb{R}$ 满足

- 反常积分 $\displaystyle\int_a^\infty f(x)dx$ 存在。
- g 是单调函数并且 g 有界。

那么反常积分 $\displaystyle\int_a^\infty f(x)g(x)dx$ 收敛。（请对比级数收敛的 Abel 判别法以及相应的证明）

B5) 判断下列积分的收敛性 (绝对收敛、条件收敛、发散)

(1) $\displaystyle\int_0^{+\infty}\frac{\log(1+x)}{x^p}dx$　(2) $\displaystyle\int_1^{+\infty}\frac{\sin x}{x^p}dx$　(3) $\displaystyle\int_1^{+\infty}\sin x^2dx$

(4) $\displaystyle\int_0^{+\infty}\frac{\sin^2 x}{x}dx$　(5) $\displaystyle\int_0^{2\pi}\frac{dx}{\cos^p x\cos^q x},\ p,q>0$　(6) $\displaystyle\int_0^{+\infty}x^p\sin(x^q)dx$

(7) $\displaystyle\int_0^{+\infty}\frac{x^p\sin x}{1+x^q},\ q\geq 0$　(8) $\displaystyle\int_0^{\pi/2}\frac{\log\sin x}{\sqrt{x}}dx$　(9) $\displaystyle\int_e^{+\infty}\frac{\log\log x}{\log x}\sin xdx$

8 字班的考试题（之一）（不交作业，鼓励讨论）

题目 C（振荡积分）

我们采取下面的约定：$F(t)$ 和 $G(t)$ 是 $[1,+\infty)$ 上定义的两个函数，$\lim\limits_{t\to\infty}G(t)=0$。我们假设对任意的 $t\geqslant 1$，$G(t)\neq 0$。如果

$$\lim_{t\to\infty}\frac{F(t)}{G(t)}=1,$$

我们就说 $F(t)$ 和 $G(t)$ 是**均阶的**（即收敛到 0 的速度是一样的）并记作 $F(t)\sim G(t)$。

第一部分：相函数是实值的情形

C1) $d>0$ 是给定的实数。假设 $g\in C^1([0,d])$。证明，存在常数 C（可能依赖于 g 和 d），使得

$$\Big|\int_0^d e^{-tx}g(x)dx\Big|\leqslant\frac{C}{t}.$$

C2) $d>0$ 是给定的实数。假设 $g\in C([0,d])$ 并且 $g(0)\neq 0$。证明，

$$\int_0^d e^{-tx}g(x)dx\sim\frac{g(0)}{t}.$$

（提示：利用积分换元 $y=tx$）

C3) $d>0$ 是给定的实数。假设 $g\in C([0,d])$ 并且 $g(0)\neq 0$。证明，

$$\int_0^d e^{-tx^2}g(x)dx\sim\frac{\sqrt{\pi}\cdot g(0)}{2\sqrt{t}}.$$

（提示：我们课堂上已经证明了 $\int_0^\infty e^{-x^2}dx=\frac{\sqrt{\pi}}{2}$）

对于 $t\geqslant 1$，$f\in C([a,b])$ 和 $\varphi\in C([a,b])$，我们定义函数

$$F(t)=\int_a^b e^{-t\varphi(x)}f(x)dx,$$

我们的目标是研究当 $t\to+\infty$ 时，$F(t)$ 的行为。

C4) 假设 $\varphi\in C^1([a,b])$，并且对任意 $x\in[a,b]$，$\varphi'(x)\neq 0$。为了简单起见，我们进一步假设 $\varphi'(x)>0$。令 $d=\varphi(b)-\varphi(a)$。证明，映射

$$\Phi:[a,b]\to[0,d],\quad x\mapsto\varphi(x)-\varphi(a),$$

在 $[a,b]$ 上连续可微的并且是双射。

C5) 假设 $\varphi\in C^1([a,b])$，并且对任意 $x\in[a,b]$，$\varphi'(x)\neq 0$。证明，如果 $f(a)\neq 0$，当 $t\to+\infty$ 时，

$$F(t)\sim\frac{f(a)}{\varphi'(a)}\frac{e^{-t\varphi(a)}}{t}.$$

（提示：利用 C4) 中的函数做积分换元）

C6) 假设 $\varphi\in C^2([a,b])$，$\varphi'(a)=0$，$\varphi''(a)>0$ 并且对任意 $x\in(a,b]$，$\varphi'(x)>0$。令 $d=\sqrt{\varphi(b)-\varphi(a)}$。证明，映射

$$\Psi:[a,b]\to[0,d],\quad x\mapsto\sqrt{\varphi(x)-\varphi(a)},$$

在 $[a,b]$ 上连续可微的并且是双射。

特别地，试计算 $\Psi'(a)$。

（提示：考虑 φ 的二阶的 Taylor 展开。）

C7) 假设 $\varphi \in C^2([a,b])$，$\varphi'(a)=0$，$\varphi''(a)>0$ 并且对任意 $x\in(a,b]$，$\varphi'(x)>0$。证明，如果 $f(a)\neq 0$，当 $t\to+\infty$ 时，

$$F(t)\sim \frac{\sqrt{\pi}f(a)}{\sqrt{2\varphi''(a)}}\frac{e^{-t\varphi(a)}}{\sqrt{t}}.$$

（提示：利用 C6) 中的函数做积分换元）

C8) 给定两个函数 $f\in C\big((0,+\infty)\big)$，$\varphi\in C^2\big((0,+\infty)\big)$。我们假设

- 存在唯一的 $a\in(0,\infty)$，使得 $\varphi'(a)=0$；
- $\varphi''(a)>0$，$f(a)\neq 0$；
- 反常积分 $\displaystyle\int_0^\infty e^{-\varphi(x)}|f(x)|dx$ 收敛。

证明，当 $t\to+\infty$ 时，函数 $G(t)=\displaystyle\int_0^\infty e^{-t\varphi(x)}f(x)dx$ 满足

$$G(t)\sim \frac{\sqrt{2\pi}f(a)}{\sqrt{\varphi''(a)}}\frac{e^{-t\varphi(a)}}{\sqrt{t}}.$$

C9) 对于 $x>0$，定义函数

$$\Gamma(x)=\int_0^\infty t^{x-1}e^{-t}dt.$$

对于正整数 n，试计算 $\Gamma(n)$。

C10) 证明 Stirling 公式：

$$n!\sim\sqrt{2\pi}\frac{n^{n+\frac{1}{2}}}{e^n}.$$

（提示：我们有 $\Gamma(n+1)=n^{n+1}\displaystyle\int_0^\infty e^{-n(x-\log x)}dx$。）

第二部分：振荡积分

对于 $\lambda\geqslant 1$，$f\in C^\infty([a,b])$ 和 $\varphi\in C^\infty([a,b])$，我们定义函数

$$I(\lambda)=\int_a^b e^{i\lambda\varphi(x)}f(x)dx,$$

我们的目标是研究当 $\lambda\to+\infty$ 时，$I(\lambda)$ 的行为。

C11) 假设对任意 $x\in[a,b]$，$\varphi'(x)\neq 0$。我们定义映射

$$L:C^\infty([a,b])\to C^\infty([a,b]),\quad h\ \mapsto \frac{1}{i\lambda\varphi'(x)}h'(x),$$
$$M:C^\infty([a,b])\to C^\infty([a,b]),\quad h\mapsto -\Big(\frac{h}{i\varphi'}\Big)'(x),$$

假设 $g, h \in C^\infty([a,b])$。证明，如果 h 在 a 和 b 附近恒为 0（即存在 $c > 0$，使得当 $x \in [a, a+c] \cup [b-c, b]$ 时，$h(x) = 0$），那么

$$\int_a^b h \cdot Lg = \frac{1}{\lambda} \int_a^b g \cdot Mh.$$

C12) 假设对任意 $x \in [a,b]$，$\varphi'(x) \neq 0$ 并且 f 在 a 和 b 附近恒为 0。证明，对任意的 $n \in \mathbb{Z}_{\geqslant 1}$，都存在**不依赖**于 λ 的常数 c_n，使得

$$|I(\lambda)| \leqslant \frac{c_n}{\lambda^n}.$$

（提示：注意到 $Le^{i\lambda\varphi(x)} = e^{i\lambda\varphi(x)}$。）

C13) 假设存在 $\delta > 0$，使得对任意 $x \in [a,b]$，$|\varphi'(x)| \geqslant \delta$ 并且 $\varphi'(x)$ 是 $[a,b]$ 上的单调函数。证明，存在**不依赖**于 λ，φ，a 和 b 的常数 C_1，使得

$$\Big| \int_a^b e^{i\lambda \cdot \varphi(x)} dx \Big| \leqslant \frac{C_1}{\lambda\delta}$$

C14) 假设对任意 $x \in [a,b]$，$|\varphi''(x)| \geqslant 1$。证明，存在唯一的 $c \in [a,b]$，使得

$$|\varphi'(c)| = \inf_{x \in [a,b]} |\varphi'(x)|.$$

进一步证明，对任意的 $x \in [a,b]$，我们的都有

$$|\varphi'(x)| \geqslant |x - c|.$$

*C15) 假设对任意 $x \in [a,b]$，$|\varphi''(x)| \geqslant 1$。证明，存在**不依赖**于 λ，φ，a 和 b 的常数 C_2，使得

$$\Big| \int_a^b e^{i\lambda \cdot \varphi(x)} dx \Big| \leqslant \frac{C_2}{\sqrt{\lambda}}.$$

C16) 假设对任意 $x \in [a,b]$，$|\varphi''(x)| \geqslant 1$。证明，存在不依赖**于 λ，φ，f，a 和 b 的常数 C_2，使得

$$\Big| \int_a^b e^{i\lambda \cdot \varphi(x)} f(x) dx \Big| \leqslant \frac{C_2}{\sqrt{\lambda}} \Big(|f(a)| + \int_a^b |f'(x)| dx \Big).$$

28 一元微积分拾遗：Baire 纲定理及其应用，原函数的初等函数表示定理：Liouville 定理

二零一九年十二月二十三日，星期一，晴

Baire 定理以及应用

定理 160 (Baire)**.** (X,d) 是完备的距离空间，那么任意可数个稠密的开集的交仍然是稠密的，即若 $\{U_n\}_{n\geqslant 1}$ 是可数个稠密的开集，那么 $U_\infty=\bigcap\limits_{n\geqslant 1}U_n$ 是稠密的。

在第二课中，我们就已经给出了稠密性的定义：$U\subset X$ 是子集。如果对任意的 $x\in X$ 和任意的 $\varepsilon>0$，都存在 $y\in U$，使得 $d(y,x)<\varepsilon$，我们就称 U 在 X 中是稠密的。我们注意到，U 未必是开集，因为开集只在有限的交的操作下封闭。

证明: 任选 $x\in X$ 和 ε_0，我们将在 U_∞ 中找到一个点 x_∞，使得 $d(x_\infty,x)<2\varepsilon_0$。为此，我们将归纳地构造 X 中的点列 $\{x_n\}_{n\geqslant 1}$。

首先，根据 U_1 的稠密性，存在 $x_1\in U_1$，使得 $d(x_1,x)<\varepsilon_0$。再根据 U_1 是开集，我们可以找到 $\varepsilon_1>0$，使得 $\overline{B(x_1,2\varepsilon_1)}\subset U_1$，其中 $\overline{B(x_1,2\varepsilon_1)}$ 是闭球，即 $\overline{B(x_1,2\varepsilon_1)}=\big\{y\in X\big|d(y,x)\leqslant 2\varepsilon_1\big\}$。另外，通过缩小 ε_1，我们还可以要求 $2\varepsilon_1<\varepsilon_0$。

我们用 x_1 代替 x，用 ε_1 代替 ε_0，重复上面的过程：根据 U_2 的稠密性，存在 $x_2\in U_2$，使得 $d(x_2,x_1)<\varepsilon_1$。根据 U_2 是开集，我们可以找到 $\varepsilon_2>0$，使得 $\overline{B(x_2,2\varepsilon_2)}\subset U_2$ 并且可以进一步要求 $2\varepsilon_2<\varepsilon_1$。

重复以上过程，我们就得到了 $\{x_n\}_{n\geqslant 0}$ 和数列 $\{\varepsilon_n\}_{n\geqslant 0}$（其中 $x_0=x$，$\varepsilon_0=\varepsilon$），使得对任意的 $n\geqslant 0$，有

1) $x_{n+1}\in U_{n+1}$ 并且 $d(x_{n+1},x_n)<\varepsilon_n$；

2) $\overline{B(x_{n+1},2\varepsilon_{n+1})}\subset U_{n+1}$；

3) $0<2\varepsilon_{n+1}<\varepsilon_n$。

特别地，根据第三条，我们有 $\varepsilon_{n+k}<2^{-k}\varepsilon_n$。

我们现在说明 $\{x_n\}_{n\geqslant 1}$ 是 Cauchy 列。对任意的自然数 n 和 p，我们有

$$\begin{aligned}d(x_{n+p},x_n)&\leqslant d(x_{n+p},x_{n+p-1})+d(x_{n+p-2},x_{n+p-2})+\cdots+d(x_{n+1},x_n)\\&<\varepsilon_{n+p}+\varepsilon_{n+p-1}+\cdots+\varepsilon_{n+1}\\&<2^{-p-1}\varepsilon_n+2^{-p-2}\varepsilon_n+\cdots+\varepsilon_n\\&<2\varepsilon_n.\end{aligned}$$

这表明，对一切 $p\geqslant 0$，$\{x_{n+p}\}_{p\geqslant 1}$ 都落在 $\overline{B(x_n,2\varepsilon_n)}$ 中。我们注意到，上面的不等式直接给出

$$d(x_{n+p},x_n)<2^{-n+p}\varepsilon+2^{-n+p-1}\varepsilon+\cdots+2^{-n+1}\varepsilon=2^{-n}\varepsilon.$$

所以 $\{x_n\}_{n\geqslant 1}$ 是 Cauchy 列，根据 X 的完备性，存在 $x_\infty \in X$，使得 $\lim\limits_{n\to\infty} x_n = x_\infty$。特别地，根据 $\{x_{n+p}\}_{p\geqslant 1} \subset \overline{B(x_n, 2\varepsilon_n)}$，我们知道 $x_\infty \in \overline{B(x_n, 2\varepsilon_n)}$（这是闭集）。从而对任意的 $n \geqslant 1$，$x_\infty \in U_n$，所以，$x_\infty \in U_\infty$。特别地，在最后一个不等式中取 $n = 0$，我们有

$$d(x_p, x_0) < \varepsilon.$$

令 $p \to \infty$，我们就得到 $d(x_\infty, x) < 2\varepsilon$。 □

对于任意的集合 $Y \subset X$，如果对 $y \in Y$，存在 $\varepsilon > 0$，使得 $B(y, \varepsilon) \subset Y$，我们就称 y 是 Y 的一个**内点** 。我们用 $\mathring{Y}$ 表示 Y 的内点所组成的集合并称之为 Y 的**内部**。

练习. *假设 $Y, Z \subset X$ 互为补集，那么 $\mathring{Y} = \emptyset$ 等价于 Z 在 X 中稠密。*

推论 161. *(X, d) 是完备的距离空间，若 $\{F_n\}_{n\geqslant 1}$ 是可数个闭集并且对每个 $n \geqslant 1$，其内部 $\mathring{F}_n = \emptyset$，那么 $F_\infty = \bigcup\limits_{n\geqslant 1} F_n$ 的内部为空集，即 $\mathring{F}_\infty = \emptyset$。*

证明: 令 $U_n = X - F_n$，这是开集。那么，$\mathring{F}_n = \emptyset$ 意味着 U_n 是稠密的。根据 Baire 的定理，$U_\infty = \bigcap\limits_{n\geqslant 1} U_n$ 是稠密的，从而 $X - U_\infty$ 的内部为空集。最后，我们注意到 $F = \bigcup\limits_{n\geqslant 1} F_n$ 的补集恰好为 U_∞，所以命题成立。 □

我们给出 Baire 纲定理的几个应用。

命题 162. *不存在可微函数 $f \in C([a, b])$，使得对任意的开区间 $(c, d) \subset [a, b]$，f' 在 (c, d) 上无界。*

证明: 假设 f 是 $[a, b]$ 上的可微函数，对任意的 $n \geqslant 1$，定义

$$f_n(x) = \frac{f(x + \frac{1}{n}) - f(x)}{\frac{1}{n}}.$$

在上面的式子中，如果 $x + \frac{1}{n} > b$，我们可以考虑修改为 $f_n(x) = \dfrac{f(b) - f(x)}{\frac{1}{n}}$，不过这对证明没有影响。很明显，由于 f 是可微的，所以 $\{f_n\}_{n\geqslant 1} \subset C([a, b])$ 逐点收敛，即对于每个点 $x \in [a, b]$，$f'(x) = \lim\limits_{n\to\infty} f_n(x)$ 存在。

对于每个 $x \in [a, b]$，我们令 $M(x) = \sup\limits_{n\geqslant 1} |f_n(x)|$。因为 f' 的存在性，我们知道 $M(x)$ 是良好定义的。我们定义集合

$$F_k = \big\{x \in [a, b] \big| M(x) \leqslant k\big\} = \bigcap_{n\geqslant 1} \underbrace{\Big\{x \in [a, b] \Big| |f_n(x)| \leqslant k\Big\}}_{Fn,k}.$$

由于每个 $F_{n,k}$ 都是闭集，所以 $F_k \subset [a, b]$ 是闭集。又因为在每一个点处 $M(x) < \infty$，所以

$$[a, b] = \bigcup_{k\geqslant 1} F_k.$$

根据 Baire 定理的推论，不可能每一个 F_k 的内部都是空集，所以存在 $k_0 \geqslant 1$，使得

$$\mathring{F_{k_0}} \neq \emptyset.$$

所以，存在开区间 $(c,d) \subset [a,b]$，使得 $(c,d) \subset \big\{x \in [a,b] \big| M(x) \leqslant k_0\big\}$，即对任意的 $x \in (c,d)$，对任意的 $n \geqslant 1$，我们都有

$$|f_n(x)| = \Big|\frac{f(x+\frac{1}{n}) - f(x)}{\frac{1}{n}}\Big| \leqslant k_0.$$

令 $n \to \infty$，这表明在 (c,d) 上，我们有 f' 有界。 □

命题 163. *不存在 $[0,1]$ 上的连续函数序列 $\{f_n\}_{n\geqslant 1} \in C([0,1])$，使得 f_n 逐点收敛到 $\mathbf{1}_{\mathbb{Q}}$（即有理数的示性函数）。*

证明: 如若不然，我们假设存在 $\{f_n\}_{n\geqslant 1} \in C([0,1])$，使得对任意的 $x \in [0,1]$，我们都有

$$\lim_{n\to\infty} f_n(x) = \begin{cases} 1, & x \in \mathbb{Q}; \\ 0, & x \notin \mathbb{Q} \end{cases}$$

对于每个 $k \geqslant 1$，我们令

$$F_k = \bigcap_{n\geqslant k} f_n^{-1}([-\frac{1}{2}, \frac{1}{2}]).$$

由于 f_n 连续，所以

$$f_n^{-1}\big([-\frac{1}{2}, \frac{1}{2}]\big) = \big\{x \in [0,1] \big| -\frac{1}{2} \leqslant f_n(x) \leqslant \frac{1}{2}\big\}$$

是闭集，从而 F_k 是闭集。根据定义，我们知道对任意的 $x \in F_k$，

$$\lim_{n\to\infty} f_n(x) = 0.$$

从而，对于任意的 $x \in F_k$，x 是无理数。特别地，$\mathbb{Q} \cap F_k = \emptyset$，所以 $\mathring{F_k} = \emptyset$。

类似地，对于每个 $k \geqslant 1$，我们令

$$G_k = \bigcap_{n\geqslant k} f_n^{-1}([\frac{1}{2}, \frac{3}{2}]).$$

这也是闭集。根据定义，我们知道对任意的 $x \in G_k$，

$$\lim_{n\to\infty} f_n(x) = 1.$$

从而，对于任意的 $x \in G_k$，x 是有理数，所以 $\mathring{G_k} = \emptyset$。

然而，根据 f_n 逐点收敛到 $\mathbf{1}_{\mathbb{Q}}$，我们知道

$$[0,1] = \bigcup_{k\geqslant 1} (F_k \cup G_k) = F_1 \cup G_1 \cup F_2 \cup G_2 \cup \cdots.$$

上式左边的集合的内部非空，这和 Baire 定理矛盾。 □

练习. 试构造函数 $\{f_{m,n}\}_{m,n\geqslant 1}$，使得对任意的 $x\in[0,1]$，我们有

$$\lim_{m\to\infty}\Big(\lim_{n\to\infty} f_{m,n}(x)\Big)=\mathbf{1}_{\mathbb{Q}}(x).$$

Baire 定理在线性代数上有如下的应用：

命题 164. 假设 $(V,\|\cdot\|)$ 是完备的赋范线性空间（不妨假设是实数域上的线性空间），如果 $\dim V=\infty$，那么 V 的任意一组（代数）基 $\{e_i\}_{i\in I}$ 都是不可数集。

我们回忆一下，如果 $\{e_i\}_{i\in I}$ 是 V 的一组（代数）基，那么对任意的 $v\in V$，存在有限个指标 $i_1,i_2,\cdots,i_m\in I$ 以及有限个实数 $\lambda_{i_1},\lambda_{i_2},\cdots,\lambda_{i_m}$，使得

$$v=\lambda_{i_1}e_{i_1}+\lambda_{i_2}e_{i_2}+\cdots+\lambda_{i_m}e_{i_m}.$$

为了证明这个命题，我们不加证明地接受如下的事实：

有限维的赋范线性空间是完备的。

上面的陈述是泛函分析中的标准事实，证明实际上很简单，我们在下个学期学起多元微积分的时候（应该）会证明。

证明: 如若不然，存在 $\{e_i\}_{i=1,2,\cdots}$ 是 V 的一组可数的基，那么对任意的 $k\geqslant 1$，定义 V 的子空间

$$V_k=\operatorname{span}\{e_1,e_2,\cdots,e_k\}.$$

我们在 V_k 上用 $\|\cdot\|$ 所诱导的度量，从而 $(V_k,\|\cdot\|)$ 是完备的。特别地，我们知道 $V_k\subset V$ 是闭集，这是因为 V_k 中的收敛的序列都在 V_k 中收敛（连续性）。我们现在说明对任意的 k，V_k 的内部为空集，实际上，假设 $v\in V$，那么任意选取 $\varepsilon>0$，我们有 $v+\varepsilon e_{k+1}\notin V_k$，这表明 v 不是 V_k 的内点。根据基的定义，我们自然有

$$V=\bigcup_{k\geqslant 1}V_k.$$

然而 V 的内部非空，这和 Baire 的定理矛盾。 □

注记. 考虑 $\Big(C\big([a,b]\big),\|\cdot\|_\infty\Big)$ 的子空间 D，它是由在至少一点处可微的函数构成的，即

$$D=\Big\{f\in C\big([a,b]\big)\Big|\text{存在 } x\in[a,b],\ f \text{ 在 } x \text{ 处可微}\Big\}.$$

那么，利用稍微精细一点的分析，我们可以证明 $\mathring{D}=\emptyset$。这说明我们有很多的处处连续处处不可微分的函数。请有兴趣的同学自己参考网络或者有关的文献。

原函数的初等表示

我们现在来说明不能找到一个初等函数 f，使得 $f'=e^{x^2}$。为此，我们先引入一些记号，当然，如果同学们学习了代数学中的域理论，这里的很多定义就会很自然。任意给定函数 $f_1,\cdots,f_n$，我们令

$$\mathbb{C}(f_1,\cdots,f_n)=\left\{\frac{\displaystyle\sum_{0\leqslant i_1,i_2,\cdots,i_n\leqslant k} a_{i_1,i_2,\cdots,i_k}(f_1)^{i_1}(f_2)^{i_2}\cdots(f_k)^{i_k}}{\displaystyle\sum_{0\leqslant j_1,j_2,\cdots,j_n\leqslant \ell} b_{j_1,j_2,\cdots,j_\ell}(f_1)^{j_1}(f_2)^{j_2}\cdots(f_\ell)^{j_\ell}}\middle| k,\ell\in\mathbb{Z}_{\geqslant 0}, a_{i_1,i_2,\cdots,i_k}, b_{j_1,j_2,\cdots,j_\ell}\in\mathbb{C}\right\}.$$

（其中，$\mathbb{C}(X)$ 是有理函数的集合，即两个多项式的商）换句话说，这是由 $f_1,\cdots,f_n$ 经过有限次四则运算所得到的所有可能的函数的集合（可以用上面的形式对四则运算的次数进行归纳即可）。特别地，根据定义，我们知道 $\mathbb{C}(f_1,\cdots,f_n)$ 在四则运算下是封闭的。

很明显，我们有如下的包含关系：

$$\mathbb{C}\subset\mathbb{C}(x)\subset\mathbb{C}(x,f_1)\subset\cdots\subset\mathbb{C}(x,f_1,\cdots,f_k)\subset\mathbb{C}(x,f_1,\cdots,f_k,f_{k+1})\subset\cdots\subset\mathbb{C}(x,f_1,\cdots,f_n).$$

假设对于任意的 k，上面的从 $\mathbb{C}(x,f_1,\cdots,f_k)$ 到 $\mathbb{C}(x,f_1,\cdots,f_k,f_{k+1})$ 所新添进去的元素 f_k 满足如下的关系之一：

1) $f_{k+1}=e^f$，其中 $f\in\mathbb{C}(x,f_1,\cdots,f_k)$；

2) $f_{k+1}=\log(f)$，其中 $f\in\mathbb{C}(x,f_1,\cdots,f_k)$；

3) 存在 $\mathbb{C}(x,f_1,\cdots,f_k)$ 上的多项式 $P(X)$ 使得 $P(f)=0$，即存在 $m\geqslant 1$，$\alpha_0,\cdots,\alpha_m\in\mathbb{C}(x,f_1,\cdots,f_k)$，使得

$$\alpha_m(f_{k+1})^m+\alpha_{m-1}(f_{k+1})^{m-1}+\cdots+\alpha_1 f_{k+1}+\alpha_0=0.$$

我们就称 $\mathbb{C}(f_1,\cdots,f_n)$ 中的元素是**初等函数**，称 $\mathbb{C}(x,f_1,\cdots,f_n)$ 是一个**初等函数域**。很明显，从函数 1 和 x 出发，由指数函数，对数函数，取多项式函数的根以及四则运算进行有限次复合所得到的函数都是初等函数。比如说，e^{x^2} 和 $\big(\log x\big)^{-1}$ 这两个函数很明显是初等函数，我们想知道是否它们的原函数是否可以用初等函数表达。

命题 165. *假设 $K=\mathbb{C}(x,f_1,\cdots,f_n)$ 是初等函数域，那么求导数运算 $\dfrac{d}{dx}$ 将 K 中的元素映射成 K 中的元素，即*

$$\frac{d}{dx}:K\to K.$$

（代数上，我们把这样的域称作是一个微分域）

证明: 对 n 进行归纳即可，当 $n=0$ 时命题明显成立。假设对 $n\geqslant 0$ 命题成立，现在只要说明 $(f_{n+1})'\in K=\mathbb{C}(x,f_1,\cdots,f_n)$ 即可。我们考虑 f_{n+1} 的构造，分情况讨论：

1) 若 $f_{n+1}=e^f$，其中 $f\in\mathbb{C}(x,f_1,\cdots,f_n)$，那么 $(f_{n+1})'=f'e^f=f'f_{n+1}$。然而，根据归纳法，$f'\in\mathbb{C}(x,f_1,\cdots,f_n)\subset K$，所以利用四则运算的封闭性，我们有 $f'f_{n+1}\in K$；

2) $f_{n+1}=\log(f)$，其中 $f\in\mathbb{C}(x,f_1,\cdots,f_n)$，证明完全类似于上面的情形。

3) 存在 $m\geqslant 1$，$\alpha_0,\cdots,\alpha_m\in\mathbb{C}(x,f_1,\cdots,f_k)$，使得

$$\alpha_m(f_{k+1})^m+\alpha_{m-1}(f_{k+1})^{m-1}+\cdots+\alpha_1 f_{k+1}+\alpha_0=0.$$

我们可以要求 m 时最小的使得上面式子成立的正整数。对上式求导数，我们得到

$$\overbrace{(\alpha_m)'(f_{k+1})^m+(\alpha_{m-1})'(f_{k+1})^{m-1}+\cdots+(\alpha_1)'f_{k+1}+(\alpha_0)'}^{\text{根据归纳假设，}\alpha_i\in K\text{，所以这是 }K\text{ 中的元素}}$$
$$=-\underbrace{\left(m\alpha_m(f_{k+1})^{m-1}+(m-1)\alpha_{m-1}(f_{k+1})^{m-2}+\cdots+\alpha_1\right)}_{\text{这是 }K\text{ 中的元素，记作 }F}\left(f_{k+1}\right)'.$$

根据 m 的选取，我们知道 F 不是零。所以，通过除法我们就证明了 $(f_{n+1})'\in K$。

一般的 $f\in K$ 的导数通过 Leibniz 法则立即可以得到。 □

我们现在承认如下 Liouville 定理（这是纯代数的结果，请参见初等的介绍：M. Rosenlicht, *Integration in finite terms*, American Math. Monthly 79 (1972), 963–972）：

定理 166 (Liouville)**.** *假设 K 是一个初等函数域，$f\in K$，那么存在初等函数 F 使得 $F'=f$ 当且仅当存在常数 $c_1,\cdots,c_m\in\mathbb{C}$，存在函数 $R_0,R_1,\cdots,R_m\in K$，使得*

$$f=R_0'+\sum_{k=1}^m c_k\frac{R_k'}{R_k}.$$

这个条件自然是充分的，因为

$$f=\left(R_0+\sum_{k=1}^m c_k\log(R_k)\right)'.$$

必要性的证明请参见上述 Rosenlicht 的短文。Louville 定理有如下的有用推论：

推论 167. *假设 $f,g\in\mathbb{C}(X)$ 是有理函数，那么函数 $f(x)e^{g(x)}$ 具有初等的原函数当且仅当存在有理函数 $R\in\mathbb{C}(X)$ 使得*

$$R'+g'R=f.$$

证明： 充分性是明显的，因为如果 $R'+g'R=f'$，那么

$$\left(R\cdot e^f\right)'=(R'+g'R)e^g=fe^g.$$

现在假设 $f(x)e^{g(x)}$ 具有初等的原函数，我们在 $K=\mathbb{C}(x,e^f)$ 中工作，根据 Liouvile 的定理，我们有

$$f(x)e^{g(x)}=R_0'(x,e^{g(x)})+\sum_{k=1}^m c_k\frac{R_k'(x,e^{g(x)})}{R_k(x,e^{g(x)})},$$

其中 $R_k(X,Y)=\dfrac{P_k(X,Y)}{Q_k(X,Y)}$，这里 P_k 和 Q_k 是 $\mathbb{C}$-系数的二元多项式。给定一个二元的多项式 $P(X,Y)$，我们通过将它写成

$$P(X,Y)=p_n(X)Y^n+\cdots+p_1(X)Y+p_0(X)$$

可以将它看作是系数在 $\mathbb{C}(X)$ 中的多项式，这是一个域（和实数一样满足四则运算法则），所以可以将 $P(X,Y)$ 写成 $\mathbb{C}(X)[Y]$ 中的不可约分多项式的分解。据此，对 $k\geqslant 1$，我们可以将 $R_k(X,Y)$ 写成

$$R_k(X,Y)=r_k(X)\frac{\prod p_{k,i}(X,Y)}{\prod q_{k,j}(X,Y)},$$

其中 $p_{k,i}(X,Y)$ 和 $q_{k,j}(X,Y)$ 都是首一的 $\mathbb{C}(X)[Y]$ 中的不可约多项式，即形如

$$Y^m+a_{n-1}(X)Y^{m-1}+\cdots+a_0(x),$$

其中 $r_k(X)$ 和 $a_i(X)$ 均为 $\mathbb{C}[X]$ 中的元素。所以，

$$\frac{R_k'(x,e^{g(x)})}{R_k(x,e^{g(x)})}=\big(\log(R_k(x,e^{g(x)}))\big)'=r_k(x)+\sum\frac{p_{k,i}'(x,e^{g(x)})}{p_{k,i}(x,e^{g(x)})}-\sum\frac{q_{k,j}'(x,e^{g(x)})}{q_{k,j}(x,e^{g(x)})}.$$

据此，我们不妨假设所有的 $R_k(X,Y)$ 都是首一的 $\mathbb{C}(X)[Y]$ 中的不可约多项式，$R(X,Y)=\dfrac{P(X,Y)}{Q(X,Y)}$，其中 $P(X,Y)$ 和 $Q(X,Y)$ 也是首一的不可约多项式（可以对 f 乘一个系数来做到这一点）。按照要求，我们有

$$f(X)Y-D(R_0(X,Y))+\sum_{k=1}^{m}c_k\frac{D(R_k(X,Y))}{R_k(X,Y)}\Big|_{X=x,Y=e^g}=0.$$

其中

$$D(R_k(X,Y))=D(\sum r(X)Y^\ell)=\sum\big(r(X)'+kr(X)g'(X)\big)Y^\ell.$$

我们现在说明，这个等式意味着

$$f(X)Y-D(R_0(X,Y))+\sum_{k=1}^{m}c_k\frac{D(R_k(X,Y))}{R_k(X,Y)}=0.$$

实际上，我们只需要说明

$$P(x,e^{g(x)})=0\quad\Rightarrow\quad P(X,Y)=0,$$

其中 $P(X,Y)\in\mathbb{C}(X)[Y]$ 是首一多项式，$g(x)$ 不是常数。为此，假设

$$P(X,Y)=Y^n+\cdots+r_1(X)Y+r_0(X).$$

我们对 n 进行归纳。$n=0$ 是显然的。对一般的 n，我们将上面的式子写成

$$e^{ng(x)}+r_{n-1}(x)e^{(n-1)g(x)}+\cdots+r_1(x)e^{g(x)}+r_0(x)=0.$$

求导数，我们就有

$$ng'(x)e^{ng(x)}+\sum_{k=0}^{n-1}\big(r_k'(x)+kr_k(x)g'(x)\big)e^{kg(x)}=0.$$

所以，

$$\sum_{k=0}^{n-1}\big(r_k'(x)+kr_k(x)g'(x)-ng'(x)r_k(x)\big)e^{kg(x)}=0.$$

由归纳假设，我们有

$$\sum_{k=0}^{n-1}\big(r_k'(X)+(k-n)r_k(X)g'(X)\big)Y^k=0.$$

对于 $k=0$ 的系数我们有

$$g'(X)=\frac{r_k'(X)}{nr_k(X)}=\sum_{i=1}^{m}\frac{1}{X-a_i}.$$

对于任意的一个有理函数 $g(X)$，我们总可以将它写成

$$g(X)=\sum_{k=1}^{M}(X-a_k)^{b_k}$$

的形式，其中 $b_k\in\mathbb{Z}$，它的导数不可能是上述的形式，矛盾。

我们现在来分析

$$f(X)Y-\frac{D(P(X,Y))Q(X,Y)-P(X,Y)D(P(X,Y))}{Q(X,Y)^2}+\sum_{k=1}^{m}c_k\frac{D(R_k(X,Y))}{R_k(X,Y)}=0\quad\cdots\cdots\quad(\star)$$

中多项式的整除关系。先研究 R_k，其中 $k\geqslant 1$。按照定义，我们有

$$\begin{aligned}D(R_k(X,Y))&=D(Y^N+\sum_{\ell<N}r(X)Y^\ell)\\&=Ng'(X)Y^N+\sum_{\ell<N}\big(r(X)'+kr(X)g'(X)\big)Y^\ell.\end{aligned}$$

它的 Y-次数和 $R_k(X,Y)$ 的一致。所以，如果要分母上的 $D(R_k(X,Y))$ 与分子上的不可约多项式 $R_k(X,Y)$ 能约分的话，只能有

$$D(R_n(X,Y))=Ng'(X)R_k(X,Y).$$

于上面完全一致，我们再次比较最低项的次数就得到矛盾。所以，每个 $\dfrac{D(R_k(X,Y))}{R_k(X,Y)}$ 都是不可约分的，为了消去 $R_k(X,Y)$ 的分母，我们只能寄希望于有这样的分母来自于 $\dfrac{D(P)Q-PD(Q)}{Q^2}$ 这一项，然而，如果假设 $Q(X,Y)=R_k(X,Y)^s\widetilde{Q}(X,Y)$，其中 $s\in\mathbb{Z}_{\geqslant 1}$，$\widetilde{Q}(X,Y)$（以及 $P(X,Y)$）

和 $R_k(X,Y)$ 互素，那么

$$\begin{aligned}\frac{D(P)Q-PD(Q)}{Q^2} &= \frac{D(P)}{R_k^s\widetilde{Q}} - \frac{P}{R_k^{2s}\widetilde{Q}^2}\big(sR_k^{s-1}D(R_k)\widetilde{Q}+R_k^sD(\widetilde{Q})\big)\\ &= \underbrace{\frac{D(P)}{R_k^s\widetilde{Q}} - \frac{PD(\widetilde{Q})}{R_k^s\widetilde{Q}^2}}_{\text{分母上贡献了 } R_k^t \text{ 的项，其中 } t\leqslant s} - \frac{sP\widetilde{Q}D(R_k)}{R_k^{s+1}}\end{aligned}$$

此时，最后一项在分母上贡献了 R_k 的因子是 $s+1\geqslant 2$ 是最高的，这是不能消去的。所以为了使得 $(\star)$ 成立，我们只能有

$$f(X)Y - \frac{D(P(X,Y))Q(X,Y)-P(X,Y)D(Q(X,Y))}{Q(X,Y)^2} = 0 \quad \cdots\cdots \quad (\star)$$

同理，$Q(X,Y)$ 也不能有不可约因子，所以 $Q(X,Y)\in\mathbb{C}(X)$，它可以被吸收到 $P(X,Y)$ 中，从而假设 $Q(X,Y=1)$，据此，我们有

$$f(X)Y = D(P(X,Y)) = \sum_{\ell\leqslant N}\big(r(X)'+\ell r(X)g'(X)\big)Y^\ell.$$

比较 1 次项系数，我们得到

$$f(X) = r(X)'+r(X)g'(X).$$

这就证明了结论。 □

作为应用，我们有

1) $\displaystyle\int e^{x^2}$ 不是初等函数。

如若不然，此时 $f=1$，$g=x^2$，所以存在有理函数 $R(x)=\dfrac{P(x)}{Q(x)}$，其中 $P(x)$ 与 $Q(x)$ 互素并且 $Q(x)$ 是首一的，使得

$$1 = R(x)'+2xR(x) \quad\Leftrightarrow\quad P'Q+2xPQ-Q^2 = PQ'.$$

这表明 Q 整除 PQ'，然而，P 与 Q 互素，所以 Q 整除 Q'，这当然是不可能的，除非 $Q(x)\equiv 1$，此时，$R(X)$ 是多项式，$1=R(x)'+2xR(x)$ 自然不对（看次数）。

2) $\displaystyle\int\frac{1}{\log(x)}$ 不是初等函数。

通过变量替换 $x=e^y$，这等价于证明 $\displaystyle\int\frac{e^x}{x}$ 不是初等函数。

如若不然，此时 $f=x^{-1}$，$g=x$，所以存在有理函数 $R(x)=\dfrac{P(x)}{Q(x)}$，其中 $P(x)$ 与 $Q(x)$ 互素并且 $Q(x)$ 是首一的，使得

$$x^{-1} = R(x)'+R(x) \quad\Leftrightarrow\quad xP'Q+xPQ-Q^2 = xPQ'.$$

很明显，R 不是多项式，即 $\deg Q \geqslant 1$。上面的式子表明 Q 整除 xPQ'，然而，Q 与 PQ' 互素，所以 Q 整除 x，从而 $Q(x) = x$。代入上面的等式，我们得到

$$x^2P'Q + x^2P - Q^2 = xP.$$

这表明 x 整除 Q^2，这和 $P(x)$ 与 $Q(x)$ 互素相矛盾。

练习. 证明，$\displaystyle\int \frac{\sin x}{x}$ 不是初等函数。

29 一元微积分拾遗：振荡与衰减，Riemann-Lebesgue 引理，van der Corput 型估计

二零一九年十二月二十六日，星期四，晴

修正： 上次课中关于 Liouville 对于用初等函数表示原函数的定义中，我们应该要求所考虑的函数是所谓的**亚纯函数**，这是一个复变函数中的概念。否则的话，有同学举出了反例子：$\mathbf{1}_{\mathbb{Q}}$ 满足有理函数域上的代数方程 $f^2+f=0$，但是它没有初等原函数。但是从第一次接触和了解的角度来看，我们就不再做进一步解释。

振荡与衰减

这是我们课程的最后一个部分，我们谈论三个相关的话题：1）Riemann-Lebesgue 引理 2）van der Corput 型估计 3）某些非绝对收敛的反常积分的收敛判断（请参考第十二次习题关于 Abel 和 Dirichlet 判别法，我们不再赘述）。这三个话题统一的地方是所考虑的函数“振荡”得很快，从工具方面而言，我们要通过分部积分来利用函数振荡得很快所以正负部分可能相消这个事实。

Riemann-Lebesgue 引理

我们先从 Riemann-Lebesgue 引理入手。这大概是分析中最重要的几个定理之一，无论是定理的含义与证明都值得好好理解：

定理 168 (Riemann-Lebesgue 引理)**.** $[a,b]\subset\mathbb{R}$ 是有界闭区间，$f\in C([a,b])$，对任意的 $n\geqslant 1$，令

$$A_n=\int_a^b f(x)\sin(nx)dx.$$

那么，

$$\lim_{n\to\infty}A_n=0.$$

与积分的第二中值定理类似，可以先证明一个比较弱的版本：我们假设 f 是连续可微的，不妨假设 $f\in C^k([a,b])$ 其中 $k\geqslant 1$。

在开始证明之前，先谈论一下所谓的“直观”，当 n 很大的时候，我们注意到函数 $\sin(nx)$ 在 $[a,b]$ 上振荡的很厉害。直观上，我们可以假设在局部上 $f(x)$ 这个函数基本上是常数，所以 $\sin(nx)$ 有正有负地振荡，我们期望正负可以相消，从而对积分的贡献很小。但是如何从分析上来说明“正负相消”这个事情呢？我们考虑 $\sin(nx)$ 在任意一个区间 $[c,d]$ 上的积分

$$\int_c^d\sin(nx)dx.$$

如果用三角不等式来估计这个积分，我们就有

$$\Big|\int_c^d\sin(nx)dx\Big|\leqslant\int_c^d|\sin(nx)|dx\leqslant\int_c^d 1dx=(d-c).$$

当区间长度变大的时候，右边的界也随之变大。然而，我们可以用“相消性”，实际上就是直接计算这个积分：

$$\int_c^d \sin(nx)dx = -\frac{1}{n}\int_c^d \big(\cos(nx)\big)'dx = -\frac{1}{n}\cos(nx)\Big|_{x=c}^{x=d} = O(\frac{1}{n}).$$

这个估计很明显比之前的好很多，在这里，我们通过计算它的原函数或者说通过 Newton-Leibniz 公式／分部积分，用到了 sin 振荡的性质。从某种意义上说，一个振荡函数的原函数的大小（≈ 积分）包含了函数振荡的信息。

根据上面的讨论，我们先假设假设 $f \in C^k([a,b])$ 其中 $k \geqslant 1$ 来证明 Riemann-Lebesgue 引理：

$$\begin{aligned} A_n = \int_a^b f(x)\sin(nx)dx &= -\frac{1}{n}\int_a^b f(x) \underbrace{\big(\cos(nx)\big)'}_{\text{振荡的函数的原函数应该增长不快}} dx \\ &= \frac{1}{n}\underbrace{\int_a^b f'(x)\cos(nx)dx}_{\leqslant \int_a^b |f'|} + \frac{1}{n}f(a)\cos(na) - \frac{1}{n}f(b)\cos(nb). \end{aligned}$$

从而，我们有

$$|A_n| \leqslant \frac{1}{n}\left(\int_a^b |f'| + 2\right).$$

这说明 $A_n \to 0$。特别地，我们对 A_n 衰减的速度还有估计（$O(\frac{1}{n})$）。

我们注意到上面的证明中我们只用到了 f 的一阶导数。我们进一步假设 f 在 $[a,b]$ 端点处都消失，即有一个 $\delta > 0$，使得 $f\big|_{[a,a+\delta]} \equiv 0$ 并且 $f\big|_{[b-\delta,b]} \equiv 0$，此时，我们注意到分部积分的时候，边界项都是零，所以

$$A_n = \frac{1}{n}\int_a^b f'(x)\cos(nx)dx.$$

现在 A_n 的形式和之前是类似的，只是 f 变为 f'，sin 变为 cos，所以重复上面的过程，我们可以进行 k 次分部积分，得到

$$\begin{aligned} A_n &= \pm\frac{1}{n^2}\int_a^b f''(x)\sin(nx)dx \\ &= \pm\frac{1}{n^3}\int_a^b f'''(x)\cos(nx)dx \\ &= \cdots = \pm\frac{1}{n^k}\int_a^b f^{(k)}(x)\underbrace{\sin(nx)}_{\text{或}\cos(nx)}dx. \end{aligned}$$

我们注意到，每次分部积分我们都多获得了一个 n^{-1} 的衰减。最终，如果 $f \in C^k([a,b])$ 并且 f 在 a 和 b 附近消失，我们证明了 $|A_n| = O(\frac{1}{n^k})$。

注记. 在上面证明过程中，f 每（通过分部积分）损失一个导数，就获得了 $\frac{1}{n}$ 的衰减。这种通过牺牲导数来换取衰减的机制是下学期课程 *Fourier* 分析的部分中（包括之后在调和分析或者偏微分方程的课程中）最重要的机制之一。

我们回到 Riemann-Lebesgue 定理的证明：

证明：我们用光滑（C^1）函数来逼近连续函数，这个想法值得好好学习。

假设 $f \in C([a,b])$，我们要证明对于任意的 $\varepsilon > 0$，存在 N，使得当 $n \geqslant N$ 时，我们有 $|A_n| < \varepsilon$。实际上，根据 Weierstrasss-Stone 定理，存在多项式 $P(x)$，使得

$$\sup_{x\in[a,b]} |P(x) - f(x)| < \frac{1}{2}\frac{\varepsilon}{b-a}.$$

由于 $P(x) \in C^1([a,b])$，根据上面的讨论，我们有

$$|\int_a^b P(x)\sin(nx)dx| \leqslant \frac{1}{n}\left(\int_a^b |P'| + 2\right).$$

任意选取 N，使得 $N > \dfrac{2\left(\int_a^b |P'| + 2\right)}{\varepsilon}$，从而当 $n \geqslant N$ 时，我们有

$$|\int_a^b P(x)\sin(nx)dx| \leqslant \frac{\varepsilon}{2}.$$

特别地，我们有

$$\begin{aligned}\left|\int_a^b f(x)\sin(nx)dx\right| &\leqslant \left|\int_a^b P(x)\sin(nx)\right| + \int_a^b \big|f(x) - P(x)\big||\sin(nx)|dx \\ &< \frac{\varepsilon}{2} + \frac{1}{2}\frac{\varepsilon}{b-a}\int_a^b = \varepsilon.\end{aligned}$$

这就证明了命题。 □

我们可以将 Riemann-Lebesgue 引理加强一些：

定理 169 (Riemann-Lebesgue 引理)**.** $[a,b] \subset \mathbb{R}$ 是有界闭区间，$f \in \mathcal{R}([a,b])$，对任意的 $n \geqslant 1$，令

$$A_n = \int_a^b f(x)\sin(nx)dx.$$

那么，

$$\lim_{n\to\infty} A_n = 0.$$

证明的想法与之前的如出一辙，唯一的不同在于我们用阶梯函数／简单函数来逼近 f。先假设 $f \in \mathcal{E}([a,b])$，根据线性，我们只要对 $f = \mathbf{1}_{[c,d]}$ 来证明即可，而我们在课程的一开始就已经证明了

$$\int_a^b \mathbf{1}_{[c,d]} f = \int_c^d \sin(nx)dx = O(\frac{1}{n}).$$

对于一般的函数，按照 Riemann 积分的定义（请注意我们采取简单函数逼近的定义，下个学期我们抽象积分的定义也采取这种方式，从而在 Lebesgue 积分的范畴里 Riemann-Lebesgue 引理的证

明也是一样的）存在阶梯函数 φ 使得 $\int_a^b |\varphi - f| < \varepsilon$。所以，我们有

$$\begin{aligned}\Big|\int_a^b f(x)\sin(nx)dx\Big| &\leqslant \Big|\int_a^b \varphi(x)\sin(nx)\Big| + \int_a^b \big|f(x)-\varphi(x)\big||\sin(nx)|dx \\ &\leqslant \underbrace{\Big|\int_a^b \varphi(x)\sin(nx)\Big|}_{\text{阶梯函数的极限}=0} + \underbrace{\int_a^b \big|f(x)-\varphi(x)\big|dx}_{<\varepsilon}.\end{aligned}$$

这就证明了命题。

van der Corput 估计

给定两个函数 $\phi(x)$ 和 $a(x)$，其中 $\phi(x)$ 被称作是**相函数**，$a(x)$ 被称作是**振幅函数**，它们满足

1) $\phi(x)$ 和 $a(x)$ 都是光滑函数；

2) $a(x)$ 具有紧支集，即存在有界闭区间 $I=[-M,M]$，使得 $a(x)$ 在 I 之外都是 0。

我们要研究如下振荡积分 $I(\lambda)$ 的衰减：

$$I(\lambda) = \int_{\mathbb{R}} e^{i\lambda\phi(x)}a(x)dx.$$

根据 a 具有紧支集的条件，上面的积分实际上是在闭区间 I 上计算的。仿照前面关于 Riemann-Lebesgue 引理的讨论，我们有

命题 170. *（非驻相法 (nonstationary phase)）假设对任意的 $x\in I$，$|\phi'(x)|\geqslant 1$。那么，对任意的 $N\in\mathbb{Z}_{>0}$，存在常数 N，使得*

$$I(\lambda)| \leqslant \frac{C_N}{\lambda^N}.$$

证明： 由于 $|\phi'(x)|\neq 0$，所以我们可以考虑算子 $L\big(F(x)\big) = \dfrac{1}{i\lambda\phi'(x)}\dfrac{d}{dx}F(x)$。直接计算，我们就得到

$$L\left(e^{i\lambda\phi(x)}\right) = e^{i\lambda\phi(x)} \quad\Rightarrow\quad L^N\left(e^{i\lambda\phi(x)}\right) = e^{i\lambda\phi(x)}.$$

所以，我们有

$$\begin{aligned}I(\lambda) &= \int_{\mathbb{R}} \frac{a(x)}{i\lambda\phi'(x)}\left(e^{i\lambda\phi(x)}\right)'dx \\ &\overset{\text{分部积分}}{=} -\frac{1}{i\lambda}\int_{\mathbb{R}}\left(\frac{a(x)}{\phi'(x)}\right)' e^{i\lambda\phi(x)}dx \\ &= -\frac{1}{i\lambda}\int_{\mathbb{R}} \underbrace{\frac{a'(x)\phi'(x)-a(x)\phi''(x)}{(\phi'(x))^2}}_{=a_1(x)} e^{i\lambda\phi(x)}dx\end{aligned}$$

其中，在分部积分的时候，边界项一定为零。由此可见，我们只要用 $a_1(x)$ 来代替以前的 $a(x)$，就可以继续进行上面的分部积分的操作。从而，在进行了 N 次之后，我们就得到了 λ^{-N} 的衰减的因子。命题中要求的参数自然由 a 和 ϕ 以及它们的导数的大小决定。 □

我们考虑一种特殊的振荡积分：

$$J(\lambda)=\int_a^b e^{i\lambda\phi(x)}dx.$$

也就是说，我们用 $\mathbf{1}_{[a,b]}$ 作为振幅函数。此时，我们有

命题 171. *假设对任意的 $x\in I$，$|\phi'(x)|\geqslant 1$。如果 $\phi'(x)$ 在 $[a,b]$ 上是单调的，那么，*

$$\big|J(\lambda)\big|\leqslant\frac{3}{\lambda}.$$

练习. *如果在上面的条件中我们要求 $|\phi'(x)|\geqslant\delta$，那么结论要相应地变成*

$$\big|J(\lambda)\big|\leqslant\frac{3}{\delta\lambda}.$$

证明： 按照之前的讨论，我们有

$$\begin{aligned}J(\lambda)&=\int_a^b\frac{1}{i\lambda\phi'(x)}\left(e^{i\lambda\phi(x)}\right)'dx\\&\overset{\text{分部积分}}{=}\underbrace{-\frac{1}{i\lambda}\int_a^b\left(\frac{1}{\phi'(x)}\right)'e^{i\lambda\phi(x)}dx}_{\mathbf{I}}+\underbrace{\frac{1}{i\lambda\phi'(x)}e^{i\lambda\phi(x)}\Big|_a^b}_{|\cdot|\leqslant\frac{2}{\lambda}}\end{aligned}$$

现在只需要控制第一项，这里用到的核心事实是 $\phi'(x)$ 是单调的，从而 $\left(\frac{1}{\phi'(x)}\right)'$ 有确定的符号，所以

$$\begin{aligned}|\mathbf{I}|&\leqslant\frac{1}{\lambda}\int_a^b\Big|\left(\frac{1}{\phi'(x)}\right)'\Big|dx=\frac{1}{\lambda}\Big|\int_a^b\left(\frac{1}{\phi'(x)}\right)'dx\Big|\\&=\frac{1}{\lambda}\Big|\frac{1}{\phi'(a)}-\frac{1}{\phi'(b)}\Big|\leqslant\frac{1}{\lambda}.\end{aligned}$$

最后一个不等号是因为 $\frac{1}{\phi'(a)}$ 和 $\frac{1}{\phi'(b)}$ 的符号相同，这就证明了结论。 □

命题 172. *假设对任意的 $x\in I=[a,b]$，$|\phi''(x)|\geqslant 1$。那么，*

$$\big|J(\lambda)\big|\leqslant\frac{2\sqrt{3}}{(\lambda)^{\frac{1}{2}}}.$$

证明： 不妨假设 $\phi''(x)\geqslant 1$，那么，$\phi'(x)$ 是单调递增的函数。我们假设存在 $c\in[a,b]$ 使得 $\phi'(c)=0$。此时，我们将 $[a,b]$ 分成三个区间

$$I=I_1\cup I_2\cup I_3=[a,c-\delta]\cup[c-\delta,c+\delta]\cup[c+\delta,b].$$

此时，根据 $\phi''(x)\geqslant 1$ 以及 Newton-Leibniz 公式，我们知道在 I_1 和 I_3 上，我们有 $|\phi'(x)|\geqslant\delta$。根据前面的命题，我们有

$$\Big|\int_{I_1}e^{i\lambda\phi(x)}dx\Big|\leqslant\frac{3}{\delta\lambda},\quad\Big|\int_{I_3}e^{i\lambda\phi(x)}dx\Big|\leqslant\frac{3}{\delta\lambda}.$$

在 I_2 上，我们用平凡的估计：

$$\Big|\int_{I_2} e^{i\lambda\phi(x)}dx\Big| \leqslant \int_{I_2} 1dx = 2\delta.$$

所以，

$$\begin{aligned}|J(\lambda)| &\leqslant \Big|\int_{I_1} e^{i\lambda\phi(x)}dx\Big| + \Big|\int_{I_2} e^{i\lambda\phi(x)}dx\Big| + \Big|\int_{I_3} e^{i\lambda\phi(x)}dx\Big| \\ &= 2\delta + \frac{6}{\delta\lambda}.\end{aligned}$$

当 $\delta = \sqrt{\dfrac{3}{\lambda}}$，上面的表达式取得最小值，这恰好是要证明的结论。 □

练习. *在上面的证明中，如果对任意的 $c \in [a,b]$，$\phi'(c) \neq 0$，试证明同样的结论成立。*

命题 173 (van der Corput, $k \geqslant 2$)**.** *假设对任意的 $x \in I = [a,b]$，$|\phi^{(k)}(x)| \geqslant 1$。那么，*

$$\big|J(\lambda)\big| \leqslant \frac{2^k}{(\lambda)^{\frac{1}{k}}}.$$

注记. *van der Corput 估计的一个关键点就是右边的界不依赖于所积分区间 I 的长度。*

证明: 我们用归纳法。当 $k = 2$ 时，命题已经证明。现在假设 $k \geqslant 3$ 并且结论对 $k-1$ 成立，我们仿照上面的证明，不妨假设 $\phi^{(k)}(x) \geqslant 1$，从而 $\phi^{(k-1)}(x)$ 是单调递增的函数。假设存在 $c \in [a,b]$ 使得 $\phi'(c) = 0$。此时，我们将 $[a,b]$ 分成三个区间

$$I = I_1 \cup I_2 \cup I_3 = [a, c-\delta] \cup [c-\delta, c+\delta] \cup [c+\delta, b].$$

所以在 I_1 和 I_3 上，我们有 $|\phi'(x)| \geqslant \delta$。根据前面的命题，我们有

$$\Big|\int_{I_1} e^{i\lambda\phi(x)}dx\Big| \leqslant \frac{2^{k-1}}{(\delta\lambda)^{\frac{1}{k-1}}}, \quad \Big|\int_{I_3} e^{i\lambda\phi(x)}dx\Big| \leqslant \frac{2^{k-1}}{(\delta\lambda)^{\frac{1}{k-1}}}.$$

在 I_2 上，我们用平凡的估计：

$$\Big|\int_{I_2} e^{i\lambda\phi(x)}dx\Big| \leqslant \int_{I_2} 1dx = 2\delta.$$

所以，

$$\begin{aligned}|J(\lambda)| &\leqslant \Big|\int_{I_1} e^{i\lambda\phi(x)}dx\Big| + \Big|\int_{I_2} e^{i\lambda\phi(x)}dx\Big| + \Big|\int_{I_3} e^{i\lambda\phi(x)}dx\Big| \\ &= \frac{2^k}{(\delta\lambda)^{\frac{1}{k-1}}} + 2\delta.\end{aligned}$$

当 $\delta = 2^{\frac{(k-1)^2}{k}}\lambda^{-\frac{1}{k}}$ 时，上面的表达式取得最小值，这就完成了证明。 □

最终我们回到

$$I(\lambda)=\int_{\mathbb{R}} e^{i\lambda\phi(x)}a(x)dx.$$

命题 174. 假设 $a(x)$ 在 $[-M,M]$ 之外恒为零并且对任意的 $x\in[-M,M]$，$|\phi''(x)|\geqslant 1$。那么，

$$\big|I(\lambda)\big|\leqslant\frac{C}{\sqrt{\lambda}},$$

其中，$C=2\sqrt{3}\displaystyle\int_{\mathbb{R}}|a'(x)|dx$。

证明: 回忆一下，我们利用振荡函数的原函数可以研究它的相消性。据此，我们自然地定义

$$J_\lambda(x)=\int_a^x e^{i\lambda\phi(t)}dt.$$

所以，

$$\begin{aligned}I(\lambda)&=\int_{\mathbb{R}}J_\lambda(x)'a(x)dx\\&=-\int_{\mathbb{R}}J_\lambda(x)a(x)'dx+J_\lambda(x)a(x)\big|_{x=-M}^{x=M}.\end{aligned}$$

我们可以对 $J_\lambda(x)$ 用上面的估计，所以

$$I(\lambda)\leqslant\frac{2\sqrt{3}}{\lambda^{\frac{1}{2}}}\int_{\mathbb{R}}|a(x)'|dx.$$

这就完成了证明。 □

29.1 建议阅读：第一学期课程没有覆盖的内容

由于时间的限制，部分重要的或者经典的内容并没有被覆盖，建议同学们在假期的时候可以考虑自己查阅参考书进行学习（我们在下个学期也会尝试在作业或其它的场合补充）：

- Arzelà-Ascoli 引理

  研究一列有界连续函数在何种条件下有收敛子序列。

- Γ-函数的构造和渐近行为。

- 区间上特殊的多项式函数以及应用：Hermite 多项式等。

- 函数的卷积。

- 某些特殊类型函数的初等原函数的求解，比如有理函数等（几乎所有经典的数学分析参考书都有这样的章节）。

- 幂级数的收敛半径与幂函数定义的函数。

- Euler-MacLaurin 展开及应用。

29.2 第一学期期末考试：Gauss 的一个椭圆积分的计算，$\int_0^\infty \frac{1}{1+x^\alpha}dx$ 的计算，整系数多项式的 Chudnovsky 逼近

清华大学 数学分析一（2019 秋季学期） 期末考试

考试说明

请在清华大学考试专用纸上**正反面**用钢笔，签字笔或者圆珠笔书写，并注明自己的姓名和年级。除定理公式所涉及的人名之外，请使用**中文**。

考试时间为 下午 2:30 至 下午 6:30。考试结束请交答题纸，草稿纸和试题请带走。大题之间是相互独立的，考试中后面的问题可以使用前面问题的结论（无论答题人是否已经得到正确的证明或者答案）。

题目 A（微分与积分的基本概念考察，共 60 分）

A1)（5 分）试求出函数 $f(x) = x^3 - 12x + 1$ 在区间 $[0, 1]$ 上的最大值并说明理由。

A2)（5 分）证明，如果 $f(x) = x^3 + ax^2 + bx + c$ 是 $\mathbb{R}$ 上定义的 3 次多项式函数，那么，$f(x)$ 一定不是 $\mathbb{R}$ 上的凸函数。

A3)（5 分）在 $x = 0$ 处写下函数 $f(x) = \sqrt{1 + x}$ 的 Peano 余项的 3 阶 Taylor 展开，即余项要求是 $o(x^3)$。

A4)（5 分）定义区间 $[-1, 1]$ 上函数

$$f(x) = \begin{cases} x \sin\left(\frac{1}{x}\right), & x \in [-1, 0) \cup (0, 1]; \\ 0, & x = 0. \end{cases}$$

证明，$f \in C([-1, 1])$ 但是 f 不可微。

A5)（5 分）计算定积分

$$\int_0^1 \frac{x^2 + 4}{x^2 + 3x + 2} dx.$$

A6)（5 分）计算定积分

$$\int_0^1 x^2 e^x dx.$$

A7)（5 分）计算定积分

$$\int_0^\pi \sin^3(x) + \sin(2x) dx.$$

A8)（5 分）计算定积分：

$$\int_{-\pi}^\pi \sin(x^3 + 2x) dx.$$

A9)（5 分）计算下面的极限：

$$\lim_{n \to \infty} \sum_{k=1}^n \frac{k^{2019}}{n^{2020}}.$$

A10)（5 分）判断下面的反常积分是否收敛并说明理由（对数函数是以 e 为底的）：

$$\int_0^1 \frac{\log x}{x^{\frac{1}{2}}} dx.$$

A11)（5 分）利用面积方法证明：对任意的自然数 $n \geqslant 1$，我们有

$$\frac{1}{1} + \frac{1}{2} + \frac{1}{3} + \cdots + \frac{1}{n} \geqslant \log(n).$$

A12)（5 分）证明，如下的反常积分收敛：

$$\int_2^\infty \frac{\sin x}{\log x} dx.$$

题目 B（Gauss 的一个椭圆积分的计算，共 21 分）

B1)（2 分）假设 $a > 0$，$b > 0$。证明，如下定义的反常积分是收敛的：

$$I(a,b) = \int_0^\infty \frac{1}{\sqrt{x^2 + a^2}\sqrt{x^2 + b^2}} dx.$$

B2)（2 分）证明，$I(a,b)$ 可以用下面的定积分来表达：

$$\int_0^{\frac{\pi}{2}} \frac{1}{\sqrt{a^2 \cos^2\theta + b^2 \sin^2\theta}} d\theta.$$

（提示：利用变量换元 $x = b\tan\theta$）

B3)（2 分）证明，$I(a,b)$ 所定义的映射

$$I : (0,+\infty) \times (0,+\infty) \to \mathbb{R}, \quad (a,b) \mapsto I(a,b)$$

是连续函数。

B4)（2 分）证明，对任意的 $\lambda > 0$，我们有

$$I(a,b) = \int_0^\infty \frac{\lambda}{ab} \frac{1}{\sqrt{x^2 + \left(\frac{\lambda}{a}\right)^2}\sqrt{x^2 + \left(\frac{\lambda}{b}\right)^2}} dx.$$

（提示：利用变量换元 $y = \dfrac{\lambda}{x}$）

B5)（2 分）证明，$I(a,b)$ 可以被表达为

$$I(a,b) = 2\int_{\sqrt{ab}}^\infty \frac{1}{\sqrt{x^2 + a^2}\sqrt{x^2 + b^2}} dx.$$

（提示：变量替换 $y = \dfrac{ab}{x}$ 将 $(0, \sqrt{ab})$ 变为 $(\sqrt{ab}, \infty)$）

B6)（2 分）证明，映射

$$\varphi : (\sqrt{ab}, +\infty) \to (0, +\infty), \quad x \mapsto \varphi(x) = \frac{1}{2}\left(x - \frac{ab}{x}\right)$$

是连续可微的双射。

B7)（3 分）利用变量替换 $y = \dfrac{1}{2}\left(x - \dfrac{ab}{x}\right)$ 证明，

$$I(a,b) = \int_0^\infty \frac{1}{\sqrt{y^2 + \left(\frac{a+b}{2}\right)^2}\sqrt{y^2 + \left(\sqrt{ab}\right)^2}} dy.$$

从而，Gauss 关于上述椭圆积分的平均值公式成立：

$$I(a,b) = I(\frac{a+b}{2}, \sqrt{ab}).$$

B8)（2 分）我们归纳地定义数列：$a_1 = a$，$b_1 = b$；对于 $n \geqslant 1$，我们定义 $a_{n+1} = \dfrac{a_n + b_n}{2}$，$b_{n+1} = \sqrt{a_n b_n}$。证明，$\lim\limits_{n\to\infty} a_n$ 和 $\lim\limits_{n\to\infty} b_n$ 存在。

B9)（2 分）证明，$\lim\limits_{n\to\infty} a_n = \lim\limits_{n\to\infty} b_n$。我们用 $M(a,b)$ 表示这个值并称之为 a 和 b 的**算数几何均衡值**（arithmetic-geometric mean）。

B10)（2 分）证明，对于 $a > 0$ 和 $b > 0$，它们的算数几何均衡值 $M(a,b)$ 与椭圆积分 $I(a,b)$ 可以相互表达：

$$2M(a,b)I(a,b) = \pi.$$

题目 C（$\displaystyle\int_0^\infty \frac{1}{1+x^\alpha}dx$ 的计算，其中 $\alpha > 1$，共 24 分）

假设 $\alpha > 1$，$\beta \in (0,1)$，$\alpha\beta = 1$。

C1)（2 分）证明，如下的反常积分

$$I(\alpha) = \int_0^\infty \frac{1}{1+x^\alpha}dx$$

是收敛的。

C2)（2 分）证明，如下的两个反常积分

$$J_1(\beta) = \int_0^1 \frac{x^{\beta-1}}{1+x}dx, \quad J_2(\beta) = \int_0^1 \frac{x^{-\beta}}{1+x}dx$$

是收敛的。

C3)（2 分）证明，

$$\int_0^1 \frac{1}{1+x^\alpha}dx = \beta J_1(\beta).$$

（提示：对任意的 $\varepsilon > 0$，先证明 $\displaystyle\int_\varepsilon^1 \frac{1}{1+x^\alpha}dx = \beta\int_{\varepsilon^\alpha}^1 \frac{x^{\beta-1}}{1+x}dx$）

C4)（2 分）证明，

$$\int_1^{\infty} \frac{1}{1+x^{\alpha}} dx = \beta J_2(\beta).$$

C5)（2 分）对于整数 $n \geqslant 1$，定义

$$h_n(x) = \sum_{k=0}^{n} (-1)^k x^k.$$

证明，对任意的 $x \in [0,1]$，我们都有

$$\left| h_n(x) - \frac{1}{1+x} \right| \leqslant x^n.$$

（提示：可以考虑函数 $\dfrac{1}{1+x}$ 在 0 处的 Taylor 展开）

C6)（2 分）令

$$J_{1,n}(\beta) = \int_0^1 x^{\beta-1} h_n(x) dx, \quad J_{2,n}(\beta) = \int_0^1 x^{-\beta} h_n(x) dx$$

证明，

$$\lim_{n\to\infty} J_{1,n}(\beta) = J_1(\beta), \quad \lim_{n\to\infty} J_{2,n}(\beta) = J_2(\beta).$$

C7)（2 分）我们如下定义 $[0,\pi]$ 区间上的函数 g：

$$g(x) = \begin{cases} \frac{\cos\left(\frac{x}{\alpha}\right)-1}{\sin\left(\frac{x}{2}\right)}, & x \in (0,\pi] \\ 0, & x = 0. \end{cases}$$

证明，$g \in C^1([0,\pi])$。

C8)（3 分）对任意的整数 $n \geqslant 1$，定义积分

$$a_n = \int_0^{\pi} g(x) \sin\left((2n+1)\frac{x}{2}\right) dx.$$

证明，存在常数 C，使得对任意的 n，我们都有

$$|a_n| \leqslant \frac{C}{2n+1}.$$

C9)（2 分）对任意的整数 $n \geqslant 1$，令

$$\varphi_n(x) = \cos(x) + \cos(2x) + \cos(3x) + \cdots + \cos(nx).$$

定义积分

$$A_n = \int_0^{\pi} \varphi_n(x) \cos\left(\frac{x}{\alpha}\right) dx.$$

证明，

$$A_n = \frac{\alpha}{2} \sin\left(\frac{\pi}{\alpha}\right) \sum_{k=1}^{n} (-1)^k \left(\frac{1}{1+\alpha k} + \frac{1}{1-\alpha k}\right).$$

C10)（3 分）证明

$$\varphi_n(x) = -\frac{1}{2} + \frac{1}{2}\frac{\sin\left((2n+1)\frac{x}{2}\right)}{\sin\left(\frac{x}{2}\right)}.$$

据此，通过计算 $A_n = \int_0^\pi \varphi_n(x)\cos\left(\frac{x}{\alpha}\right)dx$ 来证明，$\lim\limits_{n\to\infty} A_n = -\frac{\alpha}{2}\sin\left(\frac{\pi}{\alpha}\right) + \frac{\pi}{2}$。

C11)（2 分）证明，$I(\alpha) = \dfrac{\pi}{\alpha\sin\left(\frac{\pi}{\alpha}\right)}$。

附加题

附加题 1：题目 D（Stone-Weierstrass 定理，共 10 分）

经典的 Stone-Weierstrass 定理讲的是任意给定 $f \in C([0,1])$，总能找到一个实系数多项式的序列 $\{P_n\}_{n\geqslant 1} \subset \mathbb{R}[X]$，使得 P_n 一致收敛到 f。下面的问题是对这个定理的进一步讨论。

此问题中，$\{D1), D2), D3)\}$，$\{D4)\}$ 和 $\{D5), D6), D7)\}$ 是三个相互独立的部分。

D1)（2 分）假设 $f \in C^1([0,1])$ 并且 f 是单调递增的，那么存在实系数多项式的序列 $\{P_n\}_{n\geqslant 1} \subset \mathbb{R}[X]$，使得

- 对任意的 n，P_n 在 $[0,1]$ 上单调递增。
- P_n 一致收敛到 f。

D2)（1 分）假设 $f \in C([0,1])$ 并且 f 是单调递增的。我们将 $f(x)$ 延拓到 $[0,2]$ 上使得对任意的 $x \in (1,2]$，$f(x) = f(1)$。定义函数 $f_n : [0,1] \to \mathbb{R}$，使得

$$f_n(x) = n\int_x^{x+\frac{1}{n}} f(y)dy.$$

证明，函数序列 $\{f_n\}_{n\geqslant 1}$ 满足

- 对任意的 n，$f_n \in C^1([0,1])$；
- 对任意的 n，f_n 在 $[0,1]$ 上单调递增；
- f_n 一致收敛到 f。

D3)（1 分）假设 $f \in C([0,1])$ 并且 f 是单调递增的，那么存在实系数多项式的序列 $\{P_n\}_{n\geqslant 1} \subset \mathbb{R}[X]$，使得

- 对任意的 n，P_n 在 $[0,1]$ 上单调递增。
- P_n 一致收敛到 f。

D4)（2 分）（Walsh）假设 $f \in C([0,1])$，$x_1, x_2, \cdots, x_m$ 是 $[0,1]$ 区间上给定的 m 个点，那么存在实系数多项式的序列 $\{P_n\}_{n\geqslant 1} \subset \mathbb{R}[X]$，使得

– 对任意的 $n \geqslant 1$ 和任意的 $1 \leqslant i \leqslant m$，$P_n(x_i) = f(x_i)$；

– P_n 一致收敛到 f。

D5)（1 分）假设 $I = [a, b] \subset (0, 1)$，多项式 $p(x) = 2x(1 - x)$，令 $Q_n = p \circ p \circ \cdots \circ p$ 为 p 与自身的 n-次复合，即

$$Q_n(x) = \underbrace{p\Big(p\big(\cdots\big(p}_{n\text{ 个 }p}(x)\big)\cdots\Big),$$

证明，$\{Q_n\}_{n\geqslant 1}$ 在 I 上一致收敛到常数值函数 $\dfrac{1}{2}$。

D6)（1 分）假设 $I = [a, b] \subset (0, 1)$，对任意的整数 $k \in \mathbb{Z}$，证明，存在整系数多项式的序列 $\{P_n\}_{n\geqslant 1} \subset \mathbb{Z}[X]$，使得 $\{P_n\}_{n\geqslant 1}$ 在 I 上一致收敛到常数值函数 2^k。

D7)（2 分）（Chudnovsky）假设 $f \in C(I)$，其中 $I = [a, b] \subset (0, 1)$。证明，存在整系数多项式的序列 $\{P_n\}_{n\geqslant 1} \subset \mathbb{Z}[X]$，使得 $\{P_n\}_{n\geqslant 1}$ 在 I 上一致收敛到 f。

注记. *关于用整系数多项式来逼近连续函数的结果，同学们可以去参考 Hervé Pépin 和 Nicolas Tosel 文章*：Approximation par des polynômes à coefficients dans $\mathbb{Z}$, *RMS, 114 ème année, 2003-2004.*

附加题 2：题目 E（一个函数方程的求解，5 分）

试找出所有的实值的 $f \in C(\mathbb{R})$，使得对任意的 $x, y \in \mathbb{R}$，我们都有

$$f(x)f(y) = \int_{x-y}^{x+y} f(t)dt.$$

附加题 3：题目 F（方程解的个数，5 分）

任意给定两两不相等的实数 $\alpha_1, \alpha_2, \cdots, \alpha_{2020}$，任意给定不全为 0 的实数 $a_1, a_2, \cdots, a_{2020}$，考虑在 $(0, +\infty)$ 上定义的函数

$$f(x) = a_1 x^{\alpha_1} + a_2 x^{\alpha_2} + \cdots + a_{2020} x^{\alpha_{2020}}.$$

证明，$f(x)$ 在 $(0, +\infty)$ 上至多有 2019 个根。

29.3 寒假作业

清华大学数学分析 1 —— 寒假作业 (冬练三九)

题目 A（Young 不等式） $f:\mathbb{R}_{\geqslant 0}\to\mathbb{R}_{\geqslant 0}$ 是连续可微的函数。假设 $f(0)=0$，$\lim\limits_{x\to\infty}f(x)=+\infty$ 并且对任意的 $x\geqslant 0$，都有 $f'(x)>0$。我们用 $g(x)$ 表示 $f^{-1}(x)$，即 f 的反函数。

A1) 证明，对任意的 $a\geqslant 0$，我们有

$$af(a)=\int_0^a f(x)dx+\int_0^{f(a)}g(y)dy.$$

A2) 证明 Young 不等式：对任意的 $a,b\geqslant 0$，我们有

$$ab\leqslant\int_0^a f(x)dx+\int_0^b g(y)dy.$$

A3) 假设 $f:\mathbb{R}_{\geqslant 0}\to\mathbb{R}_{\geqslant 0}$ 是连续严格递增的函数，并且 $f(0)=0$，$\lim\limits_{x\to\infty}f(x)=+\infty$。试重新证明 A2) 中的不等式。

A4) 假设 $a,b\geqslant 0$，$p>1$，$q>1$ 并且 $\dfrac{1}{p}+\dfrac{1}{q}=1$。证明 Hölder 不等式并确定取等号的条件：

$$ab\leqslant\frac{a^p}{p}+\frac{b^q}{q}.$$

（提示：考虑函数 $f(x)=x^{p-1}$）

题目 B（一个 Sobolev 不等式） Sobolev 不等式的一个重要的观点就是用函数以及它的导数的积分来控制函数的最大值。在这个问题中，$[a,b]$ 是任意给定的有界闭区间。

B1)（Cauchy-Schwarz 不等式）假设 $f,g\in\mathcal{R}([a,b])$，利用多项式

$$P(t)=\int_a^b\big(f(x)+tg(x)\big)^2dx$$

证明：

$$\Big|\int_a^b f(x)g(x)dx\Big|\leqslant\left(\int_a^b|f(x)|^2dx\right)^{\frac{1}{2}}\left(\int_a^b|g(x)|^2dx\right)^{\frac{1}{2}}.$$

B2) 证明，对任意的 $\varepsilon>0$，存在常数 $C_\varepsilon>0$，使得对任意的 $f\in C^1\big([a,b]\big)$，对任意的 $x\in[a,b]$，我们都有

$$\big|f(x)^2-f(a)^2\big|\leqslant C_\varepsilon\int_a^b f(x)^2dx+\varepsilon\int_a^b f'(x)^2dx.$$

B3) 证明，对任意的 $\varepsilon>0$，存在常数 $D_\varepsilon>0$，使得对任意的 $f\in C^1\big([a,b]\big)$，我们都有

$$\sup_{x\in[a,b]}|f(x)|^2\leqslant D_\varepsilon\int_a^b f(x)^2dx+\varepsilon\int_a^b f'(x)^2dx.$$

题目 C（Wirtinger 不等式） 令 $E=\big\{f\in C^1\big([0,1]\big)\big|f(0)=f(1)=0\big\}$。

C1) 对任意的 $f\in E$，我们定义反常积分

$$\mathbf{I}_1=\int_0^1\frac{f(x)f'(x)}{\tan(\pi x)}dx,\quad \mathbf{I}_2=\int_0^1\frac{f(x)^2}{\tan^2(\pi x)}(1+\tan^2(\pi x))dx.$$

证明，上述反常积分收敛并确定它们的比值。

C2) 证明（Wirtinger 不等式），对任意的 $f\in E$，我们有

$$\int_0^1 f(x)^2dx\leqslant\pi^{-2}\int_0^1 f'(x)^2dx.$$

C3) 试找出所有的 $f\in E$，使得上面的不等式取等号。

C4) 假设 $f\in\mathcal{R}\big([0,2\pi]\big)$，试找出 $A\in\mathbb{R}$，使得积分 $\displaystyle\int_0^{2\pi}|f(x)-A|^2dx$ 最小。

C5) 证明另一种版本的 Wirtinger 不等式，对任意的 $f\in C^1([0,2\pi])$，如果 $\displaystyle\int_0^{2\pi}f(x)dx=0$，那么

$$\int_0^{2\pi}f(x)^2dx\leqslant\int_0^{2\pi}f'(x)^2dx,$$

并且等号成立当且仅当 f 是 $\sin x$ 和 $\cos x$ 的一个线性组合。

C6*)（等周不等式的证明）假设 $\gamma:[0,2\pi]\to\mathbb{R}^2$ 是一条 C^1 的闭曲线（$\gamma(0)=\gamma(2\pi)$），不妨记 $\gamma(t)=\big(x(t),y(t)\big)$，我们假设 γ 不自交（即 γ 在 $[0,2\pi)$ 上是单射）并且其长度为 L，即

$$L=\int_0^{2\pi}\sqrt{x'(t)^2+y'(t)^2}dt.$$

根据 Stokes 公式（下学期），γ 所围成图形的面积为

$$A=\int_0^{2\pi}x'(t)y(t)dt.$$

证明，$L^2\geqslant 4\pi A$ 并且等号成立的当且仅当 γ 给出一个圆。

题目 D(Gauss 逼近)对任意的 $n\geqslant 1$,我们定义 Legendre 多项式 $P_n(x)=\dfrac{1}{2^n n!}\left(\dfrac{d}{dx}\right)^n\big((x^2-1)^n\big)$。为了方便起见，我们令 $P_0(x)\equiv 1$。

D1) 证明，任意给定连续函数 $\varphi(x)\in C([-1,1])$，对任意的 $\varepsilon>0$，存在正整数 N 和实数 $c_1,c_2,\cdots,c_N$，使得

$$\|\varphi(x)-\sum_{k=1}^N c_kP_k(x)\|_\infty<\varepsilon.$$

D2) 证明，对任意的 $n \geqslant 1$，$P_n(x)$ 满足下面的微分方程：

$$(1-x^2)f'' - 2xf' + n(n+1)f = 0.$$

D3) 证明，对任意的 $n, m \geqslant 1$，我们有

$$\int_{-1}^{1} P_n(x)P_m(x)dx = \begin{cases} 0, & m \neq n; \\ \frac{2}{2n+1}, & m = n. \end{cases}$$

D4) 给定 $n \geqslant 1$。证明，如果 $Q(x)$ 是次数不超过 $n-1$ 的多项式，那么

$$\int_{-1}^{1} Q(x)P_n(x)dx = 0.$$

D5) 证明，对任意的 $n \geqslant 1$，$P_n(x)$ 恰有 n 个实数根并且悉数落在 $(-1,1)$ 上。我们将这些根记作 $x_1^{(n)} < x_2^{(n)} < \cdots < x_n^{(n)}$.

D6) 证明，对任意的 $n \geqslant 1$，存在实数 $\alpha_1^{(n)}, \alpha_2^{(n)}, \cdots, \alpha_n^{(n)}$，使得对任意的次数不超过 $2n-1$ 的多项式 $Q(x)$，我们都有

$$\int_{-1}^{1} Q(x)dx = \sum_{i=1}^{n} \alpha_i^{(n)} Q(x_i^{(n)}).$$

D7) Gauss 对积分的一个逼近公式：对任意的连续函数 $\varphi(x) \in C([-1,1])$，我们令

$$G_n(\varphi) = \sum_{i=1}^{n} \alpha_i^{(n)} \varphi(x_i^{(n)}).$$

证明，$\lim\limits_{n\to\infty} G_n(\varphi) = \int_{-1}^{1} \varphi(x)dx$。

题目 E（等分布问题） 给定一个数列 $\{x_k\}_{k\geqslant 1}$，其中 $\{x_k\}_{k\geqslant 1} \subset [0,1]$。对任意的 $0 \leqslant a < b \leqslant 1$，我们令 $S_n([a,b])$ 等于该数列的前 n 个数（即 $x_1, x_2, \cdots, x_n$）中落在 $[a,b]$ 中的数的数目，即

$$S_n([a,b]) = \Big|\big\{x_k \big| k \leqslant n, x_k \in [a,b]\big\}\Big|.$$

如果对任意的 $0 \leqslant a < b \leqslant 1$，我们都有

$$\lim_{n\to\infty} \frac{S_n([a,b])}{n} = b - a,$$

我们就称数列 $\{x_k\}_{k\geqslant 1}$ 在 $[0,1]$ 上**等分布**。

E1) 证明，如果数列 $\{x_k\}_{k\geqslant 1}$ 在 $[0,1]$ 上等分布，那么 $\{x_k\}_{k\geqslant 1}$ 是 $[0,1]$ 的稠密子集。

E2) 试构造 $[0,1]$ 的稠密子集 $\{x_k\}_{k\geqslant 1}$，使得数列 $\{x_k\}_{k\geqslant 1}$ 在 $[0,1]$ 上不是等分布的。

E3) 任意给定数列 $\{x_k\}_{k\geqslant 1}\subset[0,1]$，我们令

$$D_n=\sup_{0\leqslant a<b\leqslant 1}\Big|\frac{S_n([a,b])}{n}-(b-a)\Big|,\quad D_n^*=\sup_{0<b<1}\Big|\frac{S_n([0,b])}{n}-b\Big|.$$

证明，$D_n^*\leqslant D_n\leqslant 2D_n^*$。

E4) 证明，数列 $\{x_k\}_{k\geqslant 1}$ 在 $[0,1]$ 上等分布当且仅当 $\lim\limits_{n\to\infty}D_n^*=0$。

E5) 证明，数列 $\{x_k\}_{k\geqslant 1}$ 在 $[0,1]$ 上等分布当且仅当对任意的连续函数 $f\in C\big([0,1]\big)$，我们都有

$$\lim_{n\to\infty}\frac{1}{n}\sum_{k=1}^{n}f(x_k)=\int_0^1 f(x)dx.$$

（提示：试用连续函数逼近 $\mathbf{1}_{[a,b]}$）

E6) 证明 Weyl 判定准则：数列 $\{x_k\}_{k\geqslant 1}$ 在 $[0,1]$ 上等分布当且仅当对任意 $p\in\mathbb{Z}_{\geqslant 1}$，我们都有

$$\lim_{n\to\infty}\frac{1}{n}\sum_{k=1}^{n}e^{2\pi p x_k}=0.$$

（提示：利用三角函数版本的 Weierstrass-Stone 定理）

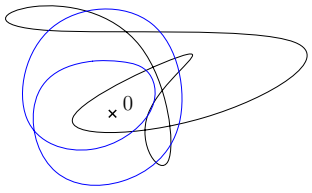
E7) 假设 $\theta>0$ 是实数，$\{x\}$ 代表一个数的小数部分，即 $\{x\}=x-\lfloor x\rfloor$。那么，数列 $\big\{\{n\theta\}\big\}_{n\geqslant 1}$ 在 $[0,1]$ 上等分布当且仅当 θ 是无理数。

E8) 证明，数列 $\{\xi_n\}_{n\geqslant 1}=\big\{\{\sqrt{n}\}\big\}_{n\geqslant 1}$ 在 $[0,1)$ 上是等分布的。

E9) 任给实数 $a\neq 0$，$\sigma\in(0,1)$。证明，数列 $\{\xi_n\}_{n\geqslant 1}=\big\{\{an^\sigma\}\big\}_{n\geqslant 1}$ 在 $[0,1)$ 上是等分布的。

E10) 证明，对任意的 $a\in\mathbb{R}$，数列 $\{\xi_n\}_{n\geqslant 1}=\big\{\{a\log n\}\big\}_{n\geqslant 1}$ 在 $[0,1)$ 上不是等分布的。

题目 F（平面闭曲线的环绕数） 令 $E=\{f:\mathbb{R}\to\mathbb{C}^\times\big|f$ 连续可微并且以 2π 为周期$\}$，其中 $\mathbb{C}^\times=\mathbb{C}-\{0\}$。对于这样的函数 f，它在 $[0,2\pi]$ 上的限制给出了 $\mathbb{C}$ 上不过原点的一条闭曲线。



对任意的 $f\in E$，我们定义

$$d(f)=\frac{1}{2\pi i}\int_0^{2\pi}\frac{f'(t)}{f(t)}dt.$$

F1) 证明 $d(f)$ 是良好定义的并且对下面这一族曲线

$$f_n : \mathbb{R} \to \mathbb{C}^\times, \quad x \mapsto f_n(x) = e^{inx}$$

计算其环绕数 $d(f_n)$（直观上，上述曲线是用 $[0, 2\pi]$ 区间将单位圆周绕 n 圈，其中 n 的正负与绕的方向相关）。

F2) 试用 $\mathbb{C}$ 上的极坐标系来写下 $d(f)$ 的定义（即将 $f(t)$ 写为 $\rho(t)e^{i\theta(t)}$），这个计算给出了 $d(f)$ 的直观含义。

F3) 利用函数 $\psi(t) = \exp\left(\int_0^t \frac{f'(s)}{f(s)} ds\right)$ 证明 $d(f)$ 是整数。（提示：证明 ψ 与 f 解同一个 1 阶常微分方程）

F4) 证明，对任意的 $f \in E$，存在 $\varepsilon > 0$，对任意的 $g \in E$，如果 $\|f - g\|_\infty < \varepsilon$，那么 $d(g) = d(f)$。

F5) 你是否可以利用 F4) 的结论对于函数 $f : \mathbb{R} \to \mathbb{C}^\times$，其中 f 连续并且以 2π 为周期来定义 $d(f)$?（提示：可以利用三角函数的 Weierstrass-Stone 定理）

F6) 证明同伦不变性：假设 $F(t, \tau) : \mathbb{R} \times [0, 1] \to \mathbb{C}^\times$ 是连续映射并且使得对任意的 $\tau \in [0, 1]$，$F(\cdot, \tau) \in E$，那么

$$d\big(F(x, 0)\big) = d\big(F(x, 1)\big).$$

F7) 假设 $P(x) = x^n + c_{n-1}x^{n-1} + \cdots + c_1 x + c_0$ 是一个 n 次的首一复系数多项式，假设 $P(0) \neq 0$，证明，存在 $\varepsilon_0 > 0$，使得对任意的 $\varepsilon \in (0, \varepsilon_0)$，函数 $f_\varepsilon(x) = P(\varepsilon e^{\pi i x}) \in E$ 并计算 $d(f_\varepsilon)$。

F8) 假设 $P(x) = x^n + c_{n-1}x^{n-1} + \cdots + c_1 x + c_0$ 是一个 n 次的首一复系数多项式，假设 $P(0) \neq 0$，证明，存在 $R_0 > 0$，使得对任意的 $R \in (R_0, \infty)$，函数 $f_R(x) = P(Re^{\pi i x}) \in E$ 并计算 $d(f_R)$。

F9) 证明代数基本定理：对任意 n 次的复系数多项式 $P(x)$，它至少有一个根。（提示：反证法）

题目 G（另一个处处连续处处不可微的函数：Bolzano 曲线） 我们按照归纳的方式定义 $[0, 1]$ 上的函数列 $\{f_n\}_{n \geqslant 0}$，其中，$f_0(x) = x$，对于 $n \geqslant 0$ 且 $0 \leqslant k \leqslant 3^n$，我们有

$$\begin{cases} f_{n+1}\left(\frac{k}{3^n}\right) & = f_n\left(\frac{k}{3^n}\right), \\ f_{n+1}\left(\frac{k}{3^n} + \frac{1}{3^{n+1}}\right) & = f_n\left(\frac{k}{3^n} + \frac{2}{3^{n+1}}\right), \\ f_{n+1}\left(\frac{k}{3^n} + \frac{2}{3^{n+1}}\right) & = f_n\left(\frac{k}{3^n} + \frac{1}{3^{n+1}}\right). \end{cases}$$

并且 f_{n+1} 在形如 $\left[\frac{k}{3^{n+1}}, \frac{k+1}{3^{n+1}}\right]$ 的区间上是线性函数。

G1) 证明，对任意的 $n \geqslant 0$，对任意的 $0 \leqslant x, y \leqslant 1$，我们都有

$$\big|f_n(x) - f_n(y)\big| \leqslant 2^n |x - y|.$$

G2) 证明，对任意的 $n \geqslant 0$，对任意的 $0 \leqslant x \leqslant 1$，我们都有

$$\left|f_{n+1}(x) - f_n(x)\right| \leqslant \frac{2^n}{3^{n+1}}.$$

G3) 证明，函数列 $\{f_n\}_{n\geqslant 0}$ 在 $[0,1]$ 上一致收敛到某个连续函数 $f \in C([0,1])$。

G4) 证明，对任意的 $n \geqslant 1$，对任意的 $0 \leqslant k \leqslant 3^n$，我们有 $f\left(\dfrac{k}{3^n}\right) = f_n\left(\dfrac{k}{3^n}\right)$。

G5) 证明，对任意的 $n \geqslant 1$，对任意的 $0 \leqslant k \leqslant 3^n$，$f$ 在 $\dfrac{k}{3^n}$ 处不可微。

G6) 证明，f 在 $[0,1]$ 上处处不可微。（提示：考虑函数 $(h,k) \mapsto \dfrac{f(x+h)-f(x-k)}{h+k}$）

题目 H 考虑有界闭区间 $[a,b]$ 上的连续函数列 $\{f_n\}_{n\geqslant 1} \subset C([a,b])$。我们假设数项的级数

$$\sum_{n=1}^{\infty}\left(\int_a^b |f_n(x)|dx\right)$$

收敛。

H1) 是否对任意的 $x_0 \in [a,b]$，级数 $\displaystyle\sum_{n=1}^{\infty}|f_n(x_0)|$ 收敛?

H2) 令 $E = \left\{x_0 \in [a,b]\middle|\text{级数}\sum_{n=1}^{\infty}|f_n(x_0)|\text{ 收敛}\right\}$。证明，$E \subset [a,b]$ 是稠密的。

H3) 令 $F = \left\{x_0 \in [a,b]\middle|\text{级数}\sum_{n=1}^{\infty}|f_n(x_0)|\text{ 发散}\right\}$。试构造 $[a,b]$ 上的连续函数列 $\{f_n\}_{n\geqslant 1} \subset C([a,b])$ 使得级数 $\displaystyle\sum_{n=1}^{\infty}\left(\int_a^b |f_n(x)|dx\right)$ 收敛并且 $F \subset [a,b]$ 是稠密的。

题目 I 在这个问题中，我们用 $\{x\}$ 代表 $x \in \mathbb{R}$ 的小数部分。

第一部分

I-1-1) 对任意的 $x \in \mathbb{R}$ 和 $q \in \mathbb{Z}_{\geqslant 1}$，我们令

$$S_q(x) = \sum_{p=1}^{q}\frac{\sin(2p\pi x)}{p\pi}.$$

证明，

$$S_q(x) = \int_0^x \frac{\sin\big((2q+1)\pi u\big)}{\sin(\pi u)}du - x.$$

I-1-2) 我们定义 $(-1,1)$ 上的函数 φ:

$$\varphi(x)=\begin{cases}\frac{1}{\sin(\pi x)}-\frac{1}{\pi x}, & x\in(-1,0)\cup(0,1),\\ 0, & x=0.\end{cases}$$

证明，$\varphi\in C^1\big((-1,1)\big)$。

I-1-3) 证明，存在常数 $A_1>0$，使得对任意的 $x\in[0,\frac{1}{2}]$ 和 $q\in\mathbb{Z}_{\geqslant 1}$，我们有

$$\Big|S_q(x)+x-\frac{1}{\pi}\int_0^{(2q+1)\pi x}\frac{\sin u}{u}du\Big|\leqslant\frac{A_1}{q}.$$

I-1-4) 计算 $\displaystyle\int_0^{\infty}\frac{\sin u}{u}du$。（提示：选取特殊的 x_0 考虑 $S_q(x_0)$）

I-1-5) 证明，存在常数 $A_2>0$，使得对任意的 $x\in\mathbb{R}$ 和 $q\in\mathbb{Z}_{\geqslant 1}$，我们有

$$|S_q(x)|\leqslant A_2.$$

I-1-6) 试写下函数 $S(x)$ 的解析表达式，使得对任意的 $x\in\mathbb{R}$，$\lim\limits_{q\to\infty}S_q(x)=S(x)$。这是一致收敛的么?

I-1-7) 证明，如果 $K\subset\mathbb{R}$ 是紧集并且 $K\cap\mathbb{Z}=\emptyset$，那么 $S_q(x)$ 在 K 上一致收敛到 $S(x)$。

I-1-8) 证明，对任意的有界闭区间 $[a,b]$ 和任意的在 $\mathbb{C}$ 中取值的连续函数 $f\in C([a,b];\mathbb{C})$，我们都有

$$\int_a^b f(x)\left(\frac{1}{2}-\{x\}\right)dx=\frac{1}{\pi}\sum_{p=1}^{+\infty}\frac{1}{p}\int_a^b f(x)\sin(2p\pi x)dx.$$

第二部分

在这一部分，我们假设 $z=s+it\in\mathbb{C}$。对于 $x\in\mathbb{R}$，我们定义 $x^z:=e^{z\log x}$。字母 n 将代表一个正整数。另外，Riemann 的 ζ 函数形式上记作

$$\zeta(z)=\sum_{n=1}^{\infty}\frac{1}{n^z}.$$

我们知道，当 $z\in(1,+\infty)$ 时，ζ 有定义。

I-2-1) 对任意给定的 $z\in\mathbb{C}$，函数 $x\mapsto x^{-z}$ 的原函数是什么?

I-2-2) 证明如下的等式:

$$\sum_{k=1}^{n}k^{-z}=1+\int_1^n x^{-z}dx-z\int_1^n\{x\}x^{-z-1}dx.$$

I-2-3) s 是 z 的实部。如果 $s>1$，证明如下的极限（说明等式两边都存在并且相等）：

$$\lim_{n\to\infty}\left(\sum_{k=1}^{n}k^{-z}-\frac{z}{z-1}\right)=-z\int_1^{\infty}\{x\}x^{-z-1}dx.$$

1-2-4) 令 $\Omega=\left\{z=s+it\in\mathbb{C}\middle|s>0,z\neq 1\right\}$，我们定义

$$\zeta_1(z)=\frac{z}{z-1}-z\int_1^{\infty}\{x\}x^{-z-1}dx.$$

证明，当 $s>1$ 时，我们有 $\zeta_1(z)=\zeta(z)$。从此，我们将 $\zeta_1(z)$ 也记作 $\zeta(z)$，它是 $\zeta(z)$ 在 Ω 中的自然扩展（解析延拓）。

1-2-5) 证明，对任意的 $z\in\Omega$，任意的 $y\geqslant n$，我们有

$$\begin{aligned}\zeta(z)-\sum_{k=1}^{n}k^{-z}=\frac{n^{1-z}}{z-1}-\frac{1}{2}n^{-z}+z\int_y^{+\infty}\left(\frac{1}{2}-\{x\}\right)x^{-z-1}dx\\+\frac{z}{\pi}\sum_{p=1}^{+\infty}\frac{1}{p}\int_n^y x^{-z-1}\sin(2p\pi x)dx.\end{aligned}$$

1-2-6) 任意给定正整数 p，任意给定 $t\in[-n,n]$，我们定义函数

$$g:[n,+\infty)\to\mathbb{R},\quad x\mapsto g(x)=\frac{-x^{-\frac{3}{2}}}{2p\pi-\frac{t}{x}}.$$

证明，存在常数 B_1，使得对任意的 $x\geqslant n$，我们都有

$$g'(x)\leqslant\frac{B_1}{p}x^{-\frac{5}{2}}.$$

证明，存在常数 B_2，使得对任意的 $y\geqslant n$，我们有

$$\left|\int_n^y x^{-\frac{3}{2}-it}e^{2ip\pi x}dx\right|\leqslant\frac{B_2}{n^{\frac{3}{2}}p},\quad\left|\int_n^y x^{-\frac{3}{2}-it}\sin(2p\pi x)dx\right|\leqslant\frac{B_2}{n^{\frac{3}{2}}p}.$$

1-2-7) 证明，存在常数 B_2，使得对任意的 $n\geqslant 1$ 和任意的 $t\in[-n,n]$，我们有

$$\left|\zeta\left(\frac{1}{2}+it\right)-\sum_{k=1}^{n}k^{-\frac{1}{2}-it}\right|\leqslant B_3\frac{\sqrt{n}}{1+|t|}.$$

第三部分

字母 n 将代表一个大于等于 2 的正整数。

I-3-1) 证明，对任意的 $x>0$，我们有 $\log(x)-1+\dfrac{1}{x}\geqslant 0$。

I-3-2) 证明，存在常数 C_1，使得如下的不等式成立（当 $n=2$ 时我们将不等式的左边取作 0）：

$$\sum_{\substack{j,k\\ 1\leqslant k\leqslant n,\frac{k}{2}<j<k}} \frac{1}{\sqrt{kj}\log\left(\frac{k}{j}\right)} \leqslant \sqrt{2}\sum_{\substack{h,k\\ 1\leqslant h\leqslant k\leqslant n}} \frac{1}{h} \leqslant C_1 n\log n.$$

I-3-3) 证明，存在常数 C_2，使得如下的不等式成立：

$$\sum_{\substack{j,k\\ 1\leqslant k\leqslant n,1\leqslant j\leqslant\frac{k}{2}}} \frac{1}{\sqrt{kj}\log\left(\frac{k}{j}\right)} \leqslant \frac{1}{\log 2}\left(\sum_{k=1}^{n}\frac{1}{\sqrt{k}}\right)^2 \leqslant C_2 n.$$

I-3-4) 证明，存在常数 C_3，使得如下的等式／不等式成立：

$$\int_0^n \Big|\sum_{k=1}^{n} k^{-\frac{1}{2}-it}\Big|^2 dt = n\sum_{k=1}^{n}\frac{1}{k} + i\sum_{\substack{j,k,j\neq k\\ 1\leqslant j,k\leqslant n}} \frac{\left(\frac{k}{j}\right)^{-in}-1}{\sqrt{kj}\log\left(\frac{k}{j}\right)} \leqslant C_3 n\log n.$$

I-3-5) 证明，存在常数 C，使得对任意的 $T>0$，都有

$$\int_0^T \left|\zeta\left(\frac{1}{2}+it\right)\right|^2 dt \leqslant CT\log T.$$

题目 J（拟周期函数） 令 E 是由 $\big\{\cos(\omega x),\sin(\omega x)\big|\omega\in\mathbb{R}\big\}$ 所生成的 $\mathbb{R}$-线性空间，很明显，$E\subset C(\mathbb{R})$ 是线性子空间。

J1) 令 $F=\big\{f\in C(\mathbb{R})\big|$存在 $\{f_n\}_{n\geqslant 1}$ 使得 f_n 一致收敛到 $f\big\}$（即 F 维 E 在 $(C(\mathbb{R}),\|\cdot\|_\infty)$ 中的闭包），证明，F 是 $\mathbb{R}$-线性空间并且对任意的 $f,g\in F$，它们的乘积 $f\cdot g\in F$。

J2) 如果 $\varphi\in C(\mathbb{R})$，$f\in F$，证明，$\varphi\circ f\in F$。

J3) 任意给定 $\omega\in\mathbb{R}$ 和 $f\in F$，证明，如下的极限存在

$$\widetilde{f}(\omega)=\lim_{T\to\infty}\frac{1}{2T}\int_{-T}^{T} e^{-i\omega x}f(x)dx.$$

J4) 任意给定 $f\in F$，证明，除了至多可数个可能的 $\omega\in\mathbb{R}$ 以外，我们有 $\widetilde{f}(\omega)=0$。

题目 K（Korovkin 定理） $E=C([0,1])$。考虑 $\mathbb{R}$-线性映射 $T:E\to E$：如果对任意的 $[0,1]$ 上的连续正函数 $f\geqslant 0$（即在每个点的取值都是非负的），Tf 也是正函数，我们就称 T 是一个**正算子**。

K1) 假设 $T:E\to E$ 是正算子。证明，存在常数 C，使得对任意的 $f\in E$，我们都有

$$\|Tf\|_\infty \leqslant C\|f\|_\infty.$$

K2) 给定 $f \in E$。证明，对任意的 $\varepsilon > 0$，存在常数 C_ε，使得对任意的 $x, y \in [0,1]$，我们有

$$|f(x) - f(y)| \leqslant \varepsilon + C_\varepsilon (x-y)^2.$$

K3)（Korovkin）对于 $n \in \mathbb{Z}_{\geqslant 0}$，$P_n(x) = x^n$ 是一族 E 中的正函数。假设 T_n 是一族 E 上的正算子。如果对任意的 $m \in \mathbb{Z}_{\geqslant 0}$，函数序列 $\{T_n(P_m)\}_{n\geqslant 1}$ 在 E 中一致收敛到 P_m，证明，对任意的 $f \in E$，函数序列 $\{T_n(f)\}_{n\geqslant 1}$ 在 E 中一致收敛到 f。

K4)（应用：用 Bernstein 多项式一致逼近连续函数）对于 $f \in E = C([0,1])$，我们定义一族算子

$$B_n : E \to E, \quad f \mapsto B_n(f)(x) = \sum_{k=0}^{n} \binom{n}{k} f\left(\frac{k}{n}\right) x^k (1-x)^{n-k}.$$

证明，函数序列 $\{B_n(f)\}_{n\geqslant 1}$ 在 E 中一致收敛到 f。

题目 L（余切函数的 Euler 展开） 对于任意的 $x \in \mathbb{R}$，我们令 $f(x) = \lim\limits_{n\to\infty} \sum\limits_{k=-n}^{n} \dfrac{1}{x+k}$。

L1) 证明，f 在 $\mathbb{R} - \mathbb{Z}$ 上是良好定义的。

L2) 证明，对任意的 $x \in \mathbb{R} - \mathbb{Z}$，我们有 $f(-x) = -f(x)$，$f(x+1) = f(x)$；对任意的 $\mathbb{R} - \dfrac{1}{2}\mathbb{Z}$，我们有 $2f(2x) = f(x) + f\left(x + \dfrac{1}{2}\right)$。

L3) 证明，函数 $x \mapsto f(x) - \pi\mathrm{cotan}(\pi x)$ 可以在 $x \in \mathbb{Z}$ 处定义从而可以被视作是 $\mathbb{R}$ 上的函数。

L4) 证明，对于 $x \notin \mathbb{Z}$，我们有 $f(x) \equiv \pi\mathrm{cotan}(\pi x)$。

题目 M（一个一致收敛的判定准则） 假设 $\{f_n\}_{n\geqslant 1}$ 是在 $[0,1]$ 上定义的一列函数（未必连续），它们满足如下的条件：对任意的收敛的数列 $\{x_n\}_{n\geqslant 1} \subset [0,1]$，数列 $\{f_n(x_n)\}_{n\geqslant 1}$ 也收敛。

M1) 证明，存在 $[0,1]$ 上定义的函数 $f : [0,1] \to \mathbb{R}$，使得对每个点 $x \in [0,1]$，$\lim\limits_{n\to\infty} f_n(x) = f(x)$。

M2) 证明，f 是连续函数。

M3) 证明，f_n 一致收敛到 f，即

$$\lim_{n\to\infty} \sup_{x\in[0,1]} |f_n(x) - f(x)| = 0.$$

题目 N（Pau Lévy，万有弦定理） 令 $E = \big\{f \in C([0,1]) \big| f(0) = f(1)\big\}$。对于任意的 $f \in E$，我们定义

$$\Lambda(f) = \big\{\sigma \in [0,1] \big| \text{存在 } x \in [0,1], \text{ 使得 } f(x+\sigma) = f(x)\big\}.$$

N1) 证明，$\bigcap_{f\in E}\Lambda(f)$ 是紧集并且

$$\{0\}\cup\big\{\frac{1}{n}\big|n\in\mathbb{Z}_{\geqslant 1}\big\}\subset\bigcap_{f\in E}\Lambda(f).$$

N2) 证明，对任意的 $f\in E$，存在 $\varepsilon>0$，使得

$$[0,\varepsilon]\cup\big\{\frac{1}{n}\big|n\in\mathbb{Z}_{\geqslant 1}\big\}\subset\bigcap_{f\in E}\Lambda(f).$$

N3) 证明，

$$\{0\}\cup\big\{\frac{1}{n}\big|n\in\mathbb{Z}_{\geqslant 1}\big\}=\bigcap_{f\in E}\Lambda(f).$$

N4)（万有弦定理）考虑连续函数 $f:[a,b]\to\mathbb{R}$ 的图像 $\Gamma_f=\big\{(x,f(x))\subset\mathbb{R}^2\big|x\in[a,b]\big\}$。$\overline{AB}$ 是图像的两个端点 $A=(a,f(a))$ 和 $B=(b,f(b))$ 所连的线段，证明，对任意的整数 $n\geqslant 1$，总存在 $c,d\in[a,b]$，使得 $C=(c,f(c))$ 和 $D=(d,f(d))$ 所连的线段与 $\overline{AB}$ 平行并且其长度恰好是 $\overline{AB}$ 长度的 $\dfrac{1}{n}$。如果 $n\geqslant 1$ 不是整数，结论是否成立?

N5) 司马迁参加 100 米赛跑，成绩是 10 秒。证明，能够找到连续的 5 秒，他恰好跑了 50 米。（我们假设人类的运动是连续的）

题目 O（除法定理） 假设 $f\in C^\infty(\mathbb{R})$ 并且 $f(0)=0$。

O1) 证明，函数

$$g:\mathbb{R}\to\mathbb{R},\quad x\mapsto g(x)=\begin{cases}\frac{f(x)}{x}, & x\neq 0;\\ f'(0), & x=0.\end{cases}$$

是 $\mathbb{R}$ 上的光滑函数。

O2) 假设对任意的 $x\neq 0$，$f(x)>0$ 并且假设 $f''(0)\neq 0$。证明，存在光滑函数 $g\in C^\infty(\mathbb{R})$，使得

$$g^2=f.$$

题目 P 证明，对任意的整数 $n\geqslant 0$，如下的反常积分均为 0：

$$\int_0^\infty x^n\sin\left(x^{\frac{1}{4}}\right)e^{-x^{\frac{1}{4}}}dx=0.$$

题目 Q 给定严格递增的实数数列 $\{a_n\}_{n\geqslant 0}$，其中 $a_0=0$。我们假设如下的级数是发散的

$$\frac{1}{a_1}+\frac{1}{a_2}+\frac{1}{a_3}+\frac{1}{a_4}+\cdots.$$

证明，对于 $f \in C([0,1])$，如果对任意的 $n \geqslant 0$，我们都有

$$\int_0^1 x^{a_n} f(x)dx = 0,$$

那么 $f \equiv 0$。

题目 R 假设 $f \in C^\infty(\mathbb{R})$，$f(0)f'(0) \geqslant 0$ 并且 $\lim\limits_{x\to\infty} f(x) = 0$。证明，存在 $0 \leqslant x_1 < x_2 < \cdots < x_n < \cdots$，使得

$$f^{(n)}(x_n) = 0.$$

题目 S（距离空间上的等距映射） 假设 (X,d) 是紧的距离空间，映射 $f : X \to X$。

S1) 如果 f 是等距映射，即对任意的 $x, y \in X$，我们都有 $d(f(x), f(y)) = d(x,y)$。证明，f 是双射。

S2) 如果对任意的 $x, y \in X$，我们都有 $d(f(x), f(y)) \geqslant d(x,y)$。证明，$f$ 是等距映射。

S3) 如果对任意的 $x, y \in X$，我们都有 $d(f(x), f(y)) \leqslant d(x,y)$ 并且 f 是满射。证明，f 是等距映射。

题目 T（不动点定理的变体） 假设 (X,d) 是完备的距离空间，映射 $f : X \to X$。

T1) 如果 (X,d) 是紧的并且对任意的 $x \neq y \in X$，都有 $d(f(x), f(y)) < d(x,y)$。证明，f 具有唯一的不动点。

T2) 当 X 非紧的时候，试举出 T1) 的一个反例。

T3) $\omega : [0,\infty) \to \mathbb{R}$ 是一个右连续的函数，其中 $\omega(0) = 0$ 并且当 $t > 0$ 时，$0 \leqslant \omega(t) < t$。假设对任意的 $x, y \in X$，有

$$d(f(x), f(y)) \leqslant \omega(d(x,y)).$$

证明，f 具有唯一的不动点。

题目 U（Hamilton-Cayley 定理） 假设 $A \in \mathbf{M}_n(\mathbb{C})$ 是 $n \times n$ 的复矩阵，$P(X)$ 是它的特征多项式，即

$$P(X) = \det(X\mathbf{I}_n - A).$$

U1) 证明，如果 A 可对角化，那么 $P(A) = 0$ 是零矩阵（将多项式中的不定元 X 换成 A）。

U2) 证明，映射

$$\mathbf{M}_n(\mathbb{C}) \to \mathbf{M}_n(\mathbb{C}), \quad A \mapsto P(A),$$

是连续映射。

U3) 证明，对任意的 $A \in \mathbf{M}_n(\mathbb{C})$，对任意的 $\varepsilon > 0$，总存在 $B \in \mathbf{M}_n(\mathbb{C})$，使得

- $\|B\|_2 < \varepsilon$；
- $A + B$ 可被对角化。

U4) 证明 Hamilton-Cayley 定理：对任意的 $A \in \mathbf{M}_n(\mathbb{C})$，我们有 $P(A) = 0$。

题目 V 考虑 $E = C(\mathbb{R})$，$F = C^\infty(\mathbb{R})$。

V1) 试找出所有的 E 的 $\mathbb{R}$-线性子空间 V，使得对任意的 $f_1, f_2 \in V$，都有 $f_1 f_2 \in V$。

V2)（求导数的平方根）求导数运算 $\dfrac{d}{dx} : F \to F$ 是线性映射。是否存在线性映射 $T : F \to F$，使得 $T \circ T = \dfrac{d}{dx}$？

题目 W（e 的超越性） 假设 $P(X)$ 是一个 n-次的实数系数的多项式。我们定义

$$I(t) = \int_0^t e^{t-x} P(x) dx.$$

W1) 证明，$I(t) = e^t \sum_{i=0}^{n} P^{(i)}(0) - \sum_{i=0}^{n} P^{(i)}(t)$。

W2) 假设存在整数 $a_0, a_2, \cdots, a_n$，$a_0 \neq 0$，使得

$$a_0 + a_1 e + a_2 e^2 + \cdots + a_n e^n = 0.$$

对于 $p \in \mathbb{Z}_{\geqslant 0}$，我们选取多项式

$$P(X) = X^{p-1}(X-1)^p (X-2)^p \cdots (X-n)^p$$

并定义

$$J = a_0 I(0) + a_1 I(1) + \cdots + a_n I(n).$$

证明，J 是整数并且 $(p-1)!$ 整除 J。

W3) 证明，如果 p 是素数并且足够大，那么 $J \neq 0$。从而，$J \geqslant (p-1)!$

W4) 证明，存在 $C > 0$，使得对一切 $p \in \mathbb{Z}_{\geqslant 1}$，我们都有 $|J| \leqslant C^p$。

W5) 证明 e 是超越数，即不存在整数 $a_0, a_2, \cdots, a_n$，$a_0 \neq 0$，使得

$$a_0 + a_1 e + a_2 e^2 + \cdots + a_n e^n = 0.$$

题目 X 给定 $3n$ 个实数 $\alpha_1,\cdots,\alpha_n,\beta_1,\cdots,\beta_n,\gamma_1,\cdots,\gamma_n$，其中，每个 β_i 都不是 0。证明，函数

$$f(x)=\sum_{k=1}^{n}\alpha_k\cos(\beta_k x+\gamma_k)$$

在 $\mathbb{R}$ 上有无限个零点。

题目 Y 假设 $\lambda_1,\cdots,\lambda_{2019}$ 是两两不同的实数，$c_1,\cdots,c_{2019}$ 是实数，满足

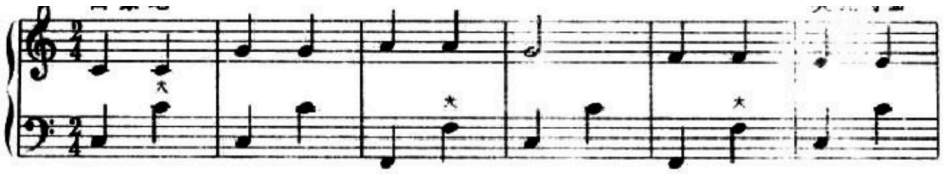
$$\lim_{x\to+\infty}\left(c_1e^{i\lambda_1 x}+c_2e^{i\lambda_2 x}+\cdots+c_{2019}e^{i\lambda_{2019}x}\right)=0.$$

证明，$c_1=c_2=\cdots=c_{2019}=0$。

题目 Z 是否能够构造一个实数的序列 $\{a_k\}_{k\geqslant 1}$，使得级数

$$a_1^\ell+a_2^\ell+\cdots+a_n^\ell+\cdots$$

在 $\ell=5$ 时是发散的并且当 ℓ 为其它正的奇数时是收敛的?



简介

数学分析二的内容分三部分：

1) $\mathbb{R}^n$ 上的微分学

   - 方向导数与微分；
   - 反函数定理与隐函数定理；
   - $\mathbb{R}^n$ 中子流形的刻画。

2) 抽象的积分学与 $\mathbb{R}^n$ 上的积分

   - 抽象测度空间的积分理论；
   - $\mathbb{R}^n$ 上的积分；
   - 变量替换公式与 Stokes 公式。

3) Fourier 级数及其应用

   - Fourier 级数的收敛理论；
   - 用频率的观点看函数；
   - Fourier 级数的应用：等分布问题与 Roth 三项等差数列定理。

30 方向导数，偏导数和微分，利用偏导数连续判断微分存在，导数与极值

二零二零年二月十七日，星期一，晴

高维的微分学

在人们谈论高维的 $\mathbb{R}^n$ 的时候，经常会有如下感觉：

- $\mathbb{R}^n$ 上的每个点就是一个数组 $(x_1, x_2, \cdots, x_n)$；
- 两个不同场合的 $\mathbb{R}^n$ 是同一个 $\mathbb{R}^n$。

我们对此加以说明（上学期我们讲过，所谓的空间是一个集合加上一个结构）：令 $\mathbb{R}$ 是实数的集合（上学期我们证明了它的存在性（唯一，在同构的意义下），我们在这学期课程中**假设取定**这样的一个实数域），我们定义

$$\mathbb{R}^n = \underbrace{\mathbb{R} \times \mathbb{R} \times \cdots \times \mathbb{R}}_{n个}.$$

按照这个定义，$\mathbb{R}^n$ 是按照确定的方式定义的（所以是唯一的）。特别地，按照集合乘积的定义，$x \in \mathbb{R}^n$ 就是一个数组 $(x_1, x_2, \cdots, x_n)$，其中 $x_i \in \mathbb{R}$。特别地，我们在 $\mathbb{R}^n$ 上指定了一个特定的坐标系统 $(x_1, \cdots, x_n)$。

从函数的观点来看，给定 $\mathbb{R}^n$ 上的一个点，它的第 i 个坐标定义出 $\mathbb{R}^n$ 到 $\mathbb{R}$ 的一个（坐标）映射（函数）：

$$\pi_i : \mathbb{R}^n \to \mathbb{R}, \quad (x_1, \cdots, x_n) \mapsto x_i.$$

注记. 微分学的一个核心话题是如何利用其它的坐标系统 $(y_1, \cdots, y_n)$ 来描述 $\mathbb{R}^n$ 上的光滑／可微函数。其中，所谓一个新的坐标系统 $(y_1, \cdots, y_n)$ 目前可以简单地想象为 n 个函数

$$y_i : \mathbb{R}^n \to \mathbb{R}, \quad (x_1, x_2, \cdots, x_n) \mapsto y_i(x_1, x_2, \cdots, x_n),$$

其中 $i = 1, 2, \cdots, n$。我们要求把它们放在一起得到的映射

$$\mathbb{R}^n \to \mathbb{R}^n, (x_1, \cdots, x_n) \mapsto \big(y_1(x_1, \cdots, x_n), \cdots, y_n(x_1, \cdots, x_n)\big).$$

是一个双射。

在这个学期的课程中，如果不另加说明，我们总假设 Ω 是 $\mathbb{R}^n$ 中的一个开集，有时候我们把它称作是一个**（开）区域**。

上学期的课程中，我们对定义在 $\mathbb{R}^1$ 上的函数定义了导数的概念（如果存在）。我们做简单的回忆：假设 f 是定义在区间 $I \subset \mathbb{R}$ 上的实值函数，$x_0 \in I$ 是一个给定的点，如果极限

$$\lim_{h \to 0} \frac{f(x_0 + h) - f(x_0)}{h}$$

存在，我们就用 $f'(x_0)$ 表示这个极限并称它为 f 的导数。我们还定义了 f 在点 x_0 处的微分 $df(x_0)$，这是一个线性映射（我们暂且不管它定义）。这两个概念都可以对 $\Omega \subset \mathbb{R}^n$ 上的函数来定义。在给出推广之前，我们必须要指出：$df(x_0)$ 是比 $f'(x_0)$ 更好的概念，因为它不依赖于坐标系统的选取。

导数在高维正确的推广是方向导数：

定义 175 (方向导数)**.** 给定函数 $f:\Omega\to\mathbb{R}$，$x_0\in\Omega$，$v\in\mathbf{R}^n$。如果下面的极限

$$\lim_{h\to 0}\frac{f(x_0+hv)-f(x_0)}{h}$$

存在，我们就称函数 f 在 x_0 处沿 v 的**（方向）导数**存在，并将它记作

$$\left(\nabla_v f\right)(x_0)=\lim_{h\to 0}\frac{f(x_0+hv)-f(x_0)}{h}.$$

习惯上，我们还把 $\nabla_v f$ 写成 $\dfrac{\partial f}{\partial v}$。

注记. 在这个定义中，我们并没有使用 $\mathbb{R}^n$ 上的坐标系。特别地，如果 $f:V\to\mathbb{R}$ 是一个赋范线性空间上的函数，我们可以同样的对 $v\in V$ 定义方向导数。

如果使用坐标系，我们就有几个特殊的向量，比如说 $e_i=(0,\cdots,0,1,0,\cdots,0)$（第 i 个位置为 1），我们约定用如下的符号代表这个向量

$$\frac{\partial}{\partial x_i}=\underbrace{(0,\cdots,0,1,0,\cdots,0)}_{\text{第 } i \text{ 个位置上为 } 0,\ \text{其余为 } 1}.$$

此时，我们用下面的符号表示方向导数：

$$\frac{\partial f}{\partial x_i}(x_0)=\left(\nabla_{\frac{\partial}{\partial x_i}} f\right)(x_0).$$

习惯上，我们称它为**偏导数**。

注记. 我们假设对任意的 $v\in\mathbb{R}^n$，f 在 x_0 处的方向导数都存在。此时，我们有映射（这个“几乎”就是微分的定义）：

$$\mathbb{R}^n\to\mathbb{R},\quad v\mapsto\left(\nabla_v f\right)(x_0).$$

这个映射是与 $\mathbb{R}^n$ 上的线性结构相容的，即这是一个线性映射：

- 对任意的 $\lambda\in\mathbb{R}$，我们有

  $$\left(\nabla_{\lambda v} f\right)(x_0)=\lambda\left(\nabla_v f\right)(x_0).$$

  特别地，$\nabla_0 f=0$。

  只要按照定义验证即可。

- 对任意的 $v, w \in \mathbb{R}^n$，我们希望有

$$(\nabla_{v+w} f)(x_0) = (\nabla_v f)(x_0) + (\nabla_w f)(x_0).$$

这个命题需要多一点的条件才可以证明，仅仅用每个方向导数存在是不够的。这个性质对我们将要证明的可微函数总是成立的。

例子. 我们有两个基本的例子

1) 假设函数 $f : \mathbb{R}^2 \to \mathbb{R}$ 只依赖于 x-坐标，我们习惯上将它写成 $f(x, y) = f(x)$，那么，

$$\frac{\partial f}{\partial y} \equiv 0.$$

为此，只需要按定义计算即可：

$$\frac{\partial f}{\partial y} = \lim_{h \to 0} \frac{f(x, y+h) - f(x)}{h} = 0.$$

2) 考虑 $\mathbb{R}^2$ 上定义的函数

$$f(x, y) = \begin{cases} \frac{xy}{x^2+y^2}, & (x, y) \neq (0, 0) \\ 0, & (x, y) = (0, 0). \end{cases}$$

那么，$f(x, y)$ 的两个偏导数 $\frac{\partial f}{\partial x}$ 和 $\frac{\partial f}{\partial y}$ 均为 0。另外，如果 $v = (v_1, v_2)$ 其中 $v_1 v_2 \neq 0$，那么

$$\nabla_v f(0) = \lim_{h \to 0} \frac{v_1 v_2}{h(v_1^2 + v_2^2)}$$

是不存在的。由此可见，偏导数存在不能保证其它的方向导数存在。

注记 (方向导数的几何解释)**.** 方向导数本质上是一维的导数，这个从定义的写法本身就不难看出。假设 $I = (-a, a)$（$a > 0$）是一个区间，假设

$$\gamma : I \to \mathbb{R}^n, \quad t \mapsto \gamma(t) = (\gamma_1(t), \gamma_2(t), \cdots, \gamma_n(t)),$$

是 C^1-的映射（上学期证明过这等价于 γ 的每个分量都是 C^1 的）。我们将这样的映射 γ 称作是 $\mathbb{R}^n$ 中的一条**曲线**。对于这样的曲线，我们称 $\gamma'(0) \in \mathbb{R}^n$ 是 γ 在 $\gamma(0)$ 处的**切向量**。我们要强调的是每个切向量都与基准点$\gamma(0)$ 相关并且我们认为不同点上的切向量是不相关的。

1) 最基本的曲线是直线：假设 $x_0 \in \Omega$，$v \in \mathbb{R}^n$，考虑曲线

$$\ell_v : (-a, a) \to \Omega, \ t \to x_0 + tv.$$

这是 Ω 中过 x_0 点以 v 为切向量的一段直线。我们记 $L = \ell((-\varepsilon, \varepsilon))$，这是直线这个**几何对象**。我们应该注意区分下面的数学对象：ℓ 的像是“直线”这个几何对象，但是我们用（区间 $(-a, a)$ 是 ℓ 的定义域）ℓ 来参数化这个对象。

现在给定函数 $f:\Omega\to\mathbb{R}$，我们用 $f\big|_L$ 表示 f 在 L 上的限制。那么，$(\nabla_v f)(x_0)$ 就是函数

$$\left(f\big|_L\right)\circ\ell : I\to\mathbb{R}$$

在 0 处的导数，即

$$\frac{d}{dt}\Big|_{t=0}\left(\left(f\big|_L\right)\circ\ell\right)=\left(\nabla_{\ell'(0)}f\right)(\ell(0)).$$

当然，我们可以采取别的方式对 L 进行参数化，比如说

$$\ell' : (-\frac{a}{\lambda},\frac{a}{\lambda})\to\mathbb{R}^n, t\mapsto x_0+\lambda t v.$$

它的像 $\mathrm{Im}(\ell')$ 也是 $\mathrm{Im}(\ell)$，但是 ℓ' 的在 $\ell'(0)=\ell(0)$ 处的切向量是 $\lambda\cdot v$。所以参数化的改变可能使得曲线的切向量发生改变。

2) 更一般的，给定 C^1-曲线

$$\gamma:(-a,a)\to\mathbb{R}^n,\quad t\mapsto\gamma(t).$$

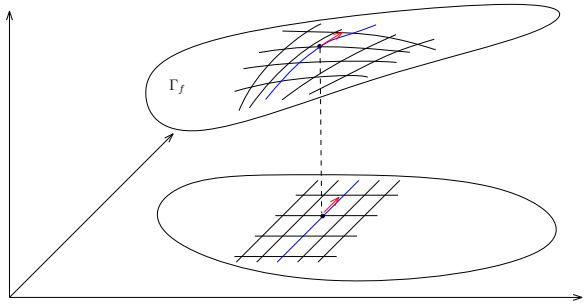
我们令 $C=\mathrm{Im}(\gamma)$，那么

$$\frac{d}{dt}\Big|_{t=0}\left(\left(f\big|_C\right)\circ\gamma\right)=\left(\nabla_{\gamma'(0)}f\right)(\gamma(0)).$$

这个结论的证明我们留作作业。

3) 有一类曲线的例子尤为重要，它们可以用来刻画高维空间中的曲面。假设函数 $f:\Omega\to\mathbb{R}$ 的所有偏导数 $\dfrac{\partial f}{\partial x_i}$ 在 Ω 上都有定义。我们考虑它的图像

$$\Gamma_f=\{(x,f(x))\in\Omega\times\mathbb{R}\subset\mathbb{R}^{n+1}|x\in\Omega\}.$$



这是 $\mathbb{R}^{n+1}$ 中的一张超曲面。给定 $x=(x_1,\cdots,x_n)\in\Omega$，我们就给定了 Γ_f 上的一个点；反之亦然，以为我们只要取 Γ_f 上的点的前 n 个坐标就给出了 Ω 上的点。我们可以形象地认为 Ω 坐标网通过 $x\mapsto(x,f(x))$ 给出了 Γ 上的一个坐标网。固定点 $p=(p_1,\cdots,p_n)$，我们考虑 Ω 上的曲线（直线）：

$$\gamma_k:(-a,a)\to\Omega,\quad t\mapsto(p_1,p_2,\cdots,p_{k-1},p_k+t,p_{k+1},\cdots,p_n).$$

这是过 p 点的 x_k 这个坐标轴的参数表示。我们可以把这条曲线提升到 Γ_f 上，从而得到

$$\widetilde{\gamma}_k : (-a, a) \to \mathbb{R}^{n+1}, \quad t \mapsto (\gamma_k(t), f(\gamma_k(t))) .$$

它们在图中对应的是两条蓝色的曲线。

按照定义，第一条曲线在 p 处的切向量为 $e_k = \dfrac{\partial}{\partial x_k}$，我们在图中用红色表示；第二条曲线在 $(p, f(p))$ 处的切向量为

$$E_k = \frac{\partial}{\partial x_k} + \frac{\partial f}{\partial x_k}(p)\frac{\partial}{\partial x_{n+1}}.$$

如果我们用 $T_{(p,f(p))}\Gamma_f$ 表示该曲面在点 $(p, f(p))$ 处的切空间（暂且不管如何定义），那么，这些 E_k 应该恰好张成这个切空间。

这里，我们通过曲线提升的方式，计算了一个曲面的切空间（利用了偏导数)。这类作为函数的图像出现的超曲面是多元微积分中最基本的几何对象，我们在后面的课程中会无限次见到它们。

我们现在来定义函数的微分：

定义 176 (函数的微分)**.** 给定函数 $f : \Omega \to \mathbb{R}$ 和 $x_0 \in \Omega$。如果存在 $\mathbb{R}$-线性映射 $A : \mathbb{R}^n \to \mathbb{R}$，使得对于 $v \to 0$ 时，其中 $v \in \mathbb{R}^n$，我们有

$$f(x_0 + v) = f(x_0) + A(v) + o(v),$$

也就是说

$$\lim_{v\to 0} \frac{|f(x_0 + v) - f(x_0) - A(v)|}{|v|} = 0,$$

其中向量的长度 $|v|$ 可以选取任意的范数（比如勾股定理所定义的 $\|\cdot\|_2$)，我们就称 f 在 x_0 处可微并且称线性映射 A 是 f 在 x_0 的**微分**。我们通常将 A 写成下面的样子：

$$df\big|_{x=x_0} = df(x_0) : \mathbb{R}^n \to \mathbb{R}.$$

如果 f 在 Ω 的每个点处都可微，我们就称 f 是 Ω 上的**可微函数**。

当 $n = 1$ 时，我们上个学期已经证明了，如果 f 在 x_0 处的导数存在，那么 $df(x_0)$ 也存在并且它可以写成

$$df(x_0) : \mathbb{R} \to \mathbb{R}, \quad v \mapsto f'(x_0) \cdot v,$$

其中 $\cdot$ 是一个数乘以一个 1 维的向量。

另外，根据定义，如果 f 在 $x_0 \in \Omega$ 处可微，那么 f 在 x_0 处连续。这几乎是显然的：

$$f(x_0 + v) - f(x_0) = A(v) + o(v) = o(1).$$

注记. 在微分的定义中，我们根本就没有用到 $\mathbb{R}^n$ 上的坐标系：我们只用到了 v 的长度的概念。所以，可以自然地对在一个赋范线性空间 $(V,\|\cdot\|)$ 上定义的函数来定义其微分：给定函数 $f:V\to\mathbb{R}$ 和 $x_0\in V$，其中 $(V,\|\cdot\|)$ 是一个（$\mathbb{R}$ 或者 $\mathbb{C}$ 上的）赋范线性空间。如果存在 $\mathbb{R}$（或者 $\mathbb{C}$）-线性映射 $A:V\to\mathbb{R}$，使得

$$\lim_{v\to 0}\frac{|f(x_0+v)-f(x_0)-A(v)|}{\|v\|}=0,$$

我们就称 f 在 x_0 处可微并且称 A 是 f 在 x_0 的微分，我们还把 A 记作 $df\big|_{x=x_0}=df(x_0)$。

命题 177 (微分的计算：微分与方向导数之间的关系)**.** 假设 $f:\Omega\to\mathbb{R}$ 在 x_0 处可微，那么 f 在 x_0 的任意方向导数都存在。特别地，对于 $v=(v_1,\cdots,v_n)=\sum_{i=1}^n v_i\frac{\partial}{\partial x_i}$（请回忆：我们约定 $\frac{\partial}{\partial x_i}$ 代表向量 $\underbrace{(0,\cdots,1,\cdots,0)}_{\text{第 } i \text{ 个位置上为 } 1}$），那么

$$df(x_0)(v)=(\nabla_v f)(x_0)=\sum_{j=1}^n\frac{\partial f}{\partial x_j}(x_0)v_j.$$

这也表明映射

$$\mathbb{R}^n\to\mathbb{R},\quad v\mapsto(\nabla_v f)(x_0)$$

是线性映射。

证明：任意给定 $v\in\mathbb{R}^n$，令 $h\to 0$，根据微分的定义，我们有

$$f(x_0+hv)-f(x_0)=(df(x_0))(hv)+o(hv).$$

所以，我们可以利用定义来计算方向导数

$$(\nabla_v f)(x_0)=\lim_{h\to 0}\frac{h\,(df(x_0))(v))+o(hv)}{h}=\big(df(x_0)\big)(v).$$

这也表明 $\mathbb{R}^n\to\mathbb{R},\ v\mapsto(\nabla_v f)(x_0)$ 是线性。为了用偏导数具体地计算 $df(x_0)$，根据上面的结果，我们有

$$\big(df(x_0)\big)(v)=(\nabla_v f)(x_0)=\sum_{i=1}^n v_i(\nabla_{\frac{\partial}{\partial x_i}}f)(x_0)=\sum_{j=1}^n\frac{\partial f}{\partial x_j}(x_0)v_j.$$

命题得证。 □

注记. 在微积分学习中，最有歧义的一个数学符号是所谓的 dx。现在我们用微分的语言定义 dx_i。假定在 $\mathbb{R}^n$ 上我们事先选好了坐标系 $(x_1,\cdots,x_n)$。此时，

$$\pi_i:\mathbb{R}^n\to\mathbb{R},\quad(x_1,\cdots,x_n)\mapsto x_i$$

是一个函数，传统上我们就把 π_i 写成 x_i，所以当给定了一个点 $p \in \mathbb{R}^n$ 之后，$dx_i(p)$ 指代的就是线性映射 $d\pi_i(p)$。很明显，通过定义线性映射

$$A : \mathbb{R}^n \to \mathbb{R}, \quad \frac{\partial}{\partial x_j} \mapsto \begin{cases} 0, & i \neq j; \\ \dfrac{\partial}{\partial x_i}, & i = j. \end{cases}.$$

这就是 $d\pi_i$ 在某个点 p 处的微分。

用传统的 *Kronecker* 符号来记，我们有

$$\big((dx_i)(p)\big)\big(\frac{\partial}{\partial x_j}\big) = \delta^i_j.$$

我们之前证明了如果函数的微分存在，那么方向导数也存在。反过来并不成立，实际上，下面的例子表明 f 甚至可以不连续，尽管其方向导数都存在：

例子. 考虑函数在 $\mathbb{R}^2$ 定义函数：

$$f(x,y) = \begin{cases} \dfrac{y^2}{x}, x \neq 0 \\ 0, \quad x = 0. \end{cases}$$

那么，f 在 $(0,0)$ 处的任意一个方向导数都存在：假设 $v = (v_1, v_2) \in \mathbb{R}^2$ 且 $v \neq 0$。那么，我们有

1) 如果 $v_1 = 0$，那么

$$\nabla_v f(0) = \lim_{t \to 0} \frac{0-0}{t} = 0.$$

2) 如果 $v_1 \neq 0$，那么

$$\nabla_v f(0) = \lim_{t \to 0} \frac{t\frac{v_2^2}{v_1}}{t} = \frac{v_2^2}{v_1}.$$

然而，f 在 $(0,0)$ 处不可微，这因为 f 在 $(0,0)$ 不连续：可以选取点列 $\left\{\left(\frac{1}{n^2}, \frac{1}{n}\right)\right\}_{n \geqslant 1}$，函数在这些点上取值不收敛到 0。

这个例子也说明，只考虑方向导数并不能对函数在一个点附近的行为有效地进行控制。

然而，如果方向导数的具有连续性，情况就大不相同。直观上，连续性允许我们从一点的信息出发理解这点附近的情况。

命题 178. 给定开区域 $\Omega \subset \mathbb{R}^n$ 和 $x_0 \in \Omega$，f 是在 Ω 上定义的函数。假设 f 的所有偏导数 $\dfrac{\partial f}{\partial x_i}$ 在 x_0 的附近（不妨设在 Ω 上）存在并且 $\dfrac{\partial f}{\partial x_i}$ 均为连续函数，其中 $i = 1, 2, \cdots, n$，那么 f 在 x_0 处可微。

这个命题可以非常方便地用来判断一个函数的可微性，比如说，考虑 $\mathbb{R}^3$ 上的函数

$$f(x,y,z) = 3y + e^{yz},$$

它的偏导数很容易计算并且很明显是连续的。所以，f 是可微的。

证明: 假设函数 f 在 p 处的微分存在，那么它可以用偏导数来表示，即

$$df(x_0)v = \sum_{j=1}^{n} \frac{\partial f}{\partial x_j}(x_0)v_j, \quad v \in \mathbb{R}^n.$$

据此，我们定义线性映射

$$D : \mathbb{R}^n \to \mathbb{R},\ v \mapsto \sum_{j=1}^{n} \frac{\partial f}{\partial x_j}(x_0)v_j.$$

我们只要证明 $r(v) = f(x_0 + v) - f(x_0) - D(v)$ 是一个 $o(v)$ 项即可，其中 $v \to 0$。我们将 x_0 用坐标写成

$$x_0 = (x_{0,1}, x_{0,2}, \cdots, x_{0,n}).$$

按定义，我们有

$$\begin{aligned}
r(v) &= f((x_{0,1} + v_1, \cdots, x_{0,n} + v_n)) - f((x_{0,1}, \cdots, x_{0,n})) - \sum_{j=1}^{n} \frac{\partial f}{\partial x_j}(x_0)v_j \\
&= f((x_{0,1} + v_1, x_{0,2}, \cdots, x_{0,n})) - f((x_{0,1}, \cdots, x_{0,n})) - \frac{\partial f}{\partial x_1}(x_0)v_1 \\
&\quad + f((x_{0,1} + v_1, x_{0,2} + v_2, x_{0,3}, \cdots, x_{0,n})) - f((x_{0,1} + v_1, x_{0,2}, \cdots, x_{0,n})) - \frac{\partial f}{\partial x_2}(x_0)v_2 \\
&\quad + \cdots \\
&\quad + f((x_{0,1} + v_1, \cdots, x_{0,n} + v_n)) - f((x_{0,1} + v_1, \cdots, x_{0,n-1} + v_{n-1}, x_{0,n})) - \frac{\partial f}{\partial x_n}(x_0)v_n \\
&= \sum_{j=1}^{n} \Big[\underbrace{\big(f(\cdots, x_{0,j} + v_j, x_{0,j+1}, \cdots) - f(\cdots, x_{0,j-1} + v_{j-1}, x_{0,j}, x_{0,j+1}, \cdots)\big)}_{\text{Lagrange 中值定理}} - \frac{\partial f}{\partial x_j}(x_0)v_j \Big] \\
&= \sum_{j=1}^{n} \frac{\partial f}{\partial x_j}(\widetilde{x}_j)v_j - \frac{\partial f}{\partial x_j}(x_0)v_j.
\end{aligned}$$

在上面我们运用 Lagrange 中值定理的时候，假设出了第 i 个位置的其它变量都是固定的，我们知道，此时求偏导数的运算就是 1 维情形时的导数运算，所以可以用中值定理。换句话说，我们在以 $(x_{0,1} + v_1, \cdots, x_{0,j-1} + v_{j-1}, x_{0,j}, \cdots, x_{0,n})$ 和 $(x_{0,1} + v_1, \cdots, x_{0,j} + v_j, x_{0,j}, \cdots, x_{0,n})$ 为端点的线段上用 Lagrange 中值定理，其中 $\widetilde{x}_j$ 是在这个线段上的一个点。当 $v \to 0$ 时，很明显 $\widetilde{x}_j \to x_0$。上面的计算表明

$$|r(v)| \leqslant \sum_{j=1}^{n} \Big| \frac{\partial f}{\partial x_j}(\widetilde{x}_j) - \frac{\partial f}{\partial x_j}(x_0) \Big| |v|.$$

根据连续性，$r(v) = o(v)$。 □

注记. 上面证明的关键想法是把问题转化成在某条直线或者曲线上讨论（维数 $= 1$，当然，我们还没有定义什么是维数），从而可以利用一元微分学的结论，这是处理多变量函数最基本的想法之一。上个学期我们证明两点之间线段最短的时候就是用的这个想法，那个场合研究的空间是所有曲线的空间，维数甚至是无穷！

我们再看上面想法的一个应用：这个例子对于实际应用非常的重要，因为我们可以用来计算多元函数的最大最小值：

命题 179. *假设 $f:\Omega\to\mathbb{R}$ 在 x_0 处可微并且 x_0 是 f 在 Ω 上的最大，那么 $df(x_0)=0$。*

证明： 任意给定方向 $v\in\mathbb{R}^m$，我们考虑通过 x_0 的直线段

$$\gamma:(-\varepsilon,\varepsilon)\to\Omega,\quad t\mapsto x_0+tv,$$

其中 ε 足够小使得线段本身落在 Ω 中，我们考虑 $f\circ\gamma$。简单而言，研究 $(-\varepsilon,\varepsilon)$ 上的函数 $g(t)=f(x_0+tv)$。之前的计算表明：

$$\frac{d}{dt}\Big|_{t=0}f\circ\gamma=\nabla_v f(x_0)=df(x_0)(v).$$

很明显，x_0 是 $f\circ\gamma$ 的最大值，所以 $(f\circ\gamma)'(0)=0$，这表明对任意的 $v\in\mathbb{R}^n$，$df(x_0)(v)=0$。所以，作为线性映射，$df(x_0)=0$。 □

另一个有趣的应用如下：

命题 180. *假设 Ω 是一个开的凸集[7]（可以替换为连通性），$f:\Omega\to\mathbb{R}$ 是可微函数。如果对任意的 $x\in\Omega$，我们都有 $df(x)=0$，那么 f 是常数。*

证明： 证明我们留作作业。基本想法如下：固定 $x_0\in\Omega$，对任意的 $x\in\Omega$，我们考虑 x_0 到 x 的线段并将 f 限制在这条线段上。这就可以帮助我们走到 1 维的情况。 □

我们引入一些符号来结束本次的课程：

注记. *给定可微 $f:\Omega\to\mathbb{R}$，对于任意的 $x\in\Omega$，$df(x)$ 都是 $\mathbb{R}^n$ 上的一个线性函数，即*

$$df(x):\mathbb{R}^n\to\mathbb{R}.$$

为了强调对 x 的依赖性，我们把 $df(x)$ 的定义域这个 $\mathbb{R}^n$ 用 $T_x\Omega$ 来表示。换句话说，$df(x)$ 是一族线性函数，他们的定义域随着空间点位置的改变而改变，这不是我们通常意义上的函数。在课程的后面，我们会考虑 $\Omega\subset\mathbb{R}^n$ 是曲面（子流形）的情形，那时这些讨论会变得异常清晰。

[7] Ω 是凸集指的是对任意的 $p,q\in\Omega$，连接它们的线段整个都落在 Ω 中。

31 映射的微分，利用偏导数表示微分，Jacobi 矩阵，微分的链式法则，逆映射微分的计算，矩阵代数中指数映射的微分

二零二零年二月二十日，星期四，阴天

我们上次课最后讲到一个函数的高阶导数是比较难定义的。

In the fall of 1972 President Nixon announced that the rate of increase of inflation was decreasing. This was the first time a sitting president used the third derivative to advance his case for re-election.

– Hugo Rossi

给定函数 $f:\Omega\to\mathbb{R}$，我们上次课定义它的微分 $df(x)$。当 x 给定的时候，这是一个 $\mathbb{R}^n$ 上的线性函数。仿照 1 维的理论，对 1 阶导数求导就应该得到 2 阶导数。然而，此时为了定义 f 的二阶导数，我们就需要对 $x\mapsto df(x)$ 求微分。此时，（不严格地讲）映射 $f:x\mapsto df(x)$ 是一个从 Ω 到 $\mathrm{Hom}(\mathbb{R}^n,\mathbb{R})\simeq\mathbb{R}^n$ 的映射，这是向量值的函数。所以，我们自然地想到对向量值的函数 $f:\Omega\to\mathbb{R}^m$ 定义微分。当然，我们还可以考虑 $df(x)$ 的每个分量（依赖于坐标系的选取）然后要求每个分量都可微，至少从形式上看，后一种做法失于简洁和美观。

定义 181 (微分)**.** 假设 Ω 是 $\mathbb{R}^n$ 中的区域，Ω' 是 $\mathbb{R}^m$ 中的区域，给定映射 $f:\Omega\to\Omega'$。如果存在线性映射

$$df\big|_{x=x_0}=df(x_0):\mathbb{R}^n\to\mathbb{R}^m,$$

使得对于 $\mathbb{R}^n$ 中的 $v\to 0$ 时，我们有

$$f(x_0+v)=f(x_0)+df(x_0)v+o(v),$$

即

$$\lim_{v\to0}\frac{|f(x_0+v)-f(x_0)-df(x_0)v|}{|v|}=0,$$

我们就称 f 在 x_0 处可微并且称线性映射 $df(x_0)$ 是 f 在 x_0 的**微分**。如果 f 在 Ω 的每个点处都可微，我们就称 f 是 Ω 上面的**可微映射**。

注记. *1)* 在上述微分的定义中，我们完全没有用到 $\mathbb{R}^n$ 和 $\mathbb{R}^m$ 上的坐标系。实际上，在下面的极限中

$$\lim_{v\to0}\frac{|\overbrace{f(\underbrace{x_0+v}_{\text{用到 }\mathbb{R}^n\text{ 上的加法结构}})-f(x_0)-df(x_0)v}^{\text{用到 }\mathbb{R}^m\text{ 上的加法结构}}|}{|v|}=0,$$

我们只用到了 $\mathbb{R}^n$ 和 $\mathbb{R}^m$ 上的线性结构和它们上面的范数（我们分别用了蓝色和红色表示，其中红色的是定义域 $\mathbb{R}^n$ 上的范数，蓝色的是值域 $\mathbb{R}^m$ 上的范数）。

据此，我们可以将上述定义进行推广：给定赋范线性空间 $(V_1, \|\cdot\|_1)$ 和 $(V_2, \|\cdot\|_2)$，$\Omega_1 \subset V_1$ 和 $\Omega_2 \subset V_2$ 是非空的开集，$f : \Omega_1 \to \Omega_2$ 是映射。如果存在线性映射

$$df\big|_{x=x_0} = df(x_0) : V_1 \to V_2,$$

使得对于 V_1 中的 $v \to 0$ 时，我们有

$$\lim_{v\to 0} \frac{\|f(x_0+v) - f(x_0) - df(x_0)v\|_2}{\|v\|_1} = 0,$$

我们就称 f 在 x_0 处可微并且称线性映射 $df(x_0)$ 是 f 在 x_0 的**微分**。

2) 假设 f 在 Ω 上可微。那么，给定 $x \in \Omega$，$df(x) \in \mathrm{Hom}(\mathbb{R}^n, \mathbb{R}^m) \approx \mathbb{R}^{mn}$（可以视作是 mn 的矩阵，这里我们用坐标比较方便）也在一个向量空间中取值的。（然而，如果在一般的（无限维的）的赋范线性空间上定义微分，我们会要求 $df(x) \in \mathrm{Hom}(V, W)$ 是所谓的连续线性映射，这里不展开讨论，有兴趣的同学可以在泛函分析的课程上学习）。所以，当 x 变化的时候，我们就得到一个映射

$$\Omega \to \mathbb{R}^{nm}, \quad x \mapsto df(x).$$

我们可以对它求导数来定义它的微分。高阶的微分不是这门课程的重点。

3) 假设 f 在 $x \in \Omega$ 处可微分，我们就有映射

$$df(x_0) : \mathbb{R}^n \to \mathbb{R}^m.$$

由于这些映射依赖于点，特别地，依赖于 $x_0 \in \Omega$（和 $f(x_0) \in \Omega'$），我们用 $T_{x_0}\Omega$ 代表它的定义域的线性空间 $\mathbb{R}^n$，用 $T_{f(x_0)}\Omega'$ 代表它的定义域的线性空间 $\mathbb{R}^m$，这样子，我们形式上就有

$$df(x_0) : T_{x_0}\Omega \to T_{f(x_0)}\Omega'.$$

符号 $T_{x_0}\Omega$ 代表的是 Ω 在 x_0 处的切空间（＝切平面），$T_{f(x_0)}\Omega'$ 代表的是 Ω' 在 $f(x_0)$ 处的切空间，我们会有专门的例子来理解这个对象，目前大家可以暂时将它们理解为好的记号。

我们上次课定义了方向导数和偏导数，这都是 1 维的对象。下面的命题表明，我们可以用偏导数这些 1 维的对象来描述 $df(x_0)$ 这个高维的对象：

命题 182 (微分的计算)**.** 假设 $V = \mathbb{R}^n$ 和 $W = \mathbb{R}^m$，我们在 V 上用坐标系 $\{x_i\}_{i=1,\cdots,n}$，在 W 上用坐标系 $\{y_j\}_{j=1,\cdots,m}$（把空间写成 V 和 W 是强调这些空间可以不用具体的坐标来描述）。考虑 $f : V \to W$（我们也可以考虑 f 定义在 V 中某个区域上），用坐标来写，我们有：

$$x \mapsto f(x) = \big(f_1(x_1, \cdots, x_n), f_2(x_1, \cdots, x_n), \cdots, f_m(x_1, \cdots, x_n)\big).$$

有时候还写成

$$y_1 = f_1(x_1, \cdots, x_n),\ y_2 = f_2(x_1, \cdots, x_n), \cdots, y_m = f_m(x_1, \cdots, x_n).$$

那么，我们有

1) 假设 f 在 x_0 处可微，那么每个分量函数 f_j 在 x_0 处都可微，其中 $j=1,2,\cdots,m$,。

2) 如果每个分量函数 f_j 在 x_0 处都可微（其中 $j=1,2,\cdots,m$），那么 f 在 x_0 处可微。

特别地，如果 f 在 x_0 处可微，那么映射 $df(x_0):\mathbb{R}^n\to\mathbb{R}^m$ 可以用 $n\times m$ 的矩阵

$$\left(\frac{\partial f_i}{\partial x_j}(x_0)\right)_{\substack{i=1,\cdots,m\\ j=1,\cdots,n}}$$

来表示（我们将这个矩阵称作是 f 在 x 处的 **Jacobi 矩阵**，并记作 $\mathrm{Jac}(f)$ 或者 $\mathbf{J}(f)$，它只是微分在一个特殊的坐标系下的表达）。

证明: 我们首先证明，f 在 x_0 可微等价于每个分量 f_j（$j=1,\cdots,m$）都可微。假设 f 在 x_0 处可微，此时 $df:\mathbb{R}^n\to\mathbb{R}^m$ 有定义并且是线性映射。由于我们在 $\mathbb{R}^n$ 上选定了基 $\{\frac{\partial}{\partial x_i}\}_{i\leqslant n}$，在 $\mathbb{R}^m$ 上选定了基 $\{\frac{\partial}{\partial y_j}\}_{j\leqslant m}$，我们可以把这个线性映射用矩阵 $\left(J_{ij}\right)_{\substack{i\leqslant n,\\ j\leqslant m}}$ 来表示。

首先，用分量表达，我们有

$$\begin{aligned}\frac{|f(x_0+v)-f(x_0)-df(x_0)v|}{|v|}&=\frac{\left|\left(\cdots,f_j(x_0+v)-f_j(x_0),\cdots\right)-\left(\cdots,\sum_{i=1}^n J_{ji}v_i,\cdots\right)\right|}{|v|}\\&=\frac{\sqrt{\sum_{j=1}^n\left|f_j(x_0+v)-f_j(x_0)-\sum_{i=1}^n J_{ji}v_i\right|^2}}{|v|}.\end{aligned}$$

由于当 $v\to 0$ 时，上述左边为 $o(1)$，所以，限制到每个分量，我们就有

$$o(1)\geqslant\frac{\left|f_j(x_0+v)-f_j(x_0)-\sum_{i=1}^n J_{ji}v_i\right|}{|v|}.$$

按定义，这表明 f_j 是可微分的（因为我们用线性映射在 x_0 附近逼近了 f_j）。反过来，假设对每个 $j\leqslant m$，我们都有

$$\frac{\left|f_j(x_0+v)-f_j(x_0)-\sum_{i=1}^n J_{ji}v_i\right|}{|v|}=o(1),$$

那么，

$$\begin{aligned}\frac{|f(x_0+v)-f(x_0)-df(x_0)v|}{|v|}&=\frac{\sqrt{\sum_{j=1}^n\left|f_j(x_0+v)-f_j(x_0)-\sum_{i=1}^n J_{ji}v_i\right|^2}}{|v|}\\&\leqslant\sum_{j=1}^n\frac{\left|f_j(x_0+v)-f_j(x_0)-\sum_{i=1}^n J_{ji}v_i\right|}{|v|}=n\times o(1)=o(1),\end{aligned}$$

所以 $df(x_0)$ 存在。

我们令 $v = t\dfrac{\partial}{\partial x_{i_0}}$，即 $v_{i_0} = t$ 而其它分量 $= 0$。此时，根据微分的定义，上面的式子的左边是 $o(1)$ 项（$t \to 0$）。计算右边，我们得到

$$o(1) = \frac{\sqrt{\sum_{j=1}^{n} \Big| f_j\big(x_0 + \overbrace{(0, \cdots, 0, t, 0 \cdots, 0)}^{\text{只有第 } i_0 \text{ 个位置非 } 0}\big) - f_j(x_0) - J_{ji_0}t \Big|^2}}{t}.$$

对于一个特定的指标 j_0，我们自然有

$$\begin{aligned} &\sqrt{\sum_{j=1}^{n} \Big| f_j\big(x_0 + (0, \cdots, 0, t, 0 \cdots, 0)\big) - f_j(x_0) - J_{ji_0}t \Big|^2} \\ \geqslant &\Big| f_{j_0}\big(x_0 + (0, \cdots, 0, t, 0 \cdots, 0)\big) - f_{j_0}(x_0) - J_{j_0i_0}t \Big|. \end{aligned}$$

所以，

$$o(1) = \Big| f_{j_0}\big(x_0 + (0, \cdots, 0, t, 0 \cdots, 0)\big) - f_{j_0}(x_0) - J_{j_0i_0}t \Big| t.$$

按照定义，这表明 f_{j_0} 的沿着 x_{i_0} 偏导数存在并且等于 $J_{j_0i_0}$，这表明

$$J_{ji} = \frac{\partial f_j}{\partial x_i}(x_0).$$

命题得证。 □

注记. 上述命题表明，映射可求微分等价于其分量可求微分，所以，我们可以通过继续对分量求微分来引入 k-次可导的概念（就是每次求完微分之后这个微分的每个分量都能再求微分）。所以，我们可以定义 $C^k(\Omega, \mathbb{R}^m)$，这是 k 次微分仍然连续的映射的空间。根据上次课程用偏导数判定微分存在性的定理，我们知道只要 f 的连续 k 次偏导数（可能是沿着不同方向的）存在并且连续，那么映射就是 C^k 的。这是一个非常方便有效的判断方式。

我们现在研究符合映射的微分，也就是所谓的链式法则。

命题 183. （链式法则）假设 $\Omega_j \subset \mathbb{R}^{m_j}$ 是开集，其中 $j = 1, 2, 3$，$f : \Omega_1 \to \Omega_2$，$g : \Omega_2 \to \Omega_3$ 是映射。假设 f 在点 $x_1 \in \Omega_1$ 处可微，g 在点 $x_2 = f(x_1) \in \Omega_2$ 处可微，那么复合映射 $g \circ f$ 在 x_0 处可微，并且

$$\big(d(g \circ f)\big)(x_0) = (dg)(f(x_0)) \circ df(x_0).$$

注记. 上述映射的复合可以用下面的交换图来表示：

$$\begin{array}{ccc} \Omega_1 & \xrightarrow{f} & \Omega_2 \\ & \searrow^{g\circ f} & \downarrow g \\ & & \Omega_3 \end{array}$$

那么，它们所对应的微分（在线性的层次上）也可以用类似的交换图来表示：

$$\begin{array}{ccc} \mathbb{R}^{m_1} & \xrightarrow{df(x_1)} & \mathbb{R}^{m_2} \\ & \searrow^{d(g\circ f)(x_1)} & \downarrow{\scriptstyle dg(f(x_1))} \\ & & \mathbb{R}^{m_3} \end{array}$$

我们之前引入的符号更好的描述了这个场景：对于映射 $df(x_1):\mathbb{R}^{m_1}\to\mathbb{R}^{m_2}$，我们将 x_1 所对应的 $\mathbb{R}^{m_1}$ 记作 $T_{x_1}\Omega_1$，将 $\mathbb{R}^{m_2}$ 记作是 $T_{f(x_1)}\Omega_2$，那么，我们有映射

$$df\big|_{x=x_1}:T_{x_1}\Omega_1\to T_{f(x_1)}\Omega_2.$$

从而，上面的交换图表可以写成

$$\begin{array}{ccc} T_{x_1}\Omega_1 & \xrightarrow{df|_{x=x_1}} & T_{f(x_1)}\Omega_2 \\ & \searrow^{d(g\circ f)|_{x=x_1}} & \downarrow{\scriptstyle dg|_{x=f(x_1)}} \\ & & T_{g(f(x_1))}\Omega_3 \end{array}$$

证明：链式法则的推导与 1 维的情形如出一辙：令 $x_2=f(x_1)\in\Omega_2$，按照定义有

$$\begin{aligned} f(x_1+h)&=f(x_1)+df(x_1)h+\delta(h),\\ g\big(f(x_1)+\ell\big)&=g\big(f(x_1)\big)+dg(x_2)\ell+\Delta(\ell), \end{aligned}$$

其中 $h\in\mathbb{R}^{m_1}$，$\ell\in\mathbb{R}^{m_2}$，$\lim\limits_{h\to0}\dfrac{|\delta(h)|}{|h|}=\lim\limits_{\ell\to0}\dfrac{|\Delta(\ell)|}{|\ell|}=0$。据此，我们有

$$\begin{aligned} g(f(x_1+h))-g(f(x_1))&=g\big(f(x_1)+df(x_1)h+\delta(h)\big)-g(f(x_1))\\ &=df(x_2)\big(df(x_1)h+\delta(h)\big)+\Delta(f'(x_0)h+\delta(h))\\ &=\underbrace{df(x_2)\big(df(x_1)h\big)}_{=dg(x_2)\circ df(x_1)(h)}+df(x_2)\big(\delta(h)\big)+\Delta(f'(x_0)h+\delta(h)). \end{aligned}$$

所以，

$$\begin{aligned} &\frac{g(f(x_1+h))-g(f(x_1))-dg(x_2)\circ df(x_1)(h)}{h}\\ =&\frac{df(x_2)\big(\delta(h)\big)}{h}+\frac{\Delta(f'(x_0)h+\delta(h))}{h}\\ \leqslant&C\Big|\frac{\delta(h)}{h}\Big|+\underbrace{\Big|\frac{\Delta(f'(x_0)h+\delta(h))}{|f'(x_0)h+\delta(h)|}\Big|}_{o(1)}+\underbrace{\Big|\frac{|f'(x_0)h+\delta(h)|}{h}\Big|}_{\leqslant C_1}. \end{aligned}$$

由此可见，这是一个 $o(1)$ 项，按照微分的定义，

$$d(g\circ f)(x_1)=dg(x_2)\circ df(x_1).$$

这就完成了证明。 □

作为推论，我们可以计算反函数（逆映射）的微分：

推论 184. 给定区域 $\Omega_1 \subset \mathbb{R}^{n_1}$ 和 $\Omega_2 \subset \mathbb{R}^{n_2}$ 和可微映射 $f : \Omega_1 \to \Omega_2$。假设 f 是双射并且其逆映射 $f^{-1} : \Omega_2 \to \Omega_1$ 是可微的，那么

- $n_1 = n_2$；
- $df(x)$ 是可逆的（等价于 $\mathrm{Jac}(f)(x)$ 的行列式是非零的）。

此时，对于任意的 $y \in \Omega_2$，我们有

$$df^{-1}(y) = \left(df\big|_{x=f^{-1}(y)}\right)^{-1}.$$

证明：我们令 $\Omega_3 = \Omega_1$，$g = f^{-1}$，$x_1 = x$，$x_2 = y$，$g \circ f = \mathrm{Id}$，其中

$$\mathrm{Id} : \Omega_1 \to \Omega_1, \quad x \mapsto x,$$

是单位映射，它的微分在每个点处都是单位映射（线性）。根据链式法则，我们就有

$$\mathrm{Id} = dg(y) \circ df(x).$$

根据矩阵的秩的理论，我们知道 $n_1 \leqslant n_2$。用 f^{-1} 替换 f，我们就得到 $n_2 \leqslant n_1$。这就证明了维数的部分。上面的等式已经蕴含了逆映射的微分的计算。 □

例子 (指数映射的微分)**.** 上个学期我们对于 $n \times n$ 的矩阵定义了指数映射

$$\exp : \mathbf{M}_n(\mathbb{R}) \to \mathbf{M}_n(\mathbb{R}),\ A \mapsto e^A = \sum_{k=0}^{\infty} \frac{A^k}{k!}.$$

我们现在计算它的微分 $d\exp$。固定 $A \in \mathbf{M}_n(\mathbb{R})$，我们要找到

$$d\exp(A) : \mathbf{M}_n(\mathbb{R}) \to \mathbf{M}_n(\mathbb{R}),$$

其中，我们把 $\mathbf{M}_n(\mathbb{R})$ 视作是 $\mathbb{R}^{n^2}$。对于任意较小的 $V \in \mathbf{M}_n(\mathbb{R})$，我们有

$$e^{A+V} - e^V = \sum_{n=0}^{\infty} \frac{1}{n!} \left((A+V)^n - A^n\right)$$

现在强行展开 $(A+V)^n - A^n$（注意矩阵 A 和 V 的乘法未必交换）。通过将 V 的二次项（以及更高次数的项）放到一起，我们得到

$$(A+V)^n - A^n = \sum_{k=0}^{n} A^k V A^{n-1-k} + Q_n(V).$$

二项式展开的一共不超过 2^n 项，所以 $Q_n(V)$ 中至多有 2^n 项。我们上学期证明过（无论你选取什么样的范数），存在常数 c（依赖于范数），使得对任意的 $n \times n$ 的矩阵 A 和 B，我们都有

$$\|A \cdot B\| \leqslant c\|A\|\|B\|.$$

上述 $Q_n(V)$ 的一个通项形如 $AAVVAA\cdots AA$，这是一个由 n 个 A 和 V 排出来的长度为 n 的字符串，其中至少有 2 个 V。我们可以要求 $\|V\| \leqslant \|A\|$，因为最终我们会令 $V \to 0$（除非 $A = 0$，此时 $Q_n(V) = V^n$，下面的结论仍然成立），所以

$$\|AAVVAA\cdots AA\| \leqslant c^n\|A\|\|A\|\|V\|\|V\|\|A\|\|A\|\cdots\|A\|\|A\|.$$

那么，我们得到

$$\|Q_v(V)\| \leqslant 2^n \times \big(c^n\|V\|^2\|A\|^{n-2}\big).$$

从而，我们有

$$\begin{aligned}&\Big\|\exp(A+V) - \exp(A) - \sum_{n=0}^{\infty}\frac{1}{n!}\Big(\sum_{k=0}^{n}A^kVA^{n-1-k}\Big)\Big\|\\ \leqslant&\sum_{n=0}^{\infty}\Big\|\frac{1}{n!}Q_n(V)\Big\| \leqslant \left(\sum_{n=0}^{\infty}\frac{(2c\|A\|)^n}{n!}\right)\|V\|^2 = e^{2c\|A\|}\|V\|^2.\end{aligned}$$

那么，我们注意到右端的项是 $o(\|V\|)$ 并且 $\displaystyle\sum_{n=0}^{\infty}\frac{1}{n!}\Big(\sum_{k=0}^{n}A^kVA^{n-1-k}\Big)$ 是收敛的。所以，

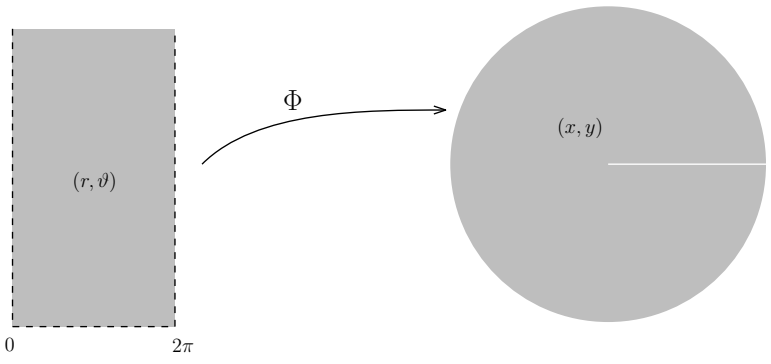
$$d\exp(A)(V) = \sum_{n=0}^{\infty}\frac{1}{n!}\Big(\sum_{k=0}^{n}A^kVA^{n-1-k}\Big)$$

特别地，如果 A 和 V 可交换，那么 $d\exp(A)(V) = \exp(A)V$。我们还有

$$d\exp(0) = \mathrm{Id}.$$

有了链式法则，我们可以讨论更换坐标系的问题。这是个核心的话题，我们在中学的时候就已经在使用这个概念，比如说我们经常在极坐标和 Descartes 坐标系之间转换。我们首先用映射的语言来描述极坐标：令

$$\Omega_1 = \mathbb{R}^2 - \big\{(x,0)\big|x \geqslant 0\big\} \subset \mathbb{R}^2,\ \ \Omega_2 = \mathbb{R}_{>0}\times(0,2\pi) = \big\{(r,\vartheta)\big|r>0,\vartheta\in(0,2\pi)\big\}.$$



我们通常用的 $x = r\cos\vartheta$ 和 $y = r\sin\vartheta$ 可以用如下的映射来写：

$$\Phi : \Omega_2 \to \Omega_2, \quad (r, \vartheta) \mapsto (r\cos\vartheta, r\sin\vartheta).$$

由于在 Ω_1 上我们给定了 (x, y) 作为坐标，在 Ω_2 上我们给定了 (r, ϑ) 作为坐标，所以我们可以用 Jacobi 矩阵来表示上述映射的微分：

$$d\Phi = \mathrm{Jac}(\Phi) = \begin{pmatrix} \frac{\partial x}{\partial r} & \frac{\partial x}{\partial \vartheta} \\ \frac{\partial y}{\partial r} & \frac{\partial y}{\partial \vartheta} \end{pmatrix} = \begin{pmatrix} \cos\vartheta & -r\sin\vartheta \\ \sin\vartheta & r\cos\vartheta \end{pmatrix}.$$

这个线性映射自然是可逆的，它的行列式是 r。

31.1 作业：齐次函数与 Euler 公式

清华大学 19-20 春季学期，数学分析二，作业 1

请用 A4 大小的纸张**正反面**用钢笔，签字笔或者圆珠笔书写，并注明自己的姓名，年级和作业的**总页数**。除定理公式所涉及的人名之外，请使用**中文**。本次作业的提交时间和地点为 **2 月 29 日**上午的课堂上，逾期视作零分。

基本习题

- **热身：多元函数的连续性**

试判断下列极限是否存在；如果存在，计算之。

(1) $\lim\limits_{(x,y)\to(0,0)} \dfrac{\sin(xy)}{x}$　　(2) $\lim\limits_{(x,y)\to(0,0)} \dfrac{\log(x+e^y)}{\sqrt{x^2+y^2}}$

(3) $\lim\limits_{(x,y)\to(0,0)} \mathbf{1}_{\{(x,y)\,|\,0<y<x^2,\ x\in\mathbb{R}\}}$.　　(4) $\lim\limits_{(x,y)\to(0,0)} (x+y)\sin\dfrac{1}{x}\sin\dfrac{1}{y}$

(5) $\lim\limits_{(x,y)\to(0,0)} \dfrac{x^2y^2}{x^2y^2+(x-y)^2}$　　(6) $\lim\limits_{(x,y)\to(0,0)} \dfrac{\sin(xy)}{\sqrt{x^2+y^2}}$

(7) $\lim\limits_{(x,y)\to(0,0)} \dfrac{xy(x^2-y^2)}{x^2+y^2}$　　(8) $\lim\limits_{(x,y)\to(0,0)} \dfrac{\sqrt{xy}}{(x^2+y^2)^p}$, $p>0$

微分的定义与计算

习题 A 我们总假设 $\Omega\subset\mathbb{R}^n$ 是开集，$x_0\in\Omega$ 是一个给定的点。如果不另加说明 $f:\Omega\to\mathbb{R}$ 将代表一个函数。

A1)（微分的唯一性）考虑映射 $f:\Omega\subset\mathbb{R}^n\to\mathbb{R}^m$。假设存在两个线性映射 $A_i:\mathbb{R}^n\to\mathbb{R}^m$（$i=1,2$），使得对 $i=1$ 和 2 和 $v\to 0$ 时，都有

$$f(x_0+v)=f(x_0)+A_i(v)+o(v).$$

证明，$A_1=A_2$。

A2) 假设 f 在 $x_0\in\Omega$ 处可微，$x_1,\cdots,x_n$ 是 n 个坐标函数。证明，$df(x_0)=\sum\limits_{i=1}^n \dfrac{\partial f}{\partial x_i}(x_0)dx_i(x_0)$。

我们通常将这个式子记作 $df=\sum\limits_{i=1}^n \dfrac{\partial f}{\partial x_i}dx_i$。

A3) 沿用上一问题的符号。证明，$\{dx_i(x_0)\}_{1\leqslant i\leqslant n}$ 给出了 $\mathrm{Hom}(T_{x_0}\Omega,\mathbb{R})$（这是 $T_{x_0}\Omega$ 的对偶空间，即它上面的线性函数所构成的线性空间）的一组基，其中 $T_{x_0}\Omega=\mathbb{R}^n$。

A4) 假设 Ω 是凸集（定义请参考课堂笔记），f 是 Ω 上的可微函数。如果对任意的 $x\in\Omega$，都有 $df(x)=0$。证明，f 是常值函数。

A5) 考察函数 $f(x,y)=\begin{cases}(x^2+y^2)\sin\frac{1}{\sqrt{x^2+y^2}}, & \text{若}(x,y)\neq(0,0),\\ 0, & \text{若}(x,y)=(0,0).\end{cases}$

试计算 f 的偏导数并证明 f 的两个偏导数在 $(0,0)$ 处均不连续但是 f 在 $(0,0)$ 处可微。

A6) 考察函数 $f(x,y)=\begin{cases}\frac{xy^2}{x^2+y^4}, & \text{若}(x,y)\neq(0,0),\\ 0, & \text{若}(x,y)=(0,0).\end{cases}$

试计算 f 在 $(0,0)$ 处所有方向导数并证明 f 在 $(0,0)$ 不可微。

A7) 假设函数 f 在 Ω 上的所有偏导数 $\dfrac{\partial f}{\partial x_i}$（$i=1,2,\cdots,n$）都存在并且有界（即存在 M，使得对任意的 $i\leqslant n$，任意的 $x\in\Omega$，我们都有 $\left|\dfrac{\partial f}{\partial x_i}\right|\leqslant M$）。证明，$f$ 在 Ω 上连续。

f 是否在 Ω 上可微？如果是，请证明；否则给出反例。

A8) $\Omega\subset\mathbb{R}^2$ 为开集，我们用 (x,y) 来表示 $\mathbb{R}^2$ 上的坐标。函数 $f:\Omega\to\mathbb{R}$ 的偏导数 $\dfrac{\partial f}{\partial x}$ 和 $\dfrac{\partial f}{\partial y}$ 处处存在。如果偏导数 $\dfrac{\partial f}{\partial y}$ 在 Ω 上连续。证明，f 在 Ω 上可微。

A9) 考虑在 $\mathbf{M}_n(\mathbb{R})$ 上定义的行列式函数

$$\det:\mathbf{M}_n(\mathbb{R})\to\mathbb{R},\quad A\mapsto\det(A).$$

任意给定 $A\in\mathbf{M}_n(\mathbb{R})$，试计算 $d\det\big|_{X=A}$。

（如果你可以证明 det 是可微的，那么可以用上学期第 6 次作业的 A9)）

A10)（偏导数的计算）计算下列函数的偏导数：

(1) $f(x,y)=e^{xy}$　　(2) $f(x,y,z)=\log(x+y^2+z^3)$

(3) $f(x,y,z)=x^{y^z}$　　(4) $f(x,y)=x+y+\sqrt{x^2+y^2}$

(5) $f(x,y)=e^{x+y^2}+\sin(x^2y)$　　(6) $f(x_1,x_2,x_3)=\tan\dfrac{x_1x_2}{x_3^2}$

导数和微分的计算

习题 B (复合函数求偏导数)（请熟练掌握）

假设 f 为可微函数，用 f 的偏导数表达下列多元函数的偏导数：

(1) $u(x,y)=f(x+y,xy)$　　(2) $u(x,y,z)=f(x,xy,xyz)$　　(3) $u(x,y,z)=f(\dfrac{x}{y},\dfrac{y}{z})$

(4) $u(x,y)=f(x^2+y^2)$　　(5) $u(x,y)=f(\log x+\dfrac{1}{y})$　　(6) $u(x,y)=f(xy+e^{f(y)})$

习题 C

C1) 求如下坐标变换 f 的 Jacobi 矩阵 $J(f)$ 并计算 $\det J(f)$:

$$(1) f:\mathbb{R}_{>0}\times(0,2\pi)\to\mathbb{R}^2,\quad (r,\theta)\mapsto(r\cos\theta,r\sin\theta);$$
$$(2) f:\mathbb{R}_{>0}\times(0,2\pi)\times\mathbb{R}\to\mathbb{R}^2,\quad (r,\theta,z)\mapsto(r\cos\theta,r\sin\theta,z);$$
$$(3) f:\mathbb{R}_{>0}\times(0,2\pi)\times(0,\pi)\to\mathbb{R}^3,(r,\theta,\varphi)\quad\mapsto(r\sin\theta\cos\varphi,r\sin\theta\sin\varphi,r\cos\theta);$$
$$(4) f:\mathbb{R}^2\to\mathbb{R}^2,\quad (u,v)\mapsto(u^2-v^2,2uv);$$
$$(5) f:\mathbb{R}^2\to\mathbb{R}^2,\quad (u,v)\mapsto(e^u\cos v,e^u\sin v);$$
$$(6) f:\mathbb{R}^2\to\mathbb{R}^2,\quad (u,v)\mapsto(\frac{u}{u^2+v^2},\frac{v}{u^2+v^2}).$$

C2) 我们考虑 $\mathbb{R}^3$ 上的柱面坐标系: $x=r\cos\theta, y=\sin\theta, z=z$。这个坐标变换用映射来写就是

$$\Phi:\mathbb{R}_{>0}\times(0,2\pi)\times\mathbb{R}\to\mathbb{R}^3,\quad (r,\theta,z)\mapsto(r\cos\theta,r\sin\theta,z).$$

设 f 是 $\mathbb{R}^3$ 上的二次可微函数。当 $(x,y)\neq(0,0)$ 时，试通过计算来证明:

$$\frac{\partial f}{\partial r}=\cos\theta\frac{\partial f}{\partial x}+\sin\theta\frac{\partial f}{\partial y},\quad \frac{\partial f}{\partial\theta}=-r\sin\theta\frac{\partial f}{\partial x}+r\cos\theta\frac{\partial f}{\partial y},$$
$$\left(\frac{\partial f}{\partial x}\right)^2+\left(\frac{\partial f}{\partial y}\right)^2+\left(\frac{\partial f}{\partial z}\right)^2=\left(\frac{\partial f}{\partial r}\right)^2+\frac{1}{r^2}\left(\frac{\partial f}{\partial\theta}\right)^2+\left(\frac{\partial f}{\partial z}\right)^2,$$
$$\frac{\partial^2 f}{\partial x^2}+\frac{\partial^2 f}{\partial y^2}+\frac{\partial^2 f}{\partial z^2}=\frac{1}{r}\frac{\partial}{\partial r}\left(r\frac{\partial f}{\partial r}\right)+\frac{1}{r^2}\frac{\partial^2 f}{\partial\theta^2}+\frac{\partial^2 f}{\partial z^2}.$$

C3) 我们考虑 $\mathbb{R}^3$ 上的球坐标系 $(x,y,z)=(r\sin\theta\cos\varphi,r\sin\theta\sin\varphi,r\cos\theta)$。证明，

$$\left(\frac{\partial f}{\partial x}\right)^2+\left(\frac{\partial f}{\partial y}\right)^2+\left(\frac{\partial f}{\partial z}\right)^2=\left(\frac{\partial f}{\partial r}\right)^2+\frac{1}{r^2}\left(\frac{\partial f}{\partial\theta}\right)^2+\frac{1}{r^2\sin^2\theta}\left(\frac{\partial f}{\partial\varphi}\right)^2,$$
$$\frac{\partial^2 f}{\partial x^2}+\frac{\partial^2 f}{\partial y^2}+\frac{\partial^2 f}{\partial z^2}=\frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial f}{\partial\theta}\right)+\frac{1}{r^2\sin\theta}\left(\sin\theta\frac{\partial f}{\partial r}\right)+\frac{1}{r^2\sin^2\theta}\frac{\partial^2 f}{\partial\varphi^2}.$$

习题 D（齐次函数与 Euler 公式）考虑函数 $f:\mathbb{R}^n-\{0\}\to\mathbb{R}$。如果存在 $k\in\mathbb{R}$，使得对任意的 $\lambda>0$ 和任意的 $x\neq0$，我们都有 $f(\lambda x)=\lambda^k f(x)$，我们就称 f 为 k **次齐次的**，其中 k 称作是它的次数。

D1) 证明，下面的函数 $f:\mathbb{R}^n\to\mathbb{R}$ 是齐次函数:

(1) f 是线性函数;

(2) $f(x_1,\cdots,x_n)=x_1x_2+\cdots+x_{n-1}x_n$;

(3) $f(x_1,\cdots,x_n)=\min\{x_1,\cdots,x_n\}$;

(4) $f(x)=\dfrac{g(x)}{h(x)}$, 其中当 $x\neq0$ 时, $h(x)\neq0$, 并且 g 和 h 都是齐次函数.

D2) (Euler)假设 f 是可微的。证明: f 是 k 次齐次函数当且仅当它满足 Euler 等式 $\displaystyle\sum_{i=1}^n x_i\frac{\partial f}{\partial x_i}=kf$。

D3) 假设可微函数 f 是 k 次齐次函数。证明，对任意 $v\in\mathbb{R}^2$，$\nabla_v f$ 是 $k-1$ 次齐次函数。

D4) 令 $f(x_1,\cdots,x_n)=\det\begin{pmatrix}1 & \cdots & 1\\ x_1 & \cdots & x_n\\ \vdots & & \vdots\\ x_1^{n-1} & \cdots & x_n^{n-1}\end{pmatrix}$。证明，$\sum_{i=1}^n x_i\frac{\partial f}{\partial x_i}=\frac{n(n-1)}{2}f$ 和 $\sum_{i=1}^n\frac{\partial f}{\partial x_i}=0$。

思考题：切空间的抽象／几何定义（这是理解微分和切空间的重要习题）

习题 T 给定开集 $\Omega\subset\mathbb{R}^n$ 和 $p\in\Omega$。我们令 $\mathcal{C}(p)$ 是所有通过 p 的 C^1 的曲线的集合。按照定义，我们有

$$\mathcal{C}(p)=\Big\{\gamma:I\to\Omega\Big|I\subset\mathbb{R}\text{ 是开区间且 }0\in I,\ \gamma(0)=p,\ \gamma\text{ 是 }C^1\text{ 的}\Big\},$$

我们在 $\mathcal{C}(p)$ 定义如下的等价关系，其中 $\gamma_1,\gamma_2\in\mathcal{C}(p)$：

$$\gamma_1\sim\gamma_2\ \text{ 当且仅当 }\ \gamma_1'(0)=\gamma_2'(0).$$

我们考虑 $\mathcal{C}(p)$ 在上述等价关系下的等价类 $\mathcal{C}(p)/\sim$：

$$\mathcal{C}(p)/\sim=\big\{[\gamma]\big|\gamma\in\mathcal{C}(p)\big\},$$

其中，如果 $\gamma\sim\gamma'$，那么在等价类的集合 $\mathcal{C}(p)/\sim$ 中，$[\gamma]=[\gamma']$。换句话说，我们把在 p 点处的切向量相同的曲线认为是同一条曲线，这就是为什么我们把这个空间称作是切空间。

T1) 证明，映射

$$\iota_p:\mathcal{C}(p)/\sim\to\mathbb{R}^n,\ \ [\gamma]\mapsto\gamma'(0)$$

是良好定义的并且是双射

T2) 证明，将 $\mathcal{C}(p)/\sim$ 定义为一个 $\mathbb{R}$-线性空间使得上述为线性空间的同构，我们把具有线性空间结构的 $\mathcal{C}(p)/\sim$（它本来仅是曲线的等价类的集合）记作是 $T_p\Omega_1$。

T3) 令 $\Omega'=\mathbb{R}^m$，$f:\Omega\to\Omega'$ 是 C^1 的映射（即其微分也连续），$p'=f(p)$。证明，映射

$$f_{\sharp,p}:\mathcal{C}(p)\to\mathcal{C}(p'),\ \ \gamma\mapsto f\circ\gamma,$$

是良好定义的。换句话说，通过和 f 复合，我们可以把 Ω 上通过 p 点的曲线映射称为 Ω' 上通过 p' 点的曲线。

T4) 证明，映射

$$f_{*p}:T_p\Omega\to T_{p'}\Omega',\ \ [\gamma]\mapsto[f_{\sharp,p}(\gamma)],$$

是良好定义线性映射。这个映射被称作是**切映射**。

T5) 在 Ω 上考虑曲线

$$\ell_k : (-1,1) \to \Omega', \quad t \mapsto p + \underbrace{(0,0,\cdots,0,t,0,\cdots,0)}_{\text{只有第 } k \text{ 个位置非 } 0}$$

证明，$\iota_p([\ell_k]) = \dfrac{\partial}{\partial x_k}$，其中，我们在 $\mathbb{R}^n$ 上用 $\{x_1,\cdots,x_n\}$ 作为坐标系。

T6) 在 Ω 上考虑曲线

$$\ell'_{k'} : (-1,1) \to \Omega', \quad t \mapsto p + \underbrace{(0,0,\cdots,0,t,0,\cdots,0)}_{\text{只有第 } k' \text{ 个位置非 } 0}$$

试用 f 的偏导数和 $[\ell'_{k'}]$ 来表达 $f_{*p}([\ell_k])$。

T7) 证明，我们有如下的交换图表：

$$\begin{array}{ccc} T_p\Omega & \xrightarrow[\approx]{\iota_p} & T_{f(x_1)}\Omega_2 \\ \Big\downarrow{\scriptstyle f_{*p}} & & \Big\downarrow{\scriptstyle df(p)} \\ T_{x_1}\Omega_1 & \xrightarrow[\approx]{\iota_{p'}} & T_{g(f(x_1))}\Omega_3 \end{array}$$

即对任意的 $[\gamma] \in T_p\Omega$，我们有

$$f_{*p}([\gamma]) = \iota_{p'}^{-1}\Big(df(p)\big(\iota_p([\gamma])\big)\Big).$$

在这个意义，f_{*p} 与 $df(p)$ 是一样的。

T8) 给定 $[\gamma_1],[\gamma_2] \in T_p\Omega$，如果它们都不是 0，我们可以计算它们的角度：

$$\cos\big(\angle([\gamma_1],[\gamma_2])\big) = \frac{\gamma_1'(0)\cdot\gamma_2'(0)}{|\gamma_1'(0)||\gamma_2'(0)|},$$

我们要求角度是在 $[0,\pi)$ 中的值。考虑反演变换：

$$f : \mathbb{R}^n - \{0\} \to \mathbb{R}^n - \{0\}, \quad x \mapsto \frac{x}{|x|}.$$

证明，f 是 C^1 的并且对任意的 $p \in \mathbb{R}^n - \{0\}$，$f_* : T_p\big(\mathbb{R}^n - \{0\}\big) \to T_{f(p)}\big(\mathbb{R}^n - \{0\}\big)$ 保持角度。

Pure mathematics is, in its way, the poetry of logical ideas.

– Albert Einstein

---

32 微分同胚，坐标变换，Clairaut-Schwarz 定理（偏导数可交换），多元函数的 Taylor 展开，$\mathbb{R}^n$ 中光滑子流形的定义

二零二零年二月二十四日，星期一，阴

微分同胚与坐标变换

给定映射 $f:\Omega_1\to\mathbb{R}^{n_2}$，其中开区域 $\Omega_1\subset\mathbb{R}^{n_1}$，用坐标分量表示，我们有

$$f=(f_1,f_2,\cdots,f_{n_2}).$$

如果对任意的 $k\in\{1,\cdots,n_2\}$ 对任意的 $N\geqslant 1$,对任意的 N 个正整数 $k_1,k_2,\cdots,k_N\in\{1,2,\cdots,n_1\}$，偏导数

$$\frac{\partial}{\partial x_{k_N}}\left(\frac{\partial}{\partial x_{k_{N-1}}}\left(\cdots\left(\frac{\partial f_k}{\partial x_{k_1}}\right)\cdots\right)\right)$$

都存在并且是连续的，那么我们就称 f 是**光滑的**。如果 $n_2=1$，我就称之为**光滑函数**，并记为 $C^\infty(\Omega_1)$。

练习. 证明以下三个关于光滑函数的基本性质：

1) $C^\infty(\Omega_1)$ 是一个 $\mathbb{R}$-代数，即对任意的 $f,g\in C^\infty(\Omega_1)$，它们的任意实线性组合以及乘积都是光滑函数。

2) 假设 $f\in C^\infty(\Omega_1)$。如果对任意的 $x\in\Omega$，$f(x)\neq 0$，那么 $\dfrac{1}{f}$ 也是光滑函数。

3) 证明，对任意的 $i\in\{1,2,\cdots,n_1\}$，对任意的 $f\in C^\infty(\Omega_1)$，有 $\dfrac{\partial f}{\partial x_i}\in C^\infty(\Omega_1)$。

对于两个开区域 $\Omega_1\subset\mathbb{R}^{n_1}$ 和 $\Omega_2\subset\mathbb{R}^{n_2}$，如果存在双射映射 $f:\Omega_1\to\Omega_2$，使得

$$f:\Omega_1\to\Omega_2,\quad f^{-1}:\Omega_2\to\Omega_1$$

都是光滑的，我们就称 Ω_1 和 Ω_2 是**微分同胚的**（光滑同胚）。

我们把 Ω_1 到 Ω_2 之间的光滑映射的全体记作是 $C^\infty(\Omega_1,\Omega_2)$。当然，光滑的映射不见得是光滑的同胚。

例子. 考虑如下映射

$$f:\mathbb{R}\to\mathbb{R},\quad x\mapsto x^3.$$

很明显，f 是光滑的并且是双射。但是 $f^{-1}(y)=\sqrt[3]{y}$ 不是光滑的。

根据上次课逆映射微分的命题，我们一定有 $n_1=n_2$（维数相同）。

注记.

1) 微分同胚是两个区域的一种等价性，这个等价性是用光滑的双射定义的。一个好的类比是研究线性空间：两个线性空间之间等价指的是用线性的双射定义的线性同构。

2) 两个区域微分同胚，但是上述的映射 f 不一定唯一。比如说，令 $\Omega_1 = \Omega_2 = \mathbb{R}^n$，任意一个可逆的线性映射都可以被选作 f。

有了微分同胚这个概念，我们可以讨论坐标变换。考虑两个开区域 $\Omega_1 \subset \mathbb{R}^{n_1}$ 和 $\Omega_2 \subset \mathbb{R}^{n_2}$，其中 $n_1 = n_2 = n$（为了区别起见）。我们在第一个 $\mathbb{R}^{n_1}$ 上面用 $(x_1, \cdots, x_n)$ 作为坐标系，在第二个 $\mathbb{R}^{n_2}$ 上面用 $(y_1, \cdots, y_n)$ 作为坐标系，假设存在光滑的同胚：

$$\Phi : \Omega_1 \to \Omega_2.$$

我们认为通过 Φ 这个映射可以用 Ω_1 上的点来**参数化** Ω_2 上的点：用 $(x_1, \cdots, x_n)$ 这个坐标系统也可以描述 Ω_2 上的点。比如说，给定坐标 $(x_1, \cdots, x_n)$，它对应的 Ω_2 中的点是 $\Phi(x_1, \cdots, x_2)$；根据双射的性质，对任意的 $p \in \Omega_2$，我们总能找到 $(x_1, \cdots, x_n) \in \Omega_1$，使得 $\Phi(x_1, \cdots, x_n) = p$。所以，我们有两种不同的方式描述 Ω_2 上的同一个点 $p \in \Omega_2$：

- 第一种方式是 p 的 y_i-坐标 $(y_1(p), \cdots, y_n(p))$；
- 第二种是 $f^{-1}(p)$ 的 x_i-坐标 $(x_1\big(\Phi^{-1}(p)\big), \cdots, x_n\big(\Phi^{-1}(p)\big))$。

我们强调之前所谈论的坐标函数的含义：x_i 和 y_i 分别视作是 Ω_1 和 Ω_2 上的函数。

所以，给定了 $(x_1, \cdots, x_n)$ 来描述 $p \in \Omega_2$，它所对应的 $(y_1, \cdots, y_n)$ 的坐标应该是

$$(\Phi_1(x_1, \cdots, x_n), \cdots, \Phi_n(x_1, \cdots, x_n)).$$

很多文献习惯上将这个写成 $(y_1(x_1, \cdots, x_n), \cdots, y_n(x_1, \cdots, x_n))$。

同学们也许会发现，上面这段讨论实际上根本没有用到 Φ 和 Φ^{-1} 是光滑的，只需要 Φ 是双射就好：实际上，光滑性保证了光滑函数的拉回还是光滑的。当给定了 Ω_2 上的一个函数 $f : \Omega_2 \to \mathbb{R}$，通过复合映射，我们可以将它看作是 Ω_1 的函数，我们通常将它记作 $\Phi^* f = f \circ \Phi$（称作是 f 被 Φ 的**拉回**），下面的交换图表给出了拉回的定义：

$$\begin{array}{ccc} \Omega_1 & \xrightarrow{\Phi} & \Omega_2 \\ & \searrow^{\Phi^* f} & \downarrow f \\ & & \mathbb{R} \end{array}$$

引理 185. *假设 $f \in C^\infty(\Omega_2)$（即 f 是 Ω_2 上的光滑函数），那么 $\Phi^* f \in C^\infty(\Omega_1)$。*

证明： 我们只需证明对任意的 Ω_2 上的光滑函数，对任意的 N，$\Phi^* f$ 的各阶偏导数

$$\frac{\partial}{\partial x_{k_N}} \left(\frac{\partial}{\partial x_{k_{N-1}}} \left(\cdots \left(\frac{\partial (f \circ \Phi)}{\partial x_{k_1}} \right) \cdots \right) \right)$$

都存在且连续即可。对所求的偏导数的次数 N 进行归纳。如果 $N=0$，由于 Φ 是连续映射，根据连续映射的符合仍然连续，我们知道此时 Φ^*f 是连续的。假设对任意的 f，Φ^*f 的连续 N 次的偏导数存在且连续，那么根据链式法则，我们知道

$$\frac{\partial(f\circ\Phi)}{\partial x_i}\Big|_{x=x_0}=\sum_{k=1}^{n}\frac{\partial\Phi^k}{\partial x^i}(x_0)\frac{\partial f}{\partial y_k}\big(\Phi(x_0)\big)=\sum_{k=1}^{n}\frac{\partial\Phi^k}{\partial x^i}\cdot\left(\Phi^*\frac{\partial f}{\partial y_k}\right)\Big|_{x=x_0}.$$

此时，每一个 $\dfrac{\partial\Phi^k}{\partial x^i}$ 都是光滑函数，对 $\dfrac{\partial f}{\partial y_k}$ 利用归纳假设（这是那个被拉回的函数），对上述函数再求 N 次偏导数也连续，所以命题成立。 □

注记. 上面的证明仅仅用到了 Φ 的光滑性，换句话说，我们证明了如下的命题，如果 f 是光滑函数，Φ 是光滑映射，那么它们的复合 $f\circ\Phi$ 是光滑函数，其中，我们假设 $f:\Omega_2\to\mathbb{R}$，$\Phi:\Omega_1\to\Omega_2$。

最终，我们考虑常见的但是容易产生混淆的一种情形：我们假设有两个坐标系统 (x_i) 和 (y_i) 来描述 Ω 中的点。此时，我们把每一个坐标函数都理解为 Ω 上的函数，那么，如果用第一个坐标系统来描述第二个坐标系统的坐标函数，我们就可以写成

$$y_1=y_1(x_1,\cdots,x_n),\cdots,y_n=y_n(x_1,\cdots,x_n).$$

反过来，我们可以用第二个坐标系统来描述第一个坐标系统的坐标函数：

$$x_1=x_1(y_1,\cdots,y_n),\cdots,x_n=x_n(y_1,\cdots,y_n).$$

假设 $f:\Omega\to\mathbb{R}$ 是 Ω 上的函数。通过利用不同的坐标，我们可以将 f 写成 $f(x_1,\cdots,x_n)$ 或者 $f(y_1,\cdots,y_n)$。如此用变量来写函数很容易产生混乱。当然，如果 f 是用第二个坐标系统写的，即 $f(y_1,\cdots,y_n)$，我们所说的 $f(x_1,\cdots,x_n)$ 应该是

$$f(x_1,\cdots,x_n)=f(y_1(x_1,\cdots,x_n),\cdots,y_n(x_1,\cdots,x_n)).$$

事实上，我们应该将 $\Omega_1=\Omega$ 和 $\Omega_2=\Omega$ 区分开，在 $\Omega_1\subset\mathbb{R}^{n_1}$ 中我们用 (x_i) 作为坐标，在 $\Omega_2\subset\mathbb{R}^{n_2}$ 中我们用 (y_i) 作为坐标，其中 $n_1=n_2=n$。我们考虑映射

$$\Phi:\Omega_1\to\Omega_2,\ \ (x_1,\cdots,x_n)\mapsto\big(y_1(x_1,\cdots,x_n),\cdots,y_n(x_1,\cdots,x_n)\big).$$

它的逆映射就是

$$\Phi^{-1}:\Omega_2\to\Omega_1,\ \ (y_1,\cdots,y_n)\mapsto\big(x_1(y_1,\cdots,y_n),\cdots,x_n(y_1,\cdots,y_n)\big).$$

此时，上面谈到的函数 f 是 Ω_2 上的函数。

另外 f 的偏导数（或者微分）也可以用 (x_i) 来表达，这就是说要计算 $f\circ\Phi$ 的偏导数（用 (y_i) 坐标来算，就是 $\dfrac{\partial f}{\partial y_i}$）。

练习 (重要)**.** 利用链式法则证明:

$$\frac{\partial f}{\partial x_i}=\sum_{j=1}^{n}\frac{\partial y_j}{\partial x_i}(x_1,\cdots,x_n)\frac{\partial f}{\partial y_j}(y_1(x_1,\cdots,x_n),\cdots,y_n(x_1,\cdots,x_n)).$$

据此可知，由偏导数的定义是依赖于坐标系的选取的。特别地，在不同的坐标系之间它们的差别由坐标变换（即上述的 Φ）的 Jacobi 矩阵决定。

注记. 偏导数的定义不是内蕴的，它依赖于具体坐标系的选取。然而，微分 df（按定义）不依赖于坐标系统的选取，是内蕴的。

我们现在回到学习的主线上来。我们现在证明偏导数运算具有可交换性（这个定理的几何表述是函数的 Hasse 算子是对称的）:

定理 186 (Clairaut-Schwarz)**.** 给定 $\mathbb{R}^n$ 上的开集 Ω（$n\geqslant 2$）和函数 $f:\Omega\to\mathbb{R}$ 是函数，$i,j\in\{1,2,\cdots,n\}$ 是两个不同的指标。假设在 Ω 上，函数 $\dfrac{\partial f}{\partial x_i}(x)$，$\dfrac{\partial f}{\partial x_j}(x)$ 和 $\dfrac{\partial}{\partial x_i}\Big(\dfrac{\partial f}{\partial x_j}\Big)(x)$ 存在并且**连续**。那么，$\dfrac{\partial}{\partial x_j}\Big(\dfrac{\partial f}{\partial x_i}\Big)(x)$ 也存在并且对任意的 $x\in\Omega$，我们有

$$\frac{\partial}{\partial x_i}\Big(\frac{\partial f}{\partial x_j}\Big)(x)=\frac{\partial}{\partial x_j}\Big(\frac{\partial f}{\partial x_i}\Big)(x).$$

为了证明这个命题，我们从一个引理开始:

引理 187. 假设函数 $f:\mathbb{R}^2-\{(0,0)\}\to\mathbb{R}$ 在 $(0,0)\in\mathbb{R}^2$ 处的极限 $\lim\limits_{(x,y)\to(0,0)}f(x,y)$ 存在。如果存在 $\varepsilon_0>0$，使得对于任意给定的 $y_0\in(-\varepsilon_0,\varepsilon_0)$，极限 $\lim\limits_{x\to 0}f(x,y_0)$ 都存在。那么，极限 $\lim\limits_{y\to 0}\big(\lim\limits_{x\to 0}f(x,y)\big)$ 存在并且

$$\lim_{y\to 0}\big(\lim_{x\to 0}f(x,y)\big)=\lim_{(x,y)\to(0,0)}f(x,y).$$

引理的证明. 按照距离空间中函数极限存在的定义，$\lim\limits_{(x,y)\to(0,0)}f(x,y)$ 存在指的是存在 L，使得对任意的 $\varepsilon>0$，存在 $\delta>0$，当 $\sqrt{x^2+y^2}<\delta$ 时，有 $|f(x,y)-L|<\varepsilon$。

任意固定 $\varepsilon>0$。在上述极限的定义中，我们现在选取 δ 使得 $\delta\leqslant\varepsilon_0$。考虑任意一个 y，其中 $|y|<\delta$。下面认为 y 是固定的。

那么，自然有 x，使得 $\sqrt{y^2+x^2}<\delta$。此时，我们令 $x\to 0$，我们就有（因为 $|y|<\varepsilon_0$）:

$$|\lim_{x\to 0}f(x,y)-L|\leqslant\varepsilon.$$

即对任意的 $\varepsilon>0$，我们找到了 $\delta>0$，使得对任意的 $|y|<\delta$，上面的不等式成立。按照极限的定义，这就是说 $\lim\limits_{y\to 0}\big(\lim\limits_{x\to 0}f(x,y)\big)$ 存在并且等于 L。 □

Clairaut-Schwarz 定理的证明. 不妨假设 $i=1$，$j=2$。为了简单期间，我们将函数 f 记为

$$f(x_1,x_2,\cdots,x_n)=f(x,y,Z),$$

其中 $(x,y)=(x_1,x_2)$，$Z=(x_3,\cdots,x_n)\in\mathbb{R}^{n-2}$。对任意的 $h,\ell\in\mathbb{R}$，定义

$$\Delta(h,\ell)=f(x+h,y+\ell,Z)-f(x+h,y,Z)-f(x,y+\ell,Z)+f(x,y,Z).$$

根据 Lagrange 中值定理，我们有

$$\begin{aligned}\frac{\Delta(h,\ell)}{h\ell}&=\frac{1}{h}\left(\frac{f(x+h,y+\ell,Z)-f(x+h,y,Z)}{\ell}-\frac{f(x,y+\ell,Z)-f(x,y,Z)}{\ell}\right)\\&=\frac{1}{h}\left(\frac{\partial f}{\partial y}(x+h,y+\theta_1\ell,Z)-\frac{\partial f}{\partial y}(x,y+\theta_2\ell,Z)\right)\end{aligned}$$

其中 $\theta_1,\theta_2\in[0,1]$。当然，简单地运用中值定理不能保证 $\theta_1=\theta_2$。为说明基本的想法，我们先假设 $\theta_1=\theta_2=\theta$。此时，根据二阶导数存在，我们可以继续运用中值定理：

$$\frac{\Delta(h,\ell)}{h\ell}=\frac{\partial}{\partial x}\left(\frac{\partial f}{\partial y}\right)(x+\theta' h,y+\theta\ell,Z),$$

其中，$\theta'\in[0,1]$。再根据 $\dfrac{\partial}{\partial x}(\dfrac{\partial f}{\partial y})$ 的连续性，我们知道 $\lim\limits_{(k,\ell)\to 0}\dfrac{\Delta(h,\ell)}{h\ell}$ 存在。

我们要对函数

$$\mathbb{R}^2-(0,0)\to\mathbb{R},\quad (h,\ell)\to\frac{\Delta(h,\ell)}{h\ell},$$

来运用前面的引理。我们可以先令 $\ell\to 0$，根据定义，我们有

$$\begin{aligned}\lim_{h\to 0}\left(\lim_{\ell\to 0}\frac{\Delta(h,\ell)}{h\ell}\right)&=\lim_{h\to 0}\left(\frac{1}{h}\left(\frac{f(x+h,y+\ell,Z)-f(x+h,y,Z)}{\ell}-\frac{f(x,y+\ell,Z)-f(x,y)}{\ell}\right)\right)\\&=\lim_{h\to 0}\frac{1}{h}\left(\frac{\partial f}{\partial y}(x+h,y,Z)-\frac{\partial f}{\partial y}(x,y,Z)\right)\\&=\frac{\partial}{\partial x}\left(\frac{\partial f}{\partial y}\right)(x,y,Z).\end{aligned}$$

反过来，我们也可以先令 $h\to 0$。类似的，我们得到

$$\lim_{\ell\to 0}\left(\lim_{h\to 0}\frac{\Delta(h,\ell)}{h\ell}\right)=\frac{\partial}{\partial y}\left(\frac{\partial f}{\partial x}\right)(x,y,Z).$$

根据引理，上面两式子左端的极限是相同的，这就证明了命题。

最终，我们解决上述 $\theta_1=\theta_2$ 的技术性问题。我们把函数写成：

$$\frac{\Delta(h,\ell)}{h\ell}=\frac{1}{h}\left(\frac{\big[f(x+h,y+\ell,Z)-f(x,y+\ell,Z)\big]-\big[f(x+h,y,Z)-f(x,y,Z)\big]}{\ell}\right).$$

我们只需要将 $f(x+h,y,Z)-f(x,y,Z)$ 作为整体视作是 y 的函数（此时，x,h,Z 都是固定的）来运用 Lagrange 中值定理即可。 □

注记. 基于这个命题，我们可以引入一个方便的记号来记多重的偏导数：令 $\alpha=(\alpha_1,\cdots,\alpha_n)$ 为一个多重指标，也就是说对于每个 $i\leqslant n$，有 $\alpha_i\in\mathbb{Z}_{\geqslant 0}$，我们用 $\partial^\alpha f$ 表示 α_1 个 $\dfrac{\partial}{\partial x_1}$，$\alpha_2$ 个 $\dfrac{\partial}{\partial x_2}$，$\cdots$，$\alpha_n$ 个 $\dfrac{\partial}{\partial x_n}$ 作用在 f 上，即

$$\partial^\alpha f=\left(\frac{\partial}{\partial x_n}\right)^{\alpha_n}\circ\left(\frac{\partial}{\partial x_{n-1}}\right)^{\alpha_{n-1}}\circ\cdots\circ\left(\frac{\partial}{\partial x_1}\right)^{\alpha_1}(f).$$

当 f 足够多次连续可微时，上述命题保证了这些算子的复合不依赖于它们作用的顺序。我们还令

$$|\alpha|=\alpha_1+\cdots+\alpha_n\leqslant k,$$

传统上我们还吧上面的多重偏导数 ∂^α 记作

$$\frac{\partial^{|\alpha|}}{\partial x_1^{\alpha_1}\cdots x_n^{\alpha_n}}.$$

比如说，我们经常看到

$$\frac{\partial^2 f}{\partial x^2},\quad\frac{\partial^2 f}{\partial x\partial y}.$$

例子 (Clairaut-Schwarz 的反例)**.** 如果 *Clairaut-Schwarz* 定理中的连续性不成立，那么命题可能并不成立。考察函数

$$f(x,y)=\begin{cases}\frac{xy(x^2-y^2)}{x^2+y^2}, & (x,y)\neq(0,0);\\ 0, & (x,y)=(0,0).\end{cases}$$

那么，我们有

$$\frac{\partial f}{\partial x}(x,y)=\frac{x^4y+4x^2y^3-y^5}{(x^2+y^2)^2}.$$

从而，我们有

$$\frac{\partial}{\partial y}\big(\frac{\partial f}{\partial x}\big)(0,0)=\lim_{y\to 0}\frac{\frac{\partial f}{\partial x}(0,y)-\frac{\partial f}{\partial x}(0,0)}{y}=-1.$$

类似地（利用对称性），我们有

$$\frac{\partial}{\partial x}\big(\frac{\partial f}{\partial y}\big)(0,0)=1.$$

这表明 *Clairaut-Schwarz* 定理并不成立。请思考定理中的哪个条件没有被满足。

多元函数的 Taylor 展开

我们现在来证明高维空间 Lagrange 余项 Taylor 公式，也就是在一个点附近用高次的多项式函数来逼近函数。证明的想法很直接：将问题沿不同的方向化为 1 维的情形。其余余项的 Taylor 公式证明是类似的。

定理 188. 假设 $\Omega\subset\mathbb{R}^n$ 是凸的开集，f 在 Ω 上 $k+1$ 次可微分（我们通常要求更多的条件：f 不超过 $k+1$ 阶的偏导数存在并且连续，这对于应用来说是足够的）。那么，对于任意的 $x,y\in\Omega$，其中 $x=(x_1,\cdots,x_n)$，$y=(y_1,\cdots,y_n)$，存在 $\vartheta\in[0,1]$（可能依赖于 x 和 y），使得

$$f(y_1,\cdots,y_n)=\sum_{|\alpha|\leqslant k}\frac{\partial^\alpha f(x)}{\alpha!}(x-y)^\alpha+\sum_{|\alpha|=k+1}\frac{\partial^\alpha f\big(x+\vartheta(y-x)\big)}{\alpha!}(x-y)^\alpha,$$

其中，

$$(x-y)^\alpha=(x_1-y_1)^{\alpha_1}\cdots(x_n-y_n)^{\alpha_n},\quad \alpha!=\alpha_1!\cdots\alpha_n!.$$

证明: 我们考虑 x 与 y 之间的连线并把函数 f 限制到这条线上。用分析的语言写，我们考虑函数

$$g:(-\varepsilon,1+\varepsilon)\to\mathbb{R},\quad g(t)=f\big(x+t(y-x)\big).$$

这里用到了 Ω 的凸性。我们首先来计算 g 的 k-次导数（$k\geqslant 0$）：

$$g^{(k)}(t)=\sum_{|\alpha|=k}\frac{k!}{\alpha!}(\partial^\alpha f)\big(x+t(y-x)\big)(y-x)^\alpha.$$

这个式子可以用归纳法来证明：$k=0$ 是显然的。假设

$$g^{(k-1)}(t)=\sum_{|\alpha|=k-1}\frac{(k-1)!}{\alpha!}(\partial^\alpha f)\big(x+t(y-x)\big)(y-x)^\alpha.$$

根据方向导数的性质，对任意的函数 f，我们有

$$\frac{d}{dt}f\big(x+tv\big)\Big|_{t=0}=v^i\frac{\partial f}{\partial x_i}(x).$$

所以，

$$\begin{aligned}
g^{(k)}(t)&=\sum_{|\alpha|=k-1}\frac{(k-1)!}{\alpha!}\Big[(\partial^\alpha f)\big(x+t(y-x)\big)\Big]'(y-x)^\alpha\\
&=\sum_{|\alpha|=k-1}\frac{(k-1)!}{\alpha!}\Big[\sum_{m=1}^n(y_m-x_m)\Big(\frac{\partial}{\partial x_m}(\partial^\alpha f)\Big)\big(x+t(y-x)\big)\Big](y-x)^\alpha\\
&=\sum_{|\beta|=k}\frac{k!}{\beta!}(\partial^\beta f)\big(x+t(y-x)\big)(y-x)^\beta.
\end{aligned}$$

最后一个等号需要搞清楚如下的组合性质：为了从前一步的某个 $|\alpha|=k-1$ 得到固定的 β，其中 $|\beta|=k$，有如下 n 中可能性：

$$\alpha_1=(\beta_1-1,\cdots,\beta_n)\longrightarrow\beta,\ \alpha_2=(\beta_1,\beta_2-1,\cdots,\beta_n)\longrightarrow\beta,\cdots,\alpha_n=(\beta_1,\cdots,\beta_n-1)\longrightarrow\beta.$$

所以，所求的系数应该是（这个等价于 $\beta_1+\cdots+\beta_n=k$）：

$$\sum_{m=1}^n\frac{(k-1)!}{\alpha_m!}=\frac{k!}{\beta!}.$$

现在对 g 用 Taylor 公式（在 0 和 1）之间，我们有 $\vartheta\in[0,1]$，使得

$$g(1)=\sum_{\ell\leqslant k}\frac{g^{(\ell)}(0)}{\ell!}+\frac{g^{(k+1)}(\vartheta)}{(k+1)!}.$$

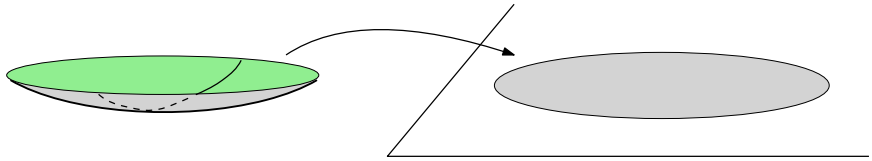
将 $g(t)$ 的值代入即可。 □

$\mathbb{R}^n$ 中的光滑子流形：曲线和曲面的定义

在 $\mathbb{R}^n$ 中，我们考虑 $\mathbb{R}^k \subset \mathbb{R}^n$，此时的 $\mathbb{R}^k$ 是按照下面精确的方式定义的

$$\mathbb{R}^k = \big\{(x_1, \cdots, x_n)\big| x_n = x_{n-1} = \cdots = x_{k+1} = 0\big\}.$$

这是由 $n-k$ 个线性函数的零点定义出来的线性子空间。这就是**最基本的**k-维子流形的例子。所谓的 k 子流形，$\mathbb{R}^n$ 中的一个子集，在每个局部上来看，它长得样子就像是线性的 $\mathbb{R}^k$ 落在 $\mathbb{R}^n$ 中一样。我们可以设想 $\mathbb{R}^3$ 中的一个曲面，在一个点的附近，这个曲面和切平面很近，从而变化不大。它的样子大概是



曲面的两面用灰色和绿色表示，其中绿色的一面上画有一条线。这个曲面落在 $\mathbb{R}^3$ 中和右边的 $\mathbb{R}^3$ 中的 $\mathbb{R}^2$ 上的圆盘很相似。

为了理解子流形的概念，我们先看两个（不平凡的）例子，它们维数或者余维数为 1（在 $\mathbb{R}^3$ 种这已经给出了所有的维数）。维数为 0 的基本例子是 $\mathbb{R}^n$ 中的点，余维数为 0 基本例子是 $\mathbb{R}^n$ 本身，这些都没有太多的意思。另外，子流形是一个局部的概念，所以我们现在只关心局部的情况。

我们首先研究曲线。

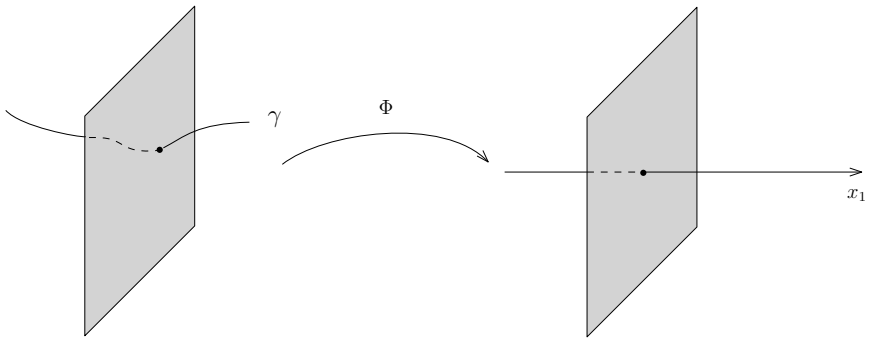
假设 $\gamma : (-\varepsilon, \varepsilon) \to \mathbb{R}^n$ 是光滑的映射（C^1 就可以），我们要求对任意的 $t \in (-\varepsilon, \varepsilon)$，$\gamma'(t) \neq 0$。这个条件称作是曲线的非退化条件，一定程度上可以认为是要排除 $\gamma(t) \equiv p$ 这种极端的例子（此时，我们得到的是一个点而不是曲线）。由于 $\gamma'(0) \neq 0$，我们不妨假设 $\gamma_n'(0) > 0$（< 0 的情况可以类似地讨论），即 $\gamma(t)$ 的 x_1 分量在 0 附近（我们总可以假设 ε 很小）是严格递增的。我们要说明 γ 在 $\gamma(0)$ 附近与 x_1 轴 $\mathbb{R}^1 = \{(x_1, 0, \cdots, 0)\}$ 非常相似。事实上，我们构造映射

$$\Phi : \mathbb{R}^n \to \mathbb{R}^n, \ \ (x_1, x_2, \cdots, x_n) \mapsto (\gamma_1^{-1}(x_1), x_2 - \gamma_2\big(\gamma_1^{-1}(x_1)\big), , \cdots, x_n - \gamma_n\big(\gamma_1^{-1}(x_1)\big)).$$

在 0 附近，这是一个良好定义并且是可逆的映射：它的逆可以用下面的公式表达

$$\Phi^{-1} : \mathbb{R}^n \to \mathbb{R}^n, \ \ (y_1, y_2, \cdots, y_n) \mapsto (\gamma_1(y_1), y_2 + \gamma_2(y_1), , \cdots, y_n + \gamma_n(y_1)).$$

不难看出，这是一个微分同胚。所在，在用一个微分同胚把曲线（的像）拉直了之后或者说换到了 $(y_1, \cdots, y_n)$ 这个坐标系下看，我们的曲线就是标准的 x_1-轴：



其次，我们研究由方程的图像定义的曲面。给定函数 $f:\mathbb{R}^{n-1}\to\mathbb{R}$。我们考虑它的图像

$$\Gamma_f=\{(x,f(x))\in\mathbb{R}^{n-1}\times\mathbb{R}\subset\mathbb{R}^n|x\in\Omega\}.$$

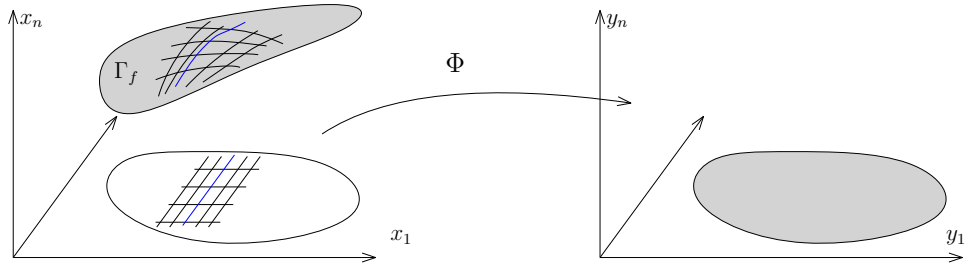
这是 $\mathbb{R}^n$ 中的超曲面（余维数为 1）。我们定义映射

$$\Phi:\mathbb{R}^n\to\mathbb{R}^n,\ \ (x_1,x_2,\cdots,x_n)\mapsto(y_1,\cdots,y_n)=(x_1,x_2,\cdots,x_n-f(x_1,\cdots,x_{n-1})).$$

很明显，Φ 是双射，其逆映射为

$$\Phi^{-1}:\mathbb{R}^n\to\mathbb{R}^n,\ \ (y_1,y_2,\cdots,y_n)\mapsto(x_1,\cdots,x_n)=(y_1,y_2,\cdots,y_n+f(y_1,\cdots,y_{n-1})).$$

这两个映射显然是光滑的，所以 Φ 是微分同胚。从而，通过用一个微分同胚把这个曲面拉直了之后或者说换到了 $(y_1,\cdots,y_n)$ 这个坐标系下看，我们的曲线就是标准的 $\mathbb{R}^{n-1}$（由一个线性方程的零点给出）：



最终，我们再研究一个经典的例子，考虑 3 维空间中的单位球面 $\mathbf{S}^2\subset\mathbb{R}^3$，也即是

$$\mathbf{S}^2=\{(x,y,z)\big|x^2+y^2+z^2=1.\}$$

很明显，$\mathbf{S}^2$ 可以被写成 6 个半球面的并：

$$\mathbf{S}^2=\mathbf{S}^2_{x>0}\cup\mathbf{S}^2_{x<0}\cup\mathbf{S}^2_{y>0}\cup\mathbf{S}^2_{y<0}\cup\mathbf{S}^2_{z>0}\cup\mathbf{S}^2_{z<0}.$$

每一个部分都是一个函数的图像，比如说，

$$\mathbf{S}^2_{y<0}=\big\{\big(x,z,f(x,z)\big|f(x,z)=-\sqrt{1-x^2-z^2},x^2+z^2<1\big)\big\},$$

其中，函数 f 定义在 (x,z) 平面的一个单位圆盘的内部。所以说，$\mathbf{S}^2$ 在局部上来看是一个微分子流形。

练习. 证明，$\mathbf{S}^2\subset\mathbb{R}^3$ 在如下的意义下永远都不是一个函数的图像：不存在微分同胚

$$\Phi:\mathbb{R}^3\to\mathbb{R}^2_{y_1,y_2}\times\mathbb{R}^1_{y_3},$$

和 $\mathbb{R}^2_{y_1,y_2}$ 中的区域（任意区域不一定是开集）$\Omega\subset$ 以及 Ω 上的光滑函数 f，使得 $\Phi(\mathbf{S}^2)$ 是 f 的图像

$$\Gamma_f=\big\{(y_1,y_2,f(y_1,y_2)\in\mathbb{R}^2_{y_1,y_2}\times\mathbb{R}^1_{y_3}\big|(y_1,y_2)\in\Omega\big\}.$$

将这些例子作为基本的图像，我们就可想办法定义 $\mathbb{R}^n$ 中的子流形了，这是所有的那些在局部上复合一个微分同胚（换一下坐标系）之后就变成了线性子空间的一部分的那样的子集合。：[8]

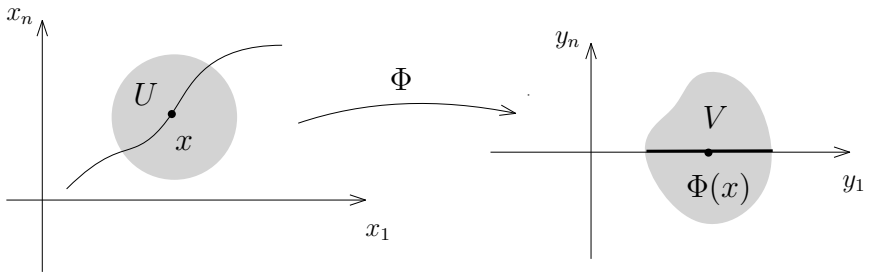
定义 189 (d-维子流形)**.** 假设 $M \subset \mathbb{R}^n$ 是非空子集，这个 $\mathbb{R}^n$ 中的坐标用 (x_i) 表示。如果存在整数 $d \geqslant 0$，使得对任意的 $x \in M$，存在开集 $U \subset \mathbb{R}^n$，$x \in U$ 以及（可能是另外一个）$\mathbb{R}^n$ 中的开集（这个 $\mathbb{R}^n$ 中的坐标用 (y_i) 表示）V 以及微分同胚

$$\Phi : U \to V$$

使得

$$\Phi : U \cap M = V \cap (\mathbb{R}^d \times \{0\}),$$

其中我们把 $\mathbb{R}^n = \mathbb{R}^{d+(n-d)}$ 写成 $\mathbb{R}^d \times \mathbb{R}^{n-d}$，上述的表达式要求后面的 $n-d$ 个坐标都取 0，我们就称 M 是 $\mathbb{R}^n$ 的一个 d-维的**（微分）子流形**。其中，d 称作是 M 的**维数**，记作 $\dim M$。



注记. 我们还把 $\operatorname{codim} M = n - \dim M$ 称作是 M 的**余维数**，它有着如下具体的含义：存在 $\operatorname{codim} M$ 个函数，使得 M 恰好是这些函数的公共零点集合。实际上，我们可以取 $f_j = \Phi^* y_j$，其中 $j = d+1, d+2, \cdots, n$。

我们需要说明维数 d 是良好的定义。换句话说，把 M 局部上看成是某个 d 维的线性子空间中的集合的方式又可能不唯一。按照定义，对于给定的 x，我们有可能有（很明显有很多）另外的 $d' \geqslant 0$ 和开集 U'，$x \in U'$ 以及（可能是另外一个）$\mathbb{R}^n$ 中的开集（这个 $\mathbb{R}^n$ 中的坐标用 (z_i) 表示）V' 以及微分同胚

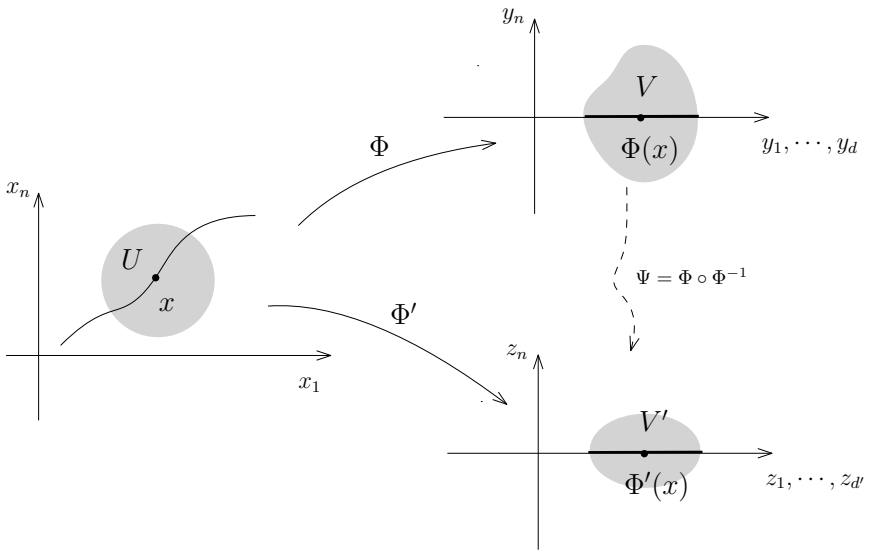
$$\Phi' : U' \to V'$$

使得

$$\Phi' : U' \cap M = V' \cap (\mathbb{R}^{d'} \times \{0\}),$$

通过考虑 $U \cap U'$，我们可以要求 $U = U'$。

[8]根据 Whitney 的定理，这个定义实际上包含了所有的微分流形。



我们考虑映射

$$\Psi = \Phi' \circ \Phi^{-1} : V \to V'.$$

这是两个微分同胚的复合，所以还是微分同胚。特别地，如果我们把映射 Ψ 限制到 $V \cap (\mathbb{R}^d \times \{0\})$ 上面，我们就得到了微分同胚

$$\Psi : V \cap (\mathbb{R}^{d'} \times \{0\}) \to V' \cap (\mathbb{R}^{d'} \times \{0\}).$$

这是 $\mathbb{R}^d$ 和 $\mathbb{R}^{d'}$ 中的两个开集之间的微分同胚，我们上周已经利用复合映射的链式法则证明了 $d = d'$。

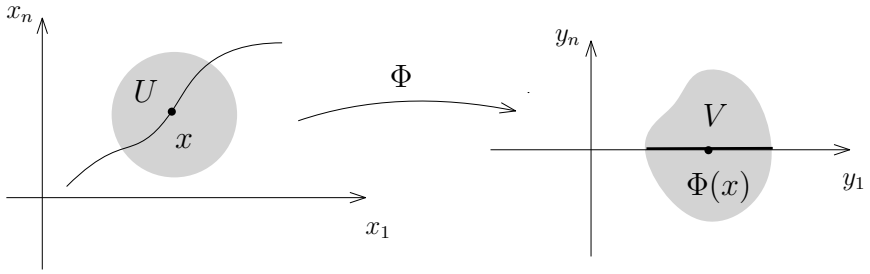
33 子流形用坐标表示，子流形作为零点集来实现，反函数定理及证明，反函数定理与坐标变换之间的关系

二零二零年二月二十七日，星期四，阴

子流形定义的一个注解

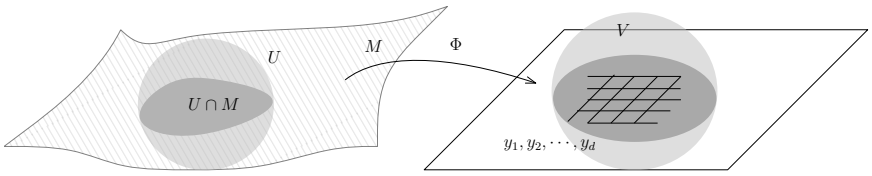
我们首先简单回顾一下子流形的定义：

定义 190 (d-维子流形)**.** 假设 $M \subset \mathbb{R}^n$ 是非空子集。如果存在整数 $d \geqslant 0$，使得对任意的 $x \in M$，存在开集 $U \subset \mathbb{R}^n$，$x \in U$ 以及 $\mathbb{R}^n$ 中的开集 V 以及微分同胚 $\Phi : U \to V$ 使得 $\Phi : U \cap M = V \cap (\mathbb{R}^d \times \{0\})$（后面 $n-d$ 个坐标为 0），我们就称 M 是 $\mathbb{R}^n$ 的一个 d-维的**（微分）子流形**。



注记. （子流形上的局部行为）

1) 由于 $V \cap (\mathbb{R}^d \times \{0\})$ 是 $\mathbb{R}^d$ 中的开集，所以，$U \cap M$ 与 $\mathbb{R}^d$ 中的一个开集是微分同胚的，所以 M 的每个点的附近（与 $\mathbb{R}^n$ 中的某个开集 U 的交）都与 $\mathbb{R}^d$ 中的某个开集是微分同胚的。我们通常说，局部上 M 是 $\mathbb{R}^n$，因为我们可以进一步取更小的邻域，使得 $U \cap M$ 与 $\mathbb{R}^d$ 中的开球是微分同胚的。



2) 由于在 $V \cap \mathbb{R}^d$ 上可以选取 $y_1, \cdots, y_d$ 作为坐标，所以 $\{\Phi^* y_i\}_{i \leqslant d}$ 可以作为 $M \cap U$ 上的一个局部坐标。所以，我们可以**局部上**用 d 个坐标来标记 M 上的点。当然，我们可能有很多种不同的方式来选取这样的坐标。

注记. （子流形局部上作为函数的公共零点）根据 $\Phi : U \to V$，在右边的 V 上有 $n-d$ 个光滑函数 $y_{d+1}, \cdots, y_d$，它们的公共零点集合

$$\left\{(y_1, \cdots, y_n) \in V \middle| y_{d+1} = y_{d+2} = \cdots = y_n = 0\right\}$$

恰好给出了 $V \cap \mathbb{R}^d$；所以，通过微分同胚 Φ，我们知道在左边的 U 上也存在 $n-d$ 个光滑函数 $\Phi^* y_{d+1}, \cdots, \Phi^* y_d$，它们的公共零点集恰好是 $M \cap U$。

这是我们基本的几何直观：每一个 $\Phi^* y_i = 0$ 给出了一个关系，使得空间的维数降低了恰好 1 维。我们在线性代数的课上很熟悉这样的例子：如果

$$L_k : V \to \mathbb{R}, \quad x \mapsto L_k(x)$$

是 n-维 $\mathbb{R}$-线性空间 V 上的 s 个线性函数，那么它们的公共零点集

$$W = \{x \in V \big| L_1(x) = \cdots = L_s(x) = 0\}$$

是 $n-s$ 维的线性空间。当然，我们必须要求这些 L_k 们是线性无关的。所谓的隐函数定理就是说如果 U 上有 s 个"无关"的光滑函数，那么它们的公共零点集就是一个 $n-s$ 维的子流形。

反函数定理

简单而言，反函数（和隐函数）定理说的是映射的线性化（微分）决定了映射的局部性质。我们先看一个有意思的例子，它讲的是在一定条件下微分是线性同构这个代数事实保证了映射的逆具有连续可微性这个分析的性质。

引理 191 (正则性引理)**.** 给定开集 $U \subset \mathbb{R}^n$ 和开集 $V \subset \mathbb{R}^n$，映射 $f : U \to V$ 是**同胚**，即 f 是双射并且 f 和 f^{-1} 都是连续的。如果 $f \in C^1(U, V)$，那么如下两条是等价的：

1) f^{-1} 是 C^1 的；

2) 对任意的 $x \in U$，$df(x) : T_x U \to T_{f(x)} V$ 是同构（即其 *Jacobi* 矩阵的行列式非零）。

注记. 当第二个条件不成立的时候，f^{-1} 可能不是可微的，比方说，我们可以考虑

$$f : \mathbb{R} \to \mathbb{R}, \quad x \mapsto x^3.$$

此时，f 和 f^{-1} 都是连续的，但是 f^{-1} 在原点处不可微。

证明: 我们注意到 1) ⇒ 2) 是显然的，因为如果 $g = f^{-1}$ 也可微，所以对 $g \circ f = \mathrm{Id}_U$ 和 $f \circ g = \mathrm{Id}_V$ 求微分，我们就得到了 dg 的逆，这实际上在推倒逆映射的微分公式时我们已经证明过了，请参考第二次课的笔记。

下面证明 2) ⇒ 1)。为了行文方便，我们先引入几个记号：对于给定的 $x \in U$，令 $y = f(x) \in V$，$f^{-1} = g$，$A = df(x)$，$B = A^{-1}$（我们假设了 A 可逆），我们要证明 g 在 y 处可微并且其微分恰为 B，即

$$\lim_{h \to 0} \frac{|\Delta(h)|}{|h|} = 0,$$

其中

$$\Delta(h) = g(y+h) - g(y) - B(h).$$

由于 f 在 x 处是可微的，所以对任意的 $k \in \mathbb{R}^n$，$k \to 0$，我们有

$$f(x+k) = f(x) + A(k) + \delta(k),$$

其中 $\delta(k)=o(k)$。根据 $g=f^{-1}$ 的连续性，我们有

$$\begin{aligned} y+h &= f\big(g(y+h)\big) = f\big(x+\overbrace{B(h)+\Delta(h)}^{o(1)}\big) \\ &= y+\underbrace{A\big(B(h)}_{=h}+\Delta(h)\big)+\delta\big(B(h)+\Delta(h)\big). \end{aligned}$$

左右两边消去相同的项，我们得到

$$0 = A\big(\Delta(h)\big)+\delta\big(B(h)+\Delta(h)\big).$$

作用上 A 的逆，我们得到

$$\begin{aligned} \big|\Delta(h)\big| &= \Big|B\left(\delta\big(B(h)+\Delta(h)\big)\right)\Big| \\ &\leqslant C_1|\delta\big(B(h)+\Delta(h)| \leqslant o(B(h))+o(\Delta(h)) \\ &= o(h)+o(\Delta(h)). \end{aligned}$$

从而，当 h 较小时，我们有

$$\big|\Delta(h)\big| \leqslant o(h)+0.01\big|\Delta(h)\big| \quad \Rightarrow \big|\Delta(h)\big| \leqslant \frac{100}{99}o(h) = o(h).$$

为了完成命题的证明，我们还需要说明 f^{-1} 的微分是连续的。然而，根据逆映射的微分公式（现在 g 可微所以可以用这个公式了），我们有

$$dg(y) = \Big(df\big(g(y)\big)\Big)^{-1}.$$

由于 df 连续，g 连续并且矩阵的求逆也是连续的，所以 dg 连续。 □

我们现在可以证明反函数定理

定理 192 (反函数定理)**.** *$\Omega\subset\mathbb{R}^n$ 是开集，$f:\Omega\to\mathbb{R}^n$ 是 C^1 的。如果对于点 $x_0\in\Omega$，f 在此点的微分 $df(x_0)$ 是可逆的，那么 f 在 x_0 的局部上是 C^1 的微分同胚，即存在开集 $U\subset\Omega$，$x_0\in U$ 和开集 $V\subset\mathbb{R}^n$，使得 f 在 U 上的限制给出的 $f:U\to V$ 双射并且 f 与 f^{-1} 均为 C^1。*

我们先回忆如下的不动点定理（已经证明，参考上学期第七课的讲义），这是我们证明的主要工具：

定理 193 (Banach/Picard 不动点定理)**.** *(X,d) 是完备的距离空间，$T:X\to X$ 是一个**压缩映射**，即存在正常数 $\gamma<1$，使得对任意的 $x,y\in X$，都有*

$$d\big(T(x),T(x')\big) \leqslant \gamma d(x,x').$$

那么，T 有唯一的不动点，即存在唯一的 $x_0\in X$，使得 $T(x_0)=x_0$。

反函数定理的证明. 问题的证明分为三步，其难点在于局部上能构造出映射 f 的逆。我们不妨假设 $x_0 = 0$，$f(x_0) = 0$ 并且 $df(x_0) = \mathrm{Id}$，也就是说作为 $\mathbb{R}^n \to \mathbb{R}^n$ 的线性映射可以用单位矩阵来表示。事实上，我们可以考虑，映射

$$F : \mathbb{R}^n \to \mathbb{R}^n, \quad x \mapsto df(x_0)^{-1}\Big(\big(f(x_0 + x) - f(x_0)\big)\Big).$$

用 F 来代替 f，那么 F 就满足上述要求。然而，如果考虑

$$df(x_0) : \mathbb{R}^n \to \mathbb{R}^n, \quad x \mapsto df(x_0)(x),$$

和

$$\tau_v : \mathbb{R}^n \to \mathbb{R}^n, \quad x \mapsto x + v,$$

那么，$f = \tau_{f(x_0)} \circ df(x_0)^{-1} \circ F \circ \tau_{-x_0}$，这是 F 和一些微分同胚的复合，所以 f 局部上也是微分同胚。

我们以下假定 f 满足这些要求。

第一步，翻译成不动点问题。先来证明局部上是满射，即对于 $y \in V$（某个 V，后面会有具体的限制），要找 $x \in \Omega$，使得

$$f(x) = y.$$

上面的式子等价于

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

所以，我们定义 $T(x) = x - f(x) + y$，T 的不动点 x 即为所求。

第二步，验证映射的压缩条件。首先注意到

$$dT(0) = df(0) - \mathrm{Id} = 0.$$

根据 df 的连续性，存在 $r > 0$，使得对任意的 $x \in B(0, r)$（这是中心在原点半径为 r 的小球）上，我们有

$$\|dT(x)\| < \frac{1}{10}.$$

其中，矩阵的范数 $\|dT(x)\|$ 我们可以任意地事先取定。

首先考虑 $T(x) - T(x')$，$x, x' \in B(0, r)$。我们定义 x 与 x' 连线上的函数，即

$$g(t) = f\big(tx + (1-t)x'\big) - (x + (1-t)x').$$

请注意，这是向量值的函数，我们仍然可以定义积分（即在每个分量上定义）。根据 Newton-Leibniz 公式，我们有

$$\begin{aligned} T(x) - T(x') &= \big(f(x) - x\big) - \big(f(x') - x'\big) \\ &= g(1) - g(0) = \int_0^1 g'(t)dt \\ &= \int_0^1 df\big(tx + (1-t)x'\big)(x - x') - (x - x')dt. \end{aligned}$$

从而，

$$\begin{aligned}\|T(x)-T(x')\| &= \int_0^1 \|T\big(tx+(1-t)x'\big)(x-x')\|dt\\ &\leqslant \int_0^1 \frac{1}{10}\|(x-x')\|dt\\ &= \frac{1}{10}\|(x-x')\|.\end{aligned}$$

从而，T 是压缩的。

另外，为了利用压缩映像定理，我们要找到距离空间 X，使得 $f(X)\subset X$。实际上，我们取 $X=B(0,r)$。根据上面的不等式，对于 $x'=0$ 和 $x\in B(0,r)$，按照定义，我们有

$$\begin{aligned}\|T(x)\| &\leqslant \|T(x)-T(0)\|+\|T(0)\|\\ &= \frac{1}{10}|x|+|y|\\ &< r.\end{aligned}$$

其中，我们要求 $y\in B(0,\frac{r}{2})$（实际上，这里就选取了 V）。这表明 $T(x)\in B(0,r)$。所以，对于 $X=\overline{B(0,r)}$（这是 $\mathbb{R}^n$ 中的闭球所以完备），我们有

$$T: X\to X.$$

根据不动点定理，对任意的 $y\in B(0,\frac{r}{2})$，都存在唯一的 $x\in B(0,r)$，使得 $f(x)=y$，即 $T(x)=x$。

综合上面的讨论，我们令

$$U=B(0,r)\cap T^{-1}\big(B(0,\frac{r}{2})\big),\quad V=B(0,\frac{r}{2}).$$

很明显，这是两个开集。按照我们的构造方式我们有满射

$$f: U\to V.$$

这实际上是一个双射。为此，我们利用 $df(x)$ 的连续性，对任意的 $\varepsilon>0$，只要 r 实现选的足够小，我们就有 $\|df(x)-\mathrm{Id}\|<0.1$。我们可以选 $\varepsilon=0.1$，此时 r 也可以确定。如果 $f(x)=f(x')$，令

$$h(t)=f\big(tx+(1-t)x'\big).$$

那么，我们通过对 $h(t)$ 用 Newton-Leibniz 公式，就有

$$\begin{aligned}0=&\|f(x)-f(x')\| = \int_0^1 df\big(tx+(1-t)x'\big)(x-x')dt\\ &= x-x'+\int_0^1 \Big(df\big(tx+(1-t)x'\big)-\mathrm{Id}\Big)(x-x').\end{aligned}$$

从而，

$$\|x-x'\|\leqslant \int_0^1 0.1\times\|x-x'\|dt = 0.1\times\|x-x'\|.$$

所以，$x = x'$。

第三步，正则性条件。为了利用正则性引理，我们需要说明 $g = f^{-1}$ 是连续的。任取 $y, y' \in V$，假设 $g(y) = x$，$g(y') = x'$，按照构造，我们有

$$\begin{aligned} |g(y) - g(y')| = |x - x'| &= |T(x) - T(x') + (y - y')| \\ &\leqslant |T(x) - T(x')| + |y - y'| \\ &\leqslant \frac{1}{10}|x - x'| + |y - y'| \\ &= \frac{1}{10}|g(y) - g(y')| + |y - y'|. \end{aligned}$$

从而

$$|g(y) - g(y')| \leqslant 1.1 \times |y - y'|.$$

这表明 g 是连续的。 □

注记. 反函数定理对任意的（无限维）完备赋范线性空间都成立，建议有兴趣的同学可以查阅并对比我们的证明。

我们还可以要求反函数定理中的映射有更高的正则性：

推论 194. 假设 $k \geqslant 2$ 是整数。在反函数定理的假设中，如果我们进一步要求 $f \in C^k(\Omega)$，那么得到的逆映射 $f^{-1} \in C^k(V)$。

证明: 根据链式法则，f^{-1} 的微分是 $d(f^{-1})(y) = (df)^{-1}(f^{-1}(y))$，其中 $y \in V$。令 $g = f^{-1}$。我们在 U 和 V 上任意选取坐标 (x_i) 和 (y_j)，假设 f 和 g 的 Jacobi 矩阵分别是 A 和 B，即

$$df(x) = A(x) = (A_{ij}(x_1, \cdots, x_n)), \quad dg(y) = B(y) = (B_{ij}(y_1, \cdots, y_n)).$$

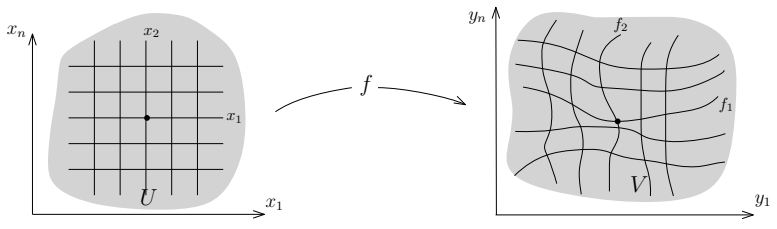
根据要求 $A_{ij} \in C^{k-1}(U)$。然而，$B(y) = A^{-1}(g(y))$。根据逆矩阵的计算，我们知道 B_{ij} 形如

$$B_{ij}(y) = \sum \frac{A_{i'j'}(g(y))A_{i''j''}(g(y)) \cdots A_{i'''j'''}(g(y))}{\det A(y)}.$$

在求 ℓ 次导数的时候，分母上只会出现 $\det A$，而导数都出现在分子上面，而由链式法导致的 g 的导数至多出现到 $\ell - 1$ 阶。据此，我们知道 $B_{ij}(y)$ 是 C^{k-1} 的，从而 $g = f^{-1}$ 是 C^k 的。 □

注记. 为了说明 f^{-1} 是高阶光滑的，这个引理说只要验证 f 的一阶导数就可以做到这一点，这在技术上提供了莫大的方便。

注记 (几何解释：坐标变换)**.** 我们可以用选取坐标系的观点来解释反函数定理：



为了描述 V 上的点，通过 f，我们实际上可以选取 $(x_, \cdots, x_n)$ 作为坐标系（而不是 $(y_1, \cdots, y_n)$），也就是说用 U 中的点（的坐标）来参数化 V。

33.1 习题课：拓扑空间

清华大学 19-20 春季学期，数学分析二，第一次习题课内容

拓扑空间的补充

关于拓扑空间的补充

定义 195 (拓扑空间)**.** *X 是集合，$\mathcal{T}=\{U|U\subset X\}$ 是 X 的某些子集所组成的集合。如果下面三个条件成立*

1) $\emptyset\in\mathcal{T}$，$X\in\mathcal{T}$。

2) 对任意的 $\{U_\alpha\}_{\alpha\in\mathcal{A}}\subset\mathcal{T}$，其中 $\mathcal{A}$ 为指标集合，我们有 $\bigcup\limits_{\alpha\in\mathcal{A}}U_\alpha\in\mathcal{T}$。

3) 对任意有限个 $U_1,U_2,\cdots,U_m\in\mathcal{T}$，我们有 $\bigcap\limits_{1\leqslant i\leqslant m}U_i\in\mathcal{T}$。

*我们就称 $\mathcal{T}$ 是 X 上的一个**拓扑**，每个 $U\in\mathcal{T}$ 都被称作是（拓扑 $\mathcal{T}$ 下的）**开集**。我们把二元组 $(X,\mathcal{T})$ 称作是一个**拓扑空间**。*

给定拓扑空间 $(X,\mathcal{T})$，考虑子集 $F\subset X$，如果其补集 $X-F$ 是开集（即 $X-F\in\mathcal{T}$），我们就称 F 是在拓扑 $\mathcal{T}$ 下的）**闭集**。对于闭集，我们有如下的性质：

1) $\emptyset$ 和 X 都是闭集。

2) 任意多闭集的交集是闭集。

3) 有限个闭集的并集是闭集。

例子 (距离空间)**.** *(X,d) 是距离空间，对任意的 $x\in X$，以 x 为中心以 r 为半径的开球我们记作*

$$B(x,r)=\{y\in X|d(y,x)<r\}.$$

*如果 $U\subset X$ 是若干开球的并，我们就称 U 是 (X,d) 中开集。一个等价的说法是对任意的 $x\in U$，都存在 $\delta>0$，使得 $B(x,\delta)\subset U$。我们用 $\mathcal{T}_d$ 表示按照这种特定方式所定义的开集所组成的集合。我们上学期已经说明了 $\emptyset$ 和 X 都是开集；任意多开集的并集还是开集；有限个开集的交集还是开集。换而言之，$(X,\mathcal{T}_d)$ 是拓扑空间。我们将 $\mathcal{T}_d$ 称作是**由距离函数 d 所定义的标准拓扑**。距离空间的例子是我们唯一关心的拓扑空间。*

我们引入两个拓扑空间之间的连续映射的概念。从范畴论的观点来看，连续映射是拓扑空间之间最自然的映射，这个与我们学过的其它空间之间的映射非常相似：在线性空间之间应该考虑线性映射，在群之间应该考虑群同态或者这个学期在两个微分子流形之间应该考虑光滑映射。

定义 196 (连续映射：开集的逆像是开集)**.** 假设 $(X,\mathcal{T})$ 和 $(X',\mathcal{T}')$ 是拓扑空间，$f:X\to X'$ 是映射。如果对每个 $U'\in\mathcal{T}'$，我们有 $f^{-1}(U')\in\mathcal{T}$，我们就称 f 是**连续映射**。

练习. 连续映射的复合是连续的，即若 $(X,\mathcal{T})$，$(X',\mathcal{T}')$ 和 $(X'',\mathcal{T}'')$ 是拓扑空间，$f:X\to X'$ 和 $f':X'\to X''$ 是连续映射，那么

$$f'\circ f:X\to X''$$

是连续映射。

练习. (X,d) 和 (X',d') 是距离空间，$\mathcal{T}_d$ 和 $\mathcal{T}_{d'}$ 分别是 X 和 X' 上由其距离函数所定义的拓扑。假设 $f:X\to X'$ 是映射，那么下面两个概念是等价的：

1) $f:(X,d)\to(X',d')$ 是距离空间之间的连续映射。（由点列的收敛定义的）

2) $f:(X,\mathcal{T}_d)\to(X',\mathcal{T}_{d'})$ 是拓扑空间之间的连续映射。

我们引入拓扑**子空间**的概念。这学期课程中将要研究 $\mathbb{R}^n$ 中曲线与曲面等都是这样的例子。

练习. $(X,\mathcal{T})$ 是拓扑空间，$Y\subset X$ 是子集。我们考虑 X 上的开集与 Y 的交所构成的集合：

$$\mathcal{T}\big|_Y=\{U\cap Y\big|U\in\mathcal{T}\}.$$

1) 证明，$\mathcal{T}\big|_Y$ 是 Y 上的一个拓扑。我们称 $(Y,\mathcal{T}_Y)$ 是 $(X,\mathcal{T})$ 的**子（拓扑）空间**。

2) 证明，包含映射

$$\iota:Y\to X,\quad y\mapsto y$$

是 $(Y,\mathcal{T}_Y)$ 与 $(X,\mathcal{T})$ 之间的连续映射。

3) 假设 $(Y,\widetilde{\mathcal{T}})$ 是拓扑空间，即 $\widetilde{\mathcal{T}}$ 是 Y 上的拓扑，使得包含映射

$$\iota:Y\to X,\quad y\mapsto y$$

是连续的，那么 $\widetilde{\mathcal{T}}\supset\mathcal{T}_Y$。换而言之，$\mathcal{T}_Y$ 是使得包含映射为连续映射的最小的拓扑。

关于距离空间我们也有子空间的概念：若 (X,d) 是距离空间，$Y\subset X$ 是子集，对任意的 $y_1,y_2\in Y$，我们定义

$$d_Y(y_1,y_2)=d(y_1,y_2).$$

这个 $d_Y:Y\times Y\to\mathbb{R}_{\geqslant 0}$ 定义出了 Y 上的距离函数。我们将 (Y,d_Y) 称作是 (X,d) 的**子（距离）空间**

练习. (X,d) 是一个度量空间，(Y,d_Y) 是 (X,d) 的子距离空间。那么，集合 Y 上有两种拓扑：

- 拓扑 $\mathcal{T}_d\big|_Y$：先用 X 的距离函数给出 X 上的开集从而使得 X 成为一个拓扑空间 $(X,\mathcal{T}_d)$，然后将这个拓扑限制到 Y 上。

- 拓扑 $\mathcal{T}_{d_Y}$：先用 X 上的距离函数限制到 Y 上使得 Y 成为距离空间，然后用个距离来定义拓扑。

证明，$\mathcal{T}_d\big|_Y = \mathcal{T}_{d_Y}$。换而言之，为了得到距离空间的子集上的拓扑，我们可以先将它视作是距离空间再取距离所诱导的拓扑，也可以先将全空间视作是拓扑空间然后取子空间拓扑，这两种操作是交换的。

最后我们引入同胚的概念：

定义 197. 假设 $(X,\mathcal{T})$ 和 $(X',\mathcal{T}')$ 是拓扑空间，$f: X \to X'$ 是连续的双射。如果逆映射 $f^{-1}: X' \to X$ 也是连续的，我们就称 f 是 $(X,\mathcal{T})$ 与 $(X',\mathcal{T}')$ 之间的**同胚**。此时，我们通常说 X 和 X' 是同胚的。

同胚是拓扑空间之间的等价性。如果两个拓扑空间同胚，除非在某些特定的场合，我们将不再区分它们。如果我们考虑的是距离空间之间的等价性，我们通常要求距离空间之间是**等距同构的**，即若 (X,d) 与 $(\overline{X},\overline{d})$ 是距离空间，如果存在双射 $f:(X,d)\to(\overline{X},\overline{d})$，使得对任意的 $x,y\in X$，

$$\overline{d}(\iota(x),\iota(y)) = d(x,y),$$

我们就称 f 是**等距同构**。

练习. 我们考虑几个拓扑空间之间的连续映射的例子：

1) $\mathbb{R}$ 和 $(0,1)$ 上我们给定最常见的绝对值定义的距离，证明，它们不等距同构但是它们是同胚的。简单来说，拓扑同胚是比等距同构要松散的多的一个等价性。

2) 考虑单位圆周 $\mathbf{S}^1 = \{(x,y)\in\mathbb{R}^2 \big| x^2+y^2=1\}$ 并将它视作是拓扑空间。证明，映射

$$f:[0,1)\to\mathbf{S}^1,\quad x\mapsto(\cos(2\pi x),\sin(2\pi x)),$$

是连续的双射但是不是同胚。

3) 证明，$\mathbf{S}^1$ 不于任何一个 $\mathbb{R}^1$ 的子集（作为拓扑子空间）同胚。

33.2 作业：反函数和隐函数定理

清华大学 19-20 春季学期，数学分析二，作业 2

请用 A4 大小的纸张**正反面**用钢笔，签字笔或者圆珠笔书写，并注明自己的姓名，年级和作业的**总页数**。除定理公式所涉及的人名之外，请使用**中文**。本次作业的提交时间和地点为 **2 月 36 日**上午的课堂上，逾期视作零分。

注记. 本次作业除了问题 C 之外，都与反函数定理相关。

A. 微分同胚与反函数定理的基本习题

A1) 考虑映射函数 $f:\mathbb{R}\to\mathbb{R}$，其中

$$f(x)=\begin{cases} x+x^2\sin\left(\frac{\pi}{x}\right), & x\neq 0;\\ 0, & x=0.\end{cases}$$

证明，f 在 $\mathbb{R}$ 上可微分并且 $df(0)\neq 0$ 但是在 0 附近任意小的邻域上 f 都不是单射。请对比反函数定理说明反函数定理中哪一个条件没有被满足。

A2) 考虑映射

$$f:\mathbb{R}^2\to\mathbb{R}^2,\ \ (x,y)\mapsto(e^x\cos(y),e^x\sin(y)).$$

证明，f 在每个 $(x,y)\neq(0,0)$ 处都满足反函数定理的要求但是 f 不是单射也不是满射。

A3) 我们考虑映射

$$f:\mathbb{R}^2\to\mathbb{R}^2,\ \ (x,y)\mapsto(2x-y+x^2y-2y^5,x+3y-4x^2y^2).$$

- 证明，存在 0 的开邻域 U 和 V，使得 $f:U\to V$ 是微分同胚。
- 假设 $W=f(\mathbb{R}^2)$ 为 f 的像，试问 $f:\mathbb{R}^2\to\mathbb{R}^2$ 是否为微分同胚?

A4) 假设 $f:\mathbb{R}^n\to\mathbb{R}^n$ 是连续可微的。如果存在 $a>0$，使得对任意的 $x,y\in\mathbb{R}^n$，我们都有

$$|f(x)-f(y)|\geqslant a|x-y|,$$

那么，$f:\mathbb{R}^n\to\mathbb{R}^n$ 是 C^1 的微分同胚。（注意：反函数定理只给出局部的同胚）

A5) 假设 $f:\mathbb{R}\to\mathbb{R}$ 是 C^∞ 的并且存在 $0\leqslant\alpha<1$，使得 $|f'(x)|\leqslant\alpha$。定义

$$F:\mathbb{R}^2\to\mathbb{R}^2,\ \ (x,y)\mapsto(x+f(y),y+f(x)).$$

- 证明，F 是 C^∞ 的映射并且对任意的 $(x,y)\in\mathbb{R}^2$，$dF(x,y)$ 可逆。

– 证明，对任意的 $(a,b)\in\mathbb{R}^2$，映射

$$g_{a,b}:\mathbb{R}^2\to\mathbb{R}^2,\ (x,y)\mapsto(a-f(y),b-f(x))$$

是压缩映像。（提示：请仔细研究反函数定理的证明）

– 证明，$F:\mathbb{R}^2\to\mathbb{R}^2$ 是微分同胚。

B. 从局部微分同胚到整体微分同胚

B1) $f,g\in C^1(\mathbb{R})$ 是实值函数，我们假设对任意的 $x,y\in\mathbb{R}$，都有 $f'(x)\neq g'(y)$。我们定义 $F:\mathbb{R}^2\to\mathbb{R}^2$：

$$F:(x,y)\mapsto(x+y,f(x)+g(y))$$

– 证明，F 的像 $U=F(\mathbb{R}^2)\subset\mathbb{R}^2$ 是开集。

– 证明，$F:\mathbb{R}^2\to U$ 是微分同胚。

B3) $\Omega=\{(x,y)\in\mathbb{R}^2|x>0\}\subset\mathbb{R}^2$ 是开集，$(x_0,y_0\in\Omega)$，$p:\Omega\to\mathbb{R}$ 是 C^1 的函数，$p_0=p(x_0,y_0)$。我们定义函数

$$F:\Omega\to\mathbb{R},\ \ (x,y)\mapsto y-y_0+\frac{p(x,y)+p_0}{2}\Big(\frac{1}{x}-\frac{1}{x_0}\Big).$$

– 证明，存在 p_0 点处的一个开邻域 U 以及 U 上的 C^1 函数 ϕ，使得

$$F(x,y)=0\text{ 等价于 }y=\phi(x).$$

– 假设 $p(x,y)=\alpha xy$，$\alpha>0$。证明下面极限的存在性并计算它（这个极限的几何意义是什么?）：

$$\lim_{x\to x_0}\frac{p(x,\varphi(x))-p_0}{x-x_0}.$$

B3) 考虑映射

$$\varphi:\mathbb{R}^2\to\mathbb{R}^2,\ \ (x,y)\mapsto\left(\sin\left(\frac{y}{2}\right)-x,\sin\left(\frac{x}{2}\right)-y\right)$$

1) 证明，这是一个光滑的单射。

2) 计算并证明在任何一个 $(x,y)\in\mathbb{R}^2$ 处，映射的 Jacobi 矩阵 $\mathrm{Jac}(\varphi)(x,y)$ 都是可逆的。

3) 证明，对任意的 $(x,y)\in\mathbb{R}^2$，总能开集 U 和 V，使得 $(x,y)\in U$，$f(x,y)\in V$，$f:U\to V$ 是光滑的微分同胚。

4) 证明，φ 的像是闭集。（提示：假设 $\{\varphi(x_n,y_n)\}_{n\geqslant1}$ 是一个收敛的点列，证明可以找到 $\{(x_n,y_n)\}_{n\geqslant1}$ 的子列也收敛）。

5) 证明，φ 是光滑的微分同胚。

6) 证明，φ^{-1} 是 Lipschitz 的映射，即存在 $C>0$，使得对任意的 $(x,y),(x',y')\in\mathbb{R}^2$，我们有

$$\left|\varphi^{-1}(x,y)-\varphi^{-1}(x',y')\right|\leqslant C\sqrt{(x-x')^2+(y-y')^2}.$$

（提示：先证明 $d\varphi^{-1}$ 是有界的）

C. 实数上 C^1 群结构的分类

对于实数集 $\mathbb{R}$，我们考虑上面的一个 C^1 的群结构 $(\mathbb{R},\star)$。按照定义，所谓的群结构指的是存在 $\mathfrak{e}\in\mathbb{R}$ 和 C^1 的映射

$$f:\mathbb{R}\times\mathbb{R}\to\mathbb{R},\quad (x,y)\mapsto x\star y,$$

满足

a) 对任意的 $x\in\mathbb{R}$，都有 $\mathfrak{e}\star x=x\star\mathfrak{e}=x$；

b) 对任意的 $x,y,z\in\mathbb{R}$，都有 $(x\star y)\star z=x\star(y\star z)$；

c) 对任意的 $x\in\mathbb{R}$，存在唯一的 $y\in\mathbb{R}$，使得 $x\star y=y\star x=\mathfrak{e}$。

我们习惯上把 $\mathfrak{e}$ 称作是单位元并且把 $c)$ 中的 y 记作是 x^{-1}（请不要和倒数搞混）。我们在 $\mathbb{R}\times\mathbb{R}$ 上用 (x_1,x_2) 作为坐标并用 $\partial_1 f$ 和 $\partial_2 f$ 来简记 f 的偏导数。另外，在 $\mathbb{R}$ 上我们所熟知的加法和乘法我们用符号 $+$ 和 $\times$ 来表示。

1) 证明，对任意的 $x,y,z\in\mathbb{R}$，我们有

$$(\partial_2 f)(x\star y,\mathfrak{e})=(\partial_2 f)(x,y)\times(\partial_2 f)(y,\mathfrak{e})$$

2) 证明，对任意的 $x\in\mathbb{R}$，$(\partial_2 f)(x,\mathfrak{e})>0$。

3) 假设 φ 是 $(\mathbb{R},\star)$ 和 $(\mathbb{R},+)$ 这两个群之间的 C^1-群同态，即

$$\varphi:\mathbb{R}\to\mathbb{R}$$

是 C^1 的并且对于任意的 $x,y\in\mathbb{R}$，我们都有

$$\varphi(x\star y)=\varphi(x)+\varphi(y).$$

证明，存在常数 $a\in\mathbb{R}$，使得

$$\varphi(x)=a\int_{\mathfrak{e}}^{x}\frac{d\tau}{(\partial_2 f)(\tau,\mathfrak{e})}.$$

4) 证明，对任意的 $a \neq 0$，映射

$$\varphi_a : \mathbb{R} \to \mathbb{R}, \quad x \mapsto a \int_{\mathfrak{e}}^{x} \frac{d\tau}{(\partial_2 f)(\tau, \mathfrak{e})},$$

从 $(\mathbb{R}, \star)$ 到 $(\mathbb{R}, +)$ 的 C^1-群同构，即 φ 是群同态并且是 C^1-的微分同胚（它的逆仍然是群同态）。

5) 证明，$\mathbb{R}$ 上任何一个 C^1 的群结构 $(\mathbb{R}, \star)$ 都是交换的，即对任意的 $x, y \in \mathbb{R}$，$x \star y = y \star x$。

6) 证明，映射

$$f : \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}^2, \quad \big((x_1, y_1), (x_2, y_2)\big) \mapsto x \star y = (x_1 + x_2, y_1 + e^{x_1} y_2),$$

是 $\mathbb{R}^2$ 上的 C^1 的群结构 $(\mathbb{R}, \star)$，它并不交换。

D. 一个矩阵问题的研究

假设 $\mathbf{M}_n$ 为 $n \times n$ 的实系数矩阵的全体。对于 $A \in \mathbf{M}_n$，我们定义

$$\|A\| = \sup_{x \in \mathbb{R}^n, x \neq 0} \frac{|A \cdot x|}{|x|}.$$

其中 $|x| = \sqrt{|x_1|^2 + \cdots + |x_n|^2}$。假设 $\lambda_1, \cdots, \lambda_n$ 是 A 的特征值。

D1) 试证明，$\|A\| = \sup_{k=1,\cdots,n} |\lambda_k|$。进一步证明这是一个范数并且和之前定义的 $\|\cdot\|_1$，$\|\cdot\|_2$ 和 $\|\cdot\|_\infty$ 都等价。

D2) 令 $\mathbf{GL}_n(\mathbb{R})$ 是 $\mathbf{M}_n$ 中可逆矩阵的全体。证明，$\mathbf{GL}_n(\mathbb{R})$ 是 $\mathbf{M}_n$（这是一个 $\mathbb{R}^{n^2}$）中的开集并且映射

$$\mathrm{inv} : \mathbf{GL}_n(\mathbb{R}) \to \mathbf{GL}_n(\mathbb{R}), \quad A \mapsto A^{-1}$$

是 C^1 的映射。进步一证明，对于 $X \in \mathbf{M}_n$，

$$d\,\mathrm{inv}(A)(X) = -A^{-1} \cdot X \cdot A^{-1},$$

（你可以将 $d\,\mathrm{inv}(A)$ 想象成是 $n^2 \times n^2$ 的矩阵，为什么）

D3) 我们定义立方运算 $\mathcal{C} : \mathbf{M}_n \to \mathbf{M}_n$，$A \mapsto A^3$。证明 $\mathcal{C}$ 是连续可微的并计算它的微分 $d\mathcal{C}(A)$。将 $d\mathcal{C}(A)$ 视为是 $n^2 \times n^2$ 的矩阵，进一步证明

$$\|d\mathcal{C}(A) - \mathbf{I}_{n^2 \times n^2}\| \leqslant 6\|A - \mathbf{I}_{n \times n}\| + 3\|A - \mathbf{I}_{n \times n}\|^2.$$

D4) 令 $B_{\frac{1}{3}}(\mathbf{I}_{n\times n})=\{X\in\mathbf{M}_n|\|X-\mathbf{I}_{n\times n}\|<\dfrac{1}{3}\}$。证明，$\mathcal{C}\big(B_{\frac{1}{3}}(\mathbf{I}_{n\times n})\big)$ 是 $\mathbf{M}_n$ 中的开集并且

$$\mathcal{C}:B_{\frac{1}{3}}(\mathbf{I}_{n\times n})\to\mathcal{C}\big(B_{\frac{1}{3}}(\mathbf{I}_{n\times n})\big)$$

是微分同胚。(提示:为了证明 $\mathcal{C}$ 是单射,可以考虑 $\mathcal{D}(A)=\mathcal{C}(A)-3A$ 并证明 $\|\mathcal{D}(A)-\mathcal{D}(B)\|\leqslant\dfrac{7}{3}\|A-B\|$)

D5) 我们定义平方运算 $\Theta:\mathbf{M}_n\to\mathbf{M}_n,\ A\mapsto A^2$。证明 Θ 是连续可微的并计算它的微分 $d\mathcal{C}(A)$。据此证明，存在 $\varepsilon>0$，当 $\|B-\mathbf{I}_{n\times n}\|<\varepsilon$ 时，存在 $A\in\mathbf{M}_n$，使得 $A^2=B$。

D6) 假设 $n=2$，$U=\left(\begin{smallmatrix}-1&0\\0&1\end{smallmatrix}\right)$，是否存在 $\mathbf{I}_{2\times2}$ 的邻域 $\mathcal{I}\subset\mathbf{M}_2$，$U$ 的邻域 $\mathcal{U}\subset\mathbf{M}_2$ 以及连续可微的映射 $\Psi:\mathcal{I}\to\mathcal{U}$，使得 $\Psi(\mathbf{I}_{2\times2})=U$ 并且对任意的 $A\in\mathcal{I}$，都有 $\Psi(A)^2=A$?（提示：考虑 $H=\left(\begin{smallmatrix}0&1\\1&1\end{smallmatrix}\right)$ 并计算 $d\Theta(U)H$）

E. 利用反函数定理来证明一个和隐函数定理相关的命题

令 $E=\mathbb{R}^m$，$F=\mathbb{R}^n$，其中 $m,n\geqslant1$。我们考虑如下的连续可微的映射

$$f:F\times E\to E,(\lambda,x)\mapsto f(\lambda,x).$$

我们假设存在非负常数 $k<1$，使得对任意的 $(\lambda,x)\in F\times E$，我们都有 $\|d_2f(\lambda,x)\|<k$，其中，我们认为 λ 是固定的，d_2 代表对后面 E 的变量求微分。

E1) 利用 Banach 不动点定理证明，对任意的 $\lambda\in F$，存在唯一的 $x_\lambda\in E$，使得 $f(\lambda,x_\lambda)=x_\lambda$。这定义出映射:

$$\varphi:F\to E,\lambda\mapsto x_\lambda.$$

E2) 选定 $\lambda_0\in F$，令 $x_0=\varphi(\lambda_0)$，证明，存在 λ_0 在 F 中的邻域 U 和 C^1 的映射 $\psi:U\to E$，使得对任意的 $\lambda\in U$，我们都有 $f(\lambda,\psi(\lambda))=\psi(\lambda)$。（提示：考虑映射 $g:F\times E\to E,(\lambda,x)\mapsto x-f(\lambda,x)$。）

E4) 证明，第一步定义的 φ 是 C^1 的并计算它的微分。

E5) 证明，对任意的 $t\in\mathbb{R}$，存在唯一的 C^1 映射 $t\mapsto(x(t),y(t))\in\mathbb{R}^2$，使得

$$\begin{cases}x&=\frac{1}{2}\sin(x+y)+t-1,\\y&=\frac{1}{2}\cos(x-y)-t+\frac{1}{2}.\end{cases}$$

进一步证明 $(x(t),y(t))=(0,0)$ 当且仅当 $t=1$ 并计算 $x'(1)$ 和 $y'(1)$。

Technical skill is mastery of complexity while creativity is mastery of simplicity.

