



固体理论进展 — 半世纪回顾

余志平^{①*}, 王燕^①, 刘晓彦^②

① 清华大学微电子学研究所, 北京 100084

② 北京大学微电子学研究所, 北京 100871

* 通信作者. E-mail: yuzhip@tsinghua.edu.cn

收稿日期: 2012-08-11; 接受日期: 2012-08-28

摘要 本文回顾了自 1964 年密度泛函理论 (DFT) 提出而引入了被称为第一原理计算电子学近半个世纪以来, 固体物理在理论、材料与实验上的进展. 所选的关注点是那些和微纳电子学与集成电路技术有关的重大进展. 内容分为固体理论与计算电子学、低维与超导材料两个部分. 硅基 CMOS 集成电路在经历了本文涵盖的同一历史时期中令人瞩目的高速发展后, 无论在研究与产业上都处于一个新发展阶段的十字路口. 通过回顾作为现代信息工程基石的固体器件与理论的进展, 可以对今后集成电路的发展趋向提供一个分析参考.

关键词 固体物理 集成电路 计算电子学 低维材料与物理 石墨烯/拓扑绝缘体/铁基超导体

1 引言

固态物理 (在广义上又可称为凝聚态物理) 是物理领域里最为活跃的一个学科, 其原因与集成电路的广泛应用不无关系. 用固体 (主要是半导体) 实现的电路不仅规模大, 性能好, 同时重复性好, 易于低成本地大批量生产. 这在很大程度上是因为固态集成电路的加工是基于类似印刷技术的平面工艺, 即薄层淀积/生长与光刻.

半导体器件与集成电路的发展与固态物理的进展相辅相成, 近半个世纪来, 两者都有了量子跃进 (quantum leap) 式的进步. 尤其在集成电路工艺加工的精度进入了纳米尺度 (可以以 Intel 在 2003 年 90 nm, 即亚 100 nm, CMOS 工艺节点^[1]为起始点) 以来, 固体物理研究的显著特点是低维 (即二维或一维, 相对于三维而言) 材料与器件的加工精度日益趋近于原子的尺度, 量子力学原理在固体器件的工作上起关键性的作用.

本文从集成电路发展的角度来回顾、总结固态物理在理论与实践方面的进展. 内容分两个方面: 能带结构与载流子量子输运计算以及低维材料物理 (包括铁基超导).

因为篇幅的限制与尽可能反映近期的有现实影响的事件与进展, 我们忽略了这半个世纪来在固体理论领域中其他一些重大的进展, 这里包括:

1. 量子霍尔效应 (1980, K. von Klitzing (U Würzburg, 德国), 1985 年获诺贝尔物理学奖);
2. 铜氧化物高温超导体 (高临界温度陶瓷材料)(1986, J. G. Bednorz, K. A. Müller (IBM Zurich)^[2], 1987 年获诺贝尔物理学奖);
3. 超晶格 (1970, L. Esaki, R. Tsu (IBM));

4. 光子晶体 (1987, E. Yablonovitch (UCLA), S. John (Princeton));
5. 量子计算与固态实现 (1998, B. Kane, U New South Wales, 澳大利亚^[3]).

2 能带结构与载流子量子输运

固体物理与原子物理的基本不同之处在于固体中的电子波函数不再局限于单个原子核附近, 而是在整个固体内都存在. 因而固体物理的一个基本任务是找到固体作为一个量子系统的电子本征态 (即波函数) 与本征能量. 固体的其他物理 (包括光学) 性质都可由有关这两个量的知识出发得到. 因为固体物理的主要研究对象是晶体, 而晶体具有晶格的周期性 (或称平移对称性), 根据布洛赫定理 (Bloch theorem)¹⁾, 电子的波函数增加了一个量子数, 即为波函数中平面波部分的波矢, 一般情况下为一个三维矢量, 记为 \mathbf{k} . 这样波函数的本征能量不仅与能带指数 (记为 n , 由原子中电子轨道能级派生而成) 有关, 而且为 \mathbf{k} (或简称为 k , 可近似视为一个连续量) 的函数. 我们一般称固体中能带 (亦称为电子结构) 计算为 $E - k$ 关系的计算, 即由此而来.

经典的能带计算方法在所有的固体物理教材中都有描述, 包括紧束缚 (TB)、平面波展开 (增强型 -APW, 或正交型 -OPW)、赝势法、格林函数法 (以三位发明人命名: KKR) 及 $k \cdot p$ (读为 k 点 p) 法. 本报告中仅以应用在纳电子材料与器件能带计算中十分广泛的紧束缚方法的进展进行讨论.

在经典的能带计算方法之后的最重要的理论进展则是考虑了多体问题 (即具有相互作用的多电子体系) 的 HK 方法. 这是由 P. Hohenberg (法国巴黎高等师范学校, ENS: Ecole Normale Supérieure) 与 W. Kohn (美国加州大学圣地亚哥分校, UCSD) (合称为 HK) 在 1964 年提出的密度泛函理论 (density functional theory, DFT)^[4]. 他们在 *Physical Review* 发表的用密度泛函方法来求得在晶格势的作用下一个有相互作用的非均匀电子气系统的基态文章奠定了被称为 “第一原理计算” (或称 *ab initio*, 即从头计算) 的计算固态物理的基础. 次年 (1965), Kohn 与 Sham (UCSD) (KS) 又在 *Physical Review* 发表文章提出了用局域密度近似 (local density approximation, LDA) 来得到 DFT 方法中最为关键 (也是最难得到的) 交换关联 (exchange and correlation, xc) 能的泛函形式^[5]. 至此, DFT-LDA 成为了求固体能带结构的一个广为流行的重要手段.²⁾

在纳米尺度的器件中, 除了能带结构与体材料会有所不同外, 考虑到其中的载流子 (电子与空穴) 的输运, 经典与半经典的输运模型, 如扩散漂移 (DD) 及水动力学 (HD) 模型, 都不再适用. 这是因为除了依赖于扩散与漂移的机理外, 量子隧穿成为载流子输运的常态. 目前广泛采用的非平衡格林函数法 (non-equilibrium Green's function, NEGF) 是用来计算开放量子系统 (半导体器件在纳米尺度下就是这样的系统) 载流子输运的理论方法, 其综合了结果 (电流电压特性 $I-V$, 电容电压特性 $C-V$) 的准确性与计算量的可实现性, 可视为近 20 年来在量子输运计算方面的最大进展事件. NEGF 方法的普及可追溯到美国普渡大学 (Purdue) 的 Datta 教授在 1995 年由剑桥大学出版社出版的《介观系统的电子输运》一书^[6].

下面, 我们分别讨论经典能带计算方法的进展、第一原理能带计算方法与量子输运模型.

1) 1928 年, Felix Bloch (布洛赫), 斯坦福大学 Max Stein 讲座教授, 1952 年因为发展了对核磁进行精确测量的方法与美国物理学家 Edward Mills Purcell 一起获得诺贝尔物理学奖

2) 这 2 篇 *Physical Review* 文章, 在 2005 年的 *Physics Today* (今日物理) 上被指出为在物理评论系列期刊 (包括 *Phys. Rev. B*, *Phys. Rev. Lett.*, *Rev. Mod. Physics*) 中被引用次数最多的 2 篇文章 (依当时统计分别是 2460 与 3227 次)

2.1 能带计算方法的进展

如前所述, 晶体的能带结构是基于晶体的周期性与布洛赫波函数理论而形成的. 晶体能带结构计算的基本思路是将要求的晶体波函数 (以波矢 \mathbf{k} 与能带指数 n 为参量, 统称为晶体量子数) 在以实空间周期矢量 \mathbf{R} 或倒格子空间周期矢量 \mathbf{K} 为指数的基函数级数上展开. 不同的计算方法以依赖的指数不同分为两大类, 在两大类中又以基函数是如何构造来进行区分.

近年来, 采用紧束缚近似计算能带结构有很大的进展. 采用紧束缚近似的优点是能够处理原胞中包含几百个原子的大系统. 这个方法具有直观性, 采用化学键来真实地描述结构与电介质特性. 传统的采用 sp^3 的 Slater-Koster 紧束缚方法^[7], 能够满意地描述价带的 $E-k$ 关系, 但不能正确地重现半导体的非直接禁带. 奥地利的 P. Vogl 在 1983 年提出将最外层激发态的 s 轨道包含在供晶体波函数展开的基函数内, 而得到了所谓的 sp^3s^* 的方法 (名称中的 s^* 表示激发态的类 s 轨道, 或反键合轨道)^[8]. 这个方法采用了 13 个参数, 能正确地重现诸如 Si, Ge, C, SiC, GaAs 一类的半导体价带与导带的主要特征.

