

引 言

1948 年巴丁 (Bardeen), 布拉顿 (Brattain) 和肖克莱 (Shockley) 发明了晶体管, 带来了现代电子学的革命, 同时也促进了半导体物理研究的蓬勃发展. 从那以后的几十年间, 无论在半导体物理研究方面, 还是半导体器件应用方面都有了飞速的发展, 半导体微电子技术和半导体光电子技术已经成为现代社会的重要技术基础, 对今后人类文明的发展将产生深远的影响.

1954 年, 半导体有效质量理论的提出^[1,2]是半导体理论的一个重大发展, 它定量地描述了半导体导带和价带边附近细致的能带结构, 给出了研究半导体中浅杂质 (施主和受主) 能级、激子能级、磁能级等的理论方法, 促进了当时的回旋共振、磁光吸收、自由载流子吸收、激子光谱等实验研究.

1958 年集成电路问世.

1959 年赝势概念的提出^[3], 使得固体能带的计算大为简化. 利用价电子态与原子核心态正交的性质, 用一个赝势代替真实的原子势, 得到了一个固体中价电子态满足的方程. 用赝势方法得到了几乎所有半导体的比较精确的能带结构.

1962 年半导体激光器发明, 1970 年半导体激光器室温连续工作得到实现.

1968 年硅 MOS 器件发明及大规模集成电路实现产业化大生产.

1970 年超高真空表面能谱分析技术相继出现, 开始了对半导体表面、界面物理的研究, 其中包括: 硅表面的 7×7 表面再构问题, 金属与 III-V 族化合物界面肖特基势垒形成起因, CoSi/硅和金属/硅界面性质, 费米能级钉扎等问题. 特别是 1982 年 G. Binnig

等发明的扫描隧道显微镜 (STM) 技术, 不仅能对物质表面原子的几何排列和表面形貌进行直接观察, 而且还可以获得表面价键、能隙等电子结构信息。

70 年代初期, 江崎和朱兆祥^[4]基于试图人为地控制半导体中电子的势分布和波函数的设想, 首次提出了半导体超晶格的新概念。与此同时, 分子束外延技术也在美国贝尔实验室和 IBM 公司开发成功。新思想和新技术的巧妙结合, 制成了第一个晶格匹配组分型的 $\text{Al}_x\text{Ga}_{1-x}\text{As}/\text{GaAs}$ 超晶格^[5], 标志着半导体材料的发展开始进入人工设计的新时代。1978 年 Dingle 等人^[6]对异质结中二维电子气沿平行于界面的输运进行了研究, 发现了电子迁移率增强现象。在以后几年中, 由于工艺的改进, 将二维电子气的迁移率提高了近三个数量级, 导致了高电子迁移率晶体管 (HEMT) 的出现和为量子霍尔效应的发现创造了条件。1980 年德国的 Von Klitzing^[7]发现了整数量子霍尔效应, 1982 年崔崎等人^[8]又在电子迁移率极高的 $\text{Al}_x\text{Ga}_{1-x}/\text{GaAs}$ 异质结中发现了分数量子霍尔效应, 这是半导体物理的重大发现, 两个发现都获得了诺贝尔物理奖。由于超晶格, 量子阱对电子运动的限制效应, 1984 年 Miller 等人^[9]观察到量子阱中激子吸收峰能量随电场强度变化发生红移的量子限制斯塔克效应, 以及由激子吸收系数或折射率变化引起的激子光学非线性效应, 为设计新一代光双稳器件提供了重要的依据。

1990 年英国的 Canham^[10]首次在室温下观测到多孔硅的可见光光致发光, 使人们看到全硅光电子集成技术的新曙光。纳米微粒、纳米固体与纳米薄膜材料开辟了材料研究的新领域。这类包含有大量表面或界面原子的新型功能材料具有许多独特的物理、化学和力学性能, 被誉为 21 世纪最有前途的材料。

近年来, 各国科学家将选择生成超薄层外延技术与精细束加工技术紧密结合起来, 研制量子线与量子点及其光电器件, 预期能发现一些新的物理现象和得到更好的器件性能。在器件长度小于

电子平均自由程的所谓介观系统中, 电子输运不再遵循通常的欧姆定律, 电子运动完全由它的波动性质决定. 人们发现了电子输运的 Aharonov-Bohm 振荡, 电子波的相干振荡以及量子点的库仑阻塞现象等. 以上这些新材料, 新物理现象的发现产生了新的器件设计思想, 促进了新一代半导体器件的发展.

