

引力

G R A V I T Y

夏草

引力

G R A V I T Y

夏草



CC0 license
NOT PUBLISHED

也许你不仅会欣赏这种文化，
甚至还想加入这场早已启航的人类心智最伟大的冒险。

——Richard P. Feynman

前言

引力是人尽皆知的吸引作用。我们自蹒跚学步起便与之相伴，身边总是随时随地存在着这种力量，没有支撑的物体因它而下坠，有支撑的物体因它而稳定。即便如此生活中仍有许多驾驭引力的高手，比如轻松掷中篮筐的职业球员；或者挥拍自如的羽毛球手，当然，他们大抵是不需要计算什么物理学的；再如鸟儿，它们和羽毛球手都得借助气流，但更胜一筹的是，鸟儿天生长有一对契合空气动力学的翼膀，从而高耸入云、自由翱翔。但我们知道，气流在大范围上的无扰压强分布仍需借助引力而稳定，因此结论或许稍显意外：引力虽可能使鸟儿丧命，但也正是引力使得鸟儿脱身。

这皆拜引力所赐，它如此普遍，世界各族文明自古以来便在思索其缘由，并给出不同色彩的描述。面对落体运动，中国古代将其解释为物体趋于回归至它原本的地方。受哲学及逻辑学的启蒙，西方世界对引力的描述要更为精准一些。面对**自由落体** (free fall)，古希腊时期的 Aristotle 尝云，“置两个物体于同一高度释放，重的那一个将率先落地”。公元 5~6 世纪，一位希腊哲学家 Ioannes Philoponus 记载了这样一个发现：在同一点的各种物体都以完全相同的方式运动着，无论大小，无论重量。虽说同一高度不一定代表着同一点，但同一点很显然属于同一高度，这即对 Aristotle 学派的物理思想做出了挑战。时间来到 16 世纪，彼时的意大利弥漫着复兴的气息。一位研究自然的学者 Galileo 得知 Aristotle 学派所秉持的思想有难以弥补的悖论，即所谓关于两个相互牵制的物体，下落加速应该归因于谁的问题。随后利用斜坡和单摆实验发现了下落距离正比于时间平方，且与其质量无关的规律。根据连续性思想以及忽略不必要阻力的思想，其可合理外推至自由落体。引入微积分学，Galileo 发现物体受引力作用的加速度之模应为一个常数。因此在同一点的各种物体都以完全相同的方式被引力加速。他在西欧致力于传播这一事实，得到了广泛回应与讨论。后来人们知道，在高度变化过大时，其对加速度的影响亦即愈发明显了起来，故加速度只能在局部近似不变。经过数十年的思索与竞争，Newton 找到了统一地上“重力”学问和天上“引力”规律的方法，于 17 世纪末在他的《自然哲学的数学原理》[1] 中阐明了一个更为精确的规律。出于宗教性的美学追求，当时已普遍承认了 Kepler 的三大定律：椭圆轨道、掠面恒速、周期关系。建立正交坐标系 $\{x, y, z\}$ ，设行星的位矢

\mathbf{x} , 恒星静止于 \mathbf{x}' , 某行星的掠面恒速说明 $\ddot{\mathbf{x}} \times (\mathbf{x} - \mathbf{x}') = \mathbf{0}$, 故若用 Newton 第二定律¹ $\mathbf{f} = m\ddot{\mathbf{x}}$ 描述行星所受的引力作用, 则 \mathbf{f} 是有心力. 如此, 椭圆轨道这一现象就可将其模长确定到 $|\mathbf{x} - \mathbf{x}'|$ 的平方反比:

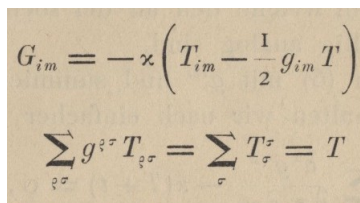
$$\mathbf{f} = -GMm \frac{\mathbf{x} - \mathbf{x}'}{|\mathbf{x} - \mathbf{x}'|^3},$$

其中 G 是万有引力常数, M 是恒星质量. 这称为 **Newton 引力**. 以此又能自洽地给出周期关系. Newton 理论取得了巨大成功, 然而, 临近 20 世纪, 它也遇到了一系列困难: 继续扩大高度, 把视线放到更远的距离去, 一个明显的问题是, 经典引力将成为一种能传播无限远的超距作用. Newton 本人拒绝这种看法, 但也不能提供合理的物理实质, 而始终认为其使用平方反比律只是在“数学地描述”吸引现象. 那么在很远的地方, 引力还会是平方反比吗? 想要解决超距作用, 一种思想便是想象整个空间弥漫着一种媒介性质的物质, 它可将相互作用传递到很远的地方. Faraday 称之为**场** (field). 比如, 空间矢量场就是给每一空间点指派一个矢量, 在插图中, 通常把某处矢量绘成起于该处的样子. 18~19 世纪, 经由 Gauss、Maxwell 等人对数学的发展, 场逐渐作为一种物理实质而在物理学里占据一席. 人们在研究电磁力时也尽力让表达式向平方反比律靠近, 如 Coulomb 静电力, 之后又用场论改写. 那么, 当然可将场论概念借至 Newton 理论里来. 经典引力是保守的, 就是说, 某个试验质点在引力场中的受力适合用标量场 ϕ (称为**引力势**) 表为 $\mathbf{f} = -m\nabla\phi$, 这样, **引力 Poisson 方程**是

$$\nabla^2\phi = \frac{\partial^2\phi}{\partial x^2} + \frac{\partial^2\phi}{\partial y^2} + \frac{\partial^2\phi}{\partial z^2} = 4\pi G\rho,$$

其中 ρ 是质量密度. 第 1.4.1 节会详细说明过程, 且要强调, 由于默认任意距离满足平方反比, 场和 Poisson 方程的引入不能根治矛盾. 引力理论仍待改进.

目前, 人们普遍接受了 Albert Einstein 的理论在引力现象上所取得的巨大成功. 他为了推相对性原理, 使之也能够囊括引力, 选择了另辟蹊径, 从引力同惯性、时空几何学的关系上入手, 终于在 1915 年 11 月完成了他 8 年奋斗的最后一幕:



$$G_{im} = -\kappa \left(T_{im} - \frac{1}{2} g_{im} T \right)$$

$$\sum_{\sigma} g^{\sigma\tau} T_{\sigma\tau} = \sum_{\sigma} T^{\tau}_{\sigma} = T$$

图: 取自《引力的场方程》(*Die Feldgleichungen der Gravitation*)

¹在较早的英语文献中, 会在人名后添后缀 (通常以 -an 结尾) 表示形容词, 甚至是名词, 如 Galilean, Newtonian, Jacobian. 这些表达沿用至今, 但对当代人物来说一般不再采取. 为迎合时代, 本书中这些旧时人物亦不添后缀.

这就是 Einstein 最引以为傲的方程。在先后经历三大实验验证后，各大报纸相继对此刊登头条：“现代物理学赢来了一次伟大胜利。”这些“神秘”数学符号固然令人费解，但我们暂时不需要完全知晓这些东西的含义，只需知道，方程的**左边关乎着引力**，而**右边关乎着物质**。这与此前 Newton 的描述方式是类似的——注意 Poisson 方程确为“引力”正比于“物质”的形式。比较 Einstein 和 Poisson 的方程，明显后者的记号要通俗些，前者那可怕的符号诚然多少“败坏”了广义相对论的名声。但不必妄自菲薄，因为它们都代表了人类智慧漫漫长征的里程碑，且都以必要的概念框架为前提：Newton 理论需要微积分，而 Einstein 理论需要额外的几何学罢了。为了陈述 Newton 理论所必须的知识，已被纳入到了本科一年级教学中，譬如 3 维欧氏空间、坐标系、函数、导数、质量和力这样一些概念，相信诸位具有一般的数学和物理背景的读者并不陌生。至于广义相对论的结构，或许人们就远不是这样熟悉了，但其所必须的知识并不比微积分难到哪儿去，一旦接受了，Einstein 理论将变得更自然、更简单。

想真正了解一个理论，仅仅知道怎样写出控制它的方程还远不够丰富。广义相对论是与 20 世纪物理学所作的一些最蔚为奇观的预言联系在一起的，也许读者或多或少在科普或通俗读物中听说过一些概念，比如引力坍缩、黑洞、奇点、宇宙膨胀……这些东西在 1915 年时还未为人知，所以它们与场方程的提出没有关系。只有当后来人们懂得围绕场方程的解，去研究整体的动力学应用时，才会注意到这些现象。这件事花的时间长得相当惊人，可能人们对其中的艰辛事迹的熟悉程度，甚至还远不及对场方程背后孤勇奋斗的了解程度。因而本书将会着重讲解广义相对论的经典应用和前沿话题，而对某些历史作简短一瞥。

场方程的解析解固然有很多，但它们都比较简单，我们只能对简单解邻域中的一般解有定性了解。所谓的宇宙监督假设、一般条件的奇点等问题都未得到普适解答。而这些问题恰恰对于广义相对论的物理意义和适用范围都有基础衡量。我们只能期望场方程会继续揭示出美丽的数学结构，帮助人类进一步地认识物理世界。

本书致力对引力及时空的话题做一个全面、深入的讲述，由其教程之出发点，仅为了帮助相关专业的学生及从业人士建立必要的数学、物理基础，为理解前沿领域中相关论文作准备。本书主要内容分为三个部分：必要的基础（主要为经典内容）、对许多话题的概述（上世纪末以及千禧年的前沿、基础概念的探讨）以及随时可供查阅的附录。书的末尾附上了参考资料及索引以飨读者。谨以此序阐明本卷内容之深度、广度，有如下考虑。

以理工类专业大一学习的微积分和线性代数为基础。笔者假定读者已完整地、熟练地掌握多元微积分的矢量、场论符号，了解一点分析学。本书尽力穿插上述课程中没有完整涉及的部分数学知识。在介绍新内容时，笔者将在教程中尽可能多地提示读者，新内容与其所学过的哪些内容相关。希望读者在比较脑海里的新旧知识后能

够同意，笔者在此教程中介绍的新知识并非完全陌生。

刨析后续物理理论所需要的基础数学概念。例如，在引入流形的概念（第 3 章）时，需要用到集合的划分知识，因此笔者会在期间穿插集合关系的知识。亦如，作为基础概念的各类张量，在本教程中会事先简化成“矢量复合出来的数组”或“满足 Lorentz 协变性的数”，并引入对偶空间后将其定义为多重线性函数。按照上一条假定，考虑到读者的数学基础，仍需要比较详尽地介绍有限维矢量空间及其上的线性变换的知识。

采用坐标依赖（详细讲解并使用求和约定）与无坐标依赖（数学系符号为主，抽象指标仅作简介）交互的表达风格。笔者希望这种做法本身能彰显广义相对论的建立者希望实现的物理客观性，既物理规律（数学表达式）不依赖包括坐标系的选择在内的任何主观选择；但同时在面对具体应用问题时，仍然保留建立坐标系的权利。

本书看似未设习题，仅讲解足够的例题，为使过程详尽将使用彩色标注，但实际上仅为证明某定理所需的数学概念和引理，要么给出简介和参考，要么放于附录。定理的证明过程，仅供有兴趣的读者参考，故未必都提供。某些省略处需读者自行补充，以当练习。笔者尽可能给出相应的参考资料。本教程用不到的数学知识不作介绍，部分放入附录。例如，在拓扑学中提到了部分集合论的公理，但又假定关于自然数、算术运算、偏序、全序、数学归纳法等知识为暂时无需，包括 Peano 公设在内的相关的内容。故关于微分几何章节的最终状态，深度上似乎要介绍完微分拓扑，但广度上又不完整。这是笔者故意为之的。因为本教程的目标不是要向读者大肆介绍数学知识，而是补充看懂广义相对论前沿进展的最少必要的、作了严格定义的数学知识。

概念的定义通过字体改变来暗示，或以带编号的方式引入。定理和例子则均带编号。定义是极其重要的，它在文中一般只出现一次，因此难免要经常反复回顾。对于不采用带编号式引入的概念，只是因为考虑到数量很多，若都带编号可能打断行文的流畅，让本就不易的内容雪上加霜。代价是使其失去引用链接之便利，故在此敬请读者学习时应不厌其烦地来回翻阅。

夏草

2024 年 4 月

目录

前言	v	1.4.5 线性近似	57
		1.4.6 Mach 原理	58
		1.4.7 洞悖论	59
I 入门	1	1.5 单位制转换	59
1 序曲	3	1.6 Lorentz 群	61
1.1 场论初步	3	2 现代框架	65
1.1.1 场及其定域性	3	2.1 拓扑学	65
1.1.2 时空度规	6	2.2 流形与度规场	79
1.1.3 矢量	13	2.3 超曲面	83
1.1.4 张量	17	2.4 仿射联络与曲率	83
1.1.5 时空结构	20	2.5 微分形式	92
1.1.6 广义不变性	25	2.6 流形的映射	92
1.2 张量代数	28	2.6.1 Lie 导数	92
1.2.1 对称与反称	28	2.6.2 Killing 矢量	93
1.2.2 协变导数	28	2.7 纤维丛	93
1.2.3 张量密度	31	2.8 Frobenius 定理	93
1.2.4 多元微分学	33	2.9 参考系	93
1.3 拉氏理论	33	2.9.1 典型运动效应	93
1.3.1 变分法和对称性	34	2.9.2 自转	96
1.3.2 粒子的拉氏理论	37	2.9.3 膨胀、剪切与扭转	97
1.3.3 场的拉氏理论	39	2.9.4 刚性系空间几何	97
1.4 经典引力场论	43	2.9.5 对钟	98
1.4.1 引力与时空几何	43	2.10 标架表述	98
1.4.2 潮汐与曲率	48	2.11 哈氏理论	98
1.4.3 Einstein 方程	53	2.11.1 3+1 分解	98
1.4.4 最小耦合	55	2.11.2 初值问题	98

3 黑洞物理	99	9 上同调论	133
3.1 黑洞	99		
3.2 引力坍缩	108		
3.3 Penrose 图	109	III 半经典理论	139
3.4 时空延拓	109		
3.5 嵌入图与虫洞	109	10 Cauchy 动力学与奇点定理	141
3.6 RN 黑洞	109	10.1 Witten 迅捷性	141
3.7 黑洞热力学	109	10.2 Penrose 过程	141
3.8 dS 与 AdS 时空	109	10.3 能量条件	141
		10.4 奇点定理	141
4 因果结构	111	10.5 宇宙监督假设	141
4.1 可定向性	111	10.6 高-Wald 定理	141
4.2 因果线	111		
4.3 因果条件	111	IV 附录	143
4.4 整体双曲时空	113		
		A 量纲理论	145
5 实验理论	115		
5.1 等效原理	115	B 量子力学	147
5.2 几何	115	B.0.1 Hilbert 空间	147
5.3 后 Newton 近似	115	B.0.2 赋范空间	152
5.4 引力波	115	B.0.3 对偶性	155
5.5 宇宙学	115	B.0.4 Fourier 变换	157
		B.0.5 Fredholm 算子	158
II 进阶	117	B.0.6 单参数 (半) 群	159
		B.0.7 谱理论	162
6 Lie 群与 Lie 代数	119	B.0.8 算子代数	163
7 规范理论	121		
		参考文献	165
8 经典力学的几何理论	123		

Part I

入门

Chapter 1

序曲

1.1 场论初步

1.1.1 场及其定域性

无论生活中还是学术里，人类总要将**物质** (matter) 或**客体**进行划分。不妨将易察觉的日常物质称为**物体**。物体受到相互作用的信息时将通过动量和能量的变化表达反馈。假若只存在单独两个物体而其周遭又无任何媒介，其间相互作用是无法传递的，故应设想除这些物体外还弥漫着一种不易察觉的物质，称为**场**，物体与它相互作用，而再由它间接作用于另一物体。相互作用强度涉及关乎物体本性的某种取实数的物理量，称为**荷**。物体与场的相互作用大致体现为两方面：1. 场对具有相应荷的物体有相应的作用；2. 具有荷的物体本身可作为场源激发并影响场。无论如何这避开寻找对场的数学语言。利用前者，设想一种**试验物**，首先可忽略体积、形变及自转，如质点、点电荷，其次最重要地，可忽略其自身产生的场对外场之干扰，这样就可通过测量试验物上的作用力来研究外场，并寻找该力与试验物的荷之联系。每单位荷的对应作用力定义为**场强**。场强通常无关于该试验物，因此可良好地用作场的数学模型。场强往往又能表为某个**势场**的某种导数，故场的数学描述就是用场强或势场指代物理场本身。当然，宏观物体也可用场去描述。

场的数学模型一般分为两类，一类是**连续介质** (continuum)，一类是点系统。连续介质是点系统的连续极限，点系统是连续介质的离散极限。以质量密度为例，总质量 M 可分为若干质元 dM 并以一点代之，这样可形式地写出 $\rho = dM/dV$ ，其中取 3 维空间正交坐标系 $\{x, y, z\}$ ，则 $dV = d^3x = dx dy dz$ 是该处体元。常用连续介质模型描述流体、固体、作用媒介场等，但须注意，目前认为物质皆由原子、分子等微粒构成，假若体元任意小，将致使其中一般不含任何微粒，这样密度似乎为零。问题

出在哪儿? 其实, 应将连续介质模型所用之概念理解为**统计平均的**, 例如湖水所产生的引力场并不取决于单个水分子的状态, 而只依赖于大量分子的平均性质. 但平均性质在各点可以不同, 例如湖水底部的压强高于上层, 温度也随着深度变化, 再如大气密度亦随位置改变而改变. 那么多少微粒的平均才合适? 首先, 微粒应足够多以使单个微粒的行为没有影响, 进而其中密度、平均速度等取值可视为相同. 微粒亦须足够少, 以保证相对地均一, 即其中各处微粒速度、动能、微粒间距等不能相差过大, 进而就能描述点与点间性质的差异. 这样在某点合适地取平均的微粒邻域称为该处一个**单元** (element). 此乃连续介质模型之根基. 极度稀薄的气体或某些非平衡态系统不存在这种单元, 故超出了此模型的适用范围, 需另辟蹊径.

严格地, 单元指在一点处直接给定某值, 整体构成关于每一点的映射, 即数学上的场. 连续介质模型就是指某属性的分布可由分段光滑¹的场确定. 对质量密度场 $\rho(\mathbf{x}, t)$, 最好还希望其非负, 且存在某个球形区域使 ρ 在球外恒为零, 即 $\rho \neq 0$ 的区域有界, 数学上称其**紧支撑的**. 进而定义总质量为 $\int \rho(\mathbf{x}) d^3x$, 这里默认取全空间体积分而省略积分域 \mathbb{R}^3 . 单元与质点不同, 质点是企图将可观质量集中于一点处, 体元无限小时, 质点的质量并不会无穷小. 不妨这样理解, 先考虑质点限制于一维空间的原点处, 具有单位质量, 设想其密度分布 $\delta(x)$ 为一系列函数 $\{\delta_n\}$ 的逼近, 只要规定每一 $\delta_n(x)$ 的无穷积分都为单位质量即可, 而最终 $\delta(x)$ 应在质点处正向发散, 且在非质点处都为零². 比如, 考虑 Gauss 分布 $\delta_n(x) := \sqrt{n/\pi} e^{-nx^2}$, 其中 “:=” 表示“定义为”. 可证任意光滑函数 $f(x)$ 的如下无穷积分满足

$$\left| \int \sqrt{\frac{n}{\pi}} e^{-nx^2} f(x) dx - f(0) \right| = \left| \int \sqrt{\frac{n}{\pi}} e^{-nx^2} (f(x) - f(0)) dx \right| \\ \leq \max \left| \frac{df}{dx} \right| \int \sqrt{\frac{n}{\pi}} |x| e^{-nx^2} dx = \frac{1}{\sqrt{n\pi}} \max \left| \frac{df}{dx} \right|,$$

因此

$$\int \delta(x) f(x) dx := \lim_{n \rightarrow \infty} \int \delta_n(x) f(x) dx = f(0). \quad (1.1.1)$$

注意极限号在最外面, 不要交换极限和积分. 这就是 **Dirac δ 函数**. 上式称为**挑选性**. 推广地定义其 3 维情形 $\delta(\mathbf{x}) := \delta(x)\delta(y)\delta(z)$. 设 $f(\mathbf{x})$ 为光滑标量场, 易得如上类似性质, 进而用简单的平移操作得

$$\int f(\mathbf{x}) \delta(\mathbf{x} - \mathbf{x}_i) d^3x = f(\mathbf{x}_i). \quad (1.1.2)$$

令质点系 $\{M_i\}$ 总位于某有界区域. 单个质点的密度表为 $\rho_i(\mathbf{x}, t) := M_i \delta(\mathbf{x} - \mathbf{x}_i(t))$, 质点系总密度为 $\rho(\mathbf{x}, t) := \sum \rho_i(\mathbf{x}, t)$, 进而给出质点系总质量 $M = \sum M_i$. 当然,

¹函数的**连续**指定义域上处处满足极限等于取值, 记 C^0 ; k **阶导函数连续**记作 C^k ; **光滑**指任意阶导函数连续, 可记 C^∞ . C^1 一定说明导数存在 (可导), 反之不然. 分段光滑指其可分为有限个光滑区间.

²这是物理上的要求. 一般的**广义函数** (generalized function) 理论可无此要求, 详见附录 B.

$\delta(x)$ 其实是象征性的记号，而非普通意义的函数. 比如， δ 函数的无穷积分为 1 与 $x \neq 0$ 时 $\delta(x) = 0$ 矛盾，因为函数值只在有限点改变甚至无意义时不影响可积性和积分值，无穷积分只能为零. 但其代数演算是严格的，因为物理学并非孤立地看待它，而是连同其积分来使用，即将 $\delta_{x_0}(f(x)) = \int f(x)\delta(x - x_0) dx = f(x_0)$ 看作一个映射. 描述一个函数所需的，就是给定义域内每一点赋予实数. 一点由 1 个独立实数描述，称**自由度**为 1，然而定义域内的点无限多，因此一个函数具有**无穷自由度**. 这种自变量并非分立实数的映射称为**泛函** (functional)，即所谓函数的函数. 泛函的自变量称为**宗量**. 换言之，我们并非在使用孤立的 $\delta(x)$ ，而是在使用一系列函数 $\{\delta_n\}$ 所满足的积分性质 (1.1.1) 式. 凡某等式中出现形式记号 $\delta(x)$ ，只需注意严格理解是先 将 $\delta(x)$ 替换为符合 (1.1.1) 式的任意实函数列 $\{\delta_n\}$ ，再令整个积分的极限成立.

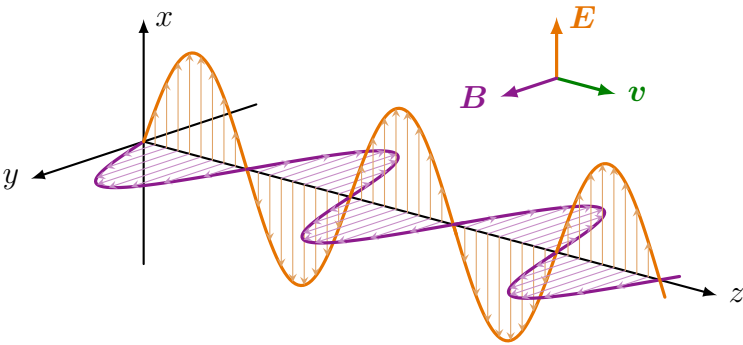


图 1.1: 电磁波

场论自诞生之初就有**局域性** (locality)，也称**定域性**，比如电磁场中相互作用的传播速度有限. 注意经典力学没有这种要求，Newton 引力就是典型的超距案例. 要将引力理论从经典力学的范畴脱离，成为真正的引力场论，就要找到电磁场论定域性的根源，这来自于经典力学和电磁学的矛盾：光在经典力学下不能定义，而在电磁学下定义为**无源电磁场**或**电磁波**，设电荷密度、电流密度均为零，则可将 Maxwell 方程化为波动方程，故光以常速率传播. 这也在实验上得到证实. 我们不再赘述这一段精彩的历史，只需指出 Maxwell 方程组成立于任意惯性系，故光在任意惯性系所测速率与光源运动无关，这一命题称为**光速不变**. 一般会强调真空光速，但要注意介质光速其实仅是光波各处相位偏移所致的表面效果，属于相速度. 光速不变固然十分挑战直觉，这意味着就光而言，其速度并不满足叠加性，而叠加性又源于**相对性原理** (principle of relativity)、绝对空间度量和绝对时间度量. 相对性原理就是说，没有实验能检测出不同惯性系之间的内在区别；后两条直接源于 Newton 时代的绝对时空观. 这样所有速度都应是相对的. 这在《原理》推论 V 和更早的 Galileo 的著作中就已有类似叙述，不过那时仅关注于力学实验，但 Einstein 对相对性原理十分笃

信, 欲把相对性原理放到更普遍的地位, 则它必须适用于全部物理定律, 而不只是适用于力学. 因此问题出在用绝对时空看待电磁学. 关键是建立一种崭新的时空观.

1.1.2 时空度规

物理学离不开背景. 比如 Newton 力学的研究背景是 \mathbb{R}^3 空间. 背景有时也可以很抽象, 比如热力学中我们会研究压强、温度的相图. 通俗地说, 背景由点连续地而非分立地分布而成. 对于粒子位置这种比较具象的概念, 一般地可推广为**位形** (configuration), 用以描述系统的空间分布或形状. 所有可能位形构成的集合称为**位形空间**, 如粒子的位形空间就是 \mathbb{R}^3 . **路径** (path) 是位形的演化, 比如粒子的位置 \mathbf{x} 随时间 t 的运动曲线 $\mathbf{x}(t)$. 因此一般地可将路径记作 $\psi(\xi)$, 其中 ξ 可以是非时间的参数. 故准确地说, 物理学研究的是各类**事件** (event). 为叙述一个生活意义上的事件所需的, 和记叙文一样需要地点和时间. 不过事件总要发生在一段时间内并占据一定空间, 但我们可以抽象出理想模型. 就质点而言, 可准确地将事件所对应的运动学模型称为**时空点** (spacetime point), 其在某坐标系下的坐标可记

$$x^0 := ct, \quad x^1 := x, \quad x^2 := y, \quad x^3 := z, \quad (1.1.3)$$

其中给时间乘上速率常数以使量纲一致¹. 由于光速不变, 选择 c 是最无歧义的. 此后, 本书一般默认几何制使 $c = G = 1$ (合理性见 1.5 节), 则 $x^0 = t$. 这样 x^0, t 都可称**坐标时**. 通过给空间再补上一个时间轴, 使我们用以讨论物理问题的 3 维背景提升至一种 4 维背景, 进而讨论各时空观的区别. 这

个 4 维背景就是所有事件按一定顺序排列的无穷集合, 称之为**时空连续统** (spacetime continuum), 简称**时空**. 可见时空就是事件的位形空间. 坐标系也不再仅为那种 3 维空间坐标系, 而可以是 4 维时空坐标系. 初学者可将时空想象为所学过的 x - t 图, 即一张 2 维白纸 \mathbb{R}^2 , 其上选有某个坐标系, 这样坐标实数具备了事件位置 (所排顺序) 的意义, 或说事件被标记为相应坐标. 事件的排序并非杂乱无章, 实际上, 我们总希望某事件的邻域 $O \subset \mathbb{R}^2$ 的确也被足够连续地标记好相邻坐标. 考察某个质点, 不考虑突然解体或其余特殊情形, 无论静止还是运动, 质点必须经历一系列

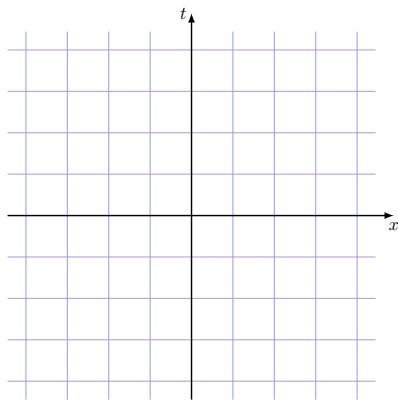


图 1.2: 2 维时空图

相邻事件, 比如就某质点的某时刻而言, 它必然要延伸至邻域 $O \subset \mathbb{R}^2$ 中. 尤其对物理上的客观微粒而言, 常认为它足够“永生”, 故其所经历的事件可无限地延伸,

¹ 计时仪器总借距离测量时间, 如钟的指针轨迹.

总之，它在时空中将划过一条曲线。质点所经历的所有事件构成的曲线就称**世界线** (worldline)，也即时空视角的路径。这或许让某些读者联想到宿命论或决定论，毕竟世界线是该粒子的所有历史和未来。以后的学习中会接触量子理论对时空因果的影响：量子观点认为粒子不可能同时确定动量与位置，故经典的轨道模型不再适用，世界线将是某种不稳定的路径而被时空上的波函数替代。现在仍期望先构建经典的相对论模型，世界线仍是很好的概念。凡示意时空及世界线的图就是**时空图** (spacetime diagram)：就 2 维时空而言，常选择时间轴正方向朝上，而空间轴正方向朝右，这一习惯与经典力学的 $x-t$ 图稍有区别；对 3 维时空，就好似将 x, y, z 中的 z 轴替换为时间轴。然而就一般的 4 维时空，固然画不出直观图像，但其实许多问题总能借助对称性，使时空可“降维”，通常省略某空间维度。并且，无明显对称性时，还可考虑空间维度是否在观测上具有必要性。譬如，研究某大范围的平面大地上的时空时，其上运动物体除可飞行物体外，其高度基本近似不变，故可省略高度这一维，即研究平面运动的 3 维时空图。

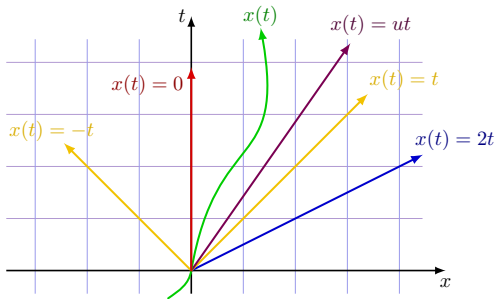


图 1.3: 世界线

背景上坐标系的选取可以非常任意，比如相比于正交坐标系，用极坐标系解行星运动将更为容易。但为避免循环定义，总需选择一种逻辑起点。注意坐标系给出若干实数，可见从 \mathbb{R}^n 出发最方便。实数集元素的坐标天然地正是其本身，即 \mathbb{R}^n 任意元素可表为 (x^1, \dots, x^n) ，各 x^i 皆称此元素的**自然坐标**。此处右上角数字只是**指标** (index) 而非乘幂。不必担心，在很多情况下是不会与乘幂相混淆的；实在冲突时，只需文字标注即可。那么什么叫坐标系？首先要捋清一下逻辑。背景是没有坐标的，而可用不同的方式，给其

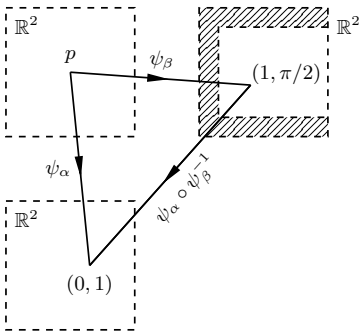


图 1.4: 坐标变换

上的点赋予不同坐标. 一种方式就对应一种坐标系. 因此, 先有一个 \mathbb{R}^n 但暂时视其“没有坐标”(可区分地记作 M), 然后再人为地给其定义坐标. 这就是 \mathbb{R}^n 具有多种不同坐标系的准确含义. 比如白纸 \mathbb{R}^2 的某点 p 可被赋予自然坐标 $(0, 1)$, 但也可被指派极坐标系的 $(1, \pi/2)$, 如图 1.4. 更准确地说, 存在两种“附着”在 \mathbb{R}^2 上的不同“坐标空间”, 一种存放正交坐标, 一种存放极坐标. 这样所谓的“指定方式”就对应于数学上的映射概念, 即**坐标映射** $\psi: M \rightarrow \mathbb{R}^n$. 自然坐标系无非是 \mathbb{R}^n 上的恒等映射. M 上的函数称为**标量场** (scalar field). 定义 $x^i: M \rightarrow \mathbb{R}$ 是 ψ 给出的第 i 个坐标, 这样坐标系又可简记 $\{x^i\}$. 记自然坐标系对应映射 ψ_α , 且 $x := x_\alpha^1, y := x_\alpha^2$. 这样 $\psi_\alpha(p) = (x(p), y(p))$. 对此自然坐标系可定义极坐标系 ψ_β 并同理设 $\psi_\beta(p) = (r(p), \theta(p))$. 则复合映射 $\psi_\alpha \circ \psi_\beta^{-1}: \mathbb{R}^2 \rightarrow \mathbb{R}^2$ 将使得 $x = r \cos \theta, y = r \sin \theta$, 其中注意标量场运算按

$$(f+g)(p) := f(p) + g(p), \quad (fg)(p) := f(p)g(p)$$

给定. 足见 $\psi_\alpha \circ \psi_\beta^{-1}$ 对应两个二元函数, 这就是**坐标变换**. 不过, 总希望坐标映射是双射¹, 故极坐标映射 ψ_β 的值域不应该是整个坐标空间 \mathbb{R}^2 , 而只能人为规定一个子集, 如 $V = (0, +\infty) \times (0, 2\pi)$. 图 1.4 中将其以外的部分涂了斜线. 且可看到极坐标的缺陷: 为使 V 为 \mathbb{R}^2 开子集, 就不能标注背景 \mathbb{R}^2 的所有点, 如原点和极轴.

背景维数的准确来源是看坐标系的独立坐标数. 换言之, 对于一个背景 M , 若存在坐标系 $\psi: M \rightarrow \mathbb{R}^n$, 就称背景维数 $\dim M$ 是有限的 n . 设两套坐标系 $\{x^i\}, \{x'^i\}$, 其间坐标变换是可逆函数 $x'^i = h^i(x^1, \dots, x^n)$. 一般考虑光滑坐标变换, 则任意点总有全微分和链式法则:

$$dx^i = \sum_{j=1}^n \frac{\partial h^i}{\partial x^j} dx^j, \quad \frac{\partial}{\partial x'^j} = \sum_{i=1}^n \frac{\partial (h^{-1})^i}{\partial x'^j} \frac{\partial}{\partial x^i}. \quad (1.1.4)$$

这里 $(\partial h^i / \partial x^j)$ 就是分析学里的 **Jacobi 矩阵**. 求和哑元称为**哑** (dummy) **标**, 而留下的指标遍历 $1, \dots, n$ 以示 n 个分量等式, 称**自由** (free) **指标**. 且上标 (upper/raised indices) 和下标 (lower indices) 位置一致, 称为**指标平衡**. 其中分母的上标看作分子的下标. 不与自由指标冲突的前提下哑标可任意替换. 由于哑标经常是一上一下成对出现, 因此可采取 Einstein 提出的**求和约定** (summation convention) 以省略求和号, 并通常混用符号地简写为 $x'^i = x'^i(x^j)$, 则

$$dx'^i = \frac{\partial x'^i}{\partial x^j} dx^j, \quad \frac{\partial}{\partial x'^j} = \frac{\partial x^i}{\partial x'^j} \frac{\partial}{\partial x^i}. \quad (1.1.5)$$

换言之, 若式子中同一指标一上一下出现两次, 则意味着要对该指标求和. 毕竟以后总和各种指标符号打交道, 初学者可能会感到复杂但又需要习惯, 不如关注分量本

¹即一一对应. 有时亦称 1-1 映射 (one-to-one map), 但它在部分文献里指单射, 故应该避免使用此术语.

身. 代数理论保证这一做法的合理性, 比如一个因式中指标的确至多出现两次; 多个因式间可出现重复哑标对, 因为是加法关系; 一个因式可以出现多对哑标, 由乘法的交换、结合及分配律, 求和顺序是无所谓的, 这体现了我们关注其求和通项的精神. 注意二者的变换矩阵互逆:

$$\frac{\partial x^i}{\partial x'^k} \frac{\partial x'^k}{\partial x^j} = \frac{\partial x'^i}{\partial x^k} \frac{\partial x^k}{\partial x'^j} = \delta_j^i, \quad (1.1.6)$$

这里 δ_j^i 表示恒等变换 I 的矩阵元, 称为 **Kronecker 符号**. 考虑时空坐标时, 指标一般用 μ, ν 等希腊字母, 可取 $0, 1, 2, 3$. 而 i, j 等拉丁字母则取 $1, 2, 3$ 或 $1, \dots, n$.

究竟什么坐标系算是正交的? 注意夹角是一个几何概念, 而几何学的研究对象均是**坐标独立的** (coordinate independent) 或**坐标选择无关的**, 如古典几何. 当然, Descartes 坐标方法的引入也有一大方便之处: 先取自然坐标系, 按勾股定理的方式定义**线元** (line element) $d\ell$ 满足

$$d\ell^2 = \delta_{ij} dx^i dx^j, \quad (1.1.7)$$

其中 δ_{ij} 也是 Kronecker 符号. 这样欧氏几何的许多涉及图像性质的概念可从中导出, 如夹角可借三角函数定义, 线长可用积分定义. Gauss 注意到一旦确立好某个二次型线元, 比如在某系 $\{x^i\}$ 给出 $d\ell^2 = g_{ij} dx^i dx^j$, 则二次型仍然保持于任何坐标变换下, 比如可任取另一系 $\{x'^k\}$ 得

$$g_{ij} dx^i dx^j = g_{ij} \frac{\partial x^i}{\partial x'^k} \frac{\partial x^j}{\partial x'^l} dx'^k dx'^l = g'_{kl} dx'^k dx'^l, \quad g'_{kl} := g_{ij} \frac{\partial x^i}{\partial x'^k} \frac{\partial x^j}{\partial x'^l}. \quad (1.1.8)$$

故他建议将任意 n 维空间中的线元按二次型的方式定义, 这称为 **Gauss 假设**, 而排除比如 $d\ell = (|dx|^3 + |dy|^3)^{1/3}$. 本书谈及几何学总采取此平方假设. 数学上称之为**度规** (metric). 式 (1.1.7) 给出了**欧氏度规**. 能保持 $g'_{kl} = g_{ij}$ 不变的特殊变换称为**等度规变换** (isometry). 比如, 对欧氏度规, 显然存在无数可不同于自然坐标的坐标系使 $g_{ij} = \delta_{ij}$, 统称 **Descartes 坐标系**. 它们之间的变换都是欧氏度规的等度规变换. 历史上还曾构造双曲度规 $d\ell^2 = (dx^2 + dy^2)/(1 - x^2 - y^2)$, 当接近欧氏圆盘 $x^2 + y^2 < 1$ 边缘时, 该线元相较于欧氏线元会越来越大, 但它除平面几何第五公理外符合其余公理. 公理的特征是可替代性, 故这说明第五公理不能由其余公理推出, 而只能作为欧氏几何的约束. 故度规决定一套几何学. 一般地对 n 维空间, 规定¹**正交归一系** $\{\xi^i\}$ 是使度规表为

$$(\tilde{g}_{ij}) = \text{diag}(\pm 1, \dots, \pm 1) \quad (1.1.9)$$

的坐标系. 设对角元中 -1 的个数为 s . 若 $s = 0$, 则称**正定的** (positive definite) 或 **Riemann 的**. 均为 $s = n$ 则称**负定的** (negative definite). 其它皆称**不定的** (indefinite).

¹对任意可逆对称矩阵, 总能找到合同变换使其对角且归一化, 这能保证该定义的存在性.

细心的读者应发现直线、正交等概念在数学中只能抽象地定义. 是的, 生活中的直线、垂直、勾股定理属于“习俗”, 我们会默认绷紧直绳呈“直线”、水面同重力方向“垂直”, 这其实是物理上的经验现象, 并用之以工业生产——读者是否也曾好奇日常中各种呈规则图形的人造产品从何而来? 而自然界中宏观的非人造规则物体确实很少. 如果抽象地看待这些概念, 则可以不要直线是“直”的. 直线和曲线不能肤浅地看作先验绝对的事物, 而只能保留相对差异. 如图 1.5,

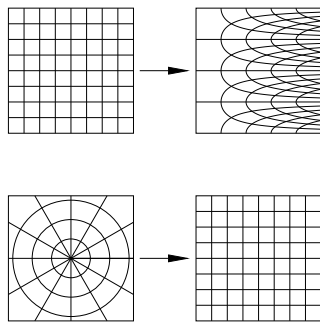


图 1.5: 线条曲直的相对性

可设想同一背景 \mathbb{R}^2 上覆盖有正交坐标系和极坐标系, 但采取另一画法, 将原先的同心圆画为竖向直线, 根据坐标变换可知正交坐标系的网格将扭曲. 将正交坐标系的网格画为生活意义的正交直线, 其实是一个为便于直观理解的习惯, 犹如读者在想象实数集 \mathbb{R} 时, 通常不会将其看作曲线, 而是一根直线. 像这样, 规定某正交坐标系按此习惯绘制图像, 就称为以此系为**基准**. 根据坐标变换可绘制出其它坐标系的网格, 其与基准系的差异自然显现. 此外, 右手系也是约定的. 一般先约定自然坐标系是右手的, 进而只要坐标变换是**保定向的** (Jacobi 行列式为正), 则称新系也是右手系. 在画图时, 基准坐标轴常常也画成右手的顺序.

再回到时空. 不同时空观涉及不同的坐标变换. 在绝对时空观中, 若两系的钟初始设定不同, 亦可能出现坐标时的不同, 但国际计量局会通过技术手段确保某时区的人们具有一致参考标准. 可约定两惯性系 S, S' 在同一位置时 (如原点), 时钟读数相同 (如零时刻), 从而在经典观念中消除任何坐标时差异, 即约定 $t = t' = 0$ 时原点重合, 还可进一步约定正交空间轴重合而消除更多不必要差异. 设 S' 相对于 S 以沿共同 x, x' 轴正方向的 u 运动, 这样 y, z 坐标总是一样的. 关注于 t, x 坐标, 可推知 **Galileo 变换**

$$t' = t, \quad x' = x - ut. \quad (1.1.10)$$

而相对论给出的结果是 (二维) **boost 变换**:

$$t' = \gamma(t - ux), \quad x' = \gamma(x - ut), \quad \gamma := \frac{1}{\sqrt{1 - u^2}}. \quad (1.1.11)$$

这里 γ 称为 **Lorentz 因子**. 容易看到, 即使做初始校准, 仍不可能使任何事件皆具一致坐标时. 当然, 国际界会考虑到卫星信号传输中的这种效应. 显然 Galileo 变换是 boost 变换的低速极限, 当速度与光速可以比拟时, 两种变换的差别才显示出来. 这套理论被 Lorentz 命名为**相对性理论** (the theory of relativity), 简称**相对论**. 当下主流称呼是**狭义相对论** (special relativity). 不同时空观往往暗示了时空背景有着不同的几何结构. 对 boost 变换做微分后整理得 $-dt^2 + dx^2 = -dt'^2 + dx'^2$, 可将 x 看

作空间点与原点的距离, 这样只要空间坐标系间相差旋转变换就可还原出 y, z 坐标:

$$-dt^2 + dx^2 + dy^2 + dz^2 = -dt'^2 + dx'^2 + dy'^2 + dz'^2, \quad (1.1.12)$$

或者

$$\eta_{\mu\nu} dx^\mu dx^\nu = \eta_{\alpha\beta} dx'^\alpha dx'^\beta, \quad (\eta_{\mu\nu}) := \text{diag}(-1, 1, 1, 1). \quad (1.1.13)$$

Minkowski 率先找到了这一几何结构 (参见 [2]), 即只要定义事件间的微小距离 ds (称**时空线元**) 满足

$$ds^2 := \eta_{\mu\nu} dx^\mu dx^\nu = -dt^2 + dx^2 + dy^2 + dz^2, \quad (1.1.14)$$

就给定了一个不定度规, 称为**闵氏度规**. 更准确地, 先用 \mathbb{R}^4 的自然坐标系定义它即可. 使得闵氏度规满足 $g_{\mu\nu} = \eta_{\mu\nu}$ 的坐标系称 **Lorentz 坐标系** 或 **惯性坐标系** (简称惯性系). 这说明任意惯性系保持线元不变, 但几何学实质上支持任意光滑变换后的坐标系都应该有线元不变. 现代观点将直接把相对论的重心放在这一前提, 即闵氏度规是新时空观的体现. 毕竟时间与空间由运动坐标系关联并相互影响, 将研究背景提升至时空是自然的. 最重要的是它能反向得到狭义相对论中的所有知识, 而度规决定一套几何学, 则狭义相对论对应于一门几何学, 冠名为**闵氏几何**.

下面简要讨论它所给出的结果. 首先, 若在某系测得光速为 $\sqrt{\delta_{ij} dx^i dx^j}/dt = 1$, 换言之 $\eta_{\mu\nu} dx^\mu dx^\nu = 0 = \eta_{\alpha\beta} dx'^\alpha dx'^\beta$, 即光速不变; 其次它可体现相对性原理, 即不能区分各惯性系. 根据定义, 惯性系间按闵氏度规的等度规变换联系, 即满足

$$\eta_{\mu\nu} = \eta_{\alpha\beta} \frac{\partial x'^\alpha}{\partial x^\mu} \frac{\partial x'^\beta}{\partial x^\nu}. \quad (1.1.15)$$

由此可证惯性系变换是**仿射的** (affine):

$$\frac{\partial^2 x'^\mu}{\partial x^\sigma \partial x^\lambda} = 0 \iff x'^\mu = x_0^\mu + \Lambda^\mu{}_\nu x^\nu, \quad x_0^\mu, \Lambda^\mu{}_\nu \text{ 无关于 } x^\mu. \quad (1.1.16)$$

其中 Λ 表示 Lorentz 变换矩阵. 证明细节见 (1.2.14) 式. 这样等度规变换可写成

$$\eta_{\mu\nu} \Lambda^\mu{}_\sigma \Lambda^\nu{}_\lambda = \eta_{\sigma\lambda}. \quad (1.1.17)$$

关于矩阵的转置, 比如对于 $\Lambda^\mu{}_\sigma$, 我们像这样左右交换, 即对调行列指标为 $(\Lambda^\top)_\sigma{}^\mu$, 这样 $(\Lambda^\top)_\sigma{}^\mu \eta_{\mu\nu} \Lambda^\nu{}_\lambda$ 是指标相靠的, 相当于按此顺序乘出矩阵 $\Lambda^\top \eta \Lambda$, 此即合同变换. $\Lambda^\top \eta \Lambda = \eta$ 说明 Λ 的确正交, 当然, 数学中正交矩阵的准确定义略有不同, 即满足 $R^\top R = I$ 的 R , 区别在于 η 在时间分量上的符号不同, 我们不妨在物理学中略微扩大这个词的范围. 对其求行列式有 $|\Lambda^\top| |\eta| |\Lambda| = |\eta|$, 即 $-|\Lambda|^2 = -1$ 或

$$|\Lambda|^2 = 1. \quad (1.1.18)$$

在空间中, 正交矩阵 R 的直观意义就是旋转变换, $|R| = -1$ 时称为**空间反射**. 同理, 4 阶 Λ 也代表“旋转”, 但无非是比空间旋转还多出“时空旋转”这一情形, 也即 boost (伪转动) 变换. 取 (1.1.17) 式 $\sigma = \lambda = 0$ 就有

$$(\Lambda^0_0)^2 = 1 + \delta_{ij} \Lambda^i_0 \Lambda^j_0 = 1 + \sum_{i=1}^3 (\Lambda^i_0)^2 \geq 1. \quad (1.1.19)$$

这里括号外的“2”当然就表示平方了. 可见我们能按 $\Lambda^0_0, |\Lambda|$ 的符号将所有满足 (1.1.17) 的 Λ 分为四类. 容易看出, Λ^0_0 的符号代表时间是否反演, 因此 $\Lambda^0_0 > 0$ 的 Λ 称**正时的**, 否则为**负时的**; $|\Lambda|$ 的符号代表整个时空的坐标系是否变成左手系 (约定最初是右手系), 故 $|\Lambda| = 1$ 时称**固有的** (proper), 否则是**非固有的**. 进而再就四种情形分别讨论:

$\Lambda^0_0 \backslash \Lambda $	≥ 1	≤ -1
1	正常时空	时空翻转
-1	空间反射	时间反演

满足 (1.1.17) 的仿射变换皆可称 **Poincaré 变换**. 固有、正时的线性 Poincaré 变换称为 **Lorentz 变换**¹. 这些定义对逆变换也合理 (见 1.6 节). 取适当条件可进一步证明 boost 变换.

现在, 我们能将相对性原理、光速不变整合为更严格的说法: 物理定律在 Lorentz 变换下保持数学形式不变, 且所有与物质状态无关的标量常数保持同样的值. 比如 $c = 1, \pi/2$ 和约化 Planck 常量 \hbar . 方程的形式不变称为**协变性** (covariance). Lorentz 协变性和这种常数不变合称 **Lorentz 不变性** (invariance). 是的, 不仅要求 Lorentz 协变性, 实际上的确还额外默认常数不变, 例如, Maxwell 方程的协变性只涉及其外观, 但并不真正给出光速不变. 进一步地, 几何学支持 Lorentz 协变性的推广, 即物理方程在任意坐标变换下形式不变, 称**广义协变性**.

量子场论教材常选择 $(\eta_{\mu\nu}) = \text{diag}(1, -1, -1, -1)$ 以直接使 $ds^2 > 0$. 定义**号差** (signature) 为 $(\eta_{\mu\nu})$ 的迹, 则该选择的号差为 -2 , 戏称**西海岸度规**. 本书一般选择 $+2$ 即**东海岸度规**. 更有甚者选择将时间分量排在最后作为第 4 分量, 这样有 $\text{diag}(1, 1, 1, -1)$ 但并不影响号差. 无论怎样这些选择纯粹只是习惯. 也有一种可能有哲学区别的习惯, 称 **Wick 技巧**: 令 $x^0 = ict$ 则 $ds^2 = \delta_{\mu\nu} dx^\mu dx^\nu$, 这样线元仿佛还原回勾股定理, boost 变换变成了 $x-ict$ 平面的旋转变换. 一般除讨论某些理论复结构方面的哲学外不使用这种做法. 推广到高维情形时, 可依旧按东海岸习惯规定 $(-1, 1, \dots, 1)$. 西海岸则取 $(1, -1, \dots, -1)$.

¹又常用“齐次”“非齐次”分别指代线性和仿射, 故也可强调地称**正齐次 Lorentz 变换**.

1.1.3 矢量

物理学自 Newton 时代就开始使用标量和矢量等来表达物理量, 自场论语言发展后还使用了标量场和矢量场. 在 \mathbb{R}^3 中我们总是用一个有向线段或一个箭头来看待矢量, 并用一个像 \mathbf{A} 这样的粗体字母或 \vec{A} 这样的箭头标注来表示. 这一观点不仅使我们对矢量有一个很好的直观图像, 并且当我们要表示矢量的恒等式, 特别是包括旋度的恒等式以及矢量积分的恒等式时, 这一观点是非常方便的, 例如

$$\begin{aligned}\nabla \times \nabla f &= 0, \quad \nabla \cdot \nabla \times \mathbf{A} = 0; \\ \nabla \times (\mathbf{B} \times \mathbf{v}) &= (\nabla \cdot \mathbf{v})\mathbf{B} - (\nabla \cdot \mathbf{B})\mathbf{v} + (\mathbf{v} \cdot \nabla)\mathbf{B} - (\mathbf{B} \cdot \nabla)\mathbf{v}; \\ \int_{\partial A} \mathbf{E} \cdot d\boldsymbol{\ell} &= \int_A \nabla \times \mathbf{E} \cdot d\mathbf{S}, \quad \int_{\partial \Omega} \mathbf{E} \cdot d\mathbf{S} = \int_{\Omega} \nabla \cdot \mathbf{E} d^3x.\end{aligned}$$

这种表述直观且简洁, 而将其展开为分量语言就略显笨拙些, 但实际上, 矢量总是需要和其分量联系起来: $\mathbf{A} = A^i \mathbf{e}_i$, 其中 $\{\mathbf{e}_i\}$ 是坐标系的正交归一基. 每一矢量都用其基底去定义, 但这又将涉及基底的定义问题. 因此在以前的学习中, 我们只能以数组的方式定义矢量及其基底, 如 $\mathbf{e}_1 = (1, 0, 0)^T$.

这种图像似乎在曲面、高维等情况下遇到了障碍. 把矢量看作是直线段的实质是利用作差得到矢量, 但假设背景是一个球面, 则两点连线将位于该球面外; 而在高维空间, 我们没办法对其中的“线段”有朴素的想象, 除非借助降维画图或直接抽象地使用计算符号; 最终, 我们还将面对无穷维的矢量, 而这个看似难以接受的概念却已被广泛应用于量子物理. 读者或许能率先想到一种解决方案, 即只利用邻近两点作差, 但如果欲得到一个大小可观的矢量, 我们还需要将该微小差距比上另一个同阶无穷小量, 这意味着我们需要用求导来产生矢量. 设背景 M 上某条曲线 C , 在某坐标系下表为 $x^i \circ C(\sigma)$. 若参数式是光滑的, 则可确定 C 上关于 σ 的**切矢场** (tangent vector field), 其在此坐标系下的分量是 $T^i := dx^i \circ C / d\sigma$. 根据全微分, 它能具备物理上的广义协变性. 正如常常混用导数和导函数一样, 矢量和矢量场也经常混用, 读者可根据上下文自行判断. 称 T^i 是切矢, 准确是想称等价类 $[T^i] = \{T^i, T'^i, \dots\}$ 是切矢, 读者可类比相似矩阵的定义. 全微分关系给出

$$T'^i = \frac{\partial x'^i}{\partial x^j} T^j. \quad (1.1.20)$$

对 p 处只要穿过不同方向、不同参数化的曲线就能给出不同切矢, 这样全体可能的切矢将构成一个集合 $T_p M$, 称为 M 在 p 处的**切空间**. 可见确实有 $\dim M = \dim T_p M$. 一个直观的图像是把它理解为背景在该点的切平面. 虽然舍弃了其它可能性, 并且一点的矢量确实被绑定在此处而不能像以前那样随意移动, 但这一做法是可以接受的: 实际上, 相对论中不会使用位矢概念, 所有的矢量都是切矢, 换言之, 对我们用到的任何矢量, 总能找到一条合理参数化的曲线与之相切; 并且, 我们研究的是场, 也即

给每一背景点赋予一种物理量, 因此这些矢量的确是有起点的, 没有以前那种平移等同性. 规定称式 (1.1.20) 这种变换为**逆变的** (contravariant), 按这种方式变换的矢量 V^i 称为**逆变矢量** (contravector), 故上指标又称**逆变指标**. 当然, 必须承认, 前文“所有矢量都是切矢”的论断不能严格正确——既然有了逆变矢量, 那也没有理由排除变换与之互逆的矢量:

$$\omega'_i = \frac{\partial x^j}{\partial x'^i} \omega_j, \quad (1.1.21)$$

我们称之为**协变矢量** (covector) 或**余切矢** (cotangent vector)¹, 下指标可称**协变指标**. 这一称呼与“协变性”有区别, 后者是指方程的形式不变. 没有物理方程“逆变性”的说法. 称它们互为**对偶矢量** (dual vectors). 协变矢量的变换不是逆变的, 不可能位于切空间中, 因此全体协变矢量将构成另一截然不同的集合, 称为**余切空间**. 然而, 我们确实有办法——也将马上介绍这一办法, 来找到切空间和余切空间之间的**同构** (线性双射), 进而总能将协变矢量对偶回逆变矢量. 这样看“所有矢量都是切矢”的说法又无伤大雅.

把矢量仅仅看成是其分量的等价类的观点, 虽然在物理和计算上没有问题, 但似乎还是缺了些什么. 我们期望的是如果有分量, 就应该有某种数学对象, 使得分量是这一数学对象在不同坐标系下的不同转述. 也就是说, 我们还是希望能像以前所学的 $\mathbf{A} = A^i \mathbf{e}_i$ 形式那样去表述矢量, 换一组基底就可给出另一组分量. 鉴于对切矢图像的充分观察, Cartan 率先给出了这一数学对象的构造: 既然切矢的图像是与该曲线相切, 而数学理论又常常使用映射来略微抽象地定义新概念, 那么, 何不把切矢直接定义为“导数”本身? 例如, 在 \mathbb{R}^3 中, 给定一个光滑标量场 $f: \mathbb{R}^3 \rightarrow \mathbb{R}$, 某点处的矢量 \mathbf{A} 总是与 f 在该点的方向导数 $\mathbf{A} \cdot \nabla f$ 一一对应. 这样对高维背景, 可以联合矢量 A^i 而构成算符 $A^i \partial_i$, 其中 $\partial_i := \partial / \partial x^i$ 只是偏导的简记, 然后我们就把 $A^i \partial_i$ 称为 A . 导数映射的运算性质酷似矢量代数, 它可相加也可数乘, 还可定义零元使 $0(f) = 0$. 这样 $\{\partial_i\}$ 的确是线性无关的: $A = 0 \Rightarrow A^i = A^j \partial_j(x^i) = A(x^i) = 0$. 因此 ∂_i 形似基底, 称为第 x^i 个**坐标基矢** (coordinate basis vector). $\{\partial_i\}$ 称为此坐标系的**坐标基**. 这符合线性代数中所学的线性空间的一般定义, 我们将其罗列于此以防读者遗忘: 实线性空间是指其中定义了加法、数乘两种封闭运算的集合 V . 具体地, 加法应满足交换律²、结合律, 且有零元存在性和负元存在性:

$$\forall x \in V, \quad \exists \bar{0} \in V \text{ 使 } x + \bar{0} = x, \quad \exists \bar{x} \in V \text{ 使 } x + \bar{x} = \bar{0};$$

数乘应满足结合律、单位元存在性 ($1x = x$) 和分配律:

$$\forall x, y \in V, \alpha, \beta \in \mathbb{R} \quad \alpha(x + y) = \alpha x + \alpha y, \quad (\alpha + \beta)x = \alpha x + \beta x,$$

¹ 简称余矢, 其英文与协变矢量皆为 covector.

² 事实上可从其余七条推出加法交换律, 因为数乘分配律暗含了交换性.

而 \mathbb{R} 自己的运算按通常方式定义. V 中的元素称为矢量或向量. 按此定义切矢是矢量, 但相比于作差的形式, 导数映射的关键性优势就是它只借助于光滑标量场在这一点处的性质. 甚至只要求 f 在该点领域有定义. 若不同函数在邻域内的取值一致 (等价类称为**函数芽**), 则给出的切矢也一致. 根据链式法则, 不同坐标系的坐标基具有协变关系, 故又称**协变基**, 进而可得 (1.1.20) 式. 与曲线及其参数相关的那类切矢 (称之**曲线切矢**) 其实就是沿线导数映射:

$$\frac{d}{d\sigma}(f) = \frac{dx^i}{d\sigma} \partial_i(f), \quad (1.1.22)$$

比如, 给定坐标系后, 定义**第 x^i 坐标线** (coordinate line) 是指除 x^i 外的坐标皆为常数的曲线, 进而就可看作是以 x^i 为参数的坐标选择有关的曲线, 其曲线切矢正是坐标基 $\delta_i^j \partial_j = \partial_i$.

总之, 在 Cartan 的理论中, 所有东西都是泛函或算符, 甚至是更多层嵌套. 而映射嵌套早已成为构建数学理论的常见手段. 比如, 当今线性代数从已有矢量构造其对偶的方法是基于如下观察: 线性空间 V 上的全体线性泛函 $\omega: V \rightarrow \mathbb{R}$ 构成的集合 V^* 也是线性空间, 因为可定义自然的加法、数乘运算及零元. 于是称 V^* 为**对偶空间**. 比如, 行向量空间就是列向量空间的**对偶空间**. 为给 V^* 构造基底, 必须从原来的 V 中产生. 方法是, 假设 $\{e_i\}$ 是 V 的一组基底, 则定义 V^* 中有一组元素 $\{e^i\}$ 满足

$$e^i(e_j) := \delta_j^i, \quad (1.1.23)$$

进而由线性性可导出它对 V 任意元素的作用情况, 即 $e^i(A) = A^i$. 现在只需证明 $\{e^i\}$ 确实是一组基: 对任意 $\omega \in V^*$ 设

$$\omega_i := \omega(e_i), \quad (1.1.24)$$

则必能展开成 $\omega = \omega_i e^i$, 因为两边作用于任意 V 中元素即可, 如 $\omega(e_j) = \omega_i e^i(e_j) = \omega_j$; 线性无关也非常显然: $0 = \omega_i e^i(e_j) = \omega_j$. 这足以说明 $\dim V^* = \dim V$. 按此关系给出的 $\{e^i\}$ 就称为 $\{e_i\}$ 的**对偶基**. 这样, 只要代入 $V = T_p M, e_i = \partial_i$ 就能利用 ω 的线性性得到 (1.1.21) 式, 故余切空间可以记作 $T_p^* M$. Cartan 还注意到余切空间可给函数微分一个更严格的理解. 标准分析学中常将 df 定义为 Δf 的线性主部, 即 $df(x, \Delta x) = f'(x) \Delta x$. 物理上的理解是在 Δx 很小时给出的微小增量. 由于 Δx 是从某点 x 出发的位移, 因此一个更合适的理解是切矢, 即

$$df(A) := A(f) = A^i \partial_i f. \quad (1.1.25)$$

可以看到 x^i 的微分正是坐标基 ∂_i 的对偶 $dx^i(\partial_j) = \partial_j(x^i) = \delta_j^i$, 称为**对偶坐标基**或**逆变基**. 故 df 的分量是 $df(\partial_i) = \partial_i f$, 即可展开为 $df = \partial_i f dx^i$. 在标准分析中, 全微分理解为各线性映射间的关系, 这里将线性映射叙述为余切矢量. 由于分量是

$\partial_i f$, 因此有时建议称标量场的**梯度**¹. 那么, 对偶操作是否有必要继续套娃下去? 设 $Q \in (T_p M)^{**}$ 在 $\{dx^i\}$ 的对偶下的分量为 $Q^i = Q(dx^i)$, 则又回到了 (1.1.20) 式. 这意味着可将它等地地视作切矢. 真正有用的是 $T_p M, T_p^* M$, 好似两面镜子可以相互反射. 究其原因, 坐标变换只有逆变、协变两种方向. 线性代数的理解是: 用基底展开后有 $\omega(A) = \omega_i A^i = A^i \omega_i$, 说明如果认为映射可交换, 矢量也应按某种办法成为对偶矢量的线性泛函, 只需定义同构 $i_V : V \rightarrow V^{**}$ 使得 $i_V A$ 满足 $i_V A(\omega) := \omega(A)$, 有时 $i_V A$ 简记作 A^{**} . 这称为**自然同构**或**典范同构**. 将 V, V^{**} 做认同后, 省略同构符号可写成 $\omega(A) = A(\omega)$, 这说明对偶矢量与矢量之间有天生的乘法. 现在, Cartan 理论、Riemann 几何等被总结为**微分几何** (differential geometry).

这套框架是如此抽象, 以至于很多人成为了 Cartan 及其后世的信徒, 常常声称其优势在于“一点都无需直接借助坐标”, 即 Cartan 关于矢量的定义是“几何”的、“坐标无关”的. 这一论断当然过分夸张. 背景是无特征且无定形的, 其中的点和曲线是不能先验识别或定义的, 除非利用坐标系. 比如, 固然希望抽象地定义标量场 f , 但实际上准确定义又总要使用坐标系 ψ , 通过 n 元函数 $f \circ \psi^{-1} : \mathbb{R}^n \rightarrow \mathbb{R}$ 就可确定 $f : M \rightarrow \mathbb{R}$, 且其光滑性也必须由 $f \circ \psi^{-1}$ 确定. 或许读者对此有这样的想法: 可通过叙述发生在一个点的一个事件来确定该点, 即“此处何物发生何事”. 但这种方式只有对该事件的目击者才有意义, 缺乏普遍的可传达性, 而可传达性是科学数据的显著特点. 若该目击者试图把事件发生的位置转告一位调查者, 或者试图为后代写下一段记录, 则不得不用明确的或含蓄的坐标 (如某栋大楼的正北一百米) 来描述这一位置. 实质上, 坐标给了我们确定点的操作程序, 这是叙述性方法所做不到的. 叙述性方法只对那些有实际发生的事件的时空点才有效, 但大多数时空点是空的². 坐标空间是一个连续的、不可数的集合, 而实际事件的集合是分离的、有限归纳的集合. 因此不可能认为这两种集合是等同的. 目前物理学更接受坐标方法. 这样, 背景从来不可能真正脱离坐标. 这种所谓的“坐标无关”, 前文已经给出准确陈述: 坐标选择无关. 比如, 只要先按某 ψ_1 定义好 f 后, 对任意坐标系 ψ_2 都有

$$f \circ \psi_1^{-1}(\psi_1(p)) = f(p) = f \circ \psi_2^{-1}(\psi_2(p)), \quad (1.1.26)$$

因此 f 在 p 处该取何值还是何值, 相当于 f 直接地给 p 赋予实数 $f(p)$; 只要给定某个 $f \circ \psi^{-1}$ 的光滑性, 由坐标变换的光滑性, f 的光滑性对任意坐标系成立. 此外, 物理学对所使用的数学工具还提出要求, 即总要把理论计算与实验结果相比较, 这样坐标是必不可少. 实验者不能直接去测量一个抽象的矢量而只能测量分量, 它通过仪器的性质被我们提取为实数信息. 因此微分几何只不过是一个巧妙的框架, 没有原则上胜过分量语言. 物理学多数还是采取分量语言, 此时将逆变矢量和协变矢量

¹为区别于标准分析学中的概念, 部分书籍会用粗体表示这些泛函, 如 $\mathbf{A}, \boldsymbol{\omega}, \boldsymbol{\partial}/\partial x^i, \mathbf{d}f$. 但其实没有必要, 新旧概念仍然等同. 看似不同是因为本科的微积分教学为了通俗性, 默认以坐标空间为基础, 而没有额外提出背景概念.

