

What is the nature of black holes?

黑洞的本质

蔡荣根^{①*}, 曹利明^{②*}^① 中国科学院理论物理研究所, 北京 100190;^② 中国科学技术大学物理学院, 合肥 230026

* 联系人, E-mail: cairg@itp.ac.cn; caolm@ustc.edu.cn

2016-02-17 收稿, 2016-02-19 修回, 2016-02-19 接受, 2016-04-18 网络版发表

摘要 什么是黑洞? 这个问题颇难回答. 对于普通大众来说, 黑洞是一种在科幻题材中经常出现的奇怪天体. 而对于相关的人员来说, 黑洞是时空中连光都逃脱不出的区域, 在宇宙中实实在在的存在. 虽然认识的水平有所不同, 但黑洞这个物理对象对他们来说都是神秘的, 神秘到至今为止我们也不能够真正地理解它. 如果人们再追问: 黑洞的本质是什么? 不同的人也许有不同的理解. 甚至有时候从事黑洞物理研究的学者们所给出的答案也会大相径庭. 本文不试图给出博人眼球的解答, 而是倾向于告诉读者学术界关于这一问题的主流理解方式.

关键词 黑洞, 引力, 量子引力

1 经典黑洞的本质

1.1 什么是黑洞?

粗略地说, “黑洞是时空中连光都逃逸不出的区域”. 这是一个朴素但又非常不平凡的关于黑洞的描述方式. 真正地理解这一描述是一件不容易的事, 原因在于人们对于时空概念理解的不同, 或者对连光都逃逸不出这一过程界定的不同. 这里我们愿意从历史发展的眼光来看待这个问题.

在介绍黑洞这个概念时, 很多人愿意提及如何在牛顿力学的框架下理解一个黑洞. 这种想法可以追溯到18世纪的英国牧师兼自然哲学家米歇尔(Mitchell). 1783年, 米歇尔在写给卡文迪许(Cavendish)的一封信中提出了暗星的概念(图1). 这封信中的内容于1784年在英国皇家学会发表^[1]. 同时代的法国著名学者拉普拉斯(Laplace)于1796年也独立地提出暗星的想法, 且将这个想法写到了其著作*Exposition du Système du Monde*的第一和第二版中, 并于1798年给出了一个光逃逸不出的证明. 拉普拉斯的工作被霍

金(Hawking)和埃利斯(Ellis)翻译成英文, 并放在他们1973所著的*The Large Scale Structure of Space-time*一书的附录中, 因此广为人知. 1979年, 剑桥大学的引力物理学家杰彭斯(Gibbons)在*New Scientist*杂志中的一文指出了米歇尔的工作. 自那时起, 米歇尔的贡献才被人们广泛知悉.

拉普拉斯关于暗星的讨论基于牛顿引力理论和光的粒子学说: 如果星体表面光子的动能小于它的引力势能, 光子便不能够逃逸到无限远处. 由此可以很容易得到质量为 M 的星体成为暗星时的最大半径为 $R=2GM/c^2$, 其中 c 是光速, G 是牛顿常数. 这就是所谓的暗星, 也是迄今为止人们能够发现的人类关于黑洞最早的一个认识. 需要指出的是: 这个半径恰好是爱因斯坦广义相对论中所预言的施瓦西黑洞的施瓦西半径. 但这只是一个巧合. 事实上, 在同时代的学者看来, 拉普拉斯等人的讨论存在着明显的漏洞, 即需假定光速不依赖于参考系. 但这和牛顿力学中任何物体的速度(包括光速)是一个相对的量相冲突. 在牛顿力学框架下, 总有一些物理过程(例如星体表

引用格式: 蔡荣根, 曹利明. 黑洞的本质. 科学通报, 2016, 61: 2083–2092

Cai R G, Cao L M. The nature of black holes (in Chinese). Chin Sci Bull, 2016, 61: 2083–2092, doi: 10.1360/N972016-00186

VII. On the Means of discovering the Distance, Magnitude, &c. of the Fixed Stars, in consequence of the Diminution of the Velocity of their Light, in case such a Diminution should be found to take place in any of them, and such other Data should be procured from Observations, as would be further necessary for that Purpose. By the Rev. John Michell, B. D. F. R. S. In a Letter to Henry Cavendish, Esq. F. R. S. and A. S.

图1 米歇尔写给卡文迪许的信中关于暗星的部分^[1]

Figure 1 The part on the dark star in the letter to Cavendish by Michell^[1]

面附近速度很大的电子发射光)使得光子的速度超过 c , 并可以逃逸到无限远处天文学家的望远镜(图2). 在牛顿力学的框架下, 物理信号可以以无限大的速度运行, 因此牛顿理论所在的时空中不存在信息逃逸不出的区域, 即不存在真正黑洞的概念. 当然我们现在知道光速不依赖于参考系是狭义相对论的一个基本假设. 可见, 若希望理解黑洞, 相对论性的时空观是必要的.

在牛顿时代或更早, 人们关于时空的认识是接近日常生活的. 先知告诉我们: 在这些时空中时间和空间是分离的(这是一种典型非动力学的, 人为加入的“背景结构”), 每一个时刻都存在一个三维的空间. 时间和空间的分离意味着我们需要两套度量, 分别来衡量时间的间隔和空间的间隔. 牛顿引力理论就

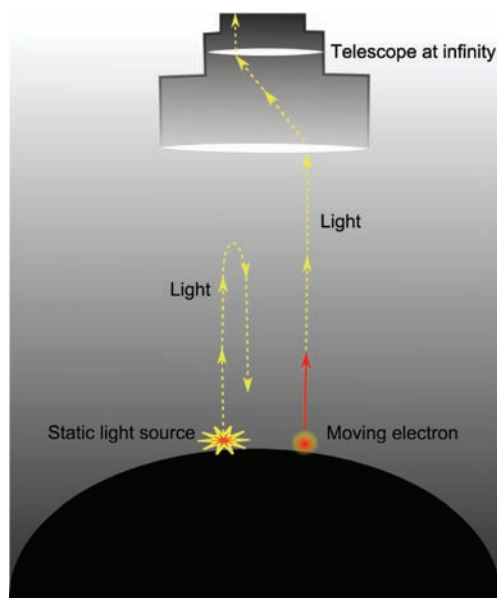


图2 (网络版彩色)米歇尔和拉普拉斯的暗星. 牛顿时空中不存在真正意义上黑洞的概念

Figure 2 (Color online) The dark stars proposed by Michell and Laplace. There is no real black hole in the Newtonian spacetimes