从以上讨论可以看到半导体物理的发展和半导体微电子技术、光电子技术的发展是密切相关, 互相促进的. 本书主要结合近几十年来半导体物理的发展, 介绍有关的和必须的基础知识. 本书作者和朱邦芬研究员曾经写过一本《半导体超晶格物理》, 在那本书里, 已经介绍了许多半导体超晶格物理的概念和最新发展, 本书不再重复. 半导体表面和界面是半导体物理的一个重要方面, 由于《北京大学物理学丛书》中另有许振嘉教授的一本关于这方面内容的专著《近代半导体材料的表面科学基础》, 本书也不再涉及. 本书的名字原来打算叫“半导体物理学”, 但是这样的书名在国际和国内都太多了. 考虑到本书的内容是作者在近年内做研究工作时感到所需要的一些基础知识, 目前科学技术发展是如此迅速, 知识不断地更新, 本书的内容只能满足当前一段时间内的需要. 因此本书的名字改叫“Current Semiconductor Physics”, 表示有当代的、流行的意思, 而不是永久的, 中文名字则叫“现代半导体物理”.

参 考 文 献

- [1] Kittel C, Mitchell A H. *Phys. Rev.*, 1954, **96**: 1488
- [2] Luttinger J M, Kohn W. *Phys. Rev.*, 1954, **97**: 869
- [3] Phillips J C, Kleinman L. *Phys. Rev.*, 1959, **116**: 287
- [4] Esaki L, Zhu R. *IBM Res. Dev.*, 1970, **14**: 61
- [5] Cho A Y. *Appl. Phys. Lett.*, 1971, **19**: 467
- [6] Dingle R, Stormer H L, Gossard A C, et al. *Appl. Phys. Lett.*, 1978, **33**: 665
- [7] von Klitzing K, Dorda G, Pepper M. *Phys. Rev. Lett.*, 1980, **45**: 494
- [8] Tsui D C, Stormer H L, Gossard A C. *Phys. Rev. Lett.*, 1982, **48**: 1562
- [9] Miller D A b, Chemla D S, Damen T C, et al. *Appl. Phys. Lett.*, 1984, **45**: 13
- [10] Canham L T. *Appl. Phys. Lett.*, 1990, **57**: 1046

第一章 半导体能带结构

1.1 键轨道理论和紧束缚方法

Ⅳ族半导体,如锗、硅等是共价晶体,Ⅲ-V族半导体基本上是共价晶体,但因为组成晶体的两类原子荷电性不同,因而具有部分离子性.而Ⅱ-VI族半导体则离子性更强一些.共价键与金属键、离子键不同,金属键的电子比较均匀地分布在固体中,离子键的电子则集中在离子上,而共价键的电子则比较集中地分布在两个原子之间的连线上.根据这一概念,提出了键轨道理论,它比较适合描写具有共价键的半导体的能带结构.

先考虑一个一维链固体,链上每个原子有 1 个 s 电子态和 1 个 p 电子态,它们的波函数空间分布如图 1.1(a)所示.在固体中,两相邻原子之间的距离较近,因此相邻原子的电子态之间会有相互作用.根据键轨道理论,将 s 态和 p 态的轨道波函数重新线性组合成

$$\begin{cases} \phi_1 = \frac{1}{\sqrt{2}}(|s\rangle + |p\rangle), \\ \phi_2 = \frac{1}{\sqrt{2}}(|s\rangle - |p\rangle). \end{cases} \quad (1.1.1)$$

它们的波函数分布如图 1.1(b)所示. ϕ_1 主要集中在原子的右边, ϕ_2 主要集中在原子的左边.也就是 ϕ_1, ϕ_2 都有了方向性,各自指向相邻的原子,我们称 ϕ_1, ϕ_2 态为杂化键.

键轨道理论认为,当原子形成半导体晶体时,只有相邻原子两相对杂化键之间的相互作用为最大,由此构成了共价晶体.如果用 ϕ_{n1}, ϕ_{n2} 分别代表第 n 个原子的两个杂化键,则有

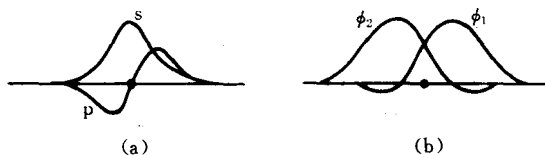


图 1.1 (a)原子波函数；(b)杂化键波函数

$$\begin{cases} \langle \phi_{n1} | H | \phi_{n1} \rangle = \langle \phi_{n2} | H | \phi_{n2} \rangle = E_0, \\ \langle \phi_{n1} | H | \phi_{n+1,2} \rangle = -V_2, \\ \langle \phi_{n1} | H | \phi_{n2} \rangle = V_1, \end{cases} \quad (1.1.2)$$

