

高二物理竞赛 • 电磁学与近代物理

学而思

目录

第一章 静电学	5
1.1 电荷与电场	5
1.1.1 电磁相互作用与电荷	5
1.1.2 库仑定律	6
1.1.3 电场	6
1.2 两个定律与电势	9
1.2.1 电场的高斯定律	10
1.2.2 电势与电场的环路定律	11
1.2.3 总结	12
1.3 静电能	13
1.3.1 静电势能	13
1.3.2 自能与相互作用能	13
1.3.3 电场能	16
1.4 电荷体系	17
1.4.1 电偶极子	17
1.4.2 电荷密度	20
1.4.3 极化强度 *	21
1.4.4 若干对称带电体系的电场	23
第二章 导体与介质	27
2.1 导体与静电平衡	27
2.1.1 绝缘体与导体	27
2.1.2 导体的特点	28
2.1.3 常见简单体系	32
2.2 电像法	34
2.2.1 半无限大空间的电像法	34
2.2.2 球面外与球面内的电像法	35
2.2.3 立体谐函数 *	37
2.3 电介质	41
2.3.1 微观角度理解极化	41
2.3.2 宏观角度理解极化	44
2.3.3 微观与宏观的联系	48
2.4 再议静电能	49

第三章 稳恒电流	53
3.1 稳恒电流描述与形成	53
3.1.1 德鲁特模型	53
3.1.2 费米气观点 *	54
3.1.3 能带论 *	56
3.1.4 惯性, 阻尼与回复力	57
3.1.5 稳恒电流与形成条件	58
3.2 电路与电路方程	60
3.3 电路分析基础	62
3.3.1 电路的整体结构与拓扑学结论	62
3.3.2 电路问题的求解套路	65
3.4 电路分析方法	66
3.5 半导体	67
第四章 静磁学	69
4.1 电流与磁场	69
4.1.1 磁场地位与电流分布	69
4.1.2 毕奥-萨伐尔定律	71
4.2 两个定律与矢势	73
4.2.1 磁场的环路定律	73
4.2.2 矢势与磁场的高斯定律	74
4.3 电流体系	77
4.3.1 磁偶极子	77
4.3.2 磁化强度	81
4.3.3 若干对称体系的磁场	81
4.4 磁介质与磁能	82
4.4.1 微观角度理解磁化	82
4.4.2 宏观角度理解磁化	83
4.4.3 磁场能量	83
第五章 磁生电	85
5.1 电磁感应	85
5.1.1 动生电动势	85
5.1.2 感生电动势	86
5.2 自感与互感	86
第六章 麦克斯韦方程组	87
6.1 麦克斯韦方程组	87
6.1.1 电生磁	87
6.1.2 麦克斯韦方程组	89
6.1.3 电荷在电磁场中的运动	93
6.2 平面电磁波	94
6.2.1 真空中的电磁波	94
6.2.2 介质中电磁波的传播	95
6.3 电磁场能量与动量	97

6.4 电磁波辐射	97
6.4.1 电磁辐射概论	97
6.4.2 偶极辐射	98
第七章 交流电路	101
7.1 相量表示	101
7.1.1 拟稳条件与交流元件	101
7.1.2 电阻, 电容, 电感特性	101
7.2 常见电路	102
7.2.1 谐振电路	102
7.2.2 滤波电路	103
7.2.3 电报线方程, 再论拟稳条件	103
7.3 变压器	103
7.3.1 磁路定律	103
7.3.2 理想变压器条件	103
7.4 电能传输	103
7.4.1 发电机与电动机: 三相绕组	103
7.4.2 变电站与电线损耗	103
7.4.3 市电规范	103
7.4.4 电源适配器	103
第八章 相对论力学	105
8.1 相对论运动学	105
8.2 相对论动力学	106
8.3 相对论连续物质	108
第九章 光的干涉	111
9.1 标量波理论	111
9.2 分波面干涉	114
9.3 分振幅干涉	115
9.4 偏振干涉	117
9.5 相干性	118
第十章 光的衍射	119
10.1 光栅与波带片	119
10.2 衍射积分公式	120
10.3 波前分析法	120
第十一章 物理光学	121
11.1 经典色散理论	121
11.1.1 复波矢与复折射率	121
11.1.2 经典电子论的解释	124
11.1.3 * 辐射阻尼与散射截面	127

第十二章 量子论	131
12.1 黑体辐射	131
12.2 光粒子性	132
12.3 玻尔原子	133
12.4 物质波与波函数	133

第一章 静电学

1.1 电荷与电场

1.1.1 电磁相互作用与电荷

物质若携带电荷(charge), 则可以参与电磁相互作用(electromagnetic interaction). 在经典情形下理解为电荷受到了一个力. 我们先研究静电学(electrostatics), 它要求受力物体与施力物体¹都处于静止状态, 或者认为一切电荷都处于静止状态.

电荷分正负, 阴阳激荡.

作为协变电磁场理论的基本要求与实验规律, 电荷是一个参考系不变的标量, 微观地看电荷具有量子化(quantized) 的奇妙特性². 质子的质量与电子的质量没有简单的整数比关系, 但电荷量确实严格相反. 如果设想两者电量有着微小差异, 那么构成自然界的原子也就携带了一个能够传递远程相互作用的库仑力, 从而导致自然界的不稳定³. 而质子(proton) 的电量被定义为基本电荷(elementary charge)⁴:

$$e = 1.602176634 \times 10^{-19} \text{ C}$$

实际上今天我们知道, 电荷量的量子化单位实际上更小, 至少是 $e/3$, 夸克的电量就是这个电量的一倍或两倍, 可正可负. 而电子(electron) 带一个单位的负电荷. 当然, 特殊场合还会有带一个单位正电荷的正电子(positron) 和一个单位负电荷的反质子(anti-proton) 甚至它们组成的反物质. 中子(neutron) 是不带电的.

电荷如同静质量, 是基本粒子的基本属性. 它只能随着基本粒子运动, 没有相互作用时不能被创造与消灭. 在存在电磁相互作用时电荷可以被创造与消灭, 典型的过程如电子与正电子湮灭(annihilation) 为两个光子, 或高能光子在原子核附近产生(production) 电子正电子对. 这些过程电荷的代数和守恒, 其实现要求反物质与相对论性运动的存在. 在经典情形下有严格的电荷守恒定律(charge conservation), 电荷可以流动形成电流, 但流动的过程中电荷的总量不能发生改变. 我们总是计算一定体积内的电荷代数和:

$$dQ = \sum_i n_i q_i dV$$

其中 n_i 是第 i 种基本构成粒子的数密度, q_i 是这种基本粒子的带电量. dV 则为微元体积. 而上式也被写为:

$$dQ = \rho dV ; \quad \rho = \sum_i n_i q_i = \sum_i \rho_i$$

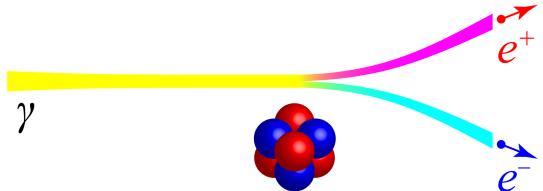
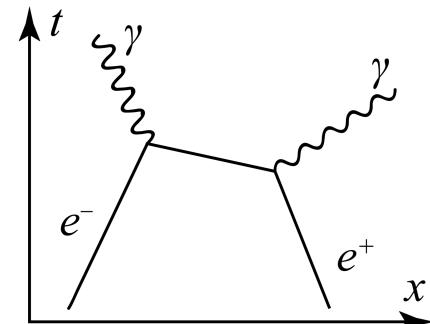


图 1.1: 电子对湮灭与产生

¹实际上物理上不允许超距的施力物体, 施力物体就是局域的电场.

²与磁单极子的存在与量子化有关.

³是的, 这种思考方式就是人择原理(anthropic principle).

⁴这个值是 2018 新国际标准单位制提出的规定值.

1.1.2 库仑定律

库仑定律(Coulomb's law) 用于描述真空中的两个孤立静止点电荷之间的相互作用力:

$$\mathbf{F}_{12} = \frac{1}{4\pi\epsilon_0} \frac{q_1 q_2}{r^2} \mathbf{e}_{12}$$

其中系数 $k_e = \frac{1}{4\pi\epsilon_0} \approx 8.99 \times 10^9 \text{ N} \cdot \text{m}^{-2} \cdot \text{C}^{-2}$ 为库仑常数.

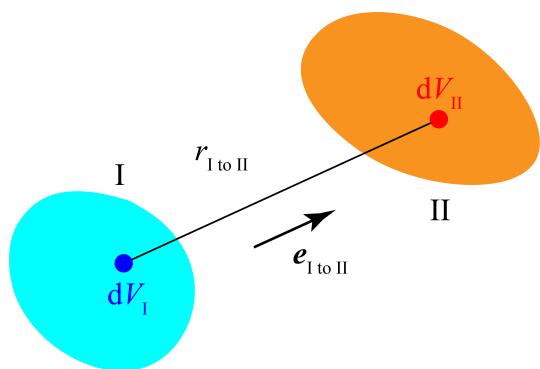


图 1.2: 电荷体系间的相互作用力
加本性. 一个复杂电荷系统 (包括构成介质的电荷) 可以用连续的电荷密度描述:

$$dQ = \rho dV$$

那么两部分电荷之间的相互作用力为:

$$\mathbf{F}_{I \text{ to } II} = \frac{1}{4\pi\epsilon_0} \int_{V_{II}} \rho(\mathbf{r}_{II}) \int_{V_I} \frac{\rho(\mathbf{r}_I)}{r_{I \text{ to } II}^2} \mathbf{e}_{I \text{ to } II} dV_I dV_{II}$$

相互作用力在定义上就是成对存在的, 所以两部分电荷体系间的相互作用力满足牛顿第三定律: 大小相等, 方向相反, 对整个体系产生的力矩的贡献为零.

1.1.3 电场

现代物理认为, 狭义相对论下的时空观是被事实验证了的. 一个重要的推论就是**局域性原理**(principle of locality): 物体仅仅被此时此刻周围的物质所影响. 所有相互作用必须被理解为**近距作用**(action upon contact) 而不能是**超距作用**(action at a distance). 这是为了能与更本质的**因果律**(principle of causality) 相适应. 如果物体这一时刻的状态及其改变取决于同一时刻的空间上相隔一定距离的另外一点 (类空间隔). 那么就可以另一个点的事件称作因, 而物体的状态与状态改变称作果. 但类空间隔在洛伦兹变换下具有可逆性. 也就是原则上可以在果的时空点也可以制造事件去对因的点产生影响. 从而造成逻辑上的紊乱. 所以我们必须规定这两个时空点间不存在因果关系, 互相独立没有影响.

局域性原理的直接结论便是, 在引力与电磁相互作用时, 作为物质的**场**(field) 的概念的引入. 场本是数学概念. 在一个三维的欧几里得空间 \mathbb{R}^3 中, 如果每一个点都定义了一个数 (标量), 代表某种物理量, 物理现象可以在不同参考系, 惯性系中观察, 所对应的时空点的位置坐标可能发生改变, 但特定的时空点的概念, 和与之相结合的物理现象与描述它的这个数, 不应该发生改变, 称为**标量**(scalar):

$$f: \mathbf{r} \mapsto f(\mathbf{r})$$

⁵而在经典电磁场论的观点下库仑定律为其矢量场论形式的推论.

如,压强场,温度场,波动的相位场,都是典型的标量场的例子.数学上定义了标量场的梯度(gradient)来描述在一点极其附近的标量场的行为,它是一个矢量,方向是使得标量上升最快的欧几里得空间中的方向,大小是沿这个方向的标量变化的方向导数.数学上可以证明:

$$\nabla f = \frac{\partial f}{\partial x} e_x + \frac{\partial f}{\partial y} e_y + \frac{\partial f}{\partial z} e_z$$

物理量在某点 r 的局部,偏离一个 dr 后值变为:

$$f(r + dr) = f(r) + \nabla f \cdot dr$$

这就是微元形式的牛顿-莱布尼茨定理(Newton-Leibniz theorem).而标量场的梯度往往也具有重要的含义.压强场的梯度决定浮力的方向与大小(反向),温度场的梯度决定热传导的方向与强弱(也是反向),波动的相位场的梯度则决定波的传播方向与波长(波矢).

而在三维的欧几里得空间中在每一个空间点也可以引入一个矢量(vector),它不占三维空间的有限体积⁶而是定义在一个点上,可以认为是这个点上扩展出了一个新的与原三维空间坐标平行的 $e_x - e_y - e_z$ 坐标系.而这一点定义的矢量就位于这个“浓缩在一点”的坐标系中⁷:

$$\mathfrak{F}: r \mapsto \mathbf{F}(r)$$

显然,标量场的梯度恰是一个矢量场.矢量不能简单理解为三个标量场的合成.因为在空间旋转(或时空的惯性系变化)下,标量应该是不变的,而矢量将发生与坐标系相似的旋转,从而三个分量之间相互转化,只有矢量的模长是不变的.另一种典型的矢量场是流体的流速场 v ,或者更普遍的,某种物理属性的流密度场 j ,前者定义为单位时间通过单位面积的体积,后者则推广为单位时间通过单位面积的某物理量 Q :

$$dQ = j \cdot dA dt$$

在一点附近的矢量场的行为则比标量场复杂很多:

$$\mathbf{F}(r + dr) = \mathbf{F}(r) + dr \cdot \nabla \mathbf{F}$$

此处 $\nabla = e_x \frac{\partial}{\partial x} + e_y \frac{\partial}{\partial y} + e_z \frac{\partial}{\partial z}$ 可以形式地视为矢量微分算符,它写在上一式的后一项中实际上是在表示:

$$\begin{aligned} d\mathbf{F} &= dF_x e_x + dF_y e_y + dF_z e_z \\ &= \left(\frac{\partial F_x}{\partial x} dx + \frac{\partial F_x}{\partial y} dy + \frac{\partial F_x}{\partial z} dz \right) e_x \\ &\quad + \left(\frac{\partial F_y}{\partial x} dx + \frac{\partial F_y}{\partial y} dy + \frac{\partial F_y}{\partial z} dz \right) e_y \\ &\quad + \left(\frac{\partial F_z}{\partial x} dx + \frac{\partial F_z}{\partial y} dy + \frac{\partial F_z}{\partial z} dz \right) e_z \\ &= \left(dx \frac{\partial}{\partial x} + dy \frac{\partial}{\partial y} + dz \frac{\partial}{\partial z} \right) (F_x e_x + F_y e_y + F_z e_z) \\ &= dr \cdot \nabla \mathbf{F} \end{aligned}$$

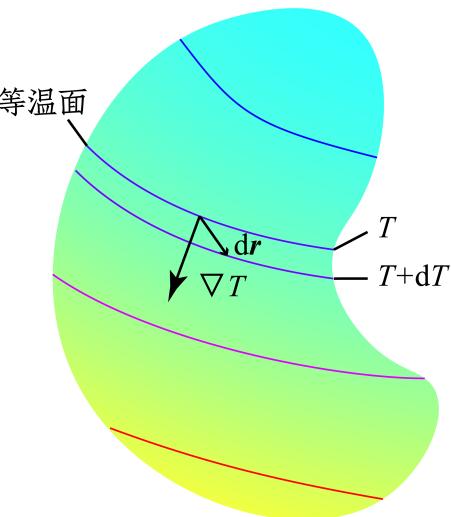


图 1.3: 温度场的梯度

⁶即,不同于位移矢量那样必须与三维空间中的不同点同时建立联系

⁷数学术语把空间叫做微分流形(differentiable manifold),而局部的矢量位于的空间叫做纤维(fiber)或切空间(tangent space),每一点都有一个这样的空间,合称纤维丛(fiber bundle).

而也可以把 ∇F 认为能完整描述这个矢量在该点处的一阶变化行为,也称为矢量的梯度,但它实际上由九个分量组成,分别描述三个分量在三个方向的变化率,一般写在一张数表中,也就是矩阵⁸:

$$\nabla \mathbf{F} = \begin{bmatrix} \frac{\partial F_x}{\partial x} & \frac{\partial F_y}{\partial x} & \frac{\partial F_z}{\partial x} \\ \frac{\partial F_x}{\partial y} & \frac{\partial F_y}{\partial y} & \frac{\partial F_z}{\partial y} \\ \frac{\partial F_x}{\partial z} & \frac{\partial F_y}{\partial z} & \frac{\partial F_z}{\partial z} \end{bmatrix}$$

显然用这样的一个矩阵去描述矢量场在一点处的行为是代数的,抽象地,不几何直观的.所以我们引入两个物理上会更实用的概念.

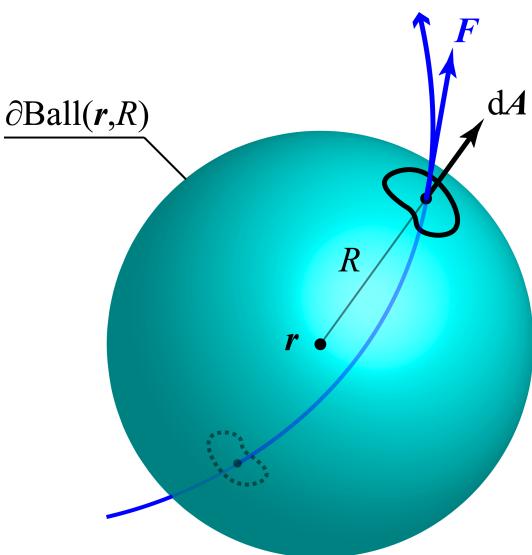


图 1.4: 矢量场的散度

一是矢量场的散度(divergence),它是一个标量,定义为矢量场的源强度:

$$\nabla \cdot \mathbf{F} = \lim_{R \rightarrow 0} \left(\oint_{\partial \text{Ball}(r, R)} \mathbf{F} \cdot d\mathbf{A} \right) / \left(\frac{4}{3}\pi R^3 \right)$$

其中 $\text{Ball}(r, R)$ 表示以待研究的点为球心的,半径为 R 的球形体,而 ∂Ball 表示它的表面.积分号 \oint 用来代表此场合下用来积分的区域是一个闭合的区域⁹.积分的量称为通量(flux),即矢量场与面积的电场,速度场情况下即单位时间通过面 $d\mathbf{A}$ 的体积.现在对闭合的面进行积分,再除以面包含的体积,代表的是该点单位体积内产生这个矢量场的场线的源的多少,即源强度.数学上可以证明,它与体积以怎样的取法趋于零无关,如果取成与坐标系平行的长方体的形状,能够在坐标系下计算出它的值为上面那个矩阵的迹:

$$\nabla \cdot \mathbf{F} = \frac{\partial F_x}{\partial x} + \frac{\partial F_y}{\partial y} + \frac{\partial F_z}{\partial z}$$

第二个概念便是旋度(curl)了.它实际上代表矢量场的变化矩阵中反对称的那一部分¹⁰.它可以被写为一个矢量.与在这一点处的无限小的环积分有关,数学上不难证明:

$$\mathbf{n} \cdot \nabla \times \mathbf{F} = \lim_{R \rightarrow 0} \left(\oint_{\partial \text{Circ}(r, n, R)} \mathbf{F} \cdot dl \right) / (\pi R^2)$$

由于半径为 R 的圆的选取还依赖于其朝向.故在上式中我们对积分的方向和垂直于积分圆面的法向单位矢量 n 采取了右手定则式的对应.而最后积分出来得到的数恰好是面法向单位矢量与从 \mathbf{F} 的各个导数中提取的旋度矢量的点乘.环积分通常被称为环量(circulation),旋度则反应矢量场局域的涡旋程度.数学上可以证明旋度可以由行列式表示为:

$$\nabla \times \mathbf{F} = \begin{vmatrix} \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ F_x & F_y & F_z \\ e_x & e_y & e_z \end{vmatrix}$$

⁸实际上,形成了一个张量场(tensor filed).张量将不在本部分教材的讨论范围内.

⁹数学上称为又闭又开(clopen)的集合,表示它没有边界.

¹⁰任何一个矩阵 A 都可以被唯一地分解为一个对称矩阵 $(A + {}^t A)/2$ 与反对称矩阵 $(A - {}^t A)/2$ 之和,前者有六个独立分量,后者仅有三个独立分量.

作为与牛顿-莱布尼茨定理对应的散度与旋度版本,根据两个量的定义可知前者发生在体积内部的积分与其闭合表面上,后者发生在曲面上面积分和它的闭合曲线边界上:

$$\int_V \nabla \cdot \mathbf{F} dV = \oint_{\partial V} \mathbf{F} \cdot d\mathbf{A}$$

$$\int_A \nabla \times \mathbf{F} \cdot d\mathbf{A} = \oint_{\partial A} \mathbf{F} \cdot dl$$

前者为[奥斯特诺格拉德斯基-高斯定理](#)(Ostrogradsky-Gauss theorem),后者则为[开尔文-斯托克斯定理](#)(Kelvin-Stokes theorem)¹¹.

现在让我们回到库仑定律,现在就必须把它理解为两个公式:

$$\mathbf{E} = \frac{1}{4\pi\epsilon_0} \frac{Q}{r^2} \mathbf{e}_r ; \quad \mathbf{F} = q\mathbf{E}$$

前一半式子说明,即使没有引入任何试探电荷 q ,在空间中距离场源电荷 Q 为 r 处依然存在某种独立于电荷而存在的“场”.否则也就不可能对引入的电荷瞬间产生力的作用.这个场由电荷激发,以后将知道它有自己的能量与动量,而且都是可以被局域地确定¹².经典物理中似乎只能由电荷产生场,但高能物理下场也能反过来产生电荷粒子,实际上此时会将电磁场理解为“光子”这样的物质.电磁场的特点是能对带电荷的物质产生作用.

而后一半的式子则给出了电磁场中一点处电场强度矢量 \mathbf{E} 的定义.带电粒子如果在电磁场中静止,那么将受到一个与电荷量成正比的力.把力与电荷量的比值定义为电场强度矢量.

显然,带电粒子的引入也会产生电场.这个电场与原来场源电荷产生的电场叠加,形成了空间的真实电场分布.但在计算这个“试探电荷”处的电场时,不应该考虑自己在自己这一点处产生的电场.因为自己对自己的力总是零.而总是计算“外场”在试探电荷这一点处的值.不过,试探电荷的引入与否可能会对外界的场源电荷部分产生改变.从而改变外场.这又需要另说了,我们下一章的导体静电感应就是在不断讨论这个问题.

电场满足[叠加原理](#)(superposition principle).我们总是一对一对地考虑电荷之间的二体作用,把相互作用力做非相干的矢量叠加.这也就是说,给定一个场源电荷分布,可以叠加计算出空间中任意一点的电场强度:

$$\mathbf{E}(\mathbf{r}) = \int_{V'} \frac{1}{4\pi\epsilon_0} \frac{\rho e_R}{R^2} dV' , \quad \mathbf{R} = \mathbf{r} - \mathbf{r}'$$

$$\mathbf{E}(\mathbf{r}) := \frac{\mathbf{F}_{\text{static}}}{q}$$

也是出于这个原因,所有由电场强度派生出来的量:电势,电通量等等都符合线性的叠加原理.马上就会介绍,这是由于描述电场与电荷关系的基本方程是一个线性的微分方程的缘故.

1.2 两个定律与电势

库仑定律是静电学最基本的公理,但我们研究电场作为场的性质,它与电荷的关系.为以后深入了解电磁场整体打好基础.

¹¹微分流形上的微分形式有一个[嘉当-斯托克斯定理](#)(Cartan-Stokes theorem) 统一了以上所有形式:

$$\oint_{\partial\Omega} \omega = \int_{\Omega} d\omega$$

¹²即,可以确定每一点的场能量动量密度.

1.2.1 电场的高斯定律

首先让我们研究电通量(electric flux),通过一个曲面(不一定闭合)的电通量定义为曲面上电场强度的积分:

$$\Phi = \int_A \mathbf{E} \cdot d\mathbf{A}$$

我们在这儿必须引入有向曲线,曲面,体积的概念.固然,曲线,曲面,体积都没有方向与正负,它们仅仅是特定空间中的几何形状.但定义各种积分时,曲线积分的线元矢量 $d\mathbf{l}$,曲面积分的面元矢量 $d\mathbf{A}$ 都有两个可能的方向.而体积元 dV 也不一定恒正,在某些情况下将取负.那么我们通常会有一个符号约定:

- 对于可定向(orientable)曲面(不可定向的比如莫比乌斯环),总是定义一个方向为正,即为每一个面元指定一个面元矢量,它取垂直于面的两个方向之一,而相邻的面元具有相同的取向.而如果涉及到面边界上闭合曲线的环积分,则线元矢量取的方向与面元构成右手定则(right-hand rule).
- 对于连通.connected)体积区域,如果取体积微元为正,那么涉及到区域边界闭合曲面上的面积分时,面元取从体积内到体积外的方向.如果取体积微元为负,则面上面元取体积外到体积内的方向.

而我们之前写出的奥-高定理和开-斯定理,都是在这个意义下正确的.

从而以上电通量也就因为我们选择的面的两种可能的正向而造成两个不同的结果,它们相差一个相反数.

电场的高斯定律(Gauss's law)指出:

$$\Phi_{\partial V} = \oint_{\partial V} \mathbf{E} \cdot d\mathbf{A} = \frac{1}{\epsilon_0} \int_V \rho dV$$

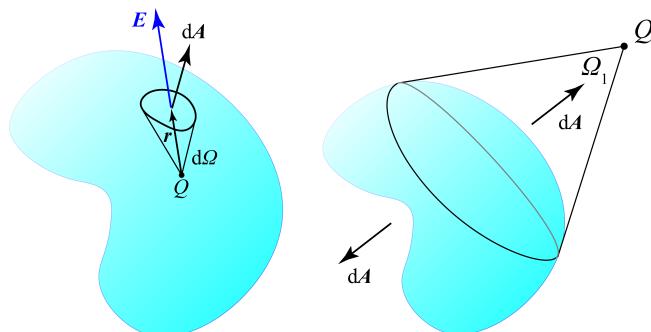


图 1.6: 证明高斯定律

我们来用库仑定律证明高斯定律.首先考虑一个在闭合曲面内部的点电荷,我们发现面元上的通量:

$$d\Phi = \mathbf{E} \cdot d\mathbf{A} = \frac{1}{4\pi\epsilon_0} \frac{Q e_r \cdot d\mathbf{A}}{r^2}$$

而这一个小面元对点电荷张的立体角恰好是:

$$d\Omega = \frac{e_r \cdot d\mathbf{A}}{r^2}$$

故对面积分总立体角给出 4π 的结果:

$$d\Phi = \frac{Q}{\epsilon_0} \cdot \frac{d\Omega}{4\pi} \Rightarrow \Phi = \frac{Q}{\epsilon_0} \cdot \oint_{\partial V} \frac{d\Omega}{4\pi} = \frac{Q}{\epsilon_0}$$

我们再考虑在面外的点电荷 Q 对面产生的电通量,此时面的正向与点电荷产生电场方向的点乘就有了正负号,积分以后得到的结果是通过点电荷去看曲面,把曲面对点电荷张的最大视角处的曲线找到,将曲面

分为两个部分,立体角都是 Ω_1 ,总积分为:

$$\Phi = \frac{Q}{\varepsilon_0} \cdot \oint_{\partial V} \frac{d\Omega}{4\pi} = \frac{Q}{\varepsilon_0} (\Omega_1 - \Omega_1) = 0$$

从而我们证明了,电荷若在闭合曲面内,则电通量贡献为 $\frac{Q}{\varepsilon_0}$.而若在闭合曲面外,则电通量贡献为 0.再由叠加原理,得任意情况下闭合曲面的电通量等于内部总电荷量除真空介电常数 ε_0 .考虑电荷连续分布或准连续分布的情况下,有:

$$\oint_{\partial V} \mathbf{E} \cdot d\mathbf{A} = \frac{1}{\varepsilon_0} \int_V \rho dV$$

此即高斯定律,由数学上的奥-高定理,我们可以把上式写为局域的微分形式:

$$\nabla \cdot \mathbf{E} = \frac{\rho}{\varepsilon_0}$$

高斯定律旨在说明:电荷产生电场(电荷是电场的源强度).

1.2.2 电势与电场的环路定律

对点电荷产生的电场 $\mathbf{E} = \frac{Q}{4\pi\varepsilon_0 r^2} \mathbf{e}_r$,再引入以下标量场是方便的:

$$\varphi = \frac{1}{4\pi\varepsilon_0} \frac{Q}{r}$$

这是因为如果计算球坐标下的梯度:

$$\nabla = \mathbf{e}_r \frac{\partial}{\partial r} + \frac{\mathbf{e}_\theta}{r} \frac{\partial}{\partial \theta} + \frac{\mathbf{e}_\varphi}{r \sin \theta} \frac{\partial}{\partial \varphi}$$

便得到以下关系式:

$$\mathbf{E} = -\nabla \varphi$$

即电场可以表示为我们引入的另一个标量场的负梯度的形式.这个标量场就叫做**电势**(electric potential).而以上关系,结合梯度的含义理解,便是沿电场线方向电势降低.可以通过牛-莱定理写为积分形式(下式与路径无关):

$$\int_A^B \mathbf{E} \cdot d\mathbf{l} = \varphi_A - \varphi_B$$

以上关系式由叠加原理,在任意电荷分部的情况下都成立,此时对应的电势标量场为:

$$\varphi(\mathbf{r}) = \int \frac{1}{4\pi\varepsilon_0} \frac{\rho}{|\mathbf{R}|} dV' , \quad \mathbf{R} = \mathbf{r} - \mathbf{r}'$$

对于电势的引入可能存在一点疑惑:我们都应该知道电场的定义是可以独立于场源电荷的,它为在一点处引入静止的试探电荷时单位电荷的受力.那么电势是否可以独立于场源电荷而定义呢?观点一比较基础,它认为在静电场情况下既然全空间电场已经被定义了,它就应该反映了该点处场的全部实际信息,那么电势作为与电场符合一个梯度关系的梯度场,也就能被相应的确定下来,两个可能的电势还可以差一个全空间都一致的常数.也就是这种观点把电场视为更加基础的场的特征.下面便可以知道电势的含义是引入试探电荷后单位试探电荷在该点处的电势能.而观点二则更加本质,我们今天知道了势比场强实际上包含更多的信息.由于微观粒子的量子特性,力的概念已然失效,从而场强反而不如用势来描述一个场更为直接而实用.电磁场论实际上也就是一个矢量场论,它由矢势和标势组成一个四维矢量来完整地描述,但取法不唯一,多出来的冗余自由度被称为**规范**(gauge),它恰好就在表示电荷与场的相互作用时,与电荷的量子相位协同变化,从而造成实际可以观测的影响.为了实际问题的方便,通常约定势有特定的取法.在静电学的常见情况下(电荷分布区域有限时),常常取无穷远处电势为零,注意上电势积分式也恰恰符合这一点.

我们进一步考察电势的物理含义,如果把以上电场视为外电场,引入试探电荷 q ,并在以上积分形式的场强与势的关系中乘以 q ,便发现 $q\mathbf{E}$ 就是电荷的受力,左边就是表示电场力的做功:

$$W_{A \rightarrow B} = q\varphi_A - q\varphi_B$$

很自然地我们发现电场力是保守力,而对应的电势能,基于它应该与电荷量成正比的观点,我们总是直接取为:

$$E_p = q\varphi$$

从而电势的效果,是为试探电荷带来一个电势能.

最后我们再考察电场的一个特性,由于任何路径上的电场积分可以根据牛-莱定理写为端点的电势差的形式,考虑闭合的曲线积分,它可以看成一个可定向曲面的边界,由于两端点重合:

$$\oint_{\partial A} \mathbf{E} \cdot d\mathbf{l} = 0$$

这就是电场的环路定律(circuital law),再根据数学上的开-斯定理,我们写出环路定律的微分形式:

$$\nabla \times \mathbf{E} = \mathbf{0}$$

也就是静电场无旋.值得指出这不过是静电场可以表示为标量电势场的负梯度的直接结果,因为任意一个标量场的梯度都是无旋的¹³:

$$\oint_{\partial A} \nabla f \cdot d\mathbf{l} = 0 \Rightarrow \nabla \times \nabla f = \mathbf{0}$$

1.2.3 总结

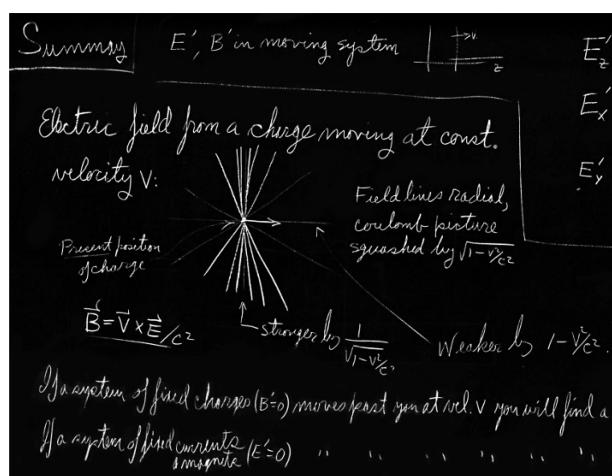


图 1.7: Feynman 讲座一瞥

我们现在有三个不同的层次去看待电荷与电荷之间的相互作用问题¹⁴:

1. 不引入场:

$$\mathbf{F} = \frac{1}{4\pi\epsilon_0} \frac{Qq}{r^2} \mathbf{e}$$

2. 引入场,用场强描述:

$$\begin{aligned} \nabla \cdot \mathbf{E} &= \frac{\rho}{\epsilon_0} ; \quad \nabla \times \mathbf{E} = \mathbf{0} \\ \mathbf{F} &= q\mathbf{E} \end{aligned}$$

3. 引入场,用势描述:

$$\nabla^2 \varphi = -\frac{\rho}{\epsilon_0}$$

$$E_p = q\varphi$$

¹³这个结论向流形推广就是著名的庞加莱引理(Poincaré lemma)了,它指出任意微分形式 ω 的二次全微分为零:

$$d\omega = 0$$

一个以后会用得上的推论可以通过考察矢量场的旋度在闭合曲面上的面积分,而把闭合曲面看成是边界曲线缩小到零的有边界曲面而得:

$$\oint_{\partial V} \nabla \times \mathbf{F} \cdot d\mathbf{A} = 0 \Rightarrow \nabla \cdot \nabla \times \mathbf{F} = 0$$

¹⁴拉普拉斯算子 $\nabla^2 = \nabla \cdot \nabla$,如作用在标量上即表示先求梯度后求散度:

$$\nabla^2 f = \nabla \cdot (\nabla f)$$

以上描述一个比一个适用范围更广,第一个仅仅适用于静电场或低速情况,第二个则是对高速运动的电荷与电磁场适用,比如匀速运动的电荷由简单相对论讨论可得:

$$\mathbf{E} = \frac{1}{4\pi\epsilon_0} \frac{Q}{r^2} \frac{1 - \beta^2}{(1 - \beta^2 \sin^2 \theta)^{3/2}} \mathbf{e}_r$$

$$\mathbf{B} = \frac{\mathbf{v} \times \mathbf{E}}{c^2}$$

它要符合第二个式子(旋度需要做相对论修正),而第三个式子更是适用于非经典场论情况.

1.3 静电能

我们将进一步深入研究电场相互作用中的能量概念.

1.3.1 静电势能

之前已经引入了电势能的概念,我们把外场中引入的试探电荷(点电荷)所具有的势能的大小计做电势能,并得出了电势能的表达式:

$$E_p = q\varphi$$

还指出了,它的作用是用来表示不变的外电场对电荷的做功的多少,基于这样的观察,我们把这个能量称为**静电势能**(electrostatic potential energy),静电的含义,是要求外场不能随着试探电荷的引入与移动而改变.而更常见的一类情况是,随着试探电荷的移动,场源电荷也在随着改变,如之后将会经常讨论的静电感应的情况.此时仍然可以计算外场的场强与电势,并引入试探电荷以后依然可以形式地写出**电势能**(electro potential energy)的值 $E_p = q\varphi$,但它不再代表实际移动试探电荷电场对电荷做的功了,它代表什么?电场对电荷做功普遍地应该怎么表示?外力做功又如何和它发生联系,这是我们之后的讨论关心的问题.

首先对于这个问题的第一步考察是,虽然试探电荷的特定位置带来特定的外场,但一旦确定了外场的场源的电荷分布,便可以独立于试探电荷地产生一个分布于全空间的电势与电场,而可以设想固定这个电荷分布不变,而把试探电荷放在空间中的任意一点,讨论其受力与能量.这些改变试探电荷位置而不改变外场的位移称为**虚位移**(virtual displacement).它是一种广泛用来建立理论模型的方法.它的强大之处在于虚位移的任意性,在每一个实际时刻试探电荷可以在任意位置,外场在缓慢随时间发生改变,而试探电荷发生的实际位移是所有虚位移中的一个特例.在这个意义上,电势能的含义便在于,试探电荷发生虚位移时,即场源电荷分布不变时,电场力对试探电荷做的虚功即为电势能的减小,此即静电场下的虚功原理:

$$\delta E_p = -q\mathbf{E} \cdot \delta\mathbf{r}$$

我们可以把这个概念推广,如果在该外场中引入不止一个点试探电荷,我们写出一个电荷体系在外场中的电势能来:

$$E_p = \sum_i q_i \varphi_i$$

这个写法存在一些问题,首先是如果没有引入任何试探电荷,体系是否固有一个能量?用场的角度来理解就会发现不对,场是一种物质,物质必定伴随着能量.或者考虑场源电荷之间的相互作用,它也一定会伴随着能量.再者,引入两个以上的点电荷时,它们之间的相互作用能没有被包含在以上项中.第三,即使是单个点电荷单独存在,也必定会带来其自能.我们需要从更基础的角度来讨论电荷体系的能量.

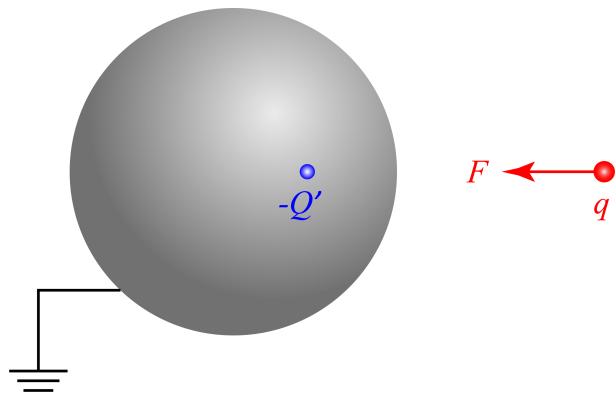


图 1.8: 情形: 场源电荷取决于试探电荷

1.3.2 自能与相互作用能

考虑真空中的两个点电荷之间的相互作用. 电荷 q_1 在电荷 q_2 处产生的电场与电势为 \mathbf{E}_{12} 与 φ_{12} , 电荷 q_2 在电荷 q_1 处产生的电场与电势为 \mathbf{E}_{21} 与 φ_{21} 无疑, 相互作用力是相互的, 满足牛顿第三定律 $q_1\mathbf{E}_{21} + q_2\mathbf{E}_{12} = \mathbf{0}$. 那么电势能是相互的吗? 注意到:

$$q_1 U_{21} = \frac{1}{4\pi\epsilon_0} \frac{q_1 q_2}{r} = q_2 U_{12}$$

从而两个电势能实际上是相等的. 根据电势能的定义, 它等于 1 不动 2 移到无穷远电场力所做的功, 也会等于 2 不动 1 移到无穷远电场力所做的功. 这两种情况平移对称, 从而我们有理由判断: 无论如何移动两个电荷, 将 1 与 2 分开到无限远时电场力做功都会是以上值, 这一点也不难在数学上证明.

从而我们把这个能量称为 12 间相互作用力对应的相互作用能(interaction energy). 它可以用任何一个电荷乘以另一个电荷在这个电荷处产生的势相乘得到, 或者找到这一对相互作用, 直接写出:

$$I_{12} = \frac{1}{4\pi\epsilon_0} \frac{q_1 q_2}{r_{12}}$$

按照这思路, 由叠加原理, 我们可以很轻易地写出点电荷体系与连续分布电荷体系的总相互作用能:

1. 点电荷体系:

$$I = \sum_{\{i,j\}!} \frac{1}{4\pi\epsilon_0} \frac{q_i q_j}{r_{ij}}$$

2. 连续分布电荷体系:

$$I = \iint_{\{\mathrm{d}V, \mathrm{d}V'\}!} \frac{1}{4\pi\epsilon_0} \frac{\rho \rho'}{r} \mathrm{d}V \mathrm{d}V'$$

关于以上求和我们做出以下解释:

首先, 点电荷体系的相互作用能表示把点电荷从计算这个能量的状态移动到点电荷彼此之间都相距无穷远时电场力所做的功. 但一定要注意每个点电荷移动到无穷远以后还会因为自身的电磁相互作用而产生一个待定的能量, 叫做自能, 我们紧接着会讨论到这个问题. 而连续分布电荷体系则表示把每一份电荷微元移动到无穷远以后电场力做的功, 移动到无穷远以后电荷无限分散(宏观意义上), 所以每一份微元电荷对应的电磁相互作用能真实地趋于零, 故我们计算的这个能量就是这个体系的总电磁相互作用对应的能量.

第二, 是这个求和表示的含义. 我们用 (i, j) 来表示特定集合中两个元素 i, j 构成的有序对(ordered pair), 而 $\{i, j\}$ 则表示交换 i, j 后表示同一代数元素的¹⁵对称无序对(symmetrized unordered pair), 即 $\{i, j\} = \{j, i\}$. 求和下加这个对符号表示对所有一对一对的情况进行求和. 每一对只需要要求一次即可. 这之后还跟着一个感叹号, 它表示 i, j 必须取不同值(而 $\mathrm{d}V$ 与 $\mathrm{d}V'$ 可以取相同值, 这些项会在体积很小时求和也很小而可以忽略). 那么由简单的代数变形, 我们发现上两式即:

$$I = \frac{1}{2} \sum_i \sum_{j \neq i} \frac{1}{4\pi\epsilon_0} \frac{q_i q_j}{r_{ij}} = \frac{1}{2} \sum_i q_i \phi_i$$

$$I = \frac{1}{2} \int \int_{\mathrm{d}V \mathrm{d}V'} \frac{1}{4\pi\epsilon_0} \frac{\rho \rho'}{r} \mathrm{d}V \mathrm{d}V' = \frac{1}{2} \int_{\mathrm{d}V} \rho \varphi \mathrm{d}V$$

注意到第一式中 ϕ_i 表示的是别的电荷在 q_i 处产生的电势, 不包括自己在自己上产生的, 我们把符号稍微变动一下以示区别. 而在后一个表达式中则不需要考虑这种区别, 这个式子将作为一个普遍的求总电磁相互作用能的表达式而使用.

¹⁵以后还会用到 n 个元素的对称无序类(symmetrized unordered setoid) 与反对称无序类(antisymmetrized unordered setoid):

$$\{i_1 \cdots i_p \cdots i_q \cdots i_n\} = \{i_1 \cdots i_q \cdots i_p \cdots i_n\}$$

$$[i_1 \cdots i_p \cdots i_q \cdots i_n] = -[i_1 \cdots i_q \cdots i_p \cdots i_n]$$

第三,是这个求和的使用方法.实际上,如果电荷连续分布系统从一个状态变化到另一个状态,电荷分布与它产生的电势分布都发生改变,那么前后的能量增加:

$$\Delta I = \frac{1}{2} \int \rho' \varphi' dV - \frac{1}{2} \int \rho \varphi dV$$

自然表示电场对电荷体系所做的负功,或者表示由电荷流向电场的能量,它使得场的能量,也就是 I ,发生了增长¹⁶.而反过来考察电荷体系,如果外力对电荷体系做功为 W ,那么根据动能定理应该有:

$$W - \Delta I = \Delta E_k$$

在十分缓慢地移动电荷的情况下,亦或是导体在载流子为电子,其动能十分的小可以忽略的情况下,动能的增长可以视为零,从而有:

$$W = \Delta I$$

即,外力做功,通过电荷为媒介,转化为了场的能量的增长.这个观点也是我们经常用到的.

我们最后回到电势能与相互作用能的关系,考虑整个体系由两部分电荷体系构成,那么由叠加原理,整个空间中每一点的电势可以写成两部分贡献的和 $\varphi_1 + \varphi_2$.我们只需要对全空间的电荷积分,就得到了整个体系的相互作用能:

$$\begin{aligned} I &= \frac{1}{2} \int_{I+II} \rho \varphi dV \\ &= \frac{1}{2} \int_I \rho_1 (\varphi_1 + \varphi_2) dV_1 + \frac{1}{2} \int_{II} \rho_2 (\varphi_1 + \varphi_2) dV_2 \\ &= \frac{1}{2} \int_I \rho_1 \varphi_1 dV_1 + \frac{1}{2} \int_{II} \rho_2 \varphi_2 dV_2 + \frac{1}{2} \left(\int_I \rho_1 \varphi_2 dV_1 + \int_{II} \rho_2 \varphi_1 dV_2 \right) \\ &= I_1 + I_2 + I_{12} \end{aligned}$$

我们于是找到了三个能量, I_1, I_2 是两个电荷分布体系单独存在时各自的自相互作用能, 我们把它简称作**自能**(self-energy), 它一定是大于零的, 这一点下一节就可以发现. 而余下的交叉求和项就是表示两个体系之间的相互作用能. 仔细一看不难发现一项其实就是电荷体系 II 产生的外场下电荷体系 I 的电势能, 而另一项就是电荷体系 I 产生的外场下电荷体系 II 的电势能, 两者仔细研究下发现其实相等:

$$I_{12} = \iint_{I II} \frac{1}{4\pi\epsilon_0} \frac{\rho_1 \rho_2}{r} dV_1 dV_2 = \int_I \rho_1 \varphi_2 dV_1 = \int_{II} \rho_2 \varphi_1 dV_2$$

这一项可以大于零也可以小于零, 表示的是相互作用. 现在我们试探电荷体系在外场中的电势能有了一个新的看法, 它其实表示试探电荷体系与场源电荷体系的相互作用能, 作为相互作用, 它也可以反过来用场源电荷体系在试探电荷体系产生的场中的电势能来计算, 或是对称地, 用两者和的一半来计算.

最后让我们来看一下点电荷的自能. 实际上在之前点电荷体系相互作用能体系中我们已经指出, 其表达式中电荷 q_i 上电势 ϕ_i 与真实电势 φ_i 的区别就在于没有计及点电荷在自己位置处产生的电势 φ_{ii} , 从而漏掉的项其实就是点电荷的自能:

$$I_i = \frac{1}{2} q_i \varphi_{ii}$$

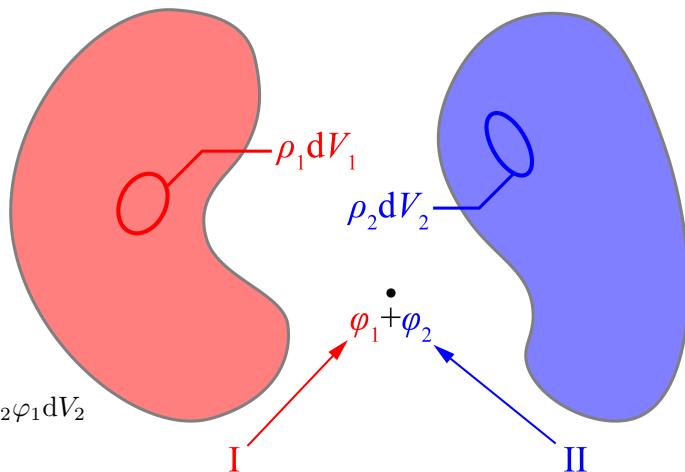


图 1.9: 自能与相互作用能

¹⁶场只能和带荷的物质发生作用, 这也是唯一的使得场能量变化的方法

这些项不会影响我们的实际计算,因为点电荷总被我们认为是没有内部改变的理想体系,我们经常把微观基本粒子做这样的抽象,因为它们都是小到可以忽略其尺寸.当粒子速度远小于光速,且不涉及到高能粒子反应的情况下,它们也的确不发生任何内部的改变,自能精确地保持不变.所以只需要计算扣除点电荷自能的相互作用能表达式,就能够知道有多少势能被存储在场中,就能够用它来写出电荷分布发生改变时释放出来的能量.

点电荷的自能如何计算呢?这取决于我们采取的模型.如果我们考虑的是一个电荷均匀分布在表面的球体,那么自能为:

$$I = \frac{1}{2}Q \cdot \frac{1}{4\pi\epsilon_0} \frac{Q}{r}$$

对于电子,如果我们认为电子的静质量全部来源于它携带的场的惯性,那么能够写出:

$$I = \frac{e^2}{8\pi\epsilon_0 r} = m_e c^2$$

事实上我们并不能确定电子内部电荷分布,所以我们更常直接定义**经典电子半径**(classical electron radius) 为:

$$\frac{e^2}{4\pi\epsilon_0 r_e} = m_e c^2 \Rightarrow r_e = \frac{1}{4\pi\epsilon_0} \frac{e^2}{m_e c^2} = 2.82 \times 10^{-15} \text{ m}$$

然而这仅仅是一个方便计算的实用常数,并不能代表电子的真实情况!事实上用不带电荷的粒子与电子发生散射实验就会发现,电子的尺寸比这要小得多,上面的计算结果甚至比质子的半径还要大几倍!而实验更是发现电子在多种意义上就是一个点粒子,根本没有内部组成结构,十分稳定不能衰变.那么显然从上式发现这也面临着自能发散的问题.双重矛盾如何解决呢?

解决方案是量子力学的引入,电子不是点粒子,在原子尺度下电子就已经必须要由波函数描述了,电子没有同时确定的位置和速度,在原子外形成“电子云”.狄拉克(P. Dirac)于1928年提出**狄拉克方程**(Dirac equation),把电子像光子那样处理为一个传播的“场”,从而从根本上摒弃了电子作为点电荷的观点,电子的质量被理解为场的质量或者静能,而电子的电荷量被理解为与电磁场(光子)的耦合常数.从而把电荷作为一种物质和场物质彻底分割开来独立描述.还同时一并给出了电子的自旋与预言了电子的反物质:**正电子**(positron)的存在.它们都被天然而巧妙的包含在一个统一的代数形式中,极大程度地推动了物理学理论的进展.

但这没有给出问题的全部解释,对于点电荷自能发散的问题这种做法仍然没有给出绝对的答案.经典地看一个匀速运动的经典无穷小点电荷会携带着一个发散的场共同向前运动,这在量子场论语言下,变成了电子与光子相互作用的无限可能性求和,从而也理所当然地引进了各种各样的发散.对于这些发散物理学家们发展出了各种各样的处理方法,但直到今天还仅仅停留在现象解释的层面,背后是否有更加简单的解释方法还是一个谜.

1.3.3 电场能

我们已经十分熟悉,电荷之间的相互作用能实际上就是场的能量.我们现在将给出这一说法的定域化描述.记得一个电荷连续分布体系的相互作用能为:

$$I = \frac{1}{2} \int_V \rho \varphi dV$$

而又根据场与电荷之间的关系:

$$\nabla^2 \varphi = -\frac{\rho}{\epsilon_0}$$

我们把以上积分写为:

$$I = -\frac{\epsilon_0}{2} \int_V \varphi \nabla^2 \varphi dV$$

再根据著名的¹⁷第一格林等式(Green's first identity):

$$\nabla \cdot (\phi \nabla \psi) = \phi \nabla^2 \psi + \nabla \phi \cdot \nabla \psi$$

我们再取 ϕ, ψ 都为 φ , 从而:

$$\varphi \nabla^2 \varphi = \nabla \cdot (\varphi \nabla \varphi) - (\nabla \varphi)^2$$

带入以上积分, 我们得到:

$$\begin{aligned} I &= \frac{\varepsilon_0}{2} \int_V (\nabla \varphi)^2 dV - \frac{\varepsilon_0}{2} \oint_{\partial V} \varphi \nabla \varphi \cdot dA \\ &= \int_V \frac{\varepsilon_0}{2} \mathbf{E}^2 dV + \frac{\varepsilon_0}{2} \oint_{\partial V} \varphi \mathbf{E} \cdot dA \end{aligned}$$

我们总是考虑电荷分布有限的情况. 此时最初的积分体积只要包含所有电荷即可, 那么我们最后得到的这个恒等式就十分的有意思, 它将一个体系的相互作用转化为了一个量的体积分和另一个量在包含所有电荷的面上的面积分的形式. 我们如果取体积为无穷大的全空间, 那么考虑到随着远离电荷体系 R 处 φ 以至多 $\frac{1}{R}$ 的方式衰减(若体系总电荷量为零则衰减更加剧烈), 而 \mathbf{E} 以至多 $\frac{1}{R^2}$ 的方式衰减, 而积分面积以 R^2 的方式增加, 故整个积分以至多 $\frac{1}{R}$ 的方式衰减, 在无穷远处这个积分的第二项就会趋于零.

从而我们得到了:

$$I = \int \frac{\varepsilon_0}{2} \mathbf{E}^2 dV$$

从而发现相互作用能可以用场来定域地计算. 电场的能量密度(energy density) 为:

$$w_e = \frac{\varepsilon_0}{2} \mathbf{E}^2$$

场的能量不再由电荷造成, 而是由场自己的场强去定义, 以后还会进一步证明, 场在一点处对电荷有力的作用, 将会使场的动量密度发生改变, 而对电荷有做功, 将会影响这一点处的能量密度.

1.4 电荷体系

讨论若干电荷体系对于之后问题的理解是大有帮助的. 这些包括:

1.4.1 电偶极子

我们都应该知道一个电荷体系如果带电, 那么尽管在离电荷体系很近处电场分布不均匀, 但离开分布电荷足够远处的场强却总是趋于一个点电荷的电场. 常见的情况发生在原子核中, 原子核由质子和中子组成, 质子带电而中子不带电. 原子核内电荷分布总是不均匀的, 但在考虑外界电子受力时总是将其产生的场视为库仑场. 这是因为电子典型半径: 波尔半径与原子核的尺寸比差了 5 个数量级. 在这个距离处电场完全可以看成是库仑场, 它与库仑场的偏离将作为一个可以忽略的小量而不去考虑.