29. If there should really exist in nature any bodies, whose density is not less than that of the sun, and whose diameters are more than 500 times the diameter of the sun, since their light could not arrive at us; or if there should exist any other bodies of a somewhat smaller size, which are not naturally luminous; of the existence of bodies under either of these circumstances, we could have no information from sight; yet, if

是建立在这样的时空之上, 相应的引力场方程是一种典型的椭圆方程, 即泊松方程. 因此牛顿引力理论中没有引力波的概念. 而引力相互作用是一种超距作用, 物理信号的传播速度可以是无限大. 虽然牛顿引力理论在物理上简单直观, 但其数学结构是相对复杂的. 除了需要引入两套退化的度规, 人们还需要额外的联络结构. 而且这种联络结构并不能由这两套度规确定. 在相对论性时空中时间和空间没有先验地分离, 而是融合在一起成为一个四维的对象. 这意味着相对论性的时空只需要一个衡量“时空间隔”的度量, 或者说只需要一个度规. 更进一步地, 很多情况下, 用来描述时空弯曲程度的联络也由度规唯一确定. 因此相对论性时空中没有“人为的背景”, 具有比牛顿引力理论更为简单的数学结构. 任何物理信号都不能超光速, 这一基本假设要求时空的每一点处都能够构造出一个光锥(图3). 换句话说, 这个度规是洛伦兹的. 这样, 一个相对论性时空可以看成是一个二元组 (M, g) , 即一个四维流形 M 配上一个洛伦兹度规 g . 或者说一个相对论性时空就是一个洛伦兹流形. 在物理上, 相对论性时空上的引力理论更为自然. 如广义相对论中的爱因斯坦场方程, 通常可以写成一个非线性(拟线性)偏微分方程组, 而且具有双曲方程的特征. 这意味着相对论性的引力理论具有传播自由度, 存在引力波的概念. 事实上, 最近位于美国路易斯安那州和华盛顿州的激光干涉引力波天文台(LIGO)已经直接观测到了引力波的存在. 这一引力波是由二个转动黑洞并合后产生的^[2]. 在相对论性引力理论中, 引力相互作用以有限的速度(如光速)传播, 而不是超距作用. 引力现象归结为时空的弯曲程度, 这表明在一个相对论性的引力理论中度规也是动力学的, 而不是简单地作为背景或舞台出现在物理理论中. 度规即是背景又是动力学变量这一特征是相对论性理论的一个核心. 可以说, 相对论

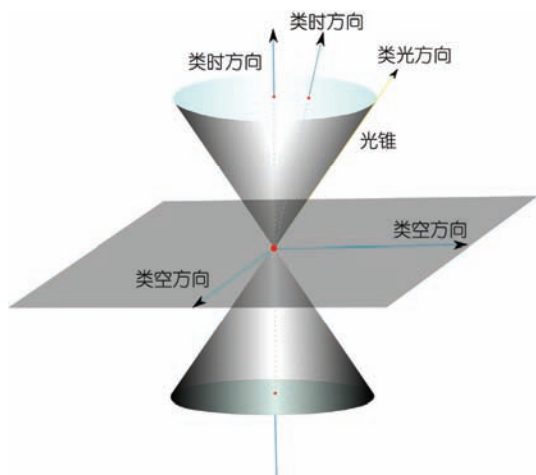


图 3 (网络版彩色)相对论性时空上每一点处的矢量可以分为三类: 类时、类光、和类空. 类时矢量可以看作过该点的质点世界线在该点的切矢量, 而类光矢量可以看成是过该点的光的世界线的切矢量, 即光的 4-波矢

Figure 3 (Color online) The vectors on a point of relativistic spacetime can be put into three classes, i.e., timelike, null, and spacelike. The timelike vector can be viewed as the tangent vector of the world line of a massive particle passed the point. While the null vector corresponds to the tangent vector of the world line of light, i.e., the 4-wave vector

性引力理论(如广义相对论)中的一系列重要的结论和困难都和这一事实密切相关。

比起牛顿或伽俐略时空, 相对论性时空除了拥有类时和类空无限远, 还拥有类光无限远的概念. 形象地说, 所谓的类光无限远可以理解成时空上光线能够延伸到的最远的“端点”的集合. 通常来说, 人们用来 $\mathcal{I}^+$ 代表未来类光无限远. 在闵氏时空上的任意一点发射的光都可以达到类光无限远. 但是不是有时空都有类似的性质呢? 答案并不是. 黑洞就是这样的时空, 在这个时空中的一些区域发出的信号无法到达类光无限远. 如果我们记时空为 (M, g) , 那么这样的区域可以记为

$$B = M - \mathcal{I}^-(\mathcal{I}^+).$$

这就是时空上的黑洞区. 其中 $\mathcal{I}^-(\mathcal{I}^+)$ 代表未来类光无限远 $\mathcal{I}^+$ 的过去. 简而言之, 所谓黑洞区就是时空上连光都逃逸不出的区域. 需要强调的是: 这里的时空是相对论性的时空, 而光逃逸不出指的是光不能够到达未来类光无限远. 黑洞区域的边界称为“黑洞事件视界”(event horizon). 因此人们常说: 所谓的“黑洞事件视界是时空未来类光无限远过去的边界”. 事件视界这个词最早由奥地利物理学家伦德勒(Rindler)^[3,4]于1956年在宇宙学的研究领域内引入.