E.C. Zeeman, *Catastrophe Theory, 1977.*

34 隐函数定理，隐函数定理的子流形叙述，子流形的判定：光滑曲线，光滑曲面超曲面的例子，子流形的参数化表示，Möbius 带的研究

二零二零年三月二日，星期一，晴

隐函数定理

我们假设 $\Omega \subset \mathbb{R}^{n+p}$ 是开集。为了定理的叙述方便，我们将 $\mathbb{R}^{n+p}$ 写成乘积结构 $\mathbb{R}^n \times \mathbb{R}^p$。在 $\mathbb{R}^n$ 上面我们用坐标 $x = (x_1, \cdots, x_n)$，在 $\mathbb{R}^p$ 上面我们用坐标 $y = (y_1, \cdots, y_p)$。对于 Ω 上的映射／函数 f，我们用 $f(x, y)$ 这样的二元函数来表示，其中 $x \in \mathbb{R}^n$，$y \in \mathbb{R}^p$。为了行文方便，我们引入下面（不标准的）符号：将 y 固定，我们就可以将 $f(x, y)$ 看作是 x 的函数，即

$$f(\cdot, y) : \mathbb{R}^n \to \mathbb{R}^m,$$

从而我们可以对 x 的微分，我们将这个微分记作 $d_x f(x, y)$；类似地，我们可以定义 $d_y f(x, y)$。

我们首先给出隐函数定理的经典叙述（和证明），然后用子流形的语言就可以有更清晰的几何表述。

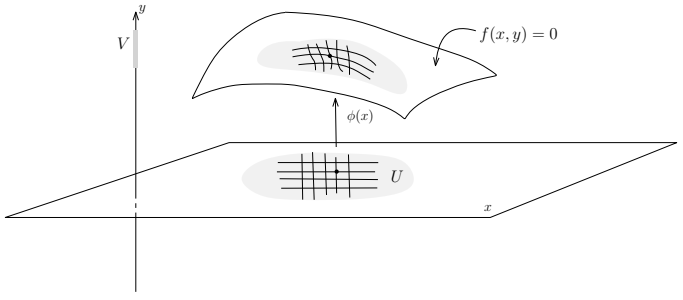
定理 198 (隐函数定理)**.** *给定正整数 n 和 p 和非空开集 $\Omega \subset \mathbb{R}^n \times \mathbb{R}^p$，映射 $f : \Omega \to \mathbb{R}^p$ 是连续可微的（C^1 的）。*

假设 $(x^*, y^*) \in \Omega$，使得 $f(x^*, y^*) = 0$，其中 $x^* \in \mathbb{R}^n$，$y^* \in \mathbb{R}^p$，$0 \in \mathbb{R}^p$。如果 $\big(d_y f\big)(x^*, y^*)$ 是可逆的，即 Jacobi 矩阵 $\left(\dfrac{\partial f_i}{\partial y_j}\right)_{1 \leqslant i,j \leqslant p}$ 在 (x^*, y^*) 点是可逆的。那么，存在开集 $U \subset \mathbb{R}^n$，$x^* \in U$，存在开集 $V \subset \mathbb{R}^p$，$y^* \in V$，存在连续可微的映射 $\phi : U \to V$，使得对任意的点 $(x, y) \in \Omega$，如下两条是等价的：

- $(x, y) \in U \times V$ *并且* $f(x, y) = 0$；
- $x \in U$ *并且* $y = \phi(x)$。

我们进一步还有 $d\phi(x) = -(d_y f(x, y))^{-1} \circ d_x f(x, y)$。

注记. *”隐”函数说的是方程 $f(x, y) = 0$ 实际上隐含地将 y 定义成 x 的函数，定理具体地将这个函数构造了出来：$y = \phi(x)$。当 $p = 1$ 时，$f(x, y) = 0$ 在空间中所定义的曲面局部上就可以表达成 $y = \phi(x)$ 的形式，也就是说它局部上可以看成是函数的图像，从而是 $\mathbb{R}^{n+1}$ 中余 1 维的子流形，下面的图很好的总结了这个情形：*



例子. 隐函数定理的叙述乍看起来晦涩难懂。最简单的就是如下的情形：$\Omega = \mathbb{R}^n \times \mathbb{R}^p$，$f$ 是投影映射：

$$f : \mathbb{R}^n \times \mathbb{R}^p \to \mathbb{R}^p, \quad (x_1, \cdots, x_n, y_1, \cdots, y_p) \mapsto (y_1, \cdots, y_p).$$

此时，$U = \mathbb{R}^n$，$V = \mathbb{R}^p$ 而

$$\phi : \mathbb{R}^n \to \mathbb{R}^p, \quad x \mapsto 0.$$

这显然满足定理的要求，因为 $f(x, y) = 0 \Leftrightarrow y = 0$。隐函数定理本质上说只要选取适当的坐标系这是唯一的例子。

证明: 我们利用反函数定理来证明隐函数定理。为此，我们通过把 f 的值域空间加一些维数来定义一个新的映射。注意到，对于 $x \in \mathbb{R}^n$, $f(x, y) \in \mathbb{R}^p$（如果 (x, y) 落在 f 的定义域里的话）。所以，以下的映射是良好定义的：

$$F : \Omega \to \mathbb{R}^n \times \mathbb{R}^p, \ (x, y) \mapsto F(x, y) = (x, f(x, y)).$$

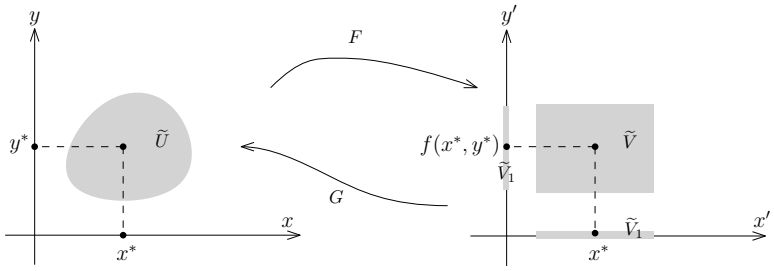
很明显，F 是连续可微的（因为它的每个分量都是连续可微的）。我们可以计算 F 在 (x, y) 处微分，用 Jacobi 矩阵来显然可以表达成：

$$dF = \begin{pmatrix} \mathrm{id} & 0 \\ d_x f & d_y f \end{pmatrix}, \tag{1}$$

这是一个 $(n + p) \times (n + p)$ 的方阵。由于 $(d_y f)(x^*, y^*)$ 可逆，所以分块矩阵 $(dF)(x^*, y^*)$ 也是可逆的。所以，我们可以对映射 F 在 (x^*, y^*) 处用反函数定理：存在开集 $\widetilde{U} \subset \Omega$，$(x^*, y^*) \in \widetilde{U}$，存在开集 $\widetilde{V} \subset \mathbb{R}^{n+p}$ 并且通过将它缩小我们可以假设它具有乘积结构 $\widetilde{V} = \widetilde{V_1} \times \widetilde{V_2}$ 以及连续可微的映射 $G : \widetilde{V} \to \widetilde{U}$（这是 F 的逆），使得

$$G \circ F = \mathrm{id}_{\widetilde{U}}, \quad F \circ G = \mathrm{id}_{\widetilde{V}}.$$

我们可以参考下面的示意图：



我们将 F 的值域（即 G 的定义域）上的坐标用 (x', y') 来表示。按照分量的记法，可以将 G 写成

$$G(x', y') = (\overbrace{G_1(x', y')}^{\in \mathbb{R}^n}, \overbrace{G_2(x', y')}^{\in \mathbb{R}^p}).$$

我们把反函数定理给出的两个等式用分量写出来，就有

$$\begin{aligned} G\circ F=\mathrm{id}_{\widetilde{U}} \quad &\Rightarrow \quad G_1(x,f(x,y))=x,\ G_2(x,f(x,y))=y;\\ F\circ G=\mathrm{id}_{\widetilde{V}} \quad &\Rightarrow \quad G_1(x',y')=x',\ f(x',G_2(x',y'))=y'. \end{aligned}$$

总结一下，我们得到了

$$\begin{cases} G(x',y')=(x',G_2(x',y')),\\ G_2(x,f(x,y))=y,\\ f(x',G_2(x',y'))=y'. \end{cases}$$

按照 F 和 G 的定义，对于给定的 $(x,y)\in\widetilde{U}$，我们有如下的等价关系

$$f(x,y)=0 \quad\Leftrightarrow\quad F(x,y)=(x,0) \quad\Leftrightarrow\quad (x,y)=F^{-1}(x,0)=G(x,0)=(x,G_2(x,0)).$$

所以，如果我们令 $\phi(x)=G_2(x,0):V\to\mathbb{R}^p$，那么，给定的 $(x,y)\in\widetilde{U}$，我们有如下的等价关系

$$f(x,y)=0 \quad\Leftrightarrow\quad y=\phi(x).$$

所以，我们可以取 $U=\widetilde{U}$，$V=\widetilde{V_1}$，这就证明了隐函数定理的主要论述。

最终，为了计算微分，我们对 $f(x,\phi(x))=0$ 求微分，即对映射

$$U\to\mathbb{R}^p,\ \ x\mapsto f(x,\phi(x))$$

求微分。我们得到

$$df_x(x,y)+df_y(x,y)\cdot d\phi(x)=0,\ \ \text{其中 } y=\phi(x).$$

在 (x^*,y^*) 的附近，$df_y(x,y)$ 是可逆的，所以 $d\phi(x)=-(d_yf(x,y))^{-1}\circ d_xf(x,y)$。 □

注记. 如果我们把反函数定理条件中映射改为 C^∞ 的，那么结论中的函数和映射也是 C^∞，这从反函数定理的叙述就可以看出来。

隐函数定理的子流形叙述

利用子流形这个概念，我们可以更形象地理解隐函数定理。反过来，隐函数定理给出了一种非常实用的判断子流形的方法。

假设 $M\subset\mathbb{R}^N$ 是一个 d-维的子流形，所以对任意的 $x\in M$，存在开集 $U\subset\mathbb{R}^N$ 包含 x，$\mathbb{R}^n$ 中的开集 V 以及微分同胚 $\varphi:U\to V$ 使得

$$\varphi:U\cap M=V\cap(\mathbb{R}^d\times\{0\}),$$

我们现在沿用隐函数定理中的符号。我们令

$$M=\big\{(x,y)\in\mathbb{R}^n\times\mathbb{R}^p\big|f(x,y)=0\big\}.$$

由于 $f(x,y)=0$ 实际上是 p 个方程（以为 $f:\mathbb{R}^{n+p}\to\mathbb{R}^p$），所以，我们将上述集合是 $\mathbb{R}^{n+p}$ 中由 p 个方程的零点集合。如果 f 的某些导数为非退化的（参考隐函数定理的叙述），那么局部上这个 M 可以表达成映射的图像的形式，即对任意的 $(x,y)\in M$，局部上都等价于 $y=\phi(x)$。从另外一种观点来看，M 可以被 x 参数化，即给定 x，就唯一地确定了 M 上的一个点。

为了说明 M 实际上是微分子流形，我们需要构造一个映射 $\Phi:\mathbb{R}^{n+p}\to\mathbb{R}^{n+p}$，使得 φ 是微分同胚（局部上）并且将 M 映射成右边 $\mathbb{R}^{n+p}$ 上的由 $y_1=\cdots=y_p=0$ 所给出来的线性子空间。这个映射就是定理证明中的 F：

$$F(x,y)=(x,f(x,y))=(x,0),\quad \text{其中}\quad (x,y)\in M.$$

这样，我们就证明了 M 实际上（局部上）是一个微分子流形。

我们把上面的讨论总结为如下的定理：

定理 199 (隐函数：子流形版本)**.** *给定正整数 n' 和 p（$n=n'+p$）和非空开集 $\Omega\subset\mathbb{R}^{n'+p}=\mathbb{R}^n$，映射 $f:\Omega\to\mathbb{R}^p$ 是光滑的。令*

$$M=\big\{(x,y)\in\mathbb{R}^n\times\mathbb{R}^p\big|f(x,y)=0\big\}.$$

假设 $(x^*,y^*)\in M$ 使得 Jacobi 矩阵 $\Big(\dfrac{\partial f_i}{\partial y_j}\Big)_{1\leqslant i,j\leqslant p}$ 在 (x^*,y^*) 点是可逆的。那么，存在开集 $U\subset\mathbb{R}^n$ 使得 $U\cap M$ 是 $\mathbb{R}^n$ 中的 n' 维子流形。

这个版本有如下重要的推论：

推论 200 (子流形的判断准则)**.** *假设 $f:\Omega\to\mathbb{R}^p$ 是光滑映射，其中 $\Omega\subset\mathbb{R}^n$ 是开集，$n\geqslant p$。对于 $c\in\mathbb{R}^p$，我们把 c 上**纤维**定义为的 c 在 f 下的逆像：*

$$f^{-1}(c)=\{x\in\mathbb{R}^{n+p}|f(x)=c\}.$$

如果对任意的 $x\in f^{-1}(c)$，$\mathrm{Rank}df(x)=p$。那么，纤维 $f^{-1}(c)$ 是余维数为 p 的子流形。

证明： 我们要在 $\mathbb{R}^n$ 上的选取坐标系 $x_1,\cdots,x_n$。根据 Rank $df(x)=p$，通过调整 $x_1,\cdots,x_n$ 的标号，我们可以使得后面 p 个坐标 $x_{n-p+1},\cdots,x_n$ 满足

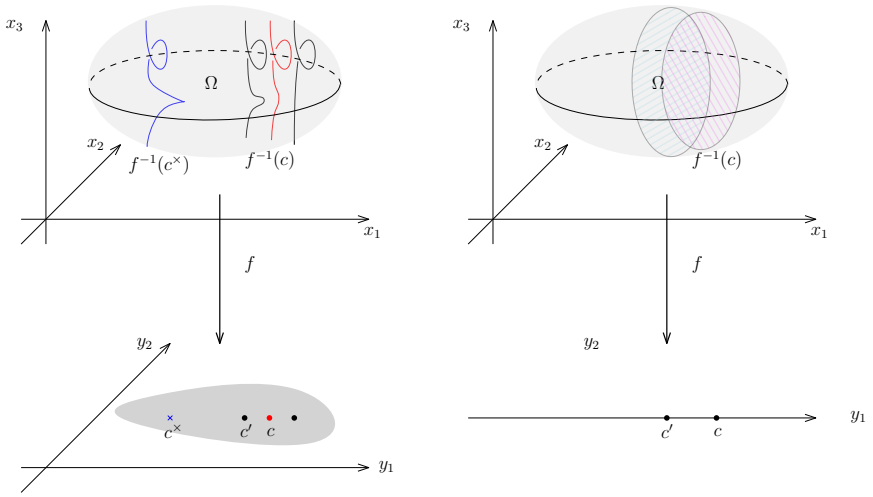
$$\Big(\frac{\partial f_i}{\partial x_{n+1-j}}\Big)_{1\leqslant i,j\leqslant p}(x)\neq 0,$$

其中 x 是任意选定的 $x\in f^{-c}$ 上的点。我们选取函数 $f-c$ 并且将后 p 个坐标用 $y_1,\cdots,y_p$ 来表示，这就化成了隐函数定理的情况。特别地，纤维 f^{-c} 在任意一点附近都是 p 维子流形，所以 $f^{-1}(c)$ 是 p-维子流形。 □

注记. *我们注意到定理本身的叙述不需要任何的坐标系，而证明本身很好地解释了不依赖于坐标和选取坐标系之间的关系。*

我们应该和线性代数的情况做类比：给定 n 维线性空间上的 p 个齐次线性方程，它们的公共零点是 $n-p$ 的线性子空间当且仅当这 p 个线性方程的秩是 p。

注记 (几何图像). 我们可以用下图形象地记忆并表示这个命题，这是关于隐函数定理最干净最形象的表述：



左图代表是余维数为 2 的情形。我们假设 $f:\Omega\to f(\Omega)$，其中 $f(\Omega)$ 是下面灰色的区域。对于每个 $c\in\Omega$，我们可以把它的原像想象成插在这个点上面的一个纤维。特别地，Ω 可以被写成这些纤维的无交并：

$$\Omega=\coprod_{c\in f(\Omega)} f^{-1}(c).$$

然而，我们需要在每个 $x\in f^{-1}(c)$ 点处要求一个非退化的条件才能保证 $f^{-1}(c)$ 是光滑的子流形。上面的图形表示，在 c 的纤维如果满足这个非退化的条件，那么它附近的 c' 处的纤维 $f^{-1}(c)$ 也是光滑的子流形：实际上，我们需要对 $f^{-1}(c)$ 限制才可以（比如纤维是紧的），但是很多情况下这一点都成立，大概的原因是 Rank $df(x)=p$ 如果成立，就在附近的一个开集上面成立。然而，离着 c 比较远的点，这个条件就可能退化，比如说上面的蓝色 $c^\times$ 点，它的纤维就“打折”了。

左图代表是余维数为 1 的情形，对于每个 c，我们形象的认为纤维是 Ω 的一个“切片”。事实上，为了研究高维子流形的结构，我们经常通过这种切片化成低维数子流形的情况来研究。在后面的课程和作业中，我们会用这个方法研究几个重要的例子。

下面的命题是隐函数定理几何版本的逆命题：子流形总是可以（在局部上）由 $\mathrm{codim}M=n-\dim M$ 个方程的零点来定义。我们已经在第四次课的开始证明了这个结论，为了完备起见（也为了再次复习子流形的定义），我们概述一下证明。

命题 201. 假设 $M^d\subset\mathbb{R}^n$ 是 d 维的子流形。那么，对任意的 $p\in M$，存在开集 $U\subset\mathbb{R}^n$，$p\in U$ 和 $n-d$ 个光滑函数 $f_1,\cdots,f_{n-d}\in C^\infty(U)$，使得

$$M^d=\{x\in U\big|f_1(x)=\cdots=f_{n-d}(x)=0\}.$$

特别地，映射的微分

$$f:U\to\mathbb{R}^{n-d},\ x\mapsto(f_1(x),\cdots,f_{n-d}(d))$$

在每个点 $x\in M$ 处的秩都是 $n-d$。

证明: 按子流形的定义，存在开集 $U \subset \mathbb{R}^n$ 包含 p，$\mathbb{R}^n$ 中的开集 V 以及微分同胚 $\varphi : U \to V$，使得

$$\varphi : U \cap M = V \cap (\mathbb{R}^d \times \{0\}).$$

我们用 $y_1, \cdots, y_n$ 表示 V 上的坐标函数，那么我们选取 $f_i = \varphi^* y_i = y_i \circ \varphi$ 即可，其中 $i = d+1, \cdots, n$。 □

注记. 我们沿用命题中的记号，如果令

$$M^{d+k} = \{x \in U \big| f_1(x) = \cdots = f_{n-d-k}(x) = 0\},$$

那么，M^{d+k} 是 $d+k$ 维的子流形（作业中会证明）。换句话说，每次加上一个 $f_j(x) = 0$ 的限制，子流形的维数恰好降低 1 维。

我们现在给出隐函数定理的几个应用，首先是用来判定子流形，现在我们只需要做一些简单的代数计算即可:

1) $\mathbb{R}^n$ 中的球面 $\mathbf{S}^{n-1} = \{x \in \mathbb{R}^n | x_1^2 + \cdots + x_n^2 = 1\}$ 是一个 $n-1$ 维的子流形。

   球面可以被视作是函数

   $$f(x_1, \cdots, x_n) = (x_1)^2 + (x_2)^2 + \cdots + (x_n)^2 - 1$$

   的零点集。所以，只要验证它的微分 $df(x)$ 的秩为 1 即可。然而，用矩阵来写

   $$\begin{aligned} df(x) &= \left(\frac{\partial f}{\partial x_1}(x), \cdots, \frac{\partial f}{\partial x_n}(x) \right) \\ &= 2(x_1, \cdots, x_n). \end{aligned}$$

   由于在 $\mathbf{S}^{n-1}$ 上，$|x| = 1$，上面的向量显然非零，所以这是光滑子流形。这比之前我们把球面 $\mathbf{S}^2$ 拆成六块（参见第三课讲义）并且将每块写成函数的图像要方便很多。

2) 我们考虑 $\mathbb{R}^4 = \mathbb{C}^2$ 中的曲面

   $$\mathbf{T}^2 = \Big\{ (z_1, z_2) \big| |z_1| = 1, |z_2| = 1 \Big\}.$$

   如果我们用 (x, y, z, w) 作为坐标，那么这个曲面实际上是

   $$f(x, y, z, w) = x^2 + y^2 - 1, \quad g(x, y, z, w) = z^2 + w^2 - 1$$

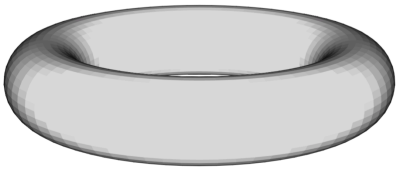
   这两个方程的公共零点集。我们需要计算

   $$\mathbb{R}^4 \to \mathbb{R}^2, \quad (x, y, z, w) \mapsto \big(f(x, y, z, w), g(x, y, z, w) \big),$$

   的 Jacobi 矩阵并判断它的秩。这个矩阵显然是

   $$\begin{pmatrix} 2x & 2y & 0 & 0 \\ 0 & 0 & 2z & 2w \end{pmatrix}.$$

   由于在 $\mathbf{T}^2$ 上面，$x^2 + y^2 = 1$，$z^2 + w^2 = 1$，所以上面这个矩阵的秩是 2，从而 $\mathbf{T}^2$ 是 2 维的子流形。它实际上是一个环面（甜甜圈）:



$\mathbb{R}^3$ 中由方程定义的曲线和曲面

在经典的微积分课程中，$\mathbb{R}^3$ 中由一个方程或者两个方程定义的曲线和曲面是最基本的几何对象。我们首先引入如下的定义：

定义 202. 假设 $M \subset \mathbb{R}^n$ 是子流形。如果 $\dim M = 1$，我们就称之为 $\mathbb{R}^n$ 中的一条**光滑曲线**；如果 $\dim M = 2$，我们就称之为 $\mathbb{R}^n$ 中的一个**光滑曲面**；如果 $\dim M = n-1$，我们就称之为 $\mathbb{R}^n$ 中的一个**光滑超曲面**。

注记. 这里曲线的概念和上个学期所讲的曲线略有不同：之前是所谓的参数化曲线，即曲线是一个映射 $\gamma : I \to \mathbb{R}^n$，其中 $I \subset \mathbb{R}$ 是一个区间。这里的曲线只是 $\mathbb{R}^n$ 中的一个子集，然而根据隐函数定理的结论，我们在局部上可以给它一个参数化。

我们关心的是如下的对象，他们有着更特殊的要求：

引理 203. 假设 $f : \mathbb{R}^n \to \mathbb{R}$ 是光滑函数，如果对于任意的 $x_0 \in f^{-1}(0)$，$df(x_0) \neq 0$，那么 $f^{-1}(0)$ 是超曲面。我们把这种超曲面称为由一个方程整体定义的光滑超曲面。

我们把（显然的）证明留作作业。特别地，在 $\mathbb{R}^3$ 中，我们有

引理 204. 假设 $f(x, y, z)$ 是 $\mathbb{R}^3$ 上的光滑函数，如果对于任意满足 $f(x_0, y_0, z_0) = 0$ 的点 (x_0, y_0, z_0)，我们有

$$\left(\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}, \frac{\partial f}{\partial z}\right)(x_0, y_0, z_0) \neq 0,$$

那么

$$\left\{(x, y, z) \in \mathbb{R}^3 \middle| f(x, y, z) = 0\right\}$$

是光滑曲面。我们把这种超曲面称为由一个方程整体定义的光滑曲面。

例子. $\mathbb{R}^3$ 中的柱面由如下方程

$$x^2 + y^2 = 1$$

定义，这是一个光滑曲面：我们取 $f(x, y, z) = x^2 + y^2 - 1$，此时，

$$\left(\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}, \frac{\partial f}{\partial z}\right) = (2x, 2y, 0) \neq 0,$$

我们还有用两个方程定义的曲线：

引理 205. 假设 $f(x,y,z)$ 和 $g(x,y,z)$ 是 $\mathbb{R}^3$ 上的光滑函数，如果对于任意同时满足 $f(x_0,y_0,z_0)=0$ 和 $g(x_0,y_0,z_0)=0$ 的点 (x_0,y_0,z_0)，如果如下两个向量线性不相关：

$$\left(\frac{\partial f}{\partial x},\frac{\partial f}{\partial y},\frac{\partial f}{\partial z}\right)(x_0,y_0,z_0),\quad \left(\frac{\partial g}{\partial x},\frac{\partial g}{\partial y},\frac{\partial g}{\partial z}\right)(x_0,y_0,z_0),$$

那么

$$\left\{(x,y,z)\in\mathbb{R}^3\middle| f(x,y,z)=0, g(x,y,z)=0\right\}$$

是光滑曲线。我们把这种曲线称为由两个方程整体定义的光滑曲线。

我们把（显然的）证明留作作业。

在结束关于隐函数定理的几何讨论之前，我们给出隐函数定理的一个对偶版本。这个版本本身和隐函数是无关的，之所以是对偶的，是因为赢函数定理研究从 $\mathbb{R}^n$ 到 $\mathbb{R}^p$ 的映射，其中 $n\geqslant p$，其结论说映射的逆项或者纤维是子流形。而这个定理（证明非常简单，只要验证定义）研究从 $\mathbb{R}^n$ 到 $\mathbb{R}^{p'}$ 的映射，其中 $n\leqslant p$，其结论说映射的像是子流形。

定理 206 (参数化子流形)**.** $\Omega\subset\mathbb{R}^n$ 是开集，$f:\Omega\to\mathbb{R}^{n+p}$ 是光滑映射。假设对于点 $x^*\in\Omega$，有 $\operatorname{rank} df(x^*)=n$，那么，存在开集 $U\subset\mathbb{R}^n$，$x^*\in U$，使得 $f(U)\subset\mathbb{R}^{n+p}$ 是 n 维的子流形。

证明: 在 Ω 上选取坐标系统 $x_1,\cdots,x_n$，在 $\mathbb{R}^{n+p}$ 上选取坐标坐标系统 $y_1,\cdots,y_n,z_1,\cdots,z_p$。我们把 f 用坐标分量写成

$$f(x)=(f_1(x),\cdots,f_{n+p}(x)),$$

其中每个 $f_i(x)$ 都是 Ω 上的光滑函数。由于有 $\operatorname{rank} df(x^*)=n$，通过调换 $\{x_i\}$ 坐标的指标，我们不妨假设

$$\det\left(\frac{\partial f_i}{\partial x_j}\right)_{i,j\leqslant n}(x^*)\neq 0.$$

我们现在考虑 $F=\pi_y\circ f$，其中 $\pi:\mathbb{R}^{n+p}\to\mathbb{R}^n$ 是取前 n 个坐标，即

$$\begin{array}{ccc}\Omega & \xrightarrow{f} & \mathbb{R}^{n+p}\\ & \searrow^{F} & \downarrow{\scriptstyle \pi_y}\\ & & \mathbb{R}^p\end{array}$$

换句话说，我们定义了 $F(x)=(f_1(x),f_2(x),\cdots,f_n(x))$。根据上述行列式非零的条件，$F$ 为局部微分同胚（反函数定理）。令 U 为 x 在 Ω 中的开集，$F:V\to F(U)$ 为这个微分同胚并令 $V=F(U)$。此时，我们考察 $f(U)\subset\mathbb{R}^{n+p}$。对于 $x\in U$，令 $y=F(x)$，$f(x)\in\mathbb{R}^{n+1}$ 可以写成

$$\begin{aligned}(f_1(x),\cdots,f_n(x);f_{n+1}(x),\cdots,f_{n+p}(x)) &= (y;f_{n+1}(x),\cdots,f_{n+p}(x))\\ &= (y,f_{n+1}(F^{-1}(y),f_{n+2}(F^{-1}(y)),\cdots,f_{n+p}(F^{-1}(y)).\end{aligned}$$

所以，它可以被视作是 V 上定义的向量值函数 $y\mapsto f_{n+1}(F^{-1}(y),f_{n+2}(F^{-1}(y)),\cdots,f_{n+p}(F^{-1}(y))$ 的图像，这是一个用 y 做为坐标来参数化的子流形（在本次作业中证明）。 □

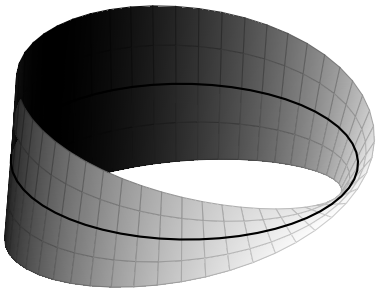
在 $\mathbb{R}^3$ 还有一个经典曲面：**Möbius 带**。我们用参数的方式可以把它写作

$$\mathcal{M}=\left\{(x,y,z)\in\mathbb{R}^3\Big|\begin{cases}x=(1+r\cos\frac{\theta}{2})\cos\theta,\\ y=(1+r\cos\frac{\theta}{2})\sin\theta,\quad \theta\in[0,2\pi],r\in(-1,1)\\ z=r\sin\frac{\theta}{2}\end{cases}\right\}.$$

适当改变一下书写方式更有启发：

$$(x,y,z)=(\cos\theta,\sin\theta,0)+(r\cos\frac{\theta}{2}\cos\theta,r\cos\frac{\theta}{2}\sin\theta,r\sin\frac{\theta}{2}).$$

当 $r=0$ 时，我们得到一个用 θ 来参数化的基本圆周 $(\cos\theta,\sin\theta,0)$。给定 ϑ，就给定了这个圆周上的一个点 $(\cos\theta,\sin\theta,0)$。固定这个 θ，当 r 在 $(-1,1)$ 之间变化的时候，我们就在这个点插上了一个小线段并且这个线段的方向的变化是 θ 的一半。所以，当 θ 从 0 变化到 2π，基本圆周上的点旅行了一周又回到了原来的点，但是它上面所插的小线段只转动了 $180°$。



我们下面证明一个不明显的命题，为此，我们先选择好的记号：把 $r=0$ 光滑圆周写成参数形式：

$$\gamma:[0,2\pi]\to\mathbb{R}^3,\quad \theta\mapsto(\cos\theta,\sin\theta,0).$$

给定 θ 处，我们用 e_0 表示 γ 在 $\gamma(\theta)$ 处的切方向 $\gamma'(\theta)$，即令 $e_0=(-\sin\theta,\cos\theta,0)$。和 e_0 垂直的方向有两个，我们分别记作

$$e_1(\theta)=(\cos\theta,\sin\theta,0),\quad e_2(\theta)=(0,0,1).$$

据此，我们把 M 的参数化写成：

$$\mathcal{M}=\left\{(x,y,z)=(\cos\theta,\sin\theta,0)+r\cos\frac{\theta}{2}e_1(\theta)+r\sin\frac{\theta}{2}e_2(\theta),\theta\in[0,2\pi],r\in(-1,1)\right\}.$$

根据刚刚证明的定理，这是 $\mathbb{R}^3$ 中的一个光滑曲面（参见本次作业）。

命题 207. 不存在光滑函数 $f\in C^\infty(\mathbb{R}^3)$，使得

$$\mathcal{M}=\big\{(x,y,z)\big|f(x,y,z)=0\big\}$$

且 $df(x,y,z)$ 在 $\mathcal{M}$ 上的每点处不为零，即 $\mathcal{M}$ 不能只用一个方程零点来**整体**定义。

注记. 我们需要运用关于由一个方程定义的光滑曲面的特殊性质（基于连续函数的介值定理）：给定 $f \in C^\infty(\mathbb{R}^3)$，假设对任意的 $x \in \Omega$，$df(x) \neq 0$，$f^{-1}(0) \subset \Omega$ 是光滑子流形并且由一个方程定义。任意选定 $x \in \Omega$，考虑 x 处的邻域 U（足够小）。由于 $df(x) \neq 0$，我们不妨选 $v \in \mathbb{R}^n$，使得

$$\nabla_v f(x) = df(x)(v) > 0.$$

从而，在当 $\varepsilon > 0$ 足够小时，$p_\pm = x \pm \varepsilon v \in U$ 并且 $f(p_+) > 0$，$f(p_-) < 0$。我们声明，任何一条连接 p_+ 和 p_- 的曲线都要经过 $f^{-1}(0)$，即对任意的连续映射 $\gamma : [0,1] \to X$，其中起点 $\gamma(0) = p_-$ 和终点 $\gamma(1) = p_+$），一定存在 $t_0 \in (0,1)$，使得 $\gamma(t_0) \in f^{-1}(0)$（等价于说 $f(\gamma(t_0)) = 0$）。实际上，$f(\gamma(t))$ 为连续函数，按照 $p_\pm$ 的选取方式，我们有

$$f(\gamma(0)) = f(p_-) < 0, \quad f(\gamma(1)) = f(p_+) > 0,$$

所以连续函数的介值定理就给出了 t_0。

我们现在证明关于 Möbius 带的命题：

证明: 我们给出证明的梗概，证明的细节留给同学们在作业中完成。如若不然，假设 f 是 $\mathcal{M}$ 的定义方程。任选 $p = (\cos\theta, \sin\theta, 0) \in \mathcal{M}$，直观上（实际上也是）这个附近的小邻域 U 被 f 的图像分成了两个部分：在一部分中，$f > 0$；在另一部分中，$f < 0$。我们选取在 p 点处和整个曲面垂直的法向量

$$\nu(\theta) = \cos\frac{\theta}{2} e_2(\theta) - \sin\frac{\theta}{2} e_1(\theta)$$

我们考虑曲线

$$\beta : [0, 2\pi] \to \mathbb{R}^3, \quad (\cos\theta, \sin\theta, 0) + \varepsilon\nu(\theta).$$

我们可以适当的选取比较小的 ε（可正可负），使得 $f(\beta(0)) > 0$。然而，$\beta(2\pi) = \beta(0)$，曲线所对应的 ε 却变换了符号，从而，$f(\beta(2\pi)) < 0$。很明显，这条曲线与曲面 $\mathcal{M}$ 不相交，矛盾。 □

35 子流形的原像定理，切空间的定义，超曲面的切空间与法向量

二零二零年三月五日，星期四，晴

上次课我们证明了隐函数定理并给出了它的几何叙述。

Equations are just the boring part of mathematics. I attempt to see things in terms of geometry.

Stephen Hawking.

我们已经证明如果 $f:\mathbb{R}^n\to\mathbb{R}^m$ 满足正确的非退化条件，那么反函数定理等价于说对于 $0\in\mathbb{R}^m$，它的逆像 $f^{-1}(0)$ 是 $\mathbb{R}^n$ 中的子流形，其中，$\operatorname{codim} f^{-1}(0)=m$。我们注意到，$0\in\mathbb{R}^m$ 是 0-维的子流形，它在 $\mathbb{R}^n$ 中的余维数是也是 m。如果 $S\subset\mathbb{R}^m$ 是子流形，我们同样地可以考虑它的逆像

$$f^{-1}(S)=\big\{x\in\mathbb{R}^n\big|f(x)\in S\big\}.$$

我们有如下漂亮的定理：

命题 208 (子流形的原像)**.** *给定开集 $\Omega\subset\mathbb{R}^n$ 和光滑映射 $f:\Omega\to\mathbb{R}^m$，$S\subset\mathbb{R}^m$ 是子流形。如果对任意的 $x\in f^{-1}(S)$，$\operatorname{rank} df(x)=m$，那么 $f^{-1}(S)\subset\Omega$ 是余维数为 $\operatorname{codim} S$ 的子流形。*

证明: 假设 $\dim S=s$。利用隐函数定理，这个命题的证明非常简单：任取 $y\in S$，S 在 y 的附近可以被 $n-s$ 个函数定义：存在 $V\subset\mathbb{R}^m$ 以及 V 上的 $m-s$ 个光滑函数 $g_1,\cdots,g_{m-s}$，使得

$$\begin{cases}1)\quad S\cap V=g_1^{-1}(0)\cap\cdots\cap g_s^{-1}(0);\\ 2)\quad \text{对任意 } y'\in S\cap V,\ dg_1(y'),\cdots,dg_{m-s}(y') \text{ 线性无关.}\end{cases}$$

现在对于任意的 $x\in f^{-1}(S)$，假设 $y=f(x)\in S$，我们要证明 $f^{-1}(S)$ 在 x 附近是一个子流形。我们可以选取包含 x 的开集 $U=f^{-1}(V)$，利用上面对 S 的描述，我们也可以用方程的零点来描述 $f^{-1}(S)$：

$$\begin{aligned}f^{-1}(S)\cap U&=\big\{x\in\Omega\big|f(x)\in S\cap V\big\}\\&=\big\{x\in\mathbb{R}^n\big|g_1(f(x)),\cdots,g_{m-s}(f(x))=0\big\}\\&=\big(g_1\circ f\big)^{-1}(0)\cap\cdots\cap\big(g_{m-s}\circ f\big)^{-1}(0).\end{aligned}$$

所以 $f^{-1}(S)\cap U$ 是函数 $f^*g_1,\cdots,f^*g_{m-s}$ 的公共零点集。为了说明这是子流形，只要验证这些 $f^*g_1,\cdots,f^*g_{m-s}$ 的微分在 x 处秩是 $m-s$ 即可。这就是链式法则的简单应用：我们考虑如下两个映射的复合：

$$\begin{array}{ccc}U & \xrightarrow{\ f\ } & V\\ & {\scriptstyle f^*G}\searrow & \downarrow{\scriptstyle G}\\ & & \mathbb{R}^{m-s}\end{array},$$

其中 $G: V \to \mathbb{R}^m$ 就是函数映射 $G(y) = (g_1(y), \cdots, g_{m-s}(y))$。所以，要求的微分的秩就是映射

$$d(f^*G)(x) = df(x) \circ dG(y)$$

的秩。根据 rank $df(x) = m$ 以及 rank $dG(x) = m - s$，所以 rank $d(f^*G)(x) = m - s$。命题得到证明。 □

隐函数定理的其它经典应用举例

最著名的例子是考虑多项式的解对多项式系数的光滑依赖性。给定 $c = (c_1, \cdots, c_n) \in \mathbb{R}^n$，考虑实系数多项式 $P_c(X) = X^n + c_1 X^{n-1} + \cdots + c_n = 0$。我们假设 $c = (c_1^*, \cdots, c_n^*) \in \mathbb{R}^n$，使得 $P_{c^*}(X)$ 恰好有 n 个两两不同的实根 $z_1(c_1^*, \cdots, c_n^*) < \cdots < z_n(c_1^*, \cdots, c_n^*)$。那么，存在 $(c_1^*, \cdots, c_n^*) \in \mathbb{R}^n$ 的开邻域 Ω 和 Ω 上的光滑函数 $z_1, \cdots, z_n$，使得对任意的 $(c_1, \cdots, c_n) \in \Omega$，我们有

$$z_1(c_1, \cdots, c_n) < z_2(c_1, \cdots, c_n) < \cdots < z_n(c_1, \cdots, c_n)$$

并且

$$P(z_1(c_1, \cdots, c_n), z_2(c_1, \cdots, c_n), \cdots, z_n(c_1, \cdots, c_n)) = 0.$$

类似地，如果一个实数系数的矩阵的特征值都是实数且两两不同，那么，这些特征值也是光滑地依赖于矩阵的系数的。

请参考本周的习题。

子流形的切空间

我们一下总是假定 $M^d \subset \mathbb{R}^n$ 是一个 d 维光滑子流形。我们要发展子流形上的微分学，这是我们对多元微分学的自然推广，在实际中有着重要的应用。为此，首先定义最重要的一个几何对象：切空间。我们应该做如下的类比：所谓的微分，是映射在一点处的线性逼近，反函数／隐函数定理讲的是映射在一点处的线性逼近决定了映射自身的若干性质。我们应该把切空间视作是流形在一点处用“线性”子流形进行逼近。

首先，我们要考虑的是通过一点 $p \in M$ 的**参数化的 C^∞-曲线**（可以要求是 C^1 的），即

$$\gamma : (-\varepsilon, \varepsilon) \to \mathbb{R}^n, \quad t \mapsto \gamma(t), \ \varepsilon > 0, \quad \gamma(0) = p.$$

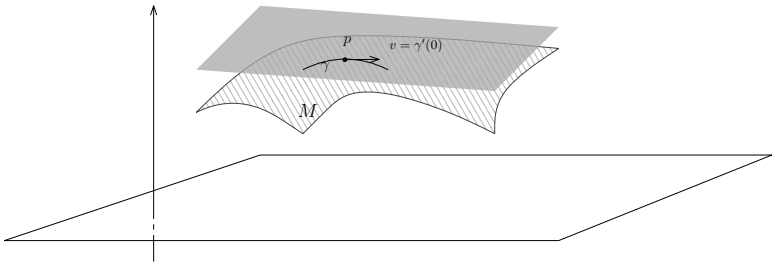
我们将

$$\gamma'(0) = \frac{d}{dt}\gamma\big|_{t=0}$$

称作该曲线在 p 处的**切向量**。如果 $\gamma\big((-\varepsilon, \varepsilon)\big) \subset M$，我们就说 γ 是M **上的参数）曲线**。所谓的 M 在 p 的切空间，就是由在 M 上的通过 p 曲线的曲线的切向量所构成的空间：

定义 209. *微分子流形在 $p \in M$ 处的***切空间** *T_pM 是 $\mathbb{R}^n$ 的子空间，它的定义如下：*

$$T_pM = \{\gamma'(0) \in \mathbb{R}^n \big| \gamma \text{ 是经过 } p \text{ 点的 } M \text{ 上的一条参数曲线}\}.$$



我们首先来看一个“线性”的例子：

例子. 假设 M 由 $x_{d+1}=x_{d+2}=\cdots=x_n=0$ 定义，即

$$M=\big\{(x_1,\cdots,x_d,0,\cdots,0)\big|x_i\in\mathbb{R},1\leqslant i\leqslant d\big\}.$$

那么，对于 $p=(p_1,\cdots,p_d)$，我们来计算 T_pM。实际上，我们有

$$T_pM=\mathbb{R}^d\times 0=\{(x_1,\cdots,x_d,0,\cdots,0)\}$$

按照定义，我们选取特殊的曲线

$$\gamma_i:(-1,1)\to\mathbb{R}^n,\quad t\mapsto(0,\cdots,t,0,\cdots),$$

其中，t 出现在第 i 个位置，$i\leqslant d$。这条曲线自然落在 M 中，所以，通过求导数，我们就知道 $(0,\cdots,1,\cdots,0)\in T_pM$，其中，1 出现在第 i 个位置。所以，$\mathbb{R}^d\times\{0\}\subset T_pM$。

为了说明 $\mathbb{R}^d\times\{0\}\subset T_pM$，我们考虑任意一条曲线落在 M 中的曲线 γ。按照 M 的定义，它可以写成

$$\gamma:(-\varepsilon,\varepsilon)\to\mathbb{R}^n,\quad t\mapsto\gamma(t)=(\gamma_1(t),\cdots,\gamma_d(t),0,\cdots,0),$$

对 t 求导数，我们得到

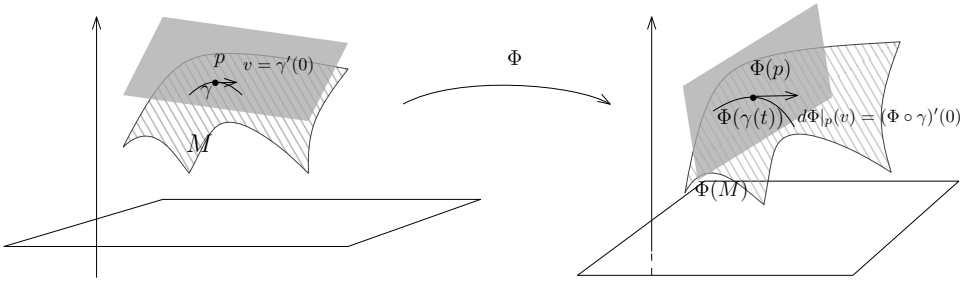
$$\gamma'(0)=(\gamma_1'(0),\cdots,\gamma_d'(0),0,\cdots,0).$$

这显然是 $\mathbb{R}^d\times\{0\}$ 中的一条曲线。

我们要**强调**关于这个例子的一个事实：上面的计算表明，T_pM 作为 $\mathbb{R}^n$ 的线性子空间是不依赖于 p 的，这只是（混淆视听的）**巧合**。

命题 210. 切空间在微分同胚下不变，即若 $\Phi:U\to V$ 是微分同胚，$M\subset U$ 是子流形，$p\in U$，其中 U 和 V 是 $\mathbb{R}^n$ 中的开集。那么，$\Phi(M)$ 是子流形并且 $\Phi(p)$ 的切空间为

$$T_{\Phi(p)}\Phi(M)=d\Phi\Big|_{x=p}(T_pM).$$



证明：$\Phi(M)\subset V$ 是子流形是平凡的，请参考本周的作业。为了说明 $T_{\Phi(p)}\Phi(M)=d\Phi\Big|_{x=p}(T_pM)$，我们考虑经过 p 点的 M 上的一条参数曲线

$$\gamma:(-\varepsilon,\varepsilon)\to M,\quad t\mapsto\gamma(t).$$

通过与 Φ 复合，我们得到了 V 中的曲线 $\Phi\circ\gamma$。按照定义，这是子流形 $\Phi(M)$ 上的通过 $\Phi(p)$ 点曲线。很明显，（通过考虑 $\Phi^{-1}\circ\alpha$）$\Phi(M)$ 上的任何一条曲线通过 $\Phi(p)$ 点的曲线都可以通过这种方式得到。所以，我们可以通过复合 Φ，M 上通过 p 点曲线与 $\Phi(M)$ 通过 $\Phi(p)$ 点曲线是一一对应的。

现在可以对曲线求导数来计算 $T_{\Phi(p)}\Phi(M)$。假设 $v\in T_pM$，按照定义，这意味着存在 M 上过 p 点的曲线 γ，使得 $\gamma'(0)=v$。我们现在证明：

$$d\Phi\big|_{x=p}(v)=\frac{d}{dt}\Big|_{t=0}\Phi(\gamma(t)),$$

即 $\Phi:U\to V$ 的微分将 γ 的切向量映射成 $\Phi\circ\gamma$ 的切向量。按照微分定义，我们有

$$\gamma(t)=p+t\gamma'(0)+o(t)$$

以及

$$\Phi(p+h)=\Phi(p)+d\Phi(p)(h)+o(h),$$

所以

$$\begin{aligned}(\Phi\circ\gamma)(t)&=\Phi\big(p+t\gamma'(0)+o(t)\big)\\&=\Phi(p)+d\Phi(p)(t\gamma'(0)+o(t))+o(\gamma'(0)t+o(t))\\&=\Phi(p)+td\Phi(p)(\gamma'(0))+o(t).\end{aligned}$$

所以，

$$(\Phi\circ\gamma)'(0)=d\Phi(p)(\gamma'(0)).$$

这就是要证明的等式。这个式子表明 $d\Phi\Big|_{x=p}(T_pM)\supset T_{\Phi(p)}\Phi(M)$。由于 $\Phi(M)$ 中的每条曲线都是这么来的，所以 $d\Phi\Big|_{x=p}(T_pM)=T_{\Phi(p)}\Phi(M)$。 □

定理 211. *T_pM 是 $\mathbb{R}^n$ 的一个 d 维线性子空间。*

进一步，按照 M 的（局部）定义方式，我们有如下两种方式刻画切空间：

1) *按子流形的定义，存在开集 $U\subset\mathbb{R}^n$ 包含 p 和 $\mathbb{R}^n$ 中的开集 V 以及微分同胚 $\Phi:U\to V$ 使得 $\Phi:U\cap M=V\cap(\mathbb{R}^d\times\{0\})$。那么，*

$$T_pM=(d\Phi)^{-1}\Big|_{\Phi(p)}(\mathbb{R}^d\times\{0\}).$$

2) 假设 M 在 p 的局部上由 $n-d$ 个光滑函数 $f_1,\cdots,f_{n-d}$ 的零点定义，即存在开集 $U\subset\mathbb{R}^n$，使得 $p\in U$，满足

$$M=\big\{x\in U\big|f_1(x)=\cdots=f_{n-d}(x)=0\big\}$$

并且对任意的 $x\in M$，$n-d$ 个线性函数 $f_1,\cdots,f_{n-d}$ 是线性无关的，那么

$$T_pM=\bigcap_{k=1}^{n-d}\ker\,df_i\big|_{x=p}=\ker\,df\big|_{x=p}.$$

其中 $f=(f_1,\cdots,f_{n-d})$。

证明: 我们首先看 1)，由于 $\Phi(M)$ 为 $\mathbb{R}^d\times\{0\}$ 中的开集，根据前面的命题以及线性情况的计算，我们有

$$\mathbb{R}^d\times\{0\}=T_{\Phi(p)}\Phi(M)=d\Phi\Big|_{x=p}(T_pM).$$

两边同时取 $d\Phi\Big|_{x=p}$ 的逆就完成了证明。特别地，T_pM 具有线性子空间的结构，就是通过这个映射:

$$\begin{array}{ccc}T_pM & \xrightarrow{d\Phi} & \mathbb{R}^d\times\{0\}\\ \downarrow & & \downarrow\\ \mathbb{R}^n & \xrightarrow{d\Phi} & \mathbb{R}^n\end{array},$$

因为我们可以把 T_pM 与 $\mathbb{R}^d\times\{0\}$ 等同，所以我们就用 $\mathbb{R}^d\times\{0\}$ 上的线性空间结构作为 T_pM 上的线性空间结构就好。特别地，因为 $\mathbb{R}^d\times\{0\}$ 的维数是 d，所以 T_pM 的维数也是 d。[9]

现在证明 2)，我们任取 $v\in T_pM$，按照定义，存在 M 上的过 p 点的曲线 γ，使得 $\gamma'(0)=v$。由于 M 是 f_i 的零点集，所以

$$f_i(\gamma(t))\equiv 0,\quad \text{对任意的 } t.$$

对 t 求导数，链式法则给出我们

$$df_i(p)\big(\gamma'(0)\big)=0,\quad\Leftrightarrow\quad df_i(p)\big(v\big)=0.$$

这表明 $T_pM\subset\bigcap_{k=1}^{n-d}\ker\,df_i$。另外，由于 $\dim\bigcap_{k=1}^{n-d}\ker\,df_i\geqslant d$，通过比较维数，我们就知道这两个空间是一致的。 □

注记. （重要） 切空间是内蕴定义的，完全不依赖于坐标系，因为在定义中我们是用 M 上的曲线来定义的。上面定理中对切空间的描述或者依赖于坐标 φ 的选取或者依赖于 M 的定义函数 f_i 的选取。然而，**因为**切空间是内蕴定义的，**所以**我们可以用任意的坐标或者定义函数来计算（我们在应用的时候总是选最容易计算的坐标或函数）。

[9]这里线性空间的结构的定义依赖于 $\Phi:U\to V$ 的选取，我们将在习题课中说明这个选取实际上不依赖于 Φ 的选取。

例子 (超曲面的切空间与法向量的概念)**.** 假设 M 是 $\mathbb{R}^n$ 中由一个函数 f 整体定义的超曲面（R^3 中的曲面是我们最常见的例子），它的定义函数为 $f \in C^\infty(\mathbb{R}^n)$。此时，我们在 $\mathbb{R}^n$ 选定坐标系。按照定义，对于 $p = (x_1, \cdots, x_n)$，那么

$$T_pM = \ker\, df(p) = \left\{ v \in \mathbb{R}^n \middle| df(p)(v) = 0 \right\}.$$

用坐标系（矩阵的语言）来写，我们有

$$v = (v_1, \cdots, v_n), \quad df(p) = \left(\frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}, \cdots, \frac{\partial f}{\partial x_n} \right)(p),$$

所以，

$$df(p)(v) = 0 \quad \Leftrightarrow \quad \sum_{i=1}^{n} v_i \frac{\partial f}{\partial x_i}(p) = 0.$$

我们假定在 $\mathbb{R}^n$ 中取定一个 *Euclid* 内积 $\langle \cdot, \cdot \rangle$，使得 $\left\{ \frac{\partial}{\partial x_i} \right\}_{i \leqslant n}$ 恰好是标准正交基。此时，我们用如下的符号表示上述向量

$$\nabla f(p) = \left(\frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}, \cdots, \frac{\partial f}{\partial x_n} \right)(p).$$

那么，上面的关系就可以表达为

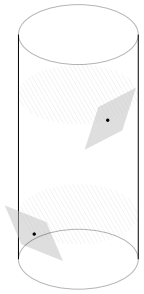
$$\langle v, \nabla f(p) \rangle = 0.$$

换句话说，$\nabla f(p)$ 是切平面的**法向量**。由于 $\nabla f(p) \neq 0$，我们经常把这个向量正规化成长度为 1 的向量：

$$\frac{\nabla f(p)}{|\nabla f(p)|}.$$

我们习惯上把 $\nabla f(p)$ 称作 f 在 p 处的**梯度**。在绝大多数的应用中，我们都（无意识地）自动假定在 $\mathbb{R}^n$ 中有上述 *Euclid* 内积。[10]

利用上面的讨论，我们很容易计算 $\mathbb{R}^3$ 中曲面的切空间，比如说，我们来看柱面的例子：



例子. 柱面的定义方程是 $f = 0$，其中

$$f(x, y, z) = x^2 + y^2 - 1$$

[10] 从微分几何的观点，这一点有待商榷，有兴趣的同学可以在学习 Riemann 几何的时候回头思考这个问题。

对任意给定的 (x, y, z)，很容易算出

$$\nabla f(p) = (2x, 2y, 0)$$

所以，我们知道 $(-x, y, 0)$ 和 $(0, 0, 1)$ 都和这个点垂直，所以，这两个向量就张成的切空间。特别地，我们也可以从图形或者这个计算上清楚的看出，切空间 T_pM 是依赖于它的基准点 p。和之前线性子空间的切空间的例子对比，我们强调说不同点的切空间是不同的。我们之所以认为它们有时候是一样的，是因为我们把它们都看成了 $\mathbb{R}^n$ 的子空间所以它们有可能相同。

35.1 作业：隐函数与反函数定理，隐函数定理在多项式和矩阵上的一个重要应用，经典群的子流形结构

清华大学 19-20 春季学期，数学分析二，作业 3

请用 A4 大小的纸张**正反面**用钢笔，签字笔或者圆珠笔书写，并注明自己的姓名，年级和作业的**总页数**。除定理公式所涉及的人名之外，请使用**中文**。本次作业的提交时间和地点为 **2 月 43 日**上午的课堂上，逾期视作零分。

A. 子流形的基本习题（课堂相关）

A1) $M \subset \mathbb{R}^n$ 和 $N \subset \mathbb{R}^m$ 是光滑子流形。证明，$M \times N \subset \mathbb{R}^{n+m}$ 也是光滑子流形。

A2) $M \subset \mathbb{R}^n$ 是光滑子流形，$\varphi : \mathbb{R}^n \to \mathbb{R}^n$ 是微分同胚。证明，M 在改映射下的像 $\varphi(M)$ 也是光滑子流形。

A3)（光滑超曲面的判定）假设 $f : \mathbb{R}^n \to \mathbb{R}$ 是光滑函数，如果对于任意的 $x_0 \in f^{-1}(0)$，存在某个 $i \leqslant n$（指标 i 可能依赖于 x_0），使得 $\dfrac{\partial f}{\partial x_i}(x_0) \neq 0$，那么 $f^{-1}(0)$ 是光滑超曲面。

A4) 假设 $\Omega \subset \mathbb{R}^n$ 是开集，$f_1, \cdots, f_k$ 是给定的 k 个光滑函数，其中 $k \leqslant n$。令

$$Z_i = \{x \in \Omega | f_i(x) = 0\}.$$

假设 $x \in \bigcap\limits_{i \leqslant k} Z_i$ 并且 $df_1(x), \cdots, df_k(x)$ 线性无关。证明，存在 x 附近的开集 U，使得对任意的 $\ell \leqslant k$，$\bigcap\limits_{i \leqslant \ell} Z_i$ 是维数为 $n - \ell$ 的子流形。

A5) 假设 $f(x, y, z)$ 和 $g(x, y, z)$ 是 $\mathbb{R}^3$ 上的光滑函数，如果对于任意同时满足 $f(x_0, y_0, z_0) = 0$ 和 $g(x_0, y_0, z_0) = 0$ 的点 (x_0, y_0, z_0)，如果如下两个向量线性不相关：

$$\left(\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}, \frac{\partial f}{\partial z}\right)(x_0, y_0, z_0), \quad \left(\frac{\partial g}{\partial x}, \frac{\partial g}{\partial y}, \frac{\partial g}{\partial z}\right)(x_0, y_0, z_0),$$

证明，

$$\left\{(x, y, z) \in \mathbb{R}^3 \middle| f(x, y, z) = 0, g(x, y, z) = 0\right\}$$

是光滑曲线。

A6) $\Omega \subset \mathbb{R}^n$ 是开集，$f : \Omega \to \mathbb{R}^m$ 是光滑映射。证明，f 的图像

$$\Gamma_f = \left\{(x, y) \in \mathbb{R}^{n+m} \middle| x \in \mathbb{R}^n, y \in \mathbb{R}^m, y = f(x)\right\}$$

是 $\mathbb{R}^{n+m}$ 中的 n 维光滑子流形。

A7) 完整证明课堂上的命题：Möbius 带 $\mathcal{M}$ 不能用一个光滑函数的零点定义（要求函数的微分在 $\mathcal{M}$ 上非零）。

A8) 假设 $M \subset \mathbb{R}^n$ 是 0-维子流形。证明，对任意的 $x \in M$，存在开集 $U \subset \mathbb{R}^n$，使得 $U \cap M = \{x\}$。（零维子流形局部上就是单点集）

A9) 假设 $M \subset \mathbb{R}^n$ 是 n-维子流形。证明，M 是 $\mathbb{R}^n$ 中的开集。

B. 隐函数与反函数定理的习题

B1) 实值函数 $f \in C^\infty(\mathbb{R})$ 满足下面的的性质：存在 $\delta > 0$，使得对任意 $x \in \mathbb{R}$，都有

$$f'(x) \geqslant \delta$$

证明，$f : \mathbb{R} \to \mathbb{R}$ 是微分同胚。如果 $\delta = 0$，试举一个反例。

B2) a 和 b 是实数，我们定义映射

$$f : \mathbb{R}^2 \to \mathbb{R}^2, \quad (x, y) \mapsto (x + a \sin y, y + b \sin x).$$

证明：f 是 $\mathbb{R}^2$ 到自身的微分同胚当且仅当 $ab \in (-1, 1)$。

B3)（隐函数定理的经典练习）假设 $f : \mathbb{R}^n \to \mathbb{R}^n$ 是光滑映射，0 是 f 的不动点，即 $f(0) = 0$。假设 1 不是 f 在 0 处的微分

$$df(0) : \underbrace{T_0\mathbb{R}^n}_{\mathbb{R}^n} \longrightarrow \underbrace{T_0\mathbb{R}^n}_{\mathbb{R}^n}$$

的特征值。

- 证明，0 是一个孤立的不动点，即存在开集 U，使得 $0 \in U$，f 在 U 有且仅有 0 为其不动点。
- 任意给定 $g \in C^1(\mathbb{R}^n, \mathbb{R}^n)$。对任意的 $\lambda \in \mathbb{R}$，我们定义映射

  $$f_\lambda : \mathbb{R}^n \longrightarrow \mathbb{R}^n, \quad x \mapsto f(x) + \lambda g(x),$$

  证明，存在 $\varepsilon > 0$，存在 0 在 $\mathbb{R}^n$ 的开邻域 U 以及 C^1 的映射

  $$\omega : (-\varepsilon, \varepsilon) \to V, \quad \lambda \mapsto \omega(\lambda),$$

  使得对任意的 $\lambda \in (-\varepsilon, \varepsilon)$，$f_\lambda$ 在 V 中恰好有一个不动点 $\omega(\lambda)$。

B4) 证明，对任意的 $t \in (-0.5, 0, 5)$，方程

$$\sin(tx) + \cos(tx) = x$$

存在唯一一个解 $x = \varphi(t) \in \mathbb{R}$。进一步证明 $\varphi(t)$ 是 t 的光滑函数并计算 $\varphi''(0)$。

B5) 证明，如下两个方程定义出 $\mathbb{R}^3$ 中的一条光滑曲线：

$$x^2+y^2+z^2=14,\ \ x^3+y^3+z^3=36.$$

对于曲线上的每一点，计算它的切空间。

B6) 在 $\mathbb{R}^2$ 中考虑如下的集合

$$M=\big\{(x,y)\big|x^3+y^3-3xy=1\big\}.$$

- 证明，在 $(0,1)\in M$ 附近，存在 $\mathbb{R}$ 上 0 的邻域 Ω 以及 Ω 上的光滑函数 $\phi\in C^\infty(\Omega)$，使得 M 是 ϕ 的图像并且 $\phi(0)=1$。
- 试计算 $\phi'(0)$ 和 $\phi''(0)$。
- 证明，M 是光滑曲线。
- 任意给定 $(x_0,y_0)\in M$，试计算 M 在此点处的切空间。

B8) 证明，如下两个方程定义的集合 Σ

$$\begin{cases} x^2+2y^2+3z^2+4t^2=4\\ x+y=z+t\end{cases}$$

是 $\mathbb{R}^4$ 中的光滑曲面。

B9)（承接上一问题）对那一些 t（t 固定），如下集合

$$\Sigma_t=\big\{(x,y,z)\in\mathbb{R}^3\big|x^2+2y^2+3z^2+4t^2=4,\ x+y=z+t\big\}$$

是 $\mathbb{R}^3$ 中的光滑曲线?（我们用时间把 Σ 切成曲线“片”）

C 子流形上生活着很多曲线

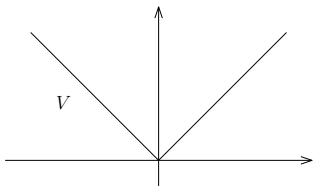
C1) 假设 $M\subset\mathbb{R}^n$ 是子流形并且 $\dim M\geqslant 1$，$p\in M$。那么，存在光滑曲线（作为映射）

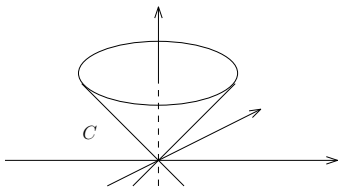
$$\gamma:(-1,1)\to\mathbb{R}^n,$$

使得 $\gamma(0)=p$，$\gamma'(0)\neq 0$ 并且 $\gamma\big((-1,1)\big)\subset M$。

C2) 令 $V=\big\{(x,y)\in\mathbb{R}^2\big|\ y=|x|\big\}$ 为 $\mathbb{R}^2$ 中的 V-型，它在原点处有一个尖，看起来不光滑。证明，V 不是 $\mathbb{R}^2$ 的光滑子流形。

注记. 实际上我们是利用切空间的概念证明 V 不是子流形。





C3) 考虑 $\mathbb{R}^3$ 中的锥

$$C = \big\{(x_0, x_1, x_2)\big| -x_0^2 + x_0^2 + x^2 = 0\big\}.$$

令 $O = (0,0,0)$，证明，$C - \{O\}$ 是子流形。

C4) 证明，C 不是（光滑）子流形。

D 隐函数定理在多项式和矩阵上的一个重要应用（熟记）

D1) 对任意的 $c \in \mathbb{R}^n$，我们可以定义一个为 n 的实系数多项式 P_c，即我们有映射

$$\mathbb{R}^{n+1} \to \mathbb{R}[X], \quad c \mapsto X^n + c_1 X^{n-1} + \cdots + c_n.$$

给定 $b \in \mathbb{R}^n$，我们假设 P_b 有一个实数根 $x_b \in \mathbb{R}$ 并且这个根的重数是 1。证明，存在 b 在 $\mathbb{R}^n$ 中的开邻域 U 和 x_b 在 $\mathbb{R}$ 中的开邻域 V，使得对任意的 $c \in U$，P_c 在 V 中恰有一个根 $z(c)$ 并且重数是 1。进一步证明映射

$$V \to \mathbb{R}, \quad c \mapsto z(c)$$

是光滑的。

D2)（根对系数的光滑依赖性）考虑实系数多项式 $P(X) = X^n + c_1 X^{n-1} + \cdots + c_n = 0$。如果对 $(c_1^*, \cdots, c_n^*) \in \mathbb{R}^n$, $P(X)$ 恰好有 n 个两两不同的实根 $z_1(c_1^*, \cdots, c_n^*) < \cdots < z_n(c_1^*, \cdots, c_n^*)$。证明，存在 $(c_1^*, \cdots, c_n^*) \in \mathbb{R}^n$ 的开邻域 Ω 和 Ω 上的光滑函数 $z_1, \cdots, z_n$，使得对任意的 $(c_1, \cdots, c_n) \in \Omega$，我们有

$$z_1(c_1, \cdots, c_n) < z_2(c_1, \cdots, c_n) < \cdots < z_n(c_1, \cdots, c_n)$$

并且

$$P(z_1(c_1, \cdots, c_n), z_2(c_1, \cdots, c_n), \cdots, z_n(c_1, \cdots, c_n)) = 0.$$

D3*) 你是否能将上面的结论推广到复系数的多项式上去?

D4) 考虑 $n \times n$ 的实矩阵 $A = (A_{ij})_{i,j \leqslant n}$，假设 A 有 n 个不同的实特征值 $\lambda_1 < \lambda_2 < \cdots < \lambda_n$。证明，存在 $\varepsilon > 0$，使得对任意的 $n \times n$ 的实矩阵 $B = (B_{ij})_{i,j \leqslant n}$，如果对每对指标 (i,j)，$|A_{ij} - B_{ij}| < \varepsilon$，那么 B 有 n 个不同的实特征值 $\lambda_1(B) < \lambda_2(B) < \cdots < \lambda_n(B)$。进一步证明，如果将每个 λ_i 看做是 B 的系数 B_{ij} 的函数，这些函数在区域 $|B_{ij} - A_{ij}| < \varepsilon$（其中 $i, j \leqslant n$）中是光滑函数。

D5*) 你是否能将上面的结论推广到复系数的矩阵上去?

E 经典群的子流形结构

$\mathbf{M}_n(\mathbb{R})$ 是全体 $n\times n$ 的实系数矩阵。我们定义**特殊线性群**

$$\mathbf{SL}_n(\mathbb{R})=\{A\in\mathbf{M}_n(\mathbb{R})\big|\det A=1\},$$

和正交群

$$\mathbf{O}(n)=\{A\in\mathbf{M}_n(\mathbb{R})\big|A\cdot{}^tA=\mathbf{I}_n\},$$

其中 $\mathbf{I}_n$ 是单位矩阵。

E1) 证明，在矩阵乘法下，上面两个对象都是群。

E2) 证明，$\mathbf{SL}_n(\mathbb{R})$ 是 $\mathbf{M}_n(\mathbb{R})=\mathbb{R}^{n^2}$ 中的光滑子流形并计算它的维数。（提示：将行列式映射 $A\mapsto\det A$ 视作是 $\mathbb{R}^{n^2}$ 上的光滑函数）

E3) 证明，$\mathbf{SL}_n(\mathbb{R})$ 在 $\mathbf{I}_n$ 处的切空间是

$$\mathfrak{sl}_n(\mathbb{R}):=\{a\in\mathbf{M}_n(\mathbb{R})\big|\mathrm{tr}a=0.\}$$

E4) 证明，$\mathbf{O}(n)$ 是 $\mathbf{M}_n(\mathbb{R})=\mathbb{R}^{n^2}$ 中的光滑子流形并计算它的维数（提示：令 $\mathrm{Sym}_n(\mathbb{R})$ 是全体对称的 $n\times n$ 的实系数矩阵，考虑映射 $f:\mathbf{M}_n(\mathbb{R})\to\mathrm{Sym}_n(\mathbb{R}),\ A\mapsto A\cdot{}^tA-\mathbf{I}_n$）。

E5) 证明，$\mathbf{O}_n(\mathbb{R})$ 在 $\mathbf{I}_n$ 处的切空间是

$$\mathfrak{o}(n):=\{a\in\mathbf{M}_n(\mathbb{R})\big|a+{}^ta=0.\}$$

E6) 令 $G=\mathbf{SL}_n(\mathbb{R})$ 或者 $\mathbf{O}(n)$。证明，映射

$$G\to G,\ g\mapsto g^{-1}$$

是光滑的。

E7) 回忆一下，$G\times G$ 可以被视作是 $\mathbf{M}_n(\mathbb{R})\times\mathbf{M}_n(\mathbb{R})$ 的子流形。证明，映射

$$G\times G\to G,\ \ (g_1,g_2)\mapsto g_1\cdot g_2$$

也是光滑的。（至此，证明了 $\mathbf{SL}_n(\mathbb{R})$ 和 $\mathbf{O}(n)$ 是 Lie 群）

E8) 令 $G=\mathbf{SL}_n(\mathbb{R})$（或 $\mathbf{O}(n)$），$\mathfrak{g}=\mathfrak{sl}_n(\mathbb{R})$（或 $\mathfrak{o}(n)$），证明，指数映射 exp 把 $\mathfrak{g}$ 映射到 G 中去。

The knowledge of which geometry aims is the knowledge of the eternal.

—— Plato

---

36 子流形之间的光滑映射，子流形之间光滑映射的微分，切向量场的定义（$\neq$ 向量值函数），切丛，切丛是光滑子流形，子流形上的微分学与极值问题：Lagrange 乘子法，子流形上的反函数定理和隐函数定理

二零二零年三月九日，星期一，晴

子流形之间的映射

给定两个子流形 $M \subset \mathbb{R}^m$ 和 $N \subset \mathbb{R}^n$，其中，M 和 N 的维数是任意的，我们定义它们之间的光滑映射 $f : M \to N$ 和以及 f 的微分 df。最常见的例子中通常 $N = \mathbb{R}$，这实际上是 M 上的光滑函数。

定义 212. *$M \subset \mathbb{R}^m$ 和 $N \subset \mathbb{R}^n$ 是子流形，$f : M \to N$ 是映射。如果 f 满足如下任意一个条件：*

1)（f 局部上是 $\mathbb{R}^m$ 上的映射的限制）对任意 $p \in M$，存在包含 p 的开集 $U \subset \mathbb{R}^m$，存在 C^∞ 的映射 $F : U \to \mathbb{R}^n$，使得

$$f\big|_{U\cap M} = F\big|_{U\cap M}.$$

2) 对任意 $p \in M$，按子流形的定义，存在包含 p 的开集 $U \subset \mathbb{R}^m$，$\mathbb{R}^m$ 中的开集 V 以及微分同胚 $\Phi : U \to V$ 使得 $\Phi(U \cap M) = V \cap (\mathbb{R}^d \times \{0\})$。那么，映射

$$f_\Phi = \left(\Phi^{-1}\right)^* f = f \circ \Phi^{-1}$$

是在 $V \cap (\mathbb{R}^d \times \{0\}) \subset \mathbb{R}^d$ 上定义并在 $\mathbb{R}^n$ 中取值的 C^∞ 映射：

$$\begin{array}{ccc} V \cap (\mathbb{R}^d \times \{0\}) & \xrightarrow{\Phi^{-1}} & M \cap U \\ & \searrow^{f_\Phi} & \downarrow f \\ & & N \subset \mathbb{R}^n \end{array}$$

我们就称 f 是 M 与 N 之间的 C^∞ 的映射并且将所有这样的映射的全体记作 $f \in C^\infty(M, N)$。

注记. *在严格的证明上述定义是规范的之前，我们先指出这个定义最经常地以如下的方式出现：$M \subset \mathbb{R}^m$ 和 $N \subset \mathbb{R}^n$ 是子流形，$F : \mathbb{R}^m \to \mathbb{R}^n$ 是光滑映射，如果 $F(M) \subset N$，那么，根据定义中的第一条，f 在 M 上的限制就给出了子流形之间的光滑映射：*

$$F\big|_M : M \to N.$$

例子. *考虑 $M = N = \mathbf{S}^2$，这是 $\mathbb{R}^3$ 中的单位球面。考虑正交矩阵 $A \in \mathbf{O}(3)$，即 3×3 的方阵 A，使得 ${}^tA \cdot A = \mathbf{I}$，其中 $\mathbf{I}$ 是单位矩阵。我们可以将 A 看作是 $\mathbb{R}^3$ 上的线性（光滑）映射：*

$$A : \mathbb{R}^3 \to \mathbb{R}^3, \quad x \mapsto A \cdot x.$$

由于正交矩阵 A 保持向量的长度，所以

$$A : \mathbf{S}^2 \to \mathbf{S}^2.$$

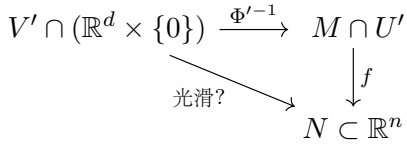
这是 $\mathbf{S}^2$ 到自身的光滑映射，我们通常把它称为是 $\mathbf{S}^2$ 上的一个旋转。

我们需要说明上述定义是规范的，这包含了如下两个内容：

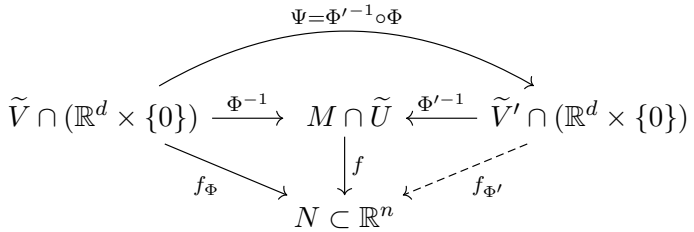
a) 定义 2) 中我们用了子流形的定义，然而，子流形定义中所选取的局部微分同胚 Φ 不是唯一的。对 $p \in M$，可以存在另外的包含 p 的开集 $U' \subset \mathbb{R}^m$，$\mathbb{R}^m$ 中的开集 V' 以及微分同胚 $\Phi' : U' \to V'$，使得 $\Phi(U' \cap M) = V' \cap (\mathbb{R}^d \times \{0\})$。那么，映射

$$f_{\Phi'} = \left(\Phi'^{-1}\right)^* f = f \circ \Phi'^{-1}$$

是否是在 $V \cap (\mathbb{R}^d \times \{0\}) \subset \mathbb{R}^d$ 上的 C^∞ 映射?



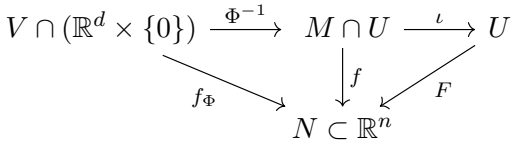
对任意的 p，我们考虑 $\widetilde{U} = U \cap U'$，此时，上述的 Φ 和 Φ' 在 $\widetilde{U}$ 上都能定义，我们令 $\widetilde{V} = \Phi(\widetilde{U})$，$\widetilde{V}' = \Phi'(\widetilde{U})$，我们就有如下的交换图表：



很明显，$f_{\Phi'}$ 是 f_Φ 与 Ψ 的复合，而 Ψ 是两个微分同胚的复合也光滑，所以 $f_{\Phi'}$ 光滑。

b) 我们要说明 1) 和 2) 是等价的。

1) $\Rightarrow$ 2). 我们考虑如下的交换图表：

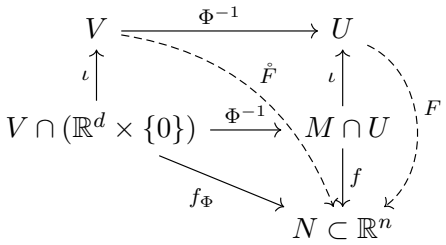


其中，映射 ι 是包含映射，即

$$\iota : M \cap U \to U, \quad x \mapsto x.$$

按照 1) 的要求，F 是光滑的。所以，$f_\Phi = F \circ \iota \circ \Phi^{-1}$ 是光滑的。

2) ⇒ 1). 我们考虑如下的交换图表：



按照 2) 的要求，下面的三角形的图表中的映射都是光滑的。特别地，f_Φ 用坐标来表达（$\mathbb{R}^2\times\{0\}$ 上用 $x_1, \cdots, x_d$ 作为坐标）可以写成

$$V \cap \mathbb{R}^m \times \{0\} \to \mathbb{R}^n, \ \ (x_1, \cdots, x_d) \mapsto f_\Phi(x_1, \cdots, x_d).$$

我们可以把 f_Φ 拓展成 V 上的函数：

$$\mathring{F} : V \to \mathbb{R}^n, \ \ (x_1, \cdots, x_d, x_{d+1}, \cdots, x_n) \mapsto f_\Phi(x_1, \cdots, x_d).$$

很明显，$\mathring{F}$ 在 $V \cap \mathbb{R}^m \times \{0\}$ 上的限制就是 f_Φ。根据上面的图表，我们令 $F = \mathring{F} \circ \Phi$ 即可。

注记. 当 $N = \mathbb{R}$ 时，我们就定义了 M 上的光滑函数 $C^\infty(M)$。与 $\mathbb{R}^n$ 中一个区域上的光滑函数类似，我们有（请参考本次作业作业）

1) $C^\infty(M)$ 是一个 $\mathbb{R}$-代数，即对任意的 $f, g \in C^\infty(\Omega_1)$，它们的任意实线性组合以及乘积都是光滑函数。

2) 假设 $f \in C^\infty(M)$。如果对任意的 $x \in M$，$f(x) \neq 0$，那么 $\dfrac{1}{f}$ 也是光滑函数。

另外，对于一般的 $N \subset M$，$C^k(M, N)$ 通常不是线性空间，因为对于任意的 $f, g \in C^k(M, N)$，$x \in M$，点 $f(x) + g(x)$ 不见得在 N 上。

注记. 我们所定义光滑函数的方式不是内蕴的：第一个定义要求子流形上的光滑函数是背景空间（$\mathbb{R}^m$）上的光滑函数的限制，这依赖于背景空间；第二个要求光滑函数在某个局部的模型下是光滑函数，这依赖于微分同胚的选取（我们当然已经证明了这个选取不重要）。然而，我们要强掉能够在子流形上定义光滑函数是子流形理论最核心的一点，比如说，通过上次的作业，我们知道如下的 V-形不是 $\mathbb{R}^2$ 中的子流形：

$$V = \big\{(x, y) \subset \mathbb{R}^2 \big| y = |x|\big\}.$$

实际上，这是一个所谓的拓扑子流形，即我们可以定义如下的同胚（参考第一次习题课）

$$\Phi : \mathbb{R}^2 \to \mathbb{R} \times \mathbb{R}, \ \ (x, y) \to (x, y - |x|).$$

这个映射是连续的但是不光滑（在 $(0, 0)$ 处）。我们可以在背景空间上选取函数 $f : \mathbb{R}^2 \to \mathbb{R}, \ (x, y) \mapsto y$ 是光滑的，但是它所对应的 f_Φ 用我们的第二个定义就不是光滑的。

从函数论的观点，子流形的定义要求局部上的同胚 Φ 是可微的原因就是容许我们在子流形上面定义光滑函数。

在我们的课程中不再会进一步的深入到如何完全内蕴地定义几何对象的层次，我们总是借助于背景空间（$\mathbb{R}^n$）来讨论。下面将要定义的光滑向量场就是一个很好的例子。

对于子流形之间的光滑映射，我们可以定义微分：

定义 213. 假设 $M \subset \mathbb{R}^m$，$N \subset \mathbb{R}^n$ 是子流形，$f : M \to N$ 是光滑映射。对任意的 $v \in T_pM$，按定义，存在（不止一条）参数化的曲线

$$\gamma : (-\varepsilon, \varepsilon) \to \mathbb{R}^m,$$

使得 $\gamma(0) = p$ 且 $\gamma'(0) = v$。通过和 f 复合，我们得到 N 上过 $f(p)$ 点的曲线：

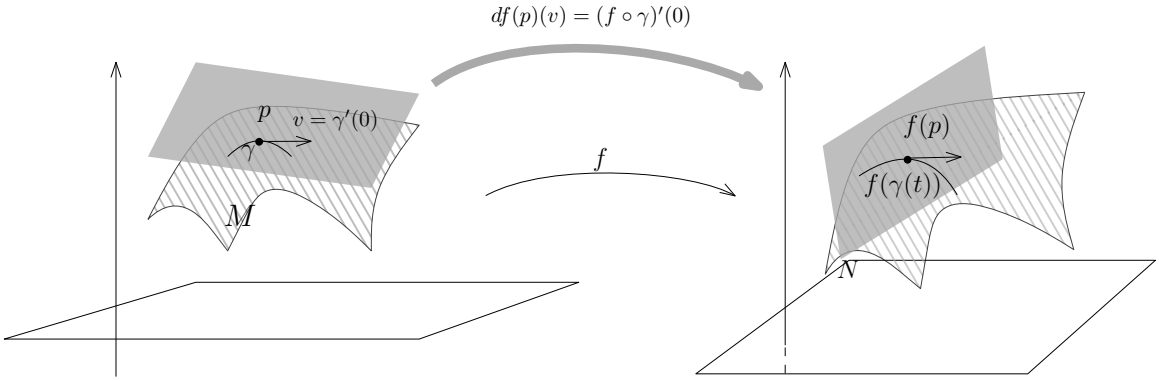
$$f \circ \gamma : (-\varepsilon, \varepsilon) \to \mathbb{R}^n, \quad t \mapsto f(\gamma(t)).$$

我们把这条曲线的在 $f(p)$ 处切向量定义为

$$df(p)(v) := \frac{d}{dt}\Big|_{t=0} f(\gamma(t)) \in T_{f(p)}N.$$

从而，我们定义了

$$df(p) : T_pM \to T_{f(p)}N.$$



注记. 如果子流形之间的映射 F 是由某个全空间上的 $F : \mathbb{R}^m \to \mathbb{R}^n$ 所诱导的，即 $F\big|_M = f$：

$$\begin{array}{ccc} \mathbb{R}^m & \xrightarrow{F} & \mathbb{R}^n \\ {\scriptstyle \iota}\uparrow & & \uparrow{\scriptstyle \iota} \\ M & \xrightarrow{f} & N \end{array}$$

那么，$df(p) : T_pM \to T_{f(p)}N$ 就是 $dF(p) : \mathbb{R}^m \to \mathbb{R}^n$ 在 T_pM 上的限制（T_pM 是 $\mathbb{R}^m$ 的线性子空间），即

$$\begin{array}{ccc} T_p\mathbb{R}^m & \xrightarrow{dF(p)} & T_{f(p)}\mathbb{R}^n \\ {\scriptstyle \iota_*}\uparrow & & \uparrow{\scriptstyle \iota_*} \\ T_pM & \xrightarrow{df(p)} & T_{f(p)}N \end{array}$$

特别地，这个观察还说明了上面定义中的映射不依赖于曲线 γ 的选取。当然，我们还可以直接证明，其参考下面的习题：

练习. 证明，上述定义不依赖于 γ 的选取，即 $\lambda(t)$ 是过 p 的 M 上的参数曲线且 $\lambda'(0)=v$，那么

$$\frac{d}{dt}\Big|_{t=0} f(\gamma(t)) = \frac{d}{dt}\Big|_{t=0} f(\lambda(t)).$$

命题 214. 假设 $M\subset\mathbb{R}^m$，$N\subset\mathbb{R}^n$ 是子流形，$f: M\to N$ 是光滑映射。那么，对任意的 $p\in M$，微分

$$df(p): T_pM \to T_{f(p)}N$$

是线性映射。

证明: 这是显然的，因为 $df(p)$ 是线性映射 $dF(p)$ 的限制，其中我们假设了 f 是背景空间上某个光滑映射 $F:\mathbb{R}^m\to N\subset\mathbb{R}^n$ 的限制。 □

在微分学这一部分，我们最后引进的一个对象叫做向量场，向量场在经典的物理中有着无比重要的应用，我们在后面的课程中会逐步展示这些相关的例子。我们来定义一个子流形上全体切向量的集合：

定义 215 (向量场与切丛)**.** 假设 $M\subset\mathbb{R}^n$ 是子流形。对任意的 $p\in M$，T_pM 是在这一点处的 M 的切向量的全体。我们把这些不同点的切向量放在一起定义为 M 的**切丛**，我们强调这是所有的 T_pM 的无交并集，即不同点处的切向量是不同的：

$$TM = \coprod_{p\in M} T_pM.$$

我们有自然的投影映射 $\pi: TM\to M$，对于 $p\in M$，这个映射把 T_pM 中的元素全部映射成 p，即 $T_pM=\pi^{-1}(p)$。

如果对于每个点 $p\in M$，我们都指定一个切向量 $X(p)\in T_pM$，我们就得到了一个映射

$$X: M\to TM,\quad p\mapsto X(p),$$

我们把这样一个映射称作是 M 上的一个**（切）向量场**。

按照定义，切向量场满足满足 $\pi\circ X=\mathrm{id}_M$：

$$\begin{array}{ccc} M & \xrightarrow{X} & TM \\ & \searrow^{\mathrm{id}} & \downarrow{\scriptstyle \pi} \\ & & M \end{array}$$

注记 (向量场光滑性的一种折衷的定义)**.** 由于所讨论的子流形 M 在 $\mathbb{R}^n$ 中，对任意的 p，我们总可以将 $X(p)$ 看作是 $\mathbb{R}^n$ 中的向量，所以我们可以 X 看作是映射 $X: M\to\mathbb{R}^n$（要求 $X(p)\in T_pM$）。此时，我们就可以谈论 X 的光滑性。如果映射 $X: M\to\mathbb{R}^n$ 是光滑的，我们就称 X 为**光滑的（切）向量场**。我们将 M 上的光滑切向量的全体记作 $\Gamma(M,TM)$。

对任意的 M 上的光滑函数 f，任意的 M 上的光滑切向量场 X 和 Y，我们可以定义它们之间的乘法和加法：

$$(fX)(p) = f(p)X(p),\ \ (X+Y)(p) = X(p) + Y(p).$$

我们有如下简单的命题：

命题 216. *假设 $M \subset \mathbb{R}^n$ 是子流形，那么，$\Gamma(M, TM)$ 是 $C^\infty(M)$-模，即对任意的 M 上的光滑函数 f，任意的 M 上的光滑切向量场 X 和 Y，fX 和 $X+Y$ 都是 M 上的光滑向量场。*

证明是平凡的，我们留作作业。

对于一个向量 $v \in TM$，我们假设 $p = \pi(v)$，即 $v \in T_pM$，这样子，我们可以认为 TM 是 $\mathbb{R}^n \times \mathbb{R}^n$ 的子集：

$$TM \hookrightarrow \mathbb{R}^n \times \mathbb{R}^n,\ \ v \mapsto (\pi(v), v).$$

我们有如下漂亮的定理：

定理 217. *假设 $M \subset \mathbb{R}^n$ 是子流形，那么*

$$TM = \{(x, v) \in \mathbb{R}^n \times \mathbb{R}^n \big| x \in M, v \in T_xM\}$$

是 $\mathbb{R}^{2n}$ 的子流形并且 $\dim TM = 2\dim M$。

证明： 令 $d = \dim M$。任意选取 $v \in T_pM$，我证明 TM 在 p 附近是子流形。先在 p 的附近用光滑函数的零点来定义 M：选取开集 $U \subset \mathbb{R}^n$，其中 $p \in M$ 以及 $n-d$ 个光滑函数 $f_i \in C^\infty(U)$，其中 $i \leqslant n-d$，使得 $M \cap U = \bigcap\limits_{i \leqslant n-1} f_i^{-1}(0)$ 并且对任意的 $q \in U$，微分 $df_1(q), \cdots, df_{n-i}(q)$ 是线性无关的。然而，我们知道对任意的 q，该点处的切空间为

$$T_qM = \bigcap_{i \leqslant n-d} \ker\, d_i f(q),$$

所以该点处的切空间可以实现为 $df_1(q), \cdots, df_{n-i}(q)$ 的公共零点集。据此，我们有

$$TM \cap (U \times \mathbb{R}^n) = \left\{(x, y) \in U \times \mathbb{R}^n \middle| \begin{cases} f_1(x) = 0, \cdots, f_{n-d}(x), \\ df_1(x)(y) = 0, \cdots, df_{n-d}(x)(y) = 0. \end{cases} \right\}.$$

所以，TM 在开集 $U \times \mathbb{R}^n$ 上由 $2(n-d)$ 个光滑函数 $f_1, \cdots, f_{n-d}, df_1(x)(y), \cdots, df_{n-d}(x)(y)$ 的公共零点定义。我们下面说明的微分是线性无关的即可。为此，我们令

$$\begin{aligned} &f : U \to \mathbb{R}^{n-d},\ \ x \mapsto (f_1(x), \cdots, f_{n-d}(x)); \\ &df : U \times \mathbb{R}^n \to \mathbb{R}^{n-d},\ \ (x, y) \mapsto (df_1(x)(y), \cdots, df_{n-d}(x)(y)). \end{aligned}$$

定义 F 为

$$F : U \times \mathbb{R}^n \to \mathbb{R}^{n-d} \times \mathbb{R}^{n-d},\ (x, y) \mapsto (f(x), df(x)(y)).$$

从而，$TM \cap U \times \mathbb{R}^n = F^{-1}(0)$。为了说明 $\mathrm{rank} dF = 2(n-d)$（满秩），我们可以计算它的 Jacobi 矩阵：

$$dF(x,y) = \begin{pmatrix} df(x) & 0 \\ * & df(x) \end{pmatrix},$$

由于 $\mathrm{rank} df = n-d$，所以 $\mathrm{rank} dF = 2(n-d)$。命题证明完毕。 □

注记. 假设 $M \subset \mathbb{R}^m$，$N \subset \mathbb{R}^n$ 是子流形，$f : M \to N$ 是光滑映射。对于固定的点 p，我们知道微分 $df(p)$ 是如下的线性映射：

$$df(p) : T_pM \to T_{f(p)}N.$$

当 p 变化时，我们得到映射：

$$df : TM \to TN.$$

我们将它称作是 f 的**微分**或者**切映射**。由于 TM 和 TN 都是子流形，这是子流形之间的映射，我们还可以证明 $df : TM \to TN$ 是光滑映射，我们把证明留作作业。

应用：在子流形上的微分学

假设 $M \subset \mathbb{R}^m$ 是子流形，我们可以利用 M 上的向量场 X 对 M 上的光滑函数求方向导数，即我们可以把向量场这种几何对象看做是求微分这种代数操作：

$$\Gamma(M, TM) \xrightarrow{\text{方向导数}} \text{一阶微分算子}$$

实际上，对于 $f \in C^\infty(M)$，$p \in M$，$X(p) \in T_pM$ 由曲线 $\gamma(t)$ 所定义。我们可以定义 f 在 p 处的方向导数为

$$\nabla_{X(p)} f(p) = \frac{d}{dt}\Big|_{t=0} f(\gamma(t)).$$

为了方便，我们还用 $X(f)(p)$ 表示上面的方向导数。我们选取向量场 $X \in \Gamma(M, TM)$。当 p 变化时，我们就得到了 M 上的函数 $X(f) = \nabla_X f$，即

$$X(f)(p) = \big(\nabla_{X(p)} f\big)(p).$$

（这实际上是用 $\dfrac{\partial}{\partial x_i}$ 来表示向量场的原因）

命题 218. 上面定义的 $X(f)$ 是光滑函数，即

$$X : C^\infty(M) \to C^\infty(M).$$

证明: 按照定义，如果 X 给定，$X(f)$ 只与 f 和 X 在 M 上的值有关系。所以，我们任意选取 $F \in C^\infty(\mathbb{R}^n)$ 和 $\mathcal{X} \in C^\infty(\mathbb{R}^n; \mathbb{R}^n)$，使得 $F\big|_M = f$，$\mathcal{X}\big|_M = X$。那么，上面的定义就是在计算 F 在 $\mathbb{R}^n$ 沿着 $\mathcal{X}$ 的方向导数，所以 $\nabla_{\mathcal{X}} F$ 光滑的，它的限制就给出了 $\nabla_X f$，从而光滑。 □

如果 M 上的可微函数有极值，那么我们也有

命题 219. *假设 $M \subset \mathbb{R}^m$ 是子流形，$f \in C^\infty(M)$，p 是 f 的最大值（局部）点，那么，$df(p) = 0$。*

证明: 根据微分的定义，我们要证明对任意的 $v \in T_pM$，$\nabla_v f(p) = 0$ 即可。实际上，我们考虑 M 上通过 p 的曲线

$$\gamma : (-\varepsilon, \varepsilon) \to M, \quad t \mapsto \gamma(t),$$

其中 $\gamma(0) = p$，$\gamma'(0) = v$。我们有

$$\frac{d}{dt}\Big|_{t=0} f \circ \gamma = (\nabla_v f)(p).$$

由于 p 是 $f \circ \gamma$ 的最大值，所以 $(f \circ \gamma)'(0) = 0$，这表明对任意的 $v \in T_pM$，$(\nabla_v f)(p) = 0$。我们注意到这个证明在形式上和在 $\mathbb{R}^n$ 的一个开集山的证明一摸一样，我们都是把证明化简为一元微分学的形式。 □

求极值问题

利用上面证明的命题，我们可以用几何的眼光来看所谓的 **Lagrange 乘子法**，这是多元微积分在求极值方面的重要方法。

我们先把要讨论的问题说清楚：假设 $\mathbb{R}^n$ 上有光滑函数 $f(x_1, \cdots, x_n)$，我们想要计算它的最大值（比如说），然而，这个问题是所谓的带有**约束条件的**。所谓的约束条件指的是点 $x \in \mathbb{R}^n$ 必须满足 m 个方程：

$$\begin{cases} g_1(x_1, x_2, \cdots, x_n) = 0, \\ g_2(x_1, x_2, \cdots, x_n) = 0, \\ \qquad \cdots \quad \cdots, \\ g_m(x_1, x_2, \cdots, x_n) = 0. \end{cases}$$

我们通常要求 $m \leqslant n$（不能有太多的约束）。我们用更简洁的语言来叙述这个问题。令

$$g : \mathbb{R}^n \to \mathbb{R}^m, \ x \mapsto g(x) = (g_1(x), \cdots, g_m(x)),$$

我们要找到 $x \in g^{-1}(0)$，使得 x 是 g 的局部极大值。以下我们进一步假定 g 是光滑的，$M = g^{-1}(0)$ 并且对任意 $x \in M$，$\mathrm{rank} dg(p) = m$。此时，M 是 $\mathbb{R}^n$ 中的子流形。那么，我们的约束条件极值问题等价于求函数 f 在子流形 M 上的极值。

经典的 Lagrange 乘子法是这样说的：为了解决上面的极值问题，我们应该考虑 $\mathbb{R}^n \times \mathbb{R}^m$ 上的函数

$$L(x_1, \cdots, x_n, \lambda_1, \cdots, \lambda_m) = f(x_1, \cdots, x_n) - \sum_{i=1}^{m} \lambda_i g_i(x_1, \cdots, x_n),$$

其中实变量 $\lambda_1, \cdots, \lambda_m$ 被称为 **Lagrange 乘子**。有些文献上将上面的函数称作是这个极值问题的 **Lagrange 函数**。Lagrange 乘子法的结论说的是 f 的约束条件极值问题等价于 L 的无条件极值问题，即我们要找 $(x_1, \cdots, x_n, \lambda_1, \cdots, \lambda_m) \in \mathbb{R}^n \times \mathbb{R}^m$ 来实现 L 的极值就可以了。

假设 $x \in M$ 是 f 在 M 上的一个局部极值，那么对任意的 $v \in T_xM$，我们有 $df(x)(v) = (\nabla_v f)(x) = 0$，即 $df\big(T_pM\big) \equiv 0$，这表明，$T_pM \subset \ker df(p)$。然而，$T_pM = \ker dg(x)$，所以，我们得到如下的结论

- $f\big|_M$ 在 $x \in M$ 处取局部极值的必要条件是

$$\ker df(x) \supset \ker dg(x) = \bigcap_{j=1}^{m} \ker dg_j(x).$$

这表明在 x 处，存在常数 $\lambda_1, \lambda_2, \cdots, \lambda_m$，使得

$$df(x) = \lambda_1 \cdot dg_1(x) + \lambda_2 \cdot dg_2(x) + \cdots + \lambda_m \cdot dg_m(x).$$

用矩阵的语言来写，对任意的 $i \leqslant n$，我们有

$$\frac{\partial f}{\partial x_i}(x_1, \cdots, x_n) - \sum_{k=1}^{m} \lambda_k \frac{\partial g_k}{\partial x_i}(x_1, \cdots, x_n) = 0.$$

换句话说，为了找到 f 在约束下的极值，一个必要条件就是找到 $(x_1, \cdots, x_n)$ 满足约束和 m 个实数 $\lambda_1, \cdots, \lambda_m$，使得上面的 n 个式子成立。为了记忆这个等式，我们可以考虑 Lagrange 函数

$$L(x_1, \cdots, x_n, \lambda_1, \cdots, \lambda_m) = f(x_1, \cdots, x_n) - \sum_{i=1}^{m} \lambda_i g_i(x_1, \cdots, x_n)$$

的微分，它对 x_i 偏导数恰好为上面的 n 个方程而对 λ_i 的偏导数恰好给出约束条件，所以，极值点的必要条件可以等价地写成：

$$dL(x, L) = 0.$$

这就是传统的 Lagrange 乘子法给出的求极值的必要条件。

我们可以给出 Lagrange 乘子法的几个经典的应用。

例子. 如果只有一个约束方程 $g : \mathbb{R}^n \to \mathbb{R}$，我们假设 $g(x) = 0$，我们要求 $f(x)$ 的最大值，此时，上述条件变成了

$$\frac{\partial f}{\partial x_i}(x_1, \cdots, x_n) - \lambda \frac{\partial g_k}{\partial x_i}(x_1, \cdots, x_n) = 0,$$

其中 $i = 1, 2 \cdots, n$。用梯度来表示，这等价于

- 在 $g^{-1}(0)$ 上找一个点，使得 $\nabla f(x)$ 与 $\nabla g(x)$ 共线。

我们令 $f(x) = d(x, z)^2 = \sqrt{(x_1 - z_1)^2 + \cdots + (x_n - z_n)^2}$，即我们要在曲面上 $g(x) = 0$ 上找一个点 x，使得该点到一个给定的点的距离是最短的。中学的经验告诉我们我们需要从 z 点到这个曲面做垂线，我们用 *Lagrange* 乘子法来解读这个问题。首先，我们计算 ∇f

$$\nabla f(x) = 2(x_1 - z_1, x_2 - z_2, \cdots, x_n - z_n).$$

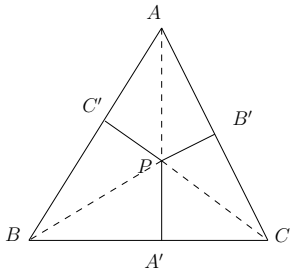
这是从 x 点到 z 的连线。我们已经证明过 $\nabla g(x)$ 是和曲面 $g^{-1}(0)$ 在这点处的切平面垂直的，所以 $\nabla f(x)$ 与 $\nabla g(x)$ 共线等价于说从 z 到 x 的连线和该曲垂直。

我们利用这个例子也可以看到，这一类问题的解可能不是唯一的，比如说如果 $g(x) = x_1^2 + \cdots + x_n^2 - 1$，这定义了单位球面。如果我们选取 $z = 0$，那么任意一个 x 都是这个极值问题的解。

例子 (一个初等几何问题)**.** 给定平面三角形 $\triangle ABC$（非退化），P 是三角形内部的点，从 P 点向三条边作垂线得到 A'，B' 和 C'，我们要找到这样的 P 使得三条线段长度的乘积

$$|PA'||PB'||PC'|$$

最大。



假设 $\triangle ABC$ 的三个高为 h_A，h_B 和 h_C，面积为 S 而 $\triangle PBC$，$\triangle PAC$ 和 $\triangle PAB$ 的面积分别为 u，v 和 w。通过简单的计算面积，我们知道这个问题等价于求如下函数的最大值：

$$f(u,v,w) = \frac{uvw}{h_A h_B h_C}, \quad u + v + w = S.$$

当然，我们要求 u, v, w 都是正数。根据 *Lagrange* 乘子法，我们要求

$$\nabla f - \lambda \nabla g = 0 \quad \Leftrightarrow \quad (vw, uw, uv) = \lambda(1,1,1).$$

这说明 $uv = vw = wz$，所以 $u = v = w$，这表明这个点到三边的距离是一样的，所以是三角形的内心。类似的，我们可以看到其实这里的推理已经可以用来证明算数－几何平均值不等式。

例子 (Hadamard 不等式)**.** 我们考虑这样的函数

$$f : \overbrace{\mathbb{R}^n \times \mathbb{R}^n \times \cdots \times \mathbb{R}^n}^{n\text{ 个}} \to \mathbb{R}, \quad (v_1, v_2, \cdots, v_n) \mapsto \det(v_1, \cdots, v_n).$$

上面最右边是把 n 个向量排成一排所组成的矩阵。我们在 $\mathbb{R}^n$ 上用标准的 *Euclid* 内积，假设 $|v_1| = |v_2| = \cdots = |v_n| = 1$，我们想求 f 的最大值。

我们观察到 f 是定义在 $\mathbb{R}^{n^2}$ 中的函数，其中，我们取坐标 $v_{ij} \in \mathbb{R}$，$1 \leqslant i, j \leqslant n$。我们现在有 n 个约束函数

$$g_i(v) = -1 + \sum_{j=1}^{n} v_{ij}^2.$$

很明显，g_i 们的公共零点定义了微分子流形，它实际上是 $\overbrace{\mathbf{S}^{n-1}\times\mathbf{S}^{n-1}\times\cdots\times\mathbf{S}^{n-1}}^{n\text{ 个}}$，这是一个紧集（有界闭集），所以连续函数 f 在它上面有最大值。我们现在用 *Lagrange* 乘子法，存在 $\lambda_1,\cdots,\lambda_n$，使得对于任意一个固定的 $1\leqslant i_0\leqslant n$，对任意的 $1\leqslant j\leqslant n$，我们有

$$\frac{\partial f}{\partial v_{i_0j}}-\lambda_{i_0}\frac{\partial g_{i_0}}{v_{i_0j}}=0\quad\Leftrightarrow\quad v_{i_0j}\text{ 的余子式与 }v_{i_0j}\text{ 成比例。}$$

如果我们用 v_{ij}^* 表示 v_{ij} 的余子式，那么（一个矩阵如果有两行一样的话它的行列式值就是 0），那么当 $i\neq i'$ 时，上面的比例关系表明

$$\sum_{j=1}^{n}v_{ij}v_{i'j}=0.$$

这表明这些 v_i 两两垂直。此时，矩阵 (v_{ij}) 是正交矩阵，所以 f 的最大值是 1 且等号成立当且仅当 $v_1,\cdots,v_n$ 两两垂直。据此，我们就得到一般情况下的 *Hadamard* 不等式：

命题 220 (Hadamard)**.** 对任意的非零向量 $v_1,v_2,\cdots,v_n\in\mathbb{R}^n$，我们有

$$|\det(v_1,v_2,\cdots,v_n)|\leqslant|v_1||v_2|\cdots|v_n|.$$

上面的不等式取等号当且仅当这些向量两两之间垂直。

子流形上的反函数定理和隐函数定理

我们还可以考虑子流形之间的反函数定理，其叙述和 $\mathbb{R}^n$ 上版本是一致的：

定理 221 (反函数定理)**.** 假设 $M\subset\mathbb{R}^m$ 和 $N\subset\mathbb{R}^n$ 是两个 d 维子流形，$f:M\to N$ 是光滑映射，$p\in M$，$f(p)=q$。假设

$$df(p):T_pM\to T_{f(p)}N$$

是可逆的，那么存在 $\mathbb{R}^m$ 中包含 p 的开邻域 $U\subset M$ 和 $\mathbb{R}^n$ 中包含 q 的开邻域 $V\subset M$，使得

$$f\big|_{U\cap M}:U\cap M\to V\cap N$$

是微分同胚，即 $f\big|_{U\cap M}$ 有逆并且也是光滑的。

证明： 我们把问题化到 $\mathbb{R}^d$ 的情况。首先，在 p 处取开集 U，微分同胚 $\Phi:U\to U'\subset\mathbb{R}^m$，使得 Φ 将 $U\cap M$ 映射为 $U'\cap\mathbb{R}^d\times\{0\}$；在 q 处取开集 V，微分同胚 $\Psi:V\to V'\subset\mathbb{R}^m$，使得 Ψ 将 $V\cap N$ 映射为 $V'\cap\mathbb{R}^d\times\{0\}$。我们考虑如下的交换图表：

$$\begin{array}{ccc}U'\cap(\mathbb{R}^d\times\{0\}) & \xrightarrow{\widehat{f}} & V'\cap(\mathbb{R}^d\times\{0\})\\ {\scriptstyle\Phi}\uparrow & & \uparrow{\scriptstyle\Psi}\\ M\cap U & \xrightarrow{\quad f\quad} & N\cap V\end{array}$$

其中，$\widehat{f}=\Psi\circ f\circ\Phi^{-1}$。

我们令 $\widehat{p}=\Phi(p)$，$\widehat{q}=\Psi(q)$，那么，$\widehat{f}(\widehat{p})=\widehat{q}$。根据链式法则，我们知道

$$d\widehat{f}(\widehat{p})=d\Psi(q)\circ df(p)\circ d\Phi^{-1}(\widehat{p})$$

是线性同构，所以我们可以在上述交换图表的第一层用反函数定理（此时，我们用 $\mathbb{R}^d$）中的反函数定理。然后用 Ψ 和 Φ 来复合就得到下面一层上的反函数定理，细节留给同学们自己来验证。 □

隐函数定理的几何形式更容易推广到子流形的情况：

命题 222 (子流形的原像)**.** *假设 $M\subset\mathbb{R}^m$ 和 $N\subset\mathbb{R}^n$ 是子流形，$f:M\to N$ 是光滑映射，$S\subset M$ 也是 $\mathbb{R}^m$ 的子流形。如果对任意的 $x\in f^{-1}(S)$，$\operatorname{rank} df(x)=\dim N$，那么 $f^{-1}(S)\subset\Omega$ 是子流形并且其位数满足等式*

$$\dim M-\dim f^{-1}(S)=\dim N-\dim S.$$

由于这个命题和后面课程的关系不大，我们在此略去它的证明。实际上，证明很简单，我们也可以仿照隐函数定理的方法化成某个 $\mathbb{R}^\ell$ 的情形。

37 多元函数 Hesse 矩阵，极值点的二阶导数判定，凸函数

二零二零年三月十二日，星期四，晴

Hesse 矩阵与极值点的二阶导数判定

为了避免张量的概念，我们考虑 $\mathbb{R}^n$ 中的开集 Ω 并假设 $\mathbb{R}^n$ 中已经选取坐标 $(x_1, \cdots, x_n)$。我们假设 f 是 Ω 上至少两次连续可微的函数。对于给定的点 $p \in \Omega$，我们定义 f 在 p 处的 Hesse 矩阵为

$$\nabla^2 f(p) = H_f(p) = \left(\frac{\partial^2 f}{\partial x_i \partial x_j}(p)\right) = \begin{pmatrix} \frac{\partial^2 f}{\partial x_1^2}(p) & \frac{\partial^2 f}{\partial x_1 \partial x_2}(p) & \cdots & \frac{\partial^2 f}{\partial x_1 \partial x_n}(p) \\ \frac{\partial^2 f}{\partial x_2 \partial x_1}(p) & \frac{\partial^2 f}{\partial x_1^2}(p) & \cdot & \frac{\partial^2 f}{\partial x_2 \partial x_n}(p) \\ \cdots & \cdot & \cdots & \cdots \\ \frac{\partial^2 f}{\partial x_n \partial x_1}(p) & \frac{\partial^2 f}{\partial x_n \partial x_2}(p) & \cdots & \frac{\partial^2 f}{\partial x_n^2}(p) \end{pmatrix}$$

根据 Clairaut-Schwarz 的定理，这是一个对阵矩阵。另外，根据线性代数中所学的知识，我们还可以将 Hesse 矩阵看成是 $\mathbb{R}^2$ 上的二次型：

$$H_f(p) : R^n \otimes \mathbb{R}^n \to \mathbb{R}, \quad (v, w) \mapsto \sum_{1 \leqslant i,j \leqslant n} \frac{\partial^2 f}{\partial x_i \partial x_j} v_i w_j.$$

我们假设在 $\mathbb{R}^n$ 给了内积 $\langle \cdot, \cdot \rangle$，使得 $\left\{\frac{\partial}{\partial x_i}\right\}_{i \leqslant n}$ 恰好是标准正交基。那么上面的二次型还可以写成

$$H_f(v, w) = \langle v, \nabla^2 f(w) \rangle.$$

另外，我们的 Taylor 公式在 2 阶的时候可以写成

$$f(x) = f(x_0) + \langle \nabla f(x_0), x - x_0 \rangle + \frac{1}{2} \langle x - x_0, \nabla^2 f(x_0)(x - x_0) \rangle + o(|x - x_0|^2).$$

对于实对称矩阵，我们可以讨论它是否正定／负定。我们简单回忆线性代数中的概念：矩阵／二次型 $\nabla^2 f$ 是正定（半正定）的，指的是它满足如下条件之一：

1) 对任意的 $v \in \mathbb{R}^n$，$H_f(v, v) > 0$（$H_f(v, v) \geqslant 0$）；

2) 矩阵 $\nabla^2 f$ 的特征值都是正（非负）的。

我们用 $\nabla^2 f > 0$ 和 $\nabla^2 f \geqslant 0$ 表示正定和半正定的二次型；负定的情形类似。

命题 223. *给定 $f \in C^2(\Omega)$，其中 $\Omega \subset \mathbb{R}^d$ 是开集，如果 $x_0 \in \Omega$ 是 f 的最小值点，那么，$\nabla f(x_0) = 0$ 并且 $\nabla^2 f(p_0) \geqslant 0$（半正定）。*

证明： 首先，$\nabla f(x_0) = 0$ 即 $df(x_0) = 0$，这是明显的。为了说明 $\nabla^2 f(x_0) \geqslant 0$，我们用反证法，假设 $\nabla^2 f$ 有一个负特征值 $\lambda < 0$，我们用 v 表示相应的一个特征向量（非零）。考虑 $x = x_0 + tv$ 并

利用 Taylor 公式（一阶导数项自动为零）：

$$\begin{aligned} f(x) &= f(x_0) + \frac{1}{2}\langle tv, \nabla^2 f(x_0)(tv)\rangle + o(|x-x_0|^2) \\ &= f(x_0) + \lambda \times \frac{|v|^2 t^2}{2} + o(t^2|v|^2) \end{aligned}$$

根据 $o(|t|^2)$ 的定义，存在 $\varepsilon > 0$，当 $|t| < \varepsilon$ 时，我们有 $o(t^2|v|^2) < \frac{1}{2}|\lambda|\frac{|v|^2t^2}{2}$。根据 $\lambda < 0$，从而

$$f(x) < f(x_0) + \lambda \times \frac{|v|^2t^2}{4} < f(x_0)$$

这与 x_0 是最小值矛盾。 □

这个证明可以用来证明一个接近于上述命题逆命题的命题：

命题 224. *给定 $f \in C^2(\Omega)$，其中 $\Omega \subset \mathbb{R}^d$ 是开集，如果 $x_0 \in \Omega$ 满足 $\nabla f(x_0) = 0$ 并且 $\nabla^2 f(p_0) > 0$（正定），那么，x_0 是 f 的局部最小值点，即存在开集 $U \subset \Omega$，使得 x_0 是 f 在 Ω 上的最小值。*

证明： 由于 $\nabla^2 f(p_0) > 0$，我们令 $0 < \lambda_1 \leqslant \lambda_2 \leqslant \cdots \leqslant \lambda_n$ 是它的特征值，其中 λ_1 是最小的特征向量。根据线性代数的知识，我们可以选取相应于上述特征值的特征向量 $v_1, v_2, \cdots, v_n$，其中它们都是单位向量并且两两正交。据此，对任意的 $w \in \mathbb{R}^n$，我们可以把 x 写成

$$w = a_1v_1 + \cdots + a_nv_n, \quad \sqrt{\sum_{i\leqslant n}(a_i)^2} = |x|.$$

由于 $\nabla^2 f(x_0)(w) = \lambda_1 a_1 v_1 + \cdots + \lambda_n a_n v_n$，所以，我们有

$$|\nabla^2 f(x_0)(w)| = \sqrt{\sum_{i\leqslant n}(\lambda_i a_i)^2} \geqslant \lambda_1 |w|.$$

根据 Taylor 公式（一阶导数项自动为零），我们有

$$\begin{aligned} f(x) &= f(x_0) + \frac{1}{2}\langle x-x_0, \nabla^2 f(x_0)(x-x_0)\rangle + o(|x-x_0|^2) \\ &\geqslant f(x_0) + \lambda_1|x-x_0|^2 + o(|x-x_0|^2) \end{aligned}$$

根据 $o(|x-x_0|^2)$ 的定义，存在 $\varepsilon > 0$，当 $|x-x_0| < \varepsilon$ 时（我们就选取 $U = B_\varepsilon(x_0)$），我们有 $o(|x-x_0|^2) < \frac{1}{2}|\lambda|\frac{|x-x_0|^2}{2}$。从而

$$f(x) > f(x_0) + \lambda \times \frac{|x-x_0|^2}{4} > f(x_0)$$

这说明 x_0 是 U 中的最小值点。 □

二阶导数与凸函数

给定凸集 $\Omega \subset \mathbb{R}^n$（即对任意的 $x, y \in \Omega$，对任意的 $t \in [0,1]$，$tx + (1-t)y \in \Omega$），$f : \Omega \to \mathbb{R}$ 是在凸集上定义的函数，如果对任意的 $x, y \in \Omega$，对任意的 $t \in [0,1]$，都有

$$f\big(tx + (1-t)y\big) \leqslant tf(x) + (1-t)f(y).$$

我们就称 f 是**凸函数**。（如果 $-f$ 是凸函数，就称 f 是**凹函数**）

另外，如果对任意的 $x, y \in \Omega$，对任意的 $t \in (0,1)$，都有

$$f\big(tx + (1-t)y\big) < tf(x) + (1-t)f(y).$$

我们就称 f 是**严格凸的**。

换而言之，对于任意 Ω 中的线段，f 在这个线段上的限制是 1 维的凸函数。另外，我们可以同样地证明 Jensen 不等式（请参见上学期第十八次课，证明完全可以照搬）：

命题 225 (Jensen 不等式)**.** *假设 $f : \Omega \to \mathbb{R}$ 是凸函数，其中 $\Omega \subset \mathbb{R}^n$ 为凸集，那么对任意的 $x_1, \cdots, x_n \in \Omega$ 和任意的 $t_1, \cdots, t_n \in [0,1]$，其中 $t_1 + t_2 + \cdots + t_n = 1$，我们有*

$$f(t_1x_1 + \cdots + t_nx_n) \leqslant t_1f(x_1) + \cdots + t_nf(x_n).$$

例子. *我们先看几个凸函数的例子（我们总假设 $\Omega \subset \mathbb{R}^n$ 是凸集）*

1) $f(x) = a_1x_1 + \cdots + a_nx_n + b$ *是 Ω 上的线性函数。*

2) 假设 $\|\cdot\| : \mathbb{R}^n \to \mathbb{R}$ 是任意一个范数，那么，这是一个凸函数，因为

$$\|tx + (1-t)y\| \leqslant \|tx\| + \|(1-t)y\| = t\|x\| + (1-t)\|y\|.$$

3) 假设 A 是一个 $n \times n$ 的半正定矩阵，那么，二次型

$$f(x) = \langle x, Ax\rangle$$

是凸函数。实际上，我们可以验证代数恒等式：

$$(1-t)f(x) + tf(y) - f((1-t)x + ty) = t(1-t)f(x-y).$$

右边的值是非负的。

定理 226 (凸函数在开集上的连续性)**.** *假设 $\Omega \subset \mathbb{R}^n$ 是凸的开集，$f : \Omega \to \mathbb{R}$ 是凸函数，那么，f 是连续函数。*

证明: 假设 $0 \in \Omega$，我们只要证明 f 在 0 处连续即可，其余的点类似。我们不妨假设如下的向量

$$\pm e_1 = (\pm 1, 0, \cdots, 0), \cdots, \pm e_n = (0, \cdots, 0, \pm 1)$$

都在 Ω 中（否则做适当的放缩即可，或者选取 $\varepsilon \pm e_i$ 来代替这些向量）。我们注意到，当 r 足够小的时，任意的 $|x| < r$，如果 x 在第一相限（即坐标都是非负的），存在 $\lambda_1, \cdots, \lambda_n \in [0, 1)$ 使得

$$x = \lambda_1 e_1 + \lambda_2 e_2 + \cdots + \lambda_n e_n.$$

如果 r 足够小，我们还可以要求 $\lambda_1 + \cdots + \lambda_n < 1$，这是因为

$$\lambda_1 + \cdots + \lambda_n \leqslant \sqrt{n(\lambda_1^2 + \cdots + \lambda_n^2)} = |x|\sqrt{n}.$$

选定这样一个 r。此时，我们令 $\lambda_0 = 1 - \lambda_1 \cdots - \lambda_n$，所以

$$x = \lambda_0 \cdot 0 + \lambda_1 e_1 + \lambda_2 e_2 + \cdots + \lambda_n e_n.$$

根据凸函数的性质（Jensen 不等式），我们就有

$$\begin{aligned} f(x) &\leqslant \lambda_0 f(0) + \lambda_1 f(e_1) + \lambda_2 f(e_2) + \cdots + \lambda_n f(e_n) \\ &= f(0) + \lambda_1 \left(f(e_1) - f(0)\right) + \cdots + \lambda_n \left(f(e_n) - f(0)\right). \end{aligned}$$

从而，

$$\begin{aligned} f(x) - f(0) &\leqslant \lambda_1 \left(f(e_1) - f(0)\right) + \cdots + \lambda_n \left(f(e_n) - f(0)\right) \\ &\leqslant \sqrt{\lambda_1^2 + \cdots + \lambda_n^2} \sqrt{|f(e_1) - f(0)|^2 + \cdots + |f(e_n) - f(0)|^2}. \end{aligned}$$

即（对其他相限类似可以证明）

$$f(x) - f(0) \leqslant C|x|, \quad \text{对任意的 } |x| < r.$$

我们还可以用 $x, -e_1, \cdots, -e_n$ 的凸组合来表示 0，实际上，利用 x 的坐标，我们选取

$$\lambda_0 = \frac{1}{x_1 + \cdots + x_n + 1}, \quad \lambda_i = \lambda_0 x_i, i \geqslant 1.$$

此时，我们有

$$0 = \lambda_0 x + \lambda_1(-e_1) + \lambda_2(-e_2) + \cdots + \lambda_n(-e_n).$$

所以，

$$\begin{aligned} f(0) &\leqslant \lambda_0 f(x) + \lambda_1 f(-e_1) + \lambda_2 f(-e_2) + \cdots + \lambda_n f(-e_n) \\ &= \lambda_0 \left(f(x) + x_1 f(-e_1) + x_1 f(-e_2) + \cdots + x_n f(-e_n)\right). \end{aligned}$$

从而，

$$f(0) - f(x) \leqslant (\lambda_0 - 1) f(x) + \underbrace{\lambda_0 \left(x_1 f(-e_1) + x_1 f(-e_2) + \cdots + x_n f(-e_n)\right)}_{\text{类似前述, } \leqslant C'|x|}.$$

对于第一项，我们有

$$\big|(\lambda_0-1)f(x)\big| \leqslant \frac{x_1+\cdots+x_2}{1+x_1+\cdots+x_n}|f(x)| \leqslant C''|x||f(x)|.$$

从而，

$$\begin{aligned} f(0)-f(x) &\leqslant C''|x||f(x)|+C'|x| \\ &\leqslant C''|x||f(x)-f(0)|+C''|x||f(0)|+C'|x|. \end{aligned}$$

综合之前的不等式，我们就得到

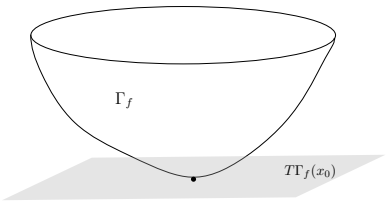
$$|f(x)-f(0)| \leqslant C''|x||f(x)-f(0)|+C''''|x|.$$

不妨假设选取的 $|x|$ 使得 $C''|x|<0.5$，从而，

$$|f(x)-f(0)| \leqslant 2C''''|x|.$$

这就说明了 $f(x)$ 是连续的（实际上是局部 Lipschitz 的）。 □

如果 f 具有一阶导数的话，我们可以用几何的方式（大约是切平面）来描述凸函数（请参考上学期第十八次课中凸函数的五个等价定义）：



命题 227. 假设 $\Omega\subset\mathbb{R}^n$ 是凸的开集，$f:\Omega\to\mathbb{R}$ 是可微函数（即微分 df 逐点存在），我们用 Γ_f 表示 f 在 $\Omega\times\mathbb{R}\subset\mathbb{R}^n\times\mathbb{R}$ 中的图像：

$$\Gamma_f=\big\{(x,y)\in\mathbb{R}^n\times\mathbb{R}\big|x\in\Omega,y=f(x)\big\}.$$

该图像在 $(x_0,f(x_0))$ 处的切平面我们定义为

$$T\Gamma_f(x_0)=\big\{(x_0,f(x_0))+df(x_0)(x-x_0)\big|x\in\mathbb{R}^n\big\}.$$

那么，如下命题是等价的：

1) f 是 Ω 上的凸函数；

2) 对任意的 $x_0\in\Omega$，函数图像 Γ_f 都在切平面 $T\Gamma_f(x_0)$，即

$$f(x)\geqslant L_{x_0}(x),$$

其中 $L_{x_0}(x)=f(x_0)+df(x_0)(x-x_0)$，$x\in\Omega$。

证明: 证明分两个方向:

1) 假设 f 是凸函数，所以对任意的 $x, y \in \Omega$，函数 $f(tx + (1-t)y)$ 是凸函数并且

$$f((1-t)x + ty) \leqslant (1-t)f(x) + tf(y) = f(x) + t(f(y) - f(x)).$$

另外，根据 Taylor 公式以及可微性，我们有

$$f((1-t)x + ty) = f(x + t(y-x)) = f(x) + tdf(x)(y-x) + o(t).$$

令 $t \to 0$，我们就得到

$$df(x)(y-x) \leqslant f(x).$$

令 $x = x_0$，$y = x$，整理就得到 2) 中的等式。

2) 反过来, 我们假设 2) 中不等式成立, 对于给定的 $x, y \in \Omega$, $t \in [0,1]$, 我们令 $z = (1-t)x + ty$。对 z 用不等式，我们有

$$f(w) \geqslant f(z) + df(z)(w-z).$$

令 $w = x$ 和 $w = y$，我们得到

$$\begin{aligned} f(x) &\geqslant f(z) - tdf(z)(y-x), \\ f(y) &\geqslant f(z) + (1-t)df(z)(y-x),. \end{aligned}$$

将第一个不等式乘 $(1-t)$，将第二个不等式乘 t，然后相加就得到所要的结论。

我们还可以将函数限制到线段上用 1 维结论直接说明等价性，这留给同学思考。 □

如果我们的函数是 2 阶可微的，类似于 1 维的情况，我们也有对凸函数的二阶导数的判定:

命题 228. *假设 $\Omega \subset \mathbb{R}^n$ 是凸的开集，$f : \Omega \to \mathbb{R}$ 是 C^2 的函数，那么，f 是凸函数当且仅当 f 在每个点处的 Hesse 矩阵都是半正定的。进一步，如果在每个点处 Hesse 矩阵都是正定的，那么 f 是严格凸的。*

证明: 首先假设 f 是凸函数。我们在 x_0 处将 f 进行 Taylor 展开:

$$\begin{aligned} f(x) &= \underbrace{f(x_0) + \langle \nabla f(x_0), x - x_0 \rangle}_{L_{x_0}(x-x_0)} + \frac{1}{2}\langle x - x_0, \nabla^2 f(x_0)(x-x_0)\rangle + o(|x-x_0|^2) \\ &\geqslant L_{x_0}(x - x_0). \end{aligned}$$

从而，对任意的 $x \to x_0$，我们有

$$\frac{1}{2}\langle x - x_0, \nabla^2 f(x_0)(x-x_0)\rangle + o(|x-x_0|^2) \geqslant 0.$$

令 $x - x_0 = tv$，$t = |x - x_0|$，其中 v 是任意的长为 1 的向量，从而

$$t^2 \frac{1}{2} \langle v, \nabla^2 f(x_0)(v) \rangle + o(t^2) \geqslant 0 \quad \Rightarrow \quad \frac{1}{2} \langle v, \nabla^2 f(x_0)(v) \rangle \geqslant o(1).$$

令 $t \to 0$，我们就得到了

$$\langle v, \nabla^2 f(x_0)(v) \rangle \geqslant 0,$$

其中 v 是任意的长为 1 的向量，所以 $\nabla^2 f(x_0)$ 是正定的。

反过来，我们假设在任何一点 $x_0 \in \Omega$ 处，我们有 $\nabla^2 f(x_0) \geqslant 0$。此时，根据 Talyor 展开公式（Lagrange 余项），我们有

$$f(x) - L_{x_0}(x - x_0) = \frac{1}{2} \langle x - x_0, \nabla^2 f(x`)(x - x_0) \rangle \geqslant 0,$$

其中 x' 位于 x_0 与 x 之间。根据前面一阶导数的命题，f 是凸的。

□

37.1 习题课：球极投影

清华大学 19-20 春季学期，数学分析二，第二次习题课内容

隐函数定理与子流形之间切映射的计算（球极投影）

隐函数定理的习题

第三次作业 B3)

假设 $f:\mathbb{R}^n\to\mathbb{R}^n$ 是光滑映射，0 是 f 的不动点，即 $f(0)=0$。假设 1 不是 f 在 0 处的微分

$$df(0):\underbrace{T_0\mathbb{R}^n}_{\mathbb{R}^n}\longrightarrow\underbrace{T_0\mathbb{R}^n}_{\mathbb{R}^n}$$

的特征值。

- 证明，0 是一个孤立的不动点，即存在开集 U，使得 $0\in U$，f 在 U 有且仅有 0 为其不动点。
- 任意给定 $g\in C^1(\mathbb{R}^n,\mathbb{R}^n)$。对任意的 $\lambda\in\mathbb{R}$，我们定义映射

  $$f_\lambda:\mathbb{R}^n\longrightarrow\mathbb{R}^n,\quad x\mapsto f(x)+\lambda g(x),$$

  证明，存在 $\varepsilon>0$，存在 0 在 $\mathbb{R}^n$ 的开邻域 U 以及 C^1 的映射

  $$\omega:(-\varepsilon,\varepsilon)\to V,\quad \lambda\mapsto\omega(\lambda),$$

  使得对任意的 $\lambda\in(-\varepsilon,\varepsilon)$，$f_\lambda$ 在 V 中恰好有一个不动点 $\omega(\lambda)$。

第三次作业 D2)

（根对系数的光滑依赖性）考虑实系数多项式 $P(X)=X^n+c_1X^{n-1}+\cdots+c_n=0$。如果对 $(c_1^*,\cdots,c_n^*)\in\mathbb{R}^n$, $P(X)$ 恰好有 n 个两两不同的实根 $z_1(c_1^*,\cdots,c_n^*)<\cdots<z_n(c_1^*,\cdots,c_n^*)$。证明，存在 $(c_1^*,\cdots,c_n^*)\in\mathbb{R}^n$ 的开邻域 Ω 和 Ω 上的光滑函数 $z_1,\cdots,z_n$，使得对任意的 $(c_1,\cdots,c_n)\in\Omega$，我们有

$$z_1(c_1,\cdots,c_n)<z_2(c_1,\cdots,c_n)<\cdots<z_n(c_1,\cdots,c_n)$$

并且

$$P(z_1(c_1,\cdots,c_n),z_2(c_1,\cdots,c_n),\cdots,z_n(c_1,\cdots,c_n))=0.$$

子流形

子流形的切空间的线性结构（参考第六课讲义的脚注）

假设 $M \subset \mathbb{R}^n$ 是子流形，我们想要说明 T_pM 是线性空间。根据定义，存在开集 $U \subset \mathbb{R}^n$，$p \in U$，存在开集 $V \subset \mathbb{R}^n$ 以及微分同胚，使得 $\Phi : U \to V$ 是微分同胚。所以，我们有

$$\Phi : M \cap U \to V \cap \mathbb{R}^d \times \{0\}.$$

对于 $V \cap \mathbb{R}^d \times \{0\} \subset \mathbb{R}^n$ 这种情况，我们课上已经计算了 $f(p)$ 处的切空间，这是一个 $T_{f(p)}\mathbb{R}^n$ 的 d 维线性子空间。另外，根据切空间的微分不变性，我们有

$$\mathbb{R}^d \times \{0\} = T_{\Phi(p)}\Big(V \cap \mathbb{R}^d \times \{0\}\Big) = T_{\Phi(p)}\Phi(U \cap M) = d\Phi\Big|_{x=p}(T_pM).$$

由于 $d\Phi(p) : T_p\mathbb{R}^n \to T_p\mathbb{R}^n$ 是线性同构，所以，T_pM 作为一个线性子空间的逆像还是线性子空间。

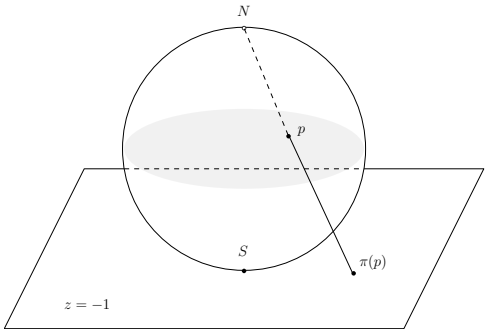
你是否能说明，T_pM 作为线性空间结构不依赖于开集 $U \subset \mathbb{R}^n$，$p \in U$，开集 $V \subset \mathbb{R}^n$ 以及微分同胚 $\Phi : U \to V$ 的选取?

球极投影（未完待续）

我们考虑 $\mathbb{R}^3$（用坐标 (x, y, z)）中的单位球面 $\mathbb{S}^2$，令 $N = (0, 0, 1)$ 为其北极，$S = (0, 0, 1)$ 为其南极。我们知道，$T_S\mathbf{S}^2$ 就是平面 $z = -1$，我们用 Σ 来表示这个平面。对任意的 $p \in S$，$p \neq N$，我们考虑从 N 到 p 的连线，它与 Σ 恰好相交于一个点 $\pi(p)$。于是，我们有映射

$$\pi : \mathbf{S}^2 - \{N\} \to \Sigma, \quad p \mapsto \pi(p).$$

这个映射被称作是从北极到南极切平面的**球极投影**。



- 证明，$\mathbb{R}^3$ 中的任意的平面（由某个线性函数 $ax + by + cz - d$ 的零点定义）与 $\mathbf{S}^2$ 上的交是圆（如果非空的话，半径可以是 0）。我们称它们为球面上的圆。
- 证明，π 是子流形之间的光滑映射并且给出了 $\mathbf{S}^2 - \{N\}$ 于 Σ 之间的微分同胚。
- 证明，对任意的 $p \in \mathbf{S}^2 - \{N\}$，试计算微分 $d\pi(p) : T_p\mathbf{S}^2 \to \Sigma$。

- 证明，这个映射是所谓的共形映射，即它保持角度：对于任意的 $v, w \in T_p\mathbf{S}^2$，我们有（用 $\mathbb{R}^3$ 中的内积）
$$\frac{\langle d\pi(p)(v), d\pi(p)(w)\rangle}{|d\pi(p)(v)||d\pi(p)(w)|} = \frac{\langle v, w\rangle}{|v||w|}.$$
- π 把 $\mathbf{S}^2$ 上不过北极的圆映成圆，把过北极的圆映射成直线并且 Σ 上的每个圆和直线都可以这么得到。
- 如果把 Σ 换成其他点 q 处的切空间，上面的结论是否成立?

37.2 作业：Lagrange 乘子法，Morse 引理，横截相交性

清华大学 19-20 春季学期，数学分析二，作业 4

请用 A4 大小的纸张**正反面**用钢笔，签字笔或者圆珠笔书写，并注明自己的姓名，年级和作业的**总页数**。除定理公式所涉及的人名之外，请使用**中文**。本次作业的提交时间和地点为 **2 月 50 日**上午的课堂上，逾期视作零分。

A. 课堂相关

A1) 对任意的 M 上的光滑函数 f，任意的 M 上的光滑切向量场 X 和 Y，我们可以定义它们之间的乘法和加法：

$$(fX)(p) = f(p)X(p),\quad (X+Y)(p) = X(p)+Y(p).$$

证明，fX 和 $X+Y$ 都是 M 上的光滑向量场。

A2) $\Omega \subset \mathbb{R}^n$ 为凸集，f 和 g 为 Ω 上的凸函数。试证明，

$$\max(f,g):\Omega\to\mathbb{R},\quad x\mapsto \max(f(x),g(x))$$

也是凸函数。

A3) $I\subset\mathbb{R}$ 是区间，$g:\mathbb{R}\to\mathbb{R}$ 是递增的凸函数，$f:\Omega\to I$ 也是凸函数。试证明，$g\circ f$ 是 U 上的凸函数。

A4)* 假设 $M\subset\mathbb{R}^m$，$N\subset\mathbb{R}^n$ 是子流形，$f:M\to N$ 是光滑映射。证明，$d:TM\to TN$ 是子流形之间的光滑映射。

（这是一个重要的命题，如果你搞清楚它的细节，那么你就搞清楚了整个子流形的理论）

A5) 在 $\mathbb{R}^3$ 中考虑方程

$$z^3+2z+e^{z-x-y^2}=\cos(x-y+z)$$

所定义的集合 Σ。证明，存在开集 $U\subset\mathbb{R}^2$ 使得 $0\in U$ 以及函数 $\varphi:U\to\mathbb{R}$，使得 $\Sigma\cap\{(x,y,z)|(x,y)\in U\}$ 是函数 $\varphi(x,y)$ 的图像 $\{(x,y,z)|z=\varphi(x,y),(x,y)\in U\}$。进一步计算 $\Sigma\cap\{(x,y,z)|(x,y)\in U\}$ 在 $(0,0,\varphi(0,0))$ 处的切空间并计算 φ 的 Hesse 矩阵。

B. Lagrange 乘子法

B1) 求下列函数的条件极值：

1. $f(x)=xy+yz+xz$, $xyz=1, x,y,z>0$。

2. $f(x)=\sin x\sin y\sin z$, $x+y+z=\pi/2, x,y,z>0$。

3. 二次型 $f(x)=\sum\limits_{i,j=1}^{n}a_{ij}x_ix_j$，其中 $a_{ij}=a_{ji}$，求 f 在单位球面 $\mathbf{S}^2$ 上的极值。

B2) 求椭球 $\dfrac{x^2}{a^2}+\dfrac{y^2}{b^2}+\dfrac{z^2}{c^2}=1$ 的最大体积的内接长方体。

B3) f 是 $\mathbb{R}$ 上的非负的连续函数并且如下的反常积分的值为 1：

$$\int_{-\infty}^{\infty}f(x)dx=1.$$

假设 $[a,b]$ 是使得 $\displaystyle\int_a^b f(x)dx=\frac{1}{2}$ 的长度最短的区间。证明，

$$f(a)=f(b).$$

B4) 假定 $n\in\mathbb{Z}_{\geqslant 1}$ 和 $a\in\mathbb{R}_{>0}$ 是给定的。考虑 n 元函数

$$f(x_1,\cdots,x_n)=\sum_{k\leqslant n}x_k\log x_k.$$

试求其最大值，其中，我们要求 $x_1+\cdots+x_n=a$。

B5) 给定 n 个正实数 $\alpha_1,\cdots,\alpha_n$，它们满足 $\alpha_1\alpha_2\cdots\alpha_n=1$。利用 Lagrange 乘子法证明：对任意的正实数 $x_1,\cdots,x_n$，我们都有

$$\prod_{k=1}^{n}x_k^{\alpha_k}\leqslant\sum_{k=1}^{n}\alpha_kx_k.$$

C. 实对称矩阵的对角化

我们在 $\mathbb{R}^n$ 上用坐标 $x_1,\cdots,x_n$ 并假设它上面配备了标准的 Euclid 内积，我们令 $e_i=\dfrac{\partial}{\partial x_i}$，它们构成了一个单位正交基。我们用 $\mathbb{S}^{n-1}$ 表示 $\mathbb{R}^n$ 中的单位球面。假设 $A=(A_{ij})_{1\leqslant i,j\leqslant n}$ 是给定的实对称矩阵。我们定义函数

$$f:\mathbb{R}^n-\{0\}\to,\quad x\mapsto f(x)=\frac{\langle A\cdot x,x\rangle}{|x|^2}.$$

C1) 证明，$\mathbf{S}^{n-1}$ 是 $\mathbb{R}^n$ 的一个紧的子流形。

C2) 证明，f 的最大值可以在 $\mathbf{S}^{n-1}$ 上取到。我们令 $v_1\in\mathbf{S}^{n-1}$ 为这样的一个最大值。

C3) 证明，$df(v_1):\mathbb{R}^n\to\mathbb{R}$ 可以如下表示：

$$df(v_1)(x)=2\langle A\cdot v_1,x\rangle-2\langle A\cdot v_1,v_1\rangle\langle v_1,x\rangle.$$

C4) 我们用 $v_1^\perp$ 表示和 v_1 垂直的向量构成的线性子空间。证明，对任意的 $x \in v_1^\perp$，我们有 $x \in \big(A \cdot v_1\big)^\perp$。

C5) 证明，$f(v_1)$ 是 A 的特征值并且 v_1 是相应的特征向量。

C6) 证明，A 有 n 个两两正交的（非零）特征向量 $v_1, \cdots, v_n \in \mathbb{R}^n$（提示：通过约化到 $v_1^\perp$ 上对 n 进行归纳）

C7) 证明，存在正交矩阵 P，使得 $P \cdot A \cdot P^{-1}$ 是对角矩阵。

D. Morse 引理

$\Omega \subset \mathbb{R}^n$ 是开集，$f \in C^\infty(\Omega)$，$\bar{x} \in \Omega$。如果 $df(\bar{x}) = 0$，我们就称 $\bar{x}$ 是 f 的一个**临界点**。如果 $\bar{x}$ 是 f 的临界点并且 Hesse 矩阵 $\nabla^2 f(\bar{x}) = \left(\dfrac{\partial^2 f}{\partial x_i \partial x_j}(\bar{x})\right)$ 是可逆的，我们就称 $\bar{x}$ 是一个**非退化的临界点**。在非退化临界点 $\bar{x}$ 附近 f 有 Taylor 展开

$$f(x) = f(\bar{x}) + \frac{\partial^2 f}{\partial x_i \partial x_j}(\bar{x})(x_i - \bar{x}_i)(x_j - \bar{x}_j) + o(|x - \bar{x}|^2).$$

这个练习的目的是证明 Morse 引理，粗略的说，我们可以换一下坐标系，使得 Hesse 矩阵是对角的。引理的精确叙述如下：

设 f 是开集 $\Omega \subset \mathbb{R}^n$ 上的光滑函数，$\bar{x}$ 是 f 的非退化临界点。那么，存在 $\bar{x}$ 的一个开邻域 U，$\mathbb{R}^n$ 原点的开邻域 V，以及微分同胚 $\varphi : V \to U$ 使得对每个 $(y_1, \cdots, y_n) \in V$，都有

$$(f \circ \varphi)(y) = f(\bar{x}) - \sum_{k=1}^{m} (y_k)^2 + \sum_{k=m+1}^{n} (y_k)^2,$$

*其中整数 m 称为临界点 $\bar{x}$ 的***指标***。*

D1) (Hadamard 引理)$\Omega \subset \mathbb{R}^n$ 是凸的开集，$0 \in \Omega$，函数 $f \in C^\infty(\Omega)$ 且 $f(0) = 0$。证明，存在光滑函数 $g_i \in C^\infty(\Omega)$，使得对任意的 $(x_1, \cdots, x_n) \in \Omega$，都有

$$f(x_1, \cdots, x_n) = \sum_{i=1}^{n} x_i g_i(x_1, \cdots, x_n)$$

并且 $g_i(0) = \dfrac{\partial f}{\partial x_i}(0)$。（提示：在射线 $t \mapsto (tx_1, \cdots, tx_n)$ 上对 f 用 Newton-Leibniz 法则）

D2)（条件同上）如果 $f(0) = 0$ 并且 0 是 f 的非退化临界点，证明，存在 0 的开邻域 $U \subset \Omega$ 以及 U 上 n^2 个光滑函数 h_{ij}（$1 \leqslant i, j \leqslant n$），它们满足

- 对任意的指标 i 和 j，$h_{ij} = h_{ji}$；
- 对任意的 $(x_1, \cdots, x_n) \in U$，矩阵 $\Big(h_{ij}(x_1, \cdots, x_n)\Big)$ 是可逆的；

– 对任意的 $(x_1,\cdots,x_n)\in U$，有

$$f(x_1,\cdots,x_n)=\sum_{i,j=1}^{n}x_ix_jh_{ij}(x_1,\cdots,x_n).$$

D3) 证明，存在 0 的邻域 $U_1\subset U$ 和 V_1（它的坐标用 $u_1,\cdots,u_n$），微分同胚 $\varphi_1:V_1\to U_1$，光滑函数 H_{ij}（$2\leqslant i,j\leqslant n$），使得

$$(f\cdot\phi_1)(u)=c_1(u_1)^2+\sum_{i,j=2}^{n}u_iu_jH_{ij}(u),\quad u=(u_1,\cdots,u_n)\in U_1.$$

其中常数 $c_1=1$ 或者 -1 并且 $H_{ij}=H_{ji}$。

D4)（用归纳法）对于 $r\leqslant n-1$，证明，存在 0 的邻域 $U_r\subset U$ 和 V_r（它的坐标用 $u_1,\cdots,u_n$），微分同胚 $\varphi_r:V_r\to U_r$，光滑函数 H_{ij}（$r+1\leqslant i,j\leqslant n$），使得

$$(f\cdot\phi_r)(u)=\sum_{k=1}^{r}c_k(u_k)^2+\sum_{i,j=r+1}^{n}u_iu_jH_{ij}(u),\quad u=(u_1,\cdots,u_n)\in U_r,$$

其中 $c_i\in\{\pm1\}$ 是常数并且 $H_{ij}=H_{ji}$。（这就证明了 Morse 引理）

D5)（非退化临界点是孤立的）证明，如果 $\bar{x}$ 是 f 的非退化临界点，那么存在 $\bar{x}$ 的开邻域 U，使得 $\bar{x}$ 是 f 在 U 上唯一的临界点。

D6) $\bar{x}$ 是 f 的非退化临界点。证明，指标 m 与局部坐标 φ 的选取无关并且等于 $\nabla^2f(\bar{x})$ 的负特征值的个数 (按重数计算)。（我们注意到，Hesse 矩阵本身是依赖于坐标选取的）

T. 横截相交性（transversality）（选作）

假设 $M_1,M_2\subset M\subset\mathbb{R}^n$ 均为是微分子流形，如果对所有的 $p\in M_1\cap M_2$，我们都有

$$T_pM_1+T_pM_2=T_pM,$$

我们就称 M_1 与 M_2 (作为 M 的子流形) **横截相交**（特别地，如果 $M_1\cap M_2=\emptyset$，它们也是横截相交），记作 $M_1\pitchfork M_2$。我们的目标是证明，如果 $M_1\pitchfork M_2$，那么 $M_1\cap M_2$ 是 M 的子流形。

T1) 试找出微分子流形 $M_1,M_2\subset\mathbb{R}^n$，使得 $M_1\cap M_2$ 不是子流形。

T2) 假设 L_1,L_2 是 $V=\mathbb{R}^n$ 的线性子空间并且 $L_1+L_2=V$。证明，$L_1\cap L_2$ 是 V 的线性子空间并且

$$\operatorname{codim}L_1\cap L_2=\operatorname{codim}L_1+\operatorname{codim}L_2.$$

T3) $C_1,C_2\subset\mathbb{R}^2$ 是光滑曲线。证明，$C_1\pitchfork C_2$ 的充要条件是 C_1,C_2 在它们的所有交点上都不相切（即它们在相交处的切空间不同）。

T4) $C_1, C_2 \subset \mathbb{R}^3$ 是光滑曲线。证明，$C_1 \pitchfork C_2$ 的充要条件是 $C_1 \cap C_2 = \emptyset$。

T5) 横截相交性的概念可以延拓到映射：假设 $M \subset \mathbb{R}^m$ 和 $N \subset \mathbb{R}^n$ 是微分子流形，$f : N \to M$ 是光滑映射，$S \subset M$ 是子流形。如果对所有的 $p \in f^{-1}(S)$，都有

$$df(T_pN) + T_{f(p)}S = T_{f(p)}M,$$

我们称光滑映射 f 与 S 是**横截相交的**并记作 $f \pitchfork S$。

微分子流形 $M_1, M_2 \subset M$，我们用 $i_1 : M_1 \subset M$ 表示包含映射。证明，$M_1 \pitchfork M_2$ 等价于 $i_1 \pitchfork M_2$ 。

T6*) 证明，如果 $M_1, M_2 \subset M \subset \mathbb{R}^n$ 是子流形并且横截相交，那么 $M_1 \cap M_2$ 是 M 的子流形并且

$$\mathrm{codim} M_1 + \mathrm{codim} M_2 = \mathrm{codim} M_1 \cap M_2.$$

（提示：在局部上用函数的零点集来定义子流形）

T7*) $M \subset \mathbb{R}^m$ 和 $N \subset \mathbb{R}^n$ 是微分子流形，$S \subset M$ 是子流形，$f : N \to M$ 是光滑映射并且 $f \pitchfork S$。证明，$f^{-1}(S)$ 是 N 的子流形并且

$$\dim N - \dim f^{-1}(S) = \dim M - \dim S.$$

(提示: 如果局部上 $S = g^{-1}(0)$，其中 $g : \mathbb{R}^m \to \mathbb{R}^{m-\dim S}$, 那么 $f^{-1}(S) = (g \circ f)^{-1}(0)$)

T8) 如果微分子流形 $M_1, M_2 \subset M$ 横截相交, 那么

$$T_p(M_1 \cap M_2) = T_pM_1 \cap T_pM_2.$$

T9) 如果 N 是紧的, $f \pitchfork S$, 证明: 对于任意的光滑映射

$$F : N \times [0, 1] \to M,$$

满足 $F(x, 0) = f(x)$ (F 称为 f 的一个光滑同伦), 存在 $\varepsilon > 0$, 使得对任意的 $t \in [0, \varepsilon)$, 都有 $f_t \pitchfork S$, 其中 $f_t(x) = F(x, t)$。

(这说明横截性在小形变下是稳定的)

T10) 证明: 设 $f : [0, 1] \to \mathbb{R}^3$ 是不自相交的光滑曲线, $S \subset \mathbb{R}^3$ 是光滑曲面。证明，对于 $\varepsilon > 0$, 存在 f 的光滑同伦 f_t, 使得

- $f_1 \pitchfork S$;
- $\sup\limits_{t,x\in[0,1]} |f_t(x) - f(x)| < \varepsilon$。

Analytical geometry has never existed. There are only people who do linear geometry badly, by taking coordinates, and they call this analytical geometry.

—— Jean Dieudonné

---

38 σ-代数，由某些子集生成的 σ-代数，Borel-代数，σ-代数的张量积，$\mathbb{R}^2$ 上的 Borel 代数与 $\mathbb{R}^1$ 上的 Borel 代数之间的关系，单调类，单调类 + 代数 $\Rightarrow\sigma$-代数，可测空间，可测映射，σ-代数的拉回，可测性的生成元判据，到乘积空间可测性的判据，距离空间的乘积空间上的 σ-代数，可测函数的性质

二零二零年三月十六日，星期一，晴

积分理论

Nature laughs at the difficulties of integration.

Pierre-Simon Laplace

我们引入一些集合论的常用记号：给定集合 X，我们用 $\mathcal{P}(X)$ 表示其幂集合（即所有子集所构成的集合）。给定 $A, B \in \mathcal{P}(X)$，我们用 A^c 表示 A 的补集并记 $A - B = A\backslash B = A \cap B^c$。

为了描述一下所要寻求的“积分”的大体轮廓，我们应该把上学期学过的 Riemann 积分理论作为原型，特别是用简单函数／阶梯函数的逼近的想法。大多数情况下，X 应该是 $\mathbb{R}^n$ 中的一个开集 Ω，我们首先要明确，我们并不是要对每个函数都定义它的积分（Riemann 积分理论中就有不可积分的函数），特别地，对于 X 中的某个子集 $A \subset X$，它的面积可能没有定义，也就是说示性函数 $\mathbf{1}_A$ 不可积分（比如说 $\mathbb{Q} \cap [0,1]$，根据 Lebesgue 定理，它的示性函数处处不连续，所以在 Riemann 的意义下不可积）。我们关心的是 X 中能定义面积的集合，它们的全体我们将用 $\mathcal{A}$ 来表示，一般而言，这是一个很大的集合，比如说，在 $\Omega \subset \mathbb{R}^n$ 上，我们希望这个集合包含所有的长方体以及所有可以用长方体铺出的集合（用这些长方体并出来）。所以，对于 $\mathcal{A}$ 中的元素（即 X）的子集，我们希望能够做一些基本的集合上的操作，比如并集等，这就是所谓的 σ-代数的结构。有了这些集合（对应于 Riemann 积分中的区间），我们就可以考虑它们所对应的示性函数的积分，这实际上要求对 $\mathcal{A}$ 中的每个元素定义它的“长度”或者“面积”（回忆上学期 Stieltjes 积分是非常有帮助的），这就是所谓的测度的概念。一旦有了测度，我们就可以对简单函数积分了，然后就可以对一切能被简单函数逼近的函数进行积分。用上学期学过的简单函数做逼近来定义 RIemann 积分的观点来看，经典意义上的 Riemann 积分也是这条路，只不过是用 Riemann 和的方式代替了简单函数逼近。我们要发展一套抽象的理论，它将囊括大部分可能的积分，比如说级数的求和与概率空间上的积分等。这个理论是在任意的集合上来构造的，从技术上而言要比 Riemann 积分更简单，从应用的角度而言会更广，从计算的角度而言它们没有太大的区别。我们也会在作业中展示和比较传统的 Riemann 积分理论和我们的理论。

σ-代数

先抽象地定义我们想定义面积的集合：

定义 229. 给定集合 X。$\mathcal{A} \subset \mathcal{P}(X)$ 是集合 X 的某些子集所构成的集合，如果它满足如下三条性质：

1) 空集 $\emptyset \in \mathcal{A}$；

2) 如果 $A \in \mathcal{A}$，那么 $A^c \in \mathcal{A}$；

3) 如果 $A_i \in \mathcal{A}$ $(i \in I)$，其中指标集 $i \in I$ 为有限集，那么 $\bigcup\limits_{i\in I} A_i \in \mathcal{A}$。

我们就称 $\mathcal{A}$ 是 X 上的一个**代数**。如果在上述条件 *3)* 中，允许 I 为可数集，那么称 $\mathcal{A}$ 为 X 上的一个 σ-**代数**。

换句话说，σ-代数在可数次并的操作下封闭。

注记. 对于 σ-代数 $\mathcal{A}$，我们很明显有 $X \in \mathcal{A}$ 以及如下性质：

4) 如果 $A_i \in \mathcal{A}$ $(i \in I)$，其中指标集 $i \in I$ 为可数集，那么 $\bigcap\limits_{i\in I} A_i \in \mathcal{A}$。

这因为并和交的操作在取补集的操作下是对偶的。

定义 230. 对于 X 上的 σ-代数 $\mathcal{A}$，如果其子集 $\mathcal{A}' \subset \mathcal{A}$ 也是 σ-代数，那么就称 $\mathcal{A}'$ 为 $\mathcal{A}$ 的**子 σ-代数**，或者成为 σ-**子代数**，也简称为子代数。

例子. 我们先给出三个接近于平凡的例子：

1) $\mathcal{A} = \mathcal{P}(X)$ 是 X 上的 σ-代数；

2) $\mathcal{A} = \{\emptyset, X\}$ 是 X 上的 σ-代数；

3) 假设 X 是可数集，对任意的 $x \in A$，$\{x\} \in \mathcal{A}$，那么 $\mathcal{A} = \mathcal{P}(X)$。

命题 231. 任意给定指标集合 J，如果对每个 $j \in J$，$\mathcal{A}_j$ 都是 X 上的 σ-代数，那么

$$\mathcal{A} := \bigcap_{j\in J} \mathcal{A}_j$$

也是 X 上的 σ-代数。

证明: 证明即为定义的验证：

1) 因为对任意的 j，$\emptyset \in \mathcal{A}_j$，所以 $\emptyset \in \mathcal{A}$。

2) 如果 $A \in \mathcal{A}$，那么，对任意的 $j \in J$，$A \in \mathcal{A}_j$，从而，$A^c \in \mathcal{A}_j$。这表明，$A^c \in \bigcap\limits_{j\in J} \mathcal{A}_j$，即 $A^c \in \mathcal{A}$。

3) 如果 $A_1, A_2, \cdots \in \mathcal{A}$（可数个），那么，对任意的 $j \in J$，$A_1, A_2, \cdots \in \mathcal{A}_j$，所以，$\bigcup_{i=1}^{\infty} A_i \in \mathcal{A}_j$。

从而，$\bigcup_{i=1}^{\infty} A_i \in \bigcup_{j\in J} \mathcal{A}_j = \mathcal{A}$。

□

注记（由某些子集生成的 σ-代数）. 根据这个命题，我们可以引入如下重要的概念：假定 $\mathcal{M} \subset \mathcal{P}(X)$ 为任意给定的子集（这是 X 中某些子集的集合），令

$$\Sigma = \{\mathcal{A} | \mathcal{A} \supset \mathcal{M},\ \mathcal{A} \text{ 为 } \sigma\text{-代数}\}.$$

很明显，Σ 不是空集，因为我们有 $\mathcal{P}(X) \in \Sigma(\mathcal{M})$。令

$$\sigma(\mathcal{M}) := \bigcap_{\mathcal{A}\in\Sigma} \mathcal{A}.$$

这是包含 $\mathcal{M}$ 的最小的 σ-代数，我们称它是**由 $\mathcal{M}$ 生成的 σ 代数**。

定义 232. 给定一个距离空间（拓扑空间）X，由 X 中一切开集所生成的 σ-代数被称作是 X 上的 **Borel-代数**，用符号 $\mathcal{B}(X)$。我们把 *Borel*-代数中的元素称作是 X 上的 **Borel-集**。

对多元微积分而言，最重要的对象是 $\mathcal{B}(\mathbb{R}^n)$，我们可以对着里面的集合定义面积／体积。如果不特别指出，我们都假定 $\mathbb{R}^n$ 上的距离就是标准的 Euclid 空间上的距离。首先研究一下 1 维的情况。按照定义，下面的性质是显然的：

注记. 给定 X 中的某些子集所组成的集合 $\mathcal{M}$ 和 $\mathcal{M}'$，如果 $\mathcal{M} \subset \mathcal{M}'$，那么 $\sigma(\mathcal{M}) \subset \sigma(\mathcal{M}')$。

按照定义，$\mathcal{B}(\mathbb{R}^1)$ 是有所有的开集生成的，实际上，它可以由更少的集合（可数个）生成：

命题 233. $\mathbb{R}$ 上的 *Borel*-代数可以由 $\{(-\infty, a) \big| a \in \mathbb{Q}\}$ 生成。

证明: 我们上学期证明过如下的命题：如果 U 是 $\mathbb{R}$ 上的开集，那么 U 可以写成可数个不相交的开区间的并集：

$$U = \bigcup_{k=1}^{\infty} (a_k, b_k),\quad \text{其中 } (a_k, b_k) \cap (a_{k'}, b_{k'}) = \emptyset,\ k \neq k'.$$

其中某个 a_k 可以是 $-\infty$，某个 $b_{k'}$ 可以是 $+\infty$。这表明 $\mathcal{B}(\mathbb{R}^1)$ 可以由所有的开区间生成。

我们现在说明 $\mathcal{B}\mathbb{R}^1$ 可以由 $\{(-\infty, a) \big| a \in \mathbb{R}\}$ 生成，为此，令 $\mathcal{A} = \sigma\left(\{(-\infty, a) \big| a \in \mathbb{R}\}\right)$。由于 σ-代数在开、并的可数操作以及取逆下是封闭的，所以，对任意的 $a, b \in \mathbb{R}$，$a < b$，我们有

$$[a, b) = (-\infty, b) \cap \big(\mathbb{R} - (-\infty, a)\big) \in \mathcal{A}.$$

从而，

$$(a, b) = \bigcup_{k\geqslant 1} [a + \frac{1}{k}, b) \in \mathcal{A}.$$

这表明，所有的开区间都落在 $\mathcal{A}$ 中，所以 $\mathcal{A}=\mathcal{B}(\mathbb{R}^1)$。

最终，为了说明 $\mathcal{B}(\mathbb{R}^1)=\sigma\left(\left\{(-\infty,a)\middle|a\in\mathbb{Q}\right\}\right)$，我们注意到对于任意的 $a\in\mathbb{R}$，我们可以选取递增的有理数列 q_k，使得 $q_k\to a$，所以

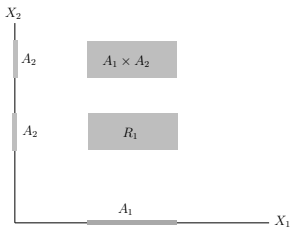
$$(-\infty,a)=\bigcup_{k\geqslant 1}(-\infty,q_k).$$

所以，$\mathcal{A}$ 中的生成元都落在 $\sigma\left(\left\{(-\infty,a)\middle|a\in\mathbb{Q}\right\}\right)$ 中，所以 $\mathcal{A}=\mathcal{B}(\mathbb{R}^1)$ 可以由这个集合生成。 □

在 $\mathbb{R}^n$ 上发展积分理论，我们要充分利用到 $\mathbb{R}^n$ 的定义（参见本学期第一次课程），它是更低维数的 $\mathbb{R}$ 通过乘积得到的。为此，我们在抽象的层次上研究两个集合的乘积上的 σ-代数：乘积空间上所对应的 σ-代数的张量积。对于指标 $i=1,2$，我们假定集合 X_i 和配备了 σ-代数 $\mathcal{A}_i$。在乘积空间 $X_1\times X_2$ 上，仿照平面上矩形的定义，我们优先考虑如下子集的集合：

$$\mathcal{R}:=\left\{A_1\times A_2\middle|A_1\in\mathcal{A}_1,A_2\in\mathcal{A}_2\right\}.$$

我们把上述集合中的的元素叫做**“矩形”**（字母 $\mathcal{R}$ 是 rectangle 的首字母）。



我们定义如下的集合

$$\widetilde{\mathcal{R}}:=\left\{A\subset X_1\times X_2\middle|A\text{ 为有限个两两不交的矩形的并}\right\}.$$

按照定义，每个 $\widetilde{\mathcal{R}}$ 中的元素 A 形如

$$A=\bigcup_{i\leqslant N}(A_i^{(1)}\times A_i^{(2)}),$$

其中，对任意的 $i,j\leqslant N$，$i\neq j$，$(A_i^{(1)}\times A_i^{(2)})\cap(A_j^{(1)}\times A_j^{(2)})=\emptyset$。

命题 234. *$\widetilde{\mathcal{R}}$ 是 $X_1\times X_2$ 上的代数。*

证明： 我们需要耐心地验证定义。我们约定 $X=X_1\times X_2$，选取 $\widetilde{\mathcal{R}}$ 中元素 $A=\bigcup\limits_{i\leqslant N}(A_i^{(1)}\times A_i^{(2)})$ 和 $B=\bigcup\limits_{j\leqslant M}(B_j^{(1)}\times B_j^{(2)})$。为了书写简洁，我们通常把它们写成

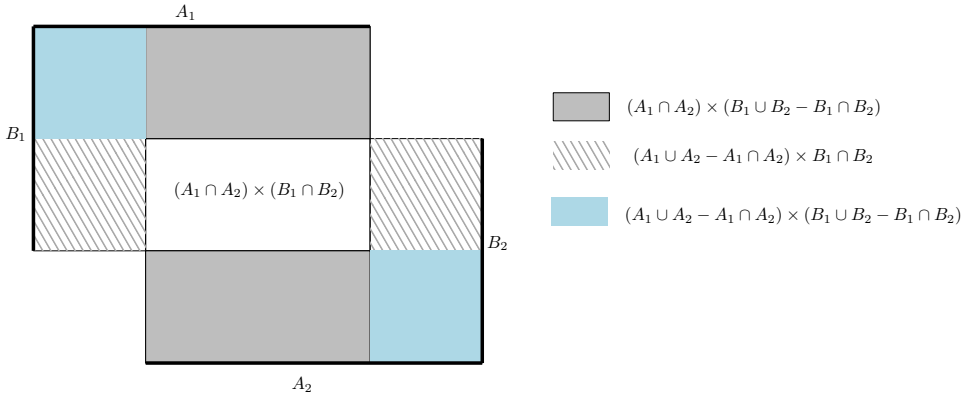
$$A=\bigcup(A_i^{(1)}\times A_i^{(2)}),\quad B=\bigcup(B_j^{(1)}\times B_j^{(2)}).$$

证明分三步：

1) $A\cup B\in\widetilde{\mathcal{R}}$。这因为

$$\begin{aligned}A\cup B&=\bigcup_{i,j}\big((A_i^{(1)}\times A_i^{(2)})\cup(B_j^{(1)}\times B_j^{(2)})\big)\\&=\bigcup_{i,j}\Big[\big((A_i^{(1)}\cap B_j^{(1)})\times(A_i^{(2)}\cap B_j^{(2)})\big)\cup\big((A_i^{(1)}\cap B_j^{(1)})\times(A_i^{(2)}-B_j^{(2)})\big)\\&\qquad\cup\big((A_i^{(1)}-B_j^{(1)})\times(A_i^{(2)}\cap B_j^{(2)})\big)\cup\big((A_i^{(1)}-B_j^{(1)})\times(A_i^{(2)}-B_j^{(2)})\big)\Big],\end{aligned}$$

上述出现的每个括号里的集合都是矩形，它们两两不交，所以 $A\cup B\in\widetilde{\mathcal{R}}$。下面的图给出了上述分解的示意图：



2) $A\cap B\in\widetilde{\mathcal{R}}$。

这也是显然的，因为

$$\begin{aligned}A\cap B&=\bigcup_{i,j}\big((A_i^{(1)}\times A_i^{(2)})\cap(B_j^{(1)}\times B_j^{(2)})\big)\\&=\bigcup_{i,j}\big((A_i^{(1)}\cap B_j^{(1)})\times(A_i^{(2)}\cap B_j^{(2)})\big),\end{aligned}$$

上述矩形很明显两两不交。

3) $A^c\in\widetilde{\mathcal{R}}$。

这因为

$$\begin{aligned}A^c&=\left(\bigcup_i(A_i^{(1)}\times A_i^{(2)})\right)^c=\bigcap_i(A_i^{(1)}\times A_i^{(2)})^c\\&=\bigcap\big(\underbrace{(A_i^{(1)c}\times A_i^{(2)c})\cup(A_i^{(1)c}\times A_i^{(2)})\cup(A_i^{(1)}\times A_i^{(2)c})}_{\text{在 }\widetilde{\mathcal{R}}\text{ 中}}\big),\end{aligned}$$

再利用刚得到的关于相交的性质即可。

至此，我们证明了代数的定义中所要求的三个条件，命题得证。 □

定义 235 (σ-代数的张量积)**.** 我们用 $\mathcal{A}_1 \otimes \mathcal{A}_2$ 表示由 $\mathcal{R}$ 生成的 σ-代数，即 $\mathcal{A}_1 \otimes \mathcal{A}_2 = \sigma(\mathcal{R})$。我们把它称作是$A_1$ 和 A_2 的张量积。

注记. $\mathcal{A}_1 \otimes \mathcal{A}_2$ 是乘积空间 $X_1 \times X_2$ 上的 σ-代数，它当然也可以由所有的矩形生成，即 $\mathcal{A}_1 \otimes \mathcal{A}_2 = \sigma(R)$。

作为例子，我们研究 $\mathbb{R}^2$ 上的 Borel 集的结构。

引理 236. $\mathbb{R}^2$ 中任一开集都可以写成可数个方块 $(a,b) \times (c,d)$ 的并（可能有交集），其中，我们可以要求 (a,b) 和 (c,d) 都是有限的区间。

证明: 假设 $\Omega \subset \mathbb{R}^2$ 是开集。我们可以把 $\mathbb{R}^2$ 写成可数个开球的并:

$$\mathbb{R}^2 = \bigcup_{n \geqslant 1} B_n(0).$$

所以，

$$\Omega = \bigcup_{n \geqslant 1} (\Omega \cap B_n(0)).$$

这是可数个有界开集的并。所以，只要对有界的开集证明我们的结论即可。在 $\mathbb{R}^2$，一个开方块 $(a,b) \times (c,d)$ 的坐标如果都是有理数的话，我们就称它是一个有理开方块，很显然，有理开方块的集合 $\mathcal{C}_{\mathbb{Q}}$ 是一个可数集。令 $F = \{C \in \mathcal{C}_{\mathbb{Q}} | C \subset \Omega\}$。由于 Ω 中的每个点都生活在某个小的（完全落在 Ω 中的）有理方块中，所以 F 中这些有理方块（至多可数个）的并集就是 Ω。 □

利用这个引理，我们可以刻画可以看出 $\mathcal{B}(\mathbb{R}^2)$ 与 $\mathcal{B}(\mathbb{R}^1)$ 之间的关系:

定理 237. $\mathbb{R}^2$ 上的 *Borel* 代数是 $\mathbb{R}^1$ 上的 *Borel* 代数与自身的张量积，即 $\mathcal{B}(\mathbb{R}^2) = \mathcal{B}(\mathbb{R}) \otimes \mathcal{B}(\mathbb{R})$。

证明: 我们先证明一个平凡的包含关系。由于 $\mathcal{B}(\mathbb{R}^2)$ 是由所有的开集生成，而根据上面的引理，开集是可数个方块 $(a,b) \times (c,d)$ 的并，所以，我们有

$$\mathcal{B}(\mathbb{R}^2) = \sigma\big(\{I \times J | I, J \subset \mathbb{R} \text{ 是开区间}\}\big).$$

另外，根据 $\mathcal{B}(\mathbb{R}) \otimes \mathcal{B}(\mathbb{R})$ 的定义，我们有

$$\mathcal{B}(\mathbb{R}) \otimes \mathcal{B}(\mathbb{R}) = \sigma\big(\{A \times B | A, B \subset \text{ 是 } \mathbb{R}^1 \text{ 上的 Borel 集}\}\big).$$

而任意开区间都是 $\mathbb{R}^1$ 上的 Borel 集，这就说明 $\mathcal{B}(\mathbb{R}^2) \subset \mathcal{B}(\mathbb{R}) \otimes \mathcal{B}(\mathbb{R})$。

为了说明反过来的包含关系，我们先证明如下的辅助命题：给定开集 $A_0 \subset R^1$，那么

- $\mathcal{B} = \{B \subset \mathbb{R} \big| A_0 \times B \in \mathcal{B}(\mathbb{R}^2)\}$ 是 $\mathbb{R}^1$ 上的 σ-代数。
  - 很明显 $A_0 \times \emptyset = \emptyset$，$A_0 \times \mathbb{R}$ 是 $\mathbb{R}^2$ 上的开集从而落在 $\mathcal{B}(\mathbb{R}^2)$ 中，所以 $\emptyset, \mathbb{R} \in \mathcal{B}$。

– 如果 $\{B_n\}_{n\geqslant 1}\subset\mathcal{B}$，按照定义，$\{A_0\times B_n\}_{n\geqslant 1}\subset\mathcal{B}(\mathbb{R}^2)$。由于 $\mathcal{B}(\mathbb{R}^2)$ 是 σ-代数，所以

$$\bigcup_{n\geqslant 1}(A_0\times B_n)=A_0\times\left(\bigcup_{n\geqslant 1}B_n\right)\in\mathcal{B}(\mathbb{R}^2).$$

按照 $\mathcal{B}$ 的定义，我们就有 $\bigcup\limits_{n\geqslant 1}B_n\in\mathcal{B}$。

– 如果 $B\in\mathcal{A}$，按照定义，$A_0\times B\in\mathcal{B}(\mathbb{R}^2)$。由于 $A_0\times\mathbb{R}\in\mathcal{B}(\mathbb{R}^2)$，所以

$$A_0\times B^c=A_0\times\mathbb{R}-A_0\times B\in\mathcal{B}(\mathbb{R}^2).$$

按照 $\mathcal{B}$ 的定义，我们就有 $B^c\in\mathcal{B}$。

至此，我们验证了 $\mathcal{B}$ 满足 σ-代数的定义

我们知道对任意的开集 $B\subset\mathbb{R}$，$A_0\times B\in\mathcal{B}(\mathbb{R}^2)$（因为 $A_0\times B$ 是 $\mathbb{R}^2$ 中的开集，请参考本次作业），所以 $\mathcal{B}$ 包含了所有的 $\mathbb{R}$ 中的开集，由于 $\mathcal{B}(\mathbb{R})$ 是包含开集的最小的 σ-代数，所以，$\mathcal{B}\supset\mathcal{B}(\mathbb{R}^1)$。

特别地，上面的证明表明，对任意的开集 $A\subset\mathbb{R}^1$，对任意的 $B\in\mathcal{B}(\mathbb{R}^1)$，我们都有 $A\times B\in\mathcal{B}(\mathbb{R}^2)$。现在固定一个Borel 集 B_0，我们再证明一个辅助命题：

- $\mathcal{A}=\{A\subset\mathbb{R}\big|A\times B_0\in\mathcal{B}(\mathbb{R}^2)\}$ 是 $\mathbb{R}^1$ 上的 σ-代数。

  – 很明显 $\emptyset\times B_0\in\mathcal{B}(\mathbb{R}^2)$。根据刚上面的结论，由于 $\mathbb{R}$ 是开集，所以 $\mathbb{R}\times B_0\in\mathcal{B}(\mathbb{R}^2)$ 中，所以 $\emptyset,\mathbb{R}\in\mathcal{A}$。

  – 如果 $\{A_n\}_{n\geqslant 1}\subset\mathcal{A}$，按照定义，$\{A_n\times B_0\}_{n\geqslant 1}\subset\mathcal{B}(\mathbb{R}^2)$。所以

$$\bigcup_{n\geqslant 1}(A_n\times B_0)=\times\left(\bigcup_{n\geqslant 1}A_n\right)\times B_0\in\mathcal{B}(\mathbb{R}^2).$$

按照 $\mathcal{A}$ 的定义，我们就有 $\bigcup\limits_{n\geqslant 1}A_n\in\mathcal{A}$。

  – 如果 $A\in\mathcal{A}$，按照定义，$A\times B_0\in\mathcal{B}(\mathbb{R}^2)$。由于 $\mathbb{R}\times B_0\in\mathcal{B}(\mathbb{R}^2)$，所以

$$A^c\times B_0=\mathbb{R}\times B_0-A\times B_0\in\mathcal{B}(\mathbb{R}^2).$$

按照 $\mathcal{A}$ 的定义，我们就有 $A^c\in\mathcal{B}$。

至此，我们验证了 $\mathcal{A}$ 满足 σ-代数的定义。

类似的，$\mathcal{A}$ 包含了所有的开集（因为对任意的开集 $A\subset\mathbb{R}^1$，对任意的 $B\in\mathcal{B}(\mathbb{R}^1)$，我们都有 $A\times B\in\mathcal{B}(\mathbb{R}^2)$），所以 $\mathcal{A}\supset\mathcal{B}(\mathbb{R}^1)$，从而，对任意的 Borel 集 A，我们都有 $A\times B_0\in\mathcal{B}(\mathbb{R}^2)$，当 B_0 也变动时，我们就证明了对任意的 A $B\in\mathcal{B}(\mathbb{R}^1)$，$A\times B\in\mathcal{B}(\mathbb{R}^2)$。这说明 $\mathcal{B}(\mathbb{R}^2)$ 包含了所有的矩形 $\mathcal{R}$，从而，$\mathcal{B}(\mathbb{R}^2)\supset\mathcal{B}(\mathbb{R})\otimes\mathcal{B}(\mathbb{R})$。 □

我们现在引入比 σ-代数略广泛的概念，这个概念在应用的时候非常有效，它可以帮助我们把大部分 σ-的东西（即可数的）转化为有限的。

所谓的 X 中**单调上升的**子集序列指的是

$$A_1 \subset A_2 \subset \cdots \subset A_n \subset \cdots .$$

我们令 $\lim\limits_{n\to\infty} A_n = \bigcup\limits_{n=1}^{\infty} A_n$。类似地，对于 X 中子集的序列

$$B_1 \supset B_2 \supset \cdots \supset B_n \supset \cdots$$

我们称它们是**单调下降的**，并记 $\lim\limits_{n\to\infty} B_n = \bigcap\limits_{n=1}^{\infty} B_n$。

定义 238. 给定集合 X，$\mathcal{M}$ 是 X 的某些子集所构成的集合。如果 $\mathcal{M}$ 中的每个单调上升或者下降的序列，其极限也在 $\mathcal{M}$ 中，我们就称 $\mathcal{M}$ 是 X 上的一个**单调类**。

注记. 根据定义，σ-代数是单调类。类似于 σ-代数的情形，我们很容易证明如下的命题（请参考作业）：假设对任意的 $j \in J$，$\mathcal{M}_j$ 都是 X 上的单调类，那么

$$\mathcal{M} := \bigcap_{j\in J} \mathcal{M}_j$$

也是 X 上的单调类。根据这个命题，我们可以定义由 X 的一些子集所生成的单调类：假设 $\mathcal{N}$ 是 X 中的一些子集所组成的集合，所有的包含 $\mathcal{N}$ 的单调类（至少包含 $\mathcal{P}(X)$）的交就是$\mathcal{N}$ **生成的单调类**。这是包含 $\mathcal{N}$ 的最小的单调类。

命题 239. 如果 X 上的代数 $\mathcal{A}$ 是单调类，那么 $\mathcal{A}$ 为 σ-代数。

我们指出，代数与 σ-代数的区别在于可数的并不一定是封闭的。

证明: 假设 $A_1, A_2, \cdots, A_n, \cdots \in \mathcal{A}$，我们要证明 $\bigcup\limits_{n\geqslant 1} A_n \in \mathcal{A}$。为此，我们定义

$$S_n = \bigcup_{k\leqslant n} A_k \in \mathcal{A}.$$

按照代数的定义，我们知道 $\{S_n\}_{n\geqslant 1} \subset \mathcal{A}$。这显然是单调上升的序列，所以，$\bigcup\limits_{n\geqslant 1} S_n \in \mathcal{A}$。按照定义，我们知道

$$\bigcup_{n\geqslant 1} A_n \quad \bigcup_{n\geqslant 1} S_n.$$

所以，$\bigcup\limits_{n\geqslant 1} A_n \in \mathcal{A}$。 □

下面的定理在理论构建上非常重要，它的证明也是非常有启发性的：

定理 240. *假设 $\mathcal{A}$ 为 X 上的代数，$\mathcal{M}$ 为 $\mathcal{A}$ 所生成单调类，那么我们有*

$$\mathcal{M} = \sigma(\mathcal{A}).$$

证明: 由于 σ-代数为单调类，按定义，我们有 $\mathcal{M} \subset \sigma(\mathcal{A})$，这是因为 $\mathcal{M}$ 是包含 $\mathcal{A}$ 的最小的单调类。另外一个包含方向的证明是不平凡的。根据上一个命题，我们只需要证明 $\mathcal{M}$ 为代数就可以了，因为此时 $\mathcal{M}$ 也是 σ-代数，它将包含 $\sigma(\mathcal{A})$。为此，对每个 $A \in \mathcal{P}(X)$，我们定义

$$\Phi_{\mathcal{M}}(A) := \big\{B \in \mathcal{P}(X) \big| A \cup B, A - B, B - A \in \mathcal{M}\big\}.$$

根据定义中 A 和 B 的对称性，我们有 $B \in \Phi_{\mathcal{M}}(A) \Leftrightarrow A \in \Phi_{\mathcal{M}}(B)$。

首先来说明集合 $\Phi_{\mathcal{M}}(A)$ 为单调类。为此，任意选取 $\{A_i\}_{i\geqslant 1}$ 和 $\{B_i\}_{i\geqslant 1}$，它们分别为 $\Phi_{\mathcal{M}}(A)$ 中单调上升和单调下降的序列，我们要证明这两个序列的极限仍然在 $\Phi_{\mathcal{M}}(A)$ 中:

1) $\lim\limits_{i\to\infty} A_i \in \Phi_{\mathcal{M}}(A)$。我们来验证定义:
   - 为了说明 $A \cup \lim\limits_{i\to\infty} A_i \in \mathcal{M}$，我们观察到
   $$A \cup \lim_{i\to\infty} A_i = A \cup \big(\bigcup_{i\geqslant 1} A_i\big) = \bigcup_{i\geqslant 1}(A \cup A_i).$$
   由于 $\{A \cup A_i\}_{i\geqslant 1}$ 为 $\mathcal{M}$ 中单调上升的序列而 $\mathcal{M}$ 为单调类，所以上式最后一项在 $\mathcal{M}$ 中，所以 $A \cup \lim\limits_{i\to\infty} A_i \in \mathcal{M}$。
   - 为了说明 $A - \lim\limits_{i\to\infty} A_i \in \mathcal{M}$，我们观察到
   $$A - \lim_{i\to\infty} A_i = A - \bigcup_{i\geqslant 1} A_i = \bigcap_{i\geqslant 1}(A - A_i).$$
   由于 $\{A - A_i\}_{i\geqslant 1}$ 为 $\mathcal{M}$ 中单调下降的序列，所以它的交也在 $\mathcal{M}$ 中。所以，$A - \lim\limits_{i\to\infty} A_i \in \mathcal{M}$。
   - 为了说明 $\lim\limits_{i\to\infty} A_i - A \in \mathcal{M}$，我们观察到
   $$\lim_{i\to\infty} A_i - A = \big(\bigcup_{i\geqslant 1} A_i\big) - A = \bigcup_{i\geqslant 1}(A_i - A).$$
   类似地，右边这一项也在 $\mathcal{M}$ 中。

2) $\lim\limits_{i\to\infty} B_i \in \Phi_{\mathcal{M}}(A)$。这里的证明和上面如出一辙:
   - 注意到 $A \cup \lim\limits_{i\to\infty} B_i = A \cup \big(\bigcap\limits_{i\geqslant 1} B_i\big) = \bigcap\limits_{i\geqslant 1}(A \cup B_i)$。由于 $\{A \cup B_i\}_{i\geqslant 1}$ 为 $\mathcal{M}$ 中单调下降的序列，它们的交在 $\mathcal{M}$ 中，从而 $A \cup \lim\limits_{i\to\infty} B_i \in \mathcal{M}$。
   - 我们有 $A - \lim\limits_{i\to\infty} B_i = A - \bigcap\limits_{i\geqslant 1} B_i = \bigcup\limits_{i\geqslant 1}(A - B_i) \in \mathcal{M}$，这因为 $\{A - B_i\}_{i\geqslant 1}$ 为 $\mathcal{M}$ 中单调上升的序列。

– 我们还有 $\lim\limits_{i\to\infty} B_i - A = \big(\bigcap\limits_{i\geqslant 1} B_i\big) - A = \bigcap\limits_{i\geqslant 1}(B_i - A) \in \mathcal{M}$。

综上所述，我们证明了 $\Phi_{\mathcal{M}}(A)$ 为单调类。

我们现在取 $A \in \mathcal{A}$，由于 $\mathcal{A}$ 为代数，所以 $\mathcal{A}$ 的元素 B 都满足 $\Phi_{\mathcal{M}}(A)$ 的定义中的要求，所以 $\mathcal{A} \subset \Phi_{\mathcal{M}}(A)$。进一步，根据 $\mathcal{M}$ 是包含 $\mathcal{A}$ 的最小的单调类，我们得到 $\mathcal{M} \subset \Phi_{\mathcal{M}}(A)$。也就是说，对每一个 $B \in \mathcal{M}$，我们有 $B \in \Phi_{\mathcal{M}}(A)$。根据对称性，我们也有 $A \in \Phi_{\mathcal{M}}(B)$。根据 A 的选取的任意性，我们知道 $\mathcal{A} \subset \Phi_{\mathcal{M}}(B)$，从而有 $\mathcal{M} \subset \Phi_{\mathcal{M}}(B)$，其中 B 可以是 $\mathcal{M}$ 中的任意元素。按照 $\Phi_{\mathcal{M}}(\cdot)$ 的定义，我们立即得到 $\mathcal{M}$ 中元素对于并和差的操作是封闭的，这就说明了 $\mathcal{M}$ 是一个代数。 □

可测空间与可测映射

定义 241. 给定一个集合 X 和它上面的一个 σ-代数 $\mathcal{A}$，我们将二元组 $(X,\mathcal{A})$ 称作是一个**可测空间**。

推论 242. 给定两个可测空间 $(X_1,\mathcal{A}_1)$ 和 $(X_2,\mathcal{A}_2)$，$X_1 \times X_2$ 上的 σ-代数 $\mathcal{A}_1 \otimes \mathcal{A}_2$ 是由 $\widetilde{\mathcal{R}}$ 所生成的单调类。

证明: 我们已经证明了 $\mathcal{A}_1 \otimes \mathcal{A}_2 = \sigma(\mathcal{R})$ 而 $\sigma(\mathcal{R})$ 是代数。 □

定义 243. $(X,\mathcal{A})$ 和 $(Y,\mathcal{B})$ 是两个可测空间，如果映射

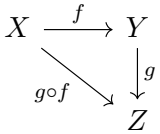
$$f : X \to Y, \quad x \mapsto f(x)$$

满足如下性质：

对每个 $B \in \mathcal{B}$，其逆像 $f^{-1}(B) \in \mathcal{A}$。(请比较拓扑空间之间的连续映射的定义)

那么，我们称 f 是这两个可测空间之间的**可测映射**。

注记. 可测映射的复合还是可测的：假设 $(X,\mathcal{A})$，$(Y,\mathcal{B})$ 和 $(Z,\mathcal{C})$ 是可测空间，$f : X \to Y$，$g : Y \to Z$ 是可测映射，那么对于 $C \in \mathcal{C}$，$g^{-1}(C) \in \mathcal{B}$，从而 $f^{-1}\left(g^{-1}(C)\right) \in \mathcal{A}$，即 $(g\circ f)^{-1}(C) \in \mathcal{A}$。



我们下面研究映射 $f : X \to Y$ 以及 σ-代数的函子性质。类似于映射和函数的拉回，我们可以定义 σ-代数的拉回：

定义 244. X 是集合，$(Y,\mathcal{B})$ 是可测空间，$f : X \to Y$。令

$$f^*\mathcal{B} = f^{-1}(\mathcal{B}) = \big\{f^{-1}(B) \subset X \big| B \in \mathcal{B}\big\}.$$

这是 X 上的 σ-代数，我们称它为 $\mathcal{B}$ 的**拉回**。

我们将在作业中证明 $f^*\mathcal{B}$ 的确是 σ-代数。根据定义，我们有如下两个显然的性质（我们沿用定义中的符号）：

1) 若 $\mathcal{B}' \subset \mathcal{B}$ 为子代数，那么 $f^*(\mathcal{B}')$ 也是 $f^*(\mathcal{B})$ 的子代数。

2) 若 $g : Z \to X$ 映射，那么 $(f \circ g)^*(\mathcal{B}) = g^*\big(f^*(\mathcal{B})\big)$。

我们现在证明，只要拉回生成元就可以生成拉回的 σ-代数了：

引理 245. *假定 $\mathcal{M} \subset \mathcal{P}(Y)$ 是 Y 中某些子集所组成的集合，$\mathcal{B} = \sigma(\mathcal{M})$ 是 $\mathcal{M}$ 生成的（Y 上的）σ-代数，$f : X \to Y$ 是映射。那么，$f^*(\mathcal{B})$ 由 $f^*(\mathcal{M})$ 生成，即*

$$f^*(\sigma(\mathcal{M})) = \sigma(f^*(\mathcal{M})).$$

证明： 由于 $\mathcal{M} \subset \sigma(\mathcal{M})$，所以 $f^*\mathcal{M} \subset f^*\sigma(\mathcal{M})$，从而 $\sigma(f^*(\mathcal{M})) \subset f^*(\sigma(\mathcal{M}))$。

另一方面，我们令

$$\mathcal{B}' = \big\{B \subset Y \big| f^{-1}(B) \in \sigma(f^{-1}(\mathcal{M}))\big\},$$

很明显，$\mathcal{M} \subset \mathcal{B}'$。

由于对任何可数个 Y 的子集 $\{B_i\}_{i\geqslant 1}$，我们都有

$$f^{-1}(\bigcup_{i\geqslant 1} B_i) = \bigcup_{i\geqslant 1} f^{-1}(B_i).$$

据此，我们很容易证明 $\mathcal{B}'$ 为 σ-代数（请参考本周的作业）。所以，$\sigma(\mathcal{M}) \subset \mathcal{B}'$。再根据 $\mathcal{B}'$ 的定义，有 $f^*(\mathcal{B}') \subset \sigma(f^*(\mathcal{M}))$，从而 $f^*(\sigma(\mathcal{M})) \subset \sigma(f^*(\mathcal{M}))$，这就完成了证明。 □

推论 246 (可测性的生成元判据)**.** *给定两个可测空间 $(X, \mathcal{A})$ 和 $(Y, \mathcal{B})$，假设 $\mathcal{B}$ 是由 $\mathcal{M} \subset \mathcal{P}(Y)$ 所生成的 σ-代数。那么，映射 $f : X \to Y$ 是可测的当且仅当 $f^*(\mathcal{M}) \subset \mathcal{A}$。*

证明： 这是因为如果 $f^*(\mathcal{M}) \subset \mathcal{A}$，那么 $\sigma(f^*(\mathcal{M})) \subset \mathcal{A}$。根据上面的命题，$f^*\big(\sigma(\mathcal{M})\big) \subset \mathcal{A}$，即 $f^*\mathcal{B} \subset A$。 □

推论 247. *$(X, \mathcal{A})$ 为可测空间，(Y, d) 是距离空间（拓扑空间），我们在 Y 上配备上 Borel 代数。那么，映射 $f : X \to Y$ 是可测的当且仅当对每个开集 $U \subset Y$，都有 $f^{-1}(U) \in \mathcal{A}$。特别地，如果 X 和 Y 均为距离空间（拓扑空间），它们上面都配备了 Borel 代数，f 为 X 与 Y 之间的连续映射，那么 f 是可测映射。*

证明： 这是显然的，因为 Borel 代数是由开集生成的。当 f 为连续映射时，开集的逆像是开集。 □

注记. *这个命题的是简单的，但是其重要性不言而喻：拓扑空间之间的连续映射一定是可测的。一般而言，连续性是非常容易验证的。特别地，我们现在有一大类可测映射的例子（连续映射）。*

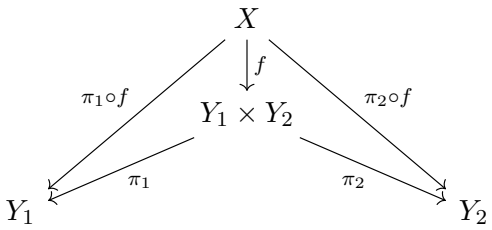
推论 248. 给定可测空间 $(Y_1, \mathcal{B}_1)$ 和 $(Y_2, \mathcal{B}_2)$，我们在乘积空间 $Y_1 \times Y_2$ 上配备 σ-代数 $\mathcal{B}_1 \otimes \mathcal{B}_2$。我们称 $(Y_1 \times Y_2, \mathcal{B}_1 \otimes \mathcal{B}_2)$ 是这两个**测度空间的乘积**。那么，自然的投影映射

$$\pi_1 : Y_1 \times Y_2 \to Y_1, \quad (y_1, y_2) \mapsto y_1,$$
$$\pi_2 : Y_1 \times Y_2 \to Y_2, \quad (y_1, y_2) \mapsto y_2,$$

是可测映射。

证明: 考虑 π_1，对任意的 $B_1 \in \mathcal{B}_1$，$\pi_1^{-1}(B_1) = B_1 \times Y_2$，这是 $Y_1 \times Y_2$ 上的" 矩形"，自然落在 $\mathcal{B}_1 \otimes \mathcal{B}_2$ 中。 □

推论 249 (到乘积空间可测性的判据)**.** 给定可测空间 $(Y_1, \mathcal{B}_1)$，$(Y_2, \mathcal{B}_2)$ 和 $(X, \mathcal{A})$。那么，映射 $f : X \to Y_1 \times Y_2$ 是可测的当且仅当每个 $f_i = \pi_i \circ f$ 均为可测的，其中 $i = 1, 2$。



证明: 如果 f 可测，那么复合映射 $f_i = \pi_i \circ f$ 自然可测；反过来对于 $Y_1 \times Y_2$ 矩形上的 $B_1 \times B_2$，其中 $B_i \in \mathcal{B}_i$，我们有

$$f^{-1}(B_1 \times B_2) = f_1^{-1}(B_1) \cap f_2^{-1}(B_2),$$

其中 $f_i = \pi_i \circ f$，$i = 1, 2$。由于 f_i 是可测的，所以 $f_i^{-1}(B_i) \in \mathcal{A}$，从而它们的交集也在 $\mathcal{A}$ 中，这就完成了证明。 □

在所谓的可分的距离空间上，要检测一个映射是否是可测的，我们可以只对开球进行检测。我们先回忆一下所谓的可分公理（定义）：

定义 250. 假设 (X, d) 是距离空间。如果 X 具有稠密的可数子集，即可以找到 $\{x_k\}_{k \geqslant 1} \subset X$，使得对任意的 $x \in X$，对任意的 $\varepsilon > 0$，存在某个 x_k，使得 $d(x, x_k) < \varepsilon$，那么我们就称 (X, d) 是**可分的**距离空间。

例子. 我们常见的几个距离空间都是可分的：

1) $\mathbb{R}^n$ 是可分的，因为所有的坐标为有理数的点所构成的集合是可数并且稠密的。

2) $C([0,1])$ 是可分的，其中，对于任意的 $f, g \in C([0,1])$，$d(f,g) = \|f - g\|_\infty = \sup_{x \in [0,1]} |f(x) - g(x)|$。事实上，根据 *Weierstrass-Stone* 的定理，所有的系数为有理数的多项式所构成的集合是可数并且稠密的。

可分距离空间中的开集可以用可数个开球并出来：

命题 251. (X,d) 是距离空间，对于 $x_0\in X$，$r>0$，我们令 $B_r(x_0)=\{x\in X|d(x,x_0)<r\}$。如果 X 是可分的，那么存在可数个开球 $\mathfrak{B}=\{B_{r_i}(x_i)\}_{i\geqslant 1}$，使得对任意开集 $U\subset X$，可以从 $\mathfrak{B}$ 中选取 $B_1,B_2,\cdots,B_m,\cdots$，使得

$$U=B_1\cup B_2\cup\cdots\cup B_m\cup\cdots.$$

证明: 根据可分性，我们在 X 中选取可数个点 $P=\{x_k\}_{k\geqslant 1}$，使得 P 在 X 中稠密。对于每个点 $x_k\in P$，选取可数个开球 $\{B(x_k,q)|q\in\mathbb{Q}\}$，我们把所有这样的小球放在一起组成了

$$\mathfrak{B}=\big\{B(x,q)\big|x\in P,q\in\mathbb{Q}_{>0}\big\}.$$

这是一个可数集（因为可数个可数集的并还是可数的）。

任取开集 U，我们定义（这里的想法与之前证明 $\mathbb{R}^2$ 上的开集都是可数个形如 $(a,b)\times(c,d)$ 的矩形的并是一样的）

$$\mathfrak{B}_U=\{B\in\mathfrak{B}|B\subset U\}.$$

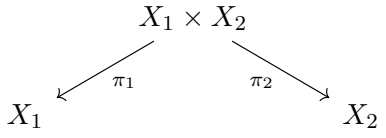
当然，$\mathfrak{B}_U$ 中只有可数个开球并且 $\bigcup\limits_{B\in\mathfrak{B}_U}B\subset U$。只需要证明 $U\subset\bigcup\limits_{B\in\mathfrak{B}_U}B$：任选 $x\in U$，由于 U 是开集，所以存在 $r>0$，使得 $B_r(x)\subset U$。根据 P 的稠密性，存在 $x_k\in P$，使得 $d(x_k,x)<\dfrac{r}{4}$，那么 $B(x_k,\dfrac{r}{2})\subset U$（从而属于 $\mathfrak{B}_U$）包含 x，这说明 $x\in\bigcup\limits_{B\in\mathfrak{B}_U}B$。 □

我们做如下的**约定**：从此往后，如果没有特别指出，每个距离空间 (X,d)（拓扑空间）都被视作是可测空间，我们默认它配有相应的 Borel-代数（由开集生成的 σ-代数）。为了方便起见，我们还把它记作 $(X,\mathcal{B}_X)$，其中 $\mathcal{B}_X$ 为 Borel-代数。

作为上面命题的推论，我们有

推论 252. 假设 $(X,\mathcal{A})$ 是可测空间，(Y,d) 是可分的距离空间。那么，映射 $f:X\to Y$ 是可测的当且仅当对每个 $B\subset\mathfrak{B}$（见上述命题的叙述），我们有 $f^{-1}(B)\in\mathcal{A}$。特别地，如果对于每个 (Y,d) 中的开球 B，我们都有 $f^{-1}(B)\in\mathcal{A}$，我们就可以断言 f 是可测的。

之前我们仔细研究了 $\mathbb{R}^2$ 上的 Borel 代数与 $\mathbb{R}^1$ 上的 Borel 代数之间的关系，这个命题可以推广到一般的可分距离空间上。我们首先回忆一下距离空间的乘积：



假设 (X_1,d_1) 和 (X_2,d_2) 是距离空间，我们在 $X_1\times X_2$ 上可以定义距离函数

$$d:\big(X_1\times X_2\big)\times\big(X_1\times X_2\big)\to\mathbb{R},\quad\big((x_1,x_2),(x_1',x_2')\big)\mapsto d\big((x_1,x_2),(x_1',x_2')\big).$$

我们通常选取

$$d\big((x_1,x_2),(x_1',x_2')\big)=\sqrt{d(x_1,x_1')^2+d(x_2,x_2')^2}.$$

（我们也可选取

$$d\big((x_1,x_2),(x_1',x_2')\big)=d(x_1,x_1')+d(x_2,x_2')$$

或者

$$d\big((x_1,x_2),(x_1',x_2')\big)=\sup\big\{d(x_1,x_1'),d(x_2,x_2')\big\}.$$

这些距离都是等价的）

另外，我们还有投影映射

$$\pi_1:X_1\times X_2\to X_1,\quad (x_1,x_2)\mapsto x_1,$$
$$\pi_2:X_1\times X_2\to X_2,\quad (x_1,x_2)\mapsto x_2.$$

定理 253. *给定距离空间 (X_1,d_1) 和 (Y_2,d_2)，我们用 $(X,d)=(X_1,d_1)\times(X_2,d_2)$ 表示它们的乘积距离空间。如果 X_1 和 X_2 是可分的，那么 X 也是可分的。进一步，X 上的 Borel 代数恰为 X_1 和 X_2 上 Borel 代数的张量积，集*

$$\mathcal{B}_Y=\mathcal{B}_{Y_1}\otimes\mathcal{B}_{Y_2},$$

（在测度空间的范畴里看距离空间，可分距离空间的乘积与测度空间的乘积是一致的）

证明: 假设 P_1 和 P_2 分别是 X_1 和 X 的可数稠密子集，那么 $P_1\times P_2$ 是 $X_1\times X_2$ 中的稠密子集：任选 $(x_1,x_2)\in X$，对任意的 $\varepsilon>0$，我们可以找到 $p_1\in P_1$，$p_2\in P_2$，使得

$$d_1(p_1,x_1)<\frac{\varepsilon}{\sqrt{2}},\quad d_1(p_2,x_2)<\frac{\varepsilon}{\sqrt{2}}.$$

所以，

$$d\big((p_1,p_2),(x_1,x_2)\big)<\sqrt{\frac{\varepsilon^2}{2}+\frac{\varepsilon^2}{2}}=\varepsilon.$$

我们现在来研究 Borel 代数。根据定义，我们有 $\mathcal{B}_X=\sigma(\{U\big|U\subset X_1\times X_2\text{ 为开集}\})$。实际上，我们可以做的更好

$$\mathcal{B}_X=\sigma(\{B_1\times B_2\big|B_1\in\mathfrak{B}_1,B_2\in\mathfrak{B}_2\}),$$

其中，$\mathfrak{B}_i$ 是 X_i 中的可数个开球的集合，使得 X_i 中的每个开集都可以表示成 $\mathfrak{B}_i$ 中若干个小球的并，这里 $i=1$ 或 2。我们把这个论断的证明留成作业题（重复命题251 的证明即可）。

根据乘积空间上矩形的定义，我们显然有

$$\mathcal{B}_X\subset\sigma(\mathcal{R})=\mathcal{B}_{X_1}\otimes\mathcal{B}_{X_2}.$$

我们现在证明上述包含关系为等式。为此，考虑恒同映射：

$$\iota:(X,\mathcal{B}_X)\to(X,\mathcal{B}_{X_1}\otimes\mathcal{B}_{X_2}),\quad x\mapsto x.$$

这是同一个集合上的映射，但是配备了不同的 σ-代数。对于每个 $i=1$ 或 2，对于任意 X_i 中的开集 U_i，我们有 $(\pi_i\circ\iota)^{-1}(U_i)=U_i\times X\in\mathcal{B}_X$。根据乘积空间可测性的判据，映射 ι 是可测的。特别的，我们有

$$\mathcal{B}_{X_1}\otimes\mathcal{B}_{X_2}=\iota^*(\mathcal{B}_{X_1}\otimes\mathcal{B}_{X_2})\subset\mathcal{B}_X.$$

这就证明了结论。 □

注记. 上述证明最后一部分表明，无论空间可分与否，$\mathcal{B}_X$ 要更细致（包含了更多的集合），即 $\mathcal{B}_{X_1} \otimes \mathcal{B}_{X_2} \subset \mathcal{B}_X$。另外，比较之前 $\mathcal{B}(\mathbb{R}^2) = \mathcal{B}(\mathbb{R}) \otimes \mathcal{B}(\mathbb{R})$ 的证明，我们看到测度空间的乘积结构使得证明简洁了很多。

可测函数的性质

对于复数域 $\mathbb{C}$ 或者实数域 $\mathbb{R}$，它们上面都有距离的结构，所以自然地配有 Borel-代数。我们称从可测空间 $(X, \mathcal{A})$ 到 $\mathcal{R}$ 或 $\mathcal{C}$ 的可测映射为**可测函数**。可测函数在代数操作和求极限操作下表现良好：

定理 254. 如果 $(X, \mathcal{A})$ 上的函数 f 和 g 是可测的，那么 $|f|, f \pm g$ 和 $f \cdot g$ 也是可测的。如果对任意的点 $x \in X$，$g(x) \neq 0$，那么 $\dfrac{f}{g}$ 是可测的。

证明： 根据映射到乘积空间的可测性判据，下面的映射

$$h : X \to \mathbb{C} \times \mathbb{C}, \quad x \mapsto (f(x), g(x)),$$

是可测的。我们将 h 与可测映射（因为它是连续的!）

$$\mathbb{C} \times \mathbb{C} \to \mathbb{C}, \quad (a, b) \mapsto a \pm b \text{ 或 } a \cdot b,$$

复合，就说明了 $f \pm g$ 和 $f \cdot g$ 是可测的。其余的情况我们留成本周的作业。 □

下一个定理说明可测函数列的极限函数也是可测的：

定理 255. $(X, \mathcal{A})$ 是可测空间，(Y, d) 是距离空间（通常是 $\mathbb{R}$ 或者 $\mathbb{C}$），给定可测映射的序列 $\{f_n\}_{n \geqslant 1}$，其中，对于每个 $n \geqslant 1$，函数 $f_n : X \to Y$。如果这个映射序列是逐点收敛的，即对每个 $x \in X$，都有

$$\lim_{n \to \infty} f_n(x) = f(x),$$

那么极限映射 $f(x)$ 也是可测的。

注记. 如果 $E \subset Y$ 是一个子集，我们可以定义如下的函数

$$d(\cdot, E) : Y \to \mathbb{R}, \quad x \mapsto d(x, E) = \inf_{e \in E} d(x, e).$$

这个函数衡量的是一个点到子集 E 的距离。

如果 F 是闭集，根据定义，F^c 是开集，那么，对于每个 $x \notin F$，$x \in F^c$，所以存在 $\varepsilon > 0$，使得 $B(x, \varepsilon) \subset F^c$，即 $B(x, \varepsilon) \cap F = \emptyset$，这表明，$d(x, F) \geqslant \varepsilon$。从而，对于闭集 F 而言，$x \notin F$ 等价于 $d(x, F) > 0$。

证明： 任取 Y 中的开集 U，我们定义 Y 中上升的（Borel 集）子集序列：

$$U_n = \left\{ x \in U \middle| d(x, U^c) > \frac{1}{n} \right\}.$$

很明显，每个 U_n 都是开集。由于 U^c 是闭集，所以

$$U = \lim_{n\to\infty} U_n = \bigcup_{n\geqslant 1} U_n.$$

另外，根据 $\lim\limits_{i\to\infty} f_i(x) = f(x)$，我们知道如果 $x \in f^{-1}(U_n) \Leftrightarrow f(x) \in U_n$（$U_n$ 为开集），那么存在 m，使得当 $q \geqslant m$ 时，$x \in f_q^{-1}(U_n) \Leftrightarrow f_q(x) \in U_n$，这表明

$$f^{-1}(U) = \bigcup_{n\geqslant 1} f^{-1}(U_n) = \bigcup_{n}\bigcup_{m}\bigcap_{q\geqslant m} f_q^{-1}(U_n).$$

由于每个 f_q 都是可测的，所以上面每个 $f_q^{-1}(U_n)$ 都是 $\mathcal{A}$ 中的元素，所以它们的可数的交和并得到的集合 $f^{-1}(U) \in \mathcal{A}$。这表明 f 是可测的。 □

39 测度，测度空间，σ-有限性，测度的基本性质，Carathéodory 扩张定理：如何利用代数定义外测度导出 σ-代数上的测度

二零二零年三月十九日，星期四，晴

约定. 从此往后，我们要用 $[0,\infty]$ 表示包含正无穷的非负实数的全体。

测度

我们要在一个可测空间 $(X,\mathcal{A})$ 上定义所谓的测度，也就是说，我们来测量 $\mathcal{A}$ 中的每个集合（成为可测集）的面积／体积。

定义 256. 假设 $\mathcal{A}$ 为 X 上的代数，如果映射 $\mu:\mathcal{A}\to[0,\infty]$ 满足

1) $\mu(\emptyset)=0$;

2) 对于 $\mathcal{A}$ 中不交的两个元素 A_1 和 A_2，我们有 $\mu(A_1\cup A_2)=\mu(A_1)+\mu(A_2)$，

那么称 μ 为代数 $\mathcal{A}$ 上的一个**加性函数**。

如果 $\mathcal{A}$ 为 X 上的 σ-代数，且映射 $\mu:\mathcal{A}\to[0,\infty]$ 满足

1) $\mu(\emptyset)=0$;

2) 对于 $\mathcal{A}$ 中不交的可数个元素 A_i，我们有

$$\mu(\bigcup_{i=1}^{\infty}A_i)=\sum_{i=1}^{\infty}\mu(A_i)$$

那么称 μ 为 σ-代数 $\mathcal{A}$ 上的一个（非负）**测度**。

一个可测空间 $(X,\mathcal{A})$ 如果配备了一个测度 μ，我们则称三元组 $(X,\mathcal{A},\mu)$ 为一个**测度空间**。

我们给出下面几个例子:

例子. *1)* X 是可数集，$\mathcal{A}=\mathcal{P}(X)$，对于任意的 $A\subset X$，我们定义 $\mu(A)=|A|$，即 A 的元素个数（可以是 ∞）。当 $A,B\subset X$ 并且 $A\cap B=\emptyset$ 时，我们自然有

$$\mu(A\cup B)=|A\cup B|=|A|+|B|=\mu(A)+\mu(B).$$

所以，$(X,\mathcal{P}(X),|\cdot|)$ 是测度空间。

2) (*Dirac* 测度) $(X,\mathcal{A})$ 是任意的可测空间，选定点 $x_0\in X$。对任意的 $A\subset\mathcal{A}$，我们定义:

$$\delta_{x_0}(A)=\begin{cases}1, & x_0\in A;\\ 0, & x_0\notin A.\end{cases}$$

当 $A, B \in \mathcal{A}$ 并且 $A \cap B = \emptyset$ 时，如果 $x_0 \in A \cup B$，不妨假设 $x_0 \in A$，那么

$$\delta_{x_0}(A \cup B) = 1 = 1 + 0 = \delta_{x_0}(A) + \delta_{x_0}(B).$$

如果 $x_0 \notin A \cup B$，那么

$$\delta_{x_0}(A \cup B) = 0 = 0 + 0 = \delta_{x_0}(A) + \delta_{x_0}(B).$$

这定义了一个测度，我们称它为 **Dirac 测度**。

另外，给出 X 中的可数个点 $\{x_k\}_{k\geqslant 1}$，给出正数的序列 $\{a_k\}_{k\geqslant 1}$，对任意的 $A \in \mathcal{A}$，我们定义

$$\left(\sum_{k=1}^{\infty} a_k \delta_{x_k}\right)(A) = \sum_{k=1}^{\infty} a_k \delta_{x_k}(A).$$

我们将在作业中证明这是一个测度。

3) 在 $\mathbb{R}$ 上，我们考虑所有的由有限个两两不交的区间的并所构成的代数 $\widetilde{\mathcal{R}}$。对于任意个的区间（无所谓开闭），我们令 $\mu([a,b]) = |b-a|$，我们将证明这可以定义出 $\widetilde{\mathcal{R}}$ 上的加性函数。从这个加性函数出发，我们将定义 *Lebesgue* 测度。

定义 257. 给定测度空间 $(X, \mathcal{A}, \mu)$，如果 $\mu(X) < \infty$，我们就称 μ 是**有限测度**或者是**有限的**。如果存在单调上升的序列 $\{A_i\}_{i\geqslant 1} \in \mathcal{A}$，使得 $\lim\limits_{i\to\infty} A_i = X$ 且对任意的 $i \geqslant 1$，$\mu(A_i) < \infty$，我们就说 μ 是σ-**有限的**。

很明显，上面给出的例子中，1) 中如果 X 是有限集，那么该测度是有限的，否则就是 σ-有限的；2) 中的 Dirac 测度是有限测度；3) 中我们还没有定义测度，但是我们将看到这会给出一个 σ-有限的测度并且这个测度是无限的。

对于一个测度空间 $(X, \mathcal{A}, \mu)$，我们有如下关于测度的基本不等式与等式：

命题 258 (熟记)**.** $(X, \mathcal{A}, \mu)$ 是测度空间，$\{A_i\}_{i\geqslant 1}$ 是给 $\mathcal{A}$ 中的序列，关于测度，我们有如下基本的性质：

1) 如果 $A_1 \subset A_2$，那么 $\mu(A_1) \leqslant \mu(A_2)$。

2) 我们有如下的等式：

$$\mu(\bigcup_{i\geqslant 1} A_i) \leqslant \sum_{i=1}^{\infty} \mu(A_i).$$

3) 如果序列 $\{A_i\}_{i\geqslant 1}$ 是上升的，那么

$$\mu(\lim_{i\to\infty} A_i) = \lim_{i\to\infty} \mu(A_i).$$

4) 如果序列 $\{A_i\}_{i\geqslant 1}$ 是下降的，并且对某个 i_0 有 $\mu(A_{i_0}) < \infty$，那么

$$\mu(\lim_{i\to\infty} A_i) = \lim_{i\to\infty} \mu(A_i).$$

证明：我们逐一的证明它们：

1) 如果 $A_1 \subset A_2$，$A_2 = A_1 \cup (A_2 - A_1)$，其中，$A_1$ 与 $A_2 - A_1$ 是不相交的，所以

$$\mu(A_2) = \mu(A_1) + \mu(A_2 - A_1) \geqslant \mu(A_1).$$

3) 序列 $\{A_i\}_{i\geqslant 1}$ 是上升的，我们定义 $A_0 = \emptyset$，$B_n = A_n - A_n - 1$，其中 $n \geqslant 1$。很明显，$\{B_n\}_{n\geqslant 1}$ 两两不相交，所以

$$\mu(\bigcup_{n\geqslant 1} B_n) = \sum_{n=1}^{\infty} \mu(B_n).$$

另外，$\lim\limits_{i\to\infty} A_i = \bigcup\limits_{i\geqslant 1} A_i = \bigcup\limits_{n\geqslant 1} B_n$，上面的等式就给出了要证明的等式。

2) 对于任意两个 $A, B \in \mathcal{A}$，我们有

$$\mu(A \cup B) = \mu(A \cup B - A) + \mu(A).$$

由于 $A \cup B - A \subset B$，所以

$$\mu(A \cup B) \leqslant \mu(B) + \mu(A).$$

通过归纳法，我们就得到

$$\mu(\bigcup_{i\leqslant n} A_i) \leqslant \sum_{i=1}^{n} \mu(A_i).$$

从而，对任意的 n，我们都有

$$\mu(\bigcup_{i\leqslant n} A_i) \leqslant \sum_{i=1}^{\infty} \mu(A_i).$$

现在我们利用 3) 的结论对 n 取极限即可。

我们把 4) 的证明留成作业，请注意，如果去掉某个 A_{i_0} 的测度是有限的条件，命题是不成立的。 □

测度构造的关键技术性定理：Carathéodory 扩张定理

一般而言，在一个“相对有限”的代数上构造测度要比在“相对无限”的 σ-代数上构造测度简单地多。比如说，在 $\mathbb{R}^n$ 上，我们知道 Borel 代数 $\mathcal{B}(\mathbb{R}^n)$ 可以由 $\widetilde{\mathcal{R}}$ 生成，其中，每个 $\widetilde{\mathcal{R}}$ 中的元素都是有限个形如 $(a_1, b_1) \times (a_2, b_2) \times \cdots \times (a_n, b_n)$ 的不相交矩形的并。不难验证，如果我们令

$$\mu\big((a_1, b_1) \times (a_2, b_2) \times \cdots \times (a_n, b_n)\big) = (b_1 - a_1)(b_2 - a_2) \cdot \cdots \cdot (b_n - a_n),$$

这给出了 $\widetilde{\mathcal{R}}$ 的一个加性函数（$\widetilde{R}$ 是一个代数）。然而，我们很想在 $\mathcal{B}(\mathbb{R}^n)$ 上定义测度，并且这个测度在 $\widetilde{\mathcal{R}}$ 与上面给的加性函数是一致的。这个就很困难，因为 $\mathcal{B}(\mathbb{R}^n)$ 的元素太多。我们下面给出一个抽象的定理以保证我们可以从代数上的加性函数（相对容易构造）来构造它所生成的 σ-代数上的测度，这就是 Carathéodory 测度扩张定理。

我们指出，尽管定理证明的本身非常值得研究（尤其是对于二年级大家学习实分析），但是在我们多元微积分这一部分完全可以略去而花更多的精力搞清楚定理的叙述和应用。

定理 259 (Carathéodory)**.** 假定 $\mathcal{A}$ 为 X 上的代数，μ 为 $\mathcal{A}$ 上 σ-有限加性函数（即存在单调上升的序列 $\{X_i\}_{i\geqslant 0}\subset\mathcal{A}$，使得 $\bigcup\limits_{i\to\infty}X_i=X$ 并且对每个 $i\geqslant 1$，都有 $\mu(X_i)<\infty$），那么至多存在一个 $\sigma(\mathcal{A})$ 上的测度 μ' 使得 $\mu'\big|_{\mathcal{A}}=\mu$，即对任意的 $A\subset\mathcal{A}$，$\mu'(A)=\mu(A)$。

如果 μ 是有限加性函数，即 $\mu(X)<\infty$，那么下述条件保证了延拓测度 μ' 的存在性：

条件 (C). 对任意单调下降的序列 $\{A_i\}_{i\geqslant 0}\subset\mathcal{A}$，如果 $\mu(A_0)<\infty$ 并且 $\lim\limits_{i\to\infty}A_i=\emptyset$，那么 $\lim\limits_{i\to\infty}\mu(A_i)\to 0$。

如果 $\mu(X)=+\infty$，还需要再加上如下条件：

条件 ($\mathbf{C}_\infty$). 存在单调上升的序列 $\{X_i\}_{i\geqslant 1}\subset\mathcal{A}$，$\lim\limits_{i\to\infty}X_i=X$ 并对任意的 $i\geqslant 1$ 都有 $\mu(X_i)<\infty$，满足对任意的 $A\in\mathcal{A}$，$\mu(A)=+\infty$，我们都有 $\lim\limits_{i\to\infty}\mu(A\cap X_i)=+\infty$。

我们先证明扩张的唯一性部分，这是单调类的一个好的应用示范：

假定 μ_1 和 μ_2 都是 μ 在 $\sigma(\mathcal{A})$ 上的延拓。我们定义如下的集合：

$$\mathcal{M}:=\big\{M\in\sigma(\mathcal{A})\big|\text{任意}A\in\mathcal{A},\mu(A)<\infty,\text{ 均有 }\mu_1(A\cap M)=\mu_2(A\cap M)\big\}.$$

我们显然有 $\mathcal{A}\subset\mathcal{M}$。我们可以和容易地验证 $\mathcal{M}$ 为单调类：假设 $\{M_j\}_{j\geqslant 0}\subset\mathcal{M}$ 是单调上升的序列，令 $M=\bigcup\limits_{j\geqslant 1}M_j$。根据 $\mathcal{M}$ 的定义，对任意的 $A\in\mathcal{A}$，$\mu(A)<\infty$，对任意的 $j\geqslant 1$，我们有

$$\mu_1(A\cap M_j)=\mu_2(A\cap M_j).$$

根据测度的基本性质，令 $j\to\infty$，上式的极限为

$$\mu_1(A\cap M)=\mu_2(A\cap M).$$

根据 $\mathcal{M}$ 的定义，$M\in\mathcal{M}$，所以 $\mathcal{M}$ 为单调类。由于 $\mathcal{M}\supset\mathcal{A}$，所以它包含 $\mathcal{A}$ 所生成的单调类，这是一个 σ 代数，从而 $\mathcal{A}$ 包含 $\sigma(A)$，这表明 $\mathcal{M}=\sigma(A)$。

我们现在来说明 $\mu_1=\mu_2$，根据 $\mu(X)$ 是否有限，我们有

- 如果 $\mu(X)$ 有限，那么在 $\mathcal{M}$ 的定义中取 $A=X$，所以对任意的 $M\in\mathcal{M}=\sigma(A)$，$\mu_1(M)=\mu_2(M)$，这表明 $\mu_1=\mu_2$。
- 如果 $\mu(X)$ 无限，根据 μ 的 σ-有限性，我们选取单调上升的序列 $\{X_i\}_{i\geqslant 1}\subset\mathcal{A}$，其中对任意的 i，$\mu(X_i)<\infty$ 并且 $\bigcup\limits_{i\geqslant 1}X_i=X$。此时，根据 $\mathcal{M}$ 的定义，对每个 i，对每个 $M\in\mathcal{M}$，我们都有
$$\mu_1(X_i\cap M)=\mu_2(X_i\cap M).$$
对 i 取极限，根据测度的性质，我们得到 $\mu_1(M)=\mu_2(M)$，这表明 $\mu_1=\mu_2$。

Carathéodory 定理的存在性部分的证明较长，特别地，证明将涉及的一系列概念本身就很有意义（它们不会在课程后面出现）。我们将证明细分七个步骤：

(第一步) 条件 **(C)** 和 $\mathbf{(C_\infty)}$ 的重新表述（这两个条件可以推出如下条件）：这两个条件强调可以“有限”到“可数”来取极限：

条件 **(D)**：假定 $\{A_i\}_{i\geqslant 0}\subset\mathcal{A}$ 是两两不交的子集并且 $\bigcup\limits_{i\geqslant 1}A_i\in\mathcal{A}$，那么

$$\mu\left(\bigcup_{i\geqslant 1}A_i\right)=\sum_{i\geqslant 1}\mu(A_i).$$

证明很简单：令 $A=\bigcup\limits_{i\geqslant 1}A_i$。

如果 A 测度有限，我们定义单调下降的序列 $\{C_n\}_{n\geqslant 1}\subset\mathcal{A}$，其中，对每个 $n\geqslant 1$，我们有

$$C_n=A-\bigcup_{i\leqslant n}A_i.$$

很明显，$\lim\limits_{i\to\infty}C_i=\emptyset$。根据条件 **(C)**，我们有

$$\lim_{n\to\infty}\mu(C_n)=0\quad\Leftrightarrow\lim_{i\to\infty}\mu(A-\bigcup_{i\leqslant n}A_i)=\mu(A)-\lim_{i\to\infty}\mu(\bigcup_{i\leqslant n}A_i)=0.$$

这显然给出了条件 **(D)**。

如果 $\mu(A)=\infty$，那先选取条件 $\mathbf{(C_\infty)}$ 中的一个 X_j（暂时固定指标 j），和上面情况一样，我们有

$$\mu(A\cap X_j)=\sum_{i\geqslant 1}\mu(A_i\cap X_j)\leqslant\sum_{i\geqslant 1}\mu(A_i).$$

条件 $\mathbf{(C_\infty)}$ 说的是当 $j\to\infty$ 时候，左边的值趋于无穷，从而 $\sum\limits_{i\geqslant 1}\mu(A_i)=\infty$，这就是条件 **(D)**。

(第二步) 外测度的定义。

对于 $E\subset X$，其中 E 不一定要在 $\mathcal{A}$ 中或者 $\sigma(\mathcal{A})$ 中，我们定义（由 μ 决定的）E **的外测度** $\mu^*(E)$：

$$\mu^*(E)=\inf_{\substack{A_i\in\mathcal{A},\\ E\subset\bigcup A_i}}\left(\sum_{i=1}^{\infty}\mu(A_i)\right).$$

在上述公式中，指标集 i 是可数的。如果我们选取

$$A_k'=A_k-\bigcup_{i\leqslant k-1}A_i$$

我们就可以假设这些 $\{A_i\}_{i\geqslant 1}$ 两两不交。我们强调，上述定义中的 A_i 个数的可数性**不能**换成有限性：

例子. 我们可以考虑 $[0,1]$ 上所有区间所生成的代数，在这个代数上我们定义加性函数为区间长度。如果我们只用有限个区间来盖住 $\mathbb{Q}\cap[0,1]$，那么这些区间一定会覆盖 $[0,1]$，从而，$\mathbb{Q}\cap[0,1]$ 外测度为至少为 1，而我们上个学期在学习 *Lebesgue* 定理的时候看到过，它的外测度"最好"是 0。

外测度满足三个基本性质：

1) 如果 $E\subset F$，那么有 $\mu^*(E)\leqslant\mu^*(F)$。

   这是显然的，因为任何盖住了 F 的 $\{A_i\}_{i\geqslant 1}$ 必然盖住了 E。

2) 对于任意可数个 $\{E_i\}_{i\geqslant 1}$，我们有 $\mu^*(\bigcup\limits_{i\geqslant} E_i)\leqslant\sum\limits_{i=1}^{\infty}\mu^*(E_i)$。

   不妨假设每个 E_i 的外测度是有限的（否则没什么需要证明的）。对任意的 $\varepsilon>0$，按照定义，我们选取 $\{A_{i,j}\}_{j\geqslant 1}\subset\mathcal{A}$，使得

$$0\leqslant\sum_{j=1}^{\infty}\mu(A_{i,j})-\mu^*(E_i)<2^{-i}\varepsilon.$$

   那么，$\{A_{i,j}\}_{i,j\geqslant 1}\subset\mathcal{A}$ 是 $\bigcup\limits_{i\geqslant} E_i$ 的覆盖并且

$$0\leqslant\sum_{i,j=1}^{\infty}\mu(A_{i,j})-\sum_{i\geqslant 1}\mu^*(E_i)<\sum_{i\geqslant 1}2^{-i}\varepsilon=\varepsilon.$$

   从而，按照外测度定义，我们有

$$0\leqslant\mu\left(\bigcup_{i\geqslant} E_i\right)-\sum_{i\geqslant 1}\mu^*(E_i)<\sum_{i\geqslant 1}2^{-i}\varepsilon=\varepsilon.$$

   令 $\varepsilon\to 0$ 我们就得到了结论。

3) 如果 $A\in\mathcal{A}$，那么 $\mu^*(A)=\mu(A)$。如果我们选 A 的覆盖为 $\{A\}$，这表明 $\mu^*(A)\leqslant\mu(A)$；任意选取 A 的覆盖 $\{A_i\}_{i\geqslant 1}\subset\mathcal{A}$，我们假设这个覆盖两两不相交。那么，

$$\sum_{i\geqslant 1}\mu(A_i)\geqslant\sum_{i\geqslant 1}\mu(A_i\cap A)=\mu(A).$$

   最后一个等号我们用了条件 **(D)**。这表明，$\mu^*(A)\geqslant\mu(A)$。

(第三步) μ^*-可测的定义。

给定 X 的子集 B，如果对所有的 $E\subset X$，我们都有如下的等式（用 B 将任何集合砍成两块，两块的外测度之和为该集合的外测度）

$$\mu^*(E\cap B)+\mu^*(E-B)=\mu^*(E),$$

那么，我们称 B 为μ^*-**可测的**。我们用 $\mathcal{B}$ 来代表一切 μ^*-可测的子集。我们注意到，由于

$$\mu^*(E\cap B)+\mu^*(E-B)\geqslant\mu^*(E),$$

为了验证 μ-可测性，只要验证不等式

$$\mu^*(E\cap B)+\mu^*(E-B)\leqslant\mu^*(E),$$

即可。

(第四步) $\mathcal{A}$ 中元素均为 μ^*-可测的。

为了说明 $\mathcal{A}\subset\mathcal{B}$，任意选取 $A\in\mathcal{A}$，$E\subset X$ 和 E 的覆盖 $\{A_i\}_{i\geqslant1}\subset\mathcal{A}$。注意到

$$\begin{aligned}\mu^*(E\cap A)&\leqslant\sum_{i\geqslant1}\mu(A_i\cap A)\leqslant\sum_{i\geqslant1}\mu^*(A_i\cap A),\\ \mu^*(E-A)&\leqslant\sum_{i\geqslant1}\mu(A_i\cap A^c)\leqslant\sum_{i\geqslant1}\mu^*(A_i\cap A^c),\end{aligned}$$

所以

$$\begin{aligned}\mu^*(E\cap A)+\mu^*(E-A)&\leqslant\sum_{i\geqslant1}\big(\mu(A_i\cap A)+\mu(A_i-A)\big)\\ &=\sum_{i\geqslant1}\mu(A_i).\end{aligned}$$

由 $\{A_i\}_{i\geqslant1}$ 选取的任意性，我们得到

$$\mu^*(E\cap A)+\mu^*(E-A)\leqslant\mu^*(E).$$

从而 A 是 μ^*-可测的。

(第五步) $\mathcal{B}$ 是 X 上的代数。

很明显，$\emptyset,X\in\mathcal{B}$。按定义，$\mathcal{B}$ 中元素对于求补集合是稳定的。我们只要证明 $\mathcal{B}$ 对于取交集是封闭的即可。任取 $B,C\in\mathcal{B}$，要证明 $D=B\cap C\in\mathcal{B}$。为此，任选 $E\subset X$，那么

$$\begin{aligned}\mu^*(E)&=\mu^*(E\cap C)+\mu^*(E-C)\\ &=\mu^*(E\cap D)+\mu^*\big((E\cap C)-B\big)+\mu^*(E-C)\end{aligned}$$

其中，我们对 $E\cap C$ 这个集合用 B 是 μ^*-可测的性质。由于

$$\big((E\cap C)-B\big)\cup(E-C)\supset E-D,$$

所以

$$\mu^*(E)\geqslant\mu^*(E\cap D)+\mu^*(E-D).$$

这表明 D 是 μ^*-可测的。

(第六步) μ^* 为 $\mathcal{B}$ 上的加性函数。

我们证明更一般的等式，对任何两个不交的 μ-可测集 $B,C\in\mathcal{B}$，我们都有

$$\mu^*\big(E\cap(B\cup C)\big)=\mu^*(E\cap B)+\mu^*(E\cap C).$$

在这个等式中取 $E=X$ 就得到 $\mu^*(B\cup C)=\mu^*(B)+\mu^*(C)$，这说明 μ^* 是加性函数。

为证明上面的等式，对任意的集合 E，我们对 $E'=E\cap(B\cup C)$ 和 B 用 μ^*-可测的定义：

$$\begin{aligned}\mu^*(E')&=\mu^*(E'\cap B)+\mu^*(E'-B)\\&=\mu^*(E\cap B)+\mu^*\big((E\cap C)-B\big)\\&\overset{B\text{ 可测}}{=}\mu^*(E\cap B)+\Big[\mu^*(E\cap C)-\mu^*\big((E\cap C)\cap B\big)\Big]\end{aligned}$$

由于最后一项是零（$C\cap B=\emptyset$），所以。

$$\mu^*(E')=\mu^*(E\cap B)+\mu^*(E\cap C).$$

这就是我们要证明的等式。

(第七步) $\mathcal{B}$ 为 σ-代数而 μ^* 为 $\mathcal{B}$ 上的测度。

由于 $\mathcal{B}$ 为 σ-代数，所以它包含 $\sigma(A)$。这样，μ^* 可以作为 μ 的扩张，这就完成了 Carathéodory 定理的证明。

为了证明上面的论断，只要对两两不交的 $\{B_i\}_{i\geqslant1}\subset\mathcal{B}$ 来证明 $\bigcup\limits_{i\geqslant1}B_i\in\mathcal{B}$ 并且

$$\mu^*\left(\bigcup_{i\geqslant1}B_i\right)=\sum_{i=1}^{\infty}\mu^*(B_i)$$

即可。为证明 $\bigcup\limits_{i\geqslant1}B_i\in\mathcal{B}$，我们令 $C_n=\bigcup\limits_{i\leqslant n}B_i$，$C_\infty=\bigcup\limits_{i\geqslant1}B_i$，需要说明 C_∞ 是 μ^*-可测的）。第六步中等式，对任意的 E，我们有

$$\begin{aligned}\mu^*(E)&=\mu^*(E\cap C_n)+\mu^*(E-C_n)\\&=\left(\sum_{i\leqslant n}\mu^*(E\cap B_i)\right)+\mu^*(E-C_n)\\&\geqslant\sum_{i\leqslant n}\mu^*(E\cap B_i)+\mu^*(E-C_\infty).\end{aligned}$$

令 $n\to\infty$，我们得到

$$\begin{aligned}\mu^*(E)&\geqslant\sum_{i=1}^{\infty}\mu^*(E\cap B_i)+\mu^*(E-C_\infty)\\&\geqslant\mu^*(E\cap C_\infty)+\mu^*(E-C_\infty),\end{aligned}$$

这给出了 $C_\infty\in\mathcal{B}$ 的证明。在上式中取 $E=C_\infty$ 也蕴含了 μ^* 为 $\mathcal{B}$ 上的测度。

至此，我们完成了 Carathéodory 定理的证明。□

40 测度的像（推出），Lebesgue 测度的构造，平移不变性刻画 Lebesgue 测度，Lebesgue 测度在尺度变换下的变换，测度空间的完备化，简单／阶梯函数的定义

二零二零年三月二十三日，星期一，晴

知识的回顾和总结

所谓的可测空间 $(X,\mathcal{A})$ 指的是一个集合 X 配上一个由它某些的子集的集合 $\mathcal{A}$，其中，我们要求 $\mathcal{A}$ 是所谓的 σ-代数，即它满足如下三个条件：

1) 空集 $\emptyset \in \mathcal{A}$；

2) 如果 $A \in \mathcal{A}$，那么 $A^c \in \mathcal{A}$；

3) 如果 $A_1, A_2, \cdots, A_n, \cdots \in \mathcal{A}$，那么 $\bigcup_{i=1}^{\infty} A_i \in \mathcal{A}$。

利用定义，我们知道 $\mathcal{A}$ 中任意可数个元素的交和并仍然落在 $\mathcal{A}$ 中。

所谓的测度空间可测空间 $(X,\mathcal{A},\mu)$ 指的是一个可测空间 $(X,\mathcal{A})$ 配备上一个测度函数

$$\mu : \mathcal{A} \to [0,\infty],$$

其中，我们要求 $\mu(\emptyset) = 0$ 并且对于 $\mathcal{A}$ 中两两不交的可数个元素 A_i，我们有

$$\mu(\bigcup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} \mu(A_i)$$

测度有一个很重要的性质，它与单调序列的极限可以交换，即若 $\{A_i\}_{i\geqslant 1}$ 是给 $\mathcal{A}$ 中的序列，那么，如果如下条件之一被满足：

1) 如果序列 $\{A_i\}_{i\geqslant 1}$ 是上升的；

2) 如果序列 $\{A_i\}_{i\geqslant 1}$ 是下降的并且对某个 i_0 有 $\mu(A_{i_0}) < \infty$，

那么，

$$\mu(\lim_{i\to\infty} A_i) = \lim_{i\to\infty} \mu(A_i).$$

对于两个可测空间 $(X,\mathcal{A})$ 和 $(Y,\mathcal{B})$ 之间的映射 $f : X \to Y$，如果对任意的 $B \in \mathcal{B}$，$f^{-1}(B) \in \mathcal{A}$，我们就称这个映射是可测的。如果只考虑映射 $f : X \to Y$ 而忘掉 X 上的 σ-代数 $\mathcal{A}$，我们可以用 f 把 Y 上的 σ-代数拉回来：

$$f^*\mathcal{B} = \left\{ f^{-1}(B) \middle| B \in \mathcal{B} \right\}.$$

这是 X 上的一个 σ-代数。所以，f 可测等价于 $f^*\mathcal{B} \subset \mathcal{A}$。

我们对于测度可以定义所谓的测度的推出：假设我们可测空间 $(X, \mathcal{A}, \mu)$ 和 $(Y, \mathcal{B})$ 以及它们之间的可测映射：

$$f : X \to Y.$$

其中，我们假定了 X 上配备了测度 μ。据此，我们可以定义 Y 上的一个测度 $f_*\mu$ 使得 $(Y, \mathcal{B}, f_*\mu)$ 称为测度空间。这个测度 $f_*\mu$ 被称作是μ **在 f 下的像测度**，其定义如下：对每个 $B \in \mathcal{B}$，根据 f 的可测性，$f^{-1}(B) \in \mathcal{A}$，所以，我们可以用 μ 来衡量 $f^{-1}(B)$ 的"面积"：

$$(f_*\mu)(B) := \mu\big(f^{-1}(B)\big).$$

在本周的作业中，我们会验证 $f_*\mu$ 是 $\mathcal{B}$ 上的测度，从而和 $(Y, \mathcal{B}, f_*\mu)$ 是测度空间。

注记. 测度的像是很重要的概念，这是积分的坐标变换公式的抽象推广。我们将要证明，如果 $\Phi : U \to V$ 是 $\mathbb{R}^n$ 中两个区域之间的微分同胚，其中，U 上的坐标我们用 $x = (x_1, \cdots, x_n)$ 表示，V 上的坐标我们用 $y = (y_1, \cdots, y_n)$ 表示，那么，它所对应的多元函数的 *Riemann* 积分的换元公式就是

$$\Phi_*\left(\Big|\det\left(\mathrm{Jac}(\Phi)\right)\Big| dx\right) = dy,$$

其中，dx 和 dy 分别是 U 和 V 上的 *Lebesgue* 测度，$\mathrm{Jac}(\Phi)$ 是 Φ 的 *Jacobi* 矩阵。这是多元微积分三大基本定理之一（另外两个是 *Fubini* 定理和 *Stokes* 公式）。

我们最后回忆上次的 Carathéodory 的测度扩张定理：

假定 $\mathcal{A}$ 为 X 上的代数，μ 为 $\mathcal{A}$ 上 σ-有限加性函数，我们想在 $\sigma(\mathcal{A})$ 上定义一个测度 μ'，使得 $\mu'\big|_{\mathcal{A}} = \mu$。如果 $\mu(X) < \infty$，那么下述条件保证了延拓测度 μ' 的存在性：

条件 (C). 对任意单调下降的序列 $\{A_i\}_{i\geqslant 0} \subset \mathcal{A}$，如果 $\mu(A_0) < \infty$ 并且 $\lim\limits_{i\to\infty} A_i = \emptyset$，那么 $\lim\limits_{i\to\infty} \mu(A_i) \to 0$。

如果 $\mu(X) = +\infty$，还需要再加上如下条件：

条件 ($\mathbf{C}_\infty$). 存在单调上升的序列 $\{X_i\}_{i\geqslant 1} \subset \mathcal{A}$，$\lim\limits_{i\to\infty} X_i = X$ 并对任意的 $i \geqslant 1$ 都有 $\mu(X_i) < \infty$，满足对任意的 $A \in \mathcal{A}$，$\mu(A) = +\infty$，我们都有 $\lim\limits_{i\to\infty} \mu(A \cap X_i) = +\infty$。

在 Carathéodory 定理的证明过程中，我们实际上给出了这个测度 μ' 的具体构造：对任意的 $E \subset \sigma(\mathcal{A})$，我们有

$$\mu^*(E) = \inf_{\substack{A_i \in \mathcal{A}, \\ E \subset \bigcup A_i}} \left(\sum_{i=1}^{\infty} \mu(A_i)\right).$$

我们还可以要求上面公式中的 $\{A_i\}_{i\geqslant 1}$ 是两两不相交的。

Carathéodory 的测度扩张定理是我们构造测度的最有效工具。

$\mathbb{R}^1$ 上的 Lebesgue 测度：长度的定义

在 $\mathbb{R}^1$ 上，我们考虑由所有的开集所生成的 σ-代数 $\mathcal{B}(\mathbb{R}^1)$，我们将要在 $\mathbb{R}^1$ 上构造 Lebesgue 测度。先规范一下术语：当我们谈论一个区间的时候，我们指的是形如 $[a,b],(a,b),(a,b],[a,b)$ 的区间，其中 a 和 b 可以分别取到 $-\infty$ 和 $+\infty$。对于任意上述一个区间 I，无论它是开的还是闭的，我们定义它的长度为

$$|I| = b - a.$$

当然，我们容许 $|I| = +\infty$。

任意给定两个区间，如果它们的并集合不构成新的区间，我们就称这两个区间是分离的，比如说，$(0,1)$ 与 $(1,2)$ 是分离的，因为 $(0,1)\cup(1,2)$ 不是区间；$(1,2)$ 与 $[1,+\infty)$ 就不是分离的，因为它们的并为 $(1,\infty)$ 是一个区间。再比如 $[-1,0]$ 与 $[2019,2020]$ 是分离的而 $[1,2020]$ 与 $[2019,2021]$ 就不是分离的。

定义 260. 令 $\mathcal{A}(\mathbb{R}^1)$ 为 $\mathbb{R}^1$ 上有限个区间的并所构成的集合的全体，即

$$\mathcal{A}(\mathbb{R}^1) = \big\{I_1 \cup I_2 \cup \cdots \cup I_m \big| m \in \mathbb{Z}_{\geqslant 1}, I_i \text{ 为区间}, 1 \leqslant i \leqslant m\big\}.$$

按照定义，$\mathcal{A}(\mathbb{R}^1)$ 是 $\mathbb{R}^1$ 上的代数，因为任意两个 $\mathcal{A}(\mathbb{R}^1)$ 中东西的并还在 $\mathcal{A}(\mathbb{R}^1)$ 中。

考虑 $\mathcal{A}(\mathbb{R}^1)$ 中的元素 $I_1\cup I_2\cup\cdots\cup I_m$，如果某个 I_k 和 I_ℓ 不是分离的，不妨假设 $k = m-1$，$\ell = m$，那么我们可以把这两个区间并在一起的到一个新的区间 $I'_{m-1} = I_{m-1}\cup I_m$，这样子令 $I'_1 = I_1, I'_2 = I_2, \cdots, I'_{m-2} = I_{m-2}$，我们就可以把这个元素重新写成 $I'_1\cup I'_2\cup\cdots\cup I'_{m-1}$。我们可以重复这种操作一直到 $I_1\cup I_2\cup\cdots\cup I_m$ 中的区间两两都是分离的。所以，$\mathcal{A}(\mathbb{R}^1)$ 中元素可以写成有限个分离的区间的并。另外，$\mathcal{A}(\mathbb{R}^1)$ 中元素只能以唯一的方式写成有限个分离的区间的并，我们把这个性质的证明留成作业。

在 $\mathcal{A}(\mathbb{R}^1)$ 上有一个自然的加性函数 m：对于 $P = I_1\cup I_2\cup\cdots\cup I_m \in \mathcal{A}(\mathbb{R}^1)$，其中，$\{I_k\}_{1\leqslant k\leqslant m}$ 是 m 个两两分离的区间的并，我们定义

$$m(P) = |I_1| + |I_2| + \cdots + |I_m| \in [0,\infty].$$

为了说明 $m:\mathcal{A}(\mathbb{R}^1)\to[0,+\infty]$ 是加性函数，我们固定 $P = I_1\cup I_2\cup\cdots\cup I_m\in\mathcal{A}(\mathbb{R}^1)$，只要说明对于任意的区间 I，$I\cap P = \emptyset$，我们有

$$m(P\cup I) = m(P) + |I|$$

即可，因为对于一般的 $Q = J_1\cup J_2\cup\cdots\cup J_n\in\mathcal{A}(\mathbb{R}^1)$，我们可以把 $P\cup Q$ 看作是每次并上一个 J_i 从而用这个结论归纳得到。为此，我们分情况讨论。我们可以假设所出现的区间的长度都是有限的，否则就没有什么可以证明的。另外，我们可以把所有的小区间的端点都去掉，这样对于区间的长度没有影响，从而，$I\cap P = \emptyset$ 意味着 I 与任意的 A_i 都是分离的，所以

$$m(P\cup I) = |I_1| + |I_2| + \cdots + |I_m| + |I|.$$

我们想要把加性函数 m 扩张到 $\mathcal{B}(\mathbb{R}^1)=\sigma(\mathcal{A}(\mathbb{R}^1))$ 上去（仍然记作 m）：

$$\begin{array}{ccc}\mathcal{A}(\mathbb{R}^1) & \xrightarrow{\ m\ } & [0,\infty] \\ \big\downarrow{\scriptstyle \iota} & \nearrow_{m} & \\ \mathcal{B}(\mathbb{R}^1) & & \end{array}$$

我们要使用 Carathéodory 扩张定理。先验证**条件** $(\mathbf{C}_\infty)$：令 $X_p=[-p,p]$，那么，$\{X_p\}_{p\geqslant 1}\subset\mathcal{A}^1$ 并且 $\lim\limits_{p\to\infty}X_p=\mathbb{R}^1$。对于任意的 $P=I_1\cup I_2\cup\cdots\cup I_m\in\mathcal{A}(\mathbb{R}^1)$，假设 $m(P)=\infty$，我们不妨设 $|I_1|=\infty$。那么，

$$m(X_p\cap P)\geqslant m(X_p\cap I_1)\to\infty,\quad p\to\infty.$$

所以**条件** $(\mathbf{C}_\infty)$ 成立。

现在来验证**条件** $(\mathbf{C})$。我们将用到用 $\mathbb{R}^1$ 上紧集的性质。我们首先观察到，如果 $P=I_1\cup I_2\cup\cdots\cup I_m\in\mathcal{A}(\mathbb{R}^1)$ 的长度是有限的，即

$$|I_1|+|I_2|+\cdots+|I_m|<\infty,$$

那么，我们可以在差一个 $\varepsilon>0$ 的意义下把所有的 I_i 都修改称有界闭区间，从而把 P 修改称略小的紧集合。实际上，对任意的 $i\leqslant m$，我们假设 I_i 的左端点为 a_i，右端点为 b_i，那么，我们考虑 $I_i'=\left[a_i+\dfrac{\varepsilon}{2^i},b_i-\dfrac{\varepsilon}{2^i}\right]\subset I_i$，从而，$P'=I_1'\cup I_2'\cup\cdots\cup I_m'\in\mathcal{A}(\mathbb{R}^1)$ 是紧集（有界闭集），并且

$$m(P)-\varepsilon<m(P')<m(P).$$

现在任取 $\mathcal{A}(\mathbb{R}^1)$ 中下降的序列 $\{A_i\}_{i\geqslant 1}$，其中 $m(A_1)<\infty$ 并且 $\lim\limits_{i\to\infty}A_i=\emptyset$，**条件** $(\mathbf{C})$ 要求我们来验证 $\lim\limits_{i\to\infty}m(A_i)=0$。用反证法：如果不然，那么 $\lim\limits_{i\to\infty}m(A_i)=a>0$（极限总是存在的因为 $\{m(A_i)\}_{i\geqslant 1}$ 是单调下降的）。利用刚才讨论的性质，对每个 $i\geqslant 1$，我们选取 $\mathcal{A}(\mathbb{R}^1)$ 中紧集 $A_i'\subset A_i$，使得 $m(A_i-A_i')<\dfrac{a}{2^{i+1}}$。当然，我们序列 $\{A_i'\}_{i\geqslant 1}$ 可能不再是下降的序列了，我们要对它们加以修正：定义

$$\widetilde{A_i}=\bigcap_{j\leqslant i}A_j'\subset A_i.$$

很明显，序列 $\{\widetilde{A_i}\}_{i\geqslant 1}\subset\mathcal{A}(\mathbb{R}^1)$ 是下降的紧集序列。我们有

$$A_i-\widetilde{A_i}\subset\bigcup_{j\leqslant i}(A_i-A_j').$$

从而，

$$\begin{aligned}m(A_i-\widetilde{A_i})&<\sum_{j\leqslant i}m(A_i-A_j')\\&\leqslant\sum_{j\leqslant i}m(A_j-A_j')\leqslant\sum_{j\leqslant i}\frac{a}{2^{j+1}}=\frac{a}{2}.\end{aligned}$$

所以，我们得到紧集 $\widetilde{A_i} \searrow \emptyset$，并且上面的不等式表明

$$\lim_{i\to\infty} m(\widetilde{A_i}) > \frac{a}{2} > 0.$$

这显然是不可能的，因为紧集序列 $\widetilde{A_i} \searrow \emptyset$ 意味着从某个 i_0 开始，后面的 A_i 全为空集！（我们将在作业中重新证明这个重要的命题，同学们应该熟记它）

根据 Carathéodory 定理，我们就有如下重要的定理：

定理 261. *在 $\mathbb{R}^1$ 配备 Borel-代数 $\mathcal{B}(\mathbb{R}^1)$，存在唯一的测度 m（称为 **Lebesgue 测度**），使得对任意的区间 $I \in \mathbb{R}$，我们有 $m(I) = |I|$。特别地，对任意的 $E \in \mathcal{B}(\mathbb{R}^1)$，它的测度可由如下公式计算*

$$m(E) = \inf_{\substack{E\subset\bigcup I_i,\\ I_i\text{是区间}}} \sum_{i=1}^{\infty} |I_i|.$$

$\mathbb{R}^n$ 上的 Lebesgue 测度：体积的定义

在 $\mathbb{R}^n$ 上，我们称形如 $C = I_1 \times \cdots \times I_n$ 的集合为方块，即

$$C = \big\{(x_1, x_2, \cdots, x_n)\big| x_1 \in I_1, \cdots, x_n \in I_n\big\},$$

其中，对每个 $i \leqslant n$，I_i 均为 $\mathbb{R}^1$ 上的区间。我们强调，这里的方块的边是和坐标轴平行的。

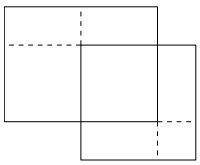
定义 262. *令 $\mathcal{A}(\mathbb{R}^n)$ 为 $\mathbb{R}^n$ 上有限个方块的并所构成的集合的全体，即*

$$\mathcal{A}(\mathbb{R}^n) = \big\{C_1 \cup C_2 \cup \cdots \cup C_m \big| m \in \mathbb{Z}_{\geqslant 1}, C_i \text{ 为方块}, 1 \leqslant i \leqslant m\big\}.$$

按照定义，$\mathcal{A}(\mathbb{R}^n)$ 是 $\mathbb{R}^n$ 上的代数。

命题 263. *$\mathcal{A}(\mathbb{R}^n)$ 中每个元素都可以被写成有限个不交的方块的并（方式不是唯一的）。*

证明: 这个命题的证明是显然，因为任意两方块的并集都可以写成更小的不交方块的并集，我们用下面的示意图来表示：



其中，两个方块用虚线切成了更小的方块。 □

我们要在 $\mathcal{A}(\mathbb{R}^n)$ 上定义一个加性函数 m。对于方块 $C = I_1 \times I_2 \times \cdots \times I_n$，我们令

$$m(C) = m(I_1) \times m(I_2) \times \cdots \times m(I_n) = |I_1| \times |I_2| \times \cdots \times |I_n| \in [0, \infty].$$

其中，$|I_j|$ 是区间 I_i 的长度。我们约定，在 $[0, \infty]$ 上，0 乘以任何数（包括 $+\infty$）都得 0，$+\infty$ 乘以任何非零的数都得 $+\infty$。

对于 $\mathcal{A}(\mathbb{R}^n)$ 中任意给定的元素 S，我们可以假设 $S = C_1 \cup C_2 \cup \cdots \cup C_m$，其中 $\{C_i\}_{i\leqslant m}$ 是两两不交的方块。我们就定义

$$m(S) = \sum_{i\leqslant m} m(C_i) \in [0, \infty].$$

由于 S 有可能可以写成其他形式的两两不交的方块的并 $S = C'_1 \cup C'_2 \cup \cdots \cup C'_{m'}$，我们必须说明 $m(S)$ 的定义不依赖于分解 $S = C_1 \cup C_2 \cup \cdots \cup C_m$ 或者 $S = C'_1 \cup C'_2 \cup \cdots \cup C'_{m'}$ 的选取。与上个学期对区间的分划进行加细是一个道理，利用上面命题的图示，我们可以选取两两不交的方块分解 $S = C''_1 \cup C''_2 \cup \cdots \cup C''_{m''}$，使得每个 C_i 或者 $C'_{i'}$ 都是若干个 C''_j 的并，从而

$$\sum_{i\leqslant m} m(C_i) = \sum_{i''\leqslant m''} m(C''_{i''}) = \sum_{i'\leqslant m'} m(C'_{i'}).$$

证明的细节我们留作本次的作业。

这样，我们就得到了 $\mathcal{A}(\mathbb{R}^n)$ 上的加性函数：

$$m : \mathcal{B}(\mathbb{R}^n) \to [0, \infty].$$

我们现在要把 m 扩张到整个 $\mathcal{B}(\mathbb{R}^n) = \sigma(\mathcal{A}(\mathbb{R}^n))$ 上，其中 Borel-代数 $\mathcal{B}(\mathbb{R}^n)$ 是 $\mathbb{R}^n$ 上所有开集所生成的 σ-代数。当然，我们可以用更少的开集来生成 $\mathcal{B}(\mathbb{R}^n)$：

命题 264. *$\mathcal{B}(\mathbb{R}^n)$ 可以由所有顶点为有理点（即每个坐标都是有理数）的闭正方体生成。*

证明： 实际上就第九次课中引理236的证明就可以给出这个结论，我们不再重复。 □

我们准备再次运用 Carathéodory 测度扩张定理来定义 $\mathcal{B}(\mathbb{R}^n)$ 上的 Lebesgue 测度。对比 $\mathbb{R}^1$ 的情形，我们先证明一个引理：

引理 265. *对任意的 $S \in \mathcal{A}(\mathbb{R}^n)$，如果 $m(S) < \infty$，那么对任意的 $\varepsilon > 0$，存在 $S' \in \mathcal{A}(\mathbb{R}^n)$，使得 S' 是有限个闭的方块的并，$S' \subset S$ 并且*

$$m(S - S') < \varepsilon.$$

证明： 证明是直截了当的：$S = C_1 \cup C_2 \cup \cdots \cup C_m$，所以只要对每个方块 C_i 考虑即可（它的测度 $m(C_i)$ 自然是有限的）。对于方块 C 而言，这是显然的，只要把每条边稍微缩短一点即可。 □

先验证**条件** $(\mathbf{C}_\infty)$：令 $X_p = [-p, p] \times [-p, p] \times \cdots \times [-p, p]$，那么，$\{X_p\}_{p\geqslant 1} \subset \mathcal{A}(\mathbb{R}^1)$ 并且 $\lim\limits_{p\to\infty} X_p = \mathbb{R}^n$。对于任意的 $S = C_1 \cup C_2 \cup \cdots \cup C_m \in \mathcal{A}(\mathbb{R}^1)$，假设 $m(S) = +\infty$，我们不妨设 $m(C_1) = \infty$。那么，

$$m(X_p \cap S) \geqslant m(X_p \cap C_1) \to \infty, \quad p \to \infty.$$

上面后面的极限是两个方体之间的相交，所以是显然的。据此，**条件** $(\mathbf{C}_\infty)$ 成立。

现在来验证**条件** $(\mathbf{C})$。现在任取 $\mathcal{A}(\mathbb{R}^n)$ 中下降的序列 $\{A_i\}_{i\geqslant 1}$，其中 $m(A_1) < \infty$ 并且 $\lim\limits_{i\to\infty} A_i = \emptyset$，我们要证明 $\lim\limits_{i\to\infty} m(A_i) = 0$。如若不然，那么 $\lim\limits_{i\to\infty} m(A_i) = a > 0$（极限存在因

为 $\{m(A_i)\}_{i\geqslant 1}$ 单调下降)。利用上述引理，对每个 $i\geqslant 1$，我们选取 $\mathcal{A}(\mathbb{R}^n)$ 中紧集 $A_i'\subset A_i$，使得 $m(A_i-A_i')<\dfrac{a}{2^{i+1}}$。和 1 维的情形一样，我们对 $\{A_i'\}_{i\geqslant 1}$ 加以修正来得到单调下降的序列。为此，定义

$$\widetilde{A_i}=\bigcap_{j\leqslant i}A_j'\subset A_i.$$

所以，$\{\widetilde{A_i}\}_{i\geqslant 1}\subset\mathcal{A}(\mathbb{R}^1)$ 是下降的紧集序列。我们有

$$A_i-\widetilde{A_i}\subset\bigcup_{j\leqslant i}(A_i-A_j').$$

从而，

$$\begin{aligned}m(A_i-\widetilde{A_i})&<\sum_{j\leqslant i}m(A_i-A_j')\\&\leqslant\sum_{j\leqslant i}m(A_j-A_j')\leqslant\sum_{j\leqslant i}\frac{a}{2^{j+1}}=\frac{a}{2}.\end{aligned}$$

由于，$\widetilde{A_i}\subset A_i$，所以 $\widetilde{A_i}\searrow\emptyset$ 并且上面

$$\lim_{i\to\infty}m(\widetilde{A_i})>\frac{a}{2}>0.$$

因为紧集序列 $\widetilde{A_i}\searrow\emptyset$ 意味着从某个 i_0 开始，后面的 A_i 全为空集，所以矛盾！在证明中，我们再次用到了如下的性质：

命题 266 (紧集区间套原理)**.** *给定 $\mathbb{R}^n$ 中紧集的序列 $\{K_i\}_{i\geqslant 1}$，如果这个序列是下降的并且 $\cap_{i\geqslant 1}K_i=\emptyset$，那么，存在 $i_0\geqslant 1$，使得当 $i\geqslant i_0$ 时，$K_i=\emptyset$。*

我们把上述命题的证明留成本次作业。

根据 Carathéodory 定理，我们得到多元微积分中最基本的一个定理：

定理 267. *在 $\mathbb{R}^n$ 配备 Borel-代数 $\mathcal{B}(\mathbb{R}^n)$，存在唯一的测度 m（称为* **Lebesgue 测度**)，*使得对任意的区间 $C\subset\mathbb{R}^n$，其中 $C=I_1\times I_2\times\cdots\times I_n$，我们有*

$$m(C)=|I_1|\times|I_2|\times\cdots\times|I_n|.$$

特别地，对任意的 $E\in\mathcal{B}(\mathbb{R}^1)$，它的测度可由如下公式计算

$$m(E)=\inf_{\substack{E\subset\bigcup C_i,\\ C_i\text{是方块}}}\sum_{i=1}^{\infty}m(C_i).$$

Lebesgue 测度的基本性质

我们来学习 Lebesgue 测度的基本性质。在 $\mathbb{R}^n$ 上，我们可以定义平移变换。这个变换本质上刻画了 Lebesgue 测度（局部紧的拓扑群也具有这样的一个测度，它在群的作用下不变，这是所谓的 Haar 测度)。对任意的 $v\in\mathbb{R}^n$，我们定义平移

$$\tau_v:\mathbb{R}^n\to\mathbb{R}^n,\quad x\mapsto x+v.$$

很明显，τ_{-v} 是 τ_v 的逆，这也是平移。另外，对于 $\mathbb{R}^n$ 的子集 $A \subset \mathbb{R}^n$，我们令

$$\tau_v(A) = \big\{x + v \big| x \in A\big\}.$$

按定义，我们有 $\big(\tau_{-v}\big)^{-1}(A) = \tau_v(A)$。

命题 268 (Lebesgue 测度的平移不变性)**.** *$\mathbb{R}^n$ 上的 Lebesgue 测度具有平移不变性，这指的是对任意的 $v \in \mathbb{R}^n$*

1) 平移 $\tau_v : (\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n)) \to (\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$ 是可测变换。

2) 对任意的 $A \in \mathcal{B}(\mathbb{R}^n))$，我们有 $m\big(\tau_v(A)\big) = m(A)$，即 $(\tau_v)_* m = m$。

反之，平移不变性在如下的意义下刻画了 Lebesgue 测度：假定 μ 是可测空间 $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$ 上的测度，使得对任意的 $v \in \mathbb{R}^n$，我们有 $\tau_{v*}\mu = \mu$。如果

$$\mu\Big(\underbrace{[0,1] \times \cdots \times [1,1]}_{n \text{ 个}}\Big) = 1,$$

那么 $\mu = m$。

这是一个值得学习的证明，它巧妙地运用了 Carathéodory 定理的唯一性部分。

证明: 我们首先证明平移变换 τ_v 是可测的（一方面，这是显然的，因为连续映射是可测的），其中 $v \in \mathbb{R}^n$ 是给定的向量。实际上，为了验证可测性，我们对 $\mathcal{B}(\mathbb{R}^n)$ 的生成元来验证即可，因为 $\mathcal{B}(\mathbb{R}^n) = \sigma(\mathcal{A}(\mathbb{R}^n))$，对于任意的 $S = C_1 \cup C_2 \cup \cdots \cup C_m \in \mathcal{A}(\mathbb{R}^n)$，我们有

$$\tau_v^{-1}(S) = \tau_{-v}(C_1) \cup \tau_{-v}(C_2) \cup \cdots \cup \tau_{-v}(C_m).$$

这仍然是一些方块的并，所以，$\big(\tau_v\big)^{-1}(S) \in \mathcal{A}(\mathbb{R}^n) \subset \mathcal{B}(\mathbb{R}^n)$，所以 τ_v 可测。这个证明还表明

$$\tau_v : \mathcal{B}(\mathbb{R}^n) \to \mathcal{B}(\mathbb{R}^n)$$

是双射，即 Borel-集经过平移之后仍是 Borel-集。

我们现在证明 Lebesgue 测度的平移不变性，即 $(\tau_v)_* m = m$。为此，我们定义一个新的测度 $\mu = (\tau_v)_* m$，即对每个 $B \in \mathcal{B}(\mathbb{R}^n)$，

$$\mu(B) = m((\tau_v)^{-1} B).$$

对于 $S = C_1 \cup C_2 \cup \cdots C_m \in \mathcal{A}(\mathbb{R}^n)$，我们假设这些方块两两不交。根据上面的公式，我们有

$$\begin{aligned}\mu(S) &= m\big(\tau_{-v}(C_1) \cup \tau_{-v}(C_2) \cup \cdots \cup \tau_{-v}(C_m)\big) \\ &= \sum_{i \leqslant m} m\big(\tau_{-v}(C_i)\big) = \sum_{i \leqslant m} m\big(C_i\big) = m(S).\end{aligned}$$

上面换行处的等号成立是因为平移不会改变一个方块的各个边的长度。由于 μ 和 m 在 $\mathcal{A}(\mathbb{R}^n)$ 上的取值是一样的，所以它们都是同一个加性函数的扩张。根据 Carathéodory 测度扩张定理的唯一性，我们知道 $m = \mu = (\tau_v)_* m$。

我们现在证明平移不变性刻画了 Lebesgue 测度：这是一个思路清晰的“几何”证明。我们研究几种情况：

1) 对任意的自然数 $k \geqslant 1$，我们考虑一个每个边都与坐标轴平行的边长为 $\frac{1}{k}$ 的（闭）正方体 $C_{\frac{1}{k}}$，根据平移不变性，无论这个正方体的位置在哪里，它的测度 $m(C_{\frac{1}{k}})$ 是固定的。

2) 我们考虑 $C_{\frac{1}{k}}$ 的一个面，比如说，

$$C_{\frac{1}{k}} = \left[0, \frac{1}{k}\right] \times \left[0, \frac{1}{k}\right] \times \cdots \times \left[0, \frac{1}{k}\right]$$

的面

$$E = \{0\} \times \left[0, \frac{1}{k}\right] \times \cdots \times \left[0, \frac{1}{k}\right]$$

很明显，对任意的 N，我们可以把 E 的 $N+1$ 个平移（两两不相交）$\left\{\tau_{\frac{i}{Nk}}(E)\right\}_{i=0,1,\cdots,N}$ 放到 E 中去，所以

$$\mu(C_1) \geqslant \mu(C_{\frac{1}{k}}) \geqslant \sum_{i=0}^{N+1} \mu\left(\tau_{\frac{i}{Nk}}(E)\right) = (N+1)E.$$

令 $N \to \infty$，我们就知道 E 的测度为零。

3) 我们可以用 k^n 个 $C_{\frac{1}{k}}$ 拼出来 C_1，其中 $C_1 = \underbrace{[0,1] \times \cdots \times [1,1]}_{n\text{ 个}}$。这 k^n 个小正方体可能在它们的边界上相交，这些所有边界给出的集合的测度为零，这是 2) 的结论，所以

$$k^n \mu(C_{\frac{1}{k}}) = \mu(C_1) = 1,$$

从而，$\mu(C_{\frac{1}{k}}) = k^{-n} = m(C_{\frac{1}{k}})$。

4) 如果方块的边长是有理数，那么我们可以用若干个 3) 中的正方体拼出来，它们只在边界上相交，而边界的测度为零，所以，对于每边长为有理数的方块 C，我们有 $\mu(C) = m(C)$。

上面的 4) 表明，那么 m 和 μ 在 $\mathcal{A}(\mathbb{R}^n)$ 上面的取值是一样的，所以，利用 Carathéodory 定理，我们就有 $m = \mu$。 □

推论 269. *对任意的 $\lambda > 0$，我们可以定义 $\mathbb{R}^n$ 上的伸缩变换：*

$$\rho_\lambda : \mathbb{R}^n \to \mathbb{R}^n, \quad x \mapsto \lambda x.$$

那么，我们有 $(\rho_\lambda)_* m = \lambda^{-n} m$，即对任意的 $A \in \mathcal{B}(\mathbb{R}^n)$，我们有 $m(\rho_\lambda(A)) = \lambda^n m(A)$。

证明: 证明很直接: 考虑测度 $\mu = \lambda^{-n}(\rho_\lambda)_* m$。由于方块在伸缩变换下还变成方块（只是每条边变长了 λ 倍), 所以 μ 和 m 在方块上的取值是一样的。我们利用 Carathéodory 定理的结论即可。 □

推论 270. 对任意的正交矩阵 $R \in \mathbf{O}(n)$，我们可以定义 $\mathbb{R}^n$ 上的旋转变换:

$$R : \mathbb{R}^n \to \mathbb{R}^n, \quad x \mapsto R \cdot x.$$

其中，$R \cdot x$ 是矩阵和列向量的乘法。那么，我们有 $R_* m = m$，即对任意的 $A \in \mathcal{B}(\mathbb{R}^n)$，我们有 $m\left(R(A)\right) = m(A)$。

根据我们的方块的定义，它们的边必须和坐标轴平行，所以旋转之后方块的体积很难计算。我们把这个有意思又有意义的结论留成本次的作业。

注记. （$\mathbb{R}^n$ 的子集） 对于 $X \subset \mathbb{R}^n$，令 $\iota : X \hookrightarrow \mathbb{R}^n$ 为标准的包含映射，即

$$\iota : X \hookrightarrow \mathbb{R}^n, \quad x \mapsto x.$$

那么，我们可以将 $\mathcal{B}(\mathbb{R}^n)$ 拉回使得 X 称为可测空间，即令

$$\mathcal{B}(X) := \iota^* \mathcal{B}(\mathbb{R}^n) = \left\{ \iota^{-1}(B) = B \cap X | B \in \mathcal{B}(\mathbb{R}^n) \right\},$$

我们称上述 σ-代数为 X 上的 *Borel*-代数。

在今后的应用中，我们只会考虑 X 为开集或闭集的情形，此时 $\mathcal{B}(X)$ 就是 X 中的开集或者闭集生成的 σ-代数。所以，当 X 为开集或闭集时，$\mathcal{B}(X)$ 中的元素就是 $\mathcal{B}(\mathbb{R}^n)$ 中的元素，我们仍然可以计算它的 *Lebesgue* 测度。

