

# 初识量子力学

陈童

2024 年 8 月 3 日

# 目录

第一章 矩阵革命	2
1.1 玻尔的新观念	2
1.1.1 线索	3
1.1.2 玻尔的新观念	4
1.2 矩阵革命	10
1.2.1 跃迁元与跃迁元乘法	11
1.2.2 一些必要的数学	19
1.2.3 海森堡运动方程	24
1.2.4 基本对易关系	26
1.2.5 求解谐振子	30
1.2.6 符号的新发展	33

# 第一章 矩阵革命

物理学是大自然最深层次的秘密，大自然从不会直接把这些秘密暴露给人类，但它偶尔会透露给我们一些蛛丝般细微的线索。而人类，往往能够凭借其天才，从这些细微的线索中，半猜测半推断地揭开大自然部分最迷人的秘密。量子力学的建立就是一个最为成功的例子。

本章当然会涉及到量子力学在历史上的一些发展，但是，值得说明的是，我们并不是要还原历史，我们借助历史的一些印迹，目的却是为了更加自然地建立起整个量子力学的理论体系，因此本章中的历史是一种经过重构以后的“历史”，它只是一个故事，或者是历史的一种可能性，而不能完全等同于真实的历史，想了解真实历史的读者应该去阅读物理学史方面的著作。

## 1.1 玻尔的新观念

### 前玻尔时代

我们的故事实际上要从玻尔才正式开始，但是在历史上，量子观念的引进却并不是从玻尔开始的，因此为了讲清楚玻尔的故事，我们需要一些前玻尔时代的背景，尤其是需要介绍一下普朗克和爱因斯坦的开创性工作。将这两个伟大物理学家的开创工作归入前玻尔时代并不意味着它们不如玻尔重要，只是从本书的理论逻辑来说，玻尔所引入的观念与我们的关系更加密切一些。不过，为了突出本书的所需要的核心观念，我们决定把对普朗克和爱因斯坦相关工作的简单介绍穿插在对玻尔工作的具体讨论中，这里只是提请读者注意，它们实际是发生在玻尔的工作之前。

### 1.1.1 线索

19世纪末20世纪初，物理学家对原子的发光行为进行了大量研究。他们发现原子的光谱是分立的，也就是每个原子都只能发出和吸收一些特定波长(或者等价的，特定频率)的光。从经典理论的角度看来，这是不可思议的，因为按照经典的电磁学理论，核外电子围绕原子核运动就会发出电磁辐射，但是这些电磁辐射的频率(或者等价地，波长)是可以连续取值的，而不是仅仅只能取一些特定的分立值。因此，如何解释原子的分立光谱就成了当时的物理学家面临的一大难题。

而那时候人们还发现了原子光谱的另一个令人吃惊的规律，这个规律我们今天已经很少提到，但它对于物理学的发展也许是极其重要的，这规律就是里兹组合规则，这规则是说，人们总是可以合适地将原子光谱的每一条分立谱线都对应到一个正整数对 $(m, n)$ ，其中一个正整数用来标记谱线属于哪一个谱线系，另外一个整数用于指明它属于此系中的第几条谱线，人们发现，相应的谱线角频率 $\omega_{mn}$ 满足如下这条规则，

$$\omega_{mn} = \omega_{ml} + \omega_{ln}. \quad (1.1)$$

也就是说，假设有一条谱线角频率是 $\omega_{ln}$ ，另外还有一条谱线角频率是 $\omega_{ml}$ ，那么这两条谱线的组合 $\omega_{ml} + \omega_{ln}$ 也必定是一条谱线，对应于整数对 $(m, n)$ 。这一规律就是所谓的里兹组合规则。并且人们发现可以进一步定义

$$\omega_{nm} = -\omega_{mn}, \quad (1.2)$$

从而引入谱线间的减法，以使得里兹组合规则对于所有的正整数对指标都成立。

历史的发展表明，里兹组合规则的这两个公式(1.1)和(1.2)是最为强有力的线索，它在指引人们建立正确的量子理论上具有重要作用。当然，经典物理根本无法解释这两条规则。

经典物理的问题还不仅仅是解释不了原子的分立光谱和其里兹组合规则，它在逻辑上更大的问题是，按照经典物理，绕核运动的电子会不断地发出电磁辐射，电磁辐射本身是有能量的，因此核外电子就会不断地损失能量，这样它就会一边绕核运动一边逐步往原子核上掉，很快所有的核外电子都会完全掉落到原子核上。原子的大小是 $10^{-10}m$ 的量级，但是原子

核的大小只有 $10^{-15}m$ 左右，因此很快所有原子的尺寸都会坍缩5个数量级。我们的世界是由原子构成的，因此这也就意味着整个世界都会坍缩。这当然是不对的，可以说当时的物理学家面临的更大难题是，如何拯救世界。

这个拯救世界的人就是玻尔。但在讨论玻尔如何拯救了世界之前，我们还是先回到当时关于原子光谱的观测。当时关于各种原子的光谱有大量的观测数据，这些数据杂乱无章，除了里兹组合规则以外，没有人能理出更多头绪。但是不久之后，通过巴尔末和里德堡等人对观测数据的归纳整理，就最简单的原子，氢原子的光谱，人们逐渐理出了头绪。人们发现，所有氢原子的光谱线都可以写成下面这个漂亮的数学公式

$$\omega_{mn} = 2\pi Rc \left( \frac{1}{m^2} - \frac{1}{n^2} \right), \quad (1.3)$$

式中 $\omega_{mn}$ 是光谱线的角频率， $c$ 是光速， $R$ 是里德伯（Rydberg）常数，用氢原子光谱的观测数据可以定出它的值是 $1.097373177 \times 10^7 m^{-1}$ 。这么多光谱线竟然可以归纳出如此简单漂亮的一个公式，简直就是一个奇迹。但是，如何解释这个奇迹呢？

首先，不难发现，巴尔末和里德堡的这个公式自动满足里兹组合规则的两个公式(1.1)和(1.2)，因为它把谱线表达成了两项之差。实际上，一般性的，对于任何原子的任何谱线，为了满足里兹组合规则的两个公式(1.1)和(1.2)，其角频率都应该是两项之差，不妨写作

$$\omega_{mn} = T_n - T_m. \quad (1.4)$$

只不过，除了氢原子光谱有漂亮的巴尔末-里德堡公式以外，人们并没有找到其它原子 $T_n$ 的公式。不过，光谱总能够写成两项之差，这对于玻尔是一个巨大的启发。

正因为氢原子是最简单的原子，而且它的光谱满足上面那个奇迹般的公式。所以当时玻尔等物理学家想到，要从理论上解释原子光谱的数据，首先就应该研究氢原子，并推导出公式(1.3)。这就是玻尔的氢原子模型解决的问题。

### 1.1.2 玻尔的新观念

为了解释氢原子光谱，玻尔提出了三条假设，这三条假设就构成了玻尔的氢原子模型。首先，针对核外电子由于辐射损失能量而不能稳定运动，

要坍缩到原子核上的问题。玻尔提出了定态假设，玻尔说，考虑到当时刚刚兴起的量子物理以后，核外电子绕核运动的轨道不能是任意的，而是只能取某些特定的轨道，当电子在一个这样的特定轨道，比方说第 $n$ 个特定轨道上运动时，它是不辐射电磁波的，因此它的能量就不会损失，而是会保持在一个确定的值 $E_n$ ，核外电子在特定轨道上运动的这种状态就称之为定态。

定态是玻尔提出的第一个全新观念！定态之间没有连续过渡的轨道就意味着，定态的能量不是一个连续值，而是分立的 $E_n$ ，或者说，原子能量是量子化的，原子能量按照从小到大形成一阶一阶的能量“阶梯”，也称为能级。当然，定态的概念，以及定态能量量子化的推论不仅适用于氢原子，而且普遍适用于所有原子！只不过，对于氢原子，核外电子只有一个，而其它原子则有多个核外电子，因此这些电子的定态轨道就更加复杂。

其实，早在玻尔之前，普朗克就提出过能量分立的假设，这就是著名的普朗克能量量子化假设，它在历史上第一次引入了量子的概念。普朗克的假设是为了解释当时的黑体辐射难题，具体内容和本书关系不大，所以我们略去，只说结论。总之，通过对黑体辐射的研究人们发现，原子辐射和吸收角频率为 $\omega$ 的电磁波时，其行为就好像一个角频率为 $\omega$ 的线性谐振子，而普朗克假设，一个线性谐振子的能量是量子化的，其必须是某个基本能量的整数倍，这个基本能量是 $\hbar\omega$ ，其中 $\hbar$ 是一个常数，称作普朗克常数，其量纲是能量量纲乘以时间量纲，或者长度量纲乘以动量量纲(和角动量的量纲一样)，在国际单位制中，其值为 $1.05457266(63) \times 10^{-34} \text{J} \cdot \text{s}$ ，非常小。换言之，普朗克假设，线性谐振子的能级 $E_n$ 由下式给出

$$E_n = n\hbar\omega, \quad (1.5)$$

$n$ 为整数。注意，普朗克并非任意作出这个假设的，而是，他发现非如此假设不足以解释实验上测出来的黑体辐射谱。

普朗克常数是自然界一个最基本的常数，其出现是量子效应的典型特征，而这个常数在国际单位制中是如此之小也正解释了为什么通常的日常生活中人们感受不到量子效应的原因，毕竟，国际单位制就是适用于日常生活的单位制。不过，单位制的选取毕竟是人为的，因此如果人们要考察的是一个纯粹量子的世界，那选取合适的单位以使得 $\hbar = 1$ 也许就更方便。

这就好像如果人们要考察的是一个高速的相对论性世界，选取光速 $c = 1$ 的单位制更方便一样。

总之，玻尔的定态假设并不是完全凭空产生的，而是在普朗克能量量子化假设的基础上又往前走了一步，一般性地将能量的量子化归结于定态的存在。

与普朗克不同的是，玻尔并不是直接假设氢原子能级 $E_n$ 的具体表达式，毕竟氢原子能级比谐振子复杂好多，并不存在一个自然而又简单的假设。相反，玻尔是通过结合其它简单假设，最终推导出了 $E_n$ 的表达式。

既然核外电子在定态上并不辐射，那么，为什么我们能观测到原子的光谱呢？为此，玻尔引入了第二条假设，跃迁假设。玻尔说，核外电子处在定态上的时候虽然不辐射，但是，它可能从一个定态跃迁到另一个定态，比方说从第 $n$ 个定态跃迁到第 $m$ 个定态，这时候电子就会辐射出特定频率 $\omega_{mn}$ 的电磁波，按照爱因斯坦此前提出的光量子假设，这个辐射出来的光子的能量就是 $\hbar\omega_{mn}$ ，因此由能量守恒，玻尔写出了如下方程，

$$\hbar\omega_{mn} = E_n - E_m. \quad (1.6)$$

同样，玻尔的这第二条假设也是普遍适用于所有的原子，而且，它自然地解释了为什么原子光谱可以写成两项之差！

这里简单解释一下爱因斯坦的光量子假设，它是爱因斯坦对普朗克量子论的发展，它提出，电磁波本身是量子化的，每一个电磁波的量子(电磁波的基本组成部分)就是一个光子，也就是今天人们熟悉的光子。一个角频率为 $\omega$ 的光子，其能量为 $E = \hbar\omega$ 。爱因斯坦利用光量子假设成功地解释了光电效应，而且康普顿散射实验也验证了光量子的正确性。