但类似的, 我们将考虑到很多不带净电荷的电荷分布体系, 它们在近处产生一个不均匀的电场, 而在远处产生的电场

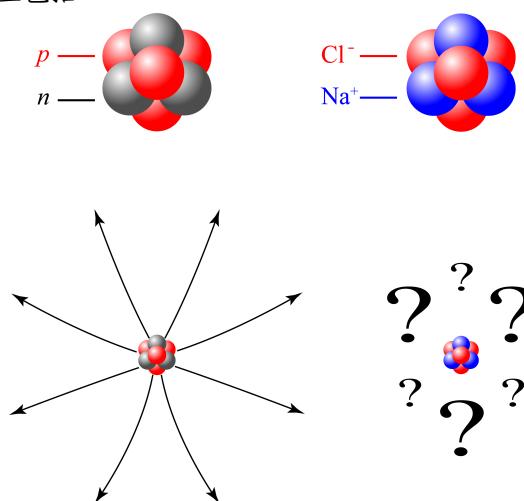


图 1.10: 远方的场

¹⁷而第二格林等式(Green's second identity) 在格林函数法解边值问题时将十分有用:

$$\nabla \cdot (\phi \nabla \psi - \psi \nabla \phi) = \phi \nabla^2 \psi - \psi \nabla^2 \phi$$

却总是存在着某些规律。这个问题有重要的实用价值。比如我们如果从 NaCl 晶体中取出一小块来，它的总电量总是为零，但由于晶体的变形等因素导致正负电荷产生的电场并不能完全抵消，这个电场将会与晶体的动力学耦合，极大地影响晶体的动力学特征。又比如很多分子实际上属于极性分子(polar molecule)，由于原子的电负性(electronegativity)差异，原子间的键合具有极性，在很多情形下一个原子对公用电子对的吸引能力更强，导致电子集中在一个原子一侧。从而负电荷的中心向这个原子位移。只要分子结构上的对称性不会使得键的极性不会相互抵消，整个分子就会显示出极性来，它是溶液中溶剂溶解机理的关键，也会影响物质的熔点、沸点、极化率等等关键的物理特性。

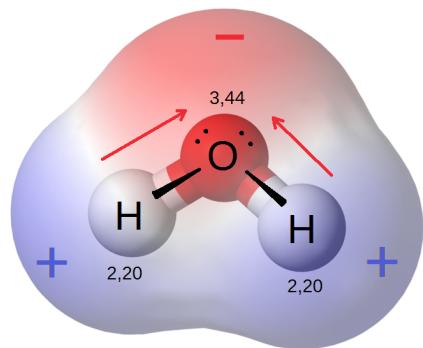


图 1.11: 水分子极性

所以我们关心总电荷量为零的体系在很远处产生的电场，理论上我们可以采取更完整的多极展开(multipole expansion)的做法，在很多情况下电偶极项(dipole term)成为了这种情形下的领头项(leading order)。我们定义小范围分布电荷体系中所有正电荷的“中心”和负电荷的“中心”：

$$\mathbf{r}_+ = \frac{\sum_{q_i > 0} q_i \mathbf{r}_i}{Q}, \quad Q = \sum_{q_i > 0} q_i$$

$$\mathbf{r}_- = \frac{\sum_{q_i < 0} q_i \mathbf{r}_i}{-Q}, \quad -Q = \sum_{q_i < 0} q_i$$

而体系的电偶极矩(dipole moment)就被定义为由负电荷指向正电荷的矢量：

$$\mathbf{d} = Q(\mathbf{r}_+ - \mathbf{r}_-)$$

或者写为：

$$\mathbf{d} = \sum_i q_i \mathbf{r}_i$$

求和遍及所有电荷，无论正负。由于总电荷量是零，不难验证这个矢量与坐标原点的选取无关，为了纪念物理学家德拜(P. Debye)在分子物理领域的相关开创性工作，它一般使用以德拜命名的单位，它是一个高斯制单位：

$$\begin{aligned} 1 \text{ D} &= 10^{-18} \text{ statC} \cdot \text{cm} \\ &= \frac{1}{299792458} \times 10^{-21} \text{ C} \cdot \text{m} \\ &= 3.34 \times 10^{-30} \text{ C} \cdot \text{m} \\ &= 0.208 \text{ e} \cdot \text{\AA} \end{aligned}$$

为何定义这样的一个矢量呢？我们考察在原点附近分布了一个电荷体系 $\{q_i, \mathbf{r}_i\}$ ，那么在很远处产生的电势应该为：

$$\varphi(\mathbf{r}) = \sum_i \frac{1}{4\pi\epsilon_0} \frac{q_i}{|\mathbf{r} - \mathbf{r}_i|}$$

而对于分母中的矢量的模，我们考虑到 $|\mathbf{r}| \gg |\mathbf{r}_i|$ ，可以做如下近似：

$$\begin{aligned} \frac{1}{|\mathbf{r} - \mathbf{r}_i|} &= [(r - r_i)^2]^{-\frac{1}{2}} = (r^2 - 2\mathbf{r}_i \cdot \mathbf{r} + r_i^2)^{-\frac{1}{2}} \\ &\approx (r^2 - 2\mathbf{r}_i \cdot \mathbf{r})^{-\frac{1}{2}} = \frac{1}{r} \left(1 - 2\frac{\mathbf{r}_i \cdot \mathbf{r}}{r^2}\right)^{-\frac{1}{2}} \\ &\approx \frac{1}{r} \left(1 + \frac{\mathbf{r}_i \cdot \mathbf{r}}{r^2}\right) = \frac{1}{r} + \frac{\mathbf{r}}{r^3} \cdot \mathbf{r}_i \\ &= \frac{1}{r} + \frac{\mathbf{e}_r}{r^2} \cdot \mathbf{r}_i \end{aligned}$$

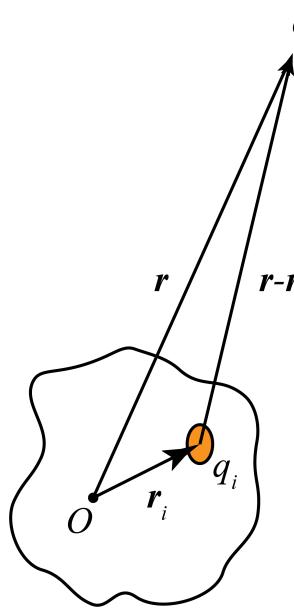


图 1.12: 偶极展开

代入求和表达式, 得:

$$\varphi(\mathbf{r}) \approx \frac{1}{4\pi\epsilon_0} \frac{\sum_i q_i}{r} + \frac{1}{4\pi\epsilon_0} \frac{\mathbf{e}_r \cdot \sum_i q_i \mathbf{r}_i}{r^2}$$

第一项求和由于总电荷量为零放弃了成为领头项的机会, 而第二项凸显了出来, 根据我们之前定义的电偶极矩, 它实际上就是:

$$\varphi(\mathbf{r}) = \frac{1}{4\pi\epsilon_0} \frac{\mathbf{e}_r \cdot \mathbf{d}}{r^2} \quad (r \gg d)$$

我们取了等号, 条件是我们考察的点的矢径 r 远大于电荷分布的平均尺寸 d . 而理论上常常构造一种十分独特的体系, 我们让一对电量为 $+Q$ 与 $-Q$ 严格相等的点电荷相离 $\mathbf{r}_+ - \mathbf{r}_- = \mathbf{l}$, 再令 $Q \rightarrow \infty$ 而 $\mathbf{l} \rightarrow \mathbf{0}$, 而保持 $\mathbf{p} = Q\mathbf{l}$ 的值保持不变, 其极限就是抽象的模型: 点电偶极子(point dipole), 它在除原点外, 全空间产生的电势都可以用上式来表示. 这也是我们以后讨论的中心. 如果在点电偶所在空间建立球极坐标系, 而把电偶极矩矢量指向的方向作为极轴, 我们可以写出电势为:

$$\varphi(r, \theta) = \frac{1}{4\pi\epsilon_0} \frac{p \cos \theta}{r^2}$$

这是一个与半径平方反比的电势, 它与角度有关. 利用球坐标下梯度算符的表示:

$$\nabla = \mathbf{e}_r \frac{\partial}{\partial r} + \frac{1}{r} \frac{\partial}{\partial \theta} + \frac{1}{r \sin \theta} \frac{\partial}{\partial \phi}$$

我们计算出点电偶极子的电场来:

$$E_r(r, \theta) = \frac{1}{4\pi\epsilon_0} \frac{2p \cos \theta}{r^3}$$

$$E_\theta(r, \theta) = \frac{1}{4\pi\epsilon_0} \frac{p \sin \theta}{r^3}$$

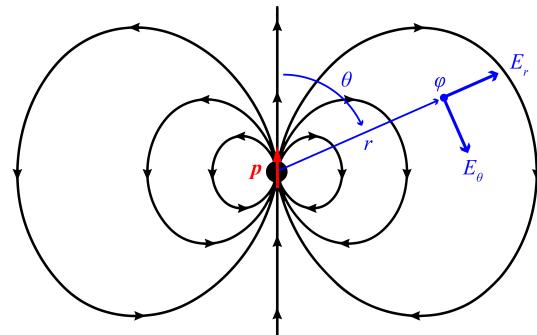


图 1.13: 点电偶极子的场

可见电场倒是三次反比了. 它比点电荷产生的场随距离减少地更快. 理论研究经常会把以上两个分量式组合为一个矢量式, 不难验证¹⁸

$$\mathbf{E} = \frac{1}{4\pi\epsilon_0} \frac{3(\mathbf{e}_r \cdot \mathbf{p})\mathbf{e}_r - \mathbf{p}}{r^3}$$

如果考虑电偶极子在外场中的情形, 采用两电荷分布极限为点电偶极子的结果, 我们很容易发现它的电势能为:

$$V = Q(\varphi_+ - \varphi_-) = Q\mathbf{l} \cdot \nabla\varphi = -\mathbf{p} \cdot \mathbf{E}$$

注意到根据电势能的含义, 这个能量是不包含构成电偶极子的两电荷之间的相互作用能的, 而在电偶极子在电场中发生位移的过程中我们也应该不能让电偶极子发生结构的改变: 即, \mathbf{p} 的大小不能改变, 但 \mathbf{p} 的方向, 位置都可以发生改变, 随着这两个剩余自由度发生变化, 电场对电偶极子做的虚功 $-\delta V = \delta[\mathbf{p} \cdot \mathbf{E}(\mathbf{r})]$ 就被理解为力矩与受力:

$$\delta\mathbf{r} = \mathbf{0}, \delta\mathbf{p} = \delta\theta \times \mathbf{p} : -\delta V = \mathbf{M} \cdot \delta\theta \Rightarrow \mathbf{M} = \mathbf{p} \times \mathbf{E}$$

$$\delta\mathbf{r} \neq \mathbf{0}, \delta\mathbf{p} = \mathbf{0} : -\delta V = \mathbf{F} \cdot \delta\mathbf{r} \Rightarrow \mathbf{F} = \mathbf{p} \cdot \nabla\mathbf{E}$$

其中表达式 $\mathbf{p} \cdot \nabla\mathbf{E}$ 的含义为取沿 \mathbf{p} 方向的矢量场 \mathbf{E} 的方向导数. 以后会提到导体或电介质颗粒将会在外电场下感应或极化出一个电偶极子来:

$$\mathbf{p} = \alpha\mathbf{E}$$

¹⁸这个表达式对原点以外的点适用, 如果考虑原点的奇异电场特性, 并引入狄拉克德尔塔函数(Dirac delta function), 则更准确的表达式为:

$$\mathbf{E} = \frac{1}{4\pi\epsilon_0} \frac{3(\mathbf{e}_r \cdot \mathbf{p})\mathbf{e}_r - \mathbf{p}}{r^3} - \frac{\mathbf{p}}{3\epsilon_0} \delta^3(\mathbf{0})$$

把此式代入受力公式, 我们发现:

$$\mathbf{M} = 0 \quad ; \quad \mathbf{F} = \frac{\alpha}{2} \nabla E^2$$

这也就是我们常说的电场有能够吸引轻小物体的本领的解释, 我们发现这些小物体在外电场中的受力方向总是指向电场强度大小增加的方向的, 而与电场强度方向关系不大.

1.4.2 电荷密度

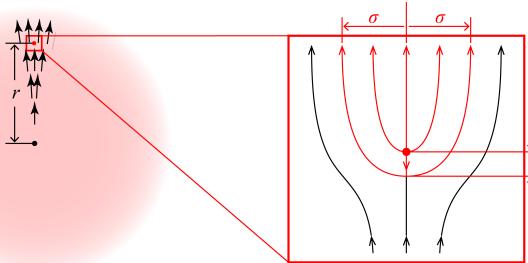


图 1.14: 连续分布点电荷产生平均电场

我们经常会用到连续分布的电荷这样的说法, 把电荷分布写成一个 $\rho(\mathbf{r})$ 的形式, 表示每一个体积元 dV 中含有的电荷量为 ρdV . 这种做法在分析理论问题时发挥了很重要的作用, 但是它却并不实际. 实际电荷分布在很小的尺度看来永远是不连续的, 经典情况下一般可以视为由基本电荷量的点粒子产生的. 那么正如前述, 在这样的情况下电荷密度被理解为:

$$\rho = \sum_i n_i q_i$$

更难理解的此时电场强度的定义. 让我们考虑一个均匀带点的球, 此时球外的电场强度自然可以简单的按 $\frac{Q}{4\pi\varepsilon_0 r^2}$ 来计算. 若内部的电场若按照完全的连续电荷分布来计算, 给出了:

$$\mathbf{E} = \frac{\rho \mathbf{r}}{3\varepsilon_0}$$

的结果. 但这个结果忽略了电荷的离散性, 我们如果认为这个球是由电荷数密度为 n , 电荷量为 q 的大量点电荷构成. 那么设在半径为 r 处的电场与单个点电荷在距中心为 r' 处的电场相等. 我们可以得到:

$$\frac{nqr}{3\varepsilon_0} = \frac{q}{4\pi\varepsilon_0 r'^2}$$

如果一个电荷的平均占据体积视为球体, 半径为 R , 那么电荷数密度与它的关系为:

$$n \cdot \frac{4}{3}\pi R^3 = 1$$

这样就得到:

$$\frac{r'}{R} = \sqrt{\frac{R}{r}}$$

或者用 r 区域内包含的电荷数量 $N = \frac{4}{3}\pi r^3 / \frac{4}{3}\pi R^3$, 得到:

$$r' = \frac{R}{N^{1/6}}$$

还有一种看法是考察一个电荷发出的电场线在远方形成匀强电场以后的截面:

$$\frac{nqr}{3\varepsilon_0} \cdot \pi \sigma^2 = \frac{q}{\varepsilon_0}$$

我们很容易发现:

$$\sigma = 2r$$

也就是说,离开中心区域越远,包含的总电荷数量越多,该处点电荷产生的电场相对与由内部大量电荷积累而来的大的几乎均匀的电场相比就足够地小.点电荷使得外电场偏离均匀的空间范围就远小于电荷与电荷之间的间距,使得电荷之间的空间就可以视为一个几乎均匀的场.

最后我们总结,在用堆放点电荷的方式来得到一个电荷密度时,尽管微观来看产生的电磁场仍然很不均匀,但我们可以抽象出来每一点的平均场强的概念.它的意义有二:一是下一节即将证明的该点局部微元体内电场的平均值.二是如果保证电荷密度不变电荷无限分散的极限情形下该点处的实际场强.它的算法依然是:

$$\mathbf{E} = \int_{V'} \frac{1}{4\pi\epsilon_0} \frac{\rho e_R}{R^2} dV'$$

而既然电场算法与真实的连续分布的电荷没有区别.电势,局部电荷受力体密度就仍然可以这样计算:

$$\varphi = \int \frac{1}{4\pi\epsilon_0} \frac{\rho}{|R|} dV' , \quad \mathbf{f} = \rho \mathbf{E}$$

1.4.3 极化强度 *

往往,平常状态下的物质的真实情况是,所有局部都不是由单一的点电荷单元堆积而成.就比如液体的水,它的重复单元是电中性的水分子.尽管电中性,对外表现出偶极子,在原场情况下近似为点电偶极子,便是其电势电场的领头项.故我们需要考虑的实际情况是:单位体积内有 $n(r)$ 个偶极矩为 $\mathbf{p}(r)$ 的点电偶极子.这里 \mathbf{p} 不会是常矢量,首先其方向往往随着空间位置逐渐在改变,而每个电偶极子的电偶极矩大小亦可能因为极化程度的不同而有所区别,下一章我们要介绍的位移极化就是这样的情形.我们将数密度 n 与电偶极矩 \mathbf{p} 相乘,得到单位体积内的电偶极矩,即**极化强度**(polarization density):

$$\mathbf{P} = np$$

对于连续情况,我们发现电势应该用电偶极子的电势进行体积分来得到:

$$\varphi = \int \frac{1}{4\pi\epsilon_0} \frac{\mathbf{P} \cdot \mathbf{e}_R}{R^2} dV' , \quad \mathbf{R} = \mathbf{r} - \mathbf{r}'$$

然而我们发现以下等式是成立的:

$$\begin{aligned} \nabla' &= e_x \frac{\partial}{\partial x'} + e_y \frac{\partial}{\partial y'} + e_z \frac{\partial}{\partial z'} \\ \nabla' \cdot \frac{\mathbf{P}}{|R|} &= \frac{\nabla' \cdot \mathbf{P}}{|R|} + \mathbf{P} \cdot \nabla' \frac{1}{|R|} \\ \nabla \frac{1}{|R|} &= \nabla \frac{1}{|\mathbf{r} - \mathbf{r}'|} = -\frac{\mathbf{r} - \mathbf{r}'}{|\mathbf{r} - \mathbf{r}'|^3} = -\frac{\mathbf{e}_R}{R^2} , \quad \nabla' \frac{1}{|R|} = \nabla' \frac{1}{|\mathbf{r} - \mathbf{r}'|} = -\frac{\mathbf{r}' - \mathbf{r}}{|\mathbf{r}' - \mathbf{r}|^3} = \frac{\mathbf{e}_R}{R^2} \\ \Rightarrow \quad \frac{\mathbf{P} \cdot \mathbf{e}_R}{R^2} &= \nabla' \cdot \frac{\mathbf{P}}{|R|} - \frac{\nabla' \cdot \mathbf{P}}{|R|} \end{aligned}$$

这样我们就得到在区域 V' 内存在极化矢量 \mathbf{P} 时在全空间产生的电势的公式:

$$\begin{aligned} \varphi &= \int_{V'} \frac{1}{4\pi\epsilon_0} \left(\nabla' \cdot \frac{\mathbf{P}}{|R|} - \frac{\nabla' \cdot \mathbf{P}}{|R|} \right) dV' \\ &= \oint_{\partial V'} \frac{1}{4\pi\epsilon_0} \frac{\mathbf{P} \cdot dA'}{|R|} + \int_{V'} \frac{1}{4\pi\epsilon_0} \frac{-\nabla' \cdot \mathbf{P} dV'}{|R|} \\ &= \oint_{\partial V'} \frac{1}{4\pi\epsilon_0} \frac{\sigma_P dA'}{|R|} + \int_{V'} \frac{1}{4\pi\epsilon_0} \frac{\rho_P dV'}{|R|} \end{aligned}$$

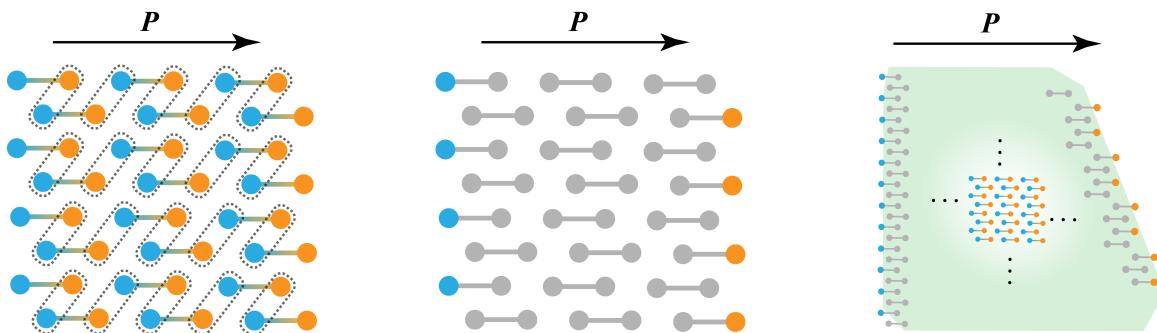


图 1.15: 极化强度导致表面电荷分布

从以上表达式看得出来, 最后电势有一种等效的求法, 就是认为在体内有一个 $\rho_P = -\nabla' \cdot \mathbf{P}$ 的电荷密度分布, 而认为在体表面的面元 dA 上带有面电荷 $dQ = \mathbf{P} \cdot d\mathbf{A}$, 或者认为有面电荷密度 $\sigma_P = \mathbf{P} \cdot \mathbf{n}$, 其中 \mathbf{n} 为体积表面的外法向量. 与其说这是一种等效, 不如说这只是对实际情况的另一种理解. 如上图所示, 这是一块均匀极化的介质, 我们这样理解介质表面上会出现的电荷: 对介质内部的电荷进行合适的分组使得一对一对的正负电荷产生的场互相抵消. 这样便只剩下了表面的电荷能产生场了. 在几何学上不难证明, 如果电偶极子长 l , 两电荷为 q , 体密度为 n , 那么在与极化矢量平行的方向看两端未配对电荷的面密度恰好为 $\sigma = nql = np = P$. 再考虑面可能的倾斜便容易发现以上极化电荷面密度 σ_P 的公式. 类似地, 体电荷密度 ρ_P 对应于沿某方向的 \mathbf{P} 分量在变化的情形. 具体的物理图像也请读者加以分析与思考.

我们有把握预期空间各处的电场强度的公式, 也可以通过上述方式等效, 毕竟我们可以通过电势计算电场:

$$\mathbf{E} = \oint_{\partial V'} \frac{1}{4\pi\epsilon_0} \frac{\sigma_P dA'}{|\mathbf{R}|^2} \mathbf{e}_R + \int_{V'} \frac{1}{4\pi\epsilon_0} \frac{\rho_P dV'}{|\mathbf{R}|^2} \mathbf{e}_R$$

然而, 若要开始考虑电偶极子在场中受到的外力, 那就势必要考虑连续分布的电荷体系产生的电场和微观非均匀分布的电荷体系的局部场之间的区别. 我们会惊人地发现以上电场只能作为平均意义上的**宏观场**(macroscopic field) 而区别于电偶极子具体感受到的**局域场**(local field). 首先让我们来计算一个概念, 一个球体内的平均场强 $\bar{\mathbf{E}}$. 它会计算球体内电荷的受力提供关键线索. 以原点为中心, 半径为 R 的球内积分直接得到:

$$\bar{\mathbf{E}} = \int_{\text{Ball}} \mathbf{E} dV / \frac{4}{3}\pi R^3$$

直接计算会具有相当的难度. 这里采用的技巧是在分子积分中引入 ρ' 常数, 把积分看做原来的电荷体系产生的电场与均匀分布在这个球体内部的电荷体系之间的相互作用力¹⁹. 这样的球体产生的电场强度为:

$$\mathbf{E}' = \begin{cases} \frac{\rho' r}{3\epsilon_0} \mathbf{e}_r & (r < R) \\ \frac{\rho' R^3}{3\epsilon_0 r^2} \mathbf{e}_r & (r > R) \end{cases}$$

从而上式被等价写作:

$$\bar{\mathbf{E}} = \int_{\text{Ball}} \rho' \mathbf{E} dV / \frac{4}{3}\pi\rho' R^3 = \int_{\text{Ball}} \mathbf{E} \nabla \cdot \mathbf{E}' dV / \frac{4}{3}\pi\rho' R^3 = \int_{\text{all}} \mathbf{E} \nabla \cdot \mathbf{E}' dV / \frac{4}{3}\pi\rho' R^3$$

最后一步 all 改为对全空间积分. 由牛顿第三定律看出, 上式通过某种化简²⁰后可以得到:

$$\int_{\text{all}} \mathbf{E} \nabla \cdot \mathbf{E}' dV + \int_{\text{all}} \mathbf{E}' \nabla \cdot \mathbf{E} dV = 0$$

¹⁹ 数学上相当于把 \mathbf{r} 视作 r 的散度的 $1/3$, 即:

$$\nabla \cdot \mathbf{r} = \frac{\partial x}{\partial x} + \frac{\partial y}{\partial y} + \frac{\partial z}{\partial z} = 3$$

²⁰ 而数学上的证明需要用到张量分析, 过程略去.

于是把原电场对应的电荷密度记做 ρ , 待求的量变成:

$$\bar{\mathbf{E}} = - \int_{\text{all}} \rho \mathbf{E}' dV / \frac{4}{3} \pi \rho' R^3 = \frac{1}{4\pi\epsilon_0} \frac{- \int_{\text{Ball}} \rho r dV}{R^3} + \int_{\text{all-Ball}} \frac{1}{4\pi\epsilon_0} \frac{\rho(-\mathbf{e}_r)}{r^2} dV$$

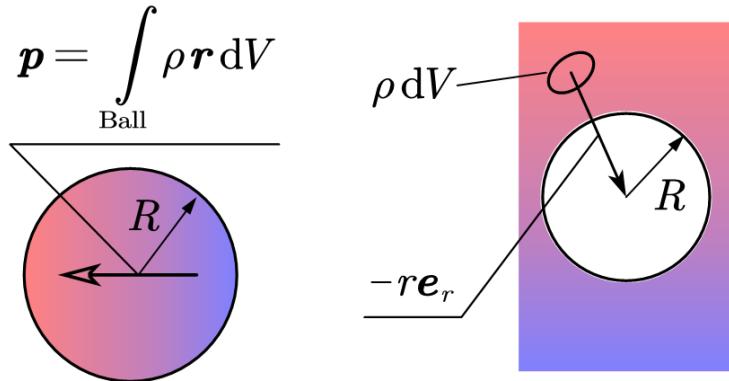


图 1.16: 球内平均电场的两项积分

通过以上结果我们发现, 在一个电荷分布体系中的一个球体内的平均电场取决于两项积分. 第一项不是别的, 就是直接取决于球内电荷分布产生的电偶极矩 \mathbf{p} . 第二项则代表外界电荷直接对球心处产生的电场²¹.

接下来可以进行以下推演: 如果外界电荷不是那么紧密地靠近球面, 那么我们可以通过球面把电荷体系分为两部分. 现在我们来关心内部这些电荷的受力. 依然从电场考虑, 由于外部电荷离球面较远, 故可以近似认为在球内产生一个均匀的电场 \mathbf{E}_{ex} , 甚至于它就可以近似地等于对球心处的场强, 这样我们就把之前的式子写为:

$$\bar{\mathbf{E}} = \frac{1}{4\pi\epsilon_0} \frac{-\mathbf{p}}{R^3} + \mathbf{E}_{\text{ex}} \Rightarrow \mathbf{E}_{\text{ex}} = \bar{\mathbf{E}} + \frac{1}{4\pi\epsilon_0} \frac{\mathbf{p}}{R^3}$$

如果让求平均的体积非常小, 而电荷又总是连续化, 那么 $\bar{\mathbf{E}}$ 又应当与之前通过极化电荷分布计算的体内点场强 \mathbf{E} 无异. 我们已经观察到一个十分关键的信息: 考虑球内电荷受力时应当使用的外界的场 \mathbf{E}_{ex} , 它与忽略电荷不连续性而计算出来的“平均场” $\bar{\mathbf{E}}$ 存在一个差. 最后考虑到极化强度的定义 $\mathbf{P} = \lim_{R \rightarrow 0} \mathbf{p} / \frac{4}{3}\pi R^3$, 最终我们得到:

$$\mathbf{E}_{\text{ex}} = \bar{\mathbf{E}} + \frac{\mathbf{P}}{3\epsilon_0}$$

这个细节对理解介质的极化十分重要. 我们将会在下一章重新回到这个问题.

1.4.4 若干对称带电体系的电场

最后还有一些常见的带电体系的电场是值得作为现成结论去记忆的. 它们易于计算的原因在于或多或少具有某些对称性. 在此逐一加以回忆与讨论.

1. 薄球壳: 以球心为原点, 三维旋转对称性. 过球心的镜面对称性.

$$\varphi = \begin{cases} \frac{1}{4\pi\epsilon_0} \frac{Q}{R} & (r < R) \\ \frac{1}{4\pi\epsilon_0} \frac{Q}{r} & (r > R) \end{cases}, \quad \mathbf{E} = \begin{cases} \mathbf{0} & (r < R) \\ \frac{1}{4\pi\epsilon_0} \frac{Q}{r^2} \mathbf{e}_r & (r > R) \end{cases}$$

我们遇到了一个问题: 球面上的电势和电场强度是如何的? 我们这时候在定义上会遇到发散困难(divergency difficulty): 论电势, 问题不大, 从球内球外的电势数据(下面的势曲线1.17)来看, 球面上的电势连续, 但是其

²¹这意味着如果球内没有电荷, 那么对球内电场做平均的结果恰好得到球心处电场. 这种性质称作调和(harmonic).

变化率不连续(不光滑). 所以很自然地球面上电势定义为双边的极限值即球内电势 $\frac{1}{4\pi\varepsilon_0}\frac{Q}{R}$. 有限面电荷情况的电势都是不存在发散困难的. 因为在面上局部划出来一个近似均匀带电 σ , 半径为 a 小圆盘, 则盘心电势:

$$\varphi = \int_0^a \frac{1}{4\pi\varepsilon_0} \frac{\sigma \cdot 2\pi r dr}{r} = \frac{\sigma a}{2\varepsilon_0}$$

是一个有限值. 而圆盘以外的电荷对圆心的电势显然也是有限的. 故无发散困难. 反观电场强度, 这情况就很不一样了. 首当其冲地从以上公式上看, 球内球外的电场显然不连续而在球面上具有不同的极限. 其更深层次的原因在于: 面电荷的电场强度积分本质上是发散的, 有奇异性(sigularity)的. 考虑一个均匀带电 σ , 圆心角为可近似的小角 $\Delta\alpha$, 半径 a 的扇形. 圆心处的场强:

$$E = \int_0^a \frac{1}{4\pi\varepsilon_0} \frac{\sigma \cdot \Delta\alpha r dr}{r^2} = \frac{\sigma\Delta\alpha}{4\pi\varepsilon_0} \ln \frac{a}{0} = \infty$$

规避掉发散²²的思想有两种. 一是通过主值积分(principal value integral)的方式去做截断(cutoff), 在需要确定场强的点处挖去一个半径为 a 的球, 计算其他部分在该点处的场强并最后让 $a \rightarrow 0$; 二是通过平均来理解发散的场强, 如之前所说的, 点电荷模型与连续分布的电荷模型是统一的, 后者计算出来的场强就是前一种不均匀场强的球体平均. 故只需要在面电荷情况也做平均就能规避发散点的计算而利用球内其他可以计算的点来计算场强. 容易验证, 两种方法最后的结果都是法向电场为内外场强的平均数, 球面上的场强分量由于对称性为零:

$$\mathbf{E} = \frac{Q}{8\pi\varepsilon_0 R^2} \mathbf{e}_r , \quad (r = R)$$

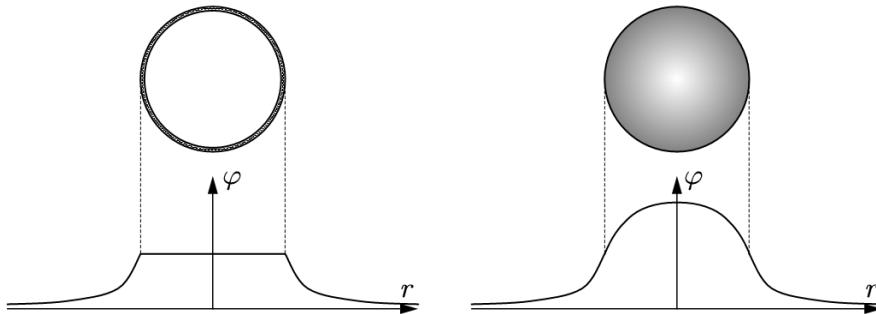


图 1.17: 薄球壳与球体的势曲线

2. 球体: 以球心为原点, 三维旋转对称性. 过球心的镜面对称性.

$$\varphi = \begin{cases} -\frac{\rho}{6\varepsilon_0} r^2 + C & (r < R) \\ \frac{1}{4\pi\varepsilon_0} \frac{Q}{r} & (r > R) \end{cases}, \quad \mathbf{E} = \begin{cases} \frac{\rho}{3\varepsilon_0} r \mathbf{e}_r & (r < R) \\ \frac{1}{4\pi\varepsilon_0} \frac{Q}{r^2} \mathbf{e}_r & (r > R) \end{cases}$$

均匀带电球体没有发散困难. 但注意在无穷远电势为零的规范下为了使得以上电势连续, 附加常数 C 应当为:

$$C = \frac{3Q}{8\pi\varepsilon_0 R}$$

3. 无限大带电平面: 保持平面位置不变的平移, 旋转对称性. 带电平面镜像对称. 竖直面镜面对称.

$$\varphi = -\frac{\sigma}{2\varepsilon_0} |z| , \quad \mathbf{E} = \frac{\sigma}{2\varepsilon_0} \text{sgn}(z) \mathbf{e}_z$$

其中 sgn 是符号函数(sign function), 即当自变量 z 取正时为 $+1$, 为负时取 -1 . $\text{sgn}(z) = z/|z|$. 从计算上看, 面上电场依然有奇异性. 面外则通过积分的方式定义的电场奇异性消失. 但是电势即使在面外此时也

²²称作正规化(regularization).

有奇异性了,不能通过电荷积分得到电势的值,只能通过与电场的关系,所以此时无法把无穷远作为势能的零点,这里的无穷远也自动被分离为下半空间无穷远和上半空间无穷远,所有涉及到无穷大带电平面的问题这样两个无穷远都必须要被区别对待.

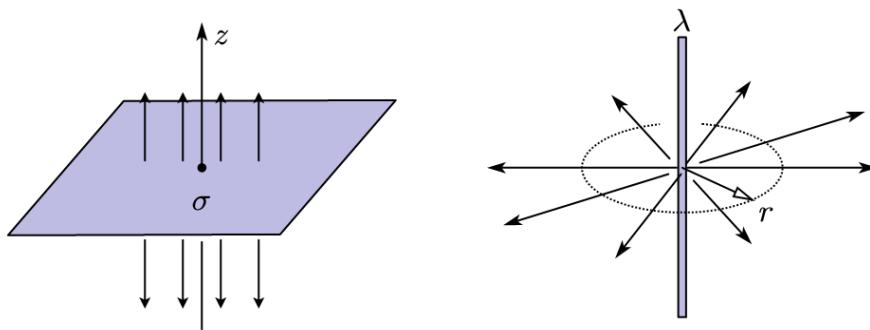


图 1.18: 无限大平面与无限长直线

4. 无限长带电直线:保持直线位置不变的平移,旋转对称性.水平面镜面对称.过直线的竖直面镜面对称.

$$\varphi = -\frac{\lambda}{2\pi\epsilon_0} \ln r \quad , \quad \mathbf{E} = \frac{\lambda}{2\pi\epsilon_0 r} \mathbf{e}_r$$

请尤其注意电势前的负号.同样地,电场在导线外是没有奇异性的,但是电势有奇异性,电势是通过与电场关系定义的,不能把无穷远电势视作零.与上一种电荷分布之所以具有不能把无穷远视作势能零点的共同点,是因为电荷分布的范围不是有限.试想在无穷远外都有电荷了怎么可能不影响无穷远的电势分布?

如果是一个半径为 R 的,带电面密度为 σ 的圆柱薄壳,那么内部也将没有电场,等势.这一点从高斯定律很容易能看出来.以及一个均匀带电的圆柱体的电势电场分布都留给读者计算.

5. 圆环:旋转对称性.过环水平面镜面对称.过轴线的竖直面镜面对称.

这样一个体系只有轴线上的电势和电场才有比较简单的表达式:

$$\varphi = \frac{\lambda}{2\epsilon_0} \frac{1}{\sqrt{R^2 + z^2}} \quad , \quad \mathbf{E} = \frac{\lambda}{2\epsilon_0} \frac{z}{(R^2 + z^2)^{\frac{3}{2}}} \mathbf{e}_z$$

但是实际应用场合中,环心附近的场也是非常需要的.其计算方法其实就是之前介绍的多极展开,不过现在需要保留到二阶项才能得到领头项:

$$\begin{aligned} \varphi &= \int \frac{1}{4\pi\epsilon_0} \frac{dQ}{|\mathbf{r} - \mathbf{R}|} \\ &= \int_0^{2\pi} \frac{1}{4\pi\epsilon_0} \frac{\lambda R d\theta}{\sqrt{(R \cos \theta - \rho)^2 + R^2 \sin^2 \theta + z^2}} \end{aligned}$$

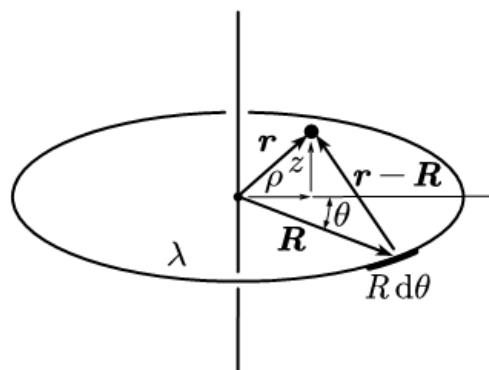


图 1.19: 圆环的中心附近电场

$$\begin{aligned}
&= \int_0^{2\pi} \frac{1}{4\pi\varepsilon_0} \frac{\lambda d\theta}{\sqrt{1 + \frac{z^2 + \rho^2 - 2R\rho \cos\theta}{R^2}}} \\
&= \frac{\lambda}{4\pi\varepsilon_0} \int_0^{2\pi} \left(1 + \frac{z^2 + \rho^2 - 2R\rho \cos\theta}{R^2}\right)^{-\frac{1}{2}} d\theta \\
&\approx \frac{\lambda}{4\pi\varepsilon_0} \int_0^{2\pi} \left[1 - \frac{1}{2} \frac{z^2 + \rho^2 - 2R\rho \cos\theta}{R^2} + \frac{3}{8} \left(\frac{z^2 + \rho^2 - 2R\rho \cos\theta}{R^2}\right)^2\right] d\theta \\
&\approx \frac{\lambda}{4\pi\varepsilon_0} \int_0^{2\pi} \left(1 + \frac{\rho \cos\theta}{R} - \frac{z^2 + \rho^2}{2R^2} + \frac{3\rho^2 \cos^2\theta}{2R^2}\right) d\theta \\
&= \frac{\lambda}{2\varepsilon_0} \left(1 - \frac{z^2}{2R^2} + \frac{\rho^2}{4R^2}\right)
\end{aligned}$$

于是原点附近的场强就是这个势的梯度:

$$E_\rho = -\frac{\partial\varphi}{\partial\rho} = -\frac{\lambda}{4\varepsilon_0 R^2} \rho, \quad E_z = -\frac{\partial\varphi}{\partial z} = -\frac{\lambda}{2\varepsilon_0 R^2} z$$

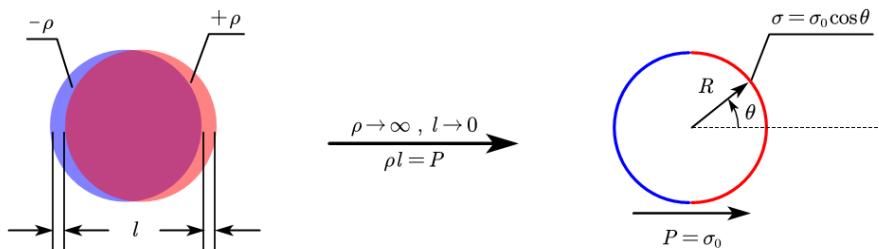


图 1.20: 均匀极化的球

6. 均匀极化球体. 旋转对称性, 过轴的平面的镜面对称性. 垂直轴的平面的镜面反对称性²³.

设球内极化强度为 P . 由之前的极化强度与电荷密度的关系知道, 这相当于一个表面带电 $\sigma = \sigma_0 \cos\theta$ 的球壳. 但是为了计算其内外的场强, 我们需要把极化强度还原为单位体积内携带的偶极子的总和. 每个偶极子可视为一对点电荷. 从而形成两个电荷密度分别为 $+\rho$ 和 $-\rho$ 的球体, 球心沿极化方向岔开 l . 只需计算极限情况:

$$\rho \rightarrow \infty, \quad l \rightarrow 0, \quad \rho l = P$$

定义球内的总电偶极矩 $\mathbf{p} = \mathbf{P} \cdot \frac{4}{3}\pi R^3$ 由球对内对外的场强公式计算得, 球内外的场强为:

$$\mathbf{E} = \begin{cases} -\frac{\mathbf{P}}{3\varepsilon_0} & (r < R) \\ \frac{1}{4\pi\varepsilon_0} \frac{2p \cos\theta}{r^3} \mathbf{e}_r + \frac{1}{4\pi\varepsilon_0} \frac{p \sin\theta}{r^3} \mathbf{e}_\theta & (r > R) \end{cases}$$

图 1.21: 均匀极化场

无奇异性, 且也可以得到电势公式:

$$\varphi = \begin{cases} \frac{P}{3\varepsilon_0} r \cos\theta & (r < R) \\ \frac{1}{4\pi\varepsilon_0} \frac{p \cos\theta}{r^2} & (r > R) \end{cases}$$

²³即需要镜面反射后进行正负电荷互换才能与变换前一致.

第二章 导体与介质

2.1 导体与静电平衡

2.1.1 绝缘体与导体

微观地看,物质由原子或分子等组成,其中导电现象通常发生在不同情况下:

1. 真空导电:一般不会说真空具有导电性(conductive),因为真空中是没有载流子(charge carrier)的¹.的确,在阴极射线管(cathode ray tube)中,加热一个阴极,并辅以合适的偏置电压可以造成真空中的电流.或是用一束能量足够的光子去轰击金属表面造成电子逸出,甚至纯粹由于阴极表面尖处十分强的电场导致电子直接克服逸出功发射出来.三种现象分别称为热发射(thermionic emission),光电效应(photoelectric effect)与场发射(field emission).这种电荷的定向移动现象被统一地称为运输现象(transport phenomenon).由于真空运输的独特性质,比如电子不会受到散射,平均自由程远大于仪器尺寸,与凝聚态物理中的一些概念对应,这被称为弹道运输(ballistic transport).

2. 绝缘体漏电:大多数非金属晶体,或是不含离子的液体与气体,原子核在晶体中都限制在点阵格子的特定位置,在气体,液体中则是可以活动的原子分子周围.电子分为两类,,一类是原子的内层电子,它极为稳定地存在于原子周围,离子晶体中的几乎完全被阴离子夺取的电子也属于这种情况.它们与原子核一起构成原子实(atomic core).而价层电子(valence shell electron),它们用来成键,将原子连接形成分子或者晶体,一般被定域在原子间的特定区域.所有这些电子的特点都是束缚态(bond state).它们不能在介质中自由传导,一个电子的区域到另一个电子的区域间存在势垒(potential barrier),经典物理认为电子的动能不足以穿过这些势垒.但是,量子理论则认为电子的波函数可以通过隧道效应(tunnelling)以小概率在不同区域间转移.这就为有电场的情况下电荷的平均定向移动提供了可能.这种现象称为漏电(leakage).

3. 绝缘体击穿:在十分高的电场下,原子与原子间,或是分子中显不同电性的部分之间的典型电压将大于阻碍电子转移的势垒.在晶体被击穿时此时电子在本质上可以被视为在周期性势能场与线性势能场的叠加中运动的经典粒子.不难看出电子不仅可以脱离单个势阱的束缚,还可以在这样的电场力下做持续的加速运动.气体液体则是在电场力作用下原子,分子被撕裂为离子而继续在电场力下做加速运动.此时碰撞就发生了:碰撞的结果往往是使得更多的原子分子发生电离.这样就会雪崩式的累积,固液气中的这些现象都称作雪崩击穿(avalanche breakdown).尤其是气体中的放电现象在不同电流条件下分别造成汤森放电(Townsend discharge),电晕放电(corona discharge),辉光放电(glow discharge),电弧放电(arc discharge)的不同情况.

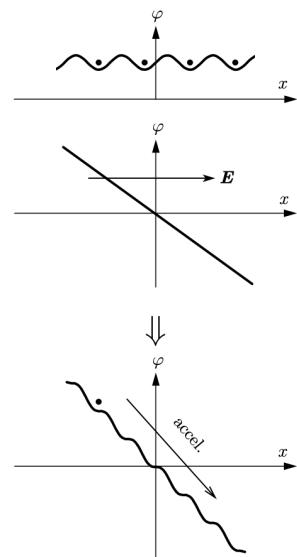


图 2.1: 晶体击穿

¹量子场论认为真空存在电子对的产生湮灭涨落从而具有一定的导电性.

4. 自然的导体: 我们将材料区分为绝缘体(insulator) 与导体(conductor), 最重要的依据就是在弱场下是否天然地具有能产生电荷输运的载流子. 所以导体就是一类事先就具有一定数密度 n 的一种或多种载流子 q 的材料, 它们往往在有电场 E 在场时发生位移, 且与散射造成的形式阻力平衡, 造成平均的匀速运动, 即定向漂移(directional drift). 常见的金属以最外层的自由电子为载流子, 半导体以热激发或掺杂形成的电子或空穴为载流子, 电解质溶液或部分气态或液态的等离子体以阴阳离子为载流子. 它们都是典型的导体. 区别于前三种情况与最后的超导体.

5. 超导体: 低温下的很多奇异的宏观现象大大拓宽了人们对量子理论在物理学各个层次中应用的认识. 其中很典型的一个就是超导(superconductivity) 现象. 最早在 1911 年由昂内斯(H. K. Onnes) 在研究低温下汞的电阻率时惊人的发现在 4.2K 下电阻率消失至零, 这就是超导的发现. 但是直到约十年后才引起足够重视. 在 1933 年迈斯纳(F. W. Meissner) 才发现超导不仅仅是电阻率的消失, 还伴随着完全抗磁性(perfect diamagnetism): 就好像理想导体内部无电场线, 在超导体中交变的磁场也只能存在于表面极小的深度内(趋肤深度), 也称迈斯纳效应(Meissner effect). 后来通过量子理论的蓬勃发展, 在 1957 年提出的BCS 理论(Bardeen-Cooper-Schrieffer theory) 终于解释了超导的成因, 原来是一对库珀电子对(Cooper pair) 通过与声子(晶格振动)相互作用而造成了没有任何散射的量子模式. 完整地解释了直到 1986 年所有低温超导(low-temperature superconductivity) 的现象. 但是随着 35K 超导的镧钡铜氧化物的发现, 宣告人类进入高温超导(high-temperature superconductivity) 纪元. 新的超导材料不断被发现, 新的理论也不断被提出, 目前最“高温”的材料是汞钡钙铜氧化物, 超导临界温度 135K. 为什么有高温超导? 高温超导的临界温度如何进一步提高? 直到今天这也还是方兴未艾的研究领域. 其在科研, 科技, 生产, 生活方面的应用是不可估量的.

我们将把研究的范围主要放在导电的导体和不导电的绝缘体上. 前者分析其静电平衡, 即本章要集中讲的情形; 下一章的稳恒电流和之后章节的拟稳的一些情况则是导体研究的其他场合. 对于绝缘体我们本章也介绍其介电特性, 光作为电磁波在介质中的传播则本章做一个抛砖引玉, 主要内容在光学色散理论中讲解.

2.1.2 导体的特点

这里说的导体, 主要指固体形式的金属.

首先界定我们研究的问题范围: 静电平衡(electrostatic equilibrium) 问题是一种稳态(steady state). 我们研究的问题由各式各样的, 有限的块状导体 $C_1, C_2 \dots C_n$ 构成, 这些导体带电情况由体电荷密度 ρ 和面电荷密度 σ 描述. 但是导体 C_i 上的总电荷量不一定为零:

$$Q_i = \int_{C_i} \rho dV + \int_{\partial C_i} \sigma dA \neq 0$$

这个电量 Q_i 一般也是可以在一定范围内人为控制的, 即给导体充电(charge). 利用电容的性质很容易可以做到这一点. 在为导体充电后由于电量守恒, 其总电量就不能改变了, 但是其电荷分布一般是未知的.

除了它们往往还有人为指定的各种点电荷 q_j 或者电荷分布 $\rho(r)$ 存在于导体外. 那么, 容易想像, 当导体外突然产生电荷分布时(如原来中和的电荷被重新分开), 或者导体被置入电荷体系中, 本质上就是当导体外的场环境发生改变时, 导体内部显然不可能保持电场强度始终为零. 只要有电场就会引起电流, 它必然使得内部和表面的电荷分布发生改变 $\rho(t), \dot{\sigma}(t) \neq 0$. 这些状态就是暂态(transient state). 但是只要外界的场不再变动, 经过特征的弛豫时间(relaxation time) 后, 一般很短只有 $10^{-14}s$ 量级, 就会接近, 直到变成稳态.

通过以上推理, 足以看出, 静电平衡下, 导体内部应当没有电场. 这就是静电平衡理论的初始命题. 它的初步推演演化出以下的九宫格式的结论, 分别讨论导体内, 导体表面, 导体外三处的电荷分布, 电场分布, 和电势分布三个特性:

一般特征	电荷	电场	电势
导体内	$\rho = 0$	$E = 0$	$\varphi = C$
导体表面	$\sigma \neq 0$	奇异	$\varphi = C$
导体外	$\rho \neq 0$	$E \neq 0$	$\varphi \neq C$

表 2.1: 静电平衡的初级结论

其实就是说,因为导体内部没有电场,所以导体内部也没有形成电场的电荷(高斯定理可证),也使得导体成为一个等势体(equipotential body),导体表面成为一个等势面(equipotential surface).注意,这里说的“内部”,“表面”,“外部”包含两种情况.导体内部就是金属成分的空间,外部一般是真空与外加电荷分布的空间,表面是它们的分界面,但是下图左的情况,导体的外部延伸到了无穷远,这种为开外界(open exterior),但是中和右两种情况,导体的外部是有限的闭外界(closed exterior).第二种典型情况这个单连通的闭外界还被称作空腔(cavity).除非特殊说明,我们一般默认导体的内部不会延伸到无穷远(总是闭的).

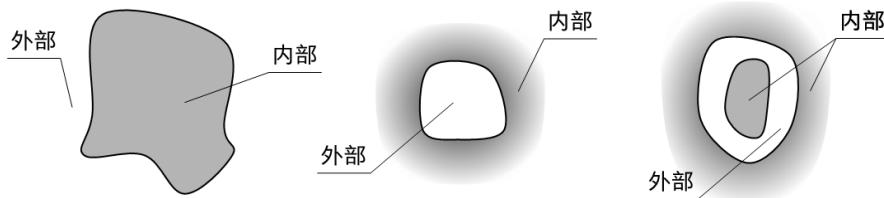


图 2.2: 导体内外

在导体表面发生了电场计算的奇异性,因为这里是面电荷分布.我们考察细节就会进一步发现以下进一步的结论: **导体表面的面电荷只朝外侧发出(或吸收)电场线,其方向垂直于表面(电场线垂直于等势面),且疏密程度正比于电荷密度(即高斯定理)**.即:

$$\mathbf{E} = -\nabla\varphi = \frac{\sigma}{\epsilon_0}\mathbf{n}$$

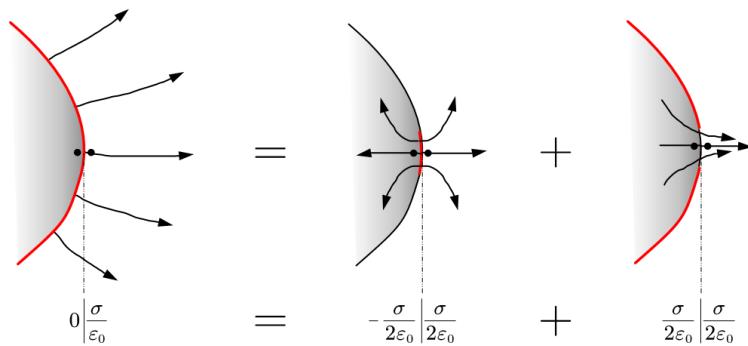


图 2.3: 导体表面

但是正如上图所示,如果在十分接近表面的内外各取一点,就会发现如果只选取相对这两点足够大的一块面电荷(但又足够小以至于可以视作平面),它对面左右的电场强度只贡献以上值的一半.那么另一半必然由去掉这一块电荷的另外的部分所贡献.所以如果要讨论面上的场强,通过上一节的球体平均的理解方

式, 它就自然地被定义为内外场强和的一半, 即:

$$\mathbf{E} = \frac{\sigma}{2\epsilon_0} \mathbf{n}$$

对于导体静电平衡问题的计算求解, 也包括模拟求解, 本质上都建立在以下基本原理的基础上:

独立叠加原理: 静电平衡问题可以拆解为独立的在导体外部一个个连通区域内的静电边值问题.

所谓的**边值问题**(boundary value problem) 本是偏微分方程的术语, 在这里特指下面说的静电情况. 每一种情况中, 我们无需分析导体的内部的电荷, 电场, 电势. 只需要求解外部的电场 $\mathbf{E}(\mathbf{r})$ 和电势 $\varphi(\mathbf{r})$. 导体外部的电荷分布 $\rho(\mathbf{r})$ 却是事先给定的. 还给定了每一种导体表面的, 要么是电势 V_i (从而导体内部的电势也是这个值), 把这种导体称作第一类导体, 对应条件称为第一类边界条件; 要么是给定了这个面上的总电量 Q_i , 把这种导体称作第二类导体, 对应条件称为第二类边界条件. 而每一个导体外部的区域的边界, 就回到了各个导体的表面 $\Sigma_1, \Sigma_2 \dots \Sigma_n, \Pi_1, \Pi_2 \dots \Pi_m$ 构成的集合. 但是表面上的电荷分布也是事先不知道的. 由于电荷, 电场都可以从电势中得到, 故只暂时考虑求电势, 我们把静电边值问题整理一下:

已知: 导体外连通区域 D 以第一类导体表面 $\Sigma_1, \Sigma_2 \dots \Sigma_n$ 和第二类导体表面 $\Pi_1, \Pi_2 \dots \Pi_m$ 为边界. 已知区域 D 内电荷分布 ρ , 已知第一类导体表面 Σ_i 上的电势 V_i . 已知第二类导体表面 Π_j 上的电量 Q_j :

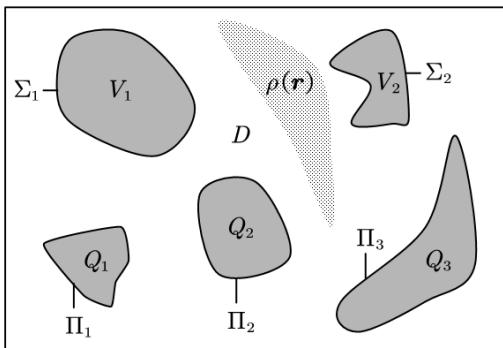
$$\partial D = \left(\bigcup_i \Sigma_i \right) \cup \left(\bigcup_j \Pi_j \right)$$

$$\rho = \rho(\mathbf{r}) \quad , \quad \mathbf{r} \in D$$

$$\varphi(\mathbf{r})|_{\Sigma_i} = V_i \quad , \quad \int_{\Pi_j} \nabla \varphi \cdot d\mathbf{A} = Q_j / \epsilon_0 \quad , \quad \varphi(\mathbf{r})|_{\Pi_j} = C$$

未知: 导体表面电荷的具体分布 $\sigma(\mathbf{r})|_{\Sigma_i, \Pi_j}$

待求: D 内的电势分布 $\varphi(\mathbf{r})|_D$



但是还要注意这种边值问题又分为**外问题**(outside problem) 和**内问题**(inside problem) 两类. 外问题是区域 D 延伸到了无穷远, 这种情况的边值问题提法就完全如上所述. 但是内问题中区域 D 存在一个最外的表面 Σ_0 , 它把整个 D 和其他所有金属表面与内部包含在其内部. Σ_0 的外部亦是金属. 对于内问题, 我们很容易发现 Σ_0 上的总带电量为内部所有带电量的和的相反数, 且仅仅由内部体系和它产生的电场线不会延伸到 Σ_0 外部. 所以我们暂时可以以这一类问题中 Σ_0 上的电势为零点:

$$\varphi(\mathbf{r})|_{\Sigma_0} = 0$$

这样所有已知未知的电势皆是相对 Σ_0 而言的电势差.

对于每一个具体的边值问题, 指导我们求解的重要思想是**广义叠加原理**(general superposition principle) 和**唯一性定理**(uniqueness theorem). 从语义上看广义叠加原理显然与之前的单纯的叠加原理有着直接的联系. 其实它更是来自于数学上具有线性结构的系统的解的叠加原理. 这个叠加原理是这样表述的:

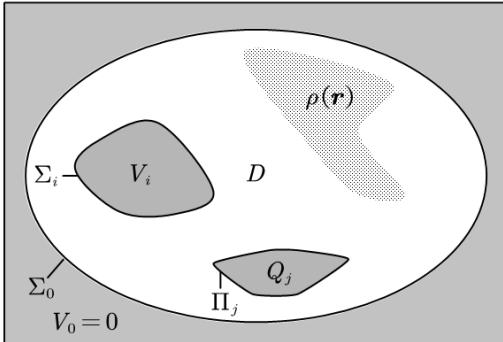


图 2.4: 边值问题

广义叠加原理: 两个边值问题如果 D, Σ_i, Π_j 一致. 第一个问题具有第一类边界条件 V_{1i} 和第二类边界条件 Q_{1j} 和内部电荷分布 $\rho_1(\mathbf{r})|_D$. 第二个问题具有第一类边界条件 V_{2i} 和第二类边界条件 Q_{2j} 和内部电荷分布 $\rho_2(\mathbf{r})|_D$. 两个问题的解分别为 $\varphi_1(\mathbf{r})|_D, \varphi_2(\mathbf{r})|_D$. 则同样以 D 为区域, Σ_i, Π_j 为 D 两类导体边界, 但是以 $V_{1i} + V_{2i}$ 和 $Q_{1j} + Q_{2j}$ 为两类边界条件的问题的解一定是:

$$\varphi(\mathbf{r}) = \varphi_1(\mathbf{r}) + \varphi_2(\mathbf{r})$$

这个定理的证明是显然的: 首先 φ 作为一个边值问题的解, 当且仅当它满足以下的一个全局条件和若干边值条件:

$$\nabla^2 \varphi = -\frac{\rho}{\varepsilon_0}$$

$$\varphi(\mathbf{r})|_{\Sigma_i} = V_i \quad , \quad \int_{\Pi_j} \nabla \varphi \cdot d\mathbf{A} = Q_j \quad , \quad \varphi(\mathbf{r})|_{\Pi_j} = C$$

那么既然 φ_1 和 φ_2 都是之前两个问题的解, 这就意味着:

$$\nabla^2 \varphi_1 = -\frac{\rho_1}{\varepsilon_0} \quad , \quad \nabla^2 \varphi_2 = -\frac{\rho_2}{\varepsilon_0}$$

$$\varphi_1(\mathbf{r})|_{\Sigma_i} = V_{1i} \quad , \quad \int_{\Pi_j} \nabla \varphi_1 \cdot d\mathbf{A} = Q_{1j} \quad , \quad \varphi_1(\mathbf{r})|_{\Pi_j} = C$$

$$\varphi_2(\mathbf{r})|_{\Sigma_i} = V_{2i} \quad , \quad \int_{\Pi_j} \nabla \varphi_2 \cdot d\mathbf{A} = Q_{2j} \quad , \quad \varphi_2(\mathbf{r})|_{\Pi_j} = C$$

那么定义 $\varphi = \varphi_1 + \varphi_2$, 它就自然也能满足:

$$\nabla^2 \varphi = -\frac{\rho_1 + \rho_2}{\varepsilon_0}$$

$$\varphi(\mathbf{r})|_{\Sigma_i} = V_{1i} + V_{2i} \quad , \quad \int_{\Pi_j} \nabla \varphi \cdot d\mathbf{A} = Q_{1j} + Q_{2j} \quad , \quad \varphi(\mathbf{r})|_{\Pi_j} = C$$

从而就的确是新问题的一个解. 但是这个解是否唯一? 答案是肯定的. 这就是著名的唯一性定理, 可以这么表述:

唯一性定理: 如果同一个边值问题具有两个解 φ_1, φ_2 . 那么 $\varphi_1 - \varphi_2 = 0$.

其证明的核心思想是利用上一章介绍的第一格林等式². 首先注意到如果 φ_1, φ_2 都是同一个边值问题的解. 那么命 $\varphi = \varphi_1 - \varphi_2$ 就是以下边值问题的解:

$$\nabla^2 \varphi = -\frac{\rho - \rho}{\varepsilon_0} = 0$$

$$\varphi(\mathbf{r})|_{\Sigma_i} = 0 \quad , \quad \int_{\Pi_j} \nabla \varphi \cdot d\mathbf{A} = 0 \quad , \quad \varphi(\mathbf{r})|_{\Pi_j} = C$$

上面这个等号右侧都是零的方程称作齐次方程(homogeneous equation). 那么问题就被转化为证明齐次方程的解必为零解. 想起第一格林等式:

$$\nabla \cdot (\phi \nabla \psi) = \phi \nabla^2 \psi + \nabla \phi \cdot \nabla \psi$$

同样是在 ϕ, ψ 都取 φ 的情况:

$$\varphi \nabla^2 \varphi = \nabla \cdot (\varphi \nabla \varphi) - (\nabla \varphi)^2$$

²其实使用简单的“电荷发出电场线, 沿电场线方向电势降低”的观点也是可以论证这个定理的. 请读者思考过程.

现在把 φ 就视作刚刚定义的 $\varphi_1 - \varphi_2$, 并把这个公式运用于 D 中的任意一点, 由于电势满足的条件 $\nabla^2\varphi = 0$, 就有:

$$(\nabla\varphi)^2 = \nabla \cdot (\varphi\nabla\varphi)$$

等号右侧会让人联想到奥-高定理. 事实上如果把它应用到 D : 左右两侧就能化为:

$$\int_D (\nabla\varphi)^2 dV = \int_D \nabla \cdot (\varphi\nabla\varphi) dV = \oint_{\partial D} \varphi\nabla\varphi \cdot dA = \oint_{(\cup_i \Sigma_i) \cup (\cup_j \Pi_j)} \dots = \sum_i \oint_{\Sigma_i} \varphi\nabla\varphi \cdot dA + \sum_j \oint_{\Pi_j} \varphi\nabla\varphi \cdot dA$$

在第一类导体边界上, 由于 $\varphi = 0$ 所以积分为零. 在第二类导体边界上, 首先 φ 也是常数所以可以提出到积分外面, 即:

$$\oint_{\Pi_j} \varphi\nabla\varphi \cdot dA = C \cdot \oint_{\Pi_j} \nabla\varphi \cdot dA = C \cdot 0 = 0$$

从而得到恰好等号的右侧就是 0. 这就相当于证明了这个电势 φ 满足:

$$\int_D (\nabla\varphi)^2 dV = 0$$

一个总是不小于零的 $(\nabla\varphi)^2$, 经过积分得到了零的结果. 这就说明它恒等于零:

$$\nabla\varphi = \mathbf{0}$$

从而电势就是一个常数, 由于在第一类边界条件下得取零, 故只能是:

$$\varphi = 0$$

以上两个原理是静电平衡问题能够被分析与解决的基础. 以下几类问题中, 区域 D 有高度对称性, 且内部都没有电荷分布 ρ , 那么其解决甚至用不上这两个原理, 下一节的电像法就非常依赖这两个原理了. 不过这里的几类情况足以解决很多实际的问题:

2.1.3 常见简单体系

1. 平行的无限大导体板

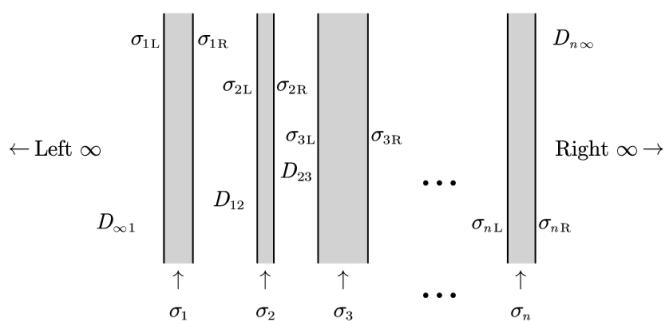


图 2.5: 平行板问题

我们先考虑无限大的导体平面板, 它可能具有一定的厚度. 把这样的 n 块板子从左至右平行放置, 如图 2.5. 这样就把空间隔离为中间的 $D_{12}, D_{23}, \dots, D_{n-1,n}$ 等区域.