其中 E_0 是键轨道自身的能量, V_2 是相对杂化键之间的相互作用能, V_1 是同一原子两杂化键之间的相互作用能, $|V_1| < |V_2|$. 键轨道之间的其他相互作用 (例如次近邻键) 很小, 我们暂时先忽略. 由紧束缚理论, 固体的波函数为各原子键轨道的线性组合,

$$\begin{cases} \psi_1 = \frac{1}{\sqrt{N}} \sum_n \phi_{n1} e^{ikR_n}, \\ \psi_2 = \frac{1}{\sqrt{N}} \sum_n \phi_{n2} e^{ikR_n}, \end{cases} \quad (1.1.3)$$

其中 k 是波矢, $-\frac{\pi}{a} < k < \frac{\pi}{a}$. $\left[-\frac{\pi}{a}, \frac{\pi}{a}\right]$ 是一维晶体的布里渊区, a 是晶格常数, 也就是相邻原子间的距离. 将波函数(1.1.3)式代入薛定谔方程, 就得到久期方程

$$\begin{vmatrix} E_0 - E & V_1 - V_2 e^{ika} \\ V_1 - V_2 e^{-ika} & E_0 - E \end{vmatrix} = 0. \quad (1.1.4)$$

解这方程就能得到本征能量

$$E = E_0 \pm \sqrt{V_1^2 + V_2^2 - 2V_1V_2 \cos ka}. \quad (1.1.5)$$

由(1.1.5)式得到的能带如图 1.2(a)所示, 在布里渊区内有上下两个能带, 分别对应于导带和价带, 中间是能隙. 能带的形成过程如图 1.2(b)所示. 当原子分离很远时, 每个杂化键的能量是 E_0 . 当原子逐渐靠近时, 相对杂化键的相互作用 V_2 形成了上、下两个带. 再由相互作用 V_1 使每个带有能量色散, 有一定宽度 $2V_1$.

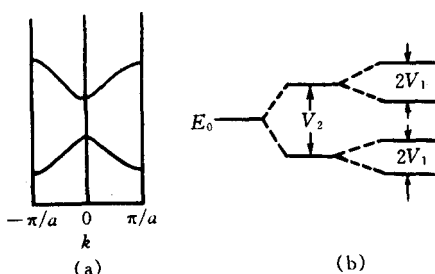


图 1.2 (a) 一维晶体的能带；(b) 能带的形成过程

可以证明,在 $k=0$ 导带底和价带顶的波函数分别为

$$\begin{cases} \Psi_c = \frac{1}{\sqrt{2}}(\psi_1 - \psi_2), \\ \Psi_v = \frac{1}{\sqrt{2}}(\psi_1 + \psi_2). \end{cases} \quad (1.1.6)$$

所以通常称价带的波函数由成键态组成,两个键轨道布洛赫函数(1.1.3)相加;而导带波函数由反键态组成,两个键轨道布洛赫函数相减。

以上是键轨道理论的一些基本概念,将它们推广到实际半导体晶体(三维晶体)是直接的.大多数半导体是金刚石或闪锌矿结构,每个原子与周围最近邻四个原子组成正四面体.为一般起见,考虑闪锌矿结构,也就是由 A, B 两种原子组成的晶体.每个原子的价电子有 1 个 s 态和 3 个 p 态.由 A 原子的轨道波函数组成它的杂化波函数

$$\begin{cases} \phi_1 = \frac{1}{2}(|s\rangle + |p_x\rangle + |p_y\rangle + |p_z\rangle), \\ \phi_2 = \frac{1}{2}(|s\rangle - |p_x\rangle - |p_y\rangle + |p_z\rangle), \\ \phi_3 = \frac{1}{2}(|s\rangle - |p_x\rangle + |p_y\rangle - |p_z\rangle), \\ \phi_4 = \frac{1}{2}(|s\rangle + |p_x\rangle - |p_y\rangle - |p_z\rangle). \end{cases} \quad (1.1.7)$$

它们也有方向性，分别指向 $[111]$ 、 $[\bar{1}\bar{1}\bar{1}]$ 、 $[\bar{1}1\bar{1}]$ 和 $[1\bar{1}\bar{1}]$ 各相邻 B 原子的方向。B 原子的杂化波函数则分别指向 $[111]$ 、 $[11\bar{1}]$ 、 $[1\bar{1}1]$ 和 $[\bar{1}11]$ 各相邻 A 原子的方向，因此它们的 p 态函数的符号与 (1.1.7) 式中的相反。