玻尔的这两条假设一起，就自然地解释了原子光谱的分立性，原因在于定态能级是分立的，而且公式(1.6)进一步告诉我们，每一条光谱线的确都对应两个正整数 $m, n$ 。

至于原子为什么稳定，玻尔给出的最终解答是这样的，根据定态假设，原子的能量会量子化为能级，假设把这些能级按照能量从低到高排列，那么低能级的核外电子轨道必定更靠近原子核，因为从经典物理就可以知道，越靠近原子核的轨道能量就越低，主要原因是原子核与核外电子之间的库仑吸引能会负得越厉害，而如果核外电子掉到核上，那这个库仑能就是负无穷。但是，玻尔最终可以推导出来，总是有一个最低能级，其能量是有

限的，也就是说，这个最低能级对应的定态轨道并不是核外电子落到原子核上。因此，当核外电子发出辐射从高能级往低能级跃迁的时候，它最多跃迁到最低能级就不可能接着往下跃迁了，因此，核外电子永远也不会掉到原子核上去，原子是稳定的！

但是，如果想进一步推导出氢原子光谱的巴尔末-里德堡公式(1.3)，玻尔就必须进一步找到定态能量 $E_n$ 的表达式。类比于行星绕太阳运动，玻尔知道决定核外电子轨道的是能量和角动量这两个守恒量，量纲分析又告诉他，考虑到量子效应，必然有 $J \sim \hbar$ 。因此玻尔又引入了第三条假设，称之为轨道角动量量子化假设，玻尔说，处在定态 $n$ 上的电子的轨道角动量 $J$ 是量子化的，它满足

$$J = rp = n\hbar. \quad (1.7)$$

式中 $r$ 是核外电子离核的距离， $p$ 是核外电子的动量。注意，这第三条假设是特殊的，它只适用于氢原子，因为其它原子的核外电子轨道涉及到多个电子。

有了这三条假设，玻尔就能导出氢原子定态能量 $E_n$ 了。玻尔的推导大致如下。首先为了简单起见，我们不妨假定核外电子绕核运动的轨道是圆轨道。因此按照牛顿定律，我们有

$$m_e \frac{v^2}{r} = \frac{e^2}{4\pi\epsilon_0 r^2}, \quad (1.8)$$

式中 $m_e$ 为电子质量， $e$ 为电子电荷， $\epsilon_0$ 为真空的介电常数。不妨引入如下无量纲常数 $\alpha$

$$\alpha = \frac{e^2}{4\pi\epsilon_0 \hbar c}, \quad (1.9)$$

称作精细结构常数，它的数值大约是1/137。从而，公式(1.8)也可以重写成，

$$m_e v^2 = \frac{\alpha \hbar}{r}. \quad (1.10)$$

另外，由角动量量子化假设我们有

$$rm_e v = n\hbar. \quad (1.11)$$



由这两个结果我们容易推导出，定态的轨道半径 $r$ 必然满足

$$r = n^2 \frac{\hbar}{m_e \alpha c}. \quad (1.12)$$

特别的，离原子核最近的第1个定态的轨道半径为 $\frac{\hbar}{m_e \alpha c} = a_0$ ， $a_0$ 称之为玻尔半径，可以算得它的数值是 $0.52917721067(12) \times 10^{-10} m$ 。

另一方面，氢原子的能量是动能和势能的和，即 $E = \frac{1}{2}m_e v^2 - \frac{e^2}{4\pi\epsilon_0 r} = \frac{1}{2}m_e v^2 - \frac{\alpha\hbar}{r}$ ，由式(1.10)就可以得到， $E = -\frac{1}{2}\frac{\alpha\hbar}{r}$ （当然它也等于 $-\frac{1}{2}m_e v^2$ ）。将定态轨道半径的结果(1.12)代入这个能量的式子，我们就能得到第 $n$ 个定态能量 $E_n$ 的表达式

$$E_n = -\frac{1}{2}m_e \alpha^2 c^2 \frac{1}{n^2} = \frac{E_1}{n^2}. \quad (1.13)$$

式中 $E_1 = -\frac{1}{2}m_e \alpha^2 c^2$ 就是氢原子第1个定态的能量，它也就是氢原子最低能态(也称之为基态)的能量，代入相应物理量的值就可以算得 $E_1 = -13.6 eV$ 。 $E_1 = -\frac{1}{2}m_e \alpha^2 c^2$ 的这个结果很好记，由于氢原子的总能量等于 $-\frac{1}{2}m_e v^2$ ，而在基态的时候电子的速度 $v = \alpha c$ ，因此就有 $E_1 = -\frac{1}{2}m_e \alpha^2 c^2$ 。

更令人高兴的是，将公式(1.13)代入跃迁假设的公式(1.6)，玻尔就能导出氢原子的光谱，

$$\omega_{mn} = \frac{m_e \alpha^2 c^2}{2\hbar} \left( \frac{1}{m^2} - \frac{1}{n^2} \right), \quad (1.14)$$

这和神奇的巴尔末里德堡公式(1.3)形式完全一样，将它和巴尔末里德堡公式比较我们就可以得到，如果里德堡常数可以由下式给出

$$R = \frac{m_e \alpha^2 c}{4\pi\hbar}, \quad (1.15)$$

那么玻尔就正确地推导出了氢原子的光谱。代入物理常数的值人们容易算出公式(1.15)给出的里德堡常数和由光谱数据定出来的里德堡常数吻合得非常好，这充分说明了玻尔的确解释了氢原子的光谱。

正如我们已经看到的，为了解决氢原子的难题，玻尔引入了一些革命性的新概念，但也可以看出玻尔理论的缺陷，在玻尔推导定态能量 $E_n$ 表达式的过程中，他还要求助于经典力学的牛顿方程，而他也没有用一个新的基本原理将牛顿方程和他新引入的三条假设统一起来。换言之，玻尔并没有证明牛顿方程和他的三条假设是相容的，相反，正如前面所说的，牛

顿方程和经典的电磁学理论结合起来会导致完全错误的结论，而玻尔的四条假设正是为了挽救这样的错误而引入的，因此这些假设实际上天然地与牛顿方程不相容。因此，玻尔的理论虽然取得了巨大的成功，但在理论结构上依然是不令人满意的，它遗留的问题几乎和它解决的问题一样大。

不过，玻尔关于定态和定态跃迁的全新观念将成为量子力学最为基本性的观念，也是本书的核心概念，值得我们作一些进一步的阐述。定态是核外电子的特定轨道，海森堡后来指出，电子轨道不可以观察，可观察的是光谱。所以，更恰当的看法是，**定态就是原子的一种确定可能性**，在这种可能性中，原子有一个分立的确定能量值。

要描述一个原子，只需要知道它的所有定态，因此，原子所有定态的集合就构成了原子的一种**可能性完备集**。由于原子存在一个最低能级，所以我们总是可以把不同定态按照能量从低到高排列，使 $E_1 \leq E_2 \leq E_3 \leq \dots$ ，相应的不同定态可能性就可以由正整数 $n$ 来标记，它相应于第 $n$ 能级。不妨将原子的第 $n$ 个定态可能性记作 $(n)$ ，从而原子的定态可能性完备集就是集合 $\{(n)|n = 1, 2, \dots\}$ 。

不过，对于氢原子，后来完整的量子力学理论表明，标记不同定态可能性的更合适方法是同时用四个数，除了决定定态能量的 $n$ 以外，还需要 $l, m, m_s$ 这三个数，其中 $l$ 和 $m$ 都是整数，而 $m_s = \pm 1/2$ ，所以，更合适的标记氢原子某个定态可能性的方法是将之记作 $(nlmm_s)$ ，当然，这个细节性的问题其实和建立量子力学的基本原理无关，只和它在原子物理中的应用有关，所以我们就不进一步阐述了。

下面讨论定态跃迁。为了更直观地阐述跃迁过程，我们以氢原子为例，画出示意图(1.1)，图中水平方向表示态可能性的空间，竖直方向代表时间，两条竖直直线分别代表定态可能性 $(n)$ 和定态可能性 $(m)$ 的电子，波浪线代表辐射出来的光子，这个图示意的就是电子从高能级 $(n)$ 跃迁到低能级 $(m)$ ，并辐射出一个光子的过程。

后来人们通过对原子发光过程的进一步研究指出，原子什么时候跃迁，跃迁到哪个定态上去，都完全是随机的。因此，**定态跃迁原则上就是随机的**，只能讨论跃迁的概率！这一点其实非常重要，虽然前文我们没怎么强调，但它是量子物理不同于经典物理的一个核心特征，**即在量子物理里面，我们只能讨论概率**，后来这成为玻恩的一个洞察，也是他得诺贝尔奖

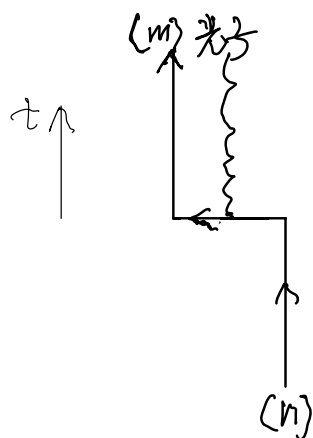


图 1.1: 水平方向表示态可能性的空间, 竖直方向代表时间, 两条直线分别代表定态可能性( $n$ )和定态可能性( $m$ )的电子, 波浪线代表辐射出来的光子。

的原因。

## 1.2 矩阵革命

差不多是夜里三点钟, 计算结果最终出来了。我深深地被震惊了。我很兴奋, 一点也不想睡。于是, 我离开房间, 坐在一块岩石上等日出。

—————海森堡

### 线索

在玻尔提出它的原子模型之后, 人们就已经完全理解里兹组合规则了,

$$\omega_{mn} = \omega_{ml} + \omega_{ln}, \quad \omega_{nm} = -\omega_{mn}. \quad (1.16)$$

之所以有这样的规则成立, 原因很简单, 因为根据玻尔的定态跃迁假设, 原子从第 $n$ 个定态跃迁到第 $m$ 个定态, 辐射(或者吸收)的光谱频率为 $\omega_{mn} = (E_n - E_m)/\hbar$ , 其中 $E_n$ 就是第 $n$ 个定态的能量, 如此一来当然就有里兹组合规则了。对里兹组合规则(1.16)的解释很简单, 那就是核外电子

从 $n$ 态跃迁到 $m$ 态所放出的能量，等于它从 $n$ 态跃迁到中间 $l$ 态所放出的能量加上它从中间 $l$ 态跃迁到 $m$ 态放出的能量。问题是，给定初始的 $n$ 态和末尾的 $m$ 态，里兹组合规则中的中间 $l$ 态可以是任意的，里兹组合规则对 $l$ 没有任何限制！

