

量子力学入门笔记

在几个月前，我接触到了量子力学，发现这是一套有趣的、美丽的，却令人捉摸不透的理论。写在最前面：一定要把量子力学当作是线性代数去看，去体会，很少有人能够真正的理解量子力学，但是在学习它是一个十分有趣的过程。我不明白为什么国内的教材喜欢把量

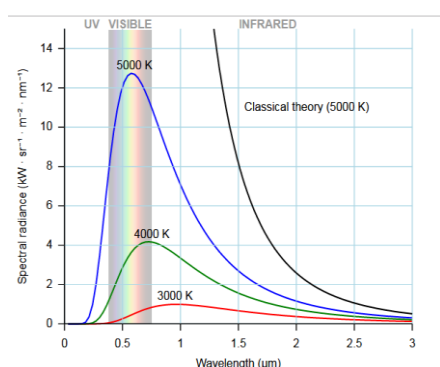


Figure 1 黑体辐射谱（来自：wikipedia）

子力学与偏微分方程联系在一起，导致很多人、包括我在看了几页书之后一头雾水，在我看来，把量子力学当作偏微分方程的课程来讲的作者他自己也没有很好的理解量子力学，所以在我看来，把量子力学当作线性代数来看是更容易理解的。量子力学很年轻，只有 100 年左右的发展时间，可是在这 100 年里，我们的生活被永远地、不可逆地改变。下面的量子力学被建立的历程在很多

物理学教材上面都可以看得到，你可以跳过下面的一大段发展历程，当你学到一些量子力学的内容后回过头来会发现，这些人的努力是多么的恰到好处。从机械钟摆到电子显示屏的手表；从马车到无人驾驶的特斯拉；从机械计算器到太湖之光，这一切都由量子力学在做基石，所以说不了解量子力学等于生活在了 100 年前。牛顿所构建的经典力学完美的解释了宏观物体的力学性质，麦克斯韦构建的麦克斯韦方程组完美的诠释了电磁场的性质。

I. 历史背景



Figure 2 Max Planck as a young man(1878) (来自: wikipedia)



Figure 3 Niels Bohr (来自: wikipedia)

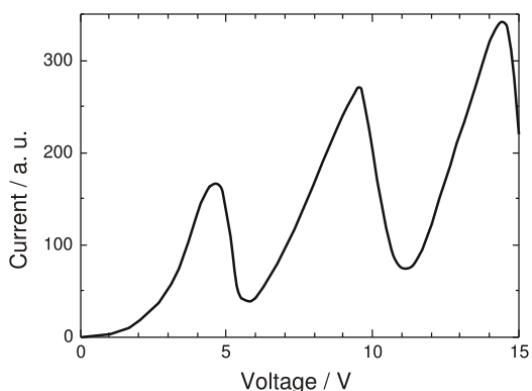
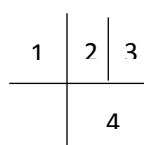
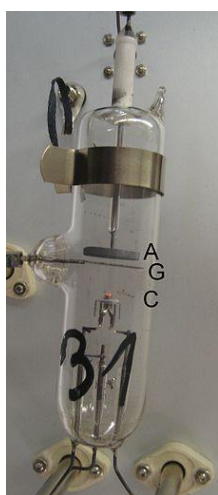


Figure 4 图一：弗兰克—赫兹实验装置；图二：James Franck；图三：Gustav Herz；图四：弗兰克—赫兹实验曲线（来源: wikipedia）

就像所有物理课本中所说的一样，在 19 世纪末、20 世纪初，物理学“晴空万里”，但是天空中飘着两朵乌云，迈克尔逊—莫雷干涉实验与以太假说的相悖以及奇怪的黑体辐射曲线。而迈克尔逊—莫雷实验已经被解决，它就是大名鼎鼎的狭义相对论。黑体辐射很好理解，在我们的宇宙中，所有具有温度的物体都会辐射出电磁波，不同的波长反映了辐射电磁波具有能量的不同，之所以叫做**黑体**辐射，是为了排除物体可能反射其他光源产生的光，所以黑体辐射出来不同电磁波的分布仅仅由温度所决定，这样所得到的谱线就叫做黑体辐射。经典电磁学所得

出来的结论与实验结果截然不同。直到 1900 年，普朗克（Max Planck）利用他强大的数学能力硬生生地“凑”出来了我们现在所看到的黑体辐射公式，普朗克自己也没有想到这样一个完全是“暴力破解”所得到的公式，竟然能够完美的解释黑体辐射谱。普朗克自己也说到：“a purely formal assumption ... actually I did not think much about it...”(来自: wikipedia)。

后来普朗克假设能量在传递过程中是一份一份的传递，也就是我们今天所看到的 $\epsilon = h\nu$ ，才把公式解释的通，如果时间允许，我会在最后推导黑体辐射的公式。在当时连普朗克自己都觉得这个想法太过于荒谬，也没有在物理界引起很大的反响。到了 1905 年，爱因斯坦根据普朗克的思想解释了光电效应，并且因此获得了诺贝尔物理学奖。当时还有一个在高中困扰



Figure 6 Louis de Broglie and Erwin Schrödinger (来自: wikipedia)

我的问题困扰着当时的人们，就是在用经典理论解释原子模型的时候，原子核外电子的运动是一个圆周运动，那么核外电子必然有一个向心加速度，一个有加速度的带电物体一定会产生电磁波，而电磁波携带有能量，那么问题来了，久而久之电子

为什么不会落到原子核上面？在 1913 年左右，一个叫做玻尔 (Bohr) 的年轻人，吸取了能量量子化的思想，他认为氢原子和电子属于满足库仑势束缚系统也应该符合能量量子化，并且用了一种不会传播能量的波来描述这个系统的能量，这种波就是驻波。他的理论很好的解释了氢原子光谱，尤其是我们所熟悉的巴尔莫系的谱线。当时还有一个比较著名的实验是我们上学期做过的物理选修实验，1914 年的弗兰克—赫兹实验，大致的实验思想是这样的：在一个密闭的玻璃管中，把电子通过电压加速，穿过玻璃管中的汞蒸气，电子形成的电流随着加速电压的增大，并不是一个线性的函数，而是有很多个峰和谷，也就是说在某些特定的能量下，阴极发射的电子可以把能量传递给汞原子，导致很少有电子穿过汞蒸气，之后当电子继续达到某个能量值后，就又会产生一个极小值，这个实验又说明了能量在传递过程中是

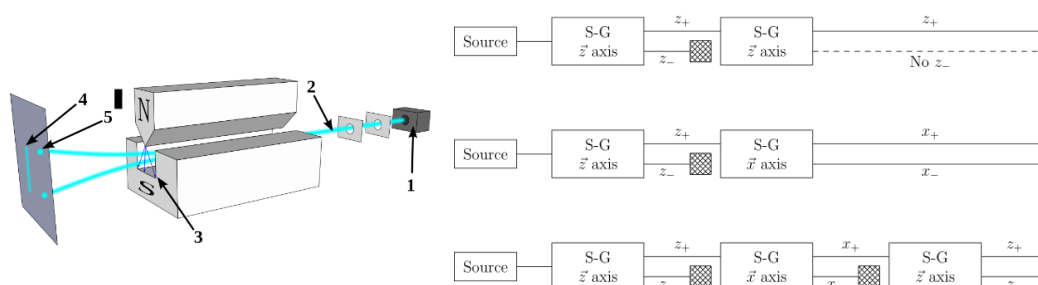


Figure 5 Stern-Gerlach experiment apparatus and sequential experiment. (来自: wikipedia)

