物理学咬文嚼字之八十
特别二的物理学

曹则贤[†]
(中国科学院物理研究所 北京 100190)

2016－09－23收到
† email:zxcao@iphy.ac.cn
DOI:10.7693/wl20161009

道生一，一生二，二生三，三生万物。
——老子《道德经》
Omnibus ex nihil ducendis sufficit unum.
——Gottfried Leibniz

摘 要 “二”贯穿物理学，two，dual，squared，binary，second-order，quadratic 等修饰的各种概念，构成了物理天地之大部分。

1 引子.二哲学与不二哲学

道家哲学谓三生万物，西洋哲人莱布尼兹(Gottfried Leibniz，1646—1716)谓 omnibus ex nihil ducendis sufficit unum(自无导出万物，一足矣)，皆属谬误之论，前者失于过繁，后者失于过简。以笔者所观数理哲诸般书籍而论，二即一切，or two is involved in everything of interest，才是正论。其实，老子和莱布尼兹应该也是明了“二才是一切”的道理的。老子在“三生万物”之后紧接着的一句是“万物负阴而抱阳，冲气以为和”，而莱布尼兹说出 sufficit unum 是在他自易经悟出二进制数字(binary number)对着 0 (nihil)和 1 (unus)的有感而发，而不管是阴阳还是 binary number，关键词都是“二”。

二的影子散见于物理、数学各处，触手可及。以英文而论，就有以 two，binary，double，second-order，rank-2，dual，以及字面是4而实际在说 2 的 squared，quadratic，quadic 等词修饰的各种概念，此外还有 interaction，coupling，correlation这种暗含2的概念。

有趣的是，人们似乎对2都有嫌多的情绪。东方有所谓的不二哲学[1)]，西方有强调整体的一元论(monism)。一旦遭遇两种面目或者选择，人们就会变得含糊、游移不定，中文谓之“有点二乎”，西文的动词怀疑 (英语的 doubt，德语的 zweifeln)，形容词令人起疑的 (dubious)，其字面上都是二。与二相关的各种概念，构成了物理天地之大部分。

2 带平方的定律

拉丁语的4是quattuor，这在法语的 quatre (第四，quatrième)，意大利语的 quattro (第四，quartto)，西班牙语的cuatro (第四，cuarto)还都能看出来。方块有四角，所以谈论方形会和4联系起来是自然而然的事情。一个边长为 x 的正四边形，面积为 x^2，德语的说法是 x quadrat，英文就说是 x squared。Square，其前缀为 ex- (不同语言中有写成 s-，es-，或者é-的)，意思为 out；词干是动词 quadrare，与数词 quattor 同源。x^2就是x的二次幂，中文读成x的平方，而平方按照字面理解就是平的方块。

数的平方在数学和物理上足以带来充分多和充分复杂的内容。在数学上，quadrature 因为历史的原因就是求面积的意思。一个为人熟知的 squaring 问题就是 squaring the circle (化圆为方)，也称为 quadrature of the circle，是古时候一段时间里非常挑战智力的几何问题。有兴趣的读者可以翻翻数学史关于这个问题的论述。以$y=kx^2$形式出现的物理定律有很多，比如焦耳定律，电阻的功率和电流平方成正比；自由(或者无摩擦沿着斜面)下落，下落高度(路程)与时间平方成正比，下落高度与获得的速度平方成正比，等等。其实，当年焦耳、伽利略和 Willem's Gravesande 分别得到这些定律时的数据都是很粗糙的。

1) 奇怪的是，中国人偏偏敬重关二爷、秦二哥和武二郎。是巧合，还是因为普适性的恐惧造成的?

