

目录

第十六章 极简宇宙学	2
16.1 从热大爆炸到宇宙学原理	3
16.1.1 热大爆炸的概念	3
16.1.2 早期宇宙简史	6
16.1.3 原初核合成	8
16.1.4 宇宙微波背景(CMB)和宇宙学原理	10
16.2 宇宙膨胀动力学	13
16.2.1 FLRW度规	13
16.2.2 联络和曲率	17
16.2.3 弗里德曼方程	19
16.2.4 宇宙学红移	21
16.2.5 宇宙学参数及其确定	24
16.2.6 宇宙学视界	30
16.3 暗物质与暗能量	33
16.4 宇宙暴胀	37
16.4.1 暴胀的基本概念	37
16.4.2 暴胀场	39
16.5 参考文献	41

第十六章 极简宇宙学

陈童

公元2世纪，托勒密创立地心说，告诉世人宇宙是一个嵌套27层的同心水晶球，地球是这个水晶球的中心。

一千多年后的1543年，哥白尼的《天体运行论》出版，它告诉世人，地球只是围绕太阳运动的行星之一。伽利略竖起天文望远镜证实了哥白尼的学说，他冲着望远镜笑道，这才是观星利器。

又过了200多年，赫歇尔架起更大号的望远镜，确定了圆盘状的银河系，并记录下了5000个星云。

时间来到20世纪20年代，哈勃叼着烟斗，瞄着胡克望远镜，测定了10万光年的银河系，看到了90万光年远的河外星系，他发现了星系间彼此远离的现象，宇宙在膨胀！

现在，请想象一下，你站在一片宁静的海滩上，脚下的沙粒细软无比。每一粒沙子都代表着宇宙中的一个星系。当你俯身拾起一粒沙子，这便是我们的银河系；而在你的指尖之间，还有无数其他的“沙粒”，它们是散布于浩瀚空间中的其他星系。这就是我们所处宇宙的一个缩影——广阔、神秘且充满了未知。

随着广义相对论的提出，以及随后哈勃发现宇宙膨胀的现象，人类对于宇宙的认识发生了革命性的变化。从大爆炸理论到暗物质与暗能量的研究，每一步进展都在帮助我们绘制出一幅更加精确的宇宙图景。

16.1 从热大爆炸到宇宙学原理

16.1.1 热大爆炸的概念

大爆炸理论的基本观察是遥远星系的红移。它们的光谱向较长波长方向偏移，离我们越远，偏移越大。这表明它们正在远离我们；它们与我们的距离在增加。根据广义相对论，这种现象不是因为星系本身在移动，而是空间本身在膨胀。随着空间的膨胀，通过空间传播的光线的波长也“膨胀”了。

在把星系看作“沙子”的大尺度上，宇宙的这种膨胀似乎是均匀的，整个宇宙以相同的速率膨胀，可以通过一个随时间变化的尺度因子 $a(t)$ 来描述， $a(t)$ 是宇宙尺寸的膨胀倍率，在宇宙的当前取1，宇宙的膨胀速率由哈勃参数 $H \equiv \dot{a}/a$ 表征。通过测量当前的膨胀速率并结合对宇宙能量内含的知识，然后应用广义相对论方程，可以推算出 $a(t)$ 作为时间 t 的函数。

大爆炸理论是伽莫夫在1948年提出的，它的基础就是刚才说的宇宙膨胀模型，这是弗里德曼等人建立的。从今天的观点来看，大爆炸理论是膨胀宇宙模型的自然延伸，但在伽莫夫提出大爆炸理论的当时，弗里德曼模型预言的宇宙年龄明显与事实不符，在这样的基础上，伽莫夫的理论就完全不被学术界所接受，这种局面持续了近20年，直到这理论预言的宇宙微波背景被发现和证实。下面我们定性地谈一谈大爆炸理论背后的物理思想。

远古的宇宙中不可能有星系

在今天的宇宙中，物质是结团的，但是从这个现状出发，根据膨胀宇宙模型往前追溯，就不可避免地得到一个非常重大的推论：星系等物质集团不是亘古恒存的，而是在宇宙演化过程中产生的！

今天，星系内的平均物质密度比全宇宙的平均密度大5个数量级，也即是说，物质是局域地结团的。根据膨胀宇宙模型往前追溯到宇宙尺度因子 $a(t) < 10^{-2}$ 时，宇宙体积是今天的 $a^3 < 10^{-6}$ 倍，从而宇宙的平均密度将比今天大6个数量级以上，比今天的星系密度还大。这表明，那时候是不能有星系存在的，今天如星系这样的物质团块只能是宇宙演化的产物。

这结论深刻地指出早期宇宙并不是今天宇宙的缩影，在其演化过程中，宇宙的面貌发生过质的变化。那么，结团前的宇宙物质以什么状态存在？

星系是均匀宇宙气体碎裂的结果

一个自然的猜测是，一切星系结构形成之前的宇宙是有微小密度差别的均匀气体，宇宙的自引力不稳定性，也就是密度偏大的那些地方对周围的引力也会更大，因此会不断吸引周围的气体，使得密度进一步增大，这就会导致原来的均匀宇宙气体碎裂成团，进而演化出今天看到的一切物质集团结构。这是伽莫夫理论的一个要点。当然，在伽莫夫提出大爆炸理论的当时，这只能是一个猜测。

膨胀的宇宙来自大爆炸

如果结构形成之前的宇宙内含物是密度和温度都均匀的气体，那随着宇宙的膨胀，不仅是它的密度会逐步降低，而且其温度也会逐步降低，相应的推论就是，越早的时候宇宙的温度越高，如果追溯到非常早的时期，宇宙内含物的密度和温度将比今天高出几十个数量级。温度的大幅度变化会导致物质状态的相变，因此，早期宇宙的演化会因温度的巨大差别而有实质性的变化。

如果往前推到极端，宇宙的膨胀应该是从密度和温度都为无穷的状态开始的，如果宇宙是有限的，它起初的体积又很小，而温度又非常高，那么宇宙的行为就很像一次爆炸，因此在历史上，伽莫夫理论的反对者讥讽地称之为“宇宙大爆炸”，这就是这个名称的来源。

关于大爆炸理论，一个最常见的误解就是认为宇宙最早开始于一个点。实际上，我们今天的可观测宇宙在刚开始的时候的确很小，在大爆炸理论框架能处理的前提下，它很可能直径小于1毫米。但是，相比于早期极端短的时间，1毫米其实并不是一个很小的尺度，至少远非一个点。另一方面，在可观测宇宙之外，我们的宇宙很可能还在延伸，甚至无限延伸，而如果今天的宇宙是无限的，那它很可能一开始就是无限的，那就更不是从一个点开始的了。当然，今天的理论还不能处理宇宙是如何开始的，因为最早的那个时刻需要量子引力才能描述，而目前还不存在完备的量子引力理论。

原子和分子是宇宙演化中产生的

宇宙大爆炸理论的另一个深刻结论是，原子分子也不是天然存在的，它们也是宇宙演化到一定阶段的产物。这里最奇妙的就是，这么小的原子和这么大的星系，两者竟然同样都是宇宙演化过程的产物，而不是永恒的天然存在。

为了理解这一点，需要追溯到星系形成前宇宙气体的构成，它是由中性原子构成的，其中氢原子的丰度约为75%，是这气体中的主要成分。氢原子是由氢核与电子所构成的束缚系统，它的结合能为13.6eV。因此这些中性原子气体的温度不能太高，当温度 T 高于 10^4K 时，粒子的平均热动能是 $k_B T \approx 1\text{eV}$ ，这时候能量超过13.6eV的光子大量存在，它们与氢原子的热碰撞将使原子电离。因此在这个温度以及更高的温度下，宇宙气体必定处于电离状态，即是等离子气体，其组分为原子核、自由电子、以及光子。原子和分子是随着宇宙膨胀降温的结果，之前是不存在的。

伽莫夫发现，在宇宙膨胀降温使等离子气体变成中性原子之后，光子与电子就几乎不散射了，因此这时候的光子将会自由地在宇宙中飞行，成为宇宙的背景，这种背景光子应该在今天依然存在而且可以被观测。这是伽莫夫理论的一个关键性预言。这个预言在1965年被观测证实，这是宇宙大爆炸理论从被否定到被重视和接受的转机。

化学元素也是演化的产物

在等离子状态的宇宙中，原子核是气体的主要成分，而它是由若干核子(也就是质子和中子)组成的，每一个核子在原子核中的平均结合能为1MeV的量级，所以，当气体的温度高于1MeV(即 10^{10}K)时，组分粒子间的热碰撞会使得原子核解离，原子核在这样的高温下将不复存在。没有原子核就没有化学元素，因此，就连化学元素也是在宇宙膨胀降温到一定程度以后的产物。

原始汤

在宇宙温度高于1MeV的阶段，物质成分是比较原子核更基本的粒子，即质子、中子、电子以及光子等。当然，按照今天的认识，质子和中子也并不是物质最基本的组元，它们是由夸克和胶子所组成的复合粒子。按照粒子物理的标准模型，如果密度超过核子，而且温度也足够高，那么核子作为复合粒子也不能存在，这时候的宇宙内含物将是由夸克、轻子以及规范玻色子所组成的粒子气体。

因此大爆炸不是指一个瞬间发生的导致宇宙开始的爆炸事件，而是描述宇宙从一个高温、高密度状态开始迅速膨胀的过程。早期宇宙充满了几乎均匀的基本粒子“原始汤”，这些粒子处于长时间的热平衡状态，因此我们可以用少量的热力学变量来描述早期宇宙的状态，这使得宇宙的时间演化可以计算出来。

事实上,应用描述高温下宇宙能量内含的粒子物理标准模型,可以得到 $a(t)$ 在约140亿年前接近于零。这标志着宇宙开始于140亿年前的这个时刻,即大爆炸开始的时刻。在这个时间点,宇宙的密度趋向于无穷大,称作大爆炸奇点,我们定义这个时刻为 $t = 0$ 时刻。但是实际上,当密度达到普朗克密度 $\rho_{\text{Pl}} \approx 10^{96} \text{kg}/\text{m}^3$ 时,广义相对论本身就失效了,需要量子引力理论(目前还不存在)来描述这一非常短的时期。因此,严格来说,包括 $t = 0$ 在内的这些最早的时间其实必须从大爆炸理论中排除。

16.1.2 早期宇宙简史

下面我们概要性地简述一下早期宇宙的简史。因为热大爆炸,早期宇宙的温度很高,从而粒子就有很高的能量。因此为了描述这个时期的物质,我们需要粒子物理的知识。粒子物理的标准模型具有 $SU(3)_c \otimes SU(2)_w \otimes U(1)_y$ 的规范对称性。在我们今天的低能世界中,标准模型的某些对称性是自发破缺了的。当宇宙的温度超过100GeV(约 10^{15} 开尔文)时,就达到了标准模型的天然能标,这大约就是宇宙的年龄小于 10^{-11} s时的情况。这时候粒子的原始汤由自由的无质量费米子(夸克和轻子)以及传递相互作用的无质量规范玻色子构成。当然,标准模型中还有一种粒子,即希格斯玻色子。

正是希格斯玻色子使得电弱对称性($SU(2)_w \otimes U(1)_y$)发生了自发破缺,这也正是早期宇宙中发生的一种相变。电弱相变之后,电弱相互作用就变成了两种相互独立的相互作用:其一是由有质量的规范玻色子 $W^\pm$ 和 Z^0 传递的弱相互作用,其二就是由无质量 γ 光子传递的电磁相互作用。通过电弱相变,费米子也获得了质量,这些质量来源于它们与希格斯玻色子的相互作用。电弱相变是当宇宙冷却到临界温度 $T_c \sim 160 \text{GeV}$ 时发生的,大约发生在 $t \sim 10^{-11}$ s时。

早期宇宙的另一种相变是QCD(量子色动力学)相变,也就是夸克-强子相变,大约发生在 $t \sim 10^{-5}$ s,其临界温度约为 $T_c \sim 150 \text{MeV}$ 。相变之前,夸克是自由的,相变之后它们就形成了强子,包括重子(例如核子中的中子 n 和质子 p)和介子(例如 π 介子和 K 介子)。QCD相变之后,宇宙的物质内容就从夸克胶子等离子体转变成了强子气。

根据量子场论,每一种粒子都有相应的反粒子,它的质量与自旋等等性质都和粒子相同,除了荷相反,这里的荷包括电荷和色荷等。如果某种粒子不带有荷,那它就是自己的反粒子,典型的比方光子。当温

度 $T \gg m$ 时, 这里 m 是粒子的质量, 这时粒子和反粒子不断地相互湮灭, 同时也不断地产生, 宇宙中粒子和反粒子的数目大致是一样的。但是, 当温度 $T \ll m$, 粒子和反粒子依然相互湮灭, 但是热能已经不足以产生粒子-反粒子对了。因此随着宇宙的不断冷却(由于宇宙膨胀), 重的粒子与反粒子就趋向于相互湮灭, 同时额外产生出更多更轻的粒子-反粒子。如果宇宙中粒子和反粒子的数目严格相当, 那么冷却到今天将只有光子和中微子¹留下来。但是今天的宇宙中显然还存有重子物质(主要由质子 p 和中子 n 构成, 比如我们的银河系), 这就表明, 在早期宇宙中, 核子和电子必然比它们相应的反粒子要稍微多一些, 这多出来的一些就存留到了今天。