我们没有必要严格写出这个问题的解. 把这个体系的解需要满足的性质罗列出来会更实用. 首当其冲地我们写出每一块板上的电量需要满足的关系. 如果板非常薄, 那么左右两侧的带电面密度 σ_{iL} 和右侧 σ_{iR} 就可以合成为统一的 σ_i . 那么以下两个特征是显著的:

$$1. \sigma_{iR} + \sigma_{i+1L} = 0$$

$$2. \sigma_{1L} = \sigma_{nR} = (\sigma_1 + \dots + \sigma_n)/2$$

而在第 i 与第 $i+1$ 块板之间的区域 D_{ii+1} 内, 如果研究场强, 以向右为正. 它应该满足两个关系:

$$3. E_{ii+1} = \sigma_{iR}/\epsilon_0 = -\sigma_{i+1L}/\epsilon_0$$

$$4. E_{ii+1} = (\sigma_1 + \dots + \sigma_i - \sigma_{i+1} - \dots - \sigma_n)/2\epsilon_0$$

请注意,在以上关系式中,1, 3 关系式仅仅依赖于上一节我们推出过的边界条件. 是无条件成立的. 但是 2, 4 却无法从边界条件中推导出来. 所以这两个式子的成立是有条件的: 那就是整个体系“仅仅”由这些带电的板组成. 事实上, 如果把第 1 块板和第 n 向左向右平移到无限远, 可以想象不改变板上的电荷分布也能保持静电平衡. 但是这样中间就只剩下 2, 3, …, $n-1$ 这些板, 它们就可以造成典型的仅仅符合 1, 3 而不符合 2, 4 的带电体系. 所以 2, 4 的正确性要求“无穷远处不能有电荷分布”. 此时我们可以通过叠加原理来解释 4 的正确性. 又反过来通过 4 计算出来的电场强度用 3 来计算各板左右两侧的带电量. 最后就能推导出 2 式来.

对于等式 2, 我们发现这实际上代表一种左右对称性: 在这些所有板的外侧区域 $D_{\infty 1}, D_{n \infty}$ 观察, 这些所有板等效于单块 $\sigma_1 + \dots + \sigma_n$ 的板. 从而左右的场强都是:

$$E_{\infty} = \frac{\sigma_1 + \dots + \sigma_n}{2\epsilon_0}$$

如果把上面的模型替换为电容器电路, 各个区域 D 变成电容内部的场强区, 而每一块导体板内部的区域被改成一根导线, 并可以通过外界引入电量. 那么需要注意, 这个体系额外多了一个性质: 一般情况下各个电容器外侧的电量都一定是零. 这是因为电容器初始状态必然不带电, 而充放电过程是拟稳³的, 故两极板积累的电量必然相反. 故按照上面的公式计算的 E_{∞} 就是零. 所以电容器极板外侧一般不允许有场强. 除非遇到电容器中间被放置了外电荷的情况, 比如在中间引入电压可以调控的栅极, 对于这种“异型”电容的问题处理方法就又回到了之前我们说的四个结论上.

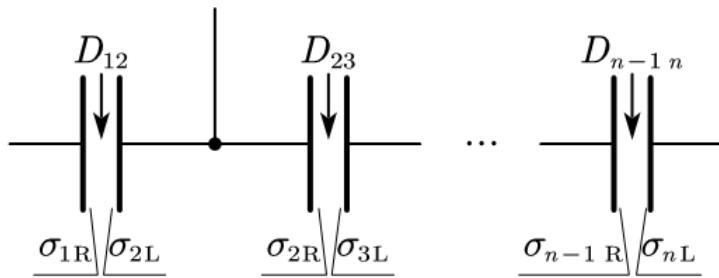


图 2.6: 电容器

对于一个面积为 S , 极板间距为 l 的平板电容器(parallel plate capacitor), 正常工作的时候根据之前的结论, 两板正对面带相反的电荷, 电场线起始于带正电的板而终止于带负电的板. 如果忽略其边缘效应, 即 $S \gg l$ 情况下, 我们可以认为其去掉边缘以后的正对面积与 S 的相对误差非常小, 电容就被定义为两板带正电或负电的绝对值与两板电势差的比值, 它等于:

$$C = \frac{Q}{U} = \frac{\epsilon_0 S}{l}$$

当平板电容器带电时, 其产生的能量可以通过静电能公式或计算外界做功, 也可以用电场能量计算得到:

$$E = \frac{1}{2}CU^2 = \frac{1}{2}QU = \frac{1}{2}\frac{Q^2}{C}$$

这个公式是普遍的, 对任何符合上一个式子定义的电容都是成立的.

2. 同心导体球与球壳

导体球与球壳反倒简单一些, 它们面上的总带电量不存在发散的情况, 而且也不需要考虑无穷远处的边界条件. 用各内表面总带电量 Q_i 和外表面带电量 Q_o 描述, 之前的四个条件现在也需要改写为:

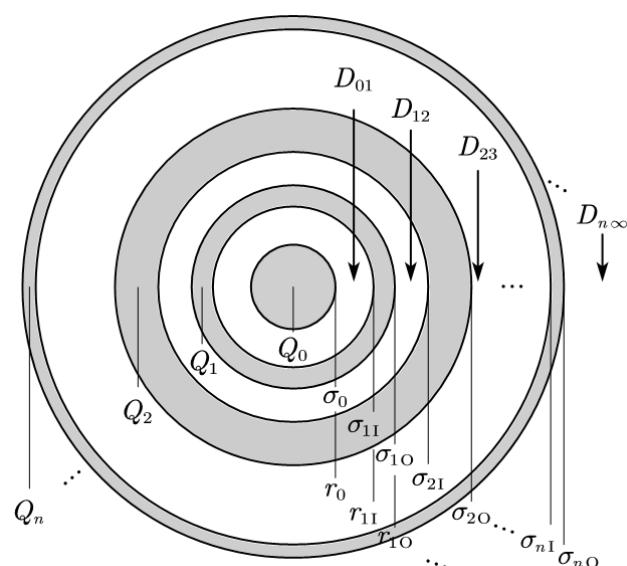


图 2.7: 球壳或柱壳

³即符合节点电流方程等, 见之后章节的说明.

1. $Q_{iO} + Q_{i+1I} = 0$
2. $Q_{nO} = Q_1 + \dots + Q_n$
3. $E_{i+1} = \frac{Q_{iO}}{4\pi\varepsilon_0 r^2} = -\frac{Q_{i+1I}}{4\pi\varepsilon_0 r^2}$
4. $E_{i+1} = \frac{Q_1 + \dots + Q_i}{4\pi\varepsilon_0 r^2}$

两个球壳构成的电容器,如果内外球壳半径分别为 r_I 和 r_O ,那么电容为:

$$C = 4\pi\varepsilon_0 \frac{r_O r_I}{r_O - r_I}$$

3. 共轴的导体圆柱与圆柱壳

我们先考虑无限长的圆柱.那么依然参考上图2.7.圆柱带电应用单位长度的电量 λ 来描述,四个关系式变为:

1. $\lambda_{iO} + \lambda_{i+1I} = 0$
2. $\lambda_{nO} = \lambda_1 + \dots + \lambda_n$
3. $E_{i+1} = \frac{\lambda_{iO}}{2\pi\varepsilon_0 r} = -\frac{\lambda_{i+1I}}{2\pi\varepsilon_0 r}$
4. $E_{i+1} = \frac{\lambda_1 + \dots + \lambda_i}{2\pi\varepsilon_0 r}$

两个柱壳构成的电容器,如果内外球壳半径分别为 r_I 和 r_O ,其长度为 l 且忽略边缘效应,即 $l \gg r_O > r_I$ 那么电容为:

$$C = 2\pi\varepsilon_0 l \ln \frac{r_O}{r_I}$$

2.2 电像法

下面考虑两种实际的边值问题的解.第一种问题中区域 D 为半无限大空间,即,其边界为一个无限大导体的表面.第二种问题中区域 D 要么是一个球面外的延伸到无穷远的区域(外问题),要么是球面内的区域(内问题),不管哪种,区域边界都是导体的球面表面,只是前一种是外表面,后一种是内表面.

我们提供的方法为电像法.其意义是,先考虑一个较简单的,在空间中 \mathbf{R} 处放置一个点电荷 q :

$$\rho(\mathbf{r}) = q\delta(\mathbf{r} - \mathbf{R})$$

的情况下,边值问题的求解.其中 $\delta(\mathbf{r} - \mathbf{R})$ 为**狄拉克德尔塔函数**(Dirac delta function),读者大可以把它视作某物理量的某种极限化的密度分布的抽象⁴,例如物理量总值为 1,分布在 $\mathbf{R} = \mathbf{0}$ 附近的一个非常小的范围内,那么其密度就是:

$$\delta(\mathbf{r}) = \begin{cases} +\infty & (\mathbf{r} = \mathbf{0}) \\ 0 & (\mathbf{r} \neq \mathbf{0}) \end{cases} \quad \& \quad \int_{\text{Ball}(\mathbf{0}, \epsilon)} \delta(\mathbf{r}) dV = 1$$

那么此前的密度函数 $\rho(\mathbf{r}) = q\delta(\mathbf{r} - \mathbf{R})$ 就可以表示点电荷的电荷密度分布.

⁴ 实则这样的广义函数数学上称作一种**分布**(distribution),它是通过积分形式的运算定义的:

$$f(\mathbf{r}) \circ g(\mathbf{r}) := \int f(\mathbf{r})g(\mathbf{r}) dV$$

$$\delta(\mathbf{r} - \mathbf{R}) \circ g(\mathbf{r}) := g(\mathbf{R})$$

2.2.1 半无限大空间的电像法

下面考虑在无限大导体平板处建立坐标系, $z > 0$ 就是待研究的 D 区域. 现在在 \mathbf{R} 处放置一个正点电荷 q , 容易证明, 由于导体板延伸到无穷远, 所以导体板的表面自然成为第一类边界条件的表面且其电势为零 $V = 0$. 表面上就会产生一个面电荷分布 $\sigma(x, y)$.

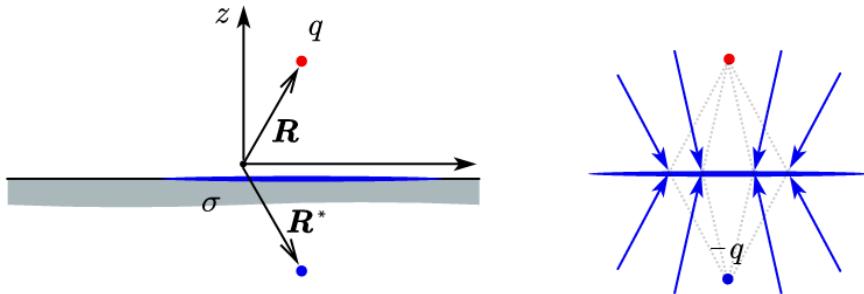


图 2.8: 半无限空间的电像

见上图右, 单独考虑这个面电荷产生的电场线, 由于在 $z < 0$ 区域内为导体, 合电场强度必然为零. 故这个区域面电荷的电场线要与原来的 \mathbf{R} 处的 q 的电场线相抵消. 故为指向 \mathbf{R} 处的汇聚的电场线. 但是, 由于这个面电荷产生的电场的上下的镜面对称性, 这就导致我们可以直接证明它对 $z > 0$ 的 D 区域产生的电场线, 实际上就相当于在 \mathbf{R} 的镜面对称点 \mathbf{R}^* 处放置一个像电荷(image charge), 值为 $-q$ 而产生的电场.

从而 $z > 0$ 区域的电场, 实际上就是由在 \mathbf{R} 的原电荷和在 \mathbf{R}^* 的像电荷共同产生的电场. 或者说, 满足以下方程的唯一解:

$$\nabla^2 \varphi = -\frac{q\delta(\mathbf{r} - \mathbf{R})}{\varepsilon_0} (z > 0) , \quad \varphi|_{z=0} = 0$$

就是:

$$\varphi(\mathbf{r}) = \frac{1}{4\pi\varepsilon_0} \frac{q}{|\mathbf{r} - \mathbf{R}|} - \frac{1}{4\pi\varepsilon_0} \frac{q}{|\mathbf{r} - \mathbf{R}^*|}$$

那么我们如何求解下列更加普遍的第一类边值问题⁵呢?

$$\nabla^2 \varphi = -\frac{\rho(\mathbf{r})}{\varepsilon_0} (z > 0) , \quad \varphi|_{z=0} = V$$

这个解答是简单的: 就是对以上点电荷电像法的情形做叠加原理. 我们把以上问题中, 如果点电荷为单位电荷时, 对应的解称作参数为 \mathbf{R} 处的格林函数(Green function):

$$\delta(\mathbf{r} - \mathbf{R}) \longrightarrow \varphi(\mathbf{r}) = G(\mathbf{R}, \mathbf{r}) = \frac{1}{4\pi\varepsilon_0} \left(\frac{1}{|\mathbf{r} - \mathbf{R}|} - \frac{1}{|\mathbf{r} - \mathbf{R}^*|} \right)$$

由于任意的电荷分布可以看成是点电荷的叠加, 下式 dV 为 \mathbf{R} 改变对应的体积元:

$$\rho(\mathbf{r}) = \int \rho(\mathbf{R}) \delta(\mathbf{r} - \mathbf{R}) dV$$

那么这个电荷分布产生的电势, 就是各个点电荷的格林函数的叠加:

$$\rho(\mathbf{r}) \longrightarrow \varphi(\mathbf{r}) = \int \rho(\mathbf{R}) G(\mathbf{R}, \mathbf{r}) dV$$

这样在导体面上产生的电势依然是零, 为了符合边界条件, 只需要整体加上常数 V , 最终结果为:

$$\varphi(\mathbf{r}) = V + \int \frac{1}{4\pi\varepsilon_0} \left(\frac{1}{|\mathbf{r} - \mathbf{R}|} - \frac{1}{|\mathbf{r} - \mathbf{R}^*|} \right) \rho(\mathbf{R}) dV$$

⁵ 第二类边值问题此时会有问题: 首先导体面上的总电荷量就是上方的自由电荷量总值的相反数而不可以任意给定, 而最终电势的结果也会产生一个未定常数的不确定性.

2.2.2 球面外与球面内的电像法

有了半无限区域的电像法求解边值问题做基础, 我们对球面内外区域的边值问题的求解就有了线索. 同样的, 在球心处建立原点和坐标系, 求解这两个问题的关键在于先求解以下两个最简单的第一类边值问题:

$$\nabla^2 \varphi = -\frac{q\delta(\mathbf{r} - \mathbf{R})}{\varepsilon_0} \quad (|\mathbf{r}| > a) \quad , \quad \varphi|_{|\mathbf{r}|=a} = 0$$

$$\nabla^2 \varphi = -\frac{q\delta(\mathbf{r} - \mathbf{R})}{\varepsilon_0} \quad (|\mathbf{r}| < a) \quad , \quad \varphi|_{|\mathbf{r}|=a} = 0$$

分别对应外问题和内问题. 即下图左和右:

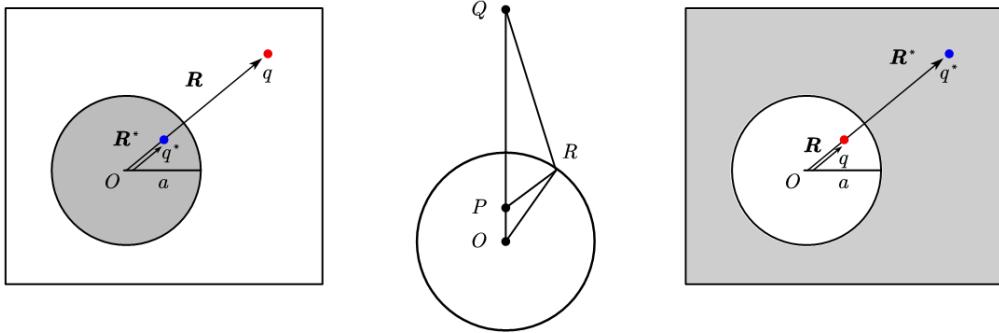


图 2.9: 球面的电像

而以上两个问题的电像解, 其实直接就都是:

$$\varphi(\mathbf{r}) = \frac{1}{4\pi\varepsilon_0} \frac{q}{|\mathbf{r} - \mathbf{R}|} + \frac{1}{4\pi\varepsilon_0} \frac{q^*}{|\mathbf{r} - \mathbf{R}^*|}$$

其中像电荷与原电荷在同一条半径上, 且其大小 q^* 和到原点的距离 R^* 与圆的半径 a 分别满足:

$$RR^* = a^2$$

$$-\frac{q^*}{q} = \frac{a}{R} = \frac{R^*}{a}$$

证明这两个解的正确性的关键在于原电荷与像电荷的位置对圆是构成了一对反演点的关系. 见上图中间的几何关系, 不难证明:

$$\overline{OP} \cdot \overline{OQ} = a^2 = \overline{OR} \Rightarrow \triangle OPR \sim \triangle ORQ$$

于是根据相似关系不难得到:

$$\frac{\overline{OP}^{1/2}}{\overline{PR}} = \frac{\overline{OQ}^{1/2}}{\overline{QR}}$$

所以只要让两个电荷分别位于 P, Q 处, 且电荷量反号且大小之比为 $\overline{OP}^{1/2} : \overline{OQ}^{1/2}$, 那么它们在 R 处的电势和就是零. 实际上上面的电势解就是取了这样的两个电荷的电势的和, 故就能满足原来的方程的边界条件, 而这样的电势恰好在给定区域 D 内也能满足对应的微分方程: 即 $\nabla^2 \varphi$ 只得到 \mathbf{R} 处电荷 q .

这就证明了以上写出来的解是正确的. 根据唯一性定理, 它也是唯一的.

下面考虑在普遍的第一类和第二类边值问题下外问题的解. 先考虑第一类边值问题, 即:

$$\nabla^2 \varphi = -\frac{\rho(\mathbf{r})}{\varepsilon_0} \quad (|\mathbf{r}| > a) \quad , \quad \varphi|_{|\mathbf{r}|=a} = V$$

同样的利用以上解写出格林函数:

$$G(\mathbf{R}, \mathbf{r}) = \frac{1}{4\pi\epsilon_0} \left(\frac{1}{|\mathbf{r} - \mathbf{R}|} - \frac{a/|\mathbf{R}|}{|\mathbf{r} - \mathbf{R}^*|} \right)$$

于是只需要写出下列积分, 就可以满足问题前一半的拉普拉斯方程:

$$\int \rho(\mathbf{R}) G(\mathbf{R}, \mathbf{r}) dV$$

这样一个势在边界上必然取零的值. 为了使得最后的解满足边界条件, 自然去思考叠加原理: 让外界 $\rho = 0$ 且边界上具有电势 V 的电势显然为:

$$\varphi = \frac{VR}{r}$$

故, 最后的解就是:

$$\varphi = \frac{VR}{r} + \int \frac{1}{4\pi\epsilon_0} \left(\frac{1}{|\mathbf{r} - \mathbf{R}|} - \frac{a/|\mathbf{R}|}{|\mathbf{r} - \mathbf{R}^*|} \right) \rho(\mathbf{R}) dV$$

但是如果换为第二类边值问题:

$$\nabla^2 \varphi = -\frac{\rho(\mathbf{r})}{\epsilon_0} (|\mathbf{r}| > a) \quad , \quad \int_{|\mathbf{r}|=a} \nabla \varphi \cdot d\mathbf{A} = Q, \quad \varphi|_{|\mathbf{r}|=a} = C$$

这就需要先考虑格林函数对应的势在边界上产生的总面电荷量, 他其实就是像电荷的值:

$$\delta(\mathbf{r} - \mathbf{R}) \longrightarrow \int_{|\mathbf{r}|=a} \nabla \varphi \cdot d\mathbf{A} = -\frac{a}{|\mathbf{R}|}$$

于是单独对格林函数进行积分得到的电势就使得面上带上电荷:

$$\rho(\mathbf{r}) \longrightarrow \int_{|\mathbf{r}|=a} \nabla \varphi \cdot d\mathbf{A} = -\int \frac{a}{|\mathbf{R}|} \rho(\mathbf{R}) dV$$

那么通过叠加原理, 为了使得面上电荷量变为 Q , 就需要引入一个电势:

$$\varphi = \frac{1}{4\pi\epsilon_0} \frac{Q + \int \frac{a}{|\mathbf{R}|} \rho(\mathbf{R}) dV}{r}$$

综上, 最后的解为:

$$\varphi = \frac{1}{4\pi\epsilon_0} \frac{Q + \int \frac{a}{|\mathbf{R}|} \rho(\mathbf{R}) dV}{r} + \int \frac{1}{4\pi\epsilon_0} \left(\frac{1}{|\mathbf{r} - \mathbf{R}|} - \frac{a/|\mathbf{R}|}{|\mathbf{r} - \mathbf{R}^*|} \right) \rho(\mathbf{R}) dV$$

球面外问题中, 所有电势的结果都是 $\rho(\mathbf{r})$ 和球面上的面电荷分布通过叠加原理产生的电势.

最后考虑内问题. 我们介绍过, 此时只有唯一的一种情况: 即内表面取为电势为零的第一类边界条件. 其问题表述和解必然为:

$$\nabla^2 \varphi = -\frac{\rho(\mathbf{r})}{\epsilon_0} (|\mathbf{r}| < a) \quad , \quad \varphi|_{|\mathbf{r}|=a} = 0$$

$$\varphi = \int \frac{1}{4\pi\epsilon_0} \left(\frac{1}{|\mathbf{r} - \mathbf{R}|} - \frac{a/|\mathbf{R}|}{|\mathbf{r} - \mathbf{R}^*|} \right) \rho(\mathbf{R}) dV$$

这个电势就是内部的电荷 $\rho(\mathbf{r})$ 和面上的感应面电荷共同对球内叠加产生的电势的和. 对于球外的电势的叠加必然为零.

2.2.3 立体谐函数 *

将一个中心在原点 $\mathbf{r} = \mathbf{0}$, 半径为 a 的球置于外场中讨论静电平衡为上一节中外问题的设置. 但是上一节选择了在距离球心 R 处放置了一个点电荷. 从而得到了球面在点电荷产生的外场下形成的格林函数 $G(\mathbf{R}, \mathbf{r})$. 实用问题中通常把外场按照**立体谐函数**(solid harmonics) 在原点进行分解, 类似于泰勒展开. 只不过立体谐函数更适合于球坐标:

$$R_{lm}(\mathbf{r}) = \sqrt{\frac{(l-m)!}{(l+m)!}} r^l P_{lm}(\cos \theta) e^{im\varphi}$$

$$P_{lm}(x) = (-)^m (1-x^2)^{m/2} \frac{d^m P_l(x)}{dx^m}$$

$$P_l(x) = \frac{(-)^l}{2^l l!} \frac{d^l}{dx^l} (1-x^2)^l = \frac{1}{2^l l!} \frac{d^l}{dx^l} (x^2 - 1)^l$$

这样的复杂定义其实非常自然地缘起于球坐标下拉普拉斯方程的解. 除了以上**正规**(regular) 部分, 还有一组对应的在原点发散的**非正规**(irregular) 解:

$$I_{lm}(\mathbf{r}) = \sqrt{\frac{(l-m)!}{(l+m)!}} r^{-(l+1)} P_{lm}(\cos \theta) e^{im\varphi}$$

非正规解作为电势代表的场由放在原点的电荷, 电偶极子乃至更高阶的对象产生. 因为我们在原点放置了一个导体球, 所以它受到的外场不可能含有非正规成分.

以上完备谐函数组非常复杂, 实际我们考虑的问题通常对球坐标 φ 的改变是对称的, 例如球外点电荷对原点附近的外场, 若以原点和点电荷的连线方向为极轴定义 θ, φ , 那么这个外场就不含任何立体谐函数中 $m \neq 0$ 的成分. 此时仅考虑 $m = 0$ 的那些函数是完备的:

$$R_l(\mathbf{r}) = \frac{1}{2^l l!} \cdot r^l \cdot \left(\frac{d}{\sin \theta d\theta} \right)^l \sin^{2l} \theta$$

$$I_l(\mathbf{r}) = \frac{1}{2^l l!} \cdot \frac{1}{r^{l+1}} \cdot \left(\frac{d}{\sin \theta d\theta} \right)^l \sin^{2l} \theta$$

对于较小的 l 我们罗列若干对称立体谐函数:

l	正规	非正规
0	1	$\frac{1}{r}$
1	$r \cos \theta$	$\frac{\cos \theta}{r^2}$
2	$\frac{r^2}{2}(3 \cos^2 \theta - 1)$	$\frac{1}{2r^3}(3 \cos^2 \theta - 1)$
3	$\frac{r^3}{2}(5 \cos^3 \theta - 3 \cos \theta)$	$\frac{1}{2r^4}(5 \cos^3 \theta - 3 \cos \theta)$
4	$\frac{r^4}{8}(35 \cos^4 \theta - 30 \cos^2 \theta + 3)$	$\frac{1}{8r^5}(35 \cos^4 \theta - 30 \cos^2 \theta + 3)$
\vdots	\vdots	\vdots

表 2.2: 对称立体谐函数表

值得注意的是,正规的立体谐函数总是可以写成多项式的形式.如果以 $\theta = 0$ 的方向为 z 轴,那么 $r \cos \theta = z$, $r^2 = x^2 + y^2 + z^2$.这样就得到了对称情况下的谐和多项式(harmonic polynomials),例如:

$$R_0(\mathbf{r}) = \mathcal{P}_0(x, y, z) = 1 \quad (2.1)$$

$$R_1(\mathbf{r}) = \mathcal{P}_1(x, y, z) = z \quad (2.2)$$

$$R_2(\mathbf{r}) = \mathcal{P}_2(x, y, z) = \frac{2z^2 - (x^2 + y^2)}{2} \quad (2.3)$$

$$R_3(\mathbf{r}) = \mathcal{P}_3(x, y, z) = \frac{2z^3 - 3z(x^2 + y^2)}{2} \quad (2.4)$$

$$R_4(\mathbf{r}) = \mathcal{P}_4(x, y, z) = \frac{35z^4 - 30z^2(x^2 + y^2) - 3(x^2 + y^2)^2}{8} \quad (2.5)$$

$$\vdots \quad (2.6)$$

不难验证它们都会符合拉普拉斯方程:

$$\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} \right) \mathcal{P}_l(x, y, z) = 0$$

就是这样的正规立体谐函数中的某项代表的电势,如果作为外场施加在置于原点的接地导体球上,应该得到怎样的静电平衡结果呢?首先我们明确这种场合下第一类边值问题表述为:

$$\nabla^2 \varphi = 0 (|\mathbf{r}| > a) \quad , \quad \varphi|_{|\mathbf{r}|=a} = 0 \quad , \quad \varphi|_{|\mathbf{r}| \rightarrow +\infty} = \lambda R_l(\mathbf{r})$$

由于正规立体谐函数电荷摆放在无穷远处,而球面上感应电荷产生的电势在无穷远处影响一次反比于半径,故此时边界条件变成了一个渐进边值条件(asymptotic boundary condition),它规定了半径趋近于无穷远时电势函数的变化趋势.

这个方程的解非常平庸地等于同阶的非正规立体谐函数与原来的正规谐函数的叠加:

$$\varphi(\mathbf{r}) = \lambda R_l(\mathbf{r}) + \mu I_l(\mathbf{r})$$

这是非常显然的.同阶的正规与非正规谐函数关于角度的部分变化趋势是一致的,为了满足在球面上电势为零的边值条件,我们只需要把半径取 a 并视作系数,再要求两个系数抵消,这只需要要求:

$$a^l \lambda + \frac{\mu}{a^{l+1}} = 0$$

从而马上得到 $\mu = -a^{2l+1} \lambda$.或者更简单地,我们把以下电势作为这一类问题的基础解:

$$\varphi = \frac{V_0}{2l!} \cdot \left[\left(\frac{a}{r} \right)^{l+1} - \left(\frac{r}{a} \right)^l \right] \cdot \left(\frac{d}{\sin \theta d\theta} \right)^l \sin^{2l} \theta$$

它代表怎样的电场呢?我们尤其关注 $l = 0, l = 1$ 两种简单情况.此时电势为:

$$l = 0 \quad : \quad \varphi = \frac{V_0 a}{r} - V_0$$

$$l = 1 \quad : \quad \varphi = \frac{V_0 a^2 \cos \theta}{2r^2} - \frac{V_0}{2a} r \cos \theta$$

作为外问题, $l = 0$ 电势中的 $-V_0$ 就必须理解为外场,而 $\frac{V_0 a}{r}$ 则理解为球面上的感应电荷产生的场,很显然,这表示一个均匀带电的球面外的场,如图2.10左的第一种情况.

作为外问题, $l = 1$ 电势中的 $-\frac{V_0}{2a} r \cos \theta = -E_0 z$ 就必须理解为外场,而 $\frac{V_0 a^2 \cos \theta}{2r^2} = \frac{E_0 a^3 \cos \theta}{r^2}$ 则理解为球面上的感应电荷产生的场,很显然,外场此时成为了 $+z$ 方向的匀强电场 $E_0 e_z$,而球面上的感应电荷产生的电势恰好与一个电偶极子产生的电势一致,电偶极矩为 $\mathbf{p} = 4\pi a^3 \epsilon_0 E_0$.如图2.10左的第二种情况.

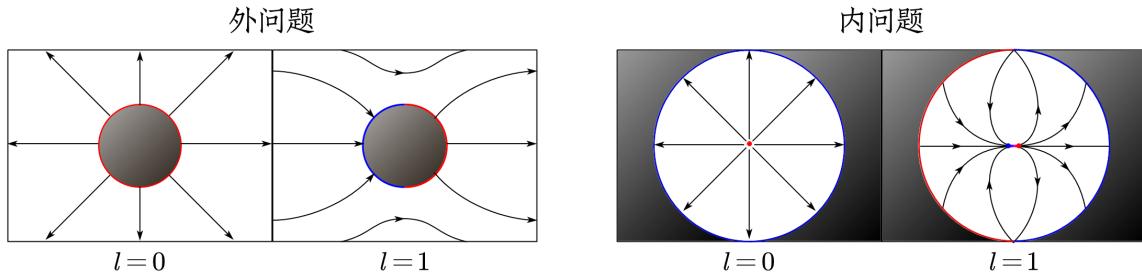


图 2.10: 低阶立体谐函数静电平衡

有意思的是,这样的解也可以适用于内问题.此时球面上的感应电荷在原点产生的场应该展开为正规项,而非正规项对应放在球心的外电荷,此前的解两项将交换意义:

作为内问题, $l = 0$ 电势中的 $\frac{V_0 a}{r}$ 就必须理解为外场,而 $-V_0$ 则理解为球面上的感应电荷产生的场,很显然,这相当于球心放了一个 $Q = 4\pi\varepsilon_0 a V_0$ 的点电荷造成的静电平衡,如图2.10右的第一种情况.

作为内问题, $l = 1$ 电势中的 $\frac{V_0 a^2 \cos\theta}{2r^2} = \frac{p_0 \cos\theta}{4\pi\varepsilon_0 r^2}$ 就必须理解为外场,而 $-\frac{V_0}{2a} r \cos\theta = -\frac{p_0}{4\pi\varepsilon_0 a^3} z$ 则理解为球面上的感应电荷产生的场,很显然,外场此时由 $+z$ 方向的电偶极子 $p_0 e_z$ 产生,而球面上的感应电荷产生的电势恰好是一个匀强电场 $E = \frac{p_0}{4\pi\varepsilon_0 a^3}$. 如图2.10右的第二种情况.

立体谐函数的意义在于,我们可以在原点附近去以球坐标的方式用正规的那些函数去展开一个电势,或者在一个球面外用非正规的那些函数去描写球面内的电荷体系产生的电势.故它天然地与多级展开挂钩.我们记得上一章我们用多级展开的思想来处理了偏移中心电荷的电势,多展开几项就是:

$$\frac{1}{4\pi\varepsilon_0} \frac{q}{|\mathbf{r} - \mathbf{R}|} = \frac{q}{4\pi\varepsilon_0} \left[\frac{1}{r} + \frac{R \cos\theta}{r^2} + \frac{R^2(3\cos^2\theta - 1)}{2r^3} + \frac{R^3(5\cos^3\theta - 3\cos\theta)}{2r^4} + \dots \right]$$

轻易就能发现,展开后各项,原来对应总电荷,电偶极矩,电四极矩,电八极矩等等,现在恰好就是一个个立体谐函数.它便是这样自然地产生于电势的多极展开中.个中原理在于以下展开式,又称**勒让德多项式的母函数**(generatin function of Legendre polynomials):

$$\frac{1}{\sqrt{1 - 2xt + t^2}} = \sum_{n=0}^{\infty} P_l(x)t^l$$

而其中 $P_l(x)$ 就是此前使用的**勒让德多项式**(Legendre polynomials).我们此前是通过**罗德里格公式**(Rodrigues' formula) 引入它的:

$$P_l(x) = \frac{(-1)^l}{2^l l!} \frac{d^l}{dx^l} (1 - x^2)^l = \frac{1}{2^l l!} \frac{d^l}{dx^l} (x^2 - 1)^l$$

我们简要介绍一种通过母函数证明罗德里格公式的方法:

容易发现,根式 $\sqrt{1 - 2xt + t^2}$ 出现在以下方程的解中:

$$tr^2 - 2r + 2x - t = 0 \quad \Rightarrow \quad r = \frac{1 - \sqrt{1 - 2xt + t^2}}{t}$$

这样 t 也可以用 r 表示:

$$t = \frac{2(x - r)}{1 - r^2}$$

但是 r 对 t 的泰勒展开为 $r = x + (x^2 - 1)t/2 + O(t^2)$. 故重新命:

$$s = r - x = \frac{1 - tx - \sqrt{1 - 2xt + t^2}}{t} = f(t)$$

这样用 s 表示 t 就是:

$$t = \frac{2s}{(s + x)^2 - 1} = g(s)$$

这样就存在两个互反函数 $f \circ g = id$. 且两者泰勒展开第一项都为一次项:

$$f(t) = f_1 t^1 + f_2 t^2 + \dots, \quad f_1 = \frac{x^2 - 1}{2}$$

$$g(s) = g_1 s^1 + g_2 s^2 + \dots, \quad g_1 = \frac{2}{x^2 - 1}$$

关于这样的函数有一个著名的拉格朗日反演定理(Lagrange inversion theorem), 如果用下角标表示泰勒展开的系数:

$$f_n = \frac{1}{n} \left(\frac{id}{g} \right)_{n-1}^n$$

写的更明确一点就是:

$$\frac{d^n}{dt^n} f(t) \Big|_{t=0} = \frac{d^{n-1}}{dA^{n-1}} \left(\frac{s}{g(s)} \right)^n \Big|_{s=0}$$

这个定理的证明如下: 把 $g(s)$ 代入 f 的泰勒展开式:

$$s = \sum_{m=1}^{\infty} f_m g^m(s)$$

两边对 s 求导数后乘以 $s^n/g^n(s)$:

$$\begin{aligned} \left(\frac{s}{g(s)} \right)^n &= s^n \sum_{m=1}^{\infty} f_m \frac{[g^m(s)]'}{g^n(s)} \\ &= s^n f_n \frac{[g^n(s)]'}{g^n(s)} + s^n \sum_{m \neq n}^{1 \rightarrow \infty} \frac{m}{m-n} f_m [g^{m-n}(s)]' \\ &= n f_n s^n \frac{g'(s)}{g(s)} + s^n \left(\sum_{m \neq n}^{1 \rightarrow \infty} \frac{m f_m}{m-n} g^{m-n}(s) \right)' \\ &= n f_n s^n \frac{g_1 + 2g_2 s + \dots}{g_1 s + g_2 s^2 + \dots} + s^n \left(\sum_{l=-n+1}^{\infty} \frac{l f_{n+l}}{l} (g_1 s + g_2 s^2 + \dots)^l \right)' \\ &= n f_n s^{n-1} \left(1 + \frac{g_2}{g_1} s + \dots \right) + s^n \left(\dots + \frac{c_{-2}}{s^2} + \frac{c_{-1}}{s} + c_0 + c_1 s + c_2 s^2 + \dots \right)' \end{aligned}$$

后一项中, 任意 s 的幂函数的导数都不会带来 s^{-1} 的结果项. 故等号右侧唯一的 s^{n-1} 项就是 $n f_n s^{n-1}$. 从而可以提取左侧的 s^{n-1} 项系数得到右侧的 $n f_n$. 原命题得证.

将这个结果重新表示 $f(t)$ 的泰勒展开, 得到:

$$f(t) = \sum_{n=1}^{\infty} \frac{t^n}{n!} \frac{d^{n-1}}{dA^{n-1}} \left(\frac{s}{g(s)} \right)^n \Big|_{s=0}$$

现在, 把此前的 $g(s)$ 代入:

$$\begin{aligned} \frac{1 - tx - \sqrt{1 - 2xt + t^2}}{t} &= \sum_{l=1}^{\infty} \frac{t^l}{2^l l!} \frac{d^{l-1}}{dA^{l-1}} [(s+x)^2 - 1]^l \Big|_{s=0} \\ &= \sum_{l=1}^{\infty} \frac{t^l}{2^l l!} \frac{d^{l-1}}{dx^{l-1}} (x^2 - 1)^l \end{aligned}$$

两边再次对 x 求导数:

$$\begin{aligned} -1 + \frac{1}{\sqrt{1 - 2xt + t^2}} &= \sum_{l=1}^{\infty} \frac{t^l}{2^l l!} \frac{d^l}{dx^l} (x^2 - 1)^l \\ \Rightarrow \quad \frac{1}{\sqrt{1 - 2xt + t^2}} &= \sum_{l=0}^{\infty} P_l(x) t^l, \quad P_l(x) = \frac{1}{2^l l!} \frac{d^l}{dx^l} (x^2 - 1)^l \end{aligned}$$

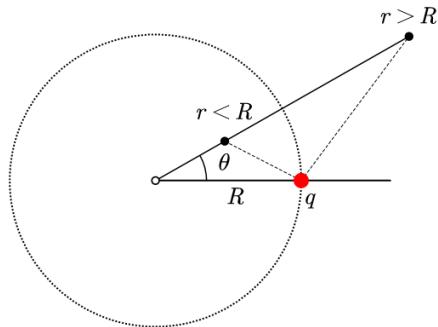


图 2.11: 电势的勒让德展开

有了这个公式就可以把点电荷的电势进行展开,如图2.11所示,在半径 r 很小时原点附近应当视 $x = r/R$ 展开为 r 的正规项,但是在半径 r 很大时的远方应当视 $x = R/r$ 展开为 r 的非正规项:

$$\varphi = \frac{q}{4\pi\epsilon_0} \frac{1}{\sqrt{R^2 + r^2 - 2Rr \cos\theta}} = \begin{cases} \sum_{l=0}^{\infty} \frac{qr^l P_l(\cos\theta)}{4\pi\epsilon_0 R^{l+1}} & (r < R) \\ \sum_{l=0}^{\infty} \frac{qR^l P_l(\cos\theta)}{4\pi\epsilon_0 r^{l+1}} & (r > R) \end{cases}$$

现在让我们来尝试一下新的工具用来解球面内外放置点电荷的第一类边值问题的结果.对于外问题,在 $r = a$ 处放置球面并接地,在 $r = R$ 处放置点电荷 q 并沿其方向建立球极坐标.则球面感受到电势:

$$\varphi_q = \sum_{l=0}^{\infty} \frac{qr^l P_l(\cos\theta)}{4\pi\epsilon_0 R^{l+1}} = \sum_{l=0}^{\infty} \frac{qa^l}{4\pi\epsilon_0 R^{l+1}} P_l(\cos\theta) \left(\frac{r}{a}\right)^l$$

根据此前的讨论,球面上的感应电荷就会在 $r > a$ 产生一个电势:

$$\varphi_{q^*} = - \sum_{l=0}^{\infty} \frac{qa^l}{4\pi\epsilon_0 R^{l+1}} P_l(\cos\theta) \left(\frac{a}{r}\right)^{l+1} = \sum_{l=0}^{\infty} \frac{(-qa/R)(a^2/R)^l P_l(\cos\theta)}{4\pi\epsilon_0 r^{l+1}}$$

恰好对应一个放置在 $r = R^*$ 处的 q^* 产生的电势:

$$q^* = -\frac{a}{R}q, \quad R^* = \frac{a^2}{R}$$

内问题的验证在这里留给读者作为一个练习.

2.3 电介质

之前介绍过,绝缘体只有考虑漏电或是大电场下击穿时,导电才得以发生.但这绝对不意味着忽略导电现象以后绝缘体对于静电场没有任何响应.这时发生的一个重要的现象为**极化**(polarization).它会造成很多独特现象:从在头发上摩擦过的塑料笔杆对小纸屑的吸引⁶,到电气工业大电容器的制作,更广义地说,在交变电场中介质的极化最终还能决定电磁波或光在介质中的传播,下一章我们将首次小小讨论以下这个问题.我们把明显能产生极化的介质称作**电介质**(dielectrics).极化的含义可以从微观和宏观两个角度来考虑,背后涉及物理学的多个方面.我们分别进行讨论.最后揭示微观角度与宏观角度极化规律的联系.

2.3.1 微观角度理解极化

固体或晶体,尤其是离子晶体或部分共价晶体,每一个原子实际上本身就带电,其极化机理非常复杂,需要用到爱因斯坦,玻恩,黄坤等人的陆续完善的理论.或是放弃研究微观机理,通过宏观的方式引入唯象的描述方法.我们下面研究的主要是由原子,分子构成的气体或者液体的极化.

在这样的场合下,每一个单元被单独研究,原子或分子整体是没有电性的,但是正如之前所说的,下式决定的原子或分子的电偶极矩不必为零:

$$\mathbf{p} = \int \rho \mathbf{r} dV$$

如果从非量子的经典物理观点(包括半经典的玻尔模型)来看,即使简单如氢原子,看作轻的电子绕近似不动的质子做匀速圆周运动,每一个瞬时也会形成一个电偶极矩.但是它随时间快速转动而导致在静电场意义下其平均值为零.事实上,量子力学认为无论哪一个时刻,氢原子都处于电偶极矩沿任何方向概率都相

⁶用金属屑的静电平衡也可以实现,不过金属屑密度大现象更不明显.

等的定态(stationary state), 并没有在随时间演变. 以无论是从理论上来看还是从实际上来看, 孤立氢原子基态下总是不体现电偶极矩的:

$$\bar{p} = 0$$

具有这样性质的分子就称作无极分子或非极性分子(nonpolar molecule), 反之则称作有极分子或极性分子(polar molecule) 一般可以通过化学原理给出分子的成键和立体构型加以判断. 例如 CO 分子和 H₂O 具有极性而 CO₂ 分子和 CH₄ 则是无极的.

现在把无极分子放在外电场中, 情况就会发生戏剧性的转变: 把无极分子感受到的局部的电场视作匀强场, 这个外场记做 \mathbf{E}' . 即使是简单如氢原子, 其平均电偶极矩也不应当再视作零. 这一个现象有若干解释方法:

- 最唯象的解释方法把氢原子视作一个由劲度系数 k 的弹簧连接的作为正电负电中心的质点构成的体系, 带电量为 $\pm e$. 但是忽略正负电荷之间的电相互作用力, 只用去考虑受到的外力与等效电相互作用出来的线性弹性力的平衡, 那么自然平衡时从负电荷到正电荷具有位移:

$$\mathbf{r} = \frac{e\mathbf{E}'}{k}$$

从而得到分子平均电偶极矩:

$$\bar{p} = er = \frac{e^2}{k} \mathbf{E}'$$

- 一个具有可证伪性的结果由以下考虑描述. 通过量子理论的计算, 把原子外的电子看作电子云, 它是一个总电量为 $-e$ 的连续电荷密度分布:

$$\rho = \frac{-e}{\pi a_0^3} e^{-\frac{2r}{a_0}}$$

其中 $a_0 = 5.3 \times 10^{-11} \text{ m}$ 为玻尔半径. 进一步假设这个电荷分布为彻底刚性的. 在外电场作用下, 仅仅是原来在原点的质子现在位移到了 \mathbf{r} 处. 质子对电子云的力就是电子云对质子的力, 它与外电场力平衡. 这样计算可以得到:

$$\mathbf{r} = \frac{3\pi\varepsilon_0 a_0^3}{e} \mathbf{E} \quad , \quad \bar{p} = 3\pi\varepsilon_0 a_0^3 \mathbf{E}'$$

- 以上结果仅仅是与实验测量结果数量级一致, 但是差了 6 倍, 即实验结果表明:

$$\bar{p} = 18\pi\varepsilon_0 a_0^3 \mathbf{E}'$$

这是因为上一个做法中认为电子云不会发生任何改变, 反倒是质子位移了的做法不符合事实. 正确的做法反而需要依然以质子为原点去解库仑场与弱匀强电场下电子的薛定谔方程的微扰解. 这样才可以给出电子云的新分布. 同样的计算也会带来著名的斯托克效应(Stark effect). 即玻尔模型对应的第 n 能级(第 $n - 1$ 激发态)现在将分裂为 $2n - 1$ 个能级. 进而影响谱线.

不管用何种方式解释, 以上现象就是第一种极化机理, 称作位移极化(distortion polarization). 它的结果为每一个分子产生了一个量子平均的电偶极矩, 在弱场下这个电偶极矩应当正比于电场强度:

$$\bar{p} = \alpha \mathbf{E}'$$

其中 α 称作分子极化率(molecular polarizability). 它是表征微观极化难易程度的重要参数. 对于位移极化, 分子极化率是一个量子行为的结果, 从氢原子的极化率就可以看出, 因为玻尔半径 a_0 的计算依赖于普朗克常数的取值:

$$\alpha = 18\pi\varepsilon_0 a_0^3 = \frac{9(4\pi\varepsilon_0)^4 \hbar^6}{2m_e^3 e^6} = 7.42 \times 10^{-41} \text{ F} \cdot \text{m}^2$$

而对于有极分子, 分子即使在没有外场作用时, 由于其电荷分布本身就会造成一个电偶极矩, 称作**固有电偶极矩**(permanent dipole moment), 我们记做 p_0 , 但是由于无规则的热运动, 其方向 e 是随机的, 所以作为每个分子的电偶极矩矢量 $\mathbf{p} = p_0 e$ 的统计平均, 其结果依然应当为 $\mathbf{0}$:

$$\bar{\mathbf{p}} = \mathbf{0}$$

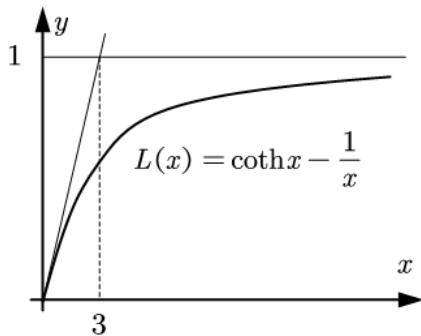


图 2.12: 郎之万函数 $\bar{p} = p_0 L\left(\frac{pE'}{kT}\right)$, $L(x) = \coth x - \frac{1}{x}$

现在施加外电场 E' 以后, 如果认为分子依然保持刚性, 即 p_0 不变, 但方向可以自由转动, 则与电场保持同向的电偶极子的势能得到降低, 电偶极子在动力学上倾向于与外电场同向排列. 但是其热运动越剧烈, 转动动能就越大, 不同电偶极子相互影响就会使得其排列趋于杂乱无章. 忽略可能的量子效应, 由玻尔兹曼分布可以得到在温度为 T 时分子沿电场方向的平均电偶极矩:

$$\bar{p} = p_0 \left(\coth \frac{pE'}{kT} - 1 \right)$$

一般把上面这个函数称作**郎之万函数**(Langevin function), 即:

这个函数的图像如2.12所示. 它有两条渐近线, 分别对应两种具体的物理情形:
如果低温 (T 很小) 或者强场 (E' 很大) 但不至于漏电或击穿. 使得宗量 $x = pE'/kT \gg 1$, 那么就使用水平 $y = 1$ 渐近线, 即:

$$\bar{p} \approx p_0$$

对应几乎所有电偶极子都沿电场方向排列的**饱和极化**(saturated polarization) 情况.

如果高温 (T 很大) 或者弱场 (E' 很小). 使得宗量 $x = pE'/kT \ll 1$, 那么就使用原点切线 $y = x/3$ 的渐近线, 即:

$$\bar{p} \approx \frac{p_0^2}{3kT} E'$$

写成矢量式, 我们发现与位移极化得到了相似的结果:

$$\bar{\mathbf{p}} = \alpha \mathbf{E}' , \quad \alpha = \frac{p_0^2}{3kT}$$

这被称作**郎之万-德拜定律**(Langevin-Debye law). 同样地把 α 称作分子极化率. 而这种极化机理称作**取向极化**(orientation polarization). 取向极化的极化率是一个热学行为的结果, 从极化率反比于温度这一点就可以看出来.

关于取向极化和位移极化我们总结以下几点:

一是要注意两者的不同: 前者是外电场诱导产生或改变固有电偶极矩, 是一个量子行为, 几乎不依赖于温度. 后者是外电场诱导固有电偶极矩选取优势方向, 是一个热学现象, 定性解释不依赖于量子特性. 但是严格说来温度对前者也会有非常微小的影响, 低温下量子特性对后者也会有不小的改变.

二是要注意两者可以同时出现: 无极分子当然只能采取位移极化. 但有极分子的取向极化和位移极化是并存的. 所以分子极化率的完整计算方法为分别考虑两种极化方式带来的系数的和:

$$\alpha = \alpha_{\text{dis}} + \alpha_{\text{ort}}$$

三是要注意具体情况主要考虑哪一种极化方式: 一般说来, 有极分子的极化率中取向极化远大于位移极化的贡献⁷, 故有极分子的极化主要是取向极化. 但是无极分子, 位移极化就取代取向极化称为唯一需要

⁷使分子转向导致的能量差小, 使电子在不同能级间跃迁需要的能量差大, 前者易后者难.

考虑的极化因素了,故静电场下其分子极化率一般比有极分子要小一个量级.但是对于高频电场,**弛豫时间**(relaxation time)成为了另一个需要考虑的重要因素.大约对于超过红外线频率的电场下,取向的改变就难以跟上电场的变化频率了,这导致介质的极化转而成为了电子的位移极化,考虑介质对光的色散和吸收都是基于这一点的.

2.3.2 宏观角度理解极化

之前章节已经介绍过极化强度 \mathbf{P} 的定义与相关的性质.宏观上的极化规律指的是,如果介质内部某点处局域地存在(平均意义下的)场强 \mathbf{E} ,那么该点就局域地产生一个极化强度 \mathbf{P} ,做为 \mathbf{E} 的响应函数:

$$\mathbf{P} = f(\mathbf{E})$$

这样的函数实际上或是应用上都不一定是线性的,但是在各向同性介质,弱场,低频的三个条件下,理论或实验都能给出线性的结果,一般写为:

$$\mathbf{P} = \chi_e \varepsilon_0 \mathbf{E}$$

这样 χ_e 就是只与材料和环境(温度,压强)有关的,而与外加电场大小与方向无关的无量纲标量常数.称作**极化率**(susceptibility).

类似于导体的静电平衡问题,介质的极化是如何达到平衡的?,这一个问题也应当动态地去思考:对比导体中的平衡条件 $\mathbf{E} = \mathbf{0}$,介质中的平衡条件也不过是一个最终才能达到的结果,达到平衡前的弛豫过程,等号不一定要成立.导体未平衡时,内部正是因为有电场才驱使电荷定向运动从而转移,从而电场才被改变,在改变中逐渐趋于零.那么介质也是类似:如果把一块未被极化的介质置于外电场 \mathbf{E}_0 中,一开始 $\mathbf{P} = \mathbf{0}$ 将小于预期的 $\chi_e \varepsilon_0 \mathbf{E}_0$ 值,那么极化就会慢慢发生, \mathbf{P} 开始增大,位移极化和取向极化在逐渐发生,宏观看会产生**极化电流**(polarization current),不难验证其电流密度为 $\mathbf{J}_P = \partial \mathbf{P} / \partial t$.但是由于 \mathbf{P} 的变化,势必会在介质表面甚至内部产生极化电荷分布 σ_P, ρ_P .这些电荷产生的场将部分抵消原来的外电场,称作**退极化场**(depolarization field).于是 \mathbf{E}_0 也将动态的变化到最终的 \mathbf{E} ,使得以上平衡方程成立.这样极化强度就不再增加.

所以同样的以极化平衡条件 $\mathbf{P} = \chi_e \varepsilon_0 \mathbf{E}$ 为出发点,应当期望可以类似于导体静电平衡那样得到有效的求解带电介质的静电平衡问题.但是由于平衡条件更复杂,我们的方法也需要进行改进.事实上不难验证改进的方法取 $\chi_e \rightarrow +\infty$ 就会退化为导体的情况.所以改进的方法反而更具有普遍性.

最核心的一点是引入辅助的**电位移**(electric displacement)矢量来简化问题.这个矢量被定义为:

$$\mathbf{D} = \varepsilon_0 \mathbf{E} + \mathbf{P}$$

用 \mathbf{D} 的值反过来计算反映极化的 \mathbf{P} ,这会带来两个好处,第一是,极化规律的平衡条件被等价地写作:

$$\mathbf{D} = \varepsilon_0 \mathbf{E} + \chi_e \varepsilon_0 \mathbf{E} = \varepsilon_r \varepsilon_0 \mathbf{E} = \varepsilon \mathbf{E}$$

同样的 $\varepsilon_r = 1 + \chi_e$ 也会作为一个基本常数,与原来的极化规律的结构复杂度一致.这个常数称作**相对介电常数**(relative permittivity),其与真空介电常数 ε_0 的乘积 $\varepsilon = \varepsilon_r \varepsilon_0$ 则称作**绝对介电常数**(absolute permittivity)或简单称作**介电常数**(permittivity).合理的极化率应当为正数,而根据相对介电常数的定义,它则要是大于 1 的正数.而介电常数的实用程度大于极化率的实用程度.

第二个关键的好处是,我们已经知道, \mathbf{E} 与电荷通过高斯定理联系,而 \mathbf{P} 与极化电荷也可以通过高斯定理联系:

$$\nabla \cdot \mathbf{E} = \frac{\rho}{\varepsilon_0}, \quad \nabla \cdot \mathbf{P} = -\rho_P$$

代入刚才的电位移的定义马上就可以发现:

$$\nabla \cdot \mathbf{D} = \rho_f$$

即, 电位移通过高斯定理联系的电荷就是“自由”电荷. 这里的“自由”不是指导体中或导体表面的电荷, 恰恰相反, 如果把导体静电平衡问题中的导体也视作 $\epsilon_r \rightarrow +\infty$ 的特殊电介质, 这些电荷反倒要被排除. ρ_f 的含义通过推导上式使用的关系式 $\rho_f = \rho - \rho_P$ 就很容易看出来了, 表示除去介质极化产生的面电荷或体电荷以外的一切电荷, 一般都是置于介质外部区域的给定电荷, 但对于介质情况它也可以分布在介质内.

对于带介质的静电平衡问题的求解就有了两种思路, 第一种思路仅仅适用于比较简单的情况, 这种情况不需要借助我们引入的电位移矢量 D 就能得到结果. 分别是把无限大块状介质和介质球放在匀强电场 E_0 中的求解:

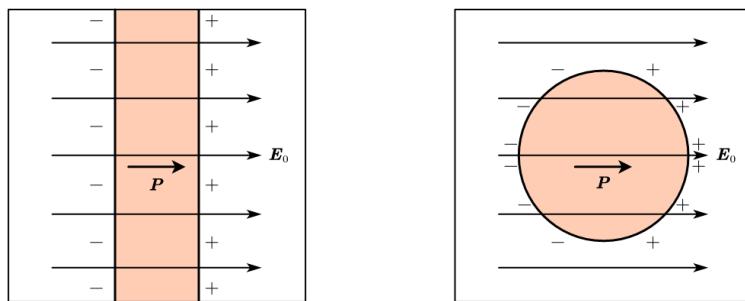


图 2.13: 两种简单情况

这两个问题的核心在于大胆猜测介质内部是均匀极化的. 设极化强度为 P . 对于平板介质块, 极化电荷在两侧产生的电荷面密度 $\pm\sigma_P$ 就是:

$$\sigma_P = P$$

从而在介质内部形成的退极化场就是:

$$\mathbf{E}_P = -\frac{\mathbf{P}}{\epsilon_0}$$

从而极化规律导致:

$$\mathbf{P} = \chi_e \epsilon_0 \mathbf{E} = (\epsilon_r - 1) \epsilon_0 \left(\mathbf{E}_0 - \frac{\mathbf{P}}{\epsilon_0} \right)$$

这就解出了平衡时的极化强度:

$$\mathbf{P} = \frac{\epsilon_r - 1}{\epsilon_r} \epsilon_0 \mathbf{E}_0$$

我们关心的是介质内部的电场强度:

$$\mathbf{E} = \mathbf{E}_0 - \frac{\mathbf{P}}{\epsilon_0} = \frac{\mathbf{E}_0}{\epsilon_r}$$

这个结论很贴切地反应了 ϵ_r 的物理意义: 它是简单的平板情况下介质对电场的削弱. 故对于导体情况, 内部电场强度在静电平衡时被削弱为零, 故可以视作 $\epsilon_r \rightarrow +\infty$.