²进一步的讨论见 1.4.7 节.

看成 (1.1.21) 和 (1.1.20) 式所约束出的等价类即可. 只有在概念上遇到重大歧义时, 我们才会讨论严格的数学.

1.1.4 张量

同一度规 $ds^2 = g_{ij}dx^i dx^j$ 在不同系有不同的 g_{ij} , 因此度规也有一种 Cartan 式的准确提法, 尽管物理学用等价类 $[g_{ij}]$ 去看待已经足够, 其间按式 (1.1.8) 相联系. 注意度规是用来定义线元的, 而线元的积分就是路程, 也即“速率”对“时间”的积分, 这涉及矢量的模长; 其次还用来衡量正交性, 比如坐标系也即坐标基底的正交. 故定义度规可从定义**内积** (inner product) 出发, 即它是两个矢量的线性函数 $g: V \times V \rightarrow \mathbb{R}$. 首先有对称性 $g(A, B) = g(B, A)$; 其次, 若 $\forall B \in V$ 都有 $g(A, B) = 0$, 我们希望这只在 $A = 0$ 时才能发生. 这称为**非退化性**. 规定度规的坐标分量是 $g_{ij} := g(\partial_i, \partial_j)$, 这样立即得出了 (1.1.8) 式. 这样 g_{ij} 可以看成 g 在某坐标系下具体表达出的若干实数, (g_{ij}) 是 g 的一个分量矩阵, 且是对称的: $g_{ij} = g_{ji}$. 内积的分量写法是 $g(A, B) = g_{\mu\nu}A^\mu B^\nu$. 定义 A 的**模方**为 $g(A, A)$, 这样曲线 $C(\sigma)$ 上任意两点 $C(\sigma_1), C(\sigma_2)$ 间的线长为

$$\ell(C) := \int_C \sqrt{|d\ell^2|} = \int_{\sigma_1}^{\sigma_2} \left| g\left(\frac{d}{d\sigma}, \frac{d}{d\sigma}\right) \right|^{1/2} d\sigma. \quad (1.1.27)$$

这里考虑到 $d\ell^2$ 可能为负. 若恰以 ℓ 为参数则称**线长参数**, 则易证切矢模方恒定且归一: $g(\frac{d}{d\sigma}, \frac{d}{d\sigma}) = \pm 1$. 使切矢模方恒定的参数 λ 统称**仿射参数** (affine parameter), 容易证明仿射参数总可表为线长参数的线性函数 $a\ell + b$. 内积为零则称两矢量**正交**. 这样正交归一的坐标基就是使 $g(\partial_i, \partial_j) = \tilde{g}_{ij}$. 度规可以展开成线性组合, 这涉及从矢量的基底构造一种新基底, 将运算符记作 \otimes , 并期望度规能表示成 $g = g_{ij}dx^i \otimes dx^j$. 为寻找其合理定义, 尝试两边作用于切矢, 则式左是 $g_{ij}A^i B^j = g_{ij}dx^i(A)dx^j(B)$, 式右是 $g_{ij}dx^i \otimes dx^j(A, B)$, 故直接希望 $dx^i \otimes dx^j(A, B) := dx^i(A)dx^j(B)$, 即 \otimes 的结果与相应矢量的作用相当于各自分开作用再相乘. 电动力学、量子力学中的**并矢** (dyadic) 其实也是这种运算, 如 $\boldsymbol{vw}, |\psi\rangle\langle\phi|$ 等, 即 $A \otimes B(\omega_1, \omega_2) := A(\omega_1)B(\omega_2)$, 这样的分量是只是简单拼凑 $A^i B^j$. 给背景每一点指定一个度规就形成了**度规场**. 总需先借助一个坐标系才能说清度规分量的定义, 比如欧氏度规可用自然坐标系写作 $\delta = \delta_{ij}dx^i \otimes dx^j$, 而闵氏度规是 $\eta = \eta_{\mu\nu}dx^\mu \otimes dx^\nu$. **欧氏空间** \mathbb{E}^n 就是 \mathbb{R}^n 配以欧氏度规, 记法上可将二者打包成 (\mathbb{R}^n, δ) . **闵氏时空** 就是 $\mathbb{R}^{1,3} := (\mathbb{R}^4, \eta)$.

的确, 所有东西现在都成了切矢、余切矢的函数, 只不过可以有多个输入端口, 且关于每个变量都线性, 称为**多重线性函数**. 可类比从单元函数拓展到多元函数. 如图 1.6, 设想度规是一台具有两个槽 (slat) 的售货机, 往槽中投入足够数量的一元纸币可获得一件商品 (比如一听可乐). 设某件商品需要两元纸币, 则投入过程中, 显

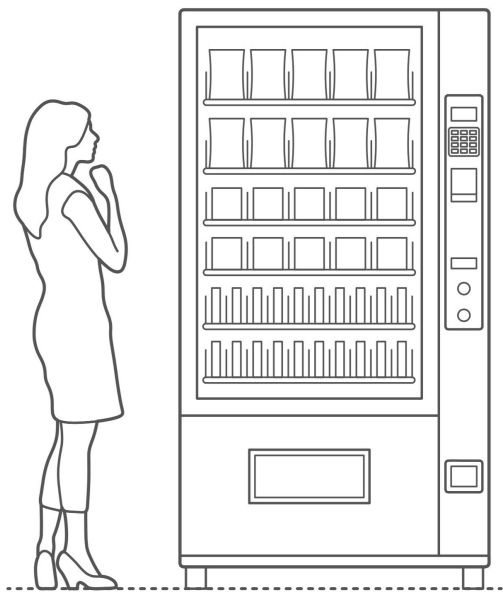


图 1.6: 售货机

示屏会不断改变数值以提示用户待支付的剩余金额，直至最终获得商品。这个过程类似于：

$$g(\cdot, \cdot) \implies g(A, \cdot) \in V^* \implies g(A, B) \in \mathbb{R}.$$

关键在于，第一步输出的 $g(A, \cdot)$ 将和之后输入的切矢 B 线性地作用，最终输出实数，故相当于 $g(A, \cdot) \in V^*$ 。类似地也有 $g(\cdot, B) \in V^*$ 。可见， g 其实给出了 V, V^* 之间的一种同构：

$$A_b := g(A, \cdot) = g(\cdot, A) \implies A_i := g_{ij} A^j = g_{ji} A^j. \quad (1.1.28)$$

后者是分量上的简写。注意分量语言是实数等式，写成 $A^j g_{ij}$ 等无所谓。因此 $g_{ij} A^i B^j$ 可简洁地写作 $A^i B_i$ 或 $A_j B^j$ ，这就将内积和“天生的乘法”等同了起来。可证 V, V^* 间不存在这种自然同构（见第 2 章），其间的同构总是人为的。换言之，欲给出一个自然的同构，需要依赖其它东西。约定使用度规干这件事。这个同构关系称为**度规对偶**。从分量上看，度规相当于给出了**降指标**操作。对切矢来说，其度规对偶称为**协变形式**。那么如何对偶回去给出**升指标**呢？应该期望有一种“上指标”度规，而对偶回去又应保持原型，因此升降指标是 V 上的恒等变换：

$$g^{ik} g_{kj} = g_{jk} g^{ki} := \delta_j^i, \quad (1.1.29)$$

换言之 (g^{ij}) 是 (g_{ij}) 的逆矩阵（且对称），而可逆性显然受到非退化性保证。比如对

闵氏度规（取惯性系），很容易证明 $(\eta_{\mu\nu}), (\eta^{\mu\nu})$ 是相等矩阵，即

$$\eta^{\mu\nu} = \delta^{\mu\alpha} \delta^{\nu\beta} \eta_{\alpha\beta}. \quad (1.1.30)$$

定义**度规逆**是 $g^{-1} := g^{ij} \partial_\mu \otimes \partial_\nu$ ，称 $\omega \in T_p^* M$ 的**逆变形式**是 $\omega^\sharp := g^{-1}(\omega, \cdot)$ ，分量写法是 $\omega^\mu = g^{\mu\nu} U_\nu$ 。

这种多重线性函数是否可以混合地输入矢量、对偶矢量？就好比售货机可以混合地投入硬币和纸币，只是对应的槽不同。比如， V 上的线性变换就是输入矢量且输出矢量，而矢量是对偶矢量的函数，因此线性变换又可以看成混合输入的函数。恒等变换 I 就是一例，它给出 $\delta_j^i A^j = A^i$ ，而 δ_j^i 单独看有上下槽。这种推广无非是将定义域扩展到若干个 V, V^* 及其混合卡氏积上。我们把这个概念称为**张量** (tensor)。该名称最初由 Cauchy 提出，它在研究弹性力学时意识到，一点处质元与质元的作用力很难只由一个矢量数组取描述，而似乎需要用到矩阵。比如应力就涉及某个面上的压力，因此必定要先给定一个面法矢，这样就产生对应的一个作用力。一个 (k, l) 型**张量**（简称 (k, l) -**张量**）就是可与有限个矢量、对偶矢量作用出标量的多重线性函数：

$$T : \underbrace{V^* \times \cdots \times V^*}_k \times \underbrace{V \times \cdots \times V}_l \rightarrow \mathbb{R} \quad (1.1.31)$$

$k + l$ 叫作张量的**阶数**。设 $\{e_\mu\}$ 及对偶基 $\{e^\mu\}$ ，张量总可用基底展开：

$$T = T^{\mu_1 \cdots \mu_k}_{\sigma_1 \cdots \sigma_l} e_{\mu_1} \otimes \cdots \otimes e_{\mu_k} \otimes e^{\sigma_1} \otimes \cdots \otimes e^{\sigma_l}. \quad (1.1.32)$$

其中

$$T^{\mu_1 \cdots \mu_k}_{\sigma_1 \cdots \sigma_l} := T(e^{\mu_1}, \dots, e^{\mu_k}; e_{\sigma_1}, \dots, e_{\sigma_l}), \quad (1.1.33)$$

这里用一个分号区分上下槽。注意，常习惯于先把张量分量的上指标写完，再错开地写下指标¹。证明只需两边作用于对应基矢为

$$e_{\mu_1} \otimes \cdots \otimes e_{\mu_k} \otimes e^{\sigma_1} \otimes \cdots \otimes e^{\sigma_l} (e^{\mu'_1}, \dots, e^{\mu'_k}; e_{\sigma'_1}, \dots, e_{\sigma'_l}) = \delta_{\mu_1}^{\mu'_1} \cdots \delta_{\mu_k}^{\mu'_k} \delta_{\sigma'_1}^{\sigma_1} \cdots \delta_{\sigma'_l}^{\sigma_l},$$

故 \otimes 的准确称呼是**张量积**。取坐标基之后，根据坐标基的变换得

$$T'^{\mu_1 \cdots \mu_k}_{\nu_1 \cdots \nu_l} = T^{\rho_1 \cdots \rho_k}_{\sigma_1 \cdots \sigma_l} \frac{\partial x'^{\mu_1}}{\partial x^{\rho_1}} \cdots \frac{\partial x'^{\mu_k}}{\partial x^{\rho_k}} \frac{\partial x^{\sigma_1}}{\partial x'^{\nu_1}} \cdots \frac{\partial x^{\sigma_l}}{\partial x'^{\nu_l}}, \quad (1.1.34)$$

这就是**张量变换律**。可记忆为是满足指标平衡的 Jacobi 矩阵元乘积，这样就可快速写出公式。 $(n, 0), (0, n)$ 型分别称 n 阶**逆变**、**协变张量**。可见有如下特例： $(0, 0)$ -张量称 0 阶**张量**，也就是标量；逆变矢量是 1 阶逆变张量，协变矢量是 1 阶协变张量；

¹部分量子场论教材选择统一使用下指标而只强调求和约定“总共重复两次”，更有甚者不区分指标顺序。这在讨论更深刻问题时出现缺陷：前者不区分逆变、协变，后者说不清升降指标后的位置。建议只在特殊情况采取简化，比如 δ_ν^μ 是考虑到 I 是对称矩阵；之后学的克氏符 $\Gamma_{\mu\nu}^\lambda$ 是考虑到一般不会混淆且它也不是张量。

度规是对称、非退化的 2 阶协变张量；Kronecker 符号是 2 阶张量。由张量的变换律，若存在坐标系使其分量为零，其在任何坐标系下分量都为零，故不论何型都称**零张量**。

张量的基本运算可总结如下：1. 一般地， $(k, l), (k', l')$ 型张量 T, T' 的张量积 $T \otimes T'$ 是一个 $(k + k', l + l')$ -张量，使得

$$\begin{aligned} T \otimes T'(\omega^1, \dots, \omega^k, \omega^{k+1}, \dots, \omega^{k+k'}; v_1, \dots, v_l, v_{l+1}, \dots, v_{l+l'}) \\ := T(\omega^1, \dots, \omega^k; v_1, \dots, v_l) T'(\omega^{k+1}, \dots, \omega^{k+k'}; v_{l+1}, \dots, v_{l+l'}). \end{aligned} \quad (1.1.35)$$

需注意张量积有顺序，一般不满足交换律。在分量上，张量积就是直接拼凑，比如将 T^μ 和 $S_{\lambda\sigma}$ 拼成 $T^\mu S_{\lambda\sigma}$ ，其变换式只是将各自的 Jacobi 矩阵元按实数的规律乘起来。2. (k, l) -张量 T 的第 i 上标 ($i \leq k$) 与第 j 下标 ($j \leq l$) 的**缩并** (contraction)¹为

$$C_j^i T := T\left(\cdot, \dots, \overset{\text{第 } i \text{ 上槽}}{\color{red}e^h}, \cdot, \dots; \cdot, \dots, \underset{\text{第 } j \text{ 下槽}}{\color{blue}e_h}, \cdot, \dots\right), \quad (1.1.36)$$

其中注意要对 h 求和。缩并结果属于 $(k - 1, l - 1)$ -张量。在分量上，这无非是将张量中一对指标换为哑标。它无论发生在张量积的不同张量间（如 $F^\mu{}_{\sigma\nu} W^\nu$ ），还是发生在单个张量中（如 $G^\nu{}_{\lambda\nu\sigma}$ ），变换式只是消去一对 Jacobi 矩阵元。3. 若干张量线性相加仍是张量。映射语言上就是指同型张量组合（如 $T + S$ ），显然在分量上保证了指标平衡。物理学更注重分量语言，常常把满足张量变换律的物理量统称张量，映射语言则以张量的分量等价类以区分之。张量的映射语言虽能帮助我们严格构造更高指标的物理量，但并非所有量都满足张量变换律，这种量通常称**赝张量**，它们也有类似于张量的运算，但结果不是张量。对于赝张量，映射语言需要强行定义其张量映射表示，显得非常累赘；分量语言反而成了聪明的选择。当然，映射语言有时能便捷且优美地推导某些定理，对其余情形，本书一律使用分量语言。

1.1.5 时空结构

坐标系、参考系这两个概念有微妙区别。欲研究某物理过程，总借助参考系来建立某种坐标系，进而用坐标进行观测与记录。实际的参考物应是实验仪器或观测人员。无论是仪器还是人，先抽象成理想的质点模型，称为**观者** (observer)，其仅在时空中划过一条世界线。所谓观测，就是持有空间、时间的测量工具（尺子和钟），方才能对事件赋予坐标。质点总能以自身为参考，视自己静止，因此其确定出的局域的坐标系要规定空间坐标恒定，故世界线上的线元将在该系下只剩下时间分量。这也

¹ 缩并其实是求迹的映射语言，在数学上记作 tr 。但两套符号的逻辑稍许不同： tr 的逻辑是写出缩并后张量类型， C 的逻辑是指明所缩指标位置。假设 T 是一个 $(1, 3)$ -张量，定义 $S = C_2^1 T$ ，分量写法是 $S_{ij} = C^k{}_{ik} T_j{}^k$ ，但数学上记作 $S = \text{tr}_2^0 T$ 。并且 tr 的用法有时并不唯一，比如 S_{ij} 求迹必须借助升指标得 $g^{ij} S_{ij}$ ，数学上写作 $\text{tr}_g S$ 。

说明质点世界线可用该系的坐标时进行参数化. 曲线的参数化表征着运动, 换言之, 此坐标时是质点自身演化的象征. 满足这种要求的坐标时也许很多, 不过为使系数一致, 可取

$$d\tau^2 := -ds^2 = dt^2 - dx^2 - dy^2 - dz^2 \implies \gamma = dt/d\tau, \quad (1.1.37)$$

这个 τ 就称为**固有时参数**, 简称**固有时** (proper time). 若观者所持的钟以其固有时为读数, 这样就能衡量自然演化, 称**标准钟** (standard clock). 总规定它是均匀流逝、走时准确的. 钟的标准性只对其**走时率** (rate) 提出约束, 而对**初始设定** (setting) 即固有时的零点选取未做要求. **惯性观者**就是世界线与某一惯性坐标系 t 轴重合的观者, 其世界线总是一条直线. 这和经典力学意义的惯性观者相相应, 即在所引进的惯性系中, 自由粒子之轨迹或是一点, 或是一条以不变速度运动的直线. 在相对论中, 除研究质点的世界线外还可讨论**光子** (photon) 的世界线, 因此通常将二者统称为**粒子** (particle), 注意区别于粒子物理学中的粒子. 相对论中, **质量**是指粒子在相对静系所测得的质量. 质点又可称**有质粒子**.

我们遇到的第一个矢量从速度开始. 设质点世界线 $C(\tau)$, 其**4 维速度**是

$$U^\mu := dx^\mu/d\tau, \quad (1.1.38)$$

Cartan 写法是 $U := d/d\tau$, 简称**4-速**. 4-速的模方满足 $U_\nu U^\nu = -1$, 可见 4-速模方在世界线上恒定. 取线上任意两事件 $C(\tau_1), C(\tau_2)$, 其间所经历的 (elapsed along) **固有时间**为

$$T(C) := \tau_2 - \tau_1 = \int_C \sqrt{-ds^2} = \int_{\tau_1}^{\tau_2} \sqrt{-U^\mu U_\mu} d\tau, \quad (1.1.39)$$

相当于世界线线长. 但由于光速不变, 光子不存在相对静系, 故光子没有质量 (甚至不可定义), 可称**无质粒子**. 世界线现推广为就任意粒子而言的情形. 但光总满足 $ds^2 = 0$ (无论号差如何), 因此对光子也不能定义固有时 (因为光对其世界线上的事件标记为一致坐标时), 即光在物理意义上就不能作为观者. 通常会借助某个正定度规定义它的仿射参数, 这样就可以描述光子世界线. 观者首先得是一个质点.

借助时空图可直观理解 boost 变换. 类似于去规定某正交归一系的右手性, 通常先规定某个观者的 4-速所指方向是时空的**未来**, 反方向为**过去**, 并在时空图中画成朝上的. 以 \mathbb{R}^2 上某 x, t 惯性系为基准, 直线 $t' = 0$ (x' 轴) 按照变换应是 $t = ux$. 同理 t' 轴就是 $t = (1/u)x$. 可见 t', x' 轴将关于 45° 直线对称. 以转动变换类比之, 相应的平方和关系应替换为双曲关系, 故代表轴线斜率的 u 与倾角 θ (称为**快度**) 的关系为 $u =: \tanh \theta$, 其中 “ $=:$ ” 表示 “记作”. 这样有 $\cosh \theta = \gamma, \sinh \theta = \gamma u$. “boost” 一词取时间轴上的推动之意, 但可见一个较好的称呼是**伪转动** (pseudo rotation):

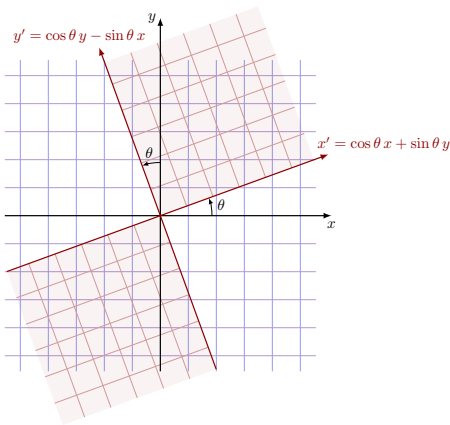
$$t' = t \cosh \theta - x \sinh \theta, \quad x' = x \cosh \theta - t \sinh \theta. \quad (1.1.40)$$

如图 1.7, 在生活中, 光速相对于我们较大, 因此图上的光锥应当“下压”, 空间轴的变化并不明显, 可见我们退回到 Galileo 变换, 即大家都共用一个绝对时. 需注意, 从图上看伪转动的坐标轴似乎并不正交, 这是因为我们还在秉持欧氏几何的直觉, 此处正交要用闵氏度规衡量.

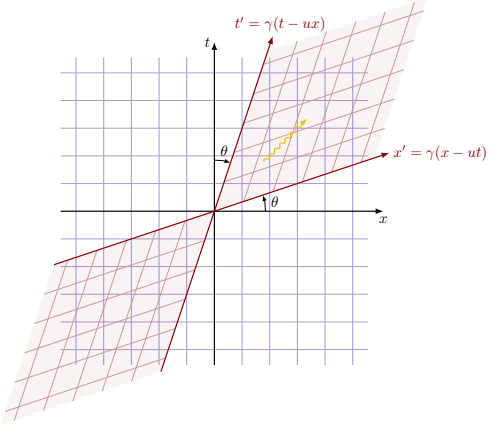
我们总感觉周围物体与自己是同时的, 这相当于截取时空而扔掉时间, 只观察一个 3 维空间. 一个坐标系中坐标时等于 t 的所有事件 (称**同时事件**) 所构成的集合 Σ_t 称为某时刻 t 的**同时面** (simultaneity surface). “面”的称呼是因为它在 3 维时空图中可用一张 2 维曲面形象表示, 但准确地是**超曲面** (hypersurface), “超”旨在强调它实际是 3 维的. 以某惯性系为基准, 同时面形如一张平面, 作某事件的同时面可投影得其坐标时. 如图 1.8, 某事件的 t 坐标可做平行于 x 轴的直线得到 (此即 S 系同时面). 而 t' 坐标可做平行于 x' 轴的直线而截得 (此即 S' 系同时面). 所以在一个惯性系同时的一对不同事件在另一惯性系中不同时. 没办法把特定的同时面区别出来, 此即**同时的相对性**的直观图像. 相比于代数计算, 可见几何语言的优雅. 注意, 这些同时面直线实际上表示 3 维空间, 故同时面交点实际是 2 维平面, 其上任意两个事件被两个观者都看做同时.

但我们总是用信号感知周遭, 无论可见光还是神经元生物电, 传播都是有限快的, 而严格来讲, 一个观者只能对发生在自己世界线上的事件做**直接观测**, 故此处的我们不能真正感受同时. 欲对一定时空范围内的事件进行观测, 必须处处设置参照物或观者 (标准钟确保每个空间点上放置的钟是全同的) 以形成参考系¹. 这就涉及如何给同时一个操作性定义. 一族世界线所组成的**线汇** (congruence) 称为**世界管** (worldtube), 换言之, 就是一个由若干世界线构成的集合 \mathcal{C} , 使一定时空范围内每一点 $\exists! \alpha_i \in \mathcal{C}$ ($!$ 表示唯一) 与之对应. 若世界管内任意世界线都是观者世界线, 则该世界管就是**参考系**, 记作 \mathcal{R} . 一系列惯性观者就构成**惯性参考系**. 通常来说世界管占据了一定时空体积, 也就是说其占据的时空点需对应四个坐标, 如对参考系而言可讨论其中点所对应的观者编号 (三维坐标) 和“固有时”. 若只需两个坐标, 则称为**世界面** (worldsheet), 可视作一条空间曲线在时空中所划过的痕迹, 它的确是 2 维的. 尽管各观者的标准钟走时率一致, 但初始设定却无特殊要求, 则应怎样合理、有用地规定同时面? 注意, 对惯性参考系而言答案很自然, 那就是选择惯性坐标系, 其具备正交性. 因此, 要从参考系确定出坐标系, 一个自然的要求是规定同时面与任意观者世界线正交, 此即**钟同步** (clock synchronization) 或**对钟**. 需要说明, 并非所有参考系都能办到良好的钟同步 (见第 2 章), 但惯性系可轻松办到, 甚至有如下操作意义上的步骤. 由光速不变, Einstein 选择借光信号对同时性进行定义. 注意, 光速不变在实验事实上应理解为双程平均真空光速, 而非单向真空光速, 也即, 应通过

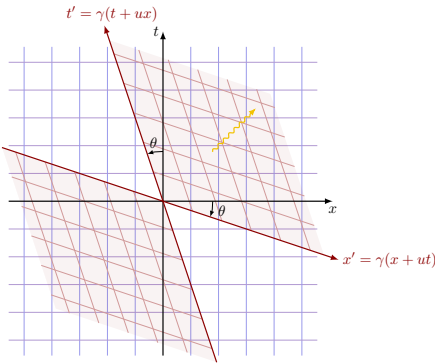
¹或依靠观者和被观测物的信息交流. 一般用光信号, 因为光速不变是优良性质. 可见光能给予人或相机以视觉图像, 这就涉及相对论视觉或图形学. 我们暂时只谈利用参考系的观测.



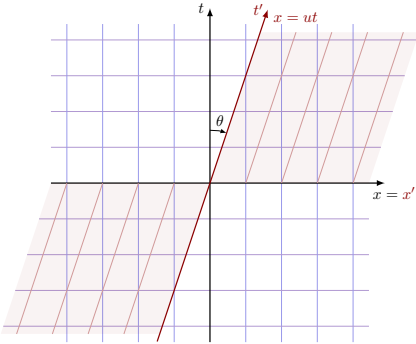
(a) 空间转动变换



(b) 时空伪转动变换

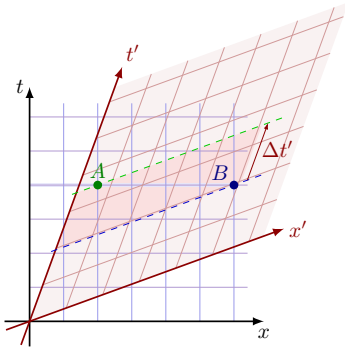


(c) 负方向伪转动

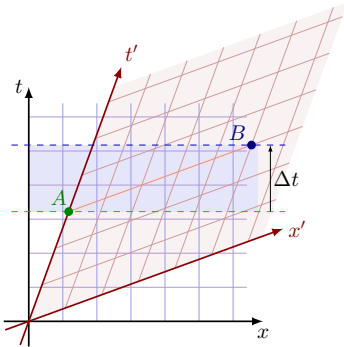


(d) Galileo 变换

图 1.7: 四类变换的几何直观



(a) A, B 在 S 中同时



(b) A, B 在 S' 中同时

图 1.8: 同时的相对性

如下的 **Fizeau 程序** 测量真空光速. 以某一观者为坐标原点, 光可相对于该坐标系在任意方向上沿直线运动. 观者在某时刻发射光, 然后由放置在某处的镜子反射, 随后观者记录光回到原点的时刻, 这样由二者距离可计算光速. 迄今为止, 单向光速从未被实验测量过, 也不可能实验测量, 除非找到光信号之外的对钟方式. 真空光速在各方向上相同是假定性的原理. 通过从原点发射的球面电磁波 (所有方向上都有光传播) 来同步各点时钟, 即事先约定某处接收到该电磁波时, 此处时钟的读数就调零. 原点处观者只要在发出、接受事件的中点处调零即可. 这样, 惯性参考系按对钟所确定出的坐标系的确是惯性坐标系, 该惯性坐标系的 t 轴与惯性参考系中某个惯性观者世界线相重合, 因此此二概念将紧密联系, 不必认真区分时可笼统称为**惯性系**.

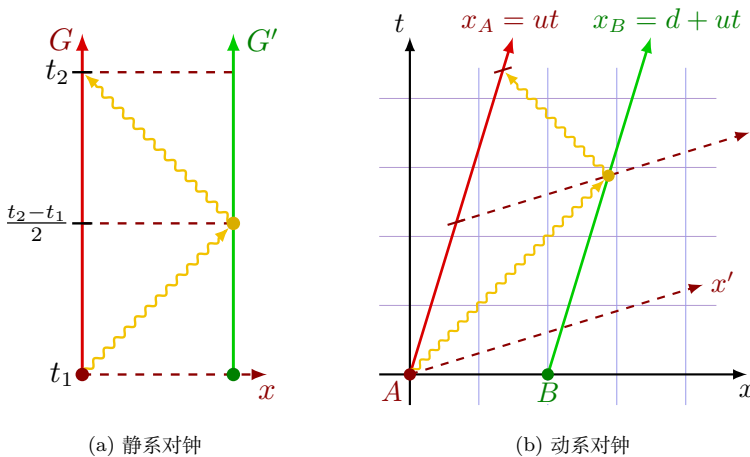


图 1.9: 对钟以确定正交同时面

在号差 +2 下, 若切矢 A 的模方为正、负或零, 那就分别称为**类空** (spacelike)、**类时** (timelike) 或**类光** (null) 矢量. 比如, 质点的 4-速是类时的, 光子的切矢是类光的. 所有类光矢量之集 $C_N \subset T_p \mathbb{R}^{1,3}$ 称为**光锥** (light cone). 注意 -2 号差习惯的模方符号刚好相反. 从三维 Cauchy-Schwartz 不等式可以看到, 若 U^μ 类时, 而 W^ν 是非零类时或类光矢量, 则当且仅当在任意按 Lorentz 变换 ($\Lambda^0_0 > 0$) 相联系的惯性系中满足 $U^0 W^0 > 0$ 时有 $\eta_{\mu\nu} U^\mu W^\nu < 0$. 因此可在光锥及其内部定义等价关系, 称 U^μ, W^ν 具有相同的**时间指向**. 容易发现存在两种等价方式. 先任意规定其中一种是**指向未来的** (future directed) 矢量, 那另一种方式指定的矢量就称为**指向过去的** (past directed). 忽略掉零矢量, 光锥被分成了两部分, 即**未来光锥**和**过去光锥**. 一般来讲总规定惯性系的 4-速指向未来, 而只在讨论某些反演对称性时才考虑指向过去的. $\mathbb{R}^{1,3}$ 中任意两个事件 x_1^μ, x_2^μ 间的直线线长 $\sqrt{|\eta_{\mu\nu} \Delta x^\mu \Delta x^\nu|}$ 称为**距离或间隔**, 可类似地分为三类. 一般而言, 不考虑切矢模方符号变化的世界线 (这涉及无穷大的

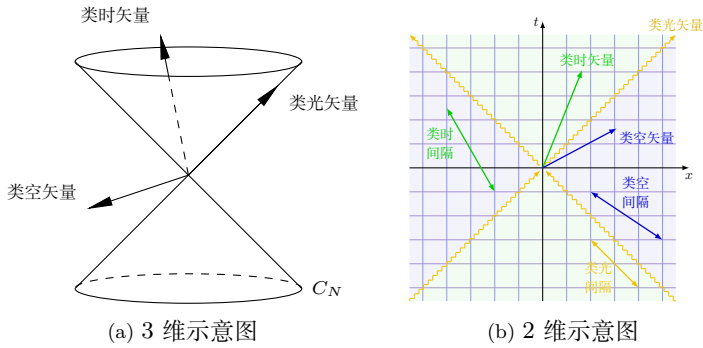


图 1.10: 光锥

作用力)，因此世界线可类似分为三类。若间隔类光或类时，则各惯性系将一致认同两个事件的先后顺序，因为 Lorentz 变换保证这种事件的四个坐标符号一致。但对于类空间隔，其事件的坐标时差可以在不同惯性系下取任意实数。即存在一些观者判定两个事件同时，另一些观者则断言 x_1^μ 先于 x_2^μ ，剩下的观者反之。这在图 1.8 和图 1.11 中一目了然。类空间隔的事件先后次序没有意义，说明类空矢量没有未来或过去的分类。综上，称能用一条类时或类光线联系的事件具有**因果联系**，在 $\mathbb{R}^{1,3}$ 中这等价于用直线间隔去联系。数学上不妨设曲线分段光滑，并使用高数里的中值定理即可。必须存在质点或光子经历两个事件才能讨论因果，换言之，因果就是信号的传递。

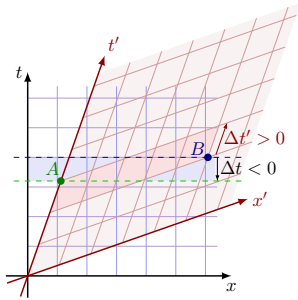


图 1.11: 类空间隔

1.1.6 广义不变性

物理中的各种原理常因其模糊的语言叙述而困扰学者。闵氏时空似乎仍在承认绝对时空的存在，一个釜底抽薪的办法是认为所有参考系皆平权，即一种“广义相对性”。这是否是指物理定律在任何参考系下形式不变呢？注意，广义协变性远比这一命题“广义”，因为其中我们允许任意的光滑坐标变换，而不要求这个坐标系是由参考系对钟所确定出的。可是，物理定律应该涉及事件的客体实在，及其间的时间、逻辑因果关系，坐标只是描述事件的记录工具，并不限制关于事件关系的定律内容。这样，对一个物理定律的描述当然要与坐标独立才能是合理的，即合理的物理定律自动独立于坐标系。换言之，应将广义协变性理解为**坐标冗余**。这太过自然以至于我们从建立物理学之初就作为了公设，或者说人类能用数学描述自然的基石，再用广义协变性的方式重申坐标独立性仿佛许许多多，因此“广义相对性”不能表述为这种说

法. 因此广义协变性没有实质的物理内容, 只对其书写的数学形式施加了限制.

一个只有 Lorentz 协变性的等式, 通过纯粹的坐标变换就可轻松地改写为有广义协变性的等式, 只要我们在正交归一系中写出公式, 再去找其它坐标系下的表述即可. 4-速能符合广义协变性, 故只需借助某正交归一系 $\{\xi^\alpha\}$ 来定义 4-速 $\tilde{U}^\alpha := d\xi^\alpha/d\tau$ 即可. 但对所谓的 **4-加速** 来说情况就微妙许多. 取正交归一系定义 $\tilde{A}^\alpha = d\tilde{U}^\alpha/d\tau = d^2\xi^\alpha/d\tau^2$, 则

$$\frac{d\tilde{U}^\alpha}{d\tau} = \frac{d}{d\tau} \left(\frac{\partial \xi^\alpha}{\partial x^\mu} U^\mu \right) = \frac{\partial \xi^\alpha}{\partial x^\mu} \frac{dU^\mu}{d\tau} + \frac{\partial^2 \xi^\alpha}{\partial x^\mu \partial x^\nu} U^\mu U^\nu,$$

可见 $d\tilde{U}^\alpha/d\tau$ 只在 Lorentz 变换下才是逆变矢量, 即只有 Lorentz 协变性而无广义协变性. 解决该问题的关键是, 将 4-加速理解为对 4-速求导是没错的, 但结果应是任意坐标下的逆变矢量, 而只是在惯性系下才还原为 $d\tilde{U}^\alpha/d\tau$ 的形式. 它们的确是一种导数, 只是在一般坐标系下不写作 $d/d\tau$, 为此不妨记 $D/d\tau$, 则

$$A^\lambda := \frac{DU^\lambda}{d\tau} = \frac{\partial x^\lambda}{\partial \xi^\alpha} \frac{d\tilde{U}^\alpha}{d\tau} = \frac{dU^\lambda}{d\tau} + \Gamma_{\mu\nu}^\lambda U^\mu U^\nu, \quad \Gamma_{\mu\nu}^\lambda := \frac{\partial x^\lambda}{\partial \xi^\alpha} \frac{\partial^2 \xi^\alpha}{\partial x^\mu \partial x^\nu}. \quad (1.1.41)$$

则 $DU^\mu/d\tau$ 就确实是广义协变的. 这个 D 称为**绝对微分**或**协变微分**, 日文所翻译的**共变**或许更无歧义. $D/d\tau$ 称为**沿线协变导数**. 换言之 $D/d\tau$ 总保持其作用对象的广义协变性, 而只在曲线上的正交归一系中还原回 $d/d\tau$. 数学上称 $\Gamma_{\mu\nu}^\lambda$ 为 **Christoffel 符号**, 简称**克氏符**, 它是典型的坐标依赖的赓张量:

$$\begin{aligned} \Gamma_{\mu\nu}^{\lambda'} &= \frac{\partial x^{\lambda'}}{\partial \xi^\alpha} \frac{\partial^2 \xi^\alpha}{\partial x'^\mu \partial x'^\nu} = \frac{\partial x^{\lambda'}}{\partial x^\rho} \frac{\partial x^\rho}{\partial \xi^\alpha} \frac{\partial}{\partial x'^\mu} \left(\frac{\partial x^\sigma}{\partial x'^\nu} \frac{\partial \xi^\alpha}{\partial x^\sigma} \right) \\ &= \frac{\partial x^{\lambda'}}{\partial x^\rho} \frac{\partial x^\rho}{\partial \xi^\alpha} \left(\frac{\partial x^\sigma}{\partial x'^\nu} \frac{\partial x^\tau}{\partial x'^\mu} \frac{\partial^2 \xi^\alpha}{\partial x^\tau \partial x^\sigma} + \frac{\partial^2 x^\sigma}{\partial x'^\mu \partial x'^\nu} \frac{\partial \xi^\alpha}{\partial x^\sigma} \right) \\ &= \frac{\partial x^{\lambda'}}{\partial x^\rho} \frac{\partial x^\tau}{\partial x'^\mu} \frac{\partial x^\sigma}{\partial x'^\nu} \Gamma_{\tau\sigma}^\rho + \frac{\partial x^{\lambda'}}{\partial x^\rho} \frac{\partial^2 x^\rho}{\partial x'^\mu \partial x'^\nu}, \end{aligned} \quad (1.1.42)$$

可见多出一项. 克氏符仅在惯性系全零而在任意坐标系下存在非零, 但 $\mathbb{R}^{1,3}$ 的测地线仍是直线, 只是在任意坐标系下直线的表达式更复杂. 因此, 两种坐标系的表述在物理性质完全一致. 按此思想甚至能写出 Newton 力学的广义协变性, 类似地改写一个仅具 “Galileo 协变性” 的方程 $d^2\xi^i/dt^2 = \tilde{f}^i/m$, 即

$$\frac{d^2 x^\lambda}{dt^2} + \frac{\partial x^\lambda}{\partial \xi^\alpha} \frac{\partial^2 \xi^\alpha}{\partial x^\mu \partial x^\nu} \frac{dx^\mu}{dt} \frac{dx^\nu}{dt} = \frac{f^\lambda}{m}, \quad (1.1.43)$$

其中**绝对时间** t 等于任何经典力学意义的惯性坐标时 x^0 , 但却不等于一般坐标系的 x^0 . 读者可自行尝

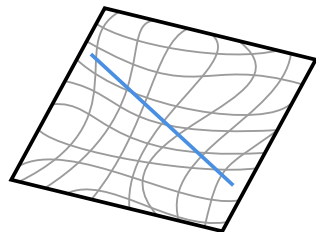


图 1.12: \mathbb{R}^2 任意坐标系及直线

试直接从 3 维语言证明,但也可预料,因为 Newton 力学是非相对论的低速近似. 这个例子也被 Kretschmann(1917) [5] 和 Weinberg [6] 提及过. 可类比地理解为惯性系的 $m\ddot{\mathbf{x}} = \mathbf{f}$ 无广义协变性, 改写坐标后有 $m\ddot{\mathbf{x}}' = \mathbf{f}' + \mathbf{f}_I$ 的形式, 这确有广义协变性, 只是取坐标系为惯性系时有 $\mathbf{f}_I = \mathbf{0}$. 假设某坐标系可对应于某参考系 (经典力学不考虑对钟), 这样就可理解为非惯性系中出现惯性力. 但这牵扯一个微妙之处: 经典力学中速度分量一般不保持 \dot{x}^i 的形式, 如二维速度 $\dot{\mathbf{x}}$ 在正交系下表为 $\dot{x}\mathbf{i} + \dot{y}\mathbf{j}$, 但极坐标系下是 $\dot{r}\mathbf{e}_r + r\dot{\theta}\mathbf{e}_\theta$. 原因是经典力学总带着基底整体求导, 速度是直接对位矢 \mathbf{x} 求导, 变换坐标系后基底要跟着换并默认总是正交归一的. 但我们是引入坐标基以保证广义协变性, 坐标基一般不正交归一.

可见广义协变性其实并不意味着定律是相对论性的. Einstein 靠着广义协变性发展了他的引力理论 (见 1.4.1 节), 故对相关批判他回应, 可认为自然似乎偏好当用坐标无关语言表述时更简单的方程, 因此引力理论将倾向于广义相对论, 而非 Newton 理论. 这一美学原则固然是主观而牵强的, 我们愿意用更严格的方式将他的想法重述——事实上, 要推广的应该是 Lorentz 不变性, 即其中不仅有广义协变性, 还要将常数推广到所谓的**绝对量** (absolute object), 即与物质状态无关的各种量, 除了常数 $c = G = 1, \hbar/2$, 最重要的是还包括更高指标的, 比如 Jacobi 矩阵元 $\partial x^\mu / \partial x'^\nu$. 与物质状态有关而可变的称**动力学量** (dynamical object), 例如粒子的位置和动量、场的场强和能量密度等. 给定任何物理方程, 按定义总能把其中的量分为绝对量和动力学量. **广义不变性**就是说, 任何物理定律有广义协变性, 且方程中的绝对量在任意坐标变换下相同, 就能决定了哪些绝对量才可以出现于方程中. 比如, 可以假设 $\Gamma_{\mu\nu}^\lambda$ 是与物质有关的动力学量 (见 1.2.2 和 1.4.1 节), 从而使式 (1.1.41) 具有广义不变性. 相比之下, 从根本上不能将式 (1.1.43) 中的偏导项理解为动力学量, 因为 Newton 力学所基础的几何形态是固定且不可改变的, 这些偏导必须视作与物质无关的绝对量, 但它又在坐标变换下可变, 因此尽管它是广义协变的, 但它不是广义不变的.

广义不变性所提出的要求高于 Lorentz 不变性. $c = 1, \hbar$ 和 $\eta_{\mu\nu}$ 有资格成为 Lorentz 不变性下的绝对量, 但 $\eta_{\mu\nu}$ 不能在任意坐标变换下严格不变, 它可变成某个 $g_{\mu\nu}$. 总之, 不变性才是严格限制物理定律的对称性. 我们对大自然有简单质朴的信念, 即在自然定律的理论公式表述中, 应当在可用的实验证据所给定的限制条件下, 使对称性最大化. 只有当实验提供了与对称性相反的证据时, 才应当考虑对称性被破坏了. 尽管自然规律本有其它可能, 毕竟不是被先验告知的, 但目前没有任何实验否认这一对称性. 我们总是想尽一切办法排除人为因素, 这才是最实质的精神. 真正给不变性提供背书的数学工具是微分几何的张量. 由张量构成的满足指标平衡的方程, 总可在坐标变换下消除 Jacobi 矩阵元. 如果描述物理规律的物理量都是张量, 则方程就能保持广义不变性. 这就是 Einstein 起初所认同的那种对广义协变性的理解, 我们已将它准确叙述为广义不变性.

1.2 张量代数

1.2.1 对称与反称

对腰张量和张量都可定义

$$T_{\dots(\mu_1 \dots \mu_k) \dots} := \frac{1}{k!} \sum_{\sigma \in S_k} T_{\dots \mu_{\sigma(1)} \dots \mu_{\sigma(k)} \dots}, \quad (1.2.1)$$

$$T_{\dots[\mu_1 \dots \mu_k] \dots} := \frac{1}{k!} \sum_{\sigma \in S_k} \text{sgn } \sigma T_{\dots \mu_{\sigma(1)} \dots \mu_{\sigma(k)} \dots}, \quad (1.2.2)$$

其中 S_k 是全体 $(1 \dots k)$ 排列之集; $\sigma \in S_k$ 为偶排列时 $\text{sgn } \sigma = 1$, 否则为 -1 ; $\sigma(k)$ 为 σ 的第 k 项. 上式分别称为 T_{\dots} 的**对称部分**和**反称部分**. 对两个指标, 我们在代数学中就学过矩阵的对称、反称分解, 比如

$$\Gamma_{\mu\nu}^\lambda = \Gamma_{(\mu\nu)}^\lambda + \Gamma_{[\mu\nu]}^\lambda, \quad \Gamma_{(\mu\nu)}^\lambda = \frac{\Gamma_{\mu\nu}^\lambda + \Gamma_{\nu\mu}^\lambda}{2}, \quad \Gamma_{[\mu\nu]}^\lambda = \frac{\Gamma_{\mu\nu}^\lambda - \Gamma_{\nu\mu}^\lambda}{2}, \quad (1.2.3)$$

对上标定义同理. 上下标之间的对称、反称性要用度规升降后再讨论. 有时会横跨地标注, 如 $V^{[\mu} W^{\nu]}$, $F_{(\mu|\sigma\lambda|\nu)}$ 都标记在了 μ, ν 指标. 这套“圆方括号”指标系统的具体性质如下. 缩并时两种括号都有传递性, 如

$$T_{(\mu_1 \dots \mu_k)} S^{\mu_1 \dots \mu_k} = T_{(\mu_1 \dots \mu_k)} S^{(\mu_1 \dots \mu_k)} = T_{\mu_1 \dots \mu_k} S^{(\mu_1 \dots \mu_k)}; \quad (1.2.4)$$

嵌套括号时, 同种子括号可随意添删, 但异种子括号会直接为零, 如

$$T_{[\mu_1 \dots [\mu_m \dots \mu_n] \dots \mu_k]} = T_{[\mu_1 \dots \mu_k]}, \quad T_{[\mu_1 \dots (\mu_m \dots \mu_n) \dots \mu_k]} = 0. \quad (1.2.5)$$

这两条证明不难, 可以留给读者. 由此还可推出异种括号缩并为零, 如

$$T_{(\mu_1 \dots \mu_k)} S^{[\mu_1 \dots \mu_k]} = 0; \quad (1.2.6)$$

对三个以上指标只能 $T_{\mu_1 \dots \mu_k} \neq T_{(\mu_1 \dots \mu_k)} + T_{[\mu_1 \dots \mu_k]}$, 但仍可在**全对称**或**全反称**时有

$$T_{\mu_1 \dots \mu_k} = T_{(\mu_1 \dots \mu_k)} \Rightarrow T_{[\mu_1 \dots \mu_k]} = 0, \quad T_{\mu_1 \dots \mu_k} = T_{[\mu_1 \dots \mu_k]} \Rightarrow T_{(\mu_1 \dots \mu_k)} = 0. \quad (1.2.7)$$

今后会经常遇到某些指标具有对称、反称性, 这往往是物理量的具体定义导致的. 因此本段讨论将简化大量运算.

1.2.2 协变导数

偏导也可记作添加逗号和下标, 如 $\omega_{\mu,\lambda} = \partial_\lambda \omega_\mu$. 相应于 $d/d\tau$ 可按全微分关系表为若干坐标偏导 ∂_ν 的线性组合, 对 $D/d\tau$ 而言亦有**协变导数** ∇_μ 的概念¹. 一般只

¹量子场论也有协变导数概念, 记作 D_μ , 但定义在表面上略有差异, 其间联系待第 7 章.

讨论张量的协变导数. 比如对逆变矢量而言有 $DV^\mu/d\tau =: \nabla_\nu V^\mu dx^\nu/d\tau = U^\nu \nabla_\nu V^\mu$, 即

$$V^\mu{}_{;\nu} := \nabla_\nu V^\mu = \partial_\nu V^\mu + \Gamma^\mu_{\sigma\nu} V^\sigma = V^\mu{}_{,\nu} + \Gamma^\mu_{\sigma\nu} V^\sigma, \quad (1.2.8)$$

这里类似地也有添加分号和下标的简记法. 但换坐标系时不会给协变导数加上标记, 所加标记只代表所作用张量的分量改变, 如 $V^\mu{}_{;\nu} = \nabla_\nu V^\mu$. 总之, 这相当于在通常不具广义协变性的偏导上加上克氏符修正项, 得到具有广义协变性的导数算符. 协变导数的优势是无需借助于曲线及其参数, 但劣势是某点的 $V^\mu{}_{;\nu}$ 要求 V^μ 在该点邻域都有定义, 而沿线求导只要求曲线上的邻域有定义. 其它张量的协变导数仍可从正交归一系按广义协变性的方式推出, 但更简单地可从如下事实出发: 首先, 就标量 f 而言, 由全微分或链式法则可知, 偏导已有广义协变性, 故

$$f_{;\nu} = f_{,\nu}; \quad (1.2.9)$$

其次, 从正交归一系的偏导可导出协变导数具有 **Leibniz 法则**, 便可从标量、逆变矢量的协变导数及 Leibniz 法则推出一般情形. 首先以协变矢量为例, 构造标量 $V^\mu \omega_\mu$, 则对任意 V^μ 有 $(V^\mu \omega_\mu)_{;\nu} = (V^\mu \omega_\mu)_{,\nu}$, 用 Leibniz 法则拆项并代入矢量协变导数定义有 $V^\mu \omega_{\mu;\nu} + \omega_\mu \Gamma^\mu_{\nu\sigma} V^\sigma = V^\mu (\omega_{\mu;\nu} + \Gamma^\sigma_{\mu\nu} \omega_\sigma) = V^\mu \omega_{\mu;\nu}$, 因此

$$\omega_{\mu;\nu} = \omega_{\mu,\nu} - \Gamma^\sigma_{\mu\nu} \omega_\sigma. \quad (1.2.10)$$

按此方法可得 (k, l) -张量的协变导数为

$$\begin{aligned} T^{\mu_1 \cdots \mu_k}{}_{\nu_1 \cdots \nu_l;\lambda} &= T^{\mu_1 \cdots \mu_k}{}_{\nu_1 \cdots \nu_l,\lambda} \\ &+ \sum_{i=1}^k \Gamma^\mu_{\lambda\sigma} T^{\mu_1 \cdots \sigma \cdots \mu_k}{}_{\nu_1 \cdots \nu_l} - \sum_{j=1}^l \Gamma^\sigma_{\lambda\nu_j} T^{\mu_1 \cdots \mu_k}{}_{\nu_1 \cdots \sigma \cdots \nu_l}. \end{aligned} \quad (1.2.11)$$

在记忆时注意, 此即先填上一个偏导, 再对所有指标依次添加相应的克氏符修正项 (正负取决于指标上下), 并满足指标平衡.

不过, 直接根据定义取计算克氏符将涉及坐标变换, 但注意正交归一系是和度规有关的, 因此其实有办法从度规计算克氏符. 根据变换式

$$g_{\mu\nu} = \tilde{g}_{\alpha\beta} \frac{\partial \xi^\alpha}{\partial x^\mu} \frac{\partial \xi^\beta}{\partial x^\nu}, \quad (1.2.12)$$

显然需对上式求坐标偏导才能构造出定义式 (1.1.41), 即

$$g_{\mu\nu,\lambda} = \Gamma^\rho_{\lambda\mu} g_{\rho\nu} + \Gamma^\rho_{\lambda\nu} g_{\rho\mu}. \quad (1.2.13)$$

其中注意线上的正交归一系给出 $\tilde{g}_{\alpha\beta,\lambda} = 0$. 为从中提出克氏符, 做一些指标替换便可合并同类项. 注意对任意坐标有 **Fubini 定理**: 无论作用于谁都有 $\partial_{[\mu} \partial_{\nu]} = 0$, 因

此有对称性 $\Gamma_{\mu\nu}^\lambda = \Gamma_{(\mu\nu)}^\lambda$, 进而根据 $g_{\mu\nu} = g_{(\mu\nu)}$ 做轮换 $\mu \rightarrow \nu \rightarrow \sigma \rightarrow \mu$ 再消元得 $g_{\sigma\nu,\mu} + g_{\mu\sigma,\nu} - g_{\mu\nu,\sigma} = 2g_{\sigma\lambda}\Gamma_{\mu\nu}^\lambda$, 两边与 $g^{\sigma\alpha}$ 缩并¹, 再将 α 换回 λ 就有

$$\Gamma_{\mu\nu}^\lambda = \frac{1}{2}g^{\sigma\lambda}(g_{\sigma\nu,\mu} + g_{\mu\sigma,\nu} - g_{\mu\nu,\sigma}). \quad (1.2.14)$$

由此式可看出 $\mathbb{R}^{1,3}$ 等度规变换的仿射性, 因为取任意惯性系 $\{x^\mu\}$ 下有 $g_{\mu\nu,\lambda} = \eta_{\mu\nu,\lambda} = 0$, 进而克氏符全零. 式 (1.2.13) 说明度规的协变导数为零:

$$g_{\mu\nu;\sigma} = g_{\mu\nu,\sigma} - g_{\mu\lambda}\Gamma_{\sigma\nu}^\lambda - g_{\nu\lambda}\Gamma_{\sigma\mu}^\lambda = 0, \quad (1.2.15)$$

称为**度规适配性** (metric-compatibility). 这给出许多有趣结论. 沿某世界线各点均有

$$A^\mu U_\mu = g_{\mu\nu} U^\nu \frac{DU^\mu}{d\tau} = \frac{1}{2} \frac{D}{d\tau} (g_{\mu\nu} U^\nu U^\mu) = 0,$$

其中注意度规适配性对任意曲线给出 $Dg_{\mu\nu}/d\tau = 0$ 且 U^ν 模方恒定, 故 A^μ, U^ν 正交, 的确迎合我们早在经典力学中就熟悉的直觉. 实际上在 2 章会细谈, 世界线在 A^μ 的作用下可视作一系列的微小伪转动之和. 数学上称满足 $DV^\mu/d\tau = 0$ 的逆变矢量 V^μ 为**平行移动** (parallel transport) 或 **Levi-Civita 移动**. 切矢沿其自身平行移动的曲线称为**测地线** (geodesic), 即满足

$$A^\lambda = \frac{\partial x^\lambda}{\partial \xi^\alpha} \tilde{A}^\alpha = 0 \iff \frac{d^2 x^\lambda}{d\tau^2} + \Gamma_{\mu\nu}^\lambda \frac{dx^\mu}{d\tau} \frac{dx^\nu}{d\tau} = 0, \quad (1.2.16)$$

上式称为**测地线方程**. 假设 V^μ, W^ν 沿某条线平行移动 (满足 $DV^\mu/d\tau = DW^\nu/d\tau = 0$), 其内积的沿线求导满足

$$\frac{D}{d\tau} (V^\mu W_\mu) = V^\mu W^\nu \frac{Dg_{\mu\nu}}{d\tau} + g_{\mu\nu} W^\nu \frac{DV^\mu}{d\tau} + g_{\mu\nu} V^\mu \frac{DW^\nu}{d\tau} = 0,$$

说明内积恒定. 就单个 V^μ 而言说明模方恒定, 的确可看作一种“平行移动”. 需要注意, 设定时空上的逆变矢量场是给各时空点赋予其相应的逆变矢量, 一个逆变矢量是“绑定”于一点的, 故任意两点 p, q 处各自的逆变矢量本无关联. 然给定某 p, q 间之路径, 可通过矢量的平行移动, 将两个切空间联系起来. 因此 ∇_ν 又被称为**联络** (connection). 在第 2 章会谈数学上一般的联络定义. 每种联络都相应地有一组**联络系数** $\hat{\Gamma}_{\mu\nu}^\lambda$, 相应测地线用 $U^\nu \hat{\nabla}_\nu U^\mu = 0$ 定义. 设**挠率** (torsion) 为 $T^\lambda_{\mu\nu} := 2\Gamma_{[\mu\nu]}^\lambda$, 它靠着克氏符作差消除了这一项, 故是张量. 按式 (1.1.41) 定义的联络系数满足 $T^\lambda_{\mu\nu} = 0$, 故这种联络准确地称为**无挠** (torsion-free) **联络**. 度规适配的无挠联络称为 **Levi-Civita 联络**. 由此二条件可证联络系数正是克氏符, 故 Levi-Civita 联络存在且唯一.

¹ 不同于代数等式有“乘过去”“除过来”等说法, 在指标运算中, 必须理解为是用缩并得到 δ_ν^μ 以消除相应项.

1.2.3 张量密度

关于张量的一个典型例子是在某系的**度规行列式** $g := \det(g_{\mu\nu})$. 对 $g_{\mu\nu}$ 的坐标变换律取行列式有

$$g' = \det\left(\frac{\partial x^\alpha}{\partial x'^\mu}\right) \det\left(\frac{\partial x^\beta}{\partial x'^\nu}\right) g = J^{-2} g, \quad (1.2.17)$$

其中 $J = \det(\partial x'^\mu / \partial x^\alpha)$ 是 Jacobi 行列式. 可见它不是标量场, 变换式多出了 J^{-2} 因子. 这种在张量变换式基础上还多出 Jacobi 行列式因子的量称为**张量密度**, 即

$$T'^{\mu_1 \cdots \mu_k}_{\nu_1 \cdots \nu_l} = J^w \frac{\partial x'^{\mu_1}}{\partial x^{\rho_1}} \cdots \frac{\partial x'^{\mu_k}}{\partial x^{\rho_k}} \frac{\partial x^{\sigma_1}}{\partial x'^{\nu_1}} \cdots \frac{\partial x^{\sigma_l}}{\partial x'^{\nu_l}} T^{\rho_1 \cdots \rho_k}_{\sigma_1 \cdots \sigma_l}, \quad (1.2.18)$$

其中 w 称为张量密度的**权** (weight). 例如度规行列式的权为 -2 . 下面再举一例很重要的张量密度. 设指标可取 $1, \cdots, n$, 某系的 **Levi-Civita 符号**是指

$$\varepsilon_{i_1 \cdots i_n} := \begin{vmatrix} \delta_{i_1}^1 & \cdots & \delta_{i_n}^1 \\ \vdots & \ddots & \vdots \\ \delta_{i_1}^n & \cdots & \delta_{i_n}^n \end{vmatrix} = \prod_{1 \leq l < m \leq n} \operatorname{sgn}(i_m - i_l), \quad (1.2.19)$$

换言之若 $(i_1 \cdots i_n) \in S_n$, 则 $\varepsilon_{i_1 \cdots i_n} = \operatorname{sgn}(i_1 \cdots i_n)$, 否则为零, 因此 $\varepsilon_{i_1 \cdots i_n} = \varepsilon_{[i_1 \cdots i_n]}$. 这样二维方阵 (A^{ij}) 的行列式总可表成

$$\det(A^{ij}) = \sum_{(i_1 \cdots i_n) \in S_n} \operatorname{sgn}(i_1 \cdots i_n) A^{i_1 1} \cdots A^{i_n n} = \varepsilon_{i_1 \cdots i_n} A^{i_1 1} \cdots A^{i_n n}. \quad (1.2.20)$$

$\varepsilon_{i_1 \cdots i_n}$ 下标数一定与指标所允许的维度一致, 因此独立分量只有 $\varepsilon_{1 \cdots n}$ 一种. 不妨设其变换是

$$\varepsilon'_{i_1 \cdots i_n} = \varepsilon_{j_1 \cdots j_n} \frac{\partial x^{j_1}}{\partial x'^{i_1}} \cdots \frac{\partial x^{j_n}}{\partial x'^{i_n}} J^w,$$

只需取 $(i_1 \cdots i_n) = (1 \cdots n)$ 即可, 这样 $1 = \varepsilon'_{1 \cdots n} = J^{-1} J^w$ 即 $w = 1$. 按度规升指标后, 有**上标 Levi-Civita 符号**

$$\varepsilon^{j_1 \cdots j_n} := g^{i_1 j_1} \cdots g^{i_n j_n} \varepsilon_{i_1 \cdots i_n}, \quad (1.2.21)$$

容易验证 $\varepsilon^{j_1 \cdots j_n} = \varepsilon^{[j_1 \cdots j_n]}$, 故只需取 $\varepsilon^{1 \cdots n} = \det(g^{\mu\nu}) = 1/g$. 类似地, 假设其变换式后可取 $1/g' = \varepsilon'^{1 \cdots n} = J^w J(1/g)$, 则 $J^2 = J^w J$, 故权也为 1. 注意 g 尽管与坐标选择有关, 但在变换下一定保号, 因为 $g = \det(\tilde{g}_{ij}) \tilde{J}^{-2}$, 其中 $\tilde{J} := \det(\partial x^\mu / \partial \xi^\alpha)$, 则

$$\operatorname{sgn} g = \det(\tilde{g}_{ij}) = (-1)^s, \quad (1.2.22)$$

这里 sgn 只表示取符. 仿照协变导数的定义构造 **Levi-Civita 张量** $\epsilon_{i_1 \cdots i_n}$, 令它只在正交归一右手系 $\{\xi^i\}$ 下表为 $\varepsilon_{i_1 \cdots i_n}$:

$$\epsilon_{i_1 \cdots i_n} = \varepsilon_{j_1 \cdots j_n} \frac{\partial \xi^{j_1}}{\partial x^{i_1}} \cdots \frac{\partial \xi^{j_n}}{\partial x^{i_n}}. \quad (1.2.23)$$

容易验证 $\epsilon_{i_1 \dots i_n} = \epsilon_{[i_1 \dots i_n]}$, 则可取 $\epsilon_{1 \dots n} = 1/\tilde{J} = \text{sgn } \tilde{J} \sqrt{|g|}$. 这样 Levi-Civita 张量在任意右手系下表为

$$\epsilon_{i_1 \dots i_n} = \sqrt{|g|} \varepsilon_{i_1 \dots i_n}. \quad (1.2.24)$$

进而上标 Levi-Civita 张量就是 $\epsilon^{j_1 \dots j_n} := g^{i_1 j_1} \dots g^{i_n j_n} \epsilon_{i_1 \dots i_n}$, 在右手系下表为

$$\epsilon^{j_1 \dots j_n} = \sqrt{|g|} \varepsilon^{j_1 \dots j_n} = (-1)^s \frac{1}{\sqrt{|g|}} \frac{\varepsilon^{j_1 \dots j_n}}{g^{-1}}. \quad (1.2.25)$$

此外, 还可这样得到 Levi-Civita 符号. 广义 Kronecker 符号是指如下张量

$$\delta_{i_1 \dots i_n}^{j_1 \dots j_n} := n! \delta_{i_1}^{[j_1} \dots \delta_{i_n}^{j_n]} = \begin{vmatrix} \delta_{i_1}^{j_1} & \dots & \delta_{i_n}^{j_1} \\ \vdots & \ddots & \vdots \\ \delta_{i_1}^{j_n} & \dots & \delta_{i_n}^{j_n} \end{vmatrix}. \quad (1.2.26)$$

最后一步是根据行列式定义, 进而 $\delta_{i_1}^{[j_1} \dots \delta_{i_n}^{j_n]} = \delta_{[i_1}^{j_1} \dots \delta_{i_n]^{j_n]} = \delta_{[i_1}^{j_1} \dots \delta_{i_n]^{j_n}}$ 或 $\delta_{i_1 \dots i_n}^{j_1 \dots j_n} = \delta_{[i_1 \dots i_n]}^{[j_1 \dots j_n]}$. 假设 $\{i_k\}, \{j_k\}$ 内均无相同指标 (否则为零), 可以看出二者互为偶排列时为 1, 互为奇排列时为 -1. 这样, 取 $j_k = k$ 就可有 $\epsilon_{i_1 \dots i_n} = \delta_{i_1 \dots i_n}^{1 \dots n}$, 但要注意 $\delta_{i_1 \dots i_n}^{1 \dots n}$ 已不再是张量. 可以证明

$$\epsilon_{i_1 \dots i_n} \epsilon^{j_1 \dots j_n} = (-1)^s \delta_{i_1 \dots i_n}^{j_1 \dots j_n}, \quad (1.2.27)$$

只要注意等式两边都只有一种独立分量, 因此只要取特例确定出系数即可. 经常还要涉及到其指标缩并:

$$\delta_{j_k i_{k+1} \dots i_n}^{j_k j_{k+1} \dots j_n} = k \delta_{i_{k+1} \dots i_n}^{j_{k+1} \dots j_n} \implies \delta_{j_1 \dots j_k i_{k+1} \dots i_n}^{j_1 \dots j_k j_{k+1} \dots j_n} = k! \delta_{i_{k+1} \dots i_n}^{j_{k+1} \dots j_n}. \quad (1.2.28)$$

读者可由归纳法证明前者, 后者是立即可得的结论. 进而

$$\epsilon_{j_1 \dots j_k i_{k+1} \dots i_n} \epsilon^{j_1 \dots j_k j_{k+1} \dots j_n} = (-1)^s k! \delta_{i_{k+1} \dots i_n}^{j_{k+1} \dots j_n}. \quad (1.2.29)$$

现在取常用情形, 使读者对以上的抽象性质有直观理解. 比如取 \mathbb{E}^3 , 设逆变矢量 A^i, B^j (任意坐标系), 构造协变矢量 $\epsilon_{ijk} A^j B^k$, 为了看出这确实是叉乘, 只需取正交归一右手系有 $\varepsilon_{ijk} \tilde{A}^j \tilde{B}^k$, 则比如

$$\varepsilon_{1jk} \tilde{A}^j \tilde{B}^k = \tilde{A}^2 \tilde{B}^3 - \tilde{A}^3 \tilde{B}^2 = (\mathbf{A} \times \mathbf{B})_1, \quad (1.2.30)$$

之后用欧氏度规升指标即可. 其次有

$$\epsilon_{kij} \epsilon^{kmn} = \delta_i^m \delta_j^n - \delta_i^n \delta_j^m, \quad (1.2.31)$$

请读者找一些电动力学教材, 试着用分量语言证明常见的矢量代数或微分恒等式, 如第 1.1.3 节开头的那些. 该式多用于连续叉乘运算. 注意取正交归一右手系和偏导即可, 这样微分算子的分量语言是 ∂_i .