### 1.2.1 跃迁元与跃迁元乘法

故事的下一个主角是海森堡，对于当时的海森堡来说，玻尔氢原子模型中所涉及到的核外电子轨道概念是没有道理的，因为当时的海森堡在物理学哲学上深受玻尔的影响，他坚持物理学应该用可观测量建立起来，或者换句话说，他认为物理学理论中所用到的量应该都是可观测的。我们这里不讨论这一哲学有没有道理，但是当时年轻的海森堡深信这一哲学，因此他觉得核外电子轨道的概念没有道理，因为不可观测。对于原子物理来说，可观测的是什么呢，是光谱，是谱线频率和谱线的强度，不是电子轨道，电子轨道不过是对太阳系模型的一种类比，而不是原子物理的观测事实。因此海森堡觉得关键是要研究谱线频率和谱线强度。但是海森堡也知道，原子光谱的分立性和里兹组合规则意味着，虽然核外电子轨道的概念没有道理，但是玻尔的定态和定态跃迁假设必定是对的。

以上也许就是导致海森堡产生他的原创性新思想的一些基本事实，以及他从玻尔那里继承的革命性新观念。当然，还有一个具体的物理学结论对于海森堡验证他的想法非常重要，那就是通过普朗克对黑体辐射公式的推导，海森堡知道，一个线性谐振子的能量是量子化的，并且量子化为 $n\hbar\omega$ 的形式， $n$ 就是线性谐振子的第 $n$ 个能级。海森堡当然也知道，玻尔的定态和定态跃迁概念不只是适用于氢原子，也同样适用于一个可以发出电磁辐射的线性谐振子。其实，很可能在海森堡看来，定态和定态跃迁是普适性的物理学观念。

前面说过，海森堡坚持原子物理应该回到可观测量，也就是谱线频率和谱线强度。玻尔的原子模型当然对谱线频率的解释很不错，尤其是完全推导出了氢原子的所有光谱线。但是玻尔氢原子模型并没有涉及谱线的强度，因此海森堡想推进当时的物理学，那就要研究谱线的强度。注意到定态跃迁是**随机的**，只能谈论跃迁概率，而核外电子从 $n$ 态跃迁到 $m$ 态时所发出的这条谱线的强度当然取决于单位时间之内它从 $n$ 态自发跃迁到 $m$ 态的

概率 $A_{mn}$ ，而这是一个当时的物理学家(爱因斯坦)已经提出来的物理量，只是当时还没有人会计算这个量，海森堡要做的，就是研究如何计算出这个量。但是这个量涉及到原子和电磁场的相互作用，而当时对电磁场的量子理论还只是知道一个光量子假设再加上爱因斯坦的受激辐射理论，对于原子和电磁场如何相互作用的量子理论还一无所知，毕竟，量子力学的正确理论还正有赖于海森堡目前正在进行的这一工作呢。唯一知道的是，辐射过程涉及的原子和电磁波的相互作用对于原本的原子而言是一个微扰。因此海森堡只能转而去经典理论是如何计算电磁波辐射的，通过经典电动力学我们知道，一个加速运动的电子会辐射电磁波，电动力学也推导出了其辐射功率为

$$P = \frac{e^2}{3\pi\epsilon_0 c^3} |\ddot{\mathbf{x}}|^2, \quad (1.17)$$

式中 $\ddot{\mathbf{x}}$ 就是电子的加速度。还没有学过电动力学从而不知道这个公式也没有关系，把加速运动的电子会辐射出电磁波当成经典物理的一个基本事实来接受就可以了。

但是，海森堡知道这个经典公式(1.17)肯定是不对的，不仅因为它的表达式中涉及到了核外电子轨道 $\mathbf{x}(t)$ 这样一个在海森堡看来无法观测的量，更重要的是，它和定态以及定态跃迁的概念是相矛盾的。当然，玻尔引入定态和定态跃迁的原因之一就是为了避免由于(1.17)这一经典公式而导致的原子稳定性问题。因此海森堡首先要做的就是改造这个经典公式(1.17)，使得它和定态以及定态跃迁的概念相容。

根据定态跃迁假设，核外电子只在从一个定态 $n$ 跃迁到另一个定态 $m$ 时才辐射，注意，跃迁总是涉及到两个定态，初态 $n$ 和末态 $m$ 。因此海森堡想到，经典意义上的 $\mathbf{x}(t)$ 是不可观测的，是没有意义的，取而代之的应该是一个类似于 $[\mathbf{x}]_{mn} \equiv [\mathbf{x}_0]_{mn} e^{-i\omega_{mn}t}$ 这样的涉及两个定态 $m, n$ 的量。其中 $[\mathbf{x}_0]_{mn}$ 与时间无关，而海森堡将这个量对时间的依赖因子写成 $e^{-i\omega_{mn}t}$ 的原因在于，通过普朗克对黑体辐射公式的推导，海森堡知道，原子定态跃迁发出电磁辐射的过程可以看成是一种角频率为 $\omega_{mn}$ 的简谐振动，而一个作简谐振动的物理量，其时间依赖就是这样的。至于写成指数形式而不是经典物理中更常见的只保留实部的 $\cos(\omega_{mn}t)$ 形式的原因是，在经典物理中，根据欧拉公式，实的 $\cos$ 形式总可以写成指数形式与其复共轭的和，而海森堡认为在量子对应中，这两部分(指数及其复共轭)是相互独立的，海

森堡将定态 $n$ 到定态 $m$ 的跃迁对应到了 $[\mathbf{x}_0]_{mn}e^{-i\omega_{mn}t}$ ，其复共轭部分则对应到了反向的定态 $m$ 到定态 $n$ 的跃迁(即后面的厄米条件)，当然，代价就是引入了复数，而物理量的实数性就是通过额外规定厄米条件补回来的。更何况，正如后文会看到的，当将两个物理量相乘时，如果这两个量对时间的依赖因子是 $\cos$ 形式，那是不可能满足里兹组合规则的，而写成指数形式则很轻易就可以满足这一点。由于相关的量变成了复数，所以这时候(1.17)式中的矢量模长 $||$ 当然就必须同时是求复数模长，因为最终物理上可观测的辐射功率当然是实数。总之，海森堡将 $[\mathbf{x}_0]_{mn}e^{-i\omega_{mn}t}$ 代入公式(1.17)就得到

$$P(n \rightarrow m) = \frac{e^2 \omega_{mn}^4}{3\pi \epsilon_0 c^3} |[\mathbf{x}]_{mn}|^2. \quad (1.18)$$

注意，现在的这个新公式(1.18)不仅用到了定态跃迁的概念，它和定态假设也是相容的，由于 $\omega_{nn} = 0$ ，因此就有 $P(n \rightarrow n) = 0$ ，这当然与定态假设一致(指定态上不辐射)。经典的轨道 $\mathbf{x}(t)$ 不可观测，但是现在的 $[\mathbf{x}]_{mn}$ 是可以观测的，这是因为，从 $n$ 态跃迁到 $m$ 态的辐射功率等于单位时间辐射概率 $A_{mn}$ 乘以辐射出来的光子能量 $\hbar\omega_{mn}$ ，因此由(1.18)我们就可以得到 $A_{mn}$ 的表达式， $A_{mn} \propto |[\mathbf{x}]_{mn}|^2$ 。 $A_{mn}$ 当然可以观测，因此 $[\mathbf{x}]_{mn}$ (严格来说是其模 $|[\mathbf{x}]_{mn}|$ )就可以观测。另外，前面说过，辐射本身对原子是一个微扰，很显然这个微扰与 $[\mathbf{x}]_{mn}$ 密切相关。

当然，由于原子自发跃迁是从高能级跃迁到低能级，所以 $[\mathbf{x}]_{mn} = [\mathbf{x}_0]_{mn}e^{-i\omega_{mn}t}$ 这个量仅当 $E_n \geq E_m$ 时(即 $n \geq m$ 时)才有定义。不过，海森堡注意到 $e^{-i\omega_{nm}t} = e^{i\omega_{mn}t} = [e^{-i\omega_{mn}t}]^*$ ，所以他进一步规定

$$[\mathbf{x}_0]_{nm} = [\mathbf{x}_0]_{mn}^*. \quad (1.19)$$

这样一来， $[\mathbf{x}]_{mn}$ (也就是 $[\mathbf{x}_0]_{mn}e^{-i\omega_{mn}t}$ )就对任意的正整数指标 $n, m$ 均有定义了。而海森堡规定的条件(称作厄米条件)就成为

$$[\mathbf{x}_0]_{nm}e^{-i\omega_{nm}t} = ([\mathbf{x}_0]_{mn}e^{-i\omega_{mn}t})^*, \quad (1.20)$$

也就是

$$[\mathbf{x}]_{nm} = [\mathbf{x}]_{mn}^*. \quad (1.21)$$

也就是说，反向的 $m$ 到 $n$ 的跃迁等于正向 $n$ 到 $m$ 跃迁的复共轭。正如前面提过的，这条件对应于经典物理量的实数性。

根据前文的讨论，在一个正比于物理量 $\mathbf{x}$ 的微扰的影响下(这个微扰就是核外电子与电磁场之间的相互作用)，原子从定态 $n$ 跃迁到定态 $m$ 的概率(即前面的 $A_{mn}$ )将与 $|\mathbf{x}_{mn}|^2$ 成比例，这就是 $\mathbf{x}_{mn}$ 这个量的物理诠释。

这样一来，对应经典物理里面的位置坐标 $\mathbf{x}(t)$ ，海森堡就引入了一个新的物理量 $\mathbf{x}_{mn}$ 。在这个物理量中，如果给定 $m$ 和 $n$ ，那 $\mathbf{x}_{mn}$ 就是一个复数，但是，指标 $m, n$ 也可以看作变量，跑遍所有正整数，所以，这时候 $\mathbf{x}_{mn}$ 就是一张有无穷行无穷列的数据表，排列出来就像下面这样

$$\mathbf{x}_{mn} = \begin{bmatrix} \mathbf{x}_{11} & \mathbf{x}_{12} & \mathbf{x}_{13} & \cdots \\ \mathbf{x}_{21} & \mathbf{x}_{22} & \mathbf{x}_{23} & \cdots \\ \mathbf{x}_{31} & \mathbf{x}_{32} & \mathbf{x}_{33} & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix} \quad (1.22)$$

这个数据表可以向右方和下方无限延伸。 $\mathbf{x}_{mn}$ 的两个指标 $mn$ 中，左边的指标 $m$ 用来指示表中的元素处于第几行，因此称作行指标，而右边的指标 $n$ 用来指示元素位于第几列，因此称作列指标。

读者不应该混淆符号 $\mathbf{x}_{mn}$ 的两种含义，当指标 $m, n$ 看作可变时，它代表一整张数据表，而当指标 $m, n$ 看作固定时，它又只代表这表中第 $m$ 行第 $n$ 列的元素，从而是一个复数。具体应该怎么看待这两个指标则应该根据上下文。

请想象一张无穷无尽的数据表，任何纸张都是有限的，都写不下它，所以只能用省略号代之！而将位置坐标这样的概念变成一张这样的数据表，这本身就蕴含了最高的革命性。数学上，这样的数据表有一个专门的名称，叫做矩阵。不过，物理上，由于海森堡是在定态跃迁思想的指引下引入 $\mathbf{x}_{mn}$ 的，所以我们不妨称之为跃迁元。