如果再追问早期宇宙中重子与反重子这种不对称的来源, 则最自然的回答可能是: 宇宙开始时重子和反重子是一样多的, 只是随之出现了破坏重子数守恒的物理过程, 导致重子数略微超过了反重子数。当然, 在今天的粒子物理标准模型中, 重子数始终是守恒的, 但是1970年代人们也提出了电、弱、强大统一理论(GUT), 其中重子数并不守恒。研究表明, 重子数不守恒加上宇宙极早期膨胀中对热平衡的短暂偏离就可以从原本正反粒子对称的宇宙中产生出过剩的质子和中子。当然, 大统一理论目前还只是理论探讨, 远没有得到实验的证实。

已知的, 有质量但是质量最轻的, 同时又参与强或者电弱相互作用的粒子, 是电子, 因此宇宙中最后的正反物质湮灭事件是电子-正电子湮灭, 这大约发生在 $t \sim 1\text{s}$ 时, 这时的温度约为 $T \sim m_e \sim 0.5\text{MeV} \sim 10^{10}\text{K}$ 。在这之前宇宙中仍然存在大量的正反电子对, 然而由于温度和密度相比更早的时候已经明显降低, 中微子以及反中微子与其它粒子的相互作用率远小于宇宙膨胀率, 它们的平均自由飞行时间显著变长, 近似成为与其它粒子没有相互作用的自由粒子, 从而不再与其它粒子共处于热平衡中, 这称之为中微子退耦, 这些退耦的中微子一直存留到今天。常常把中微子退耦的时刻记作 $t_{\nu d}$, 相应的温度记作 $T_{\nu d}$ 。因此, 在最后的电子-正电子湮灭事件之后, 宇宙中的反物质就只有反中微子了, 因此之后的宇宙原始汤将由大量的光子、中微子、反中微子, 以及少量剩余下来的质子、中子以及电子组成。

值得说明的是, 除了粒子物理标准模型中的粒子以外, 宇宙中很可能还含有暗物质粒子, 它们的性质目前人们还不太清楚。我们暂且先忽略暗物质粒子。

¹对中微子振荡的观测表明, 中微子也有很小的质量, 但是, 中微子的相互作用是如此微弱, 以至于即使在低于中微子质量的温度下, 中微子与反中微子也很少相互湮灭。

宇宙诞生之后几分钟，约在温度 $T \sim 100\text{keV}$ 时，质子和中子就结合成了轻元素的原子核。这也就是所谓的大爆炸核合成(Big Bang Nucleosynthesis, BBN)，以占通常物质的质量百分比衡量，它大约产生了75% ^1H ，25% ^4He ， 10^{-42}H ， 10^{-43}He ，以及 10^{-97}Li (其它化学元素是后来主要在恒星中形成的)。这时候这些物质是完全离子化的，所有的电子都还是自由电子。在这种等离子体中，光子不断地被电子所散射，相邻两次散射之间的平均自由程是短的。这就意味着，这时候的宇宙对于光来说是不透明的。

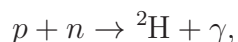
直到诞生之后的37万年，宇宙才变得透明。这时候的温度大约为 $T \sim 3000\text{K}$ ($\sim 0.25\text{eV}$)，这时电子和原子核结合形成了电中性的原子，从而使得光子的平均自由程变得比可观测宇宙的半径更大，宇宙就透明了。这个事件就称作宇宙复合过程(recombination)。复合之后，原初光子就可以自由地在空间飞行而几乎不被散射了，因此也就从与其它粒子的热平衡中退耦出来了。今天我们可以观察到这些原初光子，也就是所谓的宇宙微波背景(cosmic microwave background, CMB)，它们是从早期宇宙来的光，通过对它的观测，我们可以“看见”大爆炸。

复合过程之后，宇宙充满氢原子和氦原子气，同时还有微量的锂。大约在宇宙诞生之后的几亿年，从这个气体中形成了第一代恒星，但是大多数气体形成了星际气体。大约在宇宙诞生之后的7亿年，恒星发出的辐射重新离子化了这些星际气体。

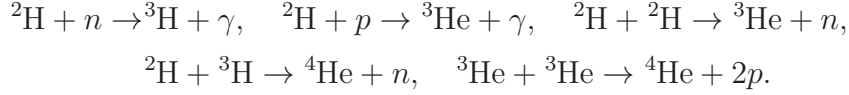
16.1.3 原初核合成

观测表明当前宇宙中氦约占宇宙总质量的25%，正如前面提到过的，这主要是由早期宇宙的原初核合成(大爆炸核合成)所决定的，恒星内部的核反应虽然也在生成氦，但其值只占上述丰度的很小一部分。原初核合成涉及的宇宙温度大约在 10^{10}K 到稍低于 10^9K 之间。温度高于 10^{10}K 时，即使质子中子结合成氦核也会被高能光子所击碎，这也就是所谓的光分裂。而在 10^{10}K 到稍低于 10^9K 这个温度区间所涉及的物理知识早已在地球实验室内得到了确证，因此人们对原初核合成的理论是很有信心的。

由于核合成时核子数密度不是很高，同时宇宙的快速膨胀又导致合成反应时间很短(核合成大约发生在 $t \sim 100\text{s}$ 时)，因此在原初核合成中只有快速的两粒子反应可以发生。首先是质子和中子结合成氘核 ^2H ，多余的能量和动量由一个光子带走，反应式就是



继而是一系列生成 ^3H 、 ^3He 、 ^4He 的后续反应，反应式分别为



由于不存在质量数为5的稳定核素，反应链至此中断。 ^4He 作为主要产物逐渐累积，直到数量足够时核反应才得以继续，结果就是在 ^4He 的基础上生成了很少一点 ^7Li 。由于不存在质量数为8的稳定核素，反应链就终结了！至于今天宇宙中各种比 ^7Li 更重的元素，则是后来在恒星内部的核反应以及超星系爆发中形成的，具体过程上一章已经讲过了。

原初核合成的整个反应链必须经过第一步质子和中子结合为氘核，这就是所谓的“氘关卡”。氘核的结合能比氦核小很多，氘核在温度降低到 $3 \times 10^9\text{K}$ 时就可以保持稳定，但氘核在这一温度下刚形成便会碎裂，因此真正有意义的核合成过程只在温度略低于 10^9K 时才能由于通过了“氘关卡”而开始发生。最后的产物是大量的 ^4He 和微量的 ^2H 、 ^3He 以及 ^7Li (^3H 不稳定，很快会转变成 ^3He)。若以 ^4He 的产量为单位，则 ^2H 、 ^3He 的产量约为 10^{-5} ，而 ^7Li 约为 10^{-10} 。

氦的丰度约为25%在理论上是如何确定的呢？为此假设以 N_n 和 N_p 分别代表宇宙中的总中子数与总质子数，以 n_n 和 n_p 分别代表原初核合成的时候中子和质子的数密度，记这两个数密度的比值为 σ ，很显然

$$\sigma = \frac{n_n}{n_p} = \frac{N_n}{N_p}. \quad (16.1)$$

由于核合成之后宇宙的主要成分就是氦元素与氢元素(含一个质子，不含中子)，所以很显然 $\sigma < 1$ ，具体是多少我们稍后再来确定。则，宇宙中的总核子数 $N = N_n + N_p = (\sigma + 1)N_p$ ，其中全部中子都与同数量的质子结合成氦元素，所以氦元素中所含的核子数 $N_{\text{He}} = 2N_n = 2\sigma N_p$ ，因此，以质量百分比计，原初核合成所产生的氦丰度为

$$Y = \frac{N_{\text{He}}}{N} = \frac{2\sigma N_p}{(\sigma + 1)N_p} = \frac{2\sigma}{\sigma + 1} = 2\left(\frac{n_n}{n_p}\right)\left(1 + \frac{n_n}{n_p}\right)^{-1}. \quad (16.2)$$

所以，宇宙中氦的丰度完全取决于核合成时 n_n 与 n_p 的比值。

这个比值可以这样来确定。在中微子退耦前(注意 $t_{\nu d} \approx 1\text{s}$ ，远在核合成的时间之前)，中子和质子可以通过以下弱作用过程相互转化：

$$p + e \leftrightarrow n + \nu_e, \quad p + \bar{\nu}_e \leftrightarrow n + e^+, \quad (16.3)$$

式中 e, e^+ 分别代表电子和正电子, ν_e 和 $\bar{\nu}_e$ 分别代表相应的电子中微子及其反中微子。因为中子质量略大于质子质量($m_n \simeq m_p + 2.5m_e$), 所以质子变中子的过程相比于其逆过程较难发生。比如, 由于 $m_p + m_e \simeq m_n - 1.5m_e < m_n$, 所以一个静止的质子和一个静止的电子不能变成一个哪怕是静止的中子(能量守恒不允许), 但是反过来却没有这个问题。当然, 在温度高达 10^{10}K 以上时, 电子远远不是静止的, 所以 $p + e \rightarrow n + \nu_e$ 可以发生, 但无论如何其发生的概率总小于逆过程的发生概率, 根据玻尔兹曼统计可知, 当正反过程达到统计平衡时, 必定有

$$\sigma = \frac{n_n}{n_p} = e^{-\frac{\Delta m}{k_B T}}, \quad (16.4)$$

式中 $\Delta m \equiv m_n - m_p$. 当宇宙温度降至 $T_{\nu d} \simeq 10^{10}\text{K}$ 时, 中微子退耦, 此后上述 n, p 相互转变的弱作用过程就基本停止了, 因此 $\frac{n_n}{n_p}$ 的值便大致冻结在了 $e^{-\frac{\Delta m}{k_B T_{\nu d}}}$ 。当然, 以上讨论略去了自由中子的自发衰变 $n \rightarrow p + e + \bar{\nu}_e$, 因为其半衰期约为10分钟, 远大于中微子退耦时的宇宙年龄 $t_{\nu d} \sim 1\text{s}$. 不过, 等到氢合成时宇宙年龄($t \simeq 100\text{s}$)相比中子半衰期来说已经可观了。因此, 等到氢合成时, $\sigma = \frac{n_n}{n_p}$ 的值将比之前的冻结值 $e^{-\frac{\Delta m}{k_B T_{\nu d}}}$ 略低, 大约为1/7弱一点。

将 $\sigma = \frac{n_n}{n_p} = 1/7$ 代入(16.2)式, 就得到氢丰度约为 $Y \simeq 0.25$, 与观测相当吻合。不过, 考虑到恒星内部的核反应也会产生 ^4He , 虽然比原初核合成的产量少很多, 但也会使得原初核合成时的氢丰度略低于今天观测到的氢丰度, 考虑这个因素以后由今天观测的氢丰度推断出来的原初氢丰度约为0.23, 与上述理论值0.25也基本吻合。

原初核合成的其它产物 ^2H 、 ^3He 以及 ^7Li , 它们虽然丰度很小, 但对于验证理论也很重要, 定量计算原初核合成产物的丰度还需要用到一个重要的物理参数 η , 它就是宇宙中重子与光子数密度之比, 即 $\eta = n_b/n_\gamma$, η^{-1} 就代表每个重子周围的光子数, 它通过影响光分裂的难易程度而影响核合成开始的时刻, 进而影响产物的丰度。 ^4He 的丰度只微弱地依赖于 η , 然而 ^2H 、 ^3He 以及 ^7Li 的丰度却对 η 值相当敏感。计算表明, 只要假定 η 值在 $3.4 \times 10^{-10} \sim 5 \times 10^{-10}$ 的范围之内, 则上述四种产物的理论丰度都能与观测丰度符合。

16.1.4 宇宙微波背景(CMB)和宇宙学原理

伽莫夫的热大爆炸理论预言了有弥漫整个宇宙的微波背景存在。

1960年代美国普林斯顿大学的Dicke和Peebles 认识到背景辐射问题对宇宙学的重要性，因此就开始组织团队准备观测仪器，以探寻这个辐射。然而，当他们还在做准备时，贝尔实验室的彭齐亚斯和威尔逊意外地发现了宇宙微波背景辐射。1965年这两组人同时在一期《天体物理杂志》上分别发表论文，彭齐亚斯和威尔逊的论文很简短，只客观地指出，在4080MHz发现有一个3.5K的过剩天线温度，在测量限度内它是各向同性的，并与季节的变化无关。Dicke和Peebles等人的论文则详细解释了这信号的宇宙学意义，也就是指出它很可能是宇宙微波背景。1978年彭齐亚斯和威尔逊因为他们的发现获得了诺贝尔物理学奖。

各向同性的背景辐射的发现，其首要意义是确证了星系形成之前的早期宇宙的确充满了均匀并且各向同性的热气体。当然，其中均匀性并不直接来自于对微波背景的观测，而是来自于我们对所谓哥白尼原理的信念，即地球并不是宇宙中心，同样的，宇宙中任何一处都不是宇宙中心，而是处处平权的。结合这个均匀性以及观测到的微波背景相对于地球的各向同性，就能确立早期宇宙是处处均匀并且处处都各向同性的。由此而来的结果就是，星系形成以后的宇宙在宇观大尺度上也是处处均匀且各向同性的。这也就是所谓的**宇宙学原理**。

根据1989年11月升空的宇宙背景探测者（COBE, Cosmic Background Explorer）测量到的结果，宇宙微波背景辐射谱非常精确地符合温度为 $2.726 \pm 0.010\text{K}$ 的黑体辐射谱，证实了银河系相对于背景辐射有一个相对的运动速度，并且还验证，扣除掉这个速度对测量结果带来的影响，以及银河系内物质辐射的干扰，宇宙背景辐射具有高度各向同性，温度涨落的幅度只有大约百万分之五。目前公认的理论认为，这个温度涨落起源于宇宙在形成初期极小尺度上的量子涨落，它随着宇宙的暴胀而放大到宇宙学的尺度上，并且正是由于温度的涨落造成宇宙物质分布的不均匀性，最终得以形成诸如星系团等的一类大尺度结构。