一份一份地传递。这个实验也很好的说明了能量量子化的现象。时间到了 1920 年，施特恩和格拉赫提出了著名的，最能体现量子现象的实验——施特恩—格拉赫实验 (Stern-Gerlach experiment)。他们把一束银原子通过一个不均匀的磁场，他们发现银原子会分成离散的两束粒子，这里可以不去纠结磁场与粒子作用具体的原理或者过程，根据经典物理中的观点，一

束粒子中各粒子的“取向”（这里用取向代指自旋取向）应该是均匀的，所以我们得到的图谱应该是连续的一条线段，但在实验中却是分离的、离散的两束。这说明在银原子内部的自旋的取向只能取离散的两个值，Figure 5.右侧的图是一组实验装置，后面我们会大篇幅的讲解施特恩—格拉赫实验和量子态，这里只需一个直观的感受即可。从普朗克研究黑体辐射发现能量量子化的现象到施特恩—格拉赫实验已经过去了 20 年，在这 20 年里，物理学家们发现了很多量子化现象，但是在理论上还不能完全解释他们，我们唯一有的就是玻尔的氢原子驻波模型，但是这个模型还是十分经典的。后来一个叫做德布罗意的年轻人，他想既然波可以像物质一样，那么物质是不是也具有波的性质呢？于是他提出了物质波的概念并且在 1924 年当作博士论文提交了，他的导师朗之万也很难理解，于是把他的论文交给了爱因斯坦，爱因斯坦看后赞叹道：“他已经掀起了面纱的一角！”。整个 20 世纪 20 年代是量子力学发展迅速的十年，这十年中，海森堡到哥廷根大学访问学习，而哥廷根大学有玻尔在坐镇，并且泡利（Pauli）在跟着玻尔学习，形成了哥本哈根学派。在 1925 年，海森堡（Werner Heisenberg）、波恩（Max Born）和乔丹（Pascual Jordan）创立了矩阵力学，他们认为，我们根本不用去纠结真正的机理是什么，我们只需要解释我们的观测值就好了，只要解释我们可观测的东西就好。之后就是大家很熟悉的历史，薛定谔在拿着德布罗意的论文度假的时候想着，既然物质

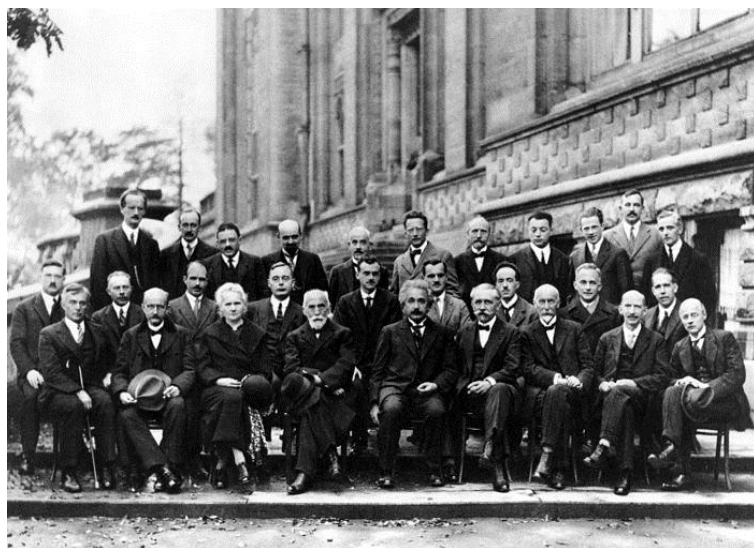


Figure 7 5th Solvay onference(1927) (来自：wikipedia)

有了波，我就给他个波动方程好了。后来泡利利用海森堡的矩阵力学解出了氢原子的定态解，薛定谔利用自己的波动方程也解出了一样的解，然后大家就都不淡定了！量子力学来到了辉煌的时代，泡利发现了电子有了自旋，于是就有了泡利矩阵（有机会后面会谈到的），

海森堡推出了不确定原理（Uncertainty principle），狄拉克（Dirac）着手研究相对论的量子力学。第五次索尔维大会也在那一年召开。之后就有了我们物理考试必拜的一张图片。量子力学在解决问题的时候从来没有错过，所以在当时，所有的物理学家都承认量子力学，只是他们并不能很好的理解量子力学，这其实和物理学的发展很相似，大家都是小打小闹的猜出、提出一些解释现象的理论，到后面才慢慢的发展出能够解释物质本质的理论。我们对量子力

学发展的历史的解释到这里就告一段落了，下面我们来讲讲怎么去入门地理解量子力学。没关系，费曼曾经说过：“I can safely said (that) no body understands quantum theory.”。也许我们并不能完全的理解量子力学，但在学习它的过程中，我们会收获到很多有趣的、有价值的东西。

II. 数学基础

虽然学习量子力学并不需要很强大的数学功底，但还是需要一部分数学基础的，如果对线性代数十分了解的话那么就比较容易了。这里主要补充些从欧几里德空间到希尔伯特空间的一些内容，还包括一小部分关于爱因斯坦求和记号的介绍。这些工具都能够在很大程度上帮助理解量子力学或者简化计算。

第一个想向大家介绍的数学概念是**空间**。可以把空间简单的想象成为一间“房间”，在空间中容纳东西就是能够进行运算的东西，在同一个空间中，这些运算都统一遵循着相同的代数法则。我们最常见的也是极其重要的空间叫做线性空间（Linear Space）或者向量空间。大家都有这样的直觉，如果想描述某个空间中的某个运动或者状态就必须有合适的数学语言，这便是基底（Base）。例如：我们想描述某个点的位置，我们就要先选择一个合适的坐标系，然后只需知道这个点与原点的相对方向和距离我们就可以很快的找出这个点了，这就是矢量的概念，当然，这些都很小儿科。更一般的，为了适应描述不同空间中的变换、运动，我们可以把矩阵、函数抽象成为基底。例如：我们可以定义一个多项式空间，它的基底是无限多的，由 $\{1, x, x^2, x^3, x^4, \dots\}$ 组成，在这个空间中的某个“点”，可以这样来表示：

$$f(x) = a_0 + a_1x + a_2x^2 + a_3x^3 + \dots$$

它与我们平时看到的欧几里德空间中的矢量 $\mathbf{v} = a\mathbf{i} + b\mathbf{j} + c\mathbf{k}$ 是不是很像？还有一个熟悉的例子是傅立叶级数（Fourier Series），只不过我们选取的基底是 $\{1, \cos x, \cos 2x, \cos 3x, \sin x, \sin 2x, \sin 3x, \dots\}$ 我们的矢量（函数）就可以表示成：

$$f(x) = \sum_0^{\infty} (a_n \cos nx + b_n \sin nx)$$

甚至我们可以把一个 2x2 矩阵（举例说明，更高维度也可以）表示成这样：

$$\mathbf{A} = a_0 \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} + a_1 \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} + a_2 \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} + a_3 \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$$