在经历了 sp^3s^* , sp^3d^2 TB 模型的发展之后, 意大利比萨的 J. M. Jancu 又在 1998 年提出了一个针对 IV 族与 III-V 族半导体 (立方体晶体结构) 的经验型紧束缚方法, 称为 $sp^3d^5s^*$ 的紧束缚方法^[9]. 该方法包括了原子所有的 d 电子轨道作为展开的基, 因而包含有单原子的 10 个电子轨道. 其用到的经验参数是原子本身的 (on-site) 能量与最近邻间的双中心积分值. 可以很好地重现由能带结构决定的对载流子输运有影响的参数, 如态密度、约化质量、与形变势. 这个模型也重现了价带与两个最低的导带的主要特征. 在纳电子材料与器件, 如硅纳米线, 的能带结构计算上得到了十分广泛的应用.

2.2 第一原理计算与密度泛函方法

2.2.1 Hohenberg-Kohn 密度泛函理论 (DFT, 1964)

Hohenberg-Kohn 密度泛函理论, 简称 HK 理论, 的要点是具有相互作用的一个多电子系统在外场 (通常是指晶格势 $v(\mathbf{r})$) 作用下, 该系统的基态能量可以表示为一个依赖于电子浓度 $n(\mathbf{r})$ (本身为空间位置的函数) 的泛函 $F[n(\mathbf{r})]$. 这个方法因此被称为密度泛函理论 (density-functional theory, DFT), 或 DFT 方法.

DFT 方法由两个定理 (theorems) 组成:

1. 一个多电子系统的基态性质为一个在三维空间的电子浓度唯一确定. 通过解依赖于这个电子浓度 (以泛函形式表示) 的单粒子薛定谔方程就可以得到该多电子系统的基态能量.
2. 定义了依赖于电子空间密度的能量泛函, 并且证明了正确的基态电子密度必然导致该能量泛函的最小化.

2.2.2 Kohn-Sham DFT 方法 (1965)

在文献 [5] 中, 当时都在 UCSD 的 Kohn 与 Sham (KS), 提出了一套自洽的方程组, 这是基于 HK 理论来处理有相互作用的非均匀电子气系统的一个近似方法, 被称为 KS-DFT. DFT 方法中的主要困难是如何确定单电子薛定谔方程中的有效电势 (effective potential). 这个有效电势包括电子间的因库仑力引起的各种相互作用, 例如交换关联相互作用. 对于这个交换 (exchange) 关联 (correlation) 能 E_{xc} 的计算, 一个最简单的方法是采用局域密度近似 (local-density approximation, LDA). LDA 交换关联能的得到, 只要求知道一个均匀的有相互作用的电子气的真实化学势 $\mu[n(\mathbf{r})]$, 注意到, 这个量也是电子密度的泛函.

KS-DFT 方法中一共有两组方程, 分别类似于固体物理中用到 Hartree(哈特里) 与 Hartree-Fock(哈特里-福克) 方程组. 这些方程组中用到了局域的有效势. 对于实际的、非均匀的金属与合金, 我们关心的量有相干能, 弹性常数等. 通过解 KS-DFT 方程组可以得到这些宏观量.

类似于 HK 在文献 [4] 得到的对于非相互作用电子气的结果, 对于有相互作用的电子气, 只需解单粒子薛定谔方程就可得到多粒子系统的基态能量与电子密度分布函数.

因为这个单粒子薛定谔方程中的交换关联能本身又依赖于空间电子密度, KS-DFT 方法包含一个方程组, 需要通过迭代求解. 其步骤是:

1. 给 $n(\mathbf{r})$ 一个初始值 (猜解);
2. 构造 $E_{xc}[n(\mathbf{r})]$;
3. 求得一个新的 $n(\mathbf{r})$.

如此循环, 直到自洽的 $n(\mathbf{r})$ 达到. 据此, 就可求得系统的基态能量.

密度泛函理论作为第一性原理计算的理论框架, 对多粒子体系的基态性质给出了非常成功的描述. 在固体中, 对价带结构、体模量、晶格类型与晶格常数、形变势与晶格振动色散关系等都有满意的结果. 但是对于多粒子的激发态, 特别是半导体与绝缘体的带隙, 总是呈现出系统性的偏小, 这被称为“带隙偏小”问题 (还有个名称叫做剪刀 — scissor — 效应, 即如果将带隙剪开, 人为地增加 0.5 eV, 就可以得到大致正确的带隙值). 但对于金属, 则存在着“带隙偏大”的问题. 究其原因, 是因为 DFT 本来就是一个基态理论, 用这个理论得到的本征能量与能带结构没有严格的对应关系. 同时实际的多粒子量子系统的能带结构是与系统内粒子数相关的.

解决这个问题的方法之一是建立一个适合于激发态 (被称为准粒子) 的运动方程, 这就是下面要介绍的 GW 近似 (或称 GWA, A 是近似的意思, GW 的含义下面会介绍).

2.2.3 DFT 第一原理计算的 GW 方法 (1965, 1986)

应用 DFT 方法的关键步骤是找到系统能量作为电子密度 $n(\mathbf{r})$ 的泛函形式, 而这其中的核心问题又是将多电子系统的交换关联能表示成为电子密度的泛函.

美国 Argonne 国家实验室的 Lars Hedin 于 1965 年提出的针对电子气的 GW 方法^[10] 是解决这个问题的一大进步. 但真正成为实用的计算方法是由美国加州大学伯克利分校 (UCB) 的 Hybertsen 於 1986 年在此基础上实现的, 他对半导体与绝缘体的禁带宽度作了计算, 发现结果与实验数据符合极好^[11]. GW 的主要方程为

$$(T + V + U)\psi_{n\mathbf{k}}(\mathbf{r}) + \int d\mathbf{r}' \Sigma(\mathbf{r}, \mathbf{r}'; E_{n\mathbf{k}})\psi_{n\mathbf{k}}(\mathbf{r}') = E_{n\mathbf{k}}\psi_{n\mathbf{k}}(\mathbf{r}), \quad (1)$$

其中 U 是由于电子间的相互库伦力作用的平均哈特里电势; Σ 是电子自能算符, 包含了电子间的交换关联效应. Σ 是一个非局域、依赖于能量的非厄密 (non-hermitian) 矩阵, 由上述方程解得的本征波函数 $\psi_{n\mathbf{k}}(\mathbf{r})$ 代表准粒子 (quasiparticle) 态, 而本征值 $E_{n\mathbf{k}}$ 是一个复数, 具有准粒子的激发 (excitation) 能的物理意义, 其虚部代表准粒子的寿命. 需要注意的是, 因为方程中的自能本身又依赖于本征能量 $E_{n\mathbf{k}}$, 所以该方程的解要通过迭代求得.

现在的难点是找到一个恰当的对于自能算符的近似. Hedin 采用了如下方法: 将 Σ 展开为一个因带电粒子间库伦相互作用引起的微扰的一个级数, 且只保留级数的第一项:

$$\Sigma \propto GW, \quad (2)$$

其中 G 是单粒子着装 (dressed) 的格林函数, W 代表动态粒子间被屏蔽后的相互作用势³⁾.

GW 方法也被称为随机相位近似 (RPA: random phase approximation). 因为该方法是基于一个精确的多体场理论的, 因此被普遍认为比 LDA 更为自然的方法. 但 GW 方法也存在着一个问题, 即是对同一个物理量计算得到的结果在元素周期表中的不同族 (group) 间会有变化. 这个情况与 LDA 方法应用的早期是类似的. 其原因是因为在计算过程中作了许多种的近似, 尤其是假设自能是由 $G^{\text{LDA}}W^{\text{LDA}}$ 计算得到的. 其他一些严重影响结果的精度的近似包括忽略了原子核芯部分的电子态 (core state) 以及用赝势近似来取代核部分等.

2.2.4 Faleev-Schilfgaarde 准粒子自洽 GW 方法 (QPscGW)(2004)

2004 年美国 Sandia 国家实验室的 Sergey V. Faleev 与亚利桑那州州立大学 (ASU) 的 Mark van Schilfgaarde 提出了一个被称为准粒子自洽 GW 的方法 (quasiparticle self-consistent GW: QPscGW)^[12]. 这个方法消除了上面提到的在经典的 GW 方法中所做的一些近似, 使得 GW 方法更为可靠与准确.

采用 DFT-LDA 方法计算得到的带隙几乎对所有的半导体都偏小. 小的带隙使得 G 增加了屏蔽效应, 因而低估了对带隙的纠正. 而由大带隙的准粒子构造的 G 则减少了屏蔽效应, 导致 GW 产生宽的带隙.

QPscGW 这个方法采用了包括全部电子与完全电势自洽的 GW 方法的实现, 在最大限度上保留了独立粒子的图像. 因此这个方法被命名为准粒子自洽 GW 方法, 即 QPscGW 近似. 该方法的一个基本点是除了 RPA 近似外, 不再做其他关键性的近似.