当然他研究的是所谓的观测者的事件视界, 不同于我们这里的黑洞事件视界(图4). 1969年英国数学物理学家彭罗斯(Penrose)将这一概念发展成所谓的“绝对事件视界”, 也就是我们这里的黑洞事件视界^[5]. 当然黑洞事件视界也可以理解为一族逃逸到无限远处的观测者共有的事件视界^[5]. 因此我们也可将黑洞区定义为 $M - \mathcal{I}^-(\mathcal{S})$, 其中 $\mathcal{S}$ 是上述的所有观测者所形成的集合. 黑洞事件视界是时空中的类光超曲面(时空的3维子流形), 也就是说它的母线是类光曲线.

1.2 广义相对论中黑洞的小历史

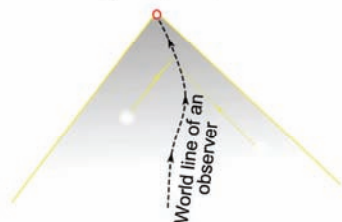
1915年爱因斯坦建立了广义相对论. 这是一种典型的相对论性引力理论, 时空度规满足的引力场方程就是著名的爱因斯坦场方程. 20世纪60年代以前, 人们关于广义相对论的研究主要集中在时空的局部性质. 典型的做法是在某个局部坐标系中求解爱因斯坦场方程和研究一些物理现象. 例如: 1916年, 在广义相对论诞生3个月左右, 德国物理学家和天文学家施瓦西(Schwarzschild)发现了爱因斯坦场方程的一个球对称解, 也就是著名的施瓦西解. 1918年, 考虑电磁场后, 莱斯纳(Reissner)和诺斯特朗(Nordström)发现了球对称的解, 也就是所谓的莱斯纳-诺斯特朗解. 在某个坐标系中求得的解只是爱因斯坦方程的一个局部解. 它极有可能是更大的一个时空的一部分. 二十世纪二三十年代, 爱丁顿(Eddington)和勒梅特(Lemaître)将施瓦西解扩张到图4中的I区和II区. 到了50年代, 辛格(Synge)意识到这种扩张的不完备性, 并将时空扩张到包含了III区和IV区. 1960年克鲁斯卡(Kruskal)和塞凯赖什(Szekeres)分别独立地实现了这种扩张. 这样人们实现了施瓦西时空的最大扩张. 1963年, 新西兰物理学家克尔(Kerr)又发现了轴对称稳态的转动解, 即克尔解. 1965年, 纽曼(Newman)发现了带电荷的克尔解, 称为克尔-纽曼解. 1967年, 博耶尔(Boyer)和伦德奎斯特(Lindquist)实现了克尔解的最大扩张. 随后, 1968年, 卡特(Carter)完成了克尔-纽曼解的最大扩张. 在这些研究中, 人们发现扩张后的时空中存在着因果不连通的区域. 例如: 1958年芬克尔斯坦(Finkelstein)意识到施瓦西时空的I区和II区是因果不连通的. 它们中间存在着因果壁垒, 物理信号无法从II区传播到I区. 这就是黑洞事件视界这个概念的雏形. 物理上我们并不要求所有的时空都是最大扩张的, 例如: 在现实世界中, 很多时候最大

(a) VISUAL HORIZONS IN WORLD-MODELS

W. Rindler

(Received 1956 November 23)

We shall define a horizon as a frontier between things observable and things unobservable. (The vague term *things* is here used deliberately.) There are then two quite different horizon concepts in cosmology which satisfy our definition and to which cosmologists have at various times devoted their attention. The first, which I shall call an *event-horizon*, is exemplified by the de Sitter model-universe. It may be defined as follows: An *event-horizon*, for a given fundamental observer A , is a (hyper-) surface in space-time which divides all events into two non-empty classes: those that have been, are, or will be observable by A , and those that are forever outside A 's possible powers of observation. It was this horizon,

(b) Gravitational Collapse: The Role of General Relativity¹

R. Penrose

Department of Mathematics, Birkbeck College - London
[Current address: Mathematical Institute, University of Oxford, 24-29 St. Giles,
Oxford, OX 13 LB, UK]
(Rivista del Nuovo Cimento, Numero Speciale 1, 257 (1969))

appropriate is the term "event horizon", since $r = 2m$ represents the absolute boundary of the set of all events which can be observed in principle by an external inertial observer. The term "event horizon" is used also in cosmology for essentially the same concept (cf. Rindler [5]). In the present case the horizon is less observer-dependent than in the cosmological situations, so I shall tend to refer to the hypersurface $r = 2m$ as the *absolute event horizon*² of the space-time

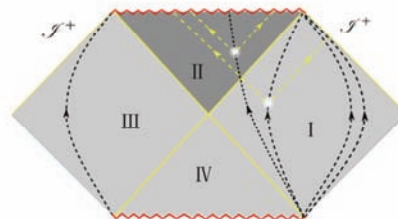


图4 (网络版彩色) (a)为某个观测者的事件视界. 观测者世界线的过去是图中的阴影部分, 也是观测者有可能探测到的时空的最大区域. 而他(她)的事件视界是这个区域的边界. 图(b)中的II区为最大扩张施瓦西时空中黑洞区B. 黑洞的事件视界是这个区域的边界. 上图中的每一点代表一个二维的曲面. 图(b)中的各种无限远已经通过共形映射拉到有限处. 这样, 压缩掉两维后, 我们可以将时空画在一张纸上. 这种图称为彭罗斯-卡特图. 在这种图中光的世界线都是和竖直方向成45°夹角的直线, 如图中的直虚线. 图(b)上下的锯齿线代表最大扩张施瓦西时空中的奇点

Figure 4 (Color online) (a) The event horizon for an observer. The past of the observer is the shadow region of figure (a) and is the largest region can be observed by the observer. The dark part of figure (b), i.e., the region II, is the black hole region of the maximally extended Schwarzschild spacetime. The event horizon of black hole is the boundary of this region. In the above figures, every point represents a two dimensional surface. In figure (b), various infinities have been pulled into a finite region by some conformal mapping, and we can draw the spacetime on a paper. These kinds of figures are standard Penrose-Carter diagrams in general relativity. In these diagrams, the world lines of light (straight dash lines) are straight lines, and the angle between the direction of light and the vertical line has degrees of 45. The jaggieszigzag lines in figure (b) are singularities of the spacetimes

扩张施瓦西时空的I区和II区所形成的小一些的时空就足够了. 在1967年之前, “黑洞”这个名词并没有在物理学界广泛使用(那时人们用“冻结星体”或“塌缩的星”等来描述黑洞). 唯一可知的记录是1964年美国女科学记者尤因(Ann Ewing)在*Science News Letter*上发表的一篇关于“黑洞”的短文(图5). 1967年12月29日, 美国著名物理学家惠勒(Wheeler)在哥伦比亚大学的一次题为“*Our Universe: the Known and Unknown*”的公众讲座中使用了“黑洞”这个词, 从那时起“黑洞”逐渐成为物理学中的一个专有名词.