如果仅仅知道一个矢量可以分解成基底的线性和还不够，我们真正关心的是基矢量前面的系数是多少？回想一下矢量是怎么用基底表示的？比如说在欧几里德空间中，我们有矢量

\vec{A} , 和基底 $v_1, v_2, v_3, v_4, \dots$, 所以我们可以把 \vec{A} 表示为

$$\vec{A} = a_1 v_1 + a_2 v_2 + a_3 v_3 + \dots$$

其中

$$a_n = \frac{\langle \vec{A}, v_n \rangle}{\langle v_n, v_n \rangle}$$

\langle, \rangle 表示求两个矢量的内积。我们可以将这样的性质推广, 于是就有了**内积空间** (inner product space) 的定义:

定义: 一个向量空间 V 上的**内积** (inner product) 为 V 上的运算, 它将 V 中的向量 x 和 y 与一个实数 $\langle x, y \rangle$ 关联, 并满足以下条件:

- I. $\langle x, x \rangle \geq 0$, 等号成立的条件是 $x=0$.
- II. 对 V 中所有的 x 和 y , 有 $\langle x, y \rangle = \langle y, x \rangle$.
- III. 对 V 中所有的 x, y, z 及所有的标量 a 和 b , 有 $\langle ax + by, z \rangle = a \langle x, z \rangle + b \langle y, z \rangle$.

所以我们可以向量空间 $C[a, b]$ 定义内积为:

$$\langle f, g \rangle = \int_a^b f(x)g(x)dx$$

可以根据上面对内积空间的定义来证明上面的式子满足内积空间的定义。

类似与欧几里德空间内的单位向量 $(1, 0)$ 或 $(0, 1)$, 我们可以找到向量空间 C 相应的规范正交集 (Orthonormal Set) $\{1, \cos x, \cos 2x, \cos 3x, \sin x, \sin 2x, \sin 3x, \dots\}$, 并且可以用它来逼近一个周期为 2π 的连续周期函数 $f(x)$, 写成

$$f(x) = \sum_0^{\infty} (a_n \cos nx + b_n \sin nx)$$

所以我们可以根据内积的定义求出系数 a_n, b_n .

$$a_0 = \frac{\langle f, 1 \rangle}{\langle 1, 1 \rangle} = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) dx$$

以及

$$a_n = \frac{\langle f, \cos nx \rangle}{\langle \cos nx, \cos nx \rangle} = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \cos nx dx$$

$$b_n = \frac{\langle f, \sin nx \rangle}{\langle \sin nx, \sin nx \rangle} = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \sin nx dx$$

这样一来, 就非常容易理解傅立叶展开了, 而且利用相同的手法可以证明我们所选的基底是相互正交的, 类似这样的空间叫做**希尔伯特空间** (Hilbert Space)。我们还可以推广到复数域中, 只不过讨论复数时我们的问题在**复希尔伯特空间** (Complex Hilbert Space) 中。这里我省略了很多东西, 比如范数 (norm), 线性赋范空间 (Normed Linear Space) 等等一些的

说明，因为我们只需要理解到这里就够了。这里主要的目的是想让大家明白很多东西都可以抽象成为矢量，并且我们可以根据需求定义内积等等概念，这些思想对后面量子力学的态矢量的理解有着十分重要的作用。上面仅仅是对线性代数的一瞥，有个很严重的问题就是学校教我们的线性代数总是从行列式开始，甚至从始至终都没有告诉我们线性代数的核心是线性空间和线性变换。这一切都是十分有趣的东西，我也很想写很多东西，只不过我们的重点是量子力学而不是线性代数。到后面的特征值，奇异值分解等等概念都是人类发明的伟大的工具，而这些工具在很多领域都有广泛的应用，最贴近生活的 JPEG2000 是电脑存储图片的基于小波变换的一种算法等等，这里就不过多废话了。

下面简单介绍一下爱因斯坦求和标记，一种很实用的矢量运算工具。如果你在之前接触过一些场论，那么你一定对矢量微积分的印象十分深刻，比如说 nabla 算符，laplace 算符等等。你一定见过这样的公式：

$$\vec{\nabla} \times \vec{\nabla} \times \vec{A} = \nabla(\vec{\nabla} \cdot \vec{A}) - \nabla^2 \vec{A}$$

想象一下如果只是简单的按照行列式求矢量叉乘的方法，这样的等式要化简的话要等到什么时候啊，于是大名鼎鼎的爱因斯坦发明了这种求和助记符，我们根本不需要写出矢量的全部维度来进行计算，因为所有的方向都是轮换对称的，我们仅仅知道一个方向就可以了！我们先来回忆一下有关“场”的概念，物理学家把某一物理量在某一区域中的分布成为场。如果物理量是标量，那么称作标量场；如果是矢量就叫做矢量场，比如说：

温度在某一特定区域内的分布叫做温度场 $T(\mathbf{r})$ ，它是一个标量场

密度在某一特定区域内的分布叫做密度场 $\rho(\mathbf{r})$ ，它是一个标量场

速度在某一特定区域内的分布叫做速度场 $\mathbf{v}(\mathbf{r})$ ，它是一个矢量场

诸如此类的还有很多。还有一个比较常见的算符，nabla 算符 $\vec{\nabla} = \left(\frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z} \right)$ ，以及相应的运算公式

$$\begin{aligned}\vec{\nabla} \times \vec{V} &= \left(-\frac{\partial}{\partial z} V_y + \frac{\partial}{\partial y} V_z, -\frac{\partial}{\partial x} V_z + \frac{\partial}{\partial z} V_x, -\frac{\partial}{\partial y} V_x + \frac{\partial}{\partial x} V_y \right) \\ \vec{\nabla} \cdot \vec{V} &= \frac{\partial}{\partial x} V_x + \frac{\partial}{\partial y} V_y + \frac{\partial}{\partial z} V_z\end{aligned}$$

为了方便表示我们把矢量写成标量的形式：

$$\begin{aligned}\vec{u} &= (u_x, u_y, u_z) \rightarrow u_i (i = x, y, z) \\ \vec{u} + \vec{v} &= \vec{k} \rightarrow u_i + v_i = k_i \\ \vec{u} \cdot \vec{v} &= \sum u_i v_i \rightarrow u_i v_i\end{aligned}$$

上式右侧的表示方法就是爱因斯坦指标求和的方法，在与标量不发生混淆的时候我们完

全可以用这种简单的方式来进行演算。当情况比较复杂的时候，比如说遇到叉乘怎么办？这时我们引进两种符号 Kronecker 符号和 Epsilon 符号

$$\delta_{ij} = \begin{cases} 1, i = j \\ 0, i \neq j \end{cases} \quad \varepsilon_{ijk} = \begin{cases} 1, ijk = xyz, zxy, yzx \\ -1, ijk = yxz, xzy, zyx \\ 0, else \end{cases}$$

这样我们的点乘和叉乘可以分别表示为

$$\vec{u} \cdot \vec{v} = \sum u_i v_i \rightarrow \delta_{ij} u_i v_j$$

$$\vec{u} \times \vec{v} = (-u_z v_y + u_y v_z, -u_x v_z + u_z v_x, -u_y v_x + u_x v_y) \rightarrow \varepsilon_{ijk} u_j v_k$$