有时候，我们也示意性地将 $\mathbf{x}_{mn}$ 记成如下样子，

$$\mathbf{x}_{mn} = \left\{ (m) \overset{\mathbf{x}}{\leftarrow} (n) \right\}. \quad (1.23)$$

其中 $(m) \leftarrow (n)$ 表示从定态 $n$ 到定态 $m$ 的跃迁，箭头上方的 $\mathbf{x}$ 表示跃迁是在物理量 $\mathbf{x}$ 的作用下产生的，而外面的花括号则表示取这个跃迁的跃迁元。

而海森堡规定的条件(1.21)用矩阵的形式来写就成了,

$$\begin{bmatrix} [\mathbf{x}]_{11} & [\mathbf{x}]_{12} & [\mathbf{x}]_{13} & \cdots \\ [\mathbf{x}]_{21} & [\mathbf{x}]_{22} & [\mathbf{x}]_{23} & \cdots \\ [\mathbf{x}]_{31} & [\mathbf{x}]_{32} & [\mathbf{x}]_{33} & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix} = \begin{bmatrix} [\mathbf{x}]_{11}^* & [\mathbf{x}]_{21}^* & [\mathbf{x}]_{31}^* & \cdots \\ [\mathbf{x}]_{12}^* & [\mathbf{x}]_{22}^* & [\mathbf{x}]_{32}^* & \cdots \\ [\mathbf{x}]_{13}^* & [\mathbf{x}]_{23}^* & [\mathbf{x}]_{33}^* & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix}.$$

也就是说, 这样的矩阵满足一个特殊的性质, 即将其行与列的相应元素对调(第1行与第1列对调, 第2行与第2列对调, 依次类推), 同时把每一个对调之后的元素换成其复数共轭, 结果得到的矩阵等于原来的矩阵。数学上称这样的矩阵为厄米矩阵!

习惯上也将整张跃迁元数据表 $[\mathbf{x}(t)]_{nm}$ 简记作 $\mathbf{x}(t)$ , 也即是说

$$\mathbf{x}(t) \equiv [\mathbf{x}(t)]_{nm} \equiv [\mathbf{x}_0]_{nm} e^{-i\omega_{nm}t}. \quad (1.24)$$

在这个记号中 $\mathbf{x}(t)$ 就不再表示经典的位置矢量了, 而是表示与之相应的整张跃迁元数据表。

当我们有了位置物理量 $[\mathbf{x}(t)]_{mn} = [\mathbf{x}_0]_{mn} e^{-i\omega_{mn}t}$ 以后, 将它对时间求导当然就得到速度 $[\mathbf{v}]_{mn} = -i\omega_{mn}[\mathbf{x}_0]_{mn} e^{-i\omega_{mn}t}$ , 进而就有动量 $[\mathbf{p}]_{mn} = m_e[\mathbf{v}]_{mn} = -im_e\omega_{mn}[\mathbf{x}_0]_{mn} e^{-i\omega_{mn}t}$ , 这里 $m_e$ 表示电子质量, 而且不难验证

$$[\mathbf{p}]_{nm} = [\mathbf{p}]_{mn}^*. \quad (1.25)$$

所以, 动量也变成了一张跃迁元数据表, 作为矩阵也是一个厄米矩阵。通常也将动量的跃迁元数据表简记如下

$$\mathbf{p}(t) \equiv [\mathbf{p}]_{nm} \equiv [\mathbf{p}_0]_{nm} e^{-i\omega_{nm}t}. \quad (1.26)$$

同样, 在这种记号中 $\mathbf{p}(t)$ 不再代表经典的动量变量, 而是代表与之相应的跃迁元数据表。

在海森堡想来, 一切物理量都应该变成跃迁元数据表, 只有跃迁元才是可观测的。由于任何物理量均是位置和动量的函数, 所以, 有了位置的跃迁元数据表和动量的跃迁元数据表以后, 原则上应该就可以构造出其它一切物理量的跃迁元数据表。由于定态跃迁可以看成是一种谐振动, 海森堡知道所有这些跃迁元应该都具有 $[A]_{mn} \equiv [A_0]_{mn} e^{-i\omega_{mn}t}$ 这样的形式, 式



中 $[A_0]_{mn}$ 与时间无关。当然，作为对经典物理量实数性的量子对应，这些跃迁元数据表也要满足类似的海森堡厄米条件

$$[A]_{mn} = [A]_{nm}^*. \quad (1.27)$$

从而作为矩阵也是厄米矩阵。和前面类似，厄米条件的含义是，反向的从 $n$ 到 $m$ 的跃迁元，是正向从 $m$ 到 $n$ 跃迁元的复数共轭，也可以示意性地记为

$$\left\{ (m) \xleftarrow{A} (n) \right\} = \left\{ (m) \xrightarrow{A} (n) \right\}^*. \quad (1.28)$$

而且，完全类似于前面对 $[\mathbf{x}]_{mn}$ 的物理诠释， $[A]_{mn}$ 的物理诠释即是：如果设法引进一个正比于物理量 $A(t)$ 的微扰相互作用，使之引起原子系统发生定态跃迁，则原子从定态 $n$ 跃迁到定态 $m$ 的概率将与 $|[A]_{mn}|^2$ 成比例。

值得说明的是，习惯上常常将整个跃迁元数据表 $[A]_{mn}$ 简记作 $A$ ，即

$$A \equiv [A]_{mn} \equiv [A_0]_{mn} e^{-i\omega_{mn}t}. \quad (1.29)$$

在这种记号规则中，同一个记号 $A$ 既代表物理量，又代表与这个物理量相应的跃迁元数据表(或者说矩阵)。

对于一个矩阵 $M$ ，当将其行与列的相应元素对调(第1行与第1列对调，第2行与第2列对调，依次类推)，同时把每一个对调之后的元素换成其复数共轭，我们就得到一个新的矩阵，称作 $M$ 的厄米共轭矩阵，记作 $M^\dagger$ ，所以

$$[M^\dagger]_{mn} = [M]_{nm}^*. \quad (1.30)$$

不难看出，厄米共轭再厄米共轭，一定等于本身，即

$$(M^\dagger)^\dagger = M. \quad (1.31)$$

有了厄米共轭的概念以后，物理量 $A$ 的跃迁元数据表需要满足的海森堡条件(1.27)，就可以重记为

$$A = A^\dagger, \quad (1.32)$$

也即是说，厄米矩阵就是厄米共轭以后等于其自身的矩阵。

但是，这里有一个重要的问题，比方说动能 $T$ ，它在经典物理里面是这样的 $T = \frac{\mathbf{p}^2}{2m_e}$ ，这里涉及到动量的平方。更一般地，假设物理量 $C$ 等于两个物理量 $A$ 和 $B$ 的乘积，即 $C = AB$ 。那么根据海森堡的跃迁元思想，就需要将 $C$ 的跃迁元 $[C]_{mn} \equiv [C_0]_{mn}e^{-i\omega_{mn}t}$ ，用 $A$ 的跃迁元 $[A]_{mn} \equiv [A_0]_{mn}e^{-i\omega_{mn}t}$ 与 $B$ 的跃迁元 $[B]_{mn} \equiv [B_0]_{mn}e^{-i\omega_{mn}t}$ 的乘积的形式表达出来。但现在 $A$ 和 $B$ 的所有跃迁元都分别构成了一张无穷行无穷列的数据表，两张这样的数据表该怎么乘，就是海森堡需要解决的大难题。

当然，学过线性代数的读者都知道(我们并不假定读者学过线性代数)，跃迁元数据表作为矩阵，可以按照标准的矩阵乘法来乘。但是，矩阵乘法只是数学家规定的一种运算，而现在的跃迁元数据表却是物理，因此物理上，你怎么知道跃迁元数据表应该按照矩阵乘法来乘呢？毕竟，两个数据表之间数据相乘的可能性可太多了，为什么一定要按照矩阵乘法来乘呢？更何况，在海森堡的时代，和今天不同，那时候物理学家并不学线性代数，所以海森堡根本就不知道存在矩阵这样的数学对象，更不知道矩阵乘法。所以，对于海森堡而言，他要处理的问题就是，给两个跃迁元数据表定义一种合适的乘法，使得它能够正确地把两个物理量乘在一块。

正是在这里，里兹组合规则(1.16)，或者等价地说玻尔的公式 $\omega_{mn} = (E_n - E_m)/\hbar$ 给了海森堡必要的指引。由定义， $[C_0]_{mn}e^{-i\omega_{mn}t} = [(AB)_0]_{mn}e^{-i\omega_{mn}t}$ ，而按照里兹组合规则， $\omega_{mn}$ 总是可以分裂成两部分的组合， $\omega_{mn} = \omega_{ml} + \omega_{ln}$ ，所以海森堡想到可以令 $[(AB)_0]_{mn}e^{-i\omega_{mn}t} = [A_0]_{ml}e^{-i\omega_{ml}t}[B_0]_{ln}e^{-i\omega_{ln}t}$ (从这里可以清楚地看到，如果不是将跃迁元数据表对时间的依赖写成 $e^{-i\omega_{mn}t}$ ，而是写成仅仅只有实部的 $\cos$ 形式，那就无法满足里兹组合规则了)，也就是说，应该有 $[(AB)_0]_{mn} = [A_0]_{ml}[B_0]_{ln}$ ，或者相应的 $[AB]_{mn} = [A]_{ml}[B]_{ln}$ ，看来问题就解决了。但是海森堡最天才的地方就在这里，他注意到里兹组合规则对中间的 $l$ 态没有任何限制，也就是说 $l$ 有无穷多种可能性，怎么对付这无穷多种可能性呢？海森堡的处理办法很简单，他说，既然我们不能确定中间的可能性 $l$ 具体是哪个，那就应该平等地对待每一种可能性，因此最自然的处理办法就是把所有的可能性都加起来！因此海森堡给出

$$[C]_{mn} = [AB]_{mn} = \sum_l [A]_{ml}[B]_{ln}, \quad (1.33)$$

式中的求和是对所有的正整数 $l$ 求和。我们可以示意性地把这种乘法记成如

下样子

$$\{(m) \xleftarrow{AB} (n)\} = \sum_l \{(m) \xleftarrow{A} (l) \xleftarrow{B} (n)\}. \quad (1.34)$$

当然学过线性代数的读者容易知道，这的确就是标准的矩阵乘法，但是海森堡并没有学过矩阵乘法，对他来说，之所以要这么乘，是因为有无穷多种中间可能性都满足要求，我们需要把所有的中间可能性都加起来！今天看来，这是一个极其原创性的想法，因为它把握住了量子力学相干叠加的本质！