为了计算实际半导体的能带，首先必须确定各杂化键（包括同一原子和相邻原子，以至次近邻原子杂化键）间的相互作用矩阵元。从单纯计算能带的角度，用杂化轨道 (1.1.7) 式或者原子轨道 $|s\rangle, |p_x\rangle, |p_y\rangle, |p_z\rangle$ 作为基函数是完全等价的。它们之间由一个幺正变换可以互相变换。但是在某些情况下，如计算半导体原子团簇或半导体中深杂质能级所用的集团模型等，用杂化轨道，或杂化轨道组成的键轨道（成键轨道或反键轨道）作为基函数是比较方便的。因为首先它们具有方向性，直观，容易根据空间对称性将它们组成不可约表示的基函数。其次在团簇或集团边界的悬键容易进行饱和和处理。

杂化轨道或原子轨道间的相互作用矩阵元一般不是由第一原理计算得到的，而是通过与其他方法（如赝势方法）计算所得到较精确的能带相比较而确定的，因此是经验的。即使是同一种半导体，不同作者得到的相互作用参数可以是完全不同的。对此 Harrison 有这样的观点^[1]：“这些参数没有一个最好的选择，因为必须在简单性（通用性）和得到能带的精确性之间作出交易。如果对每一种材料的每一种性质采用不同的矩阵元值，就能把这种性质的精确值调整得很好。但是我们将采用几乎是相反的极端，只引进四个参数，所有系统的 s 和 p 态之间的相互作用矩阵元都可以计算。”Harrison 的做法适用于研究各种材料性质之间的定性比较，它们随着原子量变化的化学趋势等。

下面简单介绍一下 Harrison 的方法。他首先引入四个量纲一的原子轨道间的相互作用矩阵元，见表 1.1.

表 1.1 确定原子轨道矩阵元的量纲一的系数

系数	$\eta_{ss\sigma}$	$\eta_{sp\sigma}$	$\eta_{pp\sigma}$	$\eta_{pp\pi}$
数值	-1.40	1.84	3.24	-0.81

相邻原子间的原子轨道矩阵元为

$$V_{ll'm} = \eta_{ll'm} \frac{\hbar^2}{m_0 d^2}, \quad (1.1.8)$$

其中 $ll'm$ 标著量子态, m_0 是自由电子质量, d 是原子间距离. 关于同一原子上原子轨道能量列于表 1.2.

表 1.2 原子轨道能量 (单位: eV)

2s, 2p	Li	Be	B	C	N	O	F
$-\epsilon_s$	5.48	8.17	12.54	17.52	23.04	29.14	35.80
$-\epsilon_p$		4.14	6.64	8.97	11.47	14.13	16.99
3s, 3p	Na	Mg	Al	Si	P	S	Cl
$-\epsilon_s$	5.13	6.86	10.11	13.55	17.10	20.80	24.63
$-\epsilon_p$		2.99	4.86	6.52	8.33	10.27	12.31
4s, 4p	K	Zn	Ga	Ge	As	Se	Br
$-\epsilon_s$	4.19	8.40	11.37	14.38	17.33	20.32	23.35
$-\epsilon_p$		3.38	4.90	6.36	7.91	9.53	11.20
5s, 5p	Ag	Cd	In	Sn	Sb	Te	I
$-\epsilon_s$	6.41	7.70	10.12	12.50	14.80	17.11	19.42
$-\epsilon_p$	2.05	3.38	4.69	5.94	7.24	8.59	9.97

利用表 1.1 和表 1.2 可以得到任意一个半导体材料的原子轨道相互作用矩阵元. Slater 和 Koster^[2]在二心近似的基础上, 得出不同原子的轨道矩阵元可以表示为

$$\begin{cases} H_{ss'} = V_{ss\sigma}, \\ H_{sx'} = lV_{sp\sigma}, \\ H_{xx'} = l^2V_{pp\sigma} + (1 - l^2)V_{pp\pi}, \\ H_{xy'} = lmV_{pp\sigma} - lmV_{pp\pi}. \end{cases} \quad (1.1.9)$$

其中 $V_{ss\sigma}, V_{sp\sigma}, \dots$ 就是(1.1.8)式中定义的矩阵元, $V_{ss\sigma}$ 代表一个原子的 s 态和近邻原子的 s' 态轨道波函数之间的矩阵元, 余此类推. l, m, n 是两原子连线的方向余弦.