直到20世纪60年代, 人们关于时空整体性质的

研究才获得实质性的进展. 1962年底, 彭罗斯引入了时空的共形完备技术. 自此, 人们关于时空渐近无限远的行为研究和时空因果结构的研究拥有了相对系统的手段. 伴随着研究方法的改进和克尔解的发现, 从1960年初到1970年初, 广义相对论和黑洞物理的研究取得了辉煌的成就. 有时候, 这10年左右的时间被称为黑洞物理的“黄金年代”. 在这段时间内人们给出了黑洞的数学定义(见上一小节)和关于黑洞的一些重要性质. 例如: 奇异性定理、唯一性定理(无毛定理)、黑洞的力学四定律, 和霍金辐射等等. 这里我们做一点简单的介绍.

20世纪30年代初, 钱德拉塞卡(Chandrasekhar)和朗道(Landau)已经意识到星体塌缩成白矮星应该有个质量上限. 事实上, 1931年钱德拉塞卡的计算给出了恒星塌缩形成白矮星的质量上限为大约1.4个太阳质量. 但钱德拉塞卡和朗道的讨论是基于牛顿力学. 1939年, 基于爱因斯坦方程, 奥本海默(Oppenheimer)和斯奈德(Snyder)研究了球对称流体的引力塌缩问

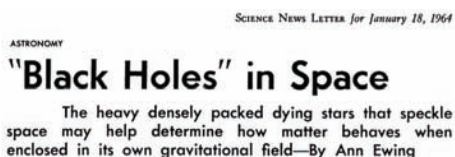


图5 安·尤因最先使用了黑洞这个名词

Figure 5 Ann Ewing first used the terminology “black hole”

题, 并发现: 当星体的质量很大时, 没有什么可以阻止星体的塌缩, 最终它必然塌缩到一个点(即施瓦西黑洞的奇点, 见图3(b)中的锯齿线). 但奇点的出现是不是和球对称的假设有关呢? 答案是否定的. 彭罗斯、霍金、盖罗奇(Geroch)的一系列奇异性定理回答了这个问题——奇点的形成和对称性没有关系. 1965年彭罗斯证明: 只要时空中存在2-维的俘获面, 奇点就不可避免地形成(但这个证明中用到了时空的整体双曲条件). 1970年霍金和彭罗斯改进了证明, 所需的条件被放松, 建立了他们著名的奇异性定理. 这些奇异性定理告诉我们: 如果广义相对论正确, 大质量星体塌缩形成奇点是必然的.

1967年, 物理学家以色列(Israel)证明了静态(无转动)黑洞的无毛定理: 真空爱因斯坦方程的静态渐近平坦解必然是球对称的, 因此是施瓦西解, 从而只需要用一个质量参数便可唯一确定. 以色列在其文章结束部分指出了将这个唯一性定理推广到转动解的情况的重要性. 随后, 在1971年, 卡特给出了一个关于克尔解唯一性的证明. 那时候人们已经逐渐相信大质量恒星引力塌缩的终态必然是克尔-纽曼解, 而且只需要用三个参数确定, 即质量、角动量、和电荷. 因此, 1971年惠勒和鲁菲尼(Ruffini)用“黑洞无毛假设”来描述黑洞的这种性质. 黑洞无毛是一个让人很吃惊的结论: 不具备任何对称性的初始引力系统, 经过引力塌缩最终会形成一个稳态轴对称的简单系统——黑洞. 对于这一现象, 1969年彭罗斯最先给出了一个解释: 在引力塌缩过程中, 初始的多极矩都被辐射殆尽, 只留下质量、角动量、和电荷这三个量^[6]. 这种想法最终被普赖斯(Price)在1972年关于引力塌缩的研究中证实. 需要指出的是: 实际上1971年卡特的证明是不完备的. 1972年霍金证明了刚性定理和拓扑唯一性定理: 稳态黑洞要么是静态, 要么是轴对称的, 以及黑洞事件视界截面在拓扑上是一个球面. 基于卡特和霍金等人的工作, 1975年罗宾森(Robinson)完善了卡特的证明. 一直到1982年克尔-纽曼解的唯一性才被澳大利亚学者邦廷(Bunting)和波兰学者马祖尔(Mazur)分别独立地证明.

霍金在1971~1972年研究了黑洞事件视界截面面积的演化, 并发现这个面积是不减的. 这就是所谓黑洞面积定理. 黑洞无毛定理表明黑洞的形成会导致熵的丢失. 为了解决熵丢失的问题, 1972年贝肯斯坦(Bekenstein)提出黑洞的熵应该正比于黑洞事件视