一般来说，两个物理量相乘，结果当然是一个新的物理量。但是，有一种特别重要的物理量例外，那就是单位1，1和任何物理量 $A$ 相乘，结果当然还是 $A$ 。因此，相对应的，1的跃迁元数据表和任何跃迁元数据表 $A$ 相乘，结果必须还是数据表 $A$ 。按照我们的记号规则，1的跃迁元数据表应该记作 $[1]_{mn}$ ，其物理含义就是在不受任何微扰作用(1的意思就是不受任何微扰)的情况下，系统自发从 $n$ 态跃迁到 $m$ 态的跃迁元。因此，对于任何 $A$ ，均有

$$[A]_{mn} = \sum_l [1]_{ml} [A]_{ln}. \quad (1.35)$$

不难验证满足这个条件的 $[1]_{mn}$ 必定为

$$[1]_{mn} = \delta_{mn}, \quad (1.36)$$

式中 $\delta_{mn}$ 称作克罗内克符号，它的定义就是当 $m = n$ 时取1，否则都取0，从而(1.35)式对 $l$ 的求和中，只有一项有非零贡献，就是 $l = m$ 的那项，这时候 $[1]_{ml}$ 当然取1，而 $[A]_{ln}$ 就成了 $[A]_{mn}$ ，从而(1.35)式右边等于左边。 $\delta_{mn}$ 排成矩阵的话就是

$$\delta_{mn} = \begin{bmatrix} 1 & 0 & 0 & \dots \\ 0 & 1 & 0 & \dots \\ 0 & 0 & 1 & \dots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix}. \quad (1.37)$$

这种矩阵称作单位矩阵，它只在对角线上的元素为1，其它元素都是0。单位矩阵乘以任何矩阵 $A$ 都还等于 $A$ 。人们也常常简单地把单位矩阵就记作1，

根据上下文不难把它和数字1相区分开来。不难验证, 对于任何矩阵 $A$ , 不仅 $1 \cdot A = A$ , 而且也有 $A \cdot 1 = A$ 。

根据跃迁元的概率诠释,  $|[1]_{mn}|^2$ 应该正比于系统不受任何微扰作用时自发从 $n$ 态跃迁到 $m$ 态的概率, 但是,  $[1]_{mn}$ 不是等于0就是等于1, 无论是0还是1, 其模方都是本身。因此这就说明 $[1]_{mn} = \delta_{mn}$ 本身就正比于系统不受任何微扰作用时自发从 $n$ 跃迁到 $m$ 的概率。又由于给定初态 $n$ , 这种跃迁只有一种可能性, 即仅当末态 $m = n$ 时, 才有 $[1]_{mn}$ 等于1, 其它 $m \neq n$ 的情况都为零。这说明, 当系统不受任何微扰作用时, 给定它的一个初态 $n$ , 你去测量它末态能跃迁到所有定态可能性中的哪个, 则会发现它只能百分之百从 $n$ 跃迁到 $n$ 自身, 而跃迁到一个与初态不同的 $m \neq n$ 的末态的概率为零。当然,  $n$ 态跃迁到 $n$ 态自身这种说法, 实际上已经稍微拓宽玻尔的定态跃迁概念了。之所以作这种拓展, 是因为它会带来语言表达上的适当简化。

以上讨论了海森堡引入的新思想, 海森堡的想法是如此自然, 却又如此新鲜大胆, 问题是, 这想法起作用吗? 这个问题我们稍后再来解决, 稍后我们将通过用海森堡的想法推导出线性谐振子的能级(就是普朗克能量量子化假设的那个), 进而验证这想法的确是起作用的。现在, 让我们先发展一点必要的数学, 已经熟悉相关数学的读者可以略过, 或者只需简单浏览一下这个数学部分。

### 1.2.2 一些必要的数学

可能有些读者并没有学过线性代数, 因此并不熟悉海森堡所制定的这种奇怪乘法规则

$$[AB]_{mn} = \sum_l [A]_{ml}[B]_{ln}, \quad (1.38)$$

所以我们对其作进一步的阐述。首先, 看这个式子各项涉及的指标的时候, 都是从右往左读,  $[AB]_{mn}$ 描述的是从初始 $n$ 态到末尾 $m$ 态的跃迁元, 它可以拆分成先从 $n$ 出发在物理量 $B$ 的作用下跃迁到中间可能性 $l$ , 再在物理量 $A$ 的作用下从中间可能性 $l$ 跃迁到 $m$ , 中间可能性 $l$ 要求和。其次, 注意等式(1.38)右边, 数据表 $A$ 和 $B$ 相乘,  $A$ 的列指标和 $B$ 的行指标是一样的, 都是中间可能性 $l$ , 而且它是要被求和的。这样被求和的指标我们称之为哑标,

所谓哑标，就是无需与等式(1.38)左边的指标对应存在的指标(等式左边的指标里面没有 $l$ )，而剩下的两个指标 $m$ 和 $n$ 在等式(1.38)的两边是要对应存在的，称作自由指标。

不妨进一步举例说明这种矩阵乘法是如何进行的。为了简单起见，假设我们考虑的不是无穷行无穷列的数据表，而是两行两列的数据表，比方说

$$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}, \quad B = \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix}. \quad (1.39)$$

那么按照乘法规则(1.38)， $AB$ 就是

$$\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix} = \begin{bmatrix} a_{11}b_{11} + a_{12}b_{21} & a_{11}b_{12} + a_{12}b_{22} \\ a_{21}b_{11} + a_{22}b_{21} & a_{21}b_{12} + a_{22}b_{22} \end{bmatrix}. \quad (1.40)$$

有时候人们也说， $AB$ 第1行第1列的元素是 $A$ 第1行乘以 $B$ 第1列的结果，同样， $AB$ 第1行第2列的元素是 $A$ 第1行乘以 $B$ 第2列的结果， $AB$ 第 $m$ 行第 $n$ 列的元素是 $A$ 第 $m$ 行乘以 $B$ 第 $n$ 列的结果。其中，当将 $A$ 的某一行抽出来乘以 $B$ 的某一列时，我们是像下面这样乘的，

$$\begin{bmatrix} a & b \end{bmatrix} \begin{bmatrix} c \\ d \end{bmatrix} = ac + bd, \quad (1.41)$$

也就是要把行和列中对应的各元素相乘起来，并求和。

注意，当我们考察的不是无穷行无穷列的数据表时，通常我们只会考察行数等于列数的数据表，比方 $N$ 行 $N$ 列的数据表，通常称之为 $N \times N$ 矩阵。而且，必须要确保两个数据表(矩阵)的行列数相同(比方两者都是 $N \times N$ )，我们才能够将它们进行运算，比如乘起来。我们不能够将一个 $N \times N$ 的数据表与一个 $M \times M$ (设 $N \neq M$ )的数据表乘起来，因为在这种情况下乘法规则(1.38)无法定义(乘法规则要求 $A$ 的列数与 $B$ 的行数相同)。

不难验证跃迁元数据表的乘法满足结合律，具体来说，假设有三个跃迁元数据表 $[A]_{mn}$ 、 $[B]_{mn}$ 以及 $[C]_{mn}$ ，则

$$[(AB)C]_{mn} = \sum_{l,k} [A]_{ml} [B]_{lk} [C]_{kn} = [A(BC)]_{mn}. \quad (1.42)$$

但是，跃迁元数据表的乘法，或者说矩阵乘法，通常不满足交换律。为了看清楚这一点，不妨以两个两行两列的数据表为例。假设

$$A = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \quad B = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}. \quad (1.43)$$

则不难按照上面的乘法规则算出

$$AB = \begin{bmatrix} i & 0 \\ 0 & -i \end{bmatrix}, \quad BA = \begin{bmatrix} -i & 0 \\ 0 & i \end{bmatrix}. \quad (1.44)$$

很显然， $AB \neq BA$ 。所以，一般来说，跃迁元数据表的乘法并不满足乘法交换律！而如果两个跃迁元数据表 $X$ 和 $Y$ 乘法可交换，即满足 $XY = YX$ ，则我们就称它们对易，反之就称为不对易。

两个物理量除了可以相乘当然还可以相加，因此我们还要定义相应跃迁元数据表的加法，相应的定义是显然的，如下

$$[A + B]_{mn} = [A]_{mn} + [B]_{mn}. \quad (1.45)$$

不妨以 $2 \times 2$ (两行两列)的数据表为例，说明这个加法是如何进行的，

$$\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} + \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix} = \begin{bmatrix} a_{11} + b_{11} & a_{12} + b_{12} \\ a_{21} + b_{21} & a_{22} + b_{22} \end{bmatrix}. \quad (1.46)$$

有时候还需要将一个常数乘以一个物理量，因此还需要定义常数与跃迁元数据表的乘法，叫做数乘，相应的定义也是显然的，如下

$$[\lambda A]_{mn} = \lambda [A]_{mn}, \quad (1.47)$$

式中 $\lambda$ 为某个常数。不妨以 $2 \times 2$ 的数据表为例进行具体说明，

$$\lambda \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} = \begin{bmatrix} \lambda a_{11} & \lambda a_{12} \\ \lambda a_{21} & \lambda a_{22} \end{bmatrix}. \quad (1.48)$$

定义了加法、乘法，以及数乘，这样一个原子系统所有可能跃迁元数据表的集合就构成了一个数学上所谓的代数。问题是，与物理量对应的跃迁元数据表还需要额外满足海森堡的条件，即它们必须是厄米矩阵，但是正如稍后将会说明的，所有厄米矩阵的集合并不是乘法封闭的，也就是两

个厄米矩阵的乘积可能不是厄米矩阵。所以，为了数学上的方便性，我们常常扩大考察范围，即不仅仅只考察厄米的跃迁元数据表，而是考察所有数学上可能的跃迁元数据表，即使它不厄米，因为这样就可以形成一个乘法封闭代数。当然，真正物理的(厄米的)跃迁元数据表只是这个乘法封闭代数中的一个子集。

为了说明为什么厄米矩阵的集合并不乘法封闭，我们来考察两个矩阵 $A$ 和 $B$ ，它们不一定厄米，我们考察 $AB$ 的厄米共轭矩阵 $(AB)^\dagger$ ，按照厄米共轭矩阵的定义，它应该满足

$$\begin{aligned} [(AB)^\dagger]_{mn} &= [AB]_{nm}^* = \left( \sum_l [A]_{nl} [B]_{lm} \right)^* = \sum_l [A]_{nl}^* [B]_{lm}^* \\ &= \sum_l [A^\dagger]_{ln} [B^\dagger]_{ml} = \sum_l [B^\dagger]_{ml} [A^\dagger]_{ln} = [B^\dagger A^\dagger]_{mn}, \end{aligned} \quad (1.49)$$

推导过程中我们多次使用了乘法规则。最终一头一尾的相等告诉我们

$$(AB)^\dagger = B^\dagger A^\dagger. \quad (1.50)$$

这是一个很有用的公式。下面我们假设 $A, B$ 均为厄米矩阵，即满足 $A^\dagger = A$ ,  $B^\dagger = B$ ，则利用上面的公式我们有 $(AB)^\dagger = B^\dagger A^\dagger = BA$ ，但是，由于矩阵乘法一般来说没有交换律，所以，一般来说， $BA \neq AB$ ，也即是 $(AB)^\dagger \neq AB$ ，从而一般来说 $AB$ 不是厄米矩阵。