1.2.4 多元微分学

协变导数对标量场的作用和梯度无异. 协变导数与所作用的张量的某一指标缩并, 称为**时空散度**或**散度**. 根据协变导数的克氏符项, 这涉及**缩并克氏符**

$$\Gamma_{\sigma\mu}^{\sigma} = \frac{1}{2}g^{\sigma\kappa}(g_{\sigma\kappa,\mu} + 2g_{\mu[\kappa,\sigma]}) = \frac{1}{2}g^{\sigma\kappa}g_{\sigma\kappa,\mu} = \partial_{\mu}g/2g = \partial_{\mu}\ln\sqrt{|g|}, \quad (1.2.32)$$

其中使用了线性代数中的 Jacobi 公式, 此处具体是

$$\partial_{\lambda}g = gg^{\mu\nu}g_{\mu\nu,\lambda}. \quad (1.2.33)$$

进而可计算任意张量的散度. 这里只举常用例子: 矢量的散度为

$$V^{\mu}_{;\mu} = V^{\mu}_{,\mu} + \Gamma_{\mu\nu}^{\mu}V^{\nu} = V^{\mu}_{,\mu} + V^{\mu}\partial_{\mu}\ln\sqrt{|g|} = \frac{1}{\sqrt{|g|}}\partial_{\mu}(\sqrt{|g|}V^{\mu}); \quad (1.2.34)$$

2 阶张量的散度为

$$T^{\mu\nu}_{;\nu} = T^{\mu\nu}_{,\nu} + \Gamma_{\lambda\nu}^{\mu}T^{\lambda\nu} + \Gamma_{\lambda\nu}^{\nu}T^{\mu\lambda} = \frac{1}{\sqrt{|g|}}\partial_{\nu}(\sqrt{|g|}T^{\mu\nu}) + \Gamma_{\lambda\nu}^{\mu}T^{\lambda\nu}, \quad (1.2.35)$$

特别当 $T^{\mu\nu}$ 反称时, 结合无挠性有 $\Gamma_{\lambda\nu}^{\mu}T^{\lambda\nu} = 0$. 我们知道三维旋度 $\epsilon^{ijk}\nabla_j B_k$ (正交系下是 $\epsilon^{ijk}\partial_j B_k$), 在多维中这个概念稍稍微妙一些, 这是因为其中的 Levi-Civita 张量要取成 $\epsilon^{i_1\cdots i_n}$, 而 $\epsilon^{i_1\cdots i_n}\nabla_{i_1}B_{i_2}$ 并不能给出一阶张量. 故可尝试丢掉 Levi-Civita 张量来直接定义旋度, 但切记要靠缩并消项, 即

$$\epsilon_{imn}\epsilon^{ijk}\nabla_j B_k = (\delta_m^j\delta_n^k - \delta_m^k\delta_n^j)\nabla_j B_k = 2\nabla_{[m}B_{n]}. \quad (1.2.36)$$

其中用到 (1.2.31) 式. 无挠性使它在任意坐标系等价于 $2\partial_{[m}B_{n]}$. 实际上 $2\nabla_{[m}B_{n]}$ 和 $\epsilon^{ijk}\nabla_j B_k$ 之间通过 Levi-Civita 张量构建的这种关系也是一种对偶, 见 ?? 节. 在四维时, 我们直接称矢量场 V^{μ} 的**时空旋度**就是 $2\nabla_{[\mu}V_{\nu]}$, 而二阶反称张量 $F_{\mu\nu}$ 的旋度则记作 $3\nabla_{[\mu}F_{\nu\lambda]}$, 无挠性使它等于偏导表述. 我们所熟悉的三维 Laplace 算子 ∇^2 只是简单的 $\partial^i\partial_i$, 而 $\mathbb{R}^{1,3}$ 惯性系下的 $\partial^{\mu}\partial_{\mu} = \eta^{\mu\nu}\partial_{\mu}\partial_{\nu}$ 就是电动力学中的 d'Alembert 算子 \square . 用协变导数表述则称 **Laplace-Beltrami 算子**, 对标量场来说很简单:

$$\nabla_{\mu}\nabla^{\mu}f = \frac{1}{\sqrt{|g|}}\partial_{\mu}(\sqrt{|g|}\nabla^{\mu}f) = \frac{1}{\sqrt{|g|}}\partial_{\mu}(\sqrt{|g|}g^{\mu\nu}\partial_{\nu}f). \quad (1.2.37)$$

1.3 拉氏理论

本节若无特别指出, 默认取惯性系. 目前许多经典物理方程可以靠这样一种信念得到: 相互作用的实际规律总能使系统的作用量最小, 这称为**最小作用量原理** (principle of least action)¹. 这一哲思最初由 18 世纪 Lagrange, Hamilton 等人发展: 给

¹当然, 最小只是美学说法, 我们可填上负号使之变成“最大”, 甚至只要求导数为零, 故不一定在该处取最值, 甚至也非极值而是鞍点. 故这又称**稳恒 (stationary) 作用量原理**. 我们仅在分析解的稳定性时再关注最优化问题.

定某质点的初末位置 $\mathbf{x}(t_1), \mathbf{x}(t_2)$, 作用量若表为沿运动轨迹 $\mathbf{x}(t)$ 的如下积分泛函

$$S(\mathbf{x}(t)) = \int_{t_1}^{t_2} L(\mathbf{x}, \dot{\mathbf{x}}) dt, \quad L(\mathbf{x}, \dot{\mathbf{x}}) = T - V(\mathbf{x}), \quad T = \frac{1}{2} m \dot{\mathbf{x}}^2, \quad (1.3.1)$$

其中标量 L 称作**拉氏量** (Lagrangian), T 是动能, V 是保守势能, 就可推得 Newton 第二定律 $m\ddot{\mathbf{x}}(t) = -\nabla V$. 换言之, 这样设定的 $S(\mathbf{x}(t))$ 就给出了关于 $\mathbf{x}(t)$ 的经典力学方程. 可以看到, 改变拉氏量的表达就给出的一套不同理论. 一个系统通常具备相应的作用量, 它是路径的泛函 $S(\psi(\sigma))$, 且总保为拉氏量积分的形式. 固定端点 p, q 间的各种路径构成的集合 Ω_p^q 称为 **Fréchet 空间**, 故作用量严格来说是 Ω_p^q 上的函数. 注意, 未给出实际规律之前, 所考虑的路径可以非常任意, 比如质点的 Fréchet 空间可包括任意 $\mathbf{x}(t_1), \mathbf{x}(t_2)$ 间的曲线. 一般地对含 N 个粒子的系统, 设各粒子的位形 q_i 而 $q_i(t)$ 是用 t 参数化的路径, 规定作用量总能表为

$$S(q_1(t), \dots, q_N(t)) = \int_{t_0}^{t_1} L(q_1, \dots, q_N, \dot{q}_1, \dots, \dot{q}_N) dt, \quad (1.3.2)$$

其中 $\dot{q}_i := dq_i/dt$. 或者简写为

$$S(q_i(t)) = \int_{t_0}^{t_1} L(q_i, \dot{q}_i) dt, \quad (1.3.3)$$

但不再研究 $L(q_i, \dot{q}_i, \ddot{q}_i, \dots, t)$, 有两点考虑: 若承认物理规律不随时间变化, L 就不应显含时间¹, 默认 $\partial L/\partial t = 0$; 若承认 **Newton 决定性原理**: 仅由位置和速度即可确定经典运动, 更高阶的导数就不会出现在方程中. 对于经典力学, 取位形空间为 \mathbb{E}^3 , 坐标系取 Descartes 系, t 视作绝对时间. 在狭义相对论中, 位形空间取 $\mathbb{R}^{1,3}$, 坐标系取 Lorentz 系, 可以用某系坐标时或固有时参数化世界线.

1.3.1 变分法和对称性

那么如何对泛函求导? 假设对某一粒子, 点 $q_i(t_0), q_i(t_1)$ 间使得泛函取极值的曲线是 $q_i(t)$. 注意极值的含义就是取邻近自变量时因变量近似不变. 考虑 $q_i(t)$ 的一条邻近曲线 $\hat{q}_i(t)$, 其参数式与原先的差是微小的 $\delta q_i = \hat{q}_i - q_i$, 但要求端点不动, 故有 $\delta q_i(t_0) = \delta q_i(t_1) = 0$. 这个 δ 记号就称为**变分** (variation), 泛函的这套运算称为**变分法**. 极值曲线的变分所引起的泛函变分满足 $\delta S = 0$, 即相对于路径变化是高阶无穷小. 最小作用量原理就是要求实际路径是 Ω_p^q 中使 $\delta S = 0$ 的元素, 这称为**正路**, 否则称**旁路**. 注意

$$\delta S = \int_{t_0}^{t_1} L(q_i + \delta q_i, \dot{q}_i + \delta \dot{q}_i) dt - \int_{t_0}^{t_1} L(q_i, \dot{q}_i) dt = \int_{t_0}^{t_1} \delta L dt, \quad (1.3.4)$$

¹ L 只通过 x^μ, \dot{x}^μ 的关系隐含时间. 按理来说, 只要给出 $x^\mu(t)$, L 就应能用 t 表示, 似乎与 $\partial L/\partial t = 0$ 矛盾. 其实这涉及诸多符号混淆, 我们在第 8 章再一并解决. 简单来说, 在给出实际运动方程前我们并不知道 $x^\mu(t)$.

因此端点和参数 t 的不变使变分就像挪进了积分内一样. 在极限 $\delta x \rightarrow 0$ 下有

$$\delta L = \frac{\partial L}{\partial q_i} \delta q_i + \frac{\partial L}{\partial \dot{q}_i} \delta \dot{q}_i,$$

此处注意对 $i = 1, \dots, N$ 求和. 可见变分的代数运算与微分极为相似. 其中的 $\delta \dot{q}_i$ 不便于处理, 注意

$$\delta \dot{q}_i = \frac{d}{dt}(q_i + \delta q_i) - \dot{q}_i = \frac{d}{dt}(\delta q_i), \quad (1.3.5)$$

因此可交换顺序. 变分法的关键是利用**分部积分法**, 将一部分积分转化为**边界项**. 这样, 结合变分的边界限制就可消去边界项为

$$\int_{t_0}^{t_1} \frac{\partial L}{\partial \dot{q}_i} \delta \dot{q}_i dt = \left. \frac{\partial L}{\partial \dot{q}_i} \delta q_i \right|_{t_0}^{t_1} - \int_{t_0}^{t_1} \delta q_i \frac{d}{dt} \frac{\partial L}{\partial \dot{q}_i} dt = - \int_{t_0}^{t_1} \frac{d}{dt} \frac{\partial L}{\partial \dot{q}_i} \delta q_i dt,$$

代回去就可提出 δq_i 为

$$\delta S = \int_{t_0}^{t_1} \left(\frac{\partial L}{\partial q_i} - \frac{d}{dt} \frac{\partial L}{\partial \dot{q}_i} \right) \delta q_i dt = 0,$$

该式对任意变分 δq_i 成立, 当且仅当括号恒为零, 即

$$\frac{d}{dt} \frac{\partial L}{\partial \dot{q}_i} - \frac{\partial L}{\partial q_i} = 0, \quad (1.3.6)$$

这称为 **Euler-Lagrange 方程**, 简称 **E-L 方程**. 请读者验证, 借助惯性系定义 **3-速** $u^i := dx^i/dt$ (对偶记 $u_i = \delta_{ij} u^j$), 代入 $L = \frac{1}{2} m u_i u^i - V$ 可计算得 $m \dot{u}^i = -\partial^i V$. 不过需强调, 对简单的 L , 可以直接计算相应偏微分再代入 E-L 方程, 而更多时候会遇上更复杂的 L , 这时我们采取上述流程直接变分, 再用分部积分和变分限制消除边界项. 基本上, 我们会找各种理由消除边界项, 则本书会经常省略积分域, 少数情况才去讨论泛函变分的边界项.

读者熟知动量、能量概念及其守恒, 现代观点下, 许多守恒量可解释为对称性的结果. 所谓对称性, 字面意思即系统作用量长得很对称, 好比圆具有左右对称、旋转对称等, 其实质是圆在反射变换、旋转变换下长相没变. 这些变换即指背景点的变动, 称为**点变换**¹, 一种变换对应一个参数 α , 可表示为新旧点之间的坐标关系 $\tilde{q} = h_\alpha(q)$. 若参数只取有限值, 则称**离散变换**, 比如反射变换只包含两种操作; 参数可连续取值, 则称**连续变换**, 如平移、旋转按平移量、旋转角来描述可有无穷多种. 可将连续变换视作无限个分段的累积, 故只需研究所谓的**无穷小变换**, 即先取 $\alpha = \epsilon$, 最后往往可同除以 ϵ 再令 $\epsilon \rightarrow 0$, 这一极限操作使我们能放心地在“取极限前”就将变换保留到一阶, 也即物理学常在理论计算中就开始做近似的理由. 下面为了方便只考虑一个粒子. 无论某条路径是否是正路, 变换导致其本身的改变, 一种是轨迹的改变

$$\delta_s q(t) := \tilde{q}(t) - q(t), \quad (1.3.7)$$

¹与坐标变换有联系, 区别也较微妙. 一般对不变性总讨论坐标变换, 对称性总讨论点变换. 细节留到第 2 章.

一种是时间自己的重参数化

$$\delta_s t := \tilde{t} - t. \quad (1.3.8)$$

下面均保留至一阶. 将总的变化记作

$$\Delta q(t) := \tilde{q}(\tilde{t}) - q(t) = \delta_s q(t) + \left(\dot{q}(t) + \frac{d}{dt}(\delta_s q(t)) \right) \delta_s t = \delta_s q(t) + \dot{q}(t) \delta_s t, \quad (1.3.9)$$

而 $\dot{q}(t)$ 的总变化为

$$\begin{aligned} \Delta \dot{q}(t) &:= \frac{d\tilde{q}(\tilde{t})}{d\tilde{t}} - \dot{q}(t) = \frac{dt}{d\tilde{t}} \left(\dot{q}(t) + \frac{d}{dt}(\delta_s q(t)) + \ddot{q}(t) \delta_s t + \dot{q}(t) \frac{d}{dt}(\delta_s t) \right) - \dot{q}(t) \\ &= \end{aligned} \quad (1.3.10)$$

一般来说, 在某个点变换 $q(t) \rightarrow \tilde{q}(\tilde{t})$ 下, 整个系统的新位形将改变作用量的值, 产生的变化是.

$$\Delta S := \int L\left(\tilde{t}, \tilde{q}(\tilde{t}), \frac{d\tilde{q}(\tilde{t})}{d\tilde{t}}\right) d\tilde{t} - \int L(t, q(t), \dot{q}(t)) dt, \quad (1.3.11)$$

满足 $\Delta S = 0$ 的变换称为**对称变换**¹, 或称作用量具有这种对称性, 即作用量在该变换下的形式不变. 一种能轻松办到这一要求的方法是让 L 不显含这个坐标. 比如, 默认 $\partial L / \partial t = 0$ 就使得时间的任何重参数化 $\tilde{t} = t + C$ 毫无意义, 因此才总称它是**时间平移对称性**. 如果 q 是线坐标, 则 $\partial L / \partial q = 0$ 意味着**空间平移对称性**; 同理, 取角坐标则给出**空间各向同性**. 将 ΔS 展开到一阶有

$$\Delta S =$$

比如, 空间上的无穷小平移可以写成 $\delta_s q = \epsilon$.

其中关于 ϵ 的一阶导 ξ^μ 取决于具体变换.

这种连续整体对称性和守恒律之间的关系被称为 **Noether 定理**.

E-L 方程立即给出动量守恒 $dp_i/dt = 0$

动量可由 $p_i = \partial L / \partial u^i = m u_i$ 给出. 由于它具有下标, 故其实是动量对偶, 我们称为**正则动量**.

能量则是, 因为仅考虑保守作用, 则机械能等同能量. 非保守力的实质是统计效应, 不在纯粹的经典力学范围中. 可推广为 $H = 2T - L = p_i u^i - L$, 这样做的优势是可从一般理论的拉氏量直接定义能量. 这称为 **Legendre 变换**, 此语境下 H 称为**哈氏量** (Hamiltonian), 第 8 章对此给出了直观理解.

$$H = T + V$$

¹当然, 也可让二者只相差边界项 (称为**准不变性**), 详细内容留到第 8 章.

并且, 从该表述能显然推得 H 的守恒, 即

$$\dot{H} = \frac{\partial L}{\partial u^i} \dot{u}^i + u^i \frac{d}{dt} \frac{\partial L}{\partial u^i} - \frac{\partial L}{\partial x^i} u^i - \frac{\partial L}{\partial u^i} \dot{u}^i = u^i \left(\frac{d}{dt} \frac{\partial L}{\partial u^i} - \frac{\partial L}{\partial x^i} \right) = 0, \quad (1.3.12)$$

当然, 由于存在不同的 L 给出相同运动方程 (比如乘以常数或加上全导数项), 这样哈氏量不一定等于能量. 若只从守恒角度考量, 非称任意哈氏量都能作为能量的定义也可以.

1.3.2 粒子的拉氏理论

但在形式上显然对质点的各种位形空间成立.

给出一个理论的关键在于写出拉氏量, 一般来说我们能按对称性、量纲等要求构造出来. 先考虑自由粒子. 经典力学是取 $L = T$, 那么相对论呢? 时空上代表物理规律的拉氏量是具有 Lorentz 不变性的标量. 基本上, 我们只能找 $d\tau, m, c = 1$ 这种量来构造, 取 $S \propto m \int d\tau$. 注意在 $u \ll 1$ 时

$$m \int \sqrt{-\eta_{\mu\nu} \dot{x}^\mu \dot{x}^\nu} dt = m \int \sqrt{1 - u^2} dt \approx m \int \left(1 - \frac{u^2}{2} \right) dt,$$

其中常数不影响变分, 可见若想还原至经典力学的自由粒子动能, 只能取

$$L = -m\sqrt{-\eta_{\mu\nu} \dot{x}^\mu \dot{x}^\nu} = -m/\gamma, \quad S = -m \int \sqrt{-\eta_{\mu\nu} \dot{x}^\mu \dot{x}^\nu} dt. \quad (1.3.13)$$

若想取参数为固有时 τ , 则默认是正路的固有时. 对于任何邻近旁路, τ 都非其固有时而只是纯粹的参数. 为何不能用每条路径自己的固有时来作参数? 因为变分的实质是像 (1.3.4) 式那样使参数本身与变分无关. 比如, 对有限积分域, Ω_p^q 中的任意邻近旁路上所用参数的取值范围必须为同一区间 $I \subset \mathbb{R}$, 若每条路径都用自己的固有时作参数就办不到. 既然该参数无论如何都只有数学意义, 我们取坐标时 t 就可避免这些讨论. 变分得到实际方程后, 再将 4-速模方 $\eta_{\mu\nu} \dot{x}^\mu \dot{x}^\nu = -1$ 代入¹, 就相当于将 $x^\mu(t)$ 重参化为 $x^\mu(\tau)$. 由 $L = -m\sqrt{-\eta_{\mu\nu} \dot{x}^\mu \dot{x}^\nu}$ 可计算得 $\partial L / \partial \dot{x}^\kappa = m\eta_{\kappa\lambda} \dot{x}^\lambda$ 和 $\partial L / \partial x^\kappa = 0$, 则 E-L 方程在约去 m 并与 $\eta^{\mu\kappa}$ 缩并后是 $\ddot{x}^\mu = 0$, 也可重参化为

$$\frac{d^2 x^\mu}{d\tau^2} = 0. \quad (1.3.14)$$

换言之, 自由粒子在 $\mathbb{R}^{1,3}$ 中走直线. 由于度规是闵氏的, 因此直线是使固有时极大而非极小, 即 “长程线”. 著名的**孪生佯谬** (twin paradox) 就是说, 设一对双胞胎从某时空点分离, 并回合于另一时空点, 则走惯性运动 (如静止) 的那个人将年长于

¹许多书籍对如下现象都未解释或作出错误解释: 若用固有时 τ 作参数, 则 $L = -m$, 得不到任何方程, 故强行解释为需还原代数形式才能进行变分. 正确的解释是 $U^\mu U_\mu = -1$ 只对正路成立, 其中参数 τ 与固有时一致. 变分必须考虑邻近旁路, 有它们自己的不同于 τ 的固有时, 因而沿着这些旁路 $U^\mu U_\mu \neq -1$.

走非惯性运动的另一人. 由于两人实际是非质点, 因此走惯性运动较为舒适, 而走非惯性运动将感受到巨大的加速度 (使相对论现象足够明显的物理量达光速级). 由 $u := \sqrt{u_i u^i}$ 知 $\partial u / \partial u^i = u_i / u$, 故

$$p_i = \partial L / \partial u^i = \gamma m u_i \implies p^i = \delta^{ij} p_j = \gamma m u^i = m u^i + \frac{1}{2} m u^2 u^i + \cdots, \quad (1.3.15)$$

后者称为 **3-动量**, 并将 γ 按二项式定理展开. 可见若 $u \ll 1$, 则 3-动量近似于经典动量 $m u^i$, 即 3-动量给出了其相对论修正. 也可用 4-速 $U^\mu = dx^\mu / d\tau = \gamma \dot{x}^\mu$ 得出相应动量, 注意 $L = m U^\mu U_\mu / \gamma = m U_\mu \dot{x}^\mu$, 则

$$P_\mu = \partial L / \partial \dot{x}^\mu = m U_\mu \implies P^\mu = m U^\mu. \quad (1.3.16)$$

后者称为 **4-动量**. 易知 $p^2 := p_i p^i = m^2 u^2 / (1 - u^2)$, 则 $u = p / \sqrt{p^2 + m^2}$, $\gamma = \sqrt{p^2 + m^2} / m$. 故在相对论中能量¹

$$E = p_i u^i - L = p u + m / \gamma = \sqrt{p^2 + m^2} = \gamma m = m + \frac{1}{2} m u^2 + \cdots, \quad (1.3.17)$$

经典动能 $(1/2) m u^2$ 的出现表明 $(\gamma - 1)m$ 是其相对论修正. 可见粒子即使静止也有能量, 也就应该能释放出来并得到应用. 对核能的开发已验证了这一点. 从上式立即看出**质壳 (mass-shell) 关系** $E^2 = p^2 + m^2$. 注意 P^μ 恰好同时具备能量和动量意义, 即 $P^0 = \gamma m = E$, $P^i = \gamma m u^i = p^i$, 则一方面 $P^\mu P_\mu = -m^2$, 另一方面 $P^\mu P_\mu = -E^2 + p^2$, 故质壳关系实质代表 4-动量的模方恒等式.

现在讨论试验粒子在场中的运动. 这一主题称为外场与粒子的**耦合 (coupling)**. 目前除了接受实验筛选, 没有先验原理支撑唯一的耦合方式. 有时可在自由粒子的 S 上做修改并保证无场时还原, 如 **Dilaton 标量场** ϕ 的耦合为 $S = -m \int e^\phi d\tau$, 请读者证明在低速和弱场 ($|\phi| = |V|/m \ll 1$) 下还原出 $L = T - V$. 更多时候可考虑简单相加, 矢量场耦合就恰行此路, 比如经典电磁场. 电磁场强有两个分明的部分 E^i, B^i , 这六个信息很难单纯用一个四维矢量去涵盖, 除非考虑其势: 取惯性系, 给定电标势 ϕ 和磁矢势 \mathbf{a} 就确定了电磁场 $\mathbf{B} = \nabla \times \mathbf{a}$, $\mathbf{E} = -\nabla \phi - \partial_t \mathbf{a}$. 这样可按 $\phi =: A^0, \mathbf{a} =: A^i$ 凑成 **4-势** A^μ . 为耦合出标量的 S , 最简单的方法是用与粒子 4-速的内积 $A_\mu U^\mu$, 即

$$S = \int (-m + q A_\mu U^\mu) d\tau, \quad (1.3.18)$$

其中常数 q 是给场与粒子的耦合项引入的系数, 称为**耦合强度**. 这无非是给理论中的参数更多自由的余地. 下面将看到它确实是粒子电荷, 这样我们对此旧概念又有了新认知. 现在先发挥物理学家独到的意识流计算: $P_\mu \propto \partial(A_\nu U^\nu) / \partial U^\mu = A_\mu$, 进

¹ $E = \gamma m$ 就是读者熟知的**质能方程** “ $E = mc^2$ ”, 但该语境下的 “ m ” 是指动质量, 即 $m_k = \gamma m$, 而称本书的质量 m 为静质量, 能量 E 为总能. 静能即本书 $E_0 = mc^2$. 笔者遵循现代学术习惯, 只保留 m, E 两个概念. 故哪怕取 $c = 1$, 严格来说也不能有 “质量就是能量” 这种说法. 但无论如何 “静质量是静能” “动质量是总能” 是可以的.

而 $dP_\mu/d\tau \propto \partial_\nu A_\mu U^\nu$, 故预料此即运动方程的形式. 而实际上, 第一项的变分我们已知, 关键是对第二项变分并默认消除边界项 (常用 “ \simeq ” 标注):

$$\begin{aligned}\delta \int A_\mu U^\mu d\tau &= \int \left(\partial_\nu A_\mu \delta x^\nu U^\mu + A_\mu \frac{d\delta x^\mu}{d\tau} \right) d\tau \\ &\simeq \int (\partial_\mu A_\nu - \partial_\nu A_\mu) U^\nu \delta x^\mu d\tau = \int F_{\mu\nu} U^\nu \delta x^\mu d\tau,\end{aligned}$$

其中 $F_{\mu\nu} := 2\partial_{[\mu} A_{\nu]}$ 是 4 势的时空旋度, 称**电磁张量**. 结合自由粒子项的变分就有电磁场中的粒子运动方程

$$m \frac{dP^\mu}{d\tau} = m \frac{d^2 x^\mu}{d\tau^2} = q F^\mu{}_\nu U^\nu. \quad (1.3.19)$$

由定义 (注意 $\mathbb{R}^{1,3}$ 下 $A_0 = -\phi, A_i = \delta_{ij} a^j$) 可知

$$E_i = F_{i0}, \quad B_i = \varepsilon_{ijk} \partial^j A^k = \varepsilon_{ijk} \partial^{[j} A^{k]} = \frac{1}{2} \varepsilon_{ijk} F^{jk} \Leftrightarrow F^{ij} = \varepsilon^{ijk} B_k. \quad (1.3.20)$$

可使用矩阵表示为

$$(F_{\mu\nu}) = \begin{bmatrix} 0 & -E^1 & -E^2 & -E^3 \\ E^1 & 0 & B^3 & -B^2 \\ E^2 & -B^3 & 0 & B^1 \\ E^3 & B^2 & -B^1 & 0 \end{bmatrix}, \quad (1.3.21)$$

它确实有场强之意义, 故取 $f^i = m\dot{u}^i$ 部分可看出该运动方程正是 Lorentz 力定律. 称 $qF^\mu{}_\nu U^\nu$ 为 **4-Lorentz 力**. 类似地, 欲直接得 3 维表述可用惯性坐标时 t 作为参数, 位形空间取同时面 Σ_t , 这样可表为 $L = -m\sqrt{1-u^2} - q\phi + qu_i a^i$, 对 x^i 变分可同样得到 Lorentz 力定律. 可见, $A_\mu U^\mu$ 的简单性¹迫使场强成为 4-速的线性变换. 从定义立即得到一个微分恒等式

$$F_{[\mu\nu,\lambda]} = 2A_{[[\nu,\mu],\lambda]} = 2A_{[\nu,(\mu,\lambda)]} = 0, \quad (1.3.22)$$

即 $F_{\mu\nu}$ 的定义给出其无旋性. 这已给出 Maxwell 方程组中的

$$\nabla \cdot \mathbf{B} = 0, \quad \nabla \times \mathbf{E} + \partial_t \mathbf{B} = \mathbf{0}.$$

也可从 $U^\mu{}_{,\nu} U^\nu, U_\mu$ 正交出发, 由 $F_{\mu\nu} U^\mu U^\nu = 0$ 通过分配律理解其反称性.

1.3.3 场的拉氏理论

我们总是选择式 (1.3.13) 作为时空中自由粒子的拉氏量. 对于存在相互作用的粒子系, 在粒子间距很大从而无相互作用的极限下, 仍把拉氏量取为式 (1.3.13) 的形

¹ 追求尽可能简单不意味着自然规律 (若有且能被人总结) 是简单的, 换言之, 从简美学纯粹是出于人为便利, 能迎合目前技术所能允许的实验即可.

式. 于是可做如下推广: 任何粒子和场的相互作用系统, 其总哈氏量等于粒子和场的总能量. 但在证明这一论断前, 我们必须得出关于场的拉氏理论. 换言之, 要从粒子的作用量推广到场的作用量, 进而可通过 Dirac 函数又囊括粒子情形, 况且确定的轨道概念几乎只能在经典范围内建立. 区别在于, 场的“位形”可取相应空间场, 但场的“路径”应是空间场往时空上的延伸, 即它正是时空上的张量 (分量) 场 $\psi(x^\mu)$, 而“位形”视作它在某同时面 Σ_t 上的空间截面 $\psi_t(\mathbf{x}) := \psi(t, \mathbf{x})$. 不妨考虑从离散的粒子系统逼近. 假设 \mathbf{x} 的自由度是有限的, 则空间截面 $\psi(\mathbf{x}_i)$ 形同位形 q_i , 因此可将作用量写成

$$S(q_i(t)) = \int L(q_i(t), \dot{q}_i(t)) dt,$$

其中 L 包含对 i 的求和, 则连续极限即对 \mathbf{x} 的积分:

$$L = \int \mathcal{L}(\psi, \psi_{,\mu}) d^3x \implies S(\psi(x^\mu)) = \int \mathcal{L} d^4x, \quad (1.3.23)$$

这个 \mathcal{L} 称为**拉氏密度**, 毕竟泛函宗量是 $\psi(x^\mu)$, 故更彻底的做法是用时空坐标作为积分变量. 可见 \mathcal{L} 在场论中的地位和 L 在轨道理论中的地位几乎一致, 因此不加区分地称为拉氏量也是可以的. ψ 及其时间导数 $\partial_t \psi$ 来源于对 q_i, \dot{q}_i 的类比, 而空间导数可以理解为场的相邻单元之间的势能. 同理, 默认许多理论能存在 \mathcal{L} 表述; \mathcal{L} 一般不显含时空点的坐标 x^μ , 且至少含宗量的一阶导. 在粒子情形我们通过决定性原理排除高阶导, 但现在是关于时空坐标的导数, 且空间项有相互作用势的含义, 冒然排除高阶是激进且危险的做法. 一般地应该写作

$$S(\psi(x^\mu)) = \int \mathcal{L}(\psi, \partial_{\mu_1} \psi, \dots, \partial_{\mu_k} \dots \partial_{\mu_1} \psi) d^4x. \quad (1.3.24)$$

下一节将看到引力理论就必须至少有二阶导. 故实操中我们总是直接计算变分. 在粒子情形, 边界项即函数作差, 而在高维情形中这就是 Gauss 定理 (见 ?? 节). 因此要想消除边界项, 可以有三类方法: 1. 若无任何条件, 可“耍赖”地取无穷积分, 认为场在无穷远没有贡献 (被积项至少比 $1/r^2$ 更快地趋于零) 或不存在边界; 2. 仅在 $\delta\psi = 0$ 条件下, 可通过修改 \mathcal{L} 使其消除边界项; 3. 对积分域 U , 认为 U 外的旁路 ψ 恒定地与边界一致, 故边界上任意阶都有 $(\delta\psi)_{,\mu_1, \dots, \mu_k} = 0$, 这一更强的条件可使 \mathcal{L} 无论含多少阶导数都能消除边界项. 一般来说, 消除边界项总采取第三类理由. 总之在变分时, 碰到全微分项可以舍去. 若 \mathcal{L} 至多含二阶导, 则类似可得场的 E-L 方程

$$\frac{\partial}{\partial x^\mu} \frac{\partial \mathcal{L}}{\partial \psi_{,\mu}} - \frac{\partial \mathcal{L}}{\partial \psi} = 0, \quad (1.3.25)$$

注意这只是在 $\mathcal{L}(\psi, \psi_{,\mu})$ 的情形下之结论. 一个典型例子是 \mathbb{R}^{3+1} 上的自由标量场

$$S = -\frac{1}{2} \int ((\partial^\mu \phi)(\partial_\mu \phi) + m^2 \phi^2) d^4x = 0 \implies \partial^\mu \partial_\mu \phi - m^2 \phi = 0, \quad (1.3.26)$$

其中动能项是用 $\partial_\mu \phi$ 构造的最简单标量, 后者称为 **Klein-Gordon 方程**, 最初是修改量子力学以迎合相对论的一种尝试, 在后来更为精确的量子场论中用以描述自旋为零的粒子 (如 π 介子), 而 m 表示量子化后的粒子静质量.

场论 Noether 定理

设某荷电系统的 3-速场 u^i , 以 $\rho_e =: J^0, j^i = \rho_e u^i =: J^i$ 构造出 **4-电流密度** J^μ 如电荷守恒. 满足 $J^\mu{}_{,\mu} = 0$

荷电系统在任意惯性系满足 $\partial_t \rho_e + \nabla \cdot \mathbf{j} = 0$

场的哈氏量类似, 但 L 包含的求和代之以积分:

$$H = \int \left(\psi_{,t} \frac{\partial \mathcal{L}}{\partial \psi_{,t}} - \mathcal{L} \right) d^3x \quad (1.3.27)$$

为了符号简便, 记张量

$$T_\mu{}^\nu := \psi_{,\mu} \frac{\partial \mathcal{L}}{\partial \psi_{,\nu}} - \delta_\mu^\nu \mathcal{L}, \quad (1.3.28)$$

注意

$$\partial_\nu T_\mu{}^\nu = \psi_{,\mu,\nu} \frac{\partial \mathcal{L}}{\partial \psi_{,\nu}} + \psi_{,\mu} \frac{\partial}{\partial x^\nu} \frac{\partial \mathcal{L}}{\partial \psi_{,\nu}} - \partial_\mu \mathcal{L},$$

假设 \mathcal{L} 至多含一阶导且不显含时空点, 则

$$\partial_\nu T_\mu{}^\nu = \psi_{,\mu,\nu} \frac{\partial \mathcal{L}}{\partial \psi_{,\nu}} + \psi_{,\mu} \left(\frac{\partial}{\partial x^\nu} \frac{\partial \mathcal{L}}{\partial \psi_{,\nu}} - \frac{\partial \mathcal{L}}{\partial \psi} \right) - \psi_{,\nu,\mu} \frac{\partial \mathcal{L}}{\partial \psi_{,\nu}} = 0, \quad (1.3.29)$$

因此 $T_\mu{}^\nu$ 无散, 进而哈氏量的时间导数为

$$\frac{dH}{dt} = \int \partial_t T_0^0 d^3x = - \int \partial_k T_0^k d^3x, \quad (1.3.30)$$

其中交换导数的前提是规定边界与参数无关. 如果边界无穷远或系统封闭, 则该边界项为零, H 是守恒的, 把 T_0^0 称为**能量密度** ρ_{EM} 是合理的, 进而把 T_0^k 理解为 **3-动量密度**或**能流密度**. 注意, 密度总是形式上含有体积分母, 关键在于相对论中体积的测量会因观者而变, 这体现在动系将观察到体积运动方向上的尺缩效应, 即 $\Delta V = \Delta V_0 / \gamma$, ΔV_0 表示静系体积. 设惯性系下某处单元的 3-速为 u^i , 则 3-动量微元是 $\Delta p^i = \gamma \Delta m u^i = \Delta E u^i$. 以此除全式便知 $\Delta p^i / \Delta V = \rho_{EM} u^i$, 可见 3-动量密度等于能流密度. 而 T_i^j 被解释为**动流密度**, 也即通常讲的应力张量. 综上, 我们将 $T_\mu{}^\nu$ 称为场的 **4-动流密度**或(正则) **能动张量** (energy-momentum tensor).

类似于试验粒子与外场的耦合, 场和源之间也有耦合. 再一次用经典电磁场举例. 假设我们知道电荷间存在一种长程相互作用, 但并不知道支配这一相互作用的 Maxwell 方程, 如何才能找到它呢? 为了迎合 Lorentz 不变性, 这必须是一组至少能做 Lorentz 变换的张量方程, 因此仅仅包含电荷密度的相互作用是不可能存在的. 任何包含 4-电流密度的某个分量的方程, 必须也包含其他分量. 例如, 假设在某系¹来

¹关于观测的细节见第 2 章.

看存在电荷但无电流, 则在新的运动系中电荷和电流都将出现. 故虽然旧系把电磁作用完全归因于电荷, 但新系却将把它归因于电荷与电流两者. 所以任何自称在所有参考系中都适用的规律, 必须同时包括电荷密度和电流密度的贡献. 故若电荷密度激发了场, 电流密度也将如此, 简而言之, J^μ 是电磁场的源并涵盖了全部信息.

它是否确实是逆变矢量?

ρ_e 形式上可表为元电荷乘数密度, 假设电荷量、粒子数在惯性系下一致, 设系统 4-速场 U^μ , 称相对静止系所测电荷密度 ρ_{e0} 是**固有电荷密度**, 则 $J^0 = \gamma\rho_{e0}$, 进而 $J^i = \gamma\rho_{e0}u^i = \rho_{e0}U^i$, 即

$$J^\mu = \rho_{e0}U^\mu. \quad (1.3.31)$$

类似于静质量, ρ_{e0} 是荷电系统上的标量场, 故它是一个逆变矢量. 这样, 场源的矢量性也给我们用矢量描述电磁场提供了充分理由, 进而必须迫使 $A^\mu, F_{\mu\nu}$ 是张量. 类似于粒子动力学, 场的动力学方程至少是二阶的 $\partial^\nu\partial_\nu A^\mu + \cdots \propto J^\mu$, 这意味着自由电磁场的作用量必须出现 $F_{\mu\nu}$, 最简单的标量构造是只使用 $F_{\mu\nu}$, 即 $F_{\mu\nu}F^{\mu\nu}$, 而无需和 A^μ 缩并. 对 A_ν 变分有

$$\delta \int F^{\mu\nu} F_{\mu\nu} d^4x = \int F^{\mu\nu} \delta(\partial_\mu A_\nu) d^4x \simeq - \int \partial_\mu F^{\mu\nu} \delta A_\nu d^4x.$$

而场源方面, 注意 J^μ 也属速度矢量, 故最简构造是 $A_\nu J^\nu$, 它对 A^ν 的变分直接是 $J^\nu \delta A_\nu$, 故

$$\partial_\mu F^{\mu\nu} = 4\pi J^\nu, \quad (1.3.32)$$

其中耦合强度采取几何 Gauss 制 (见 1.5 节), 也可从物理量的重定义或实验测量等角度去理解它. 这给出了 Maxwell 方程组的剩下两条

$$\nabla \cdot \mathbf{E} = 4\pi\rho_e, \quad \nabla \times \mathbf{B} - \partial_t \mathbf{E} = 4\pi\mathbf{j}.$$

对于非真空的情况, 就需要往泛函中加入关于物质的部分, 即

$$S = S_G + S_M, \quad (1.3.33)$$

这里 S_M 表示物质场作用量, 写作

$$S_M = \int \mathcal{L}_M d^4x, \quad (1.3.34)$$

其关于 $g^{\mu\nu}$ 的变分为

欲叠加二者以得到场方程, 以使

$$\delta S = \frac{1}{2\kappa} \int (G_{\mu\nu} - \kappa T_{\mu\nu}) \delta g^{\mu\nu} d^4x = 0,$$

说明可从作用量的视角定义能动张量，满足

$$\delta S_M = -\frac{1}{2} \int T_{\mu\nu} \delta g^{\mu\nu} d^4x, \quad (1.3.35)$$

因而

$$T_{\mu\nu} = -2 \frac{\delta}{\delta g^{\mu\nu}} (\mathcal{L}_M). \quad (1.3.36)$$

对 $g_{\mu\nu}$ 变分时，由 (1.3.35) 式亦有

$$\delta S_M = \frac{1}{2} \int T^{\mu\nu} \delta g_{\mu\nu} d^4x, \quad T^{\mu\nu} = 2 \frac{\delta}{\delta g_{\mu\nu}} (\mathcal{L}_M).$$

1.4 经典引力场论

1.4.1 引力与时空几何

根据引力的特征尝试构建引力场的定量叙述。首先是两类普适性：引力是**万有的**，即物质间皆存在引力；就物体受引力场作用方面还具 **Galileo 性**，即引力场中任何两个试验质点，只要初始位置、速度相同，且除引力外不受力，则此后运动完全相同，与内部结构及组成无关。换言之，就是任意两个试验物体在引力场中的同一点有相同加速度 $\ddot{\mathbf{x}}$ 。注意到并非所有相互作用都有此二性质。电中性物体就无静电力，不满足万有性；考虑静电场中某试验电荷，Galileo 性也不适用：设电荷 q ，则 $\ddot{\mathbf{x}} = (q/m)\mathbf{E}$ 一般不同，除非比荷相同。Newton 的解释是， $\mathbf{f} = m_I \ddot{\mathbf{x}}$ 定义的质量 m_I 应称**惯性质量**，Mach 提出了一种可行的测量性定义：若承认动量守恒（经典范围内可由 Newton 运动学定律导出）并规定好某一物体的质量数值，只需使二者产生相互作用（如碰撞），根据加速度比值便可测得另一物体的质量； $\mathbf{f} = m_G \mathbf{g}$ 定义的质量 m_G 应叫**引力质量**，当然，应还分**主动** (active) 与**被动** (passive) 两类，前者指作为场源时的质量，后者指作为试探质点的质量。此处指被动。引力质量是引力的荷，表征引力作用强弱，而惯性质量表征惯性的大小。联立有加速度 $\ddot{\mathbf{x}} = (m_G/m_I)\mathbf{g}$ ，说明 Galileo 性即指比例 m_G/m_I 与物质本身无关，重新定义 G, m_G 还可使此二概念不再区分： $m_I = m_G \Rightarrow \mathbf{g} = \ddot{\mathbf{x}}$ ，这应对任意物体在任意时刻、地点成立。Newton 将此作为其引力理论的实验事实。

接着 Newton 断言：引力有**保守性**，故可直接用标量的引力势 $\phi(\mathbf{x}, t)$ 指代引力场，满足 $\mathbf{g} = -\nabla\phi$ ；**平方反比律**处处成立¹，即质量密度场 $\rho(\mathbf{x}, t)$ 作为场源所激发的引力势场为

$$\phi(\mathbf{x}, t) = -G \int \frac{\rho(\mathbf{x}', t)}{|\mathbf{x} - \mathbf{x}'|} d^3x', \quad (1.4.1)$$

¹历史上出现过诸多修正公式，如 Yukawa 势 $\phi \propto e^{-r/C}/r$ (C 是常数)，但无论从实验精度还是理论简洁性上都是平方反比律占优。Yukawa 势更多用于量子场论。

其中规定自然边界条件 $\lim_{|\mathbf{x}| \rightarrow \infty} \phi = 0$. 记 $\mathbf{r} := \mathbf{x} - \mathbf{x}'$ 和 $r = |\mathbf{r}|$, 注意 $r \neq 0$ 时 $\nabla(1/r) = -\mathbf{r}/r^3 \neq \mathbf{0}$, 这就退回了平方反比律. 下证保守性、平方反比律必导致 Poisson 方程 $\nabla^2 \phi = 4\pi G\rho$. 试图交换导数和积分将会出现 $\nabla^2(1/r)$, 然而 $r \neq 0$ 时 $\nabla^2(1/r) = 0$ 且 $r = 0$ 处发散, 故必须用极限过程严格定义它的体积分. 考虑

$$\nabla^2 \frac{1}{\sqrt{r^2 + a^2}} = \frac{-3a^2}{(r^2 + a^2)^{5/2}}, \quad \frac{1}{r} = \lim_{a \rightarrow 0} \frac{1}{\sqrt{r^2 + a^2}}. \quad (1.4.2)$$

选择最标准的反常积分定义, 即取球形区域 $B(R) = \{\mathbf{r} | 0 < r < R\}$ 的极限:

$$\int \nabla^2 \frac{1}{\sqrt{r^2 + a^2}} d^3r = \lim_{R \rightarrow \infty} \int_{B(R)} \frac{-3a^2 d^3r}{(r^2 + a^2)^{5/2}} = -4\pi \int_0^\infty \frac{3a^2 r^2 dr}{(r^2 + a^2)^{5/2}} = -4\pi.$$

类似地,

$$\begin{aligned} \left| \int f(\mathbf{r}) \nabla^2 \frac{1}{\sqrt{r^2 + a^2}} d^3r + 4\pi f(\mathbf{0}) \right| &= \left| \int \frac{3a^2(f(\mathbf{r}) - f(\mathbf{0}))}{(r^2 + a^2)^{5/2}} d^3r \right| \\ &\leq \max |\nabla f| \int \frac{3a^2 r d^3r}{(r^2 + a^2)^{5/2}} = 4\pi \max |\nabla f| \int_0^\infty \frac{3a^2 r^3 dr}{(r^2 + a^2)^{5/2}} = 8\pi |a| \max |\nabla f|, \end{aligned}$$

因此

$$\lim_{a \rightarrow 0} \int f(\mathbf{r}) \nabla^2 \frac{1}{\sqrt{r^2 + a^2}} d^3r = -4\pi f(\mathbf{0}), \quad (1.4.3)$$

这符合 δ 函数定义¹, 记作 $\nabla^2(1/r) = -4\pi\delta(\mathbf{r})$, 立即有 $\nabla^2\phi(\mathbf{x}, t) = 4\pi G\rho(\mathbf{x}, t)$. 若还原为场强表述, 则 Poisson 方程等价于两个方程

$$\nabla \cdot \mathbf{g}(\mathbf{x}, t) = -4\pi G\rho(\mathbf{x}, t), \quad \nabla \times \mathbf{g}(\mathbf{x}, t) = \mathbf{0}, \quad (1.4.4)$$

而物体受场作用体现在试验质点满足 $\ddot{\mathbf{x}} = \mathbf{g}$, 此即 Newton 引力场论的框架. 上式与电磁学相似但又与其处境不同. 在电磁学中, 象征平方反比律的 Coulomb 定律仅是静态解, 电场、电势应分别记作 $\mathbf{E}(\mathbf{x}), \varphi(\mathbf{x})$, 但 (1.4.1) 式却是 $\phi(\mathbf{x}, t)$ 形式, 这说明 \mathbf{x} 处 t 时的 ϕ 同时由空间各点在 t 时的 ρ 决定, 引力场传播无限快. 电磁学能自然地排除超距解, 因为还有涉及磁场、时间的方程参与约束. Newton 引力场论没有这种外援, 因为默认保守性、平方反比律成立. 我们必须适当放弃保守性、平方反比律 (现在只作为近似结论), 保留引力更为深刻的万有性、Galileo 性.

注意到, 引力场强就是试验质点的加速度, 加速度又正恰是运动学概念, 仅涉及时空坐标, 此即引力同时空之联系. 时空如何从更深层的角度解释 Galileo 性? 考虑到物理学的目的是研究客体演化规律, 手段是对客体建立定量模型. 客体的演化总是在某种时空上进行, 而 Galileo 性说明任何物体在引力场中将齐步地运动, 这种无关个性的集体行为强烈暗示着引力不仅与时空有联系, 还直接是时空的自身性质, 因为齐步意味着引力场中的任意自由落体者步调一致, 彼此没有相对加速度, 故就自

¹ 许多电磁学教材会随意地讨论这种非连续的质点密度场, 其严格基础就是该式在形式上符合 Gauss 定理.

由落体者 A 而言, 总可找到其附近的自由落体系 B , 使得 A 相对于 B 做的任何事情都仿佛像是在无引力场的惯性系中一样. 这里巧妙在, 不需要研究二者相对于某个整体参考系 (如大地) 的加速度, 这涉及该参考系是否“惯性”的问题, 只研究相对运动就可避免这一点. 设想在引力场中自由落体的一部足够小的封闭电梯及其内部的一位乘客, 则乘客将感到失重. 假设无法看到外界, 则在电梯系中无法做任何力学实验来分辨电梯到底是引力场中的自由落体系, 还是无引力场的惯性系, 因为实验结果是等效的. 注意, 要求力学实验只涉及经典内容, 而无电磁或量子性质, 且其本身应是非引力的: 实验室内物体间的引力可忽略, 但室内可存在由室外物体产生的

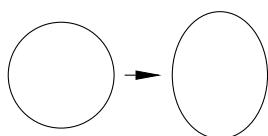


图 1.13: 潮汐效应

引力场; 电梯系亦应足够小, 这种等效性是局域的. 有多局域呢? 考虑到 m_I 代表着“惯性”质量, 故还有一种等价说法: 引力可局域等效为**惯性力** (inertial force). 考虑线加速度致使的惯性力, 其力线是均匀且平行的, 而引力场的力线在较大尺度上往往是不均或会聚的. 若稍微扩大某个电梯的范围, 以观测到两个相距足够远的空间点, 则二者的加速度在大小和方向上将明显不同. 将一非质点的物体置于该电梯中, 其不同部分受力不均, 可以观察到相对形变现象, 表现为水平挤压和竖直拉伸. 这与月球引力使地表海洋产生**潮汐效应**的原因一致, 故邻近质元间相对形变力也称**潮汐力** (tidal force), 相对加速度称**潮汐加速度**. 显然电梯系无法用惯性现象解释它. 可见, 等效性局域成立准确地指其只能成立于单个空间点. 总之, Galileo 性等价于这种局域等效性, 这一点被 Einstein 于 1907 年察觉, 他称之为**等效原理** (equivalence principle)¹.

在 Newton 体系中, 引力场的自由落体系是非惯性系, 因而和无引力场的惯性系的动力学应截然不同, 然而要想检测出一个参考系是否惯性, 又必须观察惯性力. 自由落体系下同时出现的引力和惯性力精确抵消, 故现在看来是不可物理观测的. 干脆就抛弃这些累赘, 即扔掉 Newton 所偏好的整体绝对静止系, 而关注自由落体质点的**局域惯性系**, 即其中没有可观测惯性效应的参考系. 宇宙中所有物质都处于引力场之中, 严格来讲找不到真正的惯性系, 但这也无伤大雅, 因为首先我们有广义协变性, 其次等效原理使仅存在引力的任何物体, 无论在时空何处都有其局域惯性系, 不外乎能够近似为惯性系的范围可能不同: 若试验质点处于无引力场中, 则该局域惯性系的大小可无限拓展, 即整体惯性系; 若物质集团的尺度较大, 密度较稀疏, 则引力场较弱, 局域惯性系的范围仍然较大, 如地球、太阳系、银河系的局域惯性系范围应依次增大; 无论引力场强弱, 局域惯性系总可定义在一个自由落体的质点所占

¹顺便提一句, Newton 这个天才是知道这件事的: “无论诸物体彼此间以何种方式运动, 若它们被沿着平行线的相等加速力所推动, 则它们都将继续彼此间的运动, 遵循的方式就如同没有那些力作用一样.” ——《原理》, 推论 VI “运动定律” [1]. Newton 用这个推论来计算太阳系中月球的复杂运动. 在地球系下, 惯性力和太阳引力在良好近似下相互抵消, 那么月球就遵循一个 Kepler 轨道.

据的空间点上. 这说明, 应由引力场决定参考系是否惯性, 进而决定不同参考系相对彼此运动的方式, 并阐明不同参考系间的关系. Newton 的绝对运动不是相对于绝对空间的运动, 而是相对于引力场所决定的局域惯性系的运动. 那么能否仅考虑使用 3 维空间语言呢? 首先, 注意物体的运动要么是空间线, 要么是空间点, 这般看待未免太过割裂, 毕竟我们希望研究对象是统一的, 应尝试将点情形也推广为线——静止物体仍在时空中划出一条世界线, 便成功统一. 其次, 局域惯性系一般只成立于一个空间点, 在 3 维语言下何谈坐标系呢? 但考虑时空就可解决这一问题, 我们仍可在某自由落体的质点世界线上建立坐标系.

假设在时空中设置了引力场, 由等效原理, 对任意做自由落体的世界线 G , 总存在局域惯性系 $\{\xi^\alpha\}$, 虽然其空间坐标只能邻域延伸, 但时间坐标却可较大尺度地延伸, 比如可采用固有时, 用它参数化 G , 这样 G 在 $\{\xi^\alpha\}$ 系表为 $\xi^\alpha(\tau)$. 我们在局域惯性系 $\{\xi^\alpha\}$ 中按 $\mathbb{R}^{1,3}$ 的方式去定义 4-加速 $\tilde{A}^\alpha = d\tilde{U}^\alpha/d\tau = d^2\xi^\alpha/d\tau^2$, 再按逆变矢量的性质变换到任意坐标系. 这样, 等效原理直接给出引力场中的质点运动规律 (不考虑其它相互作用): 自由落体运动在局域惯性系下恒有 $\tilde{A}^\alpha = 0$, 进而 $A^\mu = 0$, 即自由落体质点走测地线. 可见局域惯性系首先是四维时空中的正交归一系, 其次还要在线上处处使克氏符为零, 因而局域惯性系准确来说是测地线上的正交归一系, 简称**测地系**. 注意闵氏度规对应于无引力场: 若时空是 $\mathbb{R}^{1,3}$, 不妨取 $\{x^\mu\}$ 为惯性系, 自由落体退化为直线, 而我们的确承认 $\mathbb{R}^{1,3}$ 中不受力质点走直线, 因此无引力; 反之若无引力, 总存在某个坐标系 $\{x^\mu\}$, 其中由于惯性直线运动是自由落体在无引力时的特例, 则联络系数全零, 进而可坐标重标度将度规归一化到 $\eta_{\mu\nu}$, 则度规是闵氏. 而若度规不是闵氏的, 则必存在非零 $g_{\mu\nu,\lambda}$, 否则总可归一化到 $\eta_{\mu\nu}$ 而与“度规不是闵氏的”矛盾, 故存在非零联络系数, 取自然坐标系 $\{x^\mu\}$ 就可知运动不为直线, 说明存在引力. 可见在时空中设置引力场需要改变度规. 一个度规对应一套几何体系, 因此引力是时空的几何效应. 等效原理要求度规在正交归一系下的分量为 $\eta_{\mu\nu}$, 这种度规就称为 **Lorentz 度规**. Lorentz 度规所对应的体系称 **Lorentz 几何**. Poincaré 和 Riemann 都尝试构建过仅用空间几何的引力理论, 但未能成功. 引力的几何理论的关键就在于补上时间这一维. 时间维度对引力亦有贡献.

尝试不改变度规而用某个其它物理量表征引力场一定程度上也行, 但这几乎等价于定义新度规, 见 1.3 节.

这一事实表明, 引力场的确不能用矢量场描述, 这是因为从形式上 A^μ 这套理论的经典贡献是 $A^\mu U_\mu$, 这必将导致这种假象 A^μ 的性质与经典电磁场高度一致, 比如这必然使具有相同场荷之间的粒子导致斥力. 此外, 后面将看到密度作为源在相对论下几乎只能用二阶张量表示, 我们只能用二阶张量场作为引力相互作用的载体, 然而假设 $\kappa h_{\mu\nu}$ 是对于引力场中的试验粒子, 若尝试直接相加,

欲用非闵氏度规建立引力场论, 则引力定律中必须让研究对象 $g_{\mu\nu}$ 出现, 我们只

能期望 $g_{\mu\nu}$ 是动力学量, 它所满足的某个方程才可能有广义不变性. 其次, 正如质点运动方程是常微分方程, 场的 (运动) 方程是一个偏微分方程, 故包括 $g_{\mu\nu}g_{\alpha\beta}$, $g_{\mu\nu,\lambda,\kappa}$ 等 $g_{\mu\nu}$ 的衍生量也必须存在于方程中. 同理, 它们都是动力学量. 我们把度规及其衍生量统称为**时空几何量**, 这样物质状态将的确影响时空几何量¹, 我们彻底地推出了引力万有性的另一方面, 即物质作为源产生引力场. 给定物质后, 这种方程的未知量就是度规, 因此度规应看成先验存在但尚未算出的. 这称为**引力场方程**或**场方程** (field equation). 从狭义相对论到广义相对论, 就相当于把闵氏度规换成一个尚未算出的度规, 而场方程在处处无物质时应自洽地给出已有的闵氏度规.

1912 年 Einstein 也开始全面转向这一目标, 不过, $g_{\mu\nu} = g_{\nu\mu}$ 说明度规的独立分量数为 $\binom{4+1}{2} = 10$, 这个数量使他在数学上犯难, 因而求助于其在苏黎世高工的同班同学兼朋友 Grossmann. 他们欲仿造已有的微分几何的蓝图并意外发现工具非常齐备. Einstein 说 “这几乎就像是命中注定的工具”.

因此, 一个描述引力 (甚至涉及其它相互作用) 的物理方程必须与时空有所关联, 这只能通过时空几何量办到. 并且为了让方程有广义不变性, 我们关注的是时空几何量中的张量. 此即广义不变性的要害.

观察相对性原理最原始的说法, 在 $\mathbb{R}^{1,3}$ 中谈论绝对位置或绝对速度当然没有意义, 因为 $\mathbb{R}^{1,3}$ 到处都是完全一样, 没有可辨别的特征, 这样 Lorentz 不变性的确可以支撑足够的相对性. 类似地, 一种观点是认为 “广义相对性” 可以表述为没有实验能检测出非惯性系和惯性系的内在区别, 进而各非惯性系间也不能区分, 换言之, 不能谈及绝对加速度, 正如在普通的运动学中不能谈论一个参考系的绝对速度一样. 究竟什么叫不能区分? 一个非惯性系的产生是靠着相互作用产生, 然而假设物质间只靠引力联系, 这样所有观者一定自由落体, 那么前文对等效原理的分析已表明仍不能做任何力学实验以区分, 进而无法判断一个参考系是否惯性, 不如摒弃整体参考系而只关注局域惯性系. 但注意, 不可区分的局域性已经严苛到, 如果引力场非常强, 那么只能在一条世界线上成立. 这样, 观者只需参照其时空邻域内的几何特征, 如潮汐力, 以确定其位置和速度, 这样在不同的时空邻域间可作出绝对的区分, 因此消除了相对性. 可见 Einstein 称其理论为 “广义相对论” 的理由不能完美成立, 尽管这一名称已广为流传. 实际上, 狭义相对论已然是最大限度的 “相对论” 了——它已具有最大的时空对称性. 广义不变性不是一种相对性原理; 更确切地说它是一种动力学原理, 从而对引力与物质场规律加以可能的限制 (见 1.4.4 节). 目前来讲, 很难找到一个真正的 “广义相对性” 理论, 除非找到一种既不同于 Aristotle 又不同于 Galileo 和 Newton 的对惯性的诠释, 这至今不能完满解答 (见 1.4.6 节). 当然, 这也取决于究竟如何陈述 “广义相对性”, 目前人们对其理解仍存分歧.

¹ 一个典型例子是, 尽管 $\Gamma_{\mu\nu}^\lambda$ 是坐标系自身信息, 但在有引力场时, $\Gamma_{\mu\nu}^\lambda$ 却能体现绝对的引力现象, 因为对任意度规, $\Gamma_{\mu\nu}^\lambda$ 只能在一处处, 或至多在自由落体世界线上找到局部坐标系使其全零. 非闵氏度规场下, 不可能找到整体坐标系使 $\Gamma_{\mu\nu}^\lambda$ 处处为零. 因此常说 $\Gamma_{\mu\nu}^\lambda$ 代表引力场强, 进而 $g_{\mu\nu}$ 代表引力势.

虽然上述讨论清楚表明 Einstein 理论很难彻底具备某种“广义相对性”，但它具有“狭义相对性”，因为广义不变性包括了 Lorentz 不变性. 这样 Einstein 理论是一个像狭义相对论那样的相对性理论. 在上面的例子中这意味着，观者可以利用几何特征来测量其相对地球的速度，但不可能测量出地球的绝对速度，因为各星系及宇宙平均质量分布所产生的引力场在大尺度上很弱，太阳系处在一个近似的闵氏时空中. 综上，Einstein 引力理论是“相对论”，但并不比狭义相对论更“相对论”. Wheeler 很早就提议将该理论称为**几何动力学** (geometrodynamics) 以强调其本质，即与物质相互作用的动力学几何.

1.4.2 潮汐与曲率

应当存在一种代表时空非闵程度的张量，从而可直接衡量引力场强弱. 首先读者就应注意到，潮汐现象并不适用于等效原理，换言之，潮汐不能用坐标系消除. 这意味着，如果我们寻求一个尽可能简单的定义，那么该张量场在一点就应具备潮汐意义. 在 $\mathbb{R}^{1,3}$ 中，一般用惯性系表出某个物理定律的偏微分方程，但这样仅有 Lorentz 不变性. 为了构造一个广义不变的偏微分方程，我们得尝试直接使用协变导数而非单个偏导数，来构造该张量. 设 $g_{\mu\nu}$ 在某点邻域内有定义，以此为基础研究一点处可能的张量，但 $g_{\mu\nu}$ 可用坐标系消为 $\eta_{\alpha\beta}$ ，而 $\Gamma_{\mu\nu}^\lambda$ 是 $g_{\mu\nu,\lambda}$ 级的，可类似消成零，这意味着我们至少要考虑二阶导数，即场方程至少是二阶的. 不妨先考虑¹二阶情形，这样甚至能借助恰为二阶的 Poisson 方程找到引力场方程的度规语言. 这涉及二阶协变导数，但它作用于什么呢？我们不能立即知晓答案，因此要回到潮汐理论中去.

设物体由连续介质描述，它在时空上延伸为一个世界管，假设其内部的相互作用较弱，则各质元仅受外引力场作用而走各自的测地线. 这样，世界管就成了所谓的**测地参考系**或**测地管**. 要研究相对形变，就是取其中某条测地线为参考，称为**基准观者** (fiducial observer)，它邻域内的测地线称为**邻居** (neighbor). 研究它与某邻居的间距随固有时的变化规律. 二者若初态平行，则发生相对位移意味着存在相对加速度，故无论初态如何，研究邻近间距沿基准观者的测地线的二阶协变导数即可. 如图 1.14，假定该参考系可按某种方式（比如对钟）在一定范围内约定各测地线的时钟设定，规定坐标时都记为基准观者 p 的固有时 τ ，这样基准观者 p 的坐标式记作 $x^\lambda(\tau)$ ，某邻居 q 的坐标式记作 $x^\lambda(\tau) + \delta x^\lambda(\tau)$ ，其中 $\delta x^\lambda =: \epsilon S^\lambda$ 代表微小间距， S^λ

¹的确可能存在这样一种情况，需要用高阶微分方程去描述自然，但 Taylor 展开思想支撑着物理学所用的几乎任何近似技术，不妨先研究低阶理论.