但是，真实的科学就是这样做的：当把两个数据的关联(correlation)画到一张纸上时，人们首先看是否成线性关系；如果不是，就会想到平方关系。这种数据之间的线性或者平方关系只要大致有个差不多就行了，才不需要在意有多大误差呢。由下落高度与获得的速度平方成正比，Émilie du Châtelet 女士引入了 vis viva (活力)，mv^2，这个量，其后贝努里(Johann Bernoulli)引入了因子1/2，这才有了 $\frac{1}{2}mv^2$ 这种形式的动能概念。

整数之于平方似乎情有独钟，方程 $x^2+y^2=z^2$ 有整数解，但换成 $x^3+y^3=z^3$ 或者更高次幂的形式就没有整数解了，这就是所谓的费马大定理。在近代数学家几百页的证明和费马所说的书边角写不下的证明之间，我选择后者。即便费马当时吹牛了，我也相信存在费马大定理的简单证明。

人们初学物理时就会学到两个平方反比律(inverse square law)，一个是物体之质量造成的牛顿万有引力 $f=G\frac{m_1m_2}{r^2}$，一个是物体之电荷造成的库仑力 $f=\frac{1}{4\pi\varepsilon_0}\frac{q_1q_2}{r^2}$。初学时，觉得这两个力的表达可高深了。后来知道，这只不过是强调了我们是生活在3D空间这个事实而已。在3D空间中，一个半径为 r 的球其表面积为 $S=4\pi r^2$。任何一个物质流自一点各向同性地向外辐射，其流是守恒的，则其流密度总是和 $4\pi r^2$ 成反比(图1)。至于流的源对应的物理量，比如一颗桂花树上桂花的数目，同造成的场之强度，比如在远处闻到的花香浓度，可以引入一个比例因子(它的功能是平衡方程两边的量纲，它的数值比较起来就不那么重要)从而将它们写成一个方程的形式。把牛顿引力写成 $f_{1\to 2}=4\pi G\frac{m_1}{4\pi r^2}m_2$ 的形式，库仑力写成 $f_{1\to 2}=\frac{1}{\varepsilon_0}\frac{q_1}{4\pi r^2}q_2$ 的形式，场的概念就跃然纸上了，且由系数的选择能看出来后者要比前者晚。后者理解了那 4π 的来由，一个比例因子 $1/\varepsilon_0$ 简洁轻便，而万有引力常数不得不经常和 4π 结伴而行，比如爱因斯坦引力方程 $R_{\mu\nu}-\frac{1}{2}Rg_{\mu\nu}=-8\pi GT_{\mu\nu}$ 就有它的身影。

在一些文献中有库仑用挂在马槽上的铝球壳实验得到库仑力表达式的说法。想想 Charles-Augustin de Coulomb 生活的年代(1736—1806。为啥不译成“德·库仑”呢)，马槽上的铝球壳能得到什么精密的结果？更有甚者，验证牛顿万有引力是否是严格的平方反比形式 $f=G\frac{m_1m_2}{r^2}$ 在某些地方堂而皇之地成为了实验研究的课题。倘若依据所谓的精密的实验数据，得到的牛顿引力形式为 $f=G\frac{m_1m_2}{r^{1.9999999964\pm(18)}}$，库仑力形式为 $f=\frac{1}{4\pi\varepsilon_0}\frac{q_1q_2}{r^{2.00000072\pm(23)}}$，这样的物理学你愿意学？类似地，在动能定义 $\frac{1}{2}mv^2$ 中，1/2是分数，2是整数，不是实数。对于如今追求所谓的精确测量然后用计算机拟合曲线的赝物理学家来说，他们的世界里动能的形式说不定是这样的，$E_k=0.4999872(\pm 32)mv^{2.0000000013(\pm 27)}$。这种丑陋的公式如今随处可见，让人无语。