界截面的面积. 霍金并不认同贝肯斯坦的想法, 认为贝肯斯坦错误地理解了他的面积不减定理, 且于1973年与巴丁(Bardeen)和卡特建立了黑洞的力学四定律, 并强调: 虽然这四条定律和热力学中的四条定律非常相似, 但黑洞不是一个热力学系统. 这是因为: 通常来说, 如果有熵的话, 就需要一个温度的概念. 有温度就会存在热辐射, 而这对于经典黑洞来说是不可能的. 盖罗奇甚至在1971年设计了一个模型来反驳贝肯斯坦的说法. 贝肯斯坦处境尴尬. 但前苏联物理学家泽尔多维奇(Zel'dovich)、斯塔罗宾斯基(Starobinsky)和美国物理学家米斯纳(Misner)等人认为黑洞辐射粒子也不是不可能的, 例如: 转动黑洞彭罗斯过程的“波”版本, 即“超辐射”就是从转动黑洞的无限红移面内的能层(ergosurface)内辐射出粒子. 但超辐射只局限于转动黑洞. 1973年霍金访问莫斯科, 与泽尔多维奇和斯塔罗宾斯基讨论了超辐射的问题. 霍金认为泽尔多维奇和斯塔罗宾斯基关于超辐射的讨论在物理上是可靠的, 但不太喜欢他们计算超辐射的方式. 他希望用一个更好的方式来研究这个问题, 并与当年11月份找到了处理办法. 但他发现即使是不转动的静态黑洞也会辐射出粒子, 而且无限远处观测者看到的辐射粒子的谱是一个热谱, 相应的温度正比于黑洞的表面引力, 这就是霍金辐射, 而这个温度被后人称为霍金温度. 霍金并不希望支持贝肯斯坦的想法, 因此反复检验了计算, 但并没有发现错误^[6]. 最终, 霍金不得不接受了贝肯斯坦关于黑洞熵的想法, 并给出了黑洞熵和面积的比例系数 $-1/4$. 这个黑洞的熵的表达式被人们称为贝肯斯坦-霍金熵. 量子效应被考虑进来之后, 黑洞的力学四定律变成了真实的热力学四定律. 但霍金辐射的存在也表明黑洞会因为这种量子效应丢失能量, 并最终变成一堆辐射物质, 即黑洞完全蒸发了. 可见在黑洞形成并蒸发的过程中, 形成黑洞的物质的信息丢失了. 这就是所谓的黑洞信息丢失疑难.

1974年, 随着霍金辐射的发现, 至今黑洞物理最为辉煌的时代逐渐落下了帷幕. 在这个黄金时代, 群星璀璨, 广义相对论中许多重要的问题被解决, 同时也提出了很多新的问题. 这里我们只列举了与黑洞密切相关的部分. 其中的很多问题仍然是当前引力理论中重要的研究方向. 例如: 黑洞信息丢失问题. 再例如: 在引入宇宙常数时, 前面所述的那些定理会有什么样的改变?

1.3 黑洞理解的演变

至此,相信读者已经知道什么是黑洞,什么是黑洞事件视界,也已经对黑洞物理有了一定的认识.但需要指出的是:黑洞事件视界依赖于时空的整体因果结构.或者换句话说,为了知道黑洞的事件视界我们需要知道时空的一些整体信息,例如时空的类光无限远.这使得黑洞事件视界在实际应用上非常不方便,因为一般来说我们不可能知道我们的未来,不可能知道我们所处时空的整体因果结构.原则上我们无法局部地定义黑洞事件视界,更有甚者,在局部的观测者看来黑洞事件视界会具有一些反直觉的行为.这就是事件视界的“目的论”(teleological)特性.霍金和埃利斯很早就注意黑洞事件视界的非局部特征,并在其著作*The Large Scale Structure of Space-time*一书中引入了表观视界(apparent horizon)的概念^[7].这个视界的定义依赖于4-维时空中2-维闭合曲面的外几何与内禀几何,而不是时空的整体因果结构.但霍金和埃利斯的表观视界仍然依赖于时空的一种整体结构,即时空的分层.因此在实际应用中可操作性依然不强.20世纪90年代,人们将表观视界的定义推广,建立了一些准局域视界.其中颇具影响的是阿什塔卡(Ashteka)^[8]提出的孤立视界和动力学视界^[7].

前文提到的彭罗斯的俘获面(trapped surface)在定义这些准局域视界时起到了重要的作用.考虑时空中的一个2-维的闭合曲面.现在我们考察这个曲面上向外和向内发射的光.如果时空是通常的闵氏时空,那么向外发射的光是发散的,而向内发射的光是汇聚的(图6).这表明向外发射的光所成线汇的扩张标量为正,而向内发射光所成线汇的扩张标量为负.但如果曲面所处的时空弯曲的比较厉害,向外发射光所成的线汇也会变成汇聚!即相应扩张标量变为负.这样的曲面称为是俘获面.这表明曲面所在处的引力场如此之强,以至于连光都逃逸不出.向外发射光所成线汇的扩张标量为零的曲面是很特别的,称为边缘俘获面(marginally trapped surface).粗略地说,所谓的准局域视界就是这些(无穷多个)边缘俘获面所组合成的3维超曲面.和事件视界不同的是:这些准局域视界可以是类空的(动力学视界),可以是类光的(孤立视界),也可以是类时的.甚至一些准局域视界可以一部分是类空的,一部分是类光的,而另外一部分是类时的.这是因为准局域视界和时空的整体因果结构没有必然的联系,故它们和黑洞事件

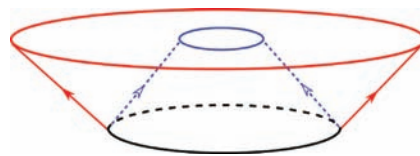


图6 (网络版彩色)黑色圆圈表示2-维类空曲面,实线箭头表示向外发射的光,而虚线箭头表示向内发射

Figure 6 (Color online) The black circle represents the 2-dimensional spacelike surface. The solid arrows denote the outward light, while the dash arrows depict the inward light

视界的这种不一样是一件很自然的事.但对于稳态黑洞这种理想情况,这些准局域视界和黑洞事件视界是重合的.俘获面一旦形成,时空中便会出现奇点,因此这些新的视界更能反映出黑洞的强引力场性质,且更为接近我们现实世界中对黑洞的描述.另外一方面,准局域视界(如孤立视界)在圈量子引力中具有重要的地位.因此近20年来,基于准局域视界的物理也得到较大的发展,并引起了很多学者的重视.这些新的发展都是在广义相对论的框架下进行的(当然也容易推广到其他的相对论性引力理论).它们极大地丰富了广义相对论中黑洞物理的内涵,并加深了人们对于黑洞的认识.

