计算机科学 COMPUTER SCIENCE
http://www.jsjkx.com
DOI:10.11896/jsjkx.190400432

量子计算与不确定性原理

Renata WONG
1 南京大学计算机科学与技术系　南京 210093
2 南京大学软件新技术国家重点实验室　南京 210093

摘　要　对量子计算的计算潜力的高度期望源于量子力学的各种特性,如叠加原理、纠缠现象、破坏性和建设性的量子干扰。相对于经典计算,量子计算具有某些假定的优势,例如量子算法的运行速度比经典算法快;但另一方面却似乎存在影响经典算法但不影响量子算法的障碍,障碍之一是传统上归因于 Werner Heisenberg 的两个不确定性原理。Heisenberg 最初制定的不确定性原理涉及用于测量量子系统的非量子仪器必然会对该系统造成影响。这个原理与其后来的发展有所不同,因为后来发现的不确定性所假定的是不交换可观察量在测量方面存在固有的不能精确测量的特性。在目前的技术发展状况以及当前对量子力学的形式表述与诠释的情况下,这两种不确定性皆有可能对量子计算的速度造成不良影响。近年来,针对这两种不确定性原理有了新的研究成果:1)Ozawa 对 Heisenberg 原理提出了修改,将两种不确定性纳入其内进行并列考虑,从而可以减小 Heisenberg 原理的不确定性程度;2)在考虑到熵不确定性的情况下,Heisenberg 不确定性可被视为 Hirschmann 不确定性的下界,因此除了在测量上的不确定性之外,量子计算还必须考虑来自其他如信息学的不确定性因素。

关键词:量子计算;不确定性原理;不确定性关系;熵不确定性
中图法分类号　TP3-0

Uncertainty Principle as Related to Quantum Computation

Renata WONG
1 Department of Computer Science and Technology, Nanjing University, Nanjing 210093, China
2 State Key Laboratory for Novel Software Technology at Nanjing University, Nanjing 210093, China

Abstract　The high expectations regarding the computational potential of quantum computation stem from quantum mechanical features, such as the principle of superposition, the phenomenon of entanglement, the destructive and constructive interference. Besides the presumed advantages of quantum computation over classical computation, there exist impediments that appear to be affecting the former but not the latter. One of them are the two uncertainty principles traditionally ascribed to Werner Heisenberg. The uncertainty principle formulated originally by Heisenberg pertains to the inability of measuring a quantum system with non-quantum instruments without affecting it. This principle is different from the later development postulating an inherent inability of non-commuting observables to be measured precisely. At present state of technological development and within the current formulation and interpretation of quantum mechanics, both versions of the uncertainty affect the speed attainable by a quantum computer. Recently, the two uncertainty principles have received more attention. In his improvement to Heisenberg's principle, Ozawa took into account both types of uncertainty mentioned above. Furthermore, research into entropic uncertainty has shown that Heisenberg's uncertainty can be seen as a lower bound of Hirschmann's uncertainty, thereby indicating that quantum computation may need to consider other types of uncertainties, such as information uncertainty, as well.

Keywords　Quantum computing, Uncertainty principle, Uncertainty relations, Entropic uncertainty

当前以随机存取机模型或图灵机模型为基础的经典计算机在理论上和实践上都遵循经典物理学的定律。自量子力学诞生起,物理学界就普遍认为量子物理定律为综合物理世界的理论,而经典力学定律必须从其中推导出来(由于量子物理等式包含经典物理未有的 Planck 常数 $\hbar$,学界一般尝试将极限值 $\hbar \to 0$ 应用于量子等式来推导出经典等式);换句话说,现代量子物理学认为宇宙是一个量子系统,其按照与经典物理学定律不同的量子物理定律运行。量子系统与经典系统有着不同的特征,包括:一个量子系统可同时处于多个经典态,即所谓的叠加态(superposition);一个量子系统在演化过程中可受量子干扰而产生所谓的量子干涉(quantum interference);在空间上隔离的两个量子系统之间亦可以有若干交互关系

到稿日期:2019-04-10　返修日期:2019-06-04
基金项目:国家重点研发计划(2019YFA0308700);江苏省自然科学基金(BK20191249)
This work was supported by the National Key R&D Program of China (2019YFA0308700) and Natural Science Foundation of Jiangsu Province (BK20191249).
通信作者:Renata Wong(renata.wong@protonmail.com)

(correlation),即所谓的纠缠关系(entanglement)。上述涉及基础理论的现象促进量子计算领域发展的主要动因,而非仅仅在于计算速度的快慢或加密能力的强弱。

在量子力学定律的基础上,量子计算研究量子计算机的计算能力、量子算法、计算复杂性等。这方面的研究始于20世纪80年代初期,当时的经典计算机模型为模拟计算机。最初的量子计算机模型亦为量子模拟计算机,由Manin[1],Feynman[2]和Benioff[3]同时期提出。冷战造成的阻隔,使得Manin出版于1980年的*Vychislimoe i nevychislimoe*(俄语,书名可翻译为《可计算与不可计算》)一书虽然先于Feynman和Benioff提出了量子模拟计算机的可能性,但到最近才获得了国际学界的认可。20世纪80年代,经典计算机模型逐步向数字计算模型转化,Deutsch在1985年设计了通用量子图灵机模型[4],从而建立了一个量子数字计算范式。

随后,20世纪90年代初Simon[5]和Deutsch等[6]提出了用于判定函数若干特性的量子算法,以及Bernstein等发展出来的量子复杂性理论[7]。但是,直到1994年Shor提出求解整数的质因子分解问题和离散对数问题的多项式量子算法[8],量子计算研究才有了明显的进度。Shor提出的这两种算法的重要性在于其在密码学上具有实际用途,并给目前经典计算系统普遍使用的加密算法(如RSA算法)带来了挑战。RSA是一个基于大数分解的公钥加密系统,能否解决其公钥的问题取决于所使用的数字能否在合理的时间内进行质因子分解。随着计算机处理器速度的提升,以及求解数字分解时间的缩短,公钥长度必须及时延长。目前,推荐公钥长度为2048 bits(等于617个十进制数位)或以上。因为尚未发现一种仅需多项式时间的经典分解算法,所以RSA的假设是求解质因子分解问题所需的时间应在多项式时间之上。一般认为,Shor的量子算法能够在多项式时间内基于一个量子计算机求解大数质因子分解问题;为了应对量子加密算法对经典加密技术所造成的挑战,目前学界已展开了以此为对象的经典密码学研究,即RSA后加密学(post-RSA cryptography)[9]。

除了上述的理论性动机,就应用而言,量子计算的主要驱动力来自两个方面:1)需要突破硬件的限制;2)希望在计算时间复杂度上超越经典算法,即对多项式以上的时间复杂度问题提供解决方法。

就硬件的限制而言,1965年Intel的Moore估计,因电子器件的体积逐渐变小,集成电路上的晶体管数量将每年增加一倍[10]。晶体管的稀有特性是管件越小其计算能力就越强,因为与大型晶体管相比,小型晶体管可以以更小的功率和更快的速度操作开关,这意味着使用更多、更快的晶体管无需更大功率或产生更多的余热。Moore对这种计算能力的指数式增长进行了预测,坊间非正式地称其为"摩尔定律"。Moore在1975年对摩尔定律进行了修改,认为晶体管数量将每两年增加一倍,这显示了芯片的微型化在过去10年内已明显减速。然而,2000年初,高端芯片的运行速度已进入坪曲线,而生产芯片的成本一直上升。其中一个有趣的观察是,电子组件正逐渐接近原子尺度,因此使储存与操作比特变得越来越难;再进一步的观察是,当电子器件的尺度变得与原子一样小时,其功能将受到量子物理效应的干扰,从而导致经典计算机无法正常运算。在这两个观察的基础上,对经典系统计算能力的指数式增长的预期将在未来10~20年内完全失效。这显然是发展量子计算的一个强烈动机。

从摩尔定律推出的结果被认为有两个:1)小型化使器件密度增加,导致计算机存储量亦随同增加,因此为运算提供了更大的储存空间;2)处理器的运行速率也有可能随之增加,使之能够加速计算过程[1)]。计算机科学家Sedgewick对此却另有异议。他指出,随着计算机的加速,现有的算法反而会减速。Sedgewick的逻辑推论如下:假设有一个时间复杂度为$O(n\log n)$的算法,其中n为问题的规模,如果将处理器速率增加一倍,便可以以$\frac{n\log n}{2}$的时间运行该算法;如果再将存储器空间加倍,同一算法所需的空间便有可能扩大至填满有限的储存器;原来的储存器只能处理n输入数据,现在则能处理$2n$数据,因此算法的运行时间即为$\frac{2n\log 2n}{2}=n\log n+n$,比原来的复杂度多了一个线性因素$n$。