测度空间的完备化（补充材料）

类似于度量空间，我们可以对测度空间进行完备化。我们先引入一个概念:

定义 271. $(X, \mathcal{A}, \mu)$ 是测度空间。对于 $A \subset \mathcal{A}$，如果 $\mu(A)\, 0$，我们就称它是μ-**零测集**并且大部分情况都称它为**零测集**。如果对任意的零测集，它的子集仍然是 $\mathcal{A}$ 中的元素，我们就称测度空间 $(X, \mathcal{A}, \mu)$ 是完备的。

定理 272. 对每个测度空间 $(X, \mathcal{A}, \mu)$，存在一个**完备的**测度空间 $(X, \mathcal{A}', \mu')$，使得

1) $\mathcal{A} \subset \mathcal{A}'$ 并且 $\mu'\big|_{\mathcal{A}} = \mu$;

2) 对每个 $A' \in \mathcal{A}'$，存在 $A_1, A_2 \in \mathcal{A}$，使得 $A_1 \subset A \subset A_2$ 且 $\mu(A_2 - A_1) = 0$。

证明: 进行完备化的基本想法是手动加入所有零测集的子集。为此，我们定义

$$\mathcal{A}' := \left\{ A' \subset X \middle| \text{存在 } B, C \in \mathcal{A}, \text{ 使得 } B \subset A' \subset C \text{ 且 } \mu(C - B) = 0 \right\}.$$

证明分为三步:

1) $\mathcal{A}'$ 为 σ-代数。

对于 $A' \in \mathcal{A}'$，有 $B, C \in \mathcal{A}$，使得 $B \subset A' \subset C$ 且 $\mu(C - B) = 0$，所以，$B^c \supset A'^c \subset C^c$ 并且 $\mu(B^c - C^c) = 0$，从而，$A'^c \in \mathcal{A}'$。对于任意的可数个 $A'_i \in \mathcal{A}'$，其中 $i = 1, 2, \cdots$，我们可以选取相应的 $B_i, C_i \in \mathcal{A}$，使得 $B_i \subset A'_i \subset C_i$ 并且 $\mu(C_i - B_i) = 0$。那么，对于 $A' = \bigcup\limits_{i \geqslant 1} A'_i$ 而言，我们令 $B = \bigcup_{i \geqslant 1} B_i$，$C = \bigcup_{i \geqslant 1} C_i$，从而 $B, C \in \mathcal{A}$ 并且 $B \subset A' \subset C$。另外，

$$m(C - B) \leqslant \sum_{i=1}^{\infty} m(C_i - B_i) = 0,$$

从而，$A' \in \mathcal{A}'$。这就说明了 $\mathcal{A}'$ 为 σ-代数。

2) 测度 μ' 的定义。

对于 $A' \in \mathcal{A}'$，存在 $B, C \in \mathcal{A}$，使得 $B \subset A' \subset C$ 且 $\mu(C - B) = 0$，我们令

$$\mu'(A') := \mu(B).$$

当然，我们可能有别的选择，即存在 $\overline{B}, \overline{C} \in \mathcal{A}$，使得 $\overline{B} \subset A' \subset \overline{C}$ 且 $\mu(\overline{C} - \overline{B}) = 0$。为了说明上式是良好定义的，我们要说明

$$\mu(B) := \mu(\overline{B}).$$

实际上，我们有

$$\mu(B - B \cap \overline{B}) \leqslant m(C \cup \overline{C} - B \cap \overline{B}) \leqslant \mu(C - B) + \mu(\overline{C} - \overline{B}) = 0.$$

所以，$\mu(B) = \mu(B \cap \overline{B})$。同理，$\mu(\overline{B}) = \mu(B \cap \overline{B})$。至此，我们定义了映射

$$\mu' : \mathcal{A}' \to [0, \infty].$$

为了验证这是良好定义测度，对于任意的可数个两两不交的 $A'_i \in \mathcal{A}'$，其中 $i = 1, 2, \cdots$，我们选取相应的 $B_i, C_i \in \mathcal{A}$，使得 $B_i \subset A'_i \subset C_i$ 并且 $\mu(C_i - B_i) = 0$。根据 1)，对于 $A' = \bigcup\limits_{i \geqslant 1} A'_i$ 和 $B = \bigcup_{i \geqslant 1} B_i$，$C = \bigcup_{i \geqslant 1} C_i$ 是相应的集合。我们注意到 $\{B_i\}_{i \geqslant 1}$ 两两不交，所以

$$\begin{aligned} \mu'(A') = \mu(B) &= \mu\left(\bigcup_{i \geqslant 1} B_i\right) \\ &= \sum_{i \geqslant 1} \mu(B_i) = \sum_{i \geqslant 1} \mu'(A'_i). \end{aligned}$$

这说明 μ 是测度。

3) $(X, \mathcal{A}', \mu')$ 是完备的。

任选零测集 $A' \in \mathcal{A}'$，我们取 $B = \emptyset$，$C \in \mathcal{A}$ 是零测集。那么，对于任意的 $A'' \subset A$，我们仍然选取 $B = \emptyset$ 和 C 作为 $\mathcal{A}'$ 定义中的集合，所以 $A'' \in \mathcal{A}$。

综上所述，命题得证。 □

注记. 上述证明给出了测度空间完备化的一个特定的构造，我们将按照这种方式给出的测度空间 $(X,\mathcal{A}',\mu')$ 称为原来测度空间的完备化。我们在二年级要学习的实变函数／实分析理论，实际上研究的是 $(\mathbb{R}^n,\mathcal{B}(\mathbb{R}^n),m)$ 的完备化上的各种分析，这对我们将要学习的多元微积分没有任何影响，我们就不再展开叙述。从此往后，我们只关心 $(\mathbb{R}^n,\mathcal{B}(\mathbb{R}^n),m)$。

积分理论

我们先做如下的约定：**从此往后，除非特别强调，所有的映射都是可测的。**

给定一个测度空间 $(X,\mathcal{A},\mu)$，我们先定义阶梯函数或者说简单函数的概念：

定义 273. 给定可测函数 $f:X\to\mathbb{C}$，如果 f 满足如下两个条件：

1) f 只取有限多个值，即 $|f(X)|<\infty$；

2) f 的非零值的逆像的测度有限，即对任意的 $\lambda\in\mathbb{C}$ 并且 $\lambda\neq 0$（根据可测性，$f^{-1}(\lambda)\in\mathcal{A}$），那么 $\mu(f^{-1}(\lambda))<\infty$，

我们就称 f 是一个**简单函数**或者**阶梯函数**。我们用 $\mathcal{E}(X,\mathcal{A},\mu)$ 表示 $(X,\mathcal{A},\mu)$ 上简单函数的全体，为了简单起见，我们通常把它写成 $\mathcal{E}(X)$。

注记. 同学们应该查阅上学期的笔记对比两边简单函数定义的异同。

任意给定 $f\in\mathcal{E}(X)$，假设 $\lambda_0,\lambda_1,\cdots,\lambda_m$ 是 f 所有可能的取值，其中，当 $i\geqslant 1$ 时，$\lambda_i\neq 0$。所以，f 可以写成下面的形式：

$$f(x)=\sum_{i=0}^{n}\lambda_i\cdot\mathbf{1}_{A_i}(x),$$

其中，$A_i\in\mathcal{A}$ 并且对于 $i\geqslant 1$，$\mu(A_i)<\infty$。在上面的表达式中，我们用 $\mathbf{1}_{A_i}$ 表示 A_i 的示性函数。

反过来，对任意的 $m+1$ 个（可能相同）值 $\lambda_0,\lambda_1,\cdots,\lambda_m$（要求当 $i\geqslant 1$ 时，$\lambda_i\neq 0$），对任意的 $A_i\in\mathcal{A}$（$i=0,1,\cdots,m$）并且当 $i\geqslant 1$ 时，$\mu(A_i)<\infty$，我们考虑函数

$$f(x)=\sum_{i=0}^{n}\lambda_i\cdot\mathbf{1}_{A_i}(x).$$

这是显然是一个简单函数。

综上所述，我们得到下面的命题

引理 274. 一个简单函数 $f\in\mathcal{E}(X)$ 总是可以表示成

$$f(x)=\sum_{i=0}^{n}\lambda_i\cdot\mathbf{1}_{A_i}(x).$$

的形式，其中，$\lambda_0,\lambda_1,\cdots,\lambda_m\in\mathbb{C}$ 并且当 $i\geqslant 1$ 时，$\lambda_i\neq 0$；对每个 $i=0,1,\cdots,m$，$A_i\in\mathcal{A}$ 并且当 $i\geqslant 1$ 时，$\mu(A_i)<\infty$。

推论 275. 简单函数 $\mathcal{E}(X)$ 是 $\mathbb{C}$-线性空间。

证明: 这是显然的，因为任取两个简单函数

$$f(x)=\sum_{i=0}^{m}\alpha_i\cdot\mathbf{1}_{A_i}(x),\quad g(x)=\sum_{i=0}^{n}\beta_i\cdot\mathbf{1}_{B_i}(x),$$

它们的和形如

$$\sum_{i=0}^{m}\alpha_i\cdot\mathbf{1}_{A_i}(x)+\sum_{i=0}^{n}\beta_i\cdot\mathbf{1}_{B_i}(x).$$

这显然还是简单函数。 □

41 简单函数的积分，测度空间上积分的定义，零测集，几乎处处，Beppo Levi 定理

二零二零年三月二十五日，星期四，晴，大风

上节课的最后，我们在测度空间 $(X,\mathcal{A},\mu)$ 上定义了简单函数空间 $\mathcal{E}(X,\mathcal{A},\mu)$，其中，对于简单函数 f，按照定义，我们要求它只能取有限多个值并且非零值的逆像的测度有限。另外，每一个简单函数都可以写成下面的形式：存在 $m+1$ 个（可能相同）复数 $\lambda_0,\lambda_1,\cdots,\lambda_m$（要求 $\lambda_0=0$，当 $i\geqslant 1$ 时，$\lambda_i\neq 0$）和 $A_i\in\mathcal{A}$（$i=0,1,\cdots,m$）并且当 $i\geqslant 1$ 时，$\mu(A_i)<\infty$，使得

$$f(x)=\sum_{i=0}^{n}\lambda_i\cdot\mathbf{1}_{A_i}(x).$$

据此，我们也知道简单函数空间 $\mathcal{E}(X)$ 是 $\mathbb{C}$-线性空间。

在简单函数的定义中，我们要求逆像的测度有限，其目的是来定义积分的。

定义 276. *对任意的的 $f\in\mathcal{E}(X)$，我们用符号表示 $\int_X\cdot d\mu$ 积分，其定义为*

$$\int_X f d\mu:=\sum_{\lambda\in\mathbb{C}}\lambda\cdot\mu\left(f^{-1}(\lambda)\right).$$

从而，我们定义映射

$$\int_X:\mathcal{E}(X,\mathcal{A},\mu)\to\mathbb{C},\quad f\mapsto\int_X f d\mu.$$

和我们学习过的 Riemann 积分一样，积分算子满足线性：

命题 277. *积分算子*

$$\int_X:\mathcal{E}(X,\mathcal{A},\mu)\to\mathbb{C},\quad f\mapsto\int_X f d\mu.$$

是简单函数空间 $\mathcal{E}(X)$ 上的 $\mathbb{C}$-线性映射（泛函）。进一步，如果 $f,g\in\mathcal{E}(X)$ 是实数值的函数且对任意 $x\in X$，$f(x)\leqslant g(x)$，那么

$$\int_X f d\mu\leqslant\int_X g d\mu.$$

证明： 我们首先证明，如果 $f(x)=\sum_{i=0}^{m}\alpha_i\cdot\mathbf{1}_{A_i}(x)$，其中 $\alpha_0,\alpha_1,\cdots,\alpha_m\in\mathbb{C}$ 并且当 $i\geqslant 1$ 时，$\alpha_i\neq 0$；对每个 $i=0,1,\cdots,m$，$A_i\in\mathcal{A}$ 并且当 $i\geqslant 1$ 时，$\mu(A_i)<\infty$，那么

$$\int_X f(x)d\mu=\sum_{i=0}^{m}\alpha_i\mu(A_i).$$

注意到，我们不妨要求 $\alpha_0=0$ 并且 $\bigcup_{i=0}^{m}A_i=X$。实际上，我们首先把 f 写成

$$f(x)=\sum_{i=0}^{N}\lambda_i\cdot\mathbf{1}_{X_i}(x),$$

其中 $\lambda_0 = 0, \lambda_1, \cdots, \lambda_N \in \mathbb{C}$ 是两两不同的，$f\big|_{X_i} \equiv \lambda_i$ 并且 $\bigcup_{i=0}^{N} X_i = X$（这是一个无交并），从而，按照积分的定义，我们有

$$\int_X f(x)d\mu = \sum_{i=0}^{N} \lambda_i \mu(X_i).$$

通过把 f 写成

$$f(x) = \sum_{i=0}^{m}\sum_{j=0}^{N} \alpha_i \cdot \mathbf{1}_{A_i \cap X_j}(x)$$

我们不妨假设对于每个 A_i，都存在 X_j，使得 $A_i \subset X_j$。根据这些 X_j，我们重新把 $\{A_i\}_{i\leqslant m}$ 分类，那么，重新编号之后，我们得到

$$f(x) = \sum_{j=0}^{N}\left(\sum_{i=0}^{m} \alpha_{ij} \cdot \mathbf{1}_{A_{i,j}}(x)\right),$$

其中 $A_{i,j} \subset X_j$。按照 f 的定义，我们自然有

$$\sum_{i=0}^{m} \alpha_{ij} \cdot \mathbf{1}_{A_{i,j}}(x) = \lambda_j \mathbf{1}_{X_j}(x),$$

我们只要证明

$$\sum_{i=0}^{m} \alpha_{ij}\mu(A_{i,j}) = \lambda_j \mu(X_j) \quad \cdots\cdots \quad (\star)$$

即可，其中 $j \leqslant N$ 是固定的。为此，我们继续对 $A_{i,j}$ 进行分划，考虑所有的形如

$$A_{i_1 j} \cap A_{i_2 j} \cap \cdots \cap A_{i_s j} \cap A^c_{i'_1 j} \cap A^c_{i'_2 j} \cap \cdots \cap A^c_{i'_t j}$$

的集合，它们是两两不相交的，每个 $A_{i,j}$ 都可以写成它们的并，据此，我们只需要在 $(\star)$ 中假设 $A_{i,j}$ 是两两不交的即可，此时，命题是显然的。

回到线性的证明，根据上面的结论，任取两个简单函数

$$f(x) = \sum_{i=0}^{m} \alpha_i \cdot \mathbf{1}_{A_i}(x), \quad g(x) = \sum_{j=0}^{n} \beta_j \cdot \mathbf{1}_{B_j}(x),$$

那么

$$f(x) + g(x) = \sum_{i=0}^{m} \alpha_i \cdot \mathbf{1}_{A_i}(x) + \sum_{j=0}^{n} \beta_i \cdot \mathbf{1}_{B_j}(x).$$

所以，

$$\begin{aligned}\int_X f + g d\mu &= \sum_{i=1}^{m} \alpha_i \mu(A_i) + \sum_{j=1}^{m} \beta_j \mu(B_j) \\ &= \int_X f d\mu + \int_X g d\mu.\end{aligned}$$

对于最后一个论断，根据线性，我们证明

$$\int_X h d\mu \geqslant 0$$

即可，其中，$h = f - g \geqslant 0$。根据积分的定义，

$$\int_X j d\mu := \sum_{\lambda \in \mathbb{C}} \lambda \cdot \mu\left(h^{-1}(\lambda)\right).$$

由于 $h \geqslant 0$，上式出现的非零的 $\mu\left(h^{-1}(\lambda)\right)$ 所对应的 λ 都是非负的，所以显然成立。 □

注记 (正简单函数积分的定义)**.** 给定可测函数 $f : X \to [0, +\infty]$（我们约定 $[0, \infty]$ 上的 σ-代数取为由 $[0, \infty)$ 中的开集和点 $\{+\infty\}$ 所生成的 σ-代数），如果 f 只取有限多个值，即 $|f(X)| < \infty$，我们就称它为**正的简单函数**或者**非负简单函数**。与之前的定义相比，我们并不对 $\mu(f^{-1}(c))$ 有限制。此时，我们可以强行定义正简单函数的积分：

$$\int_X f d\mu := \sum_{\lambda \in [0,\infty]} \lambda \cdot \mu(f^{-1}(\lambda)).$$

如果上面的求和出现了 $+\infty$，我们就说这个积分值是正无穷大并记作 $\int_X f d\mu = +\infty$。

类似上面简单函数的定义，如果映射函数 $f : X \to [0, +\infty]$ 是可测的，那么，我们就称 f 为**正函数**或者**非负函数**。我们现在可以定义测度空间上可测函数的积分了：

定义 278 (积分的定义)**.** 对于正可测函数 $f : X \to [0, +\infty]$，我们把它的积分 $\int_X f d\mu$ 定义不超过该函数的正简单函数的积分的上确界，即

$$\int_X f d\mu := \sup_{\substack{\varphi \in \mathcal{E}(X) \\ 0 \leqslant \varphi \leqslant f}} \int_X \varphi d\mu.$$

对于一般的可测函数 $f : X \to \mathbb{C}$，如果 $\int_X |f| d\mu < \infty$，我们就称 f 为**μ-可积的**或**可积的**。

当 $f : X \to \mathbb{R}$ 是实值可积函数时，我们按照它的正负两部分来定义积分：

$$\int_X f d\mu = \int_X f^+ d\mu - \int_X f^- d\mu,$$

其中，$f = f^+ - f^-$，它们的定义为

$$f^+(x) = f(x)\mathbf{1}_{\{x | f(x) \geqslant 0\}}, \quad f^-(x) = -f(x)\mathbf{1}_{\{x | f(x) \leqslant 0\}}.$$

当 $f : X \to \mathbb{C}$ 是复值可积函数时，我们按其实部和虚部来定义积分：

$$\int_X f d\mu = \int_X \mathrm{Re}(f) d\mu + i \int_X \mathrm{Im}(f) d\mu.$$

我们先来证明两个简单的性质用来熟悉一下定义：

1) 如果实值函数 f 可积，那么 $f^{\pm}$ 可积并且它们的积分都是有限的。

实际上，我们有 $f^{\pm} \leqslant |f|$，所以只要证明下面的 2) 即可。

2) 如果 f 和 g 是非负函数并且对任意的 $x \in X$，$f(x) \leqslant g(x)$，那么

$$\int_X f d\mu \leqslant \int_X g d\mu.$$

我们需要用积分最原始的定义，首先考虑如下两个集合

$$\mathcal{E}_f = \big\{\varphi \in \mathcal{E}(X) \big| 0 \leqslant \varphi \leqslant f\big\},$$
$$\mathcal{E}_g = \big\{\varphi \in \mathcal{E}(X) \big| 0 \leqslant \varphi \leqslant g\big\}.$$

由于 $f \leqslant g$，所以，$\mathcal{E}_f \subset \mathcal{E}_g$，从而，

$$\sup_{\varphi \in \mathcal{E}_f} \int_X \varphi d\mu \leqslant \sup_{\varphi \in \mathcal{E}_g} \int_X \varphi d\mu.$$

所以，$\int_X f d\mu \leqslant \int_X g d\mu$。

注记 (零测集)**.** 积分理论中人们最常用的一句黑话就是"几乎处处"。如果一个 $A \in \mathcal{A}$ 的测度为零，即 $\mu(A) = 0$，我们称它为**零测集**。我们现在考虑某个性质 **(P)**，如果存在零测集 A，使得

$$\big\{x \in X \big| \textbf{(P)} \text{ 在 } x \text{ 处不成立}\big\} \subset A,$$

我们就称 **(P) 几乎处处**成立。由于上述定义中我们考察了一个零测集的子集，所以定义更适合的场合是要求 $(X, \mathcal{A}, \mu)$ 是完备的测度空间，即要求它的每个零测集的子集仍为零测集。这一点要求相对于本课程来说也是零测集，忽略它对理解课程没有影响。

我们有如下常用的事实:

练习. 可数个零测集的并还是零测集。

证明留给同学。我们现在证明如下命题:

命题 279. 假设可测函数 $f : (X, \mathcal{A}, \mu) \to \mathbb{C}$ 几乎处处为零，即 $\mu\big(\big\{x \in X \big| f(x) \neq 0\big\}\big) = 0$，那么

$$\int_X f d\mu = 0.$$

在证明之前，我们首先指出集合 $\big\{x \in X \big| f(x) \neq 0\big\}$ 是可测的，因为 f 是可测的。下面的证明的步骤是标准的，即先研究正函数，再研究实值函数，再研究复值函数，一如积分的定义。

证明: 首先假设 f 为正函数，即对任意的 $x \in X$，$f(x) \geqslant 0$。我们用反证法。如果不然，假设 $\int_X f d\mu > 0$，那么由于

$$\int_X f d\mu := \sup_{\substack{\varphi \in \mathcal{E}(X) \\ 0 \leqslant \varphi \leqslant f}} \int_X \varphi d\mu,$$

所以，存在简单函数 $0 \leqslant \varphi \leqslant f$，使得 $\int_X \varphi d\mu > 0$。这显然是不可能的，否则，根据简单函数积分的定义，我们有

$$\int_X f(x)d\mu = \sum_{i=0}^{m} \lambda\mu(f^{-1}(\lambda)) > 0.$$

所以，有某个 $\lambda > 0$，使得 $\mu(f^{-1}(\lambda)) > 0$，所以在测度非零集合 $f^{-1}(\lambda)$ 上，

$$f(x) \geqslant \varphi(x) = \lambda > 0.$$

这与 f 几乎处处是零矛盾。

如果 f 是实值函数，我们把 f 写成 $f = f_+ - f_-$，其中 $f_\pm$ 分别为 f 的正部和负部分。很明显，f 几乎处处为零意味着 $f_\pm$ 几乎处处为零，从而

$$\int_X f d\mu = \int_X f_+ d\mu - \int_X f_0 d\mu = 0 - 0 = 0.$$

如果 f 是复值函数，我们把 f 写成 $f = \mathrm{Re}(f) + \mathrm{Im}(f)$，其中 $\mathrm{Re}(f)$ 和 $\mathrm{Im}(f)$ 分别为 f 的实部和虚部。那么，f 几乎处处为零意味着 $\mathrm{Re}(f)$ 和 $\mathrm{Im}(f)$ 都几乎处处为零。所以，

$$\int_X f d\mu = \int_X \mathrm{Re}(f) d\mu + i \int_X \mathrm{Im}(f) d\mu = 0 + 0 = 0.$$

这个证明本质上是在运用积分的线性（目前还没有证明）。 □

我们现在研究正函数的积分。下面的 Beppo Levi 定理是积分理论中最重要的定理之一，在证明之前，我们先回忆测度的一个重要性质：它与单调上升序列的极限可以交换，即若 $\{A_i\}_{i\geqslant 1}$ 是给 $\mathcal{A}$ 中的上升序列，那么，

$$\mu(\lim_{i\to\infty} A_i) = \lim_{i\to\infty} \mu(A_i).$$

我们用积分的语言来讲这件事情，为此，我们要把上面的每一个对象都“函数化”（本来它们是集合）。令 $f_i(x) = \mathbf{1}_{A_i}(x)$，这个集合的序列是上升的指的是函数列 $\{f_i\}_{i\geqslant 1}$ 是上升的，即对任意的 $x \in X$，数列 $\{f_i(x)\}_{i\geqslant 1}$ 是递增的。根据积分的性质，我们有

$$\mu(A_i) = \int_X f_i d\mu.$$

令 $A = \lim\limits_{i\to\infty} A_i$ 并且令 $f(x) = \mathbf{1}_A(x)$，那么，这个集合的极限可以做如下的翻译：对任意的点 $x \in X$，我们都有

$$\lim_{i\to\infty} f_i(x) = f(x),$$

即上升的正函数列 $\{f_i\}_{i\geqslant 1}$ 逐点地收敛到 f。所以，上面的极限说的是如果上升的正函数列 $\{f_i\}_{i\geqslant 1}$ 逐点地收敛到 f，那么，积分与极限可交换：

$$\lim_{i\to\infty} \int_X f_i d\mu = \int_X f d\mu.$$

这就是 Beppo Levi 定理的内容，它把测度与极限可交换性“函数化”了（我们更喜欢函数是因为在函数空间上我们可以做更多的“线性”操作）。

定理 280 (Beppo Levi)**.** 假设 $\{f_n\}_{n\geqslant 1}$ 为是 $(X,\mathcal{A},\mu)$ 上定义的上升的正函数列，即对任意的 $x\in X$，对任意的 $i\geqslant 1$，我们有

$$f_i(x)\leqslant f_{p+1}(x),\quad \forall i,\quad \forall x.$$

对任意的 $x\in X$，令 $f(x)=\lim\limits_{i\to\infty} f_i(x)$（可以取 $+\infty$），我们通常它简写为 $f_i(x)\nearrow f(x)$。那么，f 为可测函数并且

$$\lim_{i\to\infty}\int_X f_i d\mu=\int_X f d\mu.$$

特别地，对于任意的正函数列 $\{g_i\}_{i\geqslant 1}$，函数级数 $\sum\limits_{i=1}^{\infty} g_i$ 是 X 上良好定义的可测正函数并且

$$\int_X\sum_{i=1}^{\infty} g_i d\mu=\sum_{i=1}^{\infty}\int_X g_i d\mu.$$

证明: 由于对任意的 $i\geqslant 1$，我们有 $f(x)\geqslant f_i(x)$，所以（已经证明），

$$\lim_{i\to\infty}\int_X f_i d\mu\leqslant\int_X f d\mu.$$

我们只要证明下面的不等式即可:

$$\lim_{i\to\infty}\int_X f_i d\mu\geqslant\int_X f d\mu.$$

利用积分的定义，对任意的 $\varepsilon>0$，选取简单函数 $\varphi(x)=\sum\limits_{i=0}^{m}\lambda_i\mathbf{1}_{X_i}\leqslant f(x)$，其中 $0=\lambda_0<\lambda_1<\lambda_2<\cdots<\cdots<\lambda_m$ 并且当 $i\neq 0$ 时，$\mu(X_i)<\infty$。我们还可以假设 $\bigcup\limits_{i\leqslant m} X_i=X$ 并且

$$|\int_X f d\mu-\int_X\varphi d\mu|<\frac{1}{2}\varepsilon.$$

根据 $\varphi(x)\leqslant f(x)$，所以在每个 X_i 上，$f(x)\geqslant\lambda_i$。令 $A=\sum\limits_{i=1}^{m}\mu(X_i)$。如果 $i\neq 0$，我们可以将 φ 中的 λ_i 替换成 $\lambda_i-\dfrac{\varepsilon}{2A}$。此时，在每个 X_i 上 $f(x)>\lambda_i$ 并且

$$|\int_X f d\mu-\int_X\varphi d\mu|<\varepsilon.$$

现在，对每个指标 $i\geqslant 0$，定义集合

$$A_i=\big\{x\big|f_i(x)\leqslant\varphi(x)\big\}.$$

由于 $A_i\subset\bigcup\limits_{k=1}^{n} X_k$，所以 $\mu(A_i)$ 的测度有限。另外，根据 $\{f_i\}_{i\geqslant 1}$ 是上升的，所以 $\{A_i\}_{i\geqslant 1}$ 是下降的子集序列。由于在 $f(x)\neq 0$ 处，$f(x)>\varphi(x)$ 并且 $f(x)=\lim\limits_{i\to\infty} f_i(x)$，所以

$$\lim_{i\to\infty} A_i=\emptyset\quad\Leftrightarrow\quad\lim_{i\to\infty}(A_i)^c=X.$$

对固定的指标 i，我们定义简单函数

$$\varphi_i(x) = \mathbf{1}_{(A_i)^c}(x)\varphi(x).$$

很明显，我们有 $f_i \geqslant \varphi_i$。按照积分的定义，我们有

$$\int_X \varphi_i d\mu \geqslant \int_X \sum_{k=1}^{n} \lambda_k \mathbf{1}_{X_k - A_i} d\mu = \sum_{i=1}^{n} \lambda_k \mu(X_k - A_i).$$

当 $i \to \infty$，我们可以用集合取极限和测度可交换的性质，所以右边的极限恰好就是 $\int_X \varphi d\mu$。从而，

$$\int_X \varphi_i d\mu \geqslant \int_X \varphi d\mu.$$

再根据

$$\Big| \int_X f d\mu - \int_X \varphi d\mu \Big| < \varepsilon$$

并且 ε 是任选的，不等式得证。

函数级数的情形我们留作作业。 □

推论 281. *对正函数可测函数 $f : (X, \mathcal{A}, \mu) \to [0, +\infty]$，存在上升的简单正函数序列 $\{f_i\}_{i\geqslant 1}$，使得对每个 $x \in X$，我们都有 $f(x) = \lim\limits_{i\to\infty} f_i(x)$。特别地（根据 Beppo Levi），*

$$\lim_{i\to\infty} \int_X f_i d\mu = \int_X f d\mu.$$

证明： 由于 $(X, \mathcal{A}, \mu)$ 是 σ-有限的，我们选取上升序列 $\{X_i\}_{i\geqslant 1}$，使得对每个 i 都有 $\mu(X_i) < \infty$ 并且 $\bigcup\limits_{i\geqslant 1} X_i = X$。我们定义

$$\varphi_i(x) = \begin{cases} 0, & x \notin X_i; \\ i, & x \in X_i \text{ 且 } f(x) \geqslant i; \\ \frac{k}{2^i}, & x \in X_i, f(x) < i \text{ 且 } \frac{k}{2^i} \leqslant f(x) < \frac{k+1}{2^i}. \end{cases}$$

这是上升的函数序列。对任给定的 $x \in X$，不妨设 $x \in X_{i_0}$ 并且 $f(x) < i_0$。根据 φ_i 的定义，对任意的 $i \geqslant i_0$，我们有

$$0 \leqslant f(x) - \varphi_i(x) \leqslant 2^{-i}.$$

这说明 $\{\varphi_i\}_{i\geqslant 1}$ 逐点收敛到 f。 □

注记. *用一列上升的简单函数来逐点地逼近正函数是积分理论中非常很有用的技巧，因为在逐点逼近的同时，积分也收敛（Beppo Levi）。我们今后可以看到，这个技巧可以把个关于函数的问题转化为关于简单函数的问题，从而极大地简化了证明。*

注记 (对比 Riemann 积分). 考虑 $\mathbb{R}^n$ 中的开集 Ω，$f:\Omega\to\mathbb{R}$ 是连续的正函数。可以证明，Ω 总是可以分解为一些（可数个）闭方块的并，即

$$\Omega=\bigcup_{k=1}^{\infty}C_k.$$

在研究 *Lebesgue* 测度的平移不变性时，我们证明了方块的边界的测度是零。上面这些方块有可能会相交，但它们只能在边界处相交，由于边界是零测集，这不会影响 f 的积分的值。此时，我们可以考虑 *Darboux* 下和的类比，即如下简单函数的积分：

$$\varphi(x)=\sum_{k=1}^{\infty}\left(\inf_{x\in C_k}f(x)\right)\cdot\mathbf{1}_{C_k}(x).$$

当我们让方块越来越小的时候（类比为分划的加细），我们自然期盼 $\varphi(x)$ 给出上升到 f 的简单函数序列。根据 *Beppo Levi* 定理，我们就用 $\varphi(x)$ 的积分的极限来定义 f 的积分。除去一些枝节的论证，这基本上就是 *Riemann* 积分的定义。

另外，*Riemann* 意义下的积分需要对 Ω 的分划作比较严格的要求，而我们有了更大的自由，对于简单函数所对应的集合不做太多的要求。

41.1 作业：子流形与零测集，Stieltjies 测度的构造，Borel-Cantelli 定理和无理数的逼近

清华大学 19-20 春季学期，数学分析二，作业 5

请用 A4 大小的纸张**正反面**用钢笔，签字笔或者圆珠笔书写，并注明自己的姓名，年级和作业的**总页数**。除定理公式所涉及的人名之外，请使用**中文**。本次作业的提交时间和地点为 **2 月 57 日**上午的课堂上，逾期视作零分。

课堂基本概念：σ-代数与可测映射

A0) 给定映射 $f : X \to Y$，验证如下集合论的基本事实：由于可数个（这个条件没有必要）Y 的子集 $\{B_i\}_{i\geqslant 1}$，我们都有

$$f^{-1}(\bigcup_{i\geqslant 1} B_i) = \bigcup_{i\geqslant 1} f^{-1}(B_i),\quad f^{-1}(\bigcap_{i\geqslant 1} B_i) = \bigcap_{i\geqslant 1} f^{-1}(B_i).$$

A1) 假设 $U_1 \in \mathbb{R}^m$，$U_2 \in \mathbb{R}^n$ 是开集，$F_1 \in \mathbb{R}^m$，$F_2 \in \mathbb{R}^n$ 是闭集。证明，在 $\mathbb{R}^{m+n}$ 中，

$$U_1 \times U_2 = \big\{(x,y)\big|x \in U_1, y \in U_2\big\},\quad F_1 \times F_2 = \big\{(x,y)\big|x \in F_1, y \in F_2\big\}$$

分别也是开集和闭集。

A2) J 是一个指标集。如果对每个 $j \in J$，$\mathcal{M}_j$ 都是 X 上的单调类，那么

$$\mathcal{M} := \bigcap_{j\in J} \mathcal{M}_j$$

也是 X 上的单调类

A3) $(Y, \mathcal{B})$ 是可测空间，X 是集合，$f : X \to Y$ 是映射。证明，

$$f^*\mathcal{B} = f^{-1}(\mathcal{B}) = \{f^{-1}(B)|B \in \mathcal{B}\}.$$

是 X 上的 σ-代数。利用这个结论证明如下的结论：如果 $Y' \subset Y$ 是子集，那么，

$$\mathcal{B}\Big|_{Y'} = \big\{B \cap Y'\big|B \in \mathcal{B}\big\}$$

是 Y' 上的 σ-代数（我们称它为 $\mathcal{B}$ 在 Y' 上的**限制**）。

A4) $(X, \mathcal{A})$ 是可测空间，f 和 g 是 X 上的可测函数。证明，$|f|$ 是可测映射函数；令 $X_g = \{x \in X|g(x) \neq 0\}$，证明，函数 $\dfrac{f}{g}$ 是 $\left(X_g, \mathcal{A}\Big|_{X_g}\right)$ 上的可测函数。

A5) $(X, \mathcal{A})$ 是可测空间，$f : X \to Y$ 是映射，令

$$\mathcal{B} = \big\{B \subset Y\big|f^{-1}(B) \in \mathcal{A}\big\},$$

证明，$\mathcal{B}$ 为 σ-代数。

A6) (X_1, d_1) 和 (X_2, d_2) 是距离空间，(X, d) 是它们的乘积距离空间。证明，单位映射

$$\iota : (X, \mathcal{B}_X) \to (X, \mathcal{B}_{X_1} \otimes \mathcal{B}_{X_2}), \quad x \mapsto x$$

是可测的。进一步证明 $\mathcal{B}_{X_1} \otimes \mathcal{B}_{X_2} \subset \mathcal{B}_X$。

A7) 给定测度空间 $(X, \mathcal{A}, \mu)$，$\{A_i\}_{i\geqslant 1}$ 是 $\mathcal{A}$ 中下降的序列。假设对某个 i_0 有 $\mu(A_{i_0}) < \infty$，试证明

$$\mu(\lim_{i\to\infty} A_i) = \lim_{i\to\infty} \mu(A_i).$$

A8) 给定测度空间 $(X, \mathcal{A}, \mu)$ 和可测空间 $(Y, \mathcal{B})$ 以及它们之间的可测映射 $f : X \to Y$，对每个 $B \in \mathcal{B}$，我们定义

$$(f_*\mu)(B) := \mu\big(f^{-1}(B)\big).$$

证明，$f_*\mu$ 是 $(Y, \mathcal{B})$ 上的测度。我们将 $f_*\mu$ 称作是 μ **在** f **下的像测度**。

A9) $(X, \mathcal{A})$ 是可测空间，给定 X 中可数个点 $\{x_k\}_{k\geqslant 1}$ 和正数的序列 $\{a_k\}_{k\geqslant 1}$。对任意的 $A \in \mathcal{A}$，我们定义

$$\mu(A) = \left(\sum_{k=1}^{\infty} a_k \delta_{x_k}\right)(A) = \sum_{k=1}^{\infty} a_k \delta_{x_k}(A).$$

其中，δ_{x_k} 是 x_k 点定义的 Dirac 测度。证明，μ 是 $(X, \mathcal{A})$ 上的测度。

A10) 给定 $S \in \mathcal{A}(\mathbb{R}^n)$，$S$ 有多种方式写成两两不交的方块的并，例如

$$S = C_1 \cup C_2 \cup \cdots \cup C_m, \quad S = C_1' \cup C_2' \cup \cdots \cup C_{m'}'.$$

证明，

$$\sum_{i\leqslant m} m(C_i) = \sum_{i'\leqslant m'} m(C_{i'}').$$

（这验证了课堂上关于 $m : \mathcal{A}(\mathbb{R}^n) \to [0, \infty]$ 是良好定义的论断）

A11)（经典必做题目）对任意正交矩阵 $R \in \mathbf{O}(n)$，我们定义 $\mathbb{R}^n$ 上的旋转变换：

$$R : \mathbb{R}^n \to \mathbb{R}^n, \quad x \mapsto R \cdot x.$$

其中，$R \cdot x$ 是矩阵和列向量的乘法。那么，我们有 $R_* m = m$，即对任意的 $A \in \mathcal{B}(\mathbb{R}^n)$，我们有 $m\,(R(A)) = m(A)$。

A12)（必须熟记的结论：紧集的区间套原理）假设 $\{K_i\}_{i\geqslant 1}$ 是 $\mathbb{R}^n$ 中一列下降的紧集并且 $\cap_{i\geqslant 1} K_i = \emptyset$。证明，存在 i_0，使得当 $i \geqslant i_0$ 时，$K_i = \emptyset$。

A13) 对于 $\mathbb{R}^n$，我们用 Euclid 距离所定义的拓扑。证明，

$$\mathcal{B}(\mathbb{R}^n) = \underbrace{\mathcal{B}(\mathbb{R}) \otimes \mathcal{B}(\mathbb{R}) \otimes \cdots \otimes \mathcal{B}(\mathbb{R})}_{n\ 个}.$$

可测函数

证明如下关于可测函数的结论，其中我们假设 $(X, \mathcal{A})$ 是可测空间。

B1)（可测函数的极限也可测）给定实值可测函数的序列 $\{f_i\}_{i\geqslant 1}$，其中对于每个 i，$f_i : X \to \mathbb{R}\}$ 是可测函数。假设该序列逐点收敛，即对任意的 $x \in X$，都有 $\lim\limits_{i\to\infty} f_i(x) = f(x)$。对任意的 $c \in \mathbb{R}$，$n \in \mathbb{Z}_{\geqslant 1}$，我们令

$$U(c) = (c, \infty), \quad U(c)_n = (c + \frac{1}{n}, \infty).$$

证明，

$$f^{-1}(U(c)) = \bigcup_{n=1}^{\infty} f^{-1}(U(c)_n) = \bigcup_{n=1}^{\infty} \bigcup_{m=1}^{\infty} \bigcap_{q\geqslant m} f_q^{-1}(U(c)_n),$$

据此证明 f 是可测的。

B2) 我们在 $\mathbb{R}$ 配备 Borel-代数。证明，如果 $f : \mathbb{R} \to \mathbb{R}$ 是单调函数，那么 f 是可测的。

B3) 假设 $f : (X, \mathcal{A}) \to \mathbb{R}$ 和 $g : (X, \mathcal{A}) \to \mathbb{R}$ 均为可测函数，我们定义

$$\begin{aligned} A_< &= \big\{x \in X \big| f(x) < g(x)\big\}, \\ A_> &= \big\{x \in X \big| f(x) > g(x)\big\}, \\ A_= &= \big\{x \in X \big| f(x) = g(x)\big\}. \end{aligned}$$

证明，$A_<, A_>, A_+ \in \mathcal{A}$。

B4) 假设 $\{f_n : (X, \mathcal{A}) \to \mathbb{R}\}_{n\geqslant 1}$ 是一列可测实值函数，我们定义

$$\begin{aligned} A_\infty &= \big\{x \in X \big| \lim_{n\to\infty} f_n(x) = +\infty\big\}, \\ A_b &= \big\{x \in X \big| \text{数列 } \{f_n(x)\}_{n\geqslant} \text{ 是有界的}\big\}. \end{aligned}$$

证明，$A_\infty, A_b \in \mathcal{A}$。

子流形与零测集

考虑 $E \subset \mathbb{R}^n$，如果对任意的 $\varepsilon > 0$，存在可数个方块 C_k，使得

$$E \subset \bigcup_{k=1}^{\infty} R_k \quad \text{并且} \quad \sum_{k=1}^{\infty} |R_k| < \varepsilon,$$

我们就称 E 是（$\mathbb{R}^n$ 中的）**零测集**。（这个定义没用到我们刚刚学过的测度理论）

C1) 证明，可数个 $\mathbb{R}^n$ 中的零测集的并还是 $\mathbb{R}^n$ 中的零测集。

C2) 举例说明，$\mathbb{R}^n$ 中的零测集可以是稠密的。

C3) 证明，$\mathbb{R}^n$ 中零测集在微分同胚下的像也是零测集，即若 $\Phi:\mathbb{R}^n\to\mathbb{R}^n$ 是微分同胚，$E\subset\mathbb{R}^n$ 是 $\mathbb{R}^n$ 中的零测集，那么 $\Phi(E)$ 也是 $\mathbb{R}^n$ 中的零测集。

C4) 假设 $M\subset\mathbb{R}^n$ 是子流形并且 $\mathrm{codim}M\geqslant 1$。证明，$M$ 是 $\mathbb{R}^n$ 中的零测集。

C5) 如果 $f:\mathbb{R}\to\mathbb{R}$ 是连续函数，我们用 $\Gamma_f=\{(x,y)\in\mathbb{R}^2|y=f(x)\}$ 表示它的图像。证明，Γ_f 是 $\mathbb{R}^2$ 中的零测集。

C6)* 给定 n 个变元的实系数非零多项式 $F(x_1,\cdots,x_n)$。证明，$F^{-1}(0)$ 是 $\mathbb{R}^n$ 中的零测集。

Carathéodory 测度扩张定理应用举例：Stieltjies 测度的构造

如果 $(\mathbb{R},\mathcal{B}(\mathbb{R}))$ 上的测度 μ 在所有紧集上都有限，即对任意的有界闭集 $K\subset\mathbb{R}$，$\mu(K)<\infty$，我们就称 μ 为 Stieltjies 测度。我们要刻画 $(\mathbb{R},\mathcal{B}(\mathbb{R}))$ 上的所有 Stieltjes 测度。

D1) 假设 μ 为 Stieltjies 测度。如果点 $x\in\mathbb{R}$ 满足 $\mu(\{x\})>0$，我们就称 x 是 μ 的一个**原子**。证明，μ 只有可数个原子。

D2) 假设 μ 为 Stieltjies 测度，我们定义 $\mathbb{R}$ 上的函数 F：

$$F(x)=\begin{cases}\mu([0,x)), & x\geqslant 0;\\ -\mu([x,0)), & x<0.\end{cases}$$

证明，F 是 $\mathbb{R}$ 上增函数并且是左连续的。

D3) 证明，$x\in\mathbb{R}$ 是 F 的不连续点当且仅当 x 是原子。据此，给出 D1) 的一个新证明。

D4) 证明，任给左连续的增函数 $F(x)$，存在唯一的 $(\mathbb{R},\mathcal{B}(\mathbb{R}))$ 上的测度 μ，使得

$$\mu([a,b))=F(b)-F(a).$$

D5) 证明，上述构造的 μ 是 Stieltjies 测度。

（这个习题把 $(\mathbb{R},\mathcal{B}(\mathbb{R}))$ 上的 Stieltjes 测度的空间等价为 $\mathbb{R}^1$ 上的左连续的增函数的空间，同学们一应该对照上学期所学的 Stieltjes 积分）

测度理论的其它有趣应用举例一：Borel-Cantelli 定理和无理数的逼近

给定 X 中可数个子集 $\{A_n\}_{n\geqslant 1}$，我们定义：

$$A^{\infty}=\{x|x \text{ 属于无穷多个 } A_n\},\quad A^{\infty+}=\{x|\text{只有有限多个 } A_n\text{，使得 } x\notin A_n\}.$$

很明显，我们有 $\bigcap\limits_n A_n\subset A^{\infty+}\subset A^{\infty}\subset\bigcup\limits_n A_n$。

E1) 证明，$A^{\infty}=\bigcap\limits_{n=1}^{\infty}\bigcup\limits_{m\geqslant n}A_m$，$A^{\infty+}=\bigcup\limits_{n=1}^{\infty}\bigcap\limits_{m\geqslant n}A_m$。

E2) 证明，如果序列 A_n 是上升或者下降的，那么 $A^{\infty}=A^{\infty+}=\lim\limits_{n\to\infty}A_n$。

E3) 令 $B_n=(A_n)^c=X-A_n$。证明，$B^{\infty}=\left(A^{\infty+}\right)^c$ 和 $A^{\infty}=\left(B^{\infty+}\right)^c$。

从此往后，假设 $(X,\mathcal{A},\mu)$ 是测度空间并且 $\{A_n\}\subset\mathcal{A}$。

E4) 证明，$A^{\infty}\in\mathcal{A}$ 和 $A^{\infty+}\in\mathcal{A}$。

E5) 证明如下不等式：

$$\lim_{n\to\infty}\mu(\bigcap_{m\geqslant n}A_m)=\mu(A^{\infty+})\leqslant\mu(A^{\infty})\leqslant\lim_{n\to\infty}\sum_{m\geqslant n}\mu(A_m).$$

E6) 证明 **Borel-Cantelli 定理**：如果 $\{A_n\}\subset\mathcal{A}$ 满足 $\sum\limits_n\mu(A_n)<\infty$，那么对几乎所有的 $x\in X$，它只出现在有限个 A_n 中，即出现在无限多个 A_n 中的点的测度是零：$\mu(A^{\infty})=0$。

E7) 证明，如果存在 n_0 使得 $\mu(\bigcup\limits_{m\geqslant n_0}A_m)<\infty$，那么有

$$\mu(A^{\infty})=\lim_{n\to\infty}\mu(\bigcup_{m\geqslant n}A_m).$$

E8) 证明（Dirichlet 的一个定理），对每个实数 α，有无限多不同的有理数 $\frac{p}{q}$，使得

$$|\alpha-\frac{p}{q}|<\frac{1}{q^2}.$$

（提示：用抽屉原理，请查阅文献。Hurwitz 对于上述逼近有一个改进，可以证明有无限多不同的有理数 $\frac{p}{q}$，使得 $|\alpha-\frac{p}{q}|<\frac{1}{\sqrt{5}q^2}$。我们下面要说明这个逼近的精度是最好的。）

E9) 我们用 I 表示开区间 $(0,1)$，用 m 表示 I 上的 Lebesgue 测度。假设正实数的序列 $\{a_q\}_{q\geqslant 1}$ 满足 $\sum\limits_q\frac{1}{a_q}<\infty$。对每个 q，定义 Borel 集

$$A_q=I\cap\Big(\bigcup_{0\leqslant p\leqslant q}(\frac{p}{q}-\frac{1}{q\cdot a_q},\frac{p}{q}+\frac{1}{q\cdot a_q})\Big).$$

证明，$\sum\limits_q m(A_q)<\infty$。

E10) 序列 $\{a_q\}_{q\geqslant 1}$ 如上所述。令 $D=\left\{\alpha\in I\middle|\text{存在无穷多个有理数 }\frac{p}{q},\text{ 使得 }|\alpha-\frac{p}{q}|<\frac{1}{q\cdot a_q}\right\}$。证明，$D$ 是零测集。

E11) 任给 $\varepsilon>0$ 和 $a>0$。考虑满足下面性质的实数 α：存在无穷多个有理数 $\frac{p}{q}$，使得 $|\alpha-\frac{p}{q}|<\frac{a}{q^{2+\varepsilon}}$。证明，这样的实数的测度为零。（这表明 Dirichlet 定理给出的逼近在某种意义上是最佳的）

I remember one occasion when I tried to add a little seasoning to a review, but I wasn't allowed to. The paper was by Dorothy Maharam, and it was a perfectly sound contribution to abstract measure theory. The domains of the underlying measures were not sets but elements of more general Boolean algebras, and their range consisted not of positive numbers but of certain abstract equivalence classes. My proposed first sentence was: "The author discusses valueless measures in pointless spaces."

—— *I want to be a Mathematician* by Paul R. Halmos

---

42 Lebesgue 积分与 Riemann 积分之间的关联，可积函数空间，积分的线性，L^1-空间，Fatou 引理，Lebesgue 控制收敛定理，积分对参数的连续依赖性

二零二零年三月三十日，星期一，晴

抽象积分理论的一些具体例子与讨论

给定测度空间 $(X,\mathcal{A},\mu)$，正可测函数 $f:X\to[0,+\infty]$ 的积分被定义为：

$$\int_X f d\mu := \sup_{\substack{\varphi\in\mathcal{E}(X)\\ 0\leqslant\varphi\leqslant f}} \int_X \varphi d\mu.$$

对于一般的可测函数 $f:X\to\mathbb{C}$，如果 $\int_X |f| d\mu<\infty$，我们就说 f 是可积的。

我们先看一个例子：

例子. 假设 $X=\mathbb{Z}_{\geqslant 1}=\{1,2,\cdots,n,\cdots\}$，$\mathcal{A}=\mathcal{P}(X)$ 是 X 上所有的子集所组成的 σ-代数。我们考虑数集合元素个数的测度：

$$\mu:\mathbb{Z}_{\geqslant 1}\to[0,\infty],\quad A\mapsto\mu(A)=|A|.$$

从而，我们得到了测度空间 $\big(\mathbb{Z}_{\geqslant 1},\mathcal{P}(\mathbb{Z}_{\geqslant 1},\mu)\big)$。这个空间上函数

$$f:\mathbb{Z}_{\geqslant 1}\to\mathbb{C},\quad n\mapsto f(n),$$

就是一个复数的数列。很明显，任意这样的函数都是可测的（因为 $\mathcal{A}=\mathcal{P}(X)$）。我们将要在本次作业中证明，$f$ 可积分当且仅当

$$\sum_{n=1}^{\infty}|f(n)|<\infty$$

并且此时

$$\int_{\mathbb{Z}_{\geqslant 1}} f d\mu=\sum_{n=1}^{\infty} f(n).$$

所以，我们所熟悉的级数求和实际上是一种积分理论。

类似于上学期我们所学的关于级数收敛或者反常积分的判别法，我们对于积分的收敛性（即一个函数是否可积分）有如下的判别准则：

命题 282. f 和 h 是测度空间 $(X,\mathcal{A},\mu)$ 上的可测函数。如果 h 正的可积函数，并且不等式

$$|f(x)|\leqslant h(x)$$

几乎处处成立，那么 f 是可积函数。

证明: 令 $B=\{x\in X\big|f(x)>g(x)\}$，按照命题中的要求，$\mu(B)=0$。我们令 $f_1(x)=f(x)\mathbf{1}_{B^c}(x)$，那么，对所有的 $x\in X$，我们都有

$$|f_1(x)|\leqslant h(x).$$

作为正函数，我们自然有

$$\int_X|f_1|d\mu\leqslant\int_X|h|d\mu<\infty.$$

所以 $f_1(x)$ 是可积的。为了说明 $f(x)$ 是可积分的，我们比较 $|f(x)|$ 与 $|f_1(x)|$: 对于任意的 x，都有 $|f_1(x)|\leqslant|f(x)|$，所以

$$\int_X|f_1|d\mu\leqslant\int_X|f|d\mu.$$

这两个函数只在零测集 B 上有差别。按照积分的定义，我们有

$$\int_X|f|d\mu:=\sup_{\substack{\varphi\in\mathcal{E}(X)\\0\leqslant\varphi\leqslant|f|}}\int_X\varphi d\mu.$$

对于简单函数而言，我们在一个零测集上改变函数值（全改为零）是不影响它的积分的，这可以用简单函数的积分定义直接看出，所以，我们总是可以假设上面积分定义中的函数 φ 在 B 上取零。此时，$\varphi\in\mathcal{E}(X)$ 并且 $0\leqslant\varphi\leqslant|f|$ 意味着 $\varphi\in\mathcal{E}(X)$ 并且 $0\leqslant\varphi\leqslant|f_1|$，这说明

$$\int_X|f_1|d\mu=\int_X|f|d\mu.$$

这就给出了命题的证明。 □

尽管我们现在还没有足够好的工作计算积分（目前只会计算简单函数的积分!），但是通过这个比较的判别法，我们可以判断函数的可积性。下面的几个例子很有启发性:

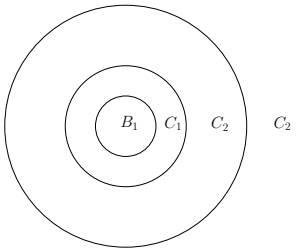
例子. *1)* 给定 $\mathbb{R}^n$ 上的可测函数 f（我们用 *Lebesgue* 测度），假设对几乎处处的 $x\in\mathbb{R}^n$，f 满足如下的控制:

$$|f(x)|\leqslant\frac{C}{1+|x|^{n+\varepsilon}},$$

其中 $\varepsilon>0$。那么，f 是可积函数。

根据上面的判断法则，我们只要证明函数 $f(x)=\dfrac{C}{1+|x|^{n+\varepsilon}}$ 可积即可。由于工具的限制，这个性质的证明目前并不简单，我们这里给出证明的细节（借此机会复习 *Lebesgue* 测度的积分性质）。我们把 $\mathbb{R}^n$ 分成一些环面的并（这个分解和所谓的 *Littlewood-Paley* 分解相关，我们将在数学分析三（如果存在的话）中学习）。令 $B_R=\{x\in\mathbb{R}^n\big||x|\leqslant R\}$，我们令 $C_k=B_{2^{k+1}}-B_{2^k}$，那么，

$$\mathbb{R}^n=B_1\cup C_1\cup C_2\cup\cdots\cup C_k\cup\cdots$$



尽管我们现在还不能计算这些换面或者球面的体积，但是，由于 B_1 和 C_1 都被一个边长为 4 的方块覆盖，所以我们知道存在常数 a，使得 $m(C_1)+m(B)\leqslant a$。另外，每个 C_k 都可以由 C_1 通过伸缩变换

$$\mathbb{R}^n\to\mathbb{R}^n,\quad x\mapsto 2^k x,$$

得到，按照 *Lebesgue* 测度在伸缩变换下的性质，我们知道

$$m(C_k)=(2^k)^n m(C_1)\leqslant 2^{kn}a.$$

在每个 C_k 上，我们有

$$f(x)=\frac{C}{1+|x|^{n+\varepsilon}}\leqslant\frac{C}{1+|2^k|^{n+\varepsilon}}\leqslant\frac{C}{|2^k|^{n+\varepsilon}}.$$

我们构造简单函数

$$h(x)=C\mathbf{1}_{B_1}+\sum_{k=1}^{\infty}\frac{C}{|2^k|^{n+\varepsilon}}\mathbf{1}_{C_k}.$$

很明显，$f(x)\leqslant h(x)$。而 h 的积分可以直接计算：

$$\begin{aligned}\int_{\mathbb{R}^n}h dm&=Cm(B_1)+\sum_{k=1}^{\infty}\frac{C}{|2^k|^{n+\varepsilon}}m(C_k)\\&\leqslant Cm(B_1)+\sum_{k=1}^{\infty}\frac{C}{|2^k|^{n+\varepsilon}}2^{kn}a=Cm(B_1)+Ca\sum_{k=1}^{\infty}\frac{1}{2^{k\varepsilon}}.\end{aligned}$$

上面的级数显然是有限的。

2) 我们考虑 $\mathbb{R}$ 上的函数 $f(x)=\dfrac{\sin(x)}{x}$。我们上学期证明了作为 *Riemann* 积分的反常积分，f 是可积分的。然而，在 *Lebesgue* 的意义，f 的可积性说的是 $\left|\dfrac{\sin(x)}{x}\right|$ 可积分，这个积分是无限大，我们可以仿照 *1)* 中的想法来证明，同学们会在本周的作业中完成。

下面的命题很有用，它的证明也很值得学习：

命题 283. *f 测度空间 $(X,\mathcal{A},\mu)$ 上的正可测函数。那么如下命题是等价的：*

1) f 几乎处处为零；

2) $\displaystyle\int_X f d\mu=0$。

证明: 我们上次课证明了 1)⇒2)。反之，假设 $\int_X f d\mu = 0$。对每个自然数 n，我们考虑集合

$$A_n = \{x \in X \big| f(x) \geqslant \frac{1}{n}\}.$$

根据定义，我们自然有 $f(x) \geqslant \frac{1}{n}\mathbf{1}_{A_n}(x)$。所以，我们有如下积分不等式:

$$0 = \int_X f d\mu \geqslant \int_X \frac{1}{n}\mathbf{1}_{A_n} d\mu = \frac{1}{n}\mu(A_n).$$

从而对所有 $n \geqslant 1$，$\mu(A_n) = 0$。然而，由于 f 是正函数，我们显然有

$$\{x|f(x) \neq 0\} = \{x|f(x) > 0\} = \bigcup_{n\geqslant 1} A_n.$$

所以，$\mu(\{x|f(x) \neq 0\}) = 0$，即 f 几乎处处为零。 □

在 $\mathbb{R}^1$ 上配备 Lebesgue 测度 m，如果一个可测函数 $f : \mathbb{R}^1 \to \mathbb{C}$ 在这个测度下是可积的，我们就称这个函数是 *Lebesgue 可积的*。我们可以探讨我们刚刚建立的抽象积分理论与上学期 Riemann 积分的联系。为了简单期间，我们在下面的定理中假设 f 是实值的。

定理 284. *假定 f 是闭区间 $[a, b]$ 上的 Riemann 可积的函数（有界），那么 f 是 Lebesgue 可测的并且其 Lebesgue 积分恰为其 Riemann 积分。*

证明: 我们利用 Darboux 上下和来逼近函数的积分。为此，任取 $[a, b]$ 区间的分划 $\mathfrak{I} = [x_0, x_1] \cup [x_1, x_2] \cup \cdots \cup [x_{n-1}, x_n]$，其中 $x_0 = a$，$x_n = b$，$x_0 < x_1 < \cdots < x_n$。令

$$M_k = \sup_{x\in[x_{k-1},x_k]} f(x), \quad m_k = \inf_{x\in[x_{k-1},x_k]} f(x).$$

我们现在用 Borel-可测的简单函数来代替 Darboux 上下和:

$$\overline{F}_{\mathfrak{I}}(x) = \sum_{k=1}^{n} M_k \mathbf{1}_{[x_{k-1},x_k]}, \quad \underline{F}_{\mathfrak{I}}(x) = \sum_{k=1}^{n} m_k \mathbf{1}_{[x_{k-1},x_k]}.$$

那么，

$$\int_{[a,b]} \overline{F}_{\mathfrak{I}} dm, \quad \int_{[a,b]} \underline{F}_{\mathfrak{I}} dm$$

恰好是分划 $\mathfrak{I}$ 的 Darboux 上下和，其中，上面积分是用我们的测度对简单函数进行积分（Lebesgue 积分的意义下）。

我们现在选取特殊分划：对任意的 n，分划 $\mathfrak{I}_n$ 对应为区间 $[a, b]$ 的 2^n 等分。此时，我们将对应的简单函数 $\overline{F}_{\mathfrak{I}_n}(x)$ 和 $\underline{F}_{\mathfrak{I}_n}(x)$ 简记为 $\overline{F}_n$ 和 $\underline{F}_n$。很明显，我们就得到两个单调的序列 $\{\overline{F}_n\}_{n\geqslant 1}$ 和 $\{\underline{F}_n\}_{n\geqslant 1}$。由于有界单调数列一定有极限，所以，存在可测函数 $\overline{F}$ 和 $\underline{F}$，使得

$$\overline{F}_n \searrow \overline{F}, \quad \underline{F}_n \nearrow \underline{F}.$$

由于 $\overline{F}_n \geqslant \underline{F}_n$，所以 $\overline{F} \geqslant \underline{F}$。按照 Riemann 积分的定义，上和的极限等于下和的极限（都等于该函数的积分），即

$$\lim_{n\to\infty} \int_{[a,b]} \overline{F}_n dm = \lim_{n\to\infty} \int_{[a,b]} \underline{F}_n dm.$$

根据 Beppo Levi 定理，我们就有

$$\int_{[a,b]} \overline{F} dm = \int_{[a,b]} \underline{F} dm = \underbrace{\int_a^b f(x)dx}_{\text{Riemann 积分}}.$$

所以，（利用积分的线性，即将证明）

$$\int_{[a,b]} \overline{F} - \underline{F} = 0.$$

从而，由于 $\overline{F} = \underline{F} \geqslant 0$，所以，在一个零测集之外，我们有

$$\overline{F} = \underline{F}, \text{ 几乎处处.}$$

另外，按照定义，我们还有

$$\overline{F} \geqslant f \geqslant \underline{F}$$

这说明，在差一个零测度的子集意义下（要用到测度的完备性，我们不去追究这个细节），我们有 $\overline{F} = f = \underline{F}$。所以，上述的论证说明了 f 是可测的并且它的 Riemann 积分与 Lebesgue 积分是一致的。 □

注记. 如果 f 是区间 $[a,b]$ 上可测有界的函数，它明显是可积的（抽象积分意义下），所以，*Lebesgue* 积分中有更多的可积函数。然而，我们通常关心的函数大多都有很好的连续性，所以这两种积分理论没有区别，这正是上面定理的内容。

Beppo Levi 定理的应用：Fatou 引理与 Lebesgue 控制收敛定理

我们上周课程最后讲到了 Beppo Levi 定理以及一个技术性的逼近定理（存在上升的简单函数列逼近正可测函数）。利用这个两个结论，我们有很多有趣有意义的推论：

推论 285. f 和 g 是测度空间 $(X, \mathcal{A}, \mu)$ 上的正可测函数。那么，对任意的非负实数 $a, b \in \mathbb{R}_{\geqslant 0}$，我们有

$$\int_X af + bg d\mu = a\int_X f d\mu + b\int_X g d\mu.$$

证明: 我们选取单调上升的简单函数序列 $\{f_i\}_{i\geqslant 1}$ 和 $\{g_i\}_{i\geqslant 1}$，使得它们分别逐点收敛到 f 和 g。所以，$\{af_i + bg_i\}_{i\geqslant 1}$ 也是单调上升的简单函数序列并且逐点收敛到 $af + bg$。根据积分的对于简单函数的线性，对每个 i，我们有

$$\int_X af_i + bg_i d\mu = a\int_X f_i d\mu + b\int_X g_i d\mu.$$

令 $i \to \infty$，由 Beppo Levi 定理，左右两端分别收敛到推论中要证明等式的左右两端。 □

Beppo Levi 定理的一个重要推论是 Fatou 引理：

定理 286 (Fatou)**.** $\{f_i\}_{i\geqslant 1}$ 是测度空间 $(X,\mathcal{A},\mu)$ 上的正可测函数序列。那么，我们有如下的积分不等式：

$$\int_X \liminf_{i\to\infty} f_i d\mu \leqslant \liminf_{i\to\infty} \int_X f_i d\mu,$$

其中，对任意的 $x\in X$，$\left(\liminf_{i\to\infty} f_i\right)(x) = \liminf_{i\to\infty} f_i(x)$。特别地，如果函数列 $\{f_i\}_{i\geqslant 1}$ 逐点地收敛到 f，即对任意的 x，$\lim_{i\to\infty} f_i(x) = f(x)$，那么，我们有

$$\int_X f d\mu \leqslant \liminf_{i\to\infty} \int_X f_i d\mu.$$

注记. *Fatou* 引理叙述中的不等式一般而言不能取到等号，同学们可以（作业）构造函数列 $\{f_i\}_{i\geqslant 1}$ 逐点地收敛到 f，使得

$$\int_X f d\mu < \liminf_{i\to\infty} \int_X f_i d\mu.$$

证明之前，有必要回忆“下极限”的概念：给定实数序列 $\{a_i\}_{i\geqslant 1}$，序列 $\{\inf_{i\geqslant p} a_i\}_{p\geqslant 1}$ 是递增的。我们定义

$$\liminf_{i\to\infty} a_i = \lim_{p\to\infty}\left(\inf_{i\geqslant p} a_i\right).$$

证明： 令 $f(x) = \liminf_{p\to\infty} f_p(x)$。我们定义函数列 $\{g_p(x)\}_{p\geqslant 1}$，其中

$$g_p(x) = \inf_{i\geqslant p} f_i(x).$$

那么，$\{g_p(x)\}_{p\geqslant 1}$ 为上升的正函数序列并且逐点地收敛到 $f(x)$。根据 Beppo Levi 定理，我们有

$$\lim_{p\to\infty}\int_X g_p d\mu = \int_X f d\mu.$$

然而，根据 $g_p(x)$ 的定义，对每个 $x\in X$，我们有 $g_p(x)\leqslant f_p(x)$。从而，

$$\int_X g_p d\mu \leqslant \int_X f_p d\mu.$$

这表明

$$\lim_{p\to\infty}\int_X g_p d\mu = \liminf_{p\to\infty}\int_X g_p d\mu \leqslant \liminf_{p\to\infty}\int_X f_p d\mu.$$

上面不等式左边就是 $\int_X f d\mu$，这就完成了 Fatou 引理第一部分的证明。第二部分是第一部分的直接推论。 □

定义 287 (可积函数空间与 $L^1(X,\mathcal{A},\mu)$ 空间)**.** 给定测度空间 $(X,\mathcal{A},\mu)$，我们把这个空间上所有可积函数的全体记作 $\mathcal{L}^1(X,\mathcal{A},\mu)$。和传统的微积分课程中处理的函数有所不同，在可积函数的定义中，由于函数被写成正函数的组合，这里的"函数"的取值可以取 $\pm\infty$ 或者 $\sqrt{-1}\times\pm\infty$。

这是线性子空间：对于任意的 $f,g\in\mathcal{L}^1(X,\mathcal{A},\mu)$，$\alpha,\beta\in\mathbb{C}$，按照定义，正函数 $|f|$ 和 $|g|$ 是可积分的。根据上面的推论，$|\alpha||f|+|\beta||g|$ 也是可积的。此时，我们注意到，对任意的 $x\in X$，有

$$|\alpha f(x)|+|\beta g(x)|\leqslant|\alpha||f(x)|+|\beta||g(x)|,$$

所以，$\alpha f+\beta g\in\mathcal{L}^1(X,\mathcal{A},\mu)$。

考虑几乎处处为零的函数

$$\mathcal{N}=\big\{f\in\mathcal{L}^1(X,\mathcal{A},\mu)\big|f\text{ 几乎处处为零}\big\}.$$

很明显，N 是 $\mathcal{L}^1(X,\mathcal{A},\mu)$ 的线性子空间：对任意的 $f,g\in\mathcal{N}$，按照定义，存在零测集 A 和 B，使得 $f\big|_{A^c}\equiv0$，$g\big|_{B^c}\equiv0$。那么，对任意的 $\alpha,\beta\in\mathbb{C}$，我们有

$$\alpha f+\beta g\big|_{(A\cup B)^c}\equiv0.$$

而 $A\cup B$ 还是零测集，所以，$\alpha f+\beta g\in\mathcal{N}$。

我们定义 $L^1(X,\mathcal{A},\mu)$ 空间为如下的等价类的集合（商空间）：

$$L^1(X,\mathcal{A},\mu)=\mathcal{L}^1(X,\mathcal{A},\mu)/\mathcal{N}.$$

也就是说，$L^1(X,\mathcal{A},\mu)$ 中的元素是函数的等价类，同一个类中的任两个函数之差为一个几乎处处为零的函数（这将对积分没有贡献）。我们在多元微积分的课程中不对此做进一步的解读，同学们会在实分析的课上对这个空间做深入的了解。

注记. 对于 $f\in L^1(X,\mathcal{A},\mu)$，我们把它看成一组函数的等价类 $[f]\subset\mathcal{L}^1(X,\mathcal{A},\mu)$，$f\in[f]$ 是这里面的一个代表。那么，在 $L^1(X,\mathcal{A},\mu)$ 中 $[f]=0$ 的意思是 $f\in\mathcal{N}$ 或者 $[f]=\mathcal{N}$。另外，f 几乎处处为零等价于 $|f|$ 几乎处处为零。所以，对于 $f\in L^1(X,\mathcal{A},\mu)$，$f=0$ 当且仅当 $\int_X|f|d\mu=0$。

我们现在终于可以证明积分算子的线性了：

定理 288. $\mathcal{L}^1(X,\mathcal{A},\mu)$ 为 $\mathbb{C}$-线性空间且积分算子

$$\int_X-d\mu:\mathcal{L}^1(X,\mathcal{A},\mu)\to\mathbb{C},\quad f\mapsto\int_X fd\mu,$$

是 $\mathbb{C}$-线性映射。进一步，对于 $f\in\mathcal{L}^1(X,\mathcal{A},\mu)$，我们

$$\Big|\int_X fd\mu\Big|\leqslant\int_X|f|d\mu.$$

证明: 我们已经证明了 $\mathcal{L}^1(X,\mathcal{A},\mu)$ 是 $\mathbb{C}$-线性空间。我们详细分情形来证明积分算子的线性。任选 $f,g\in\mathcal{L}^1(X,\mathcal{A},\mu)$。

1) 如果 f 和 g 都是正函数并且 $a, b \geqslant 0$，我们已经证明了

$$\int_X af + bgd\mu = a\int_X fd\mu + b\int_X gd\mu.$$

2) f 和 g 都是正函数，$a = 1$，$b = -1$ 的情况。我们要证明

$$\int_X f - gd\mu = \int_X fd\mu - \int_X gd\mu.$$

为此，令 $h = f - g$。我们可以把 h 写成正部和负部的差，即 $h = h_+ - h_-$。所以，

$$h^+ + g = h^- + f.$$

根据正函数的线性，我们得到

$$\int_X h^+d\mu + \int_X gd\mu = \int_X h^-d\mu + \int_X fd\mu.$$

再根据 $\int_X hd\mu$ 的定义，我们有

$$\int_X hd\mu = \int_X h^+d\mu - \int_X h^-d\mu,$$

所以

$$\int_X f - gd\mu = \int_X hd\mu = \int_X h^+d\mu - \int_X h^-d\mu = \int_X fd\mu - \int_X gd\mu.$$

3) f 和 g 都是实值函数，$a, b \in \mathbb{R}_{\geqslant 0}$ 的情况。此时，我们有

$$\begin{aligned}\int_X af + bgd\mu &= \int_X (af_+ + bg_+) - (af_- + bg_-)d\mu \\ &= \int_X (af_+ + bg_+)d\mu - \int_X (af_- + bg_-)d\mu \\ &= a\int_X f_+d\mu + b\int_X g_+d\mu - a\int_X f_-d\mu - b\int_X g_-d\mu.\end{aligned}$$

4) f 和 g 都是实值函数，$a, b \in \mathbb{R}$ 的情况。按照积分的定义，我们自然有

$$\int_X fd\mu = \int_X f_+d\mu - \int_X f_-d\mu = -\left(\int_X f_-d\mu - \int_X f_+d\mu\right) = -\int_X fd\mu.$$

此时，对与 a 和 b 可能的正负情况逐一讨论即可。

5) f 和 g 都是复值函数，$a, b \in \mathbb{C}$ 的情况。我们 f，g，a 和 b 都用它们的实部和虚部写出来，展开即可，这些繁琐且无启发性的细节留给同学们在课下验证。

至此，我们完整地证明了积分的线性。为了证明定理中的不等式，选取复数 $e^{i\theta_0}$，使得

$$e^{i\theta_0}\int_X fd\mu = |\int_X fd\mu|.$$

根据积分的线性，我们有

$$|\int_X f d\mu| = \int_X e^{i\theta_0} f d\mu.$$

所以，函数 $e^{i\theta_0} f$ 的虚数部分对积分没有贡献，从而

$$\int_X e^{i\theta_0} f d\mu = \int_X \mathrm{Re}\big(e^{i\theta_0} f\big) d\mu \leqslant \int_X \mathrm{Re}\big(e^{i\theta_0} f\big)_+ d\mu \leqslant \int_X |f| d\mu.$$

最后一步，我们用到了 $\mathrm{Re}\big(e^{i\theta_0} f\big)_+ \leqslant |f|$，这是显然的。 □

一旦有了积分的线性，我们就可以方便地证明很多命题了

推论 289. *（积分的区域可加性）给定测度空间 $(X, \mathcal{A}, \mu)$，f 是 X 上的可积函数。我们定义 f 的* **支集**$\mathrm{supp} f$ *为*

$$\mathrm{supp} f = \{x \in X \big| f(x) \neq 0\}.$$

假设 $\mathrm{supp} f \subset A \in \mathcal{A}$，我们也把 $\int_X f d\mu$ 写成 $\int_A f d\mu$。对于任意的 $A, B \in \mathcal{A}$，$A \cap B = \emptyset$，我们有 $f \cdot \mathbf{1}_A$ 和 $f \cdot \mathbf{1}_B$ 均可积并且

$$\int_{A \cup B} f = \int_A f d\mu + \int_B f d\mu.$$

证明： 由于 $|f \cdot \mathbf{1}_A| \leqslant |f|$，所以 $f \cdot \mathbf{1}_A$ 可积。推论中的等式就是对 $f \cdot \mathbf{1}_A + f \cdot \mathbf{1}_B$ 的积分应用线性。 □

推论 290. *在零测集上改变函数的值不会影响其积分，也就是说，如果 $f_1, f_2 \in \mathcal{L}^1(X, \mathcal{A}, \mu)$，使得 $f_1 = f_2$ 几乎处处，那么*

$$\int_X f_1 d\mu = \int_X f_2 d\mu.$$

证明： 按照要求，$f = f_1 - f_2$ 几乎处处是 0，所以其积分为 0，从而

$$0 = \int_X f d\mu = \int_X f_1 - f_2 d\mu = \int_X f_1 d\mu - \int_X f_2 d\mu.$$

□

推论 291 ($L^1(X, \mathcal{A}, \mu)$ 是赋范线性空间)**.** *映射*

$$\| \cdot \|_{L^1(X,\mathcal{A},\mu)} : L^1(X, \mathcal{A}, \mu) \to \mathbb{R}_{\geqslant 0}, \quad [f] \mapsto \int_X |f| d\mu,$$

是范数。我们这个范数称作是L^1 **范数**，*对于 $f \in L^1(X, \mathcal{A}, \mu)$，我们把它的范数简记为 $\|f\|_{L^1}$。从而，$\big(L^1(X, \mathcal{A}, \mu), \| \cdot \|_{L^1}\big)$ 是赋范线性空间。*

证明： 首先,对于等价类 $[f]$ 中的任何两个代表元 $f_1, f_2 \in [f]$,它们几乎处处相等,所以 $|f_1| = |f_2|$ 几乎处处，从而，

$$\|[f]\|_{L^1} = \int_X |f_1| d\mu = \int_X |f_2| d\mu.$$

这表明映射是良好定义的。另外，我们知道 $\left|\int_X |f|d\mu\right| = 0$ 等价于在 $L^1(X,\mathcal{A},\mu)$ 中 $f=0$。所以，$\|f\|_{L^1}=0$ 等价于在 $L^1(X,\mathcal{A},\mu)$ 中 $f=0$。对于任意的 $a,b\in\mathbb{C}$ 和 $a,b\in L^1(X,\mathcal{A},\mu)$，根据积分的线性，我们有

$$\begin{aligned}\|af+bg\|_{L^1} &= \int_X |af+bg|d\mu \leqslant \int_X |a||f|+|b||g|d\mu \\ &= |a|\|f\|_{L^1}+|b|\|g\|_{L^1}.\end{aligned}$$

这就说明了 $\|\cdot\|_{L^1}$ 是范数。 □

注记. 我们将证明 $\left(L^1(X,\mathcal{A},\mu),\|\cdot\|_{L^1}\right)$ 是完备的赋范线性空间。

注记. （子空间上的积分）假定 $Y\in\mathcal{A}$，我们定义 $\mathcal{A}\big|_Y=\{A\in\mathcal{A}|A\subset Y\}$，这 σ-代数：实际上，如果令

$$\iota: Y\to X,\ y\mapsto y,$$

那么，$\mathcal{A}\big|_Y=\iota^*\mathcal{A}$。我们可以将测度 μ 限制到 $\mathcal{A}|_Y$ 上得到一个测度 $\mu\big|_Y$：对任意的 $A\cap Y\subset\mathcal{A}|_Y$，我们定义

$$\mu\big|_Y(A\cap Y)=\mu(A\cap Y)$$

。这样，我们就得到了测度空间 $(Y,\mathcal{A}\big|_Y,\mu\big|_Y)$ 从而可以对 Y 上定义的函数进行积分。另外，对于 Y 上定义函数，还可以将它用零延拓成 X 上的函数从而将该函数视作是整个空间 X 上的函数，然后我们可以在 X 上积分。这两种做法是等价的，我们会在作业中证明。

我们还必须做出如下的澄清：当 $X=\mathbb{R}^n$，Y 是子流形时（余维数至少是 1），那么 $\mu\big|_Y$ 在任何集合上取值都是 0（因为此时子流形的测度是零）。多元微积分的课程要专门研究子流形上的积分理论，通常的微积分教材里把这些积分称作是曲线和曲面上的积分。

我们现在可以证明积分理论中最重要的收敛定理了，它的应用渗透到近代分析的每个角落：

定理 292 (Lebesgue 控制收敛定理)**.** 假定测度空间 $(X,\mathcal{A},\mu)$ 上的可测函数列 $\{f_i\}_{i\geqslant 1}$ 几乎处处收敛到函数 f，即存在零测集 $N\in\mathcal{A}$，使得对任意的 $x\in N^c$，我们有 $\lim\limits_{i\to\infty} f_i(x)=f(x)$。如果存在所谓的控制函数 $h\in\mathcal{L}^1(X,\mathcal{A},\mu)$，使得对每个 i，$|f_i(x)|\leqslant h(x)$ 几乎处处成立（即存在零测集 $N_i\in\mathcal{A}$，使得对任意的 $x\in(N_i)^c$，$|f_i(x)|\leqslant h(x)$），那么，我们有

$$\lim_{i\to\infty}\int_X |f_i-f|d\mu=0.$$

特别地，我们有

$$\lim_{i\to\infty}\int_X f_i d\mu=\int_X f d\mu.$$

证明: 我们要应用 Beppo Levi 定理。首先，定义正函数序列 $\{g_i\}_{i\geqslant 1}$：

$$g_i(x)=2h(x)-\sup_{k\geqslant i}|f(x)-f_k(x)|.$$

按照构造方式，这是单调上升的函数序列。根据定理的条件，$\{g_i\}_{i\geqslant 1}$ 几乎处处收敛到 $2h$，即对于任意的 $x\notin N$，我们有 $\lim\limits_{i\to\infty} g_i(x)=2h(x)$。

我们考虑集合 $N\cup\bigcup\limits_{i\geqslant 1}N_i$，这是一个零测集。在 $X-N\cup\bigcup\limits_{i\geqslant 1}N_i$ 上，我们有

$$g_i(x)\geqslant 0,\ \ g_i(x)\to 2h(x).$$

利用 Beppo Levi 定理，我们知道

$$\begin{aligned}\lim_{i\to\infty}\int_X\big(2h(x)-\sup_{k\geqslant i}|f(x)-f_k(x)|\big)d\mu&=\lim_{i\to\infty}\int_X g_i d\mu\\&=2\int_X h d\mu.\end{aligned}$$

在方程的两边同时减掉 $2\int_X h d\mu$，我们就得到

$$\lim_{i\to\infty}\int_X\sup_{k\geqslant i}|f(x)-f_k(x)|d\mu=0.$$

我们只要简单地去掉 sup 就证明了 Lebesgue 控制收敛定理。 □

注记. *在 Lebesgue 控制收敛定理的证明过程中，我们得到了更强的结论：*

$$\lim_{i\to\infty}\int_X\sup_{k\geqslant i}|f(x)-f_k(x)|d\mu=0.$$

下面的两个推论有着众多的应用，我们会在作业和考试中展现它们：

推论 293 (积分对参数的连续依赖性)**.** *假定参数空间 Ω 为距离空间（一般而言，Ω 是 $\mathbb{R}^n$ 中的一个开集），$(X,\mathcal{A},\mu)$ 是测度空间。函数*

$$f:X\times\Omega\to\mathbb{C},\ (x,t)\mapsto f(x,t),$$

满足如下条件：

1) 对每个固定的 $t\in\Omega$，函数 $x\mapsto f(x,t)$ 是可测的；

2) 对几乎处处的 x，映射 $t\mapsto f(x,t)$ 在 $t_0\in\Omega$ 处连续（即存在零测集 N，使得对任意的 $x\in N^c$，映射 $t\mapsto f(x,t)$ 在 $t_0\in\Omega$ 处连续）；

3) 存在正函数 $h\in\mathcal{L}^1(X,\mathcal{A},\mu)$，使得对每个 $t\in\Omega$，我们有

$$|f(x,t)|\leqslant h(x)$$

对几乎处处的 x 成立(即存在存在零测集 N_t，使得对任意的 $x\notin N_t$，我们有 $|f(x,t)|\leqslant h(x)$)。

那么，函数

$$F : \Omega \to \mathbb{C}, \quad t \mapsto F(t) = \int_X f(x,t) d\mu(x)$$

是良好定义的并且在 t_0 处连续。

证明: 由于 $|f(x,t)| \leqslant h(x)$，所以对于固定的 t，$f(x,t)$ 是可积的，从而函数 $F(t)$ 是良好定义的。我们来证明 F 在 t_0 处的连续性。为此，任取点列 $t_k \to t_0$，我们希望证明

$$\int_X f(x,t_k) d\mu(x) \to \int_X f(x,t_0) d\mu(x).$$

这就是 Lebesgue 控制收敛定理的内容，因为我们可以选取 h 作为控制函数。 □

43 积分与求导数的可交换性，乘积测度的构造

二零二零年四月二日，星期四，晴

我们上次课证明了 Lebesgue 控制收敛定理和它的一个推论（对参数的连续依赖性）：

定理 294 (Lebesgue 控制收敛定理)**.** 假定测度空间 $(X,\mathcal{A},\mu)$ 上的可测函数列 $\{f_i\}_{i\geqslant 1}$ 几乎处处收敛到函数 f。如果存在控制函数 $h\in\mathcal{L}^1(X,\mathcal{A},\mu)$，使得对每个 i，$|f_i(x)|\leqslant h(x)$ 几乎处处成立。那么，我们有

$$\lim_{i\to\infty}\int_X |f_i-f|d\mu=0.$$

特别地，我们有

$$\lim_{i\to\infty}\int_X f_i d\mu=\int_X f d\mu.$$

推论 295 (积分对参数的连续依赖性)**.** 假设 Ω 为距离空间，$(X,\mathcal{A},\mu)$ 是测度空间。函数

$$f:X\times\Omega\to\mathbb{C},\ (x,t)\mapsto f(x,t),$$

满足如下条件：

1) 对每个固定的 $t\in\Omega$，函数 $x\mapsto f(x,t)$ 是可测的；

2) 对几乎处处的 x，映射 $t\mapsto f(x,t)$ 在 $t_0\in\Omega$ 处连续；

3) 存在正函数 $h\in\mathcal{L}^1(X,\mathcal{A},\mu)$，使得对每个 $t\in\Omega$，我们有 $|f(x,t)|\leqslant h(x)$ 对几乎处处的 x 成立。

那么，函数

$$F:\Omega\to\mathbb{C},\quad t\mapsto F(t)=\int_X f(x,t)d\mu(x)$$

是良好定义的并且在 t_0 处连续。

我们这次证明另一个推论：

推论 296 (积分与求导数的可交换性)**.** 假定参数空间 I 为 $\mathbb{R}$ 中的开区间，$(X,\mathcal{A},\mu)$ 是测度空间，$N\in\mathcal{A}$ 为零测集合，$X'=X-N$。我们假设函数

$$f:X\times I\to\mathbb{C},\quad (x,t)\mapsto f(x,t)$$

满足如下条件：

1) 对任意固定的 $t\in I$，函数 $x\mapsto f(x,t)$ 是可积的；

2) 对任意固定的 $x\in X'$，函数 $t\mapsto f(x,t)$ 对 t 的导数 $\dfrac{\partial f}{\partial t}(x,t)$ 存在并且存在正函数 $h\in\mathcal{L}^1(X,\mathcal{A},\mu)$，使得对任意的 $(x,t)\in X'\times I$，都有

$$|\frac{\partial f}{\partial t}(x,t)|\leqslant h(x).$$

那么，函数

$$F : I \to \mathbb{C},\ t \mapsto F(t) = \int_X f(x,t)d\mu(x)$$

是良好定义的并且在 I 上可微。进一步，我们有如下的导数公式

$$F'(t) = \int_X \frac{\partial f}{\partial t}(t,x)d\mu(x).$$

注记. 我们用 $d\mu(x)$ 表明是对 x 来积分。

证明: 任意选取实数数列 $\{\varepsilon_k\}_{k\geqslant 1}$，使得 $\lim\limits_{k\to\infty}\varepsilon_k = 0$。对于任意的 $x \in X'$，按定义，我们知道 $\dfrac{\partial f}{\partial t}(t,x)$ 是可测函数序列 $\left\{\dfrac{f(x,t+\varepsilon_k)-f(x,t)}{\varepsilon_k}\right\}_{k\geqslant 1}$ 在 x 点的极限，从而 $\dfrac{\partial f}{\partial t}(t,x)$ 是可测函数。

由于 $\left|\dfrac{\partial f}{\partial t}(x,t)\right| \leqslant h(x)$，所以 $\dfrac{\partial f}{\partial t}(t,x)$ 是可积的，从而函数

$$t \mapsto \int_X \frac{\partial f}{\partial t}(t,x)d\mu(x).$$

是良好定义的。

现在证明 F 的导数存在并计算它的导数。固定 $t \in I$，根据积分的线性，我们有

$$\begin{aligned}\frac{F(t+\varepsilon_k)-F(t)}{\varepsilon_k} &= \int_X \frac{f(x,t+\varepsilon_k)-f(x,t)}{\varepsilon_k}d\mu(x)\\ &= \int_{X'} \frac{f(x,t+\varepsilon_k)-f(x,t)}{\varepsilon_k}d\mu(x).\end{aligned}$$

根据中值定理，我们知道最后一式的被积函数可写为 $f'(x,t+\theta\varepsilon_k)$，其中 $\theta \in [0,1]$。据此，我们有

$$\left|\frac{f(x,t+\varepsilon_k)-f(x,t)}{\varepsilon_k}\right| = |f'(x,t+\theta\varepsilon_k)| \leqslant h(x).$$

我们现在可以对函数序列 $\left\{\dfrac{f(x,t+\varepsilon_k)-f(x,t)}{\varepsilon_k}\right\}_{k\geqslant 1}$ 应用 Lebesgue 控制收敛定理，因为 h 就是控制函数:

$$\begin{aligned}\lim_{k\to\infty}\frac{F(t+\varepsilon_k)-F(t)}{\varepsilon_k} &= \int_X \frac{f(x,t+\varepsilon_k)-f(x,t)}{\varepsilon_k}d\mu(x)\\ &= \int_X \frac{\partial f(x,t)}{\partial t}d\mu(x).\end{aligned}$$

这就给出了证明。 □

乘积测度与 Fubini 公式

所谓的 Fubini 定理，就是将乘积空间上的积分化为分量空间上的积分的公式，它将保证我们能把 $\mathbb{R}^n$ 上的积分化为 $\mathbb{R}^1$ 上的积分来计算。为此，我们的第一件事情是建立乘积测度。

给定两个测度空间 $(X,\mathcal{A},\mu)$ 和 $(Y,\mathcal{B},\nu)$，我们要求 μ 和 ν 是 σ-有限的。我们要在可测空间 $(X\times Y,\mathcal{A}\otimes\mathcal{B})$ 上定义一个测度 $\mu\otimes\nu$。

根据定义，$X\times Y$ 上的 σ-代数 $\mathcal{A}\otimes\mathcal{B}$ 是按照如下的方式构造的：我们称 $X\times Y$ 上形如 $A\times B$ 的子集为矩形，其中 $A\in\mathcal{A},B\in\mathcal{B}$；我们用 $\widetilde{\mathcal{R}}$ 表示有限个两两不交的矩形之并所构成集合的全体并且证明了 $\widetilde{\mathcal{R}}$ 是代数；$\mathcal{A}\otimes\mathcal{B}$ 被定义为 $\widetilde{\mathcal{R}}$ 所生成的 σ-代数，即 $\mathcal{A}\otimes\mathcal{B}=\sigma(\widetilde{\mathcal{R}})$；$\mathcal{A}\otimes\mathcal{B}$ 也可以由所有的矩形生成。

与 Lebesgue 测度的构造类似，我们要在 $\widetilde{\mathcal{R}}$ 上构造一个加性函数并利用 Carathéodory 定理把它扩张成 $\mathcal{A}\otimes\mathcal{B}$ 上的测度。为此，在矩形 $A\times B\in\widetilde{\mathcal{R}}$，我们定义

$$\theta(A\times B)=\mu(A)\nu(B).$$

我们规定 0 乘任何数都得 0，∞ 乘任何非零正数都得 ∞。

对于 $\widetilde{\mathcal{R}}$ 中的元素 S，按照 $\widetilde{\mathcal{R}}$ 的定义，我们可以将 S 写成有限个不交的矩形的并：

$$S=\coprod_{i\leqslant m}A_i\times B_i,\quad A_i\in\mathcal{A},\ B_i\in\mathcal{B}.$$

此时，我们定义

$$\theta(S)=\sum_{i\leqslant m}\theta(A_i\times B_i)=\sum_{i\leqslant m}\mu(A_i)\nu(B_i).$$

然而，把 S 写成有限个不相交的矩形的并的方式不是唯一的，我们还可能有别的方式：

$$S=\coprod_{j\leqslant n}A'_j\times B'_j\in\widetilde{\mathcal{R}}.$$

这样，我们得到 S 的两种不同的分解，为了说明 $\theta(S)$ 是良好定义的（不依赖于分解的选取），我们必须证明

$$\sum_{i\leqslant m}\mu(A_i)\nu(B_i)=\sum_{j\leqslant n}\mu\big(A'_j\big)\nu\big(B'_j\big).$$

我们要考虑这两种分解共同的“加细”：对每个 $1\leqslant i\leqslant m$，我们有

$$\begin{aligned}A_i\times B_i&=\coprod_{j\leqslant n}\big(A_i\times B_i\big)\cap\big(A'_j\times B'_j\big)\\&=\coprod_{j\leqslant n}\big(A_i\cap A'_j\big)\times\big(B_i\cap B'_j\big).\end{aligned}$$

所以，上述两种分解可以加细为如下的关于 S 的新的分解：

$$S=\coprod_{\substack{i\leqslant m,\\ j\leqslant n}}\big(A_i\cap A'_j\big)\times\big(B_i\cap B'_j\big).$$

利用这个分解，我们有

$$\begin{aligned}\sum_{i\leqslant m,j\leqslant n}\mu\big(A_i\cap A'_j\big)\nu\big(B_i\cap B'_j\big)&=\sum_{i\leqslant m}\Big(\sum_{j\leqslant n}\mu\big(A_i\cap A'_j\big)\nu\big(B_i\cap B'_j\big)\Big)\\&=\sum_{i\leqslant m}\mu\big(A_i\big)\nu\big(B_i\big).\end{aligned}$$

上面最后一个红色等号我们用到了等式

$$A_i \times B_i = \coprod_{j\leqslant n} \big(A_i \cap A'_j\big) \times \big(B_i \cap B'_j\big).$$

我们暂且假设这个等号成立。类似地，如果我们先对 $i \leqslant m$ 求和，我们就得到了

$$\sum_{\substack{i\leqslant m,\\ j\leqslant n}} \mu\big(A_i \cap A'_j\big)\nu\big(B_i \cap B'_j\big) = \sum_{j\leqslant n} \mu\big(A'_j\big)\nu\big(B'_j\big).$$

这就证明了两种分解所对应的 θ 的值都是一样的！

当然，我们必须证明如下的等式：如果（所有的元素都来自 $\mathcal{A}$ 或者 $\mathcal{B}$），

$$A \times B = \coprod_{j\leqslant n} \big(A \cap A'_j\big) \times \big(B \cap B'_j\big),$$

那么

$$\sum_{j\leqslant n} \mu\big(A \cap A'_j\big)\nu\big(B \cap B'_j\big) = \mu(A)\nu(B).$$

为此，我们不妨假设 A_j 均为 A 的子集并且 $\bigcup\limits_{j\leqslant n} A_j = A$，$B_j$ 均为 B 的子集并且 $\bigcup\limits_{j\leqslant n} B_j = B$，而且

$$A \times B = \coprod_{j\leqslant n} A_j \times B_j.$$

我们要证明

$$\sum_{j\leqslant n} \mu\big(A_j\big)\nu\big(B_j\big) = \mu(A)\nu(B).$$

如果 $\{A_j\}_{j\leqslant n}$ 是两两不交的，$\{B_j\}_{j\leqslant n}$ 也是两两不交的，那么命题上述公式是显然的，因为

$$\begin{aligned}\mu(A)\nu(B) &= \left(\sum_{j\leqslant n} \mu\big(A_j\big)\right)\left(\sum_{j\leqslant n} \mu\big(B_j\big)\right)\\ &= \sum_{j,j'\leqslant m} \mu\big(A_j\big)\mu\big(B_{j'}\big).\end{aligned}$$

对于一般的情形，我们要 A_j 和 B_j 进一步分解成不相交的集合即可。我们考虑

$$\big\{X_1 \cap X_2 \cap \cdots \cap X_n \big| X_i = A_i \text{ 或 } (A_i)^c\big\},$$

它们是两两不交的并且可以把 A_j 都并出来。

据此，我们有

命题 297. 映射

$$\theta : \widetilde{\mathcal{R}} \to [0,\infty], \quad S \mapsto \sum_{i\leqslant m} \mu(A_i)\nu(B_i).$$

是代数 $\widetilde{\mathcal{R}}$ 上的 σ-有限的加性函数。

证明: 根据定义，这显然是加性函数。为了说明 σ-有限性，由于 X 和 Y 是 σ-有限的，所以存在上升的序列 $\{A_i\}_{i\geqslant 1}\subset\mathcal{A}$，$X_i\nearrow X$，$\{B_i\}_{i\geqslant 1}\subset\mathcal{B}$，$B_j\nearrow Y$，其中对每个 i，$\mu(A_i)$ 和 $\nu(B_j)$ 是有限的。所以，我们就有上升的序列 $\{A_i\times B_i\}_{i\geqslant 1}\subset\widetilde{\mathcal{R}}$，使得 $A_i\times B_i\nearrow X\times Y$。按照 θ 的定义，对每个 i，$\theta(A_i\times B_i)=\mu(A_i)\nu(B_i)$ 是有限的。这就完成了证明。

我们注意到，这里的构造和 Lebesgue 测度的构造在形式上一摸一样。 □

我们将使用 Carathéodory 测度扩张定理来构造 $\mathcal{A}\otimes\mathcal{B}$ 上的乘积测度:

$$\begin{array}{ccc}\widetilde{\mathcal{R}} & \xrightarrow{\theta} & [0,\infty] \\ \downarrow{\scriptstyle \iota} & \nearrow_{\mu\otimes\nu} & \\ \mathcal{A}\otimes\mathcal{B} & & \end{array}$$

为此，先证明一个粗糙版本的 Fubini 定理:

引理 298. *对任意 $S\in\widetilde{\mathcal{R}}$，$x\in X$，我们定义 S 的 x-截面为:*

$$S_x=\{y\in Y|(x,y)\in S\}\subset Y.$$

那么，$S_x\in\mathcal{B}$。进一步，函数

$$X\to[0,\infty],\quad x\mapsto\nu(S_x),$$

为 $(X,\mathcal{A})$ 上的可测函数并且

$$\theta(S)=\int_X\nu(S_x)d\mu(x).$$

证明: 我们考虑 S 的分解 $S=\coprod_{i\leqslant m}A_i\times B_i$，其中矩形们 $A_i\times B_i$ 两两不交，通过对 A_i 进一步进行细分（如同上一个命题中所证明的），我们可以进一步要求 A_i 们两两不交。此时，我们可以精确的计算 S 的 x-截面:

$$S_x=\begin{cases}B_i, & \text{如果 } x\in A_i;\\ \emptyset, & \text{如果 } x\notin\bigcup_{i\leqslant m}A_i.\end{cases}$$

这显然是 $\mathcal{B}$ 中的元素。我们现在研究 $\nu(S_x)$ 的可测性，实际上

$$\nu(S_x)=\begin{cases}\nu(B_i), & \text{如果 } x\in A_i;\\ 0, & \text{如果 } x\notin\bigcup_{i\leqslant m}A_i.\end{cases}$$

这是一个简单函数，自然可测。所以，我们有

$$\begin{aligned}\int_X\nu(S_x)d\mu(x)&=\sum_{i\leqslant m}\int_X\nu(B_i)\mathbf{1}_{A_i}(x)d\mu(x)\\&=\sum_{i\leqslant m}\mu(A_i)\nu(B_i).\end{aligned}$$

根据定义，上式右边即为 $\theta(S)$，命题得证。 □

利用上述引理，我们来验证 θ 满足 Carathéodory 扩张定理中的**条件** $(\mathbf{C})$ 和**条件** $(\mathbf{C}_\infty)$：

条件 $(\mathbf{C}_\infty)$ 的验证：

根据 σ-有限性，存在上升的序列 $\{A_i\}_{i\geqslant 1}\subset\mathcal{A}$，$A_i\nearrow X$，$\{B_i\}_{i\geqslant 1}\subset\mathcal{B}$，$B_j\nearrow Y$，其中对每个 i，$\mu(A_i)$ 和 $\nu(B_j)$ 是有限的。令 $Z=X\times Y$ 为整个乘积空间，考虑上升的集合的序列 $\{Z_i=A_i\times B_i\}_{i\geqslant 1}$，很明显，我们有 $Z_i\nearrow Z$ 并且每个 Z_i 的 θ 体积有限。**条件** $(\mathbf{C}_\infty)$ 要求我们证明：对于任意的 $S\in\widetilde{R}$，如果 $\theta(S)=\infty$，那么

$$\lim_{i\to\infty}\theta(S\cap Z_i)=\infty.$$

由于 $S\in\widetilde{R}$ 为有限个矩形的并，所以只要对某个 $S=A\times B$ 的情形证明 $\theta(S\cap X_i)\to\infty$ 即可（这里和 Lebesgue 测度的构造又是一样的）。由于我们假设了 $\theta(A\times B)=\infty$，不妨 $\mu(A)=\infty$ 而 $\nu(B)>0$（可以是 ∞）。根据 $Z_i=A_i\times B_i$，所以

$$\theta\big(Z_i\cap(A\times B)\big)=\mu(A_i\cap A)\nu(B_i\cap B).$$

利用测度 μ 和 ν 的性质，我们有

$$\mu(A_i\cap A)\to\infty,\ \ \nu(B_i\cap B)\geqslant\delta>0,\text{（当 } i \text{ 足够大时候）}.$$

从而，我们有

$$\lim_{i\to\infty}\theta\big(Z_i\cap(A\times B)\big)\to\infty.$$

条件 $(\mathbf{C})$ 的验证：

任意选取 $\widetilde{R}$ 中下降的序列 $\{S_i\}_{i\geqslant 1}$，如果 $\theta(S_0)<\infty$ 并且 $S_i\searrow\emptyset$，那么，**条件** $(\mathbf{C})$ 要求我们证明

$$\lim_{i\to\infty}\theta(S_i)=0.$$

我们运用粗糙版本的 Fubini 定理。首先，研究 X 上的可测函数列 $\{f_i(x)\}_{i\geqslant 1}$，其中，

$$f_i(x):X\to[0,\infty],\ \ x\mapsto\nu\left((S_i)_x\right).$$

由于 S_i 是下降集合序列，所以，$\{f_i(x)\}_{i\geqslant 1}$ 是 X 上下降的正可测函数序列。根据定义，我们有按照定义，我们有

$$f_i(x)=\nu\left((S_i)_x\right).$$

当 $x\in X$ 固定时，由于 $S_i\searrow\emptyset$，所以 $(S_i)_x\searrow\emptyset$，从而，$\nu\big((S_i)_x\big)\searrow 0$。那么，从而对任意的 x，$f_i(x)\to 0$，即函数列 $\{f_i\}_{i\geqslant 1}$ 逐点收敛到 0。我们对函数列 $\{f_i(x)\}_{i\geqslant 1}$ 运用 Lebesgue 控制收敛定理，其中，这列函数中的第一项 f_0 可以作为控制函数。根据上面粗糙版本的 Fubini 定理，我们就有

$$\theta(S_i)=\int_X f_i(x)d\mu(x)\to 0,$$

这就证明了**条件** $(\mathbf{C})$。

综上所述，我们证明了如下重要的定理：

定理 299 (乘积测度)**.** 给定 σ-有限的测度空间 $(X, \mathcal{A}, \mu)$ 和 $(Y, \mathcal{B}, \nu)$，在可测空间 $(X \times Y, \mathcal{A} \otimes \mathcal{B})$ 上存在唯一的 σ-有限的测度 $\mu \otimes \nu$，使得对任意的 $A \in \mathcal{A}$ 和 $B \in \mathcal{B}$，我们都有

$$(\mu \otimes \nu)\big(A \times B\big) = \mu(A)\nu(B).$$

特别地，对任意的 $E \in \mathcal{A} \otimes \mathcal{B}$，它的测度可由如下公式计算

$$(\mu \otimes \nu)\big(E\big) = \inf_{\substack{E \subset \bigcup_{i=1}^{\infty}(A_i \times B_i) \\ A_i \in \mathcal{A}, B_i \in \mathcal{B}}} \sum_{i=1}^{\infty} \mu(A_i)\nu(B_i).$$

注记. 在之前的课程中，我们已经证明了 $\mathbb{R}^2 = \mathbb{R}^1 \times \mathbb{R}^1$ 上 *Borel* 代数是 $\mathbb{R}^1$ 上的 *Borel* 代数的乘积，即 $\mathcal{B}(\mathbb{R}^2) = \mathcal{B}(\mathbb{R}^1) \otimes \mathcal{B}(\mathbb{R}^1)$。由于我们上面关于乘积测度的构造过程和 *Lebesgue* 测度的构造是一样的，所以 $\mathbb{R}^2$ 上的 *Lebesgue* 测度 m_2 就是两个 $\mathbb{R}^1$ 上的 *Lebesgue* 测度 m_1 的张量积。这还可以通过观察它们在方块上的取值以及测度的唯一性来证明。所以，我们有

$$\big(\mathbb{R}^2, \mathcal{B}(\mathbb{R}^2), m_2\big) = \big(\mathbb{R}^2, \mathcal{B}(\mathbb{R}^1) \otimes \mathcal{B}(\mathbb{R}^1), m_1 \otimes m_1\big).$$

类似的结论对于 $\mathbb{R}^n$ 也成立，我们不再赘述。

43.1 作业：Lebesgue 控制收敛，十进制小数的研究

清华大学 19-20 春季学期，数学分析二，作业 6

请用 A4 大小的纸张**正反面**用钢笔，签字笔或者圆珠笔书写，并注明自己的姓名，年级和作业的**总页数**。除定理公式所涉及的人名之外，请使用**中文**。本次作业的提交时间和地点为 **2 月 64 日**上午的课堂上，逾期视作零分。

课堂内容的补充与复习

A1)（积分与级数之间的关联）考虑正整数集上的测度空间 $(\mathbb{Z}_{\geqslant 1}, \mathcal{P}(\mathbb{Z}_{\geqslant 1}), \mu)$，其中 μ 是数元素个数的测度：

$$\mu(A) = |A|.$$

考虑这个空间上复数值函数 $f(x)$。证明，f 可积分当且仅当

$$\sum_{n=1}^{\infty} |f(n)| < \infty$$

并且此时

$$\int_{\mathbb{Z}_{\geqslant 1}} f d\mu = \sum_{n=1}^{\infty} f(n).$$

A2)（Beppo Levi 定理的标准应用）考虑测度空间 $(X, \mathcal{A}, \mu)$ 上的正函数列 $\{g_i\}_{i\geqslant 1}$。证明，函数级数 $\sum\limits_{i=1}^{\infty} g_i$ 是 X 上良好定义的可测正函数并且

$$\int_X \sum_{i=1}^{\infty} g_i d\mu = \sum_{i=1}^{\infty} \int_X g_i d\mu.$$

A3) 考虑 $\mathbb{R}$ 上的函数 $f(x) = \dfrac{\sin(x)}{x}$。证明，在 Lebesgue 的意义下，$f$ 不可积。

A4) f 是测度空间 $(X, \mathcal{A}, \mu)$ 上的简单函数。证明，f 的积分与 f 作为简单函数所定义的积分是一致的。（提示：利用已经证明的关于积分的线性）

A5) $(X, \mathcal{A}, \mu)$ 是测度空间，$Y \in \mathcal{A}$，我们定义 $\mathcal{A}\big|_Y = \{A \in \mathcal{A} | A \subset Y\}$。对任意的 $A \cap Y \subset \mathcal{A}|_Y$，我们定义

$$\mu\big|_Y(A \cap Y) = \mu(A \cap Y).$$

证明，$\mu\big|_Y$ 是 $\mathcal{A}|_Y$ 上的测度。对于任意给定的可测函数 $f : Y \to \mathbb{C}$，令

$$\widehat{f} : X \to \mathbb{C}, \quad x \mapsto \widehat{f}(x) = \begin{cases} f(x), & x \in Y; \\ 0, & x \notin Y. \end{cases}$$

证明，f 在 $(Y,\mathcal{A}\big|_Y,\mu\big|_Y)$ 上可积当且仅当 $\widehat{f}$ 在 $(X,\mathcal{A},\mu)$ 上可积，并且此时我们有

$$\int_X \widehat{f}d\mu = \int_Y f d\mu\big|_Y.$$

A6)（Fatou 引理中的不等式的等号未必能取到）试构造某个测度空间 $(X,\mathcal{A},\mu)$ 上的可测正函数序列 $\{f_i\}_{i\geqslant 1}$，使得该序列逐点地收敛到 f 但是

$$\int_X f d\mu < \liminf_{i\to\infty}\int_X f_i d\mu.$$

A7) $(X,\mathcal{A},\mu)$ 是测度空间。证明，对任意可测函数 $f:X\to\mathbb{C}$，都存在阶梯函数序列 $\{\varphi_i\}_{i\geqslant 1}$，使得对每个 $x\in X$，都有

$$\lim_{k\to\infty}\varphi_k(x) = f(x).$$

A8) 试构造例子，说明如果 f 和 g 是测度空间 $(X,\mathcal{A},\mu)$ 上的可积函数，它们的乘积 $f\cdot g$ 未必是。

Lebesgue 控制收敛定理（必须掌握的内容）

如果不加说明，那么 $\mathbb{R}^n$ 上的测度 dx 总假定是 Lebesgue 测度。

问题 B：基本的 Lebesgeu 控制收敛习题

B1) f 是 $\mathbb{R}^n$ 上的可积函数，$\{X_k\}_{k\geqslant 1}$ 是单调上升的 Borel 集的序列并且 $\lim\limits_{k\to\infty}X_k=\mathbb{R}^n$。试证明

$$\lim_{k\to\infty}\int_{X_k} f(x)dx = \int_{\mathbb{R}^n} f(x)dx.$$

B2) 对于自然数 n，令

$$\Gamma_n = \int_0^n (1-\frac{x}{n})^n e^{\frac{x}{2}}dx,$$

试计算 $\lim\limits_{n\to\infty}\Gamma_n$。

B3) 利用积分 $\int_0^1 \frac{dx}{1+x}$ 来计算 $\sum\limits_{n=1}^{\infty}\frac{(-1)^{n+1}}{n}$。

B4) f 是 $\mathbb{R}^n$ 上的可积函数，$\{X_k\}_{k\geqslant 1}$ 是 Borel-集的序列并且

$$\lim_{k\to\infty} m(X_k) = 0.$$

试证明：

$$\lim_{k\to\infty}\int_{X_k} f(x)dx = 0.$$

B5) f 是 $\mathbb{R}$ 上的可积函数。我们令

$$\widetilde{f}:[0,1]\mapsto\mathbb{C},\quad x\mapsto\sum_{k\in\mathbb{Z}}f(x+k).$$

证明，f 是良定义的可测函数并且 f 在 $[0,1]$ 上可积，并进一步证明

$$\int_{[0,1]}\widetilde{f}(x)dx=\int_{\mathbb{R}}f(x)dx.$$

B6) 对于每个 $x\geqslant 0$，我们定义 $F(x)=\displaystyle\int_0^\infty\frac{1-\cos(u)}{u^2}e^{-xu}du$。

B6-1) 证明，对于 $x\in[0,\infty)$，$F(x)$ 是良定义的并计算 $\lim\limits_{x\to\infty}F(x)$。

B6-2) 证明，$F\in C^2\big((0,\infty)\big)$ 并计算其两阶导数 F''。

B6-3) 给出 F 的解析表达式。

B6-4) 计算积分 $\displaystyle\int_0^\infty\frac{1-\cos(u)}{u^2}du.$

B6-5) $F(x)$ 在 0 处的右导数是否存在?

问题 C：卷积初步

f 是 $\mathbb{R}$ 上的可积函数。

C1) 如果 g 是 $\mathbb{R}$ 上的有界的连续函数，证明

$$f*g(x)=\int_{\mathbb{R}}f(x-y)g(y)dy=\int_{\mathbb{R}}f(y)g(x-y)dy$$

定义了 $\mathbb{R}$ 上的一个有界的连续函数。

C2) 如果 g 是 $\mathbb{R}$ 上连续可微的函数，并且 g 和 g' 均有界，证明 $f*g$ 是连续可微的，并且

$$(f*g)'=f*g'.$$

问题 D：Fourier 变换初步

D1)（有限测度的 Fourier 变换）我们在 $\mathbb{R}$ 上取定 Borel-代数（由开集生成的 σ-代数）。

D1-1) 如果 μ 是一个有限测度（即 $\mu(\mathbb{R})<\infty$），那么

$$\hat{\mu}(\xi)=\int_{\mathbb{R}}e^{-ix\xi}d\mu(x)$$

定义了 $\mathbb{R}$ 上的一个连续有界的函数（称作是该有限测度的 Fourier 变换）。

D1-2) 我们定义 Dirac 测度 δ_0 如下：对于 Borel-集 A，如果 $0 \in A$，则 $\delta_0(A) = 1$；否则 $\delta_0(A) = 0$。证明，δ_0 是一个有限测度并计算它的 Fourier 变换。

D2)（可积函数的 Fourier 变换） f 是 $\mathbb{R}$ 上的可积函数。

D2-1) 对于 $\xi \in \mathbb{R}$，我们定义

$$\hat{f}(\xi) = \int_{\mathbb{R}} e^{-ix\xi} f(x) dx.$$

证明，上式定义了 $\mathbb{R}$ 上的一个连续有界的函数（称作是 f 的 Fourier 变换）。

D2-2) 如果 $g(x) = xf(x)$ 也是 $\mathbb{R}$ 上的可积函数，证明 $\hat{f}(\xi)$ 连续可微并且

$$\frac{d}{d\xi}\hat{f}(\xi) = -i\hat{g}(\xi).$$

D2-3)（Riemann-Lebesgue 引理）证明，对任意的 $f \in \mathcal{L}^1(\mathbb{R}, dx)$，我们都有

$$\lim_{\xi \to \infty} \widehat{f}(\xi) = 0.$$

（提示：先对特殊的阶梯函数来证明）

选做题：测度理论

十进制小数的研究

对 $x \in [0,1)$，我们将它按照十进制小数展开写成 $x = \sum_{n=1}^{\infty} \frac{a_n(x)}{10^n}$，其中 $a_n(x) \in \{0, 1, \cdots, 9\}$。

E1) 证明，对 $x \in [0,1)$，下面的四个表述等价：

- x 允许至少两种不同的上述小数展开；
- x 有一种小数展开使得存在 n_0，当 $n \geqslant n_0$ 时，$a_n = 0$；
- x 有一种小数展开使得存在 n_0，当 $n \geqslant n_0$ 时，$a_n = 9$；
- x 可写为 $\frac{m}{10^k}$ 的形式，其中 m, k 为整数。

在此之后，我们用第二种表示，从而上述十进制小数展开是唯一的。

E2) 对 $\alpha_1, \alpha_2, \cdots, \alpha_n \in \{0, 1, \cdots, 9\}$，我们定义

$$I(\alpha_1, \cdots, \alpha_n) = \{x \in [0,1) | a_i(x) = \alpha_i, i = 1, 2, \cdots, n\}.$$

证明，当 n 固定的时候，不同的 $(\alpha_1, \cdots, \alpha_n)$ 所定义的 $I(\alpha_1, \cdots, \alpha_n)$ 两两不交并且恰好覆盖 $[0,1)$。进一步说明 $I(\alpha_1, \cdots, \alpha_n)$ 是形如 $[a,b)$ 的区间并确定 a, b 及该区间的长度。

E3) 给定子集 $A\subset[0,1)$ 和 $n\geqslant 1$，我们定义

$$N_n(A)=\#\big\{I(\alpha_1,\cdots,\alpha_n)\big|I(\alpha_1,\cdots,\alpha_n)\cap A\neq\emptyset\big\}.$$

证明，极限 $\lim\limits_{n\to\infty}\dfrac{N_n(A)}{10^n}$ 存在。

E4) 假定 A 是 Borel 集合，$m(A)$ 为其 Lebesgue 测度。证明，$m(A)\leqslant\lim\limits_{n\to\infty}\dfrac{N_n(A)}{10^n}$。

E5) 举例子使得上述不等式中等号成立或者不成立。

E6) 我们定义

$$A_{1^\times}=\{x\in[0,1)|\text{对每个 } n,a_n(x)\neq 1\}.$$

证明，$A_{1^\times}$ 是 Borel 集并计算其的 Lebesgue 测度。

E7) 证明，$N_{n+1}(A_{1^\times})\leqslant 9N_n(A_{1^\times})$ 并利用这个不等式来检验上一步测度的计算。

E8) 对于 $\alpha_1,\cdots,\alpha_n\in\{0,1,2,\cdots\}$，定义

$$A_{(\alpha_1,\cdots,\alpha_n)^\times}=\{x\in[0,1)|(\alpha_1,\cdots,\alpha_n)\text{ 不作为 }x\text{ 小数展开的连续的 }n\text{ 项出现}\}.$$

证明，$A_{(\alpha_1,\cdots,\alpha_n)^\times}$ 是 Borel 集。

E9) 证明，$N_{m+n}(A_{(\alpha_1,\cdots,\alpha_n)^\times})\leqslant(10^n-1)N_m(A_{(\alpha_1,\cdots,\alpha_n)^\times})$ 并计算 $A_{(\alpha_1,\cdots,\alpha_n)^\times}$ 的测度。

E10) 证明，对几乎处处的 $x\in[0,1)$，任何由 $0,1,\cdots,9$ 构成的有限长度的字符串都会在 x 的十进制小数展开中出现无限多次。

一个 Borel 集的研究

我们把有理数写成 $\dfrac{p}{q}$ 的形式并要求 $q>0$ 并且 p 和 q 互素。给定 $n\in\mathbb{Z}_{\geqslant 0}$ 和 $\varepsilon>0$，定义

$$A_{n,\varepsilon}=\bigcup_{\frac{p}{q}\in\mathbb{Q}}\big(\frac{p}{q}-\frac{\varepsilon}{q^n},\frac{p}{q}+\frac{\varepsilon}{q^n}\big),\quad \widetilde{A_{n,\varepsilon}}=A_{n,\varepsilon}\cap[0,1],\quad A_n=\bigcap_{\varepsilon>0}A_{n,\varepsilon},\quad \widetilde{A_n}=\bigcap_{\varepsilon>0}\widetilde{A_{n,\varepsilon}}.$$

F1) A_n 是 $\mathbb{R}$ 上稠密的 Borel 集。

F2) 证明, 如果 $n\geqslant 3$，那么 A_n 是零测集（Lebesgue 测度为零）。

F3) 证明，$\bigcap\limits_n A_n\neq\mathbb{Q}$。

F4) 证明，$\bigcap\limits_n A_n$ 是不可数集

F5) 证明，$\bigcap_n A_n - \mathbb{Q}$ 在 $\mathbb{R}$ 中稠密。

F6) 证明，当 $n = 2$ 并且 $\varepsilon = 1$ 时，$A_{2,1} = \mathbb{R}$。

F7) 证明，$A_1 = \mathbb{R}$。

When I was about thirteen, the library was going to get *Calculus for the Practical Man*. By this time I knew, from reading the encyclopedia, that calculus was an important and interesting subject, and I ought to learn it.

—— Richard P. Feynman

43.2 习题课：硬币空间的测度理论

清华大学 19-20 春季学期，数学分析二，第四／五次习题课内容

测度理论：一个例子的研究

Happy Hunger Games! And may the odds be ever in your favor.

硬币空间的测度理论

令 $\Omega = \{0,1\}^{\mathbb{Z}_{>0}}$，其元素 ω 可用无限长的数列 $\omega = (\omega_1, \omega_2, \cdots)$ 来表示，其中 ω_i 为 0 或者 1。

A1) 按二进制方式定义映射

$$\Phi : \Omega \to [0,1], \quad \omega \mapsto \sum_{k=1}^{\infty} \frac{\omega_k}{2^k}.$$

证明：这个映射是满射并且除了形如 $\frac{n}{2^k} \in (0,1)$ 的数之外，每个 $[0,1]$ 中的数的原像只有一个，并确定所有数的原像。

A2) 对于每个 $s = (s_1, \cdots, s_n) \in \{0,1\}^n$，我们定义

$$C_s = \{\omega \in \Omega | \omega_1 = s_1, \cdots, \omega_n = s_n\}.$$

证明，当 s 遍历 $\{0,1\}^n$ 时，C_s 构成了集合 Ω 的一个拆分（也就是说 C_s 们两两不相交而且它们的并恰好是 Ω）。再令 $\mathcal{F}_n$ 为由 $\{C_s | s \in \{0,1\}^n\}$ 所生成的代数（称作是*前 n 次扔硬币的代数*）。证明，我们可以建立一个从 $\mathcal{F}_n$ 到 $\mathcal{P}(\{0,1\}^n)$（幂集合）的一一对应。

A3) 证明，$\mathcal{F}_1 \subset \mathcal{F}_2 \subset \cdots$ 是上升的序列并且 $\mathcal{F}_\infty = \bigcup_{n \geqslant 1} \mathcal{F}_n$ 也是代数但是 $\mathcal{F}_\infty \neq \Omega$。

A4) 证明，对每个 $A \in \mathcal{F}_\infty$，总存在 n 和 $s_1, \cdots, s_k \in \{0,1\}^n$，使得 $A = C_{s_1} \cup \cdots \cup C_{s_k}$。定义 $P(A) = \dfrac{k}{2^n}$，证明，这个定义不依赖于 n 的选取。

A5) 证明，P 是 $\mathcal{F}_\infty$ 上的一个加性函数且全空间测度为 1（概率测度）。

A6) （非等概率分布，用来模拟第 k 次扔硬币正面朝上的概率是 p_k）对于每个 $k \in \mathbb{Z}_{>0}$，我们给一个数 $0 \leqslant p_k \leqslant 1$。证明，存在唯一一个 $\mathcal{F}_\infty$ 上的加性函数 $\widetilde{P}$ 使得，对 $s = (s_1, \cdots, s_n)$，我们有

$$\widetilde{P}(C_s) = \prod_{\substack{k \leqslant n, \\ s_k = 1}} p_k \prod_{\substack{k \leqslant n, \\ s_k = 0}} (1 - p_k).$$

A7) 我们定义 Ω 上的 Borel-代数 $\mathcal{B}(\Omega) = \sigma(\mathcal{F}_\infty)$（我们后面会看到为什么叫它 Borel-代数）。证明，$\mathcal{B}(\Omega)$ 可以由所有的 C_s 生成，其中 $s = (s_1, \cdots, s_n)$ 为一个有限长的字符串。试说明，对于任何一个 $\omega \in \Omega$，我们有 $\{\omega\} \notin \mathcal{F}_\infty$ 但是 $\{\omega\} \in \mathcal{B}(\Omega)$。

A8) 假设 $\{M_p\}_{p\geqslant 1}\in\mathcal{F}_\infty$ 是下降的序列并且 $\bigcap\limits_{p\geqslant 1}M_p=\emptyset$，证明，存在 p_0 使得 $M_{p_0}=\emptyset$。

A9) 证明存在唯一一个 $\mathcal{B}(\Omega)$ 上的概率测度 P，使得对任何 $s=(s_1,\cdots,s_n)$ 我们都有 $P(C_s)=2^{-n}$。

A10) 证明，

$$\Phi:(\Omega,\mathcal{B}(\Omega))\to([0,1],\mathcal{B}([0,1]))$$

是可测映射（其中 $\mathcal{B}([0,1])$ 是 $[0,1]$ 上的 Borel-代数）并且

$$\Phi_*P=m\text{（Lebesgue 测度）}$$

A11) 我们定义

$$\Psi:([0,1],\mathcal{B}([0,1]))\to(\Omega,\mathcal{B}(\Omega))$$

如下：$\Psi(x)$ 为 x 的二进制展开，如果展开方式超过一种，则取从某位开始都是零的那种展开。证明 Ψ 可测映射且 $\Psi_*m=P=m$。

注记. 从而 $(\Omega,\mathcal{B}(\Omega),P)$ 作为测度空间与 $[0,1]$ 加上 Lebesgue 测度同构（在 Ω 上去掉一个零测集后上述映射为双射）。

A12) 证明，第 6) 问中的加性函数也可以唯一地延拓成为 $(\Omega,\mathcal{B}(\Omega))$ 上的一个概率测度。

A13) 令

$$\begin{aligned}A_m&=\{\omega\in\Omega|\omega_k=1,k\geqslant m\},\\ B&=\{\omega\in\Omega|\text{有无限个 }k\text{ 使得 }\omega_k=0\},\\ C&=\{\omega\in\Omega|\text{有无限个 }k\text{ 使得 }\omega_k=0\text{ 还有无限个 }l\text{ 使得 }\omega_l=1\}.\end{aligned}$$

证明，它们都是 $\mathcal{B}(\Omega)$ 中的元素并计算它们的测度。

A14) 假设 $\{m_k\}_{k\geqslant 1}$ 为单调上升的自然数序列（可能是有限项的数列），$\{s_{m_k}\}_{k\geqslant 1}$ 为 0 或者 1 构成的序列，定义

$$C(s_{m_1},s_{m_2},\cdots)=\{\omega\in\Omega|\omega_{m_k}=s_{m_k},k=1,2,\cdots\}.$$

证明，$C(s_{m_1},s_{m_2},\cdots)\in\mathcal{B}(\Omega)$。证明，如果 $\{m_k\}_{k\geqslant 1}$ 仅含有 N 项，那么其测度为 2^{-N}；如果 $\{m_k\}_{k\geqslant 1}$ 有无限项，那么其测度为零。

A15) 给定两个序列 $\{m_k\}_{k=1,2,\cdots}$ 和 $\{m'_k\}_{k=1,2,\cdots}$，如果它们没有公共的项（也就是说 $\{m_k\}\cap\{m'_k\}=\emptyset$，证明，

$$P\big(C(s_{m_1},s_{m_2},\cdots)\cap C(s_{m'_1},s_{m'_2},\cdots)\big)=P\big(C(s_{m_1},s_{m_2},\cdots)\big)P\big(C(s_{m'_1},s_{m'_2},\cdots)\big).$$

A16) 给定两个有限长序列 $\{m_k\}_{k=1,2,\cdots,N}$ 和 $\{m'_k\}_{k=1,2,\cdots,N'}$，假定它们有 $\widetilde{N}$ 个共同的元素，证明，

$$P\big(C(s_{m_1},s_{m_2},\cdots)\cap C(s_{m'_1},s_{m'_2},\cdots)\big)=2^{-(N+N'-\widetilde{N})}\text{ 或 }0.$$

硬币空间上的移位算子及其遍历性

我们定义移位算子 $S:\Omega\to\Omega$：对于 $\omega=(\omega_1,\omega_2,\cdots)\in\Omega$，有

$$S(\omega)=(\omega_2,\omega_3,\cdots),\ \text{即}\ S(\omega)_i=\omega_{i+1}.$$

再定义 σ-代数 $\mathcal{G}_n$ 如下：对于每一个 $k\geqslant n$，令 $\Omega_k=\{\omega\in\Omega\big|\omega_k=1\}$。定义

$$\mathcal{G}_n=\sigma\big(\Omega_n,\Omega_{n+1},\Omega_{n+2},\cdots\big),$$

即包含所有 Ω_k（其中 $k\geqslant n$）的最小的 σ-代数。

B1) 证明，对于 $\mathcal{B}(\Omega)$ 而言，S 是可测映射。

B2) 证明，$\mathcal{G}_2=S^{-1}\big(\mathcal{B}(\Omega)\big)$。

B3) 令 $S^n=\underbrace{S\circ S\circ\cdots\circ S}_{\text{复合 n 次}}$，证明 $\mathcal{G}_{n+1}=\big(S^n\big)^{-1}\big(\mathcal{B}(\Omega)\big)$。

B4) 证明，如下两个论断等价：

(a) $f:(\Omega,\mathcal{G}_{n+1})\to[0,\infty]$ 可测；

(b) $f:(\Omega,\mathcal{B}(\Omega))\to[0,\infty]$ 可测并且 $f(\omega)$ 不依赖于 $\omega_1,\omega_2,\cdots,\omega_n$。

B5) 令 $\mathcal{G}_\infty=\bigcap_n\mathcal{G}_n$。证明，

$$f:\Omega\to[0,\infty],\ \omega\mapsto\limsup_{n\to\infty}\frac{\omega_1+\cdots+\omega_n}{n},$$

是 $\mathcal{G}_\infty$-可测的。

B6) 对每个 $\alpha\in[0,1]$，定义

$$X_\alpha=\big\{\omega\in\Omega\big|\lim_{k\to\infty}\frac{\omega_1+\omega_2+\cdots+\omega_k}{k}=\alpha\big\}.$$

证明：X_α 是一个不可数集合并且 $X_\alpha\in\mathcal{B}(\Omega)-\mathcal{F}_\infty$。

B7) 给定 $s=(s_1,s_2,\cdots,s_n)\in\{0,1\}^{\mathbb{Z}_{>0}}$，我们定义算子

$$T_s:\Omega\to\Omega,\ \omega\mapsto(s_1,\cdots,s_n,\omega_1,\omega_2,\cdots).$$

证明，对每个 $B\in\mathcal{B}(\Omega)$，我们都有 $T_s(B)\in\mathcal{B}(\Omega)$ 并且 $P(T_s(B))=2^{-n}P(B)$。

B8) 证明，如果 $A\in\mathcal{B}(\Omega)$ 使得 $S^{-1}(A)=A$，那么 $P(A\cap C_s)=2^{-n}P(A)$。

B9) 证明，给定 $A\in\mathcal{B}(\Omega)$ 使得 $S^{-1}(A)=A$，证明，

$$\widetilde{\mathcal{B}}=\{B\in\mathcal{B}|P(A\cap B)=P(A)P(B)\}$$

是 σ-代数。

B10) 证明 S 的遍历性，即对于满足 $S^{-1}(A)=A$ 的 $A\in\mathcal{B}(\Omega)$，我们一定有 $P(A)=0$ 或者 $P(A)=1$。

44 乘积测度的构造，Fubini 公式的证明和应用，$\mathbb{R}^n$ 上积分的降维计算：Arichimedes 的墓碑

二零二零年四月九日，星期四，阴转晴

上次课里，我们证明了乘积测度的存在性，即给定 σ-有限的测度空间 $(X,\mathcal{A},\mu)$ 和 $(Y,\mathcal{B},\nu)$，在可测空间 $(X\times Y,\mathcal{A}\otimes\mathcal{B})$ 上存在唯一的 σ-有限的测度 $\mu\otimes\nu$，使得对任意的 $A\in\mathcal{A}$ 和 $B\in\mathcal{B}$，我们都有

$$(\mu\otimes\nu)\big(A\times B\big)=\mu(A)\nu(B).$$

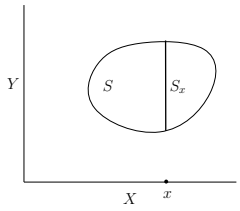
特别地，对任意的 $E\in\mathcal{A}\otimes\mathcal{B}$，它的测度可由如下公式计算

$$(\mu\otimes\nu)\big(E\big)=\inf_{\substack{E\subset\bigcup_{i=1}^{\infty}(A_i\times B_i)\\ A_i\in\mathcal{A},B_i\in\mathcal{B}}}\sum_{i=1}^{\infty}\mu(A_i)\nu(B_i).$$

证明过程的一个重要辅助工具是所谓的粗糙版本的 Fubini 定理，我们现在就它加以改进：

引理 300. *任意给定 $S\in\mathcal{A}\otimes\mathcal{B}$，$x\in X$，我们定义 S 在 x 处的截面S_x 为*

$$S_x=\{y\in Y|(x,y)\in S\}\subset Y.$$



那么，如下三条性质成立

1) $S_x\in\mathcal{B}$;

2) 函数

$$X\to[0,\infty],\quad x\mapsto\nu(S_x),$$

为 $(X,\mathcal{A})$ 上的可测函数;

3) S 的测度可以用截面的测度通过积分得到：

$$\mu\otimes\nu(S)=\int_X\nu(S_x)d\mu(x).$$

注记. *粗糙版本的 Fubini 定理讲的是对于 $S\in\mathcal{R}$，即 S 是有限个两两不交的矩形的并，上面的结论 1)，2) 和 3) 都成立。*

证明: 用集合 $\mathcal{M}$ 表示所有 $\mathcal{A}\otimes\mathcal{B}$ 中满足引理三个结论 1)，2) 和 3) 的集合 S 所组成的集合。粗糙版本的 Fubini 表明 $\widetilde{\mathcal{R}}\subset\mathcal{M}$。如果可以证明 $\mathcal{M}$ 为单调类，根据 $\widetilde{\mathcal{R}}$ 是代数，我们就可以推出 $\mathcal{M}$ 包含了 $\widetilde{\mathcal{R}}$ 所生成的 σ-代数，即 $M\supset\mathcal{A}\otimes\mathcal{B}$，这就说明了 $\mathcal{M}=\mathcal{A}\otimes\mathcal{B}$，从而证明了引理。

我们如下来证明 $\mathcal{M}$ 是单调类。根据单调类的定义，我们需要对上升和下降的子集序列分别进行验证它们的极限的仍然在 $\mathcal{M}$ 中:

(↗) 对任意上升的集合序列 $\{S_i\}_{i\geqslant1}\in\mathcal{M}$，如果当 $i\to\infty$ 时，$S_i\nearrow S$，验证 $S\in\mathcal{M}$。

首先固定的 $x\in X$。根据 $S_i\nearrow S$，我们推出 $(S_i)_x\nearrow S_x$。由于 $S_i\in\mathcal{M}$，根据定义，S_i 满足引理叙述中的 1)，所以 $(S_i)_x\in B$ 所以在 Y 中，从而（因为 $\mathcal{B}$ 是 σ-代数）

$$S=\bigcup_{i\geqslant1}(S_i)_x\in\mathcal{B}.$$

上述对于任意的 $x\in X$ 都成立，这表明 S 满足引理叙述中的 1)。

由于对任意的 $i\geqslant1$，$(S_i)_x\in\mathcal{B}$，$S_x\in\mathcal{B}$，我们可以定义 X 上的函数:

$$f_i(x)=\nu\left((S_i)_x\right),\quad f(x)=\nu\left(S_x\right).$$

由于 $(S_i)_x\nearrow S_x$，所以对任意的 x，$\{f_i(x)\}_{i\geqslant1}$ 是递增的序列并且极限为 $f(x)$，即递增函数列 $\{f_i\}_{i\geqslant1}$ 逐点地收敛到 $f(x)$。根据 $S_i\in\mathcal{M}$ 以及 $\mathcal{M}$ 的定义，f_i 是 $(X,\mathcal{A})$ 上的可测函数。所以，$f:X\to[0,\infty]$ 也是可测的，这表明 S 满足引理叙述中的 2)。

为了证明 S 满足引理叙述中的 3)，我们利用 Beppo Levi 定理:

$$\begin{aligned}\int_X\nu(S_x)d\mu(x)&=\int_X f(x)d\mu(x)\\&\overset{Beppo\ Levi}{=}\lim_{i\to\infty}\int_X f_i(x)d\mu(x)\\&=\lim_{i\to\infty}\int_X\nu(S_{ix})d\mu(x)\\&=\lim_{i\to\infty}\mu\otimes\nu(S_i)\\&=\mu\otimes\nu(S).\end{aligned}$$

其中，倒数第二个等号成立是因为 $S_i\in\mathcal{M}$ 满足引理叙述中的 3)；倒数第一个等号可以取到是因为 $\mu\otimes\nu$ 是测度。

(↘) 对任意下降的集合序列 $\{S_i\}_{i\geqslant1}\in\mathcal{M}$，如果当 $i\to\infty$ 时，$S_i\searrow S$，验证 $S\in\mathcal{M}$。我们分两种情形讨论:

- μ 和 ν 均为有限测度，即 $\mu(X)<\infty$，$\nu<\infty$。

  此时，我们可以参照上升序列的情况如法炮制:

对任意的 $x \in X$，我们有 $(S_i)_x \searrow S_x$。根据 $S_i \in \mathcal{M}$，所以 $(S_i)_x \in B$ 所以在 Y 中，从而 $S \in \mathcal{B}$。我们仍定义函数

$$f_i(x) = \nu\left((S_i)_x\right), \quad f(x) = \nu\left(S_x\right).$$

那么，函数列 $\{f_i\}_{i\geqslant 1}$ 递减并且逐点地收敛到 $f(x)$，这一点是由测度有限保证的。根据 f_i 在 $(X, \mathcal{A})$ 上可测，我们推出 f 也可测。最终，我们可以用 f_1 作为控制函数从而利用 Lebesgue 控制收敛定理：

$$\begin{aligned}\int_X \nu(S_x)d\mu(x) &= \int_X f(x)d\mu(x)\\ &\overset{Lebesgue}{=} \lim_{i\to\infty}\int_X f_i(x)d\mu(x)\\ &= \lim_{i\to\infty}\int_X \nu(S_{ix})d\mu(x)\\ &= \lim_{i\to\infty}\mu\otimes\nu(S_i)\\ &= \mu\otimes\nu(S).\end{aligned}$$

其中，倒数第二个等号成立是因为 $S_i \in \mathcal{M}$ 满足引理叙述中的 3)；倒数第一个等号可以取到是因为 $\mu\otimes\nu$ 是测度。

– μ 和 ν 是 σ-有限的测度。

利用 σ-有限性，先取有限测度的子集 $\{X_p\}_{p\geqslant 1} \subset \mathcal{A}$，使得 $X_p \nearrow X$ 和有限测度的子集 $\{Y_p\}_{p\geqslant 1} \subset \mathcal{B}$，使得 $Y_p \nearrow Y$。利用已经证明的有限测度的情况，我们有

$$S \cap \left(X_p \times Y_p\right) \in \mathcal{M} = \lim_{i\to\infty} S_i \cap \left(X_p \times Y_p\right) \in \mathcal{M}$$

另外，当 $p \to \infty$ 时，$S\cap\left(X_p\times Y_p\right) \nearrow S$。此时，对任意的 p，我们已经有 $S\cap\left(X_p\times Y_p\right) \in \mathcal{M}$。再利用前面关于上升序列的结论，我们就说明了 $S \in \mathcal{M}$。

综上所述，这种版本的 Fubini 定理成立，实际上，这个命题表明，对于形如 $\mathbf{1}_S$，其中 $S \in \mathcal{A}\otimes\mathcal{B}$ 的示性函数，Fubini 定理是成立的，请参考后面 Fubini 定理的叙述。 □

我们可以利用上升的阶梯函数列来逼近正函数，加上 Beppo Levi 定理，我们就可以证明正函数版本的 Fubini 定理：

定理 301 (Fubini 定理：正函数版本)**.** *给定 σ-有限的测度空间 $(X, \mathcal{A}, \mu)$ 和 $(Y, \mathcal{B}, \nu)$，对于可测的正函数*

$$f : (X \times Y, \mathcal{A}\otimes\mathcal{B}, \mu\otimes\nu) \to [0, \infty],$$

我们有

1) – *对任意的 $x \in X$，函数*

$$Y \longrightarrow [0, \infty], \quad y \mapsto f(x, y)$$

是 $\mathcal{B}$-可测的；

– 函数

$$X \longrightarrow [0,\infty], \quad x \mapsto \int_Y f(x,y)d\nu(y),$$

是 $\mathcal{A}$-可测的；

– 上面两个叙述中 X 和 Y 是对称的（可以把 X 换成 Y）。

2) 我们有如下的积分等式：

$$\begin{aligned}\int_{X\times Y} f(x,y)\, d\mu\otimes\nu &= \int_X \Big(\int_Y f(x,y)d\nu(y)\Big)d\mu(x)\\ &= \int_Y \Big(\int_X f(x,y)d\mu(x)\Big)d\nu(y).\end{aligned}$$

注记. 在正函数版本的 *Fubini* 定理中，我们并不要求函数是可积的，也就是说可以容许正函数的积分是无限的。

证明: 这个命题对于形如 $f=\mathbf{1}_S$ 的示性函数（$S\in\mathcal{A}\otimes\mathcal{B}$）是成立的，这是前面引理的内容。利用积分的线性，这个命题对于形如

$$f=\sum_{i\leqslant n}\lambda_i\mathbf{1}_{S_i}$$

的简单函数也成立，其中对任意的 $1\leqslant i\leqslant m$，$S_i\in\mathcal{A}\otimes\mathcal{B}$，系数 λ_i 是非负的。

对于 $(X\times Y,\mathcal{A}\otimes\mathcal{B})$ 上的正可测函数 f，我们选取单调上升的正简单函数列 $\{f_i(x,y)\}_{i\geqslant 1}$，使得该函数列逐点地收敛到 f。对固定的 x，作为 Y 上的函数，我们自然有如下的逐点收敛性：

$$\lim_{i\to\infty} f_i(x,y)=f(x,y).$$

根据 $f_i(x,y)$ 是 $\mathcal{B}$-可测的，所以 $y\mapsto f(x,y)$ 也是 $\mathcal{B}$-可测的。利用 Beppo Levi 定理，我们得到

$$\int_Y f(x,y)d\nu(y)=\lim_{i\to\infty}\int_Y f_i(x,y)d\mu(y).$$

由于 $\displaystyle\int_Y f_i(x,y)d\mu(y)$ 是 $\mathcal{A}$-可测的，所以 $\displaystyle x\mapsto\int_Y f(x,y)d\mu(y)$ 是 $\mathcal{A}$-可测的，至此，定理的第一部分中关于可测性的论断得到了证明。

为了证明第二部分，我们可以将

$$\int_Y f(x,y)d\nu(y)=\lim_{i\to\infty}\int_Y f_i(x,y)d\mu(y).$$

右端视作是为单调上升的（X 上的）正函数列。由于要证明的等式对于 f_i 是成立的，由 Beppo Levi 定理，当 $i\to\infty$ 时，我们就有

$$\begin{aligned}\int_{X\times Y} f(x,y)\, d\mu\otimes\nu &\overset{Beppo\ Levi}{=} \lim_{i\to\infty}\int_{X\times Y} f_i(x,y)\, d\mu\otimes\nu\\ &= \lim_{i\to\infty}\int_X\Big(\int_Y f_i(x,y)d\nu(y)\Big)d\mu(x)\\ &\overset{Beppo\ Levi}{=} \int_X\Big(\int_Y f(x,y)d\nu(y)\Big)d\mu(x).\end{aligned}$$

所以命题成立。 □

最终，我们证明可积函数版本的 Fubini 定理：

定理 302 (Fubini 定理：可积函数版本)**.** 给定 σ-有限的测度空间 $(X,\mathcal{A},\mu)$ 和 $(Y,\mathcal{B},\nu)$，$f(x,y)$ 是 $(X\times Y,\mathcal{A}\otimes\mathcal{B},\mu\otimes\nu)$ 上的定义的复值可积函数（允许取 ∞）。那么，存在两个零测集 $N_X\in\mathcal{A}$ 和 $N_Y\in\mathcal{B}$，使得

1) – 对任意的 $x\in X-N_X$，Y 上的函数

$$Y\to\mathbb{C},\quad y\mapsto f(x,y)$$

是 $\mathcal{B}$-可测的并且这个函数是还是 ν-可积的；

– X 上的函数

$$X\to\mathbb{C},\quad x\mapsto\int_Y f(x,y)d\nu(y)$$

在 X 上几乎处处有定义，这个函数是 $\mathcal{A}$-可测的同时还是 μ-可积的；

– 上述性质对于 X 和 Y 是对称的。

2) 我们有如下的积分等式：

$$\begin{aligned}\int_{X\times Y} f(x,y)\,d\mu\otimes\nu &= \int_X\Big(\int_Y f(x,y)d\nu(y)\Big)d\mu(x)\\ &= \int_Y\Big(\int_X f(x,y)d\mu(x)\Big)d\nu(y).\end{aligned}$$

证明*:* 我们观察到，上面的所有叙述在 $\mathbb{C}$-线性操作下是保持的，所以，通过实部和虚部的分解，我们可以只要对 f 在 $\mathbb{R}\cup\{\pm\infty\}$ 中取值的情况进行证明即可。此时，我们将 f 分解为正和负的两部分：

$$f=f^+-f^-.$$

其中，$f_\pm\leqslant|f|$，所以 $f_\pm$ 的积分（对测度 $\mu\otimes\nu$ 而言）是有限的。根据正函数版本的 Fubini 定理，X 上的函数

$$F_X^+:X\to[0,\infty],\quad x\mapsto\int_Y f^+(x,y)d\nu(y)$$

可测并且其积分有限。然而，我们并不能排除这个函数的取值是 $+\infty$ 的可能。类似的，我们定义

$$F_X^-:X\to[0,\infty],\quad x\mapsto\int_Y f^-(x,y)d\nu(y)$$

令

$$N_X^+=(F_X^+)^{-1}(+\infty),\quad N_X^-=(F_X^-)^{-1}(\infty).$$

由于 $F_X^\pm$ 的积分有限，所以，$N_X^\pm$ 是 μ-零测集。最终，我们令

$$N_X=N_X^+\cup N_X^-.$$

这是 μ-零测集。

现在任意选取 $x \in X - N_X$，作为 Y 上的函数，我们自然有 $f(x,\cdot) = f^+(x,\cdot) - f^-(x,\cdot)$。按照 N_X 的定义，积分 $\int_Y f^{\pm}(x,y)d\nu(y)$ 是有限的，所以，作为 Y 上的函数，$f^{\pm}(x,\cdot)$ 是 ν-可积的。所以，对任意的 $x \in X - N_X$，$f(x,y)$ 作为 Y 上的函数对测度 $\nu(y)$ 是可积的，即 $\int_Y |f(x,y)|d\nu(y) < \infty$。再者，按照定义，对于 $x \in X - N_X$，函数

$$x \mapsto \int_Y f(x,y)d\nu(y) = \int_Y f^+(x,y)d\nu(y) - \int_Y f^-(x,y)d\nu(y),$$

是两个 μ-可积函数之差，所以也是 μ-可积。

最后，定理中的积分等式是通过正函数版本的 Fubini 定理及积分的线性直接得到。至此，我们完成了 Fubini 定理的证明。 □

注记. 在可积版本的 *Fubini* 定理的叙述中，很多地方都有几乎处处的限制，这是不可避免的，除非能排除证明过程中类似于 $N_X^{\pm}$ 类型的集合。

注记 (应用)**.** 在使用正函数版本的 *Fubini* 定理的时候，我们对函数没有要求（只要求可测），这非常得方便；对于可积情形的 *Fubini* 定理，我们要求

$$\int_{X\times Y} |f(x,y)|d(\mu\otimes\nu) < \infty,$$

这是必须验证的条件。

$\mathbb{R}^n$ 上积分自此始

如果不加说明，$\mathbb{R}^n$ 上的测度总是 Lebesgue 测度 m，我们还用 $dx_1 \cdot dx_n$ 表示 dm。

推论 303 (积分的降维计算)**.** 给定 $\mathbb{R}^n$ 中的开集（或者闭集）$\Omega \subset \mathbb{R}^n$。对任意的 $x_n \in \mathbb{R}$，我们定义 x_n 对于在 Ω 的截面：

$$\Omega_{x_n} = \big\{(x_1,\cdots,x_n) \in \mathbb{R}^{n-1} \big| (x_1,\cdots,x_{n-1},x_n) \in \Omega\big\}.$$

那么，对每个几乎处处的 $x_n \in \mathbb{R}$，函数

$$\Omega_{x_n} \to \mathbb{R},\ \ (x_1,\cdots,x_{n-1}) \mapsto f(x_1,\cdots,x_{n-1},x_n)$$

是 $\Omega_{x_n} \subset \mathbb{R}^n$ 上的可积函数。我们还有如下的积分等式：

$$\int_\Omega f(x_1,\cdots,x_{n-1},x_n)dx_1\cdots dx_n = \int_{\mathbb{R}} \left(\int_{\Omega_{x_n}} f(x_1,\cdots,x_{n-1},x_n)dx_1\cdots dx_{n-1} \right) dx_n,$$

其中，当 $\Omega_{x_n} = \emptyset$ 时，无论 $f(x_1,\cdots,x_{n-1};x_n)$ 是否有定义，我们都规定

$$\int_{\Omega_{x_n}} f(x_1,\cdots,x_{n-1},x_n)dx_1\cdots dx_{n-1} = 0.$$

这个推论的证明很简单：

证明: 我们定义 $\mathbb{R}^n$ 上的函数 F:

$$F:\Omega\to\mathbb{R},\quad x\mapsto f(x)\mathbf{1}_\Omega(x).$$

由于 $|F(x)|\leqslant|f(x)|$，所以，F 是可积的。我们对 $\mathbb{R}^n=\mathbb{R}^{n-1}\times\mathbb{R}^1$ 运用 Fubini 定理就给出了上述定理。 □

注记. 这个推论在积分的计算中是最有用的: *Fubini* 定理把一个高维数的积分转化为两个低维数的积分，从而使我们可以用一元微积分的理论归纳地计算多元微积分。我们需要特别地指出，大多数情况下，通过降维得到的都是传统的 *Riemann* 积分（我们已经证明它和 *Lebesgue* 积分是一码事），所以我们可以用任何学习过的关于 *Riemann* 积分的性质。

我们看几个例子:

例子. 计算如下的积分 $\int_\Omega f dx_1\cdots dx_n$，其中

1) $\Omega=[1,2]\times[0,3]\subset\mathbb{R}^2$，$f(x,y)=xy+y^2+1$:

这个函数是有界的，自然可积。按照 *Fubini* 公式，我们知道

$$\begin{aligned}\int_{[1,2]\times[0,3]}f(x,y)dxdy&=\int_0^3\left(\int_1^2 xy+y^2+1dx\right)dy\\&=\int_0^3\underbrace{\frac{y}{2}x^2\Big|_1^2}_{=\frac{3y}{2}}+y^2+1dy=18\frac{3}{4}.\end{aligned}$$

2) $\Omega=[1,2]\times[0,2]\subset\mathbb{R}^2$，$f(x,y)=ye^{xy}$: 这个函数是有界的，自然可积。按照 *Fubini* 公式，我们知道（计算过程中我们可能会去掉零测集 $y=0$）

$$\begin{aligned}\int_{[1,2]\times[0,2]}f(x,y)dxdy&=\int_0^2 y\left(\int_1^2 e^{xy}dx\right)dy\\&=\int_0^3 e^{2y}-e^y dy=\frac{1}{2}e^{6y}+e^y-1.5.\end{aligned}$$

3) Ω 由 $x=1$，$x=4$，$y=(x-2)^2-4$，$y=-(x-3)^2+4$ 围成，$f(x,y)=3x-2y+1$:

对于任意的 $x\in[1,4]$，不难看出，

$$\Omega_x=\left\{y\in\mathbb{R}\middle|(x-2)^2-4\leqslant y\leqslant-(x-3)^2+4\right\}.$$

所以，

$$\begin{aligned}\int_\Omega f(x,y)dxdy&=\int_1^4\left(\int_{(x-2)^2-4}^{-(x-3)^2+4}3x-2y+1dy\right)dx\\&=\int_1^4(3x+1)\left(10x-2x^2-5\right)-y^2\Big|_{(x-2)^2-4}^{-(x-3)^2+4}dx=\text{某个数}.\end{aligned}$$

例子 (Arichimedes 的墓碑). 我们计算如下三个基本的几何对象的测度:

1) 计算单位圆盘 $\Omega = B(1) \subset \mathbb{R}^2$ 的面积 $m_2(B(1))$。

对任意的 $|x| \leqslant 1$，我们有

$$\Omega_x = \left\{ y \in \mathbb{R} \middle| -\sqrt{1-x^2} \leqslant y \leqslant \sqrt{1-x^2} \right\}.$$

从而，

$$\begin{aligned} m_2(B(1)) &= \int_\Omega 1 dxdy = \int_{-1}^1 |\Omega_x| dx \\ &= \int_{-1}^1 2\sqrt{1-x^2} dx. \end{aligned}$$

这就是上学期我们来计算曲线下面积的做法，答案是 π。

2) 计算柱体 $C = B(1) \times [-1,1] = \left\{ (x,y,z) \in \mathbb{R}^3 \middle| (x,y) \in B(1), -1 \leqslant z \leqslant 1 \right\} \subset \mathbb{R}^3$ 的体积 $m_3(C)$。

由于 $B(1) \subset \mathbb{R}^2$ 和 $[-1,1] \subset \mathbb{R}$ 都是 *Borel* 集，所以，$C = B(1) \times [-1,1]$ 是 $\mathbb{R}^3$ 中的一个“矩形”，由于 $m_3 = m_2 \otimes m_2$，它的体积自然是

$$m_3(C) = m_2(B(1)) m_1([-1,1]) = 2\pi.$$

3) 计算单位球 $\Omega = B(1) \subset \mathbb{R}^3$ 的体积 $m_3(B(1))$。

对任意的 $|z| \leqslant 1$，我们有

$$\Omega_z = \left\{ (x,y) \in \mathbb{R} \middle| x^2 + y^2 \leqslant 1 - z^2 \right\}.$$

这是一个半径为 $\sqrt{1-z^2}$ 的球面，它的面积是 $\pi(1-z^2)$. 从而，

$$\begin{aligned} m_3(B(1)) &= \int_\Omega 1 dxdy = \int_{-1}^1 m_2(\Omega_z) dx \\ &= \pi \int_{-1}^1 1 - z^2 dx = \frac{4}{3}\pi. \end{aligned}$$

特别地，我们得到了

$$\frac{m_3(B(1))}{m_3(C)} = \frac{2}{3}.$$

这是传说中 *Archimedes* 墓碑上的图形和数字。

44.1 作业：Archimedes 对抛物线面积的计算，Gauss 积分

清华大学 19-20 春季学期，数学分析二，作业 7

请用 A4 大小的纸张**正反面**用钢笔，签字笔或者圆珠笔书写，并注明自己的姓名，年级和作业的**总页数**。除定理公式所涉及的人名之外，请使用**中文**。本次作业的提交时间和地点为 **2 月 71 日**上午的课堂上，逾期视作零分。

Fubini 定理与积分的计算

A. 经典习题，计算与反例

我们假设 $(X,\mathcal{A},\mu)$ 和 $(Y,\mathcal{B},\nu)$ 是测度空间并且 μ 和 ν 是 σ-有限的。

A1)（Fubini 定理的反例）函数 $f(x,y)$ 在 $[0,1]\times[0,1]$ 上定义：

$$f(x,y)=\begin{cases}\frac{x^2-y^2}{(x^2+y^2)^2}, & (x,y)\neq(0,0);\\ 0, & (x,y)=(0,0).\end{cases}$$

证明，

$$\int_0^1\left(\int_0^1 f(x,y)dy\right)dx\neq\int_0^1\left(\int_0^1 f(x,y)dx\right)dy.$$

Fubini 定理的那个条件没有被满足?

A2)（Fubini 定理的反例）我们考虑测度空间 $([0,1],\mathcal{B},m)$（Borel 代数加上 Lebesgue 测度）和 $([0,1],\mathcal{P}([0,1]),\mu)$，其中 $\mathcal{P}([0,1])$ 是 $[0,1]$ 上所有的子集所构成的 σ-代数，$\mu(A)$ 为 A 中的元素个数。

- 证明，$([0,1],\mathcal{P}([0,1]),\mu)$ 是测度空间但不是 σ 有限的。
- 我们考虑在 $[0,1]\times[0,1]$ 上定义的函数：

$$f(x,y)=\mathbf{1}_{x=y}(x,y).$$

证明，

$$\int_0^1\left(\int_0^1 f(x,y)dy\right)dx\neq\int_0^1\left(\int_0^1 f(x,y)dx\right)dy.$$

Fubini 定理的那个条件没有被满足?

A3) 假设 f 和 g 是 $\mathbb{R}^1$ 上的局部上 Riemann 可积（即在每个闭区间上面都可积）的实值函数。对任意的 $x\in\mathbb{R}$，我们定义

$$F(x)=\int_0^x f(s)ds,\quad G(x)=\int_0^x g(s)ds.$$

证明，对任意的 $a<b$，我们都有

$$F(b)G(b)=F(a)G(a)+\int_a^b F(t)g(t)dt+\int_a^b G(t)f(t)dt.$$

A4) 假设 $(X,\mathcal{A},\mu)$ 和 $(Y,\mathcal{B},\nu)$ 是测度空间并且 μ 和 ν 是 σ-有限的。

- 证明，集合 $D_\mu=\{x\in X|\mu(\{x\})>0\}$ 是可数的。
- 证明，集合 $\Delta=\{(x,y)\in\mathbb{R}^1\times\mathbb{R}^1|x=y\}\subset X\times Y$ 是 $\mathcal{A}\otimes\mathcal{B}$-可测的。
- 证明如下的公式

$$\mu\otimes\nu(\Delta)=\sum_{x\in D_\mu}\mu(\{x\})\nu(\{x\}).$$

A5)（Archimedes，公元前三世纪）假设 Γ 是 $y=-cx^2$ 在平面上所定义的抛物线，其中 $c>0$。$A,B\in\Gamma$ 是给定的两点，AB 是这两点所连的线段（我们称它为 Γ 的一条弦）。

- 证明，Γ 是 $\mathbb{R}^2$ 中的 1 维光滑子流形。进一步证明，存在唯一的点 $C\in\Gamma$，使得 Γ 在 C 处的切线与 AB 平行。
- 假设 C 的坐标是 (C_1,C_2)，那么由 $x=C_1$ 定义的直线与 AB 相交于 AB 的中点。
- 证明，由 Γ 与 AB 所围成的图形的面积是三角形 ABC 面积的 $\dfrac{4}{3}$ 倍。

A6)（Gauss 积分的计算）试计算下面的积分：

- $\displaystyle\int_{x^2+y^2\leqslant R^2}e^{-(x^2+y^2)}dxdy$, 其中 $R>0$;
- $\displaystyle\int_{\mathbb{R}}e^{-ax^2+bx+c}dx$, 其中 $a>0$;
- $\displaystyle\int_{\mathbb{R}^n}\exp(-\frac{1}{2}\sum_{i,j=1}^{n}A_{ij}x_ix_j)dx$, 其中 $A=(A_{ij})$ 是 $n\times n$ 的正定（对称）矩阵;
- $\displaystyle\int_{\mathbb{R}^n}\exp(-\frac{1}{2}\sum_{i,j=1}^{n}A_{ij}x_ix_j+\sum_{i=1}^{n}B_ix_i)dx$, 其中 $A=(A_{ij})$ 是 $n\times n$ 的正定（对称）矩阵.

A7)* 假设 $f:\mathbb{R}^n\to\mathbb{R}$ 是 Borel-可测的函数，证明，存在零测集（Lesbegue 意义下）$N\subset\mathbb{R}$，使得对任意的 $y\notin N$，我们有

$$m\left(\{x\in\mathbb{R}^n|f(x)=y\}\right)=0.$$

积分的计算

注记 (符号约定). 人们通常用 $\iint$ 表示在 $\mathbb{R}^2$ 上的区域上积分，用 $\iiint$ 表示在 $\mathbb{R}^3$ 上的区域上积分，这和平时的一个积分号 $\int$ 表达的意思是一致的。

B. 计算积分：第一组

$$(B1)\iint_{[0,\pi/2]\times[0,1]} x\cos(xy)dxdy \qquad (B2)\iint_{[0,\pi]^2}\sin(x+y)dxdy$$

$$(B3)\iint_{[3,4]\times[1,2]}\frac{dxdy}{(x+y)^2} \qquad (B4)\iint_{[0,2]^2}[x+y]dxdy$$

$$(B5)\iint_{[0,\pi]^2}|\cos(x+y)|dxdy \qquad (B6)\iint_{\{|x|+|y|\leqslant 1\}}e^{x+y}dxdy$$

$$(B7)\iint_{\{x^2+y^2\leqslant a^2\}}|xy|dxdy \qquad (B8)\iint_{\{x^2+y^2\leqslant a^2\}}x\cos(xy)dxdy.$$

$$(B9)\iint_{\{\frac{x^2}{a^2}+\frac{y^2}{b^2}\leqslant 1\}}\sqrt{1-\frac{x^2}{a^2}-\frac{y^2}{b^2}}dxdy \qquad (B10)\iint_{\{x^4+y^4\leqslant 1\}}|xy|dxdy$$

$$(B11)\iint_{\{x^2+y^2\leqslant x+y\}}\sqrt{x^2+y^2}dxdy \qquad (B12)\iint_{\{x^2+y^2\leqslant Rx\}}\sqrt{R^2-x^2-y^2}dxdy.$$

C. 计算积分：第二组

$(C1)\iint_D y^2dxdy,\quad D$由旋轮线$\{(a(t-\sin t),a(1-\cos t)),t\in[0,2\pi]\}$与$y=0$围成.

$(C2)\iint_D e^{\frac{y}{x+y}}dxdy,\quad D$是以$(0,0),(0,1),(1,0)$为顶点的三角形.

$(C3)\iint_D(x-y)^2\sin(x+y)dxdy,\quad D$是以$(\pi,0),(2\pi,0),(\pi,2\pi),(0,\pi)$为顶点的正方形.

$(C4)\iint_D x^2+y^2dxdy,\quad D$由曲线$x^2-y^2=1,x^2-y^2=2,xy=1,xy=2$围成.

$(C5)\iint_D\frac{1}{xy}dxdy,\quad D$由四条抛物线$y^2=px,y^2=qx,x^2=ay,x^2=by(0<p<q,0<a<b)$围成.

$(C6)\iiint_D\frac{dxdydz}{(1+x+y+z)^3},\quad D=\{(x,y,z)\in\mathbb{R}^3:x+y+z\leqslant 1,x,y,z\geqslant 0\}.$

$(C7)\iiint_D(x^2+y^2+z^2)dxdydz,\quad D$由球面$x^2+y^2+z^2=a^2$与锥面$z=\sqrt{x^2+y^2}$围成.

$(C8)\iiint_D zdxdydz,\quad D$由两个球面$x^2+y^2+z^2=2az,x^2+y^2+z^2=az$围成.

D. 计算体积

计算下列集合的体积:

D1) n 维单形 $\{(x_1,\cdots,x_n)\in\mathbb{R}^n : x_1,\cdots,x_n\geqslant 0, x_1+\cdots x_n\leqslant a\}, a>0.$

D2) 心脏线 $\{(x,y)\in\mathbb{R}^2 : (x^2+y^2)^2+4ax(x^2+y^2)-4a^2y^2=0\}$ 围成的区域面积.

D3) 球体 $x^2+y^2+z^2\leqslant a^2$ 被圆柱体 $x^2+y^2=ax$ 截下的立体的体积.

I never failed in mathematics. Before I was fifteen I had mastered differential and integral calculus.

—- Albert Einstein

---

45 抽象换元积分公式，Borel 测度的正则性引理，微分同胚下换元积分公式的证明

二零二零年四月十三日，星期一，晴

我们上一周证明了 Fubini 定理，它讲的是，给定 σ-有限的测度空间 $(X,\mathcal{A},\mu)$ 和 $(Y,\mathcal{B},\nu)$，$X\times Y$ 上的函数 f 是正可测的或者是复值可积函数，那么 f 在 $X\times Y$ 上的积分可以通过每个分量的积分来计算:

$$\begin{aligned}\int_{X\times Y} f(x,y)\,d\mu\otimes\nu &= \int_X \Big(\int_Y f(x,y)d\nu(y)\Big)d\mu(x)\\ &= \int_Y \Big(\int_X f(x,y)d\mu(x)\Big)d\nu(y).\end{aligned}$$

注记 (Fubini 用来交换积分顺序)**.** 除了在计算积分时可以降维之外，*Fubini* 定理还有其它的应用: 它表明在 X 上的积分运算与在 Y 上的积分运算是可交换的，即

$$\int_X \Big(\int_Y f(x,y)d\nu(y)\Big)d\mu(x) = \int_Y \Big(\int_X f(x,y)d\mu(x)\Big)d\nu(y),$$

这个可以把在某些空间上积分运算转化为在另一个空间上的积分运算（可能更简单），下面的命题是一个典型的（重要）例子，它在基本的调和分析理论中有很多应用。

推论 304. *f 是测度空间 $(X,\mathcal{A},\mu)$ 上的正函数，其中*

$$f: X\to[0,\infty)$$

在 $\mathbb{R}_{\geqslant 0}$ 上取值。一元函数 $g:[0,\infty)\to[0,\infty)$ 是递增的并且连续可微，它满足 $g(0)=0$，那么

$$\int_X g\circ f d\mu = \int_{[0,\infty)} g'(t)\mu(\{x|f(x)\geqslant t\})dt.$$

特别地，我们有

$$\int_X f d\mu = \int_{[0,\infty)} \mu\left(\{x|f(x)\geqslant t\}\right)dt.$$

证明: 根据 Newton-Leibiniz 公式，我们有

$$\begin{aligned}\int_X g\circ f d\mu &= \int_X \left(g(f(x))-g(0)\right)d\mu\\ &= \int_X \left(\int_{[0,f(x)]} g'(x)ds\right)d\mu\\ &= \int_X \left(\int_{[0,\infty)} g'(s)\mathbf{1}_{[0,f(x)]}(s)ds\right)d\mu.\end{aligned}$$

根据正函数版本的 Fubini 定理，我们有

$$\begin{aligned}\int_X g\circ f d\mu &= \int_{X\times[0,\infty)} \left(g'(s)\mathbf{1}_{[0,f(x)]}(s)\right) ds\otimes d\mu \\ &= \int_{[0,\infty)} \left(\int_X g'(s)\mathbf{1}_{[0,f(x)]}(s)d\mu\right) ds \\ &= \int_{[0,\infty)} g'(s)\mu(\{x|f(x)\geqslant s\})ds.\end{aligned}$$

特别地，当 $g(s)=s$ 时，我们就得到了第二个等式。 □

换元积分公式

换元积分公式是计算积分的另一个重要手段，为了给出一个相对漂亮的表述，我们先进行一些抽象的表述。

给定测度空间 $(X,\mathcal{A},\mu)$，我们总是假设 μ 是 σ-有限的。我们考虑 X 上的正可测函数

$$\rho: X\to[0,\infty].$$

我们假设这个函数是几乎处处有界的，即在一个零测集之外，ρ 是有界的，即存在 $N\in\mathcal{A}$，$\mu(N)=0$，使得对任意的 $x\notin N$，我们都有

$$\rho(x)<\infty.$$

我们要定义一个新的测度：

$$\nu:=\rho\mu:\mathcal{A}\to[0,\infty],\quad A\mapsto\int_X\rho(x)\mathbf{1}_A(x)d\mu(x).$$

如果 $A,B\in\mathcal{A}$，$A\cap B=\emptyset$，根据积分的线性，我们有

$$\begin{aligned}\nu(A\cup B) &= \int_X\rho(x)\mathbf{1}_{A\cup B}(x)d\mu(x) \\ &= \int_X\rho(x)\mathbf{1}_A(x)d\mu(x)+\int_X\rho(x)\mathbf{1}_B(x)d\mu(x) \\ &= \nu(A)+\nu(B).\end{aligned}$$

所以，ν 是加性函数。

我们再任意选取单调上升的序列 $\{A_j\}_{j\geqslant1}\subset\mathcal{A}$，使得 $A_j\nearrow A$。根据 Beppo Levi 定理，我们有

$$\nu(A_j)=\int_X\rho\cdot\mathbf{1}_{A_i}d\mu\to\int_X\rho\cdot\mathbf{1}_A d\mu=\nu(A),$$

这表明 ν 是测度。

我们还可以说明这是一个 σ-有限性，要点是用 ρ 是几乎处处有界的。根据 μ 的 σ-有限性，我们选取 $\{X_p\}_{p\geqslant1}\subset\mathcal{A}$，使得 $X_p\nearrow X$ 使得并且对每个 p，我们都有 $\mu(X_p)<\infty$。我们现在定义如下的集合

$$Y_p=\{x|\rho(x)\leqslant p\}\cap X_p.$$

很明显，$\{Y_p\}_{p\geqslant 1}\subset\mathcal{A}$ 是上升的并且 $Y_p\nearrow X$。为了说明 $\nu(Y_p)\leqslant\infty$，我们只要注意到

$$\begin{aligned}\nu(Y_p)&=\int_X\rho(x)\mathbf{1}_{Y_p}(x)d\mu(x)\\&\leqslant\int_X p\mathbf{1}_{X_p}(x)d\mu(x)\\&=p\mu(X_p)<\infty.\end{aligned}$$

总结上面的证明，我们有如下的结论：

定义 305. 给定 σ-有限的测度空间 $(X,\mathcal{A},\mu)$，正可测函数 ρ 几乎处处有界，那么 $\nu=\rho\mu$ 是 $(X,\mathcal{A})$ 上 σ-有限的测度。我们将 ν 称作是 $(X,\mathcal{A},\mu)$ 上**以 ρ 为密度的测度**。

对于以 ρ 为密度的测度的测度 ν 以及 $(X,\mathcal{A},\mu)$ 上的可测函数。我们可以证明 $f\rho$ 对于测度 μ 可积当且仅当 f 对于测度 ν 可积，并且此时有

$$\int_X f(x)d\nu(x)=\int_X f(x)\rho(x)d\mu(x).$$

我们把这个性质的证明留作习题。

我们先证明一个抽象版本的换元积分公式（漂亮但是用途不大）。给定测度空间 $(X,\mathcal{A},\mu)$ 和可测空间 $(Y,\mathcal{B})$，考虑可测映射

$$\Phi:(X,\mathcal{A},\mu)\to(Y,\mathcal{B}).$$

我们已经证明过，我们可以将测度 μ 用 Φ 推出来定义 $(Y,\mathcal{B})$ 的测度 $\Phi_*\mu$：对任意的 $B\in\mathcal{B}$，我们定义

$$\left(\Phi_*\mu\right)(B)=\mu\left(\Phi^{-1}(B)\right).$$

我们要指出，这样得到的测度未必是 σ-有限的，比如考虑映射

$$\Phi:\mathbb{R}^2\to\mathbb{R},(x,y)\mapsto x.$$

那么，$\mathbb{R}^1$ 上的 Lebesgue 测度 Φ_*m_2 就不是 σ-有限的，因为对任意的 $A\subset\mathbb{R}^1$，如果 $m_1(A)>0$，那么 $\left(\Phi_*\mu\right)(A)=+\infty$，同学们会在本次作业中完成这个证明。抽象的换元积分公式如下：

定理 306. 给定测度空间 $(X,\mathcal{A},\mu)$，可测空间 $(Y,\mathcal{B})$，$(Y,\mathcal{B})$ 上可测函数 f 以及这两个空间之间的可测映射

$$\Phi:(X,\mathcal{A},\mu)\to(Y,\mathcal{B}).$$

那么，f 在 $(Y,\mathcal{B},\Phi_*(\mu))$ 上可积当且仅当 $\big(f\circ\Phi\big)(x)$ 在 $(X,\mathcal{A},\mu)$ 上可积。在此情形下，我们还有

$$\int_Y f(y)d\nu(y)=\int_X\big(f\circ\Phi\big)(x)d\mu(x).$$

其中 $\nu=\Phi_*\mu$。

证明: 这个证明过程只需要照章办事：首先，如果 $f=\mathbf{1}_B$ 是示性函数，其中 $B\in\mathcal{B}$，那么，

$$\begin{aligned}\int_Y \mathbf{1}_B d\nu &= \nu(B)=\mu(\Phi^{-1}(B))\\ &=\int_X \mathbf{1}_{\Phi^{-1}(B)}(x)d\mu(x)=\int_X \mathbf{1}_B\circ\Phi d\mu.\end{aligned}$$

所以，命题明显成立。所以，通过线性命题对一切简单正函数上式都成立。对于一般的正函数，f，我们选取单调上升的简单正 $\{f_i\}_{i\geqslant 1}$ 序列，使得它们逐点收敛到 f。那么，根据 Beppo Levi 定理，我们有

$$\begin{aligned}\int_Y f d\nu &= \lim_{i\to\infty}\int_Y f d\nu\\ &=\lim_{i\to\infty}\int_X f_i\circ\Phi d\mu\\ &=\int_X f\circ\Phi d\mu.\end{aligned}$$

从而，该定理对正函数也成立。特别地，由于 $|f\circ\Phi|=|f|\circ\Phi$，从而 f 在 $(Y,\mathcal{B},\Phi_*(\mu))$ 上可积当且仅当 $\big(f\circ\Phi\big)(x)$ 在 $(X,\mathcal{A},\mu)$ 上可积。为了验证可积函数的等式，我们只要将函数分解为正负部分或这实虚部利用线性即可，我们略去冗长无聊的细节。 □

我们现在正式进入 $\mathbb{R}^n$ 上的换元积分公式（对 Lebesgue 测度而言）。首先，我们引入必要的记号。

假定 Ω_1 和 Ω_2 是 $\mathbb{R}^n$ 中的两个开集，映射

$$\Phi:\Omega_1\to\Omega_2$$

是微分同胚（只要要求是 C^1-同胚即可，即 Φ 与 Φ^{-1} 都是 C^1 的）。如果用坐标来写，我们把 Φ 的坐标函数写成

$$\Phi:\Omega_1\to\Omega_2,\ \ (x_1,\cdots,x_n)\mapsto\big(\Phi_1(x_1,\cdots,x_n),\cdots,\Phi_n(x_1,\cdots,x_n)\big)$$

映射 Φ 的微分可以用它的 Jacobi 行列式来写，为了后面方便起见，我们把它简记为:

$$\mathbf{J}_\Phi(x_1,\cdots,x_n)=\big|\mathrm{Jac}(\Phi)(x_1,\cdots,x_n)\big|=\det\left(\frac{\partial\Phi_i}{\partial x_j}\right)\Big|_{x=(x_1,\cdots,x_n)}.$$

定理 307 (换元积分公式)**.** 假定 Ω_1 和 Ω_2 是 $\mathbb{R}^n$ 中的两个开集，映射

$$\Phi:\Omega_1\to\Omega_2$$

是微分同胚。我们用 dx 和 dy 分别表示开集 Ω_1 和 Ω_2 上的 *Lebesgue* 测度。

对于 Ω_2 上的可测函数

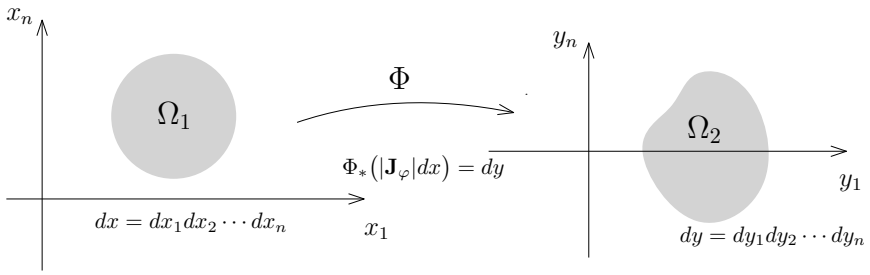
$$f:\Omega_2\to\mathbb{C}$$

而言，它在 Ω_2 上对测度 dy 是可积的当且仅当 $f\circ\Phi$ 在 Ω_1 对带密度的测度 $|\mathbf{J}_\Phi(x)|dx$ 是可积的。此时，我们进一步有

$$\Phi_*\big(|\mathbf{J}_\Phi|dx\big)=dy.$$

用积分的语言表达：对任意的 Ω_2 上对 dy 可积的函数 f，我们有

$$\int_{\Omega_2} f(y)dy=\int_X \big(f\circ\Phi\big)(x)|\mathbf{J}_\Phi(x)|dx.$$



注记. 记住（不是证明）上面的公式可以用如下的窍门：将 $y=\Phi(x)$ 直接代入左边，$f(y)$ 就变成了 $\big(f\circ\Phi\big)(x)$；另外，对于微分而言，我们有 $dy=d\Phi\circ dx$，我们然后将 $d\Phi$ 替换成它的行列式的绝对值 $|\mathbf{J}_\Phi(x)|$ 即可。

换元积分公式是本学期最困难的证明之一，我们要分若干步骤来完成。在考察一般的微分同胚之前，我们先研究比较特殊的一种微分同胚：仿射变换。

我们假定 Φ 为仿射变换，也就是说它是一次函数：

$$\Phi:\Omega_1\to\Omega_2,\quad x\mapsto\Phi(x)=A\cdot x+x_0,$$

这里我们把 $x\in\mathbb{R}^n$ 看作是列向量，其中 $x_0\in\mathbb{R}^n$，A 为 $n\times n$ 的实系数可逆矩阵。此时，

$$\Phi^{-1}(y)=A^{-1}\cdot y-A^{-1}\cdot x_0.$$

如果 A 为单位矩阵，此时 Φ 就是平移变换 τ_{x_0}，此时 $\Omega_2=\tau_{x_0}\Omega_1$。根据 Lebesgue 测度的平移不变性，我们有

$$\big(\tau_{x_0}\big)_*m=m.$$

此时，我们显然有 $\big|\mathbf{J}_\Phi\big|=1$。所以，利用抽象版本的换元积分公式就有

$$\begin{aligned}\int_{\Omega_1} f\circ\Phi dx&=\int_{\Omega_1} f(x+x_0)dx\\&=\int_{\Omega_2} f(y)d\big(\tau_{x_0}\big)_*m=\int_{\Omega_2} f(y)dy.\end{aligned}$$

这表明，对于 f（和 Ω_1）复合上任何一个平移都不会改变其积分。所以，通过对 Φ 复合上某个平移变换，我们只需要考虑 $x_0=0$ 的情形即可。

我们现在假设

$$\Phi(x) = A \cdot x.$$

根据矩阵的极分解定理，A 可以写为

$$A = O \cdot S,$$

其中 O 为正交矩阵，S 为对称矩阵[11]。从几何上来看，O 对应着旋转而 A 对应着不同方向上的伸缩（需要进一步用正交矩阵来对角化）。我们已经证明了 Lebesgue 测度在正交变换下不变（作业五 A11)），我们可以照搬上述关于平移的论证，通过对 Φ 复合上某个正交变换，从而将命题约化为 A 是对称的情况。另外，每个对称矩阵都可以通过正交矩阵对角化，所以，再次通过复合正交矩阵，可以进一步假 A 为对角矩阵：（在下图中，R 是正交矩阵，Λ 是对角矩阵）

$$\begin{array}{ccc} \Omega_1 & \xrightarrow{A} & \Omega_2 \\ \downarrow{\scriptstyle R} & & \downarrow{\scriptstyle {}^tR} \\ R(\Omega_1) & \xrightarrow{\Lambda} & {}^tR(\Omega_2) \end{array},$$

根据上面的讨论，最终，我们只要对如下的映射来证明命题即可：

$$\Phi : (x_1, \cdots, x_n) \mapsto (\lambda_1 x_1, \cdots, \lambda_n x_n),$$

其中 $\lambda_1, \cdots, \lambda_n$ 都是正实数。

我们要用 Fubini 公式降低维数 n 进行计算。为此，我们先考虑 $n = 1$ 的情形。首先，我们已经证明过对于伸缩变换

$$\rho_\lambda : \mathbb{R}^1 \to \mathbb{R}^1, x \mapsto \lambda x,$$

其中 $\lambda > 0$，我们有

$$(\rho_\lambda)_* m = \lambda^{-1} m.$$

我们现在假设 $\Omega_2 = \rho_\lambda(\Omega_1)$。根据抽象版本的换元积分公式，我们有

$$\int_{\Omega_1} f(\lambda x) dx = \int_{\Omega_2} f(y)(\rho_\lambda)_* m = \lambda^{-1} \int_{\Omega_2} f(y) dy.$$

所以命题成立。对于一般的维数 n，我们用 Fubini 公式。为了书写简洁，我们用 $d\bar{x}$ 表示 $dx_1 \cdots dx_{n-1}$，

[11]实际上，由于 ${}^tA \cdot A$ 为正定矩阵，我们可以取对称矩阵 S 使得 $S^2 = {}^tA \cdot A$，此时 $O = A \cdot S^{-1}$。这个分解是唯一的，称作是可逆矩阵的**极分解**

用 $d\bar{y}$ 表示 $dy_1\cdots dy_{n-1}$

$$\begin{aligned}
&\int_{\Omega_1} f(\lambda_1 x_1,\cdots,\lambda_n x_n)d\bar{x}\otimes dx_n\\
=&\int_{\mathbb{R}^n} f(\lambda_1 x_1,\cdots,\lambda_n x_n)\mathbf{1}_{\Omega}(x_1,\cdots,x_n)d\bar{x}\otimes dx_n\\
=&\int_{\mathbb{R}}\left(\int_{\mathbb{R}^{n-1}} f(\lambda_1 x_1,\cdots,\lambda_n x_n)\mathbf{1}_{\Omega_1}(x_1,\cdots,x_n)d\bar{x}\right)dx_n\\
=&(\lambda_1\cdots\lambda_{n-1})^{-1}\int_{\mathbb{R}}\left(\int_{\mathbb{R}^{n-1}} f(y_1,\cdots,y_{n-1},\lambda_n x_n)\mathbf{1}_{\Omega_1}\left(\frac{y_1}{\lambda_1},\cdots,\frac{y_{n-1}}{\lambda_{n-1}},x_n\right)d\bar{y}\right)dx_n\\
=&(\lambda_1\cdots\lambda_n)^{-1}\int_{\Omega_2} f(y)dy.
\end{aligned}$$

注意到 $(\lambda_1\cdots\lambda_n)^{-1}$ 恰好是 $\mathbf{J}_\Phi$ 的倒数，所以命题成立。

综上所属，当 Φ 为仿射变换时，我们就证明了换元积分公式。为了证明一般的情况，需要一个关于证明 $\mathbb{R}^n$ 中 Borel-集上的正则性（对于 Lebesgue 测度而言），这是一个技术性的引理，本身也很有意义：

定理 308 (正则性定理)**.** *我们在 $\mathbb{R}^n$ 上的 Borel-代数 $\mathcal{B}(\mathbb{R}^n)$ 上给定满足如下条件的测度 μ:*

- *如果 $K\subset\mathbb{R}^n$ 是紧集，我们有 $\mu(K)<\infty$。*

那么，对于任意的 $A\in\mathcal{B}(\mathbb{R}^n)$ 和任意的 $\varepsilon>0$，存在开集 U 包含 K 和被 K 包含的闭集 F（即 $F\subset A\subset U$）使得

$$\mu(U-F)<\varepsilon.$$

证明: 我们定义

$$\mathcal{A}=\left\{A\in\mathcal{B}(\mathbb{R}^n)\middle|\text{对任意 }\varepsilon>0\text{，存在开集 }U\supset A\text{ 和闭集 }F\subset A\text{，使得 }\mu(U-F)<\varepsilon\right\}.$$

我们注意到，如果 K 是紧集，那么 $K\in\mathcal{A}$：由于紧集是闭集，我们取 $F=K$；对任意的 $k\in\mathbf{Z}_{\geqslant 1}$，考虑开集

$$U_k=\left\{x\in\mathbb{R}^n\middle|d(x,K)<\frac{1}{k}\right\}.$$

其中，距离函数 $d(x,K)$ 的定义如下：

$$d(x,K)=\inf_{y\in K}d(x,y).$$

由于 K 是紧集，所以函数

$$K\to\mathbb{R},\quad y\mapsto d(x,y)$$

的最小值实际上可以取到。特别地，$x\in K$ 当且仅当 $d(x,K)=0$。另外，由于 K 是有界的，所以，对任意的 k，U_k 也有界（包含在某个有界闭球中），从而 $\mu(U_1)<\infty$。很明显，我们有 $U_k\searrow K$，根据测度与极限可交换性，我们就有 $\mu(U_k)\to\mu(K)$。从而，存在 k_0，使得 $\mu(U_{k_0}-K)<\varepsilon$，我们选取 $U=U_{k_0}$ 即可。

为了证明这个命题，我们要说明 $\mathcal{A}$ 包含了所有的 Borel 集，为此，只要证明 $\mathcal{A}$ 是 σ-代数即可。

很明显，$\emptyset \in \mathcal{A}$。另外，$\mathcal{A}$ 在取补集的操作下封闭，这非常容易证明：假设 $A \in \mathcal{A}$。对任意 $\varepsilon > 0$，存在开集 $U \supset A$ 和闭集 $F \subset A$，使得 $\mu(U - F) < \varepsilon$，所以，对于其补集 A^c，我们有 $F^c \supset A^c \supset U^c$，此时，$F^c$ 为开集，U^c 为闭集。另外，

$$\mu(F^c - U^c) = \mu(U - F) < \varepsilon.$$

所以，$A^c \in \mathcal{A}$。

现在来证明 $\mathcal{A}$ 在取可数并的操作下封闭。任意给定序列 $\{A_i\}_{i\geqslant 1} \subset \mathcal{A}$。根据 $\mathcal{A}$ 的定义，对任意的 $i \geqslant 1$，存在开集 U_i 和闭集 F_i，使得

$$F_i \subset A_i \subset U_i, \quad \mu(U_i - F_i) < \frac{\varepsilon}{2^i}.$$

我们定义

$$\widetilde{F} = \bigcup_{i\geqslant 1} F_i, \quad U = \bigcup_{i\geqslant 1} U_i,$$

那么，我们显然有

$$\widetilde{F} \subset \bigcup_{i\geqslant 1} A_i \subset U,$$

并且

$$\begin{aligned} \mu(U - \widetilde{F}) &\leqslant \sum_{i\geqslant 1} \mu(U_i - F_i) \\ &< \sum_{i\geqslant 1} \frac{\varepsilon}{2^i} = \varepsilon. \end{aligned}$$

U 显然是开集。但是，我们并不能保证 $\widetilde{F}$ 为闭集。为了对 $\widetilde{F}$ 进行一定的改造，我们只要能证明下述引理即可（从而完成正则性定理的证明）：

引理 309. *测度 μ 在 $\mathcal{B}(\mathbb{R}^n)$ 上定义，它在任意的紧集上取值有限。集合 $G = \bigcup_{i\geqslant 1} F_i$ 是可数个闭集的并，那么对任意的 $\varepsilon > 0$，存在闭集 $F \subset G$，使得*

$$\mu(G - F) < \varepsilon.$$

分两种情况来证明引理：

1) G 的测度有限，即 $\mu(G) < \infty$。

对任意的 $j \geqslant 1$，我们定义

$$F_j = \bigcup_{i\leqslant j} F_i.$$

这是一列上升的闭集序列并且 $F_j \nearrow G$。特别地，我们有 $\lim\limits_{j\to\infty} F_j = G$，从而当 $j \to \infty$ 时，$\mu(G - F_j) \to 0$。据此，只需要取 $F = F_{k_0}$，其中 k_0 比较大即可。

2) G 的测度无限，即 $\mu(G)=\infty$。

我们考虑 G 和环面的交

$$G_k = G \cap \big\{ x \in \mathbb{R}^n \big| k-1 \leqslant |x| \leqslant k \big\}.$$

我们注意到 $\mu(G_k)<\infty$ 并且 G_k 也是可数个闭集的并：$G_k = \bigcup\limits_{i\geqslant 1} G_k \cap F_i$。根据上一情形，对任意的 $k \geqslant 1$，存在闭集 $H_k \subset G_k$，使得

$$\mu(G_k - H_k) < \frac{\varepsilon}{2^k}.$$

我们现在令

$$F = \bigcup_{k\geqslant 1} H_k.$$

我们自然有

$$\mu(G-F) \leqslant \sum_{k\geqslant 1} \mu(G_k - H_k) < \varepsilon.$$

为了说明 F 为闭集，我们现在利用分解 $G = \bigcup\limits_{k\geqslant 1} G_k$ 的最重要的性质：对任意的 k, k'，如果 $|k-k'| \geqslant 2$，那么 $G_k \cap G_{k'} = \emptyset$。任意一个 F 中的收敛点列一定会落在某个 $G_k \cup G_{k+1}$ 中，从而落在 $H_k \cup H_{k+1}$ 中（这是闭集），所以 F 是闭集。

这就完成了正则性定理的证明。 □

我们现在正式开始换元积分公式的证明。

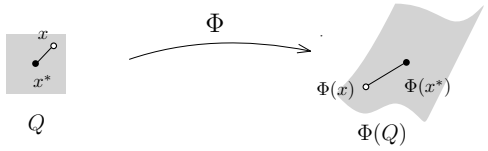
换元积分公式的证明. 我们分成五个步骤来完成这一任务：

第一步，正方体的体积在 Φ 下变换的控制：假定 Q 是一个边长为 $h>0$ 的闭正方体，那么

$$m(\Phi(Q)) \leqslant \big(\sup_{x\in Q} \|\mathrm{Jac}(\Phi)(x)\| \big)^n m(Q),$$

其中，对任意 $x=(x_1,\cdots,x_n)\in\mathbb{R}^n$ 和 $n\times n$ 的矩阵 $A=(A_{ij})$，我们用如下的范数：

$$\|x\| = \sup_{i\leqslant n} |x_i|, \quad \|A\| = \sup_{i\leqslant n} \big(\sum_{j=1}^n |A_{ij}| \big).$$



我们对 Φ 的每个分量用 Lagrange 中值定理。假设 x^* 为 Q 的中心，x 为 Q 中任意一点，那么，对于指标 $i \leqslant n$，我们有

$$\Phi_i(x) - \Phi_i(x^*) = \sum_{j \leqslant n} \frac{\partial \Phi_i}{\partial x_j}(\xi_i)(x_j - x_j^*),$$

其中 ξ_i 为线段 $\overline{xx^*}$ 上的一点（从而，$\xi_i \in Q$）。这样，我们得到

$$\begin{aligned}
\|\Phi(x) - \Phi(x^*)\| &= \sup_{i \leqslant n} |\Phi_i(x) - \Phi_i(x^*)| \\
&\overset{Lagrange}{\leqslant} \sup_i \sum_{j \leqslant n} |\frac{\partial \Phi_i}{\partial x_j}(\xi_j)||(x_j - x_j^*)| \\
&\leqslant \|x - x^*\| \sup_{i \leqslant n} \sum_{j \leqslant n} |\frac{\partial \Phi_i}{\partial x_j}(\xi_j)| \\
&\leqslant \frac{h}{2} \sup_{x \in Q} \|\mathrm{Jac}(\Phi)(x)\|.
\end{aligned}$$

最后一步，我们还用到了 $\|x - x_*\| \leqslant 0.5 \times h$。从而，$\Phi(Q)$ 落在以 $\Phi(x^*)$ 为中心且边长不超过 $\sup_{x \in Q} \|\mathrm{Jac}(\Phi)(x)\| h$ 的正方体里面，这个方块的体积自然不超过 $\sup_{x \in Q} \|\mathrm{Jac}(\Phi)(x)\|^n m(Q)$。

第二步，假定 Q 是闭正方体，那么我们有如下不等式：

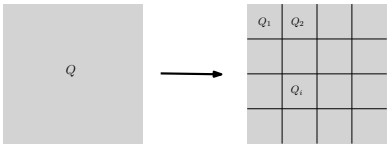
$$m(\Phi(Q)) \leqslant \int_Q |\mathbf{J}_\Phi(x)| dx$$

由于 Φ 是 C^1 同胚，所以映射

$$Q \to \mathbb{R}^{n^2}, \quad x \mapsto \mathrm{Jac}(\Phi)(x)$$

是连续的。根据连续映射在紧集上的一致连续性，对任意给定的 $\varepsilon > 0$，我们将 Q 分解为有限个足够小的闭的正方体 Q_i 的并集，其中 $i \leqslant N$，我们要求 Q_i 的内部两两不交并且并且对任意的 i 和 $x, x' \in Q_i$，我们都有

$$\| \left(\mathrm{Jac}_\Phi(x')\right)^{-1} \cdot \mathrm{Jac}_\Phi(x)\| < (1+\varepsilon)^{\frac{1}{n}}.$$



现在选定一个小正方体 Q_i 以及一个点 $q_0 \in Q_i$，我们对映射 $(\mathrm{Jac}_\Phi(q_0))^{-1} \circ \Phi$ 应用第一步的结论：

$$\begin{aligned}
m\left((\mathrm{Jac}_\Phi(q_0))^{-1}(\Phi(Q_i))\right) &= m\left((\mathrm{Jac}_\Phi(q_0))^{-1} \circ \Phi(Q_i)\right) \\
&\leqslant \left(\sup_{x \in Q_i} \|\left((\mathrm{Jac}_\Phi(q_0))^{-1} \circ \mathrm{Jac}_\Phi(x)\right\|\right)^n m(Q_i) \\
&\leqslant (1+\varepsilon) m(Q_i).
\end{aligned}$$

利用仿射变换的换元积分公式，我们有

$$\begin{aligned} m(\Phi(Q_i)) &= m\big((\mathrm{Jac}_\Phi(q_0)) \circ (\mathrm{Jac}_\Phi(q_0))^{-1} \circ \Phi \big)(Q_i)\big) \\ &= |\mathbf{J}_\Phi(q_0)| m\big((\mathrm{Jac}_\Phi(q_0))^{-1} \circ \Phi \big)(Q_i)\big) \\ &\leqslant (1+\varepsilon)|\mathbf{J}_\Phi(q_0)| m(Q_i). \end{aligned}$$

现在允许 q_0 变化，对上式两边在 Q_i 上积分，我们得到

$$m(\Phi(Q_i))m(Q_i) \leqslant (1+\varepsilon)m(Q_i)\int_{Q_i} |\mathbf{J}_\Phi(x)|dx,$$

从而（约掉共同的因子）

$$m(\Phi(Q_i)) \leqslant (1+\varepsilon)\int_{Q_i} |\mathbf{J}_\Phi(x)|dx.$$

另外，这些小正方体 Q_i 内部两两不相交，从而 $\Phi(Q_i)$ 内部两两不相交，并且对任意的 i 和 j，$Q_i \cap Q_j$ 和 $\Phi(Q_i \cap Q_j)$ 都是零测集（零测集在微分同胚下的像还是零测集）。下面我们对 Q_i 的指标求和来得到 Q 上的积分：

$$\begin{aligned} m(Q) &= \sum_i m(\Phi(Q_i)) \\ &\leqslant \sum_i (1+\varepsilon)\int_{Q_i} |\mathbf{J}_\Phi(x)|dx \\ &= (1+\varepsilon)\int_{\bigcup_i Q_i} |\mathbf{J}_\Phi(x)|dx \\ &= (1+\varepsilon)\int_{Q} |\mathbf{J}_\Phi(x)|dx. \end{aligned}$$

令 $\varepsilon \to 0$，这完成了第二步的证明。

第三步，$U \subset \Omega_1$ 是开集，我们有不等式

$$m(\Phi(U)) \leqslant \int_U |\mathbf{J}_\Phi(x)|dx.$$

我们已经证明过每个开集 U 都可以表示成可数个正方体块 Q_i 的并集 $U = \bigcup\limits_{i=1}^{\infty} Q_i$ 并且这些 Q_i 的内部两两不交（利用 2^{-k} 大小的网格来实现）。根据第二步的结论，我们有

$$\begin{aligned} m(\Phi(U)) &= \lim_{N\to\infty} m(\bigcup_{i\leqslant N} \Phi(Q_i)) \\ &\leqslant \lim_{N\to\infty} \sum_{i\leqslant N} m(\Phi(Q_i)) \\ &\leqslant \sum_{i=1}^{\infty} \int_{Q_i} |\mathbf{J}_\Phi(x)|dx \\ &= \int_U |\mathbf{J}_\Phi(x)|dx. \end{aligned}$$

第四步，$B \subset \Omega_1$ 是 Borel 集，我们有不等式

$$m(\Phi(B)) \leqslant \int_B |\mathbf{J}_\Phi(x)|dx.$$

我们假设 $\int_B |\mathbf{J}_\Phi(x)|dx < \infty$（否则没有什么可以证明的）。考虑测度带有密度的测度

$$\mu = |\mathbf{J}_\Phi(x)|dx.$$

每个紧集在这个测度下是有限值（因为 $|\mathbf{J}_\Phi(x)|$ 在紧集上有界）。根据我们刚证明的正则性定理，存在开集 U，$B \subset U \subset \Omega_1$，使得

$$\mu(U) \leqslant (1+\varepsilon)\mu(B).$$

也就是说，

$$\int_U |\mathbf{J}_\Phi(x)|dx \leqslant (1+\varepsilon)\int_B |\mathbf{J}_\Phi(x)|dx.$$

从而，

$$\begin{aligned} m(\Phi(B)) \leqslant m(\Phi(U)) &\leqslant \int_U |\mathbf{J}_\Phi(x)|dx \\ &\leqslant (1+\varepsilon)\int_B |\mathbf{J}_\Phi(x)|dx. \end{aligned}$$

其中，倒数第二个不等号我们用了第三步的结论。令 $\varepsilon \to 0$，第四步的结论成立。

第五步，f 是 Ω_1 上定义的正可测函数，那么，我们有

$$\int_{\Omega_2} f(y)dy = \int_{\Omega_1} \big(f\circ\Phi\big)(x)|\mathbf{J}_\Phi(x)|dx.$$

我们首先来证明不等式：

$$\int_{\Omega_2} f(y)dy \leqslant \int_{\Omega_1} \big(f\circ\Phi\big)(x)|\mathbf{J}_\Phi(x)|dx.$$

根据第四步，上面的不等式对示性函数 $f = \mathbf{1}_B$ 成立，其中 B 是 Borel-集。所以，根据积分的线性，上述不等式对正的简单函数也成立。另外，我们可以去单调上升的正简单函数序列 $\{f_i\}_{i\geqslant 1}$ 使得该函数列逐点地收敛到 f，从而，利用 Beppo Levi 定理，我们就有

$$\begin{aligned} \int_{\Omega_2} f(y)dy &= \lim_{i\to\infty}\int_{\Omega_2} f_i(y)dy \\ &\leqslant \lim_{i\to\infty}\int_{\Omega_1} \big(f_i\circ\Phi\big)(x)|\mathbf{J}_\Phi(x)|dx \\ &= \int_{\Omega_1} \big(f\circ\Phi\big)(x)|\mathbf{J}_\Phi(x)|dx. \end{aligned}$$

我们现在说明上面的不等号实际上是等号。此时，要用到 Φ 有逆：对 $\Psi=\Phi^{-1}$ 同样成立上述不等式。所以，

$$\begin{aligned}\int_{\Omega_1}\big(f\circ\Phi\big)(x)|\mathbf{J}_\Phi(x)|dx &\leqslant \int_{\Omega_2}\big(f\circ\Phi\circ\Psi\big)(y)|\mathbf{J}_\Phi(\Psi(y))||\mathbf{J}_\Psi(y)|dy\\ &=\int_{\Omega_2}f(y)dy.\end{aligned}$$

最后一步，我们用到了 $|\mathbf{J}_\Phi(\Psi(y))||\mathbf{J}_\Psi(y)|=1$，根据链式法则，这是明显的。

第六步，对于一般可积函数换元积分公式也成立。我们只需要把函数拆为正负和实部虚部的和，利用线性即可。这就完成了定理的证明。 □

注记. *在 1 维 Riemann 积分情形下，换元积分公式的表达有所不同。假设 $\varphi:[a,b]\to[c,d]$ 是微分同胚（严格单调的 C^1 函数），那么我们有*

$$\int_{\varphi(a)}^{\varphi(b)}f(y)dy=\int_a^b f(\varphi(x))\varphi'(x)dx.$$

我们注意到，φ 的 Jacobi 行列式是没有加绝对值符号的。这当然和积分的区域相关，因为我们要求了

$$\int_{\varphi(a)}^{\varphi(b)}f(y)dy=-\int_{\varphi(b)}^{\varphi(a)}f(y)dy.$$

这和我们刚证明的换元积分公式是一致的。

46 常用的换元积分，行列式的几何含义，子流形上的测度与积分：第一型曲线／曲面积分，函数图像上的积分公式

二零二零年四月十六日，星期四，晴

上次课，我们证明了换元积分公式：Ω_1 和 Ω_2 是 $\mathbb{R}^n$ 中开集，我们用 dx 和 dy 分别表示开集 Ω_1 和 Ω_2 上的 Lebesgue 测度。映射

$$\Phi : \Omega_1 \to \Omega_2.$$

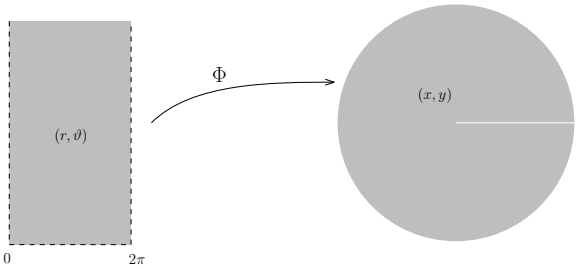
那么，对任意的 Ω_2 上对 dy 可积的函数 f，我们有

$$\int_{\Omega_2} f(y)dy = \int_X \big(f \circ \Phi\big)(x)|\mathbf{J}_\Phi(x)|dx.$$

例子. 我们要熟记如下两个最常用的换元积分公式，极坐标换元和球坐标换元：

1) 我们现在考虑 $\Omega_1 = (0,\infty)\times(-\pi,\pi)$ 和 $\Omega_2 = \mathbb{R}^2 - \{(x,0)|x \geqslant 0\}$。注意到，$\Omega_2$ 和全空间 $\mathbb{R}^2$ 只相差一个零测集，所以在 $\mathbb{R}^2$ 上的积分可以在 Ω_2 上来就算。我们考虑极坐标变换：

$$\Phi : \Omega_1 = (0,\infty)\times(-\pi,\pi) \to \Omega_2, \quad (r,\theta) \mapsto (r\cos(\theta), r\sin(\theta)).$$



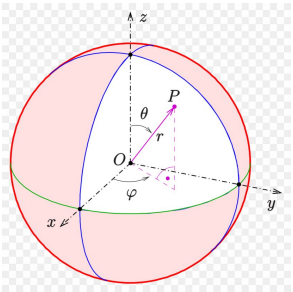
我们之前已经在作业中计算过这个坐标变换的行列式，这给出

$$\iint_{\mathbb{R}^2} f(x,y)dxdy = \iint_{\Omega_1} f(r\cos(\vartheta), r\sin(\vartheta))rdrd\vartheta.$$

我们可以看到，Ω_1 具有很好的乘积结构，所以我们可能可以利用 *Fubini* 公式来计算某些积分。

2) 令 $\Omega_1 = (0,\infty)\times(-\pi,\pi)\times(-\frac{\pi}{2},\frac{\pi}{2})$，$\Omega_2 = \mathbb{R}^3 - \big\{(x,y,z)\big|x=0, y\geqslant 0, z\in\mathbb{R}\big\}$。我们考虑球坐标变换：

$$\Phi : \Omega_1 \to \Omega_2, \quad (r,\theta,\varphi) \mapsto (r\cos(\varphi)\cos(\vartheta), r\cos(\varphi)\sin(\vartheta), r\sin(\varphi)).$$



我们有

$$\begin{aligned}&\iiint_{\mathbb{R}^3} f(x,y,z)dxdydz\\ =&\iiint_{\Omega_1} f(r\cos(\varphi)\cos(\theta), r\cos(\varphi)\sin(\theta), r\sin(\varphi))r^2\cos(\varphi)drd\theta.\end{aligned}$$

我们给出换元积分公式的两个经典应用:

例子 (Gauss 积分的计算). 我们要重新计算 $I=\int_{\mathbb{R}} e^{-x^2}dx$。为此，根据 *Fubini* 公式，我们考虑

$$\begin{aligned}I^2&=\left(\int_{\mathbb{R}} e^{-x^2}dx\right)\left(\int_{\mathbb{R}} e^{-y^2}dy\right)\\&=\iint_{\mathbb{R}^2} e^{-(x^2+y^2)}dxdy\\&=\int_0^{\infty}\int_0^{2\pi} e^{-r^2}rdrd\vartheta\\&=2\pi\int_0^{\infty} e^{-r^2}rdr=2\pi.\end{aligned}$$

所以，$I=\sqrt{2\pi}$。

例子 (行列式的几何含义). 很多的线性代数课本上来就声明行列式表示的空间中平行多面体的体积，这实际上是扑风捉影的叙述。只有定义了什么是体积，才能作出这样的论断，我们现在给出这个说法的解释。给定线性映射（我们直接用矩阵来表示）

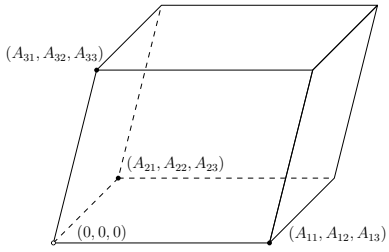
$$A:\mathbb{R}^n\to\mathbb{R}^n,$$

首先考虑由顶点在原点，以 $e_1,e_2,\cdots,e_n$ 作为边的平行多面体（正方体）Ω，它的测度为 1。另外，$A(\Omega)$ 是顶点在原点，以 $A(e_1),A(e_2),\cdots,A(e_n)$ 作为边的平行多面体，根据换元积分公式，它的体积为

$$\begin{aligned}m\big(A(\Omega)\big)&=\int_{\Omega}1dx=\int_{A(\Omega)}1|\mathbf{J}_{A^{-1}}|dy\\&=|\det A|^{-1}\int_{A(\Omega)}1dy=|\det A|^{-1}m\big(A(\Omega)\big).\end{aligned}$$

所以，

$$m\big(A(\Omega)\big)=|\det A|m\big(A(\Omega)\big).$$



另外，我们有

$$\begin{cases} A(e_1) & = (A_{11}, A_{12}, \cdots, A_{1n}), \\ A(e_2) & = (A_{21}, A_{22}, \cdots, A_{2n}), \\ \quad \cdots & \quad \cdots \quad \cdots \\ A(e_n) & = (A_{n1}, A_{n2}, \cdots, A_{nn}). \end{cases}$$

这就是我们平时所说的以 $(A_{i1}, A_{i2}, \cdots, A_{in})$ 为顶点的平行多面体的体积，其中 $i \leqslant n$。

注记. 我们所寻求的高维的积分理论一定要具备两个基本的工具：第一，*Fubini* 定理，它可以把高维积分（乘积空间上）转化为低维的来计算；第二，换元积分公式，它可以把一个不规则区域上的积分转化为乘积型区域上的积分，从而用 *Fubini* 来计算。我们现在建立的积分理论满足这两点，传统的 *Riemann* 积分理论也要建立这两个公式，所以，从计算积分的观点来看，我们的所得到的理论和 *Riemann* 积分实际上差别不大。

子流形上的积分（第一型曲线／曲面积分）

我们从测度论的观点来研究传统的曲面积分，也就是在子流形上进行积分。我们从最简单的子流形开始研究这个问题（即 $\mathbb{R}^n$ 中的一个线性子空间或者仿射子空间）。我们固定一个背景空间 $\mathbb{R}^n$，在 $\mathbb{R}^n$ 上我们将使用 Lebesgue 测度 $m = dy_1 \cdots dy_n$。我们用 E 表示 $\mathbb{R}^n$ 的一个 d-维仿射子空间 $E \subset \mathbb{R}^n$。最基本的问题是：如何测量 E 上集合的体积（面积），即如何在 E 上定义测度？对于这个测度，我们还需要它和**直观**是一致的（和 Euclid 几何的直观相符）。为此，我们研究一个更简单明了的例子：

例子. 考虑映射

$$\Phi : \mathbb{R} \to \mathbb{R}^2, \quad x_1 \mapsto (x_1, x_1).$$

它的像就是 $\mathbb{R}^2$ 中的直线 45° 的直线 E。

我们强调一点，$E \subset \mathbb{R}^2$ 是我们关心的几何对象，它还可以通过另外的方式来参数话，比如

$$\Psi : \mathbb{R} \to \mathbb{R}^2, \quad x_2 \mapsto (2x_2, 2x_2).$$

这个映射的像也是 E。

我们考虑 E 上的 $(0,0)$ 点和 $(1,1)$ 点之间的一段线段 L，我们**直观**上希望它的长度是 $\sqrt{2}$。

然而，根据参数化 Φ 或者 Ψ，我们有若干中不同的方式给出 E 上的测度：

1) 我们用 $\Phi_* m$ 作为 E 上的测度，即我们定义 L 的测度为

$$m\left(\Phi^{-1}(L)\right) = m([0,1]) = 1.$$

此时，用 x_1 表示参数，L 上的测度 $\Phi_* m_1$ 为 dx_1。我们希望这条直线上的测度是 $m_1 = \sqrt{2} dx_1$。

2) 我们用 $\Psi_* m_1$ 作为 E 上的测度，即我们定义 L 的测度为

$$m\left(\Phi^{-1}(L)\right) = m([0, \frac{1}{2}]) = \frac{1}{2}.$$

此时，用 x_2 表示参数，L 上的测度 $\Phi_* m$ 为 $\frac{1}{2}dx_2$。我们希望这条直线上的测度是 $m_2 = 2\sqrt{2}dx_1$。

我们希望说明 m_1 或者 m_2 是我们在不同的参数化下同一个测度的不同表示而已（它们都是相对于参数化的一个带有密度的测度)!

仿射子空间 E 上的测度，我们也有类似的困境。我们取仿射变换（一次函数）来参数化我们的仿射子流形:

$$\Phi : \mathbb{R}^d \to E \subset \mathbb{R}^n, \quad x = (x_1, \cdots, x_d) \mapsto J \cdot x + y_0$$

其中，$y_0 \in \mathbb{R}^n$，J 是 $n \times d$ 的矩阵并且 $\mathrm{rank}(J) = d$（J 为映射 Φ 的 Jacobi 矩阵)。我们强调在 $\mathbb{R}^d$ 和 $\mathbb{R}^n$ 上我们选取了坐标 $\{x_1, \cdots, x_d\}$ 和 $\{y_1, \cdots, x_n\}$ 并且假定了这两个空间上有和这两个坐标系相对应的标准的内积结构（选定了标准正交基)。

为了定义 E 上的测度，最干脆（但是错误）的做法是用 $\mathbb{R}^d$ 上的 Lebesgue 测度在 Φ 下的像，即 $\Phi_* m_p$，作为 E 上的测度，前面的例子表明，这个不符合直观。我们可以用 J 自身来修正这一点:

定义 310 (Gram 矩阵)**.** 我们定义一个 $d \times d$ 的矩阵

$$G = {}^t J \cdot J.$$

我们把它称作是参数化 Φ 的 **Gram 矩阵**。我们定义 E 上的测度 m_E 为:

$$m_E = |\det(G)|^{\frac{1}{2}} \cdot \Phi_* m_p.$$

换而言之，对任意的 $A \subset \mathcal{B}(E)$，我们有

$$m_E(A) = |\det(G)|^{\frac{1}{2}} \cdot m_p\big(\Phi^{-1}(A)\big).$$

我们用之前例子中的映射

$$\Phi : \mathbb{R} \to \mathbb{R}^2, \quad x_1 \mapsto (x_1, x_1).$$

此时，$J = (1, 1)$，$G = 2$，所以，相应的测度为 $\sqrt{2}dx_1$。如果我们换成

$$\Psi : \mathbb{R} \to \mathbb{R}^2, \quad x_2 \mapsto (2x_2, x_2).$$

此时，$J = (2, 2)$，$G = 8$，所以，相应的测度为 $\sqrt{8}dx_2 = 2\sqrt{2}dx_2$。

按照我们在例子中的讨论，这两个测度实际上是同一个，也恰好是我们直观上想要找的测度。

引理 311. E 上的测度 m_E 不依赖于仿射线性化 Φ 的选取。

证明：我们给两个仿射参数化（忘掉平移部分）：

$$\Phi : \mathbb{R}^d \to E \subset \mathbb{R}^n, \quad x \mapsto J_\Phi \cdot x + y_1$$
$$\Psi : \mathbb{R}^d \to E \subset \mathbb{R}^n, \quad z \mapsto J_\Psi \cdot z + y_2.$$

此时，存在一个同构的仿射变换

$$L : \mathbb{R}^d \to \mathbb{R}^d, \quad z \mapsto J_L \cdot z + x_1,$$

使得 $\Psi = \Phi \circ L$，其中 J_L 是 $d \times d$ 的可逆矩阵。我们自然有（比如求微分）

$$J_\Psi = J_\Phi \cdot J_L.$$

所以，

$$\begin{aligned}
\det\big(G_\Psi\big) &= \det\Big({}^t\big(J_\Psi\big) \cdot J_\Psi\Big) \\
&= \det\Big({}^t\big(J_\Phi \cdot J_L\big) \cdot J_\Phi \cdot J_L\Big) \\
&= \big|\det J_L\big|^2 \det\big(G_\Phi\big).
\end{aligned}$$

从而，对于任意的 $A \in E$，根据在 $\mathbb{R}^d$ 上的换元积分公式，我们有

$$\begin{aligned}
|\det(G_\Psi)|^{\frac{1}{2}} \Psi_* m_p(A) &= |\det(G_\Phi)|^{\frac{1}{2}} \big|\det J_L\big| m_p\Big(L^{-1}\big(\Phi^{-1}(A)\big)\Big) \\
&= |\det(G_\Phi)|^{\frac{1}{2}} \big|\det J_L\big| \big|\det J_{L^{-1}} m_p\big(\Phi^{-1}(A)\big)\big) \\
&= |\det(G_\Phi)|^{\frac{1}{2}} m_p\big(\Phi^{-1}(A)\big) \\
&= |\det(G)|^{\frac{1}{2}} \Phi_* m_p(A).
\end{aligned}$$

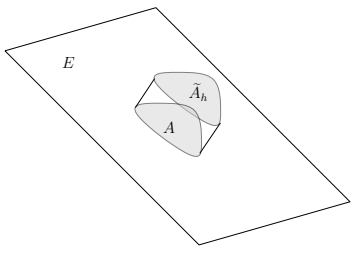
这说明 m_E 的定义不依赖于仿射参数化的选取。 □

我们给出两个例子来说明我们的构造和我们的几何直观是相符的：

例子. 这两个例子都是关于超平面的（余维数是 1）：

1) 假定 $E \subset \mathbb{R}^n$，$\dim E = n-1$，令 ν 是 E 上的单位法向量（场）（与 E 在每一点处的切空间垂直），$A \subset E$ 是 *Borel* 集。我们定义以 A 为底高为 h 的柱体 $\widetilde{A}_h$ 为：

$$\widetilde{A}_h = \big\{x + t \cdot \nu \big| x \in A, t \in [0, h]\big\}.$$



我们证明

$$m_n(\widetilde{A}) = m_E(A) \times h.$$

根据 *Lebesgue* 测度的旋转不变性以及测度 m_E 是参数不变的，我们可以在 E 的仿射参数上符合一个 $\mathbb{R}^n$ 中的正交变换（不改变 *Gram* 矩阵的行列式），从而不妨假设

$$E = \{y_n = 0\}$$

并且

$$\Phi : \mathbb{R}^{n-1} \to \mathbb{R}^n, \quad (x_1, \cdots, x_{n-1}) \mapsto (x_1, \cdots, x_{n-1}, 0).$$

这时，命题显然成立（乘积测度的定义）。

2) 我们将 E 实现为函数的图像：给定线性函数

$$f : \mathbb{R}^{n-1} \to \mathbb{R},$$

它的图像 $\Gamma_f = \left\{(x, f(x)) \in \mathbb{R}^{n-1} \times \mathbb{R} \middle| x \in \mathbb{R}^{n-1}\right\}$ 为 $\mathbb{R}^n$ 中的超平面。由于 f 是线性的，所以我们有

$$f(x) = f_1 x_1 + \cdots + f_n x_n + a, \quad f_i \in \mathbb{R}, a \in \mathbb{R}, i = 1, \cdots, n.$$

从而，$\nabla f = (f_1, \cdots, f_n)$。子流形 $E = \Gamma_f$ 的（共有两个）单位法向量（场）为

$$\begin{aligned} \nu &= \frac{1}{(1 + |\nabla f|^2)^{\frac{1}{2}}}(-\nabla f, 1) \\ &= \frac{(-f_1, \cdots, -f_{n-1}, 1)}{(1 + |f_1|^2 + \cdots + |f_{n-1}|^2)^{\frac{1}{2}}}. \end{aligned}$$

为了给出 m_E，我们现在找出 E 的参数化：

$$\Phi : \mathbb{R}^{n-1} \to \mathbb{R}^n, \quad x \mapsto \Phi(x) = (x, f(x)).$$

它所对应的 $J = (1, f_1, f_2, \cdots, f_{n-1}) = (1, \nabla f)$，所以

$$\det\Big({}^t J \cdot J_\Phi\Big) = 1 + |\nabla f|^2.$$

从而，相应的测度为

$$m_E = \sqrt{1 + |\nabla f|^2} dx_1 \cdots dx_{n-1}.$$

所以，如果你要在 E 上积分一个函数 $F : E \to \mathbb{C}$，我们先把 F 写成复合的形式：

$$F = F(x, f(x)), \quad x \in \mathbb{R}^{n-1}.$$

从而，

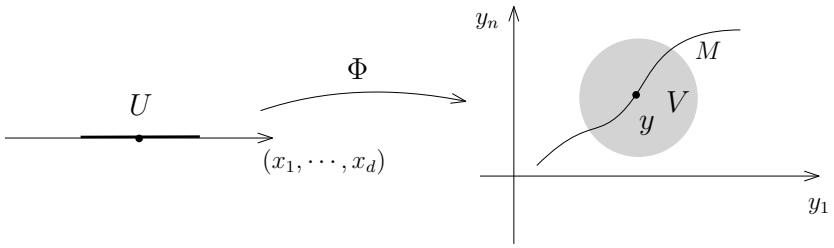
$$\int_E F dm_E = \int_{\mathbb{R}^{n-1}} F(x, f(x)) \sqrt{1 + |\nabla f|^2} dx_1 \cdots dx_{n-1}.$$

这就化成了 $\mathbb{R}^{n-1}$ 上的积分。

上面关于线性子流形的构造很容易推广到一般的子流形上去，基本的想法是局部上把每个子流形看作是参数化的流形，尽管这只是局部的构造，但是局部上能定义测度也就可以整体上定义测度了，所以子流形的积分的理论本质上是局部的问题。给定一个子流形 $M \subset \mathbb{R}^n$，我们假设 $\dim M = d$。对于任意的 $y \in S \subset \mathbb{R}^n$，按照定义，存在开集 $V \subset \mathbb{R}^n$，$y \in V$，开集 $U \subset \mathbb{R}^d$ 和光滑映射 $\Phi : U \to \mathbb{R}^n$，使得

1) $\Phi : U \to V \cap M$ 是微分同胚；

2) 对每个 $x \in U$，$\operatorname{rank}(d\Phi(x)) = d$，从而，$G_\Phi(x) = {}^t\big(\operatorname{Jac}(\Phi)\big) \cdot \operatorname{Jac}(\Phi)$ 是非退化的 $d \times d$ 的矩阵。



（上面的映射方向和我们的子流形的定义中的映射方向是相反的，但这个不妨碍证明）

仿照仿射子流形的情况，对于任意的 Borel 集 $A \subset M \cap V$，我们定义其测度为

$$\begin{aligned}\sigma(A) &= \int_U \mathbf{1}_{\Phi^{-1}(A)}(x)\sqrt{\det\big({}^t\big(\operatorname{Jac}(\Phi)\big) \cdot \operatorname{Jac}(\Phi)\big)}dx_1 \cdots dx_d \\ &= \int_U \mathbf{1}_{\Phi^{-1}(A)}(x)\sqrt{\det\left(G_\Phi(x)\right)}dx_1 \cdots dx_d.\end{aligned}$$

换而言之，我们将 M 上的子流形测度（也叫做曲面测度）定义为

$$d\sigma = \sqrt{\det\big({}^t\big(\operatorname{Jac}(\Phi)\big) \cdot \operatorname{Jac}(\Phi)\big)}\Phi_*(dx_1 \cdots dx_d),$$

其中 $dx_1 \cdots dx_d$ 是 U 上的 Lebesgue 测度。当然，如果 M 是仿射子流形并且 Φ 是仿射变换，这与之前的构造一致。

我们现在说明这个测度是良好定义的：

引理 312. *在 $V \cap M$ 上定义的测度*

$$\sigma = \sqrt{\det\big({}^t\big(\operatorname{Jac}(\Phi)\big) \cdot \operatorname{Jac}(\Phi)\big)} \cdot \Phi_*(m)$$

不依赖于参数化 $\Phi : U \to V$ 的选取。其中，我们在 $V \cap M$ 配有 Borel-代数（它和 U 是微分同胚的）。特别地，这个测度是局部上定义的，从而也在子流形的整体上有了定义。

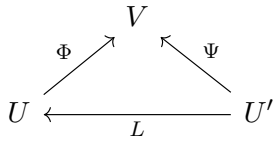
证明：这个证明和仿射情形的证明如出一辙：假设 $\Psi : U' \to V \cap M$ 是另一个参数化，即存在开集 $U \subset \mathbb{R}^d$ 和光滑映射 $\Phi : U \to \mathbb{R}^n$，使得

$$\Psi : U \to V \cap M$$

是微分同胚。所以，如果我们定义 $L = \Phi^{-1} \circ \Psi$，即 $\Psi = \Phi \circ L$，我们就得到一个微分同胚

$$L : U' \to U$$

（C^1 同胚即可）。这可以用如下的交换图表来表示：



利用复合映射的求微分公式，我们就有

$$\mathrm{Jac}(\Psi) = \mathrm{Jac}(\Phi) \circ \mathrm{Jac}(L),$$

所以，

$$\sqrt{\det\big({}^t\big(\mathrm{Jac}(\Psi)\big) \cdot \mathrm{Jac}(\Psi)\big)} = |\det\left(\mathrm{Jac}(L)\right)| \sqrt{\det\big({}^t\big(\mathrm{Jac}(\Phi)\big) \cdot \mathrm{Jac}(\Phi)\big)}.$$

我们在 U 和 U' 上分别用 x_i 和 x_i' 作为坐标，其中 $i = 1, \cdots, d$。那么，根据对 L 的换元积分公式，我们有

$$\begin{aligned}
\sigma'(A) &= \int_{U'} \mathbf{1}_{\Psi^{-1}(A)}(x') \sqrt{\det\big({}^t\big(\mathrm{Jac}(\Psi)\big) \cdot \mathrm{Jac}(\Psi)\big)} dx_1' \cdots dx_d' \\
&= \int_{L^{-1}(U)} \mathbf{1}_{L^{-1}(\Phi^{-1}(A))}(x') \sqrt{\det\big({}^t\big(\mathrm{Jac}(\Phi)\big) \cdot \mathrm{Jac}(\Phi)\big)} |\det\left(\mathrm{Jac}(L)\right)| dx_1' \cdots dx_d' \\
&= \int_{U} \mathbf{1}_{\Phi^{-1}(A)}(x') \sqrt{\det\big({}^t\big(\mathrm{Jac}(\Phi)\big) \cdot \mathrm{Jac}(\Phi)\big)} dx_1 \cdots dx_d \\
&= \sigma(A).
\end{aligned}$$

这表明测度 σ 不依赖于参数化的选取。 □

注记. 如果 f 是 $\mathbb{R}^n$ 上定义的函数并且 f 在 M 上的限制是可积的（对测度 σ 而言），那么

$$\int_S f d\sigma = \int_U f\big(\Phi(x)\big) \sqrt{\det\big({}^t\big(\mathrm{Jac}(\Phi)\big) \cdot \mathrm{Jac}(\Phi)\big)} dx_1 \cdots dx_d.$$

这将是非常有用的公式，因为它把曲面积分转化为通常的在 $\mathbb{R}^d$ 上的一个区域上的积分，定要熟练记忆和运用。

例子. 我们考虑 *1* 维的子流形对它进行参数化，这就是参数化的曲线

$$\gamma : (-1, 1) \to \mathbb{R}^n,$$

其中对任意的 $t\in(-1,1)$，$\gamma'(t)\neq 0$。我们固定 $-1<a<b<1$，我们来研究曲线上 $\gamma(a)$ 和 $\gamma(b)$ 之间的曲线段 C_a^b 的测度／长度。首先，$\mathrm{Jac}(\gamma)(t)=\gamma'(t)$，所以，按照定义：

$$\begin{aligned}\sigma(C_a^b)&=\int_{(-1,1)}\mathbf{1}_{\gamma^{-1}(S_a^b)}\sqrt{\gamma'(t)\cdot\gamma'(t)}dt\\&=\int_a^b|\gamma'(t)|dt.\end{aligned}$$

这和上个学期所定义的参数曲线的长度的定义一致。特别地，我们用 C 表示该曲线（即 γ 的像），$f:C\to\mathbb{C}$ 是给定的（可积）函数，那么

$$\int_{S_a^b}fd\sigma=\int_a^b f(\gamma(t))|\gamma'(t)|dt.$$

这将是非常有用的公式，定要熟练记忆和运用。

下面所展示的是另一个核心的例子：

例子. 给定光滑函数

$$f:\mathbb{R}^{n-1}\to\mathbb{R},$$

我们考虑它的图像所定义的 $\mathbb{R}^n$ 中的超曲面 $\Gamma_f=\big\{(x,f(x))\big|x\in\mathbb{R}^{n-1}\big\}$。该曲面在 $(x,f(x))$ 法向量为

$$\nu=\frac{1}{(1+|\nabla f|^2)^{\frac{1}{2}}}(-\nabla f,1).$$

实际上，很容易看出，$T_{(x,f(x))}$ 由如下 $n-1$ 个向量

$$v_i=(\underbrace{0,\cdots,0,1,0,\cdots,0}_{\text{共 }n-1\text{ 个，第 }i\text{ 个位置为 }1},\frac{\partial f}{\partial x_i}),\quad i=1,2,\cdots,n-1,$$

所张成，直接计算就得到 $v_i\cdot\nu=0$，其中 $i\leqslant n-1$。

我们用如下的映射来参数化 Γ_f：

$$\Phi:\mathbb{R}^{n-1}\to\mathbb{R}^n,\quad x\mapsto(x,f(x)).$$

所以，我们有

$${}^t\mathrm{Jac}(\Phi)=\begin{pmatrix}1&0&\cdots&0&\frac{\partial f}{\partial x_1}\\0&1&\cdots&0&\frac{\partial f}{\partial x_2}\\0&\cdots&\cdots&0&\cdots\\0&0&\cdots&1&\frac{\partial f}{\partial x_{n-1}}\end{pmatrix}$$

这是一个 $(n-1)\times n$ 的矩阵，如果用分块矩阵的写法，我们有 ${}^t\mathrm{Jac}(\Phi)=(\mathbf{I},\nabla f)$，其中 $\mathbf{I}$ 是 $(n-1)\times(n-1)$ 的单位矩阵，∇f 为列向量。所以，

$${}^t\mathrm{Jac}(\Phi)\cdot\mathrm{Jac}(\Phi)=\mathbf{I}+\nabla f\otimes\nabla f=\begin{pmatrix}1+\left(\frac{\partial f}{\partial x_1}\right)^2&\frac{\partial f}{\partial x_1}\frac{\partial f}{\partial x_2}&\cdots&\frac{\partial f}{\partial x_1}\frac{\partial f}{\partial x_{n-1}}\\\frac{\partial f}{\partial x_2}\frac{\partial f}{\partial x_1}&1+\left(\frac{\partial f}{\partial x_2}\right)^2&\cdots&\frac{\partial f}{\partial x_1}\frac{\partial f}{\partial x_{n-1}}\\\cdots&\cdots&\cdots&\cdots\\\frac{\partial f}{\partial x_n}\frac{\partial f}{\partial x_1}&\frac{\partial f}{\partial x_n}\frac{\partial f}{\partial x_2}&\cdots&1+\left(\frac{\partial f}{\partial x_{n-1}}\right)^2\end{pmatrix}$$

所以，（这是一个很有意义的行列式的计算，我们把它留作作业）

$$^t\mathrm{Jac}(\Phi)\cdot\mathrm{Jac}(\Phi)=1+|\nabla f|^2.$$

所以，我们所求的曲面测度为

$$\sigma=\sqrt{1+|\nabla f|^2}dx_1\cdots dx_{n-1}.$$

那么，对于任意的 $\mathbb{R}^n$ 上定义的函数，我们有（只要下面的等式有意义）

$$\int_{\Gamma_f}\varphi d\sigma=\int_{\mathbb{R}^{n-1}}\varphi\big(x,f(x)\big)\sqrt{1+|\nabla f(x)|^2}dx.$$

这将是非常有用的公式，必须要熟练记忆和运用。

46.1 期中考试：非 Borel 集的构造

清华大学 2020 年春季学期数学分析二 期中考试

考试说明

本次考试为开卷宅考（take-home exam），请将解答在 A4 大小的白纸上誊写并扫描成 pdf 格式，尽量控制 pdf 文件的大小。务必于 4 月 20 日上午 9:50 之前将解答上传至网络学堂，否则考试按零分处理。

考试中后面的问题可以使用前面问题的结论（无论答题人是否已经得到正确的证明或者答案），试题出现的先后顺序与其难度毫无关联。

在本次考试中所出现的测度均为 Lebesgue 测度。

Student: Dr. Einstein, Aren't these the same questions as last year's final exam?

Dr. Einstein: Yes! But this year the answers are different.