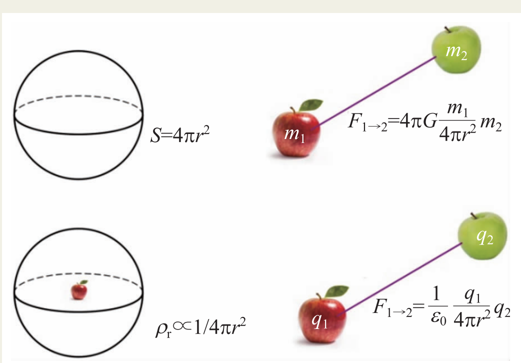

图1 平方反比律。图中公式的表达形式反映的是笔者自己的理解

掀开了量子时代的面纱的是一个和自然数平方有关的规律。氢原子在可见光范围内的光谱，从右向左可见间距渐小的四条分立的谱线(眼尖的读者可以看出第五条，等有了理论的帮助有些人能看见第六条。见图2)。这分立的四条谱线的波长分别为6562.10 Å (红色)，4860.74 Å(绿色)，4340.10 Å(蓝色)和4101.2 Å (紫色)。1885年，瑞士人巴尔莫发现这四个波长近似地是常数值 b=3645.6 Å的9/5，4/3，25/21和9/8倍。进一步地，可将这四个分数写成 $3^2/(3^2-2^2)$，$4^2/(4^2-2^2)$，$5^2/(5^2-2^2)$，和 $6^2/(6^2-2^2)$，这是公式 $n^2/(n^2-2^2)$ 的 n=3，4，5，6四项，由此可见光谱线是有规律的[1]。有了这个公式，可以计算 n=7对应的波长为 $b*7^2/(7^2-2^2)$ =3969.6 Å，回头去看光谱图，这第5条谱线赫然就在那个地方。现在换个角度看，会发现波数，就是波长的倒数，可以表示为 $\tilde{\nu}_B=R\left(\frac{1}{2^2}-\frac{1}{n^2}\right)$，$n$=3，4，5…，其中的 $R\approx 10973731\ \mathrm{m}^{-1}$ 被称为里德堡常数。这个公式后来被进一步扩展成更一般的形式 $\tilde{\nu}=R\left(\frac{1}{m^2}-\frac{1}{n^2}\right)$，其中 m，n ($>m$) 都是正整数。1913年，玻尔看出了巴尔莫公式或者说里兹公式中的奥秘，两项相减的光频率公式可以诠释为发光是电子从一个能级到一个较低能级的跃迁过程的伴随现象，

光的能量量子为原子中两电子能级之差。在经典角动量量子化的假设下，可以得到氢原子中电子的能级确实是正比于 $1/n^2$ 的。玻尔的原子理论完美地解释了氢原子谱线(的位置)。量子力学的时代真正开始了。

四次方或四次方倒数形式的公式也有。前者有Stefan—Boltzmann公式，谓黑体辐射积分能量之体积密度正比于温度的四次方，后者是说两平行板之间的Casmir 力与板间距的四次方成反比。同样是公式，前者是扎实的物理，后者是来自量子场论的比较率意的推导。后者被诠释为来自真空零点能的神奇效应。当有个实验者用扫描力学显微镜探针的振荡信号往这上凑，说有一段是满足这个关系从而可能是来自 Casmir 效应时 2)，这个不靠谱的结果反过来被一些对实验不熟悉的人当成了救命稻草，被说成了这个推导成立的实验证据。

3 二元数、两体势与相互作用

形容词 binary，来自 bis，放在名词前面表示该事物 made up of two parts or things，twofold，double (由两部分组成的，或者是双重的)。任何由两个单元组成的体系都是 binary system 3)，比如 binary star system，即由绕重心转动的两颗恒星组成的体系。两个通过万有引力相互作用的物体的运行轨迹是经典力学的标准问题。只考虑两个单元相互作用的碰撞是binary collision。

图2 现代光谱学获得的氢原子在可见光范围的光谱，可以清晰地看到五条谱线

只有两个数字0和1的数系是二进制数，英文为 binary numbers。在数学中一个拥有 two-component 的数是复数。复数定义为 $z=x+\mathrm{i}y$，其中 i 是单位虚数(imaginary number)，ii=－1，x 和 y 都是实数，分别是复数的实部和虚部。其实，没有理由认为这其中有什么古怪的虚数，将复数 z 看作是具有两个部分(part，segment，component)的数，记为 (x, y)，则只要满足算法 $(x_1, y_1)+(x_2, y_2)=(x_1+x_2, y_1+y_2)$，$(x_1, y_1)\times(x_2, y_2)=(x_1x_2-y_1y_2, x_1y_2+x_2y_1)$，那就再现了复数的算法。可见，所谓的复数，尤其是那个单位虚数 i，带入数学的是一种新的代数。这样的二元数，就是 binarions[2]。复数有两个单元，还暗含全新的代数，这样我们就容易理解了，为什么一个扩散型的方程 $\partial_t C=D\partial_x\partial_x C$，扩展成了 $\mathrm{i}\hbar\frac{\partial\psi}{\partial t}=\left[\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x^2}+V(x)\right]\psi$ 形式的复函数方程，现在它被称为(1D)薛定谔方程，就演绎出了那么多的幺蛾子。