就计算时间复杂度而言,经典计算机使用n bits来同时储存n个信息数字(information digits)或一个经典状态。而量子计算中的叠加原则指出,n个量子比特可用来同时储存2^n个信息数字或2^n个量子经典状态。假设有一个用于n个经典比特(一个经典状态)的函数f,则需要$T(1)=s$个时间单位,当将f应用于n个量子比特(2^n经典状态)时,运算所需时间同样是$T(2^n)=s$个时间单位,即在同一时段内,一个量子计算机能够同时处理2^n个值,而一个经典计算机只能处理n个值。这种内在并行性似乎表示使用量子系统计算问题比使用经典系统计算问题具有时间上的优势。但这个说法存在两方面的问题。1)一个并行算法的计算速度不一定比串行算法快。*NC*复杂度指能够在并行计算机上使用$O(n^k)$个处理器,并在综合对数(polylogarithmic)时间$O(\log^c n)$内求解的一类难题,其中n为输入规模,k和c皆为常数。*NC*并行性复杂度是计算复杂性理论中的P复杂性类(意即多项式时间算法)的一个子集合。正如复杂度科学中的$P=NP$问题为未知,$P=NC$的真假亦为未知。如果答案是肯定的,所有多项式时间算法都可以有并行性版本,并能够使用多项式数量的处理器来求解。2)但在20世纪70年代有学者发现,有些难题难以进行并行化处理[11]。这些难题被称为"P固有串行难题",并有自己独特的复杂性,称为"完全复杂性"。譬如,电路求值问题已被证明隶属此类[12]。这类计算的问题似乎有一个共同点:没有并行算法能够求解这些问题或并行算法的并行性很低。因此,对整个P完全类来说,即使使用更多处理器,也不会有效地提高计算速度。

量子计算固有并行性的另一个障碍来自于量子计算输出相关的概率问题。量子计算主要由3个步骤组成。首先,要对量子系统进行初始设置。然后,使量子系统处于叠加态,并

1) 处理器速率取决于晶体管密度和时钟速度。即使晶体管密度有所增加,时钟速度越高处理器的温度还是会越高,并需冷却。因此当前时钟速度最高约为8.5~9 GHz,并需用液态氮进行冷却,预防处理器损坏。

对叠加态的概率幅进行操作。因为量子幅度允许负数值，所以在量子计算过程中，系统的一些幅度值有可能会消失，从而减少了计算熵(computational entropy)。这是相消性量子干涉现象(destructive quantum interference)。反过来，相长性量子干涉(constructive quantum interference)可以提高概率幅的值，增加相关结果的概率。第三个步骤最为关键，即要在所得到的结果中选取正确的答案。一个量子计算的结果，是在测量过程中根据它的概率，即幅度的平方值，而得出的。与经典并行算法不同，量子计算不保证所得出的结果是正确的。在某些情况下，如当每两个潜在输出之间的概率差值相同时，得出正确结果的概率较低。为了确认一个结果是正确的，一个量子计算过程必须重复进行足够的次数。这就是为什么量子计算所提供的并行性并不是对每个问题都能带来明显计算加速的原因。

1 量子计算的物理背景

任何计算装置都是一个物理系统。虽然现代计算机硬件所使用的半导体技术依赖量子物理现象(如运用量子穿隧来做开关)，但是经典计算机的计算过程却遵循经典物理的定律。至于量子计算机，其计算过程显然要遵循量子物理的定律。量子力学理论所依赖的4条基础规律(量子力学公设)适用于封闭的量子系统，而量子计算机被视为一个封闭的量子系统，因此服从量子力学的规律。

1905年，Einstein在Max Planck有关光谱中颜色分布的研究成果之上发表了一篇论文[13]，设想光不是一种连续性现象，而是能量的若干离散(定)量，即现今所谓的“光量子”(light quantum)或“光子”(photon)。这种能量子可以以整体被吸收或释放。几年后，电子在一个原子里转换能级的离散过程便是用这个概念来解释的。在这种模式下，Einstein的能量子包含不同能级之间的能量差异。将该能量差异除以Planck常数 h 时，该能量差即确定了该量子所携带的光色。

Planck和Einstein的光量子概念认为光子具有波本性和粒子本性。1924年，de Broglie在Planck和Einstein有关光的研究基础之上提出了电子波理论，并假设所有物质都具有波本性，其结果为波粒二象性理论，即所谓的德布罗意假说(de Broglie hypothesis)：任一运动粒子皆有相应的正弦波。1926年，Thomson通过薄金属衍射实验观察到所预测的干涉模式，因此确认了de Broglie假说中的电子波本性。基于此结果，Schrödinger发展出了一个量子物理模型(称为波动力学)，使用波函数来描述物理系统的状态。

取决于粒子的离散性，量子计算机在理论上可使用任何的粒子作量子硬件，譬如电子、光子、原子、原子核等，而量子信息单位则为电子能级或自旋、光子偏振、原子自旋或原子荷值、原子核自旋等。这样的一个量子系统的量子状态可以用一个波函数来代表。

1.1 量子位及其物理表示

经典计算机中，基本信息单元为位元或二进制位(bit)。量子计算机中量子信息的基本单元是量子位元(quantum bit, qubit)。双态系统的两个状态通常分别以“0”和“1”表示，按此记法，一个二进制位可处于“0”态或“1”态，而一个量子位元同样可处于“0”态或“1”态，但除此之外，一个量子位元还可以处于这两态(或多于两态)的任何一个线性组合之中。该组合称为这两态的叠加或态叠加：

$$\sum_{i\in\{0,1\}} \alpha_i |i\rangle \tag{1}$$

任何的两态量子系统都有可能用作量子位元的物理实现。宇宙中最普遍的粒子——具有单一电子的氢原子，是最早被研究用作量子位元系统之一的粒子。电子的空间表示称为“轨域”(orbital)，而在不同的多电子原子中，不同的轨域一般具有不同的能级。轨域数量在理论上是无穷的。能量最低的轨域是球形1s轨域。以氢原子为例，1s为氢电子的基态(ground state)。向电子输入能量，如让电子吸收可见光或红外光(即波长约在400 nm～1 mm之间)的光子之后，便能提高该电子的能级。就能级而言，从1s轨域跃迁至球形2s轨域所需能量最少，因此氢电子的空间表示多半为球形2s轨域。2s为氢电子的第一激发态(first excited state)。一个处于2s轨域的电子可通过释放电磁辐射而跃迁回1s轨域，即回到其原来具有的能级。

按照量子物理普遍使用的及由Dirac引进的bra-ket向量标记法，一般性量子位元 ψ 标记为 $|\psi\rangle$。$|\psi\rangle$是一个列向量，称为“ket”。$\langle\psi| = \overline{|\psi\rangle}^{\mathrm{T}}$ 是一个行向量，是 $|\psi\rangle$的共轭转置(conjugate transpose)，称为“bra”。一般性量子位元通常如式(2)所示：

$$|\psi\rangle = \alpha_0|0\rangle + \alpha_1|1\rangle = \begin{bmatrix}\alpha_0\\ \alpha_1\end{bmatrix} \tag{2}$$

其中，$|0\rangle$和$|1\rangle$分别为“0”态和“1”态的Dirac标记，$\alpha_0\in\mathbb{C}$和$\alpha_1\in\mathbb{C}$为复数概率幅，$|\alpha_0|^2$和$|\alpha_1|^2$分别为量子位元$|\psi\rangle$(测量后)处于$|0\rangle$态或$|1\rangle$态的概率。由于叠加态具有概率性质，量子系统的概率幅必须符合归一化条件，即$|\alpha_0|^2+|\alpha_1|^2=1$，以消除概率幅中的复数和负数。用以表达量子位元的简单方法是在单位圆上画二维向量，单位圆位于欧几里得空间中以(0,0)点为中心的笛卡儿坐标系上。使用在单位圆上定义的正弦和余弦三角函数，便可以 $\alpha_0=\cos\theta$ 和 $\alpha_1=\sin\theta$ 表示量子位元的概率幅值。概率幅值的表述 $\psi^{\theta}=\cos\theta+i\sin\theta$ 建基于欧拉公式所建立的三角函数和复数指数函数之间的关系(见图1)。

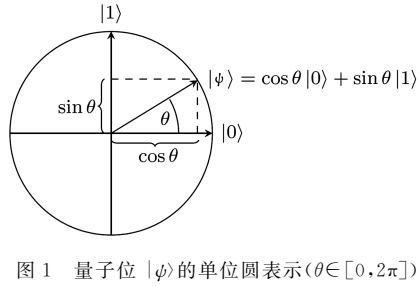

图1 量子位 $|\psi\rangle$的单位圆表示($\theta\in[0,2\pi]$)

Fig. 1 Representation of a qubit $|\psi\rangle$ on the unit circle($\theta\in[0,2\pi]$)

1.2 Hilbert空间的特性

在Neumann[14]对量子物理做出的数学表述中，量子系统所能采取的状态位于Hilbert空间。当今量子物理和量子计算亦大都使用Hilbert空间作为量子态的空间。

一个Hilbert空间 H 为复向量内积空间。两个向量$|\psi\rangle=\alpha_0|0\rangle+\alpha_1|1\rangle$和$|\phi\rangle=\beta_0|0\rangle+\beta_1|1\rangle$的内积为：

$$\langle\psi|\phi\rangle=\alpha_0^*\beta_0+\alpha_1^*\beta_1 \tag{3}$$

其中,α^* 为 α 的复共轭数。由内积概念诱导(induced)的向量范数(norm,统称向量长度)为:

$$||\psi\rangle|=\sqrt{\langle\psi|\psi\rangle} \tag{4}$$

Hilbert 空间 H 亦为完备度量空间。在一个完备度量空间中,每个 Cauchy 系列都收敛于 H 中的一个极限值。在这样的一个空间内,每个元素皆有明确的定义,因此可以使用微积分。譬如,用(偏)微分来求解 Schrödinger 方程,或用积分来归一化波函数。量子物理中自然出现的 Hilbert 空间是无限的,譬如,任一粒子的位置具有连续性,其值可为任何实数。在大多数情况下,量子计算中的 Hilbert 空间为有限维空间,但有些无限维量子系统也可以截短为有限维系统,譬如 Rangan 等[15]便通过禁止若干系统转换的手段来限制空间维度。这种系统能用有限维方法来分析可控性。1983 年,Huang 等[16]提出了"使用逐段常元控制方式无法在有限数量操作的情况下实现全局可控性"的观点;这个观点近期受到了挑战,譬如 2017 年有学者提出了量子机器学习无限维状态空间的理论[17],虽然这方面的研究才刚起步。