其中对银河系相对微波背景运动速度的确立来自于观测到的微波背景辐射温度在偶极上的微小各向异性。具体来说，星系作为膨胀宇宙中的一粒“沙子”，除了跟着宇宙膨胀以外，必定还有其自身的运动，通常称之为本动。银河系当然也不例外，因此我们是在运动的参考系中观测背景辐射的，这样就会有多普勒效应。不妨设想银河系在向右运动，设出发时频率为 ν_1 的光子在没有本动的星系上观测到的频率是 ν_0 （因为宇宙膨胀带来的宇宙红移， ν_0 当然小于 ν_1 ），则由于银河系在向右运动，右边来的光子的宇宙红移会被银河系的本动抵消一些，从而它被接收到时的频率将比 ν_0 略大，

相反，左边来的光子的红移将被本动加剧一些，从而测到的频率将比 ν_0 略小，这就等效于改变了不同方向上宇宙微波背景的温度，使之产生一个微小的各向异性，可以证明，这种由于银河系本动所产生的各向异性是一种偶极各向异性。

除了偶极各向异性以外，微波背景的温度还有微小的多极各向异性，这种多极各向异性常常用所谓的角功率谱来刻画(关于角功率谱的简短介绍，见参考文献)。正如前面所说，它起源于宇宙在形成初期极小尺度上的量子涨落，它随着宇宙的暴胀而放大到宇宙学的尺度上，最终导致星系等结构的形成。

最后讨论一点引申出来的物理观念。一百多年前，人们曾经认为有一种弥漫在全空间的介质，称作以太，它可以作为静止或运动的绝对基准。天体相对于以太的运动就称之为绝对运动，迈克尔逊和莫雷曾试图测出地球的绝对速度，然而始终得到零结果。从那以后人们就抛弃了以太的概念，并认为绝对运动是没有意义的，这也就是相对性原理的基础。然而，现在宇宙学肯定了弥漫全宇宙的背景辐射的存在，那么以背景辐射为基准的“绝对”运动就有了意义，比如COBE就测出了银河系相对于背景辐射的“绝对”速度。

实际上，在宇宙任何地点的任何一个像银河系这样的局域参考系中都能测到宇宙微波背景。如果发现从某个局域参考系中看来没有温度的偶极各向异性，那就可以认为这个局域参考系是“绝对静止”的，或者说本动速度为零。而上面所谈的银河系的“绝对速度”(本动速度)就是相对于这样的静止局域参考系而言的。当然，值得注意的是，由于宇宙本身在膨胀，所以宇宙中两个不同地点的静止局域系之间的距离是在膨胀增加的，或者说，空间不同点上的局域静止系之间是相对运动的。这就好比吹气球，气球表面任意一点都是局域静止的，但是随着气球的膨胀，两点之间的距离却在增加，我们的宇宙空间就好比气球的表面。这也说明，这种局域静止系的存在其实并不意味着牛顿绝对时空观的复活。

宇宙学原理告诉我们，不同空间点的局域静止系是完全平权的，也就是均匀的，同时，整个空间相对于任何一个局域静止系都是各向同性的。

他们用坚硬的黑
透明的亮
在太空建起一座金字塔
空中飘来百亿年前的那幅名画

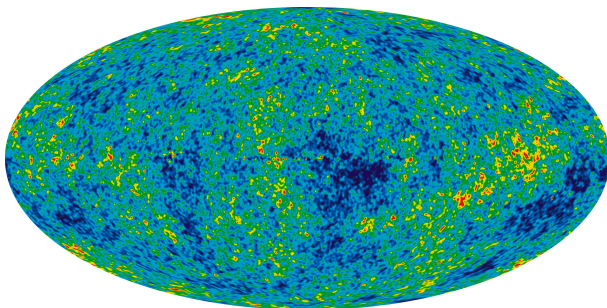


图 16.1: 宇宙微波背景辐射。

塔心那颗小小芥子
是生在画师之前的宇宙吧？

16.2 宇宙膨胀动力学

16.2.1 FLRW度规

在宇宙学中，需要定义一些基本观察者，也就是相对于宇宙微波背景局域静止的观察者，也称作共动观察者。这些基本观察者的世界线互不相交，因此时空中每一个非奇异的点(奇异点比如有大爆炸开始的点)都有唯一一条世界线穿过，因此所有这些世界线的线束就构成了整个时空。利用这些共动观察者，我们可以定义 $t = \text{常数}$ 的空间超曲面，这只需要它与基本观察者的协变速度相正交即可。接下来我们可以取基本观察者的固有时为整个宇宙同步的时间坐标，记作 t ，此外我们还可以给每一条世界线指定三个空间坐标 $x^i, i = 1, 2, 3$ (沿着世界线是常值)，因此每一个共动观察者都有一个固定的 x^i ，称作**共动坐标**，它参数化了 $t = \text{常数}$ 的空间超曲面。由于这些空间超曲面与共动观察者世界线相垂直，因此宇宙的时空度规可以写成

$$ds^2 = -dt^2 + g_{ij}dx^i dx^j, \quad (16.5)$$

式中 g_{ij} 是 t, x^i 的函数，它为空间超面上的诱导度规。

根据宇宙学原理，宇宙的空间几何(不包括时间)是均匀各向同性的，而根据第十一章关于最大对称空间的知识又可以知道，均匀各向同性的空

间必定是所谓的最大对称空间，因此宇宙的空间几何是一个最大对称空间。进一步，根据第十二章的知识可知，3维最大对称空间只有平坦的欧几里德空间 $\mathbb{R}^3$ 、闭合的三维球面 S^3 、以及开放的三维双曲空间 H^3 这三种。

对于 S^3 ，我们可以把它嵌入在4维欧几里德空间 $\mathbb{R}^4$ 中作为一张超曲面，满足方程

$$(z^1)^2 + (z^2)^2 + (z^3)^2 + (z^4)^2 = L^2, \quad (16.6)$$

S^3 上的度规就是将 $(dz^1)^2 + (dz^2)^2 + (dz^3)^2 + (dz^4)^2$ 限制在这张超曲面上。下面对前三个分量引入球坐标，即令 $(z^1)^2 + (z^2)^2 + (z^3)^2 = L^2 r^2$ ，从而

$$(dz^1)^2 + (dz^2)^2 + (dz^3)^2 = L^2(dr^2 + r^2 d\Omega^2), \quad (16.7)$$

其中 $d\Omega^2 = d\theta^2 + \sin^2 \theta d\phi^2$ 为两维球面上的立体角。另外，根据 S^3 的方程(16.6)，有 $(z^4)^2 = L^2(1 - r^2)$ ，从而 $z^4 dz^4 = -L^2 r dr$ ，进而即有

$$(dz^4)^2 = \frac{(-L^2 r dr)^2}{(z^4)^2} = \frac{(-L^2 r dr)^2}{L^2(1 - r^2)} = L^2 \frac{r^2}{1 - r^2} dr^2. \quad (16.8)$$

将(16.7)式与(16.8)式加起来，即将 S^3 上的度规 $ds^2 = [(dz^1)^2 + (dz^2)^2 + (dz^3)^2 + (dz^4)^2]|_{S^3}$ 写成

$$ds_{S^3}^2 = L^2 \left(\frac{1}{1 - r^2} dr^2 + r^2 d\Omega^2 \right). \quad (16.9)$$

对于 H^3 ，根据第十二章的知识，我们可以把它嵌入1+3维闵可夫斯基时空 $\mathbb{R}^{1,3}$ ，作为其中一张双曲面，满足如下方程

$$-(z^0)^2 + (z^1)^2 + (z^2)^2 + (z^3)^2 = -L^2, \quad (16.10)$$

H^3 上的度规就是将 $\mathbb{R}^{1,3}$ 的度规 $-(dz^0)^2 + (dz^1)^2 + (dz^2)^2 + (dz^3)^2$ 限制在这张双曲面上而得到。类似的，对后三个分量引入球坐标，即令 $(z^1)^2 + (z^2)^2 + (z^3)^2 = L^2 r^2$ ，因此当然也有(16.7)式。另一方面，根据(16.10)式，显然有 $(z^0)^2 = L^2(1 + r^2)$ ，从而 $z^0 dz^0 = L^2 r dr$ ，类似的，进而即有

$$-(dz^0)^2 = -L^2 \frac{r^2}{1 + r^2} dr^2. \quad (16.11)$$

将(16.11)式与(16.7)式加起来，即将 H^3 上的度规写成

$$ds_{H^3}^2 = L^2 \left(\frac{1}{1 + r^2} dr^2 + r^2 d\Omega^2 \right). \quad (16.12)$$

对于平坦的 $\mathbb{R}^3$ ，通过引入球坐标，其度规显然可以写成

$$ds_{\mathbb{R}^3}^2 = L^2(dr^2 + r^2 d\Omega^2). \quad (16.13)$$

因此我们可以将这三种情况综合性地写作

$$ds_{\text{Space}}^2 = h_{ij} dx^i dx^j = L^2 \left(\frac{1}{1 - kr^2} dr^2 + r^2 d\Omega^2 \right), \quad (16.14)$$

式中 h_{ij} 只是空间坐标 x^i 的函数，其中对于正曲率的 S^3 , $k = 1$ ，对于平坦的 $\mathbb{R}^3$, $k = 0$ ，对于负曲率的 H^3 , $k = -1$ 。综合三种情况，空间的截面曲率 K 可以写成

$$K = \frac{k}{L^2}. \quad (16.15)$$

对于 $k = 1$ 情形。此时，当 $r \rightarrow 1$ 时，度规(16.14)看起来是奇异的。我们可以定义

$$r = \sin \chi, \quad 0 \leq \chi \leq \pi. \quad (16.16)$$

进而将度规(16.14)重写成

$$ds_{S^3}^2 = L^2(d\chi^2 + \sin^2 \chi d\Omega^2). \quad (16.17)$$

对于 $k = 0$ 情形。可以直接令 $r = \chi$ ，进而将度规(16.14)重写成

$$ds_{\mathbb{R}^3}^2 = L^2(d\chi^2 + \chi^2 d\Omega^2). \quad (16.18)$$

对于 $k = -1$ 情形。我们令

$$r = \sinh \chi, \quad (16.19)$$

则可以将度规(16.14)重写成

$$ds_{H^3}^2 = L^2(d\chi^2 + \sinh^2 \chi d\Omega^2). \quad (16.20)$$

常常也将这三种情形统一写作

$$ds_{\text{Space}}^2 = L^2(d\chi^2 + S_k^2(\chi) d\Omega^2), \quad (16.21)$$

式中

$$S_k(\chi) = \begin{cases} \sin \chi, & k = 1, \\ \chi, & k = 0, \\ \sinh \chi, & k = -1. \end{cases} \quad (16.22)$$

当然，由于宇宙是在膨胀的，其空间大小在整体性地随时间膨胀，所以(16.14)在作为宇宙的空间几何时，要乘上一个随时间演化的缩放因子 $a^2(t)$ ，其中 $a(t)$ 就是所谓的尺度因子，它在宇宙的今天(即 t_0 时刻)取1，即 $a(t_0) = 1$ 。从而(16.5)中的空间度规 g_{ij} 必定可以写成

$$g_{ij}(t, x) = a^2(t)h_{ij}(x). \quad (16.23)$$

根据(16.5)式，从而可以将整个宇宙时空的度规取成

$$\begin{aligned} ds^2 &= -dt^2 + a^2(t)L^2\left(\frac{1}{1-kr^2}dr^2 + r^2d\Omega^2\right) \\ &= -dt^2 + a^2(t)L^2[d\chi^2 + S_k^2(\chi)d\Omega^2] \end{aligned} \quad (16.24)$$

这就是著名的弗里德曼-勒梅特-罗伯逊-沃尔克度规，简称FLRW度规。其中， t 时刻宇宙空间的尺寸即是 $a(t)L$ ，由于 $a(t_0) = 1$ ，所以 L 即是宇宙在今天的尺寸。

哈勃定律

由于宇宙空间的均匀各向同性，所以我们可以取任意一个共动观察者为中心，令其 $\chi = 0$ ，则另一共动观察者到他的共动距离就是 χ ，很显然， t 时刻这两者之间的固有距离 $d_P(t)$ 则是

$$d_P(t) = a(t)L\chi. \quad (16.25)$$

很明显，随着宇宙的膨胀，这两个共动观察者之间的固有距离在增加，换言之，一个观察者在相对于另一个观察者退行，不难得到，退行速度 $u(t) = \dot{d}_P(t)$ 为

$$u(t) = \dot{d}_P(t) = \dot{a}L\chi = \frac{\dot{a}}{a}d_P(t). \quad (16.26)$$

习惯上常常定义

$$H(t) \equiv \frac{\dot{a}}{a}, \quad (16.27)$$

称之为哈勃参数，它具有时间倒数的量纲。则退行速度满足

$$u(t) = H(t)d_P(t). \quad (16.28)$$

也即是说，固有距离越远则相对退行的速度就越大，这就是哈勃定律。哈勃正是通过观察到远处的星系在远离我们(也就是相对于我们退行)，从而发现宇宙在膨胀的。哈勃参数 $H(t)$ 依赖于宇宙时间 t ，当 $t = t_0$ 时，也就是宇宙今天，相应的 $H(t_0)$ 常常记作 H_0 ，称之为哈勃常数。