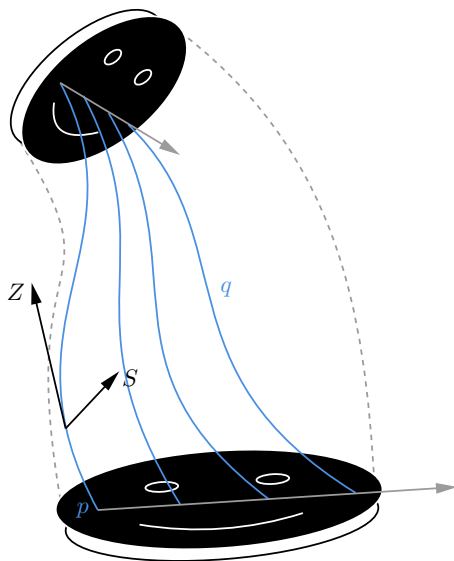


图 1.14: 测地偏离

称为**偏离矢量** (separation vector). 记 $Z^\lambda = dx^\lambda/d\tau$, 则二者的运动方程为

$$\frac{dZ^\lambda}{d\tau} + \Gamma_{\mu\nu}^\lambda|_p Z^\mu Z^\nu = 0, \quad (1.4.5)$$

$$\frac{dZ^\lambda}{d\tau} + \epsilon \frac{d^2 S^\lambda}{d\tau^2} + \Gamma_{\mu\nu}^\lambda|_q \left(Z^\mu + \epsilon \frac{dS^\mu}{d\tau} \right) \left(Z^\nu + \epsilon \frac{dS^\nu}{d\tau} \right) = 0. \quad (1.4.6)$$

将 $\Gamma_{\mu\nu}^\lambda|_q$ 按 p 展开为

$$\Gamma_{\mu\nu}^\lambda|_q = \Gamma_{\mu\nu}^\lambda|_p + \Gamma_{\mu\nu,\kappa}^\lambda|_p \delta x^\kappa = \Gamma_{\mu\nu}^\lambda|_p + \epsilon \Gamma_{\mu\nu,\kappa}^\lambda|_p S^\kappa, \quad (1.4.7)$$

现在就可省去时空点下标. 将两式相减并略去高阶有

$$\frac{d^2 S^\lambda}{d\tau^2} = -\Gamma_{\mu\nu,\kappa}^\lambda Z^\mu Z^\nu S^\kappa - 2\Gamma_{\mu\nu}^\lambda \frac{dS^\mu}{d\tau} Z^\nu, \quad (1.4.8)$$

则 S^λ 沿测地线的二阶协变导数为

$$\begin{aligned} \frac{D^2 S^\lambda}{d\tau^2} &= \frac{d}{d\tau} \frac{DS^\lambda}{d\tau} + \Gamma_{\mu\nu}^\lambda \frac{DS^\mu}{d\tau} Z^\nu \\ &= \frac{d}{d\tau} \left(\frac{dS^\lambda}{d\tau} + \Gamma_{\kappa\mu}^\lambda S^\kappa Z^\mu \right) + \Gamma_{\mu\nu}^\lambda \left(\frac{dS^\mu}{d\tau} + \Gamma_{\kappa\sigma}^\mu S^\kappa Z^\sigma \right) Z^\nu \\ &= \frac{d^2 S^\lambda}{d\tau^2} + \Gamma_{\kappa\mu,\nu}^\lambda Z^\mu Z^\nu S^\kappa + 2\Gamma_{\mu\nu}^\lambda \frac{dS^\mu}{d\tau} Z^\nu + \Gamma_{\kappa\nu}^\lambda S^\kappa \frac{dZ^\nu}{d\tau} + \Gamma_{\mu\nu}^\lambda \Gamma_{\kappa\sigma}^\mu Z^\sigma Z^\nu S^\kappa \\ &= (\Gamma_{\kappa\mu,\nu}^\lambda - \Gamma_{\mu\nu,\kappa}^\lambda) Z^\mu Z^\nu S^\kappa - \Gamma_{\kappa\nu}^\lambda S^\kappa \Gamma_{\mu\sigma}^\nu Z^\mu Z^\sigma + \Gamma_{\sigma\nu}^\lambda \Gamma_{\kappa\mu}^\sigma Z^\mu Z^\nu S^\kappa \\ &= R_{\mu\nu\kappa}^\lambda Z^\mu Z^\nu S^\kappa, \end{aligned} \quad (1.4.9)$$

其中

$$R^\lambda_{\mu\nu\kappa} := \Gamma^\lambda_{\kappa\mu,\nu} - \Gamma^\lambda_{\mu\nu,\kappa} + \Gamma^\lambda_{\sigma\nu}\Gamma^\sigma_{\kappa\mu} - \Gamma^\lambda_{\kappa\sigma}\Gamma^\sigma_{\mu\nu}. \quad (1.4.10)$$

式 (1.4.9) 称为**测地偏离** (geodesic deviation) **方程**. $R^\lambda_{\mu\nu\kappa}$ 是 $g_{\mu\nu,\lambda,\kappa}$ 级的, 且确实是张量, 因为 (1.4.9) 的式左是协变导数, 式右 Z^μ, S^κ 是逆变矢量. 这称为**潮汐张量**, 表征潮汐形变程度, 亦即度规的非闵程度: 只需取某个惯性系使克氏符处处为零, 进而潮汐张量在该系和任意坐标系都为零. 数学上称作 **Riemann 内禀曲率** (intrinsic curvature) 或**曲率**, 曲率为零时称度规为**平直的** (flat), 非零时称**弯曲的** (curved), 物理学可以借用这些词而称**时空曲率**, 时空分为**平直时空**和**弯曲时空**. 下面解释这些称呼的意义. 按 1.1.2 节的思想, 我们预先规定 \mathbb{E}^3 是生活意义的平直, 以它为绘图基准可判断某条线、某张面的弯曲. 考虑一个球面, 其上任意切矢沿自身平移所得测地线即过球心的大圆. 设从赤道某两处垂直延伸两条大圆弧, 它们的间距逐渐缩短并交于极点. 在一张平面上, 初态平行的直线不可能改变间距. 另一方面, 按如下操作可计算球面上的线元: 先用球坐标 $\{r, \theta, \phi\}$ 改写欧氏度规为

$$d\ell^2 = dr^2 + r^2 d\Omega^2, \quad d\Omega^2 := d\theta^2 + \sin^2 \theta d\phi^2, \quad (1.4.11)$$

当然, 此处 Ω 不是微积分学的立体角 (solid angle), 立体角满足的是 $d\Omega = \sin \theta d\theta d\phi$. 再取 $r = R$ 就有所谓的**球面度规** $d\ell^2 = R^2 d\Omega^2$, 由此可计算球面曲率不为零. 像这样, 约定**背景空间的嵌入** (imbedded) **曲面上的长度**是用**背景度规**衡量的, 这种度规称为背景度规的**诱导** (induce) **度规**. 可见度规在非欧/闵时, 内禀曲率非零而使测地间隔变化. 但内禀曲率其实只由度规和坐标系计算而来, 甚至无需嵌入某个高维背景来定义度规和曲率——对于四维时空, 我们就只能抛弃嵌入思想来想象它. 内禀曲率只反应内禀的弯曲程度而与曲面本身无关: 球面在内禀意义下弯曲, 但圆柱面在内禀意义下平直 (借柱坐标发现诱导度规平直), 尽管它在直观上弯曲. 数学上, 描述圆柱面这种对象的弯曲程度就需要嵌入思想, 称为**外** (extrinsic) **曲率**. 一个直观理解是, 凡能连续地展平的曲面, 尽管外曲率可不为零, 但内禀曲率为零, 真正不可连续展平的曲面内禀曲率、外曲率都非零. 读者可观察纸塑地球仪是如何尝试将平面纸张大致贴合球面的.

为获得物质与曲率的联系, 还需退回到 Newton 理论中. 在测地参考系上建立这样的坐标系 $\{\xi^\alpha\}$, 其中以每一质元的世界线为时间轴, 规定它们的固有时零点形成的超曲面处处与 S^κ 相切, 这就是局域惯性系的一种延伸. 显然 $\tilde{Z}^\mu = (1, 0, 0, 0)$ 且 $\tilde{S}^0 = 0$. 进而再根据定义得反称性

$$R^\lambda_{\mu\nu\kappa} = -R^\lambda_{\mu\kappa\nu}, \quad (1.4.12)$$

测地偏离方程在该坐标系下写作

$$\frac{D^2 S^\lambda}{d\tau^2} = -R^\lambda_{\mu\kappa\nu} Z^\mu S^\kappa Z^\nu \implies \frac{d^2 \tilde{S}^i}{dt^2} = -\tilde{R}^i_{0j0} \tilde{S}^j. \quad (1.4.13)$$

为了得到 Newton 理论, 我们要求每一根测地线足够低速, 故该参考系成为了惯性系. 这样 $\tilde{S}^0 = 0$ 可以理解为在同一时间测量质元间的相对加速度¹. 在惯性系中, Newton 引力理论给出

$$\frac{d^2\xi^i}{dt^2} = -\delta^{ij} \frac{\partial\phi}{\partial\xi^j} = -\tilde{\partial}^i\phi, \quad \frac{d^2\xi^i}{dt^2} + \epsilon \frac{d^2\tilde{S}^i}{dt^2} = -\tilde{\partial}^i\phi - \epsilon \tilde{\partial}_j \tilde{\partial}^i\phi \tilde{S}^j, \quad (1.4.14)$$

其中右侧是按相同理由将附近的 $\tilde{\partial}^i\phi$ 展开到一阶. 用 Poisson 方程就可得到曲率与密度的联系为 $\tilde{R}^\sigma_{0\sigma 0} = \tilde{R}^i_{0i0} \approx \tilde{\partial}_i \tilde{\partial}^i\phi = 4\pi\rho$ (几何制), 其中注意反称性给出 $R^\mu_{\nu 00} = 0$, 进而 $R^\sigma_{0\sigma 0} = R^i_{0i0}$. 可见

$$R_{\mu\nu} := R^\sigma_{\mu\sigma\nu} \quad (1.4.15)$$

很有意义, \tilde{R}_{00} 在 Newton 条件下近似为 $4\pi\rho$, 数学上称之为 **Ricci 曲率**, 它当然也是张量.

现在就可给等效原理的局域性一个更为宽松的实验说法. 测地线 $G(\tau)$ 的弯曲使线上曲率不可能处处为零, 因此局域坐标系只能保证 $\tilde{\Gamma}^\lambda_{\mu\nu}$ 严格在线上为零. 实际上, 假若在邻域 U 内等于零, 则借助邻域按 (1.4.10) 式计算曲率, 则显然在线上处处为零且与坐标系无关, 矛盾. 因此, 只要邻域足够大, 时空的弯曲便可体现. 此即数学上对“局域指空间点”的解释. 然而, 从操作层面看, 观者 G 往往又需要借助时空邻域才能做实验, 毕竟实验室终归不能完全缩成一点, 实验亦不能瞬间做成. 好在只要 U 尺寸远小于时空的曲率半径², 同时也远小于对曲率半径的测量精度时, 这样即使时空范围未严格至一点, 但曲率也不会被实验观测到. 此即实验所理解的局域. 以圆为例, 微小弯曲圆弧可用该处相切的平直线段近似代替, 尽管该处曲率的确不为零. 若实验范围较大, 可通过测量潮汐 (只需按 Newton 方式, 见第 ?? 章) 来确定曲率张量.

为何选择曲率? 我们已然从前文窥见了 Riemann 曲率具备各种代数性质. 下面将看到, 正是这些代数性质致使我们选择用曲率研究引力. 实际上, Riemann 曲率是唯一可由度规二阶导线性组合而成的张量. Riemann 曲率的分量有 $4^4 = 256$ 之多, 但它具有若干性质, 使其许多分量往往重复或为零. 首先前文已提过反称性 (1.4.12) 式, 其次是所谓的 **Ricci 循环恒等式** (cyclic identities)

$$R^\lambda_{\mu\nu\kappa} + R^\lambda_{\nu\kappa\mu} + R^\lambda_{\kappa\mu\nu} = 0. \quad (1.4.16)$$

只需默认无挠性并运用取局域坐标系的技巧, 这样只会留下 $\tilde{R}^\lambda_{\mu\nu\kappa} = \tilde{\Gamma}^\lambda_{\mu\kappa,\nu} - \tilde{\Gamma}^\lambda_{\mu\nu,\kappa}$.

证明同样只需默认无挠性并取局域惯性系. 此外, **降指标 Riemann 曲率** $R_{\lambda\mu\nu\kappa} = g_{\lambda\sigma} R^\sigma_{\mu\nu\kappa}$ 是一个常用概念, 它与原先 $R^\lambda_{\mu\nu\kappa}$ 有等价的独立分量数, 因此有时也称

¹一般的 S^κ 十分任意而不一定满足对钟的正交要求.

²经典力学中曲线可视为一系列圆周运动, 半径便是曲率半径; 曲面曲率可用球面去衡量. 此概念可推至高维.

Riemann 曲率. 它使我们能看到更多性质, 如反称性

$$R_{\lambda\kappa\mu\nu} = -R_{\kappa\lambda\mu\nu}. \quad (1.4.17)$$

直接证明它亦可从局域坐标系入手, 故又可推出 $R^\sigma{}_{\sigma\mu\nu} = g^{\lambda\sigma} R_{\lambda\sigma\mu\nu} = 0$.

实际上, $R_{\mu\nu}$ 的确是 $R^\lambda{}_{\mu\nu\kappa}$ 唯一非平凡缩并. 首先反称性给出 $R^\sigma{}_{\mu\nu\sigma} = -R_{\mu\nu}$. 最后最重要的是前一对指标和后一对指标的对称性

$$R_{\mu\nu\lambda\kappa} = R_{\lambda\kappa\mu\nu}. \quad (1.4.18)$$

同理默认无挠性并取局域惯性系即可. 包括式 (1.4.12) 在内的这四条性质是我们能得到的所有代数性质了. 注意, 若 Ricci 恒等式存在相同指标, 则总与另三条性质重复, 故常考虑指标各不相同的非平凡 Ricci 恒等式. 假设时空维数是 n , 显然至少需要两个不同指标才有非零分量. 给定 a, b , 由反称性知只有 R_{abab} ; 给定 a, b, c , 由反称性和对称性知有 $R_{abac}, R_{abbc}, R_{acbc}$ 三类; 给定 a, b, c, d , 前一对指标选完就确定了后一对, 由对称性知选法有 $R_{abcd}, R_{acbd}, R_{adbc}$, 但给定两个就能按非平凡 Ricci 恒等式确定第三者, 故独立分量有两类. 最后乘上各自组合数, 则独立分量数共为

$$\binom{n}{2} + 3\binom{n}{3} + 2\binom{n}{4} = \frac{n^2(n^2 - 1)}{12}.$$

对于四维时空, 这说明 Riemann 曲率只有 20 个独立分量. 另一方面, 若曲率非零, 我们可以预料无法通过任何坐标变换使 $g_{\mu\nu, \lambda, \kappa}$ 全部消除. 具体有多少种不能消除? 以某点 p 为某系 $\{x^\mu\}$ 原点, 考虑 $x^\mu(x'^\alpha)$ 形式的函数 (而非反之) 以迎合度规的协变变换, 任意原点不变的光滑坐标变换可展为

$$x^\mu = \frac{\partial x^\mu}{\partial x'^\alpha} x'^\alpha + \frac{1}{2} \frac{\partial^2 x^\mu}{\partial x'^\alpha \partial x'^\beta} x'^\alpha x'^\beta + \frac{1}{6} \frac{\partial^3 x^\mu}{\partial x'^\alpha \partial x'^\beta \partial x'^\gamma} x'^\alpha x'^\beta x'^\gamma + \cdots \quad (1.4.19)$$

将这些偏导数简记作 $A^\mu_\alpha, B^\mu_{\alpha\beta}, C^\mu_{\alpha\beta\gamma}$ 等. 为排除掉 $g_{\mu\nu}, g_{\mu\nu, \lambda}$ 的改变使 $g_{\mu\nu, \lambda, \kappa}$ 改变的那部分干扰, 我们要考虑能在 p 点保持 $g'_{\mu\nu} = g_{\mu\nu}, g'_{\mu\nu, \lambda} = g_{\mu\nu, \lambda}$ 的坐标变换, 因此只需取 $A^\mu_\alpha = \delta^\mu_\alpha$ 且 $B^\mu_{\alpha\beta} = 0$. 注意 $C^\mu_{\alpha\beta\gamma}$ 的协变指标均有对称性, 独立分量数为 $n\binom{n+2}{3}$, 而 $g_{\mu\nu, \lambda, \kappa}$ 的 μ, ν 和 λ, κ 有各自对称性, 独立分量数为 $\binom{n+1}{2}^2$. 可见, $C^\mu_{\alpha\beta\gamma}$ 最多只能自由调节 $n\binom{n+2}{3}$ 种 $g_{\mu\nu, \lambda, \kappa}$, 还剩下

$$\binom{n+1}{2}^2 - n\binom{n+2}{3} = \frac{n^2(n^2 - 1)}{12}.$$

因此存在 $n^2(n^2 - 1)/12$ 种独立的 $g_{\mu\nu, \lambda, \kappa}$, 它们不受 $C^\mu_{\alpha\beta\gamma}$ 选择的影响¹. 若我们想构造基于 $g_{\mu\nu, \lambda, \kappa}$ 的张量, 我们只能使用这些进行线性组合, 因为张量变换律只包括

¹ 请读者按此思想, 尝试证明高维空间中测地坐标的存在性. 注意这并不显然, 我们在四维时空中是用物理学理解它的. 这样, 总能在某测地坐标 $\{\xi^\alpha\}$ 基础上找到无限多测地坐标, 形如 $\xi^\mu_0 + \Lambda^\mu{}_\alpha \xi^\alpha + \frac{1}{6} C^\mu_{\alpha\beta\gamma} \xi^\alpha \xi^\beta \xi^\gamma + \cdots$, 其中 $\xi^\mu_0, \Lambda^\mu{}_\alpha$ 等无关 ξ^α , 即只要求 A^μ_α 是 Lorentz 变换且 $B^\mu_{\alpha\beta} = 0$, 从而保持 $g_{\mu\nu} = \eta_{\mu\nu}$ 且 $g_{\mu\nu, \lambda} = 0$.

$A^\mu_\alpha = \delta^\mu_\alpha$ 而不涉及 $C^\mu_{\alpha\beta\gamma}$. 利用这 $n^2(n^2-1)/12$ 个可用的 $g_{\mu\nu,\lambda,\kappa}$ 能线性构造出的张量有且仅有 Riemann 曲率, 它恰好具有 $n^2(n^2-1)/12$ 个独立分量, 用完了所有可用的 $g_{\mu\nu,\lambda,\kappa}$.

四维时空 Riemann 曲率的 20 个分量或许在处理上仍然不便, 我们可用缩并将 Riemann 曲率进行压缩从而保留有用数据. 前文已说明, 初步缩并的唯一有用结果是 Ricci 曲率. 根据 $R_{\mu\nu\lambda\kappa} = R_{\lambda\kappa\mu\nu}$, 立即有对称性

$$R_{\mu\nu} = g^{\sigma\lambda} R_{\sigma\mu\lambda\nu} = g^{\sigma\lambda} R_{\lambda\nu\sigma\mu} = R_{\nu\mu}. \quad (1.4.20)$$

这样, 可有进一步的非平凡缩并

$$R := g^{\mu\nu} R_{\mu\nu} = R^\mu{}_\mu. \quad (1.4.21)$$

这称为 **Ricci 曲率标量**, 即 $R_{\mu\nu}$ 的迹. 虽然这不是唯一能从 $g_{\mu\nu}, R^\lambda{}_{\mu\nu\kappa}$ 得到的非平凡标量 (见 2.4 节), 但确实是足够简单的一个.

除了上述纯代数性质之外, Riemann 曲率还满足某些微分恒等式. 最重要的例子是 **Bianchi 恒等式**

$$R^\lambda{}_{\mu\sigma\nu;\kappa} + R^\lambda{}_{\mu\nu\kappa;\sigma} + R^\lambda{}_{\mu\kappa\sigma;\nu} = 0. \quad (1.4.22)$$

直接取局域惯性系证明之. 这立即导致 $R_{\mu\nu}, R$ 间的重要性质: 缩并 λ, σ 再用 $g^{\mu\kappa}$ 升指标, 且注意 $g^{\mu\kappa} R_{\lambda\mu\nu\kappa} = g^{\mu\kappa} R_{\mu\lambda\kappa\nu} = R^\kappa{}_{\lambda\kappa\nu} = R_{\lambda\nu}$, 则

$$\nabla^\mu R_{\mu\nu} + \nabla^\lambda R_{\lambda\nu} - \nabla_\nu R = 0 \implies \nabla^\mu R_{\mu\nu} = \frac{\nabla_\nu R}{2} \implies \nabla^\mu \left(R_{\mu\nu} - \frac{R}{2} g_{\mu\nu} \right) = 0.$$

可以定义

$$G_{\mu\nu} := R_{\mu\nu} - \frac{1}{2} g_{\mu\nu} R \implies \nabla^\mu G_{\mu\nu} = 0. \quad (1.4.23)$$

这称为 **Einstein 张量**, 我们将马上看到命名由来. 这说明 $R_{\mu\nu}$ 的散度是 R 的梯度, 或者 $G_{\mu\nu}$ 无散.

1.4.3 Einstein 方程

其中 D^4x 称为**适配体元**, 这只是考虑到单独的 d^4x 具有变换 $d^4x' = J d^4x$, 因此应仿照协变导数之定义, 规定积分体元只在局域惯性右手系 $\{\xi^\mu\}$ 中表为 $d^4\xi$, 就能得到 $D^4x = \sqrt{|g|} d^4x$.

相应泛函是

$$S_G(g_{\mu\nu}(x^\lambda)) = \frac{1}{2\kappa} \int R \sqrt{-g} d^4x = \frac{1}{2\kappa} \int g^{\mu\nu} R_{\mu\nu} \sqrt{-g} d^4x, \quad (1.4.24)$$

该泛函称为 **Hilbert 作用量**.

现在考虑时空上的度规场, 而根据广义协变性, 可将拉氏密度表为 $\mathcal{L}(g^{\mu\nu}, \partial_\lambda g^{\mu\nu})$ 的形式, 这里为方便选择 $g^{\mu\nu}$ 而非 $g_{\mu\nu}$ (比如可使场方程的指标在下方). 开始对 Hilbert 作用量变分. 类似于极值曲线固定端点, 我们要求在边界上 $\delta g^{\mu\nu} = 0$. 变分得到 3 项:

$$\delta S_G = \frac{1}{2\kappa} \int \left(\sqrt{-g} R_{\mu\nu} \delta g^{\mu\nu} - \frac{R}{2\sqrt{-g}} \delta g + \sqrt{-g} g^{\mu\nu} \delta R_{\mu\nu} \right) d^4x, \quad (1.4.25)$$

首先计算 δg . 根据 Jacobi 公式有

$$\delta g = g g^{\mu\nu} \delta g_{\mu\nu} = -g g_{\mu\nu} \delta g^{\mu\nu}. \quad (1.4.26)$$

其中注意由 $\delta_\nu^\mu = g^{\mu\lambda} g_{\lambda\nu}$ 得 $0 = g^{\mu\lambda} \delta g_{\lambda\nu} + g_{\lambda\nu} \delta g^{\mu\lambda}$. 其次计算 $\delta R_{\mu\nu}$. 依据定义有

$$\begin{aligned} \delta R_{\mu\nu} &= \partial_\lambda (\delta \Gamma_{\mu\nu}^\lambda) - \partial_\nu (\delta \Gamma_{\mu\lambda}^\lambda) + \Gamma_{\mu\nu}^\kappa \delta \Gamma_{\kappa\lambda}^\lambda + \Gamma_{\kappa\lambda}^\lambda \delta \Gamma_{\mu\nu}^\kappa - \Gamma_{\mu\lambda}^\kappa \delta \Gamma_{\kappa\nu}^\lambda - \Gamma_{\kappa\nu}^\lambda \delta \Gamma_{\mu\lambda}^\kappa \\ &= \partial_\lambda (\delta \Gamma_{\mu\nu}^\lambda) + \Gamma_{\kappa\lambda}^\lambda \delta \Gamma_{\mu\nu}^\kappa - \Gamma_{\mu\lambda}^\kappa \delta \Gamma_{\kappa\nu}^\lambda - \Gamma_{\lambda\nu}^\kappa \delta \Gamma_{\mu\kappa}^\lambda - (\partial_\nu (\delta \Gamma_{\mu\lambda}^\lambda) - \Gamma_{\mu\nu}^\kappa \delta \Gamma_{\kappa\lambda}^\lambda). \end{aligned}$$

注意, 虽然克氏符不是张量, 但克氏符变分是张量, 因为作差会消去非协变部分. 再由 $\Gamma_{\lambda\nu}^\kappa = \Gamma_{\nu\lambda}^\kappa$ 得

$$\delta R_{\mu\nu} = \nabla_\lambda (\delta \Gamma_{\mu\nu}^\lambda) - \nabla_\nu (\delta \Gamma_{\mu\lambda}^\lambda), \quad (1.4.27)$$

于是 (1.4.25) 式中 $\delta R_{\mu\nu}$ 的积分为

$$\begin{aligned} \int \sqrt{-g} g^{\mu\nu} \delta R_{\mu\nu} d^4x &= \int \sqrt{-g} g^{\mu\nu} (\nabla_\lambda (\delta \Gamma_{\mu\nu}^\lambda) - \nabla_\nu (\delta \Gamma_{\mu\lambda}^\lambda)) d^4x \\ &= \int (\nabla_\lambda (g^{\mu\nu} \delta \Gamma_{\mu\nu}^\lambda) - \nabla_\nu (g^{\mu\nu} \delta \Gamma_{\mu\lambda}^\lambda)) \sqrt{-g} d^4x, \end{aligned}$$

由高维 Gauss 定理, 可将协变散度的体积分转化为边界项. 若体积分范围为全时空, 则边界就是无穷远处.

也可在作用量中引入适当的边界项, 使得边界项的变分正好与上述边界上的变分相消. 综上有

$$\begin{aligned} \delta S_G &= \frac{1}{2\kappa} \int \left(\sqrt{-g} R_{\mu\nu} \delta g^{\mu\nu} - \frac{R}{2} \sqrt{-g} g_{\mu\nu} \delta g^{\mu\nu} \right) d^4x \\ &= \frac{1}{2\kappa} \int G_{\mu\nu} \delta g^{\mu\nu} \sqrt{-g} d^4x, \end{aligned}$$

边界项

因此 $G_{\mu\nu} = 0$,

一方面

$$\kappa T = R - \frac{1}{2} \delta_\mu^\mu R = -R, \quad (1.4.28)$$

因此

$$R_{\mu\nu} = \kappa T_{\mu\nu} + \frac{1}{2}g_{\mu\nu}(-\kappa T) = \kappa \left(T_{\mu\nu} - \frac{1}{2}g_{\mu\nu}T \right), \quad (1.4.29)$$

这称为场方程的**反迹** (trace-reversed) 表达.

根据 (1.4.29) 式知其等价于 $R_{\mu\nu} = 0$.

令 $T_{\mu\nu} = 0$, 当然也就有 $T = 0$, 根据 (1.4.29) 式立即得这样一种特例:

$$R_{\mu\nu} = 0, \quad (1.4.30)$$

称为真空方程 (*vacuum field equation*). 闵氏度规显然是一个解.

1.4.4 最小耦合

若读者修习过电动力学, 可以察觉到电中性物质在许多方面与电磁场的相似性, 故其往往又与电磁场统称为**物质场**.

假设已知闵氏时空中的某物理定律, 它在任意时空是怎样的形式? 换言之, 引力如何影响或限制其它相互作用?

场与场之间发生的耦合

引力场与相互作用场

在这一方面, Lorentz 不变性和广义不变性 (配合等效原理) 有本质区别, 因为它们所含的对称性具有不同性质. Lorentz 不变性将时空几何限制为闵氏的, 它告诉我们张量的分量可按何种代数组合出现于方程中, 从而限制了物理学中可能发生的所有类型的相互作用, 这称**代数学对称性**. 相比之下, 广义不变性并不限制时空为闵氏, 告诉我们非闵度规 $g_{\mu\nu}$ 怎样出现在方程中, 从而只对引力施加限制, 并且它帮助确定引力与 (自身或其他形式的) 物质间可能的动力学耦合, 但它并不限制其它相互作用, 这称为**动力学对称性**. 两种对称性的基本区别与这样一个事实有关: Lorentz 变换在任意时空点进行相同操作, 因为这时 Jacobi 矩阵元是与时空点无关的常数, 显然任意坐标变换不能满足这一点. 这些对称性的区别可进一步用下面的事实来说明: 能量、动量和角动量守恒定律可从 Lorentz 不变性推导出来, 而广义不变性却不随之有这样的整体守恒定律. 作为替代的是, 广义不变性导致场量中的微分恒等式, 其成立与动力学方程无关, 即无论度规是否满足场方程, 如 Bianchi 恒等式就可由广义不变性的要求直接推导, 从而对任意度规成立.

欲使一个仅具 Lorentz 不变性的定律变为具有广义不变性的, 一个自然的想法是将偏导 ∂_ν 都替换为协变导数 ∇_ν 来构造任意时空的物理定律. 以引力对电磁作用的耦合为例, 在第 2 章中会系统学习如何用二阶张量 $F_{\mu\nu}$ 描述电磁场, 其服从的 Maxwell 方程组可表为

$$\partial_\mu F^{\mu\nu} = 4\pi J^\nu, \quad \partial_\lambda F_{\mu\nu} + \partial_\mu F_{\nu\lambda} + \partial_\nu F_{\lambda\mu} = 0, \quad (1.4.31)$$

其中 J^μ 是 4-电流密度并使用了 Gauss 制使 $1/\epsilon_0 = \mu_0 = 4\pi$. 设置引力场后我们尝试给出

$$\nabla_\mu F^{\mu\nu} = 4\pi J^\nu, \quad \nabla_\lambda F_{\mu\nu} + \nabla_\mu F_{\nu\lambda} + \nabla_\nu F_{\lambda\mu} = 0. \quad (1.4.32)$$

这样引力场 $g_{\mu\nu}$ 通过协变导数所含的克氏符而出现在方程中, 即这正是一个 $g_{\mu\nu}, F_{\mu\nu}$ 的耦合方程. 引力场因此对电磁场产生影响, 可以通过它定量预言光在引力场中的偏折. 遗憾的是, 虽然广义不变性对引力和物质可能的耦合给予了限制, 但它并不能唯一地确定耦合结果. 例如可给 $\nabla_\mu F^{\mu\nu} = 4\pi J^\nu$ 添加一个额外项:

$$\nabla_\mu F^{\mu\nu} + \alpha \nabla_\mu (F^{\mu\nu} R) = 4\pi J^\nu \quad (1.4.33)$$

方程将与广义不变性完全保持一致. 为排除这样的额外项, 通常是求助于最小耦合原理. 这一原理声称, 引力存在时的物质运动方程, 是由引力不存在时的方程 (狭义相对论方程), 通过把 g 换成 g , 把普通导数换成协变导数而得到的; 不需要做其他的改变. 因为在局域测地坐标中, 协变导数化为普通导数, 我们也可以把这一原理表述如下: 在局域测地坐标下, 运动方程就是熟悉的狭义相对论的方程.

假若存在某个可用 $F^{\mu\nu}$ 表示的复杂张量 $H^{\mu\nu}$ 满足恒等式 $\nabla_\mu H^{\mu\nu} = 0$, 则二者相加的方程也符合广义不变性.

协变导数 ∇_ν 在局域惯性系就是其偏导算符 ∂_ν , 因为局域惯性系的克氏符为零. 这迎合了等效原理, 而协变导数本身的定义便迎合广义协变性. 可见, 欲寻找物理定律在弯曲时空中的形式, 可考察相应方程在平直时空惯性系中的表达, 将其中的普通偏导替为协变导数; 此外, g 当然也应替为 η . 这称为替换 (replacement) 法则.

但最小替换原则的约束性并不强, 仍可能给出不唯一的物理方程, 特别是含有 2 阶及以上导数时, 我们在一节再讨论.

(minimal coupling)

引力理论的检验标准当然是科学的基础——实验, 为此自然需要一个关于引力实验本身的理论, 这样就可从实验结果依据实验理论来评判引力理论. Dicke 从 1960s 开始所从事的实验研究使人们对等效原理的理解逐渐深刻, 并意识到应把等效原理摆在引力理论的约束这一地位上. $m_I = m_G$ 可准确地称为弱 (weak) 等效原理, 简称 WEP. 1910 年后 Einstein 希望将力学实验推至一切物理实验, 这就是 Einstein 等效原理, 简称 EEP. 亦有学者称之强 (strong) 等效原理, 简称 SEP. 容易想象, 在局域电梯里打开手电筒, 光相对电梯将按一定频率 $\nu = E/h$ (h 为 Planck 常量) 走直线, 进而在地球上看来光线只能弯曲且频率发生变化. Einstein 便是这样预言了光在引力场中的偏折、红移. 不过, Will [3] 选择区分 WEP 和 EEP. 他将 EEP 表述为:

1. WEP 成立;
2. 局域 Lorentz 不变性: 任何局域非引力实验的结果与观者 4-速无关;

3. **局域位置不变性**: 任何局域非引力实验的结果与其时空点本身无关.

注意局域一词仍指空间点. 之后又在 WEP 的表述中强调**自引力系统**¹与测试粒子一样遵守 WEP, 且将局域 Lorentz 和位置不变性中的非引力实验换为任意实验. 可见这种要求似乎更强, 它不仅考虑系统所处的外引力场, 还考虑系统内物体所激发的自引力场, 即引力的主被动方面皆考虑. 它们对引力理论选择的具体影响如何? 目前任何引力理论都满足 WEP, 因为验证很早就已开始且精度较高, 这件近乎铁打的事实使人很难寻找一个不符 WEP 的理论. 可论证若 EEP 成立, 则引力一定是时空几何效应, 严格满足一个度规理论的基本性质:

1. 时空上能定义度规;
2. 自由的测试粒子的世界线满足该度规下的测地线方程;
3. 局域惯性系中非引力物理定律满足狭义相对论.

前文已从 WEP 窥探到 2 的成立. 又可证 **Schiff 猜想** (conjecture): 凡完备且自洽的满足 WEP 的理论皆必须满足 EEP. 当然, 引力的度规理论是很多的: 我们将马上得到 **Einstein 理论**, 当今主流称呼为**广义相对论** (general relativity); 随后 1.4.6 节将扼要介绍 **Brans-Dicke 理论**. 由于目前已知的度规理论除广义相对论均不满足 SEP, 故有人提出: 若 SEP 成立, 广义相对论可能是唯一选择. 关于这一命题的讨论目前还不够严格, 故只能算猜想. 有挠联络一般只在数学领域或某些涉及量子现象的引力理论中使用, 广义相对论所用的联络默认为 Levi-Civita 联络, 即研究**无挠时空**². 目前学术界对等效原理的分类未达成共识, Ohanian [4] 就批评用自引力系统强行区分 WEP 和 EEP. 对此无法说得更通俗了, 只有在我们学会后续知识后才能有更深理解. 上述所有论证细节和具体实验均放入第 ?? 章.

1.4.5 线性近似

规范

动力学最基本的问题便是平凡解的稳定性. 如果对初始数据作微小变化, 则对解造成的变化是否也微小? 对应地, 闵氏时空的稳定性问题的解决, 提出了所谓引力辐射定律. 迄今还未直接观察到引力辐射, 这最初是由 Hulse 和 Taylor 从双星能量损失推导出来的. 这项工作使他们得到 Nobel 奖中首次与场方程相关的奖项 (1993). Penrose 和 Bondi 对辐射做了数学工作: 对时空附加一个“在无穷远处”的理想边缘, 称为类光无穷远 \mathcal{I}^+ , 其中的点对应于很远的观者, 处在一个孤立的自引力系统里, 可以接受信号. 便可以定义某些张量为辐射.

如果把分量全部显式地展开, 就可以看到它是关于 $g_{\mu\nu}$ 的非线性微分方程.

¹ 自引力系统 (self-gravitating system) 指各部分由引力聚集或能产生引力场的系统, 大到恒星、行星, 小到日常所见的物体皆于考虑之列.

² 这例与等效原理联系不大. 虽然式 (1.1.41) 给出了无挠性 (且还是 Levi-Civita 联络), 但即使有挠, 我们仍可有等效原理, 因为 $A^\mu = 0$ 中有 4-速对称项 $\Gamma_{\mu\nu}^\lambda U^\mu U^\nu$. 本章选择无挠是出于方便.

如果在任意坐标系下把场方程显式地写出来, 就可以看到, 它并不属于哪一种通常的类型的 PDE, 诸如椭圆型方程 (像 Poisson 方程那样)、抛物型方程 (像热传导方程那样) 和双曲型方程 (像波动方程那样). 这与以下的事实有关, 即给出了一个解, 就可以把这个解与一个坐标变换复合起来, 得到一个新解. 我们可以对原来的坐标系作一个如下的坐标变换以得出一个新坐标系: 这个变换只在一个球体内与恒等变换不同. 这个事实, 即所谓洞穴论据¹, 使得 Einstein 和 Grossmann 大为困惑, 他们是用坐标下的方程“代数地”在思考, 因而竟使得他们拒绝广义协变性. 这样导致的思维滞后使得场方程的正确陈述推迟了两年之久. 对于 Einstein 的理论的几何解释给出了摆脱困境的出路: 这样的解应该看成是“相同”的解, 因为从所有的几何量度的观点看来, 它们都是相同的. 用现代的语言来说, 我们就说场方程的解 (例如真空方程的解) 就是时空的等价类: 我们说两个时空是等价的, 就是说它们之间有一个微分同胚 ϕ , 使得在任意开集合里, 如果我们把用 ϕ 互相转换的局部坐标视为相同的, 则这两个度规都看成是相同的. 结果, 当这些概念上的问题被克服以后, 场方程就可以看成是双曲型的.

看到这一点最容易的方法是作一个**规范**, 就是对坐标系作一定的限制. 具体地说, 就是要求坐标函数 x^α 满足波动方程 $\square_g x^\alpha = 0$, 这里 \square_g 是 d'Alembert 算子. 这样的坐标总是局部存在的, 称为谐和坐标, 或者波坐标. 于是场方程可以写为一个方程组

$$\square_g g_{\mu\nu} = N_{\mu\nu}(\{g_{\alpha\beta}\}, \{\partial_\lambda g_{\alpha\beta}\}),$$

这里的 $N_{\mu\nu}$ 是一个非线性表达式, 而且对于 $\{\partial_\lambda g_{\alpha\beta}\}$ 是二次式. 考虑到 Lorentz 度量的号差, 上面的方程组就是一个二阶非线性²双曲方程组. 于是方程组的波动性质就是显然的了³.

场方程的双曲性有两个重要收获. 第一即是应该存在引力波, Einstein 本人在 1918 年就看到了这一点; 第二便是适定的初值问题. 所谓适定性, 就是对于适当的初始数据, 必有唯一的解. 这里的所谓数据应该是一个 3 维 Riemann 流形 (Σ, \bar{g}) 再附带一个 2 阶对称协变张量 K . 从而 Einstein 理论就变成了经典问题, 即一个动力学问题.

1.4.6 Mach 原理

惯性起源

¹ 又称洞论证, 这是 Einstein 在发展广义相对论时提出的关于时空的一个深刻的论点, 或者说是一个悖论. 它有深刻的哲学含义, 所以下面介绍一个哲学网站上的一篇文章, 以备有兴趣的读者参阅: <http://plato.stanford.edu/entries/spacetime-holearg/>

² 实际上是拟线性.

³ 后期的 Yang-Mills 方程也具有这样的规范对称性.

Einstein 也曾试图借助于 Mach 原理, 来证明广义不变性的合理性. Mach 原理表述了一个饶有兴趣但有些含糊的看法, 即物体的惯性受到宇宙物质分布的某种影响. 虽然 Einstein 的理论确实显示出惯性对于物质分布的某种依赖关系, 但 Mach 原理在 Einstein 理论中实际上只起到一种边缘的作用.

等效一词令人十分迷惑, 意味是完全等同, 亦或是忽略了“等号”的方向.

等效原理还可有两个方向的理解. 一种是将引力场的自由落体系局域等效为无引力场的惯性系, 一种是将无引力场的加速系局域等效为引力场的惯性系. 比如, 可“颠倒”Einstein 电梯实验的过程, 设无引力场中的有一架朝乘客头顶方向做加速运动的飞船, 则乘客应感受到惯性力, 根据 $m_I = m_G$, 该惯性力场局域模拟出引力场. 再比如, 还可将飞船设置成圆盘状, 用离心效应模拟引力.

一个著名的度规理论叫 **Brans-Dicke 理论**, 其中描述引力的量除度规外还有标量场 ϕ . 这个理论能够满足 Mach 原理, 比较戏剧的是, 尽管 Einstein 最初借助 Mach 原理奠定了引力场的地位, 但广义相对论却并不能完全满足 Mach 原理.

最后, 在小尺度的微观下, 引力强度相较于静电力弱得多: 两个质子之间的静电斥力比它们之间的引力大 1.2×10^{36} 倍 (默认任意间隔上两种力都遵从平方反比律). 这在一定程度上暗示了经典引力理论在量子领域可能的失败. 现在我们知道, 自 Einstein 开始, 至今仍未完全地调和广义相对论同量子理论的各种矛盾. 由于深入讲解这个话题必将需要读者了解相关的陌生概念, 故我将不在此赘述, 而只简单地指出: 引力同经典电磁作用的统一理论实际上已经完成, 其中 Kaluza-Klein 理论和 Einstein 的 (经典) 统一场论¹是完全经典的, 而其它的诸如弦论、非交换几何 (CLCB 模型) 只部分地回答了量子引力的疑难. 现在所谓的“大一统理论”通常指的是沿着标准模型的思路, 将强相互作用同电弱作用统一的尝试. 狭义上的量子引力理论只涉足引力的量子化问题, 但这在数学上有巨大的困难; 广义的量子引力理论或许是人们所希望的“万物理论”. 此外, 受黑洞热力学的启发, 现在也有人企图寻找引力同统计物理、信息论的联系, 诸如熵力理论、黑洞信息学. 这些话题将部分地放入本书后半部分中去.

量子引力

1.4.7 洞悖论

洞悖论 (hole argument) 和**点重合** (point coincidence) 原理.

1.5 单位制转换

且使用了**国际单位制**, 简称**国际制** (SI)

¹受到 Kaluza 理论的启发的 Teleparallel 引力, 其数学语言采用 Weitzenböck 几何.

物理学是关乎实验测量的学科, 讨论物理量的单位是极为重要的. 在国际制中, 我们认为时间、空间乃不同物理量, 自然不会谈及“1s 等于多少 m”这种看似奇怪的问题. 但在相对论中, 这个问题很有意义, 因为时空乃统一的概念. 考虑令 $c = 1$ 使 $x^0 = t$, 国际制下 c 近似为 3×10^8 m/s, 故这件事相当于

$$1 \text{ s} = 3 \times 10^8 \text{ m}.$$

这不会与国际制有任何冲突, 因为在国际制中不会谈及“1s 等于多少 m”这种问题. 这是国际制中一个可以利用的自由度. 在以后学习的广义相对论中, 还经常涉及引力常量, 把它们的数值都取为 1, 即

$$c = G = 1, \quad (1.5.1)$$

可简化大量公式的书写, 这就是相对论所使用的**几何单位制** (system of geometrized units), 简称**几何制** (SG). 在国际制中 G 近似等于 $6.67 \times 10^{-11} \text{ m}^3/(\text{s}^2 \cdot \text{kg})$, 因此该操作相当于

$$1 \text{ kg} = \frac{6.67 \times 10^{-11}}{(3 \times 10^8)^2} \text{ m},$$

可见这利用的是国际制的另一自由度. 几何制便统一了长度、质量和时间的单位. 不涉及引力的量子理论经常使用如下的**自然单位制** (system of natural units), 简称**自然制** (SN):

$$c = \hbar = 1, \quad (1.5.2)$$

这里 \hbar 在国际制中近似等于 $6.58 \times 10^{-22} \text{ MeV} \cdot \text{s}$. 然而在计算物理量的数值时几何单位就不再方便, 因此我们要研究物理量、物理公式如何从几何制转换到其它单位制. 一般来讲, 所研究物理量在当前单位制是带上单位的. 比如, 几何制下粒子物理常用 eV 作为质量单位, 因为质能方程在几何制下统一了质量和能量单位. 欲将质量单位从几何制还原为国际制, 无非是在问 eV 等于多少 kg. 首先都转化到常用单位来, 即

$$1 \text{ eV} = 1.6 \times 10^{-19} \text{ C} \cdot \text{V} = 1.6 \times 10^{-19} \text{ J} = 1.6 \times 10^{-19} \text{ kg} \cdot \frac{\text{m}^2}{\text{s}^2},$$

代入几何制定义有

$$1 \text{ eV} = \frac{1.6 \times 10^{-19}}{(3 \times 10^8)^2} \text{ kg}.$$

然而, 若所研究物理量在当前单位制只是一个数字, 则其通常只是在当前单位制下无单位, 但在其它单位制中有单位, 因此严格来说不能称其无量纲. 真正无量纲的量在任何单位制下都无单位, 因此不受单位转换的影响, 如精细结构常数

$$\alpha = \frac{e^2}{4\pi\epsilon_0\hbar c} = \frac{1}{137},$$

代入自然制定义得 $e^2/4\pi\epsilon_0 = 1/137$, 这说明兼容自然制和精细结构常数的同时仍有自由度可利用. 令

$$c = \hbar = \epsilon_0 = 1, \quad (1.5.3)$$

因而 $e = \sqrt{4\pi/137}$. 由 $c = 1/\sqrt{\epsilon_0\mu_0}$ 得 $\mu_0 = 1$. 如上规定就称为 **Heaviside-Lorentz 单位制**, 简称 **HL 制** (SHL), 其下不用再写出 Maxwell 方程中的常数. 几何制和自然制还可共同构成 **Planck 制** (SP), 这在量子引力理论中用得更多. 最后, 对一个整体的物理公式而言, 最便捷的做法不是如上这样分析单位, 而是分析量纲, 即根据量纲补充缺失的常数. 另外, 有必要说明这样一种思想: 有别于传统认知, 物理公式应当视作“数”的公式而非“量”的公式. 具体细节放入附录 A 中. 简单来说, 正如定义 $x^0 = ct$ 一样, 定义新单位制下某个物理量为旧单位制中的对应物理量乘以相应常数, 这样物理方程中就再也不会出现该常数, 因为它全部被收进了新定义里, 比如, 将 G 收进场源质量里, 并不影响其它任何物理量的计算; 重新定义 $t' = x^0 = ct$, 其单位同距离单位彻底一致, 时间和距离“共享”同一单位, 因而 $c = 1$ 是真正的“数”. 不过从单位转换的角度看, 这些数据并不需要收进任何一个物理量中, 而是直接收进单位的相对关系里.

Gauss 制

μ_0 称为**真空磁导率** (vacuum permeability), ϵ_0 称为**真空介电常数** (vacuum permittivity).

几何 Gauss 制

1.6 Lorentz 群

就方阵 $\Lambda^\mu{}_\nu$ 而言可理解为左边的 μ 是行指标、右边的 ν 是列指标, 把 x^μ 看成某个列向量的分量, 则这就是对列向量左乘矩阵得到新的列向量, 矩阵乘法及顺序体现在这两个实数的相乘和列指标 ν 的“相靠”

前者容易混淆矩阵的转置和逆

由式 (??) 可得逆变换

$$x^\lambda = (\Lambda^{-1})^\lambda{}_\mu x'^\mu + x_0'^\lambda, \quad x_0'^\lambda = -(\Lambda^{-1})^\lambda{}_\mu x_0^\mu. \quad (1.6.1)$$

欲用 Λ 表示 Λ^{-1} , 需借助度规逆. 容易证明度规逆分量仍排成对称矩阵, 故转置不变. 相同惯性坐标系下, 易知 $\eta_{\mu\nu}$ 在数值上等于 $\eta^{\mu\nu}$. 由式 (1.1.17) 有

$$\eta_{\mu\nu}(\Lambda^{-1})^\mu{}_\sigma = \eta_{\sigma\lambda}\Lambda^\lambda{}_\nu, \quad (1.6.2)$$

进而

$$(\Lambda^{-1})^\lambda{}_\sigma = \eta_{\mu\sigma}\Lambda^\mu{}_\nu\eta^{\nu\lambda}. \quad (1.6.3)$$

此即用 Λ 表示 Λ^{-1} 的方法, 取其 $\sigma = \lambda = 0$ 有

$$(\Lambda^{-1})^0_0 = \Lambda^0_0 = (\Lambda^\top)^0_0, \quad (1.6.4)$$

可见 $(\Lambda^{-1})^0_0$ 与 Λ^0_0 一致, 当然按符号讨论的定义也就一致; 由 $|\Lambda^{-1}| = |\Lambda|^{-1}$ 知定义也一致.

下面尝试用两惯性系的相对速度表示 Lorentz 变换. 已证明变换仿射性, 故 Lorentz 变换总满足

$$dx'^\mu = \Lambda^\mu_\nu dx^\nu, \quad dx^\lambda = (\Lambda^{-1})^\lambda_\mu dx'^\mu. \quad (1.6.5)$$

任意世界线上某事件处在 S 系下测得的 3-速是

$$u^i := \frac{dx^i}{dt}, \quad (1.6.6)$$

同理, 任意世界线上某事件处在 S' 系下测得的 3-速分量为

$$u'^i := \frac{dx'^i}{dt'}. \quad (1.6.7)$$

我们说 S' 系相对于 S 的 3-速是 u^i , 是指 S' 任意的相对静止的世界线在 S 下测得的 3-速分量都是 u^i , 容易证明这等价于研究 S' 系的空间原点运动. 反之, 设 S 系的空间原点在 S' 下测得 u'^i . 先考虑 S' 的空间原点, 其世界线始终满足 $dx'^i = 0$, 则

$$dx^\lambda = (\Lambda^{-1})^\lambda_0 dx'^0 T_0 \Lambda^0_0 = (\Lambda^{-1})^0_0 = \frac{dt}{dt'}, \quad (\Lambda^{-1})^i_0 = \frac{u^i}{c} \frac{dt}{dt'},$$

据式 (1.6.2) 有

$$\eta_{\mu\nu} (\Lambda^{-1})^\mu_\sigma (\Lambda^{-1})^\nu_\lambda = \eta_{\sigma\lambda}, \quad (1.6.8)$$

取 $\sigma = \lambda = 0$ 有

$$(\Lambda^0_0)^2 = 1 + \delta_{ij} (\Lambda^{-1})^i_0 (\Lambda^{-1})^j_0. \quad (1.6.9)$$

再结合正时条件有

$$\Lambda^0_0 = (\Lambda^{-1})^0_0 = \text{gamma}, \quad (\Lambda^{-1})^i_0 = \text{gamma} u^i / c, \quad (1.6.10)$$

其中

$$\text{gamma} := \frac{1}{\sqrt{1 - u^2/c^2}}, \quad u := \sqrt{u^i u_i}, \quad u_i := \delta_{ij} u^j. \quad (1.6.11)$$

这里定义 u_i 以配合指标平衡, 但其在数值上与 u^i 一致; . 再考虑 S 的空间原点, 类似地有

$$\Lambda^0_0 = \frac{dt'}{dt}, \quad \Lambda^i_0 = \frac{u'^i}{c} \frac{dt'}{dt} = \text{gamma} u'^i / c, \quad (1.6.12)$$

以及

$$u'^i u'_i = u^i u_i, \quad u'_i = \delta_{ij} u'^j. \quad (1.6.13)$$

这看似是显然的, 但现在要作为等度规变换的已证结论. 然而这是用 u'^i 表示从 $\{x^\mu\}$ 到 $\{x'^\mu\}$ 的变换 $\Lambda^\mu{}_\nu$, 固然渴望用 u^i 表示, 这更为直接. 对式 (1.6.3) 其余 σ, λ 值讨论有

$$(\Lambda^{-1})^i{}_0 = -\Lambda^0{}_j \delta^{ij} = -\Lambda_0^i, \quad (\Lambda^{-1})^0{}_i = -\Lambda^j{}_0 \delta_{ij} = -\Lambda_i^0, \quad (1.6.14)$$

以及

$$(\Lambda^{-1})^i{}_j = (\Lambda^\top)^i{}_j. \quad (1.6.15)$$

可得

$$\Lambda_0^i = (\Lambda^\top)^i{}_0 = -\text{gamma} u^i / c T o \Lambda^0{}_i = -\text{gamma} u_i / c. \quad (1.6.16)$$

还差 $\Lambda^i{}_0, \Lambda^i{}_j$. 尝试用 u^i 表示 u'^i . 取 (1.1.17) 式 $\sigma = 0, \lambda = k$ (反之亦可), 有

$$\delta_{ij} \Lambda^i{}_0 \Lambda^j{}_k = \Lambda^0{}_0 \Lambda^0{}_k T o \Lambda^j{}_k u'_j = -\text{gamma} u_k. \quad (1.6.17)$$

对逆度规亦容易证明

$$\eta^{\mu\nu} \Lambda^\sigma{}_\mu \Lambda^\lambda{}_\nu = \eta^{\sigma\lambda}, \quad (1.6.18)$$

同理取 $\sigma = 0, \lambda = k$ 有

$$\Lambda^k{}_j u^j = -\text{gamma} u'^k. \quad (1.6.19)$$

现在还差 $\Lambda^i{}_j$. 对此更直观表示方法是利用空间旋转矩阵 R . 若两个坐标系仅差一个旋转变换, 则二者仍是相对静止的, 相对论的动力学效应不会体现. boost 变换实质就是指这类允许有任意方向相对速度的 Lorentz 变换, 但保持空间坐标轴相互平行. 此前我们只给了保持 $y' = y, z' = z$ 的特殊情形. boost 变换是“时空旋转”, 故任意 Lorentz 变换应可表为一个 boost 变换和一个空间旋转的复合 (因此不一定有 $u^i = -u'^i$). 为找到 R , 注意总可假设

$$R^i{}_j =: \Lambda^i{}_j + S^i{}_j,$$

其中 $\Lambda^i{}_j$ 符合取正是为符合固有条件. 且满足

$$\delta_{ik} R^i{}_j R^k{}_l = \delta_{jl}, \quad R^i{}_j u^j = -u'^i, \quad R^i{}_j u'_i = -u_j. \quad (1.6.20)$$

注意取 (1.1.17) 式 $\sigma = j, \lambda = l$ 有

$$\delta_{ik} \Lambda^i{}_j \Lambda^k{}_l = \delta_{jl} + \frac{\text{gamma} a^2}{c^2} u_j u_l, \quad (1.6.21)$$

结合式 (1.6.17)、(1.6.19) 以及上式, 经过一些繁琐计算后可得

$$S^i_j = \frac{\gamma - 1}{u^2} u'^i u_j, \quad R^i_j = \Lambda^i_j + \frac{\gamma - 1}{u^2} u'^i u_j, \quad (1.6.22)$$

最终得

$$\Lambda^i_0 = -\gamma R^i_j u^j / c, \quad \Lambda^i_j = R^i_j + \frac{\gamma - 1}{u^2} R^i_k u^k u_j. \quad (1.6.23)$$

此即任意 Lorentz 变换. 取 $R^i_j = \delta^i_j$ 就可得任意 boost 变换

$$\Lambda^0_0 = \gamma, \quad \Lambda^0_i = -\gamma u u_i / c, \quad \Lambda^i_0 = -\gamma u u^i / c, \quad \Lambda^i_j = \delta^i_j + \frac{\gamma - 1}{u^2} u^i u_j. \quad (1.6.24)$$

再令 $u^2 = u^3 = 0$ 就有

$$\Lambda^1_1 = \Lambda^0_0 = \gamma, \quad \Lambda^0_1 = \Lambda^1_0 = -\gamma u u^1 / c, \quad (1.6.25)$$

这是我们所熟知的只研究 t, x 时的情形, Λ 对另二空间坐标 y, z 而言是恒等变换 I , 故实质上成为二阶方阵.

Chapter 2

现代框架

2.1 拓扑学

即使更换背景，比如 \mathbb{R}^4 改为某种 4 维圆锥面，以母线为时间轴就仍可类似地定义闵氏度规从而使其等价于无引力场。对于一般时空，由于背景不能提供计算信息，可见坐标无关的时空物理量应与背景无关，故此物理量是内禀的，只能尝试从度规构造。

当然，背景概念在理论构建之初无需过多考虑，然而在后期发展中，人们逐渐意识到，物理现象确实与时空背景 M 有关，比如我们的时空背景很可能是方方正正的 \mathbb{R}^4 ，也可能是某个高维圆柱面或球面。

C^r 流形给拓扑对象提供微分结构，用以导出切映射并讨论场的微分性质。

大范围系统的等效原理

不论其质量和组分如何

整体拓扑性质 (global topology)²

流形涉及开覆盖概念，而给出一般集合中的开子集的学科称为拓扑学。选择从集合论知识出发构建拓扑学基本概念，称**点集拓扑**，其它方式不实用。但为讨论更为深刻的话题，还需要借助所谓**代数拓扑**。暂不涉及，尽管它在本科数学系里作为高级课程以引入。代数拓扑这一专题将在关于**上同调** (cohomology) 的部分中得以处理，在讨论拓扑分类、积分理论时 (de Rham 上同调) 比较好用。

当然，在拓扑学框架中还只能定义所谓的拓扑流形，即每一个点附近有一个邻域同胚于 \mathbb{R}^n 上开集的拓扑空间，而不能定义我们最终要用到的微分流形——那是

²拓扑学常被形象地描述成橡皮膜 (rubber-sheet) 几何。自然界存在直觉上形状相似的几何体，为抽象出共性，可试着形变某一几何体但不撕裂，观察是否能“揉”成另一几何体，如从正方体到球体。这种等价性称作同胚 (homeomorphism)，进而可将各种几何体分门别类，即把注意力集中于诸如“是否连通”“曲面是否有方向”“是否有洞”“绳子上打了几个结”这类问题上的，从而架空距离、长度、面积这些传统几何学中的概念。

更进一步的工作. 事实上, 我们讨论物理问题时一般不会特别深入地探讨所研究空间的拓扑结构, 所以仅就我们的需要而言, 最基本的点集拓扑学知识已经足够.

拓扑学除其“橡皮膜”特色外, 最有价值的作用就是叙述连续性. 拓扑是关于连续性的最重要的概念. 连续性是我们在这里所要研究的绝大部分问题的基础. 研究映射连续性以及拓扑结构在连续变换下不变的性质的学科. 点集拓扑研究的是最一般的集合上的拓扑, 讨论开集、闭集、连续映射、同胚这些基本的拓扑学概念, 是所有拓扑学分支的基础. 后续的微分流形是一个可供微积分的平台, 这在很大程度上需要借助拓扑学定义, 也就是用拓扑结构提供平台的连续性.

相信这样基础性的一节的作用纯粹在于夯实和查阅. 许多物理学家对这里绝大部分概念是熟悉的, 但通常在相当不精确的意义下. 本节将提供严格的定义和结果, 其在参阅其余内容时会很有用. 从应用于物理的角度出发, 在假定对于未加定义的更高级概念熟悉的情况下, 我们将引用大量的例子. 论题的选择也是基于同样的理由.

拓扑的基础作用就是在一般的集合上定义开集, 而不借助任何距离概念. 为此可先回顾一下分析学中的开闭概念. 在定义 \mathbb{R} 时已经借助了序结构, 因此 $>, <, \geq, \leq$ 之类的关系是良定义的. \mathbb{R} 上的任意开区间可表为若干“连通开集”之并, 所谓连通开集就是形如 (a, b) 的集合. 闭区间就是其在 \mathbb{R} 上的补集, 则边界上的括号就改为中括号, 表示可取等.

欲在一般情况下定义开集, 当然要研究一个集合的全体开集组成的集合有何性质, 这个集合就叫做拓扑. 其性质就是开集所具备的特征, 我们可从分析学中抽象出来.

比如, 全集可以看作开的, 则空集就是闭的, 但空集也可以看作开的, 则全集就是闭的. 从定义的推广上来看, 其实没有理由排除一个既开又闭, 或非开非闭的集合. 不应该仅靠直觉就否定这种存在. 不过, 考虑到拓扑是先定义开集, 不妨先规定空集、全集是开的. 在分析学中, 除了空集、全集外, 一个区间就不会存在这种兼备的现象了.

其次, 任意多甚至无限个开集的并集是开集, 这从分析学的定义就可看出. 但是, 要使得开集的交集是开集, 就只能有限多个: 在 \mathbb{R} 上, 所有开区间 $(-1/n, 1), n \in \mathbb{N}_+$ 的交集是 $[0, 1)$, 显然不是开的. 区分“任意并”、“有限交”很重要.

上述三条(足够, 不再需要更多)性质不显含任何有关集合上距离的信息, 因此可作推广之基础. 因此, 拓扑学所研究的对象可以看作脱胎于距离的一种更广义的空间. 把所有仅涉及到开集的定义和命题给抽提出来, 就得到了拓扑学中相应的定义和命题. 可见, 拓扑学并不玄乎.

定义 2.1. 设 X 为一集合. X 的一族子集 \mathcal{T} 称为一个**拓扑** (topology), 如果以下性质成立:

$$a. \emptyset, X \in \mathcal{T};$$

b. 开集的任意并是开集. 令 \mathcal{I} 为一指标集, 应有

$$A_\alpha \in \mathcal{T}, \alpha \in \mathcal{I} \Rightarrow \bigcup_{\alpha \in \mathcal{I}} A_\alpha \in \mathcal{T}.$$

也可以不借助指标集: $\forall \sigma \subseteq \mathcal{T}, \bigcup_{U \in \sigma} U \in \mathcal{T}$;

c. 开集的有限交是开集. 即

$$A_i \in \mathcal{T}, i = 1, \dots, n \Rightarrow \bigcap_{i=1}^n A_i \in \mathcal{T}.$$

X 的一个元素称为 X 在拓扑 \mathcal{T} 下的一个**开集** (open set).

例 2.1. 实数集的**通常拓扑** (usual topology) 暂且理解为分析学中开闭定义所对应的拓扑.

定义 2.2. 如果假定 \emptyset 和 X 是仅有的开子集, 则得到所谓的**非离散拓扑** (indiscrete topology) 或**最粗拓扑** (coarsest topology). 另一个极端情形是假定所有子集都是开集, 则我们得到**离散拓扑** (discrete topology) 或**最细拓扑** (finest topology).

若以实直线 \mathbb{R} 或平面 \mathbb{R}^2 为例来考虑, 这两个拓扑都不太自然. 但事实上如果我们从定义中明确地排除它们则会更不自然. 在某些方面它们显得非常有用.

定义 2.3. X 的一个子集称为**闭的** (closed), 若它在 X 中的补集是开的.

基于并集之补等于补集之交这一集合论的命题, 可以很容易地用闭集代替开集来构造一个拓扑. 因此, 也可以把拓扑看作定义闭集的过程.

定义 2.4. 一个被赋予了拓扑 \mathcal{T} 的集合 X 称为一个 (关于给定拓扑的) **拓扑空间** (topological space), 记作 (X, \mathcal{T}) . 拓扑空间中的元素有时被称为点 (point), 拓扑就称其拓扑结构. 当然, 即使没有拓扑结构的集合, 也可以把元素称作点.

此外, 还有一些看似需借助距离的分析学概念可用拓扑语言改写.

定义 2.5. 令 $x \in X$. x 的一个**邻域** (neighbourhood) 是 X 的一个子集, 该子集包含一个含有 x 的开集.

这个定义看起来有些不方便, 但实际上在一般情况下比通常那种仅限于开邻域的定义要有用. 直观上, 邻域只需要“包含 x ”即可, 但由于“大小”可任意, 可“无限接近于” x , 因此称“邻”以表附近之义.

对于一般的拓扑空间, 其开集结构往往会比较复杂, 这时我们会用所谓“拓扑基”方式来描述一个拓扑结构. 对于一个集合 X , 所谓拓扑基, 是指 X 的一个满足“有限交”性质并且能够“覆盖”整个 X 的子集族, 这个子集族通过做“任意并”操作能够生成 X 的一个拓扑结构.

定义 2.6. 一个开集的子族 $\mathcal{B} \subseteq \mathcal{T}$ 称作拓扑 \mathcal{T} 的一个**基** (*basis*), 如果每个开集都是 \mathcal{B} 中集合的并集. 也不妨考虑: $\bigcup_{U \in \mathcal{B}} U = X$; 对任意 $U_i \in \mathcal{B}$ 存在 $\mathcal{F} \subseteq \mathcal{B}$ 使得 $\bigcup_{U \in \mathcal{F}} U = \bigcap_{i=1}^n U_i$. \mathcal{B} 中成员的所有可能并集 $\mathcal{T} = \{\bigcup_{U \in \mathcal{F}} U\}$ 构成了 X 上的一个拓扑, 称为由拓扑基 \mathcal{B} 生成的拓扑.

定义 2.7. 一个开集的子族 $\mathcal{S} \subseteq \mathcal{T}$ 称作拓扑 \mathcal{T} 的一个**子基** (*sub-basis*), 如果每个开集都是 \mathcal{S} 中集合的有限交集之并.

定义 2.8. X 的子集 A 的**闭包** (*closure*) \bar{A} 是 X 中包含 A 的最小闭集;

X 的子集 A 的**内部** (*interior*) $i(A)$ 是 X 中含于 A 的最大开集;

有时可定义 $\partial A = \bar{A} \setminus i(A)$ 为 A 的**边界** (*boundary*).

定义 2.9. 设 A 为拓扑空间 X 的子集. 点 $x \in X$ 称为 A 的**极限点**, 如果每一个包含 x 的开集都包含 A 中异于 x 的点.

定义 2.10. 称 X 的子集 A 在 X 中**稠密** (*dense*), 如果 $\bar{A} = X$.

定义 2.11. 拓扑空间 X 称为 T_2 空间或 **Hausdorff 空间**, 如果对任意不同的两点 $x, y \in X$, 存在 x 的开邻域 A 与 y 的开邻域 B , 使得 $A \cap B = \emptyset$.