一个量子系统可用函数来描述。因为所测量的结果属概率性质,所以该函数必须遵循概率定律,故该函数必须为平方可积函数,即其积分必须是有限的。

$$\int||\psi(x)\rangle|^2\mathrm{d}x<\infty \tag{5}$$

式(5)亦称 Born 规则或 Born 概率诠释规则[18],用于规定函数 $|\psi\rangle$ 的概率密度。

Hilbert 空间是欧几里得空间的一个扩充。二维欧几里得空间中的单位圆可以扩充至二维 Hilbert 空间中的单位球体(如果单考虑归一化向量,即长度为 1 的向量,这些向量尖端都位于单位球体的球面上),即 Bloch 球体。Bloch 球体上可表示任何具有两态的单量子位元的量子系统。异于单位圆,Bloch 球体上的每对相对向量,如 $|0\rangle$ 和 $|1\rangle$,为相互正交向量(见图 2)。作为一个完备向量空间,Hilbert 复数向量空间能够保证每个向量的幅值以及角度都可以被测量。

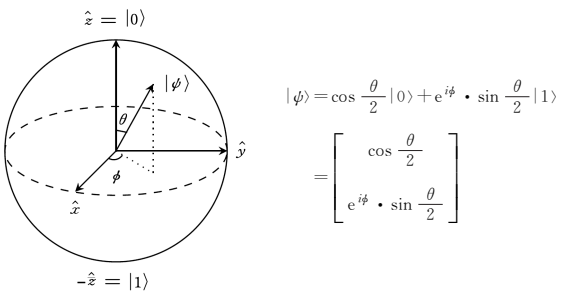

图 2 Bloch 球体上的量子位表示

Fig. 2 Representation of a qubit on Bloch sphere

每个 Hilbert 空间都具有一个规范正交基(orthonormal basis)。该空间内的每一个向量都可用基向量的线性组合构成。两个向量 $|\psi\rangle$ 和 $|\phi\rangle$ 为为正交向量,当且仅当其内积为 $\langle\psi|\phi\rangle=0$。用于表示 n 量子位元的 Hilbert 空间的维度是 2^n。2^n 维 Hilbert 空间 H 的规范正交基是一个以 2^n 个向量组成的且符合以下条件的集合 $B=\{|v_i\rangle\}\in H$:

$$\forall v_i,v_j\in B:\langle v_i|v_j\rangle=\delta_{ij} \tag{6}$$

其中,σ_{ij} 为 Kronecker σ 函数:

$$\delta_{ij}=\begin{cases}1, & i=j\\ 0, & i\neq j\end{cases} \tag{7}$$

在二维 Hilbert 空间中,通常用规范正交基向量{$|0\rangle$,$|1\rangle$}来表示一个量子位元,这个基础称为"计算基"(computational basis)。量子位元也可以用其他基础表示,譬如,将经典比特 $|0\rangle$ 或 $|1\rangle$ 置于叠加态的一种方法是利用 Hadamard 矩阵 $\frac{1}{\sqrt{2}}\begin{bmatrix}1 & 1\\ 1 & 1\end{bmatrix}$,将这个矩阵用于 $|\psi\rangle=|0\rangle$ 比特上可得出这个比特在{$|+\rangle$,$|-\rangle$} 基础中的表示 $|\psi'\rangle=\frac{|+\rangle+|-\rangle}{\sqrt{2}}$(见图 3)。不同的基础是可以变换的。如果再次将 Hadamard 矩阵应用于 $|\psi'|$ 之上,则得出 $|\psi''\rangle=|\psi\rangle=|0\rangle$。

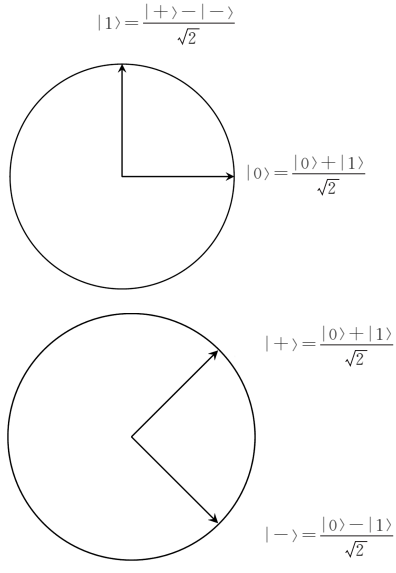

图 3 量子位单位正交基的例子及转基公式

Fig. 3 Examples of orthonormal bases of qubit and respective basis transformation formulas

在量子物理中,任何两个系统状态 $|\psi\rangle$ 和 $|\phi\rangle$ 都能够以概率 1 予以区别,当且仅当:两者相互正交,即其内积为 $\langle\psi|\phi\rangle=0$。对于两个非相互正交系统状态 $|\psi\rangle$ 和 $|\phi\rangle$,当其内积为 $\langle\psi|\phi\rangle\in(0,1)$ 时,不能在每种情况下都予以区别。状态区别,指单次测量一个系统的状态而得出的结果是独特的。假设两个量子态为 $|\psi_1\rangle=\frac{|0\rangle+|1\rangle}{\sqrt{2}}=|+\rangle$ 和 $|\psi_2\rangle=|1\rangle$,这两个状态是非相互正交的(内积为 $\langle\psi_1|\psi_2\rangle=\frac{\langle0|+\langle1|}{\sqrt{2}}|1\rangle=\frac{1}{\sqrt{2}}$),不能以概率 1 区别。只有测量结果为与状态 $|1\rangle$ 相关的本征值时,才能够明确指出量子态为 $|\psi_2\rangle$;测量结果为与状态 $|0\rangle$ 相关的本征值时,无法明确指出量子态为 $|\psi_1\rangle$ 还是 $|\psi_2\rangle$。我们在量子计算中追求的结果为完全区别状态,在计算中得到非完全区别状态的情况下,必须多次重复计算过程来明确区别所得出的结果。因此,在探测一个量子系统的任何物理量(physical quantity)(譬如测量一个电子的能级)时,必须确保该量所能采取的值是可以区别的。

在系统状态辨认性的基础上,量子物理提出了"可观察量"(observable)的概念。一个可观察量为探测器所能观察到

的量子系统的物理量状态。量子物理假设每个可观察量都存有相对的自伴算子(self-adjoint operator)。方阵 $\boldsymbol{A}$ 为自伴算子,当且仅当 $\boldsymbol{A}^{\dagger}=\boldsymbol{A}$,其中 $\boldsymbol{A}^{\dagger}$ 为 $\boldsymbol{A}$ 矩阵的复数共轭矩阵。有 k 状态的观察量相当于一个 $k\times k$ 的算子矩阵。自伴算子矩阵元素为自己的共轭,因此自伴算子的所有本征值皆为实数,从而符合一个实际存在的物理量不能包含复数的要求。量子物理工作者习惯将自伴算子与厄密算子替换使用,但这两种算子在数学中却有区别:自伴算子亦为厄密算子,但是厄密算子未必是自伴算子。只有在假设向量空间为有限维空间的情况下,这两种算子的定义才是一致的。本文将符合上述条件的算子一概称为自伴算子。

从自伴算子的条件 $\boldsymbol{A}^{\dagger}=\boldsymbol{A}$ 可得出正规算子(normal operator)的定义:$\boldsymbol{A}^{\dagger}\boldsymbol{A}=\boldsymbol{A}\boldsymbol{A}^{\dagger}$。因此,自伴算子属正规算子。每个正规算子都符合以下的谱定理。

定理 1(谱定理,spectral theorem) 每个应用于有限维 Hilbert 空间 H 的正规算子 S 都在该空间存在一个正交基,而该基的所有向量皆为 S 的本征向量。

这个定理确保了每一个自伴算子都存在至少一个正交基,因此适合用作可观察量的数学表征(与可观察物理量相关的数学算子大多是线性算子)。

算子的本征值对应可观察量所能采用的值。如果以氢原子电子的能量为可观察量,其$|0\rangle$和$|1\rangle$态分别为电子的基态和第一激发态;而与能量相连的算子叫作 Hamiltonian,标记为$\hat{H}$,用于指定量子系统的综合能量,即势能和动能的总和。Hamiltonian 矩阵必须在数理上导出或由实验结果确定。使用 Schrödinger 等式:

$$\hat{H}|\psi_n\rangle=E_n|\psi_n\rangle \tag{8}$$

即可算出与两个电子能级相关的本征值为:

$$E_n=-\frac{R\cdot Z^2}{n^2} \tag{9}$$

其中,n 为主量子数(principal quantum number),用于指定轨域的壳层(意即从原子核算起,第一个壳层为 $n=1$,第二个壳层为 $n=2$,如此类推);$R=13.6\text{ eV}$ 为 Rydberg 常数;Z 为原子系数(氢原子系数为 1)。从式(9)可以得出:氢电子作为 1s 轨域时(量子物理中粒子的位置是概率性的,而一个轨域则由测量电子位置确定,因此可以说电子位置界定了原子内的每个轨域),其能量为 $E_1=-13.6\text{eV}$;作为 2s 轨域时,其能量为 $E_2=-3.4\text{eV}$。因此,电子从基态转至激发态所需的能量为 10.2eV。