这样我们就可以用研究标量的方法研究矢量了！下面还有一个比较重要的公式用来化简做两次叉乘运算的式子，仔细观察一下下面式子的指标就可以很容易的记住它的形式，这对我们进行矢量的微积分运算是很有帮助的：

$$\varepsilon_{ijk} \varepsilon_{ipq} = \delta_{jp} \delta_{kq} - \delta_{jq} \delta_{kp}$$

通过爱因斯坦指标求和这一强大的工具，我们就可以化简许多场论中出现的公式了。比如说我们来利用麦克斯韦方程组把**真空中的光速**求出来：

$$\begin{cases} \vec{\nabla} \cdot \vec{E} = 0 \\ \vec{\nabla} \times \vec{E} = -\frac{\partial \vec{B}}{\partial t} \\ \vec{\nabla} \cdot \vec{B} = 0 \\ \vec{\nabla} \times \vec{B} = \mu_0 \varepsilon_0 \frac{\partial \vec{E}}{\partial t} \end{cases}$$

利用我们之前介绍的助记符表示就是：

$$\begin{cases} \delta_{ij} \partial_i E_j = 0 \\ \varepsilon_{ijk} \partial_j E_k = -\frac{\partial B_i}{\partial t} \\ \delta_{ij} \partial_i B_j = 0 \\ \varepsilon_{ijk} \partial_j B_k = \mu_0 \varepsilon_0 \frac{\partial E_i}{\partial t} \end{cases}$$

对 $\vec{\nabla} \times \vec{E} = -\frac{\partial \vec{B}}{\partial t}$ 等式两边同时再求一次旋度得到：

$$\varepsilon_{msi} \partial_s \varepsilon_{ijk} \partial_j E_k = -\varepsilon_{msi} \partial_s \frac{\partial B_i}{\partial t}$$

整理得

$$\varepsilon_{ims} \varepsilon_{ijk} \partial_s \partial_j E_k = -\frac{\partial(\varepsilon_{msi} \partial_s B_i)}{\partial t}$$

可以化简为：

$$(\delta_{mj} \delta_{sk} - \delta_{mk} \delta_{sj}) \partial_s \partial_j E_k = -\mu_0 \varepsilon_0 \frac{\partial^2 E_i}{\partial t^2}$$

还原到矢量微分方程即为：

$$\nabla(\vec{\nabla} \cdot \vec{E}) - \nabla^2 \vec{E} = -\mu_0 \epsilon_0 \frac{\partial^2 \vec{E}}{\partial t^2}$$

由于在真空中， $\vec{\nabla} \cdot \vec{E} = 0$ ，所以

$$\nabla^2 \vec{E} - \mu_0 \epsilon_0 \frac{\partial^2 \vec{E}}{\partial t^2} = 0$$

这样，我们就得到了真空中电磁波传播时，电场分量的波动方程，易知 $\frac{1}{c^2} = \mu_0 \epsilon_0$ 。所以

在真空中，电磁场的传播速度，或者光速为

$$c = \frac{1}{\sqrt{\mu_0 \epsilon_0}}$$

回顾一下，在这一小节中，我们主要介绍了不同的空间，把欧几里德空间中的矢量推广到了希尔伯特空间中，并介绍了一种化简矢量微分方程的爱因斯坦求和指标的用法。有了数学工具，下面让我们进入到量子力学的世界中。

III. 电子干涉实验

在这一小节中，主要为了给大家体会量子理论的一些思想，我们为什么把粒子既看作是实物粒子又看作是波。介绍研究量子力学时用到的工具左矢（Bra）和右矢（ket），以及态叠加原理。

在前面的介绍中我们知道了德布罗意提出了物质波的理论，那么如何来证明粒子具有波的属性？或者说波与宏观物质有什么不同？波很重要的一个性质就是可以发生衍射和干涉，历史上为了证明光是一种波，托马斯·杨利用了双缝干涉实验证明了光是一种波。那么我们来复制这样的实验，我们只要证明实物粒子可以发生衍射和干涉就好了。挑选实验材料是一种考验，因为根据德布罗意的假设 $\lambda = h/p$ ，这说明我们应该选用尽可能质量比较小的物体

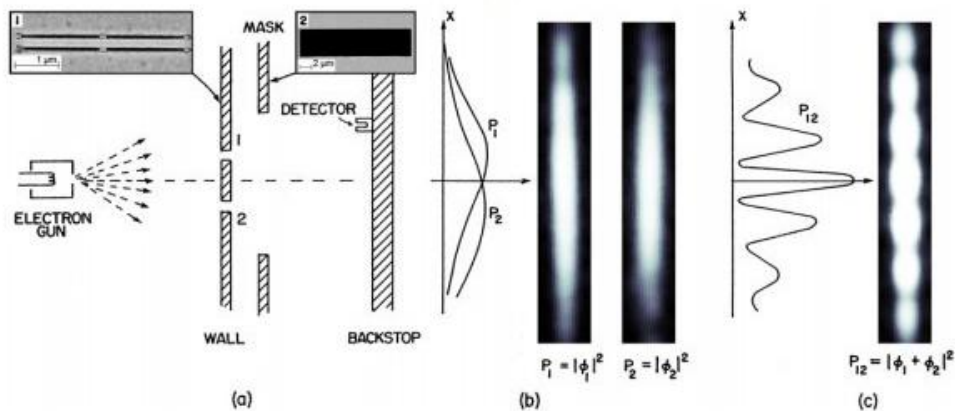


Figure 8 电子干涉实验

来做衍射或者干涉实验，因为如果波长太短，意味着我们没办法找到合适的材料来做双缝或者单缝。所以我们选择电子来作为实验材料。

在上面这幅图中，(a) 是实验装置示意图，(b) (c) 是在实验过程中，电子打到荧光屏上的图像。现在我们来考虑一种简单的情况，如果我们挡住狭缝 1 或者狭缝 2，那么我们会得到如图 (b) 所示的图像 P1 或者 P2，我们看到电子似乎是发生了衍射现象，电子在穿过狭缝后，运动方向有无数种可能，但是到达荧光屏上的概率最大的地方应该在狭缝附近，无论你是遮挡了狭缝 1 还是狭缝 2，我们得到的实验结果就是这样。可是当狭缝没有被遮挡的时候呢？直觉上我们会认为得到的图像应该是 $P1+P2$ ，就像图 (b) 左边的图那样，它有两个极大值，分别在狭缝 1 和狭缝 2 附近。可是真正的实验结果确实图 (c) 所示的图像，它有一系列的极大极小值。它的确发生了干涉现象！这意味着我们必须改变我们以前在经典力学中的思维，我们必须寻求一套新的理论来描述这样的现象。在研究光波的性质的时候，我们曾经用光强来描述干涉现象：