前面说过，矩阵(数据表)的乘法通常不满足交换律，因此两个矩阵 $A, B$ 相乘，通常有 $AB \neq BA$ ，除非这两个矩阵对易。为了衡量两个矩阵的不对易程度，人们常常引入一个新的概念，叫做两个矩阵的对易子，记作 $[A, B]$ ，它也是一个矩阵，定义是

$$[A, B] \equiv AB - BA. \quad (1.51)$$

很显然，当且仅当 $A, B$ 对易的时候，其对易子才等于零。根据定义显然有

$$[A, B] = -[B, A]. \quad (1.52)$$

很显然，任何矩阵都必定和自己对易，即 $[A, A] = 0$ 。另外，根据单位矩阵的性质，任何矩阵都必定和单位矩阵对易，即 $[A, 1] = 0$ 。

现在，假设有三个矩阵 $A, B, C$ ，则有

$$[A, BC] = [A, B]C + B[A, C]. \quad (1.53)$$

为了证明这个结果，我们把等式左边按照定义展开，得

$$ABC - BCA, \quad (1.54)$$

另一方面，我们将等式右边按照定义展开，得

$$(AB)C - (BA)C + B(AC) - B(CA) = ABC - BCA, \quad (1.55)$$

推导的过程中我们用到了结合律(但是，请特别注意，你不能用交换律)，很显然等式右边这个最终结果与等式左边一样，所以原等式得证。类似的，还可以证明如下雅可比恒等式

$$[A, [B, C]] + [B, [C, A]] + [C, [A, B]] = 0. \quad (1.56)$$

证明过程大体如下

$$[A, [B, C]] = [A, BC - CB] = [A, BC] - [A, CB], \quad (1.57)$$

然后进一步利用上面的恒等式(1.53)，得到(详细过程请读者自己写出来)

$$[A, [B, C]] = [[A, B], C] + [B, [A, C]], \quad (1.58)$$

不难发现，这就等价于要证明的(1.56)式(请读者自己思考一下)。

对易子运算的一个好处是，若 $A, B$ 均为厄米矩阵，则 $i[A, B]$ 也必定为厄米矩阵(请读者自己证明，用到 $(C + D)^\dagger = C^\dagger + D^\dagger$ ，以及 $(\lambda D)^\dagger = \lambda^* D^\dagger$ )，所以，所有厄米矩阵的集合虽然在矩阵乘法运算下不封闭，但在求对易子再乘以虚数单位 $i$ 的运算下是封闭的，进一步考虑加法和数乘(限制于乘以实数，厄米矩阵乘以实数依然得到厄米矩阵)，当然也还是运算封闭的(即对一些厄米矩阵进行这几种运算的结果依然是厄米矩阵)，因此厄米矩阵的集合在这几种运算下构成一个代数，而且由于求对易子的运算满足雅可比恒等式，求对易子再乘以 $i$ 的运算当然也满足雅可比恒等式，所以数学家也常常称它为李代数(满足雅可比恒等式的代数就叫李代数)。由于厄米矩阵对应着物理可观测量的跃迁元数据表，所以也称这个李代数为物理可观测量的李代数，实际上，这就是为什么对易子运算在量子力学中这么重要的原因。



### 1.2.3 海森堡运动方程

经典物理里有一个最重要的量，那就是系统的哈密顿量 $H$ ，也就是经典系统的能量(对哈密顿量相关内容的初步介绍，请读者参阅我的另一个简短讲义，《初识经典力学》)，等于动能加势能。但是现在，一方面系统的能量值量子化了，另一方面，作为一个物理量，哈密顿量 $H$ 当然也要变成跃迁元数据表。与哈密顿量相应的跃迁元数据表应该是什么呢？根据海森堡的一般性想法，它当然应该是 $[H]_{mn} \equiv [H_0]_{mn}e^{-i\omega_{mn}t}$ ，其中 $[H_0]_{mn}$ 与时间无关。注意到 $\omega_{nn} = 0$ ，相应的 $e^{-i\omega_{nn}t} = 1$ ，所以这个跃迁元数据表应该是这样

$$H = \begin{bmatrix} [H_0]_{11} & [H_0]_{12}e^{-i\omega_{12}t} & [H_0]_{13}e^{-i\omega_{13}t} & \dots \\ [H_0]_{21}e^{-i\omega_{21}t} & [H_0]_{22} & [H_0]_{23}e^{-i\omega_{23}t} & \dots \\ [H_0]_{31}e^{-i\omega_{31}t} & [H_0]_{32}e^{-i\omega_{32}t} & [H_0]_{33} & \dots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix}. \quad (1.59)$$

当然，它还要满足厄米条件，这里不写了。很显然，除了对角线上的元素(称作对角元)不依赖于时间之外，其它非对角元都依赖于时间。

现在一个关键的假设来了，我们知道，一个封闭的经典力学系统总能量是守恒的，其哈密顿量不会随时间演化。我们假设：**量子化以后，能量守恒定律依然成立，因此封闭系统的跃迁元数据表不能依赖于时间。**从而，根据这条假设，我们知道， $H$ 非对角元的那些 $[H_0]_{mn}, m \neq n$ 必定都为零，即 $[H_0]_{mn} = 0, m \neq n$ ，因为否则， $H$ 就会依赖于时间。所以哈密顿量的跃迁元数据表必定为

$$H = \begin{bmatrix} [H_0]_{11} & 0 & 0 & \dots \\ 0 & [H_0]_{22} & 0 & \dots \\ 0 & 0 & [H_0]_{33} & \dots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix}. \quad (1.60)$$

这样的，只有对角元非零，其它元素都为零的矩阵称作对角矩阵。厄米条件进一步告诉我们 $[H_0]_{nn} = [H_0]_{nn}^*$ ，所以这些对角元必定都是实数！然而，在另一方面，我们知道系统的能量值量子化为了 $E_n$ ，所以，最自然的可能性就是，这些对角元刚好就是这些量子化能量，即 $[H_0]_{nn} = E_n$ 。因此我们

知道，哈密顿量的跃迁元数据表最自然的可能性就是

$$H = \begin{bmatrix} E_1 & 0 & 0 & \dots \\ 0 & E_2 & 0 & \dots \\ 0 & 0 & E_3 & \dots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix}, \quad (1.61)$$

不妨称之为哈密顿矩阵，在不引起混淆的时候，也常常简称为哈密顿量。

我们知道，任意一个物理量 $A$ 所对应的跃迁元数据表可以写成 $A = [A]_{mn} = [A_0]_{mn}e^{-i\omega_{mn}t}$ ，式中 $[A_0]_{mn}$ 与时间无关，而 $\omega_{mn} = (E_n - E_m)/\hbar$ 。很显然，数据表 $A$ 是随时间演化的，现在我们想知道的是它满足什么演化微分方程。为此，我们把它对时间求导，得

$$\frac{d}{dt}[A]_{mn} = -i\omega_{mn}[A]_{mn} = \frac{-i}{\hbar}([A]_{mn}E_n - E_m[A]_{mn}). \quad (1.62)$$

排列成数据表的形式，即有

$$\begin{aligned} i\hbar \frac{dA}{dt} &= \begin{bmatrix} [A]_{11}E_1 & [A]_{12}E_2 & \dots \\ [A]_{21}E_1 & [A]_{22}E_2 & \dots \\ \vdots & \vdots & \ddots \end{bmatrix} - \begin{bmatrix} E_1[A]_{11} & E_1[A]_{12} & \dots \\ E_2[A]_{21} & E_2[A]_{22} & \dots \\ \vdots & \vdots & \ddots \end{bmatrix} \\ &= \begin{bmatrix} [A]_{11} & [A]_{12} & \dots \\ [A]_{21} & [A]_{22} & \dots \\ \vdots & \vdots & \ddots \end{bmatrix} \begin{bmatrix} E_1 & 0 & \dots \\ 0 & E_2 & \dots \\ \vdots & \vdots & \ddots \end{bmatrix} - \begin{bmatrix} E_1 & 0 & \dots \\ 0 & E_2 & \dots \\ \vdots & \vdots & \ddots \end{bmatrix} \begin{bmatrix} [A]_{11} & [A]_{12} & \dots \\ [A]_{21} & [A]_{22} & \dots \\ \vdots & \vdots & \ddots \end{bmatrix}. \end{aligned}$$

很明显，利用哈密顿矩阵的定义(1.61)，这个最终结果可以很整洁地写成

$$i\hbar \frac{dA}{dt} = [A, H]. \quad (1.63)$$

这就是著名的海森堡运动方程。

不难验证，海森堡运动方程与跃迁元数据表的乘法是相容的。即假设有两个物理量，其跃迁元数据表分别为 $A$ 和 $B$ ，它们均满足海森堡运动方程，则可以证明，两者的乘积 $AB$ 也必定满足海森堡运动方程。证明如下(注意，推导的过程中不能使用乘法交换律)

$$i\hbar \frac{d}{dt}(AB) = i\hbar \left( \frac{dA}{dt}B + A\frac{dB}{dt} \right) = [A, H]B + A[B, H] = [AB, H], \quad (1.64)$$

证明完毕！其中第2个等于号用到了 $A, B$ 均满足海森堡运动方程，最后一个等于号是用到了矩阵对易子的恒等式(1.53)。另外， $A, B$ 的和 $A + B$ 满足海森堡运动方程更是容易证明的。进而不难得到，所有满足海森堡运动方程的跃迁元数据表的集合在加法、乘法，以及数乘运算下是封闭的，从而构成一个代数。

以上证明的一个推论是，若 $A, B$ 均满足海森堡运动方程，则它们的对易子 $[A, B]$ 也必定满足海森堡运动方程。进而，所有满足海森堡运动方程的跃迁元数据表在求对易子的运算(再加上加法运算和数乘运算)下也构成一个封闭代数，由于求对易子运算满足雅可比恒等式，所以这这也是一个李代数。显然，如果将数乘限制于乘以实数，将求对易子运算再乘以 $i$ ，那所有物理可观测量的跃迁元矩阵(必然厄米且满足海森堡运动方程)的集合也是一个满足海森堡运动方程的李代数。

将海森堡运动方程中的 $A$ 取成哈密顿量 $H$ 本身，即有 $i\hbar \frac{dH}{dt} = [H, H] = 0$ ，很显然，不过是又一次验证了最终的方程会自动给出能量守恒。更一般的，如果一个物理量 $A$ (所以满足厄米条件)与哈密顿量对易， $[A, H] = 0$ ，从而 $\frac{dA}{dt} = 0$ ，我们就称 $A$ 为一个守恒量。显然，如果物理量 $A, B$ 均为守恒量，则 $AB$ 也必定为守恒量(虽然可能不满足厄米条件，从而不是真正的物理量)，同样， $i[A, B]$ 也必定为守恒量，且满足厄米条件。所以，一个量子系统所有物理守恒量的集合在求对易子再乘以虚数单位 $i$ 等几种运算下也构成一个封闭的李代数。

作为物理量，核外电子位置矢量的跃迁元数据表 $\mathbf{x}$ 以及动量的跃迁元数据表 $\mathbf{p}$ 当然也满足海森堡运动方程，从而即有

$$i\hbar \frac{d\mathbf{x}}{dt} = [\mathbf{x}, H], \quad i\hbar \frac{d\mathbf{p}}{dt} = [\mathbf{p}, H]. \quad (1.65)$$