光滑性（共 10 分）

我们用 $\mathbf{Sym}_n(\mathbb{R})$ 表示 $n\times n$ 的实对称矩阵的全体，这是一个 $\dfrac{n(n+1)}{2}$-维的 $\mathbb{R}$-线性空间。令 $\mathbf{Sym}_n^+(\mathbb{R})\subset\mathbf{Sym}_n(\mathbb{R})$ 为全体正定（对称）的矩阵。

A1)（3 分）证明，$\mathbf{Sym}_n^+(\mathbb{R})$ 是 $\mathbf{Sym}_n(\mathbb{R})$ 中的开集。

A2)（7 分）根据线性代数一个常用的命题（请自行查阅资料），对任意的 $A\in\mathbf{Sym}_n^+(\mathbb{R})$，存在唯一的 $B\in\mathbf{Sym}_n^+(\mathbb{R})$，使得

$$B^2=A.$$

我们用 $\sqrt{A}$ 来表示 B，这就我们得到了映射

$$\sqrt{\cdot}:\mathbf{Sym}_n^+(\mathbb{R})\to\mathbf{Sym}_n^+(\mathbb{R}),\quad A\mapsto\sqrt{A}.$$

证明，这个映射是 C^∞ 的。

导数的计算（共 10 分）

B1)（5 分）假设 $f\in C^2(\mathbb{R}^2,\mathbb{R})$，其中，我们用 (x,y) 表示 $\mathbb{R}^2$ 上的坐标。令

$$\begin{cases} x=au+bv, & a,b\in\mathbb{R};\\ y=cu+dv, & c,d\in\mathbb{R}.\end{cases}$$

据此，我们可以把 f 看作是 (u,v) 的函数，即 $F(u,v)=f(au+bv,cu+dv)$。试用 f 对 x 和 y 的偏导数来计算 $\dfrac{\partial^2F}{\partial u\partial v}$。

B2)（5 分）试找出所有的 $f \in C^2(\mathbb{R}^2, \mathbb{R})$，使得

$$\frac{\partial^2 f}{\partial x^2} - 3\frac{\partial^2 f}{\partial x \partial y} + 2\frac{\partial^2 f}{\partial y^2} = 0.$$

反／隐函数定理：三次多项式的实根（共 15 分）

定义 $\mathbb{R}^3$ 上的映射：

$$P : \mathbb{R}^2 \times \mathbb{R}, \quad (a, b, x) \mapsto P(a, b, x) = x^3 + bx^2 + ax + 1.$$

为了方便起见，我们还记 $p_{(a,b)}(x) = P(a, b, x)$。

C1)（2 分）我们定义 $\mathbb{R}^2$ 集合

$$\Omega = \big\{(a, b)\big| b^2 < 3a\big\}.$$

证明，这是一个开集并且对任意的 $(a, b) \in \Omega$，多项式 $p_{(a,b)}(x)$ 有实根。

C2)（4 分）对任意的 $(a, b) \in \Omega$，我们假设 $x \in \mathbb{R}$，使得 $p_{(a,b)}(x) = 0$。证明，存在 (a, b) 在 Ω 中的开邻域 $U_{(a,b)}$，x 在 $\mathbb{R}$ 中的开邻域 V_x 以及光滑映射

$$\varphi : U_{(a,b)} \to V_x,$$

使得对任意的 $(a', b', x') \in U_{(a,b)} \times V_x$，$p_{(a',b')}(x') = 0$ 等价于 $x = \varphi(a', b')$。

C3)（2 分）假设 $(a, b) \in \mathbb{R}^2$ 使得多项式 $p_{(a,b)}(x)$ 有三个不同的实根。证明，存在 (a, b) 在 $\mathbb{R}^2$ 中的开邻域 $U_{(a,b)}$，使得对任意的 $(a', b') \in U_{(a,b)}$，多项式 $p_{(a',b')}(x)$ 也有三个不同的实根。

C4)（3 分）证明，对任意的 $(a, b) \in \Omega$，$p_{(a,b)}(x)$ 恰有一个实根。

C5)（4 分）我们定义

$$\mathcal{U}_1 = \big\{(a, b) \in \mathbb{R}^2\big| p_{(a,b)}(x) \text{ 恰有一个实根}\big\},$$
$$\mathcal{U}_3 = \big\{(a, b) \in \mathbb{R}^2\big| p_{(a,b)}(x) \text{ 有三个不同的实根}\big\}.$$

证明，$\mathcal{U}_1 \cup \mathcal{U}_3$ 是 $\mathbb{R}^2$ 上稠密的集合。

凸函数（共 10 分）

假设 $\Omega \subset \mathbb{R}^n$ 是非空的凸集，$f : \Omega \to \mathbb{R}$ 是凸函数。

D1)（3 分）在这个问题中，我们加一些假设：

- Ω 是紧集并且它的内部 $\mathring{\Omega}$ 非空；[12]
- f 在 Ω 上是连续函数，在开集 $\mathring{\Omega}$ 上是 C^1 的。

如果存在 $\omega \in \mathring{\Omega}$，使得

$$f(\omega) = \sup_{x \in \Omega} f(x).$$

那么，f 是常值函数。（提示：用凸函数的几何解释）

D2)（3 分）假设 $\Omega = \{(x, y) | x^2 + y^2 \leqslant 1\}$，令

$$f(x, y) = \begin{cases} n, & \text{如果 } (x, y) = \left(\cos\left(\frac{1}{n}\right), \sin\left(\frac{1}{n}\right)\right),\ n \in \mathbb{Z}_{n \geqslant 1}; \\ 0, & \text{其余情况} \end{cases}$$

证明，这是一个凸函数并且它在 Ω 上的上确界是 $+\infty$。

D3)（4 分）我们只假设 Ω 是非空凸集，f 是凸函数。如果假设存在 $\omega \in \mathring{\Omega}$，使得

$$f(\omega) = \sup_{x \in \Omega} f(x).$$

证明，对任意的 $x \in \Omega$，我们都有 $f(\omega) = f(x)$。（提示：先证明存在 $x' \in \Omega$，使得 ω 落在 x 与 x' 的所联线段的内部）

求函数极值（共 10 分）

我们在 $\mathbb{R}_{\geqslant 0} \times \mathbb{R}_{\geqslant 0} \times \mathbb{R}_{\geqslant 0}$ 上定义函数

$$f(x, y, z) = \begin{cases} \frac{xyz}{(x+y+z)^2}, & (x, y, z) \neq (0, 0, 0); \\ 0, & (x, y, z) = (0, 0, 0). \end{cases}$$

E1)（4 分）证明，f 是 $\mathbb{R}_{\geqslant 0} \times \mathbb{R}_{\geqslant 0} \times \mathbb{R}_{\geqslant 0}$ 上的连续函数，它在开集 $\mathbb{R}_{>0} \times \mathbb{R}_{>0} \times \mathbb{R}_{>0}$ 上是连续可微的。

E2)（6 分）假设 $a > 0$。证明，f 在集合

$$K = \left\{(x, y, z) \in \mathbb{R}_{\geqslant 0} \times \mathbb{R}_{\geqslant 0} \times \mathbb{R}_{\geqslant 0} \middle| x^2 + y^2 + z^2 = a^2\right\}$$

上能取到最大值并计算这个值。

[12] Ω 的内部是 Ω 的所有内点构成的开集；对于 $x \in \Omega$，如果存在 $\varepsilon > 0$，使得 $B(x, \varepsilon) \subset \Omega$（以 x 为球心，以 ε 为半径的小球），我们就称它是 Ω 的内点。同学们还可以参考上学期第 29 次讲义。

子流形的判定：Hopf 纤维化（共 10 分）

我们用 $\mathbf{S}^3$ 和 $\mathbf{S}^2$ 表示 $\mathbb{R}^4$ 和 $\mathbb{R}^3$ 中的单位球面

$$\mathbf{S}^3 = \{(x,y,z,w)|x^2+y^2+z^2+w^2=1\},\ \ \mathbf{S}^2 = \{(a,b,c)|a^2+b^2+c^2=1\}.$$

F1)（2 分）证明，$\mathbf{S}^3$ 和 $\mathbf{S}^2$ 分别为 $\mathbb{R}^4$ 和 $\mathbb{R}^3$ 中光滑子流形。

F2)（4 分）证明，映射

$$H:\mathbb{R}^4\to\mathbb{R}^2,\ \ (x,y,z,w)\mapsto\big(x^2+y^2-z^2-w^2, 2(xw+yz), 2(yw-xz)\big)$$

将 $\mathbf{S}^3$ 映射到 $\mathbf{S}^2$ 并且对任意的 $p\in\mathbf{S}^2$，它的逆像 $H^{-1}(p)$ 是 $\mathbf{S}^3$ 中的子流形。

F3)（4 分）证明，对任意的 $p\in\mathbf{S}^2$，$H^{-1}(p)$ 与单位圆周微分同胚。

一个非 Borel 集的构造（共 10 分）

G1)（2 分）在 $[0,1]$ 上定义如下的等价关系：$x\sim y$ 当且仅当 $x-y\in\mathbb{Q}$。试验证这是一个等价关系。这样，$[0,1]$ 被分成了若干个（两两不交）的等价类，我们在每个等价类中任意选定一个元素，这些元素组成了 $[0,1]$ 的子集 A。对于任意的 $q\in[-1,1]\cap\mathbb{Q}$，我们定义

$$A_q = A+q = \{q+a|a\in A\}.$$

证明，若 $q_1\neq q_2$，那么 $A_{q_1}\neq A_{q_2}$，并且

$$[0,1]\subset\bigcup_{q\in[-1,1]\cap\mathbb{Q}} A_q\subset[-1,2].$$

G2)（4 分）证明，$A\notin\mathcal{B}(\mathbb{R}^1)$。（提示：利用 Lebesgue 测度）

G3)（4 分）我们用 $\mathcal{P}(\mathbb{R}^1)$（所有子集的集合）作为 $\mathbb{R}^1$ 上的 σ-代数，假设 μ 是 $\mathcal{P}(\mathbb{R}^1)$ 上的一个平移不变的测度。证明，要么 $\mu\equiv0$，要么对任意的非空开区间 I，$\mu(I)=+\infty$。

积分的计算（共 15 分）

对于 $a>0$ 和 $x\geqslant0$，我们定义

$$H_a(x) = \int_0^\infty e^{-(at^2+\frac{x}{t^2})}dt.$$

H1)（3 分）证明，上述积分是良好定义的并计算 $H_a(0)$。

H2)（3 分）证明，函数

$$H_a : [0,\infty) \to \mathbb{R}, \quad x \mapsto H_a(x),$$

是连续函数。

H2)（3 分）证明，函数

$$H_a : (0,\infty) \to \mathbb{R}, \quad x \mapsto H_a(x),$$

是可微的。

H3)（3 分）计算 $H_a'(x)$。

H4)（3 分）证明，

$$H_a(x) = \frac{1}{2}\sqrt{\frac{\pi}{a}} e^{-2\sqrt{ax}}.$$

积分的控制（共 20 分）

给定 $\mathbb{R}^n$ 上一个非空开集 Ω（Borel 集即可）和 Ω 上的复可积函数 f。

I1)（3 分）证明，

$$\lim_{R\to\infty} \int_\Omega |f(x)| \mathbf{1}_{[R,+\infty)}\big(|x|\big) dx_1 \cdots dx_n = 0.$$

I2)（3 分）证明，

$$\lim_{R\to\infty} \int_\Omega |f(x)| \mathbf{1}_{[R,+\infty)}\big(|f(x)|\big) dx_1 \cdots dx_n = 0.$$

I3)（4 分）证明，对任意的 $\varepsilon > 0$，存在 $\delta > 0$，使得对任意的 $A \in \mathcal{B}(\Omega)$，只要 $m(A) < \delta$，我们就有

$$\int_\Omega |f(x)| \mathbf{1}_A(x) dx_1 \cdots dx_n < \varepsilon.$$

在剩下的问题中，我们假设 Ω 的测度是有限的，即 $m(\Omega) < \infty$，$\{f_k(x)\}_{k\geqslant 1}$ 是 Ω 上的一列复可积函数。

I4)（3 分）假设

$$\lim_{R\to\infty} \left(\sup_{k\geqslant 1} \left(\int_\Omega |f_k(x)| \mathbf{1}_{[R,+\infty)}\big(|f_k(x)|\big) dx_1 \cdots dx_n \right) \right) = 0.$$

证明，对任意的 $\varepsilon > 0$，存在 $\delta > 0$，使得对任意的 $A \in \mathcal{B}(\Omega)$，只要 $m(A) < \delta$，我们就有

$$\sup_{k\geqslant 1} \left(\int_\Omega |f_k(x)| \mathbf{1}_A(x) dx_1 \cdots dx_n \right) < \varepsilon.$$

I5)（3 分）对任意的 $\varepsilon > 0$，存在 $\delta > 0$，使得对任意的 $A \in \mathcal{B}(\Omega)$，只要 $m(A) < \delta$，我们就有

$$\sup_{k \geqslant 1} \left(\int_{\Omega} |f_k(x)| \mathbf{1}_A(x) dx_1 \cdots dx_n \right) < \varepsilon.$$

证明，

$$\lim_{R \to \infty} \left(\sup_{k \geqslant 1} \left(\int_{\Omega} |f_k(x)| \mathbf{1}_{[R, +\infty)}\big(|f_k(x)|\big) dx_1 \cdots dx_n \right) \right) = 0.$$

I6)（4 分）假设 $\{f_k\}_{k \geqslant 1}$ 在 Ω 上逐点地收敛到 f 并且满足 H4) 或者 H5) 中的假设，证明：

$$\lim_{k \to \infty} \int_{\Omega} |f_k(x) - f(x)| dx_1 \cdots dx_n = 0.$$

47 子流形上积分的计算：n-维球的体积，Archimedes 公式. 截断函数的构造，周期的单位分解，有界带边光滑区域，单位外法向量，在函数图像上的计算，Stokes 公式的第一个证明

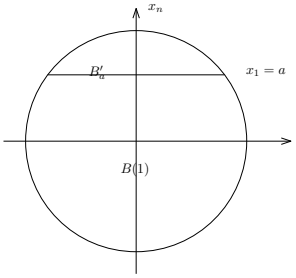
二零二零年四月二十日，星期一，晴

先补充 Fubini 公式的一个应用：

例子. 令 $c_n = m(B(1) \subset \mathbb{R}^n)$ 为 $\mathbb{R}^n$ 中单位球的体积，我们来计算所有的 c_n。对于 $n \geqslant 1$，我们令

$$B'_a = B(1) \cap \{x_n = a\}.$$

它的半径长为 $\sqrt{1-a^2}$



此时，我们有

$$\begin{aligned} c_n = m(B(1)) &= \int_{-1}^{1} m(B'_{x_n}) dx_n \\ &= 2\int_0^1 (1-x_n^2)^{n-1} c_{n-1} dx_n \\ &= 2c_{n-1} \int_0^{\frac{\pi}{2}} \sin^n \theta d\theta. \end{aligned}$$

其中，我们用了变元替换 $x_n = \cos\theta$。我们注意到，$I_n = \displaystyle\int_0^{\frac{\pi}{2}} \sin^n \theta d\theta$ 是我们上个学期研究过的 *Wallis* 积分，其中，

$$I_{2p} = \frac{(2p)!}{2^{2p}(p!)^2} \frac{\pi}{2}, \quad I_{2p+1} = \frac{2^{2p}(p!)^2}{(2p+1)!}.$$

从而，我们可以递归地计算 c_n。特别地，我们有

$$c_{2n} = \frac{\pi^n}{n!}.$$

根据递归公示，我们很容易证明

$$\lim_{n\to\infty} c_n = 0.$$

这说明当 n 变大时，$\mathbb{R}^n$ 中单位球的体积趋于 0。另外，如果借助计算器的话，我们可以很快看到 c_5 是最大的。

我们上次课讲过，如果给定函数图像 Γ_f，那么它上面的积分可以用如下的公式来计算：

$$\int_{\Gamma_f} \varphi d\sigma = \int_{\mathbb{R}^{n-1}} \varphi\big(x, f(x)\big)\sqrt{1+|\nabla f(x)|^2}dx.$$

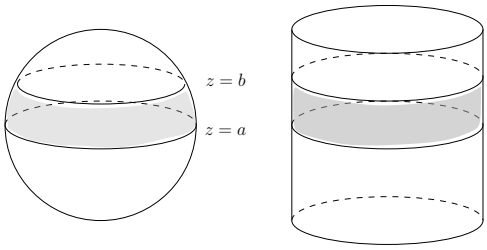
我们现在来看几个经典的例题：

例子 (Archimedes)**.** 考虑 $\mathbf{S}^2 \subset \mathbb{R}^3$ 是标准的单位球面，对于 $a, b \in [-1,1]$，我们令

$$\mathbf{S}^2(a,b) = \big\{(x,y,z) \in \mathbf{S}^2 \big| a \leqslant z \leqslant b\big\}.$$

我们令 $\mathbf{C}$ 为与 z-轴平行的圆柱面，并且这个圆柱面的直径是 2（恰好可以套在 $\mathbb{S}^2$ 上）。对于 $a, b \in [-1,1]$，我们令

$$\mathbf{C}(a,b) = \big\{(x,y,z) \in \mathbb{R}^3 \big| x^2+y^2 = 1, a \leqslant z \leqslant b\big\}.$$



这是两个平行的平面在 $\mathbf{S}^2$ 或者 $\mathbf{C}$ 上所截出的带状区域（图中的灰色区域）。*Archimedes* 的一个著名定理说，$\mathbf{S}^2(a,b)$ 的面积和 $\mathbf{C}(a,b)$ 的面积相等。我们来证明这个命题。

我们不妨假设 $a = 0, b > 0$（从这一点出发很容易证明 *Arichmedes* 的定理，这是因为我们将看到 $\mathbf{C}(a,b)$ 的面积正比于 $b-a$），此时，$\mathbf{S}^2(a,b)$ 可以写成函数的图像：

$$\mathbf{S}^2(0,b) = \big\{(x,y,z) \in \mathbb{R}^3 \big| z = \sqrt{1-x^2-y^2}, 1 \geqslant x^2+y^2 \geqslant 1-b^2\big\}.$$

此时，$f(x,y) = \sqrt{1-x^2-y^2}$，所以

$$d\sigma = \sqrt{1+|\nabla f(x)|^2}dxdy = \sqrt{\frac{1}{1-x^2-y^2}}dxdy.$$

所以，

$$\begin{aligned}
\sigma(\mathbf{S}^2(0,b)) &= \int_{1\geqslant x^2+y^2\geqslant 1-b^2} \sqrt{\frac{1}{1-x^2-y^2}}dxdy \\
&= \int_{\sqrt{1-b^2}}^1 \int_0^{2\pi} \sqrt{\frac{1}{1-r^2}} r d\theta dr \\
&= 2\pi \int_{\sqrt{1-b^2}}^1 \sqrt{\frac{1}{1-r^2}} r dr \\
&= \pi \int_{1-b^2}^1 \sqrt{\frac{1}{1-t}} dt \\
&= -2\pi\sqrt{1-t}\big|_{1-b^2}^1 = 2\pi b.
\end{aligned}$$

现在计算 $\mathbf{C}(0,b)$ 的面积，由对称性，我们把 x 看成是 y 和 z 的函数，有对称性，我们假设 $x\geqslant 0$，所以，只要算如下图形的面积即可：

$$\mathbf{C}^+(0,b)=\big\{(x,y,z)\in\mathbb{R}^3\big|x=\sqrt{1-y^2},-1\leqslant y\leqslant 1,0\leqslant z\leqslant b\big\}.$$

所以，我们有 $f(y,z)=\sqrt{1-y^2}$，从而，

$$d\sigma=\sqrt{\frac{1}{1-y^2}}dydz.$$

所以，

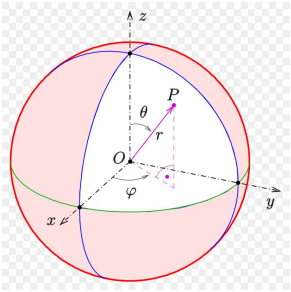
$$\begin{aligned}\sigma(\mathbf{C}^+(0,b))&=\int_{[-1,1]\times[0,b]}\sqrt{\frac{1}{1-y^2}}dydz\\&=b\int_{-1}^{1}\sqrt{\frac{1}{1-y^2}}dy\\&=b\pi.\end{aligned}$$

从而，

$$\sigma(\mathbf{C}(0,b))=2\sigma(\mathbf{C}^+(0,b))=\sigma(\mathbf{S}^2(0,b)).$$

我们还可以采取参数化的形式来计算 $\mathbf{S}^2(a,b)$ 的面积，比如说，我们可以利用球面坐标系：

$$\Phi:(0,\pi)\times(0,2\pi)\to\mathbf{S}^2,\quad(\theta,\varphi)\mapsto(\sin(\theta)\cos(\varphi),\sin(\theta)\sin(\varphi),\cos(\theta)).$$



由于 $z=\cos\theta$，所以 $\mathbf{S}^2(a,b)$ 由 $\theta\in[\arccos b,\arccos a]$ 所定义。此时，我们有

$${}^t\mathrm{Jac}(\Phi)=\begin{pmatrix}\cos\theta\cos(\varphi)&\cos\theta\sin(\varphi)&-\sin\theta\\-\sin\theta\sin(\varphi)&\sin\theta\cos(\varphi)&0\end{pmatrix}\quad\Rightarrow\quad\det G_\Phi=\sin^2\theta.$$

所以，$d\sigma=\sin\theta d\theta d\varphi$，从而

$$\begin{aligned}\sigma(\mathbf{S}^2(0,b))&=\int_{(0,\pi)\times(0,2\pi)}\sin\theta d\theta d\varphi\\&=2\pi\int_{\arccos b}^{\arccos a}\sin\theta d\theta\\&=2\pi(b-a).\end{aligned}$$

我们可以类似地计算 $\mathbf{C}(0,b)$ 的面积，考虑参数化

$$\Phi : (0,2\pi)\times(a,b)\to\sigma(\mathbf{C}(0,b)),\quad (\theta,s)\mapsto(\cos(\theta),\sin(\theta),s).$$

从而，

$${}^t\mathrm{Jac}(\Phi)=\begin{pmatrix}-\sin\theta & \cos\theta & 0\\ 0 & 0 & 1\end{pmatrix}\quad\Rightarrow\quad \det G_\Phi=1.$$

所以，$d\sigma=d\theta ds$，从而

$$\sigma(\mathbf{C}(0,b))=\int_{(0,2\pi)\times(a,b)}d\theta ds=2\pi(b-a).$$

Stokes 公式

我们先给出 Stokes 公式的第一个证明，其想法是把整体的公式转换为局部上的公式。为此，我们先证明一种特殊形式的单位分解定理，这是一个技术性的引理。

注记 (截断函数的存在性)**.** 我们在上个学期作业七习题 F 中（截断函数部分）构造了函数 $\psi(x)\in C^\infty(\mathbb{R})$，使得 $\psi(x)\geqslant 0$，在 $[-1,1]$ 上恒为 1，在 $[-2,2]$ 之外恒为零。另外，这个 ψ 还是偶函数，并且在 $[1,2]$ 之间是单调下降的。

我们现在来构造 $\mathbb{R}^n$ 上的光滑截断函数 $\varphi(x)\in C^\infty(\mathbb{R}^n)$，其中

$$\varphi(x)=\psi(x_1^2+\cdots+x_n^2)=\psi(r^2).$$

很明显，φ 只依赖于变量 r，它半径为 2 的球之外恒为 0，在半径为 1 的球之上恒为 1 并且这是一个对于半径 r 递减的函数。

我们还可以构造 $\mathbb{R}^n$ 上的光滑截断函数 $\chi(x)\in C^\infty(\mathbb{R}^n)$，其中

$$\chi(x)=\psi(x_1)\psi(x_2)\cdots\psi(x_n).$$

很明显，这个函数是光滑的正函数，它在一个中心在原点并且边长为 2 的正方体上恒为 1，在中心在原点并且边长为 4 的正方体之外恒为 0。

我们强调，这类函数具体构造并不重要，我们只用上面所列举的几条性质。

我们要构造一个与格点 $\mathbb{Z}^n\subset\mathbb{R}^n$ 相容的单位分解。为此，首先定义 $\mathbb{R}^n$ 上的函数：

$$F(x)=\sum_{(k_1,\cdots,k_n)\in\mathbb{Z}^n}\chi(x_1+k_1,\cdots,x_n+k_n).$$

这个函数是良好定义的：对于给定的 $x=(x_1,\cdots,x_n)\in\mathbb{R}^n$，这是一个有限和，因为当某个 $|k_i+x_i|\geqslant 2$ 时，$\chi(x_1+k_1,\cdots,x_n+k_n)=0$，所以，有贡献的项只是 $\mathbb{Z}^n$ 与中心在 x 处边长为 4 的正方体中的格点，这自然是一个有限集。

特别地，F 是光滑函数并且以 $\mathbb{Z}^n$ 为周期的周期函数，即对任意的 $(k_1,\cdots,k_n)\in\mathbb{Z}^n$，我们有

$$F(x_1+k_1,\cdots,x_n+k_n)=F(x_1,\cdots,x_n).$$

另外，根据 χ 的构造，对任意的 x，我们都有 $F(x)>0$。据此，对每个 $\mathbf{k}=(k_1,\cdots,k_n)\in\mathbb{Z}^n$，我们可以定义

$$F_{\mathbf{k}}(x_1,\cdots,x_n)=\frac{\chi(x_1-k_1,\cdots,x_n-k_n)}{F(x)},$$

我们得到了一组有紧支集[13]的光滑函数 $F_{\mathbf{k}}$，使得对任意的 $x\in\mathbb{R}^n$，我们有

$$\sum_{\mathbf{k}\in\mathbb{Z}^n}F_{\mathbf{k}}(x)=1.$$

对任意的 $\mathbf{k}\in\mathbb{Z}$，$F_{\mathbf{k}}$ 在某个中心在格点上面并且边长为 4 的正方体之外恒为 0（本来在某个小一点的正方体上恒为 1 的性质由于除以 F 所以不再成立）。

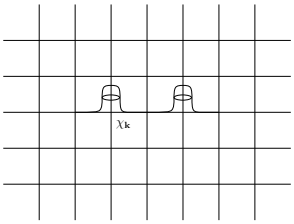
我们现在任意固定一个（很大的）正整数 $N\geqslant 1$。对每个 $\mathbf{k}\in\mathbb{Z}$，我们定义

$$\chi_{\mathbf{k}}(x)=F_{\mathbf{k}}(2^{N+1}x).$$

那么，$\chi_{\mathbf{k}}$ 在某个边长为 2^{-N} 的正方体上恒为 1，把这个正方体扩大一倍之后（边长为 2^{-N+1}），$\chi_{\mathbf{k}}$ 在这个大正方体之外恒为 0。特别地，这两个正方体的中心落在 $2^{-N+1}\mathbb{Z}^n$ 中（这是更密的格点）。我们自然还有

$$\sum_{\mathbf{k}\in\mathbb{Z}^n}\chi_{\mathbf{k}}\equiv 1.$$

综合上面的讨论，我们得到了如下的单位分解定理，其中，所谓的单位分解，指的是把 1 这个单位函数分解成若干函数的和：



引理 313 (周期的单位分解)**.** 对任意的 $N\geqslant 1$，令 $\Gamma_N=2^{-(N+1)}\mathbb{Z}^n$。那么，存在一族有紧支集的非负的光滑函数 $\{\chi_{\mathbf{k}}(x)\}_{\mathbf{k}\in\Gamma_N}$，满足如下的两个条件：

1) 对任意的 $\mathbf{k}\in\Gamma_N$，$\mathrm{supp}\chi_{\mathbf{k}}$ 落在以 $\mathbf{k}$ 为中心边长为 2^{-N+1} 的正方体中；

3) 常值函数 1 可以分解为：

$$1=\sum_{\mathbf{k}\in\Gamma_N}\chi_{\mathbf{k}}.$$

为了证明 Stokes 定理，我们还需要技术性的引理。这个引理实际上大有渊源，它和 Dirac δ-函数有关，我们会在下个学期的课程中经常用到这个引理。

[13]函数 f 的**支集**$\mathrm{supp}f$ 为

$$\mathrm{supp}f=\overline{\{x\in X\big|f(x)\neq 0\}},$$

即 f 的非零点集的闭包。

引理 314. 假设 $\chi(x)$ 是 $\mathbb{R}^1$ 上的有紧支集的光滑函数，我们假设 $\int_{\mathbb{R}^1}\chi(x)dx=1$。那么，对任意的连续函数 $f(x)$，我们都有

$$\lim_{\varepsilon\to 0^+}\left(\int_{\mathbb{R}}\frac{1}{\varepsilon}\chi\big(\frac{x}{\varepsilon}\big)f(x)dx\right)=f(0).$$

证明: 我们假设 $\mathrm{supp}(\chi)\subset[-M,M]$。根据变量替换公式，对任意的 $\varepsilon>0$，我们有

$$\int_{\mathbb{R}}\frac{1}{\varepsilon}\chi\big(\frac{x}{\varepsilon}\big)dx=1.$$

为了证明引理所要求的极限，我们做差:

$$\begin{aligned}\left|\int_{\mathbb{R}}\frac{1}{\varepsilon}\chi\big(\frac{x}{\varepsilon}\big)f(x)dx-f(0)\right|&=\left|\int_{\mathbb{R}}\frac{1}{\varepsilon}\chi\big(\frac{x}{\varepsilon}\big)f(x)dx-f(0)\int_{\mathbb{R}}\frac{1}{\varepsilon}\chi\big(\frac{x}{\varepsilon}\big)dx\right|\\&=\left|\int_{\mathbb{R}}\frac{1}{\varepsilon}\chi\big(\frac{x}{\varepsilon}\big)\big(f(x)-f(0)\big)dx\right|\\&=\left|\int_{-\varepsilon M}^{\varepsilon M}\frac{1}{\varepsilon}\chi\big(\frac{x}{\varepsilon}\big)\big(f(x)-f(0)\big)dx\right|.\end{aligned}$$

由于 f 在 0 处连续，所以对任意的 $\epsilon>0$，存在 $\delta>0$，使得当 $|x|<\delta$ 时，$|f(x)-f(0)|<\epsilon$。所以，当 $\varepsilon<\dfrac{\delta}{M}$ 时，上面的积分可以被下面的积分控制:

$$\left|\int_{\mathbb{R}}\frac{1}{\varepsilon}\chi\big(\frac{x}{\varepsilon}\big)f(x)dx-f(0)\right|\leqslant\int_{-\varepsilon M}^{\varepsilon M}\frac{1}{\varepsilon}\chi\big(\frac{x}{\varepsilon}\big)\epsilon dx=\epsilon.$$

所以，所要证明的极限成立。 □

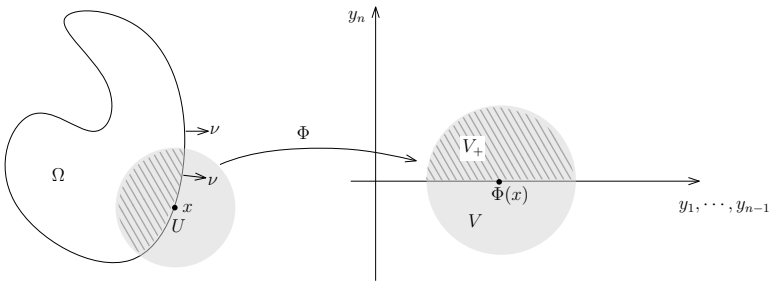
为了能够正确地陈述与证明 Stokes 公式，我们需要引入 $\mathbb{R}^n$ 中的（有界）**带边的光滑区域**（或者 C^1-光滑）的概念。很多教科书上在 Stokes 的证明方面语焉不详，很大程度上受制于没有正确地引入概念。

定义 315 (有界带边光滑区域)**.** 给定 $\mathbb{R}^n$ 中的紧集 $\Omega\subset\mathbb{R}^n$，我们假设对任意的 $x\in\Omega$，如下两种情况必居其一:

1) x 是一个**内点**，即存在开集 $U\subset\Omega$，使得 $x\in U$;

2) x 是一个**边界点**，即存在开集 $U\subset\mathbb{R}^n$，$x\in U$，存在开集 $V\in\mathbb{R}^n$，存在微分同胚 $\Phi:U\to V$，其中 U 上的坐标我们用 $x_1,\cdots,x_n$ 表示，V 上的坐标我们用 $y_1,\cdots,y_n$ 表示使得

   - $\Phi(U\cap\Omega)=V_+=V\cap\{(x_1,\cdots,x_n)|x_n\geqslant 0\}$;
   - $\Phi(x)$ 的 y_n 坐标是 0。



那么，我们把 Ω 称做是一个**有界带边光滑区域**。我们总是约定 Ω 的内点的集合非空。

注记. 根据定义，一个边界点绝对不能是内点，因为它的任何一个领域都与 Ω^c 相交。我们将 Ω 所有内点的集合记为 $\mathring{\Omega}$，边界点的集合记作 $\partial\Omega$。所以，$\Omega = \mathring{\Omega} \cup \partial\Omega$ 并且 $\mathring{\Omega} \cap \partial\Omega = \emptyset$。

在边界点的定义中，Φ 将边界点映射到 V_+ 的边界上，即 $\Phi^{-1}(V_+ \cap \{y_n = 0\}) \subset \partial\Omega$。

实际上，$\partial\Omega$ 是余维数为 1 的光滑子流形。证明是直截了当的：局部上（在上述定义中的 U 中），$x \in \partial\Omega$ 都被一个微分同胚映射成了 V 中 $y_n = 0$ 的点，按照子流形的定义，这是 $n-1$ 维的子流形。

对于 $\mathbb{R}^n$ 中的余维数为 1 的光滑子流形 M，对任意的 $x \in M$，我们可以找到两个单位长的（法）向量 $\pm\nu$，使得它们和 T_xM 是垂直的。

引理 316. 假设 Ω 是一个有界带边光滑区域。对于任意的 $p \in \partial\Omega$，存在唯一的 $\nu(p) \in T_p\mathbb{R}^n = \mathbb{R}^n$，使得

1) 这是 $\partial\Omega$ 的单位法向量，即 $\nu(p) \perp T_p\partial\Omega$ 并且 $|\nu(p)| = 1$；

2) 这个向量指向 Ω 的外部，即存在 $\varepsilon > 0$，使得对任意的 $t \in (0, \varepsilon)$，$p - t\nu(p) \in \Omega$。

我们称 $\nu(p)$ 是 Ω 在 $p \in \partial\Omega$ 处的**单位外法向量**。

证明: 按照定义，存在开集 $p \in U \subset \mathbb{R}^n$，$V \in \mathbb{R}^n$ 和微分同胚 $\Phi : U \to V$，使得 $x \in U$ 并且 $\Phi(\partial\Omega \cap U) = V \cap \{y_n = 0\}$ 并且 $\Phi(U \cap \Omega) = V_+$。如果我们令

$$f : U \to \mathbb{R}, \quad x \mapsto f(x) = \Phi^* y_n = y_n(\Phi) = \Phi_n(x),$$

其中 $\Phi(x) = (\Phi_1(x), \cdots, \Phi_n(x))$。那么，$f$ 的零点集就定义了 $\partial\Omega$ 而且

$$\Omega = \big\{x \in U \big| f(x) \geqslant 0\big\}.$$

此时，我们任意选取 $\mathbf{n}(p)$，使得 $\mathbf{n}(p) \perp T_p\partial\Omega$ 并且 $|\mathbf{n}(p)| = 1$。为了决定到底哪一个 $\pm\mathbf{n}(p)$ 是外法向量，利用如下的性质

$$\nabla_{\mathbf{n}(p)} f(p) \neq 0.$$

（否则对任意的 $v \in \mathbb{R}^n$，$\nabla_v f = 0$，从而 $df(p) = 0$，矛盾）通过选取 $\mathbf{n}$ 前面的符号，我们要求 $\big(\nabla_{\nu}(p) f\big)(p) < 0$。此时，对较小的 $\varepsilon > 0$，对任意的 $t \in (0, \varepsilon)$，我们有

$$f(p - t\nu(p)) = f(p) - t\big(\nabla_{\nu(p)} f\big)(p) + O(t^2) > 0.$$

这表明 $p - t\nu(p) \in \Omega$。 □

注记. 假设 Ω 是一个有界带边光滑区域，对任意的 $x \in \partial\Omega$，我们都可以唯一地指定它的外法向量 $\nu(x)$，所以，我们有映射

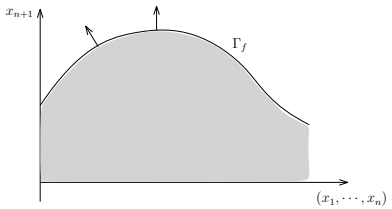
$$\nu : \partial\Omega \to \mathbb{R}^n, \quad x \mapsto \nu(x).$$

这实际上是光滑子流形 $\partial\Omega$ 上定义的光滑函数：这是因为每个余 1 维子流形局部上都可以写成函数图像的形式[14]，从而我们可以用下面例子的结论。

例子. 给定 $\mathbb{R}^n$ 上的光滑函数 $f:\mathbb{R}^n\to\mathbb{R}$，我们把 $\Omega\subset\mathbb{R}^{n+1}$ 定义为 f 的图像 Γ_f 下的区域（下面图中的灰色区域），即

$$\Omega=\big\{(x,x_n)\big|x\in\mathbb{R}^n,x_n\in\mathbb{R},x_n\leqslant f(x)\big\}.$$

我们计算 Ω 的单位外法向量（我们注意到 Ω 不是有界区域，但是这不影响我们计算法向量）。



证明: 很明显，$\partial\Omega=\Gamma_f$。给定 $(x,f(x))\in\partial\Omega$，其中 $x\in\mathbb{R}^n$，我们已经计算过

$$T_{(x,f(x))}\partial\Omega=\operatorname{span}\{e_1,\cdots,e_n\},\quad\text{其中 }e_i=(\underbrace{0,\cdots,0,1,0,\cdots,0}_{\text{第 }i\text{ 个位置处为 1，其余为 0}},\frac{\partial f}{\partial x_i}(x)).$$

据此，这个点处的单位外法向量必然是

$$\nu=\pm\frac{1}{(1+|\nabla f|^2)^{\frac12}}(-\nabla f,1)$$

我们现在确定 ν 的符号：考虑 $h(x_1,\cdots,x_n)=x_{n+1}-f(x_1,\cdots,x_n)$，那么，$\Omega=h^{-1}((-\infty,0])$。为了保证 ν 是外法向量，我们需要 $\nabla_\nu h>0$。我们计算

$$\nabla_{\frac{(-\nabla f,1)}{(1+|\nabla f|^2)^{\frac12}}}h=(1+|\nabla f|^2)^{\frac12}.$$

所以，

$$\nu=\frac{1}{(1+|\nabla f|^2)^{\frac12}}(-\nabla f,1).$$

特别地，此时，ν 对 x 是光滑依赖的。 □

我们现在叙述并证明 Stokes 公式：

定理 317. *假设 Ω 是一个有界带边光滑区域，$\nu(x)=(\nu_1(x),\cdots,\nu_n(x))$ 为 $\partial\Omega$ 的单位外法向量，$d\sigma$ 为 $\partial\Omega$ 上的曲面测度。对任意的 $\varphi\in C^1(\mathbb{R}^n,\mathbb{C})$，我们有*

$$\int_\Omega\frac{\partial\varphi}{\partial x_i}(x)\,dx=\int_{\partial\Omega}\varphi(x)\nu_i(x)\,d\sigma.$$

[14]我们总是可以假设子流形局部上是由光滑 $f(x_1,\cdots,x_n)=0$ 给出的，其中 $df(x)\neq0$。通过限制到更小的局部 U 上，我们可以假设 $\frac{\partial f}{\partial x_n}(x)\neq0$，根据隐函数定义，我们可以把 $f^{-1}(0)=M\cap U$ 写成 $x_n=h(x_1,\cdots,x_{n-1})$ 的形式这显然是函数图像。通过缩小这个区域，我们还可以假设它形如 $U'\times I$，其中 $U'\subset\mathbb{R}^{n-1}$ 为开集，$I\subset\mathbb{R}$ 为开区间

如果用向量值的函数来写，我们有

$$\int_\Omega \nabla\varphi dx = \int_{\partial\Omega} \varphi\cdot\nu d\sigma.$$

证明：证明分为四步，前两步是准备工作：

第一步，把 $\partial\Omega$ 写成若干个函数图像的并。

对于每个 $x\in\partial\Omega$，存在开集 $U_x\subset\mathbb{R}^n$，使得 $x\in U_x$，$\partial\Omega\cap U_x$ 是函数的图像，即在 U_x 上，存在 $k\leqslant n$，使得 $\partial\Omega$ 形如

$$\partial\Omega\cap U_x=\big\{(x_1,\cdots,x_n)\big|x_k=f(x_1,\cdots,x_{k-1},x_{k+1},\cdots,x_n)\big\}.$$

由于 $\partial\Omega$ 是紧集，所以存在有限个 $U_{x_1},\cdots,U_{x_m}$，使得 $\partial\Omega=U_{x_1}\cup\cdots\cup U_{x_m}$。

特别地，对于足够大的 $N>0$，如果一个边长为 2^{-N+1} 的正方体 Q 与 $\partial\Omega$ 的交非空，那么，存在 x_i，使得 $Q\subset U_{x_i}$。（请参考上学期第 11 次课关于 Lebesgue 数的讨论）

第二步，函数的局部化。

首先观察到，Stokes 公式的左右两边对于 φ 都是线性的。我们利用单位分解将 φ 限制到更小的开集上去：对任意的 $N\geqslant 1$，我们有一族有紧支集的非负的光滑函数 $\{\chi_{\mathbf{k}}(x)\}_{\mathbf{k}\in\Gamma_N}$，其中 $\Gamma_N=2^{-(N+1)}\mathbb{Z}^n$，使得对每个 $\mathbf{k}\in\Gamma_N$，$\mathrm{supp}\chi_{\mathbf{k}}$ 落在以 $\mathbf{k}$ 为中心边长为 2^{-N+1} 的正方体中并且：

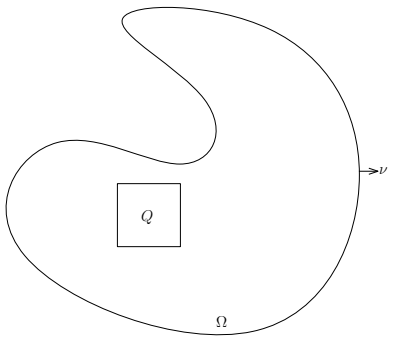
$$1=\sum_{\mathbf{k}\in\Gamma_N}\chi_{\mathbf{k}}.$$

从而，

$$\varphi=\sum_{\mathbf{k}\in\Gamma_N}\underbrace{\chi_{\mathbf{k}}(x)\varphi(x)}_{=\varphi_{\mathbf{k}}(x)}.$$

此时，每个 $\varphi_{\mathbf{k}}$ 在以 $\mathbf{k}$ 为中心边长为 2^{-N+1} 的正方体之外恒为 0。由于 Ω 是紧集，所以上述支集与 Ω 相交的函数的个数是有限个，所以，我们只要对这些函数来证明即可。我们仍然用 φ 表示 $\varphi_{\mathbf{k}}$，上面的讨论容许我们假设 $\mathrm{supp}(\varphi)$ 落在一个边长为 2^{-N+1} 的正方体 Q 中。

第三步，正方体 Q 与边界 $\partial\Omega$ 不相交的情况：$Q\subset\mathring{\Omega}$。



此时，由于 φ 在 Q 之外恒为 0，从而 $\dfrac{\partial \varphi}{\partial x_i}$ 在 Q 之外也恒为 0，所以，我们可以把积分限制到 Q 上：

$$\int_\Omega \frac{\partial \varphi}{\partial x_i} dx = \int_\Omega \frac{\partial \varphi}{\partial x_i} dx,$$

其中，$dx = dx_1 \cdots dx_n$。我们不妨假设 $Q = I_1 \times I_2 \times \cdots \times I_n$，其中，$I_i = [a_i, b_i]$。所以，根据 Fubini 公式和 Newton-Leibniz 公式，我们得到

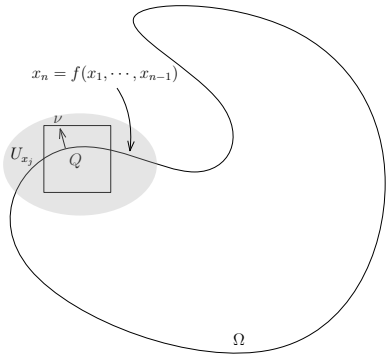
$$\begin{aligned}\int_\Omega \frac{\partial \varphi}{\partial x_i} dx &= \int_{I_1 \times \cdots \times \widehat{I_i} \times \cdots \times I_n} \left(\int_{a_i}^{b_i} \frac{\partial \varphi}{\partial x_i} dx_i \right) dx_1 \cdots \widehat{dx_i} \cdots dx_n \\ &= \int_{I_1 \times \cdots \times \widehat{I_i} \times \cdots \times I_n} \varphi(x_1, \cdots, b_i, \cdots, x_n) - \varphi(x_1, \cdots, a_i, \cdots, x_n) dx_1 \cdots \widehat{dx_i} \cdots dx_n.\end{aligned}$$

上面表达式中，一个符号上面加上 $\widehat{\ }$ 表示这个符号不在那里。由于 φ 的支集在 Q 中，所以

$$\varphi(x_1, \cdots, b_i, \cdots, x_n) = \varphi(x_1, \cdots, b_i, \cdots, x_n) = 0.$$

从而上面的积分为 0。另外，由于 $\dfrac{\partial \varphi}{\partial x_i}$ 在 $\partial\Omega$ 上为 0，所以此时 Stokes 公式成立。

第四步，正方体 Q 与边界 $\partial\Omega$ 相交的情况：$Q \cap \mathring{\Omega} \neq \emptyset$，此时，我们可以假设 $Q \subset U_{x_j}$，其中，U_{x_j} 是第一步中构造的开集。



由于 $\partial\Omega \cap U_{x_j}$ 为函数的图像，我们不妨假设

$$\partial\Omega \cap U_{x_j} = \big\{(x_1, \cdots, x_{n-1}, x_n) \big| x_n = f(x_1, \cdots, x_{n-1})\big\}.$$

我们还假设 Ω 是函数图像下的部分，即 $x_n \leqslant f(x_1, \cdots, x_{n-1})$。我们定义

$$\rho(x_1, \cdots, x_n) = x_n - f(x_1, \cdots, x_{n-1}).$$

并且选取函数光滑递增的单变量实函数 $\theta(x)$：

$$\theta(t) = \begin{cases} 1, & x \geqslant 1; \\ 0, & x \leqslant -1. \end{cases}$$

（函数的存在性请参考上个学期作业七习题 F）我们知道，当 $\varepsilon \to 0$ 时，我们有函数的逐点收敛：

$$\theta\left(\frac{\rho(x)}{\varepsilon}\right) \longrightarrow \mathbf{1}_{\Omega\cap Q}.$$

根据 Lebesgue 控制收敛定理，我们有

$$\begin{aligned}\int_\Omega -\frac{\partial\varphi}{\partial x_i}(x)dx &= \int_Q -\frac{\partial\varphi}{\partial x_i}(x)dx \\ &\overset{Lebesgue}{=} \lim_{\varepsilon\to 0}\int_Q -\varphi(x)\theta\left(\frac{\rho(x)}{\varepsilon}\right)dx \\ &= \lim_{\varepsilon\to 0}\int_Q \frac{1}{\varepsilon}\theta'\left(\frac{\rho(x)}{\varepsilon}\right)\frac{\partial\rho}{\partial x_i}(x)\varphi(x)dx.\end{aligned}$$

最后一个等号我们利用 Fubini 公式可以对 x_i 变量进行分部积分，这和第三步中的计算一模一样。

为了可以利用 Newton-Leibniz 公式，我们构造一个微分同胚（坐标变换）把 $\Omega\cap Q$ 变成矩形。我们变量替换

$$H:\Omega\cap Q\to\mathbb{R}^n_- = \big\{(y_1\cdots,y_n)\big|y_n\leqslant 0\big\},\quad \begin{pmatrix}x_1\\ \cdots\\ x_{n-1}\\ x_n\end{pmatrix}\mapsto\begin{pmatrix}y_1\\ \cdots\\ y_{n-1}\\ y_n\end{pmatrix}=\begin{pmatrix}x_1\\ \cdots\\ x_{n-1}\\ x_n-f(x_1,\cdots,x_{n-1})\end{pmatrix}.$$

显然，我们有 $\mathbf{J}_\Phi\equiv 1$，利用换元积分公式，我们有

$$\begin{aligned}\int_\Omega -\frac{\partial\varphi}{\partial x_i}(x)dx &= \lim_{\varepsilon\to 0}\int_Q \frac{1}{\varepsilon}\theta'\left(\frac{y_n}{\varepsilon}\right)\left(\frac{\partial\rho}{\partial x_i}\right)(H^{-1}(y))\varphi(H^{-1}(y))dy \\ &\overset{Fubini}{=} \lim_{\varepsilon\to 0}\int_{I_1\times\cdots\times I_{n-1}}\left(\int_{I_n}\frac{1}{\varepsilon}\theta'\left(\frac{y_n}{\varepsilon}\right)\left(\frac{\partial\rho}{\partial x_i}\right)(H^{-1}(y))\varphi(H^{-1}(y))dy_n\right)dy_1\cdots dy_{n-1}.\end{aligned}$$

蓝色项具有特殊的形式。注意到 $\theta'(y_n)$ 这个函数的积分为 1，所以，根据我们证明的第二个技术性引理（Dirac δ-函数），我们就有

$$\begin{aligned}&\lim_{\varepsilon\to 0}\int_{I_n}\frac{1}{\varepsilon}\theta'\left(\frac{y_n}{\varepsilon}\right)\left(\frac{\partial\rho}{\partial x_i}\right)(H^{-1}(y))\varphi(H^{-1}(y))dy_n \\ &= \left(\frac{\partial\rho}{\partial x_i}\right)(H^{-1}(y_1,\cdots,y_{n-1},0))\varphi(H^{-1}(y_1,\cdots,y_{n-1},0)).\end{aligned}$$

我们注意到，$y_n=0$ 实际上就是超曲面 $\partial\Omega$。为了书写简单，我们令 $Q'=I_1\times\cdots\times I_{n-1}$, $x'=(x_1,\cdots,x_{n1})$。所以，根据 Lebesgue 控制收敛定义（交换积分和极限），我们有

$$\begin{aligned}\int_\Omega -\frac{\partial\varphi}{\partial x_i}(x)dx &= \int_{Q'}\left(\frac{\partial\rho}{\partial x_i}\right)(H^{-1}(y_1,\cdots,y_{n-1},0))\varphi(H^{-1}(y_1,\cdots,y_{n-1},0))dy_1\cdots dy_{n-1} \\ &= \int_{Q'}\left(\frac{\partial\rho}{\partial x_i}\right)(x_1,\cdots,x_{n-1},x_n)\Big|_{\rho=0}\varphi(x_1,\cdots,x_n)\Big|_{\rho=0}dx_1\cdots dx_{n-1} \\ &= \int_{Q'}\left(\frac{\partial\rho}{\partial x_i}\right)(x',x_n)\Big|_{\partial\Omega}\varphi(x',f(x'))dx' \\ &= \int_{Q'}\frac{1}{\sqrt{1+|\nabla f(x')|^2}}\left(\frac{\partial\rho}{\partial x_i}\right)(x',x_n)\Big|_{\partial\Omega}\varphi(x',f(x'))\sqrt{1+|\nabla f(x')|^2}dx'.\end{aligned}$$

我们注意到，$\varphi(x', f(x'))\sqrt{1+|\nabla f(x')|^2}dx'$ 就是 $\partial\Omega$ 的曲面测度 $d\sigma$。另外，根据之前的计算，

$$\nu = \frac{1}{(1+|\nabla f|^2)^{\frac{1}{2}}}(-\nabla f, 1),$$

所以，

$$\frac{1}{\sqrt{1+|\nabla f(x')|^2}}\left(\frac{\partial \rho}{\partial x_i}\right)(x', x_n)\Big|_{\partial\Omega} = -\nu_i.$$

上式就可以写成

$$\int_\Omega -\frac{\partial \varphi}{\partial x_i}(x)dx = \int_{\partial\Omega} -\nu_i \varphi d\sigma.$$

这就对这种情况证明了 Stokes 公式，从而命题得证。 □

注记. 上述证明的一个和核心想法是把函数 φ 拆成支集很小的函数的和来进行证明而不是把 Ω 拆成更小的集合！也就是说，把 Ω 分成小块是更直观的看法，但是为了实现这个想法，我们应该走到函数的层次上。

48 Sard 型引理，Stokes 公式的微分拓扑证明

二零二零年四月二十三日，星期四，晴

Stokes 定理的另一个证明

我们先证明微分拓扑学中 Sard 定理的一个特殊形式，它的证明用到了换元积分公式证明里的基本想法和技巧：

引理 318. *$U \subset \mathbb{R}^n$ 是开集，$\Phi : U \to \mathbb{R}^n$ 是 C^1 的映射。我们定义 Φ 的临界点集或者奇异点集 $\mathrm{Sing}(\Phi)$ 为*

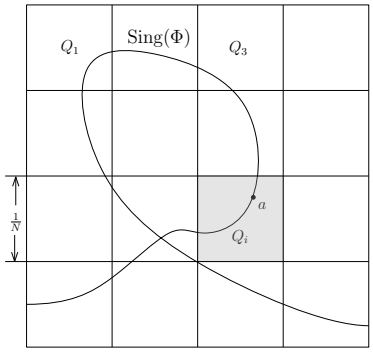
$$\mathrm{Sing}(\Phi) = \big\{x \in U \big| \mathrm{rank}\,(d\Phi(x)) < n\big\} = \big\{x \in U \big| \det\left(\mathrm{Jac}_\Phi(x)\right) = 0\big\}.$$

那么，$\Phi\left(\mathrm{Sing}(\Phi)\right) \subset \mathbb{R}^n$ 是零测集。

证明： 我们可以将 U 写成可数个闭的正方体的并集

$$U = \bigcup_{i=1}^{\infty} Q_i,$$

其中 Q_i 是闭正方体。由于可数个零测集的并还是零测集，所以只要对每个 Q_i 来证明 $\Phi\left(Q_i \cap \mathrm{Sing}(\Phi)\right)$ 是零测集即可。



我们只需要考虑单独的一个正方体 Q 即可。不妨假设 Q 的边长为 1。我们沿用在换元积分公式的证明中所用的范数，即对任意 $x = (x_1, \cdots, x_n) \in \mathbb{R}^n$ 和 $n \times n$ 的矩阵 $A = (A_{ij})$，有

$$\|x\| = \sup_{i \leqslant n} |x_i|, \quad \|A\| = \sup_{i \leqslant n} \big(\sum_{j=1}^{n} |A_{ij}|\big).$$

由于 Φ 是 C^1 的，所以映射 $d\Phi$ 在 Q 上是连续的，从而在 Q（紧集）上一致连续。据此，我们将 Q 进一步分解为边长为 $h = \dfrac{1}{N}$ 的 N^n 个的正方体 Q_i，其中 $N \in \mathbb{Z}$ 是很大的正整数，使得在每个 Q_i 上面，都有

$$\|d\Phi(x) - d\Phi(y)\| \leqslant \delta, \quad \forall x, y \in Q_i.$$

我们考察每一个 Q_i。如果 $Q_i \cap \mathrm{Sing}(\Phi) = \emptyset$，那么 Q 对 $\Phi(\mathrm{Sing}(\Phi))$ 的测度没有贡献，我们可以忽略这一类的笑正方体。现在假设 $Q_i \cap \mathrm{Sing}(\Phi) \neq \emptyset$，我们选定 $a \in Q_i \cap \mathrm{Sing}(\Phi)$。对于任意的 $x \in Q_i$ 和任意的指标 $k \leqslant n$，根据 Lagrange 中值定理（对 Φ 的分量用），我们有

$$\Phi_k(x) - \Phi_k(a) - d\Phi_k(a)(x-a) = \sum_{j \leqslant n} \left(\frac{\partial \Phi_k}{\partial x_j}(\xi_k) - \frac{\partial \Phi_k}{\partial x_j}(a) \right) (x_j - a_j),$$

其中 ξ_k 为线段 $\overline{xa}$ 上的某点（从而，$\xi_k \in Q_i$）。所以，

$$\begin{aligned}
\|\Phi(x) - \Phi(a) - d\Phi(a)(x-a)\| &= \sup_{k \leqslant n} |\Phi_k(x) - \Phi_i(a) - d\Phi_k(a)(x-a)| \\
&\overset{Lagrange}{\leqslant} \sup_{k \leqslant n} \sum_{j \leqslant n} |\frac{\partial \Phi_k}{\partial x_j}(\xi_k) - \frac{\partial \Phi_k}{\partial x_j}(a)||(x_j - a_j)| \\
&\leqslant \|x-a\| \sup_{k \leqslant n} \sum_{j \leqslant n} |\frac{\partial \Phi_i}{\partial x_j}(\xi_k) - \frac{\partial \Phi_i}{\partial x_j}(a)| \\
&= \|d\Phi(x) - d\Phi(a)\| \|x-a\| \\
&\leqslant \frac{\delta}{N}.
\end{aligned}$$

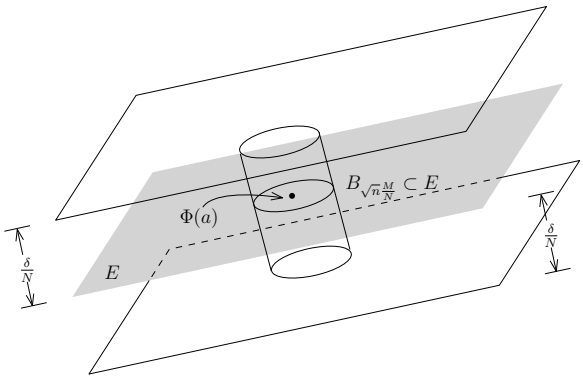
即

$$\|\Phi(x) - (\Phi(a) + d\Phi(a)(x-a))\| \leqslant \frac{\delta}{N}.$$

按照奇异集 $\mathrm{Sing}(\Phi)$ 的定义，$d\Phi(a)$ 的秩至多是 $n-1$，所以集合

$$\left\{ x \in Q_i \middle| \Phi(a) + d\Phi(a)(x-a) \right\}$$

落在某个过 $\Phi(a)$ 点的 $n-1$ 维的超平面里面 $E \subset \mathbb{R}^n$。我们刚刚在证明的不等式表明，$\Phi(Q_i)$ 落在离这个超平面的距离不超过 $\dfrac{\delta}{N}$ 距离的地方（这里我们用 Euclid 距离，上述用的范数很明显是不超过 Euclid 的范数的（= 平方和再开方））。



另外，根据上面的不等式的证明，我们有

$$\|\Phi(x) - \Phi(a)\| \leqslant \|d\Phi(a)\||x-a| + \|d\Phi(x) - d\Phi(a)\|\|x-a\| \leqslant M\|x-a\|.$$

其中，$M=\sup\limits_{x\in Q}\|d\Phi(x)\|$（整个 Q 上的最大值）。根据，$|x|\leqslant\sqrt{n}\|x\|$，我们就有（Euclid 距离）

$$|\Phi(x)-\Phi(a)|\leqslant\sqrt{n}\frac{M}{N}.$$

也就是说，$\Phi(Q_i)$ 像中的每个点到 $\Phi(a)$ 的距离都不超过 $\sqrt{n}\dfrac{M}{N}$。我们把 $\Phi(Q_i)$ 正交投影到 E 上，它的像一定包含在一个以 $\Phi(a)$ 为中心以 $\sqrt{n}\dfrac{M}{N}$ 为半径的球 $B_{\sqrt{n}\frac{M}{N}}$ 里面。所以，$\Phi(Q_i)$ 落在以 $B_{\sqrt{n}\frac{M}{N}}$ 为底（截面），以 $2\times\dfrac{\delta}{N}$ 为高的圆柱里面，它的体积不超过

$$\left(c_{n-1}\sqrt{n}\frac{M}{N}\right)^{n-1}\times\left(2\times\frac{\delta}{N}\right)=C_nM^{n-1}\frac{\delta}{N^n}.$$

由于至多有 N^n 个这样的 Q_i，所以，

$$\begin{aligned}m\left(\Phi(\mathrm{Sing}(\Phi))\right)&\leqslant\sum_{Q_i\cap\mathrm{Sing}\Phi\neq\emptyset}m(\Phi(Q_i))\\&\leqslant N^n\times\left[C_nM^{n-1}\frac{\delta}{N^n}\right]\\&=C_nM^{n-1}\delta.\end{aligned}$$

由于 δ 是任意选取的，所以 $\Phi\left(\mathrm{Sing}(\Phi)\right)$ 是零测集。 □

我们对换元积分公式的证明稍加改造，就可以证明更强一点的结论：

练习. 假设 U 和 V 是 $\mathbb{R}^n$ 中的开集，$\Phi:U\to V$ 是同胚而且 Φ 是 C^1 的映射（它的逆未必是 C^1 同胚），那么换元积分公式仍然成立（比如，我们可以考虑 $\mathbb{R}^1$ 上的变元替换 $x\mapsto x^3$）。

我们把它留作本次的作业。

另外，类似于上述引理的表述，我们还可以讨论到 $\mathbb{R}^1$ 的映射的奇异点的集合：假设 $M\subset\mathbb{R}^n$ 是 $n-1$-维的子流形，考虑投影映射

$$\pi:M\to\mathbb{R}^{n-1},\ \ (x_1,\cdots,x_{n-1},x_n)\mapsto(x_1,\cdots,x_{n-1}).$$

令

$$\mathrm{Sing}(\pi)=\left\{x\in M\middle|\mathrm{rank}\left(d\pi(x)\right)<n-1\right\}.$$

那么，$\pi(\mathrm{Sing}(\pi))\subset\mathbb{R}^{n-1}$ 是零测集。我们把它留作本次的作业。

还有一个所谓的“扭曲”版本的 Fubini 定理，尽管我们证明 Stokes 公式用不到这个定理：$\Omega\subset\mathbb{R}^n$ 是开集，$F:U\to\mathbb{R}$ 是光滑映射。假设对任意的 $x\in U$，$dF(x)\neq0$（从而，$\Sigma_t=F^{-1}(t)$ 是余 1 维子流形）。那么，对任意的 U 上的可积函数 f，我们有

$$\int_\Omega f(x)dx=\int_{\mathbb{R}}\left(\int_{\Sigma_t}\frac{f}{|\nabla F|}d\sigma_t\right)dt,$$

其中，$d\sigma_t$ 是 Σ_t 上的子流形测度。我们也将在这次作业中证明这个命题。

我们现在给出 Stoke 公式的一个新的证明：与之前的相比，这个证明更几何，更整体。

只要证明

$$\int_\Omega \frac{\partial \varphi}{\partial x_1}(x)\ dx = \int_{\partial\Omega} \varphi(x)\nu_1(x)\ d\sigma.$$

即可。

我们考虑 $\mathbb{R}^n$ 上的投影映射：

$$\pi : \mathbb{R}^n \to \mathbb{R}^{n-1},\ \ (x_1, \cdots, x_n) \mapsto (x_2, \cdots, x_n).$$

把它限制到 $\partial\Omega$ 上（仍记作 π），我们就有

$$\pi : \partial\Omega \to \mathbb{R}^{n-1}.$$

我们定义 π 的临界点集：

$$\mathrm{Sing}(\pi) = \big\{x \in \partial\Omega \big| \mathrm{rank}\big(d\pi(x)\big) < n-1\big\}.$$

由于 $\partial\Omega$ 是 $n-1$ 维的子流形，对任意的 $x \in \partial\Omega$，存在包含 x 开集 $U \subset \mathbb{R}^n$ 以及 $\mathbb{R}^n$ 中的开集 V 和微分同胚

$$\Phi : U \to V$$

使得

$$\Phi : U \cap \partial\Omega = V \cap (\mathbb{R}^{n-1} \times \{0\}).$$

从而，我们可以对

$$\pi \circ \Phi^{-1} : V \cap (\mathbb{R}^{n-1} \times \{0\}) \to \mathbb{R}^{n-1}$$

应用刚才的引理（与微分同胚复合把奇异点集映射到奇异点并且不会改变零测集的性质）。

$$\begin{array}{ccc} V \cap (\mathbb{R}^{n-1} \times \{0\}) & \xrightarrow{\Phi^{-1}} & U \cap \partial\Omega \\ & \searrow^{\pi\circ\Phi^{-1}} & \downarrow \pi \\ & & \mathbb{R}^{n-1} \end{array}$$

所以，$\pi\left(\mathrm{Sing}(\pi)\right) \subset \mathbb{R}^{n-1}$ 是零测集。所以，我们有

$$\begin{aligned} \int_\Omega \frac{\partial\varphi}{\partial x_1}(x)dx &= \int_{\mathbb{R}^m} \mathbf{1}_\Omega(x)\frac{\partial\varphi}{\partial x_1}(x)dx \\ &\overset{Fubini}{=} \int_{\mathbb{R}^{n-1}} \left(\int_{\mathbb{R}} \mathbf{1}_\Omega(x_1, \cdots, x_n)\frac{\partial\varphi}{\partial x_1}(x_1, \cdots, x_n)\ dx_1\right) dx_2 \cdots dx_n \\ &= \int_{\mathbb{R}^{n-1}-\pi(\mathrm{Sing}(\pi))} \left(\int_{\mathbb{R}} \mathbf{1}_\Omega(x_1, \cdots, x_n)\frac{\partial\varphi}{\partial x_1}(x_1, \cdots, x_n)\ dx_1\right) dx_2 \cdots dx_n. \end{aligned}$$

我们需要研究 $\mathbb{R}^{n-1} - \pi\left(\mathrm{Sing}(\pi)\right)$ 的一些简单的几何／拓扑性质。

首先注意到，对任意的 $x \in \partial\Omega - \mathrm{Sing}(\pi)$，$x$ 附近必然又一个开集 U 使得 $x \in U \subset \partial\Omega - \mathrm{Sing}(\pi)$，这表明 $\mathrm{Sing}(\pi)$ 是闭集，从而是紧集。由于紧集的在连续映射下的像还是紧集，所以，$\pi(\mathrm{Sing}(\pi)) \subset \mathbb{R}^{n-1}$ 是紧集，从而，$\mathbb{R}^{n-1} - \pi(\mathrm{Sing}(\pi))$ 是开集。

定义 319 (道路连通性)**.** $X \subset \mathbb{R}^n$ 是子集，如果对任意 $x, x' \in X$，我们都有连续映射

$$\gamma : [0,1] \to X \subset \mathbb{R}^n,$$

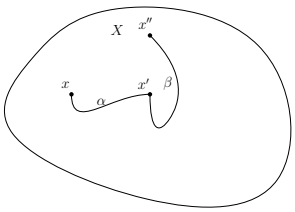
使得 $\gamma(0) = x$，$\gamma(1) = x'$，我们就称 X 是**道路连通的**。如果某两个点满足上述性质，我们就说 x 与 x' 之间存在道路。

注记. 对任意的集合 $X \subset \mathbb{R}^n$，我们定义如下的等价关系：$x \sim x'$ 当且仅当 x 与 x' 之间存在道路。我们验证这是等价关系：

- $x \sim x$：我们可以去 $\gamma(t) \equiv x$；
- $x \sim x' \Rightarrow x' \sim x$：首先，存在连续映射 $\gamma : [0,1] \to X$，使得 $\gamma(0) = x$，$\gamma(1) = x'$；我们定义

  $$\widetilde{\gamma} : [0,1] \to X, \quad t \mapsto \gamma(1-t).$$

  这显然是连续映射并且 $\widetilde{\gamma}(0) = x'$，$\widetilde{\gamma}(1) = x$；
- $x \sim x', x' \sim x'' \Rightarrow x \sim x''$：首先，存在连续映射 $\alpha : [0,1] \to X$，使得 $\alpha(0) = x$，$\alpha(1) = x'$；存在连续映射 $\beta : [0,1] \to X$，使得 $\beta(0) = x'$，$\beta(1) = x''$。



我们定义

$$\gamma : [0,1] \to X, \quad t \mapsto \gamma(t) = \begin{cases} \alpha(2t), & t \in [0, \frac{1}{2}]; \\ \beta(2t-1), & t \in [\frac{1}{2}, 1] \end{cases}.$$

这给出了从 x 到 x' 的道路。

所以，我们可以把 X 分拆为不同等价类的无交并 $X = \bigcup_i X_i$，每个 X_i 我们都称作是 X 的一个**连通分支**。比如，在 $\mathbb{R}^2$ 上，集合 $\mathbb{R}^2 - \big\{(x,y)\big|xy = 0\big\}$ 有四个连通分支。

注记. 如果 $U \subset \mathbb{R}^n$ 是开集，那么每个连通 U 的连通分支都是开集，因为每个 x 附近的点自然和 x 之间存在道路。此时，U 的连通分支的个数是可数的（因为每个连通分支当中我们都可以取一个有理点来标记）。

引理 320. *假设 $U \subset \mathbb{R}^n$ 是连通的开集，函数*

$$f : U \to \mathbb{Z} \subset \mathbb{R}$$

是连续映射，那么 f 是常值映射。

这是一个非常有用的引理。

证明: 如果不然，那么不妨假设存在两个不同的整数 k 和 ℓ 以及 $x, y \in U$，使得 $f(x) = k$，$f(y) = \ell$。我们任选一条连接 x 和 y 的道路

$$\gamma : [0,1] \to X \subset \mathbb{R}^n,$$

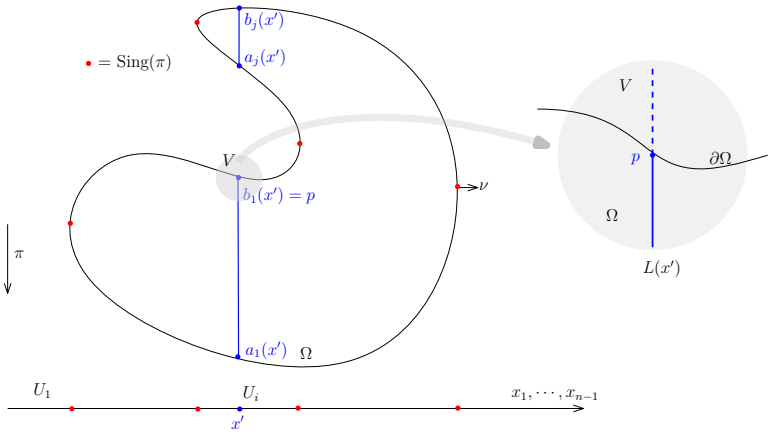
使得 $\gamma(0) = x$，$\gamma(1) = y$，那么函数 $f(\gamma(t))$ 根据介值定理，存在某个 $t' \in (0,1)$，使得 $f(\gamma(t'))$ 是 k 与 ℓ 之间的一个非整数，矛盾。 □

我们现在回到 Stokes 定理的证明。

由于 $\mathbb{R}^{n-1} - \pi\left(\mathrm{Sing}(\pi)\right)$ 是开集，所以，它可以写成（可数个）不交的（道路）连同开集 $\{U_i\}_{i\geqslant 1}$ 的并：

$$\mathbb{R}^{n-1} - \pi\left(\mathrm{Sing}(\pi)\right) = \bigcup_{i=1}^{\infty} U_i.$$

这些连通分支自然是两两不交的。



所以，我们有

$$\int_\Omega \frac{\partial \varphi}{\partial x_1}(x) dx = \sum_i \left[\int_{U_i} \left(\int_{\mathbb{R}} \mathbf{1}_\Omega(x_1, \cdots, x_n) \frac{\partial \varphi}{\partial x_1}(x_1, \cdots, x_n)\, dx_1 \right) dx_2 \cdots dx_n \right].$$

现在固定连通分支 i，对于每个 $x' \in U_i$，我们考虑 x' 在 π 下的逆向

$$L_{x'} = \pi^{-1}(x') = \left\{ (t, x') \middle| t \in \mathbb{R} \right\}.$$

考虑 Ω 的边界与 $L_{x'}$ 的一个交点（如果存在的话）$p \in \partial\Omega \cap L_{x'}$，在 p 附近的小领域 V 上（右边的图是局部上的放大）。由于 $x' \notin \pi\left(\mathrm{Sing}(\pi)\right)$，所以

$$d\pi(p) : T_p \partial\Omega \to T_{x'} \mathbb{R}^{n-1} = \mathbb{R}^{n-1}$$

是线性同构。根据

$$d\pi(p) : T_pL(x') \to T_{x'}\mathbb{R}^{n-1} = \mathbb{R}^{n-1}$$

是零映射（因为 $\pi L(x') = x'$，求微分立得），所以，$T_pL(x') \subset \ker(d\pi(p))$。这表明（数维数）：

$$T_p\mathbb{R}^n = T_p\partial\Omega \oplus T_pL(x').$$

所以，假设 Ω 在 V 上由 $f(x_1, \cdots, x_n) = 0$ 定义并且 $f(x) \leqslant 0$ 定义了 Ω。那么，对于 $v = \dfrac{\partial}{\partial x_1} \in T_pL(x')$，我们有 $\nabla_v f(p) \neq 0$（否则 $df(p) \equiv 0$，矛盾）。与 Ω 的单位外法向量的构造一致，通过缩小 V，我们知道 $L(x') \cap V$ 被 p 分成了两段，一段（实线）在 Ω 中，一段（虚线）在 Ω 外面。特别地，这表明，对于 $p \in \partial\Omega \cap L_{x'}$，$p$ 在 $L(x')$ 上面的附近（出了自己）没有点与 $\partial\Omega$ 相交。据此，我们知道 $\partial\Omega \cap L_{x'}$ 在 (x') 上没有聚点，又因为 $\partial\Omega$ 是有界的，所以，$\partial\Omega \cap L_{x'}$ 只有有限个点（可以没有）：我们把它们按照 x_1 坐标的大小依次记作

$$a_1(x'), b_1(x'), a_2(x'), b_2(x'), \cdots, a_k(x'), b_k(x') \in L(x').$$

这里我们一共有 $2k$（偶数个点），这是因为，当 x_1 坐标很负的时候，$L(x')$ 的点在 Ω 之外（因为 Ω 是有界的）。根据上面的证明，$a_1(x')$ 是第一次（如果有的话）$L(x')$ 进入 Ω，$b_1(x')$ 是接下来 $L(x')$ 离开 Ω 时与 $\partial\Omega$ 的交点，依次类推，因为最终当 x_1 很正（大）的时候，$L(x')$ 的点在 Ω 之外，所以最终 $L(x')$ 要离开 Ω，这说明有偶数个点。

这样子，$L_{x'} \cap \Omega$ 是 k 个不交的闭区间之并，即

$$L_{x'} \cap \Omega = \coprod_{j=1}^{k(x')} [a_j(x'), b_j(x')],$$

我们注意到区间的个数可能依赖于 x'。

根据隐函数定理，$a_j(x')$（或者 $b_j(x')$）是 x' 的光滑函数（因为 $\nabla \frac{\partial}{\partial x_1} f(a_j(x')) \neq 0, f(x', a_j(x')) = 0$，这就是隐函数定理的叙述），所以在 x' 附近，区间的个数 $k(x')$ 是常数，所以 $x' \to k(x')$ 是连续函数，再根据 U_i 的连通性，我们知道，在每个 U_i 上，对任意的 $x' \in U_i$，$k(x') \equiv k$。从而，我们都有

$$\begin{aligned}
&\int_{U_i} \left(\int_{\mathbb{R}} \mathbf{1}_\Omega(x_1, \cdots, x_n) \frac{\partial \varphi}{\partial x_1}(x_1, \cdots, x_n)\ dx_1 \right) dx_2 \cdots dx_n \\
=& \int_{U_i} \left(\int_{\mathbb{R} \cap L(x')} \frac{\partial \varphi}{\partial x_1}(x_1, \cdots, x_n)\ dx_1 \right) dx_2 \cdots dx_n \\
=& \sum_{j=1}^{k} \int_{U_i} \left(\int_{a_i(x')}^{b_i(x')} \frac{\partial \varphi}{\partial x_1}(x_1, x')\ dx_1 \right) dx'.
\end{aligned}$$

在每个区间 $[a_i(x'), b_i(x')]$ 上用 Newton-Leibniz 法则，我们得到

$$\begin{aligned}
&\int_{U_i} \left(\int_{\mathbb{R}} \mathbf{1}_\Omega(x_1, \cdots, x_n) \frac{\partial \varphi}{\partial x_1}(x_1, \cdots, x_n)\ dx_1 \right) dx_2 \cdots dx_n \\
=& \sum_{j=1}^{k} \int_{U_i} \big(\varphi(b_j(x'), x') - \varphi(a_j(x'), x') \big) \big) dx'.
\end{aligned}$$

我们考察其中的一项，比如说

$$\int_{U_i} \varphi(b_j(x'), x')dx'$$

我们注意到，此时，$x' \mapsto b_j(x')$ 将 $\partial\Omega \cap \pi^{-1}(U_i)$ 的一部分 Ω_{i,b_j} 实现为函数图像所定义的曲面，所以，

$$\begin{aligned}\int_{U_i} \varphi(b_j(x'), x')dx' &= \int_{\partial\Omega_{i,b_j}} \frac{1}{\sqrt{1+|\nabla b_j|^2}}\varphi(b_j(x'), x')d\sigma \\ &= \int_{\partial\Omega_{i,b_j}} \nu_1 f d\sigma.\end{aligned}$$

类似地，

$$\int_{U_i} \varphi(a_j(x'), x')dx' = \int_{\partial\Omega_{i,a_j}} \nu_1 \varphi d\sigma.$$

根据

$$\partial\Omega \cap \pi^{-1}(U_i) = \bigcup_{i\leqslant k} \left(\partial\Omega_{i,a_j} \cup \partial\Omega_{i,b_j}\right),$$

我们最终得到

$$\int_{U_i} \left(\int_{\mathbb{R}} \mathbf{1}_\Omega(x_1, \cdots, x_n) \frac{\partial\varphi}{\partial x_1}(x_1, \cdots, x_n)\ dx_1 \right) dx_2 \cdots dx_n = \int_{\partial\Omega\cap\pi^{-1}(U_i)} \nu_1 \varphi d\sigma.$$

再对所有的连通分支 U_i 求和，根据 $\pi^{-1}(\mathbb{R}^{n-1} - \pi\left(\mathrm{Sing}(\pi)\right)) = \partial\Omega - \mathrm{Sing}(\pi)$，我们就有

$$\int_\Omega \frac{\partial\varphi}{\partial x_1}(x)dx = \int_{\partial\Omega - \mathrm{Sing}(\pi)} \nu_1 \cdot \varphi d\sigma.$$

我们注意到 $\mathrm{Sing}(\pi) \subset \partial\Omega$ 未必是零测集（它在 π 的像下是零测集）。然而，当 $x \in \mathrm{Sing}(\pi)$ 时，存在 $w \in T_x\partial\Omega$，使得 $d\pi(x)(w) = 0$，由于 π 是投影映射，所以 w 必须形如 $(w_1, 0, \cdots, 0)$。根据 $\nu \perp w$，这表明 $\nu_1(x) = 0$。也即是说

$$\nu_1\big|_{\mathrm{Sing}(\pi)} \equiv 0.$$

所以，

$$\int_{\partial\Omega - \mathrm{Sing}(\pi)} \nu_1 \cdot \varphi d\sigma = \int_{\partial\Omega} \nu_1 \cdot \varphi d\sigma.$$

这就完成了 Stokes 公式的证明。

注记. *Stokes 公式第一个证明更分析，我们设法把问题转化为对支撑集很小的函数来证明，把 n 维的 Stokes 公式转化为一个方块上的问题（乘积机构），从而可以降低维数；第二个集合证明具有很强的微分拓扑的味道，直接将问题转化为 1 维的 Newton-Leibniz 公式。*

48.1 作业：曲面曲线积分的计算

清华大学 19-20 春季学期，数学分析二，作业 8

请用 A4 大小的纸张**正反面**用钢笔，签字笔或者圆珠笔书写，并注明自己的姓名，年级和作业的**总页数**。除定理公式所涉及的人名之外，请使用**中文**。本次作业的提交时间和地点为 **2 月 85 日**上午的课堂上，逾期视作零分。

课堂补充

A1) 给定 σ-有限的测度空间 $(X,\mathcal{A},\mu)$，正可测函数 ρ 几乎处处有界，我们课上证明了 $\nu=\rho\mu$ 是 $(X,\mathcal{A})$ 上 σ-有限的测度。f 是 $(X,\mathcal{A},\mu)$ 上的可测函数。证明，$f\rho\in L^1(X,\mathcal{A},\mu)$ 当且仅当 $f\in L^1(X,\mathcal{A},\nu)$（$L^1$ 表示它可积分）。进一步，在此情形下，我们有

$$\int_X f(x)d\nu(x)=\int_X f(x)\rho(x)d\mu(x).$$

（提示：证明遵循标准操作的步骤：

1) 对 $f(x)=\mathbf{1}_A$ 验证，其中 $A\in\mathcal{A}$）。

2) 对阶梯函数验证。

3) 通过取上升的简单函数逼近正函数 f 对正函数验证。

4) 对于一般的函数，将函数的实部和虚部拆成正负部分并利用线性完成证明。

）

A2) 我们在 $\mathbb{R}^2$ 上给 Lebesgue 测度 m_2。考虑投影映射

$$\pi_1:\mathbb{R}^2\to\mathbb{R},(x,y)\mapsto x.$$

证明，$\mathbb{R}^1$ 上的 Lebesgue 测度 Φ_*m_2 不是 σ-有限的，

A3) 假设 $n\geqslant 2$，$f\in C^\infty(\mathbb{R}^{n-1})$，我们用如下的方式来参数化其图像 Γ_f：

$$\Phi:\mathbb{R}^{n-1}\to\mathbb{R}^n,\quad x\mapsto(x,f(x)).$$

证明，

$$\det\left({}^t\mathrm{Jac}(\Phi)\cdot\mathrm{Jac}(\Phi)\right)=1+|\nabla f|^2.$$

A4)（换元积分公式，条件更弱）如果 $\Phi:\Omega_1\to\Omega_2$ 是同胚（即该映射及其逆都是连续的）并且 Φ 是连续可微的，对任意的可积函数 $f(y)\in L^1(\Omega_2,dy)$，我们有

$$\int_{\Omega_2}f(y)dy=\int_{\Omega_1}\big(f\circ\Phi\big)(x)|J_\Phi(x)|dx.$$

A5) 假设 $M \subset \mathbb{R}^n$ 是 $n-1$-维的子流形，考虑投影映射

$$\pi : M \to \mathbb{R}^{n-1}, \ \ (x_1, \cdots, x_{n-1}, x_n) \mapsto (x_1, \cdots, x_{n-1}).$$

令

$$\mathrm{Sing}(\pi) = \big\{x \in M \big| \mathrm{rank} d\pi(x) < n-1\big\}.$$

证明，$\pi(\mathrm{Sing}(\pi)) \subset \mathbb{R}^{n-1}$ 是零测集（用 $\mathbb{R}^{n-1}$ 上的 Lebesgue 测度来看）。对于 $x' = (x_1, \cdots, x_{n-1}) \in \mathbb{R}^{n-1}$，我们定义直线

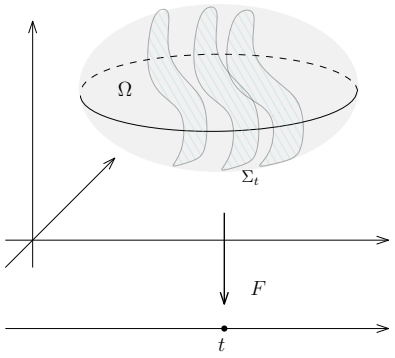
$$L_{x'} = \big\{(x_1, \cdots, x_{n-1}, t) \big| t \in \mathbb{R}\big\} = \pi^{-1}(x') \subset \mathbb{R}^n.$$

证明，

$$\pi(\mathrm{Sing}(\pi)) = \big\{x' \in \mathbb{R}^{n-1} \big| \text{直线} L_{x'} \text{ 与 } M \text{ 相切}\big\}.$$

其中，$L_{x'}$ 与 M 相切指的是存在 $p \in M$，使得 $p \in L_{x'}$ 并且 $T_p L_{x'} \subset T_p M$。

A6)（一个“扭曲”的 Fubini）$\Omega \subset \mathbb{R}^n$ 是开集，$F : U \to \mathbb{R}$ 是光滑映射。假设对任意的 $x \in U$，$dF(x) \neq 0$（从而，$\Sigma_t = F^{-1}(t)$ 是余 1 维子流形）。



那么，对任意的 U 上的可积函数 f，我们有

$$\int_\Omega f(x)dx = \int_{\mathbb{R}} \left(\int_{\Sigma_t} \frac{f}{|\nabla F|} d\sigma_t \right) dt,$$

其中，$d\sigma_t$ 是 Σ_t 上的子流形测度。

积分的计算（换元积分公式）

我们有比较常用的坐标变换（主要是前两种）：

1. 极坐标 $(r, \varphi) \mapsto (x, y) = (r\cos\varphi, r\sin\varphi)$；
2. 球坐标 $(r, \theta, \varphi) \mapsto (x, y, z) = (r\sin\theta\cos\varphi, r\sin\theta\sin\varphi, r\cos\theta)$；

3. 双曲坐标 (Rindler 坐标) $(t,x)\mapsto(T,X)=(x\sinh t, x\cosh t)$, 其中 $\sinh t=\dfrac{e^t-e^{-t}}{2}, \cosh t=\dfrac{e^t+e^{-t}}{2}$;

4. 抛物坐标 $(u,v)\mapsto(x,y)=(u^2-v^2,2uv)$;

5. 椭圆坐标 $(u,v)\mapsto(x,y)=(c\cosh u\cos v, c\sinh u, \sin v)$, $c>0$ 是常数。

B1) 计算体积或者面积（如果需要对参数分类讨论，那么请至少处理一种情形）

(a) 双纽线 $(x^2+y^2)^2=2a^2(x^2-y^2)$ 围成的面积。

(b) 曲线 $(x^3+y^3)^2=x^2+y^2$ 的内部区域面积。

(c) 曲线 $\frac{x^3}{a^3}+\frac{y^3}{b^3}=\frac{x^2}{h^2}+\frac{y^2}{k^2}$ 与坐标轴 $x=0, y=0$ 围成的面积。

(d) 曲线 $(2x+y)^4=x^2-2y^2$ 围成的面积。

(e) 曲线 $(\frac{x}{a}+\frac{y}{b})^5=\frac{x^2y^2}{c^4}$ $(x>0,y>0)$ 围成的面积.

(f) 曲面 $(x^2+y^2+z^2)^3=\frac{a^6z^2}{x^2+y^2}$ 所围立体图形的体积。

(g) 曲面 $(\frac{x^2}{a^2}+\frac{y^2}{b^2})^2+\frac{z^4}{c^4}=1$ 内部的体积。

(h) 曲面 $(\frac{x}{a}+\frac{y}{b}+\frac{z}{c})^4=\frac{xyz}{abc}$ $(x,y,z>0)$ 所围立体图形的体积。

(i) $(\frac{x}{a}+\frac{y}{b})^2+(\frac{z}{c})^2=1$ 所围立体图形的体积。

(j) 曲面 $z=x^2+y^2, z=2(x^2+y^2), xy=a^2, xy=2a^2, x=2y, 2x=y$ $(x>0,y>0)$ 所围立体图形的体积。

B2) 设物体的质量由密度函数 ρ 给出，那么在区域 V 内的总质量为 $\int_V \rho(x)dx$，V 的质心坐标 $(\bar{x}_1,\cdots,\bar{x}_n)$ 由下式确定:

$$\bar{x}_i=\frac{\int_V x_i\rho(x)dx}{\int_V\rho(x)dx}.$$

(a) 计算球 $x^2+y^2+z^2\leqslant 2az$ 的质量及质心坐标，其中密度函数 $\rho(x,y,z)=\frac{k}{\sqrt{x^2+y^2+z^2}}$, $k>0$ 是常数。

(b) 求由抛物面 $x^2+y^2=2az$ 及球面 $x^2+y^2+z^2=3a^2$ 围成的均质（$\rho\equiv 1$）物体的质心坐标。

B3) 假设 ρ 是 $\mathbb{R}^3$ 上非负可积的函数（如果一个物体的质量函数是 ρ 的话），它所对应的引力势能 $\Phi(x)$ 和引力场 $G(x)$ 由下式给出:

$$\Phi(x)=-\int_{\mathbb{R}^3}\frac{\rho(y)}{|x-y|}dy,\quad G(x)=-\nabla\Phi(x)$$

我们假设区域 $V \subset \mathbb{R}^3$ 是开区域，$\rho_0 > 0$ 是常数，

$$\rho(x) = \begin{cases} \rho_0, & x \in V \\ 0, & x \in \mathbb{R}^3 \setminus V \end{cases}$$

试计算如下情形下的 $\Phi(x)$ 和 $G(x)$：

(a) $V = \{x \in \mathbb{R}^3 : |x|^2 < R^2\}$ 是球体。

(b) $V = \{x = (x_1, x_2, x_3) \in \mathbb{R}^3 : x_1^2 + x_2^2 < R^2, 0 < x_3 < h\}$ 是圆柱体。

(c) $V = \{x = (x_1, x_2, x_3) \in \mathbb{R}^3 : x_3^2 > \frac{h^2}{R^2}(x_1^2 + x_2^2), 0 < x_3 < h\}$ 是圆锥体。

B4) (Newton 壳层 (shell) 定理：关于球对称密度分布的引力）假设

$$\rho(x) = \begin{cases} f(|x|), & r < |x| < R; \\ 0, & |x| > R \text{ 或 } |x| < r. \end{cases}$$

其中 f 是在 $[r, R])$ 上定义的非负可积函数。证明:

(a) 对球外一点 ($|x| > R$) 产生的引力与球的所有质量集中于球心的质点产生的引力一样。

(b) 球壳对球壳内任何一点 ($|x| < r$) 产生的引力 (合力) 为零。

曲面／曲线／子流形上的积分计算

C1) 计算下列的曲线积分：

(a) 计算 $\mathbb{R}^3$ 中曲线 C 的长度，其中参数曲线 C 由 $t \mapsto (a\cos(t), b\sin(t), bt)$ 给出，其中 $t \in [0, 2\pi]$；计算 $\int_C \frac{z^2}{x^+y^2} ds$，其中 ds 为 C 上的子流形测度。

(b) 假设 C 是 $\mathbb{R}^3$ 中 $x^2 + y^2 + z^2 = a^2$ 与 $x + y + z = 0$ 所截出的曲线，试计算 $\int_C xyds$。

(c) C 是参数曲线 $\big(a(t - \sin t), a(1 - \cos t)\big)$，其中 $t \in [0, 2\pi]$。计算积分 $\int_C y^2 ds$。

(d) 平面上的曲线 C 在极坐标系下的方程为 $r = f(\theta)$，其中 f 是连续可微的，$\alpha\theta\beta$。证明，C 的长度为

$$\int_\alpha^\beta \sqrt{|f(\theta)|^2 + |f'(\theta)|^2} d\theta.$$

C2) 计算下列的曲面积分：

(a) 试计算 $x^2 + y^2 + z^2 = a^2$ 与 $x^2 + y^2 = ax$ 相交得到的曲线在球面上围出的面积。

(b) 试计算参数曲面 Σ 的面积，其中 Σ 如下定义：

$$[0, 2\pi] \times [0, 2\pi] \to \mathbb{R}^3, \quad (\vartheta, \varphi) \mapsto \big((b + a\cos\vartheta)\cos\varphi, (b + a\sin\vartheta)\sin\varphi, a\sin\vartheta\big),$$

其中，我们要求 $0 < a < b$。

(c) C 是参数曲线 $\big(a(t-\sin t), a(1-\cos t)\big)$，其中 $t \in [0, 2\pi]$。将 C 在 $\mathbb{R}^3$ 中绕着 x 轴旋转一周得到的曲面是 Σ，试计算 Σ 的面积。

(d) C 是平面上的抛物线段 $x = y^2$，其中 $x \in [0, a]$。将 C 在 $\mathbb{R}^3$ 中绕着 x 轴旋转一周得到的曲面是 Σ，试计算 Σ 的面积。

(e) Σ 是锥面 $z^2 = x^2 + y^2$ 被柱面 $x^2 + y^2 = 2ax$ 所截下的部分，计算曲面积分 $\displaystyle\int_\Sigma x^2y^2 + y^2z^2 + z^2x^2 d\sigma$.

(f) Σ 是 $\mathbb{R}^3$ 中的柱面 $x^2 + y^2 = a^2$，其中 $z \in [0, h]$，计算曲面积分 $\displaystyle\int_\Sigma x^4 + y^4 d\sigma$.

(g) Σ 是 $\mathbb{R}^3$ 中的半球面 $x^2 + y^2 + z^2 = 1$，其中 $z \geqslant 0$，计算曲面积分 $\displaystyle\int_\Sigma x + y + z d\sigma$.

(h) Σ 是 $\mathbb{R}^3$ 中的球面 $x^2 + y^2 + z^2 = 1$，计算曲面积分 $\displaystyle\int_\Sigma x^2 d\sigma$.

(i) Σ 是 $\mathbb{R}^3$ 中的球面上小帽子 $x^2 + y^2 + z^2 = 1$，$z \geqslant h > 0$，计算曲面积分 $\displaystyle\int_\Sigma \frac{1}{z} d\sigma$.

(j) Σ 是 $\mathbb{R}^3$ 中的参数曲面：$(u, v) \mapsto \big(u\cos v, u\sin v, v\big)$，其中 $(u, v) \in [0, a] \times [0, 2\pi]$。计算曲面积分 $\displaystyle\int_\Sigma z d\sigma$.

(k) Σ 是 $\mathbb{R}^3$ 中的球面 $x^2 + y^2 + z^2 = 1$，f 是 $\mathbb{R}$ 上的连续函数。证明，

$$\int_\Sigma f(ax + by + cz) d\sigma = 2\pi \int_{-1}^1 f((a^2 + b^2 + c^2)^{\frac{1}{2}} t) dt.$$

In my free time I do differential and integral calculus.

—- Carl Marx

48.2 习题课：Riemann 积分的定义 1

清华大学 19-20 春季学期，数学分析二，第六次习题课内容

传统 Riemann 积分的定义：第一部分

我们在课上从测度的观点定义了积分，这次习题目的是尽可能还原历史 Riemann 积分的（严格）定义。我们不假设任何的测度理论。同学们与课堂内容作比较，就能体会为什么抽象积分实际上更简单更有效率。

———————————-

C'est pour la résolution de ces problèmes, et non par amour des complications, que j'ai introduit dans ce Livre une définition de l'intégrale plus générale que celle de Riemann et comprenant celle-ci comme cas particulier.

（此书新引入的积分定义比 Riemann 积分更广泛并且把它作为一个特例，这是为了要解决存在的那些问题，而不是因为我们更偏爱复杂的对象。）

————- Henri Lebesgue, *Leçons sur l'intégration et la recherche des fonctions primitives. Préface.*

矩形上的 Riemann 积分

所谓 $\mathbb{R}^n$ 上的一个**矩形**P 指的是形如 $[a_1,b_1]\times\cdots[a_n,b_n]$ 的集合（都是有限区间）。该矩形 P 的内部指的是集合 $(a_1,b_1)\times\cdots(a_n,b_n)$，用 $\mathring{P}$ 表示；它的边界定义为 $\partial P=P-\mathring{P}$。我们总假设矩形的内部是非空的。

对每个 $1\leqslant i\leqslant n$，给定 $[a_i,b_i]$ 的一个分划 $\sigma_i\in\mathcal{S}([a_i,b_i])$，即给定一串数

$$a_i=a_{i,0}<a_{i,1}<\cdots<a_{i,n_i-1}<a_{i,n_i}=b_i.$$

我们称

$$P_{\sigma_1,\cdots,\sigma_n}=\left\{[a_{\ell_1},a_{\ell_1+1}]\times\cdots\times[a_{\ell_n},a_{\ell_n+1}]\middle|0\leqslant\ell_1\leqslant n_1-1,\cdots,0\leqslant\ell_n\leqslant n_n-1\right\}.$$

是 P 的一个（由 $(\sigma_1,\cdots,\sigma_n)$ 所定义的）**分划**。很显然，P 的一个分划是一些更小的矩形的并。给定 P 的两个分划 $P_{\sigma_1,\cdots,\sigma_n}$ 和 $P_{\sigma_1',\cdots,\sigma_n'}$，如果对于每个 $1\leqslant i\leqslant n$，我们都有 $\sigma_i\prec\sigma_i'$（请参考上学期关于分划加细的概念），我们就称 $P_{\sigma_1,\cdots,\sigma_n}$ 是 $P_{\sigma_1',\cdots,\sigma_n'}$ 的**加细**并记作 $P_{\sigma_1,\cdots,\sigma_n}\prec P_{\sigma_1',\cdots,\sigma_n'}$。我们把 P 的所有分划的全体用 $\mathcal{S}(P)$ 来表示。

如果不加任何说明，下面的字母 P 代表一个给定的矩形。

R1) 给定矩形 $P=[a_1,b_1]\times\cdots[a_n,b_n]$，它的体积定义为

$$m(P)=\prod_{1\leqslant i\leqslant n}(b_i-a_i).$$

任给 P 的分划 $P_{\sigma_1,\cdots,\sigma_n}$，证明，

$$m(P)=\sum_{Q\in P_{\sigma_1,\cdots,\sigma_n}} m(Q).$$

R2) 给定 $A\subset\mathbb{R}^n$，如果存在有限个矩形 $P_1,\cdots,P_m$，使得 $A=\bigcup\limits_{i\leqslant m}P_i$ 并且对于 $i\neq j$，$\mathring{P}_i\cap\mathring{P}_j=\emptyset$，我们就称 A 是**可铺的**。对于可铺集 A，我们定义

$$m(A)=\sum_{i\leqslant m}m(P_i).$$

证明，$m(A)$ 的选取不依赖于上述 $\{P_i\}_{i\leqslant m}$ 的选取。

（我们在乘积空间中构造矩形上的测度，也有类似的步骤）

R3) 证明，有限个可铺集的交和并都是可铺集。

（这一点与 σ-代数或者代数类似）

R4) A 和 B 是可铺集。证明，

$$\begin{aligned}m(A\cap B)+m(A\cup B)&=m(A)+m(B),\\ m(A-\mathring{B})&=m(A)-m(A\cap B),\end{aligned}$$

其中，你需要先说明 $A-\mathring{B}$ 也是可铺集。

（这一点与测度或者加性函数类似）

R5) 任给有界函数（可以在赋范线性空间中取值）

$$f:P\to\mathbb{C},$$

如果存在 P 的分划 $P_{\sigma_1,\cdots,\sigma_n}$，使得对任意的 $Q\in P_{\sigma_1,\cdots,\sigma_n}$，函数 f 在 Q 的内部 $\mathring{Q}$ 上的限制是常数，我们就称 f 是 P 上的**阶梯函数**或者**简单函数**。P 上阶梯函数的全体记作 $\mathcal{E}(P)$。

任给 $f\in\mathcal{E}(P)$，满足阶梯函数的定义中的分划 $P_{\sigma_1,\cdots,\sigma_n}$ 称作是一个*和 f 相容的分划*。证明，任意一个和 f 相容的分划，它的加细也是和 f 相容的。进一步证明，$\mathcal{E}(P)$ 是 $\mathbb{C}$-线性空间。

R6) 假设 P 的分划 $P_{\sigma_1,\cdots,\sigma_n}$ 与 $f\in\mathcal{E}(P)$ 相容，对于 $Q\in P_{\sigma_1,\cdots,\sigma_n}$，按照定义，$f$ 在 $\mathring{Q}$ 上的取值是常数，我们将这个常数记成 $f(\mathring{Q})$。定义

$$\int_P f=\sum_{Q\in P_{\sigma_1,\cdots,\sigma_n}} f(\mathring{Q})m(Q).$$

证明，$\int_P f$ 的定义不依赖于与之相容的分划 $P_{\sigma_1,\cdots,\sigma_n}$ 的选取。

R7) 证明，映射

$$\int_P : \mathcal{E}(P) \to \mathbb{C}$$

是 $\mathbb{C}$-线性映射。

R8) 证明，对任意的 $f \in \mathcal{E}(P)$，我们有

$$\Big|\int_P f\Big| \leqslant \int_P |f|.$$

如果 $f, g \in \mathcal{E}(P)$ 是实值的并且对任意的 $x \in P$，都有 $f(x) \leqslant g(x)$，证明，

$$\int_P f \leqslant \int_P g.$$

R9) 任给简单函数 $f \in \mathcal{E}(P)$，任给 P 的分划 $P_{\sigma_1,\cdots,\sigma_n}$（未必与 f 相容）。证明，对每个 $Q \in P_{\sigma_1,\cdots,\sigma_n}$，$f$ 在 Q 上的限制 $f\big|_Q \in \mathcal{E}(Q)$ 并且

$$\int_P f = \sum_{Q \in P_{\sigma_1,\cdots,\sigma_n}} \int_Q f|_Q.$$

R10)（矩形区域上 Riemann 积分的定义）$f : P \to \mathbb{C}$ 是映射（可以在赋范线性空间中取值）证明，下面的两个命题是等价的：

1) 对任意 $\varepsilon > 0$，存在 $\Phi_\varepsilon, \Psi_\varepsilon \in \mathcal{E}(P)$，使得对任意的 $x \in P$，有

$$\big|f(x) - \Phi_\varepsilon(x)\big| \leqslant \Psi_\varepsilon(x)$$

并且

$$\int_P \Psi_\varepsilon \leqslant \varepsilon;$$

2) 存在函数序列 $\{\phi_k\}_{k\geqslant 1}, \{\psi_k\}_{k\geqslant 1} \subset \mathcal{E}(P)$，使得对任意的 $x \in P$，我们都有

$$\big|f(x) - \phi_k(x)\big| \leqslant \psi_k(x)$$

并且

$$\lim_{k\to\infty} \int_P \psi_k = 0.$$

如果一个函数 f 满足上述条件之一，我们就称 f 在 P 上是 **Riemann 可积的**函数。我们用 $\mathcal{R}(P)$ 表示 P 上所有 Riemann 可积的函数。

（请与 1 维 Riemann 积分的定义做比较）

R11) 证明，如果 $f \in \mathcal{R}(P)$，那么 f 是有界函数。

R12) 证明，$\mathcal{R}(P)$ 是 $\mathbb{C}$-线性空间。进一步证明，如果 $f, g \in \mathcal{R}(P)$，那么它们的乘积 $f \cdot g \in \mathcal{R}(P)$。

（如果 f, g 在 Lebesgue 的意义下可积分，我们不一定有 $f \cdot g$ 还是 Lebesgue 意义下可积的）

R13) 假设 $f \in \mathcal{R}(P)$，我们任选 Riemann 可积函数的定义（R19）2) 中的序列 $\{\phi_k\}_{k\geqslant 1}, \{\psi_k\}_{k\geqslant 1} \subset \mathcal{E}(P)$ 并定义

$$\int_P f = \lim_{k\to\infty} \int_P \phi_k.$$

证明，上述极限存在并且不依赖于序列的选取。从而，我们定义了积分映射

$$\int_P : \mathcal{R}(P) \to \mathbb{C}.$$

R14) 证明，映射 $\int_P : \mathcal{R}(P) \to \mathbb{C}$ 是 $\mathbb{C}$-线性映射。

（这里线性的证明比抽象积分简单，原因是我们可以用序列逼近函数，实际上，在抽象积分的理论中，我们要先建立 Beppo Levi 定理，从而也可以用简单函数列的积分逼近函数的积分）

R15) 证明，对任意的 $f \in \mathcal{R}(P)$，我们有

$$\Big|\int_P f\Big| \leqslant \int_P |f|.$$

如果 $f, g \in \mathcal{E}(P)$ 是实值的并且对任意的 $x \in P$，都有 $f(x) \leqslant g(x)$，那么

$$\int_P f \leqslant \int_P g.$$

R16)（区间可加性）任给函数 $f : P \to \mathbb{C}$ 和 P 的分划 $P_{\sigma_1,\cdots,\sigma_n}$。证明，如下两个叙述是等价的：

- $f \in \mathcal{R}(P)$；
- 对每个 $Q \in P_{\sigma_1,\cdots,\sigma_n}$，$f$ 在 Q 上的限制 $f\big|_Q \in \mathcal{R}(Q)$。

如果上面之一成立，证明，

$$\int_P f = \sum_{Q\in P_{\sigma_1,\cdots,\sigma_n}} \int_Q f\big|_Q.$$

R17)（Riemann 和）任给 P 的分划 $\sigma = P_{\sigma_1,\cdots,\sigma_n}$，其中对任意的 $i \leqslant n$，$\sigma_i \in \mathcal{S}([a_i, b_i])$ 并且对应着

$$a_i = a_{i,0} < a_{i,1} < \cdots < a_{i,n_i-1} < a_{i,n_i} = b_i.$$

我们定义分划的 σ步长$|\sigma|$ 为

$$|\sigma| = \sup_{\substack{1\leqslant i\leqslant n,\\ 1\leqslant j\leqslant n_i}} |a_{i,j-1} - a_{i,j}|.$$

任给 P 的分划 σ，我们在每个 $Q \in \sigma$ 中任意指定一个点 $\xi_Q \in Q$，据此，我们定义 **Riemann 和**：

$$S(f, \sigma, (\xi_Q)_{Q\in\sigma}) = \sum_{Q\in\sigma} m(Q) f(\xi_Q).$$

假设 $\{\sigma^{(\ell)}\}_{\ell\geqslant 1}$ 是 P 的一列分划并且对每个 $\sigma^{(\ell)}$ 都指定了一族 $(\xi_Q)_{Q\in\sigma^{(\ell)}}$。我们要求对任意的 ℓ，$\sigma^{(\ell+1)}$ 是 $\sigma^{(\ell)}$ 的加细并且

$$\lim_{\ell\to\infty}|\sigma^{(\ell)}|=0.$$

证明，如果 $f\in\mathcal{R}(P)$，那么

$$\lim_{\ell\to\infty}S(f,\sigma^{(\ell)},(\xi_Q)_{Q\in\sigma^{(\ell)}})=\int_P f.$$

（传统上，更多的教科书用这种 Riemann 和或者下面的 Darboux 上下和来定义积分）

R18) 假设 $f:P\to\mathbb{R}$ 是有界函数，我们定义

$$\overline{\mathcal{E}_f}(P)=\{\varphi\in\mathcal{E}(P)|\varphi(x)\geqslant f(x)\},$$
$$\underline{\mathcal{E}_f}(P)=\{\varphi\in\mathcal{E}(P)|\varphi(x)\leqslant f(x)\}.$$

证明，这两个集合是非空的。据此，我们定义

$$\overline{\int_P} f=\inf_{\varphi\in\overline{\mathcal{E}_f}(P)}\int_P\varphi,$$
$$\underline{\int_P} f=\sup_{\varphi\in\underline{\mathcal{E}_f}(P)}\int_P\varphi.$$

证明，上面的两个数值是有限的（称作是有界函数 f 的**上积分**和**下积分**）并且满足

$$\overline{\int_P} f\geqslant\underline{\int_P} f.$$

R19) 给定有界实值函数 $f:P\to\mathbb{R}$，σ 是 P 的一个分划。对任意 $Q\in\sigma$，我们定义

$$M_Q=\sup_{x\in Q}f(x),\quad m_Q=\inf_{x\in Q}f(x)$$

. 我们令

$$\overline{S}(f;\sigma)=\sum_{Q\in\sigma}M_Q m(Q),$$
$$\underline{S}(f;\sigma)=\sum_{Q\in\sigma}m_Q m(Q).$$

上述两个值被称为是 f 对于分划 σ 的 **Darboux 上和**和 **Darboux 下和**。

证明，对于 P 的任两个分划 σ 和 σ'，我们有

$$\underline{S}(f;\sigma)\leqslant\overline{S}(f;\sigma').$$

进一步证明，如果分划 $\sigma\prec\sigma'$，那么

$$\overline{S}(f;\sigma)\leqslant\overline{S}(f;\sigma'),\quad\underline{S}(f;\sigma)\geqslant\underline{S}(f;\sigma').$$

R20)（Darboux 上（下）和的下（上）极限是上（下）积分）假设 $f : P \to \mathbb{R}$ 是有界函数。定义

$$\overline{D}(f) = \inf_{\sigma \in \mathcal{S}(P)} \overline{S}(f;\sigma),$$
$$\underline{D}(f) = \sup_{\sigma \in \mathcal{S}(P)} \underline{S}(f;\sigma).$$

证明，

$$\overline{D}(f) = \overline{\int_Q} f, \quad \underline{D}(f) = \underline{\int_Q} f.$$

R21)（矩形上实值 Riemann 可积函数的刻画）给定实值函数 $f : Q \to \mathbb{R}$。证明，如下四条性质是等价的：

甲) $f \in \mathcal{R}(Q)$；

乙) 存在常数 $\mathbf{I}$，对任意的 $\varepsilon > 0$，存在 $\delta > 0$，使得对任意满足 $|\sigma| < \delta$ 的分划 σ，对任意的 $\xi_Q \in Q$，其中 $Q \in \sigma$，我们都有

$$\left|S(f,\sigma,(\xi_Q)_{Q\in\sigma}) - \mathbf{I}\right| < \varepsilon.$$

丙) f 是有界函数并且 $\underline{D}(f) = \overline{D}(f)$；

丁) f 是有界函数并且 $\underline{\int_P} f = \overline{\int_P} f$。

R22) 证明，如果 R21) 中某一条件成立，那么

$$\underline{D}(f) = \overline{D}(f) = \underline{\int_P} f = \overline{\int_P} f = \int_P f.$$

R23) 给定有界实值函数 $f, g : P \to \mathbb{R}$，我们假设它们是 Riemann 可积德。证明，P 上的函数

$$\max(f,g)(x) = \max(f(x),g(x)), \quad \min(f,g)(x) = \min(f(x),g(x))$$

是 Riemann 可积的并且

$$\int_P \min(f,g) \leqslant \min(\int_P f, \int_P g), \ \max(\int_P f, \int_P g) \leqslant \int_P \max(f,g).$$

R24) 假设正函数 $f : P \to \mathbb{R}_{\geqslant 0}$ 是 Riemann 可积的。如果

$$\int_P f = 0,$$

证明，对任意的 $x^* \in P$，f 在 x^* 处连续，我们都有 $f(x^*) = 0$。

R25) 证明，P 上的连续函数是 Riemann 可积的，即 $C(P;\mathbb{C}) \subset \mathcal{R}(P)$。

R26) 假设 $f, g : P \to \mathbb{R}$ 是连续函数并且对任意的 $x \in P$，我们都有 $f(x) \leqslant g(x)$。证明，如果

$$\int_P f = \int_P g,$$

那么，$f \equiv g$。

R27) 证明 Cauchy-Shwarz 不等式：对任意的 $f, g \in \mathcal{R}(P)$，我们有

$$\Big|\int_P f \cdot \overline{g}\Big| \leqslant \Big(\int_P |f|^2\Big)^{\frac{1}{2}} \Big(\int_P |g|^2\Big)^{\frac{1}{2}}.$$

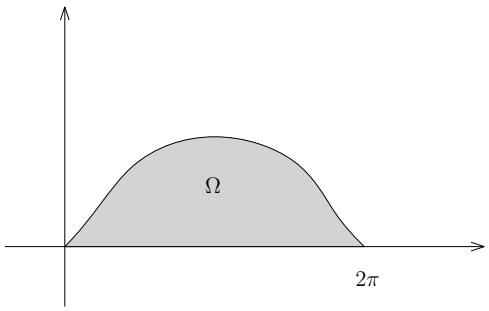
49 子流形上的积分的计算，第一／二型子流形／曲线／曲面积分，向量场的运算，散度定理，Green 公式，Gauss-Ostrogradsky 公式，散度的几何/物理解释

二零二零年四月二十七日，星期一，晴

我们首先一起计算一个积分：$\int_\Omega y^2 dxdy$，其中 Ω 是 $\mathbb{R}^2$ 上面由曲线

$$\gamma : [0, 2\pi] \to \mathbb{R}^2, \quad t \mapsto \big(a(t - \sin(t)), a(1 - \cos(t))\big)$$

与 $y = 0$ 所围成的区域。



这显然是某个函数 $y = f(x)$ 的图像下的部分（与 x 轴所围），其中 $x \in [0, 2\pi]$。所以，我们可以尝试用 Fubini 公式：

$$\begin{aligned}\int_\Omega y^2 dxdy &= \int_0^{2\pi} \big(\int_0^{f(x)} y^2 dy\big) dx \\ &= \int_0^{2\pi} \frac{1}{3} f(x)^3 dx.\end{aligned}$$

当然，我们可能需要强行决定 $f(x)$ 具体的表达式。

所以，我们很自然地想研究 y 与 x 之间的（隐函数）关系：

$$x = a \arccos(1 - \frac{y}{a}) - \sqrt{y(2a - y)}.$$

这个只需要用 y 来表达 t，然后带入 x 的表达式即可。利用这个表达式，一种计算积分的方式就是把 x 换元成 y，从而对 y 积分。当然，这个操作启发我们可以用 Fubini 公式先对 x 积分，这样再用 $x = x(y)$ 来描述函数的图像就会简单很多。

还有一种看法是用参数曲线来描述积分

$$\int_0^{2\pi} \big(\int_0^{f(x)} y^2 dy\big) dx = \int_0^{2\pi} \frac{1}{3} f(x)^3 dx.$$

我们将这个积分写的形式化一些：

$$\int_\Omega y^2 dxdy = \int_0^{2\pi} \varphi(x, f(x)) dx,$$

其中 $y=f(x)$ 就是我们要找的函数图像，$\dfrac{\partial}{\partial y}\varphi(x,y)=y^2$，所以我们可以取 $\varphi(x,y)=\dfrac{1}{3}y^3$（注意到这个函数的选取比较随意）。此时，我们可以用曲线的参数表达，因为我们通过 $(x,y)=(x,f(x))$ 的替换已经假设这个点生活在 f 的图像上。所以

$$\begin{aligned}\int_\Omega y^2dxdy&=\int_0^{2\pi}\varphi(x,f(x))dx\\&=\int_0^{2\pi}\varphi(x(t),f(x(t)))x'(t)dt\\&=\int_0^{2\pi}\varphi(x(t),y(t))x'(t)dt.\end{aligned}$$

代入参数化，我们得到

$$\int_\Omega y^2dxdy=\int_0^{2\pi}\frac{1}{3}a^3(1-\cos(t))^3a(1-\cos(t))dt.$$

这样，就得到了一个 1 维的积分，所以可以进行计算了。

如果我们坚持不用具体的参数表达式来写，我们实际上就是用了如下简单的计算：

$$\begin{aligned}\int_\Omega y^2dxdy&=\int_0^{2\pi}\frac{1}{3}f(x)^3dx\\&=\int_0^{2\pi}\frac{1}{3}y(t)^3x'(t)dt.\end{aligned}$$

这些计算实际上已经隐藏在 Stokes 公式之中，我们把 y^2 看作是 $\dfrac{\partial\varphi}{\partial y}$，并且

$$\begin{aligned}\int_\Omega y^2dxdy&=\int_0^{2\pi}\varphi(x(t),y(t))\frac{x'(t)}{|\gamma'(t)|}\underbrace{|\gamma'(t)|dt}_{d\sigma}\\&=\int_{\partial\Omega}\varphi\nu_2d\sigma.\end{aligned}$$

这恰好就是在 Stokes 公式的证明中所出现的步骤。此时，Ω 的下边界对积分没有贡献，因为 ν_2，这个边界落在向 y 轴投影的奇异点集中。

上述的讨论还建议我们实际上可以用 Stokes 公式来计算，想法是直接把积分化到边界上，从而用曲线上的积分计算：

$$\int_D y^2dxdy=\int_D\operatorname{div}(0,\frac{1}{3}y^3)dxdy.$$

我们可以参考后面的散度定理。

注记. 上周的 *Stokes* 公式的证明是值得研究的，核心的想法是把边界上的积分分解为一些函数图像上的积分，为了搞清楚证明，假设 Ω 为 $\mathbb{R}^n$ 的标准球面，然后就这个例子把证明的每一步搞清楚是很有帮助的。

我们先澄清一些出现在各种数学或物理教材文献中的各种积分的记号。我们总是假设 $\Omega \subset \mathbb{R}^n$ 是给定的一个区域（很多场合下，它是一个有界带边的光滑区域）。

假设 $M \subset \mathbb{R}^n$ 是子流形，给定函数 $f : M \to \mathbb{C}$，人们通常把子流形上的积分（如果可积的话）

$$\int_M f d\sigma$$

称作是**第一型子流形积分**。当 $\dim M = 1$ 时（通常 $n = 2$ 或 3），这个积分被称作是**第一型曲线积分**；当 $\dim M = 2$ 时（通常 $n = 3$），这个积分被称作是**第一型曲面积分**。这种分类可能不是很有意思，更多是沿用了历史上的名称。

第二型曲线积分

现在假设 X 是区域 $U \subset \mathbb{R}^n$ 上的光滑向量场。

我们首先定义第二型曲线积分。假设 $C \subset \Omega$ 是一个 1 维子流形，我们它可以用光滑的参数化

$$\gamma : [a, b] \to \Omega$$

来表示。我们定义向量场 X 在 C 上的**第二型曲线积分**为

$$\int_\gamma X \cdot d\gamma := \int_a^b \langle X(\gamma(t)), \gamma`(t)\rangle dt.$$

在上面的表达式中，左边我们认为是形式的记号，右边 $\langle X(\gamma(t)), \gamma(t)\rangle$ 表示的是向量的内积。有时候（从微分流形的观点看更自然，我们或许在下个学期会严格地定义一般流形上对微分形式的积分），人们把这个积分也写成：

$$\int_\gamma X \cdot d\gamma := \int_\gamma X_1 dx_1 + \cdots + X_n dx_n.$$

用测度的观点来写（或者用第一型曲线积分来写），我们有

$$\int_\gamma X \cdot d\gamma = \int_\gamma \langle X(\gamma(t)), \frac{\gamma'(t)}{|\gamma'(t)|}\rangle d\sigma.$$

也就是说，被积函数是 X 与 γ 的单位切向量的内积。

所以，单独地研究第二型曲线积分并没有太大的意义，因为这就是在曲线上做积分，然而，这些表达式在物理学里有着它们特殊的含义，我们会在某些例子中展现，同学们在今后其他场合遇到第二型曲线积分只要来查一下这部分笔记把它翻译成第一型积分即可。

另外，假设 C 是闭曲线，即 $\gamma(a) = \gamma(b)$，我们还把这个积分写成

$$\oint_\gamma X \cdot d\gamma,$$

其中，积分号上的圈表明这个曲线是一个“圈”（封闭），我们把它还称作是 X 沿曲线 C 的**环路积分**，物理上把它称作是 X 沿曲线 C 的**环量**。

第二型（超）曲面积分

现在 $\Omega \subset \mathbb{R}^n$ 是有界带边光滑区域，X 是 Ω 上的光滑向量场，ν 是 $\partial\Omega$ 上的单位外法向量。我们定义

$$\int_{\partial\Omega} X \cdot d\vec{\sigma} = \int_{\partial\Omega} \langle X, \nu \rangle d\sigma$$

物理学上，我们把上面的量称作是 X 穿过 $\partial\Omega$ 的**通量**，目前我们可以简单地从字面上直观认为这个积分是 X 从里向外有多少" 流" 通过了 $\partial\Omega$（我们后面会定义精确地定义它的含义）。

通常，我们研究的区域在 $\mathbb{R}^3$ 中，此时，我们也把上面的积分写成

$$\iint_{\partial\Omega} X \cdot d\vec{\sigma}.$$

其中，两个积分符号代表着 $\partial\Omega$ 的维数是 2。

另外，由于 $\partial\Omega$ 是闭曲面，我们还把这个积分写成

$$\oiint_{\partial\Omega} X \cdot d\vec{\sigma}.$$

其中，积分号上的圈表明曲面是封闭的。

另外，文献中还有另一种表达式代表的也是第二型曲面积分（$\mathbb{R}^3$）的情形，它的表达式是

$$\iint_{\partial\Omega} P dydz + Q dxdz + R dxdy,$$

其中 P, Q, R 是 $\mathbb{R}^3$（或者 $\partial\Omega$）上的光滑函数。这个积分应该按照下面的方式解读：

$$X = P\frac{\partial}{\partial x} + Q\frac{\partial}{\partial y} + R\frac{\partial}{\partial Z} = (P, Q, R)$$

为 $\mathbb{R}^3$ 上的向量场，那么

$$\iint_{\partial\Omega} P dydz + Q dxdz + R dxdy := \iint_{\partial\Omega} X \cdot d\vec{\sigma}.$$

同样地，从微分流形的观点看这些记号更有意义，但是这绝对不会影响我们学习积分，所以我们完全没有必要去记住这些记号（只需要记住在哪里查找即可）。

$\mathbb{R}^n$ 上向量场的运算

我们在 $\mathbb{R}^n$ 上固定直角坐标系 $(x_1, \cdots, x_n)$。对于 $\Omega \subset \mathbb{R}^n$ 上的光滑向量场

$$X = \sum_{i=1}^{n} X_k \frac{\partial}{\partial x_k},$$

我们定义 X 的**旋度**$\nabla \wedge X$ 是在反对称矩阵中取值的映射（这实际上是一个 2-形式）

$$\omega : \Omega \to \mathbf{M}_n(\mathbb{R}), \quad x \mapsto (\omega_{ij})_{1 \leqslant i,j \leqslant n},$$

其中

$$\omega_{ij}(x)=\frac{\partial X_j}{\partial x_i}-\frac{\partial X_i}{\partial x_j}.$$

特别地，在 $n=3$ 时，由于 ω 只有三个分量，我们将 X 的旋度用 $\mathbb{R}^3$ 中的向量场表示（正确的定义需要用到 Hodge $*$-算子），记作

$$\operatorname{curl} X=\nabla\times X=\left(\frac{\partial X_3}{\partial x_2}-\frac{\partial X_2}{\partial x_3}\right)\frac{\partial}{\partial x_1}+\left(\frac{\partial X_1}{\partial x_3}-\frac{\partial X_3}{\partial x_1}\right)\frac{\partial}{\partial x_2}+\left(\frac{\partial X_2}{\partial x_1}-\frac{\partial X_1}{\partial x_2}\right)\frac{\partial}{\partial x_3}$$

当 $n=2$ 时，由于 ω 只有一个分量，我们将 X 的旋度用 $\mathbb{R}^2$ 中的函数表示，记作

$$\operatorname{curl} X=\frac{\partial X_1}{\partial x_2}-\frac{\partial X_2}{\partial x_1}$$

我们还定义向量场 X 的**散度**为：

$$\operatorname{div} X=\sum_{k=1}^{n}\frac{\partial X_k}{\partial x_k}.$$

这是一个函数。

散度和旋度的几何意义很难从它们的定义读出，我们需要 Stokes 公式的帮助来能正确理解这两个概念。但在此之前，我们先罗列出它们满足的一些代数性质，这将为后来的众多计算提供莫大的方便：

引理 321. *我们在 $\mathbb{R}^3$ 用直角坐标系。假设 X,Y 是 $\mathbb{R}^3$ 上的光滑向量场，φ 是 $\mathbb{R}^3$ 上实值光滑函数，我们用下面经典的记号：*

$$\nabla\cdot X=\operatorname{div}X,\quad (X\cdot\nabla)Y=\nabla_X Y=\sum_{i=1}^{3}X_i\frac{\partial Y_i}{\partial x_i},\quad \triangle\varphi=\sum_{i=1}^{3}\frac{\partial^2\varphi}{\partial x_i^2},$$

（其中，$\triangle$ 被称作是 Laplace 算子）那么，我们有

$$\begin{aligned}
&\nabla\times(\nabla\varphi)=0,\quad \nabla\cdot(\nabla\times X)=0,\quad \triangle\varphi=\nabla\cdot(\nabla\varphi),\\
&\nabla\times(\varphi X)=\varphi(\nabla\times X)+\nabla\varphi\times X,\\
&\nabla\cdot(X\times Y)=(\nabla\times X)\cdot Y-(\nabla\times Y)\cdot X,\\
&\nabla\times(X\times Y)=X(\nabla\cdot Y)-Y(\nabla\cdot X)+(Y\cdot\nabla)X-(X\cdot\nabla)Y.
\end{aligned}$$

这些等式的证明可以通过直接计算得到，我们把它们留作作业。

Stokes 公式的应用

我们首先给出 Stokes 公式的散度定理形式：

定理 322 (散度定理)**.** *给定有界光滑带边区域 $\Omega\subset\mathbb{R}^n$，我们用 ν 表示 $\partial\Omega$ 的单位外法向量，$d\sigma$ 为 $\partial\Omega$ 的子流形测度。那么，对任意 $\mathbb{R}^n$ 上 C^1 的向量场 X，我们有*

$$\int_\Omega \operatorname{div}X\ dx=\int_{\partial\Omega}\langle X,\nu\rangle\ d\sigma.$$

证明: 证明几乎是平凡的：假设 $X=(X_1,\cdots,X_n)$，那么，根据 Stokes 公式，对任意的 $i\leqslant n$，我们有

$$\int_\Omega \frac{\partial X_i}{\partial x_i}\,dx=\int_{\partial\Omega} X_i\nu_i\rangle\,d\sigma.$$

对 i 求和我们就得到了公式。 □

注记. 如果我们令 $X=(0,\cdots,0,\varphi,0,\cdots,0)$，其中除了第 i 个位置之外，X 的其余分量都为 0，那么，散度定理就给出了 *Stokes* 公式，这说明这两个公式是等价的。

当 $n=2$ 时，散度定理或者 Stokes 公式被称作是 Green 公式，它经常以第二型曲线积分的形式出现：

推论 323 (Green 公式)**.** 如果 $\Omega\subset\mathbb{R}^2$ 是有界光滑带边区域，那么对任意的 $\mathbb{R}^2$ 上的光滑（C^1）P 和 Q，我们有

$$\int_\Omega(\frac{\partial Q}{\partial x}-\frac{\partial P}{\partial y})dxdy=\oint_{\partial\Omega}Pdx+Qdy.$$

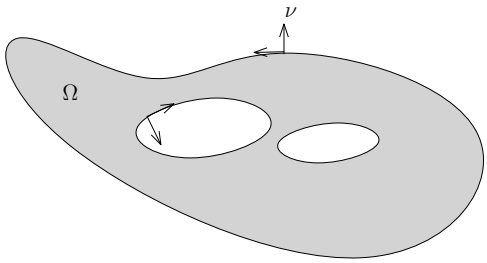
注记. 上面的公式隐含地用到了 $\partial\Omega$ 是有限个闭曲线的并，我们不打算在这里展开这一点。

证明: 我们只需要把上述公式翻译成我们熟悉的形式。我们对 $\partial\Omega$ 局部上进行参数化

$$\gamma:(-1,1)\to\mathbb{R}^2$$

. 我们利用 $\mathbb{R}^2$ 上的特点：如果 $\nu=(\nu_1,\nu_2)$ 是 $\partial\gamma$ 的外法向量，那么，我们总是可以选取参数化 $\gamma(t)$（或者 $\gamma(-t)$），使得 γ 的方向恰好是 ν 逆时针转动 90°，即

$$\frac{\gamma'(t)}{|\gamma(t)|}=(-\nu_2,\nu_1)$$



令 $X=(P,Q)$ 是 $\mathbb{R}^2$ 上的向量场，那么

$$\begin{aligned}\int_{\partial\Omega}Pdx+Qdy&=\int_{\partial\Omega}\langle X,\frac{\gamma'(t)}{|\gamma(t)|}\rangle d\sigma\\&=\int_{\partial\Omega}-P\nu_1+Q\nu_2 d\sigma.\end{aligned}$$

另外，根据散度定理，我们有

$$\begin{aligned}\int_\Omega (\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y})dxdy &= \int_\Omega (\frac{\partial Q}{\partial x_1} - \frac{\partial P}{\partial x_2})dx_1dx_2 \\ &= \int_\Omega \operatorname{div}(Q, -P)\, dx_1dx_2 \\ &= \int_{\partial\Omega} (Q, -P)\cdot(\nu_1, \nu_2)d\sigma.\end{aligned}$$

对比上下的式子，我们就得到了结论。 □

当 $n = 3$ 时，散度定理或者 Stokes 公式被称作是 Gauss-Ostrogradsky 公式，

推论 324 (Gauss-Ostrogradsky)**.** *如果 $\Omega \subset \mathbb{R}^3$ 是有界光滑带边区域，那么对任意的 $\mathbb{R}^3$ 上的光滑 (C^1) P，Q 和 R，我们有*

$$\oiint_{\partial\Omega} Pdydz + Qdxdz + Rdxdy = \int_\Omega \frac{\partial P}{\partial x} + \frac{\partial Q}{\partial y} + \frac{\partial R}{\partial z}dxdydz.$$

为了给出散度定理的几何／物理解释，我们需要研究和向量场相关的几何。我们要利用上学期第二十五课讲的关于常微分方程解的存在唯一性定理（Cauchy-Lipschitz）：

定理 325 (Cauchy-Lipschitz)**.** *给定完备的赋范线性空间 V（通常我们假设 $V = \mathbb{R}^n$），$\Omega \subset V$ 是开集，$I \subset \mathbb{R}$ 是开区间，$f : \Omega \times I \to V$ 是给定的连续函数（映射）。假设存在 $C > 0$，对任意的 $t \in I$，映射*

$$\Omega \to V, \quad x \mapsto f(x, t)$$

是 C-Lipschitz 映射，即对任意的 $x, y \in \Omega$，有如下估计

$$|f(x, t) - f(y, t)| \leqslant C|x - y|.$$

那么，对任意的（初始值）$(x_0, t_0) \in \Omega \times I$，存在 $\delta > 0$（可能依赖于 (x_0, t_0)），存在唯一的映射 $x(t) : (t_0 - \delta, t_0 + \delta) \to \Omega$，满足如下的常微分方程：

$$\begin{cases} x'(t) = f(x(t), t), \\ x(t_0) = x_0. \end{cases}$$

现在考虑更广的一个场合：给定子流形 $M \subset \mathbb{R}^n$ 上的一个光滑的向量场 X（在大部分的应用中，$M = \mathbb{R}^n$ 就足够了）。按照定义，对每个 $p \in M$，$X(p) \in T_pM$。给定 $p_0 \in M$，我们想找一条通过 p_0 曲线

$$\gamma : \mathbb{R} \to M,$$

使得

$$\begin{cases} \gamma'(0) &= X(p_0), \\ \gamma(0) &= p_0. \end{cases}$$

这样的曲线当然存在，因为这就是切向量的定义。然而我们还想要求在该曲线其他点处，曲线的切向量恰好是 X 在这点处的值。

用常微分方程的语言来讲，我们要找如下方程的一个解：

$$\begin{cases}\gamma'(t) &= X(\gamma(t)),\\ \gamma(0) &= p_0.\end{cases}$$

根据 Cauchy-Lipschitz 定理，存在 $\varepsilon > 0$，使得对于 $t \in (-\varepsilon, \varepsilon)$，上述方程有解。我们这里假设上述的解对一切 $t \in \mathbb{R}$ 都可以定义（其参考下学期的常微分方程课的进一步讨论）。我们首先注意到 γ 的像是落在 M 上的：局部上，我们可以假设 M 是 $f_1, \cdots, f_{n-d}$ 的零点所定义的，其中 $d = \dim M$，所以，为了说明 $\gamma(t) \in M$，只要说明对每个 i，$f_i(\gamma(t)) = 0$ 即可。当 $t = 0$ 时，我们有

$$f_i(\gamma(0)) = f(p_0) = 0.$$

对 t 求导数，我们就有

$$f_i(\gamma(t))' = \nabla_{\gamma'(t)} f_i(\gamma(t)) = \nabla_{X(\gamma(t))} f_i = 0,$$

最后一个等号是因为 $X \in T_pM$，所以，我们可以选一条落在 M 上的曲线来计算这个方向导数，它自然是 0。

现在假设 t_0 固定，p_0 在变化，也就是上面方程的初始值在变化，那么，我们就得到了映射

$$\Phi_{t_0} : \mathbb{R}^n \to \mathbb{R}^n, x \mapsto \Phi_{t_0}(x),$$

其中，$\Phi_{t_0}(x)$ 的定义如下：考虑满足如下条件的方程的唯一的解

$$\begin{cases}\gamma'(t) &= X(\gamma(t)),\\ \gamma(0) &= x \in M.\end{cases}$$

我们就令 $\Phi_t(p_0) = \gamma(t_0)$。

直观上来说，$\Phi_t(x)$ 就是 x 沿着 X 流了 t 这么长的时间所到达的点（至少物理学家们都采用这样的说法），我们通常把 $\Phi_t(x)$ 被称作是 X 通过 x 点的**流线**。很明显，我们有

$$\Phi_0 = \mathbf{Id}.$$

这是单位映射。

另外，利用常微分方程的解的存在唯一性定理，任意的 $s, t \in \mathbb{R}$，我们有

$$\Phi_{s+t} = \Phi_s \circ \Phi_t \quad \Leftrightarrow \Phi_{s+t}(x) = \Phi_s\left(\Phi_t(x)\right), \quad \forall x \in M.$$

证明很简单：对任意的 p_0，假设 t_0 是固定的。我们考虑考虑两条曲线：

$$\gamma_1 : \mathbb{R} \to M, \quad t \mapsto \Phi_{t_0+t}(p_0),$$
$$\gamma_2 : \mathbb{R} \to M, \quad t \mapsto \Phi_t\left(\Phi(p_0)\right).$$

利用链式法则，对每个 i，我们有

$$\gamma_i'(t) = X(\gamma_i(t)).$$

并且当 $t = 0$ 是，我们有 $\gamma_1(0) = \gamma_2(0)$，解的存在唯一性定理，我们就证明了所要的性质。特别地，我们知道

$$\Phi_t \circ \Phi_{-t} = \mathbf{Id}.$$

所以，Φ_t 是 M 到自身的双射。根据下个学期要学习的常微分方程理论（对初值的光滑依赖性），我们还可以证明 Φ_t 都是微分同胚。

我们现在给出散度的几何／物理解释：$\mathrm{div}X$ 衡量了 X 的所对应的 Φ_t 在 $t \to 0$ 时的体积变化率。

引理 326. *给定 $\mathbb{R}^n$ 上的光滑向量场 X，假设 Φ_t 是上述所构造的流，那么，*

$$\frac{d}{dt}\Big|_{t=0} \det\left(d\Phi_t(x)\right) = \left(\mathrm{div}X\right)(x).$$

证明: 根据 Newton-Leibniz 公式以及 Φ_t 的定义，对任意的 $x \in \mathbb{R}^n$，我们有

$$\Phi_t(x) = x + \int_0^t X(\Phi_\tau(x))d\tau,$$

从而（Lebesgue 控制收敛保证了积分与求导数可交换），

$$\begin{aligned} d\Phi_t(x) &= \mathbf{Id} + \int_0^t dX(\Phi_\tau(x))d\Phi_\tau(x)d\tau \\ &= \mathbf{Id} + dX(x)t + O(t^2). \end{aligned}$$

我们上个学期已经做过关于行列式导数的计算。所以，我们有

$$\frac{d}{dt}\Big|_{t=0} \det\left(d\Phi_t(x)\right) = \mathrm{tr}(dX)(x) = \left(\mathrm{div}X\right)(x).$$

□

特别地，用积分的形式来写，我们有

$$\det\left(d\Phi_t(x)\right) = e^{\int_0^t (\mathrm{div}X)(\Phi_\tau(x))d\tau}.$$

根据换元积分公式，我们有

$$\Phi_{t*}m = e^{\int_0^t \mathrm{div}(\Phi_\tau(x))d\tau}m$$

所以，$\mathrm{div}X$ 衡量了 Φ_t 局部上体积的变化。特别地，如果 $\mathrm{div}X = 0$，那么 Φ_t 是保持体积的（保持了 Lebesgue 测度）。

用积分的语言来写，我们有

$$\begin{aligned}\frac{d}{dt}\Big|_{t=0} m(\Phi_t(\Omega)) &= \frac{d}{dt}\Big|_{t=0} \int_{\Phi_t(\Omega)} 1 \; dx \\ &= \frac{d}{dt}\Big|_{t=0} \int_{\Omega} \underbrace{\det\left(d\Phi_t\right)}_{>0, t=0 \text{ 为正}} \; dx \\ &= \int_{\Omega} \frac{d}{dt}\Big|_{t=0} \left(\det d\Phi_t\right) \; dx \\ &= \int_{\Omega} \mathrm{div} X dx.\end{aligned}$$

所以散度的积分是 Ω 的体积在 Φ_t 作用下瞬时的变化率。

如果我们给定一个区域 Ω，它是不动的，那么

$$\begin{aligned}\frac{d}{dt}\Big|_{t=0} \int_{\Omega} 1 d\Phi_{t*} m &= \frac{d}{dt}\Big|_{t=0} m(\Phi_{-t}(\Omega)) \\ &= -\int_{\Omega} \mathrm{div} X dx.\end{aligned}$$

根据 Stokes 公式，我们有

$$\frac{d}{dt}\Big|_{t=0} \Phi_{t*} m(\Omega) = \int_{\partial\Omega} X \cdot (-\nu) dx.$$

上面左边代表瞬间有多少面积（质量）流入了 Ω，右边可以解释为（$-\nu$ 是内法向量，指向区域的内部）通过这个边界瞬间进入了多少面积（质量）。

50 二维的 Brouwer 不动点定理的证明，Hilbert 空间，赋范线性空间完备性的级数判定，Fischer-Riesz 定理

二零二零年五月七日，星期四，阴天

多元微积分的应用举例：二维的 Brouwer 不动点定理

我们回顾一下所谓的 Green 公式：

如果 $\Omega \subset \mathbb{R}^2$ 是有界光滑带边区域，那么对任意的 $\mathbb{R}^2$ 上的光滑（C^1）P 和 Q，我们有

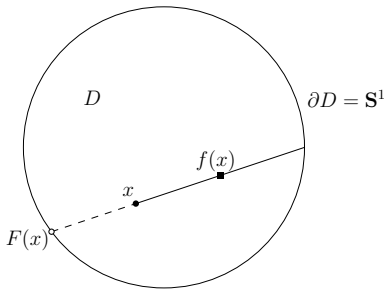
$$\int_\Omega (\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}) dxdy = \oint_{\partial\Omega} Pdx + Qdy.$$

我们现在要来证明 2 维情形的 Brouwer 不动点定理。我们用 D 表示 $\mathbb{R}^2$ 上的单位圆盘：

$$D = \left\{(x,y) \in \mathbb{R}^2 \middle| x^2 + y^2 \leqslant 1\right\}.$$

它的边界是单位圆周：$\partial D = \mathbf{S}^1$。

定理 327 (Brouwer)**.** *假设 $f : D \to D$ 是连续映射，那么一定存在某个点 $x \in D$，使得 $f(x) = x$。*



我们用反正法。如若不然，那么对任意的 $x \in D$，$f(x) \neq x$。考虑从 $f(x)$ 出发的射线，它经过 x 之后，与 ∂D 恰好有一个交点，我们把它记作 $F(x)$。这样，我们就定义了映射

$$F : D \to \mathbf{S}^1 \subset \mathbb{R}^2, \quad x \mapsto F(x).$$

这显然是连续映射（请验证）并且 $F\big|_{\mathbf{S}^1} = \mathbf{Id}$。我们做一个额外的假设：$F : D \to \mathbf{S}^1$ 是 C^∞ 映射（我们在几次课后会用函数逼近的技巧把连续的情况化归到这个情形）。另外，为了方便，我们用坐标来表示 F：

$$F(x,y) = (u(x,y), v(x,y)) \in \mathbb{R}^2.$$

我们还用

$$\gamma : [0, 2\pi] \to \mathbb{R}^2, \quad \vartheta \to (\cos\vartheta, \sin\vartheta)$$

来参数化 ∂D，其中，我们用 γ' 表示其单位切向量 $(-\sin\vartheta, \cos\vartheta)$。

证明的核心在于计算下面的积分（在下个学期学习了 de Rham 上同调理论之后，这个积分会变得很自然，目前大家把这个积分的计算看做是 Stokes 公式的应用和练习即可），我们用两种方法来做：

$$I = \oint_{\partial D} (u\nabla v - v\nabla u) \cdot \gamma'.$$

用传统的第二型曲面积分的记法，我们还可以把它写成：

$$I = \oint_{\partial D} (uv_x - vu_x)dx + (uv_y - vu_y)dy.$$

为了书写方便，我们用 u_x 表示 $\dfrac{\partial u}{\partial x}$，诸如此类。首先，

$$\begin{aligned} I &= \int_0^{2\pi} u\nabla_{\gamma'} v - v\nabla_{\gamma'} u d\vartheta \\ &= \int_0^{2\pi} x\nabla_{\gamma'} y - y\nabla_{\gamma'} x d\vartheta. \end{aligned}$$

这是沿着曲线切线方向的方向导数，所以，它只和函数在曲线上的取值相关。由于在 ∂D 上面，我们有 $u = x$，$v = y$，所以，我们有

$$\begin{aligned} I &= \oint xdy - ydx \\ &= \int_D 2dxdy \\ &= 2\pi. \end{aligned}$$

另外，F 的像落在 $\mathbf{S}^1$ 上，所以，我们有

$$u^2 + v^2 = 1 \quad \Rightarrow \begin{cases} u\frac{\partial u}{\partial x} + v\frac{\partial v}{\partial x} & = 0, \\ u\frac{\partial u}{\partial y} + v\frac{\partial v}{\partial y} & = 0. \end{cases}$$

由于 $(u, v) \neq (0, 0)$，所以上面线性方程系数的行列式为 0，即

$$u_x v_y = u_y v_x.$$

从而，

$$\begin{aligned} I &= \oint_{\partial D} (uv_x - vu_x)dx + (uv_y - vu_y)dy \\ &= \oint_{\partial D} (uv_y - vu_y)_x - (uv_x - vu_x)_y dxdy \\ &= 0, \end{aligned}$$

其中，我们用到了 $u_{xy} = u_{yx}$ 等。这是矛盾的，所以，这样的 F 不存在，从而，Brouwer 不动点定理得到了证明。

Fourier 级数理论

为了学习 Fourier 级数，我们首先复习／学习几个基本的函数空间，我们下个学期的学期也会把这几个空间上的分析作为重点。

首先回忆两个抽象的空间的概念：

a) X 为 $\mathbb{C}$-线性空间，如果它配备了某个范数

$$\|\cdot\| : X \to \mathbb{R}_{\geqslant 0},$$

（即对任意的 $x \in X$，$\|x\| = 0$ 当且仅当 $x = 0$；对任意的 $\lambda \in \mathbb{C}$，$\|\lambda x\| = |\lambda|\|x\|$；对任意的 $x, y \in X$，有 $\|x + y\| \leqslant \|x\| + \|y\|$。）

我们称 $(X, \|\cdot\|)$ 是**赋范线性空间**。此时，对任意的 $x, y \in X$，我们可以用 $d(x, y) = \|x - y\|$ 作为 X 上的距离函数，从而使得 X 称为距离空间。如果 X 在这个距离下是完备的（即 Cauchy 列必收敛），我们就称 $(X, \|\cdot\|)$ 是**完备赋范线性空间**（也叫做 Banach 空间）。

b) X 为 $\mathbb{C}$-线性空间，如果它拥有二次型

$$\langle\cdot,\cdot\rangle : X \times X \to \mathbb{C},$$

满足

1) 对任意的 $\mu, \nu \in \mathbb{C}$ 和 $x, y, z \in X$，我们有

$$\langle \mu x + \nu y, z\rangle = \mu\langle x, z\rangle + \nu\langle y, z\rangle.$$

2) 对任意的 $x, y \in X$，我们有

$$\langle x, y\rangle = \overline{\langle y, x\rangle}.$$

3) 对任意的 $x \in X$，

$$\langle x, x\rangle \geqslant 0$$

并且 $\langle x, x\rangle = 0$ 当且仅当 $x = 0$。

我们就称 $(X, \langle\cdot,\cdot\rangle)$ 是**内积空间**，其中二次型 $\langle\cdot,\cdot\rangle$ 被称作是内积。很容易验证，映射

$$\|\cdot\| : X \to \mathbb{R}_{\geqslant 0}, \quad x \mapsto \sqrt{\langle x, x\rangle}$$

是一个范数（由内积定义的范数），从而，内积空间一定是赋范线性空间。我们把完备的内积空间称作是 **Hilbert 空间**。

注记. 我们注意到，在内积空间中，对任意的 $x, y \in X$，我们都有

$$\big|\langle x, y\rangle\big| \leqslant \|x\|\|y\|$$

最简单的证明就是注意到把内积限制到由 x 和 y 生成由有限维线性空间上也得到内积空间，从而可以用有限维的结论。

在 Lebesgue 积分的理论框架下，有三个最基本的函数空间：$L^1(X,\mathcal{A},\mu)$，$L^2(X,\mathcal{A},\mu)$ 和 $L^\infty(X,\mathcal{A},\mu)$。我们现在引入它们的概念并证明它们的完备性。

首先，给定测度空间 $(X,\mathcal{A},\mu)$，它上面的可积函数全体为 $\mathcal{L}^1(X,\mathcal{A},\mu)$，这是一个 $\mathbb{C}$-线性空间（在 Fourier 级数的学习中，我们不得不研究在复数域中取值的函数）。我们把几乎处处为零的函数记作

$$\mathcal{N}=\big\{f\in\mathcal{L}^1(X,\mathcal{A},\mu)\big|f\text{ 几乎处处为零}\big\}.$$

这是 $\mathcal{L}^1(X,\mathcal{A},\mu)$ 的线性子空间。它的商空间我们定义为：

$$L^1(X,\mathcal{A},\mu)=\mathcal{L}^1(X,\mathcal{A},\mu)/\mathcal{N}.$$

从记号上而言，我们通常还是把 $f\in L^1(X,\mathcal{A},\mu)$ 看作是一个函数：实际上，它是函数的等价类 $[f]\subset\mathcal{L}^1(X,\mathcal{A},\mu)$，$f\in[f]$ 是这里面的一个代表，也就是说如果 $g\in\mathcal{L}^1(X,\mathcal{A},\mu)$ 并且与 f 几乎处处相等，那么我们就说 $f=g$。我们已经证明过：

$$\|\cdot\|_{L^1(X,\mathcal{A},\mu)}:\mathcal{L}^1(X,\mathcal{A},\mu)\to\mathbb{R}_{\geqslant0}$$

是范数，我们通常称它为**L^1-范数**。

其次，我们内积空间 $L^2(X,\mathcal{A},\mu)$。首先，我们定义

$$\mathcal{L}^2(X,\mathcal{A},\mu)=\big\{f:X\to\mathbb{C}\big|f\text{ 是可测度并且 }|f|^2\text{ 是可积函数}\big\}.$$

很容易验证这是一个 $\mathbb{C}$-线性空间并且 $\mathcal{N}$ 是其线性子空间，我们定义

$$L^2(X,\mathcal{A},\mu)=\mathcal{L}^2(X,\mathcal{A},\mu)/\mathcal{N}.$$

我们在 $L^2(X,\mathcal{A},\mu)$ 可以按照如下的方式定义内积，其中 $[f],[g]\in L^2(X,\mathcal{A},\mu)$ 是函数的等价类，f 和 g 分别是这两个等价类中的代表元：

$$\langle[f],[g]\rangle_{L^2(X,\mathcal{A},\mu)}=\langle f,g\rangle_{L^2}=\int_X f(x)\overline{g(x)}d\mu(x).$$

我们很容易看出这个定义是不依赖于等价类中的代表元的选取的并且上面的二次型是一个内积（逐条验证定义即可）。此时，$L^2(X,\mathcal{A},\mu)$ 由该内积定义的范数为

$$\|[f]\|_{L^2(X,\mathcal{A},\mu)}=\|f\|_{L^2}=\sqrt{\int_X|f|^2d\mu}.$$

特别地，根据前面的注，我们有所谓的 **Cauchy-Schwarz 不等式**：

$$\Big|\int_X fgd\mu\Big|\leqslant\|f\|_{L^2}\|g\|_{L^2}.$$

最后，我们定义赋范线性空间 $L^\infty(X,\mathcal{A},\mu)$。首先，令

$$\mathcal{L}^\infty(X,\mathcal{A},\mu)=\big\{f:X\to\mathbb{C}\big|\text{存在 }M\in\mathbb{R}_{\geqslant0}\text{，使得 }|f(x)|\leqslant M\text{ 几乎处处成立}\big\}.$$

很容易验证这是 $\mathbb{C}$-线性空间并且 $\mathcal{N}$ 是其线性子空间，我们定义

$$L^\infty(X,\mathcal{A},\mu)=\mathcal{L}^\infty(X,\mathcal{A},\mu)/\mathcal{N}.$$

我们在 $L^\infty(X,\mathcal{A},\mu)$ 定义范数 $\|\cdot\|_{L^\infty}$，其中，$[f]\in L^\infty(X,\mathcal{A},\mu)$ 是函数的等价类而 f 是其代表元：

$$\|[f]\|_{L^\infty(X,\mathcal{A},\mu)}=\|f\|_{L^\infty}=\inf_{\substack{M\in\mathbb{R}_{\geqslant 0},\\ |f(x)|\leqslant M\text{几乎处处}}} M.$$

为了证明上述空间都是完备的，我们需要一个技术性引理：

引理 328. （完备性的级数判定）假设 $(X,\|\cdot\|)$ 是赋范线性空间，那么，我们有

1) 如果 $(X,\|\cdot\|)$ 是完备的，那么绝对收敛的级数一定收敛，即给定级数 $\sum\limits_{i=1}^{\infty}x_i$，其中 $x_i\in X$，如果

$$\sum_{i=1}^{\infty}\|x_i\|<\infty,$$

那么部分和 $\sum\limits_{i\leqslant n}x_i$ 在 X 中收敛（我们把它的极限记作 $\sum\limits_{i=1}^{\infty}x_i$）。

2) 如果每个绝对收敛的级数均收敛，那么 $(X,\|\cdot\|)$ 是完备的。

证明: 1) 是平凡的：如果令 $S_n=\sum\limits_{i\leqslant n}x_i$，那么，对任意的 $n\geqslant m$，我们有

$$\|S_n-S_m\|\leqslant\sum_{k=m+1}^{n}\|x_k\|.$$

根据 $\sum\limits_{i=1}^{\infty}\|x_i\|$ 收敛，我们知道对任意的 $\varepsilon>0$，存在 $N\geqslant 1$，使得当 $n\geqslant m\geqslant N$ 时，我们有

$$\sum_{k=m+1}^{n}\|x_k\|<\varepsilon.$$

所以，$\{S_n\}_{n\geqslant 1}$ 是 X 中的 Cauchy 列，从而收敛。

为证明 2)，我们取 Cauchy 列 $\{x_i\}_{i\geqslant 1}\subset X$。根据 Cauchy 列的定义，对每个 $p\in\mathbb{Z}_{p\geqslant 1}$，存在正整数 N_p，使得当 $i,j\geq N_p$ 时，我们有

$$\|x_i-x_j\|\leqslant 2^{-p}.$$

从而，我们考虑级数 $\sum\limits_{p=1}^{\infty}(x_{N_p}-x_{N_{p-1}})$，按照指标 N_p 的选取方式，这是一个绝对收敛的级数，从而它收敛。另外，它所对应的部分和为

$$\left(x_{N_p}-x_{N_{p-1}}\right)+\left(x_{N_{p-1}}-x_{N_{p-2}}\right)+\cdots+\left(x_{N_1}-x_{N_0}\right)=x_{N_p}-x_{N_0},$$

这表明 $\{x_i\}_{i\geqslant 1}$ 的子列 $\{x_{N_i}\}_{i\geqslant 1}$ 收敛。我们知道如果 Cauchy 列的子列收敛，Cauchy 列本身就收敛，证毕。 □

我们先证明 L^1 空间的完备性：

定理 329 (Fischer-Riesz)**.** *对任意的测度空间 $(X,\mathcal{A},\mu)$，$L^1(X,\mathcal{A},\mu)$ 是完备的赋范线性空间。*

证明： 根据上述引理，只须证明绝对收敛的级数 $\sum_{i=1}^{\infty} f_i$ 是收敛的，其中 $f_i \in L^1(X,\mathcal{A},\mu)$。为此，我们首先定义函数（的等价类）

$$F : X \to \mathbb{C}, \quad x \mapsto F(x) = \sum_{i=1}^{\infty} |f_i(x)|.$$

这显然是良好定义的函数（除去一个零测集）。根据 Beppo Levi 定理，我们有

$$\begin{aligned}\int_X F(x)d\mu &= \sum_{i=1}^{\infty} \int_X |f_i(x)|d\mu \\ &= \sum_{i=1}^{\infty} \|f_i\|_{L^1} < \infty.\end{aligned}$$

特别地，我们有 $\mu(F^{-1}(+\infty)) = 0$。对任意的 $x \notin F^{-1}(+\infty)$，从而对几乎处处的 $x \in X$，我们定义 $f(x) \in \mathbb{C}$：

$$f(x) = \lim_{N\to\infty} \sum_{i=1}^{N} f_i(x).$$

我们将运用 Lebesgue 控制收敛定理，其中我们把 $F(x)$ 作为控制函数。由于对几乎处处的 x，我们有 $\sum_{i=1}^{N} f_i(x) \to f(x)$。Lebesgue 控制收敛定理表明：

$$\lim_{N\to\infty} \|\sum_{i=1}^{N} f_i(x) - f(x)\|_{L^1} = \int_X |\sum_{i=1}^{N} f_i(x) - f(x)|d\mu = 0.$$

这说明，$\sum_{i=1}^{N} f_i(x)$ 在 $\|\cdot\|_{L^1}$ 所定义的距离下收敛到 $f(x)$，根据前一个引理，$L^1(X,\mathcal{A},\mu)$ 是完备的。 □

Fischer-Riesz 定理的证明还可以给出一个很有意义的推论：

推论 330. *给定 $L^1(X,\mathcal{A},\mu)$ 中的函数序列 $\{f_i\}_{i\geqslant 1}$，我们假设它们在 L^1 范数下收敛到 f，即*

$$f_i \overset{L^1(X,\mathcal{A},\mu)}{\longrightarrow} f, \ i \to \infty.$$

那么，存在子函数序列 $\{f_{i_p}\}_{p\geqslant 1}$，使得对几乎处处的 $x \in X$，我们都有

$$\lim_{p\to\infty} f_{i_p}(x) = f(x)$$

证明：通过选取子序列，我们不妨假设 $f_0 \equiv 0$ 并且对任意的 $n \geqslant 1$，我们都有

$$\|f_n - f_{n+1}\|_{L^1} \leqslant 2^{-n-1}.$$

所以，我们可以把 f_n 写成：

$$f_n = (f_1 - f_0) + (f_2 - f_1) + \cdots + (f_n - f_{n-1})$$

这表明 f_n 可视为是 $L^1(X, \mathcal{A}, \mu)$ 中某个绝对收敛的级数的部分和，前一证明的过程表明除一个零测集外，f_n 逐点收敛到 f。 □

51 偶数维球面上的切向量场，L^2 与 L^∞ 的完备性，连续线性算子，线性映射的延拓，卷积与函数逼近

二零二零年五月九日，星期六，阴天

多元微积分的应用举例：偶数维球面上的切向量场

考虑 $\mathbb{R}^n$ 中的单位球面

$$\mathbf{S}^{n-1}=\left\{(x_1,\cdots,x_n)\in\mathbb{R}^n\middle|x_1^2+\cdots+x_n^2=1\right\}.$$

如果 $n-1=2k-1$ 是偶数，在点 $(x_1,x_2,\cdots,x_{2k-1},x_{2k})\in\mathbf{S}^{n-1}$ 处，我们指定向量

$$v(x_1,\cdots,x_{2k})=(-x_2,x_1,\cdots,-x_{2k},x_{2k-1}).$$

这是 $\mathbf{S}^{2k-1}$ 的一个单位切向量。当 $(x_1,x_2,\cdots,x_{2k-1},x_{2k})\in\mathbf{S}^{n-1}$ 变动的时候，我们就得到了 $\mathbf{S}^{2k-1}$ 上的一个处处非零的光滑向量场。我们现在证明：

定理 331. *在 $\mathbf{S}^{2k}$ 上不存在处处非零的光滑向量场，其中 $k\geqslant 1$。*

假设 $\Omega\subset\mathbb{R}^n$ 是一个有界开集并且 V 是 Ω 上的一个有界的光滑向量场，我们假设 dV 是有界的（把 V 看成映射）。我们现在定义光滑映射

$$\Phi_t:\Omega\to\mathbb{R}^n,\quad x\mapsto x+tV(x).$$

它在 x 处的微分为

$$d\Phi(x)=\mathbf{Id}+tdV(x).$$

所以，当 t 足够小时，$d\Phi(x)$ 可逆，从而，Φ_t 是局部微分同胚（反函数定理）。另外，对任意的 $x,y\in\Omega$，我们有

$$\begin{aligned}|x-y|&=|\Phi_t(x)-\Phi_t(y)-t(V(x)-V(y))|\\&\leqslant|\Phi_t(x)-\Phi_t(y)|+t|V(x)-V(y)|\\&\leqslant|\Phi_t(x)-\Phi_t(y)|+tM|x-y|.\end{aligned}$$

其中，我们用了 dV 的有界性，从而 $|V(x)-V(y)|\leqslant M|x-y|$。所以，当 $t<(2M)^{-1}$ 时，我们有

$$|\Phi_t(x)-\Phi_t(y)|\geqslant\frac{1}{2}|x-y|.$$

这表明 Φ_t 是单射，所以当 t 足够小的时，映射

$$\Phi_t:\Omega\to\Phi_t(\Omega),\quad x\mapsto x+tV(x).$$

是微分同胚。特别地，我们可以用换元积分公式计算 $\Phi_t(\Omega)$ 的体积：

$$m(\Phi_t(\Omega)) = \int_\Omega |\det(\mathbf{I} + t\left(\frac{\partial V_i}{\partial x_j}\right)|dx.$$

将行列展开，我们发现，这是关于 t 的一个次数不超过 n 的多项式。

我们现在用反正法证明这个定理：如若不然，我们假设 $X(x)$ 是 $\mathbf{S}^{n-1}$ 上的一个处处非零的光滑向量场，其中 $n-1=2k$。通过考虑 $\dfrac{X(x)}{|X(x)|}$ 这个向量场，我们不妨假设 V 在每个点处都是单位长度的。令

$$\Omega = \big\{(x_1, \cdots, x_n) \in \mathbb{R}^n \big| \frac{1}{2} < x_1^2 + \cdots + x_n^2 < 2\big\}.$$

那么，

$$\Omega = \bigcup_{\frac{1}{2}<r<2} S_r = \bigcup_{\frac{1}{2}<r<2} (r\mathbf{S}^{n-1}).$$

我们定义 Ω 上的向量场为

$$V(x) = |x| X\left(\frac{x}{|x|}\right)$$

对任意的 r，我们知道 $\Phi_t(S_r) \subset \sqrt{1+t^2}S_r$，所以，当 r 足够小的时候，Φ_t 将 Ω 映射成了

$$\Phi_t(\Omega) = \big\{(x_1, \cdots, x_n) \in \mathbb{R}^n \big| \frac{1}{2}t\sqrt{1+t^2} < x_1^2 + \cdots + x_n^2 < 2t\sqrt{1+t^2}\big\}.$$

那么，根据 Lebesgue 测度在伸缩变换下的性质，我们有

$$m(\Phi_t(\Omega)) = (1+t^2)^{\frac{2k+1}{2}} m(\Omega)$$

这不是 t 的多项式，矛盾。

函数空间的完备性：继续

定理 332. *对任意的测度空间 $(X, \mathcal{A}, \mu)$，$L^2(X, \mathcal{A}, \mu)$ 是 Hilbert 空间。*

证明： 根据之前的引理，我们只需要证明绝对收敛的级数 $\sum_{i=1}^{\infty} f_i$ 在 $L^2(X, \mathcal{A}, \mu)$ 中收敛即可，其中 $f_i \in L^2(X, \mathcal{A}, \mu)$。仿照 Fischer-Riesz 定理，我们定义

$$S_N(x) = \sum_{i=1}^{N} f_i(x), \quad |S|_N(x) = \sum_{i=1}^{N} |f_i(x)|,$$

并令

$$F(x) = \big(\lim_{N\to\infty} |S|_N\big)^2 = \big(\sum_{i=1}^{\infty} |f_i(x)|\big)^2.$$

根据 Fubini 定理和 Cauchy-Schwarz 不等式，我们有

$$\begin{aligned}\int_X |F| &= \int_X \sum_{i=1}^{\infty}\sum_{j=1}^{\infty} |f_i(x)||f_j(x)|d\mu(x)\\ &= \sum_{i=1}^{\infty}\sum_{j=1}^{\infty}\int_X |f_i(x)||f_j(x)|d\mu(x)\\ &\leqslant \sum_{i=1}^{\infty}\sum_{j=1}^{\infty}\|f_i\|_{L^2}\|f_j\|_{L^2}\\ &= \big(\sum_{i=1}^{\infty}\|f_i(x)\|_{L^2}\big)^2 < \infty.\end{aligned}$$

所以，F 是可积函数（显然可测）并且 $A = F^{-1}(+\infty)$ 是零测集。根据 F 的定义，当 $x \notin A$ 时，$|S|_N(x)$ 收敛到 $\sqrt{F(x)}$，从而存在 $S(x)$，使得当 $x \notin A$ 时，我们有 $\lim\limits_{N\to\infty} S_N(x) = S(x)$，即

$$\sum_{i=1}^{\infty} f_i(x) = S(x), \quad x \notin A.$$

我们现在选取 $4F(x)$ 作为控制函数：

$$|S_N(x) - S(x)|^2 \leqslant 2|S_N(x)|^2 + 2|S(x)|^2 \leqslant 4F(x).$$

由于 $\lim\limits_{N\to\infty}\big|S_N(x) - f(x)\big|^2 = 0$（几乎处处），根据 Lebesgue 控制收敛定理，我们有

$$\lim_{N\to\infty}\int_X |S_N(x) - S(x)|^2 d\mu(x) = 0.$$

这就完成了证明。 □

推论 333. *给定 $L^2(X,\mathcal{A},\mu)$ 中的函数序列 $\{f_i\}_{i\geqslant 1}$，我们假设它们在 L^2 范数下收敛到 f，即*

$$f_i \overset{L^2(X,\mathcal{A},\mu)}{\longrightarrow} f,\ i\to\infty.$$

那么，存在子函数序列 $\{f_{i_p}\}_{p\geqslant 1}$，使得对几乎处处的 $x\in X$，我们都有

$$\lim_{p\to\infty} f_{i_p}(x) = f(x)$$

我们把推论的证明留成作业。

定理 334. *对任意的测度空间 $(X,\mathcal{A},\mu)$，$L^\infty(X,\mathcal{A},\mu)$ 是完备的。*

证明: 对于 $L^\infty(X,\mathcal{A},\mu)$ 中的 Cauchy 列 $\{f_i\}_{i\geqslant 1}$，我们定义

$$A_{ij} = \big\{x\in X\big||f_i(x) - f_j(x)| > \|f_i - f_j\|_{L^\infty}\big\}.$$

按定义，这是一个零测集，所以 $A=\bigcup\limits_{i,j=1}^{\infty}A_{ij}$ 是零测集。对于 $x\notin A$，$\{f_i(x)\}_{i=1,2,\cdots}$ 为 Cauchy 列：

$$|f_i(x)-f_j(x)|\leqslant\|f_i-f_j\|_{L^\infty}\to 0.$$

从而，存在 $f(x)\in\mathbb{C}$，使得 $\lim\limits_{n\to\infty}f_n(x)=f(x)$，所以函数

$$f:X\to\mathbb{C}$$

几乎处处有定义。对任意的 $\varepsilon>0$，存在正整数 N，使得当 $i,j\geqslant N$ 时，我们有

$$\|f_i-f\|_{L^\infty}<\varepsilon.$$

所以，对任意的 $x\notin A$，我们有

$$|f_i(x)-f_j(x)|\leqslant\|f_i-f_j\|_{L^\infty}<\varepsilon.$$

令 $j\to\infty$，所以对任意的 $i\geqslant N$，我们都有

$$|f_i(x)-f(x)|\leqslant\varepsilon,\quad x\notin A.$$

这表明 $\{f_i\}_{i\geqslant 1}$ 在 L^∞ 范数下收敛到 f。 □

注记. （重要：在不同的意义下收敛）给定一个函数列 $\{f_i\}_{i\geqslant1}$，在我们谈论它收敛到某个函数 f 时，一定要事先界定用的哪个范数。不同的范数可能给出不同的结论，比如说，函数列

$$f_n(x)=\begin{cases}\frac{1}{n}, & 0\leqslant x\leqslant\frac{1}{n},\\ 0, & \frac{1}{n}<x\leqslant 1\end{cases}$$

它在 $L^2([0,1])$ 上收敛到 0，在 $L^1([0,1])$ 上不收敛。

另外，如果 X 是有限维赋范线性空间，X 上任意范数定义出的收敛性都是等价的。

我们将在作业中证明上面的论断。

最后，我们讨论所谓的（赋范线性空间之间）连续线性映射。

引理 335. 给定赋范线性空间 $(E,\|\cdot\|_E)$，$(F,\|\cdot\|_F)$ 和它们之间的（$\mathbb{C}$-）线性映射 $L:E\to F$。那么，如下三个论断是等价的：

1) L 在 $0\in E$ 处连续；

2) L 是连续的；

3) L 是**有界**，即存在 $C>0$，对每个 $e\in E$，我们都有

$$\|L(e)\|_F\leqslant C\|e\|_E.$$

证明: 2)⇒1) 是显然；为了说明 3) ⇒ 2)，对任意的 $e \in E$，以及 $x \in E$，我们有

$$\begin{aligned} d_F\left((L(x), L(e)\right) &= \|L(x) - L(e)\|_F = \|L(x-e)\|_F \\ &\leqslant C\|x-e\|_E = C\|x-e\|_E = d_E(x,e). \end{aligned}$$

所以，当 $x \to e$ 时，我们有 $L(x) \to L(e)$，从而，2) 成立。

现在证明 1) ⇒ 3)。由连续性，对于 $\varepsilon = 1$，存在 $\delta > 0$，当 $\|e\|_E \leqslant \delta$ 时，我们有 $\|L(e)\|_F < 1$。那么，对任意的 $e \in E$，我们有

$$\|T\big(\frac{\delta e}{\|e\|_E}\big)\|_F \leqslant 1.$$

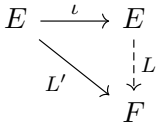
我们选取 $C = \delta^{-1}$ 即可。 □

注记. 这个命题说的是连续线性映射一定是有界线性映射，我们之后对这两个名词不加区分。

推论 336 (线性映射的延拓)**.** 给定赋范线性空间 $(E, \|\cdot\|_E)$, $(E', \|\cdot\|_{E'})$ 和 $(F, \|\cdot\|_F)$，其中 $E' \subset E$ 为线性子空间且其范数为诱导范数，即对任意的 $e' \in E$，我们有

$$\|e'\|_{E'} = \|e'\|_E.$$

假设 E' 在 E 中是稠密的。



如果存在连续线性映射 $L' : E' \to F$，那么存在唯一的连续线性映射 $L : E \to F$，使得 $L\big|_{E'} = L'$。

证明: 唯一性：对任意的 $e \in E$，根据稠密性，我们可以选取 $\{e'_i\}_{i\geqslant 1} \subset E'$，使得

$$\lim_{i\to\infty} e'_i = e.$$

所以，

$$L(e) = \lim_{i\to\infty} L(e'_i) = \lim_{i\to\infty} L'(e'_i)$$

是被唯一决定的。反之，对任意的 e，我们任取 $\{e'_i\}_{i\geqslant 1} \subset E'$，使得 $\lim\limits_{i\to\infty} e'_i = e$，我们就定义

$$L(e) := \lim_{i\to\infty} L'(e'_i)$$

为了说明这是良好定义的，假设 $\{e''_i\}_{i\geqslant 1} \subset E'$，使得 $\lim\limits_{i\to\infty} e''_i = e$，我们只要说明

$$\lim_{i\to\infty} L'(e'_i) = \lim_{i\to\infty} L'(e''_i)$$

即可。这因为

$$\|L'(e'_i) - L'(e''_i)\|_F = \|L'(e'_i - e''_i)\|_F \leqslant \|e'_i - e''_i\|_E.$$

所以，$\lim\limits_{i\to\infty} \|L'(e'_i) - L'(e''_i)\|_F = 0$，从而，$L$ 是良好定义的。上面的计算还表明

$$\|L(e)\|_F = \lim_{i\to\infty} \|L'(e'_i)\|_F \leqslant \lim_{i\to\infty} C\|e'_i\|_E = C\|e\|_E.$$

这说明 L 是连续线性映射。 □

注记. 这个不起眼的定理是贯穿分析学始终的想法：研究一个连续函数我们只需要搞清楚它在有理数（或者其他稠密子集）上的取值即可。从此往后，为了构造赋范线性空间之间的连续线性映射，我们只需要在稠密的子空间上面构造即可。我们会证明，很多函数空间都包含光滑函数作为子空间，所以，为了定义好的连续线性算子，我们只要对光滑函数进行构造即可。

函数的逼近

定理 337. 给定测度空间 $(X,\mathcal{A},\mu)$，我们仍然用 $\mathcal{E}(X)$ 表示简单函数所构成的线性空间（这是阶梯函数的商掉 $\mathcal{N}$ 的商空间）。那么，$\mathcal{E}(X)$ 是 $L^p(X,\mathcal{A},\mu)$ 的稠密子空间，其中 $p=1$ 或者 2。

注记. 对于 $p=\infty$ 也有类似的结果，然而，我们要指出，此时它依赖于我们对于阶梯函数的定义，即是否容许测度为 ∞ 的集合的示性函数作为简单函数。这些都是概念上的腾挪，没有太多的意思，所以我们忽略掉这个情况。

证明: 首先，对 $L^1(X,\mathcal{A},\mu)$ 来证明这个结论。任意选取 $f\in L^1(X,\mathcal{A},\mu)$，我们分情况讨论：

1) 如果 f 是正函数，那么存在上升的简单函数列 $\{f_i\}_{i\geqslant 1}$，它逐点地收敛到 f，从而

$$\|f_i-f\|_{L^1}=\int_X f(x)-f_i(x)d\mu\to 0.$$

2) 如果 f 是实值函数，我们将它写成正函数的差：

$$f=f_+-f_-.$$

我们可以选取简单函数序列 $\{f_{i,+}\}_{i\geqslant 1}$ 和 $\{f_{i,-}\}_{i\geqslant 1}$，使得

$$\lim_{i\to\infty}\|f_{i,\pm}-f_\pm\|_{L^1}=0.$$

从而，

$$\lim_{i\to\infty}\|(f_{i,+}-f_{i,-})-f\|_{L^1}\leqslant\lim_{i\to\infty}\|f_{i,+}-f_+\|_{L^1}+\lim_{i\to\infty}\|f_{i,-}-f_-\|_{L^1}=0.$$

这就给出了 f 的逼近。

3) 如果 f 是复值函数，我们可以类似地将它写成实部和虚部来逼近。

其次，我们对 $L^2(X,\mathcal{A},\mu)$ 来证明这个结论。我们仍然先处理 $f\in L^2(X,\mathcal{A},\mu)$ 是正函数的情形，此时，我们知道 f^2 是可积函数，所以我们选取上升的简单函数列 $\{F_i\}_{i\geqslant 1}$，它逐点地收敛到 f^2，我们令 $f_i=\sqrt{F_i}$。我们考虑函数序列 $\{|f_i-f|^2\}_{i\geqslant 1}$，很明显，这个序列逐点地收敛到 0。我们可以用 $4|f|^2$ 作为控制函数（因为 $f_i(x)\leqslant f(x)$）：

$$|f_i(x)-f(x)|^2\leqslant 4|f|^2.$$

根据 Lebesgue 控制收敛定理，我们有

$$\int_X |f_i(x) - f(x)|^2 d\mu \to 0.$$

即

$$\lim_{i\to\infty} \|f_i - f\|_{L^2} = 0.$$

这就给出了这种情形的证明。如果 f 是实值函数或者复值函数，我们把它拆成正函数的和然后利用线性即可。 □

根据积分的定义，用简单函数来逼近 L^p 中的函数算是理所当然的事情，我们现在要用光滑函数来逼近 $L^p(\mathbb{R}^n)$ 中的函数。为此，我们要引入所谓的卷积的概念。之后在不加说明的情况下，我们都用 Lebesgue 测度。

任意给定 $f, g \in L^1(\mathbb{R}^n)$，根据 Fubini 定理，我们有

$$\int_{\mathbb{R}^n\times\mathbb{R}^n} |f(x-y)||g(y)|dxdy = \left(\int_{\mathbb{R}^n} |f|dx\right)\cdot\left(\int_{\mathbb{R}^n} |g|dx\right) = \|f\|_{L^1}\|g\|_{L^1}.$$

这表明函数

$$(x, y) \mapsto |f(x-y)||g(y)|$$

是 $\mathbb{R}^n\times\mathbb{R}^n$ 上的可积函数。特别地，利用 Fubini 定理（正函数版本）的结论，对几乎处处的 $x \in \mathbb{R}^n$，关于 $y \in \mathbb{R}^n$ 的函数

$$y \mapsto f(x-y)g(y)$$

是可积的，这里可积是对 dy 而言的，我们将 x 视作是固定的。所以，对几乎处处的 $x \in \mathbb{R}^n$，函数

$$f * g : \mathbb{R}^n \to \mathbb{C}, \quad x \mapsto (f*g)(x) = \int_{\mathbb{R}^n} f(x-y)g(y)dy,$$

是良好定义的并且是 Lebesgue 可积的（对测度 dx），我们称 $f * g$ 是 f 和 g 的**卷积**。

注记. 还有一种观点来看卷积：在每个点处用 g 做为权函数，对 f 进行平均。比如说 $g(x) = \dfrac{1}{|B_1|}\mathbf{1}_{B_1}$，这是单位球的示性函数并乘了正确的因子，那么

$$f * g(x) = \frac{1}{|B_1|}\int_{B_1} f(x-y)dy.$$

这就是 f 在每个点 x 处以此点为中心以 1 为半径的球内做平均。

另外，Fubini 定理表明

$$\int_{\mathbb{R}^n} |f*g| \leqslant \big(\int_{\mathbb{R}^n} |f|dx\big)\big(\int_{\mathbb{R}^n} |g|dy\big), \ \Rightarrow \ \|f*g\|_{L^1} \leqslant \|f\|_{L^1}\|g\|_{L^1}.$$

这说明

$$L^1(\mathbb{R}^n)\times L^1(\mathbb{R}^n) \xrightarrow{*} L^1(\mathbb{R}^n).$$

这给出了 $L^1(\mathbb{R}^n)$ 上面一个新的乘法结构。这个乘法结构用到了 $\mathbb{R}^n$ 上的群结构。我们将在作业中证明，$L^1(\mathbb{R}^n)$ 上卷积没有单位元，也就是说不存在某个函数 $e \in L^1(\mathbb{R}^n)$，使得对任意的 $f \in L^1(\mathbb{R}^n)$，我们都有 $e * f = f$。我们将在作业中证明这个结论。我们在下个学期课程的开始能够更好的理解 $L^1(\mathbb{R}^n)$ 上卷积没有单位元的原因（它的单位元生活在更大的空间里）。

我们现在证明，卷积满足交换律和结合律：

引理 338. *假设 $f, g, h \in L^1(\mathbb{R}^n)$，那么，我们有*

1) 乘法交换律：对于几乎处处的 $x \in \mathbb{R}^n$，都有

$$(f * g)(x) = (g * f)(x).$$

（从而这两个函数在 $L^1(\mathbb{R}^n)$ 中对应着同一个函数）

2) 乘法结合律：对于几乎处处的 $x \in \mathbb{R}^n$，都有

$$\big((f * g) * h\big)(x) = \big(f * (g * h)\big)(x).$$

证明： 先证明交换律，根据换元积分公式，我们有

$$\begin{aligned}(f * g)(x) &= \int_{\mathbb{R}^n} f(x-y)g(y)dy \\ &= \int_{\mathbb{R}^n} f(z)g(x-z)dy \\ &= (g * f)(x).\end{aligned}$$

其中，我们令 $z = x - y$。现在证明结合律，根据 Fubini 定理（很明显，取绝对值之后所出现的函数仍然可积），我们有

$$\begin{aligned}\big((f * g) * h\big)(x) &= \int_{\mathbb{R}^n} \left(\int_{\mathbb{R}^n} f((x-z)-y)g(y)dy \right) h(z)dz \\ &= \int_{\mathbb{R}^n \times \mathbb{R}^n} f(x-z-y)g(y)h(z)dydz \\ &= \int_{\mathbb{R}^n \times \mathbb{R}^n} f(x-t)g(y)h(t-y)dydt.\end{aligned}$$

其中，我们做了坐标替换 $t = y + z$。所以，再次利用 Fubini 定理，我们得到

$$\begin{aligned}\big((f * g) * h\big)(x) &= \int_{\mathbb{R}^n} f(x-t) \left(\int_{\mathbb{R}^n} g(y)h(t-y)dy \right) dt \\ &= f * (g * h)(x).\end{aligned}$$

这就完成了证明。 □

另外，对于其它的 L^p 空间，我们有

引理 339. 对于 $p=2$ 或 ∞，我们有

$$L^1(\mathbb{R}^n)\times L^p(\mathbb{R}^n)\xrightarrow{*}L^p(\mathbb{R}^n),$$

即若 $f\in L^1(\mathbb{R}^n)$，$g\in L^p(\mathbb{R}^n)$，那么

$$(f*g)(x)=\int_{\mathbb{R}^n}f(x-y)g(y)dy$$

在 $L^p(\mathbb{R}^n)$ 中是良好定义的（几乎处处）。特别地，我们有

$$\|f*g\|_{L^p}\leqslant\|f\|_{L^1}\|g\|_{L^p}.$$

证明: 对于 $g\in L^\infty(\mathbb{R}^n)$，我们有

$$\begin{aligned}|(f*g)(x)|&=\left|\int_{\mathbb{R}^n}f(x-y)g(y)dy\right|\\&\leqslant\int_{\mathbb{R}^n}|f(x-y)||g(y)|dy\\&\leqslant\|g\|_{L^\infty}\int_{\mathbb{R}^n}|f(x-y)|dy\\&=\|f\|_{L^1}\|g\|_{L^\infty}.\end{aligned}$$

所以命题成立。

对于 $g\in L^2(\mathbb{R}^n)$，我们有考虑正函数

$$F:\mathbb{R}^n\to\mathbb{R}_{\geqslant0},\quad x\mapsto\int_{\mathbb{R}^n}|f(x-y)||g(y)|dy.$$

根据 Fubini 定理，我们有

$$\begin{aligned}\int_{\mathbb{R}^n}|F(x)|^2&=\int_{\mathbb{R}^n}\left(\int_{\mathbb{R}^n}|f(x-y)||g(y)|dy\int_{\mathbb{R}^n}|f(x-z)||g(z)|dz\right)dx\\&=\int_{\mathbb{R}^{3n}}|f(x-y)||g(y)||f(x-z)||g(z)|dxdydz\\&=\int_{\mathbb{R}^{2n}}\left(\int_{\mathbb{R}^n}|f(x-y)||f(x-z)|dx\right)|g(y)||g(z)|dydz\\&\leqslant\int_{\mathbb{R}^{2n}}\left(\int_{\mathbb{R}^n}|f(x-y)|^2dx\right)^{\frac12}\left(\int_{\mathbb{R}^n}|f(x-z)|^2dx\right)^{\frac12}|g(y)||g(z)|dydz\\&=\|f\|_{L^2}^2\int_{\mathbb{R}^{2n}}|g(y)||g(z)|dydz\\&=\|f\|_{L^2}^2\int_{\mathbb{R}^n}|g(y)|dy\int_{\mathbb{R}^n}|g(z)|dz.\end{aligned}$$

从而，

$$\int_{\mathbb{R}^n}|F(x)|^2\leqslant\|f\|_{L^2}^2\|g\|_{L^1}^2.$$

这表明，$|F(x)|^2$ 是 L^1 函数，从而几乎处处有定义，所以，函数

$$x \mapsto f * g(x) = \int_{\mathbb{R}^n} f(x-y)g(y)dy$$

也几乎处处有定义，上面的不等式自然表明

$$\|f * g\|_{L^2}^2 \leqslant \|f\|_{L^1}^2 \|g\|_{L^2}^2.$$

开方之后就证明了命题。 □

我们需要说一下光滑的这个情形是每个点都有定义的。

利用 Lebesgue 控制收敛定义，我们可以说明与光滑函数做卷积可以提升函数的正则性：

引理 340. *我们用 $\mathcal{D}(\mathbb{R}^n) = C_0^\infty(\mathbb{R}^n)$ 表示所有 $\mathbb{R}^n$ 上的光滑并且支集是紧集的函数。*

如果 $\varphi \in C_0^\infty(\mathbb{R}^n)$，$g \in L^p(\mathbb{R}^n)$，其中，$p = 1, 2$ 或者 ∞，那么 $\varphi * g \in L^p(\mathbb{R}^n) \cap C^\infty(\mathbb{R}^n)$。

证明: 给定 $g \in L^p(\mathbb{R}^n)$, 只需要证明, 对任意的 $\varphi \in C_0^\infty(\mathbb{R}^n)$, $\varphi * g \in C^1(\mathbb{R}^n)$ 并且对任意的 $i \leqslant n$, 我们有

$$\frac{\partial(\varphi * g)}{\partial x_i} = \int_{\mathbb{R}^n} \frac{\partial \varphi}{\partial x_i}(x-y)g(y)dy$$

即可：因为我们可以继续用 $\dfrac{\partial \varphi}{\partial x_i}$ 代替 φ 继续求导数，从而对任意的多重指标 α，我们有

$$\partial^\alpha(\varphi * g) = (\partial^\alpha \varphi) * g.$$

我们只对 $p = 1$ 的情形进行证明，其它情况留作作业。由于

$$\begin{aligned}\left|\frac{\partial \varphi}{\partial x_i}\left((x-y)g(y)\right)\right| &= \left|\frac{\partial \varphi}{\partial x_i}(x-y)g(y)\right| \\ &\leqslant M|g(y)|,\end{aligned}$$

其中，由于 $\dfrac{\partial \varphi}{\partial x_i}$ 也是有紧支集的光滑函数，所有存在 M，对任意的 $x \in \mathbb{R}^n$，我们都有 $\left|\dfrac{\partial \varphi}{\partial x_i}\right| \leqslant M$。此时，我们可以用 Lebesgue 控制收敛定理的第二个推论（对含参数积分的求导公式）来直接推出上面的结论。 □

注记. *在 $\varphi \in C_0^\infty(\mathbb{R}^n)$ 这种情形下，我们知道存在 A，使得对任意的 $x \in \mathbb{R}^n$，$|\varphi(x)| \leqslant A$。所以，*

$$\varphi * g(x) = \int_{\mathbb{R}^n} \varphi(x-y)g(y)dy$$

在任意一个点 x 处都有的定义，这是因为 $|\varphi(x-y)g(y)| \leqslant A|g(y)|$ 是可积函数，我们不需要用 Fubini 定理的结论来说明 $\varphi * g(x)$ 的存在性了。

51.1 作业：Stokes 公式的应用

清华大学 19-20 春季学期，数学分析二，作业 9

请用 A4 大小的纸张**正反面**用钢笔，签字笔或者圆珠笔书写，并注明自己的姓名，年级和作业的**总页数**。除定理公式所涉及的人名之外，请使用**中文**。本次作业的提交时间和地点为 **2 月 92 日**上午的课堂上，逾期视作零分。

课堂补充

A1) 证明向量场的导数公式：假设 X, Y 是 $\mathbb{R}^3$ 上的光滑向量场，φ 是 $\mathbb{R}^3$ 上实值光滑函数，那么，

$$\begin{aligned}
&\nabla \times (\nabla \varphi) = 0, \quad \nabla \cdot (\nabla \times X) = 0, \quad \triangle \varphi = \nabla \cdot (\nabla \varphi), \\
&\nabla \times (\varphi X) = \varphi (\nabla \times X) + \nabla \varphi \times X, \\
&\nabla \cdot (X \times Y) = (\nabla \times X) \cdot Y - (\nabla \times Y) \cdot X, \\
&\nabla \times (X \times Y) = X(\nabla \cdot Y) - Y(\nabla \cdot X) + (Y \cdot \nabla) X - (X \cdot \nabla) Y.
\end{aligned}$$

A2)（三维梯度、散度的几何定义）假设 X, Y 是 $\mathbb{R}^3$ 上的光滑向量场，φ 是 $\mathbb{R}^3$ 上的光滑实值函数，$B_\varepsilon(x)$ 为以 $x \in \mathbb{R}^3$ 为球心以 $\varepsilon > 0$ 为半径的球。

1. 证明，
$$\nabla \varphi(x) = \lim_{\varepsilon \to 0} \frac{1}{m(B_\varepsilon(x))} \int_{\partial B_\varepsilon(x)} \varphi \cdot \nu \, d\sigma.$$
2. 证明，
$$\operatorname{div}(X) = \lim_{\varepsilon \to 0} \frac{1}{m(B_\varepsilon(x))} \int_{\partial B_\varepsilon(x)} X \cdot \nu \, d\sigma.$$

A3)（二维旋度的几何定义）$D_\varepsilon(x)$ 表示以 $x \in \mathbb{R}^2$ 为圆心以 $\varepsilon > 0$ 为半径的圆盘，我们用 ν 表示其边界的单位法向量场。证明，

$$\operatorname{curl}(X) = \lim_{\varepsilon \to 0} \frac{1}{\sigma(D_\varepsilon)} \oint_\gamma X \cdot d\gamma$$

A4) 令 Ω 为 $\mathbb{R}^3$ 中的有界光滑带边区域，$]\nu$ 为 $\partial\Sigma$ 的单位外法向量。对于任意的 $\mathbb{R}^3$ 中的光滑函数 f，证明，

$$\int_\Omega \triangle f \, dxdydz = \int_{\partial\Omega} \nabla_\nu f \, d\sigma.$$

Stokes 公式相关问题

B1)（物理问题中，根据定义，我们写下一切都有定义）假设一个粒子 P 从 A 点运动到 B 点，它的的运动轨迹（$\mathbb{R}^n$ 中）由参数曲线

$$\gamma : [a, b] \to \mathbb{R}^n$$

给出，其中 $\gamma(a)=A$，$\gamma(b)=B$。粒子受到的力由向量场 $F(\gamma(t))$ 给出，那么我们定义粒子 P 在路径 γ 上做的**功**由第二型曲线积分

$$\int_{\gamma} F\cdot d\gamma$$

给出。

(a) 证明，做功与路径保定向的参数化无关，即若

$$\tau:[a,b]\to[\tau(a),\tau(b)]$$

是光滑函数并且对任意的 $t\in[a,b]$，都有 $\tau'(t)>0$，那么

$$\int_{\gamma} F\cdot d\gamma=\int_{\widetilde{\gamma}} F\cdot d\widetilde{\gamma}\ \left(=\int_{\tau(a)}^{\tau(b)} F(\tilde{\gamma}(\tau))\cdot\frac{d\tilde{\gamma}}{d\tau}d\tau\right)$$

其中

$$\tilde{\gamma}=\gamma\circ\tau^{-1}:[\tau(a),\tau(b)]\to\mathbb{R}^n.$$

(b) 假设 $n=2$，我们给定 $\mathbb{R}^2-\{(0,0)\}$ 上一个力场

$$F(x,y)=(-\frac{y}{x^2+y^2},\frac{x}{x^2+y^2}).$$

试计算 F 分别沿路径 γ_1 和 γ_2 所做的功，其中

$$\gamma_1:[0,2\pi]\to\mathbb{R}^2-\{(0,0)\},\quad t\mapsto(\cos t,\sin t),$$
$$\gamma_2:[0,2\pi]\to\mathbb{R}^2-\{(0,0)\},\quad t\mapsto(2+\cos t,\sin t).$$

(c) 假设

$$f:(0,+\infty)\to\mathbb{R}$$

是光滑函数，对于 $x\in\mathbb{R}^3$，我们定义 $\mathbb{R}^3-\{(0,0,0)\}$ 上的一个有心力场：

$$F:\mathbb{R}^3-\{(0,0,0)\}\to\mathbb{R},\quad x\mapsto f(|x|)x.$$

任意给定光滑曲线

$$\gamma:[0,1]\to\mathbb{R}^3-\{(0,0,0)\}$$

试计算 F 沿 γ 做的功。

B2) F 是区域 $U\subset\mathbb{R}^n$ 上的光滑向量场（力场），如果存在函数 $P\in C^\infty(U)$ ，使得 $F=-\nabla P$，我们就称 F 是**保守的**并把 P 称为 F 的**势函数**。

(a) 证明，保守力做的功只依赖于路径的起点和终点而与路径的选取无关，即若 $F = -\nabla P$，并且 $\gamma_i : [a, b] \to \mathbb{R}^n$ 和 $\gamma_2 : [a, b] \to \mathbb{R}^n$ 是光滑曲线，满足 $\gamma_1(a) = \gamma_2(a), \gamma_1(b) = \gamma_2(b)$，那么

$$\int_{\gamma_1} F \cdot d\gamma_1 = \int_{\gamma_2} F \cdot d\gamma_2.$$

(b) 证明，如果 $F = -\nabla P$ 是保守力，质点的运动满足 Newton 第二定律

$$m\ddot{x}(t) = F(x(t)),$$

其中，$\ddot{x}$ 代表两次导数，m 是常数，代表粒子的质量，那么质点运动满足能量守恒，即动能与势能之和为常数:

$$E(t) = \frac{1}{2}m|\dot{x}|^2 + \Phi(x(t)) = E(t_0).$$

(c) 如果 X 是保守的，那么 X 是无旋的，即其旋度 $\nabla \wedge X = 0$。

(d)（**重要**）给定 $\mathbb{R}^n$ 上的光滑向量场 F。如果沿任意的光滑不自交的闭曲线

$$\gamma : [a, b] \to \mathbb{R}^n$$

（不自交指的是 γ 是单射，闭指的是 $\gamma(a) = \gamma(b)$）上的环量

$$\oint_\gamma F \cdot d\gamma = 0,$$

证明，那么 F 是保守的。

B3) $\Omega \subset \mathbb{R}^2$ 是有界光滑带边区域，它的边界 $\gamma = \partial\Omega$ 是光滑的闭曲线，ν 是 γ 的单位外法向量。

(a) 证明，对任意常值向量 $v \in \mathbb{R}^2$，我们有

$$\int_\gamma \langle v, \nu \rangle d\sigma = 0$$

(b) 试计算积分

$$\int_\gamma x\nu_1 + y\nu_2 \; d\sigma$$

其中 $\nu = (\nu_1, \nu_2)$。

B5) (a) 假设空间 $\mathbb{R}^3$ 有 N 个点电荷，它们的位置和电量分别为 $x_i \in \mathbb{R}^3$ 和 $q_i \in \mathbb{R}$, 其中 $i = 1, \cdots, N$。根据 Coulomb 定律，它们在 $x \in \mathbb{R}^3$ 处产生的电场（向量场）为

$$E(x) = \sum_{i=1}^{N} \frac{q_i}{4\pi\varepsilon_0} \frac{x - x_i}{|x - x_i|^3},$$

其中 $\varepsilon_0 > 0$ 是介电常数。试证明 Gauss 定律：$\Omega \subset \mathbb{R}^3$ 是有界光滑带边区域，那么

$$\oiint_{\partial\Omega} E \cdot d\vec{\sigma} = \frac{Q}{\varepsilon_0},$$

其中 $Q = \sum_{i=1}^{N} q_i$ 是 Ω 内的总电荷并且我们假设对任意的 i，$x_i \in \mathring{\Omega}$。

(b) 假设 $\rho \in C^\infty(\mathbb{R}^3, \mathbb{R})$ 是电荷密度函数，根据 Coulomb 定律，它在 $x \in \mathbb{R}^3$ 处产生的电场（向量场）为（如果可积的话）

$$E(x) = \frac{1}{4\pi\varepsilon_0} \int_{\mathbb{R}^3} \frac{\rho(y)(x-y)}{|x-y|^3} dy.$$

试证明 Gauss 定律：如果 $\Omega \subset \mathbb{R}^3$ 是有界光滑带边区域，$\mathrm{supp}\rho \in \mathring{\Omega}$，那么

$$\oiint_{\partial\Omega} E \cdot d\vec{\sigma} = \frac{Q}{\varepsilon_0},$$

其中 $Q = \int_\Omega \rho(x)dx$ 是 Ω 内的总电荷。

(c) 证明，如果电荷密度函数 $\rho \in C_0^\infty(\mathbb{R}^3)$ 是具有紧支集的光滑函数，那么静电场

$$E(x) = \frac{1}{4\pi\varepsilon_0} \int_{\mathbb{R}^3} \frac{\rho(y)(x-y)}{|x-y|^3} dy$$

是无旋的，即 $\nabla \times E = 0$。

B6) (a)（只用区域边界的信息来计算区域的面积）假设光滑闭曲线

$$\gamma : [a, b] \to \mathbb{R}^2, t \mapsto (x(t), y(t))$$

是有界光滑带边区域 $\Omega \subset \mathbb{R}^2$ 的边界。证明，Ω 的面积有如下表达式:

$$m(\Omega) = \frac{1}{2} \int_a^b \left| \det \begin{pmatrix} x & y \\ x'(t) & y'(t) \end{pmatrix} \right| dt$$

(b) 计算下列曲线围成的面积

(1) 星形线 $(a\cos^3 t, a\sin^3 t),\ t \in [0, 2\pi]$;

(2) 双纽线 $(x^2 + y^2)^2 = a^2(x^2 - y^2)$ （提示: 令 $y = x\tan\varphi$);

(3) $\left(\frac{x}{a}\right)^{2n+1} + \left(\frac{y}{b}\right)^{2n+1} = c\left(\frac{x}{a}\right)^n \left(\frac{y}{b}\right)^n$, 其中 $a, b, c > 0,\ n \in \mathbb{Z}_{\geqslant 1}$.

(c) 求环面 $((b + a\cos\theta)\cos\varphi, (b + a\cos\theta)\sin\varphi, a\sin\theta)$ 所围成区域的体积，其中 $0 < a \leqslant b$ 是给定的，参数 $\theta, \varphi \in [0, 2\pi]$。

位势方程的基本解

我们研究 $\mathbb{R}^3$ 位势方程。我用 $x \in \mathbb{R}^3$ 表示空间中的一个点，令 $r = |x|$，$\mathbf{S}^2$ 表示 $\mathbb{R}^3$ 中的单位球面，$d\sigma$ 为 $\mathbf{S}^2$ 上的子流形测度。

C1) 证明，如果 $\varphi(x)$ 是 $\mathbb{R}^3$ 上光滑的有紧支集的函数，那么

$$\int_{\mathbb{R}^3} \varphi(x)dx = \int_0^{+\infty} r^2\big(\int_{\mathbf{S}^2} \varphi(r\vartheta)d\sigma\big)dr.$$

其中，我们用 $\vartheta \in \mathbf{S}^2$ 来表示 $\mathbf{S}^2$ 上的一个点（ϑ 是单位长的向量，所以 $r\vartheta$ 是 $\mathbb{R}^3$ 中的一个点）。

C2) 我们定义 $\mathbb{R}^3$ 上的函数

$$E(x) = -\frac{1}{4\pi|x|},$$

其中在原点处，我们可以任意定义 $E(0)$（比如说等于零。在后面的运算中，一个零测集的值将不会对问题有影响）。

证明，对任意的 $R > 0$，E 在 $B(R)$ 上是 Lebesgue 可积的，其中 $B(R)$ 是以原点为中心以 R 为半径的球。

C3) 对任意的有界的并且在 $\mathbb{R}^3$ 上 Lebesgue 可积函数 f，证明，对任意的 $x_0 \in \mathbb{R}^3$，$\dfrac{f(y)}{|x_0 - y|}$ 对于 dy 是可积函数。从而，我们可以定义函数

$$\Phi_f(x) = -\frac{1}{4\pi}\int_{\mathbb{R}^3} \frac{f(y)}{|x-y|}dy.$$

证明，

$$\lim_{|x|\to\infty} \Phi_f(x) = 0.$$

C4) 对任意的有界的并且在 $\mathbb{R}^3$ 上 Lebesgue 可积函数 f，证明，$\Phi_f(x)$ 是连续函数并且

$$\lim_{|x|\to\infty} \Phi_f(x) = 0.$$

C5*) 对任意 $\mathbb{R}^3$ 上 Lebesgue 可积函数 f，证明，除掉 $\mathbb{R}^3$ 上的一个零测集，$\Phi_f(x)$ 是良好定义的。（提示：去证明 $\Phi_f(x)$ 在每个 $B(R)$ 上都可积）

C6**) 证明，对任意的光滑的有紧支集的函数 $\varphi(x, y, z)$，我们有

$$\int_{\mathbb{R}^3} E(x)\triangle\varphi(x)dx = \varphi(0).$$

（提示：将 $\mathbb{R}^3$ 分为 $B(\varepsilon)$ 和 $\mathbb{R}^3 - B(\varepsilon)$ 两个区域来考虑）

C7) 假设 $f \in C^2(\mathbb{R}^3)$ 并且 f 在 $\mathbb{R}^3$ 上 Lebesgue 可积。证明，$\Phi_f(x) \in C^2(\mathbb{R}^3)$ 并且

$$\triangle\Phi_f(x) = f.$$

C8***) 证明, 如果 f 是 Lebesgue 可积的, 那么存在一个仅依赖于 f 的常数 C, 使得对任意的 $\lambda > 0$, 我们有

$$m\big(x \in \mathbb{R}^3 \big| |\Phi_f(x)| > \lambda\big) \leqslant \frac{C}{\lambda^3}.$$

Les mathématiciens n'étudient pas des objets, mais des relations entre les objets.
(数学家研究的是数学对象之间的关系而不是对象本身)

—- Henri Poincaré

51.2 习题课：Riemann 积分的定义 2

清华大学 19-20 春季学期，数学分析二，第七次习题课内容

传统 Riemann 积分的定义：第二部分

有界区域上的 Riemann 积分

我们沿用上次习题课（在矩形上定义了 Rimeann 积分）的记号。

这次的问题中，A 表示 $\mathbb{R}^n$ 的某个非空有界集。对于 A 上定义的复值（可在某个赋范线性空间上取值）函数

$$f : A \to \mathbb{C},$$

我们将它延拓成 $\mathbb{R}^n$ 的函数：

$$\widetilde{f} : \mathbb{R}^n \to \mathbb{R}, \quad x \mapsto \widetilde{f}(x) = \begin{cases} f(x), & x \in A; \\ 0, & x \notin A. \end{cases}$$

R28) 如果存在包含 A 的矩形 P（总存在），使得 $\widetilde{f}$ 在 P 上的限制 $\widetilde{f}\big|_P$ 是 P 上的 Riemann 可积函数，那么，我们就称 f 在 A 上是 **Riemann 可积的**并记作 $f \in \mathcal{R}(A)$。我们定义

$$\int_A f = \int_P \widetilde{f}\big|_P.$$

证明，这个定义不依赖于矩形 P 的选取，即若 P' 是另一个包含 A 的矩形，那么 $\widetilde{f}$ 在 P' 上的限制 $\widetilde{f}\big|_{P'}$ 是 P' 上的 Riemann 可积的并且

$$\int_P \widetilde{f}\big|_P = \int_{P'} \widetilde{f}\big|_{P'}.$$

（与抽象积分比较，这个积分的定义依赖于函数的延拓。抽象积分的定义是内蕴的。）

R29) 证明，$\mathcal{R}(A)$ 是 $\mathbb{C}$-线性空间。进一步证明，如果 $f, g \in \mathcal{R}(A)$，那么它们的乘积 $f \cdot g \in \mathcal{R}(A)$。

（乘积的性质对抽象积分不成立）

R30) 证明，积分算子

$$\int_A : \mathcal{R}(A) \to \mathbb{C}, \quad f \mapsto \int_A f$$

是 $\mathbb{C}$-线性映射。

（对抽象积分而言，建立线性相对困难）

R31) 证明，对任意的 $f \in \mathcal{R}(A)$，我们有

$$\Big|\int_A f\Big| \leqslant \int_A |f|.$$

如果 $f,g\in\mathcal{R}(A)$ 是实值函数并且对任意的 $x\in A$，都有 $f(x)\leqslant g(x)$，那么

$$\int_A f\leqslant\int_A g.$$

R32) 我们假设 $f,g:A\to\mathbb{R}$ 是实数值有界函数并且 Riemann 可积。证明，函数

$$\max(f,g)(x)=\max(f(x),g(x)),\quad \min(f,g)(x)=\min(f(x),g(x))$$

是 Riemann 可积的并且

$$\int_A\min(f,g)\leqslant\min(\int_A f,\int_A g),$$
$$\max(\int_A f,\int_A g)\leqslant\int_A\max(f,g).$$

R33) 证明 Cauchy-Shwarz 不等式：$f,g\in\mathcal{R}(A)$，我们有

$$\Big|\int_A f\cdot\overline{g}\Big|\leqslant\Big(\int_A|f|^2\Big)^{\frac{1}{2}}\Big(\int_A|g|^2\Big)^{\frac{1}{2}}.$$

（以上 R28)-R33) 是对矩形上的积分理论的简单应用）

R34)（积分第一中值公式）A 是一个可铺集（有限个内部不交的矩形的并集，请参考题目 R2)）假设 f 是 A 上的 Riemann 可积的实值函数。令

$$c=\inf_{x\in A}f(x),\quad C=\sup_{x\in A}f(x).$$

证明，存在 $\xi\in[c,C]$，使得

$$\int_A f=\xi\cdot m(A).$$

R35)（一致收敛与 Riemann 积分可交换）给定 A 上 Riemann 可积的函数列 $\{f_k\}_{k\geqslant1}\subset\mathcal{R}(A)$，假设 $\{f_k\}_{k\geqslant1}$ 一致收敛到函数 $f:A\to\mathbb{C}$。证明，$f\in\mathcal{R}(A)$ 并且

$$\int_A f=\lim_{k\to\infty}\int_A f_k.$$

R36)（上下积分）给定实值有界函数 $f:A\to\mathbb{R}$，任意给定矩形 $P\supset A$，我们考虑

$$\overline{\mathcal{E}_f}(P)=\{\varphi\in\mathcal{E}(P)|\varphi(x)\geqslant\widetilde{f}(x)\},$$
$$\underline{\mathcal{E}_f}(P)=\{\varphi\in\mathcal{E}(P)|\varphi(x)\leqslant\widetilde{f}(x)\}.$$

据此，我们定义

$$\overline{\int_P}f=\inf_{\varphi\in\overline{\mathcal{E}_f}(P)}\int_P\varphi,$$
$$\underline{\int_P}f=\sup_{\varphi\in\underline{\mathcal{E}_f}(P)}\int_P\varphi.$$

证明，这两个数值不依赖于 P 的选取。（我们称之为有界函数 f 的**上积分**和**下积分**并记作 $\overline{\int_A}f$ 和 $\underline{\int_A}f$）

R37) （上下积分刻画 Riemann 性）给定实值有界函数 $f : A \to \mathbb{R}$。证明，$f \in \mathcal{R}(A)$ 当且仅当 $\overline{\int}_A f = \underline{\int}_A f$。此时，我们还有

$$\overline{\int}_A f = \underline{\int}_A f = \int_A f.$$

抽象可铺集

在这一部分中，我们用 A 表示 $\mathbb{R}^n$ 中的一个非空的有界集。我们将研究一类特殊的有界集以及它们上面的 Riemann 积分理论。假设 $A \subset P$，其中 P 是一个矩形，如果示性函数 $\mathbf{1}_A$ 是 P 上的 Riemann 可积的函数，我们就称 A 是**抽象可铺的**。我们定义

$$m(A) = \int_P \mathbf{1}_A$$

为 A 的体积。很明显，这些定义不依赖于 P 的选取。我们用 $\mathscr{P}$ 表示所有的 $\mathbb{R}^n$ 上的抽象可铺集的全体（包括空集）。

R38) 证明，可铺集是抽象可铺的。

R39) 如果 A 和 B 是抽象可铺的，证明，$A \cap B$，$A \cup B$ 和 $A - B$ 也是抽象可铺的并且

$$m(A \cup B) + m(A \cap B) = m(A) + m(B),$$
$$m(A - B) = m(A) - m(A \cap B).$$

（与测度论比较，上述表明所有的 $\mathscr{P}$ 很类似于一个代数（$\mathbb{R}^n \notin \mathscr{P}$）而 m 类似于 $\mathscr{P}$ 上的加性函数）

R40) $A \subset \mathbb{R}^n$ 是非空的有界集，我们强调它未必是抽象可铺的。假设 P 是矩形并且 $A \subset P$，我们定义 A 的**上测度**和**下测度**分别为

$$\overline{m}(A) = \overline{\int}_A \mathbf{1}_A, \quad \underline{m}(A) = \underline{\int}_A \mathbf{1}_A.$$

证明，

$$\overline{m}(A) = \inf_{\substack{B \supset A, \\ B \text{ 是可铺集}}} m(B), \quad \underline{m}(A) = \sup_{\substack{C \subset A, \\ C \text{ 是可铺集}}} m(C).$$

（由此可见，$\overline{m}(A)$ 与 Carathéodory 测度扩张定理证明中的外测度类似，然而，可铺集只容许有限个矩形来覆盖，外测度容许有可数个）

R41) （抽象可铺集的测度刻画）A 是抽象可铺的当且仅当对任意的 $\varepsilon > 0$，存在可铺集 B 和 C，使得

$$B \supset A \supset C, \quad m(C) - m(B) < \varepsilon.$$

（请与课程上关于在紧集上取值有限的测度的正则性定理比较）

R42)（复习）A 是 $\mathbb{R}^n$ 中的集合（未必有界），我们定义 A 的**内部**$\mathring{A}$ 和**闭包**为 $\overline{A}$ 为：

$$\mathring{A} = \{x \in A \big| \text{存在 } \delta > 0\text{，使得以 } x \text{ 为球心以 } \delta \text{ 为半径的小球落在 } A \text{ 中}\},$$
$$\overline{A} = \{x \in \mathbb{R}^n \big| \text{对任意的 } \delta > 0\text{，使得以 } x \text{ 为球心以 } \delta \text{ 为半径的小球与 } A \text{ 的交非空}\}.$$

证明，$\mathring{A}$ 是包含在 A 中的最大的开集；$\overline{A}$ 是包含 A 的最小的闭集并且 $\overline{A}$ 恰为 A 的聚点的集合。

R43) A 是 $\mathbb{R}^n$ 中的集合（未必有界），我们定义 A 的**边界**为 $\partial A = \overline{A} - \mathring{A}$。根据上面的问题，$\partial A$ 是闭集。证明，$\mathbb{R}^n$ 上的函数 $\mathbf{1}_A$ 在每个 $x \in \mathbb{R}^n - \partial A$ 处是连续的。

R45)（抽象可铺集的拓扑刻画）A 表示 $\mathbb{R}^n$ 中的一个非空的有界集。证明，A 是抽象可铺的当且仅当 ∂A 是（Lebesgue 测度）零测集。

R46)（Riemann 意义下的零测集）$A \subset \mathbb{R}^n$，如果对每个 $\varepsilon > 0$，总存在有限个矩形 $P_1, \cdots, P_{N(\varepsilon)}$，使得

$$A \subset \bigcup_{1 \leqslant i \leqslant N(\varepsilon)} P_i, \quad \sum_{i=1}^{N(\varepsilon)} m(P_i) < \varepsilon,$$

我们称 A 是**在 Riemann 意义下的零测集**（简记为$\mathcal{R}$-零测集）。证明，$\mathcal{R}$-零测集是有界的并且是 Lebesgue 意义下的零测集；$\mathcal{R}$-零测集的子集是 $\mathcal{R}$-零测集；有限个 $\mathcal{R}$-零测集的并也是 $\mathcal{R}$-零测集。

R47) 证明，$\mathcal{R}$-零测集的闭包是 $\mathcal{R}$-零测集。进一步证明，$\mathbb{Q} \cap [0,1]$ 不是 $\mathbb{R}^1$ 上的 $\mathcal{R}$-零测集但是是 Lebesgue 意义下的零测集。

R48) 证明，$\mathbb{R}^n$ 中紧的 d-维子流形是 $\mathcal{R}$-零测集，其中 $d \leqslant n-1$。

R49) 证明，如下四个命题是等价的：

甲） A 是 $\mathcal{R}$-零测集；

乙） A 是抽象可铺的并且 $m(A) = 0$；

丙） A 是抽象可铺的并且是零测集；

丁） A 有界并且 $\overline{A}$ 是零测集。

（按照甲 $\Rightarrow$ 乙 $\Rightarrow$ 丙 $\Rightarrow$ 丁 $\Rightarrow$ 甲的顺序推理比较容易）

R50)（可铺集的刻画之三）A 表示 $\mathbb{R}^n$ 中的一个非空的有界集。证明，A 是抽象可铺的当且仅当 ∂A 是 $\mathcal{R}$-零测集。

In my opinion a mathematician, in so far as he is a mathematician, need not preoccupy himself with philosophy – an opinion, moreover, which has been expressed by many philosophers.

————- Henri Lebesgue

52 用光滑函数的逼近 L^1 函数，可分的 Hilbert 空间，Hilbert 基，Fourier 级数的定义，Fourier 级数的 L^2 理论

二零二零年五月十一日，星期一，晴

我们先证明一个经常用到的简单事实:

引理 341. *假设 $(X,\mathcal{A},\mu)$ 是有限测度空间，即 $\mu(X)<\infty$，那么，我们有*

$$L^2(X,\mathcal{A},\mu)\subset L^1(X,\mathcal{A},\mu).$$

证明: 这是 Cauchy-Schwarz 定理的标准应用：对任意的 $f\in L^2(X,\mathcal{A},\mu)$，我们有

$$\begin{aligned}\int_X |f|\mu &= \int_X |f|\cdot 1\mu\\ &\leqslant \|f\|_{L^2}\|1\|_{L^2}\\ &= \sqrt{\mu(X)}\|f\|_{L^2}.\end{aligned}$$

所以，$f\in L^1(X.\mathcal{A},\mu)$。 □

光滑逼近：接上次

为了证明光滑逼近定理，我们选取我们之前构造的非负函数 $\chi(x)\in C_0^\infty(\mathbb{R}^n)$，其中 χ 在 0 点附近恒为 1。我们不妨假设它的积分为 1（通过考虑 $\chi(ax)$ 并选择合适的 $a>0$），即

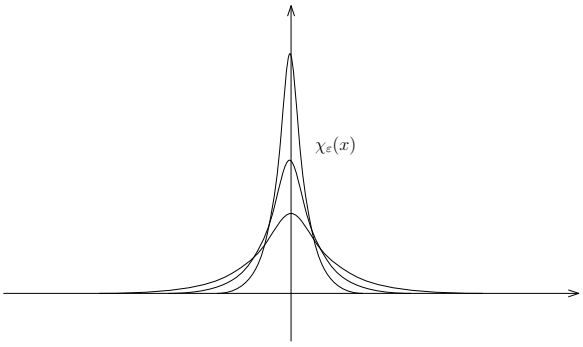
$$\int_{\mathbb{R}^n}\chi(x)dx=1.$$

令

$$\chi_\varepsilon(x)=\frac{1}{\varepsilon^n}\chi(\frac{x}{\varepsilon})$$

那么，

$$\int_{\mathbb{R}^n}\chi_\varepsilon(x)dx=1.$$



当 ε 变小的时候，函数的支集越来越小，函数本身越来越高，请参考上图。

引理 342. 对任意的 $f \in L^1(\mathbb{R}^n)$，当 $\varepsilon \to 0$，我们有

$$\chi_\varepsilon * f \xrightarrow{L^1} f,$$

即

$$\lim_{\varepsilon\to 0} \|\chi_\varepsilon * f - f\|_{L^1(\mathbb{R}^n)} = 0.$$

特别地，$C^\infty(\mathbb{R}^n) \cap L^1(\mathbb{R}^n)$ 是 $L^1(\mathbb{R}^n)$ 的稠密子空间。

证明: 我们之前证明了 $\mathcal{E}(\mathbb{R}^n) \subset L^1(\mathbb{R}^n)$ 是稠密的，我们现在说明，只要对于任意的简单函数 $\varphi \in \mathcal{E}(\mathbb{R})$ 来证明命题就足够了：对任意的 $\delta > 0$，我嗯选取简单函数 φ，使得

$$\|\varphi - f\|_{L^1} < \delta.$$

假设我们有 $\lim\limits_{\varepsilon\to 0} \|\chi_\varepsilon * \varphi - \varphi\|_{L^1(\mathbb{R}^n)} = 0$。那么，存在某个 $\varepsilon_0 > 0$，当 $\varepsilon < \varepsilon_0$ 时，我们有 $\|\chi_\varepsilon * \varphi - \varphi\|_{L^1} < \delta$。所以，

$$\begin{aligned}
\|\chi_\varepsilon * f - f\|_{L^1(\mathbb{R}^n)} &\leqslant \|\chi_\varepsilon * (f-\varphi)\|_{L^1(\mathbb{R}^n)} + \|\chi_\varepsilon * \varphi - \varphi\|_{L^1(\mathbb{R}^n)} + \|f - \varphi\|_{L^1(\mathbb{R}^n)} \\
&\leqslant \|\chi_\varepsilon\|_{L^1} \|f - \varphi\|_{L^1(\mathbb{R}^n)} + \|\chi_\varepsilon * \varphi - \varphi\|_{L^1(\mathbb{R}^n)} + \|f - \varphi\|_{L^1(\mathbb{R}^n)} \\
&< 3\delta.
\end{aligned}$$

这表明 $\lim\limits_{\varepsilon\to 0} \|\chi_\varepsilon * f - f\|_{L^1(\mathbb{R}^n)} = 0$。

我们现在对简单函数证明这个引理。根据线性，只要对 $f = \mathbf{1}_A$ 这样的示性函数证明即可，其中 A 是 Borel 集并且 $m(A) < \infty$（因为这个函数要落在 L^1 中）。另外，由于在 L^1 中，我们有

$$\mathbf{1}_{A\cap B_n} \xrightarrow{L^1} \mathbf{1}_A,$$

所以我们可以假设 A 是有界的 Borel 集合。进一步，根据 Borel 测度的正则性定理[15]，我们可以假设 $A = K$ 为紧集（因为我们可以用闭集从内部逼近 A，而 A 有界）。

我们现在构造一个光滑函数 h，使得 $\|h - \mathbf{1}_K\|_{L^1} < \varepsilon$。为此，我们要再次利用 Stokes 公式第一个证明中的单位分解的技巧：对任意的 $N \geqslant 1$，我们有一族有紧支集的非负的光滑函数 $\{\chi_{\mathbf{k}}(x)\}_{\mathbf{k}\in\Gamma_N}$，其中 $\Gamma_N = 2^{-(N+1)}\mathbb{Z}^n$，使得对每个 $\mathbf{k} \in \Gamma_N$，$\operatorname{supp}(\chi_{\mathbf{k}})$ 落在以 $\mathbf{k}$ 为中心边长为 2^{-N+1} 的正方体中并且：

$$1 = \sum_{\mathbf{k}\in\Gamma_N} \chi_{\mathbf{k}}.$$

[15]我们已经证明了如下的定理：

我们在 $\mathbb{R}^n$ 上的 Borel-代数 $\mathcal{B}(\mathbb{R}^n)$ 上给定满足如下条件的测度 μ：

- 如果 $K \subset \mathbb{R}^n$ 是紧集，我们有 $\mu(K) < \infty$。

那么，对于任意的 $A \in \mathcal{B}(\mathbb{R}^n)$ 和任意的 $\varepsilon > 0$，存在开集 U 包含 A 和被 A 包含的闭集 F（即 $F \subset A \subset U$）使得

$$\mu(U - F) < \varepsilon.$$

这里，N 是待定的。我们令

$$h(x)=\sum_{\operatorname{supp}(\chi_{\mathbf{k}})\cap K\neq\emptyset}\chi_{\mathbf{k}}(x)$$

这是个光滑函数。很明显，h 在 K 上恒为 1（关键点）。另外，如果 $d(x,K)>2\cdot 2^{-N}$，那么 $h(x)=0$。从而，

$$\int_{\mathbb{R}^b}|h-\mathbf{1}_K|\leqslant m(\{x\notin K|d(x,K)\leqslant 2\cdot 2^{-N}\}).$$

由于

$$\bigcap_{N\geqslant 1}\{x\notin K|d(x,K)\leqslant 2\cdot 2^{-N}\}=\emptyset,$$

从而，

$$\lim_{N\to\infty}m(\{x\notin K|d(x,K)\leqslant 2\cdot 2^{-N}\})=0.$$

所以，对充分大的 N，我们就有

$$\|h-\mathbf{1}_K\|_{L^1}<\varepsilon.$$

这就证明了命题。 □

推论 343. *$C_0^\infty(\mathbb{R}^n)\subset L^1(\mathbb{R}^n)$ 是稠密子空间。*

证明： 给定函数 $f\in L^1(\mathbb{R}^n)$。对任意的 $\delta>0$，存在 $R>0$，使得

$$\|f-f\cdot\mathbf{1}_{|x|\leqslant R}\|_{L^1}\leqslant\frac{1}{2}\delta.$$

实际上，我们有

$$\|f-f\cdot\mathbf{1}_{|x|\leqslant R}\|_{L^1}=\int_{|x|\geqslant R}|f(x)|dx.$$

根据 Lebesgue 控制收敛定理（用 $|f|$ 作为控制函数），当 $R\to 0$ 时，上述趋于 0。根据刚刚证明的引理，可以选取 $\varepsilon>0$，使得

$$\|\chi_\varepsilon*\left(f\cdot\mathbf{1}_{|x|\leqslant R}\right)-f\cdot\mathbf{1}_{|x|\leqslant R}\|_{L^1}\leqslant\frac{1}{2}\delta.$$

所以，

$$\|\chi_\varepsilon*\left(f\cdot\mathbf{1}_{|x|\leqslant R}\right)-f\|_{L^1}\leqslant\delta.$$

最终，我们观察到 $\chi_\varepsilon*\left(f\cdot\mathbf{1}_{|x|\leqslant R}\right)$ 的支集是紧的，这是因为 $f\cdot\mathbf{1}_{|x|\leqslant R}$ 的支集是紧的，而卷积可以看成局部上用 χ_ε 来做平均。命题得证。 □

注记. *我们还可以证明 $C_0^\infty(\mathbb{R}^n)\subset L^2(\mathbb{R}^n)$ 是稠密的；$C^\infty(\mathbb{R}^n)\cap L^\infty(\mathbb{R}^n)$ 是 $C(\mathbb{R})\cap L^\infty(\mathbb{R}^n)$ 的稠密子空间。然而，$C(\mathbb{R}^n)\subset L^\infty(\mathbb{R}^n)$ 不稠密。我们将在作业中完成它们的证明。*

由此可见，逼近的意义下，光滑函数可以用来描述这些空间，所以我们很多理论都致力于光滑函数的研究。

Hilbert 空间的基

命题 344. 假设 $(H, \langle\cdot,\cdot\rangle)$ 是 *Hilbert* 空间（完备的内积空间）。如果 $\{x_k\}_{k\geqslant 1} \subset H$ 是一族两两正交的向量，即对任意的 $i, j \geqslant 1$，$i \neq j$，都有 $\langle x_i, x_j\rangle = 0$，那么级数 $\sum\limits_{k=1}^{\infty} x_k$ 在 H 中收敛等价于 $\sum\limits_{k=1}^{\infty} \|x_k\|^2$ 收敛。

证明: 令 $S_n = \sum\limits_{k=1}^{n} x_k$。$\sum\limits_{k=1}^{\infty} x_k$ 在 H 中收敛等价于 $\{S_n\}_{n\geqslant 1}$ 在 H 中收敛，这等价于对任意的 $\varepsilon > 0$，存在 $N > 0$，使得对任意的 $n > m \geqslant N$，$\|S_n - S_m\| < \varepsilon$。这个不等式等价于

$$\langle S_{m+1} + \cdots + S_n, S_{m+1} + \cdots + S_n\rangle < \varepsilon^2.$$

利用正交性，这个不等式展开之后等价于

$$\langle S_{m+1}, S_{m+1}\rangle + \cdots + \langle S_n, S_n\rangle = \|S_{m+1}\|^2 + \cdots + \|S_n\|^2 < \varepsilon^2.$$

这自然是 $\sum\limits_{k=1}^{\infty} \|x_k\|^2$ 收敛的等价条件。 □

定义 345. 给定完备的内积空间 $(H, \langle\cdot,\cdot\rangle)$。如果 $A \subset H$ 是自己，如果它所张成的线性空间（有限个 A 中元素的线性组合）在 H 中是稠密的，即 $\overline{\mathrm{span}(A)} = H$，我们就说 A **在 Hilbert 的意义下张成** H。如果 H 可被某个可数（可以有限）子集张成，我们就称 H 是**可分的** *Hilbert* 空间。如果 $A = \{e_k\}_{k\geqslant 1}$（可数）可以张成 H，并且对任意的 $i, j \geqslant 1$，我们有 $\langle e_i, e_j\rangle = \delta_{ij}$（*Kronecker* 符号），我们就称 $\{e_k\}_{k\geqslant 1}$ 是 H 的一个 **Hilbert 基**。

注记. 请区分，*Hilbert* 基一般不是 H 作为线性空间的基。

定理 346. 每个可分的 *Hilbert* 空间都有 *Hilbert* 基。

证明: 假定 $A = \{x_k\}_{k\geqslant 1}$ 张成 H。我们可以 $x_1, x_2, \cdots, x_n, \cdots$ 依次做 Gram-Schmidt 正交化（每一步涉及到有限个 x_k，从而这和有限维线性空间的理论是一样的），这样，我们就依次得到了 $e_1, e_2, \cdots, e_n, \cdots$。很明显，这是 Hilbert 基，因为它们张成的空间包含了 A。 □

引理 347. $L^2(\mathbb{R}^n)$ 是可分的。

证明: 我们来构造 $L^2(\mathbb{R}^n)$ 得一个可数集合

$$Y = \left\{\mathbf{1}_Q \middle| Q \text{ 为顶点均为有理数坐标的正方体}\right\}.$$

由于简单函数空间 $\mathcal{E}(\mathbb{R}^n)$ 在 $L^2(\mathbb{R}^n)$ 中是稠密的，只要证明 $\mathrm{span}(Y)$ 中的元素可以逼近 $\mathbf{1}_A$ 即可，其中 A 是任意选定的 Borel 集并且 A 的测度有限。类似于本次课第一个定理的证明，我们还可以假设 A 是有界的，从而，利用 Borel 集的正则性定理，我们可以假设 $A = K$ 是紧集，此时，对于紧集，我们可以仿照本次课第一个定理的证明用形如 $\mathbf{1}_Q$ 的正方体的示性函数来逼近，其中我们要求 Q 的变长不超过 2^{-N}，N 可以选取的很大，证明的细节我们留作作业。 □

注记. 通过对上述定理中 Y 中的函数做 *Gram-Schmidt* 正交化，我们可以得到 $L^2(\mathbb{R}^n)$ 的一个 *Hilbert* 基。然而，通过这种比较随意的方式得到的 *Hilbert* 基可能不具有好的性质。分析学的一个很重要的话题就是如何对特定的函数空间构造一个具有特殊性质的基，这样的问题在几何和物理中举足轻重。比如说，通过对调和振动的研究，我们可以构造 $L^2(\mathbb{R}^n)$ 的一个好的 *Hilbert* 基，这个构造的背后既有有意思的分析，还包含了 *Lie* 代数表示的想法（我们在作业中将研究这个问题?）。再比如说，我们的 *Fourier* 级数就是周期的 L^2 函数空间上基，它们是自由振动的特征函数。

定理 348. $(H, \langle\cdot,\cdot\rangle)$ 是可分的 *Hilbert* 空间，$\{e_k\}_{k\geqslant 1}$ 是它一个 *Hilbert* 基。那么，任意的 $x \in H$ 都可以唯一地写成级数的形式：

$$x = \sum_{k=1}^{\infty} c_k e_k = c_1 e_1 + c_2 e_2 + \cdots,$$

其中 $c_i \in \mathbb{C}$。进一步，我们有 $c_k = \langle x, e_k\rangle$ 以及 *Bessel-Parseval* 等式（勾股定理）：

$$\|x\|^2 = \sum_{k=1}^{\infty} |c_k|^2.$$

在证明之前，我们先证明一个简单的引理：

引理 349. $(H, \langle\cdot,\cdot\rangle)$ 是 *Hilbert* 空间。给定 $v \in H$，我们定义映射

$$L_v : H \to \mathbb{C}, \quad x \mapsto \langle x, v\rangle.$$

那么，L_v 是连续线性泛函（从 H 到 $\mathbb{C}$ 的线性映射被称作是线性泛函）。

证明: 用内积的性质可以直接验证 L_v 是线性映射。为了说明 L_v 是连续的，我们用 Cauchy-Schwarz 不等式：

$$|L_v(x)| = \big|\langle v, x\rangle\big| \leqslant \|v\|\|x\|.$$

这就完成了证明。 □

注记. 在泛函分析的课程中，我们将证明 H 上的每个连续线性泛函都形如 L_v，其中 $v \in H$。这是所谓的 *Riesz* 表示定理。

证明: 假设我们有 $x = \sum_{k=1}^{\infty} c_k e_k$。对任意的 ℓ，通过与 e_ℓ 做内积（此时，连续性保证了内积与求和可交换），我们有

$$\begin{aligned}\langle x, e_\ell\rangle &= \sum_{k=1}^{\infty} c_k \langle e_k, e_\ell\rangle \\ &= \sum_{k=1}^{\infty} c_k \delta_k^\ell = c_\ell.\end{aligned}$$

这个计算启发我们如何证明存在性：令 $c_k = \langle x, e_k\rangle$。我们定义部分和 $x_N = \sum\limits_{k\leqslant N} c_k e_k$，其中 $N \geqslant 1$。根据勾股定理，我们有

$$\|x_N\|^2 = \sum_{i,j\leqslant N} c_i\overline{c_j}\langle e_i, e_j\rangle = \sum_{k\leqslant N} |c_k|^2.$$

另外，我们还有

$$\langle x_N, x\rangle = \sum_{k\leqslant N} c_k\langle e_k, x\rangle = \sum_{k\leqslant N} |c_k|^2.$$

所以，

$$\|x_N\|^2 = |\langle x_N, x\rangle| \leqslant \|x_N\|\|x\|.$$

从而，$\|x_N\| \leqslant \|x\|$，即

$$\sum_{k\leqslant N} |c_k|^2 \leqslant \|x\|^2.$$

根据上面的引理，部分和 $\{x_N\}_{N\geqslant 1}$ 收敛，所以，我们可以定义

$$y = \sum_{k=1}^{\infty} c_k e_k.$$

按照唯一性的计算，对任意的 $k \geqslant 1$，我们有

$$\langle y, e_k\rangle = \langle y, e_k\rangle \quad\Rightarrow\quad \langle y - x, e_k\rangle = 0 \quad\Rightarrow\quad x - y \perp e_k.$$

由于 $\{e_k\}_{k\geqslant 1}$ 是一族 Hilbert 基，所以，对任意的 $\varepsilon > 0$，存在 $b_1, \cdots, b_N \in \mathbb{C}$，使得

$$\|\sum_{k\leqslant N} b_k e_k - (y - x)\| < \varepsilon.$$

从而，

$$\begin{aligned}
\|y - x\|^2 &= \langle y - x, (y - x) - \sum_{k\leqslant N} b_k e_k\rangle + \sum_{k\leqslant N} b_k\langle y - x, e_k\rangle \\
&= \langle y - x, (y - x) - \sum_{k\leqslant N} b_k e_k\rangle \\
&\leqslant \|y - x\|\|(y - x) - \sum_{k\leqslant N} b_k e_k\| \\
&\leqslant \varepsilon\|y - x\|.
\end{aligned}$$

令 $\varepsilon \to 0$，这就证明了 $y = x$，从而，x 具有上述级数的形式。Parseval 等式实际上已经蕴含在上面的证明里面：由已经证明的关于 $\|x_N\|^2$ 的等式对 $N \to \infty$ 取极限即可。 □

上面的证明还蕴涵了如下简单的引理：

引理 350 (稠密性的判断)**.** *$(H, \langle\cdot,\cdot\rangle)$ 是 Hilbert 空间，$V \subset H$ 是线性子空间。那么，$V \subset H$ 稠密当且仅当对任意的 $x \in H$，如果 $x \perp V$（即对任意的 $v \in V$，$\langle v, x\rangle = 0$），那么 $x = 0$。*

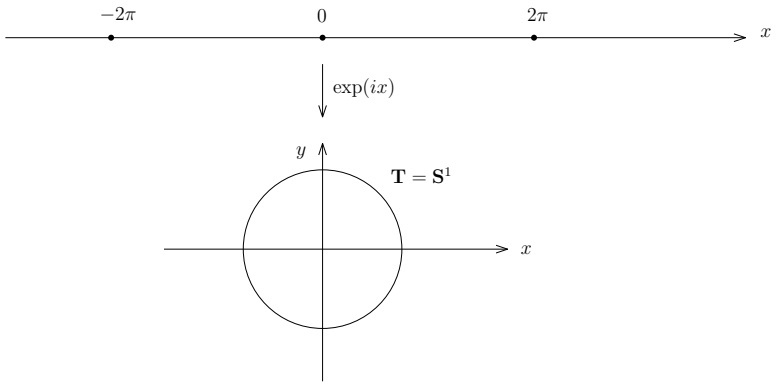
Fourier 级数

我们用 $C_{\mathrm{per},2\pi}(\mathbb{R})$ 表示以 $\mathbb{R}$ 为定义域、在 $\mathbb{C}$ 中取值并且以 2π 为周期的连续函数所构成的复线性空间（这些函数是一致连续的）。按照定义，对于 $f \in C_{\mathrm{per},2\pi}(\mathbb{R})$，对任意的 $x \in \mathbb{R}$，我们有

$$f(x+2\pi) = f(x).$$

我们考虑映射

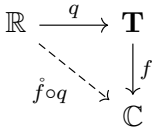
$$q : \mathbb{R} \to \mathbf{S}^1 \subset \mathbb{R}^2 = \mathbb{C}, \quad x \mapsto \exp(ix).$$



我们令 $\mathbf{T} = \mathbf{S}^1 \subset \mathbb{R}^2 = \mathbb{C}$（这是一个 1-维的环面），那么，$q$ 是从 $\mathbb{R}$ 到 $\mathbf{T}$ 的满射并且局部上是微分同胚（因为 $dq(x) \neq 0$）。

通过映射 q，我们可以将 $\mathbb{T}$ 上的连续函数 $\mathring{f}$ 拉回来得到 $\mathbf{T}$ 上的连续函数

$$f = q^*\mathring{f} = \mathring{f} \circ q.$$



很显然，这是一个以 2π 为周期的函数（因为 $e^{2\pi i} = 1$）。所以，我们得到了映射

$$q^* : C(\mathbf{T}) \longrightarrow C_{\mathrm{per},2\pi}(\mathbb{R}), \quad \mathring{f} \mapsto \mathring{f} \circ q,$$

这显然是一个单射。实际上，这还是满射：对任意的 $f \in C_{\mathrm{per},2\pi}(\mathbb{R})$，我们令

$$\mathring{f}(e^{i\theta}) = f(\theta)$$

即可，其中 $\theta \in [0, 2\pi]$。很容易看出 $\mathring{f}$ 是 $\mathbf{T}$ 上良好定义的连续函数并且 $q^*(\mathring{f}) = f$。

命题 351. 映射

$$q^* : C(\mathbf{T}) \longrightarrow C_{\mathrm{per},2\pi}(\mathbb{R})$$

是 $\mathbb{C}$-代数的同构，即这是两个 $\mathbb{C}$-线性空间之间的同构并且对任意的 $\mathring{f}, \mathring{g} \in C(\mathbf{T})$，我们有

$$q^*\left(\mathring{f} \cdot \mathring{g}\right) = q^*\left(\mathring{f}\right) \cdot q^*\left(\mathring{g}\right).$$

证明是平凡的，我们略去。所以，我们可以将 $\mathbb{R}$ 上以 2π 为周期的连续函数和圆周 $\mathbb{T}$ 的连续函数看做同一个数学对象。另外，如果不加说明，我们采取下面的约定：我们总是用 θ 来参数化 $\mathbf{T}$，即考虑

$$[0,2\pi)\to\mathbf{T},\quad \theta\mapsto e^{i\theta}.$$

我们还约定 $\mathbf{T}$ 上的测度为 $\mu=\dfrac{dx}{2\pi}$，这样子，$\mu(\mathbf{T})=1$。

类似地，我们可以考虑 $\mathbb{R}^2$ 以 2π 为周期的连续可微函数，以 2π 为周期的光滑函数，以 2π 为周期的并且在每个周期上可积的函数、以 2π 为周期的并且在每个周期上平方可积的函数和以 2π 为周期的 L^∞ 的函数，这些空间分别对应到 $\mathbf{T}$ 上的空间 $C^1(\mathbf{T})$、$C^\infty(\mathbf{T})$、$L^1(\mathbf{T})$、$L^2(\mathbf{T})$ 和 $L^\infty(\mathbf{T})$。

引理 352. *对任意的 $p=1,2$ 或 ∞，$C^\infty(\mathbf{T})\subset L^p(\mathbf{T})$ 是稠密的子空间。*

证明： 有很多可能的证明，我们采取一个从概念上最简单最直接的证明方法：将卷积推广到 $\mathbf{T}$ 上。我们注意到对于 $\mathbb{T}$ 上的参数化的 $\theta\in[0,2\pi)$（即 $e^{i\theta}\in\mathbf{T}$），我们可以定义他们之间的加法或者减法（用更几何的话说，$\mathbf{T}$ 是一个拓扑群）。首先，我们注意到 $\theta_1+\theta_2\in[0,4\pi)$，据此，我们定义

$$\theta_1\oplus\theta_2=\begin{cases}\theta_1+\theta_2, & \theta_1+\theta_2<2\pi,\\ \theta_1+\theta_2-2\pi, & \theta_1+\theta_2\geqslant 2\pi.\end{cases}$$

这对应着点 $e^{i\theta_1}\cdot e^{i\theta_2}\in\mathbf{T}$。类似的，我们可以定义

$$\theta_1\ominus\theta_2=\begin{cases}\theta_1+\theta_2, & \theta_1-\theta_2\geqslant 0,\\ \theta_1+\theta_2+2\pi, & \theta_1+\theta_2<0.\end{cases}$$

这对应着点 $e^{i\theta_1}\cdot e^{-i\theta_2}\in\mathbf{T}$。换而言之，我们在 $\mathbf{T}$ 上有自然的加法（从参数的观点来看），这个加法实际上来源于 $\mathbf{T}\subset\mathbb{C}$ 上的乘法（因为任意两个模长为 1 的复数的乘积或者商的模长仍然为 1）。

我们可以把 χ_ε（选取比较小的 ε 是的它的支撑集很小（长度不超过 2π））以 2π 为周期延拓成 $\mathbb{R}$ 上的以 2π 为周期的函数，那么，我们就可以认为 $\chi_\varepsilon\in C^\infty(\mathbf{T})$。对任意的 $f\in L^p(\mathbf{T})$，我们定义

$$\chi_\varepsilon * f(\theta)=\int_0^{2\pi}\chi_\varepsilon(\theta-\eta)f(\theta)\frac{d\theta}{2\pi}=\int_{\mathbf{T}}\chi_\varepsilon(zz'^{-1})f(z')d\mu(z').$$

我们现在可以把 $\mathbb{R}$ 上的证明一字不差地照搬过来证明这个结论，证明的细节留给对此仍然持有怀疑态度的同学去验证。我们强调，此时，这个定义对 L^∞ 也成立。 □

我们正式地开始 Fourier 级数的研究。首先，由于在相差一个零测集的情况下，积分理论没有变化，所以，为了研究 $L^2(\mathbf{T})$，我们只要研究 $L^2\big([0,2\pi]\big)$ 即可（这是一个区间的情形）。

定理 353. *在 $L^2\big([0,2\pi],\frac{dx}{2\pi}\big)$ 中，内积 $\langle\cdot,\cdot\rangle$ 的定义为*

$$\langle f,g\rangle=\int_0^{2\pi}f(x)\overline{g(x)}\frac{dx}{2\pi}.$$

函数 $\{e^{ik\cdot x}\}_{k\in\mathbb{Z}}$ 是 Hilbert 基。

证明：首先，通过直接计算可以说明 $\{e^{ik\cdot x}\}_{k\in\mathbb{Z}}$ 由两两正交的单位向量所构成：

$$\langle e^{ik\cdot x}, e^{i\ell\cdot x}\rangle = \frac{1}{2\pi}\int_0^{2\pi} e^{i(k-\ell)\cdot x}dx = \delta_k^\ell.$$

令 $V=\operatorname{span}\big(\{e^{ik\cdot x}\}_{k\in\mathbb{Z}}\big)$，根据我们证明的稠密性判定的引理，为了说明 $V\subset L^2\big([0,2\pi]\big)$ 稠密，只需证明对任意的 $f\in L^2\big([0,2\pi]\big)$，如果 $f\perp V$，即对任意的 $k\in\mathbb{Z}$，$\langle f, e^{ik\cdot x}\rangle = 0$，那么 $f=0$（作为 L^2 中的函数为 0）。

实际上，对任何形如 $\sum\limits_{|k|\leqslant N} c_k e^{ik\cdot x}$ 的函数，f 都和它垂直。根据 Weierstrass-Stone 定理，对任意的连续函数 $\varphi:[0,2\pi]\to\mathbb{C}$，如果 $\varphi(0)=\varphi(2\pi)$，那么对任意的 $\varepsilon>0$，存在形如 $\sum\limits_{|k|\leqslant N} c_k e^{ik\cdot x}$ 三角级数 $S(x)$，使得

$$\|S(x)-\varphi(x)\|_{L^\infty}\leqslant\varepsilon.$$

从而

$$\begin{aligned}|\langle f,\varphi\rangle| = |\langle f,\varphi-S\rangle| &= \Big|\int_0^{2\pi} f(x)\overline{(\varphi-S)(x)}\frac{dx}{2\pi}\Big|\\ &\leqslant \varepsilon\int_0^{2\pi}|f(x)|\frac{dx}{2\pi}\\ &\leqslant \varepsilon\|f\|_{L^2}\|1\|_{L^2} = \varepsilon\|f\|_{L^2}.\end{aligned}$$

令 $\varepsilon\to 0$，这表明对任意的 $\varphi\in C([0,2\pi])$，$\langle f,\varphi\rangle = 0$，即 $f\perp C([0,2\pi])$。然而，$C([0,2\pi])$ 在 $L^2([0,2\pi])$ 是稠密的，所以 $f=0$。 □

53 Fourier 级数的 L^2-理论，高维 Fourier 级数的 L^2-理论，绝对收敛的 Fourier 级数

二零二零年五月十四日，星期四，晴

定理 354. *函数 $\{e^{ik\cdot x}\}_{k\in\mathbb{Z}}$ 是 $L^2(\mathbf{T})$ 的 Hilbert 基。特别地，对任意的 $f(x)\in L^2(\mathbf{T})$，对任意的 $k\in\mathbb{Z}$，令*

$$\widehat{f}(k)=\int_0^{2\pi} f(x)e^{-ik\cdot x}\frac{dx}{2\pi}=\langle f,e^{ik\cdot x}\rangle,$$

那么，在 $L^2(\mathbf{T})$ 中，我们有

$$f(x)=\sum_{k\in\mathbb{Z}}\widehat{f}(k)e^{ik\cdot x}.$$

*其中，$\widehat{f}(k)\in\mathbb{C}$ 被称作是 f 的**第 k 个 Fourier 系数**。进一步，我们还有 Bessel-Parseval 等式（勾股定理）：*

$$\int_0^{2\pi}|f(x)|^2\frac{dx}{2\pi}=\sum_{k=-\infty}^{\infty}\left|\widehat{f}(k)\right|^2.$$

上次课我们用 Weierstrass-Stone 定理给出了上述定理的证明。我们还可以绕过这个定理，直接用积分给出证明：

第二个证明. 令 $V=\overline{\operatorname{span}(\{e^{ik\cdot x}\}_{k\in\mathbb{Z}})}$，只需要证明 $V^\perp=\{0\}$ 即可，现在任选 $f\in V^\perp$，那么，对任意的整数 k，我们有 $\langle f,e^{ik\cdot x}\rangle=0$。证明分为三步：

1) 把问题化归到光滑的情况。

   我们将 f 延拓成为 $\mathbb{R}$ 上以 2π 为周期的函数 $\widetilde{f}$：实际上，对任意的 $x\in\mathbb{R}-2\pi\mathbb{Z}$，存在唯一的 $k\in\mathbb{Z}$，是的 $x-2k\pi\in(0,2\pi)$，我们就定义 $\widetilde{f}(x)=f(x-2k\pi)$；在 $2\pi\mathbb{Z}$ 处，我们要求 $\widetilde{f}(2k\pi)=0$。

   我们注意到，对任意的 x_0，$\widetilde{f}$ 平移之后所得到的函数仍然与 V 垂直，即 $\widetilde{f}(\cdot-x_0)\big|_{[0,2\pi]}\in V^\perp$：实际上，根据换元积分公式以及 exp 的代数性质，我们有

$$\begin{aligned}\int_0^{2\pi} f(x-x_0)e^{ik\cdot x}\frac{dx}{2\pi}&=e^{ik\cdot x_0}\int_0^{2\pi} f(x-x_0)e^{ik\cdot(x-x_0)}\frac{dx}{2\pi}\\&=e^{ik\cdot x_0}\int_0^{2\pi} f(x)e^{ik\cdot x}\frac{dx}{2\pi}\\&=0.\end{aligned}$$

   我们把 $\widetilde{f}$ 打磨成光滑函数：

$$\widetilde{f}_\varepsilon(x)=\chi_\varepsilon*\widetilde{f}\in C^\infty(\mathbf{T}).$$

   根据我们之前的结果，我们有

$$\lim_{\varepsilon\to0}\int_0^{2\pi}|\widetilde{f}_\varepsilon(x)-f(x)|^2\frac{dx}{2\pi}=0.$$

这是 $\widetilde{f}$ 在 L^2 中的一个很好的逼近。实际上，与 V 垂直这个性质也能被保持：

$$\widetilde{f}_\varepsilon(x)|_{[0,2\pi]} \in V^\perp.$$

我们用 Fubini 定理来验证这一点：对任意的 $k \in \mathbb{Z}$，我们有

$$\begin{aligned}\langle \widetilde{f}_\varepsilon(x), e^{ik\cdot x}\rangle &= \int_0^{2\pi} \widetilde{f}_\varepsilon(x) e^{-ik\cdot x}\frac{dx}{2\pi} \\ &= \int_{\mathbb{R}} \frac{1}{\varepsilon}\chi(\frac{y}{\varepsilon}) \underbrace{\Big(\int_0^{2\pi} \widetilde{f}(x-y)e^{ik\cdot x}\frac{dx}{2\pi}\Big)}_{\text{平移之后的积分}} \frac{dy}{2\pi} \\ &= 0.\end{aligned}$$

所以如果能证明 $\widetilde{f}_\varepsilon(x)|_{[0,2\pi]}$ 恒为 0，那么通过让 $\varepsilon \to 0$，我们就可以证明 $f = 0$。从此，可以假设 $f \in C^\infty(\mathbf{T})$ 来证明这个命题即可，这里，我们认为 f 是 $\mathbb{R}$ 上的周期函数。

2) 剩下两步基本的构思如下：我们知道 $\chi_\varepsilon * f$ 逐点的收敛到 f（实际上，如果 χ 只是 L^1 的函数并且 $\int_{\mathbb{R}} \chi(x)dx = 1$，那么结论也成立，请参考本周作业），我们利用这个想法构造一族新的 χ_λ，使得 χ_λ 与 $f(x)$ 的卷积的极限逐点地收敛到 f。然后，再利用 f 在做卷积的时候可以看做是 f 做了一定的平移，从而，f 与三角级数的乘积的积分为零来证明 $f \equiv 0$，当然，我们必须要能够选取 χ_λ 使得它可以用三角级数表示。

对任意的 $\lambda \geqslant 1$，我们定义函数

$$g_\lambda(x) = \sqrt{\frac{\lambda}{2\pi}} \int_{-\pi}^{\pi} e^{-\lambda(1-\cos(y))} f(x+y)dy.$$

注记. 如果我们用 $\chi(x) = \dfrac{e^{-\frac{x^2}{2}}}{\sqrt{2\pi}}$，那么这个函数的积分是 1。此时，我们可以用令 $\lambda = \dfrac{1}{\sqrt{\varepsilon}}$，那么，利用 χ 是偶函数，我们有

$$\begin{aligned}\chi_\varepsilon * f &= \sqrt{\frac{\lambda}{2\pi}} \int_{\mathbb{R}} e^{-\frac{\lambda y^2}{2}} f(x-y)dy \\ &= \sqrt{\frac{\lambda}{2\pi}} \int_{\mathbb{R}} e^{-\frac{\lambda y^2}{2}} f(x+y)dy \\ &= \sqrt{\frac{\lambda}{2\pi}} \int_{-\pi}^{\pi} e^{-\frac{\lambda y^2}{2}} f(x+y)dy + \underbrace{\sqrt{\frac{\lambda}{2\pi}} \int_{|y|\geqslant\pi} e^{-\frac{\lambda y^2}{2}} f(x+y)dy}_{\mathbf{I}}.\end{aligned}$$

对于 I，我们来说明当 I 很大的时候，它可以被忽略：直观上，这是显然的，因为当 $\varepsilon \to 0$ 时候，几乎所有的 χ_ε 的贡献都集中在 0 附近的小邻域中。利用 f 有界，我们有

$$\begin{aligned}\mathbf{I} &\leqslant \sqrt{\frac{\lambda}{2\pi}} e^{-\frac{\frac{\lambda}{2}\pi^2}{2}} \int_{|y|\geqslant\pi} e^{-\frac{\frac{\lambda}{2}y^2}{2}} f(x+y)dy \\ &\leqslant \underbrace{\sqrt{\frac{\lambda}{2\pi}} e^{-\frac{\frac{\lambda}{2}\pi^2}{2}}}_{\text{趋于 } 0,\ \lambda\to\infty} \underbrace{\int_{|y|\geqslant\pi} e^{-\frac{\frac{1}{2}y^2}{2}} \|f\|_{L^\infty} dy}_{\text{常数}}\end{aligned}$$

所以，当 $\lambda \to \infty$ 时，我们有

$$\sqrt{\frac{\lambda}{2\pi}} \int_{-\pi}^{\pi} e^{-\frac{\lambda y^2}{2}} f(x+y)dy \to f(x).$$

（多努力一点我们还可以证明这是一致收敛的）然而，在周期的情况下（这是为什么我们要限制到 $[-\pi, \pi]$ 上积分），$e^{-\frac{y^2}{2}}$ 不是周期函数，为此，由于最终当 $\varepsilon \to 0$ 时，上面函数都集中在零点处，所以我们用 $1-\cos(y)$ 来代替 $\dfrac{y^2}{2}$（*Taylor* 展开），这基本上解释了 g_λ 的构造。

回到命题本身，证明如下的一致收敛性：

$$\lim_{\lambda \to \infty} g_\lambda(x) = f(x), \quad x \in \mathbb{R}.$$

我们把积分砍成两块：

$$g_\lambda(x) = \underbrace{\sqrt{\lambda} \int_{|y| \leqslant \frac{\pi}{4}} e^{-\lambda(1-\cos(y))} f(x+y)dy}_{\mathbf{A}} + \underbrace{\sqrt{\lambda} \int_{\pi \geqslant |y| \geqslant \frac{\pi}{4}} e^{-\lambda(1-\cos(y))} f(x+y)dy}_{\mathbf{B}}.$$

我们首先来控制第二项，它与上面注解中的 **I** 类似：

$$\begin{aligned} |\mathbf{B}| &\leqslant \sqrt{\lambda} \int_{\pi \geqslant |y| \geqslant \frac{\pi}{4}} e^{-\lambda(1-\frac{\sqrt{2}}{2})} \|f\|_{L^\infty} dy \\ &= \frac{3\pi}{2} \sqrt{\lambda} e^{-\lambda(1-\frac{\sqrt{2}}{2})} \|f\|_{L^\infty} \to 0. \end{aligned}$$

很明显，上面的收敛是不依赖于 x 的，所以是一致的。

对于 $\mathbf{A}$，由于 $1-\cos(y) > 0$，我们做变量替换 $1-\cos(y) = z^2$。令 $z_0 = \sqrt{1-\dfrac{\sqrt{2}}{2}}$，那么

$$\begin{aligned} \mathbf{A} &= \sqrt{\frac{\lambda}{2\pi}} \int_{|z| \leqslant z_0} e^{-\lambda z^2} \underbrace{f(x+y(z))y'(z)}_{F(z)} dz \\ &= \sqrt{\frac{\lambda}{2\pi}} \int_{\mathbb{R}} e^{-\lambda z^2} F(z) dz, \end{aligned}$$

其中 $F(z)$ 在 $|z| \geqslant z_0$ 处用 0 来延拓。从而，我们有

$$\begin{aligned} \mathbf{A} - F(0) &= \sqrt{\frac{\lambda}{2\pi}} \int_{\mathbb{R}} e^{-\lambda z^2} \left(F(z) - F(0)\right) dz \\ &= \underbrace{\sqrt{\frac{\lambda}{2\pi}} \int_{|z| \leqslant \varepsilon} e^{-\lambda z^2} \left(F(z) - F(0)\right) dz}_{\mathbf{C}} + \underbrace{\sqrt{\frac{\lambda}{2\pi}} \int_{|x| \geqslant \varepsilon} e^{-\lambda z^2} \left(F(z) - F(0)\right) dz}_{\mathbf{II}}. \end{aligned}$$

其中 ε 是任意选取的。类似于注解中的 **I**，如果选取好了 ε，那么 **II** 一致地收敛到 0（当 $\lambda \to \infty$ 时）；对于 **C**，利用 F 在 0 附近的 Lagrange 中值定理（F 在 0 附近可微分并且导

数有界 M，这里要求我们选取足够小的 ε)，我们有

$$\begin{aligned}\mathbf{C} &\leqslant \sqrt{\frac{\lambda}{2\pi}} \int_{|z|\leqslant\varepsilon} e^{-\lambda z^2} M|z|dz \\ &= \frac{M}{\sqrt{\lambda}} \underbrace{\sqrt{\frac{1}{2\pi}} \int_{|t|\leqslant\varepsilon} e^{-t^2} M|t|dt}_{\text{积分有限}} \to 0.\end{aligned}$$

所以也一致收敛到 0。从而，我们有如下的一致收敛：

$$\mathbf{A} \to F(0) = f(x)$$

3) 我们证明，对任意固定的 λ，我们都有 $g_\lambda(x) \equiv 0$。

实际上，可以用级数的形式将 g_λ 展开（从而化成了三角级数的形式)：

$$g_\lambda(x) = \sqrt{\frac{\lambda}{2\pi}} e^{-\lambda} \int_{-\pi}^{\pi} \sum_{k\geqslant 0} \frac{\lambda^k \cos^k(y)}{k!} f(x+y)dy.$$

利用 Lebesgue 控制收敛定理，我们可以交换积分与求和得到

$$\begin{aligned}g_\lambda(x) &= \sum_{k\geqslant 0} \sqrt{\frac{\lambda}{2\pi}} e^{-\lambda} \int_{-\pi}^{\pi} \frac{\lambda^k \cos^k(y)}{k!} f(x+y)dy \\ &= \sum_{k\geqslant 0} \sqrt{\frac{\lambda}{2\pi}} e^{-\lambda} \int_{0}^{2\pi} \frac{\lambda^k \cos^k(y)}{k!} f(x+y)dy.\end{aligned}$$

最后一步用到了周期性。然而，$\cos(y)$ 可以写为 e^{iy} 和 e^{-iy} 的线性组合，所以，$\cos^k(y)$ 是有限个 $e^{ik\cdot y}$ 的线性组合，从而，利用 $f(y+x)$（平移了一下）与它们每个都正交，我们就知道每个积分项都是 0，从而 $g_\lambda \equiv 0$

这就给出了第二个证明。 □

高维的 Fourier 级数

假设 $f(x_1, x_2, \cdots, x_n)$ 是 $\mathbb{R}^n$ 上的连续函数，其中，我们在 $\mathbb{R}^n$ 上选好了坐标系统 $(x_1, \cdots, x_n)$。如果对任意的 $(x_1, x_2, \cdots, x_n)$ 和 $i \leqslant 1$，我们都有

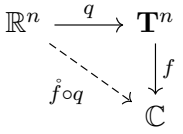
$$f(x_1, \cdots, x_{i-1}, x_i + 2\pi, x_{i+1}, \cdots, x_n) = f(x_1, \cdots, x_{i-1}, x_i, x_{i+1}, \cdots, x_n),$$

我们就称 f 是以 $2\pi\mathbb{Z}^n$ 为周期的。我们用 $C_{\text{per},2\pi}(\mathbb{R}^n)$ 表示所有这样的函数。

我们考虑映射

$$q : \mathbb{R}^n \to \mathbf{T}^n = \underbrace{\mathbf{T} \times \cdots \times \mathbf{T}}_{n\text{ 个}}, \quad (x_1, \cdots, x_n) \mapsto (\exp(ix_1), \cdots, \exp(ix_2)).$$

这里，$\mathbf{T}^n$ 是 b-维的环面并且 q 是满射并且局部上是微分同胚。类似于 1 维，通过映射 q 的拉回，我们有



命题 355. 映射

$$q^*: C(\mathbf{T}^n) \longrightarrow C_{\mathrm{per},2\pi}(\mathbb{R}^n), \quad \mathring{f} \mapsto q^*\left(\mathring{f}\right),$$

是 $\mathbb{C}$-代数的同构。

我们总是用 $(\theta_1, \cdots, \theta_n)$ 来参数化 $\mathbf{T}^n$：

$$[0, 2\pi) \times \cdots \times [0, 2\pi) \to \mathbf{T}^n, \quad (\theta_1, \cdots, \theta_n) \mapsto \left(e^{i\theta_1}, \cdots, e^{i\theta_n}\right).$$

我们约定 $\mathbf{T}^n$ 上的测度为 $\mu = \dfrac{d\theta_1 \cdots d\theta_n}{(2\pi)^n}$，这可以保证 $\mu(\mathbf{T}^n) = 1$。

类似地，$\mathbb{R}^n$ 以 $2\pi\mathbb{Z}$ 为周期的连续可微函数、光滑函数，每个周期上可积的函数、在每个周期上平方可积的函数和 L^∞ 的函数分别对应到 $\mathbf{T}^n$ 上的 $C^1(\mathbf{T}^n)$、$C^\infty(\mathbf{T}^n)$、$L^1(\mathbf{T}^n)$、$L^2(\mathbf{T}^n)$ 和 $L^\infty(\mathbf{T}^n)$。

类似于 1 维的情形，我们很容易证明

引理 356. *对任意的 $p = 1$ 或 2，$C^\infty(\mathbf{T}^n) \subset L^p(\mathbf{T}^n)$ 是稠密的子空间。*

和 1 维类似，平方可积的 $2\pi\mathbb{Z}$ 周期函数的 Fourier 级数理论具有干净的表达：

定理 357. *在 $L^2\left([0, 2\pi]^n, \frac{dx_1 \cdots dx_n}{(2\pi)^n}\right)$ 中，内积 $\langle \cdot, \cdot \rangle$ 的定义为*

$$\langle f, g \rangle = \int_{[0,2\pi]^n} f(x)\overline{g(x)} \frac{dx}{(2\pi)^n}.$$

那么，函数 $\{e^{ik\cdot x}\}_{k\in\mathbb{Z}^n}$ 是 Hilbert 基，其中

$$k = (k_1, \cdots, k_n), \quad x = (x_1, \ldots, x_n), \quad k \cdot x = \sum_{i=1}^n k_i x_i.$$

特别地，对任意的 $f(x) \in L^2\left(\mathbf{T}^n\right)$，对任意的 $k \in \mathbb{Z}^n$，令

$$\widehat{f}(k) = \int_{[0,2\pi]^n} f(x) e^{-ik\cdot x} \frac{dx}{(2\pi)^n} = \langle f, e^{ik\cdot x} \rangle,$$

那么，在 $L^2\left(\mathbf{T}^n\right)$ 中，我们有

$$f(x) = \sum_{k\in\mathbb{Z}} \widehat{f}(k) e^{ik\cdot x}.$$

*其中，$\widehat{f}(k) \in \mathbb{C}$ 被称作是 f 的**第 k 个 Fourier 系数**。进一步，我们还有 Bessel-Parseval 等式（勾股定理）：*

$$\int_{[0,2\pi]^n} |f(x)|^2 \frac{dx}{(2\pi)^n} = \sum_{k\in\mathbb{Z}^n} \left|\widehat{f}(k)\right|^2.$$

证明: 我们只要证明 $\{e^{ik\cdot x}\}_{k\in\mathbb{Z}^n}$ 是 Hilbert 基即可，其余的是显然的。令 $V=\mathrm{span}(\{e^{ik\cdot x}\}_{k\in\mathbb{Z}^n})$，假设 $f\in V^{\perp}$，我们要说明在 L^2 中，$f=0$。为了简化书写，我们采取如下简约的记号:

$$k=(k_1,\underbrace{k_2,\cdots,k_n}_{k'})=(k_1,k'),\quad x=(x_1,\underbrace{x_2,\cdots,x_n}_{x'})=(x_1,x').$$

由于 $e^{ik\cdot x}f(x)\in L^2([0,2\pi]^n)\subset L^1([0,2\pi]^n)$，所以我们可以运用 Fubini 定理从 $\langle f,e^{ik\cdot x}\rangle=0$ 推出

$$\int_{[0,2\pi]}e^{ik_1x_1}\underbrace{\Big(\int_{[0,2\pi]^{n-1}}e^{ik'\cdot x'}f(x_1,x')dx'\Big)}_{F(x_1)}dx_1=0.$$

另外，对 $|f(x)|^2$ 的积分用 Fubini 定理，我们知道 $F(x_1)$ 对几乎处处的 x_1 有定义并且 $F(x_1)\in L^2([0,2\pi])$，所以，根据 1 维的 L^2 的 Fourier 级数的结论，我们

$$F(x_1)=0.$$

利用归纳我们就完成了证明。 □

注记. *在 L^2-空间框架下的 Fourier 级数理论很大程度上类比了我们对于有限维内积空间的理解，这是用线性代数的观点来研究函数的一个重要例子: 我们可以把每个 L^2 函数 f 看成若干个更基本的函数 $e^{ik\cdot x}$（简单波）的线性叠加，从而期盼对每个 $e^{ik\cdot x}$（以及它们之间的关系）的理解能够帮助我们读出 f 的关键信息。特别地，我们希望通过研究 f 的 Fourier 系数 $\widehat{f}(k)$ 的性质来研究 f 本身，这就是 Fourier 分析。*

如果我们知道函数 $f\in L^2(\mathbf{T}^n)$，那么，级数

$$\sum_{k\in\mathbb{Z}}\widehat{f}(k)e^{ik\cdot x}$$

在 L^2 中收敛到 f。根据我们对 L^2 完备性的证明，我们知道存在下列函数列的子序列:

$$\{S_N=\sum_{k\in\mathbb{Z},|k|\leqslant N}\widehat{f}(k)e^{ik\cdot x}\}_{N\geqslant 1},$$

它几乎处处收敛到 f。然而，是否整个序列都几乎处处（不需要选取子序列）收敛到 f 呢? 这是 Carleson 证明的重要定理，他本人也因为这项工作获得了 Abel 奖。

实际上，即使我们假设了 $f\in C(\mathbf{T}^n)\subset L^2(\mathbf{T}^2)$，我们也只能说

$$\sum_{k\in\mathbb{Z}}\widehat{f}(k)e^{ik\cdot x}$$

在 L^2 中收敛到 f，这个级数不见得逐点收敛。这是 Fourier 级数的分析中重要的话题（我们将要学习），另一方面，这也表明了为什么 L^2 的理论是重要的: 因为我们总能保证基本的收敛性!

另外，对于 $f\in L^1\big(\mathbf{T}^n,\frac{dx}{(2\pi)^n}\big)$，它的 Fourier 系数

$$\widehat{f}(k)=\int_{\mathbb{R}^n}f(x)e^{-ik\cdot x}\frac{dx}{(2\pi)^n},$$

仍然有定义（可积）。但是，在 $L^1\big(\mathbf{T}^n, \frac{dx}{(2\pi)^n}\big)$ 的意义下，我们缺未必有

$$f(x) = \sum_{k\in\mathbb{Z}^n} \widehat{f}(k) e^{ik\cdot x}.$$

Kolmogorov 在大学二年级时构造了一个 L^1 的函数（1 维情形），使得其 *Fourier* 级数（上面等式的右边）几乎处处发散，我们会在作业题中研究他所构造的函数（*2019* 年数学分析二的期末考试题）。

我们通过下面的简单例子说明 L^2 理论的有效性：

例子. 假设 $f \in C(\mathbf{T}^n)$，如果 f 所有 *Fourier* 系数都是 0，那么 $f \equiv= 0$：

实际上，根据 L^2 的理论，*Fourier* 系数全部消失意味着在 L^2 中 $f = 0$，所以 f 几乎处处为 0。根据 f 的连续性，这说明 $f \equiv 0$：否则，存在 $x_0 \in \mathbb{T}^n$，$f(x_0) \neq 0$，那么，f 在 x_0 的一个小的开邻域上不为零，但是任意开邻域的测度不为 0，矛盾。

Fourier 分析

我们现在来学习古典的 Fourier 分析，大概的思路是通过一个函数的 Fourier 系数 $\{\widehat{f}(k)\}_{k\in\mathbb{Z}^n}$ 来研究函数的性质。一般而言，我们给定的周期函数 f 都是 L^1 的，此时，我们把如下的序列

$$\left\{\widehat{f}(k)\Big|\widehat{f}(k) = \int_0^{2\pi} f(x)e^{-ik\cdot x}\frac{dx}{(2\pi)^n}\right\}_{x\in\mathbb{Z}^n}$$

称作是 f 的 **Fourier 系数**。由于我们假设了 f 的可积性，所以上面序列是良好定义的（下学期我们还会研究 f 是不可积的情形）。所以，在后面的讲义中，我们一般都假设 f 是可积的（除非有特殊情况）。

注记. 我们之后所要证明的定理，大部分都可以原封不动的推广到 $\mathbf{T}^n$ 上去。为了行文方便，我们只介绍 1 维的理论。当你遇到高维的例子的时候（在作业或者考试中），你应该能够回来验证同样的论证也成立。

我们先从一个上个学期就已经（隐含地）证明过的结论开始：

定理 358. 对任意的连续函数 $f \in C(\mathbf{T})$，如果它对应的 *Fourier* 级数

$$\sum_{k\in\mathbb{Z}} \widehat{f}(k)e^{ikx}$$

是绝对收敛的，即

$$\sum_{k\in\mathbb{Z}} \big|\widehat{f}(k)\big| < \infty, \qquad \text{（注意，这只是对 } \textit{Fourier} \text{ 系数的要求）}$$

那么部分和 d 定义的函数序列 $\{S_N(x) = \sum_{-N\leqslant k\leqslant N} \widehat{f}(k)e^{ikx}\}_{N\geqslant 1}$ 是一致收敛的：对任意的 $x \in \mathbf{T}$，$\lim_{n\to\infty} S_N(x) = f(x)$ 并且

$$\lim_{N\to\infty} \|S_N - f\|_{L^\infty} = 0.$$

证明：我们在 $(C(\mathbf{T}), \|\cdot\|_{L^\infty})$ 上考虑这个问题。我们注意到：

$$\sum_{k\in\mathbb{Z}} \left|\widehat{f}(k)\right| = \sum_{k\in\mathbb{Z}} \|\widehat{f}(k)e^{ikx}\|_{L^\infty}.$$

所以，这个函数级数在 $(C(\mathbf{T}), \|\cdot\|_{L^\infty})$ 中绝对收敛，从而

$$g(x) = \lim_{N\to\infty} S_N(x) = \lim_{N\to\infty} \sum_{-N\leqslant k\leqslant N} \widehat{f}(k)e^{ikx}$$

是良好定义的并且是 $\mathbf{T}$ 上连续函数（或者 $\mathbb{R}$ 上的连续周期函数）。以下只要证明 $f = g$ 即可。

由于 f 和 g 都是连续函数（所以是 L^2 的），根据 L^2 的理论，只要验证它们具有相同的 Fourier 系数是即可。为此，我们考虑函数

我们计算 $g(x)$ 的 Fourier 系数：

$$\begin{aligned}
\widehat{g}(k) &= \int_0^{2\pi} \left(\lim_{N\to\infty} \sum_{-N\leqslant \ell\leqslant N} \widehat{f}(\ell)e^{i\ell)x} \right) e^{-ikx}dx \\
&= \int_0^{2\pi} \lim_{N\to\infty} \sum_{-N\leqslant \ell\leqslant N} \widehat{f}(\ell)e^{-i(k-\ell)x}dx \\
&= \lim_{N\to\infty} \int_0^{2\pi} \sum_{-N\leqslant \ell\leqslant N} \widehat{f}(\ell)e^{-i(k-\ell)x}dx \\
&= \lim_{\substack{N\to\infty \\ N\geqslant|k|}} \widehat{f}(k) = \widehat{f}(k).
\end{aligned}$$

上面计算中积分与极限可交换是因为上学期我们证明过一致收敛与积分可交换（也可以用 Lebesgue 控制收敛定理直接来证明，取 $\sum_{k\in\mathbb{Z}} \left|\widehat{f}(k)\right|$ 作为控制函数）。这就完成了证明。 □

53.1 作业：波动方程的局部能量估计

清华大学 19-20 春季学期，数学分析二，作业 10

请用 A4 大小的纸张**正反面**用钢笔，签字笔或者圆珠笔书写，并注明自己的姓名，年级和作业的**总页数**。除定理公式所涉及的人名之外，请使用**中文**。本次作业的提交时间和地点为 **2 月 106 日**上午的课堂上，逾期视作零分。

课堂补充

A1)（闭子空间）$(H, \langle\cdot,\cdot\rangle$ 是 Hilbert 空间，$V \subset H$ 是其线性子空间。我们假设 V 是 H 中的闭集（相对于 H 上由内积所定义的距离函数而言）。证明，对任意的 $v_1, v_2 \in V$，如果我们仍用 $\langle v_1, v_2\rangle$ 作为 V 上的内积，那么 $\big(V, \langle\cdot,\cdot\rangle\big)$ 是 Hilbert 空间。

证明，对于 $H = L^2(\mathbb{R}^n)$，$V = C(\mathbb{R}^n) \cap L^2(\mathbb{R}^n)$ 不是闭子空间。

A2) 给定 $L^2(X, \mathcal{A}, \mu)$ 中的函数序列 $\{f_i\}_{i\geqslant 1}$，我们假设它们在 L^2 范数下收敛到 f，即

$$f_i \overset{L^1(X,\mathcal{A},\mu)}{\longrightarrow} f,\ i \to \infty.$$

那么，存在子函数序列 $\{f_{i_p}\}_{p\geqslant 1}$，使得对几乎处处的 $x \in X$，我们都有

$$\lim_{p\to\infty} f_{i_p}(x) = f(x).$$

A3) 证明，$\big(L^1(\mathbb{R}^n), *\big)$ 中没有单位元，即不存在 $e \in L^1(\mathbb{R}^n)$，使得对任意的 $f \in L^1(\mathbb{R}^n)$，都有 $e * f = f$。

A4) 证明，$C_0^\infty(\mathbb{R}^n)$ 是 $L^2(\mathbb{R}^n)$ 的稠密子空间。

A5) 举例说明 $C(\mathbb{R}^n) \subset L^\infty(\mathbb{R}^n)$ 不是稠密的。

A6) 证明，$C^\infty(\mathbb{R}^n) \cap L^\infty(\mathbb{R}^n)$ 在 $C(\mathbb{R}^n) \cap L^\infty(\mathbb{R}^n)$ 中稠密（对 $\|\cdot\|_{L^\infty}$ 而言）。

A7) 假设 $(H, \langle\cdot,\cdot\rangle$ 是 Hilbert 空间，$V \subset H$ 是线性子空间。证明，V 在 H 中稠密当且仅当对任意的 $x \perp V$ 一定有 $x = 0$。

A8) 试证明，$L^2(\mathbb{R}^n)$ 是可分的。（可以参考讲义中的提示）

A9)（重要）假设 $f \in L^1(\mathbb{R}^n)$ 并且 $\int_{\mathbb{R}} \chi(x)dx = 1$，我们定义

$$f_\varepsilon(x) = \frac{1}{\varepsilon^n} f\left(\frac{x}{\varepsilon}\right)$$

证明，对任意的有紧支集的连续函数 φ，我们都有如下的一致收敛：

$$f_\varepsilon * \varphi \to \varphi,\ \ \varepsilon \to 0.$$

控制收敛与 Fubini 定理的复习

B. 含参数积分的研究

B1) 证明，对任意的 $t \in \mathbb{R}$，如下积分是良好定义的：

$$F(t) = \int_0^\infty \frac{1 + tx^5 + t^2x^{11}}{1 + t^4 + x^{13}} dx.$$

B2) 证明，$F(t)$ 是 $t \in \mathbb{R}$ 的连续函数。

B3) 证明，$F(t)$ 是可导的。

B4) 试计算 $\lim\limits_{t \to +\infty} F(t)$。

B5) 证明，对 $t \neq 0$，如下的积分是发散的：

$$\int_0^\infty \frac{1 + tx^5 + t^2x^{11}}{1 + t^4 + x^{12}} dx$$

B6) 证明，对任意的 $t \in \mathbb{R}$，如下在 Riemann 意义下的反常积分是良好定义的：

$$G(t) = \int_0^\infty \frac{1 + tx^5 + t^2x^{11}}{1 + t^4 + x^{12}} e^{ix} dx,$$

即 $\lim\limits_{T \to \infty} \int_0^T \frac{1 + tx^5 + t^2x^{11}}{1 + t^4 + x^{12}} e^{ix} dx$ 存在。

B7) 证明，$G(t)$ 是 $t \in \mathbb{R}$ 的连续函数。

B8) 试计算 $\lim\limits_{t \to +\infty} G(t)$。

C. Laplace 变换的基本性质

C1) 假设 $f(t)$ 是 $\mathbb{R}_{>0}$ 上定义的 Lebesgue 可积函数。我们定义它的 Laplace 变换为：

$$\big(\mathcal{L} f\big)(x) = \int_0^\infty f(t) e^{-xt} dt,$$

其中 $x > 0$。证明，$\big(\mathcal{L} f\big)(x)$ 是 C^∞ 的函数。

C2) 证明，对任意的 $n \in \mathbb{Z}_{\geqslant 0}$，我们都有

$$\left(\frac{d}{dx}\right)^n \big(\mathcal{L} f\big)(x) = \int_0^\infty (-t)^n f(t) e^{-xt} dt.$$

C3) 假设 $f(t)$ 和 $g(t)$ 是 $\mathbb{R}_{>0}$ 上定义的 Lebesgue 可积函数，我们定义它们的“卷积”为

$$\big(f \bullet g\big)(t) = \int_0^t f(\tau) g(t - \tau) d\tau.$$

证明，$f \bullet g = g \bullet f$。

C4) 假设 $f(t)$ 和 $g(t)$ 是 $\mathbb{R}_{>0}$ 上定义的 Lebesgue 可积函数。证明，

$$\mathcal{L}\big(f \bullet g\big)(x) = \big(\mathcal{L} f\big)(x) \big(\mathcal{L} g\big)(x).$$

Stokes 公式的应用：波动方程的局部能量估计

假设 $u(t,x)$ 是 $\mathbb{R}\times\mathbb{R}^3$ 上的光滑函数，这里空间变量 $x\in\mathbb{R}^3$，时间变凉 $t\in\mathbb{R}$。我们假设 u 满足波动方程：

$$\Box u=0,$$

其中波动算子的定义如下：

$$\Box=-\frac{\partial^2}{\partial t^2}+\triangle=-\frac{\partial^2}{\partial t^2}+\frac{\partial^2}{\partial x_1^2}+\frac{\partial^2}{\partial x_2^2}+\frac{\partial^2}{\partial x_3^2}.$$

我们令 $u_0(x)=u(0,x)$，$u_1(x)=\dfrac{\partial u}{\partial t}(0,x)$，这样的一对函数 (u_0,u_1) 被称作是波动方程的**初始值**。

我们做如下的假设：对任意的时间 $t\in\mathbb{R}$，$u(t,x)$ 对 x 而言有紧支集。

D1) 我们定义波动方程的能量函数

$$E(t)=\frac{1}{2}\int_{\mathbb{R}^3}|\frac{\partial u}{\partial t}(t,x)|^2+|\nabla u(t,x)|^2dx.$$

证明能量守恒定律：对任意的时刻 $t\in\mathbb{R}$，有

$$E(t)=E(0).$$

D2) 对于 $i=1,2,3$，我们定义波动方程的动量函数

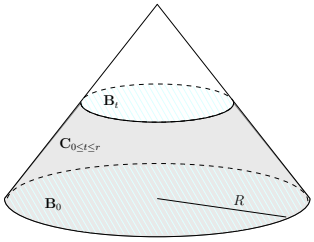
$$P_i(t)=\int_{\mathbb{R}^3}\frac{\partial u}{\partial t}\frac{\partial u}{\partial x_i}dx.$$

证明动量守恒定律：对任意的时间 t，有

$$P_i(t)=P_i(0).$$

D3)（解的唯一性）假设 $u(t,x)$ 和 $v(t,x)$ 都满足波动方程以及上面假设。证明，如果它们具有相同的的初始值，即 $u_0(x)=v_0(x)$，$u_1(x)=v_1(x)$，那么 $u(t,x)\equiv v(t,x)$。

D4) 任意给定 $R>0$，假设 $r\leqslant R$。我们在时空 $\mathbb{R}\times\mathbb{R}^3$ 中定义如下的几何对象：



– 实心锥 $\mathbf{D}=\{(t,x)|t\in[0,R],|x|\leqslant R-t\}$。这是顶点在 $(t,x)=(R,0)$ 处的一个实心锥。

- 实心圆台 $\mathbf{D}_{0\leqslant t\leqslant r}=\{(t,x)|t\in[0,r],|x|\leqslant R-t\}$（图中两个淡蓝色圆盘之间的区域）。
- 圆锥截面 $\mathbf{C}_{0\leqslant t\leqslant r}=\{(t,x)|t\in[0,r],|x|=R-t\}$（图中的灰色区域所代表的表面）。
- 不同时刻的球面 $\mathbf{B}_t=\{(t,x)||x|\leqslant R-t\}$（图中淡蓝色的圆盘区域）。

给定 $(t_0,x_0)\in\mathbf{C}_{0\leqslant t\leqslant r}$。证明，$\mathbf{C}_{0\leqslant t\leqslant r}$ 在这一点处的切空间由以下三个向量张成（球坐标系）

$$L=\frac{\partial}{\partial t}-\frac{\partial}{\partial r},\quad e_1=\frac{1}{r}\frac{\partial}{\partial\vartheta},\quad e_2=\frac{1}{r\sin\vartheta}\frac{\partial}{\partial\varphi}.$$

作为传统的记号，我们记 $|\not\nabla u|^2=|\nabla_{e_1}u|^2+|\nabla_{e_2}u|^2$。

D5)** 我们定义局部能量

$$E_{\mathrm{loc}}(t)=\frac{1}{2}\int_{\mathbf{B}_t}|\frac{\partial u}{\partial t}(t,x)|^2+|\nabla u(t,x)|^2dx.$$

证明，对任意的 $r\leqslant R$，我们都有

$$E_{\mathrm{loc}}(0)=E_{\mathrm{loc}}(r)+\int_{\mathbf{C}_{0\leqslant t\leqslant r}}|\nabla_L u|^2+|\not\nabla u|^2d\sigma.$$

（提示：对 $\Box u=0$ 两边同时乘以 $\dfrac{\partial u}{\partial t}$ 并在区域 $\mathbf{D}_{0\leqslant t\leqslant r}$ 上积分，你需要运用 Stokes 公式来得到边界上的积分）

D6)*（波方程的有限传播速度）我们现在只假设 $u(t,x)$ 是光滑的。证明，如果 u_0 和 u_1 有紧支集，那么对任意的时刻 t，$u(t,x)$ 对 x 有紧支集。

D7)* 假设 $u(t,x)$ 是光滑函数并且满足非线性波动方程

$$-\frac{\partial^2}{\partial t^2}u+\triangle u=u^3.$$

证明，如果 u_0 和 u_1 有紧支集，那么对任意的时刻 t，$u(t,x)$ 对 x 有紧支集。

In order to solve this differential equation you look at it until a solution occurs to you.

—- George Pólya

53.2 习题课：Riemann 积分的定义 3

清华大学 19-20 春季学期，数学分析二，第八次习题课内容

传统 Riemann 积分的定义：第三部分

我们沿用上次习题课的记号。

高维的 Lebesgue 定理

假设 $A \subset \mathbb{R}^n$ 是子集，对任意的函数 $f : A \to \mathbb{C}$ 和 $x \in A$，我们定义 f 在 x 处的**振幅**为

$$\omega(f, x) = \inf_{\delta > 0} \left(\sup_{y,z \in B_x(\delta) \cap A} |f(y) - f(z)| \right),$$

其中，$B_x(\delta)$ 为中心在 x 处半径为 δ 的闭球。

R51) 证明，f 在 x 处连续当且仅当 $\omega(f, x) = 0$。

（这个性质对于 A 是任意的距离空间都成立）

R52) 我们现在假设 $A = P$ 是矩形，$f \in \mathcal{R}(P)$ 是 Riemann 可积的函数。证明，对任意的 $\alpha > 0$，集合

$$A_\alpha = \{x \in P | \omega(f, x) \geqslant \alpha\}$$

是（Lebesgue 意义下的）零测集。

（请参考上学期第二十二次课的讲义中在 1 维时的证明）

R53) **（Lebesgue 定理）**假设 f 是矩形 P 上的有界函数。证明，f 是 Riemann-可积的当且仅当 f 的不连续点集是零测集（Lebesgue 意义下）。

（请参考上学期第二十二次课的讲义中在 1 维时的证明）

R54) 我们现在假设 $A = P$ 是矩形并且 Riemann-可积的函数 $f \in \mathcal{R}(P)$ 的取值是非负的。证明，$\int_P f = 0$ 当且仅当 $\{x \in P | f(x) \neq 0\}$ 是（Lebesgue 意义下的）零测集。

R55) **（Lebesgue 定理）**假设 f 是有界集合 A 上的有界函数，$\widetilde{f}$ 是 f 在 $\mathbb{R}^n$ 上用 0 所做的延拓（请参考上次习题课中的构造）。证明，f 在 A 上是 Riemann 可积的当且仅当 $\widetilde{f}$ 的不连续点集是（Lebesgue 意义下的）零测集。

R56) 假设 $A \subset \mathbb{R}^n$ 有界并且 ∂A 是（Lebesgue 意义下的）零测集。证明，A 上的连续函数在 A 上 Riemann 可积。

抽象可铺集上的积分

R57) 假设 $A \subset \mathbb{R}^n$ 是 $\mathcal{R}$-零测集，f 是 A 上的有界函数。证明，$f \in \mathcal{R}(A)$ 并且

$$\int_A f = 0.$$

R58) 假设 A 和 B 是 $\mathbb{R}^n$ 中两个有界集，$A \cap B$ 是 $\mathcal{R}$-零测集，f 是定义在 $A \cup B$ 上复值函数。证明，如果在 A，B 和 $A \cup B$ 这三个集合中的任意两个上面，f 是 Riemann-可积的，那么它在第三个集合上也是 Riemann-可积的并且下面的等式成立：

$$\int_{A\cup B} f = \int_A f + \int_B f.$$

R59) 假设 $A \subset \mathbb{R}^n$ 是有界集并且 $f \in \mathcal{R}(A)$。证明，如果 $B \subset A$ 是抽象可铺的，那么 $f\big|_B$ 是 B 上的 Riemann-可积函数。

R60) 假设 $A \subset \mathbb{R}^n$ 是抽象可铺集，$f \in \mathcal{R}(A)$，$\overline{f}$ 是 f 在 $\overline{A}$（A 的闭包）上的一个有界的延拓（即 $\overline{f}$ 是 $\overline{A}$ 上的有界函数并且 $\overline{f}\big|_A = f$）。证明，对任意的 B，满足 $\mathring{A} \subset B \subset \overline{A}$，$\overline{f}\big|_B$ 在 B 上是 Riemann 可积的并且

$$\int_B \overline{f} = \int_A f.$$

R61) 假设 $A \subset \mathbb{R}^n$ 是抽象可铺集。证明，每个满足 $\mathring{A} \subset B \subset \overline{A}$ 的集合 B 都是抽象可铺集并且 $m(B) = m(A)$。

R62)（Chasles 定理）给定有限个抽象可铺集 $A_1, \cdots, A_N \subset \mathbb{R}^n$，其中对任意的 $i \neq j$，$A_i \cap A_j$ 是 $\mathcal{R}$-零测集，给定函数

$$f : \bigcup_{i=1}^N A_i \to \mathbb{C}.$$

证明，$f \in \mathcal{R}\left(\bigcup_{i=1}^N A_i\right)$ 当且仅当对每个 i，我们都有 $f \in \mathcal{R}(A_i)$。如果上述成立，进一步证明

$$\int_{\cup_{i\leqslant N} A_i} f = \sum_{i=1}^N \int_{A_i} f.$$

R63) 假设 $A \subset \mathbb{R}^n$ 是抽象可铺集，A 上定义的有界函数 f 在 $\mathring{A}$ 上连续。证明，f 在 A 上是 Riemann-可积的。

R64) 假设 $A \subset \mathbb{R}^n$ 是紧的抽象可铺集，f 是 A 上的连续函数。证明，$f \in \mathcal{R}(A)$。

与 Lebesgue 积分的联系

R65) 假设 $A \subset \mathbb{R}^n$ 是抽象可铺集。证明，A 是 Lebesgue 可测集（按定义，一个集合 $A \subset \mathbb{R}^n$ 是 Lebesgue 可测的指的是存在 Borel 集 B_1 和 B_2，使得 $B_1 \subset A \subset B_2$ 并且 $m(B_1) = m(B_2)$。Borel 集未必是 Lebesgue 可测集。Lebesgue 可测集是 Borel 代数对于 Lebesgue 测度的完备化。）

R66) 假设 $A \subset \mathbb{R}^n$ 是抽象可铺集，$f \in \mathcal{R}(A)$。证明，f 在 A 上是 Lebesgue 可积的并且

$$\int_A f = \int_A f dm,$$

其中后者是 Lebesgue 积分。

———————————-

One should never try to prove anything that is not almost obvious.

————- Alexandre Grothendieck

———————————-

54 物理空间的光滑性与频率空间的衰减，Riemann-Lebesgue 引理，Dirichlet 核，Féjer 核，Féjer 核的收敛理论，局部化引理

二零二零年五月十八日，星期一，晴

我们现在学习一个从频率的观点看函数的重要的例子：

定理 359. *给定 $\mathbf{T}$ 上定义可积函数 f。如下两个叙述是等价的：*

1) f 是光滑函数，即 $f \in C^\infty(\mathbf{T})$；

2) $\widehat{f}(k)$ 的具有任意可能的多项式衰减，即对任意的 $N \geqslant 0$，总存在常数 C_N，使得对任意的 $k \in \mathbb{Z}$，我们有

$$\left|\widehat{f}(k)\right| \leqslant \frac{C_N}{(1+|k|)^N}.$$

注记. *函数 Fourier 系数的衰减速度决定了函数的光滑性，反之亦然。换句话说，频率空间的衰减等价于物理空间的光滑性。这是学习 Fourier 分析必须要记住的原则之一。*

证明: 我们先证明 1) ⇒ 2)。不妨假设 $k \neq 0$，证明的关键就是利用光滑性，我们可以连续做多次分部积分并且分部积分得到的在端点 0 和 2π 处的值会相互抵消（利用周期性）：

$$\begin{aligned}
\widehat{f}(k) &= \frac{1}{2\pi}\int_0^{2\pi} f(x)e^{-ikx}dx \\
&= \frac{1}{2\pi}\frac{1}{(-ik)}\int_0^{2\pi} f(x)\left(e^{-ikx}\right)' dx \\
&= \frac{1}{2\pi}\frac{1}{ik}\int_0^{2\pi} f(x)'e^{-ikx}dx \\
&\quad \cdots\ \cdots \\
&= \frac{1}{2\pi}\frac{1}{(ik)^N}\int_0^{2\pi} f^{(N)}(x)e^{-ikx}dx.
\end{aligned}$$

所以，

$$|\widehat{f}(k)| \leqslant \frac{1}{2\pi}\frac{1}{|k|^N}\int_0^{2\pi}|f^{(N)}(x)|dx \leqslant \frac{\|f^{(N)}\|_{L^\infty}}{|k|^N}.$$

证明 2) ⇒ 1) 的方法是使用 Lebesgue 控制收敛定理或者上学期证明的关于求导和求和可交换的命题。首先，根据条件 2)（选 $N=2$ 即可），我们知道 $\sum\limits_{k\in\mathbb{Z}}|\widehat{f}(k)|$ 是绝对收敛的，所以，根据刚刚证明的结论：

$$f(x) = \sum_{k\in\mathbb{Z}}\widehat{f}(k)e^{ikx}$$

是连续函数。另外，利用 $N=3$，我们知道函数项级数

$$\sum_{k\in\mathbb{Z}}\left(\widehat{f}(k)e^{ikx}\right)' = \sum_{k\in\mathbb{Z}} ik\widehat{f}(k)e^{ikx}$$

是绝对收敛的，因为

$$\sum_{k\in\mathbb{Z}}|ik\widehat{f}(k)|\leqslant\sum_{k\in\mathbb{Z}}\frac{C_3|k|}{(1+|k|)^3}<\infty.$$

根据 Lebesgue 控制收敛的第二个推论，我们有

$$f(x)'=\left(\sum_{k\in\mathbb{Z}}\widehat{f}(k)e^{ikx}\right)'=\underbrace{\sum_{k\in\mathbb{Z}}ik\widehat{f}(k)e^{ikx}}_{\text{绝对收敛，所以是连续函数}}.$$

所以，$f\in C^1(\mathbf{T})$。我们可以继续这个过程，从而证明对任意的 $\ell\in\mathbb{Z}_{\geqslant 1}$，$f\in C^\ell(\mathbf{T})$。 □

注记. 从证明中可见，对函数光滑 f 求导数等价于对每个频率 k 的 *Fourier* 系数乘以 $-ik$。把这个现象和线性代数联系起来是很有启发性的：我们考虑线性空间 $L^2(\mathbf{T})$，那么，

$$\frac{d}{dx}:L^2(\mathbf{T})\to L^2(\mathbf{T})$$

是线性算子（当然，你可以会关心 $\dfrac{d}{dx}$ 是否是良好定义的，让我们假设这个成立（其实不成立））。我们现在考虑测度空间 $(\mathbb{Z},\mathcal{P}(\mathbb{Z}),\mu)$，其中，$\mu$ 就是数元素个数的测度，我们令 $\ell^2(\mathbb{Z})=L^2(\mathbb{Z},\mathcal{P}(\mathbb{Z}),\mu)$。那么，根据关于 L^2 空间上的 *Fourier* 级数的基本定理，我们得到了（连续的）映射（同构）：

$$\mathcal{F}:L^2(\mathbf{T})\to\ell^2(\mathbb{Z}),\quad f\mapsto\left(\widehat{f}(k)\right)_{k\in\mathbb{Z}}.$$

我们把这个同构看做是在 $L^2(\mathbf{T})$ 选取了一个基。那么，上面求导可以用下面的交换图标来表达：

$$\begin{array}{ccc}L^2(\mathbf{T}) & \xrightarrow{\frac{d}{dx}} & L^2(\mathbf{T})\\ \downarrow{\scriptstyle\mathcal{F}} & & \downarrow{\scriptstyle\mathcal{F}}\\ \ell^2(\mathbb{Z}) & \xrightarrow{\times(-ik)} & \ell^2(\mathbb{Z})\end{array},$$

换句话说，换了这组基后，微分算子 $\dfrac{d}{dx}$ 被对角化了（变成了成在每个分量上乘一个数），这把求微分的操作变成了代数操作。

仿照线性代数中对一个正定的矩阵可以定义它的开方的思路，我们对一个光滑函数的 *Fourier* 系数分别乘以 $|k|^{\frac{1}{2}}$，这可以定义求二分之一次导数。这实际上是所谓的拟微分算子的基本想法（如果下学期时间足够的话，我们会讲到）。

我们上学期已经证明如下的 Riemann-Lebesgue 引理，这个引理是 Fourier 分析中的核心定理：

定理 360 (Riemann-Lebesgue)**.** 对任意的 $f\in L^1(\mathbf{T})$，我们都有

$$\lim_{|k|\to\infty}\widehat{f}(k)=0.$$

证明: 上面的定理表明，如果 $f\in C^\infty(\mathbf{T})$，那么命题成立，再利用光滑函数在 $L^1(\mathbf{T})$ 中稠密即可。我们不再重复证明的细节。 □

注记. 实际上，我们可以证明，对任意的序列 $\{a_k\}_{k\in\mathbb{Z}}$，其中 $a_k \neq 0$，其中

$$\lim_{k\to\infty} a_k = \infty,$$

总存在某个可积函数 $f \in L^1(\mathbf{T})$，使得

$$\lim_{k\to\infty} a_k|\widehat{f}(k)| = \infty.$$

换句话说，一个 L^1 函数的 *Fourier* 系数的衰减可以任意得慢。这个证明需要用到泛函分析中的 *Banach-Steinhaus* 定理，我们这里仅仅陈述结论而不给证明。(在本次的作业中，我们有一个较弱的版本，证明上面的极限对一个子序列是成立的)

注记. 证明中分部积分的技巧实际上还证明了如下的结论：如果 $f \in C^m(\mathbf{T})$，其中 $m \geqslant 1$ 是正整数，那么存在常数 C，使得对任意的 $k \in \mathbb{Z}$，我们有

$$|\widehat{f}(k)| \leqslant \frac{C_N}{(1+|k|)^m}.$$

也就是说，每多一阶可微性，*Fourier* 系数的衰减也快一阶。

Dirichlet 核和 Féjer 核

我们上周就提过：可以对 $\mathbf{T}$ 上的两个 L^1 函数定义卷积：

$$f * g : \mathbf{T} \to \mathbb{C}, \quad x \mapsto (f*g)(x) = \int_0^{2\pi} f(x-y)g(y)\frac{dy}{2\pi},$$

与 $\mathbb{R}^n$ 上的证明完全一致，利用 Fubini 定理，我们有

$$\|f * g\|_{L^1} \leqslant \|f\|_{L^1}\|g\|_{L^1}.$$

这表明，我们有如下的乘法结构

$$L^1(\mathbf{T}) \times L^1(\mathbf{T}) \xrightarrow{*} L^1(\mathbf{T}).$$

仍然根据 Fubini 定理，我们可以说明卷积满足交换律和结合律：

1) 对于几乎处处的 $x \in \mathbf{T}$，我们有 $(f*g)(x) = (g*f)(x)$。

2) 假设 $f, g, h \in L^1(\mathbb{R}^n)$，那么 $\big((f*g)*h\big)(x) = \big(f*(g*h)\big)(x)$ 几乎处处成立。

类似地，我们可以证明：对于 $p = 1, 2$ 或者 ∞，有

$$L^1(\mathbf{T}) \times L^p(\mathbf{T}) \xrightarrow{*} L^p(\mathbf{T}),$$

即对任意的 $f \in L^1(\mathbf{T})$，$g \in L^p(\mathbf{T})$，函数

$$(f*g)(x) = \int_{\mathbf{T}} f(x-y)g(y)dy$$

在 $L^p(\mathbb{R}^n)$ 中是良好定义的（几乎处处）。特别地，我们还有如下的不等式：

$$\|f * g\|_{L^p(\mathbf{T})} \leqslant \|f\|_{L^1(\mathbf{T})}\|g\|_{L^p(\mathbf{T})}.$$

卷积的定义不仅仅是概念上的推广，它有着很重要的实际意义。我们来看下面的例子：

例子. *假设 $f \in L^1(\mathbf{T})$，考虑 Fourier 分析中最基本的函数*

$$e_k(x) = e^{ikx}.$$

我们来计算 $f * e_k$：

$$\begin{aligned}\big(f * e_k\big)(x) &= \int_0^{2\pi} f(y)e^{ik(x-y)}\frac{dy}{2\pi} \\ &= e^{ikx}\int_0^{2\pi} f(y)e^{-iky}\frac{dy}{2\pi} \\ &= \widehat{f}(k)e^{ikx}.\end{aligned}$$

所以，与 e^{ikx} 做卷积就给出了 f 向 e^{ikx} 这个方向上的投影。

特别地，我们在 Fourier 技术中感兴趣的部分和就有如下的表达式：

$$S_N(f)(x) = \sum_{-N\leqslant k\leqslant N}\widehat{f}(k)e^{ikx} = f * \left(\sum_{-N\leqslant k\leqslant N} e_k\right).$$

我们定义 **Dirichlet 核** $D_N(x)$ 如下，其中，$N \geqslant 0$ 是整数：

$$\begin{aligned}D_N(x) &= \sum_{-N\leqslant k\leqslant N} e_k = \frac{e^{i(N+1)x} - e^{-iNx}}{e^{ix} - 1} \\ &= \frac{e^{i(N+\frac{1}{2})x} - e^{-i(N+\frac{1}{2})x}}{e^{i\frac{x}{2}} - e^{-i\frac{x}{2}}} = \frac{\sin\big((N+\frac{1}{2})x\big)}{\sin\big(\frac{1}{2}x\big)}.\end{aligned}$$

所以，对于 $x \in \mathbb{R}$，我们定义 **Dirichlet 核** $D_N(x)$：

$$D_N(x) = \begin{cases}\dfrac{\sin\big((N+\frac{1}{2})x\big)}{\sin\big(\frac{1}{2}x\big)}, & x \neq 2k\pi; \\ 2N+1, & x = 2k\pi.\end{cases}$$

这也是定义在 $\mathbf{T}$ 上的函数。

我们就证明了下面的引理：

引理 361. *对任意的 $f \in L^1(\mathbf{T})$，它的 Fourier 级数的部分和 $S_N(f)$ 可以表达为*

$$S_N(f)(x) = \sum_{-N\leqslant k\leqslant N}\widehat{f}(k)e^{ikx} = \big(f * D_N\big)(x).$$

在传统的 Fourier 分析中，我们还考虑一种被称作 Fejér 和的部分和。按照定义，它是部分和序列 $\{S_N\}_{N\geqslant 0}$ 的 Cesàro 求和（前 n 项的平均值），即

$$\sigma_N(f)(x)=\frac{S_0(x)+S_1(x)+\cdots+S_{N-1}(x)}{N}$$

我们将会看到，Cesàro 求和会让 Fourier 级数的收敛速度变得很快!

我们现在来计算 Féjer 核的公式。按照定义，我们有

$$\sigma_N(f)(x)=\big(f*\frac{1}{N}\sum_{k=0}^{N-1}D_k\big)(x)=\big(f*F_N\big)(x).$$

其中，Féjer 核 F_N 被定义成

$$F_N(x)=\sum_{k=0}^{N-1}D_k.$$

我们可以进一步计算 $F_N(x)$ 的解析表达式：

$$\begin{aligned}
F_N(x)&=\frac{1}{N}\sum_{k=0}^{N-1}\frac{\sin\big((k+\frac{1}{2})x\big)}{\sin\big(\frac{1}{2}x\big)}=\frac{1}{N}\frac{1}{\sin\big(\frac{1}{2}x\big)}\sum_{k=0}^{N-1}\mathrm{Im}\left(e^{\frac{i}{2}x}e^{ikx}\right)\\
&=\frac{1}{N}\frac{1}{\sin\big(\frac{1}{2}x\big)}\mathrm{Im}\left(e^{\frac{i}{2}x}\frac{e^{iNx}-1}{e^{ix}-1}\right)\\
&=\frac{1}{N}\frac{1}{\sin\big(\frac{x}{2}\big)}\mathrm{Im}\left(e^{i\frac{Nx}{2}}\frac{e^{i\frac{Nx}{2}}-e^{-i\frac{Nx}{2}}}{e^{i\frac{x}{2}}-e^{-i\frac{x}{2}}}\right)\\
&=\frac{1}{N}\left(\frac{\sin(\frac{N}{2}x)}{\sin(\frac{x}{2})}\right)^2
\end{aligned}$$

所以，对于 $x\in\mathbb{R}$，我们定义如下的 **Féjer 核**：

$$F_N(x)=\begin{cases}\frac{1}{N}\left(\frac{\sin(\frac{N}{2}x)}{\sin(\frac{x}{2})}\right)^2, & x\neq 2k\pi;\\ N, & x=2k\pi.\end{cases}$$

练习. 证明，$F_N(x)$ 的积分是 2π：

$$\int_0^{2\pi}F_N(x)\frac{dx}{2\pi}=1.$$

这是一个有趣的联系，最好的证明是观察到 $\displaystyle\int_{-\pi}^{\pi}D_N(x)\frac{dx}{2\pi}=1$，这是因为（不要用解析表达式）$\displaystyle D_N=\sum_{-N\leqslant k\leqslant N}e_k$。

我们之前卷积的核函数是 χ_ε，与之类似，我们的 Féjer 核 F_N 满足下面的性质：

1) F_N 是正的，即对任意的 $x\in\mathbb{R}$，我们有 $F_N(x)\geqslant 0$；

2) F_N 是 $\mathbf{T}$ 上的概率密度，即

$$\int_{-\pi}^{\pi}F_N(x)\frac{dx}{2\pi}=1;$$

3) 把 F_N 视作是 $[-\pi,\pi]$ 上的函数，那么，当 $N\to\infty$ 时，F_N 集中在 $x=0$ 附近，即对任意的 $\delta>0$，当 $N\to\infty$ 时，我们有

$$\int_{\delta\leqslant|x|\leqslant\pi} F_N(x)\frac{dx}{2\pi}\longrightarrow 0,\quad N\to\infty.$$

证明是直截了当的：当 $|x|\in[\delta,\pi]$ 时，由于 $\left|\sin\left(\frac{x}{2}\right)\right|\leqslant\frac{|x|}{2}$，我们有

$$|F_N(x)|\leqslant\frac{1}{N}\Big(\frac{2}{\delta}\Big)^2\Big(\sin(\frac{N}{2}x)\Big)^2\leqslant\frac{1}{N\delta^2}.$$

所以，当 $N\to 0$ 时，我们有

$$\int_{\delta\leqslant|x|\leqslant\pi} F_N(x)\frac{dx}{2\pi}\leqslant\int_{\delta\leqslant|x|\leqslant\pi}\frac{4}{N\delta^2}\frac{dx}{2\pi}\to 0.$$

传统的 Fourier 分析把 F_N 的性质提炼成所谓的好的积分核的性质：

定义 362. 给定 $\mathbb{R}$ 上的一族连续函数 $\{K_N(x)\}_{N\geqslant 0}$，我们可以认为 $x\in\mathbf{T}$。假设它们满足下面的性质：

1) K_N 在 $L^1(\mathbf{T})$ 中有界：存在常数 $M>0$，使得对任意的 $N\geqslant 0$，都有

$$\int_{-\pi}^{\pi}|K_N|\frac{dx}{2\pi}\leqslant M;$$

2) K_N 是 $\mathbf{T}$ 上的概率密度，即 $\int_{-\pi}^{\pi}K_N(x)\frac{dx}{2\pi}=1$；

3) 当 $N\to\infty$ 时，$\{K_N\}_{\geqslant 0}$ 集中在 $x=0$ 附近，即对任意的 $\delta>0$，当 $N\to\infty$ 时，我们有

$$\int_{\delta\leqslant|x|\leqslant\pi}|K_N(x)|\frac{dx}{2\pi}\to 0.$$

我们就称 $\{K_N(x)\}$ 是 $\mathbf{T}$ 上一族**好的积分核**。

按照刚才的构造，Féjer 核 $\{F_N\}_{N\geqslant 0}$ 是一族好的积分核。

定理 363. 假设 $\{K_N(x)\}_{N\geqslant 1}$ 是一族好的积分核，那么对任意的 $f\in C^0(\mathbf{T})$，我们都有一致收敛：

$$\lim_{N\to\infty}\|f*K_N-f\|_{L^\infty}=0,$$

其中

$$f*K_N(x)=\int_{-\pi}^{\pi}f(x-y)K_N(y)\frac{dy}{2\pi}.$$

证明: 利用 K_N 是概率密度函数，我们有

$$\begin{aligned}|\big(f*K_N\big)(x)-f(x)| &= \Big|\int_{-\pi}^{\pi}\big(f(x-y)-f(x)\big)K_N(y)\frac{dy}{2\pi}\Big| \\ &\leqslant \underbrace{\Big|\int_{|y|\leqslant\delta}\big(f(x-y)-f(x)\big)K_N(y)\frac{dy}{2\pi}\Big|}_{A} \\ &\quad + \underbrace{\Big|\int_{\delta\leqslant|y|\leqslant\pi}\big(f(x-y)-f(x)\big)K_N(y)\frac{dy}{2\pi}\Big|}_{B}.\end{aligned}$$

由于 f 在 $[0,2\pi]$ 上面连续，所以一致连续性。从而，对任意的 $\varepsilon>0$，存在 $\delta>0$，当 $|y|<\delta$ 时，对任意的 x，我们都有 $|f(x-y)-f(x)|<\varepsilon$。所以，根据好的积分核的性质 1)，我们有

$$A\leqslant\varepsilon\int_{|y|\leqslant\delta}|K_N(y)|\frac{dy}{2\pi}\leqslant\varepsilon M.$$

对于 B，我们用好的积分核的性质 3)：

$$B\leqslant 2\|f\|_{L^\infty}\int_{\delta\leqslant|y|\leqslant\pi}|K_N(y)|\frac{dy}{2\pi}.$$

所以，存在 N_0 依赖于 δ 从而依赖于 ε，使得当 $N\geqslant N_0$ 时，对任意的 $x\in\mathbf{T}$，我们有

$$B\leqslant\varepsilon.$$

综合上述，对任意的 $\varepsilon>0$，存在 N_0，使得当 $N\geqslant N_0$ 时，对任意的 $x\in\mathbf{T}$，我们有

$$|\big(f*K_N\big)(x)-f(x)|\leqslant(M+1)\varepsilon.$$

命题得到证明。 □

推论 364. *对任意的 $f\in C^0(\mathbf{T})$，我们都有一致收敛*

$$\lim_{N\to0}\|f*F_N-f\|_{L^\infty}=0.$$

特别地，我们注意到

$$F_N(x)=\frac{1}{N}\sum_{|k|\leqslant N}(N-|k|)e^{ikx},$$

所以，

$$\sigma_N(f)(x)=F_N*f(x)=\sum_{|k|\leq N}(1-\frac{|k|}{N})\widehat{f}(k)e^{ikx}.$$

这是个有限的三角级数，它将一致收敛到 f，所以，我们实际上再次证明了周期连续函数情形下的 Stone-Weierstrass 定理。

Féjer 核的一个应用：L^1 函数的原函数与 Fatou 的反例

根据 Riemann-Lebesgue 引理，$f \in L^1(\mathbf{T})$，那么 $\widehat{f}(k) = 0$。那么，是不是每个极限是 0 序列 $\{a_k\}_{k\in\mathbb{Z}}$，它必为某个可积函数的 Fourier 系数呢，即是否存在 $f \in L^1(\mathbf{T})$，使得对任意的 k，我们都有 $\widehat{f} = a_k$?

