

再论质量的起源

王青[†]

(清华大学物理系 北京 100084)

摘要 物质的质量是怎么来的? 能从理论上研究质量的起源吗? 人类目前所能做的实验是否到了可探索这些质量起源的程度? 本文将介绍目前物理学对这些问题的认识. 我们首先通过修改牛顿第零定律引入结合能来介绍物质的各个不同层次上的质量起源的研究链条, 然后重点讨论基本粒子质量的起源, 并介绍基本粒子质量与对称性的关系, 最后讨论三种与质量起源有关的质量产生和分裂的机制.

关键词 质量起源, 结合能, 电弱对称性, 对称性自发破缺

The origin of mass revisited

WANG Qing[†]

(Department of Physics, Tsinghua University, Beijing 100084, China)

Abstract Where does the mass of matter come from? Can we trace its origin? Does our present experimental ability reach the level needed to investigate the origin of mass? The purpose of this paper is to present our current understanding of these problems from the viewpoint of physics. We first revise Newton's Zeroth Law while introducing the binding energy to demonstrate the history and various levels of research on the origin of mass, then focus our attention on the origin of the mass of elementary particles. We will discuss the connection between mass and symmetry, and list three mechanisms of mass generation and splitting that are related to the origin of mass.

Keywords origin of mass, binding energy, electroweak symmetry, spontaneous symmetry breaking

在《物理》2008年第37卷第5期上, 中国科学院物理研究所曹则贤先生在他的系列文章“物理学咬文嚼字之十一”中, 以“质量与质量的起源”为题, 将人类对质量的认识作了十分精彩的描写和介绍. 对质量起源, 曹先生只在文章最后一段进行了简要的阐述, 提到: “质量的起源还是个未完全解决的问题……物质的质量来自原子, 原子质量主要来自核子, 核子由夸克组成, 但夸克由无惯性质量的夸克组成. 囚禁夸克的能量在核子层面上表现为质量. 这算是对核子质量起源的一个交代, 但对电子质量的起源, 目前尚无理论上的解释. 可能, 关于质量的起源, 最终还是落在无质量的存在上, …… Wheeler 教授就宣扬 ‘mass without mass’ 的观点……不知将来这一伟大壮举要着落在谁人的肩上”. 我们以曹先生

文章的这个结尾作为本文的开始, 续写一段目前人类所能获得的对质量起源更深入的认识. 需要说明的是, 质量起源是十分复杂和困难的问题, 需要很多的篇幅逐步地展开和深入. 我们努力避免展示深奥的数学, 试图只用文字或最简单的公式尽量把复杂理论背后的物理图像、思想及逻辑脉络诠释清楚.

质量是物质最重要的属性, 从这个角度上看, 质量的起源某种意义上就是物质的起源. 在宇宙发展演化的极早期, 存在哪种物质会对以后的宇宙发展演化起导向性作用. 2009年是国际天文年, 天文现象主要是由引力相互作用控制的, 而质量在传统意

2009-05-08 收到

[†] Email: Wangq@mail. tsinghua. edu. cn

义上是产生引力作用的源,在宇宙的演化中,它的作用是导致宇宙收缩(而观测到的宇宙却在加速膨胀,为此要求存在产生“负压”的所谓“暗能量”,这是宇宙学当今的热门课题,但不是本文的主题,本文只限于讨论传统意义上的质量的起源),因此探索质量的起源还可以被看成是探索产生各种各样天文现象的引力之源的起源.按照广义相对论的观点,物质的存在导致时空结构的改变,而关于质量的起源的研究也就是关于我们生存的宇宙中改变时空结构之源的起源的研究.从这几方面看,关于质量起源的讨论不仅有它自己独特的重要意义,而且和国际天文年的主题也有内在联系.

曹先生文中提到物质是有层次结构的:物质由原子组成,原子由电子和原子核组成,原子核由主要是质子和中子的强子组成,强子由夸克组成.目前人类能够了解到的物质最深层次是夸克和轻子(电子是最轻的轻子,另外还有 μ 轻子和 τ 轻子)层次.在物质的每个层次上,我们都可以同样地问质量是怎么来的?除了最深的夸克和轻子层次外,在各个物质层次上存在有类似的研究质量起源的方法.以原子层次为例,考虑到原子由若干电子和原子核组成,原子的质量可写成构成原子的所有电子的质量加上原子核的质量,再加上这些电子之间及它们和原子核之间的相互作用造成的有效质量.这里需要特别强调一下这个有效质量,它来自于电子之间以及它们和原子核之间的相互作用产生的相互作用能(通常称之为结合能).按照爱因斯坦的质能关系,这个结合能除以光速的平方就是一个质量(鉴于这种质量和能量之间的简单关系,在以下的文章介绍中不再区分其差别),也就是我们称之为的有效质量.这个有效质量在宏观的物质层次上相对于组分的质量,其取值很小,以致可以被忽略,此时我们就回到曹先生文中提到的质量所具有的可相加性,即所谓“牛顿第零定律”.它可表述为:物质的质量等于构成它的各个组分质量的叠加.当我们深入到微观的各个物质层次中时,构成物质的各个组分之间的相互作用部分对物质质量的相对贡献越来越大,例如到原子核的层次,一些质量非常大的原子核像铀、钚等,会在吸收1个中子以后分裂成2个或更多个质量较小的原子核,同时放出2个到3个中子和很大的能量,形成所谓的核裂变反应.原子核在发生核裂变反应时,释放出的就是结合能,通常称为原子核能,俗称原子能.1克铀²³⁵完全发生核裂变后放出的能量相当于燃烧2.5吨煤所产生的能量.因此,进