对于来自狭缝 1 的光波，振动的位移方程为 $h_1 e^{i\omega t}$ ，强度正比于 $I_1 = |h_1 e^{i\omega t}|^2$ ，来自狭缝 2 的光波，振动的位移方程为 $h_2 e^{i\omega t}$ ，强度正比于 $I_2 = |h_2 e^{i\omega t}|^2$ 。由于人眼所感知的恰恰是光的强度，我们并不能清楚光子或者光波的振动情况，我们只能感知到光强的平均值，所以我们才能够发现光波的干涉 $|h_1 e^{i\omega t} + h_2 e^{i\omega t}|^2$ ，正是因为其中的相位因子导致了干涉图像不同位置光强的不同。如果我们用粒子的观点去看待光的干涉现象，那么屏幕上光强的不同恰恰代表了光子到达该位置概率的平均值。这个地方比较“亮”，说明在单位时间内到达该点的光子比较多，光子到达该点的概率大于其它点。沿着这个思路一直下去，我们仿照光强定义一个新的物理量叫做概率幅 (Probability amplitude)，用字母 ϕ 来表示，有时也会用 ψ 来表示。 $|\phi|^2$ 表示该点粒子到达概率的大小。所以利用概率幅的思想，我们可以用它来描述电子干涉的现象。电子穿过狭缝 1 后打到屏幕上的概率密度用 $|\phi_1|^2$ 来表示，电子穿过狭缝 2 后打到屏幕上的概率密度用 $|\phi_2|^2$ 来表示，于是，在双缝干涉中，我们可以用 $|\phi_1 + \phi_2|^2$ 来表示电子打到屏幕上的概率密度。从这个实验中，我们窥探到了一些重要的东西，概率幅与波的振幅一样，是可以进行线性叠加的，以及导致干涉现象产生的是概率幅所具有的相位因子。这些都是我们从实验中的现象所推断出来的，我们只能去承认，因为

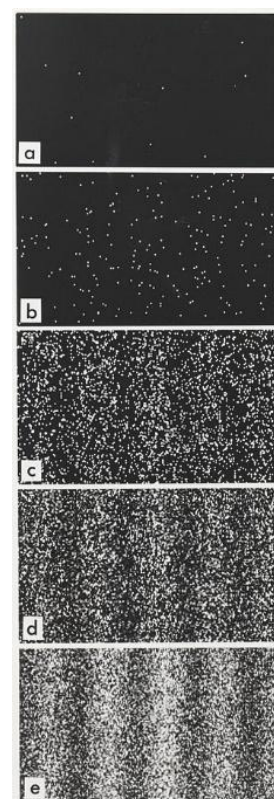


Figure 9 Double slit experiment of electrons

这样的思想在其它的量子实验中都没有错过，电子内部可能会有某些奇特的东西来控制电子的相位，我们对它一无所知，就像牛顿提出经典力学的三定律中，我们并不知道力只能改变物体的加速度，力是什么，为什么可以线性叠加，我们依然一无所知，我们只能利用数学工具来总结我们所看到的现象。关于电子的相位因子是什么并不重要，到后面我们会逐步解释这个东西，它可能是自旋，也可能是某种物理量的取向所表现出的宏观效果。

总结一下，在电子的干涉实验中，电子打到屏幕上的概率密度由一个复数 ϕ 的模长平方给出，当电子打到屏幕上由几种不同的方式组成时，这个概率密度等于这几种概率幅线性叠加后的模长平方给出。

$$\begin{cases} \phi_{\text{总}} = \phi_1 + \phi_2 \\ P = |\phi_1 + \phi_2|^2 \end{cases}$$

科学家们把概率幅抽象成为粒子的状态也叫**态 (state)**，拿电子干涉的实验来说，因为从狭缝 1 和狭缝 2 出来的电子所具有的概率幅有一个相位因子的差别，所以两种方式出射的电子处于两种不同的态。态满足态叠加原理，在不同的坐标系中，态可以表示为不同基底的线性组合，一般来说，我们通常把基底选为粒子态在这种坐标系下的本征态(Eigen state)，这和线性代数中的特征向量(Eigen vector)是一样的。我们习惯用本征态来合成粒子的态，这里我们只要明白在量子力学中，我们描述粒子的运动用态来描述，并且态是满足线性叠加原理的，像直角坐标系中那样，我们习惯用正交的本征态作为基底。

我在前面讲过，要用线性代数的思想去理解量子力学，我们刚刚讲到了用态的基底来合成态，既然涉及到了分解，就一定要规定**内积**。因为 ϕ 是一个复变量，我们要像傅立叶展开那样把它与基底(另外一个复变量)进行点乘，才能得到展开系数。我们还要拿线性代数的那一套理论来说事：

在向量空间中，我们把矢量这样分解：

$$\begin{aligned} \vec{A} &= a_1 \vec{v}_1 + a_2 \vec{v}_2 + a_3 \vec{v}_3 + \dots \\ a_n &= \frac{\langle \vec{A}, \vec{v}_n \rangle}{\langle \vec{v}_n, \vec{v}_n \rangle} \end{aligned}$$

那么，类似的在态矢空间中，我们这样分解：

$$\begin{aligned} \Phi &= a_1 \phi_1 + a_2 \phi_2 + a_3 \phi_3 + \dots \\ a_n &= \frac{\langle \Phi, \phi_n \rangle}{\langle \phi_n, \phi_n \rangle} \\ \langle \phi_n, \Phi \rangle &= \int_a^b \phi_n^*(x) \Phi(x) dx \end{aligned}$$

为了在态矢空间中，我们的态表达的清晰明了，科学家们(如果没记错的话应该是狄拉

克) 根据泊松括号 (Poisson bracket), 利用叫做 Bra-ket 的方式来描述态。比如说某个态 ϕ 用右矢 (ket) 来表示 $\phi = |\phi\rangle$, 它的复共轭用左矢 (bra) 来表示 $\phi^* = \langle\phi|$, 那么它在某点出现的概率密度, 也就是自己与本身的点积记为 $\langle\phi|\phi\rangle$. 其实 bracket 本来就是括号的意思啊, 有个比较有意思的梗是: 物理学院有很多同学不能正确拼写括号的英文单词 (这里没有嘲讽的意思啊)。

IV. 施特恩—格拉赫实验

在上面一节中我们介绍了电子的干涉实验, 以及用概率幅或者态来描述粒子的运动性质, 并且有了 bra-ket 的表示方式。有了这些我们就可以进阶的来理解施特恩—格拉赫实验, 这个实验被称为是“最量子”的实验, 因为它几乎没有经典的理论相对应。

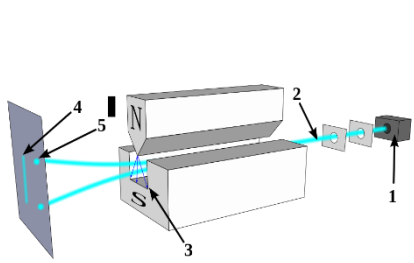


Figure 20 施特恩—格拉赫仪器

施特恩—格拉赫仪器简单来讲是一个由非均匀磁场所构成的一个孔道, 粒子穿过孔道的时候会与非均匀磁场发生作用 (注意这里的粒子是电中性的), 从而发生偏折, 其实如果这个实验更早一些就可以用来证明 AB 效应了, 至于非均匀磁场如何与粒子发生作用这里就不做过多解释了, 因为在这里我们几乎用不到这样的细节。我们只需要清楚施特恩—格拉赫仪器能够把一束原子劈裂成离散的两束, 如果我们仪器的磁场变化的方向是在 z 轴方向我们把仪器叫做 S_z , 那么粒子束就被劈裂成沿着 z 轴方向的离散的两束, 我们把两束粒子分别叫做 z_+ 和 z_- , 同理如果是 x 轴方向的仪器就叫做 S_x , 粒子束就会被劈裂成沿着 x