QPscGW 方法对几乎所有的化合物, 特别是弱关联材料如硅的光学性质都可以很好地描述. 尽管对强关联的带有 d 与 f 轨道电子的材料误差有些大, 但是这些误差是系统性的, 并且能用相比于精确理论而能确定的丢失部分来解释误差的产生. 这个方法中的自洽特点是十分重要的, 它大大地改善了与实验结果的一致性.

QPscGW 理论是一个第一性原理计算方法, 可以产生单粒子经历的真实电势, 因而能计算得到可靠的准粒子激发能.

2.2.5 KS-DFT 的其他一些改进方法

除了用 LDA 来构造基态交换关联能外, 还有一种方法称为 GGA (generalized gradient approximation).

其他的一些对 LDA 的交换关联势进行改进的方法有⁴⁾:

- 准粒子 GW 近似;
- sX-LDA (对半导体激发态);
- 自相互作用修正 (SIC);
- LDA+u (对强关联系统);
- LDA++;
- 优化有效势 (OEP) 方法 (对强关联、激发态);
- 含时密度泛函理论 (TDDFT) (原子、分子与固体的激发态).

3) W 这个符号的来源很可能是因为用 V 来表示裸露的离子引起的电势, 因而着装的电势就用英文字母表中 V 后的下一个字母, 即 W 来表示

4) 取自百度文库, 黄美纯 “DFT-8 凝聚态物理中的 GW 近似”

2.3 载流子量子输运

当纳米尺度的半导体器件, 包括一维纳米材料电子元器件如用硅纳米线、碳纳米管 (CNT)、石墨烯纳米带 (graphene nano-ribbon, GNR) 做成的二端电阻的特征尺寸与载流子的输运特征长度 (最常用的当属载流子平均自由程, 但载流子相位相干长度在决定是否需要考虑量子输运上起更重要的作用) 可比拟时, 经典的载流子扩散漂移模型不再适用. 代之以, 我们要考虑载流子的弹道输运 (平均自由程大于器件特征长度) 与量子输运 (相位相干长度大于器件特征长度). 我们将这种尺度下的器件工作状态称为“介观态”, 其相应的器件物理称为介观物理, 即介于宏观物理与微观物理之间.

这一节, 我们先描述决定量子电导的 Landauer-Büttiker 公式, 这是针对在导体中载流子弹道输运为主的情形, 同时包括了接触电阻的影响. 然后讨论近 20 年来逐渐成为计算纳米尺度器件的电特性 (I - V , C - V) 的非平衡格林函数 (NEGF) 方法.

2.3.1 弹道输运与 Landauer-Büttiker 公式

在 20 世纪 80 年代, 由于半导体器件中二维电子气的实现与微细加工能力的进步, 在实验上观察到介观尺度 (即与电子波长可比拟, 在微纳米数量级) 下导体的量子化了的电导 (简称量子电导, 为 1988 年由荷兰 Delft 大学的 van Wees 与英国剑桥大学的 Wharam 同时发现^[13,14]).

处于介观态的器件内部载流子输运呈现出明显的波动特性, 与量子力学的其他效应, 如波函数的模式数 (在量子约束维度方向), 一起决定了弹道导体的电导. 早在 1957 年 IBM 的 Landauer 就提出了对这种两端导体一个计算电导的公式^[15]. 然后在 1988 年由 IBM 的 Büttiker 扩展到多电极 (多于二个电极) 的情形^[16]. 这个公式以后统称为 Landauer-Büttiker (LB) 公式 (formalism), 该公式由于明确地考虑了波函数相位对输运的影响, 在解释介观结构中各种量子相干输运现象时获得了很大的成功.

对于一个两端接电极的弹道导体, 假设由一侧入射的电子波经过导体到另一侧时透射率为 T , 则反射率为 $R = 1 - T$ (我们此时将导体看作是一个散射势垒). 两端的电极被视为处于热平衡的“电子库”, 用各自的化学势 μ_1, μ_2 表征. LB 公式给出的整个系统的电导 (包括电极与导体的接触电阻) 为

$$G = \frac{I}{\mu_1 - \mu_2} = \frac{2e^2}{h} MT \quad (\text{假设 } \mu_1 > \mu_2, I \text{ 为正}). \quad (3)$$

上式中, e 为电子电荷, h 为普朗克常数, M 是在传输沟道 (具体来讲是连接被测导体与电极间的窄导线) 中电子波函数的模式数, 为一正整数. 系数 2 是考虑到电子的自旋简并. 这个式子传递了两个重要的与经典物理预测不同的信息: 1) 即使导体本身的电阻为零 ($T = 1$, 即在导体中完全没有散射的载流子弹道输运), 整个系统 (或称装置, set-up) 依然存在着电阻, 这个电阻当 $M = 1$ 与 $T = 1$ 时为 $12.9 \text{ k}\Omega$; 2) 与经典的欧姆定律不同, 电导不再与导体的长度相关, 与导体的截面也没有直接的成正比关系 (欧姆定律: $G = \sigma A/l$, 其中 σ 为电导率, A 为截面, l 为导体的长度), 而是由模式数 M 与透射系数 T 决定. 而且因为 M 是整数, 因此弹道导体的电导是量子化了的. 注意, 一个非零的接触电阻总是存在, 其来源是载流子在进入电极这个电子库区时发生的能量损耗而引起的.

对于有二个以上电极的情形, 上面公式中的 T 要用散射矩阵 S 来代替. 具体的方法在文献 [16] 中有详细讨论.

2.3.2 量子输运的非平衡格林函数 (NEGF) 计算方法

纳米尺度下的半导体器件 (上述的弹道导体也包括在内) 是一个开放的量子系统. 这是因为一方面, 器件与周围环境 (主要就是器件所在的电路) 的交互作用是通过电极来进行的 (这就是开放的意

思), 同时载流子在器件内部的行为由量子力学的原理 (即薛定谔方程) 所决定 (即为量子系统). 器件的电学 (甚至光学或光电) 行为原则上通过解这个具有开放边界条件的薛定谔方程就可以确定. 量子力学中有用格林函数解薛定谔方程的做法, 但并不在解封闭量子系统上常用. 对这种封闭量子系统, 通常是采用解本征值问题来求得量子系统的本征函数 (即波函数) 与本征值 (即本征能量). 但是在处理如半导体器件这样的开放量子系统, 格林函数, 特别是非平衡格林函数法是一个十分有效的计算方法. 特别是 1995 年美国普渡大学 Datta 教授的《介观系统的电子输运》^[6] 出版以来, 对普及 NEGF 这个方法起了十分有重要的作用. 目前, NEGF 已成为纳电子器件模拟的一个最通用的方法. 下面, 我们花一定篇幅来介绍 NEGF 这个理论与方法.

格林函数法是数学物理方程中一种常用的方法. 在一般的数学物理 (比如静电场) 应用上, 格林函数是用来表示一个源点上的某种信号源 (比如电激励) 在场点上产生的场. 例如, 泊松方程中点电荷在空间各处产生的静电场. 格林函数应用的一个优势是其线性叠加性. 即在不知道激励源的分布情况下, 根据无源系统的描述 (用算符与边界条件表述) 就可以得到格林函数. 然后给定源的分布, 可直接利用格林函数来积分/求和得到场的分布. 但在量子力学中应用的格林函数与上述通常的描述有较大的不同, 这也是为什么 NEGF 方法在很长一个时间 (甚至到现在也是) 没有得到广泛的应用.

量子力学中应用格林函数的目的是解薛定谔方程. 而薛定谔方程并没有独立于待求解量 (即波函数与本征能量) 的激励源这一项. 因此我们需要将格林函数满足的方程扩展成

$$(z - L)G(\mathbf{r}, \mathbf{r}'; z) = \delta(\mathbf{r} - \mathbf{r}'). \quad (4)$$

这里 z 是一个复数参量, 可记为 $z = \lambda + is$. 对比薛定谔方程的形式, 我们立即可以看出, 如果算符 L 是哈密顿量 H , 那么 z 就相当于本征能量, 只不过这里我们允许本征能量为一复数. 如果 $z = \lambda$, 为一实数, 则 λ 就代表能量. 采用这样定义的格林函数, 有许多优异的特性, 尤其是对纳电子器件而言. 除了由得到的格林函数直接知道 (薛定谔) 方程的本征函数与本征值外, 还可以知道简并能级所对应的简并态数^[17]. 与纳电子器件更有关联的是由格林函数可以引出/得到局域态密度的概念与值, 进而得到能量空间的态密度.

对于开放量子系统, 格林函数或非平衡格林函数法通过引入自能的概念, 可以将一个无穷的系统 (比如器件与其引线) 变成一个仅局限于所关心的器件本身的有限区域, 而且能理论上严格地将器件内的载流子散射 (可以由各种散射机理引起) 用自能的概念处理 (即所谓的 Büttiker probe). 从应用上来讲, NEGF 方法可以自然地隧穿电流自然地包含在解之内. 因此可以这么讲, NEGF 方法的普及将量子力学真正应用到半导体器件模拟上去.