入到微观的层次,“牛顿第零定律”必须被修改才能与实际相符,这个“修正的牛顿第零定律”的表述为:物质的质量等于构成它的各个组分质量和组分之间相互作用造成的有效质量的叠加.由于物质的组分只有深入到组分所在的层次才有可能被详细研究,在原来物质的层次上通常只能通过唯象地引入质量参数进行有效的描述,这个质量参数在本层次无法被计算出来而只能被作为理论的输入,而在每一层次能被计算的是物质组分之间的相互作用造成的有效质量,或者说在每个层次上能够并需要计算的是组分之间的结合能.因此在每个层次计算物质的质量,我们的办法是输入组分的质量参数,计算结合能,然后再按“修正的牛顿第零定律”叠加出总的质量.具体地说:在原子层次,我们通过输入电子和原子核的质量,计算原子内部电子之间以及电子与原子核之间的结合能来得到原子的质量,这就构成了对原子能谱的计算;而到原子核层次,我们通过输入强子的质量,计算原子核内强子之间的结合能来得到原子核的质量,它构成了对原子核能谱的计算……等等.从这样的分析看,质量起源的研究贯穿了物质的每个层次,各个层次上的研究主要是输入组分的质量参数并计算结合能,每个层次上的结合能的产生对应着不同的物理机制,而不同层次的物理都给我们理解“质量”时以一定程度的解释.从探索质量起源的角度看,这样做不够彻底,因为组分的质量参数不是计算出来的,它的起源有赖于下一个更深层次的研究.也就是说,每个层次对质量的研究都会有遗留到下个更深层次才能解释的质量参数:在原子层次,我们遗留了电子和原子核的质量;为计算原子核的质量,我们进入原子核的层次,但又遗留了强子的质量;为计算强子的质量,我们进入强子的层次,但又遗留了夸克的质量(需要指出,在夸克和轻子层次,不像曹先生文中所说夸克无质量,夸克是有质量的,而且不同的夸克有不同的质量,我们说的质量参数通常对应粒子物理专业上所说的夸克的“流质量”).令人奇怪的是,电子跨越了原子核和强子的层次,从原子层次直接进入到了夸克和轻子的层次.如果我们把质量起源的研究广义地理解为试图从某个理论中计算出质量的值,我们将要涉及到构成物质的所有层次;如果我们把质量起源的研究进一步狭义化理解为追究这些组分的质量是怎么来的,我们关于质量起源的讨论就可以沿着物质的层次一层层深入地追究,最后追溯到目前人类了解到的物质最小的组分的质量起源,也就是被称为基本粒子的夸克和轻子的质量的起源.这

是质量起源的真正的源头,本文以下将集中讨论这些基本粒子的质量的起源.

对基本粒子质量的起源的研究有两种思路:第一种思路是继承以往各个物质层次的做法,认为基本粒子并不基本,它们仍是由其组分——“更基本”的粒子组成的(这可以引发哲学方面的各种讨论).这样就可以沿用以前的做法,将基本粒子的质量表述为其组分的质量和组分之间的结合能对质量的贡献的叠加.在基本粒子这一层,我们只能研究结合能,而必须像以往一样把其组分的质量留给更深层次的研究.按此种思路,对基本粒子质量起源的研究只不过是从小到大的物质各层次研究质量起源的环环相扣的链条中的一环,相对于其他物质层次的研究,并没有什么特别之处.第二种思路是不认为基本粒子有更深层次的结构,也就是它不再由更小的组分构成,这样在基本粒子的层次,就会终结以往其他层次质量研究的链条,因为不再存在有更深的层次了.对第一种思路,现实是目前实验上并未发现基本粒子具有更深层次结构的迹象,具体体现在如下三个事实上:(1)若基本粒子有结构,它就会有所谓的“形状因子”.例如:质子有内部结构,物理上就能测量到质子的“形状因子”.对基本粒子,目前实验上尚未测量到它们的形状因子.(2)如果基本粒子有结构,它的质量谱应该有一定的规律性.例如原子的能谱、原子核的能谱、强子的能谱都存在有明显的规律性.而目前的基本粒子的能谱,好像相当不规则.(3)如果基本粒子是由更基本的组分构成的,这些组分的各种排列组合应该可以构成更多可能的新的基本粒子,而不只是目前实验上发现三代的基本粒子,应会有第四代、第五代……,等等,但目前人们从粒子物理的实验和宇宙学的讨论中都分别得到了对基本粒子只能有三代的限制.鉴于这些不利于基本粒子有结构或组分的事实,目前认为基本粒子是由更小的组分构成的观点已经不再是学术界的主流思想,因此我们后面不再对这种可能性进行更多的讨论,而集中对第二种思路进行更深入的探讨.