从哈勃定律(16.28)可知，只要 d_P 足够大，则退行速度 $u(t)$ 完全有可能超过光速1。那这是不是违反了相对论的基本原理呢？并非如此，因为相对论原理说的其实是，任何粒子的世界线都应该是类时或者类光曲线，不能类空。这句话换种说法其实应该是，任何粒子相对于当地观察者的速率都应该小于等于光速1。注意，这里的关键是，这个速度是相对于当地观察者而言的。而遥远星系的退行速度并不是相对星系当地的观察者的，而是相对于远方的我们而言的，那这当然是可以超过1的，并不违反相对论原理。

关于哈勃常数 H_0 的测量值，目前不同的测量方法会得到不一致的测量结果(主要是通过对晚期宇宙的观测(如通过对Ia型超新星的观测)得出来的结果，与宇宙微波背景(CMB)数据对宇宙学标准模型的全局拟合得出来的结果不一致)，为什么不一致今天人们还没搞清楚，有时候人们也称之为哈勃常数危机。典型的测量值是在 $67 \sim 74(\text{km/s})/\text{Mpc}$ 之间。Mpc即是百万秒差距，这是天文测量中常用的距离单位，1Mpc大约为330万光年。如果取 $H_0 = 73(\text{km/s})/\text{Mpc}$ ，那这就表示相距330万光年的两个星系，在以大约每秒73公里的速度彼此远离。这一速度是地球围绕太阳公转速度的两倍多。

16.2.2 联络和曲率

FLRW度规中只有尺度因子 $a(t)$ 随时间的演化需要求解， $a(t)$ 是由爱因斯坦场方程决定的，为了应用场方程，我们又需要先计算出与FLRW度规对应的联络系数以及里奇曲率。下面就是具体计算过程。

为了计算出联络系数，我们对如下世界线作用量(λ 为世界线参数)进行变分

$$S = \frac{1}{2} \int d\lambda \left[- \left(\frac{dt}{d\lambda} \right)^2 + a^2(t) h_{ij} \frac{dx^i}{d\lambda} \frac{dx^j}{d\lambda} \right]. \quad (16.29)$$

比方说对 $t(\lambda)$ 的变分会给出

$$\delta S = \int d\lambda \left[\frac{d^2 t}{d\lambda^2} + a \dot{a} h_{ij} \frac{dx^i}{d\lambda} \frac{dx^j}{d\lambda} \right] \delta t + \dots \quad (16.30)$$

式中 $\dot{a} = \frac{da}{dt}$, 由这个变分结果即有方程

$$\frac{d^2 t}{d\lambda^2} + a \dot{a} h_{ij} \frac{dx^i}{d\lambda} \frac{dx^j}{d\lambda} = 0, \quad (16.31)$$

进而即可以读出如下非零的联络系数

$$\Gamma_{ij}^0 = a \dot{a} h_{ij}. \quad (16.32)$$

类似的, 对 x^k 的变分将给出如下方程

$$\frac{d^2 x^k}{d\lambda^2} + 2 \frac{\dot{a}}{a} \frac{dt}{d\lambda} \frac{dx^k}{d\lambda} + \frac{1}{2} h^{kl} \{ \partial_i h_{lj} + \partial_j h_{li} - \partial_l h_{ij} \} \frac{dx^i}{d\lambda} \frac{dx^j}{d\lambda} = 0. \quad (16.33)$$

式中矩阵 h^{ij} 为矩阵 h_{ij} 的逆。进而又可以读出如下非零联络系数

$$\Gamma_{0j}^i = \frac{\dot{a}}{a} \delta_j^i, \quad \Gamma_{ij}^k = \frac{1}{2} h^{kl} \{ \partial_i h_{lj} + \partial_j h_{li} - \partial_l h_{ij} \}. \quad (16.34)$$

其余联络系数均为零。很显然, 其中 Γ_{ij}^k 就是按照空间度规 $h_{ij}(x)$ 算出来的标准联络系数。

特别的, 对于共动观察者, 其本动速度 $\frac{dx^i}{d\lambda} \equiv 0$, 由此再求导一次, 当然也有 $\frac{d^2 x^i}{d\lambda^2} \equiv 0$, 从而自动满足测地线方程(16.33), 而且这时候测地线方程(16.31)简化成了 $\frac{d^2 t}{d\lambda^2} = 0$, 由于 t 是共动观察者的固有时, 而根据测地线方程的定义, λ 的自然选取也是共动观察者的固有时, 即取 $\lambda = t$, 所以测地线方程(16.31)也是自动满足的。由此可见, 在FLRW时空中, 共动观察者走的就测地线。

有了联络系数以后, 我们就可以进一步计算黎曼曲率张量, 当然, 由于我们的最终目的是应用爱因斯坦场方程, 所以实际上只需计算里奇曲率张量。根据黎曼曲率张量的计算公式, 有

$$R_{\mu\nu} = R^\rho_{\mu\rho\nu} = \partial_\rho \Gamma^\rho_{\nu\mu} - \partial_\nu \Gamma^\rho_{\rho\mu} + \Gamma^\rho_{\rho\sigma} \Gamma^\sigma_{\nu\mu} - \Gamma^\rho_{\nu\sigma} \Gamma^\sigma_{\rho\mu}. \quad (16.35)$$

代入上面得出的联络系数, 立即有

$$R_{0i} = 0. \quad (16.36)$$

以及

$$R_{00} = -\partial_t \Gamma_{0i}^i - \Gamma_{0j}^i \Gamma_{0i}^j = -3 \frac{\ddot{a}}{a}. \quad (16.37)$$

以及

$$R_{ij} = \partial_t \Gamma_{ji}^0 + \Gamma_{k0}^k \Gamma_{ji}^0 - \Gamma_{jk}^0 \Gamma_{0i}^k - \Gamma_{j0}^k \Gamma_{ki}^0 + \tilde{R}_{ij}, \quad (16.38)$$

式中 $\tilde{R}_{ij}$ 为按照空间度规 $h_{ij}(x)$ 算出来的里奇张量, 它由下式给出

$$\tilde{R}_{ij} = \partial_k \Gamma_{ji}^k - \partial_j \Gamma_{ki}^k + \Gamma_{lk}^l \Gamma_{ji}^k - \Gamma_{jk}^l \Gamma_{li}^k. \quad (16.39)$$

由于空间度规 h_{ij} 描述的是一个最大对称空间, 其截面曲率为 $K = k/L^2$, 根据第十一章的知识可知, $h_{ij}(x)$ 给出的黎曼曲率张量为

$$\tilde{R}_{lkij} = K(h_{li}h_{kj} - h_{lj}h_{ki}) \Rightarrow \tilde{R}_{ij} = 2Kh_{ij}. \quad (16.40)$$

将这个结果代入(16.38)式, 同时代入 Γ_{ij}^0 和 Γ_{0j}^i 的结果, 即可以得到

$$R_{ij} = \left[\frac{\ddot{a}}{a} + 2\left(\frac{\dot{a}}{a}\right)^2 + 2\frac{K}{a^2} \right] g_{ij}, \quad (16.41)$$

式中 $g_{ij} = a^2(t)h_{ij}$ 。

根据以上结果, 不难得到里奇标量为

$$R = 6 \left[\frac{\ddot{a}}{a} + \left(\frac{\dot{a}}{a}\right)^2 + \frac{K}{a^2} \right]. \quad (16.42)$$

进而即能得到爱因斯坦张量 $G_{\mu\nu}$ 的非零分量如下

$$G_{00} = 3 \left[\left(\frac{\dot{a}}{a}\right)^2 + \frac{K}{a^2} \right], \quad G_{ij} = - \left[2\frac{\ddot{a}}{a} + \left(\frac{\dot{a}}{a}\right)^2 + \frac{K}{a^2} \right] g_{ij}. \quad (16.43)$$

16.2.3 弗里德曼方程

在大尺度上宇宙的物质内容可以看成是理想流体, 其能量动量张量为第十四章中引入的

$$T^{\mu\nu} = (\rho + p)u^\mu u^\nu + pg^{\mu\nu}. \quad (16.44)$$

由于星系的本动速度相对于宇宙膨胀速度来说都很小, 所以我们可以假定, 在共动观察者看来, 宇宙物质的这理想流体是局域静止的, 也即是说,

在FLRW坐标中，这流体的协变速度 $u^\mu = (1, 0, 0, 0)$ 。从而 $T^{00} = \rho$, $T^{0i} = 0$, $T^{ij} = g^{ij}p$, 换言之，

$$T^\mu_\nu = \text{diag}(-\rho, p, p, p). \quad (16.45)$$

宇宙空间的均匀性意味着 ρ 和 p 均只依赖于时间 t 。

能动量张量的守恒方程 $D_\mu T^{\mu\nu} = 0$ 实际上反映了流体在时空中的演化，其中 $D_\mu T^\mu_i = 0$ 会给出物质组分满足的测地线方程，也就是共动观察者的测地线方程，正如前面说过的，这是自动满足的。而能量守恒方程 $D_\mu T^\mu_0 = 0$ 将给出

$$\begin{aligned} 0 &= \partial_\mu T^\mu_0 + \Gamma^\mu_{\mu\nu} T^\nu_0 - \Gamma^\rho_{\mu 0} T^\mu_\rho \\ &= -\dot{\rho} + \Gamma^i_{i0} T^0_0 - \Gamma^i_{j0} T^j_i \\ &= -\dot{\rho} - 3\frac{\dot{a}}{a}(\rho + p). \end{aligned} \quad (16.46)$$

这个能量守恒方程有一个直观解释，为此我们把它改写成

$$d(\rho a^3) = -p d(a^3), \quad (16.47)$$

考虑到一个空间区域的体积正比于 $a^3(t)$ ，所以这个方程的意思就是，某空间区域膨胀向外做的功等于区域内物质总能量的减少。

另一方面，根据爱因斯坦场方程，我们有

$$G_{00} = 8\pi G T_{00}, \Rightarrow 3\left[\left(\frac{\dot{a}}{a}\right)^2 + \frac{K}{a^2}\right] = 8\pi G \rho. \quad (16.48)$$

重新整理一下，即可以写成

$$\left(\frac{\dot{a}}{a}\right)^2 = \frac{8\pi G}{3}\rho - \frac{K}{a^2}. \quad (16.49)$$

这就是著名的弗里德曼方程，它告诉我们宇宙是如何膨胀的。

爱因斯坦场方程的空间分量 $G_{ij} = 8\pi G T_{ij}$ 将给出

$$\begin{aligned} -2\left[\frac{\ddot{a}}{a} + \left(\frac{\dot{a}}{a}\right)^2 + \frac{K}{a^2}\right] &= 8\pi G p \\ \Rightarrow \frac{\ddot{a}}{a} &= -\frac{4\pi G}{3}(\rho + 3p). \end{aligned} \quad (16.50)$$

但是这不是一个独立的方程，因为你可以通过将弗里德曼方程对时间求导，然后代入能量守恒方程(16.46)来得到它，虽然人们有时候也称这个方程为第二弗里德曼方程。

总之，要描述宇宙的演化，我们需要两个相互独立的基本方程，分别是能量守恒方程(16.46)，以及弗里德曼方程(16.49)。

16.2.4 宇宙学红移

前面说过, 哈勃观测到远处的星系在远离我们(退行), 从而知道了宇宙在膨胀。但是, 远处星系的退行速度是如何观测的呢? 当然, 是通过发现远处星系发给我们的光信号有红移, 如果这星系离我们的距离不是过于遥远, 从而退行速度不大的话, 那么这红移可以解释为多普勒红移。但是, 一般来说, 宇宙学红移和多普勒红移并不是一回事。

为了讲清楚远方星系光信号的宇宙学红移, 不妨假设我们的观察者位于共动坐标的原点 $\chi = 0$, 远方的星系位于共动坐标 χ_A , 假设光信号沿着径向从 χ_A 传播到原点处的观察者。由于光信号走的是类光测地线, 满足 $ds^2 = 0$, 进而根据FLRW度规, 我们有

$$-dt^2 + a^2(t)L^2d\chi^2 = 0 \Leftrightarrow \frac{dt}{a(t)} = -Ld\chi, \quad (16.51)$$

式中注意到了沿着光路径 χ 在减少。假设光信号是 t_1 时刻发出的, t_2 时刻被接收到, 则将上式积分, 即有

$$\int_{t_1}^{t_2} \frac{dt}{a(t)} = L \int_0^{\chi_A} d\chi = L\chi_A. \quad (16.52)$$

经过一个周期以后, 信号发出的时刻变成了 $t_1 + \delta t_1$, 接收时刻变成了 $t_2 + \delta t_2$, 其中 δt_1 就是光发出时的时间周期, δt_2 则是光被接收时的时间周期。很显然, 依然有

$$\int_{t_1 + \delta t_1}^{t_2 + \delta t_2} \frac{dt}{a(t)} = L\chi_A. \quad (16.53)$$

将(16.53)式减去(16.52)式, 并考虑到时间周期 δt_1 和 δt_2 均很小, 从而即有

$$\frac{\delta t_2}{a(t_2)} - \frac{\delta t_1}{a(t_1)} = 0. \quad (16.54)$$

假设光发出时的波长为 λ_1 , 则很显然 $\lambda_1 = \delta t_1$, 同样, 光被接收到时的波长 $\lambda_2 = \delta t_2$, 则根据上面的式子我们有

$$\frac{\lambda_2}{\lambda_1} = \frac{\delta t_2}{\delta t_1} = \frac{a(t_2)}{a(t_1)}, \quad (16.55)$$