这看起来更像我们所期待的. 但是, 某些略微非 Hausdorff 的空间会很有用.

例 2.2. 在扭量 (*twistor*) 理论中. 一个“口袋” (*pocket*) 提供了这样的一个例子. 考虑实平面的子集 X , 由实轴上的区间 $[-1, 1]$ 和直线 $y = 1$ 上的区间 $[0, 1]$ 构成, 并且等同如下的点: $(x, 0) \approx (x, 1), 0 < x \leq 1$. 这样, 点 $(0, 0)$ 与 $(0, 1)$ 就没有不相交的开邻域. 严格说来, 我们需要下文中间拓扑的概念.

例 2.3. 一个更为地道的非 Hausdorff 空间: 考虑正整数组成的空间 \mathbb{N}_+ , 开集取成 \emptyset, \mathbb{N}_+ 以及集合 $\{1, 2, 3, \dots, n\}$. 这个空间既非 Hausdorff 亦非紧致的 (见后面关于紧致的定义).

定义 2.12. 设 X, Y 为拓扑空间. 我们称映射 $f: X \rightarrow Y$ 是**连续的** (*continuous*), 如果“开集的原像是开集”, 即对于 Y 中任意开集 A , $f^{-1}[A]$ 是 X 中的开集.

连续性是这里最为重要的概念. 该拓扑学定义也不显含任何关于距离的信息. 在这个一般的框架下, 它看起来与 ε - δ 定义略有不同. 但是后者只对于定义了距离的距离空间有效, 我们马上将要谈到.

定义 2.13. 一个双射 $f: X \rightarrow Y$ 称作**同胚** (*homeomorphism*), 如果它是连续的且逆映射 f^{-1} 也是连续的. 两个同胚的空间从拓扑学的观点来看是不可区分的. 在这个意义下, 同胚是拓扑空间之间自然的映射. 拓扑不变量就是拓扑空间在同胚下保持不变的性质.

定义 2.14. 设 $B \subseteq A$, 可定义 B 的**相对拓扑** (*relative topology*) 或**诱导拓扑** (*induced topology*): 一个子集 $C \subseteq B$ 是开集当且仅当存在 A 的开集 D 使得 $C = D \cap B$. 一个赋予了诱导拓扑的子集 $B \subseteq A$ 称为 A 的**拓扑子空间** (*topological subspace*).

这样, 对于实直线的子集, 如果我们令 $A = [0, 3], B = [0, 2]$, 则在由 \mathbf{R} 上通常拓扑的诱导拓扑下, $C = (1, 2]$ 是 B 中的开集.

定义 2.15. 给定两拓扑空间 X, Y , 可定义**乘积拓扑空间** (*product topological space*) $Z = X \times Y$. 这里 Z 作为集合是 X 和 Y 的卡氏积. 令 A 为 X 中开集, B 为 Y 中开集, 则形如 $A \times B$ 的集合构成 Z 的一个拓扑基.

当然, $X \times Y$ 中的开集并不一定都具有乘积 ($A \times B$) 的形式.

定义 2.16. 设对于指标集 \mathcal{I} , 有一个将 X 分成互不相交的子集 $A_\alpha, \alpha \in \mathcal{I}$ 的分解. 或者等价地说, 在 X 上定义了一个等价关系. 则我们可以按以下方式在等价类的集合 $\{A_\alpha, \alpha \in \mathcal{I}\}$ 上定义商拓扑 (*quotient topology*), 该等价类称作**商空间** (*quotient space*) $X/\sim = Y$. 考虑正则投影映射 $\pi: X \rightarrow Y$, 它把 X 中的一个元素 x 映到它的等价类 $[x]$. 则 Y 中的子集 $U \subseteq Y$ 是开集当且仅当 $\pi^{-1}(U)$ 是开集.

定理 2.1.1. 令 \mathcal{T} 为商空间 Y 上的商拓扑. 设 \mathcal{T}' 为 Y 上另一个拓扑使得正则投影是连续映射, 则 $\mathcal{T}' \subseteq \mathcal{T}$.

定义 2.17. X 的一个**开覆盖** (*open cover*) $\{U_\alpha: \alpha \in \mathcal{I}\}$ 是指 X 的一族开子集 $U_\alpha \subseteq X$, 它们的并集是 X . 一个开覆盖的子覆盖是指这族集合的一个子族, 它本身也是 X 的一个开覆盖.

定义 2.18. 一个拓扑空间 X 称为**紧致的** (*compact*) 或**紧的**, 如果 X 的每个开覆盖都有有限的子覆盖.

对于紧致空间来说, 无论我们选择怎样的开覆盖, 用有限个开子集就足够了. 这是开区间 (非紧致) 和闭区间 (紧致) 的根本区别. 前者实际上同胚于整个实直线.

定义 2.19. 仿紧

定义 2.20. 第二可数

给出有用的拓扑结构, 使点足够稠密但又不至于太怪异, 称为几何背景.

定义 2.21. 一个拓扑空间 X 称为**连通的** (*connected*), 如果它不能写成两个互不相交的非空开集的并. 有用的两条等价说法是: 任何一个从 X 到赋予离散拓扑的两点集合 $\{0, 1\}$ 的连续映射都不是满射; 连通拓扑空间只有两个既开又闭的子集.

定义 2.22. 一个拓扑空间 X 称为**路径连通的** (*path-connected*) 或**弧连通的** (*arcwise connected*), 如果 X 中的任意两点都可被一条完全在 X 中的路径连接.

定理 2.1.2. 路径连通的拓扑空间一定是连通的. 反之不一定, 存在“擦边”的反例¹. \mathbf{R}^n 中的连通开子空间是路径连通的.

定义 2.23. 给定一拓扑空间 X , 定义一等价关系如下: $x \sim y$ 当且仅当 x 和 y 属于 X 的同一个连通子空间. 则每个等价类称为 X 的一个**连通分支** (*connected component*)、**分支** (*branch*) 或**支集**.

例 2.4. 3×3 正交矩阵构成的群 $O(3)$ 具有两个连通分支. 单位连通分支是子群 $SO(3)$.

例 2.5. 闵氏时空中的 *Lorentz* 变换构成 *Lorentz* 群的单位连通分支.

距离空间

距离空间是具有距离定义的特殊集合. 在 \mathbf{R}^n 上用勾股定理定义距离, 就可从拓扑学转向分析学, 并一路走向微积分. 本节只简述很少的分析学, 可能对流形的学习暂时无用, 但对因果结构、泛函分析等的理解很有裨益.

距离空间和内积空间的区别在于, 后者是定义矢量空间上的双线性函数 (度规), 前者是直接定义任意集合中两元素的距离函数 (度量). 二者本无依赖, 但亦可有所关联.

定义 2.24. 一个**距离空间**或**度量空间** (*metric space*) 是一个集合 X 以及一个函数 $d: X \times X \rightarrow \mathbf{R}$, 满足

(对称) $d(x, y) = d(y, x)$

(正定) $d(x, y) \geq 0$, 当且仅当 $x = y$ 时取等;

(三角不等式) $d(x, z) \leq d(x, y) + d(y, z)$.

函数 d 称作两点间的度量函数或距离. 在普通几何中度量的概念一般被理解为欧氏度量. 显著的特征是正定性 (以及三角不等式). 可以引入具有不同符号的不定度量, 就像不定度规一样, 但是这些度量不被用来在所关心的空间上诱导拓扑.

定义 2.25. 给定距离空间 X 中一点 $x \in X$, 我们定义以 x 为中心, 正实数 r 为半径的**开球** (*open ball*) 为 $B(x, r) = \{y \in X : d(x, y) < r\}$.

可以验证, 给定距离空间 X , 取 X 中全体开球组成的集族 $\{B(x, \varepsilon) : x \in X, \varepsilon > 0\}$ 构成了 X 上的一个 (x 处的局部) 拓扑基. 就是说 x 的一组邻域使得 x 的每个

¹流形的路径连通与连通等价, 稍后叙述之.

邻域都包含这一组中的一个元. 这样就可以定义 X 中的序列 $\{u_n\}$ 收敛于 $u \in X$ 当且仅当 $d(u_n, u)$ 收敛于 0. 此拓扑基所生成的拓扑, 即

$$\mathcal{T} := \{\text{空集或 } X \text{ 中能表为开球之并的子集}\}$$

称为 X 上的给定度量所诱导的拓扑, 或称与距离相容的拓扑, 这个时候称拓扑空间是可距离化的. 如上这样从开球定义的拓扑可称为自然拓扑或通常拓扑.

定理 2.1.3 (分析学的 ε - δ 语言). 令 f 为距离空间 X 到距离空间 Y 的映射. 则 f 在 $x \in X$ 处连续 (对所诱导的拓扑) 当且仅当对任意 $\varepsilon > 0$, 存在 $\delta > 0$, 使得

$$B(x, \delta) \subseteq f^{-1}[B(f(x), \varepsilon)].$$

即 $f^{-1}[B(f(x), \varepsilon)]$ 是 x 点邻域.

上述说法显然等价于: 对任意 $\varepsilon > 0$, 存在 $\delta > 0$, 使得若 $d_X(x, x') < \delta$, 则有 $d_Y(f(x), f(x')) < \varepsilon$. 注意, 这里给两个距离空间以各自不同的度量函数.

定理 2.1.4. 如果 f 在 X 上任意一点处连续, 则 f 是连续映射.

证. 改用拓扑语言就容易发现, 如果 f 在 X 上任意点连续, 那么对于任意 $x \in X, \varepsilon > 0$, $f^{-1}[B(f(x), \varepsilon)]$ 是 X 上的开集, 再结合开集的“任意并”性质和原像的保并运算性质, 容易得到: 对于任意 Y 中的开集 U , $f^{-1}[U]$ 是 X 中的开集. 此即连续的拓扑定义. \square

拓扑定义的好处就是脱离了距离、某点连续性的概念而直接定义一个映射整体的连续性, 并且可以对更为任意的拓扑空间、映射定义相应的连续, 尽管可能会更反直觉.

定理 2.1.5. 容易证明距离空间是 Hausdorff 的.

定义 2.26. 一个距离空间之间的映射 $f: X \rightarrow Y$ 是**一致连续的** (uniformly continuous), 如果对任意 $\varepsilon > 0$, 存在 $\delta > 0$, 使得对于所有 $x_1, x_2 \in X$, 若 $d(x_1, x_2) < \delta$, 则有 $d(f(x_1), f(x_2)) < \varepsilon$.

注意连续和一致连续的差别: 后者要更强而且对整个空间要求一致的 δ .

定义 2.27. 在 X 上定义的两个度量 d_1 和 d_2 是等价 (equivalent) 的, 如果存在正常数 a 和 b , 使得对任意两点 $x, y \in X$, 有 $ad_1(x, y) \leq d_2(x, y) \leq bd_1(x, y)$.

这显然是一个等价关系. 两个等价的度量诱导出相同的拓扑.

例 2.6. 给定集合 X , 我们定义离散度量如下: $d_0(x, y) = 1$ 若 $x \neq y$. 它在 X 上诱导出离散拓扑. 这是一个十分方便的描述离散拓扑的方法.

例 2.7. 在 \mathbf{R} 上, 通常度量是 $d(x, y) = |x - y|$, \mathbf{R} 上的通常拓扑由这个度量所诱导.

例 2.8. 更一般地, 在 \mathbf{R}^n 上, 我们可以定义一个度量

$$d_p(x, y) = \left(\sum_{k=1}^n |x_k - y_k|^p \right)^{1/p}, \quad p \geq 1,$$

这里 $x = (x_1, x_2, \dots, x_n), y = (y_1, y_2, \dots, y_n)$. 特别地, 对于 $p = 2$, 我们得到欧氏度量, 但是其他情形也是有用的. 继续这个序列, 可以定义

$$d_\infty = \max_{1 \leq k \leq n} \{|x_k - y_k|\}.$$

因为这是 $p \rightarrow \infty$ 的极限. 这一定义在量子力学很有用, 因为量子力学所采用的泛函分析相当于无穷维矢量的线性代数. 所有这些度量都在 \mathbf{R}^n 上诱导出相同的拓扑, 也就是通常拓扑. 当然, 亦可视作 \mathbb{R} 通常拓扑的乘积拓扑.

定义 2.28. 给定一个距离空间 X , 一个点列 $\{x_1, x_2, \dots\}$ 称为**基本列** (*fundamental sequence*) 或 **Cauchy 列**, 如果对任意 $\varepsilon > 0$, 存在一个正整数 N , 使得对任意 $k, l > N$, 有 $d(x_k, x_l) < \varepsilon$.

定义 2.29. 在距离空间 X 中给定一个点列 $\{x_1, x_2, \dots\}$, 一点 $x \in X$ 称为这个序列的**极限** (*limit*), 如果对任意 $\varepsilon > 0$, 存在一个正整数 N , 使得对任意 $n > N$, 有 $d(x, x_n) < \varepsilon$. 我们称这个序列**收敛** (*converge*) 到 x .

定义 2.30. 一个距离空间 X 称为**完备的** (*complete*), 如果其中每个 Cauchy 列都在 X 中有极限.

例 2.9. 实直线上的闭区间 $[0, 1]$ 是完备的, 而开区间 $(0, 1)$ 则不是. 例如, Cauchy 列 $\{1/n, n = 2, 3, \dots\}$ 在这个开区间上没有极限点 (看成实直线上的序列, 它当然有极限点 0).

例 2.10. 空间 \mathbf{R}^n 是完备的.

例 2.11. 由所有使得求和 $\sum_{k=1}^\infty x_k^2$ 收敛的实数序列 $\{x_1, x_2, \dots\}$ 构成的 Hilbert 空间 ℓ^2 在前面定义的 d_2 的无穷维推广所得度量下是完备的. 对任意 $p \geq 1$, 可以类似地定义 ℓ^p , 它们也是完备的, 从而是 Banach 空间.

完备性不是拓扑不变量. 例如, 开区间 $(0, 1)$ 与整个实直线同胚 (在通常拓扑下), 但是前者不完备而后者完备. 同胚可以方便地由正切函数给出.

定义 2.31. 距离空间 X 的一个子集 B 称为**有界的**, 如果存在一个半径为正实数 R 的球完全包含 B .

定理 2.1.6 (Heine-Borel). \mathbf{R}^n 的任何有界闭子集是紧致集. 其逆命题亦成立.

从而我们有一个非常好的关于 \mathbf{R}^n 的紧致子集的刻画: 有界闭集.

定理 2.1.7 (列紧性). \mathbf{R}^n 中的任何有界序列有一个收敛子列.

这些是分析学中老生常谈的定理, 故证明不再赘述 (包括下文), 感兴趣的读者可自行查阅.

定义 2.32. 考虑定义在 \mathbf{R} 的子集 A (通常为一区间) 上的一系列实值函数 $\{f_n\}$. 我们说 $\{f_n\}$ 在 A 上**逐点收敛** (*converge pointwise*), 如果对于每个 $x \in X$, 实数列 $\{f_n(x)\}$ 收敛. 从而我们可以定义函数 $f: A \rightarrow \mathbf{R}$ 为

$$f(x) = \lim_{n \rightarrow \infty} f_n(x),$$

并记成 $f_n \rightarrow f$.

定义 2.33. 一系列函数 $f_n: A \rightarrow \mathbf{R}, A \subseteq \mathbf{R}$ 称为**一致收敛** (*converge uniformly*) 到函数 $f: A \rightarrow \mathbf{R}$, 如果对任意 $\varepsilon > 0$, 存在一个正整数 N , 使得对所有 $x \in X$, 只要 $n > N$, 就有 $|f_n(x) - f(x)| < \varepsilon$.

定理 2.1.8. 令 $f_n: (a, b) \rightarrow \mathbf{R}$ 为一列函数, 在 $c \in (a, b)$ 处连续. 假设 f_n 在 (a, b) 上一致收敛到 f , 则 f 在 c 处连续.

例 2.12. 连续函数的逐点极限不一定连续. 令 $f_n(x) = x^n, x \in [0, 1]$, 而极限函数 f 不连续.

定义 2.34. 设 X 为一距离空间. 一个映射 $f: X \rightarrow X$ 称为**(严格) 压缩映射** (*contraction*), 如果存在固定的 $0 < c < 1$, 使得对所有 $x \neq y \in X$, 有 $d(f(x), f(y)) \leq cd(x, y)$.

定理 2.1.9 (Banach). 若 X 为一完备距离空间, f 为 X 上一压缩映射, 则 f 在 X 上有唯一不动点 x_0 , 即 $f(x_0) = x_0$.

Banach 不动点定理是非线性泛函分析中的一个有用工具.

例 2.13. 考虑微分方程 $dy/dt = F(y, t)$, 其初始条件 $y(0) = y_0$, 而 F 是 *Lipschitz* 连续的, 它提供了解的局部存在性和唯一性.

现在要拾遗一些代数

群这个概念相比于**集合**稍微多了一种元素间的乘法运算, 用以更方便地描述对称性. 当然, 这些概念可能对本节所述无用, 不过为读者以后的学习考虑, 分清这些也有必要.

代数学拾遗

接着, 按线性代数方式, 给定矢量、对偶矢量、张量的映射定义 (最好用范畴论、群环模域的知识弄清动机和定义合理性, 如逆、协变称呼的由来).

前文谈及矢量空间 \mathbb{R}^4 可理解为所有矢量所构成的集合, 但读者大抵是不太喜欢这般搪塞的, 毕竟什么又是矢量呢? “满足叠加原理的箭头” 这样的说法恐怕还不够.

范畴论

群环模域

我们循序渐进, 对此可先介绍简单的 “二元函数”:

定义 2.35 (双矢量线性映射). 设 $f: V \times V \rightarrow W$ 对每个 V 而言都是线性映射, 即 $\forall \alpha_1, \alpha_2, \beta_1, \beta_2 \in V, k_1, k_2 \in \mathbb{R}$ 有

$$f(k_1\alpha_1 + k_2\alpha_2, \beta_1) = k_1f(\alpha_1, \beta_1) + k_2f(\alpha_2, \beta_1),$$

$$f(\alpha_1, k_1\beta_1 + k_2\beta_2) = k_1f(\alpha_1, \beta_1) + k_2f(\alpha_1, \beta_2),$$

称 f 是 V 上的一个双矢量线性映射, 目前暂不混淆, 可简称双线性映射. $W = V$ 时为双线性变换; $W = \mathbb{R}$ 时为双线性函数.

上述表明若固定某一输入的矢量, 则 f 就成为另一矢量的 “单元” 线性映射, 当然, 由于二者都隶属于同一矢量空间 V , 则哪怕 f 是 “双元” 的, 它当然也是 V 上的 “单元” 线性映射.

定义 2.36 (降映射). 固定 α , f 相当于把 $\beta \in V$ 映到 $f(\alpha, \beta) \in W$ 的线性映射, 记作 $f_b(\alpha)$, 即¹

$$f_b(\alpha)(\beta) = f(\alpha, \beta). \quad (2.1.1)$$

以下专注于 $W = \mathbb{R}$ 的情况. 由于 $f_b(\alpha)$ 是 V 的线性函数, 因而是一个对偶矢量, 则 f_b 就是一个 V 到 V^* 的线性映射.

在 V 中取一组基 $\{e_\mu\}$, 任取两个 V 中的矢量 x, y , 设 $x = x^i e_i, y = y^j e_j$, 则

$$\begin{aligned} f(x, y) &= f(x^i e_i, y^j e_j) \\ &= x^i y^j f(e_i, e_j), \end{aligned}$$

把上式的系数 $f(e_i, e_j)$ 作为矩阵 A 的第 i 行第 j 列元素, 即令

$$A = \begin{pmatrix} f(e_1, e_1) & \cdots & f(e_1, e_n) \\ \vdots & \ddots & \vdots \\ f(e_n, e_1) & \cdots & f(e_n, e_n) \end{pmatrix},$$

¹ “b” 是五线谱里的降记号, 读作 “降” (flat).

称上述 A 为双线性函数 f 在基 $\{e_\mu\}$ 下的 (分量) 矩阵, 它由 f 和基 $\{e_\mu\}$ 唯一决定. 这样有

$$f(x, y) = X^\top AY. \quad (2.1.2)$$

其中 $X = (x^1, \dots, x^n)^\top, Y = (y^1, \dots, y^n)^\top$ 分别是 x, y 在 $\{e_\mu\}$ 下的坐标. 可以证明, 设两个 $n \times n$ 方阵 A, B , 若对任意两个 n 维列向量 X, Y 都有 $X^\top AY = X^\top BY$, 则 $A = B$, 因此, 在基 $\{e_\mu\}$ 下, 任何一个 $n \times n$ 矩阵也按照 $x^i y^j f(e_i, e_j)$ 的方式唯一决定了一个线性映射 f .

在 V 上任取两组基 $\{\alpha_\mu\}$ 和 $\{\beta_\mu\}$, 若 $\{\alpha_\mu\}$ 到 $\{\beta_\mu\}$ 的变换矩阵是 P , f 在 $\{\alpha_\mu\}$ 下和 $\{\beta_\mu\}$ 下的矩阵分别为 A, B , 则由 (2.1.2) 不难导出

$$B = P^\top AP. \quad (2.1.3)$$

从而同一双线性函数在不同的基底下的矩阵是合同的¹. 由于任意两个合同矩阵有相同的秩, 一个双线性函数 f 的所有可能矩阵都有一个固定的秩, 将其定义为 f 本身的秩, 记作 $\text{rank } f$. 事实上, 双线性函数 f 的秩与 f_b 作为线性映射的秩相等.

定义 2.37 (非退化). 如果 f 在 V 的某一组基下的矩阵是满秩的, 那么它在 V 的其它任何基底下也是满秩的, 此时我们说 f 是非退化的.

大家学过其充要条件是行列式非零 (矩阵非奇异), 或者说

$$\forall w \in V, f(v, w) = 0 \iff v = 0. \quad (2.1.4)$$

因而 f 是非退化的 $\iff f_b$ 是 $V \rightarrow V^*$ 的同构. 也就是说, 我们可以借助 f 这个 V 上的新结构建立 V, V^* 间 “最特殊” 的同构映射.

定义 2.38 (对称与反称). 若 f 对 $\forall \alpha, \beta \in V$ 满足

$$f(\alpha, \beta) = f(\beta, \alpha), \quad (2.1.5)$$

则称 f 是对称的; 而若

$$f(\alpha, \beta) = -f(\beta, \alpha), \quad (2.1.6)$$

则称 f 是反称的.

容易验证, 对称双线性函数的矩阵全都是对称的, 反称双线性函数的矩阵也全都是反称的.

定义 2.39 (度规). 矢量空间 V 上的一个对称、非退化的双线性函数 $g: V \times V \rightarrow \mathbb{R}$ 称为 V 上的一个度规. 另一方面 g_b 是 V, V^* 的同构映射, 称作**音乐同构** (canonical isomorphism, 或译为典范同构).

¹合同变换 (matrix congruence) 即 $A \mapsto X^\top AX$ 这类变换, 其中 X 是可逆的; 若两个对称矩阵合同, 则他们有相同的正负惯性指数, 即有相同的正、负特征值个数.

在 V 上规定一个度规后, $g_b(\alpha) \in V^*$ 就可看作专门与 α 对应的对偶矢量, 记为 α_b , 这样就与另一矢量 β 产生天生的乘法 $\alpha_b(\beta)$, 因此度规相当于把矢量和对偶矢量的天生乘法引入到两个矢量之间, 即 $g(\alpha, \beta)$, 可记为 $\alpha \cdot \beta$ 或 $\langle \alpha, \beta \rangle$, 对称双线性运算 $\langle \cdot, \cdot \rangle$ 称为 V 上的一个内积¹. 定义了内积的矢量空间称为内积空间 (*inner product space*). 内积空间上可以定义长度、夹角等复合直观的几何概念, 这各位高中就学过, 比如矢量 $\alpha \in V$ 的模或范数就是

$$|\alpha| := \sqrt{|\langle \alpha, \alpha \rangle|}, \quad (2.1.7)$$

有时为与行列式区分, 写作 $\|\alpha\|$. 若不追求虚实性, 最常用的写法是 $\alpha^2 = \langle \alpha, \alpha \rangle$, 这样允许“模方”是负的. 注意, 不要求内积正定性时, 非零矢量的内积或者模长也是可以为零的, 下一节立即就会学到.

定理 2.1.10 (Cauchy-Schwartz 不等式). 可以证明

$$|\langle \alpha, \beta \rangle| \leq |\alpha| |\beta|. \quad (2.1.8)$$

有限维矢量空间定义了正定内积后, 即有限维的正定内积空间, 就称为欧氏矢量空间, 不混淆时可称欧氏空间.

定义 2.40 (正交基). 基 $\{e_\mu\}$ 在 g 的衡量下是正交的, 若 $g(e_i, e_j) = 0, i \neq j$; 在其衡量下为归一的, 若 $g(e_i, e_i) = \pm 1$. 正交且归一的基底称为正交基, 不会混淆时可简称正交基.

比如在欧氏度规的正交基下, 对 $\forall x = x^i e_i, y = y^i e_i \in V$ 有 $\langle x, y \rangle = \delta_{ij} x^i y^j$.

定理 2.1.11. 任意内积空间总存在正交基, 且度规在此基的对角矩阵中对角元正负个数与所选正交基无关. 证明见 *Schutz(1980) P.65-66*.

由此, 欧氏空间总存在正交基, 故也可把正定内积空间 V 就看成配有“勾股定理”的 \mathbb{R}^n .

定义 2.41 (多重矢量线性映射). 设 $f: \underbrace{V \times \cdots \times V}_k \rightarrow W$ 对每个 V 都是线性映射, 即 $\forall v_1, \cdots, v_k, v \in V, \lambda \in \mathbb{R}$ 有

$$f(v_1, \cdots, v_i + \lambda v, \cdots, v_k) = f(v_1, \cdots, v_k) + \lambda f(v_1, \cdots, v, \cdots, v_k), \quad (2.1.9)$$

其中 v 处于第 i 个分量的位置, $i \in [1, k] \cap \mathbb{N}$, 称 f 是 V 上的一个 k 重矢量线性映射. $W = V$ 时为 k 重矢量线性变换; $W = \mathbb{R}$ 时为 k 重矢量线性函数.

¹注意, 不涉及物理学的数学教材会强调内积应当满足正定性: $\forall \alpha \in V, \langle \alpha, \alpha \rangle > 0$, 但我们会讨论不定度规, 因此不作这个要求. 此外, 正定内积必然是非退化的, 因此若只定义了正定性, 就可略去非退化这个限制.

函数这个词, 而统一叫作张量. 这样, 矢量、对偶矢量就视作一种 1 阶张量¹. 此前所谓的“多重矢量线性函数”其实是强调只与矢量作用, 其身份就相当于协变矢量 (对偶矢量), 因此只与矢量作用的张量又称作协变张量. 比如说度规就应是一个 $(0, 2)$ 型协变张量, 或者 2 阶协变张量.

定义 2.42 (张量的对称). 张量 T 关于某两个下槽对称, 若

$$T(\cdots, v, \cdots, w, \cdots) = T(\cdots, w, \cdots, v, \cdots), \quad \forall v, w \in V.$$

在分量上, 这意味着交换此二指标结果不变. 若

$$T(\cdots, v, \cdots, w, \cdots) = -T(\cdots, w, \cdots, v, \cdots), \quad \forall v, w \in V.$$

则称其关于此二下槽反称, 这样交换指标会使得分量反号. 为上槽时同理. 称多个槽对称或反称, 若其中任意两槽是对称或反称的.

这样, 度规和 Ric, T 都是 2 阶协变对称张量. 类似地, 度规的逆就应当看作 2 阶逆变对称张量, 而 Riemann 曲率 $R^\lambda_{\mu\nu\sigma}$ 就是个 $(1, 3)$ 型张量 (的分量).

定理 2.1.12. 设 $\mathcal{T}_V(k, l)$ 表示 V 上全体 (k, l) 型张量之集. $\mathcal{T}_V(k, l)$ 是矢量空间, $\dim \mathcal{T}_V(k, l) = (\dim V)^{k+l}$.

证. 仿照 V^* 的逐点加法、数乘、零元的定义来构造矢量空间 $\mathcal{T}_V(k, l)$. 下只证其一组基矢的个数为 $(\dim V)^{k+l}$.

可以预料, 张量的基矢应是这些基矢的张量积组合, 即

而每个指标取便 $\dim V$, 共有 $(\dim V)^{k+l}$ 种. 下证它们构成 $\mathcal{T}_V(k, l)$ 的一组基. 容易验证其线性无关. 再证 $\forall T \in \mathcal{T}_V(k, l)$ 可表为 □

下面用映射语言准确定义张量的缩并. 回顾一下, $(1, 1)$ 型张量 T 可看作从 V 到 V 的线性映射, 这正是线性代数所讲的线性变换. T 在任一基底 $\{e_\mu \otimes e^\nu\}$ 的分量排成的矩阵 (T^μ_ν) 显然与基底有关, 不难证明同一 T 在任意两个基底的分量对应的两个矩阵 (T^μ_ν) 和 (T'^μ_ν) 互为相似矩阵, 证明拆到基底即可:

$$\begin{aligned} T'^\mu_\nu &= T(e'^\mu; e'_\nu) = T((A^{-\top})^\mu_\rho e^\rho; A^\sigma_\nu e_\sigma) = (A^{-\top})^\mu_\rho A^\sigma_\nu T(e^\rho; e_\sigma) \\ &= (A^{-\top})^\mu_\rho A^\sigma_\nu T^\rho_\sigma = (A^{-1})^\mu_\rho T^\rho_\sigma A^\sigma_\nu = (A^{-1}TA)^\mu_\nu, \end{aligned}$$

于是有矩阵等式 $T' = A^{-1}TA$ (其中 T', A, T 都代表矩阵²), 可见 T' 与 T 互为相似矩阵. 以 T'^μ_μ ($\sum_{\mu=1}^n T'^\mu_\mu$ 的简写) 及 T^ρ_ρ 分别代表矩阵 T' 和 T 的迹, 则易得

$$T'^\mu_\mu = (A^{-1})^\mu_\rho T^\rho_\sigma A^\sigma_\mu = A^\sigma_\mu (A^{-1})^\mu_\rho T^\rho_\sigma = \delta^\sigma_\rho T^\rho_\sigma = T^\rho_\rho.$$

¹当然, 张量构成矢量空间, 因此张量也是一种矢量. 看起来有点怪, 其实是自然语言混淆了这个地方的矢量和先前所说的普通矢量空间的逆变矢量.

² T 有时代表张量有时代表矩阵, 读者应能根据上下文识别.

这就证明了同一 $(1, 1)$ 型张量在不同基底的矩阵有相同的迹. 在关心张量时, 应该抓住其与基底无关的性质, $(1, 1)$ 型张量 T 的迹 T^μ_μ 就是这样一种性质, 通常称为 T 的缩并, 记作 CT , 即

$$CT := T^\mu_\mu = T(e^\mu; e_\mu).$$

再讨论 $(2, 1)$ 型张量 T 的缩并. T 可记作 $T(\cdot, \cdot; \cdot)$, 它有两个上槽和一个下槽, 故有两种可能缩并: (1) 第一上槽与下槽的缩并 $C_1^1 T := T(e^\mu, \cdot; e_\mu)$; (2) 第二上槽与下槽的缩并 $C_1^2 T := T(\cdot, e^\mu; e_\mu)$. 若改用另一基底 $\{e'_\rho\}$ 定义这两种缩并, 分别记作 $(C_1^1 T)'$ 和 $(C_1^2 T)'$, 则易证 $(C_1^1 T)' = C_1^1 T$, $(C_1^2 T)' = C_1^2 T$. 由映射语言可知 $C_1^1 T$ 和 $C_1^2 T$ 都是 $(1, 0)$ 型张量, 它们在任一基底的分量可用 T 在该基底的分量表为 $(C_1^1 T)^\nu = T(e^\mu, e^\nu; e_\mu) = T^{\mu\nu}_\mu$, $(C_1^2 T)^\nu = T^{\nu\mu}_\mu$ (已略去求和号). 不难推广上述讨论而得出 (k, l) 型张量的缩并定义如下:

注意, $C_j^i T$ 与基底选择无关. 联合使用张量积和缩并运算可从原有张量得到各种类型的新张量. 例如, 设 $v \in V, \omega \in V^*$, 则 $v \otimes \omega$ 是 $(1, 1)$ 型张量, 而 $C(v \otimes \omega)$ 则是 $(0, 0)$ 型张量 (标量).

后面经常遇到先求张量积再做缩并的运算, 其结果可看作张量对矢量 (或对偶矢量) 的作用.

例 2.14. 用 v^μ, ω_μ 表示 v, ω 在同一基底的分量, 有

- (a) $C(v \otimes \omega) = \omega_\mu v^\mu = \omega(v) = v(\omega), \quad \forall v \in V, \omega \in V^*$;
- (b) $C_2^1(T \otimes v) = T(\cdot, v), \quad \forall v \in V, T \in \mathcal{T}_V(0, 2)$;
- (c) $C_2^2(T \otimes \omega) = T(\cdot, \omega; \cdot), \quad \forall \omega \in V^*, T \in \mathcal{T}_V(2, 1)$.

下面只给出 (b) 的证明, 其他结论的证明由此类推.

证. 待证的等式 (b) 左边的 $T \otimes v$ 是 $(1, 2)$ 型张量, 是一部有 1 个上槽、2 个下槽的售货机, 可表为 $T \otimes v(\cdot; \cdot, \cdot)$, 故

$$C_2^1(T \otimes v) = T \otimes v(e^\mu; \cdot, e_\mu),$$

所以欲证等式 (b) 只须证明下式

$$T \otimes v(e^\mu; \cdot, e_\mu) = T(\cdot, v),$$

而上式是对偶矢量等式, 欲证上式只须证明两边作用于任一 $u \in V$ 给出相同实数.

$$\begin{aligned} \text{左边作用于 } u &= T \otimes v(e^\mu; u, e_\mu) = T(u, e_\mu) v(e^\mu) \\ &= T(u, e_\mu) e^\mu(v) = T(u, e_\mu) v^\mu = T(u, v) = \text{右边作用于 } u. \end{aligned}$$

可见式 (b) 成立. □

除以上三式外还有许多类似等式. 这些等式是如下规律的表现: “ T 对 ω (或 v) 作用就是先求 T 与 ω (或 v) 的张量积再缩并”, 或者粗略地说, “作用就是先积后并”. 对两个张量先求张量积再缩并的操作常又简称为对它们做缩并, 因此上述粗略提法还可简化为 “作用就是缩并”.

定义 2.43 (升降指标). 张量的升降指标操作就是利用音乐同构, 故是相对于某度规而言的, 比如对 $(1,1)$ 型张量 T 降指标即 $C_1^1(g \otimes T) = T_b$, 而升指标用度规的逆¹, 即 $C_1^1(g^{-1} \otimes T) = T^\sharp$. 由度规非退化性知, 其矩阵是可逆的, 因而一定存在对应的度规的逆.

当运算过程中使用的度规较为单一时, 可略去对度规的声明. 使用音乐同构记号默认是升降全部指标.

抽象指标是一套契合于广相需要的几何语言, 并不广泛采用于当代文献和大部分书籍, 读者可只做了解, 因此笔者将本节标注 “*” 以示选读. 本书谈及几何语言, 实际上均指前几节的数学系表述.

2.2 流形与度规场

单位分解

闵氏几何

但这还不够, 我们往往要借助距离 (长度) 概念研究各图形的几何性质. 最接近日常生活的几何学称为**欧氏几何**, 其中两点距离定义为勾股定理的形式:

定义 2.44. 两元素 $a, b \in \mathbb{R}^n$ 的**欧氏距离**定义为

$$d(a, b) = \left(\sum_{i=1}^n (x_a^i - x_b^i)^2 \right)^{1/2}, \quad (2.2.1)$$

其中 x_a^i, x_b^i 分别为 a, b 的自然坐标.

定义 2.45. 设集合 \mathbb{R}^n , 其子集

$$B(x_0, r) = \{x \in \mathbb{R}^n \mid d(x, x_0) < r\} \quad (2.2.2)$$

称为**开球** (*open ball*), 其中 x_0 称为球心, $r > 0$ 称为半径. 进而可对 \mathbb{R}^n 定义其**开子集** $O \subset \mathbb{R}^n$, 简称**开集**, 若存在一族 $\{B(x_\alpha, r_\alpha)\}$ 使得

$$O = \bigcup_{\alpha} B(x_\alpha, r_\alpha), \quad (2.2.3)$$

或者 $O = \emptyset$. 所有按此方式定义的开集 O 组成的集合 \mathcal{T}_u 称为 (\mathbb{R}^n) 的**通常拓扑** (*usual topology*), 称 $(\mathbb{R}^n, \mathcal{T}_u)$ 为 (\mathbb{R}^n) 的**通常拓扑空间** (*usual topological space*).

¹这里五线谱符号 \sharp 读作 “升” (*sharp*).

不过, 这些理论必然涉及对任意集合“开性质”的讨论, 因此也可说, 拓扑学的作用就是定义开子集.

注. 在分析学中将 α 准确地定义为指标集的元素. 通常来讲, 若指标集是有限集, 则我们会强调地写作 $\bigcup_{\alpha=1}^n$, 若无任何要求 (可以有限也可以无限), 则记作上述形式.

例 2.15. \mathbb{R} 的开球称为开区间. \mathbb{R}^2 的开球称为开圆盘.

定义 2.46. 设有一集合 M , 若存在双射 $\psi: M \rightarrow V$, 使得 V 是 $(\mathbb{R}^n, \mathcal{T}_u)$ 的开集, 则称为 M 的一个**整体坐标映射**, 其上域 V 称为**坐标值域**. 就整体坐标映射而言, 又可直接称为**整体坐标系**. 坐标值域所含于的 \mathbb{R}^n 可被强调为**坐标空间**. 通常将背景的元素记作字母 (称为**点**), 而将坐标空间中的元素记作数组 (x^μ) (称为**坐标**), 这样整体坐标系又可记作 $\{x^\mu\}$.

注. 对于一般的 M , 其上很可能不存在以 M 为定义域的整体坐标系, 我们以后再讨论. 但就 $M = \mathbb{R}^n$ 而言, 总存在整体坐标映射, 并且还可以是连续的, 其连续性由分析学定义. 显然, 任意 $p \in \mathbb{R}^n$ 都存在自然坐标 x^μ , 这便给出了连续的整体坐标映射 $\psi: \mathbb{R}^n \rightarrow \mathbb{R}^n$, 称为 \mathbb{R}^n 的**自然坐标系**. \mathbb{R}^n 的整体坐标系显然不唯一, 比如, 总可在自然坐标系基础上给 x^1 分量加 1. 这就是读者所熟悉的“平移”操作.

定义 2.47. 若集合 M 存在两套整体坐标系 ψ_α, ψ_β , 它们一般对 $p \in M$ 赋予不同坐标. 由于坐标映射都是双射, 因此复合映射 $\psi_\alpha \circ \psi_\beta^{-1}, \psi_\beta \circ \psi_\alpha^{-1}$ 皆是双射, 且互逆, 均称**坐标变换**.

定义 2.48. 令 $M = \mathbb{R}^n$ 及自然坐标系 $\{x^i\}$, **欧氏线元**定义为

$$d\ell = \left(\sum_{i,j=1}^n \delta_{ij} dx^i dx^j \right)^{1/2}, \quad (2.2.4)$$

这里 δ_{ij} 表示 n 阶单位阵的矩阵元, x^i 的微分 dx^i 由分析学给出. 凡能使

$$d\ell^2 = \sum_{i,j=1}^n \delta_{ij} dx^i dx^j = \sum_{i,j=1}^n \delta_{ij} dx'^i dx'^j \quad (2.2.5)$$

的坐标系 $\{x'^i\}$ 均称 **Descartes 坐标系**.

例 2.16. *Descartes* 坐标系可包括自然坐标系及其平移、旋转.

定义 2.49. 取 $M = \mathbb{R}^n$ 及自然坐标系 $\{x^\mu\}$. 凡能使闵氏线元平方为

$$ds^2 = \sum_{\mu,\nu=0}^3 \eta_{\mu\nu} dx^\mu dx^\nu = \sum_{\mu,\nu=0}^3 \eta_{\mu\nu} dx'^\mu dx'^\nu \quad (2.2.6)$$

的坐标系 $\{x'^\mu\}$ 均称为 **Lorentz 坐标系**或**惯性坐标系**.

同理, 惯性坐标系可包括自然坐标系的平移或“旋转”, 用矩阵 Λ 描述“旋转”的部分, 换言之, 就是使得 Λ 为正交矩阵的仿射变换.

由 \mathbb{R}^4 的线元不变性, 可严格导出变换的仿射性, 便绕过了对各种“公设”的争辩:

定理 2.2.1. \mathbb{R}^4 上的任意两个惯性坐标系间的变换仿射.

证. 已经证明过

□

定义 2.50. 设映射 $g: A \rightarrow B$ 和 $f: B \rightarrow C$, 令 $A = B = C$, 若 $f(g(x)) = g(f(x))$, 则称 f, g 是**对易** (commutative) 的. 称映射 $[f, g]$ 为 f, g 的**对易子** (commutator), 使得 $[f, g](x) = f(g(x)) - g(f(x))$. f, g 对易即说 $[f, g](x) = 0$, 简写为 $[f, g] = 0$ 或 $f \circ g = g \circ f$.

例 2.17. Fubini 定理

$$\left[\frac{\partial}{\partial x^\mu}, \frac{\partial}{\partial x^\nu} \right] = 0 \quad (2.2.7)$$

又称**偏导对易性**.

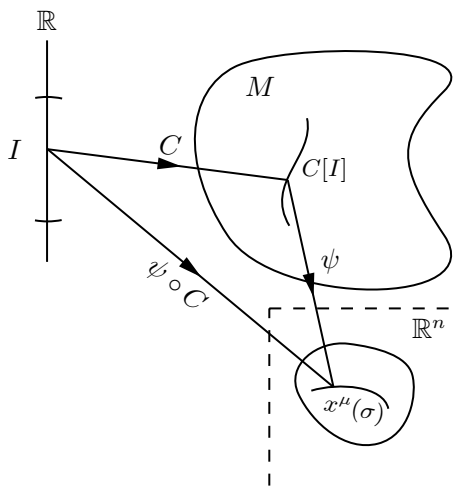


图 2.1: 曲线

相较于方程表述, 曲线的参数表达是更为方便的手段:

定义 2.51. 设 I 为 \mathbb{R} 的一个区间 (无论开闭) 和一个集合 M , 若存在 C^r 映射 $C: I \rightarrow M$, 就称 M 上的一条 C^r **曲线** (curve), 此处映射 C^r 性是指借助整体坐标系 $\psi: M \rightarrow V \subset \mathbb{R}^n$ 后, 所表的 $\psi \circ C: I \rightarrow \mathbb{R}^n$ 在分析学中的 C^r 定义. 变量 $\sigma \in I$ 称为曲线的**参数**, 映射的像 $C[I] \subset X$ 称为**路径** (path), 不混淆时亦可称曲线.

注. 常考虑至少分段光滑的曲线. 由于参数具有方向, 因此曲线也“有向”.

定义 2.52. 在 \mathbb{R}^4 中, 若曲线的参数方程满足

a

(2.2.8)

惯性观者世界线

用时空图计算

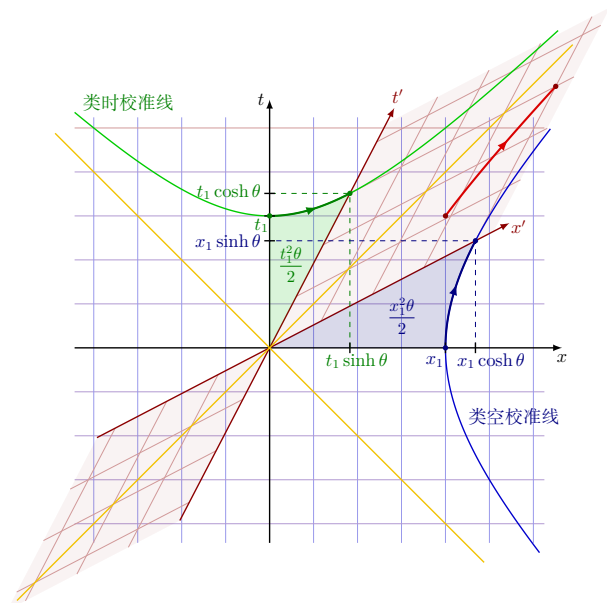


图 2.2: 校准线

在时空图中比较线长的最佳方法是用双曲线校准. 令双曲线 $x^2 - t^2 = x_1^2$, 其与 x 轴交于 $(t, x) = (0, x_1)$. 根据闵氏度规, 双曲线上的点

$$t = x_1 \sinh \theta, \quad x = x_1 \cosh \theta$$

与原点连线是类空间隔且恒定, 可用来比较类空间隔, 故称类空校准线. 从 $(0, x_1)$ 至 (t, x) 扫过角度 θ , 易得扫过面积为 $x_1^2 \theta / 2$ (类比欧式几何的圆). 同理推知类时校准线的结论.

切空间上一个对称、非退化的 2 阶协变 (即 (0,2) 型) 张量称为度规. 可见度规选择无穷, 十分任意. 借此定义切空间元素

的内积、正交性、模长、正交归一性. 度规在正交归一基的分量一定是对角归一矩阵, 其迹称为号差. 号差将度规分为正定 (Riemann 的)、负定和不定. 非正定就称伪 Riemann 的. 只有一个对角元符号不同的不定度规称 Lorentz 的. 对 4 维情况, 号差为 ± 2 . 按号差 +2 的习惯, 一般定义 $\eta_{\mu\nu}$ 为 $\text{diag}(-1, 1, 1, 1)$ 的矩阵元. 推广到切丛上, 度规场除对称、处处非退化外, 还要求号差处处一致. 显然度规场的号差与坐标系选择无关. 类比线性代数的合同变换, 可见总存在坐标系使其分量矩阵为对角归一矩阵.

固有时需要将切空间的模长概念搬到切丛上来, 再由单位分解搬到时空上来. 即首先在拓扑空间那里就应定义好至少分段可导 (以定义切矢) 的曲线为严格单调映射 (以排除折返情况) $C : [t_1, t_2] \rightarrow M$ (路径为像 $\text{Im } C$), 再定义线长是其上 Riemann

积分泛函¹

$$L(C) = \int_{t_1}^{t_2} \sqrt{\left| g \left(\frac{d}{dt}, \frac{d}{dt} \right) \right|} dt,$$

这里 d/dt 表示曲线按 t 参数化时的切矢. 可证可导性与坐标系选择无关, 线长与参数选择无关. 单调映射排除了一部分使 d/dt 为零元的情形, 但还应要求 $[t_1, t_2]$ 中不能有 Lebesgue 测度大于零的区间使像点 $C(t)$ 一致. 测度为零区间是允许的, 如 t 的鞍点或曲线自交的情形. 借助于至少囊括曲线段的坐标系 $\{x^\mu\}$, 则 x^μ 在曲线上关于 $[t_1, t_2]$ 至少分段可导, 故 d/dt 可展开至坐标基

$$\frac{d}{dt} = \frac{dx^\mu}{dt} \frac{\partial}{\partial x^\mu},$$

则很容易发现上式改写为

$$L(C) = \int_C \sqrt{|g_{\mu\nu} dx^\mu dx^\nu|}.$$

因此可引入形式记号 $ds^2 = g_{\mu\nu} dx^\mu dx^\nu$ 而不论及线元 ds 的“虚实性”. 此外, 很容易发现参数单调性还迎合了物理学的因果要求.

Wald、Hawking 等喜欢将切矢记作 $\partial/\partial t, \partial_t$. 这是因为当考虑多种曲线参数 (尤其是若干不同曲线) 时, 所需的 “ dx^μ/dt ” 具体是指 $\partial x^\mu/\partial t$, 而由

$$\frac{\partial}{\partial t} = \frac{\partial x^\mu}{\partial t} \frac{\partial}{\partial x^\mu},$$

故应记 ∂_t . 只有一个参数时两种写法全微分同偏微分无异. 但一般来讲建议用最准确的写法 d/dt , 因为 Hawking 等所用符号 $dx^\mu/dt, \partial x^\mu/\partial t$ 有歧义.

2.3 超曲面

2.4 仿射联络与曲率

定义导数算符、丛上仿射联络. 给出无挠性、度规适配性定义, 二者确定出存在且唯一的 Levi-Civita 联络 ∇ . 迎合等效原理的需求, 除涉及量子、奇点等极端情形, 物理学选择无挠联络, 而迎合广义协变性 (即张量理论), 又期望有度规适配性, 因此我们选择 Levi-Civita 联络. 在任何情况下, 建议将克氏符定义为分量形式的 Levi-Civita 联络系数 $\Gamma_{\mu\nu}^\lambda = \langle e^\lambda, \nabla_{e_\mu} e_\nu \rangle$, 其中 \langle, \rangle 尤指对偶矢量和矢量的缩并. 这样避免讨论其张量性之类的无伤大雅的问题.

¹或视作 $\text{Im } C$ 的 Lebesgue 测度, 反正在物理上两种积分无区别.

定义协变微分要求背景附加额外结构——仿射联络（这区别于外微分），以下简称联络¹.

可类分析学的方向导数一步步地引出可作用于张量的联络概念.

定义 2.53 (方向导数). 背景上的一个联络 ∇ 是指一种法则，其给任意矢量场 X 赋以算子 ∇_X ，使得任意可导²矢量场 Y 映射为另一矢量场 $\nabla_X Y$ ，满足：

(张量性) $\nabla_X Y$ 是以 X 为变元的张量. 即对任意函数 f, g 和可导矢量场 X, Y, Z 有 $\nabla_{fX+gY} Z = f\nabla_X Z + g\nabla_Y Z$. 此即要求 ∇_X 仅依赖于某点处 X 的方向；

(线性性) $\nabla_X Y$ 对 Y 是线性的. 即 $\nabla_X(\alpha X + \beta Y) = \alpha\nabla_X X + \beta\nabla_X Y$ ，这里 α, β 为实数；

(对函数) 对任意可导函数 f 和可导矢量场 Y 有 $\nabla_X(fY) = X(f)Y + f\nabla_X Y$ ，这里 $X(f)$ 在坐标基上表示为 f 的方向导数 $X^\mu \frac{\partial f}{\partial x^\mu}$.

如上 $\nabla_X Y$ 称为 Y 沿 X 方向关于 ∇ 的导数.

由张量性，还可直接定义 Y 关于 ∇ 的导数 ∇Y 定义为 $(1,1)$ 型张量场，而它与 X 缩并后产生矢量 $\nabla_X Y$. 于是亦可定义为

定义 2.54 (协变微分). C^{r+2} 流形上的一个 C^r 联络是指一种法则，其给任意点赋以协变微分（取 Lorentz 协变性之意） ∇ ，使得任意 C^{r+1} 矢量场 Y 映射为 $(1,1)$ 型 C^r 张量场 ∇Y ，满足：

(线性性) $\nabla(\alpha X + \beta Y) = \alpha\nabla X + \beta\nabla Y$ ，这里 α, β 为实数；

(对函数) 对任意可导函数 f 和可导矢量场 Y 有 $\nabla(fY) = df \otimes Y + f\nabla Y$ ，这里为绕开外微分概念， d 可仅理解为函数的微分.

在坐标域上，给定任意 C^{r+1} 矢量基 $\{e_\mu\}$ ，则 ∇Y 的展开就写为

$$\nabla Y = \nabla_\mu Y^\nu e^\mu \otimes e_\nu, \quad (2.4.1)$$

这里 ∇_μ 称为协变导数；而方向导数就在相同基下写作

$$\nabla_X Y = X^\mu \nabla_\mu Y^\nu e_\nu. \quad (2.4.2)$$

可见 ∇ 实际上就相当于给其作用对象添加一个下标，于是再结合一些微分学的特性即可推广至任意张量：

定义 2.55 (联络). 联络 ∇ 是背景上的一个算符场，其将 (k, l) 型 C^r 张量场映为 $(k, l+1)$ 型 C^{r-1} 张量场，满足

¹第 4 章会讲到，一般的联络可定义在各种“空间”上，定义在流形矢量丛上的称为仿射联络，此即本章所讨论的，且一律简称联络.

²矢量场可导指矢量场按任意坐标基展开的分量为坐标域上的可导函数.

(线性性) $\nabla(\alpha T + \beta S) = \alpha \nabla T + \beta \nabla S$, 这里 α, β 为实数而 T, S 同型;

(对函数) 对任意可导函数 f 有 $\nabla f = df$ (注意此处非梯度);

(Leibniz 律) 对任意张量 T, S 有 $\nabla(T \otimes S) = T \otimes \nabla S + S \otimes \nabla T$;

(缩并对易性) $[\nabla, C] = 0$, 这里 C 代表缩并算子.

后两条保证我们可推导任意张量的方向导数或协变微分. 一般的所谓“微分算符 (operator)”或其相应的“导数算符”并不要求其作用结果仍为张量场 (但满足联络其它要求), 因此可能算符自身就与坐标系选择 (或其它人为因素) 有关¹. 有时为区别于微分算符, 可将协变微分记作 D .

定义 2.56 (沿曲线求导). 若 T 是沿 C^1 曲线 $\lambda(t)$ 定义的 C^1 张量场 (只在路径上定义), 则我们还可定义 T 沿 $\lambda(t)$ 关于所给联络的导数为

$$\frac{DT}{dt} = \nabla_{\partial_t} \bar{T}, \quad (2.4.3)$$

这里 \bar{T} 表示将 T 以任意方式延拓至路径某开邻域的 C^1 张量场, 即选取任意范围包含路径的张量场 \bar{T} , 但满足在路径上取值为 T 且可导. $\frac{DT}{dt}$ 只在路径上定义.

定义 2.57 (平行移动). 若 $DT/dt = 0$, 则称 T 沿曲线关于所给联络平行移动.

给定具有端点 p, q 的曲线 λ , 将任意给定张量由 p 出发沿曲线平行移动, 即可得到 q 点的唯一张量. 因此, 沿曲线 λ 的平行移动是从 p 处张量空间到 q 处同型张量空间的线性映射, 且保持所以张量积、缩并不变. 此即说, 沿给定曲线从 p 至 q 平行移动可决定从 V_p 至 V_q 的一个同构 (若曲线自相交, 则 p, q 也可以重合), 这是联络名称的由来. 通常考虑这样的移动曲线已使同构映射唯一, 即沿曲线 λ 的切矢量自身关于联络的导数为零:

定义 2.58 (测地线). 若沿 $\lambda(t)$ 的 $\nabla_X X$ 平行于 $X = \frac{\partial}{\partial t}$, 即若存在函数 f 使任意坐标有 $X^\nu \nabla_\nu X^\mu = f X^\mu$, 则称 $\lambda(t)$ 为测地线. 总可取一个新参数 ξ 使

$$X^\nu \nabla_\nu X^\mu = 0, \quad X = \frac{\partial}{\partial \xi}, \quad (2.4.4)$$

则该参数称为仿射参数. 容易验证仿射参数可确定到线性变换 $\xi' = a\xi + b$, 此处 a, b 为常数. 此后谈及测地线均默认仿射参数化.

这样, 只要沿着测地线做关于给定联络的平行移动, 就可确定一点附近矢量空间的一致同构².

¹除非强行将结果理解为“坐标依赖的张量场”, 但只有少数书秉持这种思想, 如 Wald, 见第 4 章抽象指标简介一节. 本书将张量场定义为“背景上像为张量的单值映射”, 在一个背景点只赋以唯一张量, 因此所谈张量场默认“坐标无关”.

²第 4 章证明小的邻域内两点间只存在一条测地线.

可以证明, 任意流形必定存在满足定义的联络. 并且由如下讨论可知, 其数量无限多. 为此先给出联络系数的严格定义:

定义 2.59 (联络系数). 给定背景上的联络 ∇ , 设矢量基 $\{e_\mu\}$ 及其对偶, 则联络系数定义为 $\Gamma^\lambda_{\mu\nu} = e^\lambda(\nabla_{e_\mu} e_\nu)$, 或亦可记作

$$\Gamma^\lambda_{\mu\nu} = \langle e^\lambda, \nabla_{e_\mu} e_\nu \rangle. \quad (2.4.5)$$

上式等价于

$$\nabla e_\nu = \Gamma^\lambda_{\nu} \otimes e_\lambda, \quad (2.4.6)$$

这里 $\Gamma^\lambda_{\nu} = \Gamma^\lambda_{\mu\nu} e^\mu$ 称为联络 1-形式.

定理 2.4.1. ∇Y 在坐标基上的表达为

$$\nabla_\mu Y^\lambda = \frac{\partial Y^\lambda}{\partial x^\mu} + \Gamma^\lambda_{\mu\nu} Y^\nu.$$

证. 对任意矢量场 Y 有

$$\begin{aligned} \nabla Y &= \nabla(Y^\nu e_\nu) = dY^\nu \otimes e_\nu + Y^\nu \Gamma^\lambda_{\mu\nu} e^\mu \otimes e_\lambda \\ &= \frac{\partial Y^\lambda}{\partial x^\mu} dx^\mu \otimes e_\lambda + Y^\nu \Gamma^\lambda_{\mu\nu} e^\mu \otimes e_\lambda, \end{aligned}$$

取矢量基为坐标基即可得到分量. □

同理求得任意张量的协变导数.

定理 2.4.2. 设 $e'_{\mu'} = A_{\mu'}^\mu e_\mu$, $e'^{\mu'} = A^{\mu'}_\mu e^\mu$, 则

$$\Gamma'^{\lambda'}_{\mu'\nu'} = A^{\lambda'}_{\lambda} (e'_{\mu'}(A_{\nu'}^\lambda) + A_{\mu'}^\mu A_{\nu'}^\nu \Gamma^\lambda_{\mu\nu}). \quad (2.4.7)$$

其在坐标基下就是此前所学的联络系数的坐标变换式. 由于 $e'_{\mu'}(A_{\nu'}^\lambda)$ 的存在, 其不满足张量变换律. 多出来的部分称为非协变部分.

联络是坐标无关的微分算符. 下面给出“坐标依赖的微分算符”一例:

定义 2.60 (普通微分). 设 $\{x^\mu\}$ 的坐标基及其对偶分别为 $\{\partial_\mu\}, \{dx^\mu\}$, 定义普通 (ordinary) 微分是一个微分算符 d (不是外微分), 其将一点处的 (k, l) 型张量映为 $(k, l+1)$ 型张量, 使得

$$dT = \frac{\partial}{\partial x^\sigma} T^{\mu_1 \cdots \mu_k}_{\nu_1 \cdots \nu_l} dx^\sigma \otimes \partial_{\mu_1} \otimes \cdots \otimes \partial_{\mu_k} \otimes dx^{\nu_1} \otimes \cdots \otimes dx^{\nu_l}. \quad (2.4.8)$$

普通微分只在函数 f 作用时才等同于外微分.

定理 2.4.3. $d(\partial_\nu) = 0, d(dx^\nu) = 0$.

证. 注意 dT 的分量就是 T 分量对坐标求偏导数, 但坐标基在该系的分量都是常数 0 或 1, 当然结果也就是个零张量. \square

定理 2.4.4. $ddT(u, v, \dots) = ddT(v, u, \dots), \forall$ 矢量场 u, v .

证. 注意 ddT 的分量就是对 T 分量求两次坐标偏导, 由分析学知偏导是对易的, 因此分量关于此二指标对称. \square

不难验证, 其满足联络其它要求, 但偏偏就不满足“将张量场映为张量场”, 因其自定义起就依赖于坐标系的微分算符, 且只在该坐标系的坐标域上有定义, 则张量场的普通微分一般不为张量场. 因而 d 可算作一种微分算符, 但不属于联络.

下面说明联络无限多. 从一般的微分算符谈起, 若 ∇ 是微分算符, 假设存在至少一个不同的微分算符 $\tilde{\nabla}$, 我们欲分辨其差异. 二者作用于同一函数 f 的结果均为 df , 故不能考虑对函数作用. 不妨考虑作用于矢量 v 或对偶矢量 ω .

考虑点 p 处对偶矢量在邻域上的任意两种延拓场 ω', ω , 则场 $\nabla\omega'$ 在 p 处取值一般不等于 $\nabla\omega$ 的¹. 但下面证明任意两联络的差 $\tilde{\nabla} - \nabla$ 作用于任意延拓 ω 是相同的.

定理 2.4.5. 设 $\omega_p = \omega'_p$, 则

$$(\tilde{\nabla} - \nabla)\omega|_p = (\tilde{\nabla} - \nabla)\omega'|_p,$$

这里 $(\tilde{\nabla} - \nabla)\omega$ 视作 $\tilde{\nabla}\omega - \nabla\omega$.

证. 只需证明

$$\tilde{\nabla}\Omega|_p = \nabla\Omega|_p,$$

其中 $\Omega = \omega' - \omega$. 取邻域上任意坐标系, 将 Ω 表为 $\Omega_\mu dx^\mu$. 则由 Leibniz 律有

$$\nabla\Omega = \Omega_\mu \nabla dx^\mu + dx^\mu \nabla \Omega_\mu,$$

由 $\Omega_\mu|_p = 0$, 其于 p 处取值为

$$\nabla\Omega|_p = dx^\mu \nabla \Omega_\mu|_p.$$

同理 $\tilde{\nabla}\Omega|_p = dx^\mu \tilde{\nabla}\Omega_\mu|_p$, 但 Ω_μ 是函数, 任意微分算符 ∇ 作用结果均为 $d\Omega_\mu$, 因此二者相等, 当然 p 处取值也相等. \square

虽然 $\nabla\omega$ 和 $\tilde{\nabla}\omega$ 都依赖于 ω 在一点某邻域的延拓情况, 但 $(\tilde{\nabla} - \nabla)\omega$ 只依赖于 ω 在一点的值, 且 $(\tilde{\nabla} - \nabla)$ 将一点 ω 线性地映为该处 $(0, 2)$ 型张量, 故其是该处矢量空间上的 $(1, 2)$ 型三重线性映射, 写作分量为

$$(\tilde{\nabla}_\mu - \nabla_\mu)\omega_\nu = C^\lambda_{\mu\nu}\omega_\lambda,$$

¹可类分析学中, 两一元函数在同一处取值相同但切线斜率一般不同.

此关系对任意点成立, 因此我们说任意联络之间在任意坐标系差一个 $(1, 2)$ 型矩阵场¹, 其任意性导致微分算符无限多. 取两微分算符都为联络, 则 $(1, 2)$ 型矩阵任意性导致联络无限多.

该矩阵场是否可称为张量场的分量呢? 答案是不一定, 注意, 我们在定义此矩阵时使用的是分量语言, 而非映射语言. 使用映射语言在一点给出的一定为张量, 如

$$(\tilde{\nabla} - \nabla)\omega = C_3^1(C \otimes \omega),$$

其中 C 是该点处与之对应的 $(1, 2)$ 型张量, 但不一定构成张量场. 假设二者都是联络, 则 C 是不变的, 因而可构成 $(1, 2)$ 型张量场. 实际上, 由此前讨论已知, 任意联络之差相当于联络系数之差, 而联络系数之差刚好消去了非协变部分, 因此联络系数之差满足张量变换律!

若其一恰为坐标依赖的, 比如取 $\tilde{\nabla}$ 为某坐标系的普通微分 d , 则

定理 2.4.6. d 和 ∇ 在此坐标系所差的 $(1, 2)$ 型矩阵即为 ∇ 在该系的联络系数. 部分书将此作为联络系数定义.

即使可对微分算符定义联络系数, 亦成立相关的协变导数公式, 但最好只考虑能映出张量场的联络, 这样就可以说联络系数修正了一定不协变的普通微分, 从而得到协变微分 ∇ , 此时联络系数同普通微分的不协变性相抵消.

设 $\{x^\mu\}$ 系下有

$$(d - \nabla)\omega = C_3^1(\Gamma \otimes \omega),$$

取另一 $\{x'^\mu\}$ 可给出不同张量:

$$(d' - \nabla)\omega = C_3^1(\Gamma' \otimes \omega),$$

由此发现一点处 Γ 本身就依赖于坐标系, 当然也就不构成张量场. 分量语言所给出的联络系数在 $\{x^\mu\}$ 系下是 Γ 在 $\{x^\mu\}$ 的分量, 而在 $\{x'^\mu\}$ 系下是 Γ' 在 $\{x'^\mu\}$ 的分量, 因此联络系数不满足张量变换律. 这与此前的讨论是一致的.