至今,单使用 Schrödinger 等式只能为单个电子系统(二体问题)提供精确结果,却不能精确描述有两个或以上电子的系统。总体而言,原子系数越高,该等式所需的修正就越多。

1.3 量子力学的公设

量子计算装置以及量子计算都是基于量子力学原理的。其中最重要的 4 个公设由 Neumann 于 1955 年提出[14]。这些公设建立物理系统与(用于为该系统建立模型的)量子力学的数学联系,即用于定义量子力学的数学结构。

1)状态空间公设

任意的封闭物理系统都存在一个复数向量内积空间,称为该系统的状态空间。状态空间描述系统所能行使的状态,该系统完全由系统的状态向量描述。量子计算最关键的系统为量子位,其状态空间为二维 Hilbert 空间,其状态向量$|\psi\rangle$为单位向量(取决于归一化条件)。此公设建立了封闭物理系统与 Hilbert 空间之间的关系。

2)系统演变公设

一个封闭物理系统随时间经由幺正变换演化,即该系统从其处于时间 t_1 的状态$|\psi(t_1)\rangle$转换为其处于时间 t_2 的状态$|\psi(t_2)\rangle$,而该转换操作由幺正算子 $U(t_1,t_2)$执行:

$$|\psi(t_2)\rangle=U(t_1,t_2)|\psi(t_1)\rangle \tag{10}$$

式(10)为 Heisenberg 矩阵力学中的表示。由于 U 为幺正算子,意即符合 $U^{\dagger}U=U^{\dagger}U=\boldsymbol{I}$($\boldsymbol{I}$ 为单位矩阵)条件,因此它所诱发的运算为可逆转运算:$U^{\dagger}U|\psi\rangle=|\psi\rangle$。

从概念上讲,系统演变使用 Heisenberg 的矩阵力学比较容易解释;若要计算这个演变,则一般会使用 Schrödinger 的波动力学。在 Schrödinger 的波动力学中,量子系统$|\psi\rangle$的时间演变可用 Schrödinger 方程来计算:

$$i\hbar\frac{\partial}{\partial t}|\psi(t)\rangle=\hat{H}|\psi(t)\rangle \tag{11}$$

式(11)为 Schrödinger 波动力学中的表示。其中,$\hbar=h/2\pi$ 是约化普朗克常数,$\hat{H}$ 是代表量子系统能量的 Hamiltonian 算子。根据 Planck 原来的方程 $E=h\nu$,常数 h 是光子能量 E 与其相连的频率 ν 的比率,即$\frac{E}{\nu}$。

幺正演变的一个好处是它保存了系统状态空间的内积。两个向量$|\psi\rangle$和$|\phi\rangle$之间的角 θ 可按式(12)进行计算:

$$\cos\theta=\frac{\langle\psi|\phi\rangle}{||\psi\rangle||\phi\rangle||}=\langle\psi|\phi\rangle \tag{12}$$

其中,两个向量都符合归一化要求,其范数皆为 1,因此公式里的分母亦为 1。因此,在量子计算中,两个向量之间的角为两个向量的内积。在几何学上,每个幺正算子(幺正演变)相当于 Hilbert 向量空间的若干旋转,从而成为一种不改变状态向量范数的变换。因此,用于操作单量子位的幺正算子相当于二维 Hilbert 空间(Bloch 球)的若干旋转,没有改变量子位的单位范数和两个量子位之间的角度。

此公设建立了物理系统演变与幺正变换之间的关系。由于量子系统在测量过程中必须与测量仪器发生互动,因此该量子系统不再是一个封闭的系统,演变亦不再是一个正交操作。显然,如果要在量子物理框架内描述测量过程,则必须对量子测量另做独立公设。

3)量子测量公设

一个封闭系统的测量由作用于其状态空间的若干算子集合 $\{M_m\}_{m=1}^{n}$ 描述,其中的 m 指数用来指称个别测量结果。假设紧接在测量之前的系统状态是$|\psi\rangle$,则测量结果为 m 的概率为:

$$P(m)=\langle\psi|\boldsymbol{M}_m^{\dagger}\boldsymbol{M}_m|\psi\rangle \tag{13}$$

其中,$\boldsymbol{M}_m^{\dagger}$ 为 $\boldsymbol{M}_m$ 矩阵的伴随矩阵。因为所有测量结果的总概率必须总和为 1,所以式(13)可表述为:

$$\sum_m\boldsymbol{M}_m^{\dagger}\boldsymbol{M}_m=\mathrm{I} \tag{14}$$

其中,I 为单位算子(identity operator)。

在假设测量得出 m 结果的情况下,测量后的系统量子态为:

$$|\psi'\rangle=\frac{\boldsymbol{M}_m|\psi\rangle}{\sqrt{\langle\psi|\boldsymbol{M}_m^{\dagger}\boldsymbol{M}_m|\psi\rangle}} \tag{15}$$

式(15)通常称为波函数的"坍缩"。测量后,系统不再处于叠加态,而减缩至量子态 m。假设量子位为 $|\psi\rangle=\alpha_0|0\rangle+\alpha_1|1\rangle$,测量算子则为 $M_0=|0\rangle\langle0|$ 和 $M_1=|1\rangle\langle1|$。测量导致测量后的量子位 ψ 按 $|\alpha_0|^2$ 的概率处于 $|0\rangle$ 态,而按 $|\alpha_1|^2$ 的概率处于 $|1\rangle$ 态。测量后,量子态 $|\psi\rangle$ 不再处于向量 $|0\rangle$ 和 $|1\rangle$ 的叠加态。量子测量一般用于读取计算结果,不可逆,即测量后的量子系统不可恢复测量前的叠加状态。

假设有向量空间 V。在数学中,投影 P 为符合 $P^2=P$(幂等性)条件的线性变换 $P:V\to V$。量子物理定义符合 $M_m^2=M_m$ 要求的测量算子称为"投影算子"。在测量算子同时为自伴算子($\boldsymbol{M}_m^{\dagger}=\boldsymbol{M}_m$)的情况下,该测量称为"投影测量",测量算子则称为"正交投影算子"。因此,式(13)中的算子成为 $\boldsymbol{M}_m^{\dagger}\boldsymbol{M}_m=\boldsymbol{M}_m$。用正交投影算子 $\boldsymbol{M}_m$ 来进行测量时,所得出的每个结果的概率皆为:

$$P(m)=\langle\psi|\boldsymbol{M}_m|\psi\rangle \tag{16}$$

而测量后的系统量子态为:

$$|\psi'\rangle=\frac{\boldsymbol{M}_m|\psi\rangle}{\sqrt{\langle\psi|\boldsymbol{M}_m|\psi\rangle}} \tag{17}$$

测量过程中将量子态 $|\psi\rangle$ 投影于 m 子状态空间,此为投影测量名称的来源。投影算子描述若干可观察量,所测量得出的状态都是可以明确辨识的。此外,取决于幂等关系,投影测量是重复性的测量。将算子连续应用于一个量子态上时,不会改变第一次测量所得出的结果,也不会改变该量子态:第一次测量后重复进行测量得出的结果概率皆为 1。

在量子力学中,很多重要测量都属于非投影测量,如用银幕来测量一个光子的位置时会导致该光子被毁灭。一般来说,测量后的量子系统的状态是不确定的。这类测量中,最普通的一种叫做正算子值测量(positive operator valued measurement,POVM),使用的算子是正数自伴算子(positive self-adjoint operator)。此种算子可以进行弱测量,即在不导致波函数塌缩的情况下获得不完整的信息。投影测量与 POVM 之间的一个重要区别在于,前者的正交投影算子一概为相互交换,而 POVM 的算子有可能为非交换算子。

物理上,测量是对量子系统的一种外部观察。用于测量一个量子系统的装置本身亦为一个量子系统,故此测量系统与被测量的系统一有接触即发生互动。测量过程中,所测量的系统或子系统的幺正演化会受到干扰。测量过程导致的对系统波形的干扰可能是一种非幺正现象。当前物理学界尚未探测到量子波函数受到测量的影响而产生坍缩的过程。这个称为"测量问题"的现象尚未得到物理理论上的解释。