格林函数方法也是处理相互作用的多粒子体系、乃至多体物理的一个有力工具. 该方法将粒子间相互作用作为对单粒子哈密顿量的一个微扰^[18].

3 低维材料物理 (包括铁基超导)

场效应晶体管的出现, 在半导体衬底的表面自然地形成了二维电子气. 这是低维物理与器件研究的开端. 量子电导现象就是在二维结构上首先发现的^[13,14]. 上世纪 90 年代后期兴起的碳纳米管、半导体 (主要是硅、锗) 纳米线研究又将一维结构与物理的研究推向了堪称为纳米技术研究一度的核心地位. 但 2004 年石墨烯的发现^[19,20] 又将二维结构置于纳米舞台的中心.

与一维结构的研究不同, 二维结构的行为与磁场的关系密切 (霍尔效应), 石墨烯的发现是一个例子. 2006 年以来的拓扑绝缘体研究^[21] 也与电子自旋和轨道磁矩的相互作用相关. 即使是我们

下面将要讨论的铁基超导体 (在该领域中国的研究人员十分活跃, 作出了重要的贡献) 领域, 磁矩与二维层状结构也起了关键性的作用.

这一节, 我们讨论一维与二维纳米结构的研究历史与现状. 内容涵盖石墨烯、拓扑绝缘体与铁基超导.

3.1 碳纳米管^[22]

1991 年, 日本 NEC 公司基础研究实验室的电子显微镜专家 Iijima 在高分辨透射电子显微镜下检验石墨电弧设备中产生的球状碳分子时, 意外发现了由管状的同轴纳米管组成的碳分子, 这可以被认为是世界上最早有意识地发现的碳纳米管 (carbon nanotube, CNT)^[22]. 1993 年, Iijima 和 Bethune 同时报道了采用电弧法, 在石墨电极中添加一定的催化剂, 可以得到仅仅具有一层管壁的碳纳米管, 这也成为了最早关于碳纳米管制备的文献报道.

碳纳米管由于其独特的电学、热学与机械性质, 其相关的研究在 90 年代中后期以及 21 世纪初期的几年里, 呈现过异常迅猛的态势. 但在近几年来, 出现了下降的趋势. 这与材料制备的可控性与一致性差有关, 而且 CNT 作为电子器件的应用必须要强调在集成的环境下实现, 这是至今几乎所有纳米材料相比于半导体材料的致命的弱点: 不易集成有规模的电路, 集成后的器件性能比单个器件性能要差很多, 更谈不上与进入 32/22 nm 技术节点的硅 CMOS 器件相比. 下面简单总结一下碳纳米管的结构与性质.

3.1.1 碳纳米管的基本结构

碳纳米管可以分为单壁 (SWNTs) 和多壁 (MWNTs) 两种结构. 绝大多数单壁碳纳米管的直径大约只有一个纳米左右, 而长度却能达到微米量级. 单壁碳纳米管可以简单地被认为是有一层碳原子层, 即石墨烯, 卷曲成管状圆柱体. 由石墨烯而成的碳纳米管可以用一对整数 (n, m) 表征不同的卷曲方式, 从而使得单壁碳纳米管依其结构特征可以分为 3 种类型: 即扶手椅式纳米管、锯齿形纳米管和不属于这两个类型的其他手性型纳米管. 单壁碳纳米管的重要性表现在其多数性质性能与 (n, m) 相关, 尤其是其禁带宽度可以在 0 到 2eV 范围内调制, 从而使纳米管的电学性能表达为金属特性或者半导体特性.

3.1.2 碳纳米管的性质

碳纳米管的特殊结构使其具有多种特异的性质, 并表现出在各领域的不同应用前景. 如前所述, 碳纳米管具有纳米量级的孔直径, 同时长度能够达到微米量级. 有很多研究工作关注于单壁纳米管的性质与结构上的关系, 并由此发现了很多新的现象. 多壁碳纳米管的性质也与其结构和质地有关. 在这里, 我们简单总结一下碳纳米管的各个方面的性质性能.

1. 力学性质:

在抗拉强度和弹性模量上, 碳纳米管是目前发现的最强最硬的材料. 这一力学上的显著特点来源于碳原子之间的 sp^2 杂化轨道. 相比 sp^3 杂化, sp^2 杂化杂化中 s 轨道成分比较大, 从而使碳纳米管具有高模量、高强度. 碳纳米管的硬度与金刚石相当, 却拥有良好的柔韧性.

2. 热学性质:

由于存在弹道输运, 而垂直于管轴方向却有良好的绝缘性, 因此几乎所有纳米管结构的材料都被认为具有良好的热学传导性质. 实验测量表明, 室温下的单壁碳纳米管在轴方向上的热传导能力与铜相当.

3. 光学性质:

碳纳米管的光学性质包括光吸收、光致发光性质以及拉曼光谱中的信号响应. 碳纳米管, 尤其是多壁碳纳米管具有良好的微波频段波的吸收特性. 通过多壁碳纳米管中填入金属材料, 能够提高微波区的波吸收效率.

4. 电学性质:

前文中提到, 表征碳纳米管不同构成方式的一对整数 (n, m) 往往和其材料性质有关, 包括电学性质. 如果 $n = m$, 碳纳米管表现为金属性; 当 $n - m$ 为 3 的整数倍时, 纳米管为带隙极小的半导体特性; 除此之外的纳米管为一般的半导体特性. 当纳米管孔径很小的时候, 曲率效应也会改变上述碳纳米管电学性质的规则. 碳纳米管的电学特性也表现在超导性质上. 内部单元纳米管相连的多壁管能够在相对较高的转变温度下 (12 K) 表现出超导特性.

3.1.3 基于碳纳米管的电子器件

1. 场效应晶体管:

1998 年, 第一篇基于碳纳米管的场效应晶体管的文献发表. 具有半导体特性的纳米管作为构成沟道的功能材料. 由于纳米管孔径很小, 有较好的栅极与沟道的耦合, 使得栅极可以控制整个沟道的电势, 从而能够制备不受短沟道效应影响的更小尺寸的器件. 另一方面, 碳纳米管的良好表面也避免了界面态和粗糙度引起的载流子散射, 因而通常具有高的迁移率. 此外, 也有研究工作关注于基于碳纳米管晶体管结构器件的交流电状态下的性质. 理论预测短的纳米管能作为太赫兹信号源.

2. 集成电路:

碳纳米管器件目前已经突破单一器件研究与开发的局限, 可以用来制备更加复杂的电路. 美国 Stanford 大学也有三维集成的有益尝试.

3. 交叉点结构电阻型存储器:

基于交叉点结构 (cross-bar) 的存储器件由于其非常简单的器件构成而受到广泛的关注. 通过对碳纳米管掺杂硼和氮, 之后分散在聚苯乙烯的介质中, 可以得到具有电学转换行为的功能层材料, 从而实现交叉点结构的存储器件. 此外, 这一类方法往往能够将器件制备在柔性衬底上.

4. 其他电子器件:

碳纳米管在电子器件中的潜在应用非常的广泛, 除了上述几种常见的重要基本器件之外, 碳纳米管还被应用在太阳能电池与发光二极管等器件中.

3.2 石墨烯^[19]

石墨烯 (graphene) 在 2004 年的发现与物理性质的确认^[19], 无疑是固体物理本世纪以来的一个重大里程碑事件. 尽管在此前, 碳纳米管 (CNT), 自 1991 年由日本 NEC 公司的 Iijima 确认以来, 在 20 世纪末期 (1990 年代后期) 与 21 世纪初期占据了纳米技术研究的核心舞台, 但石墨烯的出现以其独特的理论价值与优异的电学、光学性质很快地形成了势头可以与 CNT 比拟、甚至压到的世界范围内的研究热潮. 其发明者, 英国曼切斯特大学 (Univ. Manchester) 的 Novoselov 与 Geim(N-G) 也在发表他们的第一篇 Science 文章 (2004)^[19] 的 6 年后就得到了诺贝尔物理学奖 (2010 年).

石墨烯的发现是有其强烈的集成电路发展驱动背景的. 为了克服场效应晶体管 (FET) 的短沟效应 (现在 CMOS 已处在 22 nm 技术节点), 一个有效的手段是减薄沟道区的厚度. 当然最薄可能达到的就是单原子层二维材料. 另一方面, 金属有大量的电子作为沟道载流子, 但问题是如何用场效应的

手段来实现栅控. 这两个目标是驱动石墨烯研究原始动力的重要因素 (尽管还可能有别的因素). 石墨烯的发现有其偶然成分 (如置于长在硅片上 300 nm 厚二氧化硅层上的石墨烯恰好能用光学显微镜分辨出来), 但 N-G 及其研究小组, 做了十分细致的实验工作与理论分析, 包括用量子霍尔效应 (QHE) 来定性地区分石墨烯 (单层, 半整数 QHE) 与二层石墨材料 (整数 QHE), 这在次年 (2005) 的 Nature 文章上得以近似完美的表述 [20].