由于杂化轨道(1.1.7)式是原子轨道的线性组合, 已知了原子轨道的矩阵元, 杂化轨道的矩阵元可直接求得. 由于杂化轨道的方向性, 可以得出杂化键矩阵元之间的一系列关系. 将相邻原子 A 和 B 的杂化键分别表示为 1, 2, 3, 4 和 5, 6, 7, 8, 如图 1.3 所示. 由空间的对称性可以它们之间的矩阵元有下列的关系:

$$\begin{cases} V_2 = \langle 1 | H | 5 \rangle, \\ V_3 = \langle 1 | H | 6 \rangle = \langle 1 | H | 7 \rangle = \langle 1 | H | 8 \rangle, \\ V_4 = \langle 2 | H | 5 \rangle = \langle 3 | H | 5 \rangle = \langle 4 | H | 5 \rangle, \\ V_5 = \langle 2 | H | 6 \rangle = \langle 3 | H | 7 \rangle = \langle 4 | H | 8 \rangle, \\ V_6 = \langle 2 | H | 7 \rangle = \langle 2 | H | 8 \rangle = \langle 3 | H | 6 \rangle = \dots \end{cases} \quad (1.1.10)$$

因此只有五个独立参量. 其中以两个相对杂化键之间的矩阵元 V_2 为最大. 利用以上各式可求得

$$V_2 = \frac{1}{4} [V_{ss\sigma} - 2\sqrt{3}V_{sp\sigma} - 3V_{pp\sigma}] = -4.37 \frac{\hbar^2}{m_0 d^2}. \quad (1.1.11)$$

其他的矩阵元可类似地求得.

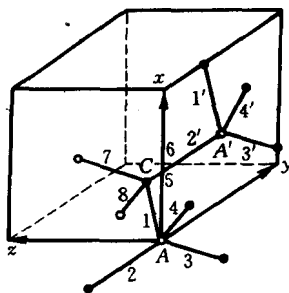


图 1.3 相邻原子 A 和 B 的杂化键

已知了原子轨道间的矩阵元或杂化键间的矩阵元就可以计算能带, 按照布洛赫原理, 组成晶体的紧束缚波函数

$$\phi_{ij} = \frac{1}{\sqrt{N}} \sum_n e^{ik \cdot R_{nj}} \phi_{ij}(R_{nj}), \quad (1.1.12)$$

其中

$$R_{nj} = R_n + \tau_j, \quad (1.1.13)$$

R_n 是第 n 个元胞的坐标, τ_j 是第 j 类原子在元胞中的位置坐标. $\phi_{ij}(R_{nj})$ 是位于 R_{nj} 上的第 j 类原子的第 i 类原子轨道或杂化轨道. 将波函数 (1.1.12) 式代入薛定谔方程, 得到久期方程

$$|H_{ij,i'j'}(k) - E\delta_{ii'}\delta_{jj'}| = 0, \quad (1.1.14)$$

其中 $H_{ij,i'j'}(k)$ 是哈密顿量矩阵元. 由波函数 (1.1.12) 式可求得

$$H_{ij,i'j'}(k) = \sum_n e^{ik \cdot (R_{n'j'} - R_{nj})} \langle \phi_{ij}(R_{nj}) | H | \phi_{i'j'}(R_{n'j'}) \rangle, \quad (1.1.15)$$

注意其中的 R_{nj} 是任意的, 可以取为坐标原点. 在求和号中求相邻原子的轨道对矩阵元的贡献. 由久期方程 (1.1.14) 可见, 它应该是 8×8 维的. 因为每个元胞有两个原子, 每个原子有四个轨道, 一共有 8 个基函数. 在一般情况下是不能得到解析结果的, 只能数值求解. 但是在布里渊区的某些特殊点, 例如 Γ, X, L 点, 利用群论方法可以将久期方程简化为低维的, 得到解析结果. 利用这方法可以由其他方法求得的在布里渊区特殊点上的能带精确值, 反推定出轨道矩阵元.

在实际应用中, 很少用紧束缚方法去计算半导体能带. 现在用紧束缚方法或键轨道方法主要计算由于平移对称性破坏而形成的局域电子态, 如表面电子态、深杂质、缺陷电子态以及半导体量子阱、量子线、量子点的电子态等. 方法是大元胞法或集团法, 将在下面的有关章节中介绍. 用紧束缚方法计算的局域电子态波函数是各原子轨道的线性组合, 因此直接反映了这些态的电子空间分布情况, 在物理上很直观.

紧束缚方法的一个主要的缺点是，它虽然能求得很好的价带结构，但是它总不能求得与实验符合很好的导带结构。其根源是“紧束缚”近似，而导带态更接近于自由电子近似，用紧束缚基函数很难得到正确的导带态。为此人们想出了许多补救办法。如：除了 s, p_x, p_y, p_z 轨道外，还人为地增加一个激发 s^* 轨道^[3]，或者增加几个 d 轨道。除了考虑最近邻原子的相互作用外，还考虑次近邻，以至第三近邻原子间的相互作用。除了二心近似的相互作用矩阵元外，还引入一些非二心近似的矩阵元。