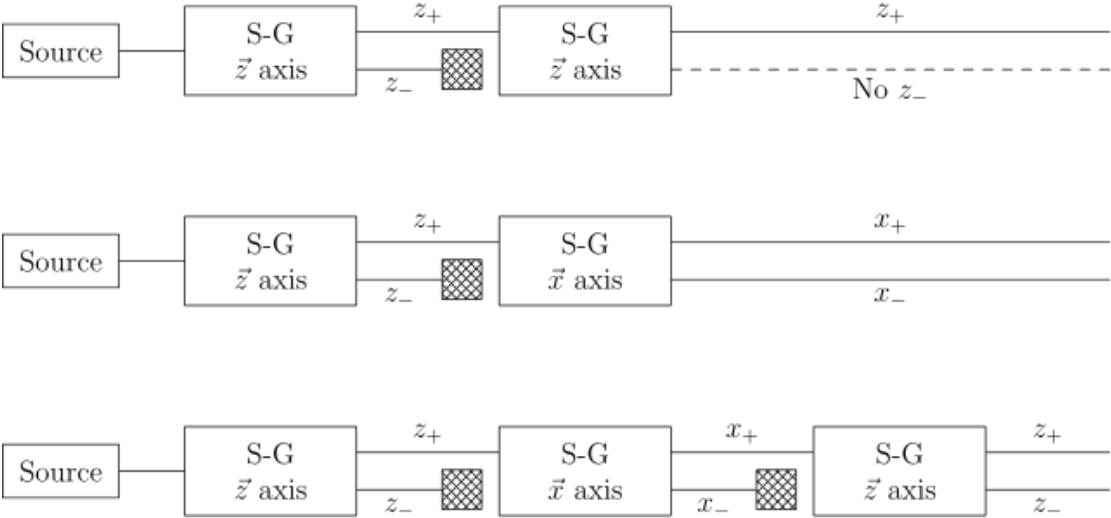


Figure 11 施特恩—格拉赫实验结论

轴方向的离散的两束，叫做 $x+$ 和 $x-$ ， y 轴亦同。理论上来说，我们可以把粒子送入一个实验装置，我们也可以把粒子送入串联的几个实验装置中。让我们来看一下实验结果：

实验是比较好理解的，在第一个实验中，粒子通过一个 z 轴方向的装置被劈裂成两束粒子 $z+$ 和 $z-$ ，接着我们把 $z-$ 的部分遮住，把 $z+$ 的部分继续通入 z 轴方向的装置，发现 $z-$ 的部分没有出现，只有 $z+$ 的部分。第二个实验中我们把劈裂成 $z+$ 和 $z-$ 的粒子束去除 $z-$ 的部分通入 x 轴取向的装置，发现出现了 $x+$ 和 $x-$ 的两束。第三个实验中，我们把第一个实验装置产生的 $z-$ 部分丢掉，把 $z+$ 的部分通入 x 轴取向的装置，此时我们得到了 $x+$ 和 $x-$ 的粒子束，我们继续把 $x-$ 的部分丢掉，把 $x+$ 的粒子束再通入一个 z 轴取向的装置中，奇怪的现象发生了，我们又得到了 $z-$ 的部分的粒子。在经典力学中我们用数来描述物体的各种性质比如说位置、动量、速度等等，它就是一个确定的数值，这些数字叠加是没有意义的。但在量子力学中，我们没有办法用确定的一个数去描述一个粒子，我们只能说它处于某种状态。有了态这个物理语言或者工具我们可以更方便的来描述施特恩-格拉赫实验。对于量子力学中奇怪的现象，我们只能先去接受他，我不赞同用经典力学中的经典量去与量子力学中的某些性质去类比，因为我们从来没有接触过微观粒子，我们所积累的仅仅是经典力学中的一些经典思想和一些偏见，微观世界对我们来说是一个全新的领域，我们必须像最初接受经典力学那样来接受量子力学的思想。

这个实验告诉我们原子的自旋取向只能有两个方向而且是离散的，换句话说自旋只有两个态，这两个态具体是什么不重要，它可能是 $|z + \rangle$ 或者 $|z - \rangle$ ， $|x + \rangle$ 或者 $|x - \rangle$ ， $|y + \rangle$ 或者 $|y - \rangle$ 。而且在第三个实验中，我们从 $|x + \rangle$ 中又分出来了 $|z + \rangle$ 和 $|z - \rangle$ ，这说明 $|x + \rangle$ 是 $|z + \rangle$ 和 $|z - \rangle$ 的一种叠加态。我们来尝试建立原子自旋方向中 x 和 z 取向的一种关系，我们可以这样写出表达式：

$$|x + \rangle = a|z + \rangle + b|z - \rangle$$

$$\langle x + | x + \rangle = a^2 \langle z + | z + \rangle + b^2 \langle z - | z - \rangle + 2ab \langle z + | z - \rangle$$

根据概率论的知识，概率密度最终是要归一化的，所以

$$\langle x + | x + \rangle = 1$$

$$\langle z + | z + \rangle = 1$$

$$\langle z - | z - \rangle = 1$$

而当我们把 $|z + \rangle$ 通入 z 取向的实验装置中，最终出来的并没有 $|z - \rangle$ ，所以 $|z + \rangle$ 中并不含有 $|z - \rangle$ ，或者说 $|z + \rangle$ 和 $|z - \rangle$ 是互不相干的，所以：

$$\langle z + | z - \rangle = 0$$

$$\langle x+ | x- \rangle = 0$$

$$\langle y+ | y- \rangle = 0$$

代入上面的式子可得：

$$\langle x+ | x+ \rangle = a^2 \langle z+ | z+ \rangle + b^2 \langle z- | z- \rangle = a^2 + b^2 = 1$$

由于我们从 $|x+ \rangle$ 中分出来了 $|z+ \rangle$ 和 $|z- \rangle$ ，而且各占一半，所以

$$a = b = \frac{\sqrt{2}}{2}$$

我们就得到了 x 取向与 z 取向的关系：

$$|x+ \rangle = \frac{\sqrt{2}}{2} |z+ \rangle + \frac{\sqrt{2}}{2} |z- \rangle$$

同理我们还可以得到类似的一些等式：

$$|x- \rangle = \frac{\sqrt{2}}{2} |z+ \rangle - \frac{\sqrt{2}}{2} |z- \rangle$$

$$|y+ \rangle = \frac{\sqrt{2}}{2} |z+ \rangle + \frac{\sqrt{2}i}{2} |z- \rangle$$

$$|y- \rangle = \frac{\sqrt{2}}{2} |z+ \rangle - \frac{\sqrt{2}i}{2} |z- \rangle$$

这里也许你会奇怪为什么出现了虚数单位 i ，这是因为我们想用二维的东西来表示三维空间中的东西，代价就是多出来一个虚数单位来描述另外一个维度，这些可以参考群论中 SU_2 群与 SO_3 群。到这里我们所干的事情就是利用态矢，以及态的叠加来解释我们所看到的现象，这些都是无法避免的，客观的实验结论摆在了我们面前，我们只能用态和复数的方式去解释这样的现象。