有关石墨烯的研究, 可以从物理性质、理论价值与材料制备三个方面来说明. 我们主要是从集成电路的应用方面来加以阐述. 石墨烯是单层的碳原子以正六角形晶格排列, 一个碳原子的 4 个价电子有 3 个与位于同一平面中最邻近的碳原子成键 (原子间的键长是 1.42 Å). 剩下的一个价电子在平行于碳原子所在二维平面的近邻平面上形成大 π 键. 这些 π 电子是石墨烯的活动载流子, 其运动所受的阻力很小. 因此石墨烯的迁移率很高. 对架空 (suspended) 的石墨烯测量的迁移率最高的可达 $2 \times 10^5 \text{ cm}^2/\text{V}\cdot\text{s}$ [23] (相比之下电子在体硅材料的迁移率为 $1450 \text{ cm}^2/\text{V}\cdot\text{s}$). 石墨烯的能带结构决定了以下的载流子特性:

- 载流子的饱和速度为 10^8 cm/s (为光速的 $1/300$), 相比之下, 电子在硅中的饱和速度为大约 10^7 cm/s .

- 在第一布里渊区的狄拉克点 (有 6 个), 带隙为零, 因此石墨烯是半金属.

- 处于中性状态 (即电子与空穴浓度均为零, 或称费米能级通过狄拉克点) 的石墨烯, 载流子的静止有效质量为零, 其输运由狄拉克相对论方程决定, 而相应的 “光速” 即是上面提到的饱和速度 (又称费米速度, 也是由在狄拉克点的能带结构所决定). 因而石墨烯是研究相对论物理的一个很好的平台, 因为其要达到的光速仅为实际光速的 $1/300$.

- 石墨烯中费米能级相对于狄拉克点的位置 (在电子能量空间中) 可以由栅电极控制/调节. 当费米能级在狄拉克点之上时, 石墨烯为 n 型, 是电子导电; 而费米能级在狄拉克点之下时, 石墨烯为 p 型, 是空穴导电. 这就是被称为双极性 (ambipolarity) 导电的行为.

所以, 从理论上讲, 石墨烯提供了一个在普通实验室条件下研究相对论的平台: 在狄拉克点附近的载流子是静止有效质量为零的费米子, 而其要呈现相对论效应所对应的光速仅为 10^8 cm/s . 从集成电路的应用角度来看, 石墨烯令人吸引的优点是高载流子迁移率及容易实现场效应晶体管 (栅可控沟道载流子的面密度). 但石墨烯的半金属 (带隙为零) 特性, 不利于其作为数字电路器件的应用, 因为漏电流太大, 导致器件的开/关电流比太小.

迄今在石墨烯作为电子器件的应用研究方面, 主要集中在两个方面: 1) 设法在石墨烯作为沟道材料时打开带隙; 2) 在石墨烯集成到器件时 (主要是场效应管), 能最大限度地保持高迁移率. 打开带隙的一个典型办法是将片状石墨烯切成纳米带 (nano-ribbon, 典型的宽度为几个纳米), 但这样做的一个负面效果是沟道载流子的迁移率会变得非常差, 相对于硅 MOSFET, 毫无优势可言. 在经历了数年的研究之后, 石墨烯的电子器件应用逐渐转向模拟、射频器件的研究. 因为在模拟、射频电路的应用中, 器件的漏电流影响并不大, 所以可以用石墨烯片来维持高的沟道载流子迁移率. 最新的研究结果 (IBM, Nature Lett., 2011) 是 CVD-graphene (置于类金刚石碳基衬底) 射频晶体管, 顶栅长 40 nm, 其截止频率可达 155 GHz.

3.2.1 石墨烯电学特性的基础理论研究

关于石墨烯的电学特性的理论研究, 实际上早在 20 世纪 40 年代就有相关科技文献. 当然, 那时二维单层原子膜还仅仅是作为理论研究的简单模型出现. 目前, 关于石墨烯电学特性研究值得关注的领域是, 石墨烯中起到电荷载流子作用的激子是零质量且具有手性的狄拉克费米子, 这带来了各

种新奇的物理现象. 研究表明, 中性石墨烯的化学势位于狄拉克点处, 使得在低能量时, 可以在石墨烯中模拟零质量费米子的各种量子电动力学的物理效应. 同时由于其饱和速度又比光速低 300 倍, 因此在实验中能观测到很多有趣且具有重大科学意义的实验现象. 在磁场的作用下, 狄拉克费米子有与普通电子相异的表现, 并导致新物理现象, 如室温下就可出现整数量子霍尔效应等. 此外, 对石墨烯电学特性研究的另一重要性是基于狄拉克费米子的克莱恩佯谬 (Klein Paradox). 这一特殊电学行为能够以数值为 1 的概率隧穿过经典电动力学架构下的高宽势垒, 并因此使得狄拉克费米子不受外加电场的影响. 而同时, 狄拉克费米子又会在约束势的作用下表现出 Zitterbewegung 现象, 即波函数的抖动行为. 由于材料中的缺陷能产生约束势, 因而很多研究工作在此理论模型上深入研究了缺陷对于石墨烯中电子物理行为的影响. 例如, 不同的缺陷对石墨烯中电子传输具有不同的影响特性, 如点缺陷、线缺陷和表面杂质缺陷, 并且迁移率主要受到颗粒界面处线缺陷的影响, 这也正是用化学气相沉积 (CVD) 方法制备的石墨烯样品的载流子迁移率大大低于通过微机械分离方法制备的样品的原因. 类似的工作有助于不断深入理解微尺度下相对论粒子的量子力学行为, 同时也为进一步提高石墨烯器件性能奠定坚实的理论基础.

3.2.2 基于石墨烯的晶体管器件与禁带开启问题

对于传统场效应晶体管来说, 沟道部分的半导体材料须要具有 0.4 eV 或者更高的禁带宽度. 而大尺寸的石墨烯却体现出半金属性质, 禁带宽度为零, 其导带和价带于能带图中是锥形结构, 在布里渊区的 K 点重合. 由于零禁带宽度, 倘若用大尺寸的石墨烯作为功能层, 相应的晶体管器件无法被调制到关闭状态, 从而无法将石墨烯用于逻辑电路的应用中. 尽管如此, 石墨烯的能带结构是可以被调制的, 而且目前也已经有多种方法可以打开石墨烯禁带宽度, 例如双层石墨烯上加载电场、制备小尺寸的石墨烯纳米带、附加应力在石墨烯上, 以及通过有机物与石墨烯的相互作用, 等等.

3.2.3 石墨烯的其他应用

这里, 仅举几个例子:

1. 高电导性能的透明电极:

因为石墨烯是透明的, 用它制造的电极比其他材料具有更优良的透光性. 通过使用石墨烯, 得到的薄膜电导率可以达到 550 S/cm, 同时 1000~3000 nm 波长光的透过率高于 70%, 因而可以用于太阳能电池和触摸显示屏等器件应用中. 此外, 作为电极的石墨烯薄膜还具有较好的化学和热学稳定性.

2. 高速电池:

美国俄亥俄州 Nanotek 仪器公司的研究人员利用锂离子可在石墨烯表面和电极之间快速大量穿梭运动的特性, 开发出一种新型储能设备, 可以将充电时间从过去的数小时之久缩短为不到一分钟^[24]. 新型石墨烯电池实验阶段的成功, 无疑将成为电池产业的一个新的发展点.

3.3 拓扑绝缘体^[21]

3.3.1 引言

整数与分数量子霍尔效应揭示了物质的一种量子态, 即在外加磁场下, 一个二维电子气系统在“体内”是绝缘的, 而在体的表面或边缘存在着来回导电的量子态. 本世纪开始以来, 物理学家逐渐意识到, 即使没有外加磁场, 对于某些具有强的电子自旋-轨道耦合效应的绝缘体, 一种鲁棒的导电表面/边缘态也可能存在. 这种特别的绝缘体被称为“拓扑绝缘体”(topological insulator, TI). 在这种绝缘体中,

自旋-轨道耦合效应就起了外加磁场的作用, 在体的边缘, 不同朝向的自旋电子沿着边缘反向运动, 而形成导电的路径.