定理 2.4.7. $DT/dt = \nabla_{\partial_t} T$. 在坐标系下可表为

$$\frac{D}{dt} T^{\mu_1 \cdots \mu_k}_{\nu_1 \cdots \nu_l} = \frac{dx^\sigma}{dt} \nabla_\sigma T^{\mu_1 \cdots \mu_k}_{\nu_1 \cdots \nu_l}. \quad (2.4.9)$$

证. 按理来说, ∇T 应该还要取决于附近 \bar{T} 的情况, 因此要求曲线外应有定义, 但沿曲线方向导数 ∇_{∂_t} 事实上与 \bar{T} 无关. 为行文简洁, 不妨设张量 T 为一矢量 v . 将沿线导数在曲线上一点 p 处按坐标基展开为

$$\frac{dx^\nu}{dt} \nabla_\nu \bar{v}^\mu|_p = \frac{dx^\nu}{dt} \frac{\partial \bar{v}^\mu}{\partial x^\nu} \Big|_p + \Gamma^\mu_{\nu\sigma} \frac{dx^\nu}{dt} \bar{v}^\sigma|_p = \frac{dx^\nu}{dt} \frac{\partial \bar{v}^\mu}{\partial x^\nu} \Big|_p + \Gamma^\mu_{\nu\sigma} \frac{dx^\nu}{dt} v^\sigma|_p,$$

¹ 矩阵场由于给每一点赋以的是坐标分量数表, 因此不能把矩阵场直接理解为定义在背景上, 而是依附于坐标系上. 注意同张量场相区别.

任意不同延拓 \bar{v}, \bar{v}' 实际上给出了同一沿线导数, 只需证明

$$\frac{d\bar{v}^\mu}{dt} = \frac{d\bar{v}'^\mu}{dt}.$$

上式显然成立, 因为任意延拓的 $\bar{v}^\mu(t)$ 实为领域内的 \bar{v} 同曲线映射复合的一元函数, 该一元函数正是 $v^\mu(t)$. 以后凡谈及只在曲线上定义的张量场时, 不会讨论协变微分 (因为没意义), 但可求其沿线导数 (沿线导数同延拓无关). \square

定理 2.4.8. 切矢平行移动直接给出使用仿射参数的测地线.

证. 取 $X = \frac{\partial}{\partial \xi}$, $DX/d\xi = \nabla_X \bar{X} = 0$ 等价于 $\nabla_X X = 0$. 这表明, 测地线可不用协变微分定义, 而是用沿线导数定义. \square

在第 1 章我们学过挠率为联络系数反称部分 (的两倍), 其按映射语言可重新定义为

定义 2.61 (挠率). 给定 C^r 联络 ∇ , 定义 $(1,2)$ 型 C^{r-1} 张量场 \mathcal{T} 为挠率, 满足

$$\mathcal{T}(X, Y) = \nabla_X Y - \nabla_Y X - [X, Y], \quad (2.4.10)$$

这里 $[X, Y]$ 是 X, Y 的对易子¹. 这里 X, Y 为任意 C^r 矢量场.

定理 2.4.9. 坐标基下 $\mathcal{T}^\mu_{\lambda\nu} = 2\Gamma^\mu_{[\lambda\nu]}$.

证. 只需注意坐标基对易子 $[\partial_\mu, \partial_\nu] = 0$, 因为对任意函数的偏导对易. \square

有挠率的联络一般只应用在某些其它引力理论中. 广相理论默认使用无挠联络, 即 $\mathcal{T} = 0$, 其物理内涵在第 3 章详细讨论. 无挠联络使我们计算矢量对易子时可写为

$$[X, Y] = \nabla_X Y - \nabla_Y X. \quad (2.4.11)$$

定理 2.4.10. $\mathcal{T} = 0$ 等价于对任意函数 f 的协变微分对易.

度规介入后, 就可谈及一种特殊而实用的联络. 设任意 u, v 沿曲线平行移动, 则为保证二者内积沿曲线不变, 将等价于说

$$\nabla_T g = 0,$$

这里 T 为曲线切矢. 若对任意曲线成立, 则又等价于 $\nabla g = 0$. 进而, 若再要求 ∇ 的无挠性, 便可将联络确定到一种上. 这是因为, 给定度规场 g , 可以用任意微分算符 $\tilde{\nabla}$ 及差 $C^\lambda_{\mu\nu}$ 构造出 ∇ 使得 $\nabla g = 0$, 利用无挠性可求解 $C^\lambda_{\mu\nu}$ 的表达式. 不妨用普通微分 d , 这样便有 $\Gamma^\lambda_{\mu\nu}$ 同 g 的关系 (我们已经学过), 而该表达式存在且唯一.

¹ 回忆一下, 在坐标系那一节我们说流形上使用的“逆变矢量”皆为函数的算符, 故其对易子按算符那样定义.

定义 2.62 (Levi-Civita 联络). 满足

$$\nabla g = 0 \quad (2.4.12)$$

且无挠的联络 (场) 称为 *Levi-Civita* 联络. 上式称为与度规 g 的适配性或相容性.

与度规适配的联络可保持度规的“均匀性”, 而广相理论又一般默认无挠性, 这样, 我们通常选择 Levi-Civita 联络研究广相物理学.

曲率张量不因真空方程而一定为零的部分, 称为 Weyl 张量. 这个张量量度了测地线组的“潮汐”扭曲. 所以真空区域里的引力场的“局部强度”在 Newton 极限之下是与宏观的试验物质的潮汐力相关的, 而不是与引力的范数相关的.

测地偏离方程

Raychaudhuri 方程

对时空 (M, g) 中一点 q , 对偶矢量和矢量的缩并其实就是将对偶矢量作用于矢量, 即 $\langle \mathbf{e}^\mu, \mathbf{e}_\nu \rangle = \mathbf{e}^\mu(\mathbf{e}_\nu)$. 张量积用映射语言定义就是将分别作用所得实数相乘.

一个 2 阶协变张量 T 其实就是线性函数

$$T : T_q M \times T_q M \rightarrow \mathbb{R}, \quad (2.4.13)$$

只要给定任意两个矢量便能给出唯一实数

$$T(\mathbf{v}, \mathbf{w}) = T_{\mu\nu} v^\mu w^\nu, \quad (2.4.14)$$

其中用到 \mathbf{v}, \mathbf{w} 在坐标基下的分量.

这样

$$T_{\mu\nu} v^\sigma w^\lambda \mathbf{e}^\mu \otimes \mathbf{e}^\nu (\mathbf{e}_\sigma, \mathbf{e}_\lambda) = T_{\mu\nu} v^\sigma w^\lambda \langle \mathbf{e}^\mu, \mathbf{e}_\sigma \rangle \langle \mathbf{e}^\nu, \mathbf{e}_\lambda \rangle.$$

容易看出, 张量分量等于张量作用于坐标基所得

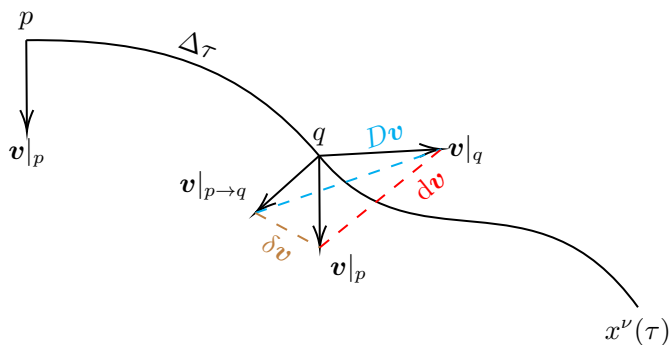
$$T_{\mu\nu} = T(\mathbf{e}_\mu, \mathbf{e}_\nu). \quad (2.4.15)$$

读者可同矢量点乘基底的投影操作相比较. 选择另一坐标基可给定另一分量 $\hat{T}_{\hat{\mu}\hat{\nu}} = T(\hat{\mathbf{e}}_{\hat{\mu}}, \hat{\mathbf{e}}_{\hat{\nu}})$. 将坐标基的变换代入其中有

$$\hat{T}_{\hat{\mu}\hat{\nu}} = T \left(\frac{\partial x^\mu}{\partial \hat{x}^{\hat{\mu}}} \mathbf{e}_\mu, \frac{\partial x^\nu}{\partial \hat{x}^{\hat{\nu}}} \mathbf{e}_\nu \right) = \frac{\partial x^\mu}{\partial \hat{x}^{\hat{\mu}}} \frac{\partial x^\nu}{\partial \hat{x}^{\hat{\nu}}} T(\mathbf{e}_\mu, \mathbf{e}_\nu) = T_{\mu\nu} \frac{\partial x^\mu}{\partial \hat{x}^{\hat{\mu}}} \frac{\partial x^\nu}{\partial \hat{x}^{\hat{\nu}}},$$

这便用映射语言推出了 T 分量的协变性.

度规适配性便于我们研究协变导数的直观含义. 以 $Dv^\mu/d\tau$ 为例, 克氏符项究竟是如何修正的? 如图 2.3, 设某曲线 $x^\nu(\tau)$ 上 p 为参数零点, 而邻近点 q 的参数记

图 2.3: 图中 $\Delta\tau$ 等差值均夸大

作 $\Delta\tau$ (可正可负). 若任意矢量场 $v(x^\mu)$ 至少在 p, q 线段 (即线上 p 邻域) 上有定义, 便可讨论 p 处沿线导数算符. 设矢量场在 p, q 分别取值 $v|_p, v|_q$. 根据定义,

$$\left. \frac{dv^\mu}{d\tau} \right|_p = \lim_{\Delta\tau \rightarrow 0} \frac{v^\mu|_q - v^\mu|_p}{\Delta\tau} \quad (2.4.16)$$

即直接比较 $v^\mu|_q, v^\mu|_p$, 但 $v|_p, v|_q$ 隶属不同切空间, 其差 dv 不存在于任何切空间里. 的确, $dv^\mu/d\tau$ 在 p, q 都无协变性, 故既非 p 点的矢量, 亦非 q 点的, 可见这是普通导数不协变缘由的直观理解. 但我们仍欲形象地画出差值, 需将矢量都 (以某种方式) 移至同一点作差. 不妨考虑 q 点. 图中 q 点画有“相同”矢量 $v|_p$, 表示将分量 $v^\mu|_p$ 照搬至 q , 故 dv 可标注为图中红线. 而另一方面,

$$\begin{aligned} \left. \frac{Dv^\mu}{d\tau} \right|_p &= \lim_{\Delta\tau \rightarrow 0} \frac{v^\mu|_q - v^\mu|_p}{\Delta\tau} + \Gamma_{\lambda\nu}^\mu|_p v^\nu|_p \lim_{\Delta\tau \rightarrow 0} \frac{\Delta x^\lambda}{\Delta\tau} \\ &= \lim_{\Delta\tau \rightarrow 0} \frac{v^\mu|_q - (v^\mu|_p - \Gamma_{\lambda\nu}^\mu|_p v^\nu|_p \Delta x^\lambda)}{\Delta\tau}, \end{aligned}$$

可见 $Dv^\mu/d\tau$ 是求 $v^\mu|_q$ 与 q 处一个暂不明意义的新矢量之差. 但其又是从 p 衍生的, 即以某种方式移至 q , 我们可记

$$v^\mu|_{p \rightarrow q} = v^\mu|_p - \Gamma_{\lambda\nu}^\mu|_p v^\nu|_p \Delta x^\lambda. \quad (2.4.17)$$

这样 Dv 可表为二者箭端连线, 如图中蓝线. 图中所标棕线为 $v|_{p \rightarrow q}, v|_p$ 之差 δv . 由 (2.4.17) 或图示知 $\delta v^\mu = dv^\mu - Dv^\mu = -\Gamma_{\sigma\nu}^\mu v^\sigma dx^\nu$, 这便是克氏符项所给出的修正. 那 $v^\mu|_{p \rightarrow q}$ 究竟代表怎样一种移动? 令 v 自己就处处按这种方式移动, 即 Dv 始终为零.

取仿射参数 λ , 可发现测地线上的切矢 $T^\mu = dx^\mu/d\lambda$ 满足

$$\frac{DT^\mu}{d\lambda} = T^\nu \nabla_\nu T^\mu = 0, \quad (2.4.18)$$

故切矢是平行移动的. 换句话说, 矢量沿自身平行移动所得轨迹就是测地线.

Fermi-Walker 联络路径及附近

在平直背景上, 平行移动就是平移, 因此平行移动与路径无关. 但在弯曲背景上亦是这样吗? 我们

考虑一个平直的扇形, 其可围成圆锥面, 尖端就是.

此外平行移动还有如下直觉: 若能将弯曲背景上含有路径的部分“撕下”并, 便会发现矢量的确按平直空间的方式平移. 由于这与“撕下”背景的哪一块儿有关.

η 场就是存在整体坐标系使其分量处处为 $\eta_{\mu\nu}$, 显然导致相应的克氏符处处为零, 从而按式 (??) 使曲率分量处处为零. 可见 η 场的确等价于曲率处处为零.

2.5 微分形式

2.6 流形的映射

定义流形间映射便可谈及微分同胚. 微分同胚映射可自然诱导张量丛的映射, 故可谈及张量变换的主动、被动观点. 所有微分同胚构成微分同胚群. 任意光滑矢量场是该群的生成元. 显然微分同胚有无穷多生成元, 故维数无穷. 一个光滑矢量场在群中挑出一个单参子群, 称为单参微分同胚局部群 (无穷小微分同胚). 若光滑矢量场还有完备性, 则给出一个单参微分同胚群. 紧致流形上任意矢量场可延拓至完备的. 保持度规分量不变的变换对应的微分同胚映射称为等度规映射. 诱导此映射的矢量场称为 Killing 矢量. 代数定义可从 Lie 导数出发, 也可从 Killing 方程出发以跳过 Lie 导数. 此定义便告知可用 Killing 矢量场讨论度规 (时空) 的对称性.

2.6.1 Lie 导数

微分同胚

2.6.2 Killing 矢量

2.7 纤维丛

2.8 Frobenius 定理

2.9 参考系

就某一时空点而言可设置 4-速为 Z^μ 的观者, 称为**瞬时观者**. 若瞬时观者与观测对象在该处的 4-速 U^μ 相同, 则称它是观测对象的**瞬时固有观者**. 这样, 设荷电系统的运动由 U^μ 描述, 瞬时共动观者在其中某处测到的电荷密度 ρ_{e0} 就是**固有电荷密度**. 处处设置瞬时共动观者就形成了**固有系**, 则固有电荷密度构成一个标量场.

$U_\mu Z^\mu = g_{00} U^0 Z^0$. 为保持模方 -1 , 在此坐标系下有 $Z^0 = 1/\sqrt{-g_{00}}$.

的 3-速差距可通过

其等价定义是

$$J^\mu = \rho_{e0} U^\mu. \quad (2.9.1)$$

以瞬时观者为参考.

尺缩效应, 有 $\rho_{e0} = \rho_e / \gamma$, 其中 $\gamma = U^0 = -U_\mu Z^\mu$.

以后会证明, g 可在任意世界线上表为 $\eta_{\alpha\beta}$, 这种坐标系准确称为固有坐标系, 但若还要在该系下满足 $\partial_\lambda \eta_{\alpha\beta} = 0$ (进而 $\Gamma_{\mu\nu}^\lambda = 0$), 则世界线一定是测地线, 且固有坐标系无自转.

$$U^0 = \frac{1}{\sqrt{-g_{00}}} \approx 1, \quad (2.9.2)$$

其中系数是为使 $g_{00} U^0 U^0 = -1$

2.9.1 典型运动效应

狭相所有熟知的运动学效应均可从以下讨论定量地获得. 设 θ_1 和 θ_2 是两个快度, 考虑 boost 变换 $\Lambda(\theta_1)$ 和 $\Lambda(\theta_2)$. 根据双曲函数的性质, 有 $\Lambda(\theta_1)\Lambda(\theta_2) = \Lambda(\theta_1 + \theta_2)$. 代入 $v_i = \tanh \theta_i$, 以及 $v = \tanh(\theta_1 + \theta_2)$, 那么根据 $\tanh \theta$ 求和公式有

$$v = \frac{v_1 + v_2}{1 + v_1 v_2}. \quad (2.9.3)$$

其物理解释很简单. 设存在 3 个惯性观者, 其空间轴的设置如前文所述. 若第二个观者相对于第一个的速度是 v_1 , 而第三个相对于第二个的速度是 v_2 , 那第三个相对于第一个的速度就不是经典力学的线性叠加所给出的 $v_1 + v_2$, 而是 $(v_1 + v_2)/(1 + v_1 v_2)$. 这就是相对论的速度叠加公式.

考虑静止在 S 系中的原点, 其世界线上的两个事件, 由同一时

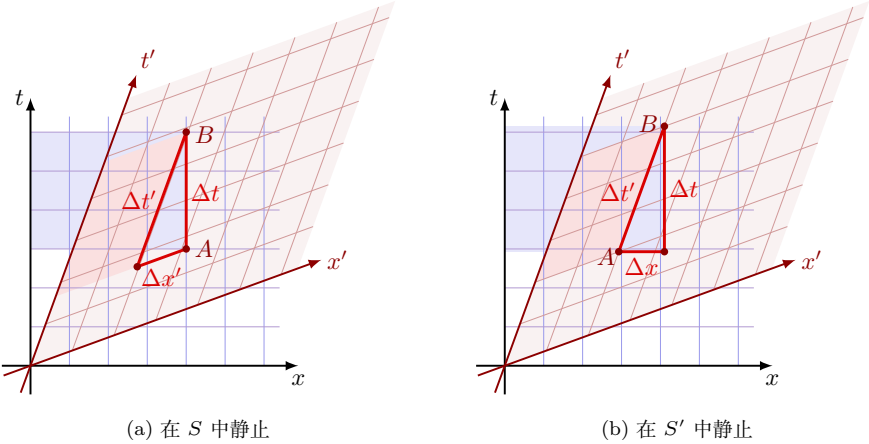


图 2.4: 钟慢效应

钟分别测得两个读数. 这种情况 $\Delta x = 0$, 所以

$$\Delta t' = \gamma \Delta t > \Delta t. \tag{2.9.4}$$

这种效应完全对称, 因为如果 $\Delta x' = 0$, 那么 boost 给出 $\Delta t = \gamma \Delta t' > \Delta t'$. 每个观者都认为其他的相对运动的时钟走时慢. 这种现象称为钟慢效应或时间膨胀. 这在几何上用校准线便一目了然. 例如, 两个观者都看到 S 系原点上的时钟读数“0”, 但直线 $t' = t_1$ 与该时钟世界线 (t 轴) 在 $(t, x) = (t_1, 0)$ 下方的 $(t_1/\gamma, 0)$ 相交. 须强调, 这种现象的物理意义是完全真实的. 例如, 在宇宙射线内发现的介子的寿命是如此之短, 以至于即使以光速运动, 它们穿透大气层所需时间也要比其通常寿命大几十倍. 按理来说, 其不可能到达地面, 但其实还是可以到达, 因为钟慢效应“使其保持年轻”, 即在我们看来, 其实际寿命比通常寿命长很多.

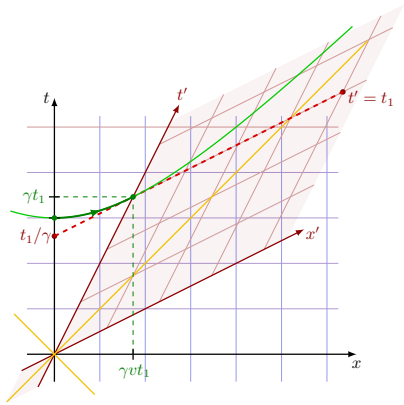


图 2.5: S' 系的钟根据同时面 $t' = t_1, t' = 0$ 认为 S 的钟慢

由于不同的惯性观者一般来说, 对事件的同时性各抒己见, 而测量一运动物体“长度”的唯一方法, 就是要利用一把量尺“同时”测量物体两端; 所以我们不应当诧异, 不同的惯性观者所测量的长度和时间的确会不同. 考虑一把沿 S' 系 x' 轴静置的测量尺子. 在这个坐标系中它的“长度”是 $\Delta x'$. 尺子两端的世界线是平行于 t' 轴的两条直线. 若 S 观者在这些世界线上同时观测两个事件, 那么他们的坐标将满足 $\Delta t = 0$, 且由 boost 有 $\Delta x' = \gamma \Delta x$, 则

$$\Delta x = \Delta x' / \gamma < \Delta x', \tag{2.9.5}$$

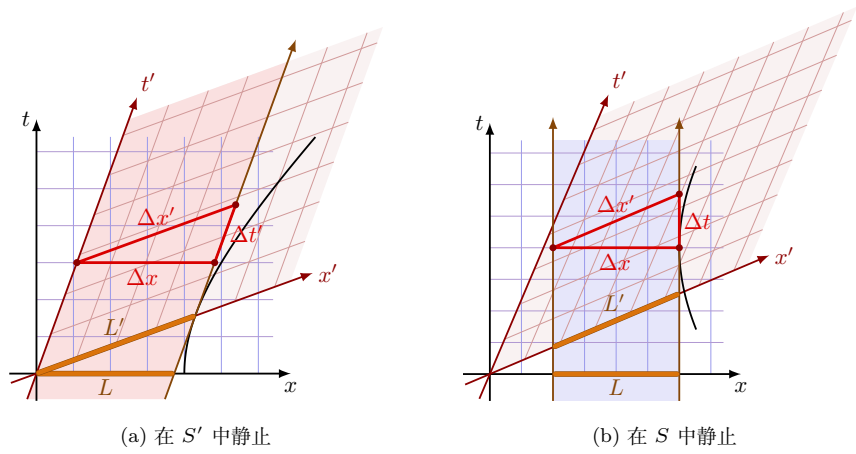


图 2.6: 尺缩效应

即量尺在其运动方向上缩短了 $1/\gamma$ 倍. 这种现象称为尺缩效应或长度收缩. 就像钟慢效应那样, 尺缩效应完全对称且实在. 沿 S 系 x 轴静置尺子的情况由上图道尽.

尺缩效应导致了一个有趣的车库佯谬¹. 设汽车与车库静长相等. 汽车朝着库匀速前进, 司机认为“动库变短, 不能放下”, 司库认为“动车收缩, 放下有余”. 二人矛盾吗? 为画出直观的时空图, 设车库并无后墙 (其“后墙”类似于隧道洞口). 注意, 画图时可借用校准曲线以保证车和库有相等静长. 由图易见, 以司库所在惯性系的同时面衡量, 车短于库; 以司机所在惯性系的同时面衡量, 车长于库. 两人同时面看法都对, 关键是同时性的相对性导致结论的相对性. “到底放下还是放不下?” 这种绝对式的问题没有意义, 正如在尺缩问题中“到底哪把尺子较长”亦无意义. 我们若真要“比较”, 能让各位意见一致的做法只能是都放入静系中去. 当车库有坚硬后墙时, 车头固然要撞墙, 因而停止. 然而撞墙信息 (以相互作用纵波) 传至车身各部分乃至车尾皆需时间, 只当车尾获悉后整个车身才能停住, 因而汽车将被压

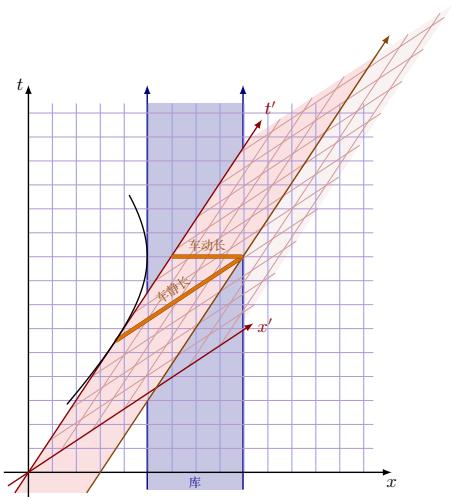


图 2.7: 汽车匀速进库

¹ 在英语世界里常称 ladder paradox, 因为可借西式农耕文化里的梯子 (ladder) 和谷仓 (barn) 表述. 一个经典的表述是用火车和隧道, 这都大同小异.

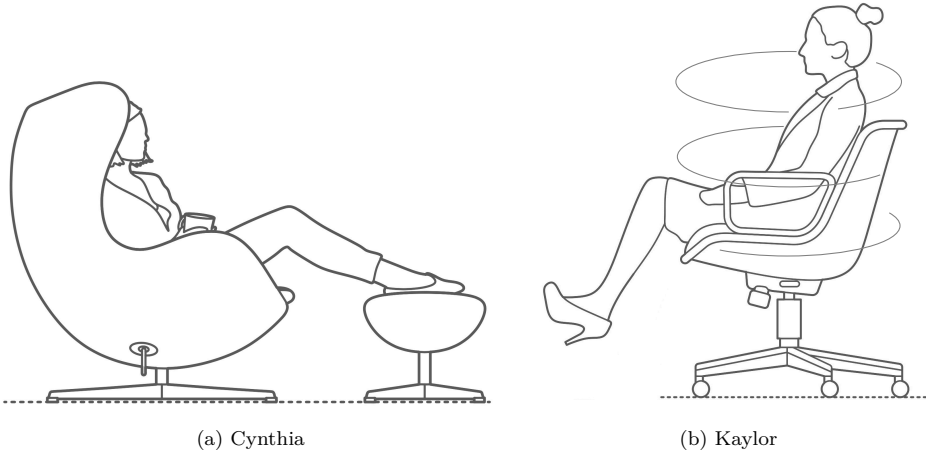


图 2.9: Cynthia 可视为惯性观者而 Kaylor 不能

缩到比司库期初所认为的长度还要短的程度. 这时, 车的确在库中装下有余, 谁看都装得下. 假设材料性质理想, 使得信息传递为光速 (在谁看来都是) 且质点获悉信号后立即停止, 有兴趣的读者不妨在图上补出撞车后的时空图.

2.9.2 自转

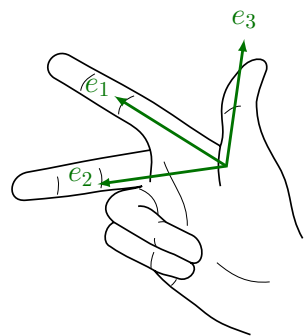


图 2.8: 空间右手 3-标架

还需处处配以 4-标架 (tetrad), 也即其上任意点矢量的正交归一基 $\{e_\mu\}$, 满足 $e_\mu \cdot e_\nu = \eta_{\mu\nu}$. 直观上, 3-标架 (frame) 可以是由三根单位长的短直杆或刻度尺焊成的正交架子, 每根直杆代表一个观测方向, 其在每一时刻的指向由该观者选定. 读者亦可尝试用右手大拇指、食指及中指按右手螺旋比划出 3-标架. 数学上, 3-标架就是空间正交归一基 $\{e_i\}$, “空间”指皆与观者世界线切矢 e_0 正交. 今后谈及 4-标架时默认为右手标架.

因而 4-标架场又称作观者所处处配备的“局部实验室”, 所测物理量无非就是将相关的 4 维量投影到 3-标架上.

惯性观者的 4-标架又称惯性基矢或惯性标架. 我们还应希望惯性标架无自转, 否则可能观测到赝力. 如图 2.9, 设 Cynthia、Kaylor 两人坐在地面的两把椅子上: Cynthia 坐底座固定的蛋壳椅; Kaylor 坐底座静置的办公椅且不停自转. 二者世界线都走直线, 但 Cynthia 可视为惯性观者而 Kaylor 不能. 请注意, 虽然观者概念本

身已要求把此二人看成没有大小的点 (于是由一世界线代表), 但自转涉及的仅是线上各点每一空间基矢的方向沿线是否改变, 故仍有明确意义. 然而在惯性观者所对应的惯性坐标系中, 若令 4-标架 $\{e_\mu\}$ 就是相应的坐标基, 则直观看来就是无自转的. 一般默认将 4-标架取为惯性坐标基. 借助惯性坐标系便可定量地定义无自转观者.

无自转

Fermi-Walker 移动

瞬时观测

测地线切矢就是时间轴的方向, 那么与之正交就代表着空间方向,

动力学

2.9.3 膨胀、剪切与扭转

2.9.4 刚性系空间几何

本节定量讨论 Einstein 转盘作为非惯性参考系的性质.

可见正是转盘的扭转导致转盘系不存在处处同观者正交的同时面. 下面定量说明转盘的扭转非零. 取惯性坐标系: 坐标时 t 、空间柱坐标 r, θ, z , 使盘面位于平面 $z = 0$, 盘心 r 为零. 设 ω 为转盘角速率, 转盘系 s 中任一观者世界线以 τ 为固有时参数, 默认所有观者标准钟零点设置保证 $t = 0$ 时 $\tau = 0$, 则世界线可表为

$$t = \gamma\tau, \quad r = \text{常数}, \quad \theta = \gamma\omega\tau, \quad z = \text{常数}. \quad (2.9.6)$$

其中 $\gamma = (1 - \omega^2 r^2)^{-1/2}$. 只当 r 满足 $\omega r < 1$ 时上述曲线才能表示类时观者, 故取转盘系定义域为 \mathbb{R}^{3+1} 中满足 $\omega r < 1$ 的开集. 求导知 4-速表达式为

$$Z = \frac{\partial}{\partial \tau} = \gamma \frac{\partial}{\partial t} + \gamma\omega \frac{\partial}{\partial \theta}. \quad (2.9.7)$$

闵氏度规按柱坐标表达 (只需把平面坐标部分替换为极坐标), 则线元表为

$$ds^2 = -dt^2 + dr^2 + r^2 d\theta^2 + dz^2. \quad (2.9.8)$$

故

$$Z_b = -\gamma dt + \gamma\omega r^2 d\theta. \quad (2.9.9)$$

对 t, r, θ, z 按 0, 1, 2, 3 标号, 可从线元或度规求得非零克氏符有 $\Gamma^1_{22} = -r, \Gamma^2_{12} = \Gamma^2_{21} = 1/r$. 由此得非零的 $\nabla_\mu Z_\nu$ (记作 $Z_{\nu;\mu}$ 便于书写) 有

$$Z_{0;1} = -\gamma^3 \omega^2 r, \quad Z_{1;2} = -\gamma\omega r \quad Z_{2;1} = \gamma^3 \omega r.$$

故

$$\nabla Z = -\gamma^3 \omega^2 r dr \otimes dt - \gamma\omega r d\theta \otimes dr + \gamma^3 \omega r dr \otimes d\theta.$$

综上可知 $A = \nabla_Z Z = -\gamma^2 \omega^2 r \, dt$, 因此

$$A \otimes Z = \gamma^3 \omega^2 r \, dr \otimes dt - \gamma^3 \omega^3 r^3 \, dr \otimes d\theta.$$

由 $B_{\mu\nu} = \nabla_\nu Z_\mu + A_\mu Z_\nu$ 知

$$B = 2\gamma^3 \omega^2 r A(dr \otimes dt) - 2\gamma^3 \omega r A(dr \otimes d\theta) = AB.$$

故 $\omega_{\mu\nu} = B_{\mu\nu} \neq 0$, 除非角速率为零.

转盘系的非超曲面正交性使其空间几何的定量研究难以进行. 幸好, 对转盘系而言, 这一困难可借其另一性质克服: 以上讨论亦表明 $\theta_{\mu\nu} = 0$, 而无膨胀就是刚性.

直观上转盘系确实是个刚体, 因此即使转动非零, 刚性性仍可帮助我们研究空间几何.

2.9.5 对钟

2.10 标架表述

2.11 哈氏理论

2.11.1 3+1 分解

2.11.2 初值问题

正则

Chapter 3

黑洞物理

3.1 黑洞

式 (??) 仅在 $r = 2M$ 发散, 说明当 $r > 2M$ 时观者可静止于时空中, 换句话说, 这里的 t 轴曲线是类时的, 因为其可作为一个静态观者的世界线. 而在 $r < 2M$ 中开方会出现虚数, 不能定义模长, 一定程度上意味着在 $r < 2M$ 中无法静止. 这样曲面 $r = 2M$ 又称静界 (*static surface*). 可见这个曲面值得进一步研究. 它是等 r 面, 即取 $f = x^1 = r$, 则法矢 $n^\mu = \partial^\mu f$ 的模方为

$$n^\mu n_\mu = g^{\mu\nu} \delta_\nu^1 \delta_\mu^1 = g^{11} = 1/g_{11} = 1 - 2M/r,$$

可见 $r > 2M$ 时 n^μ 类空, 与其正交且的未来切矢必类时, 故称超曲面 $r = C$ 类时; $r < 2M$ 时 n^μ 类时, 同理称 $r = C$ 类空; 而在视界 $r = 2M$ 上, n^μ 类光, 称 $r = C$ 类光. 其物理意义是, $r > 2M$ 的类时超曲面上与光锥相交, 这说明光可自由穿梭; 然而 $r < 2M$ 的类空超曲面位于光锥之外, $r = 2M$ 的类光超曲面刚好切于光锥, 故光只能进入而不能逃脱. 可见, $r = 2M$ 就仿佛是由光锥面构成的, 一旦进入, 连光都至多只能沿着 $r = 2M$ 传播, 任何质点的宿命皆在光锥之内. 可设想如下物理情形: 存在一个球形星体, 或许其密度实在过大, 以至于很大的质量集中于很小的体积内, 且这个球体半径满足 $r < 2M$. 其地表的任何光线无法逃出曲面 $r = 2M$. 由于看不到它的本貌, 因而称之为黑洞 (*black hole*), 这样曲面 $r = 2M$ 又称为事件视界, 简称视界, 因为它代表着一种屏障. 视界所包围的时空区域称为黑洞区域, 进而视界又可称为黑洞表面. 注意这并非星体表面, 因为视界内可以极为空旷, 毕竟质点在 $r < 2M$ 内不能静止, 如果这个星体不够坚固, 极为可能坍缩, 甚至近乎成为一点 $r = 0$. 如此, 对于黑洞而言, 一般可不区分球体和质点. 假设有一闯入 $r < 2M$ 的探险者, 对其而言, 外部空间于其而言仿若永不能逾越的时间上的过去, 而内部空



图 3.1: 电影 *Interstellar* 中黑洞 Gargantua 被设定为一颗绕有吸积盘 (accretion disk) 的巨型旋转黑洞. 黑洞背对观众处的吸积盘所发的光受引力影响, 而绕过了黑洞并射向观众, 故其视觉上仿佛包裹了视界.

间乃至 $r = 0$ 处仿若必将走向的时间上的未来. 这在度规上体现为时间分量、半径分量会变号, 而 $ds^2 = -dr^2$ 告诉我们只有系数为负的那一项具有坐标时的意义, 说明探险者对时间和空间的感知仿若互换角色, 空间表现出了类似时间箭头的特征.

广相有三大经典验证, 我们将依次介绍.

星光偏折 Eddington(1882-1944, 英国天体物理学家) 在 1919 年的一次日食观测中摄下了光线被引力“弯曲”的图片. 这一点已经在几何光学近似中从理论上计算过. 从那以后, 用太阳系的种种检验对场方程作了所谓的后 Newton 理论 (post-Newtonian theory), 这些预测和实验都在这个物理条件下以很高的精度证实了广相.

如果我们考虑黑洞附近的光线, 一路走向测地线方程, 所描述的轨迹曲线可以作为引力透镜 (gravitational lensing) 的定量解释.

引力红移

奇点

Schwarzschild 度规的一个重要的“毛病”就是奇性 (singularity). 所谓奇性, 通常是指一些数值的发散, 比如度规分量, 包括度规逆矩阵分量的发散 (即度规行列式 g 为零). 该度规在 $r = 0$ 处时间分量就发散, 当然, 凡同 $\phi = -M/r$ 牵扯的引力场都躲不过这一点. 其实本应不予关心, 毕竟总要回到 Newton 引力, 不使用反比关系或许过于严苛; 并且我们在科学素养中, 就专门区分过质点理想模型和物理客体的区别, 因此对 $r = 0$ 的讨论将无意义. 然而在广相中, 情况略微不同, 即这里 $r = 0$ 或许还真与物理客体有所关联. 经过冗长的计算后 (但思路简洁, 读者可按部就班),

可得到如下的 *Kretschmann* 标量

$$R_{\mu\nu\sigma\lambda}R^{\mu\nu\sigma\lambda} = \frac{48M^2}{r^6}, \quad (3.1.1)$$

上式在 $r = 0$ 发散. 这问题其实很严重, 因为广相的主要研究对象是连续流体, 然而我们只是框定一堆物质, 便使得光滑的时空上戳出一个“洞”, 这是很难接受的. 其次, 空间分量为倒数导致另一奇性: g_{11} 在 $r = 2M$ 发散. 但两处的奇性有区别: 上式在 $r = 2M$ 有意义. 注意上式是由张量构造的标量, 其在坐标变换下不变, 因此该结论不受坐标系影响. 既然曲率在 $r = 2M$ 处并不奇异, 那我们就有可能找到一种坐标系使得度规分量不再发散. 该过程一般称为“消除”度规奇性, 因而将度规分量在视界处的这种奇性称为坐标奇性或假奇性. 我们在第 ?? 章中讨论上述概念的微妙之处. 相较而言, $r = 0$ 处的奇性就称为时空奇性或真奇性, 因为曲率发散的事实与参考系无关. 这个点就叫做奇点¹.

我们用时空图来理解上述奇性. 视界在观感上是一个 3 维空间里的 2 维球面, 但在时空图里, 它应沿时间方向延伸为一张 3 维类时超曲面. 若将时空图绘制为 3 维的, 则可以用一个 2 维圆柱面代表视界. 若在时空图中标注各点微小光锥的方向, 就能将各点时空结构直观地呈现出来, 因此欲求出时空中的光线轨迹. 光没有质量而走类光世界线, 因此我们不能定义固有时. 实际上, 描述光线时一般取任意仿射参数 λ . 比如, 考虑 \mathbb{R}^{3+1} 中一点 x_0 , 此处一个指向未来的类光矢量 N 就可代表光的走向. 由 $\alpha(\lambda) = x_0 + \lambda N$ 定义的直线 α 就是自由光子经过 x_0 的世界线. 可以预料, 上述直线推至一般坐标系甚至一般时空时, 应得仿射参数化的测地线方程, 对 $ds^2 = 0$ 求导很容易证明这一点. 实际上, 正因为用间隔不变形可得测地线, 我们在求某时空的类光测地线时, 总是直接处理度规即可, 这样便省去克氏符的计算. 考虑到球对称性, 我们只用 2 个维度

$$0 = ds^2 = -(1 - 2M/r) dt^2 + (1 - 2M/r)^{-1} dr^2,$$

故曲线斜率满足

$$\left(\frac{dt}{dr}\right)^2 = \frac{1}{(1 - 2M/r)^2}.$$

积分得²

$$t = \pm (r + 2M \ln |r - 2M|) + C. \quad (3.1.2)$$

接着要在图中画出上述类光测地线. 图上每一点都可确定两条不同的类光测地线, 分别属于两族 (*family*), 因此可均匀地取积分常数以绘制网格. 每一条类光测地线都代

¹其英文亦是 singularity. 为了表述性质和地点两个概念, 中文语境里可区分使用.

²严格来讲, $1/x, x \neq 0$ 的不定积分应按定义域分别写出两个可以不同的积分常数, 而非直接写 $\ln |x| + C$, 但物理学一般偏好用绝对值形式, 以尽可能简化研究对象.

表光路, 因而每一点的光锥面必定切于两族在此处的类光测地线, 因此不妨在每个格点都画上一个椭圆, 构成一个小“妙脆角”, 以代表微小未来光锥 (完整光锥用小“沙漏”表示), 用以揭示该处的时空结构. 所有类光测地线可分为两类. 作为未来光锥右侧的称为外向 (*outgoing*) 族, 反之为内向 (*incoming*) 族. 比如, 注意 $r > 2M$ 时, (3.1.2) 中正负号取正后斜率为正, 只能作为未来光锥右侧, 故为外向族. 其他情况以此类推.

如图所示, 圆柱面为事件视界, 波浪线为奇点. 可见直接用 Schwarzschild 系使视界出现奇性, 生硬地割裂了时空的内外部分. 取一个 $r > 2M$ 的静态观者, 这样由 $ds^2 = -d\tau^2$ 知其固有时与坐标时的关系¹为

$$d\tau = \sqrt{1 - 2M/r} dt, \quad (3.1.3)$$

这说明, 静态观者借自己的标准钟观察一个探险者, 会发现其落入 $r = 2M$ 的过程逐渐缓慢, 因为

$$\frac{dt}{d\tau} = \frac{1}{\sqrt{1 - 2M/r}}$$

会从 1 逐渐递增至 $r = 2M$ 处发散, 进而其在 $r = 2M$ 附近“毫无动静”, 仿佛时间停滞流逝, 最终将永远定格于 $r = 2M$ 而不会坠入视界. 然而在探险者看来, 他进入视界“确实”是顺理成章的. $r > 2M$ 的部分让人误解为任何质点永远位于视界之外, 因为我们容易把 $t \rightarrow \infty$ 混同于“永远”. 古希腊时期有一著名的 Zeno 悖论讨论过类似效应. Zeno 为讨论无穷的含义设计了这样的情形: 一位运动员 Achilles 欲同一只乌龟 (tortoise) 赛跑, 且乌龟起点先于他若干路程 L . 按常理来说, 他的速度 v_1 快于乌龟的 v_2 , 明显会在有限时间 $L/(v_1 - v_2)$ 内超过乌龟. 但设想这样计时: 他欲超过乌龟, 必须先追至乌龟的起点, 但一旦如此, 乌龟又能前进若干步, 这样便有新的起点等着他, 如此这般的有无穷多这样的起点, 则他“永远”也赶不上乌龟. 实际上这并不存在孰是孰非的问题, 因为答案是相对于所选取的计时方式而言的. 乌龟均匀向前, 它未“想得”那么复杂, 因此可将“乌龟时”理解为和观众一致认同的普通时 t . Achilles 时 (或称 Zeno 时) 记作 t' , 两种时间皆起于零. 当 $t = L/v_1$ 时, Achilles 到达乌龟在 $t' = 0$ 时的起点, 则 t' 加一, 以此类推有

$$t = \sum_{i=0}^{t'-1} \frac{L}{v_1} \left(\frac{v_2}{v_1} \right)^i = \frac{L}{v_1 - v_2} \left(1 - \left(\frac{v_2}{v_1} \right)^{t'} \right),$$

¹至此可重新观察 Schwarzschild 度规, 若类比钟慢和尺缩效应, 我们可将 $1/\sqrt{1 - 2M/r}$ 理解为 γ , 而这种速度就是 $v = \sqrt{2M/r}$. Newton 理论能给予一个直观理解. 物体从引力场某处逃逸的条件是有足够的动能, 使得在力场范围内都不为零, 这样一旦越过范围就可用剩余动能逃走. 最坏的情况是力场范围延至无穷远, 则物体动能要足够大, 以至于最差只能在无穷远处为零, 这样由能量守恒推知速度至少为 $\sqrt{2M/r}$. 1799 年, Laplace 假设引力可以吸引光, 则对光而言逃脱条件便是 $r > 2M$, 因此从内部发出的光无法到达无穷远处. 这当然不严谨, 因为真空光速参考系不变; 不能定义光的质量, 故光受引力的影响只能由广相解释; 内部光线只是无法到达无穷远, 仍可超出 $r = 2M$ 而至有限远, 这与黑洞有本质不同. 用 γ “推导”引力度规的方式可见文献 [21].

上式自然将 t' 从 \mathbb{N} 延拓至 \mathbb{R}_+ , 因此从 Achilles 时 t 到乌龟时 t' 的变换为

$$t' = \frac{1}{\ln(v_2/v_1)} \ln \left(1 - \frac{v_2 - v_1}{L} t \right),$$

可见二者走势一致, 区别在于 t' 会在 $t = L/(v_2 - v_1)$ 发散.

这给了我们寻找新坐标的启发. 我们亦可给视界赋予发散的坐标, 这样就可将度规的奇性隐藏在坐标变换中. 先想办法消除度规中的因子 $(1 - 2M/r)^{-1}$. 常常采取因式分解, 即

$$ds^2 = (1 - 2M/r) \left(-dt^2 + (1 - 2M/r)^{-2} dr^2 \right),$$

定义一个新半径使得

$$dr_* = \frac{dr}{1 - 2M/r} \Rightarrow r_* = r + 2M \ln \left| \frac{r}{2M} - 1 \right|, \quad (3.1.4)$$

这里取积分常数使得 $r = 0$ 时 $r_* = 0$, 可见其与 r 走势一致而仅在 $r = 2M$ 处发散. 类比 Zeno 悖论, 可称乌龟坐标. 度规就可表为

$$ds^2 = (1 - 2M/r) (-dt^2 + dr_*^2),$$

这样就消除了发散因子. 然而这个变换并不够, 因为其行列式 g 在 $r = 2M$ 为零, 意味着逆矩阵元发散. 但是这种变换很常用, 值得单独命名并继续使用. 对物理量 x , 变换 $x' = x + (r_* - r)$ 在 $r < 2M$ 时 $x' < x$, 故称超前 (*advanced*) 变换, 而 $x' = x - (r_* - r)$ 在 $r < 2M$ 时 $x' > x$, 故称推迟 (*retarded*) 变换. 考虑到奇性导致了类似于 Zeno 悖论的现象, 应将坐标时 t 超前变换至 \tilde{t} , 则度规表为

$$ds^2 = - \left(1 - \frac{2M}{r} \right) d\tilde{t}^2 + \frac{4M}{r} d\tilde{t} dr + \left(1 + \frac{2M}{r} \right) dr^2, \quad (3.1.5)$$

便可使时空交叉项为正, 符合黑洞演化的方向. 上式在 $r = 2M$ 处所有分量皆收敛, 而行列式 $g = -(g_{01})^2 \neq 0$, 这才算消去奇性. 该坐标系称为超前 Eddington 系. 利用十字相乘法可分解出两个类光测地线族, 其一是

$$\frac{d\tilde{t}}{dr} = \frac{r + 2M}{r - 2M} \Rightarrow \tilde{t} = r + 4M \ln |r - 2M| + C_1, \quad r > 2M, \quad (3.1.6)$$

其二是

$$\frac{d\tilde{t}}{dr} = -1 \Rightarrow \tilde{t} = -r + C_2. \quad (3.1.7)$$

显然前者为外向族而后者为内向族. 进而可画得图 3.2. 其中设有一个静态观者和一个位于 \tilde{t} - r 平面的探险者. 可见视界外的质点可以穿越视界进入黑洞, 而一旦进入就无法逃逸, 并最终落入奇点. 探险者若想摆脱黑洞的引力束缚, 只能采用某种喷气火

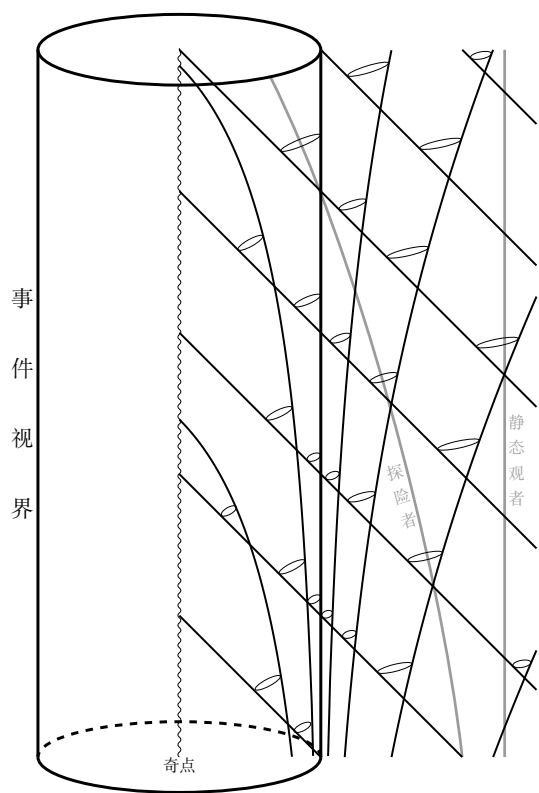


图 3.2: 静态黑洞时空图

箭，以大于该处静态观者所需力量向右飞行. 一旦燃料储备不足，便连静态观者也当不了. 假设其燃料充足，若探险者企图接近视界而迅速掉头离开，则越近风险越大，因为从 $r = 2M$ 附近逃离须使飞船接近光速. 并且，当视界内部很空旷时，对探险者而言，其进入视界时并无触觉，若稍有不慎，未能警觉自己所处位置，则很可能多走一步而遁入厄运. 当然，探险者还是能通过视觉上“眼前一黑”来观察，但更可怕的是，探险者并非真正的质点，而不同位置所受引力在此极端地带将差距巨大，因此潮汐力（对应于曲率大小）极有可能撕碎探险者，除非视界半径很大而稀释了曲率. 此外，前文用测量效应说明了静态观者对探险者的观察，亦可用视觉效应这种现象. 若静态观者借光信号观察探险者，则从图 3.2 容易看出，随探险者逐渐接近视界，探险者所发出的光将越来越晚地到达静态观者，而坠入的画面无法传播出来.

能描述 Schwarzschild 时空的坐标系有很多，超前 Eddington 系并非完美无缺. 比如，Schwarzschild 系能保证 $g_{01} = 0$ 而超前 Eddington 系不然. 可能存在一个“更好”的坐标系既能消除奇性，又能使 $g_{01} = 0$. 第 ?? 节就提过研究弯曲时空引力性

质的技巧,那就是到平直时空的加速度系去,而 Rindler 坐标系正好能满足 $g_{01} = 0$, 我们可从中获取灵感. 考虑到 Rindler 系的坐标具备时空的性质 (尤其是远处的加速系), 现在就将其改记 (t, x) . 不妨先讨论 $x > 0$, 反之情形同理. 原来的闵氏系改记 (T, X) . 这样闵氏度规 $ds^2 = -dT^2 + dX^2$ 在坐标变换 $X = x \cosh t, T = x \sinh t$ 下变为 “Rindler 度规” $ds^2 = -x^2 dt^2 + dx^2$, 二者无实质区别, 为同一度规. 现在我们不再按 T, X 正交来绘制时空图, 而是按 t, x 正交. 再次声明, 读者无需感到诧异, 同一时空由于所用坐标系不同, 时空图可千差万别. 同理求出内、外向族分别为

$$t = -\ln x + C_1, \quad t = \ln x + C_2.$$

故可绘制如下图 ?? . 注意, 闵氏度规在 Rindler 系是有奇性的, 因为其行列式 $g = -x^2$ 在 $x = 0$ 处为零. 可见, Rindler 系和闵氏系的相对地位, 就像是 Schwarzschild 系或 Eddington 系和 “更好” 坐标系的相对地位.

类光测地线意味着 $ds^2 = 0$, 而张量为零的事实不受坐标系影响, 因此上述类光测地线族其实 “对应于” 闵氏时空里的类光测地线族

$$T = -X + C_1, \quad T = X + C_2.$$

第 ?? 节中将双曲线、射线视作坐标线而衍生出 Rindler 坐标系, 仿照这一思想, 欲视类光测地线族为坐标网格 (称为类光坐标). 在 Rindler 图 ?? 里, 这是在令

$$v = t + \ln x, \quad u = t - \ln x. \quad (3.1.8)$$

度规表为 $ds^2 = -e^{v-u} dv du$. 而在闵氏图里是令

$$V = T + X, \quad U = T - X, \quad (3.1.9)$$

或

$$T = \frac{V+U}{2}, \quad X = \frac{V-U}{2}, \quad (3.1.10)$$

这有双曲函数的形式, 而 T, X 与 t, x 的变换正是双曲函数, 故易知如下指数式

$$V = e^v > 0, \quad U = -e^{-u} < 0. \quad (3.1.11)$$

度规现表为 $ds^2 = -dV dU$, 不存在任何坐标奇性. 如图 3.3 所示, 则从直观上, 闵氏时空图就像是把 Rindler 图 ?? 里竖直的 $x = 0$ 从 $t = 0$ 处向右 “掰弯” 成如光锥面一般的 Rindler 视界 $T^2 = X^2$, 即 $VU = 0$, 由上式知 V, U 轴正方向如图 3.3. 图中蓝色双曲线代表等 r 线, 红色射线代表等 t 线. 可见, 从 Rindler 图到闵氏图经历了如下过程:

$$(t, x) \rightarrow (v, u) \rightarrow (V, U) \rightarrow (T, X).$$

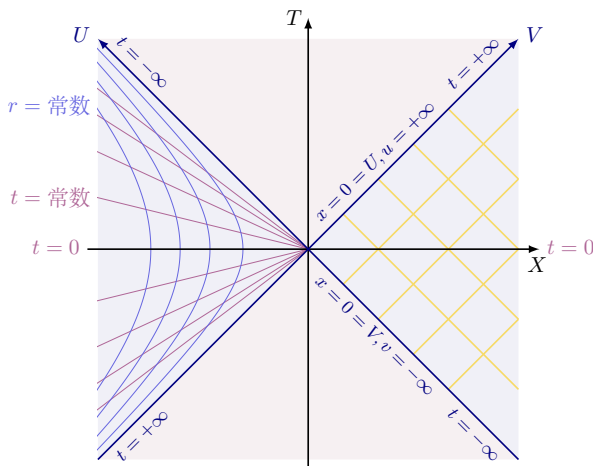


图 3.3: 消除 Rindler 系的坐标奇性

依葫芦画瓢地寻找“更好”坐标系. 在超前 Eddington 系中, 时间对角项在 $r = 2M$ 为零, 为消除奇性, 只能借由超前变换所产生的非零时空交叉项, 但我们并不想要交叉项. 可见, 必须使 T, X 坐标的时间对角项非零. 回到 Schwarzschild 系. 不妨先研究 $r > 2M$. v, u 从类光测地线族产生, 注意式 (3.1.2) 与 $t = \pm r_*$ 相差常数, 因此内、外向族分别给出

$$v = t + r_*, \quad u = t - r_*. \quad (3.1.12)$$

由于 r_* 作了超前变换, 因此前者称超前类光坐标, 后者称推迟类光坐标. 可见对 Rindler 时空来说其乌龟坐标是 $x_* = \ln x$, 读者很容易验证这一点. 度规在 v, u 下表为

$$ds^2 = -(1 - 2M/r)dvdu,$$

沿用 V, U 的指数形式但预留待定常数 (只可与 M 有关) 以使最终的时间对角项非零¹:

$$V = e^{\beta v} > 0, \quad U = -e^{-\beta u} < 0. \quad (3.1.13)$$

只预留一个是考虑到对称性. 进而度规表为

$$ds^2 = -\beta^{-2}(1 - 2M/r)e^{\beta(u-v)}dVdU,$$

考虑到 T, X 系只是 V, U 系的旋转, 我们可将其定义沿用至此, 则 $-dVdU = -dT^2 + dX^2$, 进而保证无交叉项. 现在的目标就是约掉 $(1 - 2M/r)$ 通分后所含因子 $(r - 2M)$. 这份任务只能交由 $e^{\beta(u-v)}$, 我们就要用 r 表示它 (由稳态条件可预料无 t). 注意

¹有种做法是填上 $1/\beta = 4M$ 因子以约去指数求导出的系数, 这样度规 (3.1.14) 中 $32M^3/r$ 变为 $2M/r$.

$u - v = -2r_*$, 代入 r_* 定义有

$$e^{\beta(u-v)} = e^{-2\beta r} \left(\frac{2M}{r-2M} \right)^{4\beta M},$$

说明应取 $\beta = 1/4M$, 则

$$ds^2 = -\frac{32M^3}{r} e^{-r/2M} dV dU = \frac{32M^3}{r} e^{-r/2M} (-dT^2 + dX^2).$$

这里仍留有一个不能消除的时空奇点 $r = 0$, 与 Rindler 情况有所不同. 综上, 还原另外两个空间维度后, Schwarzschild 度规就表为

$$ds^2 = \frac{32M^3}{r} e^{-r/2M} (-dT^2 + dX^2) + r^2 d\Omega^2, \quad (3.1.14)$$

这里 r 视作 T, X 的函数, 易知

$$X^2 - T^2 = -VU = e^{(v-u)/4M} = \left(\frac{r}{2M} - 1 \right) e^{r/2M}, \quad (3.1.15)$$

注意 $r > 0$, 我们可用数学中的 *Lambert* 函数

$$y = W(x) \iff x = ye^y, y \geq -1 \quad (3.1.16)$$

将其表为

$$r = 2M W \left(\frac{X^2 - T^2}{e} \right) + 2M. \quad (3.1.17)$$

这个坐标系称为 *Kruskal* 系. 我们再次按 T, X 正交绘制 2 维时空图. $(-dT^2 + dX^2)$ 形式的保留使类光测地线斜率仍为 ± 1 . 实际上, 像这样给原度规乘以一个正因子就称为 *Weyl* 重标度 (*rescaling*) 变换, 简称 *Weyl* 变换, 它能保留时空结构, 故又称为共形 (*conformal*) 变换. 从图 3.4 上可见, *Kruskal* 系也会“掰弯”Schwarzschild 时空的视界 $r = 2M$, 使之变为光锥面 $VU = 0$. 易知各坐标的取值分布同图 3.3 相似. 然而, 我们只考虑了 $r > 2M$ 的情况, 在图上标记为开区域 *A* 区, 对应 $V > 0, U < 0$. 那 $r < 2M$ 呢? Schwarzschild 系的 $r > 2M$ 和 $r < 2M$ 借超前 Eddington 系消除 $r = 2M$ 坐标奇性. 从图上看, 从 *A* 区任意点出发的内向类光线一定坠向 V 正半轴, 说明其代表事件视界, 故可记作 \mathcal{H}_V^+ . *Kruskal* 系也能消除坐标奇性, 我们可越过视界来到上方, 这里 $V, U > 0$, 说明我们自然地拓宽了 V, U 的定义. 为使 U 为正, 应定义

$$V = e^{v/4M} > 0, \quad U = e^{-u/4M} > 0, \quad (3.1.18)$$

可验证度规仍表为 (3.1.14) 式. 这里等 r 线应无限延伸, 因为等 t 线从 U 正半轴至 V 正半轴遍历了所有 t 值. 一系列等 r 线至多扫至时空奇点 $r = 0 \Leftrightarrow T^2 - X^2 = 1$ 而无法越过, 所扫区域里的任意类时、类光线确实都坠入奇点, 说明其代表黑洞, 称

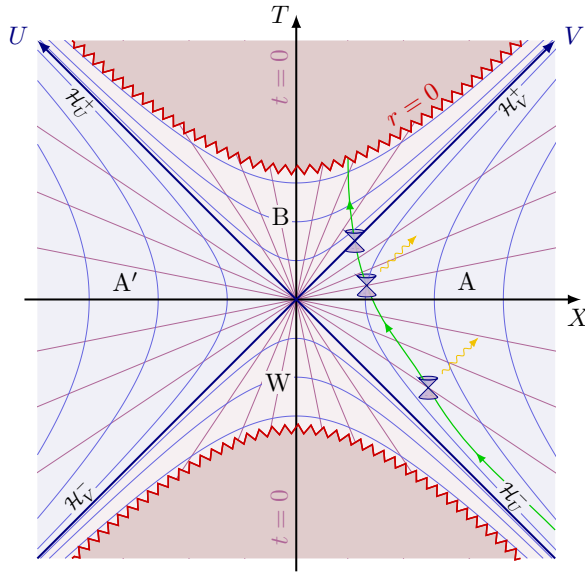


图 3.4: Kruskal 延拓

为 B 区. 可见, 超前 Eddington 系覆盖了 $A \cup \mathcal{H}_V^+ \cup B$, 即整个 Schwarzschild 时空. 奇点及其外部红色区域不属于时空. 综上, A, B 区的坐标变换可整合为

$$V = e^{(r_*+t)/4M}, \quad U = -\text{sgn}(r-2M)e^{(r_*-t)/4M}. \quad (3.1.19)$$

进而变换至 T, X 可整合为

$$\begin{aligned} T &= \frac{e^{r/4M}}{2} \sqrt{\left| \frac{r}{2M} - 1 \right|} (e^{t/4M} - \text{sgn}(r-2M)e^{t/4M}), \\ X &= \frac{e^{r/4M}}{2} \sqrt{\left| \frac{r}{2M} - 1 \right|} (e^{t/4M} + \text{sgn}(r-2M)e^{t/4M}). \end{aligned} \quad (3.1.20)$$

其仍形如双曲函数. 若从 T, X 做逆变换, 易知 r 仍满足 (3.1.17) 式, 而 t 表为

$$t = 2M \ln \frac{X+T}{|X-T|}, \quad (3.1.21)$$

形如反双曲正切函数.

3.2 引力坍缩

星体演化

这些问题花了很长时间才理清, 因为整体微分几何的流形语言是在场方程之后才发展完整的.

人们现在普遍认为星体引力坍缩 (*gravitational collapse*) 的末日归宿就是黑洞。

进入黑洞的测地线会在有限固有时内碰上奇点而“断掉”，不可无限地延伸下去，称作不可延测地线 (*inextendible geodesic*)。接近奇点的宏观观者会被差距强烈的潮汐力撕碎。

这种病态性质与高度的球对称性有关。Penrose 在 1965 年给出的奇性定理 (*singularity theorem*) 中指出，耦合适当的物质并附加某些条件后，一定会出现不可延测地线，因而就产生了这么个奇点。奇性定理给出了两个重要猜想 (conjecture) 或者假设 (hypothesis)。第一个是弱宇宙监督 (*weak cosmic censorship*) 假设：对于适当的物质方程组和一般物理意义的初始条件，奇点一定限制在黑洞区域内而“看不见”；第二个是强 (*strong*) 宇宙监督假设：解的奇性一定与其延拓遇到局部阻碍有关。后一个假设保证了动力学问题的唯一解必定是初始数据产生的经典时空，即经典决定性原理 (*Classical Newtonian determinism*)。如果抛弃条件的一般性，则这两个假设都不成立。Christodoulou 做出过标量场方程组的球对称解，其中的奇点就不在黑洞区域内，这样的时空就称作包含了裸 (*naked*) 奇点。裸奇点很容易构造出来，比如直接令 $m < 0$ ，但此时就不允许渐近平直 (*asymptotically flat*) 的 *Cauchy* 面了，因而可能不太符合咱们的物理常识。这件事与所谓的正能量定理 (*positive mass theorem*) 有关。

3.3 Penrose 图

3.4 时空延拓

3.5 嵌入图与虫洞

绘制虫洞的一个形象方法就是用嵌入图。

3.6 RN 黑洞

3.7 黑洞热力学

KN 黑洞

黑洞无毛猜想

3.8 dS 与 AdS 时空

Chapter 4

因果结构

4.1 可定向性

4.2 因果线

4.3 因果条件

接下来介绍线汇 3-速的三类变化 (膨胀、扭转和剪切), 这需要将 $\vec{u} = \nabla_{\vec{\eta}} Z$ 分解为三个有其物理意义的项. 为此, 先将 ∇Z 投影成 “空间矢量场” B , 即

$$B = C_4^1 C_3^2 (h \otimes h \otimes \nabla Z_b), \quad (4.3.1)$$

其满足 $C_1^1(Z \otimes B) = C_2^1(Z \otimes B) = 0$. 显然 B 作为 $(1, 1)$ 型张量时就是从 3-矢 $\vec{\eta}$ 到 3-矢 \vec{u} 的线性映射. 下面为行文简洁, 时常采取分量写法.

定理 4.3.1. $u^\mu = B^\mu{}_\nu \eta^\nu$.

证. 只需证明 $\nabla_\nu Z_\mu = B_{\mu\nu} - A_\mu Z_\nu$. □

将其对称和反称部分分别记作

$$\theta_{\mu\nu} = B_{(\mu\nu)}, \quad \omega_{\mu\nu} = B_{[\mu\nu]}.$$

由反称性知 ω 的任意分量矩阵的对角元为零, 当然无迹. 但 $\theta_{\mu\nu}$ 还可分解为有迹和无迹两部分. 以 θ 代表其迹, $\sigma_{\mu\nu}$ 代表无迹部分. 无非是想 $\theta_{\mu\nu} = X_{\mu\nu}\theta + \sigma_{\mu\nu}$, 这样自然有 $X^\mu{}_\mu = 1$. θ 是借 \tilde{g}^{-1} 所求之迹, 因此最简单的做法是令 $X_{\mu\nu} = \tilde{g}/3$. 这样便将 B 分解为三个部分:

$$B_{\mu\nu} = \frac{1}{3}\theta\tilde{g}_{\mu\nu} + \sigma_{\mu\nu} + \omega_{\mu\nu}. \quad (4.3.2)$$

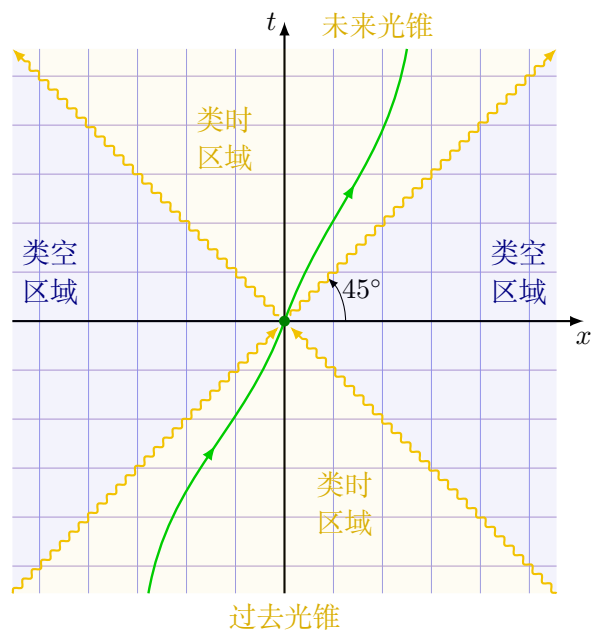


图 4.1: 平时时空的情况

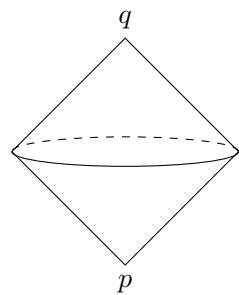


图 4.2: 3 维平时时空的因果菱形. 可见其形似一颗“钻石”, 而其 2 维截面为“菱形”状 (准确说呈平行四边形). 英语世界皆以 “diamond” 代之.