讲完了施特恩—格拉赫实验，我们来谈谈可观测量的概念。不知道有没有注意到我在开头讲海森堡发明矩阵力学的那一部分内容，海森堡想：我们根本不必去关心粒子到底处在什么样的状态，我们只要关心我们能观测到的东西就好了。所以这里我们来讲讲**算符和可观测量**的问题。

无论是宏观世界还是微观世界中，一个物理量如果不能被改变，它就一定是不可观测的。举一个不是很恰当的例子来说，当别人扔给你一个篮球，无论你是接住了它还是被它砸到，你都改变了它的运动状态，所以说有关这个篮球的一些物理量就被你“观测”到了。在量子力学中我们把能改变粒子状态的算符叫做可观测量（observable）。当对一个量子系统进行观测的时候，这个量子系统一定会发生不可逆转的改变。就像我们在研究电子干涉时那样，如果我们人为的去考察每个电子的运动状态，比如说用光去照射每个孔，当电子穿过就会有一个闪光，这样我们就能知道每个电子从哪个孔穿过，进而知道干涉图样是怎么产生的了，可

是不要忘记，根据普朗克的理论每个光子是携带一定能量的 $\varepsilon = h\nu$ ，这样每个电子就会被光子所干扰，干涉图样就再也不会出现了，当把光源撤掉的时候，干涉图样就又神奇的出现。也许你会说，让我们来把光源换成一个红外光，这样光子的动量就会变小，对电子的影响就会比较小，事实的确是如此，在这种情况下，电子的衍射图样就又会重新出现，可是由于光波的波长变长，这时不管狭缝 1 或者狭缝 2 穿过的电子都会引起一个小小的、模糊的闪光，以至于根本无法分辨狭缝 1 和狭缝 2 了。

在每次进行观测的时候都会引起这个量子系统坍缩到这个系统的本征态，在电子干涉实验中，如果我们对狭缝 1 进行观测，意味着我们要将狭缝 2 进行遮挡，此时电子的叠加态就会自动的坍缩到狭缝 1 的本征态上，同理对狭缝 2 也是这样。于是我们可以用一个数学公式来表示这样的情况：

$$\hat{O}|\phi\rangle = \lambda\phi$$

\hat{O} 是一个可观测量，或者说是一个算符， λ 是一个实数，代表了系统的本征值（通常是某个可观测量在观测时间内的平均值，到后面会继续说明）。所有的可观测量都属于厄米算符（Hermitian），这与厄米矩阵的定义是类似的，而且所有经过厄米算符得到的态都是正交的本征态，因为篇幅限制，这里就不过多赘述了，可以参考 Griffiths 的 *Introduction to Quantum Mechanics*，关于算符和可观测量的介绍就到这里。

V. 薛定谔方程

在得到薛定谔方程之前，首先声明一点，薛定谔方程完全是薛定谔根据波动方程所拼凑出来的公式，所谓推导不过是把波动方程与德布罗意波的公式拼凑出来的一个公式，在很多教科书中都可以找到所谓的推导过程。下面来介绍一种能够容易理解的含时薛定谔方程。我们首先来考虑一个算符 $U(t_2, t_1)$ ，它是一个延时算符，它的作用是把时间从 t_1 拖延到 t_2 ，当它作用到某一个量子态上时，效果是这样的：

$$|\phi(t + \Delta t)\rangle = U(t + \Delta t, t)|\phi(t)\rangle$$

如果我们想研究一个不受外界干扰的，独自演化的量子系统，我们可以把一段时间分割为很多段，就像我们在微积分第一课中写成一个类似黎曼和的东西，这样就可以得到描述一个不受外界影响的量子系统演化的规律的方程了，也叫大名鼎鼎的含时薛定谔方程（Time-dependent Schrödinger equation）。它长这个样子：

$$i\hbar \frac{\partial}{\partial t} \psi(\vec{r}, t) = \hat{H}\psi(\vec{r}, t)$$

回到我们刚刚的话题,我们构造了一个延时算符 $U(t + \Delta t, t)$ 用来给一个系统延迟一个 Δt 的时间。我们将延迟后的系统投影到基底上来表示,所以我们给等式两边同时左乘一个 $\langle i|$,可以得到:

$$\langle i|\phi(t + \Delta t)\rangle = \langle i|U(t + \Delta t, t)|\phi(t)\rangle$$

同理,我们把初态也用基底来表示:

$$\langle i|\phi(t + \Delta t)\rangle = \sum_j \langle i|U(t + \Delta t, t)|j\rangle \langle j|\phi(t)\rangle$$

我们定义: $\psi_i(t) = \langle i|\phi(t)\rangle$, 代表在 t 时刻处于第 i 个基础态, 以及我们可以把 $\langle i|U(t + \Delta t, t)|j\rangle$ 写作矩阵的一项: $U_{ij}(\Delta t) = \langle i|U(t + \Delta t, t)|j\rangle$

原式可以表示成:

$$\psi_i(t + \Delta t) = \sum_j U_{ij}(\Delta t) \psi_j(t)$$

我们知道 $U_{ij}(\Delta t)$ 是一个关于 Δt 的函数, 而且我们很容易想到当 $\Delta t \rightarrow 0$ 时, 所有的态都没有被改变, 我们可以仿照泰勒展开那样, 写出 $U_{ij}(\Delta t)$ 的一级近似:

$$U_{ij}(\Delta t) = \delta_{ij} + K_{ij}\Delta t + o(\Delta t)$$

这里不得不根据历史上的一些习惯和其它的一些原因, 我们喜欢把 K_{ij} 提出一个 $(-i/\hbar)$ 的因子, 写成 $U_{ij}(\Delta t) = \delta_{ij} + (-i/\hbar)H_{ij}\Delta t + o(\Delta t)$, 对于为什么提出一个因子稍后会进行解答, 代回到原式可得:

$$\psi_i(t + \Delta t) = \sum_j \left[\delta_{ij} - \frac{i}{\hbar} H_{ij} \Delta t \right] \psi_j(t)$$

整理后可得:

$$\psi_i(t + \Delta t) - \sum_j \delta_{ij} \psi_j(t) = -\frac{i}{\hbar} \sum_j H_{ij} \Delta t \psi_j(t)$$

$$\frac{\psi_i(t + \Delta t) - \psi_i(t)}{\Delta t} = -\frac{i}{\hbar} \sum_j H_{ij} \psi_j(t)$$

$$\frac{\partial \psi_i(t)}{\partial t} = -\frac{i}{\hbar} \sum_j H_{ij} \psi_j(t)$$

这样我们从含时薛定谔方程所代表的物理意义出发, 得到了含时薛定谔方程, 我们所选取的总是一组固定的, 不随时间变化的基础态, 而且所有粒子的性质都由 H_{ij} 来描述, 之所以把 K_{ij} 提出一个 $(-i/\hbar)$ 因子后写成 $(-i/\hbar)H_{ij}$, 是因为 H_{ij} 算符通常对应的特征值是体系的能量, 这里就不做过多解释了, 很多参考书都会教你利用求导数的方法求得薛定谔方程, 里面