此公设建立了物理系统测量与测量算子之间的关系。

4)复合系统公设

一个复合物理系统的状态空间为该系统的子系统状态空间的张量乘积。与复合量子系统关联的 Hilbert 空间 H 为其子空间 H_i 的张量乘积:

$$H=\bigotimes_i H_i \tag{18}$$

设 $\{v_i\}$ 为子空间 H_i 的基,$\{v_j\}$ 为子空间 H_j 的基,$\{v_l\otimes v_j\}$ 则为 $H_i\otimes H_j$ 的基。量子位 $|\psi_1\rangle=\alpha_0|0\rangle+\alpha_1|1\rangle$ 和 $|\psi_2\rangle=\beta_0|0\rangle+\beta_1|1\rangle$ 的复合系统则为:

$$|\psi_1\rangle\otimes|\psi_2\rangle=\begin{bmatrix}\alpha_0\\ \alpha_1\end{bmatrix}\otimes\begin{bmatrix}\beta_0\\ \beta_1\end{bmatrix}=\begin{bmatrix}\alpha_0\begin{bmatrix}\beta_0\\ \beta_1\end{bmatrix}\\ \alpha_1\begin{bmatrix}\beta_0\\ \beta_1\end{bmatrix}\end{bmatrix}=\begin{bmatrix}\alpha_0\beta_0\\ \alpha_0\beta_1\\ \alpha_1\beta_0\\ \alpha_1\beta_1\end{bmatrix} \tag{19}$$

将张量乘积算子应用于状态空间时,其维度为子空间维度的乘积:$\dim(H)=\prod_i\dim(H_i)$。

此公设建立了物理复合系统与张量乘积之间的关系。

2 Werner Heisenberg 的不确定性原理与量子计算的速度

代数的环论(ring theory)定义了一个"换位子"概念。因为复数集是一个环,所以定义在复数环上的矩阵集合亦为一个环。在这个基础上,量子物理引入了量子算子可交换的概念。$\boldsymbol{A}$ 和 $\boldsymbol{B}$ 两个复数矩阵的换位子定义为:

$$[\boldsymbol{A},\boldsymbol{B}]=\boldsymbol{AB}-\boldsymbol{BA} \tag{20}$$

从定义可见,换位子为 0,当且仅当:$\boldsymbol{A}$ 矩阵和 $\boldsymbol{B}$ 矩阵可交换。在量子力学中,换位子概念与 Heisenberg 不确定性原理有关。不确定性原理声称,同时测量粒子的两个不可交换的物理性质时,测量的精确度会存在一定的限制。Heisenberg 曾经指出,同时测量粒子的位置与动量不能获得两者的任意精确度[19]。在位置空间内,沿 x 方向的动量算子通常标记为 $P=-i\hbar\frac{\partial}{\partial x}$,沿 x 方向的位置算子标记为 $X=x$。两个算子的换位子推导如下:

$$\begin{aligned}[P,X]\psi&=(PX-XP)\psi\\&=-i\hbar\frac{\partial}{\partial x}(x\psi)+ix\hbar\frac{\partial\psi}{\partial x}\\&=-i\hbar\psi-i\hbar x\frac{\partial\psi}{\partial x}+i\hbar x\frac{\partial\psi}{\partial x}\\&=-i\hbar\psi\end{aligned} \tag{21}$$

其中,ψ 为上述量子系统的态,即物理中的波函数。移除波函数后得出的结果为 $[P,X]=-i\hbar\neq0$。不确定性原理的最普遍综合形式为 Robertson 不确定性关系[20]。Robertson 不确定性关系可表述为:

$$\Delta A\cdot\Delta B\geqslant\frac{1}{2}|\langle[A,B]\rangle| \tag{22}$$

因此,位置与动量间的不确定性关系为 $\Delta P\cdot\Delta X\geqslant\hbar/2$。在量子物理中,算子 Y 的不确定性标记为 $\Delta Y=\sqrt{\langle Y^2\rangle-\langle Y\rangle^2}$,同时 Δ 符号指值域,$\langle Y\rangle=\langle\psi|Y|\psi\rangle$ 为期待值。Δ 符号的定义与统计学中的标准偏差(以 σ 符号标志)的不同之处在于:量子物理标准偏差适用于每个粒子,而不是以统计学中的重复抽样方式得出[21]。

2.1 时间与能量的不确定性原理

对计算机科学起重要作用的另一种不确定性关系是时间 t 与能量 E(E 与上述 Hamiltonian 算子无关,因为该算子不依赖于时间;但 $E=i\hbar\partial/\partial t$ 出现在含时间的 Schrödinger 方程(11)中)的不确定性。有关时间与能量之间的不确定性原理虽由 Heisenberg 提出,但其给出的关系仅为数量级(order of magnitude)$\Delta E\cdot\Delta t\sim\hbar$,没有明确指出数值。Heisenberg[22]在

做出时间与能量不确定性原理的推测时，假设时间为一个算子，并且以 Stern-Gerlach 实验中对能量测量精确度 ΔE 的考虑为依据。在该实验中，Stern-Gerlach 装置中的原子在一定时段 Δt 内受到了偏转场(如不均匀磁场)的影响，因此在对能量进行测量时，Δt 明显不是时间测量的精确度，而是进行测量的持续时间。

当前物理学界却普遍认为时间与能量不确定性原理和位置与动量不确定性原理可能属于不同等级的现象。量子物理中，空间坐标和时间坐标都是连续性的，而不是量化的。因此，时间算子应具有连续性本征值，其范围则是从 $-\infty$ 到 $+\infty$。由此，Pauli[23]试图论证既然能量既可以具有连续的本征值，也可以具有离散的本征值，那么时间算子不能存在，并且只能把时间当作一个“经典数字”(在 Dirac 命名中，经典数字指实数或复数)[1]。但是，Pauli 的这个论证允许在一定的情况下(即当两者的本征值都是连续性的)，时间仍可用作算子。除此之外，Garrison 等[24]已在 1970 年证明(具有量化的能量本征值的)谐振子允许使用自伴算子作时间算子，并且这个算子符合 $[E,t]=-i\hbar$ 的典范交换关系(canonical commutation relation)。量子典范交换关系依照经典物理中的泊松括号(Poisson bracket)而成立。两个具有 $2n$ 个独立变量的函数 $f(p_1,\cdots,p_n,q_1,\cdots,q_n)$ 和 $g(p_1,\cdots,p_n,q_1,\cdots,q_n)$ 的泊松括号定义为 $\{f,g\}=\sum_{i=1}^{n}\left(\frac{\partial f}{\partial q_i}\frac{\partial g}{\partial p_i}-\frac{\partial f}{\partial p_i}\frac{\partial g}{\partial q_i}\right)$[25]，该括号定义了典范变换集。如果两个函数皆为可观察量，那么可以按照以下规则将其转换为交换子：$i\hbar\{A,B\}=[A,B]$。在这个关系中，H 和 t 两者皆称为“典范共轭”(canonical conjugates)。因此，Garrison 等证明了 Pauli 的论证不能成立。

在同一文献[23]中，Pauli 给出了另一个论证，从另一角度介入这个问题。他的依据是，所有已知的物理系统的能量都有下界，亦即能量算子的本征值必须远小于 $-\infty$。实际上，大多数物理系统的能量都是正值的(例外的有我们提到的氢原子电子的(负值)能级等)。在这个基础上，Pauli 论证了：对于能量谱具有下界的系统来说，不存在足以满足上述传统交换关系并因此可以用作该系统的时间算子的自伴算子。

必须注意的是，上述对能量谱的限制并不完全排斥了时间算子这个概念，而是排斥的时间算子作为一个自伴算子的可能性。因此，物理学界还是尝试将时间当作一种算子来配合不确定性原理。

Mandelstam 等[26]认为 Δt 是量子系统内的一种固有时间。这样的诠释可以把 Δt 表达为系统状态的衰变时间等，而 ΔE 为该状态的能量测量的不确定性。在这种情况下，Δt 不是时间测量的不确定性，而是衰变时间的统计分布。另一方面，Aharonov 等[27]却尝试论证，在有限时间内可以对能量进行任意精确度的测量，但必须明确区别被测量系统与时间测量装置的作用。位置与动量为同一量子系统的参数，而时间与能量的关系扩展至了两个不同的系统；能量是量子系统的参数，而时间由一个外部经典计时器(另外一个系统)测量，因此其是一种经典变量。

Briggs[28]将测量装置用作被测量量子系统的量子环境，然后使用经典极限来推导出时间与能量的不确定性关系：$\Delta E\cdot\Delta t\geqslant\frac{\hbar}{2}$。其中，$\Delta t$ 为使用经典测量装置产生的时间测量的不确定性(即时间测量的精确度而非时间的长度)，ΔE 则为被测量系统的能量。经典极限的概念来自物理学界认为量子物理可以从经典物理导出的观点。与牛顿力学相比，相对论力学中的基本方程多了一个新的基本物理常数，即光速 c。在其他方面，两个理论的数学结构是相似的。因此，将极限 $\frac{1}{c}\to 0$ 应用于相对论力学基本方程上即可将之转为牛顿力学基本方程[29]。正如相对论使用了经典物理中没有的光速常数，量子物理引入了经典物理所没有的 Planck 常数 $\hbar$。物理学界希望，将若干极限(譬如 $\hbar\to 0$)应用于量子力学方程便可以推导出经典物理方程。一个物理理论的直接推导或逼近牛顿力学的能力叫做经典极限。至今，这个方案还没有得到明确的求解。有学者甚至认为使用 $\hbar\to 0$ 极限无法从微观世界得到宏观世界[29-30]。