1.4 总结

从暗星到黑洞事件视界,再到所谓的准局域视界,我们可以看出:米歇尔和拉普拉斯的光是否能逃逸出这种朴素的想法是黑洞这个概念产生的根源.相对论性时空中的黑洞事件视界,准局域视界都是这一想法的具体实现中出现的重要物理对象.简而言之:“所谓黑洞就是相对论性时空中连光都逃逸不出的区域”.这就是“经典理论中黑洞的本质”.

2 黑洞的量子本质

但问题并非这么简单.这是因为黑洞不可避免地会产生一些量子效应.例如前面提到的霍金辐射和贝肯斯坦-霍金熵都只有在引入量子理论后才可获得解释.因此可以说没有量子理论就没有黑洞热力学.半经典的弯曲时空量子场论可以解释霍金辐射、霍金温度、昂鲁温度等等,但给不出令人信服的贝肯斯坦-霍金熵的量子统计起源.为了真正解释黑洞熵,还需要一个量子引力.因此在谈及黑洞本质时,量子引力是绕不过去的话题.可是,至今为止,还没有一个公认的量子引力理论.这里我们只准备告诉读者:

在几种可能的量子引力候选理论中黑洞是什么, 贝肯斯坦-霍金熵是怎样得到了统计解释.

若直接将引力场量子化, 无论是协变量子化(以某个最大对称空间作为背景, 如闵氏时空. 理论存在该最大对称空间的 $SO(1,3)$ 最大对称性, 如闵氏时空的庞伽莱不变性, 故称为协变)还是正则量子化(需要对时空进行 $1+3$ 分解)都会遇到很大的困难. 协变量子化过程中遇到的最严重的问题是理论不可重整化, 而正则量子化面临的问题是最后的波动方程是一个数学上没有严格定义, 几乎无法求解, 且具有无限多个自由度的泛函微分方程(Wheeler-DeWitt方程)^[9]. 协变量子化最终放弃了点粒子模型, 发展成为弦理论. 而沿着正则量子化这条路, 人们发现了圈量子引力.

2.1 弦理论中的黑洞

弦理论是一个天然的量子引力理论, 因为引力子自然地出现在闭弦(拓扑上是个圆环)的无质量振动模式中. 当然, 弦理论不仅仅是量子引力, 还能描述自然界中其他的3种相互作用. 除了无质量振动模式, 弦还拥有无穷多的有质量震动模式. 在协变量子化的框架中, 从微扰的角度来看, 时空是引力子的相干态. 因此在弦理论中, 弯曲时空可以看成是闭弦的“相干态”^[10]. 弦理论中除了1维的基本弦, 还包含了很多非微扰对象, 也就是所谓的膜(brane), 如各种空间维数的D-膜^[10]. 这些膜是一些反对称张量场的源, 并带有相应的荷. 弦理论中牛顿常数不是一个基本量, 它由弦的张力所给出的长度标度和弦与弦之间的耦合强度来决定. D-膜的质量反比于弦的耦合常数, 但其生成的引力场正比于弦的耦合常数. 1993年萨斯坎德(Susskind)基于微扰弦提出一个假设: 弦的激发态与黑洞存在一个1-1对应关系(质量很大的激发态塌缩成黑洞). 但这种假设很快被证明有问题——因为黑洞的熵和质量的平方成正比, 而弦的熵与质量成正比. 1997年泡尔钦斯基(Polchinski)和霍洛维茨(Horowitz)改进了这一想法, 并论证弦的确具有足够的自由度来形成黑洞. 但这一图像并不能给出准确的黑洞微观自由度数目, 无法实现黑洞熵的量子统计解释^[11]. 从非微扰的角度来看, 这个目标是可以实现的. 弦论中的黑洞可以看成是一些膜的构型. 而由一些D-膜形成的构型可以看成是没有霍金温度的极端黑洞(粗略地说, 质量 M 和荷 Q 相等). 当然因为黑洞的引力场是较强的, 弦处于强耦合区. 假定这些

D-膜保持了理论的一部分超对称(超弦理论), 并想象弦的耦合强度被逐渐调弱到可以忽略不计, 这样D-膜的引力效应随之可以忽略. 在这种弱耦合极限下, 系统变成闵氏时空上的一个D-膜束缚态, 也可以看成是弦理论的一个荷为 Q 的BPS态(图7). 因为在弱耦合极限下, 荷为 Q 的BPS态的数目是可以计算的. 又由于部分超对称的存在, 这个数目不依赖于弦的耦合常数. 这样, 人们便可以计算出荷为 Q 的极端黑洞所拥有的微观自由度数目, 从而计算出黑洞的统计熵. 当 Q 很大时, 这个统计熵恰好是贝肯斯坦-霍金熵! 这就是斯特罗明格(Strominger)和瓦法(Vafa)1996年完成的工作. 之后, 人们也发现了近极端黑洞和中性转动黑洞熵的统计起源. 但需要指出的是: 弦理论中的这种黑洞和现实的黑洞还是有很大的距离. 例如: 现实的黑洞荷可以忽略不计. 弦论学者们仍在努力寻找施瓦西黑洞熵的统计起源.

2.2 圈量子引力中的黑洞

圈量子引力是一种比较单纯的量子引力理论. 它并没有试图统一自然界的四种相互作用. 广义相对论相空间上的约束是比较复杂的, 这使得传统正则量子化后的约束方程几乎无法求解(Wheeler-DeWitt方程). 但基于1986年阿什提卡提出的“阿什塔卡新变量”, 人们可以将广义相对论相空间上的正则变量写成类似于SU(2)规范理论的形式, 从而得到一个关于

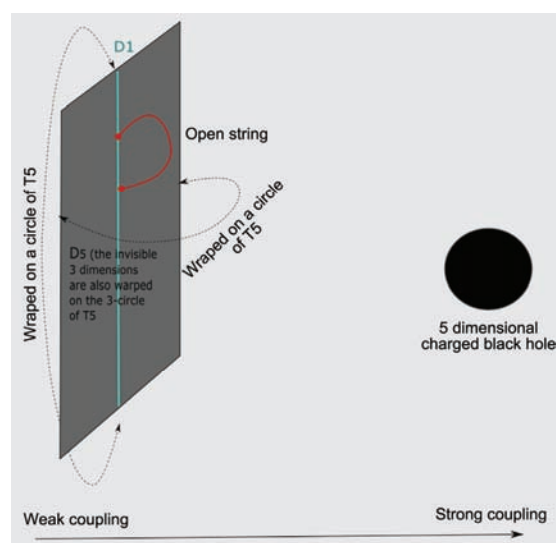


图7 (网络版彩色)弦理论中的黑洞示意图, D_1 和 D_5 系统和 5-维黑洞
Figure 7 (Color online) A simple sketch of the black hole in string theory. D_1 and D_5 system and 5 dimensional charged black hole