所有的参数对应的物理意义都比较明确，所以不妨自己试着把德布罗意波与波函数相结合，动手操作一下怎么去得到薛定谔方程。

VI. 不确定原理

在之前介绍算符和可观测量的部分我们知道，如果对一个量子系统进行测量，则必定会改变这个系统，除非它已经处在这个算符的本征态上，那么就存在这样一个问题，如果我想对一个系统同时观测两个物理量，怎么办？当我在测量第一个物理量后，系统坍缩到了一个本征态，这对第二个观测测量有没有影响？海森堡解决了这个问题。答案是互相**对易**的算子可以同时测量到精确值，而不对易的算子则不可。在证明不确定性原理之前，先让我们来看看科学家们是如何区分对易与不对易的。

定义：

设有 A, B 两个算符，它们的对易关系由 $[A, B] \equiv AB - BA$ 给出。

其实很容易理解，无非就是研究两种算符以不同次序作用后，对系统作用的区别有多大。所以，根据动量算符 $\hat{p} = -i\hbar \frac{\partial}{\partial x}$ ，和位置算符 x ，即可知道它们之间的对易关系：

$$[x, p] = xp - px = \frac{\left(x \left(-i\hbar \frac{\partial}{\partial x}\right) f - \left(-i\hbar \frac{\partial}{\partial x}\right) x f\right)}{f} = i\hbar$$

所以说位置和动量算符不是对易算符，所以当电子穿过一个狭缝时（位置被确定），它会向各个方向衍射（平行于狭缝动量的不确定性很大）。

现在，有了对易算符的概念，我们可以证明不确定性原理了，对于任意的可观测量 \hat{A} ，它的方差可以写成：

$$\sigma_A^2 = \langle (\hat{A} - \langle \hat{A} \rangle) \psi | (\hat{A} - \langle \hat{A} \rangle) \psi \rangle = \langle f | f \rangle$$

其中 $f \equiv (\hat{A} - \langle \hat{A} \rangle) \psi$ ，同理，我们可以构造另外一个可观测量 \hat{B} 的方差：

$$\sigma_B^2 = \langle (\hat{B} - \langle \hat{B} \rangle) \psi | (\hat{B} - \langle \hat{B} \rangle) \psi \rangle = \langle g | g \rangle$$

进行下一步前，让我们回忆一下柯西不等式（Cauchy-Schwarz Inequality）：

$$\int f(x) g^*(x) dx \leq \int |f(x)|^2 dx \int |g(x)|^2 dx$$

所以我们有：

$$\sigma_A^2 \sigma_B^2 = \langle f | f \rangle \langle g | g \rangle \geq |\langle f | g \rangle|^2$$

对于任意的复数我们有：

$$|z|^2 = [\text{Re}(z)]^2 + [\text{Im}(z)]^2 \geq [\text{Im}(z)]^2 = \left[\frac{1}{2i}(z - z^*)\right]^2$$

令 $z = \langle f|g \rangle$:

$$\sigma_A^2 \sigma_B^2 \geq \left[\frac{1}{2i} (\langle f|g \rangle - \langle g|f \rangle) \right]^2$$

注意到:

$$\langle f|g \rangle = \langle (\hat{A} - \langle \hat{A} \rangle) \psi | (\hat{B} - \langle \hat{B} \rangle) \psi \rangle = \langle \psi | (\hat{A} - \langle \hat{A} \rangle) (\hat{B} - \langle \hat{B} \rangle) | \psi \rangle$$

$$\langle f|g \rangle = \langle \psi | \hat{A} \hat{B} - \hat{A} \langle \hat{B} \rangle - \hat{B} \langle \hat{A} \rangle + \langle \hat{A} \rangle \langle \hat{B} \rangle | \psi \rangle$$

$$\langle f|g \rangle = \langle \psi | \hat{A} \hat{B} | \psi \rangle - \langle \psi | \hat{A} \langle \hat{B} \rangle | \psi \rangle - \langle \psi | \hat{B} \langle \hat{A} \rangle | \psi \rangle + \langle \psi | \langle \hat{A} \rangle \langle \hat{B} \rangle | \psi \rangle$$

$$\langle f|g \rangle = \langle \hat{A} \hat{B} \rangle - \langle \hat{B} \rangle \langle \hat{A} \rangle - \langle \hat{A} \rangle \langle \hat{B} \rangle + \langle \hat{A} \rangle \langle \hat{B} \rangle$$

$$\langle f|g \rangle = \langle \hat{A} \hat{B} \rangle - \langle \hat{B} \rangle \langle \hat{A} \rangle$$

同理:

$$\langle g|f \rangle = \langle \hat{B} \hat{A} \rangle - \langle \hat{A} \rangle \langle \hat{B} \rangle$$

所以:

$$\langle f|g \rangle - \langle g|f \rangle = \langle \hat{A} \hat{B} \rangle - \langle \hat{B} \hat{A} \rangle$$

所以我们的可以得到最终的不确定性原理:

$$\sigma_A^2 \sigma_B^2 \geq \left[\frac{1}{2i} (\langle \hat{A} \hat{B} \rangle - \langle \hat{B} \hat{A} \rangle) \right]^2 = \left[\frac{1}{2i} [\hat{A}, \hat{B}] \right]^2$$

现在, 我们就可以得到海森堡最初的不确定性原理了:

$$\sigma_x^2 \sigma_p^2 \geq \left(\frac{\hbar}{2} \right)^2$$

或者写成:

$$\sigma_x \sigma_p \geq \frac{\hbar}{2}$$

这里你会发现, 我们所常说的不确定关系其实不是 $\Delta x \Delta p \geq \frac{\hbar}{2}$, 而是它们的标准差的乘积!

确切的来说应该是协方差! 正是由于量子系统客观存在着不确定性关系, 我们才不能无限精确地得到动量和位置的信息。

VII. 结语

物理是一门很奇妙的学科, 它渗透到生活中的方方面面, 量子力学作为 20 世纪中物理这门学科的最大突破, 可以说没有了解量子力学, 与生活在 100 年前的人没什么不同。这篇小文章我利用了三天来完成, 基本所有的历史、证明方法都来自在维基百科, 费曼物理学讲义 (第三卷), 格里菲斯的量子力学概论和樱井纯的现代量子力学中找到, 本来量子力学发展历程那一部分可以复制粘贴一下的, 但是又觉得用自己的话叙述一遍可能会给自己找到新

的知识，不出所料，我又读了一遍海森堡发明矩阵力学时那一段有关可观测量的话，发现短短几句就把可观测量的物理学意义阐述的十分明白。本来还有算子法求解谐振子系统的方法，X 光散射研究晶体结构和倒格子的关系等等，有很多东西我想写却碍于没有时间，只能以后再说了，由于自己也是刚刚入门量子力学，所以有些地方未免不严谨，恳请指证。

参考书目：

1. 《量子力学概论》Griffths 著
2. 《费恩曼物理学讲义（第三卷）》 费恩曼著
3. 《现代量子力学》 樱井纯著
4. 《代数》Michael Artin 著
5. 《线性代数及其应用》David C.Lay 著
6. 所有图片均来自 wikipedia