二维的拓扑绝缘体 (具体的是由碲化汞, HgTe, 形成的量子阱) 最早为美国 Stanford 大学张首晟教授的研究组于 2006 年在理论上预测到 [21], 在 2007 年在实验上被观察到 [25]. 而 Univ. Penn(宾夕法尼亚大学) 的 Liang Fu 与 Charles Kane 在 2007 年预测了铋锑合金 ($\text{Bi}_{1-x}\text{Sb}_x$) 作为一个三维固体, 可以存在导电表面态, 因而是 TI 材料 [26]. 在同一篇文章中, 他们也预测了在铋锑合金体内有三维狄拉克费米子的存在. 次年 (2008), 这两个预测都在普林斯顿大学的光发射实验中得以证实 [27].

拓扑绝缘体与一般绝缘体的差别是由晶体布里渊区中的 Z_2 不变性来表征的, 其体内的激发态带隙因自旋-轨道相互作用而产生. 这被称为量子自旋霍尔相 (quantum spin-Hall, QSH). 拓扑绝缘体, 或称 QSH 相的特征是在绝缘体的表面 (三维) 或边界 (二维) 上存在着带隙为零的能态. 从细致的能带结构上看, 二维的 TI 有单个 Z_2 不变性, 而对三维 TI, 则有 4 个. 三维的 TI 又分为“弱”与“强”两类. 对于二维与三维强拓扑绝缘体, 这些导电态因为没有带隙是鲁棒的, 它们不因弱的无序性 (disorder) 与相互作用而发生变化.

文献 [26] 中提出了对 Z_2 不变性求值的方法来预测哪些材料具备强 TI 特征. 这些材料包括具备半导体性质的铋锑合金以及在单轴应力下的 α -Sn 与 HgTe.

作为 TI 材料, 其晶体结构要具有反演对称性, 同时显现出宇称本征值. 在拓扑绝缘体中, 导电的表面态是得到保护的, 不会因固体表面形变而发生变化. 这个独立于拓扑变化的原因是来源于数学上, 描述表面态的哈密顿量不因小的扰动而发生变化. 这些导电的表面态可以被测量到, 而且在量子自旋霍尔态的情况下, 这些态是自旋极化的.

下面, 我们给出相关拓扑学的基础知识, 拓扑绝缘体的研究现状与中国研究人员在这个课题研究中起到的重要作用与贡献.

3.3.2 拓扑、拓扑数和拓扑绝缘体

1. 拓扑学相关知识:

在拓扑学里不讨论两个图形全等的概念, 而是讨论拓扑等价的概念. 比方说, 尽管圆和方形、三角形的形状不同, 但在拓扑学中它们是等价图形, 通过拓扑变换可以使它们互相转化. 因此, 从拓扑学的角度来看, 它们是完全一样的. 如有几种形状不同的三维多面体, 虽然它们的形状不一样, 但是可以通过计算发现: 顶点数 (v) 减去棱数 (e) 再加上上面的数目 (f), 对于这些多面体都是 2. 那么是不是对所有的多面体而言, $v - e + f = 2$ 都成立呢? 答案是否定的, 只有正多面体和凸多面体满足上面的公式. 定义多面体的欧拉 (Euler) 数为 $v - e + f$, 具有相同欧拉数的多面体之间是拓扑等价的, 它们被称为同胚. 假设有一个光滑球面, 在上面任选一些点用不相交的线把它们连接起来, 这样球面就被这些线分成许多块, 然后让球面变成光滑的平面, 就可以变成一系列的 4 种图样, 因此其拓扑数一致是不难理解的. 在拓扑变换下, 点、线、面的数目仍和原来的数目一样, 这就是拓扑等价. 一般来说, 对于任意形状的闭曲面, 只要不把曲面撕裂或割破, 它的变换就是拓扑变换, 就存在拓扑等价. 所以可以说在拓扑学中, 一个实心小球无论怎么进行拓扑渐变, 它都不可能变成甜甜圈. 对于甜甜圈, 我们可以通过拓扑等价得到一个中空的立方体, 通过计算可以得到, 对于甜甜圈类型的物体, 其欧拉数是 0. 因此, 对于一个实际物体无论怎么进行拓扑变换, 其欧拉数都是恒定, 即欧拉数守恒是受拓扑学保护的, 它是一个拓扑不变量. 对于类似欧拉数的其他拓扑不变量, 只要是拓扑变换, 那么拓扑不变量的值就不会改变. 这种性质对于能带中的拓扑结构也是相似的, 描述能带中的拓扑结构的拓扑不变量在拓扑变换下是不会改变的.

2. 拓扑绝缘体的引入:

但是并不是所有的绝缘体都具有和真空一样的拓扑结构, 拓扑绝缘体就是这样一个反例. 由于其具有不同于一般绝缘体的表面态特性, 这种特性受拓扑性质保护, 所以这种绝缘体就称为拓扑绝缘体 (topological insulator). 在拓扑绝缘体中, 其体内的态与普通绝缘体中并没有不同, 费米面处于导带和价带之间, 但是我们能看见有两条能带穿越了禁带, 它们对应的是拓扑绝缘体的金属型表面态. 事实上, 拓扑绝缘体在边界上会维持稳定的低维金属型态, 这些无能隙的边缘激发态对应的能带处在禁带之中, 连接着价带顶和导带底, 表现出类似金属的性质. 为什么说这种性质是受拓扑学保护呢? 由于强烈的自旋轨道耦合, 拓扑绝缘体电子轨道会发生 sp 电子反转, 体态内的拓扑结构就必然不同于真空. 如前面已阐明的, 拓扑变换并不能改变一种材料内部的拓扑结构, 所以在体态和外界的交界处存在奇异的拓扑学现象, 这就是受拓扑学保护的拓扑表面态. 物理学界之所以对拓扑绝缘体充满兴趣, 其原因之一就是因其表面态存在的这种特别性质.

3.3.3 三维拓扑绝缘体的发现及实验证据

1. HgTe/CdTe 量子阱 — 拓扑绝缘体的初次发现:

为了得到更大的能隙, 需要更强的自旋轨道耦合, 因此 Z_2 奇异结构必须在重原子材料中寻找. 2006 年, Bernvig, Hughes 和张首晟指出^[21] 可以在 HgTe 和 CdTe 形成的量子阱结构中实现这一拓扑转变, 其中由自旋轨道耦合导致的能带反转是实现量子自旋霍尔效应的关键. CdTe 中, 导带的电子是由 s 轨道贡献的, 价带则是由 p 轨道的电子形成. 但在 HgTe 中, 由于自旋轨道耦合很强, 引起 s - p 轨道之间的反转. 这种能带反转会导致材料能带出现非平庸的拓扑结构, 从而实现二维量子自旋霍尔效应. 在 CdTe-HgTe-CdTe 量子阱中, 可以通过调节 HgTe 层的厚度来实现正常能带结构到反转能带结构的转变. 2007 年, König 等人制备了这种量子自旋霍尔绝缘体并测量了其输运性质^[25]. 这种低温下的弹道边缘态可以用 Landauer 等人的工作解释, 这些边缘态是由化学势导致的. 这样每个边缘就产生了量子化的电导 e^2/h . 这是拓扑绝缘体的最早的实验观察.

2. 时间反演对称性保护 (Z_2) 的三维强拓扑绝缘体 — 理论预言和实验实现:

(1) 第一类三维拓扑绝缘体:

2008 年, 普林斯顿的 Hsieh 等人通过角分辨光电子谱 (angle-resolved photoemission spectroscopy, ARPES) 观测到 $\text{Bi}_{1-x}\text{Sb}_x$ 材料表面的表面态电子态和狄拉克型色散关系^[27], 从而证实了以上理论预言, 这是实验上首次对三维拓扑绝缘体的报道. 从温度电阻关系上看, $x = 0.1$ 时, 材料电阻具有负温度系数, 这说明材料已经具有一个带隙, 约为 0.36 eV. 他们还利用角分辨光电子能谱描绘了表面态的色散曲线. 可以看到, 表面能级与费米面交叉 5 次, 这说明该材料就是之前预言的强拓扑绝缘体. 然而, 这种材料的缺点是一种合金, 其结构和表面态性质较为复杂.

(2) 第二代拓扑绝缘体:

2009 年, 方忠和张首晟等人合作报道了他们对 Bi_2Te_3 , Bi_2Se_3 以及其他一些材料的能带计算结果^[28], 表明前面两种材料都是强拓扑绝缘体, 并给出了其表面的单个狄拉克能谱的低能有效模型. 这些材料的价带和导带被线性的表面能带所连接, 并且具有单个狄拉克锥. 与此同时, 普林斯顿的研究小组也报道了他们对 Bi_2Te_3 , Bi_2Se_3 的角分辨光电子谱的实验结果, 确定了这两种材料都是表面只有一支的狄拉克能谱的强拓扑绝缘体. 结果显示, 表面态具有狄拉克锥型结构.