如果基本粒子不再由更小的组分构成,我们以往建立在组分及其相互作用基础上的研究质量的方法就不再适用了,需要有全新的机制.正像曹先生文中提到的那样,质量的起源要落脚在无质量的存在上,即现在探索基本粒子的质量的起源的想法主要是研究如何使无质量的基本粒子通过某种机制得到质量,或者称之为质量的产生,意指质量从无中生成出来.这种从无到有的思想,相对以往组分质量叠加

的想法,对质量起源问题看起来更像一个彻底的根本解决方案,其中核心问题是解决如何把一个质量参数从其取值为零变为取某个非零的有限值.从一个更高的角度看,某个量取值为零和非零往往可以和某种抽象对称性的存在与否相联系,即:对称性存在要求这个量取值为零,而当这个量不为零时,对称性就不存在或称之为被破坏了.人们经常把这种通过取值为零与否来标志对称性的有无的量称之为对称性的“序参量”.因此在目前的解决方案中一种可能的做法是把质量参数看成是某种对称性的“序参量”,这样质量的产生就与某种隐藏在背后的支配性的对称性的破缺联系起来.对称性及其破缺是2008年诺贝尔物理学奖的主题,作者在《物理与工程》杂志2009年首期受邀专门撰文对此进行过介绍,本刊和《现代物理知识》等杂志也曾分别刊登过专家们的介绍文章.这是作者为什么先写文介绍对称性及其破缺再写本文的原因之一.

为了说清楚基本粒子质量与对称性的关系,需要先从小相对论对基本粒子的分类说起.狭义相对论要求单粒子态具有自旋量子数 j ,并且这个 j 只能取分立的正整数或半正整数,而从量子力学的基本假设和狭义相对论结合而成的量子场论的自洽性出发可以导出著名的自旋统计关系,它指出自旋为整数的粒子必须是玻色子,自旋为半整数的粒子必须是费米子.量子场论进一步对非零自旋的粒子提出要求:(1)若其无质量,它只能由两个态构成(例如自由空间的光子只能有两个独立的横向极化状态),这两个态通常被选为自旋平行和反平行于粒子动量的态,分别被叫做右手态和左手态.右手态代表粒子的自旋和动量的方向之间呈现右手螺旋关系,左手态代表自旋和动量的方向之间呈现左手螺旋关系.这两个态之间可以通过空间反射相互转化.狭义相对论要求零质量粒子只能以光速运动,这导致不存在有参考系能使粒子的运动反向,因此右手态和左手态之间是不可能通过观测者的参考系选择相互转化的.(2)若其有质量,它将由 $2j+1$ 个态构成,在粒子静止的参考系中, $2j+1$ 个态的自旋在空间中的指向各不相同.不同于零质量粒子,这些态是通过观测者的参考系选择相互转化的,因为总可以选择参考系使粒子反向运动.除了自旋为零的粒子(它不出现在目前业已被发现的基本粒子的队伍中,但后面将要提到的与基本粒子质量起源有密切关系的尚未被发现的 Higgs 粒子却是自旋为零的粒子),自旋为 $1/2$ 的粒子与自旋大于 $1/2$ 的粒子在其

质量与对称性的关系方面有如下本质的不同:

(1) 对自旋为 $1/2$ 的粒子(基本粒子中的夸克和轻子都属于这类粒子),无论其有质量与否,它都只由两个态构成(因 $2j+1$ 在 $j=1/2$ 时为 2),唯一的差别是无质量时,两个态之间不能通过参考系选择相互转化,而有质量时,两个态可以通过参考系选择相互转化.因此对自旋为 $1/2$ 的粒子,质量成为其两个状态能否通过参考系选择相互转化的关键,而在量子场论中,质量参数就具体体现为把左手态和右手态耦合在一起的所谓质量项的系数.左右手态因存在质量而可以通过参考系选择相互转化,导致理论中不可能存在左手态或右手态自己独立的对称性,因为体系的左手态和右手态是关联在一起的,态之间要存在对称性就必须是左手态和右手态联合在一起做同样变换的对称性,而非左手态或右手态各自单独或做不同变换的对称性.这种左手态或右手态各自独立或做不同变换的对称性在粒子物理中被称为手征对称性,其中手征一词反映对称性是带左手或右手的手性这一特点.因此结论是:若一些自旋为 $1/2$ 的粒子态具有质量,则它们之间一定不能具有手征对称性.或者说手征对称性一定是一个被破缺了的对称性.这个结论的逆反表述为:若一些自旋为 $1/2$ 的粒子态具有手征对称性,则它们的质量必然为零.这样,自旋为 $1/2$ 的粒子质量从无到有的产生就和它可能具有的手征对称性的从有到无,或者说手征对称性的破缺联系起来.研究自旋为 $1/2$ 的粒子质量起源转化成了探索自旋为 $1/2$ 粒子所具有的手征对称性及其破缺的机制.除了手征对称性外,自旋为 $1/2$ 的粒子如果有质量,由于左手态和右手态之间可以通过观测者所在参考系的选择相互转化,因而是对称的.我们还可以建立所谓的左右手对称性,这是一个分立的对称性,它把左手态变成右手态,或把右手态变成左手态.反过来,如果自然界没有这种左右手对称性,自旋为 $1/2$ 的粒子必须没有质量,因此自旋为 $1/2$ 的粒子质量从无到有的产生还会导致体系左右手对称性的产生.进一步可以证明,对另一种分立的对称性变换—电荷共轭变换,它可以把左手正荷态变换成右手的负荷态,而对纯中性粒子的负荷态与正荷态,因为荷为零是一样的,因此自旋为 $1/2$ 的粒子如果是纯中性的(例如中微子),其左手态经过电荷共轭变换后得到的电荷共轭态也可被看成是另一种的右手态,因而存在把纯中性的左手态与其电荷共轭态耦合在一起的一种新的质量参数,叫 Majorana 质量.中微子因其质量可以