### 1.2.4 基本对易关系

下面要推导一个重要的结论，这个结论在一维系统中看得比较清楚。另外，下面我们要同时考虑经典的变量和相应的量子跃迁元数据表，所以，我们稍微改变一下符号约定：我们以小写的 $x$ 和 $p$ 分别表示一维经典系统的位置变量和动量，因此， $x$ 和 $p$ 都是通常的数，不是矩阵，而相应的量子跃迁元数据表，我们分别记作大写的 $X$ 和 $P$ 。

对于这样的一维系统，方程(1.65)就变成了

$$i\hbar \frac{dX}{dt} = [X, H], \quad i\hbar \frac{dP}{dt} = [P, H]. \quad (1.66)$$

为了得到哈密顿矩阵的表达式，我们先写出经典系统的哈密顿量  $\frac{p^2}{2m_e} + V(x)$ ，式中  $V(x)$  为势能函数， $m_e$  为粒子的质量。然后，我们将表达式中的  $x, p$  分别替换成相应的跃迁元矩阵，即  $x \rightarrow X, p \rightarrow P$ ，从而就得到

$$H = \frac{P^2}{2m_e} + V(X), \quad (1.67)$$

一个细节的问题是，由于现在  $X$  是跃迁元数据表，而不是普通的数，所以其函数  $V(X)$  需要一个合适的定义，一种最简单的定义是，先把普通的函数  $V(x)$  泰勒展开成如下幂级数

$$V(x) = a_0 + a_1x + a_2x^2 + a_3x^3 + \dots \quad (1.68)$$

再把  $x \rightarrow X$ ，就可以了，也即是说

$$V(X) = a_0 + a_1X + a_2X^2 + a_3X^3 + \dots, \quad (1.69)$$

当然常数项  $a_0$  应该理解成  $a_0$  乘以单位矩阵 1。很显然这个幂级数的每一项都只涉及到矩阵乘法，因此是有定义的。

现在，我们立即可以从方程(1.66)得出一个推论：即  $X$  和  $P$  不可能对易！具体的推理是这样的：首先  $[P, P] = 0$ ，从而  $[P, P^2] = 0$ ，这是因为  $[P, P^2] = [P, PP]$ ，利用恒等式(1.53)，即有  $[P, PP] = P[P, P] + [P, P]P = 0$ 。同样， $[X, X] = 0$ ，从而必有  $[X, V(X)] = 0$  (想想为什么)。类似的，如果  $X$  和  $P$  对易，则  $X$  与  $P^2$  也必定对易，因为根据恒等式(1.53)， $[X, P^2] = P[X, P] + [X, P]P$ 。另外，如果  $P$  和  $X$  对易，则  $P$  与  $V(X)$  也必定对易 (想想为什么)。但，如此一来 (即假如  $[X, P] = 0$ )，那  $X$  就和  $H$  中的每一项都对易，从而  $[X, H] = 0$ ，同样  $[P, H] = 0$ ，那方程(1.66)就会给出  $\frac{dX}{dt} = 0, \frac{dP}{dt} = 0$ ，也即是说，任何一维系统的运动方程都是这个，而与系统的势能无关！这当然是不可能的！所以，必定有  $[X, P] \neq 0$ 。

关键是  $[X, P]$  等于什么？不妨暂记

$$[X, P] = C. \quad (1.70)$$

所以，关键是要导出矩阵 $C$ 具体是什么？

为了找到这个答案，我们先回到熟悉的经典系统，这时候相应的方程当然不是海森堡运动方程，而是哈密顿正则方程(请参阅我的《初识经典力学》)，

$$\frac{dx}{dt} = \frac{\partial H}{\partial p} = \frac{p}{m_e}, \quad \frac{dp}{dt} = -\frac{\partial H}{\partial x} = -V'(x), \quad (1.71)$$

式中我们代入了经典的 $H = \frac{p^2}{2m_e} + V(x)$ ,  $V'(x)$ 表示 $V(x)$ 对自变量 $x$ 的导数，用上面的幂级数也就是

$$V'(x) = a_1 + 2a_2x + 3a_3x^2 + \dots \quad (1.72)$$

下面我们作一个重要的假设：**即量子系统要和相应的经典系统能够对应得上！**换言之，方程(1.66)最终给出的结果，应该等价于将上面经典系统运动方程做 $x \rightarrow X, p \rightarrow P$ 替换之后的结果，也就是应该等价于

$$\frac{dX}{dt} = \frac{P}{m_e}, \quad \frac{dP}{dt} = -V'(X). \quad (1.73)$$

下面我们来看(1.66)式的第一个式子，由于 $[X, V(X)] = 0$ ，所以这个式子告诉我们

$$\begin{aligned} \frac{dX}{dt} &= \frac{1}{i\hbar}[X, H] = \frac{1}{i\hbar}[X, \frac{P^2}{2m_e}] = \frac{1}{i\hbar} \frac{1}{2m_e} ([X, P]P + P[X, P]) \\ &= \frac{1}{m_e} \frac{1}{i\hbar} \frac{1}{2} (CP + PC). \end{aligned} \quad (1.74)$$

而我们想要的结果是 $\frac{dX}{dt} = \frac{1}{m_e}P$ ，不难看出，为了达到这个想要的结果，最简单的就是取 $C = i\hbar \cdot 1$ ，式中1当然表示单位矩阵，常常省略。所以我们试探性地取

$$[X, P] = i\hbar. \quad (1.75)$$

下面来看这种取法能否使得(1.66)式的第二个式子也同时达到我们想要的结果，也就是等价于 $\frac{dP}{dt} = -V'(X)$ 。为此，我们注意到 $[P, P^2] = 0$ ，所以，(1.66)式的第二个式子告诉我们

$$\frac{dP}{dt} = \frac{1}{i\hbar}[P, H] = \frac{1}{i\hbar}[P, V(X)], \quad (1.76)$$

为了计算 $[P, V(X)]$ , 我们代入 $V(X)$ 的级数展开, 从而

$$[P, V(X)] = a_1[P, X] + a_2[P, X^2] + a_3[P, X^3] + \dots \quad (1.77)$$

不断利用恒等式(1.53), 以及我们的假设 $[P, X] = -[X, P] = -i\hbar$ , 不难计算出结果为(具体计算留给读者作练习)

$$[P, V(X)] = -i\hbar(a_1 + 2a_2X + 3a_3X^2 + \dots) = -i\hbar V'(X). \quad (1.78)$$

代入(1.76)式, 即有

$$\frac{dP}{dt} = -V'(X). \quad (1.79)$$

正是我们想要的结果。

所以, (1.75)式的确是正确的! 不过, 我们不妨进行最后一个检验, 即注意到(1.75)式右边是一个与时间无关的常数, 而左边的 $X, P$ 都是随时间演化的, 所以我们还要检验左边求对易子的结果不随时间演化。这的确是成立的, 验证如下

$$\begin{aligned} \frac{d}{dt}([X, P]) &= \dot{X}P + X\dot{P} - \dot{P}X - P\dot{X} \\ &= (PP/m_e - XV'(X) + V'(X)X - PP/m_e) = 0. \end{aligned} \quad (1.80)$$

式中第二行的等于是代入了方程(1.73), 而最后的等号是利用了 $[X, V'(X)] = 0$ 。

以上推导过程也说明了一件事情, 即海森堡运动方程(1.66)就是经典哈密顿力学哈密顿正则方程的量子对应物, 如果读者学习过经典哈密顿力学的泊松括号运算的话, 将会进一步看清两者的对应, 大体来说就是经典的泊松括号运算, 在量子里面就对应求对易子运算。当然, 这不是我们想强调的内容, 因此就不再展开了。

以上得到的结果(1.75)在量子力学中有基本的重要性, 所以也叫做量子力学基本对易关系, 在历史上, 它最早是玻恩等人得到的。实际上在我看来, 海森堡、玻恩和若尔当三个人对这个结果都有重要的贡献, 海森堡的原始论文实际上暗含了推导的方向, 当然, 最终确立这个结果的是玻恩和若尔当。不过, 玻恩等人用的不是我们这里给出的推导方法。

以上考察的是一个自由度的一维问题，下面把重要结果(1.75)推广到三个自由度的三维情形。这时候我们恢复原来的符号约定，以小写的 $\mathbf{x}$ 和 $\mathbf{p}$ 分别代表粒子的位置和动量的跃迁元矩阵。当然，它们不仅是矩阵，同时还是三维空间的向量，有三个分量，因此实际上可以记作

$$\mathbf{x} = (x_1, x_2, x_3), \quad \mathbf{p} = (p_1, p_2, p_3). \quad (1.81)$$

当然，每一个分量的位置 $x_a$ 和动量 $p_a$  ( $a = 1, 2, 3$ )本身都是跃迁元矩阵。

注意到运动的不同自由度是相互独立的，也就是说，不同自由度的这些基本变量之间应该和普通数的乘法一样，是相互对易的，因此有，当 $a \neq b$ 时， $[x_a, x_b] = 0$ ,  $[p_a, p_b] = 0$ , 以及 $[x_a, p_b] = 0$ 。另外，上面对单自由度的讨论又告诉我们，对于同一个自由度， $[x_a, p_a] = i\hbar$ ,  $[x_a, x_a] = 0$ ,  $[p_a, p_a] = 0$ 。综上，我们可以把结果概括如下

$$[x_a, p_b] = i\hbar\delta_{ab}, \quad [x_a, x_b] = 0, \quad [p_a, p_b] = 0, \quad a, b = 1, 2, 3 \quad (1.82)$$

式中 $\delta_{ab}$ 就是克龙内克符号，当 $a = b$ 时取1，当 $a \neq b$ 时取0。由于任何物理可观测量都可以用一些 $x_a$ 和一些 $p_a$ 通过数乘、加法以及矩阵乘法生成(当然，同时要满足厄米条件)，因此任何两个物理可观测量进行求对易子再乘以 $i$ 运算的结果都可以通过上面的基本结果(1.82)计算出来，我们称(1.82)式生成了物理可观测量的李代数。通常称这个李代数为海森堡代数。

### 1.2.5 求解谐振子

然而，这一切对吗？海森堡的想法虽然并不牵强，但是，他真是在玩一个疯狂的游戏，经典物理里那些习以为常的物理量怎么就变成无穷大的矩阵了呢？怎么它们就要按照那种奇怪的乘法乘起来呢？是的，对于没有学过相关数学知识的物理学家而言，海森堡发明的乘法虽然并不牵强，但它可太奇怪了！这就是具有最高创造性的工作的一个特征，很自然，同时也很疯狂。

然而，这一切正确吗？

最直接的检验当然就是把它应用到氢原子上，然后解出氢原子能级，也就是哈密顿矩阵(正如我们讲过的，在到此为止的处理中，它都必然是一个对角矩阵)的对角元 $E_n$ ，甚至解出每一个 $[\mathbf{x}]_{mn}$ ，进而算出谱线的强度。然

而，这在数学上太难了一点，海森堡解不出来，这个问题是稍后泡利解决的。

因此海森堡想到将他的思想用到一个稍微简单一点的问题上，也就是线性谐振子上。也即是说，假设发出电磁辐射的是一个带电线性谐振子(而不是更复杂的氢原子)，一个线性谐振子从一个定态能级跃迁到另外一个定态能级从而辐射出电磁波。根据普朗克对黑体辐射公式的推导，线性谐振子的能量应该是量子化的，而且其量子化能级具有  $E_n = n\hbar\omega$  这样的形式。但是，普朗克是直接假设了这个结果(非如此假设不足以解释黑体辐射的实验曲线)，他可没有具体推导出这个结果，假设海森堡能用它的矩阵想法推导出这个，那不就是他理论正确的强烈证据吗？