紧束缚矩阵元参数通常是由拟合能带经验确定的，即使对同一种材料，不同作者采用不同近似拟合出来的参数可以相差很大。因此就产生了这样的问题，紧束缚矩阵元参数有没有物理意义？现在又回到 Harrison 的理论^[1]。Harrison 给出的最近邻原子间轨道矩阵元的普适关系(1.1.8)式和表 1.1 是由近自由电子能带推算出来，并根据拟合 Si, Ge 能带修正后得到的。自身原子轨道的能量，表 1.2 是由 Herman, Skillman 自治原子计算得到的原子能项值。由表 1.1, 1.2 和 (1.1.8) 式得到的几种半导体的紧束缚矩阵元参数列于表 1.3，为了比较还列出了拟合能带得到的参数值。

表 1.3 由 Harrison 理论得到的紧束缚参量 (单位: eV)
(第二行的值是拟合能带得到的，一格内两个数分别对应于非金属和金属性原子)

矩阵元	C	Si	Ge	GaAs	ZnSe
$-V_{ss\sigma}$	4.50	1.93	1.79	1.79	1.79
	5.55	2.03	1.70	1.70	1.54
V_{spo}	5.91	2.54	2.36	2.36	2.36
	5.91	2.55	2.30	2.4, 1.9	2.6, 1.4
V_{ppo}	10.41	4.47	4.15	4.15	4.15
	7.78	4.55	4.07	3.44	3.20
$-V_{pp\pi}$	2.60	1.12	1.04	1.04	1.04
	2.50	1.09	1.05	0.89	0.92
$(\epsilon_p - \epsilon_s)/4$	2.14	1.76	2.01	1.62, 2.36	1.26, 2.70
	1.70	1.80	2.10	1.61, 2.41	1.47, 3.10
$(\epsilon_p^1 - \epsilon_s^2)/2$	0	0	0	1.50	3.08
	0	0	0	0.96	3.29

由表可见, 两者大致是相近的. *Harrison* 理论反映了半导体能带结构的一些共性, 但它不能很好描述每一种半导体的个性. 特别是对三种晶格常数相等的半导体 $Ge, GaAs, ZnSe$, 它给出了相同的 $V_{II'm}$, 这当然不符合实际.

1.2 赝势方法

1.2.1 赝势概念的提出

在赝势概念提出以前, 人们都用正交平面波法计算晶体的能带. 由于晶体具有平移对称性, 因此可以将晶体波函数用平面波展开,

$$\psi_k(\mathbf{r}) = \frac{1}{\sqrt{V}} \sum_i C_i e^{i(\mathbf{k} + \mathbf{G}_i) \cdot \mathbf{r}}, \quad (1.2.1)$$

其中 V 是晶体体积, 并且倒格矢量

$$\mathbf{G}_i = 2\pi(m_1\mathbf{b}_1 + m_2\mathbf{b}_2 + m_3\mathbf{b}_3), \quad (1.2.2)$$

$\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3$ 是倒格矢基矢. 因此波函数(1.2.1)具有空间对称性, 也就是沿元胞三基矢方向平移是不变的. 将波函数(1.2.1)代入薛定谔方程, 就得到确定能量本征值和本征波函数的久期方程

$$\begin{cases} |H_{ij} - E\delta_{ij}| = 0, \\ H_{ij} = \langle W_i | H | W_j \rangle, \end{cases} \quad (1.2.3)$$

其中 W_i 是波函数(1.2.1)中某一个平面波分量. 实际上直接用平面波展开方法计算晶体能带是不可能的, 因为组成晶体的原子除了外层价电子外, 还有许多内壳层电子态. 这些态称为核心态(core states), 具有空间很局域的电子轨道, 相邻原子的核心态重叠很小. 由于核心态的波函数在空间是很局域的, 需要非常多的平面波才能正确地给出核心态的波函数.

由于晶体中核心态之间的相互作用可以忽略, 因此认为在晶体中的核心态与孤立原子中的核心态是一样的. *Herring*^[4]由此提出了正交平面波法(OPW). 由于晶体中较高能量的电子态(包括

导带和价带) 必须与低能量的核心态正交, 因此用一个与核心态正交的平面波来代替原来的波函数(1. 2. 1),

$$\chi_i(\mathbf{r}) = W_i(\mathbf{r}) - \sum_c \langle \phi_c | W_i \rangle \phi_c. \quad (1. 2. 4)$$

其中 ϕ_c 是核心态. 将(1. 2. 4)式代入薛定谔方程, 得到类似于(1. 2. 3)式的久期方程, 只是其中的哈密顿矩阵元的基函数由 W_i, W_j 换成了 χ_i, χ_j . 解这个久期方程, 得到的最低能量的态是价带的态, 因为波函数(1. 2. 4)已经自动地与核心态正交了. 一般用几百个正交平面波就能得到收敛的结果. 用这种方法已经计算了许多晶体的能带.