理论上，时间未必可当作量子系统(函数)的参数。在非相对论力学中，即符合伽利略相对性原理的量子力学，时间是绝对的。因为一个可观察量是一个以系统量子态为变量的函数，而量子态是一个以时间为变量的函数，即时间是一个自变量而不是算子。在相对论力学中，即符合狭义相对论的量子力学，时间与空间具有平等地位。如果位置是算子，那么时间必须也是算子；如果时间是参数，那么位置必须也是参数；因此两者皆可同时成为算子或同时成为参数。但是，相对论力学中的时间概念可有不同的意义，从而使得时间算子的定义变得极为复杂。另一方面，就计算效率而言，把位置与时间都用作参数则较为有利。虽然算子与参数的两种方式不同，但在相对论力学中使用任何方式进行计算所得的结果均一致[31]。

时间与能量不确定性原理对量子计算的重要性来自其强加于量子系统从某一量子态转换为另一个量子态的限制。作为一个量子物理系统的量子计算机，用量子态的转换来进行计算操作。因此，时间与能量不确定性原理对量子计算的速度将产生影响。

在 19 世纪末、20 世纪初，由科学哲学界发展出来的“观察者”概念进入了刚刚萌芽的量子物理[32]。量子物理对不确定性原理的最初理解来自 Heisenberg 自己。根据他于 1930 年提出的意见[33]：测量不准的原因基于若干观察者效应，即在测量量子系统的位置时，所使用的测量设备会改变量子系统的动量；相反，在测量量子系统的动量时，测量设备会改变系统的位置。当今，物理学界普遍认为不确定性原理为所有具有物质波性质的系统的本质属性，与测量设备的技术水平无关。这个本质属性可由数学解释，即一个可观察量所能有的本征值未必为另一个可观察量的本征值。我们做两个假设：1)有 A,B 两个可观察量，并且两者不可交换；2)测量后的量子系统处于可观察量 A 中的一个本征向量。因为两个可

[1] Wir schließen also, dass auf die Einführung eines Operators t grundsötzlich verzichtet und die Zeit t in der Wellenmechanik notwendig als gewöhnliche Zahl (“c-Zahl”) betrachtet werden muss.

观察量不可交换,上述的这个本征向量并非可观察量 B 的本征向量之一,而是可观察量 A 的本征向量的线性组合。由于量子物理的叠加现象是概率性的,因此未必能够准确测量可观察量 B 的值。

除了给予时间概念另外的一个解释(见上文)之外,Mandelstam 等[26]还(最早)提出了量子系统从一个状态转换到另一个状态所需的最短时间 τ:

$$\tau \geqslant \frac{\pi}{2}\tau \cdot \frac{\hbar}{\Delta H} \tag{23}$$

随后,Uffink[34]发现,在某些情况下,使用 ΔH 可导致对量子系统演化速度的估计变得很不合理,因为即使系统的平均能量是有穷的,能量算符的方差 ΔH 仍然有可能是无穷的。例如,在概率密度函数很狭窄的情况下,Hamiltonian 的方差 ΔH 便很高,从而导致式(23)中的下限值可以是任意小,于是在某种程度上移除了约束量子计算的任何限制。因此,在估计两态的转换时间时,不应依赖标准偏差或方差。基于 Uffink 所提出的异议,Margolus 等[35]提出了使用平均能量 $\langle H\rangle$ 来替代不等式(23)中的 ΔH。他们推导出的不等式(24)可用于任何长度的幺正状态的演变:

$$\tau \geqslant \frac{N-1}{N} \cdot \frac{h}{2\langle H\rangle} \tag{24}$$

其中,$\langle H\rangle$ 为一个量子系统的平均能量,N 为幺正状态的数量。按照式(24),从一个幺正状态演化至另一个幺正状态所需的最短时间 τ 为:

$$\tau \geqslant \frac{1}{2} \cdot \frac{h}{2\langle H\rangle} = \frac{\pi}{2} \cdot \frac{\hbar}{\langle H\rangle} \tag{25}$$

假设量子系统的固定平均能量 $\langle H\rangle$ 为一个单位,即 1J。按照式(25)中的不等式,幺正状态演化所需的时间至少是 1.6×10^{-34} s。由式(25)可知,演化所需的最短时间随着演化中状态数量的增加而变短。

同样,假设平均能量为 1J,可以按照式(26)中的不等式估算一个计算过程中每一秒所需的操作数量为 3×10^{33}[35]。由式(26)中的不等式可知,平均能量越高,计算速度越快。再者,如果单独考虑两个幺正状态,则该速率则为 6×10^{33},增加了一倍。

$$\nu \leqslant \frac{2\langle H\rangle}{h} = \frac{\langle H\rangle}{\pi\hbar} \tag{26}$$

当前,通常假定上述两个限制(即式(23)和式(24))为一个统一限制[36]:

$$\tau \geqslant \max\left\{\frac{\pi\hbar}{2\Delta H}, \frac{\pi\hbar}{2\langle H\rangle}\right\} \tag{27}$$

Levitin 等[37]证明了式(27)中的下界是紧的,即这个界是一个量子系统演化的最快速率。同时必须注意,上述两个下限都涉及幺正演化,即具有识别不同量子状态的能力的演化。实际的量子计算过程中有可能出现幺正转换,但也有可能出现非幺正转换。另外,必须考虑的因素是未对各个不同的操作的复杂度进行区别,所有操作一律算作一个操作。此为复杂度理论为了简化算法复杂度的分析而引进的一种限制。

值得强调的是,上面提倡的速率专门为量子计算而设,但却与 Lloyd[38]推导出的通用速率相似。Lloyd 估计一个经典或量子计算机在 1s 时间中能够进行的逻辑操作的上界为:

$$\sum_l \frac{1}{\Delta t_l} \leqslant \sum_l \frac{2E_l}{\pi\hbar} = \frac{2E}{\pi\hbar} \tag{28}$$

其中,l 指逻辑门,E_l 指逻辑门所需的能量,Δt_l 指一个逻辑门在 1s 中能够进行的操作数量。

3 超越 Heisenberg 不确定性原理

Heisenberg 不确定性原理于 1927 年提出。由于这个原理对了解量子过程的程度施加了限制,因此一直有学者试图证明或反驳。近年有研究成果显示,不确定性原理与波粒二象性为同一属性的两个表现[39-42];因此,如果波粒二象性是一个本体现象,不确定性原理则同样是一个事实。基于对双缝实验[43]的解释,当今学术界普遍认为波粒二象性亦为量子系统的本质属性,但是在发展量子物理的最早阶段,有学者却对这个假设持有疑问。Broglie 提出"相波"概念时,把相波视为与物质点运动相关联的波。Schrödinger 使用了 Broglie 的相波来构建他的波方程。Schrödinger 其实比 Broglie 多走了一步,把物质点假设为一个波系统,而非与物质点运动相关联的波系统[22]。当时,他给出的原因是,不使用相波概念的原子理论和分子理论都已经遇到了严重的障碍,因此暂时不应考虑不包含相波概念的框架。由此可见,物质波概念进入量子物理的条件不是因为 Schrödinger 认为这个概念是正确的,而是因为不使用这个概念的理论遭遇到了不少的困难。因此,在考虑不确定性原理时,必须注意它是以波粒二象性为真的假设。

3.1 Ozawa 对不确定性原理的修改(测量的不确定性)

上述不确定性原理涉及量子系统的内在不确定性(参考不等式(22))。不等式(22)中的 Δ 指的是标准偏差,与测量概念无关。但是,Heisenberg 最初是根据测量精度和测量必然产生的干扰两者之间的关系来阐述他的不确定性原理。按照这个理解,我们可重新界定式(22)中的原理[44]:

$$\varepsilon(A)\eta(B) \geqslant \frac{1}{2}|\langle[A,B]\rangle| \tag{29}$$

其中,$\varepsilon(A)$ 是测量 A 的误差,$\eta(B)$ 是测量 A 对 B 所造成的干扰。就标准偏差 $\Delta(A)$ 与测量误差 $\varepsilon(A)$ 的关系而言,Margenau[45]和 Popper[46]已指出,只有在误差远小于标准偏差($\varepsilon(A)\ll\Delta(A)$)时,才能够确定标准偏差的值。基于量子物理的概率性本质,Heisenberg 在提出位置与动量的不确定性原理时,借用了一个思想实验来描述使用 γ 射线显微镜观察电子所造成的不确定性[19]:

"Sei q_1 die Genauigkeit, mit der der Wert q bekannt ist (q_1 ist etwa der mittlere Fehler von q), also hier die Wellenlänge des Lichtes, p_1 die Genauigkeit, mit der der Wert p bestimmbar ist, also hier die unstetige Änderung von p beim Comptoneffekt …"

为符合本文记法,设 $\varepsilon(Q)$ 为 q_1,$\eta(P)$ 为 p_1。上述原理可做如下翻译:"假设 $\varepsilon(Q)$ 是 Q 的精确度(譬如 $\varepsilon(Q)$ 是 Q 的平均误差),即光的波长,$\eta(P)$ 是确定 P 值的精度,亦即 Compton 效应中的对 P 的不连续性变化…"。