物理上两粒子间的势能被选为如下形式的粒子间距离的函数，即 $V_{12}=V(\vec{r}_1, \vec{r}_2)=V(|\vec{r}_1-\vec{r}_2|)=V(r_{12})$，而一个多粒子体系的总势能是两体势能之和，即 $V=\sum_{i<j}V_{ij}$。这样的势能形式被称为binary potential，我们的物理学中采用的都是binary potential。对于三粒子体系，$V=V_{12}+V_{23}+V_{31}$；不知道存在三体势 $V_{123}=V(\vec{r}_1, \vec{r}_2, \vec{r}_3)$ 的世界，物理是什么样儿的。

Binary potential 反映的是 interaction 的思想。牛顿第三定律云：Actioni contrariam semper et æqualem esse reactionem: sive corporum duorum actiones in se mutuo semper esse æquales et in partes contrarias dirigi (反作用“总是相反”且等于作用，或者说两物体相互间的作用总是相等且作用到相反的参与方)。既然 action 和 reaction 是 mutual 的，那就是在谈论 inter-action。**相互作用不是粒子间影响对方运动的游戏，而是其自身存在的方式。**

认识到 interaction 在物理学中的地位是物理学上的一大革命。重力(gravity)问题，早期关切的是物体的 weight or lightness，此前被认为是物体自身的性质。不知是否是航海贸易发现货物重量莫名其妙的增减让人们认识到了物体的重量是物体和地球两者的事情。Interaction，在牛顿和库仑的力表示中，都落实为物质(粒子)某个指标的乘法，后来引入的关于基本粒子的强、弱相互作用也是换汤不换药。这似乎能解释群论在物理学中的地位——群论是只保留了乘法部分的代数。

4 二次型

前面已经说过，quadrature 与四

2) 把一条实验曲线分成几段解释是实验文章中常见的错误处理方式，虽然并不总是错误的。把一个实验结果往某个理论的东西上靠是实验工作的通病，反过来，非要到哪里找个实验证据是理论研究心虚的表现。一个不好的消息是，当一个理论和一个实验结果 fit very well时，两者都错的几率最大。

3) 一个常见的化学概念是binary compounds (二元化合物)。Binary compounds 的定义有点乱。如果只指由两种元素组成的化合物，$NaNO_3$就不是二元的。如果是指由两种构成单元构成的物质，酸根算一个单元，那$NaNO_3$就是二元的，考虑其晶体结构的几何单元更能认同这种观点。不过，若从晶体结构的几何单元来看，Fe_3O_4就不是二元的。

边形和面积有关。振子相位差为90°，即2π/4，从方位上讲是从指向正方形的一边转到邻边，则被说成是 in quadrature。形容词quadratic 和 quadrature，quadrate同源，也被理解为平方的意思。Univariate quadratic equation，$ax^2+bx+c=0$，就是一元二次方程。如果一个多项式的每一项都是二次幂的，即 homogeneous polynomials of order two (二阶齐次多项式)，比如 $x^2+axy+by^2+cyz+dxz+z^2$，这就是 quadratic form (二次型)。二次型对物理学具有特别的意义。

一个 n-变量的二次型有标准形式 $q(x_1, x_2\cdots x_n)=x^{\mathrm{T}}Ax$，其中 A 是实的对称矩阵。矩阵 A 总可以被对角化，也就是说二次型总可以化成只包含各个变量平方的形式，即 $q(x_1, x_2\cdots x_n)=\lambda_1x_1^2+\lambda_2x_2^2+\cdots+\lambda_nx_n^2$。如果再引入坐标尺度变换，则总可以使得系数 λ_i 是1，−1，或者0。此三个值在一个给定二次型中各自的个数，是这个二次型的标签(signature)。一个重要的定律是 Sylvester's law of inertia，它说的是一个二次型的标签是变换的不变量[3]，即将二次型 $q(x_1, x_2\cdots x_n)$ 变换成 $q(y_1, y_2\cdots y_n)=\gamma_1y_1^2+\gamma_2y_2^2+\cdots+\gamma_ny_n^2$ 形式，系数 λ_i 和 γ_i 中总有相同数目的1，−1，和0。