给定 $f \in L^1(\mathbf{T})$ 是周期函数，我们假设 $\widehat{f}(0) = 0$。（$\widehat{f}(0) = 0$ 等价于 f 的积分为零）我们定义 $\mathbb{R}$ 上的函数：

$$F : \mathbb{R} \to \mathbf{C}, \quad F(x) = \int_0^x f(t)dt.$$

根据 f 是可积的，$F(x)$ 是连续函数（比如，可以用 Lebesgue 控制收敛来证明）。由于 f 在一个周期上的积分消失，所以 $F(x)$ 也是以 2π 为周期的。我们可以计算 $F(x)$ 的 Fourier 系数：

引理 365. *给定 $f \in L^1(\mathbf{T})$，我们假设 $\widehat{f}(0) = 0$。那么，以 2π 为周期的连续函数 $F(x) = \int_0^x f(t)dt$ 的 Fourier 系数可以用 f 的信息表示如下：*

$$\widehat{F}(k) = \begin{cases} -\int_0^{2\pi} xf(x)\frac{dx}{2\pi}, & k = 0; \\ \frac{1}{ik}\widehat{f}(k), & k \neq 0. \end{cases}$$

注记. *如果 f 是连续函数，那么 F 是 C^1 的，所以，$F' = f$。此时，由于求导在频率空间来看是乘以 $-ik$，所以当 $k \neq 0$ 时，命题时显然的（可以直接分部积分）。*

证明: 当 $k = 0$ 时，利用 Fubini 定理，我们有

$$\begin{aligned} \widehat{F}(0) &= \int_0^{2\pi} \left(\int_0^x f(t)dt \right) \frac{dx}{2\pi} \\ &= \int_{[0,2\pi]^2} f(t)\mathbf{1}_{t\leqslant x}(t,x)\frac{dxdt}{2\pi} \\ &= \frac{1}{2\pi}\int_0^{2\pi} \left(\int_t^{2\pi} f(t)dx \right) dt \\ &= \frac{1}{2\pi}\int_0^{2\pi} (2\pi - t)f(t)dt. \end{aligned}$$

由于 $\widehat{f} = 0$，所以

$$\widehat{F}(0) = -\int_0^{2\pi} xf(x)\frac{dx}{2\pi}.$$

当 $k \neq 0$ 时，我们有

$$\begin{aligned} \widehat{F}(k) &= \int_0^{2\pi} \left(\int_0^x f(t)dt \right) e^{-ikx}\frac{dx}{2\pi} \\ &= \frac{1}{2\pi}\int_0^{2\pi} \left(\int_t^{2\pi} e^{-ikx}f(t)dx \right) dt \\ &= \frac{1}{2\pi}\int_0^{2\pi} \frac{1}{-ik}\left(1 - e^{-ikt}\right) f(t)dt \\ &= \frac{1}{ik}\widehat{f}(k). \end{aligned}$$

命题成立。 □

在同样的假设下，由于 F 是连续的，根据 Féjer 核的理论，当 $N\to\infty$ 时，$F_N * F$ 一致收敛到 F。另外，由于

$$F_N * F(x) = \sum_{|k|\leqslant N}(1-\frac{|k|}{N})\widehat{F}(k)e^{ikx},$$

所以，在 $x=0$ 处，我们有

$$\begin{aligned} F_N * F(0) &= \sum_{|k|\leqslant N}(1-\frac{|k|}{N})\widehat{F}(k) \\ &= -\frac{1}{2\pi}\int_0^{2\pi} xf(x)dx + \sum_{1\leqslant|k|\leqslant N}(1-\frac{|k|}{N})\frac{1}{ik}\widehat{f}(k) \\ &= -\frac{1}{2\pi}\int_0^{2\pi} xf(x)dx - i\sum_{1\leqslant|k|\leqslant N}\frac{\widehat{f}(k)}{k} + \underbrace{\frac{i}{N}\sum_{1\leqslant k\leqslant N}\widehat{f}(k)}_{A_N} - \underbrace{\frac{i}{N}\sum_{-N\leqslant k\leqslant -1}\widehat{f}(k)}_{B_N}. \end{aligned}$$

根据 Riemann-Lebesgue 引理，我们知道 $\lim\limits_{k\to\infty}\widehat{f}(k)=0$，所以，$\lim\limits_{N\to\infty}A_N=0$（这是第一学期关于数列极限的标准习题：如果一个数列的极限是 0，那么它所对应的 Cesàro 和（前 n 项的平均）的极限也是 0）；类似地，$\lim\limits_{N\to\infty}B_N=0$。由于 $F(0)=0$，所以，当 $N\to\infty$ 时，上面的等式给出了

$$\lim_{N\to\infty}\left(\sum_{1\leqslant|k|\leqslant N}\frac{\widehat{f}(k)}{k}\right) = \frac{i}{2\pi}\int_0^{2\pi} xf(x)dx.$$

特别地，这表明对于一个 L^1 函数的 Fourier 系数，数列 $\left\{\frac{\widehat{f}(k)}{k}\right\}_{k\in\mathbb{Z}}$ 具有一定的“可求和性”。Fatou 根据这个性质构造了如下的反例：

命题 366 (Fatou)**.** *对于如下定义的数列 $\{a_k\}_{k\geqslant 1}$，其中*

$$a_k = \begin{cases} 0, & |k|\leqslant 1; \\ \frac{1}{2i\log k}, & k\geqslant 2; \\ \frac{-1}{2i\log|k|}, & k\leqslant -2. \end{cases}$$

不存在 $f\in L^1(\mathbf{T})$，使得对任意的 $k\in\mathbb{Z}$，$\widehat{f}(k)=a_k$。

证明: 我们用反证法：如果不然，那么，f 的 Fourier 级数的部分和为

$$S_N(f)(x) = \sum_{|k|\leqslant N}a_k e^{ikx}\left(=\sum_{k\geqslant 2}\frac{\sin(kx)}{\log k}\right).$$

上面的计算表明，$\left\{\sum_{1\leqslant|k|\leqslant N}\frac{a_k}{k}\right\}_{N\geqslant 1}$ 是有极限的，然而，根据 a_k 的定义，我们有

$$\sum_{1\leqslant|k|\leqslant N}\frac{a_k}{k}=\frac{1}{i}\sum_{2\leqslant k\leqslant N}\frac{1}{k\log k}.$$

我们可以用上学期最后学习的面积方法来用 $(\log(x))^{-1}$ 的积分估计上面求和的大小，这表明

$$\sum_{2\leqslant k\leqslant N}\frac{1}{k\log k}\sim\log\log N\to\infty.$$

然而，上面的积分应该收敛到 $\frac{i}{2\pi}\int_0^{2\pi}xf(x)dx$，矛盾。 □

尽管 Féjer 核的理论更简洁，我们却更想知道是否 $S_N(f)$ 能够足够好地逼近 f，因为这是最自然的部分和。这使得我们要研究 Dirichlet 核函数 D_N 的性质。实际上，尽管 $D_N(x)$ 的积分是 1，但是，当 N 很大的时候，它的振荡很厉害，所有有很多正的部分对积分的贡献可以被负的部分对积分的贡献消掉，而 $D_N(x)$ 并不满足好的积分核的定义中的第一条。实际上，我们有

引理 367. *当 $N\to\infty$ 时，我们有*

$$\int_{-\pi}^{\pi}|D_N(x)|\frac{dx}{2\pi}=\frac{4}{\pi^2}\log(N)+O(1).$$

证明: 我们重新写下 $D_N(x)$ 的表达式

$$D_N(x)=\begin{cases}\frac{\sin\left((N+\frac{1}{2})x\right)}{\sin\left(\frac{1}{2}x\right)}, & x\neq 2k\pi;\\ 2N+1, & x=2k\pi.\end{cases}$$

证明的基本想法是把分母中的 sin 函数替换成更容易控制的函数。我们首先说明：对任意的 $x\in[0,\infty)$，我们有

$$x-\frac{1}{6}x^3\leqslant\sin(x)\leqslant x.$$

后面一个不等式是熟知的；为了说明前一个，我们知道

$$1-\frac{1}{2}x^2\leqslant\cos(x),\quad x\geqslant 0.$$

对 x 积分即可。

所以，对 $\pi\geqslant x\geqslant 0$，我们有

$$\frac{2}{x}\leqslant\frac{1}{\sin\left(\frac{x}{2}\right)}\leqslant\frac{2}{x}\frac{1}{1-\frac{1}{24}x^2}\leqslant\frac{2}{x}\big(1+2\times\frac{1}{24}x^2\big)\leqslant\frac{2}{x}+\frac{x}{6}.$$

即

$$\frac{2}{x}\leqslant\frac{1}{\sin\left(\frac{x}{2}\right)}\leqslant\frac{2}{x}+\frac{x}{6}.$$

从而，通过把 $[-\pi,\pi]$ 上的积分变成 $[0,\pi]$ 上的积分，我们有

$$\begin{aligned}0 &\leqslant 2\int_0^{\pi}|D_N(x)|\frac{dx}{2\pi} - \overbrace{2\int_0^{\pi}\frac{\left|\sin\left((N+\frac{1}{2})x\right)\right|}{\frac{x}{2}}\frac{dx}{2\pi}}^{I_N}\\ &\leqslant \frac{1}{6\pi}\int_0^{\pi}x\left|\sin\left((N+\frac{1}{2})x\right)\right|dx\\ &\leqslant \frac{1}{6\pi}\int_0^{\pi}xdx = \frac{\pi}{12} = O(1).\end{aligned}$$

所以，我们只需要对

$$I_N = \frac{2}{\pi}\int_0^{\pi}\frac{\left|\sin\left((N+\frac{1}{2})x\right)\right|}{x}dx$$

的增长进行估计即可。此时，我们有

$$\begin{aligned}I_N &= \frac{2}{\pi}\int_0^{(N+\frac{1}{2})\pi}\frac{|\sin x|}{x}dx\\ &= \frac{2}{\pi}\int_0^{N\pi}\frac{|\sin x|}{x}dx + \underbrace{\frac{2}{\pi}\int_{N\pi}^{(N+\frac{1}{2})\pi}\frac{|\sin x|}{x}dx}_{O(N^{-1})}\\ &= \frac{2}{\pi}\sum_{k=0}^{N-1}\int_{k\pi}^{(k+1)\pi}\frac{|\sin x|}{x}dx + O(1)\\ &= \frac{2}{\pi}\sum_{k=1}^{N-1}\int_0^{\pi}\frac{|\sin x|}{x+k\pi}dx + O(1).\end{aligned}$$

我们进一步放缩分母，我们有

$$\frac{1}{(k+1)\pi} \leqslant \frac{1}{x+k\pi} \leqslant \frac{1}{k\pi}.$$

从而

$$\frac{2}{\pi}\sum_{k=1}^{N-1}\int_0^{\pi}\frac{|\sin x|}{(k+1)\pi}dx \leqslant I_N - O(1) \leqslant \frac{2}{\pi}\sum_{k=1}^{N-1}\int_0^{\pi}\frac{|\sin x|}{k\pi}dx.$$

然而，左右两边我们可以直接计算，从而

$$\frac{4}{\pi^2}\sum_{k=2}^{N}\frac{1}{k}dx \leqslant I_N - O(1) \leqslant \frac{4}{\pi^2}\sum_{k=1}^{N-1}\frac{1}{k}.$$

最终利用

$$\sum_{1}^{N-1}\frac{1}{k} = \log N + \gamma + O(\frac{1}{N}),$$

其中 γ 是 Euler 常数。命题得证。 □

Dirichlet 核函数的研究将是我们之后学习的重点。我们先证明 Fourier 级数的部分和或者说 Dirichlet 核的局部化的性质。这是一个非常有启发意义的命题，它的证明是 Riemann-Lebesgue 的一个漂亮的应用：

引理 368 (局部化引理)**.** *对任意的 $x_0 \in \mathbb{R}$ 任意的小正数 $\delta > 0$，如果函数 $f, g \in L^1(\mathbf{T})$ 满足*

$$f\big|_{(x_0-\delta, x_0+\delta)} = g\big|_{(x_0-\delta, x_0+\delta)},$$

那么，我们有

$$\lim_{N\to\infty} \big(S_N(f)(x_0) - S_N(g)(x_0)\big) = 0.$$

注记. *据此，为了研究 f 在某点 x_0 处的 Fourier 级数是否收敛，只要将注意力集中在这个点的邻域就好！这是绝对是不平凡的结论：按照定义，Fourier 系数是依赖于 f 在整个 $\mathbf{T}$ 的积分的，然而 Fourier 级数的收敛性却是局部的！*

证明： 首先，由于 $D_N(x)$ 是偶函数，通过变量替换，我们有

$$\begin{aligned}
&S_N(f)(x_0) - S_N(g)(x_0)\\
&= \int_{-\pi}^{\pi} D_N(y)\big(f(x_0-y) - g(x_0-y)\big)\frac{dy}{2\pi}\\
&= \int_{0}^{\pi} D_N(y)\big(f(x_0-y) + f(x_0+y) - g(x_0-y) - g(x_0+y)\big)\frac{dy}{2\pi}\\
&= \int_{\delta}^{\pi} D_N(y)\big(f(x_0-y) + f(x_0+y) - g(x_0-y) - g(x_0+y)\big)\frac{dy}{2\pi}.
\end{aligned}$$

根据 $D_N(x)$ 的解析表达式，我们有

$$\begin{aligned}
S_N(f)(x_0) - S_N(g)(x_0) &= \int_{\delta}^{\pi} \sin\big((N+\frac{1}{2})y\big)\underbrace{\frac{\big(f(x_0-y)+f(x_0+y)-g(x_0-y)-g(x_0+y)\big)}{\sin\big(\frac{x}{2}\big)}}_{F(y)}\frac{dy}{2\pi}\\
&= \int_{-\pi}^{\pi} \sin\big((N+\frac{1}{2})y\big)\cdot F(y)\mathbf{1}_{\delta\leqslant|y|\leqslant\pi}(y)dy.
\end{aligned}$$

由于在 $x \geqslant \delta$ 时，$\sin(\frac{x}{2}) \geqslant \sin(\frac{\delta}{2})$ 有正的下界，所以 $F(x)\mathbf{1}_{\delta\leqslant|x|\leqslant\pi}(x)$ 实际上在 $L^1(\mathbf{T})$ 中。根据 Riemann-Lebesgue 引理，当 $N \to \infty$ 时，我们有

$$|S_N(f)(x_0) - S_N(g)(x_0)| \to 0.$$

命题得证。 □

注记. *对上面的证明稍加修改，我们就可以证明：对任意的 $h \in L^1(\mathbf{T})$ 和任意的小正数 $\delta > 0$，当 $N \to \infty$ 时，我们有*

$$\begin{aligned}
&\int_{\delta\leqslant|x|\leqslant\pi} D_N(x)h(x)\frac{dx}{2\pi} \to 0,\\
&\int_{\delta\leqslant|x|\leqslant\pi} F_N(x)h(x)\frac{dx}{2\pi} = O(\frac{1}{N}).
\end{aligned}$$

我们观察到，尽管上面的的极限都消失，但是后者有固定的衰减速度而前者并没有（由 *Riemann-Lebesgue* 引理给出的衰减是没有衰减速率的）。

55 du Bois-Reymond 反例，Fourier 级数收敛的 Dirichlet 定理，Fourier 级数收敛的 Dini 定理，Hölder 连续函数 Fourier 级数收敛理论，有界变差函数的 Jordan 定理 Fourier 级数收敛理论

二零二零年五月二十一日，星期四，晴

经典的 Fourier 级数收敛理论

首先，我们要明确，Fourier 级数在一点处是否收敛并不是明显的，比如，对于连续函数，我们有 du Bois-Reymond 的著名反例（1876）。这个构造背后的想法是基于下面不等式（$x \neq 0$ 时，利用级数的 Abel 判别法／求和法；$x = 0$ 时，这是显然的，然而这个情况很重要）：存在正常数 C，使得

$$|\sum_{1\leqslant|k|\leqslant N}\frac{e^{ikx}}{k}|\leqslant C.$$

另外，我们现在不再把每个频率 e^{ikx} 看成基本的单元，而是把若干个频率的叠加（在有控制的情形下）看成基本的单元（物理上讲这样的叠加称作是**波包**）。

对于 $K\in\mathbb{Z}_{\geqslant 1}$，我们定义波包函数

$$W_K(x)=e^{i(2K)\cdot x}\sum_{1\leqslant|k|\leqslant K}\frac{e^{ikx}}{k},$$
$$W_K^-(x)=e^{i(2K)\cdot x}\sum_{-K\leqslant k\leqslant -1}\frac{e^{ikx}}{k}.$$

那么，

- $W_K(x)$ 在频率 $k=K,K+1,\cdots,3K$ 之外都是零。

  也就是说，如果 $k\notin[K,3K]$，那么 $\widehat{W_K}(k)=0$。

- $W_K^-(x)$ 在频率 $k=K,K+1,\cdots,2K$ 之外都是零。

我们选取依次选取 $K_1,K_2,K_3,\cdots$，使得对任意的 $\ell\geqslant 1$，我们都有 $K_{\ell+1}>3K_\ell$；序列 $\{K_\ell\}_{\ell\geqslant 1}$ 将在下面具体地构造。这样得到的所有 W_{K_ℓ} 和 $W_{K_\ell}^-$ 在频率空间上面的支集是两两不相交的。

根据我们提到的不等式（用 Abel 求和法来证明的），下面级数是绝对收敛的：

$$f(x)=\sum_{\ell=1}^{\infty}\frac{1}{\ell^2}W_{K_\ell}(x).$$

从而，$f\in C^0(\mathbf{T})$ 是连续函数。

我们现在考察 f 的部分和，特别是 $S_{2K_{\ell_0}}(f)$ 这一项。很明显，

$$S_{2K_{\ell_0}}f(x)=\Big(\frac{1}{\ell_0}\Big)^2W_{K_{\ell_0}}^-(x)+\sum_{\ell=1}^{\ell_0-1}\frac{1}{\ell^2}W_{K_\ell}(x).$$

由于 W_{K_ℓ} 是有界的，所以后面一项的总是有限的（小于一个固定的常数），我们不必担心；对于第一项，我们在 $x=0$ 处取值，从而，我们得到

$$\Big(\frac{1}{\ell_0}\Big)^2 W^-_{K_{\ell_0}}(0) = \Big(\frac{1}{\ell_0}\Big)^2 \sum_{k=-1}^{-K_{\ell_0}} \frac{1}{k} \sim \frac{\log(K_{\ell_0})}{\ell_0^2}.$$

为了让这一项发散（当 $\ell_0 \to \infty$ 时），我们选取 $K_\ell = 3^{\ell^3}$，其中 $\ell \geqslant 1$ 即可（此时，$K_{\ell+1} > 3K_\ell$ 自动满足）。

总结上面的的构造，我们就有

命题 369 (du Bois-Reymond)**.** *存在连续函数 $f \in C(\mathbf{T})$，使得其 Fourier 级数的部分和序列 $\{S_N(f)(x)\}_{N\geqslant 1}$ 在 $x=0$ 处发散。*

历史上有很多定理给出了充分条件保证 Fourier 级数在一点处的收敛。

定理 370. *给定函数 $f \in L^1(\mathbf{T})$ 和 $x_0 \in \mathbf{T}$。我们假设它们满足下面的要求：*

1) f 在 x_0 处的左右极限 $f_-(x_0)$ 和 $f_+(x_0)$ 都存在；

2) 存在 $\delta > 0$，使得

$$\left|\int_0^\delta \frac{|f(x_0-t)-f_-(x_0)|}{t}dt\right| + \left|\int_0^\delta \frac{|f(x_0+t)-f_+(x_0)|}{t}dt\right| < \infty.$$

那么，数列 $\{S_N(f)(x)\}_{N\geqslant 0}$ 在 $x=x_0$ 处收敛并且

$$\lim_{N\to\infty} S_N(f)(x_0) = \frac{f_-(x_0)+f_+(x_0)}{2}.$$

证明： 证明的关键是将积分改写为可以运用 Riemann-Lebesgue 引理的形式：

$$\begin{aligned}
&S_N(f)(x_0) - \frac{f_-(x_0)+f_+(x_0)}{2} \\
&= \int_{-\pi}^{\pi} D_N(x)\big(f(x_0-x)-f_-(x_0)-f_+(x_0)\big)\frac{dy}{2\pi} \\
&= \int_0^{\pi} D_N(y)\big(f(x_0-y)-f_-(x_0)+f(x_0+y)-f_+(x_0)\big)\frac{dy}{2\pi} \\
&= \int_0^{\pi} \sin\big((N+\frac{1}{2})y\big)\cdot \underbrace{\frac{y}{\sin\big(\frac{1}{2}y\big)}\cdot\frac{\big(f(x_0-y)-f_-(x_0)+f(x_0+y)-f_+(x_0)\big)}{y}}_{\in \mathbf{L}^1}\frac{dy}{2\pi}
\end{aligned}$$

从而，我们可以再次利用 Riemann-Lebesgue 引理。 □

历史上第一个关于 Fourier 级数收敛的严格数学定理是 Dirichlet 证明的（1847），它很明显是上面定理的推论：

定理 371 (Dirichlet)**.** 假设函数 f 是 $\mathbf{T}$ 上的分段 C^1 函数，即存在 $0 = x_1 < x_2 < \cdots < x_{\ell-1} < x_\ell = 2\pi = x_1$，使得 f 限制在每个区间 $[x_k, x_{k+1}]$ 上都是 C^1 的（而函数 f 本身在 x_k 处可能不连续）。那么，对任意的 $x_0 \in \mathbf{T}$，$\{S_N(f)(x)\}_{N\geqslant 0}$ 在 $x = x_0$ 处收敛并且

$$\lim_{N\to\infty} S_N(f)(x_0) = \frac{f_-(x_0) + f_+(x_0)}{2}.$$

证明: 只要对 f 用 Lagrange 中值定理就可以验证前一个定理的条件。 □

定理 372 (Dini)**.** 给定函数 $f \in L^1(\mathbf{T})$ 和 $x_0 \in \mathbf{T}$。假设他们满足

$$\left| \int_0^\pi \frac{|f(x_0 - t) + f(x_0 + t) - 2f(x_0)|}{t} dt \right| < \infty.$$

那么，$\{S_N(f)(x)\}_{N\geqslant 0}$ 在 $x = x_0$ 处收敛并且

$$\lim_{N\to\infty} S_N(f)(x_0) = f(x_0).$$

证明: 证明是类似的，我们有

$$\begin{aligned}
&S_N(f)(x_0) - f(x_0) \\
&= \int_0^\pi D_N(x)\big(f(x_0 - x) + f(x_0 + x) - 2f(x_0)\big)\frac{dx}{2\pi} \\
&= \int_0^\pi \sin\big((N + \frac{1}{2})y\big) \cdot \underbrace{\frac{y}{\sin\left(\frac{1}{2}y\right)} \cdot \frac{\big(f(x_0 - y) + f(x_0 + y) - 2f(x_0)\big)}{y}}_{\in \mathbf{L}^1} dy
\end{aligned}$$

现在可以使用 Riemann-Lebesgue 引理。 □

定义 373 (Hölder 空间)**.** 给定 $0 < \alpha < 1$。如果对于 $\mathbf{T}$ 上的连续函数 f，存在常数 $C > 0$，使得对任意的 $x, y \in \mathbb{R}$，我们都有

$$|f(x) - f(y)| \leqslant C|x - y|^\alpha,$$

我们就说 f 是**α-Hölder 连续的**。我们把所有 α-*Hölder* 连续的周期函数记作 $C^\alpha(\mathbf{T})$。

推论 374. 假设 $f \in C^\alpha(\mathbb{T})$，其中 $0 < \alpha < 1$。那么，对任意的 $x \in \mathbf{T}$，$\{S_N(f)(x)\}_{N\geqslant 0}$ 收敛并且

$$\lim_{N\to\infty} S_N(f)(x) = f(x).$$

证明: 根据 Hölder 连续性，我们有

$$\begin{aligned}
&\left| \int_0^\pi \frac{|f(x_0 - t) + f(x_0 + t) - 2f(x_0)|}{t} dt \right| \\
&\leqslant \int_0^\pi \frac{|f(x_0 - t) - f(x_0)| + |f(x_0 + t) - f(x_0)|}{t} dt \\
&\leqslant \int_0^\pi 2Ct^{\alpha-1} dt.
\end{aligned}$$

由于 $\alpha > 0$，所以上面的积分是有限的。利用 Dini 定理我们就完成了证明。 □

注记. 上述的所有命题（包括定义）对 $\alpha = 1$ 也成立。回忆一下，$\alpha = 1$ 时，我们称这样的函数 *Lipschitz* 连续的。

注记. 上述收敛未必是一致收敛，之后我们将证明，如果 $\alpha > \dfrac{1}{2}$，那么上面的收敛是一致收敛的。另外，对于 $f \in C^\alpha$，类似于可以求导数的情形，我们稍后还会证明

$$\left|\widehat{f}(k)\right| \leqslant \frac{1}{(1+|k|)^\alpha}.$$

有一个关于有界变差函数的 Fourier 级数收敛的命题和上面证明的结论也很类似。我们先引入有界变差函数的概念。给定一个有限的闭区间 $[a,b]$ 和它上面所定义的**实值**函数 f，对任意一个分划 $\sigma \in \mathcal{S}$（$\mathcal{S}$ 为分划的集合），即我们选取 $a = x_0 < x_1 < \cdots < x_m = b$，我们定义这个分划所对应把**变差**为：

$$V(f,\sigma) = \sum_{k=1}^{m} |f(x_k) - f(x_{k-1})|.$$

我们令

$$V\left([a,b],f\right) = \sup_{\sigma\in\mathcal{S}} V(f,\sigma).$$

如果 $V\left([a,b],f\right) < \infty$，我们就称 f 是 $[a,b]$ 上的**有界变差函数**。直观上，有界变差函数在区间上整体的振荡比较小。我们有如下几个显然的性质：

1) 如果 f 是 $[a,b]$ 上的单调函数，那么

$$|V([a,b],f)| = |f(a) - f(b)|.$$

2) 如果 f 和 g 是有界变差函数，那么 $f \pm g$ 也是的。

这是因为对任意分划 σ，我们都有

$$V(f \pm g,\sigma) \leqslant V(f,\sigma) + V(g,\sigma).$$

3) 如果 f 是有界变差函数，那么 $|f|$ 也是的。

这是因为对任意分划 σ，根据 $\big||x| - |y|\big| \leqslant |x-y|$，我们都有

$$V(|f|,\sigma) \leqslant V(f,\sigma).$$

给定一个 $[a,b]$ 上的右边变差函数 f，在子区间上的有界变差可以给出如下的函数：

$$[a,b] \to \mathbb{R}_{\geqslant 0},\quad x \mapsto V([a,x],f).$$

另外，假设区间 $[a,b]$ 和 $[b,c]$ 上分别有分划 σ 和 σ'，其中

$$\sigma = (a = x_0 < x_1 < \cdots < x_m = b),\quad \sigma' = (b = x_0' < x_1' < \cdots < x_{m'}' = c),$$

那么，我们可以把这两个分划首尾相接得到 $[a,c]$ 上的分划 $\sigma\cup\sigma'$，即

$$\sigma\cup\sigma'=(a=x_0<x_1<\cdots<x_m=b=x'_0<x'_1<\cdots<x'_{m'}=c).$$

很明显，如果 f 在 $[a,c]$ 上有定义，那么

$$V(f,\sigma\cup\sigma')=V(f,\sigma)+V(f,\sigma').$$

按照定义，

$$V(f,\sigma)+V(f,\sigma')=V(f,\sigma\cup\sigma')\leqslant V([a,c],f).$$

再对左边的 σ 和 σ' 取上确界，我们得到

$$V([a,b],f)+[V([b,c],f)\leqslant[V([a,c],f).$$

特别地，当 $c\geqslant b$ 时，我们有

$$V([a,b],f)\leqslant[V([a,c],f),$$

所以函数 $x\mapsto V([a,x],f)$ 是单调上升的。

定理 375 (有界变差函数的结构定理)**.** *假设 h 是 $[a,b]$ 上的有界变差函数，那么，存在 $[a,b]$ 上的单调递增函数 f 和 g，使得*

$$h=f-g.$$

证明: 只要验证 $V([a,x],h)-h(x)$ 是单调递增的就可以了，因为我们可以选取 $f=V([a,x],h)$，$g(x)=V([a,x],h)-h(x)$。任意给定 $x_1,x_2\in[a,b]$，其中 $x_1\leqslant x_2$，我们有

$$\begin{aligned}&\big(V([a,x_2],h)-h(x_2)\big)-\big(V([a,x_1],h)-h(x_1)\big)\\&=\big(V([a,x_2],h)-\big(V([a,x_1],h)\big)-(h(x_2)-h(x_1))\\&\geqslant V([x_1,x_2],h)-(h(x_2)-h(x_1))\\&\geqslant 0.\end{aligned}$$

证毕。 □

推论 376. *界变差函数是 Riemann 可积的（所以是 L^1 的，从而可以定义其 Fourier 系数）。*

证明: 这因为单调函数只有可数多个不连续点，所以是 Riemann 可积的。 □

定理 377 (Jordan)**.** *实值函数 f 是 $\mathbf{T}$ 上的有界变差函数。那么，对任意 $x_0\in\mathbf{T}$，$\{S_N(f)(x)\}_{N\geqslant1}$ 在 x_0 处收敛并且*

$$\lim_{N\to\infty}S_N(f)(x_0)=\frac{f_-(x_0)+f_+(x_0)}{2}.$$

证明：根据有界变差函数的结构定理，我们不妨假设 f 是单调函数。仿照之前定理的证明，我们有

$$\begin{aligned}
&S_N(f)(x_0)-\frac{f_-(x_0)+f_+(x_0)}{2}\\
=&\int_0^\pi D_N(x)\big(f(x_0-y)-f_-(x_0)+f(x_0+y)-f_+(x_0)\big)\frac{dy}{2\pi}\\
=&\underbrace{\int_0^\pi \frac{\sin\big((N+\frac12)y\big)}{\sin\big(\frac12 y\big)}\big(f(x_0-y)-f_-(x_0)\big)\frac{dy}{2\pi}}_{I_-}+\underbrace{\int_0^\pi \frac{\sin\big((N+\frac12)y\big)}{\sin\big(\frac12 y\big)}\big(f(x_0+y)-f_+(x_0)\big)\frac{dy}{2\pi}}_{I_+}.
\end{aligned}$$

只要处理 I_+ 即可，另外一项可以类似地处理。我们可以将 I_+ 写为

$$I_+=\underbrace{\int_0^\delta \frac{\sin\big((N+\frac12)y\big)}{\sin\big(\frac12 y\big)}\big(f(x_0+y)-f_+(x_0)\big)\frac{dy}{2\pi}}_{I_{+,\delta}}+\underbrace{\int_\delta^\pi \sin\big((N+\frac12)y\big)\overbrace{\frac{f(x_0+y)-f_+(x_0)}{\sin\big(\frac12 y\big)}}^{\in L^1}\frac{dy}{2\pi}}_{I_{+,>\delta}},\to 0.$$

上式的第二项可以再次用 Riemann-Lebesgue 引理处理：当 $N\to\infty$ 时候，我们有 $I_{+,>\delta}\to 0$。为了处理第一项，我们用之前在估计 $\int_0^{2\pi}|D_N(x)|\frac{dx}{2\pi}$ 时所用的不等式：

$$\frac{2}{t}\leqslant\frac{1}{\sin\big(\frac{t}{2}\big)}\leqslant\frac{2}{t}+\frac{t}{6},\quad 0\leqslant t\leqslant\pi,$$

从而对任意的 $g(t)$（分正负来讨论），我们都有

$$\Big|\frac{1}{\sin\big(\frac{t}{2}\big)}g(t)-\frac{2}{t}g(t)\Big|\leqslant\frac{t}{6}|g(t)|.$$

据此，

$$\begin{aligned}
&\Big|I_{+,\delta}-\underbrace{\int_0^\delta\frac{2\sin\big((N+\frac12)y\big)}{y}\big(f(x_0+y)-f_+(x_0)\big)\frac{dy}{2\pi}}_{\widetilde{I_{+,\delta}}}\Big|\\
\leqslant&\underbrace{\frac16\int_0^\delta\Big|\sin\big((N+\frac12)y\big)\Big|\times y\times\big(f(x_0+y)-f_+(x_0)\big)\frac{dy}{2\pi}}_{\widetilde{\widetilde{I_{+,\delta}}}}.
\end{aligned}$$

综合上面的所有的不等式，我们有

$$|I_+|\leqslant|\widetilde{I_{+,\delta}}|+\widetilde{\widetilde{I_{+,\delta}}}+I_{+,>\delta}.$$

现在任意给定 $\varepsilon>0$，我们要找一个足够大的 N_0，使得当 $N\geqslant N_0$ 时，有

$$|I_+|<\varepsilon.$$

首先，我们有

$$\begin{aligned}\widetilde{\widetilde{I_{+,\delta}}} &\leqslant \int_0^\delta y \times |f(x_0+y)-f_+(x_0)|\frac{dy}{12\pi}\\ &\leqslant \int_0^\delta |f(x_0+y)-f_+(x_0)|\frac{dy}{12\pi}.\end{aligned}$$

被积分项是有界的，所以存在 $\delta_1>0$，使得当 $\delta<\delta_1$ 时，

$$\widetilde{\widetilde{I_{+,\delta}}} \leqslant \frac{1}{3}\varepsilon.$$

其次，我们考虑 $\widetilde{\widetilde{I_{+,\delta}}}$。利用 f 的单调性，我们可以运用积分第二中值定理。[16] 从而，存在 $c\in(0,\delta)$，使得

$$\begin{aligned}\widetilde{\widetilde{I_{+,\delta}}} &= \int_0^\delta \frac{\sin\left((N+\frac{1}{2})y\right)}{y}\big(f(x_0+y)-f_+(x_0)\big)\frac{dy}{2\pi}\\ &= \big(f(x_0+\delta)-f_+(x_0)\big)\underbrace{\int_{(N+\frac{1}{2})c}^{(N+\frac{1}{2})\delta}\frac{\sin t}{t}\frac{dt}{2\pi}}_{<\infty}.\end{aligned}$$

由于 $\int_1^\infty \frac{\sin t}{t}dt<\infty$，所以上面的积分项是有界的。根据右极限定义，存在 $\delta_2>0$，使得当 $\delta<\delta_2$ 时，我们有

$$\widetilde{\widetilde{I_{+,\delta}}} < \frac{1}{3}\varepsilon.$$

我们现在选取 $\delta<\min(\delta_1,\delta_2)$，从而

$$|I_+|<\frac{\varepsilon}{3}+\frac{\varepsilon}{3}+I_{+,>\delta}.$$

此时，我们可以对 $I_{+,>\delta}$ 用 Riemann-Lebesgue 引理，所以存在 N_0，使得当 $N\geqslant N_0$ 时，我们有

$$I_{+,>\delta}<\frac{\varepsilon}{3}.$$

所以，

$$|I_+|<\varepsilon.$$

这就完成了 Jordan 定理的证明。 □

[16] 假设 $I=[a,b]$，$g\in R(I)$，f 是 I 上单调递增的函数（未必严格）。那么，存在 $c\in[a,b]$，使得

$$\int_a^b fg = f(a)\int_a^c g + f(b)\int_c^b g(t)dt.$$

55.1 作业：Fourier 级数的计算，三角函数与球谐函数

清华大学 19-20 春季学期，数学分析二，作业 11

请用 A4 大小的纸张**正反面**用钢笔，签字笔或者圆珠笔书写，并注明自己的姓名，年级和作业的**总页数**。除定理公式所涉及的人名之外，请使用**中文**。本次作业的提交时间和地点为 **2 月 113 日**上午的课堂上，逾期视作零分。

Fourier 级数的计算

F1) 给定 $f \in L^1(\mathbf{T})$。证明，f 是实值函数（几乎处处）当且仅当对任意的 $k \in \mathbb{Z}$，我们有

$$\overline{\widehat{f}(k)} = \widehat{f}(-k).$$

F2) f 是 $\mathbb{R}$ 上以 2π 为周期的奇函数，它在 $[0, \pi]$ 上的定义为

$$f(x) = x(\pi - x).$$

证明，对任意的 $x \in \mathbb{R}$，我们都有

$$f(x) = \frac{8}{\pi} \sum_{k=1}^{\infty} \frac{\sin\big((2k-1)x\big)}{(2k-1)^3}.$$

F3) f 是 $\mathbb{R}$ 上以 2π 为周期的函数，它在 $[-\pi, \pi]$ 上的定义为

$$f(x) = |x|.$$

试计算 f 的 Fourier 系数。通过考虑 $f(0)$ 证明：

$$\sum_{k=1}^{\infty} \frac{1}{k^2} = \frac{1}{6}\pi^2.$$

进一步证明，

$$\sum_{k=1}^{\infty} \frac{1}{k^4} = \frac{1}{90}\pi^4.$$

F4) $\alpha \in \mathbb{R} - \mathbb{Z}$，$f$ 是 $\mathbb{R}$ 上以 2π 为周期的函数，它在 $[0, 2\pi]$ 上的定义为

$$f(x) = \frac{\pi}{\sin(\pi\alpha)} e^{i(\pi - x)\alpha}.$$

通过计算它的 Fourier 系数，证明，

$$\sum_{k \in \mathbb{Z}} \frac{1}{(k+\alpha)^2} = \frac{\pi^2}{\sin(\pi\alpha)^2}.$$

F5) 利用 F2) 中的函数，证明

$$\sum_{k=1}^{\infty} \frac{1}{k^6} = \frac{1}{945}\pi^6.$$

F6) 证明，连续函数的 Fourier 系数可以衰减的任意慢，即任给正数的序列 $\{b_k\}_{k \geqslant 1}$，$b_k \to 0$，总存在连续函数 $f \in C(\mathbf{T})$，使得存在无限多个 $k \geqslant 0$，$|\widehat{f}(k)| \geqslant b_k$。

三角函数与球谐函数

我们用 $\mathbf{S}^1 = \mathbf{T}$ 表示 $\mathbb{R}^2$ 中的单位球面

$$\mathbf{S}^1 = \{(x, y) \in \mathbb{R}^3 | x^2 + y^2 = 1\}.$$

我们可以用极坐标系 $(r = 1, \theta)$ 来表示 $\mathbf{T}$ 上的点，其中 $\theta \in [0, 2\pi]$。

s1) 证明，在 θ 的参数表示下，$\mathbf{T}$ 上的曲面测度为 $\sigma = d\theta$。我们用上述测度来定义 $L^2(\mathbf{T})$（和课堂上有区别）：对于 $f, g \in L^2(\mathbf{T})$，它们的内积是

$$(f_1, f_2) = \int_{\mathbf{T}} f_1(\theta)\overline{f_2(\theta)}\, d\theta.$$

s2) 对于 $\mathbb{R}^2$ 上的 Laplace 算子 $\triangle = \dfrac{\partial^2}{\partial x^2} + \dfrac{\partial^2}{\partial y^2}$，证明（上学期已经证明），

$$\triangle = \frac{\partial^2}{\partial r^2} + \frac{1}{r}\frac{\partial}{\partial r} + \frac{1}{r^2}\triangle_{\mathbf{T}},$$

其中 $\triangle_{\mathbf{T}} = \dfrac{\partial^2}{\partial \theta^2}$。

s3) 证明，$\triangle_{\mathbf{T}}$ 对于上述的 L^2 内积具有自伴性：对任意的 $f, g \in C^\infty(\mathbf{T})$，我们有

$$(\triangle_{\mathbf{T}} f_1, f_2) = (f_1, \triangle_{\mathbf{T}} f_2).$$

（自伴性可以类比成线性代数中的对称矩阵）

s4) 任给 $g \in \mathbf{SO}(2)$（即行列式为 1 的 2×2 正交矩阵）和 $\mathbb{R}^2$ 上的函数 f，我们定义

$$g \bullet f(x) = f(g^{-1}x),$$

其中 $g^{-1}x$ 是 2×2 的矩阵在 2 维的向量上的乘法。证明，对任意的 $f_1, f_2 \in L^2(\mathbf{T})$ 和任意的 $g \in \mathbf{SO}(2)$，我们有

$$(g \bullet f_1, g \bullet f_2) = (f_1, f_2).$$

s5) 我们将 $\mathbb{R}^2$ 中满足 $\triangle f = 0$ 的 C^2 函数称作是**调和函数**。证明，如果 f 是调和函数，那么对任意的 $g \in \mathbf{SO}(2)$，$g \bullet f$ 也是调和函数。

s6) 我们用 $\mathbf{P}^{(\ell)}$ 表示复系数的次数为 ℓ 的关于 x 和 y 的二元齐次多项式，这是线性空间，其中的元素形如

$$\sum_{\substack{\ell_1 + \ell_1 = \ell \\ 0 \leqslant \ell_1, \ell_2 \leqslant \ell}} c_{\ell_1, \ell_2} x^{\ell_1} y^{\ell_2}, \quad c_{\ell_1, \ell_2} \in \mathbb{C}.$$

证明，$\dim_{\mathbb{C}} \mathbf{P}^{(\ell)} = \ell + 1$ 并且对任意的 $P \in \mathbf{P}^{(\ell)}$ 和任意的 $g \in \mathbf{SO}(2)$，$g \bullet P \in \mathbf{P}^{(\ell)}$。

s7) 证明，对于 $\ell \geqslant 2$，Laplace 算子给出的映射

$$\triangle : \mathbf{P}^{(\ell)} \to \mathbf{P}^{(\ell-2)}$$

是线性映射并且是满射。

s8) 我们定义 $\mathbb{R}^2$ 上 ℓ-次的调和多项式为：

$$\mathbf{H}^{(\ell)} = \{P \in \mathbf{P}^{(\ell)} | \triangle P = 0\}.$$

证明，如果 $|\ell| \geqslant 1$，那么 $\dim_{\mathbb{C}} \mathbf{H}^{(\ell)} = 2$ 并且 $\mathbf{H}^{(\ell)}$ 在 $\mathbf{SO}(2)$ 的作用下不变：对任意的 $P \in \mathbf{H}^{(\ell)}$ 和任意的 $g \in \mathbf{SO}(2)$，$g \bullet P \in \mathbf{H}^{(\ell)}$。

注记. 上述证明了 $\mathbf{H}^{(\ell)}$ 是 $\mathbf{SO}(2)$ 的一个 2 维的表示。

s9) 证明，对任意的 ℓ，如果 $P \in \mathbf{H}^{(\ell)}$，那么 $p = P\big|_{\mathbf{T}}$ 满足

$$-\triangle p = \ell^2 p.$$

s10) 证明，$\mathbf{H}^{(\ell)}$ 可以由 $e_\ell = (x+iy)^\ell$ 和 $e_{-\ell} = (x-iy)^\ell$ 生成。

s11) 证明，$\left\{\frac{1}{\sqrt{2\pi}} e_k \middle| k \in \mathbb{Z}\right\}$ 是 $L^2(\mathbf{T})$ 的一个 Hilbert 基。

注记. 上述表明，$L^2(\mathbf{T})$ 的 *Hilbert* 基 $\left\{e^{ikx} \middle| k \in \mathbb{Z}\right\}$ 可以由 $\mathbb{R}^2$ 上的齐次调和多项式限制得到。我们下面要把这个结论推广到 $\mathbf{S}^2$。

我们用 $\mathbf{S}^2$ 表示 $\mathbb{R}^3$ 中的单位球面

$$\mathbf{S}^2 = \{(x_1, x_2, x_3) \in \mathbb{R}^3 | (x_1)^2 + (x_2)^2 + (x_3)^2 = 1\}.$$

我们可以用球面坐标系 $(r = 1, \vartheta, \varphi)$ 来表示 $\mathbf{S}^2$ 上的点，其中 $(\vartheta, \varphi) \in [0, \pi] \times [0, 2\pi]$。

S1) 证明，在 (ϑ, φ) 的参数表示下，$\mathbf{S}^2$ 上的曲面测度为 $\sigma = \sin\vartheta \, d\vartheta d\varphi$。我们用上述测度来定义 $L^2(\mathbf{S}^2)$：对于 $f, g \in L^2(\mathbf{S}^2)$，它们的内积是

$$(f_1, f_2) = \int_{\mathbf{S}^2} f_1(\vartheta, \varphi) \overline{f_2(\vartheta, \varphi)} \sin\vartheta \, d\vartheta d\varphi.$$

S2) 对于 $\mathbb{R}^3$ 上的 Laplace 算子 $\triangle = \sum_{i=1}^{3} \frac{\partial^2}{\partial x_i^2}$，证明（之前的作业已经证明），

$$\triangle = \frac{\partial^2}{\partial r^2} + \frac{2}{r}\frac{\partial}{\partial r} + \frac{1}{r^2} \triangle_{\mathbf{S}^2},$$

其中 $\triangle_{\mathbf{S}^2} = \frac{\partial^2}{\partial \vartheta^2} + \frac{\cos\vartheta}{\sin\vartheta}\frac{\partial}{\partial\vartheta} + \frac{1}{(\sin\vartheta)^2}\frac{\partial^2}{\partial\varphi^2}$。

S3) 证明，$\triangle_{\mathbf{S}^2}$ 对于上述的 L^2 内积具有自伴性：对任意的 $f, g \in C^\infty(\mathbf{S}^2)$，我们有

$$(\triangle_{\mathbf{S}^2} f_1, f_2) = (f_1, \triangle_{\mathbf{S}^2} f_2).$$

S4) 任给 $g \in \mathbf{SO}(3)$（即行列式为 1 的 3×3 正交矩阵）和 $\mathbb{R}^3$ 上的函数 f，我们定义

$$g \bullet f(x) = f(g^{-1}x),$$

其中 $g^{-1}x$ 是 3×3 的矩阵在 3 维的向量上的乘法。证明，对任意的 $f_1, f_2 \in L^2(\mathbf{S}^2)$ 和任意的 $g \in \mathbf{SO}(3)$，我们有

$$(g \bullet f_1, g \bullet f_2) = (f_1, f_2).$$

S5) 我们将 $\mathbb{R}^3$ 中满足 $\triangle f = 0$ 的 C^2 函数称作是**调和函数**。证明，如果 f 是调和函数，那么对任意的 $g \in \mathbf{SO}(3)$，$g \bullet f$ 也是调和函数。

S6) 我们用 $\mathbf{P}^{(\ell)}$ 表示复系数的次数为 ℓ 的关于 x_1, x_2, x_3 的齐次多项式，这显然是一个线性空间，其中的元素形如

$$\sum_{\substack{\ell_1+\ell_1+\ell_3=\ell \\ 0\leqslant \ell_1,\ell_2,\ell_3 \leqslant \ell}} c_{\ell_1,\ell_2,\ell_3} (x_1)^{\ell_1} (x_2)^{\ell_2} (x_3)^{\ell_3}, \quad c_{\ell_1,\ell_2,\ell_3} \in \mathbb{C}.$$

证明，$\dim_{\mathbb{C}} \mathbf{P}^{(\ell)} = \dfrac{1}{2}(\ell+1)(\ell+2)$ 并且对任意的 $P \in \mathbf{P}^{(\ell)}$ 和任意的 $g \in \mathbf{SO}(3)$，$g \bullet P \in \mathbf{P}^{(\ell)}$。

S7)* 证明，对于 $\ell \geqslant 2$，由 Laplace 算子给出的映射

$$\triangle : \mathbf{P}^{(\ell)} \to \mathbf{P}^{(\ell-2)}$$

是线性映射并且是满射。

S8) 我们定义调和多项式的空间：

$$\mathbf{H}^{(\ell)} = \{P \in \mathbf{P}^{(\ell)} | \triangle P = 0\}.$$

证明，$\dim_{\mathbb{C}} \mathbf{H}^{(\ell)} = 2\ell + 1$ 并且 $\mathbf{H}^{(\ell)}$ 在 $\mathbf{SO}(3)$ 的作用下不变：对任意的 $P \in \mathbf{H}^{(\ell)}$ 和任意的 $g \in \mathbf{SO}(3)$，$g \bullet P \in \mathbf{H}^{(\ell)}$。

注记. 上述实际上证明了 $\mathbf{H}^{(\ell)}$ 是 $\mathbf{SO}(3)$ 的一个 $2\ell+1$ 维的表示，利用一点点关于 $\mathbf{SU}(2)$ 和 $\mathbf{SO}(3)$ 的表示论（这是一种关于群作用的线性代数），我们容易证明 $\mathbf{H}^{(\ell)}$ 是 $\mathbf{SO}(3)$ 的不可约表示并且当 ℓ 遍历 $0, 1, 2, \cdots$ 时，我们得到了 $\mathbf{SO}(3)$ 的所有不可约表示。

S9) 对任意的 $P \in \mathbf{P}^{(\ell)}$ 和非负整数 k，证明，

$$\triangle(r^{2k}P) = 2k(2\ell - 2k + 1) r^{2k-2} P + r^{2k} \triangle P.$$

（提示：利用齐次函数的 Euler 等式）

S10)* 证明，如果 $P \in \mathbf{P}^{(\ell)}$ 并且 $r^2 P \in \mathbf{H}^{(\ell+2)}$，那么 $P = 0$。（提示：考虑 k_0 是使得 r^{2k_0} 整除 P 的最大整数）

S11) 证明，$\mathbf{P}^{(\ell)} = \mathbf{H}^{(\ell)} \oplus r^2 \mathbf{P}^{(\ell-2)}$，$\mathbf{P}^{(\ell)} = \mathbf{H}^{(\ell)} \oplus r^2 \mathbf{H}^{(\ell-2)} \oplus r^4 \mathbf{H}^{(\ell-4)} \oplus \cdots$。

S12) 令 $\mathbf{H}^\ell = \big\{ P\big|_{\mathbf{S}^2} \big| P \in \mathbf{H}^{(\ell)} \big\}$，即齐次调和多项式在球面上的限制，我们将这些函数称作是**球谐函数**。证明，在内积

$$(f_1, f_2) = \int_0^\pi \int_0^{2\pi} f_1(\vartheta, \varphi) \overline{f_2(\vartheta, \varphi)} \, \sin\vartheta \, d\vartheta d\varphi$$

下，$\mathbf{H}^\ell$ 是有限维的 Hermite 内积空间，这个内积是 $\mathbf{SO}(3)$ 不变的（即对任意的 $g \in \mathbf{SO}(3)$，任意的 $f_1, f_2 \in \mathbf{H}^\ell$，$(g \bullet f_1, g \bullet f_2) = (f_1, f_2)$）并且作为线性空间，我们有同构

$$\mathbf{P}^\ell = \mathbf{H}^\ell \oplus \mathbf{H}^{\ell-2} \oplus \mathbf{H}^{\ell-4} \oplus \cdots,$$

其中 $\mathbf{P}^\ell = \Big\{ P\big|_{\mathbf{S}^2} \Big| P \in \mathbf{P}^{(\ell)} \Big\}$。

S13) 我们用 $\mathrm{Res} : \mathbf{P}^{(\ell)} \to \mathbf{P}^\ell$ 代表到球面的限制映射，用 $\mathrm{Ext} : \mathbf{P}^\ell \to \mathbf{P}^{(\ell)}$ 表示齐次扩张映射，其中

$$\big(\mathrm{Ext}(f)\big)(r, \vartheta, \varphi) = r^\ell f(\vartheta, \varphi),$$

这两个映射互为逆映射。令

$$\begin{aligned} X_1 &= x_3 \frac{\partial}{\partial x_2} - x_2 \frac{\partial}{\partial x_3}, \\ X_2 &= x_1 \frac{\partial}{\partial x_3} - x_3 \frac{\partial}{\partial x_1}, \\ X_3 &= x_2 \frac{\partial}{\partial x_1} - x_1 \frac{\partial}{\partial x_2}, \end{aligned}$$

为 $\mathbb{R}^3$ 上的三个向量场（它们落在 $\mathbf{SO}(3)$ 的 Lie 代数中），对于 $P \in \mathbf{H}^\ell$，我们定义

$$D^i P = \mathrm{Res}\Big(\nabla_{X_i} \big(\mathrm{Ext}(P) \big) \Big), \quad i = 1, 2, 3.$$

证明，$D_i : \mathbf{P}^\ell \to \mathbf{P}^\ell$，$D_i : \mathbf{H}^\ell \to \mathbf{H}^\ell$ 并且

$$\begin{aligned} D_1 &= \sin\varphi \frac{\partial}{\partial\vartheta} + \cos\varphi \frac{\cos\vartheta}{\sin\vartheta} \frac{\partial}{\partial\varphi}, \\ D_2 &= -\cos\varphi \frac{\partial}{\partial\vartheta} + \sin\varphi \frac{\cos\vartheta}{\sin\vartheta} \frac{\partial}{\partial\varphi}, \\ D_3 &= -\frac{\partial}{\partial\varphi}. \end{aligned}$$

S14) 我们令

$$\begin{aligned}
J_1 &= i\Big(\sin\varphi\frac{\partial}{\partial\vartheta}+\cos\varphi\frac{\cos\vartheta}{\sin\vartheta}\frac{\partial}{\partial\varphi}\Big),\\
J_2 &= i\Big(-\cos\varphi\frac{\partial}{\partial\vartheta}+\sin\varphi\frac{\cos\vartheta}{\sin\vartheta}\frac{\partial}{\partial\varphi}\Big),\\
J_3 &= -i\frac{\partial}{\partial\varphi},\\
J_\pm &= J_1\pm iJ_2.
\end{aligned}$$

证明，对任意的 $f_1,f_2\in C^\infty(\mathbf{S}^2)$，我们有

$$(J_kf_1,f_2)=(f_1,J_kf_2),\ \ (J_+f_1,f_2)=(f_1,J_-f_2),\ \ (J_-f_1,f_2)=(f_1,J_+f_2).$$

（提示：可能将 D_k 写成 $\mathbf{SO}(3)$ 中一条曲线的切方向 $D_k=\frac{d}{dt}\big|_{t=0}g(t)$，其中 $g(t)\in\mathbf{SO}(3)$，会使得证明在概念上更清晰）

S15) 我们定义如下的 Casimir 算子：

$$J^2=J_1\circ J_1+J_2\circ J_2+J_3\circ J_3=J_+\circ J_-+J_3\circ J_3-J_3.$$

证明，$J^2=-\triangle_{\mathbf{S}^2}$。

S16) 对于 $Y(\vartheta,\varphi)\in\mathbf{P}^\ell$，我们令

$$P(r,\vartheta,\varphi)=\mathrm{Ext}(Y)(r,\vartheta,\varphi)=r^\ell Y(\vartheta,\varphi).$$

证明，$\triangle P=0$ 等价于 $\triangle_{\mathbf{S}^2}Y=-\ell(\ell+1)Y$。

S17) 证明，$-\triangle_{\mathbf{S}^2}:\mathbf{H}^\ell\to\mathbf{H}^\ell$ 并且 $-\triangle_{\mathbf{S}^2}=\ell(\ell+1)\mathbf{Id}$。

S18)* 证明，$\mathrm{Span}\big\{\mathbf{H}^\ell\big|\ell=1,2,3,\cdots\big\}$ 在 $L^2(\mathbf{S}^2)$ 是稠密的。

（提示：利用关于多项式（多元）的 Stone-Weierstrass 定理）

注记. 上述表明，$L^2(\mathbf{S}^2)$ 可以写成直和（*Hilbert* 空间意义下的）$L^2(\mathbf{S}^2)=\bigoplus_{\ell0}^{\infty}\mathbf{H}^\ell$，其中，$\mathbf{H}^\ell$ 恰为 $-\triangle_{\mathbf{S}^2}$ 以 $\ell(\ell+1)$ 为特征值的特征子空间，它的维数是 $\ell(\ell+1)$。以下，我们要在每个 $\mathbf{H}^\ell$ 中选取一组单位正交基 $\big\{Y_m^\ell|m=-\ell,\cdots,-1,0,1,\cdots,\ell\big\}$，此时所有的 $\big\{Y_m^\ell|\ell\geqslant0,-\ell\leqslant m\leqslant\ell\big\}$ 构成了 $L^2(\mathbf{S}^2)$ 的一个 *Hilbert* 基。线性算子 $-\triangle_{\mathbf{S}^2}$ 在这组基下被对角化。

我们可以将这些要找的函数想象成 $\mathbf{T}$ 上的三角函数 e^{ikx} 的类比。

S19) 对于 $m=0,1,2,\cdots,\ell$，我们定义两组函数

$$\begin{aligned}
P_m^\ell &= \left(\frac{d}{dx}\right)^{\ell+m}\Big((1-x^2)^\ell\Big),\\
Z_m^\ell(\vartheta) &= \big(\sin\vartheta\big)^m P_m^\ell(\cos\vartheta),
\end{aligned}$$

和一组常数

$$C_{m,\ell}=\frac{(-1)^{\ell+m}}{2^\ell \ell!}\sqrt{\frac{2\ell+1}{4\pi}}\sqrt{\frac{(\ell-m)!}{(\ell+m)!}}.$$

据此，我们定义

$$Y_m^\ell(\vartheta,\varphi)=C_{m,\ell}Z_m^\ell(\vartheta)e^{im\varphi}.$$

当 $m=-\ell,-\ell+1,\cdots,-1$ 时，我们定义

$$Y_m^\ell(\vartheta,\varphi)=\overline{Y_{-m}^\ell(\vartheta,\varphi)}.$$

证明，上述定义的 Y_m^ℓ（$-\ell\leqslant m\leqslant\ell$）都是 $\mathbf{S}^2$ 上的非零的光滑函数。

S20) 试写下 Y_m^ℓ（$\ell=0,1,2$，$-\ell\leqslant m\leqslant\ell$）这九个函数并证明它们 $L^2(\mathbf{S}^2)$ 是长度为 1 且相互正交的。

S21) 证明，对任意的 ℓ 和 m（$-\ell\leqslant m\leqslant\ell$），我们有下面的恒等式：

$$\begin{aligned}-\triangle_{\mathbf{S}^2}Y_m^\ell&=\ell(\ell+1)Y_m^\ell,\\ J_3Y_m^\ell&=mY_m^\ell,\\ J_+Y_m^\ell&=\sqrt{(\ell-m)(\ell+m+1)}Y_{m+1}^\ell,\\ J_-Y_m^\ell&=\sqrt{(\ell+m)(\ell-m+1)}Y_{m-1}^\ell.\end{aligned}$$

S22) 证明，对任意的 ℓ 和 m（$-\ell\leqslant m\leqslant\ell$），$Y_m^\ell\in\mathbf{H}^\ell$。

S23) 证明，对任意的 $(\ell,m)\neq(\ell',m')$（$-\ell\leqslant m\leqslant\ell$，$-\ell'\leqslant m'\leqslant\ell'$），函数 Y_m^ℓ 和 $Y_{m'}^{\ell'}$ 是垂直的，即 $(Y_m^\ell,Y_{m'}^{\ell'})=0$。

S24)* 固定 ℓ，证明，对任意的 $-\ell\leqslant m\leqslant\ell$，$Y_m^\ell$ 的长度 $\sqrt{(Y_m^\ell,Y_m^\ell)}$ 不依赖于 m。

（提示：利用 S21) 的结论）

S25) 证明，对任意的 ℓ 和 m（$-\ell\leqslant m\leqslant\ell$），$(Y_m^\ell,Y_m^\ell)=1$。

注记. 上述表明，我们可以取 $\left\{Y_m^\ell\middle|\ell\geqslant 0,-\ell\leqslant m\leqslant\ell\right\}$ 作为 $L^2(\mathbf{S}^2)$ 的 *Hilbert* 基。

上述构造的一个核心想法就是 $-\triangle_{\mathbf{S}^2}$ 与算子 J_3 是交换的，从而它们应该可以被同时（正交）对角化。这个问题的计算对应着量子力学中的 *Zeeman* 效应。

Derrière la série de Fourier, d'autres séries analogues sont entrées dans la domaine de l'analyse; elles y sont entrées par la même porte; elles ont été imaginées en vue des applications.

—- Henri Poincaré

---

56 Bernstein 定理，Fourier 级数在等分布问题上的应用

二零二零年五月二十五日，星期一，晴

Hölder 连续函数的 Bernstein 定理

我们考虑 $f \in C^{\alpha}(\mathbf{T})$，其中 $\alpha \in (0,1)$。按照定义，存在 $C>0$，使得对任意的 $x,y \in \mathbf{T}$，我们有

$$|f(x)-f(y)| \leqslant C|x-y|^{\alpha}.$$

令 $\|f\|_{C^{0,\alpha}}$ 是上述可能的常数 C 的下确界，即

$$\|f\|_{C^{0,\alpha}} = \inf_{x,y\in\mathbf{T},x\neq y} \frac{|f(x)-f(y)|}{|x-y|^{\alpha}}.$$

从而，对任意的 $x,y \in \mathbf{T}$，我们有

$$|f(x)-f(y)| \leqslant \|f\|_{C^{0,\alpha}}|x-y|^{\alpha}.$$

命题 378. *假设 $f \in C^{\alpha}(\mathbf{T})$，其中我们假设 $\alpha \in [0,1]$（当 $\alpha=0$ 时，我们只是假设 f 是连续的；当 $\alpha=1$ 时，我们假设 f 是 Lipschitz 函数）。那么，我们有*

$$\left|\widehat{f}(k)\right| = O(|k|^{-\alpha}),$$

即序列 $\left\{|k|^{\alpha}\left|\widehat{f}(k)\right|\right\}_{k\geqslant 1}$ 是有界的。

证明: 我们不妨假设 $\alpha>0$（否则用 Riemann-Lebesgue 引理即可），按照定义，对任意的 $x,y\in\mathbf{T}$，我们有

$$|f(x)-f(y)| \leqslant \|f\|_{C^{0,\alpha}}|x-y|^{\alpha}.$$

我们将满足对任意的 $k\neq 0$，我们有

$$\begin{aligned}
\widehat{f}(k) &= \int_{-\pi}^{\pi} f(x)e^{-ikx}\frac{dx}{2\pi} \\
&= -\int_{-\pi}^{\pi} f(x)e^{-ikx+i\pi}\frac{dx}{2\pi} \\
&= -\int_{-\pi}^{\pi} f(x)e^{-ik(x-\frac{\pi}{k})}\frac{dx}{2\pi} \\
&= -\int_{-\pi}^{\pi} f(x+\frac{\pi}{k})e^{-ixk}\frac{dx}{2\pi}.
\end{aligned}$$

从而，

$$\begin{aligned}
\widehat{f}(k) &= \frac{1}{2}\left(\int_{-\pi}^{\pi} f(x)e^{-ikx}\frac{dx}{2\pi} - \int_{-\pi}^{\pi} f(x+\frac{\pi}{k})e^{-ixk}\frac{dx}{2\pi}\right) \\
&= \int_{-\pi}^{\pi} \big(f(x+\frac{\pi}{k})-f(x)\big)e^{-ixk}\frac{dx}{4\pi}.
\end{aligned}$$

所以，利用 Hölder 函数的定义，我们有

$$\begin{aligned}|\widehat{f}(k)| &\leqslant \frac{1}{4\pi}\int_{-\pi}^{\pi}|f(x+\frac{\pi}{k})-f(x)|dx\\ &\leqslant \frac{\|f\|_{C^{0,\alpha}}}{4\pi}\int_{-\pi}^{\pi}\left(\frac{\pi}{|k|}\right)^{\alpha}dx=\frac{\pi^{\alpha}\|f\|_{C^{0,\alpha}}}{2}\cdot\frac{1}{|k|^{\alpha}}.\end{aligned}$$

□

注记. 上面命题中证明的关于 C^{α} 函数的 *Fourier* 系数的衰减估计是最佳的，其中 $\alpha\in(0,1)$（实际上，对于 $\alpha\in(0,1)$，这个衰减估计刻画了 $C^{\alpha}(\mathbf{T})$，我们这里不给出证明了）。

假设 $0<\alpha<1$，我们考虑如下的 *Fourier* 级数：

$$f_{\alpha}(x)=\sum_{k=1}^{\infty}\frac{1}{2^{k\alpha}}e^{i2^{k}x}.$$

很显然，这个级数是绝对收敛的，所以，$f_{\alpha}\in C^{0}(\mathbf{T})$。特别地，$\widehat{f_{\alpha}}(\ell)=|\ell|^{-\alpha}$ 并且 $\widehat{f_{\alpha}}(\ell)=o(|\ell|^{-\alpha})$ 不成立。

如果我们可以说明 $f_{\alpha}\in C^{\alpha}(\mathbf{T})$，这表明上述性质中对于 *Fourier* 系数的衰减估计是最佳的：

对于任意的 x 和 h，以 h 为频率的二进制尺度，我们有

$$f_{\alpha}(x+h)-f_{\alpha}(x)=\underbrace{\sum_{2^{k}\leqslant\frac{1}{100}|h|^{-1}}\frac{1}{2^{k\alpha}}\left(e^{i2^{k}(x+h)}-e^{i2^{k}x}\right)}_{I_1}+\underbrace{\sum_{2^{k}>\frac{1}{100}|h|^{-1}}\frac{1}{2^{k\alpha}}\left(e^{i2^{k}(x+h)}-e^{i2^{k}x}\right)}_{I_2}.$$

我们注意到第二项可以如下控制

$$|I_2|\leqslant\sum_{2^{k}>\frac{1}{100}|h|^{-1}}\frac{2}{2^{k\alpha}}\leqslant C|h|^{\alpha}.$$

对于第一项，我们有

$$\begin{aligned}|I_1|&\leqslant\sum_{2^{k}\leqslant\frac{1}{100}|h|^{-1}}\frac{1}{2^{k\alpha}}|e^{i2^{k}h}-1|\\ &\leqslant\sum_{2^{k}\leqslant\frac{1}{100}|h|^{-1}}\frac{C}{2^{k\alpha}}2^{k}|h|.\end{aligned}$$

我们用到了当 $|y|<1$ 时，$|e^{iy}-1|\leqslant C|y|$。从而，

$$\begin{aligned}|I_1|&\leqslant C|h|\sum_{2^{k}\leqslant\frac{1}{100}|h|^{-1}}2^{k(1-\alpha)}\\ &\leqslant C|h|\times C'\left(\frac{1}{100}|h|^{-1}\right)^{1-\alpha}\\ &=C''|h|^{\alpha}.\end{aligned}$$

在这个估计中，我们用到了 $\alpha<1$。所以，

$$|f_\alpha(x+h)-f_\alpha(x)|\leqslant |I_1|+|I_2|\leqslant C'''|h|^\alpha.$$

这表明，$f\in C^\alpha(\mathbf{T})$。

我们已经介绍的几个关于 Fourier 级数逐点收敛的经典定理，都没有涉及到一致收敛性。如果 $\alpha>\dfrac{1}{2}$，尽管对于 $f\in C^\alpha(\mathbf{T})$，它的 Fourier 系数衰减的并不迅速（不是绝对收敛的），但是 Bernstein 有如下令人惊讶的结果：

定理 379 (Bernstein)**.** 如果 $\alpha\in(\dfrac{1}{2},1]$，那么，对于任意的 $f\in C^\alpha(\mathbf{T})$，其的 *Fourier* 级数

$$\sum_{k\in\mathbb{Z}}\widehat{f}(k)e^{ikx}$$

是绝对收敛的。特别地，对于 $\alpha\in(\dfrac{1}{2},1]$，函数序列 $\{S_N(f)\}_{N\geqslant 1}$ 一致收敛到 f。

注记. 我们要强调，$\{S_N(f)\}_{N\geqslant 1}$ 不是在 $C^{0,\alpha}$ 的范数下收敛到 f。

证明*:* 受到上面评注里证明的启发，我们考虑

$$g_h(x)=f(x+h)-f(x-h).$$

从而，我们可以计算

$$\widehat{g_h}(k)=(e^{ikh}-e^{-ikh})\widehat{f}(k)=2i\sin(kh)\widehat{f}(k).$$

很明显，对任意的 $h\in\mathbb{R}$，我们有 $g_h(x)\in L^2(\mathbf{T})$。根据勾股定理，我们有

$$\int_{-\pi}^{\pi}|g_h(x)|^2\frac{dx}{2\pi}=4\sum_{k\in\mathbb{Z}}|\sin(kh)|^2|\widehat{f}(k)|^2.$$

另外，根据 $f\in C^\alpha$，我们还知道

$$|g_h(x)|\leqslant\|f\|_{C^{0,\alpha}}(2h)^\alpha.$$

所以，我们得到

$$\sum_{k\in\mathbb{Z}}|\sin(kh)|^2|\widehat{f}(k)|^2\leqslant Ch^{2\alpha}$$

特别地，如果令 $h=\dfrac{\pi}{2^{p+1}}$，我们得到

$$\sum_{k\neq 0}|\sin(\frac{k\pi}{2^{p+1}})|^2|\widehat{f}(k)|^2\leqslant\frac{C}{2^{2p\alpha}},$$

其中，常数 C 可能有所改变，但是这还是一个不依赖于 f 的常数。在这个求和中，我们只选取一部分的和：

$$\sum_{2^{p-1}\leqslant|k|<2^p}|\sin\left(\frac{k\pi}{2^{p+1}}\right)|^2|\widehat{f}(k)|^2\leqslant\frac{C}{2^{2p\alpha}}.$$

此时，由于 $\frac{k\pi}{2^{p+1}} \in [\frac{\pi}{4}, \frac{\pi}{2}]$，从而 $|\sin\left(\frac{k\pi}{2^{p+1}}\right)|^2 \geqslant 0.5$。从而，上面的不等式给出

$$0.5 \times \sum_{2^{p-1}\leqslant|k|<2^p} |\widehat{f}(k)|^2 \leqslant \frac{C}{2^{2p\alpha}}.$$

所以，

$$\sum_{2^{p-1}\leqslant|k|<2^p} |\widehat{f}(k)|^2 \leqslant \frac{C}{2^{2p\alpha}}.$$

当然，这里的 C 也改变了。此时，我们利用 Cauchy-Schwarz 不等式（因为我们想得到 $|\widehat{f}(k)|$ 的和）可以得到

$$\sum_{2^{p-1}\leqslant|k|<2^p} |\widehat{f}(k)| \leqslant 2^{\frac{p-1}{2}}\Big(\sum_{2^{p-1}\leqslant|k|<2^p} |\widehat{f}(k)|^2\Big)^{\frac{1}{2}} \leqslant \frac{C}{2^{p(\alpha-\frac{1}{2})}}.$$

所以，当 $\alpha > \frac{1}{2}$ 的时候，我们有

$$\sum_{k\in\mathbb{Z}} |\widehat{f}(k)| \leqslant \sum_{p=1}^{\infty}\Big(\sum_{2^{p-1}\leqslant|k|<2^p} |\widehat{f}(k)|\Big) \leqslant \sum_{p=1}^{\infty} \frac{C}{2^{p(\alpha-\frac{1}{2})}} < \infty.$$

这表明，$\sum_{k\in\mathbb{Z}} \widehat{f}(k)e^{ikx}$ 绝对收敛。 □

Fourier 级数的应用：等分布问题

考虑 $[0,1)$ 区间上的数列 $\{\xi_k\}_{k\geqslant 1}$，我们要用数学的语言来描述这些数是“平均地”分布在 $[0,1)$ 区间上。

定义 380. 如果对任意的 $a, b \in [0,1)$，我们有

$$\lim_{n\to\infty} \frac{\big|\{k \leqslant n \big| \xi_k \in [a,b]\}\big|}{n} = b - a,$$

即该序列中有百分之 $100(b-a)$ 那么多数落在 $[a,b)$ 中，那么我们就说 $\{\xi_k\}_{k\geqslant 1}$ 在 $[0,1)$ 上是**等分布的**。

注记. 如果 $\{\xi_k\}_{k\geqslant 1}$ 在 $[0,1)$ 上是等分布的，那么 $\{\xi_k\}_{k\geqslant 1}$ 在 $[0,1)$ 上是稠密的：否则，存在开区间 $(a,b) \subset [0,1)$，使得 $\{\xi_k\}_{k\geqslant 1} \cap (a,b) = \emptyset$，此时，对任意的 $n \in \mathbb{Z}_{\geqslant 1}$，我们有

$$\frac{\big|\{k \leqslant n \big| \xi_k \in [a,b]\}\big|}{n} = 0,$$

这与定义不符。

反之，如果 $\{\xi_k\}_{k\geqslant 1}$ 在 $[0,1)$ 上是稠密的，$\{\xi_k\}_{k\geqslant 1}$ 在 $[0,1)$ 上是不一定是等分布的，我们把反例的构造留作本次作业。

例子. 如下的两个例子不是等分布的：

1) 我们用 $\{x\}$ 表示实数 x 的小数部分，即 $\{x\}=x-\lfloor x\rfloor$。对任意的有理数 q，数列 $\big\{\{k\cdot q\}\big\}_{k\geqslant 1}$ 不是等分布的，这是因为这个数列实际上只有有限项；

2) 对任意的 $k\geqslant 1$，我们定义

$$\xi_k=\left(\frac{1+\sqrt{5}}{2}\right)^k.$$

那么，数列 $\{\xi_k\}_{k\geqslant 1}$ 不是等分布的：如果令

$$F_{k+1}=\left(\frac{1+\sqrt{5}}{2}\right)^k+\left(\frac{1-\sqrt{5}}{2}\right)^k$$

我们就得到了 *Fibonacci* 数列。由于 $\frac{1-\sqrt{5}}{2}\approx -0.618$，所以，$\xi_k\to 1$，从而不是稠密的。

给定无理数 $\alpha\in\mathbb{R}-\mathbb{Q}$，数列 $\big\{\{k\cdot\alpha\}\big\}_{k\geqslant 1}$ 是等分布的，这是等分布理论中最基本的例子，我们仔细研究这个例子并发展一般的理论。令 $\xi_k=k\cdot\alpha$，其中 $k\geqslant 1$。为了判定 $\{\xi_k\}_{k\geqslant 1}$ 在 $[0,1)$ 上是否为等分布的，我们有如下平凡但是重要的观察：等分布性等价于

$$\lim_{n\to\infty}\frac{1}{n}\sum_{k=1}^{n}\mathbf{1}_{[a,b]}(\xi_k)=\int_0^1\mathbf{1}_{[a,b]}(x)dx.$$

这是从集合到它的示性函数的过渡。这使得我们联想到我们在积分理论中学到的东西，从集合的测度过渡到简单简单函数再过渡到一般的可积函数。据此，我们猜想可能下面的结论也成立：

$$\lim_{n\to\infty}\frac{1}{n}\left(\sum_{k=1}^{n}f(\xi_k)\right)=\int_0^1 f(x)dx,\qquad\qquad(\star).$$

其中 f 是连续函数或者 Riemann 可积的函数（请想一下对于 Lebesgue 意义下的 L^1 函数会有什么问题?）。这是这类问题的大思路：把算术问题转化为分析问题，从而可以尝试微积分和函数论中的工具。

根据上面的分析，我们尝试对 $f\in C[0,1)$ 来证明上面的极限 $(\star)$。由于 $(\star)$ 左右两边对 f 都是线性的，根据 Fourier 分析的基本想法，如果 f 是 $\mathbb{R}$ 上以 1 为周期的函数（和我们之前的 Fourier 分析差了一个常数），可以先考虑用有限的三角级数来逼近 f 进行证明。作为出发点，可以先尝试用最基本的频率函数 $e_\ell(x)=e^{2\pi i\ell x}$（周期为 1）来验证命题（之后用线性组合以及逼近来证明一般的情况）。

当 $\ell=0$ 时，$(\star)$ 结论是显然的。

如果 $\ell\neq 0$，我们注意到 $e_\ell(\{k\alpha\})=e_\ell(k\alpha)$（这是因为 e_ℓ 以 1 为周期）。利用 $\xi_k=k\cdot\alpha$，我们可以直接计算：

$$\frac{1}{n}\sum_{k=1}^{n}e_\ell(\xi_k)=\frac{1}{n}\frac{1-e^{2\pi in\ell\alpha}}{1-e^{2\pi i\ell\alpha}}e^{2\pi i\ell\alpha}.$$

由于 $\alpha \notin \mathbb{Q}$，所以上式中的分母不是 0。从而，

$$\Big|\frac{1}{n}\sum_{k=1}^{n} e_\ell(\xi_k)\Big| \leqslant \frac{2}{n}\frac{1}{|1-e^{2\pi i\ell\alpha}|},$$

它的极限显然是 0。我们还要指出，这个证明唯一用到了 $\{\xi_k\}_{k\geqslant 1}$ 算术性质的地方是 $\alpha \notin \mathbb{Q}$（三角级数对此类问题之所以有效就是因为它和这些算术性质有关联）。

利用 $(\star)$ 的线性，对任何有限的三角级数 Q，结论都成立。对于任意给定的 $\mathbb{R}$ 上的周期连续函数 f，根据 Weierstrass-Stone 定理，对任意的 $\varepsilon > 0$，存在有限的三角级数 $Q = \sum_{|\ell|\leqslant K} c_k e^{2\pi ikx}$，使得

$$\|Q - f\|_{L^\infty} < \frac{1}{2}\varepsilon.$$

此时，根据上面的计算，选择 N，使得当 $n > N$ 时，对每个 $\ell \in \mathbb{Z}\cap[-K,K]$，我们都有

$$|c_k| \times \Big|\frac{1}{n}\sum_{k=1}^{n} e_\ell(\xi_k)\Big| < \frac{1}{2}\varepsilon.$$

所以，

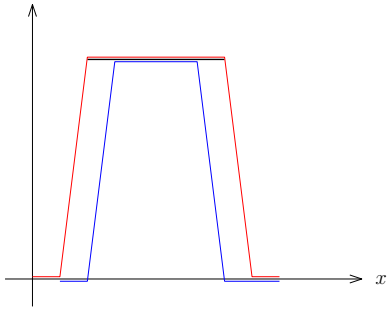
$$\Big|\frac{1}{n}\sum_{k=1}^{n} Q(\xi_k)\Big| < \frac{1}{2}\varepsilon.$$

从而，当 $n > N$ 时，

$$\Big|\frac{1}{n}\sum_{k=1}^{n} f(\xi_k)\Big| < \Big|\frac{1}{n}\sum_{k=1}^{n} Q(\xi_k)\Big| + \frac{1}{n}\sum_{k=1}^{n}|Q(\xi_k)-f(\xi_k)| < \varepsilon,$$

从而 $(\star)$ 对于周期连续函数成立。

我们之前要研究的等分布问题是对函数 $f = \mathbf{1}_{[a,b]}$ 来陈述的，其中 $a \in [0,1)$，$b \in (0,1)$，$a < b$。这个函数并连续函数，但是我们可以逼近它：对任意的 $\delta > 0$（很小），我们定义



$$f_+(x) = \begin{cases} 1, & a \leqslant x \leqslant b; \\ 0, & x \leqslant a-\delta \text{ 或 } x \geqslant b+\delta; \\ \frac{1}{\delta}(x-b)+1, & b \leqslant x \leqslant b+\delta; \\ -\frac{1}{\delta}(x-a)+1, & a-\delta \leqslant x \leqslant a. \end{cases}$$

和

$$f_-(x)=\begin{cases}1, & a+\delta\leqslant x\leqslant b-\delta;\\ 0, & x\leqslant a \text{ 或 } x\geqslant b;\\ \frac{1}{\delta}(x-b-\delta)+1, & b-\delta\leqslant x\leqslant b;\\ -\frac{1}{\delta}(x-a+\delta)+1, & a\leqslant x\leqslant a+\delta.\end{cases}$$

很明显，我们有

$$f_-\leqslant \mathbf{1}_{[a,b]}\leqslant f_+,$$

所以

$$\frac{1}{n}\sum_{k=1}^{n}f_-(\xi_k)\leqslant\frac{1}{n}\sum_{k=1}^{n}\mathbf{1}_{[a,b]}(\xi_k)\leqslant\frac{1}{n}\sum_{k=1}^{n}f_+(\xi_k).$$

由于

$$0\leqslant\int_0^1 f_+(x)dx-\int_0^1 f_-(x)dx\leqslant 2\delta.$$

所以，当 $n\to\infty$ 时，我们有

$$\lim_{n\to\infty}\frac{1}{n}\left(\sum_{k=1}^{n}\mathbf{1}(\xi_k)\right)-\int_0^1\mathbf{1}(x)dx=O(\delta).$$

令 $\delta\to 0$，命题成立。

我们可以进一步对 Riemann 可积的函数 $f\in\mathcal{R}$（根据线性，只要对实值函数证明）证明

$$\frac{1}{n}\sum_{k=1}^{n}f(\xi_k)\longrightarrow\int_0^1 f(x)dx,\quad f\in\mathcal{R}([0,1])$$

为此，选取 $[0,1]$ 的分划 $\sigma:0=x_0<x_1<\cdots<x_N=1$，我们进一步要求分划点 x_i 都是有理数（仍然可以保证 $|\sigma|\to 0$）。考虑如下两个阶梯函数

$$f_+(x)=\sum_{k=0}^{N-1}\left(\sup_{y\in[x_k,x_{k+1}]}f(y)\right)\mathbf{1}_{[x_k,x_{k+1}]}(x),$$
$$f_-(x)=\sum_{k=0}^{N-1}\left(\inf_{y\in[x_k,x_{k+1}]}f(y)\right)\mathbf{1}_{[x_k,x_{k+1}]}(x).$$

按照定义，我们有

$$\frac{1}{n}\sum_{k=1}^{n}f_-(\xi_k)\leqslant\frac{1}{n}\sum_{k=1}^{n}f(\xi_k)\leqslant\frac{1}{n}\sum_{k=1}^{n}f_+(\xi_k).$$

由于结论对于 $\mathbf{1}_{[a,b]}$ 型的函数成立，所以对于 $f_\pm$ 也成立。所以，对任意的 $\varepsilon>0$，先选较小的步长 $|\sigma|$，使得

$$\left|\int_0^1 f_\pm-\int_0^1 f\right|<\frac{1}{2}\varepsilon,$$

然后选择足够大的 N，使得 $n \geqslant N$ 时，

$$\Big|\frac{1}{n}\left(\sum_{k=1}^{n} f_{\pm}(\xi_k)\right) - \int_0^1 f_{\pm}\Big| < \frac{1}{2}\varepsilon.$$

从而，

$$\Big|\frac{1}{n}\left(\sum_{k=1}^{n} f(\xi_k)\right) - \int_0^1 f\Big| < \varepsilon.$$

这就证明了结论。

事实上，上面证明的后半部分与 $\{\xi_k\}_{k\geqslant 1}$ 的具体选择没有关系，这是一个更一般的结论：

定理 381 (Weyl 的等分布判别准则)**.** *$\{\xi_k\}_{k\geqslant 1}$ 是 $[0,1)$ 区间上的数列。那么，$\{\xi_k\}_{k\geqslant 1}$ 在 $[0,1)$ 上等分布当且仅当对任意的 $\ell \neq 0$，我们有*

$$\lim_{n\to\infty}\left(\frac{1}{n}\sum_{k=1}^{n} e^{2\pi i\ell\xi_k}\right) = 0.$$

这个定理的进一步的推广就是动力系统理论中的 Birkhoff 遍历性定理。

有了 Weyl 判别准则，我们可以相对轻松地证明一些数列在 $[0,1)$ 区间上是等分布的：

例子. *我们有其他几个等分布或者非等分布的例子：*

1) *任给非零实数 a，任意的实数 $\sigma \in (0,1)$，我们令 $\xi_k = \{ak^\sigma\}$（小数部分）。那么，$\{\xi_k\}_{k\geqslant 1}$ 在 $[0,1)$ 上等分布。*

根据 Weyl 判别准则，我们需要控制 $\sum_{k=1}^{n} e^{2\pi i\ell ak^\sigma}$ 的大小。如果令 $b = 2\pi\ell a$，我们要证明

$$\sum_{k=1}^{n} e^{ibk^\sigma} = o(n).$$

我们用积分来逼近求和：

$$\sum_{k=1}^{n} e^{ibk^\sigma} - \int_0^n e^{ibx^\sigma}dx = \sum_{k=1}^{n-1}\int_k^{k+1}\left(e^{ibk^\sigma} - e^{ibx^\sigma}\right)dx.$$

根据 Lagrange 中值定理（对函数 e^{ibx^σ} 的实部和虚部分别来做），对 $x \in [k, k+1]$，存在依赖于 b 和 σ 的常数 C，使得

$$\left|e^{ibk^\sigma} - e^{ibx^\sigma}\right| \leqslant Ck^{-1+\sigma}.$$

从而，

$$\begin{aligned}\Big|\sum_{k=1}^{n-1} e^{ibk^\sigma} - \int_1^{n-1} e^{ibx^\sigma}dx\Big| &\leqslant C\sum_{k=1}^{n}\int_k^{k+1} k^{-1+\sigma}dx \\ &= O(n^\sigma)(= o(n)).\end{aligned}$$

下面我们估计积分项 $\int_1^n e^{ibx^\sigma}dx$（对于 n 和 $n-1$ 的差别可以忽略，因为这是 $O(1)$-项）：

$$\begin{aligned}\int_1^n e^{ibx^\sigma}dx &= \frac{1}{i\sigma b}\int_1^{n-1}\left(e^{ibx^\sigma}\right)' x^{1-\sigma}dx \\ &= \underbrace{e^{ibx^\sigma}x^{1-\sigma}\Big|_{x=1}^{x=n}}_{O(n^{1-\sigma})} - \frac{1-\sigma}{i\sigma b}\int_1^n e^{ibx^\sigma}x^{-\sigma}dx.\end{aligned}$$

通过将指数函数用 1 来控制，我们有

$$\Big|\int_1^n e^{ibx^\sigma}x^{-\sigma}dx\Big| \leqslant \int_1^n x^{-\sigma}dx = O(n^{1-\delta}).$$

最终，我们证明了

$$\sum_{k=1}^n e^{2\pi i\ell a k^\sigma} = O(n^\sigma) + O(n^{1-\sigma}).$$

根据 *Weyl* 判据，我们就说明了 $\{\xi_k\}_{k\geqslant 1}$ 在 $[0,1)$ 上等分布。

2) 作为一个例子的推论（$a=1$，$\sigma=0.5$），数列 $\left\{\sqrt{k}\right\}_{k\geqslant 1}$ 的小数部分在 $[0,1)$ 上等分布。

3) 任意给定 $a\in\mathbb{R}$，对任意的 $k\geqslant 1$，我们定义

$$\xi_k = a\log(k).$$

那么，$\{\xi_k\}_{k\geqslant 1}$ 在 $[0,1)$ 上不是等分布的。

根据 *Weyl* 判别准则，我们估计 $\sum_{k=1}^n e^{2\pi i\ell a\log(k)}$。我们令 $b=2\pi\ell a$，现在用积分逼近求和

$$\sum_{k=1}^n e^{ib\log(k)} - \int_0^n e^{ib\log(x)}dx = \sum_{k=0}^{n-1}\int_k^{k+1}\left(e^{i\log(k)} - e^{ib\log(x)}\right)dx.$$

根据 *Lagrange* 中值定理，对 $x\in[k,k+1]$，存在依赖于 b 和 σ 的常数 C，使得

$$\left|e^{ib\log(k)} - e^{ib\log(x)}\right| \leqslant \frac{C}{k}.$$

所以，

$$\Big|\sum_{k=1}^n e^{ib\log(k)} - \int_0^n e^{ib\log(x)}dx\Big| \leqslant \sum_{k=0}^{n-1}\int_k^{k+1}\frac{C}{k}dx = O(\log(n)).$$

下面我们来估计积分 $\int_1^n e^{ib\log x}dx$，而这个积分可以通过分部积分直接计算：

$$\int_1^n e^{ib\log x}dx = xe^{ib\log x}\Big|_{x=1}^{x=n} - ib\int_1^n e^{ib\log x}dx.$$

从而，我们可以算出

$$\int_1^n e^{ib\log x}dx = \frac{ne^{ib\log(n)}}{1+ib} + O(1).$$

所以，

$$\frac{1}{n}\int_1^n e^{ib\log x}dx + o(1) = \frac{e^{ib\log(n)}}{1+ib}.$$

这个复数的模长是固定的，所以 $\int_1^n e^{ib\log x}dx \neq o(1)$。

57 Roth 三项等差数列定理

二零二零年五月二十八日，六月一日，晴

Roth 的三项等差数列定理：Fourier 级数的另一个应用

我们给出 Fourier 级数的基本思想的另一个应用。经典的 Roth 三项等差数列定理讲的是：

定理 382 (Roth)**.** *任给自然数的集合 $A \subset \mathbb{Z}_{\geqslant 0}$，$A$ 在自然数中的*上密度*定义为*

$$\overline{\rho}(A) = \limsup_{n\to\infty} \frac{|\{x \in A | x \leqslant A\}|}{n}.$$

如果 $\overline{\rho}(A) > 0$，那么 A 中一定包含 3 项的等差数列，即存在 $x, y, z \in A$，使得 $y - x = z - y$。

假设 $x < z < y$ 是这样的一个等差数列，那么，$x + y = 2z$；反之，在整数集中，$x + y = 2z$ 意味着 $\{x, z, y\}$ 构成等差数列。

我们将证明这个定理的一个量化形式（Roth 的定理是下面命题的推论），即

命题 383. *存在常数 $C > 0$（$C = 99$ 就可以），使得对任意给定 $\delta > 0$，只要*

$$N > e^{e^{C\delta^{-1}}},$$

对任意的 $A \subset \{0, 1, \cdots, N-1\}$，如果 $|A| \geqslant \delta N$，那么 A 中必然有 3 项的等差数列。

在后面的证明中，我们假设 N 是奇数（纯技术性假设）。

离散集合 $\{0, 1, 2, \cdots, N-1\}$ 上的 Fourier 级数

给定自然数 N，我们令 $\mathbf{T}_N = \{0, 1, 2, \cdots, N-1\}$，这是 $\mathbb{Z}$ 在除 N 的意义下的一个完全剩余类（余数的集合）。我们现在考虑 $\mathbb{Z}$ 上以 N 为周期的函数 f（复值函数），即 $f(x+N) = f(x)$，其中 $x \in \mathbb{Z}$。这个函数可以被看作是 $\mathbf{T}_N$ 上的函数。

很明显，$\mathbf{T}_N$ 上复值函数的全体是有限维复线性空间，我们把它记做 $L^2(\mathbf{T}_N)$。另外，$\dim_{\mathbb{C}} L^2(\mathbf{T}_N) = N$。在 $L^2(\mathbf{T}_N)$ 我们可以定义如下自然的内积：

$$(f, g) = \frac{1}{N} \sum_{x=0}^{N-1} f(x)\overline{g(x)}.$$

与 Fourier 级数类似，我们有一组标准正交基：$\left\{e^{i\frac{2\pi i}{N}kx} \middle| k = 0, 1, \cdots, N-1\right\}$（请自己验证这一点）。

对任意的 $f \in L^2(\mathbf{T}_N)$，我们定义 f 的 Fourier 变换：

$$\widehat{f} : \mathbf{T}_N \to \mathbb{C}, \quad \widehat{f}(k) = \sum_{x=0}^{N-1} f(x) e^{-\frac{2\pi i}{N}kx}.$$

对任意的 $f, g \in L^2(\mathbf{T}_N)$，我们定义它们的卷积：

$$f * g(x) = \sum_{y=0}^{N-1} f(x-y)g(y).$$

我们总结关于 Fourier 变换和内积的性质：

1) 我们有 Fourier 逆变换公式（用 Fourier 系数来写出函数本身）：

$$f(x) = \frac{1}{N} \sum_{k=0}^{N-1} \widehat{f}(k) e^{\frac{2\pi i}{N} kx}.$$

实际上，上面的表达式可以写成

$$f = \sum_{k=0}^{n-1} \left(f, e^{i\frac{2\pi i}{N} kx} \right) e^{i\frac{2\pi i}{N} kx}.$$

这就是在内积空间中把向量用标准正交基来表示。

2) 勾股定理仍然成立：

$$\frac{1}{N} \sum_{k=0}^{N-1} |\widehat{f}(k)|^2 = \sum_{x=0}^{N-1} |f(x)|^2.$$

根据 1) 中的表达式，我们应该有

$$(f, f) = \sum_{k=0}^{N-1} \left| \frac{\widehat{f}(k)}{N} \right|^2.$$

整理即得。

3) 在差一个常数的意义下，Fourier 变换保持内积，即

$$(\widehat{f}, \widehat{g}) = N(f, g) \quad \Leftrightarrow \quad \frac{1}{N} \sum_{k=0}^{N-1} \widehat{f}(k)\overline{\widehat{g}(k)} = \sum_{x=0}^{N-1} f(x)\overline{g(x)}.$$

根据 1) 和 $e^{\frac{2\pi i}{N} kx}$ 之间的相互正交性，我们有

$$\begin{aligned}
(f, g) &= \left(\sum_{k=0}^{N-1} \frac{\widehat{f}(k)}{N} e^{\frac{2\pi i}{N} kx}, \sum_{\ell=0}^{N-1} \frac{\widehat{g}(\ell)}{N} e^{\frac{2\pi i}{N} \ell x} \right) \\
&= \sum_{k=0}^{N-1} \sum_{\ell=0}^{N-1} \frac{\widehat{f}(k)}{N} \frac{\widehat{g}(\ell)}{N} \underbrace{\left(e^{\frac{2\pi i}{N} kx}, e^{\frac{2\pi i}{N} \ell x} \right)}_{\ell=k} \\
&= \sum_{k=0}^{N-1} \frac{\widehat{f}(k)}{N} \frac{\widehat{k}(\ell)}{N} \\
&= \frac{1}{N} (\widehat{f}, \widehat{g}).
\end{aligned}$$

4) 在频率空间来看，卷积就是正常的乘法：

$$\widehat{f * g} = \widehat{f} \cdot \widehat{g}.$$

我们利用 Fubini 公式（交换求和顺序）：

$$\begin{aligned}
\widehat{f * g}(k) &= \sum_{x=0}^{N-1} \left(\sum_{y=0}^{N-1} f(x-y) g(y) e^{-\frac{2\pi i}{N} kx} \right) \\
&= \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} f(x-y) e^{-\frac{2\pi i}{N} k(x-y)} g(y) e^{-\frac{2\pi i}{N} ky} \\
&= \sum_{y=0}^{N-1} \left(\sum_{x=0}^{N-1} f(x-y) e^{-\frac{2\pi i}{N} k(x-y)} \right) g(y) e^{-\frac{2\pi i}{N} ky} \\
&= \sum_{y=0}^{N-1} \widehat{f}(k) g(y) e^{-\frac{2\pi i}{N} ky} \\
&= \widehat{f}(k) \cdot \widehat{g}(k).
\end{aligned}$$

5) 卷积之后的 L^2 范数：

$$\frac{1}{N} \sum_{k=0}^{N-1} |\widehat{f}(k)|^2 |\widehat{g}(k)|^2 = \sum_{x=0}^{N-1} |f * g(x)|^2.$$

这是前面几个结论的直接推论。

Roth 定理的 Fourier 化

现在给定 $A \subset \mathbf{T}_N$，$|A| = \delta N$，我们用 $\mathfrak{N}$ 表示 A 中的 3 项的等差数列的数目（总假定公差不是 0）。我们现在 mod N 的意义下考虑三项等差数列的方程

$$x + y \equiv 2z \bmod N \quad \Leftrightarrow \quad z - x \equiv y - z \bmod N.$$

我们想要在 $A^3 \subset \mathbf{T}_N^3$ 中来找到 (x, y, z)，使得上面的方程成立。为此，我们定义

$$\mathfrak{N}_1 = \big|\{(x, y, z) \in A^3 \big| x + y \equiv 2z \bmod N\}\big|.$$

我们把注意力集中在 $\mathfrak{N}_1$ 上，暂且不考虑它与 $\mathfrak{N}$ 之间的关联。根据频率函数之间的正交性，我们有

$$\mathfrak{N}_1 = \sum_{x \in A} \sum_{y \in A} \sum_{z \in A} \frac{1}{N} \sum_{k=0}^{N-1} e^{-\frac{2\pi i}{N} k(x+y-2z)}$$

这因为如果当 $x+y \neq 2z \bmod N$ 时，对 k 求和给出零。所以，

$$\begin{aligned}\mathfrak{N}_1 &= \sum_{x\in\mathbf{T}_N}\sum_{y\in\mathbf{T}_N}\sum_{z\in\mathbf{T}_N}\frac{1}{N}\sum_{k=0}^{N-1} e^{-\frac{2\pi i}{N}k(x+y-2z)}\mathbf{1}_A(x)\mathbf{1}_A(y)\mathbf{1}_A(z)\\ &= \frac{1}{N}\sum_{k=0}^{N-1}\Big(\big(\sum_{x\in\mathbf{T}_N} e^{-\frac{2\pi i}{N}kx}\mathbf{1}_A(x)\big)\big(\sum_{y\in\mathbf{T}_N} e^{-\frac{2\pi i}{N}ky}\mathbf{1}_A(y)\big)\big(\sum_{z\in\mathbf{T}_N} e^{\frac{2\pi i}{N}2kz}\mathbf{1}_A(z)\big)\Big)\\ &= \frac{1}{N}\sum_{k=0}^{N-1}\widehat{\mathbf{1}_A}(k)^2\widehat{\mathbf{1}_A}(-2k),\end{aligned}$$

其中 $\mathbf{1}_A$ 是集合 $A\subset\mathbf{T}_N$ 的示性函数。利用 $\widehat{\mathbf{1}_A}(0)=\delta N$，最终，我们得到如下的等式：

$$\mathfrak{N}_1 = \delta^3N^2 + \frac{1}{N}\sum_{k=1}^{N-1}\widehat{\mathbf{1}_A}(k)^2\widehat{\mathbf{1}_A}(-2k).$$

注记. 我们有如下两个直观（并不直观的直观）：

1) 上面计数问题的首项有可以利用随机性给出如下的解释：假设 A 在 N 中是随机选取的，也就是说我们随机地从 $0,1,\cdots,N-1$ 中选取数，以概率为 δ 地机会选到了 A。此时，x 和 z 可以随便从 A 中挑选，所以一共有 δ^2N^2 种选择，此时 $y=2z-x$ 还要落在 A 中，这个概率就只有 δ，所有一共有 δ^3N^2 中选择。根据这种看法，如果 A 在 $\mathbf{T}_N$ 中是非常均匀分布的，那么我们期盼有很多三项 mod N-等差数列（大约有 δ^2N^3 个）。

2) 假设 $A\subset\mathbf{T}_N$ 是随机分布的（不管怎么定义），当 N 很大的时候，如果把 A 看成是 $\mathbf{T}_N$ 中的数列，我们直观上可以认为它们是**等分布**的，根据 *Weyl* 的等分布判据，$\mathbf{1}_A$ 这个函数的 *Fourier* 系数（任意一个非零系数）是 $o(|A|)$ 的，$|A|$ 算是这个序列的项数。

3) 假设 $\mathbf{1}_A$ 的非零 *Fourier* 系数很小，即存在 $\varepsilon>0$ 很小，使得对任意的 $k\neq 0$，我们都有 $|\widehat{\mathbf{1}_A}(k)|\leqslant\varepsilon N$。

此时，我们有

$$\begin{aligned}\Big|\sum_{k=1}^{N-1}\widehat{\mathbf{1}_A}(k)^2\widehat{\mathbf{1}_A}(-2k)\Big| &\leqslant \big(\max_{k\neq0}|\widehat{\mathbf{1}_A}(-2k)|\big)\sum_{k=1}^{N-1}|\widehat{\mathbf{1}_A}(k)|^2\\ &\leqslant \varepsilon N\sum_{k=0}^{N-1}|\widehat{\mathbf{1}_A}(k)|^2\\ &= \varepsilon\delta N^2.\end{aligned}$$

所以，如果 ε 足够的小，我们就得到

$$\mathfrak{N}_1 \geqslant \delta^3N^2 - \varepsilon\delta N^2 = \delta N^2(\delta^2-\varepsilon) > 0.$$

那么，$\mathfrak{N}_1$ 是大于 0 的（实际上有 δ^3N^2 那么多，如果 $\varepsilon\approx 0.5\times\delta^2$）。

根据上面的讨论，我们引入如下概念：

定义 384. 给定 $\mathbf{T}_N$ 的子集 A。如果存在 $\varepsilon>0$，使得对任意的 $k\neq 0$，都有

$$|\widehat{\mathbf{1}_A}(k)|\leqslant \varepsilon N,$$

我们就称 A 是 ε-**随机的**或者是 ε-**等分布**的。

我们现在给出 Roth 定理证明的梗概。有两种情形需要处理：

1) A 是随机分布的。

我们先证明有很多三项的 mod N-等差数列，然后在证明等差数列也存在。

2) A 的分布不随机。

在这个条件下，我们证明能在 $\mathbf{T}_N$ 中（不是 A 中）找到很长的一个等差数列 P_1，使得 $A_1=A\cap P_1$ 在 P_1 中的密度要相对于 A 在 $\mathbf{T}_N$ 中的密度大一些，即存在 $\varepsilon'>0$ 只依赖于 δ，使得

$$|A_1|\geqslant \big(\delta+\varepsilon'\big)|P|,$$

我们把 $\mathbf{T}_N$ 替换为 P，把 A 替换为 A_1。此时，子集密度增长了 ε'。

我们不停地进行 2) 这个操作：如果 1) 在某一步成立，我们就完成了 Roth 定理的证明；否则，经过 m 次之后，最终导致某个 A_m（$m\gg 1$）在 P_m 的密度 >1，矛盾。

A 是相对随机分布的情形：$|\widehat{\mathbf{1}_A}(k)|\leqslant\varepsilon N$，其中 $\varepsilon<\dfrac{1}{8}\delta^2$

在这种情况下，根据前面的计算，我们有

$$\mathfrak{N}_1\geqslant \delta^3N^2-\varepsilon\delta N^2\geqslant \frac{7}{8}\delta^3N^2.$$

其中，$\mathfrak{N}_1$ 是三项的 mod N-等差数列的个数，我们想得到真正的三项等差数列的个数的下界估计。

我们有如下的观察：

令 $A_{\frac{1}{3}}=A\cap[\frac{N}{3},\frac{2N}{3})$，如果 $x,z\in A_{\frac{1}{3}}$，那么满足 $x+y\equiv 2z \bmod N$ 的 $y\in A$，一定满足 $x+y=2z$。

所以，我们令

$$\begin{aligned}\mathfrak{N}_0&=\Big|\big\{(x,y,z)\in A^3\big|x+y=2z\big\}\Big|,\\ \mathfrak{N}_2&=\Big|\big\{(x,y,z)\in A_{\frac{1}{3}}\times A\times A_{\frac{1}{3}}\big|x+y\equiv 2z \bmod N\big\}\Big|.\end{aligned}$$

那么

$$\mathfrak{N}_0\geqslant\mathfrak{N}_2.$$

利用三角函数之间的相互正交性，我们仿照 $\mathfrak{N}_1$ 的计算，就有

$$\begin{aligned}\mathfrak{N}_2 &= \sum_{x\in A_{\frac13}}\sum_{y\in A}\sum_{z\in A_{\frac13}}\frac{1}{N}\sum_{k=0}^{N-1}e^{-\frac{2\pi i}{N}k(x+y-2z)}\\ &= \frac{1}{N}\sum_{k=0}^{N-1}\widehat{\mathbf{1}_{A_{\frac13}}}(k)\widehat{\mathbf{1}_A}(k)\widehat{\mathbf{1}_{A_{\frac13}}}(-2k)\\ &= \delta\big|A_{\frac13}\big|^2+\frac{1}{N}\sum_{k\neq0}^{N-1}\widehat{\mathbf{1}_{A_{\frac13}}}(k)\widehat{\mathbf{1}_A}(k)\widehat{\mathbf{1}_{A_{\frac13}}}(-2k)\end{aligned}$$

我们对后一项进行下面的控制（N 是奇数，所以 $k\mapsto -2k$ 是 $\mathbf{T}_N$ 到自身的双射（$\mathrm{mod}N$））：

$$\begin{aligned}&\Big|\sum_{k\neq0}^{N-1}\widehat{\mathbf{1}_{A_{\frac13}}}(k)\widehat{\mathbf{1}_A}(k)\widehat{\mathbf{1}_{A_{\frac13}}}(-2k)\Big|\\ &\leqslant \max_{k\neq0}|\widehat{\mathbf{1}_A}(k)|\times\sum_{k\neq0}^{N-1}\Big|\widehat{\mathbf{1}_{A_{\frac13}}}(k)\widehat{\mathbf{1}_{A_{\frac13}}}(-2k)\Big|\\ &\leqslant \varepsilon N\Big(\sum_{k\neq0}^{N-1}\big|\widehat{\mathbf{1}_{A_{\frac13}}}(k)\big|^2\Big)^{\frac12}\Big(\sum_{k\neq0}^{N-1}\big|\widehat{\mathbf{1}_{A_{\frac13}}}(-2k)\big|^2\Big)^{\frac12}\\ &= \varepsilon N\sum_{k\neq0}^{N-1}\big|\widehat{\mathbf{1}_{A_{\frac13}}}(k)\big|^2\\ &\leqslant \varepsilon N\sum_{k=0}^{N-1}\big|\widehat{\mathbf{1}_{A_{\frac13}}}(k)\big|^2\\ &= \varepsilon N^2|A_{\frac13}|.\end{aligned}$$

从而，

$$\mathfrak{N}_2\geqslant\delta\big|A_{\frac13}\big|^2-\varepsilon N|A_{\frac13}|\geqslant\delta\big|A_{\frac13}\big|^2-\frac{\delta^2}{8}N|A_{\frac13}|.$$

最终，我们得到

$$\mathfrak{N}_0\geqslant\delta\big|A_{\frac13}\big|\Big(|A_{\frac13}|-\frac{\delta}{8}N\Big).$$

我们根据 $A_{\frac13}$ 的元素数目的多少，作以下的考虑：

- 假设 $|A_{\frac13}|\geqslant\frac14\delta N$。
  此时，我们有
  $$\mathfrak{N}_0\geqslant\frac{1}{32}\delta^3N^2.$$
  我们注意到 $\mathfrak{N}_0$ 中可能还计算了公差是零的三项等差数列（即 $x=y=z$），这样的等差数列个数至多有 $|A|=\delta N$ 个，所以，
  $$\mathfrak{N}\geqslant\frac{1}{32}\delta^3N^2-\delta N=\delta N\big(\frac{1}{32}\delta^2N-1\big).$$

所以，为了保证有三项等差数列，我们总是假设 $N > 33\delta^{-2}$，此时，

$$\mathfrak{N} \geqslant \frac{1}{32}\delta N > 0.$$

- 假设 $|A_{\frac{1}{3}}| < \frac{1}{4}\delta N$。

  根据定义，$A_{\frac{1}{3}} = A \cap [\frac{N}{3}, \frac{2N}{3})$。我们考虑两个等差数列 $P_1 = [0, \frac{N}{3})$ 和 $P_2 = [\frac{2N}{3}, N)$（中的整数），从而

  $$|A \cap P_1| + |A \cap P_2| \geqslant \frac{3}{4}\delta N.$$

  所以至少有一个数列，不妨设是 $P = P_1$，使得 $|A \cap P| \geqslant \frac{3}{8}\delta N$。$P$ 中的元素个数不超过 $\frac{N}{3}$ 个，从而，$A \cap P$ 在 P 中的占的比例至少是

  $$\frac{3}{8}\delta N \div \frac{1}{3}N = \delta + \frac{1}{8}\delta.$$

作为总结，我们证明了

引理 385. *假设 N 是奇数，$A \subset \{0, 1, \cdots, N-1\}$，$|A| \geqslant \delta N$ 并且 $N > 33\delta^{-2}$。如果 A 中不含三项等差数列，那么如下两种情形必居其一：*

(a) 对任意的 $\varepsilon \leqslant \frac{\delta^2}{8}$，$A$ 在 $\{0, 1, \cdots, N-1\}$ 中不是 ε-随机的；

(b) 存在一个长度至少是 $\frac{1}{3}N$ 的等差数列 P，使得

$$|A \cap P| \geqslant \delta(1 + \frac{1}{8})|P|.$$

即 A 在 P 中密度比以前增长了至少 12.5%。

A 不是相对随机分布的情形：对任意的 $\varepsilon < \frac{1}{8}\delta^2$，存在 $k = k(\varepsilon) \neq 0$，使得 $|\widehat{\mathbf{1}_A}(k)| > \varepsilon N$

先证明一个引理：

引理 386. *对于任意的 $f, g \in L^2(\mathbf{T}_N)$，我们有*

$$|\widehat{f}(k)||\widehat{g}(k)| \leqslant \sum_{x \in \mathbf{T}_N} \Big| \sum_{y \in \mathbf{T}_N} f(y)\overline{g(y-x)} \Big|.$$

证明： 令 $F(x) = \sum_{y \in \mathbf{T}_N} f(y)\overline{g(y-x)}$，我们计算它的 Fourier 系数（和卷积的 Fourier 系数的计算

一样)：

$$\begin{aligned}\widehat{F}(k) &= \sum_{x\in\mathbf{T}_N}\sum_{y\in\mathbf{T}_N} f(y)\overline{g(y-x)}e^{-\frac{2\pi i}{N}kx}\\ &= \sum_{x\in\mathbf{T}_N}\sum_{y\in\mathbf{T}_N} f(y)\overline{g(y-x)e^{\frac{2\pi i}{N}k(x-y)}}e^{-\frac{2\pi i}{N}ky}\\ &= \sum_{y\in\mathbf{T}_N} f(y)e^{-\frac{2\pi i}{N}ky}\\ &= \Big(\sum_{x\in\mathbf{T}_N}\overline{g(y-x)e^{\frac{2\pi i}{N}k(x-y)}}\Big)\\ &= \widehat{f}(k)\overline{\widehat{g}(k)}.\end{aligned}$$

根据 $\|\widehat{F}\|_{L^\infty}\leqslant\|F\|_{L^1}$（因为 $\big|\widehat{f}(k)\big|\leqslant\sum_{x=0}^{N-1}|F(x)|.$），我们就得到了结论。 □

注记. *这个不等式表明（利用抽屉原理），如果 f 和 g 在某个共同的频率上都是大的（$\geqslant cN$），那么 f 和 g 的某个平移的内积就会很大（$\geqslant c^2$）。*

我们的目标是说明在此情形下（即上面引理 A 中的 (a) 情形），上一步引理中的 (b) 也成立（密度的增长可能比 12.5% 要少），即我们要找一个 L 项等差数列 P_0（公差为 d），使得 $|A\cap P_0|\geqslant(\delta+\varepsilon')|P_0|$，其中 L 要相对比较大。当 ε 给定的情况下，我们将会取 $\varepsilon'=\frac{1}{4}\varepsilon$，所以，根据 $\varepsilon<\dfrac{1}{8}\delta^2$，我们知道 δ 将增长 $O(\delta^2)$。

我们分两步来完成这个情况的分析：

第一) 对任意的非零频率 $k\neq0$，我们都能找到 $\mathbf{T}_N$ 中的准等差数列（mod$N$ 的意义下）：

$$B=\{\xi_i\}_{i=1,2,\cdots,m+m'},$$

其中

$$0\leqslant\xi_1<\xi_2<\cdots\xi_m\leqslant N-1,\ 0\leqslant\xi_{m+1}<\xi_{m+2}<\cdots<\xi_{m+m'}\leqslant N-1,$$

并且除了 $i=m$ 之外，$\xi_{i+1}=\xi_i+d$；$\xi_{m+1}=\xi_m+d-N$，$d(m+m')\leqslant N-1$，使得该准等差数列的项数

$$m+m'\geqslant\frac{1}{4}\sqrt{N}$$

而且

$$|\widehat{\mathbf{1}_B}(k)|\geqslant\frac{1}{2}|B|.$$

———

现在证明上面的论述：对于固定的 k，我们知道存在 $\ell_1,\ell_2\in\mathbf{T}_N$，使得下面两个不等式同时成立：

$$0<\ell_2-\ell_1\leqslant\sqrt{N},\ \ k(\ell_2-\ell_1)\bmod N\ \leqslant\sqrt{N}.$$

(考虑平面上 N 个点 $\{(s, sk \bmod N)|s = 0, 1, \cdots, N-1\}$，它们都落在 $[0, N-1] \times [0, N-1]$ 中，我们将 $[0, N-1] \times [0, N-1]$ 平均分成 $\lfloor\sqrt{N-1}\rfloor^2$ 份，总共小于 N 份，那么必有两个点落在同一个小格子中，它们就满足要求）令 $d = \ell_2 - \ell_1$，我们现在定一个 $\mathbf{T}_N$ 中的准等差数列（在考虑 $\bmod N$ 的意义下）：

$$\begin{aligned} B &= \{\cdots, -2d, -d, 0, d, 2d, \cdots\} \\ &= \{\xi_1 = N - (m-1)d, \xi_2 = N - (m-2)d, \cdots, \xi_m = N\} \\ &\quad \cup \{\xi_{m+1} = d, \xi_{m+2} = 2d, \cdots, \xi_{m+m'} = m'd\}, \end{aligned}$$

其中 $m = m' = \frac{1}{8}\lfloor\frac{N}{d}\rfloor$，从而上面两个集合不相交（看长度），这表明

$$|B| \geqslant \frac{1}{4}\lfloor\frac{N}{d}\rfloor = \frac{1}{4}\sqrt{N}.$$

我们现在将 B 选取的短一点，使得 $|B| = \lfloor\frac{1}{4}\sqrt{N}\rfloor + 1$。

现在计算

$$\begin{aligned} \big|\widehat{\mathbf{1}_B}(k) - |B|\big| &\leqslant \Big|\sum_{x\in B} \mathbf{1}_B(x)(e^{-\frac{2\pi i}{N}kx} - 1)\Big| \leqslant \sum_{x\in B} \big|e^{-\frac{2\pi i}{N}kx} - 1\big| \\ &= \sum_{1\leqslant|j|\leqslant m=\frac{|B|}{2}} \big|e^{-\frac{2\pi i}{N}jkd} - 1\big| \end{aligned}$$

我们下面运用不等式 $|e^{ix} - 1| \leqslant |x|$（看弧长并利用两点之间线段最短，上学期已经证明）。由于（请参考 d 的构造）

$$0 \leqslant dk \bmod N \leqslant \sqrt{N},$$

所以存在整数 s，使得 $0 \leqslant dk - sN \leqslant \sqrt{N}$，从而 $e^{-\frac{2\pi i}{N}jkd} = e^{-\frac{2\pi i}{N}j(kd-sN)}$，从而

$$\begin{aligned} \big|\widehat{\mathbf{1}_B}(k) - |B|\big| &\leqslant \sum_{|j|\leqslant\frac{|B|}{2}} \frac{2\pi|j|(kd - sN)}{N} \\ &\leqslant \frac{2\pi}{\sqrt{N}} \sum_{|j|\leqslant\frac{|B|}{2}} |j| \\ &\leqslant \frac{2\pi}{\sqrt{N}} \frac{1}{2}\big(\frac{|B|}{2}\big)^2 = \frac{1}{2}|B| \times \frac{\pi}{4}\frac{|B|}{\sqrt{N}}. \end{aligned}$$

根据 B 的构造，后一个因子 $\frac{\pi}{4}\frac{|B|}{\sqrt{N}} < 1$，从而

$$\big|\widehat{\mathbf{1}_B}(k) - |B|\big| \leqslant \frac{1}{2}|B|.$$

这就给出了第一步的证明。

第二) 对于上一步中构造出来的 B，我们能找到它的一个平移 $B+x$（仍然落在 $\mathbf{T}_N$ 中，$N-1$ 被平移成了 $x-1$），使得

$$|A\cap(B+x)|\geqslant(\delta+\frac{1}{4}\varepsilon)|B|.$$

———

我们首先将上面的这个叙述翻译成分析的语言，要证明的不等式等价于存在 $x\in\mathbf{T}_N$，使得

$$\sum_{y\in\mathbf{T}_N}\mathbf{1}_A(y)\mathbf{1}_B(y-x)\geqslant(\delta+\frac{1}{4}\varepsilon)|B|.$$

如果令 $f_A(x)=\mathbf{1}_A(x)-\delta$，那么，

$$|A\cap(B+x)|\geqslant(\delta+\frac{1}{4}\varepsilon)|B|\quad\Leftrightarrow\quad\sum_{y\in\mathbf{T}_N}f_A(y)\mathbf{1}_B(y-x)\geqslant\frac{1}{4}\varepsilon|B|.$$

另外，f_A 除了零频率以外，和 $\mathbf{1}_A$ 是一样的。特别地，对某个给定的 $k=k(\varepsilon)\neq0$，我们有

$$|\widehat{f_A}(k)|>\varepsilon N.$$

我们现在令 $H(x)=\sum_{y\in\mathbf{T}_N}f_A(y)\mathbf{1}_B(y-x)$，根据已经证明的引理，我们得到（在频率 k 上考虑）：

$$\sum_{x\in\mathbf{T}_N}|H(x)|\geqslant|\widehat{f_A}(k)||\widehat{\mathbf{1}_B}(k)|\geqslant\frac{\varepsilon N}{2}|B|.$$

利用 $\sum_{x\in\mathbf{T}_N}H(x)=0$（因为 f_A 没有频率为零的分量），我们得到

$$\sum_{x\in\mathbf{T}_N}|H(x)|+H(x)\geqslant\frac{\varepsilon N}{2}|B|.$$

从而存在某个 x，使得

$$|H(x)|+H(x)\geqslant\frac{\varepsilon}{2}|B|.$$

所以，

$$H(x)\geqslant\frac{\varepsilon}{4}|B|.$$

我们注意到，经过平移之后，$B+x$ 仍然形如

$$B=\{\xi_i\}_{i=1,2,\cdots,m+m'},$$

其中，

$$0\leqslant\xi_1<\xi_2<\cdots\xi_m\leqslant N-1,\ 0\leqslant\xi_{m+1}<\xi_{m+2}<\cdots<\xi_{m+m'}\leqslant N-1.$$

综合上面两个步骤的结论，我们就得到了 $\mathbf{T}_N$ 中的准等差数列

$$B = \{\xi_i\}_{i=1,2,\cdots,m+m'},$$

其中

$$0 \leqslant \xi_1 < \xi_2 < \cdots \xi_m \leqslant N-1,\ 0 \leqslant \xi_{m+1} < \xi_{m+2} < \cdots < \xi_{m+m'} \leqslant N-1.$$

并且，除了 $i = m$ 之外，$\xi_{i+1} = \xi_i + d$；$\xi_{m+1} = \xi_m + d - N$，$d(m+m') \leqslant N-1$ 并且该准等差数列的项数 $m + m' \geqslant \frac{1}{4}\sqrt{N}$，使得

$$|A \cap B| \geqslant (\delta + \frac{1}{4}\varepsilon)|B|.$$

然而，我们需要找到等差数列而不是准等差数列，我们再修正一下。B 可以分成两个真正的等差数列：

$$P_1 = \{\xi_1 < \xi_2 < \cdots \xi_m\},\ \ P_2 = \{\xi_{m+1} < \xi_{m+2} < \cdots < \xi_{m+m'}\}.$$

我们再分两种情况来讨论：

- P_1 和 P_2 之中有某个序列特别短，即长度不超过 $\frac{1}{8}\varepsilon|B|$。

  不妨假设 $|P_1| \leqslant \frac{1}{8}\varepsilon|B|$，此时

  $$|P_2 \cap A| \geqslant |B \cap A| - |P_1 \cap A| \geqslant (\delta + \frac{1}{8}\varepsilon)|B| \geqslant (\delta + \frac{1}{8}\varepsilon)|P_2|.$$

  此时，A 在 P_2 中占了相当大的比重并且 P_2 也足够长：

  $$|P_2| \geqslant \Big(\frac{1}{4} - \frac{\varepsilon}{8}\Big)\sqrt{N} \approx \frac{1}{4}\sqrt{N}.$$

- P_1 和 P_2 都比较长，即 $\min\{|P_1|, |P_2|\} \geqslant \frac{1}{8}\varepsilon|B|$。

  此时，因为 A 在 $P_1 \cup P_2$ 中的密度是至少是 $(\delta + \frac{1}{8}\varepsilon)|B|$，所以，$A$ 至少在某一个 P_i 中的密度是 $(\delta + \frac{1}{8}\varepsilon)|B|$。

作为总结，我们证明了

引理 387. *如果有某个非零的频率 $k \neq 0$，使得 $|\widehat{\mathbf{1}_A}(k)| > \varepsilon N$，那么一定存在一个长度至少是 $\frac{\varepsilon}{32}\sqrt{N}$ 等差数列 P，使得*

$$|A \cap P| \geqslant (\delta + \frac{1}{8})|P|.$$

综合上面两种情况的结论（取 $\varepsilon = \frac{1}{8}\delta^2$），我们证明了：

命题 388. *假设 $A \subset \{0, 1, \cdots, N-1\}$，$|A| \geqslant \delta N$ 并且 $N > 33\delta^{-2}$。那么，如下两种情形必居其一：*

(a) A 中含三项等差数列；

(b) 存在一个长度至少是 $\frac{\delta^2}{256}\sqrt{N}$ 的等差数列 P，使得 $|A \cap P| \geqslant (\delta + \frac{\delta^2}{64})|P|$。

Roth 定理证明的完成

我们要对之前得到结论进行迭代。我们不停地利用上面的性质并假设第一种情况不能发生。第一次迭代，我们得到一个等差数列 P_1，$N_1 = |P_1| \geqslant \dfrac{\delta^2}{256}\sqrt{N}$，$|A \cap P_1| = \delta_1 |P_1|$，其中新的密度 $\delta_1 \geqslant \delta + \frac{1}{64}\delta^2$。只要新得到的等差数列的长度不是零，我们就可以迭代。所以，连续做 $k = \dfrac{62}{\delta}$ 次，密度就从 δ 变成了 $\delta_k \geqslant 2\delta$。

我们再做 $k = \dfrac{62}{2\delta}$ 次，密度就从 2δ 变成了 4δ；再做 $k = \dfrac{62}{4\delta}$ 次，密度就从 4δ 变成了 8δ；以此类推，通过不超过

$$\frac{64}{\delta}\Big(1 + \frac{1}{2} + \cdots + \frac{1}{2^{\ell-1}}\Big) \longrightarrow \frac{128}{\delta}$$

次，密度就变成了 $2^\ell \delta$。

当然，我们不可能做 $K = \dfrac{128}{\delta}$ 次操作，否则密度变成了无穷大（超过 1）就不行。

第一次操作后，等差数列的长度 N 变成了 $N_1 \geqslant \dfrac{\delta^2}{256}\sqrt{N}$；第二次操作，我们得到的等差序列的长度是

$$N_2 \geqslant \frac{\delta_1^2}{256}\sqrt{N_1} \geqslant \frac{\delta^2}{256}\sqrt{N_1}$$

我们这里用到了每一次操作，密度 δ_k 是增长的。所以，做了 K 次操作之后，等差数列的长度至少是 $\dfrac{1}{256^2}\delta^4 N^{\frac{1}{2^K}}$。

我们希望做到第 K 次时，$\dfrac{1}{256^2}\delta^4 N^{\frac{1}{2^K}} \geqslant 1$，这就给出了矛盾。这等价于

$$\log N \geqslant 2^{\frac{128}{\delta}}\big(16\log 2 - 4\log\delta\big).$$

从而，$N \geqslant e^{e^{C\delta^{-1}}}$（可以具体算出 C 的大小）。这就完成了定理的证明。

Fourier 级数的另一个应用：Poisson 核

我们首先叙述如下关于调和函数的经典问题，这个问题在数学和物理中都有举足轻重的地位（然而，我们没有太多的时间深入了解）：

假定 f 是 $\mathbf{T} = \mathbf{S}^1$ 上的（复数或者实数值）连续函数，能否在单位圆盘 $D = \{z \in \mathbb{C} \big| |z| \leqslant 1\}$ 上面找到连续函数 u，使得 u 在 D 的内部 $\mathring{D}$ 是 C^2 的并且

$$\begin{cases} \triangle u(x) &= 0, \quad x \in \mathring{D}, \\ u\big|_{\partial D} &= f, \end{cases}$$

其中 $\triangle u = \dfrac{\partial^2 u}{\partial x^2} + \dfrac{\partial^2 u}{\partial y^2}$。

我们的基本想法来自于 Fourier 分析和线性代数。Fourier 级数的理论表明，f 可以用一组基 $e^{2\pi i k x}$ 的线性组合来表示，我们先假设 $f = e^{2\pi k x}$，如果能对这样的情形解决问题，也就说对任意的 $k \in \mathbb{Z}$，能找到

$$\begin{cases} \triangle u_k &= 0, \quad x \in \mathring{D}, \\ u_k\big|_{\partial D} &= e^{2\pi i k x}. \end{cases}$$

那么，由于 $f(x)=\sum_{k\in\mathbb{Z}}\widehat{f}(k)e^{2\pi ikx}$（当然，更准确的说法需要对 f 的正则性有限制或者在 L^2 中谈论这个等式，但是我们直观上可以这样来考虑问题），我们有理由相信

$$f(x)=\sum_{k\in\mathbb{Z}}\widehat{f}(k)u_k(x).$$

我们先假设 $k\in\mathbb{Z}_{\geqslant 0}$，此时，从复变量函数的观点，我们有

$$z^k\big|_{r=1}=e^{2\pi ikx}.$$

（参考第十一次作业中的 s 与 S 组习题，特别是 s10)）首先，我们有

$$\triangle z^k=0.$$

这个直接计算可得（从复变函数的观点这个很自然）。实际上，对于上述计算在 $z\neq 0$ 处，但是，z^k 在 0 处是不良好定义。为此，当 $k<0$ 时，我们考虑复共轭 $\overline{z^{|k|}}$，此时，我们仍然有

$$\overline{z^{|k|}}\Big|_{r=1}=e^{2\pi ikx},\quad k<0.$$

综合上述，我们有

$$\begin{cases}\triangle u_k &=0,\\ u_k\big|_{\partial D} &=e^{2\pi ikx},\end{cases}\Rightarrow\quad u_k=\begin{cases}z^k, & k\geqslant 0;\\ \overline{z^{|k|}}, & k<0.\end{cases}.$$

所以，我们期望有

$$\begin{aligned}u(z)&=-\widehat{f}(0)+\sum_{k\geqslant 0}\widehat{f}(k)z^k+\sum_{k\geqslant 0}\widehat{f}(-k)\bar{z}^k\\&=-\widehat{f}(0)+\sum_{k\geqslant 0}\int_0^{2\pi}f(t)\big(e^{-it}z\big)^k+f(x)\big(e^{it}\bar{z}\big)^k\frac{dt}{2\pi}\\&\overset{|z|<1}{=}-\widehat{f}(0)+\int_0^{2\pi}f(t)\Big(\frac{1}{1-e^{-it}z}+\frac{1}{1-e^{it}\bar{z}}\Big)\frac{dt}{2\pi}\\&=\int_0^{2\pi}f(t)\Big(\frac{1}{1-e^{-it}z}+\frac{1}{1-e^{it}\bar{z}}-1\Big)\frac{dt}{2\pi}\end{aligned}$$

如果令 $z=re^{i\theta}$，那么

$$u(z)=\int_0^{2\pi}f(t)\frac{1-r^2}{1+r^2-2r\cos(\theta-t)}\frac{dt}{2\pi}.$$

据此，我们引入 **Poisson 核**：

$$P_r(\theta)=\frac{1-r^2}{1+r^2-2r\cos(\theta)},\ r\in[0,1)$$

那么，

$$u(z)=P_r*f.$$

根据构造方式，我们知道 $P_r(\theta) > 0$ 并且

$$P_r(\theta) = \sum_{k\in\mathbb{Z}} r^{|k|} e^{ik\theta}.$$

利用一致收敛性与 Lebesgue 控制收敛，我们知道

$$\int_0^{2\pi} P_r(\theta)\frac{d\theta}{2\pi} = 1.$$

我们声明，实际上 $P_r(\theta)$ 满足下面的性质：

$$\int_{\delta\leqslant|\theta|\leqslant\pi} |P_r(\theta)|\frac{dx}{2\pi} \to 0, \quad r \to 1.$$

按照定义，我们知道

$$\begin{aligned} P_r(\theta) &= \frac{1-r^2}{1+r^2-2r\cos(\theta)} = \frac{1-r^2}{(1-r)^2+2r(1-\cos\theta)} \\ &\leqslant \frac{1-r^2}{2r(1-\cos\theta)} \leqslant \frac{1-r^2}{2r(1-\cos\delta)} \to 0. \end{aligned}$$

这说明，$P_r(\theta)$ 是一个好的积分核，所以有一致收敛：

$$\lim_{r\to 1} P_r * f = f.$$

这表明，$P_r * f\big|_{\mathbf{S}^1} = f$。

下面证明当 $r < 1$ 的时候，如果将 $P_r(\theta)$ 视作是 (x, y)（或者 (r, θ)）的函数，我们有

$$\triangle P_r(\theta) = 0.$$

我们知道

$$\triangle = \frac{\partial^2}{\partial r^2} + \frac{1}{r}\frac{\partial}{\partial r} + \frac{1}{r^2}\frac{\partial^2}{\partial\theta^2}.$$

然而，要直接验证 $\triangle P_r(\theta) = 0$ 的计算并不是很简单。我们可以采取有两种方式来验证，一种是用 $P_r(\theta)$ 的定义：

$$P_r(\theta) = \sum_{k\in\mathbb{Z}} r^{|k|} e^{ik\theta}.$$

每个分量 z^k 或 $\overline{z}^k$ 均为调和函数。为了证明 $\triangle$ 与求和可交换（从而可以放到和式里面），我们利用 Lebesgue 控制收敛的推论即可；第二种方法是观察到，如果 $\triangle f = 0$，那么将 f 旋转之后仍然成立（参考第十二次习题的 s），从而，我们只要对对 $\theta = 0$ 验证即可，此时直接计算会简单很多。

最终，为了证明在 $|z| < 1$ 处有 $\triangle P_r * f = 0$，再用 Lebesgue 控制收敛的推论。

综上所述，我们找到了

$$\begin{cases} \triangle u(x) &= 0, \quad x \in \mathring{D}, \\ u\big|_{\partial D} &= f, \end{cases}$$

的（唯一解）$u(r, \theta) = P_r * f$。

57.1 作业：Fourier 级数几乎处处发散的 L^1-函数

清华大学 19-20 春季学期，数学分析二，作业 12

请用 A4 大小的纸张**正反面**用钢笔，签字笔或者圆珠笔书写，并注明自己的姓名，年级和作业的**总页数**。除定理公式所涉及的人名之外，请使用**中文**。本次作业的提交时间和地点为 **2 月 120 日**上午的课堂上，逾期视作零分。

Gibbs 现象，Fourier 级数

A1) 试构造数列 $\{\xi_k\}_{k\geqslant 1}\subset[0,1)$，使得它在 $[0,1)$ 上稠密却不是等分布的。

A2) 对任意的 $f,g\in L^1(\mathbf{T})$，证明，卷积在频率空间来看就是正常的乘积：

$$\widehat{f*g}(k)=\widehat{f}(k)\widehat{g}(k).$$

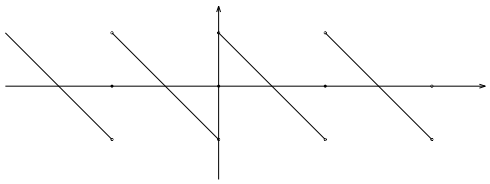
A3) f 是 $\mathbb{R}$ 上以 2π 为周期的实值函数，它在 $[-\pi,\pi]$ 上是有界的单调函数。证明，存在常数 C，使得对任意的 $k\in\mathbb{Z}$，我们都有

$$\left|\widehat{f}(k)\right|\leqslant\frac{C}{1+|k|}.$$

（先证明 $[-\pi,\pi]$ 上有界的单调函数可以被形如 $\sum_{k=1}^{N}\alpha_k\mathbf{1}_{[a_{k-1},a_k]}$ 的函数逼近，其中 $-\pi=a_0<a_1<\cdots<a_N=\pi$，$\alpha_1\leqslant\alpha_2\leqslant\cdots$）

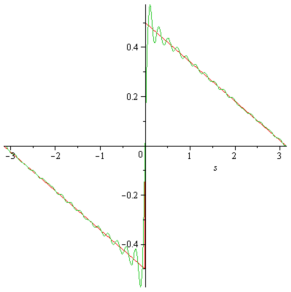
A4)（Gibbs 现象）我们定义墙头草形状的以 2π 为周期的函数 f，其中，当 $x\in[0,2\pi)$ 时，我们有

$$f(x)=\begin{cases}0, & x=0;\\ \frac{\pi-x}{2}, & \end{cases}$$



此时，在间断点 0 处，我们有 $f_+(0)-f_-(0)=\pi$。证明，f 的 Fourier 级数可以写为

$$\sum_{k\neq 0}\frac{1}{2ik}e^{ik\cdot x}=\sum_{k=1}^{\infty}\frac{\sin(kx)}{k}.$$



证明，对于部分和 $S_N(f)$，我们有

$$\max_{x\in(0,\frac{\pi}{N}]} S_N(f)(x) - \frac{\pi}{2} = \int_0^\pi \frac{\sin(y)}{y} dy - \frac{\pi}{2} > 0.$$

这是分段 C^1 函数在第一类间断点处 Fourier 系数收敛的典型行为，Fourier 级数在间断点处起伏的最大值接比间断本身（$f_+(0) - f_-(0) = \pi$）要大 8.9%。

A5) 假设 $f \in C(\mathbf{T})$，证明，级数

$$\sum_{k\in\mathbb{Z}} \frac{\widehat{f}(k)}{k}$$

是绝对收敛的。

A6) 我们定义级数

$$S(x) = \sum_{n=1}^{\infty} \frac{\sin^3(nx)}{n!}.$$

证明，$S(x)$ 是 $\mathbb{R}$ 上以 2π 为周期的光滑函数。利用

$$A(x) = \sum_{n=1}^{\infty} \frac{\sin(nx)}{n!}, \quad B(x) = \sum_{n=1}^{\infty} \frac{\cos(nx)}{n!},$$

证明，

$$S(x) = \frac{3}{4}\sin\left(\sin(x)\right) e^{\cos(x)} - \frac{1}{4}\sin\left(\sin(3x)\right) e^{\cos(3x)}.$$

一个含参数积分的研究

B1) 证明，对任意的 $t \in \mathbb{R}$，如下积分是良好定义的：

$$F(t) = \int_0^\infty \frac{1 + tx^5 + t^2x^{11}}{1 + t^4 + x^{13}} dx.$$

B2) 证明，$F(t)$ 是 $t \in \mathbb{R}$ 的连续函数。

B3) 证明，$F(t)$ 是可导的。

B4) 试计算 $\lim\limits_{t\to+\infty} F(t)$。

B5) 证明，对 $t \neq 0$，如下的积分是发散的：

$$\int_0^\infty \frac{1+tx^5+t^2x^{11}}{1+t^4+x^{12}}dx$$

B6) 证明，对任意的 $t \in \mathbb{R}$，如下在 Riemann 意义下的反常积分是良好定义的：

$$G(t) = \int_0^\infty \frac{1+tx^5+t^2x^{11}}{1+t^4+x^{12}}e^{ix}dx,$$

即 $\lim\limits_{T\to\infty}\int_0^T \frac{1+tx^5+t^2x^{11}}{1+t^4+x^{12}}e^{ix}dx$ 存在。

B7) 证明，$G(t)$ 是 $t \in \mathbb{R}$ 的连续函数。

B8) 试计算 $\lim\limits_{t\to+\infty} G(t)$。

Laplace 变换的基本性质

C1) 假设 $f(t)$ 是 $\mathbb{R}_{>0}$ 上定义的 Lebesgue 可积函数。我们定义它的 Laplace 变换为：

$$\big(\mathcal{L}f\big)(x) = \int_0^\infty f(t)e^{-xt}dt,$$

其中 $x > 0$。证明，$\big(\mathcal{L}f\big)(x)$ 是 C^∞ 的函数。

C2) 证明，对任意的 $n \in \mathbb{Z}_{\geqslant 0}$，我们都有

$$\left(\frac{d}{dx}\right)^n \big(\mathcal{L}f\big)(x) = \int_0^\infty (-t)^n f(t)e^{-xt}dt.$$

C3) 假设 $f(t)$ 和 $g(t)$ 是 $\mathbb{R}_{>0}$ 上定义的 Lebesgue 可积函数，我们定义它们的“卷积”为

$$\big(f \bullet g\big)(t) = \int_0^t f(\tau)g(t-\tau)d\tau.$$

证明，$f \bullet g = g \bullet f$。

C4) 假设 $f(t)$ 和 $g(t)$ 是 $\mathbb{R}_{>0}$ 上定义的 Lebesgue 可积函数。证明，

$$\mathcal{L}\big(f \bullet g\big)(x) = \big(\mathcal{L}f\big)(x)\big(\mathcal{L}g\big)(x).$$

Fourier 级数几乎处处发散的 L^1-函数

注记. *1922* 年，*Kolmogorov*（*19* 岁，还在莫斯科大学读本科）在论文 Une série de Fourier-Lebesgue divergente presque partout 中构造一个 L^1 函数，这个函数的 *Fourier* 级数的部分和几乎处处是发散的。这个问题的目的就是重现这个著名反例的构造。

先固定一个自然数 $n \geqslant 10000$（在接下来的过程中它会变化）。我们任意选取 n 个正的奇数 $\lambda_1 < \lambda_2 < \cdots < \lambda_n$（我们将在接下来的过程中确定它们的大小）。定义如下的数组，函数和区间：

- $\{m_k\}_{1\leqslant k\leqslant n}$。我们令 $m_1=n$，并且归纳地定义 $2m_k+1=\lambda_k(2n+1)$。
- $\{A_k\}_{1\leqslant k\leqslant n}$，其中 $A_k=\dfrac{4k\pi}{2n+1}$。
- $\phi_n(x)=\displaystyle\sum_{k=1}^{n}\frac{m_k^2}{n}\mathbf{1}_{\Delta_k}(x)$，其中区间 $\Delta_k=[A_k-m_k^{-2},A_k+m_k^{-2}]$。
- 对于 $k=2,3,\cdots,n$，我们定义区间 $\sigma_k=[A_{k-1}+2n^{-2},A_k-2n^{-2}]$。

根据 Dirichlet 核函数的基本性质，我们知道 Fourier 级数的部分和可以由下面的公式计算：

$$S_{m_k}\phi_n(x)=\sum_{|j|\leqslant m_k}\widehat{\varphi_n}(j)e^{\sqrt{-1}jx}=\int_0^{2\pi}\phi_n(t)\frac{\sin\big((m_k+\frac12)(x-t)\big)}{\sin\big(\frac12(x-t)\big)}\frac{dt}{2\pi}.$$

仿照这个形式，我们定义

$$I_{m_k,\Delta_s}(x)=\int_{\Delta_s}\frac{m_s^2}{n}\frac{\sin\big((m_k+\frac12)(x-t)\big)}{\sin\big(\frac12(x-t)\big)}\frac{dt}{2\pi}$$

K1) 证明，

$$\int_0^{2\pi}\phi_n(x)=2\quad\text{并且}\quad S_{m_k}\phi_n(x)=\sum_{s=1}^{n}I_{m_k,\Delta_s}(x).$$

K2**) 证明，$\sup\limits_{\substack{N\in\mathbb{Z}_{\geqslant0}\\x\in[0,2\pi]}}\big|\big(S_N\phi_n\big)(x)\big|<\infty$。（提示：可以参考关于有界变差函数的 Fourier 级数的 Jordan 定理的证明过程）

K3*) 证明，我们可以归纳地选取（足够大的）正奇数 $\lambda_1<\lambda_2<\cdots<\lambda_n$，使得区间 $\{\sigma_\ell,\Delta_j\}_{\ell\leqslant n-1,j\leqslant n}$ 两两不相交并且对任意的 $k\geqslant2$ 和任意的 $x\in\sigma_k$，我们都有

$$\Big|\sum_{s=1}^{k-1}I_{m_k,\Delta_s}(x)\Big|\leqslant1.$$

（提示：如果 $\lambda_1,\cdots,\lambda_{k-1}$ 已经选择好了，那么我们通过 λ_k 的选择就可以使得 m_k 足够大）

K4) 证明，Dirichlet 核函数 $D_{m_k}(x)=\begin{cases}\dfrac{\sin\big((m_k+\frac12)x\big)}{\sin\big(\frac{x}{2}\big)},&x\neq0;\\2m_k+1,&x=0\end{cases}$ 满足

$$\big|D'_{m_k}(x)\big|\leqslant2m_k^2.$$

K5) 证明，当 $s\geqslant k$ 的时候，对任意的 $x\in\sigma_k$，我们可以将 $I_{m_k,\Delta_s}(x)$ 写成

$$I_{m_k,\Delta_s}(x)=\frac{1}{n\pi}\frac{\sin\big((m_k+\frac12)(x-A_s)\big)}{\sin\big(\frac12(x-A_s)\big)}+\overbrace{\int_{\Delta_s}\frac{m_s^2}{n}\left(\frac{\sin\big((m_k+\frac12)(x-t)\big)}{\sin\big(\frac12(x-t)\big)}-\frac{\sin\big((m_k+\frac12)(x-A_s)\big)}{\sin\big(\frac12(x-A_s)\big)}\right)\frac{dt}{2\pi}}^{\mathbf{I}_{k,s}},$$

并且我们有如下的控制：$|\mathbf{I}_{k,s}|\leqslant\dfrac{4}{n}$。

K6) 证明，对任意的 $x \in \sigma_k$，我们都有

$$\Big|\sum_{s=k}^{n} I_{m_k,\Delta_s}(x)\Big| \geqslant \frac{\big|\sin\big((m_k+\frac{1}{2})(x-A_k)\big)\big|}{n\pi}\Big|\sum_{s=k}^{n}\frac{1}{\sin\big(\frac{1}{2}(x-A_s)\big)}\Big| - 4.$$

K7) 证明，对任意的 $x \in \sigma_k$，我们有

$$\frac{1}{n}\sum_{s=k}^{n}\frac{1}{\sin\big(\frac{1}{2}(x-A_s)\big)} \geqslant \frac{1}{\pi}\sum_{\ell=1}^{n-k}\frac{1}{\ell}.$$

K8) 证明，如果 $2 \leqslant k \leqslant n-\sqrt{n}$，对任意的 $x \in \sigma_k$，我们有

$$\Big|\sum_{s=k}^{n} I_{m_k,\Delta_s}(x)\Big| \geqslant \frac{\big|\sin\big((m_k+\frac{1}{2})(x-A_k)\big)\big|}{2\pi^2}\log(n) - 8.$$

K9) 证明，如果 m_2 选取的足够大，那么如下集合的 Lebesgue 测度满足

$$m\Big(\big\{x \in [A_{k-1},A_k]\Big|\big|\sin\big((m_k+\frac{1}{2})(x-A_k)\big)\big| < \frac{2\pi^2}{\sqrt{\log(n)}}\big\}\Big) = O(\frac{1}{n\sqrt{\log(n)}}).$$

K10) 我们定义

$$E_n = \bigcup_{2\leqslant k\leqslant n-\sqrt{n}} E_{n,k},\ \text{其中}\ E_{n,k} = \Big\{x\in\sigma_k\Big|\big|\sin\big((m_k+\frac{1}{2})(x-A_k)\big)\big| \geqslant \frac{2\pi^2}{\sqrt{\log(n)}}\Big\}$$

证明，$\lim\limits_{n\to\infty} m\big(E_n\big) = 2\pi$。

K11) 证明，对任意的 $2 \leqslant k \leqslant n-\sqrt{n}$ 和 $x \in E_{n,k}$，我们都有

$$|S_{m_k}\phi_n(x)| \geqslant \sqrt{\log(n)} - 9.$$

K12) 证明，存在以 2π 为周期的函数列 $\{\phi_n(x)\}_{n\geqslant 10000}$，使得

a) 对任意的 n，我们有 $\phi_n(x) \geqslant 0$，$\int_0^{2\pi}\phi_n = 2$ 并且 $\sup\limits_{\substack{N\in\mathbb{Z}_{\geqslant 0},\\ x\in[0,2\pi]}} \big|\big(S_N\phi_n\big)(x)\big| < \infty$；

b) 存在子区间序列 $\{E_n\}_{n\geqslant 10000}$，正整数序列 $\{q_n\}_{n\geqslant 10000}$ 以及单调上升的正数序列 $\{M_n\}_{n\geqslant 10000}$，使得 $\lim\limits_{n\to\infty} M_n = \infty$，$\lim\limits_{n\to\infty} m\big(E_n\big) = 2\pi$ 并且对任意的 $x \in E_n$，存在正整数 $N_x \leqslant q_n$，使得 $|S_{N_x}\phi_n(x)| \geqslant M_n$。

（提示：选取 $q_n = m_n$ 即可）

K13) 证明，存在以 2π 为周期的函数列 $\{\varphi_n(x)\}_{n\geqslant 1}$，子区间序列 $\{E_n\}_{n\geqslant 1}$，正整数序列 $\{q_n\}_{n\geqslant 1}$ 以及单调上升的正数序列 $\{M_n\}_{n\geqslant 1}$，使得

a) $M_1 \geqslant 2$ 并且对任意的 n，我们有 $M_n \geqslant 2M_{n-1}$；

b) 对任意的 n，我们有 $\varphi_n(x)\geqslant 0$，$\displaystyle\int_0^{2\pi}\varphi_n=2$ 并且 $\displaystyle\sum_{1\leqslant k\leqslant n-1}\sup_{\substack{N\in\mathbb{Z}_{\geqslant 0}\\ x\in[0,2\pi]}}\big|\big(S_N\varphi_k\big)(x)\big|<\frac{\sqrt{M_n}}{2}$；

c) $\displaystyle\sup_{1\leqslant k\leqslant n-1}q_k\leqslant\frac{\sqrt{M_n}}{2^n}$；

d) $\displaystyle\lim_{n\to\infty}m\big(E_n\big)=2\pi$ 并且对任意的 $x\in E_n$，存在正整数 $N_x\leqslant q_n$，使得 $|S_{N_x}\varphi_n(x)|\geqslant M_n$。

（提示：选取上面的题目中的函数的子序列）

K14) 利用上一个小题的结论，我们构造

$$f(x)=\sum_{n=1}^{\infty}\frac{\varphi_n(x)}{\sqrt{M_n}}.$$

证明，f 在 $L^1([0,2\pi])$ 中是良好定义的并且这个级数对几乎处处的 $x\in[0,2\pi]$ 都收敛。

K15) 证明，$f(x)$ 的 Fourier 变换的部分和可以写成

$$(S_Nf)(x)=\Big(S_N\frac{\varphi_n(x)}{\sqrt{M_n}}\Big)(x)+\sum_{s\leqslant n-1}\Big(S_N\frac{\varphi_s(x)}{\sqrt{M_s}}\Big)(x)+\sum_{s\geqslant n+1}\Big(S_N\frac{\varphi_s(x)}{\sqrt{M_s}}\Big)(x).$$

K16) 证明，对任意的 $x\in E_n$，都存在一个 $N_x\leqslant\dfrac{\sqrt{M_n}}{2^n}$，使得 $\left|\Big(S_{N_x}\frac{\varphi_n(x)}{\sqrt{M_n}}\Big)(x)\right|\geqslant\sqrt{M_n}$。

K17) 证明，对任意的 N，我们都有 $\left|\displaystyle\sum_{s\leqslant n-1}\Big(S_N\frac{\varphi_s(x)}{\sqrt{M_s}}\Big)(x)\right|\leqslant\frac{1}{2}\sqrt{M_n}$。

K18) 证明，$\left|\displaystyle\sum_{s\geqslant n+1}\Big(S_N\frac{\varphi_s(x)}{\sqrt{M_s}}\Big)(x)\right|\leqslant\frac{100}{2^n}$。（提示：把 Dirichlet 核函数用它的最大值来控制）

K19) 证明，对几乎处处的 $x\in[0,2\pi]$，$\displaystyle\lim_{N\to\infty}(S_Nf)(x)$ 都发散。

K20) 证明，$f\notin L^2([0,2\pi])$。

注记. *Lennart Carleson 在 1965 年证明了一个著名的定理，这个定理完整地解决了 Lusin 猜想：如果 $f\in L^2([0,2\pi])$，那么 $\displaystyle\lim_{N\to\infty}(S_Nf)(x)$ 几乎处处收敛到 $f(x)$。这一项工作是 Carleson 获得了 Wolf 奖和 Abel 奖的最主要贡献之一。Lusin 是 Kolmogorov 在莫斯科读大学时数学分析的老师。*

The purpose of computing is insight, not numbers.

—- Richard Hamming

57.2 期末考试：Maass 波函数的展开

清华大学 2020 年春季学期数学分析二 期末考试

考试说明

考试时间为下午 14:30 至下午 18:30，请将解答在 A4 大小的白纸上誊写并扫描成 pdf 格式，尽量控制 pdf 文件的大小。务必于下午 18:45 之前将解答上传至网络学堂，否则考试按零分处理。

考试中后面的问题可以使用前面问题的结论（无论答题人是否已经得到正确的证明或者答案），试题出现的先后顺序与其难度毫无关联。

在本次考试中所出现的测度均为 Lebesgue 测度。

多重积分与曲面积分的计算（共 100 分）

A1)（10 分）令 C 为 $\mathbb{R}^3$ 中顶点在 $(0,0,1)$ 处，对称轴为 z-轴并且底在 $z=0$ 平面上的圆锥：

$$C=\left\{(x,y,z)\in\mathbb{R}^3\middle|z\in[0,1],\sqrt{x^2+y^2}=1-z\right\}.$$

试计算 C 的曲面面积。

A2)（10 分）计算二重积分：

$$\iint_{[0,1]\times[0,1]}\left|xy-\frac{1}{4}\right|dxdy.$$

A3)（10 分）考虑 $\mathbb{R}^3$ 中的一个实心圆柱与球面的截面：

$$S=\left\{(x,y,z)\middle|x^2+\left(y-\frac{1}{2}\right)^2\leqslant\frac{1}{4},\ x^2+y^2+z^2=1,z\geqslant 0\right\}.$$

试计算 S 的曲面面积。

A4)（10 分）令 D 为 $\mathbb{R}^3$ 中球心在原点处的单位实心球，即

$$D=\left\{(x,y,z)\middle|x^2+y^2+z^2\leqslant 1\right\}.$$

试计算三重积分：

$$\iiint_D z^2\,dxdydz.$$

A5)（10 分）令 D 为 $\mathbb{R}^3$ 中球心在原点处的单位实心球，即

$$D=\left\{(x,y,z)\middle|x^2+y^2+z^2\leqslant 1\right\}.$$

试计算三重积分：

$$\iiint_D z^3\,dxdydz.$$

A6)（10 分）考虑 $\mathbb{R}^3$ 中球心在原点处的八分之一单位实心球，即

$$D_+ = \left\{(x,y,z)\middle| x^2+y^2+z^2 \leqslant 1, x \geqslant 0, y \geqslant 0, z \geqslant 0\right\}.$$

试计算三重积分：

$$\iiint_{D_+} xyz\, dxdydz.$$

A7)（20 分）给定 $(a,b,c) \in \mathbb{R}^3$，$R>0$，考虑球面

$$S = \left\{(x,y,z)\middle|(x-a)^2+(y-b)^2+(z-c)^2 = R^2\right\}.$$

我们将 S 上的外法向量取做指向球面的外部。试计算第二型曲面积分

$$\iint_S x^2 dydz + y^2 dxdz + z^2 dxdy.$$

A8)（20 分）假设 X $(X_1(x,y,z), X_2(x,y,z), X_3(x,y,z))$ 是 $\mathbb{R}^3$ 中的光滑向量场并且散度为零，即 $\mathrm{div}X = 0$。我们用 $\mathbf{S}_R$ 表示 $\mathbb{R}^3$ 中中心在原点半径为 R 的球面，即

$$\mathbf{S}_R = \left\{(x,y,z)\middle| x^2+y^2+z^2 = R^2\right\}.$$

证明，以下的量不依赖于 R 的选取：

$$I(R) = \frac{1}{R}\int_{\mathbf{S}_R} xX_1(x,y,z) + yX_2(x,y,z) + zX_3(x,y,z) d\sigma,$$

其中 $d\sigma$ 是 $\mathbf{S}_R$ 的曲面测度。

具有非负 Fourier 系数的连续函数（共 25 分）

给定 $\mathbb{R}$ 上以 2π 为周期的连续函数 $f \in C^0(\mathbf{T})$，如果对任意的 $m \in \mathbb{Z}_{\geqslant 1}$，任意的 $x_a \in \mathbb{R}$，$c_a \in \mathbb{C}$，$a = 1, 2, \cdots, m$，我们都有

$$\sum_{a=1}^{m}\sum_{b=1}^{m} c_a \overline{c_b} f(x_a - x_b) \geqslant 0,$$

我们就称 f 是**正定的**。

B1)（2 分）给定有限的 Fourier 级数

$$f(x) = \sum_{|k|\leqslant N} \widehat{f}(k) e^{ik\cdot x},$$

如果对任意的 $|k| \leqslant N$，$\widehat{f}(k) \geqslant 0$，证明，$f$ 是正定的。

B2)（8 分）假设 $f \in C^0(\mathbf{T})$ 并且对任意的 $k \in \mathbb{Z}$，$\widehat{f}(k) \geqslant 0$。证明，$f$ 是正定的。（提示：利用 Féjer 核）

B3)（7 分）假设 $f \in C^0(\mathbf{T})$ 是正定的。证明，对任意的 $g \in C^0(\mathbf{T})$，我们都有

$$\int_{-\pi}^{\pi}\int_{-\pi}^{\pi} f(x-y) g(x) \overline{g(y)} \frac{dxdy}{(2\pi)^2} \geqslant 0.$$

B4)（8 分）假设 $f \in C^0(\mathbf{T})$ 是正定的。证明，对任意的 $k \in \mathbb{Z}$，$\widehat{f}(k) \geqslant 0$。

Maass 波函数（共 45 分）

我们用 $\mathbb{H}$ 表示上半平面

$$\mathbb{H} = \big\{(x,y) \in \mathbb{R}^2 \big| y > 0\big\}.$$

对任意的 $\varepsilon > 0$，令

$$\mathbb{X}_\varepsilon = \big\{(x,y) \in \mathbb{R}^2 \big| x \in (0,1), y \in (\varepsilon, +\infty)\big\} \subset \mathbb{H}$$

对于 $\mathbb{H}$ 上定义的复值函数 f，如果对任意的 $z = (x,y) \in \mathbb{H}$，都有

$$f(x+1, y) = f(x,y),$$

我们就称 f 是**以 $\mathbb{Z}$ 为周期的**。我们定义

$$\mathscr{E} = \big\{f : \mathbb{H} \to \mathbb{C} \big| f \text{ 以 } \mathbb{Z} \text{ 为周期并且对任意的 } \varepsilon > 0,\ f\big|_{\mathbb{X}_\varepsilon} \in L^2(\mathbb{X}_\varepsilon, dxdy)\big\}.$$

在 $\mathbb{H}$ 上，我们定义双曲的 Laplace 算子：

$$\triangle_{\mathbb{H}} = y^2 \left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}\right).$$

对任意的 $s \in \mathbb{R}$，我们定义

$$\mathscr{E}_s = \big\{\psi \in C^2(\mathbb{H}) \cap \mathscr{E} \ \big|\ \triangle_{\mathbb{H}}\psi = (s^2 - \frac{1}{4})\psi\big\}.$$

这个题目的目标是给出 $\mathscr{E}_s$ 的刻画。

M1)（2 分）任给 $f \in \mathscr{E}$。证明，对几乎处处的 $y \in (0, +\infty)$，一元函数

$$(0,1) \to \mathbb{C},\ x \mapsto f(x,y)$$

落在 $L^2\big((0,1), dx\big)$ 中。

M2)（3 分）任给 $f \in \mathscr{E}$，对任意的 $k \in \mathbb{Z}$，证明，积分

$$a_k(f,y) = \int_0^1 f(x,y) e^{-2\pi i k x} dx$$

对几乎处处 $y > 0$ 有定义。将 $a_k(f,y)$ 视作是 y 的函数，证明，对任意 $\varepsilon > 0$，$a_k(f,y) \in L^2\big((\varepsilon, \infty), dy\big)$。

M3)（4 分）对任意的 $y > 0$ 和 $s \in \mathbb{R}$，我们定义

$$K_s(y) = \int_0^\infty e^{-y(u+u^{-1})} u^{s-1} du.$$

证明，$K_s(y)$ 对一切 $y > 0$ 有定义并且是光滑函数。我们还有公式

$$K_s^{(n)}(y) = \left(\frac{d}{dy}\right)^n K_s(y) = (-1)^n \int_1^\infty e^{-y(u+u^{-1})} (u^{s-1} + u^{-s-1})(u + u^{-1})^n du.$$

M4)（3 分）证明，对任意的 $n \geqslant 0$，$K_s(y)$ 在 ∞ 附近有如下的渐近行为：

$$\lim_{y\to+\infty} \sqrt{y} e^{2y} K_s^{(n)}(y) = (-2)^n \sqrt{\pi}.$$

（提示：可以考虑变量替换：$u \mapsto 1 + \frac{v}{\sqrt{y}}$）

M5)（4 分）定义微分算子 D：

$$Df(y) = y^2 \frac{d^2 f}{dy^2} + y \frac{df}{dy} - 4y^2 f.$$

证明，

$$DK_s = s^2 K_s.$$

（所有代数运算，如果无相对具体的步骤而说经整理得到，视作无效）

M6)（4 分）假设在 $(0, +\infty)$ 上的定义的 C^2 函数 f 满足

$$Df = s^2 f.$$

证明，存在复常数 c_1 和 c_2，使得

$$f = \left(c_1 + c_2 \int_0^y \frac{dy'}{K_s(y')^2 \cdot y'} \right) K_s(y).$$

M7)（3 分）假设在 $(0, +\infty)$ 上的定义的 C^2 函数 f 满足

$$Df = s^2 f.$$

证明，极限 $\lim\limits_{y\to+\infty} \sqrt{y} e^{-2y} f(y)$ 存在。进一步证明，$\lim\limits_{y\to+\infty} \sqrt{y} e^{-2y} f(y) = 0$ 当且仅当存在常数 $c_3 \in \mathbb{C}$，使得

$$f(y) \equiv c_3 \cdot K_s(y).$$

以下我们假设 $s > 1$。

M8)（2 分）证明，任给的 $f \in \mathscr{E}_s$，对任意的 $(x, y) \in \mathbb{H}$，级数 $\sum\limits_{n\in\mathbb{Z}} a_k(f, y) e^{2\pi i k x}$ 收敛并且

$$\sum_{n\in\mathbb{Z}} a_k(f, y) e^{2\pi i k x} = f(x, y).$$

M9)（2 分）任给的 $f \in \mathscr{E}_s$。证明，对任意的 $k \in \mathbb{Z}$，$a_k(f, y) e^{2\pi i k x} \in \mathscr{E}_s$。

M10)（4 分）证明，如果 $k \in \mathbb{Z}$ 且 $k \neq 0$，那么，存在常数 $a_k(f) \in \mathbb{C}$，使得

$$a_k(f, y) = a_k(f) \cdot y^{\frac{1}{2}} \cdot K_s(\pi \cdot |k| \cdot y).$$

M11)（4 分）证明，存在常数 $a_0(f) \in \mathbb{C}$，使得

$$a_0(f,y) = a_0(f) \cdot y^{\frac{1}{2}-s}.$$

M12)（5 分）证明，对任意的 $z \in \mathbb{C}$，$|z| < 1$，级数

$$\sum_{k=1}^{\infty} \left(|a_k(f)| + |a_{-k}(f)|\right) z^k$$

收敛。

M13)（5 分）任意给定 $\{a_k\}_{k\in\mathbb{Z}} \subset \mathbb{C}$，假定对任意的 $z \in \mathbb{C}$，$|z| < 1$，级数

$$\sum_{k=1}^{\infty} \left(|a_k| + |a_{-k}|\right) z^k$$

收敛。

证明，对任意的 (x,y)，级数

$$a_0 y^{\frac{1}{2}-s} + \sum_{k\neq 0} a_k y^{\frac{1}{2}} K_s(\pi \cdot |k| \cdot y) e^{2\pi i k x}$$

所定义的函数落在 $\mathscr{E}_s$ 中。

What is important is to deeply understand things and their relations to each other. This is where intelligence lies. The fact of being quick or slow isn't really relevant.

—— Laurent Schwartz

简介

数学分析三课程主要涵盖如下的内容：

1) $\mathbb{R}^n$ 及其开集上的分布

   - 分布的定义及基本例子；
   - 分布的操作与分布的 Stokes 公式；
   - 分布的支集和卷积；
   - 初步应用：偏微分方程的基本解与求解。

2) $\mathbb{R}^n$ 上的 Fourier 分析

   - Schwartz 空间与缓增分布；
   - 缓增分布的 Fourier 变换；
   - Fourier 变换的计算；
   - 初步应用：偏微分方程的基本解与求解。

3) $\mathbb{R}^n$ 上的 Sobolev 空间

   - Sobolev 空间的物理空间定义与频率空间定义；
   - Sobolev 空间在微分算子作用下的代数性质；
   - Sobolev 空间的限制（迹）定理：半空间与有限区域；
   - 应用：Dirichlet 问题的正则性定理。

4) 应用举例

   - $-\triangle$ 的特征函数与特征值；
   - 热核的构造与 Weyl 关于特征值渐近公式
   - 微局部分析一瞥：波前集，微局部椭圆正则性与，奇性传播定理。

58 分布的定义与例子：Radon 测度，局部可积函数到分布的嵌入，主值部分

二零二零年九月十四日，星期一，晴

分布理论是法国数学家 Laurent M. Schwartz 在上个世界四十年代引入的，他在 1950 出版了 *Théorie des distributions* 一书总结了分布理论的精要与应用，这项工作也是他获得 1950 年 Fields 奖的核心贡献。中文**分布**二字是书名中 distributions 的直译，却没有说明这个理论究竟讲了什么。Schwartz 在分布理论方面的第一篇论文的题目可以解答这个疑惑：这篇文章是 *Généralisation de la notion de fonction, de dérivation, de transformation de Fourier et applications mathématiques et physiques*，即对函数、微分和 Fourier 变换的推广及在数学和物理中的应用。我们在第一学年已经详细地学习了函数、微分和 Fourier 变换，这个学期就秉承 Schwartz 的观点通过对之前概念的推广去欣赏分析学在数学和物理中的应用。

分布的定义与例子

如果不加说明，我们总假设 $\Omega\subset\mathbb{R}^n$ 是非空开集，其中 $n\geqslant 1$。给定紧集 $K\subset\mathbb{R}^n$，我们用 $C_K^\infty(\Omega)$ 表示 Ω 上的支集在 K 中的光滑函数所组成的集合，即

$$C_K^\infty(\Omega)=\big\{f\in C^\infty(\Omega)\big|\mathrm{supp}(f)\subset K\big\}.$$

给定一个多重指标 $\alpha=(\alpha_1,\cdots,\alpha_n)$，其中每个分量都是非负整数，符号 $\partial^\alpha\varphi$ 表示如下多重偏导数

$$\partial_{x_1}^{\alpha_1}\partial_{x_2}^{\alpha_2}\cdots\partial_{x_n}^{\alpha_n}\varphi,$$

其中 $\partial_{x_i}^{\alpha_i}=\dfrac{\partial^{\alpha_i}}{\partial x_i^{\alpha_i}}$。另外，我们令

$$|\alpha|=\sum_{i=1}^n\alpha_i.$$

我们用 $\mathcal{D}(\Omega)=C_0^\infty(\Omega)$ 表示在 Ω 上定义并且有紧支集的光滑函数所组成的集合，我们把它称作是为**试验函数空间**。按照定义，对于任意的 $\varphi\in\mathcal{D}(\Omega)$，存在紧集 $K\subset\Omega$（这是 $\mathbb{R}^n$ 中的紧集），使得 $f\big|_{\Omega-K}\equiv 0$，即对任意的 $x\in\Omega-K$，$f(x)=0$。

定义 389 (试验函数空间). 在空间 $\mathcal{D}(\Omega)$ 上，我们规定如下的收敛性（拓扑）：给定函数序列 $\{\varphi_p\}_{p\geqslant 1}\subset\mathcal{D}(\Omega)$，所谓的该序列收敛到 $0\in\mathcal{D}(\Omega)$（记作 $\varphi_p\xrightarrow{\mathcal{D}(\Omega)}0$），指的是

1) 存在紧集 $K\subset\Omega$，使得对每个 $p\geqslant 1$，都有 $\mathrm{supp}(\varphi_p)\subset K$；

2) 对每个多重指标 α，函数序列 $\{\partial^\alpha\varphi_p\}_{p\geqslant 1}$ 在 K 上一致收敛到 0，即

$$\lim_{p\to\infty}\|\partial^\alpha\varphi_p\|_{L^\infty(K)}=0.$$

定义 390 (分布)**.** 所谓 Ω 上的一个**分布**（也称作**广义函数**）指的是 $\mathcal{D}(\Omega)$ 上的一个线性泛函（线性映射）：

$$u : \mathcal{D}(\Omega) \to \mathbb{C}, \quad \varphi \mapsto \langle u, \varphi \rangle,$$

满足如下两个条件

1) 对任意的 $\varphi, \psi \in \mathcal{D}(\Omega)$ 和 $\alpha, \beta \in \mathbb{C}$，我们有

$$\langle u, \alpha\varphi + \beta\psi \rangle = \alpha\langle u, \varphi \rangle + \beta\langle u, \psi \rangle.$$

2) 对任意的紧集 $K \subset \Omega$，存在非负整数 p 和正常数 C（p 和 C 依赖于 K），使得对任意的 $\varphi \in C_K^\infty(\Omega)$，都有

$$|\langle u, \varphi \rangle| \leqslant C \sup_{|\alpha| \leqslant p} \|\partial^\alpha \varphi\|_{L^\infty(K)}.$$

如果上述的 p 的选取不依赖于紧集 K 的选取，那么，我们就把最小的这样的非负整数 p 称作是分布 u 的**阶**。

我们用 $\mathcal{D}'(\Omega)$ 表示在 Ω 上定义的分布的全体，并规定如下的收敛性（拓扑）：所谓的分布序列 $\{u_p\}_{p\geqslant} \subset \mathcal{D}'(\Omega)$ 在分布的意义下收敛到 $0 \in \mathcal{D}'(\Omega)$（这是把所有函数映射成 0 的线性映射），记作 $u_p \xrightarrow{\mathcal{D}'(\Omega)} 0$，指的是对任意的试验函数 $\varphi \in \mathcal{D}(\Omega)$，我们都有

$$\lim_{p\to\infty} \langle u_p, \varphi \rangle = 0.$$

我们通常说一个分布可以和一个（有紧支集的）光滑函数配对得到一个数，即

$$\mathcal{D}'(\Omega) \times \mathcal{D}(\Omega), \quad (u, \varphi) \mapsto \langle u, \varphi \rangle.$$

注记. 任意给定分布 $u \in \mathcal{D}'(\Omega)$，它实际上是 $\mathcal{D}(\Omega)$ 上的**连续**线性泛函。我们没有详细讨论和分布有关的拓扑线性空间的理论，所以我们不打算对这一点做太多的展开（这对于理解分布理论也没有影响）。所谓的连续性，可以用下面的序列的语言来描述：对任意的试验函数序列 $\varphi_p \xrightarrow{\mathcal{D}} 0$，我们都有

$$\lim_{p\to\infty} \langle u, \varphi_p \rangle = 0.$$

这个命题的证明只需要用到分布概念，我们把它留作作业。

作为例子，我们先学习一些重要的分布：

例子 (Dirac 函数)**.** 对任意的 $a \in \Omega$，我们可以定义分布 $\delta_a \in \mathcal{D}'(\Omega)$。其中，对于任意的 $\varphi \in \mathcal{D}(\Omega)$，我们定义

$$\langle \delta_a, \varphi \rangle = \varphi(a).$$

我们来验证 δ_a 实际上是分布：

对任意的紧集 $K \subset \Omega$，如果 $a \notin K$，那么，对任意的 $\varphi \in C_K^\infty(\Omega)$，我们都有

$$\langle \delta_a, \varphi \rangle = 0.$$

如果 $a \in K$，那么，使得对任意的 $\varphi \in C_K^\infty(\Omega)$，我们有

$$\big|\langle \delta_a, \varphi\rangle\big| = \big|\varphi(a)\big| \leqslant 1 \cdot \sup_{|\alpha|\leqslant 0} \|\partial^\alpha \varphi\|_{L^\infty(K)}.$$

所以，我们在分布的定义中取 $q = 0$，$C = 1$ 即可。

特别地，我们还知道 δ_a 的阶为 0。

例子 (局部可积的函数)**.** 给定开集 Ω（总是装配了 *Borel* 代数和 *Lebesgue* 测度），所谓局部可积的函数指的是在每个紧的局部上都可积的函数，即可测函数 f（所对应的几乎处处相等的函数的等价类），对于任意紧集 $K \subset \Omega$，函数 $f \cdot \mathbf{1}_K \in L^1(\Omega)$。我们用 $L^1_{\mathrm{loc}}(\Omega)$ 表示 Ω 上局部可积的函数。

对于任意的 $f \in L^1_{\mathrm{loc}}(\Omega)$，我们定义 $\mathcal{D}(\Omega)$ 上的线性泛函：

$$T_f : \mathcal{D}(\Omega) \to \mathbb{C}, \varphi \mapsto \langle T_f, \varphi\rangle = \int_\Omega f(x)\varphi(x)dx.$$

由于 φ 在它的支集 K 上有届，所以，上面的积分是良好定义的。我们证明 T_f 是 Ω 上的阶为 0 的分布：对任意的紧集 $K \subset \Omega$，对任意的 $\varphi \in C_K^\infty(\Omega)$，我们有

$$\begin{aligned}\big|\langle T_f, \varphi\rangle\big| &= \Big|\int_K f(x)\varphi(x)dx)\Big| \\ &\leqslant \|f\|_{L^1(K)}\|\varphi\|_{L^\infty(K)}.\end{aligned}$$

所以，我们在分布的定义中取 $q = 0$，$\|f\|_{L^1(K)}$ 即可。

注记. 为了方便起见，我们通常把 $\langle T_f, \varphi\rangle$ 直接写成 $\langle f, \varphi\rangle$。

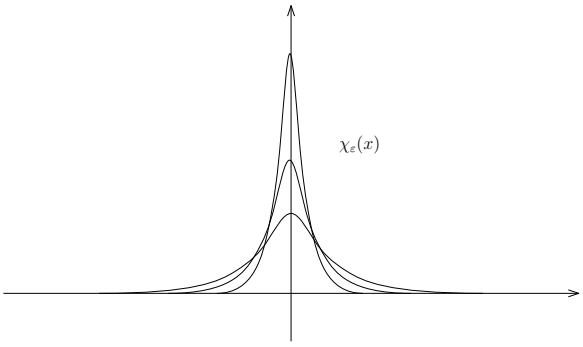
命题 391. 任意选定 $\chi(x) \in \mathcal{D}(\mathbb{R}^n)$（我们通常偏爱之前所构造的那个 $\chi(x)$），我们假定

$$\int_{\mathbb{R}^n} \chi(x)dx = 1.$$

对任意的 $\varepsilon > 0$，我们定义

$$\chi_\varepsilon(x) = \frac{1}{\varepsilon^n}\chi(\frac{x}{\varepsilon}).$$

那么，在分布的意义下，当 $\varepsilon \to 0$ 时，我们有 $\chi_\varepsilon \xrightarrow{\mathcal{D}'} \delta_0$。



注记. 由于 χ_ε 是局部可积的（因为在紧集上有最大最小值），我们通过

$$\int \chi_\varepsilon(x)\varphi(x)dx$$

把 χ_ε 视作分布，请参考上一个例子（之后不再重复）。

证明: 通过换元积分公式，我们知道对任意的 $\varepsilon > 0$，我们都有

$$\int_{\mathbb{R}^n} \chi_\varepsilon(x)dx = \int_{\mathbb{R}^n} \chi(x)dx, \quad \int_{\mathbb{R}^n} |\chi_\varepsilon(x)|dx = \int_{\mathbb{R}^n} |\chi(x)|dx.$$

任意选定试验函数 φ，我们要证明

$$\lim_{\varepsilon\to 0}\langle \chi_\varepsilon - \delta_0, \varphi\rangle = 0.$$

利用 $\int_{\mathbb{R}^n} \chi_\varepsilon(x)dx = 1$，我们有

$$\begin{aligned}\langle \chi_\varepsilon - \delta_0, \varphi\rangle &= \int_{\mathbb{R}^n} \chi_\varepsilon(x)\varphi(x)dx - \varphi(0) = \int_{\mathbb{R}^n} \chi_\varepsilon(x)(\varphi(x)-\varphi(0))dx \\ &= \underbrace{\int_{|x|\leqslant\delta} \chi_\varepsilon(x)(\varphi(x)-\varphi(0))dx}_{\mathbf{I}_1} + \underbrace{\int_{|x|\geqslant\delta} \chi_\varepsilon(x)(\varphi(x)-\varphi(0))dx}_{\mathbf{I}_2}\end{aligned}$$

对于 $\mathbf{I}_1$ 而言，当 $\varepsilon \to 0$ 时，$\varphi(x)$ 一致收敛到 $\varphi(0)$（因为 $|x| \leqslant \varepsilon$），所以，我们有

$$\lim_{\delta\to 0} \mathbf{I}_1 \leqslant o(1)\int_{|x|\leqslant\delta} |\chi_\varepsilon(x)|dx = o(1)\|\chi\|_{L^1}.$$

所以，可以先选 δ，使得

$$\mathbf{I}_1 \leqslant \epsilon,$$

其中，ϵ 是任意给定的正实数。对于 $\mathbf{I_2}$，现在已经固定了 δ，我们有

$$\begin{aligned}\mathbf{I}_2 &\leqslant 2\|\varphi\|_{L^\infty}\int_{|x|\geqslant\delta} |\chi_\varepsilon(x)|dx = 2\|\varphi\|_{L^\infty}\int_{|x|\geqslant\frac{\delta}{\delta}} |\chi(x)|dx \\ &= 2\|\varphi\|_{L^\infty}\int_{\mathbb{R}^n} |\chi(x)|\mathbf{1}_{|x|\geqslant\frac{\delta}{\delta}}dx.\end{aligned}$$

由于当 $\varepsilon \to 0$ 时，$|\chi(x)|\mathbf{1}_{|x|\geqslant\frac{\delta}{\varepsilon}}$ 是逐点收敛到 0 的，所以，根据 Lebesgue 控制收敛定理，

$$\lim_{\varepsilon\to 0} \mathbf{I}_2 = 0.$$

特别地，对足够小的 $\varepsilon > 0$，我们就有

$$\mathbf{I}_2 \leqslant \epsilon.$$

这表明，

$$\mathbf{I}_1 + \mathbf{I}_2 \leqslant 2\epsilon.$$

从而命题成立。 □

例子 (Radon 测度)**.** 假设 μ 是 $(\Omega, \mathcal{B}(\Omega))$ 上的测度，其中，$\Omega \subset \mathbb{R}^n$ 是开集，$\mathcal{B}(\Omega)$ 是 *Borel* 代数（包含所有开集的最小 σ-代数）。如果每个紧集 $K \subset \Omega$，$\mu(K) < \infty$，我们就把这种测度称作是一个 *Radon* 测度。

比如说，对任意的正函数（几乎处处）$f \in L^1_{\mathrm{loc}}(\Omega)$，对任意的 $B \in \mathcal{B}(\Omega)$，我们可以定义

$$\mu_f(B) = \int_\Omega \mathbf{1}_B \cdot f(x)dx.$$

这就是一个 *Radon* 测度。

任意给定一个 *Radon* 测度 μ，我们可以定义一个分布 T_μ：对于 $\varphi \in \mathcal{D}(\Omega)$，我们要求

$$\langle T_\mu, \varphi\rangle = \int_\Omega \varphi(x)d\mu(x).$$

我们证明 T_μ 是 Ω 上阶为 0 的分布：

对任意的紧集 $K \subset \Omega$，对任意的 $\varphi \in C^\infty_K(\Omega)$，我们有

$$\begin{aligned}\big|\langle T_\mu, \varphi\rangle\big| &= \Big|\int_K \varphi(x)d\mu(x))\Big| \\ &\leqslant \mu(K)\|\varphi\|_{L^\infty(K)}.\end{aligned}$$

所以，我们在分布的定义中取 $q = 0$，$C = \mu(K)$ 即可。

特别地，我们可以把 $L^1_{\mathrm{loc}}(\Omega)$ 中的元素看作是某个 *Radon* 测度的密度函数，从而，定义出了同样的分布。

注记. 利用所谓的 *Riesz* 表示定理，我们可以证明，Ω 上所有的 0 阶分布都是（由如上方式给出的）*Radon* 测度。

我们自然还有阶非零的分布，比如说，我们可以定义 $\mathcal{D}(\mathbb{R})$ 上的线性泛函：

$$\langle \delta', \varphi\rangle = -\varphi'(0),$$

和

$$\langle u, \varphi\rangle = \sum_{k=0}^{\infty} \varphi^{(k)}(k).$$

我们将在作业题中证明两个定义给出了分布，第一个的阶为 1，而第二个的阶不能定义（无穷大）。

命题 392 (局部可积的函数)**.** 给定开集 $\Omega \subset \mathbb{R}^n$，我们已经定义如下的线性映射（把局部可积函数视为分布）

$$T : L^1_{\mathrm{loc}}(\Omega) \to \mathcal{D}'(\Omega), \quad f \mapsto T_f.$$

这是单射。

注记. 根据这个命题，局部可积的函数可以看做是分布的子集合。在分析中，我们把 $L^1_{\mathrm{loc}}(\Omega)$ 的元素称作是 Ω 上的“函数”（这个类已经足够大了），由于某些分布不是“函数”，所以我们也经常把分布称作是“广义函数”。我们在课程中不使用这个名称。

证明：假设 $T_f \overset{\mathcal{D}'}{=} 0$，即对任意的 $\varphi \in \mathcal{D}(\Omega)$，我们都有

$$\int_\Omega f(x)\varphi(x)dx = 0,$$

我们要说明 $f = 0$（几乎处处）。为此，只需要说明对每个紧集 $K \subset \Omega$，我们都有 $f\big|_K = 0$（几乎处处）即可。

我们定义函数

$$\varphi_K(x) = \begin{cases} \frac{\overline{f(x)}}{|f(x)|}\mathbf{1}_K(x), & \text{如果 } f(x) \neq 0; \\ 0, & \text{如果 } f(x) = 0. \end{cases}$$

这是一个有紧支集 K 的函数。当 ε 较小时，$\chi_\varepsilon * \varphi_K$ 的支集仍然在 Ω 中：按照卷积的定义，对任意的 $x \in \mathrm{supp}(\chi_\varepsilon * \varphi_K)$，这个点距离 K 的不超过 ε，然而，距离 K 的不超过 ε 点是在 Ω 中的（利用紧性）。特别地，我们得到一个试验函数 $\chi_\varepsilon * \varphi_K \in \mathcal{D}(\Omega)$，从而

$$\langle T_f, \chi_\varepsilon * \varphi_K \rangle = 0.$$

我们令 $\varepsilon < \delta$ 并且要求距离 K 的不超过 ε 点都落在 Ω 中。我们令

$$K + B(\delta) = \big\{x + y | x \in K, |y| \leqslant \delta\big\}.$$

根据定义，我们就有

$$\begin{aligned} 0 &= \int_\Omega f(x)\big(\chi_\varepsilon * \varphi_K\big)(x)dx = \int_{\mathbb{R}^n} f(x)\big(\chi_\varepsilon * \varphi_K\big)(x)dx \\ &= \int_{\mathbb{R}^n} \big(f \cdot \mathbf{1}_{K+B(\delta)}\big)\big(\chi_\varepsilon * \varphi_K\big)dx \\ &\overset{Fubini}{=} \iint_{\mathbb{R}^n \times \mathbb{R}^n} \big(f(x)\mathbf{1}_{K+B(\delta)}(x)\big)\chi_\varepsilon(x-y)\varphi_K(y)dxdy \\ &\overset{Fubini}{=} \int_{\mathbb{R}^n} \left(\int_{\mathbb{R}^n} \big(f(x)\mathbf{1}_{K+B(\delta)}(x)\big)\chi_\varepsilon(x-y)dx\right)\varphi_K(y)dy \\ &= \int_{\mathbb{R}^n} \Big(\big(f\mathbf{1}_{K+B(\delta)}\big) * \chi_\varepsilon\Big)(y)\varphi_K(y)dy. \end{aligned}$$

然而，我们上个学期证明过 $\big(f\mathbf{1}_{K+B(\delta)}\big) * \chi_\varepsilon \xrightarrow{L^1} f\mathbf{1}_{K+B(\delta)}$，从而（由于 $\varphi_K(y)$ 有界）当 $\varepsilon \to 0$，我们有

$$\begin{aligned} 0 &= \lim_{\varepsilon \to 0} \langle T_f, \chi_\varepsilon * \varphi_K \rangle \\ &= \int_{\mathbb{R}^n} \big(f\mathbf{1}_{K+B(\delta)}\big)(y)\varphi_K(y)dy = \int_K |f(y)|dy. \end{aligned}$$

所以对几乎处处的 $x \in K$，我们有 $f(x) = 0$，命题得证。 □

例子 ($\frac{1}{x}$ 的主值部分)**.** 我们注意到

$$\frac{1}{x} \notin L^1_{\mathrm{loc}}(\mathbb{R}).$$

这是因为我们不能在 0 附近对 x^{-1} 积分（积出来无穷大）。所以，我们不能通过直接与试验函数配对积分的方式定义一个分布。我们想用取极限（逼近）的方式来定义：对任意的 $n \geqslant 1$，我们显然有

$$\frac{1}{x}\mathbf{1}_{|x|\geqslant\frac{1}{n}}(x) \in L^1_{\mathrm{loc}}(\mathbb{R}).$$

作为分布，我们就有

$$\begin{aligned}\langle \frac{1}{x}\mathbf{1}_{|x|\geqslant\frac{1}{n}}, \varphi\rangle &= \int_{-\infty}^{\frac{1}{n}} \frac{\varphi(x)}{x}dx + \int_{\frac{1}{n}}^{\infty} \frac{\varphi(x)}{x}dx \\ &= \int_{\frac{1}{n}}^{\infty} \frac{\varphi(x)-\varphi(-x)}{x}dx.\end{aligned}$$

其中，φ 是 $\mathbb{R}$ 上的一个试验函数。

我们现在有个简单但是重要的观察：

$$\varphi(x)-\varphi(-x)\big|_{x=0} = 0.$$

根据下面的引理（证明参考第一次作业）

引理 393. 假设 $\psi(x) \in C^\infty(\mathbb{R})$ 并且 $\psi(0) = 0$，那么，$\dfrac{\psi(x)}{x}$ 也是光滑函数。

所以 $\dfrac{\varphi(x)-\varphi(-x)}{x}$ 是光滑函数（自然是局部可积的）。从而，当 $n \to \infty$ 时，上述积分的极限存在：

$$\lim_{n\to\infty}\int_{\frac{1}{n}}^{\infty} \frac{\varphi(x)-\varphi(-x)}{x}dx = \int_0^{\infty} \frac{\varphi(x)-\varphi(-x)}{x}dx.$$

我们现在定义

$$\langle \mathrm{vp}\frac{1}{x}, \varphi\rangle = \int_0^{\infty} \frac{\varphi(x)-\varphi(-x)}{x}dx.$$

为了证明这是分布，我们利用中值定理：对任意的紧集 $K = [-M, M] \subset \mathbb{R}$，对任意的支集在 K 上的光滑函数 φ，我们有

$$|\varphi(x)-\varphi(-x)| = |2x\varphi'(\xi)| \leqslant 2\|\varphi'\|_{L^\infty(K)}x.$$

所以，

$$|\langle \mathrm{vp}\frac{1}{x}, \varphi\rangle| \leqslant \int_0^M 2\|\varphi'\|_{L^\infty(K)}dx = 2M\|\varphi'\|_{L^\infty(K)}.$$

所以 $\mathrm{vp}\dfrac{1}{x}$ 是一个阶不超过 1 的分布。在作业中，我们将证明 $\mathrm{vp}\dfrac{1}{x}$ 的阶恰好是 1。

另外，vp 是法语 *valeur principale* 的缩略，英文文献经常用 $\mathrm{pv}\dfrac{1}{x}$，因为他们把主值写为 *principal value.*

59 分布的操作：限制，求导数，与微分同胚复合，链式法则.Stokes 公式的分布形式

二零二零年九月十六日，星期三，晴

分布的操作

在微积分的学习中，我们可以对一个函数做特定的操作，比如可以把一个函数限制到比较小的定义域上、可以对一个函数求导数、两个函数可以相乘等等。我们现在讨论如何对分布做一些特定的操作。

分布的限制

假设 $\Omega' \subset \Omega$ 是开子集，那么，我们可以定义限制映射

$$\mathrm{Res} : \mathcal{D}'(\Omega) \to \mathcal{D}'(\Omega'), u \mapsto \mathrm{Res}(u).$$

其中，对于每个 $\varphi \in \mathcal{D}(\Omega')$，它自然可以看作是 $\mathcal{D}(\Omega)$ 中的元素，从而，我们可以要求

$$\langle \mathrm{Res}(u), \varphi \rangle = \langle u, \varphi \rangle.$$

为了方便起见，除了个别场合，我们总是用 u 来直接表示 $\mathrm{Res}(u)$。

求偏导数

我们先考虑一个足够光滑的函数，比如说，$u \in C^1(\mathbb{R})$。u 和 u' 都是局部可积分的函数，我们可以把它们视作是分布。通过分部积分，我们有

$$\begin{aligned}\langle u', \varphi \rangle &= \int_{\mathbb{R}} u'(x)\varphi(x)dx = -\int_{\mathbb{R}} u(x)\varphi'(x)dx \\ &= -\langle u, \varphi' \rangle.\end{aligned}$$

这个计算启发我们对于 $u \in \mathcal{D}'(\mathbb{R})$，我们可以用下面的等式来定义它的导数：

$$\langle u', \varphi \rangle := -\langle u, \varphi' \rangle.$$

定义 394. 假设 $\Omega \subset \mathbb{R}^n$ 是有界开集，给定 $u \in \mathcal{D}'(\Omega)$，对于任意的多重指标 α，我们定义

$$\langle \partial^\alpha u, \varphi \rangle = (-1)^{|\alpha|} \langle u, \partial^\alpha \varphi \rangle.$$

我们必须说明 $\partial^\alpha u$ 定义了 Ω 上的分布：根据分布的定义，对任意的紧集 $K \subset \Omega$，存在非负整数 p 和正常数 C（p 和 C 依赖于 K），使得对任意的 $\varphi \in C_K^\infty(\Omega)$，都有

$$\begin{aligned}|\langle \partial^\alpha u, \varphi\rangle| &= |\langle u, \partial^\alpha \varphi\rangle| \\ &\leqslant C \sup_{|\beta|\leqslant p} \|\partial^\beta \partial^\alpha \varphi\|_{L^\infty(K)} \\ &\leqslant C \sup_{|\gamma|\leqslant p+|\alpha|} \|\partial^\gamma \varphi\|_{L^\infty(K)}\end{aligned}$$

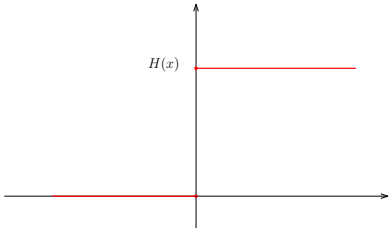
这表明我们对任意的分布都可以求导数，即对任意的多重指标 α，我们有（连续）线性映射：

$$\partial^\alpha : \mathcal{D}'(\Omega) \to \mathcal{D}'(\Omega).$$

我们研究两个经典的例子：

例子 (Heaviside 函数)**.** 我们定义 *Heaviside* 函数

$$H(x) = \mathbf{1}_{x\geqslant 0}.$$



直观上，这个函数在 0 之外的导数是 0，在 0 处函数有很大的跳跃，导数应该是无穷大。我们证明

$$H(x)' = \delta_0.$$

按照定义，和我们有

$$\begin{aligned}\langle H'(x), \varphi(x)\rangle &= -\langle H(x), \varphi'(x)\rangle = \int_0^\infty \varphi'(x)dx \\ &= -\left(\varphi(\infty) - \varphi(0)\right) = \varphi(0) \\ &= \langle \delta_0, \varphi\rangle.\end{aligned}$$

例子. 我们考虑 $\mathbb{R}$ 上的 log 函数，注意到

$$\log|x| \in L^1_{\mathrm{loc}}(\mathbb{R}).$$

我们证明，作为分布，有

$$\left(\log|x|\right)' \overset{\mathcal{D}'}{=} \mathrm{vp}\frac{1}{x}.$$

实际上，

$$\begin{aligned}\langle\big(\log|x|\big)',\varphi\rangle &= -\lim_{\varepsilon\to 0}\Big(\int_{-\infty}^{-\varepsilon}\varphi'\log|x|dx+\int_{\varepsilon}^{\infty}\varphi'\log|x|dx\Big)\\ &= \lim_{\varepsilon\to 0}\Big(\varphi(-\varepsilon)\log(\varepsilon)+\int_{-\infty}^{-\varepsilon}\frac{1}{x}\varphi dx-\varphi(\varepsilon)\log(\varepsilon)+\int_{\varepsilon}^{\infty}\varphi'\log|x|dx\Big).\end{aligned}$$

我们注意到

$$\varphi(-\varepsilon)-\varphi(\varepsilon)=O(\varepsilon).$$

所以，

$$\lim_{\varepsilon\to 0}\varphi(-\varepsilon)\log(\varepsilon)-\varphi(\varepsilon)\log(\varepsilon)=0.$$

从而，

$$\begin{aligned}\langle\big(\log|x|\big)',\varphi\rangle &= \lim_{\varepsilon\to 0}\Big(\int_{-\infty}^{-\varepsilon}\frac{1}{x}\varphi dx+\int_{\varepsilon}^{\infty}\varphi'\log|x|dx\Big)\\ &= \langle \mathrm{vp}\frac{1}{x},\varphi\rangle.\end{aligned}$$

例子. 我们计算 $\mathbb{R}$ 上的 *Dirac* 函数 δ_a 的导数，其中 $a\in\mathbb{R}$。任给 $\varphi\in\mathcal{D}(\mathbb{R})$，我们有

$$\langle\delta_a',\varphi\rangle=-\langle\delta_a,\varphi'\rangle=-\varphi'(a).$$

$C^\infty(\Omega)$-模结构

对光滑函数 $f\in C^\infty(\Omega)$ 和分布 $u\in\mathcal{D}'(\Omega)$，我们可以定义它们的乘积 $f\cdot u$:

$$\langle f\cdot u,\varphi\rangle:=\langle u,f\varphi\rangle.$$

其中，φ 是试验函数。由于 $f\varphi$ 仍然是 $\mathcal{D}'(\Omega)$ 中的函数，所以上面的等式是良好定义的。为了说明 $f\cdot u\in\mathcal{D}'(\Omega)$，要对 u 来用分布的定义：对任意的紧集 $K\subset\Omega$，存在非负整数 p 和正常数 C（p 和 C 依赖于 K），使得对任意的 $\varphi\in C_K^\infty(\Omega)$，都有

$$|\langle f\cdot u,\varphi\rangle|\leqslant C\sup_{|\alpha|\leqslant p}\|\partial^\alpha(f\cdot\varphi)\|_{L^\infty(K)}.$$

根据 Leibniz 法则，我们有

$$\partial^\alpha(f\cdot\varphi)=\sum_{\beta+\gamma=\alpha}\partial^\beta f\cdot\partial^\gamma\varphi.$$

所以，

$$\begin{aligned}|\langle f\cdot u,\varphi\rangle| &\leqslant C\sup_{|\alpha|\leqslant p}\sum_{\beta+\gamma=\alpha}\|\partial^\beta f\|_{L^\infty(K)}\|\partial^\gamma\varphi\|_{L^\infty(K)}\\ &\leqslant \underbrace{C\sum_{|\beta|\leqslant p}\|\partial^\beta f\|_{L^\infty(K)}}_{\text{新的常数 } C'}\times\sup_{|\gamma|\leqslant p}\|\partial^\gamma\varphi\|_{L^\infty(K)}.\end{aligned}$$

这表明 $f\cdot u\in\mathcal{D}'(\Omega)$。

例子. 我们有

$$x \cdot \mathrm{vp}\frac{1}{x} = 1.$$

对任意的 $\varphi \in \mathcal{D}(\mathbb{R})$，我们知道 $x\varphi(x)\big|_{x=0} = 0$。根据第一次作业题的 *A6)*，我们有

$$\begin{aligned}\langle x \cdot \mathrm{vp}\frac{1}{x}, \varphi\rangle &= \langle \mathrm{vp}\frac{1}{x}, x\varphi\rangle \\ \int_{\mathbb{R}} \frac{1}{x} \cdot x\varphi(x)dx &= \int_{\mathbb{R}} 1 \cdot \varphi(x)dx.\end{aligned}$$

这就证明了命题。

分布的平移和变量替换

对于 $x_0 \in \mathbb{R}^n$，我们有如下的平移变换：

$$\tau_{x_0} : \mathbb{R}^n \to \mathbb{R}^n, \quad x \mapsto x + x_0.$$

对于局部可积的函数 $f \in L^1_{\mathrm{loc}}(\mathbb{R}^n)$，我们可以定义（这是一种特殊的变量替换）：

$$\big(\tau_{x_0} f\big)(x) = \varphi(x + x_0).$$

如果将 f 视为是分布，对于试验函数 φ 而言，我们有

$$\begin{aligned}\langle f(x + x_0), \varphi(x - x_0)\rangle &= \int_{\mathbb{R}^n} f(x + x_0)\varphi(x)dx = \int_{\mathbb{R}^n} f(x)\varphi(x - x_0)dx \\ &= \langle f(x), \varphi(x - x_0)\rangle = \langle f(x), \tau_{-x_0}\varphi\rangle.\end{aligned}$$

对于一般的分布 $u \in \mathcal{D}'(\mathbb{R}^n)$，$x_0 \in \mathbb{R}^n$，我们定义

$$\langle \tau_{x_0} u, \varphi\rangle := \langle u, \varphi(x - x_0)\rangle.$$

容易验证，$\tau_{x_0} u$ 给出了一个分布。实际上，我们稍后在进行变量替换的时候，也会给出这个命题的证明。

下面的命题给出了在分布意义下方向导数的另一个（直观）刻画：

命题 395. 给定分布 $u \in \mathcal{D}'(\mathbb{R}^n)$，给定向量 $v = (v_1, \cdots, v_n) \in \mathbb{R}^n$，在分布意义下，我们有

$$\lim_{t\to 0} \frac{\tau_{tv} u - u}{t} \overset{\mathcal{D}'}{=} \sum_{j=1}^n v_j \partial_j u.$$

证明： 按照定义，我们有

$$\langle \frac{\tau_{tv} u - u}{t} - \sum_{j=1}^n v_j \partial_j u, \varphi\rangle = \langle u, \underbrace{\frac{\varphi(x - tv) - \varphi(x)}{t} + \sum_{j=1}^n a_j \partial_j \varphi(x)}_{\varphi_t}\rangle.$$

根据带有积分余项的 Taylor 展开，我们有

$$\varphi_t(x) = -\sum_{j=1}^{n} \int_0^1 v_j \Big(\partial_j \varphi(x - tsv) - \partial_j \varphi(x) \Big) ds.$$

由于 t 很小，上述函数 $\varphi_t(x)$ 的支撑集是紧的，从而第一个等式的右边是良好定义的。我们来证明在 $\mathcal{D}(\mathbb{R}^n)$ 中，上述 $\varphi_t(x)$ 的极限是 0，也就是说对任意的多重指标 α，我们有 $\partial^\alpha \varphi_t(x) \xrightarrow{L^\infty} 0$：根据 Lebesgue 控制收敛定理（的推论，积分与求导数可交换，请参考上学期笔记），我们有

$$\begin{aligned} \|\partial^\alpha \varphi_t(x)\|_{L^\infty} &= \Big\| \sum_{j=1}^{n} \int_0^1 v_j \Big(\partial_j \partial^\alpha \varphi(x - ts\frac{v}{|v|}) - \partial_j \partial^\alpha \varphi(x) \Big) ds \Big\|_{L^\infty} \\ &\leqslant \Big\| \sum_{j=1}^{n} \int_0^1 |v_j| \|\partial^{|\alpha|+2} \varphi\|_\infty t ds \Big\|_{L^\infty}. \end{aligned}$$

其中，最后一步我们用到了 Lagrange 中值定理。再次利用 Lebesgue 控制收敛定理，上面的极限为 0。命题得证。 □

给定 $\mathbb{R}^n$ 的两个开集 Ω_1 和 Ω_2，我们假定

$$\Phi : \Omega_1 \to \Omega_2$$

是微分同胚。对任意一个 Ω_2 上的局部可积的函数 f 和 Ω_1 上的试验函数 φ，根据换元积分公式，我们有

$$\begin{aligned} \langle \Phi^* f, \varphi(x) \rangle &= \int_{\Omega_1} f(\Phi(x)) \varphi(x) dx = \int_{\Omega_2} f(y) \varphi(\Phi^{-1}(y)) |\mathbf{J}_{\Phi^{-1}}(y)| dy \\ &= \int_{\Omega_2} f(y) \frac{\varphi(\Phi^{-1}(y))}{|\mathbf{J}_\Phi(\Phi^{-1}(y))|} dy \\ &= \left\langle f, \frac{\varphi(\Phi^{-1}(y))}{|\mathbf{J}_\Phi(\Phi^{-1}(y))|} \right\rangle. \end{aligned}$$

根据这个计算，我们定义：

$$\Phi^* : \mathcal{D}'(\Omega_2) \to \mathcal{D}'(\Omega_1), \quad u \mapsto \Phi^* u.$$

其中，对于 $u \in \mathcal{D}'(\Omega_2)$ 和 $\varphi \in \mathcal{D}(\Omega_1)$，对于我们定义 $\Phi^* u$ 如下：

$$\langle \Phi^* u, \varphi(x) \rangle := \langle u, \varphi(\Phi^{-1}(y)) |\mathbf{J}_{\Phi^{-1}}(y)| \rangle = \left\langle u, \frac{\varphi(\Phi^{-1}(y))}{|\mathbf{J}_\Phi(\Phi^{-1}(y))|} \right\rangle.$$

为了证明 $\Phi^* u$ 的确定义了 Ω_1 上的一个分布，我们利用定义：对任意的紧集 $K \subset \Omega$，存在非负整数 p 和正常数 C（p 和 C 依赖于 K），使得对任意的 $\varphi \in C_K^\infty(\Omega)$，都有

$$|\langle \Phi^* u, \varphi \rangle| \leqslant C \sup_{|\alpha| \leqslant p} \left\| \partial^\alpha \left(\frac{\varphi(\Phi^{-1}(y))}{|\mathbf{J}_\Phi(\Phi^{-1}(y))|} \right) \right\|_{L^\infty(K)}.$$

我们需要利用链式法则多次求导数。重要的观察是上述导数最重给出 φ 的不超过 p 次的导数的线性组合，其中，这些线性系数都是 $\Phi(y)$ 的不超过 $p+1$ 阶的导数的多项式。由于我们限制在紧集 K 上，所以这些系数是一致有界的（这个界可能依赖于 K），从而，通过改变一下上面不等式中的常数 C，我们最终得到

$$|\langle \Phi^* u, \varphi\rangle| \leqslant C' \sup_{|\alpha|\leqslant p} \|\partial^\alpha \varphi\|_{L^\infty(K)}.$$

注记 (记号)**.** 因为在光滑函数情况下，$\Phi^* u$ 就是函数的复合，我们还把上面的拉回映射写成

$$u \circ \Phi = \Phi^* u.$$

给定微分同胚 $\Phi : \Omega_1 \to \Omega_2$，它把 Ω_2 上的 Dirac 函数拉回，得到 Ω_1 上的 Dirac 函数，这给出了 Jacobi 行列式的一个精确的解释：这是一点处体积的变化。

例子. 假设 $x_0 \in \Omega_1$，$y_0 \in \Omega_2$ 并且 $\Phi(x_0) = y_0$。那么，我们有

$$\Phi^* \delta_{y_0} = \frac{1}{|\mathbf{J}_\Phi(x_0)|} \delta_{x_0}.$$

实际上，利用定义，对于任意的 $\varphi \in \mathcal{D}(\Omega_1)$，我们有

$$\langle \Phi^* \delta_{y_0}, \varphi(x)\rangle = \left\langle \delta_{y_0}, \frac{\varphi(\Phi^{-1}(y))}{|\mathbf{J}_\Phi\big(\Phi^{-1}(y)\big)|} \right\rangle = \frac{\varphi(\Phi^{-1}(y_0))}{|\mathbf{J}_\Phi\big(\Phi^{-1}(y_0)\big)|} = \frac{\varphi(x_0)}{|\mathbf{J}_\Phi(x_0)|}.$$

很明显，

$$\frac{\varphi(x_0)}{|\mathbf{J}_\Phi(x_0)|} = \left\langle \frac{1}{|\mathbf{J}_\Phi(x_0)|} \delta_{x_0}, \varphi \right\rangle.$$

给定微分同胚 Φ，如果 u 是 C^1 的函数，我们可以对求导数运算运用链式法则:

$$\partial_j (u \circ \Phi) = \sum_{k=1}^n \partial_j \Phi_k(x_1, \cdots, x_n) \cdot \big(\partial_k u \circ \Phi\big)(x_1, \cdots, x_n).$$

对于一般的分布 u，我们实际上（后来）可以先用光滑函数逼近这个分布，然后上面的链式法则在极限的情况下仍然成立。

我们现在给出一个直接的证明：设 $\Psi = \Phi^{-1}$ 是 Φ 的逆映射。按照分布与一个微分同胚的复合的定义，我们有

$$\begin{aligned}
\left\langle \sum_{k=1}^n \frac{\partial \Phi_k}{\partial x_j} \cdot \frac{\partial u}{\partial y_k} \circ \Phi, \varphi \right\rangle &= \sum_{k=1}^n \left\langle \frac{\partial u}{\partial y_k} \circ \Phi, \frac{\partial \Phi_k}{\partial x_j} \cdot \varphi \right\rangle \\
&= \sum_{k=1}^n \left\langle \frac{\partial u}{\partial y_k}, \Big(\frac{\partial \Phi_k}{\partial x_j} \circ \Psi\Big)(\varphi \circ \Psi)|J_\Psi| \right\rangle \\
&= -\sum_{k=1}^n \left\langle u, \frac{\partial}{\partial y_k}\Big[\Big(\frac{\partial \Phi_k}{\partial x_j} \circ \Psi\Big)(\varphi \circ \Psi)|J_\Psi|\Big] \right\rangle
\end{aligned}$$

另外，对任意 $g \in C_0^\infty(\Omega)$，我们有

$$\begin{aligned} 0 &= \int_{\Omega_1} \frac{\partial}{\partial x_j}(g \circ \Phi) dx = \sum_{k=1}^n \int_{\Omega_1} \frac{\partial \Phi_k}{\partial x_j} \cdot \frac{\partial g}{\partial y_k} \circ \Phi \\ &= \sum_{k=1}^n \int_{\Omega_2} \frac{\partial \Phi_k}{\partial x_j} \circ \Psi \cdot \frac{\partial g}{\partial y_k} \cdot |\mathbf{J}_\Psi(y)| dy \\ &= -\sum_{k=1}^n \int_{\Omega_2} g \frac{\partial}{\partial y_k}\Big(\frac{\partial \Phi_k}{\partial x_j} \circ \Psi \cdot |\mathbf{J}_\Psi(y)|\Big) dy \end{aligned}$$

根据 $L^1_{\mathrm{loc}}(\Omega)$ 到 $\mathcal{D}'(\Omega)$ 嵌入的单射性，上面的等式等价于说

$$\sum_{k=1}^n \frac{\partial}{\partial y_k}\Big(\frac{\partial \Phi_k}{\partial x_j} \circ \Psi \cdot |\mathbf{J}_\Psi(y)|\Big) = 0.$$

从而

$$\begin{aligned} \left\langle \sum_{k=1}^n \frac{\partial \Phi_k}{\partial x_j} \cdot \frac{\partial u}{\partial y_k} \circ \Phi, \varphi \right\rangle &= -\sum_{k=1}^n \left\langle u, \frac{\partial}{\partial y_k}(\varphi \circ \Psi)\Big[\Big(\frac{\partial \Phi_k}{\partial x_j} \circ \Psi\Big)|\mathbf{J}_\Psi|\Big] \right\rangle \\ &= \left\langle u, \frac{\partial \varphi}{\partial x_j} \circ \Psi |\mathbf{J}_\Psi| \right\rangle \\ &= -\left\langle u \circ \Phi, \frac{\partial \varphi}{\partial x_j} \right\rangle. \end{aligned}$$

这就证明如下关于分布的链式法则：

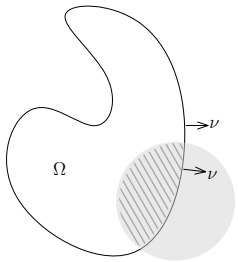
$$\partial_j(\Phi^* u) = \sum_{k=1}^n \partial_j \Phi_k \cdot \Phi^*\left((\partial_k u)\right).$$

关于分布的 Stokes 公式

我们用分布的语言来表述 Stokes 公式。我们上个学期证明了：

定理 396 (Stokes 公式)**.** *假设 Ω 是一个有界带边光滑区域，$\nu(x) = (\nu_1(x), \cdots, \nu_n(x))$ 为 $\partial\Omega$ 的单位外法向量，$d\sigma$ 为 $\partial\Omega$ 上的曲面测度。对任意的 $\varphi \in C^1(\mathbb{R}^n, \mathbb{C})$，我们有*

$$\int_\Omega \frac{\partial \varphi}{\partial x_i}(x)\ dx = \int_{\partial\Omega} \varphi(x)\nu_i(x)\ d\sigma.$$



给定了上面 Stokes 公式中所述的 Ω，它的边界 $\partial\Omega$ 的曲面测度 $d\sigma$ 在如下的意义下定义了 $\mathbb{R}^n$ 上一个分布：

$$d\sigma : \mathcal{D}(\mathbb{R}^n) \to \mathbb{C}, \quad \varphi \mapsto \int_{\partial\Omega} \varphi(x) d\sigma(x).$$

这是一个 0 阶的分布，我们把证明的细节留给不放心的同学来验证。类似地，对每一个 $i \leqslant n$，如下的公式也定义了一个分布：

$$\nu_i d\sigma : \mathcal{D}(\mathbb{R}^n) \to \mathbb{C}, \quad \varphi \mapsto \int_{\partial\Omega} \varphi(x)\nu_i(x) d\sigma(x).$$

我们可以把 Stokes 公式改写成如下的形式：

$$\langle \mathbf{1}_\Omega, \partial_i \varphi \rangle = \int_{\mathbb{R}^n} \mathbf{1}_\Omega(x) \frac{\partial \varphi}{\partial x_i}(x)\ dx = \int_{\partial\Omega} \varphi(x)\nu_i(x)\ d\sigma.$$

所以，用分布的语言来写，我们有

定理 397 (Stokes 公式)**.** *假设 Ω 是一个有界带边光滑区域，$\nu(x) = (\nu_1(x), \cdots, \nu_n(x))$ 为 $\partial\Omega$ 的单位外法向量，$d\sigma$ 为 $\partial\Omega$ 上的曲面测度，作为分布，我们有等式*

$$\partial_i \mathbf{1}_\Omega \overset{\mathcal{D}'(\mathbb{R}^n)}{=} -\nu_i d\sigma.$$

如果用向量值分布的语言（可以望文生义地定义）来写，我们有

$$\nabla \mathbf{1}_\Omega \overset{\mathcal{D}'(\mathbb{R}^n)}{=} -\nu\, d\sigma.$$

我们现在回到 1 维的情形，此时的 Stokes 公式就是 Newton-Leibniz 公式。

引理 398. *对于 $f(x) \in L^1\big((a,b)\big)$，我们定义其原函数为*

$$F(x) = \int_a^x f(y)dy.$$

那么，$F(x)$ 是连续函数。在分布的意义下，我们有

$$F(x)' \overset{\mathcal{D}'}{=} f(x).$$

证明: 为了证明 F 是连续的，我们把它写成

$$F(x) = \int_{[a,b]} f \cdot \mathbf{1}_{(a,x]} d\mu,$$

其中 $d\mu$ 是 Lebesgue 测度。对任意的 $x_k \to x$，我们知道 $f \cdot \mathbf{1}_{(a,x_k]}$ 逐点地收敛到 $f \cdot \mathbf{1}_{(a,x]}$，利用 $|f|$ 作为控制函数，Lebesgue 控制收敛定理告诉我们

$$\lim_{k\to\infty} F(x_k) = F(x).$$

我们用定义计算 F'。对于任意的试验函数 $\varphi \in \mathcal{D}\big((a,b)\big)$，我们有

$$\begin{aligned}\langle F', \varphi\rangle = -\langle F, \varphi'\rangle &= \int_{(a,b)} \left(\int_{(a,b)} f \cdot \mathbf{1}_{(a,x]}(y) dy \right) \varphi'(x) dx \\ &= \iint_{(a,b)\times(a,b)} \mathbf{1}_A(x,y) f(y) \varphi'(x) dx dy.\end{aligned}$$

这里，$A = \big\{(x,y) \in (a,b)\times(a,b) \big| y \leqslant x\big\}$ 并且我们是利用了 Fubini 定理把它化成了 2 维的积分。再次利用 Fubini 定理，我们先对 x 积分再对 y 积分，就有

$$\begin{aligned}\langle F', \varphi\rangle &= \int_{(a,b)} f(y) \left(\int_a^y \varphi'(x) dx \right) dy \\ &= \int_a^b f(y)\varphi(y) dy = \langle T_f, \varphi\rangle.\end{aligned}$$

其中，我们用到了 $\varphi(a) = 0$（因为 $\operatorname{supp}(\varphi) \subset (a,b)$）。这就是说，在分布的意义下，$F' = f$。 □

60 1 维的跳跃公式，Cauchy 积分公式，Cauchy-Riemann 算子的基本解，单位分解

二零二零年九月二十一日，星期一，晴

上次课的最后利用 Fubini 定理，我们证明了：对于 $f(x)\in L^1\big((a,b)\big)$，我们定义其原函数为

$$F(x)=\int_a^x f(y)dy.$$

那么，$F(x)$ 是连续函数。在分布的意义下，我们有

$$F(x)'\overset{\mathcal{D}'}{=}f(x).$$

引理 399. *给定分布 $u\in\mathcal{D}'\big((a,b)\big)$，如果在分布的意义下 $u'=0$，那么，在分布的意义下，u 为常数，即存在 $c\in\mathbb{C}$，使得*

$$u\overset{\mathcal{D}'}{=}c.$$

注记. *请参考第一学期第 15 次课的推论 94。*

证明: 我们任意选定一个 $\chi\in C_0^\infty\big((a,b)\big)$，使得 $\int_a^b\chi(x)dx=1$。令 $c=\langle u,\chi\rangle$。

对于任意一个试验函数 $\varphi\in\mathcal{D}\big((a,b)\big)$，我们定义

$$g(x)=\varphi(x)-\big(\int_{-\infty}^{\infty}\varphi(y)dy\big)\chi(x),$$

那么，我们有

$$\int_{\mathbb{R}}g(x)dx=0.$$

我们定义函数

$$\psi(x)=\int_a^x g(y)dy.$$

利用 g 的积分消失的性质，我们容易说明（证明支集是紧的）$\psi\in\mathcal{D}\big((a,b)\big)$。特别地，我们可以用 ψ 作为试验函数。由于 $u'=0$，所以

$$\begin{aligned}0&=-\langle u,\psi'\rangle=-\langle u,\varphi(x)-\big(\int_{\mathbb{R}}\varphi(y)dy\big)\chi(x)\rangle\\&=\langle u,\varphi\rangle-c\int_{\mathbb{R}}\varphi(x)dx.\end{aligned}$$

所以，

$$\langle u,\varphi\rangle=\int_{\mathbb{R}}c\cdot\varphi(x)dx=\langle c,\varphi\rangle.$$

这说明，作为分布，我们有 $u=c$。 □

综合上面的两个命题，我们有如下的性质：

命题 400. 假定 $f \in L^1\big((a,b)\big)$，那么关于分布的常微分方程

$$u' = f$$

在 $\mathcal{D}'\big((a,b)\big)$ 中所有的解均形如

$$u = c + \int_a^x f(y)dy,$$

其中 $c \in \mathbb{C}$ 为常数。

我们现在来证明所谓的跳跃公式：

定理 401 (跳跃公式)**.** 给定连续函数 $g \in C(\mathbb{R})$，我们假设它的分布导数 $g' \in L^1_{loc}\big(\mathbb{R}\big)$。那么，对任意的 $-\infty < a < b < +\infty$，在分布的意义下，我们有

$$\frac{d}{dx}\big(g(x)\mathbf{1}_{(a,b)}(x)\big) \overset{\mathcal{D}'}{=} g'(x)\mathbf{1}_{(a,b)}(x) - g(b)\delta_b + g(a)\delta_a.$$

证明: 根据上一个命题，我们有（作为分布或者连续函数）

$$g(x) = g(a) + \int_a^x g'(y)dy.$$

按定义，对任意的试验函数 φ，我们还有

$$\begin{aligned}
\left\langle \frac{d}{dx}\big(g(x)\mathbf{1}_{(a,b)}(x)\big), \varphi \right\rangle &= -\int_a^b g(x)\varphi'(x)dx \\
&= -g(a)\int_a^b \varphi'(x)dx + \int_a^b \left[\int_a^x g'(y)dy\right](-\varphi'(x))dx \\
&\overset{Fubini}{=} -g(a)\int_a^b \varphi'(x)dx - \int_a^b \left[\int_y^b \varphi'(x)dx\right] g'(y)dy \\
&= -g(a)\underbrace{\int_a^b \varphi'(x)dx}_{\varphi(b)-\varphi(a)} - \varphi(b)\underbrace{\int_a^b g'(x)dx}_{g(b)-g(a)} + \int_a^b \varphi(y)g'(y)dy.
\end{aligned}$$

将上面最后一行合并同类项即得。 □

分布与 Stokes 理论的第一个应用：Cauchy 积分公式

考虑复平面 $\mathbb{C}$ 上的开区域 Ω。函数（映射）F 在 Ω 上定义并且在 $\mathbb{C}$ 中取值：

$$F : \Omega \to \mathbb{C}$$

我们假设 F 连续微的（C^1）。

如果把 F 视作是从 Ω 到 $\mathbb{R}^2$ 的映射，我们通常将它写作

$$F(z) = F(x,y) = f(x,y) + ig(x,y).$$

我们定义微分算子:

$$\overline{\partial} = \frac{\partial}{\partial \overline{z}} = \frac{1}{2}\left(\frac{\partial}{\partial x} + i\frac{\partial}{\partial y}\right).$$

所以，$\overline{\partial}$ 对 F 的作用为:

$$\overline{\partial} F = \frac{1}{2}(\frac{\partial f}{\partial x} - \frac{\partial g}{\partial y}) + \frac{i}{2}(\frac{\partial g}{\partial x} + \frac{\partial f}{\partial y}).$$

我们现在定义 F在复解析意义下的导数（如果极限存在的话）:

$$F'(z_0) = \lim_{z \to z_0} \frac{F(z) - F(z_0)}{z - z_0}.$$

其中，$z_0 \in \Omega$ 是给定的一点。

定义 402. 如果 $F'(z)$ 在 Ω 上处处有定义并且是连续函数，我们就成称 F 为 Ω 上的**复解析函数**。

我们现在证明

定理 403. 函数 $F : \Omega \to \mathbb{C}$ 是复解析函数当且仅当 $\overline{\partial} F = 0$。特别地，$F(x, y) = f(x, y) + ig(x, y)$ 在 Ω 上是复解析的当且仅当如下的偏微分方程组在 Ω 上成立

$$\begin{cases} \dfrac{\partial f}{\partial x} - \dfrac{\partial g}{\partial y} = 0; \\ \dfrac{\partial g}{\partial x} + \dfrac{\partial f}{\partial y} = 0. \end{cases}$$

这个方程组是所谓的 *Cauchy-Riemann* 方程。

证明: 假设 F 在 z_0 处的复解析导数可定义，我们另它为

$$F'(z_0) = \lim_{z \to z_0} \frac{F(z) - F(z_0)}{z - z_0} = a + bi.$$

根据极限的定义，当 $z \to z_0$ 时，我们有

$$F(z) - F(z_0) - F'(z_0)(z - z_0) = o(|z - z_0|)$$

把 F 看作是从 Ω 到 $\mathbb{R}^2$ 的映射，用映射微分的语言来写（参考上学期第一部分内容），我们有

$$F(x, y) - F(x_0, y_0) - \begin{pmatrix} a & -b \\ b & a \end{pmatrix} \begin{pmatrix} x - x_0 \\ y - y_0 \end{pmatrix} = o(|x - x_0| + |y - y_0|),$$

其中 $z = x + y\sqrt{-1}$，$z_0 = x_0 + y_0\sqrt{-1}$。这表明矩阵 $\begin{pmatrix} a & -b \\ b & a \end{pmatrix}$ 是映射 F 在 (x_0, y_0) 处的微分，从而是这个映射在这个点处的 Jacobi 矩阵。所以，

$$\begin{pmatrix} a & -b \\ b & a \end{pmatrix} = \text{Jacobi 矩阵} = \begin{pmatrix} \frac{\partial f}{\partial x} & \frac{\partial f}{\partial y} \\ \frac{\partial g}{\partial x} & \frac{\partial g}{\partial y} \end{pmatrix}.$$

比较系数，我们有

$$\frac{\partial f}{\partial x} = a = \frac{\partial g}{\partial y} = 0, \quad \frac{\partial g}{\partial x} = b = -\frac{\partial f}{\partial y} = 0.$$

所以，Cauchy-Riemann 方程成立。

反之，上述计算表明我们只要取

$$F'(z_0) = \frac{\partial f}{\partial x}(x_0, y_0) + i\frac{\partial g}{\partial x}(x_0, y_0)$$

作为极限

$$\lim_{z \to z_0} \frac{F(z) - F(z_0)}{z - z_0}$$

的取值即可。 □

我们现在研究 Cauchy-Riemann 方程的的基本解，关于基本解这个概念我们后面会进一步阐明。我们考虑在 $\mathbb{C}$ 上几乎处处定义的映射

$$\mathbb{C} - \{0\}, \quad z \mapsto \frac{1}{2}.$$

我们把这个复值函数就记作 z^{-1} 或者 $\frac{1}{z}$。由于 $\left|\frac{1}{z}\right| \leqslant \frac{1}{r}$，其中，$r$ 为平面上的极坐标系中的半径函数，根据换元积分公式，我们很容易看出

$$\frac{1}{z} \in L^1_{\mathrm{loc}}(\mathbb{C}).$$

引理 404. *将 $\frac{1}{z}$ 视作是 $\mathbb{C} = \mathbb{R}^2$ 上的分布。那么，在分布的意义下，我们有*

$$\overline{\partial}\left(\frac{1}{\pi z}\right) \overset{\mathcal{D}'}{=} \delta_0.$$

特别地，对于任意的 $z_0 \in \mathbb{C}$，我们有

$$\overline{\partial}(\frac{1}{\pi(z - z_0)}) = \delta_{z_0}.$$

用偏微分微分方程的语言来描述，$\frac{1}{\pi z}$ 为算子 $\overline{\partial}$ 的基本解。

在证明之前，我们先陈述两个关于分布收敛的基本事实，它们的证明留作的作业（请参考第 2 次作业）：

- 假设 $\{f_p\}_{p \geqslant 1}$ 是 Ω 上的一列局部可积的函数，$f \in L^1_{\mathrm{loc}}(\Omega)$。如果对任意的紧集 $K \subset \Omega$，我们有

  $$\lim_{p \to} f_p\big|_K \overset{L^1(K)}{=} f\big|_K,$$

  那么，作为分布，我们有

  $$\lim_{p \to} f_p \overset{\mathcal{D}'}{=} f.$$

- 假设 $\{u_p\}_{p\geqslant 1}$ 是 Ω 上分布的序列，$u\in\mathcal{D}'(\Omega)$。如果

$$\lim_{p\to} u_p \overset{\mathcal{D}'}{=} u.$$

那么，对任意的多重指标 α，在分布的意义下，我们有

$$\lim_{p\to} \partial^\alpha u_p \overset{\mathcal{D}'}{=} \partial^\alpha u.$$

另外，我们还有

- 对于任意的 $f\in C^\infty(\Omega)$ 和 $u\in\mathcal{D}'(\Omega)$，对任意的 $k\leqslant n$，我们有

$$\partial_k(f\cdot u)\overset{\mathcal{D}'}{=}\partial_k f\cdot u+f\cdot\partial_k u.$$

证明也留作习题。

证明： 为此，我们定义

$$f_\varepsilon(z)=\begin{cases}\dfrac{1}{\pi z},\ |z|\geqslant\varepsilon;\\[2ex] \dfrac{\overline{z}}{\pi\varepsilon^2},\ |z|<\varepsilon.\end{cases}$$

所以，

$$g_\varepsilon(z)=f_\varepsilon(z)-\frac{1}{\pi z}=\mathbf{1}_{|z|<\varepsilon}(z)\cdot\left(\frac{|z|^2-\varepsilon^2}{\pi\varepsilon^2 z}\right).$$

当 $\varepsilon\to 0$ 时，g_ε 是逐点（几乎处处）收敛到 0 的；我们还可以用 $\dfrac{2}{\pi|z|}\mathbf{1}_{|z|\leqslant 1}(z)$ 作为 $\varepsilon\leqslant 1$ 时的控制函数。所以，根据 Lebesgue 控制收敛定理，我们有

$$g_\varepsilon(z)\xrightarrow{L^1}0,\quad \varepsilon\to 0.$$

所以，

$$f_\varepsilon\xrightarrow{\mathcal{D}'}\frac{1}{\pi z},\quad \varepsilon\to 0.$$

从而，

$$\overline{\partial}f_\varepsilon\xrightarrow{\mathcal{D}'}\overline{\partial}\left(\frac{1}{\pi z}\right),\quad \varepsilon\to 0.$$

现在我们来计算 $\overline{\partial}f_\varepsilon$。我们把 f_ε 写作

$$f_\varepsilon(z)=\frac{\overline{z}}{\pi\varepsilon^2}\cdot\mathbf{1}_{|z|\leqslant\varepsilon}(z)+\frac{1}{\pi z}\cdot\mathbf{1}_{|z|\geqslant\varepsilon}(z).$$

根据分布意义下的 Stoke 公式，我们有

$$\begin{aligned}\overline{\partial}f_\varepsilon(z)&=\frac{1}{\pi\varepsilon^2}\cdot\mathbf{1}_{|z|\leqslant\varepsilon}(z)-\frac{\overline{z}}{\pi\varepsilon^2}\cdot d\sigma_{|z|=\varepsilon}+\frac{1}{\pi z}\cdot d\sigma_{|z|=\varepsilon}\\&=\frac{1}{\pi\varepsilon^2}\cdot\mathbf{1}_{|z|\leqslant\varepsilon}(z).\end{aligned}$$

其中，我们用到了

$$\overline{\partial}(\overline{z}) = 1, \ \ \overline{\partial}\left(\frac{1}{z}\right) = 0, z \neq 0.$$

我们把这两个计算留作习题。我们现在对

$$\overline{\partial} f_\varepsilon(z) = \frac{1}{\pi\varepsilon^2}\mathbf{1}_{|z|\leqslant\varepsilon}(z),$$

取极限。我们发现这就是

$$\frac{1}{\varepsilon^2}h\left(\frac{1}{\varepsilon}\right), \ \ h(z) = \frac{1}{\pi}\mathbf{1}_{|z|\leqslant 1}(z),$$

的极限，其中 $h \in L^1(\mathbb{R}^2)$ 并且积分为 1。根据我们之前已经学过的例子，我们有

$$\frac{1}{\pi\varepsilon^2}\mathbf{1}_{|z|\leqslant\varepsilon}(z) \xrightarrow{\mathcal{D}'} \delta_0, \ \ \varepsilon \to 0.$$

这就证明了引理。 □

这个公式是所谓的 Cauchy 积分公式在一个点处的情况。为了叙述 Cauchy 积分公式，我们先讨论一下复变量（复值）函数在曲线上的积分。

假定 $\gamma : [0, a] \to \mathbb{C}$ 是一个分段光滑（C^1）的连续曲线，也就是说，存在 $t_0 = 0 < t_1 < t_2 < \cdots < t_{n-1} < a = t_n$，使得 $\gamma : (t_{k-1}, t_k) \to \mathbb{C}$ 是 C^1 的。在每段上面，我们有

$$s \mapsto \gamma(s) = x(s) + iy(s).$$

对于复函数 $F(z)$，我们定义

$$\begin{aligned}\int_\gamma F(z)dz &= \sum_{k=1}^n \int_{t_{k-1}}^{t_k} F(\gamma(s))\gamma'(s)ds \\ &= \sum_{k=1}^n \Big(\int_{t_{k-1}}^{t_k} F(\gamma(s))x'(s)ds + i\int_{t_{k-1}}^{t_k} F(\gamma(s))y'(s)ds \Big).\end{aligned}$$

我们要区分一下这个定义与我们上学期定义的曲线积分的关系。实际上，如果我们假设 $\gamma'(s) \neq 0$，当我们取 s 为弧长参数的时候，我们有

$$\int_\gamma F(z)d\sigma = \int_\gamma F(\gamma(s))|\gamma'(s)|ds.$$

所以，这里的区别在于考虑的曲线的方向：比如说，我们考虑另一条曲线 γ^{-1}，它是 γ 沿着反方向来走，即

$$\gamma^{-1} : [0, a] \to \mathbb{C}, \ \ s \mapsto \gamma(a - s).$$

那么，

$$\int_\gamma F(z)dz = -\int_{\gamma^{-1}} F(z)dz.$$

例子. 作为例子，我们来计算

$$\frac{1}{2\pi i}\int_{|z|=r_0} z^k dz, \quad z\in\mathbb{Z}, r_0>0.$$

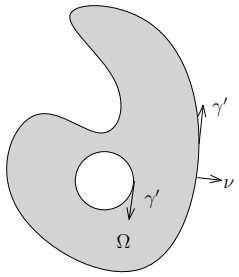
我们用

$$[0,2\pi]\to\mathbb{C}, \quad \vartheta\mapsto r_0 e^{i\vartheta},$$

来参数化 $|z|=r_0$，所以，我们有

$$\begin{aligned}\frac{1}{2\pi i}\int_{|z|=r_0} z^k dz &= \frac{1}{2\pi i}\int_0^{2\pi} (r_0)^k e^{ik\vartheta}(-r_0\sin\vartheta + r_0 i\cos\vartheta)d\vartheta\\ &= \frac{1}{2\pi}\int_0^{2\pi}(r_0)^{k+1}e^{i(k\vartheta+1)}d\vartheta\\ &= \begin{cases}1, & k=-1;\\ 0, & k\neq -1.\end{cases}\end{aligned}$$

我们可以对一个复函数用 Green 公式。假定紧集 Ω 的边界是 γ，其中我们是按照逆时针或者顺时针来标记其方向的，但是要求区域要在这个切向量的左手边。



我们用弧长参数来参数化边界曲线。我们有

$$\begin{aligned}\int_\Omega \overline{\partial}F(z)dxdy &= \frac{1}{2}\int_\Omega \overbrace{\frac{\partial}{\partial x}F(z)+\frac{\partial}{\partial y}(iF(z))}^{\text{散度形式}}dxdy\\ &= \frac{1}{2}\int_\gamma (F,iF)\cdot \underbrace{(-i)(\gamma_x'+i\gamma_y')}_{\text{外法向量}}d\sigma\\ &= \frac{1}{2}\int_\gamma (F,iF)\cdot(\gamma_y',-\gamma_x')d\sigma\\ &= \frac{1}{2i}\int_\gamma F(z)dz.\end{aligned}$$

我们现在来证明 Cauchy 积分公式：

定理 405 (Cauchy 积分公式)**.** 假设 $\Omega\subset\mathbb{C}$ 是开集，$K\subset\Omega$ 是有界带边区域（特别地，K 是紧的），其边界 $\gamma=\partial K$ 是 C^1 曲线（可以有多个连通分支）。$F(z)$ 是 Ω 上的复解析函数。那么，对于 $z_0\in\mathring{K}$（K 的内部），我们有

$$F(z_0)=\frac{1}{2\pi i}\int_\gamma \frac{F(z)}{z-z_0}dz.$$

证明: 选取一个支集在 z_0 附近小领域中的试验函数 θ，要求在 θ 在 z_0 的一个邻域内恒为 1 并且 $\mathrm{supp}(\theta)\subset \mathring{K}$。利用上面的 Stokes 公式，我们有

$$\frac{1}{2i}\int_\gamma \frac{(1-\theta(z))F(z)}{z-z_0}dz=\int_K \overline{\partial}\Big(\big(1-\theta(z)\big)\frac{F(z)}{z-z_0}\Big)dxdy.$$

由于 $\mathrm{supp}(\theta)\subset \mathring{K}$，所以 $1-\theta$ 在 γ 上恒等于 1，从而

$$\frac{1}{2i}\int_\gamma \frac{F(z)}{z-z_0}dz=\int_K \overline{\partial}\Big(\big(1-\theta(z)\big)\frac{F(z)}{z-z_0}\Big)dxdy.$$

根据 Leibiniz 公式，我们就有

$$\begin{aligned}\frac{1}{2i}\int_\gamma \frac{F(z)}{z-z_0}dz&=\int_K -\overline{\partial}\theta(z)\frac{F(z)}{z-z_0}dxdy+\int_K (1-\theta(z))\underbrace{\overline{\partial}\left(\frac{F(z)}{z-z_0}\right)}_{\equiv 0}dxdy\\&=-\left\langle \frac{F(z)}{z-z_0},\overline{\partial}\theta(z)\right\rangle=\left\langle \overline{\partial}\big(\frac{F(z)}{z-z_0}\big),\theta(z)\right\rangle\\&=\pi\langle F(z_0)\delta_{z_0},\theta(z)\rangle=\pi F(z_0).\end{aligned}$$

这就给出了证明。 □

分布的局部刻画

我们来说明分布是**局部**上可定义的数学对象。为此，我们先回忆一下所谓的单位分解。

定理 406 (单位分解). *任意给定 $\mathbb{R}^n$ 中的紧集 K，假设 K 被有限个开集 $\{U_1,\cdots,U_N\}$ 所覆盖。那么，对每个 $j\leqslant N$，存在光滑函数 $\chi_j\in C_0^\infty(U_j)$，满足*

1) 对任意 $x\in\mathbb{R}^n$，有 $0\leqslant\chi_j(x)\leqslant 1$;

2) 存在包含 K 的开集 V，对任意 $x\in V$，我们有

$$\chi_1(x)+\cdots+\chi_N(x)=1.$$

为了证明单位分解定理，我们先证明如下的引理:

引理 407. *假设 $\Omega\subset\mathbb{R}^n$ 是开集，$K\subset\Omega$ 是紧集，那么，存在 $\varphi\in C_0^\infty(\Omega)$ 和开集 V，使得*

1) $K\subset V\subset U$;

2) 对任意的 $x\in\Omega$，$0\leqslant\varphi_i(x)\leqslant 1$;

3) $\varphi\big|_V\equiv 1$。

证明梗概. 首先，我们可以选去 $\delta > 0$，使得

$$K_{3\delta} = \left\{ x \in \mathbb{R}^n \middle| |x-k| < 2\delta, \text{ 对任意的 } k \in K \right\} \subset U.$$

令 $\phi(x) = \mathbf{1}_{K_{2\delta}}(x)$。我们选取我们常用的 $\chi(x)$，其中，我们要求它的积分为 1。再令

$$\chi_\varepsilon(x) = \frac{1}{\varepsilon^n} \chi\left(\frac{x}{\varepsilon}\right).$$

那么，我们可以选取

$$\varphi = \phi * \chi_\varepsilon, \quad V = K_\delta,$$

其中 $\varepsilon < \delta$。证明的细节留作作业。 □

单位分解的证明梗概. 对于每个开集 U_i，我们可以选取紧集 $K_i \subset U_i$，使得 $\bigcup_{i \leqslant N} K_i$ 仍然包含 K。此时，我们对每个 $K_i \subset U_i$ 运用上面的引理，那么，我们可以找到 $\varphi_i \in C_0^\infty(U_i)$ 和开集 V_i，使得 $K_i \subset V \subset U$，φ_i 的值域落在 $[0,1]$ 中并且 $\varphi_i\big|_{V_i} \equiv 1$。

最终，对每个 $i \leqslant N$，我们令

$$\chi_i(x) = \frac{\varphi_i(x)}{\varphi_1(x) + \cdots + \varphi_N(x)}$$

即可。证明的细节留作作业。 □

注记. 单位分解的证明并没有任何启发性的意义，我们只要能够运用该结论即可。

61 分布的局部刻画，分布的支集，支集在一点处的分布的结构定理

二零二零年九月二十三日，星期三，阴

定理 408 (单位分解)**.** 任意给定 $\mathbb{R}^n$ 中的紧集 K，假设 K 被有限个开集 $\{U_1,\cdots,U_N\}$ 所覆盖。那么，对每个 $j\leqslant N$，存在光滑函数 $\chi_j\in C_0^\infty(U_j)$，满足

1) 对任意 $x\in\mathbb{R}^n$，有 $0\leqslant\chi_j(x)\leqslant 1$；

2) 存在包含 K 的开集 V，对任意 $x\in V$，我们有

$$\chi_1(x)+\cdots+\chi_N(x)=1.$$

我们现在证明一个分布 u 在（所有）小的开集上的限制决定了 u，这表明分布是可以局部定义的：

定理 409. 给定开集 $\Omega\subset\mathbb{R}^n$。任意给定一族开集 $\left\{\Omega_i\middle|i\in I\right\}$，其中对任意的 $i\in I$，$\Omega_i\subset\Omega$。我们假定

$$\Omega=\bigcup_{i\in I}\Omega_i.$$

对每个 $i\in I$，我们在 Ω_i 上指定一个分布 $u_i\in\mathcal{D}'(\Omega_i)$。如果这一族分布 $\{u_i\}_{i\in I}$ 满足如下的相容关系：

$$u_i\big|_{\Omega_i\cap\Omega_j}=u_j\big|_{\Omega_i\cap\Omega_j},\quad \text{对任意的 } i,j\in I,$$

那么，存在唯一的 $u\in\mathcal{D}'(\Omega)$，使得对任意的 $i\in I$，我们都有

$$u\big|_{\Omega_i}=u_i.$$

进一步，$u\in L^1_{\mathrm{loc}}(\Omega)$ 当且仅当对每个 $i\in I$，$u\big|_{\Omega_i}\in L^1_{\mathrm{loc}}(\Omega)$；$u\in C^k(\Omega)$ 当且仅当对每个 $i\in I$，$u\big|_{\Omega_i}\in C^k(\Omega)$，其中 $k\in\mathbb{Z}_{\geqslant 0}$。

证明: 任意选定紧集 $K\subset\Omega$。由于 $\left\{\Omega_i\right\}_{i\in I}$ 也是 K 的开覆盖，我们选取有限开覆盖 $\Omega_{i_1},\cdots,\Omega_{i_N}$ 以及相应的单位分解 $\chi_1,\cdots,\chi_N$（与前面的定理相对应）。对任意的 $\varphi\in C_K^\infty(\Omega)$，我们定义

$$\langle u_K,\varphi\rangle:=\sum_{l=1}^N\langle u_{i_l},\chi_l\cdot\varphi\rangle.$$

我们必须说明 $\langle u_K,\varphi\rangle$ 的定义不依赖于覆盖 $\{\Omega_{i_l}\}$ 和单位分解 $\{\chi_l\}$ 的选取。实际上，如果我们再

选取 K 的有限开覆盖 $\Omega_{i'_1},\cdots,\Omega_{i'_{N'}}$ 以及相应的单位分解 $\chi'_1,\cdots,\chi'_{N'}$（其中 $i'_k\in I$），我们有

$$\begin{aligned}\sum_{l=1}^{N}\langle u_{i_l},\chi_l\cdot\varphi\rangle&=\sum_{l=1}^{N}\left\langle u_{i_l},\overbrace{\left(\sum_{l'=1}^{N'}\chi'_{l'}\right)}^{\text{在 supp}\varphi\text{ 上 }\equiv 1}\chi_l\cdot\varphi\right\rangle\\&=\sum_{l=1}^{N}\sum_{l'=1}^{N'}\langle u_{i_l},\underbrace{\chi_l\chi'_{l'}\cdot\varphi}_{\operatorname{supp}(\chi_l\chi'_{l'})\subset\Omega_{i_l}\cap\Omega_{i'_{l'}}}\rangle\\&=\sum_{l=1}^{N}\sum_{l'=1}^{N'}\left\langle u_{i_l}\big|_{\Omega_{i_l}\cap\Omega_{i'_{l'}}},\chi_l\chi'_{l'}\cdot\varphi\right\rangle\\&=\sum_{l=1}^{N}\sum_{l'=1}^{N'}\left\langle u_{i'_{l'}}\big|_{\Omega_{i_l}\cap\Omega_{i'_{l'}}},\chi_l\chi'_{l'}\cdot\varphi\right\rangle.\end{aligned}$$

根据上述计算的对称性，我们知道它一定等于 $\sum\limits_{l'=1}^{N'}\left\langle u_{i'_{l'}},\chi'_{l'}\varphi\right\rangle$。所以，我们最终证明了

$$\sum_{l=1}^{N}\langle u_{i_l},\chi_l\cdot\varphi\rangle=\sum_{l'=1}^{N'}\left\langle u_{i'_{l'}},\chi'_{l'}\varphi\right\rangle.$$

这就说明线性映射 u_K 是良好定义的。

所以，我们就把 u 定义为

$$\langle u,\varphi\rangle=\langle u_K,\varphi\rangle,$$

其中 $\varphi\in C_K^\infty(\Omega)$。为了说明 u 是良好定义的，我们要说明这个定义与紧集 K 的依赖关系，即如果 $\operatorname{supp}\varphi\subset K$ 并且 $\operatorname{supp}\varphi\subset K'$，其中 K' 是另一个紧集，我们需要证明

$$\langle u_K,\varphi\rangle=\langle u_{K'},\varphi\rangle.$$

事实上，$\operatorname{supp}\varphi\subset K\cap K'=K''$。所以，关于 K 的覆盖以及其相应的单位分解也可视作是 K'' 的单位分解，所以，利用单位分解所定义的 u_K 与 $u_{K''}$ 是一致的，即

$$\langle u_K,\varphi\rangle=\langle u_{K''},\varphi\rangle.$$

同理，

$$\langle u_{K'},\varphi\rangle=\langle u_{K''},\varphi\rangle.$$

所以，

$$\langle u_K,\varphi\rangle=\langle u_{K'},\varphi\rangle.$$

最终，我们要说明说明 $u=u_K$ 满足分布定义中所要求的不等式：给定紧集 $K\subset\Omega$，选取有限开覆盖 $\Omega_{i_1},\cdots,\Omega_{i_N}$ 以及相应的单位分解 $\chi_1,\cdots,\chi_N$。我们用 K_l 表示 χ_{i_l} 的支集，那么，$K_i\subset\Omega_{i_l}$

是紧集，其中 $l \leqslant N$。对任意的 $\varphi \in C_K^\infty(\Omega)$，我们有

$$\begin{aligned}|\langle u, \varphi\rangle| = |\langle u_K, \varphi\rangle| &= \left|\sum_{l=1}^N \langle u_{i_l}, \chi_l \cdot \varphi\rangle\right| \\ &\leqslant \sum_{l=1}^N |\langle u_{i_l}, \chi_l \cdot \varphi\rangle| \\ &\leqslant \sum_{l=1}^N C_i \sup_{|\alpha_l| \leqslant p_l} \|\partial^{\alpha_l}(\chi_l \cdot \varphi)\|_{L^\infty(K_l)} \\ &\leqslant C \sup_{|\alpha| \leqslant \sup_{l\leqslant N} p_l} \|\partial^\alpha \varphi\|_{L^\infty(K)}.\end{aligned}$$

命题中另外两个论断是平凡的。 □

注记. 这个定理表明，在 $\mathbb{R}^n$ 的开集上所定义的分布可以构成 $\mathbb{R}^n$ 上的一个层。

分布的支集

给定开集上的分布 $u \in \mathcal{D}'(\Omega)$，我们来定义它的支集。

假设 $\Omega' \subset \Omega$ 是开子集，如果 $u\big|_{\Omega'} = 0$，我们就说u 在 Ω' 上为零，也就是说，对于每个 $\varphi \in \mathcal{D}(\Omega')$，我们有

$$\langle u, \varphi\rangle = 0.$$

我们现在来说明，存在 Ω 中使得 u 在其上为零的最大开集。为此，我们定义

$$I = \left\{\Omega' \middle| \Omega' \subset \Omega \text{ 为开集}, u\big|_{\Omega'} = 0\right\}.$$

令

$$U = \bigcup_{\Omega' \in I} \Omega'.$$

按照定义，U 为开集。我们要证明 $U \in I$，为此，只要证明对任意的 $\varphi \in C_0^\infty(U)$，我们都有

$$\langle u, \varphi\rangle = 0,$$

即可。实际上，令 $K = \operatorname{supp}(\varphi)$，那么存在限个 $\Omega_1, \cdots, \Omega_N \in I$ 覆盖 K。我们取与这个覆盖相应的单位分解 $\chi_1, \cdots, \chi_N$。从而，

$$\langle u, \varphi\rangle = \sum_{j=1}^N \langle u, \overbrace{\chi_i \varphi}^{\text{支集在 } \Omega_i \text{ 中}} \rangle = 0.$$

最后一步，我们用到了 $u\big|_{\Omega_i} = 0$。很明显，U 是这种开集中最大的。

定义 410. 我们把 $\Omega - U$ 称作是 u 的**支集**并仍然用符号 $\operatorname{supp}(u)$ 表示。如果 $\operatorname{supp}(u)$ 是紧集，我们就说 u 是有紧支集的分布。我们用 $\mathcal{E}'(\mathbb{R}^n)$ 来表示 $\mathbb{R}^n$ 上有紧支集的分布的全体。

我们现在研究 $\mathcal{E}'(\mathbb{R}^n)$ 中的分布的基本性质：

命题 411. *在此命题中出现的分布都是有紧支集的*

1) $\mathcal{E}'(\mathbb{R}^n)$ *是线性空间，并且*

$$\operatorname{supp}(u_1+u_2)\subset\operatorname{supp}(u_1)\cup\operatorname{supp}(u_2).$$

2) 对任意的多重指标 α*，我们有*

$$\operatorname{supp}(\partial^\alpha u)\subset\operatorname{supp}(u).$$

3) 对任意的 $f\in C^\infty(\mathbb{R}^n)$*，我们有*

$$\operatorname{supp}(f\cdot u)\subset\operatorname{supp}(u)\cap\operatorname{supp}(f).$$

4) 任意给定试验函数 $\varphi\in\mathcal{D}(\varphi)$*，如果* $\operatorname{supp}(u)\cap\operatorname{supp}(\varphi)=\emptyset$*，那么*

$$\langle u,\varphi\rangle=0.$$

5) 有紧支集的分布的阶是有限的。

证明: 前三条性质按定义立得，我们留作作业。

为了证明 4)，我们注意到存在开集 $\Omega'\supset\operatorname{supp}(\varphi)$，使得 $\Omega'\cap\operatorname{supp}(u)=\emptyset$。按照支集的定义，$u\big|_{\Omega'}=0$，所以，$\langle u,\varphi\rangle=0$。

为了证明 5)，我们令 $K=\operatorname{supp}(u)$。此时，存在 $\psi\in C_0^\infty(\mathbb{R}^n)$，使得 $\psi\big|_K\equiv 1$。我们令 $K'=\operatorname{supp}(\psi)$。特别地，我们有

$$K\subset K'.$$

对于任意的试验函数 $\varphi\in\mathcal{D}(\mathbb{R}^n)$，我们有

$$\langle u,\varphi\rangle=\langle u,\psi\varphi\rangle.$$

这是因为根据 4)，$\operatorname{supp}((1-\psi)\varphi)\cap\operatorname{supp}(u)=\emptyset$。我们注意到 ψ 是事先给定的，所以对任意的试验函数 φ，我们都有

$$\operatorname{supp}(\psi\varphi)\subset K'.$$

这是一个固定的紧集。那么，按照分布的定义，在紧集 K' 上所对应的 C 和 p，使得

$$\begin{aligned}|\langle u,\varphi\rangle|&=|\langle u,\psi\varphi\rangle|\\&\leqslant C\sup_{|\alpha|\leqslant p}\|\partial^\alpha(\psi\cdot\varphi)\|_{L^\infty(K')}.\end{aligned}$$

上述右边至多出现了 u 的 p-次导数，所以 p 可以作为 u 的阶的上界。 □

利用证明中的技巧，我们有如下有趣的注解：

注记. 对于有紧支集的分布，我们可以扩大试验函数的空间来定义如下的配对：

$$\mathcal{E}'(\mathbb{R}^n) \times C^\infty(\mathbb{R}^n) \longrightarrow \mathbb{C}, \quad \langle u, \varphi \rangle := \langle u, \psi \cdot \psi \rangle,$$

其中，ψ 是在 $K = \mathrm{supp}(u)$ 上恒为 1 的有紧支集的光滑函数。我们将在作业中证明这个定义不依赖于 ψ 的选取。

这种用乘以一个函数的方式来记住一个分布的支集是一个很有用的技巧，之后会经常看到。

我们现在引入一种平坦性的概念：给定非空的闭集 $F \subset \mathbb{R}^n$，$\varphi \in \mathcal{D}(\mathbb{R}^n)$ 是试验函数。如果存在非负整数 p，使得对每个满足 $|\alpha| \leqslant p$ 的多重指标 α，我们都有

$$\partial^\alpha \varphi \big|_F = 0,$$

我们就说试验函数 φ 在 F 上是 p-次平坦的。

引理 412 (平坦性引理)**.** 给定是阶为 p 的分布 $u \in \mathcal{D}'(\mathbb{R}^n)$，如果试验函数 $\varphi \in \mathcal{D}(\mathbb{R}^n)$ 在 $\mathrm{supp}(u)$ 上 p-次平坦，那么

$$\langle u, \varphi \rangle = 0.$$

证明: 我们定义如下的集合：

$$K = \mathrm{supp}(\varphi), \quad F = \mathrm{supp}(u), \quad L = K \cap F.$$

我们还定义

$$L_{k\varepsilon} = \{x \in \mathbb{R}^n \big| d(x, L) < k\varepsilon\}, \quad k = 1, 2, 3.$$

我们仍然用 χ_ε 来记我们心爱的截断函数。此时，我们定义

$$\psi_\varepsilon = \mathbf{1}_{L_{2\varepsilon}} * \chi_\varepsilon.$$

这个函数满足如下几条显然的性质

1) $\mathrm{supp}(\psi_\varepsilon) \subset L_{3\varepsilon}$；

2) $\psi_\varepsilon \big|_{L_\varepsilon} \equiv 1$；

3) 对每个多重指标 α，我们都有

$$\|\partial^\alpha \psi_\varepsilon\|_\infty \leqslant C_\alpha \varepsilon^{-|\alpha|}.$$

其中，$C_\alpha = \|\partial^\alpha \chi\|_{L^1}$：

实际上，我们有

$$\begin{aligned} |\partial^\alpha \psi_\varepsilon| &= |\mathbf{1}_{L_{2\varepsilon}} * (\partial^\alpha \chi_\varepsilon)| \\ &\leqslant \varepsilon^{-|\alpha|} \int_{L_{2\varepsilon}} \varepsilon^{-n} \left| (\partial^\alpha \chi) \left(\frac{x}{\varepsilon} \right) \right| dx \\ &= \varepsilon^{-|\alpha|} \|\partial^\alpha \chi\|_{L^1}. \end{aligned}$$

我们利用 ψ_ε 将 φ 切成两个部分：

$$\begin{aligned}\langle u,\varphi\rangle &= \langle u,\ \overbrace{(1-\psi_\varepsilon)\varphi}^{\text{支集与 } F \text{ 不交}}\ \rangle + \langle u,\ \overbrace{\psi_\varepsilon\varphi}^{\text{支集 } \subset L_{3\varepsilon}}\ \rangle \\ &= \langle u,\psi_\varepsilon\varphi\rangle.\end{aligned}$$

由于 φ 在 F 上是 p-阶平坦的并且 $L_{3\varepsilon}$ 中的每个点距离 F 不超过 3ε，根据 Taylor 公式，对于任意的 $y\in L_{3\varepsilon}$，存在 $x\in F$ 并且 $|x-y|<3\varepsilon$，以及 $\vartheta\in[0,1]$，使得

$$\begin{aligned}\varphi(y) &= \sum_{|\alpha|\leqslant p}\frac{\partial^\alpha\varphi(x)}{\alpha!}(x-y)^\alpha + \sum_{|\alpha|=p+1}\frac{\partial^\alpha\varphi\big(x+\vartheta(y-x)\big)}{\alpha!}(x-y)^\alpha \\ &\overset{\text{平坦性}}{=} \sum_{|\beta|=p+1}\frac{\partial^\beta\varphi\big(x+\vartheta(y-x)\big)}{\beta!}(x-y)^\beta.\end{aligned}$$

所以，

$$\partial^\alpha\varphi(y) = \sum_{|\beta|=p+1}\sum_{\alpha_1+\alpha_2=\alpha}\frac{\alpha!}{\alpha_1!\alpha_2!}\frac{\partial^{\alpha_1}\left(\partial^\beta\varphi\big(x+\vartheta(y-x)\big)\right)}{\beta!}\underbrace{\partial^{\alpha_2}\left((x-y)^\beta\right)}_{|\cdot|\leqslant C\varepsilon^{p+1-|\alpha_2|}}.$$

所以，在 $L_{3\varepsilon}$ 上，当 $|\alpha|\leqslant p$ 时，存在常数 C（依赖于 φ，可能很大，但是这个常数不依赖于 ε），使得

$$|\partial^\alpha\varphi|\leqslant C\varepsilon^{p+1-|\alpha|}.$$

从而对于 $|\alpha|\leqslant p$，我们就有（利用 Leibniz 公式）

$$\begin{aligned}|\partial^\alpha\big(\psi_\varepsilon\cdot\varphi\big)| &\leqslant 2^{|\alpha|}\sum_{|\beta|\leqslant|\alpha|}|\partial^{\alpha-\beta}\psi_\varepsilon||\partial^\beta\varphi| \\ &\leqslant C'\sum_{|\beta|\leqslant|\alpha|}\varepsilon^{-|\alpha-\beta|}\varepsilon^{p+1-|\beta|} \\ &\leqslant C''\varepsilon.\end{aligned}$$

根据分布的定义以及 u 的阶不超过 p，我们知道

$$\begin{aligned}|\langle u,\psi_\varepsilon\cdot\varphi\rangle| &\leqslant C'''\sup_{|\alpha|\leqslant p}\|\partial^\alpha\big(\psi_\varepsilon\cdot\varphi\big)\|_{L^\infty(L_{3\varepsilon})} \\ &\leqslant C''''\varepsilon\end{aligned}$$

令 $\varepsilon\to 0$，这就完成了该定理的证明。 □

定理 413 (支集为单点集的分布的结构定理)**.** *任意给定分布 $u\in\mathcal{D}'(\mathbb{R}^n)$，如果它的支集满足*

$$\operatorname{supp}(u)=\{0\},$$

那么，存在有限多个多重指标 α_i 和复数 C_{α_i}，其中 $i=1,\cdots,N$，使得

$$u=\sum_{j=1}^N C_{\alpha_j}\cdot\partial^{\alpha_j}\left(\delta_0\right).$$

证明： 由于 u 的支集是紧的，我们可假设其阶为 $p \in \mathbb{Z}_{\geqslant 0}$。我们选取有紧支集的函数 χ，使得它在包含 0 的一个开集上为 1。我们用类似于多项式的近试验函数 $\left(\sum\limits_{|\alpha|\leqslant p} \dfrac{\partial^\alpha \varphi(0)}{\alpha!} x^\alpha\right)\chi(x)$ 来逼近 φ，即定义

$$r(x) = \varphi(x) - \left(\sum_{|\alpha|\leqslant p} \frac{\partial^\alpha \varphi(0)}{\alpha!} x^\alpha\right)\chi(x).$$

很明显，$r(x)$ 是试验函数。

由于 $\left(\sum\limits_{|\alpha|\leqslant p} \dfrac{\partial^\alpha \varphi(0)}{\alpha!} x^\alpha\right)\chi(x)$ 和 $\varphi(x)$ 在 0 处一直到 p 阶导数是相等的，所以，$r(x)$ 在 $\{0\}$ 上是 p-阶平坦的，所以

$$\langle u, r\rangle = 0.$$

从而，

$$\langle u, \varphi\rangle = \sum_{|\alpha|\leqslant p} \frac{\partial^\alpha \varphi(0)}{\alpha!} \langle u, x^\alpha \chi\rangle,$$

所以，只要取 $C_\alpha = (-1)^{|\alpha|}\dfrac{\langle u, x^\alpha\chi\rangle}{\alpha!}$，我们就有

$$\langle u, \varphi\rangle = \sum_{|\alpha|\leqslant p} C_\alpha \langle \partial^\alpha(\delta_0), \varphi\rangle.$$

命题得证。 □

61.1 作业：齐次分布，Hadamard 有限部分，分布除以多项式

清华大学 19-20 春季学期，数学分析三，作业 1

请用 A4 大小的纸张**正反面**用钢笔，签字笔或者圆珠笔书写，并注明自己的姓名，年级和作业的**总页数**。除定理公式所涉及的人名之外，请使用**中文**。本次作业的提交时间和地点为 **9 月 30 日**（周三）上午的课堂上，逾期视作零分。

分布的定义、例子与操作

习题 A. 我们总假设 $\Omega \subset \mathbb{R}^n$ 是非空的开集。

A1) 假设 $\psi(x) \in C^\infty(\mathbb{R})$ 并且 $\psi(0) = 0$，那么，$\dfrac{\psi(x)}{x}$ 也是光滑函数。

A2) 对任意给定的分布 $u \in \mathcal{D}'(\Omega)$，对任意的试验函数序列 $\{\varphi_p\}_{p\geqslant 1} \subset \mathcal{D}(\Omega)$，如果 $\varphi_p \xrightarrow{\mathcal{D}} 0$，试证明，

$$\lim_{p\to\infty} \langle u, \varphi_p\rangle = 0.$$

A3) 对任意整数 $k \geqslant 0$，我们定义 $\mathcal{D}(\mathbb{R})$ 上的线性泛函：

$$\langle \delta^{(k)}, \varphi\rangle = (-1)^k \varphi^{(k)}(0),$$

其中，$\varphi^{(k)}$ 指的是求 k 次导数。证明，这定义了 $\mathbb{R}$ 上的分布并且其阶恰好为 k。

A4) 我们定义 $\mathcal{D}(\mathbb{R})$ 上的线性泛函：

$$\langle u, \varphi\rangle = \sum_{k=0}^{\infty} \varphi^{(k)}(k).$$

证明，这定义了 $\mathbb{R}$ 上的分布然而我们不能定义它的阶。

A5) 证明，$\mathrm{vp}\dfrac{1}{x}$ 的阶恰为 1。

A6) 证明，如果试验函数 $\varphi \in \mathcal{D}(\mathbb{R})$ 满足 $\varphi(0) = 0$，那么

$$\langle \mathrm{vp}\frac{1}{x}, \varphi\rangle = \int_{\mathbb{R}} \frac{\varphi(x)}{x} dx.$$

A7) 证明，$\mathbb{R}_{>0}$ 上的函数

$$f(x) = e^{\frac{1}{x}} \mathbf{1}_{(0,+\infty)}(x)$$

不能延拓为 $\mathbb{R}$ 上的一个分布，即不存在 $u \in \mathcal{D}'(\mathbb{R})$，使得对任意的 $\varphi \in \mathcal{D}((0,\infty))$，我们都有

$$\langle u, \varphi\rangle = \int_0^{\infty} e^{\frac{1}{x}} \varphi(x).$$

A8) 对任意的 $\varepsilon > 0$，我们令

$$f_\varepsilon(x) = \frac{1}{x}\mathbf{1}_{(-\infty,-\varepsilon)\cup(2\varepsilon,+\infty)}(x).$$

证明，当 $\varepsilon \to 0$ 时，f_ε 在分布（$\mathcal{D}'(\mathbb{R})$）的意义下有极限并计算该极限。

A9) 对任意的 $\varepsilon > 0$，我们定义

$$g_\varepsilon(x) = \frac{1}{x}\mathbf{1}_{(-\infty,-\varepsilon)\cup(\varepsilon^2,+\infty)}(x).$$

我们把它们看作是 $\mathcal{D}'(\mathbb{R})$ 中的一族分布。证明，当 $\varepsilon \to 0$ 时，g_ε 在分布的意义下是没有极限的。

习题 B（$\mathbb{R}$ 上的概率分布函数）.

B1) 假设 μ 是 $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ 上的概率测度，即 $\mu(\mathbb{R}) = 1$，其中 $\mathcal{B}(\mathbb{R})$ 是 Borel 代数。证明，函数

$$f(x) = \mu((-\infty, x])$$

是递增的、右连续的函数并且满足

$$\lim_{x\to-\infty} f(x) = 0, \quad \lim_{x\to+\infty} f(x) = 1.$$

B2) 假设 μ 是 $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ 上的概率测度，$f(x) = \mu((-\infty, x))$，证明，在分布的意义下

$$f' \overset{\mathcal{D}'(\mathbb{R})}{=} \mu.$$

B3) 任给定 $\mathbb{R}$ 上的单调递增的函数 $f(x)$，满足

$$\lim_{x\to-\infty} f(x) = 0, \lim_{x\to+\infty} f(x) = 1.$$

证明，其分布意义下的导数 f' 可以被视为是 $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ 上的概率测度，即存在概率测度 μ，使得对任意的 $\varphi \in \mathcal{D}(\mathbb{R})$，我们有

$$\langle f', \varphi\rangle = \int_{\mathbb{R}} \varphi \, d\mu.$$

习题 C（Hadamard 的有限部分（parties finies）$\mathrm{pf}\, x_+^\alpha$）.

C1) 给定参数 $\alpha \in (-1, 0)$，证明，函数 x^α 在 $[0, 1)$ 上的可积函数。

C2) 给定参数 $\alpha \in (-2,-1)$。证明，对任意 $\varphi \in \mathcal{D}(\mathbb{R})$，对任意的 $\varepsilon > 0$，我们都可以把积分 $\int_\varepsilon^\infty x^\alpha \varphi(x)dx$ 写成下面的形式：

$$\int_\varepsilon^\infty x^\alpha \varphi(x)dx = A\varepsilon^{\alpha+1} + R_\varphi(\varepsilon),$$

其中，A 是由 φ 决定的常数并且当 $\varepsilon \to 0^+$ 时，$R_\varphi(\varepsilon)$ 有极限。

C3) 证明，下述公式定义出 $\mathcal{D}'(\mathbb{R})$ 中的分布：

$$\langle \mathrm{pf}\, x_+^\alpha, \varphi\rangle = \lim_{\varepsilon\to 0^+} R_\varphi(\varepsilon).$$

C4) 给定参数 $\alpha \in (-\infty,-1) - \mathbb{Z}$（$\alpha$ 不是整数）。证明，对任意 $\varphi \in \mathcal{D}(\mathbb{R})$，对任意的 $\varepsilon > 0$，我们都可以把积分 $\int_\varepsilon^\infty x^\alpha \varphi(x)dx$ 写成下面的形式：

$$\int_\varepsilon^\infty x^\alpha \varphi(x) = P_\varphi(\varepsilon) + R_\varphi(\varepsilon),$$

其中，$P_\varphi(\varepsilon)$ 是有限个 ε 的负次方的常数系数（依赖于 φ）线性组合，而当 $\varepsilon \to 0^+$ 时，$R_\varphi(\varepsilon)$ 有极限。

C5) 证明，下述公式定义出 $\mathcal{D}'(\mathbb{R})$ 中的分布：

$$\langle \mathrm{pf}\, x_+^\alpha, \varphi\rangle = \lim_{\varepsilon\to 0^+} R_\varphi(\varepsilon).$$

C6) 计算如下的分布

$$x \cdot \mathrm{pf}\, x_+^\alpha, \quad \frac{d}{dx}\left(\mathrm{pf}\, x_+^\alpha\right),$$

其中 $\alpha \in (-\infty, 0) - \mathbb{Z}$。

C7) 如果 $\alpha = -2$，你是否可以定义出 $\mathrm{pf}\, x_+^{-2}$?