显然, 波长变了, 也就是发生了红移, 这就是宇宙学红移。

在实际应用中, 接收光信号的当然就是我们, 我们是在今天(即 t_0 时刻)接收到光信号的, 而光信号发出的时刻是过去的某个 t 时刻, 发出时的

波长可以记为 λ , 接收时的波长记为 λ_0 , 则可以通过下式定义光信号的宇宙学红移 z

$$1 + z \equiv \frac{\lambda_0}{\lambda} = \frac{a(t_0)}{a(t)} = \frac{1}{a(t)}, \quad (16.56)$$

式中最后一个等于号利用了 $a(t_0) = 1$ 。这个式子告诉我们, 只要我们观测到远方星系光信号的红移量 $z > 0$, 那就说明过去的 $a(t) < 1$, 从而说明宇宙是在膨胀! 特别的, 宇宙开始于大爆炸奇点, 即 $a(0) = 0$, 这对应于 $z = \infty$ 的红移。而宇宙今天对应 $a(t_0) = 1$, 这相应于 $z = 0$ 。不仅如此, 星系形成以后的宇宙相应于 $z < 10$, 而光子脱耦形成宇宙微波背景辐射的时刻相应于 $z \simeq 1000$ 。

有了宇宙学红移的概念以后, 就可以讨论哈勃如何通过天文观测发现哈勃定律的了。实际上, 哈勃是对远方星系进行观测, 而且这个远方星系不是太远(相比于宇宙的大小), 因此光子从它发出的时间 t 可以写成 $t = t_0 - \delta t$, t_0 就是我们接收到光子的时间, 也就是宇宙的今天, 由于被观测的星系离我们不是太远, 所以它发出光子的时刻也不会比 t_0 早很多, 因此 $\delta t \ll t_0$ 。由此即可以对尺度因子进行展开

$$\begin{aligned} a(t) &= a(t_0 - \delta t) \\ &= a(t_0) - \delta t \dot{a}(t_0) + \frac{1}{2}(\delta t)^2 \ddot{a}(t_0) + \dots \\ &= a(t_0) \left[1 - \delta t H(t_0) - \frac{1}{2}(\delta t)^2 q(t_0) H^2(t_0) + \dots \right], \end{aligned} \quad (16.57)$$

当然式中的 $a(t_0)$ 其实就是1, 而 $q(t)$ 是所谓的减速参数, 其定义是

$$q(t) = -\frac{\ddot{a}(t)a(t)}{\dot{a}^2(t)}. \quad (16.58)$$

也即是说, 如果 $q(t) > 0$, 那就说明宇宙在减速膨胀, 反之, 如果 $q(t) < 0$, 那宇宙就在加速膨胀。习惯上常常记 $a_0 = a(t_0) = 1$, $H(t_0) = H_0$, $q(t_0) = q_0$, 其中 q_0 就是宇宙当前的减速参数。

利用 H_0, q_0 , 就可以把远方星系的红移因子表示为

$$\begin{aligned} z &= \frac{a(t_0)}{a(t)} - 1 = \left[1 - \delta t H(t_0) - \frac{1}{2}(\delta t)^2 q(t_0) H^2(t_0) + \dots \right]^{-1} - 1 \\ &= \delta t H_0 + (\delta t)^2 \left(1 + \frac{1}{2} q_0 \right) H_0^2 + \dots \end{aligned} \quad (16.59)$$

由于 δt 为一个小量, 所以红移 z 也为一个小量, $z \ll 1$, 进而即有

$$\delta t = \frac{1}{H_0} \left[z - \left(1 + \frac{1}{2} q_0 \right) z^2 + \dots \right], \quad (16.60)$$

注意, 这个式子只依赖于当前的参数 H_0 和 q_0 , 与宇宙的演化历史无关。

另一方面, 根据(16.52)式, 可以得到这个星系离我们的共动距离 $d_P = a_0 L_{\chi_A}$ 为

$$\begin{aligned} d_P = a_0 L_{\chi_A} &= a_0 \int_t^{t_0} \frac{dt}{a(t)} = \int_t^{t_0} dt [1 - (t_0 - t)H_0 + \dots]^{-1} \\ &= \left[\delta t + \frac{1}{2} (\delta t)^2 H_0 + \dots \right]. \end{aligned} \quad (16.61)$$

代入 δt 的表达式(16.60), 即有

$$d_P = \frac{1}{H_0} \left[z - \frac{1}{2} (1 + q_0) z^2 + \dots \right]. \quad (16.62)$$

特别的, 对于不太远的星系, $z \ll 1$, 所以作为一阶近似, 即有

$$z = H_0 d_P. \quad (16.63)$$

对于这种不太远的星系, 我们可以将其红移看作是多普勒红移, 从而可得其退行速度 $u = z$, 进而(16.63)式给出的就是哈勃定律。

但是, 如果我们保留到距离红移关系(16.62)的二阶近似, 我们就会发现它和宇宙减速参数 q_0 有关, 给定远方星系的红移量 z , 如果宇宙在减速膨胀 $q_0 > 0$, 则相对于 $q_0 = 0$ 的情形, 远方星系的距离就会偏小, 相反, 如果宇宙在加速膨胀, $q_0 < 0$, 则远方星系的距离就会偏大。1998年, Perlmutter团队和Riess与Schmidt团队观测了红移大约在0.5左右的Ia型超新星(一种亮度非常稳定, 可以用作“标准烛光”的天体), 发现这些超新星比预期要暗, 意味着它们离我们的距离偏大, 进而推断出宇宙在加速膨胀。具体来说, 就是通过拟合这些超星系的距离红移关系(严格来说, 拟合的是光度距离和红移的关系, 但与上面的共动距离其实密切相关), 进而推断出 $q_0 < 0$ 。

更一般地, 由于哈勃参数 $H(t)$ 是时间的函数, 而通过尺度因子与时间的关系 $a(t)$ 又可以把时间看成是红移的函数, 因此哈勃参数完全可以看作是红移的函数, 记作 $H(z)$ 。进而有

$$\begin{aligned} dz &= d(1 + z) = d\left(\frac{a_0}{a}\right) = -\frac{a_0}{a} \frac{\dot{a}}{a} dt \\ &= -(1 + z)H(z)dt. \end{aligned} \quad (16.64)$$

因此,

$$t_0 - t = \int_t^{t_0} dt = \int_0^z \frac{dz'}{(1+z')H(z')}. \quad (16.65)$$

特别的, 如果取 $t = 0$, 相应 $z = \infty$, 则有宇宙年龄的计算公式

$$t_0 = \int_0^\infty \frac{dz}{(1+z)H(z)}. \quad (16.66)$$

类似的, 也有精确的距离-红移关系

$$d_P = a_0 \int_t^{t_0} \frac{dt}{a(t)} = \int_0^z \frac{dz'}{H(z')}. \quad (16.67)$$

很显然, 无论是为了计算出宇宙年龄也好, 还是为了得到精确的距离-红移关系也好, 我们都需要先知道 $H(z)$ 的表达式, 为此就需求助于弗里德曼方程, 下一小节我们再来讨论这个问题。

16.2.5 宇宙学参数及其确定

物质、辐射以及宇宙学常数

宇宙的不同物质组分有不同的物态方程, 通常可以刻画为

$$p = w\rho. \quad (16.68)$$

其中 w 是一个与时间无关的常数, 称之为物态参数。如果要求物质内容满足主能量条件, 那么 $|w| \leq 1$ 。一般来说, 这些物质组分可以分成如下三类:

1. $w = 0$, 即 $p = 0$, 这是一些非相对论性的物质, 即所谓的尘埃物质, 后文也简称为**物质**(和前文的物质概念有所不同)。比如所有星系构成的理想流体就可以看成是这样的物质, 一个星系就是这尘埃中的一个“颗粒”。**2.** $w = 1/3$, 根据第十四章的知识, 这种物质组分是极端相对论性的, 后文简称为**辐射**。比如充满宇宙空间的微波背景辐射。**3.** $w = -1$, 即 $p = -\rho$, 比方说宇宙学常数(根据第十四章的知识)。

将物态方程(16.68)代入能量守恒方程(16.46), 即可以得到 $\dot{\rho}/\rho = -3(1+w)\dot{a}/a$, 积分即可以得到

$$\rho \propto a^{-3(1+w)}. \quad (16.69)$$

很显然，对于物质

$$\rho_M \propto a^{-3}, \quad (16.70)$$

由于空间区域的体积 $\propto a^3$ ，所以很显然，对于物质，随着宇宙的膨胀，一个空间区域内的总能量保持为常数。但是，对于辐射

$$\rho_R \propto a^{-4}, \quad (16.71)$$

很显然，随着宇宙的膨胀(即 a 的增加)，辐射的能量密度衰减得更快，以至于一个空间区域内的辐射总能量不可能保持为常数。这是因为随着宇宙的膨胀，光子的波长也膨胀了，从而光子的频率就降低了，也就是所谓的红移了。而对于宇宙学常数 Λ ，由于 $w = -1$ ，所以很显然，

$$\rho_\Lambda \propto a^0, \quad (16.72)$$

当然，这是我们早就知道的，宇宙学常数的能量密度是一个常数，不随宇宙膨胀而衰减。

从以上分析可知，随着宇宙的膨胀，辐射的能量密度衰减得最快，物质次之，而宇宙学常数(真空能)的能量密度则是不变的。当今的天文学观测发现，宇宙中的可见物质约占4%，而暗能量(可以用宇宙学常数来代表)则约占68%。由此可以想见，在极早期宇宙中物质和辐射的能量应该占据绝大部分的比例，而暗能量的比重则微乎其微，但是随着宇宙的膨胀，最终却是暗能量取得了主导地位。进一步推论，如果宇宙继续膨胀下去，暗能量的比重将越来越大，其它物质成分将可以被忽略，从而宇宙将进入一个渐进德西特时空。但是，由于辐射随着宇宙膨胀衰减得最快，所以倒推回去，最早的时候，宇宙必定是由极端相对论性的辐射主导的，然后过渡到物质主导，最后可能再过渡到暗能量主导。物质和辐射相等的时刻大约对应于红移 $z \simeq 10^4$ 。

在辐射主导的时期， $\rho \sim a^{-4}$ ，代入弗里德曼方程(16.49)，由于此时 a 很小，而空间曲率能 $-\frac{k}{a^2 L^2} \sim a^{-2}$ ，因此相比于辐射能可以忽略，进而可将弗里德曼方程(16.49)近似为

$$\left(\frac{\dot{a}}{a}\right)^2 \sim a^{-4} \Rightarrow \dot{a} \sim a^{-1}, \quad (16.73)$$

积分即可以得到

$$a(t) \sim t^{1/2}. \quad (16.74)$$

特别的, $t = 0$ 时, $a(0) = 0$, 相应的空间曲率($\sim a^{-2}$)即为无穷大, 物质的能量密度也为无穷大。这说明 $t = 0$ 时, 时空是奇性的, 这就是所谓的大爆炸奇点。

在物质主导的时期, $\rho \sim a^{-3}$, 代入弗里德曼方程(16.49)并忽略空间曲率能, 即有

$$\left(\frac{\dot{a}}{a}\right)^2 \sim a^{-3} \Rightarrow \dot{a} \sim a^{-1/2}, \quad (16.75)$$

积分即可以得到

$$a(t) \sim t^{2/3}. \quad (16.76)$$

不难看出, 无论是辐射主导的时期还是物质主导的时期, 虽然宇宙都在膨胀, 但是均有

$$\ddot{a} < 0, \quad (16.77)$$

因此宇宙是减速膨胀的。

但是在宇宙学常数主导的时期, $\rho = \frac{\Lambda}{8\pi G}$, 代入弗里德曼方程(16.49)并忽略空间曲率能, 即有

$$\left(\frac{\dot{a}}{a}\right)^2 = \frac{\Lambda}{3}, \quad (16.78)$$

积分即得

$$a(t) \sim e^{\sqrt{\frac{\Lambda}{3}}t}. \quad (16.79)$$

很显然这时宇宙不仅加速膨胀, 而且是指数膨胀。实际上, 正是因为观测到当前宇宙在加速膨胀, 人们才推断出暗能量存在并进一步推断出其在当前约占68%的, 至于这个值如何可以推断出来我们后文再讨论。

利用 $a^{-1} = 1 + z$, 我们也可以将上述各组能量密度对 a 的关系式重新表示成

$$\rho_M = \rho_{M,0}(1+z)^3, \quad \rho_R = \rho_{R,0}(1+z)^4, \quad \rho_\Lambda = \rho_{\Lambda,0}, \quad (16.80)$$

式中 $\rho_{M,0}, \rho_{R,0}, \rho_{\Lambda,0}$ 分别表示当今的物质能量密度、当今的辐射能量密度、以及当今的宇宙学常数能量密度。从而宇宙内含物的总能量密度 $\rho = \rho_R + \rho_M + \rho_\Lambda$ 可以表示成

$$\rho = \rho_{R,0}(1+z)^4 + \rho_{M,0}(1+z)^3 + \rho_{\Lambda,0}. \quad (16.81)$$

宇宙学参数

下面我们把弗里德曼方程(16.49)写成如下形式

$$H^2 = \frac{8\pi G}{3}\rho - K(1+z)^2. \quad (16.82)$$

习惯上常常定义临界密度 $\rho_{c,0}$ 如下

$$\rho_{c,0} \equiv \frac{3H_0^2}{8\pi G} \approx 9.2 \times 10^{-27} \text{kg} \cdot \text{m}^{-3}, \quad (16.83)$$