因此

$$u^\mu = \frac{1}{3} \theta \eta^\mu + \sigma^\mu{}_\nu \eta^\nu + \omega^\mu{}_\nu \eta^\nu. \tag{4.3.3}$$

若只有 θ 非零, 则 $\vec{u}, \vec{\eta}$ 将平行, 因此 θ 只影响 $\vec{\eta}$ 长度, 基准观者附近将有辐向伸缩变化, 故 θ 称为膨胀系数或膨胀 (expansion); 若只有 $\omega_{\mu\nu}$ 非零, 则易证 $\vec{\eta}$ 长度不变, 其成为一个纯转动变换, 故 $\omega_{\mu\nu}$ 称为转动 (rotation) 张量或扭转 (twist); 若只有 $\sigma_{\mu\nu}$ 非零, 则可类比连续介质力学的切应力概念, 称其为剪切 (shear) 张量或剪

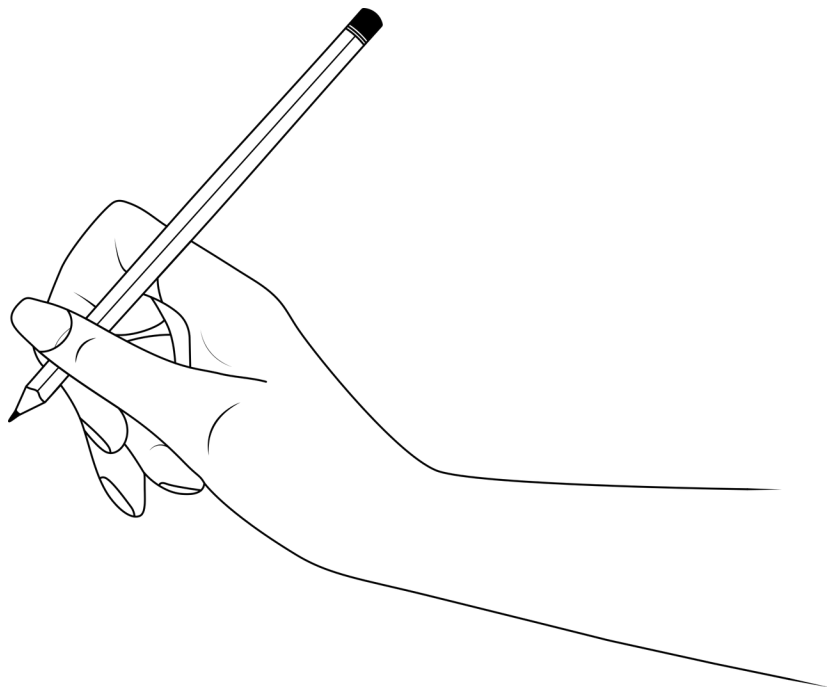


图 4.3: 空间矢量

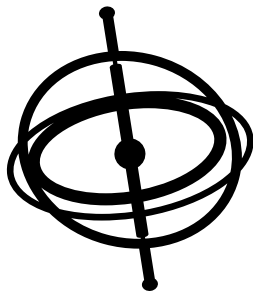


图 4.4: 陀螺仪

切. 我们将在下一章详细研究潮汐效应时再解释¹.
Cauchy 发展

4.4 整体双曲时空

渐近平直时空的共性变换

¹比如, 可证若 η 终点在单位球上, 则该单位球会变为一个椭圆.

Raychaudhuri 方程
引力能量非定域性

Chapter 5

实验理论

5.1 等效原理

WEP 目前最精确的检验是 21 世纪于太空中完成的, 得到铝、铂间 m_G/m_I 的差异小于 10^{-14} , 且数据在地球上空各处基本一致; 关于 EEP, 对量子非引力实验的描述需用波函数, 而波函数至少散布在一个区域内, 这样强引力场的潮汐效应会显现在波函数中, 但可选择对弱引力场中的小区域波函数进行实验, 已有数据显示这种量子系统可满足; SEP 目前最精确的检验利用了地月测距 (lunar-laser-ranging), 测得两天体间 m_G/m_I 的差异不大于 5.5×10^{-13} , 引力结合能的贡献差异不大于 1.3×10^{-3} .

5.2 几何

5.3 后 Newton 近似

5.4 引力波

5.5 宇宙学

Part II

进阶

Chapter 6

Lie 群与 Lie 代数

Chapter 7

规范理论

Chapter 8

经典力学的几何理论

对想要大致习得现代物理学的新人的建议之一，便是“要懂一些微分几何、要懂一点泛函分析”，这在很多情况下都是最有可能的“捷径”。尤其是几何学，对物理本科阶段的几个冤大头板块都是一个非常合适的概述工具，经典力学也不例外。经典力学在现代物理学的处境更接近于一门方法论，而不仅仅再是曾经的那种引领工业发展的力学。这种情况表现在，力学范式的精髓被抽象为了所谓的动力学，从而在物理学其他领域发挥思想上的启迪以及具体应用。

初看起来，现代物理学的理论多到令人眼花缭乱。我们把经典力学同量子力学比较，把非相对论物理与相对论物理比较，或者把粒子物理和统计力学比较，就可以看到一些微妙的联系。有许多强大的统一框架把它们联系起来，其中之一就是在所有这些理论中，系统随时间的演化（及其定常状态）在很大程度上是受到一个单个的对象的控制，即 Hamiltonian 函数，这个函数时常可以解释为系统的总能量。大略地说，每一个物理现象（电磁现象、原子键、势阱中的粒子等等）都对应于一个 Hamiltonian，而每一种类型的力学（经典、量子、统计力学）都相应于用这个 Hamiltonian 来描述一个物理系统的不同方法。例如，在经典物理中，Hamiltonian 确实是这个系统的位置和动量的函数，而按照所谓 Hamiltonian 方程演化。在量子力学中，Hamiltonian “函数” 实际上是一个线性算子（位置算子和动量算子的一个形式组合），而系统的波函数则满足 Schrödinger 方程。在统计物理中，Hamiltonian 则是系统的微观态的函数，而系统在给定的温度 T 下处于这个微观态的概率则是 $e^{-H/kT}$ 。

数学的许多领域都与自己的物理对应物密切地缠在一起，所以 Hamiltonian 的概念也出现在纯粹数学里面。例如，受到经典物理的启发，Hamiltonian 及其推广，如动量映射²在动力系统、辛几何里面都起重要作用。受到量子力学的启发，Hamiltonian 及其推广，如可观测量、拟微分算子在算子代数、谱论、表示论、微分方程和微局部

²矩映射, moment map.

分析中同样也很突出.

因为 Hamiltonian 在那么多物理和数学领域里都出现, 它在建立表面上似乎无关的领域间的桥梁中就很有用, 例如, 建立经典力学与量子力学、辛几何和算子代数的联系. 一个给定的 Hamiltonian 的性质时常会揭示出与其相关的物理或数学对象的众多性质. 例如, Hamiltonian 的对称性时常会诱导出其表示的对象的对称性. 虽然不能说一个数学或物理对象的所有有趣的性质都可以直接从它们的 Hamiltonian 读出, 但是在理解这些对象的性质和动态时它仍然起着基本的作用.

Newton 时空

回顾 Newton 的基本思路, 那就是在承认空间的存在和量度、时间的流逝和量度分别具有各自的绝对性的基础上, 利用决定性原理研究物理学.

决定性原理是一个实验结论, 即首先认同人们日常生活中所接触的时空是一个仿射的 \mathbf{R}^4 , 其次认同二阶微分方程足够描述力学规律: 从原理上来说, 经典力学是一门关于质点动力学的理论. 设若整个质点系统在惯性坐标系中的位置为 $\mathbf{X} = (x_1, \dots, x_n)$, 则运动由粒子的质量 (m_1, \dots, m_n) , $m_j > 0$, 以及相互作用势 $V(x_1, \dots, x_n)$ 通过 Newton 第二定律

$$m_i \ddot{x}_i = -\partial_{x_i} V(x_1, \dots, x_n), \quad i = 1, \dots, n$$

所决定, 这里 $x_i = (x_i^1, \dots, x_i^d)$ 为第 i 个粒子的坐标, ∂_{x_i} 为梯度 $(\partial_{x_i^1}, \dots, \partial_{x_i^d})$; d 是空间维度 (一般 $d = 3$). 势能函数假设是“光滑”的, 即除了在可能的两粒子位置重合的情形下它是解析的¹. 此外, 上式中的 $-\partial_{x_i} V$ 也称为作用在粒子上的“力”². 狭义上, 决定性原理就是在强调, 如果 V 是有下界的, 则给定了初始条件 $\mathbf{X}_0 = \mathbf{X}(0)$, $\dot{\mathbf{X}}_0 = \dot{\mathbf{X}}(0)$, 第二定律有唯一的全局解 $t \mapsto \mathbf{X}(t)$, $t \in (-\infty, \infty)$; 否则, 当且仅当解在有限时间内到达无穷远点或落入奇点 (即两个或多个粒子占据同一点那样的情况, 称作碰撞) 时, 解可以不是全局的.

Newton 时代往下, 都较为盛行使用这样的 3 维矢量的语言书写物理学.

- 让我们解释一下这些细节. 首先, 自然是人们从长期的生产实践中逐渐产生了“物”与“空间”的朴素构思. 通俗而言, 一个物体必然是从一堆物质中得以区分的部分实在. 其定义的关键是可以被区分. 如果研究对象是单一的, 而不存在任何其它物体, 那么首先不会有区分的需求, 讨论物理学将不会有意义, 因

¹为了包括引力势, 或在处理带电粒子所涉及的静电力势, 上述例外是必需的.

²更一般形式的力模型也会经常遇到, 比如依赖于速度的摩擦力. 我们这里不考虑摩擦力, 因为其本质上是关于微观现象的唯象模型 (即使在宏观的层面上摩擦力也是不完全的模型, 它们应该和伴随着的热产生现象一同考虑, 原则上是可以保守力解释的. 另一类有趣的、不对应于一个势能的力是类似科氏力 (非惯性系中) 和 Lorentz 力 (电磁学) 那样的依赖于速度的力, 它们很容易在力学的 Hamilton 表述的框架中处理, 我们可能以后会聊这个话题.

为不存在具有相互作用的物理量。如果研究对象是难以区分的，那么经典力学体系的一些物理量很难使用，比如流体是难以区分的，这个时候使用质量、位移反而并不方便。

- 其次，是所谓时间概念，它的内涵至今仍然是学术界难以解释完备的，但即使如此，粗糙的讨论也是非常具有启迪性的。人可以显然地感受到同种对象不同状态的存在，比如说天空的昼夜景色，通过记忆能比较出它们的不同。出于技术需求，总是期望知晓一个状态产生的原因，最好是能用一个参数描述状态间的运动。在经典力学中，这个参数一般定义为时间。时间的测量依附于运动中的任意状态之比较。经典观点下不再深入，认为时间的均匀流逝是“显然的”，只须规定一些稳定现象或者信号传递的周期。现代观点取某些核半衰期之倍数为时间均匀测量的基础，亦或者规定光速为信号传播的基础。后者造成的哲学思辨是时间和状态测量相互影响，但这不在本文的讨论范围内。

物体的相对移动对象的状态就是所谓的事件，自然也就是 \mathbf{R}^4 里的元素了。于是，一个质点其实就是时空里的一条世界线。在数学工具足够的情况下，可以认为世界线汇能够研究我们日常生活中的各类可区分对象。用来确定世界线之间相对空间关系的标架称作坐标系，而线汇也就称为参考系了，二者紧密关系足以彰显。Newton 时空的特点体现在同时的绝对性，这也就是说总有办法让时间 t 直接地成为 \mathbf{R}^4 的标量场，而不依赖于任何坐标系。那么，就可以定义商空间 $\mathbf{R}_t^3 \simeq \mathbf{R}^4/\mathbf{R}^1$ 为（绝对）同时面。

经典力学讨论的运动是指 \mathbf{R}^3 中的曲线 $\alpha: I \rightarrow \mathbf{R}^3, I \subset \mathbf{R}$ 。如果曲线可微，就可以定义切向量为速度，当然，由于 Galilean 空间如此简单， $T_\alpha \mathbf{R}^3$ 也可以认同为 \mathbf{R}^3 的一个嵌入。如果速度可微，就可以定义切向量为加速度，类似地， $T_\alpha T_\alpha \mathbf{R}^3$ 也可以认同为 \mathbf{R}^3 。这也就是说，这些三维矢量总可以认为是仿射的。决定性原理告诉我们，这些定义目前就够用了。

于是 Newton 第一定律就可以表述为：存在这样一类平权¹的参考系，它的运动是 \mathbf{R}^3 的测地线，称为惯性系。惯性系这种平权性就称作 (Galilean) 相对性原理²。

Galilean 变换

不能在整个 Newton 时空上定义度规³。举个例子，由于无法判断惯性系的运动状态，“某一世界线是静止孤立系”是没有意义的。于是不同事件的距离只能要么谈

¹从实验角度来说，如果两个参考系下，同一孤立世界线的 3-加速相同，那么这两个参考系就是平权的。

²Newton 提出的绝对空间，其实就是在默认一个正交归一的“绝对静止”的刚性标架，这当然在很多时候都是很方便的。

³只能定义半度规，Galilei 群正是这个半度规的等距群。

及时间差, 要么谈及同时面上的欧氏度规. 于是, 背景流形 \mathbf{R}^4 就不能是 $\mathbf{R} \times \mathbf{R}^3$, 这是因为各个同时面不能共同构成类.

因此, 讨论 Newton 时空的对称性时, 同样只能分开研究: 时间上显然有 1 个独立 Killing 矢量场, 反映其平移; 同时面上则有 6 个, 反映其平移和旋转, 组成 Euclidean 群 $E(3) = \mathbf{R}^3 \rtimes O(3)$. Galilean 变换正是 Galilean 空间的 (线性) 自同构 $g: (x, t) \mapsto (R\mathbf{x} + \mathbf{c} + \mathbf{v}t, t + s)$, 其中除了对称性还要加上惯性系变换 (由空间平移自然诱导), 这些变换组成了 Galilei 群.

另外, 实际上可以认为 Newton 时空是纤维丛 $(\mathbf{R}^4, \mathbf{R}, \pi, \mathbf{R}^3)$, 其中底流形 \mathbf{R} 给时间, 纤维 \mathbf{R}^3 给空间, 对时间的赋值是投影 π . 那么世界线的运动就是这个 Galilei 丛的截面了.

非惯性系

这种情况会自然引入一些与加速度相关的物理量. 角速度 ω 很简单, 它是 1-形式位矢和 1-形式速度的外积的对偶. $O(3)$ 中 $|R| = 1$ 的那一个连通分支¹为 $SO(3)$, 因此角速度就是其无穷小生成元. 令 2-形式 $\omega \in \mathfrak{so}(3)$ 满足 $\omega_{ij} = \omega_k$, 那么它对位矢的作用就是线速度. 因此旋转其实就是 $R = e^{\omega t}$.

Lagrange 观点

Lagrange、Euler 等提出了与第二定律等价的表述: 满足第二定律的、在时间间隔 $[t_1, t_2]$ 内从点 $\mathbf{X}_1 = \mathbf{X}_0(t_1)$ 到点 $\mathbf{X}_2 = \mathbf{X}_0(t_2)$ 的运动 $t \mapsto \mathbf{X}_0(t)$ 可以使泛函

$$S(\{\mathbf{X}\}) = \int_{t_1}^{t_2} \left(\sum_{i=1}^n \frac{1}{2} m_i \dot{\mathbf{X}}_i(t)^2 - V(\mathbf{X}(t)) \right) dt$$

在 $\Omega_{t_1, t_2}(\mathbf{X}_1, \mathbf{X}_2)$ 中平稳, 这里 Ω 定义为在 $t \in [t_1, t_2]$ 间隔内从点 \mathbf{X}_1 到点 \mathbf{X}_2 的光滑 (即解析的) “运动” $t \mapsto \mathbf{X}(t)$ 组成的集合. 中间的括号

$$L(\mathbf{Y}, \mathbf{X}) = \frac{1}{2} \sum_{i=1}^n m_i y_i^2 - V(\mathbf{X}) \stackrel{\text{定义为}}{=} T(\mathbf{Y}) - V(\mathbf{X}),$$

称为 Lagrange 函数, 其中 $\mathbf{Y} = (y_1, \dots, y_n)$, $T(\dot{\mathbf{X}}(t))$ 是动能. 这样作用量就可写成 $\int_{t_1}^{t_2} L(\dot{\mathbf{X}}(t), \mathbf{X}(t)) dt$. 容易证明这样定义的 “总能量” 是不变的: $T(\dot{\mathbf{X}}(t)) + V(\mathbf{X}(t)) = E$, 因此, 作用量原理可以直观上认为是说: 运动在保持能量守恒的同时, 尽可能均等地分配 (在时间平均的意义上) 对能量的贡献.

¹ 为防混淆, 不考虑反演的那些旋转. 因为 1-形式的反演反号, 但 2-形式不变.

流形

这个发现导致的一个效果就是，即使把这种狭义的质点坐标任意推广（比如角坐标），第二定律的表述可能会很复杂但这个原理“形式仍然没变”，这样在复杂的力学、约束系统中，Lagrange 体系可能会更加方便。即然坐标之间也只是按照几何约束进行一种变换，那这也说明，我们是在含有所有系统状态的空间讨论问题，这个空间是个流形。

定义位形空间（构型空间） Q 是一个流形，其上的局部坐标称作广义坐标，其维数称作自由度，其切向量称作虚位移或者广义速度。进而地，由决定性原理， TQ 的每一个元素就可以称作一个状态。一个 Lagrange 函数 L 定义为乘积流形 $TQ \times \mathbf{R}$ 上的标量场，其唯一确定一个 TQ 上的 Lagrangian 向量场。 (Q, L) 称作一个 Lagrangian 系统。以下，我们就可以不再用 \mathbf{X} 而是用 q 表述坐标。

作用量

设 Q 上边界为 $q_1 - q_0 = \gamma(t_1) - \gamma(t_0)$ 的曲线 γ 构成无穷维 Fréchet 流形 $\Omega(Q)$ ，其中存在 Γ 经过 γ ，称为 γ 的变分。这样的变分构成一族 $\gamma_\epsilon(t) = \Gamma(t, \epsilon)$ ，其中 ϵ 在 0 的某邻域内，且 $\gamma_0 = \gamma$ 。如前文提及，其自然诱导 TQ 里的 $\gamma'(t) := \gamma_*(\frac{\partial}{\partial t})$ 。于是，最小作用量原理就是说运动曲线是 Lagrange 矢量场的积分曲线，即等价于任意 γ_ϵ 都是作用量 $S(\gamma_\epsilon) = \int_{t_0}^{t_1} L(\gamma'_\epsilon, t) dt$ 的极值点，也即 $dS/d\epsilon|_{\epsilon=0} = 0$ 。取一图 (U, q_n) ，就可确定切丛上 $\pi^{-1}(U)$ 的自然坐标 $(q, \dot{q}) = (q^n, \dot{q}^n)$ 。于是 γ' 就是 (q, \dot{q}) ，那么对应的微分方程就是 Lagrange 方程。

Lagrangian

考虑一惯性系中的自由质点，根据对称性假设， $\frac{\partial L}{\partial \mathbf{q}} = \frac{\partial L}{\partial t} = 0$ ，且在欧氏空间里 L 只依赖于速率而非方向。 Q 上可定义切向量大小的度规称为质量 m 。根据相对性原理，令质点对另一惯性系的速度是 $\dot{\mathbf{q}} + \mathbf{v}$ ，那么两个参考系的 Lagrangian 至多差一个项 $\dot{f}, f: Q \times \mathbf{R} \rightarrow \mathbf{R}$ 。设 q 是某局部法坐标，则质量可以展开为 $\delta_{ij} + \frac{1}{3}R_{kijl}q^kq^l + O|q|^3$ ，则其余项均是高阶小量，于是 $|\dot{\mathbf{q}}| = \sqrt{\dot{\mathbf{q}} \cdot \dot{\mathbf{q}} + 2\mathbf{v} \cdot \dot{\mathbf{q}} + \mathbf{v} \cdot \mathbf{v}}$ ，那么另一参考系的 $L'(|\dot{\mathbf{q}} + \mathbf{v}|)$ 按 \mathbf{v} 展开，就得到 $L = \frac{1}{2}m(\dot{\mathbf{q}}, \dot{\mathbf{q}})$ 。

刚体

Lagrange 解就是测地线，这使得我们可以讨论 Lie 群上的 Lagrangian。考虑绕固定点旋转的刚体，其构型空间是 $SO(3)$ ，由于 m 是 $SO(3)$ 上的左不变度规，则 $L = \frac{1}{2}m(v, v)$ ，其中 $v \in TSO(3)$ 。对任意 $\dot{\omega} \in T_\omega SO(3)$ ，定义角速度 $\Omega = (L_{\omega^{-1}})_*\dot{\omega} \in \mathfrak{so}(3) = T_e SO(3)$ ，同理 $L = \frac{1}{2}m_e(\Omega, \Omega)$ 。注意 $\mathfrak{so}(3)$ 是半单的，所以可以考

考虑 Killing-Cartan 形式 $\kappa(x, y) = -\frac{1}{2}\text{tr}(\text{ad}_x \text{ad}_y)$, 且存在线性算子 $A : \mathfrak{so}(3) \rightarrow \mathfrak{so}(3)$ 使 $m_e(\Omega, \Omega) = \kappa(A(\Omega), \Omega)$. 这里 $A(\Omega)$ 就是角动量, A 被称为刚体的惯量算子, 它的三个正交归一化的本征向量叫做惯量主轴, 对应的本征值 I_1, I_2, I_3 叫做主轴的惯量矩. 展开 $\Omega = \Omega_i e_i, i = 1, 2, 3$, 则 $L = \frac{1}{2} I_i \Omega_i^2$, Lagrange 方程变为刚体的 Euler 方程 $I_i \dot{\Omega}_i = (I_j - I_k) \Omega_j \Omega_k$.

Noether 定理的 Lagrange 表述

称 $I : TQ \rightarrow \mathbf{R}$ 是守恒律, 若任意实际运动 γ 都使得 $\dot{I}(\gamma'(t)) = 0$. 若同胚 $\Phi : Q \rightarrow Q$ 使得 $\forall v \in TQ, L(\Phi_*(v)) = L(v)$, 则称 L 在 Φ 下不变, 则 Φ 为对称性. 若 Lie 群 G 在 Q 上的光滑左作用使得 $\forall g \in G, q \mapsto g \cdot q$ 是对称性, 那么 G 就是对称群. 考虑一族连续变化的对称性构成的单参群 $\{\Phi_s\}$, 按 $a^i(q) = \left. \frac{d\Phi_s^i(q)}{ds} \right|_{s=0}$ 给定 $\mathbf{a} : Q \rightarrow Q$, 再按 $I(\mathbf{q}, \dot{\mathbf{q}}) = \frac{\partial L}{\partial \dot{\mathbf{q}}}(\mathbf{q}, \dot{\mathbf{q}})(a^i(\mathbf{q})) = \frac{\partial L}{\partial \dot{\mathbf{q}}}(\mathbf{a})$, 它就是不含时系统的守恒律的一般形式. 证明只需在 $s = 0$ 处把 $L((\Phi_s)_*(\gamma'(t))) = L(\gamma'(t))$ 对 s 求导, 取 $\mathbf{a}(t) = (\mathbf{a}^n(\gamma(t)))$, 然后再代入 Lagrange 方程即可.

广义动量 $\frac{\partial L}{\partial \dot{\mathbf{q}}}$ 不依赖于坐标的选取. 取两图 $(U, \varphi), (U, \psi)$, 设坐标变换 $F = \psi \circ \varphi^{-1}$, $(\mathbf{q}, \dot{\mathbf{q}})$ 是第一个图中 TQ 的自然坐标, 那么 $(\tilde{\mathbf{q}} = F(\mathbf{q}), \dot{\tilde{\mathbf{q}}} = dF_{\mathbf{q}}(\dot{\mathbf{q}}))$ 是第二个图中的. 易得 $\frac{\partial L}{\partial \dot{\tilde{\mathbf{q}}}} \circ dF_{\mathbf{q}} = \frac{\partial L}{\partial \dot{\mathbf{q}}}$, 证毕.

有三类守恒律:

能量 $E = \dot{\mathbf{q}} \frac{\partial L}{\partial \dot{\mathbf{q}}} - L$. 能量看作守恒律时, 必须考虑含时体系, 此时的构型空间是 $Q \times \mathbf{R}$, 含时 $L'(\mathbf{q}, \tau, \dot{\mathbf{q}}, \dot{\tau}) = L(\mathbf{q}, \frac{\dot{\mathbf{q}}}{\dot{\tau}}, \tau) \dot{\tau}$, 这里 L 是把 $T(Q \times \mathbf{R})$ 仍看做不含时系统得出的 Lagrangian, 而能量对应的单参对称群是 $\{\tau \mapsto \tau + s\}$. 由于 $\frac{\partial L'}{\partial \dot{\tau}} = \frac{\partial L}{\partial(\dot{\mathbf{q}}/\dot{\tau})} \frac{\partial(\dot{\mathbf{q}}/\dot{\tau})}{\partial \dot{\tau}} \dot{\tau} + L = -\dot{\mathbf{q}} \frac{\partial L}{\partial \dot{\mathbf{q}}} + L$, 知单参群诱导的守恒律是 $-E$.

动量 $I = v^i \frac{\partial L}{\partial \dot{q}^i}$. 其单参群为 $\Phi_s(q) = q + sv$, 其中 $v \in Q$, 而为了使这种加法成立, 不妨取 Q 为线性空间. 最简单的情形是 $Q = \mathbf{R}^{3N}$, 且为简便, 取 v 每个分量都是单位向量, 那么总动量为 $I = \sum_{k=1}^N \left(\frac{\partial L}{\partial \dot{r}_k^1} + \frac{\partial L}{\partial \dot{r}_k^2} + \frac{\partial L}{\partial \dot{r}_k^3} \right)$.

角动量 \mathbf{J} . 仍取 $Q = V$ 是线性空间, $G = \text{SO}(V)$, $\exp : \mathfrak{g} \rightarrow G$ 是指数映射, 守恒律 $I = (R \cdot q)^i \frac{\partial L}{\partial \dot{q}^i}$, 其中 $R \in \mathfrak{g}, R \cdot q$ 是 Lie 代数在 Q 上的左作用¹, 单参群为 $\Phi_s(q) = \exp(sR) \cdot q$. 同上, 把 $R \in \mathfrak{so}(3)$ 按其标准基底展开为 $R = R^i X_i$, 那么 $I = R^i J_i$, 其中 $\mathbf{J} = \sum_{k=1}^N \mathbf{r}_k \times \frac{\partial L}{\partial \mathbf{r}_k}$.

若中心立场为最特殊的 $\ddot{\mathbf{r}} = -\alpha \mathbf{r} |\mathbf{r}|^{-3}$, 则额外有一个守恒的 Runge-Lenz 矢量 $\epsilon = \mathbf{P} \times \mathbf{J} - \alpha \mathbf{r} |\mathbf{r}|^{-1}$. 椭圆、抛物和双曲三种轨道中该矢量的对称群分别是 $\text{SO}(4), \text{SE}(3)$ 和 $\text{SO}(3, 1)$, 单参群的形式与角动量一致.

¹ 即矩阵作用在向量上.

Hamilton 观点

Hamilton 观点最大的特点就是使得力学变成了几何学. 位形空间的余切向量称为广义动量, 因此 T^*Q 就称作相空间. 考虑其上自然坐标 $(\mathbf{q}, \mathbf{p}) = (q^n, p_n)$ 使得 $p_i(df) = \frac{\partial f}{\partial q^i}$ 成立, 其中 f 是函数. 作用量可以用微分 1-形式表示出来. 定义 Lagrange 1-形式 $dL = \frac{\partial L}{\partial \dot{q}^i} dq^i$ 和 Liouville 典范 1-形式 $\tau = p_i dq^i$. 根据前面的论述, 我们知道 dL 的定义是不依赖于坐标选取的. 那 Liouville 1-形式呢? 实际上, Liouville 1-形式可以用内蕴的形式定义: $\tau_{(\mathbf{q}, \mathbf{p})}(v) = \mathbf{p}(d\pi_{(\mathbf{q}, \mathbf{p})}(v))$, 其中 $\mathbf{p} \in T_{\mathbf{q}}^*Q$, $\pi: T^*Q \rightarrow Q$ 是标准投影, $v \in T_{(\mathbf{q}, \mathbf{p})}(T^*Q)$. 由于 $d\pi^*(dq^i) = dq^i$, 立即可有 $\tau_{(\mathbf{q}, \mathbf{p})} = d\pi_{(\mathbf{q}, \mathbf{p})}^*(p_i dq^i) = p_i dq^i$. 因此从 Lagrangian 到 Hamiltonian 的 Legendre 变换或纤维变换 $\tau_L: TQ \rightarrow T^*Q$ 就是满足 $dL = \tau_L^*(\tau)$ 的丛态射.

τ_L 至少是局部微分同胚. 由于在自然坐标下 $\tau_L(\mathbf{q}, \dot{\mathbf{q}}) = (\mathbf{q}, \frac{\partial L}{\partial \dot{\mathbf{q}}}(\mathbf{q}, \dot{\mathbf{q}}))$, 所以我们只需考虑 $\frac{\partial L}{\partial \dot{\mathbf{q}}}$ 在局部是否可逆. 判别的最简单方法, 就是再对其求偏导, 得到的 $\frac{\partial^2 L}{\partial \dot{q}^i \partial \dot{q}^j}$ 是 L 的 Hessian 矩阵. 所以, 若其处处可逆, 则 τ_L 是局部微分同胚. 从而我们可以按照 $H \circ \tau_L = \dot{\mathbf{q}} \frac{\partial L}{\partial \dot{\mathbf{q}}} - L$ 给定 Hamiltonian $H: T^*Q \rightarrow \mathbf{R}$, 即把切丛上能量函数变换成余切丛上的. 此时, 余切丛的自然坐标是事先给定的, 切丛上的 $\dot{\mathbf{q}}$ 是需要推导出来的. 当然, 直接取 $\mathbf{p} = \frac{\partial L}{\partial \dot{\mathbf{q}}}$, 就利用隐函数定理唯一地给定了, 此时 $H(\mathbf{q}, \mathbf{p}) = \mathbf{p}\dot{\mathbf{q}} - L(\mathbf{q}, \dot{\mathbf{q}})$, 从而

$$\begin{aligned} dH &= \frac{\partial H}{\partial \mathbf{p}} d\mathbf{p} + \frac{\partial H}{\partial \mathbf{q}} d\mathbf{q} \\ &= \mathbf{p} d\dot{\mathbf{q}} + \dot{\mathbf{q}} d\mathbf{p} - \frac{\partial L}{\partial \mathbf{p}} d\mathbf{p} - \frac{\partial L}{\partial \dot{\mathbf{q}}} d\dot{\mathbf{q}} \\ &= \dot{\mathbf{q}} d\mathbf{p} - \frac{\partial L}{\partial \mathbf{p}} d\mathbf{p} \\ &= \dot{\mathbf{q}} d\mathbf{p} - \dot{\mathbf{p}} d\mathbf{q}, \end{aligned}$$

即有 Hamilton 正则方程. 易知 Hamiltonian 是守恒量. 给定了 H 和 m 后, 立刻可得 $L(\xi) = (j_m(\xi))(\xi) - H(j_m(\xi))$, 其中 $j_m: TQ \rightarrow T^*Q$ 是 m 诱导的自然丛同构或音乐同构, 这样又可以回到 Lagrange 观点.

辛结构

相空间是一个自然的辛流形, 于是就可以讨论正则变换. τ 的外微分 $\omega = dp_i \wedge dq^i$ 是非退化的, 即 $\forall q \in Q, v \in T_q Q, \exists w \in T_q Q$ 使得 $\omega_q(v, w) \neq 0$. 显然 ω 是闭的. 据此, 称 ω 其为辛形式 (辛内积), 给定了辛形式 (辛结构) 的流形叫辛流形. 对于两个辛流形 $(M_1, \omega_1), (M_2, \omega_2)$, 若同胚 $F: M_1 \rightarrow M_2$ 满足 $F^*\omega_2 = \omega_1$, 则称 F 是辛态射; 若额外地 $M_1 = M_2$, 则 F 称作正则变换. 不妨把 ω 写成 $dq_i \wedge dp^i$, 即流形坐标在前, 纤维坐标在后, 进而 $\omega = -d\tau$.

非退化意味着 $\dim M$ 为偶数, 这一点 $M = T^*Q$ 显然满足. 证明: 称定义了形式 ω 的线性空间 V 是辛线性空间, 容易证明 $\dim V \geq 2$. 若 $S \subset V$ 是线性子空间, 则 $S^\perp = \{v \in V | \omega(v, w) = 0, \forall w \in S\}$ 称为 S 的辛补. 令 $\Phi(v)(w) = \omega(v, w)$, 对其利用秩-零化度定理, 知 $\dim S + \dim S^\perp = \dim V$. 若 $S \cap S^\perp = \{0\}$ 则 $\omega|_S, \omega|_{S^\perp}$ 均非退化, 再令 $\dim S = 2$, 对 $\dim V$ 应用归纳法, 可知 $\dim S^\perp$ 一定是偶数.

Darboux 定理

辛流形上任意局部都存在坐标 (x^n, y^n) 使 $\omega = \sum_{i=1}^n dx^i \wedge dy^i$ 成立.

Poisson 括号

现在可以定义任意两个力学量的 Poisson 括号, 它仍然是一个函数. 设辛流形 (M, ω) , 我们说满足 $X_f \lrcorner \omega = df$ 的向量场 X_f 称为 f 的 Hamiltonian 向量场. M 上全体 Hamiltonian 向量场的集合记作 $HT(TM)$. 考虑 Darboux 坐标, 即使得 $\omega = \sum_{i=1}^n dq^i \wedge dp^i$ 的局部坐标, 用待定系数法设 $X_f = a^i \frac{\partial}{\partial q^i} + b^i \frac{\partial}{\partial p^i}$, 那么 $X_f \lrcorner \omega = \sum_{i=1}^n (a^i dp^i - b^i dq^i)$, 而 $df = \frac{\partial f}{\partial q^i} dq^i + \frac{\partial f}{\partial p^i} dp^i$, 对比系数知 $X_f = \sum_{i=1}^n (\frac{\partial f}{\partial p^i} \frac{\partial}{\partial q^i} - \frac{\partial f}{\partial q^i} \frac{\partial}{\partial p^i})$, 于是 Poisson 括号的定义非常自然了: $\{f, g\} := \omega(X_f, X_g) = X_g f$, 所以

$$\{f, g\} = \sum_{i=1}^n \left(\frac{\partial f}{\partial q^i} \frac{\partial g}{\partial p^i} - \frac{\partial f}{\partial p^i} \frac{\partial g}{\partial q^i} \right).$$

不论从内蕴定义还是局部坐标, 易得 $X_{\{f, g\}} = -[X_f, X_g]$. 辛形式的闭性等价于 Poisson 括号的 Jacobi 恒等式.

辛流形上的所有向量场, 都至少局部上是某函数的 Hamiltonian 向量场. 不妨考虑使得 $\mathcal{L}_X \omega = 0$ 的向量场, 由 Cartan 同伦式得 $d(X \lrcorner \omega) = 0$, 若存在 f 使得 $X = X_f$, 则其显然是闭的, 所以 Hamiltonian 向量场一定使得 $\mathcal{L}_X \omega = 0$. 反过来, 由 Poincaré 引理, 局部存在 f 使得 $X \lrcorner \omega = df$, 证毕. 局部 Hamiltonian 向量场延拓到全域的拓扑障碍是 de Rham 上同调 $H_{dR}^1(M)$, 当且仅当其为 0, 局部都为全域.

在给定相空间 M 上的函数 $H: M \rightarrow \mathbf{R}$ 之后, (M, ω, H) 就叫做 Hamiltonian 系统. 引入 Poisson 括号最大的好处就是守恒律的判断很简单. 若函数 I 在 X_H 的积分曲线 γ 上为常值, 则称其为守恒律. 首先 $\gamma'(t) = X_H|_{\gamma(t)}$, 那么 $\{I, H\}|_{\gamma(t)} = X_H|_{\gamma(t)} I = (I \circ \gamma)'(t) = 0$. 由于 γ 是任取的, 所以 $\{I, H\} = 0$; 反之也能推出 I 守恒. 若 I_1, I_2 均守恒, 由 Jacobi 恒等式, $\{I_1, I_2\}$ 也守恒.

设 Lie 群 G , 给定向量场 X , G 在 M 上的光滑左作用有逐点定义的向量场 $\hat{X}_q := \frac{d}{dt}|_{t=0} (\exp(tX) \cdot q)$ 和无穷小生成元 $\hat{\theta}(X) = \hat{X}$. 无穷小生成元是 $\mathfrak{g} \rightarrow \Gamma(TM)$ 的反 Lie 代数同态, 那么 $-\hat{\theta}$ 就是 Lie 代数同态. 我们有从函数导出 Hamiltonian 向

量场的步骤 $f \mapsto X_f$, 记这个同态为 HV . 总能找到一个同态 $\hat{J}: \mathfrak{g} \rightarrow C^\infty(M)$ 使得如下交换图成立:

$$\begin{array}{ccc} C^\infty(M) & \xrightarrow{HV} & H\Gamma(TM) \\ & \nwarrow \hat{J} & \uparrow -\hat{\theta} \\ & & g \end{array}$$

若任意 $A \in \mathfrak{g}$ 都使得 $-\hat{\theta}(A)$ 是全域 Hamiltonian 向量场¹, 则肯定找得到的但不唯一, 它们可以相差一个 $\alpha \in \mathfrak{g}^*$ 项. 给定 \hat{J} 之后, 按 $J(x)(A) = \hat{J}(A)(x)$ 定义 $J: M \rightarrow \mathfrak{g}^*$, 其中 $x \in M, A \in \mathfrak{g}$. 称 J 为 Hamiltonian 左作用的动量映射. 在引入 \mathfrak{g}^* 上的 Lie-Poisson 结构后, \mathfrak{g}^* 成为 Poisson 流形, 动量映射成为 Poisson 态射.

Noether 定理的 Hamilton 表述

设 H 在 G 的左作用下是不变的, 即任意 $g \in G$ 都使得 $H(g \cdot x) = H(x)$. 则

$$\begin{aligned} \{\hat{J}(A), H\}(x) &= (X_{\hat{J}(A)}H)(x) \\ &= \left. \frac{d}{dt} \right|_{t=0} H(\exp(tA) \cdot x) \\ &= \left. \frac{d}{dt} \right|_{t=0} H(x) \\ &= 0, \end{aligned}$$

即 J 是守恒律.

作用量作为 1-形式

在 $T^*Q \times \mathbf{R}$ 上, 定义 Poincaré-Cartan 1-形式 $dS := \mathbf{p}d\mathbf{q} - Hdt$, 考虑 $\Omega(T^*Q)$, 其中 $\pi(\gamma(t_i)) = q_i, i = 0, 1$. 这里 π 是典范投影. 以及另一 $\tilde{\Omega}(T^*Q) = \{\sigma: [t_0, t_1] \rightarrow T^*Q \times \mathbf{R}\}$, 其中 $\exists \gamma \in \Omega(T^*Q), \sigma(t) = (\gamma(t), t)$. σ 叫做 γ 的提升. 那么, 以下就是相空间中的最小作用量原理, 它的证明同样分部积分即可: σ 是作用量 $S(\sigma) = \int_\sigma dS$ 的极值点, 当且仅当 σ 是 γ 的提升. 这里 dS 就是把 S 看作坐标和时间的函数后外微分得到的.

Hamilton-Jacobi 方程

把初值条件为 $\gamma(t_0) = \mathbf{q}_0, \gamma'(t_0) = v_0$ 的 Lagrange 解记作 $\gamma(t; \mathbf{q}_0, v_0)$. 若存在 v_0 的邻域 $V_0 \subset T_{v_0}Q$ 和 $t_1 > t_0$ 使得 $\forall t \in (t_0, t_1)$, 任意 $v \in V_0$ 所对应的 $\gamma(t; \mathbf{q}_0, v)$

¹ 这样的左作用被称为 Hamiltonian 左作用.

在扩展位形空间 $Q \times \mathbf{R}$ 中是两两不相交的, 则称所有这些解构成 $\gamma(t; \mathbf{q}_0, v_0)$ 的中心场. ODE 理论告诉我们, 只要 t_1 足够小, 中心场总是存在. 在给定 $\gamma(t; \mathbf{q}_0, \dot{\mathbf{q}})$ 的中心场之后, 定义经典作用量 $S(\mathbf{q}, t; \mathbf{q}_0, t_0) := \int_{t_0}^t L(\gamma'(s))ds$, 其中 $\gamma(s)$ 是中心场中终点为 \mathbf{q} 的曲线. 令 $\mathbf{q}(t)$ 的初始速度为 $\mathbf{v} \in T_{\mathbf{q}}Q$, 选取中心场里的曲线使得其变分为 $\delta\gamma(t) = \mathbf{v}$, 那么变分式 $\int_{t_0}^{t_1} \left(\frac{\partial L}{\partial \mathbf{q}} - \frac{d}{dt} \frac{\partial L}{\partial \dot{\mathbf{q}}} \right) \mathbf{v} dt + \left[\frac{\partial L}{\partial \dot{\mathbf{q}}} \mathbf{v} \right]_{t_0}^{t_1}$ 给出 $dS(\mathbf{v}) = \frac{\partial L}{\partial \dot{\mathbf{q}}} \mathbf{v}$, 从而 $\frac{\partial S}{\partial \mathbf{q}} = \mathbf{p} = \frac{\partial L}{\partial \dot{\mathbf{q}}}$. 又注意 $\frac{d}{dt} S(\mathbf{q}(t), t) = \frac{\partial S}{\partial \mathbf{q}} \dot{\mathbf{q}} + \frac{\partial S}{\partial t} = L$, 得 $\frac{\partial S}{\partial t} = L - \mathbf{p}\dot{\mathbf{q}} = -H$, 即得 Hamilton-Jacobi 方程.

Liouville 定理

辛体积形式对应于 Q 上任意一个区域的辛体积, 因此 Liouville 定理就是说 Hamiltonian 向量场相流的辛体积不变.

动力系统

从物理观点看来, 一个动力系统就是一个随时间变化的物理系统, 如行星系或渠道里的水流. 典型情况是, 这个系统的各个部分在时刻 t 的位置和速度只依赖于恰好这个时刻以前各个部分的位置与速度, 这意味着这个系统的行为是由一组 PDE 来控制的. 一组非常简单的 PDE 时常会导致物理系统的非常复杂的行为.

从数学观点看来, 一个动力系统则是按照精确的法则随时间演化的数学对象, 这个法则从这个系统恰好在时刻 t 以前的行为决定了它在时刻 t 的行为. 这里说的“恰好在时刻 t 以前”有时是讲的在无穷小时间段以前, 所以就涉及了微积分. 但是在离散动力系统理论中, 时间只取离散值, 而“恰好在时刻 t 以前”就是指的 $t-1$. 如果一个函数 f 告诉我们这个系统在时刻 t 是如何依赖于它在时刻 $t-1$ 的情况的, 这个系统从整体上看就是 f 的迭代, 即反反复复地用 f 作用于这个系统.

和连续动力系统一样, 非常简单的函数 f , 如果迭代了充分多次, 会导致非常复杂的行为. 有些最有趣的动力系统, 会展现出一种对于初始条件的极端的敏感性, 这就叫做混沌. 例如控制天气的方程式. 我们不能期望精确地确定地球表面的每一个点处的风速 (更不说是在高处的风速了), 就是说只能是有了近似的风速就过得去了. 因为相关的方程式是混沌的, 所得出的不准确的误差, 哪怕开始很小, 迅速地就传播开来, 马上就征服了整个系统, 可以从另一个同样很好的近似开始, 就会发现, 只要经过了很短的时间, 系统就会以一种完全不同的方式演化. 这就是为什么时间超出几天以后的准确预报是不可能的.

Chapter 9

上同调论

一些函数与算子空间

矢量空间，尤其是以后会学到的赋范空间（比如 Hilbert 空间），之间的映射称为算子。通常考虑线性算子。

函数和算子空间上可以赋予不同的拓扑，它们由不同的收敛和范数（或半范数）的概念给出。通常对于同一个空间就有这些不同的概念。如前所述，向量空间上的一个范数给出一个度量，从而给出一个拓扑。类似地，利用函数和算子序列的收敛性的概念，我们可以知道什么是极限点，从而知道什么是闭集。

但是在这之前，让我们先在一个稍微不同的情形下引入一个拓扑。这个拓扑在某种意义上是空间之间连续映射所构成的空间上的自然拓扑。

定义 9.1. 考虑从拓扑空间 X 到拓扑空间 Y 的一族连续映射 F ，并定义 $W(K, U) = \{f : f \in F, f(K) \subseteq U\}$ 。则由所有具有形式 $W(K, U)$ 的集合构成的族构成 F 上的紧致开拓扑 (compact open topology) 的一个子基。这里 K 是 X 中紧致集， U 是 Y 中开集。

考虑拓扑空间 X 和其上的一列函数 (f_n) 。令 $D \subseteq X$ 。就像对于实直线的子集上的函数那样，我们可以定义逐点收敛和一致收敛。

定义 9.2. 设 $X, D, (f_n)$ 如上，

(i) 函数 f_n 在 D 上逐点收敛 (converge pointwise) 到 f ，如果对于任意 $x \in D$ ，数列 $f_n(x)$ 收敛， $f_n(x) \rightarrow f(x)$ 。

(ii) 函数 f_n 在 D 上一致收敛 (converge uniformly) 到 f ，如果对任意 $\varepsilon > 0$ ，存在一个正整数 N ，使得对所有 $x \in D, n > N$ ，有 $|f_n(x) - f(x)| < \varepsilon$ 。

下面我们考虑 Lebesgue 空间 L^p ，即定义在 \mathbf{R}^n 的子集上， $|f(x)|^p$ Lebesgue 可积的函数 f 。这里实数 $p \geq 1$ 。在定义这些空间时，我们隐含地使用了几乎处处（即

相差一个零测集) 相等的函数的等价类. 但通常我们总可以取这些等价类的代表元, 从而和真正的函数打交道. 注意对所有的 L^p , 只有 L^2 是 Hilbert 空间.

定义 9.3. 在空间 L^p 上定义范数 $\|f\| = (\int |f(x)|^p dx)^{1/p}$.

现在我们转到一般的赋范空间及其上的算子.

定义 9.4. 按范数收敛也称为强收敛. 换言之, 赋范空间 X 中的序列 (x_n) 称为强收敛 (*converge strongly*) 到 x , 如果 $\lim_{n \rightarrow \infty} \|x_n - x\| = 0$.

定义 9.5. 赋范空间 X 中的序列 (x_n) 称为弱收敛 (*converge weakly*) 到 x , 如果对所有有界线性泛函 f , 有 $\lim_{n \rightarrow \infty} f(x_n) = f(x)$.

考虑从 X 到 Y 的有界线性算子空间 $B(X, Y)$. 通过定义如下范数我们可以使之成为一赋范空间.

$$\|T\| = \sup_{x \in X, \|x\|=1} \|Tx\|, \quad (9.0.1)$$

于是我们可以在 $B(X, Y)$ 上定义 3 种不同的收敛性. 实际上, 在当前的泛函分析中还有更多的收敛性.

定义 9.6. 设 X, Y 为赋范空间, (T_n) 为一列算子, $T_n \in B(X, Y)$.

- (i) (T_n) 一致收敛 (*uniformly convergent*), 如果它按范数收敛.
- (ii) (T_n) 强收敛 (*strongly convergent*), 如果对任意 $x \in X$, $(T_n x)$ 强收敛.
- (iii) (T_n) 弱收敛 (*weakly convergent*), 如果对任意 $x \in X$, $(T_n x)$ 弱收敛.

显然有一致收敛 \Rightarrow 强收敛 \Rightarrow 弱收敛, 并且在这 3 种情况下极限相同. 但反过来通常不成立.

同伦群

拓扑空间 X 最基本和明显的性质是它的连通分支的个数. 下一个这样的性质, 在某种意义上, 是 X 中洞的个数. 存在更高阶的类似物, 称为同伦群 (homotopy groups). 它们是拓扑不变量, 即在同胚下保持不变. 它们在场论和其他物理问题的拓扑考虑中具有重要的地位. 文章 *Topological Defects and Their Homotopy Classification* 和 *Electric-Magnetic Duality* 含有一些这样的例子.

定义 9.7. 给定拓扑空间 X , 它的第 0 个同伦集 (*homotopy set*) 是它的连通分支的集合, 记做 $\pi_0(X)$. 当 X 是连通的时, 有时记做 $\pi_0(X) = 0$.

为了定义 X 的基本群, 或 $\pi_1(X)$, 我们需要环路的概念. 它在其他方面也是有用的. 为简单起见, 我们将考虑带基点的环路 (即通过 X 上一个固定点的环路). 在绝大部分的应用中, 它们是相关的对象. 我们可以考虑具有各种不同光滑性的环路

(当 X 是一流形时), 但从量子场论应用的角度来看, 我们将考虑连续的环路. 它们也是拓扑中相关的对象.

定义 9.8. 给定拓扑空间 X 和一点 $x_0 \in X$, 一个闭 (带基点) 的环路 (closed based loop) 是一个从参数化的圆周到 X 的连续映射 $\xi: [0, 2\pi] \rightarrow X$ 满足 $\xi(0) = \xi(2\pi) = x_0$.

定义 9.9. 给定连通拓扑空间 X 和一点 $x_0 \in X$, 所有闭的带基点的环路构成的空间称为 X 的 (参数化的, 带基点的 (parametrized based)) 环路空间 (loop space), 记为 ΩX .

环路空间 ΩX 上具有从闭区间 $[0, 2\pi]$ 到 X 的连续映射空间诱导的紧致开拓扑. 它也有一个自然的基点: 将整个 $[0, 2\pi]$ 映到 x_0 的常值函数. 故而容易迭代这个构造而定义 $\Omega^k X, k \geq 1$.

这里我们选择用 $[0, 2\pi]$ 来参数化圆周. 如果用幅角来解释, 将更为自然. 我们也可以简单地选择单位区间 $[0, 1]$. 这或许和前面对于路径的定义和后面对于同伦的定义更为一致.

定理 9.0.1. 拓扑空间 X 的基本群 (fundamental group), 记做 $\pi_1(X)$, 由 X 中的那些闭环路的类组成. 这些类不能在保持基点不动的连续地形变下从一个变到另一个.

定义 9.10. 空间 X 称为单连通的 (simply-connected), 如果 $\pi_1(X)$ 平凡.

为了定义高阶同伦群, 让我们涉及一点同伦论的细节.

定义 9.11. 给定拓扑空间 X, Y 和映射 $p, q: X \rightarrow Y$, 我们说 h 是映射 p, q 间的一个同伦 (homotopy), 如果 $h: X \times I \rightarrow Y$ 是一个连续映射, 使得 $h(x, 0) = p(x), h(x, 1) = q(x)$. 这里 I 是单位区间. 在这种情况下我们记做 $p \approx q$.

定义 9.12. 一个映射 $f: X \rightarrow Y$ 是一个同伦等价 (homotopy equivalence), 如果存在映射 $g: Y \rightarrow X$, 使得 $g_0 f \approx \text{id}_X, f_0 g \approx \text{id}_Y$.

这是一个等价关系.

定义 9.13. 对于带基点 x_0 的拓扑空间 X , 我们定义 $\pi_n(X) (n \geq 0)$ 为从 n 维球面 S^n 到 X 的保持基点映射的同伦等价类的集合.

这和前面关于 π_0 和 π_1 的定义相容.

在同伦类和环路空间之间有一个非常好的关系.

定理 9.0.2. $\pi_n(X) = \pi_{n-1}(\Omega X) = \cdots = \pi_0(\Omega^n X)$.

当我们考虑 Yang-Mills 理论中的规范群 (gauge group) G 时, 它的基本群分类了这个理论可能出现的磁单极子 (monopole).

对于 $n \geq 1$, $\pi_n(X)$ 是一个群, 群运算由连接两条环路构成一条新环路而得. 另一方面, $\pi_0(X)$ 一般不是群. 但当 X 是李群时, $\pi_0(X)$ 从 X 继承了一个群结构. 这是因为它可以等同于 X 模掉单位连通分支而得的商群. 例如, $O(3)$ 的两个连通分支可以等同于群 \mathbf{Z}_2 的两个元素. 行列式等于 1 的分支对应于 \mathbf{Z}_2 中的 0, 行列式等于 -1 的分支对应于 \mathbf{Z}_2 中的 1.

对于 $n \geq 2$, $\pi_n(X)$ 总是一个 Abel 群.

某些 Riemann 面 (surface) 的基本群是非 Abel π_1 的例子.

由于 π_1 不一定是 Abel 群, 我们用于同伦群的许多直和记号应该被更正为乘法记号. 但是, 在数学、物理文献中, 人们普遍采用加法记号.

例 9.1. $\pi_n(X \times Y) = \pi_n(X) + \pi_n(Y), n \geq 1$

例 9.2. 对于球面, 我们有以下结果:

$$\begin{aligned} \pi_i(S^n) &= \begin{cases} 0 & \text{若 } i > n \\ \mathbf{Z} & \text{若 } i = n \end{cases} \\ \pi_i(S^1) &= 0 & \text{若 } i > 1 = 0 \\ \pi_{n+1}(S^n) &= \mathbf{Z}_2 & \text{若 } n \geq 3 \\ \pi_{n+2}(S^n) &= \mathbf{Z}_2 & \text{若 } n \geq 2 \\ \pi_6(S^3) &= \mathbf{Z}_{12} \end{aligned}$$

例 9.3. 从球丛理论可以得到:

$$\begin{aligned} \pi_i(S^2) &= \pi_{i-1}(S^1) + \pi_i(S^3) & \text{若 } i \geq 2 \\ \pi_i(S^4) &= \pi_{i-1}(S^3) + \pi_i(S^7) & \text{若 } i \geq 2 \\ \pi_i(S^8) &= \pi_{i-1}(S^7) + \pi_i(S^{15}) & \text{若 } i \geq 2 \end{aligned}$$

第一个关系给出以下更简洁的结果:

$$\pi_i(S^3) = \pi_i(S^2) \quad \text{若 } i \geq 3$$

例 9.4. Serre 的一个结果说, 事实上, 球面的所有同伦群都是有限的, 除了 $\pi_n(S^n)$ 和 $\pi_{4n-1}(S^{2n}), n \geq 1$.

定义 9.14. 给定一个连通空间 X , 一个映射 $\pi: B \rightarrow X$ 称为一个覆叠 (covering), 如果 (i) $\pi(B) = X$, (ii) 对于每个 $x \in X$, 存在 x 的一个连通开邻域 V , 使得 $\pi^{-1}(V)$ 的每个分支在 B 中开, 并且限制在每一个分支上 π 是一个同胚. 空间 B 称为覆叠空间 (covering space).

例 9.5. 实直线 \mathbf{R} 是群 $U(1)$ 的一个覆叠.

例 9.6. 群 $SU(2)$ 是群 $SO(3)$ 的一个二重覆叠.

例 9.7. 群 $SL(2, \mathbf{C})$ 是 Lorentz 群 $SO(1, 3)$ 的一个二重覆叠.

例 9.8. $SU(2, 2)$ 群是四维共形群 (conformal group) 的一个四重覆叠. 这个局部同构在扭量理论中非常重要.

通过考虑 X 中的闭环路以及它们在 B 中的覆叠, 容易看出基本群 $\pi_1(X)$ 作用在 X 的覆叠上. 如果我们进一步假设这个作用是可迁的 (transitive), 则有如下好的结果: X 的覆叠和 $\pi_1(X)$ 的正规子群 (normal subgroup) 1-1 对应.

给定一个连通空间 X , 总存在惟一的一个单连通的覆叠空间 \tilde{X} , 称为万有覆叠空间 (universal covering space). 而且, \tilde{X} 覆叠 X 的其他覆叠空间. 对于高阶同伦群, 有

$$\pi_n(X) = \pi_n(\tilde{X}), n \geq 2.$$

一类非常重要的同伦群是李群的同伦群. 为简单起见, 我们只考虑连通群, 即 $\pi_0(G) = 0$. 而且我们将主要考虑典型群, 特别地, 正交群和酉群.

定理 9.0.3. 设 G 为一个连通李群,

(i) 如果 G 是紧致、半单的, 则 $\pi_1(G)$ 是有限的. 这意味着 \tilde{G} 还是紧致的.

(ii) $\pi_2(G) = 0$.

(iii) 若 G 是紧致、单、非 Abel 的, 则 $\pi_3(G) = \mathbf{Z}$.

(iv) 若 G 是紧致、单连通、单的, 则 $\pi_4(G) = 0$ 或 \mathbf{Z}_2 .

例 9.9. $\pi_1(SU(n)) = 0$.

例 9.10. $\pi_1(SO(n)) = \mathbf{Z}_2$.

例 9.11. 因为酉群 $U(n)$ 在拓扑上是 $SU(n)$ 和圆周 S^1 的乘积, 它们的同伦群容易通过乘积公式计算. 我们知道 $U(1)$ 在拓扑上是圆周, $SU(2)$ 在拓扑上是 S^3 .

例 9.12. 对于 $n \geq 2$, 我们有

$$\pi_i(SO(3)) = \pi_i(SU(2))$$

$$\pi_i(SO(5)) = \pi_i(Sp(2))$$

$$\pi_i(SO(6)) = \pi_i(SU(4))$$

纯属为了兴趣, 并且展示这一领域的丰富, 我们在表 1 中给出一些同伦群间的同构, 在表 2 中给出低阶 $SU(n)$ 和 $SO(n)$ 的一些同伦群.

表 1 同伦群间的一些同构

同构	范围
$\pi_i(SO(n)) \approx \pi_i(SO(m))$	$n, m \geq i + 2$
$\pi_i(SU(n)) \approx \pi_i(SU(m))$	$n, m \geq 1/2(i + 1)$
$\pi_i(SP(n)) \approx \pi_i(SP(m))$	$n, m \geq 1/4(i - 1)$
$\pi_i(G_2) \approx \pi_i(SO(7))$	$2 \leq i \leq 5$
$\pi_i(F_4) \approx \pi_i(SO(9))$	$2 \leq i \leq 6$
$\pi_i(SO(9)) \approx \pi_i(SO(7))$	$i \leq 13$

表 2 低阶 $SU(n)$ 和 $SO(n)$ 的一些同伦群

	π_4	π_5	π_6	π_7	π_8	π_9	π_{10}
$SU(2)$	\mathbf{Z}_2	\mathbf{Z}_2	\mathbf{Z}_{12}	\mathbf{Z}_2	\mathbf{Z}_2	\mathbf{Z}_3	\mathbf{Z}_{15}
$SU(3)$	0	\mathbf{Z}	\mathbf{Z}_6	0	\mathbf{Z}_{12}	\mathbf{Z}_3	\mathbf{Z}_{30}
$SU(4)$	0	\mathbf{Z}	0	\mathbf{Z}	\mathbf{Z}_{24}	\mathbf{Z}_2	$\mathbf{Z}_{120} + \mathbf{Z}_2$
$SU(5)$	0	\mathbf{Z}	0	\mathbf{Z}	0	\mathbf{Z}	\mathbf{Z}_{120}
$SU(6)$	0	\mathbf{Z}	0	\mathbf{Z}	0	\mathbf{Z}	\mathbf{Z}_3
$SO(5)$	\mathbf{Z}_2	\mathbf{Z}_2	0	\mathbf{Z}	0	0	\mathbf{Z}_{120}
$SO(6)$	0	\mathbf{Z}	0	\mathbf{Z}	\mathbf{Z}_{24}	\mathbf{Z}_2	$\mathbf{Z}_{120} + \mathbf{Z}_2$
$SO(7)$	0	0	0	\mathbf{Z}	$\mathbf{Z}_2 + \mathbf{Z}_2$	$\mathbf{Z}_2 + \mathbf{Z}_2$	\mathbf{Z}_{24}
$SO(8)$	0	0	0	$\mathbf{Z} + \mathbf{Z}$	$\mathbf{Z}_2 + \mathbf{Z}_2 + \mathbf{Z}_2$	$\mathbf{Z}_2 + \mathbf{Z}_2 + \mathbf{Z}_2$	$\mathbf{Z}_{24} + \mathbf{Z}_{24}$
$SO(9)$	0	0	0	\mathbf{Z}	$\mathbf{Z}_2 + \mathbf{Z}_2$	$\mathbf{Z}_2 + \mathbf{Z}_2$	\mathbf{Z}_{24}
$SO(10)$	0	0	0	\mathbf{Z}	\mathbf{Z}_2	$\mathbf{Z} + \mathbf{Z}_2$	\mathbf{Z}_{12}

Part III

半经典理论

Chapter 10

Cauchy 动力学与奇点定理

10.1 Witten 迅捷性

10.2 Penrose 过程

10.3 能量条件

10.4 奇点定理

陷俘面

10.5 宇宙监督假设

10.6 高-Wald 定理

Part IV

附录

附录 A

量纲理论

附录 B

量子力学

量子力学主要使用的数学知识是泛函分析，特别是有关 Hilbert 空间、线性算子的内容。泛函分析研究的是函数和函数空间，它的研究综合了来自经典分析和代数的技巧。近代泛函分析的发展是围绕那些其解为函数的方程问题展开的。在 18 世纪研究了常微分、偏微分方程之后，19 世纪出现了对积分方程和其他泛函方程的讨论，在它的后期就产生了对一种新的分析的需要，这种分析所针对的是具有无穷个变量的函数而不是通常的函数。1887 年，Volterra 受到变分法的启发，提出一种新的无穷小方法，它将通常的函数用泛函（从函数空间到 \mathbb{R} 或 \mathbb{C} 的映射）来代替，但是他和他的后继者缺少了以后才出现的一些代数和拓扑的工具来继续发展。近代分析的产生伴随着无穷维线性代数的进展，其与经典的线性代数密切相关联的，后者到了 1890 年（直到对偶的概念，这是以后才有的）已经扎实地建立起来了。Fredholm 强烈地受到代数方法的启示，在 19 世纪时的工作中出现了算子的核概念，这成为了近代积分方程理论的一个基石。Hilbert 对于对称核比 Fredholm 又进了一步，他揭示了与实二次型理论的类似性，从而使得平方可积函数这一概念的重要性清晰了。Hilbert 的工作《积分方程一般理论之基础》(*Grundzüge einer allgemeinen Theorie der Integralgleichung*)跨出了从无穷代数到无穷几何的一步。Fréchet 的贡献则在于引入了距离空间的抽象概念，从而使得从欧氏几何转向从那时起就称之为 Hilbert 空间的框架成为可能。

泛函分析在量子系统研究中的价值是在 20 世纪 50 年代明确的，当时 Kato 证明了原子的 Hamilton 量的自共轭性，随后 Garding 和 Wightman 刻画了量子场理论中的公理。从此泛函分析就与量子场理论许多方面的核心相联系了。

B.0.1 Hilbert 空间

定义 B.1 (复内积空间). 复矢量空间 V 称为复内积空间，简称内积空间，若存在内积映射 $(\cdot, \cdot) : V \times V \rightarrow \mathbb{C}$ ，对任意 $f, g, h \in V$ 和任意 $c, d \in \mathbb{C}$ 满足：

(第二槽线性性) $(f, cg + dh) = c(f, g) + d(f, h)$;

(交换律) $(f, g) = \overline{(g, f)}$, 其中 $\overline{(g, f)}$ 代表复数 (g, f) 的共轭复数;

(非退化) $(f, f) \geq 0$, 且 $(f, f) = 0 \Leftrightarrow f = 0$. 注意, 只含零元的矢量空间也可构成内积空间, 但下面只讨论维数大于零的情况.

定理 B.0.1 (共轭线性). 非退化条件配以线性性和交换律可得到 $(f+g, h) = (f, h) + (g, h)$ 及 $(cf, g) = \bar{c}(f, g)$. 这说明内积对第一槽有“反线性”, 称为共轭线性.

定理 B.0.2. 对非退化条件中取 $g = f$ 可有: $\forall g \in V, (f, g) = 0 \implies f = 0$.

定义 B.2. 由于 $[a, b] \in \mathbb{R}$ 有无穷多自由度,

$$C[a, b] = \{[a, b] \subset \mathbb{R} \text{ 上连续复值函数 } f(x)\}$$

是无限维复矢量空间. 定义内积

$$(f, g) := \int_a^b \overline{f(x)}g(x)dx, \quad \forall f, g \in C[a, b], \quad (\text{B.0.1})$$

请读者自行验证 $C[a, b]$ 是内积空间.