C8)（选做）如果 $\alpha = -2, -3, -4, \cdots$，你能否对上述问题做出相应的改动和回答?

习题 D（关于分布的一个除法问题，$\div x$）.

D1) 任意给定试验函数 $\varphi \in \mathcal{D}(\mathbb{R})$，证明，如果 $\varphi(0) = 0$，那么函数

$$\frac{\varphi(x)}{x} \in \mathcal{D}(\mathbb{R}).$$

D2) 证明，对每个 $u \in \mathcal{D}'(\mathbb{R})$，存在 $\widetilde{u} \in \mathcal{D}'(\mathbb{R})$，使得在分布的意义下

$$x \cdot \widetilde{u} = u.$$

D3) 试找出一切 $u \in \mathcal{D}'(\mathbb{R})$，使得在分布的意义下

$$x \cdot u = 0.$$

D4) 试在分布的意义下找出如下四个方程的所有解：

$$x \cdot u = 1, \quad x \cdot u = \delta_0, \quad x \cdot u = \mathrm{pv}\frac{1}{x},$$

和

$$x \cdot u = \mathrm{pf}\, x_+^\alpha,$$

其中 $\alpha < -1$ 且不是整数。

D5) 试找出所有的 $u \in \mathcal{D}'(\mathbb{R})$，使得在分布的意义下，我们有

$$u + x \cdot u' = 0.$$

习题 E（一个 $\mathbb{R}^2$ 上的微分方程）.

现在在 $\mathbb{R}^2$ 上面研究问题，我们用 (x, y) 作为标准的坐标函数。

E1) 假设分布 $u \in \mathcal{D}'(\mathbb{R}^2)$ 满足

$$\partial_x u \overset{\mathcal{D}'}{=} 0.$$

证明，存在 $v \in \mathcal{D}'(\mathbb{R})$，使得对任意的 $\varphi \in \mathcal{D}(\mathbb{R}^2)$，我们都有

$$\langle u, \varphi \rangle = \left\langle v, \int_{\mathbb{R}} \varphi(x, y) dx \right\rangle.$$

E2)（正则性）假设分布 $u \in \mathcal{D}'(\mathbb{R}^2)$ 满足

$$\partial_x u \in C^0(\mathbb{R}^2), \quad \partial_y u \in C^0(\mathbb{R}^2).$$

也就是说，$\partial_x u$ 和 $\partial_y u$ 分别对应着连续函数所定义的分布。证明，$u \in C^1(\mathbb{R}^2)$。

习题 F（齐次分布）.

给定试验函数 $\varphi \in \mathcal{D}(\mathbb{R}^n)$，对任意的实数 $\lambda \neq 0$，我们定义

$$\varphi_\lambda(x) = \varphi\left(\frac{x}{\lambda}\right).$$

对于 $u \in \mathcal{D}'(\mathbb{R}^n)$，我们定义分布 u_λ 如下：

$$\langle u_\lambda, \varphi \rangle = \left\langle u, |\lambda|^n \varphi_{\frac{1}{\lambda}} \right\rangle.$$

如果对每个 $\lambda > 0$，下面的等式成立：

$$u_\lambda \overset{\mathcal{D}'}{=} \lambda^{-d} u,$$

我们就称 u 为次数是 d 的**齐次分布**，其中 $d \in \mathbb{R}$。

F1) 假设分布 $u \in \mathcal{D}'(\mathbb{R}^n)$ 是由某个局部可积的函数 f 定义。证明，u 是 d 次齐次分布当且仅当对几乎处处的 x，我们有

$$f(\lambda x) = \lambda^d f(x), \quad \lambda > 0.$$

F2) 证明，$\mathrm{vp}\dfrac{1}{x}$ 和 $\mathrm{pf}\, x_+^\alpha$（其中 $\alpha < -1$ 且非整数）是 $\mathcal{D}'(\mathbb{R})$ 上的齐次分布并确定它们的次数。

F3) 证明，δ_0 是 $\mathcal{D}'(\mathbb{R}^n)$ 上的齐次分布并确定其次数。

F4) 假设 $u \in \mathcal{D}'(\mathbb{R}^n)$ 上的 d 次齐次分布，证明。$\partial_{x_1} u \in \mathcal{D}'(\mathbb{R}^n)$ 是 $d-1$ 次的齐次分布。

F5) 任意给定 $u_1, u_2, \cdots, u_N \in \mathcal{D}'(\mathbb{R}^n)$ 上 N 个齐次分布，如果它们的次数互不相同，证明，它们在 $\mathcal{D}'(\mathbb{R}^n)$ 中是 $\mathbb{C}$-线性无关的。

F6) 任意给定试验函数 $\varphi \in \mathcal{D}(\mathbb{R}^n)$，任意给定正实数 λ 和 λ_0，假设 $\lambda \neq \lambda_0$。我们定义

$$\psi_\lambda(x) = \frac{\varphi_{\frac{1}{\lambda}}(x) - \varphi_{\frac{1}{\lambda_0}}(x)}{\lambda - \lambda_0}$$

证明，当 $\lambda \to \lambda_0$ 时，上述函数在 $\mathcal{D}(\mathbb{R}^n)$ 中有极限并计算该极限。

F7)（齐次分布的 Euler 公式）任意给定 $u \in \mathcal{D}'(\mathbb{R}^n)$，$d \in \mathbb{R}$。证明，$u$ 是 d 次齐次分布当且仅当

$$\sum_{k=1}^{n} x_k \cdot \frac{\partial u}{\partial x_k} \overset{\mathcal{D}'}{=} d \cdot u.$$

F8) 试找出 $\mathcal{D}'(\mathbb{R}^1)$ 中所有次数为 0 和 1 的次齐次分布。

Je suis mathématicien. Les mathématiques ont rempli ma vie.

— Laurent Schwartz

---

62 分布的卷积：分布与微分和积分的交换，分布与紧支集分布的卷积，Dirac 分布与平移，卷积与求导数的交换

二零二零年九月二十八日，星期一，多云

我们要研究分布的正则化问题，也就是用如何用光滑函数来逼近一个给定的分布。我们在上个学期已经学习关于光滑逼近的一个重要工具：卷积。对于 $u \in L^1(\mathbb{R}^n)$，$\varphi \in \mathcal{D}(\mathbb{R}^n)$，我们回忆一下卷积的定义：

$$u * f = \int_{\mathbb{R}^n} u(y)\varphi(x-y)dy.$$

据此，对于 $u \in \mathcal{D}'(\mathbb{R}^n)$ 和 $\varphi(x) \in \mathcal{D}(\mathbb{R}^n)$，我们如下定义它们的**卷积**：对任意的 $x \in \mathbb{R}^n$，令

$$\big(u * \varphi\big)(x) := \langle u, \varphi(x - \cdot)\rangle,$$

我们下面证明，这个定义给出了双线性的映射：

$$\mathcal{D}'(\mathbb{R}^n) \times \mathcal{D}(\mathbb{R}^n) \xrightarrow{*} C^\infty(\mathbb{R}^n), \quad (u, \varphi) \mapsto u * \varphi.$$

定理 414. *对于 $u \in \mathcal{D}'(\mathbb{R}^n)$ 和 $\varphi(x) \in \mathcal{D}(\mathbb{R}^n)$，如上定义的卷积 $u * \varphi$ 是 $\mathbb{R}^n$ 上的光滑函数。进一步，对每个多重指标 α，我们都有*

$$\partial^\alpha(u * \varphi) = u * \partial^\alpha \varphi.$$

证明： 我们证明 $u * \varphi$ 是连续可微的函数并且对每个 $1 \leqslant k \leqslant n$，都有

$$\partial_k(u * \varphi) = u * (\partial_k \varphi).$$

实际上，只要证明了这个命题，根据归纳法，我们每次都把导数作用在 φ 上，这就给出了命题的证明。

先证明 $u * \varphi$ 是良好定义的。对任意的 $x \in \mathbb{R}^n$，把它视作是给定的参数，再把 $\varphi(x-y)$ 看作是 y 的函数，从而，$\varphi(x-y) \in \mathcal{D}(\mathbb{R}^n_y)$，其中我们在 $\mathbb{R}^n$ 上用下标 y 表明这里的函数都是以 y 为变量的。特别地，这表明分布与试验函数之间的配对 $\langle u, \varphi(x-\cdot)\rangle$ 是良好定义的，所以 $u * \varphi(x)$ 作为 x 的函数是良好定义的。

再证明 $u * \varphi$ 是连续函数。任意固定 $x \in \mathbb{R}^n$。任意选取点列 $\{a_j\}_{j\geqslant 1} \subset \mathbb{R}^n$，假设当 $j \to \infty$ 时，$a_j \to 0$。我们要证明

$$\lim_{j\to\infty} (u * \varphi)(x + a_j) = (u * \varphi)(x).$$

这等价于证明

$$\lim_{j\to\infty} \langle u, \varphi(x + a_k - \cdot)\rangle = \langle u, \varphi(x - \cdot)\rangle.$$

由于 $a_k + x \to x$，所以，我们只要证明对任意的 $\{b_j\}_{j\geqslant 1} \subset \mathbb{R}^n$，当 $j \to \infty$ 时，$b_j \to b$，那么，

$$\varphi(y + b_j) \xrightarrow{\mathcal{D}} \varphi(y),$$

即可（最终，我们选取 $b_j = a_j + x$，$y = -y$）。这是显然的：令 $K = \mathrm{supp}(\varphi)$，那么，$\mathrm{supp}(\varphi(\cdot + b_j))$ 显然都落在一个共同的紧集中 K'（如果 $\delta = \sup\limits_{j\geqslant 1} |b_j|$，那么，对任意的 j，$\mathrm{supp}(\varphi(\cdot + b_j)) \subset K_\delta$）。对任意的多重指标 α，我们自然有

$$\lim_{j\to\infty} \|\partial^\alpha \varphi(x + b_j) - \partial^\alpha \varphi(x)\|_{L^\infty(K')} = 0.$$

这就说明了在 $\mathcal{D}(\mathbb{R}^n)$ 中，$\varphi(\cdot + b_j) \to \varphi$。

我们最终证明，对任意的 $a \in \mathbb{R}^n$（$|a| = 1$），都有

$$\nabla_a(u * \varphi) = u * (\nabla_a \varphi).$$

固定 $x_0, a \in \mathbb{R}^n (|a| = 1)$，根据 Taylor 公式：

$$f(x) = \sum_{|\alpha|\leqslant m} \frac{\partial^\alpha f(x_0)}{\alpha!}(x - x_0)^\alpha + (m+1)\sum_{|\alpha|=m+1} \frac{(x - x_0)^\alpha}{\alpha!} \int_0^1 (1-t)^m \partial^\alpha f\left(x_0 + t(x - x_0)\right) dt.$$

我们有

$$\varphi(x_0 - y + \varepsilon a) - \varphi(x_0 - y) = \varepsilon \sum_{j=1}^n a_j \partial_j \varphi(x_0 - y) + \varepsilon^2 r(y, \varepsilon, a),$$

其中，

$$r(y, \varepsilon, a) = 2 \sum_{|\alpha|=2} \frac{a^\alpha}{\alpha!} \int_0^1 (1-t) \partial^\alpha \varphi(x_0 - y + t\varepsilon a) dt.$$

将上面的式子与分布 u 进行配对，我们得到

$$u * \varphi(x_0 + \varepsilon a) - u * \varphi(x_0 + \varepsilon a) - \varepsilon \sum_{j=1}^n a_j u * \partial_j \varphi(x_0) = \varepsilon^2 \langle u, r(\cdot, \varepsilon, a)\rangle,$$

其中，我们将 a, ε 和 x_0 视作是给定的参数。

根据 $r(\cdot, \varepsilon, a)$ 的表达式，我们很容易看出 $r(\cdot, \varepsilon, a)$ 的支集是紧的，实际上，它的支集在 $\tau_{\pm x_0}(\mathrm{supp}(\varphi))$ 距离不超过 1 的附近）。另外，对于任意的多重指标 β，我们有

$$\partial_y^\beta r(y, \varepsilon, a) = 2 \sum_{|\alpha|=2} \frac{a^\alpha}{\alpha!} \int_0^1 (1-t) \partial^{\alpha+\beta} \varphi(x_0 - y + t\varepsilon a) dt.$$

所以，$r(\cdot, \varepsilon, a)$ 的各阶导数被 φ 在它的支集上的各阶导数所控制，即

$$\|\partial_y^\beta r(y, \varepsilon, a)\|_{L^\infty} \leqslant C \sup_{|\alpha|\leqslant|\beta|+2} \|\partial^\alpha \varphi\|_{L^\infty}.$$

根据分布的定义，我们就有

$$\langle u, r(\cdot, \varepsilon, a)\rangle = O(1).$$

（实际上，我们需要说明这里的 $O(1)$ 是不依赖于参数 ε 的，我们把这一点的证明留作作业，这是一道很重要的习题。）从而，当 $\varepsilon \to 0$ 时，我们得到

$$\frac{|u * \varphi(x_0 + \varepsilon a) - u * \varphi(x_0) - \varepsilon u * (\nabla_a \varphi(x_0))\,|}{\varepsilon |a|} = o(1),$$

这个等式就给出了要证明的等式。 □

当试验函数含有参数时，我们要考虑对参数的微分或积分与分布的配对是否可以交换。这就是下面的定理。证明的主要工具就是 Lebesgue 控制收敛定理。

定理 415 (交换积分或者微分)**.** 给定分布 $u \in \mathcal{D}'(\mathbb{R}^n)$ 和双变量的试验函数 $\varphi(x,y) \in \mathcal{D}(\mathbb{R}^n_x \times \mathbb{R}^p_y)$，其中 $p \geqslant 1$。对于每个参数 $y \in \mathbb{R}^p$，(由于 $\mathrm{supp}_x \psi(\cdot, y)$ 是紧的)，我们定义函数

$$\psi(y) = \langle u, \varphi(\cdot, y) \rangle.$$

那么，$\psi(y)$ 是 $\mathbb{R}^p$ 上的光滑有紧支集的函数，即 $\psi(y) \in \mathcal{D}(\mathbb{R}^p_y)$。进一步，对任意的多重指标 α，我们有

$$\partial^\alpha \psi(y) = \left\langle u, \partial_y^\alpha \varphi(\cdot, y) \right\rangle.$$

另外，$\int_{\mathbb{R}^n} \cdot \, dy$ 与分布的配对可以交换：

$$\int_{\mathbb{R}^p} \psi(y) dy = \left\langle u, \int_{\mathbb{R}^p} \varphi(\cdot, y) dy \right\rangle.$$

证明: 我们首先证明 ψ 是光滑有紧支集的函数并且满足定理中的求导公式。我们可以假设 $\mathrm{supp}(\varphi(x,y)) \subset K_1 \times K_2$，其中，$K_1 \subset \mathbb{R}^n$ 和 $K_2 \subset \mathbb{R}^p$ 均为紧集。那么，对任意的 $y \in \mathbb{R}^p$，$\mathrm{supp}\phi(\cdot, y) \subset K_1$。从而，$\mathrm{supp}\psi \subset K_2$（因为当 $y \notin K_2$ 时，按照定义，$\psi(y) = 0$）。

我们观察到，这一部分剩下的证明实际上和上面的定理的证明是一致的：我们只需要按照 Taylor 公式，把 $\varphi(x, y_0 + \varepsilon a)$ 写成：

$$\varphi(x, y_0 + \varepsilon a) - \varphi(x, y_0) - \varepsilon \sum_{j=1}^{p} a_j \partial_{y_j} \varphi(x, y_0) = \varepsilon^2 r(x, y_0, \varepsilon, a),$$

然后重复之前的证明过程即可。

为了证明积分与分布的配对可以交换，我们先处理 $p = 1$ 的情形。我们利用 Riemann 积分的定义（我们上学期证明过，光滑情况下，Riemann 积分与 Lebesgue 积分是一样的）：假设 $K_2 \in [-A, A]$，其中 A 为正整数。对任意的正整数 k，我们定义部分和

$$r_k(x) = \sum_{|j| \leqslant kA} \frac{1}{k} \varphi(x, \frac{j}{k}).$$

很显然，对每个 $k \geqslant 1$，$r_k(x) \in \mathcal{D}(\mathbb{R}^n_x)$。我们现在证明，在 $r_k(x) \in \mathcal{D}(\mathbb{R}^n_x)$ 的拓扑下，有

$$r_k(x) \overset{\mathcal{D}(\mathbb{R}^n)}{\longrightarrow} \int_{\mathbb{R}^1} \varphi(x, y) dy, \quad k \to \infty.$$

通过做差，我们得到

$$\int_{\mathbb{R}^1} \varphi(x, y) dy - r_k(x) = \sum_{|j| \leqslant kA - 1} \int_{\frac{j}{k}}^{\frac{j+1}{k}} \Big(\varphi(x, y) - \varphi(x, \frac{j}{k}) \Big) dy,$$

对任意的多重指标 α，利用 Lebesgue 控制收敛定理的推论，交换积分与求导数在此情况下总可以交换：

$$\partial_x^\alpha\Big(\int_{\mathbb{R}^1}\varphi(x,y)dy-r_k(x)\Big)=\sum_{|j|\leqslant kA-1}\int_{\frac{j}{k}}^{\frac{j+1}{k}}\partial_x^\alpha\Big(\varphi(x,y)-\varphi(x,\frac{j}{k})\Big)dy.$$

根据 Rolle 中值定理，就有

$$\begin{aligned}\Big|\partial_x^\alpha\Big(\int_{\mathbb{R}^1}\varphi(x,y)dy-r_k(x)\Big)\Big|&\leqslant\Big(\sum_{|j|\leqslant kA-1}\int_{\frac{j}{k}}^{\frac{j+1}{k}}\underbrace{|y-\frac{j}{k}|}_{\leqslant\frac{1}{k}}dy\Big)\sup_{|\beta|\leqslant1}\|\partial_x^\alpha\partial_y^\beta\varphi\|_\infty\\&\leqslant\frac{2A}{k}\sup_{|\beta|\leqslant1}\|\partial_x^\alpha\partial_y^\beta\varphi\|_\infty.\end{aligned}$$

按照定义，这表明当 $k\to\infty$ 时，我们有 $r_k(x)\xrightarrow{\mathcal{D}(\mathbb{R}^n)}\int_{\mathbb{R}^1}\varphi(x,y)dy$。从而，

$$\begin{aligned}\Big\langle u,\int_{\mathbb{R}^p}\varphi(\cdot,y)dy\Big\rangle&=\lim_{k\to\infty}\langle u,r_k(x)\rangle\\&=\lim_{k\to\infty}\sum_{|j|\leqslant kA}\frac{1}{k}\langle u,\varphi(x,\frac{j}{k})\rangle\\&=\lim_{k\to\infty}\sum_{|j|\leqslant kA}\frac{1}{k}\psi(\frac{j}{k}).\end{aligned}$$

根据 Riemann 积分的定义，我们就得到

$$\Big\langle u,\int_{\mathbb{R}^p}\varphi(\cdot,y)dy\Big\rangle=\int_{\mathbb{R}}\psi(y)dy.$$

这就完成了 $p=1$ 时的证明。对于一般的 p，我们用归纳法和 Fubini 定理。假设 $y=(y',y_p)\in\mathbb{R}^p$，其中 $y'\in\mathbb{R}^{p-1}$，$y_p\in\mathbb{R}$。那么，

$$\begin{aligned}\Big\langle u,\int_{\mathbb{R}^p}\varphi(\cdot,y)dy\Big\rangle&=\Big\langle u,\int_{\mathbb{R}^{p-1}}\Big(\int_{\mathbb{R}}\varphi(\cdot,y',y_p)dy'\Big)dy_p\Big\rangle\\&=\int_{\mathbb{R}^{p-1}}\Big\langle u,\int_{\mathbb{R}}\varphi(\cdot,y',y_p)dy'\Big\rangle dy_p\\&=\int_{\mathbb{R}^{p-1}}\int_{\mathbb{R}}\langle u,\varphi(\cdot,y',y_p)\rangle dy'dy_p\end{aligned}$$

我们对最后的等式利用 Fubini 定理可以把累次积分化作对 $y\in\mathbb{R}^p$ 的积分，这就完成了证明。 □

下面的命题总结了卷积

$$\mathcal{D}'(\mathbb{R}^n)\times\mathcal{D}(\mathbb{R}^n)\xrightarrow{*}C^\infty(\mathbb{R}^n),\quad(u,\varphi)\mapsto u*\varphi.$$

的大部分性质：

命题 416 ($\mathcal{D}'(\mathbb{R}^n) * \mathcal{D}(\mathbb{R}^n)$ 的性质)**.** 任意给定分布 $u \in \mathcal{D}'(\mathbb{R}^n)$ 和试验函数 $\varphi, \psi \in \mathcal{D}(\mathbb{R}^n)$，那么，我们有

1) 卷积之后的支集满足：$\operatorname{supp}(u * \varphi) \subset \operatorname{supp}(u) + \operatorname{supp}(\varphi)$；[17]

2) 卷积的结合律：

$$u * (\varphi * \psi) = (u * \varphi) * \psi;$$

3) 卷积在 0 处的取值：

$$\big(u * \varphi\big)(0) = \langle u, \check{\varphi} \rangle,$$

其中 $\check{\varphi}(x) = \varphi(-x)$；

4) 卷积的导数：对任意的多重指标 α，我们有

$$\partial^\alpha (u * \varphi) = \partial^\alpha u * \varphi = u * \partial^\alpha \varphi.$$

证明*:* 我们逐条来证明：

1) 我们定义

$$F = \operatorname{supp}(u) + \operatorname{supp}(\varphi) = \{x + y \big| x \in \operatorname{supp}(u), y \in \operatorname{supp}(\varphi)\}.$$

很明显 F 是闭集，因为对于任意的 Cauchy 列 $\{x_i + y_i\}_{i \geqslant 1} \subset F$，其中，$x_i \in \operatorname{supp}(u), y_i \in \operatorname{supp}(\varphi)$，由于 $\operatorname{supp}(\varphi$ 是紧集，所以可以先选取子列（仍然记作）$\{y_i\}_{i \geqslant 1}$，使得 $y_i \to y$ 收敛，并且 $y \in \operatorname{supp}(\varphi)$；由于 $\{x_i + y_i\}_{i \geqslant 1}$ 是 Cauchy 列，所以 $\{x_i\}_{i \geqslant 1}$ 在闭集 $\operatorname{supp}(u)$ 中收敛到 $x \in \operatorname{supp}(u)$，这说明 $\{x_i + y_i\}_{i \geqslant 1}$ 的极限点仍然在 F 中。

为了说明该命题，我们只要证明对任意的 $x \notin F$（从而 x 附近的一个开领域中的点也不在 F 中），我们有

$$u * \varphi(x) = \langle u(y), \varphi(x - y) \rangle = 0.$$

我们证明上式的一个充分条件：

$$\operatorname{supp}(u) \cap \operatorname{supp}\left(\varphi(x - \cdot)\right) = \emptyset.$$

这是明显的，因为 $\operatorname{supp}\left(\varphi(x - \cdot)\right) = -\operatorname{supp}(\varphi) + x$。

[17]两个集合 $A, B \subset \mathbb{R}^n$ 的加法 $A + B$ 指的是：

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

2) 结合律的证明用到了上一个定理中分布与积分可交换的性质：

$$\begin{aligned}(u*\varphi)*\psi(x) &= \int_{\mathbb{R}^n}(u*\varphi)(y)\psi(x-y)dy\\ &= \int_{\mathbb{R}^n}\langle u(z),\varphi(y-z)\rangle\psi(x-y)dy\\ &= \int_{\mathbb{R}^n}\langle u(z),\varphi(y-z)\psi(x-y)\rangle dy\\ &= \left\langle u(z),\int_{\mathbb{R}^n}\varphi(y-z)\psi(x-y)dy\right\rangle.\end{aligned}$$

另外，

$$\begin{aligned}\big(u*(\varphi*\psi)\big)(x) &= \langle u(z),(\varphi*\psi)(x-z)\rangle\\ &= \left\langle u(z),\int_{\mathbb{R}^n}\varphi(x-z-y)\psi(y)dy\right\rangle.\end{aligned}$$

做变量替换 $y\mapsto x-y$ 就可以看到上面的两个等式是相等的。

3) 这是显然的：利用定义即可。

4) 我们证明第一个等号。事实上，我们有

$$\begin{aligned}\partial^\alpha u*\varphi(x) &= \Big\langle u(y),(-1)^{|\alpha|}\partial_y^\alpha\big(\varphi(x-y)\big)\Big\rangle\\ &= \big\langle u(y),\big(\partial^\alpha\varphi\big)(x-y)\big\rangle\\ &= u*\partial^\alpha\varphi.\end{aligned}$$

由于我们已经证明了第三项与第一项相等，所以命题成立。

□

作为应用卷积的应用，我们证明，在分布的意义下，任意一个分布可以被光滑函数所逼近：

定理 417. 我们任意选取标准的单位逼近 χ_ε，其中 $\varepsilon>0$。那么，对任意的 $u\in\mathcal{D}'(\mathbf{R}^n)$，当 $\varepsilon\to 0$ 时，我们都有

$$u*\chi_\varepsilon\xrightarrow{\mathcal{D}'}u.$$

另外，如果 $u\in\mathcal{E}'(\mathbb{R}^n)$，那么逼近序列 $u*\chi_\varepsilon\in C_0^\infty(\mathbb{R}^n)$。

证明：我们只要证明，对任意的试验函数 $\varphi(x)\in\mathcal{D}'(\mathbb{R}^n)$，有

$$\lim_{\varepsilon\to 0}\langle u*\chi_\varepsilon,\varphi\rangle=\langle u,\varphi\rangle.$$

根据上面第三个和第四个命题，我们有

$$\begin{aligned}\langle u*\chi_\varepsilon,\varphi\rangle &= \big((u*\chi_\varepsilon)*\check{\varphi}\big)(0)=\big(u*(\chi_\varepsilon*\check{\varphi})\big)(0)\\ &= \langle u,(\chi_\varepsilon*\check{\varphi})^{\check{}}\rangle=\langle u,(\chi_\varepsilon)^{\check{}}*\varphi\rangle.\end{aligned}$$

容易证明（请参考本次作业），当 $\varepsilon \to 0$ 时，我们有

$$(\chi_\varepsilon)^\vee * \varphi \xrightarrow{\mathcal{D}} \varphi.$$

从而定理得证。 □

注记. 对任意的分布 $u \in \mathcal{D}'(\mathbb{R}^n)$，我们通过如下的式子来定义 $\check{u}$:

$$\langle \check{u}, \varphi \rangle = \langle u, \check{\varphi} \rangle,$$

其中 φ 是试验函数。

我们对卷积做推广，来定义一个分布与一个有紧支集的分布的卷积:

$$\mathcal{D}'(\mathbb{R}^n) * \mathcal{E}'(\mathbb{R}^n) \xrightarrow{*} \mathcal{D}'(\mathbb{R}^n).$$

为此，我们选取 $u \in \mathcal{D}'(\mathbb{R}^n)$，$\psi \in \mathcal{D}(\mathbb{R}^n) \subset \mathcal{E}'(\mathbb{R}^n)$ 来做计算（实际上，我们选取所有的函数都在 $\mathcal{D}(\mathbb{R}^n)$ 中也可以）。此时，我们有如下等式

$$\langle u * \psi, \varphi \rangle = \langle u, \check{\psi} * \varphi \rangle.$$

这个等式实际上在前一个定理的证明中已经证明了，其中 $\psi = \chi_\varepsilon$。

根据上面等式的启发，我们定义

定理 418. 任意给定分布 $u \in \mathcal{D}'(\mathbb{R}^n)$ 和有紧支集的分布 $c \in \mathcal{D}'(\mathbb{R}^n)$，如下定义的 $u * c$:

$$\langle u * c, \varphi \rangle = \langle u, \check{c} * \varphi \rangle,$$

是 $\mathbb{R}^d$ 上的分布，其中 $\varphi \in \mathcal{D}(\mathbb{R}^n)$ 是任意的试验函数。

注记. 这个定义给出了卷积的定义:

$$\mathcal{D}'(\mathbb{R}^n) * \mathcal{E}'(\mathbb{R}^n) \xrightarrow{*} \mathcal{D}'(\mathbb{R}^n).$$

证明: 根据分布的定义，我们要证明，任意给定紧集 $K \subset \mathbb{R}^n$，存在常数 C 和 p，使得对每个试验函数 $\varphi \in C_K^\infty(\mathbb{R}^n)$，都有不等式

$$|\langle u, \check{c} * \varphi \rangle| \leqslant C \sup_{|\alpha| \leqslant p} \|\partial^\alpha \varphi\|_{L^\infty}.$$

首先，我们注意到 $\operatorname{supp}(\check{c} * \varphi) \subset L$，其中，$L = \operatorname{supp}(\check{c}) + K$。我们注意到，$L$ 为紧集（有界闭集）。从而，根据 u 是分布的定义，存在常数 D 和 q（只依赖于 L，从而只依赖于 K），使得

$$|\langle u, \check{c} * \varphi \rangle| \leqslant D \sup_{|\beta| \leqslant q} \|\partial^\beta \big(\check{c} * \varphi\big)\|_{L^\infty}.$$

我们计算 $\partial^\beta\big(\check{c}*\varphi\big)$：

$$\partial^\beta\big(\check{c}*\varphi\big)(x)=\check{c}*\partial^\beta\varphi=\Big\langle c,\partial^\beta\varphi(x+\cdot)\Big\rangle.$$

此时，由于 c 是有紧支集的分布，所以存在 D' 和 q'，使得

$$|\partial^\beta\big(\check{c}*\varphi\big)|\leqslant D'\sup_{|\beta|\leqslant q'}\|\partial^{\beta+\beta'}\varphi\|_{L^\infty}.$$

最终，我们选取 $C=DD'$ 和 $p=q+q'$ 就证明了 $u*c$ 是分布。 □

例子. (Dirac 函数与平移算子) 给定分布 $u\in\mathcal{D}'(\mathbb{R}^n)$，给定 $a\in\mathbb{R}^n$，我们定义了该分布的平移 $\tau_a u$：

$$\langle\tau_a u,\varphi\rangle=\langle u,\varphi(\cdot+a)\rangle.$$

我们来证明：

$$u*\delta_a=\tau_a u.$$

实际上，我们有 $\check{\delta}_a=\delta_{-a}$，所以

$$\langle u*\delta_a,\varphi\rangle=\langle u,\delta_a*\varphi\rangle.$$

所以，$\delta_{-a}*\varphi=\varphi(x+a)$ 即可。根据定义，

$$\big(\delta_a*\varphi\big)(x):=\langle\delta_{-a},\varphi(x-\cdot)\rangle=\varphi(x+a).$$

所以，命题成立。

特别地，对任意的分布 $u\in\mathcal{D}'(\mathbb{R}^n)$，我们有

$$u*\delta_0=u.$$

另外，我们还可以用卷积来还原 $\mathbb{R}^n$ 上的群结构：

$$\delta_a*\delta_b=\delta_{a+b}.$$

对于分布之间的卷积，我们也有类似于之前的性质：

命题 419. 对任意的分布 $u\in\mathcal{D}'(\mathbb{R}^n)$ 和 $c\in\mathcal{E}'(\mathbb{R}^n)$，我们有

1) 卷积之后的支集满足：$\operatorname{supp}(u*c)\subset\operatorname{supp}(u)+\operatorname{supp}(c)$。特别地，我们有

$$\mathcal{E}'(\mathbb{R}^n)*\mathcal{E}'(\mathbb{R}^n)\xrightarrow{*}\mathcal{E}'(\mathbb{R}^n).$$

2) 卷积与求导数可交换：对任意的多重指标 α，我们有

$$\partial^\alpha(u*c)=(\partial^\alpha u)*c=u*(\partial^\alpha c).$$

证明：我们逐条来证明：

1) 我们令 $F=\operatorname{supp}(u)+\operatorname{supp}(c)$，这是一个闭集（证明用到了 $\operatorname{supp}(c)$ 是紧的，请参考之前的论证）。考虑任意一个试验函数 φ，我们令 $K=\operatorname{supp}(\varphi)$，假设 $K\cap F=\emptyset$。那么，

$$\operatorname{supp}(u)\cap(K+\operatorname{supp}(\check{c}))=\emptyset.$$

这表明 $\operatorname{supp}(u)$ 与 $\operatorname{supp}\check{c}*\varphi$ 是不想交的，所以

$$\langle u*c,\varphi\rangle=\langle u,\check{c}*\varphi\rangle=0.$$

2) 我们利用定义直接计算：

$$\begin{aligned}\langle\partial^{\alpha}(u*c),\varphi\rangle&=\langle u*c,(-1)^{|\alpha|}\partial^{\alpha}\varphi\rangle=\langle u,(-1)^{|\alpha|}\check{c}*\partial^{\alpha}\varphi\rangle\\&=\langle u,(-1)^{|\alpha|}\partial^{\alpha}\big(\check{c}*\varphi\big)\rangle=\langle\partial^{\alpha}u,\check{c}*\varphi\rangle\\&=\langle(\partial^{\alpha}u)*c,\varphi\rangle.\end{aligned}$$

另外，我们还可以如下计算：

$$\begin{aligned}\langle\partial^{\alpha}(u*c),\varphi\rangle&=\langle u,(-1)^{|\alpha|}\check{c}*\partial^{\alpha}\varphi\rangle=\langle u,(-1)^{|\alpha|}\big(\partial^{\alpha}\check{c}\big)*\varphi\rangle\\&=\langle u,\big(\partial^{\alpha}c\big)\check{}*\varphi\rangle\\&=\langle u*\left(\partial^{\alpha}c\right),\varphi\rangle.\end{aligned}$$

命题得证。

□

63 卷积的连续性，基本解，Cauchy-Riemann 方程的椭圆正则性，热方程的基本解

二零二零年九月三十日，星期三，晴

我们对于

$$\mathcal{E}'(\mathbb{R}^n) * \mathcal{E}'(\mathbb{R}^n) \xrightarrow{*} \mathcal{E}'(\mathbb{R}^n).$$

来证明交换律：

命题 420. *对任意的 $c_1, c_2 \in \mathcal{E}'(\mathbb{R}^n)$，我们有*

$$c_1 * c_2 = c_2 * c_1.$$

证明： 首先定义

$$c_{1,\varepsilon} = c_1 * \chi_\varepsilon.$$

根据之前证明的逼近引理，在分布的意义下，我们有极限：

$$c_{1,\varepsilon} \xrightarrow{\mathcal{D}'} c_1, \quad \varepsilon \to 0.$$

另外，我们知道 $\check{c}_2 * \varphi$ 是有紧支集的光滑函数，所以可以作为试验函数。从而，

$$\begin{aligned}
\langle c_{1,\varepsilon}, \check{c}_2 * \varphi \rangle &= \int_{\mathbb{R}^n} c_{1,\varepsilon}(x) \check{c}_2 * \varphi(x) dx = \int_{\mathbb{R}^n} c_{1,\varepsilon}(x) \langle \check{c}_2, \varphi(x - \cdot) \rangle dx \\
&= \left\langle \check{c}_2, \int_{\mathbb{R}^n} c_{1,\varepsilon}(x) \varphi(x - \cdot) dx \right\rangle \\
&= \left\langle \check{c}_2, \underbrace{\int_{\mathbb{R}^n} c_{1,\varepsilon}(x) \check{\varphi}(\cdot - x) dx}_{\text{写成卷积的形式}} \right\rangle \\
&= \langle \check{c}_2, c_1 * (\chi_\varepsilon * \check{\varphi}) \rangle \\
&= \langle c_2, \big(c_1 * (\chi_\varepsilon * \check{\varphi})\big)\check{} \rangle \\
&= \langle c_2, \check{c}_1 * (\check{\chi}_\varepsilon * \varphi) \rangle \\
&= \langle c_2 * c_1, \check{\chi}_\varepsilon * \varphi \rangle.
\end{aligned}$$

另外，当 $\varepsilon \to 0$ 时，我们有极限

$$\begin{aligned}
\langle c_{1,\varepsilon}, \check{c}_2 * \varphi \rangle &\to \langle c_1 * c_2, \varphi \rangle, \\
\langle c_2 * c_1, \check{\chi}_\varepsilon * \varphi \rangle &\to \langle c_2 * c_1, \varphi \rangle.
\end{aligned}$$

这就证明了命题。 □

定理 421 (卷积的连续性)**.** 给定有紧支集的分布的序列 $\{c_k\}_{k\geqslant 1} \subset \mathcal{E}'(\mathbb{R}^n)$，我们假设存在紧集 $K \subset \mathbb{R}^n$，使得对任意的 $k \geqslant 1$，有 $\operatorname{supp}(c_k) \subset K$；给定分布的序列 $\{u_k\}_{k\geqslant 1} \subset \mathcal{D}'(\mathbb{R}^n)$。假设存在分布 $c, u \in \mathcal{D}'(\mathbb{R}^n)$，使得在分布的意义下，我们有极限

$$c_k \xrightarrow{\mathcal{D}'} c, \quad u_k \xrightarrow{\mathcal{D}'} u, \quad k \to \infty.$$

那么，当 $k \to \infty$ 时，在分布的意义下，我们有

$$u * c_k \xrightarrow{\mathcal{D}'} u * c, \quad u_k * c \xrightarrow{\mathcal{D}'} u * c.$$

证明: 先证明第二个极限。任意选取试验函数 $\varphi \in \mathcal{D}(\mathbb{R}^n)$，那么，$\check{c} * \varphi \in \mathcal{D}(\mathbb{R}^n)$ 也是试验函数。所以，以下的计算是良好定义的:

$$\langle u_k * c, \varphi\rangle = \langle u_k, \check{c} * \varphi\rangle \to \langle u, \check{c} * \varphi\rangle = \langle u * c, \varphi\rangle.$$

这就证明了第二个极限。

为了证明第一个极限，我们先选定一个有紧支集的光滑函数 ψ，使得 ψ 在紧集 $\operatorname{supp}(\varphi) - K$ 上恒等于 1（我们要用这个函数来记住 $\check{c}_k * \varphi$ 的支集，这是常用的技巧）。那么，

$$\langle u * c_k, \varphi\rangle = \langle u, \check{c}_k * \varphi\rangle = \langle u, \psi \cdot \check{c}_k * \varphi\rangle.$$

根据有紧支集分布的卷积的交换性，我们就有

$$\langle u * c_k, \varphi\rangle = \langle (\psi u) * c_k, \varphi\rangle = \langle c_k * (\psi u), \varphi\rangle \to \langle c * (\psi u), \varphi\rangle.$$

其中，最后一步的极限我们用到了刚证明的结论。最终，我们再对 $\langle c * (\psi u), \varphi\rangle$ 把上面的计算反着做一遍:

$$\begin{aligned}\langle c * (\psi u), \varphi\rangle &= \langle (\psi u) * c, \varphi\rangle = \langle u, \psi \cdot \check{c} * \varphi\rangle \\ &= \langle u, \check{c} * \varphi\rangle = \langle u * c, \varphi\rangle.\end{aligned}$$

这就证明了定理。 □

注记. 这个定理证明了对卷积的每个分量的连续性而不是对两个分量的连续性。直观上，我们可以不介意这一点，通过把每个分量与 χ_ε 做卷积，我们可以想象卷积

$$\mathcal{D}'(\mathbb{R}^n) \times \mathcal{E}'(\mathbb{R}^n) \xrightarrow{*} \mathcal{D}'(\mathbb{R}^n)$$

可以由它在子空间

$$C^\infty(\mathbb{R}^n) \times C_0^\infty(\mathbb{R}^n) \to \mathcal{D}'(\mathbb{R}^n)$$

上的限制决定。在这个字空间上，所有的计算都是直接用积分计算的。

推论 422 (卷积的结合律)**.** 对于 $c_1, c_2 \in \mathcal{E}'(\mathbb{R}^n)$ 和 $u \in \mathcal{D}'(\mathbb{R}^n)$，我们有

$$(u * c_1) * c_2 = u * (c_1 * c_2).$$

证明: 我们定义 $c_{2,\varepsilon} = c_2 * \chi_\varepsilon$，从而，我们有

$$\lim_{\varepsilon\to 0} c_{2,\varepsilon} \overset{D'}{=} c_2.$$

根据上述逼近定理，我们只要证明

$$(u * c_1) * c_{2,\varepsilon} = u * (c_1 * c_{2,\varepsilon})$$

即可。所以，我们不妨假设 $c_2 \in C_0^\infty(\mathbb{R}^n)$。

类似地，我们定义 $c_{1,\varepsilon} = c_1 * \chi_\varepsilon$，那么，

$$\lim_{\varepsilon\to 0} c_{1,\varepsilon} \overset{D'}{=} c_1.$$

所以，

$$\lim_{\varepsilon\to 0} u * c_{1,\varepsilon} \overset{D'}{=} u * c_1.$$

再用上述逼近定理，我们只要证明

$$(u * c_{1,\varepsilon}) * c_2 = u * (c_{1,\varepsilon} * c_2)$$

即可。所以，我们又可以假设 $c_1 \in C_0^\infty(\mathbb{R}^n)$。

当 $c_1, c_2 \in C_0^\infty(\mathbb{R}^n)$ 时，定理已经被证明过。 □

我们可以把 $C^\infty(\mathbb{R}^n)$ 中的函数看作是分布，从而可以于 $\mathcal{E}'(\mathbb{R}^n)$ 中的分布做卷积。我们现在说明，这样得到的分布是光滑函数:

定理 423. *对任意的 $c \in \mathcal{E}'(\mathbb{R}^n)$ 和 $f \in C^\infty(\mathbb{R}^n)$，分布 $f * c$ 是光滑函数，即*

$$C^\infty(\mathbb{R}^n) * \mathcal{E}'(\mathbb{R}^n) \xrightarrow{*} C^\infty(\mathbb{R}^n).$$

特别地，我们有如下的公式:

$$f * c(x) = \langle c, f(x - \cdot)\rangle.$$

其中，我们用到了有紧支集的分布可以和任意的光滑函数进行配对。

证明: 首先回忆一下，在上个学期关于 Stokes 公式的第一个证明中，我们用到了如下的单位分解定理（参考第 47 次讲义）:

引理 424 (周期的单位分解)**.** *对任意的 $N \geqslant 1$，令 $\Gamma_N = 2^{-(N+1)}\mathbb{Z}^n$。那么，存在一族有紧支集的非负的光滑函数 $\{\chi_{\mathbf{k}}(x)\}_{\mathbf{k}\in\Gamma_N}$，满足如下的两个条件:*

1) 对任意的 $\mathbf{k} \in \Gamma_N$，$\mathrm{supp}\chi_{\mathbf{k}}$ 落在以 $\mathbf{k}$ 为中心边长为 2^{-N+1} 的正方体中;

3) 常值函数 1 可以分解为:

$$1 = \sum_{\mathbf{k}\in\Gamma_N} \chi_{\mathbf{k}}.$$

令 $N=0$，那么我们就可以得到非负的光滑函数 $\{\psi_q\}_{q\in\mathbb{Z}^n}$，使得

$$\sum_{q\in\mathbb{Z}^n}\psi_q\equiv 1,$$

并且对任意的 $q\in\mathbb{Z}^n$，$\mathrm{supp}(\psi_q)$ 落在以 q 为中心边长为 2 的正方体中。利用这个单位分解，我们把 f 写成

$$f=\sum_{q\in\mathbb{Z}^n}f_q=\sum_{q\in\mathbb{Z}^n}\psi_q\cdot f.$$

我们强调一下，以上的求和在每个点处都是有限求和。从而，

$$\langle f*c,\varphi\rangle=\langle f,\overbrace{\check{c}*\varphi}^{\text{紧支集}}\rangle=\left\langle\sum_{q\in\mathbb{Z}}f_p,\check{c}*\varphi\right\rangle=\left\langle\sum_{\text{有限和}}f_p,\check{c}*\varphi\right\rangle.$$

有限和自然与分布（= 积分）是可交换的，所以

$$\begin{aligned}\langle f*c,\varphi\rangle&=\sum_{\text{有限和}}\langle f_p*c,\varphi\rangle\\&=\sum_{\text{有限和}}\langle c*f_p,\varphi\rangle\\&=\sum_{\text{有限和}}\int_{\mathbb{R}^n}\langle c,f_p(x-\cdot)\rangle\varphi(x)dx\\&=\int_{\mathbb{R}^n}\langle c,f(x-\cdot)\rangle\varphi(x)dx.\end{aligned}$$

再次交换求和与积分，我们得到

$$\begin{aligned}\langle f*c,\varphi\rangle&=\int_{\mathbb{R}^n}\left\langle c,\sum_{\text{有限和}}f_p(x-\cdot)\right\rangle\varphi(x)dx\\&=\int_{\mathbb{R}^n}\langle c,\sum_{p\in\mathbb{Z}}f(x-\cdot)\rangle\varphi(x)dx\\&=\int_{\mathbb{R}^n}\langle c,f(x-\cdot)\rangle\varphi(x)dx.\end{aligned}$$

这就给出了所要证明的公式。

最终，为了证明 $f*c$ 的光滑性，我们只要说明 $x\mapsto\langle c,f(x-\cdot)\rangle$ 是光滑的即可，这就是对参数求导的定理的简单应用，我们把证明留作作业。 □

分布的初步应用：某些微分方程的基本解

应用举例：解析函数的正则性

我们已经证明了，在分布的意义下，有如下的等式

$$\overline{\partial}\left(\frac{1}{\pi z}\right)\overset{\mathcal{D}'}{=}\delta_0.$$

其中，

$$\overline{\partial} = \frac{\partial}{\partial \overline{z}} = \frac{1}{2}\left(\frac{\partial}{\partial x} + i\frac{\partial}{\partial y}\right).$$

另外，对于 $\Omega \subset \mathbb{C}$ 上的复值连续可微的函数 f，如果它满足 $\overline{\partial} f \equiv 0$，我们就称 f 是复解析的函数。我们下面证明，复解析的函数一定是光滑函数。这是一个不平凡的事实，因为我们一开始只假设了它是 C^1 的。事实上，我们有更强的结论：

定理 425 (椭圆正则性)**.** *给定非空的开区域 $\Omega \subset \mathbb{C}$。假设分布 $u \in \mathcal{D}'(\Omega)$ 满足 Cauchy-Riemann 方程*

$$\overline{\partial} u = 0.$$

那么，u 是光滑函数，即 $u \in C^\infty(\Omega)$。

在证明 Cauchy-Riemann 方程的椭圆正则性定理之前，我们先证明如下的引理：

引理 426. *给定有紧支集的分布 $c \in \mathcal{E}'(\mathbb{R}^2) = \mathcal{E}'(\mathbb{C})$，那么，$\frac{1}{z} * c$ 在 c 的支集之外是光滑的，即*

$$\frac{1}{z} * c \in C^\infty\left(\mathbb{R}^2 - \operatorname{supp}(c)\right).$$

证明: 我们选取非负的 χ，使得它的支集在半径为 1 的小球之内并且其积分恰好等于 1。利用这个截断函数，我们把 $\frac{1}{z} * c$ 写成两部分：

$$u = \frac{1}{z} * c = \underbrace{\left(\frac{\chi_\varepsilon}{z}\right) * c}_{\text{支集} \subset B_\varepsilon + \operatorname{supp}(c)} + \underbrace{\left(\frac{1-\chi_\varepsilon}{z}\right) * c}_{\text{光滑}}.$$

第一部分中，根据支集在卷积下的关系，我们有

$$\operatorname{supp}\left[\left(\frac{\chi_\varepsilon}{z}\right) * c\right] \subset \operatorname{supp}\left(\frac{\chi_\varepsilon}{z}\right) + \operatorname{supp}(c) \subset B_\varepsilon + \operatorname{supp}(c),$$

其中，B_ε 是中心在原点半径为 ε 的开球。

第二部分中，由于 $\frac{1-\chi_\varepsilon}{z}$ 是光滑函数，所以，这一部分的卷积实际上是光滑函数，它对问题本身没有影响，可以忽略。

综合上面的叙述，u 至少在第一部分的支集之外（等于 0 加上一个光滑函数）是光滑的。也就是说，u 在 $\mathbb{R}^2 - B_\varepsilon - \operatorname{supp}(c)$ 上光滑。令 $\varepsilon \to 0$，我们就说明了 u 在 $\mathbb{R}^2 - \operatorname{supp}(c)$ 上光滑。 □

椭圆正则性定理的证明. 光滑性是一个局部性质，所以我们只要证明 u 在每个点的附近为光滑函数即可：我们任选点 $z_0 \in \Omega$ 以及 z_0 处的一个小开球 $B(z_0, 2\varepsilon)$，使得 $B(z_0, 2\varepsilon) \subset \Omega$，其中 $\varepsilon > 0$。然后，选取 $\mathbb{C}$ 上的光滑函数 $\theta(z)$，使得

$$\begin{cases} 0 \leqslant \theta(z) \leqslant 1, \text{ 对任意的 } z \in \mathbb{C} \\ \theta\big|_{B(z_0,\varepsilon)} \equiv 1; \\ \theta\big|_{\mathbb{C} - B(z_0,2\varepsilon)} \equiv 0. \end{cases}$$

根据上面的构造，我们知道，

$$\partial_x\theta\big|_{B(z_0,\varepsilon)} \equiv \partial_y\theta\big|_{B(z_0,\varepsilon)} \equiv 0.$$

我们只需要证明 $\theta\cdot u$ 光滑即可，因为这就说明 u 在 $B(z_0,\varepsilon)$ 上光滑。

利用卷积的基本性质，我们有如下的计算

$$\begin{aligned}\theta\cdot u &= \delta_0 * (\theta\cdot u) = \overline{\partial}(\frac{1}{\pi z}) * (\theta\cdot u)\\ &= \frac{1}{\pi z} * \big(\overline{\partial}(\theta\cdot u)\big)\\ &= \frac{1}{\pi z} * \big(\overline{\partial}\theta\cdot u\big)\end{aligned}$$

根据上面的引理，$\theta\cdot u$ 在 $\operatorname{supp}\big(\overline{\partial}\theta\cdot u\big)$ 之外是光滑的。由于 $\partial\theta\cdot u\big|_{B(z_0,\varepsilon)} \equiv 0$，从而，$\theta\cdot u$ 在 $B(z_0,\varepsilon)$ 上光滑，这就证明了定理。 □

线性微分算子与基本解

给定区域 $\Omega\subset\mathbb{R}^n$。给定非负整数 $m\geqslant 0$。对每个满足 $|\alpha|\leqslant m$ 的多重指标 α，我们指定一个复值的光滑函数 $a_\alpha(x)\in C^\infty(\Omega)$（关于这些函数的光滑性在很多应用中可能会放松为其他的条件，比如说，只要求它们是 L^∞ 的函数等）。我们还要求至少有一个 $a_\alpha(x)\neq 0$，其中，$|\alpha|=m$。此时，我们称如下的算子

$$P=\sum_{|\alpha|\leqslant m} a_\alpha(x)\partial^\alpha$$

为一个次数为 m 的**线性微分算子**。如果 $a_\alpha(x)$ 均为常数，我们就把 P 称作是**常系数（线性）微分算子**。另外，如果对所有的 $|\alpha|<m$，$a_\alpha(x)\equiv 0$，我们就说 P 是m**-次齐次的线性微分算子**。

作为例子，我们上面研究的 $P=\overline{\partial}$ 就是常系数微分算子。

很明显，对任意的分布 $u\in\mathcal{D}'(\Omega)$，我们可以用 P 来作用：

$$P:\mathcal{D}'(\Omega)\to\mathcal{D}'(\Omega),\quad u\mapsto Pu=\sum_{|\alpha|\leqslant m} a_\alpha(x)\partial^\alpha u.$$

我们定义 P 的伴随算子为 ${}^tP$，其中，按照定义，它写为

$${}^tP=\sum_{|\alpha|\leqslant m}(-1)^{|\alpha|}\partial^\alpha\circ a_\alpha(x).$$

也就是说，对任意的分布 u，我们有

$${}^tPu=\sum_{|\alpha|\leqslant m}(-1)^{|\alpha|}\partial^\alpha\left(a_\alpha(x)u\right)$$

通过 Leibniz 公式，我们很容易验证这也是一个线性微分算子，并且其次数也是 m。很明显，对任意的分布 $u\in\mathcal{D}'(\Omega)$，对任意的试验函数 $\varphi\in\mathcal{D}(\Omega)$，我们都有

$$\langle Pu,\varphi\rangle=\langle u,{}^tP\varphi\rangle.$$

定义 427. 给定 $\mathbb{R}^n$ 上的常系数微分算子 P。如果分布 $E \in \mathcal{D}'(\mathbb{R}^n)$，满足

$$P(E) = \delta_0,$$

我们就称 E 是 P 的一个**基本解**。

注记. 在上个世纪 *50* 年代末，*B.Malgrange* 和 *L.Ehrenpreis* 各自独立证明了：$\mathbb{R}^n$ 每一个常数系数的微分算子都有基本解。我们在课程上不去证明这个定理。

例子. 我们先给出三个例子：

1) Cauchy-Riemann 算子 $\overline{\partial}$。

这是 $\mathbb{C}$ 上的线性微分算子，我们已经证明了 $\dfrac{1}{\pi z}$ 是 $\overline{\partial}$ 的一个基本解。

2) Laplace 算子 $\triangle$。

按照定义，

$$\triangle = \sum_{k=1}^{n} \frac{\partial^2}{\partial x_k^2}.$$

我们将在作业中证明，如下的函数是 $\triangle$ 的基本解：

$$\begin{cases} E(x) = -\dfrac{1}{2\pi} \log |x|, & n = 2; \\ E(x) = \dfrac{1}{(2-n)|\mathbf{S}^{n-1}|} \dfrac{1}{|x|^{n-2}}, & n \geqslant 3. \end{cases}$$

在上式中，$|\mathbf{S}^{n-1}|$ 表示的是 $\mathbb{R}^n$ 中单位球面的测度。这些函数都是局部可积的。

我们将在作业中重点研究这个例子。

3) 热算子 $\partial_t - \triangle$。

热算子是定义在 $\mathbb{R} \times \mathbb{R}^n$ 上的算子，其中，第一个坐标是时间 t 的坐标。它作用在以 $(t, x) \in \mathbb{R} \times \mathbb{R}^n$ 为变量函数上的。如果用 $H(t)$ 表示 *Heaviside* 函数（即 $\mathbf{1}_{\mathbb{R}_{\geqslant 0}}$），那么如下的**热核**函数是 $\partial_t - \triangle$ 的基本解：

$$E(t, x) = \frac{H(t)}{(4\pi t)^{\frac{n}{2}}} e^{-\frac{|x|^2}{4t}}.$$

热算子基本解的研究

我们来研究热算子 $\partial_t - \triangle$ 的基本解 $E(t, x)$ 的性质，其中

$$E(t, x) = \frac{H(t)}{(4\pi t)^{\frac{n}{2}}} e^{-\frac{|x|^2}{4t}}.$$

至于如何找到 E 的表达式，同学们要稍有耐心：我们将从 Fourier 变换的观点来给出 E 的构造。历史上，Fourier 也正是为了研究热传导才发明了 Fourier 变换这个工具。

首先，我们证明 $E(t,x)\in L^1_{\mathrm{loc}}(\mathbb{R}\times\mathbb{R}^n)$，从而，这是一个良好定义的分布，即 $E(t,x)\in\mathcal{D}'(\mathbb{R}\times\mathbb{R}^n)$。

实际上，我们观察到这时一个正函数并且在 $t<0$ 处恒为零；在 $t>0$ 处光滑。我们在 $t=0$ 附近研究它的性质即可。实际上，对于任意的 $t>0$，做变量替换 $x=\sqrt{t}y$ 我们知道

$$\int_{\mathbb{R}^n}E(t,x)dx=\int_{\mathbb{R}^n}\frac{1}{(4\pi)^{\frac{n}{2}}}e^{\frac{|y|^2}{4}}dy$$

是常数。所以，$E(t,x)$ 在 $[-A,A]\times\mathbb{R}^n$ 上可积分，这说明 E 是局部可积的。

另外，我们知道对任意的 $t\neq0$，$E(t,x)$ 是光滑的。所以，通过直接计算，我们就有

$$(\partial_t-\triangle)(E(t,x))=0,\quad t\neq0.$$

另外，固定 $(0,x_0)$，其中，$x_0\neq0$，此时，由于当 t 从 >0 的地方趋向于 0 时，$e^{-\frac{|x_0|^2}{4t}}$ 趋于 0 的速度比任意的 t 的幂次都要快，所以，$E(0,x_0)$ 在这个点处的 t 的所有阶的偏导数都是 0。据此，不难说明，$E(t,x)$ 实际上在 $\mathbb{R}\times\mathbb{R}^2-\{(0,0)\}$ 上是光滑的。利用连续性以及上面的计算，我们得到

$$(\partial_t-\triangle)(E(t,x))=0,\quad \text{对任意的 } (t,x)\neq(0,0).$$

我们把这个细节留作作业（这是一个经典的求导数题目，每位学过微积分的同学都应该掌握。）

为了证明在分布的意义下有

$$(\partial_t-\triangle)(E(t,x))\overset{\mathcal{D}'}{=}\delta_{(0,0)},$$

我们对任意的试验函数 $\varphi(t,x)\in\mathcal{D}(\mathbb{R}\times\mathbb{R}^n)$，计算

$$\begin{aligned}\langle(\partial_t-\triangle)(E(t,x)),\varphi(t,x)\rangle&=-\int_0^\infty\int_{\mathbb{R}^n}E(t,x)\partial_t\varphi(t,x)dtdx-\int_0^\infty\int_{\mathbb{R}^n}E(t,x)\triangle\varphi(t,x)dtdx\\&=-\lim_{\varepsilon\to0}\underbrace{\int_\varepsilon^\infty\int_{\mathbb{R}^n}E(t,x)\partial_t\varphi(t,x)dtdx}_{I_\varepsilon}-\lim_{\varepsilon\to0}\underbrace{\int_\varepsilon^\infty\int_{\mathbb{R}^n}E(t,x)\triangle\varphi(t,x)dtdx}_{J_\varepsilon}.\end{aligned}$$

在上面的极限中，我们用到了 $E(t,x)\geqslant0$ 在 $[0,A]$ 上是可积分的，从而（乘以一个常数）可以作为控制函数来用 Lebesgue 控制收敛定理。

我们来计算 I_ε。这里，最重要的想法是一旦我们在 $t>0$ 处，这些计算实际上是光滑的，从而，我们可以利用 $E(t,x)$ 所满足的方程：

$$\begin{aligned}I_\varepsilon&\overset{Funibi}{=}\int_{\mathbb{R}^n}\left(\int_\varepsilon^\infty E(t,x)\partial_t\varphi(t,x)dt\right)dx\\&=-\int_{\mathbb{R}^n}\left(\int_\varepsilon^\infty\partial_tE(t,x)\varphi(t,x)dt\right)dx-\int_{\mathbb{R}^n}E(\varepsilon,x)\varphi(\varepsilon,x)dx\\&\overset{\text{热方程}}{=}-\int_{\mathbb{R}^n}\left(\int_\varepsilon^\infty\triangle E(t,x)\varphi(t,x)dt\right)dx-\int_{\mathbb{R}^n}E(\varepsilon,x)\varphi(\varepsilon,x)dx\\&\overset{Funibi}{=}-J_\varepsilon-\int_{\mathbb{R}^n}E(\varepsilon,x)\varphi(\varepsilon,x)dx.\end{aligned}$$

所以，

$$\begin{aligned} I_\varepsilon + J_\varepsilon &= -\int_{\mathbb{R}^n} E(\varepsilon, x)\varphi(\varepsilon, x)dx \\ &= -\int_{\mathbb{R}^n} \frac{1}{(4\pi\varepsilon)^{\frac{n}{2}}} e^{-\frac{|x|^2}{4\varepsilon}} \varphi(\varepsilon, x)dx \\ &= -\frac{1}{(4\pi)^{\frac{n}{2}}} \int_{\mathbb{R}^n} e^{-\frac{|y|^2}{4}} \varphi(\varepsilon, \sqrt{\varepsilon}y)dy. \end{aligned}$$

由于

$$\frac{1}{(4\pi)^{\frac{n}{2}}} \int_{\mathbb{R}^n} e^{-\frac{|y|^2}{4}} dy = 1,$$

所以，

$$-(I_\varepsilon + J_\varepsilon) = -\frac{1}{(4\pi)^{\frac{n}{2}}} \int_{\mathbb{R}^n} e^{-\frac{|y|^2}{4}} \left(\varphi(\varepsilon, \sqrt{\varepsilon}y) - \varphi(0,0)\right) dy + \varphi(0,0).$$

根据 Lagrange 中值定理，我们就有

$$\left|\varphi(\varepsilon, \sqrt{\varepsilon}y) - \varphi(0,0)\right| \leqslant 2\|\nabla\varphi\|_{L^\infty} \left(\varepsilon + \sqrt{\varepsilon}|y|\right).$$

所以，

$$\begin{aligned} &\left| \frac{1}{(4\pi)^{\frac{n}{2}}} \int_{\mathbb{R}^n} e^{-\frac{|y|^2}{4}} \left(\varphi(\varepsilon, \sqrt{\varepsilon}y) - \varphi(0,0)\right) dy \right| \\ \leqslant & 2\|\nabla\varphi\|_{L^\infty} \left(\frac{\varepsilon}{(4\pi)^{\frac{n}{2}}} \int_{\mathbb{R}^n} e^{-\frac{|y|^2}{4}} dy + \frac{\sqrt{\varepsilon}}{(4\pi)^{\frac{n}{2}}} \int_{\mathbb{R}^n} |y| e^{-\frac{|y|^2}{4}} dy \right) \end{aligned}$$

由于上面两个积分都有有限的，所以上面的式子是一个 $O(\sqrt{\varepsilon})$ 项，从而

$$\lim_{\varepsilon \to 0} -(I_\varepsilon + J_\varepsilon) = \varphi(0,0) = \langle \delta_{0.0}, \varphi \rangle.$$

这就证明了

$$(\partial_t - \triangle)(E(t,x)) \overset{\mathcal{D}'}{=} \delta_{(0,0)}.$$

利用基本解和卷积，我们可以解常系数的偏微分方程：

定理 428. *给定 $\mathbb{R}^n$ 上的常数系数微分算子 P，我们假设它有基本解 $E \in \mathcal{D}'(\mathbb{R}^n)$。那么，对任意的 $c \in \mathcal{E}'(\mathbb{R}^n)$，存在分布 $u \in \mathcal{D}'(\mathbb{R}^n)$，使得*

$$P(u) = c.$$

证明： 我们定义 $u = E * c$，这是一个分布。利用基本解的定义以及卷积的性质，我们有

$$P(u) = P(E * c) = \big(P(E)\big) * c = \delta_0 * c = c.$$

这表明，u 是一个解。 □

63.1 作业：分布的例子，Laplace 算子、位势方程与分布

清华大学 19-20 春季学期，数学分析三，作业 2

请用 A4 大小的纸张**正反面**用钢笔，签字笔或者圆珠笔书写，并注明自己的姓名，年级和作业的**总页数**。除定理公式所涉及的人名之外，请使用**中文**。本次作业的提交时间和地点为 **10 月 14 日**（周三）上午的课堂上，逾期视作零分。

分布的定义、例子与操作

习题 A.（课堂细节的补充） 我们总假设 $\Omega \subset \mathbb{R}^n$ 是非空的开集。

A0) 假设 $K \subset \Omega$ 是紧集，那么，存在 $\varphi \in C_0^\infty(\Omega)$ 和开集 V，使得

$$K \subset V \subset U;\ \varphi\big|_V \equiv 1.$$

这是课上用到的一个技术性引理。如果你觉得不安心的话，请证明它。

A1) 证明如下两个常用的收敛定理：

- 假设 $\{f_p\}_{p\geqslant 1}$ 是 Ω 上的一列局部可积的函数，$f \in L^1_{\mathrm{loc}}(\Omega)$。如果对任意的紧集 $K \subset \Omega$，我们有

  $$\lim_{p\to\infty} f_p\big|_K \overset{L^1(K)}{=} f\big|_K,$$

  那么，作为分布，我们有

  $$\lim_{p\to\infty} f_p \overset{\mathcal{D}'}{=} f.$$

- 假设 $\{u_p\}_{p\geqslant 1}$ 是 Ω 上分布的序列，$u \in \mathcal{D}'(\Omega)$。如果

  $$\lim_{p\to\infty} u_p \overset{\mathcal{D}'}{=} u.$$

  那么，对任意的多重指标 α，在分布的意义下，我们有

  $$\lim_{p\to\infty} \partial^\alpha u_p \overset{\mathcal{D}'}{=} \partial^\alpha u.$$

A2) 证明，$\mathcal{D}'(\Omega)$ 和 $\mathcal{E}'(\mathbb{R}^n)$ 是线性空间。如果你认为这是显然的，可以跳过它的证明。

A3) 证明分布的支撑集所满足的性质，其中 u，u_1，u_2 都是 Ω 上的分布。

- 两个分布的和的支集：

  $$\operatorname{supp}(u_1 + u_2) \subset \operatorname{supp}(u_1) \cup \operatorname{supp}(u_2).$$

– 对任意的多重指标 α，我们有

$$\operatorname{supp}(\partial^\alpha u) \subset \operatorname{supp}(u).$$

– 对任意的 $f \in C^\infty(\mathbb{R}^n)$，我们有

$$\operatorname{supp}(f \cdot u) \subset \operatorname{supp}(u) \cap \operatorname{supp}(f).$$

A4) 证明如下几个精神上类似的命题：分布是传统意义上函数的推广。

– 假设 $\Phi : \Omega_1 \to \Omega_2$ 是 $\mathbb{R}^n$ 上两个开集之间的微分同胚。证明，对任意的 $u \in L^1_{\mathrm{loc}}(\Omega_2)$，如果我们用 $\Phi^* u$ 表示 u 在分布的意义下被 Φ 所拉回的像，那么，在分布的意义下，我们有

$$\Phi^* u \overset{\mathcal{D}'}{=} u(\Phi(x)),$$

其中，我们把复合函数 $u(\Phi(x))$（仍然局部可积）看作是分布。

– 假设 $u \in C^1(\Omega)$。证明，u 作为分布的导数和它通常意义下导数（作为分布）是一致的。

– 假设 $u \in L^1_{\mathrm{loc}}(\Omega)$，$f \in \mathcal{D}'(\mathbb{R}^n)$。证明，在分布意义下的乘积 $f \cdot u$（先把 u 视作是分布）等于现在传统意义下得到 $f \cdot u$ 然后再把这个乘积函数看作是分布。

A5) 对于任意的 $f \in C^\infty(\Omega)$ 和 $u \in \mathcal{D}'(\Omega)$，对任意的 $k \leqslant n$，证明，

$$\partial_k(f \cdot u) \overset{\mathcal{D}'}{=} \partial_k f \cdot u + f \cdot \partial_k u.$$

A6) （（重要的计算），参考第五次课第一个定理的证明）对任意的 $\varphi \in \mathcal{D}(\mathbb{R}^n)$，对任意的 $x_0, a \in \mathbb{R}^n$，$\varepsilon \in (0,1)$，$|a| = 1$，根据 Taylor 公式，我们有

$$\varphi(x_0 - y + \varepsilon a) - \varphi(x_0 - y) = \varepsilon \sum_{j=1}^n a_j \partial_j \varphi(x_0 - y) + \varepsilon^2 r(y, \varepsilon, a),$$

其中，

$$r(y, \varepsilon, a) = 2 \sum_{|\alpha|=2} \frac{a^\alpha}{\alpha!} \int_0^1 (1-t) \partial^\alpha \varphi(x_0 - y + t\varepsilon a) dt.$$

给定分布 $u \in \mathcal{D}'(\mathbb{R}^n)$，通过对上面的式子作用，我们得到

$$u * \varphi(x_0 + \varepsilon a) - u * \varphi(x_0 + \varepsilon a) - \varepsilon \sum_{j=1}^n a_j u * \partial_j \varphi(x_0) = \varepsilon^2 \langle u, r(\cdot, \varepsilon, a) \rangle.$$

证明，存在不依赖于 ε 的常数 C，使得

$$\left| \langle u, r(\cdot, \varepsilon, a) \rangle \right| \leqslant C.$$

A7)（重要）我们选取最心爱的截断函数 χ。证明，当 $\varepsilon \to 0$ 时，对任意的试验函数 $\varphi \in \mathcal{D}(\mathbb{R}^b)$，我们都有

$$\chi_\varepsilon * \varphi \xrightarrow{\mathcal{D}} \varphi.$$

A8) 假设 f 是 $\mathbb{R}^n$ 上的光滑函数，$c \in \mathcal{E}'(\mathbb{R}^n)$，证明，函数

$$x \mapsto \langle c, f(x - \cdot)\rangle$$

是光滑函数，其中，上面的配对是一个有紧支集的分布与一个光滑函数的配对。

A9)（重要）证明，$\mathbb{R}_t \times \mathbb{R}^n_x$ 上定义的函数

$$E(t,x) = \frac{\mathbf{1}_{t>0}}{(4\pi t)^{\frac{n}{2}}} e^{-\frac{|x|^2}{4t}},$$

在 $(t,x) = (0,0)$ 之外是光滑的并且满足

$$(\partial_t - \triangle)(E(t,x)) = 0, \quad \text{对任意的 } (t,x) \neq (0,0).$$

A10) 利用卷积的连续性，重新证明（概念上更清晰简洁）关于分布的链式求导法则：

假设 $\Phi : \Omega_1 \to \Omega_2$ 是 $\mathbb{R}^n$ 上两个开集之间的微分同胚。证明，对任意的 $u \in \mathcal{D}'(\Omega_2)$，

$$\partial_j(\Phi^* u) = \sum_{k=1}^{n} \partial_j \Phi_k \cdot \Phi^* \left(\partial_k u\right),$$

其中 Φ_k 是 Φ 的第 k 个分量。

A11) 证明如下几个关于复解析函数的命题：

- 计算下列的等式：
$$\overline{\partial}(\overline{z}) = 1; \quad \overline{\partial}\left(\frac{1}{z}\right) = 0, \quad z \neq 0.$$
- 假设 $F(z)$ 单位圆盘上的复解析函数，那么，对任意的 $0 < r < 1$，我们都有
$$\int_{|z|=r} F(z)dz = 0.$$
- 证明，对任意的 $z_0 \in \mathbb{C}$ 和 $r > 0$，我们都有
$$\frac{1}{2\pi i}\int_{|z-z_0|=r} \frac{1}{z - z_0} dz = 1.$$

习题 B（分布的计算）.

B1) 证明，当 $\varepsilon \to 0+$ 时，如下分布的极限（在 $\mathcal{D}'(\mathbb{R})$ 中收敛）定义了 $\mathbb{R}$ 上的分布：

$$\lim_{\substack{\varepsilon\to 0\\ \varepsilon>0}} \frac{1}{x+i\varepsilon}.$$

我们把这个分布记作

$$\frac{1}{x+i0} := \lim_{\substack{\varepsilon\to 0\\ \varepsilon>0}} \frac{1}{x+i\varepsilon}.$$

B2) 试找一个 $u_\varepsilon \in \mathcal{D}'(\mathbb{R})$，使得

$$u_\varepsilon' \overset{\mathcal{D}'}{=} \frac{1}{x+i\varepsilon}.$$

（你需要小心定义 log 函数）

B3) 证明，分布 $\dfrac{1}{x+i0}$ 有如下的表达式：

$$\frac{1}{x+i0} \overset{\mathcal{D}'}{=} \mathrm{vp}\frac{1}{x} - i\pi\delta_0.$$

B4) 证明，当 $\varepsilon \to 0+$ 时，如下分布的极限（在 $\mathcal{D}'(\mathbb{R})$ 中收敛）定义了 $\mathbb{R}$ 上的分布：

$$\lim_{\substack{\varepsilon\to 0\\ \varepsilon>0}} \frac{1}{x-i\varepsilon}.$$

我们把这个分布记作

$$\frac{1}{x-i0} := \lim_{\substack{\varepsilon\to 0\\ \varepsilon>0}} \frac{1}{x-i\varepsilon}.$$

B5) 证明，分布 $\dfrac{1}{x-i0}$ 有如下的表达式：

$$\frac{1}{x-i0} \overset{\mathcal{D}'}{=} \mathrm{vp}\frac{1}{x} + i\pi\delta_0.$$

B6) 证明，在分布的意义下，我们有

$$\lim_{\varepsilon\to 0} \frac{\varepsilon}{\pi(x^2+\varepsilon)} \overset{\mathcal{D}'(\mathbb{R})}{=} \delta_0.$$

B7) 当 $n \to \infty$ 时，试计算下面分布序列 $\{u_n\}_{n\geqslant 1}$ 在 $\mathcal{D}'(\mathbb{R})$ 中的极限：

- $u_n(x) = \sin(nx)$;
- $u_n(x) = x^{-1}\sin(nx)$;
- $u_n(x) = n\sin(nx)\mathbf{1}_{x\geqslant 0}$。

习题 C （Laplace 算子、位势方程与分布） (这是分布理论乃至微积分中最重要的习题)

假设 $n \geqslant 2$，我们在 $\mathbb{R}^n$ 上考虑问题。令

$$r = \sqrt{x_1^2 + \cdots + x_n^2}$$

为标准的半径函数。我们用 $\mathbf{S}^{n-1}$ 表示 $\mathbb{R}^n$ 中的单位球面，即

$$\mathbf{S}^{n-1}=\{x\in\mathbb{R}^n\big|r(x)=1\}.$$

我们用 $d\sigma$ 表示 $\mathbf{S}^{n-1}$ 上的曲面测度并用 $\vartheta=\frac{x}{|x|}$ 表示单位球面上的一个点。我们回忆 Laplace 算子的定义：

$$\triangle=\sum_{i=1}^{n}\partial_{x_i}^2.$$

C1) 证明，对任意 $\varphi(x)\in\mathcal{D}(\mathbb{R}^n)$，我们有

$$\int_{\mathbb{R}^n}\varphi(x)dx=\int_0^{+\infty}r^{n-1}\left(\int_{\mathbf{S}^{n-1}}\varphi(r\cdot\vartheta)d\sigma(\vartheta)\right)dr.$$

C2) 当 $n=2$ 时，我们令

$$E=\frac{1}{2\pi}\log(r).$$

对任意的 $\varepsilon>0$，定义

$$f_\varepsilon(x)=\begin{cases}\log|x|, & x\geqslant\varepsilon;\\ \log\varepsilon, & |x|<\varepsilon.\end{cases}$$

试利用 f_ε 证明：

$$\triangle E\overset{\mathcal{D}'}{=}\delta_0$$

C3) 当 $n=2$ 时，令

$$g_\varepsilon(x)=\begin{cases}\log|x|, & x\geqslant\varepsilon;\\ a_\varepsilon|x|^2+b_\varepsilon, & |x|<\varepsilon.\end{cases}$$

证明，存在常数 a_ε 和 b_ε，使得函数 g_ε 在 $\mathbb{R}^2$ 上连续可微。

C4) 当 $n=2$ 时，试利用函数 g_ε 证明：

$$\triangle E\overset{\mathcal{D}'}{=}\delta_0$$

C5) 当 $n\geqslant3$ 时，我们用 $\left|\mathbf{S}^{n-1}\right|$ 表示 $\mathbf{S}^{n-1}$ 的测度。令

$$E=\frac{|x|^{2-n}}{(2-n)\left|\mathbf{S}^{n-1}\right|}.$$

证明，$E\in L^1_{\mathrm{loc}}(\mathbb{R}^n)$ 并且作为分布，我们有

$$\partial_jE=\frac{x_j}{\left|\mathbf{S}^{n-1}\right||x|^n}\in L^1_{\mathrm{loc}}(\mathbb{R}^n).$$

C6)（接 C6)）证明，

$$\triangle E\overset{\mathcal{D}'}{=}\delta_0$$

C7)（作为常识）试计算（或者查出）$\left|\mathbf{S}^{n-1}\right|$ 的数值。

在后面的习题中，我们假设 $n=3$（这些结论对 $n\geqslant 3$ 稍作明显的改动也成立，但是我们更关心具有物理意义的情形）。此时，$\triangle$ 的一个基本解可以写成：

$$E(x)=-\frac{1}{4\pi|x|}\in L^1_{\mathrm{loc}}(\mathbb{R}^3),$$

C8) 对任意的 $u\in\mathcal{D}'(\mathbb{R}^3)$ 和 $f\in\mathcal{E}'(\mathbb{R}^3)$，如果它们满足方程

$$\triangle u=f,$$

证明，u 在 $\mathrm{supp}(f)$ 之外是光滑的，即 $u\in C^\infty\big(\mathbb{R}^3-\mathrm{supp}(f)\big)$。

C9) 将 C8) 中的假设放松为 $f\in\mathcal{D}'(\mathbb{R}^3)$。证明同样的结论：$u\in C^\infty\big(\mathbb{R}^3-\mathrm{supp}(f)\big)$。

C10) 证明 Laplace 算子的椭圆正则性定理：假设 $\Omega\subset\mathbb{R}^3$ 是开集，$u\in\mathcal{D}'(\Omega)$，满足

$$\triangle u=f,$$

其中 $f\in C^\infty(\Omega)$ 是光滑函数，那么，$u\in C^\infty(\Omega)$。（通常要得到 u 的正则性需要知道 u 的各个方向的导数，但是椭圆正则性定理说只要知道一个特殊的二阶导数即可！）

C11) 假定 $f\in\mathcal{E}'(\mathbb{R}^3)$。证明，当 $|x|\to\infty$ 时，存在 $C_1>0$（依赖于分布 f），使得

$$|E*f(x)|\leqslant\frac{C}{|x|}.$$

C12) 假定 $f\in\mathcal{E}'(\mathbb{R}^3)$。证明，$u=E*f$ 是下面的位势方程的解：

$$\triangle u=f.$$

并且存在常数 $C_2>0$，使得当 $|x|\to\infty$ 时，我们有

$$\left|u(x)+\frac{\langle f,1\rangle}{4\pi|x|}\right|\leqslant\frac{C_2}{|x|^2}.$$

根据不同的物理情景，$\langle f,1\rangle$ 表示的是 u 的**总质量**或者**总电量**。

C13) 假定 $u\in C^\infty(\mathbb{R}^3;\mathbb{R})$ 是实值函数。我们定义符号函数 $\mathrm{sign}(u)$ 为

$$\mathrm{sign}(u)(x)=\begin{cases}1, & u(x)>0;\\ 0, & u(x)=0;\\ -1, & u(x)<0.\end{cases}$$

证明，在分布的意义下，我们有

$$\triangle|u|\overset{\mathcal{D}'}{\geqslant}\triangle u\cdot\mathrm{sign}(u),$$

即对任意的（实值）非负的 $\varphi \in \mathcal{D}(\mathbb{R}^3)$，我们都有

$$\langle \triangle |u|, \varphi \rangle \geqslant \langle \triangle u \cdot \operatorname{sign}(u), \varphi \rangle,$$

（提示：先在分布的意义下用 $\sqrt{u^2+\varepsilon^2}$ 来逼近 u）

The topology in $C^\infty(X)$ is the inductive limit of the topology in $C_0^\infty(K)$ when the compact set K increases to X, so it is a LF topology[18]. We have avoided this terminology in order not to encourage the once current misconception that familiarity with LF space is essential for the understanding of distribution theory.

— Lars Hörmander



中秋快乐！国庆快乐!

[18]LF 指的是 limit of Fréchet，这是一列 Fréchet 空间的极限。

64 可卷集，$\mathbb{R}^3$ 上的波动算子的基本解，$\mathbb{R}^3$ 上的波动方程

二零二零年十月五日，星期一，晴

定义 429. 假设 F_1 和 F_2 是 $\mathbb{R}^n$ 中的两个闭集，如果对任意的 $R>0$，存在 $R'>0$（可能依赖于 R），使得对任意的 $x_1\in F_1$，$x_2\in F_2$，我们有

$$|x_1+x_2|\leqslant R \quad\Rightarrow\quad |x_1|<R, |x_2|<R,$$

那么，我们称 F_1 与 F_2 是**可卷的**。

更一般地，假设 $\{F_i\}_{i\in I}$ 是 $\mathbb{R}^n$ 中的一族闭集，如果对任意的 $R>0$，存在 $R'>0$（可能依赖于 R），使得对任意的可数子集 $J\subset I$，对任意的 $x_j\in F_j$，其中 $j\in J$，我们有

$$|\sum_{j\in J} x_j|\leqslant R \quad\Rightarrow\quad |x_j|<R' \text{ 对所有 } j \text{ 成立},$$

那么，我们称闭集族 $\{F_i\}_{i\in I}$ 是**可卷的**。

引理 430. 假设闭集 F_1 与 F_2 是可卷的，那么，F_1+F_2 是闭集。

证明: 令 $F=F_2+F_2$ 是闭集，考虑 F 中的 Cauchy 列 $\{x_i+y_i\}_{i\geqslant 1}\subset F$，其中，$x_i\in F_1, y_i\in F_2$。由于 $\{x_i+y_i\}_{i\geqslant 1}$ 是 Cauchy 列，所以是有界的，即存在 $R>0$，使得 $|x_i+y_i|\leqslant R$，其中 $i\geqslant 1$。根据可卷集的性质，存在 R'，使得对每个 i，$|x_i|\leqslant R'$，$|y_i|\leqslant R'$。从而，存在 $\{x_i\}_{i\geqslant 1}$ 的子列 $\{x_{i_p}\}_{p\geqslant 1}$ 和 $\{y_i\}_{i\geqslant 1}$ 的子列 $\{y_{i_p}\}_{p\geqslant 1}$，它们都是 Cauchy 列，从而，存在 $x\in F_1$，$y\in F_2$（这是两个闭集），使得

$$\lim_{p\to\infty} x_{i_p}=x,\quad \lim_{p\to\infty} y_{i_p}=y.$$

所以，$x+y\in F$ 并且

$$\lim_{i\to\infty} x_i+y_i=x+y.$$

这说明 $\{x_i+y_i\}_{i\geqslant 1}$ 的极限点仍然在 F 中，即 F 是闭集。 □

例子. 如下的集合是可卷的：

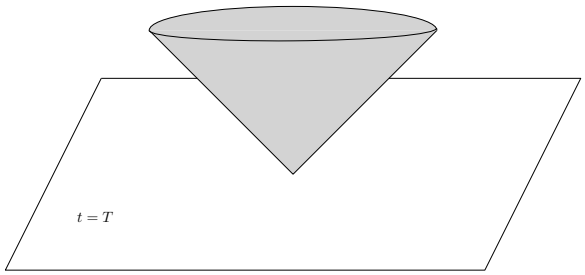
1) $K_1,\cdots,K_m\subset\mathbb{R}^n$ 是闭集，K_0 是闭集，那么，$\{K_i\}_{0\leqslant i\leqslant m}$ 是可卷的。

2) 对 $i=1,2,\cdots,m$，我们令

$$F_i=[x_i,+\infty).$$

那么，$\{F_i\}_{1\leqslant i\leqslant m}$ 在 $\mathbb{R}$ 上是可卷的。

3) 我们在时空 $\mathbb{R}^{1+3}=\mathbb{R}^1\times\mathbb{R}^3$ 上考虑如下两个集合：



- 实心的**未来光锥**:

$$\widehat{C_+} = \big\{(t,x) \in \mathbb{R}^{1+3} \big| t \geqslant |x|\big\}.$$

- 平面 $t = T$ 的未来：$T \in \mathbb{R}$，我们定义

$$\mathbb{R}^{1+3}_{t\geqslant T} = \big\{(t,x) \in \mathbb{R}^{1+3} \big| t \geqslant T\big\}.$$

那么，$\widehat{C_+}$ 和 $\mathbb{R}^{1+3}_{t\geqslant T}$ 是可卷的。

命题 431. 对于 $u, v \in \mathcal{D}'(\mathbb{R}^n)$，如果 $\mathrm{supp}(u)$ 和 $\mathrm{supp}(v)$ 是可卷的，那么，我们可以定义

$$u * v \in \mathcal{D}'(\mathbb{R}^n).$$

实际上，任意选取一组 $\chi_k \in C_0^\infty(\mathbb{R}^n)$，使得对任意的 $|x| \leqslant k$，$\chi_k(x) = 1$，其中，$k \geqslant 1$。那么，对任意的试验函数 $\varphi \in \mathcal{D}(\mathbb{R}^n)$，如下的极限定义了 $u * v$:

$$\langle u * v, \varphi\rangle = \lim_{k\to\infty} \langle (\chi_k u) * (\chi_k v), \varphi\rangle.$$

证明: 我们首先证明，当给定了 $\{\chi_k \in C_0^\infty(\mathbb{R}^n)\}_{k\geqslant 1}$ 之后，其中对任意的 $|x| \leqslant k$，$\chi_k(x) = 1$，极限

$$\lim_{k\to\infty} \langle (\chi_k u) * (\chi_k v), \varphi\rangle.$$

存在。实际上，我们证明存在 N，使得当 $k \geqslant N$ 时，

$$\langle (\chi_k u) * (\chi_k v), \varphi\rangle = \langle (\chi_k u) * (\chi_k v), \varphi\rangle.$$

考虑差

$$\langle (\chi_k u) * (\chi_k v), \varphi\rangle - \langle (\chi_j u) * (\chi_j v), \varphi\rangle = \underbrace{\langle \big(\chi_k - \chi_j)u\big) * (\chi_k v), \varphi\rangle}_{I_1} + \underbrace{\langle (\chi_j u) * \big(\chi_k - \chi_j)v\big), \varphi\rangle}_{I_2}.$$

先考虑 I_1，我们假设它不是 0，此时，

$$\mathrm{supp}(\varphi) \cap \Big\{\big(\mathrm{supp}\big(\chi_k - \chi_j)u\big) + \mathrm{supp}(v)\Big\} \neq \emptyset.$$

特别地，

$$\mathrm{supp}(\varphi) \cap \Big\{\big(\mathrm{supp}(u) + \mathrm{supp}(v)\Big\} \neq \emptyset.$$

假设 $\mathrm{supp}(\varphi)$ 在半径不超过 R 的球之内，那么，根据可卷集的性质，存在 R'，对于任意的 $x+y\in \mathrm{supp}(\varphi)\cap\Big\{(\mathrm{supp}(u)+\mathrm{supp}(v)\Big\}$，其中，$x\in\mathrm{supp}(u)$，$yin\mathrm{supp}(v)$，我们有

$$|x|\leqslant R',\ |y|\leqslant R'.$$

我们取 N 足够大，使得 $\chi_N\big|_{B_{R'}}\equiv 1$。那么，对于 $k,j\geqslant R$，

$$\mathrm{supp}(\varphi)\cap\Big\{(\mathrm{supp}\big(\chi_k-\chi_j)u\big)+\mathrm{supp}(v)\Big\}\neq\emptyset$$

与 $|x|\leqslant R'$ 矛盾。从而，$I_1\equiv 0$。类似的，$I_2\equiv 0$。

上面的论证还表明

$$\langle u*v,\varphi\rangle=\lim_{k\to\infty}\langle(\chi_k u)*(\chi_k v),\varphi\rangle.$$

不依赖于 $\{\chi_k\}_{k\geqslant 1}$ 的选取，实际上，我们只要令

$$\langle u*v,\varphi\rangle=\lim_{R\to\infty}\langle(\chi_R u)*(\chi_R v),\varphi\rangle$$

即可，其中 $\chi_R\in C_0^\infty(\mathbb{R}^n)$ 在 B_R 上恒为 1。

我们还要说明 $u*v$ 是分布。实际上，假设 $\mathrm{supp}(u)\subset B_R$，我们有

$$\begin{aligned}\langle u*v,\varphi\rangle&=\lim_{k\to\infty}\langle u,\chi_k\cdot((\chi_k v)^\vee*\varphi)\rangle\\&=\langle u,\chi_N\cdot((\chi_N v)^\vee*\varphi)\rangle,\end{aligned}$$

其中，N 如前述。我们把分布定义所要证明的不等式留作作业。 □

命题 432. *对于 $u,v,w\in\mathcal{D}'(\mathbb{R}^n)$，假设 $\mathrm{supp}(u)$，$\mathrm{supp}(v)$ 和 $\mathrm{supp}(w)$ 是可卷的，那么，我们*

1) $\mathrm{supp}(u*v)\subset\mathrm{supp}(u)+\mathrm{supp}(v)$;

2) $u*v=v*u$;

3) $(u*v)*w=u*(v*w)$;

4) 对任意的多重指标 α，我们有

$$\partial^\alpha(u*v)=(\partial^\alpha u)*v=u*(\partial^\alpha v);$$

证明: 我们只证明第四条，其余的留作习题。根据归纳法，我们不妨假设 $|\alpha|=1$。首先观察到，$\mathrm{supp}\big(\partial^\alpha u\big)$ 和 $\mathrm{supp}(v)$ 还是可卷的。我们选取试验函数 φ。根据卷积的构造，我们知道存在比较大

的 k，使得

$$\begin{aligned}\langle\partial^\alpha(u*v),\varphi\rangle &= (-1)^{|\alpha|}\langle u*v,\partial^\alpha\varphi\rangle\\ &= (-1)^{|\alpha|}\langle\chi_k u,(\chi_k v)^\vee*\partial^\alpha\varphi\rangle\\ &= (-1)^{|\alpha|}\left\langle\chi_k u,\partial^\alpha\left((\chi_k v)^\vee*\varphi\right)\right\rangle\\ &= \langle\partial^\alpha(\chi_k u),(\chi_k v)^\vee*\varphi\rangle\\ &= \langle\partial^\alpha(\chi_k)u,(\chi_k v)^\vee*\varphi\rangle+\underbrace{\langle\chi_k\cdot\partial^\alpha(u),(\chi_k v)^\vee*\varphi\rangle}_{=\langle\partial^\alpha(u)*v,\varphi\rangle}\\ &= \left\langle\left(\partial^\alpha(\chi_k)u\right)*(\chi_k v),\varphi\right\rangle+\langle\partial^\alpha(u)*v,\varphi\rangle.\end{aligned}$$

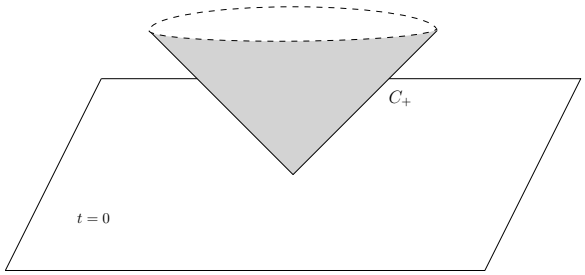
上面第一项由于 $\partial^\alpha(\chi_k)u$ 的支集在 $|x|>k$ 上，重复定理中的证明（利用可卷性），我们知道这一项是 0，所以，

$$\langle\partial^\alpha(u*v),\varphi\rangle=\langle\partial^\alpha(u)*v,\varphi\rangle.$$

另外一个等号用交换律即可。 □

$\mathbb{R}^3$ 中的波动方程

波动算子 $\Box=-\partial_t^2+\triangle$ 定义在 $\mathbb{R}^{1+3}=\mathbb{R}\times\mathbb{R}^3$ 上的算子，其中，第一个坐标是时间 t 的坐标。它作用在以 $(t,x)\in\mathbb{R}\times\mathbb{R}^n$ 为变量函数上的。



我们在 $\mathbb{R}^{1+3}$ 中定义未来光锥 C_+：

$$C_+=\left\{(t,x)\in\mathbb{R}^{1+3}\middle|t=|x|\right\}.$$

除去 $(0,0)$ 点外，C_+ 是 $\mathbb{R}^4$ 中的一个光滑超曲面（请证明这一点!），我们用 $d\sigma$ 表示 C_+ 的曲面测度。那么，对于任意的 $\varphi\in\mathcal{D}(\mathbb{R}^{1+3})$，$d\sigma$ 给出了一个 0 阶的分布：

$$\langle d\sigma,\varphi\rangle=\int_{C^+}\varphi(t,x)\Big|_{C_+}d\sigma.$$

由于 C_+ 可以看作是函数图像（$t=|x|$），所以，

$$\langle d\sigma,\varphi\rangle=\sqrt{2}\int_{\mathbb{R}^3}\varphi(|x|,x)dx.$$

定理 433. 分布 $W=-\dfrac{d\sigma}{4\pi\sqrt{t^2+|x|^2}}$ 是 $\Box$ 的一个基本解，即

$$\Box\left(-\frac{d\sigma}{4\pi\sqrt{t^2+|x|^2}}\right)=\delta_0.$$

证明：按照定义，对任意的试验函数 $\varphi\in\mathcal{D}(\mathbb{R}^{1+3})$，我们有

$$\left\langle\Box\left(\frac{d\sigma}{\sqrt{t^2+|x|^2}}\right),\varphi\right\rangle=\int_{\mathbb{R}^3}\frac{\big(\Box\varphi\big)(|x|,x)}{|x|}dx.$$

我们注意到，右边是局部可积的。

根据上学期的作业 11，在球坐标系下，我们有

$$\Box=-\partial_t^2+\triangle=-\partial_t^2+\frac{\partial^2}{\partial r^2}+\frac{2}{r}\frac{\partial}{\partial r}+\frac{1}{r^2}\triangle_{\mathbf{S}^2},$$

其中 $r=|x|$，$\vartheta=\dfrac{x}{|x|}\in\mathbf{S}^2$ 在单位球面上。所以，利用球坐标系，我们有

$$\begin{aligned}\int_{\mathbb{R}^3}\frac{\big(\Box\varphi\big)(|x|,x)}{|x|}dx&=\int_{\mathbb{R}^3}\frac{\big(\Box\varphi\big)(|x|,x)}{|x|}r^2drd\sigma_{\mathbf{S}^2}\\&=\int_0^\infty\int_{\mathbf{S}^2}\left(-\partial_t^2\varphi+\partial_r^2\varphi+2\partial_r\varphi+\frac{1}{r}\triangle_{\mathbf{S}^2}\varphi\right)(r,r,\vartheta)rd\sigma_{\mathbf{S}}dr.\end{aligned}$$

根据上学期的作业 11 的结论（球面上的散度公式），对于固定的 $r>0$，我们有

$$\int_{\mathbf{S}^2}\left(\triangle_{\mathbf{S}^2}\varphi\right)(r,r,\vartheta)d\sigma_{\mathbf{S}}=0.$$

所以，

$$\begin{aligned}\int_{\mathbb{R}^3}\frac{\big(\Box\varphi\big)(|x|,x)}{|x|}dx&=\int_0^\infty\int_{\mathbf{S}^2}\left(-\partial_t^2\varphi+\partial_r^2\varphi+2\partial_r\varphi\right)(r,r,\vartheta)rd\sigma_{\mathbf{S}^2}dr\\&=\int_0^\infty\int_{\mathbf{S}^2}\left(-\partial_t^2\varphi+\partial_r^2\varphi+2\partial_r\varphi\right)(r,r,\vartheta)rd\sigma_{\mathbf{S}^2}dr\\&=\int_0^\infty\int_{\mathbf{S}^2}-\big((\partial_t+\partial_r)(\partial_t-\partial_r)(r\varphi)\big)(r,r,\vartheta)d\sigma_{\mathbf{S}^2}dr.\end{aligned}$$

如果我们令 $\psi=r\varphi$，那么，

$$\left\langle\Box\left(-\frac{d\sigma}{\sqrt{t^2+|x|^2}}\right),\varphi\right\rangle=\int_0^\infty\int_{\mathbf{S}^2}\big((\partial_t+\partial_r)(\partial_t-\partial_r)(\psi)\big)(r,r,\vartheta)d\sigma_{\mathbf{S}^2}(\vartheta)dr.$$

我们令 $L=\partial_t+\partial_r$，$\underline{L}=\partial_t-\partial r$。对任意的光滑函数，$\vartheta$ 固定，我们计算

$$\frac{d}{dr}\left(f(r,r,\vartheta)\right)=(\partial_tf)(r,r,\vartheta)+(\partial_rf)(r,r,\vartheta).$$

所以，

$$\begin{aligned}\left\langle \Box\left(-\frac{d\sigma}{\sqrt{t^2+|x|^2}}\right),\varphi\right\rangle &= \int_0^\infty \int_{\mathbf{S}^2} \frac{d}{dr}\Big[\big(\underline{L}(\psi)\big)(r,r,\vartheta)\Big] d\sigma_{\mathbf{S}^2}(\vartheta)dr \\ &= \int_0^\infty \frac{d}{dr}\Big[\int_{\mathbf{S}^2}\big(\underline{L}(\psi)\big)(r,r,\vartheta)d\sigma_{\mathbf{S}^2}(\vartheta)\Big]dr \\ &= \lim_{\varepsilon\to 0}\int_{\mathbf{S}^2}(\underline{L}\psi)(\varepsilon,\varepsilon,\vartheta)d\sigma_{\mathbf{S}^2}(\vartheta)\end{aligned}$$

另外，

$$(\underline{L}\psi)(t,x) = \varphi(t,x) + |x|\underline{L}(\varphi)(t,x).$$

所以，

$$\lim_{\varepsilon\to 0}\int_{\mathbf{S}^2}(\underline{L}\psi)(\varepsilon,\varepsilon,\vartheta)d\sigma_{\mathbf{S}^2}(\vartheta) = \int_{\mathbf{S}^2}\varphi(0,0)d\sigma_{\mathbf{S}^2}(\vartheta) = 4\pi\varphi(0,0).$$

这就证明了结论。 □

给定 $f\in\mathcal{D}'(\mathbb{R}^{1+3})$，如果存在 $T\in\mathbb{R}$，使得 $\mathrm{supp}(f)\subset\mathbb{R}^{1+3}_{t\geqslant T}$，我们就说 f 的**过去是零**。

命题 434. *假设 $f\in\mathcal{D}'(\mathbb{R}^{1+3})$ 的过去是零，那么，存在唯一的过去为零的 $u\in\mathcal{D}'(\mathbb{R}^{1+3})$，使得*

$$\Box u = f.$$

特别地，u 可以表示为

$$u = -\frac{d\sigma}{4\pi\sqrt{t^2+|x|^2}} * f.$$

证明： 我们注意到 $\mathbb{R}^{1+3}_{t\geqslant T}$ 与 C_+ 是两个可卷的闭集，所以，

$$u = -\frac{d\sigma}{4\pi\sqrt{t^2+|x|^2}} * f$$

是良好定义的（因为 $\mathrm{supp}(d\sigma)\subset C_+$）。由于

$$\mathbb{R}^{1+3}_{t\geqslant T} + \mathrm{supp}(d\sigma)\subset\mathbb{R}^{1+3}_{t\geqslant T},$$

所以，u 的过去为零。

现在证明唯一性：假设 $v\in\mathcal{D}'(\mathbb{R}^{1+3})$ 的过去为零，并且

$$\Box v = f.$$

那么，我们有

$$\begin{aligned}v = v*\delta_0 &= v*\Box\left(-\frac{d\sigma}{4\pi\sqrt{t^2+|x|^2}}\right) \\ &= \Box(v)*\left(-\frac{d\sigma}{4\pi\sqrt{t^2+|x|^2}}\right) \\ &= f*\left(-\frac{d\sigma}{4\pi\sqrt{t^2+|x|^2}}\right) = u.\end{aligned}$$

这就证明了唯一性，其中，为了使得上面每一个式子都有定义，我们用到了 v 的过去为零这个条件。 □

推论 435. 分布 $-\dfrac{d\sigma}{4\pi\sqrt{t^2+|x|^2}}$ 是 □ 的唯一一个过去为零的基本解。

证明: 这是因为这样的基本解都满足

$$\Box u = \delta_0,$$

其中，δ_0 的过去为零。 □

我们现在研究所谓的 Cauchy（初值）问题，这和常微分方程类似，我们要在 $t=0$ 这个时刻给定 u 的初始值（此时，由于方程对时间是 2 阶的，我们需要给定 $u(0,x)$ 和 $(\partial_t u)(0,x)$），然后解方程。我们假设 u 是 $\mathbb{R}^{1+3}$ 上的光滑函数，它满足如下的波动方程：

$$\begin{cases} \Box u(t,x) = 0, \quad (t,x) \in \mathbb{R}^{1+3}; \\ u\big|_{t=0} = u_0(x), \quad x \in \mathbb{R}^3; \\ \partial_t u\big|_{t=0} = u_1(x), \quad x \in \mathbb{R}^3. \end{cases}$$

我们要在 $t \geqslant 0$ 来研究这个问题：给定了 $t=0$ 的初始值，我们想知道 $u(t,x)$ 在未来 $t \geqslant 0$ 处的演化。考虑过去为零的分布

$$H(t)u(t,x) \in \mathcal{D}'(\mathbb{R}^{1+3}).$$

那么，利用 Heaviside 函数 $H(t)$ 的导数的计算，我们有

$$\begin{aligned} \Box\left(H(t)u(t,x)\right) &\overset{\mathcal{D}'}{=} -\partial_t\big(u(t,x)\delta_0(t) + H(t)\partial_t u\big) + H(t)\triangle u \\ &= -\partial_t\big(u_0(x)\delta_0(t) + H(t)\partial_t u\big) + H(t)\triangle u \\ &= -\partial_t\big(u_0(x)\delta_0(t)\big) - u_1(x)\delta_0(t) + H(t)\Box u. \end{aligned}$$

按照定义，$u_1(x)\delta_0(t)$ 是如下的分布：

$$\langle u_1(x)\delta_0(t), \varphi(t,x)\rangle = \int_{\mathbb{R}^3} \varphi(0,x)u_1(x)dx.$$

这实际上是 $t=0$ 所定义的曲面测度乘以 $u_1(x)$ 所定义的分布。

从而，

$$\Box\left(H(t)u(t,x)\right) = -\partial_t\big(u_0(x)\delta_0(t)\big) - u_1(x)\delta_0(t).$$

上式的右边是一个过去为零的分布，所以，对于 $t \geqslant 0$，我们有

$$u(t,x) = -\partial_t\left(W * \big[u_0(x)\delta_0(t)\big]\right) - W * \big[u_1(x)\delta_0(t)\big].$$

注记. 以上的公式是在假设波动方程有光滑解的情况下所给出的解的表达式！不难看出，我们只要假设 $u(t,x)$ 是 C^2 的，上面的计算就成立。

下面我们把上面解的表达式显式地用微积分写清楚。我们用 v 表示一个支集在 $t=0$ 上的分布，在应用的时候，我们将会选取

$$v=-u_0(x)\delta_0(t)\ \text{ 或者 }\ -u_1(x)\delta_0(t).$$

为了一下计算明了，我们这里不妨假设 u_0 和 u_1 都有紧支集（否则，我们将对下面的 v_ε 加一个 x 方向的截断函数，请参考本次作业）。我们再令

$$v_\varepsilon=\chi_\varepsilon(t)u_i(x),\ \ i=0,1.$$

其中 χ_ε 是对 δ 的一个逼近。很容易验证，当 $\varepsilon\to 0$，我们有

$$v_\varepsilon\xrightarrow{\mathcal{D}'}v.$$

从而，

$$W*v_\varepsilon\xrightarrow{\mathcal{D}'}W*v.$$

现在来计算 $W*v_\varepsilon$。按照定义，我们有

$$\begin{aligned}
W*v_\varepsilon(x)&=-\left\langle\frac{d\sigma}{4\pi\sqrt{|t'|^2+|x'|^2}},\chi_\varepsilon(t-t')u_i(x-x')\right\rangle\\
&=-\int_{\mathbb{R}^3}\frac{\chi_\varepsilon(t-x')u_i(x-x')}{4\pi|x|}dx\\
&=-\int_0^\infty\int_{\mathbf{S}^2}\frac{\chi_\varepsilon(t-|r'|)u_i(x-x')}{4\pi}r'd\sigma_{\mathbf{S}^2}(\vartheta)dr'\\
&=-\int_0^\infty\chi_\varepsilon(t-r')\left(r'\int_{\mathbf{S}^2}\frac{u_i(x-x')}{4\pi}d\sigma_{\mathbf{S}^2}(\vartheta')\right)dr'\\
&\longrightarrow -t\int_{\mathbf{S}^2}\frac{u_i(x-t\vartheta')}{4\pi}d\sigma_{\mathbf{S}^2}(\vartheta').
\end{aligned}$$

其中，我们假设了 $t\geqslant 0$ 并且 $x'=r'\vartheta'$。在最后一步中，我们用到了

$$\chi_\varepsilon(t-r')\xrightarrow{\mathcal{D}'}\delta_t(r').$$

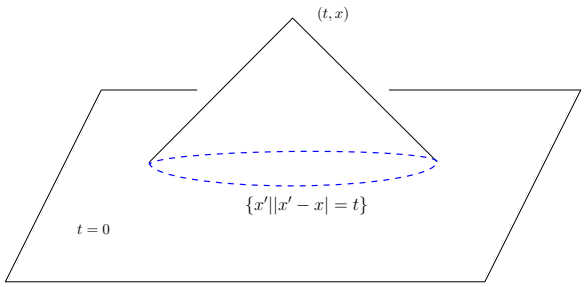
最终，我们得到

$$W*v(x)=-tH(t)\int_{\mathbf{S}^2}u_i(x-t\vartheta')\frac{d\sigma_{\mathbf{S}^2}(\vartheta')}{4\pi}.$$

特别地，我们得到了波动方程的解的表达式：对于 $t\geqslant 0$，我们有

$$u(t,x)=-t\int_{\mathbf{S}^2}u_1(x-t\vartheta')\frac{d\sigma_{\mathbf{S}^2}(\vartheta')}{4\pi}-\partial_t\left[t\int_{\mathbf{S}^2}u_0(x-t\vartheta')\frac{d\sigma_{\mathbf{S}^2}(\vartheta')}{4\pi}\right].$$

在 $u(t,x)$ 的表达式中，由于 $t\vartheta'$ 是长度为 1 的向量，所以，积分项只与 u_0 和 u_1 在以 x 为中心以 t 为半径的圆上的值有关系。



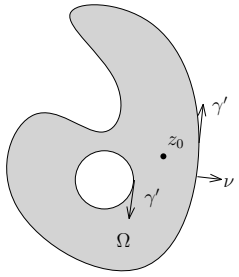
我们还有一个更几何一点的表述，u 在 (t,x) 处的值只和初始值在顶点在 (t,x) 处的倒向的光锥和 $t=0$ 相截得到的球面 $\{x'\big||x'-x|=t\}$ 上的值有关。

65 选读部分：复分析相关知识简介，极大值原理，Laurent 展开，留数定理（在 Fourier 变换的计算中要用）

二零二零年十月七日，星期三，晴

复分析中的留数定理

假设 $\Omega \subset \mathbb{C}$ 是开集，$K \subset \Omega$ 是有界带边区域（特别地，K 是紧的），其边界 $\gamma = \partial K$ 是 C^1 曲线（可以有多个连通分支）。



我们已经对复解析函数 $F(z)$ 证明了 Cauchy 积分公式：如果 $F(z)$ 在 Ω 上是复解析的，那么，对于 $z_0 \in \mathring{K}$（K 的内部），我们有

$$F(z_0) = \frac{1}{2\pi i}\int_\gamma \frac{F(z)}{z - z_0}dz.$$

利用 Cauchy 公式，我们证明，复解析函数是“解析”的，也就是说可以在如下意义下写成幂级数的形式：

定理 436. *假设 F 为开集 $\Omega \subset \mathbb{C}$ 上的复解析函数并且以 z_0 为圆心以 R 为半径的开球 $B_R(z_0) \subset \Omega$。那么，$F(z)$ 在 z_0 处的*解析半径*至少是 R，也就是说，在 $B_R(z_0)$ 上，我们有*

$$F(z) = a_0 + a_1(z - z_0) + a_2(z - z_0)^2 + \cdots + a_n(z - z_0)^n + \cdots,$$

其中

$$a_k = \frac{1}{2\pi i}\int_{|\xi - z_0| = r} \frac{F(\xi)}{(\xi - z_0)^k}dz, \quad r < R.$$

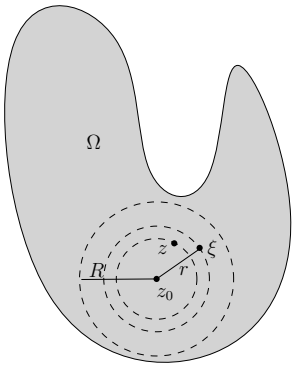
这里，等式右边的幂级数对任意的 $z \in B_R(z_0)$ 都是收敛的。上面系数定义中的 r 可以是 $(0, R)$ 中的任意一个数。。

证明的想法很简单：我们把 Cauchy 积分公式

$$F(z) = \frac{1}{2\pi i}\int_\gamma \frac{F(\xi)}{\xi - z}dz.$$

的右边的积分项强行展开即可。

证明： 我们选取一个正实数 r，使得 $|z - z_0| < r < R$。我们任意给定 ξ，使得 $|\xi - z_0| = r$。



此时，我们有

$$\begin{aligned}\frac{1}{\xi - z} &= \frac{1}{(\xi - z_0) - (z - z_0)} = \frac{1}{\xi - z_0}\frac{1}{1 - \frac{z - z_0}{\xi - z_0}} \\ &= \sum_{k=0}^{\infty} \frac{(z - z_0)^k}{(\xi - z_0)^{k+1}}.\end{aligned}$$

在上面的式子中，由于 $\left|\dfrac{z - z_0}{\xi - z_0}\right| < 1$，所以，我们有（级数收敛）：

$$\frac{1}{1 - \frac{z - z_0}{\xi - z_0}} = \sum_{k=0}^{\infty} \left(\frac{z - z_0}{\xi - z_0}\right)^k.$$

这些级数显然是绝对收敛的，从而与积分可交换（也可以利用 Lebesgue 控制收敛定理）。所以，代入 Cauchy 积分公式，我们就得到

$$\begin{aligned}F(z) &= \frac{1}{2\pi i}\int_{|\xi - z_0| = r} \sum_{k=0}^{\infty} F(\xi)\frac{(z - z_0)^k}{(\xi - z_0)^{k+1}} d\xi \\ &= \sum_{k=0}^{\infty} \left(\frac{1}{2\pi i}\int_{|\xi - z_0| = r} \frac{F(\xi)}{(\xi - z_0)^{k+1}} d\xi\right)(z - z_0)^k\end{aligned}$$

比较系数，这就给出了定理的证明。 □

推论 437 (零点的离散性)**.** 这里的假设与定理中是一致的。我们进一步假设 Ω 是道路连通的，即对任意的 $z_1, z_2 \in \Omega$，存在连续映射

$$\gamma : [0, 1] \to \Omega, \quad t \mapsto \gamma(t),$$

使得 $\gamma(0) = z_1, \gamma(1) = z_2$。那么，复解析函数 F 在 Ω 中的零点是离散的（即如果 z_0 是 F 的一个零点，那么，存在 $\varepsilon > 0$，使得对任意的 z，$|z - z_0| < \varepsilon$，$F(z) \neq 0$），除非 $F \equiv 0$。

特别地，给定 Ω 上的两个复解析函数 F 和 G，如果 F 和 G 在 $Z \subset \Omega$ 上取值相同，并且 Z 在 Ω 中有聚点，那么，$F \equiv G$。

证明: 假设 z_0 是 F 的一个零点，即 $F(z_0)=0$。根据 F 的解析表达式，在 $B_R(z_0)\subset\Omega$ 上，我们有

$$F(z)=a_0+a_1(z-z_0)+a_2(z-z_0)^2+\cdots+a_n(z-z_0)^n+\cdots.$$

如果这些系数 a_i 全部为 0，那么，F 在 $B_R(z_0)$ 上恒为 0；否则，假设 a_m 是第一个不是 0 的系数，那么，我们有

$$F(z)=(z-z_0)^m\big(a_m+a_{m+1}(z-z_0)+a_{m+2}(z-z_0)^2+\cdots\big)=(z-z_0)^mG(z).$$

根据定理的证明，级数

$$G(z)=a_m+a_{m+1}(z-z_0)+a_{m+2}(z-z_0)^2+\cdots$$

也是绝对收敛的，特别地，这是连续的，所以，当 $z=z_0$ 时，$G(z_0)=a_m\neq 0$。从而，存在 δ，使得 G 在 $B_\delta(z_0)$ 上不会等于 0，此时，我们知道，F 在 z_0 的附近（$B_\delta(z_0)$ 上）恰好有一个零点。

我们现在证明，如果 F 在 z_0 的一个领域 $B_\delta(z_0)$ 上恒为 0，那么，F 在 Ω 上恒为 0：任意选取 $z_1\in\Omega$ 和曲线

$$\gamma:[0,1]\to\Omega,\quad t\mapsto\gamma(t),$$

使得 $\gamma(0)=z_0,\gamma(1)=z_1$。令

$$t_*=\sup\big\{t\in[0,1]]\big|\ F(\gamma(t))\big|_{[0,t]}\equiv 0\big\}.$$

由于 F 在 z_0 的一个领域 $B_\delta(z_0)$ 上恒为 0，所以，$t_*>0$。我们现在证明 $t=1$：如果假设 $t_*<1$，根据连续性，我们知道 $F(\gamma(t_*))=0$。根据前面的构造，由于 F 在 $\gamma(t_*)$ 的任意一个小邻域中都有零点（因为和 $\gamma^{-1}([0,t_*])$ 相交），根据之前的推导，存在 $\gamma(t_*)$ 一个领域 $B_{\delta_1}\big(\gamma(t_*)\big)$，使得 F 在 $B_{\delta_1}\big(\gamma(t_*)\big)$ 上恒为 0，根据连续性，那么，存在 $\varepsilon>0$，使得 $F(\gamma(t))\big|_{[0,t_*+\varepsilon]}\equiv 0$，这和 t_* 的最大性矛盾。

由于 $t_*=1$，所以，$F(z_1)=0$，这就证明了在 Ω 上，$F\equiv 0$。

定理中的第二个结论考虑 $F-G$ 即可。 □

推论 438. 这里的假设与定理中是一致的。那么，F 的 n 次导数 $F^{(n)}$ 仍然是解析函数。我们进一步有如下的公式:

$$F^{(n)}(z)=\frac{n!}{2\pi i}\int_{|\xi-z|=r}\frac{F(\xi)}{(\xi-z)^{n+1}}d\xi.$$

特别地，我们有如下的导数估计:

$$|F^{(n)}(z)|\leqslant\frac{n!}{r^n}\sup_{|\xi-z|=r}|f(\xi)|.$$

证明: 我们将 F 在 $z_0=0$（不妨）处展开为级数

$$F(z)=a_0+a_1z+a_2z^2+\cdots$$

其中，我们假设上面的级数的收敛半径至少是 $R>0$，即对于 $|z|<R$ 都是绝对收敛的。根据定理中系数的计算，我们知道

$$|a_k| \leqslant \frac{1}{r^k} \sup_{|\xi|=r} |F(\xi)| \leqslant \frac{1}{r^k} M, \quad r < R.$$

其中，M 为 $|F|$ 在半径为 r 的圆圈上的最大值。从而，对 $|z| \leqslant r' < r$，对任意的 $k \geqslant 0$，我们有

$$|k a_k z^{k-1}| \leqslant M \underbrace{\frac{k}{r}\left|\frac{r'}{r}\right|^k}_{b_k}.$$

当 k 足够大的时候，比如说

$$k > k_0 = \lfloor \left(1 - \frac{r'}{r}\right)^{-1} \rfloor$$

时，我们有

$$\frac{b_{k+1}}{b_k} = \left(1 + \frac{1}{k}\right)\frac{r'}{r} < 1,$$

所以，$\{b_k\}_{k \geqslant k_0}$ 可以被一个公比小于 1 几何级数来控制，从而，级数

$$a_1 + 2a_2 z^1 + 3a_3 z^2 + \cdots$$

在 $|z| < r'$ 时是绝对收敛，这表明可以逐项求微分（根据 Lebesgue 控制收敛定理的推论）。这说明

$$F'(z) = a_1 + 2a_2 z^1 + 3a_3 z^2 + \cdots.$$

对于 $|z| < r' < r < R$ 成立。由于 r' 和 r 是任意选取的，所以，上面的的式子对于 $|z| < R$ 都成立。特别地，我们证明了

$$F'(z_0) = a_1.$$

由归纳法，对任意的 k，我们就有

$$F^{(k)}(z_0) = k! a_k.$$

再根据定理中的计算，这就证明了这个推论叙述中的公式。导数估计是显然的。 □

定理 439 (Liouville)**.** *假设 $F(z)$ 是在整个 $\mathbb{C}$ 上定义的复解析函数[19]。如果 F 是有界函数，那么 F 是常值函数。*

证明: 我们只要证明 $F'(z) \equiv 0$ 即可：根据 F 在一点处的展开，$F'(z) \equiv 0$，意味着定理中的系数

$$a_1 = 2a_2 = 3a_3 = \cdots = 0.$$

所以，F 在一点附近恒为 a_0，从而 F 为常数（一点附近的邻域有聚点）。

我们利用导数估计：

$$|F'(z)| \leqslant \frac{1}{r} \sup_{|\xi - z| = r} |F(\xi)| \leqslant \frac{\|F\|_\infty}{r}.$$

由于 F 在整个 $\mathbb{C}$ 上定义，从而可以将 r 取得任意大，这表明对任意的 $z \in \mathbb{C}$，$F'(z) = 0$。 □

[19]这样的函数被称作是**整函数**。

推论 440 (代数基本定理)**.** *对任意的次数非零的复系数多项式*

$$P(z) = z^n + a_{n-1}z^{n-1} + \cdots + a_1 z + a_0,$$

它在 $\mathbb{C}$ 上必有一个根。

证明: 我们观察到 P 在整个 $\mathbb{C}$ 上是复解析的，并且当 $|z| \to 0$ 时，我们有

$$|P(z)| \to \infty.$$

我们用反证法：如若不然，

$$F(z) = P(z)^{-1}$$

是在全平面 $\mathbb{C}$ 上良好定义的函数。另外，我们有

$$\overline{\partial}(F) = -P(z)^{-2}\overline{\partial}P = 0.$$

所以，F 是复解析的。另外，当 $|z| \to \infty$ 时，我们还有

$$|F(z)| \to 0.$$

这说明，F 是有界的。根据 Liouville 定理，F 为常数，从而 P 也是，那么它的次数是 0，矛盾。 □

Cauchy 积分公式是复分析中最重要的公式，除了用来证明解析性，它还有其他众多重要的推论，比如说关于复解析函数的极大模原理:

定理 441 (极大模原理)**.** *假设 $F(z)$ 是区域 $\Omega \subset \mathbb{C}$ 上的复解析函数，那么 $|F(z)|$ 的最大值，如果能取到的话，一定在 Ω 的边界 $\partial\Omega$ 上取到。进一步，如果 $|F|$ 在 Ω 的内部有最大值点，那么 F 一定是常值函数。*

证明: 假设 $z_0 \in \mathring{\Omega}$ 是 $|F|$ 的最大值点，即

$$|F(z_0)| = \sup_{z\in\Omega} |f(z)|.$$

根据 Cauchy 积分公式，（选取较小的 ε，使得 $B_\varepsilon(z_0) \subset \Omega$，下面的的 r 只要满足 $r < \varepsilon$ 即可）我们有

$$\begin{aligned}
|F(z_0)| &= \Big|\frac{1}{2\pi}\int_0^{2\pi} f(z_0 + re^{i\theta})d\theta\Big| \\
&\leqslant \frac{1}{2\pi}\int_0^{2\pi} \|F\|_{L^\infty} d\theta \\
&= \frac{1}{2\pi}\int_0^{2\pi} |F(z_0)| d\theta \\
&= |F(z_0)|.
\end{aligned}$$

这表明，上述不等式必须处处取等号，这表明在 z_0 的整个邻域 $B_\varepsilon(z_0)$ 上，$|F(z)|$ 为常数。

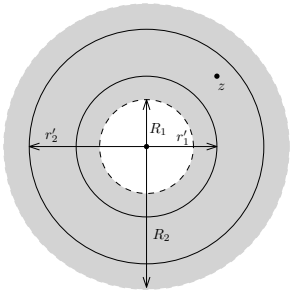
我们现在证明 F 在 $B_\varepsilon(z_0)$ 上为常数，不妨假设 $z_0 = 0$。如若不然，那么，我们当 ε 较小时，我们有解析表达式

$$F(z) = a_0 + z^m(a_m + a_{m+1}z + a_{m+2}z^2 + \cdots),$$

其中 $a_m \neq 0$。通过对 $F(z)$ 乘以一个常数，我们还可以假设 $a_0 > 0$（如果 $a_0 = 0$，那么 $|F|$ 在 z_0 附近恒为零）。选取 ξ，使得 $\xi^m = -\overline{a_m}$，所以，对于比较小的 δ，我们有

$$F(\delta) = a_0 - \delta^m|a_m|^2 + O(\delta^{m+1}).$$

令 $\delta \to 0$，这就和 $|F(z_0)|$ 最大矛盾。所以，F 在 $B_\varepsilon(z_0)$ 上为常数，所以 F 为常数。 □



定理 442 (Laurent 展开)**.** *F 是环面 $\{z \in \mathbb{C} | R_1 < |z| < R_2\}$（我们经常取把 $R_1 = 0$）上的复解析函数。对每个整数 $k \in \mathbb{Z}$，对任意的 $R_1 < r < R_2$，我们定义*

$$a_k = \frac{1}{2\pi i}\int_{|z|=r} \frac{F(z)}{z^k}dz = \frac{1}{2\pi r^k}\int_0^{2\pi} e^{-ik\theta}F(re^{ik\theta})d\theta.$$

那么，对任意满足 $R_1 < |z| < R_2$ 的 z，我们有（对每个固定的 z，以下的级数绝对收敛）

$$F(z) = \sum_{k=-\infty}^{\infty} a_k z^k.$$

证明: 根据 Cauchy 积分公式，a_k 与 r 在 (R_1, R_2) 中的选择无关。我们选取

$$R_1 < r_1' < r_1 < |z| < r_2 < r_2' < R_2$$

并令 C_i' 为半径为 r_i' 并且中心在原点的圆，其中 $i = 1, 2$。根据 Cauchy 积分公式，我们有

$$F(z) = \frac{1}{2\pi i}\int_{C_2'} \frac{F(\xi)}{\xi - z}d\xi - \frac{1}{2\pi i}\int_{C_1'} \frac{F(\xi)}{\xi - z}d\xi.$$

我们重复之前证明复解析函数能做解析展开的做法。

对于第一项，由于对任意的 $\xi \in C_2'$，我们有 $|z| < |\xi|$，我们有

$$\frac{1}{\xi - z} = \frac{1}{\xi}\frac{1}{1 - \frac{z}{\xi}} = \frac{1}{\xi}\sum_{k=0}^{\infty}\left(\frac{z}{\xi}\right)^k.$$

第二项之中，由于 $|z|>|\xi|$，其中 $\xi\in C_1'$，我们有

$$\frac{1}{\xi-z}=\frac{1}{z}\frac{1}{1-\frac{\xi}{z}}=\frac{1}{z}\sum_{k=0}^{\infty}\left(\frac{\xi}{z}\right)^k.$$

将上面的两个展开代入 $F(z)$ 的表达式，我们就有

$$F(z)=\frac{1}{2\pi i}\int_{C_2'}\frac{F(\xi)}{\xi}\sum_{k=0}^{\infty}\left(\frac{z}{\xi}\right)^k d\xi-\frac{1}{2\pi i}\int_{C_1'}\frac{F(\xi)}{z}\sum_{k=0}^{\infty}\left(\frac{\xi}{z}\right)^k d\xi.$$

将求和与积分交换就得到了要证明的结论。 □

定义 443. 假定 $F(z)$ 在区域 $\{z|0<|z-z_0|<r\}$ 上是复解析的，其中 $r>0$，它的 *Laurent* 展开为

$$F(z)=\sum_{k=-\infty}^{\infty}a_k z^k.$$

我们称其中的 z^{-1} 的系数 a_{-1} 为 F 在 z_0 处的**留数**，并记作 $\mathrm{Res}(F;z_0)$。

注记. 按照定义，我们有

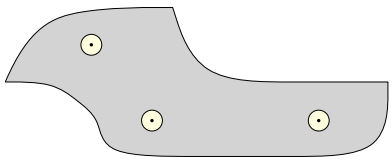
$$\mathrm{Res}(F;z_0)=\frac{1}{2\pi i}\int_{|z-z_0|=r}F(z)dz.$$

这因为 *Laurent* 展开中其它幂次的积分都是 0。

定理 444 (留数定理)**.** 假定 $\Omega\subset\mathbb{C}$ 是一个紧区域，Ω 边界为分段光滑的 C^1-曲线。除去点 $z_1,\cdots,z_N\in\mathring{\Omega}$ 之外，F 为 $\mathbb{C}-\{z_1,\cdots,z_N\}$ 上的复解析函数，那么，

$$\sum_{k=1}^{N}\mathrm{Res}(F;z_k)=\frac{1}{2\pi i}\int_{\partial\Omega}F(z)dz.$$

证明: 对每个 $k\leqslant N$，先将每个 z_k 附近的小圆盘 $B_\varepsilon(z_k)$ 从 Ω 上抠掉，使得 $z\in\Omega-\bigcup_{k\leqslant N}B_\varepsilon(z_k)$。我们在区域 $\Omega-\bigcup_{k\leqslant N}B_\varepsilon(z_k)$ 上用 Cauchy 积分公式。



考虑到曲线的定向，我们有

$$\frac{1}{2\pi i}\int_{\partial\Omega}F(z)dz-\sum_{k=1}^{N}\frac{1}{2\pi i}\int_{\partial B_\varepsilon(z_k)}F(z)dz=0.$$

所以，

$$\frac{1}{2\pi i}\int_{\partial\Omega}F(z)dz=\sum_{k=1}^{N}\frac{1}{2\pi i}\int_{\partial B_\varepsilon(z_k)}F(z)dz=\sum_{k=1}^{N}\mathrm{Res}(F;z_k).$$

这就证明了留数定理。 □

我们试举一个有趣的应用，其它在计算上的应用我们将在 Fourier 变换的一部分再做演示。

例子. 我们在第一学期已经定义了三角函数

$$\cos(z)=\frac{1}{2}\big(e^{iz}+e^{-iz}\big),\quad \sin(z)=\frac{1}{2i}\big(e^{iz}-e^{-iz}\big).$$

我们研究 $\sin(z)$ 的零点，即找到 z_0，使得

$$e^{iz_0}-e^{-iz_0}=0 \quad\Leftrightarrow\quad e^{2iz_0}=1 \quad\Leftrightarrow z_0\in\mathbb{Z}.$$

对任意的 $a\notin\mathbb{Z}$，我们考虑

$$F(z)=\frac{\pi\cot(\pi z)}{(z-a)^2}=\frac{\pi\cos(\pi z)}{\sin(\pi z)(z-a)^2}.$$

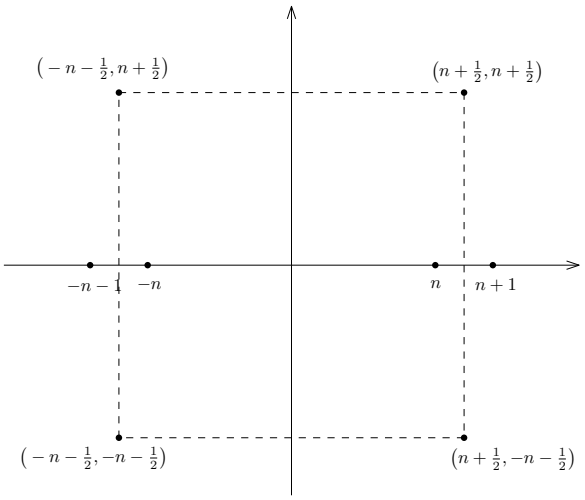
那么，$F(z)$ 不解析的地方只能是 a 和 $n\in\mathbb{Z}$。

在 $z=a$ 处，要想有非平凡的留数，$\cot(\pi z)$ 需要贡献一个 $z-a$ 的因子，所以

$$\mathrm{Res}(F;a)=(\pi\cot(\pi z))'|_{z=a}=-\frac{\pi^2}{\sin^2(\pi a)}.$$

在 $z=n$ 处，由于 $\sin(z)'\big|_{z=n}\neq 0$，所以，$\sin(z)$ 的零点是 1 阶的，据此，我们知道

$$\mathrm{Res}(F;n)=\frac{1}{(n-a)^2}.$$



我们对于顶点在 $(\pm(n+\frac{1}{2}),\pm(n+\frac{1}{2}))$ 的正方形 Q_n 上用留数定理，其中，n 足够大使得 $a\in Q_n$。所以，

$$\sum_{|k|\leqslant n}\frac{1}{(k-a)^2}-\frac{\pi^2}{\sin^2(\pi a)}=\frac{1}{2\pi i}\int_{\partial Q_n}\frac{\pi\cos(\pi z)}{\sin(\pi z)(z-a)^2}dz.$$

利用 $z=x+iy$，我们很容易看出

$$\frac{\pi\cos(\pi z)}{\sin(\pi z)(z-a)^2}=O(\frac{1}{n^2}).$$

所以上面的积分项的贡献为

$$\frac{1}{2\pi i}\int_{\partial Q_n}\frac{\pi\cos(\pi z)}{\sin(\pi z)(z-a)^2}dz = O(\frac{1}{n}).$$

令 $n\to\infty$，我们就得到了

$$\sum_{n\in\mathbb{Z}}\frac{1}{(n-a)^2} = \frac{\pi^2}{\sin^2(\pi a)}.$$

利用上面例子中的分析，我们还可以证明 Euler 的著名公式。首先，在 $\mathbb{C}-\mathbb{Z}$ 上的任意一个紧集上，级数

$$f(z) = \sum_{n\in\mathbb{Z}}\frac{1}{(z-n)^2}$$

是一致收敛的，所以，$f(z)$ 是 $\mathbb{C}-\mathbb{Z}$ 上的复解析函数（可逐项求导数）。Euler 观察到，函数

$$g(z) = \frac{\pi^2}{\sin^2(\pi z)}$$

也是 $\mathbb{C}-\mathbb{Z}$ 上的复解析函数。进一步，由于 $\sin(\pi z)$ 的零点是单零点，所以，$g(z)$ 再每个 $n\in\mathbb{Z}$ 处的 Laurent 展开的负幂和 $f(z)$ 的是一致的。所以，

$$F(z) = g(z) - f(z) = \frac{\pi^2}{\sin^2(\pi z)} - \sum_{n\in\mathbb{Z}}\frac{1}{(z-n)^2}$$

是全平面上定义的复解析函数并且具有周期性 $F(z+1)=F(z)$。

对于任意的 $x\in[0,1]$，我们很容易证明下面的极限：

$$\lim_{|y|\to\infty} F(x+iy) = 0.$$

利用周期性，我们就知道 F 是有界的，从而根据 Liouville 定理，我们得到 $F\equiv 0$。这表明

$$\frac{\pi^2}{\sin^2(\pi z)} = \sum_{n\in\mathbb{Z}}\frac{1}{(z-n)^2}$$

稍加变形，我们有

$$\frac{\pi^2}{\sin^2(\pi z)} - \frac{1}{z^2} = \sum_{n\neq 0}\frac{1}{(z-n)^2}$$

左右在 $z=0$ 处取极限，我们就证明了著名的 Euler 公式：

$$\frac{1}{1^2}+\frac{1}{2^2}+\frac{1}{3^2}+\frac{1}{4^2}+\cdots = \frac{1}{6}\pi^2.$$

65.1 Fourier 分析：$L^1(\mathbb{R}^n)$ 上的 Fourier 变换，Riemann-Lebesgue 引理，频率空间与物理空间的基本观点，Gauss 函数的 Fourier 变换，Fourier 变换的逆变换

二零二零年十月十二日，星期一，晴

$L^1(\mathbb{R}^n)$ 上的 Fourier 变换

我们在之后会经常用到下面的命题，这在上学期 5 月 1 日的课程中已经证明。

命题 445 (线性映射的延拓)**.** 给定赋范线性空间 $(E,\|\cdot\|_E)$, $(E',\|\cdot\|_{E'})$ 和 $(F,\|\cdot\|_F)$，其中 $E'\subset E$ 为线性子空间且其范数为诱导范数，即对任意的 $e'\in E$，我们有

$$\|e'\|_{E'}=\|e'\|_E.$$

假设 E' 在 E 中是稠密的。

$$\begin{array}{ccc} E & \xrightarrow{\iota} & E \\ & \searrow^{L'} & \downarrow{\scriptstyle L} \\ & & F \end{array}$$

如果存在连续线性映射 $L':E'\to F$，那么存在唯一的连续线性映射 $L:E\to F$，使得 $L\big|_{E'}=L'$。

定义 446 (L^1 函数的 Fourier 变换)**.** 对于函数 $f\in L^1(\mathbb{R}^n)$，我们定义它的 *Fourier* 变换 $\widehat{f}$（或者 $\mathcal{F}(f)$）为如下的 $\mathbb{R}^n$ 上的函数：

$$\widehat{f}(\xi)=\mathcal{F}(f)(\xi)=\int_{\mathbb{R}^n}e^{-ix\cdot\xi}f(x)dx.$$

其中 $\xi\in\mathbb{R}^n$。我们通常把 $\widehat{f}$ 的定义域 $\mathbb{R}^n$ 称作是**频率空间**，它的变量通常用 ξ 来表示。

我们考虑 $C(\mathbb{R}^n)$ 的线性子空间

$$C_\circ(\mathbb{R}^n)=\big\{f\in C(\mathbb{R}^n)\big|\lim_{|x|\to\infty}f(x)=0\big\}.$$

我们先说明 $(C_\circ(\mathbb{R}^n),\|\cdot\|_{L^\infty}$ 是一个完备的赋范线性空间：对任意的一致收敛 $\{f_k\}_{k\geqslant1}\subset C(\mathbb{R}^n)$，即

$$\lim_{k\to\infty}\|f_k-f\|_{L^\infty}=0,$$

其中，f 明显是连续函数（因为局部上一致收敛给出的极限是连续的）。对任意的 $\varepsilon>0$，存在 N，使得对任意的 $k\geqslant N$，对任意的 $x\in\mathbb{R}^n$，我们有

$$|f_k(x)-f(x)|<\varepsilon.$$

我们取 $k=N$。由于 $f_N\in C_\circ(\mathbb{R}^n)$，所以，存在 $M>0$，使得当 $|x|>M$ 时，我们有

$$|f_N(x)|<\varepsilon.$$

所以，当 $|x|>M$ 时，我们有

$$|f(x)|<2\varepsilon.$$

这表明 $f\in C_\circ(\mathbb{R}^n)$。

作为总结，我们有

引理 447. $(C_\circ(\mathbb{R}^n),\|\cdot\|_\infty)$ 是完备赋范线性空间。

定理 448. 对任意的 $f\in L^1(\mathbb{R}^n)$，$\widehat{f}\in C_\circ(\mathbb{R}^n)$。特别地，线性映射

$$\mathcal{F}:L^1(\mathbb{R}^n)\longrightarrow C_\circ(\mathbb{R}^n),$$

是连续线性映射，即存在常数 C，使得对任意的 $f\in L^1(\mathbb{R}^n)$，我们有

$$\|\widehat{f}\|_{L^\infty}\leqslant C\|f\|_{L^1}.$$

证明: 我们上学期已经证明了 $C_0^\infty(\mathbb{R}^n)\subset L^1(\mathbb{R}^n)$ 是稠密子空间，根据开始提到的命题，我们只要对 $\varphi\in C_0^\infty(\mathbb{R}^n)$ 进行研究即可。

根据定义，对任意的 $f\in L^1(\mathbb{R}^n)$，我们有

$$\left|\widehat{f}(\xi)\right|\leqslant\int_{\mathbb{R}^n}\left|e^{-ix\cdot\xi}f(x)\right|dx=\int_{\mathbb{R}^n}|f(x)|\,dx.$$

所以，对任意的 $f\in L^1(\mathbb{R}^n)$，我们都有

$$\|\widehat{f}\|_{L^\infty}\leqslant\|f\|_{L^1}.$$

下面只需要证明 $\widehat{\varphi}\in C_0(\mathbb{R}^n)$ 的连续性。为了说明连续性，我们只要用 Lebesgue 控制收敛定理即可：假设 $\mathbb{R}^n$ 中的点列 $\{\xi_k\}_{k\geqslant1}$ 满足 $\xi_k\to x$，我们要证明 $\widehat{f}(\xi_k)\to\widehat{f}(\xi)$，这等价于证明

$$\int_{\mathbb{R}^n}e^{-ix\cdot\xi_k}f(x)dx\to\int_{\mathbb{R}^n}e^{-ix\cdot\xi}f(x)dx.$$

我们只要用 $|f(x)|$ 作为控制函数即可。

最终，我们只需要证明当 $\xi\to\infty$ 时，有 $\varphi(\xi)\to0$。我们只需要分部积分即可（这里我们用到了 φ 的光滑性），这里证明的想法与第一学期我们学过的 Riemann-Lebesgue 引理。我们首先证明如下两个重要的性质／观点：

- 物理空间的求导等价于频率空间的乘法：对任意的 $\varphi\in C_0^\infty(\mathbb{R}^n)$，对任意的 $k\leqslant n$，我们有

$$\widehat{\partial_k\varphi}(\xi)=i\xi_k\widehat{\varphi}(\xi).$$

实际上，我们可以分部积分：

$$\begin{aligned}\widehat{\partial_k\varphi}(\xi)&=\int_{\mathbb{R}^n}e^{-ix\cdot\xi}\partial_k\varphi(x)dx\\&=-\int_{\mathbb{R}^n}(-i\xi_k)e^{-ix\cdot\xi}\varphi(x)dx=i\xi_k\widehat{\varphi}(\xi).\end{aligned}$$

- 物理空间的乘法等价于频率空间的求导：对任意的 $\varphi \in C_0^\infty(\mathbb{R}^n)$，对任意的 $k \leqslant n$，我们有

$$\widehat{-ix_k\varphi}(\xi) = \partial_{\xi_k}\widehat{\varphi}(\xi).$$

  实际上，根据 Lebesgue 控制收敛的推论（积分与求导数可交换，请自行验证细节），我们有

$$\partial_{\xi_k}\widehat{\varphi}(\xi) = \int_{\mathbb{R}^n}(-ix_k)e^{-ix\cdot\xi}\varphi(x)dx = \widehat{-ix_k\varphi}(\xi).$$

- 物理空间的光滑性等价于频率空间的衰减，光滑性越高衰减就越快。

  实际上，对任意的 $\varphi \in C_0^\infty(\mathbb{R}^n)$，对任意的正整数 N，利用第一个原则，我们有

$$(1+|\xi|^2)^N\widehat{\varphi}(\xi) = \mathcal{F}\left((1-\triangle)^N\varphi\right)(\xi),$$

  其中 $\triangle$ 是 Laplace 算子。

  很明显，$(1-\triangle)^N\varphi \in L^1(\mathbb{R}^n)$，所以，我们有

$$|\widehat{\varphi}(\xi)| \leqslant \frac{\|(1-\triangle)^N\varphi\|_{L^1}}{(1+|\xi|^2)^N}.$$

  这表明 $\widehat{\varphi}$ 是衰减的并且我们对于 φ 求的导数越多，那么衰减速度就越快。

  特别地，我们还证明了 $\widehat{\varphi} \in C_\circ(\mathbb{R}^n)$。

□

注记. 通常为了说明 $f \in L^1(\mathbb{R}^n)$ 时，$\widehat{f}$ 也在无穷远处趋向于 0，我们任选 $\varepsilon > 0$，再选取 $\varphi \in C_0^\infty(\mathbb{R}^n)$，使得

$$\|f-\varphi\|_{L^1} < \frac{1}{2}\varepsilon.$$

从而，对任意的 ξ，我们有

$$\begin{aligned}
|\widehat{f}(\xi)| &\leqslant |\widehat{f}(\xi)-\widehat{\varphi}(\xi)| + |\widehat{\varphi}(\xi)| \\
&\leqslant \|\widehat{f}-\widehat{\varphi}\|_\infty + \frac{\|(1-\triangle)^N\varphi\|_{L^1}}{(1+|\xi|^2)^N} \\
&\leqslant \|f-\varphi\|_{L^1} + \frac{\|(1-\triangle)^N\varphi\|_{L^1}}{(1+|\xi|^2)^N} \\
&\leqslant \frac{\varepsilon}{2} + \frac{\|(1-\triangle)^N\varphi\|_{L^1}}{(1+|\xi|^2)^N}.
\end{aligned}$$

所以，当 $|\xi|$ 很大的时候，我们可以使得 $|\widehat{f}(\xi)| < \varepsilon$。

然而，这些过程实际上都被包装在开始的命题上，所以我们通常只对光滑有紧支集的函数（在 L^1 中稠密）来证明即可。

我们现在研究卷积与 Fourier 变换的关系：

命题 449. 对任意的 $f,g\in L^1(\mathbb{R}^n)$，我们有

$$\widehat{f*g}(\xi)=\widehat{f}(\xi)\widehat{g}(\xi).$$

也就是说，卷积对应于频率空间的乘积。

证明: 我们要运用 Fubini 定理。注意到映射

$$\mathbb{R}^n\times\mathbb{R}^n\to\mathbb{C},\quad (x,y)\mapsto e^{-ix\cdot y}f(y)g(x-y)$$

是 $\mathbb{R}^n\times\mathbb{R}^n$ 上的可积函数，所以，

$$\begin{aligned}\widehat{f*g}(\xi)&=\int_{\mathbb{R}^n}e^{-ix\cdot\xi}\left(\int_{\mathbb{R}^n}f(y)g(x-y)dy\right)dx\\&=\int_{\mathbb{R}^n}f(y)e^{-iy\cdot\xi}\underbrace{\left(\int_{\mathbb{R}^n}e^{-i(x-y)\cdot\xi}g(x-y)dx\right)}_{\widehat{g}(\xi)}dy\\&=\widehat{f}(\xi)\widehat{g}(\xi).\end{aligned}$$

所以命题成立。 □

注记. 我们上个学期已经证明了 $\left(L^1(\mathbb{R}^n),*\right)$ 是乘法封闭的。我们现在说明，这个乘法在 $L^1(\mathbb{R}^n)$ 中没有乘法单位元。假设 $e\in L^1(\mathbb{R}^n)$，使得对任意的 $f\in L^1(\mathbb{R}^n)$，我们都有

$$e*f=f.$$

那么，我们有

$$\widehat{e}\widehat{f}=\widehat{f}.$$

我们下面将构造 L^1 的函数 f（*Gauss* 函数的 *Fourier* 变换），使得 $\widehat{f}>0$，那么，$\widehat{e}\equiv 1$，这与 $e\in C_\circ(\mathbb{R}^n)$ 矛盾。

命题 450. 假设 $f\in L^1(\mathbb{R}^n_x)$，$g\in L^1(\mathbb{R}^n_\xi)$。那么，

$$\int_{\mathbb{R}^n}\widehat{f}(\xi)g(\xi)d\xi=\int_{\mathbb{R}^n}f(x)\widehat{g}(x)dx$$

证明: 注意到映射

$$\mathbb{R}^n\times\mathbb{R}^n\to\mathbb{C},\quad (x,\xi)\mapsto e^{-ix\cdot\xi}f(\xi)g(x)$$

是 $\mathbb{R}^n\times\mathbb{R}^n$ 上的可积函数，对此函数运用 Fubini 定理，我们得到

$$\begin{aligned}\int_{\mathbb{R}^n}\widehat{f}(\xi)g(\xi)d\xi&=\int_{\mathbb{R}^{2n}}e^{-ix\cdot\xi}f(\xi)g(x)dxd\xi\\&=\int_{\mathbb{R}^n}\left(\int_{\mathbb{R}^n}g(\xi)e^{-ix\xi}d\xi\right)f(x)dx\\&=\int_{\mathbb{R}^n}f(x)\widehat{g}(x)dx.\end{aligned}$$

命题得证。 □

注记. 这个命题在形式上建议了如何对分布定义 *Fourier* 变换：对于 $u \in \mathcal{D}'(\mathbb{R}^n_x)$ 和 $\varphi \in \mathcal{D}(\mathbb{R}^n_\xi)$，我们想定义

$$\langle \widehat{u}, \varphi \rangle = \langle u, \widehat{\varphi} \rangle.$$

然而，对于 $\varphi \in \mathcal{D}(\mathbb{R}^n)$，我们知道 $\widehat{\varphi}(\xi) \notin \mathcal{D}(\mathbb{R}^n)$，这是因为它的支集不可能是紧的：对于 $\mathbb{R}^1$ 上的 $\varphi(x)$，如果 $\widehat{\varphi}(\xi)$ 有紧支集，那么，如果把 $\widehat{\varphi}(\xi)$ 视作是 $\mathbb{C}$ 上的函数（即 $\xi \in \mathbb{C}$，此时 *Fourier* 变换仍然是良好定义），那么，这是一个复解析函数，因为

$$\overline{\partial}\,(\widehat{\varphi}) \equiv 0.$$

然而，这个复解析函数的 $\mathbb{R}$ 上的零点集不是离散的，所以只能恒为 0，我们也将证明，*Fourier* 变换为 0 的函数只有零函数。

综上所说，我们不能用上述的方式来定义分布的 *Fourier* 变换。为了解决这个问题，我们将找一个比 $\mathcal{D}(\mathbb{R}^n)$ 大一些的由光滑函数的空间 $\mathcal{S}(\mathbb{R}^n)$（*Schwartz* 空间），使得

$$\mathcal{F} : \mathcal{S}(\mathbb{R}^n) \to \mathcal{S}(\mathbb{R}^n).$$

（之前的问题出在 $\mathcal{F}(\mathcal{D}(\mathbb{R}^n)) \not\subset \mathcal{D}(\mathbb{R}^n)$。）

命题 451 (仿射坐标变换与 Fourier 变换之间的关系)**.** 任给 $f \in L^1(\mathbb{R}^n)$，$x_0 \in \mathbb{R}^n$ 和可逆线性变换

$$A : \mathbb{R}^n \to \mathbb{R}^n,$$

我们有

1) 物理空间的平移对应频率空间乘相应的频率，即

$$\mathcal{F}\big(f(\cdot + x_0)\big)(\xi) = e^{ix_0 \cdot \xi} \widehat{f}(\xi).$$

2) 频率空间应该视作是余切丛，即

$$\mathcal{F}\big(f \circ A\big)(\xi) = |\det(A)|^{-1} \widehat{f}({}^tA^{-1}\xi).$$

证明: 第一部分是显然的；为了证明第二部分，我们直接利用换元公式：

$$\begin{aligned}
\mathcal{F}\big(f \circ A\big)(\xi) &= \int_{\mathbb{R}^n} e^{-ix\cdot\xi} f(Ax) dx \\
&= |\det(A)|^{-1} \int_{\mathbb{R}^n} e^{-i(A^{-1}y)\cdot\xi} f(y) dy \\
&= |\det(A)|^{-1} \int_{\mathbb{R}^n} e^{-iy\cdot {}^tA^{-1}\xi} f(y) dy.
\end{aligned}$$

命题成立。 □

命题 452 (Gauss 函数的 Fourier 变换)**.** *对任意的正数 $\lambda>0$，我们有*

$$\mathcal{F}\left(e^{-\lambda\frac{|x|^2}{2}}\right)(\xi)=\left(\frac{2\pi}{\lambda}\right)^{\frac{n}{2}}e^{-\frac{|\xi|^2}{2\lambda}}.$$

证明: 通过作用变量替换 $y=\sqrt{\lambda}x$，上一个命题表明只要对 $\lambda=1$ 证明命题即可，即

$$\mathcal{F}\left(e^{-\frac{|x|^2}{2}}\right)(\xi)=(2\pi)^{\frac{n}{2}}e^{-\frac{|\xi|^2}{2}}.$$

根据 Fubini 定理，我们进一步有

$$\begin{aligned}\mathcal{F}\big(e^{-\frac{|x|^2}{2}}\big)(\xi)&=\int_{\mathbb{R}^n}e^{-ix\cdot\xi}e^{-\frac{|x|^2}{2}}dx\\&=\prod_{k\leqslant n}\int_{\mathbb{R}}e^{-ix_k\cdot\xi_k}e^{-\frac{x_k^2}{2}}dx_k\end{aligned}$$

所以，只要对 $n=1$ 来证明即可。此时，我们就必须去计算

$$F(\xi)=\int_{\mathbb{R}}e^{-ix\xi-\frac{x^2}{2}}dx.$$

我们一下给出两种方式进行计算:

- 按实变量的观点。

  我们对 ξ 求导。根据 Lebesgue 控制收敛定理的推论，以下导数与积分之间的交换都是合理的：根本的原因是 $e^{-|x|^2}$ 的衰减足够快。所以

  $$\begin{aligned}F'(\xi)&=\int_{\mathbb{R}}-ixe^{-ix\xi}e^{-\frac{x^2}{2}}dx=\int_{\mathbb{R}}ie^{-ix\xi}\big(e^{-\frac{x^2}{2}}\big)'dx\\&=-\int_{\mathbb{R}}\big(ie^{-ix\xi}\big)'e^{-\frac{x^2}{2}}dx=-\xi F(\xi).\end{aligned}$$

  由于一个函数的积分就是其 Fourier 变换在 0 处取值，所以 $F(\xi)$ 满足如下的线性常微分方程（因为 $\int_{\mathbb{R}}e^{-\frac{x^2}{2}}dx=\sqrt{2\pi}$）

  $$\begin{cases}F'(\xi)+\xi F(\xi)=0,\\F(\xi)=\sqrt{2\pi}\end{cases}$$

  根据常微分方程的理论，这个方程的解是唯一的。我们只要验证

  $$F(\xi)=\sqrt{2\pi}e^{-\frac{\xi^2}{2}}$$

  是解即可，这是平凡的。

- 按复变量的观点。

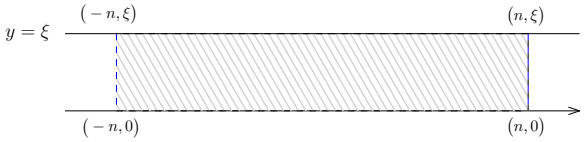
  注意到，我们可以按照复变函数的观点来看这个积分:

  $$F(\xi)=e^{-\frac{\xi^2}{2}}\int_{\mathbb{R}}e^{-\frac{(x+i\xi)^2}{2}}dx=e^{-\frac{\xi^2}{2}}\int_{\mathrm{Im}(z)=\xi}e^{-\frac{z^2}{2}}dz.$$

也就是说，我们需要计算在 $\mathbb{C}$ 上 $y=\xi$ 这条直线上的积分

$$I(\xi)=\int_{\mathrm{Im}(z)=\xi} e^{-\frac{z^2}{2}}dz.$$

当 $\xi=0$ 时，这就是 Guass 积分，我们早就知道 $I(0)=\sqrt{2\pi}$。当 $\xi\neq 0$ 时候，我们要用 Cauchy 积分公式把 $y=\xi$ 上的积分化成 $y=0$ 上的积分。



我们考虑复解析函数 $e^{-\frac{z^2}{2}}$ 以 $(\pm n,0)$ 和 $(\pm n,\xi)$ 作为长方形的上的积分（逆时针）。由于 $e^{-\frac{z^2}{2}}$ 是解析的（根据留数定理，在长方形内部没有极点），所以

$$\left(\int_{-n\leqslant x\leqslant n,y=0}+\int_{x=n,0\leqslant y\leqslant\xi}-\int_{-n\leqslant x\leqslant n,y=\xi}-\int_{x=-n,0\leqslant y\leqslant\xi}\right)e^{-\frac{z^2}{2}}dz=0.$$

由于

$$e^{\frac{z^2}{2}}=e^{-\frac{x^2-y^2}{2}}e^{xyi},\quad |y|\leqslant|\xi|,$$

这是一个对于 x 是指数衰减的函数，所以上面积分中的第二项和第四项当 $n\to\infty$ 时，为零。从而，对 n 取极限，我们就得到

$$I(\xi)=I(0)=\sqrt{2\pi}.$$

这同样证明了结论。

□

利用上面 Guass 函数的 Fourier 变换，我们可以证明关于 Fourier 逆变换的定理。

定义 453. *对任意的 $g\in L^1(\mathbb{R}^n_\xi)$，对任意的 $x\in\mathbb{R}^n$ 我们定义*

$$\left(\mathcal{F}^{-1}(g)\right)(x)=\frac{1}{(2\pi)^n}\int_{\mathbb{R}^n}e^{i\xi\cdot x}g(\xi)d\xi.$$

与 Fourier 变换的证明一样，上面的得到的 $\left(\mathcal{F}^{-1}(g)\right)(x)$ 是一个良好定义的 $C_\circ(\mathbb{R}^n)$ 中的函数，我们不再赘述其证明。

定理 454 (Fourier 逆变换)**.** *给定 $f\in L^1(\mathbb{R}^n_x)$，如果 $\widehat{f}\in L^1(\mathbb{R}^n_\xi)$，那么，我们有*

$$f=\mathcal{F}^{-1}\widehat{f},$$

其中，上面等号成立是在 $L^1(\mathbb{R}^n)$ 的意义下的。

证明: 我们考虑如下的 Gauss 分布函数

$$G_\lambda(x) = \big(\frac{1}{2\pi\lambda^2}\big)^{\frac{n}{2}} e^{-\frac{|x|^2}{2\lambda^2}}$$

其中 $\lambda > 0$。根据上面的计算，我们有

$$\widehat{G_\lambda}(\xi) = e^{-\frac{\lambda^2|\xi|^2}{2}}.$$

所以，

$$\begin{aligned}
\mathcal{F}^{-1}\left(\widehat{G_\lambda}(\xi)\right)(-x) &= \frac{1}{(2\pi)^n}\int_{\mathbb{R}^n} e^{-\frac{\lambda^2|\xi|^2}{2}} e^{-ix\cdot\xi} d\xi \\
&= \frac{1}{(2\pi)^n}\left(\mathcal{F}(e^{-\frac{\lambda^2|\xi|^2}{2}})\right)(x) \\
&= G_\lambda.
\end{aligned}$$

最后一步我们再一次用到了之前的计算。通过观察 $\widehat{G_\lambda}(\xi)$ 的表达式，我们注意到，对任意的 $\xi \in \mathbb{R}^n$，有

$$\lim_{\lambda\to 0}\widehat{G_\lambda}(\xi) = 1.$$

根据 Lebesgue 控制收敛定理，我们有

$$\begin{aligned}
\mathcal{F}^{-1}\big(\widehat{f}\big)(x) &= \frac{1}{(2\pi)^{2n}}\int_{\mathbb{R}^n} e^{i\xi\cdot x}\widehat{f}(\xi)d\xi \\
&= \lim_{\lambda\to 0}\frac{1}{(2\pi)^n}\int_{\mathbb{R}^n}\widehat{G_\lambda}(\xi)e^{i\xi\cdot x}\widehat{f}(\xi)d\xi,
\end{aligned}$$

其中，控制函数就选 $\widehat{f}$。利用 Fubini 定理可知

$$\begin{aligned}
\mathcal{F}^{-1}\big(\widehat{f}\big)(x) &= \lim_{\lambda\to 0}\frac{1}{(2\pi)^n}\int_{\mathbb{R}^n}\underbrace{\int_{\mathbb{R}^{2n}}\widehat{G_\lambda}(\xi)e^{i\xi\cdot(x-y)}d\xi}_{\text{Fourier 逆变换}} f(y)dy \\
&= \lim_{\lambda\to 0}\frac{1}{(2\pi)^n}\int_{\mathbb{R}^n} G_\lambda(x-y)f(y)dy \\
&= \lim_{\lambda\to 0} G_\lambda * f.
\end{aligned}$$

由于 G_λ 的积分恰好为 1，所以当 $\lambda \to 0$ 时，在 L^1 的意义下，

$$\lim_{\lambda\to 0} G_\lambda * f \overset{L^1}{=} f.$$

命题成立。 □

66 L^2 上的 Fourier 变换，Planchrel 公式，Schwartz 空间，Schwartz 函数的例子，试验函数在 Schwartz 空间中的稠密性，Schwartz 空间上的 Fourier 变换，缓增分布的定义

二零二零年十月十四日，星期三，晴

$L^2(\mathbb{R}^n)$ 上的 Fourier 变换

利用 Fourier 逆变换，我们可以在 L^2 上定义 Fourier 变换。注意到，对于 $f \in L^2(\mathbb{R}^n)$，积分

$$\int_{\mathbb{R}^n} e^{-ix\cdot\xi} f(x)dx$$

可能并没有定义，比如 $\mathbb{R}^1$ 上的函数

$$f(x) = \frac{1}{1+|x|}.$$

注意到，$C_0^\infty(\mathbb{R}^n) \subset L^2(\mathbb{R}^n)$ 是稠密的子空间。我们任意选取 $f \in C_0^\infty(\mathbb{R}^n)$，很明显，$\widehat{f} \in L^1(\mathbb{R}^n)$（因为光滑性意味着衰减很快，所以可积）。另外，我们有

$$\overline{\widehat{f}}(\xi) = \int_{\mathbb{R}^n} e^{ix\cdot\xi}\overline{f(x)}dx = (2\pi)^n \mathcal{F}^{-1}(\overline{f})(\xi).$$

所以，我们有

$$\begin{aligned}
\|\widehat{f}(\xi)\|_{L^2}^2 &= \int_{\mathbb{R}^n} \widehat{f}(\xi)\overline{\widehat{f}(\xi)}d\xi \\
&= (2\pi)^n \int_{\mathbb{R}^n} \widehat{f}(\xi)\mathcal{F}^{-1}\big(\overline{f(\xi)}\big)d\xi \\
&= (2\pi)^n \int_{\mathbb{R}^n} f(x)\mathcal{F}\left(\mathcal{F}^{-1}\big(\overline{f(\xi)}\big)\right)(x)d\xi \\
&= (2\pi)^n \int_{\mathbb{R}^n} f(x)\overline{f}(x)d\xi.
\end{aligned}$$

从而，

$$\|\widehat{f}(\xi)\|_{L^2} = (2\pi)^{\frac{n}{2}}\|f\|_{L^2}.$$

这表明定义在 $L^2(\mathbb{R}^n)$ 是稠密的子空间 $C_0^\infty(\mathbb{R}^n)$ 上的 Fourier 变换

$$\mathcal{F} : C_0^\infty(\mathbb{R}^n) \to L^2(\mathbb{R}^n), \quad f \mapsto \widehat{f}$$

是连续的。

$$\begin{array}{ccc}
C_0^\infty(\mathbb{R}^n) & \xrightarrow{\ \iota\ } & L^2(\mathbb{R}^n) \\
& \searrow^{\mathcal{F}} & \big\downarrow{\scriptstyle \mathcal{F}} \\
& & L^2(\mathbb{R}^n)
\end{array}$$

根据连续线性映射扩张的定理，我们就证明了

定理 455 (L^2 上的 Fourier 变换与 Planchrel 公式)**.** 我们可以定义 *Fourier* 变换 $\mathcal{F}$:

$$\mathcal{F}: L^2(\mathbb{R}^n) \to L^2(\mathbb{R}^n)$$

使得

$$(2\pi)^{\frac{n}{2}}\mathcal{F}: L^2(\mathbb{R}^n) \longrightarrow L^2(\mathbb{R}^n)$$

是等距同构。

特别地，对于 $f \in L^1(\mathbb{R}^n) \cap L^2(\mathbb{R}^n)$，我们有

$$\|\widehat{f}\|_{L^2}^2 = (2\pi)^n \|f\|_{L^2}^2.$$

通过极化，我们有对任意的 $f, g \in L^2(\mathbb{R}^n)$，我们有

$$(\widehat{f}, \widehat{g})_{L^2} = (2\pi)^n (f, g)_{L^2}.$$

证明*:* 上述一切叙述对 $C_0^\infty(\mathbb{R}^n)$ 是成立的。对一般的 $f \in L^2(\mathbb{R}^n)$, 用 $C_0^\infty(\mathbb{R}^n)$ 中函数逼近即可。 □

注记. 上面的定理定义了 $f \in L^2(\mathbb{R}^n)$ 的 *Fourier* 变换，为了行文清楚，我们暂且把 L^2 意义下的 *Fourier* 变换记作 $\mathcal{F}_2$。另外，对于 $g \in L^1(\mathbb{R}^n)$，它的 *Fourier* 变换是可以用 *Fourier* 积分表示的，我们把它记做 $\mathcal{F}_1$，也就是说

$$\mathcal{F}_1(g) = \int_{\mathbb{R}^n} g(x) e^{-ix\xi} dx.$$

那么，对于 $f \in L^1(\mathbb{R}^n) \cap L^2(\mathbb{R}^n)$，我们有

$$\mathcal{F}_1(f) = \mathcal{F}_2(f).$$

这个有趣的验证我们留作作业。

例子. 考虑 $\mathbb{R}^1$ 上的 L^2-函数

$$u(x) = (1 + |x|)^{-\alpha},$$

其中，$\frac{1}{2} < \alpha \leqslant 1$。很明显，我们知道 $e^{-ix\xi}u(x) \notin L^1(\mathbb{R}_x)$，所以我们不能直接用 L^1-函数的 *Fourier* 积分来写它的 *Fourier* 变换。然而，我们知道

$$u_n(x) = u(x)\mathbf{1}_{|x|\leqslant n}$$

是 L^1 的，我们可以先显式写下 u_n 的 *Fourier* 变换。由于序列 $\{u_n\}_{n\geqslant 1}$ 在 L^2 中逼近 u，所谓，我们有

$$\widehat{u} \overset{L^2}{=} \lim_{n\to} \widehat{u_n}.$$

练习. 证明，对任意的 $f \in L^2(\mathbb{R}^n)$，我们都有

$$\widehat{\check{f}} = \check{f} \quad \Leftrightarrow \quad \mathcal{F}^2(f) = \check{f}.$$

Schwartz 空间

对任意的给定的函数 f，对任意的多重指标 α，我们采用如下的符号：

$$x^\alpha f(x) = x_1^{\alpha_1} \cdots x_n^{\alpha_n} f(x_1, \cdots, x_n).$$

定义 456 (Schwartz 空间)**.** 函数 φ 是 $\mathbb{R}^n$ 上的光滑函数。如果 φ 满足如下的条件：对任意的多重指标 α, β，我们都有

$$x^\alpha \partial^\beta \varphi(x) \in L^\infty(\mathbb{R}^n),$$

那么，我们就称 φ 是一个 **Schwartz 函数**或者是一个**速降的函数**。我们把 $\mathbb{R}^n$ 上所有的 *Schwartz* 函数所构成的线性空间称作是 **Schwartz 空间**，并记作 $\mathcal{S}(\mathbb{R}^n)$。

对于每个非负整数 $p \in \mathbb{Z}_{\geqslant 0}$，我们定义如下的（一族）范数：

$$N_p(\varphi) = \sum_{\substack{|\alpha| \leqslant p, \\ |\beta| \leqslant p}} \|x^\alpha \partial^\beta \varphi(x)\|_{L^\infty(\mathbb{R}^n)}.$$

在 $\mathcal{S}(\mathbb{R}^n)$ 上，我们规定如下的收敛性（拓扑）：给定 *Schwartz* 函数的序列 $\{\varphi_k\}_{p=1,2,\cdots} \subset \mathcal{S}(\mathbb{R}^n)$，它收敛到 *Schwartz* 函数 $\varphi \in \mathcal{S}(\mathbb{R}^n)$，指的是对任意的非负整数 p，我们都有

$$\lim_{n\to\infty} N_p(\varphi_n - \varphi) = 0.$$

我们把这个极限简写成

$$\varphi_k \xrightarrow{\mathcal{S}(\mathbb{R}^n)} \varphi.$$

例子. 我们已经见过很多 *Schwartz* 函数

- $\mathcal{D}(\mathbb{R}^n) \subset \mathcal{S}(\mathbb{R}^n)$；
- $e^{-x^2} \in \mathcal{S}(\mathbb{R}^n)$；
- 对于 *Schwartz* 函数 $\varphi \in \mathcal{S}(\mathbb{R}^n)$，对它求若干次导数或者乘以一个多项式仍然是一个 *Schwartz* 函数，即对任意的多重指标 α, β，我们有

$$x^\alpha : \mathcal{S}(\mathbb{R}^n) \to \mathcal{S}(\mathbb{R}^n),$$
$$\partial^\beta : \mathcal{S}(\mathbb{R}^n) \to \mathcal{S}(\mathbb{R}^n).$$

上面例子的验证我们留作作业。

给定一个 Schwartz 函数，我们对它有如下的估计：对于任何的多重指标 α 和 β，其中 $|\alpha|, |\beta| \leqslant p$，我们有

$$|(1+|x|)^{n+1} x^\alpha \partial^\beta \varphi(x)| \leqslant N_{p+n+1}(\varphi),$$

其中，n 是空间的维数。从而，对任意的 $x \in \mathbb{R}^n$，我们有

$$|x^\alpha \partial^\beta \varphi(x)| \leqslant \frac{N_{p+n+1}(\varphi)}{(1+|x|)^{n+1}}.$$

上式右边的函数是可积的，所以，

$$\|x^\alpha \partial^\beta \varphi(x)\|_{L^1(\mathbb{R}^n)} \leqslant C_n N_{p+d+1}(\varphi).$$

特别地，我们可以对 $x^\alpha \partial^\beta \varphi(x)$ 用 Fourier 积分来定义其 Fourier 变换。作为推论，我们还知道

$$\mathcal{S}(\mathbb{R}^n) \subset L^1(\mathbb{R}^n).$$

另外，以上的估计是常用的技巧，在后面的不少场合都会用到。

定理 457. *$\mathcal{D}(\mathbb{R}^n)$ 在 $\mathcal{S}(\mathbb{R}^n)$ 中是稠密的，即对任意的 $\varphi \in \mathcal{S}(\mathbb{R}^n)$，存在函数序列 $\{\varphi_n\}_{n\geqslant 1} \subset \mathcal{D}(\mathbb{R}^n)$，使得*

$$\varphi_k \xrightarrow{\mathcal{S}(\mathbb{R}^n)} \varphi.$$

证明： 我们选取有紧支集的光滑函数 $\chi(x)$，使得

$$\begin{cases} \chi(x) = 1, & |x| = 1; \\ 0 \leqslant \chi(x) \leqslant 1. \end{cases}$$

对于 $\varphi(x) \in \mathcal{S}(\mathbb{R}^n)$，我们令

$$\varphi_k(x) = \chi\left(\frac{x}{k}\right)\varphi(x) \in \mathcal{D}(\mathbb{R}^n).$$

我们只要证明，对任意的非负整数 p，我们有

$$N_p(\varphi_k - \varphi) \to 0,$$

即可。对于满足 $|\alpha| \leqslant p$, $|\beta| \leqslant p$ 的多重指标，我们有

$$\begin{aligned} x^\alpha \partial^\beta(\varphi - \varphi_k) &= x^\alpha \partial^\beta \left(\left(1 - \chi\left(\frac{x}{k}\right)\right)\varphi\right) \\ &= \left(1 - \chi(\frac{x}{k})\right) x^\alpha \partial^\beta \varphi - \underbrace{\sum_{0 \neq \gamma \leqslant \beta} \frac{1}{k^{|\gamma|}} \frac{\beta!}{\gamma!(\beta-\gamma)!} \underbrace{x^\alpha \partial^{\beta-\gamma}\varphi(x)}_{|\cdot| \leqslant N_p(\varphi)} (\partial^\gamma \chi)\left(\frac{x}{k}\right)}_{\longrightarrow 0,\ k\to\infty} \end{aligned}$$

上式的第二个求和部分有 k^{-1} 这样的衰减因子，所以极限为 0。对于第一项，由于 χ 在半径为 1 的球内部为 1，所以然而，我们有

$$\begin{aligned} |(1 - \chi(\frac{x}{k}))x^\alpha \partial^\beta \varphi| &\leqslant \mathbf{1}_{|x|\geqslant k}(x) \cdot |x|^{-2} \cdot |x^{\alpha+2}\partial^\beta \varphi| \\ &\leqslant \frac{1}{k^2} N_{p+2}(\varphi) \to 0. \end{aligned}$$

这就完成了证明。 □

我们现在研究 Schwartz 函数的 Fourier 变换。我们已经证明了 $\mathcal{S}(\mathbb{R}^n) \subset L^1(\mathbb{R}^n)$，所以，$\varphi \in \mathcal{S}(\mathbb{R}^n)$ 的 Fourier 变换还是可以用积分公式

$$\mathcal{F}(\varphi)(\xi) = \widehat{\varphi}(\xi) = \int_{\mathbb{R}^n} \varphi(x) e^{-ix\cdot\xi} dx$$

来表示。

定理 458 ($\mathcal{S}(\mathbb{R}^n)$ 上的 Fourier 变换)**.** *如果 $\varphi \in \mathcal{S}(\mathbb{R}^n)$ 是 Schwartz 函数，那么，$\widehat{\varphi} \in \mathcal{S}(\mathbb{R}^n)$。在 Schwartz 函数空间上的 Fourier 变换：*

$$\mathcal{F} : \mathcal{S}(\mathbb{R}^n) \longrightarrow \mathcal{S}(\mathbb{R}^n), \quad \varphi \mapsto \widehat{\varphi}(\xi),$$

满足如下的性质：对任意的 $p \in \mathbb{Z}_{\geqslant 0}$，存在常数 $C_p > 0$，使得对每个 $\varphi \in \mathcal{S}(\mathbb{R}^n)$，我们都有

$$N_p(\widehat{\varphi}) \leqslant C_p N_{p+n+1}(\varphi).$$

特别的，$\mathcal{F} : \mathcal{S}(\mathbb{R}^n) \longrightarrow \mathcal{S}(\mathbb{R}^n)$ 是连续的线性同构，即对任意的在 $\mathcal{S}(\mathbb{R}^n)$ 中收敛的函数序列

$$\varphi_k \xrightarrow{\mathcal{S}} \varphi, \quad k \to \infty,$$

我们有那

$$\widehat{\varphi_k} \xrightarrow{\mathcal{S}} \widehat{\varphi}, \quad k \to \infty.$$

另外，对任意的 $\varphi \in \mathcal{S}(\mathbb{R}^n)$，我们还有公式

$$\widehat{\partial_k \varphi} = i\xi_k \widehat{f}, \quad \widehat{x_k \varphi} = i\partial_k \widehat{\varphi}.$$

证明： 我们首先证明叙述中的最后两个恒等式。对任意的 Schwartz 函数 φ，对任意的 $k \leqslant n$，利用分部积分，我们有

$$\widehat{\partial_k \varphi}(\xi) = -\int_{\mathbb{R}^n} (-i\xi_k) e^{-ix\cdot\xi} \varphi(x) dx = i\xi_k \widehat{\varphi}(\xi).$$

第二个等式要用 Lebesgue 控制收敛的推论（积分与求导数可交换），我们有

$$\partial_{\xi_k} \widehat{\varphi}(\xi) = \int_{\mathbb{R}^n} (-ix_k) e^{-ix\cdot\xi} \varphi(x) dx = \widehat{-ix_k \varphi}(\xi).$$

现在证明定理中的不等式（从而证明了 Fourier 变换 $\mathcal{F}$ 的像也落在 $\mathcal{S}(\mathbb{R}^n)$ 中）。固定两个多重指标 α 和 β，其中 $|\alpha|, |\beta| \leqslant p$。利用已经证明的公式，我们就有

$$\begin{aligned} |\xi^\alpha \partial_\xi^\beta \widehat{\varphi}(\xi)| &= |\widehat{\partial^\alpha (x^\beta \varphi)}(\xi)| \leqslant \|\partial^\alpha (x^\beta \varphi)(x)\|_{L^1} \\ &\leqslant C_p N_{p+n+1}(\varphi). \end{aligned}$$

Fourier 变换的连续性可以通过这个不等式得到：对任意给定 p，我们有

$$N_p(\widehat{\varphi_k} - \widehat{\varphi}) \leqslant C_p N_{p+n+1}(\varphi_k - \varphi) \to 0,$$

按照定义，我们就有

$$\widehat{\varphi_k} \xrightarrow{\mathcal{S}} \widehat{\varphi}, \quad k \to \infty.$$

最后，我们来说明 $\mathcal{F}$ 是同构。实际上，我们可以定义直接考虑 Fourier 变换的逆 $\mathcal{F}^{-1}$，因为 $\widehat{\varphi} \in L^1$，所以之前定义的 $\mathcal{F}^{-1}$ 在此也是良好定义的。此时，我们已经证明了 $\mathcal{F}$ 与 $\mathcal{F}^{-1}$ 互为逆映射，所以命题得证（$\mathcal{F}^{-1}$ 也是连续的）。 □

定义 459 (缓增的分布)**.** 假设 $u \in \mathcal{D}'(\mathbb{R}^n)$ 是一个分布。如果存在非负整数 p 和常数 $C > 0$，使得对每个 $\varphi \in \mathcal{D}(\mathbb{R}^n)$，我们都有

$$|\langle u, \varphi \rangle| \leqslant C N_p(\varphi),$$

我们就说 u 是一个**缓增的**分布。我们用 $\mathcal{S}'(\mathbb{R}^n)$ 来表示所有缓增分布的集合，很明显

$$\mathcal{S}'(\mathbb{R}^n) \subset \mathcal{D}'(\mathbb{R}^n)$$

是线性子空间。

定理 460 ($\mathcal{S}'(\mathbb{R}^n) \times \mathcal{S}(\mathbb{R}^n) \to \mathbb{R}$)**.** 对任意的 $u \in \mathcal{S}'(\mathbb{R}^n)$，存在唯一一个线性泛函

$$T_u : \mathcal{S}(\mathbb{R}^n) \to \mathbb{C},$$

使得存在非负整数 p 和常数 $C > 0$，对每个 $\varphi \in \mathcal{S}'(\mathbb{R}^n)$，我们都有

$$\big|T_u(\varphi)\big| \leqslant C N_p(\varphi),$$

并且对于任意的 $\psi \in \mathcal{D}(\mathbb{R}^n)$，都有

$$T_u(\psi) = \langle u, \psi \rangle.$$

在后面的场合，为了简单起见，我们仍将此线性泛函的作用记作 $\langle u, \cdot \rangle$。

证明: 利用 $\mathcal{D}(\mathbb{R}^n) \subset \mathcal{S}(\mathbb{R}^n)$ 的稠密性，我们用极限的形式来定义 T_u：对任意给定的 $\varphi \in \mathcal{S}(\mathbb{R}^n)$，我们选取试验函数序列 $\{\varphi_k\}_{k \geqslant 1} \subset \mathcal{D}(\mathbb{R}^n)$，使得

$$\varphi_k \xrightarrow{\mathcal{S}} \varphi.$$

我们令

$$T_u(\varphi) = \lim_{k \to \infty} \langle u, \varphi_k \rangle.$$

当然，我们需要说明上述极限存在并且不依赖于逼近序列的选取。

利用 $u \in \mathcal{S}'(\mathbb{R}^n)$ 的定义，存在 C_0 和 p_0，使得对任意的 $k, \ell \geqslant 1$，我们都有

$$|\langle u, \varphi_k - \varphi_\ell \rangle| \leqslant C_0 N_{p_0}(\varphi_k - \varphi_\ell).$$

再利用 $\varphi_k \xrightarrow{\mathcal{S}} \varphi$ 的定义，我们有

$$N_{p_0}(\varphi_k - \varphi) \to 0.$$

所以，我们有

$$\begin{aligned} |N_{p_0}(\varphi_k) - N_{p_0}(\varphi_\ell)| &= N_{p_0}(\varphi_k - \varphi_\ell) \\ &\leqslant N_{p_0}(\varphi_k - \varphi) + N_{p_0}(\varphi_\ell - \varphi) \xrightarrow{k,\ell \to \infty} 0. \end{aligned}$$

这说明 $\{\langle u, \varphi_k \rangle\}_{k \geqslant 1}$ 是 Cauchy 列，所以所讨论的极限是存在的。

下面再说明这个极限不依赖于逼近序列 $\{\varphi_k\}_{k\geqslant 1}\subset\mathcal{D}(\mathbb{R}^n)$ 的选取。假设 $\{\psi_k\}_{k\geqslant 1}\subset\mathcal{D}(\mathbb{R}^n)$ 是另一个逼近序列，那么，我们有

$$\begin{aligned}|N_{p_0}(\varphi_k)-N_{p_0}(\psi_k)| &= N_{p_0}(\varphi_k-\psi_k)\\ &\leqslant N_{p_0}(\varphi_k-\varphi)+N_{p_0}(\varphi_\ell-\varphi)\overset{k,\ell\to\infty}{\longrightarrow}0.\end{aligned}$$

所以，

$$|\langle u,\varphi_k\rangle-\langle u,\psi_k\rangle|\leqslant C_0N_{p_0}(\varphi_k-\psi_k)\to 0.$$

所以，

$$\lim_{k\to\infty}\langle u,\varphi_k\rangle=\lim_{k\to\infty}\langle u,\psi_k\rangle.$$

最后，我们证明 T_μ 所满足的不等式：

$$|\langle u,\varphi_k\rangle|\leqslant C_0N_{p_0}(\varphi_k)\leqslant C_0\left(\underbrace{N_{p_0}(\varphi_k-\varphi)}_{\to 0,\ k\to\infty}+N_{p_0}(\varphi)\right)$$

两边同时取极限，这就完成了定理的证明。 □

66.1 作业：Fourier 逆变换的另一个计算，一个分布扩张的问题，分布的张量积

清华大学 19-20 春季学期，数学分析三，作业 3

请用 A4 大小的纸张**正反面**用钢笔，签字笔或者圆珠笔书写，并注明自己的姓名，年级和作业的**总页数**。除定理公式所涉及的人名之外，请使用**中文**。本次作业的提交时间和地点为 **10 月 28 日**（周三）上午的课堂上，逾期视作零分。

习题 A.（课堂细节的补充） 我们总假设 $\Omega \subset \mathbb{R}^n$ 是非空的开集。

A1) 在第七课中（命题 43），对于 $u, v \in \mathcal{D}'(\mathbb{R}^n)$，其中 $\operatorname{supp}(u)$ 和 $\operatorname{supp}(v)$ 是可卷的，我们定义了

$$u * v \in \mathcal{D}'(\mathbb{R}^n).$$

我们课程上没有验证 $u * v$ 满足分布定义中所要求的不等式。试补充这部分的细节。

A2) 对于 $u, v, w \in \mathcal{D}'(\mathbb{R}^n)$，假设 $\operatorname{supp}(u)$，$\operatorname{supp}(v)$ 和 $\operatorname{supp}(w)$ 是可卷的。试证明：

1) $\operatorname{supp}(u * v) \subset \operatorname{supp}(u) + \operatorname{supp}(v)$；
2) $u * v = v * u$；
3) $(u * v) * w = u * (v * w)$；

A3) 在第七课中，对于给定的光滑初始值，我们证明 $\mathbb{R}^{1+3}$ 上的波动方程的解的表示公式：对于 $t \geqslant 0$，有

$$u(t,x) = -t \int_{\mathbf{S}^2} u_1(x - t\vartheta') \frac{d\sigma_{\mathbf{S}^2}(\vartheta')}{4\pi} - \partial_t \left[t \int_{\mathbf{S}^2} u_0(x - t\vartheta') \frac{d\sigma_{\mathbf{S}^2}(\vartheta')}{4\pi} \right].$$

证明的过程中我们假设了初始值 u_0 和 u_1 都有紧支集。证明，这个紧支集的要求是冗余的。

A5) 证明，对任意的 $f \in L^2(\mathbb{R}^n)$，我们都有

$$\widehat{\hat{f}} = \check{f} \quad \Leftrightarrow \quad \mathcal{F}^2(f) = \check{f}.$$

A4) 把 L^2 意义下的 Fourier 变换记作 $\mathcal{F}_2$；L^1 函数的 Fourier 变换可以用 Fourier 积分表示的，我们把它记做 $\mathcal{F}_1$。证明，对于 $f \in L^1(\mathbb{R}^n) \cap L^2(\mathbb{R}^n)$，我们有

$$\mathcal{F}_1(f) = \mathcal{F}_2(f).$$

A6) 对于 Schwartz 函数 $\varphi \in \mathcal{S}(\mathbb{R}^n)$，对它求若干次导数或者乘以一个多项式仍然是一个 Schwartz 函数，即对任意的多重指标 α, β，我们有

$$x^\alpha : \mathcal{S}(\mathbb{R}^n) \to \mathcal{S}(\mathbb{R}^n),$$
$$\partial^\beta : \mathcal{S}(\mathbb{R}^n) \to \mathcal{S}(\mathbb{R}^n).$$

A7)（平移不变算子与卷积）给定连续 $\mathbb{C}$-线性映射（线性算子）

$$T : C_0^\infty(\mathbb{R}^n) \to C_0^\infty(\mathbb{R}^n),$$

其中，连续性指的是对任意的紧集 $K \subset \mathbb{R}^n$，对任意的非负整数 p，存在紧集 $L \subset \mathbb{R}^n$、非负整数 q 和常数 $C > 0$，使得对任意的 $\varphi \in C_0^\infty(\mathbb{R}^n)$，我们都有

$$\sup_{|\alpha| \leqslant p} \|\partial^\beta (T\varphi)\|_{L^\infty(K)} \leqslant C \sup_{|\beta| \leqslant q} \|\partial^\beta \varphi\|_{L^\infty(L)}.$$

我们假设 T 与 $\mathbb{R}^n$ 上的平移算子交换，即对任意的 $a \in \mathbb{R}^n$，对任意的 $\varphi \in C_0^\infty(\mathbb{R}^n)$，我们都有

$$T\left(\tau_a \varphi\right) = \tau_a \left(T(\varphi)\right).$$

证明，存在 $u \in \mathcal{E}'(\mathbb{R}^n)$，使得对任意的 $\varphi \in C_0^\infty(\mathbb{R}^n)$，我们都有

$$T(\varphi) = u * \varphi.$$

习题 B. 一个分布的问题

对于 $\lambda \in (-2, +\infty)$ 和 $\varphi \in \mathcal{D}(\mathbb{R})$，我们定义

$$\langle u_\lambda, \varphi \rangle = \int_0^1 x^\lambda \big(\varphi(x) - \varphi(0)\big) dx.$$

B1) 证明，$u_\lambda \in \mathcal{D}'(\mathbb{R})$ 并确定它的阶和支集。

B2) 证明，$u_\lambda \notin L^2(\mathbb{R})$。

B3) 证明，$x u_\lambda \in L^2(\mathbb{R})$。

B4) 证明，当 $\lambda \in (-2, +\infty)$ 时，我们有

$$(u_{\lambda+1})' = (\lambda + 1) u_\lambda - \delta_1 + \delta_0 - \frac{1}{\lambda + 2} \delta_0'.$$

B5) 对于 $\lambda \in (-2, -1) \cup (-1, +\infty)$，试找出 u_λ 的一个原函数，即某个 $v_\lambda \in \mathcal{D}'(\mathbb{R})$，使得

$$v_\lambda' = u_\lambda.$$

B6) 对于 $\lambda_0 \in (-2, +\infty)$，证明，当 $\lambda \to \lambda_0$ 时，我们有

$$u_\lambda \xrightarrow{\mathcal{D}'(\mathbb{R})} u_{\lambda_0}.$$

B7) 证明，当 $\lambda \to -2$ 时（$\lambda > -2$），$\lim\limits_{\lambda \to -2} (\lambda + 2) u_\lambda$ 作为分布的极限存在并计算该分布。

B8) 证明，存在 $v_{-1} \in L^2(\mathbb{R})$，使得 $u_{-1} = (v_{-1})'$。

习题 C. Fourier 逆变换的另一个计算

我们对 $f \in L^1(\mathbb{R})$，定义其 Fourier 逆变换为

$$\big(\mathcal{F}^{-1} f\big)(x) = \frac{1}{2\pi} \int_{\mathbb{R}} e^{ix\xi} f(\xi) d\xi.$$

C1) 对正数 $a > 0$，试计算积分

$$\int_0^{\infty} \cos(ax) e^{-x} dx.$$

C2) 我们定义函数

$$h(\xi) = e^{-|\xi|}, \quad d(x) = \frac{1}{\pi(1+x^2)}.$$

证明，

$$d(x) = \big(\mathcal{F}^{-1} h\big)(x).$$

C3) 给定 $\varepsilon > 0$，定义

$$h_\varepsilon(\xi) = h(\varepsilon\xi), \quad d_\varepsilon(x) = \frac{1}{\varepsilon} d\left(\frac{x}{\varepsilon}\right).$$

证明，

$$\big(\mathcal{F}^{-1} h_\varepsilon\big)(x) = d_\varepsilon(x).$$

C4) 证明，

$$(f * d_\varepsilon)(x) = \frac{1}{2\pi} \int_{\mathbb{R}} h_\varepsilon(\xi) \widehat{f}(\xi) e^{ix\xi} d\xi.$$

C5) 证明，函数

$$g(y) = \int_{\mathbb{R}} |f(x-y) - f(x)| dx$$

是 $\mathbb{R}$ 上的有界连续函数。

C6) 证明，当 $\varepsilon \to 0$ 时，我们有

$$\|f * \delta_\varepsilon - f\|_{L^1} \to 0.$$

C7) 证明，存在序列 $\{\varepsilon_k\}_{k \geqslant 1}$，使得 $\varepsilon_n \to 0$ 并且

$$(f * \delta_\varepsilon)(x) \to f(x)$$

几乎处处成立。

C8) 假设 $\widehat{f}(\xi) \in L^1(\mathbb{R})$，证明，

$$\big(\mathcal{F}^{-1} \widehat{f}\big)(x) = f(x)$$

几乎处处成立。

习题 D. 关于分布扩张的一个问题

D1) 证明，线性映射

$$\mathcal{D}(\mathbb{R}) \to \mathbb{R}, \quad \varphi \mapsto \sum_{k=1}^{\infty} \Big(\varphi(\frac{1}{k}) - \varphi(0) \Big),$$

定义了 $\mathbb{R}$ 上的分布并确定该分布的支集。

D2) 给定实数序列 $\{\alpha_k\}_{k\geqslant 1} \subset \mathbb{R}$，考虑如下两个线性映射

$$u : \mathcal{D}\big((0,+\infty)\big) \to \mathbb{R}, \quad \varphi \mapsto u(\varphi) = \sum_{k=1}^{\infty} \alpha_k \varphi\left(\frac{1}{k}\right),$$

$$v : \mathcal{D}\big((0,+\infty)\big) \to \mathbb{R}, \quad \varphi \mapsto v(\varphi) = \sum_{k=1}^{\infty} \alpha_k \varphi^{(k)}\left(\frac{1}{k}\right),$$

其中 $\varphi^{(k)}$ 为 φ 的 k-阶导数。证明，u 和 v 是 $(0,+\infty)$ 上的分布。

D3) 证明，如下两个命题等价：

- u 是 $\mathbb{R}$ 上某个分布的限制，即存在 $\widetilde{u} \in \mathcal{D}'(\mathbb{R})$，使得 $u = \widetilde{u}\big|_{(0,\infty)}$；
- 序列 $\{\alpha_k\}_{k\geqslant 1}$ 是多项式增长的，即存在正实数 $C > 0$ 和正整数 ℓ，使得对任意的 $k \geqslant 1$，我们都有

$$|\alpha_k| \leqslant C k^{\ell}.$$

D4) 证明，如下两个命题等价：

- v 是 $\mathbb{R}$ 上某个分布的限制，即存在 $\widetilde{v} \in \mathcal{D}'(\mathbb{R})$，使得 $v = \widetilde{v}\big|_{(0,\infty)}$；
- 序列 $\{\alpha_k\}_{k\geqslant 1}$ 是具有紧支集，即存在正整数 k_0，使得当 $k \geqslant k_0$ 时，有 $\alpha_k = 0$。

D5) 给定 $f \in L^1_{\mathrm{loc}}\big((0,\infty)\big)$，其中 $f \geqslant 0$ 几乎处处成立。证明，如下两个命题等价：

- f 是 $\mathbb{R}$ 上某个分布的限制，即存在 $w \in \mathcal{D}'(\mathbb{R})$，使得 $f = w|_{(0,\infty)}$，亦即对任意 $\varphi \in \mathcal{D}\big((0,\infty)\big)$，我们有

$$\langle w, \varphi \rangle = \int_0^{+\infty} f(x)\varphi(x)dx;$$

- 存在实数 $C > 0$ 和正整数 ℓ，使得对一切 $\varepsilon \in (0,1)$，我们都有

$$\int_{\varepsilon}^{1} f(x)dx \leqslant \frac{C}{\varepsilon^{\ell}}.$$

习题 E. 分布的张量积

假设 $\Omega_1 \subset \mathbb{R}^{n_1}$ 和 $\Omega_2 \subset \mathbb{R}^{n_2}$ 是两个非空的开集。

E1) 对于 $u_1 \in C(\Omega_1)$ 和 $u_2 \in C(\Omega_2)$，我们定义它们的张量积 $u_1 \otimes u_2$ 为 $\Omega_1 \times \Omega_2$ 上的连续函数

$$(u_1 \otimes u_2)(x_1, x_2) = u_1(x_1)u_2(x_2).$$

证明，对任意的 $\varphi(x_1, x_2) \in \mathcal{D}(\Omega_1 \times \Omega_2)$，其中 $x_1 \in \Omega_1$，$x_2 \in \Omega_2$，我们有

$$\langle u_1 \otimes u_2, \varphi\rangle = \big\langle u_1(x_1), \langle u_2(x_2), \varphi(x_1, x_2)\rangle\big\rangle.$$

特别地，（证明）$\varphi_1 \in \mathcal{D}(\Omega_1)$ 和 $\varphi_2 \in \mathcal{D}(\Omega_2)$，那么，$\varphi_1 \otimes \varphi_2 \in \mathcal{D}(\Omega_1 \times \Omega_2)$ 并且

$$\langle u_1 \otimes u_2, \varphi_1 \otimes \varphi_2\rangle = \langle u_1, \varphi_1\rangle\langle u_2, \varphi_2\rangle.$$

E2) 对于 $u_1 \in \mathcal{D}'(\Omega_1)$ 和 $u_2 \in \mathcal{D}'(\Omega_2)$，我们定义 $\mathcal{D}(\Omega_1 \times \Omega_2)$ 上的连续线性泛函：

$$u : \mathcal{D}(\Omega_1 \times \Omega_2) \to \mathbb{C}, \quad \varphi \mapsto \big\langle u_1(x_1), \langle u_2(x_2), \varphi(x_1, x_2)\rangle\big\rangle.$$

证明，$u \in \mathcal{D}'(\Omega_1 \times \Omega_2)$。我们把它记做是 $u_1 \otimes u_2$。

进一步验证，如果 $\varphi_1 \in \mathcal{D}(\Omega_1)$，$\varphi_2 \in \mathcal{D}(\Omega_2)$，那么

$$\langle u_1 \otimes u_2, \varphi_1 \otimes \varphi_2\rangle = \langle u_1, \varphi_1\rangle\langle u_2, \varphi_2\rangle.$$

对于 $a_1 \in \Omega_1$，$a_2 \in \Omega_2$，试计算 $\delta_{a_1} \otimes \delta_{a_2}$。

E3*) 我们定义 $\mathcal{D}(\Omega_1 \times \Omega_2)$ 的子空间：

$$\mathcal{D}(\Omega_1) \otimes \mathcal{D}(\Omega_2) = \mathrm{span}_{\mathbb{C}}\big\{\varphi_1 \otimes \varphi_2 \big| \varphi_1 \in \mathcal{D}(\Omega_1), \varphi_2 \in \mathcal{D}(\Omega_2)\big\}.$$

证明，$\mathcal{D}(\Omega_1) \otimes \mathcal{D}(\Omega_2)$ 在 $\mathcal{D}(\Omega_1 \times \Omega_2)$ 中是稠密的。

（提示：先证明对任意的 $\varphi \in \mathcal{D}(\Omega_1 \times \Omega_2)$，存在关于 $x_1 \in \mathbb{R}^{n_1}, x_2 \in \mathbb{R}^{n_2}$ 的多项式序列 $\{P_k(x_1, x_2)\}_{k\geqslant 1}$，使得对任意的多重指标 α，在 $\mathrm{supp}(\varphi)$ 上，当 $k \to \infty$ 时，$\partial^\alpha P_k$ 一致收敛到 $\partial^\alpha \varphi$）

E4) 对于 $u_1 \in \mathcal{D}'(\Omega_1)$ 和 $u_2 \in \mathcal{D}'(\Omega_2)$，证明，存在唯一的分布 $u \in \mathcal{D}'(\Omega_1 \times \Omega_2)$，使得对任意的 $\varphi_1 \in \mathcal{D}(\Omega_1)$ 和 $\varphi_2 \in \mathcal{D}(\Omega_2)$，我们有

$$\langle u, \varphi_1 \otimes \varphi_2\rangle = \langle u_1, \varphi_1\rangle\langle u_2, \varphi_2\rangle.$$

E5) 我们用 $(x_1, \cdots, x_{n_1})$ 表示 Ω_1 上的坐标，用 $(y_1, \cdots, y_{n_2})$ 表示 Ω_2 上的坐标。对于 $u_1 \in \mathcal{D}'(\Omega_1)$ 和 $u_2 \in \mathcal{D}'(\Omega_2)$ 和任意的 $1 \leqslant k \leqslant n_1$，证明，

$$\partial_{x_k}(u_1 \otimes u_2) \overset{\mathcal{D}'}{=} (\partial_{x_k} u_1) \otimes u_2.$$

E6) 对任意的非负整数 k，考虑 $\mathbb{R}^n$ 上的常系数微分算子

$$P = \partial_1^{k+2}\partial_2^{k+2}\cdots\partial_n^{k+2}.$$

证明，函数

$$E(x_1,\cdots,x_n) = \left(\frac{x_1^{k+1}H(x_1)}{(k+1)!}\right)\otimes\left(\frac{x_2^{k+1}H(x_2)}{(k+1)!}\right)\otimes\cdots\otimes\left(\frac{x_n^{k+1}H(x_n)}{(k+1)!}\right),$$

是 C^k 的并且是 P 的基本解，其中 $H(x)$ 是 Heavisde 函数。

A great deal of my work is just playing with equations and seeing what they give.

— Paul Dirac

67　缓增分布的 Fourier 变换：定义与基本例子的计算

二零二零年十月十九日，星期一，晴

我们先补充一个关于 Schwartz 函数的命题：任意的多重指标 α, β，我们有连续映射

$$x^\alpha : \mathcal{S}(\mathbb{R}^n) \to \mathcal{S}(\mathbb{R}^n),$$
$$\partial^\beta : \mathcal{S}(\mathbb{R}^n) \to \mathcal{S}(\mathbb{R}^n).$$

其中，连续性指的是收敛的函数序列的像仍然是收敛的函数序列。

我们对求导数来证明这个命题：对任意的多重指标 α' 和 β'，我们有

$$\left| x^{\beta'} \partial^{\alpha'} \partial^\alpha \varphi(x) \right| \leqslant N_{|\alpha|+|\alpha'|+|\beta'|}(\varphi).$$

所以，$x^\alpha \varphi \in \mathcal{S}(\mathbb{R}^n)$。

为了说明连续性，我们任意选取 $\varphi_k \xrightarrow{\mathcal{S}} \varphi$，那么，对任意的 $p \geqslant 0$，我们有

$$N_p(\partial^\alpha \varphi_k - \partial^\alpha \varphi) \leqslant N_{p+|\alpha|}(\varphi_k - \varphi) \to 0.$$

所以，$\partial^\alpha \varphi_k \xrightarrow{\mathcal{S}} \partial^\alpha \varphi$。

关于乘以 x^β 的验证类似，我们还可以利用 Fourier 变换（保持函数列的收敛）把乘法化成求倒数即可。

我们上次定义了缓增的分布。所谓的缓增的分布 $u \in \mathcal{S}'(\mathbb{R}^n)$，就是满足如下条件的分布：存在非负整数 p 和常数 $C > 0$，使得对每个 $\varphi \in \mathcal{D}(\mathbb{R}^n)$，我们都有

$$|\langle u, \varphi \rangle| \leqslant C N_p(\varphi),$$

我们上次证明了，给定 $u \in \mathcal{S}'(\mathbb{R}^n)$，对任意的 $\varphi \in \mathcal{S}(\mathbb{R}^n)$，我们都可以定义 $T_u(\varphi) = \langle u, \varphi \rangle$：

$$T_u(\varphi) = \langle u, \varphi \rangle = \lim_{k \to \infty} \langle u, \varphi_k \rangle_{\mathcal{D}' \times \mathcal{D}},$$

其中，$\{\varphi_k\}_{k \geqslant 1} \subset \mathcal{D}(\mathbb{R}^n)$ 是 φ 在 $\mathcal{S}(\mathbb{R}^n)$ 中的逼近序列。有时候，为了行文清楚，我们还把上面的配对记作

$$T_u(\varphi) = \langle u, \varphi \rangle_{\mathcal{S}' \times \mathcal{S}}.$$

给定 $u \in \mathcal{S}'(\mathbb{R}^n)$ 和多重指标 α，β，作为 $\mathcal{D}'(\mathbb{R}^n)$ 中的元素，我们自然有 $\partial^\alpha u, x^\beta u \in \mathcal{D}'(\mathbb{R}^n)$。我们现在要说明 $\partial^\alpha u, x^\beta u \in \mathcal{S}'(\mathbb{R}^n)$。实际上，对任意的 φ，按照缓增分布的定义，我们有

$$|\langle \partial^\alpha u, \varphi \rangle| = |\langle u, \partial^\alpha \varphi \rangle| \leqslant C N_{p+|\alpha|}(\varphi).$$

这表明 $\partial^\alpha u \in \mathcal{S}'(\mathbb{R}^n)$。类似地，我们有 $x^\beta u \in \mathcal{S}'(\mathbb{R}^n)$。

另外，对任意的 $\varphi \in \mathcal{S}'(\mathbb{R}^n)$，我们选取 $\{\varphi_k\}_{k\geqslant 1} \subset \mathcal{D}(\mathbb{R}^n)$ 为 φ 在 $\mathcal{S}(\mathbb{R}^n)$ 中的逼近序列，那么，

$$\begin{aligned}\langle \partial^\alpha u, \varphi\rangle_{\mathcal{S}'\times\mathcal{S}} &= (-1)^{|\alpha|}\lim_{k\to\infty}\langle u, \partial^\alpha \varphi_k\rangle \\ &= (-1)^{|\alpha|}\langle \partial^\alpha u, \partial^{|\alpha|}\varphi\rangle_{\mathcal{S}'\times\mathcal{S}}.\end{aligned}$$

所以，我们仍然有类似于分布情形的公式：

$$\langle \partial^\alpha u, \varphi\rangle_{\mathcal{S}'\times\mathcal{S}} = (-1)^{|\alpha|}\langle u, \partial^\alpha \varphi\rangle_{\mathcal{S}'\times\mathcal{S}}.$$

类似地，我们还有

$$\langle x^\beta u, \varphi\rangle_{\mathcal{S}'\times\mathcal{S}} = \langle u, x^\beta \varphi\rangle_{\mathcal{S}'\times\mathcal{S}}.$$

但是，我们注意到，如果 f 仅仅是光滑函数，那么

$$\langle f\varphi\rangle_{\mathcal{S}'\times\mathcal{S}} = \langle u, f\varphi\rangle_{\mathcal{S}'\times\mathcal{S}}.$$

并不成立，因为通常而言 $f\varphi \notin \mathcal{S}(\mathbb{R}^n)$。

我们现在规定 $\mathcal{S}'(\mathbb{R}^n)$ 中序列的收敛性：

定义 461 (收敛性)**.** 给定缓增分布的序列 $\{u_k\}_{k\geqslant 1} \subset \mathcal{S}'(\Omega)$，我们说它在 $\mathcal{S}'(\mathbb{R}^n)$ 的意义下收敛到 $u \in \mathcal{S}'(\Omega)$，记作 $u_k \xrightarrow{\mathcal{S}'} u$，指的是对每个 *Schwartz* 函数 $\varphi \in \mathcal{S}(\mathbb{R}^n)$，都有

$$\lim_{k\to\infty}\langle u_k, \varphi\rangle = \langle u, \varphi\rangle.$$

命题 462. 对任意的多重指标 α 和 β，我们有如下的连续映射

$$\begin{aligned}x^\alpha &: \mathcal{S}'(\mathbb{R}^n) \to \mathcal{S}'(\mathbb{R}^n), \\ \partial^\beta &: \mathcal{S}'(\mathbb{R}^n) \to \mathcal{S}'(\mathbb{R}^n).\end{aligned}$$

换而言之，我们有

1) 如果 $u \in \mathcal{S}'(\mathbb{R}^n)$，那么 $\partial^\alpha u, x^\beta u \in \mathcal{S}'(\mathbb{R}^n)$。

2) 如果我们有缓增分布的收敛序列 $u_k \xrightarrow{\mathcal{S}'} u$，那么，它在求导数和乘多项式下被保持：

$$\partial^\alpha u_k \xrightarrow{\mathcal{S}'} \partial^\alpha u, \quad x^\beta u_k \xrightarrow{\mathcal{S}'} x^\beta u.$$

证明: 证明是简单（乏味）的：我们已经说明了 $\partial^\alpha u, x^\beta u \in \mathcal{S}'(\mathbb{R}^n)$。为了说明连续性，我们有

$$\langle \partial^\alpha u_k, \varphi\rangle = \lim_{k\to\infty}\langle u_k, (-1)^{|\alpha|}\varphi\rangle = \langle u, (-1)^{|\alpha|}\varphi\rangle = \langle \partial^\alpha u, \varphi\rangle.$$

关于乘法的证明是类似的。 □

我们现在看一些缓增的分布的例子：

1) 假设 $p=1,2$ 或 ∞，我们可以将函数空间 $L^p(\mathbb{R}^n)$ 视为分布（因为这都是局部可积的），那么它们是缓增的分布：

假设 $f\in L^\infty(\mathbb{R}^n)$，那么，对任意的 $\varphi\in\mathcal{D}(\mathbb{R}^n)$，我们有

$$\begin{aligned}|\langle f,\varphi\rangle| &= \int_{\mathbb{R}^n}\frac{f}{(1+|x|^2)^{\frac{n+1}{2}}}\cdot\left((1+|x|^2)^{\frac{n+1}{2}}\cdot\varphi\right)dx\\ &\leqslant\left(\int_{\mathbb{R}^n}\frac{1}{(1+|x|^2)^{\frac{n+1}{2}}}dx\right)\|f\|_{L^\infty}\cdot N_{n+1}(\varphi).\end{aligned}$$

我们把 $p=1$ 或 2 的情况留作作业。

2) 有紧支集的分布 $\mathcal{E}'(\mathbb{R}^n)$ 是缓增的分布：这是因为每个紧支集的分布都有限阶的分布。

3) 分布 $\mathrm{pv}\frac{1}{x}$ 是缓增的分布：实际上，我们可以把它写成

$$\mathrm{pv}\frac{1}{x}=\underbrace{\mathbf{1}_{x|<1}\cdot\mathrm{pv}\frac{1}{x}}_{\in\mathcal{E}'(\mathbb{R}^n)}+\underbrace{\mathbf{1}_{|x|\geqslant1}\mathrm{pv}\cdot\frac{1}{x}}_{\in L^\infty}.$$

4) 局部可积函数并且具有多项式增长速度的函数是缓增的分布，即对于 $f\in L^1_{\mathrm{loc}}(\mathbb{R}^n)$，如果存在常数 C 和 m，使得对任意的 $x\in\mathbb{R}^n$，我们都有

$$|f(x)|\leq C(1+|x|)^m,$$

那么，$f\in\mathcal{S}'(\mathbb{R}^n)$。

实际上，对任意的试验函数 $\varphi\in\mathcal{D}'(\mathbb{R}^n)$，我们有

$$\begin{aligned}|\langle f,\varphi\rangle| &= \left|\int_{\mathbb{R}^n}f(x)\varphi(x)dx\right|\\ &\leqslant\int_{\mathbb{R}^n}\underbrace{\frac{|f(x)|}{(1+|x|)^{m+n+1}}}_{\text{可积}}\cdot\underbrace{(1+|x|)^{m+n+1}|\varphi(x)|}_{\leqslant C'N_{m+n+1}(\varphi)}dx\\ &\leqslant C''N_{m+n+1}(\varphi).\end{aligned}$$

5) 函数 e^x 所定义的分布不是缓增的分布。

我们选取 $\chi\in C_0^\infty(\mathbb{R}^n)$，使得 $\chi\big|_{[0,1]}\equiv1$ 并且 $\chi\geqslant0$。令 $\chi_n=\chi(x-n)$。如果 e^x 是缓增的分布，那么，存在 p，使得

$$\int_{\mathbb{R}}e^x\chi_n(x)dx\leqslant C\sup_{|\alpha|,|\beta|\leqslant p}\|x^\alpha\partial^\beta\varphi_n\|_{L^\infty}\leqslant Cn^p.$$

然而，

$$\int_{\mathbb{R}}e^x\chi_n(x)dx=\int_{\mathbb{R}}e^x\chi(x-n)dx\geqslant\int_{[n,n+1]}e^xdx\geqslant e^n.$$

令 $n\to\infty$，我们就得到了矛盾。

6) 指数增长速度的函数也可以是缓增的分布。

我们令 $u = e^x e^{ie^x}$，这个函数是在 $\mathbb{R}$ 上是指数增长的，但是它所定义的分布是缓增的：因为 $u = \left(e^{ie^x}\right)'$，而 $e^{ie^x} \in L^\infty(\mathbb{R})$ 所定义的分布是缓增的。

定义 463 (缓增分布的 Fourier 变换)**.** 对任意的 $u \in \mathcal{S}'(\mathbb{R}^n)$，我们用以下的公式来定义其 *Fourier* 变换（记作 $\mathcal{F}u$ 或者 $\widehat{u}$）：对任意的 $\varphi \in \mathcal{S}(\mathbb{R}^n)$，令

$$\langle \widehat{u}, \varphi \rangle = \langle u, \widehat{\varphi} \rangle.$$

我们可以类似地定义 *Fourier* 逆变换 $\mathcal{F}^{-1}$。实际上，我们只要定义

$$\langle \mathcal{F}^{-1}(u), \varphi \rangle = \langle u, \mathcal{F}^{-1}(\varphi) \rangle.$$

即可。

在陈述下面的定理之前，我们先回忆在 Schwartz 函数空间上的 Fourier 变换（以及 Fourier 逆变换）：

$$\mathcal{F} : \mathcal{S}(\mathbb{R}^n) \longrightarrow \mathcal{S}(\mathbb{R}^n), \quad \varphi \mapsto \widehat{\varphi}(\xi),$$

是连续的同构，并且满足如下的性质：对任意的 $p \in \mathbb{Z}_{\geqslant 0}$，存在常数 $C_p > 0$，使得对每个 $\varphi \in \mathcal{S}(\mathbb{R}^n)$，我们都有

$$N_p(\widehat{\varphi}) \leqslant C_p N_{p+n+1}(\varphi).$$

据此，我们来说明如果 $u \in \mathcal{S}'(\mathbb{R}^n)$，那么上述所定义的 $\widehat{u}$ 也是缓增的分布：这是因为对任意速降的函数 $\varphi \in \mathcal{S}(\mathbb{R}^n)$，我们有如下的不等式

$$|\langle \widehat{u}, \varphi \rangle| = |\langle u, \widehat{\varphi} \rangle| \leqslant C N_p(\widehat{\varphi}) \leqslant C C_p N_{p+n+1}(\varphi).$$

这就验证了 $\widehat{u} \in \mathcal{S}'(\mathbb{R}^n)$。

定理 464. *Fourier* 变换

$$\mathcal{F} : \mathcal{S}'(\mathbb{R}^n) \longrightarrow \mathcal{S}'(\mathbb{R}^n),$$

是连续线性同构，这里连续性指的是对任意的在 $\mathcal{S}'(\mathbb{R}^n)$ 中收敛的缓增分布的序列 $u_k \xrightarrow{\mathcal{S}'} u$，我们都有 $\widehat{u_k} \xrightarrow{\mathcal{S}'} \widehat{u}$。进一步，对每个 $u \in \mathcal{S}'(\mathbb{R}^n)$，如下的公式成立：

$$\widehat{\partial_k u} = i\xi_k \widehat{u}, \quad \widehat{x_k u} = i\partial_k \widehat{u}, \quad \mathcal{F}^{-1}(u) = \frac{1}{(2\pi)^n} \mathcal{F}(u)\check{}.$$

证明： 先验证连续性：假设 $u_k \xrightarrow{\mathcal{S}'} u$，那么，对任意的 Schwartz 函数 φ，我们都有

$$\langle \widehat{u_k}, \varphi \rangle = \langle u_k, \widehat{\varphi} \rangle \to \langle u, \widehat{\varphi} \rangle = \langle \widehat{u}, \varphi \rangle.$$

所以，$\widehat{u_k} \xrightarrow{\mathcal{S}'} \widehat{u}$。另外，由于

$$\langle \mathcal{F}^{-1}\left(\mathcal{F}(u)\right), \varphi \rangle = \langle u, \mathcal{F}^{-1}\left(\mathcal{F}(\varphi)\right) \rangle = \langle u, \varphi \rangle.$$

所以，$\mathcal{F}$ 与 $\mathcal{F}^{-1}$ 互为逆，从而，$\mathcal{F}$ 是缓增分布上的同构。

定理中的三个公式的验证也是平凡的，因为我们总是可以在 $\mathcal{S}'(\mathbb{R}^n) \times \mathcal{S}(\mathbb{R}^n)$ 的配对时，把 $\mathcal{S}'(\mathbb{R}^n)$ 中的运算全部挪到 $\mathcal{S}\mathbb{R}^n)$ 上来验证。对于 $\mathcal{S}(\mathbb{R}^n)$ 中的 φ，这三个公式我们已经证明过了。 □

Fourier 变换的几个例子

我们计算一些常见函数的 Fourier 变换：

例子. *1) Dirac 函数 δ_0。我们有*

$$\widehat{\delta_0} = 1.$$

假设 $\varphi(x) \in \mathcal{S}(\mathbb{R}^n)$，那么，

$$\langle \widehat{\delta_0}, \varphi \rangle = \langle \delta_0, \widehat{\varphi} \rangle = \int_{\mathbb{R}^n} \varphi(x) dx = \langle 1, \varphi \rangle.$$

这就完成了计算

2) 对任意的 $a \in \mathbb{R}^n$，任意的多重指标 α，我们有

$$\widehat{\partial^\alpha \delta_a} = (i\xi)^\alpha e^{-ix\cdot\xi}, \quad \widehat{x^\alpha} = (2\pi)^n (i\xi)^\alpha \delta_0.$$

特别地，

$$\widehat{1} = (2\pi)^n \delta_0.$$

这个计算留作作业。

作为应用，我们研究 $\mathbb{R}^n$ 上的调和的缓增分布，即 $u \in \mathcal{S}'(\mathbb{R}^n)$ 并且满足

$$\triangle u = 0.$$

命题 465. 假设 $u \in \mathcal{S}'(\mathbb{R}^n)$ 是调和的缓增分布，那么，u 必然是（$x_1, \cdots, x_n$ 的）多项式函数。

证明: 由于 $u \in \mathcal{S}'(\mathbb{R}^n)$，所以我们可以对它做 Fourier 变换。从而，

$$\triangle u = 0 \quad \Rightarrow \quad -|\xi|^2 \widehat{u} = 0.$$

这表明（练习）$\mathrm{supp}(\widehat{u}) \subset \{0\}$，从而，

$$\widehat{u} = \sum_{|\alpha| \leqslant m} c_\alpha \partial^\alpha \delta_0(\xi).$$

对上式作 Fourier 逆变换：

$$u = \sum_{|\alpha| \leqslant m} c_\alpha \left(\frac{\xi}{i}\right)^\alpha.$$

这就得到了要证明的结论。 □

注记. 由于众所周知的原因，我们通常把 $\mathrm{supp}(\widehat{u})$ 称作是 u 的**谱**并记作 $\mathrm{spec}(u)$。

注记. 要求 u 是缓增的分布是非常重要条件：如果不对 u 在无穷远处的增长加以限制，那么调和的分布是很多的。比如说，在 $\mathbb{R}^2 = \mathbb{C}$ 上，任何一个在整个 $\mathbb{C}$ 上定义的复解析函数都是调和的。比如说，e^z，它就不是多项式函数，它在 ∞ 附近增长的很快，所以它所定义的分布不是缓增的。

为了计算一些特殊的分布的 Fourier 变换，我们经常用到如下两个技巧：

命题 466. 假设 $u \in \mathcal{S}'(\mathbb{R}^n)$ 是缓增的分布。我们有：

1) 给定可逆的 $n \times n$ 的实系数矩阵 A，我们把它看作是 $\mathbb{R}^n$ 到自身的线性变换，那么，A^*u 也是缓增的分布（证明几乎是显然的）并且

$$\widehat{A^*u} = |\det(A)|^{-1} \left({}^tA^{-1}\right)^* \widehat{u}.$$

通常，我们把这个公式写作

$$\widehat{u(Ax)}(\xi) = |\det(A)|^{-1} \widehat{u}\left({}^tA^{-1}\xi\right).$$

2) 如果 u 是次数为 s 的齐次分布，那么，$\widehat{u}$ 是次数为 $-s-n$ 的齐次分布。

3) 如果 u 是奇分布，即满足 $\check{u} = u$，那么 $\widehat{u}$ 也是；如果 u 是偶分布，即满足 $\check{u} = -u$，那么 $\widehat{u}$ 也是。

4) 如果 u 是旋转对称的，即对任意的 $A \in \mathbf{SO}(n)$，

$$A^*u = u,$$

那么，$\widehat{u}$ 是旋转对称的，

证明*:* 我们把 1) 的证明留作作业。为了证明 2)，我们用 1) 的结论。令 A_λ 为对角线上均为 λ 的对角矩阵，其中 $\lambda > 0$。那么（请参考第一次作业），u 是次数为 s 的齐次分布等价于对任意的 $\lambda > 0$，我们都有

$$(A_\lambda)^* u = \lambda^s u.$$

对上面的等式作 Fourier 变换，我们就得到

$$\lambda^{-n} (A_{\lambda^{-1}})^* \widehat{u} = \lambda^s \widehat{u}.$$

令 $\gamma = \lambda^{-1}$，所以，

$$(A_\gamma)^* \widehat{u} = \gamma^{-s-n} \widehat{u}.$$

所以，$\widehat{u}$ 是次数为 $-s-n$ 的齐次分布。

为了证明 3)，我们只需要把 2) 中的 λ 选为 -1 即可。

为了证明 4)，我们注意到对于任意的正交矩阵，我们都有

$${}^tA^{-1} = A.$$

所以，公式

$$\widehat{u(Ax)}(\xi) = |\det(A)|^{-1} \widehat{u}\left({}^tA^{-1}\xi\right)$$

可以写成

$$\widehat{u(Ax)}(\xi) = \widehat{u}\,(A\xi).$$

这就是要验证的。 □

例子. *Dirac* 函数 δ_0 是次数为 $-n$ 的齐次分布，因为它的 *Fourier* 变换 1 是次数为 0 的齐次分布。

例子. *3)* 我们计算 $\mathrm{vp}\frac{1}{x}$ 的 *Fourier* 变换：

$$\widehat{\mathrm{vp}\frac{1}{x}} = -2\pi i H(x) + \pi i,$$

我们注意到

$$x \cdot \mathrm{vp}\frac{1}{x} = 1.$$

所以，

$$\widehat{x \cdot \mathrm{vp}\frac{1}{x}} = 2\pi\delta_0 \quad \Rightarrow \quad \frac{d}{d\xi}\left(\widehat{\mathrm{vp}\frac{1}{x}}\right) = -i2\pi\delta_0.$$

我们知道 *Heaviside* 函数 $H(x)$ 是上述方程的一个解，所以，

$$\widehat{\mathrm{vp}\frac{1}{x}} = -2\pi i H(x) + C,$$

其中，C 是待定的常数。

我们注意到 $\mathrm{vp}\frac{1}{x}$ 是一个奇分布，所以，$\widehat{\mathrm{vp}\frac{1}{x}}$ 也是，这说明 $C = \pi i$。

4) *Heaviside* 函数 $H(x)$ 的 *Fourier* 变换为

$$\widehat{H}(\xi) = -i \cdot \mathrm{vp}\frac{1}{x} + \pi\delta_0.$$

这个计算留作作业。

5) 我们考虑 $\mathbb{R}^n$ 上 *Laplace* 算子 $\triangle$ 的一个基本解 u，其中 $n \geqslant 3$，

$$\triangle u = \delta_0.$$

在第二次作业中，我们给出了 $\triangle$ 的基本解，但是，这种方式不是令人信服的：因为我们想知道如何能够合理地动手来构造一个基本解而不是仅仅验证某些分布是基本解。

我们做如下的预设[20]：u 是缓增的分布。那么，利用 *Fourier* 变换，在频率空间中，我们就有

$$-|\xi|^2\widehat{u} = 1 \quad \Rightarrow \quad \widehat{u}(\xi) = -\frac{1}{|\xi|^2}.$$

我们注意到 $n \geqslant 3$，所以 $|\xi|^{-2}$ 是局部可积的，从而定义了一个分布。另外，我们可以把它写成

$$\frac{1}{|\xi|^2} = \frac{1}{|\xi|^2}\mathbf{1}_{|\xi|\leqslant 1}(\xi) + \frac{1}{|\xi|^2}\mathbf{1}_{|\xi|>1}(\xi).$$

[20]在找到问题的解答之前，我们总是可以做各种（相对合理的）假设。这些额外的假设可能给出问题的一类解。通过做这样的假设得到的解有时候恰好是问题的所有解，也有可能不是所有的解，但是总是比没有找到解更令人欣慰。在英文的文献中，这种假设叫做 ansatz。在分析问题中，所谓的分离变量法就是这样的一种方法。我们将会看到，利用另一种预设，我们也可以得到 $\triangle$ 的基本解。

这是一个有紧支集的分布与 L^∞ 函数的和，从而是缓增的分布。

上面给出了 $\widehat{u}$ 的一个可能的解（其他的解与 u 相差一个复系数多项式)。为了在物理空间中表达 u，我们要计算 $-\dfrac{1}{|\xi|^2}$ 的 *Fourier* 逆变换。我们注意到这是一个旋转对称的次数为 -2 的齐次分布，从而，它的 *Fourier* 逆变换是旋转对称的次数为 $-n+2$ 的齐次分布。通过这个观察，一个合理的猜测自然是

$$u = \frac{C}{|x|^{n-2}}.$$

这样子，我们就回到了第二次作业中的形式，通过在原点处计算，我们可以将待定的系数 C 算出来。

在那次作业中，我们已经证明了当 $n \geqslant 3$ 时，我们用。令

$$E = \frac{|x|^{2-n}}{(2-n)\left|\mathbf{S}^{n-1}\right|}.$$

是 $\triangle$ 的基本解，其中 $\left|\mathbf{S}^{n-1}\right|$ 表示 $\mathbf{S}^{n-1}$ 的测度。由于 E 在 0 附近可积并且在 ∞ 处衰减，所以这是一个缓增的分布，据此，我们知道

$$\widehat{E}(\xi) = -\frac{1}{|\xi|^2}.$$

从而，

$$\mathcal{F}^{-1}\left(-\frac{1}{|\xi|^2}\right) = E(x).$$

所以，

$$\mathcal{F}\left(-\frac{1}{|\xi|^2}\right) = \left((2\pi)^n \mathcal{F}^{-1}\left(-\frac{1}{|\xi|^2}\right)\right)\check{} = (2\pi)^n E(x),$$

即

$$\widehat{|x|^{-2}}(\xi) = \frac{(2\pi)^n}{(n-2)\left|\mathbf{S}^{n-1}\right|}|\xi|^{2-n}.$$

68 缓增分布的 Fourier 变换与卷积

二零二零年十月二十一日，星期三，晴

我们是首先研究球面测度的 Fourier 变换，这是一个重要计算，我们在课程的后面会看到它在解波动方程时的应用，这个公式在调和分析中也是重要例子（限制定理）：

例子. 令 $d\sigma_R$ 为 $\mathbb{R}^3$ 上中心在原点半径为 R 的球面 $\mathbf{S}_R^2$ 上的球面测度，其中 $R>0$。我们（已经见过）可以把它视作是一个 0 阶的分布：对任意的 $\varphi\in\mathcal{D}'(\mathbb{R}^3)$，我们令

$$\langle d\sigma_R,\varphi\rangle=\int_{\mathbf{S}_R^2}\varphi\big|_{\mathbf{S}_R^2}d\sigma_R.$$

这是一个具有紧支集的分布，我们要计算它的 *Fourier* 变换 $\widehat{d\sigma_R}(\xi)$。这是一个关于 ξ 的光滑函数。由于分布 $\langle d\sigma_R$ 是旋转不变的（作业），所以 $\widehat{d\sigma_R}(\xi)$ 是也是旋转对称的。特别地，$\widehat{d\sigma_R}(\xi)$ 是只与 $|\xi|$ 相关的函数。

根据上面的讨论，我们只要做如下计算即可：

$$\begin{aligned}\widehat{d\sigma_R}(0,0,|\xi|)&=\langle d\sigma_R,e^{i(x,y,z)\cdot(0,0,|\xi|)}\rangle=\int_0^{2\pi}\left(\int_0^{\pi}e^{-i|\xi|R\cos\vartheta}R^2\sin\vartheta d\vartheta\right)d\phi\\&=2R^2\pi\int_0^{\pi}e^{-i|\xi|R\cos\vartheta}\sin\vartheta d\vartheta=2R^2\frac{e^{-iR|\xi|}t}{-iR|\xi|}\Big|_{-1}^{1}\end{aligned}$$

最终，我们得到

$$\frac{\widehat{d\sigma_R}}{4\pi R}=\frac{\sin(R|\xi|)}{|\xi|}.$$

在进一步研究缓增分布的性质之前，我们先说明对于一个 L^1 或者 L^2 的函数，我们把它 Fourier 变换视作是一个缓增的分布，也可以先将它视作是缓增的分布再做 Fourier 变换来得到一个缓增的分布，这两种方式是一致的：

命题 467. 假定 $u\in L^1(\mathbb{R}^n)$（或者 $L^2(\mathbb{R}^n)$），那么，它作为缓增分布的 *Fourier* 变换与它作为 L^1（或 L^2）函数的 *Fourier* 变换是一致的。

证明: 先假设 $f\in L^1(\mathbb{R}^n)$，我们把它在 L^1 意义下的 Fourier 变换（可用积分来写）记为 $\mathcal{F}_1(f)$。由于 $\mathcal{F}_1(f)\in C_\circ(\mathbb{R}^n)$，所以，对任意的 Schwartz 函数 φ，我们有（容易验证 Fubini 定理的条件总是满足的）

$$\begin{aligned}\langle\mathcal{F}_1(f),\varphi\rangle&=\int_{\mathbb{R}^n}\mathcal{F}_1(f)(\xi)\varphi(\xi)d\xi=\int_{\mathbb{R}^n}\int_{\mathbb{R}^n}e^{-ix\cdot\xi}f(x)\varphi(\xi)dxd\xi\\&=\int_{\mathbb{R}^n}f(x)\left(\int_{\mathbb{R}^n}e^{-ix\cdot\xi}\varphi(\xi)d\xi\right)dx=\int_{\mathbb{R}^n}f(x)\widehat{\varphi}(x)dx=\langle f,\widehat{\varphi}\rangle\\&=\langle\widehat{f},\varphi\rangle.\end{aligned}$$

其中，最后的等号用的是缓增分布中的 Fourier 变换。根据局部可积函数到分布嵌入的唯一性，我们知道 $\mathcal{F}_1(f)=\widehat{f}$。

现在假设 $f \in L^2(\mathbb{R}^n)$，我们把它在 L^2-意义下的 Fourier 变换记为 $\mathcal{F}_2(f)$。对任意的 $k \geqslant 1$，我们令

$$f_n = f \cdot \mathbf{1}_{|x| \leqslant n} \in L^1(\mathbb{R}^n) \cap L^2(\mathbb{R}^n),$$

那么，我们已经证明过

$$\mathcal{F}_2(f) \overset{L^2}{=} \lim_{k \to \infty} \mathcal{F}_1(f_k),$$

其中，极限是在 L^2 的收敛的意义下取的。所以，作为分布，我们也有（利用 Cauchy-Schwartz，请参考作业题）

$$\mathcal{F}_2(f) \overset{\mathcal{D}'}{=} \lim_{k \to \infty} \mathcal{F}_1(f_k),$$

从而，根据上述，我们知道

$$\mathcal{F}_2(f) \overset{\mathcal{D}'}{=} \lim_{k \to \infty} \widehat{f_k},$$

其中，后面戴帽子的符号代表 $\mathcal{S}'$ 中的 Fourier 变换。由于

$$f \overset{L^1}{=} \lim_{k \to \infty} f_k,$$

所以，

$$f \overset{\mathcal{S}'}{=} \lim_{k \to \infty} f_k,$$

从而，根据连续性

$$\widehat{f_k} \xrightarrow{\mathcal{D}'} \widehat{f}.$$

所以，

$$\mathcal{F}_2(f) \overset{\mathcal{D}'}{=} \lim_{k \to \infty} \widehat{f_k} \overset{\mathcal{D}'}{=} \widehat{f}.$$

这就验证了这些 Fourier 变换的概念都是相容的。 □

我们强调过，Fourier 分析中的一个直观是物理空间的衰减意味着频率空间的光滑性，这对分布也是成立的：

定理 468. *假设 $u \in \mathcal{E}'(\mathbb{R}^n)$（自然在无穷远处衰减地足够快），那么，$\widehat{u} \in C^\infty(\mathbb{R}^n_\xi)$ 是光滑函数并且*

$$\widehat{u}(\xi) = \left\langle u, e^{-ix\cdot\xi} \right\rangle.$$

证明： 我们选取截断函数 $\chi \in C_0^\infty(\mathbb{R}^n)$ 来帮助我们记住 u 具有紧支集，其中 $\chi|_{\mathrm{supp}(u)} \equiv 1$。按照定义以及分布与积分可交换的命题，我们有

$$\begin{aligned}
\langle \widehat{u}, \varphi \rangle &= \langle u, \widehat{\varphi} \rangle = \langle \chi \cdot u, \widehat{\varphi} \rangle = \langle u, \chi\widehat{\varphi} \rangle \\
&= \left\langle u, \int_{\mathbb{R}^n} e^{-ix\cdot\xi} \chi(\xi)\varphi(x)dx \right\rangle \\
&= \int_{\mathbb{R}^n} \left\langle u(\xi), e^{-ix\cdot\xi}\chi(\xi) \right\rangle \varphi(x)dx \\
&= \int_{\mathbb{R}^n} \left\langle u(x), e^{-ix\cdot\xi}\chi(x) \right\rangle \varphi(\xi)d\xi \\
&= \int_{\mathbb{R}^n} \left\langle u(x), e^{-ix\cdot\xi} \right\rangle \varphi(\xi)d\xi.
\end{aligned}$$

这就证明了 $\widehat{u}(\xi)=\left\langle u, e^{-ix\cdot\xi}\right\rangle$。

我们将 ξ 视为参数，那么，

$$\widehat{u}(\xi)=\left\langle u, \chi(x)e^{-ix\cdot\xi}\right\rangle.$$

所以，函数 $\widehat{u}(\xi)$ 光滑性可以利用对分布与求导数可交换的命题立即得到。 □

注记. 特别的，上面的命题表明，如果将 ξ 视作是复变量，那么，$\widehat{u}(\xi)$ 是复解析函数。特别的，它的支集不可能是紧集。

我们再来研究 $\mathcal{S}'(\mathbb{R}^n)$ 上的卷积运算。为此，我们先考虑在 $\mathcal{S}(\mathbb{R}^n)$ 上的卷积。

命题 469. 任意给定有紧支集的分布 $c\in\mathcal{E}'(\mathbb{R}^n)$。那么，

1) 对任意的 *Schwartz* 函数 $\varphi\in\mathcal{S}(\mathbb{R}^n)$，我们有

$$\varphi * c\in\mathcal{S}(\mathbb{R}^n).$$

进一步，我们能找到正整数 q（可能依赖于 c），使得对于任何非负整数 p，都存在正常数 C_p，使得对任意的 $\varphi\in\mathcal{S}(\mathbb{R}^n)R$，我们有

$$N_p(\varphi * c)\leqslant C_pN_{p+q}(\varphi).$$

2) 对每个 *Schwartz* 函数 $\varphi\in\mathcal{S}(\mathbb{R}^n)$，我们有

$$\widehat{f * c}(\xi)=\widehat{f}(\xi)\widehat{c}(\xi),\quad \forall\xi\in\mathbb{R}^n.$$

注记. 在本次课的最后一个定理的证明过程中，我们会证明 $\widehat{c}(\xi)$ 是一个多项式增长的函数。

证明: 我们已经证明过，如果 f 是光滑函数，那么 $f*c$ 是光滑函数并且可以表达为

$$(f*c)(x)=\langle c, f(x-\cdot)\rangle.$$

所以，对于 Schwartz 函数 φ，我们有

$$(\varphi*c)(x)=\langle c(y),\varphi(y-x)\rangle.$$

从而，利用分布与求导数可交换的性质，对任意的多重指标 α 和 β，我们就有

$$(x^\alpha\partial^\beta\varphi * c)(x)=\left\langle c(y), x^\alpha\big(\partial^\beta\varphi\big)(y-x)\right\rangle$$

由于 c 是有紧支集的分布，我们令 $K=\mathrm{supp}(u))$；u 是有限阶的的分布，它的阶记作 q。以下我们认为 x 是固定的。从而，

$$\begin{aligned}|(x^\alpha\partial^\beta\varphi * c)(x)|&\leqslant C|x|^{|\alpha|}\sup_{|\gamma|\leqslant q}\left|\partial_y^\gamma\big(\partial^\beta\varphi\big)(y-x)\right|\\&\leqslant C\big(|x-y|+|y|\big)^{|\alpha|}\sup_{|\gamma|\leqslant q}\left|\partial_y^\gamma\big(\partial^\beta\varphi\big)(y-x)\right|.\end{aligned}$$

在上面的不等式中 $y \in K$，所以 $|y|$ 的因子只能贡献一个常数，只贡献一个常数 M。所以，

$$|(x^\alpha \partial^\beta \varphi * c)(x)| \leqslant C\big(|x-y|+M\big)^{|\alpha|} \sup_{|\gamma|\leqslant q} \Big|\partial_y^\gamma \big(\partial^\beta \varphi\big)(y-x)\Big|.$$

通过对 x 取 sup，从而，

$$\sup_{x\in\mathbb{R}^n} |(x^\alpha \partial^\beta \varphi * c)(x)| \leqslant C' N_{|\alpha|+|\beta|+q}(\varphi).$$

从而，1) 中的不等式成立。

现在证明 2)。我们选取截断函数 $\chi \in C_0^\infty(\mathbb{R}^n)$，使得 $\chi\big|_{\mathrm{supp}(c)} \equiv 1$。我们先假定 $\varphi \in \mathcal{D}(\mathbb{R}^n)$（从而，如下进行的分布与积分号可以交换）。我们有

$$\begin{aligned}
\widehat{\varphi * c}(\xi) &= \int_{\mathbb{R}^n} e^{-ix\cdot\xi} \varphi * c(x) dx = \int_{\mathbb{R}^n} e^{-ix\cdot\xi} \left\langle c(y), \varphi(y-x)\right\rangle dx \\
&= \left\langle c(y), \int_{\mathbb{R}^n} e^{-ix\cdot\xi} \varphi(y-x) dx \right\rangle \\
&= \left\langle c(y), \chi(y) \int_{\mathbb{R}^n} e^{-ix\cdot\xi} \varphi(y-x) dx \right\rangle \\
&= \left\langle c(y), \chi(y) e^{-iy\cdot\xi} \widehat{\varphi}(\xi) \right\rangle \\
&= \left\langle c(y), \chi(y) e^{-iy\cdot\xi} \right\rangle \widehat{\varphi}(\xi) \\
&= \left\langle c(y), e^{-iy\cdot\xi} \right\rangle \widehat{\varphi}(\xi) \\
&= \widehat{c}(\xi)\widehat{\varphi}(\xi).
\end{aligned}$$

对于一般情形，我们选取 $\varphi_k \xrightarrow{\mathcal{S}} \varphi$，其中 $\{\varphi_k\}_{k\geqslant 1} \in \mathcal{D}(\mathbb{R}^n)$。根据 1) 的证明，对任意的非负整数 p，我们有

$$N_p(\varphi * c - \varphi_k * c) \leqslant C_p N_{p+q}(\varphi - \varphi_k).$$

从而，

$$\varphi_k * c \xrightarrow{\mathcal{S}} \varphi * c,$$

从而根据 Fourier 变换的连续性，我们有

$$\widehat{\varphi_k * c}(\xi) = \widehat{\varphi_k}(\xi)\widehat{c}(\xi) \xrightarrow{\mathcal{S}} \widehat{\varphi * c}(\xi).$$

另外，我们还有 $\widehat{\varphi_k}(\xi)\widehat{c}(\xi) \xrightarrow{\mathcal{S}} \widehat{\varphi}(\xi)\widehat{c}(\xi)$（因为 $\widehat{c}$ 是多项式增长的，请参考本次作业），这就证明了定理。 □

定理 470. *对任意的 $u \in \mathcal{S}'(\mathbb{R}^n)$，$c \in \mathcal{E}'(\mathbb{R}^n)$，我们有 $u * c \in \mathcal{S}'(\mathbb{R}^n)$，即*

$$\mathcal{S}'(\mathbb{R}^n) \times \mathcal{E}'(\mathbb{R}^n) \xrightarrow{*} \mathcal{S}'(\mathbb{R}^n), \quad (u, c) \mapsto u * c.$$

进一步，我们还有

$$\widehat{u * c} = \widehat{c} \cdot \widehat{u},$$

其中，$\widehat{c}$ 是多项式增长的（实际上有衰减）光滑函数。

证明：首先证明 $u*c\in\mathcal{S}'(\mathbb{R}^n)$。对任意的 $\varphi\in\mathcal{D}(\mathbb{R}^n)$，存在常数 C 和 p，使得有

$$|\langle u*c,\varphi\rangle|=|\langle u,\check{c}*\varphi\rangle|\leqslant CN_p(\check{c}*\varphi)$$

利用上一个命题，我们有

$$|\langle u*c,\varphi\rangle|\leqslant C'N_{p+q}(\varphi).$$

所以，$u*c$ 是缓增的分布。

下面计算 $u*c$ 的 Fourier 变换：对任意给定的 $\varphi\in\mathcal{D}(\mathbb{R}^n)$，那么 $\psi=\widehat{\varphi}\in\mathcal{S}(\mathbb{R}^n)$。从而，

$$\begin{aligned}\langle\widehat{u*c},\psi\rangle&=\langle u*c,\widehat{\psi}\rangle=\langle u*c,(2\pi)^n\check{\varphi}\rangle\\&=\langle u,(2\pi)^n\check{c}*\check{\varphi}\rangle=\langle u,\widehat{\widehat{c*\varphi}}\rangle=\langle\widehat{u},\widehat{c*\varphi}\rangle\\&=\langle\widehat{u},\widehat{c}\widehat{\varphi}\rangle=\langle\widehat{c}\cdot\widehat{u},\psi\rangle.\end{aligned}$$

利用 Fourier 变换的连续性，$\mathcal{D}(\mathbb{R}^n)$ 在 $\mathcal{S}(\mathbb{R}^n)$ 的稠密性意味着上述所有可能的 ψ 在 $\mathcal{S}(\mathbb{R}^n)$ 中稠密，（通过逼近）所以上面的等式就给出了定理完整证明。 □

利用我们学过的知识，我们可以给出 $\mathcal{E}'(\mathbb{R}^n)$ 的一个漂亮的刻画：

定理 471 (有紧支集的分布的结构定理)**.** *对任意的有紧支集的分布 $c\in\mathcal{E}'(\mathbb{R}^n)$，存在有限多个多重指标 $\alpha_1,\cdots,\alpha_m$ 和有限多个有紧支集的连续函数 $f_{\alpha_1}(x),\cdots,f_{\alpha_m}(x)$，使得*

$$c=\sum_{k\leqslant m}\partial^{\alpha_k}f_{\alpha_k}(x).$$

注记. *我们甚至可以进一步要求 $|\alpha_1|=|\alpha_2|=\cdots=|\alpha_m|$。*

证明：我们选取截断函数 $\chi\in C_0^\infty(\mathbb{R}^n)$，使得 $\chi|_{\mathrm{supp}(c)}\equiv1$。那么，我们有

$$\widehat{c}(\xi)=\left\langle c,e^{-ix\cdot\xi}\right\rangle=\left\langle c,\chi(x)e^{-ix\cdot\xi}\right\rangle.$$

由于 c 的阶是有限的（记作是 p），所有存在常数 C，使得

$$|\widehat{c}(\xi)|\leqslant C\sup_{|\alpha|\leqslant p}\|\partial^\alpha\big(\chi(x)e^{-ix\cdot\xi}\big)\|_{L^\infty_x}\leqslant C(1+|\xi|)^p\sup_{|\beta|\leqslant|\alpha|}\|\partial^\beta\varphi\|_{L^\infty}.$$

所以，只要选取 $N\geqslant p+n+1$，我们就有

$$F(\xi)=\frac{1}{(1+|\xi|^2)^N}\widehat{c}(\xi)\in L^1(\mathbb{R}^n),$$

特别的，$\mathcal{F}^{-1}F\in C_\circ(\mathbb{R}^n)$ 是连续函数。通过导数与乘法在 Fourier 变换下的关系，我们就得到

$$(1-\triangle)^N\left(\mathcal{F}^{-1}F\right)=c.$$

从而，

$$\chi(x)(1-\triangle)^N\left(\mathcal{F}^{-1}F\right)=c.$$

利用 $\chi\cdot\partial_j=\partial_j(\chi\cdot)-\partial_j\chi$，我们可以把 χ（以及其导数）的乘法全部放到求导数运算里面去，这就给出了定理的证明。 □

69 分布理论与 Fourier 变换在数学物理方程上的应用：求波动方程与热方程的基本解. Sobolev 空间的定义，Sobolev 空间的映射性质

二零二零年十月二十六日，星期一，晴

我们现在考虑函数 $\varphi \in \mathcal{S}(\mathbb{R}^m \times \mathbb{R}^n)$，其中，我们用 $t \in \mathbb{R}^m$ 表示前面的 m 个变量，$x \in \mathbb{R}^n$ 表示前面的 n 个变量，我们可以考虑只对后面 n 个变量的 Fourier 变换：

$$\mathcal{F}_{x\to\xi}(\varphi)(t,\xi) = \int_{\mathbb{R}^n} e^{-ix\cdot\xi}\varphi(t,x)dx.$$

有时候，为了方便起见，我们把这个变换记做 $\widetilde{\varphi}(t,\xi)$ 或者干脆记做 $\widehat{\varphi}(t,\xi)$。我们在不同的场合总会指出具体对哪些变量做 Fourier 变换。类似地，我们可以定义对后面 n 个变量的 Fourier 逆变换：

$$\mathcal{F}^{-1}_{\xi\to x}(\psi(t,\xi))(t,x) = \frac{1}{(2\pi)^n}\int_{\mathbb{R}^n} e^{ix\cdot\xi}\psi(t,\xi)d\xi.$$

我们注意到 $\mathcal{F}_{x\to\xi}$ 把 $\mathbb{R}^m \times \mathbb{R}^n$ 上的 Schwartz 函数映射成为 $\mathbb{R}^m \times \mathbb{R}^n$ 上的 Schwartz 函数，即

$$\mathcal{F}_{x\to\xi} : \mathcal{S}(\mathbb{R}^m_t \times \mathbb{R}^n_x) \to \mathcal{S}(\mathbb{R}^m_t \times \mathbb{R}^n_x).$$

这是一个连续的线性同构，证明和之前关于 Fourier 变换的证明是一致的，我们留作作业来验证。

根据 $\mathcal{F}_{x\to\xi}$ 在 Schwartz 函数上的作用，我们就可以定义它在缓增分布上的作用：

$$\mathcal{F}_{x\to\xi} : \mathcal{S}'(\mathbb{R}^m_t \times \mathbb{R}^n_x) \to \mathcal{S}'(\mathbb{R}^m_t \times \mathbb{R}^n_x),\quad u \mapsto \mathcal{F}_{x\to\xi}(u),$$

其中，对于任意的 $\varphi \in \mathcal{S}(\mathbb{R}^m \times \mathbb{R}^n)$，我们有

$$\langle \mathcal{F}_{x\to\xi}(u), \varphi\rangle = \langle u, \mathcal{F}_{x\to\xi}(\varphi)\rangle.$$

我们同样可以证明，这是一个连续的线性同构。

例子. 考虑 $\delta_{0,0} \in \mathcal{S}'(\mathbb{R}^m_t \times \mathbb{R}^n_x)$。按照定义，我们有

$$\begin{aligned}\langle \mathcal{F}_{x\to\xi}(\delta_{0,0}), \varphi\rangle &= \langle u, \int_{\mathbb{R}^n} e^{-ix\cdot\xi}\varphi(t,x)dx\rangle = \int_{\mathbb{R}^n} e^{-ix\cdot\xi}\varphi(t,x)dx\Big|_{(t,\xi)=(0,0)}\\ &= \int_{\mathbb{R}^n}\varphi(0,x)dx = \int_{\mathbb{R}^n}\langle \delta_{t=0}, \varphi(t,x)\rangle dx.\end{aligned}$$

利用我们在第三次作业中定义的张量积，我们就有

$$\mathcal{F}_{x\to\xi}(\delta_{0,0})(t,\xi) = \delta_{t=0}\otimes 1_\xi.$$

热核的推导

我们已经证明过：定义在 $\mathbb{R}\times\mathbb{R}^n$ 上的热算子 $\partial_t-\triangle$ 以如下的热核函数作为其基本解：

$$E(t,x)=\frac{H(t)}{(4\pi t)^{\frac{n}{2}}}e^{-\frac{|x|^2}{4t}}.$$

然而，之前只是被动地验证了这个事实，我们现在给出它的推导。

我们要解如下的方程：

$$(\partial_t-\triangle)\,E=\delta_{0,0}.$$

我们做如下的**预设**：$E\in\mathcal{S}(\mathbb{R}\times\mathbb{R}^n)$。对上面方程的 x 变量做 Fourier 变换（仍然用$\widehat{\ }$来表示），我们就有

$$\left(\partial_t+|\xi|^2\right)\widehat{E}(t,\xi)=\delta_{t=0}\otimes 1_\xi.$$

我们要解上面这个常微分方程。注意到，右边分布的支集在 $\left\{(t,\xi)\middle|t=0\right\}$ 这个超平面上，所以，在 $t=0$ 之外，对于每个固定的 $\xi\in\mathbb{R}^n$，这个常微分方程的通解形如

$$c(\xi)e^{-t|\xi|^2}.$$

为了在 $t=0$ 处得到关于 t 的 Dirac 分布，我们很容易猜测

$$c(\xi)e^{-t|\xi|^2}H(t)$$

满足要求。现在计算热算子在这个分布上的作用：

$$\left(\partial_t+|\xi|^2\right)\left(c(\xi)e^{-t|\xi|^2}H(t)\right)=c(\xi)e^{-t|\xi|^2}\delta_{t=0}=\delta_{t=0}\otimes c(\xi).$$

我们只要取 $c(\xi)\equiv 1$ 即可。综上所述，我们就有

$$\widehat{E}(t,\xi)=e^{-t|\xi|^2}H(t).$$

对 ξ 做 Fourier 逆变换，（根据我们已有的计算）我们得到

$$E(t,x)=\frac{H(t)}{(4\pi t)^{\frac{n}{2}}}e^{-\frac{|x|^2}{4t}}.$$

$\mathbb{R}^n$ 中波动方程基本解的推导

我们考虑定义在 $\mathbb{R}^{1+n}=\mathbb{R}\times\mathbb{R}^n$ 上的算子波动算子 $\Box=-\partial_t^2+\triangle$，其中，第一个坐标是时间 t 的坐标。波动算子作用在以 $(t,x)\in\mathbb{R}\times\mathbb{R}^n$ 为变量函数上的。我们仍然预设我们要找的基本解 $W\in\mathcal{S}(\mathbb{R}\times\mathbb{R}^n)$，所以，

$$\Box W=\delta_{0,0}\quad\Leftrightarrow\quad-\left(\partial_t^2+|\xi|^2\right)\widehat{W}(t,\xi)=\delta_{t=0}\otimes 1_\xi.$$

这是一个二阶的线性常微分方程，与热核的情况相似，在 $t=0$ 之外，对于每个固定的 $\xi\in\mathbb{R}^n$，它的通解形如

$$a(\xi)\cos(t|\xi|)+b(\xi)\sin(t|\xi|).$$

在 $n=3$ 的时，我们寻找的基本解满足“在过去为零”的要求（这是物理上因果性的要求），与热方程的情况比较，我们对 $\widehat{W}(t,\xi)$ 的形状先做出如下的猜测：

$$\widehat{W}(t,\xi)=H(t)\left(a(\xi)\cos(t|\xi|)+b(\xi)\sin(t|\xi|)\right).$$

所以，

$$\begin{aligned}\partial_t\left(\widehat{W}(t,\xi)\right)&=H(t)\left(-a(\xi)|\xi|\sin(t|\xi|)+b(\xi)|\xi|\cos(t|\xi|)\right)+\delta_{t=0}\left(a(\xi)\cos(t|\xi|)+b(\xi)\sin(t|\xi|)\right)\\&=H(t)\left(-a(\xi)|\xi|\sin(t|\xi|)+b(\xi)|\xi|\cos(t|\xi|)\right)+\delta_{t=0}\otimes a(\xi).\end{aligned}$$

当然，如果我们再对这个式子求导数，最后一项可能会贡献出 $\delta'_{t=0}$，这不会满足基本解的公式，所以，我们假设 $a(\xi)\equiv 0$。据此，我们有

$$\begin{aligned}\widehat{W}(t,\xi)&=H(t)b(\xi)\sin(t|\xi|),\\\partial_t\left(\widehat{W}(t,\xi)\right)&=H(t)b(\xi)|\xi|\cos(t|\xi|).\end{aligned}$$

从而，

$$\partial_t^2\left(\widehat{W}(t,\xi)+|\xi|^2\right)=\delta_{t=0}\cdot b(\xi)|\xi|\cos(t|\xi|).$$

据此，我们令

$$\widehat{W}(t,\xi)=-H(t)\frac{\sin(t|\xi|)}{|\xi|}.$$

在 $n=3$ 的情形，我们已经证明了

$$\widehat{\frac{d\sigma_{\mathbb{S}_R}}{4\pi R}}=\frac{\sin(R|\xi|)}{|\xi|}.$$

所以，当 $n=3$ 时，我们有

$$\widehat{W}(t,\xi)=-H(t)\frac{d\sigma_{\mathbb{S}_t^2}}{4\pi t}.$$

最后，我们可以和之前已经给出的基本解做比较：

$$W=-\frac{d\sigma}{4\pi\sqrt{t^2+|x|^2}},$$

其中，$d\sigma$ 是正向光锥的测度。这两个计算是一致的。

分布理论与 Fourier 变换的应用：Sobolev 空间及应用

在后面的课程中，我们会经常用所谓的 **Planchrel 公式**：对任意的 $f\in L^2(\mathbb{R}^n)$，我们有

$$\|\widehat{f}\|_{L^2}^2=(2\pi)^n\|f\|_{L^2}^2\quad\Leftrightarrow\quad\int_{\mathbb{R}^n}|f(x)|^2dx=\frac{1}{(2\pi)^n}\int_{\mathbb{R}^n}|\widehat{f}(\xi)|^2d\xi.$$

它的另外一个版本是说对任意的 $f,g\in L^2(\mathbb{R}^n)$，我们有

$$(\widehat{f},\widehat{g})_{L^2}=(2\pi)^n(f,g)_{L^2}\quad\Leftrightarrow\quad\int_{\mathbb{R}^n}f(x)\overline{g(x)}dx=\frac{1}{(2\pi)^n}\int_{\mathbb{R}^n}\widehat{f}(\xi)\overline{\widehat{g}(\xi)}d\xi.$$

这个公式我们之前已经证明过。

我们现在引入 $\mathbb{R}^n$ 上的 Sobolev 空间的定义。

定义 472. 给定 $s \in \mathbb{R}$，我们将把这个数称为是 *Sobolev* 空间的指标。我们考虑满足如下性质的缓增分布 $u \in \mathcal{S}'(\mathbb{R}^n)$：

1) $\widehat{u} \in L^1_{\text{loc}}(\mathbb{R}^n)$ 是局部可积的函数；

2) $(1+|\xi|^2)^{\frac{s}{2}}\widehat{u}(\xi)$ 是平方可积的函数。

对于这样的函数，我们定义其 *Sobolev* 范数为：

$$\|u\|_{H^s} = \left(\int_{\mathbb{R}^n} (1+|\xi|^2)^s |\widehat{u}(\xi)|^2 d\xi\right)^{\frac{1}{2}}.$$

我们把所有满足上述条件的缓增分布的集合称作是一个**指标为 s 的 Sobolev 空间**，这显然是一个复线性空间，我们用 $H^s(\mathbb{R}^n)$ 来表示。

在 $H^s(\mathbb{R}^n)$ 上所赋予的范数与下面的内积是相容的：对任意的 $u, v \in H^s(\mathbb{R}^n)$，令

$$(u,v)_{H^s} = \int_{\mathbb{R}^n} (1+|\xi|^2)^s \widehat{u}(\xi)\overline{\widehat{v}(\xi)} d\xi.$$

所以，$\left(H^s(\mathbb{R}^n), (\cdot,\cdot)_{H^s}\right)$ 是内积空间。

我们注意到，当 $s=0$ 时，我们 $H^0(\mathbb{R}^n)$ 实际上就是 $L^2(\mathbb{R}^n)$，这由 Planchrel 公式立即就可以得到：

$$u \in L^2(\mathbb{R}^n) \quad \Leftrightarrow \quad \widehat{u} \in L^2(\mathbb{R}^n).$$

所以，

$$H^0(\mathbb{R}^n) = L^2(\mathbb{R}^n).$$

类似的，如果我们在频率空间 $\mathbb{R}^n_\xi$ 上考虑测度

$$\mu_s = (1+|\xi|^2)^s d\xi.$$

那么，$u \in H^s(\mathbb{R}^n)$ 当且仅当 $\widehat{u} \in L^2(\mathbb{R}^n, d\mu_s)$。利用这个观察，我们现在证明：

定理 473. 对任意的 $s \in \mathbb{R}$，$\left(H^s(\mathbb{R}^n), (\cdot,\cdot)_{H^s}\right)$ 是 *Hilbert* 空间（即完备的内积空间）。

证明： 假设 $\{u_k\}_{k\geqslant 1} \subset H^s(\mathbb{R}^n)$ 是 Cauchy 列，那么，根据定义，$\{\widehat{u_k}\}_{k\geqslant 1} \subset L^2(\mathbb{R}^n, d\mu_s)$ 是 Cauchy 列。利用 L^2-空间的完备性，存在 $v(\xi) \in L^2(\mathbb{R}^n, d\mu_s)$ 作为上述序列的极限。我们用 $u(x) \in \mathbb{S}'(\mathbb{R}^n)$ 表示它的 Fourier 逆变换，即

$$\widehat{u} = v.$$

那么，

$$\lim_{k\to\infty} \|\widehat{u_k} - \widehat{u}\|^2_{L^2(\mathbb{R}^n, d\mu_s)} = 0 \quad \Leftrightarrow \quad \lim_{k\to\infty} \|u_k - u\|^2_{H^s(\mathbb{R}^n)}) = 0.$$

这就证明了完备性。 □

根据 Sobolev 空间的定义，我们知道 $\{H^s(\mathbb{R}^n)\}_{s\in\mathbb{R}}$ 构成了一个下降的链，即对任意的 $s, s' \in \mathbb{R}$

$$s \geqslant s' \quad \Rightarrow \quad H^{s'}(\mathbb{R}^n) \subset H^s(\mathbb{R}^n).$$

我们观察到，Schwartz 函数生活在所有的 Sobolev 空间中：

$$\mathcal{S}(\mathbb{R}^n) \subset \bigcap_{s\in\mathbb{R}} H^s(\mathbb{R}^n).$$

为了说明这个一点，我们再次运用我们熟悉的一个技巧。对任意的 $\varphi \in \mathcal{S}(\mathbb{R}^n)$，我们只要说明它的 H^s-范数是有界的即可：

$$\begin{aligned}
\|\varphi\|_{H^s}^2 &= \int_{\mathbb{R}^n} \underbrace{(1+|\xi|^2)^{s+\frac{n+1}{2}}|\widehat{\varphi}|^2}_{\in L^\infty} \underbrace{(1+|\xi|^2)^{-\frac{n+1}{2}}}_{\in L^1} d\xi \\
&\leqslant C N_{2s+n+1}(\widehat{\varphi})^2 \int_{\mathbb{R}^n} (1+|\xi|^2)^{-\frac{n+1}{2}} d\xi \\
&\leqslant C' N_{2s+2n+2}(\varphi)^2.
\end{aligned}$$

我们在之后会证明

$$\mathcal{D}'(\mathbb{R}^n) \subset \varprojlim_{s\in\mathbb{R}} H^s(\mathbb{R}^n) = \bigcup_{s\in\mathbb{R}} H^s(\mathbb{R}^n).$$

命题 474 (正整数阶的 Sobolev 空间)**.** *假设 $m \in \mathbb{Z}_{\geqslant 1}$ 为正整数，那么，$H^m(\mathbb{R}^n)$ 有如下的等价刻画：*

$$H^m(\mathbb{R}^n) = \left\{ u \in \mathcal{S}'(\mathbb{R}) \middle| \text{对任意的多重指标} \alpha,\ |\alpha| \leqslant m,\ \partial^\alpha u \in L^2(\mathbb{R}^n) \right\}.$$

证明： 这个命题的证明基于如下的一个简单的观察：给定正整数 m，存在常数 C_1 和 C_2，使得对任意的 $\xi \neq 0$，我们有

$$C_1(1+|\xi|^2)^m \leqslant \sum_{|\alpha|\leqslant m} |\xi^\alpha|^2 \leqslant C_2(1+|\xi|^2)^m,$$

其中 $|\xi^\alpha| = |\xi_1^{\alpha_1}\xi_2^{\alpha_2}\cdots\xi_n^{\alpha_n}|$。这个证明是初等的，我们留作作业来验证。

所以，在差一个常数的意义下，我们就有

$$\begin{aligned}
\int_{\mathbb{R}^n} (1+|\xi|^2)^s |\widehat{u}|^2 d\xi &\approx \sum_{|\alpha|\leqslant m} \int_{\mathbb{R}^n} |\xi^\alpha \widehat{u}|^2 d\xi \\
&\overset{Planchrel}{\approx} \sum_{|\alpha|\leqslant m} \int_{\mathbb{R}^n} |\partial^\alpha u|^2 dx.
\end{aligned}$$

上式最后一个积分有限就等价于说对任意的多重指标 α，我们有 $\partial^\alpha u \in L^2(\mathbb{R}^n)$，其中 $|\alpha| \leqslant m$，这就证明了命题。 □

Sobolev 空间的映射性质

为了研究 Sobolev 空间的映射性质，我们先引入一类比微分算子更广的算子。首先，我们回忆一下，对于任意的缓增分布，对任意的 $k \leqslant n$，我们有

$$\widehat{\frac{1}{i}\partial_k u}(\xi) = \xi \widehat{u}(\xi).$$

我们定义算子

$$D_k = \frac{1}{i}\partial_k = -i\partial_k, k = 1, 2, \cdots, n.$$

为了简单起见，我们还把它写成

$$D = \frac{1}{i}\partial.$$

形式上，D 对一个分布的作用在频率空间上来看就是乘以 ξ。

定义 475 (Fourier 乘子)**.** 给定频率空间上的函数 $m(\xi)$，我们假设它是多项式增长的。对于任意的缓增分布 $u \in \mathcal{S}'(\mathbb{R}^n)$，我们定义

$$m(D)u = \mathcal{F}^{-1}\left(m(\xi)\widehat{u}(\xi)\right) \quad \Leftrightarrow \quad \widehat{m(D)u} = m(\xi)\widehat{u}(\xi).$$

由于 $m(\xi)$ 是多项式增长的，所以，$m(\xi)\widehat{u}(\xi)$ 仍然是缓增分布，所以，如下的算子是良好定义的：

$$m(D) : \mathcal{S}'(\mathbb{R}^n) \to \mathcal{S}'(\mathbb{R}^n).$$

例子. 我们先看几个简单的例子：

1) 当 $m(\xi) = \xi_k$ 时，其中 $k = 1, 2, \cdots, n$，我们有

$$m(D) = \frac{1}{i}\partial_k = D_k.$$

2) 当 $m(\xi) = |\xi|^2$ 时，我们有

$$m(D) = -\triangle.$$

3) 给定线性微分算子

$$P = \sum_{|\alpha| \leqslant m} a_\alpha \partial^\alpha,$$

它可以被视作是一个 *Fourier* 乘子 $m(D)$，其中

$$m(\xi) = \sum_{|\alpha| \leqslant m} i^{|\alpha|} a_\alpha \xi^\alpha,$$

4) 算子 $(1 - \triangle)^s$ 表示的是函数 $(1 + |\xi|^2)^s$ 所对应的 *Fourier* 乘子。

70 Sobolev 空间的基本性质，Fourier 乘子，Sobolev 空间的稠密子空间，有紧支集分布的结构定理，Sobolev 嵌入定理，高指数的 Sobolev 空间是一个代数

二零二零年十月二十八日，星期三，晴

我们上次课对多项式增长的乘子函数 $m(\xi)$ 定义了 Fourier 乘子：

$$m(D)u = \mathcal{F}^{-1}\left(m(\xi)\widehat{u}(\xi)\right) \quad \Leftrightarrow \quad \widehat{m(D)u} = m(\xi)\widehat{u}(\xi),$$

其中，u 是缓增的分布。有同学认为这个符号具有迷惑性，我们采取另一种说法：

对任意的在频率空间 $\mathbb{R}^n_\xi$ 多项式增长的乘子函数 $m(\xi)$，它决定了一个算子 $P = P_{m(\xi)}$：

$$P : \mathcal{S}'(\mathbb{R}^n) \to \mathcal{S}'(\mathbb{R}^n),$$

其中，$u \in \mathcal{S}'(\mathbb{R}^n)$ 的像在频率空间上为

$$\widehat{P(u)}(\xi) = m(\xi)\widehat{u}(\xi).$$

很重要的一类 Fourier 乘子是线性常系数的微分算子

$$P = \sum_{|\alpha| \leqslant m} a_\alpha \partial^\alpha,$$

如果我们选取

$$m(\xi) = \sum_{|\alpha| \leqslant m} i^{|\alpha|} a_\alpha \xi^\alpha,$$

那么，$P = m(D)$。

注记 (定义的订正)**.** 我们上次课程还引入了 *Sobolev* 空间。在课堂上，我们的定义有一定的问题，现在我们将它略加改正：给定 $s \in \mathbb{R}$，如果缓增分布 $u \in \mathcal{S}'(\mathbb{R}^n)$ 满足

1) $\widehat{u} \in L^1_{\mathrm{loc}}(\mathbb{R}^n)$ 是局部可积的函数；

2) $(1 + |\xi|^2)^{\frac{s}{2}}\widehat{u}(\xi)$ 是平方可积的函数。

我们就说 $u \in H^s(\mathbb{R}^n)$。

我们在此给出两个重要的例子：

例子. 下面的 *Sobolev* 空间都定义在 $\mathbb{R}^n$ 上。

1) 对任意的 $s < -\dfrac{n}{2}$，我们有

$$\delta_0 \in H^s(\mathbb{R}^n).$$

实际上，我们只要说明下面的积分有限即可：

$$\int_{\mathbb{R}^n}(1+|\xi|^2)^s\cdot 1d\xi.$$

这在 $s<-\dfrac{n}{2}$ 时是成立的。同样的推理表明，当 $s\geqslant -\dfrac{n}{2}$ 时，$\delta_0\notin H^s(\mathbb{R}^n)$。

2) 常数值函数 1 不在任何的 $H^s(\mathbb{R}^n)$ 中。特别地，这表明

$$\bigcup_{s\in\mathbb{R}}H^s(\mathbb{R}^n)\subsetneqq \mathcal{S}'(\mathbb{R}^n).$$

我们首先说明 Sobolev 空间在 Fourier 乘子下的映射性质。

命题 476 (Sobolev 空间的映射性质)**.** 给定多项式增长的乘子函数 $m(\xi)$，其中，我们假设存在常数 C 和 p，使得对任意的 $\xi\in\mathbb{R}^n$，我们都有

$$|m(\xi)|\leqslant C(1+|\xi|)^p.$$

那么，对任意的 $s\in\mathbb{R}$，对任意的 $u\in H^s(\mathbb{R}^n)$，$m(D)u\in H^{s-p}(\mathbb{R}^n)$。这就定义出有界（连续）线性映射：

$$m(D):H^s(\mathbb{R}^n)\to H^{s-p}(\mathbb{R}^n).$$

特别地，对任意的 d-阶的微分算子 P，对任意的 $s\in\mathbb{R}$，我们有连续线性映射

$$P:H^s(\mathbb{R}^n)\to H^{s-d}(\mathbb{R}^n),\quad \forall s.$$

另外，对任意的 s，我们还有连续的线性同构*:*

$$(1+\triangle)^{\frac{p}{2}}:H^s(\mathbb{R}^n)\to H^{s-p}(\mathbb{R}^n).$$

其中，上述映射的逆映射是 $(1+\triangle)^{-\frac{p}{2}}$。

证明*:* 对任意的 $u\in H^s(\mathbb{R}^n)$，我们首先证明 $m(D)u\in H^{s-p}(\mathbb{R}^n)$，其中 $m(\xi)$ 具有命题中所要求的多项式增长。根据 Planchrel 公式，我们有

$$\begin{aligned}\|m(D)u\|^2_{H^{s-p}}&=\int_{\mathbb{R}^n}(1+|\xi|^2)^{s-p}|m(\xi)\widehat{u}(\xi)|^2d\xi\\&\leqslant C\int_{\mathbb{R}^n}(1+|\xi|^2)^{s-p}(1+|\xi|)^{2p}|\widehat{u}(\xi)|^2d\xi\\&\leqslant C'\int_{\mathbb{R}^n}(1+|\xi|^2)^{s}|\widehat{u}(\xi)|^2d\xi.\end{aligned}$$

所以，存在常数 C_1，使得

$$\|m(D)u\|_{H^{s-p}}\leqslant C_1\|u\|_{H^s}.$$

这表明 $m(D)$ 是从 $H^s(\mathbb{R}^n)$ 到 $H^{s-p}(\mathbb{R}^n)$ 的连续线性映射。

微分算子的情形是一个特例。为了说明 $m(D)=(1+\triangle)^{\frac{p}{2}}$ 有逆，我们用

$$n(\xi)=(1+|\xi|)^{-\frac{p}{2}}$$

作为乘子即可，这是因为

$$\widehat{m(D)n(D)u}=(1+|\xi|)^{\frac{p}{2}}(1+|\xi|)^{-\frac{p}{2}}\widehat{u}(\xi)=u(\xi).$$

命题得证。 □

利用 Sobolev 空间之间的映射性质，我们可以证明如下的稠密性定理：

命题 477. *对每个指标 $s\in\mathbb{R}^n$，光滑有紧支集的函数 $C_0^\infty(\mathbb{R}^n)$ 在 $H^s(\mathbb{R}^n)$ 是稠密的。*