数值上就是大约每立方米5.5个质子。进而即可以把弗里德曼方程进一步改写成

$$H^2 = H_0^2 \left[\frac{\rho}{\rho_{c,0}} - \frac{K}{H_0^2} (1+z)^2 \right] \quad (16.84)$$

将前面的(16.81)代入这个式子, 即有

$$H^2 = H_0^2 [\Omega_R(1+z)^4 + \Omega_M(1+z)^3 + \Omega_\Lambda + \Omega_k(1+z)^2]. \quad (16.85)$$

式中

$$\Omega_R \equiv \frac{\rho_{R,0}}{\rho_{c,0}}, \quad \Omega_M \equiv \frac{\rho_{M,0}}{\rho_{c,0}}, \quad \Omega_\Lambda \equiv \frac{\rho_{\Lambda,0}}{\rho_{c,0}}, \quad \Omega_k \equiv -\frac{K}{H_0^2}. \quad (16.86)$$

(16.85)式正是我们一直想找的 $H(z)$ 的表达式。特别的, 在(16.85)中取 $z = 0$ (今天), 则有

$$\Omega_R + \Omega_M + \Omega_\Lambda + \Omega_k = 1. \quad (16.87)$$

所以, 参数 $\Omega_R, \Omega_M, \Omega_\Lambda, \Omega_k$ 可以解释成今天宇宙各组分能量密度所占的百分比, 称作宇宙的密度参数。特别的, Ω_k 是空间曲率所占的百分比。不过值得注意的是, Ω_k 可以取负数, 当空间曲率为正时, Ω_k 就是负数。

很显然, Ω_k 是正是负完全取决于 $\Omega_R + \Omega_M + \Omega_\Lambda$ 是大于1还是小于1, 也就是取决于 $\rho_0 = \rho_{R,0} + \rho_{M,0} + \rho_{\Lambda,0}$ 是大于 $\rho_{c,0}$ 还是小于 $\rho_{c,0}$ 。如果 $\rho_0 > \rho_{c,0}$, 则 $\Omega_k < 0$, 空间曲率为正, 宇宙为闭宇宙。如果 $\rho_0 = \rho_{c,0}$, 则 $\Omega_k = 0$, 空间曲率为零, 宇宙为平坦宇宙。如果 $\rho_0 < \rho_{c,0}$, 则 $\Omega_k > 0$, 空间曲率为负, 宇宙为开宇宙。对宇宙微波背景角功率谱第一声学峰位置(见参考文献)的拟合告诉我们 Ω_k 非常接近于0, 即 $\Omega_k \approx 0$, 宇宙非常接近于平坦的。

所以，不同的宇宙学模型基本可以通过如下四个参数来标志

$$H_0, \quad \Omega_R, \quad \Omega_M, \quad \Omega_\Lambda. \quad (16.88)$$

观测宇宙学的一个重要目标就是确定这四个参数的值，一般来说，这是一件非常复杂的事情，要结合多种不同的观测结果。 H_0 的观测结果前面已经说过了，剩下三个密度参数的观测结果大约为

$$\Omega_R \approx 5 \times 10^{-5}, \quad \Omega_M \approx 0.3, \quad \Omega_\Lambda \approx 0.7, \quad (16.89)$$

这也就是前面所说的暗能量约占68%的来源，但观测如何可以定出这个结果呢？考虑到辐射的百分比很低，可以忽略，那观测是如何定出 $\Omega_M \approx 0.3$, $\Omega_\Lambda \approx 0.7$ 的呢？这个问题我们稍后再讨论。

不仅如此，当前的减速参数 q_0 也可以用这些密度参数表达出来。为此我们先把第二弗里德曼方程(16.50)写成如下形式

$$\frac{a\ddot{a}}{\dot{a}^2} = -\frac{4\pi G}{3H^2}(\rho + 3p) = -\frac{4\pi G}{3H^2}(2\rho_R + \rho_M - 2\rho_\Lambda), \quad (16.90)$$

式中最后一个等于号我们代入了各组分的物态方程。进而根据减速参数的定义，即有

$$q_0 = \frac{1}{2}\Omega_M + \Omega_R - \Omega_\Lambda. \quad (16.91)$$

很显然，如果宇宙中没有暗能量(宇宙学常数)，即 $\Omega_\Lambda = 0$ ，则必定有 $q_0 > 0$ ，从而宇宙必定减速膨胀。因此观测到宇宙加速膨胀，就意味着宇宙学常数必定大于零，或者说必定有大于零的暗能量存在。实际上，由 $q_0 < 0$ ，就必定有 $\Omega_\Lambda > \frac{1}{2}\Omega_M$ 。

另一方面，利用(16.85)式给出的 $H(z)$ 表达式，代入宇宙年龄的计算公式(16.66)，即有

$$t_0 = \frac{1}{H_0} \int_0^\infty \frac{dz}{(1+z)E(z)}, \quad (16.92)$$

式中

$$E(z) \equiv \sqrt{\Omega_R(1+z)^4 + \Omega_M(1+z)^3 + \Omega_k(1+z)^2 + \Omega_\Lambda}. \quad (16.93)$$

代入上面给出的各宇宙学参数值(其中 $\Omega_k = 0$)并进行数值积分，就可以得到宇宙年龄 $t_0 \approx 138$ 亿年，常常也记作 $t_0 \approx 13.8\text{Gyr}$ ，Gyr即是Gigayear的缩写，1Gyr=十亿年。粗略来说，也有 $t_0 \approx 1/H_0$ 。

同样的, 假如把 $H(z)$ 的表达式代入距离-红移关系的计算公式(16.67), 即有

$$d_P(z) = \frac{1}{H_0} \int_0^z \frac{dz'}{E(z')}. \quad (16.94)$$

特别的, 如果所考察的红移不太大, 比如 $z < 6$ (通常观测的Ia型超新星都处于这个红移范围), 那注意到 Ω_R 和 Ω_k 都很接近于零, 进而即可以在上述积分中忽略 $E(z)$ 中的 Ω_R 和 Ω_k 相关项, 进而得到

$$d_P(z) = \frac{1}{H_0} \int_0^z \frac{dz'}{\sqrt{\Omega_M(1+z')^3 + \Omega_\Lambda}}, \quad z < 6. \quad (16.95)$$

以上结果告诉了我们一个如何通过拟合Ia型超新星的距离-红移关系, 进而确定 $\Omega_\Lambda \approx 0.7$ 的办法。为此还需要注意到, 由于 Ω_R 和 Ω_k 相比可以忽略, 所以(16.87)式可以近似为

$$\Omega_M + \Omega_\Lambda = 1. \quad (16.96)$$

代入上面的(16.95)式, 即可以得到

$$d_P(z) = \frac{1}{H_0} \int_0^z \frac{dz'}{\sqrt{(1-\Omega_\Lambda)(1+z')^3 + \Omega_\Lambda}}, \quad z < 6. \quad (16.97)$$

也即是说, Ia型超新星的距离红移关系取决于 Ω_Λ 的值, 因此只需要对它们的距离-红移关系进行拟合, 就能够定出 Ω_Λ 的值, 如图(16.2)。这就是确定 Ω_Λ 观测值的重要方法之一。

宇宙命运

当然, 如果是想预言宇宙的未来, 那红移这个概念就不好用了, 这时候我们可以把方程(16.85)恢复成

$$H^2 = H_0^2 [\Omega_R a^{-4} + \Omega_M a^{-3} + \Omega_\Lambda + \Omega_k a^{-2}]. \quad (16.98)$$

利用哈勃参数 H 的定义, 显然也有

$$\frac{1}{2} \dot{a}^2 - \frac{1}{2} H_0^2 [\Omega_R a^{-2} + \Omega_M a^{-1} + \Omega_\Lambda a^2 + \Omega_k] = 0. \quad (16.99)$$

可以把这个方程看作是一个沿着 a 轴运动的一维粒子的机械能守恒方程, 这个粒子的质量为1, 势能为 $V(a) = -\frac{1}{2} H_0^2 [\Omega_R a^{-2} + \Omega_M a^{-1} + \Omega_\Lambda a^2 + \Omega_k]$, 总机械能为零。

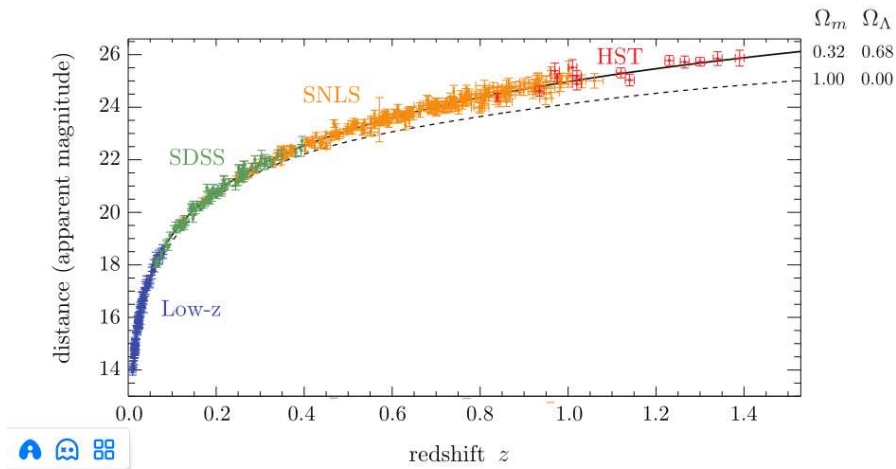


图 16.2: Ia型超新星的距离红移关系(当然严格来说这里的距离是光度距离,但它其实完全可以由共动距离 d_P 决定)与暗能量的发现。如果我们假设宇宙是平坦的,那么超新星显然比在仅由物质构成的宇宙 $\Omega_M = 1$ 中预测的要暗淡(或更远)。(SDSS = Sloan Digital Sky Survey; SNLS = SuperNova Legacy Survey; HST = Hubble Space Telescope.)

通过分析这个一维粒子在势能曲线 $V(a)$ 中的运动,我们就能搞明白宇宙未来的命运(当前这个粒子位于 $a_0 = 1$ 处,速度 $\dot{a} > 0$),是继续膨胀下去,还是被 $V(a)$ 的势垒部分挡住,然后调头开始收缩(在物理上即是由于物质之间的万有引力(为吸引力)占据了主导地位,使得宇宙在某个时刻开始收缩)。由于这种分析不过就是一道略微复杂些的力学习题,所以我们留给读者自行完成。

16.2.6 宇宙学视界

宇宙学视界 (Cosmological Horizon) 是现代宇宙学中的一个重要概念,它描述了我们在宇宙中能够观测到的最远距离。由于宇宙的膨胀和光速的限制,我们无法观测到无限远的宇宙,因此宇宙学视界定义了可观测宇宙的边界。视界的存在是因为宇宙的膨胀速度可能超过光速,因此某些区域的光可能永远无法到达我们。

为了讲清楚视界的概念,让我们考虑一个位于 $\chi = 0$ 处的共动观察者,光信号在 t_E 时刻从远处的 χ_A 处发送过来,在 t_R 时刻被这个观察者接收到,

则根据前面的(16.52)式, 必定有

$$\int_{t_E}^{t_R} \frac{dt}{a(t)} = L_{\chi_A}. \quad (16.100)$$

当然, 这个式子也可以解释成, $\chi = 0$ 处的这个共动观察者在 t_E 时刻向远处的 χ_A 位置发送光信号, t_R 时刻这个光信号到达 χ_A 。

粒子视界

所谓的粒子视界, 指的就是 $\chi = 0$ 位置的这个共动观察者在 t 时刻能够看到的最远距离, 很显然这个最远距离处的光一定是在宇宙诞生之初的 $t = 0$ 时刻发出的, 因此根据(16.100)式, 这个最远距离的 χ_A 值为

$$\int_0^t \frac{dt'}{a(t')} = L_{\chi_A}. \quad (16.101)$$

由于这是共动观察者 t 时刻观测到的, 所以对应的固有距离就应该是 $d_{ph}(t) = a(t)L_{\chi_A}$, 根据上式这也就是

$$d_{ph}(t) = a(t) \int_0^t \frac{dt'}{a(t')}. \quad (16.102)$$

$d_{ph}(t)$ 就是 $\chi = 0$ 处的这个共动观察者在 t 时刻的粒子视界!