式(29)被称为"测量与干扰的关系",并已在 1970 年被证明为无效[44,47]。证明很简单,使用了光子极化状态。极化的

各个成分可分别用 Pauli 矩阵表示。我们用 $\sigma_X=\begin{pmatrix}0&1\\1&0\end{pmatrix}$ 来表示极化的 X 成分，用 $\sigma_Y=\begin{pmatrix}0&-i\\i&0\end{pmatrix}$ 来表示极化的 Y 成分，用 $\sigma_Z=\begin{pmatrix}1&0\\0&-1\end{pmatrix}$ 来表示极化的 Z 成分。我们测量其中的一个成分，譬如 Z，然后观察该测量过程对另一个成分，譬如 X，所造成的影响，从而显示测量与干扰关系被违反。首先，我们算出 Heisenberg 的两个边界中的换位子：

$$\begin{aligned}\frac{1}{2}|\langle[Z,X]\rangle|&=\frac{1}{2}\left|\left\langle\begin{pmatrix}0&2\\-2&0\end{pmatrix}\right\rangle\right|=\left|\left\langle\begin{pmatrix}0&1\\-1&0\end{pmatrix}\right\rangle\right|\\&=\left|\left\langle i\begin{pmatrix}0&-i\\i&0\end{pmatrix}\right\rangle\right|=|\langle i\cdot Y\rangle|\\&=|\langle Y\rangle|\end{aligned}\tag{30}$$

如果所研究的量子系统的状态是 $|\psi\rangle=(|H\rangle+i|V\rangle)/\sqrt{2}$，则下界为最大：

$$\begin{aligned}|\langle Y\rangle|&=|\langle\psi|Y|\psi\rangle|\\&=\frac{1}{2}(\langle H|-i\langle V|)(i|H\rangle\langle V|-i|V\rangle\langle H|)(|H\rangle+\\&\quad i|V\rangle)=1\end{aligned}\tag{31}$$

量子系统处于这个状态时，两个不确定性 $\Delta X=\Delta Z=1$ 也最大，并且符合 Robertson 的不确定性关系(式(22))。但是如果要同时测量 Z 成分，这个测量可以是随意精确的，亦即能够 100%确定这个成分。如果假设测量 Z 的精确度为 1，其测量误差则为 0，因此无法符合式(29)中的要求，亦即 $|\langle Y\rangle|$ 不能为 1。此为对式(29)中下界的违反。

因为上述的下界不属于普遍下界，所以有学者尝试挑战 Heisenberg 的测量与干扰关系(不确定性原理)。到目前为止，对原来的不确定性原理进行修改的研究中，最有影响力的观点来自于小泽正直(Masanao Ozawa)[48-49]。这个修改方案已经得到了实验证明[50]。Ozawa 将平均误差更精确地界定为均方根误差 $\varepsilon(A)=\langle(A_--A)^2\rangle^{0.5}$，$\eta(B)=\langle(B_+-B)^2\rangle^{0.5}$。其中，$A$ 和 B 是需要测量的观察量，而 A_- 和 B_+ 是实际测量的观察量(值)。Ozawa 在 Heisenberg 不等式的基础上提出的测量与干扰关系如下：

$$\varepsilon(A)\eta(B)+\varepsilon(A)\Delta(B)+\eta(B)\Delta(A)\geqslant\frac{1}{2}|\langle[A,B]\rangle|\tag{32}$$

Rozema 等[44]和 Baek 等[50]分别进行了几场实验，并且确定了在不同的实验设置中，当式(29)中的 Heisenberg 原理论关系失效时，式(32)中的 Ozawa 理论关系保持有效。

就测量与干扰关系而言，Heisenberg 的式(29)和 Ozawa 的式(32)之间有什么区别呢？假设两个观察量不交换，即换位子大于 0。在一定情况下，不等式(29)的左面可以小于所预测的 $1/2|\langle[A,B]\rangle|$，它甚至可以为 0(例如当 $\varepsilon(A)=0$ 或 $\eta(B)=0$)[44]。在这些情况下，Heisenberg 的测量与干扰关系是无效的。与此不同，式(32)中的 $\varepsilon(A)$ 和 $\eta(B)$ 任意一个可以为 0，但是两者不能同时为 0。这意味着，Ozawa 没有推翻测量与干扰关系的真实性。但是，Ozawa 的边界小于 Heisenberg 原来的不确定性，因此也降低了量子计算的不确定性。

3.2 Hirschman 的不确定性(熵不确定性，entropic uncertainty)

除了 Heisenberg 不确定性之外，量子物理还涉及其他不确定性现象，其中比较重要的一个是 Hirschman 提出的不确定性[51]。因为它的定义基于 Shannon 熵的总和，所以当前称其为熵不确定性。这种不确定性在加密学上的应用尤为常见。

1957 年，Hirschman 使用函数 $f(x)$ 及其 Fourier 变换函数 $g(x)$ 来提出熵不确定性关系[51]：

$$H(|f(x)|^2)+H(|g(x)|^2)\geqslant 0\tag{33}$$

其中，$H(|h(x)|^2)=-\int_{-\infty}^{\infty}|h(x)|^2\cdot\log|h(x)|^2\mathrm{d}x$ 是概率频率函数 $h(x)$ 的熵，并且必须是收敛的。式(33)中的下界由 Maassen 和 Uffink 进一步证明为以下关系[52-53]：

$$H(A)+H(B)\geqslant\log\frac{1}{c}\tag{34}$$

其中，$H(X)=-\sum_i p_i\cdot\ln(p_i)$，$P=(p_1,\cdots,p_n)$ 是概率分布，而 $p_i=\langle x_i|\psi|x_i\rangle$；$c$ 是可观察量 A 和 B 的本征向量之间的最大重叠，定义为 $c=\frac{1+c_1}{2}$，并且 $c_1=\max_{ij}|\langle a_i|b_j\rangle|^2$，而 $|a_i\rangle$ 和 $|b_j\rangle$ 分别为可观察量 A 和 B 的本征向量。由等式可知，两个可观察量的共同本征向量越多，熵的不确定性越小。这个不等式应用于使用规范正交基的测量，因此对量子计算有着重要影响。式(34)中的结果以 Deutsch[54]于 1983 年证明的不确定性(式(35))为依据。

$$H(A)+H(B)\geqslant-2\ln(c)\tag{35}$$

这里，两个可观察量的本征值必须是离散的。与 Robertson 综合不确定性关系的不同之处在于上述两个下界是绝对的，不依赖量子系统的状态。

Deutsch 和 Maassen 的不等式依赖于与个别可观察量所能采取的本征向量相关的变量 c；但在完全不考虑这个因素的情况下界定熵不确定性也是可能的，Białynicki-Birula 等[55]在 1975 年即按此方向为位置与动量导出了以下关系：

$$H(Q)+H(P)\geqslant\log(\mathrm{e}\pi\hbar)\tag{36}$$

对于这个关系，Coles 等[52]随后证明了它蕴涵使用标准偏差的 Heisenberg 位置与动量不确定性原理：

$$H(Q)+H(P)\geqslant\log(\mathrm{e}\pi)\Rightarrow\Delta(Q)\Delta(P)\geqslant 1/2\tag{37}$$

这里，Heisenberg 的原理可以表示为熵的总和的下界。显然，熵不确定性比 Heisenberg 原理强，因此给量子计算的速度带来的影响也相应较大。

物理学界曾经认为，波粒二象性与 Heisenberg 不确定性相同。但是据 Coles 等[39]所提供的证明，比 Heisenberg 不确定性更强的熵不确定性才等同于波粒二象性。如文献[55]所述，不确定性的适用程度与所使用的物理对象有关。譬如，使用光子进行计算时，如果所考虑的是光子相和光子数，则熵不确定性将特别明显。本文已经提及，用光子进行计算和测量是目前最成功的计算模型之一；但是因为文献[39]的研究成果也是受熵不确定性影响较大的模型，所以量子计算在选择量子系统来运行测量时，除了考虑输入输出和操作的难度之外，还必须考虑该系统的熵不确定性程度。目前，由于熵不确定性高于 Heisenberg 不确定性，因此对熵不确定性的研

究亦已超出依赖标准差的原不确定性原理的研究。

结束语 量子计算属于计算机科学领域的较新方向,于1980年被提出,一直到20世纪90年代中期才出现第一个量子算法。

量子计算的挑战来自多个方面。一方面,量子系统状态内设的概率性,致使量子算法有可能具有概率算法的本质。这类算法必须多次执行,以期在某程度上确定正确的答案,因而导致时间上的耽误。必须强调的是,理论上,如果要确定一个答案为正确的,那么按照大数定律,需要无穷次重复运算程序。另一方面,量子计算机用来实现相对于经典计算机硬件的物理载体的性质是多种多样的,不是每个物理媒体都在同样程度上适用于量子计算。最近有学者认为使用光学方法与仪器(诸如用于光分束器和移相器的线性光学元件)进行计算的前景是可行的。光子通常由光电探测器计数,其操作模式是吸收光子并产生电流脉冲,因此在测量过程中,即便被测量的系统没有被破坏,也至少会产生物理性改变。这就是Heisenberg最早提出的不确定性原理的依据。他的原理指出,用于测量量子系统的仪器会对该系统引入干扰,从而导致所测量出的值存在一定的不确定性。目前,这个原理多半被称为测量与干扰关系。干扰指的是一个可观察量的测量过程对另外一个可观察量所造成的干涉。

Heisenberg的这个原理随后经过有关工作者的研究而获得修改。其中最受关注之一的修改来自Ozawa。Ozawa不确定性原理的下界小于Heisenberg不确定性原理的下界。这个结果显示出,因为量子计算过程中的干扰减小了,所以计算速度便相对有所提高。

但当今学术界在谈论Heisenberg不确定性原理时,大多指的是两个不交换的可观察量内含的不确定性,与上述情况无关。这个"新"原理来自于量子物理的数学形式化。任何两个可观察量一旦不交换,便出现不确定性。就位置与动量两个可观察量而言,已经有学者证明这个不确定性可以视为熵不确定性的下界。

目前,上述两个原理之间的关系还没有被清楚界定,需要进一步的研究。在考虑到当前量子物理的诠释与形式化,以及当前技术的发展程度时,两个不确定性都必须予以参照,这也是Ozawa构建中很明显的一点。

致谢 本文得到了南京大学计算机科学与技术系的宋方敏教授和吴楠副教授的帮助,在此表示衷心的感谢!