证明： 我们首先证明 $\mathcal{S}(\mathbb{R}^n)\subset H^s(\mathbb{R}^n)$ 是稠密的，其中 $s\in\mathbb{R}$：这个论断对 $s=0$ 是正确的，因为 $C_0^\infty(\mathbb{R}^n)$ 在 $H^0(\mathbb{R}^n)=L^2(\mathbb{R}^n)$ 中是稠密的。由于

$$(1-\triangle)^{-\frac{s}{2}}:H^0(\mathbb{R}^n)\to H^2(\mathbb{R}^n)$$

是连续可逆的线性映射（是同胚），所以 $\mathcal{S}(\mathbb{R}^n)$ 在这个算子下的像也是稠密的，然而，

$$(1-\triangle)^{-\frac{s}{2}}\left(\mathcal{S}(\mathbb{R}^n)\right)\subset\mathcal{S}(\mathbb{R}^n),$$

所以，$\mathcal{S}(\mathbb{R}^n)\subset H^s(\mathbb{R}^n)$ 是稠密的。

（注意到，在上面的论证中，$(1-\triangle)^{-\frac{s}{2}}$ 不一定把有紧支集的函数映射为有紧支集的函数，所以，我们的推理是对 $\mathcal{S}(\mathbb{R}^n)$ 进行的。）

为了证明命题，我们只要说明在 $H^s(\mathbb{R}^n)$ 的意义下，$\mathcal{S}(\mathbb{R}^n)$ 中的任意一个函数 f 都可以可以被 $C_0^\infty(\mathbb{R}^n)$ 的函数逼近。我们上次证明了存在常数 C，使得对任意的 $\psi\in\mathcal{S}(\mathbb{R}^n)$，我们有不等式

$$\|\psi\|_{H^s}\leqslant CN_{s+n+1}(\psi).$$

所以，对任意的 $f\in H^s(\mathbb{R}^n)$，我们先选取 $\psi\in\mathcal{S}(\mathbb{R}^n)$，使得

$$\|f-\psi\|_{H^s}<\frac{\varepsilon}{2}.$$

再利用 $C_0^\infty(\mathbb{R}^n)\subset\mathcal{S}(\mathbb{R}^n)$ 的稠密性，选取 $\varphi\in C_0^\infty(\mathbb{R}^n)$，使得

$$N_{s+n+1}(\psi-\varphi)<\frac{\varepsilon}{2C}.$$

此时，我们有

$$\|f-\varphi\|_{H^s}<\varepsilon.$$

这就证明了命题。 □

定理 478 (有紧支集分布的结构定理，H^s-版本)**.** 对任意的有紧支集的分布 $c \in \mathcal{E}'(\mathbb{R}^n)$，假设 p 是 c 的阶。那么，对任意的 $s < -p - \dfrac{n}{2}$，我们有

$$c \in H^s(\mathbb{R}^n).$$

特别地,存在有限多个多重指标 $\alpha_1, \cdots, \alpha_m$ 和有限多个有紧支集的平方可积函数 $f_{\alpha_1}(x), \cdots, f_{\alpha_m}(x) \in L^2(\mathbb{R}^n)$，使得

$$c = \sum_{k \leqslant m} \partial^{\alpha_k} f_{\alpha_k}(x).$$

证明: 我们之前证明过 $\widehat{c}(\xi)$ 是具有多项式增长的，即存在常数 $c > 0$，使得

$$|\widehat{c}(\xi)| \leqslant C(1 + |\xi|)^p.$$

据此，我们知道对任意的 $\varepsilon > 0$，$(1 + |\xi|^2)^{-\frac{p}{2} - \frac{n}{4} - \varepsilon} \widehat{c}(\xi)$ 是平方可积的，这因为

$$(1 + |\xi|^2)^{-p - \frac{n}{2} - 2\varepsilon} |\widehat{c}(\xi)|^2 \leqslant C'(1 + |\xi|)^{-n - 2\varepsilon}.$$

右边的函数是可积的。据此，对任意的 $s < -p - \dfrac{n}{2}$，我们有 $c \in H^s(\mathbb{R}^n)$。

类似地，我们知道

$$g(\xi) = (1 + |\xi|^2)^{-\lfloor \frac{p}{2} \rfloor - n - 1} \widehat{c}(\xi) \in L^2(\mathbb{R}^n).$$

令 $N = \lfloor \dfrac{p}{2} \rfloor - n - 1$，所以，

$$(1 - \triangle)^N f = c.$$

其中 $\widehat{f} = g$ 是平方可积的函数。再选取 χ 是在 c 的支集上恒为 1 的光滑的有紧支集的函数，所以，

$$\chi \cdot (1 - \triangle)^N f = c.$$

类似于之前对有紧支集分布的结构定理的证明，将这个式子展开即可。 □

我们下面证明著名的 Sobolev 嵌入定理（的一种形式）:

定理 479 (Sobolev 嵌入定理)**.** 假设指标 $s > \dfrac{n}{2}$，那么，每个 $u \in H^s(\mathbb{R}^n)$ 都落在 $L^\infty(\mathbb{R}^n)$ 中。进一步，我们有连续的线性嵌入

$$\iota : H^s(\mathbb{R}^n) \hookrightarrow L^\infty(\mathbb{R}^n), \quad u \mapsto u,$$

即存在 C_s，使得对任意 $u \in H^s(\mathbb{R}^n)$，我们都有

$$\|u\|_{L^\infty} \leqslant C_s \|u\|_{H^s}.$$

进一步，u 是连续函数（可以在它的代表类中选到一个连续函数）并且在无穷远处的极限为零，即 $u \in C_\circ(\mathbb{R}^n)$。

注记. 证明的想法比较简单：我们只要说明 $\widehat{u}$ 是一个 L^1 函数即可，因为 *Fourier* 逆变换就把它还原成一个在 ∞ 处衰减的连续函数，从而是 L^∞ 的函数。

证明: 根据 $u = \mathcal{F}^{-1}(\widehat{u})$，我们知道

$$\|u\|_{L^\infty} \leqslant \frac{1}{(2\pi)^n}\|\widehat{u}\|_{L^1}.$$

我们现在说明 $\|\widehat{u}(\xi)\|_{L^1}$ 被 $\|u\|_{H^s}$ 所控制。根据 $s > \frac{n}{2}$，我们可以 $\widehat{u}$ 写成两个平方可积的函数的乘积:

$$\widehat{u}(\xi) = \underbrace{(1+|\xi|^2)^{\frac{s}{2}}\widehat{u}(\xi)}_{L^2} \cdot \underbrace{(1+|\xi|^2)^{-\frac{s}{2}}}_{L^2},$$

前一部分根据 $u \in H^s(\mathbb{R}^n)$ 所以是 L^2 的; 后一部分根据 $s > \frac{n}{2}$ 所以是 L^2 的。利用 Cauchy-Schwarz 不等式，我们有

$$\|\widehat{u}(\xi)\|_{L^1} \leqslant \|u\|_{H^s}\|(1+|\xi|^2)^{-\frac{s}{2}}\|_{L^2} = C\|u\|_{H^s}.$$

所以，

$$\|u\|_{L^\infty} \leqslant \frac{C}{(2\pi)^n}\|u\|_{H^s}.$$

连续性的部分是明显的，因为

$$\mathcal{F}^{-1} : L^1(\mathbb{R}^n_\xi) \longrightarrow C_\circ(\mathbb{R}^n_x).$$

证明完毕。 □

推论 480. 假设 $s > \dfrac{n}{2}+k$, 其中 k 为非负整数, 那么, 对任意的 $u \in H^s(\mathbb{R}^n)$, 我们都有 $u \in C^k(\mathbb{R}^n)$。

证明: 对任意的多重指标 α, 如果 $|\alpha| \leqslant k$, 那么 $\partial^\alpha u \in H^{s-|\alpha|}(\mathbb{R}^n)$ 是连续函数, 从而, $u \in C^k(\mathbb{R}^n)$ (用归纳法来证明会更严格一点)。 □

注记. 这个版本的 *Sobolev* 嵌入定理说的是，如果指标 s 足够大，那么，函数 u 就会非常光滑。

注记. 我们在作业中将构造函数局部可积的 $u \in H^{\frac{n}{2}}(\mathbb{R}^n)$，使得 $u \notin L^\infty(\mathbb{R}^n)$。换句话说，如下的嵌入并不成立:

$$H^{\frac{n}{2}}(\mathbb{R}^n) \not\hookrightarrow L^\infty(\mathbb{R}^n).$$

这表明 *Sobolev* 嵌入的指标至少是 $\dfrac{n}{2}+\varepsilon$，其中 $\varepsilon > 0$ 可以任意小。

下一个定理说的是如果指标 s 足够大，那么，两个 H^s 的函数的乘积也是 H^s 的。这个定理在证明非线性偏微分方程的解的局部存在性时很有用。

定理 481. 如果 $s > \frac{n}{2}$，那么，$H^s(\mathbb{R}^n)$ 是一个代数，即对任意的 $u, v \in H^s(\mathbb{R}^n)$，我们有 $u \cdot v \in H^s(\mathbb{R}^n)$:

$$H^s(\mathbb{R}^n) \times H^s(\mathbb{R}^n) \xrightarrow{\times} H^s(\mathbb{R}^n).$$

实际上，存在常数 C_s，使得对任意的 $u, v \in H^s(\mathbb{R}^n)$，我们有

$$\|u \cdot v\|_{H^s} \leqslant C_s\|u\|_{H^s}\|v\|_{H^s}.$$

证明: 我们来计算 $u\cdot v$ 的 H^s-范数。由于在 Fourier 变换下，乘积变化为卷积，所以按照定义，我们有

$$\begin{aligned}\|u\cdot v\|_{H^s}^2 &= \int_{\mathbb{R}^n}(1+|\xi|^2)^s\left|\int_{\mathbb{R}^n}\widehat{u}(\xi-\eta)\widehat{v}(\eta)d\eta\right|^2 d\xi\\ &\leqslant \int_{\mathbb{R}^n}(1+|\xi|^2)^s\left(\int_{\mathbb{R}^n}|\widehat{u}(\xi-\eta)||\widehat{v}(\eta)|d\eta\right)^2 d\xi.\end{aligned}$$

我们要把因子 $(1+|\xi|^2)^s$ 进行拆分。首先，对任意的 $s>0$，$a,b\geqslant 0$，我们显然有

$$(a+b)^s\leqslant 2^s(a^s+b^s).$$

所以，对任意的 $\xi,\eta\in\mathbb{R}^n$，我们有如下的不等式

$$\begin{aligned}(1+|\xi|^2)^{\frac{s}{2}} &\leqslant (1+2|\xi-\eta|^2+2|\eta|^2)^{\frac{s}{2}}\leqslant 2^{\frac{s}{2}}\left((1+|\xi-\eta|^2)+(1+|\eta|^2)\right)^{\frac{s}{2}}\\ &\leqslant 2^{2s}\left((1+|\xi-\eta|^2)^{\frac{s}{2}}+(1+|\eta|^2)^{\frac{s}{2}}\right).\end{aligned}$$

所以，我们就得到了

$$\begin{aligned}\|u\cdot v\|_{H^s}^2 &\leqslant 2^{2s}\int_{\mathbb{R}^n}\left(\int_{\mathbb{R}^n}\big(1+|\xi-\eta|^2)^{\frac{s}{2}}|\widehat{u}(\xi-\eta)||\widehat{v}(\eta)|+(1+|\eta|^2)^{\frac{s}{2}}|\widehat{v}(\eta)||\widehat{u}(\xi-\eta)|\big)d\eta\right)^2 d\xi\\ &\leqslant 2^{2s}\int_{\mathbb{R}^n}\left(\int_{\mathbb{R}^n}\underbrace{(1+|\xi-\eta|^2)^{\frac{s}{2}}|\widehat{u}(\xi-\eta)|}_{f(\xi-\eta)}\underbrace{|\widehat{v}(\eta)|}_{g(\eta)}\right)^2+\left(\int_{\mathbb{R}^n}(1+|\eta|^2)^{\frac{s}{2}}|\widehat{v}(\eta)||\widehat{u}(\xi-\eta)|\big)d\eta\right)^2 d\xi.\end{aligned}$$

上面的表达式中本质上是两项，它们的结构是类似的，我们只要处理一项就好。我们现在利用第一项中的卷积结构来控制它。根据 H^s 的定义，我们有

$$f(\xi)=\big(1+|\xi|^2\big)^{\frac{s}{2}}|\widehat{u}(\xi)|\in L^2(\mathbb{R}^n_\xi).$$

根据 $s>\dfrac{n}{2}$，我们在 Sobolev 不等式的证明中已经证明了 $g(\xi)=\widehat{v}\in L^1(\mathbb{R}^n)$。特别地，存在常数 C_1 和 C_2，使得

$$\|f\|_{L^2}\leqslant C_1\|u\|_{H^s},\quad \|g\|_{L^2}\leqslant C_2\|v\|_{H^s}.$$

我们观察到，上面就是控制 $f*g$ 的 L^2 范数的大小。我们回忆上学期（5 月 9 日的课程，利用 Fubini 定理）已经证明了

$$L^1(\mathbb{R}^n)\times L^2(\mathbb{R}^n)\xrightarrow{*}L^2(\mathbb{R}^n).$$

其中对任意的 $\varphi\in L^1(\mathbb{R}^n)$，$\psi\in L^2(\mathbb{R}^n)$，我们有

$$\|\varphi*\psi\|_{L^2}\leqslant\|\varphi\|_{L^1}\|\psi\|_{L^2}.$$

我们对 $\varphi=g$ 和 $\psi=f$ 运用这个不等式，就得到

$$\|f*g\|_{L^2}\leqslant\|g\|_{L^1}\|f\|_{L^2}\leqslant C_1C_2\|u\|_{H^s}\|v\|_{H^s}.$$

所以，我们就有

$$\|u \cdot v\|_{H^s}^2 \leqslant 2^{2s} \left(C_1 C_2 \|u\|_{H^s}^2 \|v\|_{H^s}^2 + C_1 C_2 \|u\|_{H^s}^2 \|v\|_{H^s}^2 \right).$$

这就给出了命题的证明。 □

实际上，我们还可以证明更强的结论：对任意的 $s > 0$，$H^s(\mathbb{R}^n) \cap L^\infty(\mathbb{R}^n)$ 是一个代数。我们之后将利用频率空间的二进分解进行证明。

70.1 作业：Fourier 变换的计算，Heisenberg 测不准原理，分数次 Sobolev 空间的物理空间刻画，1 维的 Poisson 求和公式

清华大学 19-20 春季学期，数学分析三，作业 4

请用 A4 大小的纸张**正反面**用钢笔，签字笔或者圆珠笔书写，并注明自己的姓名，年级和作业的**总页数**。除定理公式所涉及的人名之外，请使用**中文**。本次作业的提交时间和地点为 **11 月 11 日**（周三）上午的课堂上，逾期视作零分。

习题 A.（课堂细节的补充） 我们总假设 $\Omega \subset \mathbb{R}^n$ 是非空的开集。

A1) 假设 $p = 1, 2$ 或 ∞。证明，如果把函数空间 $L^p(\mathbb{R}^n)$ 视为分布（这些空间中的函数都是局部可积的函数），那么它们是缓增的分布。从而，我们有良好定义的嵌入

$$\iota_p : L^p(\mathbb{R}^n) \hookrightarrow \mathcal{S}'(\mathbb{R}^n).$$

进一步证明 ι_p 都是连续的。

A2) 给定光滑函数 $p(x)$，如果对任意的多重指标 α，存在正数 C_α 和 d_α，使得

$$|\partial^\alpha p(x)| \leqslant C_\alpha (1 + |x|)^{d_\alpha},$$

我们就称 $p(x)$ 是**具有多项式增长的**函数。

证明，给定具有多项式增长的函数 $p(x)$，那么，对任意的 $\varphi \in \mathcal{S}(\mathbb{R}^n)$，$p \cdot \varphi \in \mathcal{S}(\mathbb{R}^n)$；对任意的 $u \in \mathcal{S}'(\mathbb{R}^n)$，$p \cdot \varphi \in \mathcal{S}'(\mathbb{R}^n)$。

证明，如下两个映射都是连续的：

$$\mathcal{S}(\mathbb{R}^n) \to \mathcal{S}(\mathbb{R}^n), \quad \varphi \mapsto p \cdot \varphi;$$
$$\mathcal{S}'(\mathbb{R}^n) \to \mathcal{S}'(\mathbb{R}^n), \quad u \mapsto p \cdot u.$$

A3) 给定分布 $u \in \mathcal{D}'(\mathbb{R}^n)$ 和 $\mathbb{R}^n$ 上的光滑函数 $f(x)$。假设 $f \cdot u \overset{\mathcal{D}'}{=} 0$，那么，

$$\operatorname{supp}(u) \subset Z(f) = \big\{x \in \mathbb{R}^n \big| f(x) = 0\big\}.$$

A4) 证明，对任意的 $a \in \mathbb{R}^n$，任意的多重指标 α，我们有

$$\widehat{\partial^\alpha \delta_a} = (i\xi)^\alpha e^{-ix\cdot\xi}, \quad \widehat{x^\alpha} = (2\pi)^n (i\xi)^\alpha \delta_0.$$

特别地，

$$\widehat{\partial^\alpha \delta_a} = (i\xi)^\alpha e^{-ix\cdot\xi}, \quad \widehat{x^\alpha} = (2\pi)^n (i\xi)^\alpha \delta_0.$$

A5) 证明，分布 $d\sigma_R$ 是旋转不变的，其中 $d\sigma_R$ 为 $\mathbb{R}^n$ 上中心在原点半径为 R（> 0）的球面的曲面测度。

A6) 给定 $u \in \mathcal{S}'(\mathbb{R}^n)$，给定可逆的 $n \times n$ 的实系数矩阵 A 并把它视作是 $\mathbb{R}^n$ 到自身的线性变换。证明，A^*u 也是缓增的分布并且

$$\widehat{A^*u} = |\det(A)|^{-1} \left({}^tA^{-1}\right)^* \widehat{u}.$$

A7) 证明，Heaviside 函数 $H(x)$ 的 Fourier 变换为

$$\widehat{H}(\xi) = -i \cdot \mathrm{vp}\frac{1}{x} + \pi\delta_0.$$

A8) 考虑 $\mathbb{R}$ 上的分布 $u = \sum_{n=0}^{\infty} c_n\delta_n$。证明，$u \in \mathcal{S}'(\mathbb{R})$ 当且仅当存在常数 C 和 p，使得对所有的非负整数 n，我们有

$$|c_n| \leqslant C(1 + n^p).$$

A9) 试找出一切 $f \in L^1(\mathbb{R})$，使得 $f * f = f$。

A10)（如果你不确定的话）试验证如下的初等不等式：给定正整数 m，存在常数 C_1 和 C_2，使得对任意的 $\xi \neq 0$，我们有

$$C_1(1 + |\xi|^2)^m \leqslant \sum_{|\alpha| \leqslant m} |\xi^\alpha|^2 \leqslant C_2(1 + |\xi|^2)^m,$$

其中 $|\xi^\alpha| = |\xi_1^{\alpha_1}\xi_2^{\alpha_2}\cdots\xi_n^{\alpha_n}|$。这个不等式我们之后一直会引用。

A11)（某些分量的 Fourier 变换）证明，

$$\mathcal{F}_{x\to\xi} : \mathcal{S}(\mathbb{R}^n_t \times \mathbb{R}^m_x) \to \mathcal{S}(\mathbb{R}^n_t \times \mathbb{R}^m_x).$$

是连续的线性同构并进一步证明：

$$\mathcal{F}_{x\to\xi} : \mathcal{S}'(\mathbb{R}^m_t \times \mathbb{R}^n_x) \to \mathcal{S}'(\mathbb{R}^m_t \times \mathbb{R}^n_x),$$

是连续的线性同构。

A12)（重要）试构造函数局部可积的 $u \in H^{\frac{n}{2}}(\mathbb{R}^n)$，使得 $u \notin L^\infty(\mathbb{R}^n)$，即证明

$$H^{\frac{n}{2}}(\mathbb{R}^n) \not\hookrightarrow L^\infty(\mathbb{R}^n).$$

习题 B.（Fourier 变换的计算） 证明，前八个问题中的分布都是 $\mathbb{R}^1$ 上的缓增分布并计算它们的 Fourier 变换。

B1) $u = \dfrac{1}{1 + x^2}$;

B2) $u = e^{-|x|}$;

B3) $u = \mathbf{1}_{x\geqslant a}$，其中 $a \in \mathbb{R}$；

B4) $u = \dfrac{1}{x+i0}$；

B5) $u = P(x)\mathbf{1}_{x\geqslant 0}$，其中 $P(x)$ 为复系数多项式；

B6) 假设 $\alpha > -1$。证明，

$$\frac{1}{(i\xi+0)^{1+\alpha}} \overset{\mathcal{D}'}{=} \lim_{\varepsilon\to 0+}(i\xi+\varepsilon)^{-1-\alpha}$$

定义了 $\mathbb{R}$ 上的一个分布。

其中 $\alpha > -1$。请用 $\Gamma(z)$（Γ 函数）[21] 和 $\dfrac{1}{(i\xi+0)^{1+\alpha}}$ 来表示 $u = \mathrm{pf}\, x_+^\alpha$ 的 Fourier 变换；

B7) $u = \mathrm{pf}\, x_+^\alpha$，其中 $\alpha < -1$ 且 α 不是整数；

B8) $u = e^{-ix^2}$；

B9) 证明，$\mathbb{R}^n$ 上定义的函数

$$f(x_1,\cdots,x_n) = e^{-\sum_{i,j=1}^n A_{ij}x_ix_i}$$

为 Schwartz 函数并计算其 Fourier 变换，其中 A 为实（对称）正定矩阵。

习题 C.（Heisenberg 测不准原理）

C1) 对任意的 $f \in \mathcal{S}(\mathbb{R})$，证明如下的不等式

$$\left(\int_{\mathbb{R}} x^2|f(x)|^2dx\right)\left(\int_{\mathbb{R}}\xi^2|\widehat{f}(\xi)|^2\frac{d\xi}{2\pi}\right) \geqslant \left(\int_{\mathbb{R}} x\mathrm{Re}(\overline{f}f')dx\right)^2.$$

C2) 对任意的 $f \in \mathcal{S}(\mathbb{R})$，证明，

$$\left(\int_{\mathbb{R}} x^2|f(x)|^2dx\right)\left(\int_{\mathbb{R}}\xi^2|\widehat{f}(\xi)|^2\frac{d\xi}{2\pi}\right) \geqslant \frac{1}{4}\left(\int_{\mathbb{R}}|f|^2dx\right)^2.$$

C3) 对任意的 $f \in \mathcal{S}(\mathbb{R})$，证明，如果 $\|f\|_{L^2} = 1$，那么，对任意的 $(x_0,\xi_0) \in T^*\mathbb{R}^n = \mathbb{R}^n_x \times \mathbb{R}^n_\xi$，我们都有

$$\left(\int_{\mathbb{R}}(x-x_0)^2|f(x)|^2dx\right)\left(\int_{\mathbb{R}}(\xi-\xi_0)^2|\widehat{f}(\xi)|^2\frac{d\xi}{2\pi}\right) \geqslant \frac{1}{4}.$$

C4) 证明，当且仅当

$$(x_0,y_0) = \left(\int_{\mathbb{R}} x|f(x)|^2dx, \int_{\mathbb{R}}\xi|\widehat{f}(\xi)|^2\frac{d\xi}{2\pi}\right)$$

时，上式左边取得最小值。

[21] Γ 函数的定义为

$$\Gamma(z) = \int_0^{+\infty} e^{-t}t^{z-1}dt$$

C5) 试将 C3) 和 C4) 的结论推广到 $\mathbb{R}^n$ 上。

C6) 证明，上述不等式可以取到等号。（提示：利用 Gauss 函数）

C7) 当 $n \geqslant 2$ 时，证明，能找到函数列 $\{f_k(x)\}_{k\geqslant 1} \subset L^2(\mathbb{R}^n)$，使得对任意的 k，我们有 $\|f_k\|_{L^2} = 1$ 并且

$$\lim_{k\to\infty} \Big(\int_{\mathbb{R}} x_1^2 |f_k(x)|^2 dx \Big) \Big(\int_{\mathbb{R}} \xi_2^2 |\widehat{f_k}(\xi)|^2 \frac{d\xi}{2\pi} \Big) = 0.$$

习题 D. $H^s(\mathbb{R}^n)$ 的物理空间刻画

我们要求 Sobolev 指标 $0 < s < 1$。假定 $u \in L^2(\mathbb{R}^n)$，我们定义

$$\mathbf{I}(u) = \iint_{\mathbb{R}^n \times \mathbb{R}^n} \frac{|u(x) - u(y)|^2}{|x-y|^{n+2s}} dxdy.$$

我们注意到 $\mathbf{I}(u)$ 可以是无穷大。

D1) 证明如下的等式：

$$\begin{aligned} \mathbf{I}(u) &= \int_{\mathbb{R}^n} \frac{1}{|t|^{n+2s}} \Big(\int_{\mathbb{R}^n} |u(y+t) - u(y)|^2 dy \Big) dt \\ &= \int_{\mathbb{R}^n} |\widehat{u}(\xi)|^2 \Big(\int_{\mathbb{R}^n} \frac{|e^{it\cdot\xi} - 1|^2}{|t|^{n+2s}} \frac{dt}{(2\pi)^n} \Big) d\xi. \end{aligned}$$

D2) 令

$$F(\xi) = \int_{\mathbb{R}^n} \frac{|e^{it\cdot\xi} - 1|^2}{|t|^{n+2s}} \frac{dt}{(2\pi)^n}.$$

证明，F 是旋转对称的 $2s$-次的齐次函数。

D3) 证明，存在常数 $c_0 > 0$，使得

$$F(\xi) = c_0 |\xi|^{2s}.$$

D4) 假设 $0 < s < 1$。证明，对任意的 $u \in L^2(\mathbb{R}^n)$，如下等价

a) $u \in H^s(\mathbb{R}^n)$；

b) $\displaystyle\iint_{\mathbb{R}^n \times \mathbb{R}^n} \frac{|u(x) - u(y)|^2}{|x-y|^{n+2s}} dxdy < \infty$。

习题 E. 一维情形的 Poisson 求和公式 给定 $\lambda > 0$，我们定义分布 $\mathbb{R}^1$ 上的分布

$$W_\lambda = \sum_{k\in\mathbb{Z}} \delta_{\lambda k}.$$

E1) 证明，$W_\lambda \in \mathcal{S}'(\mathbb{R})$。

E2) 证明，W_λ 满足下面关于分布的方程组

$$\begin{cases}(\delta_\lambda - \delta_0) * u = 0\\ (e^{\frac{2\pi i x}{\lambda}} - 1)u = 0\end{cases}$$

E3) 证明，满足上述方程组的缓增分布一定是 W_λ 的倍数。

E4) 证明，存在常数 C_λ，使得 $\widehat{W_\lambda} = C_\lambda W_\lambda$ 并计算 C_λ。

E5) 证明 Poisson 求和公式：对每个 $\varphi \in \mathcal{S}(\mathbb{R})$，我们有

$$\sum_{k\in\mathbb{Z}} f(2\pi k) = \frac{1}{2\pi}\sum_{k\in\mathbb{Z}} \widehat{f}(k).$$

La vie n'est bonne que pour deux choses: découvrir les mathématiques et enseigner les mathématiques

— Siméon-Denis Poisson

---

71 Riesz 表示定理，Hilbert 空间中的正交投影，Sobolev 空间的对偶性，Sobolev 空间的限制定理（迹定理）

二零二零年十一月二日，星期一，晴

最后我们来证明 Sobolev 空间之间的对偶性，为此，我们需要证明关于完备的内积空间的 Riesz 表示定理。我们之后还会运用这个定理来求解偏微分方程。

引理 482 (向闭子空间的正交投影)**.** *给定完备的内积空间 $\big(H,(\cdot,\cdot)\big)$，$F\subset H$ 是闭线性子空间。那么，存在唯一的连续的线性映射*

$$\pi: H\to F,$$

使得

$$\|x-\pi(x)\|=\min_{y\in F}\|x-y\|.$$

*我们称 $\pi(x)$ 为 x 在 F 上的***正交投影***。特别地，我们还有 $x-\pi(x)\perp F$，即对任意的 $y\in F$，我们有 $(x-\pi(x),y)=0$。*

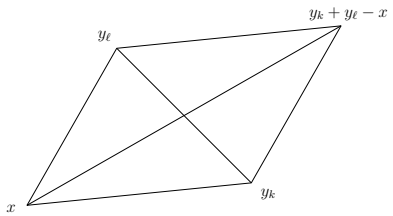
证明: 我们考虑

$$\inf_{y\in F}\|x-y\|.$$

根据下确界的定义，存在点列 $\{y_k\}_{k\geqslant 1}\subset F$，使得

$$\lim_{k\to\infty}\|x-y_k\|=\inf_{y\in F}\|x-y\|=I.$$

（我们经常把这样的一个序列称作是 $\inf\limits_{y\in F}\|x-y\|$ 的一个极小化子序列。）



我们回忆平面几何中的平行四边形等式：一个平行四边形的对角线的平方和等于四条边的平方和。现在考虑由 x,y_k,y_ℓ 和 $y_k+y_\ell-x$ 所组成的平行四边形，我们有

$$4|x-\frac{y_k+y_\ell}{2}|^2+|y_k-y_\ell|^2=2(|x-y_k|^2+|x-y_\ell|^2).$$

另外，我们还有

$$|x-\frac{y_k+y_\ell}{2}|\leqslant\frac{1}{2}(|x-y_k|+|x-y_\ell|).$$

由于，当 $k,\ell\to\infty$ 时，我们有

$$|x-y_k|\to I,|x-y_\ell|\to I.$$

所以，当 $k,\ell\to\infty$，平行四边形等式表明

$$|y_k-y_\ell|^2=2(|x-y_k|^2+|x-y_\ell|^2)-4|x-\frac{y_k+y_\ell}{2}|^2\to 0.$$

这说明，$\{y_k\}_{k\geqslant 1}$ 为 Cauchy 列，利用 F 是闭的，这证明存在 $y\in F$，使得 I 可以被实现。

为了说明，$x-\pi(x)=x-y\perp F$，我们利用变分的想法（请参考第一学期我们证明两点之间线段最短）。对任意的 $z\in F$，对任意的 $s\in\mathbb{R}$，按照定义，我们知道

$$\|x-y\|^2\leqslant\|x-y+sz\|^2.$$

所以，$s=0$ 是二次函数

$$f(s)=\|x-y+sz\|^2=\|x-y\|^2+2\mathrm{Re}\big((x-y,z)\big)s+\|z\|^2s^2$$

的最小值点。从而，$f'(0)=0$，这就说明

$$\mathrm{Re}\big((x-y,z)\big)=0.$$

类似地，如果把 s 换成 is，我们就有

$$\mathrm{Im}\big((x-y,z)\big)=0.$$

这就证明了对任意 $z\in F$，$x-y\perp z$，从而完成了证明。 □

注记. 我们注意到，F 中满足 $x-y\perp F$ 的向量是唯一的，就是 $\pi(x)$：实际上，如果 $y\in F$ 是另一个这样的向量，那么，

$$y-\pi(x)=\big(x-\pi(x)\big)-\big(x-y\big)$$

也与 F 垂直，从而，$y-\pi(x)\perp y-\pi(x)$，即

$$\|y-\pi(x)\|^2=1,$$

所以，$y=\pi(x)$。据此，对任意的 $x\in H$，我们都可以把它唯一地写成

$$x=x_{_F}+x_\perp,$$

其中，$x_{_F}\in F$，$x_\perp\perp F$。这被称作是 x 对于 F 的正交分解。

假设 H 是完备的内积空间，我们用 H^* 表示 H 上的连续线性泛函所构成的空间并把它称作是 H 在 Hilbert 空间的意义下的对偶：

$$H^*=\big\{\text{线性映射 }\ell:H\to\mathbb{C}\;\big|\;\ell\text{ 连续（有界）的}\big\}.$$

我们需要强调，在线性代数的意义下，所谓的对偶空间指的是 H 上的所有的线性函数所构成的线性空间，此处不同之处在于我们要求这些线性函数还是连续的。当 H 是有限维的内积空间时，这两个概念是一致的。

例子. 对任意的 $v \in H$，我们考虑如下的线性泛函：

$$\ell_v : H \to \mathbb{C}, \quad x \mapsto (x, v).$$

根据内积的性质，这是线性映射。另外，根据 *Cauchy-Schwarz* 不等式，我们有

$$|\ell(x)| = |(x, v)| \leqslant \|v\| \cdot \|x\|.$$

所有 ℓ_v 是有界的，从而连续，即 $\ell_v \in H^*$。

我们现在来证明，上述的例子给出了所有的 H^* 的元素。

定理 483 (Riesz 表示定理)**.** 完备内积空间 $\big(H, (-,-)\big)$ 上的连续线性泛函都可以用内积来实现，即对每个 $\ell \in H^*$，存在唯一的 $v \in H$，使得 $\ell = \ell_v$。换而言之，

$$H \longrightarrow H^*, \quad v \mapsto \ell_v,$$

是同构。

证明: 给定 ℓ，我们考虑

$$F = \mathrm{Ker}(\ell) = \big\{x \in H \big| \ell(x) = 0\big\}.$$

由于 ℓ 是连续的，所以，F 是闭的线性子空间。任选 $x \notin F$（这样的 x 总是存在的，除非 $\ell = 0$），那么，$x - \pi(x)$ 就给出了 F 的一个法向量。通过对这个法向量乘以一个系数，我们可以得到 $v \in F^\perp$，

$$\|v\|^2 = \ell(v) > 0.$$

我们声明 $\ell = \ell_v$。

为此，我们先说明 $\dim\left(F^\perp\right) = 1$。实际上，如果 $w \in F^\perp$，我们选取 λ，使得 $\ell(\lambda w) = \ell(v)$。此时，$\lambda w - v \in F$，从而，$\lambda w - v$ 与自身是垂直的，从而，$v = \lambda w$，这就表明 v 可以作为 $F^\perp$ 的基。

我们现在可以完成定理的证明：对任意的的 $x \in H$，利用正交投影（及其唯一性），我们有

$$x = \pi(x) + (x - \pi(x)) = x_{_F} + x_\perp = x_{_F} + \lambda v..$$

所以，

$$\ell_v(x) = (x_{_F}, v) + (\lambda v, v) = \lambda \ell(v) = \ell\left(x_{_F} + \lambda v\right).$$

证毕。 □

利用 Riesz 表示定理，我们来证明 $H^{-s}(\mathbb{R}^n)$ 与 $H^s(\mathbb{R}^n)$ 之间的对偶性。

固定一个 Sobolev 指标 s。首先，我们考虑分布 $u \in \mathcal{D}'(\mathbb{R}^n)$，它满足如下的性质：存在常数 C，使得对任意的试验函数 $\varphi \in \mathcal{D}(\mathbb{R}^n)$，我们都有

$$|\langle u, \varphi\rangle| \leqslant C\|\varphi\|_{H^s}.$$

由于 $\mathcal{D}(\mathbb{R}^n) \subset H^s(\mathbb{R}^n)$ 是稠密的，所以，分布 u 可以被看作是 $H^s(\mathbb{R}^n)$ 上的连续线性泛函。

反之，对任意的 $H^s(\mathbb{R}^n)$ 上的连续线性泛函 ℓ，存在常数 C，使得对任意的 $f \in H^s\mathbb{R}^n$，我们有

$$|\ell(f)| \leqslant C\|f\|_{H^s}.$$

特别地，对任意的试验函数 $\varphi \in \mathcal{D}(\mathbb{R}^n)$，我们都有

$$|\ell(\varphi)| \leqslant C\|\varphi\|_{H^s}.$$

由于

$$\|\varphi\|_{H^s} \leqslant C' N_{2s+2n+2}(\varphi).$$

所以，对任意的试验函数 $\varphi \in \mathcal{D}(\mathbb{R}^n)$，我们有

$$|\ell(\varphi)| \leqslant C'' N_{2s+2n+2}(\varphi).$$

所以，ℓ 可以被视作是一个 Schwartz 分布。

上面的推理表明，我们可以把 $(H^s(\mathbb{R}^n))^*$ 刻画为

$$(H^s(\mathbb{R}^n))^* = \left\{u \in \mathcal{S}'(\mathbb{R}^n) \middle| \text{ 存在常数 } C\text{，对任意 } \varphi \in \mathcal{D}(\mathbb{R}^n)\text{，有 } |\langle u, \varphi\rangle| \leqslant C\|\varphi\|_{H^s}\right\}.$$

命题 484 (H^{-s} 与 H^s 的对偶)**.** *对任意的 Sobolev 指标 $s \in \mathbb{R}$，我们有*

$$\begin{aligned} H^{-s}(\mathbb{R}^n) &= \left\{u \in \mathcal{D}'(\mathbb{R}^n) \middle| \text{ 存在常数 } C\text{，对任意 } \varphi \in \mathcal{D}(\mathbb{R}^n)\text{，有 } |\langle u, \varphi\rangle| \leqslant C\|\varphi\|_{H^s}\right\} \\ &= (H^s(\mathbb{R}^n))^*. \end{aligned}$$

证明： 后面的一个等号我们已经证明。

对任意的 $u \in H^{-s}(\mathbb{R}^n) \subset \mathcal{S}'(\mathbb{R}^n)$，对任意的试验函数 $\varphi \in \mathcal{D}(\mathbb{R}^n)$，我们有

$$\langle u, \varphi\rangle = \left\langle \mathcal{F}^{-1}\widehat{u}, \varphi\right\rangle = (2\pi)^{-n}\left\langle \widehat{u}, (\widehat{\varphi})^{\check{}}\right\rangle.$$

由于 $\widehat{u}$ 是局部可积的函数，所以上述的配对就是通常意义下的积分，从而

$$\begin{aligned} \langle u, \varphi\rangle &= (2\pi)^{-n}\int_{\mathbb{R}^n} \widehat{u}(\xi)\widehat{\varphi}(-\xi)d\xi \\ &= (2\pi)^{-n}\int_{\mathbb{R}^n} (1+|\xi|^2)^{-\frac{s}{2}}\widehat{u}(\xi)\cdot(1+|\xi|^2)^{\frac{s}{2}}\widehat{\varphi}(-\xi)d\xi. \end{aligned}$$

根据 Cauchy-Schwarz 不等式，我们就有

$$|\langle u, \varphi\rangle| \leqslant C\|(1+|\xi|^2)^{-\frac{s}{2}}\widehat{u}(\xi)\|_{L^2}\|(1+|\xi|^2)^{\frac{s}{2}}\widehat{\varphi}(\xi)\|_{L^2} = C\|u\|_{H^{-s}}\|\varphi\|_{H^s}.$$

这就证明了

$$H^{-s}(\mathbb{R}^n) \subset (H^s(\mathbb{R}^n))^*.$$

反之，对任意的分布 $u \in \mathcal{D}'(\mathbb{R}^n)$，如果存在常数 C，对任意 $\varphi \in \mathcal{D}(\mathbb{R}^n)$，都有

$$|\langle u, \varphi\rangle| \leqslant C\|\varphi\|_{H^s}.$$

那么，我们有

$$|\langle u, \varphi\rangle| \leqslant C' N_{s+n+1}(\varphi),$$

这说明，u 是缓增分布。特别地，我们得到

$$\widehat{v} = (1+|\xi|^2)^{-\frac{s}{2}}\widehat{u} \in \mathcal{S}'(\mathbb{R}^n),$$

其中缓增分布 v 由它的 Fourier 变换定义（在频率空间定义）。从而，

$$v = (1-\triangle)^{-\frac{s}{2}}u.$$

根据 Sobolev 空间在 Fourier 乘子下的映射性质，我们只需要证明 $v \in L^2(\mathbb{R}^n)$ 即可，根据 Planchrel 等式，这也等价于证明 $\widehat{v} \in L^2(\mathbb{R}^n)$。

对任意的 $\widehat{\varphi} \in \mathcal{S}(\mathbb{R}^n)$，我们有

$$\begin{aligned}
|\langle \widehat{v}, \widehat{\varphi}\rangle| &= \left|\langle \widehat{u}, \mathcal{F}\big((1-\triangle)^{-\frac{s}{2}}\varphi\big)\right| \\
&= (2\pi)^n \left|\langle u, (1-\triangle)^{-\frac{s}{2}}\varphi\rangle\right| \\
&\leqslant C\|(1-\triangle)^{-\frac{s}{2}}\varphi\|_{H^s} = C\|\varphi\|_L^2 \\
&= C'\|\widehat{\varphi}\|_L^2.
\end{aligned}$$

将 $\widehat{\varphi}$ 换为 $\psi(\xi)$，这说明分布所定义的映射

$$L_v : \mathcal{S}(\mathbb{R}^n) \to \mathbb{C}, \quad \psi \mapsto \langle \widehat{v}, \psi\rangle$$

可以延拓（根据稠密性）称为 $L^2(\mathbb{R}^n)$ 上的连续线性泛函。

$$\begin{array}{ccc}
\mathcal{S}(\mathbb{R}^n) & \xrightarrow{\iota} & H^s(\mathbb{R}^n) \\
& \searrow^{L_v} & \downarrow{\scriptstyle L_v} \\
& & \mathbb{C}
\end{array}$$

由 Riesz 表示定理，存在 $f(\xi) \in L^2(\mathbb{R}^n)$，使得对任意的 $\varphi(\xi) \in \mathcal{D}(\mathbb{R}^n)$，我们都有

$$\langle \widehat{v}, \varphi\rangle = \int_{\mathbb{R}^n} \varphi(\xi) f(\xi) d\xi.$$

所以，作为分布，我们有

$$\widehat{v} \overset{\mathcal{D}'}{=} f.$$

根据 $L^1_{\mathrm{loc}}(\mathbb{R}^n) \hookrightarrow \mathcal{D}'(\mathbb{R}^n)$ 是单射，我们知道 $\widehat{v}(\xi) = f(\xi)$ 几乎处处成立，命题得证。 □

注记. 根据上面的描述，从 $H^{-s}(\mathbb{R}^n)$ 到 $(H^s(\mathbb{R}^n))^*$ 的同构可以如下的构造：

$$\begin{aligned}
H^{-s}(\mathbb{R}^n) \longrightarrow (H^s(\mathbb{R}^n))^*, \quad u \mapsto \ell_u : H^s(\mathbb{R}^n) &\to \mathbb{C}, \\
v &\mapsto \int_{\mathbb{R}^n} u(x)v(x)dx.
\end{aligned}$$

Sobolev 空间的物理空间描述

当 Sobolev 指标 $0 < s < 1$ 时，假设 $u \in L^2(\mathbb{R}^n)$。我们可以证明 $u \in H^s(\mathbb{R}^n)$ 当且仅当

$$\iint_{\mathbb{R}^n \times \mathbb{R}^n} \frac{|u(x) - u(y)|^2}{|x - y|^{n+2s}} dxdy < \infty.$$

这就直接在物理空间刻画了 $H^s(\mathbb{R}^n)$。这个命题的证明请参考第四次作业。

Sobolev 空间的限制性定理

这一部分我们证明一个令人惊讶的结果。给定函数 $u \in H^s(\mathbb{R}^n)$，其中，$s > 0$。考虑 $\Sigma \subset \mathbb{R}^n$ 为一个超平面（余维数是 1 的线性子空间）。由于 $u \in L^2(\mathbb{R}^n)$，所以，u 只是几乎处处定义的。特别地，在一个零测集上改变 u 的取值不会改变 $u \in H^s(\mathbb{R}^n)$。所以，我们可以任意地改变 u 在 Σ 上的值（因为 Σ 是一个零测集）。然而，我们将证明，当 $s > \dfrac{1}{2}$ 时，我们可以把 u“限制”到 Σ 上来得到一个落在 $H^{s-\frac{1}{2}}(\mathbb{R}^n)$ 的函数。当然，我们也可以这么想象这个结果，当 s 足够大的时候（大于空间的维数的一半），此时，根据 Sobolev 嵌入定理，u 是连续函数，自然可以在 Σ 上限制。

为了把这个命题说清楚，我们不妨假设

$$\Sigma = \big\{(x_1, \cdots, x_{n-1}, x_n) \big| x_n = 0\big\}.$$

为了简单起见，我们用 $x' = (x_1, \cdots, x_{n-1})$ 表示前面 $n-1$ 个坐标；类似地，我们可以在频率空间上用 ξ' 表示 $(\xi_1, \cdots, \xi_{n-1})$。我们首先考虑限制映射

$$\mathrm{Res} : \mathcal{S}(\mathbb{R}^n) \longrightarrow \mathcal{S}(\mathbb{R}^{n-1}), \quad \varphi(x', x_n) \mapsto \varphi(x', 0).$$

当然，我们需要说明对任意的 $\varphi \in \mathcal{S}(\mathbb{R}^n)$，$\mathrm{Res}(\varphi) \in \mathcal{S}(\mathbb{R}^{n-1})$。这是明显的，因为对任意的多重指标 α 和 β，我们都有

$$\left| x'^{\alpha} \partial_{x'}^{\beta} \varphi(x', 0) \right| \leqslant \sup_{x \in \mathbb{R}^n} \left| x^{\alpha} \partial_x^{\beta} \varphi(x', x_n) \right| \leqslant N_{|\alpha|+|\beta|}(\varphi).$$

通过复合，

$$\begin{array}{ccc} \mathcal{S}(\mathbb{R}^n) & \xrightarrow{\mathrm{Res}} & \mathcal{S}(\mathbb{R}^{n-1}) \\ & \searrow^{\mathrm{Res}} & \big\downarrow \iota \\ & & H^{s-\frac{1}{2}}(\mathbb{R}^{n-1}) \end{array}$$

我们就得到

$$\mathrm{Res} : \mathcal{S}(\mathbb{R}^n) \longrightarrow H^{s-\frac{1}{2}}(\mathbb{R}^{n-1}), \quad \varphi(x', x_n) \mapsto \varphi(x', 0).$$

其中，我们假设了 $s > \dfrac{1}{2}$。如果我们可以证明，存在 $C > 0$，使得对任意的 $\varphi \in \mathcal{S}(\mathbb{R}^n)$，如下的不等式成立（连续性）：

$$\|\mathrm{Res}\,(\varphi)\|_{H^{s-\frac{1}{2}}(\mathbb{R}^{n-1})} \leqslant C \|\varphi\|_{H^s(\mathbb{R}^n)},$$

那么，根据 $\mathcal{S}(\mathbb{R}^n)$ 在 $H^s(\mathbb{R}^n)$ 中的稠密性，Res 就可以延拓成为连续线性映射：

$$\mathrm{Res} : H^s(\mathbb{R}^n) \longrightarrow H^{s-\frac{1}{2}}(\mathbb{R}^{n-1}).$$

这就给出了 Res 的含义。

定理 485. 假设 $n \geqslant 1$ 并且 *Sobolev* 指标 $s > \dfrac{1}{2}$。限制映射

$$\mathrm{Res} : \mathcal{S}(\mathbb{R}^n) \longrightarrow H^{s-\frac{1}{2}}(\mathbb{R}^{n-1})$$

可以唯一地延拓成连续线性映射

$$\mathrm{Res} : H^s(\mathbb{R}^n) \longrightarrow H^{s-\frac{1}{2}}(\mathbb{R}^{n-1})$$

使得如下的图表是交换的：

$$\begin{array}{ccc} \mathcal{S}(\mathbb{R}^n) & \xrightarrow{\mathrm{Res}} & \mathcal{S}(\mathbb{R}^{n-1}) \\ \iota\downarrow & & \downarrow\iota \\ H^s(\mathbb{R}^n) & \xrightarrow{\mathrm{Res}} & H^{s-\frac{1}{2}}(\mathbb{R}^{n-1}) \end{array}$$

特别地，存在常数 $C > 0$，对任意的 $u \in H^s(\mathbb{R}^n)$，我们有

$$\|\mathrm{Res}(u)\|_{H^{s-\frac{1}{2}}(\mathbb{R}^{n-1})} \leqslant C\|u\|_{H^s(\mathbb{R}^n)}.$$

注记. 习惯上，对于 $u \in H^s(\mathbb{R}^n)$，我们把 $\mathrm{Res}(u)$ 写成 $u(x', 0)$。

证明: 根据之前的讨论，我们只要对 $\varphi \in \mathcal{S}(\mathbb{R}^n)$，证明

$$\|\varphi(x', 0)\|_{H^{s-\frac{1}{2}}(\mathbb{R}^{n-1})} \leqslant C\|\varphi\|_{H^s(\mathbb{R}^n)}$$

即可，其中 C 是一个待定的常数（不依赖于 φ 的选取）。

我们用 $\mathcal{F}'$ 表示仅对前面 $n-1$ 个坐标的 Fourier 变换或者表示在 $\mathbb{R}^{n-1}$ 上的 Fourier 变换，用 $\widehat{\varphi}$ 表示对 n 元函数的 Fourier 变换，那么，我们需要控制如下的积分

$$\int_{\mathbb{R}^{n-1}} \left(1 + |\xi'|^2\right)^{s-\frac{1}{2}} \left|\mathcal{F}'(\varphi(\cdot, 0))(\xi')\right|^2 d\xi'.$$

为此，我们先用 n 维的 Fourier 变换来表示 $\mathcal{F}'$：

$$\begin{aligned} \mathcal{F}'(\varphi(\cdot, 0))(\xi') &= \int_{\mathbb{R}^{n-1}} \varphi(x', 0) e^{-ix'\cdot\xi'} dx' \\ &= \frac{1}{2\pi} \int_{\mathbb{R}} \widehat{\varphi}(\xi', \xi_n) d\xi_n. \end{aligned}$$

所以，

$$\begin{aligned} \left|\mathcal{F}'(\varphi(\cdot, 0))(\xi')\right|^2 &\leqslant \left(\frac{1}{2\pi}\right)^2 \left|\int_{\mathbb{R}} \widehat{\varphi}(\xi', \xi_n) d\xi_n\right|^2 \\ &= \left(\frac{1}{2\pi}\right)^2 \left|\int_{\mathbb{R}} \frac{1}{(1 + |\xi|^2)^{\frac{s}{2}}} \cdot (1 + |\xi|^2)^{\frac{s}{2}} \widehat{\varphi}(\xi', \xi_n) d\xi_n\right|^2 \end{aligned}$$

根据 Cauchy-Schwarz 不等式，我们得到

$$\left|\mathcal{F}'(\varphi(\cdot,0))(\xi')\right|^2 \leqslant \left(\frac{1}{2\pi}\right)^2 \left(\int_{\mathbb{R}} \frac{1}{(1+|\xi|^2)^s} d\xi_n\right) \left(\int_{\mathbb{R}} (1+|\xi|^2)^s \left|\widehat{\varphi}(\xi',\xi_n)\right|^2 d\xi_n\right).$$

在上面计算中，$\xi = (\xi', \xi_n)$ 中的 ξ' 是固定的。从而，

$$\begin{aligned}
\int_{\mathbb{R}} \frac{1}{(1+|\xi|^2)^s} d\xi_n &= \int_{\mathbb{R}} \frac{1}{(1+|\xi'|^2+\xi_n^2)^s} d\xi_n \\
&= \frac{1}{(1+|\xi'|^2)^{s-\frac{1}{2}}} \int_{\mathbb{R}} \frac{1}{\left(1+\left(\frac{\xi_n}{\sqrt{1+|\xi'|^2}}\right)^2\right)^s} \cdot \frac{d\xi_n}{\sqrt{1+|\xi'|^2}} \\
&= C_s \frac{1}{(1+|\xi'|^2)^{s-\frac{1}{2}}}.
\end{aligned}$$

其中，利用 $s > \dfrac{1}{2}$，我们知道右边的积分是有限的，我们用 C_s 表示积分

$$C_s = \int_{\mathbb{R}} \frac{dy}{(1+y^2)^s}.$$

综合上述，我们得到

$$\left|\mathcal{F}'(\varphi(\cdot,0))(\xi')\right|^2 \leqslant \frac{C_s}{4\pi^2} \frac{1}{(1+|\xi'|^2)^{s-\frac{1}{2}}} \left(\int_{\mathbb{R}} (1+|\xi|^2)^s \left|\widehat{\varphi}(\xi',\xi_n)\right|^2 d\xi_n\right).$$

据此，我们得到

$$\int_{\mathbb{R}^{n-1}} \left(1+|\xi'|^2\right)^{s-\frac{1}{2}} \left|\mathcal{F}'(\varphi(\cdot,0))(\xi')\right|^2 d\xi' \leqslant \frac{C_s}{4\pi^2} \int_{\mathbb{R}^n} (1+|\xi|^2)^s \left|\widehat{\varphi}(\xi',\xi_n)\right|^2 d\xi' d\xi_n.$$

这等价于

$$\|\varphi(x',0)\|^2_{H^{s-\frac{1}{2}}(\mathbb{R}^{n-1})} \leqslant \frac{C_s}{4\pi^2} \|\varphi\|^2_{H^s(\mathbb{R}^n)}.$$

命题得证。 □

注记. 我们通常说一个 H^s 的函数限制到余 1 维的子流形上会丢失 $\dfrac{1}{2}$ 个导数。

注记. 上面的证明实际上表明了限制映射

$$\mathrm{Res} : H^s(\mathbb{R}^n) \twoheadrightarrow H^{s-\frac{1}{2}}(\mathbb{R}^{n-1})$$

是满射。

根据

$$\mathcal{F}'(\varphi(\cdot,0))(\xi') = \frac{1}{2\pi} \int_{\mathbb{R}} \widehat{\varphi}(\xi',\xi_n) d\xi_n,$$

的提示，对任意的 $u \in H^{s-\frac{1}{2}}(\mathbb{R}^n)$，我们构造一个 $U \in H^s(\mathbb{R}^n)$，使得

$$\mathcal{F}'(u)(\xi') = \frac{1}{2\pi} \int_{\mathbb{R}} \widehat{U}(\xi',\xi_n) d\xi_n.$$

实际上，我们定义

$$\widehat{U}(\xi',\xi_n)=\widehat{u}(\xi')\frac{2\pi(1+|\xi'|^2)^{\frac{s}{2}-\frac{1}{2}}}{C_s(1+|\xi|^2)^{\frac{s}{2}}}.$$

据此，我们首先有 $U\in H^s(\mathbb{R}^n)$，这因为

$$\begin{aligned}\|U\|^2_{H^s(\mathbb{R}^n)}&=\frac{4\pi^2}{C_s^2}\int_{\mathbb{R}^n}(1+|\xi|^2)^s\left|\widehat{u}(\xi')\right|^2\frac{2\pi(1+|\xi'|^2)^{2s-1}}{C_s^2(1+|\xi|^2)^{2s}}d\xi\\&=\frac{4\pi^2}{C_s^2}\int_{\mathbb{R}^{n-1}}\left|\widehat{u}(\xi')\right|^2(1+|\xi'|^2)^{2s-1}\int_{\mathbb{R}}\frac{d\xi_n}{(1+|\xi|^2)^s}d\xi'\\&=\frac{4\pi^2}{C_s}\int_{\mathbb{R}^{n-1}}\left|\widehat{u}(\xi')\right|^2(1+|\xi'|^2)^{s-\frac{1}{2}}d\xi'\\&=\frac{4\pi^2}{C_s}\|u\|^2_{H^{n-1}}<\infty.\end{aligned}$$

所以，$U\in H^s(\mathbb{R}^n)$。

现在证明 $U(x',0)=u(x')$。实际上，我们只要说明

$$\begin{aligned}\frac{1}{2\pi}\int_{\mathbb{R}}\widehat{U}(\xi',\xi_n)d\xi_n&=\frac{1}{2\pi}\int_{\mathbb{R}}\widehat{u}(\xi')\frac{2\pi(1+|\xi'|^2)^{\frac{s}{2}-\frac{1}{2}}}{C_s(1+|\xi|^2)^{\frac{s}{2}}}\xi_n\\&=\frac{1}{C_s}\widehat{u}(\xi')(1+|\xi'|^2)^{\frac{s}{2}-\frac{1}{2}}\int_{\mathbb{R}}\frac{1}{C_s(1+|\xi|^2)^{\frac{s}{2}}}\xi_n\\&=\widehat{u}(\xi').\end{aligned}$$

这就完成了证明。

72 有界区域上的 Sobolev 空间，1 维区间上的 Sobolev 空间的经典表述，椭圆边值问题，Poincaré 不等式

二零二零年十一月二日，星期一，晴

椭圆边值问题的研究

假定 $\Omega \subset \mathbb{R}^n$ 是有界光滑的开区域，我们要在这个区域上解如下的位势方程：

$$\begin{cases} -\triangle u = f, \\ u\big|_{\partial\Omega} = g. \end{cases}$$

关于这个方程在经典引力理论或者电磁理论的描述，同学们可以在物理书里找到详细的叙述，我们专注于它的数学理论。

在应用中，我们要找的函数 u 属于所谓的能量类中，即 $u \in H^1(\Omega)$。根据限制定理（的类比），我们要求边界值 $g \in H^{\frac{1}{2}}(\partial\Omega)$。函数 f 可以落在所谓的 $H^{-1}(\Omega)$ 中，然而在大多应用中 $f \in L^2(\Omega)$。为了做到这些，我们将仿照全空间 $\mathbb{R}^n$ 的情形，构造上述所提到的空间。

区域上的 Sobolev 空间

定义 486 (整数指标的 Sobolev 空间)**.** 给定开区域 $\Omega \subset \mathbb{R}^n$，给定非负整数 $k \geqslant 0$，我们定义 *Sobolev* 空间 $H^k(\Omega)$：

$$H^k(\Omega) = \left\{ u \in L^2(\Omega) \middle| \text{对任意的多重指标 } \alpha, \text{ 其中 } |\alpha| \leqslant k, \text{ 我们有} \partial^\alpha u \in L^2(\Omega) \right\}.$$

我们在 $H^k(\Omega)$ 定义如下的范数：

$$\|u\|_{H^k} = \left(\sum_{|\alpha| \leqslant k} \|\partial^\alpha u\|_{L^2} \right)^{\frac{1}{2}}.$$

这个范数与如下的内积是相容的：

$$\begin{aligned} (u, v)_{H^k} &= \sum_{|\alpha| \leqslant k} (\partial^\alpha u, \partial^\alpha v)_{L^2} \\ &= \sum_{|\alpha| \leqslant k} \int_\Omega \partial^\alpha u(x) \cdot \overline{\partial^\alpha v(x)} dx. \end{aligned}$$

定理 487. 内积空间 $\left(H^k(\Omega), (\cdot,\cdot)_{H^k}\right)$ 是完备的。

证明： 假设 $\{u_p\}_{p \geqslant 1} \subset H^k(\Omega)$ 是 Cauchy 列。根据定义，对任意的 $\varepsilon > 0$，存在 N，使得当 $p, q > N$ 时，

$$\|u\|_{H^k}^2 = \sum_{|\alpha| \leqslant k} \|\partial^\alpha u\|_{L^2} < \varepsilon.$$

所以，对于每个多重指标 $|\alpha| \leqslant k$，函数列 $\{\partial^\alpha u_p\}_{p\geqslant 1} \subset L^2(\Omega)$ 也是 Cauchy 列。根据 L^2 的完备性，存在 $u^{(\alpha)} \in L^2(\Omega)$，使得

$$\partial^\alpha u_p \xrightarrow{L^2} u^{(\alpha)}.$$

我们把 $u^{(0)}$ 记作 u，那么，

$$u_p \xrightarrow{L^2} u \quad \Rightarrow \quad u_p \xrightarrow{\mathcal{D}'} u.$$

所以，对任意的多重指标 α，其中 $|\alpha| \leqslant k$，我们有

$$\partial^\alpha u_p \xrightarrow{\mathcal{D}'} \partial^\alpha u.$$

另外，

$$\partial^\alpha u_p \xrightarrow{L^2} u^\alpha \quad \Rightarrow \quad \partial^\alpha u_p \xrightarrow{\mathcal{D}'} \partial^\alpha u.$$

所以，作为分布，我们有

$$\partial^\alpha u \overset{\mathcal{D}'}{=} u^\alpha.$$

根据局部可积函数到分布嵌入的唯一性，我们知道 $\partial^\alpha u$ 是平方可积的函数，从而 $u \in H^k(\Omega)$。特别地，我们还有

$$\partial^\alpha u_n \xrightarrow{L^2} \partial^\alpha u.$$

这就完成了证明。 □

给定开区域 $\Omega \subset \mathbb{R}^n$，$k \geqslant 0$，我们明显有 $C_0^\infty(\Omega) \subset H^k(\Omega)$。我们考虑它们在 $H^k(\Omega)$ 中的闭包：

定义 488. 我们把 $H_0^k(\Omega)$ 定义为 $C_0^\infty(\Omega)$ 在 $H^k(\Omega)$ 中（用 $\|\cdot\|_{H^k}$ 范数）的闭包，即 $u \in H^1_(\mathbb{R}^n)$ 当且仅当存在序列 $\{\varphi_p\}_{p\geqslant 1} \subset C_0^\infty(\Omega)$，使得

$$\lim_{p\to\infty} \|\varphi_p - u\|_{H^k} = 0.$$

注记. 按照定义，$C_0^\infty(\Omega)$ 在 $H_0^k(\Omega)$ 中是稠密的，其中我们用的度量是由 H^k-范数诱导的。

在数学物理方程中，$H_0^1(\Omega)$ 这个空间的应用最为广泛。由于 $C_0^\infty(\Omega)$ 中的函数在 $\partial\Omega$ 上都消失，所以，直观上我们可以将 $H_0^1(\Omega)$ 解读成 $H^1(\Omega)$ 中在边界 $\partial\Omega$ 上限制为零的那些函数。实际上，我们后面会证明存在正合序列：

$$0 \to H_0^1(\Omega) \to H^1(\Omega) \to H^{\frac{1}{2}}(\partial\Omega) \to 0.$$

注记. 假定 $\Omega_1 \subset \Omega_2$ 是两个开区域，那么，*Sobolev* 空间具有如下的函子性质：

1) 限制映射

$$\mathrm{Res} : H^k(\Omega_2) \to H^k(\Omega_1), \quad u \mapsto u\big|_{\Omega_1},$$

是连续线性映射。

实际上，对任意的 $u \in H^k(\Omega_2)$，对任意的 α，其中 $\alpha| \leqslant k$，那么

$$\left\|\partial^\alpha \left(u\big|_{\Omega_1}\right)\right\|_{L^2(\Omega_1)} \leqslant \|\partial^\alpha u\|_{L^2(\Omega_2)}.$$

2) 用零做延拓，我们明显有

$$\mathrm{Ext} : C_0^\infty(\Omega_1) \to C_0^\infty(\Omega_2), \quad \varphi \mapsto \varphi \cdot \mathbf{1}_{\Omega_1}.$$

所以，利用稠密性，

$$\begin{array}{ccc} C_0^\infty(\Omega_1) & \xrightarrow{\mathrm{Ext}} & C_0^\infty(\Omega_2) \\ \iota\big\downarrow & & \big\downarrow\iota \\ H_0^k(\Omega_1) & \dashrightarrow & H_0^k(\Omega_2) \end{array}$$

我们得到连续的嵌入映射:

$$\mathrm{Ext} : H_0^1(\Omega_1) \to H_0^1(\Omega_2).$$

我们来研究两个特殊（具有启发性）的例子:

例子. *1)* (全空间的情况)

我们已经证明过 $C_0^\infty(\mathbb{R}^n) \subset H^k(\mathbb{R}^n)$ 是稠密的，其中，$k \in \mathbb{Z}_{\geqslant 0}$，所以，

$$H^k(\mathbb{R}^n) = H_0^k(\mathbb{R}^n).$$

2) (*1* 维有限区间的情形的 *Sobolev* 不等式)

假定 $\Omega = I = (-a, a) \subset \mathbb{R}$，其中 a 是正实数。与全空间情况类似，(因为 $1 = s > \dfrac{n}{2} = \dfrac{1}{2}$)
我们有所谓的 *Sobolev* 不等式（嵌入定理）: 对任意的 $f \in H^1(I)$，我们有

$$|f(x)| \leqslant \frac{1}{\sqrt{2a}}\|f\|_{L^2(I)} + \sqrt{2a}\|f'\|_{L^2(I)}.$$

如果用 $|I|$ 表示区间的长度，一个更几何的表达为

$$|f(x)| \leqslant \frac{1}{\sqrt{|I|}}\|f\|_{L^2(I)} + \sqrt{|I|}\|f'\|_{L^2(I)}.$$

实际上，给定 $b \in I$，我们有

$$f(x) = f(b) + \int_b^x f'(y)dy.$$

其中，b 待定。我们注意到这个表达式表明 f 为连续函数，因为 f' 是局部可积的。

根据抽屉原理，我们总是可以选取 $b \in (-a, a)$，使得

$$\sqrt{2a}|f(b)| \leqslant \|f\|_{L^2(I)},$$

利用 *Cauchy-Schwarz* 不等式，我们有

$$\begin{aligned} |f(x)| &\leqslant \frac{1}{\sqrt{2a}}\|f\|_{L^2(I)} + \left|\int_b^x 1 \cdot f'(y)dy\right| \\ &\leqslant \frac{1}{\sqrt{2a}}\|f\|_{L^2(I)} + \sqrt{2a}\|f'\|_{L^2(I)}. \end{aligned}$$

这就给出了证明。

3) (*1* 维有限区间的情形的 H_0^1 的刻画)

对任意的 $f \in H^1(I)$，任意选定 $b \in I$，那么，公式

$$f(x) = f(b) + \int_b^x f'(y)dy$$

表明极限

$$f\left(a^-\right) = \lim_{x \to a} f(x), \quad f\left((-a)^+\right) = \lim_{x \to -a} f(x)$$

存在，事实上，f' 是平方可积的，所以，是可积的，从而上面的极限存在。

对于 $H_0^1(I)$，我们有如下的刻画：

$$H_0^1\big(I\big) = \big\{f \in H^1\big(I\big) \big| f\left(a^-\right) = f\left((-a)^+\right) = 0\big\}.$$

首先说明：

$$H_0^1\big(I\big) \subset \big\{f \in H^1\big(I\big) \big| f\left(a^-\right) = f\left((-a)^+\right) = 0\big\}.$$

利用定义，存在 $\varphi_k \in C_0^\infty(I)$，使得在 H^s 的意义下，$\varphi_k \overset{H^s}{\to} f$。*Sobolev* 嵌入，我们有

$$|\varphi(x) - f(x)| \leqslant C\|\varphi_k(x) - f_k(x)\|_{H^s} \to 0.$$

由于 $\varphi(a) = 0$，上面的估计对 x 是一致的，所以当 $x \to a^-$ 时，我们就有 $f\left(a^-\right) = 0$。

反之，对任意的 $f \in H^1\big(I\big)$，假设 $f\left(a^-\right) = f\left((-a)^+\right) = 0$，我们定义全空间 $\mathbb{R}$ 上的函数 g (我们要在全空间上用卷积运算来构造光滑的逼近)：

$$g(x) = \begin{cases} f(x), & |x| < a; \\ 0, & |x| \geqslant a. \end{cases}$$

这显然是一个连续函数。根据分布版本的 *Stokes* 公式，它在分布意义下的导数为

$$g'(x) \overset{\mathcal{D}'}{=} \mathbf{1}_{[-a,a]}(x) \cdot f'(x).$$

所以，$g \in H^1(\mathbb{R})$。

我们现在将 g 的支集变小：选取 $\lambda < 1$，我们考虑

$$g_\lambda(x) = g\left(\frac{x}{\lambda}\right).$$

很明显，这是一个 $H^1(\mathbb{R})$ 中的函数并且 $\mathrm{supp}(g_\lambda) \subset (-a, a)$。特别地（容易验证），我们有

$$g_\lambda \xrightarrow{H^1} g.$$

对每个固定的 λ，我们可以进行光滑化：

$$\chi_\varepsilon * g_\lambda \xrightarrow{H^1} g_\lambda.$$

上面的极限是在 $\varepsilon \to \infty$ 时取的：由于 $g_\lambda, \partial_k g_\lambda \in L^2$，所以，

$$\chi_\varepsilon * g_\lambda \xrightarrow{L^2} g_\lambda,$$
$$\partial_k (\chi_\varepsilon * g_\lambda) = \chi_\varepsilon * (\partial_k g_\lambda) \xrightarrow{L^2} g_\lambda.$$

这里，我们可以选取 $\varepsilon < a(1-\lambda)$，所以，$\operatorname{supp}(\chi_\varepsilon * g_\lambda) \subset \big((-a,a)\big)$。所以，通过适当地选取 ε 和 λ，我们就有

$$\chi_\varepsilon * g_\lambda \xrightarrow{H^1} f.$$

这就证明了命题。

利用对偶性，我们定义负指数的 Sobolev 空间：

定义 489. 对于整数 $k \geqslant 0$，我们定义

$$H^{-k}(\Omega) = \big\{u \in \mathcal{D}'(\Omega) \big| \text{存在常数 } C>0, \text{使得对任意的 } \varphi \in \mathcal{D}(\Omega), \text{都有 } |\langle u, \varphi\rangle| \leqslant C\|\varphi\|_{H^k}\big\}.$$

对于 $u \in H^{-k}(\Omega)$，我们假设 $\mathscr{C}$ 是使得上面的不等式成立的所有的常数 C 的集合，我们定义范数

$$\|u\|_{H^{-k}(\Omega)} = \inf(\mathscr{C}).$$

由于 $C_0^\infty(\Omega)$ 在 $H_0^1(\Omega)$ 中稠密，

$$\begin{array}{ccc} C_0^\infty(\Omega) & \xrightarrow{\iota} & H_0^1(\Omega) \\ & \searrow^{u} & \downarrow{\scriptstyle u} \\ & & \mathbb{C} \end{array}$$

线性映射 $\varphi \mapsto \langle u, \varphi\rangle$ 可以唯一地扩张成连续线性映射

$$u : H_0^k(\Omega) \longrightarrow \mathbb{C}.$$

反之，对任意给定的连续线性泛函

$$T : H_0^k(\Omega) \longrightarrow \mathbb{C},$$

我们要证明：存在唯一的分布 $u \in H^{-k}(\Omega)$，使得对每个 $\varphi \in \mathcal{D}(\Omega)$，有

$$\langle u, \varphi\rangle = T(\varphi).$$

实际上，先证明 T 可以被视作是一个分布：对于支集在 $K \subset \Omega$ 中的试验函数 φ，我们有

$$\begin{aligned} |\langle u, \varphi\rangle| &\leqslant C\|\varphi\|_{H^k} = C \sum_{|\alpha| \leqslant k} \|\partial^\alpha \varphi\|_{L^2(K)} \\ &\leqslant C|K|^{\frac{1}{2}} \sum_{|\alpha| \leqslant k} \|\partial^\alpha \varphi\|_{L^\infty(K)} \\ &\leqslant C|K|^{\frac{1}{2}} \sup_{|\alpha| \leqslant k} \|\partial^\alpha \varphi\|_{L^\infty(K)}. \end{aligned}$$

按照定义，T 是阶不超过 k 的分布。按照定义，我们自然有 $T \in H^{-k}(\Omega)$。

综上所述，我们实际上得到了

$$H^{-k}(\Omega) = \left(H_0^1(\Omega)\right)^*.$$

对于空间 $H_0^1(\Omega)$ 中的函数，最重要的不等式是如下的 Poincaré 不等式：

定理 490 (Poincaré 不等式)**.** *如果 $\Omega \subset \mathbb{R}^n$ 是有界区域，那么，存在常数仅依赖于 Ω 的常数 C，使得对任意的 $u \in H_0^1(\Omega)$，如下的不等式成立：*

$$\int_\Omega |u|^2 dx \leqslant C \int_\Omega |\nabla u|^2 dx.$$

证明： 我们选取 $R > 0$，使得

$$\Omega \subset \left\{x = (x_1, \cdots, x_n) \middle| |x_n| \leqslant R\right\}.$$

我们首先证明对于 $\varphi \in C_0^\infty(\Omega)$，Poincaré 不等式成立。利用 Newton-Leibniz 公式，我们有

$$\begin{aligned}
|\varphi(x)|^2 = |\varphi(x', x_n)|^2 &= \left| \int_{-R}^{x_n} \partial_{x_n} \varphi(x', \tau) d\tau \right|^2 \\
&\leqslant \left| \int_{-R}^{R} |\partial_{x_n} \varphi(x', \tau)| \cdot 1 d\tau \right|^2 \\
&\overset{Cauchy-Schwarz}{\leqslant} 2R \int_{-\infty}^{\infty} |\partial_{x_n} \varphi(x', \tau)|^2 d\tau.
\end{aligned}$$

其中，我们用 x' 表示 $(x_1, \cdots, x_{n-1})$。上面的不等式的右边是不依赖于 x_n 的。对 x_n 积分在 $[-R, R]$ 上积分，我们得到

$$\int_{\mathbb{R}} |\varphi(x', x_n)|^2 dx_n \leqslant 4R^2 \int_{-\infty}^{\infty} |\partial_{x_n} \varphi(x', \tau)|^2 d\tau.$$

再对 x' 积分，我们就得到

$$\int_{\mathbb{R}} |\varphi(x)|^2 dx \leqslant 4R^2 \int_{\mathbb{R}^{n-1}} \int_{-\infty}^{\infty} |\partial_{x_n} \varphi(x', \tau)|^2 d\tau dx' = 4R^2 \int_{R^n} |\nabla \varphi(x)|^2 dx$$

为了说明不等式对任意的 $u \in H_0^1(\Omega)$ 成立，我们选取一列 $\{\varphi_k\}_{k \geqslant 1} \subset H_0^1(\mathbb{R}^n)$，使得 $\varphi_k \overset{H_0^1(\Omega)}{\longrightarrow} u$，翻译成数量的表达，我们有

$$\lim_{k \to \infty} \|\varphi_k - u\|_{L^2} + \|\nabla \varphi_k - \nabla u\|_{L^2} = 0.$$

所以，

$$\|\varphi_k\|_{L^2} \leqslant \|\nabla \varphi_k\|_{L^2} \quad \Rightarrow \quad \|u\|_{L^2} - \|u - \varphi_k\|_{L^2} \leqslant \|u\|_{L^2} + \|\nabla \varphi_k - \nabla u\|_{L^2}.$$

令 $k \to \infty$ 即可。 □

注记. 这个证明不等式的方法是典型的：一般而言，很多函数空间中的函数可以用光滑函数逼近。我们先对光滑函数证明不等式，此时，我们可以对光滑函数进行求导等预算；然后，我们再用光滑函数来逼近一般的函数。

注记. *Poincaré* 不等式的证明过程表明，只要区域 Ω 可以被两个平行的超平面夹住，那么，不等式仍然成立。

推论 491. 我们在 $H_0^1(\Omega)$ 上规定新的内积：

$$\begin{aligned}(u,v)_{H_0^1} &= (\nabla u, \nabla v)_{L^2} \\ &= \sum_{k=1}^{n} \int_\Omega \partial_k u(x) \cdot \overline{\partial_k v(x)} dx.\end{aligned}$$

那么，这个新的内积所定义的范数 $\|\cdot\|_{H_0^1}$ 与 $\|\cdot\|_{H^1}$ 是等价的，即存在非负常数 C_1 和 C_2，使得对任意的 $u \in H_0^1(\Omega)$，我们有

$$C_1\|u\|_{H_0^1} \leqslant \|u\|_{H^1} \leqslant C_2\|u\|_{H_0^1}.$$

证明: 按照定义，我们就有 $\|u\|_{H_0^1} \leqslant \|u\|_{H^1}$，所以，选取 $C_1 = 1$ 即可；根据 Poincaré 不等式，我们可以选取 $C_2 = C + 1$。 □

注记. 在之后的很多场合下，当我们提及 $H_0^1(\Omega)$，我们指的是配备有 $(\cdot,\cdot)_{H_0^1}$ 作为内积的（完备）内积空间。

72.1 期中测验

清华大学 19-20 秋季学期，数学分析三，期中小测试

请在清华大学考试专用纸上**正反面**用钢笔，签字笔或者圆珠笔书写，并注明自己的姓名和年级。除定理公式所涉及的人名之外，请使用**中文**。本次考试为开卷考试，可以携带一切手写的资料以便查阅。考试中不能使用任何印刷，打印或复印的材料，不能使用手机，违反者试卷视为零分，并按照作弊遵循学校规定处理。

考试时间为 11 月 9 日下午第四大节课，以铃声为准。考试结束请交答题纸，草稿纸和试题请带走。

本次考试中，所有的测度都是 Lebesgue 测度。

对任意的试验函数 $\varphi \in \mathcal{D}(\mathbb{R})$，我们定义如下的 $\langle u_0, \varphi \rangle$ 以及线性映射

$$\mathcal{D}(\mathbb{R}) \to \mathbb{C}, \quad \varphi \mapsto \langle u_0, \varphi \rangle = \int_{\mathbb{R}} \frac{\varphi(x) - \varphi(-x) - 2x\varphi'(0)}{2x^3} dx.$$

1) 证明，$u_0 \in \mathcal{S}'(\mathbb{R})$。

2) 找出满足方程

$$x^3 u = 1$$

的分布 $u \in \mathcal{S}'(\mathbb{R})$。

3) 证明，$(u_0)^{\check{}} = -u_0$。

4) 证明，u_0 是次数为 -3 的齐次分布。

5) 计算 u_0 的 Fourier 变换 $\widehat{u_0}$。

73 Dirichlet 问题的初步研究：存在性与解算子，用全空间上的连续函数逼近上半空间上的 H^1 函数，H^1 函数从半空间到全空间的扩张

二零二零年十一月十一日，星期三，晴

在上次课上，我们引入了 Sobolev 空间 $H_0^1(\Omega)$。在这个空间上，我们有两个内积：对任意的 $u, v \in H_0^1(\Omega)$，我们有

$$(u,v)_{H^1} = (u,v)_{L^2} + (\nabla u, \nabla v)_{L^2}, \quad (u,v)_{H_0^1} = (\nabla u, \nabla v)_{L^2}.$$

当 Ω 是有界区域时，根据 Poincaré 不等式，这两个内积所定义的范数是等价的。

我们首先证明一个在 $H_0^1(\Omega)$ 中的分部积分的结果：

引理 492. *对任意的 $u \in H_0^1(\Omega)$，$v \in H^1(\Omega)$，对任意的 $k = 1, 2, \cdots, n$，我们有*

$$\int_\Omega (\partial_k u) \cdot v \, dx = -\int_\Omega u \cdot (\partial_k v) \, dx.$$

注记. *直观上，u 在边界上等于 0，所以，边界项的贡献是 0。*

证明: 当 $u = \varphi \in C_0^\infty(\Omega)$ 时，这个等式就是分布的导数的定义，所以成立。对于一般的 u，我们选取一列 $\{\varphi_p\}_{p \geqslant 1} \subset H_0^1(\mathbb{R}^n)$，使得 $\varphi_p \xrightarrow{H_0^1(\Omega)} u$。所以，

$$\int_\Omega (\partial_k \varphi_p) \cdot v \, dx = -\int_\Omega \varphi_p \cdot (\partial_k v) \, dx,$$

从而，推出来

$$\int_\Omega (\partial_k u) \cdot v \, dx + \int_\Omega u \cdot (\partial_k v) \, dx = -\int_\Omega (\partial_k(\varphi_p - u)) \cdot v \, dx - \int_\Omega (\varphi_p - u)(\partial_k v) \, dx.$$

由于 $v, \partial_k \in L^2(\Omega)$ 而 $\varphi_p - u \xrightarrow{L^2(\Omega)} 0$，$\partial_k(\varphi_p - u) \xrightarrow{L^2(\Omega)} 0$，所以当 $p \to \infty$ 时，上面的等式的极限就给出了要证明的等式。 □

另外，这些 Sobolev 空间还有如下的映射性质：

$$H^1(\mathbb{R}^n) \xrightarrow{\nabla} L^2(\Omega) \xrightarrow{\nabla} H^{-1}(\Omega).$$

实际上，对任意的 $u \in L^2(\Omega)$，考虑试验函数 $\varphi \in \mathcal{D}(\Omega)$，那么，对任意的 $k \leqslant n$，我们有

$$|\langle \partial_k u, \varphi \rangle| = |\langle u, \partial_k \varphi \rangle| \leqslant \|u\|_{L^2} \|\nabla \varphi\|_{L^2} = C \|\nabla \varphi\|_{H_0^1}.$$

这就说明 $\nabla u \in H^{-1}(\Omega)$（定义为 $H_0^1(\Omega)$ 的对偶）。

我们现在利用 Poincaré 不等式以及 $H_0^1(\Omega)$ 空间的结构来研究 Dirichlet 问题，即找一个函数 u，满足如下微分方程的边值问题：

$$\begin{cases} -\triangle u = f, \\ u\big|_{\partial\Omega} = 0. \end{cases}$$

我们采取一种迂回折中的方式来研究上述问题。我们首先给出所求解的一个预设：寻找 $u \in H_0^1(\Omega)$，使得 $-\triangle u = f$。直观上，H^{-1} 可以视作是对 H^1 中的函数求两次导数，此时，我们先要求 $f \in H^{-1}(\Omega)$。我们要求 $u \in H_0^1(\Omega)$ 是出于两个方面的考虑：第一，物理上这个问题来源于经典引力理论或者静电磁理论，这些问题通常要求函数 u 是有限能量的，即 $\nabla u \in L^2(\Omega)$；第二，$H_0^1(\Omega)$ 中的函数可以直观上理解为在 $\partial\Omega$ 上的限制为 0 的函数，所以，在这个类中寻找解可以不再去考虑边界条件（或者说，我们可以自欺欺人地认为这个解满足了边界条件，我们之后会对这个部分做严格讨论）。

由于 u 是分布（按照定义），我们先寻找一个在分布意义下的解，即对任意的试验函数 $\varphi \in \mathcal{D}(\Omega)$，如下的等式成立（这种解在偏微分方程理论中被称作是**弱解**）：

$$-\langle \triangle u, \varphi\rangle = \langle f, \varphi\rangle \ \Leftrightarrow \langle \nabla u, \nabla\varphi\rangle = \langle f, \varphi\rangle.$$

由于我们预设的 $u \in H_2^1(\Omega)$，我们把上面的分布等式改写成内积意义下等式：

$$(\nabla\varphi, \overline{\nabla u})_{L^2} = f(\varphi).$$

其中，我们把 $f \in H^{-1}(\Omega)$ 直接写成 $H_0^{(}\Omega)$ 上的函数的线性泛函的形式。所以，利用 $H_0^1(\Omega)$ 上的内积，这个问题转化为寻找 $\overline{u} \in H_0^1(\Omega)$，使得

$$(\varphi, \overline{u})_{H_0^1(\Omega)} = f(\varphi).$$

实际上，由于 f 是有界线性泛函，根据 Riesz 表示定理，存在唯一的 $\overline{u} \in H_0^1(\Omega)$，使得对任意的 $v \in H_0^1(\Omega)$，我们都有

$$(v, \overline{u})_{H_0^1(\Omega)} = f(v).$$

所以，我们构造了 Dirichlet 问题的唯一解（在 $H_0^1(\Omega)$ 中的唯一解）。

综上所述，我们得到如下的定理：

定理 493. *假设 Ω 是有界区域（夹在平行的两个超平面之间即可）。那么，对任意的 $f \in H^{-1}(\Omega)$，Dirichlet 问题*

$$\begin{cases} -\triangle u = f, \\ u\big|_{\partial\Omega} = 0. \end{cases}$$

在 $H_0^1(\Omega)$ 中存在唯一的解，即存在唯一的 $u \in H_0^1(\Omega)$，使得在分布的意义下

$$-\triangle u \overset{\mathcal{D}'}{=} f.$$

注记. *通过解上述的 Dirichet 问题，我们可以定义算子*

$$(-\triangle)^{-1} : H^{-1}(\Omega) \longrightarrow H_0^1(\Omega), \quad f \mapsto u,$$

其中 u 是 f 所对应的那个唯一的解。对任意的 $v \in H_0^1(\Omega)$，由于

$$\big|(v, \overline{u})_{H_0^1(\Omega)}\big| = \big|f(v)\big| \leqslant C\|v\|_{H_0^1},$$

我们总是可以选取 $C = 2\|f\|_{H^{-1}}$，所以，当 $v = u$ 时，我们有

$$\|u\|_{H_0^1}^2 \leqslant 2\|f\|_{H^{-1}}\|u\|_{H_0^1} \quad \Rightarrow \quad \|u\|_{H_0^1} \leqslant 2\|f\|_{H^{-1}},$$

即

$$\|(-\triangle)^{-1}(f)\|_{H_0^1} \leqslant 2\|f\|_{H^{-1}}.$$

所以，

$$(-\triangle)^{-1} : H^{-1}(\Omega) \longrightarrow H_0^1(\Omega)$$

是连续线性映射。

引理 494. 我们可以将 $L^2(\Omega)$ 实现为 $H^{-1}(\Omega) = \left(H_0^1(\Omega)\right)^*$ 的子空间：对任意的 $f \in L^2(\Omega)$，我们定义

$$T_f : H_0^1(\Omega) \to \mathbb{C}, \quad \psi \mapsto (\psi, \overline{f})_{L^2} = \int_\Omega f(x)\psi(x)dx.$$

证明: 根据 Cauchy-Schwarz 不等式和 Poincaré 不等式，对任意的 $\psi \in H_0^1(\Omega)$ 我们有

$$|T_f(\psi)| \leqslant \|\psi\|_{L^2}\|f\|_{L^2} \leqslant (C\|f\|_{L^2}) \cdot \|\nabla\psi\|_{L^2}.$$

所以，$T_f \in H^{-1}(\Omega)$。 □

我们可以把 $(-\triangle)^{-1}$ 限制在 $L^2(\Omega)$ 上

$$\begin{array}{ccc} L^2(\Omega) & \xhookrightarrow{\iota} & H^{-1}(\Omega) \\ & \searrow^{(-\triangle)^{-1}} & \downarrow{\scriptstyle (-\triangle)^{-1}} \\ & & H_0^1(\Omega) \end{array}$$

这就得到

$$(-\triangle)^{-1} : L^2(\Omega) \longrightarrow H_0^1(\Omega).$$

换而言之，我们把 f 在 $L^2(\Omega)$ 中选取来解 Dirichlet 问题就给出了上述映射。我们还可以进一步把 $H_0^1(\Omega)$ 嵌入到 $L^2(\Omega)$ 中，从而得到

$$\begin{array}{ccc} H^{-1}(\Omega) & \xrightarrow{(-\triangle)^{-1}} & H_0^1(\Omega) \\ {\scriptstyle \iota}\uparrow & & \downarrow{\scriptstyle \iota} \\ L^2(\Omega) & \xrightarrow{(-\triangle)^{-1}} & L^2(\Omega) \end{array}$$

这样，我们可以定义 $L^2(\Omega)$ 到自身的**连续**线性算子（因为它是一系列连续算子的复合）：

$$(-\triangle)^{-1} : L^2(\Omega) \longrightarrow L^2(\Omega).$$

我们之后会研究 $(-\triangle)^{-1}$ 的谱理论，即它的特征值和特征向量的性质。

命题 495. 算子

$$(-\triangle)^{-1}: L^2(\Omega) \longrightarrow L^2(\Omega),$$

是连续的正线性算子，即对任意的 $f \in L^2(\Omega)$，我们有

$$\left((-\triangle)^{-1} f, f\right)_{L^2} \geqslant 0.$$

这个不等式中等号成立当且仅当 $f = 0$。

证明: 令 $u = (-\triangle)^{-1} f$，从而 $u \in H_0^1(\Omega)$。按照 $(-\triangle)^{-1}$ 的定义，对任意的 $v \in H_0^1(\Omega)$，我们有

$$(v, \overline{u})_{H_0^1(\Omega)} = f(v).$$

令 $v = \overline{u}$，从而，

$$0 \leqslant (\overline{u}, \overline{u})_{H_0^1(\Omega)} = f(\overline{u}) = \int_\Omega f \cdot \overline{(-\triangle)^{-1} f} dx.$$

所以，命题中的不等式成立。

进一步，上面不等式中的等号成立当且仅当 $\overline{u} \in H_0^1(\Omega)$ 为 0，这个范数与 $H^1(\Omega)$ 中的范数等价，所以 $u \equiv 0$。 □

注记 (总结)**.** 我们构造了 *Dirichlet* 问题的解算子 $(-\triangle)^{-1}$。在这个过程中，我们只是假定了 $H_0^1(\Omega)$ 中的函数满足 $u\big|_{\partial\Omega}$ 的边界条件。我们下面要把这个边界条件讲清楚。为此，我们仍然要回到 $\mathbb{R}^n$ 研究 *Sobolev* 函数的限制理论。

我们用 $\mathbb{H}^n$ 代表上半空间:

$$\mathbb{H}^n = \left\{(x', x_n) \in \mathbb{R}^n \middle| x' = (x_1, \cdots, x_{n-1}) \in \mathbb{R}^{n-1}, x_n > 0\right\}.$$

引理 496 (用全空间上的连续函数逼近上半空间上的 H^1 函数)**.** 对任意的 $u \in H^1(\mathbb{H}^n)$，存在 $\{\varphi\}_{k\geqslant 1} \subset C^\infty(\mathbb{R}^n)$，使得

1) 对任意的 $k \geqslant 1$，$\varphi_k\big|_{\mathbb{H}^n} \in H^1(\mathbb{H}^n)$;

2) u 可以被 $\varphi_k\big|_{\mathbb{H}^n}$ 逼近:

$$\lim_{k\to\infty} \left\|\varphi_k\big|_{\mathbb{H}^n} - u\right\|_{H^1(\mathbb{H}^n)} = 0.$$

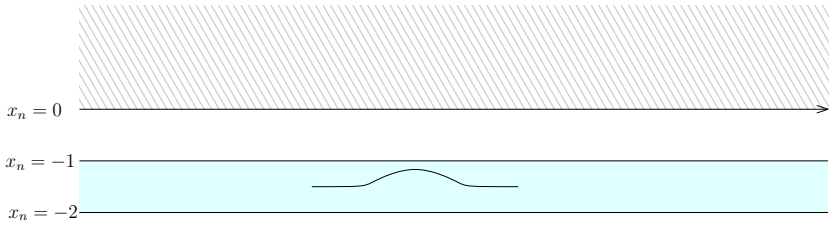
证明: 我们首先把 u 延拓成为 $\mathbb{R}^n$ 上的平方可积函数:

$$\widetilde{u}(x) = \begin{cases} u(x), & x_n > 0; \\ 0, & x_n \leqslant 0. \end{cases}$$

我们再把 $\partial_j u$ 延拓成为 $\mathbb{R}^n$ 上的平方可积函数，其中 $j \leqslant n$:

$$\widetilde{\partial_j u}(x) = \begin{cases} \partial_j u(x), & x_n > 0; \\ 0, & x_n \leqslant 0. \end{cases}$$

现在选取截断函数 χ 使得 $\mathrm{supp}(\chi) \subset \left\{(x', x_n) \middle| -2 \leqslant x_n \leqslant -1\right\}$ 并且 $\displaystyle\int_{\mathbb{R}^n} \chi(x) dx = 1$（这个积分的条件保证了各种卷积的收敛性）。



对任意的 $\varepsilon > 0$，令

$$\chi_\varepsilon(x) = \frac{1}{\varepsilon^n}\chi\left(\frac{x}{\varepsilon}\right).$$

我们定义

$$\varphi_\varepsilon(x) = \chi_\varepsilon * \widetilde{u}.$$

那么，对任意的 $j \leqslant n$，我们有

$$\left(\partial_j\varphi_\varepsilon\right)(x) = \left(\chi_\varepsilon * \left(\partial_j\widetilde{u}\right)\right)(x) = \int_{\mathbb{R}^n} \chi_\varepsilon(x-y)\partial_j\widetilde{u}(y)dy.$$

当 $x \in \mathbb{H}^n$ 时，即 $x_n > 0$ 时，上面的积分中由于 $-2\varepsilon < x_n - y_n < -\varepsilon$（否则贡献为 0），所以，

$$y_n > x_n + \varepsilon.$$

这表明积分项中我们可以只考虑 $y_n > 0$ 的情况。此时，我们有

$$\left(\partial_j\varphi_\varepsilon\right)(x) = \int_{\mathbb{R}^n} \chi_\varepsilon(x-y)\widetilde{\partial_j u}(y)dy = \left(\chi_\varepsilon * \widetilde{\partial_j u}\right)(x).$$

由于

$$\varphi_\varepsilon = \chi_\varepsilon * \widetilde{u} \xrightarrow{L^2(\mathbb{R}^n)} \widetilde{u},$$
$$\partial_j\varphi_\varepsilon = \chi_\varepsilon * \widetilde{\partial_j u} \xrightarrow{L^2(\mathbb{R}^n)} \widetilde{\partial_j u},$$

所以，限制到 $\mathbb{H}^n$ 上，我们就有

$$\varphi_\varepsilon\big|_{\mathbb{H}^n} \xrightarrow{L^2(\mathbb{H}^n)} u, \quad \partial_j\varphi_\varepsilon\big|_{\mathbb{H}^n} \xrightarrow{L^2(\mathbb{R}^n)} \partial_j u.$$

这就给出了证明。 □

我们现在再证明一个 H^1 的函数从 $\mathbb{H}$ 到 $\mathbb{R}^n$ 的扩张定理：

$$\begin{array}{ccc} H^1\left(\mathbb{H}^n\right) & \xrightarrow{\mathrm{Ext}_{\mathrm{Sym}}} & H^1\left(\mathbb{R}^n\right) \\ & \searrow^{\mathrm{id}} & \downarrow{\scriptstyle \mathrm{Res}} \\ & & H^1\left(\mathbb{H}^n\right) \end{array}$$

引理 497 (从 $H^1(\mathbb{H})$ 到 $H^1(\mathbb{R}^n)$ 扩张)**.** 限制映射

$$\mathrm{Res} : H^1\left(\mathbb{R}^n\right) \to H^1\left(\mathbb{H}^n\right)$$

是连续的满的线性映射击并且存在连续线性的扩张映射

$$\mathrm{Ext}_{\mathrm{Sym}} : H^1\left(\mathbb{R}^n\right) \to H^1\left(\mathbb{H}^n\right)$$

使得

$$\mathrm{Res} \circ \mathrm{Ext}_{\mathrm{Sym}} = \mathrm{id}_{H^1(\mathbb{H}^n)}.$$

为了证明这个命题，我们将 Stokes 公式略加推广：假设 Ω 是一个有界带边光滑区域，$\nu(x) = (\nu_1(x), \cdots, \nu_n(x))$ 为 $\partial\Omega$ 的单位外法向量，$d\sigma$ 为 $\partial\Omega$ 上的曲面测度，作为分布，我们有等式

$$\partial_k \mathbf{1}_\Omega \overset{\mathcal{D}'(\mathbb{R}^n)}{=} -\nu_k d\sigma,$$

其中 $k \leqslant n$。

命题 498 (跳跃公式)**.** 假设 $\Omega \subset \mathbb{R}^n$ 是一个有界带边光滑区域，$\nu(x)$ 为 $\partial\Omega$ 的单位外法向量，$d\sigma$ 为 $\partial\Omega$ 上的曲面测度。假设 $f \in C^{\mathbb{R}^n}$（只要在 Ω 的一个邻域上定义即可）并且对某个 $k_0 \leqslant n$，我们有 $\partial_{k_0} f \in L^1_{\mathrm{loc}}$。那么，在分布的意义下，我们有

$$\partial_{k_0}\left(f \cdot \mathbf{1}_\Omega\right) \overset{\mathcal{D}'(\mathbb{R}^n)}{=} \left(\partial_{k_0} f\right) \cdot \mathbf{1}_\Omega - \left(\nu_{k_0} \cdot f\right) d\sigma.$$

证明: 定理的证明是一系列逼近的过程。首先，我们选取 $\psi \in C_0^\infty(\mathbb{R}^n)$，使得 $\psi\big|_{\Omega_\delta} \equiv 1$（$\Omega_\delta$ 是到 Ω 的距离小于 δ 的点所组成的集合），我们再令 $g = \psi \cdot f$，这是一个连续函数，我们首先说明，在 $L^1(\mathbb{R}^n)$ 中，有

$$\partial_{k_0} g \overset{L^1(\mathbb{R}^n)}{=} f\partial_{k_0}\psi + \psi\partial_{k_0} f.$$

实际上，左右两边的项都落在 $L^1(\mathbb{R}^n)$，所以只要在分布的意义下验证上述式子即可，这是平凡的。

根据上面的公式，在 Ω 上，我们有

$$f \equiv g, \quad \partial_{k_0} f \cdot \mathbf{1}_\Omega \equiv \partial_{k_0} g \cdot \mathbf{1}_\Omega,$$

所以只要对 g 证明跳跃公式即可。

我们用 $g_\varepsilon = \chi_\varepsilon * g$ 来逼近 g，其中 χ_ε 是最常用的单位逼近，此时，由于 g 是连续有紧支集的函数，所以，g_ε 一致地收敛到 g。由于 g_ε 是光滑的，所以，根据 Stokes 公式，我们很容易得到（把 g_ε 放到分布配对的右边直接验证即可）：

$$\partial_{k_0}\left(g_\varepsilon \cdot \mathbf{1}_\Omega\right) \overset{\mathcal{D}'(\mathbb{R}^n)}{=} \left(\partial_{k_0} g_\varepsilon\right) \cdot \mathbf{1}_\Omega - \left(\nu_{k_0} \cdot g_\varepsilon\right) d\sigma.$$

根据 g_ε 一致地收敛到 g，我们有

$$g_\varepsilon \cdot \mathbf{1}_\Omega \overset{\mathcal{D}'(\mathbb{R}^n)}{\longrightarrow} g \cdot \mathbf{1}_\Omega, \quad \left(\nu_{k_0} \cdot g_\varepsilon\right) d\sigma \overset{\mathcal{D}'(\mathbb{R}^n)}{\longrightarrow} \left(\nu_{k_0} \cdot g\right) d\sigma,$$

所以，

$$\partial_{k_0} g_\varepsilon \cdot \mathbf{1}_\Omega \overset{\mathcal{D}'(\mathbb{R}^n)}{\longrightarrow} \partial_{k_0} g \cdot \mathbf{1}_\Omega, \quad (\nu_{k_0} \cdot g_\varepsilon)\, d\sigma \overset{\mathcal{D}'(\mathbb{R}^n)}{\longrightarrow} (\nu_{k_0} \cdot g)\, d\sigma,$$

只要证明上面等式中右边第一项收敛即可，实际上，我们只要证明下面的等式即可：

$$\partial_{k_0} g_\varepsilon \overset{L^1(\mathbb{R}^n)}{\longrightarrow} \partial_{k_0} g.$$

这是因为

$$\partial_{k_0} g_\varepsilon = (\partial_{k_0} g) * \chi_\varepsilon,$$

而与 χ_ε 卷积在 L^1 中是收敛的。 □

扩张引理的证明. 限制映射的连续性是显然的。

先假设 u 是 $\mathbb{R}^n$ 上的一个光滑函数 F 的限制（根据上一个引理，这样的函数在 $H^1(\mathbb{H}^n)$ 里面是稠密的）：$u = F\big|_{\mathbb{H}^n}$。利用对称扩张，我们现在构造

$$\mathrm{Ext}_{\mathrm{Sym}} : C^\infty(\mathbb{R}^n) \cap H^1(\mathbb{H}^n) \to H^1(\mathbb{H}^n), \quad u \mapsto \widetilde{u}(x', x_n) = \begin{cases} u(x', x_n), & x_n \geqslant 0; \\ u(x', -x_n), & x_n \leqslant 0. \end{cases}$$

我们注意到 $\widetilde{u}$ 是连续映射并且：

- $\widetilde{u}\big|_{x_n \geqslant 0}$ 是光滑函数 $F(x', x_n)$ 的限制；
- $\widetilde{u}\big|_{x_n \leqslant 0}$ 是光滑函数 $F(x', -x_n)$ 的限制；

特别地，上面这两个限制都将满足跳跃公式中的要求。那么，

$$\mathrm{Res} \circ \mathrm{Ext}_{\mathrm{Sym}} = \mathrm{id}_{H^1(\mathbb{H}^n)},$$

明显成立。特别地，如果 $\mathrm{Ext}_{\mathrm{Sym}}$ 是连续映射，这个等式表明限制映射 Res 是满射。

我们下面证明 $\mathrm{Ext}_{\mathrm{Sym}}$ 是连续线性映射。

此时，根据跳跃公式，我们显然有

$$\partial_j \left(\mathrm{Ext}_{\mathrm{Sym}}(u)\right) = \mathrm{Ext}_{\mathrm{Sym}}(\partial_k u), \quad k = 1, 2, \cdots, n-1.$$

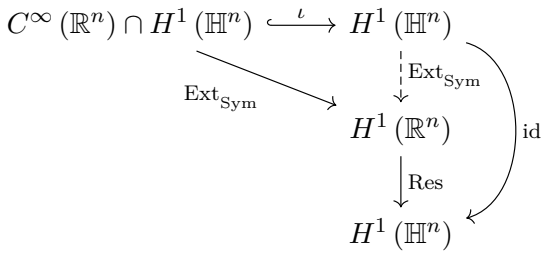
我们现在计算 $\partial_n \left(\mathrm{Ext}_{\mathrm{Sym}}(u)\right)$：

$$\begin{aligned} \partial_n \left(\mathrm{Ext}_{\mathrm{Sym}}(u)\right) &= \partial_n \left(\widetilde{u}(x', x_n)\mathbf{1}_{x_n>0} + \widetilde{u}(x', -x_n)\mathbf{1}_{x^n<0}\right) \\ &= \left(\partial_n F(x', x_n)\mathbf{1}_{x_n>0} + F dx'\right) - \left(\partial_n F(x', -x_n)\mathbf{1}_{x_n>0} + F dx'\right) \\ &= \left(\mathbf{1}_{x_n>0} - \mathbf{1}_{x_n<0}\right)\left(\mathrm{Ext}_{\mathrm{Sym}}(\partial_n u)\right). \end{aligned}$$

综合上面的两个等式，我们自然有

$$\|\mathrm{Ext}_{\mathrm{Sym}}(u)\|_{H^1(\mathbb{R}^n)} \leqslant \sqrt{2}\|u\|_{H^1(\mathbb{H}^n)}.$$

由于 $C^\infty(\mathbb{R}^n)\cap H^1(\mathbb{H}^n)$ 在 $H^1(\mathbb{H}^n)$ 是稠密的，



所以上面 $\mathrm{Ext}_{\mathrm{Sym}}$ 可以被延拓成 $H^1(\mathbb{H}^n)$ 的连续线性映射。证毕。 □

我们下面要定义到边界上的限制映射

$$\mathrm{Res}: H^1(\mathbb{H}^n)\to H^{\frac{1}{2}}\left(\partial(\mathbb{H}^n)\right).$$

由于 $H^1(\mathbb{H}^n)$ 中的函数在边界 $\partial\mathbb{H}^n$（由 $x_n=0$ 所定义）上就没有定义，所以，如何正确地写下这个映射需要一些仔细的论证。

实际上，对于每一个 $u\in H^1(\mathbb{H}^n)$，按照上面的扩张定理，我们可以找到它在全空间上的扩张

$$\mathrm{Ext}_{\mathrm{Sym}}(u)\in H^1(\mathbb{R}^n),$$

所以，我们可以定义

$$\mathrm{Res}(u)=\mathrm{Ext}_{\mathrm{Sym}}(u)\big|_{x_n=0}.$$

我们注意到，尽管我们在定义中要求 $\mathrm{Ext}_{\mathrm{Sym}}(u)(x',0)=0$，但是，$\mathrm{Ext}_{\mathrm{Sym}}(u)\big|_{x_n=0}$ 未必是 0。当然，这样的定义也带来了其他的疑问：我们的扩张映射是人为选取的（用对称的方法），如果换一个扩张，那么得到的结论可能不一样。

我们还有另外一种定义限制映射的方式：直观上，对于 $t>0$，我们应该不难定义

$$\mathrm{Res}(u)=u\big|_{x_n=t}\in H^1(\mathbb{R}^{n-1}).$$

这样，我们得到了映射

$$\mathbb{R}_{t>0}\to H^1(\mathbb{R}^{n-1}),\quad t\mapsto u\big|_{x_n=t},$$

如果我们可以说明这个映射对 t 是连续的，那么，我们希望能用下面的极限定义限制映射：

$$\mathrm{Res}(u)\overset{H^1(\mathbb{R}^{n-1})}{=}\lim_{t\to0^+}u\big|_{x_n=t}.$$

我们将看到，这两种途径在一定意义上是一致的（第二种途径显然不依赖于扩张的选取而第一种途径更容易计算）。

我们定义如下的空间：

$$C_b^0\left(\mathbb{R},H^s(\mathbb{R}^{n-1})\right)=\left\{u:\mathbb{R}\to H^s(\mathbb{R}^{n-1})\text{连续}\middle|\text{对任意 }t\in\mathbb{R}\text{，存在 }M\text{，使得 }\|u(t)\|_{H^s(\mathbb{R}^{n-1})}<M\right\}.$$

这里，由于 $H^s(\mathbb{R}^{n-1})$ 是（完备）距离空间，这里连续性指的就是两个距离空间之间的连续映射。

我们还可以在 $C_b^0\left(\mathbb{R}, H^s(\mathbb{R}^{n-1})\right)$ 上定义如下的范数，其中，对任意的 $u \in C_b^0\left(\mathbb{R}, H^s(\mathbb{R}^{n-1})\right)$，我们令

$$\|\!|u|\!\|_s = \sup_{t\in R} \|u(t)\|_{H^s(\mathbb{R}^{n-1})}.$$

在这个范数下，我们得到了一个完备的赋范线性空间 $\left(C_b^0\left(\mathbb{R}, H^s(\mathbb{R}^{n-1})\right), \|\!|\cdot|\!\|_s\right)$。实际上，这是如下抽象结论的推论（取 $X = H^s(\mathbb{R}^{n-1})$）（这就是一年级关于有界闭区间上连续函数空间是完备的简单推广）：

引理 499. *假设 $(X, \|\cdot\|)$ 是完备的赋范线性空间，我们定义*

$$C_b^0(\mathbb{R}, X) = \left\{u : \mathbb{R} \to X \text{连续} \middle| \text{对任意 } t \in \mathbb{R}, \text{存在 } M, \text{使得 } \|u(t)\| < M\right\}.$$

我们在 $C_b^0(\mathbb{R}, X)$ 上定义范数：

$$\|\!|u|\!\| = \sup_{t\in R} \|u(t)\|_X,$$

其中 $u \in C_b^0(\mathbb{R}, X)$，那么，$\left(C_b^0(\mathbb{R}, X), \|\!|\cdot|\!\|\right)$ 是完备的赋范线性空间。

证明： 验证 $\left(C_b^0(\mathbb{R}, X), \|\!|\cdot|\!\|\right)$ 是赋范线性空间是平凡的，我们现在来证明它的完备性。根据第二学期（五月七日的课）关于完备性的级数判定，我们任选绝对收敛的级数 $\sum\limits_{k=1}^{\infty} u_k(t)$，其中，对任意的 $k \geqslant 1$，$u_k(t) \in C_b^0(\mathbb{R}, X)$，并且

$$\sum_{k=1}^{\infty} \|\!|u_k(t)|\!\| < \infty,$$

我们要证明 $\sum\limits_{k=1}^{\infty} u_k(t)$ 在 $C_b^0(\mathbb{R}, X)$ 收敛即可。

首先，对任意固定的 $t_0 \in \mathbb{R}$，按照 $\|\!|\cdot|\!\|$ 的定义，我们有

$$\sum_{k=1}^{\infty} \|u_k(t_0)\| \leqslant \sum_{k=1}^{\infty} \|\!|u_k|\!\|,$$

所以，

$$u(t_0) = \sum_{k=1}^{\infty} u_k(t_0) \in X$$

是良好定义的（用到了 X 的完备性）。

现在证明映射

$$u : \mathbb{R} \to X, \quad t \mapsto u(t)$$

是连续的：实际上，对任意的 $t_0, \delta \in R$，我们有

$$\begin{aligned} u(t_0+\delta)-u(t_0) &= \sum_{k=1}^{\infty}\left(u_k(t_0+\delta)-u_k(t_0)\right) \\ &= \underbrace{\sum_{k=1}^{N}\left(u_k(t_0+\delta)-u_k(t_0)\right)}_{\text{对 } \delta\to 0 \text{ 是连续的}} + \underbrace{\sum_{k\geqslant N+1}^{\infty}\left(u_k(t_0+\delta)-u_k(t_0)\right)}_{\|\cdot\|\leqslant 2\sum\limits_{k\geqslant N+1}|||u_k|||}. \end{aligned}$$

对任意的 $\varepsilon>0$，利用级数 $\sum\limits_{k=1}^{\infty}u_k$ 的一致收敛性，存在 N，使得

$$\sum_{k\geqslant N+1}|||u_k||| < \frac{\varepsilon}{4}.$$

固定这个 N。利用 $u_k(t)$ 在 t_0 处的连续性，存在 $\delta_0>0$，使得当 $|\delta|<\delta_0$ 时，对每个 $k\leqslant N$，我们有

$$\|u_k(t_0+\delta)-u_k(t_0)\| < \frac{\varepsilon}{2N}.$$

据此，对任意的 $\varepsilon>0$，存在 $\delta_0>0$，使得当 $|\delta|<\delta_0$ 时，我们有

$$\|u(t_0+\delta)-u(t_0)\| < N\times\frac{\varepsilon}{2N}+2\times\frac{\varepsilon}{4}=\varepsilon.$$

这就证明了 $u(t)$ 的连续性。很明显，对任意的 t，我们还有

$$\|u(t)\|\leqslant\sum_{k=1}^{\infty}|||u_k|||.$$

从而，$u(t)\in C_b^0(X)$。

最终，我们还要说明部分和的收敛：

$$\lim_{N\to\infty}\left|\left|\left|u(t)-\sum_{k=1}^{N}u_k(t)\right|\right|\right|=0.$$

这是因为对于每个 t，我们有一致的上界：

$$\left\|u(t)-\sum_{k=1}^{N}u_k(t)\right\|=\left\|\sum_{k\geqslant N+1}u_k(t)\right\|\leqslant\sum_{k\geqslant N+1}|||u_k|||.$$

命题得证。 □

我们下次课程将要证明，对任意的 $u\in H^s(\mathbb{R}^n)$，通过把它限制到 $x_n=t$ 这个超平面上，我们就得到映射

$$u:\mathbb{R}\to H^{s-\frac{1}{2}}(\mathbb{R}^{n-1}),\quad t\mapsto u\big|_{x_n=t}.$$

这个映射是连续并且一致有界的。换而言之，我们有如下的连续嵌入

$$\iota: H^s(\mathbb{R}^n)\hookrightarrow C_b^0\left(\mathbb{R}, H^{s-\frac{1}{2}}(\mathbb{R}^{n-1})\right), u\mapsto\left(t\mapsto u\big|_{x_n=t}\right).$$

74 限制（迹）定理的连续性，半空间上 Sobolev 空间的限制（迹）定理，限制的正合列

二零二零年十一月十六日，星期一，晴

我们下面说明 $C_b^0\left(\mathbb{R}, H^s(\mathbb{R}^{n-1})\right)$ 中元素可以被视作是缓增的分布，实际上，我们有嵌入

$$C_b^0\left(\mathbb{R}, H^s(\mathbb{R}^{n-1})\right) \hookrightarrow \mathcal{S}'(\mathbb{R}^n).$$

首先，对任意的 $u \in C_b^0\left(\mathbb{R}, H^s(\mathbb{R}^{n-1})\right)$，对任意的 $\varphi \in \mathcal{D}(\mathbb{R}^{n-1}\times\mathbb{R})$，我们定义

$$\langle u, \varphi\rangle = \int_{\mathbb{R}} \langle u(x_n), \varphi(\cdot, x_n)\rangle dx_n.$$

根据 H^s 与 H^{-s} 之间的对偶性，我们有

$$|\langle u, \varphi\rangle| \leqslant \int_{\mathbb{R}} \|u(x_n)\|_{H^s(\mathbb{R}^{n-1})} \|\varphi(x', x_n)\|_{H^{-s}(\mathbb{R}^{n-1}_{x'})} dx_n.$$

此时，我们知道

$$\|\varphi(x', x_n)\|_{H^{-s}(\mathbb{R}^{n-1}_{x'})} \leqslant \frac{1}{x_n^2+1} N_{\lfloor -s\rfloor+2n+4}(\varphi).$$

所以，

$$|\langle u, \varphi\rangle| \leqslant \int_{\mathbb{R}} |||u(x_n)|||_s \frac{1}{x_n^2+1} N_{\lfloor -s\rfloor+2n+4}(\varphi) dx_n = C|||u(x_n)|||_s N_{\lfloor -s\rfloor+2n+4}(\varphi).$$

这说明上述定义的配对 $\langle u, \varphi\rangle$ 不仅是分布还是缓增的分布。

下面证明，$C_b^0\left(\mathbb{R}, H^s(\mathbb{R}^{n-1})\right)$ 中元素到缓增分布的映射是嵌入，即给定 $u \in C_b^0\left(\mathbb{R}, H^s(\mathbb{R}^{n-1})\right)$，如果对任意的 $\varphi \in \mathcal{D}(\mathbb{R}^n)$，我们都有 $\langle u, \varphi\rangle = 0$，那么，对任意的 $x_n \in \mathbb{R}$，$u(x', x_n) \overset{H^s(\mathbb{R}^{n-1})}{=} 0$。我们选取 φ 形如

$$0 = \varphi(x', x_n) = \psi(x')\phi(x_n), \quad \psi(x') \in \mathcal{D}(\mathbb{R}^{n-1}),\ \phi(x_n) \in \mathcal{D}(\mathbb{R}).$$

所以，

$$\langle u, \varphi\rangle = \int_{\mathbb{R}} \langle u(x_n), \psi\rangle \phi(x_n) dx_n.$$

另外，根据

$$|\langle u(x_n), \psi\rangle| \leqslant \|u(x_n)\|_{H^s(\mathbb{R}^{n-1})} \|\psi(x')\|_{H^{-s}(\mathbb{R}^{n-1}_{x'})}$$

所以，$\langle u(x_n), \psi\rangle$ 是 x_n 的有界函数（从而局部上是可积的）。由于这个函数和任意的 $\phi(x_n)$ 配对积分得 0，所以，对任意的 $\psi(x') \in \mathcal{D}(\mathbb{R}^{n-1})$，有

$$\langle u(x_n), \psi\rangle \overset{L^1_{\mathrm{loc}}}{=} 0.$$

所以，对几乎处处的 $x_n \in \mathbb{R}$，我们有

$$u(x_n) \overset{\mathcal{D}'}{=} 0.$$

从而，对几乎处处的 $x_n \in \mathbb{R}$，我们有

$$u(x_n) \overset{H^s(\mathbb{R}^{n-1})}{=} 0.$$

再利用 u 对 x_n 的连续性，我们就证明了 $u \overset{C_b^0\left(\mathbb{R},H^s(\mathbb{R}^{n-1})\right)}{=} 0$。

我们之前已经证明了，对于 $n \geqslant 1$，$s > \dfrac{1}{2}$，通过对 $\mathcal{S}(\mathbb{R}^n)$ 上定义的函数进行扩张，我们可以得到有界的限制映射

$$\mathrm{Res} : H^s(\mathbb{R}^n) \longrightarrow H^{s-\frac{1}{2}}(\mathbb{R}^{n-1}_{x_n=0}).$$

这个证明自然和 $x_n = 0$ 的选取没有关系，所以，对任意的 t，存在一致的常数 $C > 0$，对任意的 $u \in H^s(\mathbb{R}^n)$，我们有

$$\|\mathrm{Res}(u)\|_{H^{s-\frac{1}{2}}(\mathbb{R}^{n-1}_{x_n=t})} \leqslant C\|u\|_{H^s(\mathbb{R}^n)}.$$

这个 $\mathrm{Res}(u)$ 和 $x_n = t$ 相关，我们把它记做是 $u\big|_{x_n=t}$，于是，我们得到了有界的映射

$$u : \mathbb{R} \to H^{s-\frac{1}{2}}(\mathbb{R}^{n-1}), \quad t \mapsto u\big|_{x_n=t}.$$

定理 500. *对任意的 $u \in H^s(\mathbb{R}^n)$，映射*

$$u : \mathbb{R} \to H^{s-\frac{1}{2}}(\mathbb{R}^{n-1}), \quad t \mapsto u\big|_{x_n=t}.$$

是连续的，换而言之，我们有如下的连续嵌入

$$\iota : H^s(\mathbb{R}^n) \hookrightarrow C_b^0\left(\mathbb{R}, H^{s-\frac{1}{2}}(\mathbb{R}^{n-1})\right), u \mapsto \left(t \mapsto u\big|_{x_n=t}\right).$$

特别地，存在常数 C，使得对任意的 $u \in H^s(\mathbb{R}^n)$，我们有

$$|||\iota(u)|||_{s-\frac{1}{2}} \leqslant C\|u\|_{H^s(\mathbb{R}^n)}.$$

证明: 根据之前已有的结论，我们只需要证明

$$u : \mathbb{R} \to H^{s-\frac{1}{2}}(\mathbb{R}^{n-1}), \quad t \mapsto u\big|_{x_n=t},$$

对 t 的连续性。仿照之前在 $x_n = 0$ 上的限制定理的证明，我们有（我们可以将下面的计算理解为对 Schwartz 函数来做的，一般的情况需要利用逼近来得到，因为这只是例行公事，所以我们不再给出细节）:

$$\mathcal{F}'(u(x',t))(\xi') = \frac{1}{2\pi}\int_{\mathbb{R}} \widehat{u}(\xi',\tau)e^{it\tau}d\tau,$$

其中，$\mathcal{F}'$ 是对前面 $n-1$ 个变量的 Fourier 变换。所以，

$$\mathcal{F}'(u)(\xi',t_1) - \mathcal{F}'(u)(\xi',t_2) = \frac{1}{2\pi}\int_{\mathbb{R}} \widehat{u}(\xi',\tau)\left(e^{it_1\tau} - e^{it_2\tau}\right)d\tau.$$

所以，如果令 $\sigma = s - \dfrac{1}{2}$，那么

$$\begin{aligned}\|u(x',t_1)-u(x',t_2)\|^2_{H^{s-\frac{1}{2}}(\mathbb{R}^{n-1})} &= \int_{\mathbb{R}^{n-1}} (1+|\xi'|^2)^{\sigma} \left|\mathcal{F}'(u)(\xi',t_1)-\mathcal{F}'(u)(\xi',t_2)\right|^2 d\xi' \\ &= \frac{1}{4\pi^2}\int_{\mathbb{R}^{n-1}} (1+|\xi'|^2)^{\sigma} \left|\int_{\mathbb{R}} \widehat{u}(\xi',\tau)\left(e^{it_1\tau}-e^{it_2\tau}\right)d\tau\right|^2 d\xi'.\end{aligned}$$

根据 Cauchy-Schwarz 不等式，我们有

$$\left|\int_{\mathbb{R}} \widehat{u}(\xi',\tau)\left(e^{it_1\tau}-e^{it_2\tau}\right)d\tau\right|^2 \leqslant \left(\int_{\mathbb{R}} \frac{\left|e^{it_1\xi_n}-e^{it_2\xi_n}\right|^2}{(1+|\xi|^2)^s} d\xi_n\right)\left(\int_{\mathbb{R}} (1+|\xi|^2)^s \left|\widehat{u}(\xi',\xi_n)\right|^2 d\xi_n\right).$$

在上面计算中，$\xi = (\xi', \xi_n)$ 中的 ξ' 是固定的。从而，

$$\begin{aligned}\int_{\mathbb{R}} \frac{\left|e^{it_1\xi_n}-e^{it_2\xi_n}\right|^2}{(1+|\xi|^2)^s} d\xi_n &= \int_{\mathbb{R}} \frac{\left|e^{it_1\xi_n}-e^{it_2\xi_n}\right|^2}{(1+|\xi'|^2+\xi_n^2)^s} d\xi_n \\ &= \frac{1}{(1+|\xi'|^2)^{s-\frac{1}{2}}}\int_{\mathbb{R}} \frac{\left|e^{it_1\sqrt{1+|\xi'|^2}\cdot\frac{\xi_n}{\sqrt{1+|\xi'|^2}}}-e^{it_2\sqrt{1+|\xi'|^2}\cdot\frac{d\xi_n}{\sqrt{1+|\xi'|^2}}}\right|^2}{\left(1+\left(\frac{\xi_n}{\sqrt{1+|\xi'|^2}}\right)^2\right)^s}\cdot\frac{d\xi_n}{\sqrt{1+|\xi'|^2}} \\ &= \frac{1}{(1+|\xi'|^2)^{s-\frac{1}{2}}}\underbrace{\int_{\mathbb{R}} \frac{\left|e^{it_1\sqrt{1+|\xi'|^2}y}-e^{it_2\sqrt{1+|\xi'|^2}y}\right|^2}{(1+|y|^2)^s} dy}_{I(\xi;t_1,t_2)}.\end{aligned}$$

所以，

$$\begin{aligned}\|u(x',t_1)-u(x',t_2)\|^2_{H^{s-\frac{1}{2}}(\mathbb{R}^{n-1})} &\leqslant \frac{1}{4\pi^2}\int_{\mathbb{R}^{n-1}} I(\xi;t_1,t_2)\int_{\mathbb{R}} (1+|\xi|^2)^s \left|\widehat{u}(\xi',\xi_n)\right|^2 d\xi_n d\xi' \\ &= \frac{1}{4\pi^2}\int_{\mathbb{R}^n} I(\xi;t_1,t_2)(1+|\xi|^2)^s \left|\widehat{u}(\xi)\right|^2 d\xi.\end{aligned}$$

我们注意到，按照定义，

$$I(\xi;t_1,t_2) \leqslant \int_{\mathbb{R}} \frac{4}{(1+|y|^2)^s} dy = 4C_s.$$

是一致有界的（对 t_1, t_2 而言），所以，根据 Lebesgeu 控制收敛定理，当 $t_2 \to t_1$ 时，对任意的 $\xi \in \mathbb{R}^n$，我们都有

$$\lim_{t_2\to t_2} I(\xi;t_1,t_2) = 0.$$

再根据

$$\|u(x',t_1)-u(x',t_2)\|^2_{H^{s-\frac{1}{2}}(\mathbb{R}^{n-1})} = \frac{1}{4\pi^2}\int_{\mathbb{R}^n} I(\xi;t_1,t_2)(1+|\xi|^2)^s \left|\widehat{u}(\xi)\right|^2 d\xi.$$

由于 $I(\xi;t_1,t_2) \leqslant 4C_s$，所有右边是一个可积函数（因为 $u \in H^s(\mathbb{R}^n)$），由于 $\lim\limits_{t_2\to t_2} I(\xi;t_1,t_2) = 0$，再次利用 Lebesgue 控制收敛定理，上面不等式的右边就趋向于 0。这就完成了连续性的证明。 □

我们现在可以完整地陈述并证明如下的限制性定理

$$\mathrm{Res} : H^1(\mathbb{H}^n) \to H^{\frac{1}{2}}\left(\partial(\mathbb{H}^n)\right).$$

我们提过，由于 $H^1(\mathbb{H}^n)$ 中的函数在边界 $\partial\mathbb{H}^n$ 上就没有定义，这个映射有时候（在英文和法文的文献中总是）被称作是**迹映射**，迹大约代表的是从 $x_n > 0$ 上面取极限所留下的痕迹。

为了定义 $u \in H^1(\mathbb{H}^n)$ 在 $x_n = 0$ 上的迹，根据扩张定理，我们选取 u 在全空间上的扩张

$$\mathrm{Ext}_{\mathrm{Sym}}(u) \in H^1(\mathbb{R}^n),$$

所以，我们令

$$\mathrm{Res}(u) = \mathrm{Ext}_{\mathrm{Sym}}(u)\big|_{x_n=0}.$$

根据我们证明的连续性，我们自然有

$$\mathrm{Res}(u) = \lim_{t\to 0^+} \mathrm{Ext}_{\mathrm{Sym}}(u)\big|_{x_n=t}.$$

如果我们能说明，当 $t > 0$ 时，

$$\mathrm{Ext}_{\mathrm{Sym}}(u)\big|_{x_n=t} = u\big|_{x_n=t},$$

那么，我们就有

$$\mathrm{Res}(u) = \lim_{t\to 0^+} u\big|_{x_n=t}.$$

这就给出了 $\mathrm{Res}(u)$ 的定义，而且，这个定义不依赖于扩张的选取。

最终，为了说明当 $t > 0$ 时，我们有

$$\mathrm{Ext}_{\mathrm{Sym}}(u)\big|_{x_n=t} = u\big|_{x_n=t},$$

我们注意到这个等式对于满足 $u = U\big|_{\mathbb{H}^n}$ 是成立的，其中 $U \in C^\infty(\mathbb{R}^n)$。由于这样的函数在 $H^1(\mathbb{H}^n)$ 中是稠密的并且要证明的等式对于 $u \in H^1(\mathbb{H}^n)$ 也是连续的，所以根据连续性，这个等式对于所有 $u \in H^1(\mathbb{H}^n)$ 成立。

定理 501. *迹映射*

$$\mathrm{Res} : H^1(\mathbb{H}^n) \to H^{\frac{1}{2}}\left(\partial(\mathbb{H}^n)\right),$$

是连续线性映射。进一步，我们有正合列

$$0 \to H^1_0(\mathbb{H}^n) \xrightarrow{\iota} H^1(\mathbb{H}^n) \xrightarrow{\mathrm{Res}} H^{\frac{1}{2}}\left(\partial(\mathbb{H}^n)\right) \to 0,$$

也就是说 Res *是满射并且对任意的* $u \in H^1(\mathbb{H}^n)$，$\mathrm{Res}(u) = 0$ *当且仅当* $u \in H^1_0(\mathbb{H}^n)$。

证明: 首先证明 Res 的连续性。实际上，根据

$$\mathrm{Res}(u) = \lim_{t\to 0^+} u\big|_{x_n=t},$$

我们知道

$$\|\mathrm{Res}(u)\|_{H^{\frac{1}{2}}(\mathbb{R}^{n-1})} \leqslant C \lim_{t\to 0^+} \|u\big|_{x_n=t}\|_{H^{\frac{1}{2}}(\mathbb{R}^{n-1})} = C\|u\|_{H^1(\mathbb{R}^n)}.$$

下面上面的序列是正合的：我们注意到 ι 显然是单射；另外，如果 $u \in H_0^1(\mathbb{H}^n)$，那么，它可以被 $C_0^\infty(\mathbb{H}^n)$ 中的函数逼近，这些函数的迹显然是 0，所以，利用 Res 的连续性，我们就知道 $\mathrm{Res}\big|_{H_0^1(\mathbb{H}^n)} \equiv 0$。

下面假设 $\mathrm{Res}(u) \overset{H^{\frac{1}{2}}(\mathbb{R}^{n-1})}{=} 0$，其中 $u \in H^1(\mathbb{H}^n)$，我们来说明 $u \in H_0^1(\mathbb{H}^n)$。这一部分的证明并不容易，我们分成两步来完成。

第一步，构造 $\underline{u}$ 如下：

$$\underline{u}(x) = \begin{cases} u(x', x_n), & x_n > 0; \\ 0, & x_n \leqslant 0. \end{cases}$$

最重要的观察是

$$\mathrm{Res}(u) = \lim_{t\to 0^+} u\big|_{x_n=t} \overset{H^{\frac{1}{2}}(\mathbb{R}^{n-1})}{=} 0$$

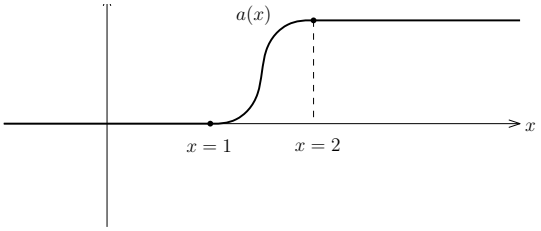
意味着

$$\underline{u}(\cdot, x_n) \in C_b^0(\mathbb{R}, H^{\frac{1}{2}}(\mathbb{R}^{n-1})).$$

我们声明，$\underline{u} \in H^1(\mathbb{R}^n)$ 并且对任意的 $k \leqslant n$，我们有

$$\partial_k \underline{u} = \partial_k u \cdot \mathbf{1}_{x_n>0}.$$

实际上，我们选取一个阶段函数 $a(x)$ 使得对任意的 $x \in \mathbb{R}$，$0 \leqslant a(x) \leqslant 1$；如果 $x \geqslant 2$，那么 $a(x) \equiv 1$；如果 $x \leqslant 1$，那么 $a(x) \equiv 0$。



为了计算 $\partial_k \underline{u}$，我们任选试验函数 $\varphi \in \mathcal{D}(\mathbb{R}^n)$，此时，我们有

$$\begin{aligned}
\langle \partial_k \underline{u}, \varphi\rangle &= -\langle \underline{u}, \partial_k \varphi\rangle = -\int_{\mathbb{R}^n} \underline{u}(x)\partial_k\varphi(x)dx \\
&= -\lim_{\varepsilon\to 0}\int_{\mathbb{R}^n} \underbrace{\underline{u}(x)a\left(\frac{x_n}{\varepsilon}\right)}_{\text{support } \subset \mathbb{H}^n} \partial_k\varphi(x)dx \\
&= -\lim_{\varepsilon\to 0}\int_{\mathbb{H}^n} u(x)a\left(\frac{x_n}{\varepsilon}\right)\partial_k\varphi(x)dx \\
&= \lim_{\varepsilon\to 0}\int_{\mathbb{H}^n} \partial_k u(x)a\left(\frac{x_n}{\varepsilon}\right)\varphi(x)dx + \lim_{\varepsilon\to 0}\int_{\mathbb{H}^n} u(x)\frac{1}{\varepsilon}a'\left(\frac{x_n}{\varepsilon}\right)\delta_k^n\varphi(x)dx.
\end{aligned}$$

根据 Lebesgue 控制收敛定理，我们有

$$\lim_{\varepsilon\to 0}\int_{\mathbb{H}^n}\partial_k u(x)a\left(\frac{x_n}{\varepsilon}\right)\varphi(x)dx=\int_{\mathbb{H}^n}\partial_k u(x)\varphi(x)dx=\langle\partial_k u\cdot\mathbf{1}_{x_n>0},\varphi\rangle.$$

我们需要证明第二个极限消失：

$$\begin{aligned}\left|\int_{\mathbb{H}^n}u(x)\frac{1}{\varepsilon}a'\left(\frac{x_n}{\varepsilon}\right)\delta_k^n\varphi(x)dx\right|&=\left|\int_{\varepsilon}^{2\varepsilon}\int_{\mathbb{R}^{n-1}}u(x',x_n)\frac{1}{\varepsilon}a'\left(\frac{x_n}{\varepsilon}\right)\delta_k^n\varphi(x',x_n)dx\right|\\&\leqslant\frac{1}{\varepsilon}\int_{\varepsilon}^{2\varepsilon}\underbrace{\left|a'\left(\frac{x_n}{\varepsilon}\right)\right|}_{O(1)}\underbrace{\|u(x',x_n)\|_{H^{\frac{1}{2}}}}_{o(1),\ \varepsilon\to 0}\underbrace{\|\varphi(x',x_n)\|_{H^{-\frac{1}{2}}}}_{O(1)}dx'dx_n\\&=o(1).\end{aligned}$$

这里，我们用到了 $\mathrm{Res}(u)=0$ 的条件。这就完成了第一步的证明。

第二步，对任意的 $\delta>0$，我们定义

$$\underline{u}_\delta(x)=\underline{u}(x',x_n-\delta).$$

很容易证明，在 $H^1(\mathbb{R}^n)$ 中，我们有

$$\underline{u}_\delta(x)\overset{H^1(\mathbb{R}^n)}{\longrightarrow}\underline{u}(x).$$

所以，

$$v_\delta(x)=\underline{u}_\delta(x)\big|_{\mathbb{H}^n}\overset{H^1(\mathbb{R}^n)}{\longrightarrow}\underline{u}(x)\big|_{\mathbb{H}^n}=u(x).$$

按照定义，我们有

$$v_\delta(x)=\begin{cases}u(x',x_n-\delta), & x_n>\delta;\\0, & x_n\leqslant\delta.\end{cases}$$

所以，只要证明 $v_\delta\in H_0^1(\mathbb{H}^n)$ 既可，其中，$\delta>0$ 是固定的。为此，我们先选取 $\{\varphi_p\}_{p\geqslant 1}\subset C_0^\infty(\mathbb{R}^n)$，使得

$$\varphi_p\overset{H^1}{\longrightarrow}\underline{u}_\delta,\quad p\to\infty.$$

我们考虑

$$\widetilde{\varphi}_p(x)=a\left(\frac{2x_n}{\delta}\right)\varphi_p(x).$$

很显然，我们有

$$\widetilde{\varphi}_p\overset{L^2(\mathbb{R}^n)}{\longrightarrow}\underline{u}_\delta,\quad p\to\infty.$$

由于

$$\partial_k\widetilde{\varphi}_p=a\left(\frac{2x_n}{\delta}\right)\partial_k\varphi_p(x)+\delta_k^n a'\left(\frac{2x_n}{\delta}\right)\varphi_p(x),$$

所以，

$$\partial_k\widetilde{\varphi}_p\to a\left(\frac{2x_n}{\delta}\right)\partial_k\underline{u}_\delta(x)+\delta_k^n a'\left(\frac{2x_n}{\delta}\right)\underline{u}_\delta(x),$$

后一项的为 0 因为 $\operatorname{supp}\left(a'\left(\dfrac{2x_n}{\delta}\right)\right)\cap\operatorname{supp}(\underline{u}_\delta)=\emptyset$。所以，

$$\widetilde{\varphi}_p \overset{H^1(\mathbb{R}^n)}{\longrightarrow} \underline{u}_\delta,\quad p\to\infty.$$

从而，（看支集）

$$H_0^1(\mathbb{H}^n)\supset \widetilde{\varphi}_p \overset{H^1(\mathbb{H}^n)}{\longrightarrow} v_\delta,\quad p\to\infty.$$

这就说明了 $\operatorname{Res}(u)=0$ 当且仅当 $u\in H_0^1(\mathbb{H}^n)$。

最终，我们还需要证明迹映射

$$\operatorname{Res}: H^1(\mathbb{H}^n)\twoheadrightarrow H^{\frac{1}{2}}\left(\partial(\mathbb{H}^n)\right)$$

是满射。实际上，我们已经证明了

$$\operatorname{Res}: H^1(\mathbb{R}^n)\twoheadrightarrow H^{\frac{1}{2}}\left(\partial(\mathbb{H}^n)\right)$$

是满射。所以，对任意的 $v\in H^{\frac{1}{2}}\left(\partial(\mathbb{H}^n)\right)$，我们选取 $u\in H^1(\mathbb{H}^n)$ 为它的一个原像，那么 $u\big|_{\mathbb{H}^n}$ 就是 v 在 $H^1(\mathbb{H}^n)$ 中的一个原像。

综合上述，命题得证。 □

75 高指数 Sobolev 空间从半空间到全空间的扩张，Sobolev 函数的局部刻画，从 $H^1(\Omega)$ 到 $H^1(\mathbb{R}^n)$ 扩张，用全空间上的光滑函数逼近的 $H^1(\Omega)$，曲面上的 Sobolev 空间的定义

二零二零年十一月十八日，星期三，晴

前面的课程中，我们对 $H^1(\mathbb{H}^n)$ 中的函数构造了到边界的限制映射并且证明如下的正合列：

$$0 \to H_0^1(\mathbb{H}^n) \xrightarrow{\iota} H^1(\mathbb{H}^n) \xrightarrow{\mathrm{Res}} H^{\frac{1}{2}}\left(\partial(\mathbb{H}^n)\right) \to 0.$$

实际上，这些结论对于更高的正则性 H^k 也成立，其中 $k \geqslant 2$。我们回忆一下证明的要点：

- 限制映射（连续线性，满射）：

  $$\mathrm{Res} : H^s(\mathbb{R}^n) \twoheadrightarrow H^{s-\frac{1}{2}}(\mathbb{R}^{n-1})$$

  对所有的正则性 $s > \dfrac{1}{2}$ 都成立。

- 用全空间上的连续函数逼近上半空间上的 H^1 函数的引理：对任意的 $u \in H^1(\mathbb{H}^n)$，存在 $\{\varphi\}_{k\geqslant 1} \subset C^\infty(\mathbb{R}^n)$，使得

  1) 对任意的 $k \geqslant 1$，$\varphi_k\big|_{\mathbb{H}^n} \in H^1(\mathbb{H}^n)$；

  2) u 可以被 $\varphi_k\big|_{\mathbb{H}^n}$ 逼近：

  $$\lim_{k\to\infty} \left\|\varphi_k\big|_{\mathbb{H}^n} - u\right\|_{H^1(\mathbb{H}^n)} = 0.$$

  这个引理的证明对于任意的 $u \in H^k(\mathbb{H}^n)$ 是一样的（同学们可以自行验证细节）。所以，我们说

  $$H^k(\mathbb{H}^n) \cap C^\infty(\mathbb{R}^n) \subset H^k(\mathbb{H}^n)$$

  是稠密的。

- 关于 H^k 的函数从 $\mathbb{H}$ 到 $\mathbb{R}^n$ 的扩张定理，其中 $k \geqslant 1$。

  在 $k = 1$ 时，我们用了对称的延拓 $\mathrm{Ext}_{\mathrm{Sym}}$。

  $$\begin{array}{ccc} H^1\left(\mathbb{H}^n\right) & \xrightarrow{\mathrm{Ext}_{\mathrm{Sym}}} & H^1\left(\mathbb{R}^n\right) \\ & \searrow^{\mathrm{id}} & \downarrow{\scriptstyle \mathrm{Res}} \\ & & H^1\left(\mathbb{H}^n\right) \end{array}$$

  我们的证明是假设 u 是 $\mathbb{R}^n$ 上的一个光滑函数 F 的限制（一般情况逼近即可）：$u = F\big|_{\mathbb{H}^n}$。所谓的对称延拓是：

  $$\mathrm{Ext}_{\mathrm{Sym}} : C^\infty\left(\mathbb{R}^n\right) \cap H^1\left(\mathbb{H}^n\right) \to H^1\left(\mathbb{H}^n\right), \quad u \mapsto \widetilde{u}(x', x_n) = \begin{cases} u(x', x_n), & x_n \geqslant 0; \\ u(x', -x_n), & x_n \leqslant 0. \end{cases}$$

为了计算它的导数，我们需要验证跳跃公式中的要求，其中，$\widetilde{u}$ 在 $x_n=0$ 处的连续性是重要的。根据跳跃公式，我们证明了

$$\partial_j\left(\mathrm{Ext}_{\mathrm{Sym}}(u)\right)=\mathrm{Ext}_{\mathrm{Sym}}\left(\partial_k u\right),\quad k=1,2,\cdots,n-1.$$

并且

$$\partial_n\left(\mathrm{Ext}_{\mathrm{Sym}}(u)\right)=\left(\mathbf{1}_{x_n>0}-\mathbf{1}_{x_n<0}\right)\left(\mathrm{Ext}_{\mathrm{Sym}}(\partial_n u)\right).$$

从而，我们有

$$\|\mathrm{Ext}_{\mathrm{Sym}}(u)\|_{H^1(\mathbb{R}^n)}\leqslant\sqrt{2}\|u\|_{H^1(\mathbb{H}^n)}.$$

由于 $C^\infty(\mathbb{R}^n)\cap H^1(\mathbb{H}^n)$ 在 $H^1(\mathbb{H}^n)$ 是稠密的，

然而，我们观察到

$$\partial_n\left(\mathrm{Ext}_{\mathrm{Sym}}(u)\right)=\left(\mathbf{1}_{x_n>0}-\mathbf{1}_{x_n<0}\right)\left(\mathrm{Ext}_{\mathrm{Sym}}(\partial_n u)\right)$$

在 $x_n=0$ 出并不连续，所以，如果要对 H^k 的函数做同样的延拓，那么，在继续进行第二次求导数的时候，跳跃公式已经不再适用。所以，我们需要适当地修改延拓方式。不管怎么样，如下的定理是成立的：

引理 502 (从 $H^k(\mathbb{H})$ 到 $H^k(\mathbb{R}^n)$ 扩张)**.** *假设 $k\geqslant 1$ 是整数，那么限制映射*

$$\mathrm{Res}:H^k\left(\mathbb{R}^n\right)\to H^k\left(\mathbb{H}^n\right)$$

是连续的满的线性映射击并且存在连续线性的扩张映射

$$\mathrm{Ext}:H^k\left(\mathbb{R}^n\right)\to H^k\left(\mathbb{H}^n\right)$$

使得

$$\mathrm{Res}\circ\mathrm{Ext}=\mathrm{id}_{H^k(\mathbb{H}^n)}.$$

证明： 我们构造一种延拓方式：假设 u 是 $\mathbb{R}^n$ 上光滑函数的限制

$$\mathrm{Ext}:C^\infty\left(\mathbb{R}^n\right)\cap H^1\left(\mathbb{H}^n\right)\to H^1\left(\mathbb{H}^n\right),\quad u\mapsto\widetilde{u}(x',x_n)=\begin{cases}u(x',x_n), & x_n\geqslant 0;\\ \displaystyle\sum_{0\leqslant j\leqslant k-1}a_j u(x',-b_j x_n), & x_n<0.\end{cases}$$

在这个表达式中，我们要求 $0<b_1<b_2<\cdots<b_k$ 是任意给定的正实数，$\{a_j\}_{0\leqslant j\leqslant k}$ 目前待定。

此时，对关于前面 $n-1$ 个分量的任意多重指标 α'，我们都有

$$\partial_{x'}^{\alpha'}\widetilde{u}(x',x_n)=\begin{cases}\partial_{x'}^{\alpha'}u(x',x_n), & x_n\geqslant 0;\\ \displaystyle\sum_{0\leqslant j\leqslant k-1}a_j\partial_{x'}^{\alpha'}u(x',-b_j x_n), & x_n<0.\end{cases}$$

再对 x_n 份量求 ℓ 次导数（后面的证明我们来验证这一点），我们就得到

$$\partial_{x_n}^{\ell}\partial_{x'}^{\alpha'}\widetilde{u}(x',x_n)=\begin{cases}\partial_{x_n}^{\ell}\partial_{x'}^{\alpha'}u(x',x_n), & x_n\geqslant 0;\\ \displaystyle\sum_{0\leqslant j\leqslant k-1}a_j(-b_j)^{\ell}\partial_{x_n}^{\ell}\partial_{x'}^{\alpha'}u(x',-b_jx_n), & x_n<0.\end{cases}$$

为了保证这个函数在 $x_n=0$ 处是连续的，我们需要

$$\partial_{x_n}^{\ell}\partial_{x'}^{\alpha'}u(x',0)=\sum_{0\leqslant j\leqslant k-1}a_j(-b_j)^{\ell}\partial_{x_n}^{\ell}\partial_{x'}^{\alpha'}u(x',0),$$

即

$$\sum_{0\leqslant j\leqslant k-1}a_j(-b_j)^{\ell}=1.$$

我们要求上面的等式对 $\ell=0,1,\cdots,k-1$ 都成立，此时，利用 Vandermonde 行列式，我们知道存在唯一的 $a_0,\cdots,a_{k-1}$ 使得上面的等式均成立。

此时，对任意的 α' 和 ℓ，只要 $0\leqslant\ell\leqslant k-1$，$|\alpha'|+\ell\leqslant k$，$\partial_{x_n}^{\ell}\partial_{x'}^{\alpha'}\widetilde{u}(x',x_n)$ 在 $x_n=0$ 处都是连续的，那么，我们就可以用跳跃公式求它的导数并且保证它们在边界上的贡献恰好消掉。我们下面只计算它们对于 ∂_n 方向的导数，其余方向的导数是更加简单的：

$$\begin{aligned}&\partial_n\left(\partial_{x_n}^{\ell}\partial_{x'}^{\alpha'}\widetilde{u}(x',x_n)\right)\\ =&\partial_n\left(\partial_{x_n}^{\ell}\partial_{x'}^{\alpha'}u(x',x_n)\mathbf{1}_{x_n\geqslant 0}\right)+\partial_n\left(\sum_{0\leqslant j\leqslant k-1}a_j(-b_j)^{\ell}\partial_{x_n}^{\ell}\partial_{x'}^{\alpha'}u(x',-b_jx_n)\mathbf{1}_{x_n<0}\right)\\ =&\partial_{x_n}^{\ell+1}\partial_{x'}^{\alpha'}u(x',x_n)\mathbf{1}_{x_n\geqslant 0}-\partial_{x_n}^{\ell}\partial_{x'}^{\alpha'}u(x',0)d\sigma_{x_n=0}\\ &+\sum_{0\leqslant i\leqslant k-1}a_i(-b_i)^{\ell+1}\partial_{x_n}^{\ell+1}\partial_{x'}^{\alpha'}u(x',-b_jx_n)\mathbf{1}_{x_n<0}+\sum_{0\leqslant i\leqslant k-1}a_i(-b_i)^{\ell}\partial_{x_n}^{\ell}\partial_{x'}^{\alpha'}u(x',0)d\sigma_{x_n=0}\\ =&\partial_{x_n}^{\ell+1}\partial_{x'}^{\alpha'}u(x',x_n)\mathbf{1}_{x_n\geqslant 0}+\sum_{0\leqslant i\leqslant k-1}a_i(-b_i)^{\ell+1}\partial_{x_n}^{\ell+1}\partial_{x'}^{\alpha'}u(x',-b_jx_n)\mathbf{1}_{x_n<0}.\end{aligned}$$

这就给出对所有的 $|\alpha'|+\ell\leqslant k$ 的导数的计算，很明显，

$$\begin{aligned}&\left\|\partial_{x_n}^{\ell}\partial_{x'}^{\alpha'}u(x',x_n)\mathbf{1}_{x_n\geqslant 0}+\sum_{0\leqslant j\leqslant k-1}a_j(-b_j)^{\ell}\partial_{x_n}^{\ell}\partial_{x'}^{\alpha'}u(x',-b_jx_n)\mathbf{1}_{x_n<0}\right\|_{L^2(\mathbb{R}^n)}\\ \leqslant&C\left\|\partial_{x_n}^{\ell}\partial_{x'}^{\alpha'}u(x',x_n)\right\|_{L^2(\mathbb{H}^n)}.\end{aligned}$$

这就证明了延拓的连续性。 □

- 基于上面几个引理，我们就可以重复上次的证明来说明对任意的整数 $k\geqslant 1$，我们都有如下的正合列：

$$0\to H_0^k(\mathbb{H}^n)\xrightarrow{\iota}H^k(\mathbb{H}^n)\xrightarrow{\mathrm{Res}}H^{k-\frac{1}{2}}(\partial(\mathbb{H}^n))\to 0.$$

我们现在给出 Sobolev 函数的局部刻画。假设 Ω 是有界光滑带边的区域（$\overline{\Omega}$ 是紧集），$\mathcal{U}=\{U_j\}_{j\leqslant N}$ 是 $\overline{\Omega}$ 的一个开覆盖，$\{\chi_j\}_{j\leqslant N}$ 是相应的一族单位分解。

命题 503. *给定 $u\in\mathcal{D}'(\Omega)$，那么如下等价：*

1) $u\in H^1(\Omega)$；

2) 对任意的 $j\leqslant N$，$\chi_j\cdot u\in H^1(\Omega)$。

证明： 先证明 1) $\Rightarrow$ 2)：很明显，我们有

$$\|\chi_j u\|_{L^2}\leqslant C\|u\|_{L^2},$$

根据分布求导的 Leibniz 公式，我们有

$$\|\nabla(\chi_j u)\|_{L^2}=\|(\nabla\chi_j)u+\chi_j\nabla u\|_{L^2}\leqslant C\left(\|u\|_{L^2}+\|\nabla u\|_{L^2}\right).$$

所以，这个方向是成立的。

再证明 2) $\Rightarrow$ 1)：由于

$$u=\sum_{j\leqslant N}\chi_j\cdot u,$$

是有限个 $H^1(\Omega)$ 中元素的线性组合，所以落在 $H^1(\Omega)$ 中。 □

注记. *这个命题对于空间 $H_0^1(\Omega)$ 也成立：假设 $\{\varphi_p\}_{p\geqslant 1}\subset C_0^\infty(\Omega)$，使得 $\lim\limits_{p\to\infty}\varphi_p\overset{H^1}{=}u$，那么，*

$$\lim_{p\to\infty}\chi_j\cdot\varphi_p\overset{H^1}{=}\chi_j\cdot u.$$

证明是平凡的。

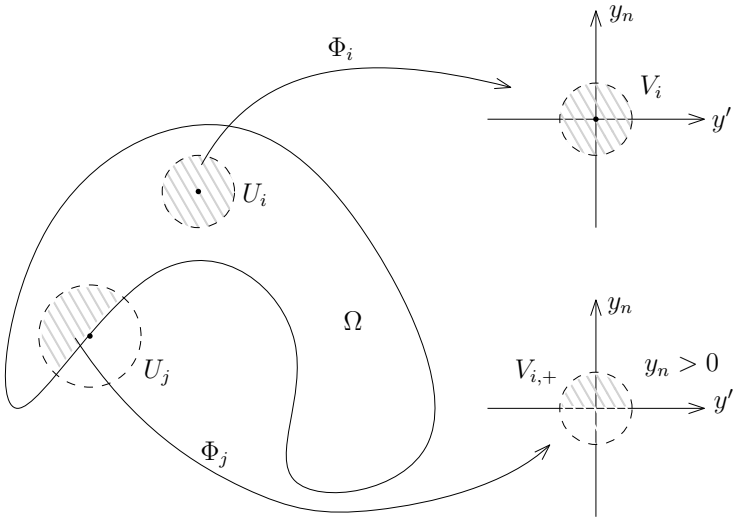
对于 $x\in\overline{\Omega}$，我们选取包含 x 的一个开的小球 U，这些开集覆盖了 $\overline{\Omega}$。对于 $x\in\partial\Omega$，我们要求 U 选取的足够小，使得存在微分同胚

$$\Phi:U\to V\subset\mathbb{R}^{n-1}\times\mathbb{R},\quad x\mapsto(y',y_n)=(\Phi'(x),\Phi_n(x)),$$

其中，V 是 $\mathbb{R}^n_y$ 中的开球，并且

- $\Phi(U\cap\Omega)=V_+=V\cap\{y',y_n)|y_n>0\}$；
- $\Phi(x)$ 的 y_n 坐标是 0。

此时，$\Phi(U\cap\partial\Omega)\subset\{y_n=0\}$ 被实现为 $\mathbb{R}^n$ 的超平面。



从这些开集中，我们可以选取 $\overline{\Omega}$ 的一个有限开覆盖 $\{U_j\}_{j\leqslant N}$，与之对应的 Φ 我们记做是 Φ_j。我们注意到，这些 U_j 们分两种，一种是边界点给出的，另一种是内点给出的（与边界不相交）。

利用与 $\{U_j\}_{j\leqslant N}$ 相适应的单位分解 $\{\chi_j\}_{j\leqslant N}$，对任意的 $u\in H^1(\Omega)$ 和 $u_0\in H_0^1(\Omega)$，我们知道 $\chi_j\cdot u\in H^1(\Omega)$，$\chi_j\cdot u_0\in H_0^1(\Omega)$，通过复合，我们得到了 $(\chi_j\cdot u)\circ\Phi_j^{-1}$ 和 $(\chi_j\cdot u_0)\circ\Phi_j^{-1}$，这都是 V_j 或者 $V_{j,+}$ 上的函数。我们不妨研究 u 和 $V_{j,+}$ 的情形（其余类似）。我们现在来说明

$$(\chi_j\cdot u)\circ\Phi_j^{-1}\in H_0^1(V_{j,+}).$$

以下我们用 $U_{j,+}$ 表示 $U_j\cap\Omega$。根据换元积分公式，我们有

$$\int_{V_{j,+}}\left|(\chi_j\cdot u)\circ\Phi_j^{-1}(y)\right|^2dy=\int_{U_{j,+}}|(\chi_j\cdot u)(x)|^2\,|\mathrm{Jac}(\Phi_j)(x)|dx.$$

由于上述 Jacobi 行列式是有界的，所以

$$\int_{V_{j,+}}\left|(\chi_j\cdot u)\circ\Phi_j^{-1}(y)\right|^2dy\leqslant C\int_{U_{j,+}}|(\chi_j\cdot u)(x)|^2\,dx.$$

另外，根据链式求导公式，我们有

$$\left|\partial_k\left((\chi_j\cdot u)\circ\Phi_j^{-1}(y)\right)\right|=\left|\sum_{\ell\leqslant n}\partial_\ell(\chi_j\cdot u)\cdot\partial_k\left(\Phi_j^{-1}{}_\ell\right)\right|\leqslant C\left|\nabla(\chi_j\cdot u)\right|.$$

所以，重复上面的计算就知道

$$\int_{V_{j,+}}\left|\nabla\left((\chi_j\cdot u)\circ\Phi_j^{-1}(y)\right)\right|^2dy\leqslant C\int_{U_{j,+}}|\nabla(\chi_j\cdot u)(x)|^2\,dx.$$

综合上面的不等式，实际上我们证明了

$$(\chi_j\cdot u)\circ\Phi_j^{-1}\in H_0^1(V_{j,+})\quad\Leftrightarrow\quad\chi_j\cdot u\in H_0^1(U_{j,+}).$$

通过限制，显然有连续线性映射

$$\mathrm{Res} : H^1(\mathbb{R}^n) \to H^1(\Omega).$$

利用上述关于 Sobolev 函数的局部描述，我们可以构造扩张映射：

引理 504 (从 $H^1(\Omega)$ 到 $H^1(\mathbb{R}^n)$ 扩张)**.** 限制映射

$$\mathrm{Res} : H^1(\mathbb{R}^n) \to H^1(\Omega)$$

是连续的满的线性映射击并且存在连续线性的扩张映射

$$\mathrm{Ext} : H^1(\Omega) \to H^1(\mathbb{H}^n)$$

使得

$$\mathrm{Res} \circ \mathrm{Ext} = \mathrm{id}_{H^1(\Omega)}.$$

注记. 上述扩张映射未必是唯一的，它依赖于局部坐标描述的选取等等，请参考证明。

证明: 我们把上述构造的覆盖 $\mathcal{U} = \{U_j\}_{j \leqslant N}$ 分成两个部分

$$\mathcal{U} = \mathcal{U}' \cup \mathcal{U}'' = \{U'_{j'}\}_{j' \leqslant N'} \cup \{U''_{j''}\}_{j'' \leqslant N''}.$$

其中，$\mathcal{U}'$ 中的开集和边界不相交，$\mathcal{U}''$ 中的开集和边界相交。那么，对任意的 $u \in H^1(\Omega)$，我们有

$$u = \sum_{j' \leqslant N'} \chi'_{j'} u + \sum_{j'' \leqslant N''} \chi''_{j''} u.$$

我们只要能对每个 $\chi'_{j'}u$ 和 $\chi''_{j''}u$ 进行延拓即可。

首先，由于 $\chi'_{j'}u$ 的支集在 Ω 中，所以我们用 0 进行延拓即可。下面研究 $\chi''_{j''}u$，为了书写简洁，我们仍然用 $\chi_j u$ 表示这个函数，那么，

$$v_j = (\chi_j u) \circ \Phi_j^{-1} \in H^1(V_{j,+}).$$

在构造这些覆盖的时候，我们总是可以要求 V_j 是对称的，即 $V_j = V_{j,+} \cup (-V_{j,+})$。对于 v_j 而言，我们可以用之前构造的对称的扩张 $\mathrm{Ext}_{\mathrm{sym}}(v_j)$。此时，由于 $\mathrm{supp}(v_j) \subset V_{j,+}$，所以，

$$\mathrm{supp}\left(\mathrm{Ext}_{\mathrm{sym}}(v_j)\right) \subset V_j = \Phi_j(U_j)$$

并且是 H^1 的函数。那么，

$$\widetilde{u_j} = \left(\mathrm{Ext}_{\mathrm{sym}}(v_j)\right) \circ \Phi_j^{-1}$$

是支集在 U_j 中的具有紧支集的 $H^1(U_j)$ 中的函数，我们可以再把它延拓到 $\mathbb{R}^n$ 上去。

由于上述的覆盖是事先选定的，所以，上述构造的算子是线性的（每一步都是确定的）。 □

作为推论，我们有

推论 505 (用全空间上的光滑函数逼近的 $H^1(\Omega)$)**.** *对任意的 $u \in H^1(\Omega)$，存在 $\{\varphi\}_{k\geqslant 1} \subset C_0^\infty(\mathbb{R}^n)$，使得 u 可以被 $\left\{\varphi_k\big|_\Omega\right\}_{k\geqslant 1}$ 逼近:*

$$\lim_{k\to\infty} \left\|\varphi_k\big|_\Omega - u\right\|_{H^1(\Omega)} = 0.$$

证明: 首先，在全空间中，我们能够找到 $\{\varphi\}_{k\geqslant 1} \subset C_0^\infty(\mathbb{R}^n)$，使得

$$\lim_{k\to\infty} \|\varphi_k - \mathrm{Ext}(u)\|_{H^1(\mathbb{R}^n)} = 0.$$

所以，

$$\lim_{k\to\infty} \left\|\varphi_k\big|_\Omega - \mathrm{Ext}(u)\big|_\Omega\right\|_{H^1(\Omega)} = 0.$$

而 $\mathrm{Ext}(u)\big|_\Omega = u$，这就给出了证明。 □

我们现在来定义 $H^{\frac{1}{2}}(\partial\Omega)$ 空间。

一种可行的方式是利用局部覆盖：我们上面构造了一族关于 $\overline{\Omega}$ 的覆盖，我们现在只考虑那些和 $\partial\Omega$ 相交的覆盖，不妨还记作是 $\mathcal{U} = \{U_j\}_{j\leqslant N}$，那么，对每个 $i \leqslant N$，我们有微分同胚

$$\phi_j : U_i \cap (\partial\Omega) \to V' \subset \mathbb{R}^{n-1}.$$

这是之前的 Φ_i 在 $U_i \cap \partial\Omega$ 上的限制。假设 $u \in L^2(\partial\Omega)$，利用与 $\mathcal{U}$ 相适应的单位分解 $\{\chi_j\}_{j\leqslant}$，我们就有

$$u = \sum_{j\leqslant N} \chi_j \cdot u = \sum_{j\leqslant N} u_j,$$

其中 $\mathrm{supp}(u_j) \subset U_j \cap \partial\Omega$。

那么，当我们说 $u \in H^{\frac{1}{2}}(\partial\Omega)$ 指的是对任意的 $j \leqslant N$，函数 $u_j \circ \phi_j^{-1}$ 作为 V_j' 上的函数是 $H^{\frac{1}{2}}$ 的，这里，因为 $\mathrm{supp}(u_j \circ \phi_j^{-1}) \subset V_j'$，所以，我们可以通过延拓把 $u_j \circ \phi_j^{-1}$ 看作是 $\mathbb{R}^{n-1}$ 上的函数，这就可以谈论 $u_j \circ \phi_j^{-1}$ 是否是 $H^{\frac{1}{2}}$ 的。在此基础上，我们定义

$$\|u\|_{H^{\frac{1}{2}}(\partial\Omega),\mathcal{U}} = \sum_{j\leqslant N} \|u_j \circ \phi_j^{-1}\|_{H^{\frac{1}{2}}(\mathbb{R}^{n-1})}.$$

当然，这样定义的 $\|u\|_{H^{\frac{1}{2}}(\partial\Omega)}$ 依赖于覆盖 $\mathcal{U}$ 以及单位分解的选取。我们不难证明（习题），如果采取另一种覆盖 $\mathcal{U}$ 和单位分解，新定义出来的范数与这种定义是等价的，所以，它们所定义的收敛（拓扑）概念是一致的。

我们现在采取另一种方式来定义 $H^{\frac{1}{2}}(\partial\Omega)$ 并且证明上面定义的范数与将要定义的范数是等价的（这就给出了 $\|u\|_{H^{\frac{1}{2}}(\partial\Omega)}$ 不依赖于覆盖 $\mathcal{U}$ 以及单位分解的选取的一种证明）。这个定义依赖于我们已经证明过的命题：当 Sobolev 指标 $0 < s < 1$ 时，$u \in L^2(\mathbb{R}^n)$，那么 $u \in H^s(\mathbb{R}^n)$ 当且仅当

$$\iint_{\mathbb{R}^n\times\mathbb{R}^n} \frac{|u(x)-u(y)|^2}{|x-y|^{n+2s}} dxdy < \infty.$$

其中，我们

$$\|u\|^2_{H^s(\mathbb{R}^n)} \approx \|u\|^2_{L^2(\mathbb{R}^n)} + \iint_{\mathbb{R}^n\times\mathbb{R}^n} \frac{|u(x)-u(y)|^2}{|x-y|^{n+2s}} dxdy.$$

定义 506. 我们用 $d\sigma$ 表示 $\partial\Omega$ 上的曲面测度。对于任意的 $u\in L^2(\partial\Omega)$，如果如下的积分是有限的（注意到 $\dim\partial\Omega=n-1$）：

$$\iint_{\partial\Omega\times\partial\Omega}\frac{|u(x)-u(y)|^2}{|x-y|^n}d\sigma(x)d\sigma(y)<\infty,$$

我们就说 $u\in H^{\frac12}(\partial\Omega)$。对于 $u\in H^{\frac12}(\partial\Omega)$，我们定义范数

$$\|u\|^2_{H^{\frac12}(\partial\Omega)}=\|u\|^2_{L^2(\Omega)}+\iint_{\partial\Omega\times\partial\Omega}\frac{|u(x)-u(y)|^2}{|x-y|^n}d\sigma(x)d\sigma(y).$$

我们可以很容易地验证三角不等式，从而，我们得到了一个范数。

对任意的 $\delta>0$，很明显，这个范数的定义等价于

$$\|u\|^2_{H^{\frac12}(\partial\Omega),\delta}=\|u\|^2_{L^2(\partial\Omega)}+\iint_{\partial\Omega\times\partial\Omega}\frac{|u(x)-u(y)|^2}{|x-y|^n}\mathbf{1}_{|x-y|<\delta}d\sigma(x)d\sigma(y).$$

（差一个大约是 $\delta^{-(n+1)}$ 的常数）

我们选取 $\partial\Omega$ 的开覆盖 $\mathfrak{U}$，我们要求对任意的 $U\in\mathfrak{U}$，都有 $\operatorname{diam}(U)<\frac12\delta$（直径）。此时，我们有

$$\|u\|^2_{H^{\frac12}(\Omega)}\leqslant\sum_{j\leqslant N}\|\chi_j\cdot u\|^2_{H^{\frac12}(\Omega)}.$$

另外，

$$\begin{aligned}
&\iint_{\partial\Omega\times\partial\Omega}\frac{|\chi_j(x)u(x)-\chi_j(y)u(y)|^2}{|x-y|^n}d\sigma(x)d\sigma(y)\\
\leqslant&2\iint_{\partial\Omega\times\partial\Omega}\frac{|\chi_j(x)|^2|u(x)-u(y)|^2}{|x-y|^n}d\sigma(x)d\sigma(y)+2\iint_{\partial\Omega\times\partial\Omega}\frac{|\chi_j(x)-\chi_j(y)|^2}{|x-y|^n}|u(y)|^2d\sigma(x)d\sigma(y)\\
\leqslant&2\iint_{\partial\Omega\times\partial\Omega}\frac{|u(x)-u(y)|^2}{|x-y|^n}d\sigma(x)d\sigma(y)+C\iint_{\partial\Omega\times\partial\Omega}\frac{1}{|x-y|^{n-2}}|u(y)|^2d\sigma(x)d\sigma(y).
\end{aligned}$$

在上面的估计中，根据 Lagrange 中值定理，我们用 $C|x-y|$ 来控制了 $\chi_j(x)-\chi_j(y)$（这里用到了截断函数的光滑性，u 没有这种光滑性）。

所以，

$$\|\chi_j\cdot u\|^2_{H^{\frac12}(\Omega)}\leqslant C\|u\|^2_{H^{\frac12}(\Omega)}+C\int_{\partial\Omega}|u(y)|^2\left(\int_{\partial\Omega}\frac{1}{|x-y|^{n-2}}d\sigma(x)\right)d\sigma(y).$$

练习. 证明，存在常数 C，使得对任意的 y，我们都有

$$\int_{\partial\Omega}\frac{1}{|x-y|^{n-2}}d\sigma(x)\leqslant C.$$

据此，

$$\|\chi_j\cdot u\|_{H^{\frac12}(\Omega)}\leqslant C\|u\|^2_{H^{\frac12}(\Omega)}.$$

所以，如果我们定义

$$|||u|||^2_{H^{\frac{1}{2}}(\Omega),\mathfrak{U}} = \sum_{j\leqslant N} \|\chi_j \cdot u\|^2_{H^{\frac{1}{2}}(\Omega)},$$

那么，范数 $|||\cdot|||_{H^{\frac{1}{2}}(\Omega),\mathfrak{U}}$ 与范数 $\|\cdot\|_{H^{\frac{1}{2}}(\Omega)}$ 是等价的。我们下面说明它们都与下面已经定义的范数等价：

$$\|u\|_{H^{\frac{1}{2}}(\partial\Omega),\mathfrak{U}} = \sum_{j\leqslant N} \|u_j \circ \phi_j^{-1}\|_{H^{\frac{1}{2}}(\mathbb{R}^{n-1})}.$$

75.1 作业：二维波动方程的基本解，Airy 函数与线性 KdV 方程

清华大学 19-20 春季学期，数学分析三，作业 5

请用 A4 大小的纸张**正反面**用钢笔，签字笔或者圆珠笔书写，并注明自己的姓名，年级和作业的**总页数**。除定理公式所涉及的人名之外，请使用**中文**。本次作业的提交时间和地点为 **12 月 2 日**（周三）上午的课堂上，逾期视作零分。

An ocean traveller has even more vividly the impression that the ocean is made of waves than that it is made of water.

— Sir Arthur Stanley Eddington

习题 A.（二维波动方程的基本解）

我们考虑定义在 $\mathbb{R}^{1+2}=\mathbb{R}\times\mathbb{R}^2$ 上的波动算子 $\Box=-\partial_t^2+\triangle$，其中，第一个坐标是时间 t 的坐标。这个算子作用在以 $(t,x,y)\in\mathbb{R}\times\mathbb{R}^3$ 为变量函数上的。我们令

$$\rho:\mathbb{R}\times\mathbb{R}^2\to\mathbb{R},\quad (t,x,y)\mapsto t^2-x^2-y^2.$$

A1) 假设 $f=F\circ\rho$ 是 $\mathbb{R}^{1+2}$ 上只依赖于 ρ 的函数，其中 $F(s)$ 是 $\mathbb{R}$ 上的光滑函数。令

$$L=4s\frac{d^2}{ds^2}+6\frac{d}{ds}.$$

证明，

$$\Box f=-L(F)\circ\rho.$$

A2) 对任意的 $\varepsilon>0$，考虑 $\mathbb{R}^1$ 上的函数

$$H_\varepsilon(s)=\mathbf{1}_{(\varepsilon,\infty)}(s).$$

试计算 $L\left(\dfrac{H_\varepsilon(s)}{\sqrt{s}}\right)$。

A3) 我们考虑 $\mathbb{R}^{1+2}$ 上的指向未来的实心光锥

$$\mathring{C}_+=\left\{(t,x)\in\mathbb{R}^{1+3}\middle|x^2+y^2<t^2\right\}.$$

定义 $\mathbb{R}^{1+2}$ 上的函数

$$E(t,x,y)=-\frac{\mathbf{1}_{\mathring{C}_+}(t,x,y)}{2\pi\sqrt{t^2-(x^2+y^2)}}.$$

证明，$E\in L^1_{\mathrm{loc}}(\mathbb{R}^{1+2})$。

A4) 证明，在极坐标 (t,r,ϑ) 下，对任意的 $\varphi\in\mathcal{D}(\mathbb{R}^{1+2})$，我们有

$$\langle\Box E,\varphi\rangle=\frac{1}{2\pi}\int_0^\infty\left(\int_0^\infty\int_0^{2\pi}\frac{(\Box\varphi)(\sqrt{s+r^2},r\cos\vartheta,r\sin\vartheta)}{2\sqrt{s+t^2}}rdrd\vartheta\right)ds.$$

A5) 我们定义算子

$$L^*=4s\frac{d^2}{ds^2}+2\frac{d}{ds}.$$

证明，对任意的分布 $u\in\mathcal{D}'(\mathbb{R})$，对任意的试验函数 $\varphi\in\mathcal{D}(\mathbb{R})$，我们有

$$\langle Lu,\varphi\rangle=\langle u,L^*\varphi\rangle.$$

A6) 对任意的试验函数 $\varphi\in\mathcal{D}(\mathbb{R}^{1+2})$，我们定义 $\mathbb{R}$ 上的（关于变量 s）函数

$$\widetilde{\varphi}(s)=\int_0^\infty\int_0^{2\pi}\frac{(\Box\varphi)(\sqrt{s+r^2},r\cos\vartheta,r\sin\vartheta)}{2\sqrt{s+t^2}}rdrd\vartheta.$$

证明，

$$\left(\widetilde{\Box\varphi}\right)(s)=(L^*\widetilde{\varphi})(s).$$

证明，

$$\langle\Box E,\varphi\rangle=-\frac{2}{\pi}\lim_{\varepsilon\to0}\sqrt{\varepsilon}\frac{d}{ds}\widetilde{\varphi}(s).$$

A7) 对任意的试验函数 $\varphi\in\mathcal{D}(\mathbb{R}^{1+2})$，我们定义其球面平均为如下 $\mathbb{R}\times\mathbb{R}_{>0}$ 上的（关于变量 (t,r)）函数：

$$\overline{\varphi}(t,r)=\int_0^{2\pi}\varphi(t,r,\vartheta)d\vartheta.$$

证明，

$$\frac{d\widetilde{\varphi}}{ds}(\varepsilon)=\frac{1}{4}\left(\int_0^\infty\frac{\partial\overline{\varphi}}{\partial t}(\sqrt{r^2+\varepsilon},r)\frac{rdr}{r^2+\varepsilon}-\int_0^\infty\overline{\varphi}(\sqrt{r^2+\varepsilon},r)\frac{rdr}{(r^2+\varepsilon)^{\frac{3}{2}}}\right).$$

A8) 证明，

$$\frac{d\widetilde{\varphi}}{ds}(\varepsilon)=-\frac{1}{4\sqrt{\varepsilon}}\overline{\varphi}(\sqrt{\varepsilon},0)+R(\varepsilon),$$

其中 $R(\varepsilon)$ 由给定的试验函数 φ 决定并且满足

$$\lim_{\varepsilon\to0}\sqrt{\varepsilon}R(\varepsilon)=0.$$

A9) 证明，E 是波动算子的基本解。

习题 B.（Airy 函数与线性 KdV 方程）

这个问题的目的是要研究线性 KdV 方程（浅水波方程）的基本解。

B1) 证明，存在唯一的缓增分布 $\mathbf{Ai} \in \mathcal{S}'(\mathbb{R})$，使得其 Fourier 变换为 $e^{\frac{1}{3}i\xi^3}$。我们这个分布称作是 Airy 函数。

B2) 证明，对任意给定 $\eta > 0$，关于 ξ 的函数

$$\xi \mapsto e^{\frac{1}{3}i(\xi+i\eta)^3}$$

是 Schwartz 函数。

B3) 证明，对任意给定 $\eta > 0$，关于 x 的函数:

$$x \mapsto \frac{1}{2\pi}\int_{\mathbb{R}} e^{ix(\xi+i\eta)+\frac{1}{3}i(\xi+i\eta)^3}d\xi,$$

是光滑函数。

B4) 证明，当 $n \to +\infty$ 时，关于 ξ 的函数序列

$$\left\{\xi \mapsto e^{\frac{1}{3}i(\xi+\frac{i}{n})^3}\right\}_{n\geqslant 1}$$

在 $\mathcal{S}'(\mathbb{R}_\xi)$ 中收敛并计算其极限。

B5) 计算

$$(\partial_\xi + i\partial_\eta)e^{ix(\xi+i\eta)+\frac{1}{3}i(\xi+i\eta)^3}$$

B6) 给定 $R > 0$ 和 $a < b$，在 $\mathcal{D}'(\mathbb{R}\times\mathbb{R}_{>0})$ 上计算下面两个分布:

$$\partial_\xi\big(\mathbf{1}_{[-R,R]}(\xi)\mathbf{1}_{[a,b]}(\eta)\big),\quad \partial_\eta\big(\mathbf{1}_{[-R,R]}(\xi)\mathbf{1}_{[a,b]}(\eta)\big).$$

B7) 证明:

$$\begin{aligned} &i\left(\int_{-R}^{R} e^{ix(\xi+ib)+\frac{1}{3}i(\xi+ib)^3}d\xi - \int_{-R}^{R} e^{ix(\xi+ia)+\frac{1}{3}i(\xi+ia)^3}d\xi\right)\\ &= \int_a^b e^{ix(-R+i\eta)+\frac{1}{3}i(-R+i\eta)^3}d\eta - \int_a^b e^{ix(R+i\eta)+\frac{1}{3}i(R+i\eta)^3}d\eta \end{aligned}$$

B8) 你是否可以用复解析函数的理论来解释 7) 中的结论?

B9) 证明，给定 $x \in \mathbb{R}$，下面关于 η 函数

$$\eta \mapsto \frac{1}{2\pi}\int_{\mathbb{R}} e^{ix(\xi+i\eta)+\frac{1}{3}i(\xi+i\eta)^3}d\xi$$

在 $\eta \in (0,+\infty)$ 上是常值函数。

B10) 证明，对任意的 $x \in \mathbb{R}$ 和 $\eta > 0$，在 1) 中定义的函数 $\mathbf{Ai}(x)$ 还可以由如下公式定义:

$$\mathbf{Ai}(x) = \frac{1}{2\pi}\int_{\mathbb{R}} e^{ix(\xi+i\eta)+\frac{1}{3}i(\xi+i\eta)^3}d\xi.$$

B11) 证明，对任意 $C>0$，我们都有

$$\lim_{x\to+\infty}|e^{Cx}\mathbf{Ai}(x)|=0.$$

B12) 证明，

$$\mathbf{Ai}(x)''-x\mathbf{Ai}(x)=0.$$

B13) 假定 $t>0$，哪一个缓增分布的 Fourier 变换是 $e^{\frac{1}{3}it\xi^3}$？

B14) 对每个 $(t,x)\in\mathbb{R}^2$，定义

$$E(t,x)=\frac{\mathbf{1}_{(0,+\infty)}(t)}{t^{\frac{1}{3}}}\mathbf{Ai}\left(\frac{x}{t^{\frac{1}{3}}}\right).$$

证明，$E(t,x)\in\mathcal{S}'(\mathbb{R})$。

B15) 证明

$$\big(\partial_t+\frac{1}{3}\partial_x^3\big)E(t,x)=\delta_{(t=0,x=0)}.$$

B16) 证明，$E(t,x)$ 是 $\mathcal{S}'(\mathbb{R}^2_{t,x})$ 中唯一一个满足如下两个条件的分布：

- $\big(\partial_t+\frac{1}{3}\partial_x^3\big)u=\delta_{(0,0)}$；
- $\mathrm{supp}(u)\subset\mathbb{R}_{\geqslant 0}\times\mathbb{R}$.

B17) 证明，对任意的 $u_0\in\mathcal{E}'(\mathbb{R})$，如下线性 KdV 方程

$$\begin{cases}\big(\partial_t+\frac{1}{3}\partial_x^3\big)u=0,\quad (t,x)\in\mathbb{R}_{>0}\times\mathbb{R},\\ u\big|_{t=0}=u_0\end{cases}$$

在 $\mathcal{S}'(\mathbb{R}\times\mathbb{R})$ 上有唯一一个满足 $u\in C^\infty(\mathbb{R}_{>0}\times\mathbb{R})$ 的解。特别地，请陈述 $u\big|_{t=0}=u_0$ 的具体定义。

76 子流形上的 Sobolev 空间的定义与等价定义，有界区域上 Sobolev 空间的限制（迹）定理以及限制正合列，Dirichlet 问题的边值的确切含义以及 Dirichlet 问题的解决，位势方程的边值问题的高阶椭圆正则性定理

二零二零年十一月二十三日，星期一，晴

我们上次课用两种方式定义了 $\partial\Omega$ 上的 $H^{\frac{1}{2}}(\Omega)$ 空间，一种是直接在物理空间里定义的，它所对应的范数是

$$\|u\|^2_{H^{\frac{1}{2}}(\partial\Omega)} = \|u\|^2_{L^2(\Omega)} + \iint_{\partial\Omega\times\partial\Omega} \frac{|u(x)-u(y)|^2}{|x-y|^n} d\sigma(x)d\sigma(y),$$

其中，我们假设了 $u \in L^2(\partial\Omega)$。

另一种利用了局部覆盖：假设 $\mathcal{U} = \{U_j\}_{j\leqslant N}$ 是 $\partial\Omega$ 的一个开覆盖，那么，对每个 $j \leqslant N$，我们有微分同胚

$$\Phi_j : U_j \cap \partial\Omega \to V_j^+ \subset \mathbb{R}^n_{x_n\geqslant 0},$$

是 $\partial\Omega$ 附近的微分同胚，而

$$\phi_j = \Phi_j\big|_{\partial\Omega\cap U_j} : \partial\Omega \cap U_j \to V_j \subset \mathbb{R}^{n-1}$$

是微分同胚。假设 $u \in L^2(\partial\Omega)$，利用与 $\mathcal{U}$ 相适应的单位分解 $\{\chi_j\}_{j\leqslant N}$，我们定义

$$\|u\|_{H^{\frac{1}{2}}(\partial\Omega),\mathcal{U}} = \sum_{j\leqslant N} \|u_j\circ\phi_j^{-1}\|_{H^{\frac{1}{2}}(\mathbb{R}^{n-1})},$$

其中 $u_j = \chi_j \cdot u$.

利用单位分解我们还可以定义第一种范数的一个局部的版本：

$$|||u|||^2_{H^{\frac{1}{2}}(\Omega),\mathcal{U}} = \sum_{j\leqslant N} \|\chi_j\cdot u\|^2_{H^{\frac{1}{2}}(\Omega)},$$

我们上次证明了 $|||\cdot|||_{H^{\frac{1}{2}}(\Omega),\mathcal{U}}$ 与范数 $\|\cdot\|_{H^{\frac{1}{2}}(\Omega)}$ 是等价的。

我们这次证明 $|||\cdot|||_{H^{\frac{1}{2}}(\Omega),\mathcal{U}}$ 与 $\|\cdot\|_{H^{\frac{1}{2}}(\Omega),\mathcal{U}}$ 等价。为此，只要对任意的 j，证明

$$\|u_j\circ\phi_j^{-1}\|_{H^{\frac{1}{2}}(\mathbb{R}^{n-1})} \approx \|u_j\|_{H^{\frac{1}{2}}(\partial\Omega)}$$

即可。这里，$\mathrm{supp}(u_j) \subset U_j \cap \Omega$ 是紧的。通过将覆盖加细，我们总是可以假设局部上 $\partial\Omega$ 是函数的图像。

我们利用微分同胚

$$\phi_j : U_i \cap (\partial\Omega) \to V' \subset \mathbb{R}^{n-1}.$$

其中，ϕ_j 是 Φ_j 在 $\partial\Omega$ 上的限制。实际上，我们有

$$\begin{aligned}\|u_j\|^2_{H^{\frac{1}{2}}(\partial\Omega)} &= \iint_{\phi_j^{-1}(V')\times\phi_j^{-1}(V')} \frac{|u_j(x)-u_j(x')|^2}{|x-x'|^n} d\sigma(x)d\sigma(x') \\ &\approx \iint_{V'\times V'} \frac{\left|\left(u_j\circ\phi_j^{-1}\right)(y)-\left(u_j\circ\phi_j^{-1}\right)(y')\right|^2}{\left|\phi_j^{-1}(y)-\phi_j^{-1}(y')\right|^n} dydy'\end{aligned}$$

在上一个约等于号中，我们实际上先把测度用函数图像的形式表达，然后利用再乘上坐标变换的 Jacobi 矩阵。这些操作只是在 dy 之前乘了一个上下有界的正的函数，所以我们有上面的约等于号。另外，由于 ϕ_j^{-1} 是微分同胚，所以

$$\left|\phi_j^{-1}(y)-\phi_j^{-1}(y')\right| \approx \left|y-y'\right|.$$

据此，我们知道

$$\|u_j\|^2_{H^{\frac{1}{2}}(\partial\Omega)} \approx \iint_{V'\times V'} \frac{\left|\left(u_j\circ\phi_j^{-1}\right)(y)-\left(u_j\circ\phi_j^{-1}\right)(y')\right|^2}{|y-y'|^{n-1+2\times\frac{1}{2}}} dydy' \approx \|u_j\circ\phi_j^{-1}\|_{H^{\frac{1}{2}}(\mathbb{R}^{n-1})}.$$

这就证明了两个范数的等价性。

特别地，我们利用范数

$$\|u\|_{H^{\frac{1}{2}}(\partial\Omega),\mathfrak{U}} = \sum_{j\leqslant N} \|(\chi_j\cdot u)\circ\phi_j^{-1}\|_{H^{\frac{1}{2}}(\mathbb{R}^{n-1})},$$

可以很方便地证明 $H^{\frac{1}{2}}(\partial\Omega)$ 的完备性：假设 $\{u_p\}_{p\geqslant 1}\subset H^{\frac{1}{2}}(\partial\Omega)$ 是 Cauchy 列，那么，按照定义，对任意的 $j\leqslant N$，$\{(\chi_j\cdot u_p)\circ\phi_j^{-1}\}_{p\geqslant 1}\subset H^{\frac{1}{2}}(\mathbb{R}^{n-1})$ 是 Cauchy 列，所以，对每个 j，存在具有紧支集的 $v_j\in H^{\frac{1}{2}}(V')$，使得

$$\lim_{p\to\infty}\left\|(\chi_j\cdot u_p)\circ\phi_j^{-1}-v_j\right\|_{H^{\frac{1}{2}}(\mathbb{R}^{n-1})}=0.$$

我们令

$$u=\sum_{j\leqslant N} v_j\circ\phi_j.$$

利用 $\|\cdot\|_{H^{\frac{1}{2}}(\partial\Omega)}$ 这个范数，很明显每个 $v_j\circ\phi_j$ 都落在 $H^{\frac{1}{2}}(\partial\Omega)$ 中，所以，$u\in H^{\frac{1}{2}}(\partial\Omega)$。仍然利用这个范数，我们就有

$$\begin{aligned}\|u_p-u\|_{H^{\frac{1}{2}}(\partial\Omega)} &\leqslant \sum_{j\leqslant N}\|\chi_j\cdot u_p-v_j\circ\phi_j\|_{H^{\frac{1}{2}}(\partial\Omega)} \\ &\approx \sum_{j\leqslant N}\|(\chi_j\cdot u_p)\circ\phi_j^{-1}-v_j\|_{H^{\frac{1}{2}}(\mathbb{R}^{n-1})}\to 0\end{aligned}$$

从而，我们证明了

命题 507. 函数空间 $\left(H^{\frac{1}{2}}(\Omega), \|\cdot\|_{H^{\frac{1}{2}}(\Omega)}\right)$ 是完备内积空间，其中，我们用如下物理空间上的内积

$$(u_1,u_2)^2_{H^{\frac{1}{2}}(\partial\Omega)} = (u_1,u_2)_{L^2(\partial\Omega)} + \iint_{\partial\Omega\times\partial\Omega} \frac{(u_1(x)-u_2(y))\left(\overline{u_1(x)}-\overline{u_2(y)}\right)}{|x-y|^n} d\sigma(x)d\sigma(y).$$

其中，$u_1, u_2 \in H^{\frac{1}{2}}(\Omega)$。

我们现在来定义限制（迹）映射：

定理 508. 假设 $n \geqslant 1$，$\Omega \subset \mathbb{R}^n$ 是有界带边区域。那么，限制映射

$$\mathrm{Res} : \mathcal{D}(\mathbb{R}^n) \longrightarrow C^\infty(\partial\Omega)$$

可以唯一地延拓成连续线性映射

$$\mathrm{Res} : H^1(\Omega) \longrightarrow H^{\frac{1}{2}}(\partial\Omega)$$

使得如下的图表是交换的：

$$\begin{array}{ccc} \mathcal{D}(\mathbb{R}^n)\cap H^1(\Omega) & \xrightarrow{\mathrm{Res}} & C^\infty(\partial\Omega) \\ \iota\big\downarrow & & \big\downarrow\iota \\ H^1(\Omega) & \xrightarrow{\mathrm{Res}} & H^{\frac{1}{2}}(\partial\Omega) \end{array}$$

$$0 \to H^1_0(\Omega) \xrightarrow{\iota} H^1(\Omega) \xrightarrow{\mathrm{Res}} H^{\frac{1}{2}}(\partial\Omega) \to 0,$$

也就是说 Res 是满射并且对任意的 $u \in H^1(\Omega)$，$\mathrm{Res}(u) = 0$ 当且仅当 $u \in H^1_0(\Omega)$。

证明: 由于我们之前已经证明了 $\mathcal{D}(\mathbb{R}^n)\cap H^1(\Omega)$ 是 $H^1(\Omega)$ 的稠密子空间，所以，我们只要证明存在常数 $C>0$，使得对任意的光滑函数 $u \in \mathcal{D}(\mathbb{R}^n)$，我们有如下的不等式

$$\|u\big|_{\partial\Omega}\|_{H^{\frac{1}{2}}(\partial\Omega)} \leqslant C\|u\|_{H^1(\Omega)}.$$

首先，利用单位分解，我们把 u 写成

$$u = \sum_{j\leqslant N} \chi_j \cdot u = \sum_{j\leqslant N} u_j.$$

如果 $\mathrm{supp}(u_j)\cap\partial\Omega = \emptyset$，那么它对限制映射没有贡献。所以，我们可以假设所有的 $\mathrm{supp}(u_j)$ 都与边界相交。对于一个特定的 u_j，我们知道

$$u_j\big|_{\partial\Omega}\circ\phi_j^{-1} = (u_j\circ\Phi_j)\big|_{y_n=0}.$$

所以，

$$\begin{aligned} \|u_j\big|_{\partial\Omega}\|_{H^{\frac{1}{2}}(\partial\Omega)} &= \|u_j\big|_{\partial\Omega}\circ\phi_j^{-1}\|_{H^{\frac{1}{2}}(\mathbb{R}^{n-1})} = \|(u_j\circ\Phi_j)\big|_{y_n=0}\|_{H^{\frac{1}{2}}(\mathbb{R}^{n-1})} \\ &\leqslant C\|u_j\circ\Phi_j\|_{H^1(\mathbb{H}^n)} \leqslant C'\|u_j\|_{H^1(\Omega)} \leqslant C''\|u\|_{H^1(\Omega)}. \end{aligned}$$

对 j 求和，我们就得到了要证明的不等式。特别地，这就构造了连续的限制映射

$$\mathrm{Res} : H^1(\Omega) \longrightarrow H^{\frac{1}{2}}(\partial\Omega).$$

正合列的单射部分是平凡的，我们现在证明如果 $\mathrm{Res}(u) = 0$，那么 $u \in H_0^1(\Omega)$：实际上，我们有

$$\mathrm{Res}(\chi_j \cdot u) = \chi_j \mathrm{Res}(u) = 0.$$

所以，只要对 u_j 证明即可。此时，考虑 $u_j \circ \Phi_j^{-1}$，它可以被 $\{\varphi_p\}_{p\geqslant 1} \subset C_0^\infty(\mathbb{H}^n)$ 在 $H^1(\mathbb{H}^n)$ 中逼近，所以，$\{\varphi_p \circ \Phi_j\}_{p\geqslant 1}$ 在 $H^1(\Omega)$ 中逼近 u_j。这说明 $u_j \in H_0^1(\Omega)$，所以，

$$u = \sum_{j\leqslant N} u_j \in H_0^1(\Omega).$$

最后一个关于 Res 的满射性也可以同样的证明：对任意的 $f \in H^{\frac{1}{2}}(\Omega)$，我们把它写成

$$f = \sum_{j\leqslant N} \chi_j \cdot f = \sum_{j\leqslant N} f_j.$$

我们只要对任意的 j，找到一个 $u_j \in H^1(\Omega)$，使得 $u_j\big|_{\partial\Omega} = f_j$ 即可。

由于 $f_j \in H^{\frac{1}{2}}(\partial\Omega \cap U_j)$，我们要转化到 V_j^+ 上考虑问题，那么，$f_j \circ \phi_j^{-1} \in H^{\frac{1}{2}}(V_j)$。根据半空间上的限制定理，我们知道存在 $u \in H^1(V_j)$ 并且 $\mathrm{supp}(u) \in H^1(V_j)$，使得 $u\big|_{y_n=0} = f_j \circ \phi_j^{-1}$，所以，$u_j = u \circ \Phi_j$ 即为所求。 □

这个定理的证明可以原封不动地用来证明：

定理 509. *假设 $n \geqslant 1$，$k \geqslant 1$ 是整数，$\Omega \subset \mathbb{R}^n$ 是有界带边区域。那么，限制映射*

$$\mathrm{Res} : \mathcal{D}(\mathbb{R}^n) \longrightarrow C^\infty(\partial\Omega)$$

可以唯一地延拓成连续线性映射

$$\mathrm{Res} : H^k(\Omega) \longrightarrow H^{k-\frac{1}{2}}(\partial\Omega)$$

我们还有如下的正合列：

$$0 \to H_0^k(\Omega) \xrightarrow{\iota} H^k(\Omega) \xrightarrow{\mathrm{Res}} H^{k-\frac{1}{2}}(\partial\Omega) \to 0.$$

作为这些函数空间理论的应用，我们可以精确地叙述 Dirichlet 边值问题（指的是边界值为 0）：对于光滑的带边区域 $\Omega \subset \mathbb{R}^n$，对任意的 $f \in H^{-1}(\Omega)$，我们已经证明了存在 $u \in H_0^1(\Omega))$，使得

$$-\triangle u \overset{\mathcal{D}'}{=} f.$$

根据我们对 Res 的了解，我们知道

$$u\big|_{\partial\Omega} := \mathrm{Res}(u) = 0.$$

所以，我们证明了如下的定理：

定理 510. *对于有界光滑的带边区域 $\Omega \subset \mathbb{R}^n$，对任意的 $f \in H^{-1}(\Omega)$，存在唯一的 $u \in H^1(\Omega))$ 解如下的 Dirichlet 边值问题*

$$\begin{cases} -\triangle u = f, \\ u\big|_{\partial\Omega} = 0. \end{cases}$$

我们还有如下关于边值问题的定理：

定理 511. *对于有界光滑的带边区域 $\Omega \subset \mathbb{R}^n$，对任意的 $g \in H^{\frac{1}{2}}(\partial\Omega)$，存在唯一的 $u \in H^1(\Omega))$，使得*

$$\begin{cases} -\triangle u = 0, \\ u\big|_{\partial\Omega} = g. \end{cases}$$

证明： 由于

$$\mathrm{Res} : H^1(\Omega) \twoheadrightarrow H^{\frac{1}{2}}(\partial\Omega)$$

是满射，我们选取 $v \in H^1(\Omega)$，使得 $v\big|_{\partial\Omega} = g$。所以，上述的问题转化为求解

$$\begin{cases} -\triangle \widetilde{u} = \widetilde{f}, \\ \widetilde{u}\big|_{\partial\Omega} = 0. \end{cases}$$

其中，$\widetilde{u} = u - v$，$\widetilde{f} = \triangle v \in H^{-1}(\Omega)$。此时，我们运用 Dirichlet 边值问题的结论即可。 □

位势方程的椭圆正则性定理

定理 512. *$\Omega \subset \mathbb{R}^n$ 是有界光滑带边区域，$k \geqslant 1$ 是整数。假设 $u \in H^1(\Omega)$ 满足如下的边值问题：*

$$\begin{cases} -\triangle u = f, \\ u\big|_{\partial\Omega} = g. \end{cases}$$

如果 $f \in H^{k-1}(\Omega)$ 并且 $g \in H^{k+\frac{1}{2}}(\partial\Omega)$，那么，$u \in H^{k+1}(\Omega)$。

证明： 我们先对问题做如下的化简：

- 约化为 Dirichlet 边值：根据我们限制的正合列，我们可以选取 $\widetilde{g} \in H^{k+1}(\Omega)$，使得 $\widetilde{g}\big|_{\partial\Omega} = g$。通过考虑 $u - \widetilde{g}$（特别地，$\triangle\widetilde{g} \in H^{k-1}(\Omega)$），我们可以在这个问题中假设 $g \equiv 0$。
- 局部化：我们可以假设 u 的支集很小。实际上，我们可以选取之前用过的 Ω 的开覆盖以及相应的单位分解 $\{\chi_j\}_{j\leqslant N}$，只要说明 $\chi_j \cdot u \in H^{k+1}(\Omega)$ 即可，这是因为，此时我们仍有

  $$\chi_j \cdot u\big|_{\partial\Omega} = 0$$

  并且

  $$-\triangle(\chi_j \cdot u) = -\chi_j \cdot f - 2\nabla\chi_j \cdot \nabla u - \chi_j \cdot \triangle u.$$

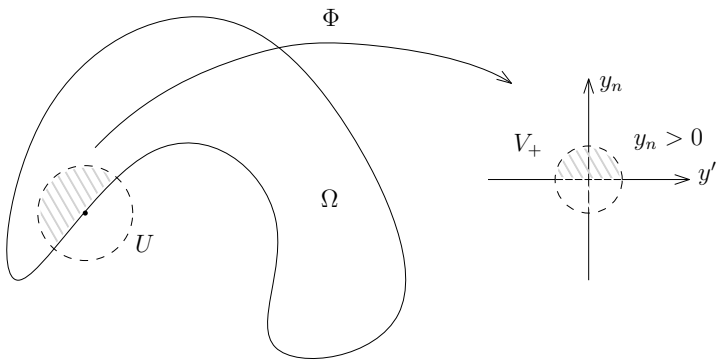
我们之后的证明是对 k 进行归纳。$k=0$ 这个命题是成立的，如果我们假设了对于一切 $\leqslant k-1$ 的整数命题成立，此时，利用归纳假设，$u\in H^k(\Omega)$，所以，$\nabla\chi_j\cdot\nabla u,\chi_j\cdot\triangle u\in H^{k-1}$，我们仍然有

$$-\triangle(\chi_j\cdot u)\in H^{k-1},$$

所以，我们只要对这种情况进行证明即可。

我们现在假设 $\mathrm{supp}(u)\subset U$，其中 $U\subset\mathbb{R}^n$ 是一个开集。利用局部化的结论，我们只要考虑两种情况：$U\cap\partial\Omega=\emptyset$；$U$ 是某个边界点处的开集。第一种情况的证明可以被第二种情况的证明过程所包含（请在下面的证明中注意这一点），所以我们只考虑第二种情况。

此时，我们选取微分同胚 $\Phi:U\cap\Omega\to V_+\subset\mathbb{H}^n$，我们要把问题转化为半空间上的问题（下面的证明是非常值得推敲）。



我们令 $v=u\circ\Phi^{-1}$，那么，$u\in H^{k+1}(\Omega)$ 等价于 $v\in H^{k+1}(V_+)$。然而，我们注意到，此时，方程的形式发生了很大的变化：v 在分布的意义下所满足的方程为

$$\sum_{1\leqslant i,j\leqslant N}\delta^{ij}\partial_i\partial_j v(\Phi(x))=f(\Phi(x)),$$

亦即

$$\sum_{1\leqslant i,j\leqslant n}\delta^{ij}\sum_{1\leqslant k,l\leqslant n}\left(\frac{\partial\Phi^k}{\partial x^i}\big(\Phi(x)\big)\frac{\partial\Phi^l}{\partial x^i}\big(\Phi(x)\big)\frac{\partial^2 v}{\partial y_k\partial y_l}\big(\Phi(x)\big)+\frac{\partial^2\Phi^k}{\partial x^i\partial x^j}\big(\Phi(x)\big)\frac{\partial v}{\partial y_k}\big(\Phi(x)\big)\right)=f(\Phi(x)).$$

这是一个变系数的二阶微分方程，它看起来很不友好。实际上，我们对与问题的理解都是在分布意义下的，所以，我需要在分布意义下（或者变分意义下）理解解的变化，上面的复杂方程实际上对我们没有用处。

我们的解有如下刻画：对任意的 $\varphi\in H_0^1(\Omega)$，我们都有

$$\int_\Omega\nabla u(x)\cdot\overline{\nabla\varphi(x)}dx=\int_\Omega f(x)\overline{\varphi(x)}dx.$$

通过转换为 V_+ 上 y 的坐标系并且令 $\psi(y)=\varphi(\Phi^{-1}(y))$，我们有

$$\int_{V^+}\nabla_k(v\circ\Phi(x))\cdot\nabla_k(\psi\circ\Phi(x))|\mathrm{Jac}_{\Phi^{-1}}(y)|dy=\int_{V^+}f(\Phi^{-1}(y))\varphi(\Phi^{-1}(y))|\mathrm{Jac}_{\Phi^{-1}}(y)|dy.$$

即

$$\sum_{1\leqslant j,k,l\leqslant n}\int_{V^+}\frac{\partial\Phi^k(x)}{\partial x_j}\frac{\partial\Phi^l(x)}{\partial x_j}\frac{\partial v(y)}{\partial y_k}\overline{\frac{\partial\psi(y)}{\partial y_l}}|\mathrm{Jac}_{\Phi^{-1}}(y)|dy=\int_{V^+}f(\Phi^{-1}(y))\overline{\varphi(\Phi^{-1}(y))}|\mathrm{Jac}_{\Phi^{-1}}(y)|dy.$$

我们令

$$\begin{aligned}b^{kl}(y)&=\sum_{1\leqslant j\leqslant n}\frac{\partial\Phi^k(\Phi^{-1}(y))}{\partial x_j}\frac{\partial\Phi^l(\Phi^{-1}(y))}{\partial x_j}|\mathrm{Jac}_{\Phi^{-1}}(y)|\in C^\infty(V_+),\\ F(y)&=f(\Phi^{-1}(y))|\mathrm{Jac}_{\Phi^{-1}}(y)|\in H^{k-1}(V_+),\end{aligned}$$

我们有

$$\boxed{\sum_{1\leqslant k,l\leqslant n}\int_{V^+}b^{kl}(y)\frac{\partial v(y)}{\partial y_k}\overline{\frac{\partial\psi(y)}{\partial y_l}}dy=\int_{V^+}F(y)\overline{\psi(y)}dy,}$$

其中，上面的等式对任意的 $\psi(y)\in H_0^1(V_+)$ 都成立，这是因为 $\psi(y)=\varphi(\Phi^{-1}(y))$ 可以表示任意 $H_0^1(V_+)$ 中的元素。上面的方框是我们对换坐标之后的表述。我们把方框的左边用一个二次型来记：

$$B(v,\psi)=\sum_{1\leqslant k,l\leqslant n}\int_{V^+}b^{kl}(y)\frac{\partial v(y)}{\partial y_k}\overline{\frac{\partial\psi(y)}{\partial y_l}}dy$$

另一个重要的观察是矩阵 $\left(b^{kl}\right)_{1\leqslant k,l\leqslant n}$ 是正定矩阵，实际上，

$$\left(b^{kl}\right)={}^t\left(\mathrm{Jac}(\Phi)\right)\cdot I\cdot\left(\mathrm{Jac}(\Phi)\right).$$

所以，存在常数 $c>0$，使得 $\left(b^{kl}(y)\right)$ 的所有特征值都至少是 c（对任意的 $y\in V_+$ 成立）。这是问题所谓的椭圆性。特别地，对于 $w\in H_0^1(V_+)$，我们就有

$$B(w,w)\geqslant c\|w\|_{H_0^1}^2.$$

我们现在来证明 v 的正则性，我们先对 $k=1$ 来证明（$k=0$ 的情况已经完成），然后对 k 进行归纳。我们要证明 $v\in H^2(V_+)$。这个证明的方法通常被称作是差分方法。

我们考虑平行方向的导数 $\partial_{y_j}v(y',y_n)$，其中，$j\leqslant n-1$（当 $j=n$ 时，我们将利用方程的结构）。不妨假设 $j=1$，那么，根据分布的知识，我们知道

$$\partial_1 v\overset{\mathcal{D}'}{=}\lim_{t\to0}\frac{\tau_t v-v}{t},$$

其中，

$$\tau_t v(y_1,\cdots,y_{n-1},y_n)=v(y_1+t,\cdots,y_{n-1},y_n).$$

我们的目标是控制 $\|\partial_1 v\|_{H^1}$（从而，v 才可能落在 H^2 中）。

我们要把 $\dfrac{\tau_t v - v}{t}$ 代入到方框中的方程里，所以，我们先计算

$$\begin{aligned}
B(\tau_t v, \psi) &= \sum_{1\leqslant k,l\leqslant n} \int_{V^+} b^{kl}(y) \frac{\partial v(y_1+t,\cdots)}{\partial y_k} \overline{\frac{\partial \psi(y)}{\partial y_l}} dy \\
&= \sum_{1\leqslant k,l\leqslant n} \int_{V^+} b^{kl}(y_1-t,\cdots) \frac{\partial v(y)}{\partial y_k} \overline{\frac{\partial \psi(y_1-t,\cdots)}{\partial y_l}} dy \\
&= \sum_{1\leqslant k,l\leqslant n} \int_{V^+} b^{kl}(y) \frac{\partial v(y)}{\partial y_k} \overline{\frac{\partial \tau_{-t}\psi(y)}{\partial y_l}} dy \\
&\quad + \sum_{1\leqslant k,l\leqslant n} \int_{V^+} \left(\tau_{-t} b^{kl}(y) - b^{kl}(y)\right) \frac{\partial v(y)}{\partial y_k} \overline{\frac{\partial \tau_{-t}\psi(y)}{\partial y_l}} dy.
\end{aligned}$$

从而，

$$B\left(\frac{\tau_t v - v}{t}, \psi\right) = B\left(v, \frac{\tau_{-t}\psi - \psi}{t}\right) + \underbrace{\sum_{1\leqslant k,l\leqslant n} \int_{V^+} \left(\tau_{-t} b^{kl}(y) - b^{kl}(y)\right) \frac{\partial v(y)}{\partial y_k} \overline{\frac{\partial \tau_{-t}\psi(y)}{\partial y_l}} dy}_{I_t}.$$

利用 b 的光滑性，$v, \psi \in H^1$ 以及 Cauchy-Schwarz 不等式，我们得到如下的一致估计：

$$|I_t| \leqslant C\|v\|_{H^1}\|\psi\|_{H^1}.$$

所以，根据

$$B\left(v, \frac{\tau_{-t}\psi - \psi}{t}\right) = \int_{V_+} F(y) \overline{\frac{\tau_{-t}\psi(y) - \psi(y)}{t}} dy,$$

我们就有

$$\begin{aligned}
B\left(\frac{\tau_t v - v}{t}, \psi\right) &\leqslant \left| B\left(v, \frac{\tau_{-t}\psi - \psi}{t}\right)\right| + I_t \\
&\leqslant C\left(\|v\|_{H^1}\|\psi\|_{H^1} + \|F\|_{L^2} \left\| \frac{\tau_{-t}\psi(y) - \psi(y)}{t} \right\|_{L^2}\right).
\end{aligned}$$

利用 Fourier 变换，我们有

$$\begin{aligned}
\left\| \frac{\tau_{-t}\psi(y) - \psi(y)}{t} \right\|_{L^2}^2 &\leqslant \int_{\mathbb{R}^n} \left| \frac{e^{it\xi_1} - 1}{t} \right|^2 |\widehat{\psi}(\xi)|^2 d\xi \\
&\leqslant \int_{\mathbb{R}^n} |\xi_1|^2 |\widehat{\psi}(\xi)|^2 d\xi \leqslant \|\psi\|_{H_0^1}^2.
\end{aligned}$$

从而，对任意的 $\psi \in H_0^1(V_+)$，我们都有如下的不等式

$$B\left(\frac{\tau_t v - v}{t}, \psi\right) \leqslant C\left(\|v\|_{H^1} + \|F\|_{L^2}\right)\|\psi\|_{H_0^1}.$$

由于 $v\in H_0^1(V_+)$，我们知道 $\dfrac{\tau_t v-v}{t}\in H_0^1(V_+)$，我们在上面把 ψ 取成 $\dfrac{\tau_t v-v}{t}$，利用椭圆性，我们就有

$$c\left\|\frac{\tau_t v-v}{t}\right\|_{H_0^1}^2\leqslant C\left(\|v\|_{H^1}+\|F\|_{L^2}\right)\left\|\frac{\tau_t v-v}{t}\right\|_{H_0^1}.$$

从而，存在常数 C'，使得对任意的 t，我们都有

$$\left\|\frac{\tau_t v-v}{t}\right\|_{H_0^1}\leqslant C'\left(\|v\|_{H^1}+\|F\|_{L^2}\right).$$

为了让 $t\to 0$，我们考虑配对

$$\left|\left\langle\frac{\tau_t\partial_k v-\partial_k v}{t},\psi\right\rangle\right|\leqslant\|\frac{\tau_t\partial_k v-\partial_k v}{t}\|_{L^2}\|\psi\|_{L^2}\leqslant C\|\psi\|_{L^2},$$

其中，ψ 是试验函数，$k\leqslant n$。从而，当 $t\to 0$ 时（此时在分布的意义下有极限），我们得到

$$|\langle\partial_1\partial_k v,\psi\rangle|\leqslant C\|\psi\|_{L^2}.$$

根据 Riesz 表示定理，$\partial_1\partial_k v\in L^2(V_+)$。

综合上述，我们证明 $\partial_j\partial_k v\in L^2(V_+)$，其中 $j\neq n$。我们现在证明 $\partial_n\partial_n v\in L^2(V_+)$：注意到，$v$ 所满足的方程为

$$\sum_{1\leqslant i,j\leqslant n}\delta^{ij}\sum_{1\leqslant k,l\leqslant n}\left(\frac{\partial\Phi^k}{\partial x^i}\big(\Phi(x)\big)\frac{\partial\Phi^l}{\partial x^i}\big(\Phi(x)\big)\frac{\partial^2 v}{\partial y_k\partial y_l}\big(\Phi(x)\big)+\frac{\partial^2\Phi^k}{\partial x^i\partial x^j}\big(\Phi(x)\big)\frac{\partial v}{\partial y_k}\big(\Phi(x)\big)\right)=f(\Phi(x)).$$

利用矩阵 $\left(b^{kl}\right)$，我们有

$$\sum_{1\leqslant k,l\leqslant n}\left(b^{kl}(y)\frac{\partial^2 v}{\partial y_k\partial y_l}(y)+\underbrace{\sum_{i=1}^n\frac{\partial^2\Phi^k}{\partial x_i^2}(y)\frac{\partial v}{\partial y_k}(y)}_{\in L^2}\right)=\underbrace{f(y)}_{\in L^2}.$$

所以，

$$b^{nn}(y)\frac{\partial^2 v}{\partial y_n^2}(y)+\sum_{k,l\neq(n,n)}b^{kl}(y)\frac{\partial^2 v}{\partial y_k\partial y_l}(y)\in L^2.$$

由于 $\partial_j\partial_k v\in L^2(V_+)$，其中 $j\neq n$ 并且 $\left(b^{kl}\right)$ 的正定性意味着 $b^{nn}(y)$ 是有正的下界的光滑函数，所以 $\partial_n^2 v\in L^2(V_+)$。

这就完成了 $k=1$ 的情况的证明。

对于一般的 k，我们进行归纳：假设对于 $\leqslant k-1$ 的时候，命题都成立，利用方程

$$\sum_{1\leqslant k,l\leqslant n}\left(b^{kl}(y)\frac{\partial^2 v}{\partial y_k\partial y_l}(y)+\sum_{i=1}^n\frac{\partial^2\Phi^k}{\partial x_i^2}(y)\frac{\partial v}{\partial y_k}(y)\right)=f(y),$$

对任意的 $j\leqslant n-1$，我们都有

$$\sum_{1\leqslant k,l\leqslant n}b^{kl}(y)\frac{\partial^2(\partial_j v)}{\partial y_k\partial y_l}(y)=-\sum_{1\leqslant k,l\leqslant n}\left(\partial_j b^{kl}\right)(y)\frac{\partial^2 v}{\partial y_k\partial y_l}(y)-\partial_j\left(\sum_{i=1}^n\frac{\partial^2\Phi^k}{\partial x_i^2}(y)\frac{\partial v}{\partial y_k}(y)\right)+\partial_j f(y).$$

根据归纳假设 $v \in H^k(V_+) \cap H_0^1(V_+)$，这表明上面表达式的右边落在空间 $H^{k-2}(V_+)$ 中；由于 ∂_j 是和边界 ∂V_+ 平行方向的导数，我们有 $\partial_j v \in H_0^1(V_+)$，这是因为对于任意的 $\varphi \in C^\infty(\mathbb{R}^n) \cap H^1(\mathbb{H})$，对于我们有

$$\mathrm{Res} \circ \partial_j = \partial_j \circ \mathrm{Res},$$

（这个等式对于 $j = n$ 不成立）利用 Res 连续性及我们所证明的正合列，如果 $v \in H_0^1(V_+)$，那么，$\mathrm{Res}\,(\partial_j v) = 0$，所以 $\partial_j v \in H_0^1(V_+)$。此时，$\partial_j v \in H_0^1(V_+)$ 满足如下的方程：

$$\sum_{1 \leqslant k,l \leqslant n} b^{kl}(y) \frac{\partial^2(\partial_j v)}{\partial y_k \partial y_l}(y) \in H^{k-2}(V_+),$$

并且这个函数满足归纳假设的要求（特别是 $\partial_j v \in H_0^1(V_+)$），从而，对任意的 $j \leqslant n-1$，我们都有 $\partial_j v \in H^k(V_+)$。最终，为了说明 $v \in H^k(V_+)$，只需要证明（只差了这一个导数的控制）$\partial_n^{k+1} v \in L^2$。我们再次利用方程：

$$b^{nn}(y) \frac{\partial^2 v}{\partial y_n \partial y_n}(y) = - \sum_{(k,l) \neq (n,n)} b^{kl}(y) \frac{\partial^2 v}{\partial y_k \partial y_l}(y) - \sum_{1 \leqslant k,l,i \leqslant n} \frac{\partial^2 \Phi^k}{\partial x_i^2}\big(y\big) \frac{\partial v}{\partial y_k}\big(y\big) + f(y).$$

对这个方程求 $k-1$ 次 ∂_n 方向的导数，除了 $b^{nn}(y)\partial_n^{k+1} v(y)$ 之外，其余的项都落在 L^2 中，这就完成了证明。 □

推论 513. $\Omega \subset \mathbb{R}^n$ 是有界光滑带边区域，$u \in H^1(\Omega)$ 满足如下的边值问题：

$$\begin{cases} -\triangle u = f, \\ u\big|_{\partial\Omega} = g. \end{cases}$$

如果 $f \in C^\infty(\Omega')$，其中 Ω' 是包含 $\overline{\Omega}$ 的开集，$g \in C^\infty(\partial\Omega)$，那么，$u \in C^\infty(\Omega)$。

76.1 习题（利用变分与 Riesz 表示定理解微分方程）：一个弹性力学的模型

一个弹性力学的模型可以用如下的偏微分方程的边值问题来描述：假设 $\Omega \subset \mathbb{R}^n$ 是光滑的有界带边的区域，$f \in L^2(\Omega)$，$c(x) \in C^0(\Omega)$，我们要找分布 $u \in \mathcal{D}'(\Omega)$，使得

$$\begin{cases} \triangle^2 u + c(x)u \overset{\mathcal{D}'(\Omega)}{=} f, \\ \triangle u\big|_{\partial\Omega} = 0, \\ u\big|_{\partial\Omega} = 0. \end{cases}$$

1) 令

$$V = H^2(\Omega) \cap H_0^1(\Omega).$$

证明，$V \subset H^2(\Omega)$ 是闭子空间（在 H^2 范数下）。

2) 假设 $c(x) > 0$，证明，

$$Q(u, v) = \int_\Omega \triangle u \overline{\triangle v} + c(x) u \bar{v} dx$$

定义出了 V 上的一个内积。

3) 利用上述内积（假设 $c(x) > 0$），给出方程的一个解。

4) 证明，如果 3) 中给出的解 $u \in H^4(\Omega)$，那么，在 L^2 的意义下，

$$\triangle^2 u + c(x)u = f.$$

5) 如果只假设 $c(x)$ 是实值函数，如何给出方程的一个解?

6) 当 $c(x)$ 是实数值的时候，试解释如何构造如下方程的解：

$$\begin{cases} \triangle^2 u + c(x)u \overset{\mathcal{D}'(\Omega)}{=} f, \\ \triangle u\big|_{\partial\Omega} = u_2, \\ u\big|_{\partial\Omega} = u_0. \end{cases}$$

其中，$f \in L^2(\Omega)$，$u_2 \in H^{\frac{3}{2}}(\Omega)$，$u_0 \in H^{\frac{7}{2}}(\Omega)$。

77 完备内积空间上的紧算子，自伴算子与弱收敛理论

二零二零年十一月二十五日，星期三，晴

为了研究有界区域上 Laplace 算子的特征值问题，我们还需要引入紧算子的概念。直观上，我们可以将紧算子理解成最接近于有限维线性映射的线性算子。假设 $(X,\|\cdot\|_X)$ 和 $(Y,\|\cdot\|_Y$ 是两个完备的赋范线性空间，给定连续线性映射

$$T:X\to Y.$$

定义 514. 如果对任何的有界集 $A\subset X$（即存在 $M>0$，使得对任意的 $a\in A$，总有 $\|a\|_X\leqslant M$），总能找到 $T(A)\subset Y$ 中的点列 $\{T(x_k)\}_{k\geqslant 1}$，使得 $\{T(x_k)\}_{k\geqslant 1}$ 在 Y 中收敛，那么，我们就称 T 是**紧算子**。

要强调是，当谈论一个算子是紧算子的时候，我们总是事先假定它是有界线性算子。

我们罗列几个关于紧算子的基本性质：

命题 515. *1)*（紧算子是双边理想）假设 $(X,\|\cdot\|_X)$，$(Y,\|\cdot\|_Y)$ 和 $(Z,\|\cdot\|_Z)$ 是完备的赋范线性空间，假设

$$T:X\to Y,\quad S:Y\to Z$$

是连续线性映射，如果 T 或者 S 其中之一为紧算子，那么它们的复合 $S\circ T:X\to Z$ 也是紧算子。

2)（紧算子的集合是闭的）给定 $(X,\|\cdot\|_X)$ 和 $(Y,\|\cdot\|_Y)$ 是完备的赋范线性空间和连续线性映射

$$T:X\to Y.$$

假设我们有一列紧算子

$$T_k:X\to Y,\quad k=1,2,\cdots,$$

并且在算子的意义下 $T_k\to T$，这里的收敛性指的是

$$\lim_{k\to\infty}\|T_k(x)-T(x)\|_Y=0$$

对所有的 X 中的单位球中的点 x（即 $\|x\|_X\leqslant 1$）一致地成立，也就是说，对任意的 $\varepsilon>0$，存在 $N>0$，当 $k\geqslant N$ 时，对任意的 X 中的单位球中的点 x，我们有

$$\|T_k(x)-T(x)\|_Y<\varepsilon,$$

那么，T 也是紧算子。

3)（有限秩算子是紧的）给定 $(X,\|\cdot\|_X)$ 和 $(Y,\|\cdot\|_Y)$ 是完备的赋范线性空间和连续线性映射

$$T:X\to Y.$$

如果 T 是**有限秩**的算子，也就是说 T 的像 $T(X)$ 是 Y 中的有限维的线性子空间，那么，T 是紧算子。

证明: 证明本身对我们后面的应用并没有影响，为了完整起见，我们还是在这里给出证明。

先证明 1)。如果 T 是紧算子，对任意的 $A\subset X$ 是有界集合，那么，存在 $\{x_k\}_{k\geqslant 1}\subset A$，使得 $\{T(x_k)\}_{k\geqslant 1}$ 在 Y 中收敛。由于连续映射把收敛的序列映射为收敛的序列，所以，$\{S\left(T(x_k)\right)\}_{k\geqslant 1}$ 在 Z 中收敛。这表明 $S\circ T$ 是紧算子。

如果 S 是紧算子，对任意的 $A\subset X$ 是有界集合，那么，利用 T 是有界线性算子，$T(A)$ 是 Y 中的有界集合，所以，利用 S 是紧算子，存在序列 $\{T(x_k)\}_{k\geqslant 1}\subset T(A)$，使得 $\{S(T(x_k))\}_{k\geqslant 1}\subset Z$ 是收敛的序列。这表明，我们可以选取 $\{x_k\}_{k\geqslant 1}\subset A$，使得 $\{S\left(T(x_k)\right)\}_{k\geqslant 1}$ 在 Z 中收敛，从而 $S\circ T$ 是紧算子。

现在证明 2)。假设 $A\subset X$ 是有界集合并且对任意的 $a\in A$，$\|a\|_X\leqslant M$。令

$$\frac{1}{M}\cdot A=\left\{\frac{1}{M}a\middle|a\in A\right\},$$

根据 T 的线性，我们知道存在 A 中点列使得其像收敛等价于存在 $\dfrac{1}{M}A$ 中点列使得其像收敛。所以，通过把 A 替换成 $\dfrac{1}{M}A$，我们总是可以假设 A 落在 X 中的单位球里。我们利用对角线法则来选取 A 中的收敛子列：

对于 T_1 而言，利用紧性，存在 $\{x_{1,k}\}_{k\geqslant 1}\subset A$，使得 $T_1(x_{1,k})$ 在 Y 中收敛；

对于 T_2 而言，由于 $\{x_{1,k}\}_{k\geqslant 1}$ 是有界的，利用紧性，存在 $\{x_{2,k}\}_{k\geqslant 1}\subset\{x_{1,k}\}_{k\geqslant 1}$ 是子序列，使得 $T_2(x_{2,k})$ 在 Y 中收敛，我们还要求 $x_{2,1}=x_{1,2}$；

$\cdots\ \cdots$；

对于 T_{m+1} 而言，由于 $\{x_{m,k}\}_{k\geqslant 1}$ 是有界的，利用紧性，存在 $\{x_{m+1,k}\}_{k\geqslant 1}\subset\{x_{m,k}\}_{k\geqslant 1}$ 是子序列，使得 $T_{m+1}(x_{m,k})$ 在 Y 中收敛，我们还要求 $x_{m+1,1}=x_{m,2}$；

$\cdots\ \cdots$；

此时，我们考虑序列 $\{x_{m,1}\}_{m\geqslant 1}$。很明显，对任意的 $k\geqslant 1$，当 $m\geqslant \ell$ 时，我们有 $\{x_{m,1}\}_{m\geqslant k}\subset\{x_{\ell,k}\}_{k\geqslant 1}$，从而，对任意的 ℓ，$\{T_\ell(x_{m,1})\}_{m\geqslant 1}$ 在 Y 中收敛。我们下面说明 $\{T(x_{m,1})\}_{m\geqslant 1}$ 是 Y 中的 Cauchy 列即可（注意到 $\{x_{m,1}\}_{m\geqslant 1}$ 落在 X 的单位球里）：对任意的 ε，存在 ℓ_1 和 ℓ_2，使得对任意的 X 中的单位球中的点 x，我们有

$$\|T_{\ell_1}(x)-T(x)\|_Y<\frac{\varepsilon}{4},\ \ \|T_{\ell_4}(x)-T(x)\|_Y<\frac{\varepsilon}{4}.$$

对于 ℓ_i 而言，其中 $i=1$ 和 2，由于 $\{T_{\ell_i}(x_{m,1})\}_{m\geqslant 1}$ 在 Y 中收敛，所以存在 $N>0$，使得当 $m_1,m_2\geqslant N$ 时，我们有

$$\|T_{\ell_i}(x_{m_1,1})-T_{\ell_i}(x_{m_2,1})\|_Y<\frac{\varepsilon}{4}.$$

所以，对任意的 $\varepsilon>0$，存在 $N>0$，使得当 $m_1,m_2\geqslant N$ 时，我们有

$$\begin{aligned}\|T(x_{m_1,1})-T(x_{m_2,1})\|_Y\leqslant&\|T(x_{m_1,1})-T_{\ell_1}(x_{m_1,1})\|_Y+\|T_{\ell_1}(x_{m_1,1})-T_{\ell_1}(x_{m_2,1})\|_Y\\&+\|T_{\ell_1}(x_{m_2,1})-T_{\ell_2}(x_{m_2,1})\|_Y+\|T_{\ell_2}(x_{m_2,1})-T(x_{m_2,1})\|_Y\\<&\varepsilon.\end{aligned}$$

所以 2) 成立。

现在证明 3)。我们在 $T(X)$ 配备上 Y 所诱导的范数，此时，$(T(X), \|\cdot\|_Y)$ 是有限维的赋范线性空间，由于有限维赋范线性空间上的范数都是等价的，所以，如果我们用常用的 Euclid 范数，那么，由于 $T(A)$ 是有界集合，所以有收敛子列。 □

例子. 我们研究一个非紧算子的例子。假设 $\big(H,(\cdot,\cdot)\big)$ 是可分的完备内积空间，从而，我们可以找到一组 *Hilbert* 基 $\{e_k\}_{k\geqslant 1}$。特别地，由于 e_k 均为单位长的，所以 $A=\{e_k\}_{k\geqslant 1}$ 是有界集。我们现在考虑恒同映射：

$$\mathrm{Id}: H\to H,\quad x\mapsto x,$$

由于当 $k\neq l$ 时，$\|e_k-e_l\|=\sqrt{2}$，我们知道 Id 是紧算子当且仅当 H 是有限维的空间：如果 $\{e_k\}_{k\geqslant 1}$ 是无限序列，那么，$\|e_{k+1}-e_k\|=\sqrt{2}$ 表明这不是 *Cauchy* 列。

上面这个例子中所选取的集合 $\{e_k\}_{k\geqslant 1}$ 通常不是紧的，为了给出这个集合的一个更好的刻画，我们引入一个与分布类似的概念：

定义 516. 假设 $\big(H,(\cdot,\cdot)\big)$ 是可分的完备内积空间，给定 $\{x_k\}_{k\geqslant 1}\subset H$ 是一个点列。如果对任意的 $y\in H$，我们都有

$$(x_k,y)\to(x,y),$$

那么，我们就称 $\{x_k\}_{k\geqslant 1}$ **弱收敛**到 x。我们把弱收敛记做是

$$x_k\rightharpoonup x,\quad \text{或}\quad x_k\xrightarrow{w}x.$$

例子. 考虑可分完备内积空间 H 的一组 *Hiblert* 基 $\{e_k\}_{k\geqslant 1}$，我们来说明 $e_k\rightharpoonup 0$：

实际上，对任意的 $x\in H$，我们可以把 x 唯一地写成

$$x=\sum_{k=1}^{\infty}a_k\cdot e_k,\quad a_k=(x,e_k),k=1,2,\cdots.$$

根据 *Parseval* 等式，我们有

$$\|x\|^2=\sum_{k=1}^{\infty}|a_k|^2<\infty.$$

所以，$\lim\limits_{k\to\infty}|a_k|=0$。所以，对任意的 $x\in H$，我们有

$$\lim_{k\to\infty}(x,e_k)=0.$$

注记. 弱极限有如下三个明显的性质：

1) (弱极限的唯一性) $\big(H,(\cdot,\cdot)\big)$ 是可分的完备内积空间，给定 $\{x_k\}_{k\geqslant 1}\subset H$ 是一个点列。假设当 $k\to\infty$ 时，我们有

$$x_k\rightharpoonup x,\quad x_k\rightharpoonup x',$$

其中 $x, x' \in H$，那么 $x = x'$。

这是因为按照定义，我们有

$$(x_k, x - x') \to (x, x - x'), \quad (x_k, x - x') \to (x', x - x').$$

取差，我们就得到

$$(x - x', x - x') = 0.$$

所以，$x = x'$。

2) (收敛意味着弱收敛) $\big(H, (\cdot,\cdot)\big)$ 是可分的完备内积空间，给定 $\{x_k\}_{k\geqslant 1} \subset H$，那么，

$$x_k \to x \quad \Rightarrow \quad x_k \rightharpoonup x.$$

实际上，对任意的 y，我们有

$$|(x_k, y) - (x, y)| = |(x_k - x, y)| \leqslant \|x_n - k\|\|y\| \to 0.$$

这表明 $\lim\limits_{k\to\infty}(x_k, y) = (x, y)$。

命题 517 (弱收敛在连续映射下被保持)**.** 给定可分的完备内积空间 $\big(H, (\cdot,\cdot)\big)$ 和 $\big(H', (\cdot,\cdot)'\big)$，$A : H \to H'$ 是连续线性映射。假设 $\{x_k\}_{k\geqslant 1} \subset H$ 弱收敛到 x，即 $x_k \rightharpoonup x$，那么，$\{A(x_k)\}_{k\geqslant 1} \subset H'$ 弱收敛到 $A(x)$，即 $A(x_k) \rightharpoonup A(x)$。

为了证明这个命题，我们需要引入对偶算子的概念：对任意给定的 $x' \in H'$，对任意的 $x \in H$，根据 Cauchy-Schwarz 不等式，我们有

$$|(Ax, x')| \leqslant \|Ax\|_{H'}\|x'\|_{H'} \leqslant C\|x\|_H\|x'\|_{H'}.$$

其中，最后一步我们用到了 A 是有界的。令 $C' = C\|x'\|_{H'}$，那么，

$$|(Ax, x')| \leqslant C'\|x\|_H.$$

所以，线性映射

$$H \to \mathbb{C}, \quad x \mapsto (Ax, x'),$$

是 H 上的连续线性泛函，根据 Riesz 表示定理，我们存在 H 中的元素 $A^*(x')$（它由 A 和 x' 所决定），使得

$$(Ax, x') = \big(x, A^*(x')\big).$$

这样，我们就构造了映射

$$A^* : H' \to H.$$

利用等式 $(Ax, x') = (x, A^*x')$，很容易看出 A^* 为线性映射。为了证明 A^* 是有界的，根据

$$\big|\big(x, A^*(x')\big)\big| = |(Ax, x')| \leqslant \|Ax\|_{H'}\|x'\|_{H'} \leqslant C\|x\|_H\|x'\|_{H'},$$

其中 x 和 x' 是任意选取的，我们可以选取 $x = A^*(x')$，从而，

$$\big|\big(A^*(x'), A^*(x')\big)\big| \leqslant C\|A^*(x')\|_H \|x'\|_{H'},$$

所以，

$$\|A^*(x')\|_H \leqslant C\|x'\|_{H'}.$$

我们注意到常数 C 这里的选取是和 A 所对应的界是一致的。

现在回到命题的证明：

证明： 为了证明 $Ax_k \rightharpoonup Ax$，我们证明对任意的 $x' \in H'$，$(Ax_k, x') \to (Ax, x')$，这等价于去证明 $(x_k, A^*x') \to (x, A^*x')$。根据 $\{x_k\}_{k\geqslant 1}$ 的弱收敛性，这是显然的。 □

我们之所以引入弱收敛的概念，是因为在可分的完备内积空间中，有界序列在弱拓扑的意义下是仍然是列紧的（和有限维的情况类似）：

定理 518. *假定 $\{x_k\}_{k\geqslant}$ 是可分的完备内积空间 $\big(H, (\cdot,\cdot)\big)$ 中的一个有界序列。那么，存在一个子序列 $\{x_{k_i}\}_{i\geqslant 1}$ 和 $x_\infty \in H$，使得*

$$x_{k_i} \rightharpoonup x_\infty.$$

证明： 选定 H 的一组 Hilbert 基 $\{e_j\}_{j\geqslant 1}$，对任意的 k，我们可以将 x_k 写成

$$x_k = x_k^1 \cdot e_1 + x_k^2 \cdot e_2 + x_k^3 \cdot e_3 + \cdots,$$

其中对任意的 k 和 j，$x_k^j \in \mathbb{C}$。根据 $\{x_k\}_{k\geqslant 1}$ 有界性和对角线法则，我们可以选取 $k_1, k_2, \cdots, k_j, \cdots$，对任意 j，极限 $\lim\limits_{i\to\infty} x_{k_i}^j$ 都存在，我们记

$$x_\infty^j = \lim_{i\to\infty} x_{k_i}^j.$$

换句话说，对任意的 $j \geqslant 1$，极限

$$\lim_{i\to\infty} (x_{k_i}, e_j)$$

存在。所以这个子序列的每个点的每个分量（在给定的 Hilbert 基下）都是收敛的。

我们现在说明，$x_\infty = \sum\limits_{j=1}^{\infty} x_\infty^j e_j$ 是在 H 中良好定义的元素。为此，只要说明 $\sum\limits_{j=1}^{\infty} |x_\infty^j|^2$ 有界即可：假设对任意的 k，$\|x_k\| \leqslant A$，那么，由于 $\lim\limits_{i\to\infty} |x_{k_i}^j|^2 = |x_\infty^j|^2$，根据 Fatou 引理，我们有

$$\sum_{j=1}^{\infty} |x_\infty^j|^2 = \sum_{j=1}^{\infty} \lim_{i\to\infty} |x_{k_i}^j|^2 \leqslant \lim_{i\to\infty} \sum_{j=1}^{\infty} |x_{k_i}^h|^2 \leqslant A^2.$$

特别地，我们还证明了 $\|x_\infty\| \leqslant A$。

最终，我们来证明弱收敛：当 $i \to \infty$ 时，$x_{k_i} \rightharpoonup x_\infty$。对任何 $y \in H$，我们把 y 按照分量写开：

$$y = \sum_{1\leqslant j\leqslant N} a_j e_j + \underbrace{\sum_{j\geqslant N+1} a_j e_j}_{=y'}.$$

其中，我们通过选取较大 N，使得

$$\|y'\| < \frac{1}{4A}\varepsilon.$$

由于 N 固定，所以通过选取足够大的 i，我们可以使得对每个 $j \leqslant N$，我们都有

$$|(x_{k_i} - x_\infty, e_j)| \leqslant \frac{\varepsilon}{2N} = \varepsilon.$$

这表明

$$\begin{aligned}|(x_{k_i} - x_\infty, y)| &\leqslant |(x_{k_i} - x_\infty, y')| + \sum_{j\leqslant N} |(x_{k_i} - x_\infty, e_j)| \\ &\leqslant 2A \times \frac{\varepsilon}{4A} + N \times \frac{\varepsilon}{2N}\end{aligned}$$

从而，当 $i \to \infty$ 时，$(x_{k_i} - x_\infty, y) \to 0$，命题得证。 □

上述运用 Fatou 引理的一段证明中表明弱极限下范数会变小：

推论 519. *假设 $x_k \rightharpoonup x$，那么*

$$\|x\| \leqslant \liminf_{k\to\infty} \|x_k\|.$$

我们知道，如果一个点列 $\{x_j\}_{k\geqslant 1}$ 是收敛到 x 的，那么，相应的范数也收敛，即 $\|x\| = \liminf_{k\to\infty} \|x_k\|$。我们之前见过，给定一组 Hilbert 基 $\{e_k\}_{k\geqslant 1}$，它们不收敛但是弱收敛到 0，而

$$\|0\| < \liminf_{k\to\infty} \|e_k\| = 1.$$

下面的命题表明，范数是否连续是从弱极限升级成为极限的唯一障碍：

命题 520. *假定 $\{x_k\}_{k\geqslant 1}$ 是可分的完备内积空间 $\big(H, (\cdot,\cdot)\big)$ 中的一个有界序列并且 $x_k \rightharpoonup x_\infty$。如果*

$$\lim_{k\to\infty} \|x_k\| = \|x_\infty\|,$$

那么，

$$\lim_{k\to\infty} x_k \overset{H}{=} x_\infty.$$

证明： 我们只要证明 $\|x_k - x_\infty\|^2 \to 0$ 即可。所以，我们计算

$$\|x_k - x_\infty\|^2 = \underbrace{\|x_k\|^2 + \|x_\infty\|^2}_{\to 2\|x_\infty\|^2} - \underbrace{2\Re\big((x_k, x_\infty)\big)}_{\to 2\|x_\infty\|^2}.$$

后面一项的极限是 $2\|x_\infty\|^2$，我们用到了 $x_k \rightharpoonup x_\infty$。所以，上面式子的极限为 0，命题得证。 □

我们回到紧算子的理论。紧算子的一个重要的作用是它也可以将弱收敛的序列升级为收敛的序列：

定理 521. *给定可分的完备内积空间 $\big(H, (\cdot,\cdot)\big)_H$ 和 $\big(H', (\cdot,\cdot)'\big)_{H'}$，$A : H \to H'$ 是连续线性映射。如下两个叙述是等价的：*

1) A 是紧算子；

2) 对 H 中任意（有界的）弱收敛序列 $x_k \rightharpoonup x_\infty$，那么在 H' 中，我们有

$$A(x_k) \xrightarrow{H'} A(x_\infty).$$

注记. 上述 *2)* 有界性假设可以去掉，这需要用到泛函分析中的共鸣定理。在其它很多场合下也不需要空间是可分的，但是我们满足于这样的叙述，因为它们后面的应用是足够的。

证明：2) ⇒ 1) 是显然的：对任意的有界集 $A \subset H$，我们总可以选出一个弱收敛的序列 $\{x_k\}_{k\geqslant 1} \subset A$，使得 $x_k \rightharpoonup x_\infty$，所以，2) 表明 $A(x_k) \to A(x_\infty)$，这说明 A 是紧算子。

现在证明 1) ⇒ 2)。由于 A 是连续的，所以在 H' 中，我们一定有

$$A(x_k) \rightharpoonup A(x_\infty).$$

因为 A 是紧算子并且 $\{x_k\}_{k\geqslant 1}$ 是有界集，所以，$\{A(x_k)\}_{k\geqslant 1}$ 中的任意子序列都包含收敛的子序列。根据弱极限的唯一性，这个收敛子序列（在 $\|\cdot\|_{H'}$ 下）必须收敛到 $A(x_\infty)$。所以，$\{A(x_k)\}_{k\geqslant 1}$ 中的任意子序列所包含的收敛子序列的极限都是 $A(x_\infty)$。我们在一年级第一学期学习极限的时候就证明了这个序列必为 Cauchy 列，从而整个序列收敛。 □

我们现在研究所谓的自伴算子，它们是线性代数中的 Hermite 矩阵或者是实对称矩阵的推广。

定义 522. 给定可分的完备内积空间 $\big(H,(\cdot,\cdot)\big)$，$A: H \to H$ 一个连续的线性自同态，如果 $A = A^*$，我们就称 A 是**自伴的**。换而言之，对任意的 $x, y \in H$，我们均有

$$(Ax, y) = (x, Ay).$$

给定一个 H 到自身的连续线性映射，我们用 $\sigma(A)$ 表示它的特征值的集合，即

$$\sigma(A) = \big\{\lambda \in \mathbb{C}\big|\ \text{存在 } x \neq 0, \text{ 使得 } A(x) = \lambda x\big\}.$$

对于 $\lambda \in \sigma(A)$，我们定义其特征子空间为

$$H_\lambda = \big\{x \in H\big|A(x) = \lambda x\big\}.$$

78 紧算子的谱理论：Hilbert-Schmidt 定理，Laplace 算子的谱分解定理，正方形区域上 Laplace 算子的特征函数与特征值的计算，正方形区域上上 Laplace 算子特征值的增长率

二零二零年十一月二十九日，星期一，晴

我们以下收集一些关于特征值和特征子空间的性质，其中，我们假设 A 是从 H 到自身的连续线性映射：

1) 对任意的 $\lambda \in \sigma(A)$，$H_\lambda \subset H$ 是闭子空间。

这个是明显的：如果 $A(x_k) = \lambda x_k$ 并且 $x_k \to x$，由于方程的两边对于变量都是连续的，所以令 $k \to \infty$，我们就得到 $A(x) = \lambda x$。

2) A 是自伴算子，那么，$\sigma(A) \subset \mathbb{R}$（只有实特征值）。

这个证明和有限维的情况是一致的：假设 $\lambda \in \sigma(A)$，$x \in H_\lambda$ 并且 $x \neq 0$。那么，

$$(\lambda x, x) = (Ax, x) = (x, Ax) = (x, \lambda x).$$

所以，

$$\lambda \|x\|^2 = \overline{\lambda} \|x\|^2.$$

从而，$\lambda \in \mathbb{R}$。

3) 假设 A 是自伴算子，那么 A 的不同特征值的特征向量是相互垂直的，即若 $\lambda, \lambda' \in \sigma(A)$ 并且 $\lambda \neq \lambda'$，那么，$H_\lambda \perp H_{\lambda'}$。

这个证明（也）和有限维的情况是一致的：任意选取 $x \in H_\lambda$，$x' \in H_{\lambda'}$，那么，

$$(\lambda x, x') = (Ax, x') = (x, Ax') = (x, \lambda' x').$$

所以，

$$\lambda (x, x') = \lambda' (x, x').$$

这里我们用到了这些特征值是实数。由于 $\lambda \neq \lambda'$，所以，$(x, x') = 0$，即 $x \perp x'$。

3) 假设 A 是紧算子，如果 $\lambda \in \sigma(A) - \{0\}$ 是非零的特征值，那么 H_λ 是有限维的线性空间。

根据特征值的定义，算子 A 在 H_λ 上的限制映射就是乘以一个非零的常数 $c \in \mathbb{C}$。由于 H_λ 是闭子空间，所以，利用诱导的内积，它也是一个可分的完备内积空间，我们就可以在 H_λ 上选取一个 Hilbert 基 $\{f_k\}_{k \geqslant 1}$。如果 $\dim_{\mathbb{C}}(H_\lambda) = \infty$，那么，$f_k \rightharpoonup 0$，从而，根据 A 是紧算子，$\{A(f_k) = c f_k\}_{k \geqslant 1}$ 收敛到 0，矛盾。

4) 给定连续线性映射 A，我们在 $H-\{0\}$ 上定义如下的一个非线性的泛函：

$$R: H-\{0\} \to \mathbb{C}, \quad u \mapsto R(u) = \frac{\big(A(u), u\big)}{\|u\|^2}.$$

实际上，通过对 u 乘以一个常数，我们知道 R 本质上是定义在 H 中的单位球面上的映射。如果 A 是自伴的算子，此时，我们还知道 R 在 $\mathbb{R}$ 中取值。

根据 A 的连续性，我们知道存在常数 $C>0$，使得 $|R(u)| \leqslant C$。我们现在进一步假设 A 是自伴的算子，我们定义

$$\lambda_1 := \sup_{u \in H-\{0\}} |R(u)|.$$

引理 523. *如果 A 是紧自伴算子，那么，上述极值 λ_1 可以被实现，即存在 $u \in H$，$\|u\|=1$，使得 $R(u)=\lambda_1$。*

证明: 按上确界的定义，存在 $\{u_k\}_{k\geqslant 1} \subset H$，使得 $\|u_k\|=1$ 并且

$$\lim_{k\to\infty} |(A(u_k), u_k)| \to \lambda_1.$$

我们还可以进一步假设存在 $u \in H$，使得 $u_k \rightharpoonup u$。我们现在比较 $R(u)$ 与 $R(u_k)$ 之间的差距。实际上，

$$(A(u_k), u_k) - (A(u), u) = \big(A(u_k - u), u_k\big) - (A(u), u - u_k).$$

由于 A 为紧算子，所以，$A(u_k-u) \to 0$，从而当 $k\to\infty$ 时，上式第一项的极限为零；第二项极限为零，这是因为 $u_k \rightharpoonup u$。所以，$(A(u_k), u_k) \to (A(u), u)$，从而，$(Au, u) = \lambda_1$（此时，我们证明了更强的结论：$(A(u_k), u_k) \to \lambda_1$）。特别地，因为 $\lambda_1 \neq 0$（否则 $A=0$ 就没什么可说的了），所以 $u \neq 0$。

另外，我们有

$$\|u\| \leqslant \lim_{k\to\infty} \|u_k\| = 1,$$

所以，

$$R(u) \geqslant \lambda_1.$$

根据 λ_1 的定义，我们必然有

$$R(u) = \lambda_1.$$

□

注记. *我们证明了更强的结论：$R(u)=\lambda_1$。*

5) 假设 A 是紧自伴算子，$u_1 \in H-\{0\}$ 使得 $R(u_1)=\lambda_1$，那么，$Au_1 = \lambda_1 u_1$，即 λ_1 是 A 的最大特征值。

我们用变分的观点来研究这个问题（这个方法和之前研究 Dirichlet 问题时的想法很类似）：我们不妨假设 $\|u\| = 1$（否则除以一个系数）。对任意的 $v \in H$。我们考虑复数 ε，其中 $\varepsilon \to 0$。此时，

$$\begin{aligned} R(u+\varepsilon v) &= \frac{\big(A(u+\varepsilon v), u+\varepsilon v\big)}{\big(u+\varepsilon v, u+\varepsilon v\big)} \\ &\overset{A=A^*}{=} \frac{(Au,u)+2(Au,v)\varepsilon+O(\varepsilon^2)}{(u,u)+2(u,v)\varepsilon+O(\varepsilon^2)} \\ &= \frac{\lambda_1+2(Au,v)\varepsilon+O(\varepsilon^2)}{1+2(u,v)\varepsilon+O(\varepsilon^2)} \\ &= \big(\lambda_1+2(Au,v)\varepsilon+O(\varepsilon^2)\big)\big(1-2(u,v)\varepsilon+O(\varepsilon^2)\big) \\ &= \lambda_1+2(Au-\lambda_1 u,v)\varepsilon+O(\varepsilon^2). \end{aligned}$$

由于 $R(u_1)$ 是最大值，如果 $a = (Au - \lambda_1 u, v) \neq 0$，我们总可以选取很小的 $\varepsilon = \delta \cdot \overline{a}$，其中 $\delta > 0$ 足够小，使得 $|R(u+\varepsilon v) > \lambda_1|$，这与 λ_1 的定义矛盾。所以，对任意的 v，我们都有

$$(A(u) - \lambda_1 u, v) = 0.$$

令 $v = A(u) - \lambda_1 u$，我们就证明了 $A(u) - \lambda_1 u = 0$。

6) 假设 A 是紧自伴算子，那么 0 是 $\sigma(A)$ 唯一可能的聚点，也就是说如果存在两两不同的 $\{\lambda_k\}_{k\geqslant 1} \subset \sigma(A)$，使得当 $k \to \infty$ 时，$\lambda_k \to \lambda$，那么，$\lambda = 0$。

我们用反证法：假设 $\lambda \neq 0$。对每个 k，我们选取 $u_k \in H$，使得 $Au_k = \lambda_k u_k$ 并且 $\|u_k\| = 1$。我们可以进一步要求 $u_k \rightharpoonup u$。由于 A 是紧算子，所以 $\{A(u_k)\}_{k\geqslant 1}$ 是收敛的，从而，我们有如下的极限：

$$\lim_{k\to\infty} \lambda_k u_k = \lambda u.$$

由于我们假设了 $\lambda \neq 0$，所以，

$$\lim_{k\to\infty} u_k = \lim_{k\to\infty} \frac{\lambda_k}{\lambda} u_k = u.$$

特别地，我们得到了 $Au = \lambda u$，从而，λ 为非零特征值。

利用 λ 是聚点这个性质，对任意的 $\varepsilon > 0$，H 的子空间

$$X = \widehat{\bigoplus}_{|\lambda_k - \lambda| < \varepsilon} H_{\lambda_k}$$

是无限维的闭子空间。其中上述直和 $\widehat{\oplus}$ 是在 H 中的闭包的意义下下取的，即

$$\widehat{\bigoplus}_{|\lambda_k-\lambda|<\varepsilon} H_{\lambda_k} = \overline{\bigoplus_{|\lambda_k-\lambda|<\varepsilon} H_{\lambda_k}}.$$

并且各个分量两两正交。我们证明，A 是紧算子意味着当 ε 足够小时，我们就必然有 $\dim_{\mathbb{C}} X < \infty$：实际上，由于 A 把 H_{λ_k} 映射到自身，所以，A 把 X 也映射到自身。我们在每个 H_{λ_k} 中（一种有无穷多个）选取一个 u_k，使得 $\|u_k\| = 1$，当 $k \neq \ell$ 时，我们知道

$$\|\lambda_k u_k - \lambda_\ell u_\ell\| = \sqrt{\lambda_k^2 + \lambda_\ell^2} \approx \sqrt{2}|\lambda|.$$

这是因为不同特征值所对应的特征向量相互垂直。此时，$\{Au_k = \lambda_k u_k\}_{k\geqslant 1}$ 中不可能有收敛的子列，与 A 的紧性矛盾。

定理 524 (Hilbert-Schmidt 谱定理)**.** *给定可分的完备内积空间 $\big(H, (\cdot,\cdot)\big)$，$A : H \to H$ 自伴的紧算子（有界）。那么，H 有如下的（拓扑）直和分解：*

$$H = \widehat{\bigoplus}_{\lambda\in\sigma(A)} H_\lambda.$$

特别地，如果 $\lambda \neq 0$，那么，$\dim_{\mathbb{C}} H_\lambda < \infty$。进一步，如果 $|\sigma(A)| = \infty$（有无限个特征值），那么，我们可以将所有的特征值 $\lambda_k \in \sigma(A)$ 排序（只有可数个）使得

$$\lim_{k\to\infty} \lambda_k \to 0.$$

证明： 除了关于直和的叙述，我们在此之前已经证明了其他的论断。

首先，我们可以假设 $\ker(A) = 0$，即 $H_0 = \{0\}$：由于 A 是自伴算子，所以，A 把 $\ker(A)$ 的正交补空间映射到自身：

$$A : \ker(A)^\perp \to \ker(A)^\perp.$$

所以，我们只要用 $\ker(A)^\perp$ 来代替 H 考虑问题即可。

我们下面采用将实对称矩阵对角化的方法进行论证。

我们令 $H_1 = H$，$A_1 = A$，其中 $A_1 : H_1 \to H_1$，我们注意到 A_1 也是紧的自伴算子。定理之前 5) 中的论证，我们可以找到 $u_1 \in H_1$，使得 $\|u_1\| = 1$，$R(u_1) = \lambda_1$ 并且 $A_1(u_1) = \lambda_1 u_1$。据此，我们将 H_1 分解为

$$H_1 = \mathbb{C}u_1 \oplus H_2,$$

其中 $H_2 = \big(\mathbb{C}u_1\big)^\perp$ 是 u_1 在 H_1 中的正交补。利用 A_1 的自伴行，我们知道 A_1 在 H_2 上的限制，我们把它记作 A_2，满足 $A_2 : H_2 \to H_2$ 并且 A_2 仍然是自伴的紧算子。

我们重复上面的构造，令 λ_2 为 R 在 H_2 上的最大值，那么，我们可以找到 $u_2 \in H_2$，使得 $\|u_2\| = 1$，$R(u_2) = \lambda_2$ 并且 $A_2(u_2) = \lambda_2 u_2$。据此，我们将 H_2 分解为

$$H_2 = \mathbb{C}u_2 \oplus H_3.$$

如此往复，我们得到点的序列 $\{u_k\}_{k\geqslant 1} \subset H$ 和子空间的下降的序列

$$H_1 \supset H_2 \supset H_3 \supset \cdots.$$

特别地，根据构造，我们还有

$$|\lambda_1| \geqslant |\lambda_2| \geqslant \cdots.$$

我们现在令

$$H' = \overline{\bigoplus_{\lambda_k\in\sigma(A)} \mathbb{C}u_k}.$$

我们只要 $H'=H$ 就完成了证明。首先，我们假设上述操作需要做无限次（否则这就是有限维关于 Hermite 矩阵对角化的过程，我们在线性代数中已经证明，实际上这里重新证明了这个结论），此时，我们已经将 $\{\lambda_k\}_{k\geqslant 1}$ 重排，使得 $\{\lambda_k|\}_{k\geqslant 1}$ 是单调下降到 0 的序列。现在考虑 A 在 $H'^{\perp}$ 上的限制：

$$A: H'^{\perp}\to H'^{\perp}.$$

这仍然是自伴紧算子并且具有特征值 $\lambda\neq 0$。所以，存在某个 k_0，使得 $|\lambda|\in(|\lambda_{k_0+1}|,|\lambda_{k_0}|]$。但是，按照构造方式，我们的变分方法应该先构造出 λ 之后才有可能构造出 λ_{k_0+1}，矛盾。 □

Laplace 算子的谱分解：有界区域上的 Fourier 级数的类比

我们可以将上述紧算子的理论应用到 Dirichlet 问题之上（我们在这个章节并不需要区域是光滑的，$\Omega\subset\mathbb{R}^n$ 只要是有界的开集即可，而我们只在 $H_0^1(\Omega)$ 中研究问题而不再关心函数在边界 $\partial\Omega$ 上的限制）。为此，我们给出一个紧算子例子。从某种意义上说，这是最重要的一类紧算子的例子：

定理 525. *任意给定有界开区域 $\Omega\subset\mathbb{R}^n$，那么，自然的嵌入映射*

$$\iota: H_0^1(\Omega)\to L^2(\Omega),\quad u\mapsto u,$$

是紧算子。

证明： 我们将利用 Fourier 级数的理论来证明这个重要的定理。首先，我们不妨假设 $\overline{\Omega}\subset\subset(0,2\pi)^n$（否则，我们可以把 $\overline{\Omega}$ 放到更大的一个 $\mathbb{R}^n$ 中的一个盒子中去，从而用一个周期与 2π 不同的 Fourier 级数即可）。

我们可以把 $C_0^\infty(\Omega)$ 中的函数在 $\mathbb{R}^n-\Omega$ 中来 0 来延拓，这给出连续映射

$$\mathrm{Ext}: C_0^\infty(\Omega)\to H^1(\mathbb{R}^n).$$

其中，上述连续性之所以成立是因为我们可以在 Ω 上运用 Poincaré 不等式。根据 $C_0^\infty(\Omega)$ 在 $H_0^1(\Omega)$ 中的稠密性，我们就得到了等距嵌入

$$\iota: \left(H_0^1(\Omega),\|\cdot\|_{H^1}\right)\hookrightarrow H^1(\mathbb{R}^n).$$

由于 $\overline{\Omega}\subset(0,2\pi)^n$ 是相对紧的，我们选取 $\chi\in C_0^\infty\left((0,2\pi)^n\right)$，使得 $\chi\big|_{\Omega}\equiv 1$。我们现在证明映射

$$T_\chi: H^1(\mathbb{R}^n)\to L^2(\mathbb{R}^n),\quad f(x)\mapsto\chi(x)f(x),$$

是紧算子。我们注意到，作为 $L^2\left((0,2\pi)^n\right)$ 中的函数，利用 Fourier 级数的展开，我们有

$$\begin{aligned}\chi(x)f(x)&=\chi(x)\sum_{k\in\mathbb{Z}^n}c_k e^{ik\cdot x}\\&=\chi(x)\sum_{k\in\mathbb{Z}^n}\left(\frac{1}{(2\pi)^n}\int_{(0,2\pi)^n}f(y)e^{-ik\cdot y}dy\right)e^{ik\cdot x}\\&=\underbrace{\sum_{|k|\leqslant N}c_k\chi(x)e^{ik\cdot x}}_{=T_{\chi,N}(f)}+\underbrace{\chi(x)\sum_{|k|\geqslant N}c_k e^{ik\cdot x}}_{=R_{\chi,N}(f)}.\end{aligned}$$

我们把 T_χ 分解成两个算子的和，其中，

$$T_{\chi,N}: H^1(\mathbb{R}^n) \to L^2(\mathbb{R}^n), \quad f(x) \mapsto \sum_{|k|\leqslant N} c_k \chi(x) e^{ik\cdot x},$$

而

$$R_{\chi,N}: H^1(\mathbb{R}^n) \to L^2(\mathbb{R}^n), \quad f(x) \mapsto \chi(x) \sum_{|k|\geqslant N} c_k e^{ik\cdot x}.$$

这里，整数 N 是待定的。由于 $T_{\chi,N}$ 的像完全落在 $\{e^{ik\cdot x}\}_{|k|\leqslant N}$ 这有限多个函数所生成的有限维线性空间中，所以它是有限秩的算子，从而，$T_{\chi,N}$ 是紧算子。

为了研究 $R_{\chi,N}$，我们将利用 ∇f 也是 L^2 的函数这个事实。实际上，我们有

$$\begin{aligned}\|R_{\chi,N}(f)\|^2_{L^2(\mathbb{R})^n} &\leqslant \left\| \sum_{|k|\geqslant N} c_k e^{ik\cdot x} \right\|^2_{L^2((0,2\pi)^{2n})} \leqslant \frac{1}{N^2} \sum_{|k|\geqslant N} |k|^2 |c_k|^2 \\ &\leqslant \frac{1}{N^2} \|\nabla f\|_{L^2}.\end{aligned}$$

其中，我们把导数转化为频率空间上衰减。这表明，T_χ 这个算子可以一个紧算子 $T_{\chi,N}$ 来逼近，其中，$N \to \infty$，所以 T_χ 为紧算子。

最后，我们把嵌入

$$\iota: H^1_0(\Omega) \to L^2(\Omega)$$

写成如下两个连续线性映射的复合：

$$H^1_0(\Omega) \to H^1(\mathbb{R}^n) \xrightarrow{T_\chi} L^2(\Omega) \subset L^2(\mathbb{R}^n).$$

由于任何算子与紧算子复合之后是紧算子，所以我们就证明了 ι 也是紧算子。 □

我们现在考虑如下的交换图表：

$$\begin{array}{ccc} H^{-1}(\Omega) & \xrightarrow{(-\triangle)^{-1}} & H^1_0(\Omega) \\ \iota\uparrow & & \downarrow\iota \\ L^2(\Omega) & \xrightarrow{(-\triangle)^{-1}} & L^2(\Omega) \end{array}$$

据此，通过复合，我们就可以定义

$$(-\triangle)^{-1}: L^2(\Omega) \longrightarrow L^2(\Omega).$$

从此之后，当我们谈论 $(-\triangle)^{-1}$ 的时候，我们总假设它的定义域和值域都是 $L^2(\Omega)$。

由于右边竖列的箭头是紧算子，复合之后的算子 $(-\triangle)^{-1}$ 也是紧的。我们现在验证 $(-\triangle)^{-1}$ 是自伴算子：对任意的 $f_1, f_2 \in L^2(\Omega)$，令 $u_1 = (-\triangle)^{-1} f_1, u_2 = (-\triangle)^{-1} f_2$，我们知道，$u_1, u_2 \in H^1_0(\Omega)$，所以，

$$((-\triangle)^{-1} f_1, f_2) = (u_1, \triangle u_2) = -(\nabla u_1, \nabla u_2),$$

这里，我们用到了 $u_1 \in H_0^1(\Omega)$ 事实，请参考引理492。类似地，我们还有

$$(f_1, (-\triangle)^{-1} f_2) = -(\nabla u_1, \nabla u_2).$$

这就证明了自伴性。此时，就可以对 $(-\triangle)^{-1}$ 使用抽象的紧算子理论了。

另外，我们之前还证明了 $(-\triangle)^{-1}$ 是连续的正线性算子，即对任意的 $f \in L^2(\Omega)$，我们有

$$\left((-\triangle)^{-1} f, f\right)_{L^2} \geqslant 0.$$

这个不等式中等号成立当且仅当 $f = 0$。这表明，$(-\triangle)^{-1}$ 的所有特征值都是正实数（没有 0!）。

运用 Hilbert-Schmidt 定理，那么，我们可把 $(-\triangle)^{-1}$ 的特征值的集合写成

$$\sigma((-\triangle)^{-1}) = \{\mu_k\}_{k \geqslant 1}$$

使得

$$\mu_1 \geqslant \mu_2 \geqslant \mu_3 \geqslant \cdots, \quad \lim_{k \to \infty} \mu_k = 0.$$

我们用 $\varphi_k(x) \in L^2(\Omega)$ 表示与 μ_k 相对应的长度（L^2-范数）为 1 的特征函数。我们再令 $\lambda_k = \mu_k^{-1}$，所以，

$$-\triangle \varphi_k = \lambda_k \varphi_k.$$

由于 $\varphi_k \in L^2(\Omega)$，根据 Dirichlet 问题的解，我们知道 $\varphi_k \in L^2(\Omega)$。综上所述，我们证明了如下关于 $(-\triangle)^{-1}$ 的谱分解定理：

定理 526. *假设 $\Omega \subset \mathbb{R}^n$ 是有界的开区域，那么，存在单调上升的无界序列*

$$0 < \lambda_1 \leqslant \lambda_2 \leqslant \lambda_3 \leqslant \cdots, \quad \lim_{k \to \infty} \lambda_k = +\infty,$$

以及 $L^2(\Omega)$ 的一组 Hilbert 基 $\{\varphi_k\}_{k \geqslant 1}$，使得

$$-\triangle \varphi_k = \lambda_k \varphi_k.$$

进一步，对任意的 $k \geqslant 1$，我们有 $\varphi_k \in H_0^1(\Omega)$。

注记. *这个定理的证明没有用到 Ω 边界的正则性。*

例子. *我们现在考虑 $\Omega = (0, d)^n$ 是一个正方体的情况，其中 $d > 0$ 是常数。（我们注意到 Ω 的边界并不是光滑的。）我们要详细地计算 $-\triangle$ 在 Ω 上的所有特征值和特征函数。根据 Fourier 级数的理论，$L^2(\Omega)$ 的一族 Hilbert 基可以取作*

$$\left\{ d^{-\frac{n}{2}} e^{\frac{2\pi i}{d} k \cdot x} \right\}_{k \in \mathbb{Z}}.$$

由于我们要求 $\varphi\big|_{\partial\Omega} = 0$，所以，我们希望选取希望 sin 型的函数。

先研究 $d = 1$ 的情形，我们选取 $\left\{ \sin\left(\frac{k\pi}{d} x\right) \right\}_{k \in \mathbb{Z}_{\geqslant 1}}$ 作为备选。我们指出这里和 Fourier 级数的不同之处：指数中的 2π 变成了 π。很明显，我们有

- $-\triangle\Big(\sin\big(\frac{k\pi}{d}x\big)\Big)=\frac{\pi^2k^2}{d^2}\sin\big(\frac{k\pi}{d}x\big)$;
- $\sin\big(\frac{k\pi}{d}x\big)\Big|_{\partial(0,d)}=0$，所以，根据一维的 H_0^1 空间的描述，我们有 $\sin\big(\frac{k\pi}{d}x\big)\in H_0^1\big((0,d)\big)$;
- 当 $k\neq l$ 时，我们有如下的正交性：

$$\Big(\sin\big(\frac{k\pi}{d}x\big),\sin\big(\frac{l\pi}{d}x\big)\Big)_{L^2}=\int_0^d\sin\big(\frac{k\pi}{d}x\big)\sin\big(\frac{l\pi}{d}x\big)dx=0.$$

为了说明这给出了 $-\triangle=-\dfrac{d^2}{dx^2}$ 的所有特征函数，我们在区间 $(0,d)$ 上解（常）微分方程

$$-u''=\lambda^2u,$$

其中 $\lambda>0$。注意到，我们要找的 u 满足 $u\in H_0^1\big((0,d)\big)$，根据 *Sobolev* 嵌入定理，$u\in C^0\big((0,d)\big)$。据此，根据方程，$u''\in C^0$，所以，$u\in C^2$，再代入方程，$u''\in C^2$，所以，$u\in C^4$。如此迭代，我们知道 u 是足够光滑的函数，从而可以能用经典的常微分方程理论（解存在唯一）。所以，这个方程的通解可写成

$$u(x)=Ae^{i\lambda x}+Be^{-i\lambda x},$$

其中，A,B 是复数。再利用 $u(0)=u(d)=0$，我们有

$$A+B=0,\quad Ae^{i\lambda d}+Be^{-i\lambda d}=0,$$

这说明

$$B=-A,\quad A\sin(\lambda d)=0.$$

由于 $A\neq0$，所以，

$$u(x)=C\sin\big(\frac{\pi}{d}kx\big).$$

这说明，我们已经列出了所有的特征函数和特征值。

现在假设维数是 n，我们假设 $k=(k_1,\cdots,k_n)\in(\mathbb{Z}_{\geqslant1})^n$，那么

$$\left\{\prod_{j=1}^n\sin\big(\frac{\pi i}{d}k_j\cdot x_j\big)\right\}_{k_1,\cdots,k_n\in(\mathbb{Z}_{\geqslant1})^n}$$

是所有的特征函数。

我们首先计算

$$-\triangle\Big(\prod_{j=1}^n\sin\big(\frac{\pi i}{d}k_j\cdot x_j\big)\Big)=\frac{\pi^2|k|^2}{d^2}\prod_{j=1}^n\sin\big(\frac{\pi i}{d}k_j\cdot x_j\big).$$

其次，我们仍然有正交关系：当 $k\neq k'$ 时，我们就有

$$\int_{(0,d)^n}\prod_{j=1}^n\sin\big(\frac{\pi i}{d}k_j\cdot x_j\big)\prod_{j=1}^n\sin\big(\frac{\pi i}{d}k'_j\cdot x_j\big)dx=0.$$

对于 $k\in(\mathbb{Z}_{\geqslant 1})^n$，我们定义

$$\varphi_k=\prod_{j=1}^n\sin\left(\frac{\pi i}{d}k_j\cdot x_j\right),$$

我们来证明 $\varphi_k\in H_0^1\left((0,d)^n\right)$：这显然是一个 $H^1\left((0,d)^n\right)$ 中的函数，下面用光滑的有紧支集的函数来逼近它：由于对每个 j, $\sin\left(\dfrac{\pi i}{d}k_j\cdot x\right)\in H_0^1((0,d))$，所以，存在 $\{\varphi_j^{(p)}(x)\}_{p\geqslant 1}\subset C_0^\infty((0,d))$，使得

$$\lim_{p\to\infty}\left\|\varphi_j^{(p)}(x)-\sin\left(\frac{\pi i}{d}k_j\cdot x\right)\right\|_{H^1}=0.$$

所以，

$$\left\|\prod_{j=1}^n\varphi_j^{(p)}(x_j)-\prod_{j=1}^n\sin\left(\frac{\pi i}{d}k_j\cdot x_j\right)\right\|_{H^1}^2=\prod_{j=1}^n\left\|\varphi_j^{(p)}-\sin\left(\frac{\pi i}{d}k_j\cdot x_j\right)\right\|_{H_0^1}^2\to 0.$$

最终，为了证明这是所有的特征函数，我们来证明 $\{\varphi_k(x)\}_{(\mathbb{Z}_{\geqslant 1})^n}$ 构成 L^2 的 *Hilbert* 基。我们假设 $n\geqslant 2$（($n=1$ 的情形已经完成)），只要证明与 $\{\varphi_k\}_{k\in(\mathbb{Z}_{\geqslant 1})^n}$ 都垂直的函数只有 0 即可（这表明这些函数所张成的空间的闭包是整个 L^2)。这与证明高维的 *Fourier* 级数是一组基是完全一样的（利用 *Fubini* 定理)，请参考上学期 *5* 月 *14* 日的讲义。

我们现在研究 $\Omega=(0,d)^n$ 上 *Laplace* 算子的特征值分布问题。我们定义

$$\Psi(\lambda)=\left|\{\lambda_k|\lambda_k\leqslant\lambda\}\right|.$$

按照上述计算，我们有

$$\Psi(\lambda)=\left|\left\{(k_1,\cdots,k_n)\middle|k_i\geqslant 1,k_1^2+\cdots+k_n^2\leqslant\frac{d^2}{\pi^2}\lambda\right\}\right|.$$