注意到这样的 $\pm\sigma_P$ 在介质块外场强为零, 故不影响外界. 从而如果把多块介质块叠放, 不同介质块就相当于按照不同的相对介电常数去削弱同一个电场. 也就是说如图2.14的三个介质块上方下方的导体板上的电荷面密度如果为 σ , 就会有:

$$E_0 = \frac{\sigma}{\epsilon_0}, \quad \epsilon_{r1} E_1 = \epsilon_{r2} E_2 = \epsilon_{r3} E_3 = E_0$$

但是要注意, 这个 E_0 不一定要与旁边的空间中的 E' 相等, 即在 S 和 S' 面上的电荷面密度没有必要是相等的, 同一个表面的带电不可以均匀分布 E_i , $i = 1, 2, 3$ 也没有必要等于 E_0 , 这不会与电场的环路定理矛盾, 因为在介

质块的边缘电场实际上已经不能视作匀强电场,而这一点作为边缘效应被我们忽略了.只有以下等势条件式是必须成立的:

$$E_1 d_1 + E_2 d_2 + E_3 d_3 = E' d = U$$

除了严格把以上量都解出来,我们讨论这个体系的静电平衡时经常会把它看做电容,每一个电容的性质和值由以下关系描述:

$$Q = CU \quad , \quad C = \frac{\epsilon S}{d}$$

注意电容器的并联等效电容应当满足直接求和的关系,串联反倒需要用电阻并联的方式来计算:

$$C_{\text{series}} = C_1 + C_2 \quad , \quad C_{\text{parallel}} = C_1 \parallel C_2 = \frac{C_1 C_2}{C_1 + C_2}$$

于是上一个问题的总电容就是:

$$C = (C_1 \parallel C_2 \parallel C_3) + C'$$

现在让我们关心一下在三个介质中的电位移矢量,就会发现:

$$D_1 = D_2 = D_3 = \epsilon_0 E_0 = \sigma$$

可以发现电位移与忽略电介质时的电场其实就差一个物理常数 ϵ_0 . 它的存在可以理解为仅仅为了统一量纲.

现在让我们看介质球的极化问题. 此时均匀的极化为球的表面带来的电荷面密度就变成了 $\sigma_P = P \cos \theta$, 而如上一章所述,这样的电荷分布带来的退极化场为:

$$\mathbf{E}_P = -\frac{\mathbf{P}}{3\epsilon_0}$$

仅为块状情况的三分之一. 可以想见,其极化会更加的强. 事实上平衡方程与解变为:

$$\mathbf{P} = (\epsilon_r - 1)\epsilon_0 \left(\mathbf{E}_0 - \frac{\mathbf{P}}{3\epsilon_0} \right) \Rightarrow \mathbf{P} = \frac{3(\epsilon_r - 1)}{\epsilon_r + 2}\epsilon_0 \mathbf{E}_0$$

那么球内的电场就被削弱到了:

$$\mathbf{E} = \mathbf{E}_0 - \frac{\mathbf{P}}{3\epsilon_0} = \frac{3}{\epsilon_r + 2}\mathbf{E}_0$$

可见此时 ϵ_r 就没有体现出它的物理意义来: 它并不能代表任意情况下介质对外加电场的削弱的比例,这里被削弱的比例变成了:

$$\epsilon = \frac{\epsilon_r + 2}{3}$$

从而这种情况下研究电位移表面上看上去没有意义:

$$\mathbf{D} = \epsilon_0 \mathbf{E} + \mathbf{P} = \frac{3\epsilon_r}{\epsilon_r + 2}\epsilon_0 \mathbf{E}_0$$

其实不然,下面我们来看带电介质的静电平衡问题的第二种思路,它也是普遍问题的处理思路,主要就是需要借助电位移矢量来考虑问题的.

首先,在同一种介电常数为 ϵ 的介质中,我们抛弃极化矢量 \mathbf{P} ,用场强度 \mathbf{E} 和电位移 \mathbf{D} 可以完备地表达出极化平衡以后的场与荷(下面默认去除极化电荷,平衡的时候极化电荷的值是有唯一解的未知量)关系,它由三个关系构成:

$$\mathbf{D} = \epsilon \mathbf{E}$$

$$\begin{cases} \nabla \cdot \mathbf{D} = \rho \\ \nabla \times \mathbf{E} = 0 \end{cases} \quad \text{或} \quad \begin{cases} \oint_{\partial V} \mathbf{D} \cdot d\mathbf{A} = \int_V \rho dV \\ \oint_{\partial A} \mathbf{E} \cdot dl = 0 \end{cases}$$

第一个式子称作电介质的**本构方程**(constitutive equation). 后两个式子就是介质中的高斯定律和环路定律. 实际上还缺少一个势与它们的关系, 势与场之间的关系不需要做任何改变:

$$\mathbf{E} = -\nabla \varphi$$

从而很容易得到势与荷之间的关系, 现在变为:

$$\nabla^2 \varphi = -\frac{\rho}{\varepsilon}$$

那么现在我们考虑电介质的静电平衡问题的表述方法. 与导体静电平衡的区别在于, 除了各真空区域 $\varepsilon = \varepsilon_0$ 的电势分布待解以外, 各个介质内部的电势分布也是待解的, 我们甚至可以把外电荷引入介质的内部. 这样我们索性就把空间按大块介质分成不同不相交的区域 D_i 的并:

$$\mathbb{R}^3 = \bigcup_i D_i$$

每一个区域内部的介电常数为 ε_i . 而待解的电势为 φ_i , 如果该区域内有外电荷分布 $\rho(\mathbf{r})$, 则在介质内部需要满足:

$$\nabla^2 \varphi(\mathbf{r})|_{\mathbf{r} \in D_i} = -\frac{\rho(\mathbf{r})}{\varepsilon_i}$$

但是凭以上方程不足以确定整个问题, 一般还需要配合**边界条件**(boundary relation). 介质静电平衡问题的边界条件有它独特的重要性, 一般会做如下表述:

如图2.15, 在介质 i, j 的分界面上, 把 i 指向 j 为正方向的场的法向分量的电位移记为 D_{ni}, D_{nj} . 而切向分量以其自然方向为正方向的电场强度记做 $E_{\tau i}, E_{\tau j}$. 如果面上没有外电荷, 只有极化电荷, 那么根据高斯定律和环路定律, 容易得到:

$$D_{ni} = D_{nj}, \quad E_{\tau i} = E_{\tau j}$$

通常称作: **法向电位移连续, 切向电场强度连续**.

但是面上如果有 σ_{ij} 的外电荷, 第一个式子就必须改为:

$$D_{nj} - D_{ni} = \sigma_{ij}$$

用这样的方程来确定平板电容器或其他具有高对称性电容器中的场分布非常方便. 即使是之前的球在匀强电场中极化的问题, 抛弃极化矢量而取而代之利用电位移也不会造成困难. 此时我们假设表面上的电荷分布为 $\sigma = \sigma_0 \cos \theta$. 那么对面内面外产生的退极化场场强为:

$$\mathbf{E}_P = \begin{cases} -\frac{\sigma_0}{3\varepsilon_0} \mathbf{e} & (r < R) \\ \frac{1}{4\pi\varepsilon_0} \frac{2p \cos \theta}{r^3} \mathbf{e}_r + \frac{1}{4\pi\varepsilon_0} \frac{p \sin \theta}{r^3} \mathbf{e}_\theta & (r > R) \end{cases}$$

其中 $p = 4\pi\sigma_0 R^3 / 3$. 现在取在 $\theta = 0, r = R$ 处的分界面, 将退极化场与外场叠加, 得法向内外场强分别为:

$$E_1 = E_0 - \frac{\sigma_0}{3\varepsilon_0}, \quad E_2 = E_0 + \frac{1}{4\pi\varepsilon_0} \frac{2p}{R^3} = E_0 + \frac{2\sigma_0}{3\varepsilon_0}$$

再由法向电位移连续条件, 即得到:

$$\varepsilon_r \varepsilon_0 E_1 = \varepsilon_0 E_2 \Rightarrow \sigma_0 = \frac{3(\varepsilon_r - 1)}{\varepsilon_r + 2} \varepsilon_0 E_0$$

代入之前的电场表达式就得到了平衡时的场强分布.

但是含介质问题的静电平衡问题难度有甚于导体的静电平衡问题:在导体静电平衡问题中我们发展了格林函数法,并对区域 D 的边界为平面和球面两种情况分别引入了电像法.但是对于介质问题,可以证明平面情况依然可以使用电像法,但是球面的静电平衡问题外电荷已经不存在单个的简单的电像.此时原则上静电平衡的求解需要用到偏微分方程求解的通用思路,这已经超出本教材的范围,在此不再加以介绍.

2.3.3 微观与宏观的联系

我们已经得到微观和宏观的极化规律并使用后者解决了一部分问题:

$$\bar{p} = \alpha E' , \quad P = \chi_e \varepsilon_0 E$$

当然如果单个分子的平均电偶极矩为 \bar{p} ,而这些分子以数密度 n 组成介质,那么介质的极化强度满足以下关系式也是正确无误的:

$$P = n\bar{p}$$

但因此而认为微观的分子极化率是通过以下关系式确定宏观的极化率 χ_e 的话就不够准确了:

$$\chi_e = \frac{n\alpha}{\varepsilon_0}$$

这是因为 E' 和 E 略有区别.但是我们先指出,以上做法对于稀疏的材料,比如气体,通常是足够准确的了.但是一旦分子稠密起来,比如低温下的液氦或是分子晶体固体,就必须考虑 E' 和 E 的区别.前者是造成分子极化的场强,而后者则是在宏观小,微观大尺度下平均化的局域场强.两者的区别实际上就在于偶极子自己的场强在前者计算中需要略去,因为自己的场强不可能极化自己.而后者则需要加以考虑.定量上的差值实际上我们在上一章末也进行过计算,关系式为:

$$E' = E + \frac{P}{3\varepsilon_0}$$

这样就得到更准确的关系式:

$$\chi_e = \frac{n\alpha}{\varepsilon_0 - n\alpha/3}$$

可见,如果 $n\alpha \ll \varepsilon_0$,那么之前的近似计算法就可以适用,此时总是有 $\chi_e \ll 1$,从而只会得到 $\varepsilon_r \approx 1$ 介质的介电特性与真空区别不大.所以对于所有介电特性明显的电介质,实际的联系微观分子极化率与宏观极化率的公式必然是以上式子,或者写为相对介电常数:

$$\varepsilon_r = \frac{\varepsilon_0 + 2n\alpha/3}{\varepsilon_0 - n\alpha/3}$$

这两个公式称为克劳修斯-莫索蒂关系(Clausius-Mossotti relation).我们对于稀薄介质的判定条件进一步加以说明.如果采用氢原子的位移极化模型,我们就知道 $\alpha = 18\pi\varepsilon_0 a_0^3$,代入条件 $n\alpha \ll \varepsilon_0$ 就得到了:

$$na_0^3 \ll 1$$

如果把每个分子占据的体积近似为 a_0^3 ,这就是说,单位体积内被分子占据的体积必须是个小量,主要的体积必须是被真空占据才能称作“稀薄”.这一点也符合物理的直觉.但是要注意,如果考虑取向极化,分子极化率要大一个量级,那么就要求 n 更小一个量级才能符合“稀薄”的条件.

2.4 再议静电能

导体静电平衡时的总静电能计算并不会带来任何问题.之前的用荷与势的公式和用场的计算公式依然适用:

$$I = \frac{1}{2} \int_V \rho \varphi dV = \int_V \frac{1}{2} \epsilon_0 \mathbf{E}^2 dV$$

唯一要注意的问题是别在使用电像法的场合,把像电荷也代入第一个电荷乘电势的公式去计算了.这很可笑,第一可笑是即使出现了像电荷,它也仅仅是作为一种替代:用来替代某一部分感应电荷在某一特定区域产生的场强与电势.但绝对不是说这个虚假的电荷和它那个位置的实际的电势会对这个像电荷产生一个实际的能量.第二可笑是因为在更多的静电平衡问题中原则上没法用电像法,就更加不存在使用上面的积分计算静电能时没有考虑像电荷上的能量的问题了.

还有一个细节可以简化:实际上导体静电平衡问题中除了第一类或第二类区域,现在统一记为 D_i 中可以有 $\rho(r)$ 分布以外,其余部分就是各个导体,每一个导体都等势,故具有带电量 Q_j 和电势 V_j .那么总的静电能就可以更具体地表达为:

$$I = \sum_i \frac{1}{2} \int_{D_i} \rho \varphi dV + \sum_j \frac{1}{2} Q_j V_j$$

如果实际电荷分布实际由微观的点电荷形成,那么在平均场意义下的上式应当代表所有点电荷之间的相互作用能的和.而点电荷的自能并不包含在上式中:它一般也不会改变.

但是如果遇到电介质极化的情况,问题就大不相同了!因为此时伴随着分子的极化,分子内部的能量也会改变.位移极化改变了分子的自能是不言而喻的.但是取向极化何以改变分子内部的自能?分子仅仅是在外电场下选取了优势方向而已.事实上,注意到极化率是温度的函数,就非常容易理解,此时有效的能量其实不是真正的系统内能,而是热力学理论中的自由能,它代表体系能量最低原理与熵最大原理的一种平衡:

$$F = U - TS$$

当分子取向杂乱无章时,极化强度 $\mathbf{P} = \mathbf{0}$,此时熵最大,自由能最小.可以取为自由能的零点.但是一旦产生极化, $\mathbf{P} \neq \mathbf{0}$,此时熵减小,自由能就因此而增加.

这样就不难理解,讨论介质极化时,此时的能量(自由能)应当具有四个组成部分:

1. 非宏观静电形式的,形成极化导致的极化能(polarization energy).
2. 单纯看介质极化自己产生的,介质内外的电场能,即介质自己的静电自能.
3. 不考虑介质时其他电荷产生的全空间的电场能,即其他电荷分布的静电自能.
4. 介质与其他电荷之间的静电相互作用能.

其中后三个部分构成了我们之前意义上的“静电能”.但是它并不是能量的全部.考虑一个能量守恒的过程时,应当有外界做的功等于体系内部所有能量的变化.如为一个电容器充电的过程.很容易发现,如果介质内部采取取向极化,在一个等温的极化过程中介质是要向外界放热的,因为体系的熵在降低.故外界需要做的功,并不是要与体系的内能挂钩,而是直接与自由能挂钩:

$$W = \Delta U - Q = \Delta U - T \Delta S = \Delta F$$

所以从这个意义上说,我们此时定义的静电能,应当包括第一项极化能,而直接就是指自由能:

$$I = F$$

我们从平板电容器的极化来思考其表达式,因为这种情况场是简单的匀强场,而且区域限制在简单的长方体的内部.我们不妨预先假设,极化能的能量密度应当为极化强度的二次函数:

$$w_p = \frac{1}{2} \alpha \mathbf{P}^2$$

做这个假设的原因是因为自由能必然可以对 \mathbf{P} 泰勒展开, 而空间的各向同性决定了它必然不含一阶的项 (一维情况就是必然是偶函数), 那么处于弱场近似可以忽略高阶项. 其实待会也可以发现, 这是唯一能够得到之前的极化规律的能量密度的取法.

那么在电容器完成极化后, 真正的静电能, 就应当为上式代表的极化能与后三项代表的之前意义上的“静电能”的和, 其能量密度为:

$$w = w_p + w_0 = \frac{1}{2}\alpha\mathbf{P}^2 + \frac{1}{2}\varepsilon_0\mathbf{E}^2$$

此时我们可以将介质的极化看作是一种自由能最小的结果: 需要将外场 \mathbf{E}_0 视作一种不变的约束, 而极化强度被视作自由能中的变量, 在极化平衡时应当有以下自由能密度函数最小:

$$w = \frac{1}{2}\alpha\mathbf{P}^2 + \frac{1}{2}\varepsilon_0(\mathbf{E}_0 - \frac{\mathbf{P}}{\varepsilon_0})^2$$

那么由导数为零的方式得到:

$$\alpha\mathbf{P} = \mathbf{E}_0 - \frac{\mathbf{P}}{\varepsilon_0}$$

对比之前的极化规律, 我们就非常容易看出来:

$$\alpha = \frac{1}{\chi_e\varepsilon_0}$$

从而得到极化能量密度的表达式:

$$w_p = \frac{\mathbf{P}^2}{2\chi_e\varepsilon_0}$$

这一点我们在用稀薄气体模型配合位移极化中的最唯象的弹簧模型进行一个验证. 位移极化使得每个弹簧偏离平衡位置 \mathbf{r} 以后, 弹簧势能造成的能力密度就是极化能的本质:

$$w_p = n \cdot \frac{1}{2}k\mathbf{r}^2$$

再结合宏观极化强度, 微观电偶极矩与位移值三者的关系:

$$\mathbf{P} = np = ner$$

得到:

$$w_p = \frac{\mathbf{P}^2}{2ne^2/k}$$

而由之前讨论过的平衡方程得到分子极化率:

$$\alpha = \frac{e^2}{k}$$

从而宏观极化率, 在稀薄条件下为:

$$\chi_e = \frac{n\alpha}{\varepsilon_0}$$

把上两式去代换之前的极化能, 就得到了:

$$w_p = \frac{\mathbf{P}^2}{2\chi_e\varepsilon_0}$$

所以这就是均匀各向同性的介质的极化能公式. 如果与之前的“静电能”合并, 就得到了体系真正的静

电能,如果利用极化规律就得到:

$$\begin{aligned}
 w &= \frac{\mathbf{P}^2}{2\chi_e\epsilon_0} + \frac{1}{2}\epsilon_0\mathbf{E}^2 \\
 &= \frac{(\chi_e\epsilon_0\mathbf{E})^2}{2\chi_e\epsilon_0} + \frac{1}{2}\epsilon_0\mathbf{E}^2 \\
 &= \frac{1}{2}(1+\chi_e)\epsilon_0\mathbf{E}^2 \\
 &= \frac{1}{2}\epsilon_r\epsilon_0\mathbf{E}^2 \\
 &= \frac{1}{2}\epsilon\mathbf{E}^2 \\
 &= \frac{1}{2}\mathbf{D} \cdot \mathbf{E}
 \end{aligned}$$

可见其实相比纯粹的静电场,能量密度就是单纯的增大了 ϵ_r 倍.

以上静电能也具有用电荷来计算的变式.在一块状均匀介质的内部区域 D_i ,由于此时电势梯度依然是场强 \mathbf{E} ,但是场强的散度与极化电荷挂钩,故如果考虑电位移 $\mathbf{D} = \epsilon\mathbf{E}$,其散度就会是外电荷,从而得到:

$$\nabla^2\varphi = -\frac{\rho}{\epsilon}$$

通过上一章类似的积分公式,在这块介质内部总是有:

$$\frac{1}{2}\int_{D_i}\rho\varphi dV = \int_{D_i}\frac{1}{2}\epsilon\mathbf{E}^2 dV + \oint_{\partial D_i}\frac{1}{2}\varphi \cdot \epsilon\mathbf{E} \cdot d\mathbf{A}$$

前一项的积分是我们熟知的,后一项在朝外为正的表面上的积分可以拆解为各个介质分界面上的项的求和,如果是介质 1, 2 之间的分界面,两项积分合成为:

$$\int_{\Sigma_{12}}\frac{1}{2}\varphi(\mathbf{D}_1 - \mathbf{D}_2) \cdot d\mathbf{A}_{12} = -\int_{\Sigma_{12}}\frac{1}{2}\varphi(D_{2n} - D_{1n})dA$$

由之前说过的介质边界条件,两个法向电位移的差就是在分界面上具有的外电荷面密度,故以上积分被简单化为:

$$-\frac{1}{2}\int_{\Sigma_{12}}\sigma_{12}\varphi dA$$

综上,在分块介质情况下,体系的总静电能就等价于求体内与分界面上两类外电荷与电势产生的以下积分值:

$$I = \sum_i \frac{1}{2} \int_{D_i} \rho\varphi dV + \sum_j \frac{1}{2} \int_{\Sigma_j} \sigma\varphi dA$$

最后,容易验证,即使有电介质,电容器具有的总能量的表达式形式依然不变,为:

$$E = \frac{1}{2}CU^2 = \frac{Q^2}{2C} = \frac{1}{2}QU$$

这一点不难理解:虽然这个能量并不必真正意义上内部的“内能”,但是它作为自由能,的的确确要与充电过程中外电路需要做的功相联系.而在这个意义上通过外电路电压对积累的电荷量积分就会得到完全一致的结果.只是要注意,如果涉及取向极化,电容器充放电的过程就总是伴随着与外界的吸放热.

第三章 稳恒电流

3.1 稳恒电流描述与形成

3.1.1 德鲁特模型

电荷的定向移动形成电流(current). 就好像一缸气体在慢慢挪动那样, 电荷的定向移动并不是纯粹的匀速运动, 而是与无规则的热运动相叠加. 1900 年前后德鲁特 (P. Drude) 和洛伦兹 (H. A. Lorentz) 等人提出德鲁特模型(Drude model) 来解释金属中的导电现象. 主要观点是金属内部自由运动的电子类似于理想气体那样做自由的运动, 称为自由电子气(free electron gas). 我们用电子电量 $-e$, 质量 m , 电子数密度 n , 和弛豫时间(relaxation time) τ , 平均速度 v 来表示其特征. 弛豫时间就是电子做匀速直线运动, 与原子实两次碰撞之间的平均间隔时间. 与之相关的另一个量还可以是平均自由程(mean free path) λ . 容易想像, 典型的情形是, 常温下电子的平均速度是和气体分子的平均热运动速度那样, 一个非常巨大的速度, 而金属原子之间的距离又是那么地短, 导致电子发生十分频繁的碰撞. 而如果在金属中加一个电场 E , 它在两次碰撞内是只可以让电子速度改变一个十分微小的量的:

$$\mathbf{a} = -\frac{e\mathbf{E}}{m}$$

$$\mathbf{v}(0) \rightarrow \mathbf{v}(t) = \mathbf{v}(0) + \mathbf{a}t \rightarrow \mathbf{v}(\tau) = \mathbf{v}(0) + \mathbf{a}\tau$$

我们根据统计力学的思想, 计算在某一时刻电子速度对所有电子取热平衡分布的平均:

$$\mathbf{u} = \langle \mathbf{v} \rangle = \langle \mathbf{v}(0) + \mathbf{a}t \rangle = \mathbf{a}\tau$$

上式 t 代表距离上次碰撞每个电子幸存的时间. $\mathbf{v}(0)$ 代表上次碰撞后其速度. 一方面, 认为碰撞使得电子速度完全随机分布, 平均的结果为零. 另一方面, 认为电子的碰撞是一个泊松过程, 其碰后幸存时间的概率分布是一个指数分布:

$$p(t) = e^{-\frac{t}{\tau}}$$

从而平均的 t 为

$$\langle t \rangle = - \int_0^\infty t dp = pt|_0^\infty + \int_0^\infty e^{-\frac{t}{\tau}} dt = \tau$$

最后结合电流密度为:

$$\mathbf{j} = -nev$$

我们得到:

$$\mathbf{j} = \frac{ne^2\tau}{m}\mathbf{E}$$

的结果. 称为微观欧姆定律(microscopic Ohm's law), 其中系数被称为电导率(electrical conductivity):

$$\sigma = \frac{ne^2\tau}{m} ; \quad \mathbf{j} = \sigma\mathbf{E}$$

十分值得指出的是,作为德鲁特模型的另一个重要结论,金属的热导率(thermal conductivity)在理论中也可以给出一个估计值.我们都应该知道金属比绝大多数其他固体都拥有好得多的导热性能.这可以用热导率与傅里叶热传导定律(Fourier's law of thermal conduction)来描述:

$$\mathbf{q} = -\kappa \nabla T$$

其中 \mathbf{q} 为热流密度, κ 即为材料的热导率.水常温下热导率只有 $0.591\text{W}/(\text{m}\cdot\text{K})$,但纯的铜却能够达到 $401\text{W}/(\text{m}\cdot\text{K})$.金属的高热导率全都得益于轻盈的电子气,它能够迅速地把局部的热运动加剧传导开来.类比理想气体的非平衡态统计方法,我们给出:

$$\kappa = \frac{1}{3}nv\lambda c$$

式中 c 为每个电子的动能与温度的比.利用 $\lambda = v\tau$, 我们把热导率和电导率做比,便可以把较难确定的散射弛豫时间消去,得到:

$$\frac{\kappa}{\sigma} = \frac{mv^2c}{3e^2}$$

德鲁特模型认为,作为类似于理想气体的电子气,理应有:

$$c = \frac{3}{2}k \quad ; \quad \frac{1}{2}mv^2 = \frac{3}{2}kT$$

从而得到:

$$\frac{\kappa}{\sigma T} = \frac{3}{2}\left(\frac{k}{e}\right)^2$$

历史上金属热导率和电导率的比值与温度相关的现象很早就被人们发现,一般温度升高时电导率会有十分明显的下降,也就是金属的电阻会上升.小灯泡在未发光时与正常发光时电阻就经常有 2 倍左右的差距.而 1853 年魏德曼 (G. Wiedemann) 和弗朗茨 (R. Franz) 观察到不同金属虽然导电导热性能差距悬殊,室温下两者之比却接近一个常数,称为魏德曼-弗朗茨定律(Wiedemann-Franz's law).而洛伦茨 (L. Lorenz)¹ 在 1872 则把这一经验规律确定到常数与绝对温度成正比的形式.最后德鲁特电子论将这个常数以微观常数的形式确定下来,等式右边现在约为 $1.11 \times 10^8 \text{W} \cdot \Omega/\text{K}^2$, 这与实验结果数量级是一致的,但却差了约 2 倍.真实的值被称为洛伦兹常数(Lorentz number):

$$\frac{\kappa}{\sigma T} = L = \frac{\pi^2}{3}\left(\frac{k}{e}\right)^2 = 2.44 \times 10^8 \text{W} \cdot \Omega/\text{K}^2$$

这个理论很成功,但最后的结果的不符合让人感到疑惑.还有一个令人感到疑惑的问题.就是对于电子气的电容.既然电子是完全独立于金属原子实的另一个热力学体系.按照经典统计物理.原子实的摩尔热容与离子晶体等类似,大概在 $3R$ 附近².而电子按只有平动自由度的理想气体考虑应贡献 $1.5R$ 热容.实验却否定了这一点,金属常温下热容仍然是在 $3R$ 附近.电子仿佛没有被热激发,然而电子又势必参与导热,因为金属的导热性能明显优于其他物质.那么以上推导过程中哪儿存在不合理之处?几年后到来的量子力学革命让人们认识到这一个经典模型从一开始就做了不止一个的与微观实际情况不符的假定,但又是十分巧合地,最终结果与真实值数量级自动一样了.

3.1.2 费米子观点 *

元素周期律引发了人们对原子核外电子排布规律的研究,人们惊奇地发现电子是费米子(fermion),符合泡利不相容原理(Pauli exclusion principle).这赋予电子独特的量子特性.具体来说,单位体积内如果电子数目越多,那么其最低平均能量就必须越大.因为能量最低的状态一旦被占据,其他电子就必须占据能量更高的状态,即使按照最低能量的方式去堆积(绝对零度时的行为),电子也将具有很高的平均能量.它符合类

¹注意,丹麦物理学家路德维希·洛伦茨 (Ludwig Lorenz)(1829-1891) 与荷兰物理学家亨德里克·洛伦兹 (Hendrik Lorentz)(1853-1928) 是两个不同的人.有一个方程以它们两人的名字共同命名:Lorenz-Lorentz 关系.

²即督龙-裴替定律(Dulong-Petit Law).

似位置-动量不确定性原理的反比率, 电子浓度的减小了单电子占据的位置尺度, 则它的动量就会增加, 从而根据色散关系 $p^2 = 2mE$ 其能量也会升高. 这一点使得我们去修改经典的德鲁特模型. 相应的量子气体称为费米气(Fermi gas).

定量计算费米气的特性需要考虑电子的波动本性. 我们暂时取金属为长宽高为 ABC 的长方体, 那么如果将 N 个电子倒入这个容器, 电子的数密度为:

$$n = \frac{N}{ABC}$$

让我们考虑一下电子对状态的填充, 长方体相当于一个谐振腔, 事实上给出了电子动量状态的量子化:

$$p_x A = ah, p_y B = bh, p_z C = ch$$

$$a, b, c \in \mathbb{N}$$

这是因为根据德布罗意关系 $p = \hbar k$ 而电子状态需要满足 $kL = 2m\pi$ 的缘故. 从而我们发现电子的三个方向的动量都是量子化的. 量子化的单位为:

$$\Delta p_x = \frac{\hbar}{A}, \Delta p_y = \frac{\hbar}{B}, \Delta p_z = \frac{\hbar}{C}$$

现在把 N 个电子倒入动量空间中, 那么电子在低温下将从最低能量态开始填充, 由于电子数目巨大, 最后电子将填充到能量为 ε_F 处. 这样的一个填满电子的动量空间中的球体称为费米海(fermi sea), 最终填充到的能量称为费米能级 ε_F , 对应的电子动量为费米动量 p_F . 注意到一个动量空间中的一个标志状态的坐标内部实际上有两个独立的状态, 它们表示即使这两个波的波矢 k 一样, 它们代表电子的自旋也不一样³. 从而我们写出:

$$2 \cdot \frac{4\pi p_F^3}{3} / (\Delta p_x \Delta p_y \Delta p_z) = N$$

从而得到:

$$p_F = \left(\frac{3Nh^3}{8\pi N}\right)^{\frac{1}{3}} = \left(\frac{3nh^3}{8\pi}\right)^{\frac{1}{3}}$$

在以上过程中总电子数和总体积被消去了, 最后电子堆积到的动量值仅仅取决于各处电子的数密度而成为了强度量. 这样一个由于泡利不相容原理所造成的动量值对应到速度上, 对一般金属估计约为 10^6 m/s . 而如果按非费米气计算, 热运动速度应为 $\sqrt{kT/m} \sim 7 \times 10^4 \text{ m/s}$, 我们发现经典结果是严重估计少了的. 但电子的热容又估计多了. 量子统计给出:

$$c = \frac{\pi^2}{2} k \cdot \frac{kT}{\varepsilon_F}$$

代入热导率公式, 得:

$$\frac{\kappa}{\sigma} = \frac{\pi^2}{6} \cdot \frac{mv_F^2}{\varepsilon_F} \cdot \left(\frac{k}{e}\right)^2 \cdot T$$

恰好, $\varepsilon_F = \frac{1}{2}mv_F^2$, 从而我们得到了正确形式的魏德曼-弗朗茨定律:

$$\frac{\kappa}{\sigma T} = \frac{\pi^2}{3} \left(\frac{k}{e}\right)^2$$

我们最后做一个经典德鲁特模型与量子费米气模型的比较. 两个理论都承认以下基本公式的成立:

$$\sigma = \frac{ne^2\tau}{m}, \quad \kappa = \frac{1}{3}nv^2\tau c$$

³电子的微观描述是旋量波, 也就是二分量波函数, 具有两种可能的自旋.

先看电导率, 经典理论与量子理论对各个参数的估计除了 τ 其他都是一致的. 而经典更倾向于对 τ 的估计更小. 因为经典理论一般认为 $\tau = \lambda/v$, 而 λ 一般就按照原子实之间的平均距离来估计. 这一点之后就知道是不妥当的, 因为电子实际上可以在严格周期性的势能场中毫无散射的传播下去. 而使得电子能够被散射的其实是晶格的缺陷, 热振动等因素, 从而一般温度升高, τ 减小, 金属的导电性能就大大减弱. 这给出了金属电阻随温度升高而升高的结论. 这里经典理论对 λ 的过低估计恰好被对 v 的过低估计所抵消掉一部分, 从而最后 τ 的值在常温下差距也不大. 而对于热导率, 经典理论对 v 的过低估计又恰好被对 c 的过高估计修正, 精确到了仅仅相差一个常数, 除了经典理论说不清楚的 τ , 经典与量子的在数量级上是基本符合的.

值得一提, 热导率描述导热, 电导率描述导电, 而导电现象的附效应便是热产生. 也就是焦耳热(Joule heating). 微观的功率密度

$$\mathbf{j} = \sigma \mathbf{E} \quad , \quad w = \mathbf{j} \cdot \mathbf{E}$$

也即:

$$w = \sigma \mathbf{E}^2 = \rho \mathbf{j}^2$$

其中 **电阻率**(resistivity) ρ 为电导率的倒数. 上为**微观焦耳-楞次定律**(microscopic Joule-Lenz law).

3.1.3 能带论 *

费米气模型在量子力学迅猛发展的几十年后回头看又是过于 naive 了. 人们发现电子终究是在晶格间传播的波动, 如果原子实与背景的电子给出了严格的周期性的势能, 那么如果电子的总机械能小于最大势能, 那么就成为了束缚态, 限制在一个原子附近运动而成为原子实的一部分. 而如果电子守恒机械能大于最大势能, 就成为不受任何阻碍的传导态. 动能虽在运动中有所波动, 但绝不会越来越小. 描述电子的波函数称为**布洛赫波函数**(Bloch wave function). 而电子如何去填充怎样的能级? 布洛赫(F. Bloch)研究并发现了**能带论**(energy band theory) 来替代旧的费米气理论, 这个理论可以统一地描述导体与绝缘体.

能带论可以看成是电子在单原子外的束缚态特征与传导的自由电子气的特征的一种综合, 一方面电子在这样的一个周期性势场中的运动其守恒的机械能并不是可以取所有值, 而是分为可以连续取值的区间: **能带**(band) 与根本取不到的能量区间: **带隙**(gap) 所构成. 而电子作为波动用其所在的能带与波矢 \mathbf{k} 来描述. 其波矢方向的物理意义不再重要, 因为每一个态一般都代表某种驻波与行波的混合, 而每一个态附近的群速度则可以代表这个态代表的真实电子运动速度. 另一方面这些能带的结构又恰好对应到了单原子的束缚态能级. 这使得我们经常是一望某元素的单原子核外电子排布, 便能从其是否具有最外层电子读出这种元素单质是否能导电的信息.

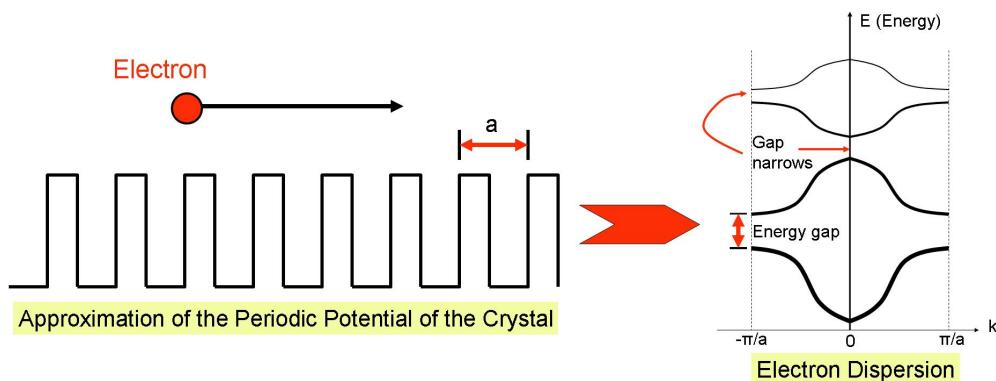


图 3.2: 电子在周期性势场中传播

某种原子形成晶体时, 电子从最低的能带开始填充, 内层电子填满了较低的那些能带, 它们不参与导电, 对应于原子实电子和在晶体共价键键合的价电子. 最外层电子较少的原子, 形成晶体时最后填充的那个能带没有填满, 或者虽然填满但与下一个能带在能量区间上重合而能够导电. 称为**导体**(conductor). 对应

的元素也就显示出明显的金属性⁴. 而若原子外层电子数目多, 则很有可能电子恰好填满一个能带后隔着一个能隙与空的上方能带相望, 则所有态都稳定不变, 材料所有电子都是束缚电子, 最后填充的那个满带称为**价带**(valence band), 代表原子间成共价键的价电子所在的能带. 而上方的那个空带称为**导带**(conduction gap). 意为一旦这个带中存在电子则会大大改善其导电性能. 此时元素体现出非金属性, 一般形成的都是**绝缘体**(insulator). 而在金属元素与非金属元素分界处则存在很多的**类金属**(metalloid), 它们形成的晶体虽然与绝缘体一样, 在最低能量时恰好把价带排满, 而隔着一个带隙与空的导带相望. 但带隙很小 (小于4eV), 导致常温的热激发和掺杂, 抑或是光照激发都将在导带中激发大量载流子而具有可观的导电性能. 这些材料称为**半导体**(semiconductor).

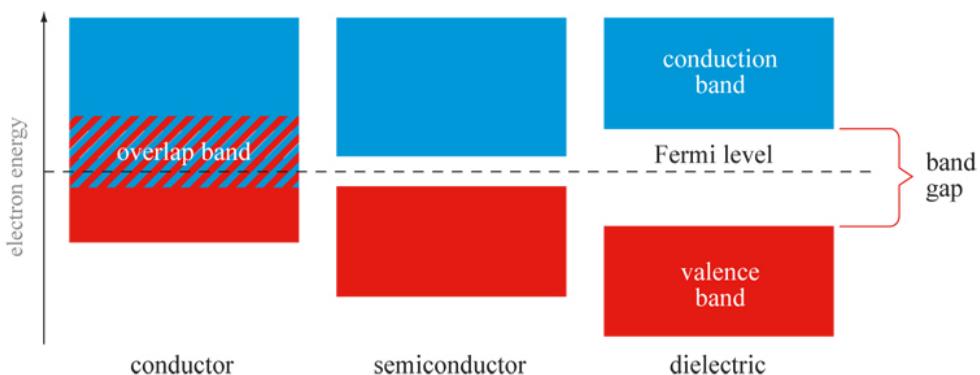


图 3.3: 能带与导电性

1 H														2 He			
3 Li	4 Be																
11 Na	12 Mg																
19 K	20 Ca	21 Sc	22 Ti	23 V	24 Cr	25 Mn	26 Fe	27 Co	28 Ni	29 Cu	30 Zn	31 Ga	32 Ge	33 As	34 Se	35 Br	36 Kr
37 Rb	38 Sr	39 Y	40 Zr	41 Nb	42 Mo	43 Tc	44 Ru	45 Rh	46 Pd	47 Ag	48 Cd	49 In	50 Sn	51 Sb	52 Te	53 I	54 Xe
55 Cs	56 Be		72 Hf	73 Ta	74 W	75 Re	76 Os	77 Ir	78 Pt	79 Au	80 Hg	81 Tl	82 Pb	83 Bi	84 Po	85 At	86 Rn
87 Fr	88 Ra		104 Rf	105 Db	106 Sg	107 Bh	108 Hs	109 Mt	110 Ds	111 Rg	112 Cn						
57 La	58 Ce	59 Nd	60 Nd	61 Pm	62 Sm	63 Eu	64 Gd	65 Tb	66 Dy	67 Ho	68 Er	69 Tm	70 Yb	71 Lu			
89 Ac	90 Th	91 Pa	92 U	93 Np	94 Pu	95 Am	96 Cm	97 Bk	98 Cf	99 Es	100 Fm	101 Md	102 No	103 Lr			

图 3.4: 金属, 类金属与非金属元素

那么能带论如何给出电导率的计算公式? 仍然是:

$$\sigma = \frac{ne^2\tau}{m}$$

与上一个费米气模型区别在于一点, 就是这里的 m 不再是电子的**裸质量**(bare mass), 而是由于与周期性势场相互作用, 电子的色散关系(能量与波矢关系)被修正后给出的**有效质量**(effective mass). 即使对于金属, 这个质量与真实电子质量相差好几倍的现象也是十分普遍的.

⁴思考: 氢是否具有金属性? 参考https://en.wikipedia.org/wiki/Metallic_hydrogen.

3.1.4 惯性, 阻尼与回复力

在量子力学还未诞生的时期, 为了解释导体的导电, 绝缘体的介电现象, 并适用于任意交变电磁场在介质中的传播, 色散, 吸收与散射. 洛伦兹提出著名的洛伦兹模型(Lorentz model) 作为德鲁特模型的补充. 在这儿电子的惯性被重视, 电子与晶格的碰撞被简化, 可能的原子实对电荷的束缚被简化为线性回复力. 也就是洛伦兹用唯象的谐振子类比来解释电子在外场下的行为:

$$m\ddot{\mathbf{r}} + \gamma\dot{\mathbf{r}} + k\mathbf{r} = -e\mathbf{E}e^{i\omega t}$$

这个模型的解和它对应的各种特性我们将在光学教材里详细讲解. 在这里我们仅仅讨论导体中的自由电子, 故 $k = 0$. 与导体导电相关的两个因素为:

一是阻尼系数 γ . 在直流电场下平衡时稳定的电子速度为:

$$\dot{\mathbf{r}} = -\frac{e}{\gamma}\mathbf{E}$$

对比更加本质的德鲁特等导电模型的电导公式, 容易发现这个唯象系数与电导率和基本参量关系为:

$$\gamma = \frac{ne^2}{\sigma} = \frac{m}{\tau}$$

于是我们可以写出在导体上突然加一个不随时间变化的匀强电场后电子的运动的微分方程:

$$m\ddot{\mathbf{r}} + \frac{m}{\tau}\dot{\mathbf{r}} = -e\mathbf{E}$$

它的解为:

$$\dot{\mathbf{r}} = -\frac{e\mathbf{E}\tau}{m}(1 - e^{-\frac{t}{\tau}})$$

我们十分自然地发现定义电子碰撞的弛豫时间恰好与导体对电场响应的弛豫时间相吻合. 这一点稍微需要一些讨论和修正. 见下.

第二点我们来讨论表示电子惯性的质量 m . 正是它导致了衰减因子 $e^{-\gamma t/m}$, 从而造成电路对电场响应的弛豫. 但如果讨论电路的弛豫, 或者说电流的惯性, 有一点是被我们过高地估计了, 便是自感现象. 电荷的加速运动形成变化的电流, 从而与之相伴的是变化的磁场. 这个变化的磁场反过来给电流一个反向的作用力. 由于该力正比于加速度, 故等价于增加了载流子的质量. 这里的弛豫时间对应直流暂态电路中的 $\tau' = L/R$. 如果一定要理解为电子的惯性质量, 那么一般比电子的裸质量要大好几个数量级 $m^* \gg m$. 从而一般有电路弛豫时间 $\tau = m^*/\gamma$ 不等于微观碰撞弛豫时间.

纯粹的导体是否有回复力项? 显然在以上电子运动方程中不含这一项. 但注意到如果有回复力项, 即使外场为零电子也会做振动. 真实情况会发生这样的现象吗? 答案是会, 即等离子体振荡(plasma oscillation). 原因是电子相对于原子实的位移实际上在金属的表面累积了电荷分布, 从而在内部激发了电场. 由于频率很高, 一般要到紫外波段的频率, 我们可以完全忽略阻尼的影响. 当电子有位移 \mathbf{r} 时将形成极化强度⁵:

$$\mathbf{P} = -ner$$

我们考虑块状金属板⁶的集体等离子体振荡, 那么由于金属表面的累积电荷在金属内部形成的电场为:

$$\mathbf{E} = \mathbf{E}_{in} = -\frac{\mathbf{P}}{\varepsilon_0} = \frac{ner}{\varepsilon_0}$$

代入原方程:

$$m^*\ddot{\mathbf{r}} = -e\mathbf{E} = -\frac{ne^2\mathbf{r}}{\varepsilon_0}$$

⁵显然此时“极化强度”描述的非极化电荷, 而是自由电荷.

⁶的确, 取不同形状的物体这个频率会有所不同.

从而得到谐振子方程. 振荡频率为:

$$\omega_p = \sqrt{\frac{ne^2}{\varepsilon_0 m^*}} ; \quad \sigma = \varepsilon_0 \omega_p^2 \tau$$

最后我们考虑低频电磁波 (内部实际场强) 下金属的电子行为, 利用光学中讨论过的洛伦兹模型, 我们计算金属的复电容率, 最后得到:

$$\varepsilon_r(\omega) = 1 - \frac{\omega_p^2 \tau^2}{1 + \omega^2 \tau^2} ; \quad \sigma(\omega) = \frac{\sigma}{1 + \omega^2 \tau^2}$$

一般 $\omega_p \tau \gg 1$, 故低频下金属的介电常数实际上是绝对值很大的负数. 电导率随着频率的增加而减小.

3.1.5 稳恒电流与形成条件

作为电荷的流动, 电流密度与电荷密度间满足电荷守恒定律:

$$\nabla \cdot \mathbf{j} + \frac{\partial \rho}{\partial t} = 0$$

而所谓**稳恒电流**(steady current) 指的是一定区域内导电物质形成的电荷, 电场, 电流分布. 其中电流的分布不能随时间变化, 自然的结论是电荷密度分布也不能随时间变化. 否则将产生随时间变化的电场, 电场造成电流的变化. 即上式必有:

$$\frac{\partial \rho}{\partial t} = 0 ; \quad \nabla \cdot \mathbf{j} = 0$$

即在稳恒电流中电流密度是一定没有散度的, 无头无尾的闭合曲线. 稳恒电流必须是**直流**(direct current) 而非**交流**(alternating current).

容易发现, 在稳恒电流中微观欧姆定律写成以下形式是不完整的:

$$\mathbf{j} = \sigma \mathbf{E}$$

在很多情况下, 上式并没有问题. \mathbf{E} 被理解为由电荷分布 ρ 形成的静电场⁷, 它是一定可以引入电势的:

$$\mathbf{E} = -\nabla \varphi$$

那么如果承认以上形式的微观欧姆定律就会引起矛盾: 电流线是环状闭合的, 电场线必须与电流线同向平行, 但电场线又不能是闭合的, 否则与静电场环路定理, 电势的可定义性矛盾. 实际上, 微观形式的欧姆定律必须被写成:

$$\mathbf{j} = \sigma (\mathbf{E} + \mathbf{K})$$

上式 \mathbf{K} 代表**非静电力场**(non-electrostatic field) 对电子产生的作用. 定义方式与电场量纲一致, 都是单位正电荷的受力. 这么写的不同之处就在于, 把驱动电流的力分成了两部分: 静电场 \mathbf{E} 是保守的, 可以引入势的, 在一个回路中不做功的, 不吞吐能量的场 (注意电流本身就会发热, 这不是场造成的而是电荷定向流动碰撞晶格造成的). 但非静电力就是源源不断做功驱动回路中电流流动的, 非保守的, 不能引入势来描述的场. 虽然不能引入电势来描述非保守力, 但可以用**电动势**(electromotive force) 来描述非静电力:

$$\mathcal{E}_\gamma = \int_{A \xrightarrow{\gamma} B} \mathbf{K} \cdot d\mathbf{r} ; \quad \mathcal{E}_\Gamma = \oint \mathbf{K} \cdot d\mathbf{r}$$

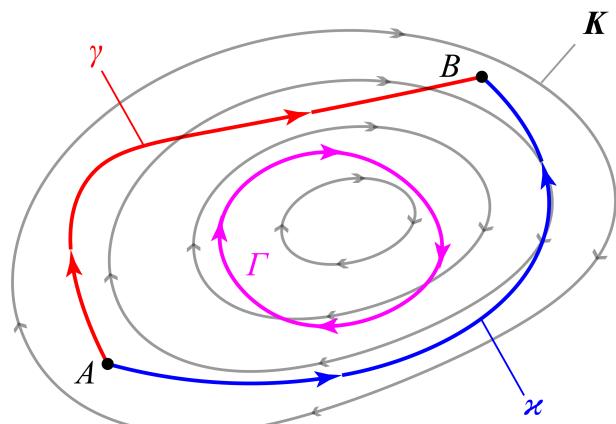


图 3.5: 非静电力场

⁷实际上, 不包括涡旋电场, 因为它被归于之后引入的非静电力. 换句话说, 这里的 \mathbf{E} 不再由静止电荷的受力 \mathbf{F}/q 去定义, 而是单纯根据静止电荷分布产生平方反比的电场叠加去定义.

不像对静电场的积分那样仅仅取决于端点,电动势与积分路径一一对应.回路电动势一般来说不为零,从而相同端点的两个路径电动势也不一定相等:

$$\mathcal{E}_\Gamma \neq 0 ; \quad \mathcal{E}_\gamma \neq \mathcal{E}_x$$

非静电力既然能够作用在电子上,其形式一般就是电磁力,万有引力,或是惯性力.磁场力对应**动生电动势**(motional emf),而由于变化的磁场激发的涡旋电场部分的电场力对应**感生电动势**(transformer emf).电子会受到引力,1967年Schiff指出电子会因为引力的原因聚集到金属的底部.而后Dessler提出了严格的考虑晶格在重力场下压缩与不均匀情况下的电子平衡情况.引力是很弱的保守力,对金属内部电荷体系的影响虽然很小但绝不是不可观测.而与引力类似的惯性力则可以人工制造出很大的数值,而且可以使它非保守.试想加速旋转一个线圈,那么在随晶体加速的参考系里电子受到的角向惯性力就和受到一个电场力造成的效果没有本质的区别.

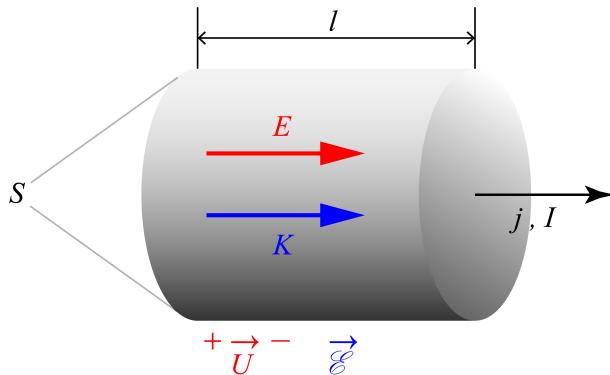


图 3.6: 宏观欧姆定律

还有一些电动势一般无法用非静电力描述.因为在一个回路中某些部分的确有能量的输入,但不是以作用在某些具体的电荷上的力的形式.而是更加广义的一些力作用在一些流上.例如温差电偶利用两种物质的电子的热扩散性质差异,从高温端吸取热量而驱动电流.化学反应更是源源不断地产生新的物质,利用反应的自发性而驱动物质做定向的移动形成电流.在这些抽象的场合,我们采用某段电路中移动单位电荷,外界向体系注入的能量来定义电动势:

$$\mathcal{E} = \frac{dW}{dQ}$$

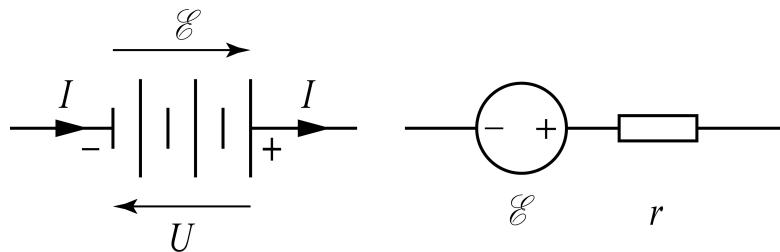


图 3.7: 电池的欧姆定律

最后,从微观到宏观,我们考虑一段长为 l ,面积为 S ,电阻率为 ρ 的导体上的欧姆定律:

$$E + K = \rho j$$

等式两边同时乘以 l ,把电流密度写成 I/S .得到:

$$U + \mathcal{E} = IR ; \quad R = \rho \frac{l}{S}$$

一般叫做部分电路的欧姆定律.而如果考虑**电池**(battery)的两端,正常工作时一般正极电压更高,故约定电动势与电压取相反的方向,电流则与电动势方向一致,却与电压方向反向:

$$U = \mathcal{E} - Ir$$

为一段含源电路的欧姆定律.此时我们把电池等效为理想电压源和一个定值电阻 r 的串联.电阻即为电池的**内阻**(internal resistance).

3.2 电路与电路方程

具体到直流电路中, 我们一般提供以下理想的电路元件:

名称	电阻	恒压源	恒流源	接地
符号	R	$\begin{array}{c} + \\ \mathcal{E} \\ - \end{array}$	$\begin{array}{c} + \\ \mathcal{I} \\ - \end{array}$	$\begin{array}{c} \parallel \\ \perp \end{array}$

图 3.8: 常见电路元件

实际电路总是由这些元件和导线构成, 导线和元件的内阻均被用实际电阻显式表示. 接地符号可以被视作特殊的导线. 一般与真实的“地面”没有任何联系, 电路中一般“地”指的是公共的零电势参考点. 电路中如果仅仅有一个接地符号, 就把接地的点的电势认为是零. 如果出现了多个接地符号, 就用导线把这些点连接到同一个点, 并把这个点电势视作零.

任何一个元件, 或者二端的部分电路, 为了研究的统一性与方便, 我们规定两端的电压的正方向与流过的电流的正方向都取“相关方向”, 沿着这个方向, 元件或部分电路表现的像一个电阻, 或者宽泛一点说, 像一个用电器. 即电流就是沿这个方向流动, 而且电压也沿这个方向降低. 比如在如图3.9所示的左侧大电池为右侧小电池充电的过程中, 右侧电路对上下两端口以向下为电压电流的相关参考方向, 则电压和电流分别为:

$$U = +2.8V, \quad I = +0.2A$$

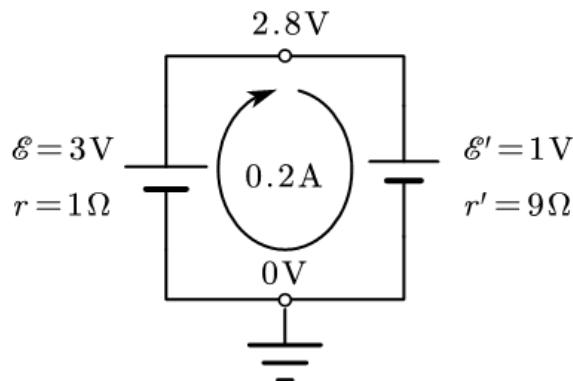


图 3.9: 相关方向

但如果对左侧电路依然取向下为电压电流的相关参考方向, 则电压和电流分别为:

$$U = +2.8V, \quad I = -0.2A$$

可以发现, 如果一个元件或部分电路按照相关方向计算的电压电流同时是正或者负, 这个元件就是类用电器的, 整体效果是消耗电能, 而如果一正一负, 就是类电源的, 整体效果是产生电能.

在具体分析复杂电路时, 顺次串联的各个部分电路相关方向一般取为同向; 讨论单个回路时, 各个部分的相关方向也是按照某种环绕方向确定的; 多个二端电路并联在两点之间也通常取各个支路为统一约定的从一个点到另一个点的相关方向.

对于之前的三种元件, 电阻由于对称性没有天生的电压电流相关方向. 无论如何取相关方向, 其电压 U , 电流 I 和其参数电阻 R 都一定要满足下式, 称作伏安特性(V-A characteristic):

$$U = IR$$

恒压源, 顾名思义, 总是能保持其正极电压比负极电压高 \mathcal{E} . 在正常工作状态时, 电流应当从正极流出, 使得其成为类电源的元件. 取此时电流方向为其相关方向 (具体问题中相关方向也可以与之相反), 则其特性为:

$$U = -\mathcal{E}, \quad I > 0$$

理想的恒压源同样能够工作在 $I < 0$ 的类用电器区域, 此时同样有 $U = -\mathcal{E}$. 但是对理想电压源短路是没有意义的, 一般恒压源所在支路一定需要串联电阻.

恒流源, 顾名思义, 总是能保持从负极向正极输出电流 \mathcal{I} . 在正常工作状态时, 正极电压应当比负极电压高, 使得其成为类电源的元件. 取此时电流方向为其相关方向 (具体问题中相关方向也可以与之相反), 则其特性为:

$$I = \mathcal{I}, \quad U < 0$$

理想的恒流源同样能够工作在 $U > 0$ 的类用电器区域, 此时同样有 $I = \mathcal{I}$. 但是对理想电压源开路是没有意义的, 一般恒流源所在支路一定需要并联电阻.

所以我们考虑的电路中, 如果出现恒压源或恒流源, 一般以以下带分压内阻或分流内阻的形式存在, 如果考虑整个元件的特性, 就给出了著名的输出特性(output characteristic):

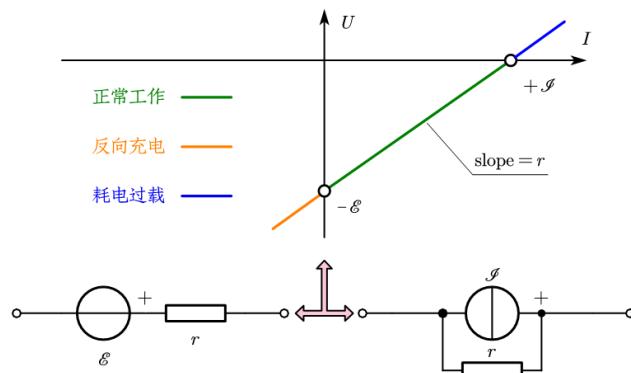


图 3.10: 源的统一输出特性
 $U = -\mathcal{E} + Ir = (I - \mathcal{I})r$

我们也从中发现了, 如果恒压源与恒流源满足内阻相等都是 r 且 $\mathcal{E} = \mathcal{I}r$, 那么它们的输出特性是完全一致的. 也就从功能上等价且可以相互替代, 这是我们之后小节要介绍的戴维南-诺尔顿定理的一个预兆. 但这主要只有理论价值, 实际情况分压内电阻总是要求很小, 分流内电阻总是要求很大, 这样在较大输出范围内才能凸显其恒压恒流的特性.

我们只有三种元件 (电阻, 带内阻的恒压源, 带内阻的恒流源), 而且已经成功给出了三种元件的特性, 它们实际上就是元件上取相关方向以后电压与电流之间的函数关系, 从而只要知道电流或电压中的一者就可以表示出另一者来. 但是如果涉及到电路的求解, 就需要把不同的元件联系起来. 此时还会涉及到两类方程. 分别称作基尔霍夫电流定律(KCL, Kirchhoff's current law) 与基尔霍夫电压定律(KVL, Kirchhoff's voltage law).

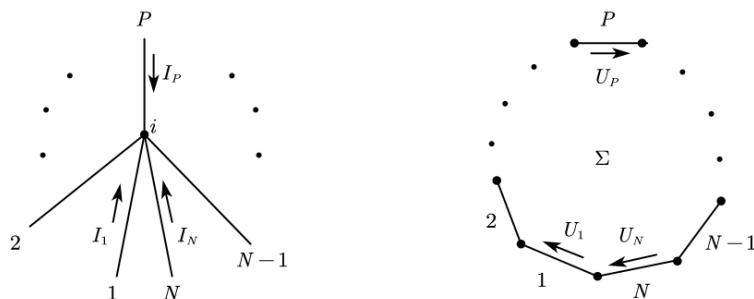


图 3.11: KCL 与 KVL

KCL 的含义是, 如果多个元件的一端连在一点. 那么取指向或背离公共节点的方向为所有元件的相关

方向,必有:

$$\sum_{P \in i} I_P = 0$$

我们约定小写拉丁字母 $i, j, k \dots$ 用来表示点,大写拉丁字母 $P, Q, R \dots$ 用来表示元件所在的支路(边),而从属符号 \in 用来表示结合关系, $P \in i$ 念做“边 P 以 i 为端点”.KCL 因此也称作**节点电流定律**.