具有 Majorana 质量的成分而导致中微子的质量与其他基本粒子的质量可能有质的不同.

(2) 对自旋大于 $1/2$ 的粒子,也就是自旋等于 1 或大于 1 的粒子(基本粒子中的传递相互作用的粒子都是自旋为 1 的规范粒子).如前所述,如果它有质量,将至少有 3 个以上的态(因 $2j+1$ 在 $j \geq 1$ 时大于等于 3);而如果无质量,则只有 2 个态.物理上并没有理由禁止我们将无质量的粒子看成是有质量的粒子把质量参数取趋于零的极限.但我们发现在取零质量的极限时,构成粒子的态的数目不对了.以自旋为 1 的粒子为例,从粒子有质量出发,在质量参数趋于零时,仍得到粒子有 3 个态,而不是零质量粒子的 2 个态,虽然这时质量参数已经取成零了.为了确保我们取零质量极限得到的确实真是物理的零质量粒子,3 个态中至少需要有 1 个态(通常是所谓的纵向态)在物理上是观测不到的,这就意味着存在某种对称性,因为对称性的特性是它会导致某种物理量不可观察.这种使粒子多余态不可观察的对称性叫规范对称性,它的存在保证了粒子只有两个可以被观察到的状态,因而是无质量的.反之如果粒子有质量,这种规范对称性就不能存在或者说必须被破坏.因此我们得到结论:对自旋大于 $1/2$ 的粒子,规范对称性的存在将保证其质量为零.反过来,对自旋大于 $1/2$ 的粒子,若要有质量,则相应的规范对称性必须被破缺.这样,自旋大于 $1/2$ 的粒子质量的从无到有的产生就和它所具有的规范对称性的从有到无,或者说规范对称性的破缺联系起来.研究自旋大于 $1/2$ 的粒子质量起源转化成了探索自旋大于 $1/2$ 的粒子所具有的规范对称性及其破缺的机制.

目前实验上业已发现的基本粒子包括夸克、轻子和规范粒子.其中除了自旋为 1 的光子和胶子是零质量外,所有其他的粒子包括自旋为 1 的电弱规范粒子 W^\pm 和 Z^0 及自旋为 $1/2$ 的夸克和轻子都是有质量的.这些质量参数构成了目前人类描写基本粒子及其相互作用的标准模型中所要输入的参数最大的部分.经过多年的研究,人们把这些有质量粒子的质量的产生机制等价为一组抽象的对称性——数学上的群论表述的 $SU(3)_C \times SU(2)_L \times U(1)_Y$ 对称性到其子对称性 $SU(3)_C \times U(1)_{em}$ 的破缺机制上.这里每个对称性群都有一个下角标用于指明它的出处:下角标 C 代表强作用的颜色群、L 代表电弱作用的左手群、Y 代表电弱作用的超荷群、em 代表电磁作用的群.被称之为强作用颜色对称性(同时也是规范对称性)的 $SU(3)_C$ 群是不破缺的,它保证

胶子没有质量,另一部分被称之为电弱对称性的 $SU(2)_L \times U(1)_Y$ 群要破缺到其子对称性——电磁对称性群 $U(1)_{em}$,剩下未破缺的电磁对称性 $U(1)_{em}$ 同样也是规范对称性,它保证光子没有质量.电弱对称性 $SU(2)_L \times U(1)_Y$ 是规范对称性,它若存在,则要求所有自旋为 1 的电弱规范粒子(包括光子、 W^\pm 和 Z^0)都没有质量;电弱对称性 $SU(2)_L \times U(1)_Y$ 同时还是手征对称性,它若存在,则要求所有自旋为 $1/2$ 的夸克和轻子都没有质量.因此研究基本粒子的质量是如何获得的实际上等价于研究电弱对称性 $SU(2)_L \times U(1)_Y$ 是通过什么机制破缺到其子电磁对称性 $U(1)_{em}$ 的.进一步讲,这种对称性的破缺如果只是通过在理论中加入粒子的质量项(这种破缺的实现被称为对称性的明显破缺)是不能保证理论的自洽性的,它将导致理论的么正性和可重整性的破坏,对称性的破缺只有是自发的破缺才有可能保证理论自洽性.关于对称性的自发破缺和质量产生的关系我们后面还会进一步讨论.