对于一个 n-变量的空间，点到原点距离的平方为 $\lambda_1x_1^2+\lambda_2x_2^2+\cdots+\lambda_nx_n^2$，其中 p 个系数是1，q 个是−1，这样空间就可以标记为 $R^{p,q}$。对于具体的物理问题，比如转动惯量矩阵，也即将绕某方向的转动惯量表示成方向余弦的二次型中的系数组成的矩阵，对角化后的本征值，都是正的。当然，这是因为我们生活的空间是欧几里得空间 $R^{3,0}$ 的原因。狭义相对论的时空是 $R^{3,1}$，点到原点距离的平方为 $-c^2t^2+x^2+y^2+z^2$ 4)。不同空间其几何性质的不同，研究一下距离平方为 x^2+y^2 和 x^2-y^2 两个2D世界就能找到一点感觉，比如试着写出这两个空间中表示矢量转动的矩阵。

只有两个变量的二次型是 binary quadratic form。谐振子的哈密顿量可写成 $H=q^2+p^2$ 的形式，是两变量的标准二次型。由麦克斯韦方程组得到的电磁场的哈密顿量也是这样的标准二次型，这是理解黑体辐射模型中引入谐振子模型的关键 5)。$H=q^2+p^2$ 形式的二次型和数学意义上的 x^2+y^2 不同，物理的坐标 q 和动量 p 在量子力学中存在对易(共轭)关系 $[q, p]=\mathrm{i}\hbar$。由谐振子的哈密顿量形式可引入产生算符和湮灭算符，这为谐振子的描述提供了一套别样的语言，这套语言被肆意滥用到各种物理问题。有人甚至断言说理论物理75%的天下不过就是谐振子模型，那么理解 $H=q^2+p^2$ 形式的二次型之重要性就不言而喻了，只是这看似简单的二次型所包含的数学还真不是一般人能掌握的。即便是整数域上的 binary quadratic form x^2+y^2，那也是数论的专门研究对象，是高斯都要花点功夫的学问。

空间是有度规的连续统。没有连续就没有几何，而度规(距离函数)则意味着二次型的计算。空间的概念中，其实不是距离，而是某个平方定义了距离，所以必涉及秩为2的度规张量，其实就是上文中二次型表示中的对称矩阵。狭义相对论中，距离微分的平方 $\mathrm{d}s^2=-c^2\mathrm{d}t^2+\mathrm{d}x^2+\mathrm{d}y^2+\mathrm{d}z^2$ 又可以写为 $\mathrm{d}s^2=g_{\mu\nu}\mathrm{d}x^\mu\mathrm{d}x^\nu$。这些是微分二次型，其变换不变性研究想必更复杂[3]。广义相对论涉及关于一般二次型度规空间的二次微分方程，不能够原谅自己不理解的人那就要多花点功夫了！二次型，微分二次型，以及它们的变换不变性，都有系统的数学工具，这些系统的工具我们都没掌握，想学到相关物理的精髓自然无从说起。

5 二阶微分方程

牛顿发明了微积分。牛顿对物理学的贡献之一，是确立了用 second-order 微分方程描述物理世界的基调。牛顿关于运动的第二定律说 Mutationem motus proportionalem esse vi motrici impressae, et fieri secundum lineam rectam qua vis illa imprimatur 6)(运动的改变正比于所受驱动力，且在驱动力所在的直线方向上)。把运动的改变表示为坐标对时间的二阶微分，而力被限制为是时空和速度的函数，则牛顿第二定律可形式地表示为 $\frac{\mathrm{d}^2x}{\mathrm{d}t^2}=f(t, x; \mathrm{d}x/\mathrm{d}t)$，这为此后的物理学定下了基调：物理学的动力学定律采二阶微分方程(second-order differential equation)的形式。为什么要采取二阶微分方程的形式呢？因为一阶微分方程太简单，而三阶微分方程太复杂。