KVL 的含义是,如果多个元件的首尾相连构成回路.那么取顺时针或逆时针方向为所有元件的相关方向,必有:

$$\sum_{P \in \Sigma} U_P = 0$$

我们约定大写希腊字母 $\Sigma, \Pi, \Phi \dots$ 用来表示回路,同样用从属符号 \in 用来表示结合关系, $P \in \Sigma$ 念做“边 P 在回路 Σ 上”.KVL 因此也称作**回路电压定律**.

而需要注意的是,KCL 与 KVL 都可以以元件电流或电压作为自变量.比如用电流做为 KVL 方程的自变量就构成了所谓的数电压法.下图所示的电路回路使用数电压法列顺时针方向的 KVL 方程就如右式:

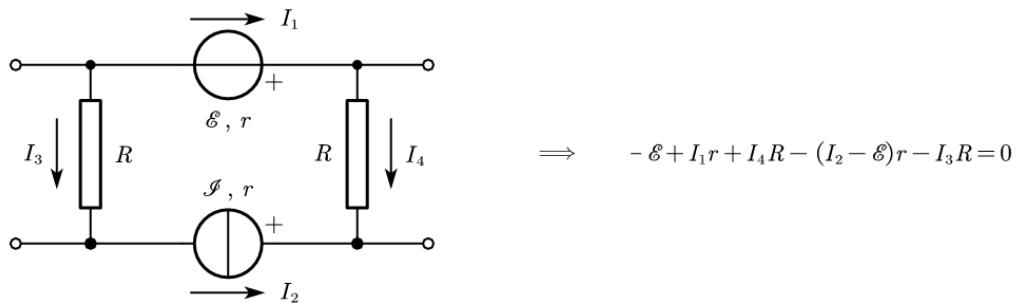


图 3.12: 回路: 数电压法

而节点电流也可以使用电压作为变量.另外还需要注意,利用数电压法我们可以得到 KVL 方程的一种非完整形式.由于我们对于节点总是喜欢定义其电势,那么如果节点 i, j 之间存在一条长的支路链,如下图所示,那么取 $i \rightarrow j$ 方向为参考方向,电势降低就满足:

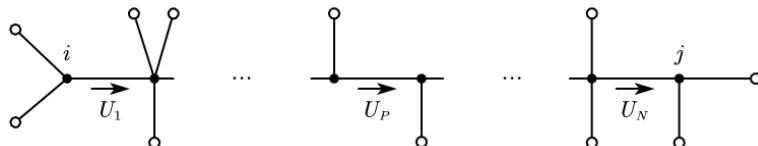


图 3.13: 路径: 数电压法

$$\varphi_i - \varphi_j = \sum_P U_p$$

尤其是对于以后用得上的星型电阻网络问题,以三端为例,KCL 或节点电流定律就可以在此时列为:

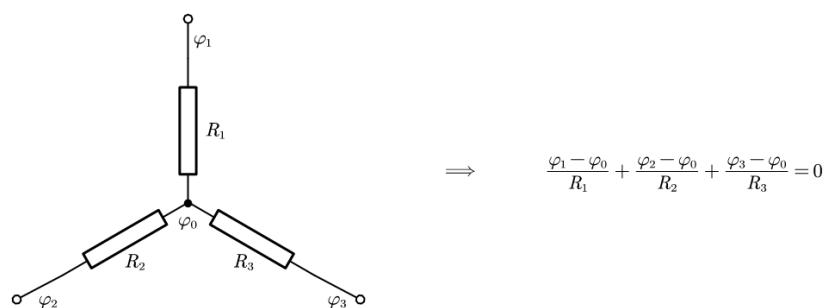


图 3.14: 节点电压法使用 KCL

一个典型的电路问题提法是如何? 它又具有哪些元素与性质? 解决方案又有哪些? 我们下面从基础开始细细分析.

3.3 电路分析基础

3.3.1 电路的整体结构与拓扑学结论

电路分析中的以下三个重要概念的术语与立体几何学中的术语是可等值替换的:

- 支路 (branch)-边 (edge)
- 节点 (node)-顶点 (vertex)
- 回路⁸(circuit)-面 (face)

它们分别是什么含义?

最基础的概念莫过于支路. 我们暂时采取最保守的做法, 认为每一个元件都占据一个独立的支路, 不同的支路之间可以是串联, 并联, 或者组合在一起的混联, 或者什么也不是 (典型的情况比如五条支路构成的桥式连接).

接下来定义节点的概念: 节点把支路连接在一起. 在定义这个概念时我们首次遇到了拓扑等价 (topological equivalence) 的概念. 众所周知, 如果是用真实导线把真实的元件连接, 那么连接方式不一定与电路图是完全对应的. 事实上就电路图本身也具有一定的变化范围, 比如对连接点的位移与导线的弯折具有等价性, 如下图左. 再比如说如果只允许电路图中出现三导线相交的情况时, 把四个元件的一端相连就有如下图右的三种方式, 当然, 如果在中间红色导线上接上其他元件或是用理想电流表去测量其电流, 那么这三个电路并不等价, 但是如果只考虑下面要阐述的电路问题的标准提法的求解那么三个电路并无区别.

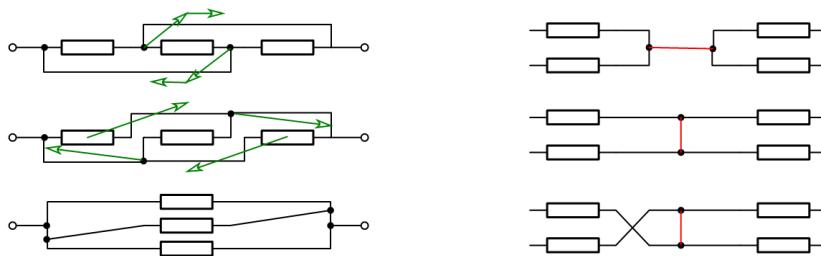


图 3.15: 拓扑等价变形

既然电路可以被拓扑变形, 那么节点的概念也就需要被重新被考量了. 事实上, 节点相当于彼此等势的导线延伸体. 导线无论如何连接与延伸, 只要是彼此连通的点就被视作属于同一个节点. 从这个视角上来看, 上面的左图从一开始实际上就只有左右两个节点, 右边的三个图中间全部都归为一个节点.

但是, 出于几何上的直观, 如果一个节点只被两个元件一端共用, 那么这两个元件就称作串联 (series connection) 的关系, 两个元件的串联或多个元件的顺次串联的情况下, 我们有理由把这一串元件全都视作在一条“强支路”上, 原来的支路被合并, 原来的中间的节点也就随着取消, 剩下的节点至少应该是三条或三条以上的支路汇合的“强节点”. 按照这种方式理解的强支路与强节点赋予电路以类似于多面体的拓扑结构.

要使得电路的结构完全类似于多面体, 还缺少至关重要的一环: 电路的回路就类似于多面体的面. 电路的回路是由多段支路首尾相连构成的不相交的环. 为了说明清楚回路分析中的重要一

⁸注意真正与面形成对应关系的应该是之后要引入的网孔的概念

个因素: 回路基. 我们还需要引入可平面化图(planar graph)的概念. 我们知道, 图是用来表示事物的连接关系的重要数学工具, 在这个意义上就无所谓图所在空间的维数的说法. 比如以右图表示的著名的三对三图(three for three graph)中, 有三户家庭分别都需要从水厂, 气厂, 电厂输送过来的资源, 那么三户家庭与三个工厂看作节点, 输送管道看作支路, 这就构成了 $K_{3,3}$ 图. 当然任何维度空间中的网络, 多面体的边与顶点自然也构成了一些图. 但是原则上抽象为一个图以后就失去了任何抽象前的维度信息. 这样就可以定义平面图(plane graph): 它是被画在一个平面上的, 同时支路与支路没有相交的图. 右边画的 $K_{3,3}$ 图就不是一个平面图. 事实上, 它也不是一个可平面化图: 可平面化图指的是虽然这个图可能以某种方式画在平面上支路会相交, 但是总能找到使边不相交的办法. 而著名的库拉托夫斯基定理⁹(Kuratowski's theorem) 指出:

任何不可平面化的图, 将串联支路合并后, 必然包含一个五阶完全图¹⁰或三对三图为子图.

我们下面只研究平面图, 但除了网孔概念以外, 其余性质都是所有电路中的图共有的.

在平面图中网孔(mesh)就是一个比较自然的概念了: 平面图的每一个格子就是一个网孔. 比如下图所示的三维立方体就产生了一个八节点十二支路的图. 它是可以平面化的, 按照右边的方式平面化以后, 就会导致变成 5 个网孔的平面图.

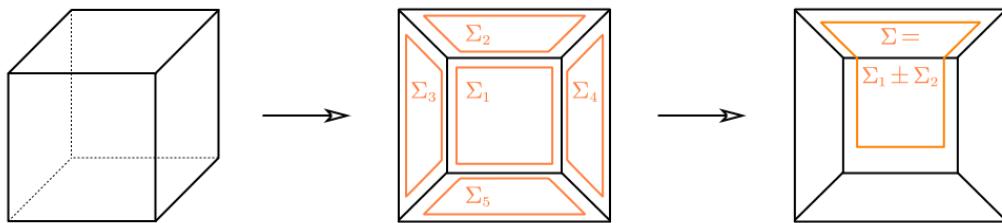


图 3.17: 立方体的平面化与网孔合并为复杂回路

可以发现, 网孔不过是特殊的一组回路. 这些回路内部的平面区域已经不再有其他支路. 如果网孔的个数为 F , 那么整个二维平面就被这些网孔回路分为 $F + 1$ 个区域, 每个网孔内部是一个区域, 而所有网孔外还有一个大的区域.

网孔作为一组特殊回路的特殊之处就在于, 网孔是一组自然的回路基(circuit basis). 所谓回路独立就是指它们对应的回路电压方程独立, 所谓回路基就是指所有可能回路的回路电压方程中选取的一组可以用来表示¹¹所有其他方程的独立方程对应的独立回路. 可以想见, 回路的取法是远多于网孔的个数的, 比如可以把上图两个相邻的网孔 Σ_1, Σ_2 合并为较大的回路 Σ^{12} . 所以回路基的取法并不是唯一的, 原则上存在很多种独立回路的取法.

⁹直到 1930 年才被严格表述与证明.

¹⁰即五个点彼此连接

¹¹通过方程的线性组合

¹²从集合论的角度来讲, 这种合并的本质是对称差(symmetric difference), 它被定义为:

$$A \Delta B = A \cup B - A \cap B$$

对于回路的合并这是很自然的: 上式代表保留各自的边, 去除公共边. 但是对于回路合并对称差符号写为 \pm 符号是非常贴切的, 因为容易理解对称差两次必然相互抵消:

$$(A \Delta B) \Delta B = A$$

从而从合并后的网孔去再次对称差合并前的网孔, 又会回到合并的另一个网孔, 像极了加减法:

$$\Sigma_1 \pm \Sigma_2 = \Sigma \Leftrightarrow \Sigma \pm \Sigma_2 = \Sigma_1$$

图论给出,对于任何一个电路图,都存在一个特征的数字: **回路秩**¹³(circuit rank),对于平面图它就是网孔的个数,对于非平面图可以用之后给出的办法计算. 我们取一组最大数目的互相独立的回路做回路基时,其个数总是会等于这个回路秩. 利用这个特性,即使是复杂的非平面电阻网络,只要我们总是避免选取的回路不独立,选够回路秩那么多回路时,就得到了一组回路基.

联系支路,节点,网孔(可换为回路基,下同)三者的最核心的定理,莫过于欧拉早在 1758 年就发现并证明的定理:如果一个三维空间中的没有洞的多面体(比如像甜甜圈那样,但是表面被磨为多个平面的多面体就是有洞的多面体)的顶点数为 V ,边数为 E ,而面数为 F ,那么必然有:

$$E = V + F - 2$$

但是如果变成电路问题,由于多面体的面数与它所对应的被平面化的图的网孔数(或者图的回路秩,下同)总是差一. 故对于电路问题,节点数 V ,支路数 E 和网孔数 F 三者满足的关系式应当为:

$$E = V + F - 1$$

默认以上结果,就意味着我们终于具备了开始研究电路问题的数学基础. 接下的第一步,就是严格的表述一个电路问题.

- 已知量:

- 首先,必须给定的是电路的拓扑,即已知电路由 V 个节点, E 条支路构成,还应当能判断每一个节点 i 与每一条支路 P 之间是否有从属关系 $i \in, \notin P$.
- 其次,每一条支路上的元件种类: 电阻, 恒压源或恒流源. 与对应的参数 $R, \mathcal{E}, \mathcal{I}, r$ 也应当事先给定.

- 待求量:

- 为每一条支路指定好方向以后,支路上的电流和电压都是待求的. 由于电压电流通过支路元件特性相关联,一共 E 个未知数.
- 节点上的电势通常也是我们关心的,此时一般要约定好某节点接地视作零电势,其他节点的电势均是相对它的电势,都是待求量,一共 $V - 1$ 个未知数.
- 平面图形成了 F 个网孔. 图论可以证明,因为满足 KCL 方程,合理的某种支路电流分布总存在唯一的某种**网孔电流**(mesh current) 的叠加与之等价. 故也会产生对应的网孔电流待求量,一共 F 个未知数.

网孔(回路)电流法是一种用来表示电路中的电流分布的快捷方法,其优点是自动会满足 KCL 方程. 且未知数个数相对较少. 其做法是: 对于每个网孔 Σ 约定一个顺时针或逆时针的参考方向,并给定一个网孔电流 I_Σ . 那么对于每一条支路 P ,再为支路约定参考方向,则其上的电流等于:

$$I_P = \sum_{\Sigma \in P} \pm I_\Sigma$$

再一次地 \in 表示结合关系, $\Sigma \in P$ 念做“ Σ 以 P 为边”. 需要注意,如果 Σ 就是取网孔,那么与任何一条支路发生结合关系的网孔最多只能有两个,即上式最多只有两项,而在图的边缘的支路则仅仅属于一个网孔. 但是,如果取更广义地取回路基,则与同一条支路发生结合的回路个数可以是任意个. 电流的符号需要根据回路电流的参考方向与后面对支路求和时的方向是否相同来决定,同向取正,反向取负. 下图对同一个平面电路图,用网孔电流法和回路电流法分别表示两条代表性支路上的电流:

¹³最早被基尔霍夫(G. Kirchhoff)在 1847 年引入,当时称作**圈数**(cyclomatic number).

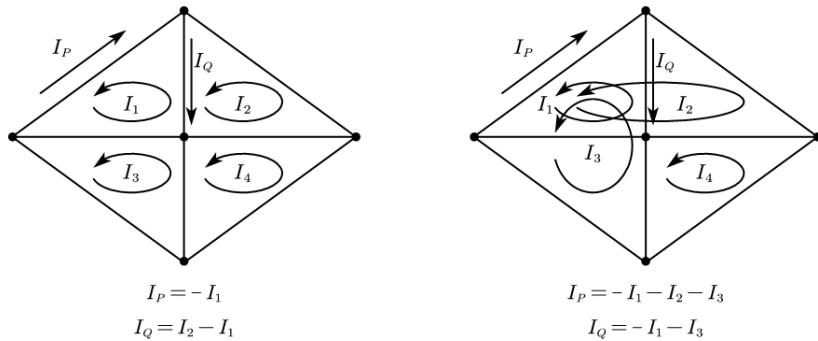


图 3.18: 网孔 (回路) 电流法

最后,通过已知量求未知量的方法,应当是各类 KCL 和 KVL 方程.它们应当如何列,列多少个?这就形成了电路问题的求解套路:

3.3.2 电路问题的求解套路

电路问题的求解有三种套路:支路电流法,节点电势法和网孔 (回路) 电流法.无论采用哪一种方法,无非是对之前梳理的三类待求量中,选取一类集中加以解决:解决的方案无非是设它们为未知数,并找到与未知数的个数恰好相等的,包含且仅包含这些未知数的 KCL 或 KVL 方程,这必然得到线性方程组.解出它们并根据它们得到另外两类待求量.下面逐一论述之.

1. 支路电流法

如果为每一条支路约定好参考方向,并设其上电流为 I_P ,则一共有 E 个未知数.

我们需要列出几乎全套的 KCL,对每一个节点 i ,这个方程写作:

$$\sum_{P \in i} \pm I_P = 0$$

正负号约定是:如果支路电流参考方向为指向该节点,则取正,若背离该节点,则取负.但是所有的 KCL 放在一起并不独立,这是因为所有的 KCL 方程求和后,每一条支路的 I_P 恰好在两端点处的 KCL 中出现两次,一正一负,故最后得到零等于零.从而需要去掉一个任意节点 j ,即可以列出 $V - 1$ 个方程:

$$\forall i \neq j, \quad \sum_{P \in i} \pm I_P = 0$$

对于每一个网孔还要列一个 KVL 方程,为了方便与统一见,我们把所有带内阻的恒流源都变成等效的带内阻恒压源,而电阻元件也视作 $\mathcal{E} = 0$ 的特殊情况,下同.这样 KVL 方程中的电压就可以用电流表示,一共是 F 个方程:

$$\forall \Sigma, \quad \sum_{P \in \Sigma} \pm U_P = 0$$

$$U_P = -\mathcal{E}_P + I_P R_P$$

正负号约定是,如果约定回路的环绕方向与支路参考方向一致则取正,相反则取负.

整理,对以上 KVL 方程把未知数置于左侧,常数(非齐次项)置于右侧,得到:

$$\forall i \neq j, \quad \sum_{P \in i} \pm I_P = 0$$

$$\forall \Sigma, \quad \sum_{P \in \Sigma} \pm R_P I_P = \sum_{P \in \Sigma} \pm \mathcal{E}_P$$

一共是 $V - 1 + F$ 个方程, 恰好等于未知数的个数 E , 因为:

$$E = V - 1 + F$$

2. 节点电势法

如果为某个节点 j 接地, 即约定其电势为 $\varphi_j = 0$, 对其他节点 $i \neq j$ 设电势 φ_i , 则一共有 $V - 1$ 个未知数. 此时 KVL 方程已经给不出有意义的方程, 只能列 KCL 方程. 任何一条支路 P , 如果两端点为 i, k , 那么如果约定 i 到 k 方向为其参考方向, 有:

$$U_P = \varphi_i - \varphi_k = -\mathcal{E}_P + I_P R_P \Rightarrow I_P = \frac{\varphi_i}{R_P} - \frac{\varphi_k}{R_P} + \frac{\mathcal{E}_P}{R_P}$$

那么根据 KCL 方程, 所有流入 k 节点的电流应当和为零:

$$\left(\sum_{P \in k} \frac{1}{R_P} \right) \varphi_k - \sum_{i \in P \in k} \left(\frac{1}{R_P} \varphi_i \right) = \sum_{P \in k} \frac{\mathcal{E}_P}{R_P}$$

其中 $i \in P \in k$ 念做 i, jP . 同样的, 这 V 个方程并不独立, 应当再去掉一个 k 的方程, 由于 $\varphi_j = 0$ 不必解, 不妨就去掉这一个, 并使左侧不出现 φ_j :

$$\forall k \neq j, \quad \left(\sum_{P \in k} \frac{1}{R_P} \right) \varphi_k - \sum_{\substack{i \in P \in k \\ i \neq j}} \left(\frac{1}{R_P} \varphi_i \right) = \sum_{P \in k} \frac{\mathcal{E}_P}{R_P}$$

这样方程数也是 $V - 1$ 个, 等于未知数个数 V_1 .

3. 网孔电流法

如果对每个网孔, 约定参考方向后设每个网孔电流为 I_Σ , 则一共有 F 个未知数. 此时 KCL 方程已经给不出有意义的方程, 只能列 KVL 方程. 任何一条支路 $P \in \Sigma$, 如果按照所在网孔 Σ 的参考方向, 有:

$$U_P = -\mathcal{E}_P + I_P R_P = -\mathcal{E}_P + R_P \sum_{\Pi \in P} \pm I_\Pi$$

求和中恰好有一项就是原来的 I_Σ , 它取正号, 而其他的回路 Π 应当根据与回路 Σ 在公共边 P 上的方向是否一致来决定正负号, 同向为正, 反向为负. 将上式求和得到 KVL:

$$\forall \Sigma, \quad \left(\sum_{P \in \Sigma} R_P \right) I_\Sigma + \sum_{P \in \Sigma} \left(R_P \sum_{\Pi \in P \in \Sigma} \pm I_\Pi \right) = \sum_{P \in \Sigma} \mathcal{E}_P$$

其中 $\Pi \in P \in \Sigma$ 念做“ Π 和 Σ 共用边 P ”. 这样方程数也是 F 个, 等于未知数个数 F .

一方面, 三种方法的未知数与方程个数分别是 $E, V - 1, F$, 而 $E = V - 1 + F$, 所以显然后两种方法的未知数和方程数都是比第一种少的, 在实际应用过程中更加省事, 原则上我们应当尽可能地使用后两种方法来解决实际问题. 但是另一方面, 第一种方法得到的方程形式上更简洁而易于分析, 所以理论推导的时候也青睐基础的支路电流法.

3.4 电路分析方法

- 叠加原理: 用上一节介绍的任何一种方法求解方程组时, 右边的非齐次项都可以进行拆分与叠加. 直接推论是, 如果电路只含电压源, 那么保留各个单个电压源, 其他电压源置零, 解出的各支路电流的和即每个电压源都存在时的电流.
- 任何二端不含源电阻网络都可以等效为一个线性电阻.

- 戴维南-诺尔顿原理: 任何二端含源网络如果把各电压源电流源置零之后形成的不含源电阻网络的等效电阻网络电阻为 r , 其开路电压 U , 其短路电流 I . 那么这个电路等效于一个电压源或电流源, 内阻即 r , 电压或电流即 U, I .
- 特勒根原理: 单一电路中若各个支路选取相关方向, 则:

$$\sum_P U_P I_P = 0$$

如果两个电路的图一样, 即具有相同的节点, 支路, 网孔结构, 但是元件可以不一样. 那么选取各个支路为相关方向, 也有:

$$\sum_P U_{1P} I_{2P} = 0 = \sum_P U_{2P} I_{1P}$$

3.5 半导体

第四章 静磁学

4.1 电流与磁场

4.1.1 磁场地位与电流分布

人们对磁的认识一共经历了几个阶段:

1. 起: 远古-**奥斯特**(H. C. Ørsted). 磁现象被认为与**磁体**(magnet) 相关. 它被设想为由**磁荷**(magnetic charge) 造成.
2. 承: 奥斯特-**麦克斯韦**(J. C. Maxwell). 奥斯特发现了电流的磁效应, **法拉第**(M. Faraday) 建立了“场”的观念, 可以定量描述电流产生的磁场, 法拉第与**亨利**(J. Henry) 等人发现了电磁感应, **安培**(A.-M. Ampère) 的**分子电流假说**解释了物质磁性的本质也是电流产生磁场.
3. 转: 麦克斯韦-**狄拉克**(P. Dirac). 麦克斯韦和**洛伦兹**(H. A. Lorentz) 给出了电磁场的普遍理论, 即**电动力学**(electrodynamics)¹, 洛伦兹和**爱因斯坦**(A. Einstein) 等人将电动力学纳入相对论框架. **赫兹**(H. R. Hertz) 的实验有效地支持了电磁场论. 洛伦兹等人的经典电子论是微观粒子与电磁场之间相互作用的第一个完备理论, 微观粒子被陆续发现, 电子的**自旋**(spin) 与**磁矩**(magnetic moment) 被发现与研究.
4. 合: 狄拉克-现今. 狄拉克主张把电子量子化为场, 提出了**量子电动力学**(quantum electrodynamics), 预言了反物质的存在, **安德森**(C. D. Anderson) 实验室发现正电子. 狄拉克提出了量子化的**磁单极子**(magnetic monopole) 模型, 但至今未找到磁单极子.

现在, 我们对**磁场**(magnetic field) 概念的理解就可以用下图来概括:

如何证明磁场是电场的相对论效应? 设想原参考系中只有电场, 那么无论原参考系中的速度 v_x, v_y, v_z 如何, 受力都不取决于速度大小, 它是:

$$F_x = qE_x$$

$$F_y = qE_y$$

$$F_z = qE_z$$

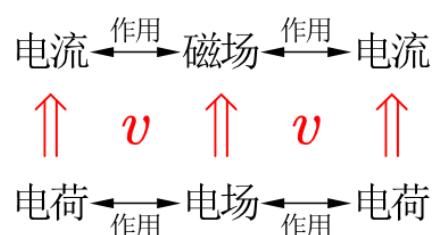


图 4.1: 磁场与电荷的关系

根据相对论的力变换公式:

$$F'_x = F_x - \frac{u}{c^2} \cdot \frac{v_y F_y + v_z F_z}{1 - \frac{u v_x}{c^2}} , \quad F'_y = \frac{F_y}{\gamma(1 - \frac{u v_x}{c^2})} , \quad F'_z = \frac{F_z}{\gamma(1 - \frac{u v_x}{c^2})}$$

代入原式, 并假设原参考系速度和变换速度相对 c 都是小量, 故新参考系速度可以使用经典的速度叠加:

$$v_x = v'_x + u , \quad v_y = v'_y , \quad v_z = v'_z$$

¹这个词语却是安培发明的.

近似到与 v'_x, v'_y, v'_z 一阶相关的领头项:

$$F'_x = qE_x - qv'_y \cdot \frac{uE_y}{c^2} - qv'_z \cdot \frac{uE_z}{c^2}$$

$$F'_y = qE_y + qv'_x \cdot \frac{uE_y}{c^2}$$

$$F'_z = qE_z + qv'_x \cdot \frac{uE_z}{c^2}$$

我们如果假设以下**洛伦兹力**(Lorentz force) 公式的成立性:

$$\mathbf{F} = q(\mathbf{E} + \mathbf{v} \times \mathbf{B})$$

它是磁场力分量 $\mathbf{F}_m = q\mathbf{v} \times \mathbf{B}$ 与电场力分量 $\mathbf{F}_e = q\mathbf{E}$ 的和. 则受力公式简化为:

$$F'_x = qE'_x + qv'_y B'_z - qv'_z B'_y$$

$$F'_y = qE'_y + qv'_z B'_x - qv'_x B'_z$$

$$F'_z = qE'_z + qv'_x B'_y - qv'_y B'_x$$

对比即知道:

$$B'_x = 0 \quad , \quad B'_y = \frac{uE_z}{c^2} \quad , \quad B'_z = -\frac{uE_x}{c^2}$$

这实际上:

$$\mathbf{B} = -\frac{\mathbf{u} \times \mathbf{E}}{c^2}$$

如果原参考系是静止点电荷产生的电场, 现在换参考系以后, 点电荷做速度为 $-\mathbf{u}$ 的匀速直线运动而产生以上磁场, 那么就能推理出, 如果点电荷以速度 \mathbf{v} 做匀速直线运动, 那么它产生的磁场必然为:

$$\mathbf{B} = \frac{1}{4\pi\epsilon_0 c^2} \frac{q\mathbf{v} \times \mathbf{e}_r}{r^2}$$

我们从中总结出:

- 只有运动的电荷才会受到磁场力.
- 只有运动的电荷才产生磁场.
- 但是运动与静止具有相对性, 所以电场和磁场的区别就像静止与运动的区别那样是相对的, 合称**电磁场**(electromagneto field).

我们还注意到来自理论推导过程中产生的重要公式:

$$\mathbf{B} = \frac{1}{4\pi\epsilon_0 c^2} \frac{q\mathbf{v} \times \mathbf{e}_r}{r^2}$$

$$\mathbf{F} = q\mathbf{v} \times \mathbf{B}$$

由于推导的近似性这两个式子似乎没有太多说服力. 但是历史上这两个式子是作为实验规律先被总结, 再有麦克斯韦的普遍电磁理论, 最后才是洛伦兹和爱因斯坦对其背后的相对论基础的研究.

4.1.2 毕奥-萨伐尔定律

强调一下本章研究的范围是重要的。我们暂时只研究静磁场(static magnetic field)，其场源为电流，且其分布不随时间，从而改变的磁场也不随时间改变。或者更简单的说，我们要研究的体系就是上一章建立的稳恒电流体系，它由不变的电荷产生不变的电场，不变的电场驱动不变的电流，现在增加了不变的电流产生不变的磁场这一要素。上一章的讨论告诉我们这就是要求电流分布是无散的：

$$\nabla \cdot \mathbf{j} = 0$$

如何把微观电荷的运动同宏观的电流 \mathbf{j} , I 建立联系？引言中我们指出来过，微观运动电荷产生的磁场正比于量：

$$\mathbf{C} = q\mathbf{v}$$

这个量被称作电流元(current element)。事实上作为微观电荷 q 通常是基本电荷 e 的量级，从而上述量是一个微元，即使取物质中一个很小的体积 dV ，其内部包含的所有电荷个数 dN 依然未达到宏观量级时，总的电流元依然是个小量，记做 $d\mathbf{C}$ 。我们知道电流密度的定义为：

$$\mathbf{j} = \rho\mathbf{v} = nq\mathbf{v}$$

那么体积产生的总的电流元：

$$d\mathbf{C} = dNq\mathbf{v} = nqvdV = \mathbf{j}dV = \rho dV \cdot \mathbf{v}$$

称为体电流元(volume current element)。同理，我们可以写出面电流元(surface current element)和线电流元(curve current element)：

$$d\mathbf{C} = i dA = \sigma v dA \quad , \quad d\mathbf{C} = I dl = \lambda v dl$$

注意上面线电流元的写法，由于我们不认为 I 是矢量，故把电流元的矢量性由线元 dl 来承载。这么做的好处之后就能体会到。我们发现电流元与电荷元形成了对应关系：

	微观单元	体元	面元	线元
磁	$q\mathbf{v}$	$\mathbf{j}dV$	$i dA$	$I dl$
电	q	ρdV	σdA	λdl

图 4.2: 电荷元与电流元对应

毕奥-萨伐尔定律(Biot-Savart law)是一个实验定律，它用于描述真空中两个孤立的电流元之间的相互作用力：

$$d\mathbf{F}_{12} = \frac{\mu_0}{4\pi} \frac{I_2 dl_2 \times (I_1 dl_1 \times \mathbf{e}_{12})}{r^2}$$

其中系数 $k_m = \frac{\mu_0}{4\pi} \approx 1 \times 10^{-7} \text{H/m}^2$ 为磁力常数，而 μ_0 称作真空磁导率(magnetic permeability of vacuum)。事实上：

$$k_e/k_m = c^2 \quad \text{或者} \quad c = \frac{1}{\sqrt{\epsilon_0 \mu_0}}$$

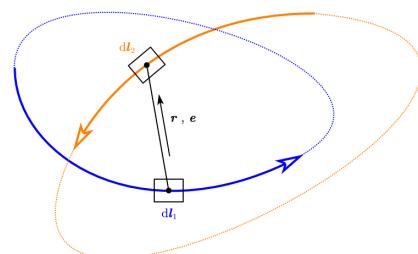


图 4.3: 电流体系间的相互作用力

²注意，2019 年 5 月 20 日之后，新的国际单位制标准下其值为 $1.000\ 000\ 000\ 82(20) \times 10^{-7} \text{H/m}$ ，不再有严格的 $\mu_0 := 4\pi \times 10^{-7} \text{H/m}$ 。

c 是真空中的光速. 这也揭示了电磁理论统一性背后的相对论本质. 我们这里还用到了电感的国际单位亨利, 而磁场的国际单位特斯拉(Tesla) 与它的关系为:

$$1\text{H} = 1\text{T} \cdot \text{m}^2/\text{A}$$

另一个常用的单位来自 CGS 单位制, 即高斯(Gauss), 以纪念这位伟大的数学家在物理学上同样瞩目的贡献. 高斯在历史上被广泛使用, 直至今天也如此, 例如地表地磁场大概就在 $0.25 - 0.60\text{G}$, 它与特斯拉的换算关系为:

$$1\text{T} = 10000\text{G}$$

根据场论的精神, 我们将上式拆解为电流产生场的定律和电流在场中受力的公式:

$$\text{d}\mathbf{F}_{12} = I_2 \text{d}\mathbf{l}_2 \times \text{d}\mathbf{B}_{12}, \quad \text{d}\mathbf{B}_{12} = \frac{\mu_0}{4\pi} \frac{I_1 \text{d}\mathbf{l}_1 \times \mathbf{e}_{12}}{r^2}$$

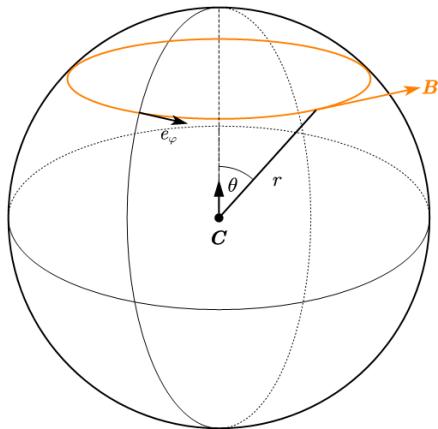


图 4.4: 电流元产生的磁场

如前所述, 毕奥-萨伐尔定律可以在相对论背景下由库仑定律推出, 故它继承了库仑定律的平方反比和叠加原理的特性. 但是它突出的特点是具有横向力(transverse force) 的特点: 力的方向与受力物体电流元垂直, 也与另一个只依赖于场源的磁场矢量垂直, 但这个矢量又与场源电流元的方向垂直. 两次叉乘这让力的计算变得比较复杂. 让我们以场源电流元 C 为中心, 其方向建立极坐标系, 那么磁场的大小与方向就可以从公式得到:

$$\mathbf{B} = \frac{\mu_0}{4\pi} \frac{C \sin \theta}{r^2} \mathbf{e}_\varphi$$

磁场符合叠加原理, 例如一个载流线圈, 源点坐标 \mathbf{r}' , 场点坐标 \mathbf{r} , 记 $\mathbf{R} = \mathbf{r} - \mathbf{r}'$, $\text{d}\mathbf{l} = \text{d}\mathbf{r}'$, 则其产生的磁场公式为:

$$\mathbf{B} = \oint \frac{\mu_0}{4\pi} \frac{I \text{d}\mathbf{l} \times \mathbf{e}_R}{R^2}$$

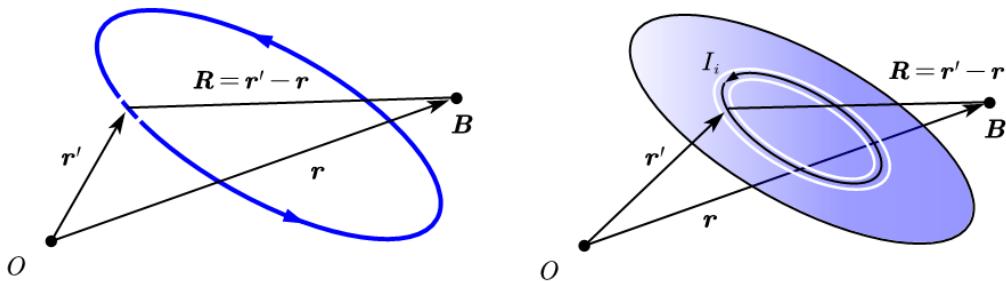


图 4.5: 连续体系产生磁场

载流线圈具有非常重要的地位, 任何体电流分布, 只要是有界区域内的稳恒电流, 电流线就必然构成闭合的回路. 那么就一定可以分解为电流圈. 如果用求和的极限表示积分但是暴力地省去极限符号, 此时磁场计算方法为:

$$\mathbf{B} = \int_{V'} \frac{\mu_0}{4\pi} \frac{\mathbf{j} \times \mathbf{e}_R}{R^2} \text{d}V = \sum_i \oint \frac{\mu_0}{4\pi} \frac{I_i \text{d}\mathbf{l}_i \times \mathbf{e}_R}{R^2}$$

而一个线圈或一个体电流分布在外磁场中的受力为:

$$\mathbf{F} = \oint I \text{d}\mathbf{l} \times \mathbf{B} \Rightarrow \mathbf{F} = \sum_i I_i \text{d}\mathbf{l}_i \times \mathbf{B}$$

在稳恒电流情况下, 讨论一个电流元产生的磁场是没有意义的, 因为空间的磁场必然是整个体系产生的磁场的叠加, 从而单个电流元的磁场并不具有可观测的意义, 除非电流可以独立的被操控以改变. 而讨论一个电流体系各个部分受到的作用力时, 尤其是一个电流线圈受到的作用力时, 区分内力和外力就变得尤其重要. 我们发现, 两个电流元之间的相互作用力之和:

$$d\mathbf{F}_{12} + d\mathbf{F}_{21} = \frac{\mu_0}{4\pi} \frac{I_2 d\mathbf{l}_2 \times (I_1 d\mathbf{l}_1 \times \mathbf{e}_{12})}{r^2} + \frac{\mu_0}{4\pi} \frac{I_1 d\mathbf{l}_1 \times (I_2 d\mathbf{l}_2 \times \mathbf{e}_{21})}{r^2}$$

根据三重矢积公式 $\mathbf{A} \times (\mathbf{B} \times \mathbf{C}) = (\mathbf{A} \cdot \mathbf{C})\mathbf{B} + (\mathbf{A} \cdot \mathbf{B})\mathbf{C}$, $\mathbf{e}_{12} = -\mathbf{e}_{21}$. 得到:

$$d\mathbf{F}_{12} + d\mathbf{F}_{21} = -\frac{\mu_0}{4\pi} \frac{(I_1 d\mathbf{l}_1 \times I_2 d\mathbf{l}_2) \times \mathbf{e}_{12}}{r^2}$$

一般来说不总为零, 可见朴素的牛顿第三定律在静磁场情况下失效了. 我们不得不把 $d\mathbf{F}_{12}$, $d\mathbf{F}_{21}$ 不再视作相互作用力, 而是需要通过场作为媒介. 事实上两个运动电荷之间的磁场力是在作为电场力的相对论修正, 从相对论的根本原理上, 相互作用传递速度的有限性导致力的传播被推迟, 原则上也不可能符合牛顿第三定律.

但是如果我们计算两个线圈之间的相互作用力之和:

$$\begin{aligned} \delta\mathbf{F} &= \oint_1 \oint_2 d\mathbf{F}_{12} + \oint_2 \oint_1 d\mathbf{F}_{21} \\ &= \frac{\mu_0 I_1 I_2}{4\pi} \left(\oint_1 \oint_2 \frac{(\mathbf{r}_2 - \mathbf{r}_1) \cdot d(\mathbf{r}_2 - \mathbf{r}_1)}{|\mathbf{r}_2 - \mathbf{r}_1|^3} d\mathbf{r}_1 + \oint_2 \oint_1 \frac{(\mathbf{r}_1 - \mathbf{r}_2) \cdot d(\mathbf{r}_1 - \mathbf{r}_2)}{|\mathbf{r}_1 - \mathbf{r}_2|^3} d\mathbf{r}_2 \right) \end{aligned}$$

但是:

$$\begin{aligned} \nabla_2 \frac{1}{|\mathbf{r}_2 - \mathbf{r}_1|} &= -\frac{\mathbf{r}_2 - \mathbf{r}_1}{|\mathbf{r}_2 - \mathbf{r}_1|^3} \quad \Rightarrow \quad \oint_2 \frac{(\mathbf{r}_2 - \mathbf{r}_1) \cdot d(\mathbf{r}_2 - \mathbf{r}_1)}{|\mathbf{r}_2 - \mathbf{r}_1|^3} = 0 \\ \nabla_1 \frac{1}{|\mathbf{r}_1 - \mathbf{r}_2|} &= -\frac{\mathbf{r}_1 - \mathbf{r}_2}{|\mathbf{r}_1 - \mathbf{r}_2|^3} \quad \Rightarrow \quad \oint_1 \frac{(\mathbf{r}_1 - \mathbf{r}_2) \cdot d(\mathbf{r}_1 - \mathbf{r}_2)}{|\mathbf{r}_1 - \mathbf{r}_2|^3} = 0 \end{aligned}$$

从而得到:

$$\delta\mathbf{F} = 0$$

当两个环重合时, 这也代表“自己给自己的力”是零. 即使是磁场给线圈的力, 也是线圈自己产生的磁场而无法对自己产生力, 这个结果其实也是空间平移对称性(而不是牛顿第三定律)的必然结果. 而空间旋转对称性则给出, 任意选取 O 点计算各个电流元受力对 O 的力矩, 总和必然为零. 或者说两个线圈之间的作用力与作用力矩之间必然符合牛顿第三定律. 其证明在此不再赘述. 这样我们就证明了此前给出的理论公式与牛顿第三定律之间是自洽的, 无散的电流体系下牛顿第三定律依然成立, 尽管表面上电流元的毕奥萨伐尔定律与牛顿第三定律有冲突.

更有甚者, 我们指出, 根据麦克斯韦方程的协变性, 即使是微观电荷做相对论性运动, 只要是稳恒电流体系, 那么毕奥萨伐尔定律可以作为一个推论而具有普适性. 而对于非稳恒体系, 后面我们将会证明, 磁场此时并不能简单地由电流确定而依赖于电场的变化率, 甚至它成为了自由存在的物质形式而摆脱了电荷造成任何约束.

4.2 两个定律与矢势

4.2.1 磁场的环路定律

静磁场的环路定律(circuital law), 又称安培定律(Ampère's law)指出, 对于一个可定向曲面 A 和它的闭合右手边界 ∂A , 有:

$$\oint_{\partial A} \mathbf{B} \cdot d\mathbf{l} = \mu_0 \int_A \mathbf{j} \cdot d\mathbf{A}$$

我们用毕奥-萨伐尔定律来证明环路定律。首先，上式中的 A 和其边界上的场点位矢取为 r ，则 $dl = dr$ ，而场源电流分解为电流环 I_i 。将电流环上坐标选为 r' 。那么以上积分可按照叠加原理分解为：

$$\Gamma = \sum_i \Gamma_i = \sum_i \oint_{\partial A} \mathbf{B}_i \cdot dl$$

而每一项环量就可以用场点和源点的两次环积分表示，将 e_R 记做 e ：

$$\begin{aligned} \Gamma_i &= \oint_r \mathbf{B}_i \cdot dr \\ &= \oint_r \left(\oint_{r'} \frac{\mu_0 I_i dr' \times e_R}{4\pi R^2} \right) \cdot dr \\ &= \frac{\mu_0 I_i}{4\pi} \oint_r \oint_{r'} \frac{(dr \times dr') \cdot e_R}{R^2} \end{aligned}$$

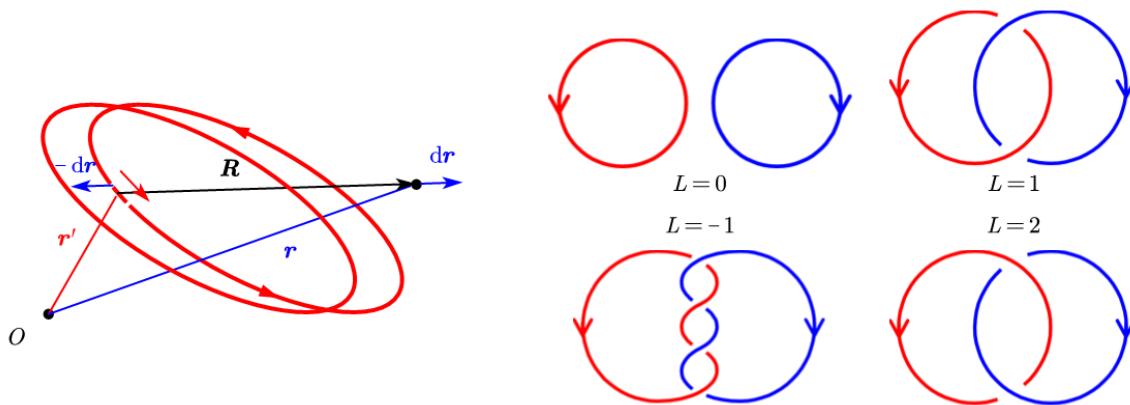


图 4.6: 环绕数

如上图所示，对上面的积分做如下诠释：当从场点 r 去看线圈时，线圈对场点产生了一个立体角 Ω 。当观察点移动 dr 时，考察立体角的改变 $d\Omega$ 。可以等效地把线圈反向移动 $-dr$ 。图中所示情况立体角将减小，算法为：

$$-d\Omega = \int \frac{dA \cdot (-e_R)}{R^2} = \oint_{r'} \frac{(-dr \times dr') \cdot (-e_R)}{R^2}$$

下面考虑以源电流圈（4.6中红圈）为边界的那个面，场点形成的安培回路（4.6中蓝圈）与这个面的交点个数为 $n_1 + n_2$ ，其中 n_1 代表蓝圈正向穿过面时方向与红圈形成右手规则的交点个数， n_2 代表蓝圈正向穿过面时方向与红圈形成左手规则的交点个数。于是再定义环绕数(linking number)：

$$\ell = n_1 - n_2$$

这是一个拓扑学不变量，它只取决于两个回路的整理性质，与面的取法无关，而且具有对称性（红对蓝与蓝对红的环绕数一致）。

从而我们注意到，如果环绕数恰好是一，那么从面上出发， Ω 从一个接近 2π 的大小一直在减小直到变为零，然后变成负的，直到回到面上直到变成接近 -2π ，虽然初末位置立体角不一定都等于 2π ，但是一定有：

$$\oint_r d\Omega = -4\pi$$

相似的，如果考虑一个环绕数为 L 的安培回路，就有：

$$\oint_r d\Omega = -4\pi\ell$$

将此前的立体角微分代入, 得到:

$$\oint_{\mathbf{r}} \oint_{\mathbf{r}'} \frac{(\mathbf{dr} \times \mathbf{dr}') \cdot \mathbf{e}_R}{R^2} = 4\pi\ell$$

最后代入此前的磁场环量, 得到:

$$\Gamma_i = \mu_0 \ell_i I_i$$

最后我们总结出:

$$\oint_{\partial A} \mathbf{B} \cdot d\mathbf{l} = \mu_0 \sum_i \ell_i I_i$$

现在做极限化处理, 注意到求和式实际上就是统计了所有以右手规则为正向通过面的电流, 统一用电流密度表示就变成了自带正负的点乘的积分:

$$\oint_{\partial A} \mathbf{B} \cdot d\mathbf{l} = \mu_0 \int_A \mathbf{j} \cdot d\mathbf{A}$$

此即环路定律, 由数学上的开尔文-斯托克斯定理, 我们可以把上式写出局域的微分形式:

$$\nabla \times \mathbf{B} = \mu_0 \mathbf{j}$$

4.2.2 矢势与磁场的高斯定律

对电流元产生的磁场 $d\mathbf{B} = \frac{\mu_0}{4\pi} \frac{\mathbf{j} dV \times \mathbf{e}_r}{r^2}$ 引入以下矢量场是方便的:

$$\mathbf{A} = \frac{\mu_0}{4\pi} \frac{\mathbf{j} dV}{r}$$

这是因为如果计算球坐标下的旋度, 考虑到 \mathbf{A} 分解到 r, θ 方向分别为:

$$A_r = \frac{\mu_0}{4\pi} \frac{j dV}{r} \cos \theta, \quad A_\theta = -\frac{\mu_0}{4\pi} \frac{j dV}{r} \sin \theta$$

那么考虑一个回路³ S : $(r, \theta) - (r + dr, \theta) - (r + dr, \theta + d\theta) - (r, \theta + d\theta) - (r, \theta)$, 利用旋度的定义式:

$$\mathbf{n} \cdot \nabla \times \mathbf{A} = \lim_{S \rightarrow 0} \oint_{\partial S} \mathbf{A} \cdot d\mathbf{l} / S$$

得到其旋度:

$$\nabla \times \mathbf{A} = \frac{1}{r} \left(\frac{\partial r A_\theta}{\partial r} - \frac{\partial A_r}{\partial \theta} \right) \mathbf{e}_\varphi$$

代入便得到以下关系式:

$$\mathbf{B} = \frac{\mu_0}{4\pi} \frac{j dV}{r^2} \sin \theta \mathbf{e}_\varphi = \nabla \times \mathbf{A}$$

\mathbf{A} 这个矢量称作**磁矢势**(magnetic vector potential). 由于叠加原理, 一个电流分布体系的磁矢势和磁场应当为:

$$\mathbf{A} = \int_V \frac{\mu_0}{4\pi} \frac{\mathbf{j} dV}{r}, \quad \mathbf{B} = \int_V \frac{\mu_0}{4\pi} \frac{\mathbf{j} dV \times \mathbf{e}_r}{r^2}$$

同样地将满足:

$$\mathbf{B} = \nabla \times \mathbf{A}$$

以上旋度关系写成积分形式意味着一种计算磁通量的特殊方法:

$$\Phi = \oint_{\partial S} \mathbf{A} \cdot d\mathbf{l} = \int_S \mathbf{B} \cdot d\mathbf{S}$$

³为了避免与磁矢势符号重复, 暂时用 S 表示面元.

对于无边界的闭合曲面，上式直接证明了磁场**高斯定律**(Gauss's law):

$$\begin{aligned}\partial\partial V = \emptyset \Rightarrow \oint_{\partial V} \mathbf{B} \cdot d\mathbf{S} &= \oint_{\partial\partial V} \mathbf{A} \cdot dl = 0 \\ \Rightarrow \nabla \cdot \mathbf{B} &= 0\end{aligned}$$

我们如果计算此前电流元产生的磁矢势的散度:

$$\nabla \cdot \mathbf{A} = \nabla \cdot \left(\frac{\mu_0 j dV}{4\pi r} \mathbf{e}_z \right) = \frac{\mu_0 j dV}{4\pi} \frac{\partial}{\partial z} \left(\frac{1}{r} \right) = \frac{\mu_0 j dV \cos \theta}{4\pi r^2} = \frac{\mu_0 j dV \cdot \mathbf{e}_r}{4\pi r^2}$$

又发现对电流元得到非零的结果。但是如果考虑线圈:

$$\nabla \cdot \mathbf{A} = \oint \frac{\mu_0 I dl \cdot \mathbf{e}_R}{4\pi R^2} = \frac{\mu_0 I}{4\pi} \oint \frac{\mathbf{e}_R}{R^2} \cdot dr' = \frac{\mu_0 I}{4\pi} \oint \nabla_{r'} \left(\frac{1}{R} \right) \cdot d\mathbf{R} = \frac{\mu_0 I}{4\pi} \oint d \left(\frac{1}{R} \right) = 0$$

从而对于任意稳恒电流体系，通过此前积分定义的磁矢势也是无散的。最后将磁矢势与磁场关系和磁场与电流关系联立，注意到:

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

即三重矢积公式。我们把梯度算符 ∇ 直接看作了一个矢量。上式右侧第一项由于磁矢势无散故为零，从而得到:

$$\nabla^2 \mathbf{A} = -\mu_0 \mathbf{A}$$

电场中的式子 $\nabla^2 \varphi = -\rho/\varepsilon_0$ 与它相呼应。

可以设想存在可以像点电荷产生电场那样的方式产生磁场的物质，称作**磁单极子**(magnetic monopole)，从而可以产生不为零的净**磁荷**(magnetic charge)，事实证明这样的物质至今都未曾找到，从而相伴的磁场形式也仅仅存在与理论中。但是，总磁荷为零的体系：**磁偶极子**(magnetic dipole)，不仅仅是纯粹的理论模型，后面可以发现线圈产生的磁场与磁偶极子是相似的。

我们记得电势可以被视作电荷在电场中运动的势能有关的物理量，即 $E_p = q\varphi$ 。那么电荷在磁场中运动时，磁矢势代表什么物理概念呢？如果空间中只有纯的磁场，那么显然磁场力不做功：

$$P = \frac{dW}{dt} = \mathbf{F}_m \cdot \mathbf{v} = (q\mathbf{v} \times \mathbf{B}) \cdot \mathbf{v} = 0$$

所以电荷一定是动能不变的。但是即使如此，我们也不认为磁场力是一个保守力，这是因为保守力的定义应该是：能够用势能的梯度来表示的作用力。我们写出一个项：

$$\boldsymbol{\pi} = q\mathbf{A}$$

这个项不再具有能量的量纲而是动量的量纲。我们知道电场除磁场的量纲为速度，而两者与对应的矢势标势都差一个长度量纲，故 $q\mathbf{A}$ 量纲等于用 $q\varphi$ 的能量量纲除速度，即动量量纲。这个动量称为**势动量**(potential momentum)。而带电粒子的旧动量 $\mathbf{p} = mv$ 就称作**机械动量**(kinetic momentum)，它是运动造成的，而势动量是位置 \mathbf{r} 造成的 $\boldsymbol{\pi}(\mathbf{r})$ ，是与磁场相互作用引起的动量。两者之和：

$$\mathbf{P} = \mathbf{p} + \boldsymbol{\pi}$$

称作**正则动量**(canonical momentum)。不像在电场中的电荷运动机械能(动能+势能)守恒，这个和也是一般不守恒的：

$$\begin{aligned}\dot{\mathbf{P}} &= \dot{\mathbf{p}} + \dot{\boldsymbol{\pi}} \\ &= \mathbf{F} + q\dot{\mathbf{A}} \\ &= q\mathbf{v} \times (\nabla \times \mathbf{A}) + q\mathbf{v} \cdot \nabla \mathbf{A} \\ &= -\nabla(-q\mathbf{v} \cdot \mathbf{A})\end{aligned}$$

最后一步把 ∇ 视作矢量使用了三重矢积公式. 注意到瞬时的 v 不被视作变量而视作常数而从偏导数符号内外移动不引起改变. 我们发现瞬时的正则动量导数似乎恰好等于某个势能的负梯度:

$$V = -q\mathbf{v} \cdot \mathbf{A}$$

这被称作**广义势**(generalized potential). 它含有速度, 从而每一个时刻有着完全不同的形式. 但不管怎样, 粒子在磁场中的运动的确是写成了一个动量导数等于势的梯度的形式(拉格朗日方程形式). 即使这样, 我们依然不认为磁场力是一个保守力.

我们来看一个具体的例子. 假设空间中的电流分布产生了一个随 y 方向变化的 x 方向的磁矢势分布:

$$\mathbf{A} = (f(y), 0, 0) = f(y)\mathbf{e}_x$$

那么可以验证这个式子的确符合:

$$\nabla \cdot \mathbf{A} = \frac{\partial f(y)}{\partial x} = 0$$

而产生它的电流和对应的磁场就是:

$$\mathbf{j} = -\nabla^2 \mathbf{A}/\mu_0 = -\frac{1}{\mu_0} \left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} \right) f(y)\mathbf{e}_x = -\frac{f''(y)}{\mu_0}\mathbf{e}_x$$

$$\mathbf{B} = \nabla \times \mathbf{A} = -\frac{\partial f(y)}{\partial y}\mathbf{e}_z = -f'(y)\mathbf{e}_z$$

这表示一个垂直 $x-y$ 平面的磁场. 那么粒子在某一瞬时的广义势:

$$V = -q\mathbf{v} \cdot \mathbf{A} = -qv_x f(y)$$

其瞬时梯度就是:

$$-\nabla V = -qv_x f'(y)\mathbf{e}_y$$

而粒子的正则动量是:

$$\mathbf{P} = (mv_x + qf(y))\mathbf{e}_x + mv_y\mathbf{e}_x$$

根据此前的论述:

$$\dot{\mathbf{P}} = -\nabla V$$

从而我们发现, 在 x 方向粒子的**广义动量守恒**(concervation of generalized momentum):

$$mv_x + qf(y) = \text{Const.}$$

而在 y 方向有一动力学方程:

$$ma_y = -qv_x f'(y)$$

读者可以用牛顿力学同样的论证以上两个式子的成立性.