参 考 文 献

[1] MANIN Y. Вычислимое и невычислимое [S]. Советское Радио, 1980.

[2] FEYNMAN R. Simulating physics with computers [J]. International Journal of Theoretical Physics, 1982, 21(6/7): 467-488.

[3] BENIOFF P A. Quantum mechanical Hamiltonian models of Turing machines [J]. Journal of Statistical Physics, 1982, 29(3): 515-546.

[4] DEUTSCH D. Quantum theory, the Church-Turing principle, and the universal quantum Turing machine [J]. Proceedings of the Royal Society of London, 1985, 404(1818): 97-117.

[5] SIMON D. On the power of quantum computation [C]// Foundations of Computer Science. IEEE, 1994: 116-123.

[6] DEUTSCH D, JOZSA R. Rapid solution of problems by quantum computation [J]. Proceedings of the Royal Society of London, 1992, 439(1907): 553-558.

[7] BERNSTEIN E, VAZIRANI U. Quantum complexity theory [C]// Proc. 25th Annual ACM Symposium on Theory of Computing. ACM, 1993: 11-20.

[8] SHOR P. Polynomial-time algorithms for prime factorization and discrete logarithms on a quantum computer [J]. SIAM Journal on Scientific and Statistical Computing, 1994, 26: 1484-1509.

[9] BERNSTEIN D, HENINGER N, LOU P, et al. Post-quantum RSA: Report 217/351[R]. Cryptology ePrint Archive, 2017.

[10] MOORE G. Cramming more components onto integrated circuits [J]. Electronics, 1965, 38(8): 114-117.

[11] GREENLAW R, HOOVER H, RUZZO W. Limits to parallel computation: P-completeness theory [M]. Oxford University Press, 1995.

[12] LADNER R. The circuit value problem is log space complete for P [J]. SIGACT News, 1975, 7(1): 18-20.

[13] EINSTEIN A. Über einen die Erzeugung und Verwandlung des Lichtes betreffenden heuristischen Gesichtspunkt [J]. Annalen der Physik, 1905, 17: 132-148.

[14] VON NEUMANN J. Mathematische Grundlagen der Quantenmechanik [M]. Springer, 1932.

[15] RANGAN C, BLOCH A M, MONROE C, et al. Control of trapped-ion quantum states with optical pulses [J]. Physical Review Letters, 2004, 92(11): 113004.

[16] HUANG G M, TARN T J, CLARK J W. On the controllability of quantum-mechanical systems [J]. Journal of Mathematical Physics, 1983, 24(11): 2608- 2618.

[17] LAU H K, POOSER R, SIOPSIS G, et al. Quantum Machine Learning over Infinite Dimensions [J]. arXiv: 1603. 06222. Physical Review Letters, 2017(118): 080501.

[18] BORN M. Zur Quantenmechanik der Stoβvorgänge [J]. Zeitschrift für Physik, 1926, 37: 863-867.

[19] HEISENBERG W. Über den anschaulichen Inhalt der quantentheoretischen Kinematik und Mechanik [J]. Zeitschrift für Physik, 1927, 43(3): 172-198.

[20] ROBERTSON H. The uncertainty principle [J]. Physical Review, 1929, 34: 163-164.

[21] BACH R, POPE D, LIOU S H, et al. Controlled double-slit electron diffraction [J]. New Journal of Physics, 2013, 15: 033018.

[22] SCHRÖDINGER E. An undulatory theory of the mechanics of atoms and molecules [J]. Physical Review, 1926, 28(6): 1049-1070.

[23] PAULI W. Die allgemeinen Prinzipien der Wellenmechanik [M]// Handbuch der Physik (24). Springer, 1933.

[24] GARRISON J, WONG J. Canonically conjugate pairs, uncertainty relations, and phase operators [J]. Journal of Mathematical Physics, 1970, 11(8): 2242-2249.

[25] POISSON J. Poisson bracket [J]. Journal de l'École Polytechnique, 1809, 8: 266.

[26] MANDELSTAM L, TAMM I. The uncertainty relation between

energy and time in nonrelativistic quantum mechanics [J]. Journal of Physics (USSR), 1945, 9(4): 249-254.
[27] AHARONOV Y, BOHM D. Time in the Quantum Theory and the Uncertainty Relation for Time and Energy [J]. Physical Review, 1961, 122(5): 1649-1658.
[28] BRIGGS J S. A derivation of the time-energy uncertainty relation [J]. Journal of Physics Conference Series, 2008, 99: 012002.
[29] KLEIN U. What is the limit of quantum theory? [J]. American Journal of Physics, 2012, 80(11): 1009.
[30] SEN D, SENGUPTA S. A critique of the classical limit problem of quantum mechanics [J]. Foundations of Physics Letters, 2006, 19: 403-412.
[31] SREDNICKI M. Quantum Field Theory [M]. Cambridge University Press, 2007.
[32] BUNGE M. The turn of the tide [M]// Quantum theory and reality. Springer, 1967.
[33] HEISENBERG W. Die physikalischen Prinzipien der Quantentheorie [M]. S. Hirzel, 1930.
[34] UFFINK J. The rate of evolution of a quantum state[J]. American Journal of Physics, 1993, 61: 935.
[35] MARGOLUS N, LEVITIN L B. The maximum speed of dynamical evolution [J]. Physica D, 1998, 120: 188.
[36] DEFFNER S, CAMPBELL S. Quantum speed limits: from Heisenberg's uncertainty principle to optimal quantum control [J]. Journal of Physics A, 2017, 50(45): 453001.
[37] LEVITIN L B, TOFFOLI Y. Fundamental limit on the rate of quantum dynamics: The unified bound is tight [J]. Physical Review Letters, 2009, 103: 160502.
[38] LLOYD S. Ultimate physical limits to computation [J]. Nature, 2000, 406: 1047-1054.
[39] COLES R, KANIEWSKI J, WEHNER S. Equivalence of wave-particle duality to entropic uncertainty [J]. Nature Communications, 2014, 5: 5814ôarXiv: 1403. 4687.
[40] COLES R, BERTA M, TOMAMICHEL M, et al. Entropic uncertainty relations and their applications [J]. Reviews of Modern Physics, 2017, 89(1): 015002-1.
[41] DüRR S, REMPE G. Can wave-particle duality be based on the uncertainty relation? [J]. American Journal of Physics, 2000, 68: 1021-1024.
[42] BUSCH P, SHILLADAY C. Complementarity and uncertainty in Mach-Zehnder interferometry and beyond [J]. Physics Reports, 2006, 435: 1-31.
[43] YOUNG T. Bakerian Lecture: Experiments and calculations relative to physical optics [J]. Philosophical Transactions of the Royal Society, 1804, 94: 1-16.
[44] ROZEMA L A, DARABI A, MAHLER D H, et al. Violation of Heisenberg's measurement-disturbance relationship by weak measurements [J]. Physical Review Letters, 2012, 109: 100404.
[45] MARGENAU H. Measurements in quantum mechanics [J]. Annals of Physics, 1963, 23(3): 469-485.
[46] POPPER K. Quantum mechanics without "the observer" [M]// Quantum theory and reality. Springer, 1967.
[47] BALLENTINE L E. The statistical interpretation of quantum mechanics [J]. Review of Modern Physics, 1970, 42(4): 358-381.
[48] OZAWA M. Uncertainty relations for noise and disturbance in generalized quantum measurements [J]. Annals of Physics, 2003, 311(2): 350-416.
[49] OZAWA M. Uncertainty relations for joint measurements of noncommuting observables [J]. Physics Letters A, 2004, 320: 367-374.
[50] BAEK S-Y, KANEDA F, OZAWA M, et al. Experimental violation and reformulation of the Heisenberg's error-disturbance uncertainty relation [R]. Scientific Reports 3, 2013.
[51] HIRSCHMAN I I. A note on entropy [J]. American Journal of Mathematics, 1957, 79(1): 152-156.
[52] COLES P J, BERTA M, TOMAMICHEL M, et al. Entropic uncertainty relations and their applications [J]. Review of Modern Physics, 2017, 89(1): 015002.
[53] MAASSEN H, UFFINK J B M. Generalized entropic uncertainty relations [J]. Physical review Letters, 1988, 60(12): 1103-1106.
[54] DEUTSCH D. Uncertainty in quantum measurements [J]. Physical Review Letters, 1982, 50(9): 631-633.
[55] BIAŁYNICKI-BIRULA I, MYCIELSKI J. Uncertainty relations for information entropy in wave mechanics [J]. Communications in Mathematical Physics, 1975, 44(2): 129-132.

Renata WONG, doctoral student. Her main research interests include quantum computing, protein structure prediction, physics and linguistics.