麦克斯韦方程组关于电场和磁场的微分方程形式看似一阶微分方程，因为那里只出现了电、磁场关于时空的一阶微分(differentiated once)。引入了标量势 φ 和磁矢势 A，麦克斯韦也被写成二阶微分方程的形式——那里出现了场对时空

4) 我总觉得“i”是物理学内禀的一个对象，蕴含许多物理学的奥秘。把物理学中的“i”简单地当作 $x^2=-1$ 的根，正如在流行的相对论文献中那样，宁愿写成太数学的 $\mathrm{d}s^2=-c^2\mathrm{d}t^2+\mathrm{d}x^2+\mathrm{d}y^2+\mathrm{d}z^2$ 而不是写成有点物理的 $\mathrm{d}s^2=(\mathrm{d}(\mathrm{i}ct))^2+\mathrm{d}x^2+\mathrm{d}y^2+\mathrm{d}z^2$ 形式，可能流失了一些物理的内容。

5) 当年笔者读到相关问题时总是疑惑，空腔里哪来的谐振子？呃，原来是类比。

6) 一次次地引用拉丁文是希望读者能对着原文理解牛顿的伟大。那些不明就里的英文和中文翻译大大减损了这个效果。

的二阶微分 (differentiated twice)。这个二阶微分方程，注定了观察者位置与观察者速度的被超越，因此就有了相对论的两层意思——无去来处，动静等观。量子力学的动力学方程为薛定谔方程 $\mathrm{i}\hbar\partial_t\psi = H\psi$ ，这是一个关于复函数的二阶微分方程。因为 $H=\frac{p^2}{2m}+V(r)$ ，且根据量子化条件 $p_x=\mathrm{i}h\partial_x$ ，薛定谔方程是关于空间二阶关于时间一阶的微分方程(扩散方程)。动能项的本征值问题，即 $\mathrm{d}^2\psi/\mathrm{d}x^2+n^2\psi=0$ ，为我们提供了傅里叶分析出现的另一场景。有趣的是，我们所谓的数学物理方程课，讨论的大多都是这个本征值问题在不同维度、不同对称下的变种，提供了可与三角函数类比的其它可作为正交归一基的 L^2-范数的各种函数。笔者一直困惑的是，为什么复函数的方程，其解的径向函数似乎总是实函数，复数的性质只由角变量部分来表现，太诡异了。

物理学的动力学方程为我们描绘了一个二阶微分方程统治的世界[7)]。自然真的是由二阶微分方程描述的吗？关于这一点，笔者有些含糊。自然是否可以未经某些物理学家的同意就遵循更高阶的微分方程或者干脆别的形式的方程呢?

6 二象性与对偶

Duality，词根 dualis 是 2 的意思，汉译根据不同的语境译为对偶性和二象性，有点拔高的意味，其实就是“二乎”。比如，量子力学关于粒子的本性就有波—粒子二象性(wave-particle duality)的说法，一般教科书会诠释为既是粒子又是波，因此还有 wavicle (wave + particle)的说法。笔者以为这种说法未说到点子上——我们构造的波与粒子的形象是电磁波在频率标上的两个极端。波长过百米的无线电波，怎么着也难以想象它是粒子，而能量为MeV的 γ 光子，很难想象它会表现出什么波动性。居于两个极端之中间地带的，比如可见光和X-射线，就会同时表现出我们以为的那种波和粒子的特性。特征X射线就同时有将其当做波的 WDX (wavelength dispersive X-ray analysis)分析模式和将其当成粒子的EDX (energy dispersive X-ray analysis) 分析模式；而可见光部分的所谓量子力学实验中，有些实验者可能无意中在实验路途的不同地带上不停地变换着关于电磁波的认识。出现连续—分立二象性的地方是关于量子力学的测量问题，一方面我们要求因为测量诱导了一个量子事件，某个测量的结果以恰当的几率突然出现；另一方面，对于一个孤立的、未置于测量之下的体系，我们又要求体系按照薛定谔方程连续地、决定性地演化着。这种翻手为云、覆手为雨的手法对应一种 duality in dynamics(动力学二象性)。存在这些有点儿“二乎”的观念，反映出量子力学是一个四面透风的理论。