4.3 电流体系

研究更进一步的磁场与电流相互作用的模型前, 熟悉一些常见的电流体系的结论是大有裨益的, 这些包括:

4.3.1 磁偶极子

同样类似于对静电学中一个分布在原点附近的电流体系在远场产生的电势进行展开的思想, 我们此时运用到静磁学情形: 计算一个分布在原点附近的电流体系在远场产生的磁矢势:

$$\mathbf{A} = \int_{V'} \frac{\mu_0}{4\pi} \frac{j dV'}{|\mathbf{r} - \mathbf{r}'|}$$

利用分解为线圈可以简化以上计算:

$$\mathbf{A} = \sum_i \oint \frac{\mu_0}{4\pi} \frac{I_i d\mathbf{l}_i}{|\mathbf{r} - \mathbf{r}_i|}$$

依然利用静电学中的展开公式:

$$\begin{aligned} \frac{1}{|\mathbf{r} - \mathbf{r}_i|} &= [(\mathbf{r} - \mathbf{r}_i)^2]^{-\frac{1}{2}} = (\mathbf{r}^2 - 2\mathbf{r}_i \cdot \mathbf{r} + \mathbf{r}_i^2)^{-\frac{1}{2}} \\ &\approx (\mathbf{r}^2 - 2\mathbf{r}_i \cdot \mathbf{r})^{-\frac{1}{2}} = \frac{1}{r} \left(1 - 2\frac{\mathbf{r}_i \cdot \mathbf{r}}{r^2}\right)^{-\frac{1}{2}} \\ &\approx \frac{1}{r} \left(1 + \frac{\mathbf{r}_i \cdot \mathbf{r}}{r^2}\right) \\ &= \frac{1}{r} + \frac{\mathbf{e}_r}{r^2} \cdot \mathbf{r}_i \end{aligned}$$

代入上式, 注意到 $d\mathbf{l}_i = d\mathbf{r}_i$ 得到磁矢势展开的前两项:

$$\mathbf{A} \approx \frac{\mu_0}{4\pi r} \sum_i I_i \oint d\mathbf{r}_i + \frac{\mu_0}{4\pi r^2} \sum_i I_i \oint (\mathbf{e}_r \cdot \mathbf{r}_i) d\mathbf{r}_i$$

第一项由于对回路积分必然为零. 第二个积分中需要注意 \mathbf{e}_r 照理应当是可以提取到积分号外面的常矢量, 但是这样积分内部的 $\mathbf{r}_i d\mathbf{r}_i$ 构成了一个不直观的张量. 我们采用一个更加简单的做法: 常矢量点乘法(method of dot product by a constant vector). 任意取一个常矢量 \mathbf{v} 来点乘积分结果:

$$\mathbf{v} \cdot \oint (\mathbf{e}_r \cdot \mathbf{r}_i) d\mathbf{r}_i = \oint (\mathbf{e}_r \cdot \mathbf{r}_i) (\mathbf{v} \cdot d\mathbf{r}_i) = \mathbf{e}_r \cdot \oint (\mathbf{v} \cdot d\mathbf{r}_i) \mathbf{r}_i$$

这样倒是可以实现 \mathbf{e}_r 移到积分外, 但代价是另一个类似的 \mathbf{v} 又到了积分内. 其实无济于事. 但是注意到中间的积分其实完全可以使用开尔文-斯托克斯定理:

$$\oint (\mathbf{e}_r \cdot \mathbf{r}_i) (\mathbf{v} \cdot d\mathbf{r}_i) = \oint [(\mathbf{e}_r \cdot \mathbf{r}_i) \mathbf{v}] \cdot d\mathbf{r}_i = \int_{S_i} \nabla \times [(\mathbf{e}_r \cdot \mathbf{r}_i) \mathbf{v}] \cdot d\mathbf{S}_i$$

根据叉乘的定义:

$$\nabla \times [(\mathbf{e}_r \cdot \mathbf{r}_i) \mathbf{v}] = \begin{vmatrix} \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ (\mathbf{e}_r \cdot \mathbf{r}_i) v_x & (\mathbf{e}_r \cdot \mathbf{r}_i) v_y & (\mathbf{e}_r \cdot \mathbf{r}_i) v_z \\ \mathbf{e}_x & \mathbf{e}_y & \mathbf{e}_z \end{vmatrix}$$

其中 $\mathbf{e}_r = (e_{rx}, e_{ry}, e_{rz})$ 和 (v_x, v_y, v_z) 都是常的. 故求偏导只发生在点乘式 $\mathbf{e}_r \cdot \mathbf{r}_i$ 中的一项上给出不为零的系数. 即:

$$\nabla \times [(\mathbf{e}_r \cdot \mathbf{r}_i) \mathbf{v}] = \begin{vmatrix} e_{rx} & e_{ry} & e_{rz} \\ v_x & v_y & v_z \\ \mathbf{e}_x & \mathbf{e}_y & \mathbf{e}_z \end{vmatrix} = \mathbf{e}_r \times \mathbf{v}$$

最终得到:

$$\mathbf{v} \cdot \oint (\mathbf{e}_r \cdot \mathbf{r}_i) d\mathbf{r}_i = \int_{S_i} (\mathbf{e}_r \times \mathbf{v}) \cdot d\mathbf{S}_i = \mathbf{v} \cdot \left[\left(\int_{S_i} d\mathbf{S}_i \right) \times \mathbf{e}_r \right] = \mathbf{v} \cdot (\mathbf{S}_i \times \mathbf{e}_r)$$

由于工具 \mathbf{v} 的任意性, 现在就可以看出以下关系:

$$\oint (\mathbf{e}_r \cdot \mathbf{r}_i) d\mathbf{r}_i = \mathbf{S}_i \times \mathbf{e}_r$$

再代入此前的磁矢势, 得到最终结果:

$$\mathbf{A} \approx \frac{\mu_0}{4\pi r^2} \left(\sum_i I_i \mathbf{S}_i \right) \times \mathbf{e}_r$$

在讨论最终结果前我们对上式出现的一种现象需要做进一步的解释. 如图4.8所示, 一个线圈对应的面矢量 \mathbf{S} 的标准定义实际上对应着以下线圈上的路径积分:

$$\mathbf{S} = \frac{1}{2} \oint \mathbf{r} \times d\mathbf{r}$$

这样定义也绝非完美, 我们自然觉得, 一个线圈对应的面矢量的确定应该是与坐标原点选取无关的, 但上式在表面上有关. 如果原点位移了 \mathbf{R} , 两个积分的差为:

$$\begin{aligned} \mathbf{S}' &= \frac{1}{2} \oint (\mathbf{r} - \mathbf{R}) \times d(\mathbf{r} - \mathbf{R}) \\ &= \frac{1}{2} \oint (\mathbf{r} - \mathbf{R}) \times d\mathbf{r} \\ &= \mathbf{S} - \frac{1}{2} \mathbf{R} \times \oint d\mathbf{r} \\ &= \mathbf{S} \end{aligned}$$

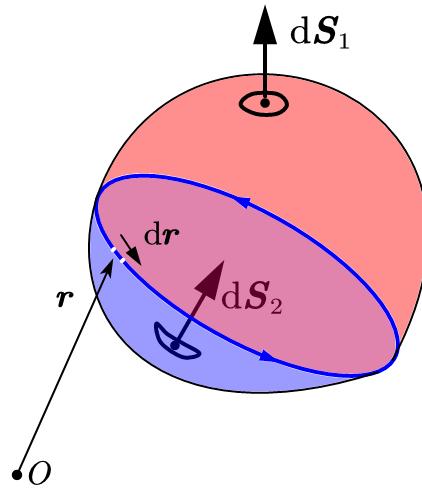


图 4.8: 线圈与面矢量

故以上定义的确不依赖于原点的选取. 但是我们要问, 为什么上式就表示以线圈为边界的面矢量呢? 而这样定义又有何好处? 事实上, 如过按照通常的方式来定义一个边界所确定的面的面矢量, 那么就面临着一个奇怪的问题: 不同的面可能共用了同一个边界, 甚至会出现莫比乌斯带那样, 一个简单的闭合曲线却以一个反常的不可定向的面的边界的形式出现了, 我们甚至可以在一个圆环作为边界的情况下构造出莫比乌斯带型的不可定向曲面来, 将莫比乌斯带的边界扯成一个圆环即可. 为了避免这种情况, 我们最好是用曲线积分来定义面矢量, 这样对于复杂曲线(比如自己打结的曲线, 比如螺旋线)也不必受难以确定以其为边界的曲面之苦. 事实上, 如果对于4.8那样的曲线较为简单的情况, 选取以其为边界的可定向曲面, 依然把右手定则确定的微元面矢量积分定义为整个面的面矢量, 那么对于4.8中两个不同的面:

$$\mathbf{S}_1 = \int_{S_1} d\mathbf{S}_1, \quad \mathbf{S}_2 = \int_{S_1} d\mathbf{S}_2$$

我们自然也要问, 以上两个积分是否会相等? 事实上如果把两个积分相减, 就得到了在一个体的闭合表面上的面元积分:

$$\mathbf{S}_1 - \mathbf{S}_2 = \oint_{\partial V} d\mathbf{S}$$

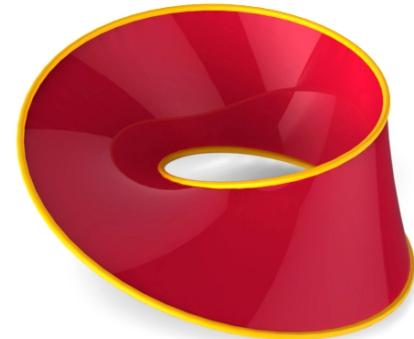


图 4.9: 莫比乌斯带

要想得到上式的值, 我们又再一次使用常矢量点乘法:

$$\mathbf{v} \cdot (\mathbf{S}_1 - \mathbf{S}_2) = \oint_{\partial V} \mathbf{v} \cdot d\mathbf{S}$$

很显然, 这表示一个常矢量场在闭合面上的通量, 如匀强电场通过一个闭合面的电通量代数和, 从而一定是零. 由 \mathbf{v} 的任意性, 总有:

$$\mathbf{S}_1 - \mathbf{S}_2 = 0 \quad , \quad \mathbf{S}_1 = \mathbf{S}_2$$

从而这个定义也是不依赖于曲面的选取的, 而为了把之前标准定义的曲线积分和对面元的积分联系起来, 我们再使用常矢量点乘法:

$$\mathbf{v} \cdot \frac{1}{2} \oint_{\partial S} \mathbf{r} \times d\mathbf{r} = \frac{1}{2} \oint_{\partial S} (\mathbf{v} \times \mathbf{r}) \cdot d\mathbf{r} = \frac{1}{2} \int_S \nabla \times (\mathbf{v} \times \mathbf{r}) \cdot d\mathbf{S}$$

利用 \mathbf{v} 为常矢量时可以在 ∇ 前后移动的特性, 得到:

$$\nabla \times (\mathbf{v} \times \mathbf{r}) = (\nabla \cdot \mathbf{r})\mathbf{v} - (\mathbf{v} \cdot \nabla)\mathbf{r} = 3\mathbf{v} - \mathbf{v} = 2\mathbf{v}$$

代入上式:

$$\mathbf{v} \cdot \frac{1}{2} \oint_{\partial S} \mathbf{r} \times d\mathbf{r} = \frac{1}{2} \int_S 2\mathbf{v} \cdot d\mathbf{S} = \mathbf{v} \cdot \int_S d\mathbf{S}$$

从而就证明的确这个曲线积分就是面矢量:

$$\frac{1}{2} \oint_{\partial S} \mathbf{r} \times d\mathbf{r} = \int_S d\mathbf{S} = \mathbf{S}$$

关于用曲线积分得到总的面矢量 \mathbf{S} 还有一种几何理解方式, 再仔细观察图4.8就会发现, 项:

$$\frac{1}{2} \mathbf{r} \times d\mathbf{r}$$

实际上就是原点 O 出发向曲线上的一段 $d\mathbf{r}$ 连线形成的三角形面元, 它恰好是从 O 出发向曲线上的所有点引射线形成的一个面元, 故积分就得到整个锥面的总面矢量且与曲线环绕方向恰好符合右手规则, 它也是一种特殊的以曲线为边界的面的总面矢量.

终于我们回到此前电流体系形成的磁矢势问题, 现在我们就定义一个载流 I_i 的线圈形成的磁矩(magnetic moment) 为:

$$\mathbf{m}_i = I_i \mathbf{S}_i$$

而这个电流体系形成的总磁矩就是:

$$\mathbf{m} = \sum_i \mathbf{m}_i = \sum_i I_i \mathbf{S}_i$$

那么在远场的磁矢势的领头项就是由磁矩产生的平方反比项:

$$\mathbf{A} = \frac{\mu_0}{4\pi} \frac{\mathbf{m} \times \mathbf{e}_r}{r^2}$$

这个式子其实非常类似电流元 C 产生的磁场公式:

$$\mathbf{B} = \frac{\mu_0}{4\pi} \frac{\mathbf{C} \times \mathbf{e}_r}{r^2}$$

但是要计算磁矩产生的磁场, 应对以上形式的磁矢势来求旋度, 计算可得:

$$B_r = \frac{\mu_0}{4\pi} \frac{2m \cos \theta}{r^3} \quad , \quad B_\theta = \frac{\mu_0}{4\pi} \frac{m \sin \theta}{r^3}$$

也许令人惊讶的是, 其关于 r, θ 变化的形式居然与电偶极子产生的电场的公式一致⁴.

磁矩产生的磁场具有非常重要的地位, 第一是因为任何情况下这都会是一个电流体系激发的磁场的领头项, 因为磁单极子不存在, 类似点电荷产生平方反比的电场那样的磁荷项, 由于磁高斯定律, 永远是不可能的. 磁矩的计算方式上面已经给出, 但是我们依然要指出, 将求和再还原为对提电流的积分, 可以得到磁矩的体积分算法:

$$\mathbf{m} = \frac{1}{2} \int_V \mathbf{r} \times \mathbf{j} dV$$

而最简单的磁矩模型就是一个平面圆环, 其电流为 I , 半径为 R . 那么磁矩就是:

$$m = I\pi R^2$$

而考虑 $I \rightarrow \infty, S \rightarrow 0$ 的模型, 就构成了点磁偶极子(point dipole)的情况. 此时体系产生的磁场将严格符合我们得出的三个计算式:

$$\mathbf{A} = \frac{\mu_0}{4\pi} \frac{\mathbf{m} \times \mathbf{e}_r}{r^2}, \quad B_r = \frac{\mu_0}{4\pi} \frac{2m \cos \theta}{r^3}, \quad B_\theta = \frac{\mu_0}{4\pi} \frac{m \sin \theta}{r^3}$$

第二, 任何稳恒电流体系可以向线圈分解, 而任何线圈又可以向磁偶极子来分解. 我们把载有电流 I 的线圈所在的可定向曲面分解为无穷多无穷小的网格, 并让每一个网格都载有同样的方向的电流 I , 那么网格形成的总电流分布实际上就是原线圈的电流分布, 因为有公共边的两个网格在公共边上的电流一定恰好反向而产生的电流互相抵消, 最后只剩下最外面的一圈边界上电流被保留下来. 从而不管是产生的场还是在外场中的受力均可以用于等效原来的线圈.

磁偶极子的磁势能, 受力与力矩:

$$V = -\mathbf{m} \cdot \mathbf{B}, \quad \mathbf{F} = \mathbf{m} \cdot \nabla \mathbf{B}, \quad \mathbf{M} = \mathbf{m} \times \mathbf{B}$$

4.3.2 磁化强度

- 磁化强度(magnetization):

$$\mathbf{M} = n\mathbf{m}$$

- 磁化强度造成的体电流与面电流分布:

$$\mathbf{j}_M = \nabla \times \mathbf{M}, \quad \mathbf{i}_M = \mathbf{M} \times \mathbf{n}$$

4.3.3 若干对称体系的磁场

- 无限长载流直线:

$$\mathbf{A} = -\frac{\mu_0 I}{2\pi} \ln r, \quad \mathbf{B} = \frac{\mu_0 I}{2\pi r} \mathbf{e}_\varphi$$

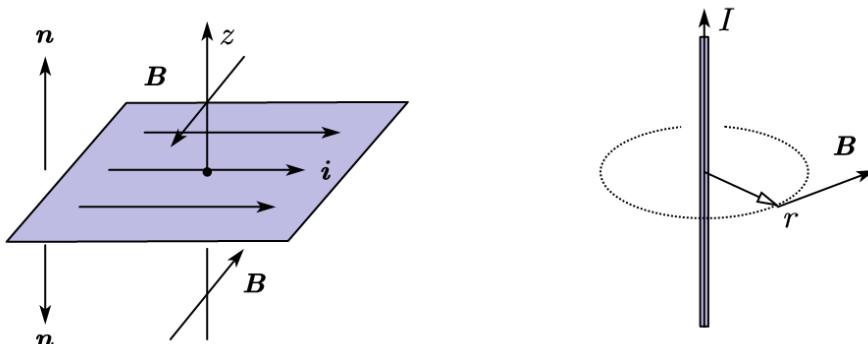


图 4.10: 无限大平面与无限长直线

⁴其实这个形式是拉普拉斯算符, 或者说, 背后的对称性导致的必然结果

- 无限大载流平面:

$$\mathbf{B} = \frac{1}{2}\mu_0 \mathbf{i} \times \mathbf{n}$$

- 圆环: 半径为 R 的标准载流线圈轴线上磁场:

$$\mathbf{B} = \frac{\mu_0 I R}{2(R^2 + z^2)^{\frac{3}{2}}} \mathbf{e}_z$$

该处的径向场强可以由磁高斯定理确定:

$$\frac{1}{2} \frac{dB_r}{dr} + \frac{dB_z}{dz} = 0 \quad \Rightarrow \quad B_r \approx -\frac{3\mu_0 I R z r}{(R^2 + z^2)^{\frac{5}{2}}}$$

- 无限长密绕螺线管: 无论其横截面形状, 总是有:

$$B = \mu_0 i = \mu_0 n I = \frac{\mu_0 N I}{l}$$

4.4 磁介质与磁能

4.4.1 微观角度理解磁化

- 分子的固有磁矩: 电子自旋磁矩 + 电子轨道磁矩.
- 如果分子具有固有磁矩, 那么按照热力学规律发生取向磁化: 磁矩倾向于取与外磁场相同的方向, 宏观上产生顺磁性(paramagnetism). 类似于电介质的极化, 微观上顺磁磁化规律可以写作:

$$\bar{\mu} = \alpha \mathbf{B}$$

- 电子的磁矩与角动量的比称为旋磁比(gyromagnetic ratio), 轨道旋磁比与自旋旋磁比分别为:

$$\gamma_L = -\frac{e}{2m_e}, \quad \gamma_S = -\frac{e}{m_e}$$

角动量量子化理论告诉我们, z 方向上电子轨道与自旋角动量只能为:

$$L_z = \pm n\hbar, \quad S_z = \pm \frac{\hbar}{2}$$

故 z 方向上的磁矩的单位就是著名的玻尔磁子(Bohr magneton):

$$\mu_B = \frac{e\hbar}{2m_e} = 0.00579 \text{ meV/T}$$

事实上由于自旋-轨道耦合, 分子磁矩可以是玻尔磁子的分数倍.

- 原子核也有磁矩, 显然由于质子质量比质子质量大三个量级, 故根据旋磁比算法, 其磁矩也小三个量级. 虽然对分子磁矩和宏观磁化几乎没有贡献, 但是利用在磁场下形成的各个能级间跃迁吸收发射电磁辐射的原理可以造成核磁共振(nuclear magnetic resonance) 现象.
- 如果分子的固有磁矩为零, 那么在外磁场下会产生一个与外磁场反向的磁矩, 宏观上造成抗磁性(diamagnetism). 从微观来看似乎与电介质极化的方向相反, 但是从宏观效果上却与电介质削弱电场的效果相同, 这是电与磁固有的区别导致的.
- 抗磁性的微观起源为电子轨道运动的拉莫尔进动(Larmor precession). 由轨道运动产生的力矩, 角动量, 磁矩三者关系可得:

$$\begin{aligned} \mathbf{M} &= \mu \times \mathbf{B}, \quad \mu = \gamma \mathbf{L}, \quad \frac{d}{dt} \mathbf{L} = \mathbf{M} \\ \Rightarrow \quad \frac{d}{dt} \mathbf{L} &= -\gamma \mathbf{B} \times \mathbf{L} = \frac{e\mathbf{B}}{2m} \times \mathbf{L} \end{aligned}$$

故在运动学上, 电子将以以下角速度发生进动:

$$\Omega = \frac{e}{2m} \mathbf{B}$$

电子原来的轨道运动半径如果按照玻尔半径 a_0 估计, 而角动量量级为 \hbar , 那么现在附加的角速度与原来的角速度的比为:

$$\Omega/\omega \sim \frac{ea_0^2}{2\hbar} B$$

故产生的磁矩与原来的玻尔磁子的比值的量级也与之相当, 得到磁矩的量级为:

$$\bar{\mu} \sim -\frac{e^2 a_0^2}{4m} \mathbf{B}$$

系数便是抗磁性磁化的分子磁化率的量级, 它一般比顺磁性磁化要小一到两个量级.

- 有一些金属, 或者特殊的金属化合物具有独特的铁磁性(ferromagnetism), 微观上它们由介观的包含大量原子的磁畴(magnetic domains)构成. 每一个磁畴由于单元之间的关联已经超过了热力学的涨落而成为了决定性的因素, 使得所有电子的自旋几乎都严格指向同一个方向, 磁化达到饱和. 而材料的磁化过程其实是不同的磁畴在外磁场下的转向. 它的磁化率将远远大于简单的顺磁和抗磁磁化. 而且体现出非线性和历史相关性.

4.4.2 宏观角度理解磁化

- 出于 \mathbf{B} 和 \mathbf{M} 的旋度分别为总电流密度和磁化电流密度, 引入磁场强度(magnetic field strength)以区别于以往一贯描述磁场的 \mathbf{B} , 一般可区别称作磁通密度(magnetic flux density):

$$\mathbf{H} = \frac{\mathbf{B}}{\mu_0} - \mathbf{M}$$

这样这个矢量的旋度就只取决于外电流:

$$\nabla \times \mathbf{H} = \mathbf{j}_f$$

- 在顺磁或铁磁的柱状介质外绕制密绕螺线管, 那么没有介质时其磁通密度为:

$$B_0 = \mu_0 i$$

那么, 由于此时加入介质时磁化电流不影响 \mathbf{H} 的计算, 故磁场强度为:

$$H = \frac{B_0}{\mu_0} = i$$

但是磁化电流的存在将产生一个与原磁场同向的附加磁场, 使得介质内部的磁场变大, 那么定义此时磁场与原磁场的比值为相对磁导率(relative magnetic permeability):

$$B/B_0 = \mu_r$$

那么把绝对磁导率定义为 $\mu = \mu_r \mu_0$, 就有以下关系式:

$$B = \mu H = \mu i$$

$$M = (\mu_r - 1)H = \chi_m H$$

其中 χ_m 就是宏观的磁化率(magnetic susceptibility). 对于抗磁性物质, 其值是小于零的.

4.4.3 磁场能量

磁场体系的能量的计算方法可以证明可以用电流与矢势和用能量密度两种计算方法:

$$I = \frac{1}{2} \int \mathbf{j} \cdot \mathbf{A} dV = \int \frac{1}{2\mu_0} \mathbf{B}^2 dV$$

同样地, 两个体系同时存在时, 体系将产生自能与相互作用能. 如两个电感元件如果之间存在互感, 那么其总能量应当为:

$$E = \frac{1}{2} L_1 I_1^2 + \frac{1}{2} L_2 I_2^2 + M I_1 I_2$$

如果存在介质, 其能量密度需要加上磁化带来的能量, 能量密度变为:

$$w = \frac{1}{2\mu} \mathbf{B}^2 = \frac{1}{2} \mu \mathbf{H}^2 = \frac{1}{2} \mathbf{B} \cdot \mathbf{H}$$

第五章 磁生电

5.1 电磁感应

静止电荷产生的静电场, 稳恒电流产生的静磁场, 这些都已经被我们研究地足够充分了. 是时候, 我们也该思考以下问题:

1. 如果电荷不是完全静止, 静电场是否会有改变?
2. 如果电荷不是完全静止, 会不会产生一个磁场?
3. 如果电流不是完全稳恒, 静磁场是否会有改变?
4. 如果电流不是完全稳恒, 会不会产生一个电场?

5.1.1 动生电动势

- 动生电动势的非静电力: 磁场力沿导线方向的分量:

$$\mathbf{K} = \mathbf{v} \times \mathbf{B}$$

- 回路的总动生电动势的计算方法:

$$\mathcal{E}_m = \oint \mathbf{K} \cdot d\mathbf{l} = - \frac{d\Phi}{dt} \Big|_{B \text{ 不变}}$$

总电动势等于磁通量变化率, 这就是法拉第电磁感应定律的动生部分.

- 一段电路动生电动势做功的功率:

$$dP_{\mathcal{E}} = d\mathcal{E} \cdot I = I(\mathbf{v} \times \mathbf{B}) \cdot d\mathbf{l}$$

该功率变成了电路内部的能量. 但是注意到磁场力的另外一个分力: 安培力也要做功, 功率为:

$$dP_A = d\mathbf{F} \cdot \mathbf{v} = I(d\mathbf{l} \times \mathbf{B}) \cdot \mathbf{v}$$

这个功将转化为导线的机械能. 那么由于磁场力本质上不会做功, 这一点很容易就能够利用三重标积公式在两个功率上得到印证:

$$dP_{\mathcal{E}} + dP_A = 0$$

从而两个功率的大小代表机械能和电能转化的快慢. 对于电动机, 电能转化为机械能, 对于发电机, 机械能转化为电能.

5.1.2 感生电动势

- 相对性原理容易说明,如果产生磁场的物质在发生运动以致使其发生变化,那么即使线圈静止,也会产生感应电动势.那么内部电荷受到的非静电力就必然不是磁场力,因为根据电磁场的定义:

$$\mathbf{F} = q(\mathbf{E} + \mathbf{v} \times \mathbf{B})$$

而现在线圈 $\mathbf{v} = \mathbf{0}$.我们只能把这个力称作电场力.这种电场称作感生电场(induced electric field).它可以理解为由变化的磁场产生的.事实上,磁场又由电流产生.故可以直接认为这个感生电场就是由变化的电流产生的.更有甚者,以下关系式均是成立的:

- 感生电场公式:

$$\mathbf{E} = -\frac{\partial \mathbf{A}}{\partial t} = -\int \frac{\mu_0}{4\pi} \frac{\partial \mathbf{j}}{r^2} dV$$

- 将上式对回路进行积分便得到感生电动势:

$$\mathcal{E}_i = \oint \mathbf{E} \cdot d\mathbf{l} = -\frac{\partial}{\partial t} \oint \mathbf{A} \cdot d\mathbf{l} = -\frac{\partial \Phi}{\partial t} = -\left. \frac{d\Phi}{dt} \right|_{S \text{ 不变}}$$

这就是法拉第电磁感应定律的感生部分.

- 将感应电动势的动生和感生部分合并,得到完整的法拉第电磁感应定律:

$$\mathcal{E} = -\frac{d\Phi}{dt}$$

5.2 自感与互感

- 自感定义与自感系数公式:

$$\Psi = LI \quad , \quad L = \frac{\mu N^2 S}{l}$$

- 两线圈互感时的相关公式:

$$E = \frac{1}{2}L_1 I_1^2 + \frac{1}{2}L_2 I_2^2 + MI_1 I_2$$

$$\Psi_1 = L_1 I_1 + MI_2 \quad \Rightarrow \quad U_1 = -L_1 \frac{dI_1}{dt} - M \frac{dI_2}{dt}$$

$$\Psi_2 = L_2 I_2 + MI_1 \quad \Rightarrow \quad U_2 = -L_2 \frac{dI_2}{dt} - M \frac{dI_1}{dt}$$

$$U_1 I_1 + U_2 I_2 = -\frac{dE}{dt}$$

第六章 麦克斯韦方程组

6.1 麦克斯韦方程组

6.1.1 电生磁

磁荷不存在,为电磁学规律带来了充分的简化.但是目前我们发现的三种现象:静止电荷产生电场,稳恒电流产生磁场,变化电流(通过产生变化的磁场)产生涡旋电场,并不是电磁学客观规律的全部.前两章我们已经充分说明,后两种现象其实都是相对性原理的后果:在一个参考系内不变的电荷分布,必将在另一个参考系看来产生电流,稳恒电流也会因为在另一个参考系而随时间变化,从而在一个参考系中纯粹的电场可以在另一个参考系内引发磁场和涡旋电场.电磁学的最后一项现象,由麦克斯韦(J. C. Maxwell)在1861提出的,位移电流产生磁场的现象,在某种程度上也由相对性原理引发.我们之后就可以发现这一点.

一切的源起,在于麦克斯韦对安培环路定律普遍性的质疑.我们不能认为在变化的荷与场的情况下,磁场的安培环路定律具有普适性.它写作:

$$\oint_{\partial S} \mathbf{B} \cdot d\mathbf{l} = \mu_0 \int_S \mathbf{j} \cdot d\mathbf{S}$$

为何质疑?考虑一个电容器不断充电的过程,如图所示.我们可以认为在距离电容较远的某点处,取一个垂直于导线的半径足够小的圆盘 S_1 后,边界环路上的磁场与导线上的电流之间符合安培环路定律的关系.但是注意到如果我们选择一个兜状的面 S_2 时,它的边界可以与圆盘边界完全一致.但是由于它把导线和一个极板完全兜住,故这个面上完全没有电流.从而就不能认为这个面和它的边界符合安培环路定律!至少, S_1 和 S_2 绝对不可能同时符合安培环路定律!这是因为,等号左侧这两个面边界上具有同一个磁场的环路积分,但是等号右侧却是不一样的电流通量: S_1 上为导线电流 I 而 S_2 上为零.两个等式是不可能同时成立的.

但是空间中的磁场又的确是客观存在的,而不是虚无缥缈的辅助物理量,我们依然可以通过对带电粒子受到的洛伦兹力的要求来确定空间中某处的磁场值:

$$\mathbf{F} = q(\mathbf{E} + \mathbf{v} \times \mathbf{B})$$

那么如何寻找这种情况下正确反映磁场计算方法的公式?麦克斯韦做到了这一点,而他的做法其实可以被简单的归纳为三点:

逻辑上要求自洽,数学上发展工具,哲学上遵从吝啬原则(law of parsimony)¹.

首先第一点,安培环路定律从逻辑上已经不能自洽了,所以必须修改它,一个合理的想法是为它增添一项:

$$\oint_{\partial S} \mathbf{B} \cdot d\mathbf{l} = \mu_0 \int_S \mathbf{j} \cdot d\mathbf{S} + (?)$$

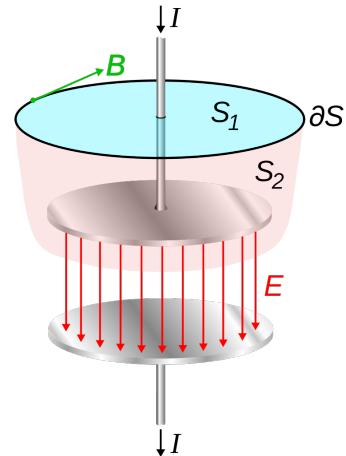


图 6.1: 修正安培环路定律

¹又做奥卡姆剃刀 (Occam's razor): 如无必要,勿增实体.认为基本规律的不同表述之间越简单的描述约能反应基本规律的本质.

增添这一项的目的就是为了能避免如之前电容器充电情况下的矛盾情况的发生. 事实上电容器充电的例子就提供了修改之后的表达式是否正确的试金石.

为了找到这个项, 矢量微分工具就发挥了不可替代的作用. 使用微分形式以上积分等式写作:

$$\nabla \times \mathbf{B} = \mu_0 \mathbf{j} + (?)$$

再两边同时取散度, 左边数学上一定变成零, 从而得到添加的项需要符合的条件:

$$\mu_0 \nabla \cdot \mathbf{j} + \nabla \cdot (?) = 0$$

再注意到电荷电流符合连续性方程:

$$\nabla \cdot \mathbf{j} + \frac{\partial \rho}{\partial t} = 0$$

从而得到:

$$\nabla \cdot (?) = \mu_0 \frac{\partial \rho}{\partial t}$$

为了实现这一点, 最简单的一种可能性为:

$$(?) = \mu_0 \varepsilon_0 \frac{\partial \mathbf{E}}{\partial t}$$

这是因为电场的散度就会回到电荷:

$$\nabla \cdot \mathbf{E} = \frac{\rho}{\varepsilon_0} \Rightarrow \nabla \cdot \left(\mu_0 \varepsilon_0 \frac{\partial \mathbf{E}}{\partial t} \right) = \mu_0 \varepsilon_0 \frac{\partial}{\partial t} (\nabla \cdot \mathbf{E}) = \mu_0 \frac{\partial \rho}{\partial t}$$

但是我们注意到, 使得之前的要求成立的添加项 (?) 不是唯一的, 事实上在刚刚的结果上添加任意一个矢量场的旋度都不改变其散度的值:

$$\nabla \cdot \left(\mu_0 \varepsilon_0 \frac{\partial \mathbf{E}}{\partial t} + \nabla \times \mathbf{F} \right) = \mu_0 \frac{\partial \rho}{\partial t} + 0$$

于是最后一个思想就发挥了作用: 如无必要, 勿增实体. 往往简单而自洽的形式就能反映真实的物理规律. 从而麦克斯韦就利用这样的推理得到了修正的安培环路定律:

$$\nabla \times \mathbf{B} = \mu_0 \mathbf{j} + \mu_0 \varepsilon_0 \frac{\partial \mathbf{E}}{\partial t}$$

同时麦克斯韦提出了以下视角去解释以上式子背后的物理图像: 我们定义位移电流(displacement current):

$$\mathbf{j}_d = \varepsilon_0 \frac{\partial \mathbf{E}}{\partial t}$$

那么我们的工作可以归纳为一句话: 通过引入虚构位移电流, 导致与真实电流的和形成一个没有散度(即连续)的电流场:

$$\nabla \cdot (\mathbf{j} + \mathbf{j}_d) = 0$$

举例来说, 在之前的电容器充电的场合下, 如果忽略电容器的边缘效应, 那么从上方往下通过的电流为 I 时, 电容器上积累的面电荷密度 σ 的增长为:

$$I = \dot{\sigma} S$$

从而电容内部的场强就会随之增长:

$$E = \frac{\sigma}{\varepsilon_0} \Rightarrow \dot{E} = \frac{I}{\varepsilon_0 S}$$

从而把电场的导数视作位移电流的话, 造成的总电流就是:

$$I_d = j_d S = \varepsilon_0 \dot{E} S = I$$

从而实际上补充位移电流后电流并没有在电容器极板上中断,这样也就不会造成之前电容器上不同面上电流不同而造成的矛盾情况.

如果认为补充这一项后形成的新安培环路定律具有在任意变化的电磁场下的普适性,我们可以发现,这样的一条规律在数学形式上具有与感生电场规律具有某种“对称性”.尤其是当空间中某处没有通过实际电流时,此时两条规律分别为:

$$\nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t}$$

$$\nabla \times \mathbf{B} = \mu_0 \epsilon_0 \frac{\partial \mathbf{E}}{\partial t}$$

即,变化的磁场(部分来源于变化的电流分布)会产生电场,变化的电场(部分来源于变化的电荷分布)会产生磁场.

我们曾经解释过第一个式子,它实际上可以解释为动生电动势现象的相对论效应.即在一个参考系中的磁场会在另一个参考系中招致电场:

$$\mathbf{E}' = \mathbf{v} \times \mathbf{B}$$

而一个参考系中的电场在另一个参考系中招致磁场这样的对称行为,实际上我们在之前静磁学章节中建立毕奥-萨伐尔定律时就已经有介绍过,得到的结果为:

$$\mathbf{B}' = -\frac{\mathbf{v} \times \mathbf{E}}{c^2}$$

这就不难看出两者之间的对称性来.而第一个式子,我们通过一个参考系中的动生电动势与另一个参考系中的感生电动势都与法拉第电磁感应定律相联系的思路,就推出了电场的旋度公式:

$$\mathbf{E}' = \mathbf{v} \times \mathbf{B} \Rightarrow \oint_{\partial S} \mathbf{E} \cdot d\mathbf{l} = - \int_S \frac{\partial \mathbf{B}}{\partial t} \cdot d\mathbf{S} \Rightarrow \nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t}$$

注意到系数 $1/c^2$ 实际上就是 $\mu_0 \epsilon_0$,完全仿照这个过程,我们也能得到磁场的环量也将等于电通量的导数,即对应有:

$$\mathbf{B}' = -\frac{\mathbf{v} \times \mathbf{E}}{c^2} \Rightarrow \oint_{\partial S} \mathbf{B} \cdot d\mathbf{l} = \mu_0 \epsilon_0 \int_S \frac{\partial \mathbf{E}}{\partial t} \cdot d\mathbf{S} \Rightarrow \nabla \times \mathbf{B} = \mu_0 \epsilon_0 \frac{\partial \mathbf{E}}{\partial t}$$

所以至此我们可以总结:电磁学的所有规律,都可以看做唯一一个规律:即库仑定律,在相对论原理下分化出来的丰富变式.

6.1.2 麦克斯韦方程组

将四个规律:电荷产生电场,电流产生磁场,变化的磁场产生电场,变化的电场产生磁场组合到一起,就得到了麦克斯韦方程(Maxwell equations),实际上指以下方程组²:

$$\nabla \cdot \mathbf{E} = \frac{\rho}{\epsilon_0}$$

$$\nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t}$$

$$\nabla \cdot \mathbf{B} = 0$$

$$\nabla \times \mathbf{B} = \mu_0 \mathbf{j} + \epsilon_0 \mu_0 \frac{\partial \mathbf{E}}{\partial t}$$

它完整而全面地描述了场与荷之间的关系.在之前的观点中,我们认为电荷是因,场是结果.这种观点之后将不再适用,这是因为在数学形式上,变化的电场和磁场可以互相产生,之后我们利用这一点能找到完全

²但是其实这四个方程是一个整体,正如电磁场本身应当视作一个整体对待

脱离电荷存在的电磁波解. 而完整的电磁理论实际上只需要再配合电场磁场本身的定义, 即洛伦兹的力公式:

$$\mathbf{F} = q(\mathbf{E} + \mathbf{v} \times \mathbf{B})$$

然而, 同样的一个现象 (电荷与场之间的相互作用) 可以有不同理论来描述. 力学我们就曾经介绍过牛顿经典力学体系和尝试过分析力学的可能性. 静电学和静磁学中, 电场和磁场不仅可以由场强来描述, 我们也研究了由势来描述的可能性. 当时的势与电荷的关系是比较简单的:

$$\varphi = \int \frac{1}{4\pi\epsilon_0} \frac{\rho}{r} dV, \quad \mathbf{A} = \int \frac{\mu_0 \mathbf{j}}{4\pi r} dV$$

那么能不能为变化的电磁场整体采取一个用势来描述的方法? 至少以上两个式子应当被摒弃, 因为他们只描述了电荷直接产生的电场与电流直接产生的磁场. 而我们知道变化的电荷与电流可以交叉地产生磁场和电场. 事实上之前成立的以下两个描述势与场强之间关系的式子也应当考虑摒弃:

$$\mathbf{E} = -\nabla\varphi, \quad \mathbf{B} = \nabla \times \mathbf{A}$$

后一个式子不会招致明显的矛盾, 因为即使在变化电磁场下, 依然有:

$$\nabla \cdot \mathbf{B} = 0$$

故无需添加新项即可成为恒等式.

矛盾集中体现在第一条上. 两边同时取旋度就能看出其中的不妥之处:

$$\nabla \times \mathbf{E} = -\nabla \times \nabla\varphi$$

其中左侧根据麦克斯韦方程, 它等于磁场的对时间变化率, 它不需要为零. 但是右侧根据数学规律它必然为零. 这就造成了矛盾.

解决这个问题的核心思想, 其实与麦克斯韦发现位移电流的思想完全一致. 我们添加一个项以使得矛盾消除, 以相同的思路推理, 过程留给读者, 可得这个项为:

$$\mathbf{E} = -\nabla\varphi - \frac{\partial \mathbf{A}}{\partial t}$$

这一点本身又不难理解. 我们知道完整的电磁理论有四个要素, 在最低阶的近似下之前的式子反映了电荷直接产生电场与电流直接产生磁场的要素. 那么我们添加的项实际上反映了变化的电流产生电场的要素: 因为磁矢势由电流产生, 那么它的导数自然反映了变化的电流, 而它构成了电场的第二项.

而变化的电荷产生磁场, 它不会在新的场与势的关系式中体现, 这是因为这样的改变直接同时改变了 \mathbf{B} 和 \mathbf{A} 的值. 前者通过位移电流, 后者之后介绍. 从而使得 $\mathbf{B} = \nabla \times \mathbf{A}$ 依然成立. 从而我们总结出任意情况下描述电磁场的标势和矢势需要符合的关系:

$$\mathbf{E} = -\nabla\varphi - \frac{\partial \mathbf{A}}{\partial t}, \quad \mathbf{B} = \nabla \times \mathbf{A}$$

但是, 接下来我们会引入一个惊人的不确定性: 其中的端倪在静电静磁学中其实已经出现苗头: 电势, 具有一个整体常数的不确定性, 这是因为电势的零点选取其实本质上是任意的, 以下变换后产生的标量场依然可以作为电势:

$$\varphi'(\mathbf{r}) = \varphi(\mathbf{r}) + C$$

只有进一步规定无穷远处的电势为零才能回到之前点电荷积分的电势公式:

$$\varphi(\infty) = 0$$

而静磁场的情况则更具有戏剧性: 我们之前通过电流元直接积分产生的磁矢势其实符合以下性质:

$$\nabla \cdot \mathbf{A} = 0$$

当然以上两个结果都还要求势不随时间改变, 比如:

$$\frac{\partial \varphi}{\partial t} = 0$$

如果仅仅希望能通过磁矢势的旋度计算磁场 $\nabla \times \mathbf{A} = \mathbf{B}$, 即使进一步要求 $\nabla \cdot \mathbf{A} = 0$, \mathbf{A} 的取法其实也不唯一. 可以证明, 如果找到所谓的三维调和函数, 即在全空间满足以下规律的标量场:

$$\nabla^2 \phi = 0$$

常数就属于调和函数, 除此之外我们还能举出很多的例子:

$$\phi = x, x^2 - y^2, xy, x^2 + y^2 - 2z^2, e^x \sin y, \dots$$

那么任意找一个调和函数梯度形成的矢量场:

$$\mathbf{F} = \nabla \phi$$

那么不难看出这个场就会符合:

$$\nabla \times \mathbf{F} = \mathbf{0}, \quad \nabla \cdot \mathbf{F} = 0$$

从而以下对磁矢势的变换就可以在不改变对磁矢势的要求下, 造成同一个磁场:

$$\mathbf{A}' = \mathbf{A} + \mathbf{F}$$

$$\nabla \cdot \mathbf{A}' = \nabla \cdot \mathbf{A} = 0 \quad \& \quad \nabla \times \mathbf{A}' = \nabla \times \mathbf{A} = \mathbf{B}$$

我们从其中归纳出两个概念. 一是我们要求电势和矢势根据具体问题不同选取不同的条件, 上面过程中出现的各种要求就有:

$$\begin{aligned} \nabla \cdot \mathbf{A} &= 0 \\ \frac{\partial \varphi}{\partial t} &= 0 \\ \varphi(\infty) &= 0 \end{aligned}$$

这些称作**规范条件**(gauge condition). 或者简单称作**规范**.

但是, 即使存在规范条件, 例如静磁场下要求 $\nabla \cdot \mathbf{A} = 0$, 或者进一步放宽干脆没有任何规范条件的限制, 我们发现电势和矢势其实不是位移确定的, 同样的一个电磁场, 在下式成立的条件下:

$$\mathbf{E} = -\nabla \varphi - \frac{\partial \mathbf{A}}{\partial t}, \quad \mathbf{B} = \nabla \times \mathbf{A}$$

势 φ, \mathbf{A} 存在不同的取法, 普遍来说, 存在以下**规范变换**(gauge transformation):

$$\varphi' = \varphi + \phi, \quad \mathbf{A}' = \mathbf{A} + \mathbf{F}$$

它能使得在特定规范下虽然描述场的势改变, 但是它不反映场的真实改变. 即同一个场对应多种相差规范变换的势的取法. 这种现象称作**非物理自由度**(nonphysical dof.). 它深刻反映了标势矢势不能直接反映可测量的物理量的实质, 正似我们在经典力学中一贯强调的势能的道理: 其本身绝对大小没有可测量的意义, 可以测量的是其在过程中的变化, 它才有实际物理意义. 但是, 对于标势矢势的哪一部分具有实际的物理意义还有待之后进行考察. 选取更多的规范可以消除势的不确定性, 这种操作称作**规范固定**(gauge fixing).

这就不难理解在普遍的电磁场方程中, 同样也存在着标势和矢势的不确定性. 更有甚者, 正如电场强度和磁场强度是描述同一个电磁场的两个侧面, 那么标势和矢势也是描述同一个电磁场的两个侧面, 两者的不确定性甚至是互相联系的. 如果没有任何实现假定的规范条件, 我们发现选取任意的随时间变换的标量场做变换:

$$\forall \phi = \phi(\mathbf{r}, t)$$

$$\varphi' = \varphi - \frac{\partial \phi}{\partial t} , \quad \mathbf{A}' = \varphi + \nabla \phi$$

可以发现它不改变通过之前式子计算的电场和磁场的强度. 从而这是一个非常普遍的规范变换.

而对于普遍的电磁理论, 它也是囊括静电学和静磁学的普遍情况的. 可以想见, 在不同的规范条件下就会产生不同的势. 例如对于静电学, 之前我们通过标势来描述电场, 这一点就不是必然的. 如果选取所谓的外尔规范 (Weyl gauge):

$$\varphi = 0$$

那么电场就必须全权由矢势 \mathbf{A} 来描述, 但这一点也毫无难度, 只需要取:

$$\mathbf{A} = - \int \frac{1}{4\pi\epsilon_0} \frac{\rho e_r}{r^3} dV \cdot t$$

这就足以使得:

$$\mathbf{E} = -\nabla\varphi - \frac{\partial \mathbf{A}}{\partial t} = \int \frac{1}{4\pi\epsilon_0} \frac{\rho e_r}{r^3} dV , \quad \mathbf{B} = \nabla \times \mathbf{A} = \mathbf{0}$$

事实上, 外尔规范是在研究磁现象占主导地位时常采用的规范, 此时电磁场整体都单由矢势 \mathbf{A} 来描述, 电场可以形象地视作由矢势变化产生的“感生电场”. 而在电现象占主导地位时, 更常采用库仑规范 (Coulomb gauge):

$$\nabla \cdot \mathbf{A} = 0$$

可以证明这样做带来的好处是标势的计算变得很简单: 只需要把电磁场和势之间的关系代入麦克斯韦方程, 我们得到两个恒等式和以下两个联系势和电荷电流的复杂关系式:

$$\nabla^2\varphi + \frac{\partial \nabla \cdot \mathbf{A}}{\partial t} = -\frac{\rho}{\epsilon_0}$$

$$\nabla^2\mathbf{A} - \nabla(\nabla \cdot \mathbf{A}) - \epsilon_0\mu_0 \frac{\partial^2 \mathbf{A}}{\partial t^2} - \epsilon_0\mu_0 \frac{\partial \nabla\varphi}{\partial t} = -\mu_0\mathbf{j}$$

代入规范条件, 首当其冲地我们发现: 第一个式子变成了:

$$\nabla^2\varphi = -\frac{\rho}{\epsilon_0}$$

这和静电学中的公式一致. 从而电势可以完全按照静电学的方式来计算, 它产生一个静电场. 而矢势则用来计算电荷分布变化带来的附加磁场和更高阶的感生电场.

但是, 除了静电或静磁现象占主导地位的两种情况, 我们之后研究的最多的是第三种情况: 那便是辐射占主导地位的情形. 此时无论是电荷密度或是电流密度, 要么其量级远远达不到辐射场对应的大小, 要么就是在一周期内平均值为零从而净效应很弱. 在这一类情况下, 使用的最多的是所谓的洛伦茨规范³ (Lorenz gauge). 写作:

$$\nabla \cdot \mathbf{A} + \frac{1}{c^2} \frac{\partial \varphi}{\partial t} = 0$$

这个神奇的条件使得之前描述势和电荷之间的关系变得非常简单, 背后原因在于可以证明这个条件也是唯一的在相对论变换下不同参考系内形式一致的规范条件. 在这个规范下:

$$\nabla^2\varphi - \frac{1}{c^2} \frac{\partial^2 \varphi}{\partial t^2} = -\frac{\rho}{\epsilon_0}$$

$$\nabla^2\mathbf{A} - \frac{1}{c^2} \frac{\partial^2 \mathbf{A}}{\partial t^2} = -\mu_0\mathbf{j}$$

这实际上构成了带非齐次项的波动方程. 一般称作达朗贝尔方程 (d'Alembert equations).

由于其最为简洁而普遍的形式. 我们之后的讨论都将以洛伦茨规范为默认规范, 并在以上达朗贝尔方程和关于荷与场强的麦克斯韦方程, 以及用势计算场强的关系式基础上展开讨论.

³这里的洛伦茨 (L. V. Lorenz) 不同于洛伦兹 (H. A. Lorentz), 前者为丹麦物理学家, 后者为荷兰物理学家

6.1.3 电荷在电磁场中的运动

在静磁学中我们简要介绍了磁矢势对于电荷与磁场相互作用的物理意义. 电势为引入场的试探电荷带来相互作用的电势能. 而磁矢势的改变似乎会改变试探电荷的某类动量. 通过之前的讨论, 我们发现电势和矢势其实内部具有非物理的自由度, 从而并不总是为试探电荷带来可以直接观测的物理量, 之前的电势带来电势能的情况, 其实仅仅是特殊情况下成立的结论, 并不具有普适性.

但是电势对应电势能的推导思路是依然可以借鉴的: 我们考虑了洛伦兹力的功, 并把电磁场用势来表示, 依然这么做给出:

$$\begin{aligned}\mathbf{F} \cdot \mathbf{v} &= q(\mathbf{E} + \mathbf{v} \times \mathbf{B}) \cdot \mathbf{v} \\ &= q\mathbf{E} \cdot \mathbf{v} \\ &= q\left(-\nabla\varphi - \frac{\partial\mathbf{A}}{\partial t}\right) \cdot \mathbf{v}\end{aligned}$$

我们考虑两个关系, 其一为:

$$\frac{d\varphi}{dt} = \frac{\partial\varphi}{\partial t} + \mathbf{v} \cdot \nabla\varphi$$

其背后的道理我们在力学中阐释过: 第一项为电荷运动过程中达到的点具有的电势的全导数, 它由不考虑电荷位移时电势随时间的局部导数和不考虑电势本身随时间改变而仅考虑随电荷位移导致的电势改变的随体导数项合成.

第二个关系为:

$$\frac{\partial}{\partial t}(\mathbf{v} \cdot \mathbf{A}) = \mathbf{v} \cdot \frac{\partial}{\partial t}\mathbf{A}$$

背后的道理为: 当我们考虑 $\mathbf{v} \cdot \mathbf{A}$ 整体作为时空 r, t 的函数时, \mathbf{v} 被我们视作了外来的参变量而不是时空的函数, 从而能够从偏导数符号中提出.

通过这两点, 我们得到了:

$$\mathbf{F} \cdot \mathbf{v} = -q\frac{d}{dt}\varphi + q\frac{\partial}{\partial t}(\varphi - \mathbf{v} \cdot \mathbf{A})$$

而左侧应当是电荷动能 T 的变化率. 那么这就给出了:

$$\frac{d}{dt}(T + q\varphi) = \frac{\partial}{\partial t}q(\varphi - \mathbf{v} \cdot \mathbf{A})$$

事实上, 单纯从洛伦兹力公式出发:

$$\mathbf{F} = q(\mathbf{E} + \mathbf{v} \times \mathbf{B})$$

考虑到等号左侧为粒子动量 \mathbf{p} 的变化率, 我们也可以通过推导 (静磁场那一章我们推导过没有电场的对应结果) 给出以下式子:

$$\frac{d}{dt}(\mathbf{p} + q\mathbf{A}) = -\nabla q(\varphi - \mathbf{v} \cdot \mathbf{A})$$

这两个结论引出了三组物理概念:

一是隶属于能量的概念: T 为动能, $q\varphi$ 现在可称作电磁势能. 而其和 $H = T + q\varphi$ 应当称作广义能量.

二是隶属于动量的概念: \mathbf{p} 为动量, $q\mathbf{A}$ 就是我们新发现的一个物理效应, 事实上磁矢势的物理意义可以理解为与位置有关的存储于场中的粒子的动量, 称作电磁势动量. 而机械动量 \mathbf{p} 与势动量 $q\mathbf{A}$ 的和 $\mathbf{P} = \mathbf{p} + q\mathbf{A}$, 称作广义动量.

事实上大家已经发现, 这种称呼其实就是来源于分析力学中的相应理论架构, 事实上电磁场中的电荷运动这一个实例我们之前分析力学正是以它举过例子, 读者可以对照着理解. 故其实我们还可以把 $H = T + q\varphi$ 称作哈密顿量(hamiltonian), 而 $\mathbf{P} = \mathbf{p} + q\mathbf{A}$ 称作正则动量(cannonical momentum).

最后还有一个概念, 它出现于之前两个等式的右侧, 它也是一种势能函数, 但是并不对应真实能量, 而是出现在电荷运动的拉格朗日量中作为势能函数. 这个项称作广义势:

$$V = q(\varphi - \mathbf{v} \cdot \mathbf{A})$$

其表达式中反常地出现了粒子的速度, 是一种速度势.

以往出现的守恒律与对称性现在就在这两个表达式中体现的淋漓尽致:

$$\frac{dH}{dt} = \frac{\partial V}{\partial t}$$

$$\frac{d\mathbf{P}}{dt} = -\nabla V$$

举例来说: 静电场中带电粒子的运动, 粒子总能量守恒; 静磁场中带电粒子的运动, 粒子干脆动能都不改变. 这都是源于广义势不随时间改变 (时间平移对称性导致能量守恒). 而在静磁场问题中还经常出现某方向动量与某个关于位置的函数的和守恒的情况, 究其本质是因为合理选取规范后, 磁矢势 \mathbf{A} 和电势 φ 都与某个位置坐标无关 (空间平移对称性导致动量守恒).

6.2 平面电磁波

6.2.1 真空中的电磁波

自由空间中的麦克斯韦方程:

$$\nabla \cdot \mathbf{E} = 0$$

$$\nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t}$$

$$\nabla \cdot \mathbf{B} = 0$$

$$\nabla \times \mathbf{B} = \epsilon_0 \mu_0 \frac{\partial \mathbf{E}}{\partial t}$$

也可以用势来描述, 此时达朗贝尔方程直接变成波动方程:

$$\nabla^2 \varphi - \frac{1}{c^2} \frac{\partial^2 \varphi}{\partial t^2} = 0$$

$$\nabla^2 \mathbf{A} - \frac{1}{c^2} \frac{\partial^2 \mathbf{A}}{\partial t^2} = \mathbf{0}$$

注意背后要求洛伦茨规范成立:

$$\nabla \cdot \mathbf{A} + \frac{1}{c^2} \frac{\partial \varphi}{\partial t} = 0$$

十分使得注意的是, 以上所有的方程, 作为齐次方程, 是具有非零解的. 最典型的解为平面波解:

$$\mathbf{E} = \mathbf{E}_0 e^{i(\omega t - \mathbf{k} \cdot \mathbf{r})}$$

$$\mathbf{B} = \mathbf{B}_0 e^{i(\omega t - \mathbf{k} \cdot \mathbf{r})}$$

$$\mathbf{A} = \mathbf{A}_0 e^{i(\omega t - \mathbf{k} \cdot \mathbf{r})}$$

$$\varphi = 0$$

$$\frac{E_0}{B_0} = c \quad , \quad \mathbf{E}_0 \perp \mathbf{k} \perp \mathbf{B}_0 \quad , \quad \mathbf{E}_0 \times \mathbf{B}_0 \parallel \mathbf{k}$$

$$\mathbf{E}_0 = -i\omega \mathbf{A}_0 \quad , \quad \mathbf{B}_0 = -ik \times \mathbf{A}_0$$

$$\omega/k = c$$

6.2.2 介质中电磁波的传播

注意在介质中,不能再仿照真空中的电磁波传播公式而写出齐次的麦克斯韦方程.这是因为介质内部完全可以存在电荷和电流,从而计算电场磁场的散度旋度时必然需要考虑到他们.但是我们此前研究过的电介质和磁介质的理论告诉我们,通过引入位移 \mathbf{D} 和磁场强度 \mathbf{H} ,可以有效地规避掉由介质产生的电荷和电流造成的电场散度和磁场旋度,即:

$$\nabla \cdot \mathbf{D} = 0$$

$$\nabla \times \mathbf{H} = \frac{\partial \mathbf{D}}{\partial t}$$

对于第二式可能需要更多的说明:我们知道如果在静磁学磁介质问题中,如果把总电流密度 \mathbf{j} 分解为磁化电流密度 \mathbf{j}_M 和自由电流密度 \mathbf{j}_f ,那么磁场旋度公式写作:

$$\nabla \times \mathbf{B} = \mu_0 \mathbf{j} = \mu_0 \mathbf{j}_M + \mu_0 \mathbf{j}_f$$

以往的做法是将磁化电流归因于磁化强度的不均匀性 $\mathbf{j}_M = \nabla \times \mathbf{M}$,并进一步定义 $\mathbf{H} = \mathbf{B}/\mu_0 - \mathbf{M}$,这样就给出:

$$\nabla \times \mathbf{H} = \mathbf{j}_f$$

但是现在作为自由介质 $\mathbf{j}_f = 0$.然而又由于是变化的电磁场,故磁场旋度需要增添电场变化带来的项.还有一点需要引起注意:极化强度 \mathbf{P} 的变化也会带来一个极化电流密度 \mathbf{j}_P ,它应当为:

$$\mathbf{j}_P = \frac{\partial \mathbf{P}}{\partial t}$$

对比极化电荷密度 $\rho_P = -\nabla \cdot \mathbf{P}$ 与极化电荷的连续性方程就不难理解这个式子.另外引人注目的一点是,这个电流与极化强度的关系完全类似与位移电流与电场强度的关系.如果做求和就会变成电位移的导数,这正是麦克斯韦当初把电场强度的导数称作“位移电流”的原因.

现在修改之后的磁场旋度写作:

$$\nabla \times \mathbf{B} = \mu_0 \mathbf{j} + \varepsilon_0 \mu_0 \frac{\partial \mathbf{E}}{\partial t} = \mu_0 \mathbf{j}_M + \mu_0 \frac{\partial \mathbf{P}}{\partial t} + \varepsilon_0 \mu_0 \frac{\partial \mathbf{E}}{\partial t}$$

最后两项十分恰好地凑出了电位移 \mathbf{D} .从而整理就得到了介质中修正的齐次磁场旋度公式.

在普遍的介质中的麦克斯韦方程就写作:

$$\begin{aligned}\nabla \cdot \mathbf{D} &= 0 \\ \nabla \times \mathbf{E} &= -\frac{\partial \mathbf{B}}{\partial t} \\ \nabla \cdot \mathbf{B} &= 0 \\ \nabla \times \mathbf{H} &= \frac{\partial \mathbf{D}}{\partial t}\end{aligned}$$

在这里介质兼具电与磁的特性,这表现在以下两个用于描述介质的结构方程上:

$$\mathbf{D} = \varepsilon \mathbf{E}$$

$$\mathbf{B} = \mu \mathbf{E}$$

其中 ε, μ 称作介质的电容率和磁导率.它们,超出我们以往的认知的是,往往是复数⁴.这一方面来自于微观极化和磁化机理上的弛豫,我们之后物理光学还将进一步研究这个问题,还有一方面来自于一个非常直接的原因:我们综合了不导电的电介质和磁介质,却忽视了介质本身完全可能导电的情况.对于导电性能较

⁴非各向同性介质中甚至抽象为复张量

差的情形,比如漏电的介质,考虑传导电流 j_c 实际上就是自由电流,那么磁场旋度的方程还需要被进一步修改为:

$$\nabla \times \mathbf{H} = j_c + \frac{\partial \mathbf{D}}{\partial t}$$

如果是考虑单一频率的电磁波,那么对时间求导的操作必然转变为乘 $i\omega$. 如果介质电导率为 σ ,那么可以把等号右侧进行以下变形:

$$j_c + \frac{\partial \mathbf{D}}{\partial t} = \sigma \mathbf{E} + i\omega \varepsilon \mathbf{E} = i\omega \left(\varepsilon - i\frac{\sigma}{\omega} \right) \mathbf{E} = i\omega \mathbf{D}' = \frac{\partial \mathbf{D}'}{\partial t}$$

这样就通过定义复电容率(complex permittivity):

$$\varepsilon' = \varepsilon - i\frac{\sigma}{\omega}$$

实现了把传导电流也吸收进了电位移中,造成了形式上的统一. 代价是电容率和相关的其他所有物理量都需要理解为复数.

这样的做法对于漏电绝缘介质是完全适用的. 但是对于导电性能非常好的导体介质中电磁波的传播,最好是单独建立新的处理方法. 有兴趣的读者可以阅读其他相关书籍. 我们指出,我们统一用复数描述不同情况下介质中电磁波传播的做法,从之后的结果上来看是符合事实的.