当然, 由于 $t \rightarrow 0$ 时, $a(t) \rightarrow 0$, 所以上面这个积分在下限处的贡献可能发散, 如果是这种情况, 那粒子视界就是无穷大, 否则粒子视界就是有限的。比方说, 按照前面的标准宇宙学模型, 最早期的宇宙是辐射主导的, 因此是作减速膨胀, 那时 $a(t) \propto t^{1/2}$, 代入(16.102)式不难发现, 积分下限处的贡献是有限的。进一步, 根据前面的(16.67)式, 可得当前的粒子视界的大小为

$$d_{ph}(t_0) = a_0 \int_0^{t_0} \frac{dt}{a(t)} = \int_0^\infty \frac{dz}{H(z)} = \frac{1}{H_0} \int_0^\infty \frac{dz}{E(z)} \quad (16.103)$$

代入 $E(z)$ 的表达式, 原则上就能算出这个结果。

另外, 将 $d_{ph}(t)$ 对时间求导, 不难导出如下结果

$$\dot{d}_{ph}(t) = H(t)d_{ph}(t) + 1, \quad (16.104)$$

也即是说, 随着宇宙的膨胀($H(t) > 0$), 粒子视界的大小也是在膨胀的, 而且膨胀的速度满足某种哈勃定律。

粒子视界的意义不仅是说明可观测范围。注意光速是一切信号速度的上限，当两个点的空间距离超过视界大小，它们之间就无法用信号建立联系，因此，这两个点上的物理情况是不可能存在因果关系的。从这个意义上讲， t 时刻粒子视界的大小也反映了此时可能有因果联系的区域的大小。

粒子视界的一个直接后果就是宇宙学模型中的视界问题。简而言之，我们观测到即使对于分得非常开的区域，它们的宇宙微波背景辐射温度也是高度一致的。然而在前面的标准宇宙学模型中，这样两个分得很开的点在光子脱耦形成微波背景的时候它们之间的距离可以超过当时的粒子视界大小，因此那时候这两个点之间就不可能有因果联系。而这两点的热平衡也不可能是此后在宇宙继续膨胀到今天的过程中形成的，因为此后光子已经脱耦变成没有相互作用的自由光子了。那为什么这样的两个点有相同的温度呢？

事件视界

虽然粒子视界会随着宇宙的膨胀而膨胀，但这并不意味着 $\chi = 0$ 处的这个共动观察者在未来可以看到 t 时刻的任何事件，实际上，即使这个观测者一直等到无穷远未来， t 时刻的宇宙也总存在一些他观测不到的区域。实际上，根据(16.100)式，这个共动观察者在无穷远将来能够观察到的最远坐标 χ_A 为

$$\int_t^\infty \frac{dt'}{a(t')} = L\chi_A. \quad (16.105)$$

(当然，这个 χ_A 坐标也可以解释成， t 时刻 $\chi = 0$ 处的观察者向远处发送光信号，这光信号在无穷远未来能够到达的最大 χ 值。)

由于 χ_A 是宇宙在 t 时刻的坐标，所以相应固有距离 $d_{eh}(t)$ 当然就是 $d_{eh}(t) = a(t)L\chi_A$ ，根据上一个式子，这也就是

$$d_{eh}(t) = a(t) \int_t^\infty \frac{dt'}{a(t')}. \quad (16.106)$$

很显然 $d_{eh}(t)$ 就是共动观察者在无穷远将来能看到的宇宙在 t 时刻的最大半径，称作这个共动观察者的事件视界。

不难导出，事件视界满足

$$\dot{d}_{eh}(t) = H(t)d_{eh}(t) - 1. \quad (16.107)$$

这也就是说, 如果某个时刻的事件视界 $d_{eh}(t) < 1/H$, 则随着时间的推移, 它是在缩小的。

哈勃视界

除了以上两种常用宇宙学视界以外, 还有一种视界也很常用, 即所谓的哈勃视界 d_H , 其定义就是退行速度等于光速的那个距离。

具体来说, 根据哈勃定律(16.28), 远处星系相对于 $\chi = 0$ 处共动观察者的退行速度为 $u(t) = H(t)d_P(t)$, $d_P(t)$ 为远处星系离共动观察者的固有距离, 所谓的哈勃视界 d_H , 就是使得 $u(t) = 1$ 的那个距离 $d_P(t)$, 换言之,

$$d_H = \frac{1}{H}. \quad (16.108)$$

当然由于我们取了光速 $c = 1$ 的单位制, 如果恢复成常用单位制的话应该是 $d_H = c/H$ 。常常也把哈勃视界称作哈勃半径。

16.3 暗物质与暗能量

暗物质

我们来考察一下当前宇宙物质的成分, 即 $\rho_{M,0}$ 的成分。暂且假定它主要由星系组成。由此就可以通过天文观测数据对 $\rho_{M,0}$ 进行一个独立的估计, 如果由此得出来的 Ω_M 的确约为前文确定的0.3, 那就说明这个假设基本是成立的。

设当今宇宙星系的数密度为 n , 每个星系的平均质量为 $\overline{M}$, 则 $\rho_{M,0} = n\overline{M}$ 。设宇宙空间中单位体积的光度(也就是光度密度)为 $\mathcal{L}$, 目前它已经有比较可靠的观测值了。再设每个星系的平均光度为 $\overline{L}$, 则 $\mathcal{L} = n\overline{L}$, 进而即有

$$\rho_{M,0} = \mathcal{L} \frac{\overline{M}}{\overline{L}}, \quad (16.109)$$

式中 $\overline{M}/\overline{L}$ 叫做星系的平均质光比。很显然, 为了估计 $\rho_{M,0}$, 我们只需估计 $\overline{M}/\overline{L}$ 。实际的测量当然是对具体星系进行的, 得到的只是该星系的质光比 M/L , 很显然, 只有当这个星系具有相当的代表性时其质光比才可以作为 $\overline{M}/\overline{L}$ 的估计值。好在虽然不同星系的质量差别很大, 可以差好几个数量

级，但是星系质光比的差别却要小很多。因此通常星系质光比的值都可以作为 $\overline{M}/\overline{L}$ 的估计值。

要测星系的质光比，关键是测出星系的总质量。下面介绍测量漩涡星系(如银河系)质量的动力学方法。在漩涡星系中，恒星除无规运动外主要就是以星系的引力为向心力绕星系中心旋转。为简化讨论，可以把漩涡星系看成一个圆盘。则由牛顿引力理论可知，距离星系中心 r 处的恒星的旋转速度 $v(r)$ 由下式决定

$$v^2(r) = G \frac{M(r)}{r}, \quad (16.110)$$

其中 $M(r)$ 是星系在半径 r 之内的质量。函数 $v(r)$ 的曲线就称之为星系的旋转曲线。天文学家把许多星系的旋转曲线都测出来了。

假设 R_0 代表星系光度消失处的半径，则 $M(R_0)$ 就是星系内发光物质的总质量。如果星系在 $r > R_0$ 之外没有质量，则由(16.110)式可知，在 $r > R_0$ 处 $v(r)$ 曲线应该按照 $r^{-1/2}$ 的规律下降，如图(16.3)中的曲线A(虚线)。然而实测的星系旋转曲线并非如此，而是先从星系中心陡然升起，然后近似水平地延伸，直到超出 R_0 以外很远以至于无法测量为止，如图(16.3)中的曲线B。这就表明，漩涡星系在发光部分之外有一个半径比 R_0 大得多的暗

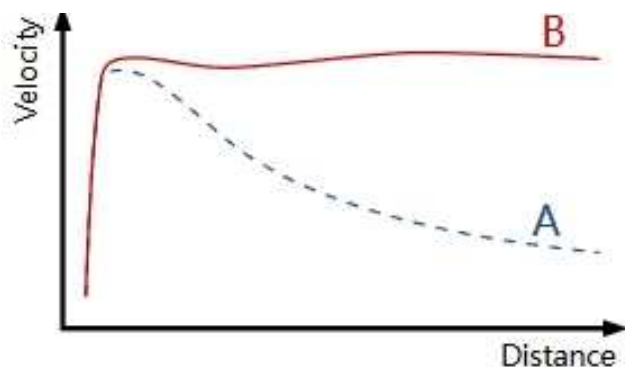


图 16.3: 星系旋转曲线示意图。其中B为实测的曲线，A为我们预期的曲线。

晕，由暗物质组成，其质量约为发光部分的3—10倍。除漩涡星系之外还有其他星系，例如椭圆星系，有证据表明椭圆星系也有数量可观的暗物质。

实际上, 如果仅仅只考虑星系中的发光物质, 并由此估算出 $\rho_{M,0}$, 那么由此得出的 Ω_M 将仅仅为

$$\Omega_M(\text{发光物质}) < 1\% \quad (\text{约为} 0.5\%), \quad (16.111)$$

很显然这远远小于0.3(即30%), 这也说明了应该有大量暗物质存在。总之, 可以肯定, 宇宙中的暗物质在质量上大大超过了发光物质。

暗物质的存在本来不值得惊讶, 因为很容易举出暗物质的例子: 行星、暗淡的小恒星、白矮星、中子星、黑洞、瓦解了的星系以及星际和星系际稀薄气体.... 这些例子中的暗物质都由重子(主要是质子和中子)组成。然而原初核合成理论对当今宇宙重子物质的百分比是有限制的。具体来说如下。

以 n_{b0} 代表当今重子(主要是核子)数密度, m_N 代表每个核子的质量, 则当今重子物质的质量密度 $\rho_{b,0} = n_{b0}m_N$, 它对应的 $\Omega_b = \rho_{b,0}/\rho_{c,0}$ 为

$$\Omega_b = \frac{8\pi G}{3H_0^2} n_{b0} m_N = \frac{8\pi G}{3H_0^2} m_N n_{\gamma 0} \eta, \quad (16.112)$$

式中 $n_{\gamma 0}$ 为当今的光子数密度, 其值可由对宇宙微波背景的观测推算出来, η 是前面原初核合成小节中引入的参数, $\eta = n_b/n_\gamma = n_{b0}/n_{\gamma 0}$ 。前面原初核合成小节中给出了 η 的范围。

由于宇宙微波背景是一个黑体辐射谱, 所以根据斯忒藩-玻尔兹曼定律, 其能量密度正比于温度的四次方, 从而光子数密度正比于温度的三次方, 根据观测到的宇宙微波背景温度为2.728K即可以推算出

$$n_{\gamma 0} \approx 4.1 \times 10^8 \text{ 个}/m^3. \quad (16.113)$$

结合前面给出的 η 的取值范围, 就可以得出

$$2.3\% < \Omega_b < 5.1\%. \quad (16.114)$$

显然 Ω_b 的值大于发光物质值(0.5%), 由此可知, 重子物质中只有一小部分是发光物质, 大部分重子物质为暗物质。

但是, $\Omega_b < 5.1\% < 30\%$ 也告诉我们, 宇宙中的暗物质大部分是非重子暗物质! 当然, 非重子暗物质远不如重子暗物质好想象, 对非重子暗物质候选者的探究已经成为当今物理学的热门课题。实际上, 我们通常所谓的暗物质问题更多指的就是探究这些非重子暗物质。

非重子暗物质大约应为非常稳定的非重子粒子。中微子如果有质量的话，将是很明显的候选者。从1997年开始人们倾向于相信某些中微子有大约0.1eV量级的质量，这些中微子对 Ω_M 的贡献约为0.3% ~ 15%。当然，它们是所谓的热暗物质，热指的是在它们退耦时运动速度依然很大。除了热暗物质以外，还有所谓的冷暗物质，冷指的是粒子以低速运动，它们曾活跃于温度很高的极早期宇宙，并随温度下降而退耦，退耦时速度已经不高，并作为背景遗迹留存至今，所以是冷暗物质。但是，虽然一些超越粒子物理标准模型的理论预言有这样的有质量非重子粒子作为候选者而存在，但目前人们还没有观测到它们。

今天的宇宙在大尺度上是均匀各向同性的，但在小一点的尺度上却呈现出有层次的结构，也就是存在恒星、星系、星系团以及超星系团。正如前面介绍热大爆炸基本概念那一小节(第16.1.1小节)中所讲的，今天的复杂结构起源于极早期宇宙中非常微小的密度涨落，然后通过引力系统的不稳定性而被放大形成。暗物质的存在对于具体解释结构形成也是必须的，因为如果认为宇宙中的物质主要是重子物质，那结构形成理论就会遇到严重困难。

根据非重子暗物质是热暗物质还是冷暗物质的区分，我们有两种结构形成模型。在热暗物质模型中，结构的形成是“自上而下”的，即先形成超星系团，然后再逐级破裂为星系团和星系。相反，在冷暗物质模型中，结构的形成是“自下而上”的，即先形成星系，然后逐级形成星系团和超团。这两类模型中，冷暗物质模型已经取得了很大的成功。今天人们广泛地认为，宇宙物质以非重子暗物质为主，而非重子暗物质中冷暗物质又是必不可少的。

不过，由于观测上一直还没有发现暗物质粒子，所以目前也有一些非主流的理论，其中人们通过修改万有引力规律来避免对暗物质的需要。

暗能量

正如前文讲过的，宇宙加速膨胀的发现，使得人们推断暗能量的存在，对暗能量的最简单也是最广为接受的解释就是宇宙学常数。通常认为宇宙学常数就是量子场的真空能，不过，正如前面的章节中提到过的，根据目前量子场论算出来的真空能比观测宇宙学确定的宇宙学常数大太多了，完全不在一个量级，这到底是怎么回事目前人们还不清楚，这就是著名的宇宙学常数问题。当然，在本书作者看来，最大的可能性是，当前的量子场论只在局域小尺度上才通过了实验检验，这样的局域量子场论很可能不能

适用于宇宙学的大尺度，因此我们通过局域场论来计算整个宇宙学尺度上的真空能是不对的，因为万有引力的存在修改了量子场论在大尺度上的性质。但，关键在于，正确的做法是什么？这就要留给勇敢的读者们自己去探索了。

不过，暗能量也可以不是宇宙学常数，比如也有人提出暗能量密度是某种动态标量场的能量密度，等等。总之，关于暗能量的理论很多，其中宇宙学常数是最简单最广为认可的。但是，根据暗能量光谱仪(Dark Energy Spectroscopic Instrument, DESI)最近的观测数据，也的确有迹象表明暗能量不是一个宇宙学常数，而是会随时间演化的东西。

如果认为暗能量是宇宙学常数，暗物质是冷暗物质(cold dark matter, CDM)，由此得到的宇宙学模型就是所谓的 Λ CDM模型，它不仅与观测结果吻合得比较好，而且也是最广为接受的宇宙学模型，通常也称之为标准宇宙学模型。