Duality 是数学中常遇到的概念，投影几何中把点与线的角色对调，欧拉多面体公式中顶点与面的对调，都是一种 duality。数学中的矢量空间，有一个对应的 dual space (对偶空间)。以简单的代数对偶空间为例，对于定义在域 F 上的空间 V，存在对偶空间 V^*，对偶空间 V^* 中泛函 φ 和空间 V 中的元素 x 定义一个非简并(退化)的双线性(bi-linear) 映射 $\varphi(x)=\langle\varphi,x\rangle: V^*\times V\to F$ 。对于 V 为有限维空间的情形，基为 $\{e_i\}$，对偶空间 V^* 有相同的维度，且可以构造相应的基 $\{e^j\}$，使得 $e^i\cdot e_j=\delta^i_j$ 。在固体物理学中，晶体中原子占据的空间是离散的空间，为 lattice结构；其对偶空间被名为倒空间(reciprocal space)或者倒格子(reciprocal lattice)。学习量子力学时会遇到的希尔伯特空间，即具有内积结构(由此可定义距离或者长度)的抽象矢量空间。希尔伯特空间是一个完备的内积空间，其基在量子力学中是自伴随算符的本征矢量；希尔伯特空间的dual space也是一个完备的内积空间。内积空间的Schwarz不等式是可以引入波函数几率诠释的数学基础。

7 二值性

二值性，two-valuedness，付诸于实验验证，是最能得到明确结论的情形。当一束银原子经过非均匀磁场后它会分成两束(图3)，分开多远、各自多宽以及边缘处根本没分开[8)]，这些都不重要。当将其中一束经过一个与前置的非均匀磁场垂直的非均匀磁场后，发现这一束又被分成了两束，物理学家凌乱了。

银原子束在磁场中暴露出的two-valuedness，大自然早就显示了，钠黄线双线(doublet)就是一例(图3)。钠的黄色双线揭示了电子波函数的基本二值性。电子波函数的二值性又不得不进行第二次倍分以便同相对论的原理相调和，扩展电子波函数的四值性被狄拉克漂亮地解释为一个新粒子——反电子，如今被称为正电子。一次又一次，当我们被意料之外的观测逼入死角的时候，一个凑巧的加倍(double)——一个理论结构到自身的折叠——又恢复了我们的理解，或者说由其而

7) 狄拉克方程形式上是一阶微分方程，但那里的主角，作为时空的函数，就不再是简单的实函数或者复函数啦——那样就太简单了。

8) 薄薄的一层银原子落到金属板上，肉眼是看不见的。烟鬼手头的劣质烟卷将它熏黄了。

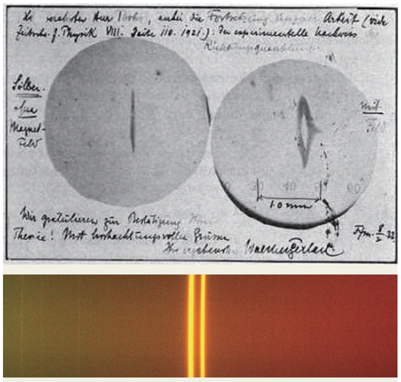

图3　银原子束在非均匀磁场中的劈裂与钠双黄线

来的幻觉。本征值为1和−1的算符有很多，包括手性(chirality)，螺旋性(helicity)，镜面反射与反演(mirror reflection and inversion)，等等，都可以归入二值问题。

Double 有个意思相近的动词 bifurcate。Bifurcation (分叉)，前缀 bi-来自bis，和binary 同源。在研究 logistic 方程 $x_{n+1}=rx_n(1-x_n)$ 时人们发现，迭代结果随着控制参数的增加，开始时稳定在一个值上，然后是稳定在2，4，8…个值上，就这样一直加倍下去。这样的分叉进行下去，由2^n达到的无穷大的后面是3，因此有周期−3意味着混沌的说法。