当时的海森堡的确做了这件事情，并且他发现，结果真是正确的！

下面我们来重复海森堡的工作。我们要考察的是单自由度的所谓一维谐振子，其哈密顿矩阵为

$$H = \frac{P^2}{2m_e} + \frac{1}{2}m_e\omega^2 X^2. \quad (1.83)$$

其中  $\frac{1}{2}m_e\omega^2 X^2$  为谐振子的势能(弹簧振子的势能)。由运动方程(1.73)，可以得到

$$\dot{X} = P/m_e, \quad \dot{P} = -m_e\omega^2 X, \quad (1.84)$$

两式联立，即有  $\ddot{X} + \omega^2 X = 0$ ，代入  $X = [X_0]_{mn}e^{-i\omega_{mn}t}$ ，即有

$$(-\omega_{mn}^2 + \omega^2)[X_0]_{mn} = 0. \quad (1.85)$$

以上结果表明，仅当  $\omega_{mn} = \pm\omega$  时， $[X_0]_{mn}$  才不为零，即除了两个矩阵元，其它矩阵元均为零，对定态进行适当编号，从而可设  $\omega_{m,m\pm 1} = \pm\omega$ ，进而  $[X_0]_{m,m\pm 1}$  非零，其余矩阵元都为零。

由量子力学基本对易关系  $XP - PX = i\hbar$ ，知  $m_e(X\dot{X} - \dot{X}X) = i\hbar$ ，即

$$m_e([X\dot{X}]_{mn} - [\dot{X}X]_{mn}) = i\hbar\delta_{mn}. \quad (1.86)$$

当  $n = m$  时，利用矩阵乘法规则，即有

$$\begin{aligned} m_e \sum_l (-[X_0]_{ml}\omega_{lm}[X_0]_{lm} + \omega_{ml}[X_0]_{ml}[X_0]_{lm}) &= \hbar \\ \Rightarrow 2m_e \sum_l (\omega_{ml}[X_0]_{ml}[X_0]_{lm}) &= \hbar. \end{aligned} \quad (1.87)$$



利用厄米条件 $[X_0]_{lm} = [X_0]_{ml}^*$ , 即有

$$2m_e \sum_l (\omega_{ml} |[X_0]_{ml}|^2) = \hbar. \quad (1.88)$$

但是我们知道, 仅仅当 $l = m \pm 1$ 时,  $[X_0]_{ml}$ 才非零, 所以上式的求和中, 仅仅只有这两项有贡献, 从而

$$2m_e \omega (|[X_0]_{m,m+1}|^2 - |[X_0]_{m,m-1}|^2) = \hbar. \quad (1.89)$$

注意到 $[X_0]_{m,m+1} = [X_0]_{m+1,m}^*$ , 也即是说, 我们有递推公式

$$|[X_0]_{m+1,m}|^2 = |[X_0]_{m,m-1}|^2 + \hbar/(2m_e \omega). \quad (1.90)$$

假设把最低能级编号为第0能级(虽然前文的习惯是把最低能级编码为第1能级), 由于不存在-1能级, 因此当然有 $[X_0]_{0,-1} = 0$ 。进而上面的递推公式就给出

$$|[X_0]_{m+1,m}|^2 = (m+1)\hbar/(2m_e \omega). \quad (1.91)$$

通常可以进一步取 $[X_0]_{m+1,m}$ 为实数, 进而即有 $[X_0]_{m+1,m} = [X_0]_{m,m+1} = \sqrt{m+1} \sqrt{\frac{\hbar}{2m_e \omega}}$ . 不妨将 $[X_0]_{mn}$ 的矩阵具体排出来(注意, $m, n$ 从零开始数的), 即是

$$[X_0]_{mn} = \sqrt{\frac{\hbar}{2m_e \omega}} \begin{bmatrix} 0 & \sqrt{1} & 0 & 0 & \dots \\ \sqrt{1} & 0 & \sqrt{2} & 0 & \dots \\ 0 & \sqrt{2} & 0 & \sqrt{3} & \dots \\ 0 & 0 & \sqrt{3} & 0 & \dots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix}. \quad (1.92)$$

谐振子的能级就是哈密顿矩阵的对角元, 从而

$$\begin{aligned} E_n &= [H]_{nn} = \frac{1}{2} m_e ([\dot{X}^2]_{nn} + \omega^2 [X^2]_{nn}) \\ &= \frac{1}{2} m_e \left( \sum_l [\dot{X}]_{nl} [\dot{X}]_{ln} + \omega^2 \sum_l [X]_{nl} [X]_{ln} \right). \end{aligned} \quad (1.93)$$

注意到 $[\dot{X}]_{nl} = -i\omega_{nl}[X]_{nl}$ 以及厄米条件, 并注意到 $|[X]_{mn}|^2 = |[X_0]_{mn}|^2$ , 即有

$$\begin{aligned} E_n &= \frac{1}{2}m_e \sum_l (\omega_{nl}^2 + \omega^2) |[X]_{nl}|^2 \\ &= m_e \omega^2 (|[X_0]_{n,n+1}|^2 + |[X_0]_{n,n-1}|^2) \\ &= (n + \frac{1}{2})\hbar\omega. \end{aligned} \quad (1.94)$$

上式最后一行我们代入了(1.91)式。

所以, 根据海森堡的推导, 线性谐振子的能级 $E_n = (n + \frac{1}{2})\hbar\omega = n\hbar\omega + \frac{1}{2}\hbar\omega$ ,  $n = 0, 1, 2, \dots$  与普朗克的量子化假设是完全吻合的。实际上, 相比于普朗克的假设, 结果只多了一个整体的常数 $\frac{1}{2}\hbar\omega$ , 而定态跃迁辐射光的过程只涉及能级差, 因此这个整体常数当然是普朗克无从预料的。所以, 海森堡知道, 他对了!

### 1.2.6 符号的新发展

跃迁元的概念, 以及跃迁元的乘法规则(1.38)当然是海森堡引领的这场矩阵革命中最具创造性, 也是最核心的成果。这一小节, 我们再次回到它们, 我们将用一种新的符号来书写它们, 这种新的符号是狄拉克所引入的。你可能觉得, 只不过是一种不同的记号而已, 不是什么本质的东西, 然而情况并非如此, 不同的符号系统有不同的表达能力, 事实证明, 狄拉克所引入的这种新记号在量子力学后来的发展中有更加强大的表达能力。

为了引入这一新的符号系统, 让我们考察跃迁元矩阵 $[A]_{mn}$ , 它表示在物理量 $A$ 的微扰之下, 系统从定态可能性( $n$ )(还记得把, 前面说过, 定态就是系统的一种特定可能性)跃迁到定态可能性( $m$ )的跃迁元, 根据前文说过的, 这个跃迁元可以示意性地记作

$$[A]_{mn} = \left\{ (m) \overset{A}{\leftarrow} (n) \right\}. \quad (1.95)$$

现在, 我们把这个示意性的符号再简化一下, 使之成为一种更好用的正式记号, 我们将跃迁元 $[A]_{mn}$ 重记为

$$[A]_{mn} = \left\{ (m) \overset{A}{\leftarrow} (n) \right\} = \langle m | \hat{A} | n \rangle, \quad (1.96)$$

$\langle m|\hat{A}|n\rangle$ 就是这个新的正式记号，它的物理意义还是和 $[A]_{mn}$ 一样，只不过记号变了。现在，我们用 $|n\rangle$ 来表示初始的定态可能性，用 $\langle m|$ 来表示末尾的定态可能性，用加了一个“帽子”的 $\hat{A}$ 来表示物理量 $A$ 对系统的作用(通过微扰引入)，注意，新记号区分了初态可能性和末态可能性，而且我们将初态可能性写在了 $\hat{A}$ 的右边，将末态可能性写在了它的左边。因此，跃迁元 $[A]_{mn}$ 就重记成了 $\langle m|\hat{A}|n\rangle$ ，读作系统从初态 $|n\rangle$ 经过物理量 $A$ 的作用跃迁到末态 $\langle m|$ ，注意，从右往左读的。

用这种新的记号，厄米共轭矩阵的定义 $[M^\dagger]_{mn} = [M]_{nm}^*$ 就应该重写为

$$\langle m|\hat{M}^\dagger|n\rangle = \langle n|\hat{M}|m\rangle^*. \quad (1.97)$$

而物理量的厄米条件 $[A]_{mn} = [A]_{nm}^*$ 就应该重写为

$$\langle m|\hat{A}|n\rangle = \langle n|\hat{A}|m\rangle^*. \quad (1.98)$$

它的含义当然还是，交换初末态之后反向的从 $(n)$ 到 $(m)$ 的跃迁元，等于正向的从 $(m)$ 到 $(n)$ 的跃迁元的复数共轭，新记号由于明显地区分了初末态，所以这个含义就更明显了。

按照这种新的记号，跃迁元的乘法规则(1.38)就应该重写为

$$\langle m|\hat{A}\hat{B}|n\rangle = \sum_l \langle m|\hat{A}|l\rangle \langle l|\hat{B}|n\rangle. \quad (1.99)$$

它表示，系统在物理量 $AB$ 的影响下从 $(n)$ 跃迁到 $(m)$ 的跃迁元，等于系统先在 $B$ 的影响下从 $(n)$ 跃迁到中间可能性 $(l)$ 的跃迁元，再乘以，在 $A$ 的影响下从 $(l)$ 跃迁到 $(m)$ 的跃迁元，注意，对于 $B$ 来说， $(l)$ 是一个末态可能性，但是，对于 $A$ 来说， $(l)$ 是一个初态可能性。当然，由于中间可能性不能确定，所以在乘的时候我们要把所有的中间可能性都平等地加起来。值得说明的是，从上面的解释可以清楚看到，这个式子的正确读法也是倒着的，即从最右边往左边读。

特别的，按照新的记号，物理量1的跃迁元 $[1]_{mn}$ 就应该记作 $\langle m|1|n\rangle$ ，不过，由于1的意思就是不受任何微扰，所以在书写中常常省略，就把这个跃迁元记作 $\langle m|n\rangle$ ，它表示系统在不受微扰时从可能性 $(n)$ 跃迁到 $(m)$ 的跃迁元。前面我们说过， $[1]_{mn} = \delta_{mn}$ ，所以，用新的记号即是

$$\langle m|n\rangle = \delta_{mn}, \quad (1.100)$$

它表示系统在不受微扰时从 $(n)$ 跃迁到 $(n)$ 自身的概率是1，而跃迁到其它可能性的概率都是零。我们称这种性质为：**系统的任何两个不同定态可能性都可以确定地相区分！**因此，系统定态可能性的完备集(即所有定态可能性的集合)是一个**可确定区分可能性完备集**。

本章所讨论的跃迁元都是定态可能性之间的跃迁元，非常依赖于定态这一概念，这实际上是一个很大的局限，下一章我们将从这种局限中获得解放。