有人说把研究基本粒子的质量起源问题转化为研究电弱对称性自发破缺的产生机制问题是换汤不换药,把一个未解决的问题转换成了另一个未解决的问题.这种评论虽不能说错,但换个角度看问题,很可能会使我们更好地理解以致促进我们解决问题.特别是在基本物理规律的研究中,人们发现对称性经常起着非常根本和第一性原理的作用,把质量起源与对称性联系起来也就把质量的起源的研究上升到了与对称性同等根本的层次.

还有人说夸克和轻子的质量对研究质量起源并不重要,理由是:原子的质量主要来自原子核中的质子和中子,而质子和中子的质量主要来自组成它的夸克的结合能,夸克质量(指流质量)对质子和中子的贡献是很小的,而电子的质量对原子质量的贡献显然也是很小的.因此即使夸克和轻子都维持没有质量,质子和中子的质量以及原子的质量与实验上观测到的数值也不会相差太大.据此认为我们研究夸克和轻子的质量实际上只是关注了原子质量中一个很小的不见得重要的部分.这种观点所说的我们研究质量起源只关注了原子质量中一个很小的部分是不错的,那个剩下的大部分主要是夸克的结合能,还有其他的结合能,如原子核内质子与质子、中子与中子、质子与中子之间的结合能,原子核与电子的结合能等等,这些结合能需要本文前面提到的物质各层次对应的学科进行专门的研究.我们之所以把目标锁定那部分相对很小的夸克和轻子质量的贡献,

除了前面提到的认为这是质量起源的源头外,还因为夸克和轻子质量有着我们通常所没有预期到的更重要的作用,具体如下:

(1)如果夸克没有质量,质子将不再会像现在那样比中子轻,而反过来会比中子重.这是因为质子带电而中子不带电,额外的电磁作用产生的能量倾向使质子更重.在现实世界中,质子比中子轻是因为夸克有质量,质子由两个 u 夸克和一个 d 夸克组成,中子由两个 d 夸克和一个 u 夸克组成, u 夸克比 d 夸克轻导致质子由于比中子多一个 u 夸克少一个 d 夸克而更轻.这个夸克质量导致的效应超过了质子带电的效应,最后使得质子比中子轻.若夸克无质量,夸克质量的效应就消失了,质子带电的效应使得质子变成比中子重.这样自然界就将允许质子向中子衰变,而中子将成为最稳定的重子.从而自然界将会允许大量自由存在不带电的中子而不是现实中观察到的质子(也就是氢原子核).这样自然界最轻的元素氢的原子核就不再是带电的质子,而应是不带电的中子了,因此,通常的氢原子也就无法形成,而只能由其同位素如氘或氚来替代了.

(2)如果轻子没有质量,注意到氢原子的玻尔半径反比于电子的质量,作为轻子一员的电子没有质量将导致无穷大的玻尔半径,也就是说,氢原子同样无法形成,因此无论怎样都导致没有现在意义上的氢原子,因而也就没有现在绝大部分的分子、无机物、有机物……以致生命就也都没有了.这样看来,夸克和轻子的质量对于我们这个世界维持目前这个样子是至关重要的.

由于质量起源与对称性及其自发破缺的密切关系,人们既从对称性自发破缺的角度,也从完全不涉及对称性的其他角度分别对质量起源开展了研究,下面我们具体罗列和讨论三种人类业已发现的与质量起源有关的质量产生和分裂的机制:

(1)粒子与真空作用,导致对称性的自发破缺,产生质量

这是人类目前所认识到质量起源的最主要的机制.自然界里的真空并不是空空如也的虚空,而是有复杂的结构.其中当体系的相互作用具有对称性时,最低能态若发生简并,就会出现所谓的对称性自发破缺现象,这正是我们谈到的质量起源的对称性表述,即电弱对称性的自发破缺所需要的.本文作者在《物理与工程》杂志 2009 年第 1 期上的第一篇文章中,对对称性自发破缺给予了详细的介绍.在那里,我们说明了对连续对称性(电弱对称性就是一种连

续对称性),在最低能态上,从真空态出发,沿每一独立的破缺对称性方向的激发将对应产生一个零质量的被称为 Nambu-Goldstone 的粒子,而垂直破缺对称性方向的激发将产生有质量的粒子,这个有质量的粒子在电弱对称性的自发破缺的情形就是被称之为“上帝的粒子”的 Higgs 粒子.它来源于电弱对称性的自发破缺,能在实验上发现它将直接显示自发破缺的附带效应,因此受到人们的极大关注.但 Higgs 粒子的性质不仅仅依赖对称性,还强烈地依赖相互作用动力学的具体细节,不同的动力学理论会给出完全不同的性质,有些相互作用的动力学理论甚至可以允许没有 Higgs 粒子存在,这些大大加大了在实验上寻找 Higgs 粒子的难度.人们有时按照 Higgs 粒子在理论中表现为基本的粒子还是复合的粒子或是根本不出现来对描述基本粒子质量起源的理论模型进行分类.从某种意义上说,研究 Higgs 粒子及其可能的替代物也就是研究基本粒子质量的起源.