在这基础上, Phillips 等人^[5]提出了赝势的概念. 假定价电子波函数为

$$\psi(\mathbf{r}) = \varphi(\mathbf{r}) - \sum_c \langle \phi_c | \varphi \rangle \phi_c(\mathbf{r}). \quad (1. 2. 5)$$

将(1. 2. 5)式代入薛定谔方程, 可以得到一个 $\varphi(\mathbf{r})$ 满足的方程

$$\left[\frac{\mathbf{p}^2}{2m_0} + V(\mathbf{r}) + \sum_c (E - E_c) |\phi_c\rangle \langle \phi_c| \right] \varphi(\mathbf{r}) = E \varphi(\mathbf{r}), \quad (1. 2. 6)$$

其中 $|\phi_c\rangle \langle \phi_c|$ 是狄拉克的并矢符号, 代表一个算符. 因此 $\varphi(\mathbf{r})$ 具有与真实波函数 $\psi(\mathbf{r})$ 相同的本征能量, 而晶体势被下列势所代替,

$$V_p(\mathbf{r}) = V(\mathbf{r}) + \sum_c (E - E_c) |\phi_c\rangle \langle \phi_c|. \quad (1. 2. 7)$$

V_p 称为晶体赝势(pseudopotential). 这样做表面上看好像仅仅是一个数学变换, 但实质上将使能带计算大为简化, 甚至比正交平面波法还简单得多. 首先考虑赝波函数 $\varphi(\mathbf{r})$, 如果真实波函数用正交平面波展开,

$$\psi(\mathbf{r}) = \sum_i a_i \chi_i = \sum_i a_i \left[W_i - \sum_c \langle \phi_c | W_i \rangle \phi_c \right], \quad (1. 2. 8)$$

则可以证明, $\varphi(\mathbf{r})$ 是系数 a_i 相同的平面波展开,

$$\varphi(\mathbf{r}) = \sum_i a_i W_i. \quad (1. 2. 9)$$

由于平面波的展开项数不多, 因此 $\varphi(r)$ 在空间是一个变化平缓的函数. 它在核心区以外与真实波函数相同, 在核心区以内, 它将真实波函数的强烈振荡平滑掉了.

在赝势理论发展的初期, 还有人从 (1.2.7) 式出发研究赝势的具体形式. 但后来发现, 赝势不是惟一的, 而且它的细节对能带结构的影响不大, 因此后来发展了经验赝势、模型离子赝势等方法.

1.2.2 经验赝势

Phillips 等^[5]已经证明, 赝势在核心区域对价电子有排斥作用, 将价电子排斥在核心区域之外. 这个排斥势几乎抵消了在核心区内很强的离子吸引势, 最后形成一个弱的吸引赝势, 这称为 Phillips 抵消原理. 图 1.4 是 Mo 离子的真实势(虚线)和赝势(实线)^[6], 由图可见, 在核心区以内赝势是弱的吸引势, 并且对不同角动量的波函数是不同的 (非局域的). 在核心区域以外, 赝势与真实势相同.

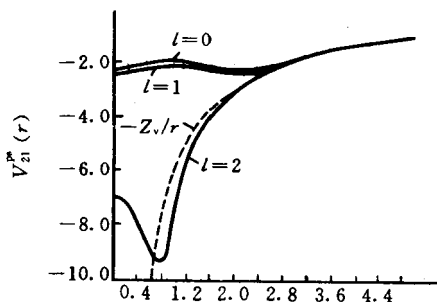


图 1.4 Mo 离子的真实势(虚线)和赝势(实线)

设晶体赝势是由组成晶体的所有原子的赝势叠加而成. 用 $v_i(r)$ 代表第 i 种原子的赝势, 则晶体赝势

$$V_p(r) = \sum_{n,i} v_i(r - R_{ni}), \quad (1.2.10)$$

其中

$$\mathbf{R}_{ni} = \mathbf{R}_n + \boldsymbol{\tau}_i, \quad (1.2.11)$$

$\mathbf{R}_n = n_1 \mathbf{a}_1 + n_2 \mathbf{a}_2 + n_3 \mathbf{a}_3$ 是元胞基矢的平移矢量, $\boldsymbol{\tau}_i$ 是第 i 类原子在元胞中的位置矢量. 由于 $V_p(\mathbf{r})$ 是沿元胞基矢三个方向的周期函数, 它可以作傅里叶展开

$$V_p(\mathbf{r}) = \sum_{\mathbf{G}} V_p(\mathbf{G}) e^{i\mathbf{G} \cdot \mathbf{r}}, \quad (1.2.12)$$