联络的动力学而不是关于度规的动力学. 这样规范场量子化的方法便可以用来研究引力场量子化的问题. 需要注意的是: 这里的正则变量(如正则动量)继承了广义相对论中度规的特征——既有动力学的角色又有背景场的角色. 这和通常闵氏时空上的规范理论中规范场是动力学场而度规是背景场的情况很不一样. 另外, 正则变量的构造并不唯一. 事实上, 人们可以构造出一族这样的正则变量. 这一族正则变量由一个正的参数(称为Barbero-Immirzi参数)来描述. 在经典理论中, 这种不确定性没有任何影响. 但在量子层面上, 不同的Barbero-Immirzi参数对应的量子理论是不同的. 数学上严格的场的量子化需将场做一个空间上的抹平(空间上的积分, 依赖于度规体元)来构造量子化所需的基本变量(以避免一点处的紫外问题). 一个量子引力理论中的基本变量的构造方式不应该是一种依赖于度规的抹平方式, 因为度规本身就是要进行量子化的对象. 如何构造出这种不依赖于度规的基本变量呢? 联络是1-形式, 因此考虑它沿着某条曲线的积分便是一种不依赖于度规的抹平方式, 这对应于规范场中的威尔逊圈(或和乐, 圈引力中圈的由来). 而正则动量可以看成是一个2-形式, 因此它在某个2-维曲面上的积分给出的通量也是一个不依赖于度规的抹平. 这样威尔逊圈(和乐)和通量便成为圈量子引力中的基本变量(对应于量子力学中的坐标 x 和动量 p). 将它们量子化后便得到和乐算子和通量算子. 藉此, 在20世纪90年代中后期阿什塔卡等人^[8]建立了严格的量子几何理论. 量子几何中的几何算子可以由通量算子构造出来, 例如某个2-维曲面的面积算子和某个3维区域的体积算子. 对于具有孤立视界的时空(黑洞), 孤立视界提供了时空的一个“内边界”. 阿什塔卡等人研究了这样时空的量子几何, 找到了希尔伯特空间中满足视界边条件的子空间(可称为黑洞的希尔伯特空间), 并得到了该子空间的维数. 这个维数对应于黑洞的微观自由度, 且可得黑洞熵的领头项正比于孤立视界(截面)的面积. 但比例系数中有一个不确定的因子, 即Barbero-Immirzi因子. 只有当这个因子取特定值的时候才能够回到贝肯斯坦-霍金熵的形式. 在圈量子引力中, 黑洞量子视界的图像是比较直观的. 视界外部的和乐可看成是一些类似于polymer的物体, 它们在与视界相交处刺穿了视界. 因此黑洞的量子视界不再是光滑的曲面, 见图8.

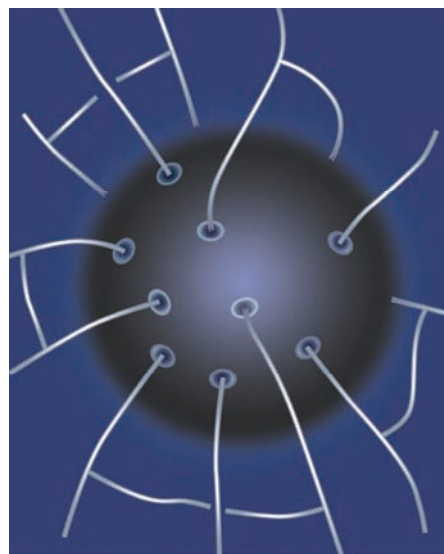


图8 (网络版彩色)圈量子引力中, 黑洞的量子视界不再是光滑曲面. 图片来源于PhysOrg.com

Figure 8 (Color online) The black hole in loop quantum gravity. The quantum horizon is not a smooth surface any more. The picture comes from PhysOrg.com

2.3 全息图像下的黑洞

受黑洞贝肯斯坦-霍金熵的启发, 在1994年, 胡夫特(t'Hooft)和萨斯坎德(Susskind)提出了引力具有全息性质的概念: 一个引力系统的独立自由度由它的边界面积来测度. 1997年底, 马尔德西纳(Maldacena)在弦理论中精确地实现了这一思想: $AdS_5 \times S^5$ 上的IIB型超弦理论等价于 AdS_5 边界上的 $\mathcal{N} = 4$ 的超对称杨-米尔斯规范场理论. 随后, 威滕(Witten)等人猜测: $(D+1)$ -维AdS时空中的量子引力理论等价于AdS边界上 D -维的共形场论. 这就是所谓的AdS/CFT对应. 按照AdS/CFT对应, 纯AdS时空对应于边界上零温的共形场论, 而AdS黑洞对应于边界上有限温度的场论, 且边界场论的温度就是AdS黑洞的霍金温度. 这样, 人们便可以通过计算边界场论的熵来确定AdS黑洞的熵. 计算表明边界场的熵和黑洞的贝肯斯坦熵只差一个 $3/4$ 因子. 这个因子的存在是因为边界场论的计算是在弱耦合下进行的. 因此在从事AdS/CFT对应研究的学者看来, AdS黑洞是边界场的热激发态. 进一步地, 按照胡夫特和萨斯坎德的想法, 全息是量子引力的一个基本特征. 它可以和弦理论没有任何关系, 也可以和AdS时空没有关系. 如果任何一个引力系统完全可以由其边界上的某种没有引力参与的量子理论来描述, 那么全息原理便可以看成是引力

量子化的另外一种实现. 它不同于前面的协变量子化或正则量子化的途径, 而是一种新的量子化方式. 黑洞, 在这种图像下, 自然应该对应于边界上的某种量子态(图9). 当然, 现在人们关于引力全息性质的研究还在继续.