如果体系的相互作用具有的对称性只是近似的,相应出现的对称性的自发破缺也就是近似的,它体现为原来沿破缺对称性方向激发的粒子从零质量的 Goldstone 粒子变成了具有小质量的所谓赝 Goldstone 粒子,这提供了一种产生小质量粒子的方法.有一类描写电弱对称性自发破缺的理论把 Higgs 粒子看成是这种具有小质量的赝 Goldstone 粒子,因而这类理论被称为小 Higgs 理论.

如果自发破缺发生在像电弱对称性这样的规范对称性上,原本的规范对称性主要是使粒子中多余的态不可观察,当对称性破坏后这些不可观察的态就变成可以观察的了.可以证明这些变成的可观察的态的自由度的来源正是前面提到的 Goldstone 粒子,即,当规范对称性发生自发破缺时,破缺产生的 Goldstone 粒子与无质量的规范粒子耦合成为新的粒子,将本身具有的一个自由度转化为这种新粒子的纵向自由度,使规范粒子获得质量,这个使自旋为 1 的规范粒子获得质量的机制被称为 Higgs 机制. Higgs 机制把自旋为 1 的规范粒子获得质量的问题与规范对称性的自发破缺问题具体关联起来了.

在对称性自发破缺的系统中,自旋为 1/2 的粒子要获得质量,就需要存在它与真空的相互作用.这体现在无质量的裸粒子在真空中做穿行运动,并与其发生相互作用.这种相互作用在对称性存在时不起作用,但在对称性发生了破缺后就开始起作用,表现为它赋予粒子以有效的惯性,因而使粒子穿上了

有质量的外衣.但粒子与真空的相互作用形式可以是多种多样的,目前在基本粒子的标准模型中采用所谓的 Yukawa 型作用来等效地描述自旋为 1/2 的粒子与真空的相互作用.代价是相应每个粒子都需要有一个参数来描写其与真空的相互作用的强度,而且这种等效的作用的起源仍是不清楚的问题.

(2) 粒子与粒子作用,导致粒子之间的混合与质量分裂

除了与对称性及其破缺发生密切联系的粒子与真空的作用会导致质量的产生外,当一个粒子与其他粒子之间发生相互作用时,参与作用的粒子之间会发生混合,导致它们的质量发生改变,通常称之为质量的分裂.以一个最简单的两粒子体系为例:如果这两个粒子之间没有相互作用,体系的哈密顿量可以写成:

$$H = \begin{pmatrix} M_1 & 0 \\ 0 & M_2 \end{pmatrix},$$

其中非对角元为零,即反映两个粒子不发生相互作用.按照量子力学的规则,哈密顿量 H 的本征态是该体系在物理上可以被看到的粒子状态,而 H 的本征值即是体系粒子的质量.因此体系具有两个互相不发生混合的粒子状态(体现为两个本征态互相之间没有关系),分别对应质量 M_1 和 M_2 的两个粒子.为讨论方便起见,我们不妨假设 $M_1 \geq M_2$,如果这两个粒子态之间有相互作用,哈密顿量的非对角元就会不为零,而哈密顿量的厄米性要求出现在两个可能的非对角元的两个粒子态之间的耦合系数必须互为复共轭,因此我们可以把有相互作用的两粒子体系的哈密顿量写为

$$\begin{pmatrix} M_1 & me^{i\alpha} \\ me^{-i\alpha} & M_2 \end{pmatrix},$$

其中参数 m 和 α 分别反映两个态之间相互作用的强度和位相.这时体系的本征态应是 $m=0$ 时的两个本征态的某种叠加,或者说两个粒子态之间发生了混合.混合后产生的两个物理粒子的质量对应现在哈密顿量的两个本征值,它们很容易通过求解久期方程(在这种情况下是一元二次方程)得到,它们分别为:

$$M_{\text{Heavy}} = \frac{1}{2}(M_1 + M_2) + \sqrt{m^2 + \frac{1}{4}(M_1 - M_2)^2},$$

$$M_{\text{Light}} = \frac{1}{2}(M_1 + M_2) - \sqrt{m^2 + \frac{1}{4}(M_1 - M_2)^2}.$$

注意分裂后的质量与两粒子态之间相互作用的位相 α 是无关的.很容易看到, $M_{\text{Heavy}} \geq M_1$ 、 $M_{\text{Light}} \leq M_2$,并且等号只在 $m=0$ 时成立.因此我们得到结论:粒子之间的相互作用使较重的粒子变得更重,较轻的