$\mathbf{G}$ 是倒格矢量 (1.2.2) 式. 可以证明, 展开系数

$$V_p(\mathbf{G}) = \frac{1}{V} \int V_p(\mathbf{r}) e^{-i\mathbf{G} \cdot \mathbf{r}} d\mathbf{r} = \frac{1}{N} \sum_i e^{-i\mathbf{G} \cdot \boldsymbol{\tau}_i} v_i(\mathbf{G}), \quad (1.2.13)$$

其中 V 是晶体体积, N 是一个元胞中原子的个数,

$$v_i(\mathbf{G}) = \frac{1}{\Omega_a} \int v_i(\mathbf{r}) e^{-i\mathbf{G} \cdot \mathbf{r}} d\mathbf{r}, \quad (1.2.14)$$

Ω_a 是一个原子占有的体积. $v_i(\mathbf{G})$ 称为原子赝势的形状因子, (1.2.13) 式中的 $e^{-i\mathbf{G} \cdot \boldsymbol{\tau}_i}$ 称为原子赝势结构因子, 因为它与原子在晶体中的位置有关. 因此晶体赝势的傅氏展开是由原子赝势的形状因子和结构因子两部分组成.

由 (1.2.14) 式可见, 如果原子赝势 $v_i(\mathbf{r})$ 是球对称的, 只是坐标绝对值 r 的函数, 则它的形状因子也是球对称的,

$$v_i(\mathbf{G}) = \frac{4\pi}{\Omega_a G} \int r v_i(r) \sin Gr \, dr. \quad (1.2.15)$$

实际上大部分赝势计算都是假定这种情形.

对闪锌矿结构的半导体, 取坐标原点在元胞中两个原子连线的中点, 则

$$\boldsymbol{\tau}_1 = -\frac{a}{8}(1, 1, 1) = -\boldsymbol{\tau}, \quad \boldsymbol{\tau}_2 = \frac{a}{8}(1, 1, 1) = \boldsymbol{\tau}, \quad (1.2.16)$$

$$\begin{aligned} V_p(\mathbf{G}) &= \frac{1}{2} [e^{i\mathbf{G} \cdot \boldsymbol{\tau}_1} v_1(\mathbf{G}) + e^{-i\mathbf{G} \cdot \boldsymbol{\tau}_2} v_2(\mathbf{G})] \\ &= v_s(\mathbf{G}) \cos(\mathbf{G} \cdot \boldsymbol{\tau}) + i v_a(\mathbf{G}) \sin(\mathbf{G} \cdot \boldsymbol{\tau}), \end{aligned} \quad (1.2.17)$$

其中

$$v_s(G) = \frac{1}{2}[v_1(G) + v_2(G)], \quad v_a(G) = \frac{1}{2}[v_1(G) - v_2(G)]. \quad (1.2.18)$$

v_s 和 v_a 分别称为对称和反对称赝势形状因子. 对金刚石结构的半导体, $v_1(G) = v_2(G)$, 对称赝势形状因子就是原子赝势形状因子, 而反对称赝势形状因子为零.

为了确定原子赝势的形状因子, 假定已知了裸离子 (价电子剥离后的原子, 如 Si^{4+} , Ga^{3+} , As^{5+} 等) 赝势的形状因子 $v_{\text{ion}}(G)$, 并且当这些原子组成晶体时, 价电子是均匀分布的, 则晶体赝势

$$V_p(r) = \sum_G \frac{v_{\text{ion}}(G)S(G)}{\epsilon(G)} e^{iG \cdot r}, \quad (1.2.19)$$

其中对 G 的求和不包括对 $G=0$ 项的求和, $S(G)$ 是晶体结构因子, $\epsilon(G)$ 是介电屏蔽函数. 因此在均匀价电子分布的近似下, 原子赝势

$$v(q) = \frac{v_{\text{ion}}(q)}{\epsilon(q)}. \quad (1.2.20)$$

这里用 q 表示一个连续变化的变量, 而在 (1.2.19) 式中, G 只取与倒格矢 (1.2.2) 对应的离散值. Sham^[7] 利用电子气理论, 考虑了相关和交换效应, 给出了介电屏蔽函数. 这是一个复杂的表示式, 这里不再列出. 此外还可以利用其他近似, 例如 Thomas-Fermi 近似给出的介电屏蔽函数.

由 Phillips 的抵消原理, 赝势在核心区域内是弱的吸引势, 在核心区以外与真实的离子势相同, Abarenkov 等^[8] 提出了一个简单的离子赝势模型:

$$v_{\text{ion}}(r) = \begin{cases} -A, & r < R_M, \\ -Z/r, & r > R_M, \end{cases} \quad (1.2.21)$$

其中 R_M 相当于核心区的半径, A 是一个小的常数. 由 (1.2.21) 式可算得它的傅氏变换