把介质的两个结构方程代入介质的齐次麦克斯韦方程,就得到了均匀介质中一般形式的麦克斯韦方程,习惯上总是以 \mathbf{E}, \mathbf{H} 作为变量:

$$\begin{aligned}\nabla \cdot \mathbf{E} &= 0 \\ \nabla \times \mathbf{E} &= -\mu \frac{\partial \mathbf{H}}{\partial t} \\ \nabla \cdot \mathbf{H} &= 0 \\ \nabla \times \mathbf{H} &= \varepsilon \frac{\partial \mathbf{E}}{\partial t}\end{aligned}$$

这两个方程具有高度的对称性,其解也与真空中的解相似,只有两点发生了改变,一是传播速度变为:

$$v = \frac{\omega}{k} = \frac{1}{\sqrt{\varepsilon \mu}}$$

这一点建立在 ε, μ 都是实数的基础上. 如果出现复数的情形,那么上式一般写作:

$$k = \sqrt{\varepsilon \mu} \omega$$

即此时将出现复波矢的情况. 意味着电磁波在介质中的传播伴随着耗散与衰减.

第二点就是电场磁场振幅强度的比例. 真空中这个比为 c ,现在由以下等式描述:

$$\varepsilon E_0^2 = \mu H_0^2$$

对于所谓良导体(good conductor)的情况,电导率 σ 特别大,以至于使得复电容率 ε' 中反倒是与电导率成正比的虚部远大于反应介电的实部.那么电磁波在良导体中的传播就存在严重的衰减. 此时有两个结论:

第一是对电磁场比值的公式,取 $\varepsilon' \rightarrow \infty$,就能发现:

$$E_0 \rightarrow 0$$

即良导体中类似于静电平衡情况下的导体,内部电场是非常微弱的. 磁场相对电场虽然很大,但是也需要迅速沿传播方向衰减.

第二是 k 的公式有以下渐进关系:

$$\frac{2\pi}{\lambda} \sim k \propto \sqrt{\frac{1}{\omega}} \cdot \omega = \sqrt{\omega}$$

即随着频率升高,传播与衰减的特征距离 λ 将越来越小. 这意味着高频交流电路中电流仅仅集中于导线表面薄薄的一层,会减小导线的有效截面积而增大线路损耗. 这个现象称作趋肤效应(skin effect).

6.3 电磁场能量与动量

真空中电磁场的能量密度为:

$$w = \frac{1}{2}\varepsilon_0 E^2 + \frac{1}{2}\mu_0 H^2$$

能流密度, 即著名的坡印亭矢量(Poynting vector) 为:

$$\mathbf{S} = \mathbf{E} \times \mathbf{H}$$

而动量密度与能流密度紧密联系:

$$\mathbf{g} = \mathbf{D} \times \mathbf{B} = \frac{1}{c^2} \mathbf{E} \times \mathbf{H} = \frac{\mathbf{S}}{c^2}$$

6.4 电磁波辐射

6.4.1 电磁辐射概论

麦克斯韦方程实际上描述了两类现象:

一类我们已经通过之前所有章节逐一学习. 表面上看包括静止的电荷产生电场, 稳恒的电流产生磁场, 变化的电荷分布与电流分布产生感应的磁场和电场四类. 但是其实无非都是第一类由库仑定律描述的基本规律在相对论变换下的衍生物. 在这一大类现象中, 场始终都是电荷的附庸, 基本的逻辑是先有荷, 再有场. 空间分布上, 场源电荷对远方的场的支配量级服从平方反比律.

通过本章的推理我们逐渐看到第二类截然不同的图像存在的可能性. 它诞生于同一个麦克斯韦方程, 但是故事不再由电荷主演, 主角阵容中新增了一颗闪亮的明星: 辐射场(radiation field). 平面波实际上就是形形色色辐射场中最单纯的一种. 显然, 它可以独立于电荷而存在. 但是依然, 电荷也愿意主动参与进麦克斯韦方程搭建的舞台中, 与辐射场这样的强势主角又将擦出怎样的火花? 简单来说, 我们将得到以下两个新的现象:

电荷发出辐射, 电荷吸收辐射.

拼借电荷吸收辐射的可能性, 第二类电磁现象也与第一类现象区分开来. 更何况电荷激发辐射场的空间分布也区别于第一类现象, 在辐射情形下, 远方的场是一次反比的. 这一点之后就能发现.

最后我们针对电荷发出辐射与吸收辐射的条件进行推理: 静止的点电荷产生静电场, 匀速运动的点电荷产生一个静电场的相对论变换, 这是毫无疑问的事实. 在这些场合下没有辐射场的产生.

所以, 显然电荷做加速运动是电荷激发辐射场的必要条件.

实际上, 电荷做加速运动其实即使就是电荷激发辐射场的充要条件, 我们之后就能给出电荷辐射场与辐射功率与其加速度的直接联系.

那么如何看待电荷吸收辐射的过程?

事实上这个过程可以看做电荷发射辐射的时间反演. 麦克斯韦方程具有明确的时间反演对称性, 只需要令 $t \rightarrow -t, \mathbf{B} \rightarrow -\mathbf{B}, \mathbf{j} \rightarrow -\mathbf{j}$. 其他量不改变就可以看出这一点. 辐射作为麦克斯韦方程的解, 其时间反演解自然也是合理的. 但需要注意先后发生的两件事总是存在因果的区别, 故时间反演后, 因果关系需要重新解释, 两个物理现象需要理解为不具有相通性⁵的独立现象. 我们可以举一个电荷吸收辐射的例子: 考虑电子的经典模型: 将一个平面电磁波照射到真空中的自由电子上, 那么电子就会在周期性的电场力(速度不大时磁场力可忽略)作用下做简谐振动, 那么根据下一小节的内容我们将知道, 电子会以一定功率向四面八方产生辐射场. 那么根据能量守恒, 电子必然是从原来入射的平面波中吸收了能量, 即平面波在一个周期内对电子做净的正功. 这种先吸收再辐射的行为就称作散射(scattering), 而以上经典的处理方法构成托马斯散射模型(Thomas scattering model). 而散射的总功率, 对应了原来平面波面积为 σ 范围内通过的功率. 这个面

⁵例如, 不能通过洛伦兹变换来颠倒因果.

积称作**散射截面**(scattering cross section). 经典理论计算结果为:

$$\sigma = \frac{8\pi}{3} r_e^2$$

其中 r_e 就是由洛伦兹提出的**经典电子半径**(classical electron radius):

$$r_e = \frac{e^2}{4\pi\epsilon_0 m_e c^2}$$

6.4.2 偶极辐射

通过点电荷产生的电势公式, 我们知道, 数学上以德爾塔函数作为泊松方程的非齐次项的解为一次反比势:

$$\nabla^2 \varphi = -\delta(\mathbf{r}) \Rightarrow \varphi = \frac{1}{4\pi r}$$

数学上容易猜想并证明, 如果将方程替换为达朗贝尔方程, 而如果将非齐次项以 ω 的角频率产生振荡, 那么这个新的方程的解应当变为球面波:

$$\nabla^2 \varphi - \frac{1}{c^2} \frac{\partial^2 \varphi}{\partial t^2} = -\delta(\mathbf{r}) e^{i\omega t} \Rightarrow \varphi = \frac{e^{i(\omega t \pm kr)}}{4\pi r}, \quad k = \frac{\omega}{c}$$

对于两个可能的解, + 数学上表示内行球面波, - 数学上表示外行球面波. 两个解在时间反演下对称. 目前我们在因果性上是考虑电荷体系的自发辐射, 故考虑负号的外行波解.

现在考虑一个这样的体系: 两个无限小的导体球用长为 l 的导线连接, 第一个球位于 $x = \frac{l}{2}$ 处, 第二个导体球位于 $x = -\frac{l}{2}$ 处. 由于两个球发生了 LC 振荡, 第一个球上电荷量为 $q_+ = q \cos \omega t$, 第二个球上电荷量为 $q_- = -q \cos \omega t$. 那么导线上的电流为 $I = -\omega q \sin \omega t$. 如果换成复数记法, 把实际值看做复数的实部, 则:

$$q_+ = q e^{i\omega t}, \quad q_- = -q e^{i\omega t}, \quad I = i\omega q e^{i\omega t}$$

定义从第二个球指向第一个球的矢量为 \mathbf{l} . 那么此时标势和矢势符合的达朗贝尔方程就写作:

$$\begin{aligned} \nabla^2 \varphi - \frac{1}{c^2} \frac{\partial^2 \varphi}{\partial t^2} &= -\frac{q}{\epsilon_0} \delta\left(\mathbf{r} - \frac{\mathbf{l}}{2}\right) e^{i\omega t} + \frac{q}{\epsilon_0} \delta\left(\mathbf{r} + \frac{\mathbf{l}}{2}\right) e^{i\omega t} \\ \nabla^2 \mathbf{A} - \frac{1}{c^2} \frac{\partial^2 \mathbf{A}}{\partial t^2} &= -i\mu_0 \omega q l \delta(\mathbf{r}) e^{i\omega t} \end{aligned}$$

那么标势和矢势的解就根据之前的数学结论得到:

$$\begin{aligned} \varphi &= e^{i\omega t} \cdot \frac{q}{4\pi\epsilon_0} \frac{e^{-ik|\mathbf{r}-\frac{\mathbf{l}}{2}|}}{|\mathbf{r}-\frac{\mathbf{l}}{2}|} - e^{i\omega t} \cdot \frac{q}{4\pi\epsilon_0} \frac{e^{-ik|\mathbf{r}+\frac{\mathbf{l}}{2}|}}{|\mathbf{r}+\frac{\mathbf{l}}{2}|} \\ \mathbf{A} &= e^{i\omega t} \cdot \frac{i\mu_0 \omega q l}{4\pi} \frac{e^{-ikr}}{r} \end{aligned}$$

这样的解称作**推迟势**(retarded potential) 解. 而对于电势可以进一步在足够远 $r \gg l$ 处进行小量近似, 我们不再考虑两项分母的区别, 这只会带来类似于电偶极子远方的电势那样的平方反比的高阶小量. 而主要关注分子的差:

$$\begin{aligned} e^{-ik|\mathbf{r}-\frac{\mathbf{l}}{2}|} - e^{-ik|\mathbf{r}+\frac{\mathbf{l}}{2}|} &\approx \frac{de^{-ikr}}{dr} \cdot \left(\left| \mathbf{r} - \frac{\mathbf{l}}{2} \right| - \left| \mathbf{r} + \frac{\mathbf{l}}{2} \right| \right) \\ &\approx ik l \cos \theta e^{-ikr} \end{aligned}$$

其中 θ 为 \mathbf{r} 方向与 \mathbf{l} 的夹角. 则在以 \mathbf{l} 方向为极轴的球坐标下, 写出近似后的标势与矢势:

$$\begin{aligned} \varphi &= e^{i(\omega t - kr)} \cdot \frac{ikql \cos \theta}{4\pi\epsilon_0 r} \\ A_r &= e^{i(\omega t - kr)} \cdot \frac{i\mu_0 \omega ql \cos \theta}{4\pi r}, \quad A_\theta = -e^{i(\omega t - kr)} \cdot \frac{i\mu_0 \omega ql \sin \theta}{4\pi r} \end{aligned}$$

最后再由势与场之间的关系:

$$\mathbf{E} = -\nabla\varphi - \frac{\partial\mathbf{A}}{\partial t}, \quad \mathbf{B} = \nabla \times \mathbf{A}$$

忽略超过一阶反比于半径的项, 就最终得到了电磁场的表达式:

$$\begin{aligned}\mathbf{E} &= E_\theta \mathbf{e}_\theta, \quad E_\theta = -\frac{\omega^2 q l \sin \theta}{4\pi\epsilon_0 c^2 r} e^{i(\omega t - kr)} \\ \mathbf{B} &= B_\varphi \mathbf{e}_\varphi, \quad B_\varphi = -\frac{\omega^2 q l \sin \theta}{4\pi\epsilon_0 c^3 r} e^{i(\omega t - kr)} = \frac{E_\theta}{c}\end{aligned}$$

从而取实部, 计算瞬时能流密度为:

$$S = \frac{\omega^4 q^2 l^2 \sin^2 \theta}{16\pi^2 \epsilon_0 c^3 r^2} \cos^2(\omega t - kr)$$

在一个周期内的平均辐射能流为:

$$\bar{S} = \frac{\omega^4 q^2 l^2 \sin^2 \theta}{32\pi^2 \epsilon_0 c^3 r^2}$$

最后计算全立体角内的总辐射功率:

$$P = \int_0^\pi \bar{S} \cdot 2\pi r^2 \sin \theta d\theta = \frac{\omega^4 q^2 l^2}{12\pi \epsilon_0 c^3}$$

考虑到辐射等价于由两个电荷量恒定, 但在 $\pm l/2$ 范围内做简谐振动, 加速度振幅为 $a = \omega^2 l/2$ 的电荷共同贡献. 那么每一个电荷加速度幅值 a 对应的辐射幅值为:

$$P = \frac{q^2 a^2}{6\pi \epsilon_0 c^3}$$

这就是著名的拉莫尔公式(Larmor formula).

第七章 交流电路

7.1 相量表示

7.1.1 拟稳条件与交流元件

- 拟稳条件: 一是忽略传输线上的推迟, 即传输线的分布电容电感电阻电导都被忽略, 如果不忽略同一导线两端电流就会不一致且导线将分去电压. 二是忽略元件内部的推迟. 不忽略会导致元件非理想.
- 拟稳条件对应数学表述:

$$\lambda \gg l \quad , \quad T \gg \frac{l}{c}$$

- 交流电路瞬时值描述: 每一瞬时类似直流电路, 但是电压电流在变化:

$$u = U_0 \cos(\omega t + \varphi_u)$$

$$i = I_0 \cos(\omega t + \varphi_i)$$

直流电路中的图, KVL, KCL 均成立. 但是多了电容电感两种元件, 其特性之后介绍.

- 有效值:

$$U = U_0/\sqrt{2} \quad , \quad I = I_0/\sqrt{2}$$

有效值存在意义是为了直接算平均功率:

$$P = I^2 R = U^2 / R = UI$$

- 交流电路的相量描述:

$$\tilde{U} = U e^{j\varphi_u}$$

$$\tilde{I} = I e^{j\varphi_i}$$

7.1.2 电阻, 电容, 电感特性

- 复阻抗:

$$\tilde{Z} = \frac{\tilde{U}}{\tilde{I}}$$

两种分解方法:

$$\tilde{Z} = |\tilde{Z}| e^{j \arg \tilde{Z}} = Z e^{j\delta}$$

Z 称为阻抗, δ 为余损耗角.

$$\tilde{Z} = R + iX$$

R 称为电阻, X 称为电抗.

- 复导纳:

$$\tilde{Y} = \frac{\tilde{I}}{\tilde{U}}$$

两种分解方法:

$$\tilde{Y} = |\tilde{Y}| e^{j \arg \tilde{Y}} = Y e^{j\delta}$$

Y 称为导纳, δ 为余损耗角.

$$\tilde{Y} = G + iB$$

G 称为电导, B 称为电纳.

- 电阻

$$\tilde{Z} = R$$

- 电感

$$\tilde{Z} = j\omega L$$

$Z_L = \omega L$ 为感抗. 任何阻抗若电抗为正称为感抗.

- 电容

$$\tilde{Z} = -\frac{j}{\omega C}, \quad \tilde{Y} = j\omega C$$

$Z_C = \omega C$ 为容抗. 任何阻抗若电抗为负称为容抗.

- 元件功率分为视在功率, 有功功率, 无功功率.

视在功率:

$$S = UI = I^2 Z$$

有功功率:

$$P = UI \cos \delta = I^2 R = \operatorname{Re}(\tilde{U} \tilde{I}^*) = \frac{1}{2} (\tilde{U} \tilde{I}^* + \tilde{U}^* \tilde{I})$$

无功功率:

$$P' = UI |\sin \delta| = I^2 |X|$$

三者关系:

$$P' = P \tan |\delta|, \quad S^2 = P^2 + P'^2$$

损耗角:

$$\tan \varepsilon = \frac{P}{P'} = \frac{R}{|X|}, \quad \pm \delta = \frac{\pi}{2} - \varepsilon$$

7.2 常见电路

7.2.1 谐振电路

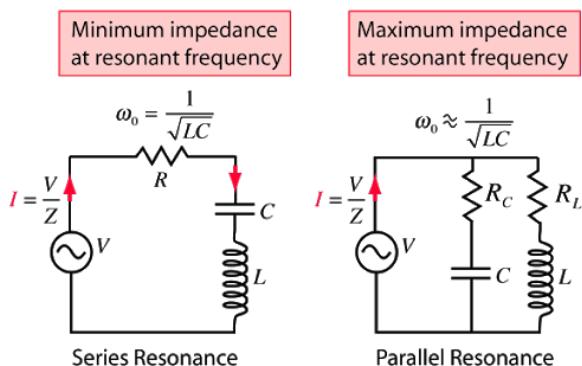


图 7.1: 串联谐振与并联谐振

- 串联谐振:

$$\tilde{Z} = R + j \left(\omega L - \frac{1}{\omega C} \right)$$

谐振条件: 电流作为响应与电压无相位差:

$$\omega = \frac{1}{\sqrt{LC}}$$

此时恰也有最小阻抗, 电流最大.

品质因数 Q :

$$Q = \frac{U_L}{U_R} = \frac{\omega_0 L}{R} = \frac{\omega_0}{\Delta\omega}$$

- 并联谐振: 谐振一般定义为相位谐振点, 即电流作为响应与电压无相位差. 此时阻抗一般接近最大值, 故干路电流反而较小. 支路电路往往远大于干路电流. 振幅达最小值的振幅谐振点一般不与相位谐振点重合.

7.2.2 滤波电路

7.2.3 电报线方程, 再论拟稳条件

7.3 变压器

7.3.1 磁路定律

7.3.2 理想变压器条件

7.4 电能传输

7.4.1 发电机与电动机: 三相绕组

7.4.2 变电站与电线损耗

7.4.3 市电规范

7.4.4 电源适配器

第八章 相对论力学

8.1 相对论运动学

- 光速不变原理: 时空的结构为平直时空: $ds^2 = c^2 dt^2 - dx_i^2$: 物质运动的“舞台”.
- 相对性原理: 物理规律的协变性:

$$\text{Sca.} = \text{Sca.}, \quad \text{Vec.} = \text{Vec.}, \quad \text{Vec.} \cdot \text{Vec.} = \text{Sca.} \text{ etc.}$$

- 四-标量(four-scalar) 与四-矢量(four-vector) 的定义: 类比三维空间标量矢量, 但按闵可夫斯基空间处理.
- 洛伦兹变换的定义: 保度规的变换. 物理上是不同共原点惯性参考者建立的参考系.
- 洛伦兹群:

$$\text{SO}(3, 1) = \langle R_x, R_y, R_z, B_x, B_y, B_z \rangle$$

其中 B_x 为 x 方向的推促(boost):

$$\begin{bmatrix} ct \\ x_1 \\ x_2 \\ x_3 \end{bmatrix}' = \begin{bmatrix} \gamma & -\gamma\beta & 0 & 0 \\ -\gamma\beta & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} ct \\ x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

- 经典物理没有协变性, 但电磁学与相对论理论具有协变性. 意为所有物理定律都由协变的对象构成, 这主要包括标量, 矢量, 张量, 旋量四类. 四-矢量尤其常见:

$$\begin{bmatrix} A_0 \\ A_1 \\ A_2 \\ A_3 \end{bmatrix}' = \begin{bmatrix} \gamma & -\gamma\beta & 0 & 0 \\ -\gamma\beta & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} A_0 \\ A_1 \\ A_2 \\ A_3 \end{bmatrix}$$

出于方便, 四-矢量也被简记为 $A_\mu = (A_0, A_i) = (A_0, \mathbf{A})$. 而 $A_i = (A_1, A_2, A_3)$ 是三维矢量的简记. 四-矢量与四-矢量可以算伪内积(pseudo-inner-product):

$$A_\mu B^\mu = A_0 B_0 - A_i B_i$$

伪内积将得到四-标量. 即参考系变换下不变的量.

- 洛伦兹变换下任意四-矢量具有不变量 $A_\mu A^\mu = A_0^2 - \mathbf{A}^2$, 对于类时四-矢量这将大于零, 即 $A_\mu A^\mu = A_{\text{proper}}^2$, 事实上以速度 $\mathbf{v} = \frac{\mathbf{A}}{A_0}c$ 做洛伦兹变换, 可以把空间分量变为零. 这叫做**本征参考系**(proper reference system). 得到本征四-矢量 $A'_\mu = (A_{\text{proper}}, \mathbf{0})$.
- 三个基本图像:
 - 尺缩: $L = L_0/\gamma$.
 - 钟慢: $t = \gamma t_0$.
 - 同时相对性: $\tau = -\beta \frac{x}{c}$.

只要是真的尺子, 真的钟, 就永远不可能错! 没有真的尺子和钟就去构造.

- 相对论质点运动:

具有四-标量: 间隔 $ds = \sqrt{dx_\mu dx^\mu}$, 本征时 $dt_0 = ds/c$, 下四-速度的不变模长 $v_\mu v^\mu = c^2$.

具有四-矢量: 四-位移 $dx_\mu = (cdt, d\mathbf{r})$, 四-速度 $v_\mu = (\gamma c, \gamma \mathbf{v})$, 四-加速度见下节.

- 速度变换公式:

$$\gamma' = \frac{1 - \frac{uv_x}{c^2}}{\sqrt{1 - u^2/c^2}} \gamma$$

$$v'_x = \frac{v_x - u}{1 - \frac{uv_x}{c^2}}$$

$$v'_{y,z} = \frac{\sqrt{1 - u^2/c^2}}{1 - \frac{uv_x}{c^2}} v_{y,z}$$

- 光行差公式: 可以用速度变换推导, 也可以之后用流变换推导:

$$\cos \theta' = \frac{\cos \theta - \beta}{1 - \beta \cos \theta}, \quad \sin \theta' = \sin \theta \cdot \frac{\sqrt{1 - \beta^2}}{1 - \beta \cos \theta}, \quad \tan \theta = \frac{\sin \theta'}{\cos \theta'}$$

- 相对论下讨论刚体? 模型存在缺陷: 隧道佯谬, 转盘佯谬. 故整个刚体运动学都没有必要建立.

8.2 相对论动力学

- 动量-能量构成四矢量的必要性: 守恒律需要被保留且协变.

- 四-动量:

$$p_\mu = \left(\frac{E}{c}, \mathbf{p} \right), p_\mu p^\mu c^2 = E^2 - p^2 c^2 = m^2 c^4, \mathbf{v} = \frac{\mathbf{p} c^2}{E}$$

$$\mathbf{p} = \frac{m \mathbf{v}}{\sqrt{1 - v^2/c^2}}, E = \frac{mc^2}{\sqrt{1 - v^2/c^2}}$$

- 普遍的过程的守恒律:

$$\sum_i p_{i,\mu} = \sum_j p_{j,\mu}$$

- 四-力与四-加速度:

$$a_\mu = \frac{dv_\mu}{dt_0} = \left(\gamma \frac{d\gamma}{dt} c, \gamma \frac{d\gamma \mathbf{v}}{dt} \right), \frac{d\mathbf{v}}{dt} := \mathbf{a}$$

$$F_\mu = \frac{dp_\mu}{dt_0} = \left(\gamma \frac{dE}{cdt}, \gamma \frac{d\mathbf{p}}{dt} \right) := \left(\gamma \frac{W}{c}, \gamma \mathbf{F} \right)$$

- 推论一: $F_\mu = a_\mu$. 但三维矢量之间不会直接符合 Newton-like 的动力学方程.
- 推论而: $F_\mu v^\mu = 0$. 这个三维继续可以得到:

$$\mathbf{F} dt = d\mathbf{p}, \mathbf{F} \cdot d\mathbf{r} = dE$$

- 单方向受力时, 在该方向的推促下力不变.
- 本征加速度与本征力: 它本意指的是四-速度的本征系中的四-加速度与四-力, 不过神奇的是恰好让两者的时间分量变为零 (类空, 与类时的四-速度正交), 空间分量为 $\mathbf{a}_0, \mathbf{F}_0$:

$$\mathbf{F}_0 = m\mathbf{a}_0$$

$$a_\mu a^\mu = -a_0^2$$

- 沿平行于本征加速度与本征力方向施以推促:

$$F = F_0, \quad a = a_0/\gamma^3$$

- 沿垂直于本征加速度与本征力方向施以推促:

$$F = F_0/\gamma, \quad a = a_0/\gamma^2$$

- 快度(rapidity): $\beta = \text{th}\vartheta, \gamma = \text{ch}\vartheta$.
- 自然坐标下的相对论性牛顿定律:

$$F_\tau = \gamma^3 m a_\tau$$

$$F_n = \gamma m a_n$$

- 相对论关于碰撞的新理解:

情形	经典	相对论
弹性	无耗散, $\mathbf{p}' = \mathbf{p}, E' = E$	粒子守恒, $\mathbf{p}' = \mathbf{p}, E' = E$
非弹性	有耗散, $\mathbf{p}' = \mathbf{p}, E$ 不可列	粒子变性, $\mathbf{p}' = \mathbf{p}, E' = E$

只有一个例外: 光子的散射若频率不变, 叫做弹性散射. 如经典的瑞利散射, 米氏散射; 但如果频率变了, 叫做非弹性散射, 如拉曼散射, 康普顿散射.

- 相对论碰撞根据具体情形一般有以下大的分类: 弹性的散射, 非弹性的散射, 粒子反应.
- 质点系的动量-能量也将构成一个四-矢量, 即:

$$(E, \mathbf{p}) = \sum_i (E_i, \mathbf{p}_i)$$

那么这个四-矢量的本征系速度:

$$\mathbf{v} = \frac{\mathbf{p}c^2}{E}$$

称作动量中心系 (center-of-momentum frame), 以代替经典物理的质心系. 简称动心系. 顾名思义, 此系中质点系总动量为零; 而总能量:

$$E_0^2 = E^2 - \mathbf{p}^2 c^2 = m_0^2 c^4$$

m_0 为动心系中的总动质量, 以后称作质点系的等效静质量. 这个式子可以解释一部分的质量起源问题.

- 一个等效静质量 m_0 的质点系若选取相对动心系以 v 运动的参考系, 动量与能量为:

$$\mathbf{p} = \sum_i \mathbf{p}_i = \frac{m_0 \mathbf{v}}{\sqrt{1 - v^2/c^2}}$$

$$E = \sum_i E_i = \frac{m_0 c^2}{\sqrt{1 - v^2/c^2}}$$

- 放能反应: $\sum_i m_i > \sum_j m_j$. 则无论取哪个系, 静能转化为动能的量都相等. 定义为反应能 Q :

$$Q = \left(\sum_i m_i - \sum_j m_j \right) c^2$$

- 吸能反应: $\sum_i m_i < \sum_j m_j$. 则具有阈能. **仅在动心系中**, 反应发生的最小动能 (阈能) 的值为:

$$Q = \left(\sum_j m_j - \sum_i m_i \right) c^2$$

8.3 相对论连续物质

- 若有一守恒荷 Q 作为四-标量, 则其空间分布与流动构成一个密度场 ρ, \mathbf{j} :

$$dQ = \rho dV$$

$$dQ = \mathbf{j} \cdot d\mathbf{S} dt$$

- 那么 $(\rho c, \mathbf{j})$ 是一个四-矢量且符合洛伦兹变换, 如 x 方向的推促:

$$\begin{bmatrix} \rho' c \\ j'_x \\ j'_y \\ j'_z \end{bmatrix} = \begin{bmatrix} \gamma & -\gamma\beta & 0 & 0 \\ -\gamma\beta & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \rho c \\ j_x \\ j_y \\ j_z \end{bmatrix}$$

- 若有一平面波场, 其必要的一个要素为相位, 它随着时空分布必然写作:

$$\varphi = \omega t - \mathbf{k} \cdot \mathbf{r}$$

ω 称作角频率, \mathbf{k} 为波矢.

- 那么由于相位为四-标量, 且可以写作以下形式, 故 $k_\mu = (\omega/c, \mathbf{k})$ 为四-矢量, 称作四-波矢:

$$\varphi = k_\mu x^\mu$$

如 x 方向的推促:

$$\begin{bmatrix} \omega'/c \\ k'_x \\ k'_y \\ k'_z \end{bmatrix} = \begin{bmatrix} \gamma & -\gamma\beta & 0 & 0 \\ -\gamma\beta & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \omega/c \\ k_x \\ k_y \\ k_z \end{bmatrix}$$

- 可以从上式出发得到光行差公式, 同时也能得到多普勒效应公式. 在前提 $\omega/k = c$ (四-波矢类光) 时:

$$\omega' = \omega \cdot \frac{1 - \beta \cos \theta}{\sqrt{1 - \beta^2}}$$

$$\cos \theta' = \frac{\cos \theta - \beta}{1 - \beta \cos \theta}$$

- 将三维的力密度和功率密度:

$$\mathbf{f} = \frac{d\mathbf{F}}{dV}$$

$$p = \frac{dP}{dV}$$

合并可以得到四-力密度 $f_\mu = (p/c, \mathbf{f})$.

第九章 光的干涉

9.1 标量波理论

标量波是指以下标量偏微分波动方程(wave equation) 的解:

$$\left(\nabla^2 - \frac{1}{c^2} \frac{\partial^2}{\partial t^2} \right) A(\mathbf{r}, t) = 0$$

分别在直角坐标, 柱坐标, 球极坐标下用所谓的分离变量法可以得到一些正确的基础解¹:

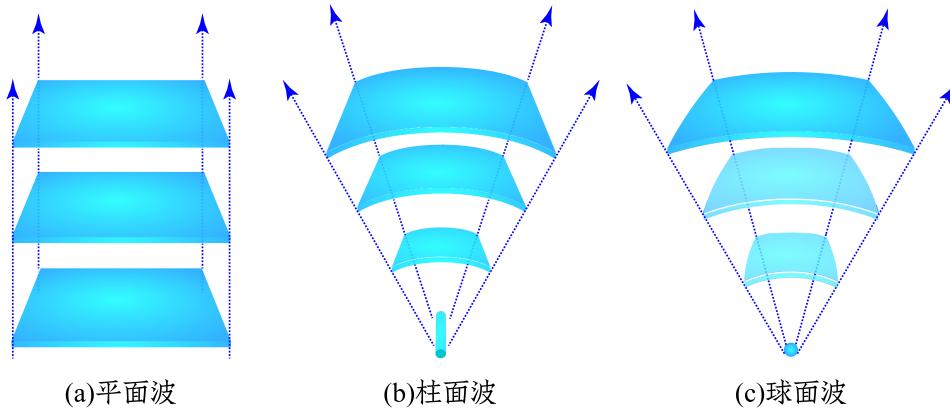


图 9.1: 三种典型的波

$$1. A(\mathbf{r}, t) = A \cos(\omega t - \mathbf{k} \cdot \mathbf{r} + \phi)$$

要求 $\frac{\omega}{|k|} = c$. 此为传播方向为 k 的平面行波.

$$2. \ A(\mathbf{r}, t) = \frac{I}{\sqrt{\rho}} \cos(\omega t - k\rho + \phi)$$

要求 $\frac{\omega}{|k|} = c \cdot k > 0$ 时为向外传播的柱面行波, k 小于零则向内传播.

$$3. A(\mathbf{r}, t) = \frac{I}{r} \cos(\omega t - kr + \phi)$$

要求 $\frac{\omega}{|k|} = c$. $k > 0$ 时为向外传播的球面行波, k 小于零则向内传播.

容易发现,以上波全都具有这样的形式:

$$A(\mathbf{r}, t) = |A|(\mathbf{r}) \cos(\omega t - \varphi(\mathbf{r}))$$

也就是说,每个点的相位都在以共同的 ω 做等速率的增加,代表一种同步的振动.而振动幅度 $|A| > 0$ 是要由于波的传播特点和能量流动的连续性而缓慢变化的.具体传播方向则看相位的梯度:

$$\boldsymbol{k} = +\nabla \varphi$$

¹以下三式都要求 $\omega > 0$, 数学上看 $\omega < 0$ 也没错, 但是利用三角函数的性质可以将 t 系数变号而不带来新的解.

在近似的意义上, $\frac{\omega}{|\mathbf{k}|} = c$ 是没有问题的. 而 c 便是波传播的速度. 在真空与介质 (写作 v) 中出现在光的波动方程中的 $\frac{1}{c^2}$ 中的 c 其实是:

$$c = \frac{1}{\sqrt{\epsilon_0 \mu_0}} , \quad v = \frac{c}{n}, \quad n = \sqrt{\epsilon_r \mu_r} \approx \sqrt{\epsilon_r}$$

约等于符号是因为考虑到了透明光学介质一般磁性质不显著故 $\mu_r \approx 1$.

我们不是很推荐用三角函数来计算波动光学. 要写成复指数的形式, 而实际三角函数理解为它的有效的实部, 虚部是为了方便计算而添加的没有物理效应的. 而且波动光学统一约定, 利用三角函数的形式, 为相位增添一个负号, 使得光在同一刻沿传播方向相位是增加的, 改为:

$$A(\mathbf{r}, t) = |A|(\mathbf{r}) e^{i(\varphi(\mathbf{r}) - \omega t)} = |A|(\mathbf{r}) e^{i\varphi(\mathbf{r})} e^{-i\omega t} = A(\mathbf{r}) e^{-i\omega t}$$

在以后情形下我们都将把这种符号约定称为“光学符号约定”, 而以往的指数上的宗量随时间增加随传播方向减小的符号约定称为“电磁学符号约定”. 以上讨论就引出了复振幅 $A(\mathbf{r})$ 的概念. 而之前找到的各点振幅 $|A|$ 和 φ 现在就被整合在了一起, 作为复振幅的模与幅角:

$$|A|(\mathbf{r}) = |A(\mathbf{r})| , \quad \varphi(\mathbf{r}) = \arg A(\mathbf{r})$$

回过头来再看, 我们在解波动方程时, 实际上从一开始就没有必要认为该标量 $A(\mathbf{r}, t)$ 是一个实数, 在复数域内解, 也能够直接得到完全与上述等效一样的结果. 最后我们发现, 该复数标量场 $A(\mathbf{r}, t)$ 如果能够符合波动方程, 其实部与虚部也是能够分别满足波动方程的:

$$\left(\nabla^2 - \frac{1}{c^2} \frac{\partial^2}{\partial t^2} \right) A(\mathbf{r}, t) = 0 \Rightarrow \begin{cases} \left(\nabla^2 - \frac{1}{c^2} \frac{\partial^2}{\partial t^2} \right) \Re A(\mathbf{r}, t) = 0 \\ \left(\nabla^2 - \frac{1}{c^2} \frac{\partial^2}{\partial t^2} \right) \Im A(\mathbf{r}, t) = 0 \end{cases}$$

如果再要求这样的 $A(\mathbf{r}, t)$ 具有形式 $A(\mathbf{r}, t) = A(\mathbf{r}) e^{-i\omega t}$, 数学上也就是对变量 t 分离变量, 物理就意味着研究单色光的传播, 那么代入便会发现实际上 $A(\mathbf{r})$ 需要满足的对空间的方程为 **亥姆霍兹方程** (helmholtz equation):

$$(\nabla^2 + k^2) A(\mathbf{r}) = 0 , \quad k = \frac{\omega}{c}$$

这就是整个波动光学理论的核心方程. 在继续阐述之前我们先说说 A 指什么. 其实真空与线性介质 (不导电) 电磁场几乎每个量 (标势, 矢势, 电场, 磁场) 都是符合波动方程的:

$$A : \{\varphi, A_x, A_y, A_z, E_x, E_y, E_z, B_x, B_y, B_z\}$$

$$\left(\nabla^2 - \frac{1}{c^2} \frac{\partial^2}{\partial t^2} \right) A(\mathbf{r}, t) = 0 , \quad (\nabla^2 + k^2) A(\mathbf{r}) = 0$$

所以作为复数的标量波理论就构成了光干涉, 衍射等波动光学课题的数学基础. 而且我们率先注意到其中的两点:

一是, 波动方程, 亥姆霍兹方程是线性方程, 也就是说, 如果两个光场都符合方程, 那么两个场直接逐点做标量加法得到的场也会自动符合方程. 这也就是之后干涉, 衍射操作的重要前提. 如果由于两个原因同时产生了两束光波, 那么只要写出各自的场, 加在一起便是整个体系现在的光场. 这就是双光束干涉时我们将会采取的做法.

二是, 平面波与球面波的重要性. 只要是行波, 在任意一点就可以按照平面波来近似. 如果取其 k 方向为 z 方向, 那么局部的光场为:

$$A(\mathbf{r}) = A e^{ikz}$$

但是旁边的点 (x_1, y_1) 的最佳平面波的近似却不一定平行于 z 轴, 它应该为:

$$A(\mathbf{r}) = A e^{i[k_z z + k_x(x - x_1) + k_y(y - y_1)]}$$

这下便有第一个近似条件了, 我们叫做**傍轴条件**(paraxial condition). 它要求光线传播方向与 z 轴夹一个小角度. 此时我们习惯把 x 和 y 方向方向余弦 $\cos \alpha_x, \cos \alpha_y$ 使用的约 90 度的角度 α_x, α_y 的余角记做方向角, 也就是说:

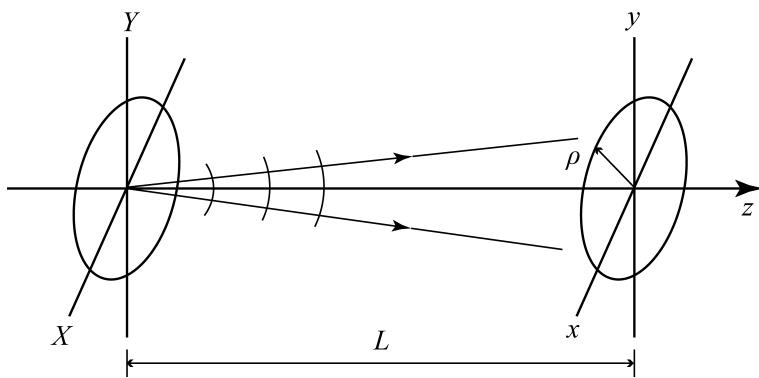


图 9.2: 球面波向平面波近似

$$\sin \theta_x \approx \theta_x = \frac{k_x}{k}, \quad \sin \theta_y \approx \theta_y = \frac{k_y}{k}, \quad k_z \approx \sqrt{1 - \theta_x^2 - \theta_y^2} k \approx k$$

在傍轴近似下, 以上以方向角 (θ_x, θ_y) 传播的平面波被方便地写为:

$$A(\mathbf{r}) = A e^{ikz} e^{i\phi} e^{ik(\theta_x x + \theta_y y)} = A' e^{ik(\theta_x x + \theta_y y)}$$

经常的, 我们只需要关心在像屏 $z = 0$ 上的场, 参考上图. 所以上式我们可以把不变的系数都统一写作 A' .

但是平面波却不总是好的近似, 我们已经发现了, 像上图那样, 如果在物屏 $z = -L$ 上原点 $X = 0, Y = 0$ 处有一点光源发出球面波, 那么光屏上 $(0, 0)$ 处和 (x_1, y_1) 处的近似公式就采取了不同的形式. 所以我们会寻求更好的球面波近似. 但且慢, 是不是任意情况下光场中的任意一点附近的场都能近似为一个球面波呢? 答案显然是否定的. 比如柱面波很显然就不能被近似为球面波, 因为其**波前**(wavefront), 即等相位面沿垂直于传播方向的一个方向弯曲而另一个方向却是平直的. 事实上球面波的光源是点, 柱面波的光源是线. 而介于柱面波与球面波中间的某种不对称波前就连光源都没法找到. 但是球面波的地位仍然是重要的. 这体现在两点上. 一是在真实的光路系统中, 点光源是切实存在且常用的. 此时物方到像方的点到点的消除了像差的理想成像是我们努力追求的方向. 故在这样的系统中光波的确一直都适合被近似为球面波, 而平面波也被当做一种光源在无穷远的特殊状态而处理. 二是, 若把波前上各个点当成是新的次波源向前发射出球面波而在前方相干叠加形成新的光场, 后面在衍射这一章我们数学上可以证明这就是计算光的传播的正确方式. 其中用到的球面波作为一种本质上重要的物理对象, 即**传播子**(propagator), 有着重要的物理地位.

故我们研究从物屏 $z = -L$ 的 (X, Y) 处传播到像屏 $z = 0$ 处 (x, y) 点的球面波的场:

$$\begin{aligned} A(x, y) &= \frac{I}{\sqrt{L^2 + (x - X)^2 + (y - Y)^2}} e^{ik\sqrt{L^2 + (x - X)^2 + (y - Y)^2}} \\ &= \frac{I}{L} e^{ikL} \left[1 + \frac{(x - X)^2}{L^2} + \frac{(y - Y)^2}{L^2} \right]^{-\frac{1}{2}} e^{ikL \left[\left[1 + \frac{(x - X)^2}{L^2} + \frac{(y - Y)^2}{L^2} \right]^{\frac{1}{2}} - 1 \right]} \end{aligned}$$

由于傍轴条件, 我们发现方向角 (u_x, u_y) , $u_x = \frac{x}{L}, u_y = \frac{y}{L}$ 和方向角 (U_x, U_y) , $U_x = \frac{X}{L}, U_y = \frac{Y}{L}$ 都是小量, 而传播到光屏上该点的光的方向角(k 方向)为 $(u_x - U_x, u_y - U_y)$. 仅考虑领头项, 把不随 x, y 变化的系数记为常数, 我们发现:

$$A(x, y) = A e^{\frac{1}{2}kL[(u_x - U_x)^2 + (u_y - U_y)^2]} = A' e^{ik\frac{x^2 + y^2}{2L}} e^{-ik\frac{Xx + Yy}{L}}$$

也就是说光屏上的波由两个相因子决定. 第一个是平方相关于 x, y 的 $\varphi_1 = k \frac{x^2 + y^2}{2L} = \pi \frac{x^2 + y^2}{\lambda L}$, 第二个恰好是平面波的相因子 $\varphi_2 = -k \frac{Xx + Yy}{L} = -k(U_x x + U_y y) = k(\theta_x x + \theta_y y)$. 就是说当点光源相对 $(x, y) = (0, 0)$ 从 $X < 0, Y < 0$ 的方向照过来能够在原点附近获得方向角为正的平面波场.

那么球面波能够被近似为球面波的条件也就不言自明了, 它要求第一个相因子的相位很小不干扰第二个相因子, 也就是 $\varphi_1 \ll \pi$, 也就是远场条件(farfield condition):

$$\rho^2 \ll \lambda L \Leftrightarrow L \gg \frac{\rho^2}{\lambda}$$

其实换一个角度理解这个条件, 如果远场条件被满足, 但是在这样的 ρ 内第二个相因子变化也非常小时, 整个 A 几乎就是常数, 这时候也形成不了有价值的光场, 所以我们如果要求 $\varphi_2 \gg \varphi_1$, 还能发现第三个条件 $X, Y \gg x, y$. 对于这个条件我们做这样的理解, 相位的绝对大小是没有观测效果的. 我们需要的其实是光束干涉时相位的差值. 故这其实是要求物屏的花样尺寸要远大于像屏上干涉花样的观测范围.

在远场条件不被满足的情况下, 考虑第一个相因子对第二个相因子的影响, 会造成干涉图样的扭曲与形变, 几何光学上会造成最基础的像差: 相散和散焦.

在介绍下两节具体的干涉之前, 对干涉的原理做一个简要的介绍是有必要的. 根据之前的介绍, 空间中如果同时存在两个光场:

$$A_1(\mathbf{r}) = |A_1|(\mathbf{r})e^{i\varphi_1(\mathbf{r})} ; A_2(\mathbf{r}) = |A_2|(\mathbf{r})e^{i\varphi_2(\mathbf{r})}$$

我们可以找到其单独存在时的光强, 一般来说, 光强为位置的缓变函数:

$$I_1(\mathbf{r}) = |A_1|^2 = A_1^* A_1 ; I_2(\mathbf{r}) = |A_2|^2 = A_2^* A_2$$

然而光场的叠加是有两种典型的方式的. 最常见的其实是非相干叠加(incoherent superposition). 此时出于下面要介绍的原因, 光强是可以逐点直接相加的:

$$I(\mathbf{r}) = I_1(\mathbf{r}) + I_2(\mathbf{r})$$

将自然界或人造的两束来源不同的光照在一次是无法观察到干涉花样的. 光的干涉条件比机械波干涉条件来的苛刻的多. 在某些细心制备的实验室条件下才能观察到干涉花样, 此时光场是相干叠加(coherent superposition). 在光学之中相干度是十分值得关注的. 我们将在之后的一节中集中讨论. 读者可能自然地认为相干叠加时为复振幅相加:

$$A(\mathbf{r}) = A_1(\mathbf{r}) + A_2(\mathbf{r})$$

但是其实无论非相干叠加还是相干叠加上式都适用, 那么是什么造成了两者的区别呢? 我们把上式模方求光强:

$$\begin{aligned} I(\mathbf{r}) &= A^* A \\ &= (A_1^* + A_2^*)(A_1 + A_2) \\ &= A_1^* A_1 + A_2^* A_2 + A_1^* A_2 + A_2^* A_1 \\ &= I_1 + I_2 + |A_1||A_2|e^{i(\varphi_2 - \varphi_1)} + |A_1||A_2|e^{i(\varphi_1 - \varphi_2)} \\ &= I_1 + I_2 + 2\sqrt{I_1 I_2} \cos(\varphi_2 - \varphi_1) \end{aligned}$$

上式中出现的交叉项 $2\sqrt{I_1 I_2} \cos(\varphi_2 - \varphi_1)$ 就叫干涉项(interference term). 它是关于位置的快变函数.乍一看这也未免太快了, 因为只要当空间位置改变一个波长, 相位 φ 就会有 π 的量级的改变, 这样干涉项在这个过程中就会有正负号的改变. 从而产生干涉花样. 在实际的干涉实验中, 合理的仪器设置可以使得这个空间特征长度被放大到人眼或助视仪器可以直接观测的程度. 从而这就是干涉实验的基础.

那么为什么对于非相干叠加情形我们直接去掉了干涉项. 那是因为在之前我们做把光场向标量复振幅简化的过程中去掉了公共的单色光时间演化项 $e^{i\omega t}$. 这并不是永远合理的做法. 真实光场由于后面要介绍的各种原因, 可以认为相位随时间的演化并不是完全线性增加的而是有随时间偏离期望值的方差越来越大的涨落. 从而两束并没有关联的自然光之间不同时刻同一点 \mathbf{r} 的相位差 $\delta = \varphi_2(\mathbf{r}) - \varphi_1(\mathbf{r})$ 是一个在随机涨落的数, 只要涨落够大, $\cos(\varphi_2 - \varphi_1)$ 项的随时间平均就会总是趋近于零. 从而干涉现象也就消失了.

在以下的两小节中, 我们先不用关心相干性的问题, 认为不同光场的叠加总是有“完美的相干性”: 完全稳定的相位差. 我们会在不同干涉装置的介绍中去强调为什么总是能保证“完美的相干性”的成立.

9.2 分波面干涉

以下几种干涉装置一般统一归为分波面干涉(interference by dividing wave-front). 因为它们实现的核心思想都是让本来沿各个方向独立传播的波前的部分改变传播方式而在空间中产生交叠.

- 杨氏双孔干涉: 单色扩展光源-单孔-双孔-光屏.

“完美的相干性”: 诚然, 单色扩展光源可以分解为大量点光源的叠加, 彼此之间是非相干的 (激光等非相干光源除外). 但是, 单孔的存在可以使得我们忽略此前的光场结构而直接把单孔视作单个点光源, 由它发出的光的同一波前上两点具有完全相干性. 实际上, 即使我们对到达单孔的来自扩展光源不同点光源的光进行分解, 那么每一个光波此后通过单孔形成新的球面波, 再通过双孔形成两个具有稳定相位差 (下面会计算) 的相干球面波, 并在光屏上形成干涉条纹, 只要不同点光源形成的条纹位置完全一致, 就等效于从单孔出来的光波整体具有“完美的相干性”.

条纹间距公式:

$$\Delta x = \frac{\lambda L}{d}$$

缺点: 只利用了点光源波面上非常有限的两个小面元. 而以下的都是改善方案, 直接将波面分成两个区域分别进行成像以实现干涉:

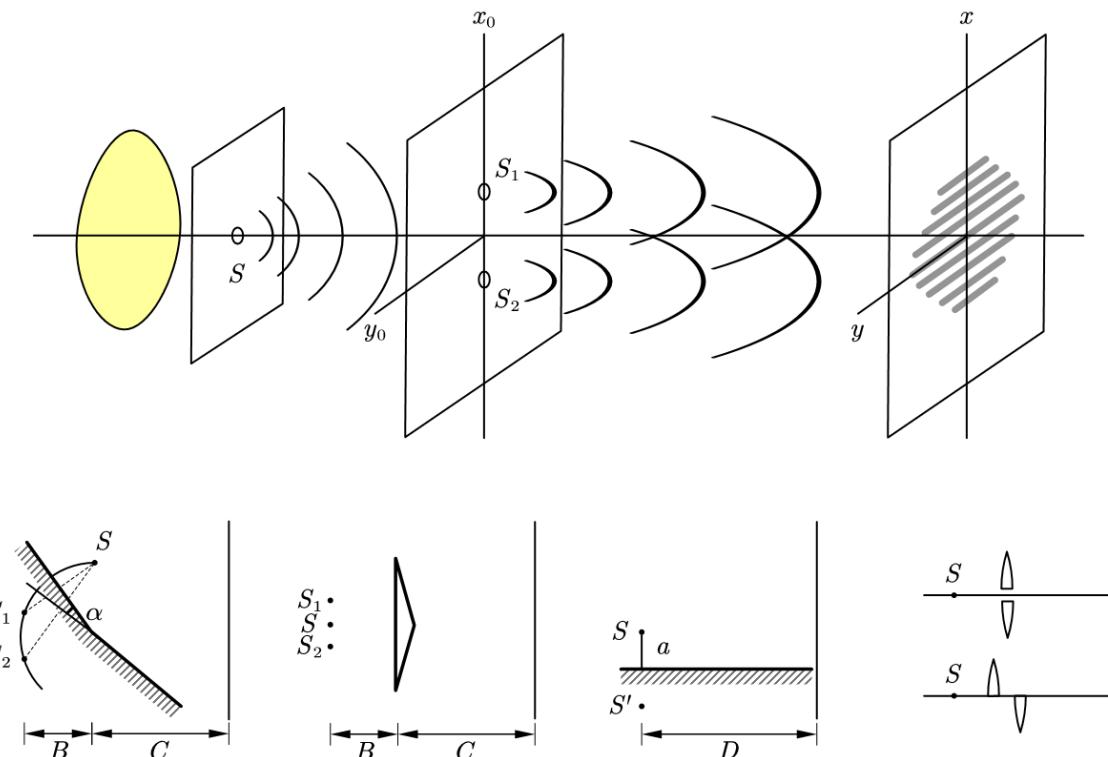


图 9.3: 分波面干涉们

- 菲涅尔双面镜:

$$\Delta x = \frac{\lambda(B + C)}{2\alpha B}$$

- 菲涅尔双棱镜:

$$\Delta x = \frac{\lambda(B + C)}{2(n - 1)\alpha B}$$

棱镜的正反放置有区别.

- 劳埃德镜:

$$\Delta x = \frac{\lambda D}{2a}$$

- 比累对切透镜: 横竖分开有区别.
- 散斑干涉: 如果在带一层玻璃板的镜面上撒粉末, 那么粉末作为散射源, 对于单色光入射情况也可以视作点光源处理. 此时也可以发生分波面干涉: 一部分吸收入射光后形成相应的散射光, 经反射后继续传播. 而一部分入射光先反射再被散射. 相当于在入射光波面上通过粉末和粉末的反射像发生了分波面干涉.

作为干涉, 一般指两个波的叠加. 但是实际情况常常遇到多光束干涉(multiple-beam interference) 的情况. 比如把杨氏实验换成多个孔, 或者把孔换成全同的散射单元. 多光束干涉的进一步推广就是衍射(diffraction), 此时一个空间区域内的整个波前应当被分解为无穷多个面微元, 各产生一个波, 在光屏上叠加, 后面会发现表达式成了一个积分.

9.3 分振幅干涉

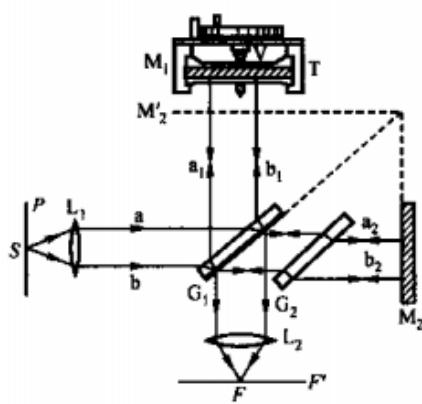


图 9.4: 迈克耳孙干涉仪

以下几种干涉装置一般统一归为分振幅干涉(interference by dividing amplitude). 因为它们实现的核心思想, 相比去把整个波前分成几个部分, 它选择保持了波前的完整性, 而是通过界面效应把其强度分解为透射与反射两个成分.

- 迈克耳孙干涉仪:

G_1 为半透半反镜, G_2 为完全透射的补偿板. M_2, M'_2 对 G_1 为物像关系. 那么在光屏中央为亮斑条件:

$$2d = j\lambda$$

事实上, 等价于薄膜干涉, 可以是等倾或者等厚.

- 薄膜等厚干涉: 光线倾角固定(往往受制于光路), 薄膜不同厚度处形成不同光程差导致干涉条纹.

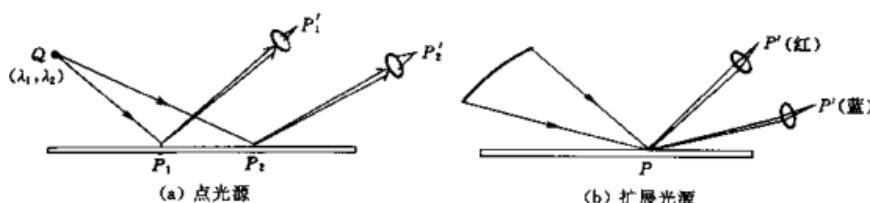


图 9.5: 薄膜等厚干涉

- 薄膜等倾干涉: 薄膜厚度固定, 不同光线倾角处造成不同光程差导致干涉条纹.

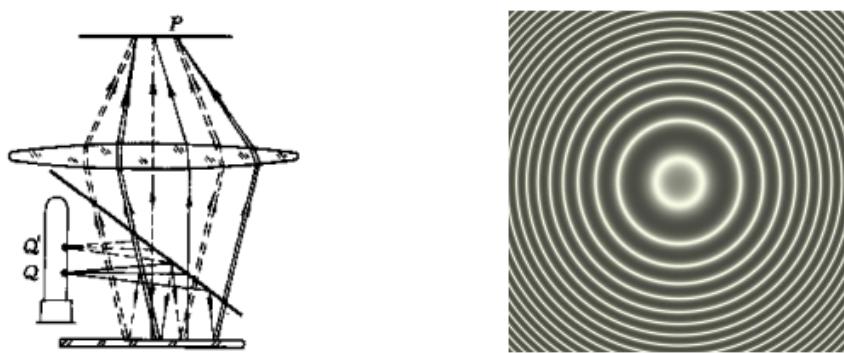


图 9.6: 薄膜等倾干涉

根据有无半波损失, 薄膜等厚等倾干涉的光程差均为:

$$\delta = 2nd \cos \theta \left(+\frac{\lambda}{2} \right)$$

- 法布里-泊罗干涉仪(Fabry-Pérot interferometer).

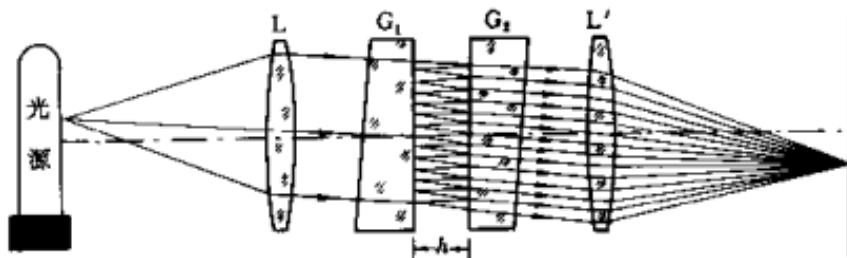


图 9.7: FP 干涉仪

真正的薄膜干涉往往是一个多光束干涉. 以 FP 干涉仪为例, 透射振幅为:

$$A_1 = t^2 A_0 = T A_0 = (1 - R) A_0$$

$$A_{i+1} = r^2 A_i e^{ik\delta} = R e^{ik\delta} \cdot A_i$$

做等比数列求和, 代入薄膜光程差公式, 模方取光强得到:

$$I = \frac{I_0}{1 + \frac{4R}{(1-R)^2} \sin^2(kd \cos \theta)}$$

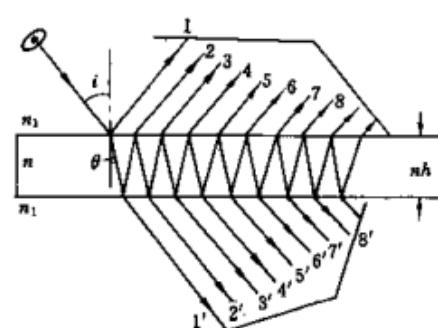


图 9.8: 多光束干涉

9.4 偏振干涉

- 相干度:

$$\gamma = \frac{\langle A_1^* A_2 \rangle}{\sqrt{\langle A_1^* A_1 \rangle \langle A_2^* A_2 \rangle}}$$

一般是个复数, 如果模为零代表完全不相干, 如果模为一代表完全相干, 但是幅角代表两个信号的相位差. 模在零一之间代表部分相干.

- 密度矩阵包含有一个偏振态的全部信息:

$$\mathbf{P} = \begin{bmatrix} \langle A_x^* A_x \rangle & \langle A_x^* A_y \rangle \\ \langle A_y^* A_x \rangle & \langle A_y^* A_y \rangle \end{bmatrix}$$

对角线元素都是实数:两个方向光强. 迹: 总光强.