粒子变得更轻,好像跷跷板一样,一个升高,另一个降低.这个由于粒子间的相互作用导致混合并使得粒子原有质量谱的进一步分裂的机制,被称之为跷跷板机制.从上面的公式可以进一步判断出, m 越大,质量的进一步分裂越大,跷跷板跷的越厉害.当粒子间的相互作用强度大到一个极限即 $m^2 = M_1 M_2$ 时,轻的粒子态变成无质量的,而重的粒子态的质量变为 $M_1 + M_2$.以上关于两粒子体系的讨论可以直接向多粒子体系推广,结果无非产生更复杂的粒子混合和质量分裂现象.总之,跷跷板机制可以使某个粒子通过与其他粒子的相互作用,使其质量升高或降低.在质量起源的研究中,一般不用它来解释粒子质量如何从零到非零(这通常要靠其他的机制),因为这将导致另一个具有负质量的粒子.通常用它来解释各种可能的粒子质量的分裂.例如:在强子物理中,如果要解释为什么中性的 π^0 介子的质量为 135MeV 比质量为 140MeV 带电的 π^\pm 介子要轻,人们可以说,因为 π^\pm 介子带电,所以要重一点;但电弱规范粒子中,带电的 W^\pm 质量为 80GeV ,却比中性的质量为 91GeV 的 Z^0 要轻.显然带电与否在这里无法同时解释这两种情形,但应用跷跷板机制考虑粒子间的混合就可以.在现实中 π^0 介子实际是 π^3 介子与 η 混合后的那个轻的粒子态,因而比 π^\pm 介子要更轻一些,而 π^3 介子和 π^\pm 介子同属一个同位旋三重态,其质量应该是近似相同的,因此 π^0 比 π^\pm 介子要轻.而 Z^0 实际是 W^3 与另一个所谓的超荷规范粒子 B 混合后的那个重的粒子态,因而比 W^3 要更重一些,而 W^3 和 W^\pm 同属一个弱同位旋三重态,其质量应该是近似相同的,因此 Z^0 比 W^\pm 要重.在基本粒子中,人们发现中微子具有不为零但十分微小的质量,利用跷跷板机制可以解释为什么中微子质量那么小,只要自然界存在另一个很重的中微子.前面曾提到中微子因为是中性的,因而存在两种右手型的中微子:一种是真正的右手中微子;另一种是左手中微子的电荷共轭态.这两种粒子态之间若有自作用和互作用,就可形成一个我们所讨论的需要用 2×2 矩阵描述其哈密顿量的两粒子体系,而实验上看到的中微子很轻就对应我们前面讨论过的这个体系接近极限的情形:两个粒子通过混合,一个变得很轻(接近于零因而被解释成实验上看到的中微子),另一个变得很重(接近于 $M_1 + M_2$),实验上目前还观察不到.

(3) 粒子与额外维空间作用产生质量

如果自然界中在三维空间之外还存在有额外维

的空间,它们到现在一直未被发现的一种解释是额外维空间被紧致化成很小的空间.这种紧致化也会产生粒子的质量.以一维额外维空间为例,现在的时空从4维变成了5维,多出的第5维被周期性地限制在一个半径为 R 的小环上.按照量子力学的观点,在无限大的无穷大空间中,粒子的动量可以连续地取值,但对限制在大小为 L 并满足周期性边界条件的区域中的粒子的动量就只能取分立值 $p = 2\pi n/L$ (n 为整数).现在额外维限制在线度为 $2\pi R$ 的小环上的粒子的动量在额外维的分量也相应地只能取分立值 $p_5 = 2\pi n/2\pi R = n/R$.如果这个粒子在5维时空是无质量的,它应该满足5维时空中的质壳条件: $E^2 = (p_1^2 + p_2^2 + p_3^2 + p_5^2)c^2$,其中 p_1, p_2, p_3 分别是动量在前三维空间的分量, p_5 是动量在第5维的分量.带入动量的第5维分量只能取的分立值 $p_5 = n/R$,这个质壳条件变为: $E^2 = (p_1^2 + p_2^2 + p_3^2)c^2 + c^2 n^2/R^2$.若把这个结果看成是4维时空中粒子的质壳条件,我们发现第5维动量在4维时空质壳条件中好像起了一个质量项的作用,因为标准的4维时空质壳条件是 $E^2 = (p_1^2 + p_2^2 + p_3^2)c^2 + c^4 m^2$,对比之后我们得到 $c^2 m^2 = n^2/R^2$.也就是说,紧致化第5维空间导致在4维时空出现了一系列的粒子,通常被称之为Kaluza-Klein粒子,它们分别对应 n 取不同的整数值,形成所谓一层层的Kaluza-Klein塔.这里Kaluza-Klein粒子中最轻的仍然是无质量的,但所有的其他的激发粒子态都有质量,并且态与态之间的质量差都是一样的,都为 $1/cR$.从这个角度看,检验质量的产生是否来自额外维,很重要的判据是要看有没有发现等间距的激发态粒子谱.现在已被实验发现的基本粒子并不符合这一等间距的要求,因此这一机制无法用来解释目前基本粒子的质量谱分布.但未来将要被发现的新的基本粒子与现有的基本粒子之间还是有形成等间距的质量差的谱的可能性的,如果确实发现了这种现象,这些形成了等间距质量差的基本粒子的质量就很可能来自额外维的空间.