人们遇到的拓扑空间最多的实际上是距离空间. 内积空间可自然诱导出矢量空间中任意两点的距离, 使之成为距离空间, 进而再诱导空间的拓扑. 线性泛函分析起始于拓扑矢量空间, 就是一个矢量空间赋予了一个拓扑, 对这个拓扑其(线性结构的)运算是连续的. 通常将任意两元素 f 和 g 的距离定义为

$$d(f, g) := \sqrt{(f - g, f - g)}.$$

用开球给 V 定义通常拓扑, V 可自然地定义为一个拓扑空间. 今后凡涉及 V 的拓扑时一律指这一拓扑.

我们曾学过有限维实矢量空间 V 的对偶空间 V^* , 它是由 V 到 \mathbb{R} 的全体线性映射的集合. 对任意复矢量空间 V , 对偶矢量可定义为由 V 到 \mathbb{C} 的线性映射. 然而内积空间除了是复矢量空间外还是拓扑空间, 因此还可讨论映射 $\eta: V \rightarrow \mathbb{C}$ 的连续性. 这时还涉及 \mathbb{C} 的拓扑, 其定义也很自然: 设 $z_1, z_2 \in \mathbb{C}$ 且 $z_1 = a_1 + ib_1, z_2 = a_2 + ib_2$ (其中 $a_1, a_2, b_1, b_2 \in \mathbb{R}$), 则 z_1, z_2 的距离可定义为

$$d(z_1, z_2) := \sqrt{(a_1 - a_2)^2 + (b_1 - b_2)^2},$$

用此距离便可定义开球, 从而定义 \mathbb{C} 的拓扑.

每个连续的线性映射(对偶矢量) $\eta: V \rightarrow \mathbb{C}$ 在泛函分析领域里又可称为 V 上的一个连续线性泛函 (continuous linear functional). 我们关心 V 上全体连续线性泛函的集合(它有许多好的性质), 并称它为 V 的对偶空间. 可以证明, 若 V 为有限维, 则 V 上的线性泛函必定连续. 因此只当 V 为无限维时连续性才是对线性泛函有实质意义的要求.

定义 B.3 (共轭空间). 内积空间 V 的对偶空间又称共轭空间, 定义为

$$V^* := \{\eta : V \rightarrow \mathbb{C} \mid \eta \text{ 为连续的线性映射}\}.$$

V^* 也可看作复矢量空间, 为此只须用如下的自然方法定义附加结构:

$$(\text{加法}) \quad (\eta_1 + \eta_2)(f) := \eta_1(f) + \eta_2(f), \quad \forall \eta_1, \eta_2 \in V^*, f \in V;$$

$$(\text{数乘}) \quad (c\eta)(f) := c \cdot \eta(f), \quad \forall \eta \in V^*, f \in V, c \in \mathbb{C};$$

$$(\text{零元}) \quad \text{零元 } 0 \in V^* \text{ 作用于任意 } f \in V \text{ 都得零.}$$

读者至此自然会联想到右矢 (*ket*) 和左矢 (*bra*), 并猜想右矢空间是内积空间 V , 而左矢空间则是 V^* . 然而事情比此略为复杂. 要保证 *Dirac* 左右矢记号 (*bra-ket*)¹ 自洽性, 左、右矢空间应该“完全对等”. 这似乎不难: 我们学过, 有限维实矢量空间 V 上的度规诱导了从 V 到 V^* 的音乐同构. 然而, 由于复空间的内积与实空间的度规少许不同, 复空间 V 上的内积自然诱导的从 V 到 V^* 的映射不是线性而是反线性的. 不过这不构成什么问题. 真正构成问题的是量子力学中用到的内积空间多数是无限维的, 而这导致上述映射未必满射, 即 V 与 V^* “不一样大”.

定理 B.0.3. 内积映射 (\cdot, \cdot) 自然诱导出一个单射的、反线性的映射 $v : V \rightarrow V^*$.

证. 设 $f \in V$, 则 $\eta_f \equiv (f, \cdot)$ 是从 V 到 \mathbb{C} 的映射. 由内积定义可知它是线性的. 还可证明它是连续的, 因此 $\eta_f \in V^*$. 具体说, η_f 是 V^* 的这样一个元素, 它作用于 $g \in V$ 的结果为 $\eta_f(g) := (f, g)$. 可见 (\cdot, \cdot) 自然诱导出一个映射 $v : V \rightarrow V^*$, 定义为 $v(f) := \eta_f$. 容易证明映射 v 的单射性和反线性性. \square

若 V 是有限维空间, 则 V^* 与 V 有相同维数, 因而上述反线性映射 $v : V \rightarrow V^*$ 一定满射, 因此双射构成同构. 若 V 是无限维的, 则 V^* 也是无限维的, 这时 $v : V \rightarrow V^*$ 不一定满射. 直观地说, V^* 有可能“比 V 大”, 即 $v[V] \subset V^*$ 但 $v[V] \neq V^*$. 然而, 为保证 *Dirac* 符号给出正确结果, 我们需要 $v[V] = V^*$.

为了用准确语言表述这些思考, 先讨论内积空间的完备性. 我们学过, 收敛列一定是 *Cauchy* 列, 但反之不然, 因为距离空间不一定完备.

定义 B.4 (Hilbert 空间). 完备内积空间称为 *Hilbert* 空间 \mathcal{H} .

定义 B.5. 不完备内积空间是很容易构造的, 比如 $C[a, b]$ 就是一例. 为使之完备化, 可把 $[a, b]$ 上虽不连续却平方可积的复值函数包含进去, 扩大后的空间记作 $L^2[a, b]$, 即

$$L^2[a, b] = \left\{ f : [a, b] \rightarrow \mathbb{C} \mid f \text{ 满足 } \int_a^b |f(x)|^2 dx < \infty \right\}.$$

¹bra-ket 由括号的英文 bracket 得来.

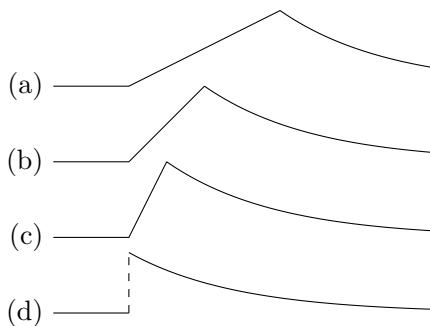


图 B.1: 不完备内积空间的一个直观解释. 图中 (a), (b), (c) 代表 $C[a, b]$ 中的某一 Cauchy 列 $\{f_n(x)\}$ 的三个元素 (其序号 n 依次增大). 序列 $\{f_n(x)\}$ 不收敛于 $C[a, b]$ 的任何元素. 或者说它收敛于 (d) 代表的函数 $f(x)$, 但 $f(x) \notin C[a, b]$, 因为它不连续.

这里 $< \infty$ 是指积分收敛. 积分具体指 Lebesgue 积分. 严格地说 $f(x)$ 还应具“可测性”, 然而物理上遇到的函数都满足这一要求. 注意, 两个函数若只在测度为零的集上 (比如有有限个点) 不同, 应视为 $L^2[a, b]$ 的同一元素. 内积仍用式 (B.0.1) 定义, 可证其完备性.

事实上, 任何不完备的内积空间 V 都可以完备化, 为此只须把它略加扩大: 把所有 Cauchy 列“应有的极限点”都补进 V 中. 可以证明, 对任何不完备内积空间 V , 总可找到完备的内积空间 \tilde{V} , 使得 $V \subset \tilde{V}$ 而且 $\bar{V} = \tilde{V}$, 其中 \bar{V} 是 \tilde{V} 作为拓扑空间时子集 V 的闭包. 因此直观地说 \tilde{V} 比 V 至多“多层皮”.

由于有限维内积空间一定完备, 因此一定是 Hilbert 空间. 然而量子力学中用到的 Hilbert 空间多数是无限维的. 无限维会让许多有限维成立的结论变得微妙. 同 $L^2[a, b]$ 相仿,

$$L^2(\mathbb{R}^n) \equiv \left\{ f: \mathbb{R}^n \rightarrow \mathbb{C} \mid f \text{ 满足 } \int_{\mathbb{R}^n} |f|^2 d^n x < \infty \right\} \quad (\text{B.0.2})$$

也是 Hilbert 空间, 这里内积定义仿式 (B.0.1), 其中 $L^2(\mathbb{R}^3)$ 就是所谓的波函数空间. 由于具有完备性, Hilbert 空间有许多良好性质, 其中有用的就是 \mathcal{H} 与 \mathcal{H}^* “一样大”, 即 $v[\mathcal{H}] = \mathcal{H}^*$.

定理 B.0.4. $\forall \eta \in \mathcal{H}^*, \exists ! f_\eta \in \mathcal{H}$ 使 $\eta(g) = (f_\eta, g), \forall g \in \mathcal{H}$.

证. 可从 Riesz 表现定理的证明中得到启发, 略. □

这表明, 对 Hilbert 空间 \mathcal{H} , 定理 B.0.3 的 $v: \mathcal{H} \rightarrow \mathcal{H}^*$ 是满射, 即 v 是反线性双射. 这个命题的重要性在于保证 \mathcal{H} 同 \mathcal{H}^* “一样大”, 请注意不完备内积空间没有此结论. 可见物理学不但需要内积空间, 而且需要 Hilbert 空间. 利用映

射 $v: \mathcal{H} \rightarrow \mathcal{H}^*$ 还可把 \mathcal{H}^* 定义为 Hilbert 空间: $\forall \eta, \xi \in \mathcal{H}^*$, 可知有唯一的 $f_\eta, f_\xi \in \mathcal{H}$ 使 $\eta = v(f_\eta), \xi = v(f_\xi)$. 定义 η 和 ξ 的内积为

$$(\eta, \xi) := (f_\xi, f_\eta), \quad (\text{B.0.3})$$

则不难验证 (η, ξ) 满足内积定义, 故 \mathcal{H}^* 是内积空间. 还可证明 \mathcal{H}^* 是 Hilbert 空间. 可见 \mathcal{H}^* 与 \mathcal{H} 有一种“认同”: 其间不但存在反线性双射, 且这一映射还在上式的意义上保内积. \mathcal{H} 和 \mathcal{H}^* 的这种认同性使我们可以把它们分别用作右矢和左矢空间, 将在后面揭晓细节.

设 V 是不完备内积空间, 下面说明 V^* 仍可被自然地定义为 Hilbert 空间. 以 \mathcal{H} 代表把 V 完备化后所得 Hilbert 空间, 则 $\mathcal{H} = \bar{V}$. 按如上讨论知 \mathcal{H}^* 也是 Hilbert 空间. \mathcal{H}^* 的任一元素 η 是 \mathcal{H} 上的连续线性泛函, 把它的作用范围限制在 $V \subset \mathcal{H}$, 使得 V 上的一个连续线性泛函, 记作 $\tilde{\eta} \in V^*$. 可见存在映射 $\beta: \mathcal{H}^* \rightarrow V^*$, 定义为 $\beta(\eta) := \tilde{\eta}$. 由 $\mathcal{H} = \bar{V}$ 并借助 Cauchy 列可证明 β 是线性双射. 于是可用 β 将 V^* 和 \mathcal{H}^* 认同, 从而使 V^* 也获得 Hilbert 空间的结构. 总之, 无论内积空间 V 是否完备, 其对偶空间 V^* 一定是 Hilbert 空间. 若 V 不完备, 则 $v[V] \neq V^*$ 而 $\overline{v[V]} = V^*$, 此即说 V^* 比 V 至多“多层皮”; 若 V 完备则 $v[V] = V^*$. 可见, 只要给 V 适当地“补上这层皮”, $v: V \rightarrow V^*$ 就可满射.

例 B.1. 第 ?? 章讨论过 δ 函数. 积分具线性性质, 可见 δ 函数对应于一个线性泛函. 实际上, δ 函数对应于一个特定的函数空间 K (带拓扑) 上的连续线性泛函. 用以定义广义函数的那个函数空间 K 定义为

$$K := \{\varphi: \mathbb{R} \rightarrow \mathbb{C} \mid \varphi \text{ 光滑且支集有界}\}. \quad (\text{B.0.4})$$

这里函数 $\varphi(x)$ 的支集定义为子集 $\{x \in \mathbb{R} \mid \varphi(x) \neq 0\}$ 的闭包. 再在其上定义适当的拓扑 (定义了这一拓扑的 K 称为基本空间), 则便可讨论 K 上的线性泛函的连续性. K 上的每个连续线性泛函称为一个广义函数, 亦称一个分布 (*distribution*). 可证 \mathbb{R} 上任一局部可积¹函数 $f(x)$ 都可看作一个广义函数 η , 更确切地说, $f(x)$ 按下式定义的泛函

$$\eta_f(\varphi) = \int_{\mathbb{R}} f(x)\varphi(x) dx, \quad \forall \varphi \in K \quad (\text{B.0.5})$$

是广义函数, 称为函数型广义函数, 可见广义函数可看作普通函数在某种意义上的推广. 另一方面, 可证 $\forall x_0 \in \mathbb{R}, f \in K, \delta_{x_0}(f) := f(x_0)$ 是一个广义函数, 而这又可形式地表为

$$\delta_{x_0}(f) = \int_{\mathbb{R}} \delta(x - x_0) f(x) dx, \quad (\text{B.0.6})$$

¹函数 $f(x)$ 在 \mathbb{R} 的每一有界区域内可积, 则称为局部可积的.

但并不存在任何局部可积函数 $f(x)$ 满足 $\eta_f = \delta_{x_0}$, 因此 δ_{x_0} 属于非函数型广义函数. 对广义函数还可定义导函数, 可以证明广义函数都有导函数且无限阶可微, 甚至导函数都是广义函数, 在此不做赘述. 由于 *Dirac* 在讨论量子力学时经常使用 δ 函数, 人们总以为 δ 函数与量子力学密切相关. 初学者往往误以为它是 *Hilbert* 空间 \mathcal{H} 上的连续线性泛函. 如果果真如此, δ 函数就属于 \mathcal{H}^* , 由定理 B.0.4 知 \mathcal{H} 必有元素与之对应, 这样一来 δ 就成为函数型广义函数, 即普通意义的函数了. 事实上, 广义函数是基本空间 K 上的连续线性泛函, 但基本空间一定不是 *Hilbert* 空间. 其实, δ 函数在量力学中并不至关, *von Neumann* 等人建立的一整套量力的严密数学基础根本不涉及 δ 函数, 详见下一节.

接下来讨论 *Hilbert* 空间的正交归一基. N 维向量空间 V 的一个基底无非是由 N 个元素组成的、满足如下两个要求的一个序列 $\{e_1, \dots, e_N\}$: (1) $\{e_1, \dots, e_N\}$ 线性独立; (2) V 的任一元素 f 可由 $\{e_1, \dots, e_N\}$ 线性表出. 我们想把基底概念推广至无限维的 *Hilbert* 空间 \mathcal{H} . 定义 \mathcal{H} 的有限子集 $\{f_1, \dots, f_N\}$ 称为线性独立的, 若

$$\sum_{n=1}^N c_n f_n = 0 \Rightarrow c_n = 0, n = 1, \dots, N.$$

\mathcal{H} 的任一子集 $\{f_\alpha\}$ 称为线性独立的, 若 $\{f_\alpha\}$ 的任一非空有限子集线性独立. 如果 \mathcal{H} 中存在满足以下两条条件的无限序列 $\{e_n\}$: (1) $\{e_n\}$ 线性独立; (2) \mathcal{H} 的任一元素 f 可由 $\{e_n\}$ 线性表出:

$$f = \sum_{n=1}^{\infty} c_n e_n, \quad c_n \in \mathbb{C}$$

就说 $\{e_n\}$ 构成 \mathcal{H} 的一个基底. (上式中的 $f = \sum_{n=1}^{\infty} c_n e_n$ 是 $\lim_{N \rightarrow \infty} \sum_{n=1}^N c_n e_n$ 的简写.)

B.0.2 赋范空间

一个拓扑向量空间赋予了一个局部基, 它的元是凸的, 就称之为局部凸的. 局部凸空间的例子是赋范空间, 即向量空间赋予了一个范数, 这个概念首次出现在 *Fréchet* 的工作中:

定义 B.6 (范数). 在一个实数或复数域上的向量空间 V 上, 一个函数 $\|\cdot\|: V \rightarrow \mathbf{R}^+$ 称为一个范数 (*norm*), 如果它满足:

(非退化) $\|x\| = 0$ 当且仅当 $x = 0$,

(数乘) $\|\alpha x\| = |\alpha| \|x\|$,

(三角不等式) $\|x + y\| \leq \|x\| + \|y\|$.

半范数是一个映射 $\rho: V \rightarrow [0, \infty)$, 它服从三角不等式和数乘性, 但不一定满足非退化.

易见利用矢量空间上的范数可以定义一个平移不变的度量

$$d(x, y) = \|x - y\|,$$

使之成为一个距离空间. 在许多情况下, 例如对于例 2.8 中定义的度量, 我们可以把一个矢量的范数定义成它到原点的距离 (一个显然的例外是离散度量). 其合理性自明, 这样, 勾股定理所定义的距离, 就可看作欧氏内积 (度规) 诱导的度量: $|x - y| = \langle x - y, x - y \rangle$. 于是在 \mathbb{R}^n 的通常拓扑下, 度规和距离就仿若同一事物了.

历史上最早的赋范空间例子之一是 $C([0, 1])$ 空间, 它是 Riesz 所讨论的区间 $[0, 1]$ 上的连续复值函数空间, 赋以最大值范数 $\|f\|_\infty := \sup_{x \in [0, 1]} |f(x)|$. 20 世纪 20 年代, 伴随着 Hahn 和 Banach 的工作产生了 Banach 空间的一般定义. 一个赋范线性空间是一个 Banach 空间, 如果作为一个由它诱导出的距离的距离空间是完备的, $C([0, 1])$ 是一个典型的 Banach 空间. 更为一般的, 对任意非负整数 k , $[0, 1]$ 上的 C^k 类函数空间 $C^k([0, 1])$, 赋予范数 $\|f\|_k = \sum_{i=0}^k \|f^{(i)}\|_\infty$, 即被表示成有限个半范数 $\|f^{(i)}\|_\infty = \sup_{x \in [0, 1]} |f^{(i)}(x)|, i = 0, \dots, k$ 的项之和, 它也是 Banach 空间.

区间 $[0, 1]$ 上的光滑函数空间 $C^\infty([0, 1])$ 就不再是 Banach 空间, 这是由于它的拓扑是由可数个半范数族 $\|f\|_k$ 所刻画的, 其中 k 在正整数范围内变化. 距离

$$d(f, g) = \sum_{k=1}^{\infty} 2^{-k} \frac{\|f - g\|_k}{1 + \|f - g\|_k}$$

使得它成为了一个 Fréchet 空间, 就是一个局部凸的完备距离空间. 对速降函数空间 $\mathcal{S}(\mathbb{R}^n)$, 其元素速降函数是 \mathbb{R}^n 上的光滑函数 f 并且

$$\|f\|_{\alpha, \beta} := \sup_{x \in \mathbb{R}^n} |x^\alpha D_x^\beta f(x)|$$

对每个多重指标 α 和 β 都是有穷的, $\mathcal{S}(\mathbb{R}^n)$ 也是 Fréchet 空间, 它的拓扑是由半范数 $\|\cdot\|_{\alpha, \beta}$ 给定的. Fréchet 空间除此之外的例子是具有一个固定的紧支集 $K \subset \mathbb{R}^n$ 的光滑函数空间 $C_o^\infty(K)$, 赋以可数的半范数

$$\|D^\alpha f\|_{\infty, K} = \sup_{x \in K} |D_x^\alpha f(x)|, \alpha \in \mathbb{N}_0^n$$

另一个空间是一个闭流形 M 上的矢量丛 E 的光滑截面空间 $C^\infty(M, E)$, 赋予一个类似的可数半范数族. 给定一个开子集 $\Omega = \bigcup_{p \in \mathbb{N}} K_p$, 其中 $K_p (p \in \mathbb{N})$ 是 \mathbb{R}^n 的紧子集, 赋予空间 $\mathcal{D}(\Omega) = \bigcup_{p \in \mathbb{N}} C_o^\infty(K_p)$ 归纳极限拓扑 $\mathcal{D}(\Omega)$ 中的序列 (f_n) 收敛到 $f \in \mathcal{D}(\Omega)$, 如果每个 f_n 的支集在某个固定的紧子集 K 中并且对每个多重指标 $\alpha, (D^\alpha f_n)$ 在 K 上一致收敛于 $D^\alpha f$. $\mathcal{D}(\Omega)$ 是一个局部凸空间.

Banach 空间中的 Hilbert 空间具有和有穷维空间很相似的性质, 它在历史上是在 20 世纪初出现在 Hilbert 工作中的第一类无穷维空间. 一个 Hilbert 空间是一个 Banach 空

(1) 原文此处有误单射译者注. 间, 它有一个从内积导出的范数 $\|\cdot\|$, 就是 $\|u\|^2 = \langle u, u \rangle$, $\langle \cdot, \cdot \rangle$ 是一个正定的观线性 (或共轭双线性, 这要根据基空间是实的还是复的来确定) 型. Hilbert 空间是量子力学结构的基本成分; 使用 (闭的) 张量乘积, 人们从一个 Hilbert 空间 H 出发构造出 Fock 空间 $\mathcal{F}(H) = \sum_{k=0}^{\infty} \otimes^k H$, 和玻色子 Fock 空间 $\mathcal{F}(H) = \sum_{k=0}^{\infty} \otimes_s^k H$ (其中 \otimes) 是 (闭的) 对称化的张量乘积) 以及费米子 Fock 空间 $\mathcal{F}(H) = \sum_{k=0}^{\infty} \wedge^k H$ (这里 \wedge 是反对称化的 (闭) 张量乘积).

一个典型的 Hilbert 空间是使得 $\sum_{n \in \mathbb{Z}} |u_n|^2$ 有穷的复值序列 $(u_n)_{n \in \mathbb{Z}}$ 的空间 l_2 (\mathbb{Z}), 这个空间已经暗含在 Hilbert 的工作 “基础” (Grundzüge) 中了. 这之后很快, 借助于 Lebesgue 引人的积分工具, Riesz 和 Rischer 证明了区间 $]0, 1[$ 上的平方可和函数. 也就是满足

$$\|f\|_{L^2} = \left(\int_0^1 |f(x)|^2 dx \right)^{1/2}$$

为有穷的函数 f , 构成的空间 $L^2(]0, 1[)$ (首次由 Riesz 所引入) 是 Hilbert 空间. 这些区进一步推广到 $]0, 1[$ 上的 p 一次可和 ($1 \leq p < \infty$) 函数 (即那些满足

$$\|f\|_{L^p} = \left(\int_0^1 |f(x)|^p dx \right)^{1/p}$$

为有穷的函数 f) 空间 $L^p(]0, 1[)$, 除了 $p = 2$ 它们不是 Hilbert 空间, 但却提供了 Banach 空间的另一些例子, $]0, 1[$ 上的按 Lebesgue 测度几乎处处有界的函数空间 $L^\infty(]0, 1[)$ 则是 Banach 空间的另一个例子.

1936 年, Sobolev 用分部积分的方法推广了函数和它们的导数的概念, 这就引导出了所谓的 Sobolev 空间 $W^{k,p}(]0, 1[)$, 它包含的函数 $f \in L^p(]0, 1[)$ 且直到 k 阶的导数也在 $L^p(]0, 1[)$ 中, 空间是作为 $C^\infty(]0, 1[)$, 以

$$f \mapsto \|f\|_{W^{k,p}} = \left(\sum_{j=0}^k \|\partial^j f\|_{L^p}^p \right)^{1/p}$$

为范数取得的闭包 ($p = 2$ 时, $W^{k,p}(]0, 1[)$ 是 Hilbert 空间, 通常记成 $H^k(]0, 1[)$). 空间 $W^{k,p}(]0, 1[)$ 不同于 Sobolev 空间 $W_0^{k,p}(]0, 1[)$, 与它相对应的是集合 $D(]0, 1[)$ 对于范数 $f \mapsto \|f\|_{W^{k,p}}$ 的闭包; 例如 $u \in W^{1,p}(]0, 1[)$ 是 $W_0^{1,p}(]0, 1[)$ 中的元当且仅当它在 0 和 1 处为零, 即它在区间的边界上满足 Dirichlet 型边界条件. 类似地, 我们可以定义 \mathbb{R} 上的 Sobolev 空间 $W_0^{k,p}(\mathbb{R}) = W^{k,p}(\mathbb{R})$, 开子集 $\Omega \subset \mathbb{R}^n$ 上的 Sobolev 空间 $W^{k,p}(\Omega)$ 和 $W_0^{k,p}(\Omega)$, 以及在一个闭流型 M 上使用单位分解以定义 M 上

矢量从 E 截面的 Sobolev 空间 $H^k(M, E) = W^{k,2}(M, E)$. 用 Fourier 变换 (以后讨论), 我们可以去掉 k 是一个整数的假设, 从而把 Sobolev 空间的概念扩充到定义 $W^{s,p}(\Omega)$ 和 $H^s(M, E)$, 其中 s 是任意实数.

Sobolev 空间现在许多数学领域; 概率理论的一个重要的例子是 Cameron-Martin 空间 $H^1([0, t])$, 它嵌入到 Wiener 空间 $C([0, t])$ 中. 这个嵌入是更为一般的 Sobolev 嵌入定理的一个特例, 定理将 Sobolev 空间 $W^{k,p}$ 嵌入到 (可能是连续的, 有时甚至是紧的 (紧算子的概念将在以后的章节中讨论)) 空间 L^q 中, 这里 $q > p$, 且有这样的连续包含关系 $W^{k,p}(\mathbb{R}^n) \subset L^q(\mathbb{R}^n)$, 其中 $1/q = 1/p - k/n$, 或者它包含在 C^l 空间中, $l \leq k$, 即对于一个有界开并充分正规的 \mathbb{R}^n 的子集 Ω 以及任意 $s \geq l + n/p, p > n$ 有连续包含关系 $W^{s,p}(\Omega) \subset C^l(\bar{\Omega})$ (是 $C^l(\Omega)$ 中的函数的集合, $D^\alpha u$ 可以对所有的 $|\alpha| < l$ 连续地延拓到闭包 $\bar{\Omega}$ 上). Sobolev 嵌入定理在讨论偏微分方程解的正则性时有重要的应用, 证明所得到的弱解实际上是光滑的. 特别地, 在一个 n 维的闭流形 M 上, 如 $s > l + n/2$, 则 Sobolev 空间 $H^s(M, E)$ 可以连续地嵌入到 C^l 类的 E 的截面空间 $C^l(M, E)$ 中, 这就意味着超椭圆偏微分方程 $Au = v, v \in L^2(M, E)$ 的解是光滑的, 例如 Seiberg-Witten 方程解.

B.0.3 对偶性

对偶性的概念 (在拓扑的意义上) 起始于 20 世纪之初的 Hadamard, 他在寻找紧区间 I 上的连续函数赋以一致拓扑的 Banach 空间 $C(I)$ 上的连续线性泛函. 它暗含在了 Hilbert 的理论之中, 并且在 Riesz 的工作中是一个重要的部分, 他设法将这个连续泛函表示成 Stieltjes 积分, 这是近代积分理论的出发点之一.

一个拓扑矢量空间 E 的拓扑对偶是 E 上的连续线性型组成的空间 E^* , 当 E 是一个赋范空间时, 连续线性型可以赋予对偶范数 $\|L\|_{E^*} = \sup_{u \in E, \|u\| \leq 1} |L(u)|$.

对偶空间经常带来新的格外的研究对象; 任意一个函数 $f \in L^p(\mathbb{R}^n) (p \geq 1)$ 和在点 $x \in \mathbb{R}^n$ 的 δ 函数, $\delta_x : f \rightarrow f(x)$ 都在空间 $\mathcal{S}(\mathbb{R}^n)$ 中, 它是与 $\mathcal{S}(\mathbb{R}^n)$ 对偶的 \mathbb{R}^n 上的缓增分布, 它自身包含在与 $\mathcal{D}(\mathbb{R}^n)$ 对偶的分布空间 $\mathcal{D}'(\mathbb{R}^n)$ 中. 进而, 核空间 E 的拓扑对偶 E^* 包含有一个概率测度的支集, 它具有由一个 E 上的连续正定函数给出的特征函数 (见下一节). 核空间中的投影极限是 $E = \bigcap_{p \in \mathbb{N}} H_p$ (一个序列 $(u_n) \in E$ 收敛于 $u \in E$ 只要在每个 H_p 中都收敛于 u), 这里可数多个 Hilbert 空间的网 $\cdots \subset H_p \subset H_{p-1} \subset \cdots \subset H_0$, 使得嵌入 $H_p \subset H_{p-1}$ 是一个迹类算子 (见 “算子代数” 一节). 如果 H_p 是 E 对于范数 $\|\cdot\|_p$ 的闭包, E 拓扑对偶 E' 是一个归纳极限 $E' = \bigcup_{p \in \mathbb{N}_0} H_{-p}$, 这里 H_{-p} 是 (H_p) 的对偶 Hilbert 空间以 $\|\cdot\|_{-p}$ 为范数 (E' 中的序列 (u_n) 收敛于 $u \in E'$ 只要它在某个 H_{-p} 中且按 H_{-p} 的拓扑收敛于 u), 我们有

$$\begin{aligned} E &\subset \cdots \subset H_p \subset H_{p-1} \subset \cdots \subset H_0 \\ &= H'_0 \subset H_{-1} \subset \cdots \subset H_{-p} \subset \cdots \subset E' \end{aligned}$$

一个闭流形上的椭圆算子理论的一个结果是: 一个闭流形 M 上的矢量丛的光滑截面 Fréchet 空间 $C^\infty(M, E)$ 是核的, 它可数多个 Sobolev 空间 $H^p(M, E)$ 的投影极限, 而它的对偶由可数多个 Sobolev 空间 $H^{-p}(M, E)$ 的归纳极限给出 $\oplus^{\mathbb{N}}$.

一个赋范线性空间 E 上非零的连续线性型的存在性是由 Hahn-Banach 定理予以确定的, 该定理说对 E 的任意闭的线性子空间 F , 存在一个非零的连续线性型, 它在 F 上为零. 当空间是一个 Hilbert 空间 $(H, \langle \cdot, \cdot \rangle_H)$ 时, 从 Riesz-Fréchet 定理得出, H 上的任一连续线性型 L 都可以惟一地用一个矢量 $v \in H$ 来表示, 使得 $L(u) = \langle v, u \rangle_H$ 对所有的 u 成立, 这样偶对的左方和右方作 Hilbert 内积, 从而将拓扑对偶 H^* 等同于了 H .

在一个赋范矢量空间上由范数导出的强拓扑单射在这个拓扑下, 序列 (u_n) 收敛于 u 是指 $\|u_n - u\| \rightarrow 0$ 单射在 E 是无穷维的情形, 这对于紧集它是太过细了, 因为对于强拓扑单位球的紧性是有穷维空间的特性. 由于紧集在存在性定理中是有用的, 人们有意弱化拓扑: E 上的弱拓扑—它在 E 是有穷维的情形是和强拓扑一致的, 在弱拓扑下一个序列 (u_n) 收敛于 u 当且仅当 $L(u_n) \rightarrow L(u) \forall L \in E^*$ 单射下单位球是紧的当且仅当 E 是自

(1) 该段中文有一些错误, 已做了改正, 不在此单射列举—译者注. 反的, 换言之, E 能典则地等同于它的二次对偶 $(E^*)^*$. 对于 $1 < p < \infty$, 给定一个开子集 $\Omega \subset \mathbb{R}^n$, $L^p(\Omega)$ 的拓扑对偶借助于 Riesz 表示能等同于 $L^{p^*}(\Omega)$, 这里 p^* 共轭于 p , 即 $1/p + 1/p^* = 1$ 并且 $L^p(\Omega)$ 是自反的, 然而 $W^{s,p}(\Omega)$ 和 $W_0^{s,p}(\Omega)$ 的拓扑对偶两者都是 $W_0^{-s,p^*}(\Omega)$, 所以只有 $W_0^{s,p}(\Omega)$ 是自反的. $L^1(\Omega)$ 和它的拓扑对偶 $L^\infty(\Omega)$ 都不是自反的, 因为 $L^1(\Omega)$ 严格地包含在 $L^\infty(\Omega)$ 的拓扑对偶中, 有 $L^\infty(\Omega)$ 上的连续线性型 L 它不具有如下的形式

$$L(u) = \int_{\Omega} uv, \forall u \in L^\infty(\Omega), \text{ 这里 } v \in L^1(\Omega)$$

类似地, 一个赋范线性空间 E 的拓扑对偶 E^* 可以赋予由对偶范数 $\|\cdot\|_{E'}$ 导出的拓扑和弱*拓扑, 后者是使得映射 $L \mapsto L(u), u \in E$ 是连续的最弱的拓扑, 在这个拓扑下 E^* 中的单位球是紧的 (Banach-Alaoglu 定理).

对偶性并不总能保留可分性单射一个拓扑矢量空间是可分的如果它有一个可数稠子空间— $L^{\infty(x)}(\Omega)$ 空间不是可分的, 它是 $L^1(\Omega)$ 的拓扑对偶, 而 $L^1(\Omega)$ 是可分的. 然而 Hahn-Banach 定理的结论是: 如果一个 Banach 空间的拓扑对偶是可分的, 那么原空间也是, 如果再加上自反的假设条件那么就一致了; 一个 Banach 空间是自反

和可分的只要它的拓扑对偶是就行了. 对于 $1 \leq p < \infty$, $L^p(\Omega)$ 和 $W_0^{s,p}(\Omega)$ 是可分的, 如果 $p \neq 1$ 还是自反的.

B.0.4 Fourier 变换

在 18 世纪中叶, Bernouilli 把一根弦的振动解释成 n 个质点的振动当 n 趋于无穷时的极限情形, Bernouilli 引入了叠加原理的新概念, 认为弦的振动可以分解成“固有振动”的叠加. 这一观点引发了一场关于能否把一个任意函数展开成三角级数的讨论. 另一个关于展开成“正交函数”(这一术语的使用始于 Hilbert) 的例子也在那时出现在关于振动问题和热理论的讨论中, 但只是在 19 世纪, 有了 Fourier 和 Dirichlet 的工作, 叠加问题才得以解决.

可分的 Hilbert 空间提供了一组可数的正交系 $\{e_n\}_{n \in \mathbb{Z}}$ (对 H 上的标量积 $\langle \cdot, \cdot \rangle_n$ 有 $\langle e_n, e_m \rangle_{II} = \delta_{nn}$), 它是完备的, 就是每个矢量 $u \in H$ 可以在这个系中惟一地展开成 $u = \sum_{n \in \mathbb{Z}} \hat{u}_n e_n$, 它的 Fourier 系数是 $\hat{u}_n = \langle u, e_n \rangle$. 系数满足 Parseval 关系 $\sum_{n \in \mathbb{Z}} |\hat{u}_n|^2 = \|u\|^2$ (这里 $\|\cdot\|$ 是和 $\langle \cdot, \cdot \rangle$ 相关的范数), 并且 Fourier 变换 $u \mapsto (\hat{u}(n))_{n \in \mathbb{Z}}$ 建立了可分 Hilbert 空间 H 和复数平方可和序列空间 $l^2(\mathbb{Z})$ 之间的一个等距同构. 特别地, 在单位圆 $S^1 = \mathbb{R}/\mathbb{Z}$ 上具有通常 Haar 测度 dt 的 L^2 函数空间 $L^2(S^1)$ 是可分的, 具有完备的正交系 $t \rightarrow e_n(t) = e^{2i\pi nt}$, $n \in \mathbb{Z}$, Fourier 变换

$$u \mapsto \left(\hat{u}(n) = \int_0^1 e^{-2i\pi nt} u(t) dt \right)_{n \in \mathbb{Z}}$$

将该空间与空间 $l^2(\mathbb{Z})$ 等同了起来. 在这一等同下投影 $p_+ : u \mapsto (u_n)_{n \in \mathbb{N}}$ 在 $l^2(\mathbb{Z})$ 中的值域, 所得到的 Hilbert 空间 $l^2(\mathbb{N})$ 就相应于 Hardy 空间 $\mathcal{H}^2(S^1)$.

Fourier 变换扩展到了空间 $\mathcal{S}(\mathbb{R}^n)$, 取 $f \in \mathcal{S}(\mathbb{R}^n)$, 映射

$$f \mapsto \hat{f}(\xi) = \frac{1}{\sqrt{(2\pi)^n}} \int_{\mathbb{R}^n} e^{-i\xi \cdot x} f(x) dx$$

将 $\mathcal{S}(\mathbb{R}^n)$ 线性连续地映到它自身之上, 它有连续逆 $f \mapsto \hat{f}(-\xi)$ $n = 1$ 时, 泊松公式建立了 $f \in \mathcal{S}(\mathbb{R})$ 和它的 Fourier 变换 \hat{f} 之间的关系 $\sum_{n=-\infty}^{\infty} f(2\pi n) = \sum_{n=-\infty}^{\infty} \hat{f}(n)$.

Fourier 变换将对一个多重指标 $\alpha = (a_1, \dots, a_n)$ 的微分运算 D_ξ^α 转换 (直到差一个乘常数因子) 成对 $\xi^2 = \xi_1^2 \cdots \xi_n^2$ 的乘法, 所以可以用来定义 Sobolev 空间 $W^{s,p}$, 其中 s 是一个实数, 这是一个具有有穷 Sobolev 范数 $\|u\|_{W^{s,p}} = (\int |l + |\xi||^2(\xi)|^p)^{1/p}$ 的 L^p 函数空间 (它和先前定义的当 $s = k$ 是一个非负整数的情形是一致的).

Fourier 变换还用来描述一个线性伪微分算子 A (见下面两个章节要讨论的有界和无界线性算子的概念), A 是 a 阶的是依据它的符号 σ_A 作用在 \mathbb{R}^n 的开子集上的

光滑函数上 $-\sigma$ 是 $U \times \mathbb{R}^n$ 的一个光滑映射在 x 中有紧支集, 对任意的多重指标 $\alpha, \beta \in \mathbb{N}_a^n$ 有常数 $C_{\alpha, \beta}$ 满足

$$|D_x^\alpha D_\varepsilon^\beta \sigma(x, \xi)| \leq C_{\alpha, \beta} (1 + |\xi|)^{\sigma - |\beta|}$$

对一切 $\xi \in \mathbb{R}^n$ 成立单射且有

Fourier 变换将一个 \mathbb{R}^n 上的高斯函数 $x \mapsto e^{-(1/2)\lambda|x|^2}$ (λ 是一个非零标量) 映成另

的一个出发点. 更为一般的, 一个 Hilbert 空间 H 上的具有协方差 C 的高斯概率测度 μ 的特征函数

$$\hat{\mu}(\xi) := \int_H e^{i(x, \xi)_H} \mu(dx)$$

是函数 $\xi \mapsto e^{-(1/2)(\xi, \xi)_H}$. 这样的概率测度在 Euclidean 量子场理论中是典型的; 在公理化量子场理论中, n 点函数的解析性质可以用 Fourier 变换从 Wightman 公理中推导出来. 这样 Fourier 变换成了量子场理论诸多不同方面的基础.

B.0.5 Fredholm 算子

一个 $[0, 1] \times [0, 1]$ 上的复值连续函数 K 产生了一个 $[0, 1]$ 上复值连续函数 (赋以极大值范数 $\|\cdot\|_\infty$) 的积分算子

$$A : f \rightarrow \int_0^1 K(x, y) f(y) dy$$

它有如下的上界性质

$$\|Af\|_\infty \leq \text{Sup}_{[0,1] \times [0,1]} |K(x, y)| \|f\|_\infty$$

换言之, A 是一个有界线性算子, 范数以 $\text{Sup}_{[0,1] \times [0,1]} |K(x, y)|$ 为上界; 一个从赋范线性空间 $(E, \|\cdot\|_E)$ 到另一个赋范线性空间 $(F, \|\cdot\|_F)$ 的线性算子 $A : E \rightarrow F$ 是有界的 (或连续的) 当且仅当 (算子) 范数 $\|A\| := \text{Sup}_{\|u\|_E \leq 1} \|Au\|_F$ 是有界的.

一个用连续核 K 定义的积分算子

$$A : f \rightarrow \int_0^1 K(x, y) f(y) dy$$

还是紧的; 一个紧算子是一个赋范空间的有界算子并且将有界集合映射成一个预紧集合, 即集合的闭包是紧的. 赋范空间上紧算子的另一些例子是有穷秩算子, 这些算子有有穷维的值域. 事实上, 一个可分 Hilbert 空间上的任意的紧算子都可以用一个有穷秩算子的序列以算子范数 $\|\cdot\|$ 导出的 \boxtimes 扑来逼近.

在上面提到的积分算子的情形, Volterra 提出方程 $f = (1 - A)\phi, f \in C([0, 1])$ 的连续解是 $\phi = (1 - A)^{-1}f$, 受到这个工作的启发, Fredholm 在 1900 年的“论一类泛函方程”(Sur une class d'équation fonctionnelles) 中引入了一个复参数 λ , 讨论方程 $f = (1 - \lambda A)\phi$. 他证明了从那时起称之为 Fredholm 的备择, 它表明或者方程 $f = (1 - \lambda A)\phi$ 对每一个 $f \in C([0, 1])$ 有惟一解, 或者齐次方程 $(I - \lambda A)\phi = 0$ 有非零解. 用现在的话来讲, 这意味着一个紧线性算子 A 的预解式 $R(A, \mu) = (A - \mu I)^{-1}$ 是满射的充要条件是它是单射的. Fredholm 的备择定理是解偏微分方程的强有力工具, 在这些方程中有 Dirichlet 问题, 它的解是在某个区域 $\Omega \subset \mathbb{R}^n$ 上的调和函数 u (即 $\Delta u = 0$, 这里 $\Delta = -\sum_{i=1}^n \partial^2/\partial x_i^2$), u 还要满足 Dirichlet 边界条件 $u|_{\partial\Omega} = f$, 这里 f 是边界 $\partial\Omega$ 上的连续函数. Dirichlet 问题有几何上的应用, 特别是对于非线性 Plateau 问题, 它是在给定边界曲线的条件下求在 \mathbb{R}^{dd} 中的最小表面积, 它化成了一个(线性)Dirichlet 问题.

从紧算子 A 构造出算子 $B = I - A$, 它是一个特殊的 Fredholm 算子, Fredholm 算子是有界线性算子 $B: E \rightarrow F$, 它“差一个紧算子”是可逆的, 就是存在一个有界线性算子 $C: F \rightarrow E$, 使得 $BC - I_F$ 和 $CB - I_E$ 两者紧. 一个 Fredholm 算子 B 有一个穷维的核 $\ker B$, 当 $(E, \langle \bullet, \cdot \rangle_E)$ 和 $(F, \langle \bullet, \cdot \rangle_F)$ 是 Hilbert 空间时算子 B 的余核 $\text{Ker } B^*$ 也是有穷维的, 这里 B^* 是 B 的共轭它定义成

$$\langle Bu, v \rangle_F = \langle u, B^*v \rangle_E \quad \forall u \in E, \forall v \in F,$$

所以就有了算子指标的定义 $\text{ind}(B) = \dim(\text{Ker } B) - \dim(\text{Ker } B^*)$, 这是指标理论的一个出发点. Töplitz 算子 T_ϕ , 这里 ϕ 是单位圆上的一个连续函数, 提供了 Fredholm 算子的第一个例子; 比作用在 Hardy 空间 $\mathcal{H}^2(S^1)$ 上:

$$Te_{-n} \left(\sum_{m \geq 0} a_m e_m \right) = \sum_{m \geq 0} a_{m+n} e_m$$

这里有等同关系 $\mathcal{H}^2(S^1) \simeq l^2(\mathbb{N}) \subset l^2(\mathbb{Z}), l^2(\mathbb{Z})$ 赋予正则完备的正交基 $(e_n, n \in \mathbb{Z})$. Fredholm 指标 $\text{ind}(Te_{-n})$ 是正整数 n , 所以它的共轭的指标是 $-n$, 由此得出结论 Fredholm 算子的指标映射是到整数映上的.

B.0.6 单参数(半)群

与有穷维的情形不同, 两个赋范线性空间 $(E, \|\cdot\|_E)$ 和 $(F, \|\cdot\|_F)$ 之间的一个线性算子 $A: E \rightarrow F$ 不能期望它是有界的. 在偏微分方程中出现的无界算子包括象在开子集 $\Omega \subset \mathbb{R}^n$ 上的拉普拉斯算子 Δ . 以下举出偏微分方程一些基本的例子, 它们出现在泊松、傅里叶和柯西研究数学物理各种问题的辛劳工作中:

$$\Delta u = 0 \quad \text{拉普拉斯方程}$$

$$\frac{\partial^2 u}{\partial t^2} + \Delta u = 0 \quad \text{波动方程}$$

$$\frac{\partial u}{\partial t} + \Delta u = 0 \quad \text{热传导方程}$$

以及以后量子力学中的 Schrödinger 方程:

$$i \frac{\partial u}{\partial t} = \Delta u$$

这里 t 是时间参数.

一个无穷维赋范空间上的一个无界线性算子通常定义在一个区域 $D(A)$ 上, 它是严格包含在 E 中的. 拉普拉斯算子 Δ 就是定义在 $L^2(\mathbb{R}^n)$ 中的稠定义域 $D(A) = H^2(\mathbb{R}^n)$ 上; 它定义了一个从 $H^2(\mathbb{R}^n)$ 到 $L^2(\mathbb{R}^n)$ 的有界算子, 但不能延拓成为 $L^2(\mathbb{R}^n)$ 上的一个有界算子. 跟这个算子一样, 我们所遇到的绝大多数无界算子 $A: E \rightarrow F$ 在 E 中有稠密的定义域 $D(A)$ 并且是闭的, 也就是它们的图像 $\{(u, Au), u \in D(A)\}$ 作为赋范线性空间 $E \times F$ 中的子集是闭的. 在算子不是真的闭的情形, 我们可以将它们闭化, 就是说, 进行闭延拓称之为算子的闭包. 由闭图像定理, 当 E 和 F 是 Banach 空间时, 一个线性算子 $A: E \rightarrow F$ 是连续的只要它的图像是闭的, 由此推出, 定义在稠密定义域上的闭线性算子, 如果它的定义域是全空间, 是有界的.

对于一个有稠密定义域 $D(A)$ 的闭算子 $A: E \rightarrow F$, 这里 E 和 F 是 Hilbert 空间, 有内积 $\langle \bullet, \bullet \rangle_E$ 和 $\langle \cdot, \cdot \rangle_F$, A 的共轭 A^* 在它的定义域 $D(A^*)$ 上定义成

$$\langle Au, v \rangle_F = \langle u, A^*v \rangle_E \quad \forall (u, v) \in D(A) \times D(A^*)$$

一个定义域是 $D(A)$ 的自共轭算子 A 是这样的: $D(A) = D(A^*)$ 和 $A = A^*$; \mathbb{R}^n 上的拉普拉斯算子 Δ 是在 Sobolev 空间 $H^2(\mathbb{R}^n)$ 上自共轭的, 但从根本上讲它只是在稠密定义域 $D(\mathbb{R}^n)$ 上才是自共轭的, 这意味着它的闭包是自共轭的.

无界自共轭算子能成为单参数有界算子半群的生成元. 一个 Hilbert 空间 H 上的单参数有界算子族 $T_t, t \geq 0 (T_t, t \in \mathbb{R})$ 是一个半群 (群) 如果满足 $T_s T_t = T_{t+s} \forall t, s \geq 0 (\forall t, s \in \mathbb{R})$ 和它是强连续的 (或简称为连续): 对任意 $t_0 \geq 0 (t_0 \in \mathbb{R})$ 和 $u \in H$ 有 $\lim_{t \rightarrow t_0} T_t u = T_{t_0} u$.

Stone 定理建立起了 Hilbert 空间上连续单参数酉 ($U_t^* U_t = U_t U_t^* = I$) 群 $U_t, t \in \mathbb{R}$ 且 $U_0 = \text{Id}$, 和自共轭算子 A 之间的单射对应, 从 $U_t, t \in \mathbb{R}$ 得到群的无穷小生成元: $iA^{(1)}$

$$iAu = \lim_{t \rightarrow 0} \frac{U_t u - u}{t}, \quad u \in D(A)$$

所以可以紧凑地写成 $U_t = e^{itA}$. 量子力学中的一个重要的例子是 $U_t = e^{itH} U_0, t \in \mathbb{R}$, H 是自共轭的 Hamilton 量, 它是 Schrödinger 方程 $du/dt = iHu$ 的解. Lie-Trotter

公式, 它对于 Feynman 的路径积分有重要的应用, 表示出了由 $A+B$ 生成的西群, 其中 A, B 和 $A+B$ 都是在它们各自定义域上的自共轭算子, 西群表示成了一个强极限

$$e^{it(A+B)} = \lim_{n \rightarrow \infty} \left(e^{\frac{iA}{n}} e^{\frac{iB}{n}} \right)^n$$

另一方面, Hilbert 空间 $(H, \langle \bullet, \cdot \rangle_H)$ 上的正定算子——就是 A 是自共轭的且 $\langle Au, u \rangle_H \geq 0, \forall u \in D(A)$ ——生成了单参数半群 $T_t = e^{-zA}, t \geq 0$. Hille 和 Yosida 证明了在一个 Hilbert 空间上, 强连续压缩 (即 $\|T_t\| \leq 1, \forall t \geq 0$) 半群且 $T_0 = \text{Id}$ 是单射对应于稠定的正定算子 $A: D(A) \subset H \rightarrow H$, 它是极大的 (即 $I+A$ 是映上的), 无穷小生成元可以从相应的半群得出

(1) 原文此处有误一译者注.

$$-Au = \lim_{t \rightarrow 0} \frac{T_t u - u}{t}, u \in D(A)$$

类似地, 一个 Hilbert 空间上的正稠定的自共轭算子 Λ 产生了一个稠定闭对称共轭双线性型 $(u, v) \mapsto (\sqrt{\Lambda}u, \sqrt{\Lambda}v)$ (见下一节 $\sqrt{\Lambda}$ 的定义; $\langle \bullet, \cdot \rangle_H$ 是 H 上的内积), 并且这一映射产生了算子和具有前述性质的共轭双线性型之间的单射对联应, 这是 Dirichlet 型理论的出发点之一. 对于一个可分的 Banach 空间 E 上的概率测度 μ , 我们可以将它与一个 Hilbert 空间 $H, E^* \subset H^* = H \subset E$, 上的一个稠定闭对称共轭双线性型 (事实上是一个 Dirichlet 型) 相关联, 这是 Wiener 空间 $E = C([0, t])$ 上标准的 Wiener 测度 μ 和由 Cameron-Martin 空间 $H = H^1([0, t])$ 给定的 Hilbert 空间的特殊情形, 双线性型是

$$(u, v) \mapsto \int \langle \nabla u, \nabla v \rangle_H$$

这里 $\bar{\nabla}$ 是闭 Malliavin 梯度.

Δ 是 \mathbb{R}^n 上的拉普拉斯算子, $-\Delta$ 生成了热算子半群 $e^{-\Delta}, t \geq 0$. 它有一个光滑的核 $K_t \in C^\infty(\mathbb{R}^n \times \mathbb{R}^n)$, 定义是

$$(e^{-\Delta x} f)(x) = \int_{\mathbb{R}^n} K_t(x, y) f(y) dy, \quad \forall f \in C_o^\infty(\mathbb{R}^n)$$

这是一个光滑算子, 将 Sobolev 函数映成光滑函数的算子. 一般地, 一个 \mathbb{R}^n 的开子集 U 上具有符号 σ_A 的一个伪微分算子 A 只能有一个分布的核

$$K_A(x, y) = \int_{\mathbb{R}^n} e^{i\langle x-y, \xi \rangle} \sigma(\xi) d\xi$$

\mathbb{R}^n 上的拉普拉斯算子的逆 $(\Delta + m^2)^{-1}$ 的核 (非负实数 m^2 表示质量) 称之为 \mathbb{R}^n 上的 Green 函数, 它在 Feynman 图像理论中起了根本性的作用.

B.0.7 谱理论

谱理论研究的是复参数 λ 的值分布, 这些 λ 使得对给定的赋范空间 E 上的一个线性算子 A , 算子 $A - \lambda I$ 有逆, 以及如果逆存在, 这个逆就是算子 A 的预解式 $R(A, \lambda) = (A - \lambda I)^{-1}$ 的性质. A 的谱集 $\text{Sp}(A)$ 是 A 的预解集在 \mathbb{C} 中的补; 它是 3 个不相交集的并; 使得 $A - \lambda I$ 不是单射的所有复数 λ 的集合称之为点谱这样的 λ 是 A 的本征值, 有相应的本征函数 $u \in D(A)$, $Au = \lambda u$; 使得 $A - \lambda I$ 有一个稠定的无界逆 $R(A, \lambda)$ 的 λ 的集合, 称之为连续谱; 以及使 $A - \lambda I$ 有一个无界但不稠定的逆 $R(A, \lambda)$ 的 λ 组成剩余谱.

一个有界算子的谱集是有界的; 作用在一个 Hilbert 空间上的自共轭算子 A 有实谱集且没有剩余谱, 因为 $A - \lambda I$ 的值域是稠的. 作为 Fredholm 备择定理的一个结果, 紧算子的谱集中只有点谱; 它是可数的以 0 为聚点. 量子力学系统的一个 Hamilton 量的点谱和连续谱两者都有, 其中的点谱特别有意义因为与之对应的本征函数是系统的稳态. 正如 Kac 首次所指出的 (你能听得出鼓的形状吗?), 作用在函数上的一个算子的谱能够反映出这些函数所在的空间的几何, 这是在微分几何中许多有趣并影响深远的问题的出发点.

一个 Hilbert 空间上的一个自共轭线性算子可以用一簇投影 $E_\lambda (\lambda \in \mathbb{R})$ 加以刻画, 这就是谱表示

$$A = \int_{\text{Sp}(A)} \lambda dE_\lambda$$

给出 \mathbb{R} 上的一个 Borel 实值函数 f , 算子

$$f(A) = \int_{\text{Sp}(A)} f(\lambda) dE_\lambda$$

是另一个自共轭算子. 某个 Hilbert 空间 $(H, \langle \bullet, \bullet \rangle_H)$ 上一个具有稠定义域 $D(A)$ 的正算子 A 有非负的谱且对任意正实数 t , 映射 $\lambda \mapsto e^{-t\lambda}$ 给出了相关联的有界热算子

$$e^{-tA} = \int_{\text{Sp}(A)} e^{-t\lambda} dE_\lambda$$

而映射 $\lambda \mapsto \sqrt{\lambda}$ 则产生了一个正算子 \sqrt{A} 满足 $\sqrt{A}^2 = A$. 预解式还可以用来从一个线性算子定义新的算子

$$f(A) = \frac{1}{2i\pi} \int_C f(\lambda) R(A, \lambda) d\lambda$$

这就是沿着围绕谱的一个围道 C 的柯西型积分; 用这种方法可以用来定义 (本质自共轭的) 正定的椭圆伪微分算子 A 的复数幂次 A^{-z} , 并定义 A 的 ζ 函数 $z \mapsto$

$\zeta(A, z)$. ζ 函数是将普通的行列式扩展到了自共轭椭圆算子的 ζ 行列式的有用工具, 从而对量子场理论的路径积分方法中的分割函数的意义提供了说明.

B.0.8 算子代数

一个 Hilbert 空间 H 上的有界线性算子组成了一个代数 $\mathcal{L}(H)$, 它对于算子范数是闭的, 并有由共轭运算 $A \mapsto A^*$ 给出的对合; 它还是 C^* 代数, 即它是 \mathbb{C} 上的代数 A , 有范数和对合 $*$, 使得对这个范数 A 是闭的以及 $\|ab\| \leq \|a\|\|b\|$ 和 $\|a^*a\| = \|a\|^2$ 对所有 $a, b \in A$ 成立, 由 Gelfand-Naimark 定理, 每个 C^* 代数同构于某个 $\mathcal{L}(H)$ 的子 C^* 代数. 谱的概念从有界算子扩充到了 C^* 代数; 一个 C^* 代数 A 中的一个元 a 的谱 $\text{Sp}(a)$ 是复数 λ 的 (紧) 集合, 这些 λ 使得 $a - \lambda \cdot 1$ 不可逆. 自共轭的概念也扩展了 ($a = a^*$), 并且如同自共轭算子 $B \in \mathcal{L}(H)$ 是非负的 (在这个情形它的谱在 \mathbb{R}^+ 中) 当且仅当 $B = A^*A$ 对某个有界算子 A 成立一样, 一个元 $b \in A$ 是非负的当且仅当对某个 $a \in A$ 有 $b = a^*a$, 在这情形 $\text{Sp}(a) \subset \mathbb{R}_0^+$.

X 是某个局部紧的 Hausdorff 空间, 连续函数 $f: X \rightarrow \mathbb{C}$ 在无穷远处为零, 这些 f 的代数 $C(X)$ 赋予极大范数以及复共轭 $f \mapsto \bar{f}$ 后也是一个 C^* 代数, 并且是一个典型的交换 C^* 代数, Gelfand 证明了如果 C^* 代数有单位元, 则它等距同构于 $C(X)$, X 紧. 对于一个 C^* 代数 A , 我们可以与之关联到一个可交换群 $K_0(A)$, 它是一个紧 Hausdorff 空间 X 上矢量丛的同构类的 Grothendieck 群 $K^0(Z)$ 的对偶.

一个 Hilbert 空间上的紧算子构成了 C^* 代数 $\mathcal{L}(H)$ 的惟一真双边理想 $\mathcal{K}(H)$, 它对于 $\mathcal{L}(H)$ 上的算子范数拓扑是闭的. Calkin 对一个可分 Hilbert 空间 H 的 $\mathcal{L}(H)$ 中的所有双边理想进行了分类, 之后, 商空间 $\mathcal{L}(H)/\mathcal{K}(H)$ 就称之为 Calkin 空间了; 我们可以在这样的理想和一定的序列空间之间建立起单射对应. 对应于复值序列 (u_n) 满足 $\sum_{n \in \mathbb{N}} |u_n| < \infty$ 的 Banach 空间 $l^1(\mathbb{Z})$ 的是迹类算子的 $*$ 理想 $\mathcal{I}_1(H)$. 一个非负算子 $A \in \mathcal{L}(H)$ 的迹 $\text{tr } A = \sum_{n \in \mathbb{Z}} \langle A e_n, e_n \rangle$ ($e_n >_H$) 在 $[0, +\infty]$ 中, 它不依赖于赋予 H 中内积 $\langle \cdot, \cdot \rangle_H$ 的完备正交基 $\{e_n, n \in \mathbb{Z}\}$ 的选取. $\mathcal{I}_1(H)$ 是满足 $\|A\|_1 = \text{tr}(|A|)$ 为有界的 H 上的有界线性算子的 Banach 空间. 给定一个 (本质自共轭的) d 阶正微分算子 D , 它作用在闭 n 维 Riemannian 流形 M 上的光滑函数上, 它的复数幂次 D^{-z} 当 $\text{Re}(z) > n/d$ 时是 M 上 L^2 函数空间上的一个迹类, 并且相应的迹 $\text{tr}(D^z)$ 扩展到整个平面上的一个半纯函数, ζ 函数 $\zeta(D, z)$ 在 0 点是全纯的.

更为一般地, 满足 $\sum_{n \in \mathbb{Z}} |u_n|^p < \infty$ 的复数值序列 $(u_n)_{n \in \mathbb{Z}}$ 的 Banach 空间 $l^p(\mathbb{Z})$, ($1 \leq p < \infty$) 关联到 Schatten 理想 $\mathcal{I}_p(H)$ ($1 \leq p < \infty$), 这里的 $\mathcal{I}_p(H)$ 是满足 $\|A\|_p = (\text{tr } |A|^p)^{1/p}$ 为有界的 H 上有界线性算子的 Banach 空间. 正如所有的 l^p 序都收敛于 0, Schatten 理想 $\mathcal{I}_p(H)$ 都在 $\mathcal{K}(H)$ 中, 我们有 $\cdots (\mathcal{I}_{p+1}(H)) \subset \mathcal{I}_p(H) \subset \cdots \subset \mathcal{K}(H)$.

紧算子和 Schatten 理想在把指标理论扩展到非交换的情形时是有用的; 一个对

合代数 A 上的一个 Fredholm 模 (H, F) 是由一个 Hilbert 空间 H 中的 A 的一个对合表示 π 和 H 上一个满足 $F^2 = \text{Id}_H$ 的自共轭有界线性算子 F 所给出, 而算子括号 $[F, \pi(a)]$ 对所有的 $a \in A$ 是紧的. 对一个 p 可和的 Fredholm 模 (H, F) , 即对所有的 $\alpha \in A, [F, \pi(a)] \in \mathcal{I}_p(H)$, 我们可以相关联到由 A 上一个循环上闭链所给出的陈-特征 $\text{ch}^*(H, F)$ 的一个表示 τ , 它与 K 理论相配合以在 K 理论上建立一个整数值指标映射 τ .

Schatten 理想在讨论无穷维空间的几何时也是有用的, 诸如环路群, 对于它 Hilbert-Schmidt 算子 ($\mathcal{I}_2(H)$ 中的算子也称为 Hilbert-Schmidt 算子) 是特别有用的. 一个 Holder 型的不等式表明两个 Hilbert-Schmidt 算子的乘积是迹类. 又进一步, 对任意两个 Hilbert-Schmidt 算子 A 和 B , “循环性质” $\text{tr}(AB) = \text{tr}(BA)$ 成立, 以及共轭双线性型 $(A, B) \mapsto \text{tr}(AB^*)$ 使 $\mathcal{L}_2(H)$ 成为一个 Hilbert 空间.

参考文献

- [1] I. Newton, *Principia Mathematica Philosophia Naturalis*, 1687, English translation *The Principia: Mathematical Principles of Natural Philosophy* (City? University of California Press, 1999). (document), 1
- [2] Einstein A et al. (1958) *The Principle of Relativity*. New York: Dover. 1.1.2
- [3] Will, Clifford M. 2018. *Theory and Experiment in Gravitational Physics*. Cambridge: Cambridge University Press. 1.4.4
- [4] Ohanian H C, Ruffini R. 2013. *Gravitation and Spacetime*(Third Edition). Cambridge: Cambridge University Press. 1.4.4
- [5] Kretschmann, Erich. 1917. *Über den physikalischen Sinn der Relativitätspostulate. A. Einsteins neue und seine ursprüngliche Relativitätstheorie*. Annalen der Physik 53: 575–614. 1.1.6
- [6] Weinberg S. 1971. *Gravitation and Cosmology*,(Principles and Applications of the General Theory of Relativity) WILEY. 1.1.6
- [7] R.Descartes, *Principia Philosophiae*(1644), Translated by V.R.Miller and R.P. Miller (Dordrecht: Reidel, 1983).
- [8] I. Newton, *De Gravitatione et Aequipondio Fluidorum*, translation in *Unpublished Papers of Isaac Newton*, ed. A.R. Hall and M.B. Hall (Cambridge: Cambridge University Press, 1962).
- [9] Synge J L (1972) *Relativity: The Special Theory*. Amsterdam: North-Holland.
- [10] Spivak M (1965) *Calculus on Manifolds*. New York: W A Benjamin.
- [11] Choquet-Bruhat Y, De Witt-Morette C, and Dillard-Bleick M (1977) *Analysis, Manifolds and Physics*. Amsterdam: North-Holland.

- [12] Parrott S (1987) *Relativistic Electrodynamics and Differential Geometry*. Berlin: Springer.
- [13] Einstein A. *Zur Elektrodynamik bewegter Körper*(*On the Electrodynamics of Moving Bodies*)[J]. *Annalen der Physik*, 1905, 17(891): 50. 论文的翻译可在此书找到: *The Principle of Relativity* by Lorentz HA, Einstein A, Minkowski H, and Weyl H, 结合笔记: Sommerfeld A. New York: Dover, 1952.
- [14] Einstein A. *Does the inertia of a body depend upon its energy-content*[J]. *Annalen der Physik*, 1905, 18(13): 639-41.
- [15] H. Minkowski. *Raum und Zeit, Jahresberichte der Deutschen Mathematiker-Vereinigung*, 1909: 75-88.
- [16] Herrmann F., *New Physics Course For Senior High School*. (德国 KPK 物理课程教材)
- [17] Chalmers AF(1975), *Maxwell and the displacement current*. *Physics Education* January 1975: 45-49.
- [18] Siegel DM(1985), *Mechanical image and reality in Maxwell's electromagnetic theory*. In: Harman PM(ed.) *Wranglers and Physicists*. Manchester: Manchester University Press.
- [19] Roche J(1998), *The present status of Maxwell's displacement current*. *European Journal of Physics* 19:155-166.
- [20] Jammer M(1989), *The Conceptual Development of Quantum Mechanics*, 2nd edn. Tomash Publishers, American Institute of Physics.
- [21] Jan Czerniawski(2006), *The possibility of a simple derivation of the Schwarzschild metric*, arxiv:gr-qc/0611104. 1