(3) 其他拓扑绝缘体:

以上工作建立了一些简单的原则, 可以用来判别一种材料是不是拓扑绝缘体. 比如, 对于一个具有反演中心的体系, 判断其宇称的变化; 对于拓扑数和 Berry 相的计算等. 所以 2009 年以来大量的

理论工作被开展用以搜索新型的拓扑绝缘体和狄拉克电子体系, 如 NaCoO_4 , AmN 和一些多元化合物等等. 其中 AmN 被认为可作为一个可利用拓扑表面态快速散射的核燃料. 也有理论同行提出在烧绿石结构中存在一种“三维的狄拉克点”—费米弧. 虽然其散热机制不一定能实现, 但是该工作本身即是拓扑绝缘体理论领域的一个重要进展. 而实验上的工作则仍然较多关注 Bi_2Se_3 家族, 近来也发展到 $\text{Bi}_2\text{Te}_2\text{Se}$ 等的制备和研究.

3.3.4 拓扑奇异表面态的输运表现

1. 扫描隧道谱:

2010 年 8 月, 清华大学薛其坤小组首次使用扫描隧道显微镜在线表征 Bi_2Se_3 薄膜, 测量其扫描隧道谱, 并在垂直于膜面 (111) 方向加上磁场, 测量到了 Bi_2Se_3 薄膜表面准二维电子气的朗道量子化现象. 并通过朗道能级和磁场的平方根关系得出表面能带的线性色散关系, 还通过拟合载流子浓度随磁场的变化关系得出 Bi_2Se_3 具有单狄拉克锥的能谱. 在垂直于膜面加上磁场后原来连续的表面能级, 变成离散的朗道能级. 此时通过改变扫描探针和薄膜之间的电压从而扫描到一个个的朗道能级.

2. AB(Aharonov-Bohm) 效应:

2009 年 12 月, 斯坦福张首晟研究组通过测量 Bi_2Se_3 纳米线的磁致电阻, 发现其振荡周期乘以截面积等于 h/e . 在考虑到费米能级处于带隙之间, 那么参与导电的必定来自于表面态. 说明这个 h/e 的周期来自于 AB 效应. 这也提供了一种研究表面态性质的方法. 在他们的工作中, 沿着一个 200 nm 宽, 50 nm 厚的纳米线的电流方向加一磁场后, 测到了其 0.4T 的振荡周期, 可以计算 $200 \text{ nm} \times 50 \text{ nm} \times 0.4 \text{ T} = h/e$, FFT 明确展示了明显的 h/e 的峰, 但是没有观察到量子干涉中常出现的 $h/2e$ 的周期. 这证实了一种表面导电态. 后来 UCLA 和南京大学课题组也开展了进一步的工作, UCLA 课题组实现了栅电压调控的 AB 振荡, 南京大学课题组则在一些精细选择过的解理样品中观察到了 AAS(Altshuler, Aronov, Spivak) 振荡, 其表面相的相干长度可以达到微米以上. AB 振荡方面的测量因受到表面能带弯曲引起的二维电子气的干扰, 需要更进一步的工作.

3.3.5 拓扑绝缘体的潜在应用

1. 奇异自旋输运载体:

拓扑绝缘体表面态受到强大的自旋轨道作用的驱动. 实验已经证明: 表面态狄拉克锥呈现显著的螺旋性, 其螺旋性比率接近 100%, 也就是说, 表面态中前行 (后行) 的电子具有向上 (向下) 的自旋. 这样的电子在经过一个杂质的时候, 如果发生完美的背散射, 则有顺时针/逆时针环绕两种路径向后散射, 这两种路径获得的相位刚好相差 π , 也即背散射被相干禁戒. 这就有可能获得低能耗高保真的自旋流, 从而使得拓扑绝缘体在自旋电子学材料中成为热点. 目前大量的相关人员在开展拓扑绝缘体材料的磁性掺杂, 这被认为是自旋输运和打开狄拉克点的必要路径. 斯坦福研究组的结果表明, 在 Bi_2Se_3 中掺杂 Mn 可以引起狄拉克锥中略小 0.1 meV 的带隙, 这可能来自于自旋中心对于表面态时间反演对称性的破坏. 这样的材料将可能允许反常霍尔效应的观察, 这也是当前半导体物理和量子物理的前沿内容.

2. Majorana(马约拉纳) 费米子:

自从中微子发现和双 β 衰变研究以来, Majorana 费米子的存在一直是个谜. 它是一类有质量的粒子而且反粒子就是其本身的基本粒子, 是量子信息、粒子物理和大统一理论等多学科研究的焦点. 近来, 有理论工作者提出, 利用超导邻近效应和拓扑绝缘体可以实现 Majorana 型激发, 在拓扑绝缘体中实现类似 p 波超导的性质, 在其中演示一类粒子和反粒子相同的元激发. 在此驱动下, 大量的工作集

中在拓扑绝缘体和超导的结合. 一种思路是在常见的拓扑绝缘体材料中实现超导, 普林斯顿研究组将 Cu 大比率的掺入 Bi_2Se_3 , 在 12%~20% 浓度时观察到 3K 附近的超导. 角分辨光电子能谱的测量还对其能带结构变化进行了描绘. 国内课题组对 Bi_2Te_3 施加高压, 也观察到了超导的出现. 另一种思路是将超导体与高自旋轨道耦合的材料接触实现 Majorana 型激发. 这一思路近来取得了突破性的进展. 荷兰 Delft 大学的研究组将超导电极施加在一类强自旋轨道耦合材料上, 并在低温下利用 STM 对其电子能态进行了详细观察, 在纳米线的两端零偏压时观察到一个零能的峰. 这一结果投给 Science 仅 20 天就被发表, 被认为是 Majorana 费米子的一个重要迹象.

3.4 铁基超导^[29]

铁基超导体是指含有铁的半金属化合物, 在低温 (几十 K) 时具有超导特性, 其中铁起了形成超导的主体作用. 2006 年, 日本东京工业大学 (Tokyo Institute of Technology) Hideo Hosono 的研究组第一个发现以铁为超导主体的化合物 LaOFeP ^[29]. 这个发现打破了以往认为铁磁元素不利于形成超导的观点. 根据 BCS 理论, 产生超导性的必要条件是材料中电子必须配对, 形成所谓的库珀 (Cooper) 对. 库珀对中的两个电子自旋相反, 所以总自旋为零. 因此, 此前通用的观点是超导性与铁磁性无法共存, 材料中加入磁性元素, 如铁、镍会大大降低超导性. 铁基超导体不但含有铁元素, 而且是产生超导的主体, 但铁与其他元素 (如砷、硒) 形成铁基平面后, 已不再具有铁磁性. 这个现象又一次显示了二维晶体在材料性质中起到的重要作用. 2008 年 2 月初, Hosono 研究组再度发表铁基层材料 $\text{La}[\text{O}_{1-x}\text{F}_x]\text{FeAs}$ ($x = 0.05 - 0.12$) 在绝对温度为 26K 时存在着超导性^[30], 此后铁基超导的研究在世界上形成一股热潮. 其之所以引起这么大的兴趣 (之后的研究工作表明, 这种铁基超导体的临界温度只有数十 K, 远不及 1986 年发现的铜氧化物^[2] (一种绝缘体) 高临界温超导体可以达到的液氮 (77K) 以上的温度) 是因为超导性发生在铁基平面上, 是属于二维超导材料, 因而与其他一种的铜氧高温超导类似. 研究铁基超导体可以有助于了解高温超导的机制. 值得提出的是尽管铜氧化物高温超导材料与铁基超导都具有层状结构, 但从电子结构上来讲, 导电的电子波函数前者是从过渡金属的 d 轨道演变过来, 而后者 (铁基) 则是由 s 轨道而来.

中国物理学家在铁基超导研究方面取得了很有时效与骄人的成绩. 中科院物理所闻海虎 (现在南京大学) 小组在 2008 年发表了临界温度为 56K ($\text{Ca}_{1-x}\text{Nd}_x\text{FeAsF}$) 的铁基超导体. 在 2009 年 3 月举行的美国物理学会年会上, 浙江大学与美国一些大学与阿拉莫斯 (Los Alamos) 国家实验室的研究人员提出了弱反铁磁 (或称磁无序、量子磁起伏) 的机理来解释这种超导的量子物相. 就在今年 (2012) 3 月中科院物理所的赵忠贤小组在 Nature Letter 上发表了铁基超导体与压力的关系, 当压力增加到一定值时 (11.5 GPa), 超导现象不但没有消失, 而且临界温度 (T_c) 可达到 48K^[31].