16.4 宇宙暴胀

16.4.1 暴胀的基本概念

热大爆炸宇宙学虽然很成功，但它本身也有一些严重的问题。其中**视界问题**我们前面介绍粒子视界时已经讲过了。简单来说，即是按照标准宇宙学模型，光子脱耦时的粒子视界不够大，以至于无法解释超出视界范围的两个点为什么有相同的温度(正如前面讲到过的，宇宙微波背景的温度涨落幅度只有大约百万分之五)。宇宙为什么会如此均匀和各向同性呢？(以至于在宇宙早期即使天空中两个因果无关的点也几乎完全均匀。)

在1970年代之前，一些物理学家说，这是由于热大爆炸开始时的初始条件非常特殊，而初始条件要用量子引力来研究，因此这个问题超出了当前理论物理的能力范围。

热大爆炸无法解释的第二个问题是，**平坦性问题**。即宇宙空间为何如此平坦，为何 Ω_k 如此接近于零，通常来说，这同样意味着需要对初始条件进行极端精细的微调。

第三个问题是所谓的**磁单极子问题**。具体来说即是，在粒子物理大统一理论的框架下，早期宇宙在温度极高时会发生对称性破缺，导致磁单极子的产生。理论预测磁单极子的数密度为 $n_{monopole} \propto T_G^3$ ，其中 T_G 是大统一相变的温度(约为 10^{16} GeV)。根据这一预测，现今宇宙中应该有大量磁单极

子。然而，尽管人们进行了大量实验和观测，至今未发现磁单极子的存在。观测上限表明，磁单极子的数密度应该远低于刚才说到的理论预言。那么，这是为什么呢？磁单极子为何如此稀少呢？

为了解决热大爆炸的这些问题，阿兰·古斯于1980年提出，在早期辐射主导的宇宙减速膨胀阶段之前，宇宙还经历了一个很短暂的指数级快速膨胀阶段，称作暴胀阶段，它将极早期宇宙的一个小区域迅速拉伸到整个可观测宇宙的范围。暴胀发生在大约 10^{-36}s 到 10^{-32}s 之间，这个阶段宇宙大约膨胀了 e^{50} 到 e^{60} 倍，这个暴胀阶段结束以后宇宙才开始按照原来的热大爆炸宇宙学演化。

由于粒子视界 $d_{ph}(t) = a(t) \int_0^t \frac{dt'}{a(t')}$ ，所以如果 $a(t)$ 指数膨胀了 e^{50} 到 e^{60} 倍，那么不难算出， $d_{ph}(t)$ 也会指数膨胀这么多倍。这就能够解决热大爆炸模型中光子脱耦时的粒子视界不够大的问题，因为原来按照热大爆炸模型算出来的没有因果联系的区域现在就可以被包含在同一个粒子视界中了，那这就解决了热大爆炸的视界问题。

同样的，暴胀阶段的存在也能够解决平坦性问题，为了看清楚这一点，假设我们定义 t 时刻的曲率密度参数 $\Omega_{k,t}$ 为 $\Omega_{k,t} \equiv -\frac{K}{a^2(t)H^2}$ ，则不难看出，一个指数膨胀的 $a(t)$ 会把 $\Omega_{k,t}$ 指数压低，从而相当于精细地把宇宙调节到了非常接近于平坦。直观上也就是急剧的膨胀把空间曲率给拉平了。

宇宙暴胀对磁单极子为何如此稀少的解释就更加简单了，因为暴胀导致的宇宙体积的急剧增大一下子把磁单极子给稀释了！

这里有一个问题，暴胀阶段宇宙急剧膨胀这么多倍，那温度应该就急剧降低到接近于零了，那暴胀结束以后宇宙怎么又按照原来的热大爆炸演化了呢，毕竟热大爆炸是需要高温的，这高温从哪里来呢？别着急，这个问题我们后文会给出解释。

先让我们来说说暴胀理论的一个巨大成就。前面讲到过，宇宙的结构形成需要密度涨落的存在，问题是原来的热大爆炸宇宙学很难解释原初的密度涨落从哪里来。暴胀理论能给出自然的解释！当然，你可能想到原初的密度涨落来自于量子涨落，但问题是量子涨落完全是微观的，它和宇宙学不在一个尺度上。但是，暴胀理论恰恰能自然地连接这两个尺度，因为暴胀恰好可以把原来的微观尺度拉大到宇宙学尺度！而且具体进行的计算也表明这个想法是完全能够成立的，与观测可以吻合。当然，这样的计算对于我们的极简宇宙学而言过于专门了，所以我们这里略过。

16.4.2 暴胀场

一个重要的问题是，什么驱动了宇宙暴胀。它不可能是我们今天观测到的暗能量，虽然暗能量的确能让宇宙加速膨胀，但是，暗能量在早期宇宙中占比是可以忽略的。但是暗能量启发我们，只要我们在那个时期的宇宙中引入一种 $w \approx -1$ 的常数能量密度，并让它占主导地位，那它就能和暗能量一样让宇宙加速膨胀甚至指数膨胀。

在粒子物理中，产生常数能量密度的最简单办法就是引入一个不随空间坐标变化的标量场，并让它有一个比较大的势能密度，那这个势能密度就会产生一个常数能量密度。基于这些想法，人们引入了暴胀场，这是一个标量场，它驱动了宇宙暴胀。

作为标量场，暴胀场的作用量为

$$S_I = \int \sqrt{-g} d^4x \left[-\frac{1}{2} g^{\mu\nu} \partial_\mu \phi \partial_\nu \phi - \mathcal{U}(\phi) \right]. \quad (16.115)$$

根据第14章的相关知识可知，作为弯曲时空中的标量场，暴胀场对应于一种理想流体，其能量密度 ρ_I 和压强 p_I 分别为

$$\rho_I = -\frac{1}{2}(\partial\phi)^2 + \mathcal{U}(\phi), \quad p_I = -\frac{1}{2}(\partial\phi)^2 - \mathcal{U}(\phi). \quad (16.116)$$

为了将注意力集中在时间演化上，可以假设暴胀场仅仅是时间 t 的函数，与空间坐标无关。代入FLRW度规，进而即有

$$\rho_I = \frac{1}{2}\dot{\phi}^2 + \mathcal{U}(\phi), \quad p_I = \frac{1}{2}\dot{\phi}^2 - \mathcal{U}(\phi). \quad (16.117)$$

不难看出，暴胀场的物态参数 $w_I = \frac{p_I}{\rho_I}$ 是在 -1 到 $+1$ 之间的，并且如果暴胀场的动能远远小于势能，即

$$\dot{\phi}^2 \ll \mathcal{U}(\phi). \quad (16.118)$$

那么就有 $w_I \approx -1$ 。从而这时候暴胀场就可能像宇宙学常数那样驱动宇宙指数膨胀。

进一步，把FLRW度规代入暴胀场作用量 S_I 的表达式，进而即有

$$S_I = \int \sqrt{h} d^3x dt a^3(t) \left[\frac{1}{2} \dot{\phi}^2 - \mathcal{U}(\phi) \right]. \quad (16.119)$$

式中 h 是空间度规 h_{ij} 的行列式。对这个作用量变分，就能得到暴胀场的运动方程

$$\ddot{\phi} + 3H\dot{\phi} + \mathcal{U}'(\phi) = 0, \quad (16.120)$$

式中 $\mathcal{U}'$ 表示函数 $\mathcal{U}$ 对自变量 ϕ 的导数。很显然，这个方程可以看成是描述一个沿着 ϕ 轴运动的粒子，它受到势能曲线 $\mathcal{U}(\phi)$ 的作用，方程中的 $3H\dot{\phi}$ 项相当于一种阻尼项，也即是说，宇宙的膨胀会阻碍暴胀场随时间的变化。

把 ρ_I 的表达式代入弗里德曼方程，并注意到暴胀场是主导的(从而可以忽略其它一些项)，即有

$$H^2 = \frac{8\pi G}{3} \left[\frac{1}{2} \dot{\phi}^2 + \mathcal{U}(\phi) \right]. \quad (16.121)$$

如果(16.118)式得到满足，那就有

$$H^2 \approx \frac{1}{3M_{pl}^2} \mathcal{U}(\phi). \quad (16.122)$$

式中 $M_{pl} = 1/\sqrt{8\pi G} = 2.4353 \times 10^{18} \text{GeV}$ ，称作约化的普朗克质量。很显然，只要 ϕ 场随时间变化得足够缓慢，从而使得 $\mathcal{U}(\phi) > 0$ 近似为常数，那宇宙近似就是指数膨胀的。为此我们进一步要求

$$\ddot{\phi} \ll \mathcal{U}'(\phi). \quad (16.123)$$

从而(16.120)就可以近似成

$$3H\dot{\phi} \simeq -\mathcal{U}'(\phi). \quad (16.124)$$

这时候 $\mathcal{U}(\phi)$ 可以看作是一个有效的宇宙学常数，只不过相比于今天观测到的有效宇宙学常数，它要大好多数量级，以至于可以在暴胀阶段占据了主导地位。

(16.118)式和(16.123)式合起来就是所谓的慢滚条件。如果慢滚条件满足，我们就可以利用(16.122)式和(16.124)式，进而有 $\dot{\phi}^2 \sim (\mathcal{U}'/H)^2 \sim M_{pl}^2 \mathcal{U}'^2 / \mathcal{U}$ ，进而即可以把条件 $\dot{\phi}^2 \ll \mathcal{U}$ 重写成 $M_{pl}^2 (\mathcal{U}'/\mathcal{U})^2 \ll 1$ 。进一步，既然 $\dot{\phi} \sim -M_{pl} \mathcal{U}'/\mathcal{U}^{1/2}$ ，则有 $\ddot{\phi} \sim M_{pl}^2 (\mathcal{U}'\mathcal{U}''/\mathcal{U}) - M_{pl}^2 (\mathcal{U}'^3/(2\mathcal{U}^2))$ ，又有 $M_{pl}^2 (\mathcal{U}'/\mathcal{U})^2 \ll 1$ ，从而条件 $\ddot{\phi} \ll \mathcal{U}'$ 相当于 $M_{pl}^2 \mathcal{U}''/\mathcal{U} \ll 1$ 。根据这些，习惯上常常定义

$$\varepsilon \equiv \frac{1}{2} M_{pl}^2 \left(\frac{\mathcal{U}'}{\mathcal{U}} \right)^2, \quad \eta \equiv M_{pl}^2 \frac{\mathcal{U}''}{\mathcal{U}}. \quad (16.125)$$

从而由慢滚条件即有 $\varepsilon \ll 1, |\eta| \ll 1$ ，这就是对暴胀场势能函数 $\mathcal{U}(\phi)$ 的一个限制。

前面说了，暴胀阶段的有效宇宙学常数很大，那它后来是怎么变成像今天观测到的那么小的呢？另一个相关的问题是，暴胀阶段是如何结束的呢？

这两个问题都可以通过选取合适的势能函数 $\mathcal{U}(\phi)$ 来进行解答。比如说，可以设想暴胀场的势能函数像图(16.4)中所示的那样，图中左边的平台区满足慢滚条件要求的 $\epsilon \ll 1, |\eta| \ll 1$ ，所以可以设想开始时暴胀场在这个平台区往下慢滚，从而驱使宇宙指数膨胀。但是，一旦暴胀场进入图中的势阱区域，慢滚条件就不满足了，暴胀也就结束了。尤其是暴胀场可以掉入阱底的极小值位置，这个位置对应的势能密度可能比平台区小好多个数量级，从而暴胀场掉到阱底的话，有效的宇宙学常数就会马上变小好多个数量级。

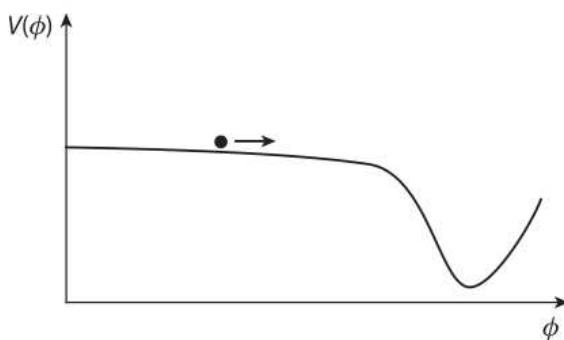


图 16.4: 暴胀场势能函数，图中的 $V(\phi)$ 即是正文中的 $\mathcal{U}(\phi)$ 。

不仅如此，暴胀场一旦掉入势阱，它原来的势能就会转化成动能，从而使得暴胀场在势阱底部来回振荡，这种振荡会耦合到辐射场以及其它物质粒子场上，从而使得振荡的动能转换成辐射场的激发，以及正反粒子对的产生，并且多余的暴胀场振荡能量还可以把辐射和物质粒子加热到很高的温度，从而开始热大爆炸。这个过程就是所谓的宇宙重加热过程，它解决了前面提到过的暴胀阶段和热大爆炸衔接的问题。

16.5 参考文献

关于什么是宇宙微波背景角功率谱，以及其第一声学峰如何确定宇宙

非常接近于平坦，最简短而清楚的介绍可以参见A. Zee, *Einstein Gravity in a Nutshell*, 章节VIII.3，或者也可以直接问人工智能，比如deepseek。

本章参考了俞允强《物理宇宙学讲义》以及梁灿彬《微分几何与广义相对论》的相关章节。尤其是俞允强老师的《物理宇宙学讲义》给了本章很大启发。