基本粒子的质量起源问题到目前为止仍是尚未解决的问题.虽然本文以上介绍了很多思想和看法,它们远没有达到能够解决质量起源问题的地步,只是一些人们对此问题的进一步的认识.这些认识能否更加深入,以致于更加趋近于真正地理解质量的起源,这还有赖于实验上的进展.物理学本质上是实验的学科,人类对自然规律的正确认识和诠释必须经得起实验的检验,否则对各种各样理论上可能

很优美并且自洽的畅想就无从判断其正确性了。物理实验特别是与基本粒子相关的高能物理实验的发展强烈地受到技术和工程的发展水平的限制(当然也推动它们的发展和进步)。很多实验过去无法做,现在能做了,还有很多实验现在还无法做,要等将来技术的进一步进步才有可能做。这样就提出了我们前面阐述的关于质量起源的各种认识能否被目前实验所检验的问题,如果无法检验,这些讨论只能是学院式的抽象理论设想;如果能够被检验,我们就有可能在实验检验结果的指导和启示下,不断地

发展和深化我们的认识。所幸的是,目前在欧洲核子中心 CERN 新开始运行的大型强子对撞机 LHC 上的高能物理实验能力已经达到了能够部分地检验人类提出的关于基本粒子质量起源的很多设想的地步,质量起源问题已经被设定为在 LHC 上要进行探索的头号科学目标。因此随着 LHC 的运行,人类对质量起源的了解将要进入到一个全新的时代,我们期待着这方面新的革命性的突破和发现,使得人类对质量起源的探索能够更加深化。

• 中国物理学会通讯 •

2008—2009 年度首届黄昆物理奖获奖者获奖成果介绍

2008—2009 年度首届黄昆物理奖获奖者是清华大学物理系姜开利研究员,他主要从事碳纳米管的生长机理、可控生长、物理性质及应用等方面的研究,取得了以下研究成果:(1)从理论和实验两个方面开展了碳纳米管生长机理的研究,提出了基于 VLS 机制的碳纳米管生长模型和成核机理;(2)在机理研究基础上,针对不同的应用需求,发展了超顺排碳纳米管阵列合成、单双壁碳纳米管阵列合成,以及可控终结、激光直写等可控生长技术;(3)在获得高质量、高性能碳纳米管材料的基础上,研究其物理性质,发现了碳纳米管场发射过程中的电子冷却效应,阐明了碳纳米管场发射体的失效机制,测量了碳纳米管线和单壁、双壁、多壁碳纳米管侧壁的逸出功;(4)在物性研究基础上,探索碳纳米管的宏观应用,研制成功基于碳纳米管的 TEM 栅网、触摸屏、同轴电缆、扬声器等,在国际上产生了较大的反响。

一句话:

姜开利研究员领导的课题组致力于碳纳米管生长机理、物性及其应用研究,发现了碳纳米管的热声效应,发明了碳纳米管扬声器。

• 封面故事 •

封面是庞大星系团 MACS J0717.5+3745(简称 MACS J0717)的合成图像。图像中弥漫的热气体是钱德拉空间 X 射线天文台观测结果的伪彩色显示,而星系的光学图像由哈勃空间望远镜获得。热气体的不同色彩表示了不同的温度,最冷的气体是红紫色,最热的气体为蓝色,处于中间温度的气体是紫色的。星系团 MACS J0717 距离地球大约 54 亿光年。

在这个遥远的庞大星系团中,一个长约 1300 万光年的由星系、气体和暗物质组成的丝带状物质流,正在不断地涌入早已充满星系的区域。好比一条高速公路上的汽车不断驶入停满车辆的停车场。物质流导致了星系团内多次、反复、连环的碰撞,有 4 个独立的星系团参与了这一无比壮观激烈的碰撞过程。

除了惊人的连环碰撞之外,MACS J0717 的温度也异常引人注目。由于星系团中的每次碰撞都会以热量的形式释放大量的能量,MACS J0717 是此类系统中迄今所见温度最高的一个。尽管通向 MACS J0717 的丝带状物质流早已被发现,但近期的观测和研究结果首次说明,它就是星系连环碰撞的源泉。星系团的运动方向,与丝带状物质流的方位基本上是一致的;星系团中温度最高的区域是丝带状物质流和星系团的相交处,表明碰撞一直在进行。

星系团是宇宙中受引力束缚的最大天体。MACS J0717 是人们见过的最复杂、最混乱的星系团,此前人们见过的其他庞大星系团,参与碰撞的仅有两个独立星系团并且显示比较简单的几何形状。MACS J0717 展现了庞大星系团在几百万光年的尺度上与周边环境相互作用的方式,让人们看到了丝带状物质流流入星系团导致其增长的绝妙情景。MACS J0717 所显示的复杂过程,给科学家提供了研究宇宙中最大的物体在广袤的宇宙空间是如何发生相互作用的、以及宇宙结构是如何生长与演化的一个新的机会。

(中国科学院国家天文台 温学诗)