4 结语

硅基集成电路技术近半个世纪来眩目的进步在很大程度上是得益于固体物理的发展. 进入 21 世纪以来, 新材料尤其是低维材料的涌现, 为固体物理与集成电路提供了极佳的持续发展空间. 本文在回顾了基于量子力学的计算电子学发展史后, 重点讨论了低维晶体材料, 包括石墨烯、拓扑绝缘体与铁基超导体. 在 CMOS 技术快速接近摩尔定律得以成立的物理极限的时间段里, 如何利用这些已发现 (与将要出现) 的材料来探索新的器件结构与机理会是今后一个时期为固态领域中的研究人员所面临的主要任务及挑战. 固体物理的发展历史不止一次地证明, 人类在这方面的知识 (及应用) 远未穷尽, 更多、更大的进步会伴随我们不懈的努力而发生.

致谢 本报告撰写过程中得到了南京大学王伯根、赵毅教授的大力协助, 尤其是南京大学物理系提供的材料, 构成了报告的重要部分. 对他们的贡献与南京大学施毅教授的全力支持, 在此一并表示感谢. 北京大学博士生曾琅同学承担了量子输运部分的撰写, 对全稿的完成起了十分积极的作用.

参考文献

- 1 Ghani T, Armstrong M, Auth C, et al. A 90nm high volume manufacturing logic technology featuring novel 45nm gate length strained silicon CMOS transistors. In: Int'l Elec Dev Meeting (IEDM), 2003. 978–980
- 2 Bednorz J G, Müller K A. Possible high T_c superconductivity in the Ba-La-Cu-O system. *Z Phys B-Condens Matt*, 1986, 64: 189–193
- 3 Kane B E. A silicon-based nuclear spin quantum computer. *Nature*, 1998, 393: 133–137
- 4 Hohenberg P, Kohn W. Inhomogeneous electron gas. *Phys Rev*, 1964, 136: B864–B871
- 5 Kohn W, Sham L J. Self-consistent equations including exchange and correlation effects. *Phys Rev*, 1965, 140: A1133–A1138
- 6 Datta S. *Electronic Transport in Mesoscopic Systems*. 1st ed. Cambridge: Cambridge University Press, 1997
- 7 Slater J C, Koster G F. Simplified LCAO method for the periodic potential problem. *Phys Rev*, 1954, 94: 1498–1524
- 8 Vogl P, Hjalmarsen H P, Dow J. A semi-empirical tight-binding theory of the electronic structure of semiconductors. *J Phys Chem Solids*, 1983, 44: 365–378
- 9 Jancu J M, Scholz R, Beltram F, et al. Empirical *spds** tight-binding calculation for cubic semiconductors: general method and material parameters. *Phys Rev B*, 1998, 57: 6493–6507
- 10 Hedin L. New method for calculating the one-particle Green's function with application to the electron-gas problem. *Phys Rev*, 1965, 139: A796–A823
- 11 Hybertsen M S, Louie S G. Electron correlation in semiconductors and insulators: band gaps and quasiparticle energies. *Phys Rev B*, 1986, 34: 5390–5413
- 12 Faleev S V, van Schilfgaarde M, Kotani T. All-electron self-consistent *GW* approximation: application to Si, MnO, and NiO. *Phys Rev Lett*, 2004, 93: 126406 (1–4)
- 13 Van Wees B J, van Houten H, Beenakker C W J, et al. Quantized conductance of point contacts in a two-dimensional electron gas. *Phys Rev Lett*, 1988, 60: 848–850
- 14 Wharam D A, Thornton T J, Newbury R, et al. One-dimensional transport and the quantisation of the ballistic resistance. *J Phys C Solid State*, 1988, 21: L209–L214
- 15 Landauer R. Spatial variation of current and fields due to localized scatters in metallic conduction. *IBM J Res Dev*, 1957, 1: 223–231
- 16 Buttiker M. Symmetry of electrical conduction. *IBM J Res Dev*, 1988, 32: 317–334
- 17 Economou E N. *Green's Functions in Quantum Mechanics*. 3rd ed. Berlin: Springer, 2005
- 18 Datta S. *Quantum Transport: Atom to Transistor*. Cambridge: Cambridge University Press, 2005
- 19 Novoselov K S, Geim A K, Morozov S V, et al. Electric field effect in atomically thin carbon films. *Science*, 2004, 306: 666–669
- 20 Novoselov K S, Geim A K, Morozov S V, et al. Two-dimensional gas of massless Dirac fermions in graphene. *Nature*, 2005, 438: 197–200
- 21 Bernevig B A, Hughes T L, Zhang S C. Quantum spin hall effect and topological phase transition in HgTe quantum wells. *Science*, 2006, 314: 1757–1761
- 22 Iijima S. Helical microtubules of graphitic carbon. *Nature*, 1991, 354: 56–58
- 23 Bolotin K I, Sikes K J, Jiang Z. Ultrahigh electron mobility in suspended graphene. *Solid State Comm*, 2008, 146: 351–355
- 24 Jang B Z, Liu C, Neff D, et al. Graphene surface-enabled lithium ion-exchanging cells: next-generation high-power energy storage devices. *Nano Lett*, 2011, 11: 3785–3791
- 25 König M, Wiedmann S, Brüne C, et al. Quantum spin hall insulator state in HgTe quantum wells. *Science*, 2007, 318: 766–770
- 26 Fu L, Kane C L. Topological insulators with inversion symmetry. *Phys Rev B*, 2007, 76: 045302 (1–17)
- 27 Hsieh D, Qian D, Wray L, et al. A topological Dirac insulator in a quantum spin Hall phase. *Nature*, 2008, 452: 970–974
- 28 Zhang H, Liu C X, Qi X L, et al. Topological insulators in Bi_2Se_3 , Bi_2Te_3 and Sb_2Te_3 with a single Dirac cone on the surface. *Nature Phys*, 2009, 5: 438–442

- 29 Kamihara Y, Hiramatsu H, Hirano M, et al. Iron-based layered superconductor: LaOFeP. *J Am Chem Soc*, 2006, 128: 10012–10013
- 30 Kamihara Y, Watanabe T, Hirano M, et al. Iron-based layered superconductor $\text{LaO}_{1-x}\text{F}_x\text{FeAs}$ ($x = 0.05 - 0.12$) with $T_c = 26$ K. *J Am Chem Soc*, 2008, 130: 3296–3297
- 31 Sun L, Chen X J, Guo J, et al. Re-emerging superconductivity at 48 kelvin in iron chalcogenides. *Nature* 2012, 483: 67–69

Progress in solid-state physics—A review for the past half a century

YU ZhiPing^{1*}, WANG Yan¹ & LIU XiaoYan²

1 *Institute of Microelectronics, Tsinghua University, Beijing 100084, China*

2 *Institute of Microelectronics, Peking University, Beijing 100871, China*

*E-mail: yuzhip@tsinghua.edu.cn

Abstract The progress of solid-state physics for the past half a century is reviewed in view of theory, materials, and experiments. The entry point for the review is the development of density-functional theory (DFT) in 1964, which ushered the so-called first-principles computational electronics. The review mainly focuses on two aspects: solid-state theory and computational electronics, and low-dimensional materials including superconductors. Their impact on and relevance to solid-state integrated circuits (ICs) are also discussed.

Keywords solid-state physics, integrated circuits, computational electronics, low-dimensional materials and physics, graphene/topological insulator/iron-based superconductors



YU ZhiPing Professor in the Institute of Microelectronics, Tsinghua University, Beijing, and an IEEE Fellow. He graduated from Tsinghua University with B.S. degree in 1967, and M.S. and Ph.D. degree from EE Dept. at Stanford University in the US in 1980 and 1985, respectively. His research interests include IC CAD (computer-aided-design) and nanoelectronics.



LIU XiaoYan received the B.S., M.S., and Ph.D. degrees in microelectronics from Peking University, Beijing, in 1988, 1991, and 2001, respectively. In 1991, she joined Peking University to work on TFT AM LCD and driver ICs for display. From 1995 to 1996, she was a Research Assistant with the Department of Electrical and Electronic Engineering, The Hong Kong University of Science and Technology, Kowloon, Hong Kong, where she worked on IC drivers for LCD display. She is currently a Professor with Institute of Microelectronics, Peking University. Her research interests include nanoscale device physics, device simulation, and nanoscale device modeling. She has authored and coauthored over 100 papers and three books on semiconductor devices.



WANG Yan received her B.S. and M.S. degrees in electrical engineering from Xi'an Jiaotong University, Xi'an, in 1988 and 1991, respectively, and the Ph.D. degree in semiconductor device and physics from the Chinese Academy of Sciences, Beijing, in 1995. Since 1999, she has been an Associate Professor with the Institute of Microelectronics at Tsinghua University (IMETU), Beijing, and promoted to Full Professor in 2004. Since 2003, She has been the vice-director of the IMETU and engaged in the student management and teaching management. Her current research interests include modeling of RF/MicroWave/MMWave devices and components, the carrier transport models in scaled-down MOS devices, CAD software development for micro- and nano-devices. She gives the course Nano-electronic Devices to the graduate students and Solid State Physics to the undergraduate students.