3 结语

在经典理论中, 黑洞的本质是明确的: 相对论性时空中连光都逃逸不出的区域. 在量子引力理论中, 黑洞是某种量子态. 但在不同的量子引力理论中, 黑洞的图像是非常不一样的. 到底哪一种量子理论真实地描述了黑洞? 对于这一问题, 我们还知之甚少. 但无论如何, 在量子层面上, 黑洞肯定不是简单的由质量、角动量和电荷这三根毛来标识. 它应该具有更多的量子毛. 事实上, 最近霍金、派瑞(Perry)和斯特罗明戈提出了黑洞具有软毛的想法, 用以解释黑洞的信息丢失问题, 并引起了人们的关注^[12]. 不可否认, 关于黑洞的量子本质的研究是现在理论物理学界重要的研究方向, 也是现有或将会出现的量

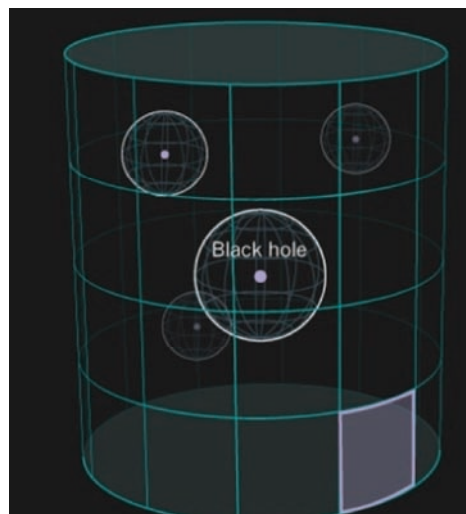


图9 (网络版彩色)全息图像下的黑洞. 图片来源于www.nature.com
Figure 9 (Color online) The black hole in the scenario of holography. Picture comes from www.nature.com

子引力理论无法回避的研究课题. 人们相信对黑洞物理解的不断加深一定会对相对论天体物理, 量子引力, 信息理论带来革命性的发展.

参考文献

- 1 Michell J. On the Means of Discovering the Distance, Magnitude, &c. of the Fixed Stars, in Consequence of the Diminution of the Velocity of Their Light, in Case Such a Diminution Should be Found to Take Place in any of Them, and Such Other Data Should be Procured from Observations, as Would be Farther Necessary for That Purpose. In: A letter to Henry Cavendish. Philosoph Trans R Soc London, 1784, 74: 35–57
- 2 Rindler W. Visual horizons in world models. Month Not R Astron Soc, 1956, 116: 662–677
- 3 Penrose R. Gravitational collapse: The role of general relativity. Riv Nuovo Cim, 1969, 1: 252–276
- 4 Penrose R. “Golden Oldie”: Gravitational collapse: The role of general relativity. Gen Rel Grav, 2002, 34: 1141–1156
- 5 Abbott B P, Abbott R, Abbott T D, et al. [The LIGO Scientific Collaboration, The Virgo Collaboration]. Observation of gravitational waves from a binary black hole merger. Phys Rev Lett, 2016, 116: 061102
- 6 Hawking S W, Israel W. Three Hundred Years of Gravitation. Cambridge: Cambridge University Press, 1987
- 7 Hawking S W. The Large Scale Structure of Space-Time. Cambridge: Cambridge University Press, 1973
- 8 Abhay A, Badri K. Isolated and dynamical horizons and their applications. Liv Rev Rel, 2004, 7: 1–91
- 9 Ashtekar A. Introduction to Loop quantum gravity and Cosmology. 2012, arXiv:1201.4598
- 10 Polchinski J. String Theory. Cambridge: Cambridge University Press, 2005
- 11 Horowitz G. Spacetimes in string theory. 2004, arXiv:gr-qc/0410049
- 12 Hawking S W, Perry M J, Strominger A. Soft hair on black holes. 2016, arXiv:1601.00921

The nature of black holes

CAI RongGen¹ & CAO LiMing²

¹ Institute of Theoretical Physics, Chinese Academy of Sciences, Beijing 100190, China;

² University of Science and Technology of China, Hefei 230026, China

What is a black hole? It is a little bit difficult to answer this question. For general public, black holes frequently appear as strange celestial bodies in science fiction movies. However, for researchers, they are real presence in our universe and the regions of spacetimes in which nobody can escape even for light. Although the level of awareness is different, the black holes are mysterious to all of us. The black holes are mysterious so much so that we cannot really understand them up to date. What is the nature of the black holes? If people further ask. Probably, the understandings are different from person to person. On occasion, one even gets distinct answers from the experts in the physics of black holes. In this article, we will not provide some particularly striking ideas on this question, but we would like to tell readers the popularly accepted viewpoints in the community of this research field.

black hole, gravity, quantum gravity

doi: 10.1360/N972016-00186



蔡荣根

1995 年 6 月获得复旦大学理学博士学位。相继于中国科学院理论物理研究所、韩国汉城国立大学理论物理中心、日本大阪大学物理系从事理论物理研究。2000 年底入选中国科学院“引进国外杰出人才计划”。现任中国科学院理论物理研究所研究员，博士生导师，中国引力和相对论天体物理学会理事

长，国际广义相对论和引力学会理事，亚太物理学会引力、天体物理和宇宙学分会副理事长。2011 年获得国家自然科学二等奖。长期从事引力理论、宇宙学、量子场论和超弦理论等方面的研究，发表篇论文 200 余篇，被引用 1 万余次。



曹利明

2008 年 1 月在中国科学院理论物理研究所获得博士学位。随后分别在韩国和日本从事博士后研究工作。从 2011 年 12 月至今为中国科学技术大学物理学院教授，博士生导师。主要研究方向是引力理论以及相关问

题，例如：黑洞物理学，引力的全息性质等。特别对时空中各种可能视界的力学、热力学做了一些研究，发表了一系列文章，引用 1500 余次。