这是在圆内的整点问题（*Gauss*)：$\Psi(\lambda)$ 是半径为 $\dfrac{d}{\pi}\sqrt{\lambda}$ 的球在第一卦限中的整点的个数，从而

$$\Psi(\lambda)\sim c_n\lambda^{\frac{n}{2}}\quad\Leftrightarrow\quad\lim_{\lambda\to\infty}\frac{\Psi(\lambda)}{c_n\lambda^{\frac{n}{2}}}=1,$$

其中 c_n 是只依赖于维数的常数，实际上，通过计算半径为 $\dfrac{d}{\pi}\sqrt{\lambda}$ 的球在第一卦限中的体积，我们知道

$$c_n=\frac{|B_n(1)|d^n}{(2\pi)^n},$$

其中 $B_n(1)=\dfrac{\pi^{\frac{n}{2}}}{\Gamma(\frac{n}{2}+1)}$ 是 $\mathbb{R}^n$ 中单位球的体积，所以，

$$\Psi(\lambda)\sim\frac{d^n}{(2\sqrt{\pi})^n\Gamma(\frac{n}{2}+1)}\lambda^{\frac{n}{2}}.$$

我们考虑 λ_k 的渐近大小，其中 $k\to\infty$。我们在上式中取 $\lambda=\lambda_k$，按照定义，$\Psi(\lambda_k)=k$，从而

$$k\sim\frac{d^n}{(2\sqrt{\pi})^n\Gamma(\frac{n}{2}+1)}\lambda_k^{\frac{n}{2}}=\frac{|\Omega|}{(2\sqrt{\pi})^n\Gamma(\frac{n}{2}+1)}\lambda_k^{\frac{n}{2}}.$$

所以，

$$\lambda_k \sim \frac{(2\pi)^2}{|B_n(1)|^{\frac{2}{n}}|\Omega|^{\frac{2}{n}}} k^{\frac{2}{n}}, \quad \lambda \to \infty.$$

我们将证明，这个公式对于一般的区域 Ω 都成立，这就是所谓的 *Weyl* 渐进公式。

注记. 如果 Ω 是连通的，我们可以证明 $\lambda_1 < \lambda_2$，也就是说第一特征值的重数是 1。

对于上面的例子，这一点很容易验证。在这个例子中，我们还看到，$\lambda_2 = \lambda_3 = \cdots = \lambda_{n+1}$。

79 Laplace 算子的特征函数的应用：利用特征函数刻画函数空间，各阶特征值的变分表述，相互包含区域的特征值比较，弱化版本的 Weyl 渐近公式：一般区域上的 Laplace 算子的特征值的增长率

二零二零年十二月二日，星期三，晴

上次课程，我们证明了如果 $\Omega \subset \mathbb{R}^n$ 是有界的开区域，那么，存在单调上升的无界序列

$$0 < \lambda_1 \leqslant \lambda_2 \leqslant \lambda_3 \leqslant \cdots, \quad \lim_{k\to\infty} \lambda_k = +\infty,$$

以及 $L^2(\Omega)$ 的一组 Hilbert 基 $\{\varphi_k\}_{k\geqslant 1}$，使得

$$-\triangle \varphi_k = \lambda_k \varphi_k.$$

进一步，对任意的 $k \geqslant 1$，我们有 $\varphi_k \in H_0^1(\Omega)$。

注记. 我们现在说明，特征函数 $\varphi_k(x)$ 是 Ω 上的光滑函数。

由于光滑性是局部性质，所以，我们只要证明对任意的 $\chi \in C_0^\infty(\Omega)$，我们都有 $\chi(x)\varphi_k(x) \in C^\infty(\Omega)$ 即可。为了书写方便，我们令 $u = \varphi_k(x)$ 并且 $\lambda = \lambda_k$，所以，

$$-\triangle u = \lambda u.$$

所以，

$$-\triangle(\chi \cdot u) = \chi\lambda u - 2\nabla\chi \cdot \nabla u - u\triangle\chi = \underbrace{(\lambda\chi - \triangle\chi)\, u}_{\in L^2} - \underbrace{2\nabla\chi \cdot \nabla u}_{\in L^2}.$$

按照定义，由于 $u \in H^1(\Omega)$，所以右边都是 L^2 的函数，根据 $\triangle$ 的正则性（注意到，由于 $\chi \cdot u$ 的支集远离 $\partial\Omega$，我们不妨假设 $\partial\Omega$ 是光滑的即可），我们知道 $\chi \cdot u \in H^2$。所以，根据 χ 选取的任意性，u 限制在任意的紧集 $K \subset \Omega$ 上都是 H^2 的。

重复这个做法，由于

$$-\triangle(\chi \cdot u) = \underbrace{(\lambda\chi - \triangle\chi)\, u}_{\in H^1} - \underbrace{2\nabla\chi \cdot \nabla u}_{\in H^1},$$

所以，u 限制在任意的紧集 $K \subset \Omega$ 上都是 H^3 的。

以此类推，对任意的 $\chi \in C_0^\infty(\Omega)$，$\chi u \in H^k(\mathbb{R}^n)$。根据 *Sobolev* 嵌入定理，我们就知道 $\chi \cdot u$ 是光滑的。

由于 $\{\varphi_k\}_{k\geqslant 1}$ 是 $L^2(\Omega)$ 的一族 Hilbert 基，所以

$$L^2(\Omega) = \Big\{ f = \sum_{k=1}^{\infty} c_k\varphi_k \Big| \sum_{k=1}^{\infty} |c_k|^2 < \infty \Big\}.$$

利用这些特征函数，我们可以刻画 $L^2(\Omega)$ 的子空间：

推论 527 (利用特征函数刻画函数空间)**.** *作为 $L^2(\Omega)$ 的子空间，我们有*

$$H_0^1(\Omega)=\left\{f=\sum_{k=1}^{\infty}c_k\varphi_k\Big|\sum_{k=1}^{\infty}|c_k|^2<\infty,\ \ \sum_{k=1}^{\infty}\lambda_k|c_k|^2<\infty\right\}.$$

进一步，$\left\{\frac{1}{\sqrt{\lambda_k}}\varphi_k\right\}_{k\geqslant 1}$ 是 $H_0^1(\Omega)$ 的 Hilbert 基，其中，我们在 $H_0^1(\Omega)$ 上用如下的内积：

$$(f,g)_{H_0^1}=\int_\Omega \nabla f\cdot\overline{\nabla g}dx.$$

证明： 这个命题的证明表面上是平凡的，其实蕴涵着分析上很容易被忽略的一个细微但是重要的细节。我们首先回忆 $(-\triangle)^{-1}$ 是自伴算子的证明：对任意的 $f_1,f_2\in L^2(\Omega)$，令 $u_1=(-\triangle)^{-1}f_1$，$u_2=(-\triangle)^{-1}f_2$，我们知道，$u_1,u_2\in H_0^1(\Omega)$，所以，

$$((-\triangle)^{-1}f_1,f_2)=(u_1,\triangle u_2)=-(\nabla u_1,\nabla u_2),$$

这里，我们用到了 $u_1\in H_0^1(\Omega)$ 事实，请参考引理492。这一部分的论断表明，对任意的 $u_1,u_2\in H_0^1(\Omega)$，我们有

$$(\triangle u_1,u_2)_{L^2}=-(\nabla u_1,\nabla u_2)_{L^2}=(u_1,\triangle u_2)_{L^2},$$

这个等式的成立强烈地依赖于 $u_1,u_2\in H_0^1(\Omega)$。

利用上面的等式，我们对每个 $f\in H_0^1(\Omega)$，我们有

$$(f,\varphi_k)_{H_0^1}=(\nabla f,\nabla\varphi_k)_{L^2}=(f,(-\triangle)\varphi_k)_{L^2}=\lambda_k(f,\varphi_k)_{L^2}.$$

令 $f=\varphi_l$，我们得到

$$\left(\frac{1}{\sqrt{\lambda_l}}\varphi_l,\frac{1}{\sqrt{\lambda_k}}\varphi_k\right)_{H_0^1}=\delta_{lk}.$$

这表明 $\left\{\frac{1}{\sqrt{\lambda_k}}\varphi_k\right\}_{k\geqslant 1}$ 是一族两两正交的向量，为了说明它们是一族 Hilbert 基，我们只要说明如果对任意的 $k\geqslant 1$，$\left(f,\frac{\varphi_k}{\sqrt{\lambda_k}}\right)_{H_0^1}=0$，那么 $f=0$。实际上，根据上面的公式（以及 $\lambda_k>0$），$\left(f,\frac{\varphi_k}{\sqrt{\lambda_k}}\right)_{H_0^1}=0$ 意味着 $(f,\varphi_k)_{L^2}=0$，由于 $\{\varphi_k\}_{k\geqslant 1}$ 是 $L^2(\Omega)$ 的 Hilbert 基，所以，$f=0$。

现在，每个 $f\in H_0^1$ 都可以写成

$$f=\sum_{k=1}^{\infty}c_k\varphi_k=\sum_{k=1}^{\infty}(\sqrt{\lambda_k}c_k)\cdot\frac{\varphi_k}{\sqrt{\lambda_k}},$$

从而，

$$\|f\|_{H_0^1}^2=\sum_{k\geqslant 1}\lambda_k|c_k|^2.$$

这就完成了证明。 □

注记. 给定 $f \in H^1(\Omega)$，我们自然可以把它写成

$$f = \sum_{k=1}^{\infty} c_k \varphi_k,$$

其中 $\|f\|_{H_0^1}^2 = \sum_{k\geqslant 1} \lambda_k |c_k|^2$，因为 f 先验的是 L^2 的。但是，我们未必有

$$\sum_{k\geqslant 1} \lambda_k |c_k|^2 < \infty.$$

比如说，当 $\Omega = (0, d)$ 时，我们有

$$\varphi_k(x) = \sin\left(\frac{k\pi}{d} x\right), \quad \lambda_k = \frac{\pi^2 k^2}{d^2}.$$

为了计算方便，我们令 $d = \pi$，即 $\Omega = (0, \pi)$，所以

$$\varphi_k(x) = \sin(kx), \quad \lambda_k = k^2.$$

令 $f(x) \equiv 1$，很明显，$f \in H^1(\Omega) - H_0^1(\Omega)$（它在端点处不为 0)。我们首先计算它在 L^2 意义下的展开:

$$(f, \varphi_k)_{L^2} = \int_0^{\pi} \sin(kx) dx = \frac{2}{k}.$$

从而，$c_k = \dfrac{2}{k}$。特别地，我们知道

$$\sum_{k\geqslant 1} \lambda_k |c_k|^2 = \sum_{k\geqslant 1} k^2 \frac{2}{k^2} = +\infty$$

注记. 对任意的 $u_1, u_2 \in H_0^1(\Omega)$，我们有

$$(\triangle u_1, u_2)_{L^2} = -(\nabla u_1, \nabla u_2)_{L^2} = (u_1, \triangle u_2)_{L^2},$$

我们令 $u_1 = f$，$u_2 = \varphi_k$，这表明

$$(\triangle f, \varphi_k)_{L^2} = (f, \triangle \varphi_k)_{L^2} = -\lambda_k c_k.$$

注意到 $\triangle f \in H^{-1}(\Omega)$ 未必是 L^2 中的元，所以，等式

$$-\triangle f \stackrel{?}{=} \sum_{k\geqslant 1} \lambda_k c_k \varphi_k,$$

未必有意义。然而，对于任意的 $g \in H_0^1(\Omega)$，由于 H^{-1} 可以实现为 $H_0^1(\Omega)$ 的对偶（连续)，所以，如果

$$\sum_{k\geqslant 1} b_k \varphi_k,$$

那么，

$$\langle -\triangle f, \overline{g} \rangle = \sum_{k\geqslant 1} \lambda_k c_k \overline{b_k},$$

这由 $(\triangle f, \varphi_k)_{L^2} = \lambda_k c_k$ 以及 H^{-1} 作为 $H_0^1(\Omega)$ 上的连续线性泛函的性质所决定。

我们可以对 $-\triangle$ 的特征值进行如下的变分表述:

定理 528. 假设 $u \in H_0^1(\Omega)$ 并且 $u \neq 0$，我们令

$$R(u) = \frac{\langle -\triangle u, \overline{u} \rangle_{L^2}}{\|u\|^2} \overset{\text{形式上}}{=} \frac{(-\triangle u, u)_{L^2}}{\|u\|^2}.$$

上面的定义之所以有意义是因为 $-\triangle u \in H^{-1}(\Omega)$ 而 $u \in H_0^1(\Omega)$。记 $H = H_0^1(\Omega)$，那么，我们有

1) 对于 λ_1，我们有

$$\lambda_1 = \inf_{u \in H, u \neq 0} R(u).$$

2) 对于 $k \geqslant 2$，我们有

$$\lambda_k = \inf_{\substack{u \perp \varphi_1, \cdots, u \perp \varphi_{k-1}, \\ u \neq 0}} R(u),$$

其中，上面的垂直关系是（可以是）用 L^2-内积来定义的。

3) 令 $\mathrm{Gr}_{k-1}(H)$ 为 H 的所有 $k-1$ 维线性子空间所构成的集合,其中 $k \geqslant 1$。对于 $P \in \mathrm{Gr}_{k-1}(H)$，令

$$\mu(P) = \inf_{u \perp P, u \neq 0} R(u).$$

那么，

$$\lambda_k = \sup_{P \in \mathrm{Gr}_{k-1}(H)} \mu(P).$$

3) 对于 $Q \in \mathrm{Gr}_k(H)$，定义

$$\nu(Q) = \sup_{u \in Q, u \neq 0} R(u).$$

那么，

$$\lambda_k = \inf_{Q \in \mathrm{Gr}_k(H)} \nu(Q).$$

证明: 我们先证明 1) 和 2):

考虑 $u = \sum_{k \geqslant 1} c_k \varphi_k \in L^2(\Omega)$。由于 $u \perp \varphi_1, \cdots, u \perp \varphi_k$，所以，

$$u = \sum_{j \geqslant k} \alpha_j \varphi_j$$

从而，（根据定理之前的注记），我们可以计算

$$\begin{aligned} R(u) &= \frac{\lambda_k |\alpha_k|^2 + \lambda_{k+1} |\alpha_{k+1}|^2 + \cdots}{|\alpha_k|^2 + |\alpha_{k+1}|^2 + \cdots} \\ &\geqslant \frac{\lambda_k |\alpha_k|^2 + \lambda_k |\alpha_{k+1}|^2 + \cdots}{|\alpha_k|^2 + |\alpha_{k+1}|^2 + \cdots} \\ &= \lambda_k. \end{aligned}$$

另外，$R(\varphi_k)=\lambda_k$，这就给出了 1) 和 2)。

为了证明 3)，我们先选取 $P_k=\varphi_1\wedge\cdots\wedge\varphi_{k-1}$，这个符号代表的是由 $\varphi_1,\cdots,\varphi_{k-1}$ 所张成的 $k-1$-维线性子空间。根据上面的计算，我们知道

$$\mu(P_k)=\lambda_k.$$

所以，

$$\sup_{P\in \mathrm{Gr}_{k-1}(H)}\mu(P)\geqslant\lambda_k.$$

为了证明反过来的不等式，我们利用下面的观察：

- 对每一个 $P\in \mathrm{Gr}_{k-1}$，总有不全为零的 $\alpha_1,\cdots,\alpha_k$，使得 $\alpha_1\varphi_1+\cdots+\alpha_k\varphi_k\perp P$。

这是一道标准的线性代数习题：为了让 $\alpha_1\varphi_1+\cdots+\alpha_k\varphi_k\perp P$，我们假设 $v_1,\cdots,v_{k-1}$ 是 P 的一组基，那么，这个垂直的条件等价于

$$\begin{cases}\alpha_1\cdot(\varphi_1,v_1)+\cdots+\alpha_k\cdot(\varphi_k,v_1)=0,\\ \cdots\quad\cdots,\\ \alpha_1\cdot(\varphi_1,v_{k-1})+\cdots+\alpha_k\cdot(\varphi_k,v_{k-1})=0.\end{cases}$$

这是 k 个未知数 $k-1$ 个方程，所以有解。

利用这个观察，我们有

$$\begin{aligned}\mu(P)&\leqslant R(\alpha_1\varphi_1+\cdots+\alpha_k\varphi_k)\\ &=\frac{\lambda_1|\alpha_1|^2+\cdots+\lambda_k|\alpha_k|^2}{|\alpha_1|^2+\cdots+|\alpha_k|^2}\\ &\leqslant\frac{\lambda_k|\alpha_1|^2+\cdots+\lambda_k|\alpha_k|^2}{|\alpha_1|^2+\cdots+|\alpha_k|^2}=\lambda_k.\end{aligned}$$

所以，

$$\sup_{P\in \mathrm{Gr}_{k-1}(H)}\mu(P)\leqslant\lambda_k.$$

这就证明了 3)。

最后，我们来证明 4)。通过选取 $Q_k=\varphi_1\wedge\cdots\wedge\varphi_k$，那么，我们有

$$\begin{aligned}\nu(Q_k)&=\sup_{\alpha_1,\cdots,\alpha_k\in\mathbb{C}}\frac{(-\triangle(\sum_{j\leqslant k}\alpha_j\varphi_j,\sum_{j\leqslant k}\alpha_j\varphi_j)}{(\sum_{j\leqslant k}\alpha_j\varphi_j,\sum_{j\leqslant k}\alpha_j\varphi_j)}\\ &=\sup_{\alpha_1,\cdots,\alpha_k\in\mathbb{C}}\frac{\lambda_1|\alpha_1|^2+\cdots+\lambda_k|\alpha_k|^2+\cdots}{|\alpha_1|^2+\cdots+|\alpha_k|^2}\\ &\leqslant\sup_{\alpha_1,\cdots,\alpha_k\in\mathbb{C}}\frac{\lambda_k|\alpha_1|^2+\cdots+\lambda_k|\alpha_k|^2+\cdots}{|\alpha_1|^2+\cdots+|\alpha_k|^2}\\ &=\lambda_k.\end{aligned}$$

又因为 φ_k 可以实现上面的最大值，所以

$$\nu(Q_k) = \lambda_k.$$

从而，

$$\inf_{Q\in \mathrm{Gr}_k(H)} \nu(Q) \leqslant \lambda_k.$$

为了证明反过来的不等式，我们利用下面的线性代数事实（证明仿照前述）：

- 对每一个 $Q \in \mathrm{Gr}_k$，总有不为零的 $u \in Q$，使得 $u \perp \varphi_1, \cdots, u \perp \varphi_{k-1}$。

据此以及 1) 和 2) 证明的过程，我们有

$$\mu(Q) \geqslant R(u) \geqslant \lambda_k.$$

也就是说，

$$\inf_{Q\in \mathrm{Gr}_k(H)} \nu(Q) \geqslant \lambda_k.$$

这就完成了证明 □

注记 (λ_1 与 Ponacaré 不等式)**.** *根据 1)，我们知道对任意的 $u \in H_0^1(\Omega)$，有*

$$\|\nabla u\|^2_{L^2(\Omega)} \geqslant \lambda_1 \|u\|^2_{L^2(\Omega)}.$$

这表明 λ_1 是使得 Poincaré 不等式成立的最佳常数。

推论 529 (相互包含区域的特征值比较)**.** *给定两个有界开区域 $\Omega_1 \subset \Omega_2 \subset \mathbb{R}^n$，对于每个 $k \geqslant 1$，我们都有*

$$\lambda_k(\Omega_1) \geqslant \lambda_k(\Omega_2).$$

证明： 通过将函数用零来延拓，我们已经构造了自然的（连续）嵌入映射：

$$\iota : H_0^1(\Omega_1) \to H_0^1(\Omega_2).$$

我们用 $\{\varphi_j\}_{j\geqslant 1}$ 表示 Laplace 算子在小区域 Ω_1 上的特征函数。根据上面的定理，我们有

$$\lambda_k(\Omega_1) = \sup_{\substack{u\in \varphi_1\wedge\cdots\wedge\varphi_k \\ u\neq 0}} R(u) = \sup_{\substack{u\in \iota(\varphi_1)\wedge\cdots\wedge\iota(\varphi_k) \\ u\neq 0}} R(u).$$

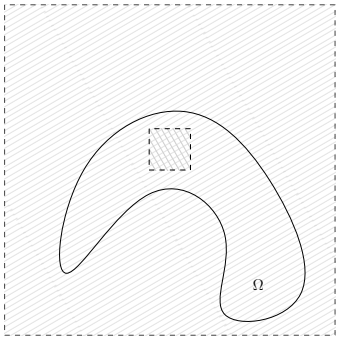
其中，第二个等号已经开始在 $H_0^1(\Omega_2)$ 中进行计算。所以，

$$\begin{aligned} \lambda_k(\Omega_1) & \sup_{\substack{u\in \iota(\varphi_1)\wedge\cdots\wedge\iota(\varphi_k) \\ u\neq 0}} R(u) \\ &= \nu\big(\iota(\varphi_1)\wedge\cdots\wedge\iota(\varphi_k)\big) \\ &\geqslant \lambda_k(\Omega_2). \end{aligned}$$

最后一个等号利用的是 $\lambda_k(\Omega_2)$ 在 4) 中的表述。证明完毕。 □

推论 530 (弱化版本的 Weyl 渐近公式)**.** 对任意的有界开区域 Ω，我们可以找到常数 c_1 和 c_2，使得当 $k \to \infty$ 时，我们有

$$c_1 k^{\frac{2}{n}} \leqslant \lambda_k(\Omega) \leqslant_2 k^{\frac{2}{n}}, \quad k \to \infty.$$



证明: 由于 Ω 是有界开区域，所以我们总能找到 $d, D > 0$，使得 $(0,d)^n \subset \Omega(0,D)^n$（通过平行移动，这不改变特征值）。从而，根据特征值的比较定理，我们有

$$\lambda_k\left((0,D)^n\right) \leqslant \lambda_k(\Omega) \leqslant \lambda_k\left((0,d)^n\right).$$

根据我们之前的例子，

$$\lambda_k\left((0,D)^n\right) \sim \frac{(2\pi)^2}{|B_n(1)|^{\frac{2}{n}}|D|^2} k^{\frac{2}{n}}, \quad \lambda_k\left((0,d)^n\right) \sim \frac{(2\pi)^2}{|B_n(1)|^{\frac{2}{n}}|d|^2} k^{\frac{2}{n}},$$

所以，我们选取

$$c_1 = \frac{(2\pi)^2}{|B_n(1)|^{\frac{2}{n}}|D|^2}, \quad c_2 = \frac{(2\pi)^2}{|B_n(1)|^{\frac{2}{n}}|d|^2},$$

即可。 □

80 $H_0^1(\Omega)\cap H^2(\Omega)$ 的刻画，保证分布落在 $H_0^1(\Omega)\cap H^m(\Omega)$ 中的必要条件，特征函数到边界的连续性（光滑边界），利用特征函数构造热核（频率空间的观点）

二零二零年十二月七日，星期一，晴

假设 $\Omega\subset\mathbb{R}^n$ 是有界的开区域，我们证明了如下的结论：存在单调上升的无界序列

$$0<\lambda_1\leqslant\lambda_2\leqslant\lambda_3\leqslant\cdots,\quad \lim_{k\to\infty}\lambda_k=+\infty,$$

以及 $L^2(\Omega)$ 的一组 Hilbert 基 $\{\varphi_k\}_{k\geqslant 1}$，使得

$$-\triangle\varphi_k=\lambda_k\varphi_k.$$

并且对于任意的 $k\geqslant 1$，我们有 $\varphi_k\in H_0^1(\Omega)$。从此往后，我们固定这样的一族特征函数（这种选取并不唯一）。

上次课，我们利用这些特征函数刻画了 $H_0^1(\Omega)$ 中的函数。如果 Ω 具有光滑的边界，我们还可以对指数更高的 Sobolev 空间进行一定的描述：

定理 531. *假设 Ω 具有光滑的边界，那么，作为 $L^2(\Omega)$ 的子空间，我们有*

$$H_0^1(\Omega)\cap H^2(\Omega)=\left\{u=\sum_{k=1}^{\infty}c_k\varphi_k\Big|\sum_{k=1}^{\infty}\lambda_k^2|c_k|^2<\infty\right\}.$$

进一步，假设 $m\geqslant 2$ 是整数并且 $u=\sum_{k=1}^{\infty}c_k\varphi_k\in L^2(\Omega)$ 满足

$$\sum_{k=1}^{\infty}\lambda_k^m|c_k|^2<\infty,$$

那么，$u\in H_0^1(\Omega)\cap H^m(\Omega)$。

对于 $u\in H_0^1(\Omega)\cap H^m(\Omega)$，如果令 $f=-\triangle u$，那么，$f\in H^{m-2}$。另外，如果 u 解如下的方程：

$$\begin{cases}-\triangle u=f,\\ u\big|_{\partial\Omega}=0.\end{cases}$$

并且 $f\in H^{m-2}(\Omega)$，那么，根据正则性理论（这里用到了 $\partial\Omega$ 是光滑的），$u\in H^m(\Omega)\cap H_0^1(\Omega)$。所以，我们就有如下的结论：

- 假设 $u\in H_0^1(\Omega)$（其中 Ω 是有界的光滑带边区域），那么，对任意的 $m\geqslant 1$，$u\in H^m(\Omega)$ 当且仅当 $\triangle u\in H^{m-2}(\Omega)$。

注记 (伪证). 我们先给出定理的一个"错误"的证明：因为

$$u=\sum_{k=1}^{\infty}c_k\varphi_k$$

在 $L^2(\Omega)$ 中成立，所以它也在 $\mathcal{D}'(\Omega)$ 的意义下成立。据此，在分布的意义下，我们有（因为求导数与分布的极限可以交换）

$$\triangle u\overset{\mathcal{D}'}{=}\sum_{k=1}^{\infty}c_k\triangle\varphi_k\overset{\mathcal{D}'}{=}-\sum_{k=1}^{\infty}\lambda_kc_k\varphi_k.$$

利用 $u\in H^2(\Omega)$，我们知道 $\triangle u\in L^2(\Omega)$，所以，

$$\sum_{k=1}^{\infty}|\lambda_kc_k|^2<\infty.$$

如果同学仔细阅读上面的证明，我们发现这个证明只用到了 $u\in H^2(\Omega)$ 而不需要 $u\in H_0^1(\Omega)$ 的条件。为了看出其中的错误，我们需要搞清楚

$$\triangle u\overset{\mathcal{D}'}{=}-\sum_{k=1}^{\infty}\lambda_kc_k\varphi_k$$

的含义：这个等式指的是

$$\triangle u\overset{\mathcal{D}'}{=}-\lim_{N\to\infty}\left(\sum_{k=1}^{N}\lambda_kc_k\varphi_k\right),$$

即对任意的 $\phi(x)\in\mathcal{D}(\Omega)$，我们有

$$\langle\triangle u,\phi\rangle=-\lim_{N\to\infty}\sum_{k=1}^{N}\lambda_kc_k\int_\Omega\varphi_k(x)\phi(x)dx.$$

这个基本等价于（弱于）在 $L^2(\Omega)$ 中的弱收敛：

$$\sum_{k=1}^{\infty}\lambda_kc_k\varphi_k\rightharpoonup-\triangle u.$$

我们知道，$x_k\rightharpoonup x_0$ 不意味着 x_k 的范数是控制的。

证明: 假设 $u\in H_0^1(\Omega)\cap H^2(\Omega)$，那么 $f=-\triangle u\in L^2(\Omega)$。我们把 f 用 $L^2(\Omega)$ 中的特征函数展开：

$$f(x)\overset{L^2}{=}\sum_{k=1}^{\infty}f_k\varphi_k(x).$$

特别地，我们知道

$$\sum_{k=1}^{\infty}|f_k|^2<\infty.$$

现在定义函数

$$v(x)=\sum_{k=1}^{\infty}\frac{f_k}{\lambda_k}\varphi_k(x).$$

由于 $\lambda_k \to +\infty$，v 显然是 L^2 的函数。另外，我们知道

$$\sum_{k=1}^{\infty} \lambda_k |\frac{f_k}{\lambda_k}|^2 < \infty.$$

所以，$v(x) \in H_0^1(\Omega)$。再者，利用 v 的表达式，我们知道

$$-\triangle v \overset{\mathcal{D}'}{=} -\sum_{k=1}^{\infty} \lambda_k \frac{f_k}{\lambda_k} \varphi_k = f.$$

所以，$v(x) \in H_)^1(\Omega)$ 和 $u \in H_0^1(\Omega)$ 都满足如下分布意义下的方程：

$$-\triangle v \overset{\mathcal{D}'}{=} f, \quad -\triangle u \overset{\mathcal{D}'}{=} f.$$

根据 Dirichlet 问题解的唯一性，我们知道

$$u = v.$$

特别地，我们有

$$u(x) \overset{L^2}{=} \sum_{k=1}^{\infty} \frac{f_k}{\lambda_k} \varphi_k(x).$$

所以，$c_k = \lambda_k^{-1} f_k$，从而

$$\sum_{k=1}^{\infty} \lambda_k^2 |c_k|^2 = \sum_{k=1}^{\infty} |f_k|^2 < \infty.$$

接下来，我们只要证明对任意整数 $m \geqslant 0$，$u = \sum_{k=1}^{\infty} c_k \varphi_k \in L^2(\Omega)$ 并且

$$\sum_{k=1}^{\infty} \lambda_k^m |c_k|^2 < \infty,$$

那么，$u \in H_0^1(\Omega) \cap H^m(\Omega)$。

我们对 m 进行归纳：$m = 0$（Poincaré 不等式）和 $m = 1$ 的情况已经证明。假设对一切小于 m 的整数命题都成立（$m \geqslant 2$），那么，

$$\triangle u \overset{\mathcal{D}'}{=} -\sum_{k=1}^{\infty} \lambda_k c_k \varphi_k.$$

我们注意到作用 $-\triangle$ 使得 Sobolev 指数降 2，所以，我们归纳的基础需要两个相邻的整数。根据 $m \geqslant 2$ 以及

$$\sum_{k=1}^{\infty} \lambda_k^m |c_k|^2 < \infty,$$

我们知道 $u \in H_0^1(\Omega)$ 并且（利用归纳假设）

$$\sum_{k=1}^{\infty} \lambda_k c_k \varphi_k \in H^{m-2}$$

所以，根据椭圆正则性，我们就知道 $u \in H_0^1(\Omega) \cap H^m(\Omega)$。 □

注记. 证明的过程表明，对任意的 $m \geqslant 1$ 时，如果

$$u \overset{L^2}{=} \sum_{k=1}^{\infty} c_k \varphi_k \in L^2(\Omega)$$

并且

$$\sum_{k=1}^{\infty} \lambda_k^m |c_k|^2 < \infty,$$

那么，$u \in H_0^1(\Omega) \cap H^m(\Omega)$ 并且存在常数 C，使得

$$\|u\|_{H^m(\Omega)} \leqslant C \sqrt{\sum_{k=1}^{\infty} \lambda_k^m |c_k|^2}.$$

证明的想法是在归纳假设中用所谓的椭圆估计：我们已经证明了如下的结论：

- $\Omega \subset \mathbb{R}^n$ 是有界光滑带边区域，$k \geqslant 1$ 是整数。假设 $u \in H^1(\Omega)$ 满足如下的边值问题：

$$\begin{cases} -\triangle u = f, \\ u\big|_{\partial\Omega} = g. \end{cases}$$

  如果 $f \in H^{k-1}(\Omega)$ 并且 $g \in H^{k+\frac{1}{2}}(\partial\Omega)$，那么，$u \in H^{k+1}(\Omega)$。

实际上，证明的过程还给出了如下的估计：存在仅依赖于 Ω 和 k 的常数，使得

$$\|u\|_{H^{k+1}(\Omega)} \leqslant C \left(\|f\|_{H^{k-1}(\Omega)} + \|g\|_{H^{k+\frac{1}{2}}(\Omega)} \right).$$

注记. 当 $m \geqslant 3$ 时，并非每个 $u \in H_0^1(\Omega) \cap H^m(\Omega)$ 中的函数都可以写成

$$u \overset{L^2}{=} \sum_{k=1}^{\infty} c_k \varphi_k \in L^2(\Omega)$$

并且

$$\sum_{k=1}^{\infty} \lambda_k^m |c_k|^2 < \infty.$$

我们考虑 $m = 3$ 的情况。任选 $f \in H^1(\Omega)$ 但是 $f \notin H_0^1(\Omega)$，考虑如下方程的解

$$\begin{cases} -\triangle u = f, \\ u\big|_{\partial\Omega} = 0. \end{cases}$$

我们知道，$u \in H_0^1(\Omega) \cap H^3(\Omega)$。如果 $u \overset{L^2}{=} \sum_{k=1}^{\infty} c_k \varphi_k \in L^2(\Omega)$ 满足

$$\sum_{k=1}^{\infty} \lambda_k^3 |c_k|^2 < \infty,$$

根据定理中最后一部分的证明，$f = -\triangle u \in H_0^1(\Omega)$，矛盾。

推论 532. 假设 Ω 具有光滑的边界，对任意的 $k \geqslant 1$，特征函数 $\varphi_k(x) \in C^\infty(\Omega) \cap C\left(\overline{\Omega}\right)$（连续到边界）。特别地，$\varphi_k$ 在边界上的限制是 0。

证明: 根据定理，我们知道 $\varphi_k \in H^N(\Omega)$ 其中 $2N > n$，所以，我们可以任意选取 φ_k 在 $H^N(\mathbb{R}^n)$ 中的扩张，从而，根据 Sobolev 嵌入定理，φ_k 在 $\overline{\Omega}$ 上连续。注意到，我们此时同时证明了 φ_k 在 Ω 内部是光滑的（一直到边界）。由于 $\varphi_k\big|_{\partial\Omega} \overset{H^{\frac{1}{2}}}{=} 0$，所以，$\varphi_k$ 在边界上的限制是 0。 □

热核的构造

对任意的 $t > 0$，对任意的 $(x, y) \in \Omega \times \Omega$，我们定义

$$p(t, x, y) = \sum_{k \geqslant 1} e^{-\lambda_k t} \varphi_k(x) \overline{\varphi_k(y)}.$$

我们已经证明了 λ_k 具有多项式的衰减并且 $\varphi_k \in C^\infty(\Omega)$，所以，上面的级数是（逐点）绝对收敛的（我们可以证明给定点 $x \in \Omega$，$\varphi_k(x)$ 对 k 是多项式增长的，不过这个级数数是逐点绝对收敛的这一点我们之后并不需要）。

我们首先证明，对任意的 $t \geqslant 0$ 时（包括 0），$p(t, x, y)$ 定义出 $\mathcal{D}'(\Omega \times \Omega)$ 中的一个元素：对任意的 $\phi(x, y) \in \mathcal{D}(\Omega \times \Omega)$，要定义

$$\left\langle \sum_{k \geqslant 1} e^{-\lambda_k t} \varphi_k(x) \overline{\varphi_k(y)}, \phi(x, y) \right\rangle_{\mathcal{D}'(\Omega\times\Omega)\times\mathcal{D}(\Omega\times\Omega)}.$$

为此，我们先理解其中一项 $\varphi_k(x)\overline{\varphi_k(y)}$ 的贡献。由于这是一个光滑函数，所以

$$I_k = \left\langle \varphi_k(x)\overline{\varphi_k(y)}, \phi(x, y) \right\rangle = \int_{\Omega\times\Omega} \varphi_k(x)\overline{\varphi_k(y)}\phi(x, y)dxdy.$$

根据 Cauchy-Schwarz 不等式，我们有

$$|I_k| \leqslant \|\varphi_k(x)\overline{\varphi_k(y)}dx\|_{L^2(\Omega\times\Omega)} \|\phi(x, y)\|_{L^2(\Omega\times\Omega)}.$$

由于 $\varphi_k(x)$ 是单位化的，所以

$$|I_k| \leqslant \|\phi(x, y)\|_{L^2(\Omega\times\Omega)}.$$

另外，我们可以把 I_k 写成

$$I_k = \frac{1}{\lambda_k^2} \int_{\Omega\times\Omega} (-\triangle_x \varphi_k(x)) \left(\overline{-\triangle_y \varphi_k(y)}\right) \phi(x, y)dxdy.$$

由于上述 $\phi(x, y)$ 的支集是紧的并且所有的函数都是光滑的，所以，我们可以进行分部积分（这恰好就是证明 Riemann-Lebesgue 引理的想法!）来得到

$$I_k = \frac{1}{\lambda_k^2} \int_{\Omega\times\Omega} \varphi_k(x)\overline{\varphi_k(y)}(\triangle_x\triangle_y\phi)(x, y)dxdy.$$

重复这个过程，对于正偶数 $2N$，我们有

$$I_k = \frac{1}{\lambda_k^{2N}} \int_{\Omega\times\Omega} \varphi_k(x)\overline{\varphi_k(y)}(\triangle_x^N \triangle_y^N \phi)(x,y)dxdy.$$

我们已经证明过 $\lambda_k \geqslant ck^{-\frac{2}{n}}$，下面我们将选取 $N = n$（这当然不是最优的）。现在假设 $\mathrm{supp}\phi(x,y) \subset K$，那么，重复上面对于 I_k 的控制，我们就有

$$|I_k| \leqslant \frac{1}{\lambda_k^{2N}} \|\triangle_x^N \triangle_y^N \phi)(x,y)\|_{L^2(K)} \leqslant \frac{1}{\lambda_k^{2N}} \sup_{\substack{(x,y)\in K,\\ |\alpha|\leqslant 2N_0}} |\partial^\alpha \phi(x,y)|$$

所以，我们定义

$$\begin{aligned}\left\langle \sum_{k\geqslant 1} e^{-\lambda_k t}\varphi_k(x)\overline{\varphi_k(y)}, \phi(x,y)\right\rangle &:= \sum_{k\geqslant 1} e^{-\lambda_k t}\left\langle \varphi_k(x)\overline{\varphi_k(y)}, \phi(x,y)\right\rangle \\ &= \lim_{m\to\infty}\sum_{k=1}^{m} e^{-\lambda_k t}\left\langle \varphi_k(x)\overline{\varphi_k(y)}, \phi(x,y)\right\rangle \\ &= \lim_{m\to\infty}\sum_{k=1}^{m} e^{-\lambda_k t} I_k.\end{aligned}$$

由于当 $N = n$ 时，λ_k^{-2N} 是绝对可和的，所以，上面是良好定义的。另外，根据 I_N 的估计，我们还有

$$\begin{aligned}\left|\left\langle \sum_{k\geqslant 1} e^{-\lambda_k t}\varphi_k(x)\overline{\varphi_k(y)}, \phi(x,y)\right\rangle\right| &\leqslant \sum_{k=1}^{\infty} e^{-\lambda_k t}|I_k| \\ &\leqslant C \sup_{\substack{(x,y)\in K,\\ |\alpha|\leqslant 2N_0}} |\partial^\alpha \phi(x,y)|.\end{aligned}$$

这表明我们定义出了 $\mathcal{D}'(\Omega\times\Omega)$ 中的分布。

我们注意到，上面所得（I_k 的）估计是不依赖于 t。所以，对于任意的试验函数

$$\phi(t,x,y) \in \mathcal{D}\big((0,\infty)\times\Omega\times\Omega\big),$$

我们可以定义

$$\begin{aligned}\left\langle \sum_{k\geqslant 1} e^{-\lambda_k t}\varphi_k(x)\overline{\varphi_k(y)}, \phi(t,x,y)\right\rangle &:= \sum_{k\geqslant 1}\int_0^\infty e^{-\lambda_k t}\left\langle \varphi_k(x)\overline{\varphi_k(y)}, \phi(t,x,y)\right\rangle dt \\ &= \lim_{m\to\infty}\sum_{k=1}^{m}\int_0^\infty e^{-\lambda_k t}\left\langle \varphi_k(x)\overline{\varphi_k(y)}, \phi(t,x,y)\right\rangle dt.\end{aligned}$$

我们假设 $\mathrm{supp}(\phi(t,x,y)) \subset J\times K$，其中 $J \subset (0,\infty)$ 是紧集，$K \subset \Omega\times\Omega$ 是紧集，那么，根据之前的证明

$$\begin{aligned}\left|\left\langle \varphi_k(x)\overline{\varphi_k(y)}, \phi(t,x,y)\right\rangle\right| &\leqslant \frac{1}{\lambda_k^{2N}} \sup_{\substack{(x,y)\in K,\\ |\alpha|\leqslant 2N_0}} |\partial_{x,y}^\alpha \phi(t,x,y)| \\ &\leqslant \frac{1}{\lambda_k^{2N}} \sup_{\substack{(t,x,y)\in J\times K,\\ |\alpha|\leqslant 2N_0}} |\partial_{x,y}^\alpha \phi(t,x,y)|\end{aligned}$$

从而，我们有如下（很粗糙）的估计：

$$\left|\int_0^\infty e^{-\lambda_k t}\left\langle \varphi_k(x)\overline{\varphi_k(y)}, \phi(t,x,y)\right\rangle\right| \leqslant \frac{|J|}{\lambda_k^{2N}} \sup_{\substack{(t,x,y)\in J\times K,\\ |\alpha|\leqslant 2N_0}} |\partial_{x,y}^\alpha \phi(t,x,y)|.$$

上面的右边对 k 是可和的，所以，我们就有

$$\left|\left\langle \sum_{k\geqslant 1} e^{-\lambda_k t}\varphi_k(x)\overline{\varphi_k(y)}, \phi(t,x,y)\right\rangle\right| \leqslant C_{J,K} \sup_{\substack{(t,x,y)\in J\times K,\\ |\alpha|\leqslant 2N_0}} |\partial_{x,y}^\alpha \phi(t,x,y)|.$$

这说明

$$p(t,x,y) \in \mathcal{D}'\big((0,\infty)\times\Omega\times\Omega\big).$$

根据定义，作为 $(0,\infty)\times\mathcal{D}\times\mathcal{D}$ 上的分布，我们有

$$\sum_{k\geqslant 1} e^{-\lambda_k t}\varphi_k(x)\overline{\varphi_k(y)} \overset{\mathcal{D}'}{=} \lim_{m\to\infty}\sum_{k=1}^m e^{-\lambda_k t}\varphi_k(x)\overline{\varphi_k(y)}.$$

由于在分布的意义下，求导数与极限是可以交换的，所以，我们可以逐项求导，从而

$$\begin{aligned}
&\left(\partial_t - \frac{1}{2}(\triangle_x+\triangle_y)\right)\left(\sum_{k\geqslant 1} e^{-\lambda_k t}\varphi_k(x)\overline{\varphi_k(y)}\right)\\
=&\lim_{m\to\infty}\sum_{k=1}^m \left(\partial_t - \frac{1}{2}(\triangle_x+\triangle_y)\right)\left(e^{-\lambda_k t}\varphi_k(x)\overline{\varphi_k(y)}\right)\\
=&\lim_{m\to\infty}\sum_{k=1}^m -\lambda_k e^{-\lambda_k t}\varphi_k(x)\overline{\varphi_k(y)} + \frac{\lambda_k}{2}\lambda_k e^{-\lambda_k t}\varphi_k(x)\overline{\varphi_k(y)} + \frac{\lambda_k}{2}\lambda_k e^{-\lambda_k t}\varphi_k(x)\overline{\varphi_k(y)}\\
=&0.
\end{aligned}$$

这就说明了作为 $(0,\infty)\times\mathcal{D}\times\mathcal{D}$ 上的分布，热核 $p(t,x,y)$ 满足如下的方程：

$$\left(\partial_t - \frac{1}{2}(\triangle_x+\triangle_y)\right)p(t,x,y) = 0.$$

注记 (热核的两个看法)**.** 到目前为止，我们对热核 $p(t,x,y)$ 有两种看法：

- $p(t,x,y)$ 是映射
$$[0,+\infty)\to\mathcal{D}'(\Omega\times\Omega),\quad t\mapsto p(t,x,y).$$
- $p(t,x,y)$ 是 $(0,\infty)\times\mathcal{D}\times\mathcal{D}$ 上的分布。

可以把热核函数 $p(t,x,y)$ 视作是映射

$$[0,+\infty)\to\mathcal{D}'(\Omega\times\Omega),\quad t\mapsto p(t,x,y).$$

我们现在证明，这个映射是连续映射，即对任意的 $\{t_j\}_{j\geqslant 1}\subset[0,+\infty)$，$t_k\to t_0$，在分布的意义下，我们有

$$p(t_k,x,y)\overset{\mathcal{D}'(\Omega\times\Omega)}{\longrightarrow}p(t_0,x,y).$$

按照定义，对任意的 $\phi(x,y)\in\mathcal{D}(\Omega\times\Omega)$，我们要证明如下的极限即可：

$$\lim_{j\to\infty}\left\langle\sum_{k\geqslant 1}e^{-\lambda_k t_j}\varphi_k(x)\overline{\varphi_k(y)},\phi(x,y)\right\rangle-\left\langle\sum_{k\geqslant 1}e^{-\lambda_k t_0}\varphi_k(x)\overline{\varphi_k(y)},\phi(x,y)\right\rangle=0.$$

这等价于证明

$$\sum_{k\geqslant 1}e^{-\lambda_k t_j}\int_{\Omega\times\Omega}\varphi_k(x)\overline{\varphi_k(y)}\phi(x,y)dxdy\to\sum_{k\geqslant 1}e^{-\lambda_k t_0}\int_{\Omega\times\Omega}\varphi_k(x)\overline{\varphi_k(y)}\phi(x,y)dxdy.$$

刚才的证明表明，

$$\begin{aligned}&\left|\sum_{k\geqslant 1}\left(e^{-\lambda_k t_j}-e^{-\lambda_k t_0}\right)\int_{\Omega\times\Omega}\varphi_k(x)\overline{\varphi_k(y)}\phi(x,y)dxdy\right|\\ \leqslant&\left|\sum_{k\geqslant 1}\left(e^{-\lambda_k t_j}-e^{-\lambda_k t_0}\right)|I_k|\right|\to 0.\end{aligned}$$

最后一步因为 $|I_k|$ 是绝对可和的（Lebesgue 控制收敛）。所以，我们证明了

$$p(t,x,y)\in C^0\left([0,+\infty),\mathcal{D}'(\Omega\times\Omega)\right).$$

我们还可以计算 $p(0,x,y)\in\mathcal{D}'(\Omega\times\Omega)$：任选 $f(x)g(y)\in\mathcal{D}(\Omega)\otimes\mathcal{D}(\Omega)$（在 $\mathcal{D}(\Omega\times\Omega)$ 中稠密），假设

$$f(x)\overset{L^2}{=}\sum_{k=1}^{\infty}a_k\varphi_k(x),\quad g(x)\overset{L^2}{=}\sum_{k=1}^{\infty}b_k\varphi_k(x),$$

那么，

$$\begin{aligned}\langle p(0,x,y),f(x)g(y)\rangle&=\sum_{k=1}^{\infty}\left\langle\varphi_k(x)\overline{\varphi_k(y)},f(x)\otimes g(y)\right\rangle\\&=\sum_{k=1}^{\infty}(g,\varphi_k)_{L^2}\overline{(\overline{f},\varphi_k)_{L^2}}\\&=\sum_{k=1}^{\infty}b_k\overline{a_k}=\int_{\Omega}f(x)g(x)dx.\end{aligned}$$

所以，

$$p(0,x,y)=\delta(x-y).$$

我们下面说明，对任意的 $t>0$，函数 $p(t,x,y)$ 是光滑函数。

我们任选 $\chi(x,y)\in C_0^\infty(\Omega\times\Omega)$，其中 $\chi(x,y)=\chi(x)\chi(y)$，我们假设 $\mathrm{supp}(\chi(x,y))=K$。那么，作为分布，我们有

$$\chi(x,y)p(t,x,y)=\sum_{k\geqslant 1}e^{-\lambda_k t}\chi(x,y)\varphi_k(x)\overline{\varphi_k(y)}$$

首先，

$$\begin{aligned}\|\chi(x,y)p(t,x,y)\|_{L^2(\Omega\times\Omega)}&\leqslant\sum_{k\geqslant 1}e^{-\lambda_k t}\|\varphi_k(x)\overline{\varphi_k(y)}\|_{L^2}\|\chi\|_{L^\infty}\\&=\|\chi\|_{L^\infty}\sum_{k\geqslant 1}e^{-\lambda_k t}\\&<\infty.\end{aligned}$$

这里，我们用到了 $e^{-\lambda_k t}$ 对于 k 是指数衰减的。

其次，

$$\nabla_x\left(\chi(x,y)p(t,x,y)\right)=\sum_{k\geqslant 1}e^{-\lambda_k t}\left(\nabla_x\chi(x,y)\varphi_k(x)\overline{\varphi_k(y)}+\chi(x,y)\nabla_x\varphi_k(x)\overline{\varphi_k(y)}\right)$$

所以，

$$\begin{aligned}\|\nabla_x(\chi(x,y)p(t,x,y))\|_{L^2(\Omega\times\Omega)}&\leqslant\sum_{k\geqslant 1}e^{-\lambda_k t}\|\varphi_k(x)\overline{\varphi_k(y)}\|_{L^2}\|\nabla\chi\|_{L^\infty}\\&\quad+\sum_{k\geqslant 1}e^{-\lambda_k t}\|\chi(x)\nabla\varphi_k(x)\|_{L^2}\\&\leqslant\sum_{k\geqslant 1}e^{-\lambda_k t}\|\nabla\chi\|_{L^\infty}+\sum_{k\geqslant 1}e^{-\lambda_k t}\|\chi(x)\nabla\varphi_k(x)\|_{L^2}\end{aligned}$$

和第一步类似，我们只需要控制 $\|\chi(x)\nabla\varphi_k(x)\|_{L^2(\Omega)}$：

$$\|\chi(x)\nabla\varphi_k(x)\|^2_{L^2(\Omega)}=\int_\Omega\chi(x)^2\nabla\varphi_k(x)\overline{\nabla\varphi_k(x)}.$$

现在都是光滑函数的等式，并且由于有了 χ 作为截断函数，上面的式子实际上与 $\partial\Omega$ 是没有关系的，所以，我们可以进行分部积分：

$$\begin{aligned}\|\chi(x)\nabla\varphi_k(x)\|^2_{L^2(\Omega)}&=-\int_\Omega\nabla(\chi(x)^2)\nabla\varphi_k(x)\overline{\varphi_k(x)}-\int_\Omega\chi(x)^2\triangle\varphi_k(x)\overline{\varphi_k(x)}\\&=-2\int_\Omega(\varphi_k\nabla\chi)\cdot\overline{(\chi\cdot\nabla\varphi_k)}+\lambda_k\int_\Omega\chi(x)^2\varphi_k(x)\overline{\varphi_k(x)}\\&\leqslant 2\|\varphi_k\nabla\chi\|^2_{L^2(\Omega)}+\frac{1}{2}\|\chi\cdot\nabla\varphi_k\|^2_{L^2(\Omega)}+\lambda_k\|\chi\varphi_k\|^2_{L^2(\Omega)}.\end{aligned}$$

在最后一步中，我们用了如下初等的不等式：

$$2ab\leqslant\lambda a^2+\frac{1}{\lambda}b^2,$$

其中 $\lambda > 0$ 可以任意选取（我们选取了 $\lambda = 2$）。注意到，上面不等式右边的第二项恰好是不等式左边的项，并且其系数小于 1，所以，它可以被左边“吃掉”，从而得到

$$\begin{aligned}\|\chi(x)\nabla\varphi_k(x)\|^2_{L^2(\Omega)} &\leqslant 4\|\varphi_k\nabla\chi\|^2_{L^2(\Omega)} + 2\lambda_k\|\chi\varphi_k\|^2_{L^2(\Omega)}\\ &\leqslant (4+2\lambda_k)C_1^2.\end{aligned}$$

其中，

$$C_m = \sup_{x\in K,|\alpha|\leqslant m}\|\partial^\alpha\chi\|_{L^\infty}.$$

代入到之前的等式中，利用 $e^{-\lambda_k t}$ 是指数衰减的，我们就知道

$$\sum_{k\geqslant 1} e^{-\lambda_k t}\|\chi(x)\nabla\varphi_k(x)\|_{L^2} < \infty.$$

这就证明了

$$\|\nabla_x(\chi(x,y)p(t,x,y))\|_{L^2(\Omega\times\Omega)} < \infty.$$

利用 x 与 y 之间的对称性，我们就知道

$$\chi(x,y)p(t,x,y) \in H^1(\Omega\times\Omega).$$

为了证明 $\chi(x,y)p(t,x,y) \in H^m(\Omega\times\Omega)$，我们先做如下的准备：利用分部积分，我们有

$$\begin{aligned}\|\chi(x)\nabla\partial^\alpha\varphi_k(x)\|^2_{L^2(\Omega)} &= -\int_\Omega \nabla(\chi(x)^2)\cdot\nabla\partial^\alpha\varphi_k(x)\overline{\partial^\alpha\varphi_k(x)} - \int_\Omega \chi(x)^2\triangle\partial^\alpha\varphi_k(x)\overline{\partial^\alpha\varphi_k(x)}\\ &= -2\int_\Omega(\partial^\alpha\varphi_k\nabla\chi)\cdot\overline{(\chi\cdot\nabla\partial^\alpha\varphi_k)} + \lambda_k\int_\Omega\chi(x)^2\partial^\alpha\varphi_k(x)\overline{\partial^\alpha\varphi_k(x)}\\ &\leqslant 2\|\partial^\alpha\varphi_k\nabla\chi\|^2_{L^2(\Omega)} + \frac{1}{2}\|\chi\cdot\nabla\partial^\alpha\varphi_k\|^2_{L^2(\Omega)} + \lambda_k\|\chi\partial^\alpha\varphi_k\|^2_{L^2(\Omega)}.\end{aligned}$$

所以，

$$\|\chi(x)\nabla\partial^\alpha\varphi_k(x)\|^2_{L^2(\Omega)} \leqslant 4\|\partial^\alpha\varphi_k\nabla\chi\|^2_{L^2(\Omega)} + 2\lambda_k\|\chi\partial^\alpha\varphi_k\|^2_{L^2(\Omega)}.$$

对 $|\alpha| = m$ 求和，我们就得到

$$\|\chi(x)\nabla^{m+1}\varphi_k(x)\|^2_{L^2(\Omega)} \leqslant 4\|\nabla\chi\cdot\nabla^m\varphi_k\|^2_{L^2(\Omega)} + 2\lambda_k\|\chi\nabla^m\varphi_k\|^2_{L^2(\Omega)}.$$

据此进行迭代，我们就得到

$$\|\chi(x)\nabla^m\varphi_k(x)\|^2_{L^2(\Omega)} \leqslant C_m\lambda_k^m\|\varphi_k\|^2_{L^2(\Omega)}.$$

此时，

$$\begin{aligned}
&\sum_{k\geqslant 1} e^{-\lambda_k t}\nabla_x^{m_1}\nabla_y^{m_2}\left(\chi(x,y)p(t,x,y)\right)\\
=&\sum_{k\geqslant 1} e^{-\lambda_k t}\sum_{\substack{|\alpha_1|+|\alpha_2|\leqslant m_1,\\ |\beta_1|+|\beta_2|\leqslant m_2}}\nabla^{\alpha_1}\chi(x)\nabla^{\alpha_2}\varphi_k(x)\overline{\nabla^{\beta_1}\chi(y)\nabla^{\beta_2}\varphi_k(y)}\\
=&\sum_{k\geqslant 1} e^{-\lambda_k t}\sum_{\substack{|\alpha_1|+|\alpha_2|\leqslant m_1,|\beta_1|+|\beta_2|\leqslant m_2,\\ |\alpha_2|<m,|\beta_2|<m}}\nabla^{\alpha_1}\chi(x)\nabla^{\alpha_2}\varphi_k(x)\overline{\nabla^{\beta_1}\chi(y)\nabla^{\beta_2}\varphi_k(y)}\\
&+\sum_{k\geqslant 1} e^{-\lambda_k t}\chi(x)^2\nabla^m\varphi_k(x)\overline{\nabla\varphi_k(y)}+\sum_{k\geqslant 1} e^{-\lambda_k t}\chi(x)^2\varphi_k(x)\overline{\nabla^m\varphi_k(y)}
\end{aligned}$$

第一个求和中的项的导数个数不超过 $m-1$，可以利用归纳法来解决，对于后面两项，它们的贡献可以被下面不等式控制

$$\begin{aligned}
&\|\sum_{k\geqslant 1} e^{-\lambda_k t}\chi(x)^2\nabla^m\varphi_k(x)\overline{\nabla\varphi_k(y)}+\sum_{k\geqslant 1} e^{-\lambda_k t}\chi(x)^2\varphi_k(x)\overline{\nabla^m\varphi_k(y)}\|_{L^2}\\
\leqslant&\sum_{k\geqslant 1} e^{-\lambda_k t}C_m\lambda_k^m\|\varphi_k\|^2_{L^2(\Omega)}.
\end{aligned}$$

利用 $e^{-\lambda_k t}$ 的指数衰减，上面的求和是有限的。

综上所述，我们证明了对任意的 $t>0$，任意的 m，$\chi(x,y)p(t,x,y)\in H^m(\mathbb{R}^n)$，所以，根据 Sobolev 嵌入定理，我们就知道 $\chi p(t,x,y)\in C^\infty(\Omega\times\Omega)$。由于光滑性是局部性质，所以我们就证明了对任意的 $t>0$，我们 $p(t,x,y)\in C^\infty(\Omega\times\Omega)$。

由于热核 $p(t,x,y)$ 在 $(0,\infty)\times\Omega\times\Omega$ 满足如下的方程：

$$\left(\partial_t-\frac{1}{2}\left(\triangle_x+\triangle_y\right)\right)p(t,x,y)=0,$$

所以，对于任意的 m 和任意的多重指标 α，我们还有

$$\partial_t^m\partial^\alpha p(t,x,y)=\left(\frac{1}{2}\left(\triangle_x+\triangle_y\right)\right)^m\partial^\alpha p(t,x,y)\in C^\infty_{x,y}(\Omega\times\Omega).$$

这就说明了

$$p(t,x,y)\in C^\infty\big((0,\infty)\times\Omega\times\Omega\big).$$

注记. 关于 $p(t,x,y)$ 的光滑性我们还可以仿照 *Cauchy-Riemann* 方程的情况进行证明：我们注意到

$$E(t,x,y)=\frac{H(t)}{(4\pi t)^n}e^{-\frac{|x|^2+|y|^2}{2t}}$$

是 $C^\infty(\mathbb{R}^1\times\mathbb{R}^n\times\mathbb{R}^n-\{0\})$ 上的光滑函数并且

$$\left(\partial_t-\frac{1}{2}\left(\triangle_x+\triangle_y\right)\right)p(t,x,y)=\delta_0,$$

其中，0 代表的是 $\mathbb{R}^1\times\mathbb{R}^n\times\mathbb{R}^n$ 上的原点。令 $P=\partial_t-\frac{1}{2}\left(\triangle_x+\triangle_y\right)$，我们先证明如下的引理：

引理 533. 给定有紧支集的分布 $c \in \mathcal{E}'(\mathbb{R}^1 \times \mathbb{R}^n \times \mathbb{R}^n)$，那么，$E * c$ 在 c 的支集之外是光滑的，即

$$E * c \in C^\infty\left(\mathbb{R}^1 \times \mathbb{R}^n \times \mathbb{R}^n - \operatorname{supp}(c)\right).$$

证明: 选取非负的 χ，使得它的支集在半径为 1 的小球（在 $\mathbb{R}^1 \times \mathbb{R}^n \times \mathbb{R}^n$ 中的）内并且其积分恰等于 1。我们把 $E * c$ 写成两部分:

$$u = E * c = \underbrace{(\chi_\varepsilon \cdot E) * c}_{\text{支集} \subset B_\varepsilon + \operatorname{supp}(c)} + \underbrace{((1-\chi_\varepsilon)E) * c}_{\text{光滑}}.$$

根据支集在卷积下的关系，以上两部分的支集有上面的表达。所以，u 至少在 $\mathbb{R}^1 \times \mathbb{R}^n \times \mathbb{R}^n - B_\varepsilon - \operatorname{supp}(c)$ 上光滑。令 $\varepsilon \to 0$，我们就说明了 u 在 $\operatorname{supp}(c)$ 之外光滑。 □

现在可以证明 $p(t,x,y)$ 在 $(0,+\infty)\times\Omega\times\Omega$ 上光滑，即证明 u 在每个点的附近为光滑函数即可: 我们任选点 $(t_0,x_0,y_0) \in (0,+\infty)\times\Omega\times\Omega$ 以及 (t_0,x_0,y_0) 处的一个半径为 2ε 小开球 $B(2\varepsilon)$，使得 $B(2\varepsilon) \subset (0,+\infty)\times\Omega\times\Omega$，其中 $\varepsilon > 0$。然后，选取 $\mathbb{R}\times\mathbb{R}^n\times\mathbb{R}^n$ 上的光滑函数 $\theta(z)$，使得

$$\begin{cases} 0 \leqslant \theta(z) \leqslant 1, \text{ 对任意的 } z \in \mathbb{R}\times\mathbb{R}^n\times\mathbb{R}^n \\ \theta\big|_{B(0,\varepsilon)} \equiv 1; \\ \theta\big|_{\mathbb{R}\times\mathbb{R}^n\times\mathbb{R}^n - B(0,2\varepsilon)} \equiv 0. \end{cases}$$

根据上面的构造，我们知道，

$$\partial_t\theta\big|_{B(\varepsilon)} \equiv= \partial_x\theta\big|_{B(\varepsilon)} \equiv \partial_y\theta\big|_{B(\varepsilon)} \equiv 0.$$

我们只需要证明 $\theta \cdot p(t,x,y)$ 光滑即可，因为这就说明 $p(t,x,y)$ 在 $B(\varepsilon)$ 上光滑。

利用卷积的基本性质，我们有如下的计算

$$\begin{aligned} \theta \cdot p(t,x,y) &= \delta_0 * (\theta \cdot p(t,x,y)) = P(E) * (\theta \cdot p) \\ &= E * \big(P(\theta \cdot p)\big) \\ &= E * \big((P\theta)\cdot p - \nabla_x\theta\nabla_x p - \nabla_y\theta\nabla_y p\big). \end{aligned}$$

所以 $\theta \cdot p$ 在 $\operatorname{supp}\big((P\theta)\cdot p - \nabla_x\theta\nabla_x p - \nabla_y\theta\nabla_y p\big)$ 之外是光滑的。由于

$$\big((P\theta)\cdot p - \nabla_x\theta\nabla_x p - \nabla_y\theta\nabla_y p\big)\Big|_{B(\varepsilon)} \equiv 0,$$

从而，$\theta \cdot p$ 在 $B(\varepsilon)$ 上光滑。

81 利用热方程刻画热核（物理空间的观点），热方程的极大值原理，热核的正性与对称性，区域上热核与全空间热核的比较定理

二零二零年十二月九日，星期三，阴

热核解线性热方程

利用热核，我们可以解热方程：

$$\begin{cases} \partial_t u - \triangle u = 0, \\ u\big|_{t=0} = u_0. \end{cases} \tag{2}$$

注记. 这个基本的想法很可能就是 *Fourier* 本人的观点：我们把 u_0 分解为最基本的波函数的组合：

$$u_0 = \sum_{k=1}^{\infty} c_k \varphi_k(x).$$

对一个基本的频率而言，我们知道

$$(\partial_t - \triangle)\left(e^{-t\lambda_k}\varphi_k(x)\right) = 0.$$

所以，我们希望 $u(t,x)$ 就是这些基本的波函数的组合，从而，

$$\begin{aligned} u(t,x) &= \sum_{k=1}^{\infty} c_k e^{-t\lambda_k}\varphi_k(x) \\ &= \sum_{k=1}^{\infty} (u_0(y), \lambda_k(y))_{L^2} e^{-t\lambda_k}\varphi_k(x) \\ &= \sum_{k=1}^{\infty} \int_\Omega u_0(y)\overline{\lambda_k(y)}dy e^{-t\lambda_k}\varphi_k(x). \end{aligned}$$

所以，在形式上，我们就有

$$\begin{aligned} u(t,x) &= \int_\Omega \sum_{k=1}^{\infty} u_0(y)\overline{\lambda_k(y)} e^{-t\lambda_k}\varphi_k(x)dy \\ &= \int_\Omega p(t,x,y)u_0(y)dy. \end{aligned}$$

我们现在做严格的推导。

如果 $u_0 \in L^2(\Omega)$，我们假设

$$u_0(x) = \sum_{k=1}^{\infty} c_k \varphi_k(x),$$

其中，$\sum_{k=1}^{\infty}|c_k|^2<\infty$。此时，对任意的 $t\geqslant 0$，我们首先定义

$$u(t,x)=\sum_{k=1}^{\infty}c_k e^{-t\lambda_k}\varphi_k(x).$$

很明显，对任意的 $t\geqslant 0$，$u(t,x)\in L^2(\Omega)$（算系数的平方和）。其次，由于 $t>0$ 时，$\{e^{-t\lambda_k}\}_{k\geqslant 1}$ 对于 k 是指数衰减的，所以，$u(t,x)\in H_0^1(\Omega)$，其中 $t>0$（用 $H_0^1(\Omega)$ 的刻画）。

我们现在把 $u(t,x)$ 视作是 $(0,+\infty)\times\Omega$ 上的分布：对于任意的试验函数 $\phi(t,x)\in\mathcal{D}\big((0,\infty)\times\Omega\big)$，我们可以定义

$$\begin{aligned}\langle u,\phi\rangle=\left\langle\sum_{k\geqslant 1}c_k e^{-\lambda_k t}\varphi_k(x),\phi(t,x)\right\rangle&:=\sum_{k\geqslant 1}c_k\int_0^{\infty}e^{-\lambda_k t}\left\langle\varphi_k(x),\phi(t,x)\right\rangle dt\\&=\lim_{m\to\infty}\sum_{k=1}^{m}c_k\int_0^{\infty}e^{-\lambda_k t}\left\langle\varphi_k(x),\phi(t,x)\right\rangle dt.\end{aligned}$$

我们假设 $\operatorname{supp}(\phi(t,x))\subset J\times K$，其中 $J=[t_*,T_*]\subset(0,\infty)$ 是紧集，$K\subset\Omega\times\Omega$ 是紧集。那么，根据 Cauchy-Schwarz 不等式

$$|\langle\varphi_k(x),\phi(t,x)\rangle|\leqslant\|\phi(t,x)\|_{L^2(K)}\leqslant|K|^{\frac{1}{2}}\sup_{x\in K}|\phi(t,x)|.$$

从而，我们有如下估计：

$$\begin{aligned}\left|\int_0^{\infty}e^{-\lambda_k t}\left\langle\varphi_k(x),\phi(t,x)\right\rangle\right|&\leqslant|K|^{\frac{1}{2}}\sup_{(t,x)\in J\times K}|\phi(t,x)|\int_{t_*}^{T_*}e^{-\lambda_k t}\\&\leqslant\frac{|K|^{\frac{1}{2}}e^{-\lambda_k t_*}}{\lambda_k}\sup_{(t,x)\in J\times K}|\phi(t,x)|.\end{aligned}$$

因为 $e^{-\lambda_k t_*}$ 仍然提供了指数衰减，所以下面式子右边对 k 是绝对可和的。从而

$$\begin{aligned}\left|\left\langle\sum_{k\geqslant 1}e^{-\lambda_k t}\varphi_k(x),\phi(t,x)\right\rangle\right|&\leqslant\sum_{k\geqslant 1}c_k\frac{|K|^{\frac{1}{2}}e^{-\lambda_k t_*}}{\lambda_k}\sup_{(t,x)\in J\times K}|\phi(t,x)|\\&\leqslant C(J,K,u_0)\sup_{(t,x)\in J\times K}|\phi(t,x)|.\end{aligned}$$

这说明

$$u(t,x)\in\mathcal{D}'\big((0,\infty)\times\Omega\big).$$

类似地，利用求导数与分布的极限可以交换，作为 $(0,\infty)\times\Omega$ 的分布，我们就有

$$(\partial_t-\triangle)u(t,x)\overset{\mathcal{D}'}{=}0.$$

我们可以采用上次课上证明 $p(t,x,y)$ 在 $(0,\infty)\times\Omega\times\Omega$ 上光滑的同样方法（仿照 Cauchy-Riemann 方程的情况）直接说明 $u(t,x)\in C^{\infty}((0,\infty)\times\Omega)$，证明的细节留给不放心的同学来验证。

我们再来证明

$$\lim_{t\to 0^+} u(t,x) \overset{L^2}{=} u_0(x).$$

实际上，我们有

$$\|u(t,x)-u_0(x)\|^2_{L^2(\Omega)} = \sum_{k=1}^{\infty} |c_k|^2 (e^{-\lambda_k t}-1)^2.$$

利用 Lebesguge 控制收敛，我们就有

$$\lim_{t\to 0^+} \|u(t,x)-u_0(x)\|_{L^2(\Omega)} = 0.$$

很明显，对任意的 $t>0$，我们有 $u(t,x)\in L^2(\Omega)$，所以，同样的证明给出了

$$u(t,x)\in C^0\left([0,+\infty), L^2(\Omega)\right).$$

如果假设 $u_0(x)\in C_0^\infty(\Omega)$，我们将 $u_0(x)$ 用特征函数展开：

$$u_0(x) = \sum_{k=1}^{\infty} c_k \varphi_k(x),$$

此时，我们可以利用 $u_0(x)$ 的光滑性得到

$$\begin{aligned} c_k &= \int_\Omega u_0(x)\overline{\varphi_k(x)}dx = \frac{-1}{\lambda_k}\int_\Omega u_0(x)\overline{\triangle\varphi_k(x)}dx \\ &= \frac{-1}{\lambda_k}\int_\Omega \triangle u_0(x)\overline{\varphi_k(x)}dx. \end{aligned}$$

所以，

$$c_k = \frac{(-1)^N}{\lambda_k^N}\int_\Omega \triangle^N u_0(x)\overline{\varphi_k(x)}dx.$$

从而，对任意的 $N\geqslant 1$，我们有

$$|c_k| \leqslant \frac{\|\triangle^N u_0\|_{L^2(\Omega)}}{\lambda_k^N}.$$

据此，对任意的自然数 m，我们都有

$$\sum_{k=1}^{\infty} \lambda_k^m |c_k|^2 < \infty.$$

注记. 如果假设 Ω 是光滑的，那么，对任意的 $t\geqslant 0$，我们都有

$$u(t,x)\in H^m(\Omega)$$

并且 $u(t,x)$ 的 H^m 范数是一致的（不依赖于 t）。特别地，我们可以重复上面的关于 L^2 的计算，这就可以证明

$$u(t,x)\in C^0\left([0,+\infty), H^m(\Omega)\right),$$

其中，m 是任意的正整数。特别地，我们知道对任意的 $t \geqslant 0$，$u \in C^0(\overline{\Omega})$（先把 $u(t,x)$ 延拓成 $H^m(\mathbb{R}^n)$ 中的函数然后用 *Sobolev* 嵌入定理）。

实际上，我们还可以说的更多：考虑复合映射

$$\begin{array}{ccc} \mathbb{R}_{\geqslant 0} & \longrightarrow & H^m(\Omega) \\ & \searrow & \downarrow \\ & & C^0(\Omega) \end{array}$$

所以，我们知道

$$u(t,x) \in C^0\left([0,+\infty), C^0(\Omega)\right),$$

由于上面的复合用到了 *Sobolev* 嵌入，所以，当 $t_j \to t_0$ 时，我们知道

$$\|u(t_j,x) - u(t_0,x)\|_{L^\infty} \to 0.$$

这个一致连续性是非常重要的。

我们现在说明，u 是 $[0,\infty)\times\overline{\Omega}$ 上的连续函数：任选 $(t_j,x_j) \in [0,\infty)\times\overline{\Omega}$，使得 $(t_j,x_j) \to (t_0,x_0)$，那么，

$$\begin{aligned} |u(t_j,x_j) - u(t_0,x_0)| &\leqslant |u(t_j,x_j) - u(t_0,x_j)| + |u(t_0,x_j) - u(t_0,x_0)| \\ &\leqslant \|u(t_j,\cdot) - u(t_0,\cdot)\|_{L^\infty} + |u(t_0,x_j) - u(t_0,x_0)|. \end{aligned}$$

对任意的 $\varepsilon > 0$，先选取 N_1，当 $j \geqslant N_1$ 时，我们有

$$\|u(t_j,\cdot) - u(t_0,\cdot)\|_{L^\infty} < \frac{1}{2}\varepsilon.$$

利用 $u(t_0,\cdot)$ 的连续性，再选取 N_2，当 $j \geqslant N_2$ 时，我们有

$$|u(t_0,x_j) - u(t_0,x_0)| < \frac{1}{2}\varepsilon.$$

所以，当 $j > \max(N_1,N_2)$ 时，我们就有

$$|u(t_j,x_j) - u(t_0,x_0)| < \varepsilon.$$

即 $u(t_j,x_j) \to u(t_0,x_0)$。

另外，如果 $u_0 \in \mathcal{D}(\Omega)$，我们还可以用热核来构造热方程的解

$$u(t,x) = \int_\Omega p(t,x,y)u_0(y)dy.$$

对于每个 t,x，这显然是良好定义的，因为 $u_0(y)$ 具有紧支集。实际上，仿照我们上次课程对热核的构造，当 (t,x) 固定时，我们有

$$p(t,x,\cdot) := \lim_{m\to\infty} \sum_{k\leqslant m} e^{-\lambda_k t}\varphi_k(x)\overline{\varphi_k(y)}.$$

搜益，上面的定义恰好是

$$u(t,x)=\langle p(t,x,y),u_0(y)\rangle=\sum_{k=1}^{\infty}c_ke^{-\lambda_kt}\varphi_k(x),$$

其中 $c_k=(u_0,\varphi_k)_{L^2}$。

利用 $p(t,x,y)\in C^0\left([0,\infty)\times\Omega\times\Omega\right)$，我们计算 $\lim\limits_{t\to0}u(t)$.

我们任选 $\phi(x)\in\mathcal{D}(\Omega)$，我们考虑

$$\begin{aligned}\lim_{t\to0}\langle u(t,x),\phi(x)\rangle&=\lim_{t\to0}\left\langle\langle p(t,x,y),u_0(y)\rangle,\phi(x)\right\rangle\\&=\lim_{t\to0}\left\langle p(t,x,y),\phi(x)\otimes u_0(y)\right\rangle\\&=\left\langle p(0,x,y),\phi(x)\otimes u_0(y)\right\rangle\\&=\sum_{k=1}^{\infty}\left\langle\varphi_k(x)\overline{\varphi_k(y)},\phi(x)\otimes u_0(y)\right\rangle\\&=\sum_{k=1}^{\infty}(u_0,\varphi_k)_{L^2}\overline{(\overline{\phi},\varphi_k)_{L^2}}\\&=(u_0,\overline{\phi})_{L^2}=\int_\Omega u_0(x)\phi(x)dx.\end{aligned}$$

所以，作为分布，我们有

$$\lim_{t\to0}u(t)\overset{\mathcal{D}'}{=}u_0(x).$$

热核的比较定理

我们用特征函数 φ_k 构造了热核。特征函数可以看作是特殊频率的波，所以，目前我们对热核的刻画是从频率空间的视角来做的。我们下面要在物理空间上刻画 $p(t,x,y)$。

我们注意到，任意给定 $u_0(x)\in\mathcal{D}(\Omega)$，我们构造的

$$u(t,x)=\int_\Omega p(t,x,y)u_0(x)dx$$

解如下的热方程：

$$\begin{cases}\partial_tu-\triangle u=0,\\u\big|_{t=0}=u_0.\end{cases}$$

并且 $u(t,x)\in C^\infty\big((0,\infty)\times\Omega\big)\cap C^0\big([0,\infty)\times\overline{\Omega}\big)$。所以，当 $p(t,x,y)$ 与试验函数 $u_0(y)$ 配对之后，我们得到的函数就有了物理空间上的描述。我们要利用这个方程来了解 $p(t,x,y)$，这就是对热核在物理空间上进行描述的基本想法。

我们试举一例来说明这个基本的想法并借此机会引入关于热传导方程极大值原理

命题 534. *假设 $u(t,x)\in C^\infty\big((0,\infty)\times\Omega\big)\cap C^0\big([0,\infty)\times\overline{\Omega}\big)$ 并且 u 在 $(0,\infty)\times\Omega$ 中满足热方程型的不等式：*

$$\partial_tu-\triangle u\leqslant0.$$

如果 $u\big|_{\partial\left[0,\infty)\times\overline{\Omega}\right)}\leqslant 0$，那么，在 $[0,\infty)\times\overline{\Omega}$ 上，$u(t,x)\leqslant 0$。

证明：我们考虑 u 的一个扰动：

$$u_\varepsilon(t,x)=u(t,x)+\varepsilon|x-a|^2.$$

此时，通过选取 $a\notin\overline{\Omega}$（然后固定这个 a），我们知道 $u_\varepsilon\big|_{\partial\left[0,\infty)\times\overline{\Omega}\right)}>0$。通过直接计算，我们还有

$$\partial_t u_\varepsilon-\triangle u_\varepsilon>0.$$

我们只要证明在 $[0,\infty)\times\overline{\Omega}$ 上，$u_\varepsilon(t,x)\geqslant 0$ 即可，因为我们令 $\varepsilon\to 0$ 就可以给出我们要证明的结论。

用反证法：如若不然，一定存在 $(t_0,x_0)\in(0,\infty)\times\Omega$，使得 $u_\varepsilon(t_0,x_0)>0$。我们现在考虑区域 $[0,t_0]\times\overline{\Omega}$，由于 u 在这个区域上是连续的，所以，存在 $(t_*,x_*)\in[0,t_0]\times\overline{\Omega}$，使得

$$u_\varepsilon(t_*,x_*)=\sup_{(t,x)\in[0,t_0]\times\Omega}u_\varepsilon(t,x).$$

根据 u_ε 的构造，我们知道 $t_*>0$ 并且 $x_*\in\Omega$。根据最大性，我们还知道 $\partial_t u_\varepsilon(t_*,x_*)\geqslant 0$ 并且 $\triangle u_\varepsilon(t_*,x_*)\leqslant 0$，所以，

$$\partial_t u_\varepsilon-\triangle u_\varepsilon\geqslant 0.$$

这就得到了矛盾。 □

类似地，我们有

推论 535. 假设 $u(t,x)\in C^\infty\big((0,\infty)\times\Omega\big)\cap C^0\big([0,\infty)\times\overline{\Omega}\big)$ 并且 u 在 $(0,\infty)\times\Omega$ 中满足热方程型的不等式：

$$\partial_t u-\triangle u\geqslant 0.$$

如果 $u\big|_{\partial\left[0,\infty)\times\overline{\Omega}\right)}\geqslant 0$，那么，在 $[0,\infty)\times\overline{\Omega}$ 上，$u(t,x)\geqslant 0$。

推论 536. 假设 $u(t,x)\in C^\infty\big((0,\infty)\times\Omega\big)\cap C^0\big([0,\infty)\times\overline{\Omega}\big)$ 并且 u 在 $(0,\infty)\times\Omega$ 中满足热方程：

$$\partial_t u-\triangle u=0.$$

如果 $u\big|_{\partial\left[0,\infty)\times\overline{\Omega}\right)}=0$，那么，$u(t,x)\equiv 0$。

注记. 假设 $u(t,x)\in C^\infty\big((0,\infty)\times\Omega\big)\cap C^0\big([0,\infty)\times\overline{\Omega}\big)$，那么，用同样的证明，我们可以说明

1) 如果 u 在 $(0,\infty)\times\Omega$ 上满足

$$\partial_t u-\triangle u\leqslant 0,$$

那么，

$$\sup_{(t,x)\in[0,\infty)\times\Omega}u(t,x)=\sup_{(t,x)\in\partial\left[0,\infty)\times\Omega\right)}u(t,x).$$

2) 如果 u 在 $(0,\infty)\times\Omega$ 上满足

$$\partial_t u - \triangle u \geqslant 0,$$

那么，

$$\inf_{(t,x)\in[0,\infty)\times\Omega} u(t,x) = \inf_{(t,x)\in\partial\left([0,\infty)\times\Omega\right)} u(t,x).$$

3) 如果 u 在 $(0,\infty)\times\Omega$ 上满足

$$\partial_t u - \triangle u = 0,$$

那么，

$$\inf_{(t,x)\in[0,\infty)\times\Omega} u(t,x) = \inf_{(t,x)\in\partial\left([0,\infty)\times\Omega\right)} u(t,x),$$

$$\sup_{(t,x)\in[0,\infty)\times\Omega} u(t,x) = \sup_{(t,x)\in\partial\left([0,\infty)\times\Omega\right)} u(t,x).$$

我们现在选取 $u_0(x)\geqslant 0$ 为处处非负的光滑的有紧支集的函数，我们并且假设 Ω 的边界是光滑的。那么，我们所构造的热方程的解 $u(t,x)$ 落在 $C^\infty\big((0,\infty)\times\Omega\big)\cap C^0\big([0,\infty)\times\overline{\Omega}\big)$ 中，并且在 $t=0$ 处是非负的，在 $\partial\Omega$ 上一直取 0。根据上面的极值原理，我们知道

$$u(t,x)\geqslant 0.$$

作为总结，我们有

$$u_0\in C_0^\infty(\Omega),\ \ u_0\geqslant 0 \quad\Rightarrow\quad \int_\Omega p(t,x,y)u_0(y)dy\geqslant 0.$$

我们固定 $(t,x)\in(0,\infty)\times\Omega$，那么，由于 $p(t,x,y)$ 对于 $y\in\Omega$ 是连续的，所以，上面的等式意味着 $p(t,x,y)\geqslant 0$，这就给出了热核的正性。特别的，$p(t,x,y)$ 是实值的。

注记. 我们也可以从代数的角度来证明这个结论：我们要说明，总是可以把 $\varphi_k(x)$ 选做实函数。对任意的特征值 λ，我们考虑 $-\triangle$ 的特征子空间 $E_\lambda\subset$。这是一个有限维的特征子空间，如果 $\varphi\in E_\lambda$，那么，

$$-\triangle\varphi=\lambda\varphi \quad\Rightarrow\quad -\triangle\overline{\varphi}=\lambda\overline{\varphi}.$$

所以，我们可以用选 $\mathrm{Re}(\varphi)$ 或者 $\mathrm{Im}(\varphi)$ 作为一个非零实特征函数。

然后在 E_λ 中考虑这个函数的正交补空间就可以用归纳法把 E_λ 中的所有特征函数都取成实特征函数。

特别的，按照定义，我们就有

$$p(t,x,y)=\sum_{k=1}^{\infty}e^{-\lambda_k t}\varphi_k(x)\varphi(y)$$

是实值的并且 $p(t,x,y)=p(t,y,x)$。

当然，我们对于热核的构造是基于特征函数的选取的，我们现在说明，即使换成另一组特征函数作为 Hilbert 基，它们所定义的

$$\widetilde{p}(t,x,y)=\sum_{k=1}^{\infty}e^{-\lambda_k t}\widetilde{\varphi}_k(x)\overline{\widetilde{\varphi}_k(y)}$$

与 $p(t,x,y)$ 是一致的。

对于任意的 $u_0(x)\in\mathcal{D}(\Omega)$，我们定义（这里我们假设 Ω 是光滑的）

$$\widetilde{u}(t,x)=\int_{\Omega}\widetilde{p}(t,y,x)u_0(x)dx.$$

那么，对这个新的热核重复之前的构造，我们就知道 $\widetilde{u}(t,x)$ 与 $u(t,x)$ 都在 $(0,\Omega)$ 上解热方程并且这两个函数在边界上的值是一样的，所以，对任意的试验函数 $u_0(x)\in\mathcal{D}(\Omega)$，我们就有

$$\int_{\Omega}p(t,y,x)u_0(x)dx=\int_{\Omega}\widetilde{p}(t,y,x)u_0(x)dx,$$

所以，$p(t,x,y)=\widetilde{p}(t,y,x)$。

综上所述，我们证明了

命题 537. *热核 $p(t,x,y)$ 的构造不依赖于具体的由 $-\triangle$ 的特征函数所给出的 $L^2(\Omega)$Hilbert 基的选取。进一步，我们有*

1) 正性：对任意的 $(t,x,y)\in(0,+\infty)\times\Omega\times\Omega$，$p(t,x,y)\geqslant 0$；

2) 对称性：对任意的 $(t,x,y)\in(0,+\infty)\times\Omega\times\Omega$，$p(t,x,y)=p(t,y,x)$。

我们现在讲热核与全空间 $\mathbb{R}^n$ 上的热核进行比较。在 $\mathbb{R}^n$ 上，我们已经构造了（物理空间上描述的）热核：

$$E(t,x)=\frac{H(t)}{(4\pi t)^{\frac{n}{2}}}e^{-\frac{|x|^2}{4t}}.$$

我们记

$$E(t,x,y)=\frac{H(t)}{(4\pi t)^{\frac{n}{2}}}e^{-\frac{|x-y|^2}{4t}}.$$

我们要比较 $p(t,x,y)$ 和 $E(t,x,y)$。类似地，我们也要将 $E(t,x,y)$ 与 $u_0(x)$ 进行配对，所以，我们定义

$$v(t,x)=\int_{\Omega}E(t,x,y)u_0(x)dx=\left(E(t,\cdot)*u_0\right)(x).$$

此时，根据 Lebesgue 控制收敛定理，当 $t>0$ 时，我们就有

$$\left(\partial_t-\triangle_x\right)v(t,x)=\int_{\Omega}\left(\partial_t-\triangle_x\right)E(t,x-y)u_0(y)dy=0.$$

另外，我们也很容易证明（利用 $\displaystyle\int_{\mathbb{R}^n}E(t,x)dx=1$）$v(t,x)$ 在 $[0,\infty)\times\overline{\Omega}$ 上连续并且

$$\lim_{t\to 0^+}\|v(t,x)-u_0(x)\|_{L^\infty}=0.$$

我们进一步要求 $u_0(x) \geqslant 0$。

那么，根据 v 的表达式，我们自然有

$$v(t,x) \geqslant 0, \quad \text{对任意的 } (x,t) \in [0,\infty) \times \overline{\Omega}.$$

此时，$u(t,x)$ 也解热方程并且 u 与 v 在 $t=0$ 处的初始值是一样的。这两个函数的不同之处可能在于对任意的 $t>0$，$u(t,\cdot)\big|_{\partial\Omega} = 0$。特别的，如果我们定义

$$U(t,x) = v(t,x) - u(t,x),$$

那么，$C^\infty\big((0,\infty)\times\Omega\big) \cap C^0\big([0,\infty)\times\overline{\Omega}\big)$，$U(t,x)$ 在 $(0,\infty)\times\Omega$ 中解热方程并且 $U(t,x)\big|_{\partial\left([0,\infty)\times\Omega\right)} \geqslant 0$。所以，对任意的 $(t,x) \in (0,\infty)\times\Omega$，我们都有 $U(t,x) \geqslant 0$，从而，

$$u_0 \in C_0^\infty(\Omega), \quad u_0 \geqslant 0 \quad \Rightarrow \quad \int_\Omega \big(E(t,x,y) - p(t,x,y)\big) u_0(y) dy \geqslant 0.$$

这表明，对任意的 $(t,x,y) \in (0,\infty)\times\Omega\times\Omega$，我们有

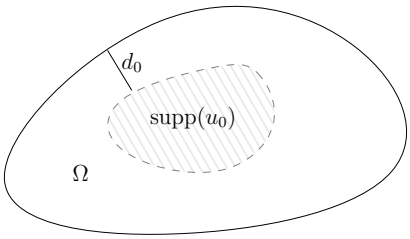
$$0 \leqslant p(t,x,y) \leqslant E(t,x,y).$$

练习. 假设 $\Omega_1 \subset \Omega_2$ 是两个光滑的有界带边光滑区域，我们用 $p_1(t,x,y)$ 和 $p_2(t,x,y)$ 分别代表它们的热核。那么，对于任意的 $(t,x,y) \in (0,\infty)\times\Omega_1\times\Omega_2$，我们有

$$0 \leqslant p_1(t,x,y) \leqslant p_2(t,x,y).$$

我们下面想对 $E(t,x,y) - p(t,x,y)$ 的上界进行控制，这样子，我们就可以对 $p(t,x,y)$ 有较为精确的控制。为此，我们需要对 $U(t,x)$ 的上界进行控制。

令 $d_0 = d\left(\mathrm{supp}(u_0), \partial\Omega\right)$，这是 u_0 的支集与 $\partial\Omega$ 之间的距离：



我们需要计算 $U(t,x)$ 在边界上 $\partial\Omega$ 的最大可能值。对任意的 $t>0$ 和 $x \in \partial\Omega$，我们有

$$\begin{aligned} U(t,x) = v(t,x) &= \frac{1}{(4\pi t)^{\frac{n}{2}}} \int_\Omega e^{-\frac{|x-y|^2}{4t}} u_0(y) dy \\ &\leqslant \frac{1}{(4\pi t)^{\frac{n}{2}}} \int_\Omega e^{-\frac{d_0^2}{4t}} u_0(y) dy. \end{aligned}$$

即

$$U(t,x) \leqslant \frac{e^{-\frac{d_0^2}{4t}}}{(4\pi t)^{\frac{n}{2}}} \int_\Omega u_0(y) dy.$$

我们考虑能使得上式右边值尽可能大的 t。通过对 t 求倒数，我们当 $t=t_0=\dfrac{d_0^2}{2n}$ 我们能取到最大值，并且在 $t<t_0$ 时，上式右边对 t 是递增的，在 $t>t_0$ 时，上式右边对 t 是递减的。所以，

$$U(t,x)\leqslant\begin{cases}\dfrac{e^{-\frac{d_0^2}{4t}}}{(4\pi t)^{\frac{n}{2}}}\int_\Omega u_0(y)dy, & \text{若 } t\leqslant\dfrac{d_0^2}{2n};\\ \dfrac{e^{-\frac{d_0^2}{4t_0}}}{(4\pi t_0)^{\frac{n}{2}}}\int_\Omega u_0(y)dy, & \text{若 } t\geqslant\dfrac{d_0^2}{2n}.\end{cases}$$

从而，

$$u_0\in C_0^\infty(\Omega),\ u_0\geqslant 0\ \Rightarrow\ \int_\Omega\big(E(t,x,y)-p(t,x,y)\big)u_0(y)dy\leqslant\begin{cases}\dfrac{e^{-\frac{d_0^2}{4t}}}{(4\pi t)^{\frac{n}{2}}}\int_\Omega u_0(y)dy, & \text{若 } t\leqslant\dfrac{d_0^2}{2n};\\ \dfrac{e^{-\frac{d_0^2}{4t_0}}}{(4\pi t_0)^{\frac{n}{2}}}\int_\Omega u_0(y)dy, & \text{若 } t\geqslant\dfrac{d_0^2}{2n}.\end{cases}$$

对任意的 $x_0,y_0\in\Omega$ 固定，我们选取 $\chi(x)$ 使得 $\chi\geqslant 0$，$\|\chi\|_{L^1}=1$ 并且 $\mathrm{supp}(\chi)$ 落在原点处半径为 1 的球中。令 $u_0(y)=\chi_\varepsilon(y-y_0)$。当 $\varepsilon_0\to 0$ 时，我们显然有

$$d\big(\mathrm{supp}(u_0),\partial\Omega\big)=d\left(y_0,\partial\Omega\right)+O(\varepsilon).$$

我们也有

$$\lim_{\varepsilon\to 0}\big(E(t,x_0,y)-p(t,x_0,y)\big)\chi_\varepsilon\left(y-y_0\right)dy=E(t,x_0,y_0)-p(t,x_0,y_)).$$

所以，我们就有（把 (x_0,y_0) 换成 (x,y)），我们就证明了如下的结论

定理 538. *假设 $\Omega\subset\mathbb{R}^n$ 是光滑的有界带边区域。那么，对任意的 $(t,x,y)\in(0,+\infty)\times\Omega\times\Omega$，我们有如下的热核比较公式：*

$$0\leqslant E(t,x,y)-p(t,x,y)\leqslant\begin{cases}\dfrac{1}{(4\pi t)^{\frac{n}{2}}}e^{-\frac{d(y,\partial\Omega)^2}{4t}}, & \text{若 } t\leqslant\dfrac{d(y,\partial\Omega)^2}{2n};\\ \dfrac{1}{(4\pi t_0(y))^{\frac{n}{2}}}e^{-\frac{d(y,\partial\Omega)^2}{4t_0(y)}}, & \text{若 } t\geqslant\dfrac{d(y,\partial\Omega)^2}{2n},\end{cases}$$

其中，$t_0(y)=\dfrac{d(y,\partial\Omega)^2}{2n}$。

特别的，当我们把 $p(t,x,y)$ 限制到对角线 $\Delta=\{(x,,x)|x\in\Omega\times\Omega\}$ 上，我们有

$$0\leqslant(4\pi t)^{-\frac{n}{2}}-p(t,x,x)\leqslant\frac{1}{(4\pi t)^{\frac{n}{2}}}e^{-\frac{d(x,\partial\Omega)^2}{4t}},\ \text{若 } t\leqslant\frac{d(x,\partial\Omega)^2}{2n}.$$

我们将证明，如果对 $x\in\Omega$ 积分，当 $t\to 0^+$ 时，

$$\int_\Omega\frac{1}{(4\pi t)^{\frac{n}{2}}}e^{-\frac{d(x,\partial\Omega)^2}{4t}}dx=O(t^{-\frac{n}{2}+1}),$$

然而，

$$\int_\Omega(4\pi t)^{-\frac{n}{2}}dx=(4\pi t)^{-\frac{n}{2}}|\Omega|$$

并且

$$\begin{aligned}\int_\Omega p(t,x,x)dx &= \sum_{k=1}^{\infty}\int_\Omega e^{-\lambda_k t}|\varphi_k(x)|^2 dx \\ &= \sum_{k=1}^{\infty} e^{-\lambda_k t}.\end{aligned}$$

所以，我们就给出了

$$\sum_{k=1}^{\infty} e^{-\lambda_k t} = (4\pi t)^{-\frac{n}{2}}\left(|\Omega| + O(t^{\frac{1}{2}})\right), \quad t \to 0^+.$$

利用这个渐近公式，我们就可以证明 Weyl 关于特征值的渐近公式。

82 热核的渐近展开，热核在对角线上的积分，Karamata 的 Tauber 型渐近公式，Weyl 渐近公式的证明，波前集的定义

二零二零年十二月十四日，星期一，晴

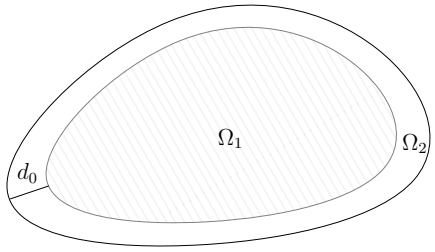
上次课我们证明了 $\mathbb{R}^n$ 中有界带边区域的热核在对角线上的限制 $p(t,x,x)$ 满足如下的估计：

$$0 \leqslant E(t,x,x) - p(t,x,x) \leqslant \begin{cases} \frac{1}{(4\pi t)^{\frac{n}{2}}} e^{-\frac{d(x,\partial\Omega)^2}{4t}}, & \text{如果 } t \leqslant \dfrac{d(x,\partial\Omega)^2}{2n}; \\ \frac{1}{(4\pi t_0(x))^{\frac{n}{2}}} e^{-\frac{d(x,\partial\Omega)^2}{4t_0(x)}}, & \text{如果 } t \geqslant \dfrac{d(x,\partial\Omega)^2}{2n}, \end{cases}$$

其中 $t_0(x) = \dfrac{d(x,\partial\Omega)^2}{2n}$。

假定 t 是给定的，我们令 $d_0 = \sqrt{2nt}$。

我们把 Ω 分成两个区域：



$$\Omega = \underbrace{\left\{x \in \Omega \middle| d(x,\partial\Omega) \geqslant d_0\right\}}_{\Omega_1} \cup \underbrace{\left\{x \in \Omega \middle| d(x,\partial\Omega) > d_0\right\}}_{\Omega_2}.$$

那么，我们把积分也分成两个部分

$$\int_\Omega E(t,x,x) - p(t,x,x)dx = \underbrace{\int_{\Omega_1} E(t,x,x) - p(t,x,x)dx}_{I_1} + \underbrace{\int_{\Omega_2} E(t,x,x) - p(t,x,x)dx}_{I_2}.$$

我们首先估计 I_1，此时，我们用热核比较估计的第二种情形，从而，

$$E(t,x,x) - p(t,x,x) \leqslant \frac{1}{(4\pi t)^{\frac{n}{2}}} e^{-\frac{d(x,\partial\Omega)^2}{4t}}.$$

此时，利用 Newton-Leibniz 公式，我们有

$$\begin{aligned} I_1 &\leqslant \frac{1}{(4\pi t)^{\frac{n}{2}}} \int_{\Omega_1} e^{-\frac{d(x,\partial\Omega)^2}{4t}} dx \\ &= -\frac{1}{(4\pi t)^{\frac{n}{2}}} \int_{\Omega_1} \left(\int_{d(x,\partial\Omega)}^{\infty} \frac{d}{d\tau}\left(e^{-\frac{\tau^2}{4t}}\right) d\tau \right) dx \\ &= \frac{1}{(4\pi t)^{\frac{n}{2}}} \frac{1}{2t} \int_{\Omega_1} \int_{d(x,\partial\Omega)}^{\infty} e^{-\frac{\tau^2}{4t}} \tau d\tau dx. \end{aligned}$$

根据 Fubini 公式以及 Ω_1 的定义（$d(x,\partial\Omega)\geqslant d_0$），我们有

$$\begin{aligned}I_1 &\leqslant \frac{1}{(4\pi t)^{\frac{n}{2}}}\frac{1}{2t}\int_{\Omega_1}\int_{d(x,\partial\Omega)}^{\infty} e^{-\frac{\tau^2}{4t}}\tau d\tau dx\\ &= \frac{1}{(4\pi t)^{\frac{n}{2}}}\frac{1}{2t}\int_{\Omega_1}\int_{0}^{\infty} \mathbf{1}_{\tau\geqslant d(x,\partial\Omega)}(\tau,x)e^{-\frac{\tau^2}{4t}}\tau d\tau dx\\ &= \frac{1}{(4\pi t)^{\frac{n}{2}}}\frac{1}{2t}\int_{0}^{\infty} e^{-\frac{\tau^2}{4t}}\tau S(\tau)d\tau.\end{aligned}$$

这里，

$$S(\tau)=\big|\{x\in\Omega|d(x,\partial\Omega)\leqslant\tau\}\big|.$$

引理 539. *假设 Ω 是光滑的有界带边区域，那么，存在只依赖于 Ω 的常数 C，使得*

$$S(\tau)\leqslant C\tau.$$

我们把引理的证明留到后面来处理。利用这个引理，我们就有

$$\begin{aligned}I_1 &\leqslant \frac{C}{(4\pi t)^{\frac{n}{2}}}\frac{1}{2t}\int_{0}^{\infty} e^{-\frac{\tau^2}{4t}}\tau^2 d\tau\\ &\leqslant \frac{C}{(4\pi t)^{\frac{n}{2}}}\frac{\sqrt{t}}{2}\underbrace{\int_{0}^{\infty} e^{-\frac{\tau^2}{4}}\tau^2 d\tau}_{\text{常数}}.\end{aligned}$$

所以，存在常数 C_1，使得

$$I_1\leqslant C_1 t^{-\frac{n}{2}}t^{\frac{1}{2}}.$$

现在来处理 I_2，也就是在区域 Ω_2 中的积分。此时，我们用热核比较估计的第二种情形，从而，

$$E(t,x,x)-p(t,x,x)\leqslant\frac{1}{(4\pi t)^{\frac{n}{2}}}e^{-\frac{d(x,\partial\Omega)^2}{4t}}\leqslant\frac{1}{(4\pi t)^{\frac{n}{2}}}.$$

所以，

$$\begin{aligned}I_2 &\leqslant \int_{\Omega_2}\frac{1}{(4\pi t)^{\frac{n}{2}}}dx=\frac{1}{(4\pi t)^{\frac{n}{2}}}S(\sqrt{2nt})\\ &\leqslant \frac{1}{(4\pi t)^{\frac{n}{2}}}C\sqrt{2nt}=C_2t^{-\frac{n}{2}}t^{\frac{1}{2}}.\end{aligned}$$

综合 I_1 和 I_2 的估计，我们就证明了存在常数 $C>0$，使得

$$\int_{\Omega}E(t,x,x)-p(t,x,x)\leqslant Ct^{-\frac{n}{2}}t^{\frac{1}{2}}.$$

我们现在来证明引理539。为此我们先证明如下的引理：

引理 540. *假设 $\Omega\subset\mathbb{R}^n$ 是光滑的有界带边区域，我们令*

$$\Omega_\alpha=\big\{x\in\Omega\big|d(x,\partial\Omega)<\alpha\big\},\quad \Sigma_\alpha=\big\{x\in\Omega\big|d(x,\partial\Omega)=\alpha\big\}.$$

那么，存在 $\alpha_0>0$，使得

1) 函数

$$\Omega_{\alpha_0} \to \mathbb{R}_{>0}, \quad x \mapsto d(x, \partial\Omega)$$

是 Ω_{α_0} 上的光滑函数。

2) 对任意的 $\alpha < \alpha_0$，Σ_α 是光滑的超曲面。

证明: 对任意的 $x_0 \in \partial\Omega$，我们考虑

$$B_{\frac{1}{2}\varepsilon}(x_0)\Omega \subset B_\varepsilon(x_0) \cap \Omega \subset \Omega.$$

我们要证明如下的论断：存在 ε，使得函数 $d(\cdot, \partial\Omega \cap B_\varepsilon)(x_0)$ 满足

1) $d(\cdot, \partial\Omega \cap B_\varepsilon)(x_0)$ 在 $B_{\frac{1}{2}\varepsilon}(x_0) \cap \Omega$ 的光滑函数。

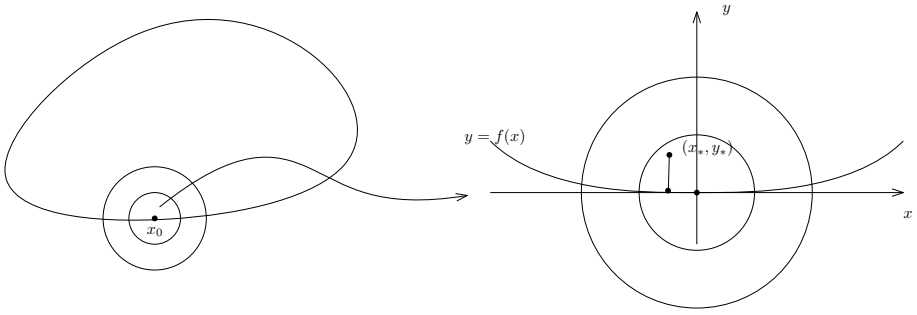
2) 对任意的 $\alpha < \dfrac{1}{2}\varepsilon$，$\Sigma_\alpha \cap B_{\frac{1}{2}\varepsilon}(x_0) \cap \Omega$ 是光滑的超曲面。

我们指出，在上面的构造中，我们之所以选取 $B_{\frac{1}{2}\varepsilon}(x_0)$ 和 $B_\varepsilon(x_0)$ 两个小球是因为对任意的 $x \in B_{\frac{1}{2}\varepsilon}(x_0) \cap \Omega$，我们有

$$d(x, \partial\Omega) = d(x, \partial\Omega \cap B_\varepsilon(x_0)).$$

假设上面的论断成立，根据紧性，我们总是可以选取有限个 $x_1, \cdots, x_m \in \partial\Omega$，对应这些点，我们有相应的 $\varepsilon_1, \cdots, \varepsilon_m > 0$，满足上面的性质并且 $B_{\frac{1}{2}\varepsilon_1}(x_1), B_{\frac{1}{2}\varepsilon_2}(x_2), \cdots, B_{\frac{1}{2}\varepsilon_m}(x_m)$ 是 $\partial\Omega$ 的开覆盖。此时，我们令 $\alpha_0 = \frac{1}{2}\min_{i\leqslant m}(\varepsilon_1, \cdots, \varepsilon_m)$ 即可。这就完成了命题的证明。

由于 ε 可以选取的足够小，所以，我们总是可以假设 $\partial\Omega$ 在 $B_\varepsilon(x_0)$ 上是函数图像（这里，我们用到了 $\partial\Omega$ 的光滑性）。另外，通过适当的选取直角坐标系（复合上一个 $\mathbb{R}^n$ 上的等距变换），我们总是可以假设 $x_0 = (0,0) = (x, y)$ 是原点，其中，$x \in \mathbb{R}^{n-1}$，$y \in \mathbb{R}$。在 $B_\varepsilon(0)$ 上，$\partial\Omega$ 是 $y = f(x)$ 的图像，其中，f 是 $\mathbb{R}^{n-1}$ 上的光滑函数并且 $df|_{x=0} = 0$（此时，$\partial\Omega$ 在 x_0 的切平面恰好是 $y = 0$ 这个平面）。



我们现在任选 $(x_*, y_*) \in \Omega \cap B_{\frac{1}{2}\varepsilon}(0)$，我们证明存在存在唯一的 $(x', y') \in \partial\Omega \cap B_\varepsilon(0)$，使得

$$d\big((x_*, y_*), \partial\Omega\big) = d\big((x_*, y_*), (x', y')\big).$$

这个问题等价于我们要找 $(x, f(x)) \in \partial\Omega$，使得如下的函数达到最小值：

$$F(x) = |x_* - x|^2 + |y_* - f(x)|^2.$$

这显然是关于 x 的光滑函数并且它的最小值一定能取到并且只能在 $B_\varepsilon(0)$ 中取到。如果 x' 是这样的一个极值点，那么，

$$dF\big|_{x=x'}=0 \quad\Leftrightarrow\quad 2(x_*-x')+2(y_*-f(x'))df\big|_{x=x'}=0.$$

这等价于

$$x_*+(y_*-f(x'))df(x')-x'=0.$$

当 $(x_*,y_*)=(0,0)$ 时，这个方程显然只有唯一的解。另外，考虑映射

$$\Phi: x'\mapsto x^*+(y^*-f(x'))df(x').$$

此时，

$$d\Phi(0)=\underbrace{(y_*-f(x'))}_{O(\varepsilon)}\ \underbrace{\nabla^2 f(x')}_{\text{在 } B_\varepsilon(0) \text{ 上有界}}\ -\ \underbrace{\nabla f(x')\cdot\nabla f(x')}_{\text{根据连续性，在 } B_\varepsilon(0) \text{ 上为 } O(\varepsilon^2)}\qquad -\mathrm{Id}.$$

所以，当 ε 足够小的时候，$d\Phi(0)$ 是可逆的。进一步缩小 ε，根据反函数定理，我们就知道存在唯一的 $(x',f(x'))\in B_\varepsilon(0)\cap\partial\Omega$，使得

$$x_*+(y_*-f(x'))df(x')-x'=0 \quad\Leftrightarrow\quad dF\big|_{x=x'}=0.$$

这表明，$(x',f(x'))$（唯一地）实现了 (x_*,y_*) 到 $\partial\Omega$ 的距离。特别地，反函数定理表明 x' 对于 (x_*,y_*) 是光滑依赖的，所以，

$$d\big((x_*,y_*),\partial\Omega\big)=d\big((x_*,y_*),(x',y')\big)=\sqrt{|x_*-x'|^2+|y_*-y'|^2}\big)$$

是 (x_*,y_*) 的光滑函数。

另外，假设 $d\big((x_*,y_*),\partial\Omega\big)=\alpha$，我们用直线段 L 连接 (x_*,y_*) 与 (x',y')，那么，在这个线段上的每个点到 $\partial\Omega$ 的距离都被 (x',y') 所实现。此时，我们知道对任意的 $v\in T_(x_*,y_*)\Sigma_\alpha$，$\nabla_v d\big((x_*,y_*),\partial\Omega\big)=0$；但是沿着 L 方向对 $d\big((x_*,y_*),\partial\Omega\big)$ 求导数必然为 ±1。从而，

$$\big|\nabla d\big(\cdot,\partial\Omega\big)\big|\equiv 1.$$

这就表明到边界距离为定值的点所构成的超曲面是光滑的。这就完成了引理的证明。 □

引理539的证明. 我们选取上面引理中的 α_0。我们只要对于 $\alpha<\alpha_0$，证明存在常数 C_1，使得

$$S(\alpha)=\big|\big\{x\in\Omega\big|d(x,\partial\Omega)<\alpha\big\}\big|\leqslant C_1\alpha$$

即可。实际上，我们可以选取

$$C=\max\left(C_1,\frac{\Omega}{\alpha_0}\right).$$

因为当 $\alpha\geqslant\alpha_0$ 时，我们有

$$S(\alpha)\leqslant|\Omega|=\frac{\Omega}{\alpha_0}\alpha.$$

对于任意的 $\alpha < \alpha_0$，我们有

$$\Omega_\alpha = \bigcup_{t<\alpha} \Sigma_t.$$

根据扭曲版本的 Fubini 定理（第二学期作业 8 问题 A6），我们有

$$\begin{aligned} S(\alpha) &= \int_{\Omega_\alpha} 1 dx = \int_0^\alpha \left(\int_{\Sigma_t} \frac{1}{|\nabla d(\cdot, \partial\Omega)|} d\sigma_t \right) dt \\ &= \int_0^\alpha |\Sigma_t| dt. \end{aligned}$$

根据反函数定理，每个 Σ_t 在局部上都可以由一族光滑依赖于 t 的函数图像实现，所以，当 $t < \alpha$ 较小的时候，我们知道 Σ_t 与 Σ_0 的差别不大，从而，

$$|\Sigma_t| \leqslant C_1.$$

$$S(\alpha) \leqslant= \int_0^\alpha C_1 dt = C_1 \alpha.$$

命题得证。 □

Karamata 的 Tauber 型渐近公式

定理 541. *假设 μ 是 $(\mathbb{R}_{>0}, \mathcal{B}$ 上的一个（正）测度，其中 $\mathcal{B}$ 为 Borel 代数。我们假设对任意的 $t > 0$，它的 Laplace 变换*

$$(\mathcal{L}\mu)(t) = \int_0^\infty e^{-\lambda t} d\mu(\lambda)$$

是良好定义的。

假设对于 $t \to 0^+$，存在正常数 C_0 和 α，使得

$$(\mathcal{L}\mu)(t) \sim C_0 t^{-\alpha}, \quad t \to 0^+.$$

即

$$\lim_{t\to 0^+} t^\alpha \int_0^\infty e^{-t\lambda} d\mu(\lambda) = C_0,$$

那么，对任意的 $f \in C^0([0,1])$，我们都有

$$\lim_{t\to 0^+} t^\alpha \int_0^\infty f\left(e^{-t\lambda}\right) e^{-t\lambda} d\mu(\lambda) = \frac{C_0}{\Gamma(\alpha)} \int_0^\infty f\left(e^{-t}\right) t^{\alpha-1} e^{-t} dt,$$

其中，

$$\Gamma(\alpha) = \int_0^\infty x^{\alpha-1} e^{-x} dx.$$

证明： 我们首先说明如下简单的观察：所要证明的不等式如果对一致收敛的函数列 $\{f_k(x)\}_{k\geqslant 1} \subset C^0([0,1])$ 成立，那么，这个不等式对于这个函数列的极限函数 $f(x)$ 也成立。实际上，作为 $\mathbb{R}_{>0}$ 上的函数列，我们也有一致收敛性：

$$f_k(e^{-t\lambda}) \xrightarrow{\text{一致}} f(e^{-t\lambda}).$$

所以，

$$\begin{aligned}&\left|\frac{C_0}{\Gamma(\alpha)}\int_0^\infty f_k\left(e^{-t}\right)t^{\alpha-1}e^{-t}dt-\frac{C_0}{\Gamma(\alpha)}\int_0^\infty f\left(e^{-t}\right)t^{\alpha-1}e^{-t}dt\right|\\ \leqslant&\frac{C_0}{\Gamma(\alpha)}\int_0^\infty \|f_k-f\|_{L^\infty}t^{\alpha-1}e^{-t}dt\\ =&C_0\|f_k-f\|_{L^\infty}\to 0.\end{aligned}$$

这说明要证明的不等式的右边是收敛的。为了处理左边，我们先研究

$$\begin{aligned}&\left|t^\alpha\int_0^\infty f_k\left(e^{-t\lambda}\right)e^{-t\lambda}d\mu(\lambda)-t^\alpha\int_0^\infty f\left(e^{-t\lambda}\right)e^{-t\lambda}d\mu(\lambda)\right|\\ \leqslant&t^\alpha\int_0^\infty \|f_k-f\|_{L^\infty}e^{-t\lambda}d\mu(\lambda)\\ =&\|f_k-f\|_{L^\infty}\times t^\alpha\mathcal{L}\mu(t).\end{aligned}$$

根据 $(\mathcal{L}\mu)(t)\sim C_0t^{-\alpha}$，我们知道当 $k\to\infty$ 时，上面的式子的极限为零，这就说明

$$\lim_{t\to0^+}t^\alpha\int_0^\infty f_k\left(e^{-t\lambda}\right)e^{-t\lambda}d\mu(\lambda)-\lim_{t\to0^+}t^\alpha\int_0^\infty f\left(e^{-t\lambda}\right)e^{-t\lambda}d\mu(\lambda)=o_k(1).$$

这就证明了上述的观察。

根据要证明的不等式的线性以及 Weiestrass-Stone 逼近定理，我们只需要对 $f(x)=x^k$ 来证明命题即可：此时，要证明的等式等价于

$$\lim_{t\to0^+}t^\alpha\int_0^\infty e^{-tk\lambda}e^{-t\lambda}d\mu(\lambda)=\frac{C_0}{\Gamma(\alpha)}\int_0^\infty e^{-kt}t^{\alpha-1}e^{-t}dt,$$

左边可以写成

$$t^\alpha\int_0^\infty e^{-((k+1)t)\lambda}d\mu(\lambda)\to\frac{C_0}{(k+1)^\alpha}.$$

右边可以如下计算

$$\int_0^\infty e^{-kt}t^{\alpha-1}e^{-t}dt=\frac{1}{(k+1)^\alpha}\int_0^\infty e^{-(k+1)t}((k+1)t)^{\alpha-1}(k+1)dt=\frac{\Gamma(\alpha)}{(k+1)^\alpha}.$$

所以，命题成立。 □

为了应用这个定理，我们假设令

$$\mu=\sum_{k=1}^\infty\delta_{\lambda_k}(\lambda).$$

那么，根据

$$\sum_{k=1}^\infty e^{-\lambda_kt}=(4\pi t)^{-\frac{n}{2}}\left(|\Omega|+O(t^{\frac{1}{2}})\right),\quad t\to0^+,$$

我们知道

$$(\mathcal{L}\mu)(t)\sim C_0t^{-\alpha},\quad t\to0^+,$$

其中

$$\alpha = \frac{n}{2}, \quad C_0 = \frac{|\Omega|}{(4\pi)^{\frac{n}{2}}}.$$

我们现在选取如下的 f：

$$f(x) = x^{-1}\mathbf{1}_{[e^{-1},1]}(x).$$

我们注意到，f 并不是连续函数，我们后面会用逼近的方式证明对于这个特殊的 f，定理的结论仍然成立。此时，定理结论的左边可以写成：

$$\begin{aligned}\lim_{t\to 0^+} t^\alpha \int_1^{t^{-1}} e^{t\lambda} \times e^{-t\lambda} d\mu(\lambda) &= \lim_{t\to 0^+} t^\alpha \left|\left\{\lambda_k \middle| \lambda_k \leqslant t^{-1}\right\}\right| \\ &= \lim_{\lambda\to\infty} \lambda^{-\alpha} \left|\left\{\lambda_k \middle| \lambda_k \leqslant \lambda\right\}\right|.\end{aligned}$$

右边可以写成

$$\begin{aligned}\frac{C_0}{\Gamma(\alpha)} \int_0^\infty f\left(e^{-t}\right) t^{\alpha-1} e^{-t} dt &= \frac{C_0}{\Gamma(\alpha)} \int_1^\infty e^t \times t^{\alpha-1} e^{-t} dt \\ &= \frac{C_0}{\alpha\Gamma(\alpha)} = \frac{C_0}{\Gamma(\alpha+1)}.\end{aligned}$$

综上所述，我们就有

$$\lim_{\lambda\to\infty} \lambda^{-\frac{n}{2}} \left|\left\{\lambda_k \middle| \lambda_k \leqslant \lambda\right\}\right| = \frac{|\Omega|}{(4\pi)^{\frac{n}{2}}\Gamma\left(\frac{n}{2}+1\right)}.$$

这就给出了 Weyl 关于特征值的渐近公式。

最后，我们用连续函数来逼近 f。我们构造两族函数：

$$f_\varepsilon^+(x) = \begin{cases} 0, & x \leqslant e^{-(1+\varepsilon)}; \\ \text{线性地连接}; & \\ \frac{1}{x}, & x \geqslant e^{-1}. \end{cases} \qquad f_\varepsilon^-(x) = \begin{cases} 0, & x \leqslant e^{-1}; \\ \text{线性地连接}; & \\ \frac{1}{x}, & x \geqslant e^{-(1-\varepsilon)}. \end{cases}$$

我们知道，

$$f^-(x) \leqslant f(x) \leqslant f^+(x).$$

所以，

$$\lim_{t\to 0^+} t^\alpha \int_0^\infty f_\varepsilon^-\left(e^{-t\lambda}\right) e^{-t\lambda} d\mu(\lambda) \leqslant \liminf_{t\to 0^+} \left(t^\alpha \int_0^\infty f\left(e^{-t\lambda}\right) e^{-t\lambda} d\mu(\lambda)\right).$$

从而，

$$\frac{C_0}{\Gamma(\alpha)} \int_0^\infty f_\varepsilon^-\left(e^{-t}\right) t^{\alpha-1} e^{-t} dt \leqslant \liminf_{t\to 0^+} \left(t^\alpha \int_0^\infty f\left(e^{-t\lambda}\right) e^{-t\lambda} d\mu(\lambda)\right).$$

类似地，我们有

$$\limsup_{t\to 0^+} \left(t^\alpha \int_0^\infty f\left(e^{-t\lambda}\right) e^{-t\lambda} d\mu(\lambda)\right) \leqslant \frac{C_0}{\Gamma(\alpha)} \int_0^\infty f_\varepsilon^+\left(e^{-t}\right) t^{\alpha-1} e^{-t} dt.$$

所以，我们只要证明

$$\lim_{\varepsilon\to 0}\left(\int_0^\infty f_\varepsilon^{\pm}\left(e^{-t}\right)t^{\alpha-1}e^{-t}dt\right)=\int_0^\infty f\left(e^{-t}\right)t^{\alpha-1}e^{-t}dt$$

即可。我们有

$$\begin{aligned}\int_0^\infty (f_\varepsilon^{\pm}-f)\left(e^{-t}\right)t^{\alpha-1}e^{-t}dt&=\int_0^\infty (f_\varepsilon^{\pm}-f)\left(e^{-t}\right)t^{\alpha-1}e^{-t}dt\\&=\int_{1-\varepsilon}^{1+\varepsilon}(f_\varepsilon^{\pm}-f)\left(e^{-t}\right)t^{\alpha-1}e^{-t}dt\\&\leqslant\int_{1-\varepsilon}^{1+\varepsilon}2e\left(e^{-t}\right)t^{\alpha-1}e^{-t}dt\to 0.\end{aligned}$$

这就完成了全部的证明。

微局部分析一瞥：微分算子的奇性传播定理

光滑性与波前集

我们知道一个分布的光滑性是一个局部的性质。给定一个分布 $u\in\mathcal{D}'(\Omega)$，u 在 $x_0\in\Omega$ 处附近是光滑当且仅当对任意的支集在 x_0 附近的光滑函数 f，我们有 $f\cdot u\in C_0^\infty(\mathbb{R}^n)$。进一步，$u$ 在 $x_0\in\Omega$ 处附近是光滑当且仅当对某一个任意的支集在 x_0 附近的光滑函数 f，$f(x_0)\neq 0$，我们有 $f\cdot u\in C_0^\infty(\mathbb{R}^n)$。从 Fourier 分析的角度来看，函数的光滑性等价于它在频率空间上的衰减。实际上，根据 Sobolev 嵌入定理（或者直接证明），在我们刚刚谈论的场合下，如下两个论述显然是等价的：

1) $f\cdot u\in C_0^\infty(\mathbb{R}^n)$；

2) 对任意的正整数 $N\geqslant 1$，存在常数 C_N，使得对任意的 $\xi\in\mathbb{R}^n$，我们有

$$\left|\widehat{fu}(\xi)\right|\leqslant\frac{C_N}{(1+|\xi|^2)^{\frac{N}{2}}}.$$

我们现在考察一个特殊的例子：$u=\mathbf{1}_{x_n\geqslant 0}$。这是半空间 $\mathbb{H}^n$ 上的示性函数。它在 $x_n>0$ 或者 $x_n<0$ 上显然是光滑的。在 $x_n=0$ 这个集合上，尽管 u 不光滑，我们发现沿着 $x_1,\cdots,x_{n_1}$ 的方向求导数总是可以的，唯一不光滑（连续）的方向实际上是变量 x_n 造成的。我们下面要引入一个概念，它不仅能说明函数在一个点处不光滑，而且能说明这个函数沿着哪些方向不光滑。

我们通常用所谓的锥型集来标记方向的集合：

定义 542. $\Gamma\subset\mathbb{R}^n_\xi$ 是频率空间中的一个子集，如果下面的两个性质成立：

1) 对任意的 $\lambda>0$，对任意的 $\xi\in\Gamma$，我们有 $\lambda\cdot\xi\in\Gamma$；

2) $\Gamma\cap\mathbf{S}^{n-1}_\xi$ 是 $\mathbf{S}^{n-1}_\xi$ 上的开集（闭集）。这里，$\mathbf{S}^{n-1}_\xi$ 是频率空间 $\mathbb{R}^n_\xi$ 中的单位球面。

那么，我们就称 Γ 是一个**锥型开集（闭集）**。

另外，为了方便起见，我们引入如下的记号：

$$(T^*\Omega)^{\times} = \big\{(x,\xi) \in \Omega \times \mathbb{R}^n \big| \xi \neq 0\big\}.$$

我们现在引入波前集的概念，目的是说明分布沿着哪些方向不光滑：

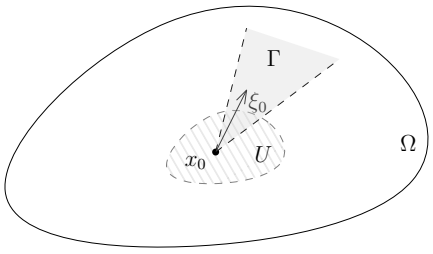
定义 543. $\Omega \subset \mathbb{R}^n$ 是开集，$u \in \mathcal{D}'(\Omega)$ 是 Ω 上的分布。给定 $(x_0,\xi_0) \in (T^*\Omega)^{\times}$，假设开集 $U \subset \Omega$，锥型开集 $\Gamma \subset \mathbb{R}^n_\xi$ 以及 $f \in C_0^\infty(\Omega)$，使得

1) $x_0 \in U$，$\mathrm{supp}(f) \subset U$ 并且 $f(x_0) \neq 0$；

2) $\xi_0 \in \Gamma$；

3) 对任意的正整数 $N \geqslant 1$，存在常数 C_N，使得对任意的 $\xi \in \Gamma$，我们有

$$\left|\widehat{fu}(\xi)\right| \leqslant \frac{C_N}{(1+|\xi|^2)^{\frac{N}{2}}}.$$



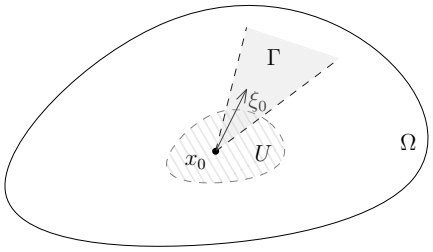
我们就称 (x_0,ξ_0)不落在 u 的**波前集**里。我们用 $WF(u)$ 表示不满足上述要求的 $(x_0,\xi_0) \in (T^*\Omega)^{\times}$ 的点的集合。按照定义，我们知道

$$WF(u) \subset (T^*\Omega)^{\times}.$$

83 波前集以及其等价定义，微局部光滑性意味着局部光滑性，一个非驻相法的引理（理论的技术核心），波前集在微分同胚下的变换

二零二零年十二月十六日，星期三，晴

假设 $\Omega \subset \mathbb{R}^n$ 是开集，$u \in \mathcal{D}'(\Omega)$ 是 Ω 上的分布。



上次课我们引入了 u 的波前集 $WF(u) \subset (T^*\Omega)^\times$：按照定义，对于 $(x_0,\xi_0) \in (T^*\Omega)^\times$，$(x_0,\xi_0) \notin WF(u)$ 当且仅当存在开集 $U \subset \Omega$，锥型开集 $\Gamma \subset \mathbb{R}^n_\xi$ 以及 $f \in C_0^\infty(\Omega)$，使得

1) $x_0 \in U$，$\mathrm{supp}(f) \subset U$ 并且 $f(x_0) \neq 0$；

2) $\xi_0 \in \Gamma$；

3) 对任意的正整数 $N \geqslant 1$，存在常数 C_N，使得对任意的 $\xi \in \Gamma$，我们有

$$\left|\widehat{fu}(\xi)\right| \leqslant \frac{C_N}{(1+|\xi|^2)^{\frac{N}{2}}}.$$

波前集的等价定义

引理 544. 给定锥型开集 $\Gamma \subset \mathbb{R}^n_\xi$，$s_0 > 0$ 和 $u \in H^{-s_0}(\mathbb{R}^n)$。如果对任意的锥型闭集 $\Gamma' \subset \Gamma$，对任意的正整数 $N \geqslant 1$，存在常数 C_N，使得对任意的 $\xi' \in \Gamma'$，我们有

$$\left|\widehat{u}(\xi')\right| \leqslant \frac{C'_N}{(1+|\xi'|^2)^{\frac{N}{2}}}.$$

那么，对任意的 $f \in C_0^\infty(\Omega)$（$\approx$ 物理空间的任意截断函数），对任意的锥型闭集 $\Gamma'' \subset \Gamma$，对任意的正整数 $N \geqslant 1$，存在常数 C''_N，使得对任意的 $\xi'' \in \Gamma''$，我们有

$$\left|\widehat{fu}(\xi'')\right| \leqslant \frac{C''_N}{(1+|\xi''|^2)^{\frac{N}{2}}}.$$

证明: 选定 $f \in C_0^\infty(\Omega)$ 和锥型闭集 $\Gamma'' \subset \Gamma$，根据 Γ'' 在 $\mathbf{S}^{n-1}$ 上的闭性，存在锥型闭集 $\Gamma' \subset \Gamma$，使得 $\Gamma'' \subset \mathring{\Gamma}'$（落在 Γ' 的内部）。我们计算 $\widehat{fu}(\xi'')$：

$$\begin{aligned}\widehat{fu}(\xi'') &= \frac{1}{(2\pi)^n}\int_{\mathbb{R}^n} \widehat{f}(\xi''-\eta)\widehat{u}(\eta)d\eta \\ &= \underbrace{\frac{1}{(2\pi)^n}\int_{\Gamma'} \widehat{f}(\xi''-\eta)\widehat{u}(\eta)d\eta}_{I} + \underbrace{\frac{1}{(2\pi)^n}\int_{\mathbb{R}^n-\Gamma'} \widehat{f}(\xi''-\eta)\widehat{u}(\eta)d\eta}_{II}.\end{aligned}$$

对于 I，我们用引理的条件，对任意的正整数 $N \geqslant 1$，存在常数 $C_N(\Gamma')$，使得对任意的 $\xi' \in \Gamma'$，我们有

$$\left|\widehat{u}(\xi')\right| \leqslant \frac{C_N(\Gamma')}{(1+|\xi'|^2)^{\frac{N}{2}}}.$$

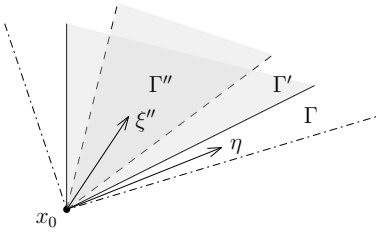
所以，

$$\begin{aligned}
(1+|\xi''|)^N I &\leqslant \frac{1}{(2\pi)^n}\int_{\Gamma'}(1+|\xi''-\eta|+|\eta|)^N \left|\widehat{f}(\xi''-\eta)\right| |\widehat{u}(\eta)|\, d\eta \\
&\leqslant \frac{1}{(2\pi)^n}\int_{\Gamma'}(1+|\xi''-\eta|)^N \left|\widehat{f}(\xi''-\eta)\right| (1+|\eta|)^N |\widehat{u}(\eta)|\, d\eta \\
&\leqslant \frac{1}{(2\pi)^n}\int_{\Gamma'}(1+|\xi''-\eta|)^N \left|\widehat{f}(\xi''-\eta)\right| C_N(\Gamma') d\eta \\
&= \frac{C_N(\Gamma')}{(2\pi)^n}\int_{\Gamma'}(1+|\eta|)^N \left|\widehat{f}(\eta)\right| d\eta.
\end{aligned}$$

由于 $f \in C_0^\infty$，所以上面的积分是有限的，从而，

$$(1+|\xi''|)^N \leqslant C_N.$$

在 II 这一项中，η 落在 Γ' 之外。



我们注意到，此时，存在 $\delta > 0$，使得

$$|\xi''| \leqslant \frac{1}{\delta}|\xi''-\eta|.$$

这是因为从 ξ'' 到 η 的距离不会小于 ξ'' 到 $\partial\Gamma'$ 的距离，而后面这个距离与 $|\xi''|$ 是成正比的（投影到 $\partial\Gamma'$ 上），这表明

$$|\xi''-\eta| \geqslant |\xi''|\delta.$$

类似地，我们还有

$$|\xi''-\eta| \geqslant |\eta|\delta.$$

据此，我们有

$$\begin{aligned}
(1+|\xi''|)^N II &\leqslant \frac{1}{(2\pi)^n}\int_{\mathbb{R}^n-\Gamma'}(1+|\xi''|)^N |\widehat{f}(\xi''-\eta)||\widehat{u}(\eta)| d\eta \\
&\leqslant C_\delta \int_{\mathbb{R}^n-\Gamma'}(1+|\xi''-\eta|)^N |\widehat{f}(\xi''-\eta)||\widehat{u}(\eta)| d\eta \\
&\leqslant C'_\delta \int_{\mathbb{R}^n-\Gamma'}(1+|\xi''-\eta|)^{-\frac{s_0}{2}-n} |\widehat{u}(\eta)| d\eta.
\end{aligned}$$

最后一步是因为

$$(1+|\xi''-\eta|)^N|\widehat{f}(\xi''-\eta)|\leqslant C''(1+|\xi''-\eta|)^{-\frac{s_0}{2}-n}.$$

从而，

$$\begin{aligned}(1+|\xi''|)^N II &\leqslant C'_\delta\int_{\mathbb{R}^n-\Gamma'}(1+|\eta|)^{-\frac{s_0}{2}}|\widehat{u}(\eta)|\times(1+|\eta|)^{-n}d\eta\\ &\leqslant C'_\delta\left(\int_{\mathbb{R}^n-\Gamma'}(1+|\eta|)^{-s_0}|\widehat{u}(\eta)|^2d\eta\right)^{\frac{1}{2}}\times\left(\int_{\mathbb{R}^n-\Gamma'}(1+|\eta|)^{-2n}d\eta\right)^{\frac{1}{2}}\\ &= C'_\delta\|u\|_{H^{-s_0}}.\end{aligned}$$

综合上面两个不等式，我们就得到了

$$I+II\leqslant\frac{C_N}{(1+|\xi|)^N}.$$

命题得到证明。 □

根据这个引理，我们有如下重要的推论：

推论 545. *给定 $\Omega\subset\mathbb{R}^n$，$u\in\mathcal{D}'(\Omega)$。对任意的 $\psi(x)\in C^\infty(\Omega)$，我们有*

$$WF(\psi\cdot u)\subset WF(u).$$

证明： 我们只要证明如果 $(x_0,\xi_0)\notin WF(u)$，那么 $(x_0,\xi_0)\notin WF(\psi\cdot u)$ 即可。根据定义，存在锥型开集 $\xi_0\in\Gamma\subset\mathbb{R}^n_\xi$，$f\in C_0^\infty(\Omega)$，$f(x_0)\neq0$ 并且对任意的正整数 $N\geqslant1$，存在常数 C_N，使得对任意的 $\xi\in\Gamma$，有

$$\left|\widehat{fu}(\xi)\right|\leqslant\frac{C_N}{(1+|\xi|)^N}.$$

我们对 $(u,f)=(fu,\psi)$ 用上面的引理：存在一个锥型开集 $\Gamma'\text{‘}\subset\Gamma$ 包含 ξ_0（这里用到了引理中 Γ'' 的任意性）并且 $\mathring{\Gamma''}\neq\emptyset$，使得对任意的正整数 $N\geqslant1$，存在常数 C''_N，使得对任意的 $\xi''\in\Gamma''$，我们有

$$\left|\widehat{f(\psi\cdot u)}(\xi'')\right|=\left|\widehat{\psi\cdot fu}(\xi'')\right|\leqslant\frac{C''_N}{(1+|\xi''|^2)^{\frac{N}{2}}}.$$

这表明 $(x_0,\xi_0)\notin WF(\psi\cdot u)$ □

作为这个引理的另一个推论，我们给出波前集的等价定义：

定理 546. *$\Omega\subset\mathbb{R}^n$ 是开集，$u\in\mathcal{D}'(\Omega)$ 是 Ω 上的分布。给定 $(x_0,\xi_0)\in(T^*\Omega)^\times$。那么，如下的论断是等价的：*

1) $(x_0,\xi_0)\notin WF(u)$；

2) 存在包含 x_0 的开集 $U\subset\Omega$，包含 ξ_0 的锥型开集 $\Gamma\subset\mathbb{R}^n_\xi$，使得对任意的 $f\in C_0^\infty(U)$，对任意的正整数 $N\geqslant1$，存在常数 $C_{N,f}$，使得对任意的 $\xi\in\Gamma$，我们有

$$\left|\widehat{fu}(\xi)\right|\leqslant\frac{C_{N,f}}{(1+|\xi|)^N}.$$

证明: 2)⇒1) 是明显的: 我们只要任意选取一个 $f\in C_0^\infty(U)$ 并且 $f(x_0)\neq 0$ 即可。反之, 我们假设 1) 成立来证明 2): 根据定义, $(x_0,\xi_0)\notin WF(u)$ 意味着存在锥型开集 $\xi_0\in\Gamma\subset\mathbb{R}^n_\xi$, $F\in C_0^\infty(\Omega)$, $F(x_0)\neq 0$ 并且对任意的正整数 $N\geqslant 1$, 存在常数 C_N, 使得对任意的 $\xi\in\Gamma$, 有

$$\left|\widehat{Fu}(\xi)\right|\leqslant\frac{C_N}{(1+|\xi|)^N}.$$

由于 $F(x_0)\neq 0$, 所以存在开集 U, $x_0\in U$ 并且对任意的 $x\in U$, $|F(x)|>\delta>0$。我们选定这个开集 U。对任意的 $f\in C_0^\infty(U)$, 我们考虑:

$$fu=\frac{f}{F}Fu.$$

我们对 $(u,f)=\left(Fu,\dfrac{f}{F}\right)$ 来运用引理: 存在一个锥型开集 $\Gamma'\subset\Gamma$ 包含 ξ_0 (这里用到了引理中 Γ'' 的任意性) 并且 $\mathring{\Gamma}''\neq\emptyset$, 使得对任意的正整数 $N\geqslant 1$, 存在常数 C''_N, 使得对任意的 $\xi''\in\Gamma''$, 我们有

$$\left|\widehat{fu}(\xi'')\right|=\left|\widehat{\frac{f}{F}Fu}(\xi'')\right|\leqslant\frac{C''_N}{(1+|\xi''|)^N}.$$

此时, 在 2) 中我们选取的锥型开集是 Γ''。命题得证。 □

推论 547 (从微局部到局部: 光滑性)**.** *给定 $\Omega\subset\mathbb{R}^n$, $u\in\mathcal{D}'(\Omega)$。假设 $x_0\in\Omega$ 使得对任意的 $\xi\in\mathbb{R}^n_\xi$, $(x_0,\xi)\notin WF(u)$ (即 u 在 (x_0,ξ) 处是在微局部意义下光滑的)。那么, 存在 x_0 在 Ω 中的开邻域 U, 使得 $u\big|_U$ 是光滑的。*

证明: 对任意的 $\xi\in\mathbf{S}^{n-1}$ (单位长的向量), $(x_0,\xi)\notin WF(u)$, 从而, 存在锥型开集 $\Gamma\subset\mathbb{R}^n_\xi$ 和开集 U, 其中 $x_0\in U$, $\xi\in\Gamma$, 对任意的 $f\in C_0^\infty(U)$, 对任意的正整数 $N\geqslant 1$, 存在常数 $C_{N,f}$, 使得对任意的 $\xi\in\Gamma$, 有

$$\left|\widehat{fu}(\xi)\right|\leqslant\frac{C_{N,f}}{(1+|\xi|)^N}.$$

当 ξ 取遍整个 $\mathbf{S}^{n-1}$ 时, 它们所对应的上述的 Γ 是 $\mathbf{S}^{n-1}$ 的一个开覆盖, 所以, 我们可以选出有限个 $\xi_1,\cdots,\xi_m\in\mathbf{S}^{n-1}$, 使得对每个 $j=1,\cdots,m$, 存在锥型开集 $\Gamma_j\subset\mathbb{R}^n_\xi$ 和开集 U_j, 其中 $x_0\in U$, $\xi_j\in\Gamma$, 对任意的 $f\in C_0^\infty(U_j)$, 对任意的正整数 $N\geqslant 1$, 存在常数 $C_{j,N,f}$, 使得对任意的 $\xi\in\Gamma$, 有

$$\left|\widehat{fu}(\xi)\right|\leqslant\frac{C_{j,N,f}}{(1+|\xi|)^N}$$

并且

$$\bigcup_{j=1}^m\Gamma_j\supset\mathbf{S}^{n-1}\quad\Rightarrow\quad\bigcup_{j=1}^m\Gamma_j\supset\mathbb{R}^n_\xi-\{0\}.$$

此时, 我们令 $U=\bigcap\limits_{j\leqslant m}U_j$, 这也是包含 x_0 的开集。我们任意选取 $f\in C_0^\infty(U)\subset C_0^\infty(U_j)$, 其中 $f(x_0)\neq 0$。那么, 对任意的 $\xi\in\mathbb{R}^n$, 存在 j, 使得 $\xi\in\Gamma_j$, 此时, 如果令 $C_N=\max\limits_{j\leqslant m}C_{j,N,f}$, 那

么，对任意的 $N \geqslant 1$，我们就有

$$\left|\widehat{fu}(\xi)\right| \leqslant \frac{C_{j,N,f}}{(1+|\xi|)^N} \leqslant \frac{C_N}{(1+|\xi|)^N}.$$

这表明 fu 是光滑函数（因为它的 Fourier 变换是衰减得比任意的多项式都要快）。 □

我们给出这个推论的一个应用。我们先回忆微分算子的定义：给定区域 Ω 上的一个 m-次微分算子（变系数），即

$$P = \sum_{|\alpha| \leqslant m} p_\alpha(x)\partial^\alpha,$$

其中，对每个 α，$p_\alpha \in C^\infty(\Omega)$ 并且至少有一个 α 使得 $|\alpha| = m$ 并且 $p_\alpha \neq 0$（不恒为 0）。我们定义 P 的**主象征**为

$$p_m(x,\xi) = \sum_{|\alpha|=m} p_\alpha(x)(-i\xi)^\alpha.$$

那么，$p_m(x,\xi)$ 是 $T^*\Omega = \Omega \times \mathbb{R}^n_\xi$ 上的光滑函数。我们把 p_m 的零点集记作

$$Z(p_m) = \big\{(x,\xi) \in T^*\Omega \big| p_m(x,\xi) = 0\big\}.$$

我们先不加证明的接受如下的定理（下节课给出完整证明）：

定理 548 (微局部椭圆正则性)**.** *给定区域 Ω 上的一个 m-次微分算子*

$$P = \sum_{|\alpha| \leqslant m} p_\alpha(x)\partial^\alpha$$

假设 $u \in \mathcal{D}'(\Omega)$ 满足微分方程

$$Pu \overset{\mathcal{D}'}{=} f,$$

其中 $f \in \mathcal{D}'(\Omega)$。那么，我们有

$$WF(u) \subset Z(p_m) \cup WF(f).$$

我们现在证明 $-\triangle$ 算子的正则性：假设分布 $u \in \mathcal{D}'(\Omega)$，满足 $-\triangle u \in C^\infty(\Omega)$，我们现在来证明 $u \in C^\infty(\Omega)$。实际上，

$$p_2(-\triangle) = |\xi|^2 \quad \Rightarrow \quad Z(p_2) = \{0\}.$$

由于 $-\triangle u$ 光滑，所以 $WF(-\triangle u) = \emptyset$。利用上述微局部正则性定理，我们就有

$$WF(u) \subset \{0\} \cup \emptyset = \{0\}.$$

然而，$WF(u) \subset (T^*\Omega)^\times$，所以，$WF(u) = \emptyset$。我们刚刚证明的推论表明 u 是光滑的。

波前集在微分同胚下的变换

我们下面研究波前集在微分同胚下的变换法则。假设

$$\Phi : \Omega \to \Omega'$$

是 $\mathbb{R}^n$ 中两个开区域之间的微分同胚，其中，我们用 $x = (x_1, \cdots, x_n)$ 和 $y = (y_1, \cdots, y_n)$ 分别表示 Ω 和 Ω' 上的坐标。假设 $u(y)$ 是 Ω' 上的分布，我们不难验证，u 在 y_0 的附近光滑当且仅当 $u \circ \Phi$ 在 x_0 处光滑，其中 $y_0 = \Phi(x_0)$。

给定了上述的微分同胚，我们定义 $T^*\Omega$ 与 $T^*\Omega'$ 之间的微分同胚：

$$(d\Phi)^* : T^*\Omega \to T^*\Omega', \ \ (x_0, \xi_0) \mapsto \left(\Phi(x_0), {}^t\left(d\Phi\big|_{x=x_0} \right)^{-1} (\xi_0) \right).$$

其中，${}^t\left(d\Phi\big|_{x=x_0} \right)^{-1}$ 表示的是 Jacobi 矩阵 $d\Phi\big|_{x=x_0}$ 转置取逆。

练习. 证明，$(d\Phi)^*$ 是微分同胚。

定理 549. 给定开区域 $\Omega \subset \mathbb{R}^n$ 和 $\widetilde{\Omega} \subset \mathbb{R}^n$，用 $x = (x_1, \cdots, x_n)$ 和 $y = (y_1, \cdots, y_n)$ 分别表示 Ω 和 $\widetilde{\Omega}$ 上的坐标，我们假设

$$\Phi : \Omega \to \widetilde{\Omega}$$

是微分同胚。对任意的 $\widetilde{u} \in \mathcal{D}'(\widetilde{\Omega})$，任意的 $x_0 \in \Omega$，令 $y_0 = \Phi(x_0) \in \widetilde{\Omega}$，那么，$(y_0, \widetilde{\xi}) \in WF(\widetilde{u})$ 当且仅当 $\left(x_0, {}^t\left(d\Phi\big|_{x=x_0} \right) \left(\widetilde{\xi} \right) \right) \in WF(\widetilde{u} \circ \Phi)$。换而言之，如果令 $u = \Phi^*(\widetilde{u})$，那么，

$$WF(\widetilde{u}) = (d\Phi)^* \left(WF(u) \right).$$

注记. 这个变换与流形上余切丛的转移函数的形式是一致的，这建议说频率空间应该被视作是余切丛的纤维。

一个非驻相法的引理

给定如下的基本数据：

(a)（相函数）假设 $\phi(x, \xi) \in C^\infty \left(\mathbb{R}^n_x \times (\mathbb{R}^n_\xi - \{0\}) \right)$ 是实值函数，使得

- $\phi(x, \xi)$ 对变量 ξ 是 1 次齐次函数；
- 对任意的 $(x, \xi) \in \mathbb{R}^n_x \times (\mathbb{R}^n_\xi - \{0\})$，$\xi \neq 0$，我们有 $\nabla_x \phi(x, \xi) \neq 0$（作为 n-维的向量）。

(b)（振幅函数）假设 $a(x, \xi) \in C^\infty \left(\mathbb{R}^n_x \times (\mathbb{R}^n_\xi - \{0\}) \right)$，使得存在 $N_0 \in \mathbb{R}$，使得对任意的多重指标 α，存在常数 C_α，使得对任意的 $(x, \xi) \in \mathbb{R}^n_x \times (\mathbb{R}^n_\xi - \{0\})$，我们都有

$$|\partial_x^\alpha a(x, \xi)| \leqslant C_\alpha \left(1 + |\xi| \right)^{N_0}.$$

(c)（被作用函数）假设 $K \subset \mathbb{R}^n$ 是紧集，$u \in \mathcal{E}'(\mathbb{R}^n)$ 使得 $\mathrm{supp}(u) \subset K$。如果存在 $\alpha_0 > 0$ 和 $\eta_0 \in \mathbf{S}^{n-1}$，使得对任意的 $N \geqslant 1$，存在常数 C_N，使得对任意的 $\xi \in \Gamma(\eta_0, \alpha_0)$，都有我们就有

$$|\widehat{u}(\xi)| \leqslant \frac{C_N}{(1+|\xi|)^N}.$$

其中，锥型开集 $\Gamma(\eta_0, \alpha_0)$ 的定义如下

$$\Gamma(\eta_0, \alpha_0) = \big\{\xi \in \mathbb{R}^n - \{0\} \big| \left| \frac{\xi}{|\xi|} - \frac{\eta_0}{|\eta_0|} \right| < \alpha_0 \big\}.$$

我们想要定义积分

$$I(\xi) = \Big\langle u(x), e^{-i\phi(x,\xi)} a(x,\xi)) \Big\rangle = \int_{\mathbb{R}^n} e^{-i\phi(x,\xi)} a(x,\xi) u(x) dx.$$

上面的第一个等号自然是良好定义的，因为 u 具有紧支集，所以这就是如下的自然配对：

$$\mathcal{E}'(\mathbb{R}^n_x) \times C^\infty(\mathbb{R}^n_x) \to \mathbb{C}.$$

由于 $u(x)$ 并没有关于 x 的控制，尽管所以上面的积分至少在形式上并不是良好定义的。我们下面用积分的形式来逼近上述的 $I(\xi)$。任意选取 $\chi(x) \in C_0^\infty(\mathbb{R}^n)$，其中 $\chi \geqslant 0$ 并且在原点附近恒为 1。那么，在频率空间，我们知道在缓增分布的意义下，有

$$\chi_\varepsilon(\eta)\widehat{u} \xrightarrow{\mathcal{S}'} \widehat{u}, \quad \varepsilon \to 0^+.$$

我们令

$$u_\varepsilon = \mathcal{F}^{-1}(\chi_\varepsilon \cdot \widehat{u}),$$

从而，

$$u_\varepsilon \xrightarrow{\mathcal{D}'} u, \quad \varepsilon \to 0^+.$$

特别地，我们就有

$$I_\varepsilon(\xi) = \Big\langle u_\varepsilon(x), e^{-i\phi(x,\xi)} a(x,\xi)) \Big\rangle \to I(\xi).$$

由于 $\widehat{u_\varepsilon}$ 具有紧支集（并且是光滑函数），所以，$u_\varepsilon(x)$ 对于 x 比任意的多项式衰减的都要快，从而，

$$\int_{\mathbb{R}^n} e^{-i\phi(x,\xi)} a(x,\xi) u_\varepsilon(x) dx$$

是良好定义的（被积分函数是 L^1 的），所以，

$$I_\varepsilon(\xi) = \int_{\mathbb{R}^n} e^{-i\phi(x,\xi)} a(x,\xi) u_\varepsilon(x) dx \to I(\xi), \quad \varepsilon \to 0^+.$$

由于一开始 u 具有紧支集，我们再利用一个有紧支集的函数来记住这个事实：选取 $\psi(x) \in C_0^\infty(\mathbb{R}^n)$，使得 $\psi(x)\big|_K \equiv 1$，我们就给出了 $I(\xi)$ 的积分表示：

$$\begin{aligned} I(\xi) &= \Big\langle u(x), e^{-i\phi(x,\xi)} \psi(x) a(x,\xi)) \Big\rangle \\ &= \lim_{\varepsilon \to 0^+} \int_{\mathbb{R}^n} e^{-i\phi(x,\xi)} \psi(x) a(x,\xi) u_\varepsilon(x) dx \\ &= \lim_{\varepsilon \to 0^+} \frac{1}{(2\pi)^n} \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} e^{-i\phi(x,\xi) + ix\cdot\eta} \psi(x) a(x,\xi) \chi_\varepsilon(\eta) \widehat{u}(\eta) dx d\eta. \end{aligned}$$

上面我们用到了 Fubini 定理，这是因为

$$\left|e^{-i\phi(x,\xi)+ix\cdot\eta}\psi(x)a(x,\xi)\chi_\varepsilon(\eta)\widehat{u}(\eta)\right| = |\psi(x)a(x,\xi)\chi_\varepsilon(\eta)\widehat{u}(\eta)|$$

是光滑的可积函数。令

$$\widetilde{a}(x,\xi) = \psi(x)a(x,\xi),$$

我们就有

$$I(\xi) = \lim_{\varepsilon\to 0^+}\frac{1}{(2\pi)^n}\int_{\mathbb{R}^n}\int_{\mathbb{R}^n} e^{-i\phi(x,\xi)+ix\cdot\eta}\,\widetilde{a}(x,\xi)\chi_\varepsilon(\eta)\widehat{u}(\eta)dxd\eta.$$

引理 550 (非驻相法)**.** *假设相函数 $\phi(x,\xi)$，振幅函数 $a(x,\xi)$ 和被作用函数 $u(x)$ 满足前面所要求的 (a), (b) 和 (c) 三个条件。*

给定 $\xi_0 \in \mathbb{R}^n_\xi - \{0\}$，$\Gamma_0$ 是包含 ξ_0 的一个锥型开集，$\alpha_1 < \alpha_0$，其中 α_0 在条件 (c) 中出现。假设对任意的 $(x,\xi) \in K \times \Gamma_0$，我们都有

$$\left|\frac{\nabla_x\phi(x,\xi)}{|\nabla_x\phi(x,\xi)|} - \eta_0\right| < \alpha_1.$$

那么，我们可以对任意的 $N \geqslant 1$，存在 $C_N > 0$，使得对任意的 $\xi \in \Gamma_0$，有

$$|I(\xi)| \leqslant \frac{C_N}{(1+|\xi|)^N}.$$

证明: 假设 $\xi \in \Gamma_0$。我们现在考虑如下的集合

$$\Omega_1 = \left\{\eta \in \mathbb{R}^n \middle| |\eta - \nabla_x\phi(x,\xi)| < \varepsilon_1|\nabla_x\phi(x,\xi)| \text{ 或者 } |\eta - \nabla_x\phi(x,\xi)| < \varepsilon_1|\eta|\right\}.$$

这显然是一个有界开集。如果 ε_1 选取的足够小，对于 $\eta \in \Omega_1$，我们知道

$$\left|\frac{\nabla_x\phi(x,\xi)}{|\nabla_x\phi(x,\xi)|} - \frac{\eta}{|\eta|}\right| = O(\varepsilon_1).$$

所以，利用 $\alpha_1 < \alpha_0$，我们知道如果 ε_1 足够小，那么，引理中的条件意味着

$$\left|\frac{\nabla_x\phi(x,\xi)}{|\nabla_x\phi(x,\xi)|} - \eta_0\right| < \alpha_1 \quad\Rightarrow\quad \eta \in \Gamma(\eta_0,\alpha_0).$$

所以，通过选取足够小的 ε_1（由 α_1 和 α_1 决定，这是一个只依赖于 α_0 和 α_1 的绝对参数），我们就有

$$\Omega_1 \subset \Gamma(\eta_0,\alpha_0).$$

另外，根据 Ω_1 的定义，我们知道（$\varepsilon_1 < 0.5$），对任意的 $\eta \in \Omega_1$，我们都有

$$\frac{1}{2}|\eta| \leqslant |\nabla_x\phi(x,\xi)| \leqslant 2|\eta|.$$

从而，利用 ϕ 对频率分量的齐次性质，我们就有

$$\frac{1}{2}|\eta| \leqslant |\nabla_x\phi\left(x,\frac{\xi}{|\xi|}\right)||\xi| \leqslant 2|\eta|.$$

由于光滑函数 $\nabla_x\phi(x,\xi')$ 在紧集 $K\times\mathbf{S}^{n-1}$ 上具有最大最小值，根据 (a) 中第二点的要求，上面这个不等式意味着存在常数 $C_1>0$，使得对任意的 $\eta\in\Omega_1$，我们都有

$$\frac{1}{C_1}|\eta|\leqslant|\xi|\leqslant C_1|\eta|.$$

我们令 $\Omega_2=\mathbb{R}^n_\eta-\Omega_1$，那么，我们将积分对 η 分量进行区域分解

$$\begin{aligned}I_\varepsilon(\xi)&=\frac{1}{(2\pi)^n}\int_{\mathbb{R}^n}\int_{\mathbb{R}^n}e^{-i\phi(x,\xi)+ix\cdot\eta}\,\widetilde{a}(x,\xi)\chi_\varepsilon(\eta)\widehat{u}(\eta)dxd\eta\\&=\frac{1}{(2\pi)^n}\left(\int_{\mathbb{R}^n}\int_{\Omega_1}+\int_{\mathbb{R}^n}\int_{\Omega_2}\right)e^{-i\phi(x,\xi)+ix\cdot\eta}\,\widetilde{a}(x,\xi)\chi_\varepsilon(\eta)\widehat{u}(\eta)dxd\eta\\&=I_{\varepsilon,1}(\xi)+I_{\varepsilon,2}(\xi).\end{aligned}$$

我们分别控制这两项：

- 控制 $I_{\varepsilon,1}(\xi)$。

  此时，积分因子中的 $\eta\in\Omega_1$，根据前面的讨论，我们就有 $\eta\in\Gamma(\eta_0,\alpha_0)$，所以，根据条件 (c)，我们就有

  $$|\widehat{u}(\eta)|\leqslant\frac{C_N}{(1+|\eta|)^N},$$

  其中 N 待定。另外，根据 ${C_1}^{-1}|\eta|\leqslant|\xi|\leqslant C_1|\eta|$，所以，我们就有

  $$|\widehat{u}(\eta)|\leqslant\frac{C_N}{(1+|\eta|+|\xi|)^N}.$$

  上面的常数 C_N 可能有所改变，但是这对我们的推理没有影响。再利用 (b) 中的界来控制 $a(x,\xi)$，我们就得到

  $$\begin{aligned}I_{\varepsilon,1}(\xi)&\leqslant\frac{1}{(2\pi)^n}\int_{\mathbb{R}^n}\int_{\Omega_1}|\psi(x)a(x,\xi)||\chi_\varepsilon(\eta)||\widehat{u}(\eta)|dxd\eta\\&\leqslant\frac{1}{(2\pi)^n}\int_{\mathbb{R}^n}\int_{\Omega_1}|\psi(x)|C_0(1+|\xi|)^{N_0}\frac{C_N}{(1+|\eta|+|\xi|)^N}dxd\eta\\&\leqslant C'\int_{\mathbb{R}^n}\int_{\Omega_1}\frac{|\psi(x)|}{(1+|\eta|+|\xi|)^{N-N_0}}dxd\eta\\&\leqslant\frac{C'}{(1+|\xi|)^m}\int_{\mathbb{R}^n}\int_{\Omega_1}\frac{|\psi(x)|}{(1+|\eta|+|\xi|)^{N-N_0-m}}dxd\eta.\end{aligned}$$

  所以，当 N 选取的足够大时（$N>N_0+m+n+1$），对任意的 m，我们就有

  $$I_{\varepsilon,1}(\xi)\leqslant\frac{C'}{(1+|\xi|)^m}.$$

  特别地，这个估计对 ε 是一致的。

- 控制 $I_{\varepsilon,2}(\xi)$。

此时，$\eta \notin \Omega_1$。所以，按照定义，我们知道

$$|\eta - \nabla_x \phi(x,\xi)| \geqslant \varepsilon_1 |\nabla_x \phi(x,\xi)| \text{ 并且 } |\eta - \nabla_x \phi(x,\xi)| \geqslant \varepsilon_1 |\eta|.$$

所以，

$$|\eta - \nabla_x \phi(x,\xi)| \geqslant \frac{\varepsilon_1}{2} \left(|\nabla_x \phi(x,\xi)| + |\eta| \right).$$

利用 $\phi(x,\xi)$ 对 ξ 的齐次性，我们就有

$$|\eta - \nabla_x \phi(x,\xi)| \geqslant \frac{\varepsilon_1}{2} \left(1 + |\xi| + |\eta| \right),$$

在这里，我们假设了 $|\xi| \geqslant 1$，实际上，我们总是可以做这个假设，因为我们只对这种情况感兴趣（证明积分对 $|\xi| \to \infty$ 的衰减）。

我们把 $I_{\varepsilon,2}$ 重新写成

$$I_{\varepsilon,2}(\xi) = \frac{1}{(2\pi)^n} \int_{\mathbb{R}^n} \int_{\Omega_2} e^{-i\widetilde{\phi}(x,\xi,\eta)} \widetilde{a}(x,\xi) \chi_\varepsilon(\eta) \widehat{u}(\eta) dx d\eta,$$

其中，

$$\widetilde{\phi}(x,\xi,\eta) = \phi(x,\xi) - x \cdot \eta.$$

所以，上面的推导表明，对任意的 $\eta \in \Omega_2$，我们有

$$|\nabla_x \widetilde{\phi}(x,\xi,\eta)| \geqslant \frac{\varepsilon_1}{2} \left(1 + |\xi| + |\eta| \right).$$

所以，这个函数对 x 的导数是有下界的，我们将用我们第一学期学过的振荡积分的想法。

我们定义微分算子

$$L = \sum_{j \leqslant n} \frac{i\partial_{x_j} \widetilde{\phi}}{|\nabla_x \widetilde{\phi}|^2} \partial_{x_j},$$

很明显，我们有

$$L \left(e^{-i\widetilde{\phi}} \right) = e^{-i\widetilde{\phi}}.$$

对任意的光滑有紧支集的函数 $f(x)$ 和 $g(x)$，通过分部积分，我们有

$$\int_{\mathbb{R}^n} (Lf)(x) \cdot g(x) dx = \int_{\mathbb{R}^n} f(x) \cdot ({}^t Lg)(x) dx,$$

其中，

$$({}^t Lg)(x) = -\sum_{j \leqslant n} \partial_{x_j} \left(\frac{i\partial_{x_j} \widetilde{\phi} g(x)}{|\nabla_x \widetilde{\phi}|^2} \right).$$

据此，我们有

$$\begin{aligned} I_{\varepsilon,2}(\xi) &= \frac{1}{(2\pi)^n} \int_{\mathbb{R}^n} \int_{\Omega_2} L^N \left(e^{-i\widetilde{\phi}(x,\xi,\eta)} \right) \widetilde{a}(x,\xi) \chi_\varepsilon(\eta) \widehat{u}(\eta) dx d\eta \\ &= \frac{1}{(2\pi)^n} \int_{\mathbb{R}^n} \int_{\Omega_2} e^{-i\widetilde{\phi}(x,\xi,\eta)} \times ({}^t L)^N (\widetilde{a}(x,\xi)) \times \chi_\varepsilon(\eta) \widehat{u}(\eta) dx d\eta \end{aligned}$$

根据 (b)，由于对任意的多重指标 α，我们都有

$$|\partial_x^\alpha a(x,\xi)| \leqslant C_\alpha(1+|\xi|)^{N_0}.$$

所以，对任意的 $|\alpha| \leqslant N$，我们就有

$$\left|\partial_x^\alpha \widetilde{a}(x,\xi)\right| = \left|\partial_x^\alpha[\psi(x)\cdot a(x,\xi)]\right| \leqslant C_\alpha(1+|\xi|)^{N_0}.$$

另外，对任意的 $|\alpha| \leqslant N$，利用 $\phi(x,\xi)$ 对频率分量的齐次性（以及 $x \in K$），我们还有

$$|\partial_x^\alpha \phi(x,\xi)| = |\partial_x^\alpha \phi\left(x, \frac{\xi}{|\xi|}\right)||\xi| \leqslant C_N|\xi|.$$

根据这些性质以及

$$|\nabla_x \widetilde{\phi}(x,\xi,\eta)| \geqslant \frac{\varepsilon_1}{2}\left(1+|\xi|+|\eta|\right),$$

其中 ε_1 是一个只依赖于 α_1 和 α_1 的常数，我们就有

$$\left|({}^tL)^N(\widetilde{a}(x,\xi))\right| \leqslant C_N(1+|\xi|)^{N_0}(1+|\xi|+|\eta|)^{-N}\mathbf{1}_K(x).$$

所以，

$$\begin{aligned} I_{\varepsilon,2}(\xi) &\leqslant C_N \int_{\mathbb{R}^n}\int_{\Omega_2}(1+|\xi|)^{N_0}(1+|\xi|+|\eta|)^{-N}\mathbf{1}_K(x)|\chi_\varepsilon(\eta)\widehat{u}(\eta)|dxd\eta \\ &\leqslant C'\int_{\Omega_2}(1+|\xi|+|\eta|)^{-N+N_0}|\widehat{u}(\eta)|dxd\eta. \end{aligned}$$

由于对任意的 $u \in \mathcal{E}'(\mathbb{R}^n) \subset \mathcal{S}'(\mathbb{R}^n) \subset \bigcup_{s\in\mathbb{R}} H^s(\mathbb{R}^n)$，所以，我们可以假设 $u \in H^{s_0}(\mathbb{R}^n)$，从而，

$$\begin{aligned} I_{\varepsilon,2}(\xi) &\leqslant \frac{C'}{(1+|\xi|)^m}\int_{\mathbb{R}^n}(1+|\xi|+|\eta|)^{-N+N_0+m-\frac{s_0}{2}} \times (1+|\xi|)^{\frac{s_0}{2}}|\widehat{u}(\eta)|dxd\eta \\ &\leqslant \frac{C'}{(1+|\xi|)^m}\left(\int_{\mathbb{R}^n}(1+|\xi|+|\eta|)^{-2N+2N_0+2m+s_0}d\eta\right)\|u\|_{H^{s_0}} \\ &= \frac{C''}{(1+|\xi|)^m}. \end{aligned}$$

其中，我们只要选取 $2N > 2N_0 + 2m - s_0 + 1$ 即可。这个估计仍然不依赖于 ε。

综合上面的论证，对任意的 m，我们都有 $C(m)$，使得

$$|I(\xi)| \leqslant \lim_{\varepsilon\to 0^+} I_{\varepsilon,1}(\xi) + I_{\varepsilon,2}(\xi) \leqslant \frac{C(m)}{(1+|\xi|)^m}.$$

这就完成了证明。 □

我们利用这个引理来证明波前集在微分同胚下的变化：

波前集在微分同胚下的变化. 根据波前集的定义，我们选取 $f(x)$ 是支撑集在 $x_0 \in \Omega$ 附近的光滑函数，我们需要计算 $f \cdot \Phi^* \widetilde{u}$ 的 Fourier 变换，其中 $u = \Phi^* \widetilde{u}$。我们有

$$\begin{aligned}
\widehat{f\Phi^* u}(\xi) &= \int_{\mathbb{R}^n_x} e^{-ix\cdot\xi} f(x)\widetilde{u}(\Phi(x))dx \\
&= \int_{\mathbb{R}^n_x} e^{-i\Phi^{-1}(y)\cdot\xi} \underbrace{f(\Phi^{-1}(y))}_{F(y)} \widetilde{u}(y)|d\Phi^{-1}(y)|dy \\
&= \int_{\mathbb{R}^n_x} e^{-i\Phi^{-1}(y)\cdot\xi} \underbrace{f(\Phi^{-1}(y))|d\Phi^{-1}(y)|}_{F(y)} \widetilde{u}(y)dy.
\end{aligned}$$

这里，$F(y)$ 是支撑集在 $y_0 \in \widetilde{\Omega}$ 附近的光滑函数。由于 F 具有紧支集，所以我们不妨假设 $\widetilde{u}$ 也具有紧支集。

我们现在假设 $(y_0, \eta_0) \notin WF(\widetilde{u})$，所以，根据定义波前集，$\widetilde{u}$ 满足引理中 (c) 的要求。同样，相函数 $\phi(y,\xi) = \Phi^{-1}(y)\cdot\xi$ 满足引理中 (a) 的要求；振幅函数 $a(y,\xi) = F(y)$ 满足引理中 (b) 的要求。我们现在要找到满足引理中的条件的 ξ。实际上，我们有

$$\nabla_y \Phi(y,\xi) = {}^t d\Phi^{-1}(y)\cdot\xi.$$

所以，对于 ${}^t d\Phi^{-1}(y)(\xi_0) = \eta_0$，只要 ξ 落在 ξ_0 附近的锥 $\Gamma_0 = \Gamma(\xi_0, \alpha_1)$，其中 α_1 足够小，那么，

$$|\nabla_y \Phi(y,\xi) - \eta_0|$$

就足够小，所以，引理的条件成立，引理的结论表明 $(x_0,\xi_0) \notin WF(u)$，这就证明了命题。 □

84 微局部的椭圆正则性定理，拟解（paramatrix）的构造，双特征曲线（bicharacteristics），奇性传播定理的叙述

二零二零年十二月二十一日，星期一，晴

给定区域 Ω 上的一个 m-次微分算子

$$P=\sum_{|\alpha|\leqslant m}p_\alpha(x)\partial^\alpha,$$

它的 P 的**主象征**$\sigma_m(P)$ 是 $T^*\Omega$ 上的光滑函数，其定义为

$$p_m(x,\xi)=\sum_{|\alpha|=m}p_\alpha(x)(-i\xi)^\alpha.$$

我们可以利用单色波来来计算 P 的主象征：

$$\begin{aligned}e^{ix\cdot\xi}P(e^{-ix\cdot\xi})&=\sum_{|\alpha|\leqslant m}e^{ix\cdot\xi}p_\alpha(x)\partial^\alpha\left(e^{-ix\cdot\xi}\right)\\&=\sum_{|\alpha|\leqslant m}p_\alpha(x)(-i\xi)^\alpha.\end{aligned}$$

所以，我们只要取 ξ 的 m-次齐次部分即可。

另外，如果我们用 ${}^tP$ 表示 P 的对偶算子（这里，我们不用 L^2 的内积），即对任意的 $f,g\in C_0^\infty$，我们有

$$\langle f,Pg\rangle=\langle {}^tPf,g\rangle,$$

那么，

$${}^tPf=\sum_{|\alpha|\leqslant m}(-1)^\alpha\partial^\alpha\big(p_\alpha(x)\cdot f(x)\big).$$

特别地，这也是 m-次的微分算子并且我们有

$$\sigma_m\left({}^tP\right)=\sum_{|\alpha|=m}p_\alpha(x)(i\xi)^\alpha=(-1)^m\sigma_m\left(P\right).$$

我们要证明如下的定理：

定理 551 (微局部椭圆正则性)**.** 给定区域 Ω 上的 m-次微分算子 P 和分布 $u\in\mathcal{D}'(\Omega)$。那么，我们有

$$WF(u)\subset Z\left(\sigma_m(P)\right)\cup WF\left(P(u)\right),$$

其中，$Z\left(\sigma_m(P)\right)$ 是 P 的主象征的零点集。

我们先来把证明的目标讲清楚: 假设 $(x_0,\xi_0)\notin Z(\sigma_m(P))\cup WF(P(u))$, 我们要证明 $(x_0,\xi_0)\notin WF(u)$。由于 $p_m(x,\xi)=\sigma_m(P)(x,\xi)$ 对 ξ 是齐次的，所以，$(x_0,\xi_0)\notin Z(\sigma_m(P))$ 意味着存在紧集 $K_1\subset\Omega$（$x_0\in K_1$），常数 $\alpha_1>0$ 和 $C_1>0$，使得对任意的 $(x,\xi)\in K_1\in\Gamma(\xi_0,\alpha_1)$，我们都有 $|p_m(x,\xi)|\geqslant C|\xi|^m$；$(x_0,\xi_0)\notin WF(P(u))$，意味着存在紧集 $K_1\subset\Omega$（$x_0\in K_2$）和 $\alpha_2>0$ 使得对任意的 $\varphi\in C_0^\infty(K_2)$，对任意的整数 $N\geqslant 1$，存在 C_N，使得对任意的 $\xi\in\Gamma(\xi_0,\alpha_2)$，使得 $(1+|\xi|)^N|\widehat{\varphi\cdot Pu}(\xi)|\leqslant C_N$。所以，综合这些表述，我们就有:

存在紧集 $K\subset\Omega$，$x_0\in\Omega$，存在常数 $\beta>0$，$C>0$，使得

- 对任意的 $(x,\xi)\in K\in\Gamma(\xi_0,\beta)$，我们有
$$|p_m(x,\xi)|\geqslant C|\xi|^m.$$
- 对任意的 $\varphi\in C_0^\infty(K)$，对任意的整数 $N\geqslant 1$，存在 $C_N>0$，使得对任意的 $\xi\in\Gamma(\xi_0,\beta)$，使得
$$|\widehat{\varphi\cdot Pu}(\xi)|\leqslant\frac{C_N}{(1+|\xi|)^N}.$$

我们要证明存在紧集合 $K'\subset K$，$x_0\in K$ 和常数 $\beta'<\beta$，使得对任意的 $\varphi\in C_0^\infty(K')$，对任意的整数 $N\geqslant 1$，存在 $C'_N>0$，使得对任意的 $\xi\in\Gamma(\xi_0,\beta')$，使得
$$|\widehat{\varphi\cdot u}(\xi)|\leqslant\frac{C'_N}{(1+|\xi|)^N},$$
这就说明了 $(x_0,\xi_0)\notin WF(u)$。

注记 (一个"直观的"证明思路). 我们实际上关心的是 $\widehat{\varphi\cdot u}(\xi)$（在某个锥形邻域里）的衰减。形式上，我们有
$$\widehat{\varphi\cdot u}(\xi)=\int_{\mathbb{R}^n}e^{-ix\cdot\xi}\varphi(x)u(x)dx.$$
我们希望能至少对 $e^{-ix\cdot\xi}\varphi(x)$ 定义 ${}^tP$ 的逆，即找到 $a(x,\xi)$，使得（只对 x 作用，ξ 视作是参数）
$${}^tP\left(e^{-ix\cdot\xi}a(x,\xi)\right)=e^{-ix\cdot\xi}\varphi(x).$$
那么，我们就有
$$\begin{aligned}\widehat{\varphi\cdot u}(\xi)&=\int_{\mathbb{R}^n}{}^tP\left(e^{-ix\cdot\xi}a(x,\xi)\right)u(x)dx\\&=\int_{\mathbb{R}^n}e^{-ix\cdot\xi}a(x,\xi)Pu(x)dx\\&=\langle e^{-ix\cdot\xi}a(x,\xi)Pu(x)\rangle.\end{aligned}$$
此时，$\widehat{Pu}$ 在特定的锥之内有衰减，我们可以选取相函数为 $-x\cdot\xi$，从而，我们可以用上次课所提到的非驻相法的引理。

另外，根据
$$\widehat{\varphi\cdot u}(\xi)=\langle e^{-ix\cdot\xi}a(x,\xi),Pu(x)\rangle.$$
我们知道 $a(x,\xi)$ 对 ξ 的衰减越快越好。

拟解（paramatrix）的构造

我们首先做一番计算上的准备：

引理 552. 假设 $c(x,\xi)\in C^\infty\big(K\times\Gamma(\xi_0,\beta)\big)$，$c(x,\xi)$ 对每个固定的 ξ 都是 ξ 的 d-次齐次函数并且如果 $x\notin K$，那么 $c(x,\xi)=0$。那么，我们有

$$ {}^tP\left(e^{-ix\cdot\xi}c(x,\xi)\right)=e^{-ix\cdot\xi}(-1)^m p_m(x,\xi)c(x,\xi)+e^{-x\cdot\xi}\sum_{1\leqslant k\leqslant m}c_k(x,\xi), $$

其中，对每个 $k=1,\cdots,m$，函数 $c_k(x,\xi)\in C^\infty\big(K\times\Gamma(\xi_0,\beta)\big)$，$c(x,\xi)$ 对每个固定的 ξ 都是 ξ 的 $d+m-k$-次齐次函数并且如果 $x\notin K$，那么 $c_k(x,\xi)=0$。

注记. 我们认为

$$ c_0(x,\xi)=p_m(x,\xi)c(x,\xi). $$

它的次数是 $d+m-k$。根据前一个注解，k 越大，$c_k(x,\xi)$ 对 ξ 的衰减就越好，所以，我们认为 c_k（$k\geqslant 1$）是比首项 $p_m(x,\xi)c(x,\xi)$ 要“好”的。

证明: 实际上，我们有

$$ \begin{aligned}
{}^tP\left(e^{-ix\cdot\xi}c(x,\xi)\right)&=\sum_{|\alpha|\leqslant m}(-1)^\alpha\partial^\alpha\big(e^{-ix\cdot\xi}\underbrace{p_\alpha(x)c(x,\xi)}_{c_\alpha(x)}\big)\\
&=\sum_{|\alpha|\leqslant m}(-1)^\alpha\sum_{\gamma\leqslant\alpha}\binom{\alpha}{\gamma}\partial^\gamma\big(e^{-ix\cdot\xi}\big)\partial^{\alpha-\gamma}\big(c_\alpha(x,\xi)\big)\\
&=\left(\sum_{|\alpha|=m}+\sum_{|\alpha|<m}\right)(-1)^\alpha\sum_{\gamma\leqslant\alpha}\binom{\alpha}{\gamma}\partial^\gamma\big(e^{-ix\cdot\xi}\big)\partial^{\alpha-\gamma}\big(c_\alpha(x,\xi)\big)\\
&=\sum_{|\gamma|=m}(-1)^m\partial^\gamma\big(e^{-ix\cdot\xi}\big)\cdot c_\gamma(x)+\sum_{\gamma\leqslant\alpha,|\gamma|\leqslant m-1,|\alpha|\leqslant m-1}\pm\binom{\alpha}{\gamma}\partial^\gamma\big(e^{-ix\cdot\xi}\big)\partial^{\alpha-\gamma}\big(c_\alpha(x,\xi)\big)\\
&=\sum_{|\alpha|=m}(i\xi)^\alpha e^{-ix\cdot\xi}p_\alpha(x)c(x,\xi)+\sum_{0\leqslant k\leqslant m-1}\sum_{|\gamma|=k}\pm\binom{\alpha}{\gamma}e^{-ix\cdot\xi}\underbrace{\xi^\gamma\partial^{\alpha-\gamma}\big(c_\alpha(x,\xi)\big)}_{\geqslant d+m-k\text{ 次}}\\
&=e^{-ix\cdot\xi}(-1)^m p_m\alpha(x,\xi)c(x,\xi)+e^{-x\cdot\xi}\sum_{1\leqslant k\leqslant m}c_k(x,\xi).
\end{aligned} $$

按照定义，每个 $c_\alpha(x,\xi)$ 与 $c(x,\xi)$ 满足同样的性质，从而命题成立。 □

注意到，我们想构造 $a(x,\xi)$，使得

$$ {}^tP\left(e^{-ix\cdot\xi}a(x,\xi)\right)=e^{-ix\cdot\xi}\varphi(x). $$

假设 $a(x,\xi)$ 是 d-次的，根据上面的计算，我们有

$$ e^{-ix\cdot\xi}(-1)^m p_m(x,\xi)a(x,\xi)+e^{-ix\cdot\xi}O(d+m-1)=e^{-ix\cdot\xi}\varphi(x), $$

其中，$O(d+m-1)$ 代表的是一些次数至多为 $d+m-1$ 次（关于 ξ）的齐次函数的和。我们如果令

$$a(x,\xi)=a_0(x,\xi)=\frac{(-1)^m}{p_m(x,\xi)}\varphi(x),$$

其中，我们假设 $(x,\xi)\in K\times\Gamma(\xi_0,\beta)$（从而，$p_m(x,\xi)\neq 0$）那么，（此时 $d=m$）

$${}^tP\left(e^{-ix\cdot\xi}a_0(x,\xi)\right)-e^{-ix\cdot\xi}\varphi(x)=e^{-ix\cdot\xi}R_0(x,\xi),$$

那么，$R_0(x,\xi)\in C^\infty\big(K\times\Gamma(\xi_0,\beta)\big)$ 并且如果 $x\notin K$，那么 $R_0(x,\xi)=0$。进一步，$R_0(x,\xi)$ 对是有限个 ξ 的次数 $\leqslant -1$ 的齐次函数的和。

我们对 $a_0(x,\xi)$ 再加上一个低一次的扰动 $a_1(x,\xi)$（从而，次数为 $-m-1$），使得

$${}^tP\left(e^{-ix\cdot\xi}[a_0(x,\xi)+a_1(x,\xi)]\right)-e^{-ix\cdot\xi}\varphi(x)=e^{-ix\cdot\xi}R_1(x,\xi),$$

并且 $R_1(x,\xi)\in C^\infty\big(K\times\Gamma(\xi_0,\beta)\big)$ 并且如果 $x\notin K$，那么 $R_1(x,\xi)=0$。进一步，$R_1(x,\xi)$ 对是有限个 ξ 的次数 $\leqslant -2$ 的齐次函数的和。为此，我们计算

$$\begin{aligned}
&{}^tP\left(e^{-ix\cdot\xi}[a_0(x,\xi)+a_1(x,\xi)]\right)-e^{-ix\cdot\xi}\varphi(x)\\
&=e^{-ix\cdot\xi}R_0(x,\xi)+{}^tP\left(e^{-ix\cdot\xi}a_1(x,\xi)\right)\\
&=e^{-ix\cdot\xi}R_0(x,\xi)+e^{-ix\cdot\xi}(-1)^m p_m(x,\xi)a_1(x,\xi)+O\big((-m-1)+m-1\big)
\end{aligned}$$

由于 $R_0(x,\xi)$ 是一些次数不超过 -1 的齐次函数的线性组合，我们单独把 -1 次（如果存在）的齐次函数部分拿出来记作是 $R_0^{(-1)}(x,\xi)$。所以，如果令

$$a_1(x,\xi)=\frac{(-1)^{m+1}R_0^{(-1)}(x,\xi)}{p_m(x,\xi)},$$

我们就消去了所有最高次（-1 次）的项（次数恰好满足要求）。我们强调，我们是在 $(x,\xi)\in K\times\Gamma(\xi_0,\beta)$ 中做的计算。

重复上述计算，对于每个整数 $k\geqslant 0$，我们都可以构造 $b_k(x,\xi)$ 和 $R_k(x,\xi)$，使得

- $b_k(x,\xi)\in C^\infty\big(K\times\Gamma(\xi_0,\beta)\big)$ 并且如果 $x\notin K$，那么 $b_k(x,\xi)=0$。进一步，$b_k(x,\xi)$ 对 ξ 是次数为 $-m-k$ 的齐次函数。
- $R_k(x,\xi)\in C^\infty\big(K\times\Gamma(\xi_0,\beta)\big)$ 并且如果 $x\notin K$，那么 $R_k(x,\xi)=0$。进一步，$R_k(x,\xi)$ 对是有限个 ξ 的次数 $\leqslant -k-1$ 的齐次函数的和。
- 我们有

$${}^tP\left(e^{-ix\cdot\xi}\sum_{0\leqslant j\leqslant k}b_j(x,\xi)\right)-e^{-ix\cdot\xi}\varphi(x)=e^{-ix\cdot\xi}R_k(x,\xi).$$

这里，

$$a_k(x,\xi)=\sum_{0\leqslant j\leqslant k}b_j(x,\xi)$$

是我们要找的 $a(x,\xi)$ 的一个好的逼近，我们称作是**拟解**。

注记. 在下面的应用中，我们把 K 将换成略小的紧集 $K'\subset K$。

微局部椭圆正则性定理的证明

我们假设 $p_m(x_0,\xi_0)\neq 0$ 并且 $(x_0,\xi_0)\notin WF(Pu)$，那么，存在紧集 $K\subset\Omega$，$x_0\in\Omega$，存在常数 $\beta>0$，$C>0$，使得

- 对任意的 $(x,\xi)\in K\in\Gamma(\xi_0,\beta)$，我们有

$$|p_m(x,\xi)|\geqslant C|\xi|^m.$$

- 对任意的 $\varphi\in C_0^\infty(K)$，对任意的整数 $N\geqslant 1$，存在 $C_N>0$，使得对任意的 $\xi\in\Gamma(\xi_0,\beta)$，使得

$$|\widehat{\varphi\cdot Pu}(\xi)|\leqslant\frac{C_N}{(1+|\xi|)^N}.$$

我们要证明存在紧集合 $K'\subset K$，$x_0\in K$ 和常数 $\beta'<\beta$，使得对任意的 $\varphi\in C_0^\infty(K')$，对任意的整数 $N\geqslant 1$，存在 $C'_N>0$，使得对任意的 $\xi\in\Gamma(\xi_0,\beta')$，使得

$$|\widehat{\varphi\cdot u}(\xi)|\leqslant\frac{C'_N}{(1+|\xi|)^N}.$$

利用前面构造的拟解，我们有

$$\begin{aligned}\widehat{\varphi\cdot u}(\xi)&=\left\langle u(x),e^{-ix\cdot x}\varphi(x)\right\rangle\\&=\left\langle u(x),{}^tP\left(e^{-ix\cdot\xi}a_k(x,\xi)\right)-e^{-ix\cdot\xi}R_k(x,\xi)\right\rangle\\&=\underbrace{\left\langle(Pu)(x),e^{-ix\cdot\xi}a_k(x,\xi)\right\rangle}_{I(\xi)}-\underbrace{\left\langle u(x),e^{-ix\cdot\xi}R_k(x,\xi)\right\rangle}_{J(\xi)}\end{aligned}$$

首先处理 $J(\xi)$。固定 ξ，我们知道 $R_k(x,\xi)$ 对于 x 是光滑的有紧支集的函数。利用 u 是分布（的定义），存在常数 C_0，使得，

$$|J(\xi)|\leqslant C\sup_{\substack{|\alpha|\leqslant m,\\ x\in K}}\left|\partial^\alpha\left(e^{-ix\cdot\xi}R_k(x,\xi)\right)\right|.$$

另外，$R_k(x,\xi)$ 是有限个 ξ 的次数 $\leqslant -k-1$ 的齐次函数的和，所以，根据 Leibniz 法则，上面右端出现的是一些次数不超过 $m-k-1$ 的齐次函数的和。利用 R_k 的光滑性，我们知道存在常数 C_m，使得

$$\sup_{\substack{|\alpha|\leqslant m-k-1,\\ (x,\xi)\in K\times\mathbf{S}^{n-1}}}|\partial^\alpha R_k(x,\xi)|\leqslant C_m.$$

所以，

$$|J(\xi)|\leqslant\frac{C'_m}{(1+|\xi|)^{m-k-1}},$$

其中，k 是任意的正整数。

最后处理 $I(\xi)$。我们有

$$\begin{aligned} I(\xi) &= \left\langle (Pu)(x), e^{-ix\cdot\xi} a_k(x,\xi) \right\rangle \\ &= \left\langle \varphi(x)(Pu)(x), e^{-ix\cdot\xi} a_k(x,\xi) \right\rangle, \end{aligned}$$

其中，$\psi(x) \in C_0^\infty(K)$ 在 K' 上恒为 1。

我们选取相函数 $\phi(x,\xi) = x\cdot\xi$，振幅函数 $a_k(x,\xi)$ 以及被作用函数 $f(x) = \psi \cdot Pu(x)$，我们来验证它们满足上次证明的非驻相法的引理所要求的基本数据的条件：

(a) $\phi(x,\xi) \in C^\infty\left(\mathbb{R}_x^n \times (\mathbb{R}_\xi^n - \{0\})\right)$ 是实值的，变量 ξ 是 1 次齐次的并且当 $\xi \neq 0$ 时，有 $\nabla_x \phi(x,\xi) \neq 0$。

(b) 按照构造，$b_j(x,\xi) \in C^\infty\left(\mathbb{R}_x^n \times (\mathbb{R}_\xi^n - \{0\})\right)$，所以，$a_k(x,\xi) \in C^\infty\left(\mathbb{R}_x^n \times (\mathbb{R}_\xi^n - \{0\})\right)$。另外，$a_k(x,\xi)$ 是一些次数不超过 $-m-k$ 的齐次函数的和，并且当 $x \notin K'$ 时，$a_k(x,\xi) = 0$。类似于上面对 $J(\xi)$ 的估计，取 $N_0 = -m$，使得对任意的多重指标 α，存在常数 C_α，使得对任意的 $(x,\xi) \in \mathbb{R}_x^n \times (\mathbb{R}_\xi^n - \{0\})$，我们都有

$$\left|\partial_x^\alpha a_k(x,\xi)\right| \leqslant C_\alpha \left(1+|\xi|\right)^{N_0}.$$

(c) 很明显，$\operatorname{supp}(f) \subset K$。又因为 $(x_0, \xi_0 \notin WF(Ph)$，根据我们已知的条件，对任意的 $N \geqslant 1$，存在常数 C_N，使得对任意的 $\xi \in \Gamma(\xi_0, \beta)$，都有我们就有

$$\left|\widehat{f}(\xi)\right| \leqslant \frac{C_N}{(1+|\xi|)^N}.$$

这里，$\eta_0 = \dfrac{\xi_0}{|\xi_0|}$，$\alpha_0 = \beta$。

现在我们任意选定 $\alpha_1 = \beta' < \beta = \alpha_0$，$\Gamma_0 = \Gamma(\xi_0, \beta')$。我们只要验证非驻相法的引理的条件即可，即证明对任意的 $(x,\xi) \in K \times \Gamma_0$，我们都有

$$\left|\frac{\nabla_x \phi(x,\xi)}{|\nabla_x \phi(x,\xi)|} - \eta_0\right| < \beta' = \alpha_1.$$

实际上，由于

$$\frac{\nabla_x \phi(x,\xi)}{|\nabla_x \phi(x,\xi)|} = \frac{\xi}{|\xi|},$$

所以，这是显然的。

此时，引理的结论表明，对任意的 $N \geqslant 1$，存在 $C_N > 0$，使得对任意的 $\xi \in \Gamma_0$，有

$$|I(\xi)| \leqslant \frac{C_N}{(1+|\xi|)^N}.$$

综合 $I(\xi)$ 和 $J(\xi)$ 的估计，我们就证明了微局部版本的椭圆正则性定理。

奇性传播定理

给定区域 Ω 上的 m-次微分算子 P 和分布 $u \in \mathcal{D}'(\Omega)$，假设 $Pu = f$ 是光滑的，那么，

$$WF(u) \subset Z\left(\sigma_m(P)\right),$$

其中，$Z\left(\sigma_m(P)\right)$ 是 P 的主象征的零点集。我们现在要进一步刻画 $WF(u)$ 的几何结构，这就是所谓的奇性传播定理。

我们先引入一些定义：

定义 553. 给定区域 Ω 上的 m-次微分算子 P，我们把

$$\mathrm{Char}(P) := Z\left(\sigma_m(P)\right) \subset T^*\Omega = \Omega \times \mathbb{R}^n$$

称作是 P 的**特征簇**（*characteristic variety*）。

给定 $(x_0, \xi_0) \in \mathrm{Char}(P)$，如果 $p_m(x,\xi) = \sigma_m(P)$ 在 (x_0,ξ_0) 的一个邻域上是实值的并且 $\nabla_{x,\xi} p_m(x_0,\xi_0)$ 与 $(\xi_0, 0)$ 是线性无关的，我们就称 $\mathrm{Char}(P)$ 在 (x_0,ξ_0) 处是**简单的**，也称 P 在 (x_0,ξ_0) 处具有**简单特征**。

如果 P 在每个 $\mathrm{Char}(P)$ 上的点处都是简单的，我们就称 P 具有**简单的特征簇**。

注记. 假设微分算子 P 具有简单的特征簇，那么，对每个 $(x_0,\xi_0) \in \mathrm{Char}(P)$，$p_m(x,\xi)$ 在 (x_0,ξ_0) 处的微分不是 0，这说明 $\mathrm{Char}(P) \subset T^*\Omega$ 是光滑子流形（余维数为 1）。

注记. 之前，我们按照如下的方式定义了主特征：如果

$$P = \sum_{|\alpha| \leqslant m} p_\alpha(x) \partial^\alpha,$$

那么，

$$p_m(x,\xi) = \sigma_m(P)(x,\xi) \sum_{|\alpha|=m} p_\alpha(x)(-i\xi)^\alpha.$$

从此之后，因为有上述关于 $p_m(x,\xi)$ 是实值的要求，我们令

$$p_m(x,\xi) = \sigma_m(P)(x,\xi) = \sum_{|\alpha|=m} p_\alpha(x)(\xi)^\alpha.$$

这对之前的证明没有影响。

定义 554. 给定区域 Ω 上的 m-次微分算子 P，我们假设它的象征 $p_m(x,\xi)$ 是实数值的函数，那么如下定义的 $T^*\Omega$ 上的向量场

$$\mathbf{H}_P = \left(\frac{\partial p_m}{\partial \xi_1}, \cdots, \frac{\partial p_m}{\partial \xi_n}, -\frac{\partial p_m}{\partial x_1}, \cdots, -\frac{\partial p_m}{\partial \xi_n} \right)$$

被称作是 P 所定义的 *Hamilton* 向量场。形式上，我们经常把这个向量场写成

$$\mathbf{H}_P = \left(\frac{\partial p_m}{\partial \xi}, -\frac{\partial p_m}{\partial x} \right).$$

我们通常把 $H(x,\xi) = p_m(x,\xi)$ 称作是 $T^*\Omega$ 上的一个 **Hamilton 作用量**。

假设

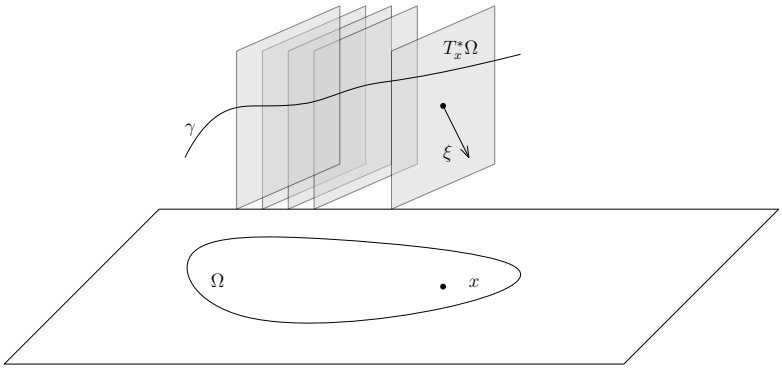
$$\gamma : (-a, b) \to T^*\Omega$$

是过 $(x_0, \xi_0) \in T^*\Omega$ 的 $\mathbf{H}_p$ 的一条积分曲线，其中 $a, b > 0$。这指的是

$$\gamma(t) = \big(x_1(t), x_2(t), \cdots, x_n(t), \xi_1(t), \xi_2(t), \cdots, \xi_n(t)\big)$$

并且

$$\begin{cases} x_1'(t) &= \frac{\partial p_m}{\partial \xi_1}\big(x_1(t), \cdots, x_n(t), \xi_1(t), \cdots, \xi_n(t)\big), \\ x_2'(t) &= \frac{\partial p_m}{\partial \xi_2}\big(x_1(t), \cdots, x_n(t), \xi_1(t), \cdots, \xi_n(t)\big), \\ & \cdots\cdots \\ x_n'(t) &= \frac{\partial p_m}{\partial \xi_n}\big(x_1(t), \cdots, x_n(t), \xi_1(t), \cdots, \xi_n(t)\big), \\ \xi_1'(t) &= -\frac{\partial p_m}{\partial x_1}\big(x_1(t), \cdots, x_n(t), \xi_1(t), \cdots, \xi_n(t)\big), \\ \xi_2'(t) &= -\frac{\partial p_m}{\partial x_2}\big(x_1(t), \cdots, x_n(t), \xi_1(t), \cdots, \xi_n(t)\big), \\ & \cdots\cdots \\ \xi_n'(t) &= -\frac{\partial p_m}{\partial x_b}\big(x_1(t), \cdots, x_n(t), \xi_1(t), \cdots, \xi_n(t)\big), \\ \gamma(0) &= (x_0, \xi_0). \end{cases}$$



在上面这个图中，我们把每个点 $x \in \Omega$ 处的 $\{x\} \times \mathbb{R}^n$ 记作是 $T_x^*\Omega$。我们经常把这个方程组简写成

$$\begin{cases} x'(t) &= \frac{\partial p_m}{\partial \xi}\big(x(t), \xi(t)\big), \\ \xi'(t) &= -\frac{\partial p_m}{\partial x}\big(x(t), \xi(t)\big), \\ \gamma(0) &= (x_0, \xi_0). \end{cases}$$

关于 Hamilton 向量场，我们有如下熟知的性质：

引理 555. *Hamilton 作用量 $p_m(x, \xi)$ 沿着 $\mathbf{H}_P$ 的积分曲线是常数，即对任意的 $\mathbf{H}_p$ 的积分曲线*

$$\gamma : (-a, b) \to T^*\Omega,$$

我们有

$$\frac{d}{dt}\left(p_m(x(t), \xi(t))\right) = 0.$$

证明: 我们只要利用 γ 的方程进行计算即可:

$$\begin{aligned}\frac{d}{dt}\left(p_m(x(t),\xi(t))\right) &= \frac{\partial p_m}{\partial x}\big(x(t),\xi(t)\big)x'(t) + \frac{\partial p_m}{\partial \xi}\big(x(t),\xi(t)\big)\xi'(t)\\ &= \frac{\partial p_m}{\partial x}\big(x(t),\xi(t)\big)\frac{\partial p_m}{\partial \xi}\big(x(t),\xi(t)\big) + \frac{\partial p_m}{\partial \xi}\big(x(t),\xi(t)\big)\left(-\frac{\partial p_m}{\partial x}\big(x(t),\xi(t)\big)\right)\\ &= 0.\end{aligned}$$

命题成立。 □

假设 γ 是 $\mathbf{H}_P$ 的一条积分曲线，如果 $\gamma \cap \mathrm{Char}(P) \neq \emptyset$，根据上面的引理，那么，$p_m$ 在 γ 的每个点上的取值都是 0，所以，$\gamma \subset \mathrm{Char}(P)$。

定义 556. 我们把落在 $\mathrm{Char}(P)$ 中的 $\mathbf{H}_P$ 的极大的一条积分曲线称作是微分算子 P 的一条**双特征曲线**（*bicharacteristics*）。

有了双特征曲线的概念，我们就可以陈述奇性传播定理了:

定理 557 (奇性传播定理)**.** 给定区域 Ω 上的 m-次微分算子 P，我们假设 P 具有简单的特征簇。如果 γ 是 P 的一条双特征曲线并且

$$\gamma \cap WF\big(Pu\big) = \emptyset,$$

其中 $u \in \mathcal{D}'(\Omega)$ 是分布，那么如下两种情形必居（且只居）其一:

- $\gamma \subset WF(u)$;
- $\gamma \cap WF(u) = \emptyset$.

推论 558 (奇性传播定理)**.** 算子 P 是区域 Ω 上的 m-次微分算子并且具有简单的特征簇,$u \in \mathcal{D}'(\Omega)$ 是分布。如果 Pu 是光滑的，那么 $WF(u)$ 是双特征曲线的无交并。

85 奇性传播定理的证明

二零二零年十二月二十三日，星期三，晴

我们上次课陈述了如下的定理：

定理 559 (奇性传播定理)**.** 给定区域 Ω 上的 m-次微分算子 P，我们假设 P 具有简单的特征簇。如果 γ 是 P 的一条双特征曲线并且

$$\gamma \cap WF\big(Pu\big) = \emptyset,$$

其中 $u \in \mathcal{D}'(\Omega)$ 是分布，那么如下两种情形必居（且只居）其一：

- $\gamma \subset WF(u)$;
- $\gamma \cap WF(u) = \emptyset$.

我们回忆一下其中出现的几个定义：

给定区域 Ω 上的 m-次微分算子 P，P 的特征簇Char(P) 是它的主象征的零点集合。我们假设 P 具有简单的特征簇，特别地，Char$(P) \subset T^*\Omega$ 是光滑子流形。利用 P 的主特征 $p_m(x,\xi)$，我们可以定义它的 Hamilton 向量场：

$$\mathbf{H}_P = \left(\frac{\partial p_m}{\partial \xi}, -\frac{\partial p_m}{\partial x}\right).$$

我们把落在 Char(P) 中的 $\mathbf{H}_P$ 的极大积分曲线称作是双特征曲线。

为了证明这个定理，我们先做一番准备。我们定义相函数

$$\phi(s,x,\xi) : \mathbb{R} \times T^*\Omega \to \mathbb{R}.$$

这是一个对于 ξ 是 1 次齐次的函数，它由如下的常微分方程所定义：

$$\begin{cases} |\xi|^{m-1}\dfrac{\partial \phi}{\partial s}(s,x,\xi) = -p_m\left(x, \nabla_x \phi(s,x,\xi)\right), \\ \phi(0,x,\xi) = x\cdot\xi. \end{cases}$$

其中，m 为 P 的次数。特别地，由于 ϕ 对于 ξ 的次数为 1，所以，ϕ 可以被下面的方程刻画：

$$\begin{cases} \dfrac{\partial \phi}{\partial s}(s,x,\xi) = -p_m\left(x, \nabla_x \phi(s,x,\xi)\right), \\ \phi(0,x,\xi) = x\cdot\xi, \end{cases}$$

其中，$\xi \in \mathbf{S}^{n-1}$。这是因为

$$p_m\left(x, \nabla_x\phi(s,x,\xi)\right) = p_m\left(x, \nabla_x\phi\left(s,x,\frac{\xi}{|\xi|}\right)|\xi|\right) = |\xi|^m p_m\left(x, \nabla_x\phi\left(s,x,\frac{\xi}{|\xi|}\right)\right).$$

利用相函数 $\phi(s,x,\xi)$，我们可以刻画双特征曲线：

引理 560. 假设 $\gamma(t)=\big(x(t),\xi(t)\big)$ 是 P 的一条双特征曲线，那么，对任意的 $t\in\mathbb{R}$，我们有

$$\xi(t)=(\nabla_x\phi)\,(t,x(t),\xi_0),$$

其中，$\gamma(0)=(x_0,\xi_0)$。

我们定义

$$\zeta(t)=(\nabla_x\phi)\,(t,x(t),\xi_0).$$

很明显，

$$\zeta(0)=\nabla_x(x\cdot\xi_0)=\xi_0=\xi(0).$$

我们将证明，对任意的 t，我们都有

$$\xi(t)=\zeta(t).$$

证明: 假设 $|\xi|=1$，我们考虑 $\zeta(t)$ 所满足的微分方程:

$$\zeta'(t)=\frac{\partial^2\phi}{\partial t\partial x}\big(t,x(t),\xi_0\big)+\frac{\partial^2\phi}{\partial x^2}\big(t,x(t),\xi_0\big)x'(t).$$

利用方程，我们有

$$\frac{\partial^2\phi}{\partial t\partial x}\big(t,x(t),\xi_0\big)=-\frac{\partial p_m}{\partial x}\left(x,\frac{\partial\phi}{\partial x}(t,x(t),\xi_0)\right)-\frac{\partial p_m}{\partial\xi}\left(x,\frac{\partial\phi}{\partial x}(t,x(t),\xi_0)\right)\frac{\partial^2\phi}{\partial x^2}\,(t,x(t),\xi_0)\,.$$

所以，

$$\zeta'(t)=-\frac{\partial p_m}{\partial x}\,(x(t),\zeta(t))+\underbrace{\frac{\partial^2\phi}{\partial x^2}\,(t,x(t),\xi_0)\cdot\left[x'(t)-\frac{\partial p_m}{\partial\xi}\,(x(t),\zeta(t))\right]}_{\text{矩阵乘法}}$$

另外，根据定义，我们还有

$$\begin{cases}x'(t) &=\dfrac{\partial p_m}{\partial\xi}\,(x(t),\xi(t))\,,\\ \xi'(t) &=-\dfrac{\partial p_m}{\partial x}\,(x(t),\xi(t))\,.\end{cases}.$$

所以，

$$\begin{aligned}\zeta'(t)-\xi'(t)=&-\left[\frac{\partial p_m}{\partial x}\,(x(t),\zeta(t))-\frac{\partial p_m}{\partial x}\,(x(t),\xi(t))\right]\\&-\frac{\partial^2\phi}{\partial x^2}\,(t,x(t),\xi_0)\cdot\left[\frac{\partial p_m}{\partial\xi}\,(x(t),\zeta(t))-\frac{\partial p_m}{\partial\xi}\,(x(t),\xi(t))\right].\end{aligned}$$

根据 Lagrange 中值定理（我们用到了 p_m 是光滑函数并且其导数在 γ 附近是有界的），我们就有

$$|\zeta'(t)-\xi'(t)|\leqslant C|\zeta(t)-\xi(t)|.$$

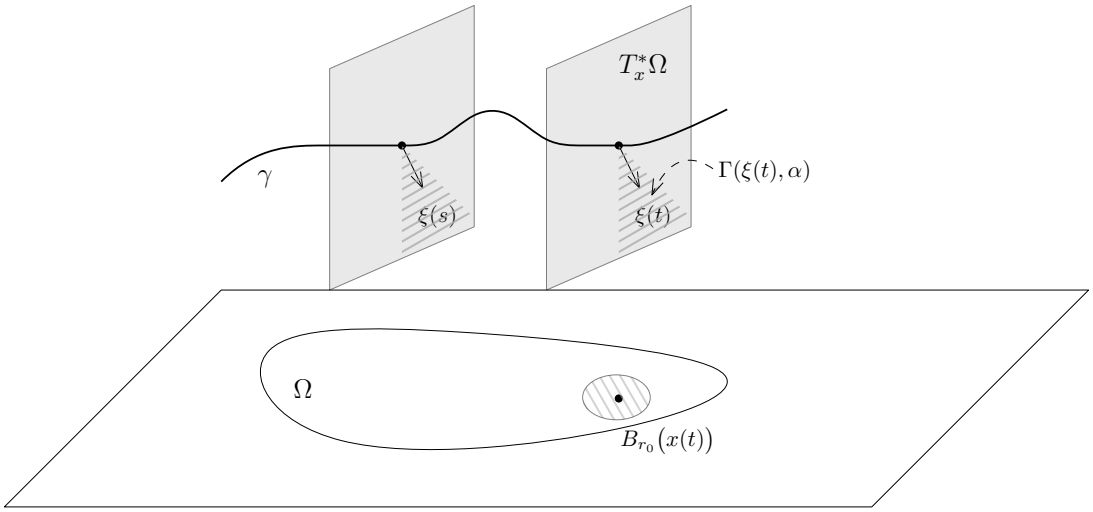
由于

$$|\zeta(t)-\xi(t)|\Big|_{t=0}=0,$$

上面的微分不等式表明

$$\zeta(t)\equiv\xi(t).$$

命题得证。 □



利用微分方程的解对初始值的光滑依赖性以及紧性，我们有如下的推论：

推论 561. *我们考虑一段双特征曲线*

$$\gamma : [0, t_0] \to T^*\Omega.$$

对任意给定 $\alpha_0 > 0$，存在 $r_0 > 0$ 和 $\beta > 0$，使得对任意的 $t \in [0, t_0]$，对任意的 $(x, \xi) \in B_{r_0}\big(x(t)\big) \times \Gamma\big(\xi(t), \beta\big)$，我们都有

$$\frac{\partial \phi}{\partial x}(t, x, \xi) \in \Gamma\big(\xi(t), \alpha_0\big).$$

证明： 对任意的 $t \in [0, t_0]$，根据定理，上述关系对于 $(x(t), \xi(t)) \in B_{r_0}\big(x(t)\big) \times \Gamma\big(\xi(t), \beta\big)$ 成立，利用连续性，就对 $B_{r_0}\big(x(t)\big) \times \Gamma\big(\xi(t), \beta\big)$ 中所有的点都成立。由于所有可能的 t 所给出的 $\big\{B_{r_0}\big(x(t)\big) \times \Gamma\big(\xi(t), \beta\big)\big\}$ 是 γ 的开覆盖，利用紧性，我们就得到了结论。 □

奇性传播定理的证明：条件

我们首先把条件

$$\gamma \cap WF\big(Pu\big) = \emptyset,$$

用分析的语言写清楚。

对任意的 $t \leqslant t_0$，$\gamma(t) = \big(x(t), \xi(t)\big) \notin WF(Pu)$，从而，存在 $\psi_0(x) \in C_0^\infty(\Omega)$，$\alpha(t) > 0$，使得 ψ_0 在 $x(t)$ 附近恒为 1 并且对任意的 $N > 0$，存在 C_N，对任意的 $\xi \in \Gamma(\xi(t), 2\alpha(t))$，我们有

$$\Big|\widehat{\psi_0 \cdot Pu}(\xi)\Big| \leqslant \frac{C_N}{(1 + |\xi|)^N}.$$

所以，对任意的 $\varphi \in C_0^\infty(B_{r(t)}(x(t)))$（半径 $r(t)$ 很小），其中 $\psi_0\big|_{B_{r(t)}(x(t))} \neq 0$，对任意的 $\xi \in$

$\Gamma(\xi(t),\alpha(t))$，我们有

$$\begin{aligned}\left|\widehat{\varphi\cdot Pu}(\xi)\right| &= \left|\mathcal{F}\left(\varphi\cdot\psi_0 Pu\right)(\xi)\right| \\ &\leqslant \frac{C'_N}{(1+|\xi|)^N}\sup_{|\mu|\leqslant M(N)}\|\partial^\mu\varphi\|_{L^\infty}.\end{aligned}$$

最后的一步，我们用了引理544，其中，$M(N)$ 是依赖于 N 的常数。当 t 遍历 $[0,t_0]$ 时（我们可以假设 $t_0=1$），我们知道 $B_{r(t)}(x(t))\times\Gamma(\xi(t),\alpha(t))$ 覆盖了 $\gamma([0,1])$，所以，我们可以选出有限个 $t_1,\cdots,t_l$，使得

$$\gamma\subset\bigcup_{j\leqslant l}B_{r(t_j)}(x(t_j))\times\Gamma(\xi(t_j),\alpha(t_j)):=\Gamma(\gamma).$$

所以，存在 $r_0>0$ 和 $\alpha_0>0$，对任意的 $t\in[0,1]$，$B_{r_0}(x(t))\times\Gamma(\xi(t),\alpha_0)\subset\Gamma(\gamma)$。特别地，根据 Lebesgue 数的存在性（第一学期第 11 次课），对任意的 t，$B_{r_0}(x(t))\times\Gamma(\xi(t),\alpha_0)$ $xi\in\Gamma(\xi(t),\alpha_0)$ 必然落在某个 $B_{r(t_j)}(x(t_j))\times\Gamma(\xi(t_j),\alpha(t_j))$ 中（这里我们可以先在 $\Omega\times\mathbf{S}^{n-1}$ 这个紧集上考虑）。通过选取最小的 C_N 和最大的 $M(N)$，我们就得到了如下的结论：

引理 562. *存在 $r_0>0$ 和 $\alpha_0>0$，对任意的 $t\in[0,1]$，使得对任意的 $N>0$，存在常数 C_N 和 M_N，使得对任意的 $\varphi\in C_0^\infty\left(B_{r_0}(x(t))\right)$，对任意的 $\xi\in\Gamma(\xi(t),\alpha_0)$，我们有*

$$\left|\widehat{\varphi\cdot Pu}(\xi)\right|\leqslant\frac{C_N}{(1+|\xi|)^N}\sup_{|\mu|\leqslant M_N}\|\partial^\mu\varphi\|_{L^\infty}.$$

我们不妨假设 $\gamma(1)\notin WF(u)$，我们只要说明在另一个端点处 $\gamma(0)\notin WF(u)$ 即可。我们先在翻译 $\gamma(1)\notin WF(u)$ 这个条件：存在 $r_1>0$ 和 $\alpha_1>0$，对任意的 $N>0$，存在常数 C_N 和 M_N，使得对任意的 $\varphi\in C_0^\infty\left(B_{r_1}(x(1))\right)$，对任意的 $\xi\in\Gamma(\xi(1),\alpha_1)$，我们有

$$|\widehat{\varphi\cdot u}(\xi)|\leqslant\frac{C_N}{(1+|\xi|)^N}\sup_{|\mu|\leqslant M_N}\|\partial^\mu\varphi\|_{L^\infty}.$$

作为总结（缩小 r_0 和 r_1 等），我们把奇性传播定理的条件总结如下：

存在 $r_0>0$ 和 $\alpha_0>0$，对任意的 $t\in[0,1]$，使得对任意的 $N>0$，存在常数 C_N 和 M_N，使得

- 对任意的 $\varphi\in C_0^\infty\left(B_{r_0}(x(t))\right)$，对任意的 $\xi\in\Gamma(\xi(t),\alpha_0)$，我们有

$$\left|\widehat{\varphi\cdot Pu}(\xi)\right|\leqslant\frac{C_N}{(1+|\xi|)^N}\sup_{|\mu|\leqslant M_N}\|\partial^\mu\varphi\|_{L^\infty}.$$

- 对任意的 $\varphi\in C_0^\infty\left(B_{r_0}(x(1))\right)$，对任意的 $\xi\in\Gamma(\xi(1),\alpha_0)$，我们有

$$|\widehat{\varphi\cdot u}(\xi)|\leqslant\frac{C_N}{(1+|\xi|)^N}\sup_{|\mu|\leqslant M_N}\|\partial^\mu\varphi\|_{L^\infty}.$$

奇性传播定理的证明：拟解的构造

我们需要计算如下形式的导数：

$$-e^{i\phi}\left(|\xi|^{m-1}i\frac{\partial}{\partial t}+{}^tP\right)\left(e^{-i\phi}c\right)$$

假设 $c(t,x,\xi)$ 是阶为 d，那么，我们要将上面的式子写成三个部分，第一个部分的阶为 $d+m$，第二个部分的阶为 $d+m-1$，第二个部分的阶 $\leqslant d+m-2$。

我们令 $Q={}^tP$，那么

$$\begin{aligned}Qf={}^tPf&=\sum_{|\alpha|\leqslant m}q_\alpha(x)\partial^\alpha f(x)\\&=\sum_{|\alpha|\leqslant m}(-1)^\alpha\partial^\alpha\big(p_\alpha(x)\cdot f(x)\big).\end{aligned}$$

利用归纳法，我们不难证明，对任意的多重指标 μ，我们有

$$e^{i\phi}\partial^m u\left(e^{-i\phi}\right)=\sum_{\substack{\mu_1+\cdots+\mu_l=\mu,\\|\mu_1|\geqslant1,\cdots,|\mu_l|\geqslant1}}(-1)^l\partial^{\mu_1}\phi\cdot\partial^{\mu_2}\phi\cdots\partial^{\mu_l}\phi.$$

所以，

$$\begin{aligned}e^{i\phi\,t}P\left(e^{-i\phi}c\right)&=e^{i\phi}\sum_{|\alpha|\leqslant m}q_\alpha(x)\partial^\alpha\left(e^{-i\phi}c\right)\\&=\sum_{|\alpha|\leqslant m}\sum_{\mu\leqslant\alpha}\binom{\alpha}{\mu}q_\alpha(x)e^{i\phi}\partial^\mu\left(e^{-i\phi}\right)\partial^{\alpha-\mu}c\\&=\sum_{|\alpha|\leqslant m}\sum_{\mu\leqslant\alpha}\sum_{\mu_1+\cdots+\mu_l=\mu,|\mu_j|\geqslant1}(-1)^l\binom{\alpha}{\mu}q_\alpha(x)\left(\partial^{\mu_1}\phi\cdots\partial^{\mu_l}\phi\right)\partial^{\alpha-\mu}c.\end{aligned}$$

在这个表达式中，每个 $\partial^{\mu_j}\phi(t,x,\xi)$ 对于 ξ 都是 1 次的，所以，上面所出现的关于 ξ 的最高阶项是 $d+m$ 阶的，最低阶项是 d 阶的。我们下面将对不同的阶数进行合并同类项。

- $d+m$ 阶项。

  此时，我们必须有 $l=|\mu|=|\alpha|=m$，所以，这些项为

$$\begin{aligned}&\sum_{|\alpha|=m}(-1)^m\sum_{\mu_1+\cdots+\mu_m=\alpha}q_\alpha(x)\left(\partial^{\mu_1}\phi\cdots\partial^{\mu_l}\phi\right)c\\=&\sum_{|\alpha|=m}(-1)^m q_\alpha(x)\left(\frac{\partial\phi}{\partial x}\right)^\alpha c(t,x,\xi)\\=&\sum_{|\alpha|=m}(-1)^m q_\alpha(x)\left(\frac{\partial\phi}{\partial x}\right)^\alpha c(t,x,\xi)\\=&p_m\left(x,\frac{\partial\phi}{\partial x}\right)\cdot c(t,x,\xi).\end{aligned}$$

- $d+m-1$ 阶项。

  此时，我们必须有 $l=m-1$。所以，$|\mu|\geqslant m-1$。

  - 如果 $|\mu|=m$，此时，$|\alpha|=m$，所以，这些项为

$$\begin{aligned}&\sum_{|\mu|=m}\sum_{\mu_1+\cdots+\mu_{m-1}=\mu}(-1)^{m-1}q_\mu(x)\left(\partial^{\mu_1}\phi\cdots\partial^{\mu_l}\phi\right)c\\=&M_{\phi,P}\cdot c(t,x,\xi).\end{aligned}$$

    这里，

$$M_{\phi,P}(t,x,\xi)=\sum_{|\mu|=m}\sum_{\mu_1+\cdots+\mu_{m-1}=\mu}(-1)^{m-1}q_\mu(x)\left(\partial^{\mu_1}\phi\cdots\partial^{\mu_l}\phi\right)$$

    对于 ξ 是一个次数为 $m-1$ 的齐次函数，它的构造只依赖于 ϕ 和 p_m（主象征!）。

  - 如果 $|\mu|=m-1$。我们定义多重指标 $\delta_j=(0,\cdots,0,1,0,\cdots,0)$，其中只有在第 j 个位置是 1。这些项为

$$\sum_{|\alpha|\leqslant m}\sum_{\mu\leqslant\alpha,|\mu|=m-1}(-1)^{m-1}\binom{\alpha}{\mu}q_\alpha(x)\left(\frac{\partial\phi}{\partial x}\right)^\mu\partial^{\alpha-\mu}c.$$

    如果 $|\alpha|=m-1$，这一类的项可以合并到上面的一项中（此时，我们要对 $M_{\phi,P}(t,x,\xi)$ 进行修正，它的构造变得依赖于 ϕ 和 P）。所以，除去这些项之外，此时没有处理的项形如

$$\begin{aligned}&\sum_{|\alpha|=m}\sum_{j=1}^n(-1)^{m-1}\binom{\alpha}{\alpha-\delta_j}q_\alpha(x)\left(\frac{\partial\phi}{\partial x}\right)^{\alpha-\delta_j}\partial^{\delta_j}c\\=&i\sum_{j=1}^n\frac{\partial p_m}{\partial\xi_j}\left(x,\frac{\partial\phi}{\partial x}\right)\frac{\partial c}{\partial x_j}.\end{aligned}$$

综上所述，我们就有

$$e^{i\phi}{}^tP\left(e^{-i\phi}c\right)=\underbrace{p_m\left(x,\frac{\partial\phi}{\partial x}\right)\cdot c}_{d+m\text{ 次}}+\underbrace{M_{\phi,P}\cdot c+i\sum_{j=1}^n\frac{\partial p_m}{\partial\xi_j}\left(x,\frac{\partial\phi}{\partial x}\right)\frac{\partial c}{\partial x_j}}_{d+m-1\text{ 次}}+c_{\leqslant d+m-2},$$

其中，$c_{\leqslant d+m-2}(t,x,\xi)$ 是一些次数不超过 $d+m-2$ 次的齐次函数的（有限）线性组合。

利用这个公式，我们就有

$$\begin{aligned}
&-e^{i\phi}\left(|\xi|^{m-1}i\frac{\partial}{\partial t}+{}^tP\right)\left(e^{-i\phi}c\right)\\
=&-|\xi|^{m-1}i\frac{\partial c}{\partial t}-|\xi|^{m-1}\frac{\partial\phi}{\partial t}c\\
&-p_m\left(x,\frac{\partial\phi}{\partial x}\right)\cdot c-M_{\phi,P}\cdot c-i\sum_{j=1}^{n}\frac{\partial p_m}{\partial\xi_j}\left(x,\frac{\partial\phi}{\partial x}\right)\frac{\partial c}{\partial x_j}+c_{\leqslant d+m-2}\\
=&\frac{1}{i}\left(|\xi|^{m-1}\partial_t+\sum_{j=1}^{n}\frac{\partial p_m}{\partial\xi_j}\left(x,\frac{\partial\phi}{\partial x}\right)\partial_{x_j}\right)c-\underbrace{\left(|\xi|^{m-1}\frac{\partial\phi}{\partial t}+p_m\left(x,\frac{\partial\phi}{\partial x}\right)\right)}_{=0,\ \text{根据 }\phi\text{ 的定义}}c\\
&-M_{\phi,P}\cdot c-c_{\leqslant d+m-2}.
\end{aligned}$$

所以，

$$-e^{i\phi}\left(|\xi|^{m-1}i\frac{\partial}{\partial t}+{}^tP\right)\left(e^{-i\phi}c\right)=\frac{1}{i}L(c)+M_{\phi,P}\cdot c+c_{\leqslant d+m-2},$$

其中，

$$L=|\xi|^{m-1}\partial_t+\sum_{j=1}^{n}\frac{\partial p_m}{\partial\xi_j}\partial_{x_j}.$$

作为总结，我们有

引理 563. 假设 $c(t,x,\xi)$ 是关于 ξ 为 d 次齐次的光滑函数，那么，我们有

$$-e^{i\phi}\left(|\xi|^{m-1}i\frac{\partial}{\partial t}+{}^tP\right)\left(e^{-i\phi}c\right)=\frac{1}{i}L(c)+\frac{1}{i}M_{\phi,P}\cdot c+c_{\leqslant d+m-2},$$

其中，$L(c)$ 是 $d+m-1$ 次的，$M_{\phi,P}$ 是次数为 $m-1$ 的齐次函数（只依赖于 ϕ 和 P）。

注记. 我们注意到 ϕ 的构造方式恰好消去了上面可能出现的 $d+m$ 次的项。

引理 564 (拟解的构造)**.** 对给定的正数 (r_0,α_0)，存在 $\rho_0>0$，使得 $\rho_0<r_0$ 并且如下成立：

对任意的 $\varphi\in C_0^\infty\left(B_{\rho_0}(x(0))\right)$，存在函数序列

$$\{b_k(t,x,\xi)\}_{k\geqslant 0},\ \{R_k(t,x,\xi)\}_{k\geqslant 0}\subset C^\infty\left([0,1]\times B(x(t),r_0)\times\Gamma(\xi(t),\alpha_0)\right),$$

- $b_k(t,x,\xi)$ 对于 ξ 是次数为 $-k$ 的齐次函数并且对任意的 t 和 ξ，如果 $x\notin B(x(t),r_0)$，那么 $b_k(t,x,\xi)=0$；
- $R_k(t,x,\xi)$ 是有限个次数不超过 $m-k-2$ 的齐次函数（对于 ξ）的和并且对任意的 t 和 ξ，如果 $x\notin B(x(t),r_0)$，那么 $R_k(t,x,\xi)=0$；

- 我们令

$$a_k(t,x,\xi)=\sum_{j=1}^{k}b_j(t,x,\xi).$$

那么，我们有

$$\begin{cases} -ie^{i\phi}|\xi|^{m-1}\dfrac{\partial}{\partial t}\left(e^{-i\phi}a_k\right)=e^{i\phi}\cdot{}^tP\left(e^{-i\phi}a_k\right)+R_k,\\ a_k(0,x,\xi)=\varphi(x). \end{cases}$$

证明： 我们归纳地来构造。先考虑 $k=0$ 的情形。此时，我们定义 $b_0(t,x,\xi)$ 为如下（常）微分方程的解：

$$\begin{cases} L(b_0)+M_{\phi,P}\cdot b_0=0,\\ b_0(0,x,\xi)=\varphi(x). \end{cases}$$

那么，b_0 对于 ξ 是 0 次的齐次函数（可以先对 $|\xi|=1$ 来解，然后进行齐次的扩张）。

假设对于 $\leqslant k$ 的指标 j，我们已经构造了 b_j，那么，我们考虑

$$\begin{aligned} &-ie^{i\phi}|\xi|^{m-1}\frac{\partial}{\partial t}\left(e^{-i\phi}\left(a_k+b_{k+1}\right)\right)-e^{i\phi}\cdot{}^tP\left(e^{-i\phi}\left(a_k+b_{k+1}\right)\right)\\ =&R_k+\left(L(b_{k+1})+M_{\phi,P}\cdot b_{k+1}\right)+c_{d+m-2}. \end{aligned}$$

其中 d 为 b_{k+1} 的次数。

由归纳假设，R_k 是次数不超过 $m-k-2$ 的齐次函数的线性组合，我们假设它的 $m-k-2$ 次分量为 $R_k^{(m-k-2)}$。那么，我们令

$$\begin{cases} L(b_{k+1})+M_{\phi,P}\cdot b_{k+1}=-R_k^{(m-k-2)},\\ b_{k+1}(0,x,\xi)=0. \end{cases}$$

那么，$d=-(k+1)$，从而 c_{d+m-2} 的次数不超过 $m-k-3$。这就完成了证明。 □

奇性传播定理的证明：完成

我们对方程

$$\frac{\partial}{\partial t}\left(e^{-i\phi}a_k\right)=i|\xi|^{1-m}\cdot{}^tP\left(e^{-i\phi}a_k\right)+i|\xi|^{1-m}e^{-i\phi}R_k$$

对 t 从 0 到 1 积分，由于 $\phi(0,x,\xi)=x\cdot\xi$，我们就得到

$$\begin{aligned} e^{-ix\cdot\xi}\varphi(x)=&e^{-i\phi(1,x,\xi)}a_k(1,x,\xi)-i|\xi|^{1-m}\int_0^1{}^tP\left(e^{-i\phi(t,x,\xi)}a_k(t,x,\xi)\right)dt\\ &-i|\xi|^{1-m}\int_0^1e^{-i\phi(t,x,\xi)}R_k(t,x,\xi)dt. \end{aligned}$$

所以，我们有

$$\begin{aligned}\widehat{\varphi\cdot u}(\xi)&=\langle u,e^{-ix\cdot\xi}\varphi(x)\rangle\\&=\left\langle u,e^{-i\phi(1,x,\xi)}a_k(1,x,\xi)\right\rangle-i|\xi|^{1-m}\int_0^1\left\langle u,\,{}^tP\left(e^{-i\phi(t,x,\xi)}a_k(t,x,\xi)\right)\right\rangle dt\\&\quad-i|\xi|^{1-m}\int_0^1\left\langle u,e^{-i\phi(t,x,\xi)}R_k(t,x,\xi)\right\rangle dt\\&=\underbrace{\left\langle u,e^{-i\phi(1,x,\xi)}a_k(1,x,\xi)\right\rangle}_{H(\xi)}-i|\xi|^{1-m}\int_0^1\underbrace{\left\langle Pu,e^{-i\phi(t,x,\xi)}a_k(t,x,\xi)\right\rangle}_{I(\xi,t)}dt\\&\quad-i|\xi|^{1-m}\int_0^1\underbrace{\left\langle u,e^{-i\phi(t,x,\xi)}R_k(t,x,\xi)\right\rangle}_{J(\xi,t)}dt.\end{aligned}$$

我们逐一地控制上面的三项：

- 控制 $H(\xi)$。

  首先，由于 $a_k(t,x,\xi)$ 对于 $x\notin B(x(t),r_0)$ 是消失的，所以，我们可以用一个示性函数 $\chi(x)$ 来记住这个事实（它在 $B(x(1),r_0)$ 上恒为 1）：

  $$\begin{aligned}H(\xi)&=\left\langle u,e^{-i\phi(1,x,\xi)}\chi(x)a_k(1,x,\xi)\right\rangle\\&=\int_{\mathbb{R}^n}e^{-i\phi(1,x,\xi)}a_k(1,x,\xi)\chi(x)u(x)dx.\end{aligned}$$

  我们选取相函数 $\phi(1,x,\xi)$，振幅函数 $a_k(1,x,\xi)$ 以及被作用函数 $f(x)=\chi\cdot u(x)$，我们来验证它们满足非驻相法引理的基本数据的条件：

  (a) $\phi(1,x,\xi)\in C^\infty\left(\mathbb{R}^n_x\times(\mathbb{R}^n_\xi-\{0\})\right)$ 是实值的，变量 ξ 是 1 次齐次的并且当 $\xi\neq0$ 时，有 $\nabla_x\phi(1,x,\xi)\neq0$：

  为了说明后者，我们用 $\phi(t,x,\xi)$ 满足的方程：

  $$\begin{cases}\dfrac{\partial\phi}{\partial s}(s,x,\xi)=-p_m\left(x,\nabla_x\phi(s,x,\xi)\right),\\\phi(0,x,\xi)=x\cdot\xi,\end{cases}$$

  其中 $|\xi|=1$。对 x 求导数，我们就有

  $$\frac{\partial^2\phi}{\partial s\partial x}(s,x,\xi)=-\frac{\partial p_m}{\partial x}(x,\nabla_x\phi(s,x,\xi))-\underbrace{\frac{\partial p_m}{\partial\xi}(x,\nabla_x\phi(s,x,\xi))}_{F(s,x,\xi)}\frac{\partial^2\phi}{\partial x^2}(s,x,\xi).$$

  所以，

  $$\left|\left(\frac{\partial}{\partial s}+F\frac{\partial}{\partial x}\right)\frac{\partial\phi}{\partial x}\right|=\left|-\frac{\partial p_m}{\partial x}(x,\nabla_x\phi(s,x,\xi))\right|\leqslant C\left|\frac{\partial\phi}{\partial x}\right|.$$

  如果 $\nabla_x\phi(1,x,\xi)=0$，上述微分不等式表明 $\nabla_x\phi(0,x,\xi)=\xi=0$，矛盾！

  上面的推理对于 $t\neq1$ 也成立。

(b) 按照构造，对于任意的 t，$b_j(t,x,\xi)$ 光滑，所以 $a_k(t,x,\xi)$ 也光滑。另外，$a_k(t,x,\xi)$ 是一些次数不超过 0 的齐次函数的和，利用紧性，使得对任意的多重指标 α，存在常数 C_α，我们都有

$$|\partial_x^\alpha a_k(t,x,\xi)| \leqslant C_\alpha.$$

(c) 由于 $(x(1),\xi(1)) \notin WF(u)$，对任意的 $N \geqslant 1$，存在常数 C_N，使得对任意的 $\xi \in \Gamma(\xi(1),\beta)$，都有我们就有

$$\left|\widehat{f}(\xi)\right| \leqslant \frac{C_N}{(1+|\xi|)^N}.$$

这里，$\eta_0 = \dfrac{\xi(1)}{|\xi(1)|}$。

根据引理561以及连续性，对任意的 $(x,\xi) \in B(x(1),r_0) \times \Gamma(\xi(1),\alpha_0)$，我们有

$$\left|\frac{\nabla_x\phi(x,\xi)}{|\nabla_x\phi(x,\xi)|} - \eta_0\right| < \beta.$$

所以我们可以运用非驻相法引理，从而，

此时，引理的结论表明，对任意的 $N \geqslant 1$，存在 $C_N > 0$，使得对任意的 $\Gamma(\xi(0),\alpha_0)$，有

$$|H(\xi)| \leqslant \frac{C_N}{(1+|\xi|)^N}.$$

- 控制 $I(\xi,t)$。

  这一部分的控制和 $H(\xi)$ 是完全一致的，唯一的不同在于把 u 换成了 Pu。另外，我们可以对 t 一致地选取 χ。重复上面的过程，我们就得到对任意的 $N \geqslant 1$，存在 $C_N > 0$，使得对任意的 $\Gamma(\xi(0),\alpha_0)$，有

$$|I(\xi,t)| \leqslant \frac{C_N}{(1+|\xi|)^N}.$$

- 控制 $J(\xi,t)$。

  利用分布的定义，存在常数 C_0，使得，

$$|J(\xi,t)| \leqslant C \sup_{\substack{|\alpha|\leqslant M,\\ x\in B_{r_0}(x(t))}} \left|\partial^\alpha\left(e^{-i\phi(t,x,\xi)}R_k(t,x,\xi)\right)\right|.$$

  $R_k(t,x,\xi)$ 是有限个 ξ 的次数 $\leqslant m-k-2$ 的齐次函数的和，所以，根据 Leibniz 法则（以及 ϕ 对 ξ 的次数为 1），上面右端出现的是一些次数不超过 $m-k-2+M$ 的齐次函数的和。利用 R_k 的光滑性，我们知道存在常数 C_m，使得

$$\sup_{\substack{|\alpha|\leqslant M,\\ (x,\xi)\in B_{r_0}(x(t))\times\mathbf{S}^{n-1}}} |\partial^\alpha R_k(t,x,\xi)| \leqslant C_M.$$

  所以，

$$|J(\xi,t)| \leqslant \frac{C'_m}{(1+|\xi|)^{k+2-M-m}},$$

  其中，k 是任意的正整数。

综合上面的所有估计，我们就证明了对任意的 $N \geqslant 1$，存在 $C_N > 0$，使得对任意的 $\Gamma(\xi(0), \alpha_0)$，有

$$|\widehat{\varphi \cdot u}(\xi)| \leqslant \frac{C_N}{(1+|\xi|)^N}.$$

这说明 $(x_0, \xi_0) \notin WF(u)$，从而完成了奇性传播定理的证明。

86 分布理论期末复习题

86.1 第一套

清华大学 19-20 春季学期，数学分析三，期末复习题一

请用 A4 大小的纸张**正反面**用钢笔，签字笔或者圆珠笔书写，并注明自己的姓名，年级和作业的**总页数**。除定理公式所涉及的人名之外，请使用**中文**。本次复习题的提交时间和地点为**12 月 31 日**的期末考试现场。
如下的四个题目是相互独立的。

题目 A. $u(x)$ 是在 $\mathbb{R}$ 上定义的以 1 为周期的函数。对于 $x \in [0,1]$，我们有

$$u(x) = e^{2\pi(x-\frac{1}{2})} + e^{-2\pi(x-\frac{1}{2})}$$

A1) 证明，

$$u(x) = \frac{e^{\pi} - e^{-\pi}}{\pi} \sum_{k=-\infty}^{+\infty} \frac{e^{2\pi i k x}}{1+k^2}$$

A2) 证明，作为 $\mathbb{R}$ 上的分布，我们有

$$u'' - 4\pi^2 u = -8\pi \sinh(\pi) \sum_{k\in\mathbb{Z}} \delta_k.$$

A3) 证明分布形式的 Poisson 求和公式：

$$\sum_{k\in\mathbb{Z}} e^{2\pi i k x} = \sum_{k\in\mathbb{Z}} \delta_k.$$

题目 B. 我们来构造微分算子

$$u(x) \mapsto u''(x) - u(x)$$

的基本解。

第一部分：在 $\mathcal{D}'(\mathbb{R})$ 中解 $u'' - u = \delta$

B1) 证明，对任意的 $a \in \mathbb{C}$，试给出方程

$$u' - au = 0$$

的所有解，其中，$u \in \mathcal{D}'(\mathbb{R})$。

B2) 对 $a \in \mathbb{C}$，我们定义分布

$$u_a(x) = e^{ax} H(x),$$

其中 $H(x)$ 是 Heaviside 函数。试计算

$$(\delta' - a\delta) * u_a,$$

其中，$\delta = \delta_0$ 是 Dirac 分布。

B3) 在 $\mathcal{D}'(\mathbb{R})$ 中解方程

$$u' - au = \delta,$$

其中 $a \in \mathbb{C}$。

B4) 在 $\mathcal{D}'(\mathbb{R})$ 中解方程

$$u'' - u = \delta.$$

B5) 在 $\mathcal{S}'(\mathbb{R})$ 中解方程

$$u'' - u = \delta.$$

第二部分：在 $\mathcal{S}'(\mathbb{R})$ 中解 $u'' - u = \delta$

B6) 证明，

$$\int_{\mathbb{R}} \frac{e^{itx}}{x^2+1} dx = \pi e^{-|t|}.$$

B7) 利用上面的积分在 $\mathcal{S}'(\mathbb{R})$ 中解方程

$$u'' - u = \delta.$$

B8) $\mathcal{D}'(\mathbb{R}_{>0})$ 中解方程

$$u'' - u = \delta.$$

注记. 我们可以把 $\mathcal{S}'(\mathbb{R})$ 看作是在无穷远处增长有限制的分布，这个问题很好的展示了边界条件对于方程的解的约束。

题目 C. 给定具有紧支集的分布 $u \in \mathcal{E}'(\mathbb{R})$，我们用 $\widehat{u}$ 表示它的 Fourier 变换（作为缓增分布）。对任意的 $z \in \mathbb{C}$，我们定义函数

$$\mathcal{L}(u)(z) := \langle u, e^{-izx} \rangle.$$

C1) 证明，$\mathcal{L}(u)(z)$ 是 $\mathbb{C}$ 上的复解析函数。

C2) 证明，如果 $\xi \in \mathbb{R}$，那么，

$$\mathcal{L}(u)(\xi) = \widehat{u}(\xi).$$

C3) 证明，如果对任意的 $k \geqslant 0$，我们有

$$\langle u, x^k \rangle = 0,$$

那么，$u = 0$。

C4) 如果 p 是分布 u 的阶，证明，存在常数 C，使得对任意的 $\xi \in \mathbb{R}$，我们都有

$$|\widehat{u}(\xi)| \leqslant C(1 + |\xi|)^p.$$

C5) 我们假设 $z \in \mathbb{H}$，其中，$\mathbb{H} = \{z \in \mathbb{C} \big| \mathrm{Im}(z) > 0\}$。证明，

$$U(z) := \int_0^\infty \widehat{u}(\xi) e^{i\xi z} d\xi$$

是良好定义的并且是 $\mathbb{H}$ 上的复解析函数。

C6) 证明，存在常数 $C > 0$，$\varepsilon > 0$ 和正整数 N，使得当 $0 < \mathrm{Im}(z) < \varepsilon$ 时，我们有

$$|U(z)| \leqslant C \cdot \mathrm{Im}(z)^{-N}.$$

C7) 对任意的 $\varepsilon > 0$，我们定义分布 $\mathbb{R}$ 上的分布：

$$u_\varepsilon(x) = U(x + i\varepsilon).$$

证明，在 $\mathcal{D}'(\mathbb{R})$ 中下面的极限存在：

$$\lim_{\varepsilon \to 0^+} u_\varepsilon.$$

C8) 对于 $z \in \mathbb{H}$，我们定义

$$V(z) := \int_\infty^0 \widehat{u}(\xi) e^{i\xi z} d\xi$$

并令

$$v_\varepsilon(x) = V(x - i\varepsilon).$$

证明，在 $\mathcal{D}'(\mathbb{R})$ 中，我们有

$$\frac{1}{2\pi} \lim_{\varepsilon \to 0^+} (u_\varepsilon + v_\varepsilon) = u.$$

题目 D.

对任意的 $\varphi(x, y) \in \mathcal{D}(\mathbb{R}^2)$，我们定义如下的线性形式：

$$\langle u, \varphi \rangle = \int_0^\infty \varphi(t, 2t) dt.$$

D1) 证明，$u \in \mathcal{D}'(\mathbb{R}^2)$ 并且确定它的阶和支集。

D2) u 是否落在自然映射

$$\iota : L^1_{\mathrm{loc}}(\mathbb{R}^2) \to \mathcal{D}'(\mathbb{R}^2)$$

的像里?

D3) 在分布的意义下计算

$$\frac{\partial u}{\partial x} + 2\frac{\partial u}{\partial y}.$$

D4) 证明，$u \in \mathcal{S}'(\mathbb{R}^2)$。

D5) $\widehat{u}(\xi)$ 是否落在自然映射

$$\iota : L^1_{\mathrm{loc}}(\mathbb{R}^2) \to \mathcal{D}'(\mathbb{R}^2)$$

的像里?

D6) 给定 $R > 0$，证明，如下的线性形式

$$\langle u_R, \varphi \rangle = \int_0^R \varphi(t, 2t) dt$$

定义了一个缓增的分布并且在缓增分布的意义下

$$\lim_{R \to +\infty} u_R \overset{\mathcal{S}'(\mathbb{R}^2)}{=} u.$$

D6) 计算 Fourier 变换 $\widehat{u_R}(\xi)$ （提示：考虑算子 $\dfrac{\partial}{\partial x} + 2\dfrac{\partial}{\partial y}$）

D7) 证明，对任意的 $s \in \mathbb{R}$，$u \notin H^s(\mathbb{R}^2)$。

D8) 给定 $R > 0$。试找出所有的 s，使得 $u_R \in H^s(\mathbb{R}^2)$。

86.2 第二套

清华大学 19-20 春季学期，数学分析三，期末复习题二

请用 A4 大小的纸张**正反面**用钢笔，签字笔或者圆珠笔书写，并注明自己的姓名，年级和作业的**总页数**。除定理公式所涉及的人名之外，请使用**中文**。本次复习题的提交时间和地点为**12 月 31 日**的期末考试现场。
如下的三个题目是相互独立的。

题目 A. 给定 $a \in \mathbb{R}$，$\lambda \in \mathbb{C}$ 并且 $\mathrm{Re}(\lambda) \leqslant 0$。我们定义 $\mathbb{R}$ 上的分布

$$H_a(x) = \mathbf{1}_{[a,+\infty)}(x),$$
$$u(x) = e^{\lambda x} H_a(x).$$

A1) 在分布的意义下计算

$$u' - \lambda u.$$

A2) 证明，$u \in \mathcal{S}'(\mathbb{R})$ 并且

$$(i\xi - \lambda)\widehat{u}(\xi) = e^{a(\lambda - i\xi)}.$$

A3) 如果 $\mathrm{Re}(\lambda) < 0$，试确定 $\widehat{u}(\xi)$。

A4) 当 $\lambda = 0$ 时，计算 $\widehat{u}(\xi)$。

A5) 当 $\lambda = i\mu$ 时，其中 $\mu \in \mathbb{R}$，计算 $\widehat{u}(\xi)$。

题目 B. 对任意的 $\varphi(x,y) \in \mathcal{D}(\mathbb{R}^2)$，我们定义如下的线性映射：

$$\langle u, \varphi \rangle = \int_0^\infty \varphi(t,t)dt.$$

我们用 L 表示如下的微分算子：

$$L = \frac{\partial}{\partial x} + \frac{\partial}{\partial y}.$$

B1) 证明，$u \in \mathcal{D}'(\mathbb{R}^2)$ 并且在分布的意义下计算 $L(u)$。

B2) 假设 $g \in C^1(\mathbb{R})$，试计算

$$u * (g \otimes \delta),$$

其中 δ 是 $\mathbb{R}$ 上的 Dirac 分布。

B3) 令

$$\overline{\mathbb{H}} = \big\{(x,y) \in \mathbb{R}^2 \big| y \geqslant 0\big\}.$$

试找出所有的 $f \in C^1\left(\overline{\mathbb{H}}\right)$，使得

$$\begin{cases} L(f) & = 0, \\ f(x,0) & = g(x), \end{cases}$$

其中 g 是上一个问题中给定的函数。

题目 C. 固定一个正实数 $\lambda > 0$。我们用如下的方式定义实值函数

$$g : \left[0, \frac{\pi}{4}\right) \cup \left(\frac{\pi}{4}, \pi\right] \to \mathbb{R},$$

其中，

$$\tan\left(g(x)\right) = \frac{\lambda}{1 - \tan(x)},$$

并且我们要求

$$g(x) \in \begin{cases} \left[-\frac{3\pi}{2}, -\frac{\pi}{2}\right], & x \in \left[0, \frac{\pi}{4}\right); \\ \left[-\frac{\pi}{2}, \frac{\pi}{2}\right], & x \in \left(\frac{\pi}{4}, \pi\right]. \end{cases}$$

C1) 证明，$g(x)$ 可以唯一地被拓成 $[0, \pi]$ 上的连续函数并计算 $g(0)$ 和 $g(\pi)$。

C2) 定义 $[0, \pi]$ 上的函数

$$h(x) = \frac{\cos(x) + (1 + i\lambda)\sin(x)}{\sin(x) - (1 + i\lambda)\cos(x)},$$
$$f(x) = \frac{1}{2}\log\left(\left(\sin(x) - \cos(x)\right)^2 + \lambda^2\cos(x)^2\right) + ig(x).$$

证明，

$$f' = h.$$

C3) 计算积分

$$\int_0^\pi h(x)dx.$$

C4) 定义 $\mathbb{R}^2$ 上的微分算子：

$$A = \frac{\partial}{\partial x} + (1 + i\lambda)\frac{\partial}{\partial y}.$$

试构造 A 的一个基本解。

C5) 证明，如果 $u \in \mathcal{D}'(\mathbb{R}^2)$（在分布意义下）满足微分方程：

$$A(u) = 0,$$

那么，$u \in C^\infty(\mathbb{R}^2)$。

86.3 第三套

清华大学 19-20 春季学期，数学分析三，期末复习题三

请用 A4 大小的纸张**正反面**用钢笔，签字笔或者圆珠笔书写，并注明自己的姓名，年级和作业的**总页数**。除定理公式所涉及的人名之外，请使用**中文**。本次复习题的提交时间和地点为**12 月 31 日**的期末考试现场。

如下的两个题目是相互独立的。

题目 A. 我们定义 $\mathbb{R}$ 上的分布

$$u(x) = e^{-|x|}.$$

A1) 证明，u 是缓增的分布并且在分布的意义下有

$$u'' - u + 2\delta_0 = 0.$$

A2) 证明，计算 $\widehat{u}$ 并且确定所有的 s，使得 $u \in H^s(\mathbb{R})$。

A3) 任意给定多项式 $P \in \mathbb{C}[x]$，证明，对任意的 $U \in \mathcal{S}'(\mathbb{R})$，在频率空间上 $\widehat{P \cdot u} \cdot \widehat{U}$ 是良好定义的缓增分布。据此，我们可以定义卷积

$$(P \cdot u) * U := \mathcal{F}^{-1}\left(\widehat{P \cdot u} \cdot \widehat{U}\right).$$

A4) 证明，对任意的 $U \in H^s(\mathbb{R})$，其中 $s \in \mathbb{R}$，我们有

$$(P \cdot u) * U \in H^{s+2}(\mathbb{R}).$$

A5) 证明，如果 $U \in L^2(\mathbb{R})$，那么

$$(P \cdot u) * U \in C^1(\mathbb{R}).$$

A6) 我们现在要绕过 Fourier 变换直接来定义卷积 $(P \cdot u) * U$。证明，存在 $v \in \mathcal{E}'(\mathbb{R})$ 和 $\varphi \in \mathcal{S}(\mathbb{R})$，使得

$$P \cdot u = v + \varphi.$$

A7) 我们定义

$$(P \cdot u) \bullet U = v * U + \varphi * U.$$

证明，这个定义所给出的分布不依赖于 A6) 中分解的选取。

A8) 证明，$(P \cdot u) \bullet U$ 在 $\operatorname{supp}(U)$ 之外是光滑函数。

A9) 证明，上面的两种卷积的定义是一致的。

A10) 考虑如下的缓增分布：

$$U = \sum_{k\in\mathbb{Z}} \delta_k,$$

证明，$T * U$ 是 $\mathbb{R}$ 上的连续周期函数并在一个周期上用初等函数表示它。

题目 B. 我们先引入如下的记号：

- A 是给定的 3×3 的实矩阵并且 $\det(A) = 1$。我们把 A 视作是线性变换（矩阵从左边去乘以一个列向量）：

$$A : \mathbb{R}^3 \to \mathbb{R}^3.$$

对任意的 $u \in \mathcal{D}'(\mathbb{R}^3)$，我们把 Au 定义为如下的分布：

$$\langle Au, \varphi\rangle = \langle u, \varphi \circ A^{-1}\rangle.$$

特别地，如果 $u \in L^1_{\mathrm{loc}}(\mathbb{R}^3)$，那么，

$$(Au)(x) = u(A(x)).$$

对于给定的 $\theta, t \in \mathbb{R}$，我们定义两类特殊的线性变换：

$$R_\theta : \mathbb{R}^3 \to \mathbb{R}^3,\ \ (x, y, z) \mapsto (x\cos(\theta) + y\sin(\theta), -x\sin(\theta) + y\cos(\theta), z),$$

和

$$S_t : \mathbb{R}^3 \to \mathbb{R}^3,\ \ (x, y, z) \mapsto (x\cosh(t) + z\sinh(t), y, z\cosh(t) + x\sinh(t)).$$

如果对任意的 $\theta, t \in \mathbb{R}$，我们都有

$$R_\theta u = u,\ \ S_t u = u,$$

我们就说 u 是**不变的**。

- 我们定义 $\mathbb{R}^3$ 上的微分算子：

$$\Box = \frac{\partial^2}{\partial z^2} - \frac{\partial^2}{\partial x^2} - \frac{\partial^2}{\partial y^2}.$$

我们定义三个变量的多项式

$$P(x, y, z) = z^2 - x^2 - y^2.$$

- 对于实数 $\alpha \neq 0$，我们定义 $\mathbb{R}^3$ 中的双曲面

$$H_\alpha = \big\{(x, y, z) \in \mathbb{R}^3 \big| z^2 - x^2 - y^2 = \alpha.\big\}$$

当 $\alpha > 0$ 时，我们定义

$$H^+_\alpha = \big\{(x, y, z) \in H_\alpha \big| z > 0\big\},\ \ H^-_\alpha = \big\{(x, y, z) \in H_\alpha \big| z < 0\big\}.$$

当 $\alpha = 0$ 时，我们定义如下的（无定点的锥）

$$H_0 = \big\{(x,y,z) \in \mathbb{R}^3 \big| z^2 - x^2 - y^2 = 0, (x,y,z) \neq (0,0,0).\big\}$$

和

$$H_0^+ = \big\{(x,y,z) \in H_0 \big| z > 0\big\}, \quad H_0^- = \big\{(x,y,z) \in H_0 \big| z < 0\big\}.$$

- 我们用 $\vec{r} = (x,y,z)$ 表示 $\mathbb{R}^3$ 中的一个向量（点），我们用如下的记号表示特定的向量：

$$\vec{e}_x = (1,0,0),\ \vec{e}_y = (0,1,0),\ \vec{e}_z = (0,0,1).$$

我们用 r 和 ρ 表示如下定义的 $\mathbb{R}^3$ 上的半径函数

$$r = \sqrt{x^2+y^2+z^2}, \quad \rho = \sqrt{x^2+y^2}.$$

当 $\alpha \neq 0$ 时，我们令

$$a = \sqrt{|\alpha|}.$$

第一部分

B1) 假设 $T \in \mathcal{D}'(\mathbb{R}^3)$ 并且在分布的意义下满足

$$z \cdot T = 0.$$

证明，存在分布 $S \in \mathcal{D}'(\mathbb{R}^2)$，使得对任意的 $\varphi(x,y,z) \in \mathcal{D}(\mathbb{R}^3)$，如果令

$$\psi(x,y) = \varphi(x,y,0),$$

那么

$$\langle T, \varphi\rangle = \langle S, \psi\rangle.$$

B2) 给定开集 $\omega \subset \mathbb{R}^2$ 和实值的光滑函数 $F \in C^\infty(\omega)$，我们用 Σ 表示 F 的图像所定义的曲面：

$$\Sigma = \big\{(x,y,z) \in \omega \times \mathbb{R} \big| z = F(x,y), (x,y) \in \omega\big\}.$$

假设 $T \in \mathcal{D}'(\omega \times \mathbb{R})$ 满足分布意义下的方程

$$\big(z - F(x,y)\big) \cdot T = 0,$$

证明，存在分布 $S \in \mathcal{D}'(\mathbb{R}^2)$，使得对任意的 $\varphi(x,y,z) \in \mathcal{D}(\omega \times \mathbb{R})$，如果令

$$\psi(x,y) = \varphi\big(x,y,F(x,y)\big),$$

那么

$$\langle T, \varphi\rangle = \langle S, \psi\rangle.$$

B3) 假设 $T \in \mathcal{D}'(\mathbb{R}^3)$ 是不变的分布。如果对实数 $\alpha > 0$，T 满足如下分布意义下的等式

$$(P - \alpha) \cdot T = 0,$$

我们就称 T 是 α-**不变的**。

证明，如果 T 是 α-不变的 ($\alpha > 0$)，那么，$\mathrm{supp}(T) \subset H_\alpha$。进一步证明，如果 $\mathrm{supp}(T) \subset H_\alpha^+$，那么，存在分布 $S \in \mathcal{D}'(\mathbb{R}^2)$，使得对任意的 $\varphi(x, y, z) \in \mathcal{D}(\mathbb{R}^3)$，我们都有

$$\langle T, \varphi \rangle = \left\langle S, \varphi(x, y, \sqrt{x^2 + y^2 + a^2}) \right\rangle.$$

我们把

$$\psi(x, y) = \varphi(x, y, \sqrt{x^2 + y^2 + a^2})$$

称作是 φ 的**伴随函数**，把 S 称作是 T 的**伴随分布**。

B4) 对任意的 $t, t' \in \mathbb{R}$，证明 Lorentz 变换满足如下的性质：

$$S_{t+t'} = S_t \circ S_{t'}.$$

（注意，不要把 S_t 和上面的伴随分布的符号弄混）

对任意的 $\varphi(x, y, z) \in \mathcal{D}(\mathbb{R}^3)$，我们令

$$\varphi^t = S_t \varphi = \varphi \circ S_t$$

试用 φ 伴随函数 $\psi(x, y)$ 的导数来计算

$$\frac{d}{dt}\Big|_{t=0} \psi^t(x, y),$$

其中，$\psi^t(x, y)$ 是 $\varphi^t(x, y, z)$ 的伴随函数。

B5) 假设 $T \in \mathcal{D}'(\mathbb{R}^3)$ 是 α-不变的（$\alpha > 0$）并且 $\mathrm{supp}(T) \subset H_\alpha^+$，$S \in \mathcal{D}'(\mathbb{R}^2)$ 是它的伴随分布。证明，S 满足如下两个条件

- S 是在 $\mathbb{R}^2$ 上旋转不变的；
- 在分布的意义下，我们有

$$\frac{\partial}{\partial x}\left(\sqrt{a^2 + x^2 + y^2} \cdot S\right) = 0.$$

B6) 如果 $\mathrm{supp}(T) \subset H_\alpha^+$ 并且 S 是它的伴随分布。证明，

- 如果 S 是在 $\mathbb{R}^2$ 上旋转不变的，那么，对任意的 $\theta \in \mathbb{R}$，我们有

$$R_\theta T = T.$$

– 如果在分布的意义下，我们有

$$\frac{\partial}{\partial x}\left(\sqrt{a^2+x^2+y^2}\cdot S\right)=0,$$

那么，对任意的 $t\in\mathbb{R}$，我们有

$$S_tT=T.$$

B7) 假设 $\alpha>0$，我们用 $d\sigma_\alpha^+$ 表示 H_α^+ 的曲面测度，证明，

$$\langle d\sigma_\alpha^+,\varphi\rangle=\int_{\mathbb{R}^2}\varphi(x,y,\sqrt{x^2+y^2+a^2})\frac{1}{|\vec{n}\cdot\vec{e}_z|}dxdy,$$

其中，$\vec{n}$ 是 H_α^+ 在 (x,y,z) 处的单位法向量，$\vec{n}\cdot\vec{e}_z$ 是两个向量在 $\mathbb{R}^3$ 中的标准内积。

B8) 试计算 $d\sigma_\alpha^+$ 的伴随分布并证明如下定义的分布是 α-不变的：

$$T_\alpha^+=\frac{1}{r}d\sigma_\alpha^+.$$

B9) 假设 $S\in\mathcal{D}'(\mathbb{R}^2)$ 是旋转不变的并且满足

$$\frac{\partial S}{\partial x}=0.$$

证明，S 是由某个常数值的函数所定义的分布。

B10) 证明，$T\in\mathcal{D}'(\mathbb{R}^3)$ 是 α-不变的分布并且 $\mathrm{supp}(T)\subset H_\alpha^+$，那么，存在 $C\in\mathbb{C}$，使得

$$T=C\cdot T_\alpha^+.$$

第二部分

B11) 证明，T_α^+ 是缓增的分布。

B12) 缓增的分布 $T\in\mathcal{S}'(\mathbb{R}^3)$ 是不变的当且仅当其 Fourier 变换是不变的。

B13) 试找出所有的不变的缓增分布 $U\in\mathcal{S}'(\mathbb{R}^3)$，使得 U 满足 Klein-Gordon 方程：

$$(\Box-\lambda)U=0,$$

其中，$\lambda<0$ 是给定的负数。

B14) 我们将换一种方式将 B13) 中的结论推广到 $\lambda=0$ 的情况（不需要重复第一部分中的计算）。我们用 $d\sigma^+$ 表示 H_0^+ 的曲面测度，用 $d\sigma^-$ 表示 H_0^- 的曲面测度，那么，

$$T_0^+=\frac{1}{r}d\sigma_+,\quad T_0^-=\frac{1}{r}d\sigma_-$$

都是 $\mathbb{R}^3-\{(0,0,0)\}$ 上的分布。证明，如下的曲面积分对任意的 $\varphi\in\mathcal{D}(\mathbb{R}^3)$ 都是收敛的

$$\int_{H_0^+}\frac{\varphi}{r}d\sigma^+,\quad\int_{H_0^-}\frac{\varphi}{r}d\sigma^-$$

并且定义出了 $\mathbb{R}^3$ 上的两个缓增的分布。我们之后仍旧把它们记作 T_0^+ 和 T_0^-。

B15) 假设 $T \in \mathcal{D}'(\mathbb{R}^3)$ 是 0-不变的（即 $P \cdot T = 0$）分布。证明，T 是 T_0^+，T_0^- 和 $\delta_{(0,0,0)}$ 的线性组合。

B16) 试找出所有的不变的缓增分布 $U \in \mathcal{S}'(\mathbb{R}^3)$，使得 U 满足波动方程：

$$\Box U = 0.$$

第三部分

我们在 $\mathbb{R}^3$ 上定义开集

$$\Omega_1 = \big\{(x,y,z) \in \mathbb{R}^3 \big| z^2 - x^2 - y^2 > 0, z > 0\big\}.$$

我们在 Ω_1 上用如下的极坐标系：

$$\begin{cases} x &= a\sinh(u)\cos(\beta), \\ y &= a\sinh(u)\sin(\beta), \quad (a,u,\beta) \in \mathbb{R}_{>0} \times \mathbb{R} \times \mathbb{R}. \\ z &= a\cosh(u). \end{cases}$$

对任意的 $\varphi \in C_0^\infty(\Omega_1)$，我们定义函数 $\chi_\varphi(a) \in C_0^\infty\big((0,+\infty)\big)$，使得

$$\chi_\varphi(a) = a\int_{H_\alpha^+} \varphi \frac{d\sigma_\alpha^+}{r}, \quad \alpha = a^2, a > 0.$$

B17) 假设 $T \in \mathcal{D}'(\Omega_1)$ 是分布，证明，存在分布 $S_T \in \mathcal{D}'\big((0,+\infty)\big)$，使得对任意的 $\varphi \in C_0^\infty(\Omega_1)$，我们都有

$$\langle T, \varphi\rangle = \langle S_T, \chi_\varphi\rangle.$$

B18) 假设 $T = F \in C^0(\Omega_1)$ 是不变的连续函数。证明，S_T 可以被如下的函数实现：对任意的 $a > 0$，

$$S_T(a) = F(\vec{r}), \quad \vec{r} = (x,y,z),$$

其中，$\vec{r}$ 是任意的一个 $H_{a^2}^+$ 上的点。

B19) 我们用 $\mathcal{D}'(\Omega_1)^{\mathrm{inv}}$ 表示 Ω_1 中所有不变的分布，用 $C^\infty(\Omega_1)^{\mathrm{inv}}$ 表示 Ω_1 中所有不变的光滑函数。证明，

$$C^\infty(\Omega_1)^{\mathrm{inv}} \subset \mathcal{D}'(\Omega_1)^{\mathrm{inv}}$$

是稠密的。

B20) 如果 $T \in \mathcal{D}'(\Omega_1)^{\mathrm{inv}}$，证明，$\Box T \in \mathcal{D}'(\Omega_1)^{\mathrm{inv}}$ 并且

$$S_{\Box T} = \frac{\partial^2}{\partial a^2} S_T + \frac{2}{a}\frac{\partial}{\partial a} S_T.$$

（提示：令 $a = \sqrt{z^2 - y^2 - x^2}$，其中 $(x,y,z) \in \Omega_1$，假设 $F(x,y,z) = f(a) \in C^\infty(\Omega_1)^{\mathrm{inv}}$，先计算 $\Box F$）

B21) 给定 $T \in \mathcal{D}'(\Omega_1)^{\mathrm{inv}}$，假设存在 $\lambda \in \mathbb{C}$，使得 T 满足如下的 Klein-Gordon 方程：

$$\Box T = \lambda T.$$

证明，如果 $\lambda \neq 0$，令 ℓ 与 $-\ell$ 为 λ 的两个平方根，那么，$T \in C^\infty(\Omega_1)^{\mathrm{inv}}$ 并且有如下的公式给出：

$$T(x,y,z) = \frac{1}{a}\left(C_+ e^{\ell a} + C_- e^{-\ell a}\right),$$

其中 $C_\pm$ 为常数并且 $a = \sqrt{z^2 - x^2 - y^2}$。

当 $\lambda = 0$ 时，T 的表达式是什么?

B22) 接 B21)，假设 $\lambda \neq 0$。我们将 T 用 0 在 $\mathbb{R}^3 - \Omega_1$ 上进行延拓得到 $F_T \in \mathcal{D}'(\mathbb{R}^3)$。证明，$F_T \in L^1_{\mathrm{loc}}(\mathbb{R}^3)$。试确定 $\lambda \neq 0$，C_+ 和 C_- 的值，使得 $F_T \in \mathcal{S}'(\mathbb{R}^3)$。

B23) 我们在 $\mathbb{R}^3$ 上定义开集

$$\Omega_2 = \left\{(x,y,z) \in \mathbb{R}^3 \middle| z^2 - x^2 - y^2 > 0, z < 0\right\},$$

和

$$\Omega_3 = \left\{(x,y,z) \in \mathbb{R}^3 \middle| z^2 - x^2 - y^2 < 0\right\}.$$

以上关于 S_T 的理论显然可以推广到 Ω_2 上。

证明，对任意的 $T \in \mathcal{D}'(\Omega_3)^{\mathrm{inv}}$，存在分布 $S_T \in \mathcal{D}'\big((0,+\infty)\big)$，使得对任意的 $\varphi \in C_0^\infty(\Omega_3)$，

$$\langle T, \varphi\rangle = \langle S_T, \chi_\varphi\rangle,$$

其中

$$\chi_\varphi(a) = a\int_{H_\alpha} \varphi \frac{d\sigma_\alpha}{r}, \quad \alpha = -a^2, a > 0,$$

这里 $d\sigma_\alpha$ 是 H_α 的曲面测度。

B24) 假设 $T = F \in C^0(\Omega_3)$ 是不变的连续函数。证明，S_T 可以被如下的函数实现：对任意的 $a > 0$，

$$S_T(a) = F(\vec{r}), \quad \vec{r} = (x,y,z),$$

其中，$\vec{r}$ 是任意的一个 H_{-a^2} 上的点。

B25) 假设 $T \in \mathcal{D}'(\mathbb{R}^3)^{\mathrm{inv}}$ 并且满足 Klein-Gordon 方程：

$$\Box T = \lambda T.$$

令 $\Omega = \Omega_1 \cup \Omega_2 \cup \Omega_3$。证明，存在函数 $F \in L^1_{\mathrm{loc}}(\mathbb{R}^3)$，使得 $F \in C^\infty(\Omega)$ 并且

$$T\big|_\Omega \overset{\mathcal{D}'}{=} F\big|_\Omega.$$

B26) 我们在 $\mathbb{R}^3$ 上定义开集

$$\Omega_{13}=\left\{(x,y,z)\in\mathbb{R}^3\middle|z>\sqrt{x^2+y^2}\right\}=\Omega_1\cup H_0^+\cup\Omega_3.$$

证明，对任意的 $T\in\mathcal{D}'(\Omega_{13})^{\mathrm{inv}}$，存在分布 $v_T\in\mathcal{D}'\big(\mathbb{R}\big)$，使得对任意的 $\varphi\in C_0^\infty(\Omega_{13})$，

$$\langle T,\varphi\rangle=\langle v_T,\psi_\varphi\rangle,$$

其中

$$\psi_\varphi(\alpha)=\frac{1}{2}\int_{H_\alpha}\varphi\frac{d\sigma_\alpha}{r},\quad\alpha\in\mathbb{R}.$$

如果 $T=f(\alpha)$，其中，f 为连续函数，$\alpha=z^2-x^2-y^2$，证明，$v_T=f(\alpha)$。

B27) 假设 $T\in\mathcal{D}'(\mathbb{R}^3)^{\mathrm{inv}}$ 并且满足 Klein-Gordon 方程：

$$\Box T=\lambda T.$$

证明，B25) 中所构造的 F 在 Ω_{13} 上的限制形如 $f(\alpha)$，其中，$\alpha=z^2-x^2-y^2$，$f\in L^1_{\mathrm{loc}}(\mathbb{R})$。进一步，证明 $v_T=f(\alpha)$ 并且满足如下的微分方程

$$4\alpha v_T''+6v_T'-\lambda v_T=0.$$

B28) 给定分布 $v\in\mathcal{D}'(\mathbb{R})$，如果对任意的试验函数 $\varphi\in\mathcal{D}(\mathbb{R})$，我们都有

$$\lim_{\varepsilon\to0}\left\langle v,\varphi\left(\frac{x}{\varepsilon}\right)\right\rangle=0,$$

我们就称 v 在 0 处是**正则的**。如果对于自然数 $n\geqslant1$，$x^n\cdot v$ 在 0 处正则但是 $x^{n-1}\cdot v$ 在 0 处不正则，我们就说 v 的正则次数为 n。如果 v 在 0 处正则，我们就说 v 的正则次数为 0。证明如下四个论断：

- 如果 $v\in L^1_{\mathrm{loc}}(\mathbb{R})$，那么，$v$ 的正则次数为 0。
- 如果 $v\in\mathcal{D}'(\mathbb{R})$ 的正则次数为 n，其中，$n\geqslant1$，那么，$x\cdot v$ 的正则次数为 $n-1$。
- 如果 $v\in\mathcal{D}'(\mathbb{R})$ 的正则次数不超过 n，其中 $n\geqslant0$，那么，v' 的正则次数不超过 $n+1$。
- Dirac 分布的 k 次导数 $\delta_0^{(k)}$ 的正则次数为 $k+1$。

B29) 假设 T 与 F 满足 B26) 与 B27)，证明，

$$w=v_T-v_F$$

是 δ_0 及其导数的线性组合。

B30) 证明，$w=0$。（提示：考虑 $4\alpha w''+6w'-\lambda w$）

B31) 证明，$T\big|_{\mathbb{R}^3-\{0\}}=F\big|_{\mathbb{R}^3-\{0\}}$。