- 线偏振光:

$$\mathbf{P} = \begin{bmatrix} I & 0 \\ 0 & 0 \end{bmatrix} \quad \text{or} \quad \mathbf{P} = I \begin{bmatrix} \cos^2 \theta & \cos \theta \sin \theta \\ \cos \theta \sin \theta & \sin^2 \theta \end{bmatrix}$$

- 圆偏振光, 左旋右旋分别为:

$$\mathbf{P} = \frac{I}{2} \begin{bmatrix} 1 & -i \\ i & 1 \end{bmatrix} \quad \text{or} \quad \mathbf{P} = \frac{I}{2} \begin{bmatrix} 1 & -i \\ i & 1 \end{bmatrix}$$

- 椭圆偏振光, 左旋右旋分别为:

$$\mathbf{P} = \begin{bmatrix} I_x & -i\sqrt{I_x I_y} \\ i\sqrt{I_x I_y} & I_y \end{bmatrix} \quad \text{or} \quad \mathbf{P} = \begin{bmatrix} I_x & i\sqrt{I_x I_y} \\ -i\sqrt{I_x I_y} & I_y \end{bmatrix}$$

- 以上三种光都具有完全相干性. 完全相干的光的任意两个分量, 或者是经过之后的偏振器件处理前后的各分量之间, 都具有模一的相干度. 但是, 两个不同的完全相干光之间又可以是不相干的. 偏振态为完全相干的充要条件为:

$$\det \mathbf{P} = 0$$

- 自然光: 相干度最差的光:

$$\mathbf{P} = \frac{I}{2} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

- 一般的光称作部分偏振光. 一般来说密度矩阵写作:

$$\mathbf{P} = \begin{bmatrix} I_x & I_{xy} e^{-i\varphi} \\ I_{xy} e^{i\varphi} & I_y \end{bmatrix}$$

可以证明可以找到“主方向”, 一方面, 这两个方向光强分别达到最大和最小 (取 x 方向最大, y 方向最小), 另一方面取这样的方向后恰好有 $\varphi = \pi/2$. 故部分偏振光在主方向坐标系中永远可以写作:

$$\mathbf{P} = \begin{bmatrix} I_{max} & -iJ \\ iJ & I_{min} \end{bmatrix}$$

即: 视作椭圆偏振光和一个与之不相干的自然光的叠加.

- 偏振片: 筛选某个特定方向的分振动:

$$A_\theta = A_x \cos \theta + A_y \sin \theta$$

- 波片: 命 x, y 方向的透射振动产生一定的相位差. 二分之一波片产生 $\varphi = \pi$ 相位差, 四分之一波片产生 $\varphi = \pi/2$ 相位差, 合理选取时间原点可以认为 y 方向相位不变, 那么:

$$A'_x = e^{i\varphi} A_x \quad , \quad A'_y = A_y$$

9.5 相干性

- 空间相干性: 在光源产生的光场同一时间 (同一波前) 上选取两个不同的空间点, 其信号相干度模不足一, 且随着点距离变大而愈发不相干的情况:

$$A_1 = A(\mathbf{r}_1, t) \quad , \quad A_2 = A(\mathbf{r}_2, t)$$

$$\gamma = \frac{\langle A_1^* A_2 \rangle}{\sqrt{\langle A_1^* A_1 \rangle \langle A_2^* A_2 \rangle}} \quad , \quad |\gamma| < 1$$

一般原因: 光源的扩展性. 对于宽度为 b 的单缝, 相干角度为:

$$\Delta\theta = \frac{\lambda}{b}$$

- 时间相干性: 在光源产生的光场同一位置 (同一波列) 上选取两个不同的时间点, 其信号相干度模不足一, 且随着时间间隔变大而愈发不相干的情况:

$$A_1 = A(\mathbf{r}, t + \Delta t) \quad , \quad A_2 = A(\mathbf{r}, t)$$

$$\gamma = \frac{\langle A_1^* A_2 \rangle}{\sqrt{\langle A_1^* A_1 \rangle \langle A_2^* A_2 \rangle}} \quad , \quad |\gamma| < 1$$

一般原因: 发光的不连续性, 存在波列长度 L . 内在原因为光源的非单色性, 即波长存在不确定度 $\Delta\lambda$. 两者符合:

$$L = \frac{\lambda^2}{\Delta\lambda}$$

第十章 光的衍射

10.1 光栅与波带片

- 光栅衍射: 光栅常数, 即空间周期为 d , 缝宽为 a , 总共刻 N 道. 那么:

$$I = I_0 \cdot f_1(\theta) \cdot \left(\frac{\sin N\beta}{\sin \beta} \right)^2, \quad \beta = \frac{\pi d}{\lambda} \sin \theta$$

其中 $f_1(\theta)$ 为之后可以算出来的与单缝衍射有关的慢变 (需要 $a \ll d$) 的单元因子:

$$f_1(\theta) = \left(\frac{\sin \alpha}{\alpha} \right)^2, \quad \alpha = \frac{\pi a}{\lambda} \sin \theta$$

体现光栅的结构的是后面的结构因子:

$$f_2(\theta) = \left(\frac{\sin N\beta}{\sin \beta} \right)^2, \quad \beta = \frac{\pi d}{\lambda} \sin \theta$$

从中可以得到三个信息: 主极大方向 $d \sin \theta_j = j\lambda$, 主极大强度 $I = N^2 I_0$, 主极大宽度 $\delta\theta \sim j\lambda/Nd$

- 圆孔菲涅尔衍射: 波带法与半波带法. 核心在于以下等式:

$$\frac{dS}{r} = \text{const.} \cdot dr$$



图 10.1: 波带法与半波带法

这说明不同位置的波前上的带状面元, 根据其对要计算的点所造成的光程差, 相等的光程差造成完全相等的振幅, 尽管面元的面积不相等. 只需要把这些振幅相等, 相位不等的光矢量进行合成.

合成方法一般有两种, 对于半波带问题把面元分解为半波带显得方便. 而复杂问题需要按照微元分解, 下一节将介绍面元对要计算的点的光矢量贡献还有一个起到调制作用的角度因子, 故合成图像一般为:

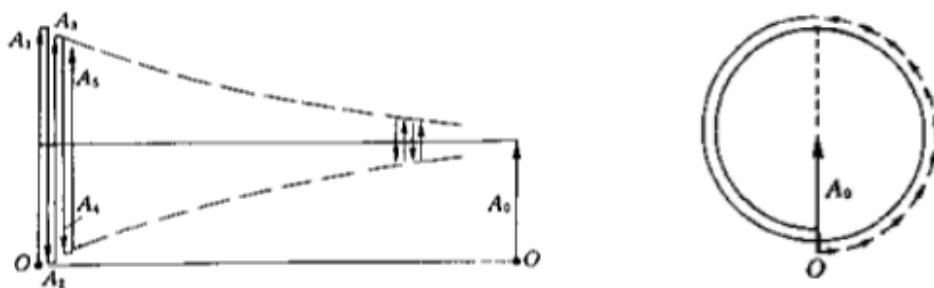


图 10.2: 分组合成与积分合成

- 波带片: 预想将一点光源和一要计算光强的点中间合适距离处的波前按照半波带分解, 再制造一光学器件, 或者遮住奇数半波带而开放偶数半波带, 或者遮住偶数半波带而开放奇数半波带. 就构成了波带片. 如果在一张圆形的片上入射平行光, 而选取距离为 f 的点计算光程, 那么各个半径为:

$$\rho_j = \sqrt{j\lambda f}$$

这样的波带片可以当做一个焦距为 f 的透镜使用. 但是成像并不唯一, 容易证明它具有一组虚焦点和实焦点.

10.2 衍射积分公式

- 基尔霍夫衍射积分公式:

$$A(\mathbf{r}) = \frac{1}{i\lambda} \int A(\mathbf{r}') \cdot F \cdot \frac{e^{ik|\mathbf{r}-\mathbf{r}'|} dS}{|\mathbf{r}-\mathbf{r}'|}$$

F 为角度因子, 若入射波波矢 \mathbf{k}' 方向为 e' , 子波波矢 \mathbf{k} 方向为 e , 面元 dS 方向为 n , 则有:

$$F = \frac{\mathbf{n} \cdot \mathbf{e} + \mathbf{n} \cdot \mathbf{e}'}{2} = \frac{\cos \theta + \cos \theta'}{2}$$

- 夫琅和费衍射公式: 如果在衍射屏后放置焦距为 f 的透镜并在后焦面上观察, 并且忽略角度因子. 把入射场在衍射屏作用后的波前写作 $A(x_0, y_0)$. 那么这个情况具有更简单的公式:

$$A(\theta_x, \theta_y) = \frac{1}{i\lambda f} \iint A(x_0, y_0) e^{-i(k_x x_0 + k_y y_0)} dx_0 dy_0$$

- 矩孔衍射:

$$I = I_0 \cdot \left(\frac{ab}{\lambda f} \right)^2 \cdot \left(\frac{\sin \alpha}{\alpha} \right)^2 \cdot \left(\frac{\sin \beta}{\beta} \right)^2, \quad \alpha = \frac{\pi a}{\lambda} \sin \theta_x, \beta = \frac{\pi b}{\lambda} \sin \theta_y$$

- 圆孔衍射: 直径 d , 艾里斑半角 θ :

$$\theta \approx \frac{1.22\lambda}{d}$$

10.3 波前分析法

第十一章 物理光学

11.1 经典色散理论

蓝天,白云,红太阳.在简单的自然现象中蕴含着光在传播过程中的另外一种典型的现象: **散射**(scattering).而散射却又是一个多么复杂的现象!要完全理解散射,我们得研究光的吸收与发射,因为光与物质中分子的相互作用其实本质上可以看成吸收与发射,这是**分子光学**(molecular optics)的研究范畴.其次,光在介质中的衰减(吸收),不同颜色的光折射率的不同(色散)其实都导致了一类散射现象(瑞利散射).但是还有不同的散射现象则表现出不用的特性,需要从更宏观或更微观的角度去建立新的理论来解释他们.

不同于以往在透明介质中的平面波传播,新的现象需要新的平面波模型.

11.1.1 复波矢与复折射率

我们在波动方程的解中引入复数是为了更好地描述振动形式的解.但这却忽视了另一种解的存在性.就是如果波矢 \mathbf{k} 本身亦为复数,由于同时又是一个矢量,可以表示为 $\mathbf{k} = \alpha + i\beta$.代入场的空间部分:

$$A(\mathbf{r}) = A_0 e^{i\mathbf{k} \cdot \mathbf{r}} = A_0 e^{-\beta \cdot \mathbf{r}} e^{i\alpha \cdot \mathbf{r}}$$

这就代表了一个沿 α 方向传播,沿 β 方向衰减的平面波.而 α 为波矢的实部, β 为波矢的虚部:

$$\alpha = \Re(k) ; \beta = \Im(k)$$

这种波可不可以真空中传播?出人意料的是这居然是可能的.因为如果代入真空中的亥姆霍兹方程:

$$\nabla^2 A + \frac{\omega^2}{c^2} A = 0$$

便会觉得这只需要要求:

$$\mathbf{k}^2 = (\alpha + i\beta) \cdot (\alpha + i\beta) = \frac{\omega^2}{c^2} \Rightarrow \alpha^2 - \beta^2 = \frac{\omega^2}{c^2}, \alpha \cdot \beta = 0$$

这样的波我们并不陌生,早在几何光学我们介绍光的全反射时便研究过另一侧界面中的隐失波场便是以上形式.更典型地,在 $a \times b$ 的矩形波导中传播的波满足:

$$\begin{aligned} \left(\frac{2\pi}{\lambda_x}\right)^2 + \left(\frac{2\pi}{\lambda_y}\right)^2 + \left(\frac{2\pi}{\lambda_z}\right)^2 &= \frac{\omega^2}{c^2} \\ \lambda_x = \frac{2b}{m}, \lambda_y = \frac{2a}{n}, m, n &= 1, 2 \dots \end{aligned}$$

此时在 x, y 方向的振动是驻波, z 方向传播的波则是行波. x, y 方向要形成驻波是因为在边界的导体上要符合相同的边界条件,所以或都是波腹,或都是波节.亦可以用同一个

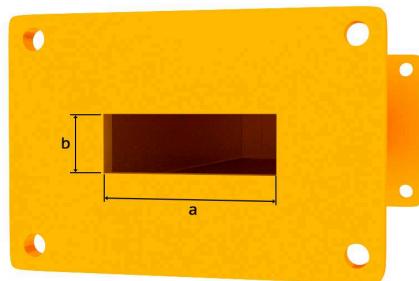


图 11.1: 矩形波导

平面波在侧面上的多次反射的叠加来理解这种兼具驻波和行波特质的场. 我们发现模式为 (m, n) 的波, 会有一个截止角频率 ω_c , 如果角频率低于此值使得波无法向前传播:

$$\omega_c = c \sqrt{\left(\frac{m\pi}{a}\right)^2 + \left(\frac{n\pi}{b}\right)^2}$$

那如果在波导的起点的天线给出信号角频率 $\omega < \omega_c$ 会发生什么情况? 此时就是 z 方向波矢 k_z 变为虚数, 意味着波沿 z 方向不是传播而是衰减:

$$\mathbf{k} = \boldsymbol{\alpha} + i\boldsymbol{\beta}$$

$$\boldsymbol{\alpha} = \left(\frac{m\pi}{a}, \frac{n\pi}{b}, 0 \right) \quad , \quad \boldsymbol{\beta} = \left(0, 0, \sqrt{\left(\frac{m\pi}{a}\right)^2 + \left(\frac{n\pi}{b}\right)^2 - \frac{\omega^2}{c^2}} \right)$$

在真空中尚可以产生衰减的波¹, 更何况在各种介质中. 事实上, 描述光在介质中随着传播的距离而衰减的现象早有一经验公式, 即**比尔-朗博-布葛律**(Beer-Lambert-Bouguer law):

$$\frac{I(l)}{I(0)} = e^{-(\sum_i n_i \sigma_i)l}$$

其中 $I(0)$ 为在传播 l 前的光强, $I(l)$ 为在传播 l 后的光强. n_i 是构成介质的第 i 种分子的数密度, σ_i 为描述其吸收本领的**衰减截面**(attenuation cross section). 例如对大气来说主要的衰减因素就来自气溶胶, 干净的海水则能提供 200m 左右的透光带, 如果有的话, 浮游植物含有的光合色素是造成衰减的主要原因. 而对于强衰减性质的介质, 我们可以唯象地认为波在介质中的传播为:

$$|A(l)| = |A(0)|e^{-\beta l} \Rightarrow I(l) = I(0)e^{-2\beta l}$$

这样的选择是有道理的. 下以导电导致的损耗为例: 在不漏电介质中, 麦克斯韦关系本应为:

$$\left\{ \begin{array}{l} \nabla \cdot \mathbf{E} = 0 \\ \nabla \times \mathbf{E} + \frac{\partial \mathbf{H}}{\partial t} = \mathbf{0} \\ \nabla \cdot \mathbf{H} = 0 \\ \nabla \times \mathbf{H} - \varepsilon \mu \frac{\partial \mathbf{E}}{\partial t} = \mathbf{0} \end{array} \right.$$

这样便有:

$$v = \frac{c}{n} = \frac{1}{\sqrt{\varepsilon \mu}} \quad , \quad k = \frac{\omega}{v} = \omega \sqrt{\varepsilon \mu}$$

但是, 如果介质漏电, 且符合 $\mathbf{J} = \sigma \mathbf{E}$, 那么以上方程就改写为:

$$\left\{ \begin{array}{l} \nabla \cdot \mathbf{E} = 0 \\ \nabla \times \mathbf{E} + \frac{\partial \mathbf{H}}{\partial t} = \mathbf{0} \\ \nabla \cdot \mathbf{H} = 0 \\ \nabla \times \mathbf{H} - \varepsilon \mu \frac{\partial \mathbf{E}}{\partial t} = \mu_0 \sigma \mathbf{E} \end{array} \right.$$

¹事实上, 只需要亚波长的结构便可造成衰减.

对最后一个式子所引发的不同在光学情况下是很好处理的. 因为单色光总是具有固定的频率 ω , 那么其实以上式子无非是把 $\frac{\partial}{\partial t}$ 变成 $-i\omega$. 在磁导率被认为几乎等于真空磁导率的情形下, 这相当于说:

$$\left\{ \begin{array}{l} \nabla \cdot \mathbf{E} = 0 \\ \nabla \times \mathbf{E} - i\omega \mathbf{H} = \mathbf{0} \\ \nabla \cdot \mathbf{H} = 0 \\ \nabla \times \mathbf{H} + i\omega \varepsilon \mu \mathbf{E} = \mathbf{0} \end{array} \right. \xrightarrow{\varepsilon \rightarrow \varepsilon + i\frac{\sigma}{\omega}} \left\{ \begin{array}{l} \nabla \cdot \mathbf{E} = 0 \\ \nabla \times \mathbf{E} - i\omega \mathbf{H} = \mathbf{0} \\ \nabla \cdot \mathbf{H} = 0 \\ \nabla \times \mathbf{H} + i\omega \varepsilon \mu \mathbf{E} - i\sigma \mu \mathbf{E} = \mathbf{0} \end{array} \right.$$

所以对于漏电介质², 通常会有复电容率(complex permittivity) 的说法³, 它就是以上把电容率和电导率合并以后的新的复常数, 用它第四个方程就与真空中的方程没有任何区别了, 除了系数是一个复数:

$$\varepsilon' = \varepsilon + i\frac{\sigma}{\omega} : \quad \nabla \times \mathbf{H} + i\omega \varepsilon' \mu \mathbf{E} = \mathbf{0}$$

这样一个方程的解是可以完全照搬之前的解的, 因为数学上可以证明复数解具有可解析延拓(analytic continuation) 的特性. 从而容易发现, 复波矢就变为:

$$\mathbf{k}^2 = \omega^2 \varepsilon' \mu$$

这样就得到:

$$\alpha^2 - \beta^2 = \omega^2 \varepsilon \mu, \quad 2\alpha \cdot \beta = 2\alpha\beta \cos \theta = \omega \sigma \mu > 0$$

此时传播波矢 α 与衰减波矢 β 就不一定要垂直了, 它们必须夹锐角, 也就是说, 如果电磁波在漏电介质或者导体中传播, 沿传播方向必须要衰减. 我们最后引入光学中使用最多的复折射率(complex refraction index), 按照原来的看法它意味着 $k^2 = n^2 k_0^2 = n^2 \frac{\omega^2}{c^2} = \frac{\omega^2}{v^2}$. 现在要更小心些, 因为 $\mathbf{k} = \alpha + i\beta$ 已经包含两个不贡献的波矢部分. 故我们先对以下表达式开方:

$$\alpha^2 - \beta^2 + 2i\alpha\beta \cos \theta = \omega^2 \varepsilon' \mu$$

$$\Rightarrow \sqrt{\alpha^2 - \beta^2 + 2i\alpha\beta \cos \theta} = n \frac{\omega}{c} = (n_1 + i n_2) \frac{\omega}{c}$$

通过以上两式对比, 我们能得到两个方面. 第一是复折射率实部 n_1 与虚部 n_2 分别是这样依赖于相对介电常数 $\varepsilon_r = \varepsilon/\varepsilon_0$ 和电导率 σ 的:

$$n_1 = \sqrt{\varepsilon_r} \cdot \sqrt{\frac{1}{2} \left(1 + \frac{\sigma^2}{\omega^2 \varepsilon^2} + \sqrt{1 + \frac{\sigma^2}{\omega^2 \varepsilon^2}} \right)}$$

$$n_2 = \frac{\sigma}{\sqrt{\varepsilon_r} \omega \varepsilon_0} \cdot \sqrt{\frac{1}{2} \left(1 + \frac{\sigma^2}{\omega^2 \varepsilon^2} + \sqrt{1 + \frac{\sigma^2}{\omega^2 \varepsilon^2}} \right)^{-1}}$$

第二组关系式如果已知了介质的两个折射率, 那么在介质中传播的波的两个波矢 α 和 β 需要满足的关系为:

$$\alpha^2 - \beta^2 = (n_1^2 - n_2^2) \frac{\omega^2}{c^2}$$

$$\alpha\beta \cos \theta = n_1 n_2 \frac{\omega^2}{c^2}$$

只有当 $\theta = 0$ 才恰有:

$$\alpha = n_1 \frac{\omega}{c}, \quad \beta = n_2 \frac{\omega}{c}$$

²即使不漏电, 也会由于有损耗而等效于有复电容率.

³对于普通的导体, 也一样可以讨论复电容率, 或者更常见地, 复电导率(complex conductivity). 它不仅包含由于原子实部分极化导致的电容性, 还要包含由于电子运动惯性导致的电感性.

11.1.2 经典电子论的解释

下面我们来介绍历史上发挥了重要作用的经典电子论。它虽然不够精确但物理图像十分重要。

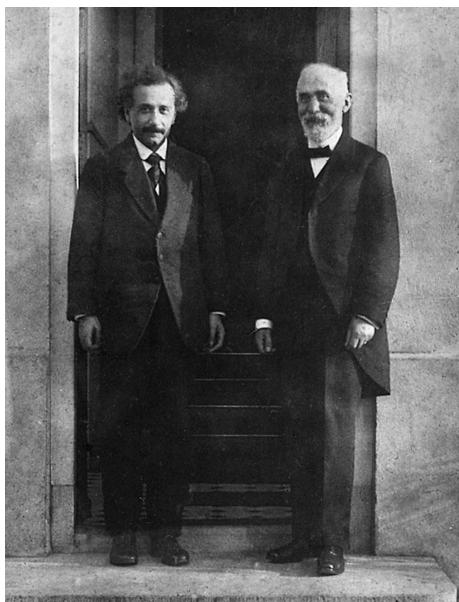


图 11.2: Einstein 与 Lorentz

历史上虽然电子作为粒子的发现是 1897 年汤姆孙 (J. J. Thomson) 的工作。但是早在半个世纪前人们就开始逐渐相信物质由两种带电粒子构成并在导电时可以移动的模型了。人们用这样的模型去想象输运现象, 电磁感应, 热电耦合等现象并取得了丰硕的成果, 经典电磁学早在 1860 年代便在麦克斯韦 (J. C. Maxwell) 的工作中达到顶点进入尾声。剩下来的几十年时间历史让位给了洛伦兹 (H. A. Lorentz), 洛伦兹对电磁学的贡献是深远的。他承前人之大成, 讨论电磁场与电磁介质的相互作用, 讨论电磁学原理与古老的伽利略式相对性原理的结合; 又开近代物理之先河, 电子论自然地导向了量子理论, 而电磁学又自然地促成了狭义相对论的诞生。事实上, 早年间狭义相对论在学界被介绍时便是以洛伦兹和爱因斯坦这一对忘年交命名的洛伦兹-爱因斯坦理论。

电子, 这样一个基本粒子可以说是妇孺皆知的概念, 它带一个单位的负电荷, 其基本参数为:

$$e = 1.602176634 \times 10^{-19} \text{C} \quad m_e = 9.10938356 \times 10^{-34} \text{kg}$$

然而出于各种原因, 在经典电子论中的电子却又是陌生的概念。根据洛伦兹的经典电子论精神, 重的原子实被认为静止, 轻的束缚电子被认为在外场中可以运动, 且主要受电场的影响⁴。电子在外电场中相对原子实的位移造成了电偶极矩。出于好的近似, 将电子视作被绑在连接中心的弹簧的另一端的质点。质量为 m 不再被认为与基本粒子的质量一致, 事实上这个质量的本质实际上是虚无的, 因为当考虑到原子与电磁场相互作用的严格理论时图像将完全是量子的而超出了此处可以讨论的范围, 这样的一个为了解释现象去构造的模型就叫唯象模型 (phenomenological model)。但是其电荷量还是 $-e$, 弹簧劲度系数为 k 。这样子我们就可以写出电子在外电场中的动力学方程:

$$m\ddot{\mathbf{r}} + k\mathbf{r} = -e\mathbf{E}$$

但是我们还忽略了一点, 加速运动的电荷会产生辐射, 也会因为辐射而带走能量和动量, 这就被视作介质对电磁波吸收, 散射的根本原因。也同样为了唯象地描述它, 我们认为电子是受到了一个辐射阻尼力 $\mathbf{f} = -\gamma\dot{\mathbf{r}}$ 。这样以上方程就被修改为:

$$m\ddot{\mathbf{r}} + \gamma\dot{\mathbf{r}} + k\mathbf{r} = -e\mathbf{E}$$

命 $\omega_0 = \sqrt{k/m}$ 代表共振频率 (resonance frequency) 或待会就会说明的吸收频率 (absorption frequency)。再命入射电磁波为频率 ω 的光, 按光学符号约定 $\mathbf{E} = \mathbf{E}_0 e^{-i\omega t}$, 那么很容易解出受迫振动解 $\mathbf{r} = \mathbf{r}_0 e^{-i\omega t}$, 振幅为:

$$\mathbf{r}_0 = \frac{-e\mathbf{E}_0/m}{\omega_0^2 - \omega^2 - i\gamma\omega/m}$$

我们现在就能算出动态的分子极化率来, 它被定义为分子中 Z 个价电子产生的偶极矩 $\mathbf{p} = -Zer$ 与外电场 \mathbf{E} 的比值, 可以发现它也是个复数, 依赖于外光场的频率:

$$\alpha(\omega) = \frac{Ze^2/m}{\omega_0^2 - \omega^2 - i\gamma\omega/m}$$

我们知道, 对于具体介质的极化总是与各个单元的极化有关。我们在此讨论稀薄的无固有偶极矩的气体

⁴根据洛伦兹力公式 $\mathbf{F} = q(\mathbf{E} + \mathbf{v} \times \mathbf{B})$, 而电磁场中 $E = cB$, 故 $v \ll c$ 时磁场力可以忽略。

在外电磁波中的极化⁵, 此时不难想到介质的介电常数应该直接依赖于分子极化率, 事实上, 极化强度为:

$$\mathbf{P} = \chi \varepsilon_0 \mathbf{E} = (\varepsilon_r - 1) \varepsilon_0 \mathbf{E} = n \mathbf{p}$$

从而得到复电容率的值:

$$\varepsilon' = \varepsilon_0 + \frac{n Z e^2 / m}{\omega_0^2 - \omega^2 - i \gamma \omega / m}$$

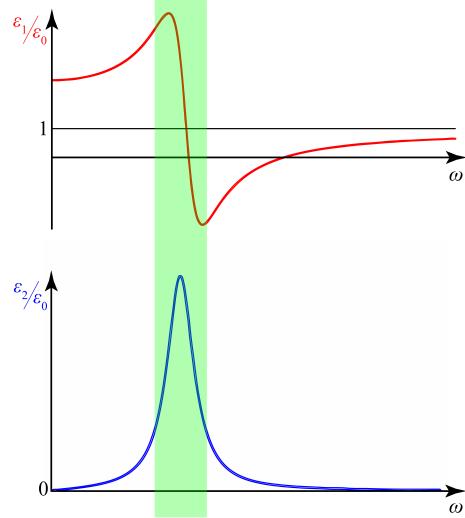
通常, 把 $n Z e^2 / \varepsilon_0 m$ 称作等离子体频率(plasma frequency)的平方 ω_p^2 . 那么上式也作:

$$\varepsilon' = \varepsilon_0 + \frac{\omega_p^2}{\omega_0^2 - \omega^2 - i \gamma \omega / m} \varepsilon_0$$

出现虚部其实就意味着损耗, 在振动的过程中电场在持续不断地对振子输入能量而导致了电磁波的衰减. 只不过这里的损耗并不是导电行为导致的. 写出复电容率的实部和虚部 $\varepsilon' = \varepsilon_1 + i \varepsilon_2$:

$$\varepsilon_1 = \varepsilon_0 \cdot \left[1 + \frac{\omega_p^2(\omega_0^2 - \omega^2)}{(\omega_0^2 - \omega^2)^2 + (\gamma \omega / m)^2} \right]$$

$$\varepsilon_2 = \varepsilon_0 \cdot \frac{\gamma \omega_p^2 \omega / m}{(\omega_0^2 - \omega^2)^2 + (\gamma \omega / m)^2}$$



两个函数看上去复杂, 其实都属于近似的洛伦兹型函数(lorentzian function). 图像如右图所示. 在绿色带状区域这个函数有着较大的转折. 这发生在 $\omega \approx \omega_0$ 处. 若做近似, 认为反应损耗的 γ 较小. 则令 $\omega^2 - \omega_0^2 = x, \gamma \omega_0 / m = a$, 那么上式在带状区域附近近似为标准的洛伦兹型:

$$\varepsilon_1 / \varepsilon_0 = 1 - \omega_p^2 \cdot \frac{x}{x^2 + a^2}$$

$$\varepsilon_2 / \varepsilon_0 = \omega_p^2 \cdot \frac{a}{x^2 + a^2}$$

电容率的实部在带状区域由区域外的缓慢增加转为剧烈减少, 而恰好在这样一个区域, 电容率的虚部突然变得很大. 我们习惯上用虚部这种洛伦兹函数增加到最大值的一半的两个点之间的间距作为特征的半高峰宽(full width at half maximum, FWHM). 对于洛伦兹型, 其值恰为 $x = \pm a$ 之间的间距 $2a$, 但是注意到要转化为 ω 对应的间距, 它恰好为:

$$\Gamma = \frac{\gamma}{m}$$

我们可以从这个模型中清楚地看到两个物理现象: 色散(dispersion)与吸收(absorption). 前者就是说折射率依赖于波长, 这和复电容率的实部依赖于频率只是说法不同本质一样. 而复电容率的虚部, 根据之前的讨论, 也就代表吸收. 我们发现在入射光的频率近似为谐振子的固有频率 ω_0 时, 对应经典力学中发生共振的区域附近, 色散和吸收都变得很显著. 我们知道, 经典电子论给出的解释仅仅是一个唯象模型, 所以相关的参数都要根据实验结果来确定. ω_0 可以根据发生强烈吸收与色散的波长来确定, 而 γ/m 则根据吸收峰的半高峰宽 Γ 来确定, 最后等离子体频率 ω_p 则比较有趣, 我们恰好可以根据零频率处的折射率 n_0 来确定它:

$$n_0^2 = \varepsilon_1(0) / \varepsilon_0 = 1 + \omega_p^2 / \Gamma^2$$

那么实际情况是否这么简单呢? 至少通过对右图的常温下水的色散的测量中我们发现, 除了在可见光波段水几乎是透明的而具有大约 1.33 的折射率, 在近红外(0.8-2.5 μm)与中红外(2.5-15 μm)波段的短波区定性上符合以上模型推导出来的结果. 但是中红外到远红外(15-1000 μm)则明显偏离以上结果.

⁵如果是液体或固体, 一方面由于不同的偶极子间的相互影响不可忽略导致公式需要修正, 也因为极化方式还有取向极化等所以需要修正, 但全都只影响定量结果, 定性的图像仍然是适用的.

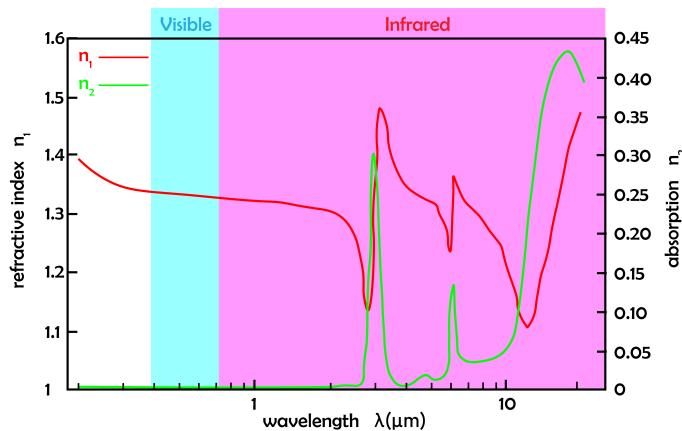


图 11.4: 水的可见-红外色散曲线

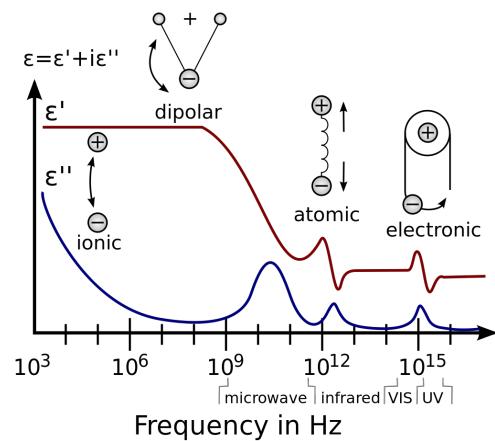


图 11.5: 光与原子作用类型

事实上,在不同的波段,适用于不同的光与物质相互作用的基本模型.微波波段光波长足够长使得晶体的正负离子做大范围的整体相对运动⁶,就是晶格振动.红外则开始使得分子振动,包括转动和振动等形式.再往下的可见波段才是电子共振,也包括原子间的能带共振和原子内的能级共振.最后在紫外波段,电子甚至能直接被电离,这就造成了光与物质相互作用问题的复杂性.

但无论哪种相互作用的机制,我们上述推导得到的描述有着特定共振频率的色散与吸收的最终公式是十分普适的.它只有三个待定的参数:共振频率 ω_0 , 半高峰宽 Γ 和反应共振强度的等离子体频率 ω_p . 我们接下来要做的,是把不同振子,不同吸收效应带来的结果进行求和:

$$\begin{aligned}\varepsilon_1/\varepsilon_0 &= 1 + \omega_p^2 \cdot \sum_i \frac{c_i(\omega_i^2 - \omega^2)}{(\omega_i^2 - \omega^2)^2 + \Gamma_i^2 \omega^2} \\ \varepsilon_2/\varepsilon_0 &= \omega_p^2 \cdot \sum_i \frac{c_i \Gamma_i \omega}{(\omega_i^2 - \omega^2)^2 + \Gamma_i^2 \omega^2}\end{aligned}$$

其中 ω_p 仍然统计各个吸收峰的总强度,而 c_i 系数的和必须为 1 则统计各个吸收峰的相对强度.我们最后计算远离吸收带处的折射率值.忽略各个 Γ_i 后得到:

$$n^2 = 1 + \omega_p^2 \cdot \sum_i \frac{c_i}{\omega_i^2 - \omega^2}$$

上式就是著名的塞尔迈耶尔方程(Selmeier's equation),如果用波长表示则为:

$$n^2 = 1 + \sum_i a_i \cdot \frac{\lambda^2}{\lambda^2 - \lambda_i^2}, \quad a_i = \frac{c_i \omega_p^2 \lambda_i^2}{4\pi^2 c^2}$$

上式依然可以做近似,把共振的各个波长 λ_i 从小到大按 i 来排列.不妨设 $\lambda_i \ll \lambda_{i+1}$,而 λ 恰好接近 λ_i ,那么忽略所有 λ_{i+1} 及之后对应的小量项并将 λ_{i-1} 及之前的项直接近似为 a_i :

$$\begin{aligned}n^2 &\approx 1 + a_1 + a_2 + \cdots + a_{i-1} + \frac{a_i \lambda^2}{\lambda^2 - \lambda_i^2} \\ &= 1 + a_1 + a_2 + \cdots + a_{i-1} + a_i + \frac{a_i \lambda_i^2}{\lambda^2 - \lambda_i^2} \\ &\approx 1 + a_1 + a_2 + \cdots + a_{i-1} + a_i + \frac{a_i \lambda_i^2}{\lambda^2} \\ &= A_0 + \frac{B_0}{\lambda^2}\end{aligned}$$

⁶十分类似于热运动的那种形式,区别在于这是外场诱导的有规律的振动,而且正负电荷位移一定反相,对应热振动的特定光学支.

而进一步近似得到:

$$\sqrt{A_0 + \frac{B_0}{\lambda^2}} = \sqrt{A_0} \left(1 + \frac{B_0}{A_0 \lambda^2}\right)^{1/2} = \sqrt{A_0} \left(1 + \frac{B_0}{2A_0 \lambda^2} - \frac{B_0^2}{8A_0^2 \lambda^4} + \dots\right)$$

一般就总结为如下柯西色散公式(Cauchy dispersion formula):

$$n(\lambda) = A + \frac{B}{\lambda^2}$$

有时为了精度原因再加上四阶项 $+C/\lambda^4$. 而从推导过程中我们发现系数之间的联系:

$$A = \sqrt{1 + \sum_1^i a_i}, \quad B = \frac{a_i \lambda_i^2}{2A}$$

11.1.3 * 辐射阻尼与散射截面

我们此前得到了电荷做加速运动时需要满足的辐射能流和拉莫尔公式:

$$S = \frac{q^2 \ddot{\mathbf{r}}^2}{8\pi^2 \varepsilon_0 c^3 R^2} \sin^2 \theta, \quad P = \frac{q^2 \ddot{\mathbf{r}}^2}{6\pi \varepsilon_0 c^3}$$

从这个式子中可以看出, 在有加速度 $\ddot{\mathbf{r}}$ 的情况下, 辐射场的动量分布是关于电荷中心对称的. 故虽然会造成电荷的动能损失, 但是似乎不造成电荷的动量改变. 但是这一点其实不够准确. 因为偶极辐射本身过程中就包含有取最低阶项的近似成分. 本来可以精确计算的非零发射动量在这个过程中完全有可能被忽略. 而洛伦兹和亚伯拉罕在 1892 年通过将点电荷模型扩大为一个半径很小的球体证明了球体自己产生的辐射场会给自己一个动量的改变. 其中一部分解释为自带场产生的质量对应的跟随电荷运动的动量改变. 另一部分则就对应一个辐射阻尼力(radiation damping force), 后人称之为洛伦兹-亚伯拉罕力(Lorentz Abraham force). 后来量子电动力学进一步将其表达式精确化, 它为:

$$\mathbf{f} = +\frac{q^2}{6\pi \varepsilon_0 c^3} \dot{\mathbf{r}}$$

这样一个力却不大可能代表电荷与场这样一个无限多自由度⁷系统的全部. 注意到以下量具有时间量纲, 我们可以把电荷受到一个外力 \mathbf{F} 并与辐射场相互作用下电荷符合的动力学方程在仅考虑辐射阻尼力的情况下简化为:

$$\tau = \frac{q^2}{6\pi \varepsilon_0 m c^3}$$

$$m\ddot{\mathbf{r}} = \mathbf{F} + m\tau \ddot{\mathbf{r}}$$

这样一个简化就会招致祸端: 若在一维情况下, 将受力设置为以下元脉冲函数(unit pulse function):

$$F(t) = \begin{cases} 0 & (t < -0.5) \\ 1 & (-0.5 \leq t \leq 0.5) \\ 0 & (t > 0.5) \end{cases}$$

通过解数值方程 $\tau \ddot{x} - \ddot{x} = -F(t)/m$, 把 $a = \ddot{x}$ 视作待解函数, 可得通解:

$$a(t) = \begin{cases} A_1 e^{t/\tau} & (t < -0.5) \\ A_2 e^{t/\tau} + \frac{1}{m} & (-0.5 \leq t \leq 0.5) \\ A_3 e^{t/\tau} & (t > 0.5) \end{cases}$$

⁷每一点的场强全都是独立的变量. 类似于一张连续的弹性膜, 膜的位移就相当于场的激发.

且三个系数满足使 a 为连续函数的条件:

$$A_1 - A_2 = \frac{e^{0.5/\tau}}{m}, \quad A_3 - A_2 = \frac{e^{-0.5/\tau}}{m}$$

从中可以看出 $A_3 \neq A_1$. 这就导致了一个尴尬的结果: 在 $t < -0.5, t > 0.5$ 两段过程中, 至少有一段过程电荷做非匀速直线运动. 通过因果律来分析, 当力尚未作用在电荷上时, 电荷自然没有道理产生加速度. 故应取 $A_1 = 0$. 那么在这个外力作用于电荷一段时间后并撤销后, 这个理论预言电荷已经被一个扰动后的电磁场挟持以致做以下运动:

$$A_3 = -\frac{e^{0.5/\tau} - e^{-0.5/\tau}}{m}$$

$$a = -\frac{e^{0.5/\tau} - e^{-0.5/\tau}}{m} \cdot e^{t/\tau}$$

无疑, 这是荒诞的, 加速度不仅不是零, 甚至还随时间趋近于无穷大, 这种病态的行为显然没有在实际观测中出现过. 其实, 根据这个辐射阻尼力的推导过程中忽略的高阶项需要足够小的要求来看, 我们上面定义的 τ 恰好界定了这个近似可以进行的条件, 那便是我们关心的过程的 $\ddot{\mathbf{r}}$ 变化的特征时间 T 需要不大于对电子以上定义的特征时间的量级:

$$T \ll (\sim) \tau = \frac{q^2}{6\pi\varepsilon_0 mc^3} = 6.27 \times 10^{-24} \text{ s}$$

在知道这个辐射阻尼力的做法的适用条件之后我们再来看一个常见的简单推导方法. 我们注意到辐射阻尼公式与辐射功率拉莫尔公式系数的高度相似性. 不难去思考两者的联系, 实际上:

$$\frac{q^2 \ddot{\mathbf{r}}^2}{6\pi\varepsilon_0 c^3} = \frac{d}{dt} \left(\frac{q^2 \dot{\mathbf{r}} \cdot \ddot{\mathbf{r}}}{6\pi\varepsilon_0 c^3} \right) - \frac{q^2 \ddot{\mathbf{r}} \cdot \dot{\mathbf{r}}}{6\pi\varepsilon_0 c^3}$$

可见辐射阻尼力对粒子做负功的功率与拉莫尔辐射功率之间仅仅相差一项. 我们指出在我们研究的电磁波照射到振子的经典色散模型中, 电子做周期性运动, 故这一项的周期积分为零:

$$\int_0^T \frac{d}{dt} \left(\frac{q^2 \dot{\mathbf{r}} \cdot \ddot{\mathbf{r}}}{6\pi\varepsilon_0 c^3} \right) dt = \frac{q^2 \dot{\mathbf{r}} \cdot \ddot{\mathbf{r}}}{6\pi\varepsilon_0 c^3} \Big|_0^T = 0$$

故如果一个表达式能代表辐射阻尼, 它必然具有上式具有的形式. 而在 $T \ll (\sim) \tau$ 时, 正如我们之前介绍的那样, 这个做法是合理的.

利用特征时间 τ , 在频率 ω 的电磁波下的谐振子的动力学方程, 考虑辐射阻尼时, 就写作:

$$m\ddot{\mathbf{r}} - m\tau\ddot{\mathbf{r}} + k\mathbf{r} = -e\mathbf{E}$$

这看上去与之前我们给出的谐振子的形式存在出入. 实际上, 在简谐 $\mathbf{E} = \mathbf{E}_0 e^{i\omega t}$ 的情况下, 对时间导数相当于乘以 $i\omega$. 于是上式写为:

$$m\ddot{\mathbf{r}} + m\omega^2 \tau \dot{\mathbf{r}} + k\mathbf{r} = -e\mathbf{E}$$

这就可以看出, 实际上之前的经典色散理论不过是将 $m\omega^2 \tau$ 整体取做 γ 而已. 引入辐射阻尼的现象之后我们就可以看出 γ 与角频率潜在的比例关系 $\gamma \propto \omega^2$.

在这个基础上我们解出单电子某维度上做振动的函数:

$$x = Ae^{i\omega t} = -\frac{eE_x/m}{\omega_0^2 - \omega^2 + i\omega^3 \tau}$$

通过这个振幅我们除了推导色散, 还可以得到两个十分有意义的结果:

第一个结果是重新去计算电子辐射的功率, 即把外电磁场能量单位时间散射的那部分能量. 根据偶极振荡公式:

$$\bar{P} = \frac{e^2 \omega^4 |x|^2}{12\pi\varepsilon_0 c^3}$$

而入射电磁波,如果就是沿 z 方向传播,沿 x 方向偏振的平面波,则具有平均能流:

$$S = \frac{1}{\mu_0 c} \cdot \frac{|E_x|^2}{2} = \frac{\epsilon_0 c |E_x|^2}{2}$$

用这个功率除以这个能流,我们得到一个面积量纲的物理量.实际上就是相当于电子吸收了这个面积里面的辐射并以散射的形式做了新的辐射发射.这个面积称作**散射截面**(scattering cross section),计算可得:

$$\sigma = \frac{\bar{P}}{S} = \frac{8}{3} \pi r_e^2 \cdot \frac{1}{\omega^2 \tau^2 + (1 - \omega_0^2/\omega^2)^2}$$

而其中 r_e 就是我们之前提过的**经典电子半径**(classical electron radius).它恰恰是在洛伦兹等人在研究辐射的问题时提出的.故也称**洛伦兹半径**(Lorentz radius):

$$r_e = \frac{e^2}{4\pi\epsilon_0 mc^2}$$

我们首先注意一点:由于 τ 是个很短的时间,故能够使得 $\omega\tau \sim 1$ 的角频率对应的光波具有以下的量级:

$$\omega \sim 10^{23} \text{ rad/s} , \quad \lambda \sim 12 \text{ fm} , \quad E \sim 105 \text{ MeV}$$

这实际上是一个极高的频率了,到这个能标,光子对电子的影响已经完全不能忽略光子的量子性,当然也要考虑电子的波动性,康普顿散射和量子电动力学的理论时必要的.而物质形成的体系总是有 $\omega_0\tau \ll 1$,故即使在 $\omega \gg \omega_0$,一般也可以把 $\omega\tau$ 视作小量.那么在这种场合下:

$$\sigma \approx \frac{8}{3} \pi r_e^2$$

这个截面称为**汤姆孙截面**(Thomson cross section).反映了在高频电磁波作用下电子接近自由电子散射电磁波的行为.

而在低频下 $\omega \ll \omega_0$,可做近似:

$$\sigma \approx \frac{8}{3} \pi r_e^2 \cdot \frac{\omega^4}{\omega_0^4}$$

这个式子后面会介绍,反映了瑞利散射的基本原理.

最后,在 $\omega \approx \omega_0$ 时,截面表达式的分母很接近零.散射截面远大于汤姆孙截面,电子强烈吸收外电磁场能量,并造成大的散射.这称作**共振散射**(resonance scattering).实验上实际上也观察到了这个现象.但是洛伦兹经典色散理论的弱点就在于它是个经典的理论,并不能很好地解释实验中得到的数据,因为它没有构建出理论模型中谐振子的量子本质.

第二个结果,我们来考察一下这个振子在外界电磁场下产生的振动能量:

$$\begin{aligned} K &= \overline{\frac{1}{2}k\Re e(x)^2 + \frac{1}{2}m\Re e(\dot{x})^2} \\ &= \frac{1}{4}m(\omega^2 + \omega_0^2)|x|^2 \\ &= \frac{e^2}{2\epsilon_0 m \omega^2} \cdot \frac{\omega^2(\omega^2 + \omega_0^2)}{(\omega^2 - \omega_0^2)^2 + \omega^6 \tau^2} \cdot \frac{1}{2}\epsilon_0 |E_x|^2 \end{aligned}$$

这样也许还不足以引起理论上的价值.但是如果设想一个复杂的电磁波场,它总可以分解为简单的平面波场(参考下一节介绍的傅里叶变换理论).那么该点的瞬时 x 方向电场就可以分解为:

$$E_x(t) = \int_0^\infty [A(\omega) \cos \omega t + B(\omega) \sin \omega t] d\omega$$

而对于波 $A(\omega) \cos \omega t$ 和 $B(\omega) \sin \omega t$,在谐振子上激发的能量是:

$$K(\omega) = \frac{e^2}{2\epsilon_0 m \omega^2} \cdot \frac{\omega^2(\omega^2 + \omega_0^2)}{(\omega^2 - \omega_0^2)^2 + \omega^6 \tau^2} \cdot \frac{1}{2}\epsilon_0 [A(\omega)^2 + B(\omega)^2]$$

作为原因,多个谐振电场的叠加产生的结果就是简单的振子的振动的简单叠加,这是因为谐振电场与振子的振动满足的微分方程是线性的.而按照简谐分解的一个数学结果就是,能量也可以直接叠加:

$$\overline{\frac{1}{2}\varepsilon_0 E_x^2(t)} = \int_0^\infty \left[\frac{1}{2}\varepsilon_0 A(\omega)^2 \overline{\cos^2 \omega t} + \frac{1}{2}\varepsilon_0 B(\omega)^2 \overline{\sin^2 \omega t} \right] d\omega = \int_0^\infty \frac{1}{2} \frac{1}{2} \varepsilon_0 [A(\omega)^2 + B(\omega)]^2 d\omega$$

从这个角度来看,振子就会因此具有 x 方向的平均能量:

$$\begin{aligned} K_x &= \int K(\omega) d\omega \\ &= \int \frac{e^2}{2\varepsilon_0 m \omega^2} \cdot \frac{\omega^2 (\omega^2 + \omega_0^2)}{(\omega^2 - \omega_0^2)^2 + \omega^6 \tau^2} \cdot \frac{1}{2} \varepsilon_0 [A(\omega)^2 + B(\omega)]^2 d\omega \end{aligned}$$

如果考虑一个热力学系统,辐射和谐振子达到了一个温度为 T 的热平衡.那么经典理论认为:每一个振子在一个方向上应当具有平均能量 kT .这一点之后由普朗克发现是错误的.但是如果我们将坚持这个做法,由于电磁辐射的内能密度亦可以根据频率展开:

$$\begin{aligned} u &= \overline{\frac{1}{2}\varepsilon_0 E_x^2(t)} + \overline{\frac{1}{2}\varepsilon_0 E_y^2(t)} + \overline{\frac{1}{2}\varepsilon_0 E_z^2(t)} + \overline{\frac{1}{2\mu_0} B_x^2(t)} + \overline{\frac{1}{2} B_y^2(t)} + \overline{\frac{1}{2\mu_0} B_z^2(t)} \\ &= 6 \cdot \overline{\frac{1}{2}\varepsilon_0 E_x^2(t)} \\ &= \int_0^\infty u(\omega d\omega) \end{aligned}$$

这就得到了:

$$u(\omega) = 3 \cdot \frac{1}{2} \varepsilon_0 [A(\omega)^2 + B(\omega)]^2$$

从而之前的积分就可用 $u(\omega)$ 表示.但是又由于 τ 很小,故仅在 $\omega \sim \omega_0$ 附近近似:

$$\begin{aligned} K_x &= \int \frac{e^2}{2\varepsilon_0 m \omega^2} \cdot \frac{\omega^2 (\omega^2 + \omega_0^2)}{(\omega^2 - \omega_0^2)^2 + \omega^6 \tau^2} \cdot \frac{u(\omega)}{3} d\omega \\ &= \int \frac{e^2}{2\varepsilon_0 m \omega_0^2} \cdot \frac{2\omega_0^4}{4\omega_0^4(\omega - \omega_0)^2 + \omega_0^6 \tau^2} \cdot \frac{u(\omega_0)}{3} d\omega \\ &= \frac{e^2 u(\omega_0)}{12\varepsilon_0 m \omega_0^2} \cdot \int \frac{dx}{x^2 + \omega_0^2 \tau^2 / 4} \\ &= \frac{e^2 u(\omega_0)}{12\varepsilon_0 m \omega_0^2} \cdot \frac{2\pi}{\omega_0 \tau} \\ &= kT \end{aligned}$$

代入 τ 的表达式,这就得到了:

$$u(\omega_0) = \frac{3kT \omega_0^3}{2\pi^2 c^3}$$

这就是辐射与振子热力学平衡的理论基础,从此建立的理论可以导出著名的瑞利-金斯公式,从而招致了著名的紫外灾难.普朗克提出了辐射的量子理论成功地解决了这个问题.参见下一章.

第十二章 量子论

我们用选自费曼(R. Feynman)先生的物理学讲义的开篇金句来作为本章与下一章近代物理内容的开头:

Each piece, or part, of the whole of nature is always merely an approximation to the complete truth. Therefore, things must be learned only to be unlearned again or, more likely, to be corrected. The test of all knowledge is **experiment**. Experiment is the sole judge of scientific “truth”.

的确,量子理论以其不直观而被近代早期物理学家们所疑惑,这其中不乏一些赫赫有名的大师.直到现在也有很多基础的问题是没有被深刻地理解的:电子的本性与内部结构,基本粒子的类别与参数,对量子非定域性与测量的理解...所以在关于可能会造成问题的领域的学习时,需要先明白实验上的事实,从中理解理论建立的必要性.

12.1 黑体辐射

对于**热辐射**(thermal radiation)的讨论是何时进入物理研究的视野的呢?可以肯定的是人类认识到热辐射现象非常的早:光芒万丈的太阳,烧红的木炭与金属都是典型的热辐射的情形.但人们掌握足够的方法去测量它则也是要到19世纪后半期了.热辐射势必涉及到电磁场与电荷的相互作用.而且深入到原子尺度,实际上就是电磁波的发射与吸收.对于电磁波的发射,我们在电磁学中粗略讲过,只要有加速运动的电荷就会导致电磁辐射.之后小节我们将认识到微观电荷不能用“加速”来描述,其状态其实是量子态,处于激发态才会自发向基态去跃迁放出电磁辐射.而对于电磁波的吸收,则在光学中我们简要介绍过洛伦兹电子论中的处理方法.

一个物体如果能够在任何温度下把照射到它上面的任何频率的光都全部吸收掉,那么这个物体就叫做**黑体**(black body).虽然在现实生活中这样的物体并不存在,但石墨和碳黑往往被视作比较理想的黑体,尽管这样,黑体“看上去”也不总是“黑的”.著名生物学家查尔斯·达尔文的一个亲戚,陶瓷师,托马斯·玮致活在1792年注意到所有物体,不论化学构成,形状,尺寸,几乎都在相同的温度下烧红.黑体也不例外.作为这个实验现象研究的推广,1859年**基尔霍夫**(G. Kirchhoff)利用热力学理论证明了著名的**基尔霍夫定律**:即

对于某角频率的光波,某温度下任意物体的辐射本领正比于吸收率:



图 12.1: 热辐射

$$e(\omega) = J(\omega, T)A(\omega)$$

其中, e 为单位面积单位时间单位角频率间隔辐射的能量,而 A 为对该频率光的吸收率. $0 < A < 1$.除了吸收以外就是反射.物体较薄时还存在透射.不同物体在相同的 T, ω 下可以有不同的 A, e ,但是其比值 J 却是一个**普适**(universal)的函数.即与物体本身没有关系.而黑体的定义为 $A(\omega)$ 恒等于1.也就是说明了

黑体的辐射本领:

$$e(\omega) = J(\omega, T)$$

之后人们的任务就非常清晰了: 测出不同 T 下的函数 $J(\omega, T)$. 而这个过程大致可以分为三个阶段:

1. 挖掘出 $J(\omega, T)$ 的整体特征.
2. 实验测量 $J(\omega, T)$ 的曲线与提出可能的理论模型.
3. 最终确定 $J(\omega, T)$ 的函数形式与理论解释.

$$J(\omega, T) = \frac{1}{4} u(\omega, T) c$$

$$u(\omega, T) = \frac{C_1 \omega^3}{e^{C_2 \omega/kT} - 1}$$

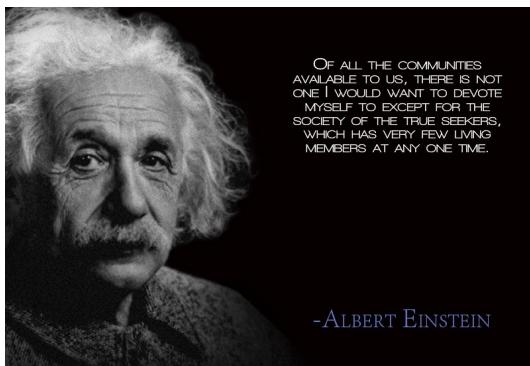


图 12.2: 真正的研究者

如何去解释这个式子呢? 从接触到黑体辐射谱的曲线开始, 普朗克用了足足六年的时间才得到一个合理的解释. 那是在 1901 年, 功夫不负有心人, 这位“真正的研究者”十分谨慎地发表了关于能量量子化的著名论文, 科学研究史首次叩响了量子理论的大门.

普朗克的工作大致可以如此解释. 我们

- 黑体辐射的性质:

– 总辐射强度:

$$J = \sigma T^4 = \frac{1}{4} u c$$

$$\sigma = 5.67 \times 10^{-8} \text{W}/(\text{m}^2 \cdot \text{K}^4) = \frac{2\pi^5 k^4}{15 h^3 c^2}$$

– 辐射频域谱: 普朗克公式:

$$J(\omega) = \frac{\hbar \omega^3}{4\pi^3 c^2} \frac{1}{e^{\frac{\hbar \omega}{kT}} - 1}$$

– 辐射空间谱: 郎伯定律:

$$\frac{dJ}{d\Omega} = \frac{J}{\pi} \cos \theta$$

- 非黑体满足基尔霍夫定律: 某角频率 ω 下发射率 $e(\omega)$ 与吸收率 $a(\omega)$ 必然相等. 发射率是

$$\begin{aligned} \frac{dE}{dSdt} &= \int E(\omega) d\omega \\ e(\omega) &= E(\omega)/J(\omega) \end{aligned}$$

故基尔霍夫定律发现:

$$e(\omega) = a(\omega)$$

12.2 光粒子性

- 光子概念:

$$E = \hbar \omega , \quad \mathbf{p} = \hbar \mathbf{k}$$

- 用于解释普朗克公式: 玻色-爱因斯坦统计.

- 用于解释光电效应:

$$E_k = h\nu - W \geq eU \rightarrow I$$

12.3 玻尔原子

- 精细结构常数(fine structure constant):

$$\alpha = \frac{e^2}{4\pi\varepsilon_0\hbar c} \approx \frac{1}{137}$$

- 半经典假设: 忽略电子的波动性, 认为仍然有轨道运动, 且为圆周运动, 符合牛顿力学:

$$m \frac{v_n^2}{r_n} = \frac{e^2}{4\pi\varepsilon_0 r_n^2}$$

- 量子化假设: 角动量是量子化的, 或者作用量量子化:

$$mv_n \cdot 2\pi r_n = nh$$

- 定态跃迁假设: 基态 $n = 1$ 是能量最低的稳定态, 其他态与基态彼此之间都有可能发生转变, 转变是瞬间发生的, 伴随着一个光子的吸收或发射, 且需要能量守恒:

$$E_n = \frac{1}{2}mv_n^2 - \frac{e^2}{4\pi\varepsilon_0 r_n}$$

$$h\nu = E_n - E_m$$

- 玻尔理论的解:

$$r_n = n^2 a_0$$

$$v_n = \frac{\alpha c}{n}$$

$$E_n = -\frac{E_0}{n^2} = -\frac{\alpha^2 mc^2}{2n^2}$$

其中 E_0 为第一电离能, 氢原子为 13.6 eV. 而 a_0 为玻尔半径, 它与另外两个特征长度: 电子康普顿波长:

$$\lambda_e = \frac{h}{mc}$$

经典电子半径:

$$r_e = \frac{e^2}{4\pi\varepsilon_0 mc^2}$$

符合;

$$\alpha = \frac{\lambda_e/2\pi}{a_0} = \frac{a_0}{r_e}$$

12.4 物质波与波函数

- 匀速运动的粒子对应一个平面波;

$$\psi = A e^{i(\omega t - \mathbf{k} \cdot \mathbf{r})}$$

且符合德布罗意关系:

$$E = \hbar\omega$$

$$\mathbf{p} = \hbar\mathbf{k}$$

- 对这个波的诠释: 模方表示粒子在某空间点被探测到的概率分布函数:

$$f(\mathbf{r}) = |\psi|^2$$

$$\int f(\mathbf{r}) dV = 1$$

- 测不准原理: 平面波情况下粒子位置是不确定的, 但是动量却唯一. 任意一个波函数可以分解为不同概率的平面波的叠加, 那么某方向上位置和动量的两个不确定度(标准差)满足:

$$\Delta x \Delta p \geq \frac{\hbar}{2}$$