8　二元一次方程组

二元一次方程组 (two-variable linear equations，system of linear equations with two unknows)是初二学生都会解的，然而，就是这样简单的二元体系，其蕴涵的科学内容却是非常深刻的。笔者此前读到所谓热力学第一定律和第二定律是耦合的，百思不得其解。近期翻译热力学早期文献，并配合着线性方程的研究，方始有所觉悟。考察卡诺循环，热机在一个循环内从高温热源T_1处吸收热量Q_1，向低温热源T_2处放出热量Q_2，其间净做功W，根据热力学第一定律，有$Q_1-Q_2=W$。因为是在两个热源间工作，做功是目的，$Q_1-Q_2=W$可看作是两变量Q_1，Q_2的一个方程。这个方程不足以把一个卡诺循环确定下来；反过来说，关于卡诺循环必然还有另一个关系，这就是1850年代克劳修斯和开尔文爵士揭示的另一个关系，$Q_1/T_1-Q_2/T_2=0$。这个关系才是对第二定律的正确表达。从二元一次方程联立求解的角度来看，热力学第一、第二定律是耦合的说法，就好理解了。对一个物理定律的正确表达，应尽可能采取严格的数学表达，那种文字上的表达，比如热力学第二定律的开尔文表述和克劳修斯表述，只会带来认识上的混乱。热力学第二定律的出现比第一定律早，从第二定律数学表达的微分形式容易导出熵(entropy)的概念。基于熵的概念，热力学第二定律在公理化的热力学里有了更严格的表述：对于1-form $TdS+y_i dX_i$，其中$y_i dX_i$是力学量定义的、量纲为能量(功)的1-form，$TdS+y_i dX_i$总可以表达为某热力学函数的全微分，即有$dU=TdS+y_i dX_i$。

笔者注意到的从二元一次方程组的角度更容易理解的一项伟大科学成就是拉瓦锡的工作。通过称量反应物与生成物的质量，拉瓦锡确立了化学反应的质量守恒定律，比如对反应A+B→C，有关系式$m_A+m_B=m_C$，这和$Q_1-Q_2=W$如出一辙。同样的问题是，一个方程$m_A+m_B=m_C$不足以确定这个化学反应，一定还存在另一个关系！拉瓦锡经过仔细分析发现，对于一个化学反应反应物和生成物的质量之比总是小的整数比，比如对于反应$C+O_2\rightarrow CO_2$，$m_A:m_B:m_C\sim 3:8:11$。这是一个普适性的观察，如何理解。拉瓦锡推论，原子的化学性质是不同的，但是原子质量的构成是由下一层次的单元构成的，而下一层次的提供质量的单元是同一的。拉瓦锡用天平称量物质的质量研究化学反应，是他天才的思维能力让他由此洞察了原子核的秘密。那种以为物理学说到底是实验科学的说法，实在是浅薄之见。

9　结语

本篇介绍了平方(反比)律、二次型、二元数、二阶微分方程、两体势、对偶性、二值性、二元一次方程组等诸多与“二”有关的数学物理内容，虽然内容繁多，可能有助于读者学到一点物理学的实质内容，实则是只能提供一个挂一漏万且必定是蜻蜓点水般的介绍。有太多的关于“二”的内容，比如对物理学很重要的dyad，quadric概念，就没能介绍。显然，有更多的内容我根本就不知道。Rank-2张量在物理学中的作用我拟专文介绍，而order-2算子的本征值问题及其在量子力学中的作用笔者则刚开始做相关的研究。就深度而论，所发议论也是令人遗憾地浅薄，象(x, y)，$x+iy$，$\langle\psi|\phi\rangle$，$(\psi_\uparrow, \psi_\downarrow)$里涉及的two components的内在关联是怎样的，polarity，duality，coupling，conjugacy，complementarity，correlation这几个二元观念的异同，笔者都没有能力加以深入讨论。有兴趣的读者，不妨于闲暇时自己参详一二。

参考文献

[1] Steinberg S. Group theory and physics. Cambridge，1994. p. 399

[2] McCrimmon K. A Taste of Jordan Algebras. Springer，2004. p. 64

[3] Veblen O. Invariants of quadratic differential forms. Cambridge University Press，1908